



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA - PPGECO

THAÍZE FERNANDES OLIVEIRA

**EMPREGO FORMAL NA REGIÃO NORDESTE BRASILEIRA: SEUS EFEITOS NA
INCLUSÃO SOCIAL**

Natal - RN
2009

Livros Grátis

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

Milhares de livros grátis para download.

THAÍZE FERNANDES OLIVEIRA

**EMPREGO FORMAL NA REGIÃO NORDESTE BRASILEIRA: SEUS EFEITOS NA
INCLUSÃO SOCIAL**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Norte como requisito parcial para obtenção do grau de Mestre em Economia.

Orientador: Prof^o Dr. Paulo César Formiga Ramos.

Co-orientadora: Prof^a Dr^a Maria Célia de Carvalho Formiga

Natal - RN
2009

THAÍZE FERNANDES OLIVEIRA

**EMPREGO FORMAL NA REGIÃO NORDESTE BRASILEIRA: SEUS EFEITOS NA
INCLUSÃO SOCIAL**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Norte como requisito parcial para obtenção do grau de Mestre em Economia.

Aprovado em : ____ / ____ / ____

BANCA EXAMINADORA

Prof^o Dr. Paulo César Formiga Ramos – UFRN - Orientador

Prof^a Dr^a. Maria Célia de Carvalho Formiga – UFRN - Co-Orientadora

Prof^o Dr. Neir Antunes Paes - UFPB

Prof^o Dr. Jorge Luiz Mariano da Silva - UFRN

DEDICATÓRIA

Aos MEUS PAIS, José Nilson de Oliveira e Tânia Maria Fernandes Oliveira por todo amor que me transmitem, referência e incentivo na continuidade de minha formação acadêmica e profissional.

AGRADECIMENTOS

A Santíssima Trindade: Pai, Filho e Espírito Santo pelo meu viver. E a Nossa Senhora pela sua intercessão e proteção de mãe. Agradeço a luz inspiradora e a força nos momentos de assumir com responsabilidade, competência e amor a missão a mim confiada.

Aos meus pais, José Nilson de Oliveira e Tânia Maria Fernandes Oliveira, pela vida, exemplo de educação, trabalho e amor.

Aos meus irmãos, Bruno Fernandes Oliveira e Victor Hugo Fernandes Oliveira que com muito amor estão sempre acompanhando meus passos.

Aos meus familiares (materno e paterno) que me apoiaram, direta ou indiretamente, nessa caminhada.

Ao meu namorado Emerson Souza de Assis pelo amor, paciência e companheirismo sempre dedicado a mim.

A meus amigos orientadores Prof^o Dr. Paulo César Formiga Ramos e Prof^a Dr^a Maria Célia de Carvalho Formiga; vocês são exemplos de competência, dedicação e amor à Profissão Docente. Agradeço a Deus por tê-los em meu caminho.

Ao Programa de Pós-Graduação em Economia da Universidade Federal do Rio Grande do Norte representado pelos Coordenadores Prof^o. Dr. João Matos Filho e Prof^a. Dr^a. Maria do Socorro Gondim Teixeira. E, também aos demais professores doutores que tive a honra de ser aluna: Jorge Luiz Mariano da Silva, André Luis Cabral de Lourenço, Maria Lussieu da Silva e Valdênia Apolinário.

Aos professores doutores que apesar de não ter pago disciplina deram grande contribuição nesta trajetória: Márcia Maria de O. Bezerra, Odair Lopes Garcia, Erick Alencar de Figueiredo e demais professores que diretamente contribuíram com seus conhecimentos.

Aos meus amigos do mestrado com quem tenho muita admiração e gratidão, em especial as amigas Elinete e Elisete pela amizade construída, com muito carinho, companheirismo, apoio e paciência acadêmica sempre demonstrada nos momentos de socialização dos conhecimentos.

Aos professores da banca examinadora pela disponibilidade e sugestões para melhoria deste trabalho.

Ao Banco do Nordeste pelo grande incentivo nesta caminhada da pesquisa.

E a todas as pessoas que apesar de não ter citado os nomes, contribuíram direta ou indiretamente no decorrer deste trabalho.

Meus sinceros agradecimentos.

LISTA DE ABREVIATURAS E SIGLAS

AESB - Atlas de Exclusão Social no Brasil

AL - Alagoas

BA – Bahia

CAGED – Cadastro Geral de Empregados e Desempregados

CE - Ceará

CO – Centro-Oeste

IBGE – Instituto Brasileiro de Geografia e Estatística

DE – Distância Euclidiana

DF – Distrito Federal

IDH – Índice de Desenvolvimento Humano

IES - Índice de Exclusão Social

IIS - Índice de Inclusão Social

IND_ALF – Índice de Alfabetização

IND_DES – Índice de Desigualdade

IND_EMP – Índice de Emprego Formal

IND_ESC – Índice de Escolaridade

IND_JUV – Índice de Juventude

IND_POB – Índice de Pobreza

IND_VIO - Índice de Violência

LISA – *Local Indice Spatial Association* (Índice Local de Associação Espacial)

MA - Maranhão

MRLMP - Modelo de Regressão Linear Múltipla Ponderada

NE - Nordeste

NO - Norte

ONU – Organização das Nações Unidas

PB – Paraíba

PE - Pernambuco

PI - Piauí

PIA - População em Idade Ativa

PP_50RTG - Percentual de pessoas com mais de 50% da renda provenientes de transferência governamental

RAIS – Relação Anual de Informações Sociais

RLP – Regressão Linear Ponderada
RN – Rio Grande do Norte
SE – Sergipe
SE' - Sudeste
SU - Sudeste
UF's – Unidades Federativas
VMPE – Variável de Maior Poder Explicativo

LISTA DE GRÁFICOS

Gráficos 3.1 a 3.1.1 – Correlação bivariada entre: IND_EMP e a Variável de Maior Poder Explicativo (VMPE) para a região Nordeste (1757 municípios).....	51
Gráficos 3.2 a 3.10 – Correlação bivariada entre: IND_EMP e VMPE para os nove estados que compõem a região Nordeste.....	52
Gráficos 4.1 a 4.10.1 - Correlação bivariada do IIS com o IND_EMP e com o PP_50RTG para a região Nordeste e para cada uma de suas respectivas Unidades Federativas.....	65

LISTA DE ILUSTRAÇÕES (FIGURAS E MAPAS)

Figura 2.1 – Diagrama de construção do Índice de Exclusão Social.....	23
Figura 3.1 – Comportamento das variáveis IND_EMP, IND_POB, IND_DES, IND_ESC e IND_JUV para as UF's da região Nordeste, 2000.....	31
Figura 3.2 – Comportamento das variáveis, IND_EMP, IND_POB, IND_DES, IND_ESC e IND_JUV para as capitais de cada UF's da região Nordeste, 2000.....	33
Figura 3.3 – <i>Box-plot</i> das Taxas de Emprego Formal dos municípios do Nordeste segundo seus respectivos estados.....	33
Figura 4.1 – Diagrama em (<i>Tree Diagram</i>) segundo as variáveis emprego formal (IND_EMP) e IIS, Unidades Federadas da região Nordeste (n=27), 2000.....	55
Figura 4.2 – Resultado gráfico da junção (<i>Two-Way joining Results</i>) das variáveis IND_EMP e IIS para as UF's brasileiras (n=27), 2000.....	56
Figura 4.3 - Diagrama em <i>Árvore (Tree Diagram)</i> , segundo as variáveis IND_EMP e IIS para as UFs da região Nordeste do Brasil (n=9), 2000.....	57
Figura 4.4 – Resultado gráfico da junção (<i>Two-Way joining Results</i>) das variáveis IND_EMP e IIS para as UF's nordestinas (n= 9), 2000.....	58
Figura 5.1 - Índice de Moran multivariado - versus IND_EMP - para o próprio IND_EMP (W) e os demais índices IND_DES (A), IND_ESC (B), IND_POB (C), IND_JUV (D) e W_IIS com IND_EMP (E), Região Nordeste, 2000.....	69

Mapa 4.1: <i>Clusters</i> dos Estados do Nordeste agrupados segundo as condições de emprego e do Novo índice de Inclusão Social, 2000.....	59
Mapa 5.1 - BoxMap da taxa de emprego formal sobre a PIA (IND_EMP) para o ano 2000, Região Nordeste.....	70
Mapa 5.2 - Mapa do desvio padrão para o IND_EMP, Região Nordeste, 2000.....	71
Mapa 5.3 - Cluster Map , representativo da distribuição espacial do IND_EMP, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.....	72
Mapa 5.4 - Cluster Map , representativo da distribuição espacial do IND_EMP e IND_POB, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000...	73
Mapa 5.5 - Cluster Map , representativo da distribuição espacial do IND_EMP e IND_DES, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000...	74
Mapa 5.6 - Cluster Map , representativo da distribuição espacial do IND_EMP e IND_ESC, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000...	75
Mapa 5.7 - Cluster Map , representativo da distribuição espacial do IND_EMP e IND_JUV, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000...	76
Mapa 5.8 - Cluster Map , representativo da distribuição espacial do IIS e IND_EMP, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.....	77

LISTA DE QUADROS

Quadro 2.1 – Grupos temáticos analisados com seus componentes Índices de Exclusão/Inclusão Social e respectivos pesos atribuídos.....	23
Quadro 2.2 – Classificação do IES.....	24
Quadro 3.1 – Estatísticas Descritivas para a variável IND_EMP, segundo UFs da região Nordeste, 2000.....	34
Quadro 3.2 – Matriz de correlações entre as variáveis do estudo, região Nordeste (1757 municípios), 2000.....	35
Quadro 3.3 – Correlações entre o IND_EMP e as demais variáveis sob estudo das UFs da região Nordeste, 2000.....	36
Quadro 4.1 – Composição do novo Índice de Inclusão Social (IIS).....	54
Quadro 4.2 - Distribuição dos grupos formados pela Análise de cluster, segundo as variáveis IND_EMP e IIS para as UFs brasileiras (n=27), 2000.....	55
Quadro 4.3 - Distribuição dos grupos formados pela Análise de cluster, segundo as variáveis IND_EMP e IIS para as UF's da região Nordeste do Brasil (n=9), 2000.....	57
Quadro 4.4 – Correlações referente as variáveis: IND_EMP, IIS e PP_50RTG para a Região Nordeste (1757 municípios).....	60
Quadro 4.5 – Correlações para as nove Unidades Federativas da região Nordeste referente as variáveis IND_EMP, IIS e PP_50RTG.....	61

