

CIRA SOUZA PITOMBO

**ESTUDOS DE RELAÇÕES ENTRE VARIÁVEIS  
SOCIOECONÔMICAS, DE USO DO SOLO, PARTICIPAÇÃO EM  
ATIVIDADES E PADRÕES DE VIAGENS ENCADEADAS  
URBANAS**

Tese apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo, como parte dos requisitos para obtenção do título de Doutor em Engenharia Civil: Planejamento e Operação de Sistemas de Transportes.

**Orientador: Prof. Titular Eiji Kawamoto**

São Carlos  
Abril  
2007

# **Livros Grátis**

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

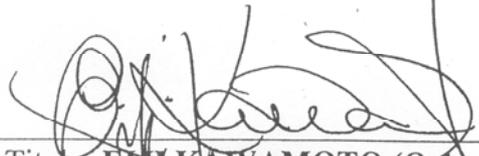
Ficha catalográfica preparada pela Seção de Tratamento  
da Informação do Serviço de Biblioteca – EESC/USP

|       |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |
|-------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| P685e | <p>Pitombo, Cira Souza</p> <p>Estudos de relações entre variáveis socioeconômicas, de uso do solo, participação em atividades e padrões de viagens encadeadas urbanas / Cira Souza Pitombo ; orientador Eiji Kawamoto. -- São Carlos, 2007.</p> <p>Tese (Doutorado) - Programa de Pós-Graduação em Engenharia de Transportes e Área de Concentração em Planejamento e Operações de Sistemas de Transportes -- Escola de Engenharia de São Carlos da Universidade de São Paulo.</p> <p>1. Variáveis de uso do solo. 2. Características socioeconômicas. 3. Participação em atividades. 4. Técnicas de análise multivariadas. 5. Padrões de viagens. 6. Modelagem de demanda. I. Título.</p> |
|-------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

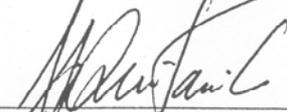
FOLHA DE JULGAMENTO

Candidata: Engenheira **CIRA SOUZA PITOMBO**

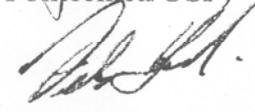
Tese defendida e julgada em 27/04/2007 perante a Comissão Julgadora:

  
\_\_\_\_\_  
Prof. Titular **EIJI KAWAMOTO (Orientador)**  
(Escola de Engenharia de São Carlos/USP)

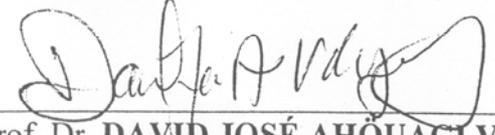
APROVADA

  
\_\_\_\_\_  
Prof. Associado **JOSE ALBERTO QUINTANILHA**  
(Escola Politécnica/USP)

APROVADA

  
\_\_\_\_\_  
Prof. Associado **ORLANDO STRAMBI**  
(Escola Politécnica/USP)

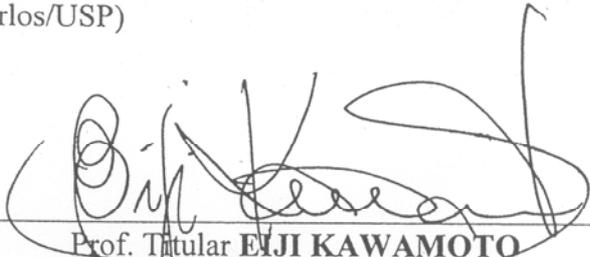
Aprovado

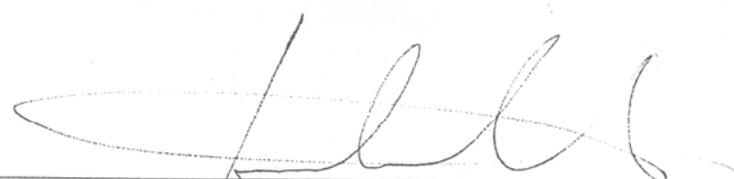
  
\_\_\_\_\_  
Prof. Dr. **DAVID JOSÉ AHÓUAGI VAZ DE MAGALHÃES**  
(Universidade Federal de Minas Gerais/UFMG)

Aprovado

  
\_\_\_\_\_  
Prof. Associado **ANTÔNIO NELSON RODRIGUES DA SILVA**  
(Escola de Engenharia de São Carlos/USP)

APROVADA

  
\_\_\_\_\_  
Prof. Titular **EIJI KAWAMOTO**  
Coordenador do Programa de Pós-Graduação em  
em Engenharia de Transportes

  
\_\_\_\_\_  
Prof. Associado **GERALDO ROBERTO MARTINS DA COSTA**  
Presidente da Comissão da Pós-Graduação da EESC

Aos meus queridos pais, Nilton e Jacina, pelo apoio, incentivo e amor irrestritos.

Ao meu irmão, Marcelo, meu grande amigo, pelas revisões, além da eterna torcida pelo meu êxito.

# AGRADECIMENTOS

---

A Deus, por tudo que tem me proporcionado.

A Nossa Senhora de Fátima, pela eterna proteção e à minha mãe, Iacina, por fortalecer diariamente na minha vida a importância da fé e da família.

Ao professor Eiji Kawamoto, toda a minha admiração e respeito, pelos seus conselhos, não somente no decorrer deste trabalho, mas da minha vida acadêmica. Agradeço sinceramente pela amizade e paciência ao longo de seis anos.

Aos professores da Escola Politécnica (Universidade de São Paulo) Orlando Strambi e José Alberto Quintanilha pela disponibilidade de sempre e pelos valiosos conselhos durante o desenvolvimento deste trabalho.

Ao professor Antônio Nelson Rodrigues da Silva (Escola de Engenharia de São Carlos), pelas sugestões ao trabalho no exame de qualificação, assim como por todo auxílio durante estes seis anos.

Ao professor David José Ahouagi Vaz de Magalhães (Universidade Federal de Minas Gerais) pelas conversas e sugestões nos encontros da ANPET.

Ao professor Wellington Figueiredo (Universidade Federal da Bahia) pelo incentivo desde 2000.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão da bolsa de doutorado nos três últimos anos.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) pela concessão da bolsa de doutorado no primeiro ano da pesquisa.

À Companhia do Metropolitan de São Paulo pela cessão dos dados das pesquisas de origem e destino de 1987 e 1997.

À Empresa Metropolitana de Transportes Urbanos (EMTU – SP), sobretudo ao Sr. Atila Sarkozy, pela cessão dos dados referentes à parte da rede de transporte público da Região Metropolitana de São Paulo.

À amiga querida Karênina Teixeira pelos conselhos e amizade diária ao longo destes seis anos.

Gostaria de deixar registrado meu agradecimento pessoal a três colegas e grandes amigos com quem tive a sorte de trabalhar diariamente - Mateus Araújo e Silva, Bruno Bertoncini e Professor Marcus Seraphico - pela grande ajuda e solidariedade.

Ao colega e amigo André Cunha pelo auxílio no desenvolvimento das macros para processamento de parte dos dados.

Aos amigos e funcionários do Departamento de Transportes: Heloísa, Beth, Suely, Alexandre, Gigante, Paulo, Toco e Beto. À Magaly, o meu carinho especial.

A vários outros amigos que fiz no Departamento de Transportes, meu agradecimento coletivo, destacando a importância de todos na minha vida pessoal e profissional: Adalberto Faxina, Ana Paula Furlan, Antonio Carlos Dinato, Andréa Júlia, Adriana Goulart, Camilla Carrara, Cida Moraes, Cynthia Medeiros, Daniel Anijar, Diogo Collela, Fabiana Arruda, Fábio Zancheta, Francis Kakuda, Giseli Ortolani, Gustavo Manzato (especialmente pelo incentivo aos projetos após o doutorado), Jesner Ildfonso, Juliana Araújo, Marcela Costa, Marcelo Aguiar, Márcia Pereira, Marta Pereira da Luz, Michael Saunders, Pablo Brilhante, Pastor, Paulo Rogério Zanoli, Ricardo Melo, Sandra Ichikawa, Simone Becker e Wesley Novaes.

A todos os professores do Departamento de Transportes que muito contribuíram para meu crescimento intelectual.

Aos meus amigos "são-carlenses" Ana Paula Larocca, Andréa Helena de Castro, Rodrigo Caruso Friedmann, Cristina Maria de Castro, Georgina Gonçalves de Castro e Mariana Passos.

Às minhas amigas inseparáveis e insubstituíveis Bruna Carraro, Juliana Melo, Lorena Amorim e Veruska Pithon.

A toda a minha numerosa família, pela torcida pelo meu sucesso, especialmente à "minha senhora" por tanto carinho mesmo que à distância.

Finalmente, a todos que contribuíram de alguma forma para o desenvolvimento e conclusão do presente trabalho.

**"A verdadeira viagem da descoberta não consiste em procurar novas paisagens, mas em ter novos olhares."**

Marcel Proust

## RESUMO

PITOMBO, C.S. **Estudos de relações entre variáveis socioeconômicas, de uso do solo, participação em atividades e padrões de viagens encadeadas urbanas.** 2007. 268 p. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos 2007.

Um dos tópicos mais importantes na análise de demanda por transportes é a relação entre as necessidades individuais de realização de atividades geograficamente distribuídas, a estrutura urbana, as características individuais e domiciliares, o sistema de transporte e as diferenças no comportamento relacionado a viagens. Isto motivou o desenvolvimento desta pesquisa, que tem como objetivo principal analisar o comportamento individual subjacente ao encadeamento de viagens sob a perspectiva de três grupos de variáveis: (1) participação em atividades; (2) características socioeconômicas; e (3) uso do solo. Há dois objetivos secundários, fundamentais para se atingir a finalidade do trabalho: (a) propor um conjunto de variáveis de uso do solo; e (b) testar a significância do grupo de variáveis ora proposto. Este trabalho baseou-se nos dados da Pesquisa Origem-Destino de 1997 da Região Metropolitana de São Paulo, sendo extraídas e analisadas seis amostras finais que foram caracterizadas por setor econômico (no caso de trabalhadores) e grau de instrução (no caso de estudantes). Com utilização conjunta de técnicas de Análise Multivariadas, confirmatórias e exploratórias, foi possível representar a variável dependente (Análise de *Cluster*), bem como encontrar relações entre variáveis envolvidas (Árvore de Decisão) e, finalmente, mensurar a significância estatística das variáveis independentes (Regressão Linear Múltipla). Através dos resultados obtidos, foi possível analisar a influência dos três grupos de variáveis na seqüência de viagens: (1) variáveis socioeconômicas (Renda familiar, Usa Vale Transporte, Nº provável de carteiras de habilitação no domicílio, Idade, Nº de automóveis no domicílio) afetam principalmente a seqüência de modos de transporte utilizados durante as viagens; (2) participação em atividades (Estuda, Trabalha) interfere na seqüência de motivos de viagem; e, enfim, (3) variáveis de uso do solo (Parcela acumulada de empregos ou escolas por faixas de distância a partir do centróide da zona de residência) influenciam a seqüência de destinos escolhidos. Espera-se que o presente trabalho constitua uma contribuição ao meio acadêmico, tanto em termos de representação da intensidade e distribuição geográfica das atividades no meio urbano (variáveis de uso do solo), quanto em relação à influência de tais variáveis nos deslocamentos dos indivíduos.

Palavras-chave: Variáveis de uso do solo, Características socioeconômicas, Participação em atividades, Técnicas de Análise Multivariadas, Padrões de viagens, Modelagem de demanda.

## ABSTRACT

PITOMBO, C.S. **Study of relationships between socioeconomic, land use, activity participation variables and trip-chaining urban patterns.** 2007. 268 p. Thesis (Doctoral) – Engineering School of São Carlos, University of São Paulo. São Carlos, Brazil, 2007.

One of the most important topics in Transportation Demand is the relationship between individual needs to carry out geographically distributed activities, urban configuration, individual and household characteristics, transportation system and travel behavior. For that reason, the main aim of this work is to analyze the individual trip-chaining behavior in terms of three variables groups: (1) activity participation; (2) socioeconomic characteristics; and (3) land use. There are also two secondary objectives derived from the main objective: (a) to propose one set of land use variables; and (b) to verify the statistical significance of the created land use group variables. This work was based on the Origin-Destination survey carried out in the São Paulo Metropolitan Area in 1997, from which six final samples were extracted, analyzed and characterized by economic sector (for workers) and level of education (for students). Applying Multivariate Analysis techniques, confirmatory and exploratory, it was possible to represent the dependent variable (Cluster Analysis), as well as to find relationships between the concerned variables (Decision Tree) and, finally, to measure the statistical significance of the independent variables (Multiple Regression). From the results, it was possible to analyze the influence of the three variables groups on trip-chaining: (1) socioeconomic variables (Household Income, Transit Voucher use, Probable number of driver licenses per household, Age, Car-ownership) affect the travel mode sequence used for the trips; (2) activity participation (Study, Work) has an effect on the trip purpose sequence; and (3) land use variables (accumulated proportion of jobs or schools by distance buffers starting from the residence zone centroid) influence the sequence of chosen destinations. It is expected that the present work could be a contribution to the scientific community for the representation of the activities level and their geographic distribution in the urban configuration (land use variables), and the influence of such variables on individuals displacements.

Keywords: Land use variables, Socioeconomic characteristics, Activity participation, Multivariate Analysis techniques, Travel patterns, Demand modelling.

# SUMÁRIO

---

|                                                                                |    |
|--------------------------------------------------------------------------------|----|
| <b>Capítulo 1: Introdução</b> .....                                            | 12 |
| 1.1 Uma breve revisão bibliográfica: contextualizando o tema.....              | 12 |
| 1.2 Objetivos .....                                                            | 15 |
| 1.3 Descrição do trabalho .....                                                | 15 |
| 1.3.1 Linha de pesquisa.....                                                   | 15 |
| 1.3.2 Síntese do método.....                                                   | 18 |
| 1.4 Justificativa .....                                                        | 20 |
| 1.5 Questões fundamentais .....                                                | 21 |
| 1.6 Estrutura do trabalho .....                                                | 22 |
| <br>                                                                           |    |
| <b>Capítulo 2: Uso do solo x padrões de viagens</b> .....                      | 24 |
| 2.1 Considerações iniciais.....                                                | 24 |
| 2.2 Uso do solo e comportamento relacionado a viagens .....                    | 26 |
| 2.2.1 Evolução da estrutura urbana: algumas teorias sobre formas urbanas ..... | 27 |
| 2.2.2 A estrutura urbana e comportamento relacionado a viagens.....            | 29 |
| 2.2.3 Uso do solo e transportes: fortes relações ou casualidade.....           | 31 |
| 2.3 Mensuração de variáveis de uso do solo .....                               | 36 |
| 2.4 Modelos de distribuição de viagens .....                                   | 40 |
| 2.4.1 Modelo gravitacional.....                                                | 41 |
| 2.4.2 Modelo de oportunidades intervenientes .....                             | 43 |
| 2.4.3 Modelos gravitacionais – de oportunidades.....                           | 46 |
| 2.5 Considerações finais.....                                                  | 48 |
| <br>                                                                           |    |
| <b>Capítulo 3: Análise Multivariada de dados</b> .....                         | 50 |
| 3.1 Escolha das técnicas de Análise Multivariada de dados .....                | 50 |
| 3.1.1 Técnicas de AM tradicionais (técnicas confirmatórias) .....              | 51 |
| 3.1.2 Técnicas de AM “emergentes” (técnicas exploratórias).....                | 52 |
| 3.1.3 Uso de técnicas exploratórias + confirmatórias .....                     | 54 |
| 3.2 Análise de Cluster (AC) .....                                              | 55 |
| 3.2.1 Definir o problema de aglomeração ou objetivo da AC na pesquisa .....    | 55 |
| 3.2.2 Escolha das variáveis .....                                              | 57 |
| 3.2.3 Escolha da medida de similaridade ou dissimilaridade.....                | 58 |
| 3.2.4 Definição do processo de aglomeração.....                                | 58 |
| 3.2.5 Definição do número de conglomerados .....                               | 61 |
| 3.2.6 Peculiaridades da AC .....                                               | 61 |
| 3.3 Árvore de Decisão (AD) .....                                               | 62 |
| 3.3.1 Principais propriedades .....                                            | 63 |
| 3.3.2 S-PLUS 6.1 e CART .....                                                  | 64 |
| 3.4 Regressão Linear Múltipla (RLM) .....                                      | 67 |
| 3.5 Considerações a respeito da aplicação da RLM em conjunto com a AD.....     | 70 |
| 3.6 Principais características das técnicas utilizadas.....                    | 71 |

|                                                                         |            |
|-------------------------------------------------------------------------|------------|
| <b>Capítulo 4: Estudo de caso</b> .....                                 | <b>73</b>  |
| 4.1 Descrição da região de estudo .....                                 | 73         |
| 4.2 Dados individuais e domiciliares .....                              | 74         |
| 4.3 Características por zonas de tráfego .....                          | 76         |
| 4.4 Tratamento dos dados .....                                          | 85         |
| <b>Capítulo 5: Variáveis independentes</b> .....                        | <b>88</b>  |
| 5.1 Considerações iniciais .....                                        | 88         |
| 5.2 Participação em atividades .....                                    | 89         |
| 5.3 Características socioeconômicas .....                               | 90         |
| 5.4 Uso do solo .....                                                   | 93         |
| 5.4.1 Análise das variáveis agregadas por zona de tráfego .....         | 94         |
| 5.4.2 Investigação da escolha dos destinos na RMSP .....                | 95         |
| 5.4.3 Proposta do grupo de variáveis de uso do solo .....               | 97         |
| 5.5 Considerações finais .....                                          | 105        |
| <b>Capítulo 6: Variável dependente</b> .....                            | <b>107</b> |
| 6.1 Considerações iniciais .....                                        | 107        |
| 6.2 Seqüência de motivos de viagem (etapa 1) .....                      | 108        |
| 6.3 Seqüência de modos de transporte (etapa 2) .....                    | 112        |
| 6.4 Seqüência de destinos (etapa 3) .....                               | 116        |
| 6.4.1 Aplicação da AC .....                                             | 116        |
| 6.4.2 Resultados do agrupamento .....                                   | 118        |
| 6.4.3 Caracterização das seqüências de destinos .....                   | 118        |
| 6.5 Padrões finais (etapa 4) .....                                      | 121        |
| <b>Capítulo 7: Aplicação de AD e resultados</b> .....                   | <b>125</b> |
| 7.1 Variáveis envolvidas e critérios adotados .....                     | 125        |
| 7.2 Trabalhadores no setor industrial .....                             | 127        |
| 7.3 Trabalhadores no setor comercial .....                              | 134        |
| 7.4 Trabalhadores no setor de serviços .....                            | 137        |
| 7.5 Estudantes até o 1º grau .....                                      | 140        |
| 7.6 Estudantes no 2º grau .....                                         | 143        |
| 7.7 Estudantes universitários .....                                     | 146        |
| 7.8 Síntese dos resultados obtidos .....                                | 149        |
| <b>Capítulo 8: Análise dos resultados</b> .....                         | <b>151</b> |
| 8.1 Análise geral dos resultados .....                                  | 151        |
| 8.2 Trabalhadores no setor industrial .....                             | 153        |
| 8.2.1 Síntese da análise de resultados – amostra de industriários ..... | 165        |
| 8.3 Trabalhadores no setor comercial .....                              | 166        |
| 8.4 Trabalhadores no setor de serviços .....                            | 170        |
| 8.5 Estudantes até o 1º grau .....                                      | 173        |
| 8.6 Estudantes no 2º grau .....                                         | 178        |
| 8.7 Estudantes universitários .....                                     | 180        |
| 8.8 Considerações finais .....                                          | 184        |

|                                                                                |            |
|--------------------------------------------------------------------------------|------------|
| <b>Capítulo 9: Significância estatística das variáveis independentes .....</b> | <b>185</b> |
| 9.1 Aplicação da RLM em conjunto com a AD.....                                 | 185        |
| 9.2 Trabalhadores no setor industrial.....                                     | 187        |
| 9.2.1 Discretização das variáveis independentes.....                           | 188        |
| 9.2.2 Representação das variáveis dependentes.....                             | 188        |
| 9.2.3 Análise de correlação entre variáveis envolvidas.....                    | 190        |
| 9.2.4 Dez modelos lineares obtidos e seus respectivos parâmetros.....          | 191        |
| 9.3 Trabalhadores no setor comercial.....                                      | 195        |
| 9.3.1 Aplicação da RLM e significância das variáveis independentes.....        | 196        |
| 9.4 Trabalhadores no setor de serviços.....                                    | 200        |
| 9.4.1 Aplicação da RLM e significância das variáveis independentes.....        | 200        |
| 9.5 Estudantes até o 1º grau.....                                              | 203        |
| 9.5.1 Aplicação da RLM e significância das variáveis independentes.....        | 203        |
| 9.6 Estudantes no 2º grau.....                                                 | 206        |
| 9.6.1 Aplicação da RLM e significância das variáveis independentes.....        | 206        |
| 9.7 Estudantes universitários.....                                             | 209        |
| 9.7.1 Aplicação da RLM e significância das variáveis independentes.....        | 209        |
| 9.8 Considerações finais.....                                                  | 212        |
| <br>                                                                           |            |
| <b>Capítulo 10: Considerações finais .....</b>                                 | <b>213</b> |
| 10.1 Descrição sucinta do trabalho atual.....                                  | 213        |
| 10.1.1 Revisão da literatura.....                                              | 214        |
| 10.1.2 Escolha das técnicas de AM.....                                         | 214        |
| 10.1.3 Estudo de caso.....                                                     | 215        |
| 10.1.4 Representação das variáveis independentes.....                          | 216        |
| 10.1.5 Representação da variável dependente.....                               | 217        |
| 10.1.6 Aplicação da AD e resultados.....                                       | 217        |
| 10.1.7 Análise de resultados obtidos com aplicação da AD.....                  | 218        |
| 10.1.8 Teste de significância das variáveis independentes.....                 | 218        |
| 10.2 Principais conclusões.....                                                | 219        |
| 10.3 Recomendações para trabalhos futuros.....                                 | 223        |
| <br>                                                                           |            |
| <b>Referências bibliográficas .....</b>                                        | <b>226</b> |
| <br>                                                                           |            |
| <b>Anexo A .....</b>                                                           | <b>236</b> |
| <br>                                                                           |            |
| <b>Anexo B .....</b>                                                           | <b>238</b> |
| <br>                                                                           |            |
| <b>Anexo C .....</b>                                                           | <b>242</b> |
| <br>                                                                           |            |
| <b>Anexo D .....</b>                                                           | <b>244</b> |
| <br>                                                                           |            |
| <b>Anexo E .....</b>                                                           | <b>247</b> |
| <br>                                                                           |            |
| <b>Anexo F.....</b>                                                            | <b>249</b> |
| <br>                                                                           |            |
| <b>Anexo G .....</b>                                                           | <b>251</b> |

# Capítulo

# 1

# INTRODUÇÃO

---

*Esta introdução responde a indagações básicas para o desenvolvimento de uma pesquisa: (1) Até onde caminhou o estado da arte do tema pesquisado? (breve revisão bibliográfica de forma a contextualizar o tema); (2) O quê fazer? (objetivos); (3) Como fazer? (descrição sucinta do método adotado); (4) Por que fazer? (justificativa, descrevendo razões para se debruçar sobre o tema e a sua possível contribuição à comunidade a que se dirige). Ao final deste capítulo há uma série de questões que motivaram a realização do presente trabalho.*

## 1.1. UMA BREVE REVISÃO BIBLIOGRÁFICA – CONTEXTUALIZANDO O TEMA

Um dos tópicos mais importantes na análise de demanda por transportes é a possível relação entre as necessidades individuais de realização de atividades espacialmente dispersas, a estrutura urbana, as características individuais e domiciliares, o sistema de transporte e as diferenças no comportamento relacionado a viagens. Os indivíduos deslocam-se no meio urbano considerando suas necessidades de realizar atividades fora do domicílio, suas próprias características individuais, atributos domiciliares e características do seu meio.

Tornou-se então um dos objetos da Análise de Viagens baseada em Atividades investigar as variáveis que influenciam a programação da seqüência de viagens individuais. Foi sugerido através de diversos trabalhos (STRATHMAN; DUEKER, 1995; KWAM, 2000; BHAT; SINGH, 2000; SUN; WILMOT; KASTURI; 1998) que indivíduos consideram alguns fatores principais ao realizar os seus respectivos itinerários de viagens: (1) participação em atividades; (2) características socioeconômicas individuais e domiciliares; (3) características das áreas urbanas; (4) atributos do sistema de transporte.

A influência das características socioeconômicas individuais e domiciliares e a participação das pessoas em diversas atividades na formação de padrões de viagens vem sendo estudada nos últimos anos (LU; PAS, 1999 ; GOLOB; MACNALLY, 1997; BHAT; KOPPELMAN, 1991). De fato, na literatura

vigente predomina a afirmação de que características dos deslocamentos pessoais podem ser determinadas pelo sexo, posse de automóveis (STRAMBI; VESPUCCI; van DE BILT, 2004; MCGUCKIN; MURAKAMI, 1999), papel do indivíduo no domicílio e alocação de tarefas (SRINIVASAN; ATHURU, 2005; SIMMA, AXHAUSEN, 2001; BALASUBRAMANIAM; GOULIAS, 1999) e participação em atividades (classificadas geralmente como subsistência, lazer, saúde, compras, etc.).

Contudo, deve-se ressaltar que características socioeconômicas e participação em atividades são apenas partes de um conjunto de variáveis que permite explicar o comportamento concernente à escolha do padrão de viagens encadeadas. Assim como a posse de automóveis influencia a escolha modal, é provável que a disponibilidade de determinados modos de viagem, numa dada zona, exerça influência na formação de padrões de viagem nessa zona. Ainda, é possível que o padrão de uso do solo das zonas ajude a explicar as escolhas de destinos.

Aos poucos, novas dimensões vêm sendo incorporadas à análise, em busca da representação mais realística do comportamento referente ao encadeamento de viagens e da construção de estruturas mais adequadas para previsão da demanda por transportes. Isso pode ser exemplificado pela introdução de variáveis que caracterizam a configuração urbana e outras, relacionadas ao sistema de transportes. Desta forma, surgiram novos trabalhos que consideram, além de características socioeconômicas e participação em atividades, o uso do solo e aspectos ligados ao sistema de transportes.

Kitamura (1985) examinou relações entre a tendência dos indivíduos encadear as viagens e as características de uma cidade hipotética linearmente disposta. A análise mostrou que a tendência de encadear viagens é função da utilidade de um conjunto de oportunidades e do tipo de atividade a ser realizada. Arruda (2005) aplicou um modelo baseado em atividades na cidade de São Carlos (SP), a fim de estudar relações entre uso do solo e comportamento de viagem/agenda de atividades. Entretanto, os resultados gerados pelo modelo não foram suficientes para fornecer evidências de que as características de uso do solo atuam de forma significativa no processo de tomada de decisão individual em relação a viagens.

Srinivan (2000) investigou como as características da vizinhança (uso do solo, rede de transporte e medidas de acessibilidade, quantificadas com auxílio de um SIG) afetam o comportamento de viagem em relação à escolha modal e cadeia de viagens. Dentre os resultados obtidos, observa-se, por exemplo, que indivíduos residentes em zonas de tráfego mistas (com altas densidades comerciais e residenciais), além de realizarem cadeias de viagens com atividades diferentes de trabalho, fazem este tipo de viagem a pé. Dong et al. (2002) desenvolveram uma medida de acessibilidade - ABA (*Activity-Based Accessibility*)-gerada através de um modelo de programação de atividades diárias, incorporando efeitos de uso do solo e sistemas de transportes a escolhas de padrões de viagens.

Assim, baseando-se na literatura, espera-se que a escolha dos padrões de viagens faça parte de um processo de tomada de decisão individual, onde se consideram as atividades a serem realizadas, características individuais e do domicílio, a localização relativa da residência e um conjunto de oportunidades e atributos do sistema de transportes. A estrutura proposta na Figura 1.1 representa um possível conjunto de variáveis relacionadas a diferentes escolhas de padrões de viagens encadeadas.

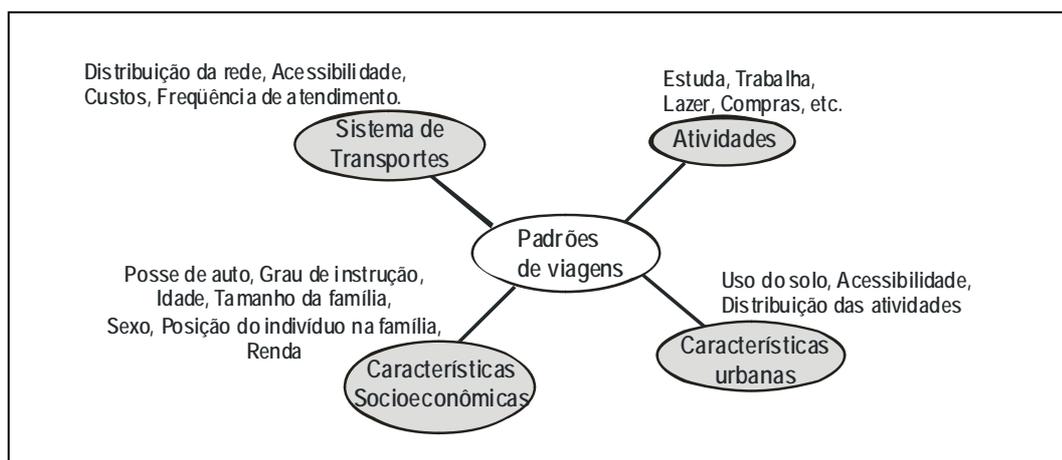


Figura 1.1 Conjunto de variáveis que influenciam o encadeamento de viagens urbanas.

## 1.2. OBJETIVOS

O objetivo da pesquisa proposta é analisar o comportamento referente ao encadeamento de viagens urbanas na Região Metropolitana de São Paulo (RMSP), sob a perspectiva de três dos quatro grupos de variáveis representados na Figura 1.1: (1) participação em atividades; (2) características socioeconômicas individuais e domiciliares e; (3) características do meio urbano (uso do solo).

Assim, propõe-se encontrar relações entre padrões de encadeamento de viagens (variável dependente) e variáveis socioeconômicas individuais e domiciliares, participação em atividades, além de uso do solo (variáveis independentes).

Há dois objetivos secundários, que fazem parte do método adotado, que são fundamentais para atingir a finalidade principal:

- Propor um conjunto de variáveis que representem características de uso do solo – vale ressaltar que neste trabalho variáveis de uso do solo representam a intensidade e a distribuição geográfica das atividades no meio urbano;
- Testar a significância do conjunto de variáveis ora proposto.

## 1.3. DESCRIÇÃO DO TRABALHO

### 1.3.1 Linha de pesquisa

O presente trabalho é uma extensão de estudos realizados no Departamento de Transportes da Escola de Engenharia de São Carlos, Universidade de São Paulo (EESC-USP), sob orientação do professor Dr. Eiji Kawamoto a partir de 2002. O grupo de pesquisa realizou trabalhos complementares nos últimos anos com objetivo de melhor compreender a diversidade de padrões de atividades e viagens realizados pelos indivíduos e suas relações com características socioeconômicas e do meio urbano. São trabalhos investigativos e exploratórios, por se tratar de uma linha de pesquisa relativamente nova, caracterizadas por dificuldades e desafios.

Inicialmente, foi analisada a influência apenas de características socioeconômicas e participação dos indivíduos em atividades na escolha de padrões de viagens urbanas. Os trabalhos iniciais,

realizados por Ichikawa (2002) e Ichikawa, Pitombo e Kawamoto (2003), buscaram avaliar o desempenho de minerador de dados na obtenção de relações entre características socioeconômicas e participação em atividades e padrões de viagens encadeadas.

Uma vez comprovada a utilidade de minerador de dados neste tipo de pesquisa, seguiram-se outros trabalhos, desta vez com o objetivo de analisar o comportamento de viajantes no que se refere à escolha dos padrões de viagens encadeadas (PITOMBO, 2003; PITOMBO; KAWAMOTO, 2003; PITOMBO; KAWAMOTO, 2004). Como resultado, foi possível mapear a forma como as características socioeconômicas e a participação em atividades, principalmente nas compulsórias, influenciam a escolha dos padrões de viagens encadeadas, definidos em termos de seqüências de motivos, de modos de viagem, de períodos do dia e duração das viagens. Realizada a análise do comportamento individual subjacente ao encadeamento de viagens, Silva (2006) utilizou um conjunto de variáveis similar em uma tentativa inicial de aplicação de técnicas de mineração de dados na previsão de demanda por transportes.

Posteriormente buscou-se identificar a possível influência de diferentes áreas urbanas na escolha de padrões de viagens encadeadas, assim como mensurar características de uso do solo a fim de incorporá-las a modelos de segmentação (CART – *Classification and Regression Tree*).

Souza (2004) identificou possíveis influências de diferentes áreas urbanas nos padrões de viagens encadeadas. Utilizando como casos de estudo a RMSP, Região Metropolitana de Belém (RMB) e a cidade de Bauru, conclui que, em geral, o comportamento de viajantes urbanos é influenciado por políticas urbanas regionais, características socioeconômicas e espaciais.

Aguiar (2005) analisou, através de Árvore de Decisão e Classificação, a influência de uma variável categórica, denominada “Macro-zona” (representando características espaciais e de uso do solo da área urbana que constitui a RMB), nos padrões de viagens encadeadas realizados pelos indivíduos. Pitombo e Kawamoto (2006a) propuseram, em um estudo preliminar, uma caracterização de uso do solo alternativa para a RMSP e investigaram a escolha dos diferentes destinos durante a

cadeia de viagens realizadas pelos indivíduos. Os resultados indicam que o grau de atividade da zona de origem influencia a escolha dos padrões de viagens.

No intuito de representar adequadamente variáveis de uso do solo, de forma que estas reproduzissem a distribuição das atividades urbanas na RMSP e fossem consideradas importantes no modelo CART, foram desenvolvidos os últimos trabalhos, utilizando técnicas de Análise Multivariada de dados (PITOMBO; KAWAMOTO, 2006b; PITOMBO; KAWAMOTO; QUINTANILHA, 2006).

O trabalho em questão, que se insere no contexto dos estudos anteriores, propõe investigar relações entre aspectos da intensidade e distribuição de atividades no meio urbano (mensurados e representados em termos de variáveis independentes), características socioeconômicas, participação em atividades (variáveis abordadas anteriormente) e padrões de encadeamento de viagens.

Um dos objetivos desta linha de pesquisa é buscar, continuamente, fundamentação teórica e compreensão das variáveis que realmente influenciam os padrões de viagens, visando brevemente trazer contribuições para aplicações práticas. Pretende-se obter um modelo de comportamento de viajante que seja aplicável a pelo menos um grupo de cidades. O modelo resultante poderia ser aplicado a um conjunto de informações, tais como as disponíveis do IBGE (recenseamento) e nas prefeituras (distribuição espacial de atividades, etc.). Desta forma, poder-se-ia associar grupos de indivíduos com determinadas características socioeconômicas e de uso do solo à escolha da seqüência de deslocamentos. Fornecendo, assim, informações úteis e alternativas baratas a programas de Planejamento de Transportes. A Figura 1.2 mostra sucintamente os trabalhos realizados no Departamento de Transportes da EESC-USP, incluindo a pesquisa proposta.

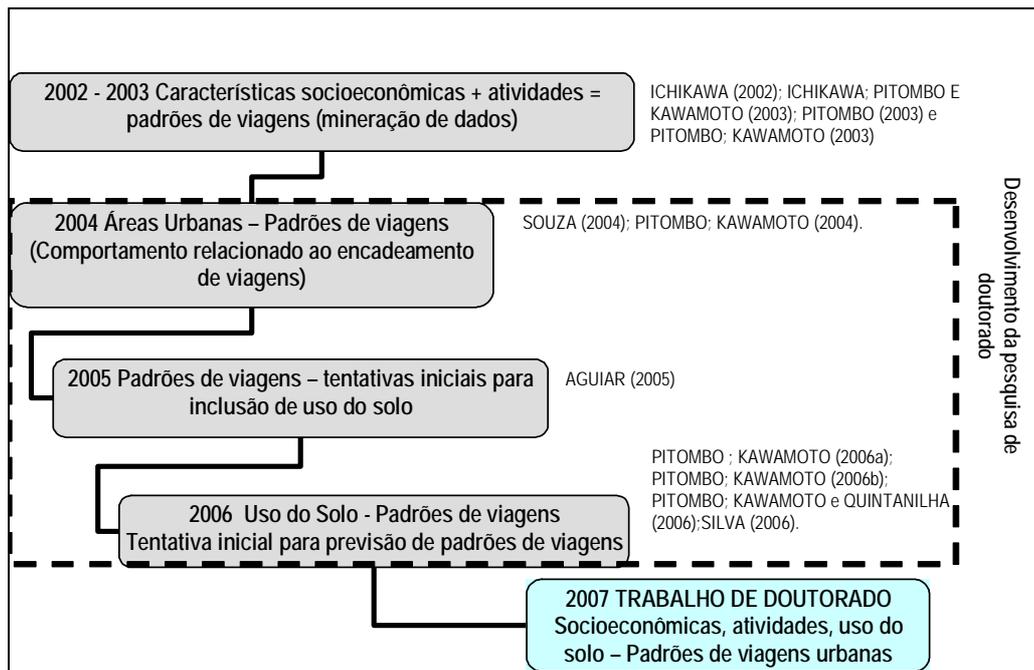


Figura 1.2 Linha de Pesquisa – Departamento de Transportes (EESC-USP)

### 1.3.2 Síntese do método

Há inúmeros conceitos sobre método científico. Abbagnano (1970) define método científico como sendo “um procedimento de investigação ordenado, repetível e auto-corrigível, que garanta obtenção de resultados válidos”. Desta forma, é adotada uma seqüência de procedimentos a fim de verificar a hipótese principal, a qual é enunciada da seguinte forma:

- É possível encontrar relações entre encadeamento de viagens urbanas (variável dependente) e características socioeconômicas, atributos do meio urbano (uso do solo) e participação das pessoas em atividades (variáveis independentes).

Esta subseção apresenta apenas uma síntese do método adotado (Figura 1.3). Cada etapa ilustrada na figura seguinte, sumariada inicialmente, corresponde aos capítulos subseqüentes do presente texto.

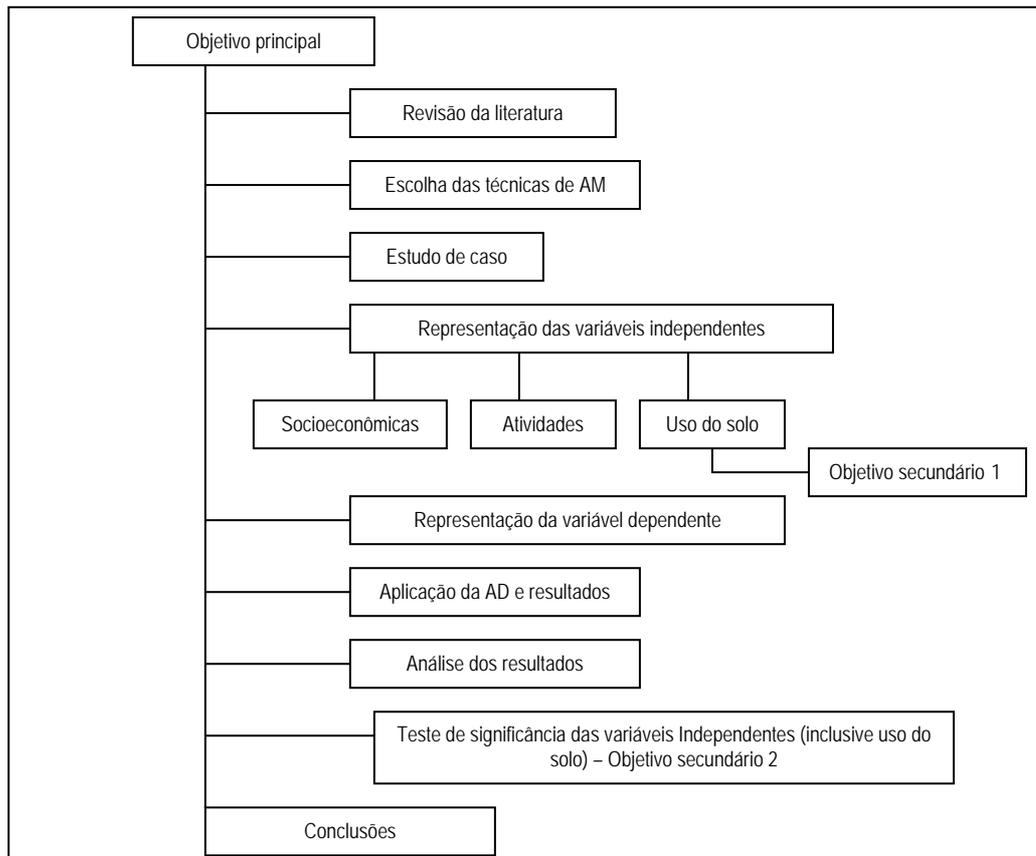


Figura 1.3 Síntese do método adotado

A primeira etapa é a revisão da literatura a respeito de uso do solo e comportamento relacionado ao encadeamento de viagens. Optou-se por dar ênfase maior à influência das características urbanas, já que o principal desafio está associado à representação do conjunto de variáveis de uso do solo. Fatores como participação em atividades e características socioeconômicas foram bastante explorados em trabalhos anteriores (PITOMBO, 2003; ICHIKAWA, 2002).

O próximo passo consiste na escolha das técnicas de Análise Multivariada de dados (AM), considerando os objetivos do trabalho. As técnicas são aplicadas tanto para representar a variável dependente adequadamente (padrões de viagem), como para encontrar relações entre estas e as variáveis independentes. São utilizadas três técnicas de AM conjuntamente: Análise de *Cluster* (AC), Árvore de Decisão (AD) e Regressão Linear Múltipla (RLM).

A seguir, observa-se o estudo de caso (RMSP) e separação de amostras para viabilizar a aplicação das técnicas. Vale ressaltar que nesta pesquisa trabalha-se com diferentes amostras (seis no

total), e não com a população da RMSP. Os dados da amostra não estão expandidos. As amostras utilizadas foram segregadas por setor de emprego (no caso de trabalhadores) ou grau de instrução (no caso de estudantes) com o intuito de melhor explicar a contribuição do uso do solo no comportamento relacionado ao encadeamento de viagens.

Em seguida são adotadas ou propostas, baseando-se na literatura vigente e na observação dos deslocamentos dos indivíduos residentes na RMSP, as variáveis independentes. A variável dependente é codificada em termos de seqüência de motivos de viagem, modo de transportes e destinos. A escolha adequada das variáveis independentes e da variável dependente é um dos pontos essenciais para obtenção de resultados suficientemente satisfatórios. Para auxílio na representação das variáveis foi feita a aplicação da técnica de AC.

Enfim, é feita aplicação da técnica AD para encontrar relações entre as variáveis independentes e variável dependente, escolhidas e representadas nos passos anteriores. É realizada posteriormente a análise dos resultados obtidos e o teste de significância das variáveis independentes, sobretudo daquelas relacionadas ao uso do solo. É aplicada uma técnica de Análise Multivariada tradicional (RLM) para obter dados a respeito da significância da variável. Optou-se por utilizar uma técnica tradicional, juntamente com a AD, pois técnicas de mineração de dados descobrem padrões em grande base de dados, elucidando várias questões. No entanto, não trazem valores estatísticos associados a cada variável. Associar técnicas estatísticas tradicionais a técnicas de mineração de dados (como AD, por exemplo) pode trazer maiores contribuições para o trabalho.

Vale ressaltar o caráter exploratório da pesquisa, devido principalmente à dificuldade de propor o grupo de variáveis de uso do solo, de forma que este fosse considerado representativo para segregação dos dados pelo algoritmo CART (técnica de AD) e também significativo estatisticamente.

#### **1.4 JUSTIFICATIVA**

Na maioria das vezes, as atividades que indivíduos desempenham durante o dia não são realizadas no mesmo local, levando a uma seqüência de deslocamentos no espaço urbano ao longo

do dia, ou seja, o encadeamento de viagens. A análise dos deslocamentos como resultado da seqüência de atividades exercidas fora do domicílio implica uma expansão no âmbito analítico, pois ela não é limitada apenas a viagens independentes. Este fato leva a maiores exigências em termos de dados e maior complexidade na análise. As vantagens oferecidas, em particular a habilidade de superar as limitações dos modelos convencionais, são de grande importância para o Planejamento de Transportes, prevalecendo sobre as desvantagens (ICHIKAWA; PITOMBO; KAWAMOTO, 2003).

O prosseguimento do estudo relacionado ao encadeamento de viagens possibilita uma melhor compreensão da diversidade de padrões de atividades e viagens realizados pelos indivíduos. A inclusão de variáveis de uso do solo vem enriquecer a análise, baseando-se no fato de que as viagens realizadas pelos indivíduos são afetadas pela maneira como as atividades encontram-se distribuídas geograficamente. Assim, este trabalho busca maior fundamentação no que se refere à investigação da escolha de padrões de viagem, já que considera a programação diária dos indivíduos em termos de necessidades, características individuais e das vizinhanças e seqüência de deslocamentos.

A principal justificativa seria aprimorar este tipo de estudo, que atualmente tem um caráter exploratório, para que possam ser utilizados futuramente modelos mais precisos que representem adequadamente a forma como os indivíduos se comportam e o que consideram para organizar seus respectivos itinerários de viagens. O presente trabalho faz parte de uma linha de pesquisa que aspira contribuir em estudos investigativos relacionados ao comportamento subjacente a viagens urbanas.

## 1.5. QUESTÕES FUNDAMENTAIS

Este estudo é uma tentativa de melhor compreender a forma como os indivíduos realizam a seqüência de viagens durante o período de um dia. Esta análise visa responder algumas indagações:

- Como mensurar padrões de uso do solo e caracterizar as zonas de tráfego da RMSPP em termos de intensidade e distribuição de atividades no meio urbano?
- Como representar características urbanas (uso do solo) em termos de variáveis independentes?

- Como o comportamento referente a viagens encadeadas está relacionado aos padrões de uso do solo dos locais de residência das pessoas?
- Quais os tipos predominantes de padrões de viagem associados a determinados grupos de indivíduos, conhecidas suas características socioeconômicas e dos seus locais de residência?
- Quais as vantagens de segregar amostras por “Ocupação do indivíduo” ou “Grau de instrução” com o intuito de melhor compreender a influência das variáveis independentes, especialmente daquelas de uso do solo?
- Como testar estatisticamente as variáveis de uso do solo adotadas?
- As técnicas de AM escolhidas, utilizadas em conjunto, são úteis considerando os objetivos do trabalho?
- De que forma estudos como estes podem trazer benefícios para o meio acadêmico, assim como contribuir futuramente para aplicações práticas?

## 1.6. ESTRUTURA DO TRABALHO

O presente texto é constituído de nove capítulos, além dessa introdução, conforme ilustrado na Figura 1.3, e descrito abaixo:

- i. O Capítulo 2 traz a etapa de revisão da literatura a respeito de uso do solo. Também mostra sucintamente modelos de distribuição de viagens, especialmente o modelo de oportunidades intervenientes, que foi considerado a fim de propor o conjunto de variáveis de uso do solo;
- ii. O Capítulo 3 apresenta fundamentação teórica das técnicas de AM utilizadas;
- iii. O Capítulo 4 constitui o estudo de caso, trazendo informações a respeito da RMSP e das amostras utilizadas;
- iv. O Capítulo 5 apresenta as variáveis independentes adotadas, dando maior ênfase às variáveis de uso do solo (principal desafio do trabalho);
- v. O Capítulo 6 mostra a codificação da variável dependente em termos de seqüência de atividades, modo de transporte e destinos;

- vi. O Capítulo 7 traz a aplicação da técnica de AD e os resultados obtidos para as diferentes amostras;
- vii. O Capítulo 8 apresenta a análise dos resultados obtidos com o processamento das árvores;
- viii. O Capítulo 9 traz o teste de significância das variáveis independentes, especialmente daquelas relacionadas ao uso do solo (RLM);
- ix. Finalmente, o Capítulo 10 traz uma síntese do trabalho, bem como as principais conclusões obtidas, objetivando responder as indagações propostas nesta introdução. Também são apresentadas sugestões para trabalhos futuros.

## Capítulo

# 2 USO DO SOLO X PADRÕES DE VIAGENS

---

*Este capítulo traz uma revisão da literatura a respeito do uso do solo e comportamento relacionado a viagens. É dividido em três partes principais que constituem a bibliografia pesquisada: (1) Uso do solo e comportamento relacionado a viagens; (2) Mensuração de variáveis de uso do solo; e (3) Modelos de distribuição de viagens. Na última parte do presente capítulo encontram-se conceitos a respeito de modelos de distribuição de viagens, dando ênfase ao modelo de oportunidades intervenientes, que fornece suporte teórico para a proposta do conjunto de variáveis de uso do solo deste trabalho.*

### 2.1 CONSIDERAÇÕES INICIAIS

A despeito da grande quantidade de trabalhos e estudos a respeito do tema, há um consenso geral em relação à investigação das escolhas das seqüências de viagens realizadas pelos indivíduos: *o estudo do comportamento relacionado ao encadeamento de viagens é de grande complexidade.* Analisar o que realmente afeta decisões em relação à seqüência de viagens é uma tarefa desafiadora.

Em termos de uso do solo, deve-se atentar para o fato de que uma vizinhança com alta acessibilidade nem sempre interfere nas escolhas dos itinerários, pois o comportamento de viagem pode ser determinado não apenas por localizações residenciais ou das vizinhanças, mas também através da localização dos centros de atividades. "Pessoas que moram em zonas periféricas podem, por exemplo, dirigir até centros de atividades de alta acessibilidade e fazer viagens a pé em tais zonas para realizar diferentes atividades" (EWING; HALIYUR; PAGE, 1994).

As pessoas geralmente possuem a tendência de combinar várias localizações em uma cadeia de viagens. A Figura 2.1, a seguir, ilustra um padrão de viagem realizado por um indivíduo que mora em um bairro periférico e utiliza o automóvel para ir ao trabalho. A zona onde se localiza o trabalho, por sua

vez, possui alta atratividade, então são realizadas viagens com múltiplos propósitos naquela zona, todas a pé. Finalmente, a última viagem (de retorno ao domicílio) é realizada com o uso do automóvel.

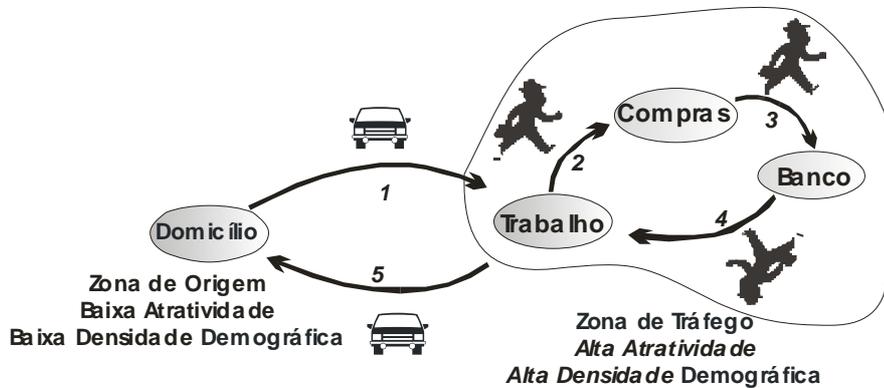


Figura 2.1 Exemplo de padrão de viagem e acessibilidade relativa entre as zonas urbanas

Considera-se um exemplo adaptado de Krizek (2003), onde se observam duas situações diferentes. Judy e Johnny residem em localizações com alta acessibilidade. Judy realiza diversas atividades durante a sua viagem de casa ao trabalho, cada uma das atividades localiza-se nas proximidades do trabalho. Apesar de morar em uma área que oferece muitos serviços, Judy acha mais conveniente realizar diversas atividades na região de seu trabalho. Em seguida, considera-se o caso de Johnny, que dirige uma vez por semana de casa até a lavanderia. A decisão de uso de automóvel para realizar tal atividade é tomada porque a lavanderia situa-se no caminho do mercado que Johnny vai uma vez por semana e localiza-se numa região mais distante. Em ambos os casos, o fato de morar em regiões de alta acessibilidade (com diversas atividades e serviços) não modificou suas decisões de realizarem atividades fora das suas vizinhanças. Isto ocorre devido ao fato de que muitos fatores são levados em conta durante a combinação das viagens realizadas pelos indivíduos, no período de um dia.

Intuitivamente, pode-se dizer que não apenas características de uso do solo das suas vizinhanças ou locais de trabalho influenciam o comportamento referente ao encadeamento de viagens. A distribuição da rede de transporte público, assim como a disponibilidade de modos de transporte e características socioeconômicas individuais contribuem substantivamente para este comportamento.

Este capítulo revisa alguns trabalhos que abordaram o tema “Uso do solo x Comportamento de viagens”, considerando o escopo principal da presente pesquisa - que visa propor um conjunto de variáveis de uso do solo e encontrar relações entre grupos de variáveis independentes ((1) uso do solo; (2) características socioeconômicas; (3) participação em atividades) e variável dependente (padrões de viagens urbanas).

Revisão da literatura a respeito da influência dos dois últimos grupos de variáveis independentes nos padrões de viagens urbanas foi apresentada no capítulo introdutório e será discutida ao longo do trabalho de forma sucinta, pois se discutiu largamente a respeito do tópico em trabalhos prévios que fazem parte da linha de pesquisa desenvolvida no Departamento de Transportes da EESC-USP (ICHIKAWA, 2002; PITOMBO; 2003; PITOMBO; KAWAMOTO; 2003; PITOMBO; KAWAMOTO; 2004).

O presente capítulo divide-se em três seções principais:

- **Uso do solo e comportamento relacionado a viagens:** discute sinteticamente a respeito da evolução das cidades (citando algumas teorias urbanas), da influência de estruturas urbanas no comportamento relacionado a viagens, assim como o eterno debate sobre a existência de relação entre uso do solo e transportes;
- **Mensuração de variáveis de uso do solo:** traz diferentes medidas ou indicadores de características de uso do solo propostos em trabalhos anteriores; e
- **Modelos de distribuição de viagens:** descreve sucintamente modelos de distribuição de viagens sintéticos, dando ênfase ao modelo de oportunidades intervenientes, o qual foi suporte teórico para a proposta do conjunto de variáveis de uso do solo do presente trabalho.

## 2.2 USO DO SOLO E COMPORTAMENTO RELACIONADO A VIAGENS

A idéia de que políticas de uso do solo podem influenciar o comportamento relacionado a viagens é proveniente da observação da correlação entre tais variáveis. Durante décadas, os modelos tradicionais de previsão de demanda por transportes previam o número de viagens originadas em uma determinada área

geográfica (zona de tráfego) ou o número de viagens atraídas à outra área geográfica, considerando variáveis que incluíam população e empregos por zonas de tráfego. No entanto, as decisões individuais acerca das viagens a serem realizadas fazem parte de uma análise mais complexa, que envolve muitos outros fatores além da densidade populacional e de empregos (SARMIENTO, 1996).

Características como distribuição das atividades no meio urbano e facilidade de acesso a determinadas atividades, possivelmente, interferem nas escolhas dos diferentes destinos. O meio onde os indivíduos residem, ou trabalham, intuitivamente, exerce influência no processo de tomada de decisões para realização de diversos padrões de viagens urbanas. No entanto, a relação entre uso do solo e transportes (e, conseqüentemente, comportamento relacionado a viagens) tem sido debatida durante anos, havendo divergências entre diversos autores. O que se discute é a casualidade ou não de tais relações (HILLIER, 1996; GORDON; RICHARDSON, 1997). Esta seção traz alguns aspectos importantes a respeito do tema: (1) evolução da estrutura urbana; (2) a estrutura urbana e comportamento relacionado a viagens; e (3) debate da existência ou não da relação entre uso do solo e comportamento referente a viagens.

### **2.2.1 Evolução da estrutura urbana: algumas teorias sobre formas urbanas**

O surgimento das cidades deveu-se à necessidade de atender a alguns critérios funcionais ligados a idéia de abrigo. Atualmente, definem-se cidades como sistemas complexos, envolvendo aglomerados humanos, compostos por múltiplas atividades, situadas em espaços distintos (SILVA, 1993).

A história do desenvolvimento dos núcleos urbanos está diretamente relacionada ao sistema de transportes e à evolução da sua infra-estrutura. Assim, é notória a relação intrínseca entre estruturas urbanas e demanda por viagens, bem como a necessidade de coordenação entre políticas de uso do solo e Planejamento de Transportes.

Diversos estudiosos e urbanistas desenvolveram teorias a respeito de diferentes estruturas urbanas considerando configurações das cidades em relação à distribuição de atividades no seu meio.

Uma das teorias urbanistas é a teoria da zona concêntrica (Burgess's *concentric zone theory*), que descreve um padrão de cidade com um centro de negócios principal, rodeado por uma zona industrial. Esta, por sua vez, é envolta em anéis de áreas residenciais que variam de rendas baixas a altas (BROWN, 1992).

Uma outra teoria, desenvolvida por Alonso (1964), pressupõe que acessibilidade seja fundamental para formação da estrutura urbana. Assim, a máxima acessibilidade é dada ao centro da cidade, onde se localizam as atividades principais. A acessibilidade das zonas industriais e residenciais é medida em relação ao centro, adquirindo menores valores ao aproximar-se da periferia. Esta teoria também constitui um padrão de uso do solo concêntrico.

Uma tendência norte-americana, apontada por Blakely (1994), é a formação de subúrbios, onde se observa uma reorganização da localização de opções de empregos e oportunidades. Pivo (1990) faz uma analogia entre uma nova teoria de estrutura urbana e um colar de pérolas. Cada pérola equivale a um centro, e muitas vezes, realiza atividades centrais da antiga estrutura de cidade. O colar significa a ligação entre os centros pela mesma via.

Apesar das diversas teorias, a estrutura da maioria das grandes cidades atualmente é composta por muitos centros, devido à descentralização de oportunidades de empregos e outras atividades. O crescimento urbano, muitas vezes desordenado, faz com que cidades adquiram diferentes formas. A distribuição das atividades no meio urbano e a forma como estas estão interligadas pela rede de transportes são essenciais para possíveis explicações de comportamento relacionado a viagens.

McNally e Kulkarni (1997) citam dois tipos de estrutura urbana: a primeira, que tem se tornado popular nas últimas décadas nos EUA, a fim de incentivar a redução do uso do automóvel, é denominada TND - *Traditional and Neotraditional neighborhood Design*, ou *Transit Oriented Design*; enquanto a segunda, que promove a dependência do automóvel, devido às suas características é denominada PUD - *Planned Unit Developments*.

Os mesmos autores enumeram diferentes características inerentes a cada uma das duas estruturas citadas (Tabela 2.1). Muitos pesquisadores acreditam que soluções estratégicas de uso do solo, como implantação de características de estruturas urbanas TND (*Design-oriented Solutions*) podem efetivamente reduzir congestionamentos, por exemplo.

**Tabela 2.1 Características de uso de solo de duas estruturas urbanas (PUD e TND)**

|                 | PUD                                                                                                            | TND                                                                                                                |
|-----------------|----------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------|
| Características | Uso do solo bem definido e separado por zonas<br>Zonas residenciais com baixa densidade<br>Grandes loteamentos | Grande quantidade de zonas de uso do solo misto<br>Zonas residenciais com alta densidade<br>Pequenos loteamentos   |
| de uso do solo  | Acesso a um número limitado de atividades<br>Grande presença de subúrbios<br>Cidades típicas norte-americanas  | Fácil acesso à vizinhança<br>Maior dispersão de atividades no meio urbano<br>Presença de muitos centros comerciais |

### 2.2.2 A estrutura urbana e comportamento relacionado a viagens

O comportamento referente a viagens é função da necessidade individual de realização de atividades, distribuídas nas cidades, determinando desta forma o padrão de viagem a ser escolhido pelo indivíduo, dentro de um conjunto de opções e restrições. A disposição geográfica dos locais das atividades (comércio, serviço, educação, etc.) nos meios urbanos determina a maior ou menor facilidade com que o indivíduo pode ter acesso a elas e realizá-las diariamente (ARRUDA, 2005).

A fim de ilustrar a interdependência entre estrutura das cidades, dispersão de atividades no meio urbano e comportamento dos indivíduos relacionado a viagens, considera-se a Figura 2.2.

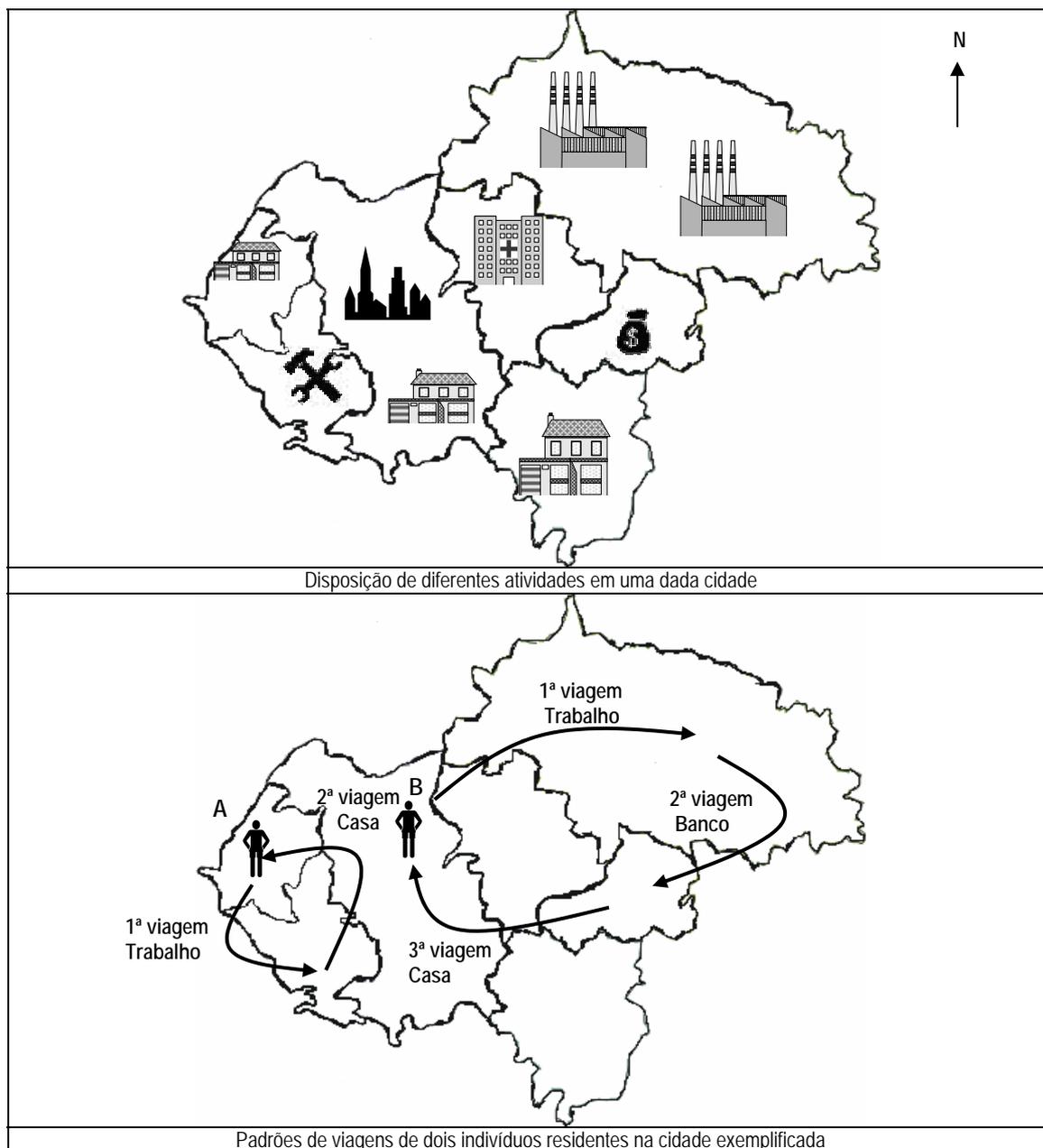


Figura 2.2 Distribuição de atividades em uma cidade fictícia e padrões de viagens

A Figura 2.2 ilustra inicialmente a disposição de diversas atividades em uma cidade fictícia (indústrias, consultórios médicos, serviços bancários, residências, etc.) e, posteriormente, padrões de viagens realizados por indivíduos residentes na referida cidade.

O indivíduo A realiza um padrão de viagem simples, apenas duas viagens, a primeira viagem com motivo trabalho e a segunda e última viagem com motivo residência. Esta pessoa mora relativamente próxima ao seu trabalho e realiza viagens curtas.

Já o indivíduo B realiza uma cadeia composta por três viagens. A primeira viagem, ao trabalho, é longa. O indivíduo B reside em uma zona residencial e trabalha na indústria. Na fictícia cidade há apenas uma zona de tráfego industrial localizada na região norte. O motivo da segunda viagem no dado exemplo seria a realização de serviços bancários. Como o banco localiza-se mais próximo à zona industrial, a viagem ao banco é realizada após o trabalho. Finalmente, a última viagem é de retorno ao domicílio.

Possivelmente, caso a disposição de determinadas atividades fosse diferente, os itinerários de viagens escolhidos pelos indivíduos poderiam variar, seja na escolha dos destinos ou modos de transportes utilizados, ou até mesmo nas escolhas a longo prazo, como localizações residenciais.

As cidades, ao longo de toda a sua história, têm sido locais de concentração de inúmeras atividades. Com o objetivo de realizar as diferentes atividades distribuídas no meio urbano, os indivíduos programam o seu itinerário de viagens, ou seja, a seqüência de viagens a serem realizadas durante o dia, levando em conta a localização relativa da residência e um conjunto de oportunidades.

### **2.2.3 Uso do solo e transportes: fortes relações ou casualidade?**

O transporte, incluindo o sistema viário e todas as redes de infra-estrutura, influencia a localização das diversas atividades nas cidades. Por sua vez, a ocupação do solo (densidades, dimensões das zonas, etc.), o uso do solo (tipos de uso, localização das zonas, vazios urbanos, etc.) são características da estrutura urbana que influenciam o sistema de transportes (ASSIS Jr, 1987). A interligação entre estes dois fatores, pode ser reforçada, em termos desagregados (individuais) devido ao fato de que demanda por viagens é derivada da necessidade de realização de atividades fora do domicílio. Assim, considera-se a disposição das atividades e o modo de transporte mais conveniente para se ter acesso a determinados destinos.

Wegener e Fürst (1999) representam a inter-relação entre uso do solo e transportes através de um ciclo de realimentação uso do solo - transportes (ilustrado na Figura 2.3), definido da seguinte forma:

1) Uso do solo: a distribuição do uso do solo (residencial, comercial, etc.) afeta a localização das atividades humanas (compras, lazer, educação, residência);

2) Atividades: por sua vez, a dispersão das atividades no meio urbano requer a realização de viagens nos sistemas de transportes;

3) Sistemas de transportes: a distribuição da rede de transportes cria oportunidade para deslocamentos, aumentando, assim, a acessibilidade de determinadas regiões;

4) Acessibilidade : as diferentes acessibilidades no meio urbano determinam decisões individuais tomadas a longo prazo, como localizações residenciais, por exemplo, alterando a distribuição de uso do solo.



Figura 2.3 Ciclo de realimentação transportes – uso do solo proposto por Wegener e Fürst (1999)

Ainda que uma veracidade intuitiva possa ser creditada a esta relação, há inúmeras discussões acerca do tema. Alguns autores acreditam que tais relações sejam casuais e que as variáveis socioeconômicas expliquem melhor a variância no comportamento de viagens do que aquelas variáveis que caracterizam uso do solo (KITAMURA; MOKHTARIAN; LAIDET, 1997; GIULIANO, 1995). Apesar das muitas discussões, o presente trabalho procura sustentar a hipótese de que uso do solo seja um dos fatores que interfiram no comportamento relacionado ao encadeamento de viagens. Nesta subseção são citados trabalhos que corroboram a forte relação existente entre comportamento

subjacente a viagens e uso do solo, assim como outros trabalhos encontrados na literatura onde não se pode afirmar real ligação entre estes dois fatores.

- *Há relações entre uso do solo e comportamento dos viajantes: trabalhos pontuais consultados na literatura*

Considerando relações entre densidades urbanas e comportamento de viagens, Newman e Kenworthy (1989) compararam diferentes cidades do mundo com diferentes densidades populacionais e seus comportamentos referentes a viagens. Os autores concluíram que cidades com mais altas densidades populacionais são menos dependentes do automóvel do que aquelas com baixas densidades demográficas.

Frank e Pivo (1994) analisaram o uso de transporte público para viagens relacionadas ao trabalho e viagens com motivo compras. Identificaram uma relação negativa entre densidade de empregos, densidade populacional e uso do solo misto e uso de automóveis, assim como uma relação positiva foi encontrada entre as mesmas variáveis e uso de transporte público e modo de transporte não motorizado.

A fim de simular efeitos de mudanças na forma urbana, aplicando características de estruturas TND (Tabela 2.1), estudos realizados pelo Middlesex-Somerset-Mercer Regional Council (MSMRC) identificaram fatores como uso do solo misto, razão empregos por residências balanceada e altas densidades demográficas como sendo fundamentais para redução de viagens por automóvel.

Friedman, Gordon e Peers (1994) compararam áreas de altas densidades, uso do solo misto e alta conexão da rede de transporte público (características de TND design) com comunidades tradicionais (PUD). Concluíram com seus resultados que estruturas urbanas com características de TND possuem menores proporções de uso de automóveis (inclusive motorista sozinho) e maior percentual de viagens utilizando modo de transporte público.

Cervero e Radisch (1996) concluíram que indivíduos que residem em regiões mais compactas, com uso do solo misto e facilidades para pedestres, em média, realizam com mais frequência viagens (que não sejam relacionadas ao trabalho) utilizando modo de transporte não motorizado ou transporte público, quando comparados àqueles que moram em típicos subúrbios norte-americanos.

Handy (1996) acredita que a forma urbana pode afetar o comportamento de viagens, caso os indivíduos percebam que modos de transporte não motorizados possam ser uma alternativa viável para eles, considerando um conjunto de opções.

Handy (1993) afirma que altos níveis de acessibilidade nas vizinhanças hipoteticamente influenciam viagens não relacionadas ao trabalho. Em algumas regiões metropolitanas dos EUA, há a suposição de que indivíduos escolham localizações residenciais próximas ao trabalho. Assim, os padrões de viagens com motivo trabalho são muitas vezes fixos (pois tendem a possuir restrições como horários ou rotas), sendo que tais viagens são tão curtas devido à proximidade entre residência e trabalho. No entanto, viagens não relacionadas ao trabalho são mais flexíveis e potencialmente influenciadas pelo nível de acessibilidade da vizinhança. Outros estudos fizeram distinções entre viagens relacionadas ao trabalho e aquelas não relacionadas a atividades de trabalho a fim de investigar influências das características das vizinhanças dos indivíduos (CERVERO; KOCKELMAN, 1997; EWING, 1995).

- *Não se pode afirmar necessariamente que existam relações entre uso do solo e comportamento dos viajantes: alguns trabalhos revisados na literatura*

Miller e Ibrahim (1998) testaram a relação entre forma urbana e viagens com motivo trabalho. Os autores não encontraram evidências de que uma distribuição balanceada entre empregos e moradias reduziria as distâncias de viagens.

Giuliano (1995) afirma que a relação entre uso do solo e transportes seja possivelmente fraca, assim como não acredita que escolhas de padrões de viagens sejam fortemente influenciadas pela

estrutura urbana e mudanças de uso do solo. O autor ainda supõe que residentes em diferentes meios urbanos (como TND e PUD, por exemplo), possuem diferentes comportamentos relacionados a viagens devido a fatores socioeconômicos e características particulares.

McNally e Kulkarni (1997) desenvolveram uma avaliação empírica a respeito da interação entre uso do solo/sistemas de transportes e comportamento relacionado a viagens. Utilizando técnica de Análise de *Cluster* (que será detalhada no Capítulo 3), um total de vinte cenários de Orange County (Califórnia) foi agrupado em três *clusters* (temas). Medidas convencionais associadas ao comportamento de viagens foram comparadas através de uma análise de variância (ANOVA). Os resultados obtidos enfatizam a idéia da fraca relação entre uso do solo e comportamento de viajantes. No entanto, os autores sugerem a realização de estudos adicionais para verificar se esta seria uma característica inerente ao sul da Califórnia. Assim, é recomendada cautela ao se afirmar que soluções estratégicas de uso do solo (*Design-oriented Solutions*) sejam alternativas à redução de congestionamentos.

Boarnet e Sarmiento (1996) ratificam a hipótese da possível não existência da relação entre uso do solo e comportamento relativo a viagens. Embora os planejadores tendam a aplicar políticas de uso do solo para gerenciamento da demanda por transportes, os autores afirmam que não há evidências de tais relações. Viagens não relacionadas ao trabalho (viagens por automóvel e quilômetros viajados por automóvel) foram modeladas em função de variáveis socioeconômicas individuais e características de uso do solo de localizações próximas às residências dos indivíduos. Resultados mostraram que raramente as variáveis de uso do solo são estatisticamente significativas.

Kitamura, Mokhtarian e Laidet (1997) examinaram os efeitos de características de uso de solo e variáveis socioeconômicas (definidas pelos autores como *attitudinal variables*) no comportamento de viajantes em cinco regiões pertencentes à área da baía de São Francisco (EUA). Embora cada grupo de variáveis estudadas tenha oferecido certo poder explicativo no modelo, as variáveis socioeconômicas explicavam alta proporção da variabilidade dos dados. A conclusão de que variáveis

socioeconômicas estão mais fortemente associadas a viagens sugere que políticas de uso do solo que promovem altas densidades e uso do solo misto podem não alterar significativamente a demanda por viagens, a menos que ocorram mudanças nas características individuais (*attitudinal variables*).

Apesar do eterno debate entre a existência de relações entre uso do solo e transportes, sabe-se que muitas vezes resultados indicam a inexistência de relações entre ambos os fatores devido a uma série de questões. Deve-se atentar primeiramente para o contexto ou o cenário em que foi realizado cada estudo. Alto nível de empregos em todas as zonas de uma determinada cidade, por exemplo, certamente não afetará escolhas de destinos. Muitas vezes observou-se também que existiam relações entre uso do solo e comportamento relacionado a viagens, no entanto, características socioeconômicas sobressaíam-se, sendo mais importantes na análise do comportamento de viajantes urbanos. Outro aspecto importante é a não representação adequada das variáveis envolvidas. Cervero e Kockelman (1997) notaram que variáveis ordinais ou *dummy* são utilizadas muitas vezes para indicar características de vizinhança (uso do solo) e podem não estar representando apropriadamente tais características.

A representação adequada de variáveis de uso do solo é um dos pontos chaves para obtenção de resultados significativos. Propor variáveis que expliquem características do meio urbano é parte fundamental no processo de investigação de inter-relações entre uso do solo e comportamento de viajantes. Mensurar características de uso do solo da RMSP, objetivo secundário do trabalho, constitui-se um desafio, caracterizando a atual pesquisa como exploratória. Na próxima seção, serão citadas algumas propostas de representação de características de uso do solo em alguns trabalhos que fazem parte da literatura a respeito do tema.

### 2.3 MENSURAÇÃO DE VARIÁVEIS DE USO DO SOLO

Para entender relações entre uso do solo e comportamento relativo a viagens é necessário, primeiramente, investigar formas de mensurar características espaciais das cidades. Atributos espaciais ou características de uso do solo podem representar distribuição de atividades como densidade de empregos,

acessibilidade a serviços diversos como compras, lazer, ou características locais como forma ou localização da zona de tráfego, padrões das ruas, configuração da rede, etc.

Muitos estudos utilizaram indicadores gerais, como dados censitários a respeito de população e empregos a fim de representar características espaciais. Foram amplamente utilizadas características como densidade demográfica, densidade de empregos ou razão empregos por residências, por exemplo.

Golob e Brownstone (2005) analisaram o impacto de densidade residencial no uso veicular e consumo de energia (combustível). Características de uso do solo de regiões do Estado da Califórnia (EUA) foram representadas através de densidade populacional em termos de zona de tráfego ou quadras. A pesquisa do ano de 2001 fornecida pelo *U.S. National Household Transportation Survey* (NHTS) continha muitas medidas de uso do solo relacionadas às localizações domiciliares.

A variável “densidade demográfica” (população/milhas ao quadrado) foi classificada em oito categorias (<100 indivíduos; 100 – 500 indivíduos; 500 – 1000 indivíduos; 1000 a 2000 indivíduos; 2000 a 4000 indivíduos; 4000 a 10000 indivíduos. 10000 a 25000 indivíduos; > 25000 indivíduos). A densidade em termos de domicílios foi representada em seis categorias (< 50 domicílios; 50 a 250 domicílios; 250 a 1000 domicílios; 1000 a 3000 domicílios; 3000 a 5000 domicílios; > 5000 domicílios).

Levinson e Kumar (1997) avaliaram a influência de densidade populacional no comportamento relativo a viagens, com motivo relacionado ao trabalho. Foram utilizadas medidas de densidades metropolitanas e locais em algumas cidades dos EUA. Densidades metropolitanas delimitam o tamanho das cidades enquanto que densidades locais mensuram localizações relativas dentro das cidades.

Outros estudos utilizaram índices que mediam configuração da rede, acessibilidade ou medidas qualitativas, tais como facilidades para pedestres ou uso de transporte público. Handy (1996) utilizou variáveis que representavam características da rede viária (densidade das ruas, tipos de interseções, número de vias arteriais) e características que representavam distribuição de atividades no meio urbano

(número de diferentes tipos de estabelecimentos comerciais e acessibilidade de cada região a centros regionais, lojas de departamento ou supermercados).

Moudon et al. (1997) estudaram efeitos de desenho urbano nas viagens realizadas por pedestres. Eles mensuraram facilidades dos pedestres através da extensão e distribuição de calçadas que são protegidas do trânsito veicular (ruas exclusivas para pedestres), bem como facilidades físicas (qualidade do caminho – superfície em bom estado; segurança das vias, etc.).

Evans, Perincherry e Douglas (1997) propuseram um índice que quantificava o acesso ao transporte público (*Transit Friendliness Factor*). Tal índice era medido em função das características da área ao redor dos terminais de transporte público. Estas características indicavam a qualidade do acesso dos pedestres (estado das calçadas) aos terminais, atributos das vias das proximidades, presença de comércio ou outros serviços nos terminais e proximidade dos potenciais destinos.

Outros trabalhos utilizaram medidas mais elaboradas derivadas de características gerais de uso do solo das regiões estudadas. Verifica-se o uso de técnicas de Análise Multivariada (descritas no Capítulo 3) para representação de tais variáveis.

McNally e Kulkarni (1997) agruparam vinte regiões pertencentes a Orange County (Califórnia) em três conglomerados que possuíam características de estruturas TND e PUD (Tabela 2.1) e MISTAS (regiões que possuíam características de ambas estruturas urbanas). Através da aplicação da Análise de *Cluster* (técnica de agrupamento que será posteriormente definida), as regiões foram agrupadas em três temas pré-definidos. Neste caso já havia uma tipologia, a qual faz parte da hipótese da pesquisa que sugeria que características de uso do solo poderiam ser representadas através de três categorias (estruturas urbanas): (1) TND; (2) PUD; (3) MIX (mista).

Utilizando a mesma técnica de Análise Multivariada (Análise de *Cluster*), Pitombo, Kawamoto e Quintanilha (2006) propuseram uma caracterização de uso do solo alternativa para RMSP. As variáveis adotadas para o agrupamento caracterizavam distribuição de atividades na região (Taxa de

industrialização, comércio, serviços, matrículas escolares e densidades populacionais), bem como localização das zonas de tráfego (latitude e longitude dos centróides das zonas). Foram obtidas 10 macro-zonas caracterizadas em função do nível de atividade, localização e população. A Figura 2.4 representa as localizações das macro-zonas obtidas e a Tabela 2.2 traz suas respectivas características em termos de valores médios, assim como a nomenclatura adotada pelos autores, considerando os critérios de localização, nível de atividade e densidade populacional.

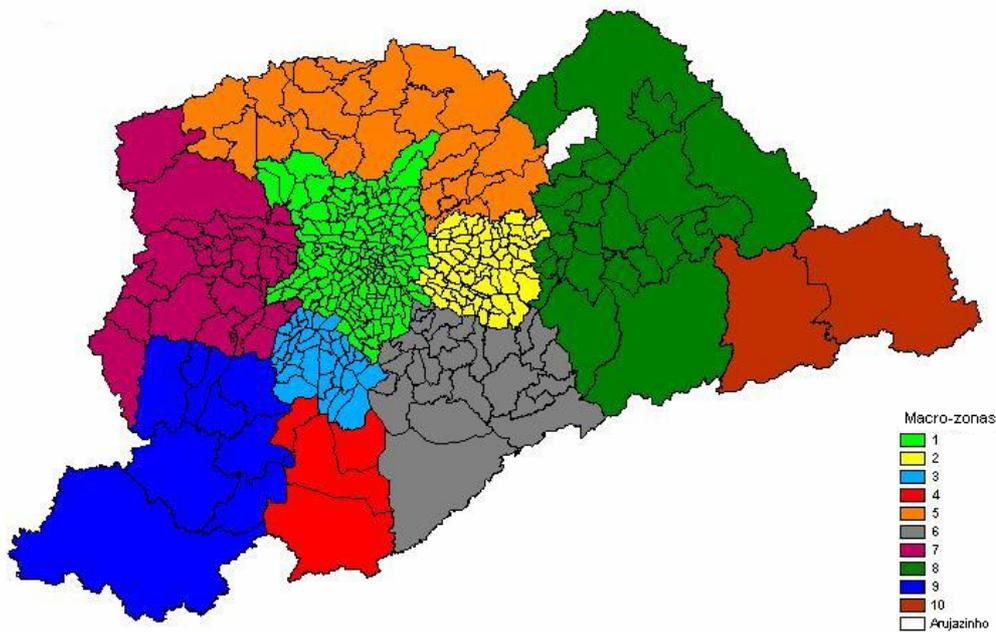


Figura 2.4 Distribuição das macro-zonas na RMSP

Tabela 2.2 Macro-zonas e Nomenclatura adotada

| Mz | Nº zonas | Municípios                                                                                                                         | Densidade | Taxa Ind | Taxa Com | Taxa Serv | Taxa ab2gr | Taxa ac2gr | Sub Região | Nomenclatura             |
|----|----------|------------------------------------------------------------------------------------------------------------------------------------|-----------|----------|----------|-----------|------------|------------|------------|--------------------------|
| 1  | 167      | São Paulo                                                                                                                          | 16846,97  | 36,33    | 47,08    | 52,89     | 27,22      | 60,51      | Centro     | AltissDensAltissAtiv     |
| 2  | 65       | São Paulo                                                                                                                          | 8079,78   | 9,72     | 14,04    | 10,12     | 18,98      | 6,83       | Centro     | AltissDensAltaAtiv       |
| 3  | 34       | São Paulo                                                                                                                          | 3077,12   | 8,03     | 8,00     | 7,15      | 9,75       | 4,32       | Centro     | AltaDensModAtiv          |
| 4  | 4        | São Paulo                                                                                                                          | 9,40      | 0,23     | 0,33     | 0,22      | 0,46       | 0,01       | Centro     | BaixaDensBaixaAtiv       |
| 5  | 24       | Guarulhos, Caieiras, Mariporó, Cajamar, Francisco Morato, Franco da Rocha                                                          | 600,70    | 10,04    | 5,03     | 5,83      | 8,34       | 5,59       | Norte      | ModbaixaDensModbaixaAtiv |
| 6  | 27       | S Caetano do Sul, Sto André, S Bernardo Diadema, Mauá, Ríb Pires, R Gde da Serra                                                   | 1597,54   | 18,66    | 11,31    | 9,73      | 14,08      | 12,44      | Sudeste    | ModDensAltaAtiv          |
| 7  | 31       | Osasco, Taboão da Serra, Carapicuíba, Embu Barueri, Cotia, Jandira, Sta do Parnaíba, Itapevi, Pirapora do B Jesus, Vg Gde Paulista | 1262,40   | 10,82    | 8,53     | 8,79      | 12,81      | 4,79       | Oeste      | ModDensModAtiv           |
| 8  | 25       | F de Vasconcelos, Arujá, Itaquaquecetuba M das Cruzes, Poá, Suzano, Guarulhos, St Isabel, Guararema                                | 531,42    | 5,37     | 4,98     | 4,25      | 7,01       | 5,20       | Nordeste   | ModbaixaDensModbaixaAtiv |
| 9  | 9        | Embu Guapu, Itb da Serra, Cotia, S Lorengo da Serra, Juquitiba                                                                     | 23,12     | 0,71     | 0,61     | 0,82      | 1,15       | 0,23       | Sudoeste   | BaixaDensBaixaAtiv       |
| 10 | 2        | Biritiba Mirim, Salesópolis                                                                                                        | 0,82      | 0,08     | 0,09     | 0,19      | 0,20       | 0,08       | Leste      | BaixaDensBaixaAtiv       |

Legenda: Mz - Macrozona; Altiss - Altíssima; Dens - Densidade Populacional; Ativ - Atividade; Mod - Moderada; Modbaixa - Moderadamente baixa.

Srinivasan (2000) utilizou Análise Fatorial Exploratória a fim de agrupar algumas variáveis independentes que representavam características espaciais da região de Boston. As 50 variáveis iniciais foram agrupadas em 8 fatores. Cada um dos fatores possuía uma característica peculiar.

Desta forma, os fatores obtidos foram as novas variáveis de uso do solo do trabalho em questão. Encontram-se, a seguir, os fatores, a nomenclatura adotada para caracterizar cada fator, assim como o número de variáveis contidas em cada um deles:

- 1) Fator 1 (Características de áreas suburbanas – *Suburban character*): 18 variáveis;
- 2) Fator 2 (Acesso ao transporte público – *Transit access*): 7 variáveis;
- 3) Fator 3 (Proximidade da rodovia – *Highway proximity*) : 7 variáveis;
- 4) Fator 4 (Balanço entre uso do solo comercial e residencial – *Commercial residential balance*): 3 variáveis;
- 5) Fator 5 (Balanço entre uso do solo residencial e espaços abertos – *Open space residential balance*): 3 variáveis;
- 6) Fator 6 (Características físicas da rede viária – *Cul-de-sac design*) : 5 variáveis;
- 7) Fator 7 (Acessibilidade por automóvel a atividades diferentes de trabalho – *Non-work accessibility by auto*): 3 variáveis;
- 8) Fator 8 (Conveniência para pedestres – *Pedestrian convenience*): 4 variáveis.

## 2.4 MODELOS DE DISTRIBUIÇÃO DE VIAGENS

O objetivo desta seção é apresentar breve fundamentação teórica a respeito de alguns modelos sintéticos de distribuição de viagens. Conceitua-se o modelo gravitacional, o modelo de oportunidades intervenientes e, o mais recente, chamado modelo gravitacional – de oportunidades intervenientes. Maior ênfase será dada ao modelo de oportunidades intervenientes (SCHNEIDER, 1959), pois este foi o suporte teórico para a proposta do conjunto de variáveis de uso do solo da atual pesquisa.

Os modelos de distribuição de viagens têm como objetivo determinar os elementos  $T_{ij}$  da matriz O-D. Cada elemento  $T_{ij}$  representa o número de viagens entre um par de origem-destino  $ij$ . Vários modelos matemáticos foram desenvolvidos durante décadas, enquadrando-se basicamente em dois grupos principais: (1) modelos de fatores de crescimento; e (2) modelos sintéticos.

Os modelos sintéticos baseiam-se nas seguintes proposições (BRUTON, 1979):

- Os fatos que causam os movimentos (deslocamentos) devem ser entendidos antes da previsão de viagens futuras;
- Pode-se melhor entender as causas dos movimentos (deslocamentos) se estas se assemelham a certas leis de comportamento físico.

A seguir, apresentam-se três modelos de distribuição de viagens sintéticos.

#### 2.4.1 Modelo gravitacional

Modelos gravitacionais talvez sejam os mais populares dentre os modelos sintéticos e foram originalmente conceituados através de uma analogia à Lei da Gravitação Universal: “Duas partículas se atraem com uma força diretamente proporcional ao produto de suas massas e inversamente proporcional ao quadrado da distância entre elas”. Apresenta-se matematicamente pela expressão:

$$F_{ij} = \frac{\alpha m_i m_j}{d_{ij}^2} \quad (2.1)$$

onde:  $F_{ij}$  = Força;  $\alpha$  = Constante;  $m_i$  = massa da partícula i;  $m_j$  = massa da partícula j;  $d_{ij}$  = distância entre as partículas i e j .

Segundo Ortúzar e Willumsen (1994) a primeira aplicação real do modelo foi realizada por Casey (1955), que analisou o fluxo de compras do comércio varejista entre várias cidades. Diversas pesquisas foram realizadas posteriormente e, genericamente, o modelo partiu da suposição de que o efeito da distância ou “separação” entre zonas de tráfego pode ser melhor modelado através de uma função

decrecente, que especifique a distância ou qualquer outra medida de impedância como custo de viagem entre zonas de tráfego. A forma genérica do modelo gravitacional é dada pela equação 2.2.

$$T_{ij} = \alpha O_i D_j f(c_{ij}) \quad (2.2)$$

onde:  $T_{ij}$  = número de viagens originadas em  $i$  e destinadas a  $j$ ;  $\alpha$  = fator de proporcionalidade;  $O_i$  = número de viagens produzidas em  $i$ ;  $D_j$  = número de viagens atraídas a  $j$ ;  $f(c_{ij})$  = função do custo de viagem generalizado com um ou mais parâmetros de calibração.

A função do custo generalizado representa o “não incentivo” a realização de viagens como o aumento da distância ou do custo da viagem percebido pelo usuário, pode ser uma função exponencial, potência ou combinação de ambas:  $f(c_{ij}) = \exp(-\beta c_{ij})$ ;  $f(c_{ij}) = c_{ij}^{-\gamma}$ ;  $f(c_{ij}) = c_{ij}^{\delta} \exp(-\lambda c_{ij})$ , onde  $\beta; \gamma; \delta; \lambda$  = constantes.

Deve-se ser observado que o número de viagens deve satisfazer a duas restrições:

$$\sum_j T_{ij} = O_i \quad \text{para } i = 1, 2, 3, \dots, n \quad (2.3)$$

e

$$\sum_i T_{ij} = D_j \quad \text{para } j = 1, 2, 3, \dots, n \quad (2.4)$$

Para garantir que as restrições (equações 2.3 e 2.4) sejam satisfeitas, o fator de proporcionalidade ( $\alpha$ ) é substituído por dois fatores de balanceamento das viagens  $A_i$  e  $B_j$ . Assim, o modelo gravitacional segue a equação 2.5, podendo ser restrito ou duplamente restrito.

$$T_{ij} = A_i O_i B_j D_j f(c_{ij}) \quad (2.5)$$

O modelo gravitacional restrito (*singly constrained*) é determinado através de restrições apenas na origem ou no destino, podendo ser produzido igualando um dos fatores de balanceamento de viagens a uma unidade. Por exemplo, fazendo  $B_j = 1$ , para  $j = 1, 2, 3, \dots, n$ , tem-se:

$$A_i = \frac{1}{\sum_j D_j f(c_{ij})} \quad \text{para } i = 1, 2, 3, \dots, n \quad (2.6)$$

No caso de modelos duplamente restritos, os valores dos fatores de balanceamento de viagens são:

$$A_i = \frac{1}{\sum_j B_j D_j f(c_{ij})} \quad \text{para } i = 1, 2, 3, \dots, n \quad (2.7)$$

$$B_j = \frac{1}{\sum_i A_i O_i f(c_{ij})} \quad \text{para } j = 1, 2, 3, \dots, n \quad (2.8)$$

#### 2.4.2 Modelo de oportunidades intervenientes

Apresenta-se nesta subsecção apenas a forma usual do modelo de oportunidades intervenientes, seus conceitos básicos e algumas deficiências. Deduções matemáticas não serão descritas, pois maior ênfase é dada à sua concepção e às razões para a escolha de tal modelo para caracterização de uso do solo da RMSP.

Ulysséa Neto e Kühlkamp (2002) partem de um exemplo hipotético que melhor explica as premissas do modelo de oportunidades intervenientes:

- Considera-se um indivíduo que reside em um bairro "A" e está com determinada necessidade que pode ser satisfeita no bairro vizinho "B";
- Esta mesma necessidade poderia ser satisfeita, com a mesma qualidade, em um bairro mais distante "C".

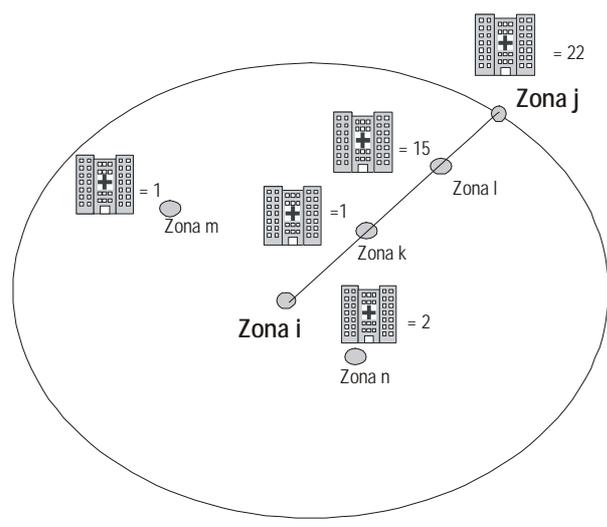
Pode-se afirmar, mesmo que intuitivamente, que caso o único objetivo da viagem seja satisfazer tal necessidade, seria irracional supor que este indivíduo viajaria até o bairro mais distante "C", arcando com os custos pertinentes a tal viagem, ao invés de simplesmente realizar a atividade necessária no bairro mais próximo "B". Da mesma forma, a decisão da escolha do destino "C" se tornará menos provável na medida em que existam outros destinos alternativos que se interponham entre "A" e "C".

A partir do exemplo, citam-se as premissas principais do modelo de oportunidades intervenientes:

- Fundamenta-se na probabilidade de um viajante, que se desloca a partir de uma zona de tráfego i, ter seu propósito de viagem satisfeito numa determinada zona de tráfego j, levando em conta as oportunidades oferecidas nas demais zonas de destinos alternativas;
- A referida probabilidade será tanto maior quanto maior for o número de oportunidades na zona de destino j;
- A mesma probabilidade será menor quanto maior for o número de oportunidades intervenientes entre i e j.

Assim, observando-se a Figura 2.5 (adaptada de Hutchinson, 1979), são consideradas intervenientes entre as zonas i e j as oportunidades oferecidas pelas zonas cujos centróides estão contidos no círculo de centro no centróide i e raio igual à distância entre os centróides de i e de j (zonas k, l, m e n).

