



FACULDADE DE ECONOMIA E FINANÇAS IBMEC  
PROGRAMA DE PÓS-GRADUAÇÃO E PESQUISA EM  
ADMINISTRAÇÃO E ECONOMIA

DISSERTAÇÃO DE MESTRADO  
PROFISSIONALIZANTE EM ADMINISTRAÇÃO

**“MINERAÇÃO DE DADOS DE HELP DESK  
USANDO RATTLE – O CASO  
PETROBRAS”**

**ULISSES CORRÊA**

ORIENTADOR: MARIA AUGUSTA MACHADO

**Rio de Janeiro, 15 de Fevereiro de 2007**

# **Livros Grátis**

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

Milhares de livros grátis para download.

“MINERAÇÃO DE DADOS DE HELP DESK USANDO RATTLE – O CASO  
PETROBRAS”

ULISSES CORRÊA

Dissertação apresentada ao curso de  
Mestrado Profissionalizante em  
Administração como requisito parcial para  
obtenção do Grau de Mestre em  
Administração.

Área de Concentração: Administração  
Geral

ORIENTADOR: Prof<sup>ª</sup> Dr<sup>ª</sup> MARIA AUGUSTA MACHADO

Rio de Janeiro, 15 de fevereiro de 2007.

“MINERAÇÃO DE DADOS DE HELP DESK USANDO RATTLE – O CASO  
PETROBRAS”

ULISSES CORRÊA

Dissertação apresentada ao curso de  
Mestrado Profissionalizante em  
Administração como requisito parcial para  
obtenção do Grau de Mestre em  
Administração.  
Área de Concentração: Administração  
Geral

Avaliação:

BANCA EXAMINADORA:

---

Professor Dr<sup>a</sup> MARIA AUGUSTA MACHADO (Orientador)  
Instituição: Faculdades IBMEC

---

Professor Dr. EDSON DALTO  
Instituição: Faculdades IBMEC

---

Professor Dr. JORGE MUNIZ BARRETO  
Instituição: Universidade Federal de Santa Catarina - UFSC

Rio de Janeiro, 15 de fevereiro de 2007.

005.741  
C824

Corrêa, Ulisses.

Mineração de dados de help desk usando rattle: o caso Petrobras / Ulisses Corrêa. - Rio de Janeiro: Faculdades Ibmecc. 2007.

Dissertação de Mestrado Profissionalizante apresentada ao Programa de Pós-Graduação em Administração das Faculdades Ibmecc, como requisito parcial necessário para a obtenção do título de Mestre em Administração.

Área de concentração: Administração Geral.

1. Mineração de dados. 2. Regras de associação. 3. Help Desk.

## **DEDICATÓRIA**

Dedico este trabalho aos meus pais, pela firme persistência frente às dificuldades e desafios impostos pela vida.

## **AGRADECIMENTOS**

À minha orientadora, Maria Augusta, pela sua dedicação e apoio durante todo o curso. Ela despertou em mim a paixão pelo Fuzzy, pela Inteligência Artificial e pela Mineração de Dados.

À minha família, em especial meu filho que soube suportar o longo tempo de minha ausência

À minha eterna namorada, sem seu incentivo e ajuda este trabalho não seria possível.

À minha grande amiga Denise, pela força nos momentos mais difíceis deste trabalho.

Aos colegas de trabalho da TI/APS/CIS que tanto ajudaram na revisão de texto e na formatação no Word.

Ao meu irmão Carlos pelo incentivo, mostrando que é possível e que devemos buscar sempre o aperfeiçoamento.

Aos ilustres colegas do Ibmecc Hime Aguiar e José Mounir. Foi uma honra cursar algumas disciplinas com pessoas tão brilhantes.

A todos os colegas e gerentes da Petrobras que direta ou indiretamente contribuíram para este trabalho.

## RESUMO

Este estudo se propõe a utilizar as técnicas e tarefas da mineração de dados para análise dos registros de atendimento aos usuários de Tecnologia da Informação de uma grande organização, com o objetivo de extrair conhecimento útil e aplicável à tomada de decisão gerencial. Estes registros de solicitação de atendimento são criados, na maioria dos casos, pelos técnicos do *Help-desk* com apoio de uma ferramenta de *workflow* que gera as bases de dados. Diariamente, são criados cerca de 8000 novos registros de atendimento, o que por si só inviabiliza qualquer análise humana sem auxílio de ferramentas computacionais. Através do processo de descoberta de conhecimento em grandes bases de dados e com o apoio de softwares de exploração e mineração de dados, estes dados foram transformados e processados. O resultado final é apresentado em conclusões extraídas utilizando-se a técnica de mineração de dados denominada regras de associação. Estas regras irão apresentar a relação entre fatos que a princípio não apresentavam relação entre si, revelando conhecimento novo.

Palavras Chave: Mineração de Dados, Regras de Associação, *Help-desk*.



## **ABSTRACT**

This study intend to use techniques and tasks of data mining for analysis the support calls of Information Technology users of a great organization, with the objective of extract useful knowledge, applicable to managerial decision making. These calls are created, in most of cases, by the Help-desk technicians with the support of a workflow tool that create the data bases. Daily, about 8000 new support calls are created, what turn into a not viable human analysis without aid of computational tools. Through the knowledge discovery process in large data bases and with de support of data mining and exploration softwares, this data were transformed and processed. The final result is presented in conclusions extracted using the data mining technique called association rules. These rules will exhibit the relation between facts that, at first, don't exhibit relation among them, disclosing new knowledge.

Keywords: Data Mining, Association Rules, Help-desk.

## LISTA DE FIGURAS

<b>FIGURA 1: ORGANOGRAMA DA PETROBRAS .....</b>	<b>3</b>
<b>FIGURA 2: NOVO POSICIONAMENTO DA FUNÇÃO TI DA PETROBRAS.....</b>	<b>5</b>
<b>FIGURA 3: ORGANOGRAMA SIMPLIFICADO DA TI DA PETROBRAS .....</b>	<b>7</b>
<b>FIGURA 4: ORGANOGRAMA DE SERVIÇOS DA ÁREA DE SERVIÇOS DA TI DA PETROBRAS.....</b>	<b>9</b>
<b>FIGURA 5: ÁREA DE ATUAÇÃO GEOGRÁFICA DA ÁREA DE SERVIÇOS .....</b>	<b>10</b>
<b>FIGURA 6: FLUXO DE ATENDIMENTO .....</b>	<b>11</b>
<b>FIGURA 7: AS ETAPAS DO PROCESSO DE KDD .....</b>	<b>16</b>
<b>FIGURA 10 – PRÉ-PROCESSAMENTO DOS DADOS DE HELP DESK .....</b>	<b>35</b>
<b>FIGURA 11 – VOLUME DE REGISTROS – COMPARATIVO 2005-2006.....</b>	<b>37</b>

## LISTA DE TABELAS

Tabela 1 – Tarefas e Técnicas de Mineração de Dados .....	23
Tabela 2 – Número total de registros do CIANS.....	36
Tabela 3 – Distribuição de Frequência – Atributo Categoria.....	38
Tabela 4 - Distribuição de Frequência – Atributo Tipo de Registro .....	39
Tabela 5 - Distribuição de Frequência – Atributo Período.....	40
Tabela 6 - Distribuição de Frequência – Atributo Período.....	40
Tabela 7 - Distribuição de Frequência – Atributo Origem do Registro .....	41
Tabela 8 - Distribuição de Frequência – Atributo Tempo Total de atendimento.....	42
Tabela 9 – Análise das Regras – Atributos Categoria X Origem.....	46
Tabela 10 – Análise das Regras – Atributos Categoria X Tempo Total .....	47
Tabela 11 – Análise das Regras – Atributos Categoria X Período.....	48
Tabela 12 – Análise das Regras – Atributos Categoria X Tipo de Grupo.....	49
Tabela 13 – Análise das Regras – Atributos Categoria X Tipo de Registro .....	50
Tabela 14 – Análise das Regras – Atributos Origem X Período.....	51
Tabela 15 – Análise das Regras – Atributos Origem X Tempo Total.....	52
Tabela 16 – Análise das Regras – Atributos Origem X Tipo de Registro.....	53
Tabela 17 – Análise das Regras – Atributos Origem X Tipo de Grupo.....	54
Tabela 18 – Análise das Regras – Atributos Período X Tipo de Grupo.....	55
Tabela 19 – Análise das Regras – Atributos Tipo de Registro X Período .....	56
Tabela 20 – Análise das Regras – Atributos Tipo de Registro X Tempo Total.....	57
Tabela 21 – Análise das Regras – Atributos Mês/Ano X Período .....	58

## LISTA DE ABREVIATURAS

ARS	<i>Action Request System</i>
HD	<i>Help-Desk</i>
IA	Inteligência Artificial
ITIL	<i>Information Technology Infrastructure Library</i>
KDD	<i>Knowledge Discovery in Databases</i>
ME	Mesa Especializada
PA	Posto Avançado
RAM	<i>Random Access memory</i>
TI	Tecnologia da Informação

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>14</b>
<b>1.1</b>	<b>CENÁRIO.....</b>	<b>2</b>
1.1.1	A empresa Petrobras .....	2
1.1.2	A área de TI da Petrobras.....	3
1.1.3	O serviço de Atendimento ao usuário de TI.....	10
<b>1.2</b>	<b>OBJETIVO .....</b>	<b>13</b>
1.2.1	Objetivo Geral.....	13
1.2.2	Objetivos específicos .....	13
<b>1.3</b>	<b>ESTRUTURA DO TRABALHO .....</b>	<b>14</b>
<b>2</b>	<b>REVISÃO DA LITERATURA .....</b>	<b>15</b>
<b>2.1</b>	<b>O PROCESSO KDD – KNOWLEDGE DISCOVERY IN DATABASES .....</b>	<b>15</b>
<b>2.2</b>	<b>AS ETAPAS DO PROCESSO KDD.....</b>	<b>17</b>
2.2.1	O Pré-processamento dos dados.....	17
2.2.1.1	Seleção dos dados.....	18
2.2.1.2	Limpeza de dados.....	19
2.2.1.3	Codificação dos dados.....	21
2.2.1.4	Enriquecimento dos dados.....	21
2.2.2	Mineração de Dados.....	21
2.2.3	O pós-processamento .....	25
<b>3</b>	<b>DESCOBRINDO CONHECIMENTO EM BASES DE HELP-DESK .....</b>	<b>27</b>
<b>3.1</b>	<b>O PRÉ-PROCESSAMENTO DOS DADOS DE HELP-DESK.....</b>	<b>27</b>
3.1.1	A seleção dos atributos.....	28
3.1.2	Limpeza e seleção dos dados .....	33
3.1.3	Transformando os dados – codificação e enriquecimento.....	33
<b>3.2</b>	<b>ANÁLISE EXPLORATÓRIA DOS DADOS .....</b>	<b>36</b>
<b>3.3</b>	<b>A SELEÇÃO DA FERRAMENTA PARA MINERAÇÃO DOS DADOS .....</b>	<b>42</b>
3.3.1	A ferramenta escolhida .....	44
<b>3.4</b>	<b>ANÁLISE DOS RESULTADOS ENCONTRADOS .....</b>	<b>45</b>
3.4.1	Consolidando alguns resultados encontrados.....	59
<b>4</b>	<b>CONCLUSÃO.....</b>	<b>60</b>

<b>4.1</b>	<b>TRABALHOS FUTUROS.....</b>	<b>61</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>63</b>
	<b>BIBLIOGRAFIA COMPLEMENTAR.....</b>	<b>65</b>
	<b>APÊNDICE A - PEQUENO GUIA PARA GERAÇÃO DE REGAS DE ASSOCIAÇÃO USANDO O RATTLE.....</b>	<b>66</b>
	<b>APÊNDICE B – REGRAS CONSOLIDADAS ATRIBUTO - CATEGORIA .....</b>	<b>84</b>

## 1 INTRODUÇÃO

O crescimento, em escala exponencial, do volume de dados armazenados nos Centros de Processamento de Dados é uma realidade comum a todas as grandes Organizações do mundo moderno. Ele pode ser atribuído a fatores externos ou internos às organizações, tais como:

- A necessidade de manter o histórico de suas operações e de seus dados;
- A complexidade das transações;
- A globalização das empresas;
- As exigências legais de registros.

Em contrapartida a todo este volume de informações disponíveis, hoje em dia, está a limitada capacidade do homem para analisar grandes conjuntos de dados. Mesmo as ferramentas de *software* disponíveis, até pouco tempo atrás, não permitiam realizar esta tarefa de forma efetiva. Este fato motivou o desenvolvimento de novos meios computacionais para auxiliar na execução desta tarefa.

Neste contexto, a Mineração de Dados surge como uma excelente ferramenta para a análise de grandes conjuntos de dados. Apoiada na estatística e na inteligência artificial, a mineração de dados apresenta eficiência na descoberta de conhecimento que se encontrava oculto nas “montanhas” de dados acumulados ao longo dos anos.

Este trabalho irá explorar o processo de descoberta de conhecimento e as técnicas de Mineração de Dados, utilizando a base de dados de *Help-Desk* da Petrobras.

## 1.1 CENÁRIO

### 1.1.1 A empresa Petrobras

A Petrobras é a maior empresa do Brasil e da América Latina. Recentemente, em seu novo posicionamento estratégico, passou a se apresentar como uma empresa de energia e não mais voltada apenas para atividades relacionadas com o petróleo. Tem presença marcante no Brasil e no exterior, principalmente na América Latina. A mudança de foco está refletida na missão e visão 2015, extraídas do site da Petrobras na internet, descritas a seguir:

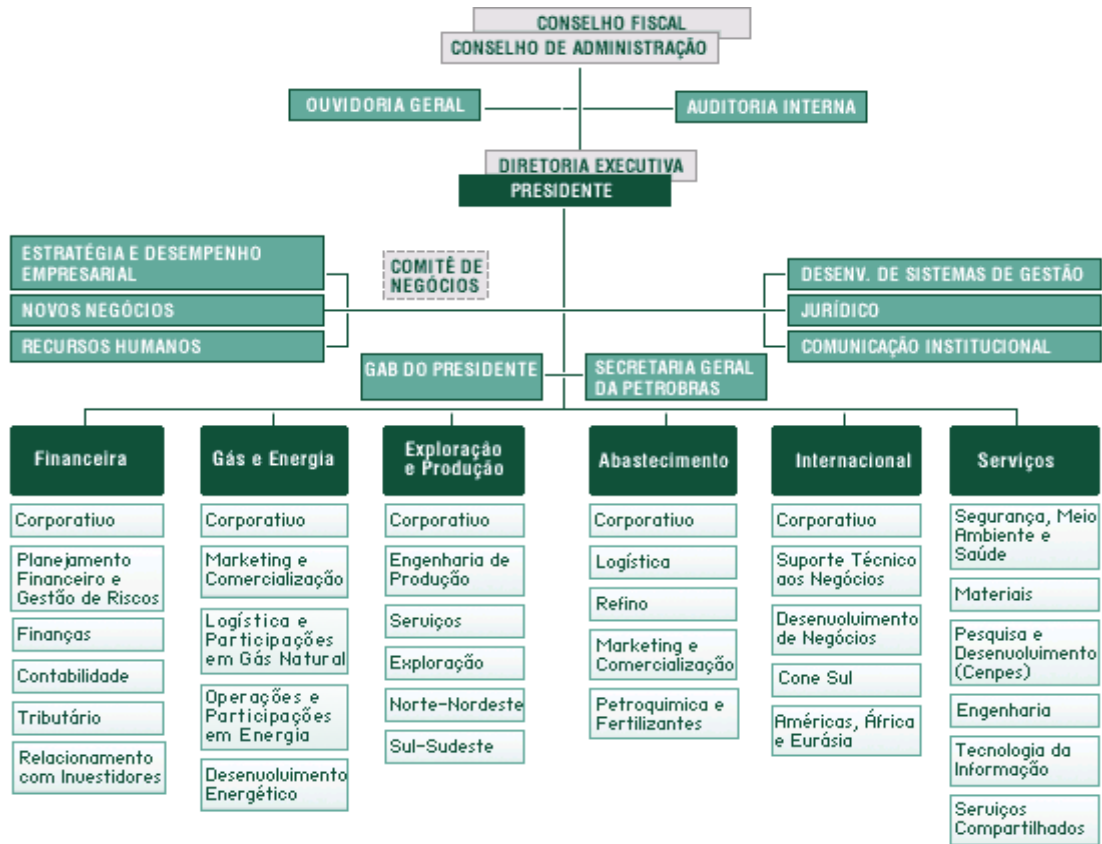
“Missão: Atuar de forma segura e rentável, com responsabilidade social e ambiental, nas atividades da indústria de óleo, gás e energia, nos mercados nacional e internacional, fornecendo produtos e serviços adequados às necessidades dos seus clientes e contribuindo para o desenvolvimento do Brasil e dos países onde atua”.

“Visão 2015: A Petrobras será uma empresa integrada de energia com forte presença internacional e líder na América Latina, atuando com foco na rentabilidade e na responsabilidade social e ambiental”.

A estrutura organizacional da empresa é dividida em áreas de negócio: Exploração & Produção, Abastecimento, Internacional, Gás & Energia; Área Corporativa: Jurídico, RH, Novos Negócios, etc; Área Financeira; Área de Serviços: Pesquisa, Tecnologia da Informação, Engenharia, Materiais, Compartilhados e Segurança Meio-Ambiente Saúde; além da alta administração formada pelos conselhos, diretoria e Presidente.



Na Figura 1, a seguir, podemos melhor visualizar esta estrutura.



**Figura 1: Organograma da PETROBRAS**

Fonte: site da Petrobras

### 1.1.2 A área de TI da Petrobras

A área de tecnologia da informação da Petrobras – TI está diretamente subordinada a diretoria de serviços. Tem tamanho e estrutura que refletem a dimensão da Petrobras, atendendo as demandas de Tecnologia da Informação de todas as áreas de negócio e corporativas da companhia.

Até o ano de 2003, a área de TI da Petrobras era descentralizada e as unidades de negócio tinham relativa autonomia para criar e manter suas estruturas de TI de forma a atender suas demandas conforme suas necessidades. Este modelo visava dar velocidade as decisões, impactando o mínimo possível os projetos de cada unidade e permitia que cada órgão de TI fosse bastante especializado no negócio de sua unidade uma vez que a área de TI fazia parte da estrutura organizacional unidade de negócio.

No entanto, este modelo teve como consequência problemas de integração, custos e compatibilidade. Não era incomum encontrar produtos diferenciados para executar uma mesma tarefa em unidades diferentes, que não podiam ser integrados ou cuja integração teria custos elevados. Não raro também era a duplicidade de esforços, nos órgãos de TI das unidades negócio, visando o mesmo objetivo. Desenvolvimento de ferramentas semelhantes, operações de sistemas idênticos, compras descentralizadas com preços diferentes para aquisição dos mesmos produtos eram alguns dos problemas que comprovavam o esgotamento deste modelo.

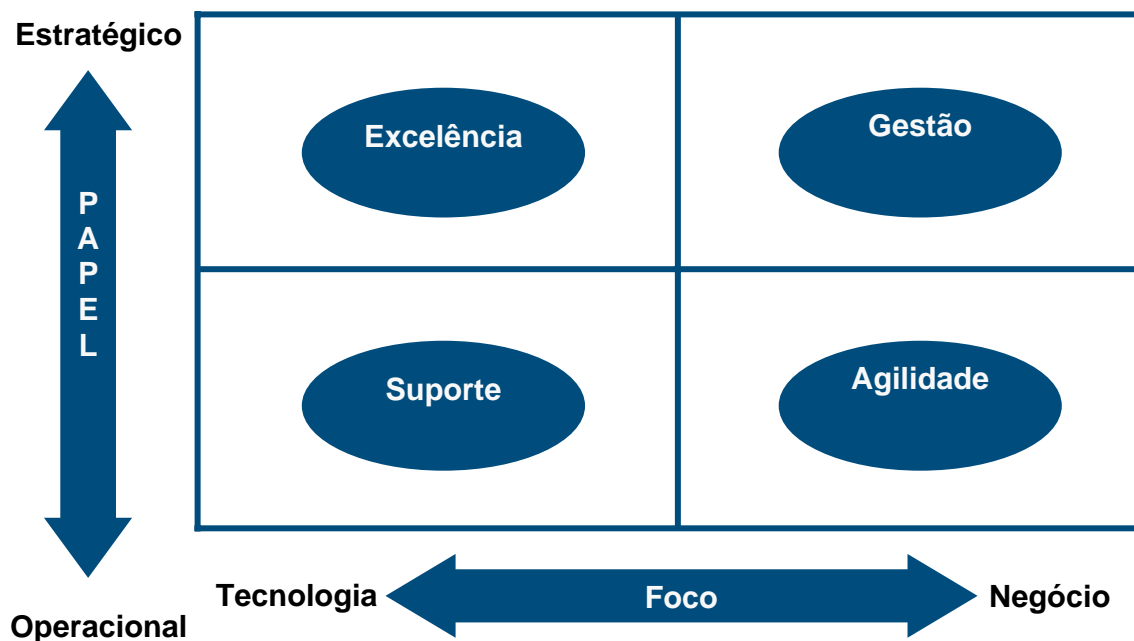
Faltava a Tecnologia de Informação da Petrobras um planejamento estratégico abrangente, para toda companhia e que ao mesmo tempo estivesse alinhado com as estratégias de negócio de cada unidade.

Diante deste quadro, no início do ano de 2004, a Petrobras implantou um novo modelo de gestão de TI, que visava à integração de todas as áreas de tecnologia de informação em um único órgão, sob uma gestão única.

O gerente executivo da Função TI sintetizou a importância deste novo órgão ao afirmar que: “O grande desafio da TI na Petrobras é uniformizar a aplicação de políticas e diretrizes, o que

proporcionará ganhos de escala e redução de custos, sem perder a agilidade e o alinhamento necessários ao negócio”.

O quadro a seguir apresenta este novo posicionamento da função TI. Envolve quatro quadrantes com modelos de atuação típicos e características bem definidas.



**Figura 2: Novo posicionamento da Função TI da PETROBRAS**

**Fonte: TI da Petrobras**

As áreas de Excelência e Gestão têm papéis estratégicos e atuam como órgãos de assessoria. A área de Gestão tem como principal foco as aquisições de bens de serviços de TI, RH, o controle contábil-financeiro e gestão de projetos. No quadrante de excelência estão as gerências que controlam a integração das demandas de TI (visando evitar a duplicação de esforços) e gerência de Tecnologia e Arquitetura que faz a prospecção de novas tecnologias e atua na padronização, na gestão do conhecimento e disseminação de modelos de metadados corporativos.

As áreas de Serviços e de Agilidade atuam com foco essencialmente operacional. A área de Agilidade atua prioritariamente como interface entre o cliente e a área de Serviços. Pode atuar também como intermediário entre serviços contratados externamente e o cliente, quando estes são específicos das áreas de negócio que representam. A área de serviços provê o atendimento às demandas internas de TI de todas as Agilidades e atua diretamente no apoio ao usuário final.

As principais atribuições da área de Agilidade:

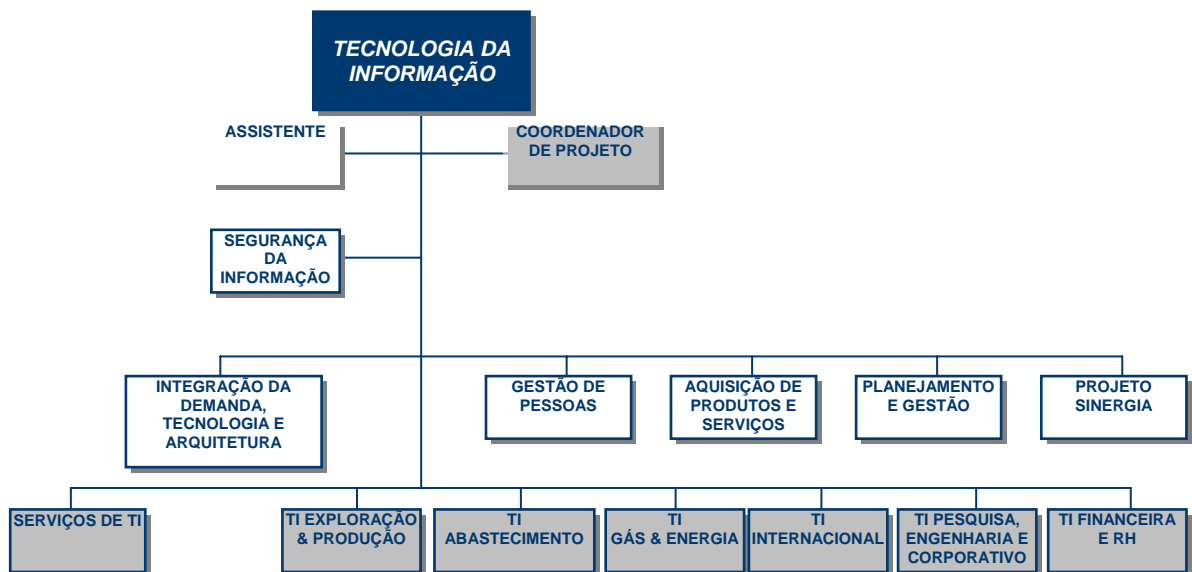
- Identifica oportunidades junto ao cliente e levanta as necessidades de soluções de TI;
- Analisa e prioriza demandas locais e departamentais;
- Encaminha demandas corporativas à Integração da Demanda;
- Lidera o provimento de soluções;
- Especifica os requisitos funcionais (macro requisitos) junto ao cliente;
- Monitora o cumprimento dos ANSs;
- Responde pelo ANS ao Cliente, garantindo a qualidade dos serviços prestados;
- Identifica junto ao cliente necessidades de prospecção de tecnologias
- Realiza a gestão da carteira de projetos de TI (custos, benefícios, riscos e progresso), dentro dos respectivos níveis de responsabilidade;
- Administra e opera a infra-estrutura específica, observando os procedimentos e padrões corporativos;
- Planeja os recursos necessários para as operações específicas.

As principais atribuições da área de Serviços:

- Planeja a capacidade e os recursos necessários para a operação dos Serviços de TI
- Planeja e desenvolve ou contrata soluções locais, departamentais e corporativas;
- Identifica e planeja demandas de infra-estrutura e de apoio ao usuário, interagindo com os centros de agilidade;
- Administra e opera a infra-estrutura regional e local (aplicativos, bases de dados, ambientes operacionais, servidores, redes);

- Executa o apoio remoto e local ao usuário;
- Interage com os Gestores de Demanda e agilidades na implementação dos projetos;
- Implementa a excelência operacional nos Serviços de TI;
- Monitora os níveis de prestação dos serviços;
- Monitora os níveis de segurança;
- Realiza a gestão da prestação dos serviços de comunicação de dados locais e de longa distância;
- Realiza a gestão de ativos de TI sob sua responsabilidade

A figura 3 apresenta o organograma simplificado da TI da Petrobras após o processo de reorganização, com os quatro quadrantes representados pelas gerências de TI. Na parte inferior, em destaque, estão as áreas de Serviço e de Agilidade cujas atribuições já foram descritas anteriormente.



**Figura 3: Organograma Simplificado da TI da PETROBRAS**

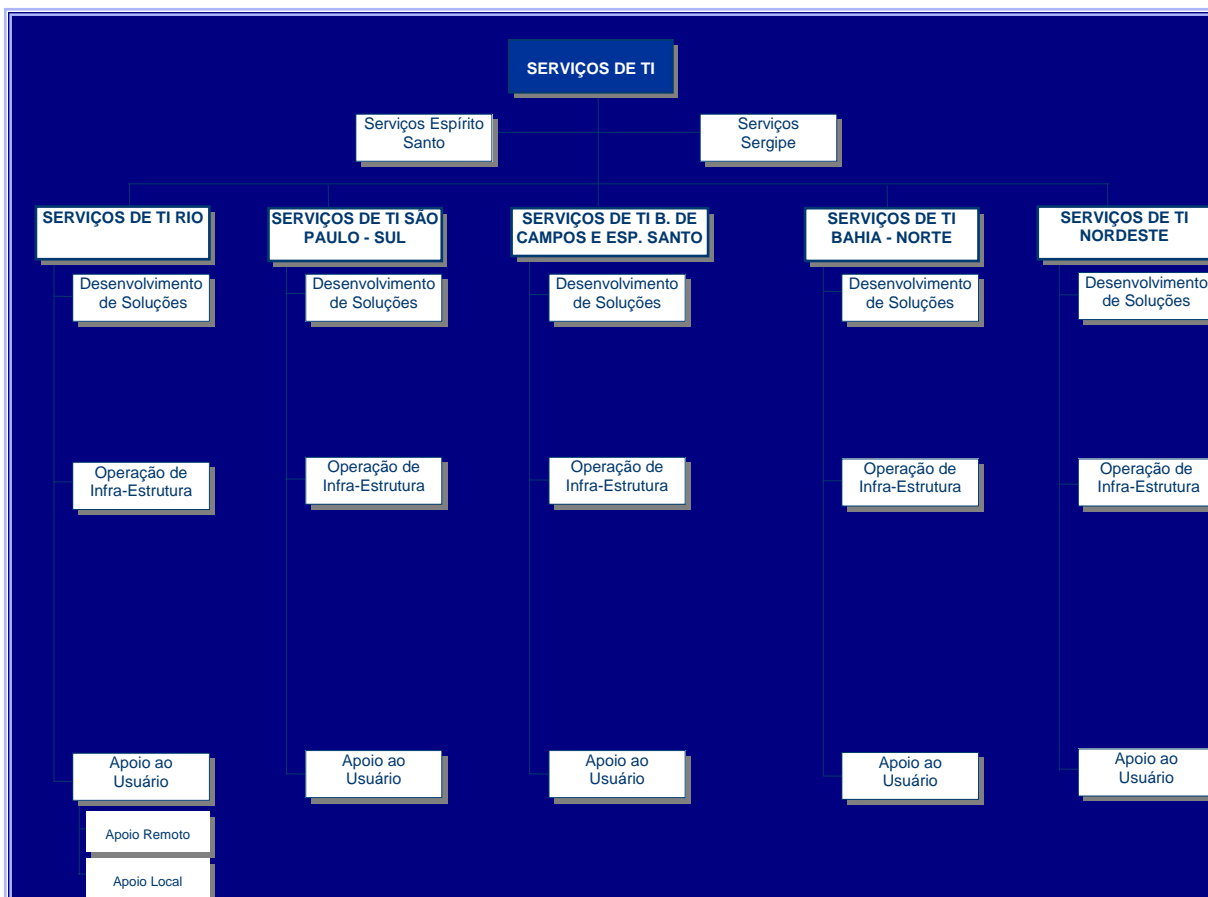
**Fonte: TI da Petrobras**

A área de Serviços de TI será o ponto focal deste estudo uma vez que é a partir dela que o usuário interage com TI. Apesar de as áreas de Agilidade serem a referência de contato para os clientes de TI das unidades de negócio, é o serviço de apoio ao usuário, estrutura

subordinada a área de Serviços, que recebe a maior parte das demandas dos usuários no dia-a-dia. Esta área de serviços que, se encontra na linha de frente com o usuário final, é o principal gerador dos registros da base de dados que será alvo de análise deste estudo.