LISTA DE TABELAS

- Tabela 3.1** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Região Nordeste, 2000 (n = 1757 municípios)**..... 38
- Tabela 3.2** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado do Maranhão, 2000 (217 municípios)**..... 39
- Tabela 3.3** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado do Piauí, 2000 (n = 220 municípios)**..... 39
- Tabela 3.4** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado do Ceará (181 municípios)**..... 40
- Tabela 3.5** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado do Rio Grande do Norte (n = 164 municípios)**..... 40
- Tabela 3.6** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado da Paraíba (n = 221 municípios)**..... 41
- Tabela 3.7** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado de Pernambuco (n = 178 municípios)**..... 41
- Tabela 3.8** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado de Alagoas (n = 101 municípios)**..... 42
- Tabela 3.9** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado do Sergipe (n = 72 municípios)**..... 42
- Tabela 3.10** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), **Estado da Bahia (n = 415 municípios)**..... 43
- Tabela 3.1.1** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), **Região Nordeste, 2000 (n = 1757)**..... 44
- Tabela 3.2.1** - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), **Estado do Maranhão, 2000 (n = 217 municípios)**..... 45

Tabela 3.3.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Piauí, 2000 (n = 220 municípios)	45
Tabela 3.4.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Ceará, 2000 (n = 181 municípios)	46
Tabela 3.5.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Rio Grande do Norte, 2000 (n = 164 municípios) ...	46
Tabela 3.6.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado da Paraíba, 2000 (n = 221 municípios)	47
Tabela 3.7.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado de Pernambuco, 2000 (n = 178 municípios)	47
Tabela 3.8.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado de Alagoas, 2000 (n = 101 municípios)	48
Tabela 3.9.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Sergipe, 2000 (n = 72 municípios)	48
Tabela 3.10.1 - Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado da Bahia, 2000 (415 municípios)	49
Tabela 4.1 – Sumário de regressão linear múltipla ponderada para a variável resposta IIS, ajustado para as variáveis independentes IND_EMP e PP_50RTG, Área Geográfica - Região Nordeste (NE) e Unidades Federativas do NE, 2000	63

LISTA DE EQUAÇÕES

Equação 2.1 – Cálculo do IES.....	22
Equação 2.2 – Cálculo do coeficiente de correlação.....	95
Equação 2.3 – Cálculo do teste t de Student.....	95
Equação 2.4 – Modelo Geral de Regressão Linear Múltipla.....	96
Equação 2.4.1 – Equação de estimação da Regressão Linear Múltipla.....	96
Equação 2.5 – Cálculo da Distância Euclidiana entre dois pontos X e Y.....	100
Equação 2.6 – Cálculo do índice de Moran.....	103

RESUMO

O trabalho visa estudar a problemática do emprego formal na região Nordeste brasileira e seus efeitos na inclusão social, tomando por base a análise de variáveis retiradas do Atlas de Exclusão Social, que tem por base o censo 2000, empregando o município como unidade de análise. Como opções metodológicas, apresentou-se, inicialmente, uma análise exploratória dos dados, seguida da utilização de técnicas estatísticas multivariadas, tais como, análise de regressão linear múltipla ponderada, análise de *cluster* e análise exploratória espacial. Os resultados apontam para baixas taxas de emprego formal na População em Idade Ativa, bem como baixos índices de inclusão social na região Nordeste brasileira. Constatou-se uma forte associação do emprego formal com os indicadores de inclusão social investigados (escolaridade, desigualdade, pobreza, juventude e renda de transferência governamental). Nas Unidades Federativas, onde há melhores níveis de emprego formal, também foram observados bons índices de inclusão social, havendo destaque para o estado do Rio Grande do Norte com melhores condições de vida e para o estado do Maranhão e Piauí com piores condições. A situação da região Nordeste é bastante precária diante dos indicadores sob estudo, remetendo para a necessidade de se dar ênfase aos programas e ações governamentais direcionadas, em especial, ao aumento no nível de emprego formal da região, refletindo, assim, em melhorias da desigualdade de renda, melhorando a inclusão social da população Nordestina.

PALAVRAS-CHAVE: Emprego formal; Região Nordeste; Inclusão Social.

ABSTRACT

This work aims to study the problem of the formal job in the Brazilian Northeast region and its effect in the social inclusion, taking for base the analysis of variables defined in the Atlas of Social Exclusion, which is based on the 2000 Brazilian Census, choosing the county as unit of analysis. As methodological options, an exploratory data analysis was performed, followed by multivariate statistical techniques, such as weighted multiple regression analysis, cluster analysis and exploratory analysis of spatial data. The results pointed out to low rates of formal job for the active age population as well as low indexes of social inclusion in the Northeast region of Brazil. A strong association of the formal job with the indicators of social inclusion under investigation, was evidenced (schooling, inequality, poverty, youth and income from government transfers), as well as a strong association of the formal job with the new index of social inclusion (IIS), modified from the IES. At the Federative Units, in which better levels of formal job had been found, good indexes of social inclusion are also observed. Highlights for the state of the Rio Grande do Norte, with the best conditions of life, and for the states of the Maranhão and Piauí, with the worst conditions. The situation of the Northeast region, facing the indicators under study, is very precarious, claiming for the necessity of emphasizing programs and governmental actions, specially directed to the raise of formal job levels of the region, reflecting, thus, in improvements on the income inequality, as well as in the social inclusion of the population of Northeastern natives.

KEYWORDS: Formal employment, Northeast Region, Social Inclusion.

SUMÁRIO

Capítulo 1	INTRODUÇÃO.....	13
1.1	O EMPREGO FORMAL NO NORDESTE BRASILEIRO.....	14
1.2	O EMPREGO FORMAL NO CONTEXTO DA INCLUSÃO/EXCLUSÃO SOCIAL.....	16
1.3	JUSTIFICATIVA E RELEVÂNCIA.....	18
1.4	OBJETIVOS.....	19
1.5	ORGANIZAÇÃO DA TESE.....	19
Capítulo 2	METODOLOGIA.....	21
2.1.	TÉCNICAS DE ANÁLISES DOS DADOS.....	25
2.1.1.	Análise Descritiva e Exploratória.....	25
2.1.2.	Análise do Coeficiente de Correlação.....	26
2.1.3.	Análise de Regressão Linear Múltipla Ponderada.....	26
2.1.4.	Análise de <i>Cluster</i> (ou de agrupamento).....	27
2.1.5.	Análise Espacial.....	28
Capítulo 3	INDICADORES SOCIOECONÔMICOS QUE SÃO INFLUENCIADOS E INFLUENCIAM O EMPREGO FORMAL NAS UNIDADES FEDERATIVAS NORDESTINAS.....	30
3.1	EXPLORANDO OS INDICADORES.....	30
3.2	RELAÇÃO ENTRE A “TAXA DE EMPREGO FORMAL SOBRE A PIA (IND_EMP)” E OUTRAS VARIÁVEIS SOCIOECONÔMICAS.....	37
3.3	RELAÇÃO ENTRE A “POBREZA DOS CHEFES DE FAMÍLIA DO MUNICÍPIO (IND_POB)” E OUTRAS VARIÁVEIS SOCIOECONÔMICAS.....	44
3.4	DISTRIBUIÇÃO DO ÍNDICE “TAXA DE EMPREGO FORMAL SOBRE A PIA” COM A VARIÁVEL DE MAIOR PODER EXPLICATIVO.....	50
Capítulo 4	ASSOCIAÇÃO DO EMPREGO FORMAL COM A INCLUSÃO SOCIAL NO CONTEXTO DAS UNIDADES FEDERATIVAS BRASILEIRAS COM ÊNFASE NO NORDESTE.....	54
4.1	AS UNIDADES FEDERATIVAS (UF’S) NO CONTEXTO DAS GRANDES REGIÕES.....	54
4.2	O CASO DAS UF’S DO NORDESTE BRASILEIRO.....	56
4.3	ASSOCIAÇÕES ENTRE INDICADORES.....	59
4.4	INVESTIGANDO RELAÇÃO ENTRE VARIÁVEIS ATRAVÉS DA APLICAÇÃO DO MODELO DE REGRESSÃO LINEAR MÚLTIPLA PONDERADA.....	62
4.5	DISTRIBUIÇÃO DO INDICADOR “NOVO ÍNDICE DE INCLUSÃO SOCIAL” COM AS DEMAIS VARIÁVEIS DE INTERESSE.....	64
Capítulo 5	ANÁLISE EXPLORATÓRIA ESPACIAL DO EMPREGO FORMAL E INCLUSÃO SOCIAL/EXCLUSÃO SOCIAL NA REGIÃO NORDESTE DO BRASIL.....	68
Capítulo 6	CONCLUSÕES E RECOMENDAÇÕES.....	78
Referências	83
Anexos	86

CAPÍTULO 1: INTRODUÇÃO

A maior pressão concorrencial que a economia brasileira vem enfrentando, em decorrência da liberalização comercial vigente desde o início da década de noventa, está forçando as empresas a buscar maior competitividade, sendo essa uma característica recente da inserção do país na economia mundial. O mercado de trabalho está exigindo, a cada dia, qualificação crescente. Segundo Oliveira (2007), o desemprego é um dos mais complexos problemas das sociedades contemporâneas, já que o trabalho representa a independência, as rendas das pessoas e o status social. Sendo assim, o desemprego produz o agravamento de problemas sociais, tais como: exclusão social, perda da auto-estima, desigualdade social, aumento dos índices de violência, e outros mais.

Segundo o "Atlas da Exclusão Social no Brasil" (AESB) (POCHMANN; AMORIM, 2003), dos 5.507 municípios brasileiros em 2000, apenas 200 (duzentos) apresentavam padrão de vida adequado, entre os quais boa parte estava apenas em condições intermediárias e a maioria ainda está em condição de exclusão. Assim, o levantamento apontou que mais de 25% dos brasileiros viviam em condições precárias, com baixa renda, sem emprego formal e baixo acesso à educação.

A discrepância fica mais evidenciada quando se observa as regiões Norte e Nordeste, em relação ao restante do país. Entre os 100 municípios com melhor índice de inclusão social, apenas um estava localizado na região Nordeste - Fernando de Noronha. No extremo oposto, a situação se inverte: apenas duas entre as 100 cidades com o pior desempenho não estavam nas regiões localizadas acima do Trópico de Capricórnio - São João das Missões e Verdelândia, ambas em Minas Gerais.