Oportunidades, neste exemplo, são definidas pelo número de consultórios médicos localizados em cada zona de tráfego. Assim, se o motivo da viagem for saúde (necessidade de visita a consultório médico), a probabilidade de uma pessoa que reside na zona i, viajar até a zona j é: (1) proporcional ao número de consultórios médicos na zona j ("22"); e (2) inversamente proporcional ao número de consultórios médicos acumulados entre as zonas i e j ( $1 + 2 + 1 + 15 = "19"$ ).



Sendo  = Número de consultórios médicos

Figura 2.5 ilustração do conjunto de oportunidades intervenientes (adaptado de Hutchinson, 1979)

O modelo usual de oportunidades intervenientes, ou modelo de Schneider, é dado por

$$T_{ij} = O_i k_i e^{-\lambda W_{ij}} (1 - e^{-\lambda V_j}) \quad (2.9)$$

onde:  $T_{ij}$  = número de viagens com origem na zona i e destino na zona j;

$O_i$  = número de viagens com origem na zona i;

$\lambda$  = probabilidade de uma oportunidade ser aceita se for considerada;

$W_{ij}$  = número de oportunidades intervenientes entre as zonas i e j;

$V_j$  = número de oportunidades oferecidas pela zona j;

$k_i$  = fator de balanceamento das origens i (constante para cada i).

Kühlkamp e Ulysséa Neto (2004) analisam o modelo de Schneider, observando que a exponencial  $e^{-\lambda W_{ij}}$  assegura que, mantidos fixos os números  $V_j$  e  $O_i$ , ocorrendo um aumento no número  $W_{ij}$  de oportunidades intervenientes entre as zonas i e j, haverá redução no número de viagens da zona i para a zona j, estimado pelo modelo. Assim, sendo  $\lambda$  e  $W_{ij}$  números positivos, a exponencial em pauta possui expoente negativo e, portanto, decresce tanto com o aumento de  $W_{ij}$  quanto com o de  $\lambda$ .

Contudo, mantidos fixos os valores de  $W_{ij}$  e  $O_i$ , o valor da expressão  $1 - e^{-\lambda V_j}$  aumenta quando se aumenta o número de  $V_j$  das oportunidades oferecidas pela zona j. Desta forma, assegura-se que, nestas condições, o número de viagens da zona i para a zona j, estimado pelo modelo, é proporcional ao número de oportunidades oferecidas na zona j.

Observando-se as escolhas dos destinos dos indivíduos residentes na RMSP, se verificou, de uma forma genérica, que as viagens mantêm-se curtas tanto quanto possível, somente aumentando o seu comprimento no caso de não ocorrer aceitabilidade (oportunidades suficientes – taxa de empregos na indústria, comércio, serviços, etc.) de um destino numa distância menor. Assim, a proposta do conjunto de variáveis referentes ao uso do solo da RMSP, baseou-se nas premissas do modelo de oportunidades intervenientes, acumulando-se as oportunidades por raios concêntricos a partir do centróide da zona de residência dos indivíduos. A proposta das variáveis que caracterizam o uso do

solo ou intensidade e distribuição de atividades da região de estudo será descrita minuciosamente no Capítulo 5.

Apesar da adoção dos conceitos do modelo de oportunidades intervenientes neste trabalho, há algumas deficiências apontadas por Ulysséa Neto e Kühlkamp (2002) e por Ortúzar e Willumsen (1994), que tornam a sua aplicação menos popular do que outros modelos sintéticos, como os gravitacionais, por exemplo:

- A idéia de matrizes de destinos ordenados pela distância em relação à origem é mais complexa (a célula “n” da matriz, não é o destino “n”, mas sim o enésimo destino mais distante da origem). Assim, cada origem terá uma linha com ordens de destinos diferentes;
- As vantagens teóricas e práticas desse modelo em relação aos modelos gravitacionais não são muito conhecidas;
- Não se consideram medidas de impedância (distância, custo da viagem) de maneira explícita. As oportunidades são acumuladas considerando que um destino é mais distante do que outro qualquer, mas não o quanto mais distante cada destino está do outro em relação a uma origem;
- Há problemas relacionados à definição do conjunto de oportunidades intervenientes entre duas zonas de tráfego i e j. Duas zonas de tráfego com distâncias iguais em relação à origem podem, muitas vezes, não competir em igualdade de condições. A escolha de uma zona em detrimento da outra leva também em consideração a sua atratividade e fatores relacionados às respectivas vizinhanças.

Na próxima subseção será definida uma forma alternativa aos modelos de oportunidades intervenientes, que combina características dos primeiros aos modelos gravitacionais.

### 2.4.3 Modelos gravitacionais – de oportunidades

Nos modelos gravitacionais, conceituados anteriormente, a impedância entre as zonas de tráfego é caracterizada não somente pela distância entre elas, mas também pelo tempo de viagem ou custo percebido pelo usuário. Já nos modelos de oportunidades intervenientes, caracteriza-se a impedância

entre as zonas pelas oportunidades intervenientes. No referido caso, a separação espacial entre as zonas não é levada explicitamente em consideração.

Partindo-se da idéia de que os deslocamentos individuais consideram tanto separação espacial quanto oportunidades acumuladas entre zonas, surgiram novos modelos de distribuição de viagens, ainda mais complexos, chamados modelos Gravitacionais – de Oportunidades.

Wills (1986) desenvolveu um modelo gravitacional – de oportunidades, de grande complexidade, combinando em um único modelo características dos modelos gravitacionais e de oportunidades intervenientes. O autor propôs um modelo flexível (*gravity-opportunity* – GO), no qual a escolha entre as abordagens “gravitacionais” ou “oportunidades intervenientes” era decidida empiricamente, permitindo a estimação de parâmetros que controlavam a função global.

Gonçalves e Ulysséa Neto (1993) desenvolveram um modelo gravitacional – de oportunidades, aplicando-o para estimação de fluxos intermunicipais de passageiros por transporte público em uma região de Santa Catarina. Assume-se que o número de viagens entre cada par de zonas:

- É proporcional ao número de viagens produzidas na zona de origem;
- É proporcional ao número de viagens atraídas à zona de destino;
- É inversamente proporcional a uma medida de impedância associada à distância entre origem e destino;
- É inversamente proporcional a uma medida de impedância relativa ao número de oportunidades existentes entre a origem e destino.

A formulação geral do modelo proposto por Gonçalves e Ulysséa Neto (1993) é dada pela equação 2.10 e as respectivas restrições são apresentadas em seguida (2.11 e 2.12). Deduções matemáticas não são descritas no presente texto. O objetivo é apresentar três modelos sintéticos de distribuição de viagens, enfatizando suas premissas básicas.

$$T_{ij} = A_i O_i B_j D_j e^{-\lambda w_{ij}} e^{-\beta c_{ij}} \quad (2.10)$$

onde:

$O_i$  = número de viagens originadas em  $i$ ;

$D_j$  = número de viagens atraídas a  $j$ ;

$A_i$  e  $B_j$  = parâmetros que garantem as equações de consistência de fluxos;

$W_{ij}$  = número de oportunidades entre  $i$  e  $j$ ;

$\lambda$  = parâmetro de impedância que está associado às oportunidades intervenientes;

$\beta$  = parâmetro de impedância associado à distância ou custo da viagem;

$c_{ij}$  = custo da viagem entre  $i$  e  $j$ .

Sendo:

$$A_i = \left( \sum_j B_j D_j e^{-\lambda w_{ij}} e^{-\beta c_{ij}} \right)^{-1} \quad \text{para } i = 1, 2, 3, \dots, n \quad (2.11)$$

$$B_j = \left( \sum_i A_i O_i e^{-\lambda w_{ij}} e^{-\beta c_{ij}} \right)^{-1} \quad \text{para } j = 1, 2, 3, \dots, n \quad (2.12)$$

## 2.5 CONSIDERAÇÕES FINAIS

O presente capítulo abordou temas considerados relevantes para o desenvolvimento da pesquisa. Inicialmente é discutida a complexidade da análise dos fatores que interferem no encadeamento de viagens, pois os indivíduos consideram uma série de questões ao planejar a seqüência de viagens a serem realizadas no período de um dia. Levando-se em conta o uso do solo, deve-se atentar que, não somente as características das regiões de origem, como também de outras localizações (trabalho, por exemplo), possivelmente influenciam os itinerários de viagem.

Em seguida são apresentados trabalhos que corroboram possíveis relações entre uso do solo e comportamento relacionado a viagens, bem como outros trabalhos em que não se pode afirmar realmente a existência de tais relações. Esta pesquisa visa encontrar relações entre padrões de viagens urbanas e conjuntos de variáveis relativas a características de uso do solo, socioeconômicas e participação em atividades. No entanto, um dos principais desafios é mensurar o conjunto de variáveis referente às características de uso do solo da RMSP. Uma série de propostas de representação de variáveis de uso do solo, revisadas na literatura a respeito do tema, foi citada no presente texto.

Finalmente, foram descritos três modelos de distribuição de viagens. Para propor o grupo de variáveis de uso do solo (Capítulo 5) foram levados em conta conceitos do modelo de oportunidades intervenientes.

A etapa do método que consiste na revisão da literatura é essencial para adquirir conhecimento e fundamentação teórica a respeito do tema, no intuito de propor soluções alternativas para se atingir o objetivo principal. No próximo capítulo será descrita a etapa subsequente do método. Serão descritas as técnicas de auxílio à análise de dados, aplicadas no trabalho. A utilização de tais técnicas foi um dos meios que viabilizou o estudo de relações das variáveis envolvidas e auxiliou interpretações de resultados obtidos.

# Capítulo 3

## ANÁLISE MULTIVARIADA DE DADOS

---

*Este capítulo apresenta definição e fundamentação teórica a respeito das ferramentas de apoio à análise, empregadas neste trabalho. Trata, especificamente, das técnicas de Análise Multivariada de Dados (AM). As técnicas de AM podem ser classificadas em dois grandes grupos: Tradicionais e "Emergentes" (Data Mining – DM). Aquelas utilizadas no trabalho e descritas no presente capítulo são: (1) Análise de Cluster (AC); (2) Regressão Linear Múltipla (RLM) – ambas classificadas como técnicas de AM tradicionais e; (3) Árvore de Decisão (AD) - técnica de DM. Também são mostradas as principais razões para sua utilização, bem como discussões a respeito de sua aplicação no trabalho.*

### 3.1. ESCOLHA DAS TÉCNICAS DE ANÁLISE MULTIVARIADA DE DADOS (AM)

Devido principalmente ao progresso computacional nas últimas décadas, houve uma crescente aplicação de um conjunto de técnicas estatísticas conhecidas como Análise Multivariada de Dados (AM), permitindo ao pesquisador sintetizar a complexidade dos dados, detectando padrões, relações e outras particularidades e auxiliando o desenvolvimento de hipóteses e modelos adequados.

Existem diferentes técnicas de AM que podem ser utilizadas para diversas finalidades específicas, sendo comum a todas elas um elevado grau de complexidade que requer conhecimento matemático relativamente sofisticado. As referidas técnicas analisam um conjunto de diversas variáveis e podem ser utilizadas principalmente para:

- Redução e simplificação de um conjunto de dados;
- Classificação;
- Investigação de dependência entre variáveis.

De uma maneira genérica, as técnicas de AM podem ser definidas como ferramentas analíticas que auxiliam na investigação acerca de fenômenos complexos envolvendo múltiplas dimensões. Hair

et al. (1998) conceituam AM como um conjunto de técnicas estatísticas utilizadas com o objetivo de explicar e prever o grau de relações existentes entre diversas variáveis independentes (inclusive entre si) e a variável dependente. Os mesmos autores classificam as técnicas de AM em dois grandes grupos: (1) Técnicas de AM tradicionais; (2) Técnicas de AM "emergentes" – *Data Mining* (DM).

### 3.1.1 Técnicas de AM tradicionais (técnicas confirmatórias):

Podem ser consideradas extensões de técnicas estatísticas tradicionais, conhecidas como univariadas ou bivariadas (regressão linear simples – regressão linear múltipla; ANOVA – MANOVA). A rigor, esse conjunto inclui muitas outras técnicas, tais como Análise Fatorial, Análise de Aglomerados (*Cluster Analysis*) e Escalonamento Multidimensional. São técnicas predominantemente confirmatórias, sendo que a inferência estatística é a base da maioria das técnicas de AM tradicionais.

Para sua aplicação, o pesquisador geralmente depara-se com uma situação bem delineada, onde se conhecem o problema, os tipos de variáveis e relações (dependência, interdependência). Assim, planeja-se e desenvolve-se o experimento, escolhendo a técnica de AM apropriada de forma que esta possa ou não confirmar as suposições iniciais. A inferência estatística, que faz parte da maioria das técnicas de AM tradicionais, dá suporte à refutação ou não de hipóteses prévias.

É cada vez mais comum o uso de técnicas de AM tradicionais em trabalhos que investigam relações complexas, envolvendo múltiplas dimensões na área de Transportes. Pode-se afirmar que nos últimos anos a grande maioria dos trabalhos a respeito do comportamento relacionado a viagens utiliza tais técnicas para corroborar relações supostamente ou intuitivamente conhecidas.

Muitas aplicações de modelos de equações estruturais no estudo de comportamento relacionado a viagens, considerando a abordagem de viagens baseadas em atividades podem ser encontradas no estudo realizado por Golob (2001). McNally e Kulkarni (1997) estudaram interações entre uso do solo e comportamento de viagem. Os autores utilizaram Análise de *Cluster* a fim de representar uso do solo e aspectos da rede de transportes. Srinivasan (2000) utilizou Análise Fatorial Exploratória e Confirmatória a fim de caracterizar aspectos do uso do solo da área metropolitana de

Boston. Obtidos grupos (fatores) caracterizados por diferentes atributos de uso do solo, utilizou-se Regressão Multinomial Logística a fim de prever relações entre escolha modal e características socioeconômicas e uso do solo.

### 3.1.2 Técnicas de AM “emergentes” (técnicas exploratórias) – *Data Mining* (DM)

Técnicas de AM “emergentes” ou Mineração de dados (Data Mining – DM) são procedimentos que muitas vezes devem ser associados às técnicas de AM tradicionais. Pode-se definir esse conjunto de técnicas como o processo de extrair informação válida, previamente desconhecida, a partir de grandes bases de dados, a qual pode ser usada como um elemento auxiliar na tomada de decisões importantes. O principal objetivo da DM é descobrir padrões ocultos em complexas fontes de informação, razão pela qual nos dias de hoje é utilizada crescentemente em diversas áreas.

Com base em estudos realizados por Bispo (1998), pode-se citar algumas técnicas de DM: Redes Neurais Artificiais, Árvores de Decisão, Algoritmos Genéticos, Regras de Associação (*Association Rules*). Pode-se observar recentemente a aplicação de tais técnicas em trabalhos referentes ao comportamento relacionado a viagens.

Com o objetivo de explorar o potencial de técnicas de DM para análise de viagens, Keuleers e Wets (2001) utilizaram regras de associação em um amplo conjunto de dados com o propósito de descobrir associações significativas entre padrões de atividades diárias e atributos individuais, domiciliares e relacionados ao meio. O interesse na utilização de regras de associação em tal trabalho foi motivado pela suposição fundamental da Análise de Viagens baseadas nas Atividades de que indivíduos e familiares aplicam regras de decisão quando programam suas atividades no tempo e no espaço, considerando algumas condições.

Yamamoto, Kitamura e Fuji (2001) desenvolveram um trabalho de aplicação de DM para investigação do comportamento de escolha de rotas por motoristas, concluindo que as técnicas de mineração de dados possuem estruturas mais flexíveis do que os modelos convencionais de escolha discreta para representar relações entre os atributos das alternativas e as escolhas.

Na literatura também são encontrados trabalhos que utilizam CHAID – (*Chi-squared Automatic Interaction Detection*) para análises de taxas de produção de viagens, estudos da mobilidade feminina, considerando um conjunto de fatores e diferenças de mobilidade entre homens e mulheres na RMSP (STRAMBI; VAN DE BILT; PIETRANTONIO, 1995; STRAMBI; VAN DE BILT, 1998).

Com a contribuição observada em diversos trabalhos que aplicaram técnicas de AM “emergentes”, pode-se afirmar que tais ferramentas são promissoras, principalmente quando se trata de estudos exploratórios. No entanto, por serem técnicas predominantemente exploratórias, há um questionamento quanto à sua inferência estatística (HAIR et al, 1998).

Embora inferência estatística para se obter, por exemplo, a significância dos parâmetros estimados seja a base da maioria das técnicas tradicionais de AM, técnicas “emergentes” de AM permitem que pesquisadores explorem modelos analíticos não necessariamente fundamentados na inferência estatística. Tais métodos, na maioria das vezes, permitem uma estimação empírica ao invés de suposição estatística.

Para aplicação das técnicas de DM não se conhece inicialmente o tipo de relações que devem ser posteriormente confirmadas. Geralmente defronta-se com uma questão muito genérica a ser respondida (por exemplo, “O que afeta o comportamento de clientes?”). Desta forma, o pesquisador encontra-se em uma situação em que necessita de técnicas exploratórias, pois não se supõe previamente o tipo de relação ou padrões existentes. Assim, as técnicas a serem utilizadas nesse caso não serão escolhidas baseando-se no número ou tipo de variáveis dependentes ou independentes.

O foco exploratório de tais técnicas não exclui, entretanto, as etapas de confirmação ou validação. No entanto, mesmo quando há base estatística (Regras de Associação, Árvores de Decisão), a significância estatística deixa a desejar se comparada a outras técnicas de AM tradicionais. Isso significa uma nova tendência de análise de dados, com características peculiares, focando em descobrir conhecimento ao invés de refutar ou não supostas hipóteses? Não necessariamente. A aplicação conjunta dos dois tipos de técnicas pode trazer benefícios e resultados mais interessantes.

Hair et al (1998) afirmam que DM pode explorar diversas opções, encontrando relações e modelos *a priori* desconhecidos. Mas como mensurar o nível de ajuste do modelo, ou a significância das variáveis?

### 3.1.3 Uso de técnicas exploratórias + confirmatórias

No trabalho em questão, o objetivo principal é encontrar relações entre padrões de viagem (variável dependente) e características socioeconômicas, participação em atividades e atributos do uso do solo (variáveis independentes). Pode-se destacar o caráter exploratório da atual pesquisa, principalmente considerando duas questões principais:

- Como mensurar padrões de uso do solo da RMSP?
- Que tipo de relações existem entre variável dependente e variáveis independentes?

Embora se tenha estabelecido inicialmente que a variável dependente seja categórica e que as independentes sejam categóricas e numéricas, não se sabe como tais variáveis exercem influência umas sobre as outras. Além disso, defronta-se com grandes bases de dados desagregadas, contendo um grande número de indivíduos e informações a respeito de cada um deles, possivelmente com padrões e relações ainda não conhecidos que fazem parte da estrutura dos dados (Capítulo 4). Assim, justifica-se a utilização de técnicas exploratórias associadas às confirmatórias. Foram utilizadas três técnicas de AM, descritas nas próximas seções:

- 1) Análise de Cluster: Para agrupar os destinos e auxiliar na representação da variável dependente;
- 2) Árvore de Decisão: A fim de encontrar relações entre as variáveis dependente e independentes, formando grupos de indivíduos, consistentes em seus atributos (socioeconômicas, atividades e uso do solo), homogêneos segundo a variável dependente (padrões de viagem) – técnica exploratória;
- 3) Regressão Linear Múltipla: Trazer inferência estatística ou estimar a significância das variáveis independentes, especialmente daquelas de uso do solo – técnica confirmatória.

A Figura 3.1 traz um esquema adaptado de Hair et al (1998), ilustrando as principais técnicas de AM. Encontram-se destacadas na figura as técnicas utilizadas no trabalho.

### 3.2. ANÁLISE DE CLUSTER (AC)

A Análise de *Cluster* (AC), também conhecida como análise de conglomerados, é um conjunto de técnicas estatísticas cujo objetivo é agrupar ou dividir objetos segundo suas características (métodos divisivos ou aglomerativos), formando grupos ou conglomerados homogêneos, de acordo com determinado critério. Os objetos em cada grupo tendem a ser semelhantes entre si e diferentes dos demais objetos contidos em outros conglomerados. Os conglomerados obtidos devem apresentar tanto uma homogeneidade interna (dentro de cada conglomerado), como uma grande heterogeneidade externa (entre conglomerados). Portanto, se a aglomeração for bem sucedida, quando representados em um gráfico, os objetos dentro dos conglomerados estarão muito próximos, e os conglomerados distintos estarão afastados (HAIR et al., 1998).

Apesar de ser classificada com uma técnica de AM tradicional, AC é aplicada na maioria das vezes em pesquisas de caráter exploratório. É uma técnica para analisar interdependência entre casos/indivíduos segundo determinadas variáveis. Não é possível determinar antecipadamente as variáveis dependentes e independentes. Ao contrário, examina relações de interdependência contidas na estrutura dos dados. Para aplicação da AC é necessário seguir algumas etapas básicas, sugeridas por Malhotra (2001) e descritas nas próximas subseções.

#### 3.2.1 Definir o problema de aglomeração ou objetivo da AC na pesquisa

Como mencionado anteriormente o objetivo da AC é dividir um conjunto de objetos em dois ou mais grupos considerando a similaridade dos objetos em relação a variáveis específicas. Obtendo-se os grupos, o pesquisador pode atingir três finalidades distintas, que devem estar relacionadas ao objetivo da pesquisa:



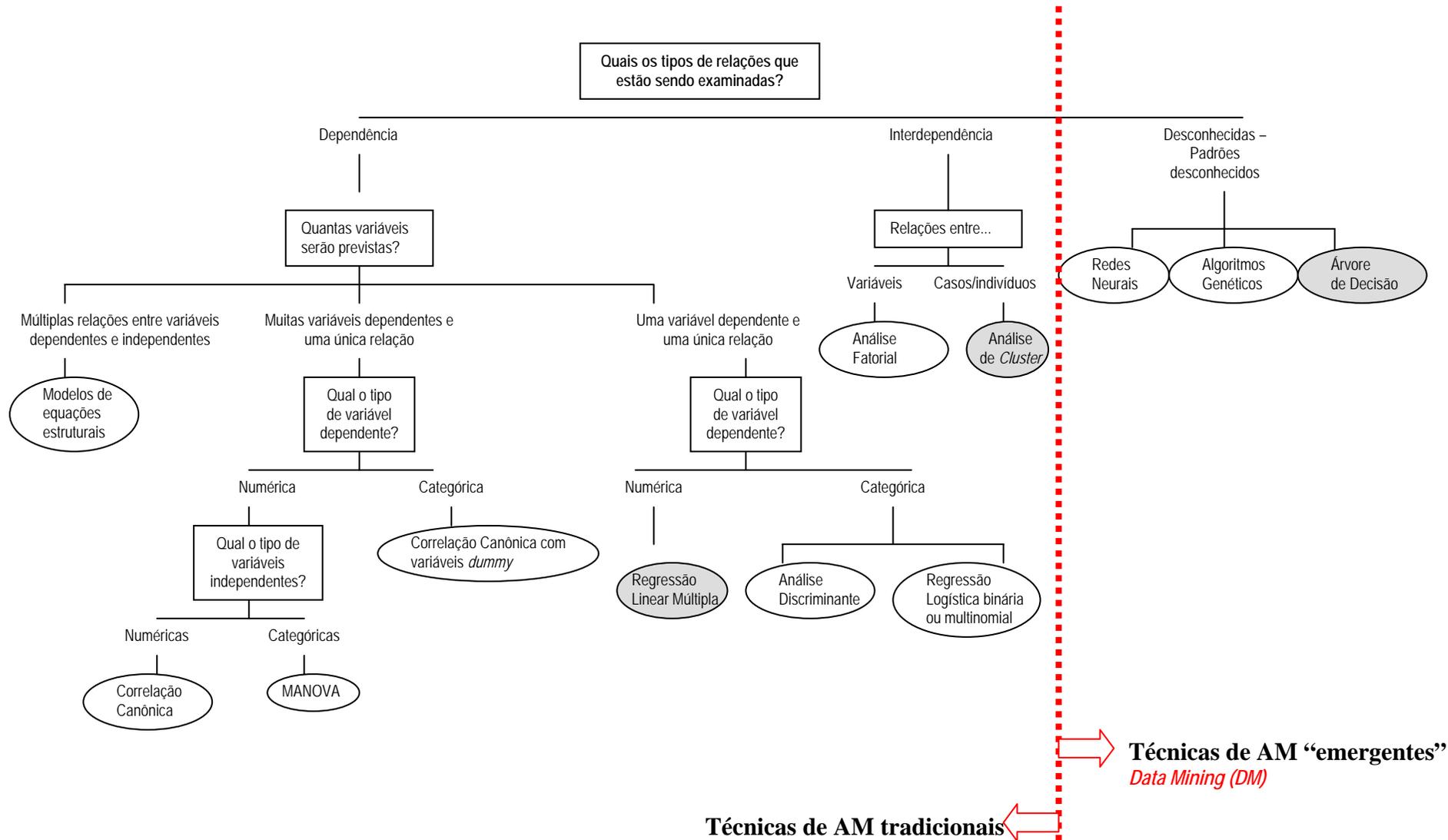


Figura 3.1 Síntese das principais técnicas de AM



- (A) Taxonomia: é considerada uma classificação empírica de objetos. O uso tradicional da AC é para fins exploratórios e formação de taxonomia. A despeito de ser preferencialmente usada para propósitos exploratórios, também pode ser utilizada em casos confirmatórios (uma tipologia proposta pode ser comparada a uma classificação derivada da aplicação da AC);
- (B) Simplificação dos dados: Assim como a Análise Fatorial reduz um conjunto de variáveis, a AC pode agrupar as observações (objetos, indivíduos, etc.). Ao invés de analisar as observações de forma individual, estas farão parte de grupos que possuem características relacionadas a cada um deles.
- (C) Identificação de relações: Com os grupos definidos, o pesquisador pode encontrar relações entre as observações que talvez não aparecessem observando-as individualmente, mas sim através dos grupos obtidos.

Na presente pesquisa, deseja-se agrupar os destinos escolhidos pelos indivíduos no período de um dia, considerando faixas de distâncias. Os destinos não poderiam ser representados na forma de zonas de tráfego, por exemplo, devido ao elevado número de zonas da RMSP (389 no total). Assim, é proposta uma classificação empírica dos destinos considerando como critério as distâncias da primeira viagem realizada (distâncias a partir das residências dos indivíduos). A aplicação da AC no trabalho será descrita de forma minuciosa no Capítulo 6.

### 3.2.2 Escolha das variáveis

Selecionar as variáveis para aplicação da AC deve ser uma tarefa cuidadosa e vinculada explicitamente ao objetivo do pesquisador. A seleção das variáveis também depende da disponibilidade e coleta de dados. Os possíveis resultados estarão restritos às variáveis escolhidas. AC pode ser fortemente afetada pela inclusão de uma ou duas variáveis inapropriadas, sem poder discriminatório entre objetos ou com presença de *outliers*.

### 3.2.3 Escolha da medida de similaridade ou dissimilaridade

Como o objetivo da AC é agrupar objetos semelhantes ou dessemelhantes, é necessária uma medida de similaridade ou dissimilaridade entre os mesmos. Medidas de distância são as mais utilizadas na AC. Os objetos com menor distância entre si são mais semelhantes, logo são aglomerados em um mesmo conglomerado. Já os mais distantes participam de conglomerados distintos. Existem várias formas de medir a distância entre os objetos, porém a mais utilizada é a distância euclidiana. A Tabela 3.1 conceitua as principais medidas de distância.

Tabela 3.1 Medidas de distância (AC)

| Tipos de medidas de distância                                                                                                                                                                                                                   |                                                                                                                                        |                                                                                                                           |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------|
| <p>Distância Euclidiana</p> <p>A medida do comprimento de uma linha reta entre dois objetos. É a medida de similaridade mais comum utilizada.</p> $D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$                                                    | <p>City-block</p> <p>Calcula distâncias através da soma das diferenças absolutas entre coordenadas</p> $D =  x_2 - x_1  +  y_2 - y_1 $ | <p>Mahalanobis</p> <p>Forma padronizada da distância euclidiana. Ajustes feitos para intercorrelações entre variáveis</p> |
| <p>Utiliza-se também o quadrado da distância euclidiana (<i>square euclidian distance</i>).</p> <p>Há outras medidas de similaridade disponíveis no SPSS 13.0 (<i>cosine, pearson correlation, chebychev, block, Minkowski, Customized</i>)</p> |                                                                                                                                        |                                                                                                                           |

$x_1; x_2; y_1; y_2$  – coordenadas de dois objetos

### 3.2.4 Definição do processo de aglomeração

Os processos de aglomeração podem ser hierárquicos e não-hierárquicos. A aglomeração hierárquica se caracteriza pelo estabelecimento de uma hierarquia ou estrutura em forma de árvore, podendo ser aglomerativos ou divisivos. Já a aglomeração não-hierárquica, exemplificada pelo processo *K-means clustering*, segue um algoritmo mais simplificado, dado pelas seguintes etapas: (i) é adotado o número de conglomerados; (ii) inicialmente assume-se um centro de aglomeração (centróide) para cada grupo; (iii) aloca-se cada objeto ao grupo que possui um centro de aglomeração mais próximo ao seu; (iv) quando todos os objetos estiverem alocados recalculam-se os centróides de cada grupo; (v) repete-se os passos "iii" e "iv" até os valores dos centróides dos grupos não mais variarem. A Tabela 3.2 sintetiza os principais processos de aglomeração segundo Hair et al. (1998). A maioria dos *softwares* inclui tais

métodos nos seus pacotes. A seguir, será descrito o processo hierárquico aglomerativo, utilizado neste trabalho.

Tabela 3.2 Processos de aglomeração

| Processo<br>de<br>aglomeração | Hierárquicos  |                              | Não-Hierárquicos            |
|-------------------------------|---------------|------------------------------|-----------------------------|
|                               | Aglomerativos | <i>single linkage</i>        | <i>(K-means clustering)</i> |
| <i>complete linkage</i>       |               | <i>Sequential Threshold</i>  |                             |
| <i>average linkage</i>        |               | <i>Parallel Threshold</i>    |                             |
| <i>Ward</i>                   |               | <i>Optimization</i>          |                             |
| <i>centroid method</i>        |               | <i>Selecting Seed Points</i> |                             |
| Divisivos                     |               |                              |                             |

- **Processo hierárquico aglomerativo**

O processo de aglomeração utilizado no estudo é chamado hierárquico aglomerativo. Tal processo envolve uma construção hierárquica em uma estrutura de árvore (dendograma), sendo compreendido por vários estágios. Em cada um dos estágios há redução do número de objetos (*clusters*) devido ao agrupamento destes, considerando como critério de similaridade a distância euclidiana.

Cada objeto no primeiro estágio forma o seu próprio *cluster* (ou conglomerado). Assim, o número de *clusters* no primeiro estágio é igual ao número de objetos do banco de dados. Em um estágio subsequente os dois *clusters* (objetos) de maior similaridade ou menor distância formam um novo *cluster*, reduzindo assim o número de *clusters* contidos anteriormente no estágio 1. No final do processo, todos os objetos ou indivíduos são agrupados em um *cluster* maior. Este processo chama-se de aglomerativo porque os *clusters* são formados sucessivamente a partir de outros existentes (HAIR et al, 1998).

A Figura 3.2 descreve as etapas do processo aglomerativo onde foi realizada uma pesquisa a respeito da satisfação dos clientes quanto aos serviços de uma rede de *fast-food*. Neste exemplo, há seis indivíduos (A,B,C,D,E,F) que atribuíram notas (variando de zero a sete) quando questionados a respeito de duas variáveis (rapidez de atendimento, no eixo x; e serviço de entrega, no eixo y). Na

primeira solução (estágio 1), cada indivíduo faz parte de um único grupo (solução com seis *clusters*), na segunda solução (estágio 2) verifica-se que os indivíduos E e F possuem maior similaridade, agrupando-se desta forma, em um novo *cluster* (solução com cinco *clusters*). No exemplo as soluções variam de seis a um *cluster*, sendo que quanto maior o número de *clusters* maior é homogeneidade interna. O exemplo apresenta seis estágios.

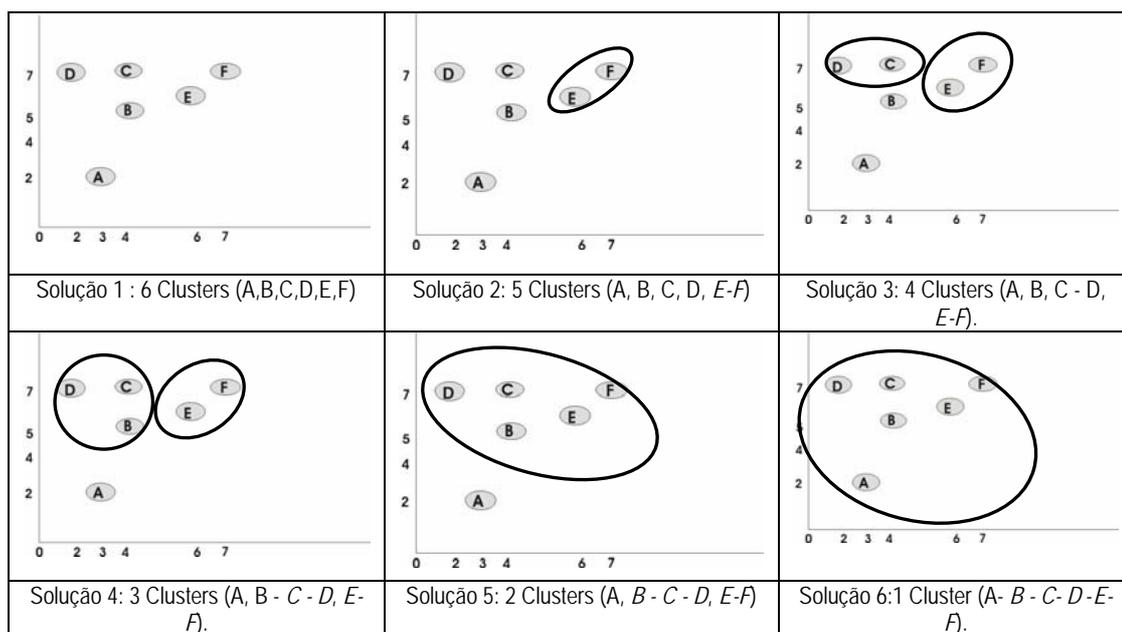


Figura 3.2 Exemplo de aplicação do processo hierárquico aglomerativo

A Figura 3.3, em seguida, ilustra o dendograma obtido para o exemplo em questão.

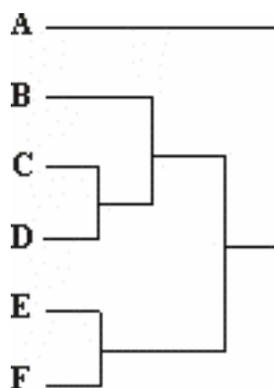


Figura 3.3 Dendograma (processo aglomerativo – 6 indivíduos e 6 soluções)

### 3.2.5 Definição do número de conglomerados

Um dos principais problemas que o pesquisador defronta-se ao trabalhar com AC é determinar um número final de *clusters* a serem formados (*stopping rule*). Não há nenhum procedimento, ou padrão a ser seguido. O mais sensato seria testar diversas soluções (um a seis grupos como no exemplo ilustrado na Figura 3.2) e decidir entre as alternativas utilizando um critério prévio, bom senso ou fundamentação teórica a respeito do problema em questão.

### 3.2.6 Peculiaridades da AC

AC pode ser caracterizada como descritiva e não inferencial. Embora seja classificada como uma técnica de AM tradicional (HAIR et al, 1998), a AC não é uma técnica de inferência estatística na qual os parâmetros de uma amostra podem ser estimados, sendo possivelmente expandidos para a população. Tal ferramenta de análise é usada principalmente como uma técnica exploratória. Pode haver diversas soluções para o problema de agrupamento, principalmente quando se modificam as variáveis ou número de grupos, por exemplo. Assim, a adição ou retirada de variáveis relevantes pode causar impactos substanciais na solução final. Por isso, o pesquisador possui um papel fundamental na escolha do critério de agrupamento.

Possui forte base matemática, mas não estatística. Suposições como normalidade, linearidade e homocedasticidade (que são tão importantes em outras técnicas, como Regressão Linear Múltipla conforme mostrado na seção 3.4) possuem pouco impacto no caso de AC. Deve-se, no entanto, atentar para dois fatores importantes:

- Representatividade da amostra: os conglomerados são obtidos partindo-se do pressuposto de que a amostra represente a estrutura da população que esta é derivada, já que não há inferência estatística que expanda parâmetros estimados da amostra para população. Assim, é importante certificar-se de que a amostra trabalhada é representativa e os resultados podem ser generalizados para a população de interesse;

- Impacto da multicolinearidade: variáveis altamente correlacionadas implicitamente recebem maior peso do que as outras na formação de conglomerados. Para exemplificar melhor tais tipos de problemas, observa-se um caso hipotético (adaptado de HAIR et al, 1998), onde indivíduos foram questionados (dez perguntas no total) a respeito de um serviço de internet rápida. As dez respostas dadas por cada indivíduo seriam as variáveis utilizadas na AC.

No entanto, observa-se que as dez perguntas fazem parte de dois grupos principais (oito relacionadas à satisfação do serviço e duas relacionadas a preços e promoções). Assim as oito variáveis do primeiro grupo são altamente correlacionadas entre si, assim como as duas do segundo grupo de perguntas.

Considerar as dez variáveis originais para aplicação da AC seria um possível erro. Variáveis contidas no primeiro grupo teriam quatro vezes mais chances (oito perguntas comparadas a duas perguntas) de afetar medidas de similaridade.

### 3.3. ÁRVORE DE DECISÃO (AD)

A técnica de DM utilizada neste trabalho foi Árvore de Decisão (AD). O principal objetivo para seu uso é encontrar relações (não necessariamente lineares) entre variável dependente (padrões de viagem) e variáveis independentes (uso do solo, socioeconômicas e atividades). A técnica é considerada uma forma simples de representação de relações existentes em um conjunto de dados. Possui este nome porque a sua estrutura, fácil de entender e assimilar, se assemelha a uma árvore. Os dados são divididos em subgrupos, com base nos valores das variáveis. O resultado é uma hierarquia de declarações tipo "Se... então..." que são utilizadas, principalmente, para classificar dados.

Esses modelos são ajustados por sucessivas divisões dos dados, com a finalidade de obter subconjuntos cada vez mais homogêneos segundo a variável dependente. O resultado é uma árvore hierárquica de regras de decisão utilizadas para prever ou classificar. Pode-se citar como principais vantagens a estrutura clara e a facilidade de compreensão dos resultados, bem como a sua habilidade

em representar relações não compensatórias entre atributos, difícil em funções algébricas (VAN DER SMAGT; LUCARDIE, 1991).

Embora a AD se assemelhe a outras técnicas tradicionais como análise discriminante, por exemplo, a sua flexibilidade a torna uma ferramenta de análise bastante atraente. No entanto, o seu uso não exclui a aplicação conjunta de técnicas tradicionais.

O algoritmo usado para dividir os dados nos modelos de árvore procura encontrar aquelas variáveis independentes mais importantes, ou seja, aquelas que fornecem máxima segregação nos dados (WESTPHAL; BLAXTON, 1998). Alguns dos algoritmos que produzem tipos similares de árvores de decisão são C4.5 (QUINLAN, 1993), CHAID (KASS, 1980) e CART (*Classification and Regression Tree* - BREIMAN et al., 1984).

### 3.3.1 Principais propriedades

AD pode ser definida como um gráfico acíclico e direto que satisfaz as seguintes propriedades (SAFAVIAN; LANDGREBE, 1991):

- A hierarquia é denominada árvore e cada segmento é denominado nó;
- Há um nó, chamado raiz, que contém todo o banco de dados;
- Este nó contém dados que podem ser subdivididos dentro de outros sub-nós, chamados de nós filhos;
- Existe um único caminho entre o nó raiz e cada nó;
- Quando os dados do nó não podem ser mais subdivididos dentro de um outro subconjunto ele é considerado um nó terminal ou folha;
- Para construção da AD deve-se considerar três elementos principais: um conjunto de perguntas que delimita as divisões dos dados, um critério que avalia a melhor divisão e uma regra para finalização das divisões (*stop-splitting rule*).

A Figura 3.4 mostra um exemplo esquemático de AD, para o caso de sucessivas divisões binárias no conjunto de dados (CART). O algoritmo CART foi aplicado a uma análise da escolha modal.

Estão representados na Figura 3.4: (1) o nó raiz (elipse sombreada); (2) nós filhos (elipses brancas); (3) folhas ou nós terminais (retângulos sombreados); (4) as variáveis independentes contínuas e discretas (Idade, Renda, Nº de Automóveis no Domicílio, Distância da Viagem e Membro da Família); e (5) variável dependente (três categorias: automóvel, transporte público, a pé).

Observando a AD, pode-se afirmar que a variável que faz a máxima divisão dos dados é a distância. A partir do nó raiz são obtidos então dois grupos principais: (A) indivíduos que realizam viagens com distâncias inferiores a 3 km; e (B) indivíduos que realizam viagens com distâncias superiores ou iguais a 3 km. Encontrando-se a melhor segregação da amostra total, é realizada novamente a busca pela máxima segregação dos nós filhos (Idade < 22 anos, Automóveis  $\leq 1$ , Renda < R\$ 500, Automóveis = 0, Chefe da família). Há alguns critérios a serem seguidos para busca da melhor divisão dos dados. Breiman et al (1984) utilizam o critério chamado de impureza ou heterogeneidade do nó. Na próxima subseção serão descritas algumas características do algoritmo de AD utilizado no trabalho, inclusive as regras para finalização das divisões.

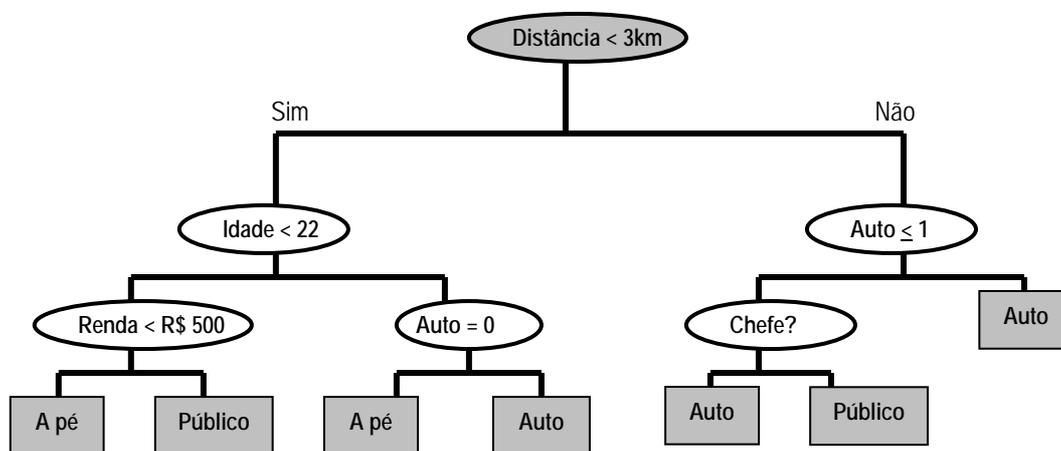


Figura 3.4 Exemplo de aplicação de AD (algoritmo CART)

### 3.3.2 S-PLUS 6.1 e CART

Neste trabalho utilizou-se o *software* S-PLUS 6.1, minerador de dados disponível no Departamento de Transporte da EESC. As principais razões para sua escolha foram primeiramente a sua disponibilidade, a natureza probabilística dos relatórios emitidos pelo programa e a viabilidade do

seu uso, comprovada em trabalhos anteriores (ICHIKAWA, 2002; PITOMBO, 2003; SOUZA, 2004; AGUIAR, 2005; SILVA, 2006).

A árvore contida no S-PLUS 6.1 é uma variante do algoritmo do CART, que estabelece uma relação entre variáveis independentes e variável dependente. O algoritmo é ajustado mediante sucessivas divisões binárias no conjunto de dados, de modo a tornar os subconjuntos resultantes cada vez mais homogêneos, em relação à variável dependente. Essas divisões são representadas por estrutura de árvore binária, sendo que cada nó corresponde a uma divisão (BREIMAN et al., 1984).

Tanto a variável dependente quanto as independentes podem assumir valores contínuos ou categóricos. Se a variável dependente for numérica, o modelo é conhecido como árvore de regressão e assume uma distribuição normal (gaussiana), caso seja categórica, o modelo é denominado árvore de classificação e segue uma distribuição multinomial.

As componentes básicas de algoritmo CART são os “nós” e as “regras de divisão”. Os nós estão associados aos subconjuntos resultantes da aplicação de uma regra de divisão a determinado conjunto de dados. As regras de divisão podem ser do tipo:

- $Idade < 32.5$ , no caso variável independente contínua, ou;
- $X \in \{A, B\}$ , para variável independente categórica.

As respostas às regras de divisão podem ser positivas ou negativas, dividindo-se nos ramos da árvore.

A função contida no S-PLUS 6.1 permite até 128 valores para variável dependente categórica e 32 para variáveis independentes categóricas, pois se há  $K$  níveis de variáveis independentes, então ocorrerão  $2^{(K-1)} - 1$  divisões, que serão examinadas, impondo alta demanda de processamento computacional.

Para o término das divisões, S-PLUS 6.1 utiliza dois critérios:

- I. Se a homogeneidade é menor do que um dado índice,  $e$ ;
- II. O número de observações no nó é menor do que um valor mínimo absoluto.

Neste trabalho foram adotados para todas as amostras o valor de “0,15” (denominado desvio mínimo no Capítulo 7), considerando-se o critério I. Para atender ao critério II, adotou-se um número de 50 observações.

O S-PLUS 6.1 trata a árvore como modelo de probabilidade, empregando o desvio como critério de divisão. Para as árvores de classificação, a heterogeneidade é medida pelo desvio global (D):

$$D = -2 \sum_t \sum_{j \in t} n(j|t) \log p(j|t) \quad (3.1)$$

Em que:  $n(j|t)$  é o número de casos no nó  $t$  com a resposta categórica  $j$ ;

$p(j|t)$  denota a proporção dos casos na categoria  $j$  do nó  $t$ .

Significa que na árvore do S-PLUS 6.1 o desvio foi definido como recíproco da função verossimilhança elevada ao quadrado. Portanto, o desvio é utilizado como uma medida da heterogeneidade dos grupos resultantes.

Outra característica inerente ao algoritmo contido no *software* utilizado é que cada divisão depende do valor de uma única variável. Assim, dentre o conjunto de variáveis independentes, é escolhida aquela que tem maior poder de segregação dos dados. Nas divisões subseqüentes são escolhidas, dentre todas as variáveis independentes, novamente aquelas que têm máxima segregação (ou forma o grupo mais homogêneo) nos sub-nós (nós filhos). Desta forma, multicolinearidade, por exemplo, deixa de ser problema na maioria dos algoritmos CART, pois as variáveis são analisadas uma a uma separadamente. Há casos em que são utilizadas combinações das variáveis independentes para divisão dos dados em sub-conjuntos. Breiman et al (1984) citam três tipos de combinações utilizados: (1) combinação linear; (2) combinações *Booleana*; e (3) adição de características (ou seja, combinações de variáveis sugeridas examinando a estrutura dos dados).

Neste trabalho, não se pretende discutir procedimentos que utilizam tais combinações de variáveis independentes. Para esses casos, multicolinearidade, por exemplo, poderia constituir problema, embora não se possa afirmar ao certo que a presença de multicolinearidade compromete os resultados (PIRAMUTHU, 2007).

Um dos principais motivos que levaram à escolha da AD contida no S-PLUS 6.1 é a sua capacidade de representar a natureza probabilística do objeto analisado, que no caso é o comportamento dos viajantes relacionado às viagens encadeadas. Ela é um modelo que reconhece que “Indivíduos homogêneos” (segundo a variável dependente) podem tomar diferentes decisões, e associa probabilidade às diferentes respostas possíveis (padrões de viagem). Na Tabela 3.3 apresenta-se um resumo das principais características do CART contido no S-PLUS 6.1.

Tabela 3.3 Características Gerais do S-PLUS 6.1 e aplicação no trabalho

| Características principais S-PLUS 6.1            |                                                                                         | Aplicação no trabalho                                                                                     |
|--------------------------------------------------|-----------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------|
| Tipo de variáveis de entrada                     | Catégoricas ou Contínuas                                                                | Catégoricas e Contínuas                                                                                   |
| Tipo de variáveis independentes                  | Catégoricas ou Contínuas                                                                | uso do solo (contínuas)<br>socioeconômicas (catégoricas e contínuas)<br>e atividades (catégoricas)        |
| Tipo de variável dependente                      | Catégorica (árvores de classificação)<br>Numérica (árvores de regressão)                | Catégoricas (padrões de viagem)                                                                           |
| Capacidade para controlar grande número de casos | Sim                                                                                     | Há grande número de casos                                                                                 |
| Capacidade para controlar falta de dados         | Não                                                                                     | Omitir dados ausentes                                                                                     |
| Segmentação                                      | Binária                                                                                 | Binária                                                                                                   |
| Regras de Divisão                                | Classificação- desvio<br>Regressão - Mínimo Quadrado<br>Apenas uma variável por divisão | Desvio (0,15)<br>Mínimo de observações nós terminais (50)<br>Apenas uma variável independente por divisão |
| Fornecimento de Árvore ótima                     | Não                                                                                     | Não                                                                                                       |
| Limitação Variável dependente                    | Até 128 valores                                                                         | Até 128 valores                                                                                           |
| Limitação Variáveis independentes                | Até 32 variáveis                                                                        | Até 32 variáveis                                                                                          |
| Apresentação dos Resultados                      | Gráfico / Relatório                                                                     | Gráfico / Relatório                                                                                       |
| Natureza dos Resultados                          | Gráfico - Determinístico<br>Relatório - Probabilístico                                  | Gráfico - Determinístico<br>Relatório - Probabilístico                                                    |

### 3.4. REGRESSÃO LINEAR MÚLTIPLA (RLM)

Sendo uma das mais utilizadas e versáteis técnicas tradicionais de AM, a Regressão Linear Múltipla (RLM), é aplicada em uma infinidade de casos, onde se deseja encontrar uma relação entre uma única variável dependente (numérica) e diversas variáveis independentes (numéricas ou *dummy*). Relação esta supostamente linear.

A RLM é uma extensão lógica dos princípios da Regressão Linear Simples (RLS). Desta vez, há um coeficiente para cada uma das variáveis independentes. Assim, a variável dependente é prevista a partir da combinação de todas as variáveis independentes multiplicadas por seus respectivos coeficientes adicionada a um termo que representa o resíduo (equação 3.2). A finalidade é encontrar a

combinação linear das variáveis independentes que forneça máxima correlação com a variável dependente.

$$Y_i = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n + \varepsilon_i \quad (3.2)$$

sendo,

$y$  = variável dependente

$b_0$  = constante

$b_1$  = coeficiente da primeira variável independente  $x_1$

$b_2$  = coeficiente da segunda variável independente  $x_2$

$b_n$  = coeficiente da  $n$ ésima variável independente  $x_n$

$\varepsilon_i$  = a diferença entre o valor previsto de  $y$  e o valor observado considerando o indivíduo/objeto  $i$ .

Um exemplo de aplicação de RLM, por exemplo, seria a previsão do número de cartões de crédito utilizados no domicílio em função do tamanho da família e da sua renda. Assim, o modelo, baseado na Equação 3.2, calcula os valores dos coeficientes para as variáveis independentes, assim como a constante. Para o exemplo em questão (adaptado de HAIR et al., 1998), tem-se a equação (3.3) genérica a seguir. Como há apenas duas variáveis independentes no modelo, pode-se representá-lo graficamente em três dimensões (Figura 3.5).

$$N_{\text{cartõesdecrédito}} = b_0 + b_1\text{tamanho da família} + b_2\text{renda familiar} + \varepsilon \quad (3.3)$$

sendo,

$N_{\text{cartõesdecrédito}}$  = número de cartões de crédito no domicílio

$b_0$  = constante

$b_1$  = coeficiente da variável tamanho da família

$b_2$  = coeficiente da variável renda familiar

$\text{tamanho da família}$  = número de pessoas na família

$\text{renda familiar}$  = renda familiar

$\varepsilon$  = termo residual

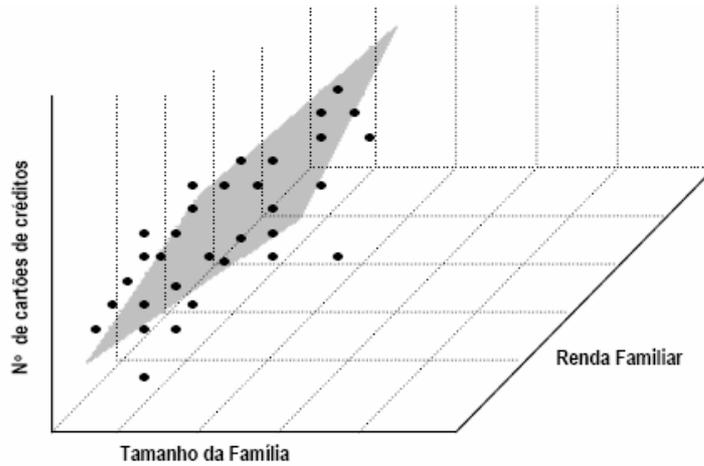


Figura 3.5 Exemplo de aplicação da RLM

A Equação 3.3 descreve o plano cinza no gráfico e os pontos representam os valores observados (Figura 3.5). O plano é ajustado com a finalidade de prever da melhor forma os dados observados. É relativamente fácil visualizar o modelo de regressão quando há apenas duas variáveis independentes. No entanto, quando se trata de múltiplas variáveis, embora não se possa visualizar graficamente o modelo, deve-se aplicar os mesmos princípios da RLS aos cenários mais complexos. Como a RLM é uma técnica vastamente utilizada e conhecida, optou-se em descrever os seus pontos básicos de forma bem sintética na Tabela 3.4. Tais descrições foram extraídas de Field (2005).

Tabela 3.4 Principais definições e suposições para modelos de RLM

| Conceitos básicos                                                                                                                                   |                                                                                                                          |                                                                                                                     |                                                                                                                                 |                                                                                                                                                                                                             |
|-----------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Soma dos quadrados                                                                                                                                  |                                                                                                                          |                                                                                                                     | Coefficiente de determinação                                                                                                    | Teste da significância dos coeficientes                                                                                                                                                                     |
| <b>SS<sub>T</sub></b><br>Diferença entre os valores observados e o valor da média da variável dependente (y)                                        | <b>SS<sub>R</sub></b><br>Diferença entre os valores de y previstos pelo modelo de regressão e os valores de y observados | <b>SS<sub>M</sub></b><br>Diferença entre os valores de y previstos pelo modelo de regressão e o valor da média de y | <b>R<sup>2</sup></b><br>Proporção da variância da variável dependente explicada pelo modelo de regressão<br>$\frac{SS_M}{SS_T}$ | <b>teste-t</b><br>Os coeficientes estimados através de um grande número de amostras são significamente diferentes de zero<br>$\frac{b}{SE_b}$ <i>b = coeficiente</i><br><i>SE<sub>b</sub> = erro padrão</i> |
| Principais Suposições                                                                                                                               |                                                                                                                          |                                                                                                                     |                                                                                                                                 |                                                                                                                                                                                                             |
| <b>Linearidade</b><br>a relação modelada deve ser linear. É difícil obter bons modelos lineares para o caso de relações não-lineares entre os dados | <b>Normalidade</b><br>Os resíduos estão normalmente distribuídos em torno de zero                                        | <b>Homocedasticidade</b><br>modelos de RLM supõem que a variância dos resíduos deve ser constante                   | <b>Não-Multicolinearidade</b><br>modelos de RLM supõem que a correlação entre variáveis independentes não deve ser alta         | <b>Resíduos independentes</b><br>modelos de RLM supõem que os resíduos não são correlacionados entre si. São independentes                                                                                  |

### 3.5 CONSIDERAÇÕES A RESPEITO DA APLICAÇÃO DA RLM EM CONJUNTO COM A AD

A RLM foi aplicada no trabalho para fins complementares. O objetivo principal da sua aplicação é associar inferência estatística às variáveis independentes. Através da RLM é possível mensurar a significância de cada uma das variáveis independentes, especialmente das variáveis de uso do solo propostas.

O uso da RLM foi realizado após aplicação da AD. A aplicação da AD traz resultados fáceis de serem interpretados, no entanto, por se tratar de uma técnica de DM, não se tem inferência estatística associada a variáveis.

Para aplicação da RLM, a variável dependente é dada pela probabilidade de ocorrência dos padrões de viagens mais frequentes em cada uma das amostras. Quanto às variáveis independentes, cada uma delas foi dividida em classes e a cada classe foi associada uma variável *dummy*. Os valores para escolha das classes das variáveis *dummy* foram obtidos com a aplicação da AD. A Figura 3.6 ilustra a aplicação da RLM de forma conjunta com a AD no trabalho.

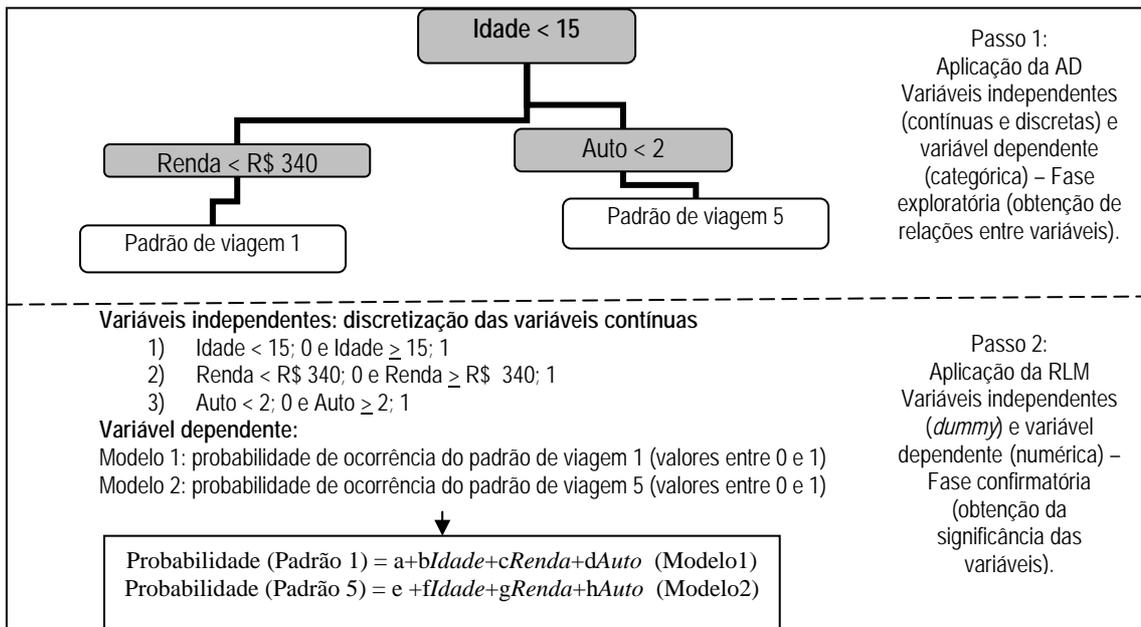


Figura 3.6 Aplicação da RLM no trabalho

A exemplo do trabalho realizado por Kawamoto (2003), as variáveis contínuas foram divididas em classes com o intuito de reduzir o efeito da não-linearidade entre variáveis independentes e variável dependente. Tanto a aplicação da AD e da RLM encontram-se em capítulos subseqüentes (Capítulo 7 e 9 respectivamente).

### **3.6 PRINCIPAIS CARACTERÍSTICAS DAS TÉCNICAS UTILIZADAS**

As técnicas utilizadas no trabalho possuem diferentes características, objetivos, vantagens e desvantagens. Para as finalidades propostas, pode-se afirmar que foram satisfatórias, mostrando-se ter viabilidade como ferramentas de auxílio à análise de dados. A Tabela 3.5 é um quadro resumo das três técnicas utilizadas. São representadas suas principais características.

Tabela 3.5 Principais características das técnicas de AM utilizadas

| <b>Técnica</b> | <b>Classificação</b> | <b>Descrição</b>                                                                                                         | <b>Caráter</b> | <b>Vantagens</b>                                                                                                                                                                                                | <b>Desvantagens</b>                                                                                                                                                    | <b>Aplicação no trabalho</b>                                                                                            |
|----------------|----------------------|--------------------------------------------------------------------------------------------------------------------------|----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------|
| AC             | Tradicional          | Formar grupos homogêneos de acordo com determinado critério                                                              | Exploratório   | Não é sensível a suposições como normalidade, linearidade e homocedasticidade. Descobrir estruturas presentes no banco de dados                                                                                 | Ausência de inferência estatística<br>Necessita de amostras representativas<br>Sensíveis à multicolinearidade, outliers<br>Cuidados especiais na escolha das variáveis | Agrupar os destinos escolhidos pelos indivíduos durante o período de um dia considerando faixas de distâncias           |
| CART           | Emergente            | Encontra relações entre variável dependente (numérica ou categórica) e variáveis independentes (numéricas e categóricas) | Exploratório   | Resultados facilmente interpretados<br>Não apresenta sensibilidade à presença de multicolinearidade (para o caso onde não há combinação de variáveis na divisão dos dados)<br>Padrões previamente desconhecidos | Ausência de inferência estatística<br>Geralmente necessitam de grandes quantidades de dados<br>Sensíveis a presença de outliers                                        | Encontrar relações entre padrões de viagens e características socioeconômicas, uso do solo e participação em atividades |
| RLM            | Tradicional          | Relação linear entre uma variável dependente (numérica) e múltiplas variáveis independentes (numéricas ou <i>dummy</i> ) | Confirmatório  | Amplamente utilizada e conhecida<br>Fácil interpretação dos resultados<br>Inferência estatística<br>Significância de cada variável independente                                                                 | Suposições de distribuição gaussiana dos dados de entrada. Sensíveis a não-linearidade, outliers, heterocedacidade, Multicolinearidade                                 | Mensurar a significância das variáveis independentes, especialmente as variáveis de uso do solo                         |

## Capítulo

# 4

## ESTUDO DE CASO

---

*Este capítulo apresenta o estudo de caso que é a Região Metropolitana de São Paulo (RMSP). Inicialmente é descrita de forma sucinta a região de estudo e as variáveis disponibilizadas pelo METRÔ-SP. Posteriormente, encontra-se a etapa de tratamento dos dados e obtenção das seis amostras finais.*

### 4.1. DESCRIÇÃO DA REGIÃO DE ESTUDO

O estudo é baseado nos dados da Pesquisa Origem-Destino (O-D) da Região Metropolitana de São Paulo (RMSP), realizada em 1997, por meio de entrevista domiciliar, pela Companhia do Metropolitano de São Paulo (METRÔ-SP). Na época, a RMSP contava com uma população de aproximadamente 17 milhões de habitantes, distribuídos em 39 municípios.

A região foi subdividida em sete sub-regiões (centro, leste, nordeste, norte, oeste, sudeste e sudoeste) e 389 zonas de tráfego. A pesquisa domiciliar, aplicada a todos os residentes nos domicílios amostrados, levantou dados socioeconômicos, características das viagens realizadas, localização do emprego e da escola de cada indivíduo residente e dados gerais do domicílio, tais como condição de ocupação e número de veículos.

O banco de dados da RMSP é composto originalmente de 98.780 indivíduos e 26.278 domicílios, representando um total de 0,61% do total de domicílios existentes na RMSP. Vale ressaltar que no presente trabalho foram utilizados os dados de indivíduos pertencentes à amostra. Os dados não foram expandidos.

Fazendo uma caracterização sucinta das sub-regiões que compõem a RMSP, pode-se afirmar que:

- As maiores concentrações populacionais ocorrem nas sub-regiões Centro e Sudeste. Tais sub-regiões também concentram as maiores rendas familiares, taxas de motorização e número de

empregos (principalmente no setor terciário). A sub-região Sudeste concentra um grande número de empregos no setor industrial;

- Os dados de circulação e transporte apontam um maior número de viagens originadas e atraídas às mesmas sub-regiões.
- Em termos de municípios, observa-se maior número de viagens, considerando modo de transporte motorizado ou não, no município de São Paulo, localizado no Centro.

A Figura 4.1 representa as sub-regiões e municípios que compõem a região de estudo

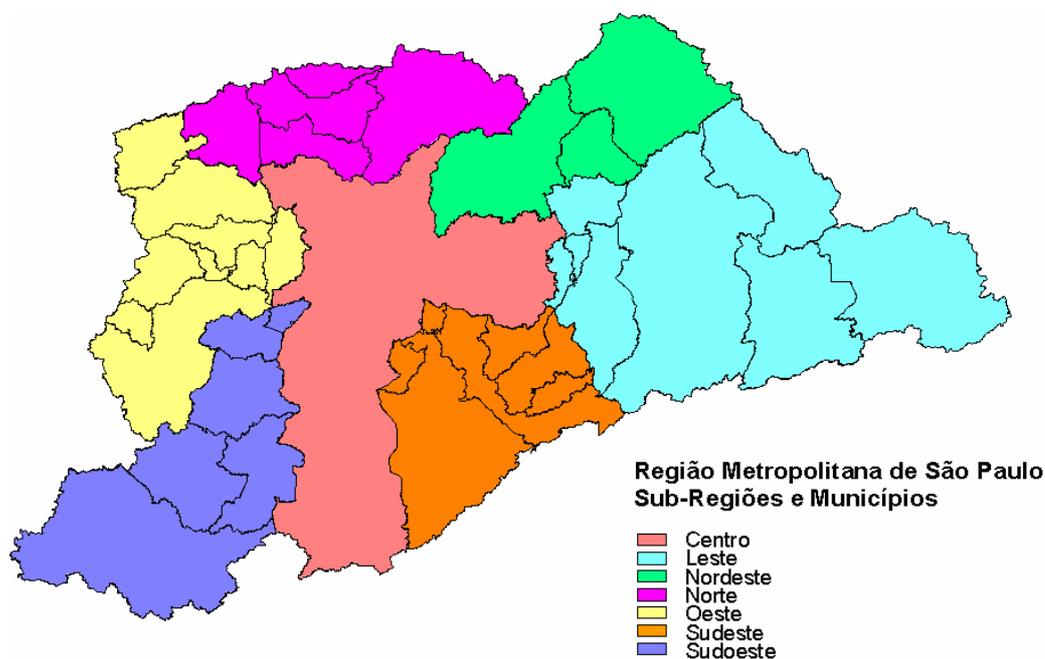


Figura 4.1 Região Metropolitana de São Paulo (Sub-Regiões e Municípios) – Fonte: METRÔ-SP

## 4.2. DADOS INDIVIDUAIS E DOMICILIARES

Inicialmente, foram cedidos pelo METRÔ-SP, dados referentes a características socioeconômicas e dos deslocamentos dos indivíduos na RMSP durante o dia anterior ao da entrevista. Em seguida, foram colocados à disposição dados relativos às zonas de tráfego, descritos na próxima seção.

Foram disponibilizadas variáveis individuais (como sexo, grau de instrução, ocupação principal, etc.), ou domiciliares (como por exemplo, número de automóveis no domicílio, total de pessoas na

família). Também foram cedidas variáveis referentes à viagem (como motivo da viagem, modo principal, etc.), além dos distritos e municípios da residência, da escola e do emprego. A descrição das variáveis individuais e domiciliares realizada pelo METRÔ-SP (1997), bem como características de viagem estão representadas respectivamente nas Tabelas 4.1 e 4.2.

Tabela 4.1 Variáveis individuais e domiciliares: fonte METRÔ-SP

| Variável                                 | Códigos                             | Variável                       | Códigos                               |                            |
|------------------------------------------|-------------------------------------|--------------------------------|---------------------------------------|----------------------------|
| Município do Domicílio                   | 1 a 39                              | Grau de Instrução              | 1 - Não-alfabetizado                  |                            |
| Distrito do Domicílio                    | 1 a 96                              |                                | 2 - Pré-Escola                        |                            |
| Tipo de Domicílio                        | 1 - Particular                      |                                | 3 - 1o. Grau Incompleto               |                            |
|                                          | 2 - Coletivo                        |                                | 4 - 1o. Grau Completo                 |                            |
|                                          | 3 - Favela                          |                                | 5 - 2o. Grau Incompleto               |                            |
| No. Moradores do Domicílio               |                                     |                                | 6 - 2o. Grau Completo                 |                            |
| Tempo de Residência no Município         |                                     |                                | 7 - Superior Incompleto               |                            |
| Tempo de Residência no Bairro            |                                     | 8 - Superior Completo          |                                       |                            |
| Condição de Moradia                      | 1 - Alugada                         | Condição de Atividade          | 1 - Ocupado                           |                            |
|                                          | 2 - Própria                         |                                | 2 - Ocupado Eventualmente             |                            |
|                                          | 3 - Cedida                          |                                | 3 - Em Licença                        |                            |
|                                          | 4 - Outros                          |                                | 4 - Não Ocupado                       |                            |
|                                          | 5 - Não Respondeu                   |                                | 5 - Aposentado/Pensionista            |                            |
| Rádios                                   |                                     |                                | 6 - Nunca Trabalhou                   |                            |
| Geladeiras                               |                                     |                                | 7 - Dona de Casa                      |                            |
| Tv. a Cores                              |                                     | 8 - Estudante                  |                                       |                            |
| Vídeo Cassetes                           |                                     | Ocupação Principal             | 01 - Assalariado com Carteira         |                            |
| Banheiros                                |                                     |                                | 02 - Assalariado sem Carteira         |                            |
| Automóveis                               |                                     |                                | 03 - Funcionário Público              |                            |
| Aspiradores de Pó                        |                                     |                                | 04 - Autônomo                         |                            |
| Máquinas de Lavar                        |                                     |                                | 05 - Empregador                       |                            |
| Empregados Domésticos                    |                                     |                                | 06 - Profissional Liberal             |                            |
| Microcomputadores                        |                                     |                                | 07 - Trab. Doméstico com Carteira     |                            |
| Telefones                                |                                     |                                | 08 - Trab. Doméstico sem Carteira     |                            |
| Celulares                                |                                     |                                | 09 - Dono de Negócio Familiar         |                            |
| Renda Familiar                           | 8 dígitos 2 casas decimais          |                                | 10 - Trabalhador Familiar             |                            |
| Classificação ABIPEME                    | 1 - A                               |                                | 11 - Não se Aplica                    |                            |
|                                          | 2 - B                               | 01 - Agrícola                  |                                       |                            |
|                                          | 3 - C                               | 02 - Construção Civil          |                                       |                            |
|                                          | 4 - D                               | 03 - Indústria                 |                                       |                            |
|                                          | 5 - E                               | 04 - Comércio                  |                                       |                            |
| Renda Familiar Atribuída                 | 8 dígitos 2 casas decimais          | Setor de Atividade             | 05 - Serviços de Transp. Carga        |                            |
| Situação Familiar                        | 1 - Chefe                           |                                | 06 - Serviços de Transp. de Passag.   |                            |
|                                          | 2 - Cônjuge                         |                                | 07 - Serviços Creditícios/Financeiros |                            |
|                                          | 3 - Filho(a)                        |                                | 08 - Serviços Pessoais                |                            |
|                                          | 4 - Parente/Agredado                |                                | 09 - Serviços de Alimentação          |                            |
|                                          | 5 - Empregado Residente             |                                | 10 - Serviços de Saúde                |                            |
|                                          | 6 - Visitante não Residente na RMSP |                                | 11 - Serviços de Educação             |                            |
| Idade                                    | (anos)                              |                                | 12 - Serviços Especializados          |                            |
| Sexo                                     | 1 - Masculino                       |                                | 13 - Serviços da Adm. Pública         |                            |
|                                          | 2 - Feminino                        |                                | 14 - Outros                           |                            |
| Estuda Atualmente ?                      | 1 - Não                             |                                | 15 - Não se Aplica                    |                            |
|                                          | 2 - Creche/Pré-Escola               |                                | Renda Individual                      | 8 dígitos 2 casas decimais |
|                                          | 3 - 1o./2o./3o. Graus               |                                | Usa Vale-Transporte ?                 | 1 - Sim                    |
|                                          | 4 - Outros                          |                                |                                       | 2 - Não                    |
| Primeiro Trabalho é igual a Residência ? | 1 - Sim                             |                                | Distrito da Escola                    | 1 a 96                     |
| Distrito do Segundo Trabalho             | 2 - Não                             | Município da Escola            | 1 a 39                                |                            |
| Município do Segundo Trabalho            | 1 a 39                              | Distrito do Primeiro Trabalho  | 1 a 96                                |                            |
| Segundo Trabalho é igual a Residência ?  | 1 - Sim                             | Município do Primeiro Trabalho | 1 a 39                                |                            |
|                                          | 2 - Não                             |                                |                                       |                            |

Tabela 4.2 Características de viagem: fonte METRÔ-SP

| Variável                     | Códigos                      | Variável                       | Códigos                      |                              |
|------------------------------|------------------------------|--------------------------------|------------------------------|------------------------------|
| Distrito de Origem           | 1 a 96                       | Motivo da Viagem a Pé          | 1 - Condução Cara            |                              |
| Município de Origem          | 1 a 39                       |                                | 2 - Condução Desconfortável  |                              |
| Pólo Gerador na Origem       |                              |                                | 3 - Ponto/Estação Distante   |                              |
| Distrito de Destino          | 1 a 96                       |                                | 4 - Condução Demora a Passar |                              |
| Município de Destino         | 1 a 39                       |                                | 5 - Condução Lotada          |                              |
| Pólo Gerador no Destino      |                              |                                | 6 - Viagem Demorada          |                              |
| Motivo na Origem             | 1 - Trabalho/Indústria       |                                | 7 - Horário Irregular        |                              |
|                              | 2 - Trabalho/Comércio        |                                | 8 - Pequena Distância        |                              |
|                              | 3 - Trabalho/Serviços        |                                | 9 - Outros Motivos           |                              |
|                              | 4 - Escola/Educação          |                                |                              |                              |
|                              | 5 - Compras                  | Hora da Saída                  |                              |                              |
|                              | 6 - Médico/Dentista/Saúde    | Minutos da Saída               |                              |                              |
|                              | 7 - Recreação/Visitas        | Minutos Andando desde a Origem |                              |                              |
|                              | 8 - Residência               | Hora da Chegada                |                              |                              |
|                              | 9 - Outros                   | Minutos da Chegada             |                              |                              |
| Motivo no Destino            | idem ao anterior             | Minutos Andando até o Destino  |                              |                              |
| Servir Passageiro na Origem  | 1 - Sim                      | Duração da Viagem (em minutos) | 01 - Ônibus                  |                              |
|                              | 2 - Não                      |                                | 02 - Ônibus Fretado          |                              |
| Servir Passageiro no Destino | 1 - Sim                      |                                | 03 - Transporte Escolar      |                              |
|                              | 2 - Não                      |                                | 04 - Dirigindo Automóvel     |                              |
| Modo 1                       | 01 - Ônibus                  |                                | Modo Principal               | 05 - Passageiro de Automóvel |
|                              | 02 - Ônibus Fretado          |                                |                              | 06 - Táxi                    |
|                              | 03 - Transporte Escolar      |                                |                              | 07 - Lotação/Perua           |
|                              | 04 - Dirigindo Automóvel     |                                |                              | 08 - Metrô                   |
|                              | 05 - Passageiro de Automóvel |                                |                              | 09 - Trem                    |
|                              | 06 - Táxi                    | 10 - Moto                      |                              |                              |
|                              | 07 - Lotação/Perua           | 11 - Bicicleta                 |                              |                              |
|                              | 08 - Metrô                   | 12 - A Pé                      |                              |                              |
|                              | 09 - Trem                    | 13 - Outros                    |                              |                              |
|                              | 10 - Moto                    | Tipo de Viagem                 |                              | 1 - Coletivo                 |
|                              | 11 - Bicicleta               |                                |                              | 2 - Individual               |
|                              | 12 - A Pé                    |                                |                              | 3 - A pé                     |
|                              | 13 - Outros                  |                                |                              |                              |
| Modo 2                       | idem ao anterior             |                                |                              |                              |
| Modo 3                       | idem ao anterior             |                                |                              |                              |
| Modo 4                       | idem ao anterior             |                                |                              |                              |

### 4.3 CARACTERÍSTICAS POR ZONAS DE TRÁFEGO

Posteriormente, foram cedidos pelo METRÔ-SP dados agregados, ou seja, características por zonas de tráfego. Dados gerais como número de domicílios, famílias, população, matrículas, empregos, automóveis, viagens produzidas, viagens atraídas, densidade populacional, densidade de emprego e taxa de motorização.

Também foram disponibilizados dados mais específicos, como população por faixa etária (em cada uma das zonas de tráfego), população por gênero, população por faixa de renda familiar,

população por ocupação principal, tipo de emprego (agrícola, construção civil, indústria, etc.), número de matrículas por grau de instrução (creche/pré-escola, 1º grau, 2º grau e superior).

As zonas com maior número de viagens atraídas e produzidas são Carapicuíba, Mauá e Pimentas. Entretanto, analisando em termos de densidade de viagens (por área) atraídas e produzidas, maiores densidades são encontradas nas zonas Sé, República e Masp. Aquela de maior taxa de motorização é Chácara Flora, seguida de Brooklin.

As Figuras 4.2, 4.3, a seguir, são mapas temáticos da densidade populacional e da densidade de empregos respectivamente. Observa-se uma concentração na região central dos maiores valores de densidade para ambos os casos.

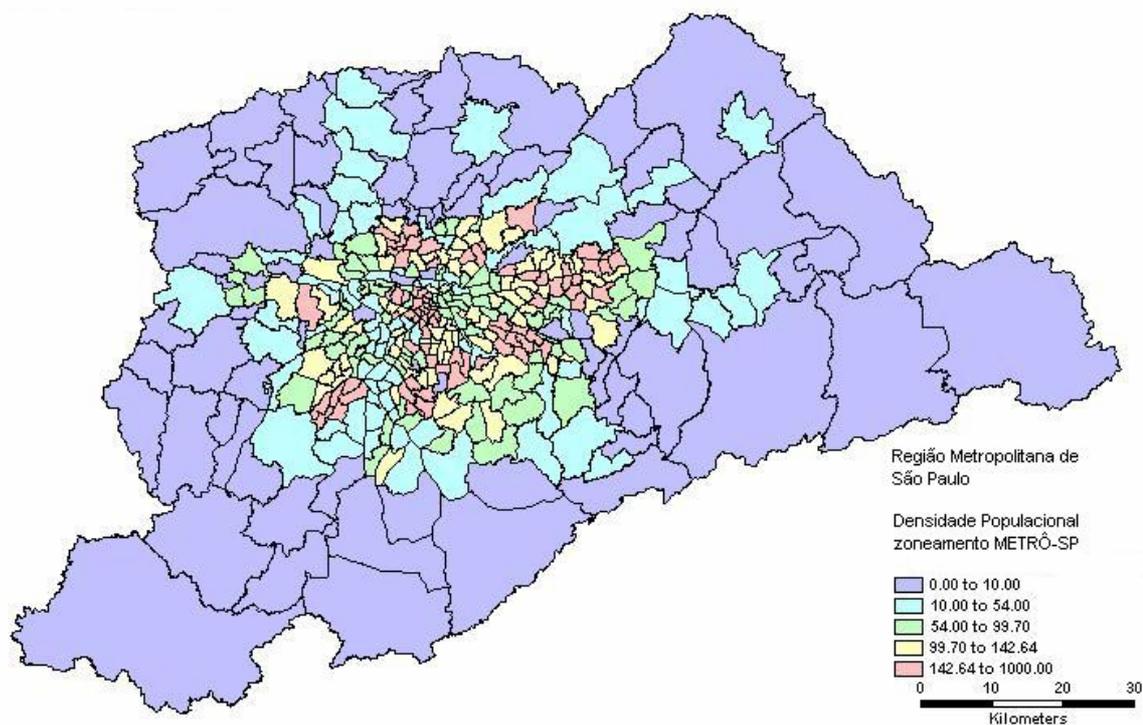


Figura 4.2 RMSP. Densidade Populacional. Pesquisa O/D 1997. Fonte: METRÔ-SP

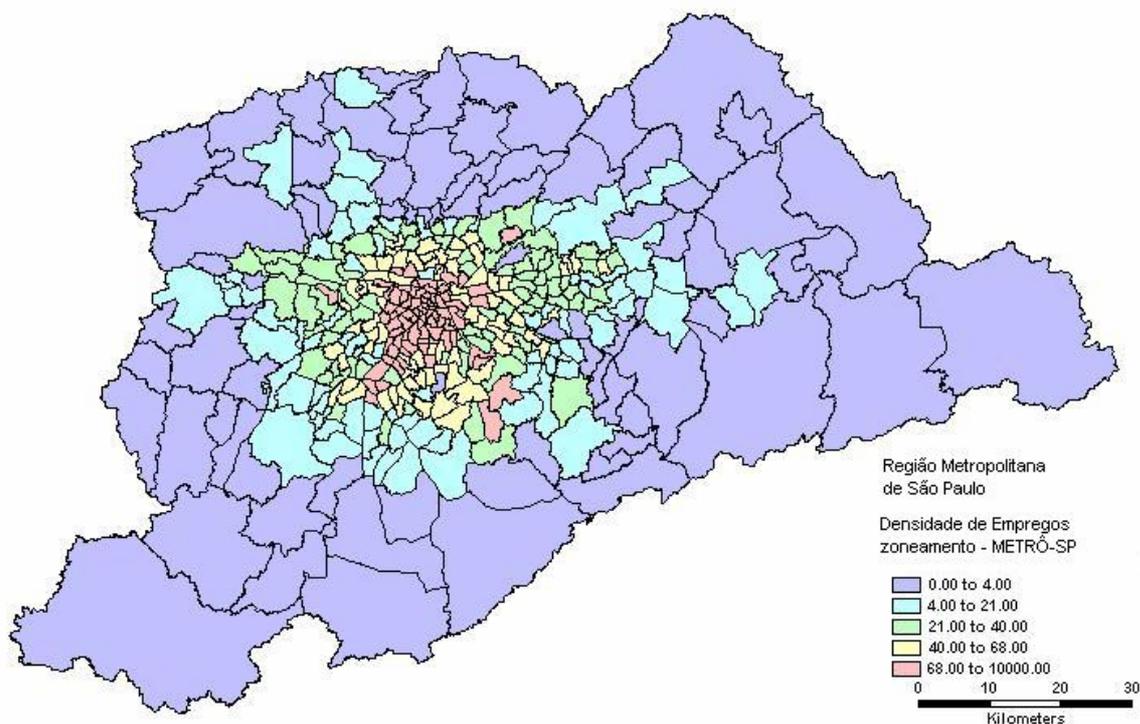


Figura 4.3 RMSP. Densidade de Empregos. Pesquisa O/D 1997. Fonte: METRÔ-SP

Já as Figuras 4.4, 4.5, 4.6, 4.7, 4.8, 4.9 apresentam respectivamente mapas temáticos do número de empregos desagregado por setor (indústria, comércio e serviços) e do número de matrículas escolares por grau de instrução (até o 1º grau, no 2º grau, universitários). As referidas figuras ilustram a distribuição espacial de tais atividades, auxiliando análises posteriores.

Observa-se que no caso de empregos na indústria (Figura 4.4), há uma concentração maior em algumas regiões: (1) a região que compreende as zonas de tráfego Pimentas, Rodovia Presidente Dutra e Guarulhos; (2) a região do grande ABCD, principalmente nas zonas de tráfego de São Bernardo do Campo, Santo André, Ruge Ramos e Diadema; e (3) a região que compreende zonas pertencentes aos municípios de Barueri e Osasco. Apesar de se verificar uma leve concentração nas zonas centrais, pode-se afirmar que há basicamente três pólos industriais principais.

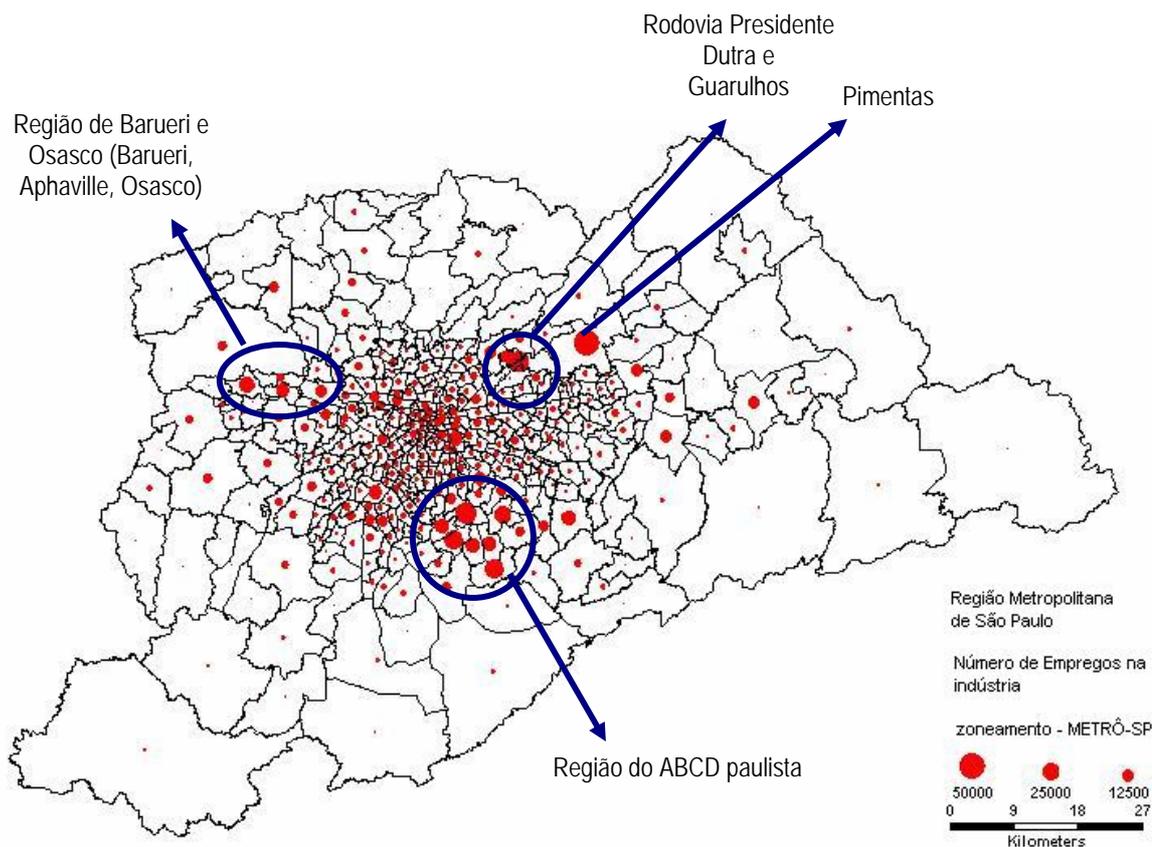


Figura 4.4 RMSP. Distribuição de empregos na indústria. Pesquisa O/D 1997. Fonte: METRÔ-SP

Verificam-se distribuições bastante parecidas na RMSP quanto ao número de empregos nos setores comerciais e de serviços (Figuras 4.5 e 4.6 respectivamente). Em ambos os casos há forte concentração de tais atividades na região central. Para o caso de empregos no comércio observa-se maior número nas zonas de tráfego República, Brás, Bresser e Parque D. Pedro. Já no caso do setor de serviços, citam-se as zonas de tráfego República, Consolação, Sé e Masp como aquelas que possuem maior concentração de tais atividades. Também observa-se grande número de empregos nas duas atividades referidas na região do ABCD paulista, principalmente nas zonas de tráfego Mauá, Santo André e São Bernardo do Campo.

Considerando a distribuição do número de empregos no comércio (Figura 4.5), pode-se citar, além dos pólos mencionados acima, mais quatro pólos de concentração na região: (1) região que compreende as zonas Itaquaquecetuba, Suzano e Mogi das Cruzes; (2) região que abrange as zonas Guarulhos e Vila Galvão; (3) região onde se encontram as zonas Carapicuíba, Jardim Veloso e Santo

Antônio; e (4) região um pouco a sudoeste do centro, abrangendo principalmente as zonas Jardim São Luis, Centro Empresarial e Vila Santa Catarina. Nota-se que a atividade comercial é relativamente dispersa pela RMSP e possui basicamente seis pólos de concentração de empregos.

Para o caso da distribuição de empregos no setor de serviços (Figura 4.6), observa-se além da região central e do ABCD, mais três pólos de concentração: (1) Mogi das Cruzes e Suzano; (2) região que compreende zonas dos municípios de Osasco e Barueri; e (3) região que abrange as zonas Vila Galvão, Jardim América e Guarulhos. Pode-se afirmar, a partir dos mapas temáticos, que atividades comerciais e de serviços possuem um maior número de pólos de concentração dispersos pela RMSP.

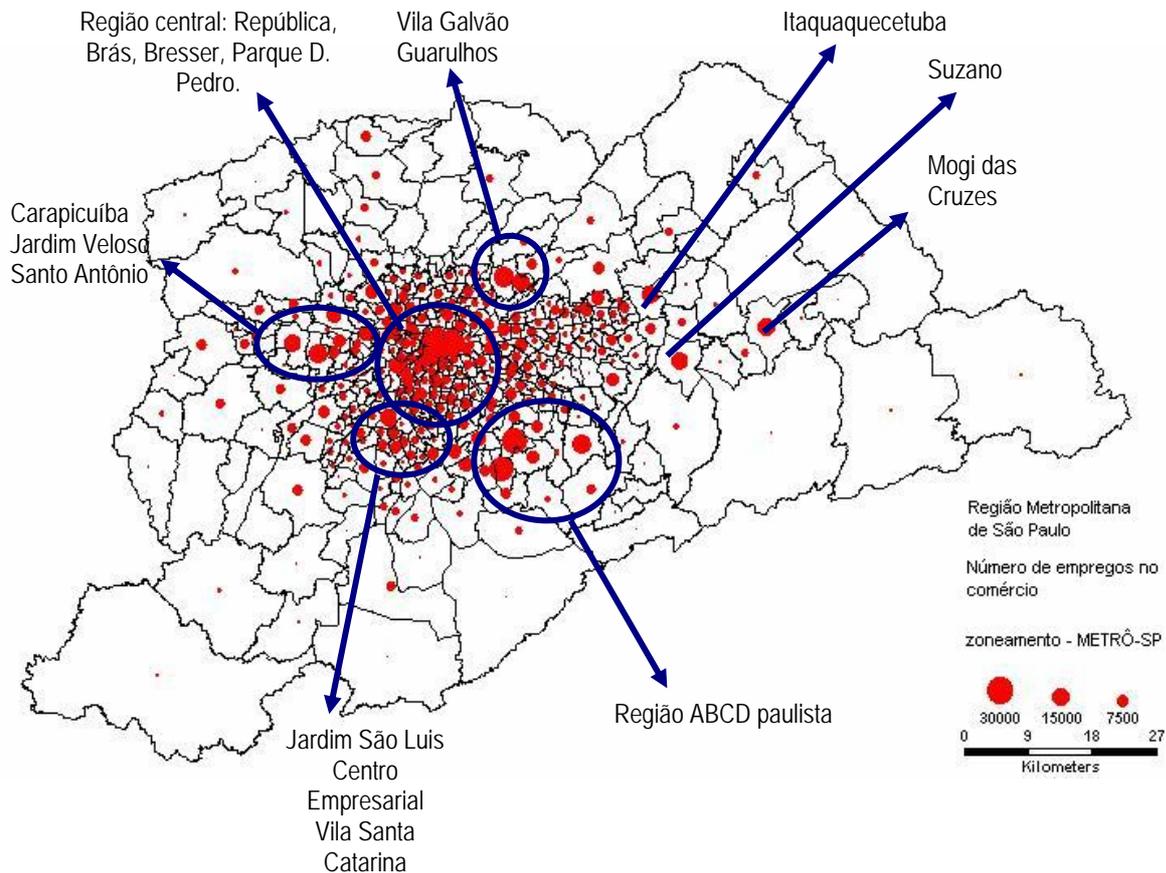


Figura 4.5 RMSP. Distribuição de empregos no comércio. Pesquisa O/D 1997. Fonte: METRÔ-SP

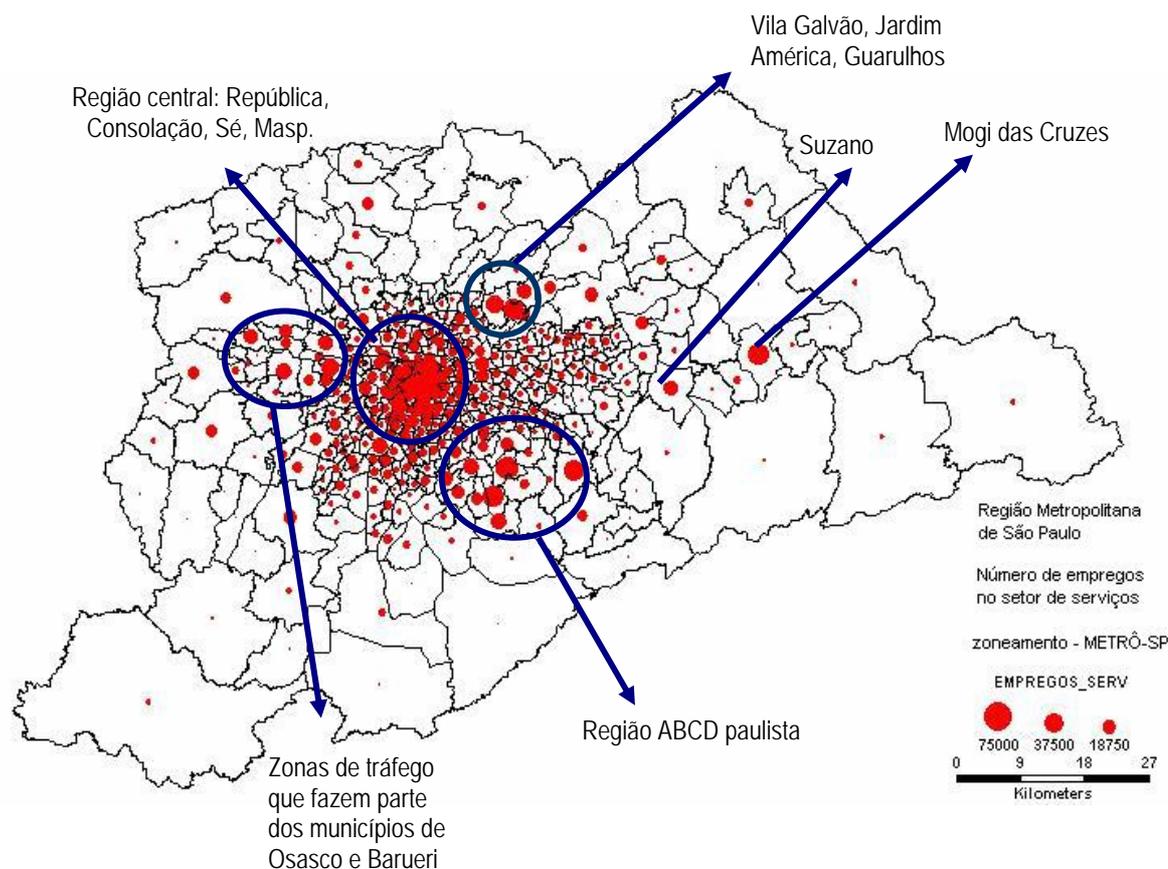


Figura 4.6 RMSP. Distribuição de empregos no setor de serviços. Pesquisa O/D 1997. Fonte: METRÔ-SP

Considerando o número de matrículas escolares segregadas por grau de instrução (Figuras 4.7, 4.8 e 4.9), pode-se observar uma concentração maior de atividades em apenas algumas zonas de tráfego, conforme aumento do nível de escolaridade. Assim, para o caso de matrículas até o 1º grau, verifica-se uma dispersão pela RMSP, ou seja, provavelmente há oferta de escolas para as crianças até o 1º grau em diversas localizações da região de estudo.