A área de Serviços da TI é responsável por atender a todas as necessidades de TI que são demandadas das áreas de agilidade ou diretamente pelo usuário final. Sua estrutura segue um modelo geográfico e por processo, segmentada nas chamadas regionais de serviços, como mostra a figura 4. Seus principais serviços estão relacionados com o Desenvolvimento de Soluções (criação de sistemas internamente ou através de fábricas de software), Infra-estrutura (operação de *Data Centers*) e Apoio ao Usuário (serviços de *Help-Desk*). A estrutura de cada regional geográfica depende, fundamentalmente, da quantidade de usuários que por ela é suportada.

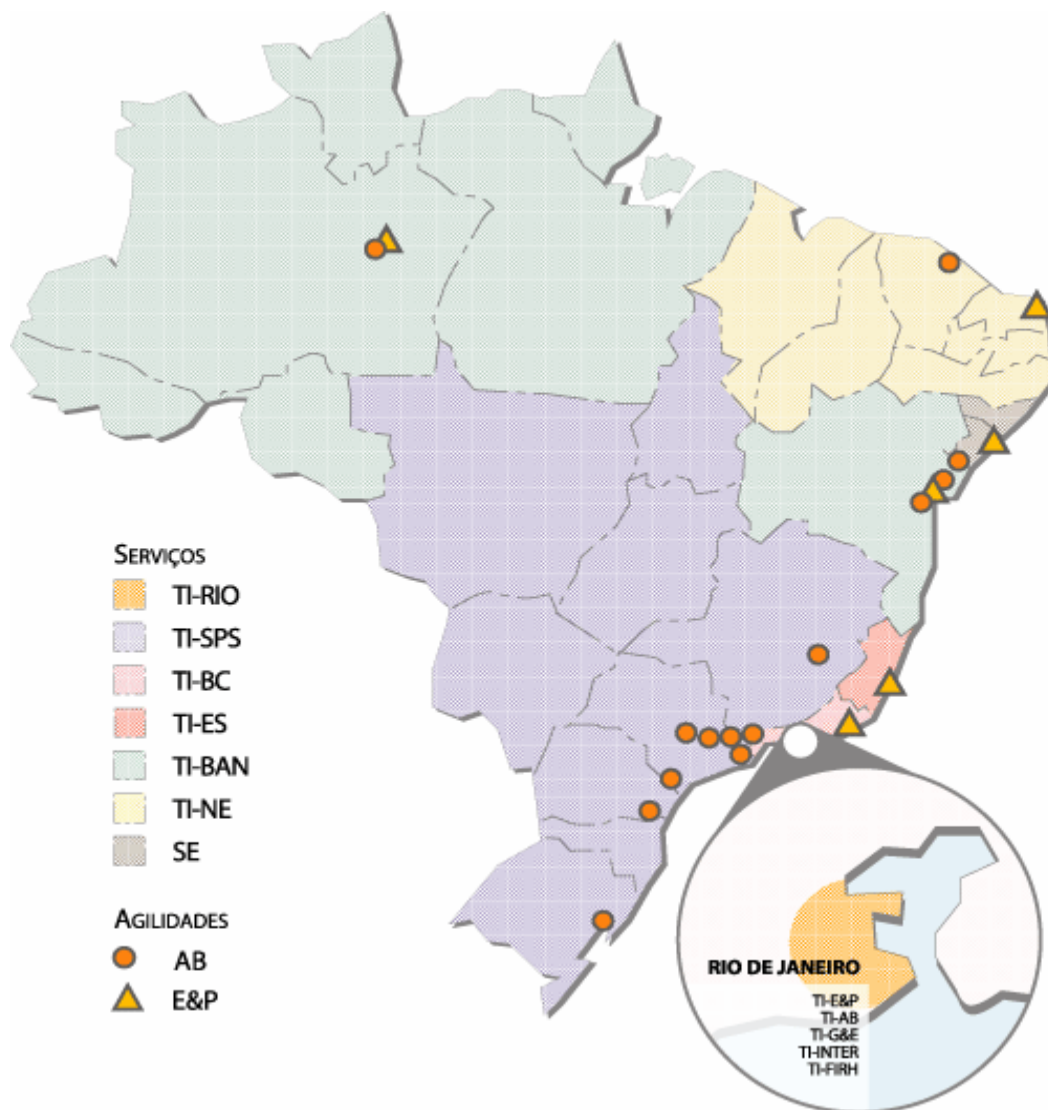
Na Figura 4 está reproduzido o organograma da área de Serviços da TI da PETROBRAS.



**Figura 4: Organograma de Serviços da área de Serviços da TI da PETROBRAS**

**Fonte: TI da Petrobras**

A Figura 5, na página seguinte, área de atuação geográfica da área de serviços, por regional e os locais onde há gerências de agilidade atuando junto ao seu cliente.



**Figura 5: Área de atuação geográfica da área de serviços**

**Fonte: TI da Petrobras**

### 1.1.3 O serviço de Atendimento ao usuário de TI

O serviço de Apoio ao Usuário (AU) da área de Serviços é o gestor do processo de atendimento ao usuário. As equipes de atendimento trabalham para encontrar a solução dos problemas de TI, no menor prazo possível, reduzindo ao mínimo os impactos no trabalho dos usuários. São realizados acordos com os clientes para o estabelecimento de níveis básicos de serviço. A Figura 6, a seguir, apresenta o fluxo de atendimento.



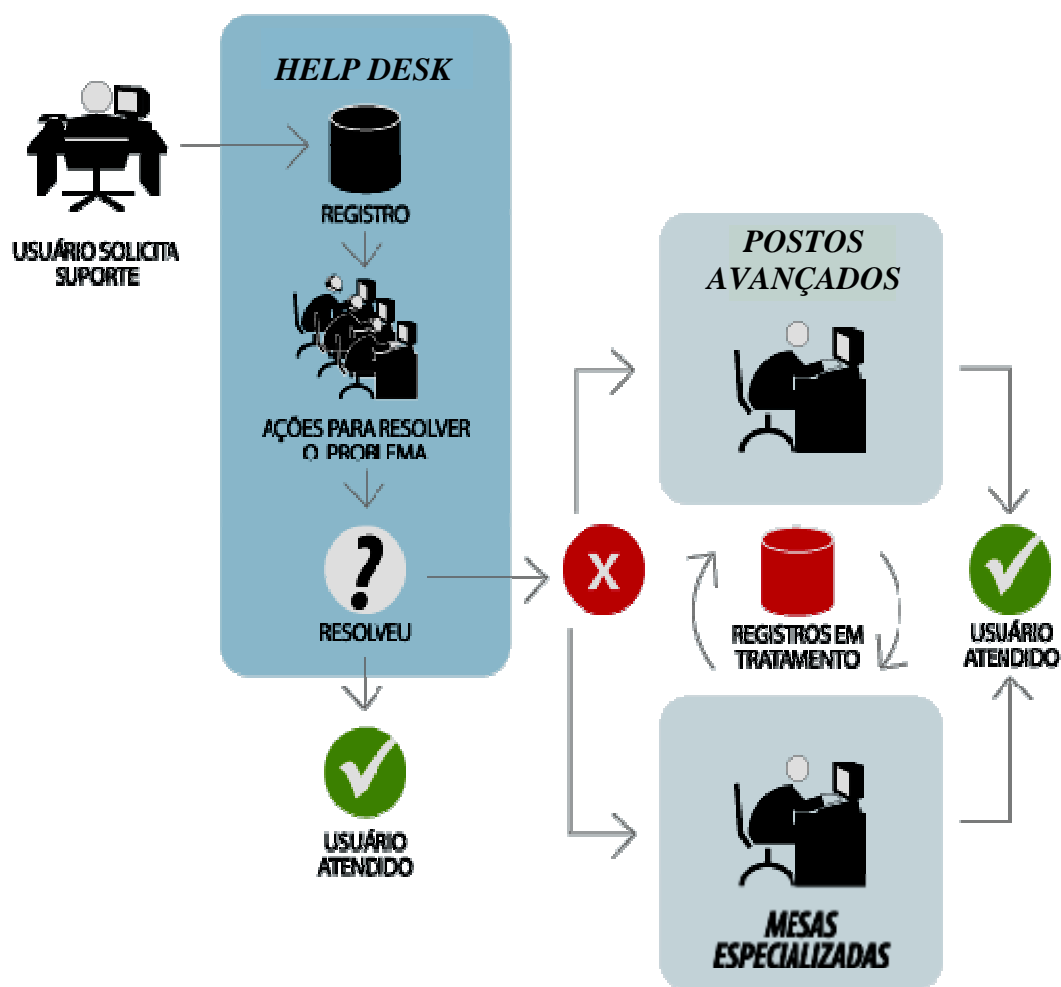


Figura 6: Fluxo de Atendimento

Fonte: TI da Petrobras

Os principais componentes da estrutura de atendimento ao usuário de TI da Petrobras, apresentados na figura anterior, têm a seguinte descrição básica:

- *Help-Desk* – É denominado atendimento de 1º nível. São equipes de atendimento remoto capaz de gerenciar, coordenar e resolver incidentes no menor tempo possível garantindo que nenhuma solicitação seja perdida, esquecida ou ignorada. É composta de técnicos de informática que recebem treinamento para resolver a maior parte dos atendimentos no primeiro contato. Quando se esgotam as alternativas para solução do atendimento executado pelo *Help-Desk*, o registro é direcionado para a equipe que

possui maior capacidade em solucionar aquele registro. A equipe poderá ser uma Mesa especializada (ME) ou Posto avançado (PA). Existem duas estruturas de Help Desks na TI um na cidade do Rio de Janeiro e outra em Macaé. Trabalham em regime de 24 x 7 para contemplar todos os clientes do sistema Petrobras, em especial os trabalham em Plataformas de Exploração & Produção, Navios e Refinarias.

- Posto Avançado (PA) – É denominado atendimento de 2º nível. São equipes de atendimento físico, ou seja, de apoio presencial. São alocados junto ao Cliente. Atendem quando o a solicitação não pode ser resolvida remotamente e não se trata de um caso que possa ser resolvido por uma mesa especializada. Um exemplo típico é um problema de hardware em um equipamento. Existem PAs espalhados por todo o país. Onde houver equipamentos de TI e usuários, teremos pelo menos um PA próximo.
- Mesa Especializada (ME) – É denominada atendimento de 3º nível. São equipes de atendimento especializado em um ou mais produtos. A TI possui MEs nas áreas de Apoio ao Usuário, Infra-estrutura, SAP R/3, Desenvolvimento e Agilidade (sistemas específicos dos clientes das áreas de negócio). As MEs estão distribuídas por todo país, de acordo com as necessidades dos clientes .

Os registros de atendimento são gerados sob a plataforma ARS – *Action Request System* na qual foi desenvolvido o principal sistema que controla todo o atendimento de TI. Denominado LIG-TI, este sistema é a interface disponível para os técnicos da TI possam manipular o registro desde a abertura, passando pelas equipes de atendimento (HD, ME e PA) até o seu encerramento. Estes registros gerados pelo ARS são, a base de dados na qual este estudo se baseia. São três bases de dados diferentes contendo os registros de atendimento, a saber:

- Base de Atendimento → contém os registros que se encontram em aberto.

- Base de Movimentação → registra todas movimentações entre as equipes de atendimento para cada chamado. Toda vez que um chamado troca de responsável ou de equipe é gerado um registro nesta base.
- Base do CIANS → é a base de dados que serve para apuração dos indicadores de desempenho das equipes de atendimento. Utilizada também para apuração dos indicadores de nível de serviço junto aos Clientes da TI. Contém apenas os chamados encerrados, consolidando todos os tempos e informações relevantes.

## 1.2 OBJETIVO

### 1.2.1 Objetivo Geral

O objetivo desta dissertação é analisar uma grande massa de registros de atendimento aos usuários de Tecnologia da Informação e extrair, a partir desta, conhecimento válido que possa ser utilizado para dar suporte à decisão na gestão do processo de apoio ao usuário.

### 1.2.2 Objetivos específicos

- Implementar o processo de KDD nas bases de dados de *help-desk* da Petrobras.
- Identificar relacionamentos entre atributos que compõem estas bases de dados.
- Analisar os resultados das relações entre os atributos.
- Gerar, a partir das informações obtidas, conhecimento que possa ser utilizado na gestão do *help-desk* da Petrobras.

### 1.3 ESTRUTURA DO TRABALHO

Neste capítulo foi descrito ambiente de onde foram extraídas as bases de dados que são objeto deste estudo

No capítulo 2 é feita uma revisão de literatura, expondo de forma resumida os principais conceitos que envolvem o processo de descoberta de conhecimentos em bases de dados, com maior ênfase na etapa de mineração de dados.

O capítulo 3 aborda o processo de descoberta de conhecimento empregado nas bases de *Help-Desk*, a técnica de mineração de dados que foi utilizada, a ferramenta escolhida e como foi desenvolvido todo o processo em cada uma de suas etapas, desde a preparação dos dados até a análise dos resultados encontrados.

No capítulo 4 são apresentadas as conclusões deste trabalho. Os resultados alcançados são comparados ao que era esperado bem, dificuldades encontradas e possibilidades de melhoria. Neste capítulo são feitas também alusões a possíveis trabalhos futuros que não foram contemplados neste estudo.

O Apêndice “A” apresenta um pequeno guia para gerar regras de associação utilizando a ferramenta Rattle.

O Apêndice “B” reúne as figuras com os resultados consolidados das regras de associação.

## 2 REVISÃO DA LITERATURA

Este capítulo descreve o processo de descoberta de conhecimento em bases de dados e suas principais etapas. Uma ênfase maior é dada para etapa de mineração de dados por ser nesta etapa onde efetivamente ocorre a descoberta de conhecimento, na forma de modelos e padrões. O detalhamento desta fase é apresentado através da descrição das principais tarefas e técnicas que podem ser utilizadas na mineração de dados.

### 2.1 O PROCESSO KDD – KNOWLEDGE DISCOVERY IN DATABASES

Descobrir significa deixar visível, mostrar, expor (AURÉLIO, 1999). Quando se fala em descobrimento indubitavelmente está se buscando o novo, o inesperado. O conhecimento, neste contexto, está relacionado ao mesmo tempo com a experiência e a informação. Descobrir conhecimento em base de dados é então a investigação minuciosa, em arquivos de dados armazenados em sistemas computacionais, com vistas a obter experiências e informações novas que a princípio estavam ocultas nestes arquivos de dados.

Segundo FAYYAD 1996, “o KDD - Knowledge Discovery in Databases - é um processo, de várias etapas, não trivial, interativo e iterativo, para identificação de padrões compreensíveis, válidos, novos e potencialmente úteis a partir de grandes conjuntos de dados”.

Como existe vasta literatura sobre o assunto e pela semelhança desta definição de KDD com a definição de mineração de dados, estes dois termos se confundem e muitas vezes são utilizados de forma indistinta. Esta visão foi assumida por RESENDE (2005) em sua obra para referenciar o processo de extrair conhecimento a partir de grandes bases de dados.

De fato, de acordo com HAN (2001), o termo mineração de dados ou *data mining* deveria ter um nome mais apropriado, indicando conhecimento através da mineração de dados. No entanto, por ser este um termo longo não é tão utilizado.

Mas, ainda segundo HAN (2001), a descoberta de conhecimento em bases de dados é um processo mais amplo que a mineração de dados. Envolve várias etapas onde a mineração de dados é parte deste processo, sendo uma destas etapas, como mostra a figura 7.

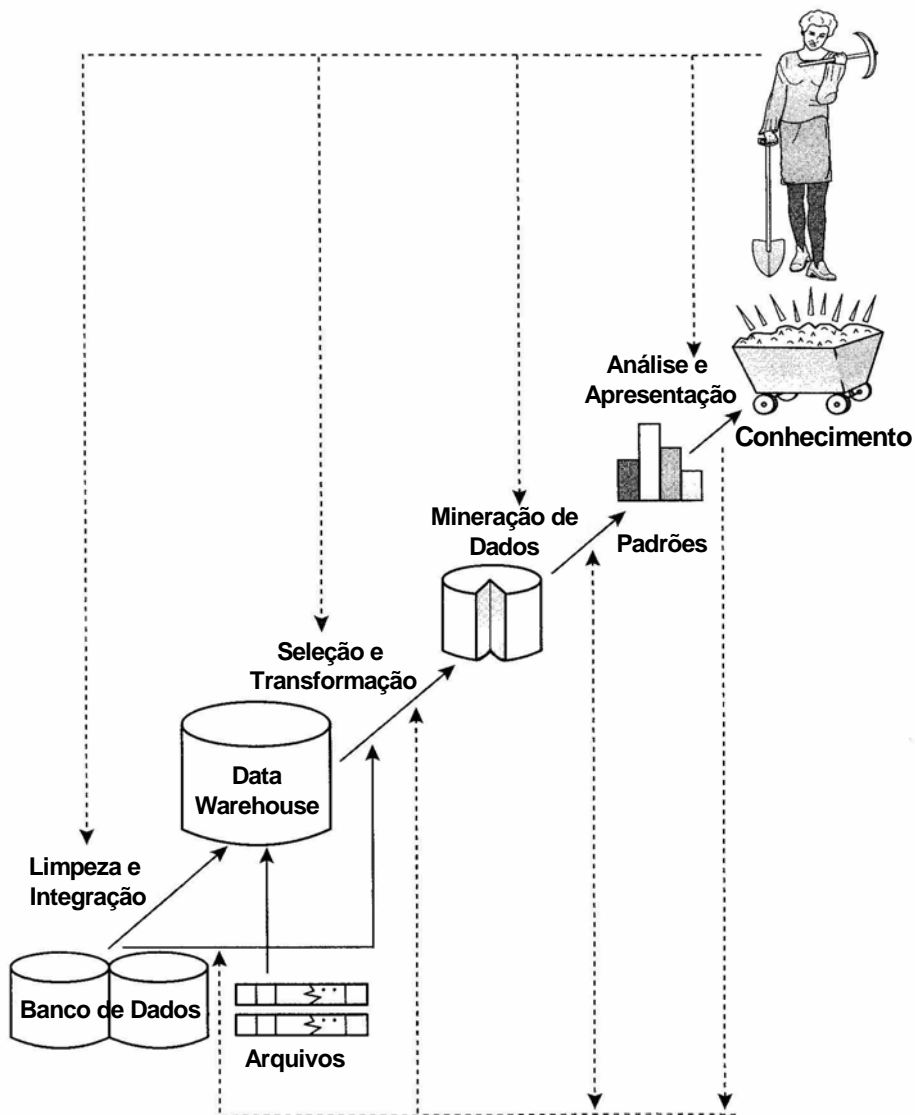


Figura 7: As etapas do processo de KDD  
Fonte: Adaptado de HAN, 2001

## 2.2 AS ETAPAS DO PROCESSO KDD

As principais etapas operacionais do processo de KDD são:

- Pré-processamento dos dados
- Mineração de dados
- Pós-processamento

A seguir, a descrição detalhada de cada uma destas etapas:

### 2.2.1 O Pré-processamento dos dados

O objetivo desta etapa é preparar os dados para que possam ser processados pelas ferramentas de mineração de dados na etapa seguinte do processo de KDD de modo a otimizarem-se os resultados. Esta etapa normalmente é a que demanda a maior parte do tempo do processo de descoberta de conhecimento e também pode se apresentar como sendo a mais árdua. O sucesso ou fracasso de um processo de descoberta de conhecimento pode, muitas vezes, estar associado ao tempo e aos critérios utilizados nesta etapa de preparação.

Na maioria dos casos, quando as bases de dados já estão organizadas em um *Data Warehouse*, a etapa de seu pré-processamento pode ser bastante facilitada. Em um *Data Warehouse* as várias bases de dados já foram reunidas e, de alguma forma, passaram por um processo de limpeza e transformação dos dados. Todavia, o fato de os dados não estarem no modelo de um *Data Warehouse* não impede que sejam analisados pelos algoritmos de mineração de dados. Significa, entretanto, que muito provavelmente será gasto um tempo maior na etapa de

preparação, principalmente nas tarefas de seleção e limpeza dos dados. Maiores esclarecimentos sobre *Data Warehouse* podem ser obtidos em BARBIERI (2001) e HAN (2001).

A etapa de Pré-processamento dos dados consiste das seguintes tarefas:

#### 2.2.1.1 Seleção dos dados

Selecionar dados significa identificar e segregar o que realmente será relevante para o processo de KDD, como, por exemplo, a escolha das bases de dados e seus respectivos registros e atributos.

Selecionar dados significa identificar e segregar o que realmente será relevante para o processo de KDD, como a escolha das bases de dados e seus respectivos registros e atributos. Assim, o primeiro passo é determinar a fonte primária de dados será utilizada. Diante da definição do problema a ser resolvido, escolhem-se os bancos de dados que serão alvos da pesquisa. Em muitos casos, as fontes de dados poderão estar em formatos diferentes, como planilhas eletrônicas, bancos de dados até mesmo em *Data Warehouse*. Para iniciar a etapa de mineração de dados, todas as fontes deverão estar reunidas em uma única base de dados ou arquivo. A maior diversidade das fontes de dados e dos formatos destes dados implicará em um maior esforço para reunir todas as fontes de dados.

Após unificação dos dados em uma base única, selecionam-se os registros dos arquivos. Cada registro de um arquivo representa um caso completo. Supondo a análise dos dados de *help-desk*, por exemplo, esta seria o conjunto de todos os atributos que descrevem um único pedido de atendimento.



Decidir que atributos serão investigados no processo de mineração de dados é também parte da tarefa de seleção. Os atributos são os valores assumidos nos diversos registros dos arquivos. Alguns atributos típicos de base de dados de *Help-desk*: hora em que um pedido de atendimento foi registrado, hora em que foi fechado, tempo total gasto no atendimento, nome do solicitante, seu telefone de contato etc. Ao analisar os atributos, tendo em vista o problema a ser resolvido, pode-se, por exemplo, não considerar necessário o atributo que contém o número do telefone do solicitante.

#### 2.2.1.2 Limpeza de dados

De acordo com GOLDSCHMIT (2005), esta tarefa consiste de ações sobre os dados selecionados visando garantir a qualidade destes quanto à completude, veracidade e integridade. Após a reunião dos dados de diferentes fontes, é provável que o conjunto de dados apresente registros duplicados, erros nos dados ou apresente dados ausentes. Até mesmo os dados coletados de uma única fonte também podem apresentar os mesmos problemas.

A limpeza dos dados busca eliminar três situações indesejáveis para o conjunto de dados que foram selecionados: dados faltantes, dados ruidosos e dados inconsistentes.

- Dados faltantes são os atributos que não possuem valor ou quando o valor dos mesmos está incompleto ou não detalhado. HAN (2001) elencou vários métodos para tratar os dados faltantes, tais como:
  - Ignorar o registro;
  - Preencher o valor faltante manualmente;

- Usar um valor constante global;
- Usar a média do atributo;
- Usar a média do atributo em todos os registros que estiverem na mesma situação;
- Usar o valor mais provável;

Todos os métodos apresentam vantagens e desvantagens. A natureza do atributo, a quantidade de registros e o número de faltantes serão determinantes para a escolha do método mais adequado.

- Dados ruidosos são aqueles que apresentam valores de atributos divergentes, errados ou fora do esperado. Podem ser denominados como outliers ou pontos fora da curva. Métodos da estatística e da computação são muito utilizados para identificar os ruídos nos dados. Após a identificação dos ruídos, são possíveis duas ações distintas: excluir todos os registros que apresentam algum tipo de ruído ou corrigir o erro atributo a atributo seja manualmente seja de forma automatizada.
- Dados inconsistentes são aqueles que por algum motivo apresentam valores que não podem ser considerados válidos em virtude de alguma característica intrínseca do dado. Neste caso os registros são eliminados ou ignorados.

### 2.2.1.3 Codificação dos dados

Esta tarefa é necessária para que os dados estejam em formato compatível com as ferramentas de mineração de dados que serão utilizadas e, dependendo do dado a ser tratado, para preservação da identificação direta do dado analisado. Segundo Goldschmit (2005), alguns algoritmos de mineração de dados necessitam que os dados sejam categóricos para que se obtenham os melhores resultados enquanto outros necessitam de dados no formato numérico.

### 2.2.1.4 Enriquecimento dos dados

A tarefa de enriquecimento dos dados consiste melhorar a informação contida nos registros dos bancos de dados através da criação de novos atributos a partir dos já existentes, agregando novas informações. A geração de totalizadores em variáveis numéricas, a criação de faixa ou classes de valores para atributos contínuos e a generalização de valores de atributos são exemplos de enriquecimento dos dados.

Essa fase de enriquecimento consiste em agregar aos dados existentes informações que contribuam ao KDD. Como se trata de informações que não estão na base de dados, porém são conhecidas e ratificadas por especialistas e analistas de dados, serão incorporadas ao processo como meta conhecimento.

## 2.2.2 Mineração de Dados

Esta é a principal etapa do processo de descoberta de conhecimento em bases de bases, pois é onde efetivamente ocorre o processamento dos dados e a geração dos resultados.

A literatura atual apresenta um grande número de definições para a mineração de dados. Algumas delas incluindo a mineração como um processo completo, outras posicionam a mineração como uma etapa do processo de KDD.

Segundo CARVALHO (2005), “é um conjunto de técnicas reunidas da Estatística e da Inteligência Artificial com objetivo específico de descobrir conhecimento novo que, por ventura, esteja escondido em grandes massas de dados armazenadas em bancos de dados empresariais”.

Para GROTH (1998), “é o processo de descoberta automática de informações”.

Segundo KREMER (1999), *Data Mining* é uma tecnologia usada para revelar informação estratégica escondida em grandes massas de dados.

Dependendo da abordagem, as definições de *Data Mining* têm maior ênfase na tecnologia, nas áreas de pesquisa da Estatística e da IA ou no processo de descoberta de conhecimento.

Assim, a mineração de dados está associada aplicação de técnicas e algoritmos de outras áreas do conhecimento para a análise de grandes volumes de dados com vistas à extração de padrões sobre estes dados. Para aplicação das técnicas e algoritmos, no entanto, se faz necessário definir a abordagem que será utilizada para a solução dos problemas sobre os dados analisados. Esta abordagem e o algoritmo ou técnica constituem as tarefas e técnicas da mineração de dados.

### **As tarefas e técnicas de Mineração de Dados**

As tarefas definem a finalidade a que se propõe a análise dos dados. Para execução destas tarefas são aplicadas algumas técnicas que podem ser específicas ou não para uma determinada tarefa. É importante destacar que as técnicas não envolvem apenas a tecnologia computacional em si, mas também a estatística, a inteligência artificial, modelos matemáticos, entre outros campos do conhecimento.

<b>TAREFAS</b>	<b>TÉCNICAS</b>
Classificação	Árvores de decisão, regressão, redes neurais etc
Estimação	Regressão e redes neurais
Previsão	Redes neurais, séries temporais, regressão, árvore de decisão, raciocínio baseado em casos
Análise de afinidades	Regras de associação
Análise de agrupamentos	Redes neurais, algoritmos genéticos e específicos

**Tabela 1 – Tarefas e Técnicas de Mineração de Dados**

Para elaboração deste trabalho foi utilizada a tarefa de análise de afinidades e a técnica regras de associação que serão descritas a seguir.

### **Análise de Afinidades**

A busca de padrões é uma das tarefas mais comuns na mineração de dados. Estes padrões são unidades de informação que se repetem seja na seqüência em que aparecem seja em sua estrutura de formação.

## **Regras de Associação**

As regras de associação procuram por associações interessantes ou pelas relações de correlação entre os atributos em grandes bancos de dados (HAN, 2001). Elas foram introduzidas por AGRAWAL (1993) através do algoritmo denominado apriori. Este algoritmo utiliza uma forma interativa na busca de itens freqüentes.

### **Definições**

Seja um conjunto de itens  $I = \{I_1, I_2, \dots, I_n\}$ .

Se  $X \subseteq I$ ,  $Y \subseteq I$ , então  $X$  e  $Y$  são chamados itensets. Seja  $D$  um conjunto de transações que contém itensets. Seja  $T$  cada transação, tal que  $T \subseteq I$ . Seja  $n$  o número total de registros em  $D$ .