Sabe-se que a região Nordeste tem uma participação muito importante em nível nacional, representando 28% da população e 33% de seus municípios, dos quais a maioria (72,1%) apresenta os maiores índices de exclusão de todo o país.

O Índice de Exclusão Social (IES) é o mais recente dos indicadores que se propõem a medir o grau de desenvolvimento humano da população. É mais complexo que o Índice de Desenvolvimento Humano (IDH), modelo elaborado pela Organização das Nações Unidas (ONU) e que leva em conta educação, renda e a longevidade populacional. Por sua vez o IES envolve sete indicadores (sub-índices) que traduzem as condições de emprego formal, pobreza, desigualdade, alfabetização, escolaridade, juventude, violência.

Em relação ao desenvolvimento humano, dentre os 174 países analisados em 2000, o Brasil encontra-se na 74^a posição, com um IDH equivalente a 0,766, classificado como Médio Desenvolvimento Humano. Analisando-se as regiões brasileiras, conclui-se que a região Nordeste foi a que mais cresceu entre 1991 e 2000. Entre os estados da região Nordeste, Maranhão, Alagoas e Piauí foram os que apresentam maior número de municípios com baixo IDH na região.

No que se refere aos indicadores componentes do IES, o de maior interesse deste estudo é o emprego formal, traduzido pelo índice de emprego (IND_EMP), o qual, em termos da região Nordeste, a exemplo dos demais indicadores socioeconômicos, está entre os mais baixos do país. Observa-se que a participação dos assalariados em ocupações formais é bastante reduzida nas cidades brasileiras (36,1%), a maioria situada nas regiões Norte e Nordeste, onde a estrutura do mercado de trabalho assalariado é pouco desenvolvida. Por sua vez, 10,3% dos municípios do país contam com uma estrutura ocupacional centrada no emprego formal, a maioria geograficamente situada nas demais regiões do país.

Neste contexto, o esforço investigativo desse trabalho teve como principal objetivo realizar uma análise do emprego formal da região Nordeste, a nível de estados e municípios e seu efeito na inclusão social, com base nos dados do censo 2000. Para tanto, foram utilizadas técnicas estatísticas que possibilitaram melhor compreensão do emprego formal na região Nordeste, de modo a propiciar uma agenda de prioridades que possam subsidiar as políticas sociais voltadas para essa população.

1.1. EMPREGO FORMAL NO NORDESTE BRASILEIRO

A abertura econômica no final do século XX acarretou, no Brasil, uma perda de emprego para muitos trabalhadores e exigente qualificação desses profissionais em sua área de atuação. A competitividade do mundo globalizado está exigindo incessante busca do conhecimento por parte dos trabalhadores em geral, para que, dessa forma, possam se manter no trabalho ou mesmo para os desempregados uma maneira de se reintegrarem nesse mercado em mutação.

A taxa de empregos não tem acompanhado a demanda. Essa situação faz com que muitos trabalhadores se sujeitem a trabalhar sem carteira assinada, para continuar no mercado de trabalho.

Segundo Nagameni (2007), assessor do ministério do Trabalho e Emprego, a geração de empregos é fundamental para o desenvolvimento de qualquer região, principalmente quando são empregos formais com “proteção social”. A criação de empregos formais é importante na medida em que traz um processo de inclusão social dos trabalhadores, e isso garante o crescimento da região de maneira sustentável.

Diante das inúmeras vantagens que o emprego formal tem e devido ao seu baixo crescimento, surge o interesse de analisar o seu perfil para a Região Nordeste que é, dentre as regiões brasileiras, a que apresenta índices mais baixos para todos os indicadores aqui estudados e, em especial, o que trata do emprego formal.

Desta Forma:

A percepção é de que se o País lograsse enveredar por um padrão de desenvolvimento em que a geração de empregos formais passasse a ter importância relativa cada vez maior, reduzindo-se a um grau residual a ocorrência de ocupações precárias, haveria modificações significativas no sentido de redução da desigualdade de rendimentos do trabalho e de diminuição de disparidades regionais – já que em tal cenário o Nordeste tenderá a ser consideravelmente beneficiado (ARAÚJO; LIMA; 2007, p.275).

Segundo Chahad (2007), utilizando dados do Instituto Brasileiro de Geografia e Estatística (IBGE) entre 1997 e 2006, o emprego formal cresceu mais no Nordeste (22,7%) que no Brasil como um todo (17,8%). Entre os estados da Região Nordeste, observa-se que alguns, como o Piauí (13,7%) e a Paraíba (17,2%), cresceram abaixo da média da região, enquanto outros, como o Ceará (35,5%) e o Rio Grande do Norte (42,3%), apresentaram um crescimento bem superior à média.

Comparando os resultados acima citados com os obtidos no trabalho, tem-se o estado do Rio Grande do Norte, no ano 2000, com melhor índice de emprego formal, mantendo, portanto, inalterada a melhor posição.

Outra característica marcante, como mostra Pires (2006) - com base nos dados do Cadastro Geral de Empregados e Desempregados (CAGED) e Relação Anual de Informações Sociais (RAIS) - é o fato dos trabalhadores nordestinos terem um rendimento inicial no setor formal de, no máximo, 1,5 salários mínimos. De 2000 a 2004, apenas o Rio Grande do Norte diminuiu o percentual (de 66,99% para 64,33%) de trabalhadores que ganham esse valor. Nos demais estados da região o percentual nesse período aumentou. O autor conclui:

É provável que a situação diferenciada do estado do Rio Grande do Norte deva-se à indústria de petróleo existente que possibilita tanto uma melhor situação salarial para os trabalhadores lotados na indústria de transformação, como também para os demais segmentos da economia, na medida em que esta atividade reproduz um efeito em cadeia muito forte no tocante à geração de empregos diretos e indiretos (PIRES, 2006, p.376).

O desenvolvimento de políticas públicas integradas permitirá a expansão do mercado de trabalho e, conseqüentemente, da economia, além do crescimento pessoal dos profissionais. Estas políticas são emergenciais e precisam ser colocadas em prática rapidamente para evitar que profissionais continuem à margem do mercado de trabalho, sem condições de inserção.

1.2. O EMPREGO FORMAL NO CONTEXTO DA INCLUSÃO/EXCLUSÃO SOCIAL

O fato de o Brasil ter mudado muito desde o início da década de noventa, ainda não acarretou em mudanças na desigualdade social, permanecendo ainda um elevado índice de excluídos no país. Diante de um quadro de diferenciações sociais e regionais, há uma relevância destacada deste estudo diante das atuais possibilidades de enfrentamento do emprego formal no contexto da inclusão/ exclusão social no território brasileiro tomando como base os indicadores do Atlas de Exclusão Social no Brasil.

Segundo Pochmann (2003), a exclusão social é entendida como a impossibilidade de parcela da população ter acesso a alguns direitos sociais básicos que, em decorrência, leva-os à condição de sub-cidadãos: sem direitos, sem consumo e sem ferramentas para superar essa condição. É o que acontece com grande parte da população que se encontra na região Nordeste brasileira.

Um diagnóstico preocupante é que no Brasil existem algumas ilhas de inclusão social, em meio a um mar de exclusão social. Este problema vem acompanhado de uma desigualdade regional marcante no nosso país. Dos 5.507 municípios brasileiros em 2000, tem-se 2.290 municípios com Índice de Exclusão Social até 0,4 e, portanto, em situação de elevada exclusão social. Verifica-se que a região Nordeste é a recordista, com 1.652 municípios nessa situação, o que representa 72,1% desses municípios no território nacional.

Contrário a esses resultados, tem-se os municípios do Sul e Sudeste do país onde se concentram os maiores focos de inclusão social e, proporcionalmente, a menor participação no número de municípios com Índice de Exclusão Social até 0,4. Já na Região Nordeste,

existem poucos municípios com índices que indicam média/baixa inclusão – ou seja, com valores entre 0,5 e 0,6 - que é o mais elevado valor que os municípios dessa região atingem, estando estes municípios localizados da seguinte forma: Alagoas, (Maceió), Maranhão (São Luiz), Ceará (Fortaleza), Piauí (Teresina), e Sergipe (Aracaju) com apenas um município que são as respectivas capitais desses estados; Rio Grande do Norte (Natal e Parnamirim) com apenas dois municípios; Paraíba (João Pessoa, Cabedelo e Campina Grande) com três municípios; Pernambuco (Fernando de Noronha, Recife, Paulista e Olinda) e Bahia (Salvador, Lauro Freitas, Madre de Deus e Dias d'Ávila) com quatro municípios. Assim, o Nordeste é caracterizado por poucos pontos de inclusão, que totalizam apenas 18 municípios, em meio a um mar de exclusão, que são os 1.769 municípios restantes, enquadrados em média exclusão ou elevada exclusão (com valores de 0,0 a 0,5).

Um ponto fundamental é que, em cada cidade do Brasil, está surgindo uma nova forma de exclusão social. Esta pode ser caracterizada por novas variáveis que passaram a ser importantes indutores de privação de direitos sociais como, por exemplo, o desemprego e a violência. Os grupos que sofrem a nova exclusão, diferentemente dos que sofrem a exclusão tradicional, tem boa escolaridade, participam de famílias pequenas, não são migrantes e, para eles, tem menos importância questões como raça e gênero. O pior é que essa nova exclusão não vem, até o momento, substituir velhas formas de exclusão social, mas sim se somar a elas. Isto é, longe de vencer o processo de exclusão social tradicional, criou-se uma nova forma, mais difícil de ser vencida. Isso aparece em todas as grandes cidades analisadas, pois, afinal, foram elas que sofreram mais fortemente com os (des) caminhos seguidos pela economia brasileira nas últimas décadas, acolhendo fechamento de empresas, desemprego e aperto fiscal. É natural, portanto, que seja exatamente nelas que a nova exclusão se faça mais visível. (POCHMANN, 2003).

Trata-se realmente de retratos melancólicos sobre a realidade brasileira. Os problemas que deles emergem, apontam para a fraqueza das medidas implementadas para a superação do quadro. Ao mesmo tempo, mais do que nunca, fica claro, sobre os territórios brasileiros, que a Região Nordeste precisa sofrer intervenções que dinamizem sua economia e suas relações sociais. Por outro modo, tem-se com o AESB, desde então, mais um instrumento capaz de ajudar no diagnóstico dos problemas e facilitar o planejamento das soluções.