No entanto, para matrículas no 2º grau (Figura 4.8), já se observa concentração em algumas zonas de tráfego: (1) nas zonas de tráfego centrais (principalmente a zona Consolação); (2) na zona correspondente a Mogi das Cruzes; e (3) Guarulhos.

Para o caso de número de matrículas observadas em universidades (Figura 4.9), a concentração é ainda em localizações mais restritas. Predominantemente na zona Cidade Universitária, seguida de Vila Gerti. Também é observado um bom número de matrículas no ensino superior em algumas zonas

centrais, principalmente Consolação e Pacaembu e na zona de tráfego correspondente à Mogi das Cruzes.

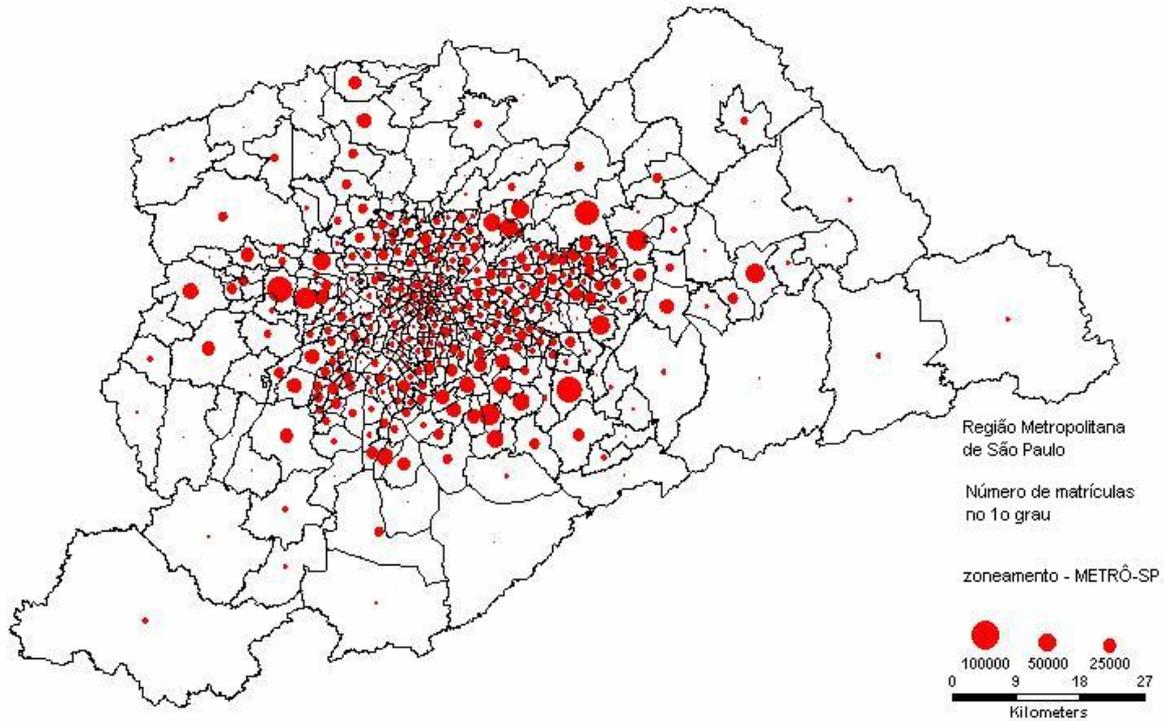


Figura 4.7 RMSP. Distribuição de matrículas até 1º grau. Pesquisa O/D 1997. Fonte: METRÔ-SP

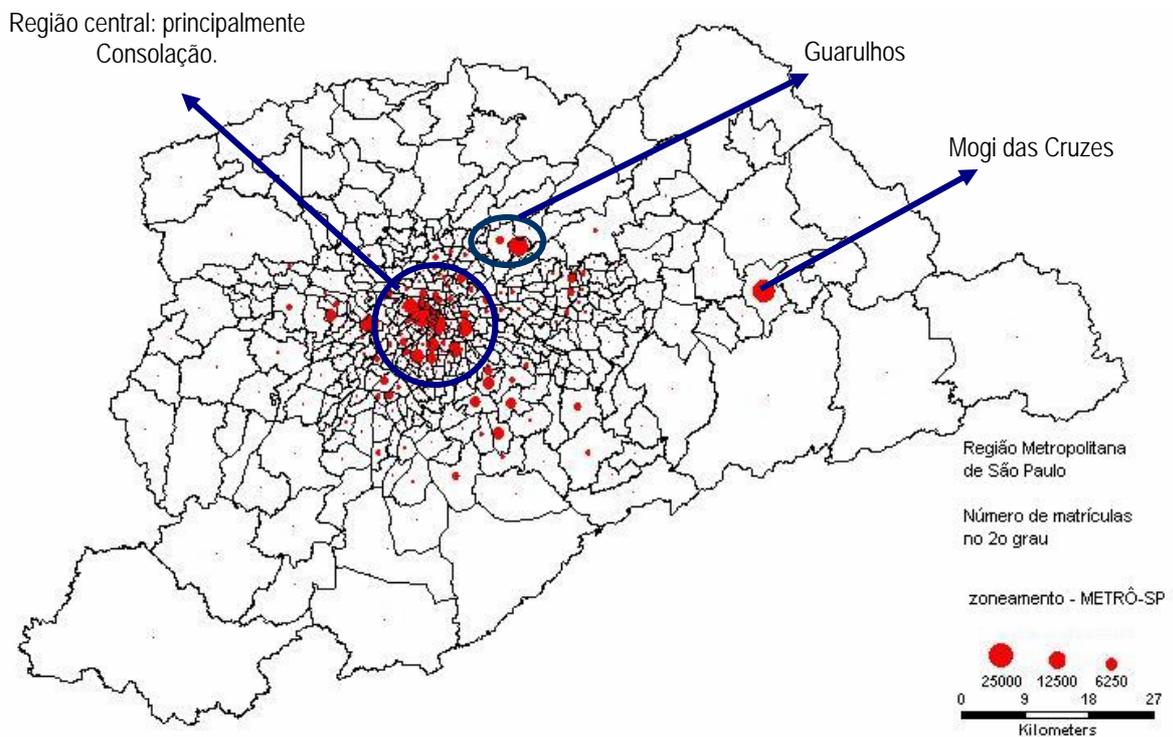


Figura 4.8 RMSP. Distribuição de matrículas no 2º grau. Pesquisa O/D 1997. Fonte: METRÔ-SP

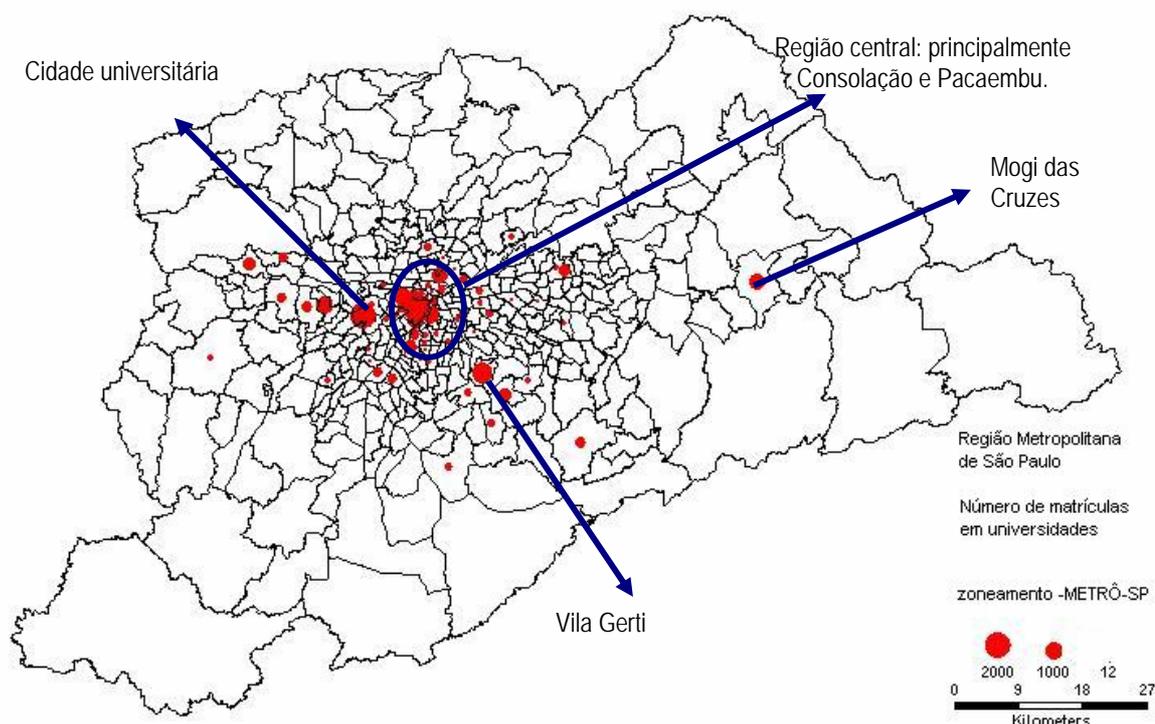


Figura 4.9 RMSP. Distribuição de matrículas no 3º grau. Pesquisa O/D 1997. Fonte: METRÔ-SP

A verificação da distribuição geográfica das atividades na região é bastante útil para demais análises. Pode-se conjecturar, por exemplo, que no caso da indústria e matrículas no ensino superior, onde há concentração em apenas algumas localizações, haja uma forte influência do uso do solo nos deslocamentos individuais.

A análise da distribuição de empregos e matrículas (em números absolutos) auxiliou também na etapa de tratamento de dados, onde optou-se por separar a amostra inicial por setor de atividade dos indivíduos. Esta etapa será descrita na próxima subseção.

A seguir, encontra-se a Tabela 4.3, a qual representa as variáveis agregadas por zona de tráfego, disponibilizadas pelo METRÔ-SP.



Tabela 4.3 Variáveis agregadas por zona de tráfego

| Zona | Dados gerais                                                |                          |                     |                                                |                      |                        |                                     |                           |                    |                |                        |                      |                      |
|------|-------------------------------------------------------------|--------------------------|---------------------|------------------------------------------------|----------------------|------------------------|-------------------------------------|---------------------------|--------------------|----------------|------------------------|----------------------|----------------------|
|      | Nome                                                        | Domicílios               | Famílias            | População                                      | Matrículas           | Empregos               | Automóveis                          | Viagens Produzidas        | Viagens Atraídas   | Área(ha)       | Densidade Populacional | Densidade Emprego    | Taxa Motorização*    |
|      | População por Faixa Etária                                  |                          |                     |                                                |                      |                        |                                     |                           |                    |                |                        | População por Gênero |                      |
|      | até 3 anos                                                  | 4 a 6 anos               | 7 a 10 anos         | 11 a 14 anos                                   | 15 a 17 anos         | 18 a 22 anos           | 23 a 29 anos                        | 30 a 39 anos              | 40 a 49 anos       | 50 a 59 anos   | mais de 60 anos        | Masculino            | Feminino             |
|      | População por Faixas de Renda (Em Reais de Outubro de 1997) |                          |                     |                                                |                      |                        | População por condição de atividade |                           |                    |                |                        |                      |                      |
|      | até 250                                                     | 251 - 500                | 501 - 1.000         | 1.001 - 1.500                                  | 1.501 - 3.600        | 3601 e mais            | Ocupado                             | Ocupado eventualmente     | Em Licença         | Não ocupado    | Aposentado             | Nunca trabalhou      | Dona de casa         |
|      | População por Ocupação Principal                            |                          |                     |                                                |                      |                        |                                     |                           |                    |                |                        | Renda por zona       |                      |
|      | Assalariado com carteira                                    | Assalariado sem carteira | Funcionário Público | Autônomo                                       | Empregador           | Profissional Liberal   | Trab.Domést. com carteira           | Trab.Domést. sem carteira | Dono Neg. Familiar | Trab. Familiar | Não se Aplica          | Renda Total          | Renda Média Familiar |
|      | Número de Empregos por Setor de Atividade                   |                          |                     |                                                |                      |                        |                                     |                           |                    |                |                        |                      |                      |
|      | Agrícola                                                    | Construção               | Indústria           | Comércio                                       | Serviços de Carga    | Transporte passageiros | Credícios Financeiros               | Pessoais                  | Alimentação        | Saúde          | Educação               | Especializado        | Adm. Pública         |
|      | Número de matrículas escolares                              |                          |                     | Produção de viagens/ Atração de viagens (tipo) |                      |                        |                                     |                           |                    |                |                        |                      |                      |
|      | Creche e Pré-Escola                                         | 1º/2º/3º Graus           | Outros              | Trabalho Indústria                             | Trabalho Comércio    | Trabalho Serviços      | Educação                            | Compras                   | Saúde              | Lazer          | Residência             | Outros               | Total                |
|      | Produção de viagens/ atração de viagens (modo)              |                          |                     |                                                |                      |                        |                                     |                           |                    |                |                        |                      |                      |
|      | Ônibus                                                      | Fretado                  | Escolar             | Dirigindo Automóvel                            | Passageiro Automóvel | Táxi                   | Lotação                             | Metrô                     | Trem               | Moto           | Bicicleta              | A pé                 | Outros               |

\* Número de automóveis particulares por 1.000 habitantes



#### 4.4 TRATAMENTO DOS DADOS

Na etapa de tratamento dos dados procura-se separar as amostras finais eliminando dados incompletos ou aqueles que não fazem parte dos objetivos da análise. Partiu-se então da amostra, disponibilizada pelo METRÔ-SP, composta por 98.780 indivíduos e foram incorporados dados agregados da zona de tráfego de origem de cada indivíduo. Também foram adicionadas as distâncias de todas as viagens (as distâncias euclidianas foram geradas com auxílio do TRANSCAD 3.6 e correspondiam às distâncias entre centróides de zonas).

Deste modo, a amostra inicial continha 98.780 linhas (indivíduos) e 204 colunas (com dados socioeconômicos individuais e domiciliares, dados de todas as viagens realizadas no dia anterior ao da entrevista e variáveis associadas às características das zonas da residência de todos os indivíduos).

A fase de tratamento de dados incluiu as seguintes etapas:

- (1) Eliminação de dados incompletos;
- (2) Eliminação de indivíduos que tenham realizado uma ou mais de quatro viagens, com intuito de limitar a complexidade da análise (4.080 indivíduos/4,13% da amostra inicial);
- (3) Eliminação dos indivíduos que não tenham tido como origem e destino final a residência;
- (4) Eliminação dos indivíduos que não viajaram no dia anterior ao da pesquisa;
- (5) Separação das amostras por setor de emprego (para o caso de trabalhadores) e grau de instrução (para o caso de estudantes), obtendo então seis amostras desagregadas:

5.1 Indivíduos que trabalham no setor industrial;

5.2 Indivíduos que trabalham no setor comercial;

5.3 Indivíduos que trabalham no setor de serviços;

5.4 Estudantes até o 1º grau;

5.5 Estudantes do 2º grau;

5.6 Estudantes de ensino superior;

(6) Seleção das variáveis agregadas por zona de tráfego com intuito de limitar a análise e o processamento dos dados. Nas amostras finais utilizaram-se apenas as variáveis relativas ao número de empregos por setor de atividade (industrial, comercial e serviços) e matrículas escolares (1º/2º/3º graus). Tais variáveis estão estritamente relacionadas às amostras finais (trabalhadores na indústria, no comércio, serviços, etc.);

(7) Eliminação dos padrões de viagem menos frequentes em cada uma das amostras devido à limitação de 128 categorias de variáveis dependentes do *software* utilizado.

Optou-se por fazer uma análise desagregada por setor de atividade dos indivíduos para melhor investigação dos resultados. O intuito era evidenciar a possível influência do uso do solo para cada uma das classes de atividades. Desta forma, deveriam ser analisadas amostras compostas por indivíduos que trabalham na indústria, comércio e serviços separadamente, assim como estudantes de diferentes escolaridades.

Espera-se que o uso do solo exerça maior influência em algumas amostras (como industriários e universitários, por exemplo), pois constata-se que há uma relativa concentração do número de empregos na indústria e número de matrículas em universidades em determinadas zonas de tráfego da RMSP (Figuras 4.4 e 4.9 respectivamente). O número de empregos na indústria da RMSP, por exemplo, concentra-se nas regiões do ABCD paulista, Guarulhos, Pimentas, Osasco e Barueri enquanto que observa-se uma distribuição predominantemente central de atividades terciárias, com diversos pólos dispersos pela região.

Assim, supôs-se que variáveis de uso do solo relacionadas à indústria, por exemplo, poderiam apresentar de forma mais clara a influência da distribuição de vagas de trabalho em relação ao deslocamento espacial dos trabalhadores.

A Tabela 4.4 traz as principais características das seis amostras analisadas.

Tabela 4.4 Características das amostras finais

| Amostras                        |                        |                 |                                                 |                                                   |
|---------------------------------|------------------------|-----------------|-------------------------------------------------|---------------------------------------------------|
|                                 | linhas<br>(indivíduos) | Colunas         |                                                 |                                                   |
|                                 |                        | Socioeconômicas | Característica agregada<br>(zona da residência) | Características das Viagens<br>(de duas a quatro) |
| Industriários                   | 4.102                  | 18              | (1) Empregos na indústria                       | 24                                                |
| Comerciários                    | 6.043                  | 18              | (1) Empregos no comércio                        | 24                                                |
| Trabalhadores<br>setor serviços | 8.328                  | 18              | (1) Emprego em serviços                         | 24                                                |
| Estudantes<br>até o 1º grau     | 13.839                 | 18              | (1) Matrículas até o 1º grau                    | 24                                                |
| Estudantes<br>no 2º grau        | 2.143                  | 18              | (1) Matrículas no 2º grau                       | 24                                                |
| Estudantes<br>universitários    | 568                    | 18              | (1) Matrículas no 3º grau                       | 24                                                |
| Total                           | 35.023                 |                 |                                                 |                                                   |

## Capítulo

# 5 VARIÁVEIS INDEPENDENTES

---

*Este capítulo apresenta a etapa do método definida como representação das variáveis independentes. É indispensável para o objetivo do trabalho representar adequadamente as variáveis envolvidas. São descritas as variáveis relacionadas à participação em atividades e características socioeconômicas. A proposta do grupo de variáveis relacionadas ao uso do solo, que abrangeu algumas etapas, é apresentada detalhadamente, já que é um dos objetivos secundários da pesquisa atual.*

### 5.1. CONSIDERAÇÕES INICIAIS

Conforme o objetivo principal da pesquisa, as variáveis independentes (categóricas ou numéricas) utilizadas no trabalho pertencem a três grupos principais de variáveis: (1) participação em atividades; (2) características socioeconômicas; e (3) características de uso do solo.

Nesta etapa são selecionadas as variáveis que, de certo modo, possuem inter-relações com os padrões de viagem. As variáveis pertencentes aos dois primeiros grupos foram escolhidas com base na literatura vigente (RECKER, 1995; YE; PENDYALA; GOTTARDI, 2007; BHAT; SIVAKUMAR; AXHAUSEN, 2003; BHAT; SRINIVASAN, 2005) e disponibilidade de dados. Para representar participação em atividades e atributos socioeconômicos, as variáveis disponibilizadas pelo METRÔ-SP foram adaptadas.

O grande desafio, integrante dos objetivos secundários do trabalho, diz respeito a como mensurar as variáveis de uso do solo, de modo que estas atendam aos seguintes critérios: (1) representem bem a intensidade e distribuição de atividades no meio urbano; (2) sejam viáveis para aplicação do algoritmo CART; e (3) estatisticamente significativas.

Nas próximas seções serão descritas as variáveis independentes utilizadas. Para propor as variáveis de uso do solo foram necessárias algumas etapas que serão apresentadas na seção 5.4.

## 5.2 PARTICIPAÇÃO EM ATIVIDADES

Um dos princípios mais conhecido e largamente aceito é o de que demanda por viagens é derivada de demanda por atividades. Essa é a principal motivação para o estudo de demanda por transportes baseado em atividades. A análise das Viagens baseada em Atividades relaciona a viagem com a necessidade de realização de atividades específicas, considerando quando, onde e por quanto tempo tal atividade é desempenhada.

Pode-se afirmar que a programação dos itinerários das viagens realizadas pelos indivíduos no período de um dia obedece a seqüências de decisões (escolha dos destinos, do modo de transporte, da rota, etc.). Um dos primeiros fatores considerados seria a necessidade de realizar diversas atividades fora do domicílio. Assim, observam-se relações estritas entre participação em atividades e padrões de viagens, já que estes são derivados diretamente de tais atividades.

Participação dos indivíduos em atividades é representada de diversas formas em diferentes trabalhos: (1) variáveis dependentes – “padrões de atividades” (*activity patterns, activity sequences, activity travel patterns*); ou (2) variáveis independentes. Muitas vezes, observa-se a separação das atividades em classes principais. A classificação das atividades em diferentes categorias pode variar de acordo com trabalhos revisados na literatura (GOLOB; MCNALLY, 1997; AXHAUSEN, 2000; BHAT; KOPPELMAN; 2000). No entanto, pode-se citar três classes principais de atividades:

(A) Subsistência (*subsistence*) – trabalho (ou atividades relacionadas ao trabalho) e estudo;

(B) Relacionadas ao domicílio ou família (*maintenance*) – refeições, cuidados médicos, obrigações domiciliares; buscar ou pegar passageiros, etc.;

(C) Lazer (*leisure*) – visitas, atividades culturais, exercícios, *hobbies*, etc.

Na presente pesquisa, foram representadas apenas participação dos indivíduos em atividades de subsistência. Assim, o grupo referente à participação em atividades é composto por duas variáveis categóricas:

- Trabalha: (1) Assalariado; (2) Autônomo; (3) Não trabalha;
- Estuda: (1) Não estuda; (2) Creche/pré-escola; (3) 1º 2º 3º graus; (4) Outros.

A primeira variável categórica "TRABALHA" foi adaptada dos dados disponibilizados do METRÔ-SP. A variável base, pertencente ao conjunto de dados disponibilizados, foi "OCUPAÇÃO PRINCIPAL". Suas onze categorias foram agrupadas em três, formando assim a nova variável "TRABALHA": (1) Assalariados (Assalariado com Carteira, Funcionário Público, Empregador, Trabalhador Doméstico com Carteira, Trabalhador Familiar); (2) Autônomos (Assalariado sem Carteira, Autônomo, Profissional Liberal, Trabalhador Doméstico sem carteira, Dono de Negócio Familiar); (3) Não Trabalha (Não se Aplica).

Optou-se por incluir a categoria "Autônomos" na variável adaptada devido ao aumento significativo de trabalhadores liberais na RMSP, principalmente comparando as Pesquisas O-D de 1987 e 1997. Observaram-se consideráveis transformações nas atividades econômicas desempenhadas durante as décadas de 80 e 90 na região. Estas transformações dizem respeito à redução do emprego industrial, expansão e consolidação das atividades de serviço e deterioração das condições de inserção no mercado de trabalho, expressas pela redução dos empregos regulamentados e pela elevação do patamar das taxas de desemprego (MONTALI, 2003). Considerando a metodologia de cálculo do IBGE, a RMSP apresentava uma taxa de desemprego de 3,9% nos anos 80, 5,8% nos anos 90 e 8,4% no primeiro semestre de 2002.

A variável categórica "ESTUDA", no entanto, não sofreu adaptações. Foi utilizada na sua forma original, conforme dados fornecidos.

### 5.3 CARACTERÍSTICAS SOCIOECONÔMICAS

Características socioeconômicas estão fortemente relacionadas ao comportamento humano de maneira geral. Alguns atributos (como renda, por exemplo) fornecem uma base apropriada para segmentação da população e compreensão de comportamentos individuais, em particular em relação às viagens.

Na literatura existente que avalia relações entre viagens e características individuais e domiciliares predomina a afirmação de que características dos deslocamentos das pessoas podem ser impostas:

- Pelo seu poder econômico: definido não somente como renda, mas também ocupação e posse de automóveis (HANSON; HANSON, 1981; DOUBLEDAY, 1977; MITCHELL; TOWN, 1977);
- Papel do domicílio e estrutura familiar: definidos como características gerais domiciliares - número de pessoas no domicílio, presença de crianças, posição do indivíduo no domicílio (SIMMA, AXHAUSEN, 2001; BALASUBRAMANIAM; GOULIAS, 1999; PITOMBO; KAWAMOTO, 2003);
- Sexo (HANSON ; HANSON, 1980; MCGUCKIN; MURAKAMI , 1999; SARMIENTO, 1996).

Baseando-se na literatura foram escolhidas algumas variáveis do banco de dados. Outras variáveis foram adaptadas das originais. A Tabela 5.1 mostra as características socioeconômicas utilizadas como variáveis independentes.

Tabela 5.1 Variáveis socioeconômicas

|                                                |                                                                                                                                                                                                                                                                                                                     |
|------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>Poder econômico</b>                         | <b>Renda Individual (R\$); Renda Familiar (R\$); N° de automóveis no domicílio; N° provável de carteiras de habilitação no domicílio*</b>                                                                                                                                                                           |
| <b>Papel do domicílio e estrutura familiar</b> | <b>Posição do indivíduo no domicílio</b> (1: chefe; 2: cônjuge; 3: filho; 4: parente/agregado; 5: Empregado residente; 6: Visitante não residente na RMSP);<br><b>Total de pessoas no domicílio; Tipo de chefe do domicílio</b> (1: Homem; 2: Mulher)*;<br><b>N° de crianças no domicílio</b> (menores de 11 anos)* |
| <b>Sexo</b>                                    | <b>Sexo</b> (1: Masculino; 2: Feminino)                                                                                                                                                                                                                                                                             |
| <b>Outras variáveis</b>                        | <b>Grau de instrução</b> (1: Não alfabetizado; 2: Pré-escola; 3: 1º grau incompleto; 4: 1º grau completo; 5: 2º grau incompleto; 6: 2º grau completo; 7: superior incompleto 8: superior completo); <b>Idade</b> (anos); <b>Usa Vale Transporte</b> (1: Sim; 2: Não)                                                |

\* variáveis adaptadas das originais

As variáveis “Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO”, “Nº DE CRIANÇAS NO DOMICÍLIO” e “TIPO DE CHEFE DO DOMICÍLIO” foram adaptadas das originais da seguinte forma:

- (1) **Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO:** É um número estimado. Tal variável foi criada associando a variável “Nº DE AUTOMÓVEIS” à variável “IDADE”. Assim, um indivíduo com idade superior ou igual a 18 anos e com pelo menos um automóvel no domicílio, provavelmente teria carteira nacional de habilitação;
- (2) **Nº DE CRIANÇAS NO DOMICÍLIO:** foram considerados crianças aqueles indivíduos com idade inferior ou igual a 11 anos. Assim, esta variável representa o número de moradores do domicílio nessa faixa etária;
- (3) **TIPO DE CHEFE DO DOMICÍLIO:** Esta variável representa o sexo do chefe da família. Observou-se que em domicílios onde a mulher exerce o papel de chefe, na maioria das vezes não há cônjuge.

Optou-se por incluir esta última variável adaptada na análise dos dados, devido às novas tendências de estrutura familiar na RMSP, principalmente considerando as mudanças que ocorreram entre as décadas de 80 e 90. As citadas novas tendências referem-se às famílias com chefia feminina sem cônjuge, nas quais a participação da chefe entre os ocupados da família aumenta em decorrência da menor absorção dos filhos e parentes jovens no mercado de trabalho. Tais mudanças na estrutura familiar são verificadas pela SEADE (2002), segundo a qual as famílias chefiadas por mulheres na RMSP, que representavam 10,2% dos arranjos familiares em 1988/89, passaram a responder por 14,3% do total de famílias em 2000/01.

Como é de se esperar, algumas das variáveis escolhidas possuem alta correlação entre si (“Nº DE AUTOMÓVEIS NO DOMICÍLIO” e “Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO”, por exemplo, com correlação igual a 0,745). No entanto, ainda que algumas das variáveis socioeconômicas sejam altamente correlacionadas, o que pode criar um possível problema de redundância dos dados, técnicas de AD não são sensíveis a problemas de multicolinearidade. Como mencionado anteriormente, a forma como o algoritmo contido no pacote do S-PLUS 6.1 faz a

segregação dos dados (testando variável por variável, e não uma combinação de variáveis) faz com que problemas desse gênero não sejam obstáculos para o processamento de dados.

#### 5.4 USO DO SOLO

As variáveis de uso do solo fazem parte do conjunto de fatores que afetam decisões individuais de realização de viagens. A complexidade da inter-relação entre atributos de uso do solo e a seqüência de viagens vem sendo estudada ao longo dos anos, havendo divergências quanto às respostas encontradas (WEE, 2002; KITAMURA; MOKHTARIAN; LAIDET, 1997).

A incorporação de tais variáveis ao estudo de viagens encadeadas envolve uma série de dificuldades como, por exemplo, representar ou mensurar tais atributos e, a forma como tais variáveis podem ser incorporadas aos modelos, já que, intuitivamente, espera-se que a estrutura urbana influencie a seqüência de viagens realizadas pelos indivíduos durante o dia.

Ao longo desta pesquisa é realizado um estudo exploratório a respeito da influência do uso do solo no encadeamento de viagens, buscando uma forma de incluir variáveis que caracterizem o uso do solo em modelos que representam a cadeia de viagens realizadas pelos indivíduos. A proposta do grupo de variáveis de uso do solo compreendeu algumas etapas ilustradas na Figura 5.1. Algumas etapas estão diretamente associadas aos objetivos secundários do trabalho (etapas "C" e "E"). A etapa D, que corresponde ao objetivo principal do trabalho será descrita posteriormente (Capítulo 7 e 8), bem como a etapa E (Capítulo 9).

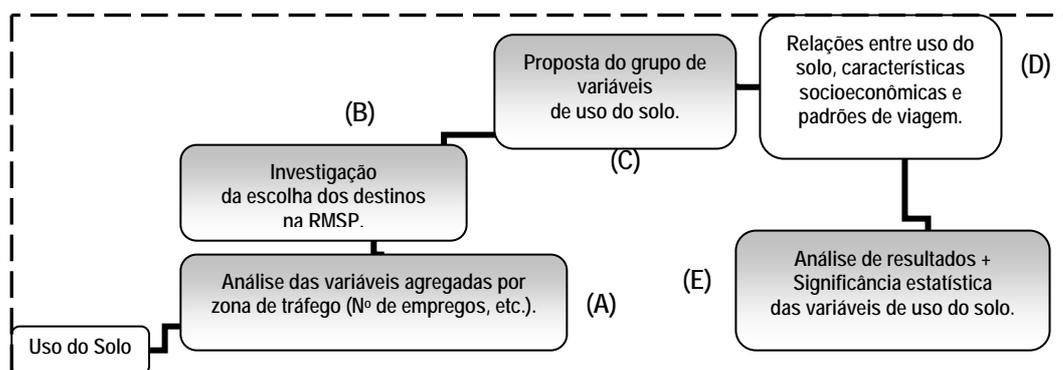


Figura 5.1 Etapas para proposta do grupo de variáveis de uso do solo

#### 5.4.1 Análise das variáveis agregadas por zona de tráfego

Nesta etapa foram analisadas as variáveis agregadas, disponibilizadas pelo METRÔ-SP, descritas no capítulo anterior. Dentre o conjunto das variáveis originais, optou-se por utilizar aquelas relacionadas ao número de empregos ou matrículas escolares, classificados por diferentes setores ou escolaridades, respectivamente (fonte METRÔ-SP/1997).

A utilização de todas variáveis desagregadas aumentaria a complexidade da análise. Além disso, imagina-se, intuitivamente, que a oferta de empregos ou escolas em determinadas localizações pode ser um importante fator de atração de viagens, bem como de caracterização do uso do solo.

Como mencionado anteriormente, foram separadas seis amostras finais de indivíduos (industriários, comerciários, trabalhadores no setor de serviços, estudantes até o 1º grau, estudantes no 2º grau e estudantes universitários). Assim, a partir das variáveis originais “Nº DE EMPREGOS” e “Nº DE MATRÍCULAS” (desagregadas por setor econômico e grau de instrução, respectivamente), foram adaptadas seis variáveis contínuas que descrevem a parcela do total de empregos (nos setores correspondentes) e parcela do total de matrículas (de escolaridades correspondentes) por zona de tráfego. As variáveis utilizadas para posteriores análises foram:

1. Parcela do total de empregos na indústria (Nº de empregos no setor industrial na zona de tráfego/ Nº de empregos no setor industrial na RMSP \* 100) (%);
2. Parcela do total de empregos no comércio (Nº de empregos no setor comercial na zona de tráfego/ Nº de empregos no setor comercial na RMSP \* 100) (%);
3. Parcela do total de empregos no setor de serviços (Nº de empregos no setor de serviços na zona de tráfego/ Nº de empregos no setor de serviços na RMSP \* 100) (%);
4. Parcela do total de matrículas até 1º grau (Nº de matrículas até o 1º grau na zona de tráfego/ Nº de matrículas até o 1º grau na RMSP \* 100) (%);
5. Parcela do total de matrículas no 2º grau (Nº de matrículas no 2º grau na zona de tráfego/ Nº de matrículas no 2º grau na RMSP \* 100) (%);

6. Parcela do total de matrículas em universidades ( $\text{N}^\circ$  de matrículas no 3º grau na zona de tráfego/  $\text{N}^\circ$  de matrículas no 3º grau na RMSP \* 100) (%).

#### 5.4.2 Investigação da escolha dos destinos na RMSP

A escolha dos diferentes destinos provavelmente sofre influência de vários fatores: necessidade de realizar atividades obrigatórias (como trabalho, por exemplo, que geralmente possui localização e horários fixos e envolve decisões tomadas a longo prazo), realização de atividades flexíveis e adequação de horários e localizações, distribuição de diversas atividades em determinadas zonas de tráfego, a inter-dependência entre todas as viagens realizadas em determinado período (considerando as distintas origens e destinos da cadeia de viagens), a localização relativa das zonas, o modo de transporte mais conveniente, as distâncias a serem percorridas e a atratividade ou acessibilidade de determinados locais.

O objetivo desta etapa é investigar a escolha de destinos dos indivíduos residentes na RMSP a fim de propor o grupo de variáveis de uso do solo, baseado em um suporte teórico. Um dos principais desafios, então, seria compreender a escolha dos destinos de diferentes indivíduos, confrontando as distâncias percorridas e as características das origens e destinos relacionadas à oferta de empregos e escolas. Este tipo de análise visa responder a algumas questões, as quais motivaram a realização desta etapa:

- Indivíduos que residem próximos a regiões de alta atividade comercial, industrial, etc. fazem viagens longas ou curtas a fim de realizar diversas atividades? Qual o destino escolhido e as características deste?
- Estudantes geralmente trocam de zona para realizar atividades de estudo? Ou freqüentam escolas próximas às suas residência? E no caso do ensino superior, onde muitas vezes não existe faculdade do seu interesse na proximidade do domicílio?

- Indivíduos que residem em zonas de tráfego de baixa atividade geralmente vão trabalhar em quais regiões? Em zonas próximas ou distantes? Eles tendem a ir a zonas mais distantes, porém com maiores ofertas de empregos ou permanecem nas regiões próximas às suas residências?

Observou-se que na RMSP, considerando todas as seis amostras analisadas, os indivíduos tendem a permanecer nas zonas de tráfego da origem ou realizam viagens curtas de forma geral. Fazem viagens mais longas somente devido à insuficiência de empregos ou escolas nas suas proximidades (principalmente no caso de universitários e industriários).

Ressalta-se que a oferta de empregos e escolas foi representada pelas variáveis agregadas definidas na subseção anterior (Parcela do total de empregos e matrículas por setor ou escolaridade). Desta forma, para análise dos industriários, por exemplo, observou-se a matriz de distribuição de viagens, confrontando as distâncias percorridas com a variável "PARCELA DO TOTAL DE EMPREGOS NA INDÚSTRIA".

Verifica-se que a maior parte das viagens concentra-se numa faixa de 5 km. Quanto mais longas as viagens, maiores as diferenças observadas entre a "PARCELA DO TOTAL DE EMPREGOS NA INDÚSTRIA" na origem e no destino. Para o caso de viagens acima de 15 km, por exemplo, verifica-se uma parcela média do total de empregos na origem com valor igual a 0,28%, enquanto que no destino a mesma parcela assume um valor médio de 0,65%. Assim, uma das possíveis explicações para realização de viagens longas é a ausência de oferta de empregos na origem. Um dos fatores importantes na etapa de análise de distribuição de viagens seria então a ausência ou não de oferta de empregos ou escolas na zona de tráfego de origem, ou nas suas vizinhanças. A Tabela 5.2 sintetiza os valores observados nesta etapa de investigação de deslocamentos.

As distâncias (euclidianas), como mencionado no Capítulo 4, foram geradas com auxílio do TRANSCAD 3.6. Para o caso de viagens intra-zonais, as distâncias adotadas têm o valor igual ao raio aproximado da zona (considerando a área igual a um círculo).

Tabela 5.2 Síntese das informações a respeito da distribuição de viagens (indústria)

| Distâncias        | Nº de pessoas | Parcela em %                                    | Parcela em %                                     |
|-------------------|---------------|-------------------------------------------------|--------------------------------------------------|
|                   |               | do total de empregos na Ind<br>(média) - origem | do total de empregos na Ind<br>(média) - destino |
| Até 5 km          | 1823          | 0,47                                            | 0,56                                             |
| 5 a 10 km         | 924           | 0,32                                            | 0,69                                             |
| 10 a 15 km        | 552           | 0,35                                            | 0,70                                             |
| acima de 15 km    | 803           | 0,28                                            | 0,65                                             |
| Amostra indústria | 4102          |                                                 |                                                  |

Assim, admite-se no presente trabalho que os indivíduos seguem o princípio que fundamenta o modelo de oportunidades intervenientes, proposto por Schneider (1959), que admite que numa área urbana todas as viagens sejam tão curtas quanto possível, sendo apenas longas o necessário para atingirem o destino aceitável mais próximo, onde o propósito do viajante é satisfeito. Então, neste trabalho, o grupo de variáveis que caracterizam o uso do solo foi representado em termos de grau de “oportunidade acumulada” por faixa de distância.

Vale ressaltar que o modelo de oportunidades intervenientes pode explicar parcialmente a distribuição de viagens das amostras que compõem indivíduos residentes na RMSP. Devem-se considerar também outras medidas de impedância, como distância, por exemplo, de forma direta, e não implicitamente como no modelo de oportunidades intervenientes. Certamente há outros fatores relacionados à investigação da distribuição de viagens, como a competitividade entre zonas, por exemplo. Não é o escopo desta pesquisa debater a respeito da investigação de destinos. Esta foi apenas uma etapa, que forneceu fundamentação teórica para proposta das variáveis de uso do solo.

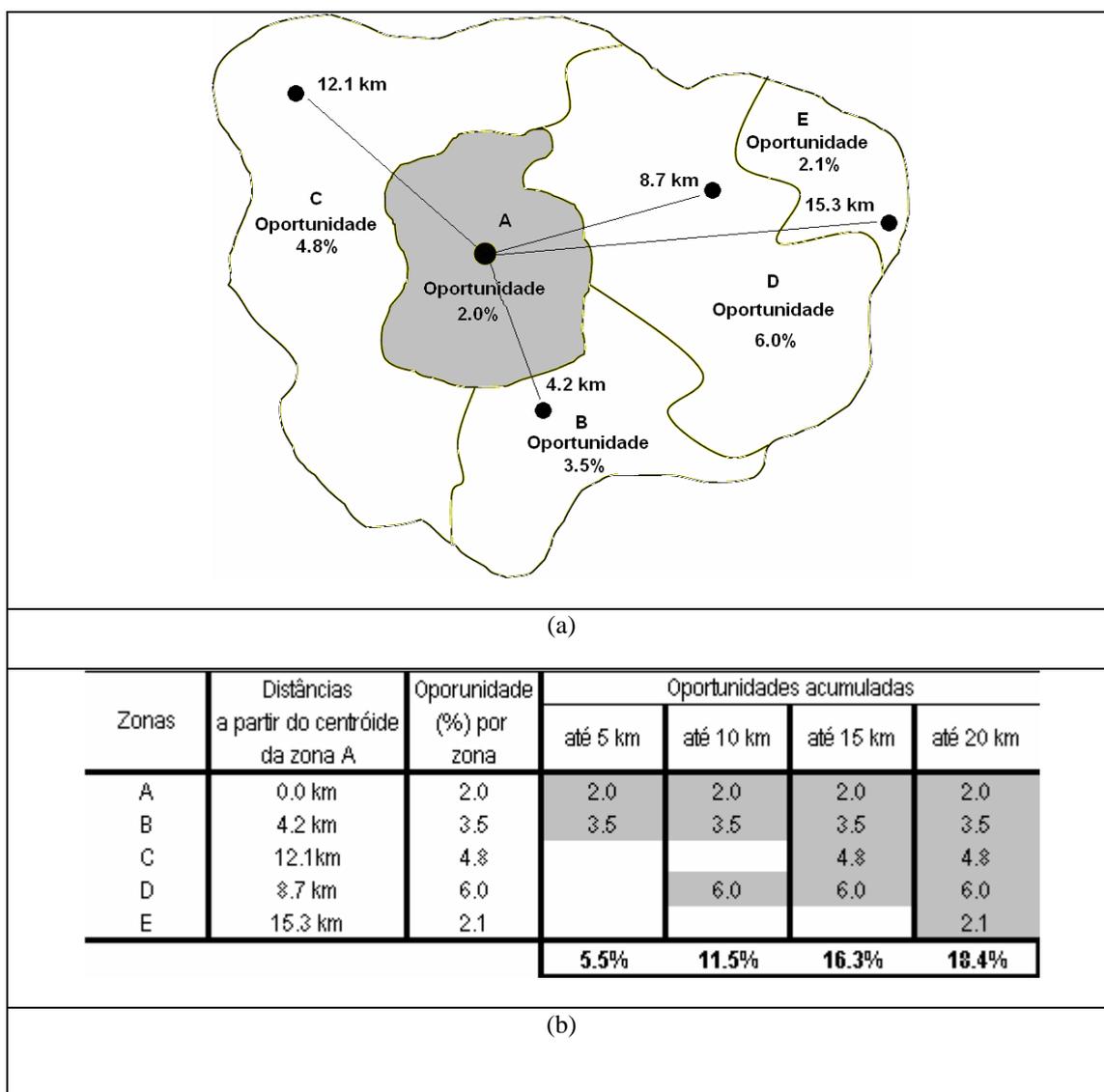
#### 5.4.3 Proposta do grupo de variáveis do uso do solo

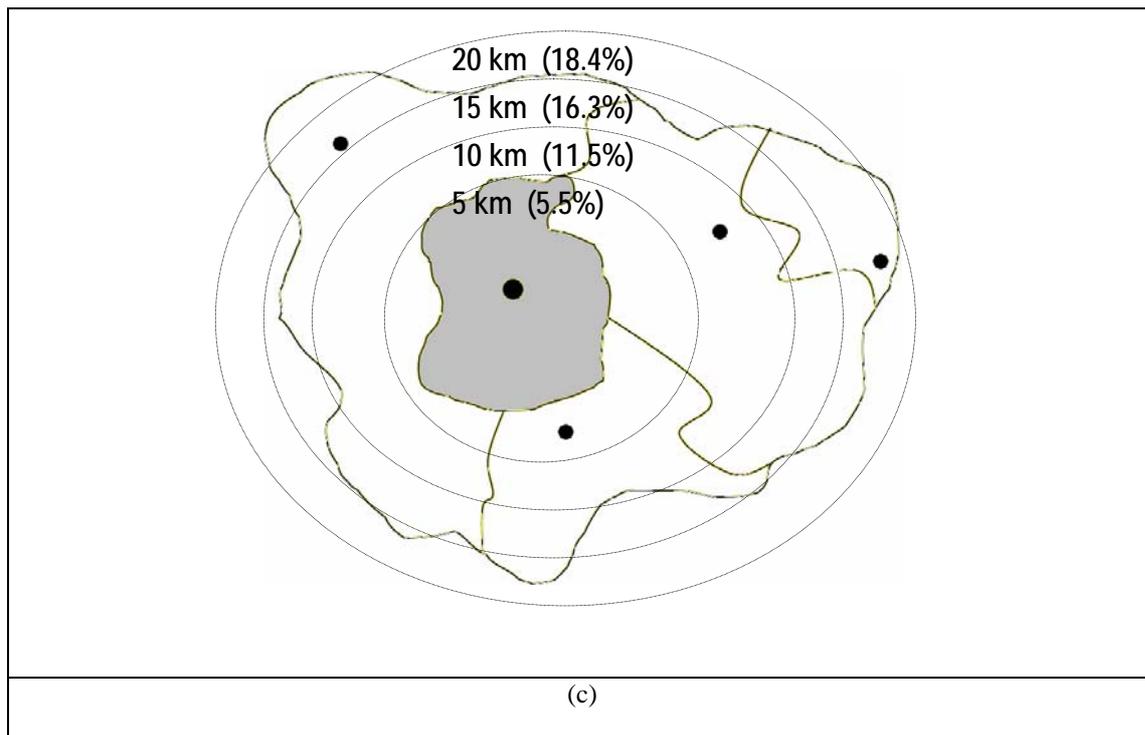
Para elaboração das variáveis de uso do solo, partiu-se dos princípios do modelo de oportunidades intervenientes. Como mencionado anteriormente, as variáveis de uso do solo foram representadas em termos de grau de “oportunidade acumulada” por faixa de distância.

Qual o significado do termo grau de “oportunidade acumulada”? No presente trabalho, “oportunidade” refere-se à oferta de empregos ou matrículas, representadas através da variável agregada descrita na subseção 5.4.1: (1) “PARCELA (EM %) DO TOTAL DE EMPREGOS” para cada

um dos três setores econômicos; e (2) "PARCELA (EM %) DO TOTAL DE MATRÍCULAS" (por grau de instrução).

Assim, acumulou-se tal variável agregada em faixas de distância (raios) a partir do centróide da zona de origem. Foram adotados valores de raios: 5 km, 10 km, 15 km e 20 km, gerando desta forma o termo "oportunidade acumulada". A Figura 5.2 (a, b e c) exemplifica a proposta das variáveis de uso do solo. A zona de origem é a central "A" (sombreada). Na etapa (a) estão representados os valores das "oportunidades" em (%) em cada zona, assim como distâncias a partir do centróide de "A". Na próxima fase (b) são mostrados os cálculos das "oportunidades acumuladas" para cada uma das quatro faixas de distância. Finalmente, na etapa (c) é ilustrada a variável de uso do solo por ora proposta, a partir da zona de tráfego A.





Desta forma, o conjunto de variáveis que representam características de uso do solo é formado por quatro variáveis. Cada variável está associada a uma das faixas de distância a partir do centróide da zona de origem.

No exemplo em questão, o termo “oportunidade acumulada” poderia ser substituído por “PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA”, “PARCELA ACUMULADA DE EMPREGOS NO COMÉRCIO”, “PARCELA ACUMULADA DE MATRÍCULAS ATÉ O 1º GRAU” e assim sucessivamente.

Para o caso específico da amostra de industriários, por exemplo, as variáveis de uso do solo são dadas por:

- (1) Parcela acumulada de empregos na indústria até um raio de 5 km, a partir do centróide da zona de residência (%);
- (2) Parcela acumulada de empregos na indústria até um raio de 10 km, a partir do centróide da zona de residência (%);
- (3) Parcela acumulada de empregos na indústria até um raio de 15 km, a partir do centróide da zona de residência (%);

- (4) Parcela acumulada de empregos na indústria até um raio de 20 km, a partir do centróide da zona de residência (%);

Para os casos de raios acima de 20 km, a parcela acumulada corresponde a 100%.

Verifica-se que as variáveis são complementares (por serem acumuladas) e também são altamente correlacionadas entre si. No entanto, como dito anteriormente, este não seria um problema potencial para aplicação da técnica de AM "emergente" utilizada no trabalho.

Considerando o estudo de caso (RMSP), a Figura 5.3 ilustra a variável de uso do solo para os trabalhadores industriais que residem na zona Sé (1). A partir da referida zona, são calculadas as parcelas acumuladas (em porcentagem).

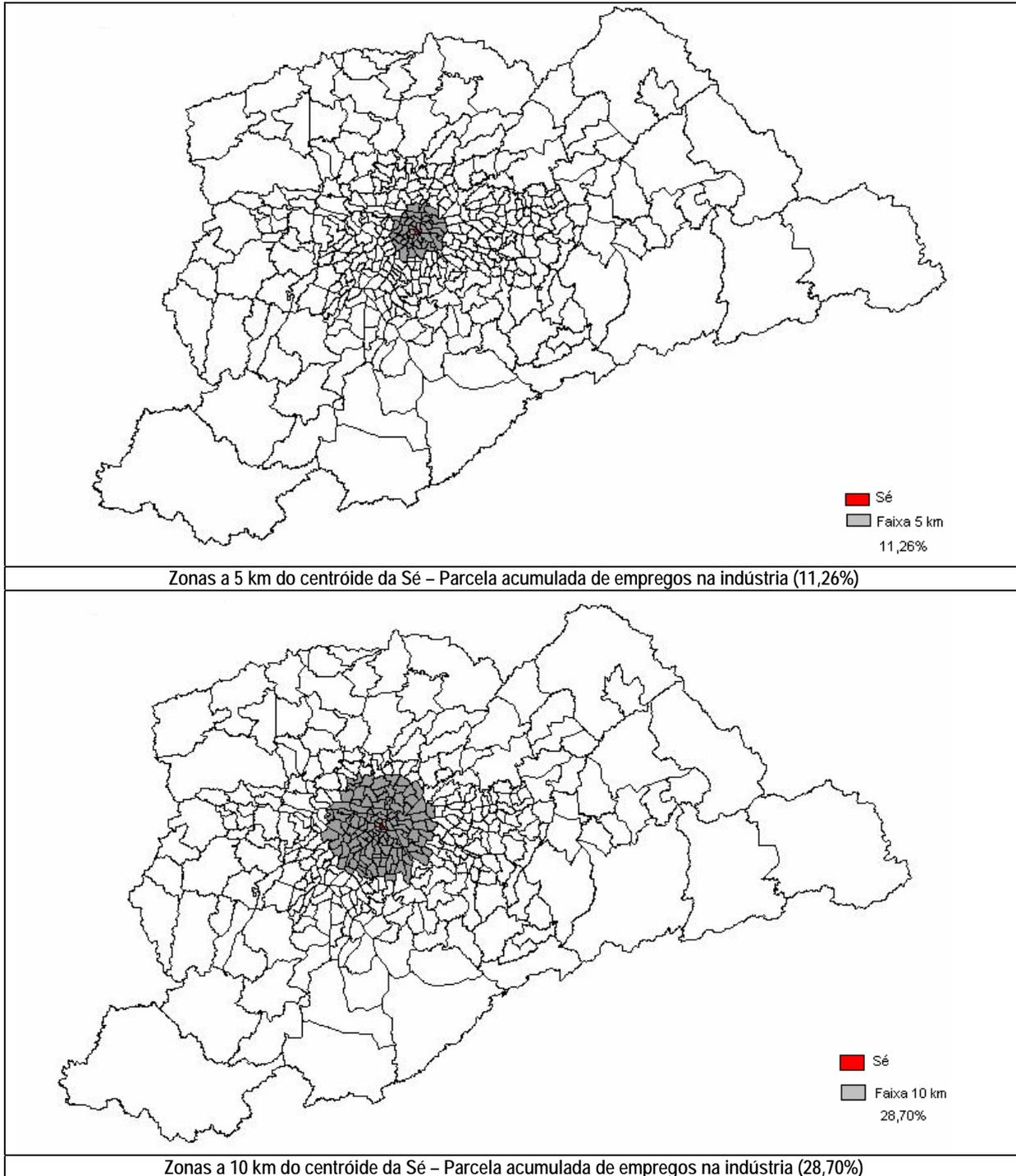


Figura 5.3 Variáveis de uso do solo (indústria) Zonas localizadas respectivamente até 5 e 10 km do centróide da zona Sé.

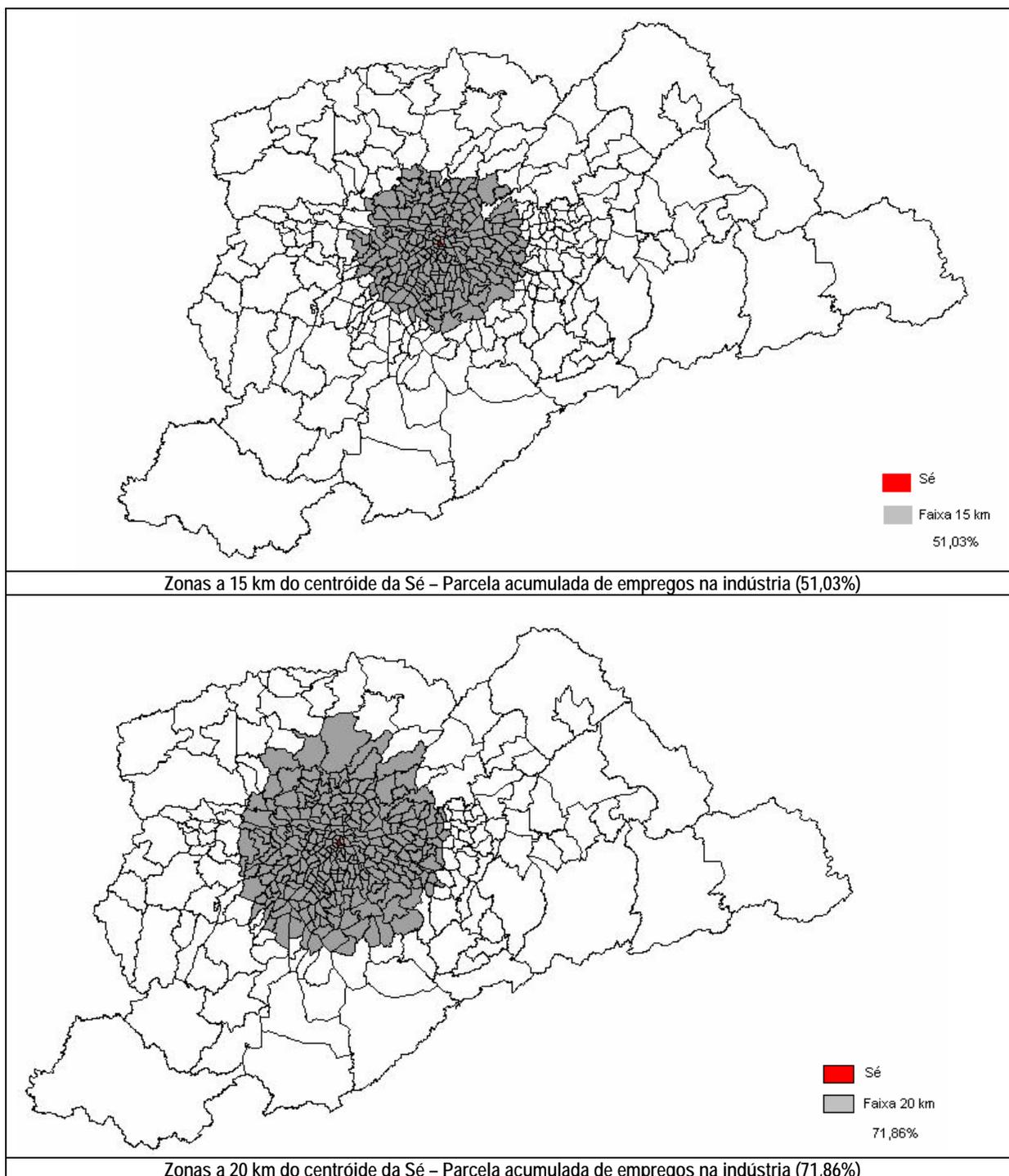


Figura 5.3 (cont.) Variáveis de uso do solo (indústria) Zonas localizadas respectivamente até 15 e 20 km do centróide da zona Sé.

O grupo de variáveis de uso do solo representa o grau de “atratividade” da zona de origem e “atratividade/oportunidade” acumulada dentro de círculos concêntricos. Pode-se interpretar a influência das variáveis propostas da seguinte forma:

(a) industriários que moram em zona de tráfego com alta parcela acumulada de indústria, até um raio de 5 km, por exemplo, provavelmente tendem a realizar viagens curtas ao trabalho;

Contudo;

(b) aqueles indivíduos que residem em zonas de tráfego com baixa parcela acumulada de indústria em raios menores e alta parcela acumulada em raios maiores, possivelmente tendem a realizar viagens mais longas com motivo TRABALHO.

A Figura 5.4 ilustra respectivamente os dois casos citados acima, considerando o caso de trabalhadores na indústria. Primeiro, verifica-se as variáveis de uso do solo a partir da zona 3 (Praça João Mendes), as quais apresentam alta parcela acumulada nas proximidades (11,54% até um raio de 5 km), crescendo de maneira homogênea. Em seguida, considera-se a zona 268 (Tamboré), que apresenta parcela acumulada relativamente baixa até o raio de 15 km (17 %), aumentando abruptamente a partir deste ponto.

Considerando a frequência de viagens, verifica-se que, no caso de residentes na zona de tráfego 3, a maioria das viagens realizadas tem distâncias até uma faixa de 5 km (59%). Já na zona 268, esse percentual cai para aproximadamente 33%.

Entretanto, observando as viagens longas (superiores a 15 km de distância), pode-se afirmar que:

- (i) Aproximadamente 9% dos industriários residentes na zona 3 realizam viagens iguais ou superiores a 15 km de distância, enquanto;
- (ii) Aproximadamente 22% dos industriários residentes na zona 268 realizam viagens com motivo TRABALHO, com distâncias iguais ou superiores 15 km.

Assim, o grupo de variáveis de uso do solo proposto pode explicar comportamento relacionado à escolha de possíveis destinos, principalmente considerando fatores como: (1) necessidade de realizar atividades de subsistência (trabalho, estudo); e (2) oferta ou não de empregos ou escolas nas proximidades da zona de origem.

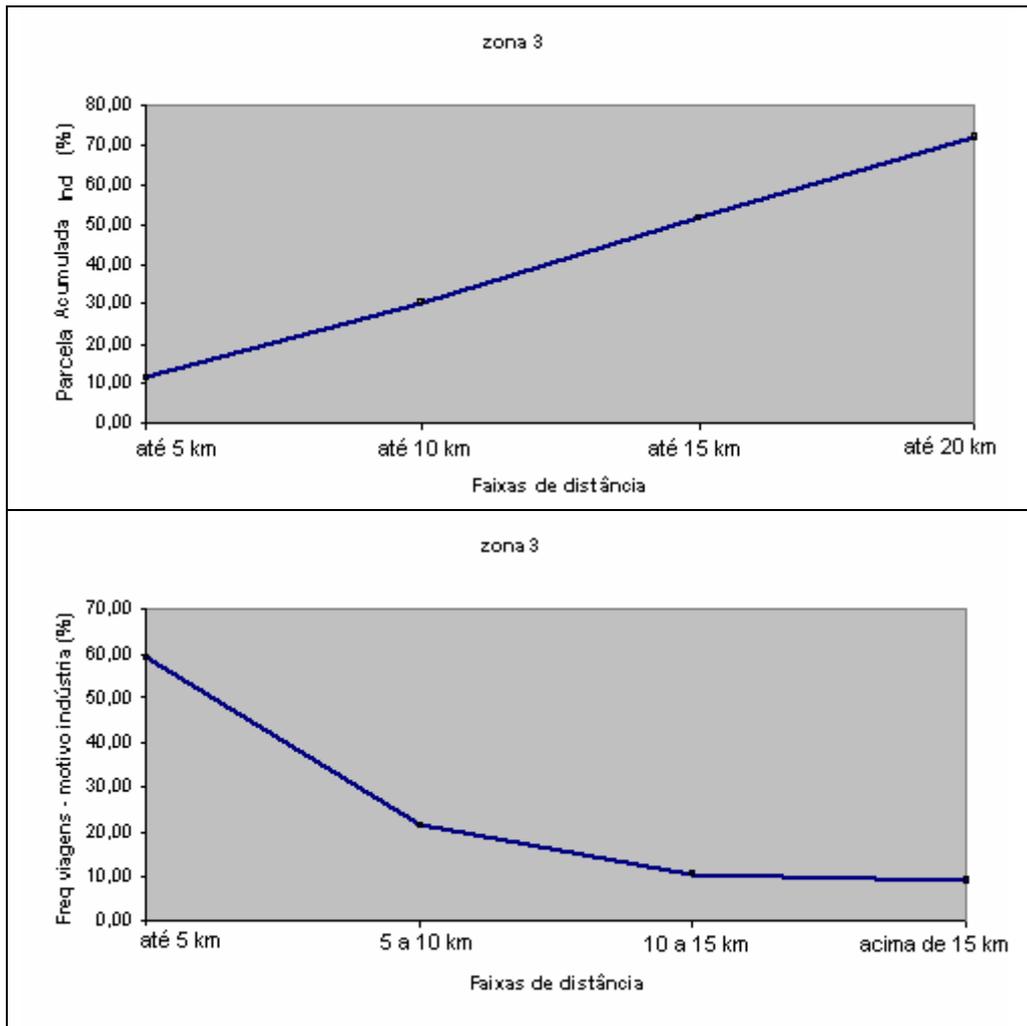


Figura 5.4 Zona 3 – Parcela acumulada de empregos na indústria e frequência de viagens (motivo trabalho na indústria).

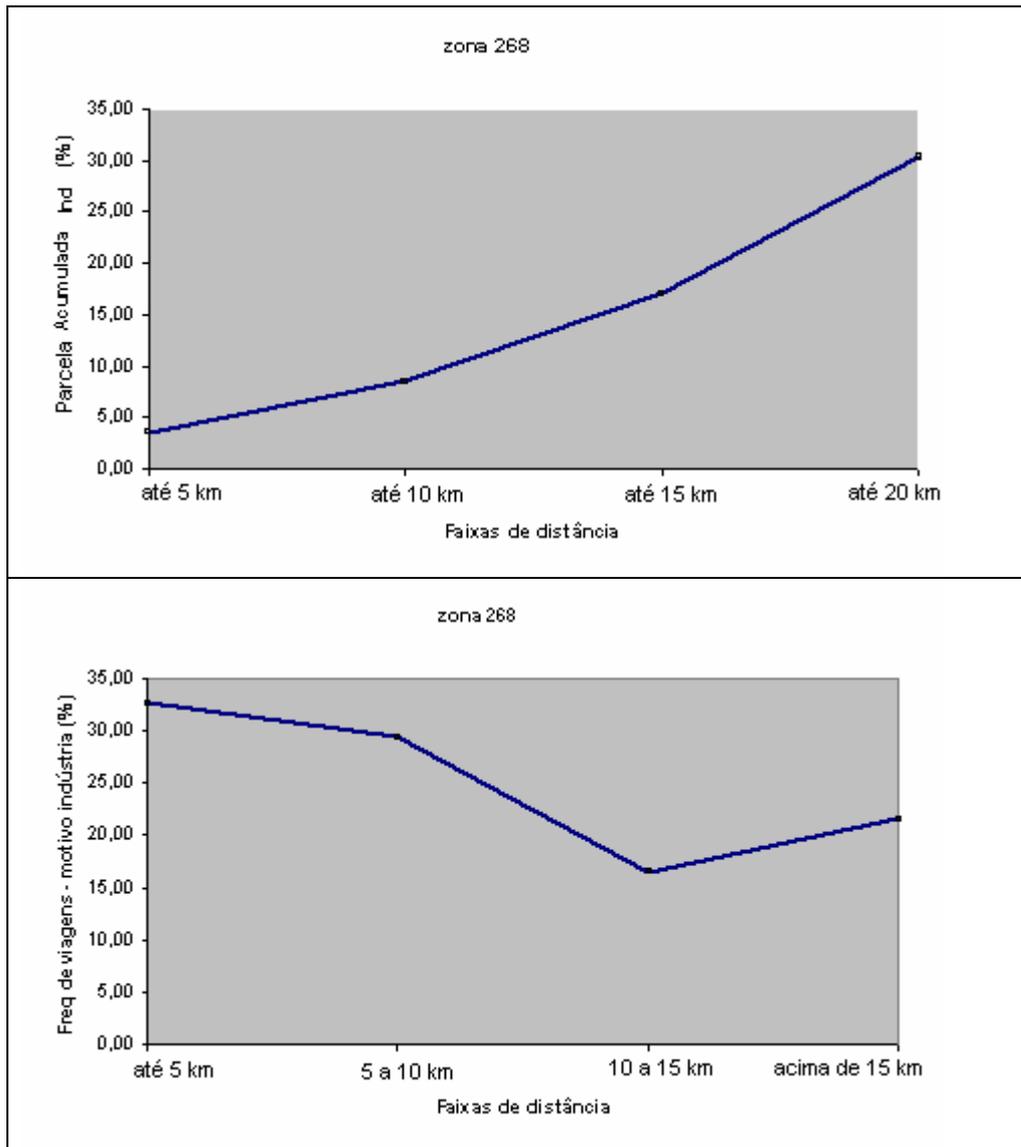


Figura 5.4 (cont.) Zona 268 – Parcela acumulada de empregos na indústria e freqüência de viagens (motivo trabalho na indústria).

## 5.5 CONSIDERAÇÕES FINAIS

A finalidade principal do trabalho, que é encontrar relações entre variáveis independentes (participação em atividades, características socioeconômicas, atributos de uso do solo) e variável dependente (padrões de viagem) depende fundamentalmente da representação adequada do conjunto de variáveis envolvidas.

Este capítulo mostrou o processo de representação das variáveis independentes utilizadas na análise. Para o caso de participação em atividades e características socioeconômicas seguiu-se a literatura vigente, adaptando as variáveis originais, disponibilizadas pelo METRÔ-SP.

Já a representação do conjunto de variáveis de uso do solo englobou algumas etapas, com o intuito de representar adequadamente fatores da oferta de empregos e escolas, associados a medidas de distância.

A proposta do grupo de variáveis de uso do solo, objetivo secundário da presente pesquisa, foi apoiada em um suporte teórico. As referidas variáveis foram criadas baseadas nos princípios do modelo de oportunidades intervenientes. A Tabela 5.3 é um quadro resumo das variáveis independentes utilizadas nas análises para as seis amostras de indivíduos consideradas.

**Tabela 5.3 Variáveis independentes envolvidas na análise**

| Amostras     | Atividades               |                        | Socioeconômicas                                                                                                                                                                                                                                                                                                                                                                  | Uso do solo                                                                                                                                                          |  |  |  |
|--------------|--------------------------|------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|--|--|
| indústria    | TRABALHA<br>(categórica) | ESTUDA<br>(categórica) | Renda Individual; Renda Familiar; Usa Vale Transporte<br>Nº de automóveis no domicílio (numéricas);<br>Nº provável de carteiras de habilitação no domicílio(numérica);<br>Sexo (categórica); Posição do indivíduo na família (categórica);<br>Total de pessoas na família (numérica); Grau de instrução (categórica);<br>Nº de crianças no domicílio(numérica); Idade (numérica) | parcela acumulada de empregos na indústria até um raio de<br>5km      10km      15km      20km<br>(numérica)      (numérica)      (numérica)      (numérica)         |  |  |  |
| comércio     | idem ao anterior         | idem ao anterior       | idem ao anterior                                                                                                                                                                                                                                                                                                                                                                 | parcela acumulada de empregos no comércio até um raio de<br>5km      10km      15km      20km<br>(numérica)      (numérica)      (numérica)      (numérica)          |  |  |  |
| serviços     | idem ao anterior         | idem ao anterior       | idem ao anterior                                                                                                                                                                                                                                                                                                                                                                 | parcela acumulada de empregos no setor de serviços até um raio de<br>5km      10km      15km      20km<br>(numérica)      (numérica)      (numérica)      (numérica) |  |  |  |
| até 1º grau  | idem ao anterior         | idem ao anterior       | idem ao anterior                                                                                                                                                                                                                                                                                                                                                                 | parcela acumulada de matrículas até o 1º grau até um raio de<br>5km      10km      15km      20km<br>(numérica)      (numérica)      (numérica)      (numérica)      |  |  |  |
| no 2º grau   | idem ao anterior         | idem ao anterior       | idem ao anterior                                                                                                                                                                                                                                                                                                                                                                 | parcela acumulada de matrículas no 2º grau até um raio de<br>5km      10km      15km      20km<br>(numérica)      (numérica)      (numérica)      (numérica)         |  |  |  |
| universidade | idem ao anterior         | idem ao anterior       | idem ao anterior                                                                                                                                                                                                                                                                                                                                                                 | parcela acumulada de matrículas em universidades até um raio de<br>5km      10km      15km      20km<br>(numérica)      (numérica)      (numérica)      (numérica)   |  |  |  |

*Representadas as variáveis independentes no capítulo anterior, a próxima etapa do método consiste na representação da variável dependente. Este capítulo apresenta as etapas de codificação da variável dependente, que são os padrões de encadeamento de viagens realizados pelos indivíduos. Ao final da codificação, é obtida a variável dependente que possui 128 categorias para cada uma das seis amostras analisadas.*

### 6.1. CONSIDERAÇÕES INICIAIS

A representação adequada das variáveis envolvidas na análise é fundamental para obtenção de relações entre as mesmas. No capítulo anterior, foram representadas as variáveis independentes que fazem parte de três grupos de variáveis (participação em atividades, características socioeconômicas e uso do solo). Na presente etapa do método, deve-se representar a variável dependente, que é definida como os padrões de encadeamento de viagens realizados pelos indivíduos.

Assim, a variável dependente é formada por categorias que representam os diferentes padrões das cadeias de viagens. O termo “cadeia de viagens” ou “encadeamento de viagens” é definido como a seqüência das viagens realizadas pelos indivíduos durante o período de um dia. Neste trabalho o número de viagens varia de dois a quatro (conforme mencionado na etapa de tratamento dos dados – Capítulo 4). Foi decidido que iriam ser representados três atributos relacionados às viagens realizadas: (1) seqüência de motivos de viagem; (2) seqüência de modos de transporte; e (3) seqüência dos destinos das viagens realizadas.

Os padrões de encadeamento das viagens urbanas foram representados em termos de variáveis dependentes em diversos trabalhos anteriores, realizados no grupo de pesquisa do Departamento de Transporte da EESC-USP, variando apenas os atributos de viagens envolvidos em suas análises

(ICHIKAWA, 2002; PITOMBO, 2003, SOUZA 2004; AGUIAR, 2005; SILVA, 2006). Para que os padrões de viagens representassem os atributos propostos neste trabalho – motivo de viagem, modo de transportes e destinos – foram seguidas quatro etapas básicas para caracterização dos padrões, como mostrado no esquema (Figura 6.1). As etapas serão descritas nas próximas seções.

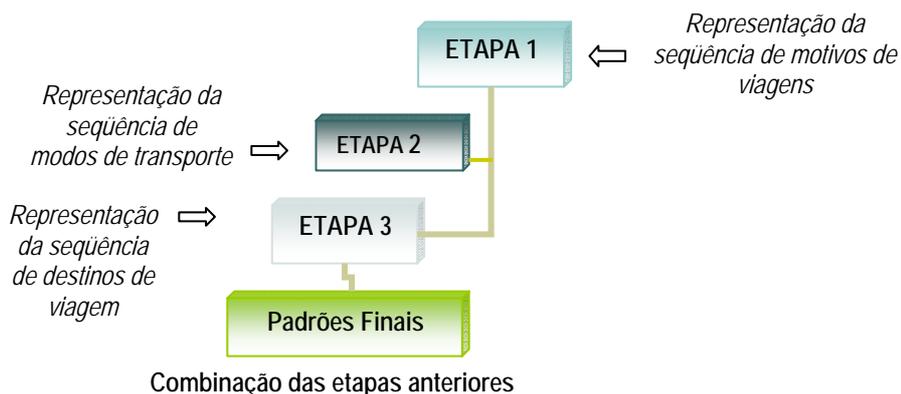


Figura 6.1 Etapas para codificação dos padrões de encadeamento de viagens

Um ponto relevante para representação dos padrões de cadeia de viagens é a limitação do *software* utilizado para aplicação da AD. Há uma limitação do S-PLUS 6.1 de 128 categorias de variável dependente. Por esta razão foram agrupados motivos de viagem, modos de transporte e também destinos (em termos de distâncias de viagem), a fim de limitar o número de combinações possíveis de atributos de viagem. Os padrões finais eram aqueles 128 mais frequentes para cada uma das seis amostradas analisadas.

## 6.2 SEQÜÊNCIA DE MOTIVOS DE VIAGEM (ETAPA 1)

Na etapa 1 foi representada a seqüência dos motivos das viagens (ou seqüência de atividades) realizadas pelos indivíduos durante o período de um dia. Os motivos das viagens contidos nos dados originais, disponibilizados pelo METRÔ-SP, foram agrupados em quatro categorias e representados através de letras. As letras indicam as atividades desempenhadas pelos indivíduos no dia anterior à entrevista e a seqüência das letras significa a ordem cronológica em que as atividades foram desenvolvidas.

As letras estão associadas às seguintes atividades (quatro classes principais de atividades ou motivos de viagem no total):

- (1) H (*Home*) – motivo de viagem RESIDÊNCIA;
- (2) W (*Work*) – motivo de viagem TRABALHO;
- (3) S (*School*) – motivo de viagem ESCOLA;
- (4) A (*Activity*) – motivo de viagem ATIVIDADE (outra atividade qualquer).

O agrupamento dos motivos das viagens nas quatro classes principais, para cada uma das seis amostras, é mostrado na Tabela 6.1. Vale ressaltar que nas três amostras correspondentes a trabalhadores nos três setores econômicos, os indivíduos podem também realizar atividades de estudos. No entanto, considerando as amostras de estudantes, independente do grau de instrução, os indivíduos não realizam atividades de trabalho. As três últimas amostras representam indivíduos que pertencem à categoria “ESTUDANTE” considerando a variável original “CONDIÇÃO DE ATIVIDADE”. Assim, a sua única atividade de subsistência seria estudo.

Tabela 6.1 Agrupamento dos motivos das viagens para as seis amostras analisadas

| Indústria                     |                       | Comércio                      |                       | Serviços                      |                       |
|-------------------------------|-----------------------|-------------------------------|-----------------------|-------------------------------|-----------------------|
| Motivos de viagem (originais) | Agrupados             | Motivos de viagem (originais) | Agrupados             | Motivos de viagem (originais) | Agrupados             |
| 1 - Trabalho/Indústria        | <b>(W)</b> Trabalho   | 2 - Trabalho/Comércio         | <b>(W)</b> Trabalho   | 3 - Trabalho/Serviços         | <b>(W)</b> Trabalho   |
| 4 - Escola/Educação           | <b>(S)</b> Escola     | 4 - Escola/Educação           | <b>(S)</b> Escola     | 4 - Escola/Educação           | <b>(S)</b> Escola     |
| 5 - Compras                   | <b>(A)</b> Atividade  | 5 - Compras                   | <b>(A)</b> Atividade  | 5 - Compras                   | <b>(A)</b> Atividade  |
| 6 - Médico/Dentista/Saúde     | <b>(A)</b> Atividade  | 6 - Médico/Dentista/Saúde     | <b>(A)</b> Atividade  | 6 - Médico/Dentista/Saúde     | <b>(A)</b> Atividade  |
| 7 - Recreação/Visitas         | <b>(A)</b> Atividade  | 7 - Recreação/Visitas         | <b>(A)</b> Atividade  | 7 - Recreação/Visitas         | <b>(A)</b> Atividade  |
| § - Residência                | <b>(H)</b> Residência | § - Residência                | <b>(H)</b> Residência | § - Residência                | <b>(H)</b> Residência |
| 9 - Outros                    | <b>(A)</b> Atividade  | 9 - Outros                    | <b>(A)</b> Atividade  | 9 - Outros                    | <b>(A)</b> Atividade  |
| Estudantes até 1º grau        |                       | Estudantes no 2º grau         |                       | Estudantes universitários     |                       |
| 4 - Escola/Educação           | <b>(S)</b> Escola     | 4 - Escola/Educação           | <b>(S)</b> Escola     | 4 - Escola/Educação           | <b>(S)</b> Escola     |
| 5 - Compras                   | <b>(A)</b> Atividade  | 5 - Compras                   | <b>(A)</b> Atividade  | 5 - Compras                   | <b>(A)</b> Atividade  |
| 6 - Médico/Dentista/Saúde     | <b>(A)</b> Atividade  | 6 - Médico/Dentista/Saúde     | <b>(A)</b> Atividade  | 6 - Médico/Dentista/Saúde     | <b>(A)</b> Atividade  |
| 7 - Recreação/Visitas         | <b>(A)</b> Atividade  | 7 - Recreação/Visitas         | <b>(A)</b> Atividade  | 7 - Recreação/Visitas         | <b>(A)</b> Atividade  |
| § - Residência                | <b>(H)</b> Residência | § - Residência                | <b>(H)</b> Residência | § - Residência                | <b>(H)</b> Residência |
| 9 - Outros                    | <b>(A)</b> Atividade  | 9 - Outros                    | <b>(A)</b> Atividade  | 9 - Outros                    | <b>(A)</b> Atividade  |

Assim, a seqüência das letras (H, W, S, A) indica a seqüência de atividades realizadas no período de um dia. A primeira e última letra representa sempre os motivos da origem e da última viagem, que, conforme comentado anteriormente (Capítulo 4) deve ser a Residência (H), ou seja, foram

considerados apenas aqueles indivíduos que inicialmente saíram do domicílio e retornaram ao domicílio ao final das viagens.

HSH, HWH, HAH são exemplos de seqüências de motivos de viagens indicando que indivíduos realizaram duas viagens, com origem em sua residência, motivo da primeira viagem respectivamente “ESCOLA”, “TRABALHO” e “ATIVIDADE”, e a segunda viagem de retorno ao domicílio. HSAH já representa aqueles indivíduos que realizaram três viagens, com origem residência, motivo da primeira viagem “ESCOLA”, motivo da segunda viagem “ATIVIDADE”, e, finalmente, a última viagem de retorno ao domicílio. Esta primeira etapa, de representação cronológica da seqüência de atividades foi baseada naquela realizada por Ichikawa (2002).

A Figura 6.2 ilustra as seqüências de atividades (motivos de viagens) para duas, três e quatro viagens respectivamente, na amostra que contém trabalhadores no setor de serviços. Os algarismos romanos representam a ordem de realização das viagens. A Tabela 6.2, a seguir, mostra as seqüências de atividades mais freqüentes para cada uma das amostras de trabalhadores e estudantes. Nota-se que a amostra formada por industriários realiza mais freqüentemente viagens com motivo TRABALHO, quando comparada à amostra de comerciários (Seqüência “HWH” – 79,89% e 67,75%, respectivamente). Este fato justifica-se principalmente pela maior presença de trabalhadores autônomos no caso do comércio. Espera-se que profissionais liberais, não necessariamente, realizem viagens diárias ao trabalho. Já no caso de industriários, onde encontra-se aproximadamente 95% de “Assalariados”, é esperado que haja uma maior necessidade de realização de seqüências de atividades relacionadas ao trabalho.

Tabela 6.2 Seqüências de atividades mais freqüentes

| Amostras                                 |                       | seqüências de atividades (motivo de viagem mais frequentes) |              |              |              |              |             |              |              |              |              |
|------------------------------------------|-----------------------|-------------------------------------------------------------|--------------|--------------|--------------|--------------|-------------|--------------|--------------|--------------|--------------|
| <b>Trabalhadores por setor econômico</b> | Indústria<br>4.102    | <b>HWH</b>                                                  | <b>HWHWH</b> | <b>HWHSH</b> | <b>HAH</b>   | <b>HWHAH</b> | <b>HWSH</b> | <b>HWAWH</b> | <b>HWAH</b>  | <b>HSH</b>   | <b>HSWH</b>  |
|                                          |                       | 79,89%                                                      | 5,29%        | 4,22%        | 3,12%        | 1,76%        | 1,46%       | 1,37%        | 1,17%        | 0,68%        | 0,41%        |
|                                          | Comércio<br>6.043     | <b>HWH</b>                                                  | <b>HAH</b>   | <b>HWHWH</b> | <b>HWSH</b>  | <b>HWHSH</b> | <b>HSH</b>  | <b>HWAWH</b> | <b>HWAH</b>  | <b>HWWH</b>  | <b>HWHAH</b> |
|                                          | 67,75%                | 7,78%                                                       | 5,49%        | 5,33%        | 3,16%        | 2,27%        | 1,87%       | 1,47%        | 1,19%        | 0,88%        |              |
|                                          | Serviços<br>8.328     | <b>HWH</b>                                                  | <b>HAH</b>   | <b>HWHWH</b> | <b>HSH</b>   | <b>HWHSH</b> | <b>HWSH</b> | <b>HWAWH</b> | <b>HWWH</b>  | <b>HWHAH</b> | <b>HWAH</b>  |
|                                          | 72,51%                | 7,30%                                                       | 3,94%        | 2,76%        | 2,56%        | 2,31%        | 2,15%       | 1,17%        | 1,12%        | 0,82%        |              |
| <b>Estudantes por Grau de instrução</b>  | Até 1º grau<br>13.839 | <b>HSH</b>                                                  | <b>HAH</b>   | <b>HSHAH</b> | <b>HSHSH</b> | <b>HSAH</b>  | <b>HSSH</b> | <b>HAHSH</b> | <b>HASH</b>  | <b>HAAH</b>  | <b>HASAH</b> |
|                                          |                       | 89,18%                                                      | 3,27%        | 2,18%        | 1,86%        | 1,16%        | 0,64%       | 0,61%        | 0,20%        | 0,16%        | 0,12%        |
|                                          | No 2º grau<br>2.143   | <b>HSH</b>                                                  | <b>HAH</b>   | <b>HSHSH</b> | <b>HSHAH</b> | <b>HAHSH</b> | <b>HSAH</b> | <b>HSSH</b>  | <b>HAAH</b>  | <b>HASH</b>  | <b>HAHAH</b> |
|                                          | 83,48%                | 4,53%                                                       | 4,25%        | 3,08%        | 1,49%        | 1,40%        | 0,65%       | 0,42%        | 0,28%        | 0,14%        |              |
|                                          | Universidade<br>568   | <b>HSH</b>                                                  | <b>HAH</b>   | <b>HSHAH</b> | <b>HSHSH</b> | <b>HSSH</b>  | <b>HSAH</b> | <b>HAAH</b>  | <b>HAHSH</b> | <b>HAHAH</b> | <b>HASH</b>  |
|                                          | 81,87%                | 4,75%                                                       | 4,23%        | 2,29%        | 2,11%        | 1,59%        | 1,23%       | 1,23%        | 0,35%        | 0,18%        |              |

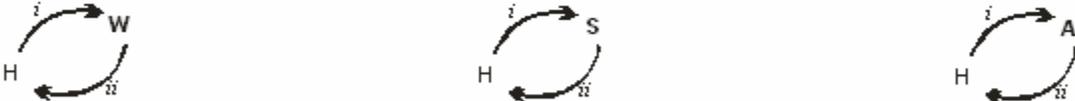
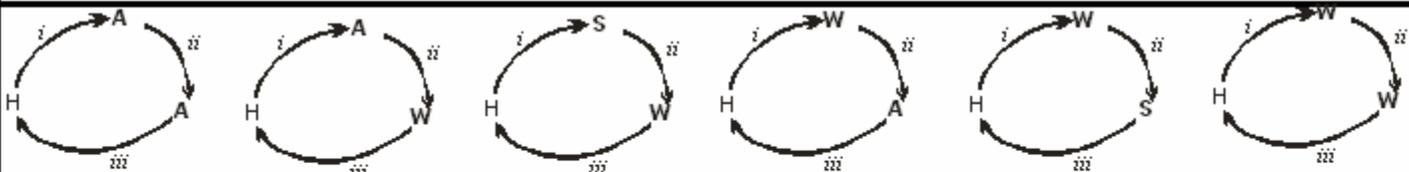
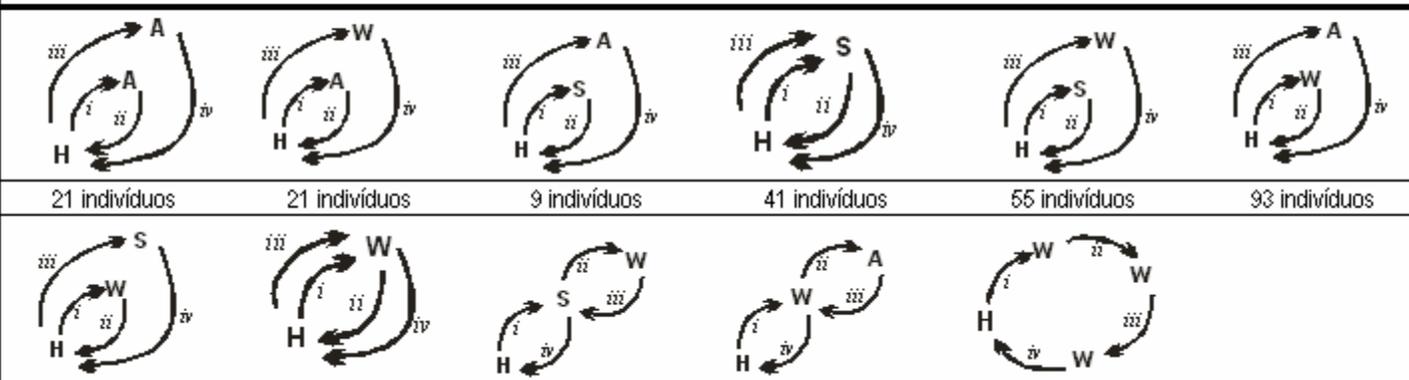
| Número de viagens      | Seqüências de motivos de viagem (atividades)                                        |                |                |                |                |               |
|------------------------|-------------------------------------------------------------------------------------|----------------|----------------|----------------|----------------|---------------|
| 2                      |   |                |                |                |                |               |
| 6.877 indivíduos       | 6.039 indivíduos                                                                    |                | 230 indivíduos |                | 608 indivíduos |               |
| 3                      |   |                |                |                |                |               |
| 452 indivíduos         | 40 indivíduos                                                                       | 7 indivíduos   | 47 indivíduos  | 68 indivíduos  | 192 indivíduos | 98 indivíduos |
| 4                      |  |                |                |                |                |               |
| 999 indivíduos         | 213 indivíduos                                                                      | 328 indivíduos | 19 indivíduos  | 179 indivíduos | 20 indivíduos  |               |
| Total 8.328 indivíduos |                                                                                     |                |                |                |                |               |

Figura 6.2 Seqüências de atividades para amostras de trabalhadores no setor de serviços

### 6.3 SEQUÊNCIA DE MODOS DE TRANSPORTE (ETAPA 2)

Analogamente à etapa anterior, as seqüências de modos de transporte utilizados na cadeia de viagens realizadas pelos indivíduos no período de um dia foram representadas através de uma seqüência de letras (P, T, N). Os modos de transportes principais, utilizados em cada uma das viagens, foram agrupados em três classes principais, representadas por tais letras:

- (1) P (*Private*) – Modo de transporte particular motorizado;
- (2) T (*Transit*) – Modo de transporte público;
- (3) N (*non motorized*) – Modo de transporte não motorizado.

O agrupamento dos modos de viagem (variável original codificada pelo METRÔ-SP) nas três classes principais para todas as amostras é mostrado na Tabela 6.3.

Tabela 6.3 Agrupamento dos modos de viagem para as seis amostras analisadas

| Todas as amostras            |                                  |
|------------------------------|----------------------------------|
| Modo de viagem (METRÔ-SP)    | Codificação do trabalho          |
| 01 - Ônibus                  | <b>(T)</b> Transporte Público    |
| 02 - Ônibus Fretado          | <b>(T)</b> Transporte Público    |
| 03 - Transporte Escolar      | <b>(T)</b> Transporte Público    |
| 04 - Dirigindo Automóvel     | <b>(P)</b> Particular motorizado |
| 05 - Passageiro de Automóvel | <b>(P)</b> Particular motorizado |
| 06 - Táxi                    | <b>(P)</b> Particular motorizado |
| 07 - Lotação/Perua           | <b>(T)</b> Transporte Público    |
| 08 - Metrô                   | <b>(T)</b> Transporte Público    |
| 09 - Trem                    | <b>(T)</b> Transporte Público    |
| 10 - Moto                    | <b>(P)</b> Particular motorizado |
| 11 - Bicicleta               | <b>(N)</b> Não motorizado        |
| 12 - A Pé                    | <b>(N)</b> Não motorizado        |

Assim, a seqüência das letras (P, T, N) representa a seqüência de modos de transporte utilizados durante a cadeia de viagens realizadas pelos indivíduos no período de um dia. A ordem das letras indica a ordem cronológica em que cada modo foi utilizado (primeira viagem, segunda viagem, etc.). A seguir, são citados exemplos de seqüência de modos de transporte para indivíduos que realizaram duas, três e quatro viagens respectivamente.

- (1) Seqüência "PT" – Representa indivíduos que realizaram duas viagens, utilizando modo de transporte particular na primeira viagem e transporte público na segunda viagem;

- (2) Seqüência “PPN” – Representa indivíduos que realizaram três viagens no total: 1ª viagem - modo de transporte particular; 2ª viagem - modo de transporte particular; e 3ª viagem - modo de transporte não motorizado.
- (3) Seqüência “PNNT” – Representa indivíduos que realizaram quatro viagens no total: 1ª viagem - modo de transporte particular; 2ª viagem - modo de transporte não motorizado; 3ª viagem - modo de transporte não motorizado; e 4ª viagem – modo de transporte público.

A exemplo da seção anterior, a Figura 6.3 ilustra as seqüências de modos de transporte para duas, três e quatro viagens respectivamente, na amostra que contém trabalhadores no setor de serviços. Os algarismos romanos representam a ordem de realização das viagens. Já a Tabela 6.4, a seguir, exibe as seqüências de modos de viagem mais freqüentes para cada uma das amostras de trabalhadores e estudantes.

Tabela 6.4 Seqüências de modos de transporte mais freqüentes

| Amostras                                 |                       | seqüências de modos de transporte |           |           |             |             |             |             |             |             |             |
|------------------------------------------|-----------------------|-----------------------------------|-----------|-----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| <b>Trabalhadores por setor econômico</b> | Indústria<br>4.102    | <b>TT</b>                         | <b>PP</b> | <b>NN</b> | <b>NNNN</b> | <b>PPPP</b> | <b>PPP</b>  | <b>TTNN</b> | <b>TTT</b>  | <b>TP</b>   | <b>TTTT</b> |
|                                          |                       | 47,34%                            | 20,82%    | 13,51%    | 5,14%       | 3,61%       | 1,58%       | 1,49%       | 1,10%       | 0,90%       | 0,80%       |
|                                          | Comércio<br>6.043     | <b>TT</b>                         | <b>PP</b> | <b>NN</b> | <b>NNNN</b> | <b>TNN</b>  | <b>PPPP</b> | <b>PPP</b>  | <b>TNT</b>  | <b>TTNN</b> | <b>TNNT</b> |
|                                          | 33,63%                | 26,17%                            | 16,10%    | 5,78%     | 4,24%       | 3,84%       | 1,90%       | 1,31%       | 1,04%       | 0,96%       |             |
|                                          | Serviços<br>8.328     | <b>TT</b>                         | <b>PP</b> | <b>NN</b> | <b>PPPP</b> | <b>NNNN</b> | <b>PPP</b>  | <b>TTT</b>  | <b>TNNT</b> | <b>TTNN</b> | <b>TP</b>   |
|                                          | 41,55%                | 26,30%                            | 12,36%    | 4,80%     | 3,00%       | 2,92%       | 2,05%       | 1,25%       | 1,09%       | 0,88%       |             |
| <b>Estudantes por grau de instrução</b>  | Até 1º grau<br>13.839 | <b>NN</b>                         | <b>PP</b> | <b>TT</b> | <b>NNNN</b> | <b>PN</b>   | <b>PT</b>   | <b>PPP</b>  | <b>PPPP</b> | <b>NNN</b>  | <b>TN</b>   |
|                                          |                       | 62,65%                            | 16,52%    | 9,30%     | 2,20%       | 1,57%       | 1,04%       | 0,96%       | 0,87%       | 0,85%       | 0,51%       |
|                                          | No 2º grau<br>2.143   | <b>NN</b>                         | <b>TT</b> | <b>PP</b> | <b>PT</b>   | <b>PPPP</b> | <b>TTNN</b> | <b>NNNN</b> | <b>PN</b>   | <b>TTTT</b> | <b>TP</b>   |
|                                          | 35,23%                | 29,82%                            | 13,77%    | 5,97%     | 2,01%       | 1,96%       | 1,68%       | 1,68%       | 0,98%       | 0,70%       |             |
|                                          | Universidade<br>568   | <b>PP</b>                         | <b>TT</b> | <b>NN</b> | <b>PPPP</b> | <b>PPP</b>  | <b>PT</b>   | <b>TTTT</b> | <b>TP</b>   | <b>NNN</b>  | <b>TTT</b>  |
|                                          | 41,20%                | 33,27%                            | 7,57%     | 5,63%     | 2,82%       | 2,29%       | 1,41%       | 1,23%       | 1,06%       | 0,88%       |             |

Observando a Tabela 6.4, pode-se verificar grande freqüência do uso de modo de transporte não motorizado para o caso de estudantes (principalmente até o 1º grau). Para explicar este fato, deve-se lembrar a distribuição geográfica de tais atividades no meio urbano exemplificada no Capítulo 4 (Figuras 4.7).

Para o caso de estudantes até o 1º grau, por exemplo, há uma grande dispersão de oferta de escolas pela RMSP. Intuitivamente espera-se que haja escolas disponíveis, próximas aos domicílios, na maioria das localizações da região de estudo. Assim, dificilmente haveria necessidade de uso de outros modos de transporte diferentes do não motorizado para viagens com motivo ESCOLA. Observa-se também uma boa frequência para a seqüência "PP" (16,52%). Provavelmente crianças viajando como passageiros de modo de transporte particular motorizado.

Já no caso dos estudantes universitários, por exemplo, onde se notou uma concentração de ofertas de universidades em pontos bem restritos da RMSP (Cidade Universitária, Vila Gerti, Mogi das Cruzes, etc.), espera-se que predominem outros modos de transportes diferentes dos não motorizados, para a maioria das viagens com motivo ESCOLA. Trata-se, neste caso, da necessidade de realizar viagens longas (devido à ausência de oferta de universidades próximas à residência) e, por consequência, uso de modos de transporte motorizados (PP – 41,20%; TT – 33,27%).