Uma regra de associação é uma expressão na forma  $X \rightarrow Y$ , onde  $X$  e  $Y$  são itensets frequentes, sendo  $X \neq \emptyset$  e  $Y \neq \emptyset$ , onde  $X$  é denominado o antecedente da regra e  $Y$  o conseqüente.

### **Suporte de uma Regra**

O suporte de uma regra é o número de ocorrências da regra em relação ao número total de transações.

$$\text{Sup}(X \rightarrow Y) = \frac{|X \cup Y|}{n}$$

## Confiança

A confiança de uma regra é dada pela relação entre número de ocorrências da regra e o número de ocorrências do itenset antecedente da regra.

$$\text{Confiança}(X \rightarrow Y) = \frac{\text{Sup}(X \rightarrow Y)}{\text{Sup}(X)}$$

Quanto maior o valor da confiança, mais interessante é a regra. Ao aplicar os algoritmos para geração das regras de associação, os parâmetros de suporte e confiança mínimos devem ser configurados, limitando assim o espaço de busca da solução. A busca do valor ideal para suporte e confiança depende de vários fatores. Em muitos casos pode ser necessário executar o processo de geração de regras várias vezes, ajustando-se os parâmetros de forma que seja reduzido o número de regras geradas sem prejuízo do resultado que se pretende alcançar.

Assim, os parâmetros de suporte e confiança são denominados medidas interesse das regras de associação, pois auxiliam na determinação do grau de interesse da regra. Embora estas sejam as mais utilizadas pelas ferramentas de mineração de dados, outras medidas de interesse foram introduzidas posteriormente, como o *lift* e o *Rule Interest*.

### 2.2.3 O pós-processamento

Nesta etapa, o conhecimento obtido através da mineração de dados é estruturado em forma de relatórios, gráficos, diagramas ou qualquer outra que apresente o conhecimento adquirido.

No entanto, a extração e a apresentação do conhecimento obtido através da mineração de dados, por si só, não encerra o processo de descoberta de conhecimento. O conhecimento descoberto deve ser avaliado sob duas óticas distintas: quanto à validade dos resultados obtidos, tendo em vista a proposição inicial e quanto à qualidade das heurísticas geradas.

No primeiro caso, trata-se de validar a precisão e a acurácia do que é produzido. Poder-se-ia, por exemplo, inquirir: As regras de associação descobertas são compreensíveis? Qual a precisão das regras geradas? Do ponto de vista do analista especializado estas regras fazem sentido? De modo análogo, outras perguntas podem ser feitas para apurar a qualidade do conhecimento obtido através das demais técnicas de mineração de dados.

No segundo caso, estamos interessados em analisar os resultados que foram obtidos pelo processo de KDD à luz dos objetivos propostos e dos problemas identificados. As heurísticas resolvem os problemas? Têm aplicação prática?

Nos dois casos pode-se concluir que os resultados encontrados não foram satisfatórios e torna-se indispensável retornar à etapa de mineração de dados, revendo parâmetros ou mesmo as técnicas que foram utilizadas. Em alguns casos é imprescindível o retorno à etapa de preparação dos dados para que todo o processo seja refeito.

Além disso, é possível integrar conhecimentos obtidos por meio de diversas abordagens a fim de aprimorar a precisão e o desempenho do sistema. Após essa etapa pode-se decidir se o conhecimento obtido atendeu as expectativas e aos resultados propostos.



### 3 DESCOBRINDO CONHECIMENTO EM BASES DE HELP-DESK

Este capítulo apresenta o processo de descoberta de conhecimento e suas etapas a partir das bases de dados de *Help-Desk* da Petrobras.

#### 3.1 O PRÉ-PROCESSAMENTO DOS DADOS DE HELP-DESK

A etapa de preparação dos dados, fundamental para que estes possam ser utilizados de forma eficiente no processo de mineração de dados, normalmente se apresenta como sendo a mais árdua e a mais longa de todo o processo de descoberta de conhecimento. Neste trabalho não foi diferente.

Inicialmente, foi necessário gerar um arquivo texto a partir da base de dados CIANS, gerenciada pela ferramenta ARS. Esta ferramenta possui um formato de banco de dados próprio que guarda alguma compatibilidade com os bancos de dados ORACLE. Através de uma *query*, os administradores dos bancos de dados do ARS geraram um arquivo texto com todos os registros do CIANS, desde seu início até meados de setembro de 2006.

O arquivo texto gerado tinha tamanho aproximado de 10 gigabytes, com quatro milhões de linhas e 70 atributos. Cada linha do arquivo continha 1909 bytes. Devido a dificuldade para manipular arquivo tão grande, o mesmo foi dividido em partes para que pudesse ser melhor analisada sua estrutura.

Após esta análise, o arquivo foi importado para a ferramenta de banco de dados Visual FoxPro onde os dados foram selecionados, limpos e transformados.

Alguns atributos foram logo eliminados por serem repetitivos ou por não serem úteis em um processo de geração de regras de associação, como por exemplo, nome do técnico que atendeu o chamado, nome do gerente, chave de identificação, sinais de controle do registro, nome de usuário etc.

### 3.1.1 A seleção dos atributos

A seleção dos atributos do arquivo CIANS para busca de relações de afinidades foi baseada nas entrevistas com os gestores do processo de apoio ao usuário. Inicialmente, dos 70 viáveis do arquivo CIANS, foram selecionados 10. Posteriormente, foram priorizados para este estudo seis atributos. Dado o elevado número de registros do arquivo CIANS, um dos critérios para seleção dos atributos foi o número de valores diferentes que o mesmo poderia assumir. Este critério visou reduzir o número de regras que seriam geradas uma vez que este é um dos maiores problemas para análise dos resultados de um processo de mineração de dados.

Desta forma, foram selecionados os atributos:

- Categoria
- Tipo de grupo
- Origem do registro
- Tipo de registro
- Data e hora da criação do registro
- Tempo total para a solução em minutos

O atributo data de criação do registro foi transformado para que pudesse ser utilizado no processo de mineração. Foram criados dois novos atributos a partir deste, denominados PERIODO e MESANO.

O atributo “Tempo total” também foi transformado, tendo como balizador o acordo de nível de serviço que a TI assina com seus clientes internos.

A seguir, uma breve descrição de cada um dos atributos e dos valores que cada um pode assumir.

CATEGORIA – corresponde a principal classe de registros. Existem subgrupos para cada uma delas que não foram tratadas neste estudo. Pode assumir nove valores distintos, como segue:

- Lotus Notes - Serviços referentes às aplicações ditas "de prateleira" (sem customização), ou customizadas para a Petrobras, ou desenvolvidas internamente, que operam na plataforma Notes.
- Aplicação - Serviços referentes às demais aplicações em geral, ditas "de prateleira" (sem customização).
- Sistemas - Serviços referentes a aplicações customizadas para a Petrobras, ou desenvolvidas internamente, as quais operem em qualquer plataforma, excetuando-se Notes.
- Hardware - Serviços referentes a problemas físicos em equipamentos de informática, excetuando-se os de infra-estrutura ou de rede.
- Impressão - Serviços referentes a aspectos de impressão.
- Segurança – Serviços que envolvam assuntos referentes à segurança da informação, ao acesso a informações, criação de chaves, assuntos relacionados com senhas, permissões de acesso etc.
- Infra-estrutura - Serviços referentes à administração de servidores e recursos de rede.

- SAP – Serviços relacionados o sistema ERP da empresa SAP, denominado R/3. Este sistema foi implantado em toda a Petrobras, contemplando todos os módulos, em outubro de 2004.
- Serviços de TI - Serviços realizados que não estejam incluídos em alguma das outras categorias

TIPO DE GRUPO – são grupos que compõe as equipes de atendimento. As funções de cada um destes grupos foram detalhadas na Introdução deste trabalho, capítulo 1. São três grupos:

- HD – *Help-desk*, equipes de atendimento remoto, geralmente através de telefone. É denominado atendimento de primeiro nível.
- PA – Posto Avançado, equipes de atendimento local, junto ao usuário. É o atendimento de segundo nível.
- ME – Mesa Especializada, equipes especializadas em um determinado sistema ou aplicação. Normalmente atende remotamente por telefone. É chamado de atendimento de terceiro nível;

ORIGEM DO REGISTRO – identifica a procedência do registro, ou seja, o meio que foi utilizado para a criação do registro. São oito origens diversas:

- Web – São as solicitações feitas através da intranet da Companhia. Qualquer usuário pode abrir sua própria solicitação sem auxílio do *Help-Desk*. Os registros de cobertura também podem ter esta origem porque os técnicos de apoio local (PA) e mesa especializada (ME), ao receberem solicitações diretas do usuário, utilizam deste meio para registrar o atendimento. Este sistema é específico para abrir registros já atendidos (há uma página Web para técnicos de TI e outra para os usuários) e os registros

gerados nessa página são visualizados automaticamente no sistema de workflow de *Help-Desk* TI, cuja plataforma é o ARS.

- ARS – São registros criados pela ferramenta.
- LIG-881 - Sistema de workflow da área de telecomunicações. Muitas vezes um problema de informática está associado a um atendimento da área de telecomunicações e vice-versa (problema em cabos de rede, hubs etc). Assim, os registros que foram criados pela área de telecomunicações e que são repassados para área de TI, apresentam esta origem.
- Aprovação Notes – São registros criados pelo usuário através de um sistema desenvolvido na plataforma Lotus Notes. Este sistema possui um fluxo de trabalho para garantir que houve aprovação gerencial. A solicitação feita pelo usuário só irá gerar um registro no ARS se o gestor aprovar a demanda. Estes registros são então designados automaticamente pelo sistema para as equipes de acordo com a sua natureza. São exemplos de solicitações com essa origem: criação de chave, permissão em diretórios, aumento de cota de correio etc.
- Telefone – são as solicitações originadas a partir de ligações para o *Help-Desk*. O usuário solicita atendimento através do ramal do apoio da TI. Este ramal é único para toda a Petrobras.
- E-mail – São solicitações feitas através de e-mail. A TI possui um endereço eletrônico como alternativa de contato com o usuário.
- Formulário – Serviços que possuem em seu fluxo de atendimento documentos que necessitam de assinatura. A TI só atenderá a solicitação mediante documentação devidamente preenchida e assinada.

- Tivoli - Ferramenta que cria registros no ARS na forma de alarmes. São avisos de possíveis problemas nos servidores de rede. Tem designação automática para as equipes de Infra-estrutura (MEs).

TIPO DE CHAMADO – Este atributo identifica o registro de acordo com a natureza do problema. Apresenta sete valores diferentes:

- Incidente – é qualquer evento que não faz parte da operação normal de um serviço, causando sua interrupção ou reduzindo sua qualidade.
- Dúvida - esclarecimento de alguma dúvida ou questionamento do solicitante.
- Solicitação – quando o usuário solicita algum serviço (Ex. instalação de um aplicativo).
- Tarefa – são as solicitações de serviços entre as equipes de TI, cuja origem é um registro onde a conclusão depende da intervenção de outras equipes. É criada sempre a partir de um registro do tipo “Solicitação”.
- Registro de Cobertura – compreende àquelas situações em que um técnico encerra um atendimento sem que tenha havido uma formalização prévia do solicitante com o Help Desk da TI; ou as situações em que o técnico está prestando algum atendimento em campo e é solicitado diretamente pelo usuário; ou quando o sistema de *Help-desk* fica indisponível. Este tipo de chamado é criado com a situação já definida como "Resolvido", diretamente na ferramenta do *Help-desk*. O técnico da TI deverá informar o tempo que foi gasto no atendimento.

DATA E HORA DA CRIAÇÃO DO REGISTRO – o campo data/hora na base de dados original do CIANS é composto de um único número inteiro que representa a quantidade em minutos desde a 0h de 01/01/1970 até a data e hora em que o registro foi aberto.

TEMPO TOTAL DE SOLUÇÃO – representa o tempo total em minutos que foi gasto, desde a criação do registro até o seu término.

### 3.1.2 Limpeza e seleção dos dados

As ações abaixo promoveram a limpeza nos dados do CIANS.

- Foram eliminados todos os registros que continham erros ou que não apresentavam informações completas.
- Ajustados os campos que continham informações semelhantes, como “sistema” e “sistemas”, “infra-estrutura” e “infra estrutura” etc. Diversas situações semelhantes foram encontradas e ajustadas, através de programação.
- Foram também segregados os registros que se referiam às opiniões do cliente e aqueles que foram cancelados pelo usuário. Apenas os registros referentes aos chamados encerrados permaneceram no arquivo.
- Somente os registros de janeiro de 2005 a agosto de 2006 permaneceram no arquivo CIANS. Os demais foram apagados por não serem representativos para este estudo.

### 3.1.3 Transformando os dados – codificação e enriquecimento

Nesta fase de preparação dos dados alguns campos serão codificados ou transformados de forma a tornar viável ou facilitar a extração de padrões. A seguir, estão listadas as ações efetuadas para transformação dos dados do arquivo CIANS.

- O número inteiro que representava a data e hora de criação do registro foi convertido para data e hora, nos formatos DD/MM/AAAA e HH:MM. Um programa na linguagem Visual FoxPro foi escrito com esta finalidade. A partir deste atributo que

foi convertido, dois novos foram criados, com objetivo de facilitar o processo de mineração de dados. São eles:

**Período** → contém o período do dia em que o registro foi aberto, podendo assumir os valores: M – manhã, T – tarde, N – Noite, O – Madrugada.

**Mesano** → representa o mês e o ano em que o registro foi aberto. Pode assumir os valores de Jan/05 até Ago/06.

- Os campos relativos à Origem do chamado, Grupo de atendimento e Tipo de Registro eram numéricos na base original. Como as ferramentas de mineração de dados de um modo geral, trabalham com campos categóricos para geração de regras de associação, todos os campos foram convertidos de numéricos para categóricos, através de codificação e programação.
- O atributo “tempo total em minutos” foi convertido para o formato categórico, refletindo o acordo de nível de serviço da TI com seus clientes, podendo assumir os seguintes valores:

1h – para registros encerrados em até 1 hora de atendimento

9h – para registros encerrados com mais de 1h e até 9h de atendimento

27h – para registros encerrados com mais de 9h e até 27h de atendimento

54h – para registros encerrados com mais de 27h e até 54h de atendimento

99h – para registros encerrados com mais de 54h e até 99h de atendimento

100h – para registros com 100 ou mais horas de atendimento.

- Criados dois arquivos de dados, separados por data. Um deles contendo todo ano de 2005 e outro com os meses de janeiro a agosto de 2006, de forma que possam ser



comparadas as informações estatísticas levantadas e as regras geradas. Os demais registros que não estavam compreendidos entre estas datas foram eliminados.

A figura a seguir resume todo o processo de limpeza e transformação dos dados desde a criação do registro até a geração dos arquivos usados no processo de mineração de dados.

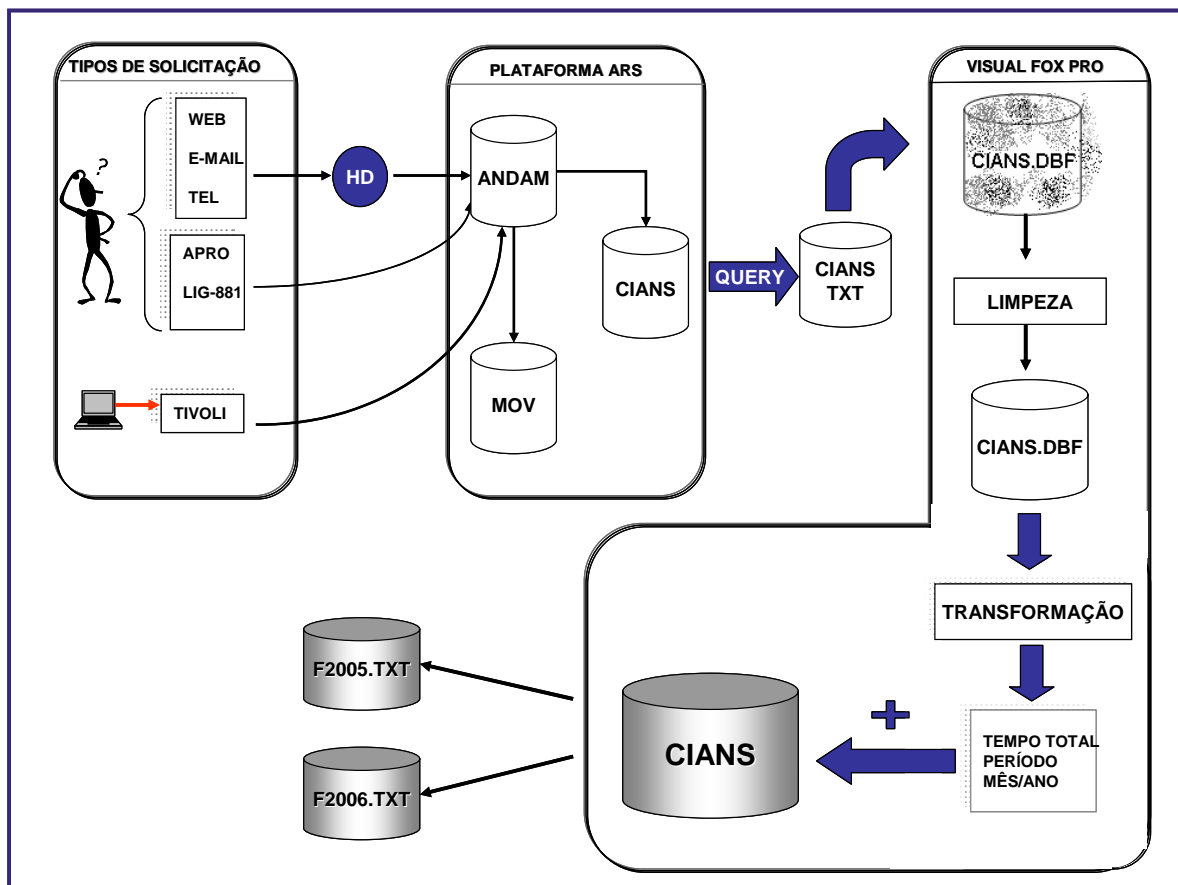


Figura 10 – Pré-processamento dos dados de Help Desk

### 3.2 ANÁLISE EXPLORATÓRIA DOS DADOS

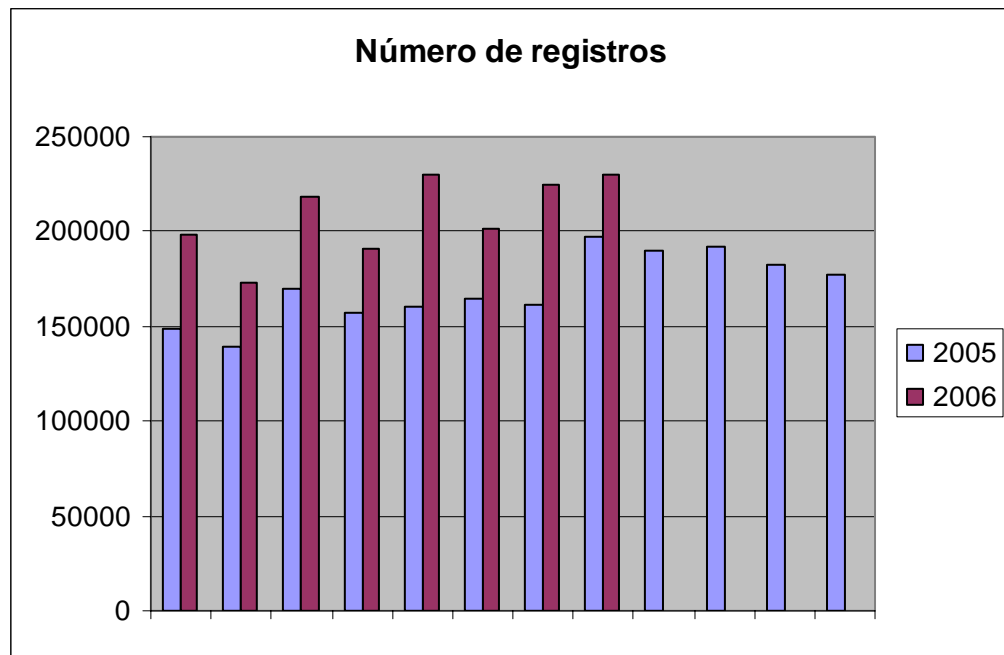
Após a seleção, limpeza e transformação dos dados, o arquivo principal do CIANS estava pronto para ser usado no processo de mineração. No entanto, dado que esta é a primeira vez que estes registros são analisados como um todo e, visando obter um maior conhecimento dos dados, optou-se por fazer algumas estatísticas dos atributos antes de se iniciar o processo de mineração. Esta etapa contribuiu sobremaneira para a análise e interpretação das regras de associação que serão geradas.

Nesta seção são apresentadas estatísticas dos dados analisados como o número total de registros e as tabelas de frequência para cada atributo selecionado. As tabelas são seguidas de comentários que procuram explicar de forma sucinta as variações de um ano para o outro.

#### Número total de Registros

MESES	2005		2006		Diferença 05->06
	Frequência	%	Frequência	%	
Janeiro	148.933	7,3	198.327	11,9	33,2
Fevereiro	139.337	6,8	173.131	10,4	24,3
Março	169.995	8,3	218.516	13,1	28,5
Abril	156.972	7,7	190.412	11,4	21,3
Mai	160.766	7,9	229.979	13,8	43,1
Junho	165.009	8,1	201.033	12,1	21,8
Julho	161.571	7,9	225.130	13,5	39,3
Agosto	196.826	9,6	229.555	13,8	16,6
Setembro	189.959	9,3			
Outubro	192.291	9,4			
Novembro	182.963	9,0			
Dezembro	176.913	8,7			
Total	2.041.535	100,0	1.666.083	100,0	28,5%

Tabela 2 – Número total de registros do CIANS



**Figura 11 – Volume de registros – Comparativo 2005-2006**

- Há um substancial crescimento do número de registros. O aumento do número de equipamentos (notebooks, microcomputadores, impressoras e estações RISC) e o crescimento do número de usuários, em função do crescimento da força de trabalho da empresa, justificam o fato.
- Os dados do ano de 2006 contemplam somente o mês de agosto. Até este mês, não foi possível apurar um fator sazonal determinante. Apenas o mês de fevereiro, por ter um número menor de dias, explica a redução no número de registros. Este evento ocorreu também em 2005.

**Atributo: CATEGORIA**

CATEGORIA	2005		2006		Dif 05->06
	Freqüência	%	Freqüência	%	
APLICACAO	337.009	16,5	262.008	15,7	-0,8
HARDWARE	250.324	12,3	197.432	11,9	-0,4
IMPRESSAO	73.429	3,6	54.404	3,3	-0,3
INFRA-ESTRUTURA	130.388	6,4	118.616	7,1	0,7
LOTUS NOTES	221.916	10,9	205.644	12,3	1,5
SAP	266.994	13,1	150.754	9,0	-4,0
SEGURANCA	170.444	8,3	185.355	11,1	2,8
SERVICOS DE TI	465.940	22,8	368.918	22,1	-0,7
SISTEMAS	125.091	6,1	122.952	7,4	1,3
<b>Total</b>	<b>2.041.535</b>	<b>100,0</b>	<b>1.666.083</b>	<b>100,0</b>	

**Tabela 3 – Distribuição de Freqüência – Atributo Categoria**

- O crescimento dos registros na categoria Segurança retrata um maior rigor das normas e procedimentos de TI quanto à Segurança de Informação e também, ao crescimento do número de usuários, pois nesta categoria estão os registros de criação de chave, senhas e permissões de acesso.
- Decréscimo significativo dos registros da categoria SAP R/3. Reflete a estabilização do sistema, tendo atingido um grau de maturidade de uso pelos usuários uma vez que sua implantação ocorreu em outubro de 2004.
- O correio eletrônico e suas aplicações têm seu crescimento associado ao aumento do número de usuários, a maior utilização desta ferramenta no dia-a-dia para recebimento e envio de mensagens, a uma maior demanda por sistemas desenvolvidos nesta plataforma e ao crescente número de aplicativos que usam a base de dados do correio como base de dados de usuários.
- A diminuição no percentual de serviços TI é positiva e indica que há uma melhor categorização dos registros por parte do *Help-Desk*.

**Atributo: TIPO DE REGISTRO**

TIPO DE REGISTRO	2005		2006		Dif 05->06
	Frequência	%	Frequência	%	
Cobertura	126.264	6,2	104.289	6,3	0,1
Dúvida	17.063	0,8	11.753	0,7	-0,1
Incidente	1.487.174	72,8	1.008.423	60,5	-12,3
Solicitação	339.680	16,6	451.942	27,1	10,5
Tarefa	71.354	3,5	89.676	5,4	1,9
<b>Total</b>	<b>2.041.535</b>	<b>100,0</b>	<b>1.666.083</b>	<b>100,0</b>	

**Tabela 4 - Distribuição de Frequência – Atributo Tipo de Registro**

- Em função do início da implantação da metodologia ITIL, o *Help-desk* foi orientado a categorizar alguns registros como Solicitação que anteriormente eram categorizados como Incidente. Isso explica a queda no número de Incidentes e o aumento do número de Solicitações.
- Os registros de Solicitação têm seu crescimento associado também ao aumento da utilização de meios eletrônicos pelos usuários para criação de registros, como por exemplo, os registros cuja origem é Aprovação Notes ou Web.
- O aumento dos registros Tipo Tarefa é explicado principalmente pelo aumento do número de solicitações (de onde é derivado), pela melhor definição dos papéis de cada equipe de atendimento e pela maior divisão das atividades entre estas equipes.

### Atributo: PERÍODO

Período	2005		2006		Dif 05->06
	Frequência	%	Frequência	%	
Manhã	1.025.278	50,2	828.107	49,7	-0,5
Tarde	884.882	43,3	716.078	43,0	-0,3
Noite	97.638	4,8	89.875	5,4	0,6
Madrugada	33.737	1,7	32.023	1,9	0,2
<b>Total</b>	<b>2.041.535</b>	<b>100,0</b>	<b>1.666.083</b>	<b>100,0</b>	

Tabela 5 - Distribuição de Frequência – Atributo Período

- Discreto crescimento dos registros abertos durante a noite e na madrugada em detrimento dos registros abertos no período da manhã e da tarde. É possível que o aumento do número de usuários que estejam trabalhando em plataformas, navios e refinarias, em turnos de revezamento tenha influenciado este resultado.
- O aumento do número de servidores de rede que passam a ser monitorados pela ferramenta Tivoli pode ser outro fator que justifique este crescimento.
- Pode estar ocorrendo um deslocamento do horário inicial da força de trabalho pela manhã, retardando seu início, mas prolongando-o à noite.

### Atributo: TIPO DE GRUPO

Tipo Grupo	2005		2006		Dif 05->06
	Frequência	%	Frequência	%	
HD	1.032.726	50,6	841.787	50,6	0,0
ME	528.886	25,9	458.965	27,5	1,6
PA	479.923	23,5	365.331	21,9	-1,6
<b>Total</b>	<b>2.041.535</b>	<b>100,0</b>	<b>1.666.083</b>	<b>100,0</b>	

Tabela 6 - Distribuição de Frequência – Atributo Período

- As MEs vêm crescendo no atendimento aos registros quando comparadas aos PAs possivelmente devido a dois fatores: a complexidade dos registros, que requerem especialistas nos produtos e a maior confiabilidade nos equipamentos, comprovada pela diminuição dos registros categorizados como Hardware.

- Por outro lado, registros mais complexos exigem mais tempo para serem resolvidos e provavelmente o tempo total de atendimento será afetado por esta mudança.

**Atributo: ORIGEM DO REGISTRO**

Origem	2005		2006		Dif 05->06
	Frequência	%	Frequência	%	
Aprovação Notes	64.465	3,2	83.327	5,0	1,8
ARS	2.128	0,1	3.519	0,2	0,1
e-mail	34.057	1,7	41.234	2,5	0,8
Formulário	19	0,0	79	0,0	0,0
LIG-881	1.278	0,1	668	0,0	0,0
Telefone	1.465.804	71,8	1.101.108	66,1	-5,7
Tivoli	67.540	3,3	79.874	4,8	1,5
WEB	406.244	19,9	356.274	21,4	1,5
<b>Total</b>	<b>2.041.535</b>	<b>100,0</b>	<b>1.666.083</b>	<b>100,0</b>	<b>-</b>

**Tabela 7 - Distribuição de Frequência – Atributo Origem do Registro**

- Apesar de o telefone ser ainda o meio mais utilizado para se abrir um registro, há expressiva mudança de comportamento por parte dos usuários que passaram a utilizar mais outros meios.
- Aumento expressivo dos registros abertos pelos usuários através de e-mail, da Web e da Aprovação Notes, comprovando que eles estão utilizando mais os meios eletrônicos para comunicação com o *Help-desk* da TI.
- Os registros abertos automaticamente pelos sistemas de controle comprovam o aumento dos registros na Categoria infra-estrutura.
- Queda do número de registros abertos pela área de Telecomunicações que são repassados para a TI. Este dado indica uma melhora na identificação inicial do problema.