Na região, há uniformidade, com a maior parte dos municípios das Unidades Federativas (UF's) em situação de exclusão e alguns focos que ilustram situação social intermediária. Esses focos de melhor situação localizam-se próximos à costa, junto às capitais e em pontos isolados do interior. Destacam-se Maranhão (97,7%) e Piauí (97,3%), pela maior

proporção de municípios com grande exclusão (valores entre 0,0 e 0,4). Embora as demais unidades da federação localizadas na região Nordeste apresentem uma proporção igual ou superior a 90% dos municípios com elevada exclusão, exceto o estado do Rio Grande do Norte (84,3%), onde esta proporção é um pouco menor.

Observados os dados apresentados neste estudo, é inevitável a sensação de urgência no enfrentamento de questões tão antigas e contemporâneas como a desigualdade regional do Brasil. Os resultados são inquestionáveis e mostram o quanto precisa ser feito.

O presente estudo tem o propósito de dar uma contribuição para que órgãos governamentais possam fazer planejar e incentivar a inserção no mercado de trabalho formal, tirando a população da informalidade.

1.3. JUSTIFICATIVA E RELEVÂNCIA

Apesar das grandes mudanças ocorridas no Brasil ao longo do século XX, as desigualdades sociais mantiveram-se quase inalteradas, com o problema de exclusão social sempre presente no decorrer das décadas e fortemente associada à questão do emprego e geração de renda. Em vista disso, a identificação do complexo quadro de diferenciações sociais e regionais assume relevância destacada diante das atuais possibilidades de enfrentamento da exclusão social no país (POCHMANN; AMORIM, 2003).

De acordo com o "Atlas da Exclusão Social no Brasil" (POCHMANN; AMORIM, 2003), no Brasil, o 5º (quinto) maior país do mundo, há alguns “oasis” de inclusão social em meio a uma “deserto” de exclusão, que se estende por praticamente todo o território brasileiro, mais fortemente revelados na região Norte e Nordeste. É importante frisar que essa inclusão/exclusão social referida pelos autores, contempla dentre seus indicadores, a variável emprego.

Tendo em vista que a região Nordeste está enquadrada em pior situação de exclusão social, surge o interesse de aprofundar os estudos dessa exclusão e de seus diversos condicionantes, especialmente o relativo ao emprego formal, objeto central de análise neste trabalho, através de métodos e técnicas estatísticas, com o propósito final de traçar um panorama da população formalmente empregada no Nordeste, e em cada um dos seus estados e seus efeitos na inclusão/exclusão social da região.

1.4. OBJETIVOS

Geral:

Analisar a situação do emprego formal da região Nordeste brasileira, a nível de estados e municípios e seu efeito na inclusão social, com base nos dados do censo 2000.

Específicos:

a) Realizar uma análise comparativa exploratória do emprego formal entre os estados da região Nordeste;

b) Identificar o poder explicativo do emprego formal nos indicadores de pobreza e no novo índice de inclusão social da região Nordeste, bem como a influência dos indicadores de inclusão/ exclusão social no indicador emprego formal, através de um modelo de regressão linear múltipla ponderada;

c) Agrupar as Unidades Federativas (UF's) investigadas segundo características consideradas: o Índice de Emprego Formal e o novo Índice de Inclusão Social, através do método da análise de *cluster* (ou de agrupamento);

d) Realizar uma análise exploratória espacial do emprego formal na região Nordeste, tomando o município como unidade de análise, bem como suas associações espaciais com cada um dos índices de inclusão/exclusão social e com o novo Índice de Inclusão Social, utilizando métodos de Estatística Espacial;

Com a consecução dos objetivos propostos pretende-se avançar no conhecimento das desigualdades sociais da Região Nordeste e de seus estados, enfocadas no emprego formal, traduzido pelo índice de emprego (IND_EMP), componente do índice de exclusão social (IES) criado por Pochmann e Amorim (2003), bem como na associação do IND_EMP com outros indicadores de inclusão/exclusão social.

1.5. ORGANIZAÇÃO DA DISSERTAÇÃO

O presente trabalho de pesquisa está estruturado em seis capítulos. O primeiro faz a abordagem do tema estudado, na qual é possível visualizar um cenário de abordagens teóricas acerca do emprego formal na região Nordeste brasileira. Em seguida, se apresenta o emprego formal no contexto da inclusão/exclusão social, a relevância do estudo e os objetivos que motivaram o trabalho abordado.

O segundo capítulo apresenta a metodologia utilizada, tratando da pesquisa quantitativa no tratamento de dados secundários: análise descritiva e exploratória dos dados, seguida de uma abordagem das técnicas multivariadas utilizadas para a inferência das informações que refletem a população estudada.

O terceiro, o quarto e o quinto capítulos dedicam-se a uma investigação, no sentido de se colher os resultados do estudo. Inicialmente. No capítulo 3, se faz uma abordagem com as observações obtidas e analisadas de forma descritiva e exploratória, na qual é possível traçar o perfil da região Nordeste brasileira em termos de estados e municípios (como unidade de análise) partindo mais adiante para os resultados das análises multivariadas utilizando as técnicas análise de regressão linear múltipla ponderada e análise bivariada, visando identificar as relações de dependência entre as variáveis que compõem o índice de inclusão/exclusão social. No capítulo quatro, é analisada a associação do emprego formal com o novo índice de inclusão social, onde se utilizou o método de análise Cluster e demais técnicas multivariadas utilizadas no capítulo anterior. No capítulo cinco, é abordada uma análise exploratória espacial para a região Nordeste, tomando sempre os municípios como unidade de análise, de forma a revelar a distribuição espacial dos dados através dos mapas temáticos, com o objetivo de identificar as posições relativas dos municípios da Região Nordeste quanto ao nível de inclusão social.

Finalmente, o sexto capítulo fecha o trabalho, tratando das Conclusões e Recomendações, onde são apresentadas as conclusões finais do trabalho e as Direções de Pesquisa para futuros trabalhos, encerrando com as Recomendações dirigidas às instituições e atores sociais que estão envolvidos na temática enfocada.

CAPÍTULO 2: METODOLOGIA

Os dados básicos a serem utilizados nesse estudo são de fonte primária do IBGE, sendo, portanto, dados secundários. A fonte intermediária é o Atlas da Exclusão Social no Brasil (POCHMANN; AMORIM, 2003).

Dada a sua complexidade, o problema, da exclusão/inclusão social (considerado um processo de natureza trans-disciplinar) que permeia todo este trabalho, muito embora o foco principal seja o emprego formal, requer o envolvimento de diferentes componentes analíticos, pelo que, estabeleceram-se várias opções metodológicas, a seguir explanadas.

Para o tratamento e organização da base de dados, foram considerados os indicadores socioeconômicos municipais da região Nordeste (1787 municípios), disponíveis no AESB, (POCHMANN; AMORIM, 2003), referentes ao censo 2000. Para tratamento e organização desse banco de dados, foram empregados os *softwares* Excel e Statistica 7.0¹.

Informações metodológicas sobre o Atlas da Exclusão Social no Brasil de Pochmann e Amorim (2003).

Para construir o Índice de Exclusão Social (IES) no Brasil, os autores citados utilizaram os dados de suas unidades territoriais, totalizando 5.507 municípios, tomando como referência o Censo 2000 do IBGE. O IES é o mais recente dos indicadores que se propõem a medir o grau de desenvolvimento humano da população, sendo mais complexo que o IDH, modelo elaborado pela ONU.

As principais vantagens deste índice é que o mesmo incorpora parâmetros que, até o momento, não foram utilizados por outros medidores, como a proporção de jovens no município, a taxa de homicídios por cada 100 mil habitantes e a proporção de trabalhadores com carteira de trabalho assinada.

Os autores selecionaram sete variáveis² consideradas fundamentais para o estudo da exclusão social, compondo três grupos de indicadores, a seguir apresentados com as denominações dos índices e as respectivas siglas a serem utilizadas no presente estudo. O primeiro grupo buscou avaliar se as pessoas do município possuíam um padrão digno de vida.

¹ Software adquirido pelo departamento de Estatística da UFRN.

² Essas variáveis foram transformadas em índices pelos autores. Na realidade, seria mais apropriado denominá-los sub-índices, uma vez que a partir deles é que são gerados os índices sintéticos, de exclusão propriamente ditos.

Três variáveis foram contabilizadas: a pobreza dos chefes de família no município (Índice de Pobreza – IND_POB), a participação do emprego formal na população em idade ativa (Índice de Emprego Formal – IND_EMP) e a desigualdade de renda percebida pelos chefes de família (Índice de Desigualdade – IND_DES). O segundo grupo preocupou-se em avaliar o padrão de educação formal dos habitantes do município e estudou: a taxa de alfabetização das pessoas acima de cinco anos de idade (Índice de Alfabetização – IND_ALF) e o número de anos de estudo do chefe de família (Índice de Escolaridade – IND_ESC). Por fim, o terceiro grupo de variáveis considerou o risco a que está exposta a juventude nas cidades quando o assunto é violência. Para tanto, duas variáveis foram estudadas: a porcentagem de jovens na população (Índice de Juventude – IND_JUV) e a taxa de homicídios por 100 mil habitantes (Índice de Violência – IND_VIO). Os grupos temáticos analisados pelo IES com as respectivas ponderações dos índices, bem como o diagrama ilustrativo de suas construções encontram-se no Quadro 2.1 e Figura 2.1, na página seguinte.

A expressão para cálculo do IES é semelhante a do IDH, utilizando-se a mesma equação que os transforma em valores que variam no intervalo de 0 a 1, dada pela:

Equação 2.1 **Cálculo do IES**

$$X_{i,p} = \frac{X_i - \text{MIN}(X_i)}{\text{MAX}(X_i) - \text{MIN}(X_i)} \quad (2.1)$$

Onde:

$X_{i,p}$: valor calculado do índice em estudo;

p : identifica qual variável de bem-estar social está em estudo;

i : índice que identifica o município em estudo;

X_i : valor assumido pela variável p ;

$\text{MIN}(X_i)$: valor mínimo da variável X ;

$\text{MAX}(X_i)$: valor máximo da variável X ;

A partir dos sete índices que refletem a exclusão social, ponderando-os adequadamente, obtém-se o Índice de Exclusão Social, uma espécie de “termômetro” do bem-estar material dos habitantes dos municípios brasileiros. Os resultados, que devido à Equação (2.1) podem variar entre 0 e 1, foram agrupados em classes, de modo a separar as melhores condições de inclusão (os mais próximos de um) dos casos mais graves de exclusão social (os

mais próximos de zero). Assim o índice 0 (zero) reflete a exclusão absoluta enquanto que o índice 1 (um) reflete a inclusão total. Uma adaptação da classificação básica do Atlas foi adaptada neste estudo, de modo a reduzir a concentração de municípios nas faixas de média e elevada exclusão, conforme mostrado no Quadro 2.2.