No caso de industriários, onde se nota também concentração de atividade industrial em apenas alguns pólos da RMSP, verifica-se uma maior frequência relacionada a modo de transporte público motorizado (TT- 47,43%). Principalmente quando comparados às demais amostras de trabalhadores (comerciários "TT" – 33,63% e trabalhadores no setor de serviços "TT" – 41,55%).



| Número de viagens      | Seqüências de modos de transporte                  |                  |                  |               |               |               |
|------------------------|----------------------------------------------------|------------------|------------------|---------------|---------------|---------------|
| 2                      |                                                    |                  |                  |               |               |               |
|                        | 3.460 indivíduos                                   | 2.190 indivíduos | 1.029 indivíduos | 73 indivíduos | 66 indivíduos |               |
| 6.877 indivíduos       |                                                    |                  |                  |               |               |               |
|                        | 20 indivíduos                                      | 14 indivíduos    | 13 indivíduos    | 12 indivíduos |               |               |
| 3                      |                                                    |                  |                  |               |               |               |
|                        | 243 indivíduos                                     | 171 indivíduos   | 30 indivíduos    | 8 indivíduos  |               |               |
| 4                      |                                                    |                  |                  |               |               |               |
|                        | 400 indivíduos                                     | 250 indivíduos   | 104 indivíduos   | 91 indivíduos | 53 indivíduos | 53 indivíduos |
|                        |                                                    |                  |                  |               |               |               |
| 999 indivíduos         | 21 indivíduos      19 indivíduos      8 indivíduos |                  |                  |               |               |               |
| Total 8.328 indivíduos |                                                    |                  |                  |               |               |               |

Figura 6.3 Seqüências de modos de transporte para amostras de trabalhadores no setor de serviços



## 6.4 SEQUÊNCIA DE DESTINOS (ETAPA 3)

A terceira etapa, necessária para codificação dos padrões de encadeamento de viagens (variável dependente), está relacionada à representação das seqüências dos destinos escolhidos pelos indivíduos durante seus respectivos itinerários de viagem.

Encontra-se dificuldade nesta etapa, principalmente devido ao alto número de zonas de tráfego contidas na RMSP (389 no total), e, conseqüentemente, a grande quantidade de destinos e origens possíveis.

Há uma importante necessidade de limitação de destinos, devido à restrição de categorias da variável dependente do *software* utilizado para aplicação da AD (S-PLUS 6.1), conforme mencionado anteriormente. Assim, foi feita a aplicação da AC com os seguintes objetivos: (1) agrupar os destinos escolhidos pelos indivíduos, baseando-se em um critério adotado pela autora (distâncias da primeira viagem); e (2) reduzir o número de destinos a serem representados na variável dependente.

A variável escolhida para agrupamento foi a distância da primeira viagem. Desta forma, os destinos agrupados caracterizam-se por faixas de distâncias. No presente trabalho, os destinos são representados pelas distâncias a partir do centróide da origem inicial (residência). Assim, pode-se definir a presente etapa como caracterização das *seqüências de distâncias a partir da zona de domicílio* ou simplesmente *seqüência de destinos*.

### 6.4.1 Aplicação da AC

Considerando a RMSP e seu número de zonas de tráfego muito alto, a representação destas como unidades base de deslocamentos, principalmente para caracterizar a variável dependente, torna-se inviável. A principal finalidade da aplicação da AC é viabilizar a representação da seqüência de destinos, juntamente com a seqüência de atividades (motivos das viagens) e modos de transporte como variável dependente.

Para aplicação da técnica de AC, utilizando o *software* SPSS 13.0, as etapas sugeridas por Malhotra (2001), descritas no Capítulo 3, foram seguidas:

### ✚ Definição do problema de aglomeração ou objetivo da aplicação da AC na pesquisa:

Encontrar grupos homogêneos, segundo o critério distância, a fim de reduzir o número de destinos possíveis e viabilizar a representação da variável dependente. A utilidade da AC na pesquisa é propor uma classificação empírica dos destinos, considerando como critério as distâncias a partir da zona de domicílio.

### ✚ Escolha das variáveis

Neste trabalho foi escolhida apenas uma variável para aglomeração dos destinos. A variável adotada foi a distância da primeira viagem. A primeira viagem geralmente é a principal (ao trabalho, à escola), por isso optou-se pela escolha das distâncias da primeira viagem. A amostra considerada levou em conta todos os indivíduos, trabalhadores dos três setores econômicos, assim como estudantes pertencentes às três escolaridades (somatório das seis amostras analisadas – 35.023 indivíduos).

### ✚ Escolha da medida de similaridade ou dissimilaridade e do processo de aglomeração

O processo de aglomeração utilizado foi o hierárquico aglomerativo (descrito no Capítulo 3) e a medida de similaridade foi a distância euclidiana.

### ✚ Definição prévia de número de *clusters*

Neste estudo adotou-se um número fixo de quatro *clusters* (*single solution*) para a solução final. O número restrito deve-se principalmente à necessidade de limitação de possíveis destinos. Para 10 destinos (*clusters*), por exemplo, haveria um grande número de combinações possíveis associadas às seqüências de destinos.

### 6.4.2 Resultados do agrupamento

Considerando-se a variável adotada, foram obtidos, através da aplicação da AC, quatro destinos (*clusters*), caracterizados por faixas de distâncias. A Tabela 6.5 contém os quatro *clusters* formados, a quantidade de indivíduos agrupados em cada *cluster*, a distância de viagem média por conglomerado, assim como os seus valores máximos e mínimos em cada um dos quatro grupos.

Tabela 6.5 Características dos quatro *clusters* obtidos

| Cluster | indivíduos | Distância_primeira_viagem média | Desvio padrão (km) | Distância Máxima (km) | Distância Mínima (km) |
|---------|------------|---------------------------------|--------------------|-----------------------|-----------------------|
| 1       | 21.739     | 1,84 km                         | 0,85               | 4,01                  | 0,09                  |
| 2       | 7.436      | 6,56 km                         | 1,79               | 10,15                 | 4,02                  |
| 3       | 4.020      | 14,13 km                        | 2,68               | 19,61                 | 10,16                 |
| 4       | 1.828      | 26,79 km                        | 7,12               | 80,72                 | 19,64                 |

### 6.4.3 Caracterização das seqüências de destinos

As seqüências de destinos escolhidos durante a cadeia de viagens realizadas pelos indivíduos no período de um dia foram representadas através de uma seqüência de números (1, 2, 3, 4). Cada um dos números representa um dos quatro conglomerados obtidos com a aplicação da AC. As faixas de distâncias foram adaptadas dos resultados da aglomeração:

- (1) Destino 1: destinos localizados até 5 km do centróide da zona de origem inicial (domicílio);
- (2) Destino 2: destinos localizados entre 5 km e 10 km do centróide da zona de origem inicial (domicílio);
- (3) Destino 3: destinos localizados entre 10 km e 15 km do centróide da zona de origem inicial (domicílio);
- (4) Destino 4: destinos localizados acima de 15 km do centróide da zona de origem inicial (domicílio).

Vale ressaltar que os destinos foram caracterizados em faixas de distâncias a partir do centróide da zona de origem inicial (domicílio). Então, no caso de realização de três viagens, por exemplo, não se

sabe a faixa de distância entre o segundo e terceiro destinos. Conhece-se sim, a faixa de distância entre o segundo destino e a residência e a faixa de distância entre o terceiro destino e a residência.

A seqüência numérica representa a seqüência de destinos escolhidos durante a cadeia de viagens realizadas pelos indivíduos no período de um dia. A ordem dos números indica a ordem cronológica em que ocorreu viagem a cada destino (primeira viagem, segunda viagem, etc.). Citam-se, a seguir, exemplos de seqüência de destinos para indivíduos que realizaram duas, três e quatro viagens respectivamente. O primeiro e último números caracterizam a origem e destino final (no caso, a residência), sempre representados pelo número 1.

(A) Seqüência "141" – Representa indivíduos que realizaram duas viagens, ambas as viagens têm distâncias superiores a 15 km;

(B) Seqüência "1241" - Representa indivíduos que realizaram três viagens no total: 1º destino – situa-se entre 5 e 10 km do centróide da zona de residência ; 2º destino – localiza-se a uma distância acima de 15 km do centróide da zona de residência (sabe-se a distância entre o segundo destino e a residência, mas não a distância entre o primeiro e o segundo destinos); e 3º destino – residência;

(C) Seqüência "13241" - Representa indivíduos que realizaram quatro viagens no total: 1º destino – localiza-se a uma distância entre 10 e 15 km do centróide da zona de residência; 2º destino – situa-se entre 5 e 10 km do centróide da zona de residência; 3º destino – está localizado a uma distância superior a 15 km do centróide da zona de residência; 4º destino – domicílio.

A Tabela 6.6 mostra as seqüências de destinos mais freqüentes para cada uma das amostras de trabalhadores e estudantes. A Figura 6.4, em seguida, ilustra a seqüência de destinos "13341" de um indivíduo que mora na zona "Sé". Estão representadas as faixas de distâncias, bem como a localização espacial (na RMSP) da seqüência de destinos exemplificada.

Tabela 6.6 Sequências destinos mais freqüentes

| Amostras                                                     |                       | seqüências de destinos |            |              |              |              |              |              |              |              |              |
|--------------------------------------------------------------|-----------------------|------------------------|------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|
| <b>Trabalhadores</b><br><b>por</b><br><b>setor econômico</b> | Indústria<br>4.102    | <b>111</b>             | <b>121</b> | <b>141</b>   | <b>131</b>   | <b>11111</b> | <b>12111</b> | <b>1111</b>  | <b>14111</b> | <b>13111</b> | <b>11121</b> |
|                                                              |                       | 34,18%                 | 19,84%     | 17,65%       | 11,95%       | 8,53%        | 1,27%        | 0,88%        | 0,68%        | 0,61%        | 0,56%        |
|                                                              | Comércio<br>6.043     | <b>111</b>             | <b>121</b> | <b>141</b>   | <b>11111</b> | <b>131</b>   | <b>12111</b> | <b>12211</b> | <b>1111</b>  | <b>12221</b> | <b>1441</b>  |
|                                                              |                       | 37,65%                 | 17,52%     | 12,84%       | 11,12%       | 9,78%        | 2,76%        | 2,12%        | 1,92%        | 0,50%        | 0,43%        |
|                                                              | Serviços<br>8.328     | <b>111</b>             | <b>121</b> | <b>141</b>   | <b>131</b>   | <b>11111</b> | <b>1111</b>  | <b>1221</b>  | <b>12221</b> | <b>1211</b>  | <b>11121</b> |
|                                                              |                       | 37,92%                 | 19,10%     | 14,07%       | 11,48%       | 8,41%        | 1,90%        | 0,77%        | 0,75%        | 0,71%        | 0,67%        |
| <b>Estudantes</b><br><b>por</b><br><b>Grau de instrução</b>  | Até 1º grau<br>13.839 | <b>111</b>             | <b>121</b> | <b>11111</b> | <b>1111</b>  | <b>131</b>   | <b>141</b>   | <b>11121</b> | <b>12111</b> | <b>13111</b> | <b>1221</b>  |
|                                                              |                       | 84,33%                 | 5,73%      | 4,41%        | 2,13%        | 1,48%        | 1,10%        | 0,33%        | 0,14%        | 0,11%        | 0,05%        |
|                                                              | No 2º grau<br>2.143   | <b>111</b>             | <b>121</b> | <b>11111</b> | <b>131</b>   | <b>141</b>   | <b>1111</b>  | <b>11121</b> | <b>12111</b> | <b>14111</b> | <b>1221</b>  |
|                                                              |                       | 68,88%                 | 11,48%     | 6,25%        | 4,11%        | 3,69%        | 1,45%        | 0,79%        | 0,79%        | 0,65%        | 0,56%        |
|                                                              | Universidade<br>568   | <b>111</b>             | <b>121</b> | <b>141</b>   | <b>131</b>   | <b>11111</b> | <b>1111</b>  | <b>1211</b>  | <b>12121</b> | <b>13111</b> | <b>14111</b> |
|                                                              |                       | 32,04%                 | 24,12%     | 17,08%       | 13,38%       | 3,87%        | 1,94%        | 1,23%        | 1,23%        | 1,06%        | 1,06%        |

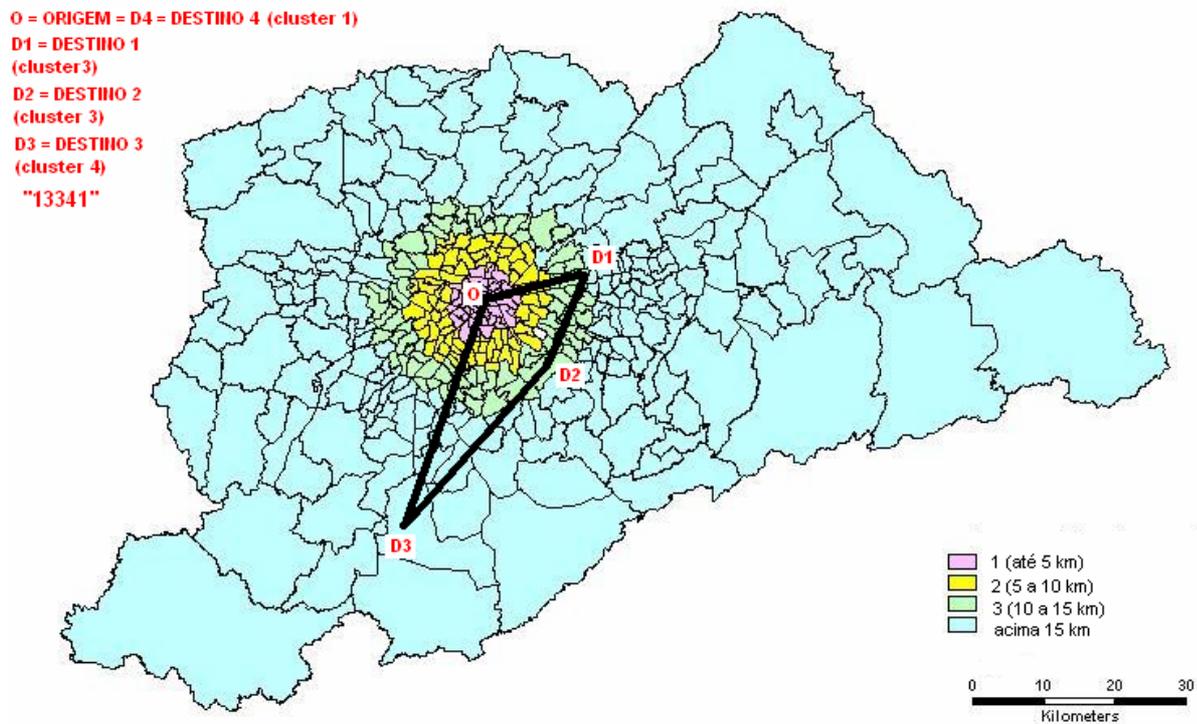


Figura 6.4 Exemplo de seqüência de destinos (origem e destino final zona Sê) – "13341"

Observando a Tabela 6.6, verificam-se altas freqüências de ocorrência para maiores faixas de distâncias de viagem (destinos mais distantes), considerando especificamente as amostras de industriários e universitários (Seqüência "141" – 17,65% e 17,08% respectivamente). Como é de se esperar, indivíduos pertencentes às duas amostras referidas, intuitivamente, realizam viagens mais longas para satisfazer a atividade de subsistência (trabalho na indústria ou estudo na universidade). A necessidade de realizar viagens longas explica-se novamente pela distribuição geográfica da oferta de

atividade industrial e de universidades na RMSP. A distribuição de tais atividades é concentrada em alguns pólos específicos na região de estudo.

Já na amostra composta de estudantes até o 1º grau é onde se observa a mais alta frequência para a seqüência de destinos (“111” – 84,33%). Como há oferta de escolas até o 1º grau dispersa por toda RMSP, muito raramente há necessidade de realização de viagens longas (destinos distantes) para satisfazer a atividade de estudo. Em tal amostra observa-se uma frequência muito baixa da seqüência “141” (1,10%), por exemplo.

Finalmente, na próxima seção será descrita a etapa final para codificação da variável dependente. Serão reunidas as três codificações (motivo de viagem ou atividades – “letras”, modo de transporte – “letras” e destinos – “números”) descritas nas etapas anteriores.

#### 6.5 PADRÕES FINAIS (ETAPA 4)

A partir da codificação atribuída às seqüências das viagens na primeira etapa (motivos de viagem), na segunda etapa (modos de transporte) e na terceira etapa (destinos) foi feita a combinação das etapas iniciais para obtenção dos padrões finais. Os padrões finais representam o número de viagens realizadas, a seqüência de atividades (motivos de viagem), a seqüência de modos de transporte e a seqüência de destinos escolhidos durante a cadeia de viagens.

Os padrões finais foram representados por três conjuntos de caracteres: o primeiro, referente à seqüência de atividades (**HWAWH**, por exemplo), o segundo correspondente à seqüência dos modos de viagens principais utilizados (**PNNP**, por exemplo) e o terceiro correspondente à seqüência de destinos (**14421**, por exemplo). Assim, obtém-se, por exemplo, o padrão final **HWAWHPNNP14421**.

O padrão final recebe uma codificação alfa-numérica. A seqüência de letras e números indica os atributos de viagem representados no presente trabalho. A Figura 6.5 ilustra um exemplo do procedimento proposto, sendo feita a combinação de caracteres atribuídos nas etapas anteriores para obtenção dos padrões finais (variável dependente).

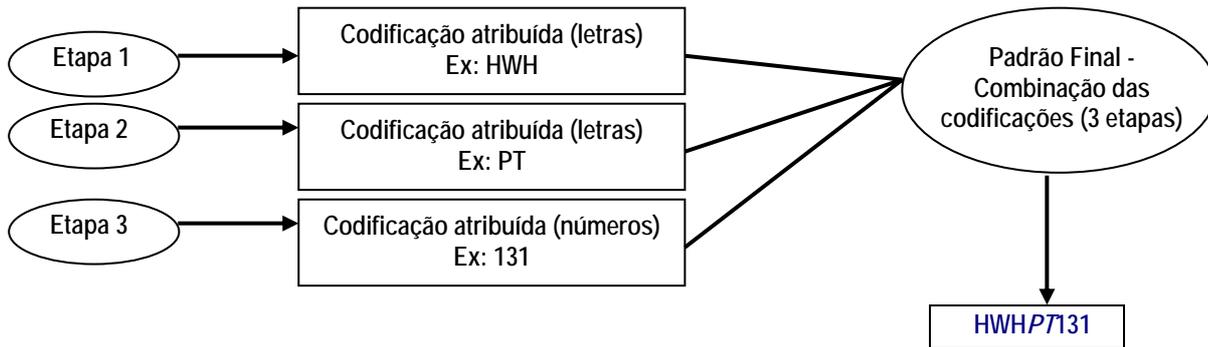


Figura 6.5 Obtenção do padrão final através da combinação das codificações atribuídas nas três primeiras etapas.

O padrão de encadeamento de viagem "HWHPT131", por exemplo, é uma das categorias da variável dependente e representa indivíduos que realizaram duas viagens no dia anterior ao da entrevista:

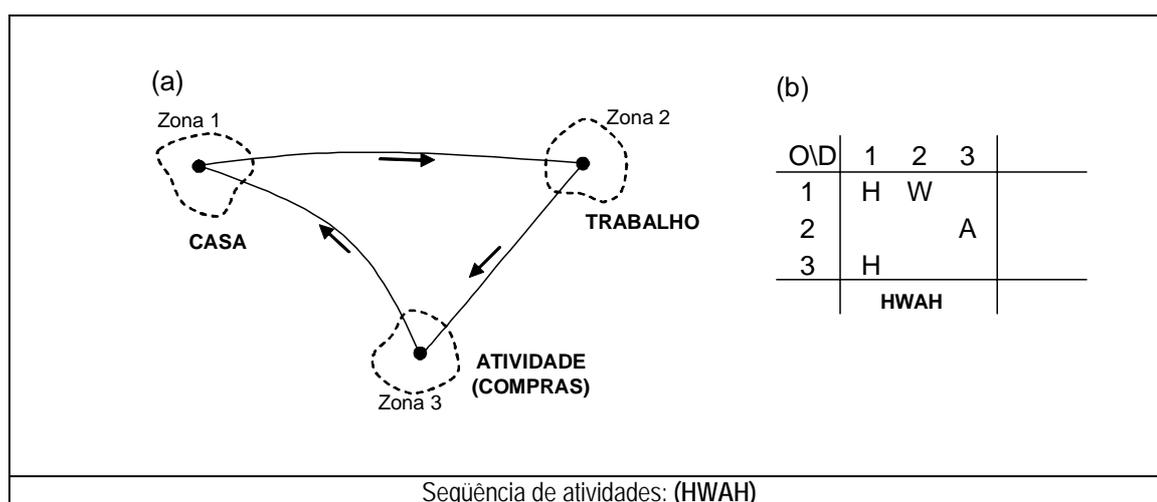
- 1ª viagem: MOTIVO DA VIAGEM: Trabalho (W); MODO DE TRANSPORTE UTILIZADO: Particular motorizado (P); DESTINO: localiza-se a uma distância entre 10 e 15 km do centróide da zona da residência (3);
- 2ª viagem: MOTIVO DA VIAGEM: Residência (H); MODO DE TRANSPORTE UTILIZADO: Público (T); DESTINO: domicílio (1).

A Tabela 6.7 mostra os oito padrões de encadeamento de viagens mais freqüentes em cada uma das seis amostras analisadas. A variável dependente possui 128 categorias em cada amostra. As categorias correspondem aos 128 padrões de encadeamento de viagens mais freqüentes em cada um dos casos (indústria, comércio, serviços, estudantes até o 1º grau, 2º grau e universidade). A eliminação daqueles indivíduos que realizaram padrões de viagens menos freqüentes (129 menos freqüente em diante) delimitou o tamanho das amostras.

Tabela 6.7 Padrões de encadeamento de viagens mais freqüentes

| Amostras                 |              | PADRÕES DE ENCADEAMENTO DE VIAGENS |                 |                 |                 |                 |                 |                 |                       |
|--------------------------|--------------|------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------------|
| <b>Trabalhadores</b>     | Indústria    | <b>HWHTT121</b>                    | <b>HWHTT111</b> | <b>HWHTT141</b> | <b>HWHNN111</b> | <b>HWHTT131</b> | <b>HWHP111</b>  | <b>HWHP121</b>  | <b>HWHP141</b>        |
|                          | 4.102        | 12,75%                             | 12,31%          | 12,24%          | 11,70%          | 8,41%           | 7,02%           | 5,24%           | 4,44%                 |
|                          | <b>por</b>   | Comércio                           | <b>HWHNN111</b> | <b>HWHP111</b>  | <b>HWHTT121</b> | <b>HWHTT111</b> | <b>HWHTT141</b> | <b>HWHP121</b>  | <b>HWHTT131</b>       |
| <b>setor econômico</b>   | 6.043        | 13,09%                             | 9,45%           | 8,29%           | 8,14%           | 7,76%           | 6,21%           | 5,46%           | 3,57%                 |
|                          | Serviços     | <b>HWHTT121</b>                    | <b>HWHTT111</b> | <b>HWHNN111</b> | <b>HWHP111</b>  | <b>HWHTT141</b> | <b>HWHTT131</b> | <b>HWHP121</b>  | <b>HWHP141</b>        |
|                          | 8.328        | 10,47%                             | 10,40%          | 10,12%          | 9,92%           | 9,26%           | 7,20%           | 5,90%           | 3,58%                 |
| <b>Estudantes</b>        | Até 1º grau  | <b>HSHNN111</b>                    | <b>HSHPP111</b> | <b>HSHTT111</b> | <b>HSHNN121</b> | <b>HSHPP121</b> | <b>HSHPN111</b> | <b>HSHTT121</b> | <b>HSHSHNNNN11111</b> |
|                          | 13.839       | 59,35%                             | 13,41%          | 5,82%           | 1,71%           | 1,61%           | 1,50%           | 1,43%           | 1,06%                 |
|                          | <b>por</b>   | No 2º grau                         | <b>HSHNN111</b> | <b>HSHTT111</b> | <b>HSHPP111</b> | <b>HSHTT121</b> | <b>HSHPT111</b> | <b>HSHTT131</b> | <b>HSHTT141</b>       |
| <b>Grau de instrução</b> | 2.143        | 33,04%                             | 16,19%          | 9,47%           | 6,67%           | 4,29%           | 2,47%           | 2,24%           | 2,01%                 |
|                          | Universidade | <b>HSHPP121</b>                    | <b>HSHPP111</b> | <b>HSHTT121</b> | <b>HSHPP141</b> | <b>HSHTT111</b> | <b>HSHTT141</b> | <b>HSHNN111</b> | <b>HSHPP131</b>       |
|                          | 568          | 12,50%                             | 11,97%          | 9,68%           | 8,27%           | 8,10%           | 7,93%           | 6,87%           | 6,51%                 |

Observando a Tabela 6.7, pode-se ter uma idéia inicial de que a seqüência de atividades varia muito pouco em todas as amostras (“HWH” – no caso de trabalhadores e “HSH” – no caso de estudantes). No entanto, foram mostrados apenas os oito padrões mais freqüentes em todos os casos. Há uma variação mais acentuada da agenda de atividades dos indivíduos, principalmente considerando aqueles padrões menos freqüentes. Devido a este fato, optou-se por representar a seqüência de atividades juntamente com os outros atributos de viagem, apesar desta não ter uma expressiva variação quando se observa predominância de padrões de viagem. A Figura 6.6, em seguida, ilustra as etapas de codificação do padrão de encadeamento de viagem. É representada a categoria “**HWAHTTT1321**”.



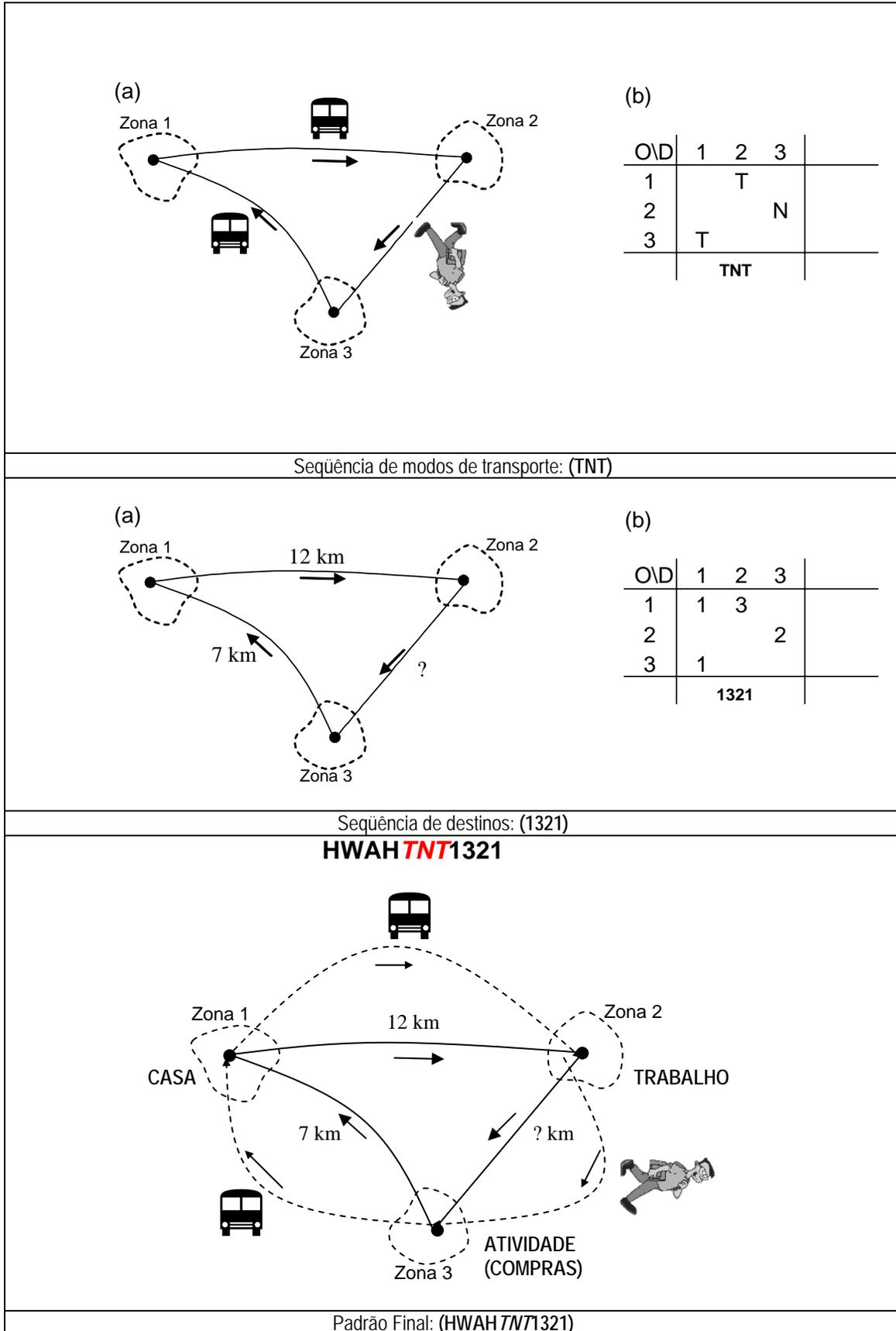


Figura 6.6 Codificação da categoria “HWAH *TNT* 1321” da variável dependente

## Capítulo

# 7 APLICAÇÃO DE AD E RESULTADOS

---

Realizadas as etapas de tratamento das amostras e representação das variáveis envolvidas, torna-se possível a aplicação da Arvore de Decisão (AD). Este capítulo apresenta as seis árvores geradas (para cada uma das amostras da análise), bem como síntese dos relatórios emitidos pelo S-PLUS 6.1. São verificadas as variáveis independentes consideradas importantes pelo algoritmo CART para segregação dos dados e os padrões de viagem predominantes em cada grupo de indivíduos. O escopo do presente capítulo é apresentar os resultados obtidos, postergando análises e discussões para o capítulo seguinte.

### 7.1. VARIÁVEIS ENVOLVIDAS E CRITÉRIOS ADOTADOS

Após o tratamento das amostras utilizadas e representação das variáveis independentes e dependente, torna-se viável a aplicação da Árvore de Decisão (AD). O processamento da AD contida no S-PLUS 6.1 é realizado considerando as etapas esquematizadas na Figura 7.1.

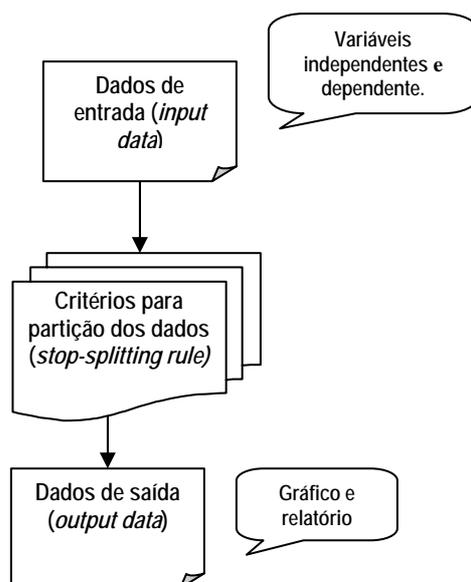


Figura 7.1 Processamento da AD contida no S-PLUS 6.1

Os dados de entrada são definidos como as variáveis envolvidas na análise, representadas nos capítulos anteriores (Capítulos 5 e 6). Foi utilizada como variável dependente o “Padrão de encadeamento de viagens” (padrões de viagens segundo motivos de viagem, modos de transporte e destinos). Cada padrão de viagem corresponde a uma categoria de variável dependente. Foi utilizado um total de 128 padrões de viagens (categorias de variável dependente) em cada uma das amostras.

Já as variáveis independentes foram representadas através de três conjuntos de variáveis:

- (1) Participação em atividades, contendo duas variáveis categóricas: TRABALHA; ESTUDA;
- (2) Características socioeconômicas, contendo sete variáveis numéricas e cinco categóricas: RENDA INDIVIDUAL; RENDA FAMILIAR; NÚMERO DE AUTOMÓVEIS NO DOMICÍLIO; NÚMERO PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO; TOTAL DE PESSOAS NO DOMICÍLIO; NÚMERO DE CRIANÇAS NO DOMICÍLIO; IDADE; TIPO DE CHEFE DO DOMICÍLIO; POSIÇÃO DO INDIVÍDUO NO DOMICÍLIO; SEXO; GRAU DE INSTRUÇÃO; USA VALE TRANSPORTE;
- (3) Uso do solo, contendo quatro variáveis numéricas: PARCELA ACUMULADA DE “OPORTUNIDADES” ATÉ UM RAIOS DE 5 km DO CENTRÓIDE DA ZONA DE RESIDÊNCIA; PARCELA ACUMULADA DE “OPORTUNIDADES” ATÉ UM RAIOS DE 10 km DO CENTRÓIDE DA ZONA DE RESIDÊNCIA; PARCELA ACUMULADA DE “OPORTUNIDADES” ATÉ UM RAIOS DE 15 km DO CENTRÓIDE DA ZONA DE RESIDÊNCIA; PARCELA ACUMULADA DE “OPORTUNIDADES” ATÉ UM RAIOS DE 20 km DO CENTRÓIDE DA ZONA DE RESIDÊNCIA.

Para a geração das árvores no S-PLUS 6.1, os critérios adotados para a sua classificação (*stop-splitting rule*) foram o mínimo de 50 observações por nó terminal e desvio mínimo de 0,15 em relação ao nó principal.

Assim, na etapa final de processamento dos resultados, foi gerado um total de seis árvores (uma para cada amostra analisada: trabalhadores nos setores industrial, comercial e de serviços e

estudantes até o 1º grau, no 2º grau e universitários). As árvores foram construídas a partir dos dados filtrados e adaptados da pesquisa O/D realizada pelo METRÔ-SP, em 1997.

O S-PLUS 6.1 gera os resultados de forma gráfica (Árvore de Classificação e Regressão) e tabular. A forma tabular consiste num relatório que apresenta detalhadamente os resultados da árvore: número do nó, total de observações no nó, desvio, padrão predominante e as probabilidades de ocorrência de todos os padrões da análise. Os valores apresentados nos nós terminais (folhas) do gráfico correspondem ao padrão predominante encontrado em cada grupo de indivíduos delimitado pela árvore.

As próximas seções trazem os respectivos gráficos (árvores) para as seis amostras, bem como síntese do relatório gerado pelo *software*, contendo todos os nós terminais encontrados, os dez padrões mais freqüentes para cada folha e o número de indivíduos contidos em cada folha. O leitor interessado em saber a probabilidade de ocorrência de todos os 128 padrões para cada nó terminal em cada uma das seis amostras deve observar o texto anexo (Anexo A, Anexo B, Anexo C, Anexo D, Anexo E, Anexo F)

## 7.2. TRABALHADORES NO SETOR INDUSTRIAL

A árvore gerada com desvio mínimo de 0,15 para amostra de industriários apresentou como resultado final 10 classes distintas de indivíduos (folhas) e seus respectivos padrões de viagem predominantes, ou seja, 10 nós terminais a partir do conjunto de dados correspondente à amostra estudada. Dados ausentes foram omitidos na etapa de processamento da AD, de modo que a amostra de industriários apresentada ao final dos resultados era composta de 4.084 indivíduos.

As variáveis independentes e dependente utilizadas como dados de entrada foram descritas anteriormente. No presente caso, as variáveis relativas a uso do solo são definidas como "PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA" nas quatro faixas de distância consideradas (até 5 km a partir do centróide da zona de residência; até 10 km a partir do centróide da zona de residência; até

15 km a partir do centróide da zona de residência; até 20 km a partir do centróide da zona de residência).

A variável de maior importância é "NÚMERO (PROVÁVEL) DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO" (CNH). A partir da raiz, a árvore se ramifica em dois grupos principais: (1)  $CNH = 0$ ; e (2)  $CNH \geq 1$ .

O ramo esquerdo da árvore corresponde às respostas positivas para construção da hierarquia da árvore, enquanto que o ramo direito corresponde às respostas negativas. O ramo esquerdo é formado por aqueles industriários que residem em domicílios sem nenhuma (provável) carteira de habilitação. Tal grupo de pessoas se subdivide então considerando a segunda variável mais relevante para segregação dos dados, que neste caso é USA VALE TRANSPORTE (Usa VTRA). A árvore novamente se ramifica, formando mais dois grupos: (3) Usa VTRA; e (4) Não Usa VTRA.

São formadas desta maneira duas classes de indivíduos que provavelmente não possuem nenhuma carteira de habilitação no domicílio ( $CNH=0$ ):

- $CNH = 0$  e Usa VTRA;
- $CNH = 0$  e Não Usa VTRA.

Já aqueles industriários que possuem pelo menos uma carteira de habilitação no domicílio ( $CNH \geq 1$ ) são subdivididos primeiramente de acordo com a variável ESTUDA, formando então duas novas classes:

- $CNH \geq 1$  e Não estudantes;
- $CNH \geq 1$  e estudantes.

O número (provável) de carteiras de habilitação no domicílio, o uso do vale transporte e a participação das pessoas em atividades de estudos foram as variáveis mais importantes para construção da AD. Um esquema com as quatro classes inicialmente formadas na amostra de trabalhadores no setor industrial é mostrado na Figura 7.2.



caracterizada pelas características socioeconômicas, de uso do solo e participação em atividades - todas indicadas na segunda coluna.

Além dos dez padrões predominantes por nó terminal, já mencionados, apresenta-se também a frequência de ocorrência de cada um deles e a quantidade de indivíduos que os realizaram.

As probabilidades de ocorrência dos demais padrões da análise (128 no total) estão apresentadas nos anexos do presente texto.

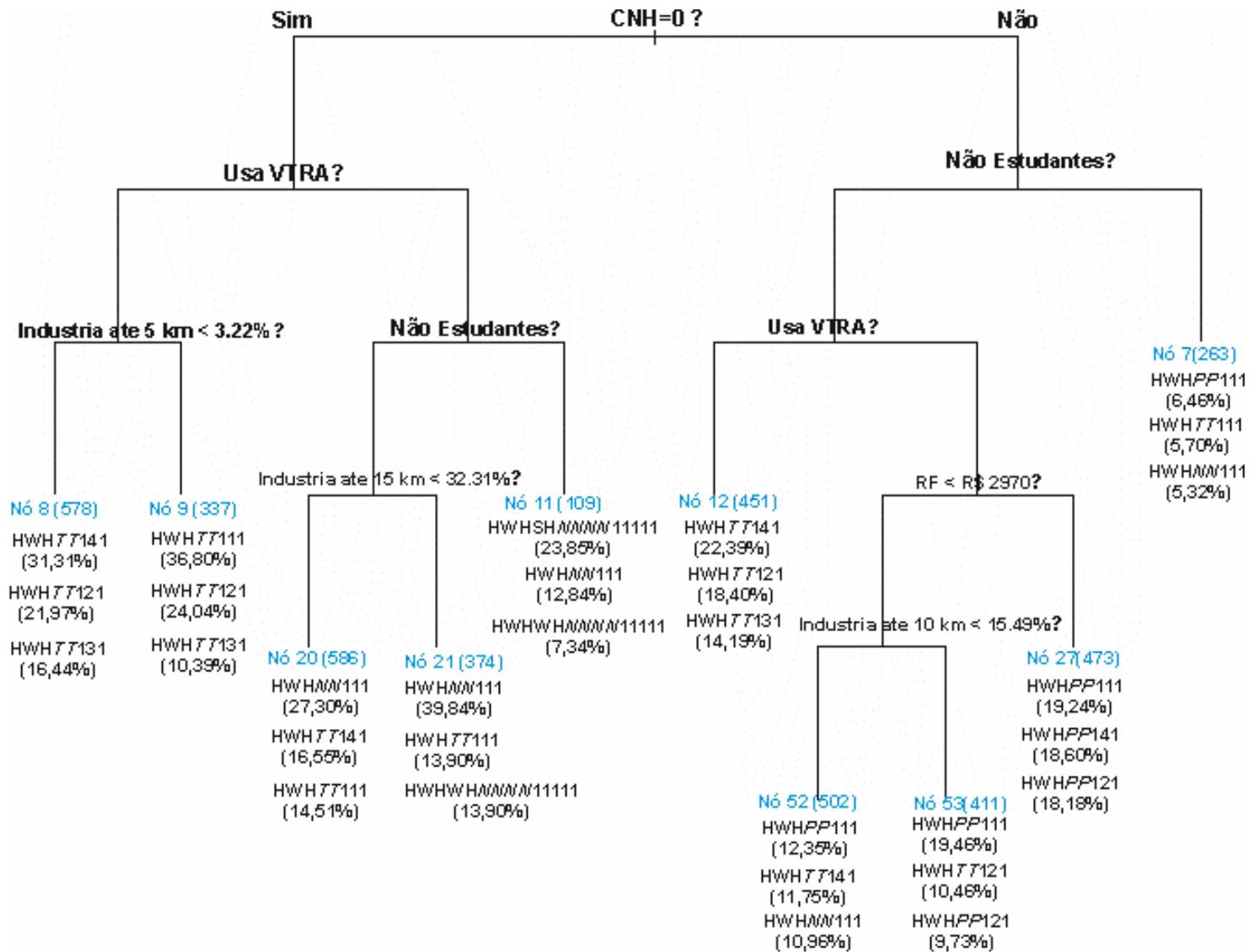


Figura 7.3 Gráfico adaptado do S-PLUS 6.1 (amostra de industriários)

LEGENDA: CNH (Nº provável de carteiras de habilitação no domicílio); VTRA (Vale Transporte); Indústria até 5 km (Parcela acumulada de empregos na indústria até 5 km do centróide da zona de residência); Indústria até 10 km (Parcela acumulada de empregos na indústria até 10 km do centróide da zona de residência); Indústria até 15 km (Parcela acumulada de empregos na indústria até 15 km do centróide da zona de residência); RF (Renda Familiar).



Tabela 7.1 Síntese do relatório (amostra de industriários)

| Nó         | Características dos grupos   | Padrões mais frequentes |           |                  |           |                        |                  |                        |                        |                        |                        |                        |
|------------|------------------------------|-------------------------|-----------|------------------|-----------|------------------------|------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
|            |                              | HWHT7141                | HWHT7121  | HWHT7131         | HWHT7111  | HWHTNN111              | HAHT7111         | HWHTSH7111111111       | HAHT7141               | HWHTSH7111111111111111 | HWHTSH7111111111111111 | HWHTSH7111111111111111 |
| 8          | CNH=0                        | HWHT7141                | HWHT7121  | HWHT7131         | HWHT7111  | HWHTNN111              | HAHT7111         | HWHTSH7111111111       | HAHT7141               | HWHTSH7111111111111111 | HWHTSH7111111111111111 | HWHTSH7111111111111111 |
| 578        | Usa VTRA                     | 31,31%                  | 21,97%    | 16,44%           | 15,05%    | 1,56%                  | 1,04%            | 0,90%                  | 0,70%                  | 0,70%                  | 0,70%                  | 0,70%                  |
| indivíduos | Parcela 5km Ind < 3.22%      | 181                     | 127       | 95               | 87        | 9                      | 6                | 5                      | 4                      | 4                      | 4                      | 4                      |
| 9          | CNH=0                        | HWHT7111                | HWHT7121  | HWHT7131         | HWHTNN111 | HWHT7141               | HWHTSH7111111111 | HWHTSH7111111111111111 | HAHT7121               | HWHT7121               | HWHTSH7111111111111111 | HWHTSH7111111111111111 |
| 337        | Usa VTRA                     | 36,80%                  | 24,04%    | 10,39%           | 5,64%     | 5,34%                  | 2,08%            | 1,78%                  | 1,19%                  | 1,19%                  | 1,19%                  | 0,89%                  |
| indivíduos | Parcela 5km Ind ≥ 3.22%      | 124                     | 81        | 35               | 19        | 18                     | 7                | 6                      | 4                      | 4                      | 4                      | 3                      |
| 20         | CNH=0; Não Usa VTRA          | HWHTNN111               | HWHT7141  | HWHT7111         | HWHT7121  | HWHT7131               | HWHTSH7111111111 | HWHTNN121              | HWHTPP111              | HWHTPP121              | HWHTPP141              | HWHTPP141              |
| 586        | Não Estuda                   | 27,30%                  | 16,55%    | 14,51%           | 13,65%    | 7,34%                  | 5,12%            | 2,73%                  | 1,71%                  | 1,20%                  | 1,02%                  | 1,02%                  |
| indivíduos | Parcela 15 km Ind < 32.31%   | 160                     | 97        | 85               | 80        | 43                     | 30               | 16                     | 10                     | 7                      | 7                      | 7                      |
| 21         | CNH=0; Não Usa VTRA          | HWHTNN111               | HWHT7111  | HWHTSH7111111111 | HWHT7121  | HWHT7131               | HWHT7141         | HAHT7111               | HAHT7121               | HWHTPP111              | HWHTPP141              | HWHTPP141              |
| 374        | Não Estuda                   | 39,84%                  | 13,90%    | 13,90%           | 9,89%     | 5,88%                  | 2,14%            | 1,34%                  | 1,34%                  | 1,34%                  | 1,34%                  | 1,34%                  |
| indivíduos | Parcela 15 km Ind ≥ 32.31%   | 149                     | 52        | 52               | 40        | 22                     | 8                | 5                      | 5                      | 5                      | 5                      | 5                      |
| 11         | CNH=0                        | HWHTSH7111111111        | HWHTNN111 | HWHTSH7111111111 | HWHT7111  | HWHT7141               | HSHTNN111        | HWHTSH7111111111       | HWHT7131               | HWHTSH7111111111111111 | HAHT7141               | HAHT7141               |
| 109        | Não Usa VTRA                 | 23,85%                  | 12,84%    | 7,34%            | 5,51%     | 5,51%                  | 4,59%            | 4,59%                  | 4,59%                  | 3,67%                  | 2,75%                  | 2,75%                  |
| indivíduos | Estuda                       | 26                      | 14        | 8                | 6         | 6                      | 5                | 5                      | 5                      | 4                      | 3                      | 3                      |
| 12         | CNH ≥ 1                      | HWHT7141                | HWHT7121  | HWHT7131         | HWHT7111  | HWHTPP111              | HWHTPP121        | HWHTPP141              | HWHTPP131              | HWHTNN111              | HAHTPP111              | HAHTPP111              |
| 451        | Não Estuda                   | 22,39%                  | 18,40%    | 14,19%           | 13,97%    | 4,66%                  | 4,44%            | 3,33%                  | 2,22%                  | 2,00%                  | 0,90%                  | 0,90%                  |
| indivíduos | Usa VTRA                     | 101                     | 83        | 64               | 63        | 21                     | 20               | 15                     | 10                     | 9                      | 4                      | 4                      |
| 52         | CNH ≥ 1; Não Estuda          | HWHTPP111               | HWHT7141  | HWHTNN111        | HWHTPP141 | HWHTPP121              | HWHT7121         | HWHT7131               | HWHTPP131              | HWHT7111               | HWHTSH7111111111111111 | HWHTSH7111111111111111 |
| 502        | Não Usa VTRA ; RF < R\$ 2970 | 12,35%                  | 11,75%    | 10,96%           | 9,76%     | 9,36%                  | 9,16%            | 8,96%                  | 5,58%                  | 5,18%                  | 3,19%                  | 3,19%                  |
| indivíduos | Parcela 10 km Ind < 15.49%   | 62                      | 59        | 55               | 49        | 47                     | 46               | 45                     | 28                     | 26                     | 16                     | 16                     |
| 53         | CNH ≥ 1; Não Estuda          | HWHTPP111               | HWHT7121  | HWHTPP121        | HWHT7111  | HWHTNN111              | HWHTPP131        | HWHTSH7111111111111111 | HWHTSH7111111111111111 | HWHT7131               | HWHT7141               | HWHT7141               |
| 411        | Não Usa VTRA ; RF < R\$ 2970 | 19,46%                  | 10,46%    | 9,73%            | 8,76%     | 8,27%                  | 7,79%            | 5,35%                  | 3,89%                  | 3,65%                  | 2,92%                  | 2,92%                  |
| indivíduos | Parcela 10 km Ind ≥ 15.49%   | 80                      | 43        | 40               | 36        | 34                     | 32               | 22                     | 16                     | 15                     | 12                     | 12                     |
| 27         | CNH ≥ 1; Não Estuda          | HWHTPP111               | HWHTPP141 | HWHTPP121        | HWHTPP131 | HWHTSH7111111111111111 | HWHTNN111        | HWHT7131               | HAHTPP111              | HWHTSH7111111111111111 | HWHT7111               | HWHT7111               |
| 473        | Não Usa VTRA ; RF ≥ R\$ 2970 | 19,24%                  | 18,60%    | 18,18%           | 9,09%     | 4,23%                  | 3,17%            | 2,54%                  | 2,11%                  | 2,11%                  | 1,69%                  | 1,69%                  |
| indivíduos |                              | 91                      | 88        | 86               | 43        | 20                     | 15               | 12                     | 10                     | 10                     | 8                      | 8                      |
| 7          | CNH ≥ 1                      | HWHTPP111               | HWHT7111  | HWHTNN111        | HWHT7141  | HWHT7121               | HSHTNN111        | HWHTPP121              | HWHTSH7111111111111111 | HWHTSH7111111111111111 | HWHT7131               | HWHT7131               |
| 263        | Estuda                       | 6,46%                   | 5,70%     | 5,32%            | 5,32%     | 4,18%                  | 3,80%            | 3,42%                  | 3,42%                  | 3,42%                  | 3,42%                  | 3,42%                  |
| indivíduos |                              | 17                      | 15        | 14               | 14        | 11                     | 10               | 9                      | 9                      | 9                      | 9                      | 9                      |

CNH = Nº provável de carteiras de habilitação no domicílio; VTRA = Vale Transporte; Parcela 5 km Ind = Parcela acumulada de empregos na indústria até 5 km do centróide da zona de residência; Parcela 15 km Ind = Parcela acumulada de empregos na indústria até 15 km do centróide da zona de residência; Parcela 10 km Ind = Parcela acumulada de empregos na indústria até 10 km do centróide da zona de residência; RF = Renda Familiar.



Os resultados da AD são relativamente fáceis de serem interpretados. Como mencionado anteriormente, a árvore é construída através de regras hierárquicas, no caso em questão perguntas e respostas. O “ramo” esquerdo da árvore corresponde sempre a respostas positivas, enquanto o “ramo” direito, àquelas negativas.

Por exemplo, a primeira pergunta para início das divisões é se não há provavelmente carteiras de habilitação no domicílio. O algoritmo então pergunta: “(CNH = 0) ?”,

- Sim, essas pessoas residem em domicílio sem nenhuma carteira de habilitação;
- Não, há pelo menos uma carteira de habilitação no domicílio.

A Figura 7.4 mostra um segmento da árvore com o caminho do nó raiz até um dos nós terminais (NÓ 8). A partir da ilustração indicada há maior facilidade em interpretar os resultados gráficos do programa.

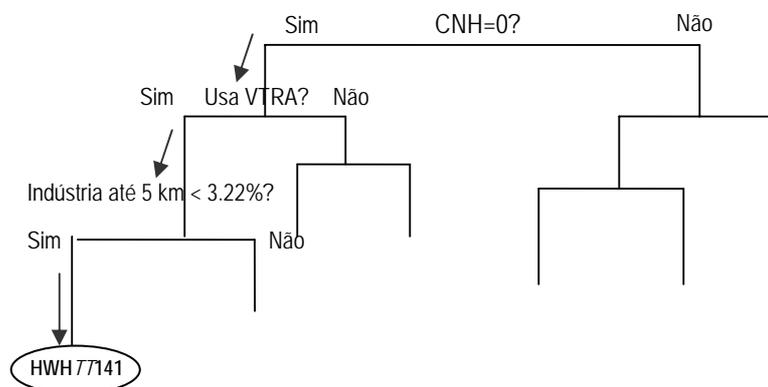


Figura 7.4 Segmento da árvore para auxílio na interpretação dos resultados.

A árvore faz sucessivas divisões até chegar à folha indicada. Este nó corresponde então ao grupo de indivíduos que residem em domicílio sem nenhuma carteira de habilitação (CNH=0), usam vale transporte (Usa VTRA) e residem em uma zona de tráfego que tem uma parcela acumulada de empregos na indústria menor que 3,22% até 5 km do seu centróide (Indústria até 5 km < 3.22%).

Vale lembrar que nem todos os indivíduos pertencentes a este grupo realizam o padrão de viagem “HWH77141”. Apesar de serem mostrados os três padrões de viagens mais frequentes no gráfico adaptado (Figura 7.3), a AD gerada pelo S-PLUS 6.1 apresenta apenas o padrão de viagem

predominante em cada um dos nós terminais (Padrão “HWH77141” – NÓ 8). As probabilidades de ocorrência de todos os outros padrões (categorias de variável dependente) são emitidas pelo relatório do *software*. O gráfico é determinístico, enquanto o relatório é probabilístico. Assim, pode-se dizer que a forma tabular dos resultados é complementar àquela gráfica.

### 7.3. TRABALHADORES NO SETOR COMERCIAL

A exemplo da amostra composta por industriários, a AD para amostra de trabalhadores no setor comercial foi gerada adotando-se os mesmo critérios: (1) desvio mínimo de 0,15; e (2) mínimo de 50 observações por nó. Após processamento da AD, foram obtidas 10 folhas e uma amostra final (excluindo dados faltantes) de 6.019 comerciários.

As variáveis independentes foram aquelas descritas na seção 7.1, sendo “PARCELA ACUMULADA DE OPORTUNIDADES” aqui definida como “PARCELA ACUMULADA DE EMPREGOS NO COMÉRCIO”.

Algumas variáveis independentes selecionadas para segregação dos dados foram as mesmas que no caso dos industriários, havendo algumas variações devido a diferenças relativas às amostras. O Capítulo 8 irá discutir diferenças obtidas entre os resultados das seis amostras.

Para o caso de trabalhadores no setor comercial a variável mais importante para segregação dos dados foi ESTUDA, dividindo a amostra total em grupos de (1) não estudantes e (2) estudantes. Assim, a primeira pergunta para segregação dos dados foi “ESTUDA = 1” (“Não Estudantes?”) A Figura 7.5 ilustra o gráfico adaptado para amostra composta por comerciários, proveniente do S-PLUS 6.1. Já a Tabela 7.2 é a síntese do relatório emitido para a mesma amostra.

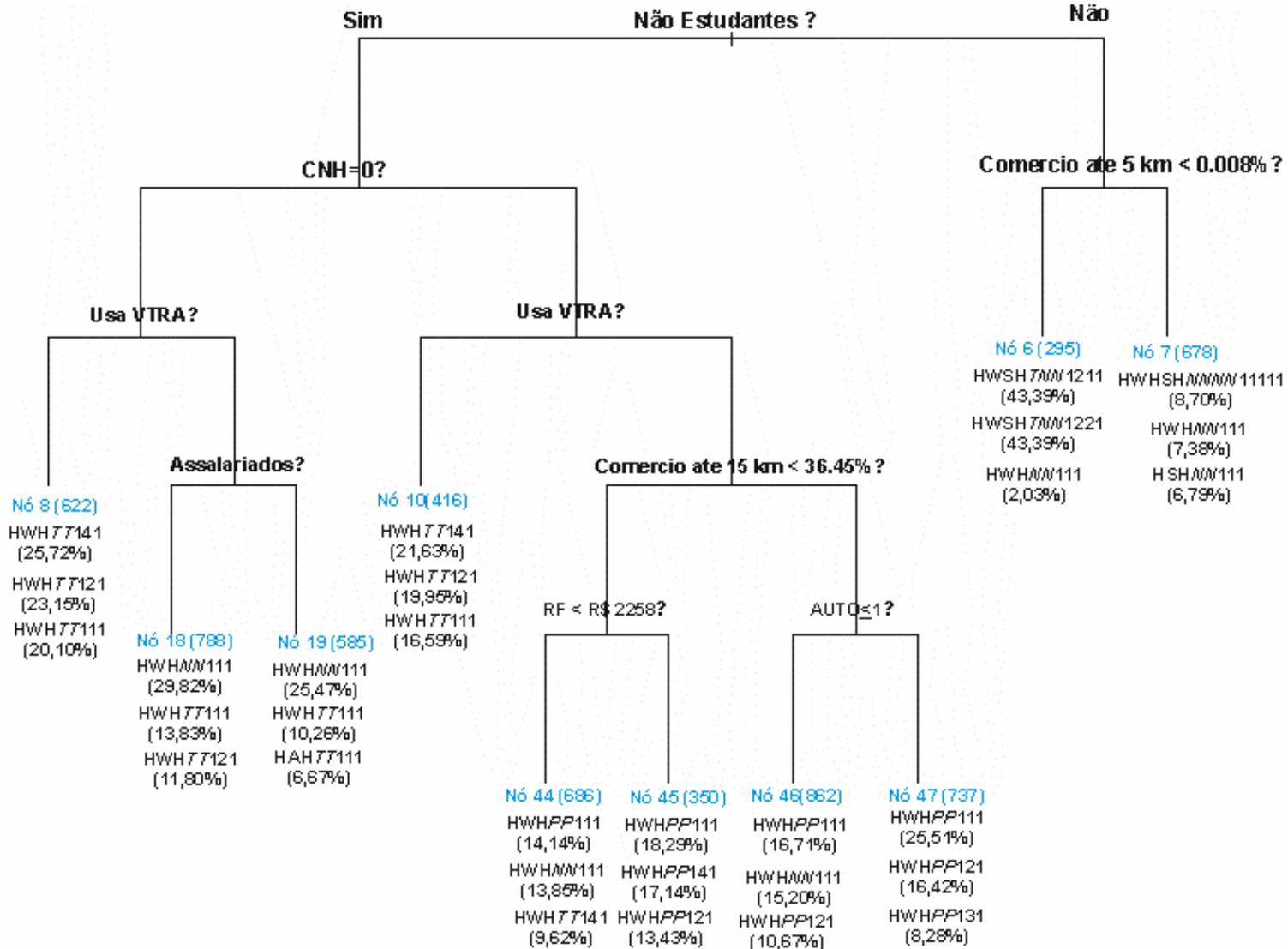


Figura 7.5 Gráfico adaptado do S-PLUS 6.1 (amostra comerciários)

LEGENDA: CNH (Nº provável de carteiras de habilitação no domicílio); VTRA (Vale Transporte); Comércio até 5 km (Parcela acumulada de empregos no comércio até 5 km do centróide da zona de residência); Comércio até 15 km (Parcela acumulada de empregos no comércio até 15 km do centróide da zona de residência); RF (Renda Familiar); AUTO (Nº de automóveis no domicílio).



Tabela 7.2 Síntese do relatório (amostra de comerciários)

| Nó                      | Características dos grupos                                                      | Padrões mais frequentes               |                                     |                                  |                                       |                                       |                                       |                                  |                                       |                                       |                                       |
|-------------------------|---------------------------------------------------------------------------------|---------------------------------------|-------------------------------------|----------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|----------------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| §<br>622<br>indivíduos  | Não Estuda<br>CNH=0<br>Usa VTRA                                                 | <b>HWH77141</b><br>25,72%<br>160      | <b>HWH77121</b><br>23,15%<br>144    | <b>HWH77111</b><br>20,10%<br>125 | <b>HWH77131</b><br>15,92%<br>99       | <b>HWHNN111</b><br>2,09%<br>13        | <b>HWAWH7NN712221</b><br>1,29%<br>§   | <b>HAH77141</b><br>1,13%<br>7    | <b>HAH77111</b><br>0,96%<br>6         | <b>HAH77121</b><br>0,80%<br>5         | <b>HWAWH7NN713331</b><br>0,64%<br>4   |
| 1§<br>7§§<br>indivíduos | Não Estuda ; CNH=0<br>Não Usa VTRA<br>Assalariados                              | <b>HWHNN111</b><br>29,82%<br>235      | <b>HWH77111</b><br>13,83%<br>109    | <b>HWH77121</b><br>11,80%<br>93  | <b>HWHWHNNNN111111</b><br>9,39%<br>74 | <b>HWH77141</b><br>9,26%<br>73        | <b>HWH77131</b><br>7,74%<br>61        | <b>HWHPP121</b><br>2,54%<br>20   | <b>HWHPP111</b><br>2,16%<br>17        | <b>HWHPP141</b><br>1,27%<br>10        | <b>HAH77111</b><br>1,14%<br>9         |
| 19<br>585<br>indivíduos | Não Estuda ; CNH=0<br>Não Usa VTRA<br>Autônomos / Não Trabalha                  | <b>HWHNN111</b><br>25,47%<br>149      | <b>HWH77111</b><br>10,26%<br>60     | <b>HAH77111</b><br>6,67%<br>39   | <b>HWH77141</b><br>6,67%<br>39        | <b>HWH77121</b><br>6,50%<br>38        | <b>HAHNN111</b><br>5,30%<br>31        | <b>HWH77131</b><br>5,30%<br>31   | <b>HWHWHNNNN111111</b><br>5,00%<br>29 | <b>HAH77121</b><br>3,76%<br>22        | <b>HWHPP111</b><br>2,91%<br>17        |
| 10<br>416<br>indivíduos | Não Estuda<br>CNH≥1<br>Usa VTRA                                                 | <b>HWH77141</b><br>21,63%<br>90       | <b>HWH77121</b><br>19,95%<br>93     | <b>HWH77111</b><br>16,59%<br>69  | <b>HWH77131</b><br>13,46%<br>56       | <b>HWHPP121</b><br>3,85%<br>16        | <b>HWHPP111</b><br>3,61%<br>15        | <b>HWHNN111</b><br>2,64%<br>11   | <b>HWHPP131</b><br>2,64%<br>11        | <b>HWHPP141</b><br>2,40%<br>10        | <b>HAHPP121</b><br>1,44%<br>6         |
| 44<br>686<br>indivíduos | Não Estuda; CNH≥1<br>Não Usa VTRA<br>Parcela 15 km Com < 36.45%<br>RF< R\$ 225§ | <b>HWHPP111</b><br>14,14%<br>97       | <b>HWHNN111</b><br>13,85%<br>95     | <b>HWH77141</b><br>9,62%<br>66   | <b>HWHPP121</b><br>7,14%<br>49        | <b>HWHPP141</b><br>6,85%<br>47        | <b>HWHPP131</b><br>5,10%<br>35        | <b>HWH77121</b><br>4,08%<br>28   | <b>HWH77131</b><br>4,08%<br>28        | <b>HWH77111</b><br>3,79%<br>26        | <b>HWHWHNNNN111111</b><br>3,64%<br>25 |
| 45<br>350<br>indivíduos | Não Estuda; CNH≥1<br>Não Usa VTRA<br>Parcela 15 km Com < 36.45%<br>RF≥ R\$ 225§ | <b>HWHPP111</b><br>18,29%<br>64       | <b>HWHPP141</b><br>17,14%<br>60     | <b>HWHPP121</b><br>13,43%<br>47  | <b>HWHNN111</b><br>10,57%<br>37       | <b>HWHPP131</b><br>8,00%<br>28        | <b>HWHWHPPPP111111</b><br>3,71%<br>13 | <b>HWH77141</b><br>3,43%<br>12   | <b>HWH77111</b><br>2,57%<br>9         | <b>HWHWHNNNN111111</b><br>2,57%<br>9  | <b>HAH77141</b><br>2,00%<br>7         |
| 46<br>862<br>indivíduos | Não Estuda; CNH≥1<br>Não Usa VTRA<br>Parcela 15 km Com ≥ 36.45%<br>AUT≤1        | <b>HWHPP111</b><br>16,71%<br>144      | <b>HWHNN111</b><br>15,20%<br>131    | <b>HWHPP121</b><br>10,67%<br>92  | <b>HWH77121</b><br>6,85%<br>59        | <b>HWHPP131</b><br>5,57%<br>48        | <b>HWH77111</b><br>5,45%<br>47        | <b>HWHPP141</b><br>4,18%<br>36   | <b>HWHWHNNNN111111</b><br>3,48%<br>30 | <b>HAHPP111</b><br>3,25%<br>28        | <b>HWH77131</b><br>2,90%<br>25        |
| 47<br>737<br>indivíduos | Não Estuda; CNH≥1<br>Não Usa VTRA<br>Parcela 15 km Com ≥ 36.45%<br>AUT>1        | <b>HWHPP111</b><br>25,51%<br>188      | <b>HWHPP121</b><br>16,42%<br>121    | <b>HWHPP131</b><br>8,28%<br>61   | <b>HWHNN111</b><br>8,14%<br>60        | <b>HWHWHPPPP111111</b><br>7,19%<br>53 | <b>HWHPP141</b><br>4,89%<br>36        | <b>HAHPP111</b><br>2,71%<br>20   | <b>HWH77121</b><br>2,58%<br>19        | <b>HWHWHNNNN111111</b><br>2,44%<br>18 | <b>HWH77111</b><br>2,04%<br>15        |
| 6<br>295<br>indivíduos  | Estuda<br>Parcela 5 km Com < 0.00§§                                             | <b>HWSH7NN1211</b><br>43,39%<br>128   | <b>HWSH7NN1221</b><br>43,39%<br>128 | <b>HWHNN111</b><br>2,03%<br>6    | <b>SHHNN111</b><br>1,70%<br>5         | <b>HWHSHNNNN111111</b><br>1,02%<br>3  | <b>HWHSHNNNT111111</b><br>1,02%<br>3  | <b>HWSH7771441</b><br>1,02%<br>3 | <b>SHH77121</b><br>0,68%<br>2         | <b>HWHPP141</b><br>0,68%<br>2         | <b>HWH77111</b><br>0,68%<br>2         |
| 7<br>678<br>indivíduos  | Estuda<br>Parcela 5 km Com ≥ 0.00§§                                             | <b>HWHSHNNNN111111</b><br>8,70%<br>59 | <b>HWHNN111</b><br>7,38%<br>50      | <b>SHHNN111</b><br>6,79%<br>46   | <b>HWAH7NT111111</b><br>4,72%<br>32   | <b>HWH77121</b><br>4,72%<br>32        | <b>HWHWH7NT111111</b><br>4,72%<br>32  | <b>HWH77111</b><br>4,43%<br>30   | <b>HWHSH77NN111111</b><br>3,69%<br>25 | <b>SHHWHNNNN111111</b><br>3,54%<br>24 | <b>HWHPP111</b><br>3,54%<br>24        |

CNH = Nº provável de carteiras de habilitação no domicílio; VTRA = Vale transporte; Parcela 15 km Com = Parcela acumulada de empregos no comércio até 15 km do centróide da zona de residência; Parcela 5 km Com = Parcela acumulada de empregos no comércio até 5 km do centróide da zona de residência.



#### 7.4. TRABALHADORES NO SETOR DE SERVIÇOS

A AD para amostra de trabalhadores no setor de serviços foi gerada adotando-se os mesmos critérios para as amostras anteriores. Contudo, o conjunto de variáveis de uso do solo, neste caso, corresponde às variáveis “PARCELA ACUMULADA DE EMPREGOS NO SETOR DE SERVIÇOS” nas quatro faixas de distâncias consideradas. Foram obtidas também 10 folhas e a amostra final, após exclusão de alguns dados faltantes, foi composta de 8.286 trabalhadores no setor de serviços.

A árvore gerada para amostra de trabalhadores no setor de serviços foi bastante similar àquela gerada para amostra de industriários. Foram observados resultados que podem ser explicados ou justificados com base nas características inerentes ao setor econômico. Espera-se, por exemplo, que industriários tenham um comportamento subjacente ao encadeamento de viagens particular, devido a características do referido setor econômico e também da concentração de tais atividades da RMSP. Esse tipo de discussão será descrito no capítulo de análises de resultados (Capítulo 8). O presente capítulo tem como finalidade apresentar os resultados obtidos com o processamento da AD para cada amostra.

Para o caso de trabalhadores no setor de serviços a variável mais importante para segregação dos dados foi NÚMERO (PROVÁVEL) DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO. A exemplo da AD obtida da amostra de industriários, os dados foram segregados em duas classes principais de indivíduos:

- (1) Pessoas que residem em domicílios onde não há nenhuma carteira nacional de habilitação (CNH=0);
- (2) Indivíduos com pelo menos um membro no domicílio com carteira nacional de habilitação (CNH  $\geq$ 1).

A Figura 7.6 ilustra o gráfico adaptado gerado pelo S-PLUS 6.1 para amostra composta por trabalhadores no setor de serviços. Já a Tabela 7.3 apresenta a síntese do relatório gerado.



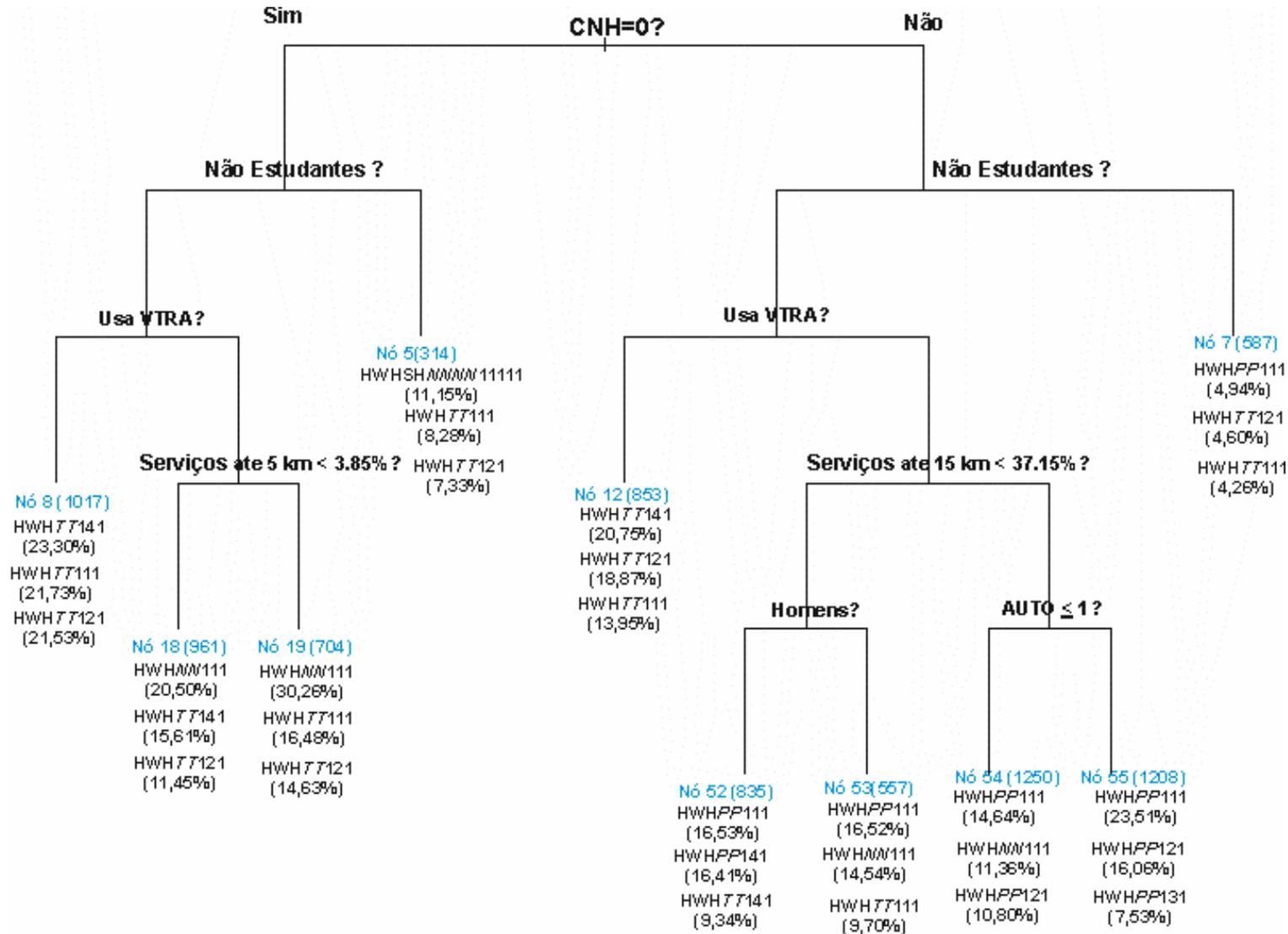


Figura 7.6 Gráfico adaptado do S-PLUS 6.1 (amostra de trabalhadores no setor de serviços)

LEGENDA: CNH (Nº provável de carteiras de habilitação no domicílio); VTRA (Vale Transporte); Serviços até 5 km (Parcela acumulada de empregos no setor de serviços até 5 km do centróide da zona de residência); Serviços até 15 km (Parcela acumulada de empregos no setor de serviços até 15 km do centróide da zona de residência); AUTO (Nº de automóveis no domicílio).



Tabela 7.3 Síntese do relatório (amostra de trabalhadores no setor de serviços)

| Nó         | Características dos grupos             | Padrões mais frequentes |                 |                 |                 |                       |                 |                       |                      |                       |                       |
|------------|----------------------------------------|-------------------------|-----------------|-----------------|-----------------|-----------------------|-----------------|-----------------------|----------------------|-----------------------|-----------------------|
| 8          | CNH=0                                  | <b>HWH77141</b>         | <b>HWH77111</b> | <b>HWH77121</b> | <b>HWH77131</b> | <b>HWHNN111</b>       | <b>HAH77111</b> | <b>HAH77121</b>       | <b>HAH77141</b>      | <b>HWAWHTNNT12221</b> | <b>HAHNN111</b>       |
| 1017       | Não Estuda                             | 23,30%                  | 21,73%          | 21,53%          | 16,32%          | 3,64%                 | 1,48%           | 1,48%                 | 0,90%                | 0,69%                 | 0,59%                 |
| indivíduos | Usa VTRA                               | 237                     | 221             | 219             | 166             | 37                    | 15              | 15                    | 9                    | 7                     | 6                     |
| 18         | CNH=0, Não Estuda                      | <b>HWHNN111</b>         | <b>HWH77141</b> | <b>HWH77121</b> | <b>HWH77111</b> | <b>HWH77131</b>       | <b>HWHPP111</b> | <b>HWHWHNNNN11111</b> | <b>HAH77111</b>      | <b>HAHNN111</b>       | <b>HWHPP141</b>       |
| 961        | Não Usa VTRA                           | 20,50%                  | 15,61%          | 11,45%          | 11,24%          | 8,32%                 | 4,06%           | 4,06%                 | 2,91%                | 2,39%                 | 2,39%                 |
| indivíduos | Parcela 5 km Serv <3.85%               | 197                     | 150             | 110             | 108             | 80                    | 39              | 39                    | 28                   | 23                    | 23                    |
| 19         | CNH=0, Não Estuda                      | <b>HWHNN111</b>         | <b>HWH77111</b> | <b>HWH77121</b> | <b>HWH77131</b> | <b>HWHWHNNNN11111</b> | <b>HAH77111</b> | <b>HWHPP111</b>       | <b>HWHPP121</b>      | <b>HAHNN111</b>       | <b>HAH77121</b>       |
| 704        | Não Usa VTRA                           | 30,26%                  | 16,48%          | 14,63%          | 6,96%           | 4,69%                 | 3,41%           | 2,70%                 | 2,27%                | 2,13%                 | 1,42%                 |
| indivíduos | Parcela 5 km Serv ≥3.85%               | 213                     | 116             | 103             | 49              | 33                    | 24              | 19                    | 16                   | 15                    | 10                    |
| 5          | CNH=0                                  | <b>HWSHNNNN11111</b>    | <b>HWH77111</b> | <b>HWH77121</b> | <b>HWHNN111</b> | <b>SHNN111</b>        | <b>HWH77141</b> | <b>HWH77131</b>       | <b>HWSH77NN12111</b> | <b>HWSH771111</b>     | <b>HWSH77NN13111</b>  |
| 314        | Estuda                                 | 11,15%                  | 8,28%           | 7,33%           | 6,69%           | 5,73%                 | 5,73%           | 3,50%                 | 3,19%                | 3,19%                 | 2,55%                 |
| indivíduos |                                        | 35                      | 26              | 23              | 21              | 18                    | 18              | 11                    | 10                   | 10                    | 8                     |
| 12         | CNH≥1                                  | <b>HWH77141</b>         | <b>HWH77121</b> | <b>HWH77111</b> | <b>HWH77131</b> | <b>HWHPP111</b>       | <b>HWHNN111</b> | <b>HWHPP131</b>       | <b>HWHPP121</b>      | <b>HAHPP111</b>       | <b>HWHPP141</b>       |
| 853        | Não Estuda                             | 20,75%                  | 18,87%          | 13,96%          | 13,13%          | 4,10%                 | 3,87%           | 2,70%                 | 2,34%                | 1,29%                 | 1,29%                 |
| indivíduos | Usa VTRA                               | 177                     | 161             | 119             | 112             | 35                    | 33              | 23                    | 20                   | 11                    | 11                    |
| 52         | CNH≥1; Não Estuda                      | <b>HWHPP111</b>         | <b>HWHPP141</b> | <b>HWH77141</b> | <b>HWHPP121</b> | <b>HWH77111</b>       | <b>HWHPP131</b> | <b>HWHNN111</b>       | <b>HWH77131</b>      | <b>HWH77121</b>       | <b>HWHWHPPPP11111</b> |
| 835        | Não Usa VTRA                           | 16,53%                  | 16,41%          | 9,34%           | 7,19%           | 7,19%                 | 5,75%           | 5,27%                 | 4,55%                | 3,83%                 | 2,99%                 |
| indivíduos | Parcela 15 km Serv <37.15%<br>Homens   | 138                     | 137             | 78              | 60              | 60                    | 48              | 44                    | 38                   | 32                    | 25                    |
| 53         | CNH≥1; Não Estuda                      | <b>HWHPP111</b>         | <b>HWHNN111</b> | <b>HWH77111</b> | <b>HWH77141</b> | <b>HWH77121</b>       | <b>HWHPP141</b> | <b>HWH77131</b>       | <b>HWHPP121</b>      | <b>HAHPP111</b>       | <b>HWHPP131</b>       |
| 557        | Não Usa VTRA                           | 16,52%                  | 14,54%          | 9,70%           | 8,44%           | 4,85%                 | 4,67%           | 4,49%                 | 4,31%                | 3,23%                 | 2,51%                 |
| indivíduos | Parcela 15 km Serv <37.15%<br>Mulheres | 92                      | 81              | 54              | 47              | 27                    | 26              | 25                    | 24                   | 18                    | 14                    |
| 54         | CNH≥1; Não Estuda                      | <b>HWHPP111</b>         | <b>HWHNN111</b> | <b>HWHPP121</b> | <b>HWH77121</b> | <b>HWH77111</b>       | <b>HWH77131</b> | <b>HWHPP131</b>       | <b>HAHPP111</b>      | <b>HWHWHPPPP11111</b> | <b>HWH77141</b>       |
| 1250       | Não Usa VTRA                           | 14,64%                  | 11,36%          | 10,80%          | 9,44%           | 7,92%                 | 4,96%           | 3,12%                 | 2,96%                | 2,72%                 | 2,32%                 |
| indivíduos | Parcela 15 km Serv ≥37.15%<br>AUTO = 1 | 183                     | 142             | 135             | 118             | 99                    | 62              | 39                    | 37                   | 34                    | 29                    |
| 55         | CNH≥1; Não Estuda                      | <b>HWHPP111</b>         | <b>HWHPP121</b> | <b>HWHPP131</b> | <b>HWHPP141</b> | <b>HWHWHPPPP11111</b> | <b>HWHNN111</b> | <b>HWH77121</b>       | <b>HWH77111</b>      | <b>HAHPP111</b>       | <b>HWH77131</b>       |
| 1208       | Não Usa VTRA                           | 23,51%                  | 16,06%          | 7,53%           | 4,55%           | 4,31%                 | 4,22%           | 3,81%                 | 3,15%                | 2,48%                 | 2,40%                 |
| indivíduos | Parcela 15 km Serv ≥37.15%<br>AUTO ≥ 2 | 284                     | 194             | 91              | 55              | 52                    | 51              | 46                    | 38                   | 30                    | 29                    |
| 7          | CNH≥1                                  | <b>HWHPP111</b>         | <b>HWH77121</b> | <b>HWH77111</b> | <b>SHNN111</b>  | <b>HWHNN111</b>       | <b>HWH77131</b> | <b>HWHPP121</b>       | <b>HWHPP131</b>      | <b>HWH77141</b>       | <b>HWSHPPPP11111</b>  |
| 587        | Estuda                                 | 4,94%                   | 4,60%           | 4,26%           | 3,92%           | 3,92%                 | 3,58%           | 3,24%                 | 3,07%                | 3,07%                 | 2,56%                 |
| indivíduos |                                        | 29                      | 27              | 25              | 23              | 23                    | 21              | 19                    | 18                   | 18                    | 15                    |

CNH= N<sup>o</sup> provável de carteiras de habilitação no domicílio; VTRA = Vale transporte; Parcela 5 km Serv = Parcela acumulada de empregos no setor de serviços até 5 km do centróide da zona de residência; Parcela 15 km Serv = Parcela acumulada de empregos no setor de serviços até 15 km do centróide da zona de residência; AUTO = N<sup>o</sup> de automóveis no domicílio.

### 7.5. ESTUDANTES ATÉ O 1º GRAU

Conforme mencionado no Capítulo 4, a amostra composta por estudantes foi dividida em três grupos, considerando o grau de instrução: (1) até o 1º grau; (2) no 2º grau; e (3) universitários. O algoritmo CART considerou para as três amostras variáveis similares para segregação dos dados (IDADE, RENDA FAMILIAR, NÚMERO DE AUTOMÓVEIS NO DOMICÍLIO, PARCELA ACUMULADA DE MATRÍCULAS, etc.). No entanto, em cada caso, há diferenças no comportamento relativo a viagens que dependem de cada tipo de amostra de estudantes.

A AD para amostra de estudantes matriculados até o 1º grau foi gerada adotando-se os mesmos critérios mencionados anteriormente. As variáveis de uso do solo foram as “PARCELAS ACUMULADAS DE MATRÍCULAS ATÉ O 1º GRAU” nas distâncias consideradas a partir do centróide da zona de origem. Foram obtidos 9 nós terminais e a amostra final, após exclusão de alguns dados faltantes, foi composta de 13.733 estudantes até o 1º grau.

Através dos resultados, observa-se que a variável mais importante para segregação dos dados foi RENDA FAMILIAR. A presente amostra foi dividida em dois grupos:

- (1) Indivíduos com Renda Familiar inferior a R\$ 1.690;
- (2) Indivíduos com Renda Familiar superior ou igual a R\$ 1.690.

A Figura 7.7 apresenta a AD gerada pelo *software* e adaptada pela autora. A Tabela 7.4, em seguida, traz a síntese do relatório emitido.



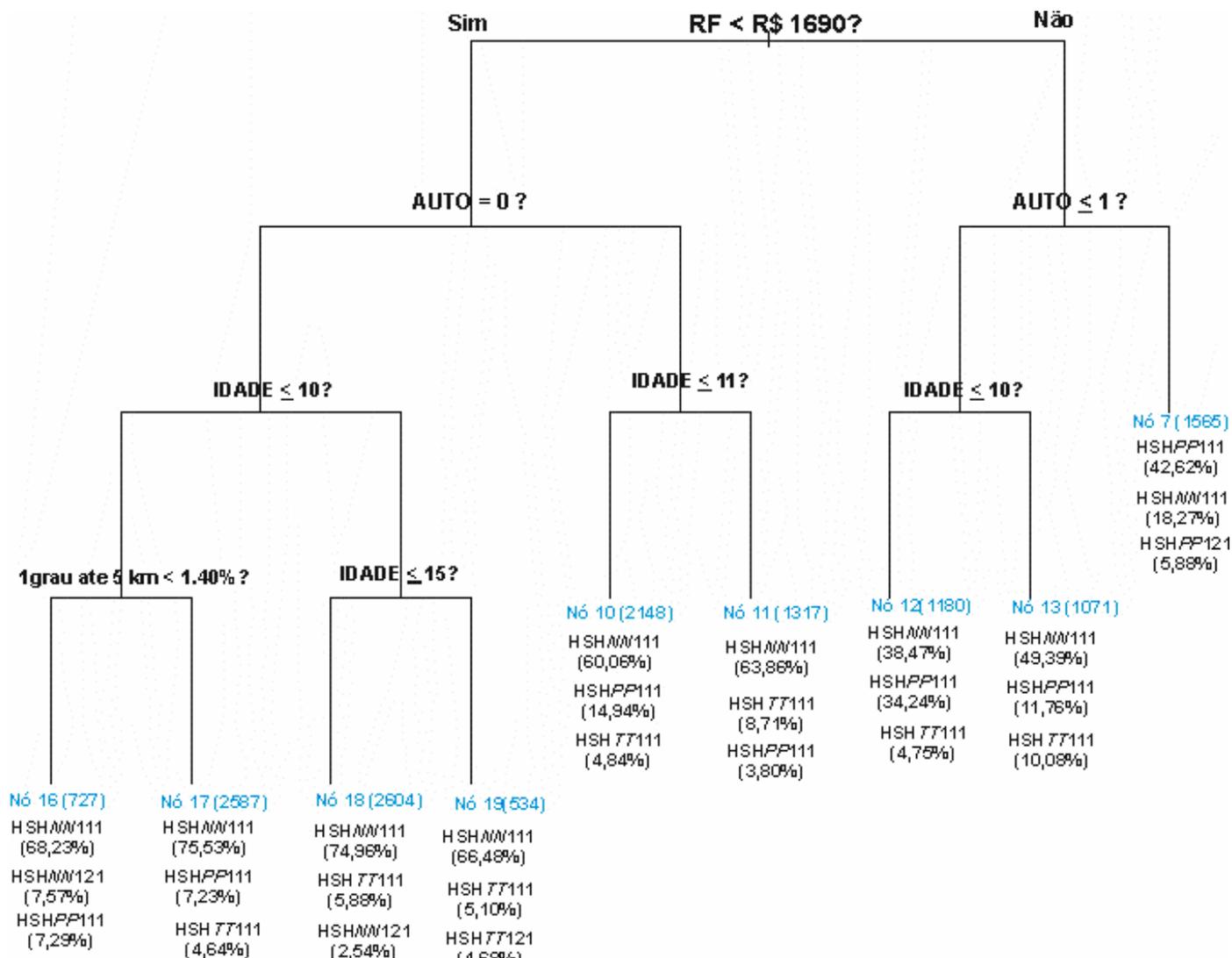


Figura 7.7 Gráfico adaptado do S-PLUS 6.1 (amostra de estudantes até 1º grau)

LEGENDA: RF (Renda Familiar); AUTO (Nº de automóveis no domicílio); 1º grau até 5 km (Parcela acumulada de matrículas escolares até o 1º grau até 5 km do centróide da zona de residência).



Tabela 7.4 Síntese do relatório (amostra estudantes até o 1º grau)

| Nó         | Características dos grupos  | Padrões mais frequentes |          |          |                |                |             |                |                |             |                |
|------------|-----------------------------|-------------------------|----------|----------|----------------|----------------|-------------|----------------|----------------|-------------|----------------|
|            |                             | HSHNN111                | HSHNN121 | HSHPP111 | HSH77111       | HSHPP121       | HSH77121    | HSHNN141       | HAHNN111       | HAH77141    | HSHNN131       |
| 16         | RF < R\$1690; AUT=0         | HSHNN111                | HSHNN121 | HSHPP111 | HSH77111       | HSHPP121       | HSH77121    | HSHNN141       | HAHNN111       | HAH77141    | HSHNN131       |
| 727        | Idade ≤ 10                  | 68,23%                  | 7,57%    | 7,29%    | 6,19%          | 1,65%          | 1,51%       | 1,10%          | 0,83%          | 0,55%       | 0,55%          |
| individuos | Parcela 5 km 1 grau < 1.40% | 496                     | 55       | 53       | 45             | 12             | 11          | 8              | 6              | 4           | 4              |
| 17         | RF < R\$1690; AUT=0         | HSHNN111                | HSHPP111 | HSH77111 | HAHNN111       | HSHNN121       | HSHNNNN1111 | HSH77121       | HSHAHNNNN11111 | HSHPN111    | HSHSHNNNN11111 |
| 2587       | Idade ≤ 10                  | 75,53%                  | 7,23%    | 4,64%    | 1,12%          | 0,97%          | 0,89%       | 0,77%          | 0,73%          | 0,73%       | 0,66%          |
| individuos | Parcela 5 km 1 grau ≥ 1.40% | 1954                    | 187      | 120      | 29             | 25             | 23          | 20             | 19             | 19          | 17             |
| 18         | RF < R\$1690; AUT=0         | HSHNN111                | HSH77111 | HSHNN121 | HSHSHNNNN11111 | HSH77121       | HAHNN111    | HSHAHNNNN11111 | HSHPP111       | HAH77111    | HSH7N111       |
| 2604       | 10 < Idade ≤ 15             | 74,96%                  | 5,88%    | 2,54%    | 2,04%          | 1,61%          | 1,31%       | 1,31%          | 1,15%          | 0,73%       | 0,65%          |
| individuos |                             | 1952                    | 153      | 66       | 53             | 42             | 34          | 34             | 30             | 19          | 17             |
| 19         | RF < R\$1690; AUT=0         | HSHNN111                | HSH77111 | HSH77121 | HSHNN121       | HSHSHNNNN11111 | HSH77131    | HAHNN111       | HAHSHNNNN11111 | HAH77111    | HAH77121       |
| 534        | Idade > 15                  | 66,48%                  | 5,10%    | 4,68%    | 3,00%          | 2,62%          | 2,06%       | 1,87%          | 1,69%          | 1,69%       | 1,12%          |
| individuos |                             | 355                     | 27       | 25       | 16             | 14             | 11          | 10             | 9              | 9           | 6              |
| 10         | RF < R\$1690;               | HSHNN111                | HSHPP111 | HSH77111 | HSHPN111       | HSHPP121       | HSHNN121    | HAHPP111       | HAHNN111       | HSHPT111    | HSH77121       |
| 2148       | AUT ≥ 1                     | 60,06%                  | 14,94%   | 4,84%    | 2,10%          | 1,72%          | 1,26%       | 0,88%          | 0,79%          | 0,75%       | 0,75%          |
| individuos | Idade ≤ 11                  | 1290                    | 321      | 104      | 45             | 37             | 27          | 19             | 17             | 16          | 16             |
| 11         | RF < R\$1690;               | HSHNN111                | HSH77111 | HSHPP111 | HSH77121       | HSHPN111       | HAHNN111    | HSHNN121       | HSHSHNNNN11111 | HSH77131    | HSHPT111       |
| 1317       | AUT ≥ 1                     | 63,86%                  | 8,71%    | 3,80%    | 2,85%          | 1,83%          | 1,68%       | 1,46%          | 1,31%          | 0,95%       | 0,88%          |
| individuos | Idade > 11                  | 841                     | 115      | 50       | 38             | 24             | 22          | 19             | 17             | 13          | 12             |
| 12         | RF ≥ R\$1690;               | HSHNN111                | HSHPP111 | HSH77111 | HSHPN111       | HSHPP121       | HSAHPPP1111 | HSHNP111       | HSHPT111       | HSAHNNN1111 | HSH77121       |
| 1180       | AUT ≤ 1                     | 38,47%                  | 34,24%   | 4,75%    | 3,31%          | 2,97%          | 0,93%       | 0,85%          | 0,76%          | 0,68%       | 0,68%          |
| individuos | Idade ≤ 10                  | 454                     | 404      | 56       | 39             | 35             | 11          | 10             | 9              | 8           | 8              |
| 13         | RF ≥ R\$1690;               | HSHNN111                | HSHPP111 | HSH77111 | HSHPT111       | HSH77121       | HSHPN111    | HSHSHNNNN11111 | HSHNN121       | HSHPP121    | HSHAHNNNN11111 |
| 1071       | AUT ≤ 1                     | 49,39%                  | 11,76%   | 10,08%   | 2,61%          | 2,52%          | 1,87%       | 1,87%          | 1,49%          | 1,49%       | 1,21%          |
| individuos | Idade > 10                  | 529                     | 126      | 108      | 28             | 27             | 20          | 20             | 16             | 16          | 13             |
| 7          | RF ≥ R\$1690;               | HSHPP111                | HSHNN111 | HSHPP121 | HSH77111       | HSHPN111       | HSHPT111    | HSHPP141       | HSHSHPPP11111  | HAHPP111    | HSHPP131       |
| 1565       | AUT ≥ 2                     | 42,62%                  | 18,27%   | 5,88%    | 4,47%          | 3,00%          | 2,49%       | 1,66%          | 1,53%          | 1,28%       | 1,28%          |
| individuos |                             | 667                     | 286      | 92       | 70             | 47             | 39          | 26             | 24             | 20          | 20             |

RF = Renda Familiar; Parcela 5 km 1 grau = Parcela acumulada de matrículas até o 1º grau até 5 km do centróide da zona de residência; AUT = Nº de automóveis no domicílio.

## 7.6. ESTUDANTES NO 2º GRAU

A AD para amostra de estudantes matriculados no 2º grau foi gerada adotando-se as variáveis descritas na seção 7.1, sendo “OPORTUNIDADE ACUMULADA” definida como “PARCELA ACUMULADA DE MATRÍCULAS ESCOLARES NO 2º GRAU”. Os critérios para segregação dos dados (desvio mínimo e mínimo de observações por nó) foram os mesmos adotados para as demais amostras.

Foram obtidos 8 nós terminais e a amostra final, omitindo dados faltantes, foi composta de 2.125 estudantes matriculados no 2º grau.

Assim como o caso de estudantes matriculados até o 1º grau, observa-se que a variável mais importante para segregação dos dados foi RENDA FAMILIAR. Os dados foram divididos em duas classes de indivíduos:

- (1) Indivíduos com Renda Familiar inferior a R\$ 2.320;
- (2) Indivíduos com Renda Familiar superior ou igual a R\$ 2.320.

A Figura 7.8 e a Tabela 7.5 apresentam o gráfico adaptado e a síntese do relatório gerado, respectivamente.