## Atributo: TEMPO TOTAL DE ATENDIMENTO

Tempo Total de Atendimento	2005		2006		Dif 05->06
	Frequência	%	Frequência	%	
1h	115.524	5,7	90.502	5,4	-0,2
9h	32.041	1,6	12.601	0,8	-0,8
27h	366.513	18,0	297.601	17,9	-0,1
54h	589.948	28,9	495.280	29,7	0,8
99h	653.774	32,0	460.414	27,6	-4,4
100h	283.735	13,9	309.685	18,6	4,7
<b>Total</b>	<b>2.041.535</b>	<b>100,0</b>	<b>1.666.083</b>	<b>100,0</b>	

Tabela 8 - Distribuição de Frequência – Atributo Tempo Total de atendimento

- O aumento dos tempos para conclusão dos registros pode estar associado a diversos fatores, como: maior complexidade dos problemas (categorias Segurança, Infra-estrutura e Sistemas); crescimento dos registros abertos à noite e na madrugada.
- Outro componente importante que pode afetar diretamente o tempo de atendimento é a disponibilidade do usuário. Os usuários que trabalham em turnos de revezamento têm pouca disponibilidade durante o trabalho e estão ausentes durante o período de folga.

### 3.3 A SELEÇÃO DA FERRAMENTA PARA MINERAÇÃO DOS DADOS

Há um número considerável de ferramentas para executar tarefas de mineração de dados no mercado. Especificamente para gerar regras de associação, existe um grande número softwares disponíveis, sejam eles de uso livre ou não. Foram testadas algumas ferramentas de uso livre ou de demonstração, tendo-se como base as seguintes premissas:

- Permitir a importação de arquivos tipo texto
- Capacidade de tratar arquivos com muitas observações (casos)



O software WIZRULE foi desenvolvido pela empresa Wizsoft de Israel. Tem custo relativamente baixo por ser voltado essencialmente para geração de regras de associação. Sua interface é muito simples, intuitiva e fácil de usar. De todos os softwares analisados este é o de aprendizado mais rápido. Suas regras são geradas no formato IF ... THEN, onde são indicadas também a probabilidade de ocorrência da regra, o número de vezes que ocorreu e os registros onde a regra não é válida. É utilizado por algumas empresas para a auditoria em bases de dados, pois o mesmo aponta claramente os desvios nos dados em relação às regras de associação. O formato de apresentação dos resultados e a forma de interação com o usuário permitem a utilização produto mesmo por pessoas que não sejam especialistas em mineração de dados. O demo do produto é limitado a 1000 observações, tornando-o inviável para este trabalho.

O software WEKA (*Waikato Environment for Knowledge Analysis*) é bastante conhecido nos meios acadêmicos e muito utilizado. Foi desenvolvido na linguagem Java pela Universidade Waikato – Nova Zelândia. Bastante completo, vem sendo aprimorado através de contribuições vindas de diversos autores. Apresenta muitos recursos para a análise de dados e para a mineração de dados, contemplando algoritmos diferentes para execução de uma mesma tarefa. Para que os dados sejam manipulados eles devem estar no formato ARFF, sendo este um arquivo ASCII com as definições dos atributos e seus valores. Seus maiores problemas estão na performance baixa da plataforma Java e no arquivo de entrada que deve ter um formato próprio. Há também um alto consumo de memória quando se está utilizando arquivos grandes. O WEKA pode ser utilizado através de linha de comando ou através de sua interface gráfica, embora esta não seja tão simples de usar como as demais ferramentas analisadas neste trabalho. Os resultados podem ser apresentados de forma gráfica, facilitando o trabalho de análise.

TANAGRA é um software de uso livre para exploração de dados, aprendizado de máquina e mineração de dados, desenvolvido na linguagem Delphi por Ricco Rakotomalala da Universidade de Lumière Lyon 2 - França. Possui uma excelente interface gráfica, de uso intuitivo e fácil aprendizado. Tem funções para trabalhar com *clustering*, estatística não-paramétrica, regressão, regras de associação, entre outras. Não pode ser utilizado neste trabalho devido ao fato de a sua estrutura de dados para geração de regras de associação estar limitada a 250.000 observações. Mesmo quando utilizado com valores próximos a este há um grande consumo de memória e a performance cai consideravelmente. Esta limitação foi confirmada pelo criador da aplicação.

O Rattle (*R Analytical Tool To Learn Easily*) é um pacote para ser utilizado a partir da linguagem estatística R. Apresenta uma interface gráfica que facilita a utilização da linguagem R para análise, manipulação e mineração de dados. É uma ferramenta desenvolvida recentemente e ainda apresenta alguns problemas de operação. No entanto, tem muitos recursos e pode manipular arquivos grandes com desempenho satisfatório. Maiores informações sobre o Rattle podem ser obtidas no Apêndice A deste trabalho.

### 3.3.1 A ferramenta escolhida

O Rattle foi selecionado como ferramenta para este trabalho por atender as premissas estabelecidas, ou seja:

- Permite a leitura direta de arquivos no formato texto (TXT), não sendo necessário nenhum processo de conversão;

- Consegue processar arquivos com muitas observações (casos), desde que se tenha memória RAM disponível.

### 3.4 ANÁLISE DOS RESULTADOS ENCONTRADOS

O Rattle utiliza o algoritmo Apriori para identificação de itens frequentes e a geração de regras de associação. Várias combinações possíveis entre os atributos escolhidos foram testadas. São apresentadas somente as regras que apresentaram os valores de confiança mais altos. Foram considerados todos os itens de todos os atributos, exceto as origens de registros relacionadas ao Formulário ao ARS, pois estão em desuso e são pouco representativas na base de dados.

Os resultados e os comentários são apresentados a seguir.

## Categoria X Origem

2005				
lhs		rhs	support	confidence
{Categoria=SERVICOS DE TI}	=>	{Origem=TEL}	0,2016	0,8843
{Categoria=SAP}	=>	{Origem=TEL}	0,1153	0,8803
{Origem=TIV}	=>	{Categoria=INFRA-ESTRUTURA}	0,0278	0,8385
{Categoria=LOTUS NOTES}	=>	{Origem=TEL}	0,0887	0,8157
{Categoria=IMPRESSAO}	=>	{Origem=TEL}	0,0280	0,7763
{Categoria=APLICACAO}	=>	{Origem=TEL}	0,1168	0,7073
{Categoria=SISTEMAS}	=>	{Origem=TEL}	0,0431	0,7044
{Origem=APO}	=>	{Categoria=SEGURANCA}	0,0190	0,6028
{Origem=EMA}	=>	{Categoria=SEGURANCA}	0,0088	0,5284
{Categoria=HARDWARE}	=>	{Origem=TEL}	0,0636	0,5183
{Categoria=SEGURANCA}	=>	{Origem=TEL}	0,0405	0,4853
{Categoria=HARDWARE}	=>	{Origem=WEB}	0,0576	0,4700
{Categoria=INFRA-ESTRUTURA}	=>	{Origem=TIV}	0,0278	0,4343
2006				
lhs		rhs	support	confidence
{Categoria=SERVICOS DE TI}	=>	{Origem=TEL}	0.1889	0.8530
{Origem=TIV}	=>	{Categoria=INFRA-ESTRUTURA}	0.0401	0.8346
{Categoria=SAP}	=>	{Origem=TEL}	0.0755	0.8342
{Categoria=LOTUS NOTES}	=>	{Origem=TEL}	0.0979	0.7940
{Categoria=IMPRESSAO}	=>	{Origem=TEL}	0.0253	0.7731
{Origem=FOR}	=>	{Categoria=SEGURANCA}	0.0000	0.7037
{Categoria=SISTEMAS}	=>	{Origem=TEL}	0.0497	0.6730
{Categoria=APLICACAO}	=>	{Origem=TEL}	0.1052	0.6679
{Origem=APO}	=>	{Categoria=SEGURANCA}	0.6608	0.6608
{Origem=EMA}	=>	{Categoria=SEGURANCA}	0.0320	0.6406
{Categoria=INFRA-ESTRUTURA}	=>	{Origem=TIV}	0.0013	0.6052
{Categoria=HARDWARE}	=>	{Origem=WEB}	0.0145	0.5851
{Categoria=SEGURANCA}	=>	{Origem=TEL}	0.0401	0.5626
{Categoria=HARDWARE}	=>	{Origem=TEL}	0.0623	0.5263

**Tabela 9 – Análise das Regras – Atributos Categoria X Origem**

- Os registros originários através de e-mail ou da Aprovação Notes têm como predominante a Categoria Segurança.
- Contrariando o esperado, os registros da Categoria Hardware têm como origem preponderante a Web.
- O telefone ainda é o meio mais utilizado para a maior parte das Categorias de registros.
- Não há mudanças significativas nas regras de 2005 para 2006.

- A principal origem dos registros relacionados à infra-estrutura é devido ao controle e monitoramento efetuado pela ferramenta Tivoli, sendo este um resultado esperado.

### **Categoria X Tempo Total**

<b>2005</b>			
lhs	rhs	support	confidence
{Categoria=SERVICOS DE TI}	=> {Tempocod= 99h}	0,1181	0,5181
{Categoria=SAP}	=> {Tempocod= 99h}	0,0598	0,4563
{Categoria=SISTEMAS}	=> {Tempocod= 54h}	0,0266	0,4346
{Categoria=SEGURANCA}	=> {Tempocod= 54h}	0,0344	0,4127
<b>2006</b>			
lhs	rhs	support	confidence
{Categoria=SISTEMAS}	=> {Tempocod= 27h}	0,0387	0,5238
{Categoria=SERVICOS DE TI}	=> {Tempocod= 99h}	0,0993	0,4483
{Categoria=SEGURANCA}	=> {Tempocod= 54h}	0,0481	0,4328
{Categoria=HARDWARE}	=> {Tempocod= 54h}	0,0496	0,4188
{Categoria=SAP}	=> {Tempocod= 99h}	0,0373	0,4124
{Categoria=INFRA-ESTRUTURA}	=> {Tempocod= 54h}	0,0291	0,4081
{Tempocod= 9h}	=> {Categoria=HARDWARE}	0,0031	0,4071

**Tabela 10 – Análise das Regras – Atributos Categoria X Tempo Total**

- Os Serviços de TI e o SAP R/3 são as categorias que demandam mais tempo para concluir um registro.
- A categoria Sistemas reduziu seu tempo de “54h” em 2005 para “27h” em 2006. Este fato pode ser atribuído ao aumento do número de Mesas Especializadas para atender sistemas específicos.
- A maior parte dos registros encerrados em “9h” pertence a Categoria Hardware. Este fato surpreende e merece uma análise mais detalhada pelos gestores do apoio ao usuário.

## Categoria X Período

2005				
lhs		rhs	support	confidence
{Categoria=HARDWARE}	=>	{Período=M}	0,0693	0,5653
{Categoria=APLICACAO}	=>	{Período=M}	0,0865	0,5235
{Categoria=LOTUS NOTES}	=>	{Período=M}	0,0567	0,5219
{Categoria=SERVICOS DE TI}	=>	{Período=M}	0,1148	0,5035
{Categoria=IMPRESSAO}	=>	{Período=M}	0,0182	0,5034
{Categoria=SEGURANCA}	=>	{Período=M}	0,0402	0,4823
{Categoria=SISTEMAS}	=>	{Período=T}	0,0294	0,4806
{Categoria=SAP}	=>	{Período=T}	0,0623	0,4761
{Categoria=SAP}	=>	{Período=M}	0,0621	0,4742
{Categoria=SISTEMAS}	=>	{Período=M}	0,0287	0,4691
{Categoria=SEGURANCA}	=>	{Período=T}	0,0378	0,4531
{Categoria=IMPRESSAO}	=>	{Período=T}	0,0157	0,4341
{Categoria=SERVICOS DE TI}	=>	{Período=T}	0,0983	0,4314
{Categoria=LOTUS NOTES}	=>	{Período=T}	0,0464	0,4263
{Categoria=APLICACAO}	=>	{Período=T}	0,0697	0,4220
{Período=O}	=>	{Categoria=INFRA-ESTRUTURA}	0,0068	0,4142
{Categoria=HARDWARE}	=>	{Período=T}	0,0498	0,4062
{Categoria=INFRA-ESTRUTURA}	=>	{Período=M}	0,0258	0,4040
2006				
lhs		rhs	support	confidence
{Categoria=HARDWARE}	=>	{Período=M}	0,0681	0,5749
{Período=O}	=>	{Categoria=INFRA-ESTRUTURA}	0,0100	0,5210
{Categoria=LOTUS NOTES}	=>	{Período=M}	0,0641	0,5197
{Categoria=APLICACAO}	=>	{Período=M}	0,0815	0,5177
{Categoria=SISTEMAS}	=>	{Período=T}	0,0375	0,5076
{Categoria=IMPRESSAO}	=>	{Período=M}	0,0166	0,5076
{Categoria=SEGURANCA}	=>	{Período=M}	0,0554	0,4987
{Categoria=SERVICOS DE TI}	=>	{Período=M}	0,1099	0,4963
{Categoria=SAP}	=>	{Período=T}	0,0431	0,4769
{Categoria=SAP}	=>	{Período=M}	0,0428	0,4728
{Categoria=SISTEMAS}	=>	{Período=M}	0,0333	0,4509
{Categoria=SEGURANCA}	=>	{Período=T}	0,0483	0,4343
{Categoria=IMPRESSAO}	=>	{Período=T}	0,0140	0,4290
{Categoria=SERVICOS DE TI}	=>	{Período=T}	0,0944	0,4262
{Categoria=APLICACAO}	=>	{Período=T}	0,0667	0,4239
{Categoria=LOTUS NOTES}	=>	{Período=T}	0,0522	0,4232
{Categoria=HARDWARE}	=>	{Período=T}	0,0486	0,4101

Tabela 11 – Análise das Regras – Atributos Categoria X Período

- Os registros relacionados às categorias Impressão, Lotus Notes, Hardware, Aplicação, Segurança e Serviços de TI são abertos, na sua maioria, no período da manhã.
- Durante a madrugada, a maior incidência de registros está relacionada à infraestrutura. Isso ocorre principalmente devido a criação automática de registros pelos sistemas de monitoramento de servidores.

- Os usuários dos sistemas específicos e do SAP R/3 abrem seus registros no período da tarde. Provavelmente estes sistemas são mais utilizados também no período da tarde.
- Não há mudanças significativas nas regras de 2005 para 2006.

### Categoria X Tipo de Grupo

2005				
lhs	rhs	support	confidence	
{Categoria=SERVICOS DE TI}	=> {Tipogrupo=HD}	0,1947	0,8541	
{Categoria=INFRA-ESTRUTURA}	=> {Tipogrupo=ME}	0,0527	0,8250	
{Categoria=HARDWARE}	=> {Tipogrupo=PA}	0,0976	0,7955	
{Categoria=SAP}	=> {Tipogrupo=HD}	0,0878	0,6705	
{Categoria=LOTUS NOTES}	=> {Tipogrupo=HD}	0,0708	0,6511	
{Categoria=IMPRESSAO}	=> {Tipogrupo=HD}	0,0222	0,6162	
{Categoria=SEGURANCA}	=> {Tipogrupo=ME}	0,0501	0,6010	
{Categoria=SISTEMAS}	=> {Tipogrupo=ME}	0,0367	0,6002	
{Categoria=APLICACAO}	=> {Tipogrupo=PA}	0,0771	0,4671	
{Categoria=APLICACAO}	=> {Tipogrupo=HD}	0,0748	0,4531	
{Tipogrupo=PA}	=> {Categoria=HARDWARE}	0,0976	0,4152	
2006				
lhs	rhs	support	confidence	
{Categoria=INFRA-ESTRUTURA}	=> {Tipogrupo=ME}	0,0625	0,8764	
{Categoria=SERVICOS DE TI}	=> {Tipogrupo=HD}	0,1828	0,8256	
{Categoria=HARDWARE}	=> {Tipogrupo=PA}	0,0869	0,7340	
{Categoria=SAP}	=> {Tipogrupo=HD}	0,0644	0,7118	
{Categoria=LOTUS NOTES}	=> {Tipogrupo=HD}	0,0842	0,6829	
{Categoria=IMPRESSAO}	=> {Tipogrupo=HD}	0,0207	0,6339	
{Categoria=SEGURANCA}	=> {Tipogrupo=ME}	0,0651	0,5860	
{Categoria=SISTEMAS}	=> {Tipogrupo=ME}	0,0362	0,4903	
{Categoria=APLICACAO}	=> {Tipogrupo=PA}	0,0747	0,4743	
{Categoria=APLICACAO}	=> {Tipogrupo=HD}	0,0709	0,4505	
{Categoria=SISTEMAS}	=> {Tipogrupo=HD}	0,0332	0,4496	

Tabela 12 – Análise das Regras – Atributos Categoria X Tipo de Grupo

- Confirmando o levantamento estatístico, o *Help-desk* responde pela maior parte dos atendimentos, seguido pelas mesas especializadas.
- Contrariando o esperado, a Categoria Aplicação é mais atendida pelos postos avançados.

- O crescimento da confiança da regra SAP R/3 → HD pode indicar a estabilização do sistema R/3 e a maior capacitação do *Help-desk*, reduzindo assim a necessidade destes registros serem atendidos pelas mesas especializadas.
- Os registros de Impressão, na sua maior parte, estão resolvidos pelo *Help-desk*, remotamente, evitando-se a presença física de um técnico. Isso reduz o tempo e o custo do atendimento.
- Não há mudanças significativas nas regras de 2005 para 2006.

### Categoria X Tipo de Registro

2005				
lhs	=>	rhs	support	confidence
{Categoria=SAP}	=>	{Tipocham=INC}	0.1148	0.8769
{Categoria=INFRA-ESTRUTURA}	=>	{Tipocham=INC}	0.0547	0.8563
{Categoria=SERVICOS DE TI}	=>	{Tipocham=INC}	0.1869	0.8198
{Categoria=LOTUS NOTES}	=>	{Tipocham=INC}	0.0806	0.7410
{Categoria=IMPRESSAO}	=>	{Tipocham=INC}	0.0263	0.7281
{Categoria=APLICACAO}	=>	{Tipocham=INC}	0.1170	0.7087
{Categoria=SISTEMAS}	=>	{Tipocham=INC}	0.0400	0.6552
{Categoria=HARDWARE}	=>	{Tipocham=INC}	0.0665	0.5420
{Categoria=SEGURANCA}	=>	{Tipocham=INC}	0.0417	0.4996
{Tipocham=DUV}	=>	{Categoria=SAP}	0.0039	0.4658
{Categoria=SEGURANCA}	=>	{Tipocham=SOL}	0.0298	0.3577
{Tipocham=COB}	=>	{Categoria=HARDWARE}	0.0190	0.3069
2006				
lhs	=>	rhs	support	confidence
{Categoria=INFRA-ESTRUTURA}	=>	{Tipocham=INC}	0.0599	0.8405
{Categoria=SAP}	=>	{Tipocham=INC}	0.0715	0.7900
{Categoria=SISTEMAS}	=>	{Tipocham=INC}	0.0521	0.7059
{Categoria=LOTUS NOTES}	=>	{Tipocham=INC}	0.0794	0.6440
{Categoria=SERVICOS DE TI}	=>	{Tipocham=INC}	0.1306	0.5897
{Categoria=APLICACAO}	=>	{Tipocham=INC}	0.0909	0.5772
{Categoria=IMPRESSAO}	=>	{Tipocham=INC}	0.0175	0.5363
{Categoria=HARDWARE}	=>	{Tipocham=INC}	0.0546	0.4613
{Categoria=SEGURANCA}	=>	{Tipocham=INC}	0.0486	0.4373
{Categoria=SEGURANCA}	=>	{Tipocham=SOL}	0.0453	0.4079
{Tipocham=DUV}	=>	{Categoria=SAP}	0.0025	0.3523
{Categoria=SERVICOS DE TI}	=>	{Tipocham=SOL}	0.0764	0.3451
{Categoria=IMPRESSAO}	=>	{Tipocham=SOL}	0.0109	0.3343
{Tipocham=COB}	=>	{Categoria=HARDWARE}	0.0203	0.3240

Tabela 13 – Análise das Regras – Atributos Categoria X Tipo de Registro



- Os Incidentes são predominantes para todas as categorias.
- Os registros do tipo “Dúvida” estão relacionados com o SAP R/3. Os registros do tipo Cobertura estão relacionados com a Categoria Hardware.
- Não há mudanças significativas nas regras de 2005 para 2006.

### Origem X Período

2005				
lhs	=>	rhs	support	confidence
{Período=M}	=>	{Origem=TEL}	0,3649	0,7265
{Período=T}	=>	{Origem=TEL}	0,3120	0,7200
{Período=N}	=>	{Origem=TEL}	0,0331	0,6913
{Origem=WEB}	=>	{Período=M}	0,1052	0,5285
{Origem=LIG}	=>	{Período=M}	0,0003	0,5091
{Origem=TEL}	=>	{Período=M}	0,3649	0,5082
{Período=O}	=>	{Origem=TEL}	0,0080	0,4860
{Origem=LIG}	=>	{Período=T}	0,0003	0,4705
{Origem=APO}	=>	{Período=T}	0,0147	0,4666
{Origem=APO}	=>	{Período=M}	0,0145	0,4592
{Origem=WEB}	=>	{Período=T}	0,0883	0,4438
{Origem=EMA}	=>	{Período=M}	0,0073	0,4399
{Origem=EMA}	=>	{Período=T}	0,0073	0,4363
{Origem=TEL}	=>	{Período=T}	0,3120	0,4345
{Período=O}	=>	{Origem=TIV}	0,0068	0,4140
2006				
lhs	=>	rhs	support	confidence
{Período=M}	=>	{Origem=TEL}	0,3337	0,6715
{Período=T}	=>	{Origem=TEL}	0,2861	0,6657
{Período=N}	=>	{Origem=TEL}	0,0335	0,6181
{Período=O}	=>	{Origem=TIV}	0,0109	0,5665
{Origem=WEB}	=>	{Período=M}	0,1142	0,5339
{Origem=TEL}	=>	{Período=M}	0,3337	0,5049
{Origem=LIG}	=>	{Período=M}	0,0002	0,4936
{Origem=EMA}	=>	{Período=M}	0,0117	0,4743
{Origem=APO}	=>	{Período=M}	0,0234	0,4675
{Origem=LIG}	=>	{Período=T}	0,0002	0,4638
{Origem=APO}	=>	{Período=T}	0,0231	0,4620
{Origem=WEB}	=>	{Período=T}	0,0947	0,4431
{Origem=EMA}	=>	{Período=T}	0,0109	0,4414
{Origem=TEL}	=>	{Período=T}	0,2861	0,4329

Tabela 14 – Análise das Regras – Atributos Origem X Período

- Crescimento dos registros originados pela ferramenta Tivoli durante a madrugada. Este dado confirma o crescimento do número de servidores e dos serviços destes que são monitorados.
- Confirmando as estatísticas o telefone é a principal origem dos registros em qualquer horário. Todavia em 2006 durante a madrugada os registros cuja origem é Tivoli superaram a telefone.

### Origem X Tempo Total

2005				
lhs	=>	rhs	support	confidence
{Tempocod= 99h}	=>	{Origem=TEL}	0,2861	0,8923
{Tempocod=100h}	=>	{Origem=TEL}	0,1229	0,8834
{Tempocod= 1h}	=>	{Origem=WEB}	0,0438	0,7747
{Tempocod= 54h}	=>	{Origem=TEL}	0,1891	0,6553
{Tempocod= 27h}	=>	{Origem=TEL}	0,1012	0,5643
{Origem=APO}	=>	{Tempocod= 54h}	0,0168	0,5310
{Tempocod= 9h}	=>	{Origem=TEL}	0,0080	0,5115
{Origem=EMA}	=>	{Tempocod= 54h}	0,0079	0,4741
{Tempocod= 9h}	=>	{Origem=WEB}	0,0068	0,4348
2006				
lhs	=>	rhs	support	confidence
{Tempocod=100h}	=>	{Origem=TEL}	0,1630	0,8742
{Tempocod= 1h}	=>	{Origem=WEB}	0,0472	0,8692
{Tempocod= 99h}	=>	{Origem=TEL}	0,2394	0,8673
{Tempocod= 9h}	=>	{Origem=WEB}	0,0058	0,7644
{Origem=EMA}	=>	{Tempocod= 54h}	0,0146	0,5925
{Tempocod= 54h}	=>	{Origem=TEL}	0,1670	0,5624
{Origem=APO}	=>	{Tempocod= 27h}	0,0239	0,4788
{Tempocod= 27h}	=>	{Origem=TEL}	0,0830	0,4647
{Origem=APO}	=>	{Tempocod= 54h}	0,0220	0,4393
{Origem=LIG}	=>	{Tempocod= 54h}	0,0002	0,4170
{Origem=TIV}	=>	{Tempocod= 54h}	0,0196	0,4085

Tabela 15 – Análise das Regras – Atributos Origem X Tempo Total

- Os registros cujo “tempo total” é “1h” tem como origem a Web. Este fato, cuja confiança cresceu significativamente de 2005 para 2006, pode estar relacionado aos registros de cobertura.
- Os registros que demandam maior tempo para solução e representam os problemas mais complexos, são na maioria dos casos, abertos pelo telefone.

- Os registros cuja origem é e-mail exibem tempo de atendimento de “54h”. Pelas regras anteriores, este tempo coincide também com os chamados da Categoria Segurança.

### Origem X Tipo de Registro

2005				
lhs	=>	rhs	support	confidence
{Tipocham=TAR}	=>	{Origem=TEL}	0,0350	0,9990
{Origem=TIV}	=>	{Tipocham=INC}	0,0330	0,9969
{Tipocham=DUV}	=>	{Origem=TEL}	0,0081	0,9724
{Origem=APO}	=>	{Tipocham=SOL}	0,0302	0,9549
{Origem=EMA}	=>	{Tipocham=INC}	0,0152	0,9100
{Origem=TEL}	=>	{Tipocham=INC}	0,5811	0,8093
{Tipocham=INC}	=>	{Origem=TEL}	0,5811	0,7976
{Origem=LIG}	=>	{Tipocham=INC}	0,0005	0,7932
{Tipocham=COB}	=>	{Origem=WEB}	0,0486	0,7859
{Tipocham=SOL}	=>	{Origem=TEL}	0,0827	0,4975
{Origem=WEB}	=>	{Tipocham=INC}	0,0981	0,4932
2006				
lhs	=>	rhs	support	confidence
{Origem=TIV}	=>	{Tipocham=INC}	0,0475	0,9886
{Tipocham=DUV}	=>	{Origem=TEL}	0,0068	0,9657
{Tipocham=COB}	=>	{Origem=WEB}	0,0536	0,8560
{Origem=LIG}	=>	{Tipocham=INC}	0,0003	0,7574
{Tipocham=INC}	=>	{Origem=TEL}	0,4529	0,7484
{Origem=APO}	=>	{Tipocham=SOL}	0,0355	0,7104
{Origem=TEL}	=>	{Tipocham=INC}	0,4529	0,6853
{Tipocham=SOL}	=>	{Origem=TEL}	0,1723	0,6346
{Origem=EMA}	=>	{Tipocham=INC}	0,0153	0,6184
{Origem=WEB}	=>	{Tipocham=INC}	0,0889	0,4159

**Tabela 16 – Análise das Regras – Atributos Origem X Tipo de Registro**

- A maior parte dos usuários prefere utilizar o telefone para dirimir suas dúvidas.
- Os registros do tipo “Cobertura” têm maior origem através da Web.
- É interessante observar a queda da confiança da relação Telefone → Incidente. Este fato coincide com a estatística que aponta queda no uso do telefone e no percentual de Incidentes de 2005 para 2006.

## Origem X Tipo de Grupo

2005				
lhs		rhs	support	confidence
{Tipogrupo=HD}	=>	{Origem=TEL}	0.4836	0.9559
{Origem=TIV}	=>	{Tipogrupo=ME}	0.0314	0.9496
{Origem=APO}	=>	{Tipogrupo=ME}	0.0288	0.9100
{Origem=TEL}	=>	{Tipogrupo=HD}	0.4836	0.6735
{Origem=WEB}	=>	{Tipogrupo=PA}	0.1173	0.5897
{Origem=EMA}	=>	{Tipogrupo=ME}	0.0085	0.5076
{Tipogrupo=PA}	=>	{Origem=WEB}	0.1173	0.4993
{Tipogrupo=PA}	=>	{Origem=TEL}	0.1117	0.4755
{Tipogrupo=ME}	=>	{Origem=TEL}	0.1227	0.4736
{Origem=LIG}	=>	{Tipogrupo=PA}	0.0003	0.4511
2006				
lhs		rhs	support	confidence
{Origem=TIV}	=>	{Tipogrupo=ME}	0.0460	0.9582
{Tipogrupo=HD}	=>	{Origem=TEL}	0.4808	0.9513
{Origem=APO}	=>	{Tipogrupo=ME}	0.0455	0.9098
{Origem=TEL}	=>	{Tipogrupo=HD}	0.4808	0.7275
{Origem=EMA}	=>	{Tipogrupo=ME}	0.0162	0.6553
{Origem=WEB}	=>	{Tipogrupo=PA}	0.1182	0.5526
{Tipogrupo=PA}	=>	{Origem=WEB}	0.1182	0.5390
{Origem=LIG}	=>	{Tipogrupo=PA}	0.0002	0.5000
{Tipogrupo=PA}	=>	{Origem=TEL}	0.0937	0.4276

Tabela 17 – Análise das Regras – Atributos Origem X Tipo de Grupo

- Os resultados nesta tabela confirmam que o *Help-desk* está muito ligado ao telefone, assim como os PAs à Web e ao LIG-TI. As demais origens estão mais ligadas as MEs.
- As regras PA → Web e PA → LIG-TI confirmam o uso destes meios para se criar os registros do tipo Cobertura.
- A regra de maior confiança em 2006 (Tivoli → ME) confirma que quase todos estes registros estão relacionados a servidores e devem ser tratados pelas Mesas Especializadas.