Quadro 2.1
Grupos temáticos analisados com seus componentes Índices de Exclusão/Inclusão Social e respectivos pesos atribuídos.

Grupos Temáticos	Índices criados	Pesos
Padrão de vida digno	Pobreza: medido pela pobreza dos chefes de família do município.	17%
	Emprego formal: medido pela taxa de emprego formal sobre a PIA.	17%
	Desigualdade: medido por uma <i>proxi</i> da desigualdade de renda.	17%
Conhecimento	Alfabetização: medido pela taxa de alfabetização de pessoas com mais de 5 anos.	5,7%
	Escolaridade: medido pelo número médio de anos de estudo do chefe de domicílio.	11,3%
Risco juvenil	Juventude: medido pela porcentagem de jovens na população.	17%
	Violência: medido pelo número de homicídios por 100 mil habitantes.	15%

Fonte: Pochmann e Amorim (2003).

COMPONENTE →	Padrão de vida digno			Conhecimento		Risco Juvenil	
INDICADOR →	Pobreza	Emprego formal	Desigualdade	Anos de estudos	Alfabetização	Concentração de jovens	Violência
	↓	↓	↓	↓	↓	↓	↓
ÍNDICE DO INDICADOR →	Índice de Pobreza	Índice de Emprego	Índice de Desigualdade	Índice de Escolaridade	Índice de Alfabetização	Índice de juventude	Índice de Violência
	↘	↓	↙	↘	↙	↘	↙
ÍNDICE DO COMPONENTE →	Índice de Padrão de vida digno			Índice de Conhecimento		Índice de Risco Juvenil	
	↘			↓		↙	
	ÍNDICE DE EXCLUSÃO SOCIAL						

Figura 2.1: Diagrama de construção do Índice de Exclusão Social.

Fonte: Pochmann e Amorim (2003), elaborado a partir do Human Development Report 2002.

Quadro 2.2
Classificação do IES

Variação do IES	Classificação da Exclusão/Inclusão Social
0,0 – 0,4	Elevada exclusão
0,4 – 0,5	Média exclusão
0,5 – 0,6	Média/baixa inclusão
0,6 – 0,8	Média/alta inclusão
0,8 – 1,0	Elevada inclusão

Fonte: Elaboração própria / Adaptado de Pochmann e Amorim (2003).

A análise exploratória preliminar dos índices de interesse, bem como dos indicadores ou sub-índices, foi realizada pela utilização das ferramentas usuais para esse tipo de análise: tabelas, gráficos, *Box-Plots*, coeficiente de correlação, etc. Em seguida foram utilizadas algumas técnicas estatísticas multivariadas, tais como: análise de regressão linear múltipla ponderada, análise de *Cluster* e análise espacial, que serão descritas posteriormente.

Na análise explanatória espacial dos dados, foram construídos mapas temáticos, utilizando-se o *software* GeoDa (ANSELIN, 2004), calculadas auto-correlações espaciais, correlações espaciais univariada e multivariada, além da construção de outros tipos de mapas espaciais.

Cabe ressaltar também que no Modelo de Regressão Linear Múltipla Ponderada (MRLMP) utilizada para analisar os indicadores que influenciam e são influenciados pelo emprego formal, ao ser aplicado de forma a identificar o poder explicativo do emprego formal, utilizou-se como variável resposta ou independente, os indicadores de pobreza e posteriormente o novo índice de inclusão social (que será visto mais adiante). Neste segundo caso, foi considerado como variável explicativa, além do emprego formal, uma nova variável referente à transferência governamental, que não foi obtida do Atlas de Pochamn e Amorim, de forma que se tem uma análise complementar ao estudo. Contudo, no MRLMP foi tomada como variável resposta tanto a população formalmente empregada, quanto a pobreza e o novo índice de exclusão/inclusão social e, como variáveis explicativas, os demais indicadores socioeconômicos (sub-índices do IES) diferentes da variável resposta. Porém, no caso específico em que a variável resposta é o novo Índice de Inclusão Social (IIS) que será detalhado no capítulo 4, utilizou-se um indicador descrito como: Percentual de pessoas com mais de 50% da renda provenientes de transferência governamental (PP_50RTG), que foi tomada como variável explicativa no modelo de regressão, junto com a variável emprego formal, sendo a nova variável retirada do Atlas de Desenvolvimento Humano para o ano

2000. A análise foi feita através do *software* Estatística 7.0. A seguir serão descritas, sucintamente, cada uma das técnicas de análises utilizadas no trabalho.

2.1. TÉCNICAS DE ANÁLISES DOS DADOS

A fase da análise dos dados, de acordo com Sellitz (1967, p.123), representa o ato de “sumariar as observações completadas, de forma que estas permitam responder às perguntas de pesquisa”. Para tanto, utilizou-se, para auxiliar essas análises, os softwares Estatística 7.0 e GeoDa, nos quais foram utilizadas as seguintes análises:

- a) Análise Descritiva e Exploratória;
- b) Análise do Coeficiente de Correlação;
- c) Análise de Regressão Linear Múltipla;
- d) Análise de *cluster* (ou de agrupamento);
- e) Análise espacial

2.1.1. Análise Descritiva e Exploratória

Os dados secundários utilizados na pesquisa foram captados do Atlas e organizados em um banco de dados na planilha eletrônica Excel, para posterior leitura nos *softwares* Estatística e GeoDa. A partir desse banco de dados, foram aplicadas técnicas de Análises Descritiva e Exploratória, que consistem em ferramentas úteis na produção de tabelas, gráficos e medidas estatísticas que descrevem o perfil socioeconômico da Região Nordeste e seus respectivos estados, tomando o município como unidade de análise,

Dessa concepção, a Análise Descritiva e Exploratória de dados procura apresentar, de forma sucinta, o comportamento das variáveis: Taxa de emprego formal sobre a População em Idade Ativa (PIA), a Pobreza dos chefes de domicílio, a *prox* da Desigualdade de Renda, número médio de Anos de Estudo dos chefes de domicílio e percentagem de Jovens na população da Região Nordeste e nos seus respectivos Estados a nível de municípios, com ênfase no estudo da Taxa de emprego formal sobre a PIA.

2.1.2 Análise do Coeficiente de Correlação

Na Estatística, pode-se quantificar relações entre variáveis de diversas maneiras, sendo que a mais comum envolve o coeficiente de correlação linear de Pearson, o qual, quantifica o quão linear é essa relação.

Os valores do coeficiente de correlação (r) variam de -1 a +1. Valores próximos a -1 e +1 indicam forte correlação e valores próximos a zero indicam fraca correlação. Esses valores também retratam se a relação entre as variáveis é direta ou inversa, ou seja, se na medida em que um índice cresce, o outro cresce ou decresce. Melhor detalhamento é visto no Anexo (Descrição dos Métodos Estatísticos).

No estudo, este método foi utilizado nos capítulos 3 e 4, para analisar o grau da correlação (e a sua direção – se direta ou inversa) entre cada duas das seguintes variáveis de interesse: IND_EMP, IND_DES, IND_POB, IND_ESC, IND_JUV, IIS e PP_50RTG.

2.1.3. Análise de Regressão Linear Múltipla Ponderada.

A Análise de Regressão é uma técnica de análise que pode ser utilizada, com o auxílio de um software estatístico, para buscar resultados que tenha como proposta, a descrição do relacionamento existente entre uma determinada variável, dita como variável resposta ou variável dependente e um conjunto de variáveis explicativas, também conhecidas como variáveis independentes (DRAPER; SMITH, 1998).

A Análise de Regressão Linear Múltipla é utilizada para examinar a relação entre uma única variável dependente e um conjunto de variáveis independentes. Por apresentar ampla aplicabilidade, a Regressão Múltipla tem sido usada para muitos propósitos. Suas aplicações recaem em duas grandes classes de problemas de pesquisa: previsão e explicação, as quais não são mutuamente excludentes, e uma aplicação da análise de regressão múltipla, no caso, a exclusão/inclusão social e o emprego formal pode abordar qualquer um desses problemas de pesquisa ou ambos (HAIR JUNIOR et al., 2006).

Como o propósito deste trabalho não envolve previsões, o Modelo de Regressão Linear Múltipla é utilizado apenas para descrever o quanto cada variável independente explica a variável resposta.

Em grande parte dos casos, encontram-se intervalos de concentração amplos onde é de se esperar que a variância em cada ponto da reta seja diferente (MODELO DE

CALIBRAÇÃO, 2008), pelo que, a condição de homogeneidade de variância, designada por homocedasticidade, não é cumprida. Contudo, são aplicadas ações corretivas, de forma que não haja violação desta suposição.

A utilização do Modelo de Regressão Linear Múltipla Ponderada é a forma mais simples e efetiva de harmonizar as discrepâncias entre as variâncias dos diferentes pontos que constituem a reta (MODELO DE CALIBRAÇÃO, 2008).

Devido à variabilidade dos índices existente entre os municípios neste trabalho, a ponderação utilizada, para os municípios foi o tamanho de sua respectiva população de forma que assim foi melhorada a discrepância entre as variâncias dos diferentes pontos, deixando os resultados do modelo mais consistentes. Melhor detalhamento se encontra no Anexo (Descrição dos Métodos Estatísticos).

Neste trabalho, o MRLMP foi utilizado primeiramente no capítulo 3, para explicar como o emprego formal é influenciado pelos demais indicadores que compõe o Índice de Exclusão/Inclusão Social e em seguida, como ele influencia no indicador pobreza, tomando também os demais indicadores do IES para explicar esta variável. Por último, no capítulo 4, foi analisado como o emprego formal influencia no novo Índice de Inclusão Social, gerado da composição do IES sem a presença do emprego formal. No entanto, para este caso, foi utilizada como variável explicativa além do emprego formal, a variável referente à transferência governamental (citado anteriormente), de forma que se pôde observar qual variável influencia mais no novo Índice de Inclusão Social (IIS): se o emprego formal ou as transferências governamentais. Ressalta-se que os dados utilizados no trabalho são do censo 2000.