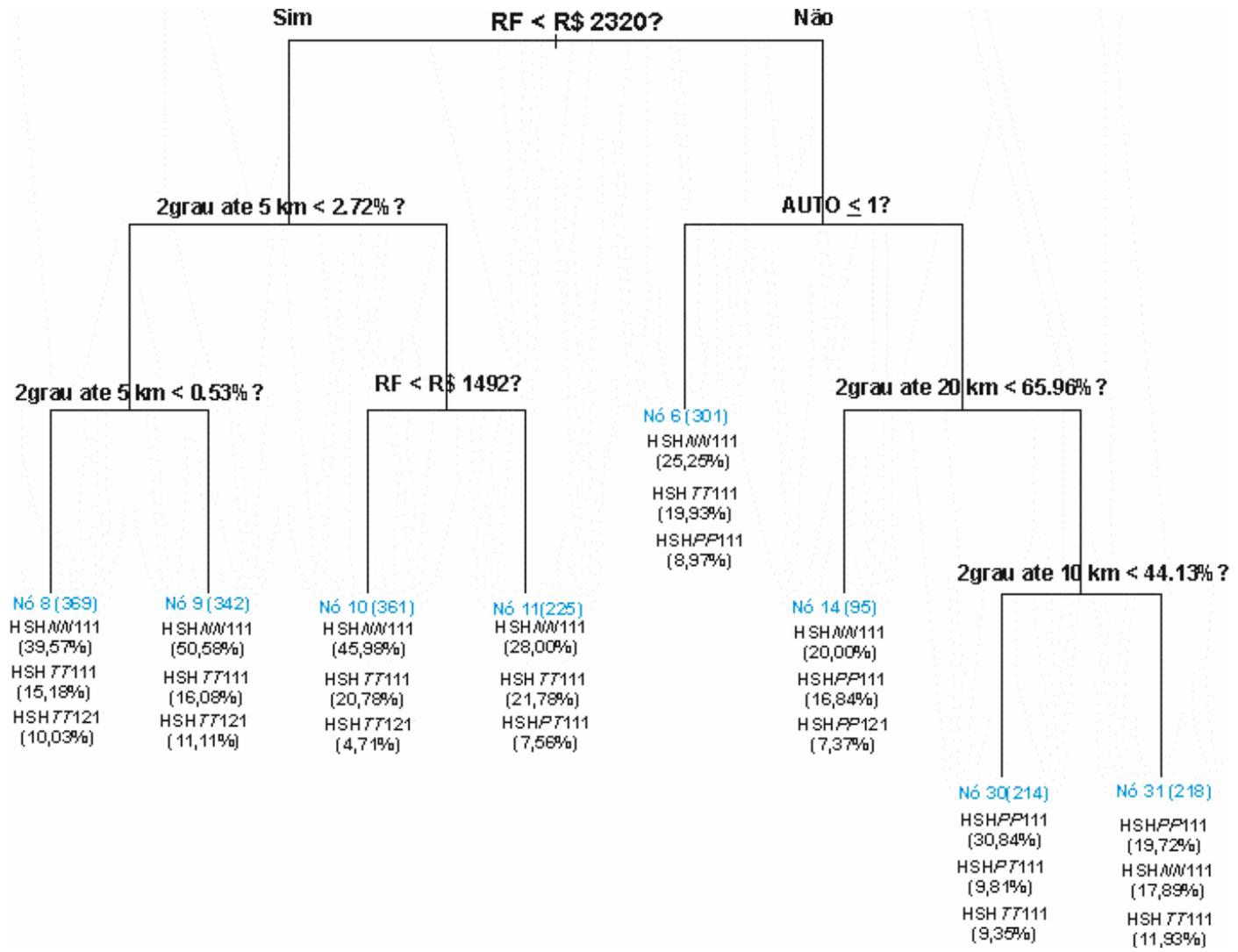


Figura 7.8 Gráfico adaptado do S-PLUS 6.1 (amostra de estudantes no 2º grau)

LEGENDA: RF (Renda Familiar); AUTO (Nº de automóveis no domicílio); 2º grau até 5 km (Parcela acumulada de matrículas escolares no 2º grau até 5 km do centróide da zona de residência); 2º grau até 10 km (Parcela acumulada de matrículas escolares no 2º grau até 10 km do centróide da zona de residência); 2º grau até 20 km (Parcela acumulada de matrículas escolares no 2º grau até 20 km do centróide da zona de residência).



Tabela 7.5 Síntese do relatório (amostra estudantes no 2º grau)

| Nó                      | Características dos grupos                                              | Padrões mais frequentes |          |          |          |          |          |          |          |          |          |
|-------------------------|-------------------------------------------------------------------------|-------------------------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
|                         |                                                                         | HSHNN111                | HSH77111 | HSH77121 | HSH77141 | HSH77131 | HSHNN121 | HAH77141 | HSHPP111 | HAHNN111 | HSHNN131 |
| 8<br>369<br>indivíduos  | RF < R\$2320                                                            | 39,57%                  | 15,18%   | 10,03%   | 5,96%    | 5,15%    | 4,07%    | 2,17%    | 2,17%    | 1,08%    | 1,08%    |
|                         | Parcela 5 km 2grau < 0.53%                                              | 146                     | 56       | 37       | 22       | 19       | 15       | 8        | 8        | 4        | 4        |
| 9<br>342<br>indivíduos  | RF < R\$2320                                                            | 50,58%                  | 16,08%   | 11,11%   | 2,92%    | 2,34%    | 1,46%    | 0,87%    | 0,87%    | 0,87%    | 0,58%    |
|                         | Parcela 5 km 2grau < 2.72%<br>Parcela 5 km 2grau ≥ 0.53%                | 173                     | 55       | 38       | 10       | 8        | 5        | 3        | 3        | 3        | 2        |
| 10<br>361<br>indivíduos | RF < R\$1492                                                            | 45,98%                  | 20,78%   | 4,71%    | 3,60%    | 2,22%    | 1,66%    | 1,66%    | 1,66%    | 1,39%    | 1,11%    |
|                         | Parcela 5 km 2grau ≥ 2.72%                                              | 166                     | 75       | 17       | 13       | 8        | 6        | 6        | 6        | 5        | 4        |
| 11<br>225<br>indivíduos | RF < R\$2320                                                            | 28,00%                  | 21,78%   | 7,56%    | 6,22%    | 4,89%    | 2,22%    | 1,78%    | 1,78%    | 1,33%    | 1,33%    |
|                         | RF ≥ R\$1492<br>Parcela 5 km 2grau ≥ 2.72%                              | 63                      | 49       | 17       | 14       | 11       | 5        | 4        | 4        | 3        | 3        |
| 6<br>301<br>indivíduos  | RF ≥ R\$2320                                                            | 25,25%                  | 19,93%   | 8,97%    | 8,97%    | 5,65%    | 3,65%    | 3,65%    | 1,99%    | 1,66%    | 1,33%    |
|                         | AUT ≤ 1                                                                 | 76                      | 60       | 27       | 27       | 17       | 11       | 11       | 6        | 5        | 4        |
| 14<br>95<br>indivíduos  | RF ≥ R\$2320                                                            | 20,00%                  | 16,84%   | 7,37%    | 5,26%    | 4,21%    | 4,21%    | 3,16%    | 3,16%    | 2,11%    | 2,11%    |
|                         | AUT ≥ 2<br>Parcela 20 km 2grau < 65.96%                                 | 19                      | 16       | 7        | 5        | 4        | 4        | 3        | 3        | 2        | 2        |
| 30<br>214<br>indivíduos | RF ≥ R\$2320                                                            | 30,84%                  | 9,81%    | 9,35%    | 8,88%    | 4,67%    | 4,67%    | 4,21%    | 3,74%    | 3,27%    | 2,34%    |
|                         | AUT ≥ 2<br>Parcela 20 km 2grau ≥ 65.96%<br>Parcela 10 km 2grau < 44.13% | 29                      | 9        | 9        | 8        | 4        | 4        | 4        | 4        | 3        | 2        |
| 31<br>218<br>indivíduos | RF ≥ R\$2320                                                            | 19,72%                  | 17,89%   | 11,93%   | 9,63%    | 6,88%    | 3,67%    | 3,21%    | 2,29%    | 1,84%    | 1,84%    |
|                         | AUT ≥ 2<br>Parcela 20 km 2grau ≥ 65.96%<br>Parcela 10 km 2grau ≥ 44.13% | 43                      | 39       | 26       | 21       | 15       | 8        | 7        | 5        | 4        | 4        |

RF = Renda Familiar; Parcela 5 km 2grau = Parcela acumulada de matrículas no 2º grau até 5 km do centróide da zona de residência; AUT = Nº de automóveis no domicílio; Parcela 20 km 2grau = Parcela acumulada de matrículas no 2º grau até 20 km do centróide da zona de residência; Parcela 10 km 2grau = Parcela acumulada de matrículas no 2º grau até 10 km do centróide da zona de residência.

### 7.7. ESTUDANTES UNIVERSITÁRIOS

Para processamento da AD, considerando a amostra de universitários, foram adotados os mesmos critérios mencionados nas seções anteriores. O grupo de variáveis de uso do solo neste caso é definido como “PARCELA ACUMULADA DE MATRÍCULAS EM UNIVERSIDADES” (variando de acordo com a distância em relação ao centróide da zona de residência).

Foram observadas 8 folhas e a amostra final, após aplicação da AD foi composta de 560 estudantes universitários.

A variável escolhida para divisão dos dados a partir do nó raiz foi PARCELA ACUMULADA DE MATRÍCULAS EM UNIVERSIDADES ATÉ 10 km DA ZONA DA RESIDÊNCIA. Os dados foram divididos em dois grupos iniciais:

- (1) Indivíduos que residem em zonas de tráfego com Parcela Acumulada de Universidades até 10 km a partir do centróide inferior a 44.72%; e
- (2) Indivíduos que residem em zonas de tráfego com Parcela Acumulada de Universidade até 10 km a partir do centróide superior ou igual a 44.72%.

Nota-se que o caso específico dos universitários foi o único em que se observou uma das quatro variáveis de uso do solo como a primeira considerada para divisão da amostra pelo algoritmo CART. O fato reflete, principalmente, a provável grande influência do uso do solo nesses casos, devido, essencialmente, à concentração dessas atividades em localizações restritas na RMSP.

A Figura 7.9 e a Tabela 7.6 apresentam respectivamente a AD adaptada e o relatório sintetizado gerados pelo S-PLUS 6.1.



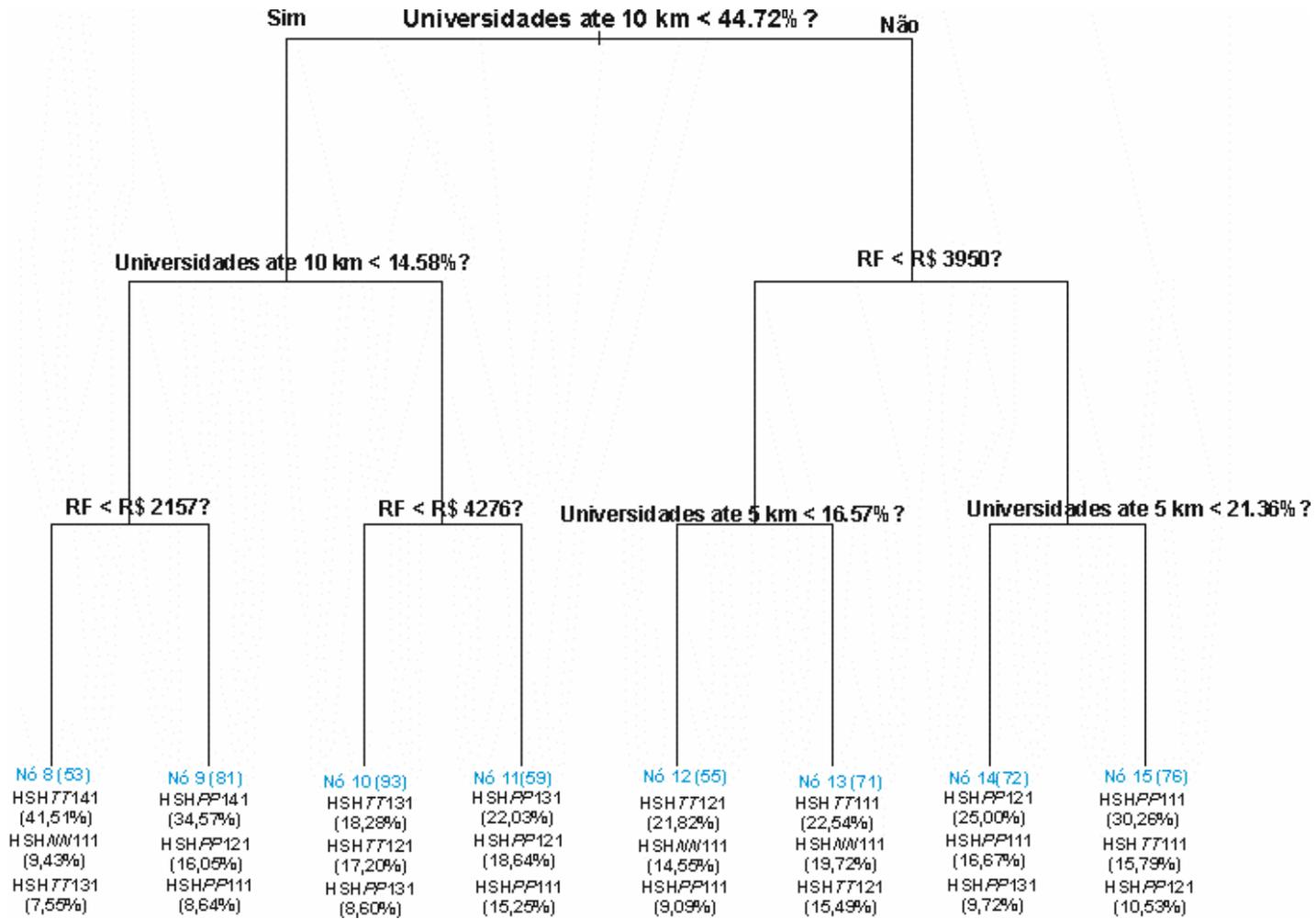


Figura 7.9 Gráfico adaptado do S-PLUS 6.1 (amostra de estudantes universitários)

LEGENDA: RF (Renda Familiar); Universidades até 5 km (Parcela acumulada de matrículas escolares em universidades até 5 km do centróide da zona de residência); Universidades até 10 km (Parcela acumulada de matrículas escolares em universidades até 10 km do centróide da zona de residência).



Tabela 7.6 Síntese do relatório (amostra de estudantes universitários)

| Nó         | Características dos grupos | Padrões mais frequentes |                 |                 |                     |                    |                       |                       |                    |                       |                       |
|------------|----------------------------|-------------------------|-----------------|-----------------|---------------------|--------------------|-----------------------|-----------------------|--------------------|-----------------------|-----------------------|
| 8          | RF < R\$2157               | <b>HSH7T141</b>         | <b>HSHNN111</b> | <b>HSHTT131</b> | <b>HSHPP121</b>     | <b>HSH7T111</b>    | <b>HSHPP111</b>       | <b>HSHPP131</b>       | <b>HAH7TT1131</b>  | <b>HAHNN111</b>       | <b>HAHSH7TNN11111</b> |
| 53         | Parcela 10 km Uni <14.58%  | 41,51%                  | 9,43%           | 7,55%           | 5,66%               | 5,66%              | 3,77%                 | 3,77%                 | 1,89%              | 1,89%                 | 1,89%                 |
| indivíduos |                            | 22                      | 5               | 4               | 3                   | 3                  | 2                     | 2                     | 1                  | 1                     | 1                     |
| 9          | RF ≥ R\$2157               | <b>HSHPP141</b>         | <b>HSHPP121</b> | <b>HSHPP111</b> | <b>HSH7T141</b>     | <b>HSHPP131</b>    | <b>HSH7T111</b>       | <b>HSH7T131</b>       | <b>HAHPP121</b>    | <b>HAH7T141</b>       | <b>HSHNN111</b>       |
| 81         | Parcela 10 km Uni <14.58%  | 34,57%                  | 16,05%          | 8,64%           | 8,64%               | 6,17%              | 4,94%                 | 3,70%                 | 2,47%              | 2,47%                 | 2,47%                 |
| indivíduos |                            | 28                      | 13              | 7               | 7                   | 5                  | 4                     | 3                     | 2                  | 2                     | 2                     |
| 10         | RF < R\$4276               | <b>HSH7T131</b>         | <b>HSH7T121</b> | <b>HSHPP131</b> | <b>HSHPP121</b>     | <b>HSH7T111</b>    | <b>HSHNN111</b>       | <b>HSHPP111</b>       | <b>HSHPP141</b>    | <b>HSH7T141</b>       | <b>HSH7P111</b>       |
| 93         | Parcela 10 km Uni ≥14.58%  | 18,28%                  | 17,20%          | 8,60%           | 6,45%               | 6,45%              | 5,38%                 | 5,38%                 | 5,38%              | 5,38%                 | 4,30%                 |
| indivíduos | Parcela 10 km Uni <44.72%  | 17                      | 16              | 8               | 6                   | 6                  | 5                     | 5                     | 5                  | 5                     | 4                     |
| 11         | RF ≥ R\$4276               | <b>HSHPP131</b>         | <b>HSHPP121</b> | <b>HSHPP111</b> | <b>HSH7T131</b>     | <b>HSH7T141</b>    | <b>HSHAHPPPP11111</b> | <b>HSHAHPPPP13111</b> | <b>HSH7T121</b>    | <b>HSHPP141</b>       | <b>HAHPPPP1211</b>    |
| 59         | Parcela 10 km Uni ≥14.58%  | 22,03%                  | 18,64%          | 15,25%          | 10,17%              | 6,78%              | 5,09%                 | 5,09%                 | 5,09%              | 3,39%                 | 1,70%                 |
| indivíduos | Parcela 10 km Uni <44.72%  | 13                      | 11              | 9               | 6                   | 4                  | 3                     | 3                     | 3                  | 2                     | 1                     |
| 12         | Parcela 10 km Uni ≥ 44.72% | <b>HSH7T121</b>         | <b>HSHNN111</b> | <b>HSHPP111</b> | <b>HSSHNNN11111</b> | <b>HSSHPPP1211</b> | <b>HSHNP111</b>       | <b>HSHPP121</b>       | <b>HSHPP131</b>    | <b>HSHSHPPPP12111</b> | <b>HSH7T141</b>       |
| 55         | RF < R\$ 3950              | 21,82%                  | 14,55%          | 9,09%           | 7,27%               | 7,27%              | 3,64%                 | 3,64%                 | 3,64%              | 3,64%                 | 3,64%                 |
| indivíduos | Parcela 5 km Uni < 16.57%  | 12                      | 8               | 5               | 4                   | 4                  | 2                     | 2                     | 2                  | 2                     | 2                     |
| 13         | Parcela 10 km Uni ≥ 44.72% | <b>HSH7T111</b>         | <b>HSHNN111</b> | <b>HSH7T121</b> | <b>HSHPP121</b>     | <b>HSHPP111</b>    | <b>HAHNNN1111</b>     | <b>HAH7T111</b>       | <b>HSH7T141</b>    | <b>HAHPPPP1211</b>    | <b>HAHPP111</b>       |
| 71         | RF < R\$ 3950              | 22,54%                  | 19,72%          | 15,49%          | 12,68%              | 4,23%              | 2,82%                 | 2,82%                 | 2,82%              | 1,41%                 | 1,41%                 |
| indivíduos | Parcela 5 km Uni ≥ 16.57%  | 16                      | 14              | 11              | 9                   | 3                  | 2                     | 2                     | 2                  | 1                     | 1                     |
| 14         | Parcela 10 km Uni ≥ 44.72% | <b>HSHPP121</b>         | <b>HSHPP111</b> | <b>HSHPP131</b> | <b>HSH7T121</b>     | <b>HSHPP141</b>    | <b>HSH7T111</b>       | <b>HSHAHPPPP12111</b> | <b>HAHPPPP1111</b> | <b>HAHPP111</b>       | <b>HAHSHPPPP11131</b> |
| 72         | RF ≥ R\$ 3950              | 25,00%                  | 16,67%          | 9,72%           | 8,33%               | 6,94%              | 6,94%                 | 2,78%                 | 1,39%              | 1,39%                 | 1,39%                 |
| indivíduos | Parcela 5 km Uni < 21.36%  | 18                      | 12              | 7               | 6                   | 5                  | 5                     | 2                     | 1                  | 1                     | 1                     |
| 15         | Parcela 10 km Uni ≥ 44.72% | <b>HSHPP111</b>         | <b>HSH7T111</b> | <b>HSHPP121</b> | <b>HSHPP141</b>     | <b>HSHPT111</b>    | <b>HSHNN111</b>       | <b>HSH7T121</b>       | <b>HAHNN111</b>    | <b>HAHPP111</b>       | <b>HSHPN111</b>       |
| 76         | RF ≥ R\$ 3950              | 30,26%                  | 15,79%          | 10,53%          | 6,58%               | 6,58%              | 5,26%                 | 3,95%                 | 2,63%              | 2,63%                 | 2,63%                 |
| indivíduos | Parcela 5 km Uni ≥ 21.36%  | 23                      | 12              | 8               | 5                   | 5                  | 4                     | 3                     | 2                  | 2                     | 2                     |

RF = Renda Familiar; Parcela 10 km Uni = Parcela acumulada de universidades até 10 km do centróide da zona de residência; Parcela 5 km Uni = Parcela acumulada de universidades até 5 km do centróide da zona de residência.

## 7.8. SÍNTESE DOS RESULTADOS OBTIDOS

A partir dos resultados das árvores geradas para cada uma das amostras, pode-se verificar que, de uma forma geral, as variáveis independentes selecionadas para segregação dos dados para amostras de trabalhadores foram similares. No entanto, houve variação na ordem de seleção de cada uma delas para divisão das amostras, assim como na faixa de valor selecionada, para o caso de variáveis contínuas (Renda Familiar, por exemplo).

Algumas variáveis foram selecionadas apenas em uma das amostras de trabalhadores, como a variável TRABALHA, por exemplo, que apareceu somente na AD gerada a partir da amostra de comerciários. Pode-se explicar tal fato, pela maior quantidade de trabalhadores autônomos no setor de comércio na RMSP, e também especificamente na amostra analisada. Conseqüentemente, a referida variável foi considerada importante para formação dos grupos de indivíduos.

Para o caso de estudantes, foram selecionadas basicamente, aquelas variáveis de uso do solo (principalmente na amostra de universitários), RENDA FAMILIAR e NÚMERO DE AUTOMÓVEIS NO DOMICÍLIO. Na amostra de estudantes até o 1º grau, verifica-se a seleção da variável IDADE. Tal variável mostrou-se importante pelo algoritmo CART, principalmente porque a amostra possui características específicas (formada basicamente por crianças e adolescentes).

A Tabela 7.7 traz as variáveis independentes selecionadas em cada uma das seis amostras.

**Tabela 7.7 Variáveis independentes selecionadas nas seis amostras analisadas**

| Amostras      | Variáveis      |                                                                                                         |
|---------------|----------------|---------------------------------------------------------------------------------------------------------|
| Trabalhadores | Indústria      | CNH; Usa VTURA; Indústrias até 5 km; ESTUDA; Renda Familiar; Indústrias até 10 km; Indústrias até 15 km |
|               | Comércio       | ESTUDA; CNH; Usa VTURA; TRABALHA; Comércio até 15 km; Renda Familiar; AUTO; Comércio até 10 km          |
|               | Serviços       | CNH; ESTUDA; Usa VTURA; Serviços até 5 km; Serviços até 15 km; Sexo; AUTO                               |
| Estudantes    | Até o 1º grau  | Renda Familiar; AUTO; Idade; 1º grau até 5 km                                                           |
|               | No 2º grau     | Renda Familiar; 2º grau até 5 km; AUTO; 2º grau até 20 km; 2º grau até 10 km                            |
|               | Universitários | Universidades até 10 km; Renda Familiar; Universidades até 5 km                                         |

Analisando apenas os padrões predominantes (o primeiro padrão de viagem mais freqüente) por folha em todas as amostras, verifica-se também certa similaridade considerando as três amostras de trabalhadores. No entanto, observa-se que para o caso dos estudantes, a amostra composta por

universitários apresentou maior variação de padrões predominantes, principalmente em relação à seqüência de destinos. Estudantes universitários realizam predominantemente viagens mais longas e também utilizam modos de transporte motorizados. Neste caso específico, a localização de ofertas de universidades possivelmente exerce grande influência na escolha de padrões de viagens.

A Tabela 7.8 apresenta os padrões mais freqüentes em cada um dos nós terminais para todas as amostras.

Tabela 7.8 Padrões de viagens predominantes nas folhas das seis amostra

| Amostras      |                | Padrões de viagens predominantes                       |                                        |                                           |  |
|---------------|----------------|--------------------------------------------------------|----------------------------------------|-------------------------------------------|--|
| Trabalhadores | Indústria      | <b>HWHTT141</b><br>(Nó 8 e 12)                         | <b>HWHTT111</b><br>(Nó 9)              | <b>HWHNN111</b><br>(Nó 20 e 21)           |  |
|               |                | <b>HWHSHNNNN11111</b><br>(Nó 11)                       | <b>HWHPP111</b><br>(Nó 52, 53, 27 e 7) |                                           |  |
|               | Comércio       | <b>HWHTT141</b><br>(Nó 8 e 10)                         | <b>HWHNN111</b><br>(Nó 18 e 19)        | <b>HWHPP111</b><br>(Nó 44, 45, 46, 47)    |  |
|               |                | <b>HWHSHNNNN11111</b><br>(Nó 6)                        | <b>HWSHTNN1211</b><br>(Nó 7)           |                                           |  |
|               | Serviços       | <b>HWHTT141</b><br>(Nó 8 e 12)                         | <b>HWHNN111</b><br>(Nó 18 e 19)        | <b>HWHPP111</b><br>(Nó 52, 53, 54, 55, 7) |  |
|               |                | <b>HWHSHNNNN11111</b><br>(Nó 5)                        |                                        |                                           |  |
| Estudantes    | Até o 1º grau  | <b>HSHNN111</b><br>(Nó 16, 17, 18, 19, 10, 11, 12, 13) |                                        | <b>HSHPP111</b><br>(Nó 7)                 |  |
|               | No 2º grau     | <b>HSHNN111</b><br>(Nó 8, 9, 10, 11, 6, 14)            | <b>HSHPP111</b><br>(Nó 30 e 31)        |                                           |  |
|               | Universitários | <b>HSHTT141</b><br>(Nó 8)                              | <b>HSHPP141</b><br>(Nó 9)              | <b>HSHTT131</b><br>(Nó 10)                |  |
|               |                | <b>HSHPP131</b><br>(Nó 11)                             | <b>HSHTT121</b><br>(Nó 12)             | <b>HSHTT111</b><br>(Nó 13)                |  |
|               |                | <b>HSHPP121</b><br>(Nó 14)                             | <b>HSHPP111</b><br>(Nó 15)             |                                           |  |

Vale ressaltar que o escopo do presente capítulo é apresentar os resultados obtidos. Análises minuciosas e discussões a respeito de diferentes escolhas de padrões de viagens devido a fatores de uso do solo, socioeconômicos e participação em atividades de subsistência serão realizadas no capítulo subsequente. No entanto, apenas observando as seis árvores geradas pelo *software*, são verificadas relações entre os conjuntos de variáveis propostos no trabalho e padrões de viagens realizados por indivíduos residentes na RMSP.

## Capítulo

# 8 ANÁLISE DOS RESULTADOS

---

*Após processamento da AD para cada uma das amostras e apresentação dos resultados na sua forma gráfica e tabular, a próxima etapa do método corresponde à análise dos resultados obtidos. Este capítulo analisa as árvores geradas para cada uma das seis amostras. Foi verificada a influência das variáveis independentes (sobretudo das variáveis de uso do solo) na seqüência das viagens realizadas pelos indivíduos, a partir da construção das árvores (ramo a ramo), assim como detalhamento de alguns nós terminais.*

### 8.1. ANÁLISE GERAL DOS RESULTADOS

Foram identificados dez nós terminais nas três amostras de trabalhadores, nove folhas na amostra de estudantes até o 1º grau e oito nós terminais nas amostras de estudantes no 2º grau e universitários. As folhas (ou nós terminais) representam grupos de indivíduos caracterizados de forma “homogênea” em relação ao padrão de viagem escolhido. Muitos dos grupos encontrados possuíam características semelhantes entre si e se diferenciavam apenas por uma das variáveis (Idade, variáveis de uso do solo, Renda Familiar, Estuda, etc.).

Desta forma, encontra-se certa variação de comportamento expressa em diferentes escolhas de padrões de viagens ou probabilidade de ocorrência destes devido à influência de determinadas características.

Uma das principais finalidades deste capítulo é comparar alguns dos grupos para identificar as variáveis que possivelmente influenciam o comportamento de viagem (seja em relação à seqüência de atividades, modos de transporte ou destinos) e analisar mais detalhadamente alguns nós para enriquecimento da análise. Acompanhar a construção da hierarquia da árvore (ramo a ramo) também foi um dos métodos para investigar a influência de cada uma das variáveis nas escolhas das seqüências de viagens.

Fazendo uma análise genérica do conjunto de resultados, pode-se dizer que as três amostras de trabalhadores apresentaram comportamentos semelhantes (especialmente trabalhadores no setor industrial e no setor de serviços), considerando principalmente as variáveis independentes selecionadas pelo algoritmo CART (“Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO”; “USA VALE TRANSPORTE”; “ESTUDA”).

Este resultado é justificável pelo fato de os três tipos de amostras terem algumas características similares (são trabalhadores assalariados ou autônomos, que eventualmente estudam, com idade na faixa de 20 a 44 anos, etc.). Assim, as principais variáveis, com maior poder discriminatório nos três conjuntos de dados são as mesmas, havendo eventualmente diferenças na ordem de seleção destas pelo algoritmo.

Há alguns casos particulares como a seleção da variável “TRABALHA” para o caso da amostra composta por comerciários. Como mencionado anteriormente, a amostra composta por comerciários possui maior número de trabalhadores autônomos, tornando assim esta variável importante quanto a escolhas das seqüências de atividades, por exemplo.

Contudo, a principal diferença de comportamento relativo a viagens considerando os três setores econômicos é na escolha da seqüência de destinos e, conseqüentemente, na importância do conjunto de variáveis de uso do solo. Trabalhadores no setor industrial fazem viagens mais longas do que aqueles que trabalham no setor comercial. Isto se justifica pela diferença de distribuição geográfica de empregos nos dois setores econômicos na área de estudo.

Quanto às três amostras de estudantes, estas têm características singulares, tanto em relação à variável “IDADE” (6 a 16 anos para o caso dos estudantes matriculados até o 1º grau; 15 a 20 anos para o caso dos estudantes matriculados no 2º grau; 18 a 28 anos para o caso dos estudantes universitários) quanto em relação à variável “GRAU DE INSTRUÇÃO”. Adicionalmente, pode-se verificar que algumas variáveis selecionadas no caso dos trabalhadores, não poderiam ser selecionadas no presente caso. As variáveis relacionadas à participação em atividades como

“TRABALHA” e “ESTUDA”, por exemplo, não são importantes no caso de estudantes, pois todos os que compõem as amostras estudam e não trabalham. Nas três amostras de estudantes observou-se a importância das variáveis “RENDA FAMILIAR” e “Nº DE AUTOMÓVEIS NO DOMICÍLIO” para segregação dos dados.

Houve seleção de outras variáveis como “IDADE”, por exemplo, apenas na amostra composta por estudantes até o 1º grau, pois esta variável mostrou-se relevante principalmente considerando a seqüência de modos de transporte escolhidos e também a seqüência de destinos (crianças mais velhas utilizam com maior freqüência modo de transporte público e fazem viagens mais longas, quando comparadas àquelas crianças mais novas). Observou-se também variação do comportamento relacionado ao encadeamento de viagens, levando-se em conta a amostra composta por universitários, que possui uma série de características particulares.

Ainda falando genericamente dos resultados obtidos, observa-se maior importância das variáveis de uso do solo na escolha dos padrões de viagens em cadeia nas amostras compostas por universitários e industriários. Este comportamento explica-se principalmente devido a características relativas a ambas as amostras. Há localizações restritas de pólos de concentração de indústrias e universidades na RMSP, influenciando fortemente a escolha de destinos mais próximos ou distantes para satisfazer tais necessidades (trabalho na indústria ou estudo em universidade).

As próximas seções trazem as análises mais detalhadas para cada uma das amostras utilizadas no trabalho.

## 8.2 TRABALHADORES NO SETOR INDUSTRIAL

Considerando os critérios para partição dos dados e uma amostra final de 4.084 industriários, foi gerada a AD através do *software* S-PLUS 6.1. Vale ressaltar que tanto o número de folhas obtidas, quanto as variáveis selecionadas pelo algoritmo dependem da escolha de tais critérios. Caso o desvio adotado tivesse menor valor, possivelmente a AD obtida seria composta por maior número de ramificações e variáveis independentes.

Para análise da influência de características socioeconômicas, de uso do solo e participação em atividades nas escolhas dos padrões de viagens urbanas realizados pelos industriários, foi observada a construção da hierarquia da árvore. A importância de cada uma das variáveis, seja na escolha da seqüência das atividades, dos modos de transporte ou dos destinos, é explicitada quando se analisa a construção da árvore (ramo a ramo).

Assim, a Figura 8.1 ilustra a primeira partição dos dados, a partir do nó raiz. A primeira variável considerada foi “Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO” (CNH), segregando os dados respectivamente em dois ramos: (1) Ramo 1 – CNH=0 (Nó terminal 8; 9; 20; 21 e 11) ; e (2) Ramo 2 – CNH ≥ 1 (Nó terminal 12; 52; 53; 27 e 7).

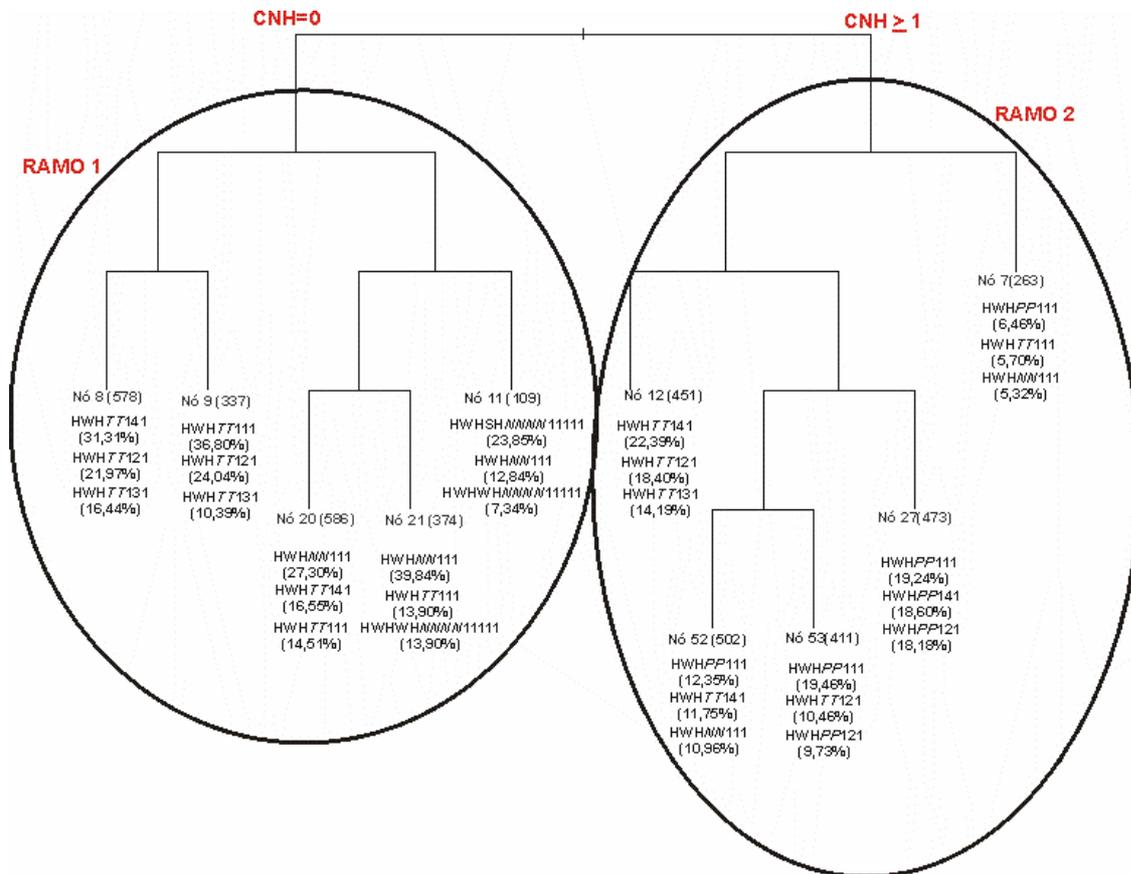


Figura 8.1 Primeira partição dos dados para trabalhadores no setor industrial (influência da variável CNH na escolha da seqüência de modos de transporte).

Espera-se intuitivamente que tal variável exerça influência principalmente em relação à seqüência de modos de transporte. A provável presença de membros da família com Carteira Nacional

de Habilitação (e, portanto, pelo menos um automóvel no domicílio) pode delimitar o uso mais freqüente ou não de automóvel na seqüência de viagens realizadas.

Observando as folhas que compõem os ramos 1 e 2, e os padrões predominantes em cada folha, verifica-se que no ramo 1 (que corresponde a domicílios sem nenhum membro com Carteira Nacional de Habilitação) predominantemente observam-se seqüências de modos de transportes não motorizados ou transporte público (TT – nó 8 e 9; NN – nó 20 e 21; NNN – nó 11). Por outro lado, no ramo 2 (que corresponde aos domicílios com pelo menos um membro familiar com Carteira Nacional de Habilitação) verifica-se o predomínio do uso de automóvel, ou modo de transporte público (TT – nó 12; PP – nó 52, 53, 27 e 7).

Em seguida os ramos 1 e 2 são novamente divididos, gerando os ramos 3, 4, 5 e 6. O ramo 1 foi segregado (ramo 3 e 4) considerando a variável “USA VALE TRANSPORTE” (Usa VTRA). Já o ramo 2 foi dividido (ramo 5 e 6) considerando a variável “ESTUDA” (Figura 8.2).

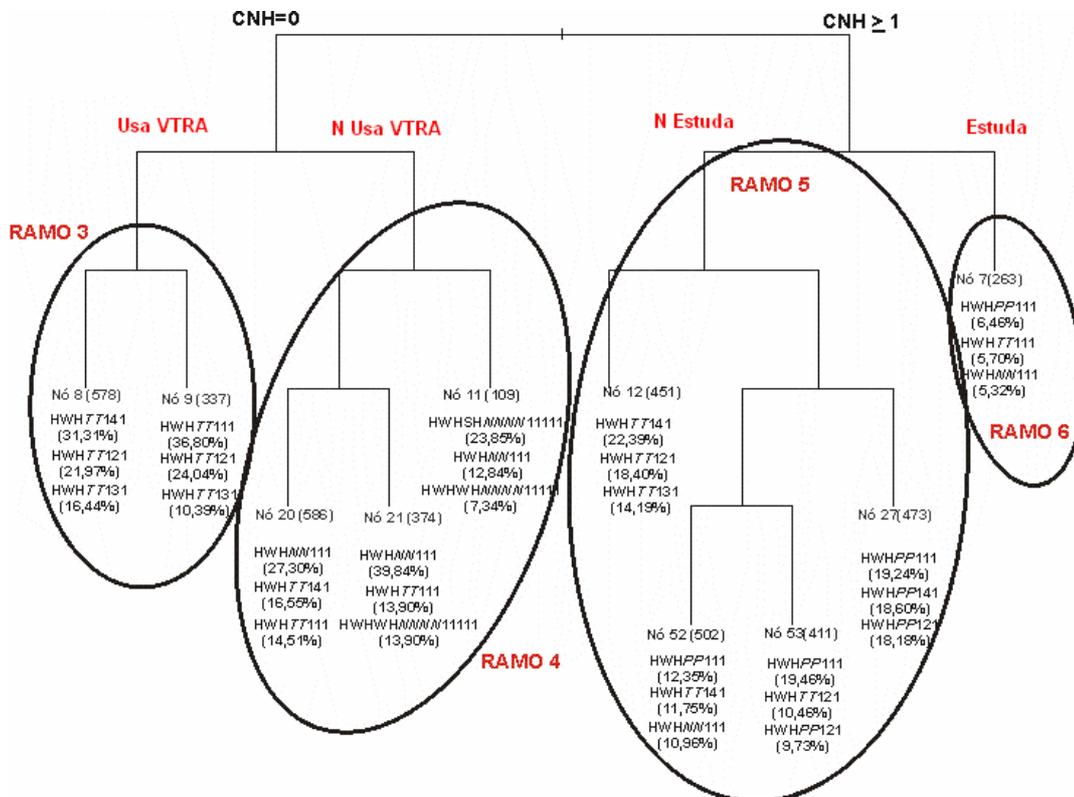


Figura 8.2 Segunda partição dos dados para trabalhadores no setor industrial (influência das variáveis Usa VTRA e ESTUDA na escolha da seqüência de modos de transporte).

Espera-se que a variável “USA VALE TRANSPORTE” interfira também na seqüência de modos de transporte utilizados durante as viagens. Observando o ramo 3 (que corresponde àqueles indivíduos que residem em domicílios provavelmente sem nenhuma carteira de habilitação e que usam Vale Transporte), pode-se verificar que há o predomínio do uso de transporte coletivo motorizado (TT – nó 8 e 9). Enquanto que o ramo 4 (que corresponde àquelas pessoas que também moram em domicílios sem nenhuma provável carteira de habilitação, mas não fazem uso de Vale Transporte) apresenta predominância de modo de transporte não motorizado para realização da seqüência de viagens (NN – nó 20 e 21; NNNN – nó 11).

Muito há de se discutir sobre a variável “USA VALE TRANSPORTE”. Não se sabe realmente a respeito da sua relação de causalidade ou consequência considerando a variável dependente. Afinal, há indagações a respeito desta variável: (1) A variável seria socioeconômica ou estaria relacionada a características do sistema de transportes da RMSP? ; (2) Pessoas utilizam modo de transporte público por possuírem Vale Transporte? ; (3) Indivíduos possuem Vale Transporte por haver necessidade de realizar viagens ao trabalho por transporte público? ; ou ainda (4) Há pessoas (que ainda que possuam Vale Transporte) não o utilizam devido à disponibilidade de outros modos de transporte alternativos ao coletivo? Uma afirmação que pode ser feita é que tal variável está relacionada fortemente ao custo do transporte.

Considerando os ramos 5 e 6, provenientes da subdivisão do ramo 2, verifica-se que, no presente caso, a variável considerada foi “ESTUDA”. Pode-se afirmar que a participação de indivíduos em atividades de estudos, possivelmente influencia a seqüência de atividades realizadas. Aqueles indivíduos que estudam (e também trabalham na indústria) realizariam atividades diferentes daqueles que apenas trabalham no setor industrial.

Embora não estejam ilustradas na Figura 8.2 diferenças na seqüência de atividades, pois se observa apenas a seqüência HWH predominantemente em todas as folhas pertencentes aos ramos 5 e 6, o nó 7 (ramo 6) apresenta outras seqüências de atividades como HSH; HWHS (apresentadas na

síntese do relatório gerado pelo S-PLUS 6.1 – Tabela 7.1, Capítulo 7), onde há motivos de viagem relacionados à escola.

Continuando a observar a construção da árvore, a Figura 8.3, a seguir, ilustra a próxima subdivisão dos dados.

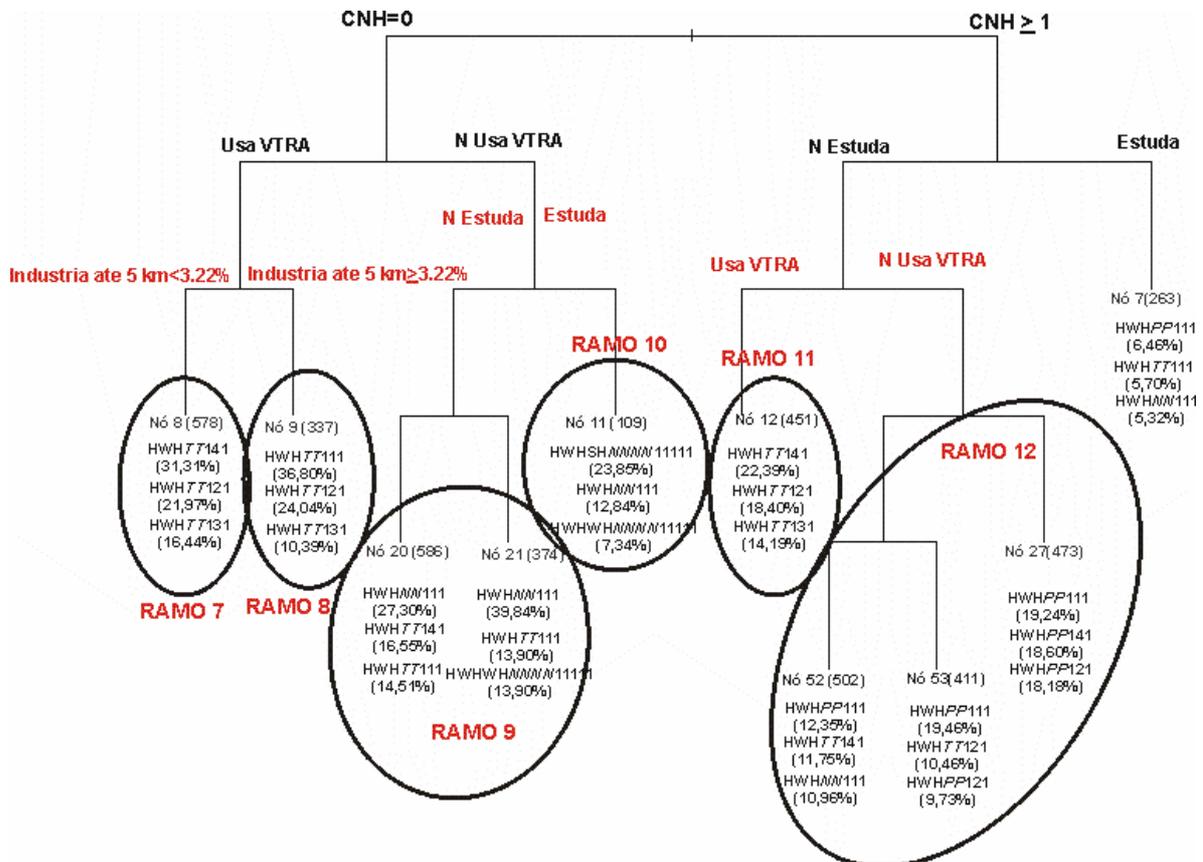


Figura 8.3 Terceira partição dos dados para trabalhadores no setor industrial (influência das variáveis PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 5 KM A PARTIR DO CENTRÓIDE DA ORIGEM; ESTUDA; USA VALE TRANSPORTE).

O ramo 3 subdivide-se nos ramos 7 e 8 (que correspondem respectivamente às folhas 8 e 9) considerando a variável de uso do solo “PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 5 KM DO CENTRÓIDE DA ORIGEM”. Analisando os padrões predominantes nas duas folhas, verifica-se a clara influência da referida variável de uso do solo na seqüência de destinos escolhidos pelos indivíduos.

Aquelas pessoas que residem em zonas de tráfego onde a “PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 5 KM DO CENTRÓIDE DA ORIGEM” é inferior a 3,22%, ou seja,

moram em zonas de tráfego com poucas oportunidades de empregos na indústria nas suas vizinhanças, realizam viagens mais longas para satisfazer a necessidade de trabalho (Nó 8 – HWH77141 (31,31%); HWH77121 (21,97%); HWH77131 (16,44%)). No caso em questão, a maioria das pessoas que compõe este grupo faz viagens acima de 15 km, para realizar atividades de trabalho, utilizando transporte público.

Considerando-se, todavia, os indivíduos que moram em zonas de tráfego onde a “PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 5 KM DO CENTRÓIDE DA ORIGEM” é superior ou igual a 3,22% (Ramo 8 – Nó terminal 9), há, predominantemente, a escolha de destinos mais próximos à origem para realização de atividades de trabalho na indústria (HWH77111 (36,80%); HWH77121 (24,04%); HWH77131 (10,39%)). Assim, a maioria dos indivíduos que compõe o grupo com essas determinadas características (Sem nenhuma provável Carteira Nacional de Habilitação no domicílio, usa Vale Transporte e mora em zonas de tráfego com uma parcela acumulada de empregos maior ou igual a 3,22% até 5 km do seu centróide) realiza viagens ao trabalho com uma distância inferior ou igual a 5 km, utilizando transporte coletivo motorizado.

Os indivíduos que compõe os nós terminais 8 e 9 exemplificam bem a diferença entre pessoas que moram em zonas de tráfego com alto número de oportunidades de empregos e aquelas que residem em zonas com baixo número de oportunidades de empregos. No último caso citado, há a necessidade de realizar viagens longas para o trabalho, pelo fato de não haver oferta de empregos na indústria suficiente nas suas proximidades.

Analisando a folha 8 mais detalhadamente e os deslocamentos predominantes dos indivíduos que compõem tal grupo em termos de zonas de tráfego, verifica-se que as viagens mais longas (acima de 5 km, por exemplo) ocorrem quando a parcela de industrialização da zona de residência é relativamente baixa (0,02% a 0,74%). Nestes casos, a parcela de industrialização da zona de tráfego de destino é bem mais alta do que na zona de tráfego da origem (0,38% a 1,26%).

A Tabela 8.1 mostra os dez padrões de viagens predominantes (em termos de motivos de viagem, modos de transporte e zonas de tráfego), as zonas de origem com suas respectivas parcelas de industrialização e as zonas de tráfego de destino com os mesmos dados. Verifica-se que a parcela média de industrialização dos destinos (0,83%) é bem mais alta do que a parcela média de industrialização das origens (0,50%). Pode-se afirmar então que, quando não há oportunidades de empregos suficientes nas origens, os indivíduos fazem viagens mais longas para o trabalho, apesar da maior impedância associada à realização de tais viagens.

Tabela 8.1 Análise detalhada dos deslocamentos dos indivíduos que compõem o nó 8

| Padrão Predominante (zonas de tráfego) | Residência |      |                       | Destino - Trabalho |      |                   | Distância<br>Km |              |
|----------------------------------------|------------|------|-----------------------|--------------------|------|-------------------|-----------------|--------------|
|                                        | zona       | Nome | Parcela_indústria     | zona               | Nome | Parcela_indústria |                 |              |
| <b>HWH</b> TT322-322-322               | 0,71%      | 322  | Cajamar               | 1,59%              | 322  | Cajamar           | 1,59%           | 1,74         |
| <b>HWH</b> TT335-335-335               | 0,71%      | 335  | Miguel Badra          | 0,59%              | 335  | Miguel Badra      | 0,59%           | 2,24         |
| <b>HWH</b> TT145-19-145                | 0,53%      | 145  | Vila Terezinha        | 0,15%              | 19   | Bom Retiro        | 0,40%           | <b>9,31</b>  |
| <b>HWH</b> TT145-80-145                | 0,53%      | 145  | Vila Terezinha        | 0,15%              | 80   | Limão             | 0,30%           | 5,77         |
| <b>HWH</b> TT219-77-219                | 0,53%      | 219  | Limoeiro              | 0,13%              | 77   | Lapa              | 0,59%           | <b>24,01</b> |
| <b>HWH</b> TT243-181-243               | 0,53%      | 243  | Eldorado              | 0,33%              | 181  | Diadema           | 1,22%           | 5,22         |
| <b>HWH</b> TT289-95-289                | 0,53%      | 289  | José Bonifácio        | 0,02%              | 95   | Vila California   | 0,38%           | <b>11,58</b> |
| <b>HWH</b> TT297-297-297               | 0,53%      | 297  | Mauá                  | 1,19%              | 297  | Mauá              | 1,19%           | 3,05         |
| <b>HWH</b> TT300-241-300               | 0,53%      | 300  | Reservatório Billings | 0,74%              | 241  | Planalto          | 1,26%           | 6,88         |
| <b>HWH</b> TT341-238-341               | 0,53%      | 341  | Ouro Fino             | 0,14%              | 338  | Suzano            | 0,80%           | <b>9,07</b>  |
|                                        |            |      | Média                 | <b>0,50%</b>       |      | Média             | <b>0,83%</b>    | <b>7,89</b>  |

Ainda observando a Figura 8.3, constata-se que o ramo 4 subdivide-se nos ramos 9 e 10 de acordo com a variável "ESTUDA". Neste caso, observa-se bem a influência da variável principalmente na seqüência de atividades realizadas. Indivíduos que possuem as características observadas na folha 11 ou ramo 10 (Não possuem Carteira Nacional de Habilitação no domicílio, não usam Vale Transporte e estudam) realizam predominantemente viagens ao trabalho e à escola (HWHSH/WWW11111 – 23,85%). Já aqueles indivíduos que possuem as mesmas características citadas, mas não estudam, somente trabalham no setor industrial, predominantemente realizam seqüências de atividades relacionadas ao trabalho apenas (HWH/WW111 – Nó 20 e 21). Há de se esperar tal comportamento, devido a diferenças de necessidades, considerando participação em atividades de estudo, além das atividades de trabalho no setor industrial.

Enfim, considera-se o ramo 5 que se ramifica formando os ramos 11 (Nó terminal 12) e 12 (Folhas 52, 53 e 27). A variável selecionada desta vez foi “USA VALE TRANSPORTE” (Usa VTRA). Constata-se novamente a influência de tal variável na sequência de modos de transportes escolhidos. Como no caso os indivíduos residem em domicílios com pelo menos uma provável Carteira Nacional de Habilitação (domicílios onde há pelo menos um automóvel), se verificam diferenças no uso de transporte coletivo ou automóvel. Pessoas que utilizam Vale Transporte usam modo de transporte coletivo mais frequentemente (HWH77141 – Nó 12 ou ramo 11), enquanto que aquelas pessoas que não fazem uso do Vale Transporte, realizam viagens por automóvel predominantemente (HWHPP111 – Nó 52, 53 e 27 – ramo 12).

Para ilustrar a quarta partição dos dados considerando respectivamente as variáveis “PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 15 KM DO CENTRÓIDE DA ORIGEM” e “RENDA FAMILIAR”, segue a Figura 8.4.

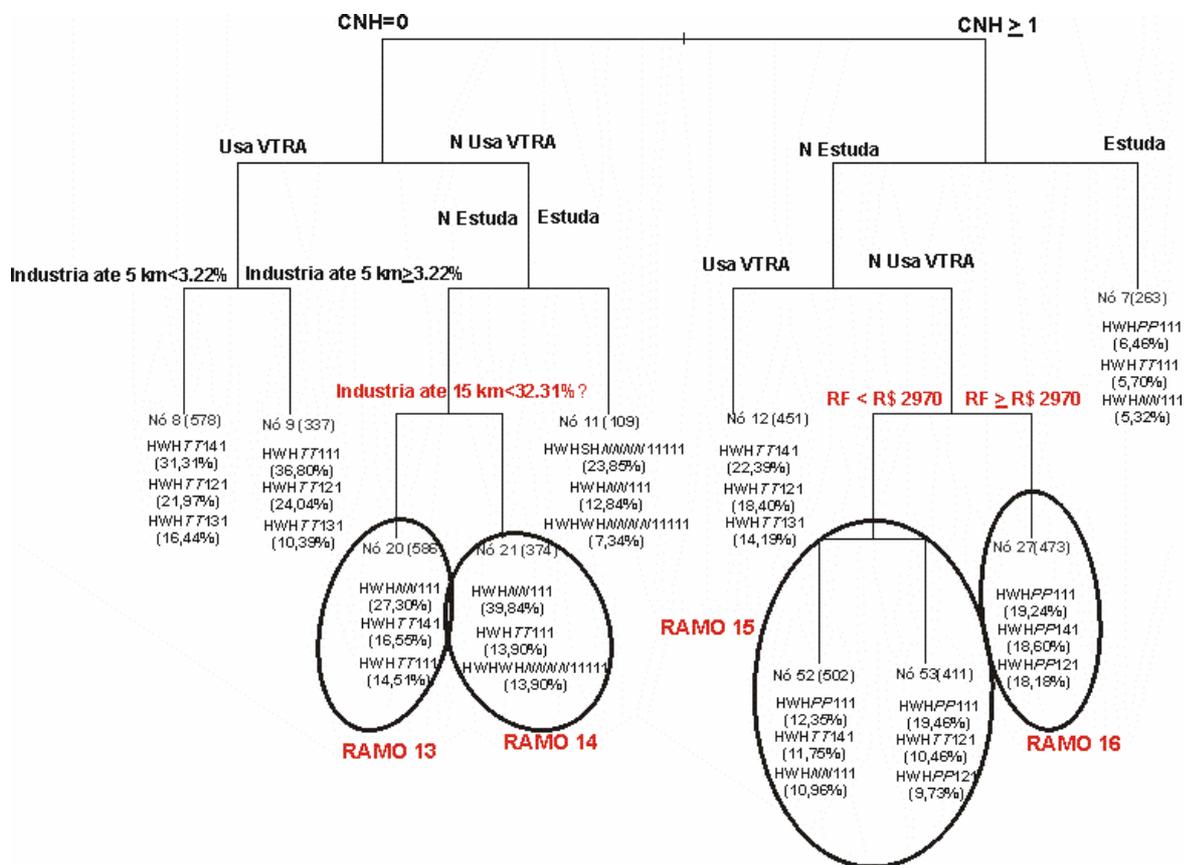


Figura 8.4 Quarta partição dos dados para trabalhadores no setor industrial (influência das variáveis PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 15 KM A PARTIR DO CENTRÓIDE DA ORIGEM; RENDA FAMILIAR).

O ramo 9 segrega-se formando os ramos 13 e 14 (que correspondem respectivamente às folhas 20 e 21) considerando a variável de uso do solo “PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 15 KM DO CENTRÓIDE DA ORIGEM”.

Assim, os grupos de indivíduos pertencentes às referidas folhas possuem as seguintes características: (1) Não possuem Carteira Nacional de Habilitação no domicílio (CNH=0); (2) Não usam Vale Transporte (N Usa VTRA); (3) Não estudam. Diferenciam-se pela parcela acumulada de empregos até uma distância de 15 km a partir do centróide da origem. Embora, apareçam apenas os três padrões predominantes em cada folha, observam-se (principalmente através da síntese do relatório – Tabela 7.1 do Capítulo 7) diferenças na frequência de padrões de viagem considerando principalmente a seqüência de destinos.

Em ambas as folhas, verifica-se o mesmo padrão predominante “HWH/M111”, havendo diferença na sua frequência de ocorrência (27,30% - nó 20 e 39,84% - nó 21). Espera-se que indivíduos residentes em zonas de tráfego com um número de oportunidades de empregos relativamente baixo até uma distância de 15 km (Nó 20) realizem viagens mais longas, com motivos relacionados ao trabalho. A Tabela 8.2 apresenta as diferentes frequências de alguns padrões de viagens, onde há escolha de destinos mais distantes considerando as folhas 20 e 21, respectivamente. Há maiores frequências de escolha de destinos mais distantes da origem (“121”; “131”; “141”) no grupo de indivíduos que compõe a folha 20 (Indústria até 15 km < 32.31%).

Tabela 8.2 Diferenças nas seqüências de destinos entre as folhas 20 e 21

| Padrões  | Folhas                       |                              |
|----------|------------------------------|------------------------------|
|          | 20                           | 21                           |
|          | Indústria ate 15 km < 32.31% | Indústria ate 15 km ≥ 32.31% |
| HWH/T141 | 16,55%                       | 2,14%                        |
| HWH/T121 | 13,65%                       | 9,89%                        |
| HWH/T131 | 7,34%                        | 5,88%                        |

Ainda considerando a folha 21, fazendo-se uma análise detalhada de tal nó em relação aos deslocamentos em termos de zonas de tráfego, verifica-se que as viagens são relativamente curtas e

intra-zonais (0,67 a 1,73 Km). Isto ocorre porque as parcelas de industrialização das origens não são tão baixas quanto no caso do nó 8, por exemplo.

Havendo empregos suficientes da zona de origem, ou nas vizinhanças, não existe necessidade de realização de viagens longas ao trabalho. Ressalta-se, além disso, que muitas vezes ocorrem decisões a longo prazo como escolhas residenciais baseadas nos locais de trabalho. Possivelmente, muitos dos indivíduos que fazem parte da folha 21 escolheram a localização das suas residências baseando-se nos seus locais de trabalho. A Tabela 8.3 traz características dos padrões de deslocamentos predominantes (em termos de motivos de viagens, modos de transporte e zonas de tráfego) dos indivíduos que compõem o nó 21.

Tabela 8.3 Análise detalhada dos deslocamentos dos indivíduos que compõem o nó 21

| Padrão Predominante (zonas de tráfego) | Residência |                    |                   | Destino - Trabalho |                    |                   | Distância Km |
|----------------------------------------|------------|--------------------|-------------------|--------------------|--------------------|-------------------|--------------|
|                                        | zona       | Nome               | Parcela_indústria | zona               | Nome               | Parcela_indústria |              |
| <b>HWHH</b> WV50-50-50-50 1,60%        | 50         | Parque Novo Mundo  | 0,58%             | 50                 | Parque Novo Mundo  | 0,58%             | 1,33         |
| <b>HWH</b> WV103-103-103 1,34%         | 103        | Moinho Velho       | 0,37%             | 103                | Moinho Velho       | 0,37%             | 1,09         |
| <b>HWH</b> WV103-103-103 1,34%         | 103        | Moinho Velho       | 0,37%             | 103                | Moinho Velho       | 0,37%             | 1,09         |
| <b>HWH</b> WV106-106-106 1,07%         | 106        | Sacomã             | 0,64%             | 106                | Sacomã             | 0,64%             | 1,48         |
| <b>HWH</b> WV19-19-19 1,07%            | 19         | Bom Retiro         | 0,39%             | 19                 | Bom Retiro         | 0,39%             | 0,93         |
| <b>HWHH</b> WV80-80-80-80 1,07%        | 80         | Limão              | 0,30%             | 80                 | Limão              | 0,30%             | 0,7          |
| <b>HWH</b> WV181-181-181 0,80%         | 181        | Diadema            | 1,22%             | 181                | Diadema            | 1,22%             | 1,73         |
| <b>HWH</b> WV21-21-21 0,80%            | 21         | Pari               | 0,31%             | 21                 | Pari               | 0,31%             | 0,67         |
| <b>HWH</b> WV28-28-28 0,80%            | 28         | Vila Independência | 0,45%             | 28                 | Vila Independência | 0,45%             | 0,7          |
| <b>HWH</b> WV57-28-57 0,80%            | 57         | Quinta Pararada    | 0,14%             | 28                 | Vila Independência | 0,45%             | 1,71         |
|                                        |            | Média              | <b>0,48%</b>      |                    | Média              | <b>0,51%</b>      | <b>1,14</b>  |

Observando-se ainda a Figura 8.4, há subdivisão do ramo 12 considerando a variável “RENDA FAMILIAR”. Há formação então do ramo 15 (Nó terminal 52 e 53) e do ramo 16 (Nó terminal 27). Observa-se o uso mais freqüente do automóvel (PP) principalmente no grupo de indivíduos que formam a folha 27 (possuem Renda Familiar superior ou igual a R\$ 2970). É de se esperar que indivíduos com rendas maiores utilizem predominantemente automóvel (já que se sabe que neste caso ambos os grupos de indivíduos residem em domicílio com pelo menos um automóvel –  $CNH \geq 1$ ).

Constata-se então que a folha 27 apresenta os três padrões predominantes com uso de automóvel para realização de viagens ao trabalho, variando apenas as escolhas dos destinos. Pelo menos 56,02% dos indivíduos (HWHPP111 (19,24%); HWHPP141 (18,60%); HWHPP121(18,18%))

utilizam transporte particular motorizado. No entanto, considerando apenas os três padrões predominantes nas folhas 52 e 53, observa-se que apenas 12,35% dos indivíduos pertencentes ao grupo que compõe a folha 52 utilizam automóvel (HWHPP111). Já analisando a folha 53, apenas 29,19% dos indivíduos que formam a folha 53 fazem uso de automóvel (HWHPP111 (19,46%); HWHPP121 (9,73%)). Salienda-se, no entanto, que neste caso foram observados apenas os três padrões predominantes. O leitor interessado em uma análise mais detalhada da frequência dos demais padrões deve observar anexos ou síntese de relatórios apresentadas no Capítulo 7.

Finalmente, a quinta e última partição dos dados considerando a variável de uso do solo “PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 10 KM DO CENTRÓIDE DA ORIGEM” é ilustrada na Figura 8.5.

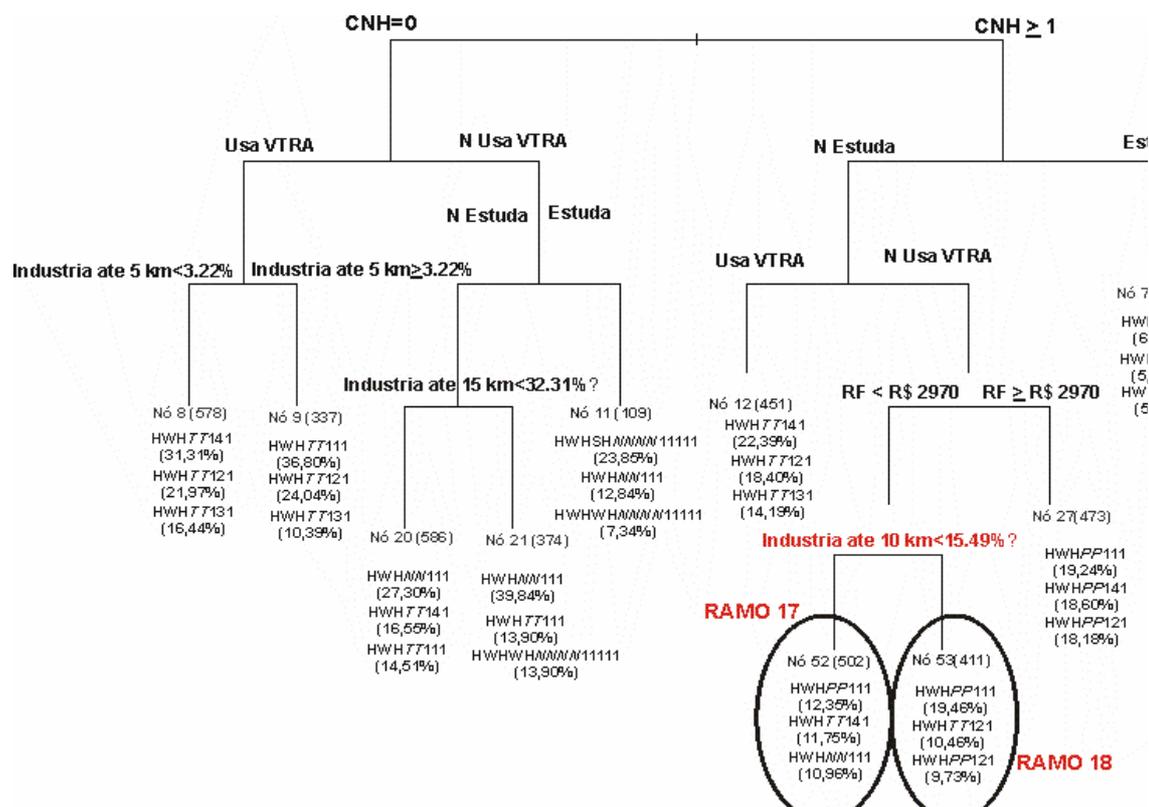


Figura 8.5 Quinta partição dos dados para trabalhadores no setor industrial (influência da variável PARCELA ACUMULADA DE EMPREGOS NA INDÚSTRIA ATÉ 10 KM A PARTIR DO CENTRÓIDE DA ORIGEM).

Assim, no último passo para a obtenção da árvore final, o ramo 15 segrega-se, dando origem aos ramos 17 e 18, os quais correspondem às folhas 52 e 53, respectivamente. Os grupos de pessoas que compõem tais folhas possuem as seguintes características: (1) Possuem pelo menos um membro da família com Carteira Nacional de Habilitação; (2) Não estudam; (3) Não utilizam Vale Transporte; e (4) Possuem Renda Familiar inferior a R\$ 2970. As folhas diferenciam-se pela Parcela de oportunidades (empregos na indústria) acumulada até um raio de 10 km a partir do centróide da zona de origem.

Como mencionado anteriormente, esta variável de uso do solo influencia especialmente a escolha de destinos. Espera-se, por exemplo, que aqueles indivíduos que formam a folha 52 (Indústria até 10 km < 15.49%), façam viagens longas (maiores que 10 km "131" ou "141") para realizar atividades relacionadas ao trabalho na indústria.

A Tabela 8.4 representa alguns padrões onde se observa escolha de destinos mais distantes nas folhas 52 e 53 e suas respectivas frequências de ocorrência. As frequências foram extraídas do relatório emitido pelo S-PLUS 6.1, sintetizado no Capítulo 7. Nos anexos apresentados ao final do presente texto encontram-se as frequências de todos os padrões de viagem da análise, provenientes do relatório gerado pelo *software*.

Tabela 8.4 Diferenças nas seqüências de destinos entre as folhas 52 e 53

| Padrões  | Folhas                       |                              |
|----------|------------------------------|------------------------------|
|          | 52                           | 53                           |
|          | Indústria até 10 km < 15.49% | Indústria até 10 km ≥ 15.49% |
| HWHTT141 | 11,75%                       | 2,92%                        |
| HWHPP141 | 9,76%                        | 2,68%                        |
| HWHTT121 | 9,16%                        | 10,46%                       |
| HWHPP121 | 9,36%                        | 9,73%                        |
| HWHTT131 | 8,96%                        | 3,65%                        |

Observa-se que aqueles indivíduos que fazem parte da folha 53 (ramo 18) escolhem mais freqüentemente destinos situados entre 5 e 10 km da origem ("121") quando comparados àqueles indivíduos que formam a folha 52. Este fato ocorre principalmente porque há maiores oportunidades de empregos na indústria (superior ou igual a 15,49%) nesta faixa de distância considerando este grupo

de pessoas (Nó terminal 53). Aqueles indivíduos que compõe o nó 52 necessitam fazer viagens acima de 10 km para satisfazer a sua necessidade de trabalho na indústria.

### 8.2.1 Síntese da análise de resultados – amostra de industriários

Verificou-se na análise de resultados obtidos a partir da amostra composta por industriários, influência de algumas variáveis na seqüência de viagens realizadas pelos indivíduos trabalhadores no setor industrial.

- Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO: influencia principalmente a seqüência de modos de transporte utilizados;
- USA VALE TRANSPORTE: também influencia a seqüência de modos de transporte utilizados;
- ESTUDA: influencia especialmente as seqüências de atividades (motivos de viagem) realizadas pelos indivíduos;
- RENDA FAMILIAR: exerce influência principalmente quando se considera a seqüência de modos de viagem utilizados;
- VARIÁVEIS DE USO DO SOLO: interferem especificamente na seqüência de destinos escolhidos pelos indivíduos, devido principalmente à localização restrita de empregos industriais na RMSP.

Um fato importante que vale ressaltar é que não houve grande variação na seqüência de atividades predominantes. Em todos os casos verificou-se que a grande maioria das atividades estava relacionada ao trabalho. Características da amostra analisada (aproximadamente 95% de assalariados) podem explicar a não diversidade da seqüência de motivos de viagem.

Em termos de relevância ou importância das variáveis de uso do solo, pode-se afirmar que estas (selecionadas três vezes pelo algoritmo CART na construção da AD) delimitaram principalmente a escolha de destinos (mais próximos ou mais distantes) para realização de viagens ao trabalho. Assim, fica implícito que a chance de indivíduos conseguirem empregos na proximidade de suas residências aumenta com a quantidade de oportunidades aí existentes, permitindo que viagens ao trabalho sejam mais curtas. A importância das variáveis de uso do solo na amostra de industriários será ratificada no

próximo capítulo, onde se encontram a aplicação da Regressão Linear Múltipla (RLM) e a medida da significância estatística de tais variáveis através dos coeficientes estimados e estatística *t*.

Esta seção, referente à discussão e análise de resultados obtidos com amostra de industriários, mostrou a influência de cada variável independente na escolha da seqüência de viagens. Nas próximas seções há discussões a respeito dos resultados obtidos nas demais amostras. No entanto, ressalta-se que as próximas análises destacarão apenas resultados diferentes dos obtidos anteriormente, que representam características particulares de cada amostra analisada.

### 8.3 TRABALHADORES NO SETOR COMERCIAL

Assim como na amostra composta por industriários, através da aplicação da AD na amostra composta por comerciários (6.019 indivíduos), utilizando os mesmos critérios para partição dos dados, foram obtidas 10 folhas.

Algumas variáveis, selecionadas no caso de trabalhadores do setor industrial, também foram consideradas importantes para o caso daqueles indivíduos que trabalham no comércio. No entanto, houve inversão na ordem de seleção das variáveis, de acordo com a importância de cada uma delas para segregação dos dados. Entende-se aqui como importante, aquela variável que consegue explicar uma maior parcela da variância observada na variável dependente. A Tabela 8.5 sintetiza a construção da hierarquia da árvore, mostrando as variáveis selecionadas em cada partição dos dados e sua influência nas seqüências das viagens.

Tabela 8.5 Variáveis selecionadas na construção da AD para comerciários

| <b>Partição</b>   | <b>Variável selecionada</b>                                                | <b>Influência principal</b>      |
|-------------------|----------------------------------------------------------------------------|----------------------------------|
| Primeira partição | Estuda                                                                     | Seqüência de atividades          |
| Segunda partição  | Nº Provável de CNH no domicílio                                            | Seqüência de modos de transporte |
|                   | Parcela acumulada de empregos no comércio até 5 km do centróide da origem  | Seqüência de destinos            |
| Terceira partição | Usa Vale Transporte                                                        | Seqüência de modos de transporte |
| Quarta Partição   | Trabalha                                                                   | Seqüência de atividades          |
|                   | Parcela acumulada de empregos no comércio até 15 km do centróide da origem | Seqüência de destinos            |
| Quinta Partição   | Renda Familiar                                                             | Seqüência de modos de transporte |
|                   | Nº de automóveis no domicílio                                              | Seqüência de modos de transporte |

A Tabela 8.5 mostra as variáveis selecionadas pelo algoritmo CART e a sua principal influência no comportamento subjacente ao encadeamento de viagens. Muitas das variáveis apresentadas foram analisadas na seção anterior. O escopo da atual seção é apresentar alguns resultados inerentes à amostra de comerciários.

Fazendo uma análise global dos resultados obtidos para a amostra de trabalhadores no setor do comércio pode-se afirmar que:

- De uma forma geral, os indivíduos realizam, com maior frequência, viagens curtas ("111"). Pode-se explicar tal resultado com base na maior dispersão geográfica de atividades comerciais na RMSP. Provavelmente, indivíduos que trabalham no setor do comércio residem próximos aos seus respectivos locais de trabalho, pois a dispersão de empregos no setor comercial na área de estudo é grande;
- São selecionadas pelo algoritmo CART, variáveis como "TRABALHA", por exemplo, exercendo influência na escolha da seqüência de motivos de viagem. Esta variável teve poder discriminante devido a características da amostra, pois há maior número de trabalhadores autônomos no setor comercial. Assim, a variável "TRABALHA" apresenta duas categorias predominantes na atual amostra (1 - Assalariados; 2 - Autônomos);
- As variáveis de uso do solo não interferem tanto na escolha de padrões de viagem (principalmente considerando escolhas de destinos). Este resultado já era esperado, uma vez que empregos no setor comercial estão bastante dispersos e não se observam pólos de concentração como no caso do setor industrial. Observa-se influência de variáveis de uso do solo no caso de indivíduos que trabalham no comércio e também estudam (Nó 6 e 7).

A Figura 8.6 ilustra a AD obtida para comerciários, onde são destacadas as folhas que serão aqui analisadas (Nó terminal 18, 19, 6 e 7).

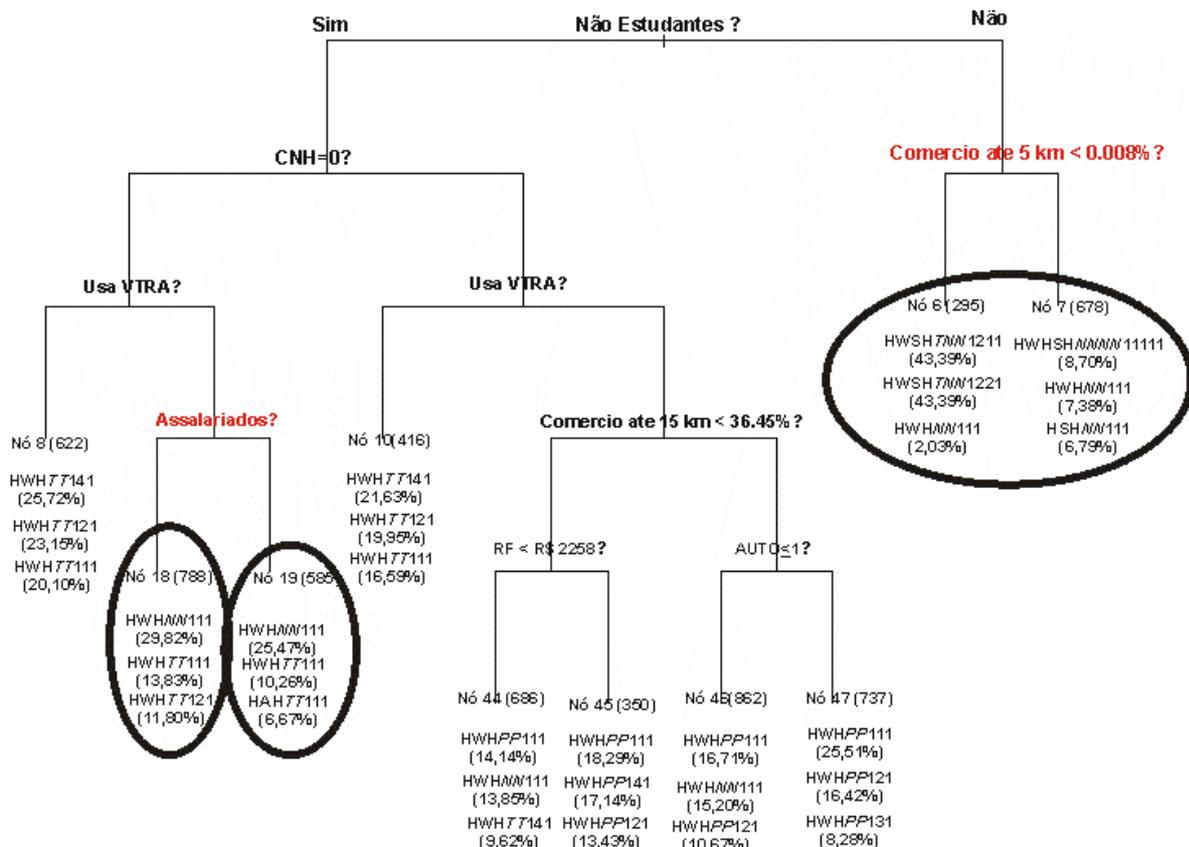


Figura 8.6 AD para amostra de comerciantes

Os indivíduos que compõem as folhas 18 e 19 possuem as seguintes características: (1) Não estudam; (2) Provavelmente nenhum membro da família possui carteira de habilitação; (3) Não utilizam Vale Transporte. Diferenciam-se pela variável "TRABALHA".

Assim, os indivíduos que formam o nó 18 são assalariados, enquanto que aqueles que fazem parte do nó 19 são autônomos. Intuitivamente, pode-se dizer que trabalhadores autônomos não têm a necessidade de realizar viagens diárias ao trabalho. Desta maneira, tais indivíduos poderão realizar outras atividades não relacionadas ao trabalho no período do dia. Comparando os dois grupos de trabalhadores, verificam-se maiores frequências considerando a seqüência de atividades HAH no grupo de trabalhadores autônomos (Nó 19), bem como maior ocorrência de seqüência de atividades HWH para o grupo de assalariados (Nó 18). A Tabela 8.6 mostra a frequência de ocorrência de alguns padrões nas duas referidas folhas. As frequências foram extraídas do relatório gerado pelo *software* e encontram-se no texto anexo.

Tabela 8.6 Diferenças nas seqüências de atividades entre as folhas 18 e 19

| Padrões  | Folhas       |           |
|----------|--------------|-----------|
|          | 18           | 19        |
|          | Assalariados | Autônomos |
| HWHWW111 | 29,82%       | 25,47%    |
| HWH7T111 | 13,83%       | 10,26%    |
| HAH7T111 | 1,14%        | 6,67%     |
| HAHWW111 | 1,02%        | 5,30%     |
| HAH7T121 | 0,25%        | 3,76%     |

Observando a Figura 8.6, verifica-se a influência da variável de uso do solo “PARCELA ACUMULADA DE EMPREGOS NO COMÉRCIO ATÉ 5 KM DO CENTRÓIDE DA ORIGEM” para o caso de indivíduos que trabalham no setor comercial e também estudam (Folhas 6 e 7).

Os indivíduos que compõem o nó terminal 6, ou seja, residem em zonas de tráfego com uma parcela muito baixa de empregos no comércio até 5 km (Comércio até 5 km < 0,008%) realizam, predominantemente, viagens ao trabalho com distância superior a 5 km. A localização da escola, no entanto pode ser próxima ao local de trabalho ou à residência. Os dois padrões predominantes neste caso, mostrados na Figura 8.6 ilustram estas situações:

- HWSH7MM1211 (43,39%): indivíduos realizam três viagens, sendo a primeira com motivo trabalho, utilizando transporte público, com distância entre 5 a 10 km da residência. Já a segunda viagem tem motivo escola e é realizada a pé. A escola, no entanto localiza-se até 5 km da residência. A última viagem é de retorno ao domicílio;
- HWSH7MM1221 (43,39%): neste caso a escola pode estar situada próxima ao trabalho, a uma distância entre 5 e 10 km da residência.

Já no caso dos indivíduos que compõe o nó 7, tanto a localização do trabalho quanto a localização da escola são próximas a residência dos indivíduos. Neste exemplo, há maior oferta de empregos no comércio nas redondezas (Comércio até 5 km  $\geq$  0,008%), assim as viagens ao trabalho têm uma distância inferior ou igual a 5 km (HWSH7MMM1111).

#### 8.4 TRABALHADORES NO SETOR DE SERVIÇOS

Utilizando os mesmos critérios para partição dos dados, através da aplicação da AD foram obtidos 10 nós terminais para a amostra composta por trabalhadores no setor de serviços (8.286 indivíduos).

Para as três amostras de trabalhadores, as variáveis mais importantes selecionadas pelo algoritmo CART foram as mesmas (Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO; ESTUDA e USA VALE TRANSPORTE). No entanto, os resultados obtidos para amostras composta por trabalhadores no setor de serviços foram bem similares aos obtidos para industriários. Os resultados provenientes da amostra de comerciários eram menos semelhantes, pois trabalhadores no setor comercial deslocam-se predominantemente nas suas proximidades, inclusive para realizar atividades relacionadas ao trabalho. A Tabela 8.7 sintetiza a construção da hierarquia da árvore, mostrando as variáveis selecionadas em cada partição dos dados e sua influência nas seqüências das viagens para o caso de trabalhadores no setor de serviços. Nota-se a semelhança dos resultados obtidos anteriormente, especialmente no caso de trabalhadores na indústria.

Tabela 8.7 Variáveis selecionadas na construção da AD para trabalhadores no setor de serviços

| <b>Partição</b>   | <b>Variável selecionada</b>                                                         | <b>Influência principal</b>                               |
|-------------------|-------------------------------------------------------------------------------------|-----------------------------------------------------------|
| Primeira partição | Nº Provável de CNH no domicílio                                                     | Seqüência de modos de transporte                          |
| Segunda partição  | Estuda                                                                              | Seqüência de atividades                                   |
| Terceira partição | Usa Vale Transporte                                                                 | Seqüência de modos de transporte                          |
| Quarta Partição   | Parcela acumulada de empregos no setor de serviços até 5 km do centróide da origem  | Seqüência de destinos                                     |
|                   | Parcela acumulada de empregos no setor de serviços até 15 km do centróide da origem | Seqüência de destinos                                     |
| Quinta Partição   | Sexo                                                                                | Seqüência de modos de transporte<br>Seqüência de destinos |
|                   | Nº de automóveis no domicílio                                                       | Seqüência de modos de transporte                          |

Considerando a construção da árvore (ramo a ramo), muito já foi discutido a respeito de cada uma das variáveis e sua influência em relação ao comportamento dos viajantes. Levando-se em conta a amostra de trabalhadores no setor de serviços, pode-se afirmar que neste caso também há importância das variáveis de uso do solo na escolha da seqüência de destinos. Tais variáveis não

foram tão relevantes para a amostra de comerciários. Contudo, para o caso de trabalhadores no setor de serviços, observa-se a sua influência quando se verificam os grupos de indivíduos que correspondem às folhas 18, 19, 52, 53, 54 e 55 (Figura 8.7).

Tais resultados podem ser explicados com base na distribuição geográfica das atividades no setor de serviços na RMSP. Apesar de os empregos no setor de serviços não estarem tão concentrados na RMSP quanto os empregos no setor industrial, pode-se afirmar que há alguns pólos principais destas atividades. De certa forma, variáveis que caracterizam oportunidades (ou oferta) no setor de serviços exercem maior influência do que no caso dos comerciários.

Um resultado particular obtido com a atual amostra foi a influência da variável "SEXO" (Folhas 52 e 53 destacadas na Figura 8.7). Os indivíduos que compõem ambas as folhas possuem as seguintes características: (1) Possuem, em sua família, pelo menos um membro com carteira de habilitação; (2) Não estudam; (3) Não usam Vale Transporte; (4) Residem em zonas de tráfego com parcela acumulada de empregos no setor de serviços até 15 km do seu centróide menor que 37.15%. Diferenciam-se apenas pelo sexo, refletindo assim, diferenças de comportamentos relacionadas a tal variável.

Apesar de as mulheres estarem com uma participação no mercado de trabalho cada vez mais forte, os papéis que tradicionalmente são alocados a homens e mulheres permanecem. Mesmo que em diferentes proporções, as mulheres ainda continuam exercendo um papel de maior responsabilidade em relação a tarefas domésticas e afazeres relacionados ao domicílio, refletindo diferentes comportamentos de viagem em cada sexo.

Pode-se observar, através do exemplo que contém os nós 52 e 53, que as mulheres utilizam com menor frequência o automóvel em relação aos homens, além de realizar predominantemente viagens mais curtas ao trabalho.

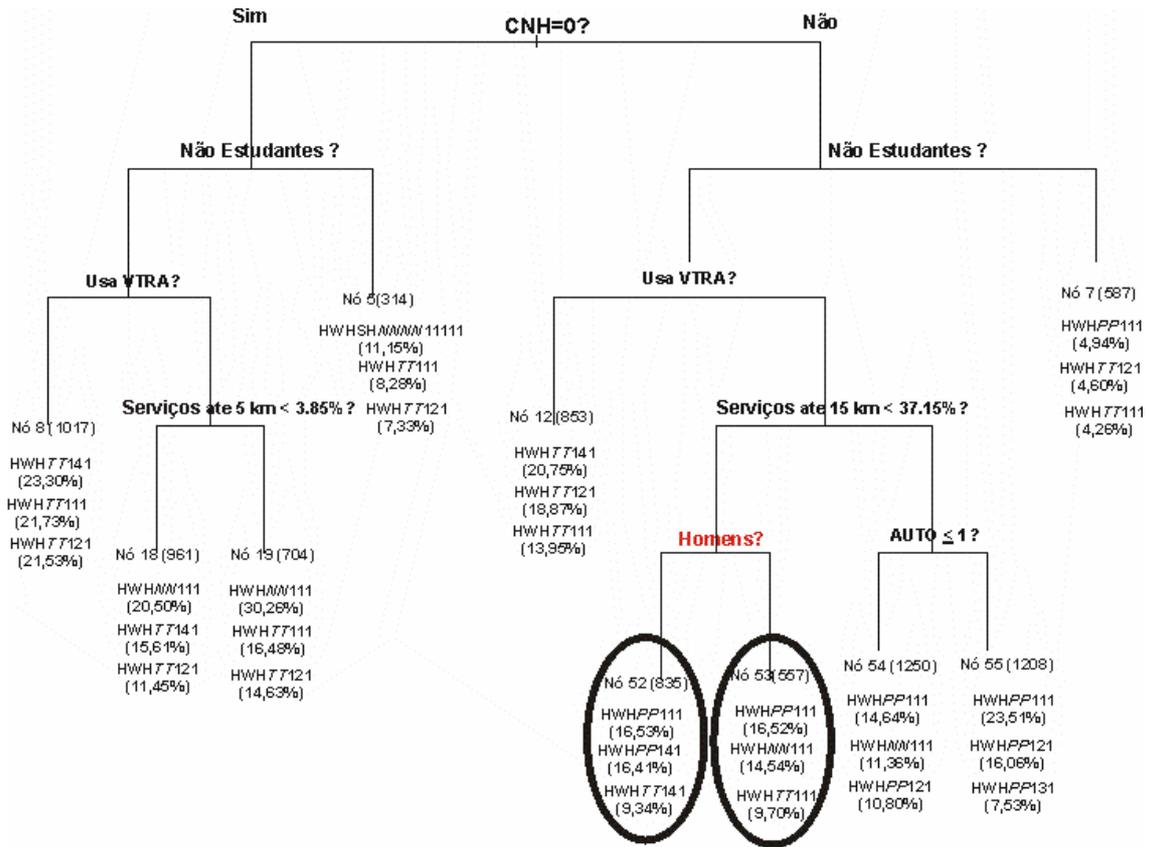


Figura 8.7 AD para amostra de trabalhadores no setor de serviços

Nas próximas seções será discutido a respeito do comportamento relacionado ao encadeamento de viagens para o caso das amostras de estudantes. Imagina-se que estudantes e trabalhadores tenham comportamentos relativos a viagens completamente distintos, a começar pela seqüência de atividades (já que a participação em atividades de subsistência é diferente para ambos os casos). Também se pode refletir a respeito da escolha da seqüência de destinos (pois há uma distribuição geográfica maior de escolas, pelo menos até o 1º grau pela RMS, sem contar com a possibilidade de escolha do local de estudo, que geralmente é próximo à residência). Assim, espera-se, de uma maneira geral, que estudantes realizem viagens mais curtas para as suas respectivas escolas.

Também há de se discutir a respeito das diferenças quanto às escolhas das seqüências de modos de transporte. Mesmo em domicílios onde há automóveis, este não pode ser considerado uma opção tão disponível quanto no caso de trabalhadores, a julgar pelas características inerentes a pelo menos duas das amostras de estudantes (até 1º grau e 2º grau): (1) Crianças ou adolescentes; (2)

Filhos; (3) Sem carteira de habilitação. Neste caso a opção automóvel estaria restrita apenas a passageiros.

Para análise mais detalhada da contribuição das variáveis de uso do solo no caso de estudantes, também se optou em desagregar amostras por grau de instrução, já que se observou que quanto maior o nível de escolaridade, maior a concentração de escolas em alguns pólos da RMSP. Assim, se há grande dispersão de escolas públicas até o 1º grau na região de estudo, há maior concentração no caso de escolas de 2º grau, e localizações ainda mais restritas no caso de universidades. Desta forma, espera-se que variáveis de uso do solo que representem parcela acumulada de escolas até determinadas distâncias a partir da zona de origem tenham grande influência no caso de universitários e uma contribuição relativamente fraca para o caso de estudantes matriculados até o 1º grau.

A fim de ratificar estas reflexões, discutidas acima, a respeito do comportamento relacionado a viagens para o caso de estudantes, encontram-se, a seguir análises realizadas para as três amostras.

### **8.5 ESTUDANTES ATÉ O 1º GRAU**

Através da aplicação da AD foram obtidas 9 folhas para a amostra composta por estudantes até o 1º grau (13.733 indivíduos). Os grupos finais obtidos são considerados “homogêneos” em relação à variável dependente e foram caracterizados segundo as variáveis “RENDA FAMILIAR”, “Nº DE AUTOMÓVEIS NO DOMICÍLIO”, “IDADE” e “PARCELA ACUMULADA DE MATRÍCULAS ATÉ O 1º GRAU ATÉ 5 KM DO CENTRÓIDE DA ZONA DE ORIGEM”.

A Tabela 8.8 sintetiza a construção da hierarquia da árvore obtida para a amostra de estudantes matriculados até o 1º grau, mostrando as variáveis selecionadas em cada partição dos dados e sua influência nas seqüências das viagens.

Tabela 8.8 Variáveis selecionadas na construção da AD para estudantes até o 1º grau

| Partição          | Variável selecionada                                                             | Influência principal                                      |
|-------------------|----------------------------------------------------------------------------------|-----------------------------------------------------------|
| Primeira partição | Renda Familiar                                                                   | Seqüência de modos de transporte                          |
| Segunda partição  | Nº de automóveis no domicílio                                                    | Seqüência de modos de transporte                          |
| Terceira partição | Idade                                                                            | Seqüência de modos de transporte<br>Seqüência de destinos |
| Quarta Partição   | Parcela acumulada de matrículas<br>até o 1º grau até 5 km do centróide da origem | Seqüência de destinos                                     |
|                   | Idade                                                                            | Seqüência de modos de transporte<br>Seqüência de destinos |

Em relação à amostra composta por estudantes até o 1º grau, pode-se fazer algumas análises gerais a respeito dos resultados obtidos com o processamento da AD:

- Em todos os casos (com exceção do nó 7) o padrão de viagem predominante é o HSH/MM11. Quanto à seqüência de atividades, é perfeitamente compreensível a maior freqüência para a seqüência HSH devido às características da amostra, que é composta exclusivamente por estudantes. Estes realizam então, com maior predominância, seqüência de atividades relacionadas à ESCOLA e à RESIDÊNCIA. Em relação à seqüência de modos de transporte (NN) e seqüência de destinos ("111") justifica-se tal comportamento principalmente pela distribuição dispersa de escolas públicas de ensino básico pela RMSP. Assim, os indivíduos geralmente podem escolher escolas próximas às suas residências. Por esta razão realizam viagens à escola a pé e tais viagens têm uma distância inferior a 5 km;
- A variável "RENDA FAMILIAR" foi considerada a mais importante para explicar a variabilidade do conjunto de dados, influenciando principalmente a seqüência de modos de transporte. Indivíduos que possuem uma renda familiar superior ou igual a R\$ 1690, com automóveis no domicílio, possuem maior probabilidade de realizar viagens à escola como passageiros de automóveis;
- A variável "IDADE" foi considerada bastante importante pelo algoritmo CART, delimitando diferenças de comportamento em relação à escolha modal e escolha de destinos. A variável foi selecionada quatro vezes pelo algoritmo, considerando os critérios adotados para partição dos dados (desvio 0.15 e mínimo de 50 observações por nó terminal). Conclui-se que crianças mais velhas

geralmente realizam viagens à escola sozinhas e muitas vezes utilizam transporte coletivo. Também se observa maior frequência de viagens acima de 5 km no grupo de crianças mais velhas, principalmente quando comparadas às crianças mais novas. Este fato será corroborado mais adiante pela análise das folhas 18 e 19;

- Como esperado, o grupo de variáveis de uso do solo não foi considerado relevante pelo algoritmo CART, foi selecionada apenas uma variável uma única vez, esta exerceu influência na escolha de destinos. O caso onde se observou a relação entre uso do solo e padrões de viagens na amostra de estudantes até o 1º grau (folhas 16 e 17) será discutido mais adiante;
- No caso do conjunto de dados aqui estudado, já se esperava maior influência das variáveis socioeconômicas nos padrões de viagens, haja vista que as variáveis relacionadas à participação em atividades não possuem nenhum poder discriminatório neste exemplo (todos os indivíduos estudam e não trabalham) e variáveis de uso do solo não são consideradas tão relevantes na escolha dos itinerários de viagens, porque além de haver oferta de escolas até o 1º grau distribuída por toda área de estudo, os indivíduos optam por estudar nas redondezas da residência. Além disso, na rede estadual de ensino do Estado de São Paulo, há recomendação para que as matrículas sejam feitas nas escolas próximas ao endereço de residência do aluno. Também é possível fazer matrículas em unidade próxima ao local de trabalho dos pais.