## Período X Tipo de Grupo

2005				
lhs	=>	rhs	support	confidence
{Período=N}	=>	{Tipogrupo=HD}	0,0289	0,6044
{Tipogrupo=PA}	=>	{Período=M}	0,1322	0,5626
{Período=T}	=>	{Tipogrupo=HD}	0,2214	0,5109
{Período=M}	=>	{Tipogrupo=HD}	0,2494	0,4964
{Tipogrupo=HD}	=>	{Período=M}	0,2494	0,4929
{Período=O}	=>	{Tipogrupo=ME}	0,0080	0,4860
{Tipogrupo=ME}	=>	{Período=M}	0,1208	0,4660
{Tipogrupo=ME}	=>	{Período=T}	0,1161	0,4481
{Tipogrupo=HD}	=>	{Período=T}	0,2214	0,4376
{Tipogrupo=PA}	=>	{Período=T}	0,0958	0,4079
2006				
lhs	=>	rhs	support	confidence
{Período=O}	=>	{Tipogrupo=ME}	0,0117	0,6087
{Período=N}	=>	{Tipogrupo=HD}	0,0310	0,5727
{Tipogrupo=PA}	=>	{Período=M}	0,1238	0,5646
{Período=T}	=>	{Tipogrupo=HD}	0,2189	0,5094
{Período=M}	=>	{Tipogrupo=HD}	0,2486	0,5002
{Tipogrupo=HD}	=>	{Período=M}	0,2486	0,4919
{Tipogrupo=ME}	=>	{Período=M}	0,1246	0,4523
{Tipogrupo=ME}	=>	{Período=T}	0,1201	0,4361
{Tipogrupo=HD}	=>	{Período=T}	0,2189	0,4332
{Tipogrupo=PA}	=>	{Período=T}	0,0907	0,4138

Tabela 18 – Análise das Regras – Atributos Período X Tipo de Grupo

- O grupo PA tem maior número de registros no horário da manhã. Este resultado confirma alguns resultados anteriores, onde os registros da categoria Hardware, atendidos quase que exclusivamente pelos PAs, ocorrem na sua maioria, pela manhã.
- O aumento da confiança dos registros na madrugada está ligado ao grupo ME e confirma o crescimento dos registros na categoria infra-estrutura bem como o aumento da confiança na regra Tivoli → ME.

## Tipo de Registro X Período

2005				
lhs		rhs	support	confidence
{Período=O}	=>	{Tipocham=INC}	0,0132	0,7995
{Período=N}	=>	{Tipocham=INC}	0,0379	0,7921
{Período=T}	=>	{Tipocham=INC}	0,3171	0,7317
{Período=M}	=>	{Tipocham=INC}	0,3604	0,7175
{Tipocham=COB}	=>	{Período=M}	0,0343	0,5546
{Tipocham=SOL}	=>	{Período=M}	0,0867	0,5215
{Tipocham=INC}	=>	{Período=M}	0,3604	0,4947
{Tipocham=TAR}	=>	{Período=T}	0,0171	0,4874
{Tipocham=TAR}	=>	{Período=M}	0,0170	0,4851
{Tipocham=DUV}	=>	{Período=M}	0,0040	0,4729
{Tipocham=DUV}	=>	{Período=T}	0,0038	0,4582
{Tipocham=INC}	=>	{Período=T}	0,3171	0,4352
{Tipocham=SOL}	=>	{Período=T}	0,0717	0,4315
2006				
lhs		rhs	support	confidence
{Período=O}	=>	{Tipocham=INC}	0,0153	0,7957
{Período=N}	=>	{Tipocham=INC}	0,0343	0,6325
{Período=T}	=>	{Tipocham=INC}	0,2589	0,6024
{Período=M}	=>	{Tipocham=INC}	0,2967	0,5971
{Tipocham=COB}	=>	{Período=M}	0,0367	0,5867
{Tipocham=DUV}	=>	{Período=M}	0,0035	0,4946
{Tipocham=SOL}	=>	{Período=M}	0,1342	0,4945
{Tipocham=TAR}	=>	{Período=T}	0,0265	0,4917
{Tipocham=INC}	=>	{Período=M}	0,2967	0,4903
{Tipocham=TAR}	=>	{Período=M}	0,0258	0,4788
{Tipocham=SOL}	=>	{Período=T}	0,1176	0,4331
{Tipocham=INC}	=>	{Período=T}	0,2589	0,4278
{Tipocham=DUV}	=>	{Período=T}	0,0030	0,4259

**Tabela 19 – Análise das Regras – Atributos Tipo de Registro X Período**

- O tipo de registro que mais ocorre à noite e na madrugada é o Incidente. Principalmente quanto aos registros da madrugada, este fato pode ser explicado pelo crescimento dos registros de infra-estrutura neste horário.
- As tarefas, por serem registros criados a partir de solicitações de forma manual, têm maior ocorrência durante a tarde.
- Não há mudanças significativas nas regras de 2005 para 2006.

## Tipo de Registro X Tempo Total

2005				
lhs	=>	rhs	support	confidence
{Tempocod= 1h}	=>	{Tipocham=COB}	0,0557	0,9848
{Tipocham=COB}	=>	{Tempocod= 1h}	0,0557	0,9011
{Tempocod= 99h}	=>	{Tipocham=INC}	0,2827	0,8817
{Tempocod=100h}	=>	{Tipocham=INC}	0,1196	0,8600
{Tempocod= 54h}	=>	{Tipocham=INC}	0,2059	0,7134
{Tempocod= 27h}	=>	{Tipocham=INC}	0,1118	0,6232
{Tipocham=DUV}	=>	{Tempocod= 54h}	0,0045	0,5353
{Tempocod= 9h}	=>	{Tipocham=INC}	0,0079	0,5059
{Tipocham=TAR}	=>	{Tempocod= 54h}	0,0161	0,4587
{Tipocham=TAR}	=>	{Tempocod= 27h}	0,0149	0,4252
2006				
lhs	=>	rhs	support	confidence
{Tempocod= 1h}	=>	{Tipocham=COB}	0,0543	0,9994
{Tempocod= 9h}	=>	{Tipocham=COB}	0,0076	0,9974
{Tipocham=COB}	=>	{Tempocod= 1h}	0,0543	0,8669
{Tempocod= 99h}	=>	{Tipocham=INC}	0,1932	0,6997
{Tempocod=100h}	=>	{Tipocham=INC}	0,1303	0,6987
{Tempocod= 54h}	=>	{Tipocham=INC}	0,1880	0,6329
{Tempocod= 27h}	=>	{Tipocham=INC}	0,0937	0,5244
{Tipocham=DUV}	=>	{Tempocod= 54h}	0,0035	0,5023
{Tipocham=TAR}	=>	{Tempocod= 54h}	0,0247	0,4584
{Tipocham=TAR}	=>	{Tempocod= 27h}	0,0245	0,4557

**Tabela 20 – Análise das Regras – Atributos Tipo de Registro X Tempo Total**

- Os registros de cobertura são aqueles criados pelos técnicos após a solução do problema. São os que apresentaram os menores tempos de solução (1h e 9h). Este dado precisa de um estudo mais profundo para que se identifique se estes tempos são reais, uma vez que os registros de cobertura têm seus tempos alimentados pelos técnicos e não pelo sistema.
- Os tipos de registro que demandam mais tempo são os incidentes, sendo este um resultado esperado.
- As dúvidas estão apresentando um alto tempo para solução, sendo este um resultado que precisará de informações complementares para que sejam identificadas as causas.

## Mês/Ano X Período

2005				
lhs		rhs	support	confidence
{Mesano=Dec/05}	=>	{Período=M}	0.0467	0.5390
{Mesano=Nov/05}	=>	{Período=M}	0.0473	0.5288
{Mesano=Oct/05}	=>	{Período=M}	0.0479	0.5084
{Mesano=Feb/05}	=>	{Período=M}	0.0344	0.5032
{Mesano=Apr/05}	=>	{Período=T}	0.0343	0.4453
{Mesano=Jan/05}	=>	{Período=T}	0.0322	0.4416
{Mesano=Mar/05}	=>	{Período=T}	0.0367	0.4407
{Mesano=Aug/05}	=>	{Período=T}	0.0423	0.4395
{Mesano=Sep/05}	=>	{Período=T}	0.0408	0.4394
{Mesano=Jul/05}	=>	{Período=T}	0.0348	0.4391
{Mesano=May/05}	=>	{Período=T}	0.0346	0.4382
{Mesano=Jun/05}	=>	{Período=T}	0.0352	0.4357
2006				
lhs		rhs	support	confidence
{Mesano=Feb/06}	=>	{Período=M}	0.0546	0.5262
{Mesano=Jan/06}	=>	{Período=M}	0.0620	0.5195
{Mesano=Jun/06}	=>	{Período=M}	0.0615	0.5093
{Mesano=May/06}	=>	{Período=T}	0.0612	0.4446
{Mesano=Jul/06}	=>	{Período=T}	0.0587	0.4347
{Mesano=Mar/06}	=>	{Período=T}	0.0569	0.4334
{Mesano=Aug/06}	=>	{Período=T}	0.0597	0.4330

Tabela 21 – Análise das Regras – Atributos Mês/Ano X Período

- O mês de fevereiro teve predominância de registros no período da manhã. Interessante observar que este fato não só se confirmou em 2006 como cresceu a confiança da regra.
- Os meses de março, julho e agosto apresentaram predominância de registros no período da tarde. Estes resultados precisam ser analisados em conjunto com outros para um melhor entendimento.
- A análise destas relações ficou prejudicada pelo fato dos dados de 2006 não contemplar todos os meses.



### 3.4.1 Consolidando alguns resultados encontrados

Os resultados das regras geradas podem ser agrupados e consolidados, criando um conhecimento novo e mais abrangente que simples relação entre os atributos analisados.

Tomando-se como exemplo as regras de maior confiança do atributo Tipo de registro = Cobertura(COB), temos: COB → HARDWARE, COB → PA, COB → WEB , COB → MANHÃ, COB → 1H

Pode-se observar que os registros de cobertura estão relacionados com os técnicos que fazem o atendimento presencial e que em grande parte são problemas de hardware. Foram identificados no período da manhã e resolvidos em menos de 1 hora. Estes registros foram criados a partir da Web, pelos próprios técnicos dos PAs. Analisando as demais regras relacionadas com a categoria Hardware, com a origem Web e com o Tempo de até 1h conclui-se que este tempo de atendimento, alimentado pelo próprio técnico, está subestimado.

O Apêndice B apresenta o resultado da consolidação de todas as regras para o atributo Categoria. Este atributo foi selecionado por ser o mais importante dentro do conjunto dos que foram analisados neste estudo. Através da análise de volume das categorias de registros e de seus respectivos subgrupos, pode ser realizado um melhor dimensionamento dos recursos de atendimento.

## 4 CONCLUSÃO

A grande massa de dados dos registros de *Help-desk* foi analisada em sua totalidade pela primeira vez desde a reorganização do órgão de Tecnologia da Informação da Petrobras. As estatísticas dos principais atributos foram úteis, antes do processo de mineração de dados, para conhecimento do seu volume e sua variação ao longo do tempo.

A ausência de um *Data Warehouse* dos dados de *Help-desk* tornou a etapa de preparação de dados mais longa e difícil. O esforço de programação poderia ter sido bastante reduzido caso os dados já estivessem preparados e limpos para serem utilizados no processo de mineração. O uso de softwares que auxiliem a exploração e a manipulação de dados como o SPSS, pode reduzir bastante o tempo e o esforço despendido nesta etapa.

A metodologia sugerida no processo de *KDD*, onde há grande foco na preparação dos dados, mostrou-se eficiente e fundamental neste estudo. A mineração de dados, independentemente da ferramenta que será usada, exige dados limpos, padronizados e codificados. O esforço despendido nesta etapa é compensado, pois torna mais simples a extração das regras e aumenta a confiabilidade dos resultados.

A ferramenta Rattle atendeu as necessidades deste estudo de forma satisfatória, com bom desempenho. A simplicidade de sua interface aliada a facilidade de operar o software, tornaram as tarefas repetitivas menos tediosas. No entanto, alguns problemas surgiram ao executar seguidas vezes o processo de geração de regras. Este problema, todavia, não comprometeu o resultado.

As regras geradas, de uma forma geral, não sofreram grandes variações de um ano para o outro. Esta informação dá consistência ao trabalho de preparação dos dados e ao mesmo

tempo, confirma o fato de que não houve mudanças significativas na estrutura e na operação do *Help-desk*.

Os resultados consolidados das regras representam conhecimento novo e útil para a gestão do *Help-desk*. Foi possível fazer uma análise detalhada deste atributo em relação aos demais atributos. As informações apresentadas poderão ser usadas para o melhor dimensionamento das diversas equipes de atendimento, de acordo com o horário de maior demanda e tempo gasto para aquele atendimento.

#### 4.1 TRABALHOS FUTUROS

Alguns atributos importantes, do ponto de vista dos gestores do Apoio ao Usuário da TI, ficaram ausentes deste estudo, como:

- A regional de serviços ou de Agilidade responsável pelo registro;
- O subgrupo de HD, ME ou PA que atendeu o registro;
- O tempo de atendimento atribuído somente a TI, ou seja, o tempo total descontando-se o tempo em que o registro ficou pendente por conta do usuário ou de algum outro fator externo;
- O órgão de origem do usuário;
- O produto (corresponde a especificidade da viável Categoria).

Estes atributos, para que possam servir de entrada em um processo de mineração de dados, deverão passar por um longo processo de preparação, padronizando seus valores possíveis e agrupando-os.

Outras tarefas de mineração de dados poderão ser utilizadas, como a classificação e análise de agrupamentos. A grande massa de dados que foi preparada poderá ser utilizada como base de entrada.

As regras extraídas para o ano de 2006 poderão ser geradas novamente quando os dados de todos os meses do ano de 2006 estiverem disponíveis para compor as bases de dados de entrada. Desta forma, uma nova avaliação de resultados poderá ser feita, comparando com o resultado deste trabalho, ressaltando-se assim as diferenças importantes encontradas.

O conhecimento gerado nos resultados consolidados deste estudo, combinado com o estudo de outras bases de dados e de conhecimento, além dos demais sistemas de *Help-desk* da Petrobras, poderão servir de base para elaboração de novas aplicações de *Help-desk*. Uma opção interessante seria a criação de um sistema inteligente, que utilizasse técnicas da Inteligência Artificial e da Mineração de Dados, para identificação automática dos registros, estimativa de tempo a ser gasto e as possíveis soluções do problema.

## REFERÊNCIAS BIBLIOGRÁFICAS

AGRAWAL, R. (1993) Mining Association rules between sets of items in large databases. In: ACM SIGMOD Conf. Management of Data, 1993. Washington: ACM SIGMOND.

BARBIERI, C. (2001) **BI – Bussiness Intelligence: Modelagem & Tecnologia**. Rio de Janeiro, Axcel Books.

CARVALHO, Luís Alfredo Vidal de (2005) **Datamining - A mineração de dados no Marketing, Medicina, Economia, Engenharia e Administração**. Rio de Janeiro, Ciência Moderna.

FAYYAD, U.; PIATETSKY-SHAPIRO, G.; SMYTH, P. (1996). **The KDD Process for Extracting Useful Knowledge from Volumes of Data**. *Communications of the ACM*, USA.

FERREIRA, A.B.H.(Ed.) (1999) **Dicionário Aurélio Eletrônico Século XXI**. Rio de Janeiro: Nova Fronteira.

GOLDSCHMIDT, Ronaldo, PASSOS Emmanuel (2005) **Data Mining: um guia prático**. Rio de Janeiro, Ed Campus/Elsevier.

GROTH, Robert (1998) **Data Mining: a Hands on Approach for Business Professionals**. Prentice-Hall PTR.

HAN, Jiawei, KAMBER, Micheline (2001) **Data Mining : Concepts and Techniques**. San Diego. Academic Press.

OLIVEIRA, Elizabeth (2005) **Determinação dos Fatores Críticos na Análise de Desempenho do Corpo Discente de Cursos de Pós-graduação Lato Sensu Utilizando Metodologia Data Mining**. Rio de Janeiro, Faculdades Ibmecc.

PETROBRAS. **Missão e Visão**.  
[http://www2.petrobras.com.br/ri/port/ApresentacoesEventos/ConfTelefonicas/pdf/Plano\\_Estrategico\\_2015\\_FINAL\\_1007.pdf](http://www2.petrobras.com.br/ri/port/ApresentacoesEventos/ConfTelefonicas/pdf/Plano_Estrategico_2015_FINAL_1007.pdf) . Acesso em: outubro/2006.

RESENDE, Solange Oliveira (2005) **Sistemas Inteligentes**. São Paulo: Manole

TANAGRA. Functionalities – Tutorials – Download. <http://eric.univ-lyon2.fr/~ricco/tanagra/en/tanagra.html>. Acesso em dezembro/2006.

TOGAWARE.(2006). **Rattle: Gnome R Datamining**. Disponível em: <http://rattle.togaware.com> Acesso em junho/2006.

WEKA. Getting started. <http://www.cs.waikato.ac.nz/ml/weka/index.html>. Acesso em dezembro/2006.

WIZRULE. Products. <http://www.wizsoft.com>. Acesso em dezembro/2006.

## **BIBLIOGRAFIA COMPLEMENTAR**

ARBEX, Eduardo C., SABOREDO, Alexandre P, MIRANDA, Dhalila (2004) **Implementação e Estudo de Caso do Algoritmo Apriori para Mineração de Dados**. Resende/RJ, Associação Educacional Dom Bosco.

ÁVILA, B. C. (1998) **Data Mining: Escola Regional de Informática da SBC**. Blumenau, Anais SBC.

CID Dante José Alexandre, PASSOS, Emmanuel P.L. (2005) **Análise de Ferramentas de Data Mining**. Rio de Janeiro, Pontifícia Universidade Católica - PUC.

COUTINHO, Fernando Vieira (2005) **Data Mining**. Disponível em: <http://www.dwbrasil.com.br/html/dmining.html>. Acesso em: julho/2006.

CHRISTIAN BOGELT. (2006) Apriori - Association Rule Induction. Disponível em: <http://fuzzy.cs.uni-magdeburg.de/~borgelt/apriori.html> . Acesso em: julho/2006.

DIAS, Maria M. (2002) **Estudo e Análise de Técnicas e Ferramentas de Mineração de Dados**. Santa Catarina, UFSC.

FERNANDES, Anita Maria da Rocha (2005) **Inteligência Artificial: noções gerais**. Florianópolis, VisualBooks.

KANTARDZIC, Mehmed (2003) **Data Mining : Concepts, Models, Methods, and Algorithms**. Nova Jersey, IEEE Press.

NAVEGA, Sérgio (2002) **Princípios Essenciais do Data Mining**. São Paulo, Anais Infoimagem, Cenadem.

## **APÊNDICE A - PEQUENO GUIA PARA GERAÇÃO DE REGAS DE ASSOCIAÇÃO USANDO O RATTLE**

### **1. Introdução**

Rattle (R Analytical Tool To Learn Easily) é uma ferramenta de data mining que utiliza a interface gráfica Gnome. Baseado na linguagem R de código aberto, o Rattle é de fácil aprendizado e uso graças a sua interface simples e objetiva. Simplifica, sobremaneira, as tarefas de data mining, principalmente quando comparado com os procedimentos que seriam necessários caso fosse utilizado somente a linguagem R.

Vale lembrar que este é um projeto novo, ainda em fase de desenvolvimento e algumas de suas funcionalidades ainda não foram implementadas. Apesar de aparecerem na interface (como desabilitadas), não podem ser selecionadas.

Existem versões para diferentes sistemas operacionais, como Linux, Macintosh OS e Windows. Este guia foi criado a partir da versão para Windows e todos os testes foram realizados executando os programas a partir do Windows XP.



Para a geração de regras de associação, o programa utiliza o algoritmo apriori. A implementação deste algoritmo apriori usado pelo Rattle foi desenvolvida por Cristian Borgelt. Este programa vem sendo utilizado em vários softwares de Data Mining, incluindo comerciais, como Clementine (programa vendido à parte do pacote Estatístico SPSS).

Maiores informações sobre o apriori desenvolvido por Cristian Borgelt podem obtidas em <http://fuzzy.cs.uni-magdeburg.de/~borgelt/apriori.html>

Para mais detalhes sobre o Rattle, visite o site <http://rattle.togaware.com>

Para conhecer a linguagem R e seu projeto, acesse o site: <http://www.r-project.org>

## **2. Instalando o Rattle**

A instalação do Rattle é demorada e requer alguns cuidados. A seqüência de passos deve ser seguida à risca sob pena de não funcionar no final e sendo necessário reiniciar todo o processo.

Para instalação do Rattle, é fundamental consultar o roteiro descrito no site <http://datamining.togaware.com/survivor/Installation.html>

Em essência, são cinco passos para instalar a versão para Windows:

1. Instalação da biblioteca GTK+
2. Instalação da linguagem R
3. Instalação do pacote RGtk2
4. Instalação de pacotes adicionais para linguagem R
5. Instalação do pacote Rattle

As versões que foram instaladas para elaboração deste guia foram as seguintes:

Biblioteca Gtk+ → 2.10.6    Linguagem R → 2.4.0    Rattle → 2.1.92

**2.1. Os passos 1 e 2 são instalações de softwares no sistema Windows. Os demais (3, 4 e 5) são instalações efetuadas a partir da interface gráfica da linguagem R.**

É fundamental que ao final da instalação o pacote “arules” tenha sido instalado, pois este é o que provê a infra-estrutura necessária para manipulação dos dados e a geração das regras de associação. Caso ele não tenha sido instalado, basta digitar o comando a partir da interface do R:

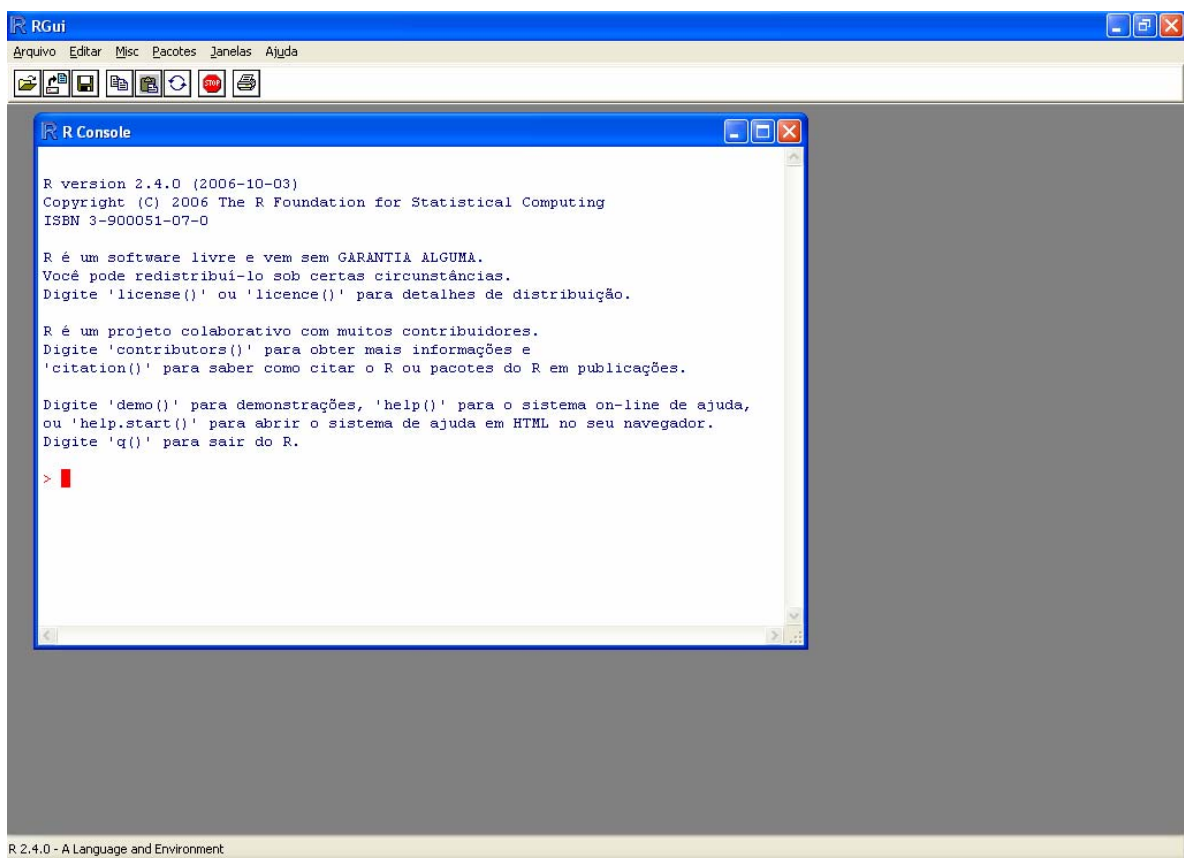
```
> install.packages (“arules”)
```

O R irá solicitar a escolha do local do download. Selecione o local e clique no botão OK.



### 3. Iniciando o Rattle

Para iniciar o Rattle é necessário antes abrir a interface gráfica da linguagem R executando um duplo clique no ícone do R que foi instalado no desktop do Windows. Aparecerá a tela a seguir:

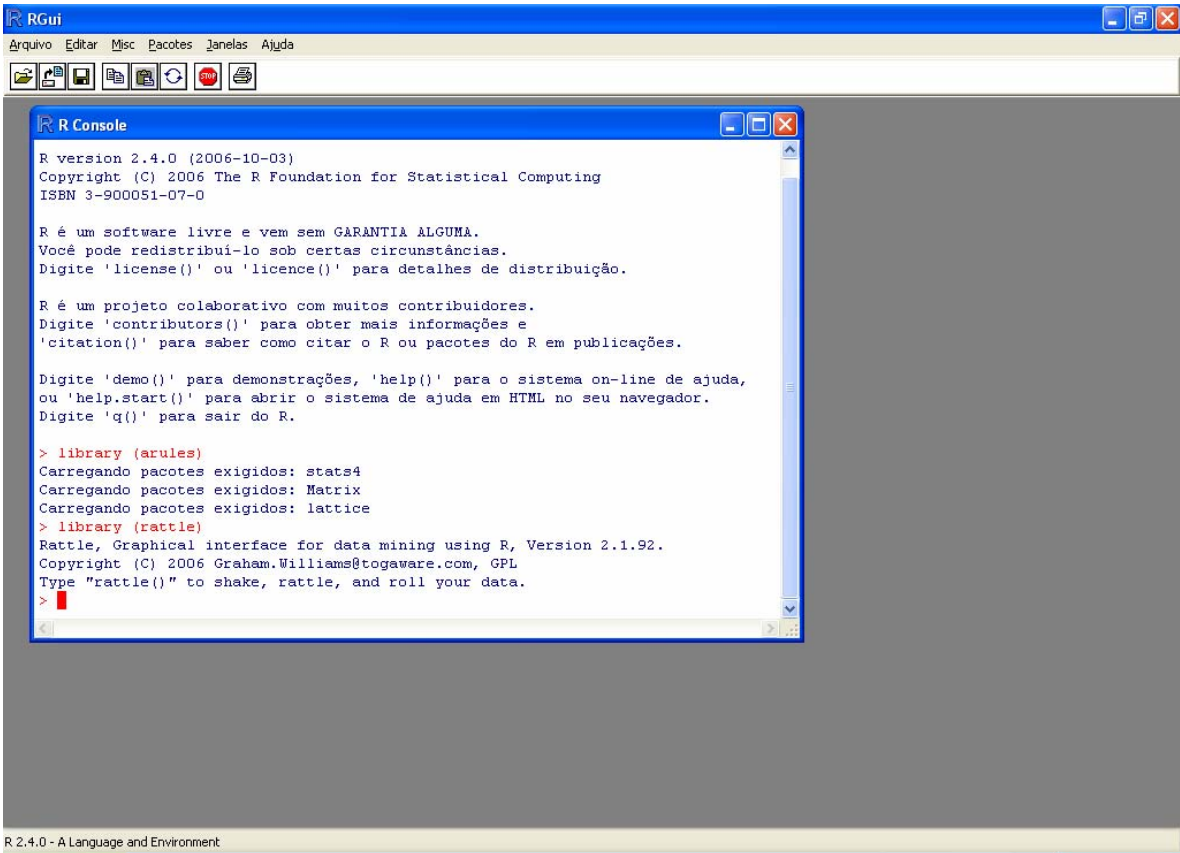


Após, devemos carregar o pacote de regras de associação (denominado “arules”) e o pacote do Rattle (denominado “rattle”), através do comando library, digitando:

```
> library(arules)
```

```
> library(rattle)
```

A tela a seguir exibe o resultado destes comandos, exibindo os pacotes que foram carregados.



```
R version 2.4.0 (2006-10-03)
Copyright (C) 2006 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R é um software livre e vem sem GARANTIA ALGUMA.
Você pode redistribuí-lo sob certas circunstâncias.
Digite 'license()' ou 'licence()' para detalhes de distribuição.

R é um projeto colaborativo com muitos contribuidores.
Digite 'contributors()' para obter mais informações e
'citation()' para saber como citar o R ou pacotes do R em publicações.

Digite 'demo()' para demonstrações, 'help()' para o sistema on-line de ajuda,
ou 'help.start()' para abrir o sistema de ajuda em HTML no seu navegador.
Digite 'q()' para sair do R.

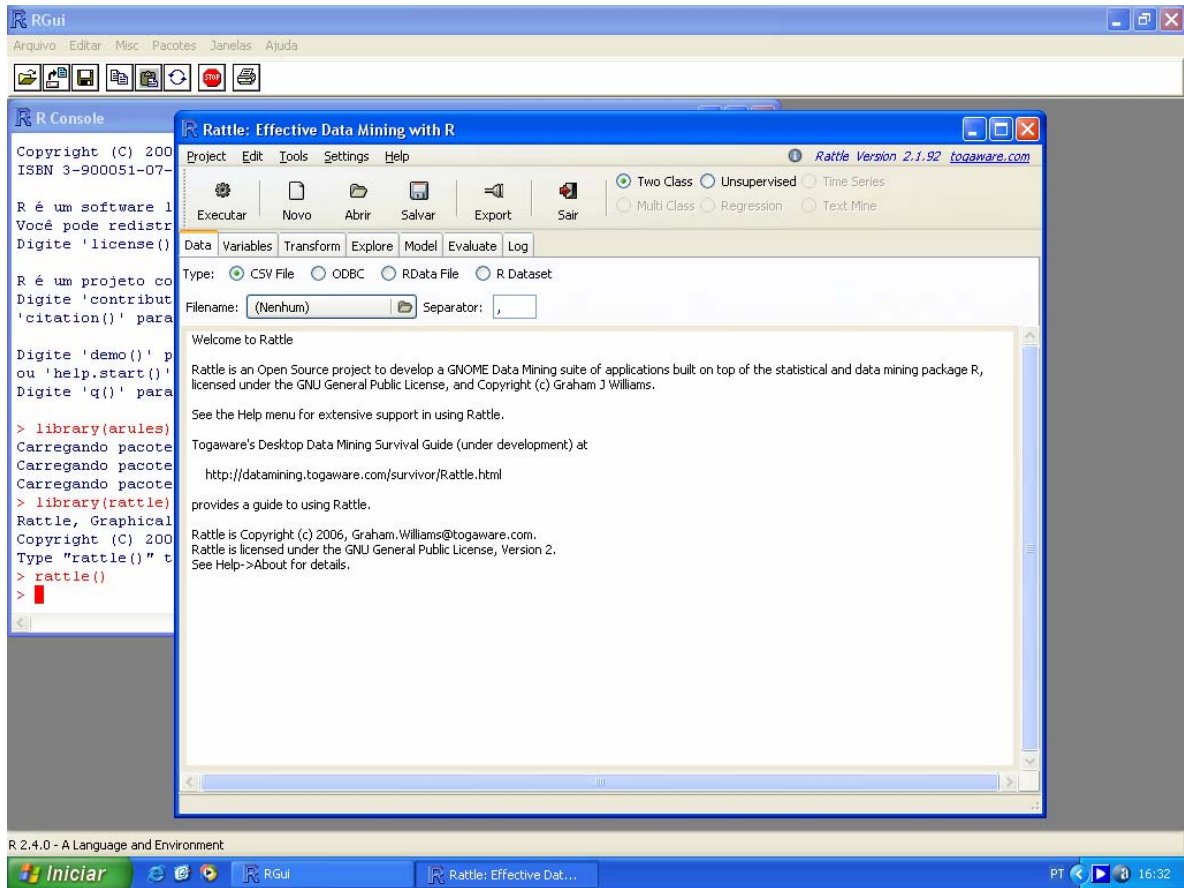
> library(arules)
Carregando pacotes exigidos: stats4
Carregando pacotes exigidos: Matrix
Carregando pacotes exigidos: lattice
> library(rattle)
Rattle, Graphical interface for data mining using R, Version 2.1.92.
Copyright (C) 2006 Graham.Williams@togaware.com, GPL
Type "rattle()" to shake, rattle, and roll your data.
>
```

Agora, basta iniciar o Rattle, executando o comando Rattle().

```
> rattle()
```

A tela inicial do programa irá aparecer, exibindo sua interface gráfica.