2.1.4. Análise de *cluster* (ou de agrupamento)

A análise de agrupamento é uma técnica exploratória de dados multivariados que tem por finalidade agrupar elementos ou variáveis. Os elementos agrupados devem ter as seguintes características: homogeneidade dentro e heterogeneidade entre os grupos, ou seja, dentro dos grupos os elementos devem ser “similares” e entre os grupos devem ser “diferentes”.

A análise de agrupamentos (ou *cluster analysis*), a ser trabalhada no Capítulo 4, foi realizada com o auxílio do *software* Statistica 7.0, fazendo-se necessária a escolha de variáveis a partir das quais os grupos similares são formados. Neste estudo fez-se opção pelas

variáveis emprego formal (IND_EMP) e o novo Índice de Inclusão social (IIS). A princípio, foram considerados os 26 estados e o Distrito Federal, que formam agregados estaduais segundo as referidas características consideradas. Em seguida, foi estabelecida a análise de agrupamento para a região Nordeste, área geográfica, objeto de estudo do trabalho. Para a formação dos agrupamentos de municípios, utilizou-se a regra da ligação simples (*single linkage*) semelhantes e medidas de distâncias Euclidianas (*Euclidean distances*), usando ainda o critério *two-way joining* para enfatizar ainda mais a associação entre essas variáveis. Dessa forma, as análises foram centradas para o país como um todo, de forma a mostrar que as regiões Norte e Nordeste permanecem em grupos isolados por apresentarem índices inferiores aos das demais regiões, e mais especificamente, para a região Nordeste. Os resultados ilustrados pelos respectivos dendogramas (no método de ligação simples) e por Gráficos dos resultados cruzados (no método *two-way joining results*) para as variáveis IND_EMP e IIS, permitiram uma boa percepção e classificação dos grupamentos obtidos, tendo sido gerados no *software* Statistica 7.0.

Para o método de ligação simples, usou-se a distância euclidiana (DE), segmento de linha reta entre cada par de observações, para medir a proximidade de cada Unidade Federativa no universo da região Nordeste; verifica-se que, quanto menor a distância entre os estados, maior a similaridade entre eles. Quanto ao método *Two-Way joining*, é representado por um gráfico de escala de cores que expressa na leitura vertical, os diferenciais entre cada índice composto (quão desigual é a realidade nos vários estados), e na horizontal, uma diferenciação dos vários índices dentro de cada estado. Melhor detalhamento sobre a análise de *cluster* é visto no Anexo (Descrição dos Métodos Estatísticos).

2.1.5. Análise espacial

A análise espacial de dados tem a finalidade de encontrar evidências de concentração espacial. Esta técnica inclui métodos estatísticos que procuram descrever a variação espacial do fenômeno em estudo, a partir de amostras disponíveis, detectando áreas onde ocorre um grande valor de risco. Com estes dados podem-se descobrir possíveis fatores de risco que explicarão este excesso nas áreas observadas. Assim, é possível estabelecer hipóteses e testá-las em diferentes conjuntos de dados e com diferentes metodologias.

A análise espacial dos dados foi feita através do GeoDa, *software* de domínio público (ANSELIN, 2004) e incluiu a definição dos critérios de vizinhança e cálculos das respectivas

matrizes de proximidade espacial, cálculo dos coeficientes de auto-correlações espaciais de Moran uni e multivariado, realização dos testes de significância e construção dos mapas ilustrativos da inclusão/exclusão social da região Nordeste e de sua associação espacial com o emprego formal.

O índice de Moran (I) tem seu equivalente no coeficiente de correlação linear r , só que avaliado em termos geoespaciais, onde é medida a correlação da própria variável em áreas vizinhas (Moran Univariado) ou entre duas variáveis especificadas (Moran Multivariado). Seu intervalo de variação, assim como o r , é entre -1 e 1. Quanto maior o valor de I, em valor absoluto, maior a correlação espacial. Os valores próximos ao zero são indicativos de ausência de correlação espacial.

Este índice indica se há tendência a agrupamentos (*clusters*) de municípios onde as taxas sejam mais baixas, ou mais altas, relativamente a outras áreas. Também se lançou mão dos mapas temáticos, especialmente o mapa representativo do *Box plot* dos dados (*Box Map*), o mapa do desvio padrão (*Standard Deviation*) e o mapa que identifica agrupamentos espaciais significantes (*Lisa clusters map*, indicadores locais de associação espacial). Melhor detalhamento da Análise Espacial é visto no Anexo (Descrição dos Métodos Estatísticos).

CAPÍTULO 3: IDENTIFICAÇÃO DOS INDICADORES SOCIOECONÔMICOS QUE SÃO INFLUENCIADOS E INFLUENCIAM O EMPREGO FORMAL NAS UNIDADES FEDERATIVAS NORDESTINAS.

3.1. EXPLORANDO OS INDICADORES.

A Análise Exploratória das informações levantadas, com base nos dados do Atlas de Exclusão Social, de Pochmann e Amorim (2003), compreendem o comportamento dos indicadores, IND_POB, IND_EMP, IND_DES, IND_ESC, IND_JUV para o contexto de suas grandes regiões, dando ênfase à região Nordeste brasileira. Ressalta-se que apenas o IND_VIO e o IND_ALF não foram analisados, por razões posteriormente apresentadas.

Cabe ainda destacar que a análise dos dados realizada, tomando por base as classificações apresentadas no Quadro 2.2, para o Índice de Exclusão Social, define os seguintes intervalos: de 0,0 a 0,4, elevada exclusão; de 0,4 a 0,5, média exclusão; de 0,5 a 0,6, média/ baixa inclusão; de 0,6 a 0,8, média/ alta inclusão e de 0,8 a 1,0, elevada inclusão. Na medida em que os indicadores estudados mostram índices elevados (acima de 0,8) implicará em uma elevada inclusão social da área estudada.

O comportamento dos indicadores socioeconômicos para as Unidades Federativas, são apresentados na Figura 3.1. Verifica-se que, há uma elevada exclusão social apresentando todos indicadores sob estudo valores entre 0,001 e 0,386, para todas as UF's. A exceção é apenas para o IND_JUV onde este intervalo aumenta. A “Desigualdade de Renda” (IND_DES) é muito elevada. Os estados que apresentaram melhores índices foram o Rio Grande do Norte (0,1), Pernambuco (0,09) e, com piores índices, os estados do Maranhão (0,001) e Piauí (0,02), ou seja, o problema da desigualdade de renda, além de ser antigo, continua sem reversão. A “Escolaridade dos chefes de família” (IND_ESC) também é bastante precária, apresentando maior índice o estado do Rio Grande do Norte (0,28) e Pernambuco (0,30) mostrando assim como também é baixo o nível de educação da população. A problemática da “Pobreza dos chefes de família no município” (IND_POB) ainda é pior do que IND_ESC e o IND_EMP, apresentando-se com maior índice o estado de Pernambuco e Rio Grande do Norte, ambos com 0,23 e com pior índice o estado do Maranhão, com 0,001. O IND_EMP apresenta seu melhor índice com 0,31 para o estado do Rio Grande do Norte e em seguida o estado de Sergipe (0,30); os demais estados estão classificados na seguinte ordem: Pernambuco (0,26), Paraíba (0,22), Bahia (0,22), Alagoas (0,21), Ceará (0,18), Piauí (0,07) e Maranhão (0,001). A Região Nordeste se apresenta de maneira bastante singular. Sabe-se que

é a região mais carente do país e seus indicadores são bastante baixos e, numa análise mais ampla, dentro do contexto nacional, a mancha da exclusão social mostra-se intensa e homogênea.

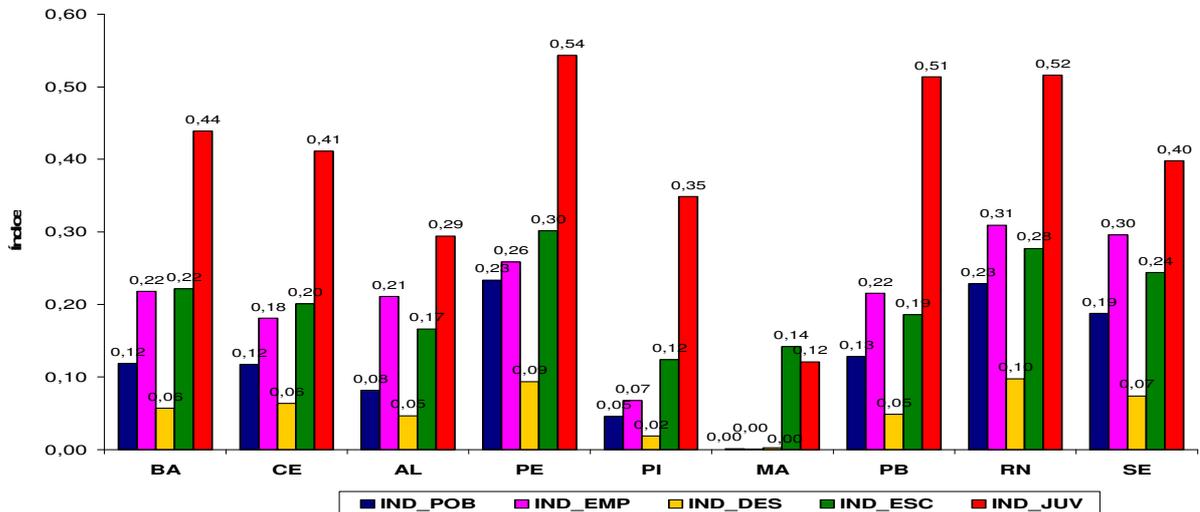


Figura 3.1: Comportamento das variáveis IND_EMP, IND_POB, IND_DES, IND_ESC e IND_JUV para as UF's da região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Quanto ao Índice de Exclusão Social (IES) em si, retirado do Atlas de Pochmann e Amorim (2003), percebe-se que, dentre os Estados da região Nordeste, o que mais se destacou, com uma menor exclusão social (melhor IES) é o Estado do Rio Grande do Norte (0,386) e o de maior exclusão (pior IES) é o Estado do Maranhão (0,197). Observe-se, contudo, que o intervalo desses valores, para as unidades federativas da região, evidencia elevada exclusão, estando todas, com índices abaixo de 0,4.