A Figura 8.8, a seguir, ilustra a AD para amostra de estudantes matriculados até o 1º grau, destacando as folhas que serão aqui analisadas (folhas 16; 17; 18 e 19).

Os indivíduos que fazem parte das folhas 16 e 17 possuem as seguintes características: (1) Renda Familiar inferior a R\$1.690; (2) Sem automóveis no domicílio; (3) Idade inferior ou igual a 10 anos. Diferenciam-se pela variável "PARCELA ACUMULADA DE MATRÍCULAS ATÉ O 1º GRAU ATÉ 5 KM DA ORIGEM". Este é o único exemplo onde se encontram relações entre variáveis de uso do solo e escolha de padrões de viagens na amostra composta por estudantes matriculados até o 1º grau.

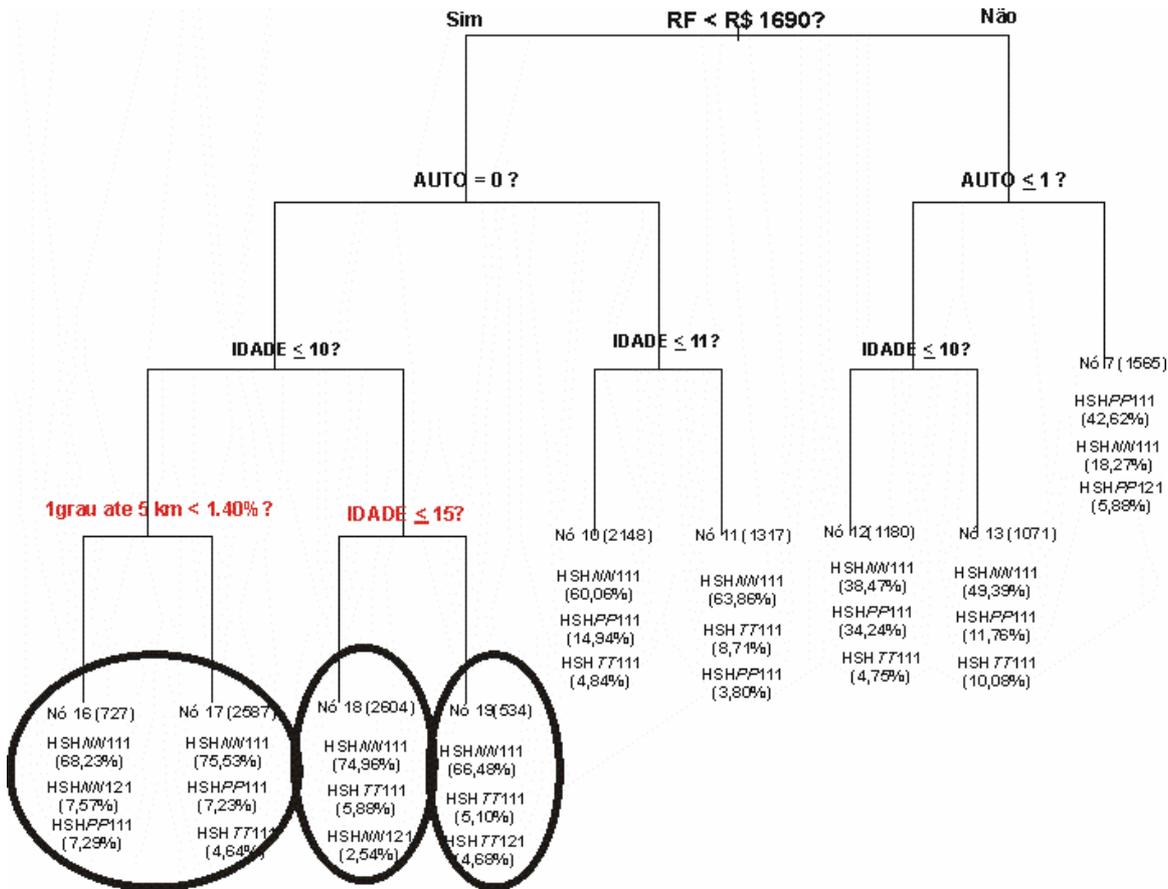


Figura 8.8 AD para amostra de estudantes matriculados até o 1º grau

Indivíduos que residem em domicílios situados em zonas de tráfego onde há parcela acumulada de escolas até o 1º grau inferior a 1.40% (nó 16) realizam viagens mais longas com motivo escola, com maior frequência, quando comparados àqueles indivíduos que residem em domicílios onde a parcela acumulada de escolas até o 1º grau é superior ou igual a 1.40% (nó 17). Em ambos os casos o padrão predominante é HWH/M11(68,23% - nó 16 e 75,53% - nó 17), contudo, ao analisar escolhas de destinos entre 5 e 10 km da zona de origem, observa-se maiores frequências de ocorrência na folha 16. Justificativa para isto é a influência da referida variável de uso do solo na escolha da seqüência de destinos.

Mesmo no caso de estudantes matriculados até o 1º grau, onde há oferta de escolas em um grande número de localizações na RMSP, a influência do uso do solo manifesta-se quando esta parcela acumulada de escolas é inferior a 1.40%. A Tabela 8.9 apresenta alguns padrões de viagens,

com distância entre 5 e 10 km, nas folhas 16 e 17 e suas respectivas freqüências de ocorrência. Tais freqüências foram extraídas do relatório gerado pelo S-PLUS 6.1 e encontram-se no texto anexo (Anexo D).

Tabela 8.9 Diferenças nas seqüências de destinos entre as folhas 16 e 17

| Padrões  | Folhas                   |                          |
|----------|--------------------------|--------------------------|
|          | 16                       | 17                       |
|          | 1º grau até 5 km < 1.40% | 1º grau até 5 km ≥ 1.40% |
| HSHWN121 | 7,57%                    | 0,97%                    |
| HSHPP121 | 1,65%                    | 0,31%                    |
| HSHJT121 | 1,51%                    | 0,77%                    |

Considerando agora a influência da variável "IDADE", tem-se o exemplo dos nós terminais 18 e 19. Os indivíduos que compõem ambos os nós possuem as seguintes características: (1) Renda Familiar inferior a R\$ 1690; (2) Sem automóveis no domicílio. Diferenciam-se pela variável "IDADE". Os indivíduos que fazem parte do nó 18 têm idade entre 10 e 15 anos, enquanto que aqueles indivíduos que compõem o nó 19 têm idade superior a 15 anos.

No caso das crianças mais velhas (nó 19) observa-se o uso de transporte coletivo com maior freqüência, além de escolha de destinos situados acima de 5 km da origem. Pitombo (2003) já havia notado diferenças comportamentais em relação a viagens através da variável "IDADE" em amostras de estudantes. Na ocasião a autora observou que crianças mais novas, em idade escolar, realizam viagens à escola durante o dia. Já as mais velhas e com o mesmo grau de instrução das mais novas estudam no período da noite. No presente caso, espera-se que crianças mais velhas tenham maior "liberdade" para realização das viagens, sem a necessidade de serem passageiros ou realizar tais viagens na companhia de pais ou responsáveis. A Tabela 8.10 mostra diferenças de escolha de alguns padrões de viagens entre os grupos (folha 18 e 19).

Tabela 8.10 Diferenças nas seqüências de destinos e de modos de transporte entre as folhas 18 e 19

| Padrões  | Folhas          |            |
|----------|-----------------|------------|
|          | 18              | 19         |
|          | 10 < Idade ≤ 15 | Idade > 15 |
| HSHTT111 | 5,88%           | 5,10%      |
| HSHTT121 | 1,61%           | 4,68%      |
| HSHTT131 | 0,58%           | 2,06%      |
| HAHTT111 | 0,73%           | 1,69%      |
| HAHTT121 | 0,31%           | 1,12%      |

Em síntese, a análise dos resultados obtidos através do processamento da AD para amostra de estudantes matriculados até o 1º grau pode ser encontrada na Tabela 8.8 apresentada no início desta seção. É observada a seleção de cada variável na partição dos dados e sua respectiva influência na seqüência de viagens. A seguir, é descrita a análise dos resultados obtidos com amostra de estudantes matriculados no 2º grau.

## 8.6 ESTUDANTES NO 2º GRAU

Foram obtidas 8 folhas para a amostra composta por estudantes no 2º grau (2.125 indivíduos). Os grupos finais de indivíduos foram caracterizados segundo as variáveis "RENDA FAMILIAR", "Nº DE AUTOMÓVEIS NO DOMICÍLIO", "PARCELA ACUMULADA DE MATRÍCULAS NO 2º GRAU ATÉ 5 KM DO CENTRÓIDE DA ZONA DE ORIGEM", "PARCELA ACUMULADA DE MATRÍCULAS NO 2º GRAU ATÉ 10 KM DO CENTRÓIDE DA ZONA DE ORIGEM" e "PARCELA ACUMULADA DE MATRÍCULAS NO 2º GRAU ATÉ 20 KM DO CENTRÓIDE DA ZONA DE ORIGEM".

A Tabela 8.11 traz o sumário da construção da AD obtida (ramo a ramo).

Tabela 8.11 Variáveis selecionadas na construção da AD para estudantes no 2º grau

| <b>Partição</b>   | <b>Variável selecionada</b>                                                 | <b>Influência principal</b>      |
|-------------------|-----------------------------------------------------------------------------|----------------------------------|
| Primeira partição | Renda Familiar                                                              | Seqüência de modos de transporte |
| Segunda partição  | Parcela acumulada de matrículas no 2º grau até 5 km do centróide da origem  | Seqüência de destinos            |
|                   | Nº de automóveis no domicílio                                               | Seqüência de modos de transporte |
| Terceira partição | Parcela acumulada de matrículas no 2º grau até 5 km do centróide da origem  | Seqüência de destinos            |
|                   | Renda Familiar                                                              | Seqüência de modos de transporte |
|                   | Parcela acumulada de matrículas no 2º grau até 20 km do centróide da origem | Seqüência de destinos            |
| Quarta Partição   | Parcela acumulada de matrículas no 2º grau até 10 km do centróide da origem | Seqüência de destinos            |

Pode-se perceber que os resultados obtidos foram muito similares àqueles anteriormente obtidos com amostra de estudantes matriculados até o 1º grau. A diferença básica é que as variáveis de uso do solo foram selecionadas mais vezes (quatro vezes no total) pelo algoritmo CART para segregação dos dados. Deve-se isso à maior concentração de escolas de 2º grau na RMSP. A oferta de escolas de 2º grau não é tão dispersa geograficamente pela região de estudo quanto a oferta de escolas de 1º grau.

No entanto, em termos de padrões de viagens predominantes, não houve grandes mudanças (HSHVM111; HSHPP111; HSHTT111). Ainda que localização de escolas de 2º grau seja mais pontual quando comparada à localização de escolas de 1º grau, estudantes continuam a optar por escolas mais próximas às suas residências. Assim, as viagens realizadas são geralmente curtas (até 5 km da origem).

Um ponto relevante é a não seleção da variável "IDADE" para a segregação dos dados. No caso de estudantes de 2º grau, estes têm idade superior ou igual a 15 anos e já realizam viagens sozinhos (sem companhia de pais ou responsáveis), com transporte coletivo. Desta forma, não se espera variações no comportamento de viagens comparando indivíduos de 15 ou 17 anos, por exemplo.

A seguir são discutidos os resultados obtidos através da amostra composta por universitários.

## 8.7 ESTUDANTES UNIVERSITÁRIOS

Foram obtidas 8 folhas para a amostra composta por universitários (560 indivíduos). As classes finais de indivíduos foram caracterizadas somente segundo as variáveis de uso do solo e renda familiar. A Tabela 8.12 sintetiza a construção da AD obtida ramo a ramo e a influência de cada uma das variáveis nas seqüências de viagens.

Tabela 8.12 Variáveis selecionadas na construção da AD para estudantes universitários

| <b>Partição</b>   | <b>Variável selecionada</b>                                                       | <b>Influência principal</b>      |
|-------------------|-----------------------------------------------------------------------------------|----------------------------------|
| Primeira partição | Parcela acumulada de matrículas em universidades até 10 km do centróide da origem | Seqüência de destinos            |
| Segunda partição  | Parcela acumulada de matrículas em universidades até 10 km do centróide da origem | Seqüência de destinos            |
|                   | Renda Familiar                                                                    | Seqüência de modos de transporte |
| Terceira partição | Renda Familiar                                                                    | Seqüência de modos de transporte |
|                   | Parcela acumulada de matrículas em universidades até 5 km do centróide da origem  | Seqüência de destinos            |

Há alguns resultados particulares em relação à amostra de universitários que são discutidos nos pontos seguintes:

- Os padrões predominantes têm a mesma seqüência de atividades (HSH), variando a seqüência de modos de transporte e destinos escolhidos. Esta amostra foi a que apresentou maior número de escolha de destinos distantes da residência ("141", "131", "121"). Pode-se justificar este fato pela ausência de oferta de universidades por toda área de estudo. As universidades concentram-se em alguns pólos (conforme mostrado no Capítulo 4), obrigando os indivíduos a se deslocarem por maiores distâncias para satisfazer necessidades de estudos em universidades. Neste caso, na maioria das vezes não há como os indivíduos optarem por universidades que sejam próximas às suas residências.
- Esta amostra foi a que apresentou maior influência das variáveis de uso do solo nas escolhas dos destinos. Há vários exemplos (nó 12, 13, 14 e 15). Este comportamento já era intuitivamente esperado devido às características da distribuição geográfica de universidades na RMSP. Assim, a parcela acumulada de universidades na zona de origem ou nas suas vizinhanças interfere fortemente

nas escolhas de padrões de viagens. Este fato será corroborado mais adiante (Capítulo 9) com a aplicação da RLM para a atual amostra.

- A amostra de universitários tem uma característica particular relacionada à variável “RENDA FAMILIAR”. Aqueles indivíduos que estudam em universidades possuem uma renda mais alta, principalmente quando comparados aos estudantes de 1º e 2º graus. Desta forma, o ponto de corte desta variável contínua foi bem mais alto do que nas duas últimas amostras (RF < R\$ 4.276; RF < 2157 e RF < 3950). Esta variável delimitou o uso mais freqüente ou não de automóveis para as viagens realizadas à universidade. Indivíduos com maior Renda Familiar (e possivelmente maior número de automóveis no domicílio) utilizam predominantemente modo particular motorizado nos seus padrões de viagens (nó 9; nó 11).
- A variável selecionada para primeira partição dos dados “PARCELA ACUMULADA DE MATRÍCULAS EM UNIVERSIDADES ATÉ 10 KM DO CENTRÓIDE DA ZONA DE ORIGEM” foi de essencial importância para escolha de destinos pelos indivíduos. Aqueles indivíduos que residem em zonas de tráfego onde a parcela acumulada de matrículas em universidades até 10 km da origem é inferior a 44.72%, predominantemente realizam viagens acima de 10 km (“141” – nó 8; “141” – nó 9; “131” – nó 10; “131” – nó 11). Já aqueles indivíduos que residem em zonas de tráfego com parcela acumulada de matrículas em universidades até 10 km superior a 44.72%, realizam viagens até 10 km da origem, já que se subentende que há boa oferta de universidades até esta distância a partir das respectivas residências (“121” – nó 12; “111” – nó 13; “121” – nó 14; “111” – nó 15).

A Figura 8.9, a seguir, ilustra a AD obtida para amostra composta por universitários, destacando-se aqueles nós terminais que serão estudados mais detalhadamente.

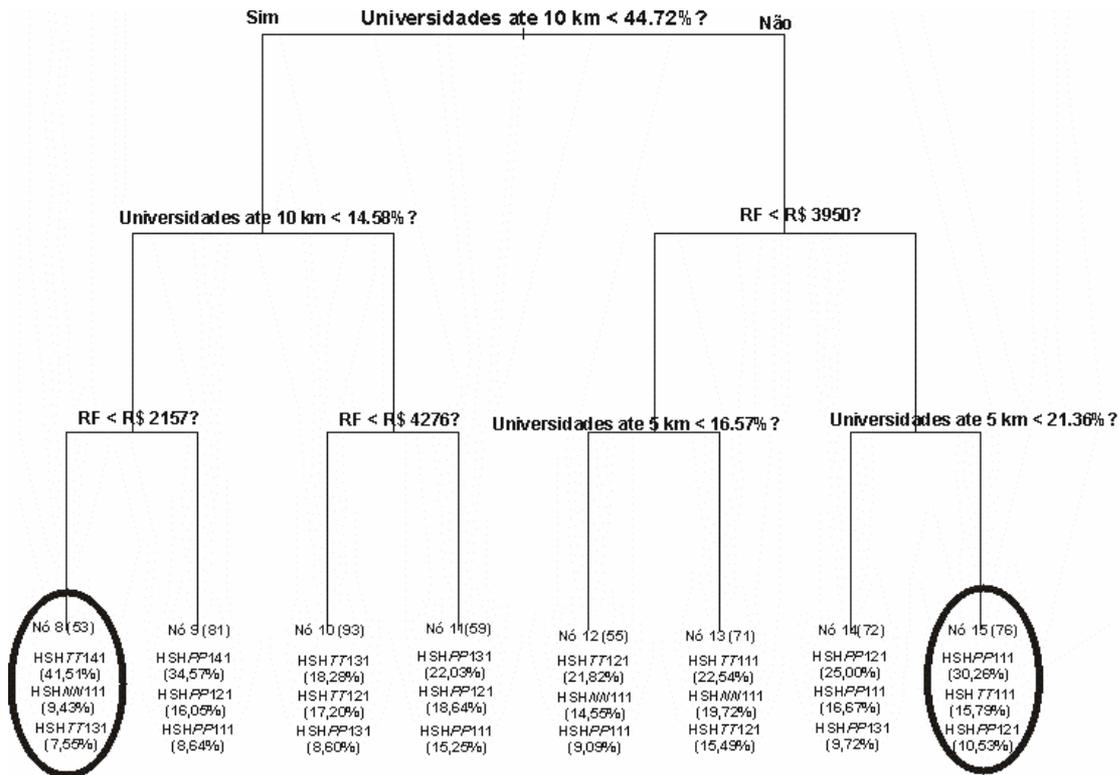


Figura 8.9 AD para amostra de estudantes universitários

A Folha 8 é formada por indivíduos que possuem as seguintes características: (1) residem em zonas de tráfego com parcela acumulada de matrículas em universidades até 10 km da origem inferior a 14,58%; e (2) Renda Familiar inferior a R\$ 2157. Como esperado, a variável “RENDA FAMILIAR” afeta principalmente a escolha da seqüência de modos de transporte. Assim, predominantemente, tais indivíduos realizam viagens à universidade através de transporte coletivo ou a pé (TTT; NN).

Quanto à baixa parcela de escolas até 10 km da origem, isto delimita a escolha de destinos. Indivíduos pertencentes a tal grupo predominantemente fazem viagens longas, superiores a 15 km (“141”). Fazendo uma análise detalhada deste nó terminal, pode-se observar claramente (em termos de zonas de tráfego) que os indivíduos residem em zonas com pouca oferta de universidades, e deslocam-se grandes distâncias para zonas com maiores oportunidades (ou seja, número de universidades). A Tabela 8.13 apresenta padrões predominantes de indivíduos que fazem parte desta classe, com características associadas às zonas de origem e destino.

Tabela 8.13 Análise detalhada dos deslocamentos dos indivíduos que compõem o nó 8

| Padrão Predominante (zonas de tráfego) |       | Residência |                     |                       | Destino - Universidade |                      |                       | Distância<br>Km |
|----------------------------------------|-------|------------|---------------------|-----------------------|------------------------|----------------------|-----------------------|-----------------|
|                                        |       | zona       | Nome                | Parcela_universidades | zona                   | Nome                 | Parcela_universidades |                 |
| <b>HSHAV158-158-158</b>                | 5,56% | 158        | Guarulhos           | 0,71%                 | 158                    | Guarulhos            | 0,71%                 | 1,34            |
| <b>HSHTT344-372-344</b>                | 3,70% | 344        | Ouro Fino Paulista  | 0,01%                 | 372                    | Mogi das cruces      | 3,20%                 | 25,33           |
| <b>HSHPP218-158-218</b>                | 1,85% | 218        | Parque Buturussu    | 0,01%                 | 158                    | Guarulhos            | 0,71%                 | 6,11            |
| <b>HSHPP229-105-229</b>                | 1,85% | 229        | Vila Camosina       | 0,00%                 | 105                    | Boa Vista            | 0,18%                 | 11,79           |
| <b>HSHPP229-224-229</b>                | 1,85% | 229        | Vila Camosina       | 0,00%                 | 224                    | Rio Verde            | 0,20%                 | 5,04            |
| <b>HSHPP349-31-349</b>                 | 1,85% | 349        | Jaceguava           | 0,00%                 | 31                     | Lins de Vasconcelos  | 0,67%                 | 24,45           |
| <b>HSHPT289-27-289</b>                 | 1,85% | 289        | José Bonifácio      | 0,00%                 | 27                     | Parque da Moóca      | 0,56%                 | 16,87           |
| <b>HSHTT176-17-176</b>                 | 1,85% | 176        | Parque Sta Madalena | 0,00%                 | 17                     | Consolação           | 5,75%                 | 16,16           |
| <b>HSHTT252-131-252</b>                | 1,85% | 252        | Centro Empresarial  | 0,00%                 | 131                    | Cidade Universitária | 7,47%                 | 11,3            |
| <b>HSHTT259-16-259</b>                 | 1,85% | 259        | Jardim Mitsutani    | 0,79%                 | 16                     | Masp                 | 2,15%                 | 15,68           |
|                                        |       |            | Média               | <b>0,15%</b>          |                        | Média                | <b>2,16%</b>          | <b>13,41</b>    |

A fim de fazer uma análise daqueles indivíduos que residem em zonas com maiores oportunidades, ou seja, maior parcela acumulada de matrículas em universidade nas zonas de origem e vizinhanças, investigou-se mais detalhadamente os deslocamentos dos indivíduos que pertencem à folha 15.

Tal grupo de indivíduos possui as seguintes características: (1) residem em zonas de tráfego com parcela acumulada de matrículas em universidades até 10 km da origem superior ou igual a 44,72%; e (2) Renda Familiar superior ou igual a R\$ 3950; (3) residem em zonas de tráfego com parcela acumulada de matrículas em universidades até 5 km da origem superior ou igual a 21.36%.

O valor alto da variável "RENDA FAMILAR" influencia a seqüência de modos de transporte escolhidos. Observa-se o uso predominante do automóvel para realização de viagens à universidade. Indivíduos que possuem maiores rendas, possivelmente também possuem automóveis no domicílio, aumentando a possibilidade de uso de modo particular motorizado.

Quanto à influência da variável de uso do solo, nota-se que os indivíduos realizam viagens curtas, pois há zonas próximas aos seus domicílios com alta oferta de universidades. A Tabela 8.14 detalha os deslocamentos (inclusive em termos de zonas de tráfego) dos indivíduos que pertencem ao nó 15. Verifica-se que a distância média das viagens é aproximadamente 2,30 km.

Tabela 8.14 Análise detalhada dos deslocamentos dos indivíduos que compõem o nó 15

| Padrão Predominante (zonas de tráfego) |       | Residência |                    |                       | Destino - Universidade |                    |                       | Distância<br>Km |
|----------------------------------------|-------|------------|--------------------|-----------------------|------------------------|--------------------|-----------------------|-----------------|
|                                        |       | zona       | Nome               | Parcela_universidades | zona                   | Nome               | Parcela_universidades |                 |
| <b>HSHPP38-37-38</b>                   | 2,56% | 38         | Higienópolis       | 0,86%                 | 37                     | Cardoso de Almeida | 4,30%                 | 1,64            |
| <b>HSHPP65-66-65</b>                   | 2,56% | 65         | Mirandópolis       | 0,26%                 | 66                     | Parque Ibirapuera  | 1,77%                 | 2,38            |
| <b>HSHPP75-37-75</b>                   | 2,56% | 75         | Vila Beatriz       | 0,00%                 | 37                     | Cardoso de Almeida | 4,30%                 | 3,09            |
| <b>HSH7710-29-10</b>                   | 2,56% | 10         | Cambuci            | 0,00%                 | 29                     | Vila Monumento     | 0,18%                 | 1,21            |
| <b>HSHMV20-21-20</b>                   | 2,56% | 20         | Canindé            | 0,00%                 | 21                     | Pan                | 0,91%                 | 0,72            |
| <b>HSHMV37-37-37</b>                   | 1,28% | 37         | Cardoso de Almeida | 4,30%                 | 37                     | Cardoso de Almeida | 4,30%                 | 0,77            |
| <b>HSHMV38-17-38</b>                   | 1,28% | 38         | Higienópolis       | 0,86%                 | 17                     | Consolação         | 5,75%                 | 1,79            |
| <b>HSHMV38-36-38</b>                   | 1,28% | 38         | Higienópolis       | 0,86%                 | 36                     | Pacaembu           | 5,63%                 | 1,88            |
| <b>HSHPP12-36-12</b>                   | 1,28% | 12         | Aclimação          | 0,00%                 | 36                     | Pacaembu           | 5,63%                 | 4,33            |
| <b>HSHPP12-37-12</b>                   | 1,28% | 12         | Aclimação          | 0,00%                 | 37                     | Cardoso de Almeida | 4,30%                 | 5,24            |
|                                        |       | Média      |                    |                       | Média                  |                    |                       | <b>2,31</b>     |
|                                        |       |            |                    |                       |                        |                    |                       | <b>0,71%</b>    |
|                                        |       |            |                    |                       |                        |                    |                       | <b>3,71%</b>    |

## 8.8 CONSIDERAÇÕES FINAIS

Neste capítulo foram analisados os resultados obtidos com o processamento das árvores de decisão para as seis amostras desagregadas. Verificou-se influência de algumas variáveis independentes nas seqüências de viagens realizadas pelos indivíduos (seja na seqüência de atividades, na seqüência de modos de transporte ou na seqüência de destinos).

Quanto ao conjunto de variáveis de uso do solo ora proposto, verificou-se a maior influência de tais variáveis nas amostras compostas por industriários e universitários. Este fato justifica-se pela concentração geográfica de oferta de empregos na indústria e vagas nas universidades na RMSP.

Visando corroborar tais afirmações e buscando associar inferência estatística às variáveis independentes, a próxima etapa do método (Capítulo 9) consiste na aplicação da RLM para estimação de parâmetros que comprovem a contribuição das variáveis independentes na realização de determinados padrões de viagens. Em especial, espera-se verificar também a forte contribuição das variáveis de uso do solo nas amostras de industriários e universitários.

Desta forma, o capítulo seguinte descreve a aplicação da RLM em conjunto com a AD. Encontradas relações entre as variáveis envolvidas na análise, a próxima etapa consiste na mensuração da significância estatística de tais variáveis (etapa confirmatória da pesquisa).

## Capítulo

# 9 SIGNIFICÂNCIA ESTATÍSTICA DAS VARIÁVEIS INDEPENDENTES

---

*Depois da análise e discussão dos resultados obtidos com as árvores de decisão, é associar inferência estatística ao estudo. Este capítulo descreve a aplicação da Regressão Linear Múltipla (RLM), que foi a maneira encontrada para mensurar a significância das variáveis independentes, especialmente daquelas relacionadas ao uso do solo, propostas neste trabalho (objetivo secundário).*

### 9.1. APLICAÇÃO DA RLM EM CONJUNTO COM A AD

Conforme mencionado no Capítulo 3, a RLM foi aplicada no trabalho para fins complementares. O objetivo principal da sua aplicação é associar inferência estatística às variáveis independentes. Através da RLM é possível mensurar a significância de cada uma das variáveis independentes, especialmente das variáveis de uso do solo propostas.

O uso da RLM foi realizado após aplicação da Árvore de Decisão (AD). A AD traz resultados facilmente interpretáveis, no entanto, por se tratar de uma técnica de *Data Mining* (DM), não se tem inferência estatística associada a variáveis.

Para aplicação da RLM, a variável dependente é dada pela probabilidade de ocorrência dos dez padrões de viagens mais frequentes em cada uma das amostras. Quanto às variáveis independentes, cada uma delas foi dividida em classes e a cada classe foi associada uma variável *dummy*. Os valores para escolha das classes das variáveis *dummy* foram obtidos com a aplicação da AD. As etapas para aplicação da RLM e obtenção dos modelos lineares serão discutidas mais detalhadamente na próxima seção, que considera a amostra de trabalhadores no setor industrial. Uma síntese dessas etapas encontra-se ilustrada na Figura 9.1.

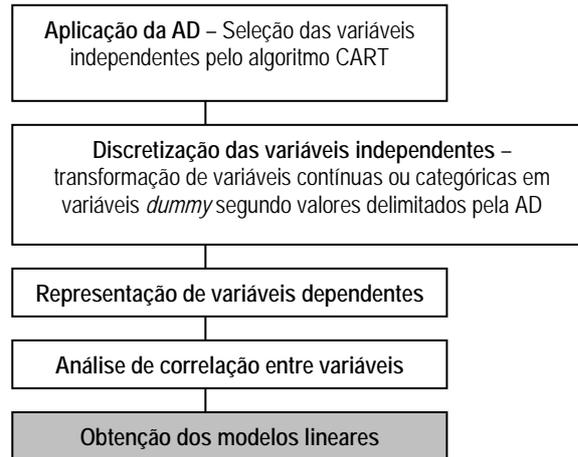


Figura 9.1 Etapas para aplicação da RLM no trabalho

Muito já foi discutido a respeito da importância da aplicação da RLM em conjunto com a AD, já que técnicas de DM, como árvores de decisão, não trazem inferência estatística associada a seus resultados. No entanto, algumas ressalvas devem ser feitas neste capítulo a respeito da apresentação dos resultados obtidos com a aplicação da RLM e os objetivos do trabalho.

Será apresentada apenas uma síntese dos modelos obtidos nas seis amostras estudadas no trabalho (60 modelos lineares no total), contendo os valores dos coeficientes estimados para cada uma das variáveis, a estatística *t* e os valores de *R* e *R*<sup>2</sup>.

A atual pesquisa não se propôs a fazer uma análise pormenorizada a respeito da qualidade dos 60 modelos lineares obtidos. Desta forma, a afirmação de que determinado modelo é razoável ou não, é baseada apenas nos valores obtidos para os coeficientes de determinação e dos demais parâmetros a serem apresentados neste capítulo.

Exemplos de análises adicionais, além do sumário dos modelos apresentados, tais como análise de resíduos para checar suposições de normalidade ou homocedasticidade, são apresentadas no texto anexo (ANEXO G).

Deve-se deixar claro que o objetivo não é checar minuciosamente a qualidade dos modelos lineares, e sim, propor um método para encontrar relações entre as variáveis envolvidas na análise, e, adicionalmente, mensurar a significância estatística de tais variáveis. Assim, a aplicação conjunta das técnicas de AD e RLM mostrou-se ser um meio viável para se atingir esse objetivo. No entanto, os

modelos devem ser robustos quanto aos pressupostos, para que valores de estatística t e coeficientes estimados sejam considerados significativos. Nas seções seguintes encontra-se a aplicação da RLM para cada uma das amostras da análise. No anexo G encontra-se análise mais detalhada de alguns modelos lineares obtidos.

## 9.2 TRABALHADORES NO SETOR INDUSTRIAL

No capítulo anterior, foram discutidos resultados obtidos, provenientes do processamento da AD para todas as amostras da análise. Naquela ocasião, observou-se a forte influência das variáveis de uso do solo, principalmente nas amostras compostas por trabalhadores no setor industrial e universitários. Também se verificou a influência de variáveis socioeconômicas como “RENDA FAMILIAR” (principalmente no caso do uso de transporte particular motorizado), “USA VALE TRANSPORTE”, “IDADE”, “Nº PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO”, etc.

A intenção é comprovar a contribuição das variáveis independentes selecionadas pelo algoritmo CART na escolha dos padrões de viagens através de coeficientes estimados e valores da estatística t associados ao conjunto de variáveis. Contudo, a finalidade principal concentra-se nas variáveis de uso do solo (objetivo secundário do trabalho).

Pelos resultados discutidos, no caso dos trabalhadores, espera-se que as variáveis de uso do solo possuam maior contribuição no caso dos industriários, seguidos dos trabalhadores no setor de serviços, e finalmente, comerciários. Já no caso dos estudantes, espera-se maior contribuição das mesmas variáveis na amostra composta por universitários, seguidos de estudantes matriculados no 2º grau, e por fim, estudantes matriculados até o 1º grau.

A seguir são descritas as representações das variáveis independentes e da variável dependente, bem como resultados obtidos com os modelos de regressão para o caso particular de trabalhadores no setor industrial. São discutidas as etapas representadas na Figura 9.1.

### 9.2.1 Discretização das variáveis independentes

Para aplicação da RLM, as variáveis contínuas foram divididas em classes com o intuito de reduzir o efeito da eventual não-linearidade entre variáveis independentes e variável dependente. Cada classe foi associada a uma variável *dummy* e os valores para escolha das classes das variáveis *dummy* foram obtidos com a aplicação da AD.

Assim, de acordo com AD final obtida e as variáveis selecionadas pelo algoritmo CART para amostra de industriários, tem-se as variáveis *dummy* para aplicação da RLM apresentadas na Tabela 9.1. São representadas as transformações das variáveis em variáveis binárias, assim como nova nomenclatura adotada para aplicação da RLM.

Tabela 9.1 Variáveis independentes para aplicação da RLM (industriários)

| Variáveis independentes                                                     |                                                 |                                                                      |                   |
|-----------------------------------------------------------------------------|-------------------------------------------------|----------------------------------------------------------------------|-------------------|
| Nomenclatura antiga                                                         | Original                                        | Dummy                                                                | Nova nomenclatura |
| Nº provável de carteiras de habitação no domicílio                          | CNH<br>(numérica - variando de 0 a 5)           | CNH = 0 (0)<br>CNH ≥ 1 (1)                                           | CNH               |
| Usa Vale Transporte                                                         | Usa VTRA<br>(categórica - 1 (Sim) e 2 (Não))    | Usa VTRA = 1 (0)<br>Usa VTRA = 2 (1)                                 | Usa VTRA          |
| Estuda                                                                      | ESTUDA<br>(categórica - 1 (Não) e 3 ou 4 (Sim)) | ESTUDA = 1 (0)<br>ESTUDA = 3 ou 4 (1)                                | ESTUDA            |
| Renda Familiar                                                              | RENDA FAMILIAR (RF)<br>(contínua - em reais)    | RF < R\$ 2970 (0)<br>RF ≥ R\$ 2970 (1)                               | RENDA             |
| Parcela acumulada de empregos na indústria até 5 km do centróide da origem  | Indústria até 5 km<br>(contínua - %)            | Indústria até 5 km < 3.22% (0)<br>Indústria até 5 km ≥ 3.22% (1)     | USO DO SOLO 1     |
| Parcela acumulada de empregos na indústria até 15 km do centróide da origem | Indústria até 15 km<br>(contínua - %)           | Indústria até 15 km < 32.31% (0)<br>Indústria até 15 km ≥ 32.31% (1) | USO DO SOLO 2     |
| Parcela acumulada de empregos na indústria até 10 km do centróide da origem | Indústria até 10 km<br>(contínua - %)           | Indústria até 10 km < 15.49% (0)<br>Indústria até 10 km ≥ 15.49% (1) | USO DO SOLO 3     |

### 9.2.2 Representação das variáveis dependentes

Para a representação das variáveis dependentes utilizadas na análise, foram realizadas as etapas ilustradas na Figura 9.2, sumariadas a seguir. O exemplo esquematizado é referente à amostra composta por trabalhadores no setor industrial (4.084 indivíduos).

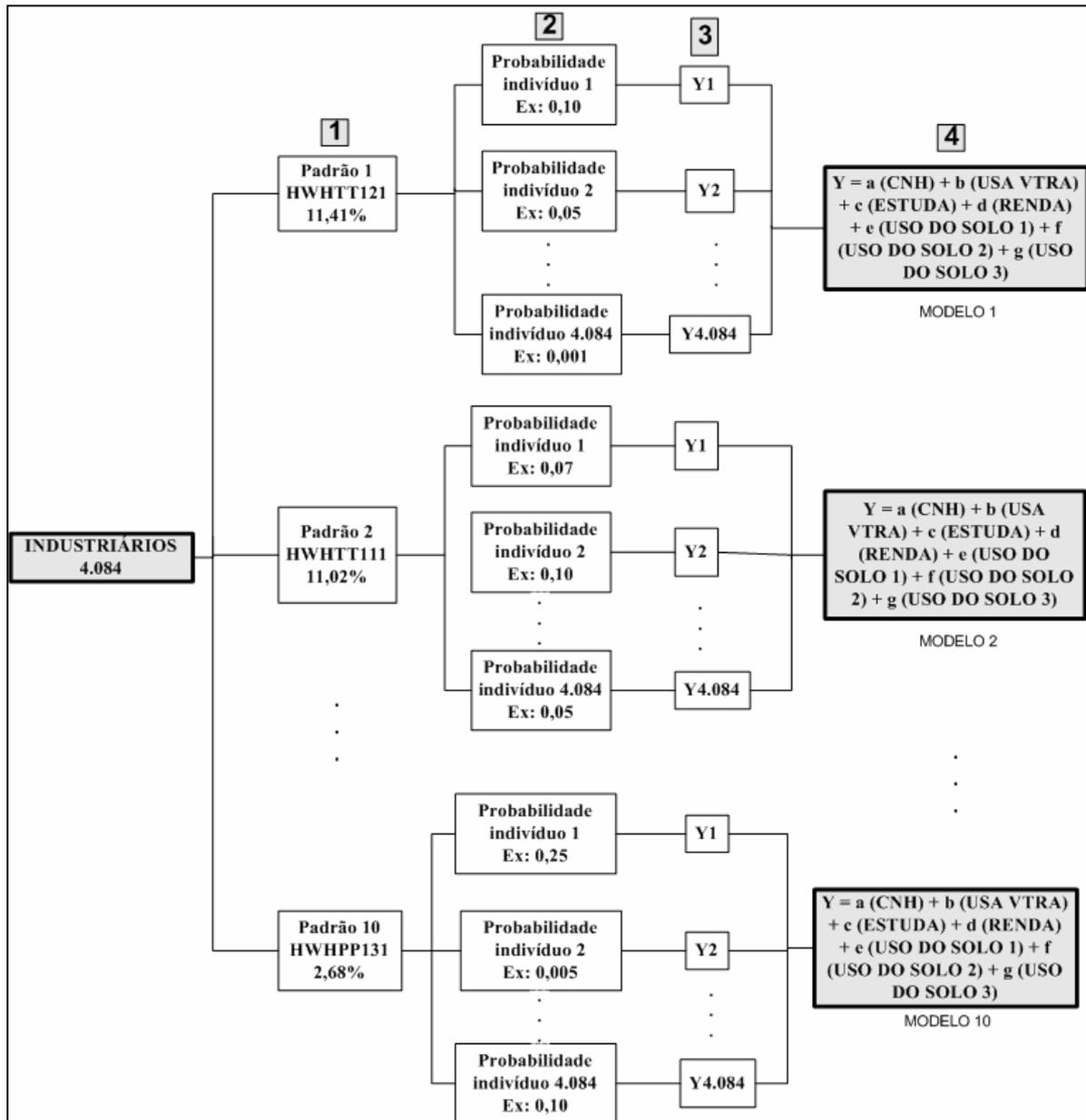


Figura 9.2 Etapas para representação das variáveis dependentes (exemplo para industriários)

1) São selecionados os dez padrões de viagens predominantes em cada amostra (observa-se que há leve variação nos valores das freqüências dos padrões considerando aqueles valores apresentados no Capítulo 6 – Tabela 6.7, pois houve uma pequena redução das amostras na aplicação da AD, devido à exclusão de dados faltantes pelo *software*). Assim, para amostra de industriários o padrão predominante é HWH77121, seguido de HWH77111 e assim sucessivamente;

2) A cada um dos dez padrões predominantes é associada a probabilidades de ocorrência de cada um deles para cada indivíduo da amostra. Nesta etapa utilizou-se a função *predict* do S-PLUS 6.1, exemplificada a seguir:

2.1) Considera-se o indivíduo 1, pertencente à amostra de industriários;

2.2) No processamento da AD, o indivíduo 1 faz parte do grupo de indivíduos referentes ao nó terminal 8;

2.3) As pessoas que compõem a folha 8 possuem a probabilidade de 0.22 de realizar o padrão HWH77121, por exemplo;

2.4) Essa probabilidade de realizar cada um dos dez padrões mais freqüentes da amostra é calculada para todos os indivíduos que compõem os dez nós terminais na amostra de industriários.

3) As variáveis dependentes ( $y_1, y_2, \dots, y_{4084}$ ) são definidas pela probabilidade de ocorrência por indivíduo (calculada na etapa anterior - variando de 0 a 1) de cada um dos dez padrões de viagens;

4) Assim, haverá dez modelos lineares. Cada um dos modelos corresponde a um dos dez padrões de viagens predominantes. A probabilidade de ocorrência por indivíduo para cada padrão de viagem corresponde aos valores "y" de cada um dos modelos. As variáveis independentes (*dummy*) foram descritas na subseção anterior.

Desta maneira, considerando a amostra de industriários, no primeiro modelo de regressão, a variável dependente seria a probabilidade de ocorrência do padrão HWH77121 (padrão mais freqüente da amostra) para cada indivíduo e as variáveis independentes seriam aquelas listadas na Tabela 9.1. No entanto, antes da aplicação da RLM, fez-se uma análise de correlação entre as variáveis envolvidas, com a finalidade de evitar problemas de multicolinearidade.

### 9.2.3 Análise de correlação entre variáveis envolvidas

Representadas as variáveis independentes (*dummy*) e as variáveis dependentes (numérica – 0 a 1), fez-se uma análise de correlação entre todas as variáveis envolvidas. Variáveis independentes altamente correlacionadas foram descartadas do modelo.

Desta forma, se duas variáveis independentes tinham alta correlação (valor superior a 0.5), era descartada da análise aquela (dentre as duas altamente correlacionadas) que possuía menor correlação com a variável dependente. Considerando as variáveis independentes utilizadas para a amostra de industriários, aquelas altamente correlacionadas foram “USO DO SOLO 1”, “USO DO SOLO 2” e “USO DO SOLO 3”.

A Tabela 9.2 traz a matriz de correlação das variáveis envolvidas, considerando a variável dependente do primeiro modelo de regressão (modelo HWH77121). Para o primeiro modelo da análise, descartou-se então as variáveis “USO DO SOLO 1” e “USO DO SOLO 3”. A variável “USO DO SOLO 2” possui maior correlação com a variável dependente (-0.225).

Tabela 9.2 Matriz de correlação entre variáveis envolvidas (indústria)

|           | CNH      | USA VTRA | RENDA    | ESTUDA   | USO_SOLO1 | USO_SOLO2     | USO_SOLO3 | HWH77121 |
|-----------|----------|----------|----------|----------|-----------|---------------|-----------|----------|
| CNH       | <b>1</b> | 0,209    | 0,369    | 0,054    | 0,139     | 0,134         | 0,136     | -0,524   |
| USA VTRA  | 0,209    | <b>1</b> | 0,234    | -0,011   | 0,155     | 0,147         | 0,154     | -0,785   |
| RENDA     | 0,369    | 0,234    | <b>1</b> | 0,046    | 0,245     | 0,245         | 0,241     | -0,546   |
| ESTUDA    | 0,054    | -0,011   | 0,046    | <b>1</b> | -0,009    | -0,013        | -0,004    | -0,287   |
| USO_SOLO1 | 0,139    | 0,155    | 0,245    | -0,009   | <b>1</b>  | 0,748         | 0,803     | -0,214   |
| USO_SOLO2 | 0,134    | 0,147    | 0,245    | -0,013   | 0,748     | <b>1</b>      | 0,856     | -0,225   |
| USO_SOLO3 | 0,136    | 0,154    | 0,241    | -0,004   | 0,803     | 0,856         | <b>1</b>  | -0,224   |
| HWH77121  | -0,524   | -0,785   | -0,546   | -0,287   | -0,214    | <b>-0,225</b> | -0,224    | <b>1</b> |

#### 9.2.4 Dez modelos lineares obtidos e seus respectivos parâmetros

Assim, para cada um dos dez padrões de viagens predominantes na amostra de industriários, foi gerado um modelo linear. A Tabela 9.3 traz cada um dos modelos, valores de R e R<sup>2</sup>, coeficientes associados a cada uma das variáveis, bem como a estatística t.

Fazendo uma análise geral, todos os dez modelos obtidos foram considerados razoáveis, com valores de R<sup>2</sup> variando de 0,691(modelo HWHWH/M/M11111) a 0,893 (modelo HWH77121). Afirma-se então que a variância das variáveis dependentes é razoavelmente explicada pelas variáveis independentes utilizadas em cada um dos modelos.

Analisando cada uma das variáveis independentes individualmente, verifica-se que foram consideradas significativas, com valores relativamente altos da estatística t associados a cada uma delas.

O modelo onde se observou maior contribuição da variável de uso do solo (USO DO SOLO 1) foi o modelo HWH77141. Apresentou-se maior valor do coeficiente relativo a tal variável (- 0,101), bem como maior valor de estatística t (-54,71), indicando que o coeficiente estimado é significativamente diferente de zero. Estes resultados são coerentes, haja vista que escolhas de destinos mais distantes ("141") estão fortemente correlacionadas à suficiência ou não de oportunidades na zona de origem e vizinhanças.

Tabela 9.3 Modelos de regressão obtidos para a amostra de industriários

| Indústria (4.084 indivíduos) |           |        |        |         |        |           |        |       |                |
|------------------------------|-----------|--------|--------|---------|--------|-----------|--------|-------|----------------|
| Padrões                      | modelos   |        |        |         |        |           |        | R     | R <sup>2</sup> |
|                              | variáveis |        |        |         |        |           |        |       |                |
| HWH7T121<br>11,41%           | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO2 |        |       |                |
|                              | B         | 0,227  | -0,038 | -0,099  | -0,065 | -0,061    | -0,004 | 0,945 | 0,893          |
|                              | t         | 315,23 | -47,78 | -124,79 | -48,2  | -52,31    | -5,25  |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |
| HWH7T111<br>11,02%           | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO1 |        |       |                |
|                              | B         | 0,207  | -0,084 | -0,09   | -0,048 | -0,033    | 0,056  | 0,847 | 0,718          |
|                              | t         | 140,44 | -52,62 | -56,5   | -21,33 | -14,06    | 37,03  |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |
| HWH7T141<br>10,95%           | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO1 |        |       |                |
|                              | B         | 0,261  | -0,019 | -0,109  | -0,025 | -0,048    | -0,101 | 0,834 | 0,695          |
|                              | t         | 145,85 | -9,98  | -56,56  | -9,09  | -16,92    | -54,71 |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |
| HWHHW111<br>10,47%           | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO2 |        |       |                |
|                              | B         | 0,077  | -0,133 | -0,175  | -0,77  | -0,36     | 0,026  | 0,951 | 0,724          |
|                              | t         | 39,64  | -61,78 | 81,37   | -25,33 | -11,26    | 12,69  |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |
| HWH7T131<br>7,53%            | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO3 |        |       |                |
|                              | B         | 0,15   | -0,006 | -0,069  | -0,022 | -0,033    | -0,026 | 0,913 | 0,834          |
|                              | t         | 254,43 | -9,81  | -106,77 | -24,16 | -34,29    | -41,76 |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |
| HWHHP111<br>6,28%            | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO3 |        |       |                |
|                              | B         | -0,021 | 0,102  | 0,047   | 0,031  | -0,045    | 0,02   | 0,92  | 0,846          |
|                              | t         | -22,68 | 101,94 | 47,3    | 22     | -30,627   | 20,672 |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |
| HWHHP121<br>4,69%            | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO3 |        |       |                |
|                              | B         | -0,005 | 0,068  | 0,027   | 0,068  | -0,045    |        | 0,932 | 0,868          |
|                              | t         | -8     | 94,27  | 37,17   | 67,41  | -41,88    |        |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      |           |        |       |                |
| HWHHP141<br>3,97%            | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO1 |        |       |                |
|                              | B         | 0      | 0,047  | 0,024   | 0,094  | -0,043    | -0,011 | 0,886 | 0,784          |
|                              | t         | 0,41   | 50,86  | 26,38   | 71     | -31,55    | -12,61 |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |
| HWHHWHNNWN11111<br>3,32%     | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO2 |        |       |                |
|                              | B         | 0,011  | -0,029 | 0,053   | -0,027 |           | 0,027  | 0,832 | 0,691          |
|                              | t         | 17,13  | -41,96 | 76      | -26,79 |           | 39,48  |       |                |
| Sig                          | 0         | 0      | 0      | 0       |        | 0         |        |       |                |
| HWHHP131<br>2,68%            | cte       | CNH    | VTRA   | RENDA   | EST    | USO_SOLO3 |        |       |                |
|                              | B         | -0,008 | 0,043  | 0,019   | 0,02   | -0,025    | 0,008  | 0,916 | 0,839          |
|                              | t         | -19,86 | 94,71  | 41,41   | 30,91  | -37,43    | 18,88  |       |                |
| Sig                          | 0         | 0      | 0      | 0       | 0      | 0         |        |       |                |

| RENDA            | CNH        | USA VTRA            | ESTUDA         | USO DO SOLO 1  | USO DO SOLO 2    | USO DO SOLO 3   |
|------------------|------------|---------------------|----------------|----------------|------------------|-----------------|
| RENDA < 2970 ; 0 | CNH=0; 0   | VTRA=1 (USA); 0     | EST=1 (Não); 0 | 5 km < 3,22; 0 | 15 km < 32,31; 0 | 10km < 15,49; 0 |
| RENDA ≥ 2970 ; 1 | CNH ≥ 1; 1 | VTRA=2 (NÃO USA); 1 | EST>1 (Sim); 1 | 5km ≥ 3,22; 1  | 15km ≥ 32,31; 1  | 10km ≥ 15,49; 1 |

Todos os valores dos coeficientes estimados para cada variável foram coerentes. Assim, considerando o modelo HWH77141, pode-se afirmar:

- **CNH (-0,019):** o fato de haver pelo menos um membro da família com Carteira Nacional de Habilitação (valor um para a variável CNH), contribui negativamente para realização do padrão HWH77141. Foi visto que indivíduos que residem em domicílios onde há pelo menos uma Carteira de Habilitação (ou seja, pelo menos um automóvel), realizam predominantemente viagens utilizando transporte particular motorizado. Assim, os coeficientes estimados para a variável CNH assumem valores maiores e positivos para padrões de viagens onde se observa uso de automóvel, como nos modelos HWHPP111, HWHPP121, HWHPP141 e HWHPP131;
- **USA VTRA (-0,109):** O não uso de Vale Transporte (valor um para a variável Usa VTRA) contribui fortemente para não realização do padrão HWH77141, por isso encontra-se um valor tão alto e negativo para o coeficiente estimado para a variável "USA VTRA". Foi vista a importância desta variável na escolha da seqüência de modos de transportes, especialmente modo de transporte público. O valor estimado do coeficiente para esta variável somente corrobora o que foi dito anteriormente a respeito dos resultados obtidos com a AD. São encontrados valores positivos para os coeficientes estimados para a referida variável no caso dos modelos associados aos padrões de viagem com uso de modo privado (HWHPP111, HWHPP121, HWHPP141 e HWHPP131);
- **RENDA (-0.025):** Foi dito anteriormente que indivíduos com maiores rendas possuem uma maior tendência ao uso de automóvel. Assim, o fato de indivíduos possuírem rendas familiares superiores ou iguais a R\$ 2970 (valor um para a variável RENDA) contribui negativamente para realização do padrão HWH77141. Pode-se verificar valores positivos para coeficientes associados à variável "RENDA" para os modelos HWHPP121; HWHPP111; HWHPP131; HWHPP141 (onde há uso de automóvel na seqüência de viagens);
- **ESTUDA (-0,048):** Os resultados obtidos com a AD sugerem que indivíduos que estudam (além de trabalharem no setor industrial) realizam seqüências de atividades relacionadas ao trabalho e

ao estudo. Desta forma, o fato de os indivíduos participarem em atividades de estudo (valor um para a variável ESTUDA) contribui de forma negativa para a realização do padrão HWH77141, pois neste caso os únicos motivos das viagens são TRABALHO e RESIDÊNCIA;

- **USO DO SOLO 1 (-0,101)** : O alto valor do coeficiente estimado para tal variável explicita a sua importância principalmente na escolha de destinos. Indivíduos que residem em zonas com maior parcela de oportunidades de empregos na indústria ( $5 \text{ km} \geq 3.22\%$ ;) possuem uma forte tendência a escolher destinos próximos. Assim, caso a variável "USO DO SOLO 1" tenha valor um (ou seja,  $5 \text{ km} \geq 3.22\%$ ), esta contribui fortemente de maneira negativa para realização do padrão HWH77141. São encontrados valores positivos de coeficiente estimados para a mesma variável no caso dos modelos associados a padrões de viagens onde há escolha de destinos próximos à residência (modelo HWH77111; modelo HWHWH/WWW/11111).

Esta seção, referente à discussão e análise de resultados obtidos através da aplicação da RLM para amostra de industriários, comprovou os resultados anteriormente gerados a partir do processamento da AD.

Os valores dos coeficientes estimados para as variáveis independentes foram considerados coerentes e ratificaram a influência de cada variável independente selecionada pelo modelo CART na sequência de viagens realizadas pelos indivíduos. Também foi mensurada a significância estatística de cada variável, especialmente daquelas relacionadas ao uso do solo da RMSP. Estas exercem forte influência na escolha das seqüências de destinos.

Nas próximas seções há discussões a respeito dos modelos lineares obtidos nas demais amostras.

### 9.3 TRABALHADORES NO SETOR COMERCIAL

Na presente seção é descrita sucintamente a aplicação da RLM para amostra de comerciários. Espera-se que as variáveis de uso do solo, apesar de serem também estatisticamente significativas, possuam uma menor contribuição nos modelos obtidos.

Este fato é justificado, como mencionando diversas vezes no capítulo anterior, pela maior dispersão geográfica de oferta de empregos no setor do comércio na RMSP. Assim, variáveis de uso do solo relacionadas a empregos no setor do comércio não são tão relevantes na escolha dos padrões de viagens, pois, na maioria das vezes as localizações de empregos dos comerciários são próximas às respectivas residências.

Além disso, observa-se também a contribuição das demais variáveis independentes, selecionadas pelo algoritmo CART, através dos valores dos coeficientes estimados nos dez modelos lineares.

### 9.3.1 Aplicação da RLM e significância das variáveis independentes

A Tabela 9.4 representa a transformação das variáveis independentes selecionadas pelo algoritmo CART em variáveis *dummy* para aplicação da RLM na amostra de trabalhadores no setor do comércio.

Tabela 9.4 Variáveis independentes para aplicação da RLM (comércio)

| Variáveis independentes                                                    |                                                           |                                                                    |                   |
|----------------------------------------------------------------------------|-----------------------------------------------------------|--------------------------------------------------------------------|-------------------|
| Nomenclatura antiga                                                        | Original                                                  | Dummy                                                              | Nova nomenclatura |
| Estuda                                                                     | ESTUDA<br>(categórica - 1 (Não) e 3 ou 4 (Sim))           | ESTUDA = 1 (0)<br>ESTUDA = 3 ou 4 (1)                              | ESTUDA            |
| Nº provável de carteiras de habilitação no domicílio                       | CNH<br>(numérica - variando de 0 a 5)                     | CNH = 0 (0)<br>CNH ≥ 1 (1)                                         | CNH               |
| Parcela acumulada de empregos no comércio até 5 km do centróide da origem  | Comércio até 5 km<br>(contínua - %)                       | Comércio até 5 km < 0.008% (0)<br>Comércio até 5 km ≥ 0.008% (1)   | USO DO SOLO 1     |
| Usa Vale Transporte                                                        | Usa VTRA<br>(categórica - 1 (Sim) e 2 (Não))              | Usa VTRA = 1 (0)<br>Usa VTRA = 2 (1)                               | Usa VTRA          |
| Trabalha                                                                   | TRABALHA<br>(categórica - 1 (Assalariados) e 2 (Autômos)) | TRABALHA = 1 (0)<br>TRABALHA = 2 (1)                               | TRABALHA          |
| Parcela acumulada de empregos no comércio até 15 km do centróide da origem | Comércio até 15 km<br>(contínua - %)                      | Comércio até 15 km < 36.45% (0)<br>Comércio até 15 km ≥ 36.45% (1) | USO DO SOLO 2     |
| Renda Familiar                                                             | RENDA FAMILIAR (RF)<br>(contínua - em reais)              | RF < R\$ 2258 (0)<br>RF ≥ R\$ 2258 (1)                             | RENDA             |
| Nº de automóveis no domicílio                                              | AUT<br>(numérica - variando de 0 a 8)                     | AUT = 1 (0)<br>AUT ≥ 2 (1)                                         | AUT               |

A exemplo da aplicação da RLM na amostra de industriários, como descrito anteriormente, houve análise de correlação entre variáveis independentes, no entanto, conforme mostrado na Tabela 9.5, apenas as variáveis que apresentaram alta correlação (*pearson correlation*) foram "AUT" e "CNH".

Desta forma, os modelos lineares tinham sete variáveis independentes, pois uma delas é descartada de acordo com a variável dependente.

Tabela 9.5 Matriz de correlação entre variáveis envolvidas (comércio)

|           | ESTUDA        | USA VTRA      | CNH           | AUT           | TRABALHA      | RENDA         | USO_SOLO1     | USO_SOLO2     |
|-----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| ESTUDA    | <b>1,0000</b> | -0,2677       | 0,0509        | 0,0708        | -0,1958       | -0,0594       | -0,0958       | -0,3073       |
| USA VTRA  | -0,2677       | <b>1,0000</b> | 0,0904        | 0,1041        | 0,3932        | 0,1831        | 0,1542        | 0,2288        |
| CNH       | 0,0509        | 0,0904        | <b>1,0000</b> | 0,6230        | 0,1254        | 0,3511        | 0,0574        | -0,0832       |
| AUT       | 0,0708        | 0,1041        | 0,6230        | <b>1,0000</b> | 0,1359        | 0,3881        | 0,0672        | -0,0936       |
| TRABALHA  | -0,1958       | 0,3932        | 0,1254        | 0,1359        | <b>1,0000</b> | 0,0931        | 0,0331        | 0,0876        |
| RENDA     | -0,0594       | 0,1831        | 0,3511        | 0,3881        | 0,0931        | <b>1,0000</b> | 0,2682        | 0,1543        |
| USO_SOLO1 | -0,0958       | 0,1542        | 0,0574        | 0,0672        | 0,0331        | 0,2682        | <b>1,0000</b> | 0,3391        |
| USO_SOLO2 | -0,3073       | 0,2288        | -0,0832       | -0,0936       | 0,0876        | 0,1543        | 0,3391        | <b>1,0000</b> |

Também foram obtidos dez modelos lineares razoáveis (a partir dos dez padrões de viagens mais frequentes na amostra), com valores do índice de determinação ( $R^2$ ) variando de 0,57 (modelo HWHPP141) a 0,794 (modelo HWH77131).

Os coeficientes estimados foram considerados significativamente diferentes de zero e como esperado, as variáveis de uso do solo não tiveram uma contribuição tão forte quanto no caso dos industriários. O maior valor de coeficiente estimado foi 0,044 para a variável “USO DO SOLO 1” no modelo HWH77141. A Tabela 9.6 apresenta os dez modelos lineares obtidos, seus parâmetros e estatísticas associadas.

Tabela 9.6 Modelos de regressão obtidos para a amostra de comerciários

| Comércio (6.019 indivíduos) |           |        |         |        |        |           |          |           |        |       |                |
|-----------------------------|-----------|--------|---------|--------|--------|-----------|----------|-----------|--------|-------|----------------|
| Padrões                     | modelos   |        |         |        |        |           |          |           |        | R     | R <sup>2</sup> |
|                             | variáveis |        |         |        |        |           |          |           |        |       |                |
| HWHHW111<br>11,47%          | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | 0,114  | 0,149   | -0,087 | -0,043 | -0,026    | -0,026   | -0,012    | 0,005  | 0,845 | 0,713          |
|                             | t         | 44,17  | 89,38   | -59,56 | -22,91 | -16,52    | -10,74   | -8,16     | 3,77   |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        | 0         |        |       |                |
| HWHHP111<br>8,28%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | -0,033 | 0,065   | 0,096  | -0,063 | 0,034     | 0,008    | 0,005     | 0,02   | 0,864 | 0,746          |
|                             | t         | -14,37 | 44,33   | 74,68  | -33,08 | 24,4      | 3,74     | 3,55      | 16,15  |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        | 0         |        |       |                |
| HWH77121<br>7,27%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | 0,173  | -0,107  | -0,039 | -0,084 | -0,016    | 0,039    | -0,017    | 0,007  | 0,881 | 0,776          |
|                             | t         | 101,54 | -98,18  | -40,55 | -68,18 | -14,89    | 24       | -18,01    | 7,55   |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        | 0         |        |       |                |
| HWH77111<br>7,14%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | 0,169  | -0,077  | -0,059 | -0,076 | -0,01     | 0,029    | -0,013    |        | 0,889 | 0,791          |
|                             | t         | 144,07 | -81,16  | -71,01 | -70,84 | -11,72    | 21,8     | -15,09    |        |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        |           |        |       |                |
| HWH77141<br>6,80%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | 0,199  | -0,132  | -0,032 | -0,107 | -0,019    | 0,044    | -0,009    | -0,032 | 0,885 | 0,783          |
|                             | t         | 100,24 | -103,83 | -28,74 | -74,07 | -15,29    | 23,33    | -8,13     | -29,8  |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        | 0         |        |       |                |
| HWHHP121<br>5,44%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | -0,006 | 0,039   | 0,052  | -0,049 | 0,03      |          | -0,002    | 0,014  | 0,859 | 0,738          |
|                             | t         | -6,98  | 41,87   | 64,5   | -47,53 | 34,11     |          | -3,07     | 18,65  |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        | 0         |        |       |                |
| HWH77131<br>4,77%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | 0,122  | -0,074  | -0,029 | -0,06  | -0,011    | 0,028    | -0,008    | -0,006 | 0,891 | 0,794          |
|                             | t         | 107,85 | -101,71 | -45,97 | -72,52 | -15,96    | 25,55    | -13,19    | -10,48 |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        | 0         |        |       |                |
| HWHHP141<br>3,12%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | 0,007  | 0,025   | 0,03   | -0,029 | 0,025     | 0,004    | 0,002     | -0,03  | 0,755 | 0,57           |
|                             | t         | 5,06   | 28,92   | 39,48  | -29,88 | 30,14     | 3,05     | 2,21      | -40,57 |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0,03     | 0         |        |       |                |
| HWHHNNNN11111<br>3,02%      | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | 0,028  | 0,044   | -0,024 | -0,008 | -0,006    | -0,005   | -0,012    |        | 0,813 | 0,66           |
|                             | t         | 34,82  | 85,76   | -52,22 | -12,98 | -12,28    | -6,28    | -27,71    |        |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0         | 0        |           |        |       |                |
| HWHHP131<br>2,92%           | cte       | VTRA   | AUT     | EST    | RENDA  | USO-SOLO1 | TRABALHA | USO-SOLO2 |        |       |                |
|                             | B         | -0,005 | 0,02    | 0,33   | -0,028 | 0,014     | 0,002    | 0,001     | 0,002  | 0,871 | 0,758          |
|                             | t         | -6,28  | 39,85   | 76,68  | -50,91 | 30,18     | 3,37     | 2,73      | 3,77   |       |                |
|                             | Sig       | 0      | 0       | 0      | 0      | 0,001     | 0,006    | 0         |        |       |                |

| ESTUDA        | TRABALHA                   | RENDA         | USA VTRA           | USO DO SOLO 1   | USO DO SOLO 2    | AUT      |
|---------------|----------------------------|---------------|--------------------|-----------------|------------------|----------|
| EST=1 (Não);0 | TRABALHA=1 (Assalariado);0 | RENDA<=2258;0 | VTRA=1 (USA);0     | 5 km < 0,008; 0 | 15 km < 36,45; 0 | AUT<=1;0 |
| EST>1 (Sim);1 | TRABALHA = 2 (Autônomo);1  | RENDA>2258;1  | VTRA=2 (NÃO USA);1 | 5km ≥ 0,008; 1  | 15km ≥ 36,45; 1  | AUT>1;1  |

Considerando o mesmo modelo (HWH77141) e os valores dos coeficientes estimados para cada variável, observam-se valores negativos para as variáveis “USA VTRA”, “ESTUDA” e “RENDA”. Conforme comentado na seção anterior, valores negativos para coeficientes estimados para as citadas variáveis são coerentes. Caso essas variáveis assumam valores iguais a um, estas contribuem de

forma negativa para realização do padrão HWH77141. Levando-se em conta as demais variáveis (diferentes daquelas consideradas na amostra de industriários), faz-se alguns comentários a respeito das suas respectivas contribuições:

- **AUT (-0,032):** o fato de haver pelo menos dois automóveis no domicílio (valor um para a variável AUT), contribui negativamente para realização do padrão HWH77141. Indivíduos que residem em domicílios com automóveis realizam predominantemente viagens utilizando transporte particular motorizado. Assim os coeficientes estimados para a variável AUT assumem valores positivos para padrões de viagens onde se observa uso de automóvel, como nos modelos HWHPP131, HWHPP121; HWHPP111; HWHPP141;
- **TRABALHA (-0,009):** Caso a variável assumira valor um (ou seja, "trabalhadores autônomos"), esta contribui negativamente para realização do padrão HWH77141. Trabalhadores autônomos podem realizar outras atividades não relacionadas ao trabalho com maior frequência, devido ao seu tipo de ocupação e não necessidade de realizar viagens diárias ao trabalho. No entanto, observa-se um valor pequeno para o coeficiente estimado, evidenciando sua fraca contribuição no referido modelo. Possivelmente, se fossem obtidos modelos associados a diferentes seqüências de atividades (como HAH ou HAHWH, por exemplo), os valores obtidos fossem maiores;
- **USO DO SOLO 1 (0,044):** O fato de residir em zonas com parcela acumulada de empregos no comércio até 5 km do seu centróide maior que 0.008% (valor um para a variável "USO DO SOLO 1") contribui de maneira positiva para a realização do padrão HWH77141. O valor considerado da parcela acumulada de empregos é muito baixo (0.008%), assim, é factível que indivíduos necessitem escolher destinos distantes para satisfazer suas necessidades de trabalho;
- **USO DO SOLO 2 (-0,032):** A contribuição negativa da referida variável pode ser justificada. Quando esta assume valor igual a um (ou seja, parcela acumulada de empregos no comércio até 15 km do centróide da origem superior ou igual a 36.45%), é mais provável que indivíduos realizem viagens ao trabalho com uma distância até 15 km (já que se admite que haja alta parcela acumulada de

empregos até esta distância). Assim, seria mais improvável a realização de viagens acima de 15 km, que é o caso do padrão HWH 77141.

## 9.4 TRABALHADORES NO SETOR DE SERVIÇOS

São descritas a seguir as etapas para aplicação da RLM para o caso de trabalhadores no setor de serviços. Espera-se que a contribuição das variáveis de uso do solo seja maior do que no caso dos comerciários, devido a fatores expostos no capítulo anterior.

### 9.4.1 Aplicação da RLM e significância das variáveis independentes

A Tabela 9.7 representa a transformação das variáveis independentes selecionadas pelo algoritmo CART (para o caso de trabalhadores no setor de serviços) em variáveis *dummy* para aplicação da RLM. Muitas variáveis assemelham-se com aquelas obtidas na amostra de industriários.

Tabela 9.7 Variáveis independentes para aplicação da RLM (serviços)

| Variáveis independentes                                                    |                                                  |                                                                    |                   |
|----------------------------------------------------------------------------|--------------------------------------------------|--------------------------------------------------------------------|-------------------|
| Nomenclatura antiga                                                        | Original                                         | Dummy                                                              | Nova nomenclatura |
| Nº provável de carteiras de habilitação no domicílio                       | CNH<br>(numérica - variando de 0 a 5)            | CNH = 0 (0)<br>CNH ≥ 1 (1)                                         | CNH               |
| Estuda                                                                     | ESTUDA<br>(categórica - 1 (Não) e 3 ou 4 (Sim))  | ESTUDA = 1 (0)<br>ESTUDA = 3 ou 4 (1)                              | ESTUDA            |
| Usa Vale Transporte                                                        | Usa VTRA<br>(categórica - 1 (Sim) e 2 (Não))     | Usa VTRA = 1 (0)<br>Usa VTRA = 2 (1)                               | Usa VTRA          |
| Parcela acumulada de empregos em serviços até 5 km do centróide da origem  | Serviços até 5 km<br>(contínua - %)              | Serviços até 5 km < 3.85% (0)<br>Serviços até 5 km ≥ 3.85% (1)     | USO DO SOLO 1     |
| Parcela acumulada de empregos em serviços até 15 km do centróide da origem | Serviços até 15 km<br>(contínua - %)             | Serviços até 15 km < 37.15% (0)<br>Serviços até 15 km ≥ 37.15% (1) | USO DO SOLO 2     |
| Sexo                                                                       | Sexo<br>(categórica - 1 (Homens) e (2) Mulheres) | SEXO = 1 (0)<br>SEXO = 2 (1)                                       | SEXO              |
| Nº de automóveis no domicílio                                              | AUT<br>(numérica - variando de 0 a 8)            | AUT = 1 (0)<br>AUT > 2 (1)                                         | AUT               |

Fazendo uma análise de correlação das sete variáveis independentes notou-se uma correlação alta apenas entre as variáveis de uso do solo (0.812). Desta forma, uma delas foi descartada em cada um dos modelos de acordo com sua correlação com a variável dependente, a exemplo das amostras anteriores. Assim, cada um dos modelos obtidos era composto por seis das sete variáveis descritas na

Tabela 9.7. A Tabela 9.8 apresenta os dez modelos lineares obtidos, seus parâmetros e estatísticas associadas para o caso da presente amostra.

Tabela 9.8 Modelos de regressão obtidos para a amostra de trabalhadores do setor de serviços

| Serviços (8.286 indivíduos) |           |        |         |        |         |        |        |           |       |                |
|-----------------------------|-----------|--------|---------|--------|---------|--------|--------|-----------|-------|----------------|
| Padrões                     | modelos   |        |         |        |         |        |        |           | R     | R <sup>2</sup> |
|                             | variáveis |        |         |        |         |        |        |           |       |                |
| HWH77121<br>8,72%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO1 | 0,93  | 0,865          |
|                             | t         | 0,209  | -0,047  | -0,064 | -0,096  | 0,003  | -0,03  | 0,017     |       |                |
|                             | Sig       | 315,2  | -78,75  | -77,26 | -158,89 | 6,05   | -45,21 | 32,58     |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWH77111<br>8,66%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO2 | 0,926 | 0,857          |
|                             | t         | 0,2    | -0,058  | -0,059 | -0,066  | 0,004  | -0,026 |           |       |                |
|                             | Sig       | 341,68 | -105,44 | -77,65 | -119,49 | 9,33   | -43,13 | 0         |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWHHW111<br>8,43%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO1 | 0,814 | 0,663          |
|                             | t         | 0,076  | -0,088  | -0,048 | 0,109   | 0,018  | -0,047 | 0,016     |       |                |
|                             | Sig       | 56,51  | -72,25  | -28,93 | 89,84   | 17,43  | -34,58 | 15,22     |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWHPP111<br>8,26%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO1 | 0,919 | 0,845          |
|                             | t         | -0,016 | 0,088   | -0,069 | 0,072   | -0,003 | 0,052  | 0,005     |       |                |
|                             | Sig       | -18,06 | 108,23  | -60,91 | 87      | -3,86  | 56,56  | 6,5       |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWH77141<br>7,71%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO1 | 0,895 | 0,801          |
|                             | t         | 0,247  | -0,033  | -0,068 | -0,126  | -0,001 | -0,009 | -0,059    |       |                |
|                             | Sig       | 233,2  | -34,53  | -51,87 | -131    | -0,66  | -8,813 | -69,57    |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWH77131<br>6,00%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO1 | 0,942 | 0,888          |
|                             | t         | 0,157  | -0,029  | -0,045 | -0,076  |        | -0,013 | -0,005    |       |                |
|                             | Sig       | 360,38 | -72,52  | -82,85 | -193,29 |        | -29,5  | -14,53    |       |                |
|                             |           | 0      | 0       | 0      | 0       |        | 0      | 0         |       |                |
| HWHPP121<br>4,91%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO2 | 0,884 | 0,781          |
|                             | t         | -0,02  | 0,047   | 0,038  | 0,041   | -0,006 | 0,033  | 0,032     |       |                |
|                             | Sig       | -27,52 | 73,26   | -43,43 | 62,78   | -11,65 | 45,68  | 56,51     |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWHPP141<br>2,98%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO2 | 0,729 | 0,531          |
|                             | t         | 0,024  | 0,036   | -0,021 | 0,029   | -0,02  | 0,01   | -0,041    |       |                |
|                             | Sig       | 26,58  | 44,87   | -19,25 | 36,82   | -28,77 | 11,02  | -58,74    |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWHPP131<br>2,56%           | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO1 | 0,887 | 0,786          |
|                             | t         | 0,003  | 0,025   | -0,01  | 0,011   | -0,005 | 0,023  | 0,003     |       |                |
|                             | Sig       | 11     | 88,63   | -26,94 | 38,42   | -22,58 | 71,75  | 11,23     |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |
| HWHHWHNNNN11111<br>1,52%    | B         | cte    | CNH     | EST    | VTRA    | SEXO   | AUT    | USO-SOLO1 | 0,89  | 0,791          |
|                             | t         | 0,011  | -0,017  | -0,005 | 0,024   | 0,001  | -0,003 | 0,003     |       |                |
|                             | Sig       | 61,26  | -104,88 | -21,25 | 145,82  | 6,25   | -15,23 | 18,46     |       |                |
|                             |           | 0      | 0       | 0      | 0       | 0      | 0      | 0         |       |                |

|            |                    |               |            |             |                      |                      |
|------------|--------------------|---------------|------------|-------------|----------------------|----------------------|
| <b>CNH</b> | <b>USA VTRA</b>    | <b>ESTUDA</b> | <b>AUT</b> | <b>SEXO</b> | <b>USO DO SOLO 1</b> | <b>USO DO SOLO 2</b> |
| CNH=0; 0   | VTRA=1 (USA);0     | EST=1 (Não);0 | AUT≤1;0    | Sexo = 1; 0 | 5 km < 3,85; 0       | 15km < 37,15; 0      |
| CNH ≥ 1;1  | VTRA=2 (NÃO USA);1 | EST>1 (Sim);1 | AUT>1;1    | Sexo = 2; 1 | 5km ≥ 3,85; 1        | 15km ≥ 37,15; 1      |

Pode-se afirmar que, fazendo uma análise apenas nos valores dos coeficientes de determinação e parâmetros obtidos, os dez modelos lineares foram considerados razoáveis. Os valores dos índices de determinação ( $R^2$ ) variaram de 0,531 (modelo HWHPP141) a 0,888 (modelo HWH77131).

Os coeficientes estimados foram considerados significativamente diferentes de zero e como esperado, as variáveis de uso do solo tiveram uma boa contribuição no caso de alguns modelos. O maior valor de coeficiente estimado foi -0,059 para a variável "USO DO SOLO 1" no modelo HWH77141. O valor obtido indica que o fato de residir em zonas com parcela acumulada de empregos na indústria até 5 km superior ou igual a 3.85% (valor um para a variável USO DO SOLO 1) contribui negativamente para a realização do padrão HWH77141.

Como já mencionado anteriormente, maiores oportunidades de empregos nas proximidades da residência contribuem positivamente para a realização de viagens curtas ("111") e negativamente para realização de viagens longas ("141"). Em todos os casos analisados até o presente momento, as maiores contribuições das variáveis de uso do solo foram nos modelos HWH77141 (nas três amostras discutidas). Pode-se justificar isto no fato de que escolha de destinos distantes ("141") para viagens relacionadas ao trabalho, está vinculada de certa forma a suficiência ou não de oferta de empregos próximos à residência.

Como esperado, a contribuição das variáveis de uso do solo foi maior no caso de trabalhadores no setor de serviços, quando comparados aos comerciários. Isto valida a discussão dos resultados obtidos com aplicação da AD, onde se afirmou que a maior concentração de empregos no setor de serviços influencia escolhas de destinos de trabalhadores neste setor econômico, especialmente quando comparados aos trabalhadores no setor de comércio.

Quanto à variável "SEXO", que foi selecionada apenas na presente amostra para segregação dos dados, pode-se citar o exemplo do modelo HWHPP141. Observa-se uma contribuição negativa desta variável (-0.02), caso esta assumo valor igual a uma unidade (SEXO = 2; Mulheres) para a realização do padrão HWHPP141. Este fato converge completamente com os resultados gerados

através do processamento da AD. Concluiu-se que mulheres utilizam o automóvel com menor frequência e também realizam viagens mais curtas, quando comparadas aos homens. Assim, observa-se valor positivo para o coeficiente estimado para variável "SEXO" para padrões de viagens associados a modos de transporte não motorizados e destinos próximos às residências (modelo HWH/M111 – 0.018).

## 9.5 ESTUDANTES ATÉ O 1º GRAU

A seguir, é descrita a aplicação da RLM e os modelos lineares obtidos para o caso particular de estudantes com escolaridade até o 1º grau completo.

### 9.5.1 Aplicação da RLM e significância das variáveis independentes

A Tabela 9.9 representa a transformação das variáveis independentes selecionadas pelo algoritmo CART, no presente caso, em variáveis *dummy* para aplicação da RLM. Observa-se que muitas das variáveis podem estar altamente correlacionadas entre si (variáveis relacionadas à idade; Renda Familiar e Nº de automóveis no domicílio). Isto foi comprovado após verificação da matriz de correlação entre as variáveis independentes.

O problema de multicolinearidade foi resolvido da mesma forma que nas amostras anteriores, foram descartadas aquelas variáveis que eram altamente correlacionadas entre si e que tinham menor correlação com a variável dependente.

Assim, das seis variáveis descritas na Tabela 9.9, foram utilizadas apenas quatro delas em cada um dos modelos obtidos.

Tabela 9.9 Variáveis independentes para aplicação da RLM (estudantes até o 1º grau)

| Variáveis independentes                                                       |                                       |                                                            |                   |
|-------------------------------------------------------------------------------|---------------------------------------|------------------------------------------------------------|-------------------|
| Nomenclatura antiga                                                           | Original                              | Dummy                                                      | Nova nomenclatura |
| Renda Familiar                                                                | RF<br>(contínua - em reais)           | RF < 1690 (0)<br>RF ≥ 1690 (1)                             | RENDA             |
| Nº de automóveis no domicílio                                                 | AUT<br>(numérica - variando de 0 a 8) | AUT = 0 (0)<br>AUT ≥ 1 (1)                                 | AUT               |
| Idade                                                                         | IDADE<br>(numérica - em anos)         | IDADE ≤ 10 (0)<br>IDADE > 10 (1)                           | IDADE 1           |
| Idade                                                                         | IDADE<br>(numérica - em anos)         | IDADE ≤ 11 (0)<br>IDADE > 11 (1)                           | IDADE 2           |
| Idade                                                                         | IDADE<br>(numérica - em anos)         | 10 < IDADE ≤ 15 (0)<br>IDADE > 15 (1)                      | IDADE 3           |
| Parcela acumulada de matrículas até o 1º grau até 5 km do centróide da origem | 1 grau até 5 km<br>(contínua - %)     | 1 grau até 5 km < 1.40% (0)<br>1 grau até 5 km ≥ 1.40% (1) | USO DO SOLO 1     |

Representadas as variáveis independentes (binárias) e variáveis dependentes, foram obtidos dez modelos lineares, a exemplo das amostras anteriores. Considerando apenas o coeficiente de determinação ( $R^2$ ) alguns modelos não explicavam muito bem a variabilidade da variável dependente, como o modelo HSH77111, por exemplo, com valor de  $R^2$  igual a 0,39. Outros foram considerados razoáveis, como o modelo HSH/M11, com valor de  $R^2$  igual a 0,886.

Os coeficientes estimados para as variáveis independentes foram considerados significativamente diferentes de zero através dos valores obtidos com a estatística t. Como esperado, a variável de uso do solo teve baixa contribuição (isto é, baixos valores de coeficientes estimados) para todos os modelos. O maior valor de coeficiente encontrado para variável de uso do solo foi 0,019 para o modelo HSH/M11. Isto pode explicar a sua contribuição positiva caso o indivíduo resida em uma zona de tráfego com parcela acumulada de escolas superior ou igual a 1.40% (valor um para variável USO DO SOLO 1) para a realização do referido padrão de viagem.

Os resultados obtidos justificam o fato de que indivíduos que residem em zonas com maiores oportunidades (escolas no 1º grau) nas suas redondezas realizam viagens curtas à escola, e, conseqüentemente vão a pé para tais destinos próximos.

A Tabela 9.10 apresenta os dez modelos lineares obtidos e o sumário dos seus principais parâmetros para o caso das amostras de estudantes matriculados até o 1º grau.

Tabela 9.10 Modelos de regressão obtidos para a amostra de estudantes matriculados até o 1º grau

| Estudantes até 1º grau (13.733 indivíduos) |           |        |         |        |         |           |       |       |
|--------------------------------------------|-----------|--------|---------|--------|---------|-----------|-------|-------|
| Padrões                                    | modelos   |        |         |        |         |           | R     | R²    |
|                                            | variáveis |        |         |        |         |           |       |       |
| HSHWW111<br>59,12%                         | B         | cte    | RF      | AUT    | IDADE 1 | USO-SOLO1 | 0,931 | 0,866 |
|                                            | t         | 0,711  | -0,302  | -0,121 | 0,024   | 0,019     |       |       |
|                                            | Sig       | 470,09 | 206,49  | -93,48 | 20,91   | 12,77     |       |       |
|                                            |           | 0      | 0       | 0      | 0       | 0         |       |       |
| HSHPP111<br>13,21%                         | B         | cte    | RF      | AUT    | IDADE 1 | USO-SOLO1 | 0,906 | 0,82  |
|                                            | t         | 0,084  | 0,213   | 0,069  | -0,089  |           |       |       |
|                                            | Sig       | 64,03  | 168,78  | 61,28  | -89,75  |           |       |       |
|                                            |           | 0      | 0       | 0      | 0       |           |       |       |
| HSH77111<br>5,73%                          | B         | cte    | RF      | AUT    | IDADE 1 | USO-SOLO1 | 0,623 | 0,39  |
|                                            | t         | 0,048  | 0,002   | 0,006  | 0,02    | -0,004    |       |       |
|                                            | Sig       | 159,16 | 6,28    | 23,18  | 88,14   | -13,7     |       |       |
|                                            |           | 0      | 0       | 0      | 0       | 0         |       |       |
| HSHWW121<br>1,62%                          | B         | cte    | RF      | AUT    | IDADE 1 | USO-SOLO1 | 0,662 | 0,44  |
|                                            | t         | 0,037  | -0,006  | -0,1   | 0,003   | -0,018    |       |       |
|                                            | Sig       | 139,84 | -22,51  | -46,12 | 13,66   | -69,37    |       |       |
|                                            |           | 0      | 0       | 0      | 0       | 0         |       |       |
| HSHPP121<br>1,59%                          | B         | cte    | RF      | AUT    | IDADE 1 | USO-SOLO1 | 0,844 | 0,712 |
|                                            | t         | 0,01   | 0,026   | 0,009  | -0,006  | -0,004    |       |       |
|                                            | Sig       | 46,72  | 129,52  | 51,87  | -35,34  | -17,42    |       |       |
|                                            |           | 0      | 0       | 0      | 0       | 0         |       |       |
| HSHPW111<br>1,42%                          | B         | cte    | RF      | AUT    | IDADE 1 | USO-SOLO1 | 0,956 | 0,913 |
|                                            | t         | 0,007  | 0,011   | 0,013  | -0,005  | 0,001     |       |       |
|                                            | Sig       | 98,45  | 158,69  | 219,28 | -85,04  | 12,44     |       |       |
|                                            |           | 0      | 0       | 0      | 0       | 0         |       |       |
| HSH77121<br>1,33%                          | B         | cte    | RF      | AUT    | IDADE 2 | USO-SOLO1 | 0,704 | 0,495 |
|                                            | t         | 0,012  | -0,003  |        | 0,014   | -0,002    |       |       |
|                                            | Sig       | 77,1   | -21,75  |        | 111,66  | -14,09    |       |       |
|                                            |           | 0      | 0       |        | 0       | 0         |       |       |
| ISHSHWWWW11111<br>1,02%                    | B         | cte    | RF      | AUT    | IDADE 1 | USO-SOLO1 | 0,85  | 0,724 |
|                                            | t         | 0,007  | -0,001  | -0,004 | 0,011   | 0,001     |       |       |
|                                            | Sig       | 77,27  | -17,18  | -56,91 | 174,29  | 16,53     |       |       |
|                                            |           | 0      | 0       | 0      | 0       | 0         |       |       |
| HAHWW111<br>0,92%                          | B         | cte    | RF      | AUT    | IDADE 2 | USO-SOLO1 | 0,904 | 0,817 |
|                                            | t         | 0,01   | -0,008  | -0,001 | 0,005   | 0,001     |       |       |
|                                            | Sig       | 210,77 | -175,73 | -24,35 | 126,92  | 18,13     |       |       |
|                                            |           | 0      | 0       | 0      | 0       | 0         |       |       |
| HSHPT111<br>0,81%                          | B         | cte    | RF      | AUT    | IDADE 3 | USO-SOLO1 | 0,888 | 0,788 |
|                                            | t         | 0,004  | 0,013   | 0,006  | -0,003  |           |       |       |
|                                            | Sig       | 42,27  | 147,37  | 74,37  | -49,83  |           |       |       |
|                                            |           | 0      | 0       | 0      | 0       |           |       |       |

|              |            |                |                |                     |                      |
|--------------|------------|----------------|----------------|---------------------|----------------------|
| <b>RENDA</b> | <b>AUT</b> | <b>IDADE 1</b> | <b>IDADE 2</b> | <b>IDADE 3</b>      | <b>USO DO SOLO 1</b> |
| RF < 1690; 0 | AUT = 0; 0 | Idade ≤ 10 ; 0 | Idade ≤ 11 ; 0 | 10 < Idade ≤ 15 ; 0 | 5 km < 1,4; 0        |
| RF ≥ 1690; 1 | AUT ≥ 1; 1 | Idade > 10 ; 1 | Idade > 11; 1  | Idade > 15; 1       | 5km ≥ 1,4; 1         |

Em relação à variável “IDADE 1”, observa-se maior contribuição no modelo HSHPP111 (-0.089). Analisando o valor do coeficiente estimado, pode-se afirmar que há contribuição negativa desta variável, caso ela assuma valor igual a um (ou seja, Idade > 10 anos) para a realização do citado padrão de viagem. Espera-se que crianças mais velhas, utilizem com maior frequência transporte público para se deslocarem até às escolas, podendo eventualmente escolher destinos mais distantes. O padrão HSHPP111 caracteriza crianças que foram à escola como passageiras de automóvel (padrão mais freqüente no caso de crianças mais novas, que vão à escola na companhia de pais ou responsáveis).

## 9.6 ESTUDANTES NO 2º GRAU

### 9.6.1 Aplicação da RLM e significância das variáveis independentes

Considerando o caso particular de estudantes matriculados no 2º grau, a Tabela 9.11 representa a transformação das variáveis independentes selecionadas pelo algoritmo CART em variáveis *dummy* para aplicação da RLM.

Observou-se, através da matriz de correlação gerada antes da aplicação da RLM, que muitas das variáveis são altamente correlacionadas entre si (como as variáveis relativas à Renda Familiar, nº de automóveis no domicílio e as variáveis de uso do solo entre si).

Como nas amostras anteriores, utilizou-se o mesmo critério para evitar multicolinearidade, ou seja, descartar aquelas variáveis que tinham menor correlação com a variável dependente e que fossem altamente correlacionadas entre si. Exatamente por isso, os modelos obtidos para o caso de estudantes matriculados no 2º grau eram compostos por apenas uma ou duas variáveis independentes na maioria das vezes.

Observando a Tabela 9.11, e baseando-se nas matrizes de correlação geradas antes dos modelos lineares, pode-se dizer que as variáveis “RENDA 1”, “RENDA 2” e “AUT” são altamente correlacionadas, assim como as variáveis “USO DO SOLO 1”, “USO DO SOLO 2”, “USO DO SOLO 3” e “USO DO SOLO 4”. Assim, de acordo com a variável dependente, foi selecionada apenas uma

variável relacionada à renda ou posse de automóveis, bem como apenas uma variável relacionada ao uso do solo.

Tabela 9.11 Variáveis independentes para aplicação da RLM (estudantes no 2º grau)

| Variáveis independentes                                                     |                                       |                                                              |                   |
|-----------------------------------------------------------------------------|---------------------------------------|--------------------------------------------------------------|-------------------|
| Nomenclatura antiga                                                         | Original                              | Dummy                                                        | Nova nomenclatura |
| Renda Familiar                                                              | RF<br>(contínua - em reais)           | RF < 2320 (0)<br>RF ≥ 2320 (1)                               | RENDA 1           |
| Nº de automóveis no domicílio                                               | AUT<br>(numérica - variando de 0 a 8) | AUT = 1 (0)<br>AUT ≥ 2 (1)                                   | AUT               |
| Parcela acumulada de matrículas no 2º grau até 5 km do centróide da origem  | 2grau até 5 km<br>(contínua - %)      | 2grau até 5 km < 2.72% (0)<br>2grau até 5 km ≥ 2.72% (1)     | USO DO SOLO 1     |
| Parcela acumulada de matrículas no 2º grau até 5 km do centróide da origem  | 2grau até 5 km<br>(contínua - %)      | 2grau até 5 km < 0.53% (0)<br>2grau até 5 km ≥ 0.53% (1)     | USO DO SOLO 2     |
| Renda Familiar                                                              | RF<br>(contínua - em reais)           | RF < 1462 (0)<br>RF ≥ 1462 (1)                               | RENDA 2           |
| Parcela acumulada de matrículas no 2º grau até 20 km do centróide da origem | 2grau até 20 km<br>(contínua - %)     | 2grau até 20 km < 65.96% (0)<br>2grau até 20 km ≥ 65.96% (1) | USO DO SOLO 3     |
| Parcela acumulada de matrículas no 2º grau até 10 km do centróide da origem | 2grau até 10 km<br>(contínua - %)     | 2grau até 10 km < 44.13% (0)<br>2grau até 10 km ≥ 44.13% (1) | USO DO SOLO 4     |

Foram obtidos dez modelos lineares, onde a maioria deles continha apenas duas variáveis independentes. Os valores dos coeficientes de determinação variavam entre 0,298 (modelo HSHPT121) e 0,94 (modelo HWHPM11).

As variáveis independentes foram consideradas significativas e na maioria dos modelos houve maior contribuição (maiores valores para os coeficientes estimados) para aquelas variáveis relacionadas à renda (RENDA 1 ou RENDA 2 ou AUT), especialmente quando há uso de transporte particular motorizado. Como citado diversas vezes no capítulo anterior, “posse de automóveis e renda” estão fortemente relacionados ao uso de automóveis para realização de viagens.

Apenas em três modelos houve maior contribuição daquelas variáveis relacionadas ao uso do solo para explicar a variabilidade da variável dependente (modelo HWH77121; modelo HWH77131; modelo HWH77141). Embora os coeficientes estimados para as variáveis de uso do solo não tenham um valor muito alto (o maior valor obtido foi -0.044 para o modelo HSH77121), tais valores foram mais altos do que no caso da amostra composta por estudantes matriculados até o 1º grau. Justifica-se este

fato pela maior concentração de escolas de 2º grau na RMSF e a maior relevância de características de uso do solo no presente caso. A Tabela 9.12 apresenta os dez modelos lineares obtidos e os parâmetros associados a cada um destes.

Tabela 9.12 Modelos de regressão obtidos para a amostra de estudantes matriculados no 2º grau

| Estudantes até 2º grau (2.125 indivíduos) |         |           |           |           |       |       |                |
|-------------------------------------------|---------|-----------|-----------|-----------|-------|-------|----------------|
| Padrões                                   | modelos |           |           |           |       | R     | R <sup>2</sup> |
|                                           |         | variáveis |           |           |       |       |                |
| HSHMW111<br>33,02%                        | B       | cte       | RENDA 1   | USO-SOL01 | 0,838 | 0,703 |                |
|                                           | t       | 0,442     | -0,226    | -0,04     |       |       |                |
|                                           | Sig     | 165,3     | -62,41    | -11,27    |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSH7T111<br>15,89%                        | B       | cte       | AUTO      | USO-SOL01 | 0,715 | 0,512 |                |
|                                           | t       | 0,158     | -0,066    | 0,04      |       |       |                |
|                                           | Sig     | 146,11    | -43,33    | 28,49     |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSHPP111<br>9,23%                         | B       | cte       | RENDA 1   | USO-SOL01 | 0,846 | 0,716 |                |
|                                           | t       | 0,025     | 0,144     | 0,022     |       |       |                |
|                                           | Sig     | 15,38     | 65,18     | 10,18     |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSH7T121<br>6,52%                         | B       | cte       | RENDA 1   | USO-SOL01 | 0,897 | 0,805 |                |
|                                           | t       | 0,099     | -0,016    | -0,044    |       |       |                |
|                                           | Sig     | 225,4     | -27,46    | -75,43    |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSHPT111<br>4,14%                         | B       | cte       | RENDA 2   | USO-SOL01 | 0,86  | 0,74  |                |
|                                           | t       | -0,004    | 0,06      | 0,022     |       |       |                |
|                                           | Sig     | -5,05     | 62,17     | 22,63     |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSH7T131<br>2,32%                         | B       | cte       | RENDA 2   | USO-SOL02 | 0,776 | 0,602 |                |
|                                           | t       | 0,05      | -0,002    | -0,03     |       |       |                |
|                                           | Sig     | 98,7      | -4,74     | -52,95    |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSH7T141<br>2,10%                         | B       | cte       | USO-SOL02 |           | 0,735 | 0,566 |                |
|                                           | t       | 0,056     | -0,042    |           |       |       |                |
|                                           | Sig     | 78,63     | -52,87    |           |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSHPP121<br>1,96%                         | B       | cte       | RENDA1    | USO-SOL03 | 0,8   | 0,64  |                |
|                                           | t       |           | 0,038     | 0,007     |       |       |                |
|                                           | Sig     |           | 40,36     | 9,08      |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSHPW111<br>1,53%                         | B       | cte       | RENDA1    | USO-SOL01 | 0,969 | 0,94  |                |
|                                           | t       |           | 0,024     | 0,01      |       |       |                |
|                                           | Sig     |           | 100,44    | 51,67     |       |       |                |
|                                           |         | variáveis |           |           |       |       |                |
| HSHPT121<br>1,11%                         | B       | cte       | AUTO      | USO-SOL01 | 0,546 | 0,298 |                |
|                                           | t       | 0,005     | 0,021     |           |       |       |                |
|                                           | Sig     | 10,65     | 29,49     | 0         |       |       |                |

| Renda 1      | AUT        | Renda 2      | Uso_Solo1       | Uso_Solo2       | Uso Solo3          | Uso_Solo4:        |
|--------------|------------|--------------|-----------------|-----------------|--------------------|-------------------|
| RF < 1690; 0 | AUT ≤ 1; 0 | RF < 1462; 0 | 5km < 2,72%; 0  | 5 km < 0,53%; 0 | 20 km < 65,96% ; 0 | 10 km < 44,13%; 0 |
| RF ≥ 1690; 1 | AUT > 1; 1 | RF ≥ 1462; 1 | 5 km ≥ 2,72%; 1 | 5 km ≥ 0,53%; 1 | 20 km ≥ 65,96%; 1  | 10 km ≥ 44,13%; 1 |

### 9.7 ESTUDANTES UNIVERSITÁRIOS

Apresenta-se a aplicação da RLM para o presente caso. O principal objetivo é comprovar a importância das variáveis de uso do solo. Estas foram consideradas fundamentais para segregação dos dados pelo algoritmo CART (Capítulo 8).