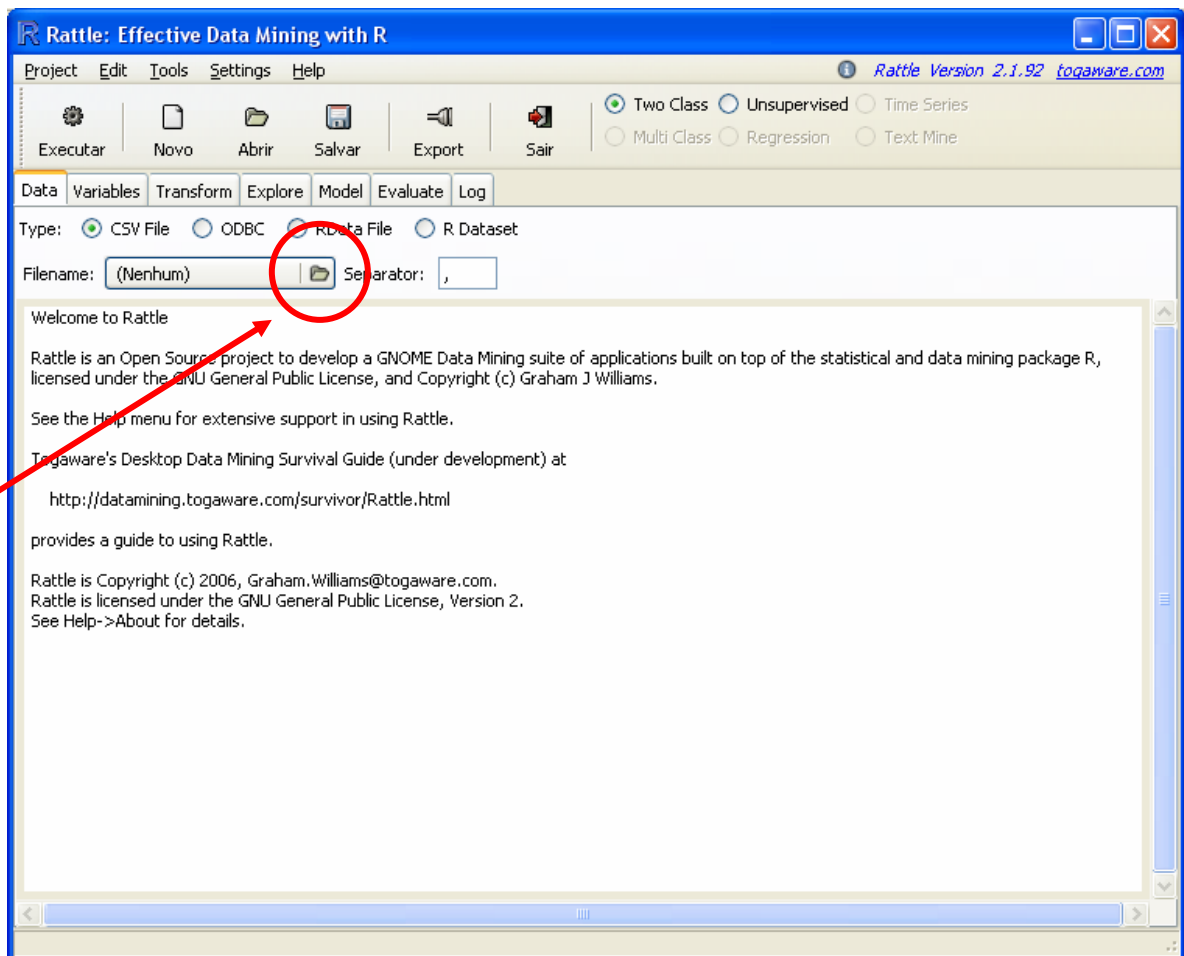
Uma vez que o pacote executa a partir da interface gráfica da linguagem R vale lembrar que algumas mensagens do Rattle serão apresentadas na interface RGui/RConsole.



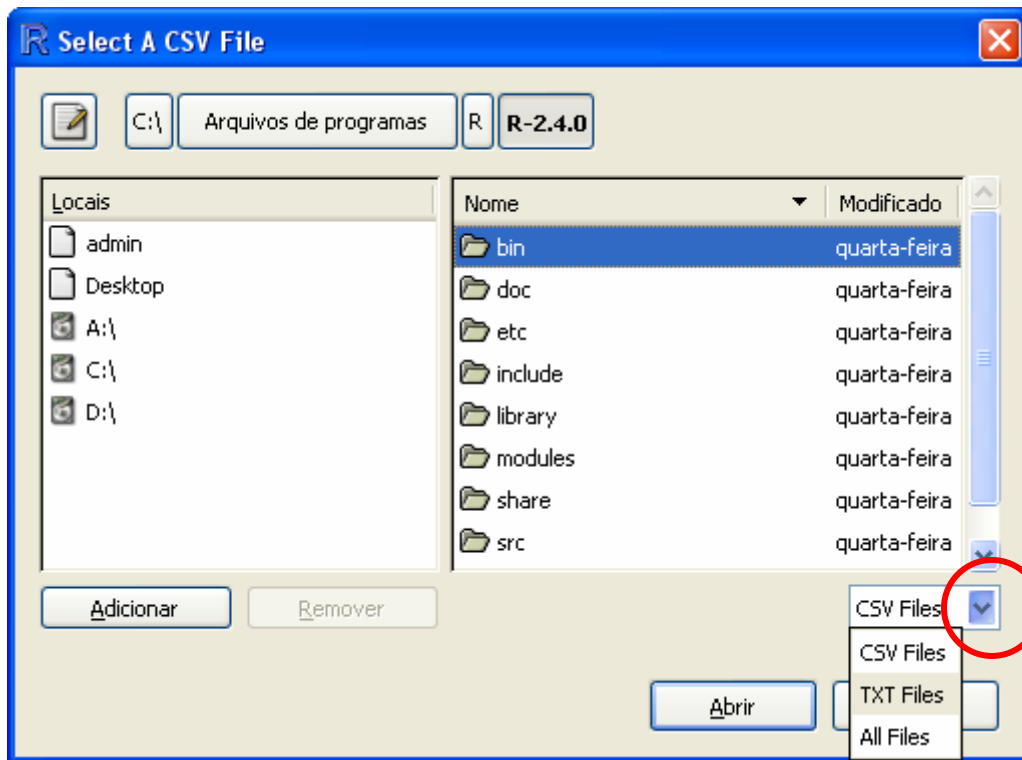
#### 4. Carregando um arquivo

Diversos formatos de arquivos podem ser utilizados para entrada de dados no Rattle. O tipo de arquivo deve ser indicado na tela principal podendo ser selecionado entre CSV , ODBC, Rdata File e R Dataset. O default é o tipo CSV.

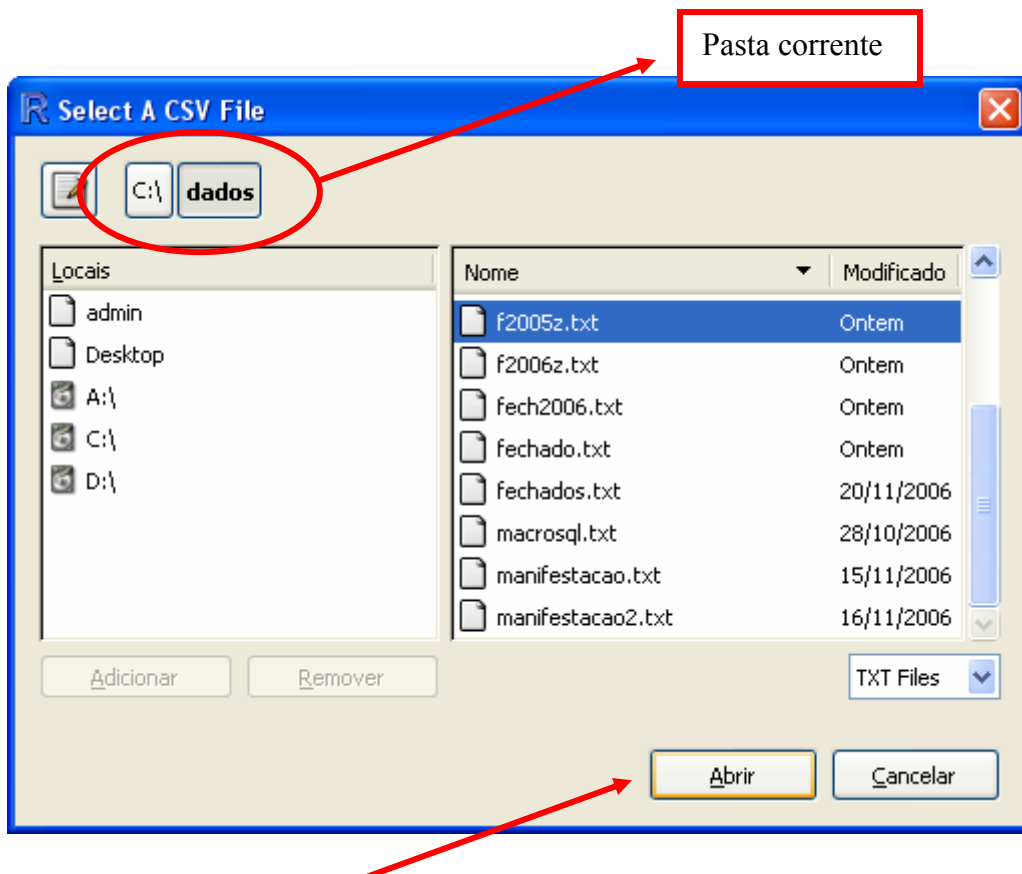
Como exemplo, descreveremos a carga de um arquivo texto (extensão TXT), com os campos dos registros separados por vírgulas. É importante ressaltar que o primeiro registro do arquivo texto deverá conter o nome de cada uma das variáveis, separado por vírgulas.



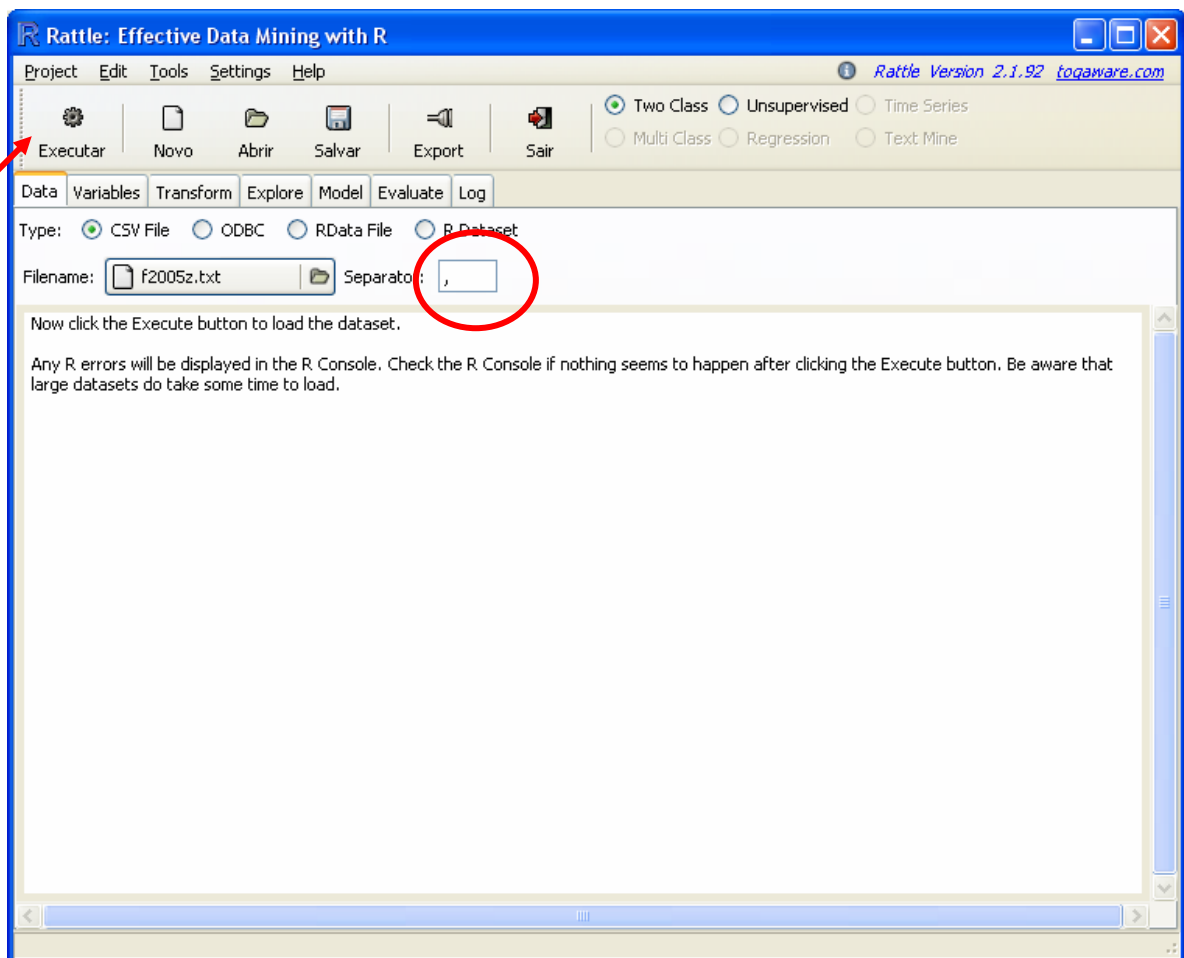
A caixa de diálogo para seleção do arquivo será aberta. Seleccionamos primeiro o tipo de arquivo (no caso TXT), como mostra a figura. Após, seleccionamos a unidade, o diretório e o arquivo que será objeto de análise.



Selecionado o arquivo, clicamos no botão Abrir, como segue.



O próximo passo será a seleção do separador de campos. Normalmente, os arquivos TXT ou CSV utilizam a vírgula para separar os campos (variáveis). O Rattle aceita outros separadores, como por exemplo, o ponto e vírgula (;). Neste exemplo, vamos manter o separador default e apenas clicar no botão Executar para iniciar a carga do arquivo.

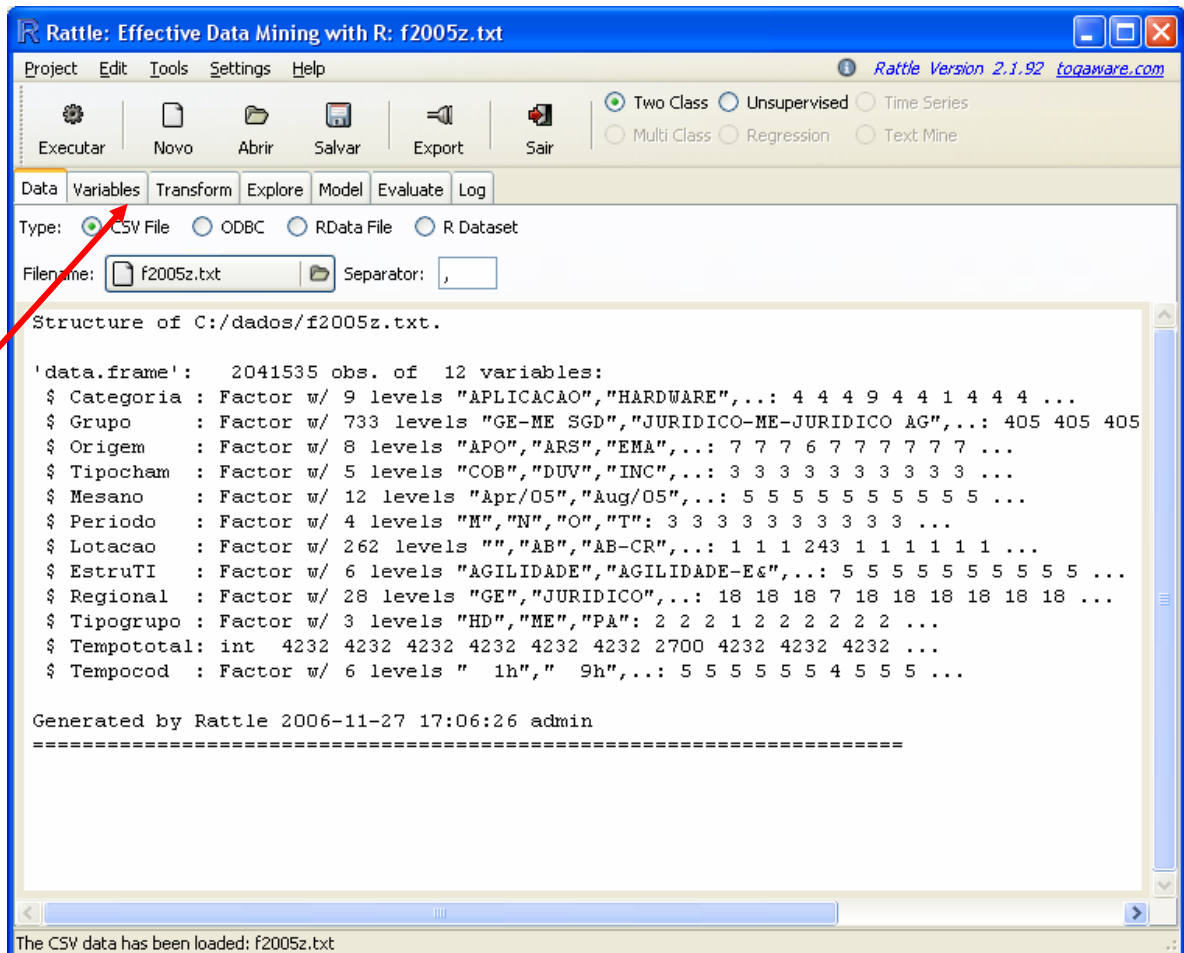


O tempo de carregamento do arquivo depende da estrutura, do tamanho do arquivo e evidentemente do computador onde está sendo executado. O Rattle suporta arquivos grandes com muitos registros (o exemplo usado aqui tem mais de 2 milhões de linhas), todavia há um grande consumo de memória. Nestes casos é importante monitorar o consumo de memória através do gerenciador de tarefas do Windows.



Quando o Rattle está operando, a tradicional ampulheta do Windows é substituída por um pequeno relógio de pulso.

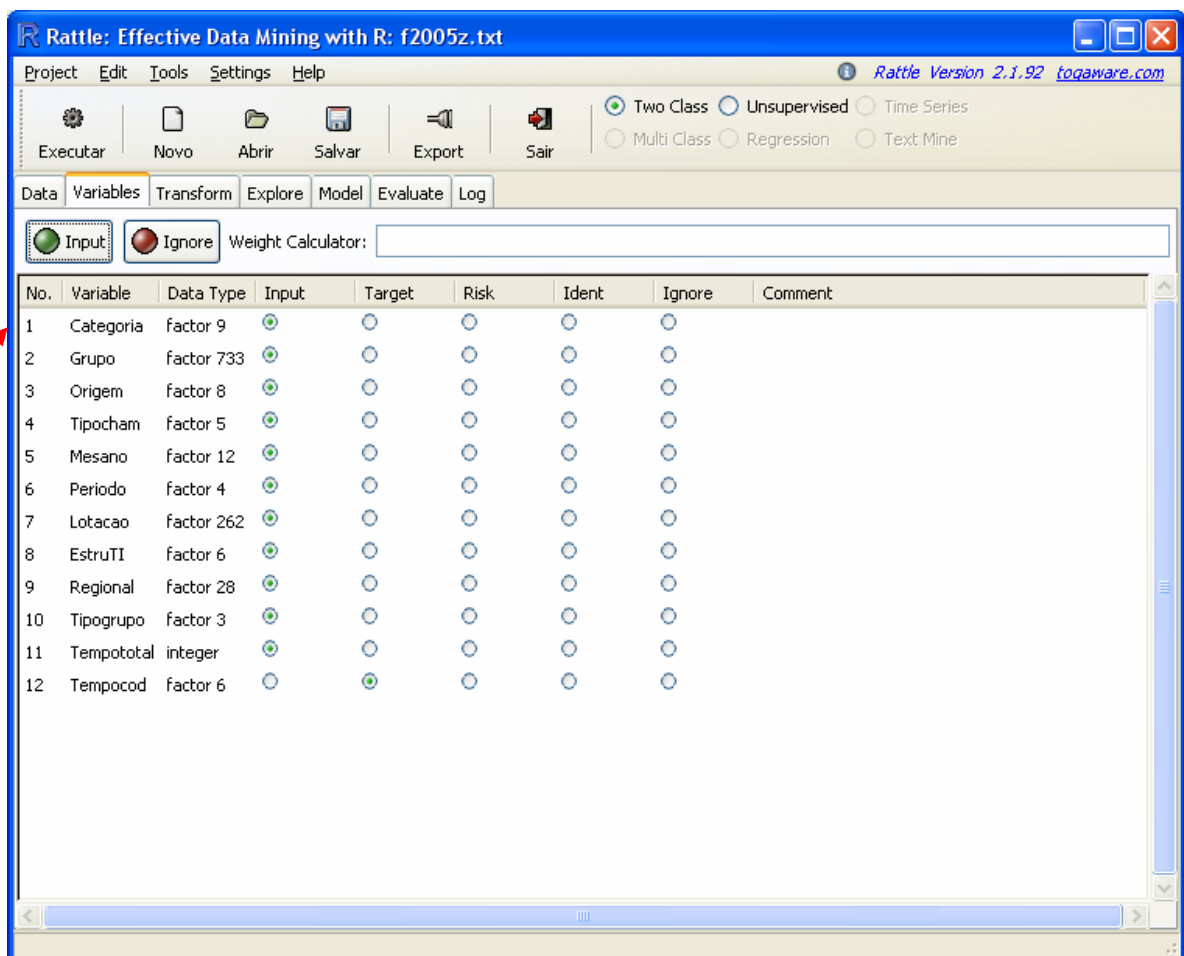
Terminado o carregamento do arquivo será exibida a tela a seguir. São listadas cada uma das variáveis (Categoria, Grupo, Origem etc), o número de observações (2041535), o número de variáveis(12), o tipo da variável (Factor, int), o número de exemplos de cada variável (“levels” – 9 para Categoria, 8 para Origem, etc) e alguns valores que cada uma das variáveis pode assumir (Categoria – “ALICACAO”, “HARDWARE”, etc).

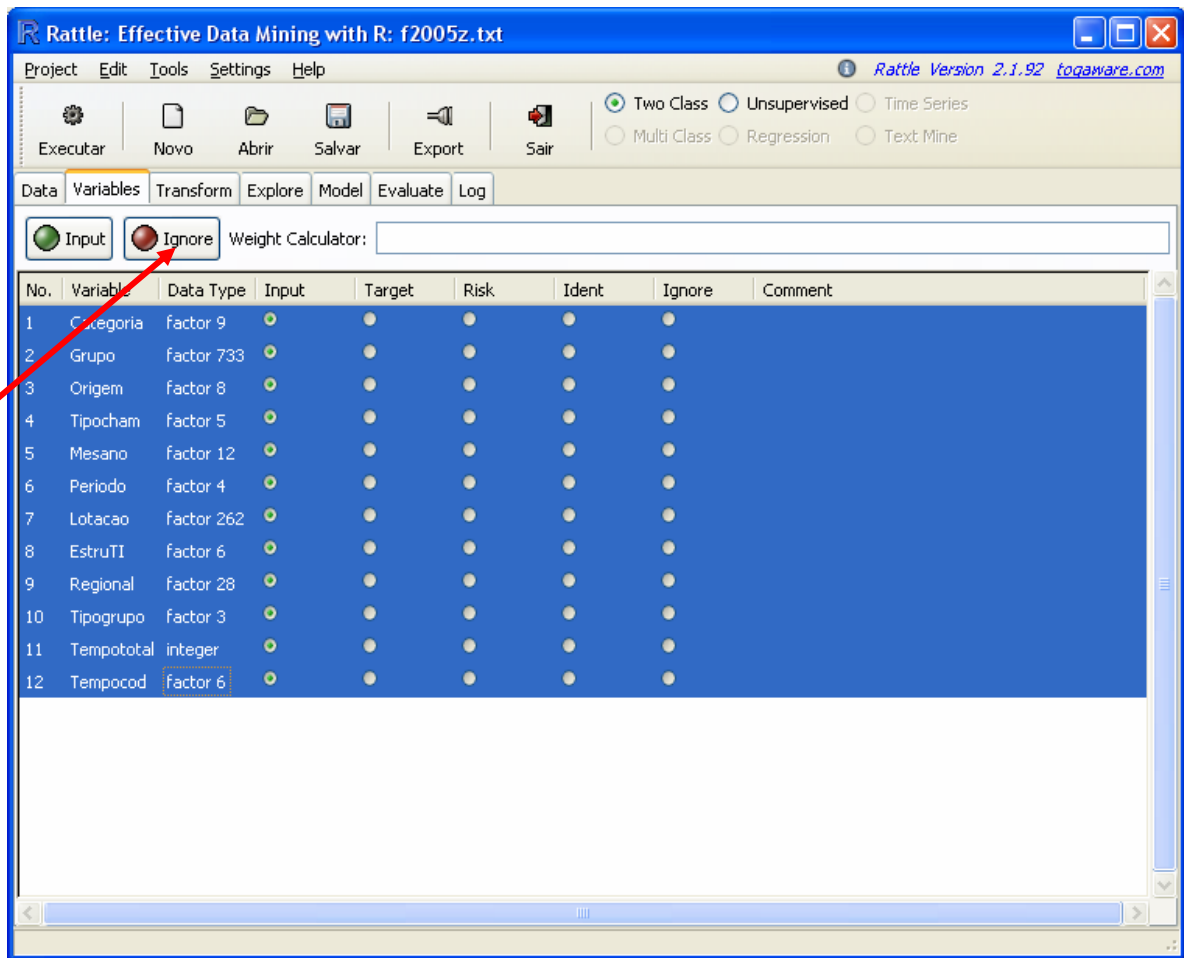


A etapa seguinte é a seleção das variáveis que serão usadas para a geração das regras de associação. Para um exame melhor das variáveis, clicamos na Guia “Variables”, como apontado na seta vermelha da figura acima.