Segundo dados do IBGE, em 1996 o Estado do Maranhão apresentava os piores indicadores sociais, estando os números bastante elucidativos. Quando se estudam os indicadores no ano 2000, o estado permanece na mesma situação, disputando o último lugar com o Piauí. (SILVA, 2001, p. 199-226).

Reconhecidamente, o Brasil, no geral, apresenta um alto índice de pobreza e elevada desigualdade na distribuição da renda. No contexto de desigualdades, o Nordeste situa-se como uma região onde prevalecem condições sociais precárias para a grande maioria da população, que refletem, sobremaneira, na sua qualidade de vida.

Para Jannuzzi apud Barbosa e Ribeiro (2006), os indicadores sociais se prestam a subsidiar as atividades de planejamento público e formulação de políticas sociais nas diferentes esferas de governo, possibilitam o monitoramento das condições de vida e bem-

estar da população por parte do poder público e da sociedade civil e permitem aprofundamento da investigação acadêmica sobre a mudança social e sobre os determinantes de diferentes fenômenos sociais (BARBOSA; RIBEIRO, 2006).

O patamar de pobreza na região Nordeste se mostra sem igual no país, em 1993, 70% era pobre passando em 2004 para 80% da população nordestina, situando-se em pior situação dentre as demais regiões do Brasil. Isto mostra que as grandes desigualdades sociais continuam a mesma, estando as regiões com melhores indicadores localizadas abaixo do Trópico de Capricórnio e as com piores localizadas acima desse Trópico, permanecendo, desde há muito tempo, com estas características (FALVO, 2007).

Quando se trata do emprego formal, tem-se que, numa região marcada pela pobreza, com grandes segmentos de informalidade no trabalho, tem tido alguma semelhança com aquilo que ocorre no país como um todo: a precarização do trabalho – uma grande parcela da população, por não ter alternativa, termina em trabalhos que não dão o retorno que querem ter. A formalidade do emprego é muito baixa em todo o país e, em especial, na região Nordeste.

Analisando as capitais dos Estados do Nordeste, ilustrado na Figura 3.2, observa-se que o IND_EMP e o IND_DES são os indicadores com índices mais baixos para todas as 9 (nove) capitais – com valores inferiores a 0,4, que indica “elevada exclusão”. Quanto ao IND_JUV, tem-se que todas as capitais estão com índice acima de 0,6, o que representa um razoável percentual de jovens nas respectivas populações, a maioria deles desempregados ou sub-empregados, fato considerado de grande vulnerabilidade social, por serem elevadas as chances de envolvimento desses jovens com o crime e as drogas. O IND_ESC também se apresenta com índices acima de 0,6 em todas as capitais e, inclusive, acima de 0,7, na sua grande maioria, ou seja, para o estudo das capitais tem-se um bom grau de escolaridade dos chefes de domicílio na região. Para o IND_POB, observa-se que apenas as capitais João Pessoa - PB e Natal - RN apresentaram índices acima de 0,6, reveladores de maior inclusão social, estando as demais capitais entre 0,5 e 0,6, ou seja, a pobreza também é muito evidenciada nas capitais dos respectivos estados da Região Nordeste. Ficou evidente, então, que a pobreza se apresenta de maneira mais precária quando comparada à escolaridade, ou seja, deve haver não só incentivos de políticas públicas de maneira a melhorar a educação, mas, ao mesmo tempo, a busca de uma diminuição da pobreza, pois este é um dos mais graves problemas da região Nordeste. Uma das saídas seria a oferta de trabalhos dignos para a população, fazendo com que melhorem sua qualidade de vida. Ou ainda, dinamizar a

economia, de forma a gerar empregos formais para os vários níveis de escolaridade da população brasileira.

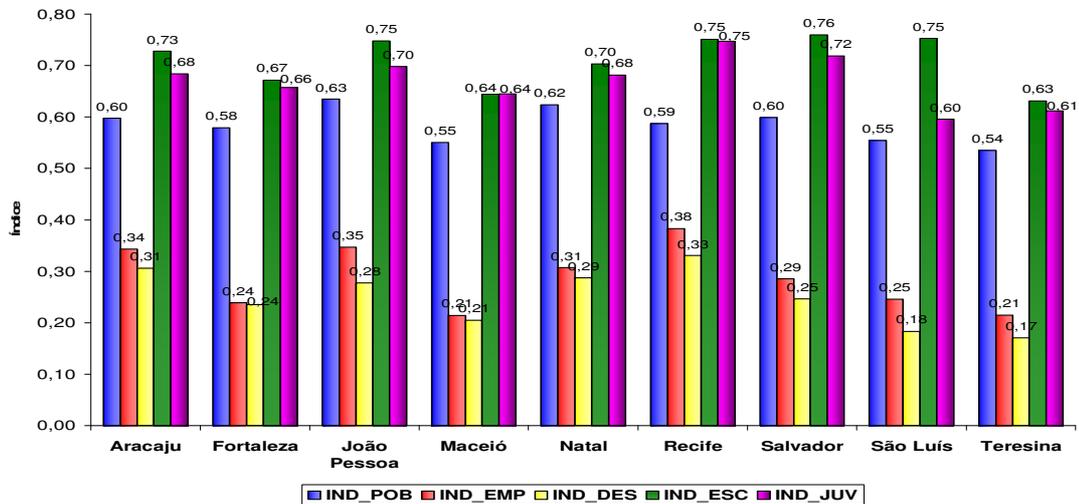


Figura 3.2: Comportamento das variáveis, IND_EMP, IND_POB, IND_DES, IND_ESC e IND_JUV para as capitais de cada UF's da região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Com ênfase no indicador IND_EMP mostra-se na Figura 3.3, os *Box-plots* da distribuição deste índice dentro de cada Unidade Federativa (UF's) da região Nordeste, tendo como unidade de análise o município. Pode-se perceber, nessa figura, que todas as UF's possuem muitos municípios com valores extremos e *outliers* para o IND_EMP, revelando uma distribuição geral quase uniforme, em termos comparativos.

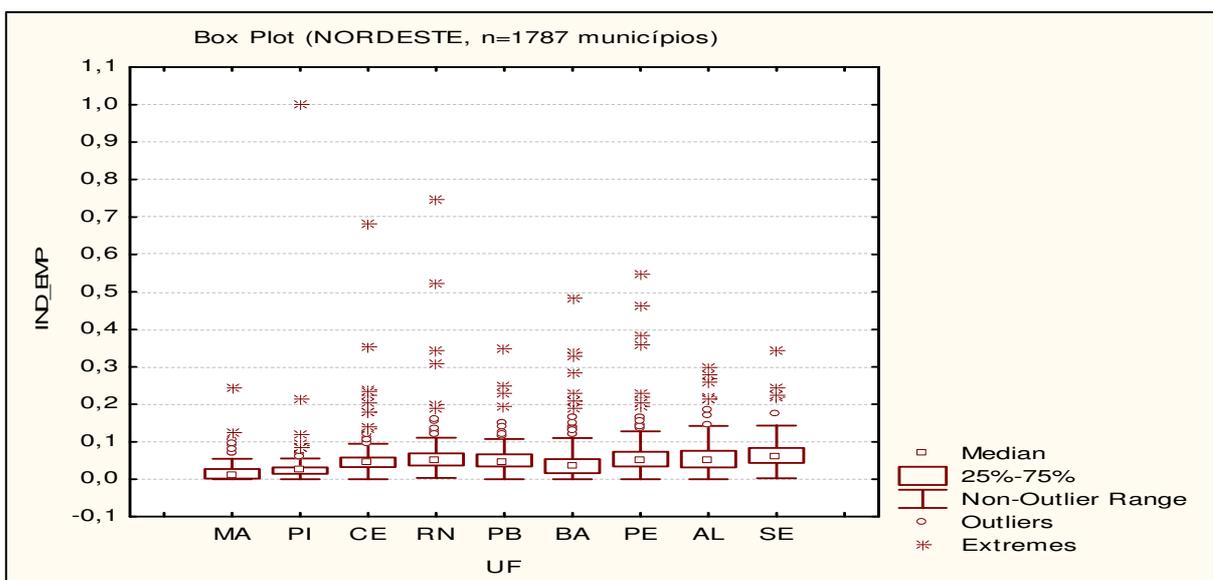


Figura 3.3: *Box-plot* das Taxas de Emprego Formal dos municípios do Nordeste segundo seus respectivos Estados.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

No contexto regional, apenas 7 municípios, dos 1787 que faziam parte da Região Nordeste, apresentaram o IND_EMP acima de 0,4. Logo, 99,6% do total de municípios tiveram índices bastante baixos (chamada de “elevada exclusão”), indicando uma precarização do trabalho na região. Os municípios que tiveram valores acima de 0,4 se apresentam classificados na seguinte ordem crescente: Porto (PI) – 1,00, Maxaranguape (RN) – 0,75, Eusébio (CE)- 0,68, Camutanga (PE) – 0,55, Baía Formosa (RN)- 0,52, Lauro de Freitas (BA)- 0,48 e Joaquim Nabuco (PE) – 0,46 . Acrescenta-se que, tanto esses sete municípios, quanto as dezenas de outros que possuem o IND_EMP entre 0,2 e 0,4, são todos valores extremos superiores, isto é, considerados muito afastados da distribuição geral dos dados municípios, dentro de cada Unidade Federativa. As estatísticas descritivas do IND_EMP encontram-se sumarizados no Quadro 3.1.

Quadro 3.1
Estatísticas Descritivas para a variável IND_EMP, segundo UFs da região Nordeste, 2000.