#### 9.7.1 Aplicação da RLM e significância das variáveis independentes

A Tabela 9.13 representa a transformação das variáveis independentes selecionadas pelo algoritmo CART em variáveis *dummy* para aplicação da RLM, considerando o caso de estudantes universitários.

Observando a Tabela 9.13, e baseando-se nas matrizes de correlação geradas antes dos modelos lineares, pode-se dizer que as variáveis “RENDA 1”, “RENDA 2” e “RENDA 3” são altamente correlacionadas, assim como as variáveis “USO DO SOLO 1”, “USO DO SOLO 2”, “USO DO SOLO 3” e “USO DO SOLO 4”. Assim, de acordo com a variável dependente, foi selecionada apenas uma variável relacionada à renda e apenas uma variável relacionada ao uso do solo. Os modelos obtidos tinham apenas duas variáveis independentes.

**Tabela 9.13 Variáveis independentes para aplicação da RLM (estudantes universitários)**

| Variáveis independentes                                                           |                         |                                      |                   |
|-----------------------------------------------------------------------------------|-------------------------|--------------------------------------|-------------------|
| Nomenclatura antiga                                                               | Original                | Dummy                                | Nova nomenclatura |
| Parcela acumulada de matrículas em universidades até 10 km do centróide da origem | universidades até 10 km | universidades até 10 km < 44.72% (0) | USO DO SOLO 1     |
|                                                                                   | (contínua - %)          | universidades até 10 km ≥ 44.13% (1) |                   |
| Parcela acumulada de matrículas em universidades até 10 km do centróide da origem | universidades até 10 km | universidades até 10 km < 14.58% (0) | USO DO SOLO 2     |
|                                                                                   | (contínua - %)          | universidades até 10 km ≥ 14.58% (1) |                   |
| Renda Familiar                                                                    | RF                      | RF < 2157 (0)                        | RENDA 1           |
|                                                                                   | (contínua - em reais)   | RF ≥ 2157 (1)                        |                   |
| Renda Familiar                                                                    | RF                      | RF < 4276 (0)                        | RENDA 2           |
|                                                                                   | (contínua - em reais)   | RF ≥ 4276 (1)                        |                   |
| Renda Familiar                                                                    | RF                      | RF < 3950 (0)                        | RENDA 3           |
|                                                                                   | (contínua - em reais)   | RF ≥ 3950 (1)                        |                   |
| Parcela acumulada de matrículas em universidades até 5 km do centróide da origem  | universidades até 5 km  | universidades até 5 km < 16.57% (0)  | USO DO SOLO 3     |
|                                                                                   | (contínua - %)          | universidades até 5 km ≥ 16.57% (1)  |                   |
| Parcela acumulada de matrículas em universidades até 5 km do centróide da origem  | universidades até 5 km  | universidades até 5 km < 21.36% (0)  | USO DO SOLO 4     |
|                                                                                   | (contínua - %)          | universidades até 5 km ≥ 21.36% (1)  |                   |

Como nos demais casos, foram obtidos dez modelos lineares. Os valores de  $R^2$  variaram entre 0.342 (HSHAHPPPP11111) e 0.778 (HSH77111). Conforme esperado, os valores dos coeficientes estimados para variáveis de uso do solo foram os mais altos obtidos levando-se em conta as seis amostras estudadas. (1) - 0.152 (HSHPP141); (2) 0.139 (HSH77111); (3) - 0.164 (HSH77141).

Os resultados obtidos com a aplicação da RLM comprovam a importância dos conjuntos de variáveis de uso do solo no caso de estudantes universitários. Considerando o modelo HSH77141, por exemplo, pode-se afirmar que o fato de indivíduos residirem em zonas com parcela acumulada de matrículas em universidades até 10 km da zona de origem superior ou igual a 14.58% ( $10 \text{ km} \geq 14.58\%$ ; valor um para a variável USO DO SOLO 2) contribui de forma forte e negativa para a realização do padrão HSH77141. Neste caso, espera-se que indivíduos realizem viagens curtas ("111"), considerando o número relativamente alto de universidades até 10 km da sua residência.

Já tomando-se como exemplo o modelo HSH77111, observa-se o valor positivo do coeficiente estimado para variável de uso do solo. Assim, o valor do coeficiente estimado para a variável "USO DO SOLO 3" indica que o fato de residir em uma zona de tráfego com parcela acumulada de matrículas em universidades até 5 km do seu centróide superior ou igual a 16.57% (valor um para a variável USO DO SOLO 3), contribui de forma positiva e forte (0.139) para realização do padrão HSH77111, ou seja, escolha de destinos próximos às residências.

A Tabela 9.14 traz os dez modelos lineares obtidos e seus principais parâmetros. A análise da amostra de estudantes universitários e os resultados gerados com a aplicação RLM, comprovaram a importância das variáveis relacionados ao uso do solo na escolha de destinos mais distantes ou próximos das residências dos indivíduos, considerando a parcela de oportunidades acumuladas na zona de origem e redondezas.

Tabela 9.14 Modelos de regressão obtidos para a amostra de estudantes universitários

| Estudantes universitários (560 indivíduos) |         |           |           |               |       |                |
|--------------------------------------------|---------|-----------|-----------|---------------|-------|----------------|
| Padrões                                    | modelos |           |           |               |       |                |
|                                            |         | variáveis |           |               | R     | R <sup>2</sup> |
| HSHPP121<br>12,36%                         | cte     | RENDA 3   | USO-SOLO1 |               |       |                |
|                                            | B       | 0,083     | 0,076     | 0,011         | 0,62  | 0,385          |
|                                            | t       | 24,47     | 18,05     | 2,67          |       |                |
| Sig                                        | 0       | 0         | 0,008     |               |       |                |
| HSHPP111<br>11,80%                         | cte     | RENDA 3   | USO-SOLO4 |               |       |                |
|                                            | B       | 0,036     | 0,132     | 0,089         | 0,823 | 0,677          |
|                                            | t       | 10,62     | 28,42     | 16,53         |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSH7T121<br>9,63%                          | cte     | RENDA 2   | USO-SOLO2 |               |       |                |
|                                            | B       | 0,081     | -0,091    | 0,071         | 0,779 | 0,61           |
|                                            | t       | 21,71     | -25,62    | 17,42         |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSHPP141<br>8,21%                          | cte     | RENDA 3   | USO-SOLO2 |               |       |                |
|                                            | B       | 0,184     | 0,03      | <b>-0,152</b> | 0,762 | 0,581          |
|                                            | t       | 37,09     | 6,43      | <b>-27,89</b> |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSH7T111<br>8,07%                          | cte     | RENDA 3   | USO-SOLO3 |               |       |                |
|                                            | B       | 0,056     | -0,031    | <b>0,139</b>  | 0,882 | 0,778          |
|                                            | t       | 26,55     | -10,82    | <b>44,12</b>  |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSH7T141<br>7,91%                          | cte     | RENDA 1   | USO-SOLO2 |               |       |                |
|                                            | B       | 0,224     | -0,027    | <b>-0,164</b> | 0,773 | 0,597          |
|                                            | t       | 36,29     | -4,83     | <b>-27,59</b> |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSHWW111<br>6,80%                          | cte     | RENDA 3   | USO-SOLO4 |               |       |                |
|                                            | B       | 0,098     | -0,099    | 0,073         | 0,82  | 0,673          |
|                                            | t       | 40,73     | -30       | 19,04         |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSHPP131<br>6,51%                          | cte     | RENDA 3   | USO-SOLO4 |               |       |                |
|                                            | B       | 0,066     | 0,046     | -0,092        | 0,64  | 0,41           |
|                                            | t       | 20,39     | 10,48     | -17,78        |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSH7T131<br>5,62%                          | cte     | RENDA 2   | USO-SOLO1 |               |       |                |
|                                            | B       | 0,109     | -0,015    | -0,095        | 0,878 | 0,771          |
|                                            | t       | 61,73     | -6,74     | -41,78        |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |
| HSHAHPPPP11111<br>1,05%                    | cte     | RENDA 2   | USO-SOLO4 |               |       |                |
|                                            | B       | 0,01      | 0,011     | -0,016        | 0,584 | 0,342          |
|                                            | t       | 14,51     | 11,54     | -13,98        |       |                |
| Sig                                        | 0       | 0         | 0         |               |       |                |

| USO SOLO 1        | USO SOLO 2        | USO SOLO 3       | USO SOLO 4      | RENDA 1      | RENDA 2      | RENDA 3      |
|-------------------|-------------------|------------------|-----------------|--------------|--------------|--------------|
| 10 km < 44,72%; 0 | 10 km < 14,58%; 0 | 5 km < 16,57%; 0 | 5 km < 21,36; 0 | RF < 2157; 0 | RF < 4276; 0 | RF < 3950; 0 |
| 10 km ≥ 44,72; 1  | 10 km ≥ 14,58%; 1 | 5 km ≥ 16,57%; 1 | 5 km ≥ 21,36; 1 | RF ≥ 2157; 1 | RF ≥ 4276; 1 | RF ≥ 3950; 1 |

## 9.8 CONSIDERAÇÕES FINAIS

Este capítulo, complementar à análise de resultados realizada no capítulo anterior, apresentou a aplicação da RLM (em conjunto com a AD), a fim de mensurar a significância estatística das variáveis independentes, selecionadas pelo algoritmo CART, especialmente do conjunto de variáveis associado às características de uso do solo da RMSP (objetivo secundário do trabalho).

Os coeficientes estimados para as variáveis independentes foram considerados significativamente diferentes de zero (considerando valores de estatística  $t$ ), comprovando a influência e importância de cada variável independente (socioeconômica, participação em atividades e uso do solo) na explicação da variabilidade da variável dependente.

Considerando a principal finalidade da aplicação da RLM, que é comprovar a importância das variáveis de uso do solo e a sua significância, pode-se verificar a sua maior contribuição no caso de industriários e universitários.

Verificou-se, assim, como a distribuição geográfica de atividades (descritas no Capítulo 4) interfere na escolha das seqüências de destinos. Desta forma, quanto maior a dispersão das atividades pela região de estudo (oferta alta em diversas localizações), menor a influência do uso do solo nos itinerários de viagens dos indivíduos.

Pôde-se verificar no presente capítulo a viabilidade do método proposto, ou seja, aplicação conjunta de técnicas de Análise Multivariada (AM) de dados, a fim de estudar relações entre padrões de viagens (variável dependente) e características de uso do solo, socioeconômicas e participação em atividades (variáveis independentes), bem como mensurar a significância estatística das variáveis envolvidas na análise.

## Capítulo

# 10 CONSIDERAÇÕES FINAIS

---

*Este último capítulo do presente texto apresenta três seções principais: (1) Descrição sucinta do trabalho; (2) Principais conclusões e; (3) Recomendações para trabalhos futuros. É descrita a última etapa do método proposto, que consiste nas conclusões a respeito do trabalho. Algumas indagações, que motivaram a realização desta pesquisa, apresentadas no capítulo introdutório, são respondidas a seguir.*

### 10.1. DESCRIÇÃO SUCINTA DO TRABALHO ATUAL

O presente trabalho, que faz parte de uma linha de pesquisa desenvolvida no Departamento de Transportes da EESC-USP sob orientação do professor Dr. Eiji Kawamoto, aspira melhor compreender a diversidade de padrões de atividades e viagens realizados pelos indivíduos e suas relações com características socioeconômicas e do meio urbano.

Com o objetivo principal de encontrar relações entre padrões de encadeamento de viagens urbanas e três grupos de variáveis independentes associados a: (1) Participação em atividades; (2) Características socioeconômicas; e (3) Uso do solo, foi proposto um método compreendido por nove etapas básicas. Oito delas foram descritas nos capítulos anteriores. A etapa final, que consiste nas conclusões a respeito do trabalho, é apresentada no capítulo atual.

Há dois objetivos secundários, que fazem parte do método adotado, que foram fundamentais para atingir a finalidade principal:

- Propor um grupo de variáveis que representem características de uso do solo (definidas como intensidade e distribuição geográfica de atividades no meio urbano);
- Testar a significância do grupo de variáveis de uso do solo ora proposto.

Estes dois objetivos fazem com que o trabalho seja definido como exploratório, pois mensurar características de uso do solo da RMSP, em termos de variáveis independentes, de forma que sua

importância seja capturada pelo algoritmo da Árvore de Decisão (CART), utilizado na pesquisa, é uma tarefa desafiadora.

As etapas que fazem parte do método seguido, anteriores às conclusões, são descritas de maneira sintética nas próximas subseções.

### 10.1.1 Revisão da literatura



A primeira etapa realizada foi a revisão da literatura a respeito de uso do solo e comportamento relacionado ao encadeamento de viagens. Optou-se por dar ênfase maior à influência das características urbanas, já que o principal desafio estava associado à representação das variáveis de uso do solo.

Assim, a revisão bibliográfica dividiu-se em três partes principais, com a finalidade de obter fundamentação teórica a respeito do tema: (1) Uso do solo e comportamento relacionado a viagens – trazendo trabalhos pontuais que corroboram a existência de fortes relações entre estes dois fatores, bem como trabalhos que discordam dessa afirmação; (2) Mensuração de variáveis de uso do solo – analisando trabalhos revisados na literatura que propuseram diferentes índices ou variáveis que representavam características de uso do solo de diversas regiões; e (3) Modelos de distribuição de viagens – dando ênfase ao modelo de oportunidades intervenientes (SCHNEIDER, 1959), que foi o suporte teórico para a proposta do conjunto de variáveis de uso do solo deste trabalho.

### 10.1.2 Escolha das técnicas de AM



O próximo passo consistiu na escolha das técnicas de Análise Multivariada de dados (AM), considerando os objetivos do trabalho. As técnicas foram aplicadas tanto para representar a variável dependente adequadamente (padrões de viagem), como para encontrar relações entre estas e as

variáveis independentes. Foram utilizadas três técnicas de AM, exploratórias e confirmatórias, de forma conjunta:

- (1) **Análise de Cluster (AC)** – Para auxiliar a representação da seqüência de destinos nos padrões de viagem. Os 389 destinos possíveis (número total de zonas de tráfego) foram agrupados segundo a variável “DISTÂNCIA DA PRIMEIRA VIAGEM”. Os destinos foram agrupados de acordo com faixas de distância em relação à origem inicial (residência). Foram obtidos 4 destinos no total (1 - até 5 km da residência; 2 - de 5 a 10 km da residência; 3 – de 10 a 15 km da residência; 4 – acima de 15 km da residência);
- (2) **Árvore de Decisão (AD)** - A fim de encontrar relações entre as variáveis dependente e independentes, formando grupos de indivíduos, consistentes em seus atributos (socioeconômicos, participação em atividades e uso do solo), homogêneos segundo a variável dependente (padrões de viagem);
- (3) **Regressão Linear Múltipla (RLM)** - Associar inferência estatística ou estimar a significância das variáveis independentes, especialmente as variáveis propostas de uso do solo.

### 10.1.3 Estudo de caso

|                       |                            |                |                         |                     |                              |                       |                                                    |            |
|-----------------------|----------------------------|----------------|-------------------------|---------------------|------------------------------|-----------------------|----------------------------------------------------|------------|
| Revisão da literatura | Escolha das Técnicas de AM | Estudo de caso | Variáveis Independentes | Variável dependente | Aplicação de AD e Resultados | Análise de Resultados | Teste de significância das variáveis independentes | Conclusões |
| Cap 2                 | Cap 3                      | Cap 4          | Cap 5                   | Cap 6               | Cap 7                        | Cap 8                 | Cap 9                                              | Cap 10     |

A seguir, observou-se o estudo de caso (RMSP) e separação de amostras para viabilizar a aplicação das técnicas. As amostras utilizadas foram segregadas por setor econômico (no caso de trabalhadores) ou grau de instrução (no caso de estudantes) com o intuito de melhor explicar a contribuição do uso do solo no comportamento relacionado ao encadeamento de viagens.

Verificou-se inicialmente a dispersão geográfica de diferentes atividades na RMSP, a partir disso, partiu-se da hipótese de que a maior ou menor concentração de pólos de empregos ou escolas poderia interferir nas escolhas dos destinos dos indivíduos. Assim fez-se o tratamento de seis amostras no total: (1) trabalhadores no setor industrial (4.102 indivíduos); (2) trabalhadores no setor comercial

(6.043 indivíduos); (3) trabalhadores no setor de serviços (8.328 indivíduos); (4) estudantes matriculados até o 1º grau (13.829 indivíduos); (5) estudantes matriculados no 2º grau (2.143 indivíduos); (6) estudantes universitários (568 indivíduos).

#### 10.1.4 Representação das variáveis independentes

|                       |                            |                |                         |                     |                              |                       |                                                    |            |
|-----------------------|----------------------------|----------------|-------------------------|---------------------|------------------------------|-----------------------|----------------------------------------------------|------------|
| Revisão da literatura | Escolha das Técnicas de AM | Estudo de caso | Variáveis Independentes | Variável dependente | Aplicação de AD e Resultados | Análise de Resultados | Teste de significância das variáveis independentes | Conclusões |
| Cap 2                 | Cap 3                      | Cap 4          | Cap 5                   | Cap 6               | Cap 7                        | Cap 8                 | Cap 9                                              | Cap 10     |

Em seguida foram adotadas ou propostas, baseando-se na literatura vigente e na observação dos deslocamentos dos indivíduos residentes na RMSP, as variáveis independentes. Esta etapa do método foi indispensável para o objetivo do trabalho, pois a representação adequada das variáveis envolvidas é essencial para se encontrar possíveis relações entre estas.

Os conjuntos de variáveis associados a características socioeconômicas e participação em atividades foram definidos baseando-se na literatura, bem como em trabalhos anteriores realizados no Departamento de Transportes da EESC-USP. Já a proposta das variáveis de uso do solo (objetivo secundário), abrangeu algumas etapas e foi fundamentada nos deslocamentos dos indivíduos residentes na RMSP e no modelo de oportunidades intervenientes. Assim, as variáveis independentes foram representadas através de três conjuntos de variáveis:

- (1) Participação em atividades, contendo duas variáveis categóricas: TRABALHA; ESTUDA;
- (2) Características socioeconômicas, contendo sete variáveis numéricas e cinco categóricas: RENDA INDIVIDUAL; RENDA FAMILIAR; NÚMERO DE AUTOMÓVEIS NO DOMICÍLIO; NÚMERO PROVÁVEL DE CARTEIRAS DE HABILITAÇÃO NO DOMICÍLIO; TOTAL DE PESSOAS NA FAMÍLIA; NÚMERO DE CRIANÇAS NO DOMICÍLIO; IDADE; TIPO DE CHEFE DO DOMICÍLIO; POSIÇÃO DO INDIVÍDUO NA FAMÍLIA; SEXO; GRAU DE INSTRUÇÃO; USA VALE TRANSPORTE;
- (3) Uso do solo, contendo quatro variáveis numéricas: PARCELA ACUMULADA DE "OPORTUNIDADES" ATÉ UM RAIOS DE 5 km DO CENTRÓIDE DA ZONA DA RESIDÊNCIA; PARCELA ACUMULADA DE "OPORTUNIDADES" ATÉ UM RAIOS DE 10 km DO CENTRÓIDE DA

ZONA DA RESIDÊNCIA; PARCELA ACUMULADA DE “OPORTUNIDADES” ATÉ UM RAIOS DE 15 km DO CENTRÓIDE DA ZONA DA RESIDÊNCIA; PARCELA ACUMULADA DE “OPORTUNIDADES” ATÉ UM RAIOS DE 20 km DO CENTRÓIDE DA ZONA DA RESIDÊNCIA.

### 10.1.5 Representação da variável dependente

|                       |                            |                |                         |                     |                              |                       |                                                    |            |
|-----------------------|----------------------------|----------------|-------------------------|---------------------|------------------------------|-----------------------|----------------------------------------------------|------------|
| Revisão da literatura | Escolha das Técnicas de AM | Estudo de caso | Variáveis Independentes | Variável dependente | Aplicação de AD e Resultados | Análise de Resultados | Teste de significância das variáveis independentes | Conclusões |
| Cap 2                 | Cap 3                      | Cap 4          | Cap 5                   | Cap 6               | Cap 7                        | Cap 8                 | Cap 9                                              | Cap 10     |

A variável dependente foi codificada em termos de seqüência de motivos de viagem, modos de transporte e destinos. A codificação da variável dependente foi composta por quatro etapas: (1) representação da seqüência de atividades (H – Residência; W – Trabalho; S – Escola; A – Atividade); (2) representação da seqüência de modos de transporte (T – modo de transporte público; P – modo de transporte particular motorizado; N – modo de transporte não motorizado); (3) representação da seqüência de destinos (1 – até 5 km da residência; 2 – entre 5 e 10 km da residência; 3 – entre 10 e 15 km da residência; 4 – acima de 15 km da residência); e (4) combinação das etapas anteriores (representação dos padrões finais abrangendo todos os três atributos de viagem propostos).

Ao final da codificação, foi obtida a variável dependente que possuía 128 categorias (limitação do *software*) para cada uma das seis amostras analisadas.

### 10.1.6 Aplicação da AD e resultados

|                       |                            |                |                         |                     |                              |                       |                                                    |            |
|-----------------------|----------------------------|----------------|-------------------------|---------------------|------------------------------|-----------------------|----------------------------------------------------|------------|
| Revisão da literatura | Escolha das Técnicas de AM | Estudo de caso | Variáveis Independentes | Variável dependente | Aplicação de AD e Resultados | Análise de Resultados | Teste de significância das variáveis independentes | Conclusões |
| Cap 2                 | Cap 3                      | Cap 4          | Cap 5                   | Cap 6               | Cap 7                        | Cap 8                 | Cap 9                                              | Cap 10     |

Enfim, foi feita aplicação da técnica AD para encontrar relações entre as variáveis independentes e variável dependente, escolhidas e representadas nos passos anteriores. O Capítulo 7 apresentou as seis árvores geradas (para cada uma das amostras da análise), assim como sínteses dos relatórios emitidos pelo S-PLUS 6.1.

Foram verificadas as variáveis independentes consideradas importantes pelo algoritmo CART para segregação dos dados (CNH; USA VTRA; ESTUDA; TRABALHA; RF; AUTO; PARCELA ACUMULADA DE OPORTUNIDADES; SEXO; IDADE).

### 10.1.7 Análise de resultados obtidos com aplicação da AD



Foi realizada posteriormente a análise dos resultados obtidos com o processamento das árvores para as seis amostras analisadas. Pôde-se notar uma forte influência das variáveis de uso do solo nas amostras compostas por industriários e universitários. Influência esta, justificada pela concentração de empregos na indústria e de universidades apenas em alguns pólos principais na região de estudo, interferindo assim na escolha de destinos mais distantes para realizar atividades de trabalho na indústria ou estudo em universidades.

Também se verificou a influência das variáveis associadas à participação em atividades (ESTUDA e TRABALHA) na seqüência de atividades realizadas pelos indivíduos, assim como relações entre características socioeconômicas (USA VTRA; RF; AUTO; IDADE) e a seqüência dos modos de transporte utilizados durante as viagens.

### 10.1.8 Teste de significância das variáveis independentes



Foi aplicada uma técnica de análise multivariada tradicional (RLM) para obter dados a respeito da significância das variáveis independentes, sobretudo daquelas relacionadas a características de uso do solo da RMSP (objetivo secundário).

Optou-se por utilizar uma técnica tradicional, juntamente com a AD, pois embora técnicas de mineração de dados descubram padrões em grandes bases de dados, estas não trazem valores estatísticos associados a cada variável. Associar técnicas estatísticas tradicionais a técnicas de mineração de dados (como AD, por exemplo) trouxe contribuições para o trabalho.

Os valores dos coeficientes estimados nos modelos lineares obtidos e valores de estatística t associados às variáveis independentes comprovaram a forte influência das variáveis de uso do solo,

especialmente no caso de industriários e estudantes universitários. A influência de demais variáveis como "RENDA", "USA VTRA", "AUT", "CNH", "IDADE", "TRABALHA", "ESTUDA" também foi ratificada pelos valores dos coeficientes estimados para as referidas variáveis nas amostras consideradas.

Enfim, a última etapa do método proposto é descrita na próxima seção.

## 10.2. PRINCIPAIS CONCLUSÕES



O texto apresentado descreve o trabalho realizado, caracterizado por ser investigativo e exploratório. É uma tentativa de melhor compreender os fatores que os indivíduos (trabalhadores de diferentes setores econômicos e estudantes de diferentes escolaridades) consideram ao escolher a seqüência de atividades, de modos de viagem utilizados e de destinos durante o período de um dia.

Algumas indagações sugeridas no início deste texto, que motivaram a realização do estudo, podem ser respondidas através das seguintes conclusões.

- *Como mensurar padrões de uso do solo e caracterizar as zonas de tráfego da RMSP em termos de intensidade e distribuição de atividades no meio urbano? Como representar características urbanas (uso do solo) em termos de variáveis independentes?*

Mensurar uso do solo, ou intensidade e distribuição de atividades no meio urbano, em termos de variáveis independentes foi um dos principais desafios do trabalho e consistia em um dos objetivos secundários.

Para entender relações entre uso do solo e comportamento referente a viagens, é necessário primeiramente, investigar formas de mensurar características do meio urbano. Alguns autores propuseram diferentes indicadores ou variáveis de uso do solo, finalizando encontrar tais relações (MOUDON et al., 1997; MCNALLY; KULKARNI; 1997; SRINIVASAN; 2000).

Na pesquisa atual, foi proposto um conjunto de variáveis de uso do solo baseado no modelo de oportunidades intervenientes. Assim, as variáveis de uso do solo consideraram parcelas acumuladas de oportunidades (empregos por setor econômico ou escolas por grau de instrução) a partir do centróide das zonas de tráfego da residência dos indivíduos.

Como conclusão a respeito das variáveis propostas, pode-se afirmar que a representação foi adequada já que tal variável sobressaiu-se no modelo CART, separando indivíduos que residiam em zonas de baixas oportunidades e aqueles que residiam em zonas de altas oportunidades. Assim, indivíduos que residem em zonas com baixas oportunidades (seja empregos ou escolas) deslocam-se maiores distâncias para satisfazer suas necessidades de trabalho ou estudo.

- *Como o comportamento referente a viagens encadeadas está relacionado aos padrões de uso do solo dos locais de residência das pessoas?*

Conforme mencionado anteriormente, a partir dos resultados obtidos no trabalho, tanto através do processamento da AD quanto com a aplicação da RLM, características de uso do solo (especialmente no caso de trabalhadores na indústria e universitários) estão relacionadas fortemente à escolha de destinos.

Dependendo da oferta de empregos ou escolas na zona de residência ou vizinhanças, as pessoas optam por viagens mais longas ou curtas a fim de realizar atividades de trabalho ou estudo.

A penúltima etapa do método, que consiste no teste de significância das variáveis independentes, comprovou que variáveis de uso do solo estão mais fortemente relacionadas a escolhas de destinos mais distantes (maiores valores de coeficientes estimados para as variáveis de uso do solo nos modelos HWH 77141, por exemplo).

- *Quais os tipos predominantes de padrões de viagem associados a determinados grupos de indivíduos, conhecidas suas características socioeconômicas e dos seus locais de residência?*

Através da aplicação da AD, foi possível encontrar relações entre variáveis socioeconômicas (*Usa Vale Transporte, Renda Familiar e Nº de automóveis no domicílio, Nº provável de carteira de habilitação no domicílio, Idade*), participação em atividades (*Estuda, Trabalha*), variável de uso do solo (*Parcela acumulada de oportunidades até 5 km do centróide da zona de origem, Parcela acumulada de oportunidades até 10 km do centróide da zona de origem, Parcela acumulada de oportunidades até 15 km do centróide da zona de origem, Parcela acumulada de oportunidades até 20 km do centróide da zona de origem*) e padrões de viagens encadeadas (*seqüência de atividades, modos de transporte e destinos*). A influência das variáveis socioeconômicas refletiu-se principalmente na seqüência de modos de transportes, já a participação em atividades de subsistência influenciou basicamente a seqüência de atividades ou motivos de viagens, enquanto que a variável uso do solo relacionou-se predominantemente com a seqüência de destinos.

Foram obtidas 10 folhas (grupos de indivíduos) nas três amostras de trabalhadores, 9 folhas na amostra de estudantes até o 1º grau e 8 folhas nas amostras de estudantes no 2º grau e universitários. Cada uma das folhas representa um grupo de indivíduos com determinadas características socioeconômicas e de uso do solo das suas zonas de residência. A AD associa a tais grupos de indivíduos, freqüências de ocorrência de todos os padrões de viagens da análise. Desta forma, através dos resultados das seis árvores, foi possível conhecer os padrões de encadeamento de viagens predominantes em cada classe de indivíduos e, conseqüentemente, diferenças relacionadas ao comportamento de viagens em diversos grupos de pessoas.

- *Quais as vantagens de segregar amostras por "Ocupação do indivíduo" ou "Grau de instrução" com o intuito de melhor compreender a influência das variáveis independentes, especialmente daquelas de uso do solo?*

Verificada a distribuição geográfica de diferentes atividades na RMSP (Capítulo 4), intuiu-se que quanto maior a concentração de determinadas atividades em pólos principais (como empregos na indústria, por exemplo), maior a influência do uso do solo nos itinerários de viagens.

Assim, optou-se por segregar amostras no intuito de melhor verificar a influência de localizações geográficas das atividades e escolhas de destinos. A análise de seis amostras, diferenciadas por setor econômico ou grau de instrução, foi proveitosa, evidenciando relações entre uso do solo e comportamento de viajantes.

- *Como testar estatisticamente as variáveis de uso do solo adotadas?*

A fim de testar a significância estatística do conjunto de variáveis de uso do solo ora proposto, optou-se por utilizar a RLM em conjunto com a AD. Desta forma, através dos coeficientes estimados para estas variáveis nos modelos lineares obtidos, bem como valores de estatística t associados, pôde-se verificar se as variáveis selecionadas pelo algoritmo CART foram escolhidas ao acaso para segregação dos dados, ou realmente possuíam forte contribuição na escolha de determinados padrões de viagens.

Corroborou-se a influência das variáveis relacionadas ao uso do solo na escolha de destinos através dos valores (inclusive positivos ou negativos) dos coeficientes estimados.

- *As técnicas de AM escolhidas, utilizadas em conjunto, foram úteis considerando os objetivos do trabalho?*

O método adotado, que consistia na aplicação de três técnicas de AM, considerando os objetivos da pesquisa, mostrou-se viável.

Através da Análise de *Cluster* (AC) foi possível reduzir o número de destinos da área de estudos, tornando-se factível a representação de atributos de viagens em forma de categorias de variável dependente.

Representada a variável dependente, pôde-se em seguida aplicar a segunda técnica de AM, conhecida como Árvore de Decisão (AD). A partir dos resultados gerados pela árvore, foi atingido o objetivo principal da pesquisa, que é encontrar relações entre variável dependente (padrões de

viagens urbanas) e características de uso do solo, socioeconômicas e participação em atividades (variáveis independentes).

Finalmente, com o uso da terceira técnica de AM – Regressão Linear Múltipla (RLM) – se conseguiu atingir o outro objetivo secundário, que foi trazer inferência estatística associada às variáveis independentes, sobretudo aquelas de uso do solo.

- *De que forma estudos como estes podem trazer benefícios para o meio acadêmico, assim como contribuir futuramente para aplicações práticas?*

Finalmente, espera-se que o presente trabalho constitua uma contribuição ao meio acadêmico, seja em termos de representação de intensidade e distribuição geográfica de atividades em áreas urbanas, seja em termos de uso de técnicas de AM para análises semelhantes, para representação das seqüências de viagens realizadas no período de um dia, ou ainda para estudos de relações entre uso do solo e comportamento subjacente ao encadeamento de viagens.

Quanto a aplicações práticas, um dos objetivos deste tipo de linha de pesquisa é obter um modelo de comportamento de viajante que seja útil ao Planejamento de Transportes. O modelo resultante poderia ser aplicado em conjunto com informações disponíveis no IBGE (recenseamento) e nas prefeituras (distribuição espacial de atividades, etc.). Desta forma, poder-se-ia associar grupos de indivíduos, com determinadas características socioeconômicas e de uso do solo a escolhas das seqüências de deslocamentos, fornecendo, assim, alternativas baratas para o Planejamento de Transportes. No entanto ainda há um longo caminho de estudos e análises até posteriores aplicações práticas.

### **10.3. RECOMENDAÇÕES PARA TRABALHOS FUTUROS**

A presente linha de pesquisa é bastante ampla, com uma série de idéias e opções para o prosseguimento do estudo da influência de determinados fatores na diversidade de padrões de

viagens ou atividades realizados pelos indivíduos no meio urbano. São citadas, a seguir, algumas sugestões para pesquisas futuras:

- Análise mais detalhada a respeito da distribuição de viagens na RMSP. No atual trabalho, notou-se que os indivíduos se deslocavam segundo premissas do modelo de oportunidades intervenientes. No entanto, há outros fatores a considerar, como introduzir outras medidas de impedância como distância, tempo de viagem ou custo de viagem percebido pelo usuário, bem como competitividade entre zonas, etc.
- Incluir medidas ou variáveis relacionadas ao sistema de transporte da RMSP. Nesta pesquisa desconsiderou-se característica da distribuição da rede de transportes, por exemplo, pois a inclusão de tais variáveis estaria sujeita à obtenção de dados, além das questões teóricas para incorporação de tais variáveis, inviabilizando a conclusão do trabalho em tempo hábil. Futuramente, sugere-se analisar formas de mensurar variáveis relacionadas a características da rede ou oferta de transporte público, por exemplo, a fim de adicionar à análise mais um conjunto de variáveis independentes.
- Incorporar estatística espacial neste tipo de análise. Com o intuito de aprimorar a compreensão do comportamento relacionado às viagens encadeadas, pode-se propor uma extensão dos estudos realizados no Departamento de Transportes da EESC-USP. Sugere-se, então, incorporar análise espacial de dados ao estudo, considerando que a dependência espacial, presente nas áreas urbanas, exerce efeitos em modelos de previsão de demanda por transportes. Além do mais, aplicar estatística espacial ao caso de amostras desagregadas pode trazer resultados interessantes.
- Adicionar modelagem multinomial ao estudo, bem como outras técnicas de AM tradicionais, a fim de não somente trazer inferência estatística a análises deste tipo, como também elucidar muitas relações inicialmente encontradas através de técnicas de mineração de dados;

- Analisar outras formas de representar variáveis dependentes, onde não haja tantas combinações de atributos e um número tão elevado de categorias (limitação de número de categorias no *software* utilizado). Uma sugestão é analisar cada etapa separadamente, inicialmente a seqüência de atividades, em seguida seqüência de modos de transporte e finalmente, seqüência de destinos.

# REFERÊNCIAS

---

ABBAGNANO, N. (1970) **Dicionário de Filosofia**. Mestre Jou, São Paulo.

AGUIAR, M.F.M. (2005) **Possibilidade de Redução no Tamanho da Amostra de Pesquisas de Entrevistas Domiciliares para Planejamento de Transportes: Uma Verificação Preliminar**. 180 f. Dissertação (mestrado). Escola de Engenharia de São Carlos, Universidade de São Paulo.

ALONSO, W. (1964) **Location and land use**. Harvard University Press. Cambridge.

ARRUDA, F. S. (2005) **Aplicação de um Modelo Baseado em Atividades para Análise da Relação Uso do Solo e Transportes no Contexto Brasileiro**. 145f. Tese (doutorado). Escola de Engenharia de São Carlos, Universidade de São Paulo, 2005.

ASSIS Jr, M.F. (1987) **Avaliação do desempenho e adequação de sistemas de transporte público, por ônibus, de cidades de porte médio em relação a sua estrutura urbana**. 152f. Dissertação (Mestrado) –Universidade de São Paulo, 1987.

AXHAUSEN, K.W. (2000) **Activity-based modeling: research directions and possibilities. New look at multi-modal modeling**. Report for the Department of Environment, Londres, Cambridge e Oxford. Disponível em <[http://e-collection.ethbib.ethz.ch/ecol-pool/incoll/incoll\\_86.pdf](http://e-collection.ethbib.ethz.ch/ecol-pool/incoll/incoll_86.pdf)> Acesso em 05/01/2007.

BALASUBRAMANIAM C.; GOULIAS, K. (1999) Exploratory longitudinal analysis of solo and joint tripmaking. IN: 78th ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 1999, Washington, D.C. Compendium of Papers CD-ROM.

BHAT, C.R.; SRINIVASAN, S. (2005) A multidimensional mixed ordered-response model for analyzing weekend activity participation. **Transportation Research Part B**, v.39, p.255-278.

BHAT, C.R.; SIVAKUMAR, A.; AXHAUSEN, K.W. (2003) An analysis of the impact of information and communication technologies on non-maintenance shopping activities. **Transportation Research Part B**, v.37, p.857-881.

BHAT, C.R.; SINGH, S.K. (2000) A comprehensive daily activity-travel generation model system for workers. **Transportation Research Part A**, v.34, p.1-22.

BHAT, C.R.; KOPPELMAN, F.S. (2000) Activity-based travel demand analysis: history, results and future directions. IN: 79th ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 2000, Washington, D.C. Compendium of Papers CD-ROM.

BHAT, C.R.; KOPPELMAN, F.S. (1991) A conceptual framework of individual activity program generation. **Transportation Research Part A**, v. 27, p.433-446.

BISPO, C.A.F. (1998) **Uma análise de nova geração de sistemas de apoio à decisão**. Dissertação (Mestrado). Escola de Engenharia de São Carlos, Universidade de São Paulo.

BLAKELY, E.J. (1994) **Transformation of the American metropolis**. Institute of urban and regional development. University of California, Berkeley.

BOARNET, M.G.; SARMIENTO, S. (1996) Can land use policy really affect travel behavior? A study of the link between non-work travel and land use characteristics. **Working Paper**. Disponível em <<http://repositories.cdlib.org/itsirvine/casa/UCI-ITS-AS-WP-96-5>>. Acesso em 15 jan. 2007.

BREIMAN, L. et al. (1984) **Classification and Regression Trees**. Wadsworth International Group, Califórnia.

BROWN, S. (1992). **Retail location theories: a micro-scale perspective**. Avebury : Aldershot (Reino Unido).

BRUTON, M.J. (1979). **Introdução ao planejamento dos transportes**. Rio de Janeiro: Interciência.

CASEY, H.J. (1955) Applications to traffic engineering of law of retail gravitation. **Traffic Quarterly**, IX (1), p.23-35.

CERVERO, R.; RADISCH, C. (1996) Pedestrian versus automobile oriented neighborhoods. **Transport Policy**, v.3, p.127-141.

CERVERO, R., KOCKELMAN, K.M. (1997) Travel demand and the three Ds.: Density, Diversity and Design. **Transportation Research Part D**, v.2, p.199-219.

DONG, X. et al. (2002) Analysis of Activity-Based Accessibility. IN: 81 th ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 1999, Washington, D.C. Compendium of Papers CD-ROM.

DOUBLEDAY, C. (1977) Some studies of the temporal stability of person trip generation models. **Transportation Research Part B**, v.11, p. 225-263.

EVANS, J.E.; PERINCHERRY, V; DOUGLAS, G.B. (1997) Transit friendliness factor: approach to quantifying transit access environment in a transportation planning model. **Transportation Research Record**, ISSN 0361-1981, nº 1604, p.32-39.

EWING,R.; HALIYUR,P.; PAGE,G.W. (1994) Getting around a traditional city, a suburban planned unit development, and everything in between. **Transportation Research Record**, 1466, 53-62.

EWING,R. (1995) Beyond density, mode choice, and single-purpose trips. **Transportation Quarterly**, v. 49, p. 16-23.

FRANK, L.D.; PIVO, G. (1994) Impacts of mixed use and density on utilization of three modes of travel: single-occupation vehicle, transit, and walking. **Transportation Research Record**, ISSN 0361-1981, nº 1466, p. 44-52.

FIELD, A. (2005) **Discovering statistics using SPSS**. 2.ed.rev. Sage Publications, Londres, Inglaterra.

FRIEDMAN, B., GORDON, S., PEERS, J. (1994) Effect of neotraditional neighborhood design on travel characteristics. **Transportation Research Record**, ISSN 0361-1981, nº 1466, p. 63-70.

GIULIANO, G. (1995) The weakening transportation - land use connection. **Access**, v.6, p.3-11.

GOLOB, T.F.; MCNALLY, M. (1997) A model of activity participation and travel interactions between household heads. **Transportation Research Part B**, v.31, p.177-194.

GOLOB, T.F. (2001) **Structural Equation Modeling for Travel Behavior Research**. Scholarship Repository, University of California. Disponível em <<http://repositories.cdlib.org/itsirvine/casa/UCI-ITS-AS-WP-01-2>>. Acesso em 15 out. 2005.

GOLOB, T.F.; BOWNSTONE, D. (2005) **The impact of residential density on vehicle usage and energy consumption**. Scholarship Repository, University of California. Disponível em <<http://repositories.cdlib.org/ucei/policy/EPE-011>>. Acesso em 01 fev 2007.

GONÇALVES, M. B.; ULYSSÉA NETO, I. (1993) The Development of a new Gravity – Opportunity Model for Trip Distribution. **Environment and Planning**, v. 25, p. 817-826

GORDON, P.; RICHARDSON, H. (1997) Are compact cities a desirable planning goal? **Journal of the American Planning Association**, v.63, p.95-105.

HAIR, J.F. et al. (1998). **Multivariate Data Analysis**. 5ª ed. Prentice-Hall. Upper Saddle River, New Jersey, 730p.

HANDY, S.L. (1996) Understanding the link between urban form and non work travel behavior. **Journal of Planning Education and Research**, v. 15, p. 183-198.

HANDY, S. (1993) Regional versus local accessibility: implications for non-work travel. **Transportation Research Record**, ISSN 0361-1981, nº 1400, p. 58-66.

HANSON S.; HANSON P. (1980) Gender and Urban Activity Patterns in Uppsala, Sweden. **Geographical Review**, v.7, p.291-299.

HANSON S.; HANSON P. (1981) The travel-activity patterns of urban residents: dimensions and relationships to sociodemographic characteristics. **Economic Geography**, v. 57, p. 332-347.

HILLIER, B. (1996). **Space is the machine: a configurational theory of architecture**. Cambridge University Press. Cambridge (Reino Unido).

HUTCHINSON, B.G. (1979) **Princípios do planejamento dos sistemas de transporte urbano**. Editora Guanabara. Rio de Janeiro.

ICHIKAWA, S.M. (2002) **Aplicação de minerador de dados na obtenção de relações entre padrões de encadeamento de viagens codificados e características sócio-econômicas**. 136 f. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, 2002.

ICHIKAWA, S. M.; PITOMBO, C. S.; KAWAMOTO, E. (2003) Aplicação de minerador de dados na obtenção de relações entre padrões de viagens encadeadas e características socioeconômicas. In: **Transporte em Transformação VII**, v. VII, Cap. 5, p. 95-111.

KASS, G.V. (1980) An exploratory technique for investigating large quantities of categorical data. **Applied Statistics** , N 29, p.119-127.

KAWAMOTO, E. (2003) Transferability of standardized regression model applied to person-based trip generation. **Transportation Planning and Technology**, no 4, v.26, p.331-359.

KEULEERS, B.; WETS, G. (2001) Using association rules to identify patterns in activity diary data. IN: 80th ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 2001, Washington, D.C. Compendium of Papers CD-ROM.

KITAMURA, R. ; MOKHTARIAN, P.L. ; LAIDET, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. **Transportation**. n. 24, p. 125-158.

KITAMURA, R. (1985) Trip chaining in a linear city. **Transportation Research Part A**, v.19, p. 115 – 167.

KRIZEK, K.J. (2003). Neighborhood services, trip purpose, and tour-based travel. **Transportation**. n. 30, p. 387-410.

KÜHLKAMP, N.; ULYSSÉA Neto, I. (2004) Uso de modelos de oportunidades intervenientes na análise da localização de novos pólos atratores de tráfego em áreas urbanas. IN: COBRAC 2004. Congresso Brasileiro de Cadastro Técnico Multifinalitário, Florianópolis.

KWAN, M. (2000) Interactive revisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. **Transportation Research Part C**, v.8, p.185-203.

LEVINSON, D.M.; KUMAR, A. (1997) Density and the journey to work. **Growth and Change**, v.28, p.147-172.

LU, X.; PAS, E.I. (1999) Socio-demographics, activity participation and travel behavior. **Transportation Research Part A**, v.33, p.1-18.

MALHOTRA, N. K. (2001) **Pesquisa de Marketing: uma orientação aplicada**. Editora Bookman, Porto Alegre.

MCGUCKIN N.; MURAKAMI E. (1999) Examining trip-chaining behavior: Comparison of travel by men and women. **Transportation Research Record**, ISSN 0361-1981, nº 1693, p. 79-85.

MCNALLY, M.G.; KULKARNI, A. (1997) Assessment of influence of land use-transportation system on travel behavior. **Transportation Research Record**, ISSN 0361-1981, nº 1607, p. 105-115.

METRÔ (1990) Pesquisa origem e destino 1987: Região Metropolitana de São Paulo, Síntese das informações.

MIDDLESEX-SOMERSET- MERCER REGIONAL COUNCIL. (1992) **The impact of various land use strategies on suburban mobility**. Report FTA-NJ-08-7001-93-1. U.S. Department of Transportation.

MILLIER, E.J.; IBRAHIM, A. (1998) Urban form and vehicular travel – some empirical findings. **Transportation Research Record**, ISSN 0361-1981, nº 1617, p.18-27.

MITCHELL, C. G. B.; TOWN S.W. (1977) **Accessibility of various social groups to different activities**. Transport and Road Research Laboratory, Report 258, Inglaterra.

MONTALI, L. (2003) Relação Família-Trabalho: reestruturação e desemprego. **Revista São Paulo em Perspectiva**, SEADE, v. 17, nº 2, Abr/Jun, p. 123-135.

MOUDON, A.V. et al. (1997) Effects of site design on pedestrian travel in mixed-use, medium-density environments. **Transportation Research Record**, ISSN 0361-1981, nº 1578, p.48-55.

NEWMAN, P., KENWORTHY, J. (1989) Gasoline consumption and cities: a comparison of U.S. cities with a global survey. **APA Journal**, v.55, n.1, p.23-39.

OTÚZAR, J. D.; WILLUM, L.G. (1994). **Modeling Transport**. 2<sup>a</sup> ed. John Wiley & Sons, 439p.

PIRAMUTHU, S. (2007) Input data for decision trees. **Expert Systems with applications**. <[www.elsevier.com/locate/eswa](http://www.elsevier.com/locate/eswa)> Acessado em 15 Jan 2007.

PITOMBO, C.S. (2003) **Análise do comportamento subjacente ao encadeamento de viagens através do uso de minerador de dados**. 148f. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2003.

PITOMBO, C.S.; KAWAMOTO, E. (2003) A posição do indivíduo na família e padrões de encadeamento de viagens urbanas. In: XVII ANPET, Rio de Janeiro. **Proceedings** Panorama Nacional de Pesquisa em Transportes, v.2 – p. 859-870.

PITOMBO, C.S.; KAWAMOTO, E. (2004) Estudo do comportamento individual relacionado ao encadeamento de viagens através do uso de minerador de dados. In: XIII PANAMERICAN CONFERENCE OF TRAFFIC AND TRANSPORTATION ENGINEERING, 2004, Albany. 1 CD-ROM.

PITOMBO, C. S.; KAWAMOTO, E. (2006a). Aplicação de técnicas de análise multivariadas de dados em um estudo exploratório da influência do uso do solo nas escolhas de padrões de encadeamento de viagens. In: **Transporte em Transformação X**. v. X, Cap 5, p. 97-115.

PITOMBO, C. S.; KAWAMOTO, E. (2006b). Uso de análise de *cluster* em uma caracterização de uso do solo alternativa para Região Metropolitana de São Paulo. In: PLURIS 2006, 2006, Braga (Portugal). PLURIS 2006 **PROCEEDINGS** – CD-ROM.

PITOMBO, C. S.; KAWAMOTO, E.; QUINTANILHA, J.A. (2006) Investigação da influência do uso do solo e das características socioeconômicas na seqüência de deslocamentos dos industriários na Região Metropolitana de São Paulo. In: XX ANPET, 2006, Brasília. **Proceedings** Panorama Nacional da Pesquisa em Transportes, v. 1. p. 541-552.

PIVO, G. (1990) The net of mixed beads: suburban office development in six metropolitan regions. **Journal of the American Planning Association**, vol. 56, 457-469.

QUINLAN, J.R. (1993) **C4.5: Programs for Machine Learning**. Morgan Kaufmann Publishers, San Mateo, California.

RECKER, W. W. (1995) The household activity pattern problem: general formulation and solution. **Transportation Research Part B**, v.29, p.61-77.

SAFAVIAN, S.R.; LANDGREBE, D. (1991) A Survey of Decision Tree Classifier Methodology, **Man at Cybernetics**, v.21, p. 660-674.

SARMIENTO, S. (1996) Household, Gender, and Travel. THE WOMEN'S TRAVEL ISSUES SECOND NATIONAL CONFERENCE. Disponível em <http://www.fhwa.dot.gov/ohim/womens/chap3.pdf>> Acessado em 05 set 2002.

SCHNEIDER, M. (1959) Gravity models and trip distribution theory. **Papers and Proceedings of the Regional Science Association**, v.5, p.51-56.

SEADE.< <http://www.seade.gov.br>.> Acessado em 22 jun. 2004.

SILVA, M.A. (2006) **Verificação da aplicabilidade da técnica de mineração de dados na previsão da demanda por transporte de passageiros urbanos usando dados da Região Metropolitana de São Paulo**, 176 f. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2006.

SILVA, A.N.R. (1993) Transporte e uso do solo. In: IV SIMPÓSIO DE ENGENHARIA CIVIL DE ILHA SOLTEIRA, Ilha Solteira (SP) p. 113-121.

SIMMMA, A.; AXHAUSEN, K.W. (2001) Within household allocation of travel: the case of upper Austria. **Transportation Research record**, ISSN 0361-1981, nº 1752, p. 69-75.

SOUSA, P. B. (2004) **Análise comparativa do encadeamento de viagens de três áreas urbanas**. 130f. Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2004.

SRINIVASAN, K.K.; ATHURU, S.R. (2005) Analysis of within-household effects and between household

differences in maintenance activity allocation. **Transportation**, v.32, p.495-521.

SRINIVASAN, S. (2000) **Linking Land Use and Transportation: Measuring the Impact of Neighborhood-scale Spatial Patterns on Travel Behavior**. 248 f. Tese (doutorado). Harvard University, 2000.

STRAMBI, O.; VESPUCCI, K. M; BILT, K. Van de (2004) Analysis of the evolution of classes of individual patterns and their relation to socio-demographic and economic variables. IN: 83RD ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 2004, Washington, D.C. Compendium of Papers CD-ROM.

STRAMBI, O.; VAN DE BILT K.; PIETRANTONIO, H. (1995) Utilização do CHAID – Método de Detecção Automática de Interações – Para Análise de Taxas de Produção de Viagens. In: IX ANPET, 1995. **Proceedings** Panorama Nacional da Pesquisa em Transportes, v. 2. p. 492-506.

STRAMBI, O.; VAN DE BILT K. (1998) Tendências sócio-demográficas das mulheres e suas implicações para análise da demanda e elaboração de políticas de transportes. In: Ingeniería de Tránsito y Transporte, X Congreso Panamericano. **Proceedings** p. 57-68

STRATHMAN, J.G.; DUEKER, K.J. (1995) **Understanding Trip Chaining**. 1990 NPTS Special Reports on Trip and Vehicle Attributes. Report FHWA-PL-95-033.

SUN, X.; WILMOT, C.G.; KASTURI, T. (1998) Household travel, household characteristics, and land use: An empirical study from the 1994 Portland Activity-Based Travel survey. **Transportation Research record**, ISSN 0361-1981, nº 1617, p.10-17.

ULYSSEÁ Neto, I.; KÜHLKAMP, N. (2002) O tratamento do espaço em modelos de interação especial – considerações sobre a determinação das oportunidades intervenientes. IN: COBRAC 2002. Congresso Brasileiro de Cadastro Técnico Multifinalitário.

VAN DER SMAGT, T.; LUCARDIE, L. (1991) Decision Making under Not-Well-Defined Conditions: From Data Processing to Logical Modelling. **Tijdschrift voor Economische en Sociale Geografie**, v.82, p.295-304.

WEE, B. VAN (2002). Land use and transport: research and policy challenges. **Journal of Transport Geography**, v. 10, n. 4, p. 259-271.

WEGENER, M.; FÜRST, F. (1999) **Land-use transport interactions: state of the art**. Institut für Raumplanung, Universität Dortmund, Dortmund. Disponível em <<http://www.inro.tro.nl/transland/Deriverable%202a.pdf>> Acesso em 22/07/2004.

WESTPHAL, C.; BLAXTON, T. (1998) **Data Mining Solutions: methods and tools for solving real-word problems**, John Wiley & Sons, Inc.

WILLS, M.J. (1986) A flexible gravity-opportunities model for trip distribution. **Transportation Research Part B**, v.20, p.89-111.

YAMAMOTO, T.; KITAMURA, R.; FUJI, J. (2001) An Analysis of Drivers' Route Choice Behavior by Data Mining Algorithms. IN: 81<sup>st</sup> ANNUAL MEETING OF TRANSPORTATION RESEARCH BOARD, 2001, Washington, D.C. Compendium of Papers CD-ROM.

YE, X.; PENDYALA, R.M.; GOTTARDI, G. (2007) An exploration of the relationship between mode choice and complexity of trip chaining patterns. **Transportation Research Part B**, v.41, p.96-113.

# ANEXO A

## Probabilidade de ocorrência de todos os padrões de viagens – proveniente do relatório gerado pelo S-PLUS 6.1 (nós terminais)

### 1) AMOSTRA DE TRABALHADORES NO SETOR INDUSTRIAL

| Nó | HAHAHPPPP11111 | HAHNN111       | HAHNT111    | HAHPP111       | HAHPP121       | HAHPP131    | HAHPP141    | HAHTP121    | HAHTT111     | HAHTT121      | HAHTT131   | HAHTT141 | HAMHTTT1441 | HSHNN111 | HSHPP141 |
|----|----------------|----------------|-------------|----------------|----------------|-------------|-------------|-------------|--------------|---------------|------------|----------|-------------|----------|----------|
| 8  | 0,0000         | 0,0017         | 0,0017      | 0,0000         | 0,0035         | 0,0000      | 0,0000      | 0,0000      | 0,0104       | 0,0000        | 0,0017     | 0,0063   | 0,0052      | 0,0052   | 0,0000   |
| 9  | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0059       | 0,0119        | 0,0059     | 0,0000   | 0,0000      | 0,0030   | 0,0000   |
| 20 | 0,0000         | 0,0063         | 0,0034      | 0,0000         | 0,0017         | 0,0000      | 0,0000      | 0,0017      | 0,0051       | 0,0051        | 0,0034     | 0,0051   | 0,0000      | 0,0000   | 0,0000   |
| 21 | 0,0000         | 0,0053         | 0,0000      | 0,0027         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0134       | 0,0134        | 0,0027     | 0,0000   | 0,0000      | 0,0027   | 0,0000   |
| 11 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        | 0,0000     | 0,0275   | 0,0000      | 0,0453   | 0,0000   |
| 12 | 0,0000         | 0,0067         | 0,0000      | 0,0083         | 0,0022         | 0,0000      | 0,0044      | 0,0000      | 0,0022       | 0,0044        | 0,0022     | 0,0022   | 0,0000      | 0,0000   | 0,0000   |
| 52 | 0,0000         | 0,0040         | 0,0000      | 0,0020         | 0,0040         | 0,0020      | 0,0000      | 0,0020      | 0,0040       | 0,0000        | 0,0040     | 0,0060   | 0,0000      | 0,0020   | 0,0000   |
| 53 | 0,0024         | 0,0073         | 0,0000      | 0,0097         | 0,0122         | 0,0024      | 0,0000      | 0,0024      | 0,0000       | 0,0000        | 0,0000     | 0,0024   | 0,0000      | 0,0024   | 0,0000   |
| 27 | 0,0021         | 0,0063         | 0,0000      | 0,0211         | 0,0021         | 0,0021      | 0,0106      | 0,0000      | 0,0000       | 0,0021        | 0,0042     | 0,0000   | 0,0000      | 0,0000   | 0,0000   |
| 7  | 0,0038         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0038      | 0,0000      | 0,0000      | 0,0038       | 0,0000        | 0,0000     | 0,0000   | 0,0000      | 0,0380   | 0,0114   |
| Nó | HSHSHNNNN11111 | HSHSHPPPP11111 | HSHTT111    | HSHWHNNTT11111 | HSSWHPPPP11111 | HSHWHPP1111 | HSHWHPP1121 | HSHWHPP1131 | HWAHNNNN1111 | HWAHPP1111    | HWAHPP1211 |          |             |          |          |
| 8  | 0,0017         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        | 0,0000     |          |             |          |          |
| 9  | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0030        | 0,0000     |          |             |          |          |
| 20 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        | 0,0017     |          |             |          |          |
| 21 | 0,0027         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0053        | 0,0027     |          |             |          |          |
| 11 | 0,0000         | 0,0000         | 0,0000      | 0,0092         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        | 0,0000     |          |             |          |          |
| 12 | 0,0000         | 0,0000         | 0,0044      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        | 0,0000     |          |             |          |          |
| 52 | 0,0000         | 0,0060         | 0,0000      | 0,0000         | 0,0040         | 0,0020      | 0,0000      | 0,0000      | 0,0000       | 0,0000        | 0,0020     |          |             |          |          |
| 53 | 0,0049         | 0,0024         | 0,0000      | 0,0024         | 0,0024         | 0,0097      | 0,0000      | 0,0024      | 0,0000       | 0,0024        | 0,0000     |          |             |          |          |
| 27 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0063      | 0,0085      | 0,0042      | 0,0000       | 0,0042        | 0,0063     |          |             |          |          |
| 7  | 0,0000         | 0,0000         | 0,0038      | 0,0076         | 0,0000         | 0,0038      | 0,0038      | 0,0000      | 0,0000       | 0,0038        | 0,0038     |          |             |          |          |
| Nó | HWAHPP11321    | HWAHTNT1221    | HWAHTNT1331 | HWAHTTN1211    | HWAHTTN1311    | HWAHTTT1111 | HWAHTTT1221 | HWAHTTT1421 | HWAHTTT1441  | HWAHNNNN11111 |            |          |             |          |          |
| 8  | 0,0000         | 0,0000         | 0,0000      | 0,0017         | 0,0017         | 0,0017      | 0,0000      | 0,0052      | 0,0035       | 0,0000        |            |          |             |          |          |
| 9  | 0,0000         | 0,0000         | 0,0030      | 0,0059         | 0,0030         | 0,0059      | 0,0030      | 0,0000      | 0,0030       | 0,0000        |            |          |             |          |          |
| 20 | 0,0000         | 0,0017         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        |            |          |             |          |          |
| 21 | 0,0000         | 0,0000         | 0,0027      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0080        |            |          |             |          |          |
| 11 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0092      | 0,0000      | 0,0000       | 0,0092        |            |          |             |          |          |
| 12 | 0,0000         | 0,0044         | 0,0000      | 0,0000         | 0,0022         | 0,0022      | 0,0022      | 0,0000      | 0,0000       | 0,0000        |            |          |             |          |          |
| 52 | 0,0080         | 0,0000         | 0,0000      | 0,0020         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0020       | 0,0020        |            |          |             |          |          |
| 53 | 0,0000         | 0,0000         | 0,0024      | 0,0000         | 0,0000         | 0,0000      | 0,0024      | 0,0000      | 0,0000       | 0,0000        |            |          |             |          |          |
| 27 | 0,0021         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        |            |          |             |          |          |
| 7  | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0038        |            |          |             |          |          |
| Nó | HWAHPP11321    | HWAHTNT1221    | HWAHTNT1331 | HWAHTTN1211    | HWAHTTN1311    | HWAHTTT1111 | HWAHTTT1221 | HWAHTTT1421 | HWAHTTT1441  | HWAHNNNN11111 |            |          |             |          |          |
| 8  | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0017       | 0,0063        |            |          |             |          |          |
| 9  | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0059      | 0,0030       | 0,0000        |            |          |             |          |          |
| 20 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0051      | 0,0000       | 0,0017        |            |          |             |          |          |
| 21 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0027      | 0,0000       | 0,0000        |            |          |             |          |          |
| 11 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000        |            |          |             |          |          |
| 12 | 0,0000         | 0,0000         | 0,0000      | 0,0000         | 0,0000         | 0,0022      | 0,0000      | 0,0067      | 0,0044       | 0,0022        |            |          |             |          |          |
| 52 | 0,0000         | 0,0000         | 0,0000      | 0,0020         | 0,0000         | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0040        |            |          |             |          |          |
| 53 | 0,0000         | 0,0024         | 0,0049      | 0,0073         | 0,0000         | 0,0024      | 0,0000      | 0,0000      | 0,0000       | 0,0000        |            |          |             |          |          |
| 27 | 0,0063         | 0,0042         | 0,0021      | 0,0042         | 0,0042         | 0,0042      | 0,0042      | 0,0000      | 0,0000       | 0,0021        |            |          |             |          |          |
| 7  | 0,0038         | 0,0076         | 0,0038      | 0,0000         | 0,0000         | 0,0000      | 0,0038      | 0,0000      | 0,0000       | 0,0038        |            |          |             |          |          |

| Nó | H00H0HNNNN11111 | H00H0HNPP11121 | H00H0HNNTT11121 | H00H0HPPPP11111 | H00H0HPPPP11121 | H00H0HPPPP12111 | H00H0HPPPP13111 | H00H0HPPPP13121 | H00H0HPPPP13141 |
|----|-----------------|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 8  | 0,0000          | 0,0000         | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          |
| 9  | 0,0030          | 0,0000         | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          |
| 20 | 0,0068          | 0,0017         | 0,0051          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          |
| 21 | 0,0107          | 0,0000         | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          |
| 11 | 0,0000          | 0,0000         | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          |
| 12 | 0,0000          | 0,0000         | 0,0000          | 0,0000          | 0,0022          | 0,0022          | 0,0067          | 0,0044          | 0,0000          |
| 52 | 0,0000          | 0,0000         | 0,0020          | 0,0040          | 0,0000          | 0,0040          | 0,0040          | 0,0020          | 0,0020          |
| 53 | 0,0000          | 0,0024         | 0,0000          | 0,0097          | 0,0024          | 0,0097          | 0,0000          | 0,0000          | 0,0000          |
| 27 | 0,0000          | 0,0021         | 0,0000          | 0,0106          | 0,0042          | 0,0085          | 0,0021          | 0,0000          | 0,0042          |
| 7  | 0,0000          | 0,0000         | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          |

| Nó | H00H0HTTNN11111 | H00H0HTTTP12111 | H00H0HTTTP14111 | H00H0HTTTT14111 | H00H0HTTTT14121 | H00H0NN111 | H00H0NN121 | H00H0NN131 | H00H0NN141 | H00H0NP111 |
|----|-----------------|-----------------|-----------------|-----------------|-----------------|------------|------------|------------|------------|------------|
| 8  | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0035          | 0,0156     | 0,0000     | 0,0017     | 0,0000     | 0,0035     |
| 9  | 0,0089          | 0,0000          | 0,0000          | 0,0030          | 0,0000          | 0,0564     | 0,0000     | 0,0000     | 0,0000     | 0,0000     |
| 20 | 0,0000          | 0,0000          | 0,0000          | 0,0017          | 0,0000          | 0,2730     | 0,0273     | 0,0017     | 0,0034     | 0,0034     |
| 21 | 0,0027          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,3984     | 0,0107     | 0,0000     | 0,0000     | 0,0027     |
| 11 | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0092          | 0,1284     | 0,0000     | 0,0000     | 0,0000     | 0,0000     |
| 12 | 0,0000          | 0,0022          | 0,0067          | 0,0000          | 0,0000          | 0,0200     | 0,0000     | 0,0000     | 0,0000     | 0,0000     |
| 52 | 0,0000          | 0,0000          | 0,0020          | 0,0020          | 0,0000          | 0,1096     | 0,0100     | 0,0040     | 0,0020     | 0,0020     |
| 53 | 0,0000          | 0,0049          | 0,0024          | 0,0000          | 0,0000          | 0,0827     | 0,0024     | 0,0000     | 0,0000     | 0,0024     |
| 27 | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0317     | 0,0000     | 0,0000     | 0,0021     | 0,0000     |
| 7  | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0000          | 0,0532     | 0,0000     | 0,0000     | 0,0000     | 0,0038     |

| Nó | H00H0NT111 | H00H0PN111 | H00H0PP111 | H00H0PP121 | H00H0PP131 | H00H0PP141 | H00H0PT111 | H00H0PT121 | H00H0PT131 | H00H0PT141 | H00H0SHNNNN11111 | H00H0SHNNPP11111 |
|----|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------------|------------------|
| 8  | 0,0000     | 0,0000     | 0,0000     | 0,0035     | 0,0000     | 0,0000     | 0,0000     | 0,0017     | 0,0017     | 0,0017     | 0,0035           | 0,0000           |
| 9  | 0,0000     | 0,0000     | 0,0059     | 0,0030     | 0,0030     | 0,0030     | 0,0030     | 0,0000     | 0,0000     | 0,0000     | 0,0000           | 0,0000           |
| 20 | 0,0034     | 0,0000     | 0,0171     | 0,0120     | 0,0000     | 0,0102     | 0,0017     | 0,0017     | 0,0000     | 0,0000     | 0,0000           | 0,0000           |
| 21 | 0,0080     | 0,0000     | 0,0134     | 0,0027     | 0,0027     | 0,0134     | 0,0000     | 0,0000     | 0,0000     | 0,0027     | 0,0000           | 0,0000           |
| 11 | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,0092     | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,2385           | 0,0092           |
| 12 | 0,0067     | 0,0022     | 0,0466     | 0,0444     | 0,0222     | 0,0333     | 0,0022     | 0,0044     | 0,0022     | 0,0067     | 0,0000           | 0,0000           |
| 52 | 0,0000     | 0,0020     | 0,1235     | 0,0936     | 0,0558     | 0,0976     | 0,0020     | 0,0000     | 0,0020     | 0,0000     | 0,0000           | 0,0000           |
| 53 | 0,0000     | 0,0049     | 0,1946     | 0,0973     | 0,0779     | 0,0268     | 0,0000     | 0,0024     | 0,0000     | 0,0000     | 0,0000           | 0,0000           |
| 27 | 0,0000     | 0,0000     | 0,1924     | 0,1818     | 0,0909     | 0,1860     | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,0000           | 0,0000           |
| 7  | 0,0000     | 0,0000     | 0,0646     | 0,0342     | 0,0152     | 0,0228     | 0,0038     | 0,0000     | 0,0000     | 0,0000     | 0,0342           | 0,0076           |

| Nó | H00H0SHNNNTT11111 | H00H0SHNNNTT11121 | H00H0SHPPNN11111 | H00H0SHPPNN12111 | H00H0SHPPPP11111 | H00H0SHPPPP11121 | H00H0SHPPPP12111 | H00H0SHPPPP12121 | H00H0SHPPPP13111 |
|----|-------------------|-------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 8  | 0,0000            | 0,0000            | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 9  | 0,0030            | 0,0000            | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 20 | 0,0000            | 0,0000            | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 21 | 0,0000            | 0,0000            | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 11 | 0,0092            | 0,0092            | 0,0092           | 0,0092           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 12 | 0,0000            | 0,0000            | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 52 | 0,0000            | 0,0000            | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 53 | 0,0000            | 0,0000            | 0,0024           | 0,0000           | 0,0000           | 0,0024           | 0,0000           | 0,0024           | 0,0000           |
| 27 | 0,0000            | 0,0000            | 0,0021           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0042           | 0,0000           |
| 7  | 0,0076            | 0,0076            | 0,0038           | 0,0076           | 0,0076           | 0,0076           | 0,0152           | 0,0228           | 0,0114           |

| Nó | H00H0SHPPPP14111 | H00H0SHTTNN11111 | H00H0SHTTNN12111 | H00H0SHTTNN13111 | H00H0SHTTNN14111 | H00H0SHTTTP11111 | H00H0SHTTTP14111 | H00H0SHTTTT11111 | H00H0SHTTTT11121 |
|----|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 8  | 0,0000           | 0,0000           | 0,0069           | 0,0087           | 0,0035           | 0,0000           | 0,0017           | 0,0035           | 0,0000           |
| 9  | 0,0000           | 0,0208           | 0,0059           | 0,0000           | 0,0000           | 0,0000           | 0,0030           | 0,0000           | 0,0030           |
| 20 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 21 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 11 | 0,0000           | 0,0367           | 0,0184           | 0,0459           | 0,0000           | 0,0000           | 0,0000           | 0,0184           | 0,0000           |
| 12 | 0,0000           | 0,0000           | 0,0022           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 52 | 0,0020           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 53 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 27 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0021           | 0,0000           |
| 7  | 0,0114           | 0,0228           | 0,0342           | 0,0114           | 0,0266           | 0,0114           | 0,0114           | 0,0152           | 0,0152           |

| Nó | HWSHTTTT12111  | HWSHTTTT13111  | HWSHTTTT14111  | HWSHTN111      | HWSHTP111      | HWSHTP121      | HWSHTP141      | HWSHTT111      | HWSHTT121      | HWSHTT131      | HWSHTT141      | HWSHTN111111   | HWSHTN12121 |
|----|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|-------------|
| 8  | 0,0017         | 0,0035         | 0,0035         | 0,0017         | 0,0000         | 0,0017         | 0,0017         | 0,1505         | 0,2197         | 0,1644         | 0,3131         | 0,0017         | 0,0000      |
| 9  | 0,0030         | 0,0000         | 0,0000         | 0,0059         | 0,0059         | 0,0119         | 0,0000         | 0,3680         | 0,2404         | 0,1039         | 0,0534         | 0,0178         | 0,0000      |
| 20 | 0,0000         | 0,0000         | 0,0000         | 0,0017         | 0,0034         | 0,0000         | 0,0000         | 0,1451         | 0,1365         | 0,0734         | 0,1655         | 0,0512         | 0,0034      |
| 21 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0027         | 0,0027         | 0,0027         | 0,1390         | 0,0989         | 0,0588         | 0,0214         | 0,1390         | 0,0000      |
| 11 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0551         | 0,0275         | 0,0459         | 0,0551         | 0,0734         | 0,0000      |
| 12 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0089         | 0,0044         | 0,0022         | 0,1397         | 0,1840         | 0,1419         | 0,2239         | 0,0044         | 0,0000      |
| 52 | 0,0000         | 0,0000         | 0,0000         | 0,0020         | 0,0040         | 0,0060         | 0,0020         | 0,0518         | 0,0916         | 0,0896         | 0,1175         | 0,0319         | 0,0020      |
| 53 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0024         | 0,0049         | 0,0000         | 0,0876         | 0,1046         | 0,0365         | 0,0292         | 0,0535         | 0,0000      |
| 27 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0021         | 0,0021         | 0,0021         | 0,0169         | 0,0169         | 0,0254         | 0,0127         | 0,0211         | 0,0000      |
| 7  | 0,0190         | 0,0038         | 0,0038         | 0,0000         | 0,0000         | 0,0038         | 0,0038         | 0,0570         | 0,0418         | 0,0342         | 0,0532         | 0,0190         | 0,0000      |
| Nó | HWSHTNPPN11111 | HWSHTNPPN11111 | HWSHTNPPN11111 | HWSHTNPPN12121 | HWSHTNPPN13131 | HWSHTNPPN11111 | HWSHTNPPN11111 | HWSHTNPPN12111 | HWSHTNPPN12111 | HWSHTNPPN12111 | HWSHTNPPN12111 | HWSHTNPPN12111 |             |
| 8  | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0017         | 0,0000         | 0,0000         |             |
| 9  | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |             |
| 20 | 0,0034         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |             |
| 21 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0027         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |             |
| 11 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0092         | 0,0092         | 0,0000         | 0,0000         | 0,0000         |             |
| 12 | 0,0000         | 0,0000         | 0,0000         | 0,0067         | 0,0022         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |             |
| 52 | 0,0020         | 0,0000         | 0,0000         | 0,0080         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |             |
| 53 | 0,0000         | 0,0049         | 0,0389         | 0,0024         | 0,0000         | 0,0000         | 0,0000         | 0,0024         | 0,0000         | 0,0049         | 0,0000         | 0,0000         |             |
| 27 | 0,0000         | 0,0021         | 0,0423         | 0,0085         | 0,0042         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0021         | 0,0000         | 0,0000         |             |
| 7  | 0,0000         | 0,0000         | 0,0114         | 0,0038         | 0,0038         | 0,0038         | 0,0038         | 0,0076         | 0,0038         | 0,0038         | 0,0152         | 0,0152         |             |
| Nó | HWSHTPPN13111  | HWSHTPPN13211  | HWSHTPPN14311  | HWSHTPPN14411  | HWSHTNT1221    | HWSHTTT1121    | HWSHTTT12111   | HWSHTTT13211   | HWSHTTT14111   | HWSHTTT14411   | HWSHTPPN11111  |                |             |
| 8  | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0035         | 0,0000         | 0,0000         | 0,0017         | 0,0035         | 0,0017         | 0,0000         |                |             |
| 9  | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0030         | 0,0000         | 0,0000         | 0,0000         |                |             |
| 20 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0017         |                |             |
| 21 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |                |             |
| 11 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0092         | 0,0092         | 0,0275         | 0,0092         | 0,0092         | 0,0092         | 0,0000         |                |             |
| 12 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |                |             |
| 52 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |                |             |
| 53 | 0,0024         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0024         |                |             |
| 27 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0063         |                |             |
| 7  | 0,0076         | 0,0114         | 0,0152         | 0,0152         | 0,0038         | 0,0114         | 0,0076         | 0,0038         | 0,0114         | 0,0190         | 0,0000         |                |             |

# ANEXO B

Probabilidade de ocorrência de todos os padrões de viagens – proveniente do relatório gerado pelo S-PLUS 6.1 (nós terminais)

## 2) AMOSTRA DE TRABALHADORES NO SETOR COMERCIAL

| Nó | HAHPPP1111  | HAHPPP1441  | HAHTNT1111    | HAHTTT1111    | HAHAHPPP1111 | HAHNN111    | HAHNN121    | HAHPP111    | HAHPP121    | HAHPP131    | HAHPP141     | HAHTT111     | HAHTT121 |
|----|-------------|-------------|---------------|---------------|--------------|-------------|-------------|-------------|-------------|-------------|--------------|--------------|----------|
| 8  | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0016      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0096       | 0,0080   |
| 18 | 0,0000      | 0,0000      | 0,0000        | 0,0013        | 0,0000       | 0,0102      | 0,0000      | 0,0013      | 0,0000      | 0,0013      | 0,0000       | 0,0114       | 0,0025   |
| 19 | 0,0017      | 0,0000      | 0,0051        | 0,0017        | 0,0000       | 0,0530      | 0,0051      | 0,0137      | 0,0034      | 0,0034      | 0,0034       | 0,0667       | 0,0376   |
| 10 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0024      | 0,0000      | 0,0000      | 0,0144      | 0,0000      | 0,0000       | 0,0048       | 0,0024   |
| 44 | 0,0015      | 0,0015      | 0,0000        | 0,0015        | 0,0029       | 0,0190      | 0,0000      | 0,0321      | 0,0190      | 0,0044      | 0,0248       | 0,0175       | 0,0131   |
| 45 | 0,0029      | 0,0057      | 0,0000        | 0,0000        | 0,0029       | 0,0029      | 0,0029      | 0,0143      | 0,0171      | 0,0086      | 0,0086       | 0,0029       | 0,0029   |
| 46 | 0,0035      | 0,0012      | 0,0000        | 0,0012        | 0,0012       | 0,0151      | 0,0000      | 0,0325      | 0,0162      | 0,0023      | 0,0046       | 0,0093       | 0,0128   |
| 47 | 0,0000      | 0,0000      | 0,0027        | 0,0000        | 0,0014       | 0,0081      | 0,0000      | 0,0271      | 0,0136      | 0,0000      | 0,0027       | 0,0041       | 0,0027   |
| 6  | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       | 0,0000   |
| 7  | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0044      | 0,0000      | 0,0059      | 0,0030      | 0,0000      | 0,0030       | 0,0044       | 0,0030   |
| Nó | HAHTT131    | HAHTT141    | HAHWHPPPP1111 | HAHWHPPPP1111 | HSNN111      | HSPP111     | HSPP121     | HSNNNN1111  | HSHPPPP1111 | HSHTT111    | HSHTT121     | HSHTT131     |          |
| 8  | 0,0000      | 0,0113      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0016       | 0,0016       |          |
| 18 | 0,0038      | 0,0038      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0025      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 19 | 0,0239      | 0,0222      | 0,0000        | 0,0000        | 0,0137       | 0,0000      | 0,0017      | 0,0205      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 10 | 0,0024      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0024      | 0,0024      | 0,0000      | 0,0000       | 0,0000       |          |
| 44 | 0,0073      | 0,0131      | 0,0000        | 0,0029        | 0,0073       | 0,0087      | 0,0000      | 0,0029      | 0,0058      | 0,0015      | 0,0015       | 0,0000       |          |
| 45 | 0,0057      | 0,0200      | 0,0029        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0029      | 0,0057      | 0,0000      | 0,0000       | 0,0000       |          |
| 46 | 0,0035      | 0,0023      | 0,0035        | 0,0035        | 0,0046       | 0,0035      | 0,0000      | 0,0046      | 0,0104      | 0,0000      | 0,0000       | 0,0000       |          |
| 47 | 0,0000      | 0,0000      | 0,0014        | 0,0014        | 0,0000       | 0,0095      | 0,0014      | 0,0000      | 0,0027      | 0,0000      | 0,0000       | 0,0000       |          |
| 6  | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0170       | 0,0034      | 0,0034      | 0,0000      | 0,0000      | 0,0034      | 0,0068       | 0,0000       |          |
| 7  | 0,0000      | 0,0000      | 0,0000        | 0,0015        | 0,0679       | 0,0133      | 0,0089      | 0,0059      | 0,0000      | 0,0192      | 0,0148       | 0,0044       |          |
| Nó | HSNNNN1111  | HSNNNTT1111 | HSNNNTT1121   | HSNNPPP1111   | HSNNPPP1121  | HSNNNN1111  | HSNNPPP1111 | HSNNPPP1121 | HSNNPPP1141 | HSNNPPP1111 | HSNNPPP1111  |              |          |
| 8  | 0,0000      | 0,0016      | 0,0016        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 18 | 0,0013      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0025      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0038       |          |
| 19 | 0,0017      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0017      | 0,0017      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 10 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 44 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0029      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 45 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 46 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0023       | 0,0046      | 0,0023      | 0,0000      | 0,0046      | 0,0023      | 0,0023       | 0,0023       |          |
| 47 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0027       | 0,0000      | 0,0000      | 0,0068      | 0,0068      | 0,0014      | 0,0000       | 0,0000       |          |
| 6  | 0,0034      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0034      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 7  | 0,0354      | 0,0059      | 0,0044        | 0,0000        | 0,0000       | 0,0015      | 0,0000      | 0,0000      | 0,0030      | 0,0000      | 0,0000       | 0,0015       |          |
| Nó | HSNNPPP1111 | HSNNPPP1111 | HSNNPPP121    | HSNNPPP1321   | HSNNPPP1331  | HSNNNTT1111 | HSNNNTT1331 | HSNNNTT141  | HSNNNTT1111 | HSNNNTT1221 | HSNNNNNN1111 | HSNNNNPP1111 |          |
| 8  | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0048      | 0,0032      | 0,0032      | 0,0048      | 0,0000       | 0,0000       |          |
| 18 | 0,0013      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0013      | 0,0000      | 0,0025       | 0,0000       |          |
| 19 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0017      | 0,0000      | 0,0051       | 0,0000       |          |
| 10 | 0,0000      | 0,0024      | 0,0000        | 0,0000        | 0,0024       | 0,0000      | 0,0000      | 0,0024      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 44 | 0,0000      | 0,0015      | 0,0000        | 0,0015        | 0,0015       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 45 | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 46 | 0,0023      | 0,0081      | 0,0035        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0023       | 0,0035       |          |
| 47 | 0,0014      | 0,0064      | 0,0041        | 0,0041        | 0,0041       | 0,0000      | 0,0000      | 0,0000      | 0,0014      | 0,0000      | 0,0014       | 0,0014       |          |
| 6  | 0,0000      | 0,0000      | 0,0000        | 0,0000        | 0,0000       | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000      | 0,0000       | 0,0000       |          |
| 7  | 0,0000      | 0,0030      | 0,0000        | 0,0000        | 0,0000       | 0,0472      | 0,0015      | 0,0015      | 0,0000      | 0,0015      | 0,0000       | 0,0000       |          |

| Nó | HWAJHPNP1111  | HWAJHPNP1222  | HWAJHPNP1333  | HWAJHPPPP1111 | HWAJHPPPP1222 | HWAJHTNNT1111 | HWAJHTNNT1222 | HWAJHTNNT1333 | HWAJHTNNT1444 |               |               |               |               |               |
|----|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| 8  | 0,0016        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0016        | 0,0129        | 0,0064        | 0,0032        |               |               |               |               |               |
| 18 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0013        | 0,0013        | 0,0025        | 0,0025        |               |               |               |               |               |
| 19 | 0,0017        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0017        | 0,0000        |               |               |               |               |               |
| 10 | 0,0024        | 0,0024        | 0,0024        | 0,0000        | 0,0000        | 0,0024        | 0,0096        | 0,0072        | 0,0144        |               |               |               |               |               |
| 44 | 0,0044        | 0,0000        | 0,0015        | 0,0015        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0015        |               |               |               |               |               |
| 45 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0029        |               |               |               |               |               |
| 46 | 0,0070        | 0,0023        | 0,0046        | 0,0035        | 0,0012        | 0,0058        | 0,0023        | 0,0000        | 0,0023        |               |               |               |               |               |
| 47 | 0,0095        | 0,0068        | 0,0014        | 0,0000        | 0,0054        | 0,0027        | 0,0014        | 0,0014        | 0,0014        |               |               |               |               |               |
| 6  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |               |               |               |
| 7  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0015        | 0,0015        | 0,0000        | 0,0044        |               |               |               |               |               |
| Nó | HWHHNNNN1111  | HWHHNNPP1111  | HWHHPPPP1111  | HWHHPPPP1121  | HWHHPPPP1211  | HWHHPPPP1411  | HWHHNN11      | HWHHNN12      | HWHHNN13      | HWHHNN14      | HWHHP11       | HWHHT11       |               |               |
| 8  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0209        | 0,0048        | 0,0016        | 0,0000        | 0,0000        | 0,0032        |               |               |
| 18 | 0,0089        | 0,0013        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,2982        | 0,0089        | 0,0051        | 0,0000        | 0,0076        | 0,0013        |               |               |
| 19 | 0,0017        | 0,0000        | 0,0000        | 0,0017        | 0,0000        | 0,0000        | 0,2547        | 0,0051        | 0,0000        | 0,0051        | 0,0034        | 0,0034        |               |               |
| 10 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0264        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 44 | 0,0000        | 0,0015        | 0,0058        | 0,0029        | 0,0000        | 0,0029        | 0,1385        | 0,0044        | 0,0000        | 0,0029        | 0,0000        | 0,0000        |               |               |
| 45 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0029        | 0,0000        | 0,1057        | 0,0000        | 0,0000        | 0,0000        | 0,0029        | 0,0000        |               |               |
| 46 | 0,0000        | 0,0023        | 0,0046        | 0,0035        | 0,0023        | 0,0023        | 0,1520        | 0,0012        | 0,0000        | 0,0000        | 0,0046        | 0,0000        |               |               |
| 47 | 0,0000        | 0,0027        | 0,0095        | 0,0054        | 0,0041        | 0,0000        | 0,0814        | 0,0014        | 0,0014        | 0,0000        | 0,0041        | 0,0000        |               |               |
| 6  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0203        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 7  | 0,0015        | 0,0015        | 0,0000        | 0,0000        | 0,0015        | 0,0015        | 0,0738        | 0,0015        | 0,0000        | 0,0000        | 0,0015        | 0,0000        |               |               |
| Nó | HWHHP11       | HWHPP11       | HWHPP12       | HWHPP13       | HWHPP14       | HWHPT11       | HWHPT12       | HWHPT13       | HWHPT14       | HWHSHNNNN1111 | HWHSHNNNT1111 | HWHSHPPNN1111 | HWHSHPPPP1111 | HWHSHPPPP1211 |
| 8  | 0,0000        | 0,0048        | 0,0048        | 0,0000        | 0,0000        | 0,0000        | 0,0016        | 0,0032        | 0,0032        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |
| 18 | 0,0000        | 0,0216        | 0,0254        | 0,0089        | 0,0127        | 0,0000        | 0,0013        | 0,0013        | 0,0038        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |
| 19 | 0,0017        | 0,0291        | 0,0188        | 0,0085        | 0,0154        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0017        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |
| 10 | 0,0024        | 0,0361        | 0,0385        | 0,0264        | 0,0240        | 0,0000        | 0,0024        | 0,0048        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |
| 44 | 0,0015        | 0,1414        | 0,0714        | 0,0510        | 0,0685        | 0,0000        | 0,0029        | 0,0000        | 0,0029        | 0,0000        | 0,0000        | 0,0015        | 0,0000        | 0,0015        |
| 45 | 0,0000        | 0,1829        | 0,1343        | 0,0800        | 0,1714        | 0,0029        | 0,0029        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |
| 46 | 0,0012        | 0,1671        | 0,1067        | 0,0557        | 0,0418        | 0,0023        | 0,0023        | 0,0012        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |
| 47 | 0,0014        | 0,2551        | 0,1642        | 0,0828        | 0,0489        | 0,0027        | 0,0014        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0014        |
| 6  | 0,0000        | 0,0000        | 0,0034        | 0,0000        | 0,0068        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0102        | 0,0102        | 0,0000        | 0,0000        | 0,0000        |
| 7  | 0,0000        | 0,0354        | 0,0162        | 0,0074        | 0,0074        | 0,0044        | 0,0059        | 0,0000        | 0,0000        | 0,0870        | 0,0221        | 0,0089        | 0,0103        | 0,0089        |
| Nó | HWHSHPPPP1311 | HWHSHTTNN1111 | HWHSHTTNN1211 | HWHSHTTNN1311 | HWHSHTTNN1411 | HWHSHTTTT1211 | HWHSHTTTT1212 | HWHSHTTTT1411 | HWHHTN11      | HWHHTP11      | HWHHTP12      |               |               |               |
| 8  | 0,0000        | 0,0032        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0032        | 0,0016        |               |               |               |
| 18 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0025        | 0,0013        |               |               |               |
| 19 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0034        | 0,0051        | 0,0017        |               |               |
| 10 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0072        | 0,0048        | 0,0096        |               |               |
| 44 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0029        | 0,0015        | 0,0015        |               |               |
| 45 | 0,0029        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0029        |               |               |
| 46 | 0,0012        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0012        | 0,0058        | 0,0035        |               |               |
| 47 | 0,0014        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0014        | 0,0027        | 0,0000        |               |               |
| 6  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0034        | 0,0000        | 0,0034        | 0,0034        | 0,0034        | 0,0068        | 0,0000        | 0,0000        |               |
| 7  | 0,0030        | 0,0369        | 0,0221        | 0,0148        | 0,0148        | 0,0118        | 0,0059        | 0,0044        | 0,0015        | 0,0000        | 0,0044        | 0,0000        | 0,0044        |               |
| Nó | HWHHTP13      | HWHHTP14      | HWHHTT11      | HWHHTT12      | HWHHTT13      | HWHHTT14      | HWHHNNNN1111  | HWHHNNNN1212  | HWHHPPPP1111  | HWHHPPPP1212  | HWHHHTTTT1111 |               |               |               |
| 8  | 0,0016        | 0,0016        | 0,2010        | 0,2315        | 0,1592        | 0,2572        | 0,0048        | 0,0000        | 0,0000        | 0,0000        | 0,0032        |               |               |               |
| 18 | 0,0000        | 0,0000        | 0,1383        | 0,1180        | 0,0774        | 0,0926        | 0,0939        | 0,0038        | 0,0025        | 0,0000        | 0,0013        |               |               |               |
| 19 | 0,0000        | 0,0000        | 0,1026        | 0,0650        | 0,0530        | 0,0667        | 0,0496        | 0,0000        | 0,0017        | 0,0017        | 0,0000        |               |               |               |
| 10 | 0,0048        | 0,0000        | 0,1659        | 0,1995        | 0,1346        | 0,2163        | 0,0048        | 0,0000        | 0,0048        | 0,0000        | 0,0048        |               |               |               |
| 44 | 0,0000        | 0,0029        | 0,0379        | 0,0408        | 0,0408        | 0,0962        | 0,0364        | 0,0044        | 0,0102        | 0,0044        | 0,0000        |               |               |               |
| 45 | 0,0057        | 0,0029        | 0,0257        | 0,0086        | 0,0143        | 0,0343        | 0,0257        | 0,0000        | 0,0371        | 0,0114        | 0,0000        |               |               |               |
| 46 | 0,0012        | 0,0000        | 0,0545        | 0,0685        | 0,0290        | 0,0093        | 0,0348        | 0,0000        | 0,0220        | 0,0012        | 0,0000        |               |               |               |
| 47 | 0,0000        | 0,0000        | 0,0204        | 0,0258        | 0,0081        | 0,0041        | 0,0244        | 0,0000        | 0,0719        | 0,0068        | 0,0000        |               |               |               |
| 6  | 0,0000        | 0,0000        | 0,0000        | 0,0034        | 0,0000        | 0,0034        | 0,0034        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |               |
| 7  | 0,0000        | 0,0015        | 0,0443        | 0,0472        | 0,0280        | 0,0251        | 0,0251        | 0,0000        | 0,0015        | 0,0000        | 0,0000        |               |               |               |
| Nó | HWHSHNN1111   | HWHSHNTT1111  | HWHSHPPP1111  | HWHSHPPP1321  | HWHSHTTNN1211 | HWHSHTTNN1221 | HWHSHNTT1111  | HWHSHTTN1111  | HWHSHTTN1211  | HWHSHTTN1411  | HWHSHTTTT1111 | HWHSHTTTT1211 |               |               |
| 8  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0016        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 18 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 19 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 10 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 44 | 0,0000        | 0,0000        | 0,0015        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 45 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 46 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 47 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        |               |               |
| 6  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,4339        | 0,4339        | 0,0000        | 0,0000        | 0,0000        | 0,0034        | 0,0000        | 0,0000        |               |               |
| 7  | 0,0089        | 0,0059        | 0,0133        | 0,0059        | 0,0000        | 0,0000        | 0,0089        | 0,0118        | 0,0044        | 0,0044        | 0,0059        | 0,0074        |               |               |

| Nó | H01SHTTT1221 | H01SHTTT1441 | H0000HNNN1111 | H0000HPPP1111 | H0000HPPP1221 | H0000HPPP1341 | H0000HPPP1441 | H0000HTNT11111 |
|----|--------------|--------------|---------------|---------------|---------------|---------------|---------------|----------------|
| 8  | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000         |
| 18 | 0,0000       | 0,0000       | 0,0025        | 0,0000        | 0,0000        | 0,0013        | 0,0000        | 0,0000         |
| 19 | 0,0000       | 0,0000       | 0,0034        | 0,0000        | 0,0000        | 0,0017        | 0,0000        | 0,0000         |
| 10 | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000         |
| 44 | 0,0000       | 0,0000       | 0,0000        | 0,0015        | 0,0029        | 0,0000        | 0,0058        | 0,0000         |
| 45 | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000        | 0,0029        | 0,0143        | 0,0000         |
| 46 | 0,0000       | 0,0000       | 0,0000        | 0,0035        | 0,0012        | 0,0012        | 0,0035        | 0,0000         |
| 47 | 0,0000       | 0,0000       | 0,0000        | 0,0054        | 0,0027        | 0,0014        | 0,0000        | 0,0000         |
| 6  | 0,0000       | 0,0102       | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000         |
| 7  | 0,0089       | 0,0030       | 0,0000        | 0,0000        | 0,0015        | 0,0000        | 0,0000        | 0,0472         |

| Nó | H0000HTTT1441 | H00000HPPP11111 | H00000HPPP11311 | H00000HPPP1441 |
|----|---------------|-----------------|-----------------|----------------|
| 8  | 0,0032        | 0,0000          | 0,0000          | 0,0000         |
| 18 | 0,0000        | 0,0000          | 0,0000          | 0,0000         |
| 19 | 0,0017        | 0,0000          | 0,0000          | 0,0000         |
| 10 | 0,0000        | 0,0000          | 0,0000          | 0,0000         |
| 44 | 0,0029        | 0,0000          | 0,0000          | 0,0015         |
| 45 | 0,0000        | 0,0029          | 0,0000          | 0,0114         |
| 46 | 0,0000        | 0,0023          | 0,0023          | 0,0000         |
| 47 | 0,0000        | 0,0014          | 0,0054          | 0,0000         |
| 6  | 0,0000        | 0,0000          | 0,0000          | 0,0000         |
| 7  | 0,0000        | 0,0015          | 0,0000          | 0,0000         |

# ANEXO C

Probabilidade de ocorrência de todos os padrões de viagens – proveniente do relatório gerado pelo S-PLUS 6.1 (nós terminais)