## 5. Selecionando as variáveis que serão analisadas

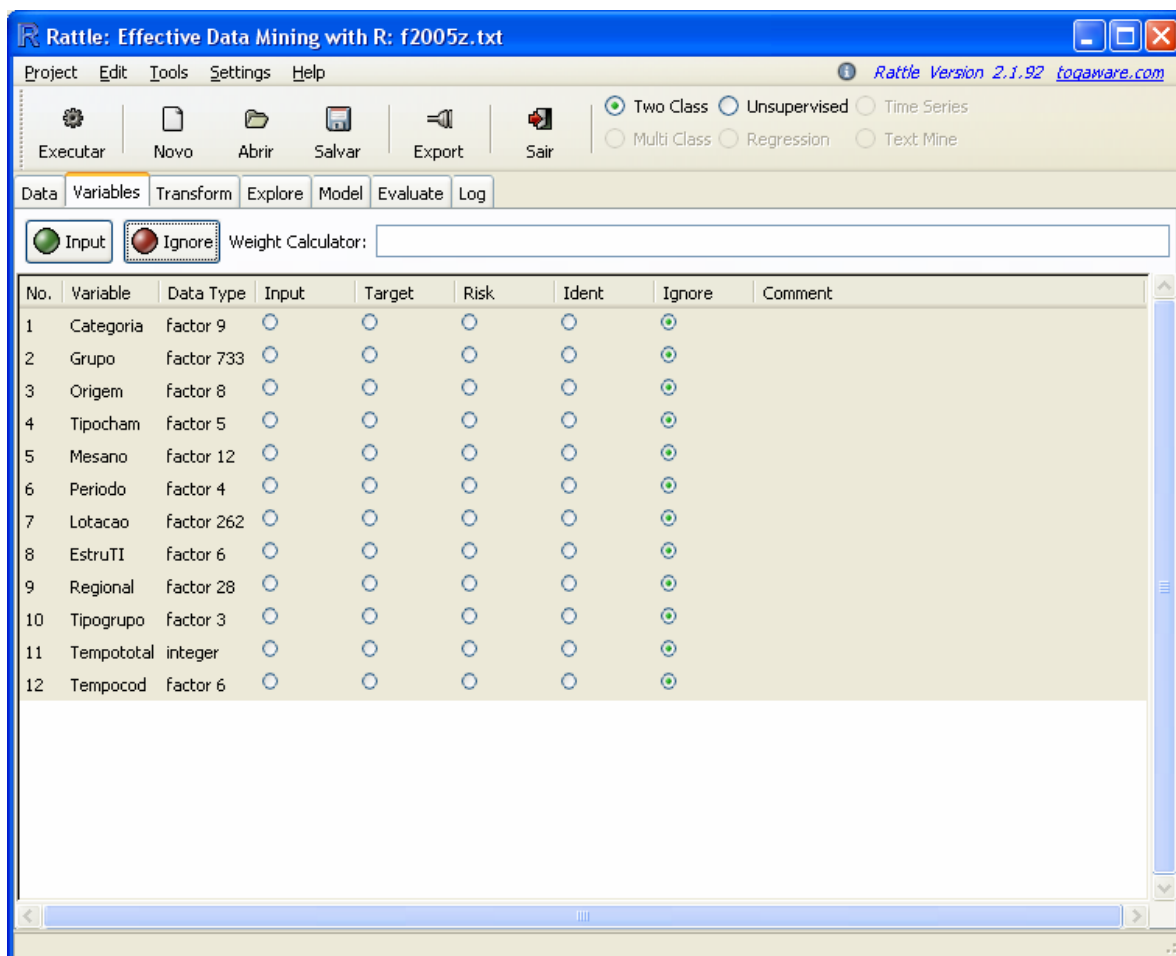
Por default, todas as variáveis aparecem selecionadas como “Input”. Como neste exemplo vamos trabalhar com regras de associação, selecionaremos apenas duas variáveis de entrada. Para isso, clicamos na primeira variável (1 Categoria) e mantendo a tecla shift e a seta para baixo pressionadas selecionamos todas as variáveis.





Após a seleção de todas variáveis, clicamos então, no botão “Ignore”, conforme figura acima, de forma que todas as variáveis mudem seu status para ignoradas.

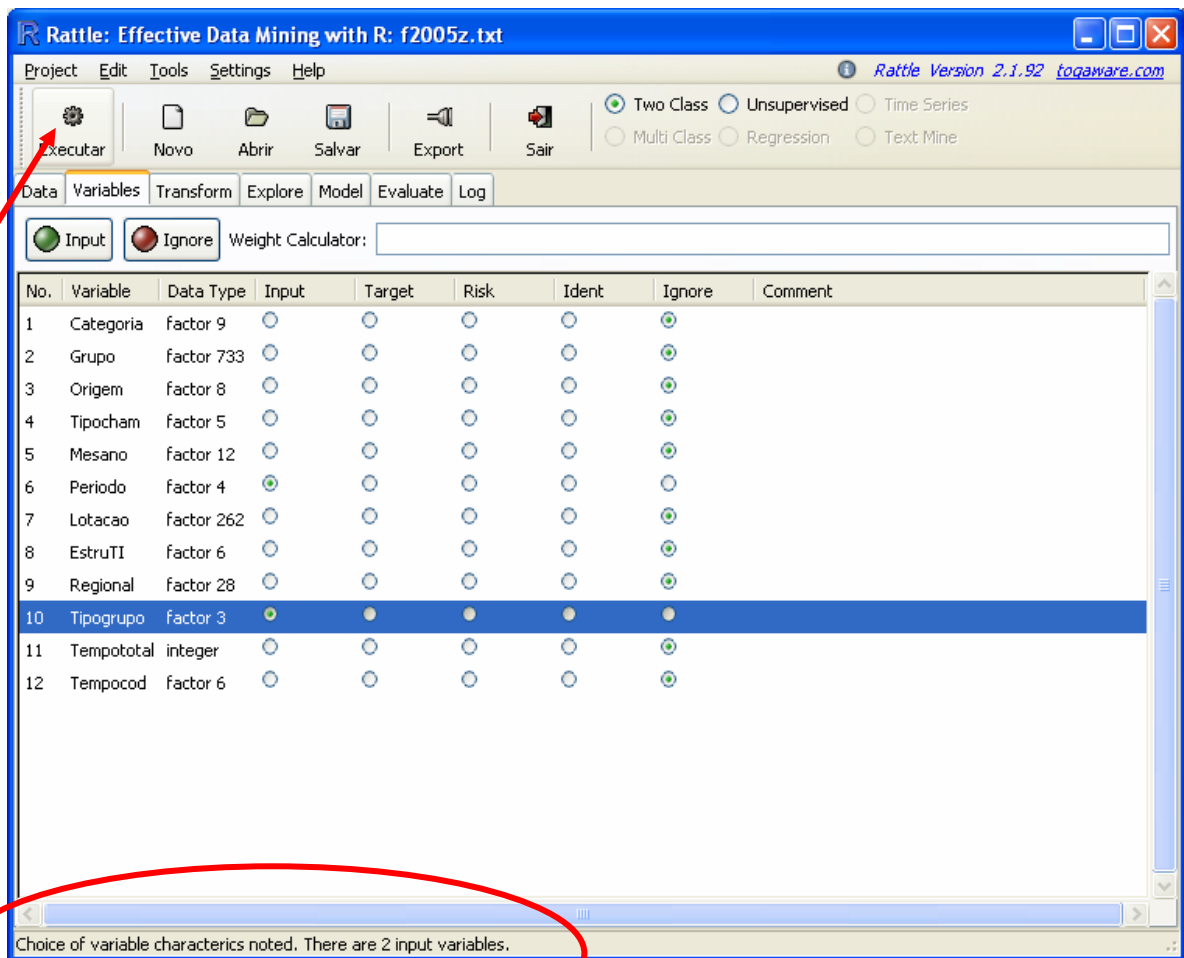
A tela abaixo todas as variáveis, já com status de Ignoradas. Esta manipulação do status pode ser feita caso a caso, clicando-se no item corresponde ao status que se quer atribuir para cada uma das variáveis individualmente.



Selecionamos então apenas as variáveis “Período” e “Tipogrupa”, marcando-as na coluna Input e em seguida, clica-se no botão “Executar” como mostrado na tela abaixo.

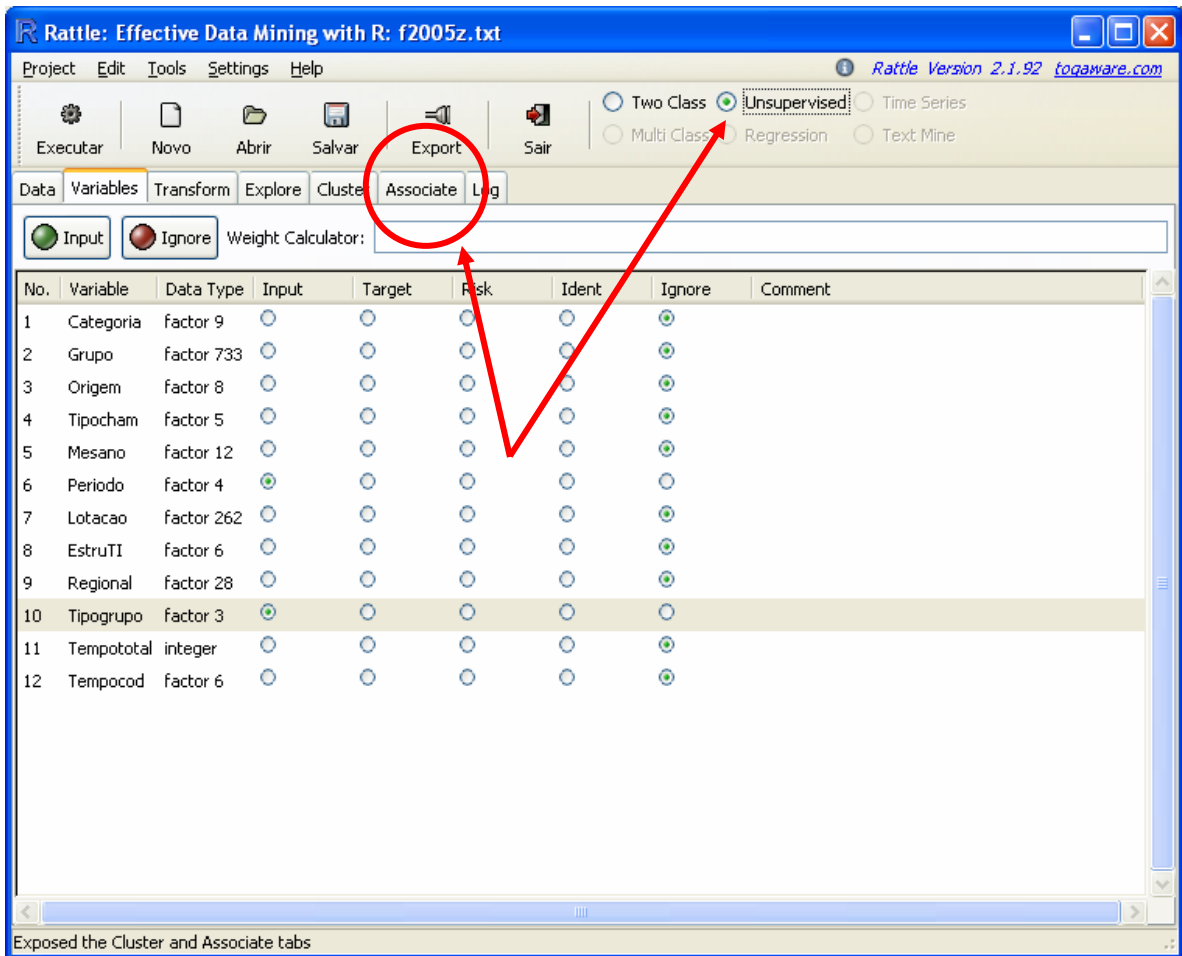
Observe a mensagem no rodapé da janela indicando que as duas variáveis foram selecionadas como “Input” (variáveis de entrada).

No Rattle o botão “Executar” tem fundamental importância. Toda e qualquer alteração nos parâmetros do Rattle ou em suas variáveis devem ser seguidas do botão “Executar” para que as mesmas sejam processadas.



## 6. Gerando as Regras de Associação

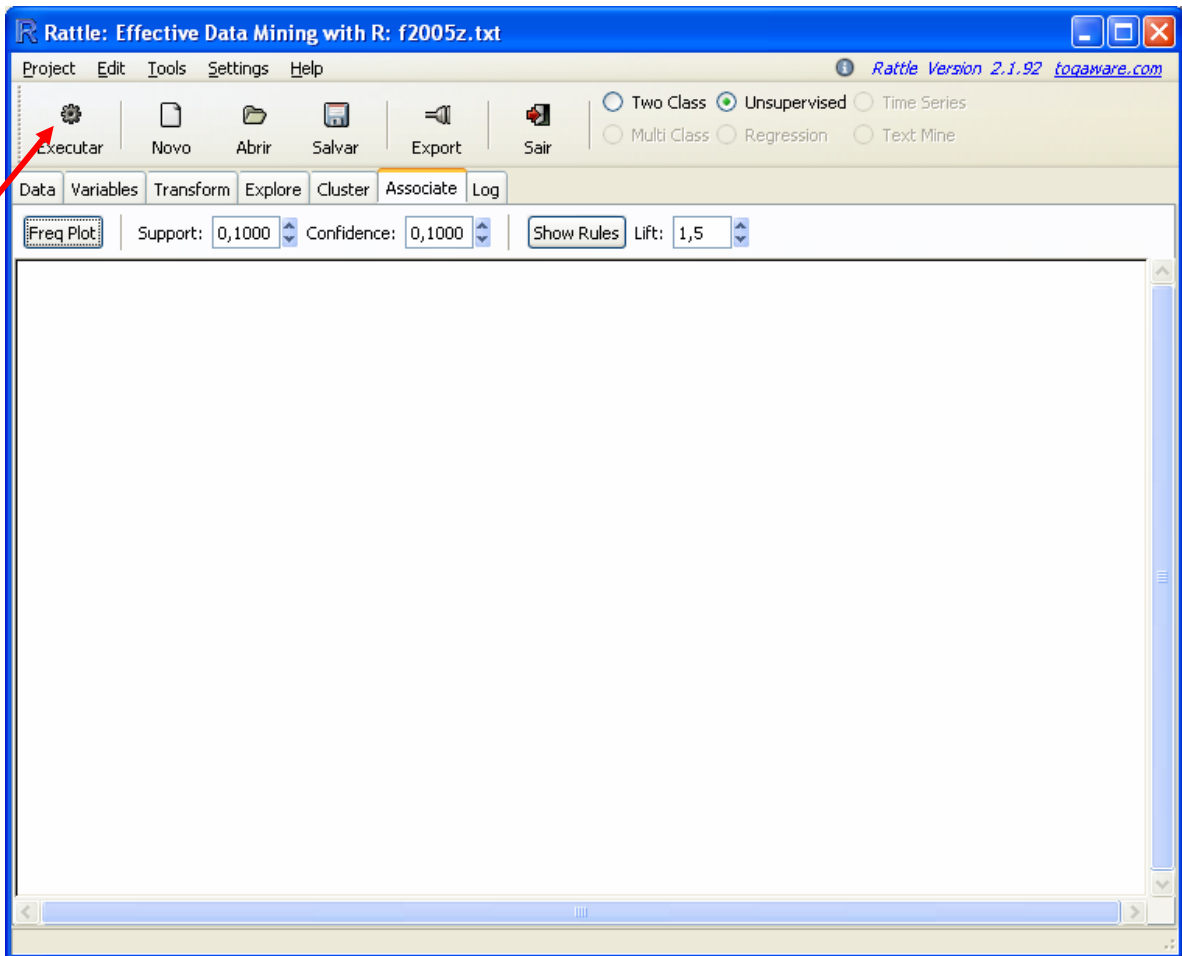
O Rattle permite trabalhar com muitas tarefas de data mining. O default é o modelo “Two Class”. Com este modelo selecionado a guia para geração de regras de associação fica oculta. Desta forma, devemos selecionar o método “Unsupervised”, para que a guia “Associate” apareça, conforme abaixo:



Com variáveis de entrada já definidas e o método escolhido, o próximo passo é a geração das Regras de Associação. Clicamos então, na Guia “Associate”. Os parâmetros de Suporte, Confiança e Lift podem ser configurados, clicando-se nas setas ao lado de cada um destes parâmetros, aumentando ou diminuindo os valores default.

Neste exemplo foi mantido o parâmetro default.

Clicamos no botão “Executar” para que o programa apriori seja iniciado. Dependendo do número de variáveis de entrada, da máquina onde está sendo feito o processamento e do tamanho do arquivo, a geração de regras pode demorar um pouco para apresentar o resultado.

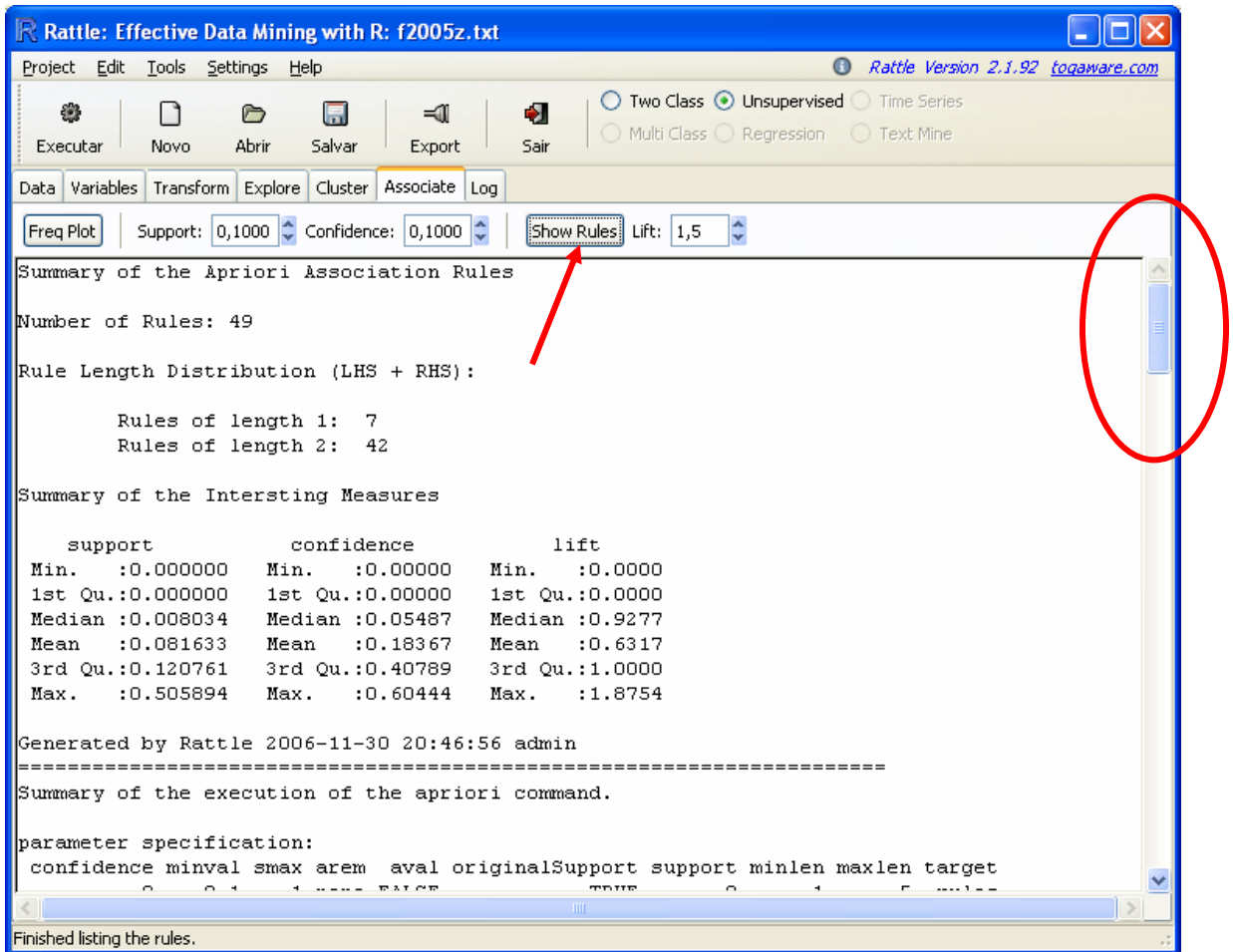


A tela seguinte apresenta o resultado da execução do programa apriori de Cristian Borgelt . Um relatório exhibe várias informações sobre o processamento, como o número de regras de associação que foram encontradas, valores mínimos, máximos e médios para os parâmetros de suporte, confiança e lift dentre outras informações.

Para exibição das regras que foram geradas, no entanto, se faz necessário clicar no botão “Show Rules”, conforme imagem abaixo.

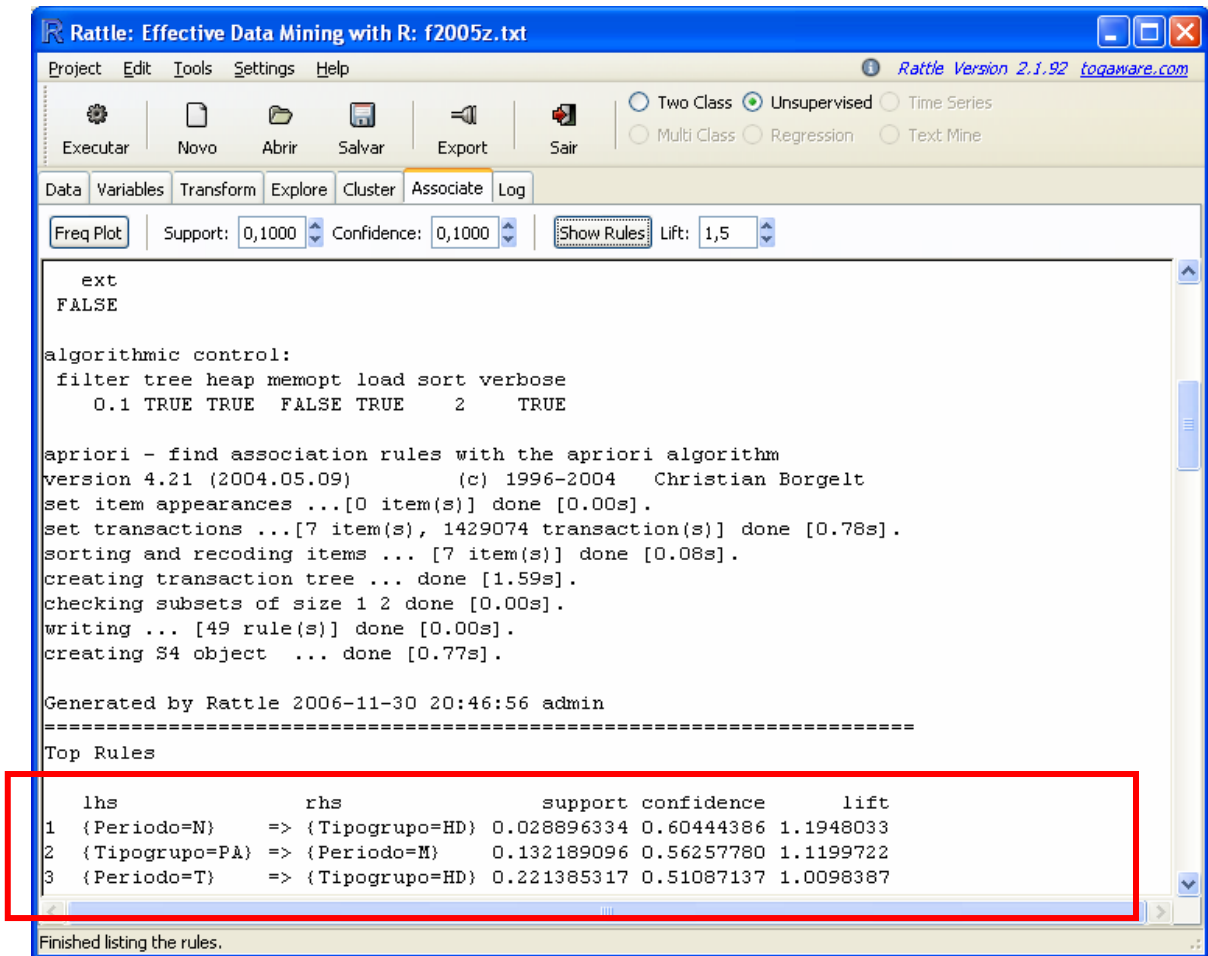
Observe que neste caso não é necessário clicar no botão “Executar”.

## 7. Exibindo as Regras de Associação geradas



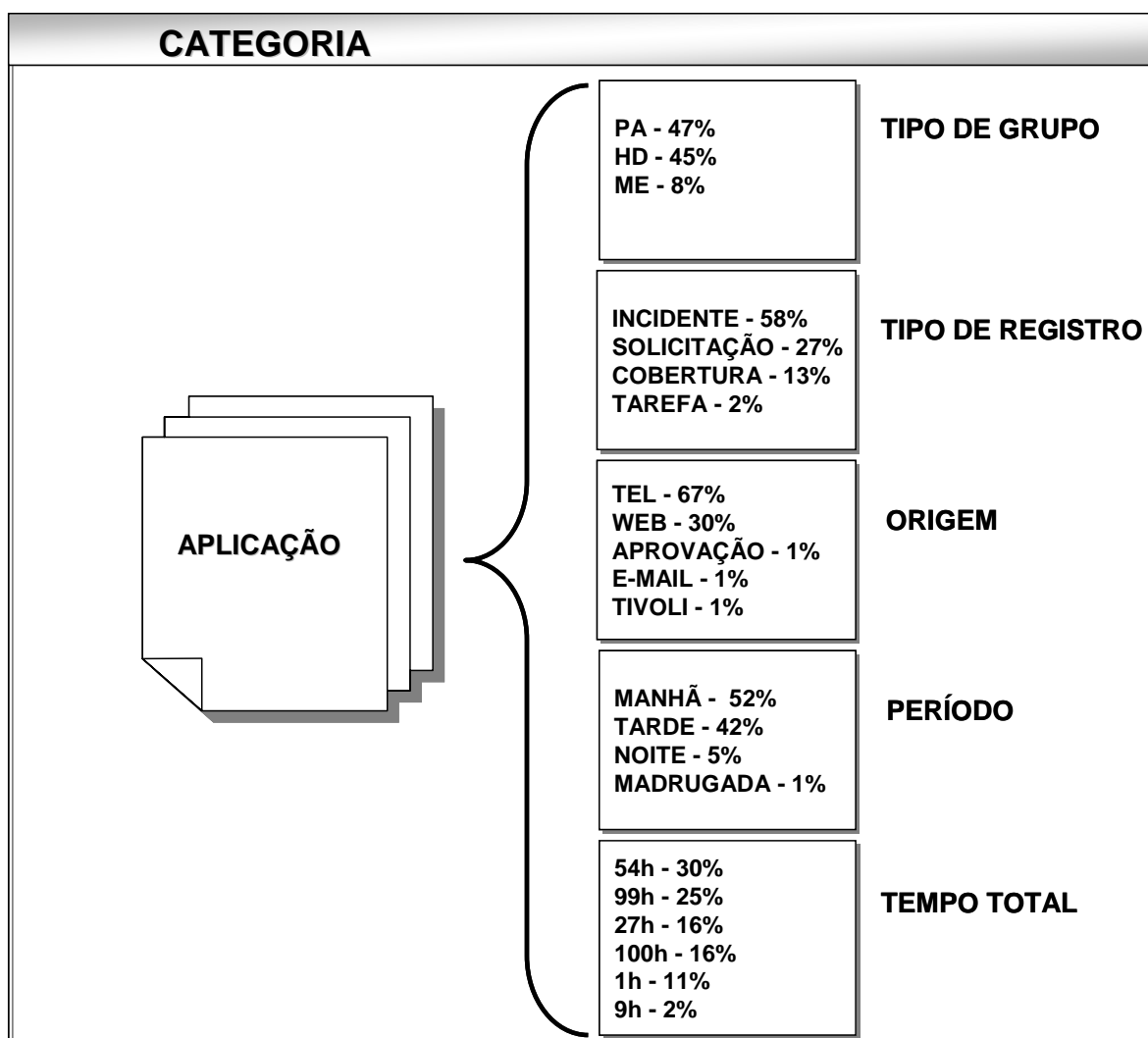
Após execução deste comando, a barra de rolagem altera de posição pois foi acrescentado ao relatório exibido anteriormente as regras de associação que foram geradas. Para exibi-las, use a barra de rolagem para baixo.