## 3) AMOSTRA DE TRABALHADORES NO SETOR DE SERVIÇOS

| Nó | HAAHPPP1111   | HAHPPP1121     | HAAHPPP1211    | HAAHTNT1331    | HAAHTTT1441   | HAHAHNNNN1111 | HAHAHPPP1111   | HAHNN111       | HAHPP111       | HAHPP121     | HAHPP131      | HAHPP141   | HAHTT111 |
|----|---------------|----------------|----------------|----------------|---------------|---------------|----------------|----------------|----------------|--------------|---------------|------------|----------|
| 8  | 0,0000        | 0,0000         | 0,0000         | 0,0010         | 0,0010        | 0,0000        | 0,0000         | 0,0059         | 0,0020         | 0,0010       | 0,0000        | 0,0000     | 0,0148   |
| 18 | 0,0000        | 0,0000         | 0,0000         | 0,0021         | 0,0000        | 0,0021        | 0,0010         | 0,0239         | 0,0062         | 0,0010       | 0,0000        | 0,0000     | 0,0291   |
| 19 | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0014        | 0,0043        | 0,0014         | 0,0213         | 0,0028         | 0,0099       | 0,0028        | 0,0028     | 0,0341   |
| 5  | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0064        | 0,0000         | 0,0064         | 0,0000         | 0,0032       | 0,0000        | 0,0000     | 0,0064   |
| 12 | 0,0012        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0000        | 0,0023         | 0,0035         | 0,0129         | 0,0047       | 0,0000        | 0,0023     | 0,0082   |
| 52 | 0,0012        | 0,0024         | 0,0024         | 0,0000         | 0,0024        | 0,0000        | 0,0012         | 0,0036         | 0,0192         | 0,0156       | 0,0120        | 0,0132     | 0,0024   |
| 53 | 0,0018        | 0,0018         | 0,0000         | 0,0000         | 0,0036        | 0,0000        | 0,0018         | 0,0144         | 0,0323         | 0,0090       | 0,0054        | 0,0054     | 0,0072   |
| 54 | 0,0048        | 0,0008         | 0,0008         | 0,0024         | 0,0000        | 0,0000        | 0,0032         | 0,0104         | 0,0296         | 0,0104       | 0,0056        | 0,0040     | 0,0112   |
| 55 | 0,0041        | 0,0017         | 0,0033         | 0,0000         | 0,0000        | 0,0008        | 0,0008         | 0,0083         | 0,0248         | 0,0108       | 0,0041        | 0,0025     | 0,0066   |
| Nó | HAHTT121      | HAHTT131       | HAHTT141       | HAHWHPPPP1111  | HAHWHPPPP1121 | HAHWHTTNN1111 | HAHWHPPPP1111  | HSWAHPPP1111   | HSNN111        | HSPP111      | HSPP121       | HSNNNN1111 |          |
| 8  | 0,0148        | 0,0059         | 0,0089         | 0,0000         | 0,0000        | 0,0010        | 0,0000         | 0,0000         | 0,0059         | 0,0000       | 0,0000        | 0,0000     |          |
| 18 | 0,0187        | 0,0167         | 0,0219         | 0,0000         | 0,0000        | 0,0031        | 0,0000         | 0,0000         | 0,0062         | 0,0010       | 0,0000        | 0,0031     |          |
| 19 | 0,0142        | 0,0043         | 0,0057         | 0,0000         | 0,0000        | 0,0014        | 0,0000         | 0,0000         | 0,0028         | 0,0014       | 0,0014        | 0,0057     |          |
| 5  | 0,0000        | 0,0000         | 0,0064         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0573         | 0,0032       | 0,0000        | 0,0000     |          |
| 12 | 0,0094        | 0,0094         | 0,0023         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0012         | 0,0023         | 0,0023       | 0,0000        | 0,0000     |          |
| 52 | 0,0036        | 0,0012         | 0,0048         | 0,0012         | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0036         | 0,0036       | 0,0000        | 0,0012     |          |
| 53 | 0,0018        | 0,0000         | 0,0126         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0215         | 0,0144       | 0,0018        | 0,0054     |          |
| 54 | 0,0120        | 0,0032         | 0,0008         | 0,0016         | 0,0016        | 0,0008        | 0,0024         | 0,0008         | 0,0048         | 0,0096       | 0,0024        | 0,0024     |          |
| 55 | 0,0017        | 0,0033         | 0,0017         | 0,0025         | 0,0041        | 0,0000        | 0,0033         | 0,0025         | 0,0025         | 0,0149       | 0,0025        | 0,0008     |          |
| Nó | HSWSHPPP1111  | HSHTT111       | HSHTT121       | HSHTT131       | HSHTT141      | HSWNNNN1111   | HSWNNNTT1111   | HSWHPPPP1111   | HSWHPPPP1121   | HSWHTTNN1111 | HSWHPPPP1111  |            |          |
| 8  | 0,0000        | 0,0010         | 0,0030         | 0,0010         | 0,0010        | 0,0000        | 0,0020         | 0,0000         | 0,0000         | 0,0000       | 0,0000        | 0,0000     |          |
| 18 | 0,0000        | 0,0021         | 0,0021         | 0,0010         | 0,0010        | 0,0021        | 0,0000         | 0,0000         | 0,0000         | 0,0000       | 0,0000        | 0,0000     |          |
| 19 | 0,0000        | 0,0000         | 0,0014         | 0,0000         | 0,0000        | 0,0014        | 0,0014         | 0,0000         | 0,0000         | 0,0000       | 0,0000        | 0,0000     |          |
| 5  | 0,0000        | 0,0159         | 0,0064         | 0,0096         | 0,0064        | 0,0096        | 0,0032         | 0,0000         | 0,0000         | 0,0000       | 0,0096        | 0,0000     |          |
| 12 | 0,0000        | 0,0047         | 0,0012         | 0,0000         | 0,0012        | 0,0000        | 0,0012         | 0,0000         | 0,0000         | 0,0000       | 0,0000        | 0,0012     |          |
| 52 | 0,0036        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0036         | 0,0000         | 0,0000       | 0,0000        | 0,0012     |          |
| 53 | 0,0090        | 0,0090         | 0,0036         | 0,0000         | 0,0000        | 0,0036        | 0,0018         | 0,0000         | 0,0018         | 0,0000       | 0,0000        | 0,0072     |          |
| 54 | 0,0056        | 0,0016         | 0,0016         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0024         | 0,0016         | 0,0016       | 0,0000        | 0,0056     |          |
| 55 | 0,0066        | 0,0000         | 0,0000         | 0,0000         | 0,0008        | 0,0000        | 0,0000         | 0,0050         | 0,0033         | 0,0000       | 0,0000        | 0,0050     |          |
| Nó | HSWHPPPP1121  | HSWHPPPP1131   | HSWHPPPP1221   | HSWSHPPP1111   | HSWAHPPP1111  | HSWAHPPP1121  | HSWAHPPP1211   | HSWAHPPP1221   | HSWAHTTT1111   | HSWAHTTT1441 | HSWAHNNNN1111 |            |          |
| 8  | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0010        | 0,0000        | 0,0000         | 0,0000         | 0,0010         | 0,0030       | 0,0000        |            |          |
| 18 | 0,0000        | 0,0010         | 0,0000         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0021         | 0,0021       | 0,0010        |            |          |
| 19 | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0028        | 0,0000         | 0,0000         | 0,0000         | 0,0000       | 0,0028        |            |          |
| 5  | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000       | 0,0000        |            |          |
| 12 | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0012        | 0,0000        | 0,0012         | 0,0000         | 0,0012         | 0,0047       | 0,0012        |            |          |
| 52 | 0,0012        | 0,0012         | 0,0000         | 0,0000         | 0,0036        | 0,0000        | 0,0012         | 0,0012         | 0,0000         | 0,0012       | 0,0000        |            |          |
| 53 | 0,0018        | 0,0000         | 0,0000         | 0,0036         | 0,0054        | 0,0000        | 0,0000         | 0,0018         | 0,0018         | 0,0000       | 0,0000        |            |          |
| 54 | 0,0008        | 0,0016         | 0,0008         | 0,0080         | 0,0016        | 0,0032        | 0,0056         | 0,0008         | 0,0008         | 0,0000       | 0,0024        |            |          |
| 55 | 0,0075        | 0,0033         | 0,0017         | 0,0058         | 0,0075        | 0,0017        | 0,0041         | 0,0025         | 0,0008         | 0,0000       | 0,0008        |            |          |
| Nó | HSWAHPNPP1111 | HSWAHPNPP12221 | HSWAHPNPP13331 | HSWAHPNPP14441 | HSWAHPPPP1111 | HSWAHTTNT1111 | HSWAHTTNT12221 | HSWAHTTNT13331 | HSWAHTTNT14441 |              |               |            |          |
| 8  | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0049        | 0,0069         | 0,0039         | 0,0039         |              |               |            |          |
| 18 | 0,0000        | 0,0010         | 0,0000         | 0,0000         | 0,0010        | 0,0021        | 0,0010         | 0,0010         | 0,0031         |              |               |            |          |
| 19 | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0028        | 0,0028         | 0,0057         | 0,0000         |              |               |            |          |
| 5  | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0064        | 0,0000         | 0,0000         | 0,0000         |              |               |            |          |
| 12 | 0,0012        | 0,0012         | 0,0000         | 0,0000         | 0,0000        | 0,0070        | 0,0106         | 0,0094         | 0,0059         |              |               |            |          |
| 52 | 0,0000        | 0,0012         | 0,0000         | 0,0084         | 0,0048        | 0,0000        | 0,0000         | 0,0000         | 0,0084         |              |               |            |          |
| 53 | 0,0036        | 0,0000         | 0,0000         | 0,0000         | 0,0000        | 0,0000        | 0,0000         | 0,0018         | 0,0000         |              |               |            |          |
| 54 | 0,0024        | 0,0040         | 0,0032         | 0,0000         | 0,0040        | 0,0040        | 0,0064         | 0,0016         | 0,0000         |              |               |            |          |
| 55 | 0,0091        | 0,0099         | 0,0017         | 0,0000         | 0,0025        | 0,0033        | 0,0033         | 0,0017         | 0,0008         |              |               |            |          |

| Nº | H00H0HNNNN11111  | H00H0HNNPP11111  | H00H0HPPPP11111  | H00H0HPPPP11121  | H00H0HPPPP12111  | H00H0HPPPP13121  | H00H0HPPPP11111  | H00H0HNN111      | H00H0HNN121      | H00H0HNP111      | H00H0HNT111      |                  |
|----|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|------------------|
| 8  | 0,0000           | 0,0010           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0020           | 0,0364           | 0,0000           | 0,0000           | 0,0030           |
| 18 | 0,0000           | 0,0021           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,2050           | 0,0073           | 0,0021           | 0,0010           |
| 19 | 0,0057           | 0,0028           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,3026           | 0,0028           | 0,0014           | 0,0014           |
| 5  | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0669           | 0,0000           | 0,0032           | 0,0032           |
| 12 | 0,0000           | 0,0012           | 0,0000           | 0,0000           | 0,0012           | 0,0035           | 0,0000           | 0,0035           | 0,0387           | 0,0023           | 0,0012           | 0,0012           |
| 52 | 0,0012           | 0,0000           | 0,0048           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0527           | 0,0036           | 0,0012           | 0,0048           |
| 53 | 0,0018           | 0,0018           | 0,0072           | 0,0054           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,1454           | 0,0036           | 0,0036           | 0,0000           |
| 54 | 0,0000           | 0,0048           | 0,0072           | 0,0008           | 0,0000           | 0,0000           | 0,0008           | 0,0016           | 0,1136           | 0,0008           | 0,0016           | 0,0016           |
| 55 | 0,0000           | 0,0008           | 0,0099           | 0,0050           | 0,0066           | 0,0000           | 0,0050           | 0,0000           | 0,0422           | 0,0000           | 0,0008           | 0,0008           |
| Nº | H00H0PN111       | H00H0PP111       | H00H0PP121       | H00H0PP131       | H00H0PP141       | H00H0PT111       | H00H0PT121       | H00H0PT131       | H00H0SHNNNN11111 | H00H0SHNNPP11111 | H00H0SHNNNT11121 | H00H0SHPPPP11111 |
| 8  | 0,0000           | 0,0039           | 0,0020           | 0,0030           | 0,0010           | 0,0020           | 0,0020           | 0,0010           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 18 | 0,0000           | 0,0406           | 0,0167           | 0,0094           | 0,0239           | 0,0010           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 19 | 0,0000           | 0,0270           | 0,0227           | 0,0099           | 0,0043           | 0,0014           | 0,0000           | 0,0000           | 0,0043           | 0,0000           | 0,0000           | 0,0000           |
| 5  | 0,0000           | 0,0064           | 0,0032           | 0,0032           | 0,0032           | 0,0000           | 0,0000           | 0,0000           | 0,1115           | 0,0032           | 0,0096           | 0,0000           |
| 12 | 0,0012           | 0,0410           | 0,0235           | 0,0270           | 0,0129           | 0,0094           | 0,0035           | 0,0047           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |
| 52 | 0,0000           | 0,1653           | 0,0719           | 0,0575           | 0,1641           | 0,0024           | 0,0012           | 0,0012           | 0,0000           | 0,0000           | 0,0000           | 0,0024           |
| 53 | 0,0018           | 0,1652           | 0,0431           | 0,0251           | 0,0467           | 0,0000           | 0,0054           | 0,0000           | 0,0000           | 0,0000           | 0,0018           | 0,0018           |
| 54 | 0,0024           | 0,1464           | 0,1080           | 0,0312           | 0,0192           | 0,0104           | 0,0072           | 0,0008           | 0,0016           | 0,0000           | 0,0000           | 0,0000           |
| 55 | 0,0058           | 0,2351           | 0,1606           | 0,0753           | 0,0455           | 0,0025           | 0,0033           | 0,0017           | 0,0000           | 0,0000           | 0,0000           | 0,0008           |
| Nº | H00H0SHPPPP12111 | H00H0SHTTNN11111 | H00H0SHTTNN12111 | H00H0SHTTNN13111 | H00H0SHTTNN14111 | H00H0SHTTTT11111 | H00H0SHTTTT11121 | H00H0SHTTTT12111 | H00H0SHTTTT14111 |                  |                  |                  |
| 8  | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0010           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| 18 | 0,0000           | 0,0010           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| 19 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| 5  | 0,0000           | 0,0223           | 0,0319           | 0,0255           | 0,0255           | 0,0127           | 0,0064           | 0,0096           | 0,0159           |                  |                  |                  |
| 12 | 0,0000           | 0,0012           | 0,0000           | 0,0000           | 0,0012           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| 52 | 0,0012           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| 53 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| 54 | 0,0016           | 0,0008           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| 55 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |                  |                  |
| Nº | H00H0HTN111      | H00H0HTP111      | H00H0HTP121      | H00H0HTP131      | H00H0HTP141      | H00H0HTT111      | H00H0HTT121      | H00H0HTT131      | H00H0HTT141      | H00H0HNNNN11111  | H00H0HPPPP11111  | H00H0HPPPP11121  |
| 8  | 0,0049           | 0,0020           | 0,0020           | 0,0000           | 0,0020           | 0,2173           | 0,2153           | 0,1632           | 0,2330           | 0,0039           | 0,0010           | 0,0000           |
| 18 | 0,0021           | 0,0031           | 0,0010           | 0,0021           | 0,0000           | 0,1124           | 0,1145           | 0,0833           | 0,1561           | 0,0406           | 0,0021           | 0,0000           |
| 19 | 0,0057           | 0,0085           | 0,0043           | 0,0000           | 0,0000           | 0,1648           | 0,1463           | 0,0696           | 0,0114           | 0,0469           | 0,0043           | 0,0014           |
| 5  | 0,0000           | 0,0000           | 0,0032           | 0,0000           | 0,0000           | 0,0828           | 0,0733           | 0,0350           | 0,0573           | 0,0159           | 0,0000           | 0,0000           |
| 12 | 0,0059           | 0,0059           | 0,0012           | 0,0012           | 0,0023           | 0,1395           | 0,1887           | 0,1313           | 0,2075           | 0,0000           | 0,0012           | 0,0000           |
| 52 | 0,0000           | 0,0036           | 0,0012           | 0,0000           | 0,0024           | 0,0719           | 0,0383           | 0,0455           | 0,0934           | 0,0132           | 0,0299           | 0,0000           |
| 53 | 0,0018           | 0,0108           | 0,0018           | 0,0036           | 0,0018           | 0,0970           | 0,0485           | 0,0449           | 0,0844           | 0,0162           | 0,0180           | 0,0018           |
| 54 | 0,0000           | 0,0072           | 0,0016           | 0,0032           | 0,0016           | 0,0792           | 0,0944           | 0,0496           | 0,0232           | 0,0200           | 0,0272           | 0,0016           |
| 55 | 0,0017           | 0,0025           | 0,0017           | 0,0000           | 0,0000           | 0,0315           | 0,0381           | 0,0240           | 0,0041           | 0,0166           | 0,0431           | 0,0025           |
| Nº | H00H0HPPPP12121  | H00H0HHTTTT11111 | H00H0SHNN1111    | H00H0SHPP1111    | H00H0SHPP1211    | H00H0SHPP1221    | H00H0SHTNT1111   | H00H0SHTNT1221   | H00H0SHTNT1331   | H00H0SHTTT1111   | H00H0SHTTT12121  |                  |
| 8  | 0,0000           | 0,0010           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |
| 18 | 0,0000           | 0,0042           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |
| 19 | 0,0000           | 0,0028           | 0,0014           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |
| 5  | 0,0096           | 0,0000           | 0,0096           | 0,0000           | 0,0000           | 0,0000           | 0,0096           | 0,0096           | 0,0096           | 0,0319           | 0,0096           |                  |
| 12 | 0,0012           | 0,0012           | 0,0000           | 0,0023           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |
| 52 | 0,0036           | 0,0024           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0012           | 0,0000           | 0,0000           |                  |
| 53 | 0,0000           | 0,0000           | 0,0000           | 0,0018           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0018           | 0,0000           |                  |
| 54 | 0,0040           | 0,0016           | 0,0008           | 0,0024           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0008           | 0,0000           | 0,0000           |                  |
| 55 | 0,0066           | 0,0008           | 0,0000           | 0,0033           | 0,0008           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |
| Nº | H00H0SHTTT1211   | H00H0SHTTT1221   | H00H0SHTTT1231   | H00H0SHTTT1311   | H00H0SHTTT1331   | H00H0SHTTT1341   | H00H0SHTTT1411   | H00H0SHTTT1441   | H00H0HPPPP1111   | H00H0HPPPP1121   | H00H0HPPPP1211   |                  |
| 8  | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0010           | 0,0000           | 0,0000           | 0,0000           | 0,0000           |                  |
| 18 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0010           | 0,0000           | 0,0000           |                  |
| 19 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0014           | 0,0000           |                  |
| 5  | 0,0255           | 0,0191           | 0,0032           | 0,0159           | 0,0127           | 0,0127           | 0,0064           | 0,0223           | 0,0000           | 0,0000           | 0,0000           |                  |
| 12 | 0,0000           | 0,0012           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0023           | 0,0012           | 0,0000           |                  |
| 52 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0072           | 0,0024           | 0,0036           |                  |
| 53 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0054           | 0,0000           | 0,0036           |                  |
| 54 | 0,0008           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0072           | 0,0016           | 0,0032           |                  |
| 55 | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0000           | 0,0058           | 0,0017           | 0,0025           |                  |

| Nó | H0000HPPP1221 | H0000HPPP1431 | H0000HTTT1111 | H0000HTTT1441 | H00000HPPPP11111 | H00000HPPPP12221 |
|----|---------------|---------------|---------------|---------------|------------------|------------------|
| 8  | 0,0000        | 0,0000        | 0,0010        | 0,0000        | 0,0000           | 0,0000           |
| 18 | 0,0000        | 0,0000        | 0,0010        | 0,0062        | 0,0000           | 0,0000           |
| 19 | 0,0000        | 0,0000        | 0,0014        | 0,0028        | 0,0014           | 0,0000           |
| 5  | 0,0000        | 0,0000        | 0,0000        | 0,0032        | 0,0000           | 0,0000           |
| 12 | 0,0035        | 0,0023        | 0,0023        | 0,0023        | 0,0012           | 0,0000           |
| 52 | 0,0024        | 0,0036        | 0,0000        | 0,0012        | 0,0024           | 0,0000           |
| 53 | 0,0000        | 0,0000        | 0,0000        | 0,0018        | 0,0000           | 0,0000           |
| 54 | 0,0040        | 0,0016        | 0,0016        | 0,0000        | 0,0008           | 0,0024           |
| 55 | 0,0041        | 0,0008        | 0,0008        | 0,0008        | 0,0058           | 0,0025           |

# ANEXO E

## Probabilidade de ocorrência de todos os padrões de viagens – proveniente do relatório gerado pelo S-PLUS 6.1 (nós terminais)

### 5) AMOSTRA DE ESTUDANTES NO 2º GRAU

|    |               |               |              |               |              |              |              |               |              |             |             |          |
|----|---------------|---------------|--------------|---------------|--------------|--------------|--------------|---------------|--------------|-------------|-------------|----------|
| Nó | HAANN1111     | HAAPPP1111    | HAAPPP1211   | HAATTT1111    | HAATTT1211   | HAHAHP1111   | HAHAHT1111   | HAHNN111      | HAHPP111     | HAHPP121    | HAHPP131    | HAHPP141 |
| 8  | 0,0027        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0108        | 0,0027       | 0,0027      | 0,0000      | 0,0081   |
| 9  | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0029        | 0,0029       | 0,0000      | 0,0029      | 0,0029   |
| 10 | 0,0028        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0222        | 0,0139       | 0,0055      | 0,0000      | 0,0000   |
| 11 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0044       | 0,0000       | 0,0089        | 0,0178       | 0,0044      | 0,0044      | 0,0000   |
| 6  | 0,0000        | 0,0000        | 0,0000       | 0,0066        | 0,0066       | 0,0000       | 0,0000       | 0,0033        | 0,0066       | 0,0033      | 0,0000      | 0,0000   |
| 14 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0105        | 0,0000       | 0,0000      | 0,0000      | 0,0000   |
| 30 | 0,0047        | 0,0047        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0140       | 0,0000      | 0,0000      | 0,0000   |
| 31 | 0,0000        | 0,0000        | 0,0046       | 0,0000        | 0,0000       | 0,0046       | 0,0000       | 0,0092        | 0,0184       | 0,0046      | 0,0000      | 0,0000   |
| Nó | HAHSHNN1111   | HAHSHPPN14111 | HAHSHPPP1111 | HAHSHPPP12121 | HAHSHTTN1111 | HAHSHTTN1211 | HAHSHTTN1311 | HAHSHTTN1411  | HAHSHTTT1121 |             |             |          |
| 8  | 0,0000        | 0,0027        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0054       | 0,0081        | 0,0027       |             |             |          |
| 9  | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0029       | 0,0058       | 0,0000        | 0,0058       |             |             |          |
| 10 | 0,0111        | 0,0000        | 0,0000       | 0,0000        | 0,0055       | 0,0000       | 0,0000       | 0,0000        | 0,0028       |             |             |          |
| 11 | 0,0000        | 0,0000        | 0,0000       | 0,0044        | 0,0133       | 0,0044       | 0,0000       | 0,0000        | 0,0000       |             |             |          |
| 6  | 0,0000        | 0,0033        | 0,0000       | 0,0000        | 0,0000       | 0,0033       | 0,0000       | 0,0000        | 0,0033       |             |             |          |
| 14 | 0,0105        | 0,0000        | 0,0105       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       |             |             |          |
| 30 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       |             |             |          |
| 31 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       |             |             |          |
| Nó | HAHSHTTT12121 | HAHSHTTT14111 | HAHTP111     | HAHTT111      | HAHTT121     | HAHTT131     | HAHTT141     | HASAHTTT12221 | HASHPPP1111  | HASHTTT1111 | HASHTTT1121 |          |
| 8  | 0,0027        | 0,0000        | 0,0000       | 0,0054        | 0,0081       | 0,0054       | 0,0217       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 9  | 0,0000        | 0,0029        | 0,0000       | 0,0029        | 0,0088       | 0,0029       | 0,0088       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 10 | 0,0000        | 0,0000        | 0,0000       | 0,0166        | 0,0055       | 0,0028       | 0,0028       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 11 | 0,0000        | 0,0044        | 0,0000       | 0,0133        | 0,0044       | 0,0044       | 0,0044       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 6  | 0,0000        | 0,0000        | 0,0033       | 0,0066        | 0,0000       | 0,0000       | 0,0000       | 0,0066        | 0,0000       | 0,0066      | 0,0066      |          |
| 14 | 0,0000        | 0,0000        | 0,0000       | 0,0211        | 0,0000       | 0,0105       | 0,0000       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 30 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 31 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0046       | 0,0000       | 0,0000       | 0,0000        | 0,0092       | 0,0000      | 0,0000      |          |
| Nó | HSAHPPP14111  | HSAHNN1111    | HSAHNTT1111  | HSAHPNT1111   | HSAHPPP1111  | HSAHPPP1211  | HSAHPPP1221  | HSAHTN1111    | HSAHTNT1111  | HSAHTNT1221 | HSAHTPP1211 |          |
| 8  | 0,0000        | 0,0027        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0027        | 0,0000       | 0,0000      | 0,0000      |          |
| 9  | 0,0000        | 0,0029        | 0,0000       | 0,0000        | 0,0029       | 0,0000       | 0,0000       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 10 | 0,0000        | 0,0111        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0028        | 0,0000       | 0,0028      | 0,0000      |          |
| 11 | 0,0000        | 0,0089        | 0,0000       | 0,0000        | 0,0000       | 0,0044       | 0,0089       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 6  | 0,0000        | 0,0000        | 0,0033       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0033        | 0,0000       | 0,0033      | 0,0000      |          |
| 14 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       | 0,0105      | 0,0211      |          |
| 30 | 0,0047        | 0,0000        | 0,0000       | 0,0033        | 0,0000       | 0,0000       | 0,0047       | 0,0000        | 0,0000       | 0,0000      | 0,0000      |          |
| 31 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0046       | 0,0000      | 0,0000      |          |
| Nó | HSAHTPP1221   | HSAHTTP1111   | HSAHTTT1231  | HSAHNN1111    | HSAHNNPP1111 | HSAHNNNT1111 | HSAHNNTT1111 | HSAHNNTT1121  | HSAHNNPP1111 | HSAHPPN1111 |             |          |
| 8  | 0,0000        | 0,0000        | 0,0000       | 0,0054        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       | 0,0000      |             |          |
| 9  | 0,0000        | 0,0000        | 0,0029       | 0,0058        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0029       | 0,0029      |             |          |
| 10 | 0,0000        | 0,0000        | 0,0000       | 0,0055        | 0,0000       | 0,0000       | 0,0000       | 0,0028        | 0,0000       | 0,0000      |             |          |
| 11 | 0,0000        | 0,0000        | 0,0000       | 0,0044        | 0,0044       | 0,0000       | 0,0133       | 0,0000        | 0,0000       | 0,0000      |             |          |
| 6  | 0,0000        | 0,0000        | 0,0000       | 0,0100        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       | 0,0033      |             |          |
| 14 | 0,0211        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       | 0,0000      |             |          |
| 30 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0047        | 0,0047       | 0,0047      |             |          |
| 31 | 0,0000        | 0,0046        | 0,0000       | 0,0092        | 0,0000       | 0,0092       | 0,0046       | 0,0000        | 0,0046       | 0,0000      |             |          |
| Nó | HSAHPPP1111   | HSAHPPP1121   | HSAHPPP1211  | HSAHPPP12121  | HSAHPTN1111  | HSAHPTP1211  | HSAHTN1111   | HSAHTPP1111   | HSAHTPP14111 |             |             |          |
| 8  | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0027       | 0,0000        | 0,0000       |             |             |          |
| 9  | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       |             |             |          |
| 10 | 0,0000        | 0,0000        | 0,0000       | 0,0028        | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000       |             |             |          |
| 11 | 0,0044        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0044       | 0,0089       | 0,0000        | 0,0000       |             |             |          |
| 6  | 0,0066        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0033        | 0,0033       |             |             |          |
| 14 | 0,0000        | 0,0000        | 0,0000       | 0,0000        | 0,0000       | 0,0000       | 0,0000       | 0,0105        | 0,0105       |             |             |          |
| 30 | 0,0140        | 0,0093        | 0,0000       | 0,0000        | 0,0047       | 0,0000       | 0,0093       | 0,0000        | 0,0000       |             |             |          |
| 31 | 0,0138        | 0,0046        | 0,0046       | 0,0000        | 0,0092       | 0,0000       | 0,0046       | 0,0000        | 0,0000       |             |             |          |

| Nó | HSHAHTTTP11111 | HSHAHTTTT11111 | HSHAHTTTT11121 | HSHAHTTTT11131 | HSHAHTTTT12121 | HSHNN111 | HSHNN121 | HSHNN131 | HSHNN141 | HSHNP111 |
|----|----------------|----------------|----------------|----------------|----------------|----------|----------|----------|----------|----------|
| 8  | 0,0000         | 0,0027         | 0,0000         | 0,0000         | 0,0000         | 0,3957   | 0,0407   | 0,0108   | 0,0027   | 0,0000   |
| 9  | 0,0000         | 0,0029         | 0,0000         | 0,0000         | 0,0000         | 0,0029   | 0,5058   | 0,0058   | 0,0029   | 0,0058   |
| 10 | 0,0028         | 0,0000         | 0,0000         | 0,0000         | 0,0055         | 0,0000   | 0,4598   | 0,0000   | 0,0000   | 0,0111   |
| 11 | 0,0000         | 0,0044         | 0,0000         | 0,0000         | 0,0000         | 0,0000   | 0,2800   | 0,0000   | 0,0000   | 0,0089   |
| 6  | 0,0033         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000   | 0,2525   | 0,0000   | 0,0000   | 0,0000   |
| 14 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000   | 0,2000   | 0,0000   | 0,0105   | 0,0000   |
| 30 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000   | 0,0888   | 0,0000   | 0,0000   | 0,0000   |
| 31 | 0,0000         | 0,0000         | 0,0046         | 0,0000         | 0,0046         | 0,1789   | 0,0000   | 0,0000   | 0,0000   | 0,0138   |

| Nó | HSHNT111 | HSHPN111 | HSHPP111 | HSHPP121 | HSHPP131 | HSHPP141 | HSHPT111 | HSHPT121 | HSHPT131 | HSHPT141 | HSHSHNNNN11111 | HSHSHNNPP11111 |
|----|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------------|----------------|
| 8  | 0,0000   | 0,0027   | 0,0217   | 0,0000   | 0,0000   | 0,0027   | 0,0027   | 0,0081   | 0,0054   | 0,0000   | 0,0054         | 0,0027         |
| 9  | 0,0029   | 0,0029   | 0,0292   | 0,0088   | 0,0029   | 0,0000   | 0,0058   | 0,0000   | 0,0000   | 0,0000   | 0,0146         | 0,0000         |
| 10 | 0,0028   | 0,0055   | 0,0360   | 0,0083   | 0,0000   | 0,0000   | 0,0055   | 0,0055   | 0,0000   | 0,0000   | 0,0111         | 0,0000         |
| 11 | 0,0000   | 0,0222   | 0,0622   | 0,0000   | 0,0000   | 0,0000   | 0,0756   | 0,0044   | 0,0000   | 0,0044   | 0,0089         | 0,0044         |
| 6  | 0,0033   | 0,0365   | 0,0897   | 0,0133   | 0,0000   | 0,0000   | 0,0897   | 0,0066   | 0,0100   | 0,0000   | 0,0100         | 0,0033         |
| 14 | 0,0211   | 0,0316   | 0,1684   | 0,0737   | 0,0421   | 0,0316   | 0,0105   | 0,0211   | 0,0000   | 0,0211   | 0,0000         | 0,0000         |
| 30 | 0,0000   | 0,0327   | 0,3084   | 0,0467   | 0,0234   | 0,0093   | 0,0981   | 0,0374   | 0,0140   | 0,0000   | 0,0047         | 0,0047         |
| 31 | 0,0000   | 0,0229   | 0,1972   | 0,0688   | 0,0046   | 0,0000   | 0,0963   | 0,0321   | 0,0000   | 0,0000   | 0,0092         | 0,0046         |

| Nó | HSHSHNPT11111 | HSHSHNNT11111 | HSHSHNNT11121 | HSHSHNNT11131 | HSHSHNNT11141 | HSHSHPPNN11111 | HSHSHPPPP11111 | HSHSHPPPP11121 | HSHSHPPPP12111 | HSHSHPPPP12121 |
|----|---------------|---------------|---------------|---------------|---------------|----------------|----------------|----------------|----------------|----------------|
| 8  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0027         |
| 9  | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |
| 10 | 0,0000        | 0,0000        | 0,0028        | 0,0000        | 0,0000        | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |
| 11 | 0,0000        | 0,0044        | 0,0000        | 0,0044        | 0,0000        | 0,0000         | 0,0000         | 0,0044         | 0,0044         | 0,0000         |
| 6  | 0,0033        | 0,0100        | 0,0000        | 0,0033        | 0,0000        | 0,0000         | 0,0166         | 0,0000         | 0,0000         | 0,0000         |
| 14 | 0,0000        | 0,0000        | 0,0211        | 0,0000        | 0,0000        | 0,0000         | 0,0105         | 0,0000         | 0,0000         | 0,0000         |
| 30 | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0000        | 0,0047         | 0,0234         | 0,0047         | 0,0000         | 0,0000         |
| 31 | 0,0046        | 0,0046        | 0,0000        | 0,0000        | 0,0092        | 0,0184         | 0,0000         | 0,0046         | 0,0000         | 0,0046         |

| Nó | HSHSHPPPT13111 | HSHSHPPPT14111 | HSHSHPPPT11121 | HSHSHPTNN11111 | HSHSHPTTT11111 | HSHSHTTNN11111 | HSHSHTTNN12111 | HSHSHTTNN13111 | HSHSHTTNN14111 |
|----|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| 8  | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0081         | 0,0000         | 0,0054         |
| 9  | 0,0000         | 0,0000         | 0,0000         | 0,0029         | 0,0000         | 0,0000         | 0,0058         | 0,0029         | 0,0000         |
| 10 | 0,0000         | 0,0000         | 0,0000         | 0,0028         | 0,0000         | 0,0028         | 0,0028         | 0,0028         | 0,0028         |
| 11 | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0044         | 0,0000         | 0,0000         | 0,0000         |
| 6  | 0,0000         | 0,0000         | 0,0033         | 0,0000         | 0,0033         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |
| 14 | 0,0000         | 0,0105         | 0,0000         | 0,0000         | 0,0000         | 0,0211         | 0,0000         | 0,0105         | 0,0000         |
| 30 | 0,0047         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         | 0,0000         |
| 31 | 0,0000         | 0,0000         | 0,0046         | 0,0000         | 0,0046         | 0,0046         | 0,0000         | 0,0000         | 0,0000         |

| Nó | HSHSHTTTT11111 | HSHSHTTTT11121 | HSHTN111 | HSHTN131 | HSHTP111 | HSHTP121 | HSHTT111 | HSHTT121 | HSHTT131 | HSHTT141 | HSSHNNNN11111 | HSSHPPPT11111 |
|----|----------------|----------------|----------|----------|----------|----------|----------|----------|----------|----------|---------------|---------------|
| 8  | 0,0000         | 0,0027         | 0,0027   | 0,0000   | 0,0027   | 0,0027   | 0,1518   | 0,1003   | 0,0515   | 0,0596   | 0,0000        | 0,0000        |
| 9  | 0,0000         | 0,0000         | 0,0029   | 0,0000   | 0,0000   | 0,0000   | 0,1608   | 0,1111   | 0,0234   | 0,0058   | 0,0029        | 0,0000        |
| 10 | 0,0028         | 0,0000         | 0,0000   | 0,0028   | 0,0055   | 0,0000   | 0,2078   | 0,0471   | 0,0166   | 0,0166   | 0,0028        | 0,0000        |
| 11 | 0,0089         | 0,0044         | 0,0089   | 0,0044   | 0,0178   | 0,0044   | 0,2178   | 0,0489   | 0,0133   | 0,0133   | 0,0000        | 0,0000        |
| 6  | 0,0000         | 0,0000         | 0,0033   | 0,0000   | 0,0100   | 0,0000   | 0,1993   | 0,0565   | 0,0199   | 0,0365   | 0,0033        | 0,0000        |
| 14 | 0,0105         | 0,0000         | 0,0000   | 0,0000   | 0,0000   | 0,0000   | 0,0421   | 0,0526   | 0,0211   | 0,0211   | 0,0000        | 0,0000        |
| 30 | 0,0000         | 0,0000         | 0,0047   | 0,0000   | 0,0093   | 0,0000   | 0,0935   | 0,0467   | 0,0421   | 0,0047   | 0,0000        | 0,0000        |
| 31 | 0,0000         | 0,0000         | 0,0000   | 0,0000   | 0,0000   | 0,0000   | 0,1193   | 0,0367   | 0,0000   | 0,0046   | 0,0000        | 0,0046        |

| Nó | HSSHPT1111 | HSSHPT1121 | HSSHNT1221 | HSSHNT1111 | HSSHNT1121 |
|----|------------|------------|------------|------------|------------|
| 8  | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,0000     |
| 9  | 0,0029     | 0,0000     | 0,0058     | 0,0029     | 0,0000     |
| 10 | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,0000     |
| 11 | 0,0000     | 0,0044     | 0,0000     | 0,0000     | 0,0000     |
| 6  | 0,0066     | 0,0000     | 0,0000     | 0,0000     | 0,0033     |
| 14 | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,0000     |
| 30 | 0,0000     | 0,0000     | 0,0000     | 0,0000     | 0,0000     |
| 31 | 0,0000     | 0,0046     | 0,0046     | 0,0000     | 0,0000     |

# ANEXO F

Probabilidade de ocorrência de todos os padrões de viagens – proveniente do relatório gerado pelo S-PLUS 6.1 (nós terminais)

## 6) AMOSTRA DE ESTUDANTES UNIVERSITÁRIOS

|    |              |              |              |              |              |               |               |             |              |              |              |              |
|----|--------------|--------------|--------------|--------------|--------------|---------------|---------------|-------------|--------------|--------------|--------------|--------------|
| Nó | HAHNIN1111   | HAHPP1111    | HAHPP1211    | HAHTT1131    | HAHAHPP1111  | HAHAHTT1111   | HAHNN111      | HAHPP111    | HAHPP121     | HAHPP131     | HAHSHPP1111  | HAHSHPP1131  |
| 8  | 0,0000       | 0,0000       | 0,0000       | 0,0189       | 0,0000       | 0,0000        | 0,0189        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 9  | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0247       | 0,0000       | 0,0000       | 0,0124       |
| 10 | 0,0000       | 0,0000       | 0,0000       | 0,0108       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0108       | 0,0108       | 0,0000       | 0,0000       |
| 11 | 0,0000       | 0,0000       | 0,0170       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0170       | 0,0170       | 0,0000       |
| 12 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0182       | 0,0182        | 0,0182        | 0,0182      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 13 | 0,0282       | 0,0000       | 0,0141       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0141      | 0,0000       | 0,0000       | 0,0141       | 0,0000       |
| 14 | 0,0000       | 0,0139       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0139      | 0,0000       | 0,0000       | 0,0000       | 0,0139       |
| 15 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0263        | 0,0263      | 0,0132       | 0,0000       | 0,0000       | 0,0000       |
| Nó | HAHSHPP12121 | HAHSHTT1111  | HAHSHTT12121 | HAHTP111     | HAHTT111     | HAHTT121      | HAHTT131      | HAHTT141    | HAHSHPP1111  | HAHSHPP11411 | HAHNTT1111   | HAHSHPP1111  |
| 8  | 0,0000       | 0,0189       | 0,0189       | 0,0000       | 0,0000       | 0,0189        | 0,0189        | 0,0189      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 9  | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0247      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 10 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0215        | 0,0000      | 0,0108       | 0,0108       | 0,0000       | 0,0000       |
| 11 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 12 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0182        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0182       | 0,0000       |
| 13 | 0,0000       | 0,0000       | 0,0000       | 0,0141       | 0,0282       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0141       |
| 14 | 0,0139       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0139       |
| 15 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0132       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| Nó | HAHSHPP1211  | HAHSHPP1221  | HAHSHPP1331  | HAHHTT1111   | HAHHTT1231   | HAHSHNPP1111  | HAHSHNNTT1111 | HAHSHPP1111 | HAHSHPP11121 | HAHSHPP12111 | HAHSHPP12121 |              |
| 8  | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 9  | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0124       | 0,0000       | 0,0000       | 0,0000       |
| 10 | 0,0000       | 0,0000       | 0,0215       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0108      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 11 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0509      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 12 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0182        | 0,0182      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 13 | 0,0000       | 0,0000       | 0,0000       | 0,0141       | 0,0000       | 0,0141        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 14 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0139       | 0,0000        | 0,0000        | 0,0139      | 0,0000       | 0,0000       | 0,0278       | 0,0139       |
| 15 | 0,0132       | 0,0132       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| Nó | HAHSHPP13111 | HAHSHPP14111 | HAHSHPT12111 | HAHSHPT14111 | HAHSHTTT1111 | HAHSHTTT11121 | HAHSHTTT12121 | HAHSHN111   | HAHSHN111    | HAHSHN111    | HAHSHN111    | HAHSHN111    |
| 8  | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0943       | 0,0000       | 0,0189       | 0,0377       |
| 9  | 0,0000       | 0,0000       | 0,0000       | 0,0124       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0247       | 0,0000       | 0,0000       | 0,0864       |
| 10 | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0108        | 0,0000      | 0,0538       | 0,0000       | 0,0000       | 0,0538       |
| 11 | 0,0509       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,1525       |
| 12 | 0,0182       | 0,0000       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,1455       | 0,0364       | 0,0000       | 0,0909       |
| 13 | 0,0000       | 0,0141       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0141      | 0,1972       | 0,0000       | 0,0000       | 0,0423       |
| 14 | 0,0000       | 0,0139       | 0,0000       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0139       | 0,0000       | 0,1667       |
| 15 | 0,0000       | 0,0132       | 0,0132       | 0,0000       | 0,0132       | 0,0000        | 0,0000        | 0,0000      | 0,0526       | 0,0000       | 0,0263       | 0,3026       |
| Nó | HAHSHPT121   | HAHSHPT131   | HAHSHPT141   | HAHSHPT111   | HAHSHPT121   | HAHSHPT131    | HAHSHPT141    | HAHSHPP1111 | HAHSHPP12111 | HAHSHPP12121 | HAHSHPP13111 | HAHSHPP14111 |
| 8  | 0,0566       | 0,0377       | 0,0189       | 0,0000       | 0,0000       | 0,0000        | 0,0189        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 9  | 0,1605       | 0,0617       | 0,3457       | 0,0000       | 0,0000       | 0,0124        | 0,0124        | 0,0124      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 10 | 0,0645       | 0,0360       | 0,0538       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0108       | 0,0000       | 0,0108       |
| 11 | 0,1864       | 0,2203       | 0,0339       | 0,0000       | 0,0170       | 0,0170        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 12 | 0,0364       | 0,0364       | 0,0000       | 0,0000       | 0,0182       | 0,0000        | 0,0000        | 0,0000      | 0,0364       | 0,0000       | 0,0182       | 0,0000       |
| 13 | 0,1268       | 0,0000       | 0,0141       | 0,0000       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 14 | 0,2500       | 0,0972       | 0,0694       | 0,0139       | 0,0000       | 0,0000        | 0,0000        | 0,0000      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |
| 15 | 0,1053       | 0,0000       | 0,0658       | 0,0658       | 0,0000       | 0,0000        | 0,0000        | 0,0263      | 0,0000       | 0,0000       | 0,0000       | 0,0000       |

| Nó | HSHTTNN13111 | HSHTTTT11111 | HSHTTTT12121 | HSHTP111 | HSHTP121 | HSHTT111 | HSHTT121 | HSHTT131 |
|----|--------------|--------------|--------------|----------|----------|----------|----------|----------|
| 8  | 0,0000       | 0,0000       | 0,0189       | 0,0000   | 0,0000   | 0,0566   | 0,0189   | 0,0755   |
| 9  | 0,0000       | 0,0000       | 0,0000       | 0,0000   | 0,0000   | 0,0494   | 0,0247   | 0,0370   |
| 10 | 0,0108       | 0,0000       | 0,0000       | 0,0430   | 0,0108   | 0,0645   | 0,1720   | 0,1828   |
| 11 | 0,0000       | 0,0000       | 0,0000       | 0,0000   | 0,0000   | 0,0000   | 0,0509   | 0,1017   |
| 12 | 0,0000       | 0,0000       | 0,0000       | 0,0000   | 0,0182   | 0,0000   | 0,2182   | 0,0000   |
| 13 | 0,0000       | 0,0141       | 0,0000       | 0,0000   | 0,0000   | 0,2254   | 0,1549   | 0,0141   |
| 14 | 0,0000       | 0,0139       | 0,0139       | 0,0000   | 0,0000   | 0,0694   | 0,0833   | 0,0139   |
| 15 | 0,0000       | 0,0000       | 0,0000       | 0,0000   | 0,0000   | 0,1579   | 0,0395   | 0,0000   |

| Nó | HSHTT141 | HSSHNNN11111 | HSSHPPP11111 | HSSHPPP1211 | HSSHTTT11111 | HSSHTTT1121 |
|----|----------|--------------|--------------|-------------|--------------|-------------|
| 8  | 0,4151   | 0,0000       | 0,0000       | 0,0000      | 0,0000       | 0,0000      |
| 9  | 0,0864   | 0,0000       | 0,0000       | 0,0000      | 0,0000       | 0,0000      |
| 10 | 0,0538   | 0,0000       | 0,0000       | 0,0000      | 0,0108       | 0,0000      |
| 11 | 0,0678   | 0,0000       | 0,0000       | 0,0000      | 0,0000       | 0,0000      |
| 12 | 0,0364   | 0,0727       | 0,0000       | 0,0727      | 0,0000       | 0,0000      |
| 13 | 0,0282   | 0,0000       | 0,0000       | 0,0000      | 0,0000       | 0,0000      |
| 14 | 0,0139   | 0,0000       | 0,0139       | 0,0000      | 0,0000       | 0,0139      |
| 15 | 0,0000   | 0,0000       | 0,0132       | 0,0000      | 0,0000       | 0,0000      |

# ANEXO G

Sumário dos modelos lineares, ANOVA, coeficientes estimados, Histograma, Curva acumulado e resíduos (SPSS 13.0).

## 1) AMOSTRA DE INDUSTRIÁRIOS

### 1.1 MODELO HWH77121

**Model Summary<sup>b</sup>**

| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1     | ,945 <sup>a</sup> | ,893     | ,893              | *****                      |

a. Predictors: (Constant), USO\_SOLO2, EST, CNH, VTRA, RENDA

b. Dependent Variable: HWHTT121

**ANOVA<sup>b</sup>**

| Model |            | Sum of Squares | df   | Mean Square | F        | Sig.              |
|-------|------------|----------------|------|-------------|----------|-------------------|
| 1     | Regression | 18,712         | 5    | 3,742       | 6855,560 | ,000 <sup>a</sup> |
|       | Residual   | 2,236          | 4096 | ,001        |          |                   |
|       | Total      | 20,948         | 4101 |             |          |                   |

a. Predictors: (Constant), USO\_SOLO2, EST, CNH, VTRA, RENDA

b. Dependent Variable: HWHTT121

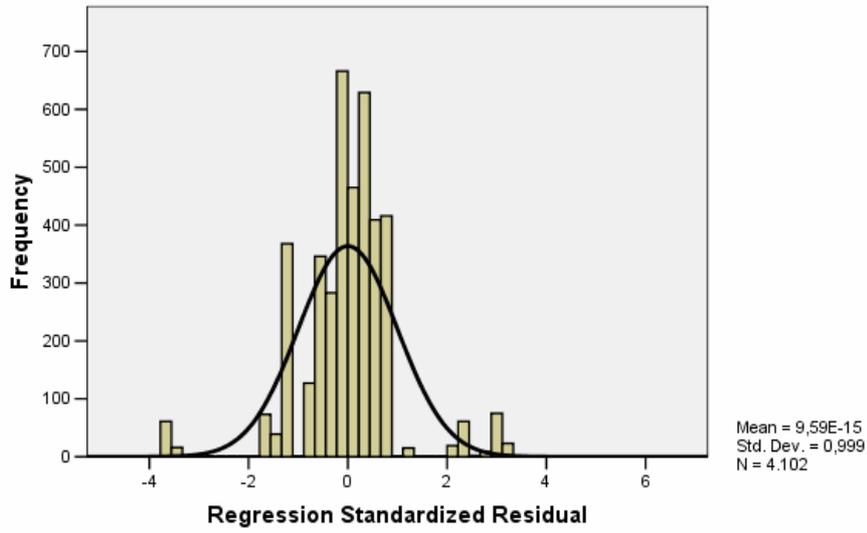
**Coefficients<sup>a</sup>**

| Model |            | Unstandardized Coefficients |            | Standardized Coefficients | t        | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|----------|------|
|       |            | B                           | Std. Error | Beta                      |          |      |
| 1     | (Constant) | ,227                        | ,001       |                           | 315,234  | ,000 |
|       | CNH        | -,038                       | ,001       | -,265                     | -47,779  | ,000 |
|       | VTRA       | -,099                       | ,001       | -,664                     | -124,785 | ,000 |
|       | RENDA      | -,055                       | ,001       | -,274                     | -48,198  | ,000 |
|       | EST        | -,061                       | ,001       | -,268                     | -52,312  | ,000 |
|       | USO_SOLO2  | -,004                       | ,001       | -,028                     | -5,254   | ,000 |

a. Dependent Variable: HWHTT121

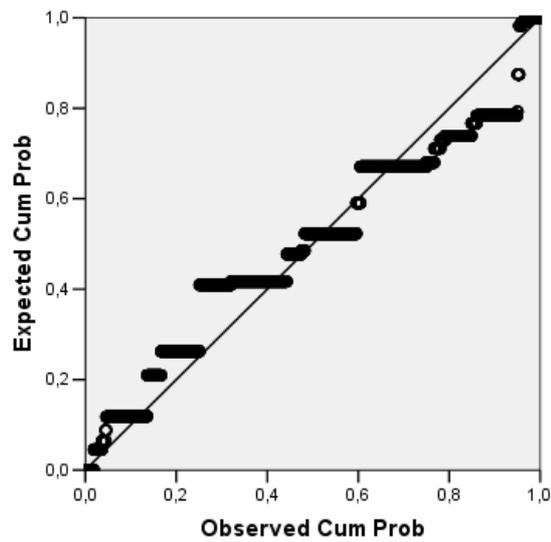
### Histogram

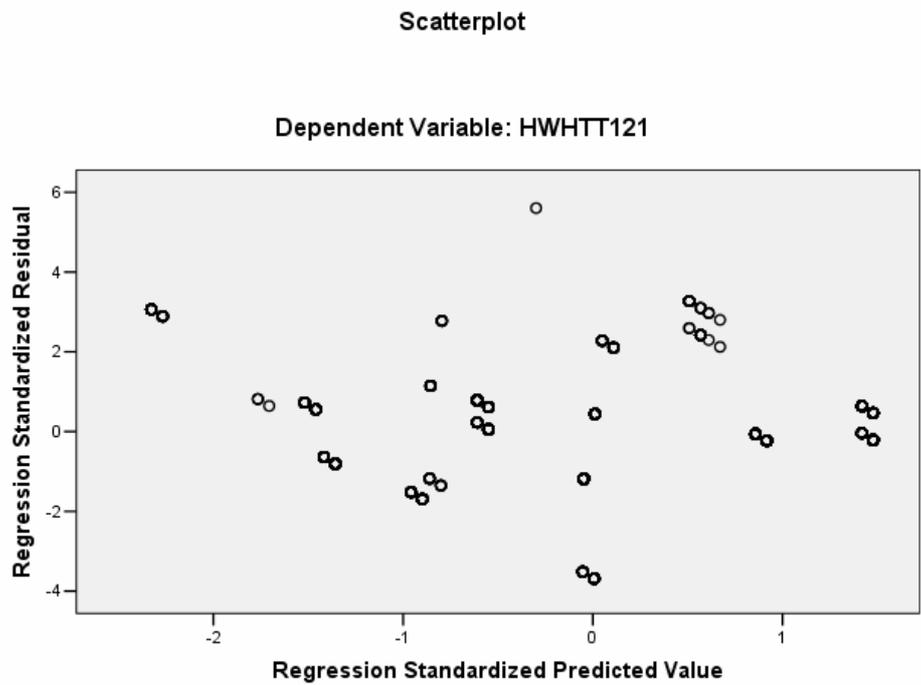
Dependent Variable: HWHTT121



### Normal P-P Plot of Regression Standardized Residual

Dependent Variable: HWHTT121





**1.2 MODELO HWHTT131**

**Model Summary<sup>b</sup>**

| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1     | ,913 <sup>a</sup> | ,834     | ,834              | *****                      |

a. Predictors: (Constant), USO\_SOLO3, EST, CNH, VTRA, RENDA

b. Dependent Variable: HWHTT131

**ANOVA<sup>b</sup>**

| Model |            | Sum of Squares | df   | Mean Square | F        | Sig.              |
|-------|------------|----------------|------|-------------|----------|-------------------|
| 1     | Regression | 7,537          | 5    | 1,507       | 4124,444 | ,000 <sup>a</sup> |
|       | Residual   | 1,497          | 4096 | ,000        |          |                   |
|       | Total      | 9,034          | 4101 |             |          |                   |

a. Predictors: (Constant), USO\_SOLO3, EST, CNH, VTRA, RENDA

b. Dependent Variable: HWHTT131

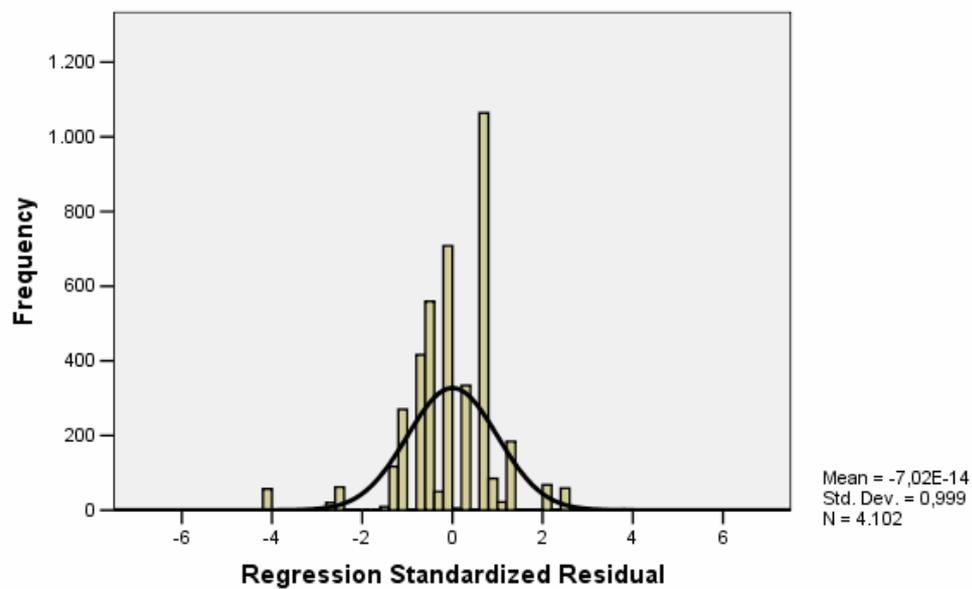
Coefficients<sup>a</sup>

| Model |            | Unstandardized Coefficients |            | Standardized Coefficients | t        | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|----------|------|
|       |            | B                           | Std. Error | Beta                      |          |      |
| 1     | (Constant) | ,150                        | ,001       |                           | 254,431  | ,000 |
|       | CNH        | -,006                       | ,001       | -,068                     | -9,811   | ,000 |
|       | VTRA       | -,069                       | ,001       | -,709                     | -106,773 | ,000 |
|       | REND A     | -,022                       | ,001       | -,171                     | -24,164  | ,000 |
|       | EST        | -,033                       | ,001       | -,219                     | -34,292  | ,000 |
|       | USO_SOLO3  | -,026                       | ,001       | -,275                     | -41,757  | ,000 |

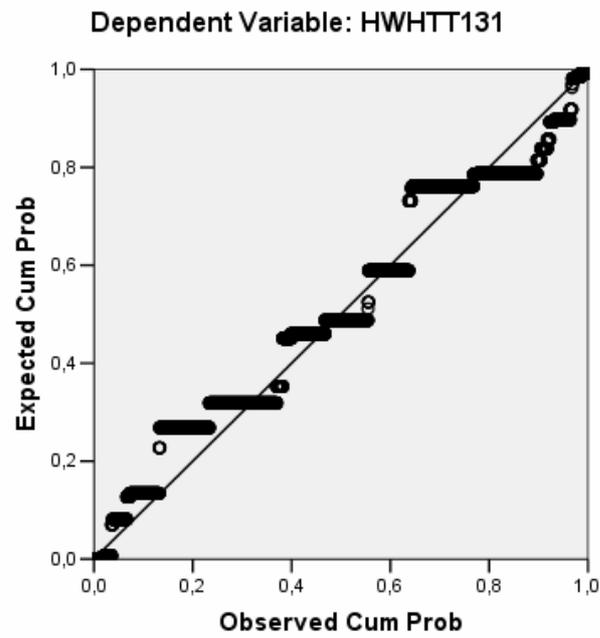
a. Dependent Variable: HWHTT131

Histogram

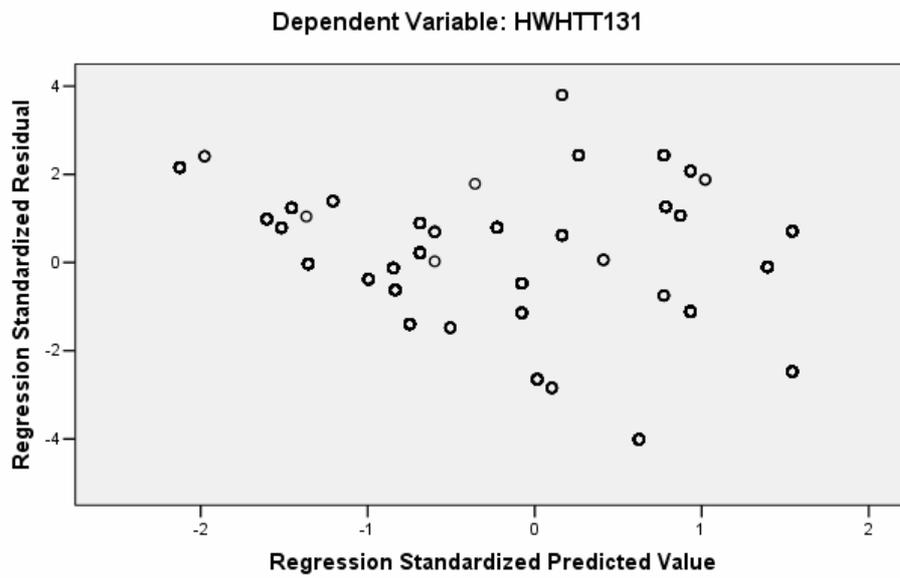
Dependent Variable: HWHTT131



### Normal P-P Plot of Regression Standardized Residual



Scatterplot



## 2) AMOSTRA DE COMERCIÁRIOS

## 2.1 MODELO HWHNN111

Model Summary<sup>b</sup>

| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1     | ,845 <sup>a</sup> | ,713     | ,713              | ,0495210                   |

a. Predictors: (Constant), USO\_SOLO2, TRABALHA, AUT, USO\_SOLO1, EST, RENDA, VTRA

b. Dependent Variable: HWHNN111

ANOVA<sup>b</sup>

| Model |            | Sum of Squares | df   | Mean Square | F        | Sig.              |
|-------|------------|----------------|------|-------------|----------|-------------------|
| 1     | Regression | 36,833         | 7    | 5,262       | 2145,659 | ,000 <sup>a</sup> |
|       | Residual   | 14,800         | 6035 | ,002        |          |                   |
|       | Total      | 51,633         | 6042 |             |          |                   |

a. Predictors: (Constant), USO\_SOLO2, TRABALHA, AUT, USO\_SOLO1, EST, RENDA, VTRA

b. Dependent Variable: HWHNN111

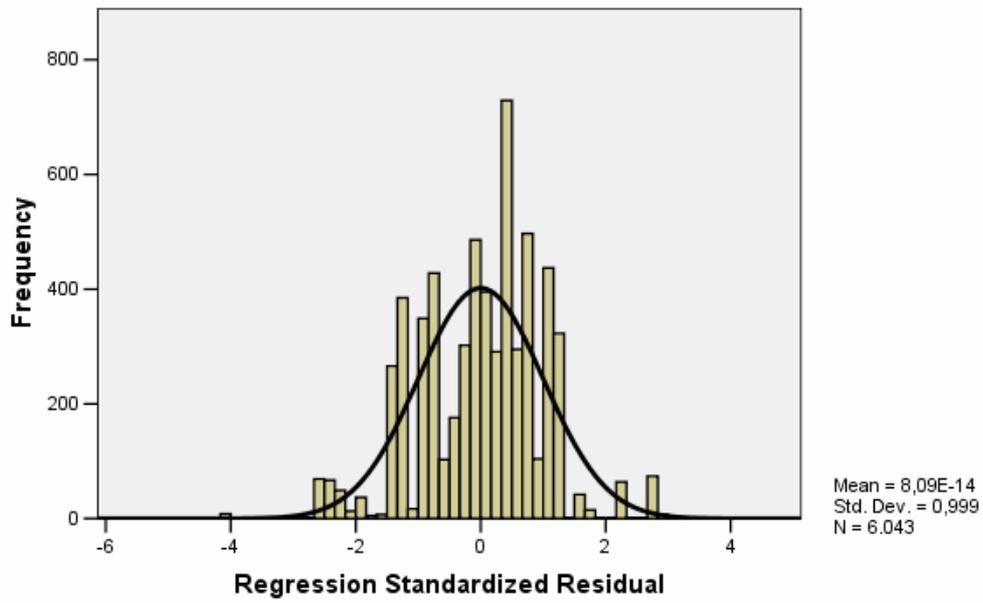
Coefficients<sup>a</sup>

| Model |            | Unstandardized Coefficients |            | Standardized Coefficients | t       | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|---------|------|
|       |            | B                           | Std. Error | Beta                      |         |      |
| 1     | (Constant) | ,114                        | ,003       |                           | 44,166  | ,000 |
|       | VTRA       | ,149                        | ,002       | ,703                      | 89,381  | ,000 |
|       | AUT        | -,087                       | ,001       | -,457                     | -59,561 | ,000 |
|       | EST        | -,043                       | ,002       | -,171                     | -22,905 | ,000 |
|       | RENDA      | -,026                       | ,002       | -,130                     | -16,524 | ,000 |
|       | USO_SOLO1  | ,005                        | ,001       | ,028                      | 3,767   | ,000 |
|       | TRABALHA   | -,012                       | ,001       | -,062                     | -8,164  | ,000 |
|       | USO_SOLO2  | -,026                       | ,002       | -,084                     | -10,740 | ,000 |

a. Dependent Variable: HWHNN111

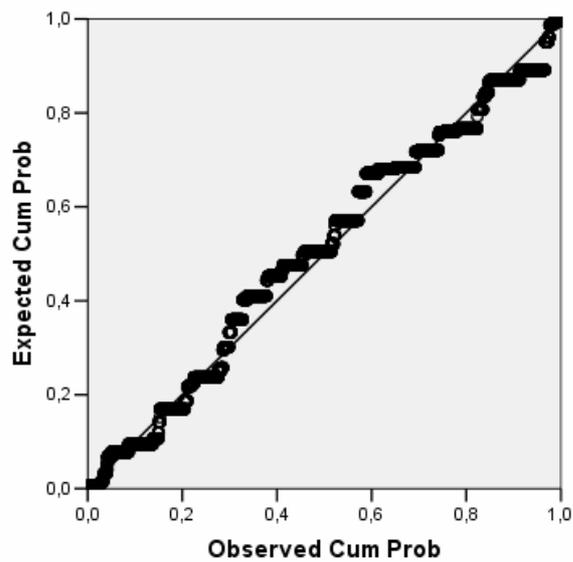
### Histogram

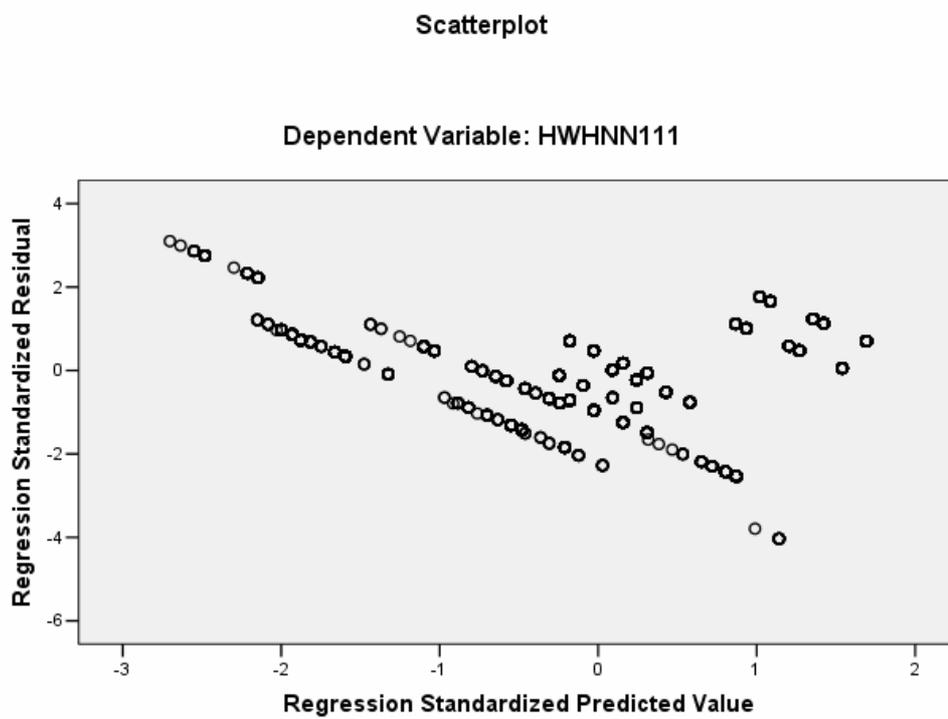
Dependent Variable: HWHNN111



### Normal P-P Plot of Regression Standardized Residual

Dependent Variable: HWHNN111





### 3) AMOSTRA DE TRABALHADORES NO SETOR DE SERVIÇOS

#### 3.1 MODELO HWH77121

**Model Summary<sup>b</sup>**

| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1     | ,930 <sup>a</sup> | ,865     | ,865              | ,0233665                   |

a. Predictors: (Constant), USO\_SOLO1, EST, SEXO, VTRA, CNH, AUT

b. Dependent Variable: HWHTT121

**ANOVA<sup>b</sup>**

| Model |            | Sum of Squares | df   | Mean Square | F        | Sig.              |
|-------|------------|----------------|------|-------------|----------|-------------------|
| 1     | Regression | 29,159         | 6    | 4,860       | 8900,958 | ,000 <sup>a</sup> |
|       | Residual   | 4,543          | 8321 | ,001        |          |                   |
|       | Total      | 33,702         | 8327 |             |          |                   |

a. Predictors: (Constant), USO\_SOLO1, EST, SEXO, VTRA, CNH, AUT

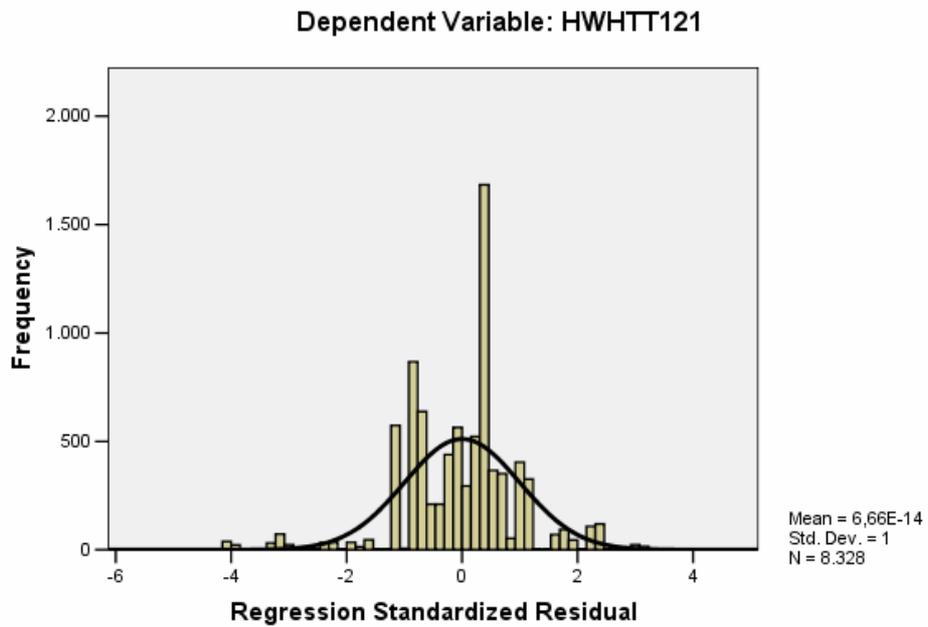
b. Dependent Variable: HWHTT121

**Coefficients<sup>a</sup>**

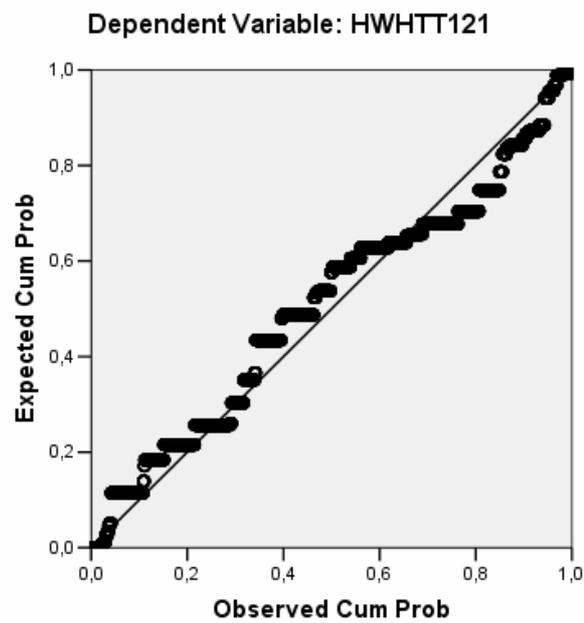
| Model |            | Unstandardized Coefficients |            | Standardized Coefficients | t        | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|----------|------|
|       |            | B                           | Std. Error | Beta                      |          |      |
| 1     | (Constant) | ,209                        | ,001       |                           | 315,204  | ,000 |
|       | CNH        | -,047                       | ,001       | -,357                     | -78,749  | ,000 |
|       | EST        | -,064                       | ,001       | -,312                     | -77,262  | ,000 |
|       | VTRA       | -,096                       | ,001       | -,662                     | -158,888 | ,000 |
|       | SEXO       | ,003                        | ,001       | ,024                      | 6,053    | ,000 |
|       | AUT        | -,030                       | ,001       | -,206                     | -45,205  | ,000 |
|       | USO_SOLO1  | ,017                        | ,001       | ,135                      | 32,575   | ,000 |

a. Dependent Variable: HWHTT121

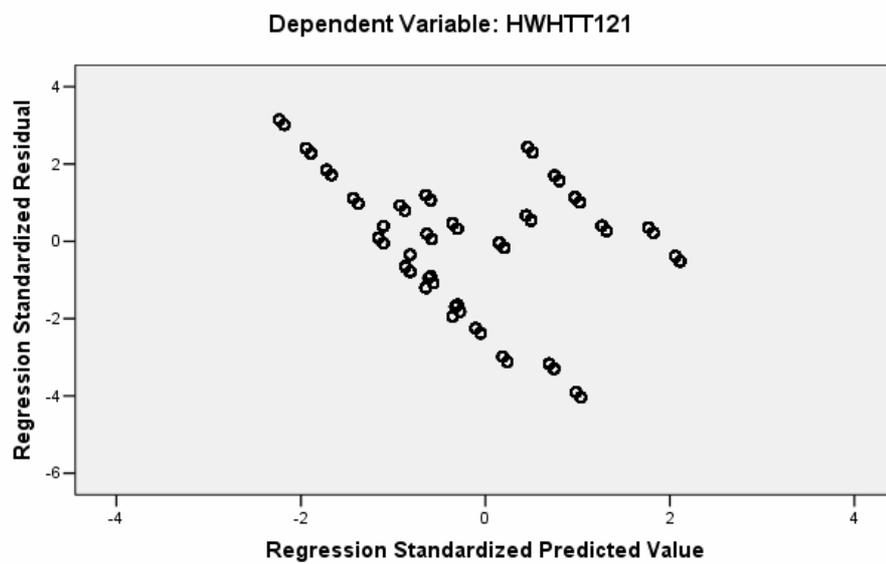
**Histogram**



### Normal P-P Plot of Regression Standardized Residual



### Scatterplot



## 4) AMOSTRA DE ESTUDANTES ATÉ O 1º GRAU

## 4.1 MODELO HSHMM11

**Model Summary<sup>b</sup>**

| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1     | ,931 <sup>a</sup> | ,866     | ,866              | ,0675061                   |

a. Predictors: (Constant), USO\_SOLO1, IDADE 1, AUT, RF

b. Dependent Variable: HSHNN111

**ANOVA<sup>b</sup>**

| Model |            | Sum of Squares | df    | Mean Square | F         | Sig.              |
|-------|------------|----------------|-------|-------------|-----------|-------------------|
| 1     | Regression | 408,968        | 4     | 102,242     | 22435,894 | ,000 <sup>a</sup> |
|       | Residual   | 63,043         | 13834 | ,005        |           |                   |
|       | Total      | 472,010        | 13838 |             |           |                   |

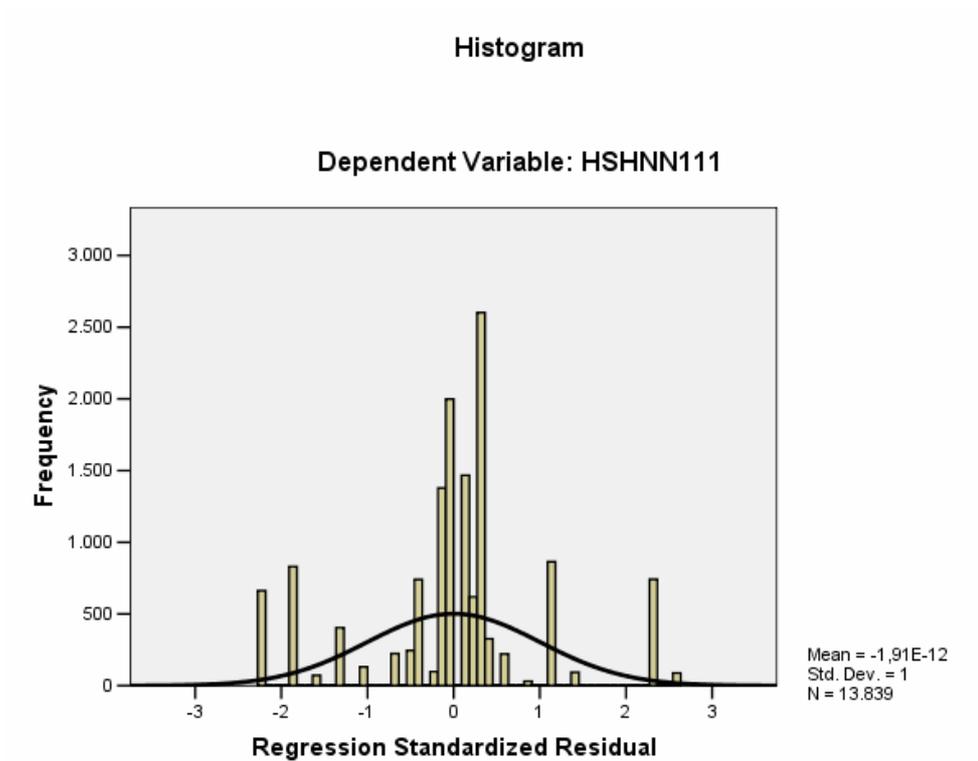
a. Predictors: (Constant), USO\_SOLO1, IDADE 1, AUT, RF

b. Dependent Variable: HSHNN111

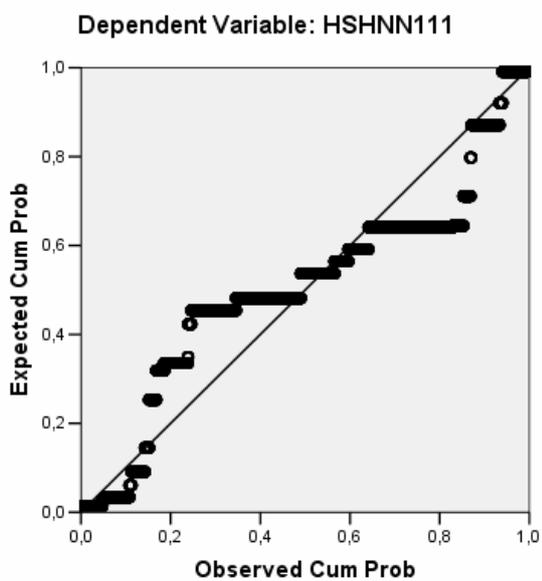
**Coefficients<sup>a</sup>**

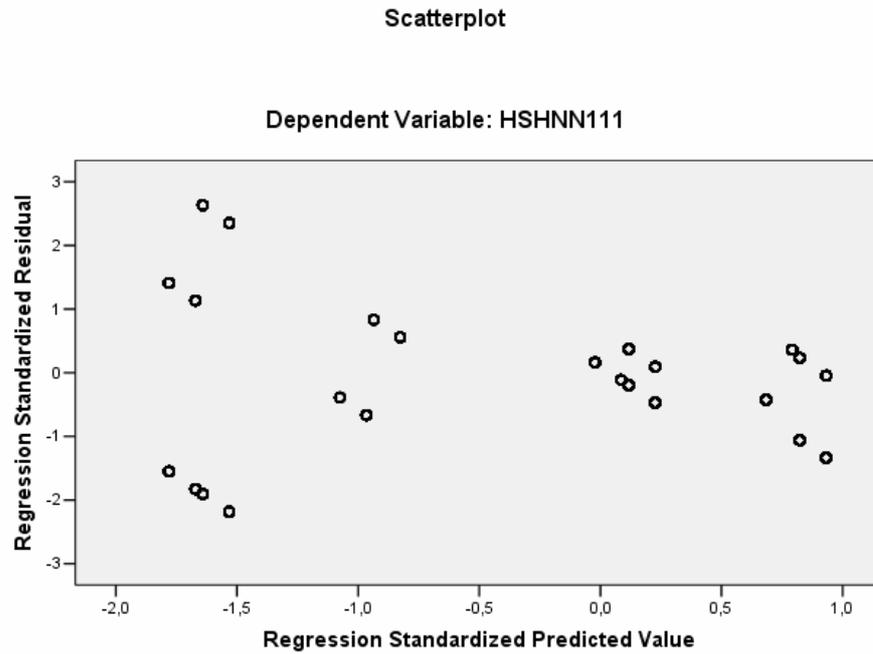
| Model |            | Unstandardized Coefficients |            | Standardized Coefficients | t        | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|----------|------|
|       |            | B                           | Std. Error | Beta                      |          |      |
| 1     | (Constant) | ,711                        | ,002       |                           | 470,086  | ,000 |
|       | RF         | -,302                       | ,001       | -,733                     | -206,494 | ,000 |
|       | AUT        | -,121                       | ,001       | -,328                     | -93,479  | ,000 |
|       | IDADE 1    | ,024                        | ,001       | ,065                      | 20,906   | ,000 |
|       | USO_SOLO1  | ,019                        | ,001       | ,040                      | 12,774   | ,000 |

a. Dependent Variable: HSHNN111



Normal P-P Plot of Regression Standardized Residual





5) AMOSTRA DE ESTUDANTES NO 2º GRAU

5.1 MODELO HSHMM11

**Model Summary<sup>b</sup>**

| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1     | ,838 <sup>a</sup> | ,703     | ,703              | ,0769716                   |

a. Predictors: (Constant), USO\_SOLO1, RENDA1

b. Dependent Variable: HSHNN111

**ANOVA<sup>b</sup>**

| Model |            | Sum of Squares | df   | Mean Square | F        | Sig.              |
|-------|------------|----------------|------|-------------|----------|-------------------|
| 1     | Regression | 29,996         | 2    | 14,998      | 2531,484 | ,000 <sup>a</sup> |
|       | Residual   | 12,679         | 2140 | ,006        |          |                   |
|       | Total      | 42,675         | 2142 |             |          |                   |

a. Predictors: (Constant), USO\_SOLO1, RENDA1

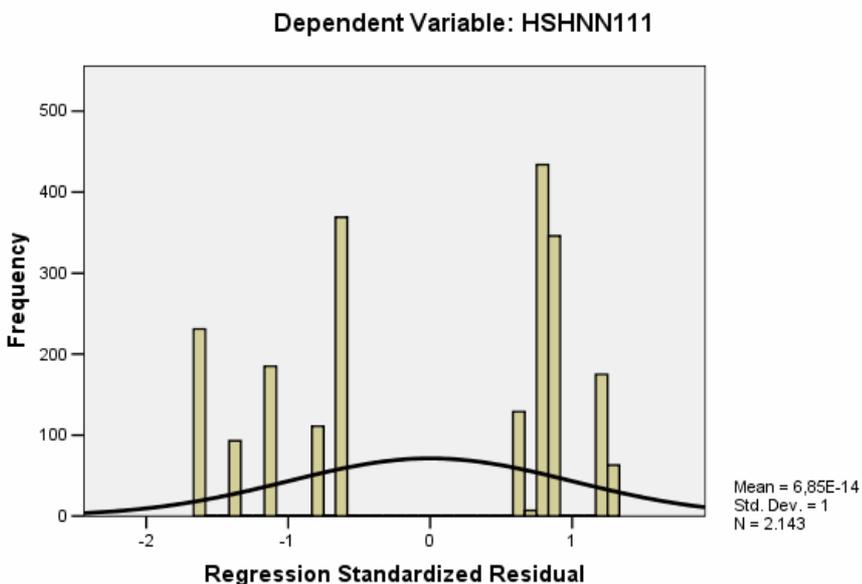
b. Dependent Variable: HSHNN111

**Coefficients<sup>a</sup>**

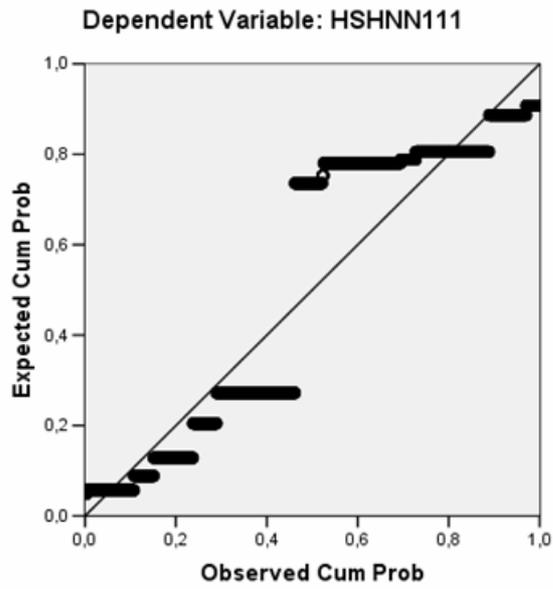
| Model |            | Unstandardized Coefficients |            | Standardized Coefficients | t       | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|---------|------|
|       |            | B                           | Std. Error | Beta                      |         |      |
| 1     | (Constant) | ,442                        | ,003       |                           | 165,303 | ,000 |
|       | RENDA1     | -,226                       | ,004       | -,781                     | -62,412 | ,000 |
|       | USO_SOLO1  | -,040                       | ,004       | -,141                     | -11,270 | ,000 |

a. Dependent Variable: HSHNN111

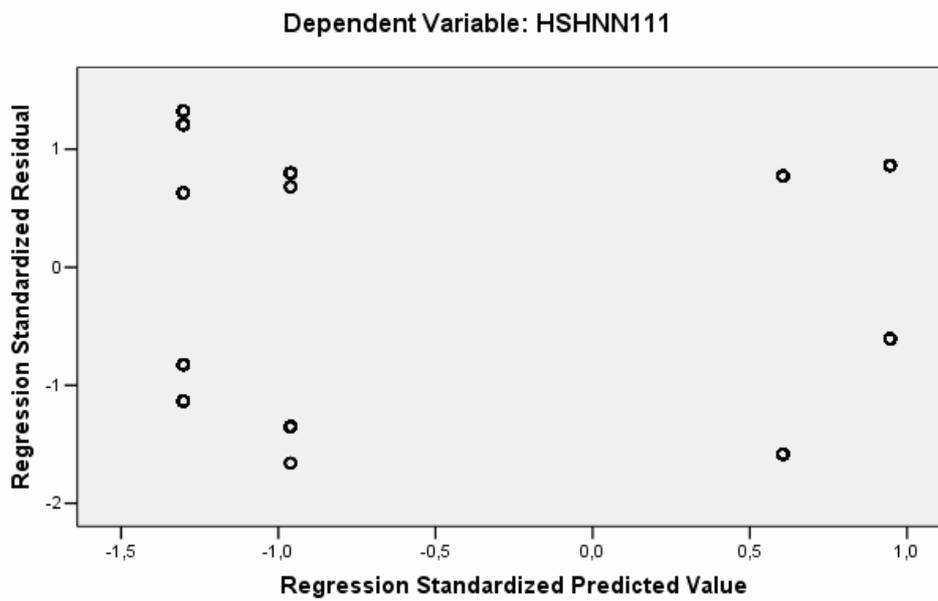
**Histogram**



Normal P-P Plot of Regression Standardized Residual



Scatterplot



## 6) AMOSTRA DE ESTUDANTES UNIVERSITÁRIOS

## 6.1 MODELO HSHPP121

**Model Summary<sup>b</sup>**

| Model | R                 | R Square | Adjusted R Square | Std. Error of the Estimate |
|-------|-------------------|----------|-------------------|----------------------------|
| 1     | ,620 <sup>a</sup> | ,385     | ,383              | ,0499321                   |

a. Predictors: (Constant), USO\_SOLO1, RENDA 3

b. Dependent Variable: HSHPP121

**ANOVA<sup>b</sup>**

| Model |            | Sum of Squares | df  | Mean Square | F       | Sig.              |
|-------|------------|----------------|-----|-------------|---------|-------------------|
| 1     | Regression | ,881           | 2   | ,440        | 176,642 | ,000 <sup>a</sup> |
|       | Residual   | 1,409          | 565 | ,002        |         |                   |
|       | Total      | 2,289          | 567 |             |         |                   |

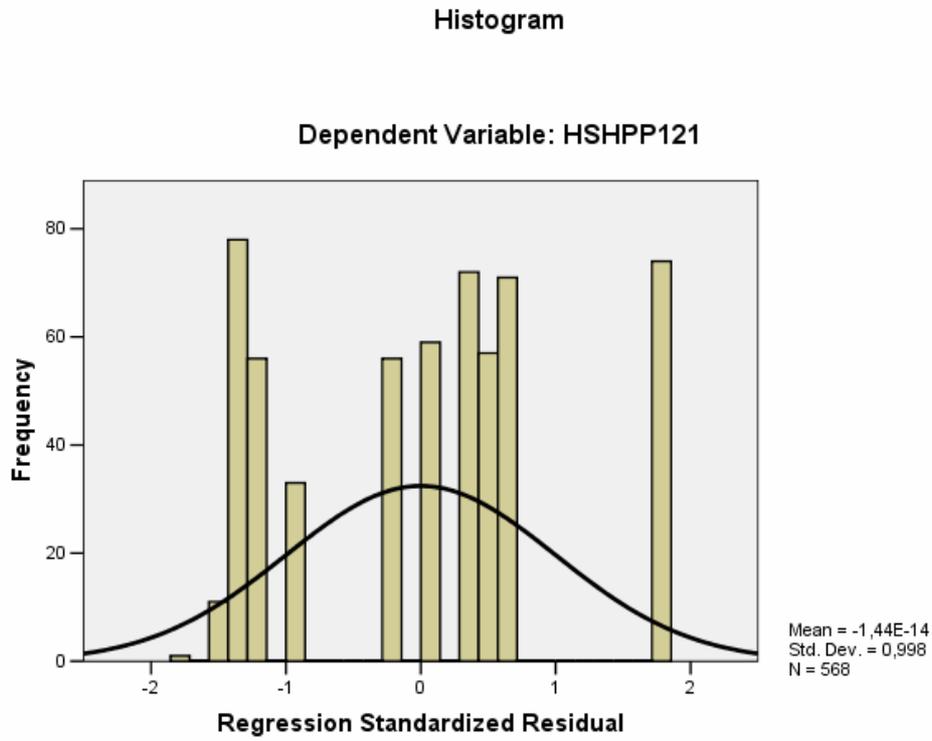
a. Predictors: (Constant), USO\_SOLO1, RENDA 3

b. Dependent Variable: HSHPP121

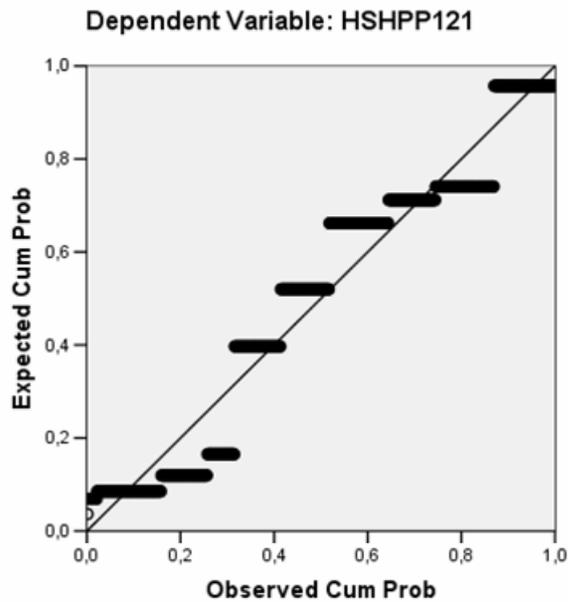
**Coefficients<sup>a</sup>**

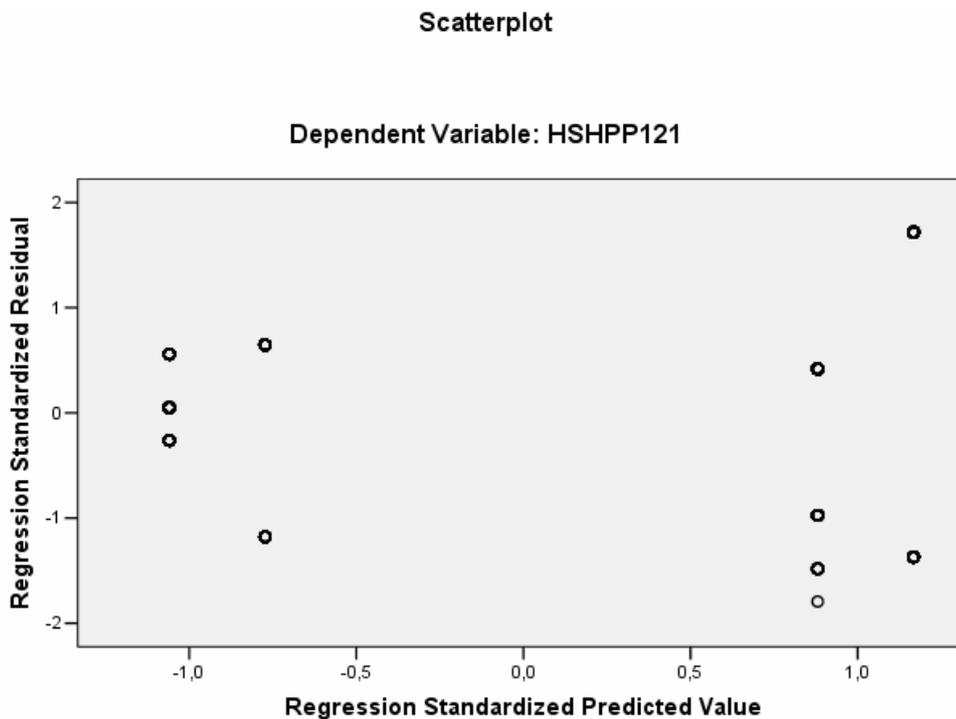
| Model |            | Unstandardized Coefficients |            | Standardized Coefficients | t      | Sig. |
|-------|------------|-----------------------------|------------|---------------------------|--------|------|
|       |            | B                           | Std. Error | Beta                      |        |      |
| 1     | (Constant) | ,083                        | ,003       |                           | 24,466 | ,000 |
|       | RENDA 3    | ,076                        | ,004       | ,601                      | 18,047 | ,000 |
|       | USO_SOLO1  | ,011                        | ,004       | ,089                      | 2,672  | ,008 |

a. Dependent Variable: HSHPP121



Normal P-P Plot of Regression Standardized Residual





Foi obtido um total de 60 modelos lineares, com intuito de associar significância estatística às variáveis independentes. A significância das variáveis envolvidas foi mensurada pelo teste t e também pelos coeficientes estimados. No entanto, a afirmação de que determinada variável é estatisticamente significativa depende da qualidade ou robustez dos modelos. A verificação das suposições básicas da Regressão Linear, como homocedasticidade, normalidade, multicolinearidade, etc., faz parte dos cuidados que devem ser levados em conta numa análise de modelos lineares.

Neste anexo, foram apresentados 7 exemplos de modelos obtidos para as seis amostras analisadas. Verifica-se que nas amostras menores (estudantes no 2º grau e universitários – itens 5.1 e 6.1) aparentemente não há normalidade, pelos menos nos exemplos mostrados.

Observa-se também um possível problema de heterocedasticidade (verificando os resíduos) no modelo exemplificado no item 2.1.

Desta forma, deve-se fazer ressalvas quanto à qualidade de todos os modelos obtidos. Alguns deles atendem àquelas suposições básicas, no entanto pode haver problemas com muitos outros. Deve-se atentar especialmente para os casos das amostras onde haja menor número de observações.

# Livros Grátis

( <http://www.livrosgratis.com.br> )

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)  
[Baixar livros de Literatura de Cordel](#)  
[Baixar livros de Literatura Infantil](#)  
[Baixar livros de Matemática](#)  
[Baixar livros de Medicina](#)  
[Baixar livros de Medicina Veterinária](#)  
[Baixar livros de Meio Ambiente](#)  
[Baixar livros de Meteorologia](#)  
[Baixar Monografias e TCC](#)  
[Baixar livros Multidisciplinar](#)  
[Baixar livros de Música](#)  
[Baixar livros de Psicologia](#)  
[Baixar livros de Química](#)  
[Baixar livros de Saúde Coletiva](#)  
[Baixar livros de Serviço Social](#)  
[Baixar livros de Sociologia](#)  
[Baixar livros de Teologia](#)  
[Baixar livros de Trabalho](#)  
[Baixar livros de Turismo](#)