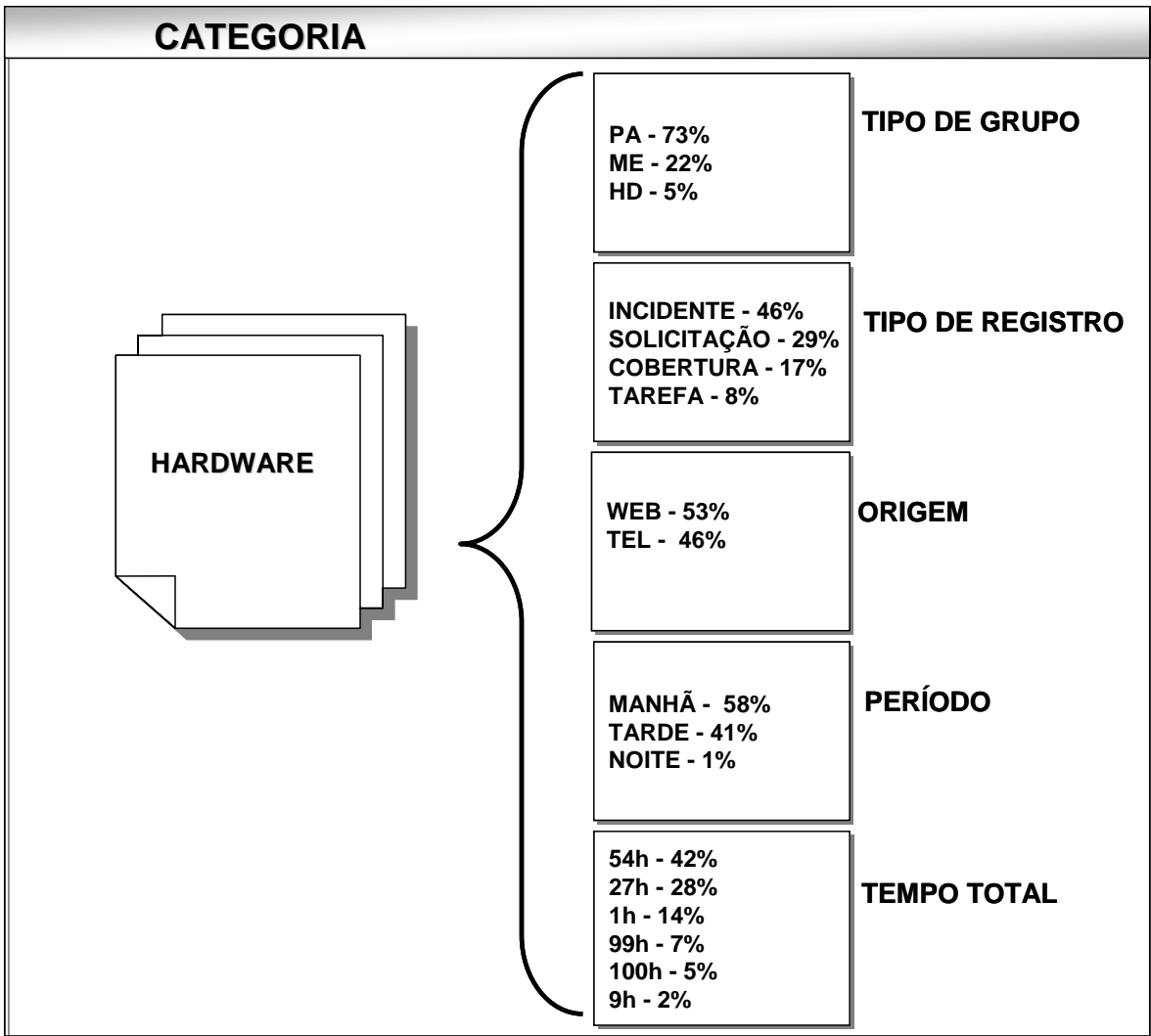


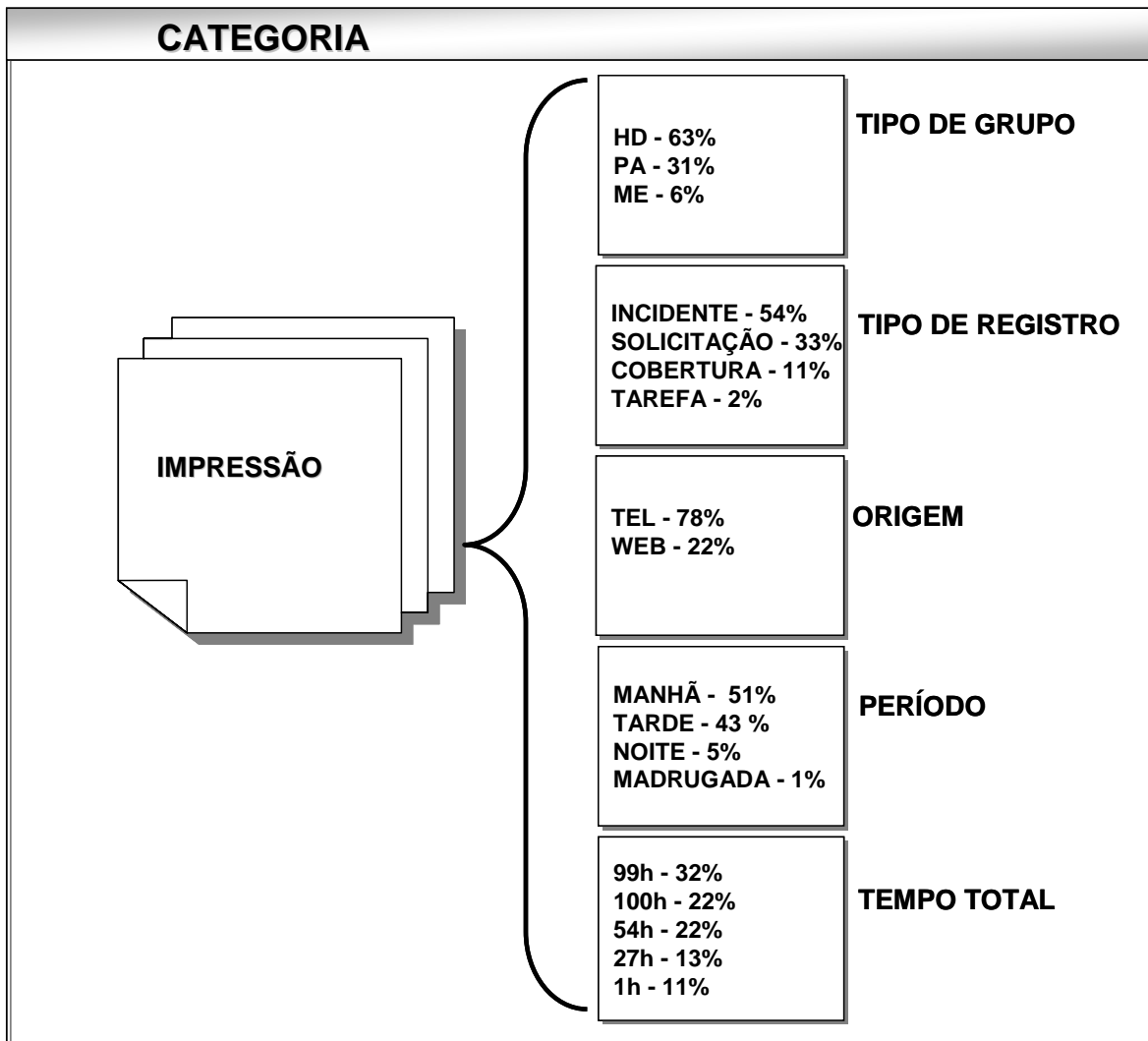


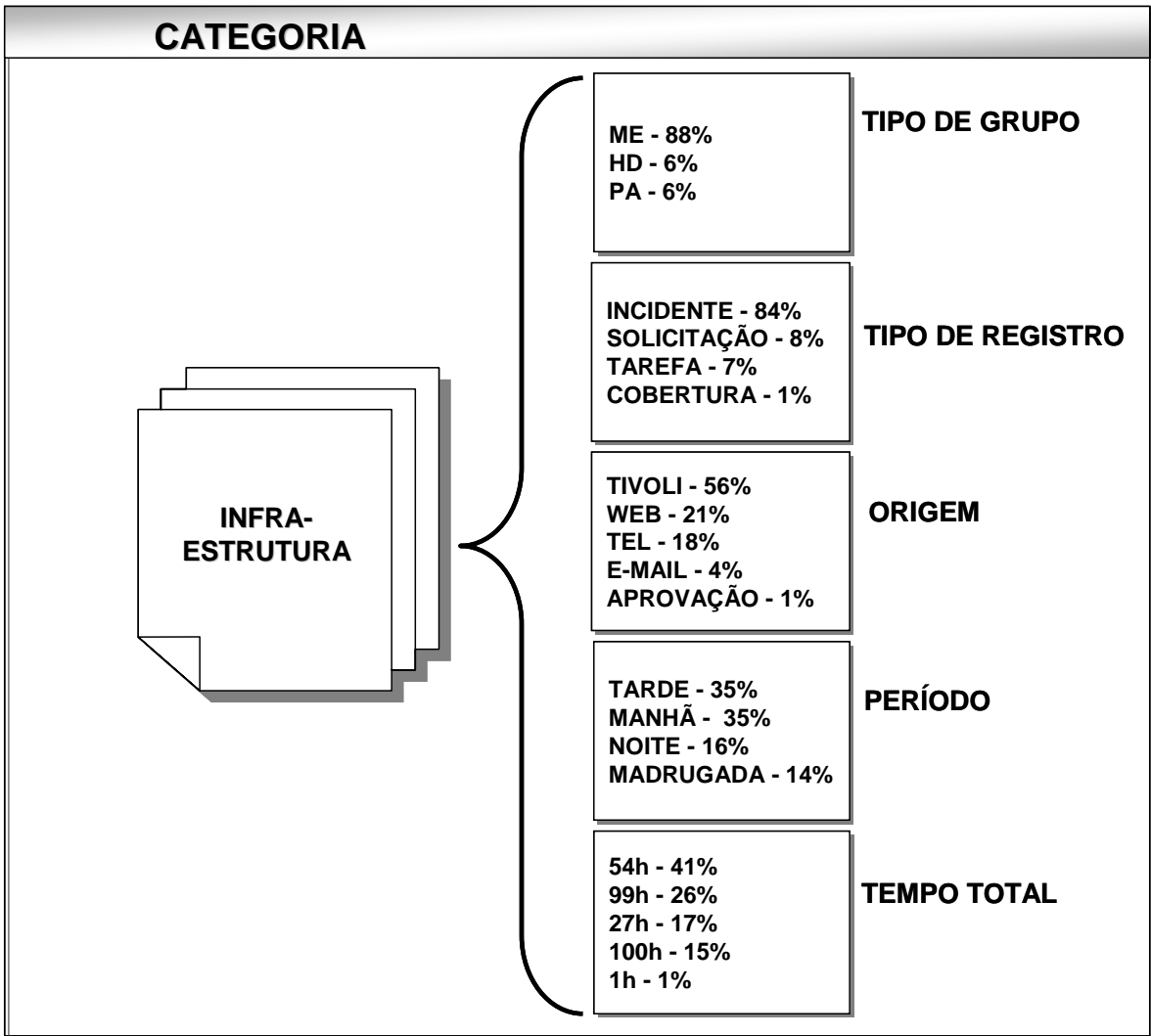
A figura acima exhibe em destaque parte das regras geradas, o item antecedente e conseqüente, bem como os parâmetros de suporte, confiança e lift para cada uma das regras encontradas.

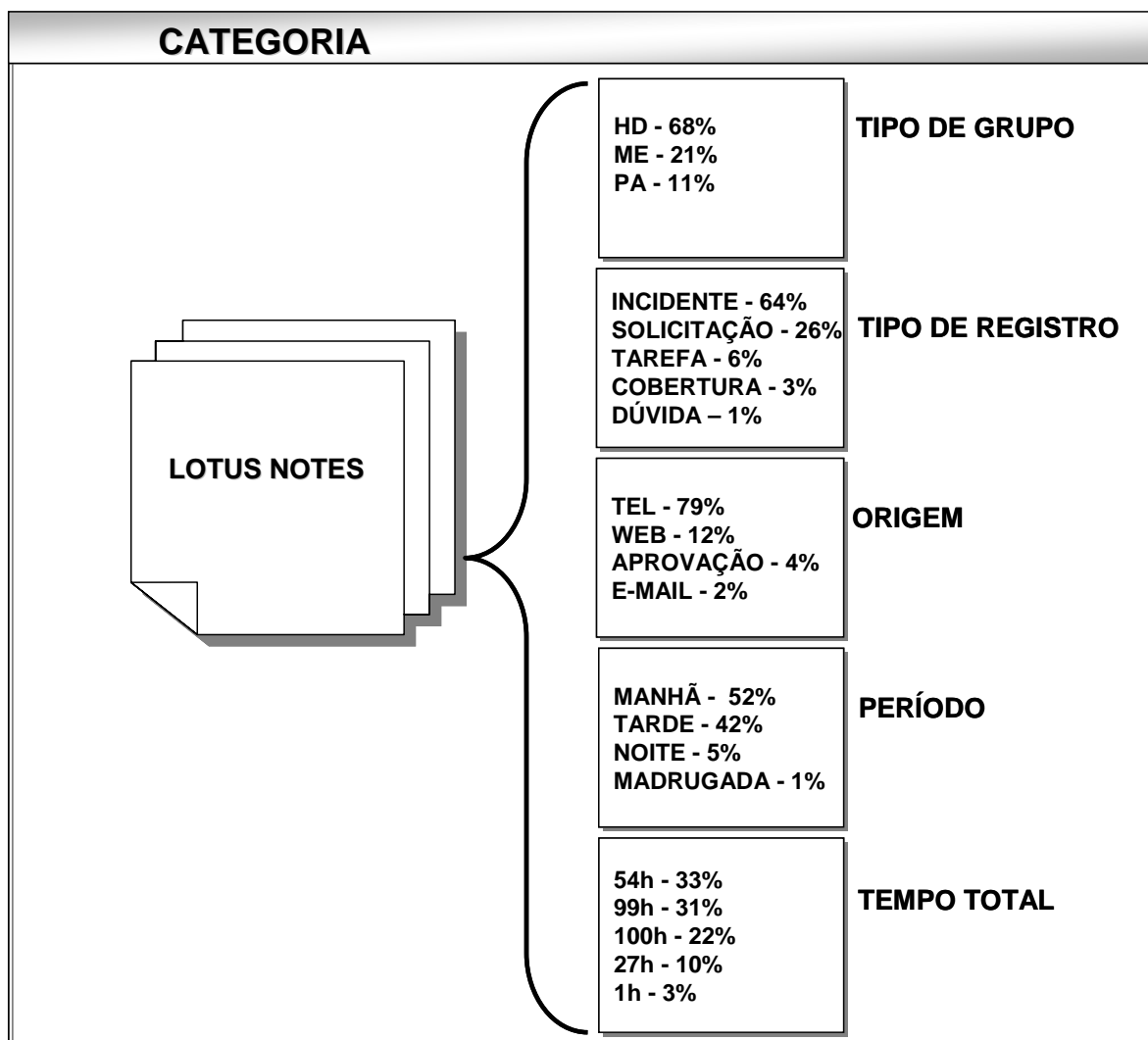
## APÊNDICE B – REGRAS CONSOLIDADAS ATRIBUTO – CATEGORIA

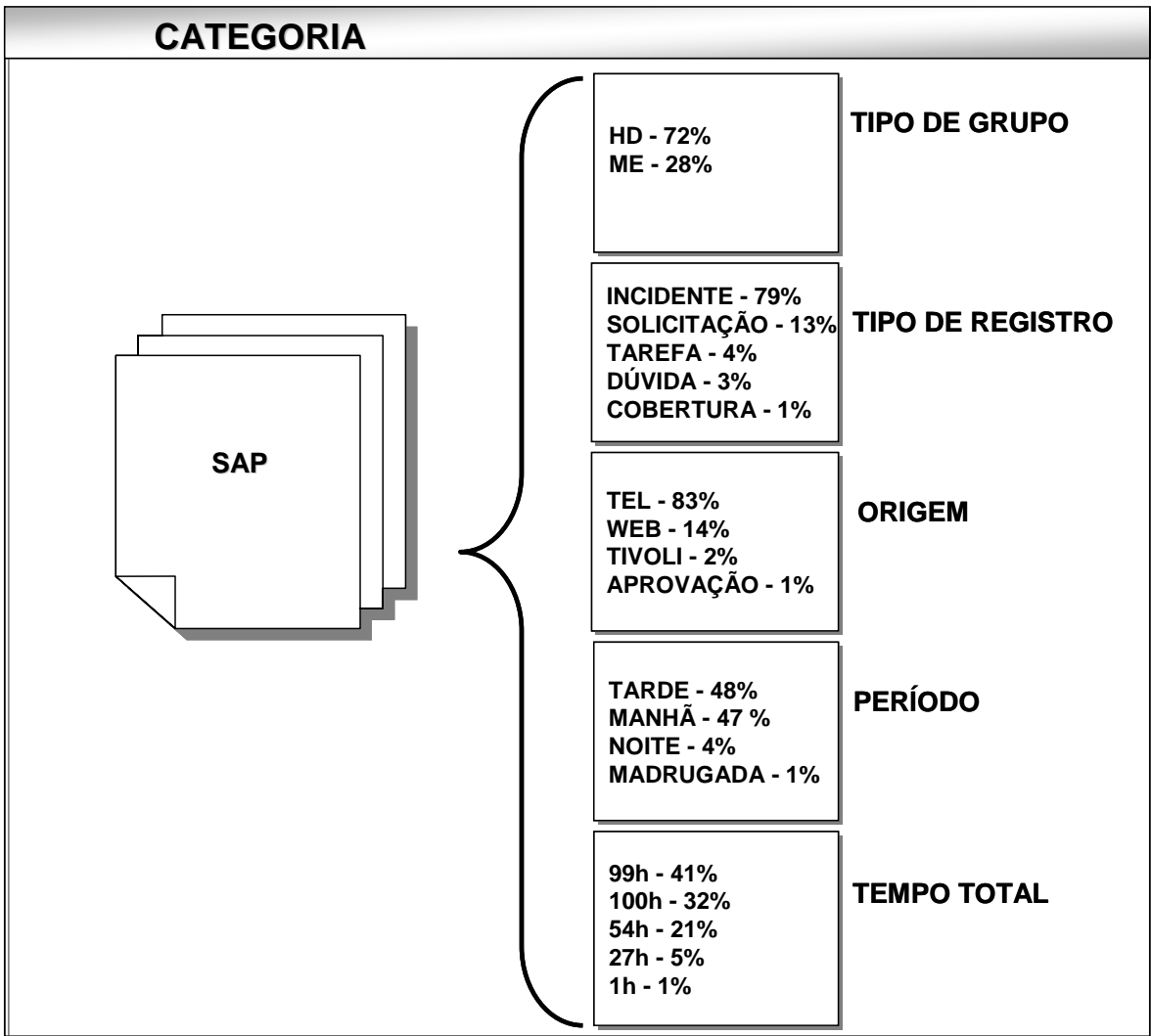


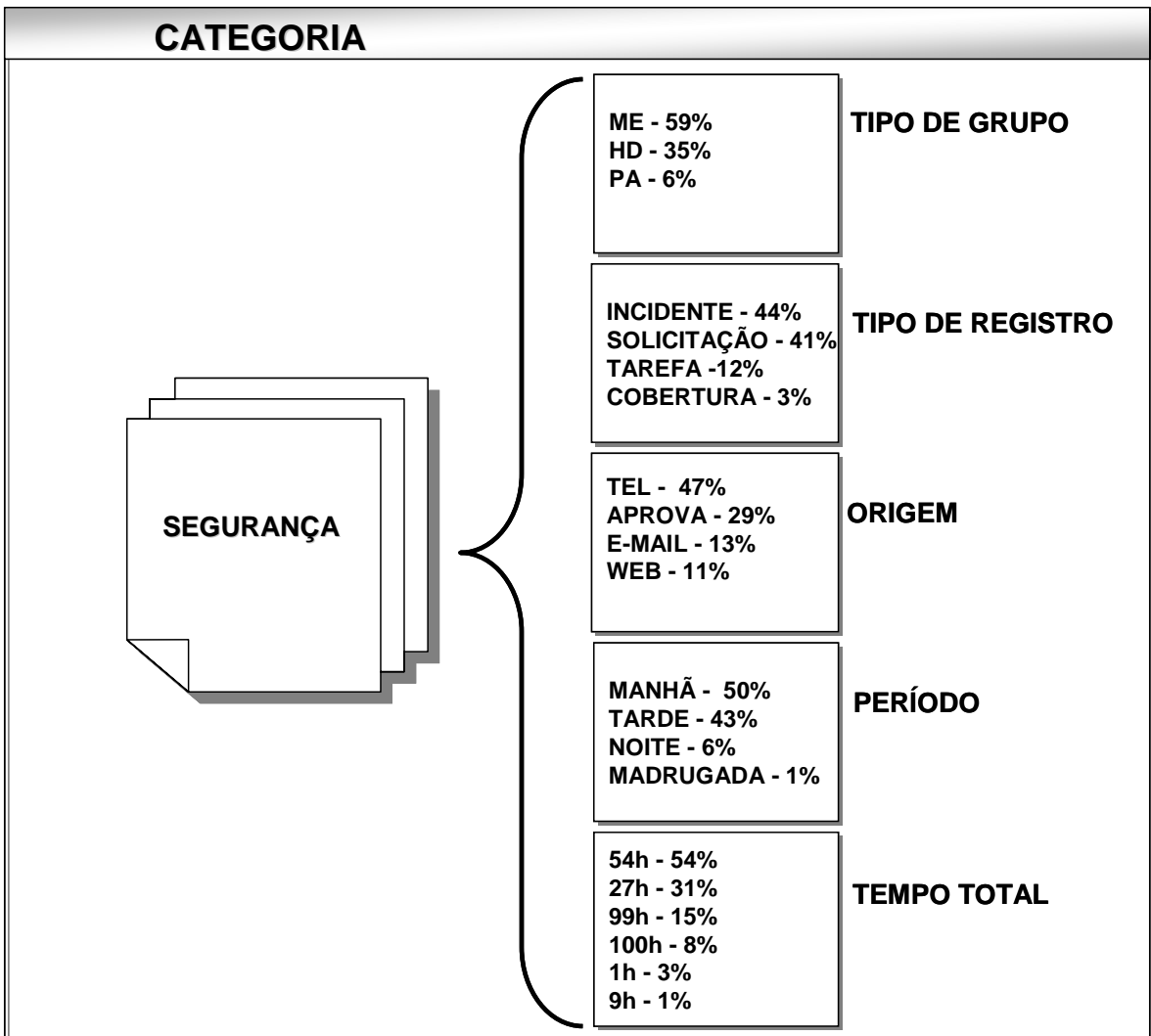




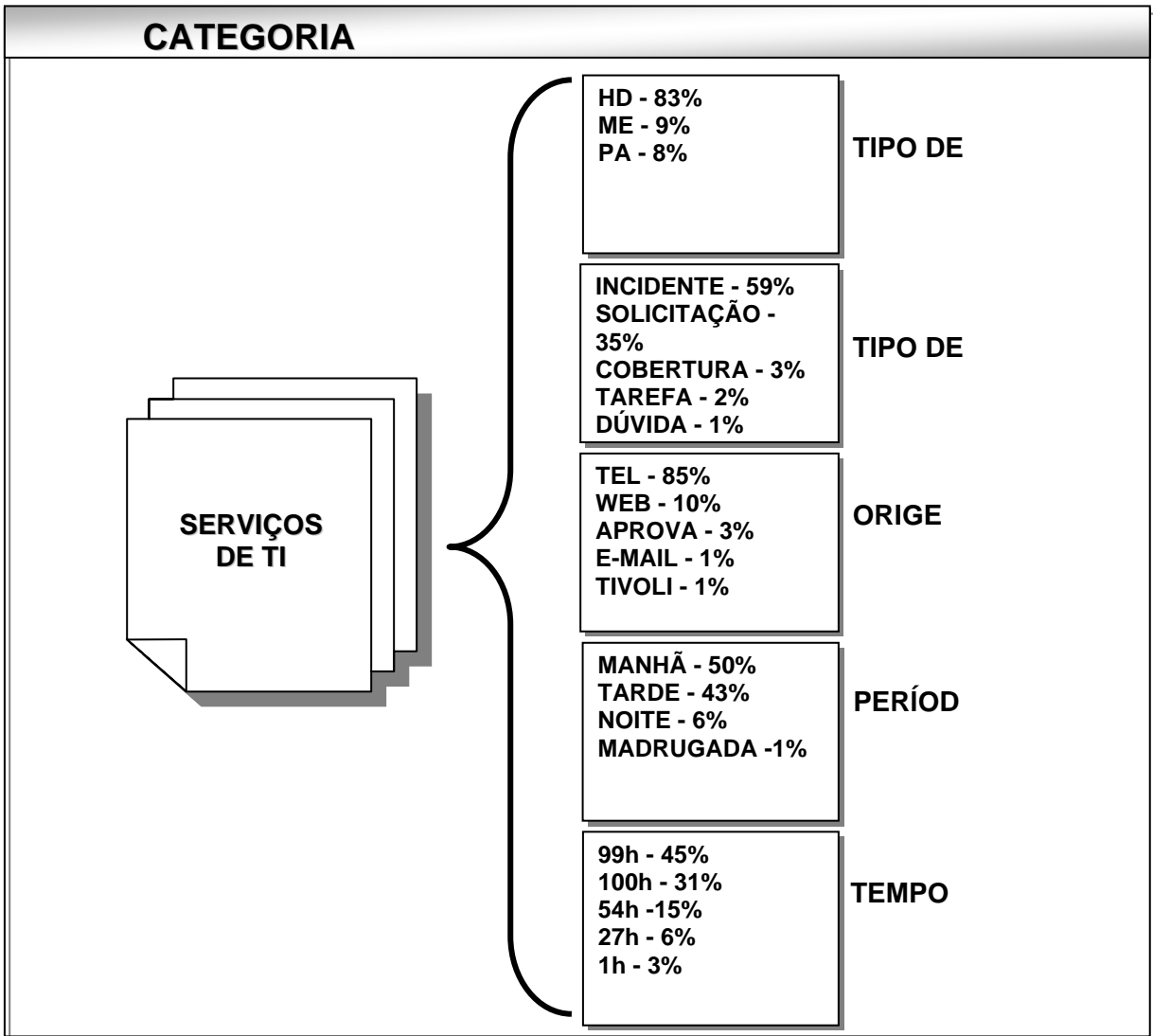


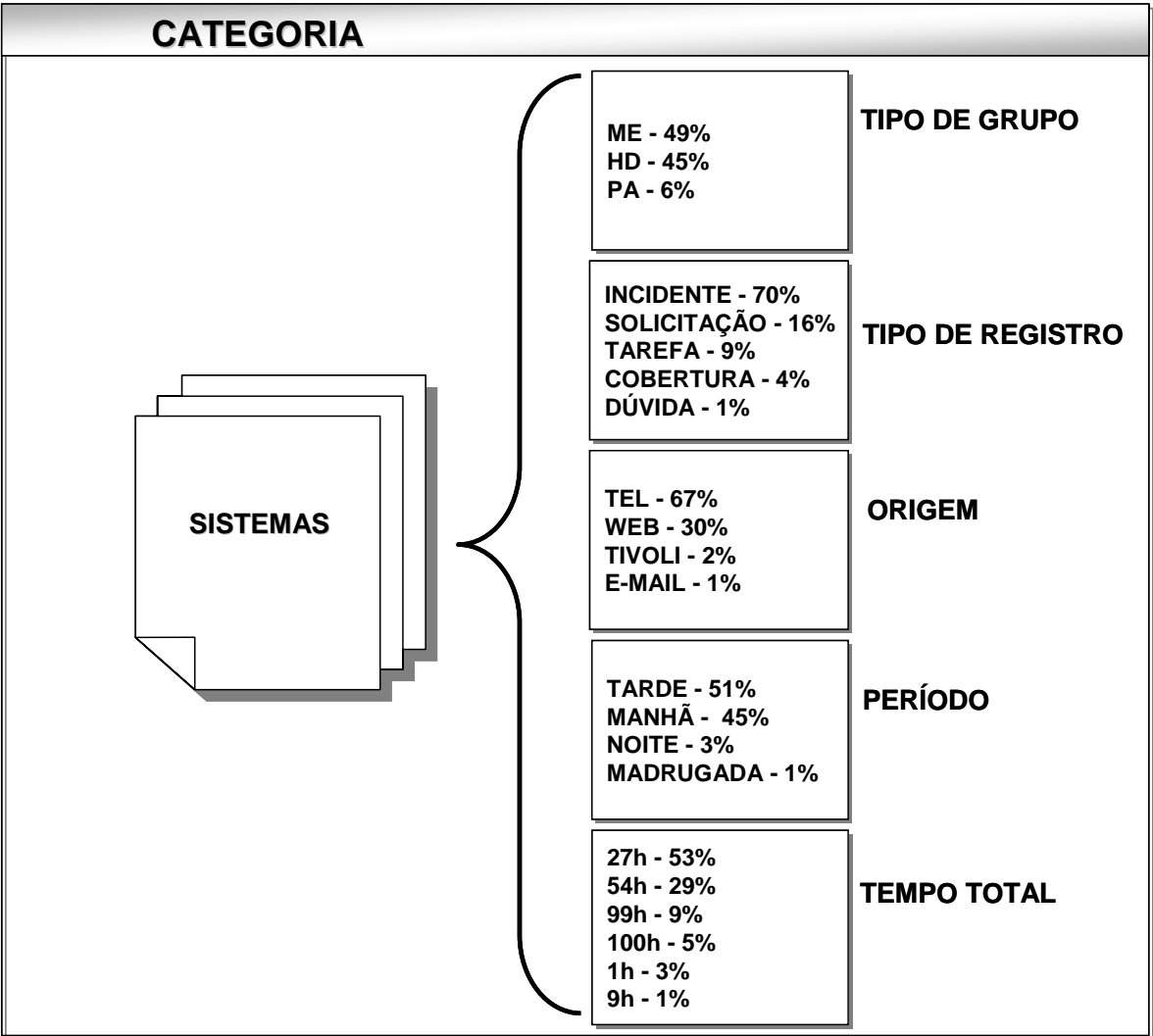












# Livros Grátis

( <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](#)