UFs da Região	Estatísticas Descritivas do IND_EMP						
	Total (N)	Mediana	Mínimo	Máximo	Quartil 1	Quartil 3	Desvio Padrão
NORDESTE	1787	0,04	0,00	1,00	0,02	0,06	0,06
MA	217	0,01	0,00	0,25	0,00	0,03	0,03
PI	221	0,02	0,00	1,00	0,01	0,03	0,07
CE	184	0,04	0,00	0,68	0,03	0,06	0,06
RN	166	0,05	0,00	0,75	0,04	0,07	0,08
PB	223	0,05	0,00	0,35	0,03	0,07	0,04
PE	185	0,05	0,00	0,55	0,03	0,07	0,07
AL	101	0,05	0,00	0,3	0,03	0,08	0,06
SE	75	0,06	0,00	0,34	0,04	0,08	0,06
BA	415	0,03	0,00	0,48	0,02	0,05	0,05

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

De acordo com a Figura 3.3 e o Quadro 3.1, o IND_EMP em todos os estados está com uma mediana em torno de 0,04, refletindo o baixo valor do índice para todas as UFs, pois 50% ou mais dos municípios apresenta o IND_EMP abaixo de 0,04. Outra característica é que apenas um município tem todas as pessoas empregadas formalmente (índice 1,00 – que indica inclusão total) estando localizado no estado da Piauí, sendo identificado como valor extremo. Destaca-se ainda, que apenas três municípios do Nordeste apresentaram índice maior que 0,6, indicando uma média/alta inclusão, sendo classificados na seguinte ordem crescente: Porto – PI (1,00), Maxaranguape - RN (0,75) e Eusébio - CE (0,68). Contudo, o estado do Piauí, juntamente com o estado o Maranhão, apresentaram a mais baixa distribuição desses dados, na qual 75% dos municípios nesses estados possuem valores menores ou iguais a 0,03. De

acordo com o desvio-padrão apresentado (Quadro 3.1), fica evidenciado que há uma baixa variabilidade dos dados, tanto para região Nordeste como para seus respectivos estados, indicando assim, uma grande concentração em baixos níveis de emprego formal. O valor que mais se repete para a região Nordeste (1787 municípios) é o índice igual a zero, correspondendo a 40 municípios, ou seja, nestes municípios nenhum indivíduo trabalha com carteira de trabalho assinada. Observa-se também que desses 40 municípios, praticamente todos estão localizados nos estados do Piauí (15) ou Maranhão (24), com exceção apenas de um que se localiza no estado da Paraíba (Gado Bravo). Conclui-se que há um baixíssimo nível de emprego formal na região como um todo, estando a situação ainda mais precária nos estados do Maranhão e Piauí. Necessitando, contudo, de uma melhora radical em todos os municípios da região Nordeste.

Nessa análise dos indicadores que influenciam e são influenciados pelo emprego formal, torna-se indispensável avaliar a magnitude das correlações entre os índices de exclusão social associados a esses indicadores, os quais, como se pode ver pela matriz de correlação (Quadro 3.2) apresentaram significância estatística para todas as associações, ao nível de 5%.

Quadro 3.2

Matriz de correlações entre as variáveis do estudo, região Nordeste (1757 municípios), 2000.

Variáveis	IND_POB	IND_EMP	IND_DES	IND_ESC	IND_JUV
IND_POB	1,00	0,57	0,78	0,80	0,39
IND_EMP	0,57	1,00	0,63	0,63	0,35
IND_DES	0,78	0,63	1,00	0,76	0,28
IND_ESC	0,80	0,63	0,76	1,00	0,39
IND_JUV	0,39	0,35	0,28	0,39	1,00

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

* Correlações significativas ao nível de 5%.

A mais forte correlação foi encontrada entre o IND_POB com o IND_ESC (0,80). Quando se compara o IND_EMP com as demais, tem-se que as maiores correlações foram entre IND_ESC e o IND_DES, ambas com um valor de 0,63. Por sua vez, o IND_DES tem sua maior correlação com o IND_POB (0,78), seguida pela correlação com IND_ESC (0,76). As mais baixas correlações são observadas para o IND_JUV com as demais variáveis (abaixo do valor 0,4). Mesmo assim, apresentam significância estatística (p-valor <0,05), onde a

variável com maior correlação com este indicador foi o IND_POB e em seguida o IND_ESC ambas com um valor de 0,39.

Observando as correlações da variável IND_EMP com os demais indicadores sob estudo, para cada Unidade Federativa da Região Nordeste (Quadro 3.3), tem-se que nos estados do Maranhão, Piauí, Paraíba, Sergipe e Bahia, a variável com maior correlação é o IND_DES com, respectivamente, 0,70; 0,56; 0,69, 0,74 e 0,72. Para os demais estados - Ceará, Rio Grande do Norte e Pernambuco - a variável que está mais fortemente associada com o IND_EMP é o IND_ESC com, respectivamente, 0,64; 0,59; 0,69. A exceção é apenas para o estado do Alagoas, que teve o IND_POB como o indicador mais associado ao IND_EMP, com um valor de 0,58. Todas as associações foram significantes, ao nível de 5%.

Quadro 3.3

Correlações entre o IND_EMP e as demais variáveis sob estudo das UFs da região Nordeste, 2000.

IND_EMP e Variáveis	UFs e Número de municípios (n)								
	MA (n=217)	PI (n=220)	CE (n=181)	RN (n=164)	PB (n=221)	PE (n=178)	AL (n=101)	SE (n=72)	BA (n=415)
IND_POB	0,57	0,31	0,55	0,54	0,57	0,59	0,58	0,56	0,70
IND_DES	0,70	0,56	0,51	0,46	0,69	0,65	0,40	0,74	0,72
IND_ESC	0,63	0,53	0,64	0,59	0,64	0,69	0,47	0,66	0,70
IND_JUV	0,36	0,29	0,22	0,25	0,25	0,38	0,28	0,28	0,23

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Contudo, observa-se que em três estados (PE, CE e RN), o IND_EMP aparece mais correlacionado com a IND_ESC e, em outros cinco estados (MA, PI, PB, SE e BA) se apresenta mais correlacionado com o IND_DES. Em todos os nove estados da região, incluindo Alagoas, que teve o IND_EMP mais associado com o IND_POB, apresentou como segundo lugar na correlação com o IND_EMP a variável IND_ESC. Portanto, percebe-se que o IND_ESC está sempre bem associado com os demais indicadores, sendo a mais forte ou segunda mais forte associação. Logo, a melhoria na educação da população irá remeter, naturalmente, a uma melhoria em outros indicadores, como o emprego formal e, conseqüentemente, na redução da pobreza.

Para apresentar uma análise ampla dos cinco indicadores estudados até o momento, será mostrada, nos próximos itens, a aplicação do modelo de regressão linear múltipla ponderada, que permitiu identificar os indicadores socioeconômicos que são influenciados e influenciam o Emprego formal, tomando o município como unidade de análise e observando

esse comportamento, tanto na Região Nordeste, como um todo, quanto para cada uma de suas nove Unidades Federativas.

3.2. RELAÇÃO ENTRE A “TAXA DE EMPREGO FORMAL SOBRE A PIA (IND_EMP)” E OUTRAS VARIÁVEIS SOCIOECONÔMICAS.

Para avaliar as possíveis relações de dependência entre as variáveis pesquisadas, empregou-se a técnica da Análise de Regressão Linear Múltipla, na qual a “taxa de emprego formal sobre a PIA (IND_EMP)”, foi definida como a variável dependente ou variável resposta, escrita em função das demais, definidas como variáveis independentes ou explicativas; tendo sempre o município como unidade de análise.

O ajuste do modelo de regressão linear múltipla ponderada (MRLMP) na Tabela 3.1 mostra que, para a Região Nordeste, todas as variáveis foram estatisticamente significantes, ao nível de 5%, para explicar o IND_EMP: “pobreza dos chefes de família do município (IND_POB)”, “*proxi* da desigualdade renda (IND_DES)”, “anos médio de estudo do chefe de domicílio (IND_ESC)” e “porcentagem de jovens na população (IND_JUV)”.

É importante destacar que esta seleção de variáveis foi feita após alguns testes do ajuste do modelo de regressão linear múltipla ponderada, para que fosse empregado um modelo representativo para o IND_EMP como variável dependente, no contexto das demais variáveis explicativas. Mediante esta relação, duas variáveis do Atlas da Exclusão Social no Brasil (POCHAMANN; AMORIM, 2003) não foram inseridas no modelo: a primeira variável foi o “número de homicídios por 100 mil habitantes (IND_VIO)”, que não foi colocada no modelo devido à sua não significância; e a segunda variável foi a “taxa de alfabetização de pessoas com mais de 5 anos (IND_ALF), pois se observou que bastava uma variável que representasse a educação (dentre as duas do Atlas IND_ESC e IND_ALF) para evitar uma multicolinearidade na análise. Além disso, constatou-se que o IND_ESC apresentou sempre uma melhor explicação para o IND_EMP do que o IND_ALF.

O sumário da RLMP para os 1787 municípios da região Nordeste com os respectivos coeficientes “ β ” (betas) e erros padrões, bem como os valores do teste “t” e os p-valores para as significâncias estatísticas de cada variável incluída no modelo em análise, está mostrado na Tabela 3.1.

Tabela 3.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Região Nordeste, 2000 (n = 1757 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			-0,01	0,01	-1,50	<0,1338
IND_POB	-0,16	0,03	-0,10	0,02	-5,32	<0,0001
IND_DES	0,84	0,02	0,92	0,02	44,29	<0,0001
IND_ESC	0,24	0,03	0,12	0,02	7,12	<0,0001
IND_JUV	0,04	0,01	0,04	0,01	3,63	<0,0003

R= 0,95 R²= 0,91 R² Ajustado = 0,91

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Como se pode ver na Tabela 3.1, o modelo apresenta um coeficiente de correlação múltipla R= 0,95, e um coeficiente de determinação ajustado R²= 0,91, indicando que 91% da variação dos dados foi explicada pelas variáveis independentes IND_DES, IND_ESC, IND_POB e IND_JUV. Tanto nesse processo de escolha, quanto na opção *forward stepwise*, todas as quatro variáveis permaneceram estatisticamente significantes (p-valor <0,05), ressaltando, na opção *stepwise*, que o IND_DES sozinho contribuiu com 0,90 da explicação total dos dados, dado por R². Pode-se dizer, portanto, que a desigualdade de renda representada pelo IND_DES foi a variável de maior poder explicativo para o emprego formal (IND_EMP) na Região Nordeste, dentre as variáveis envolvidas no modelo. Na verdade, o IND_EMP mediu a desigualdade de renda de chefe de família dos municípios em análise, estando fortemente associado à pobreza, educação e o maior contingente de jovens, na maioria, desempregados.

Dessa forma, analisando o ajuste do MRLMP para cada estado da Região Nordeste, tomando o IND_EMP como variável resposta (e o município como unidade de análise), pode-se observar resultados semelhantes, com algumas variações peculiares, conforme mostrado nas Tabelas 3.2 a 3.10.

Para o Estado do Maranhão (Tabela 3.2) observa-se que os indicadores: IND_POB, IND_DES e IND_JUV contribuíram significativamente para explicar o emprego formal do estado (p-valor < 0,01). Apresentando um coeficiente de correlação múltipla de R= 0,97 e um coeficiente de determinação ajustado de R²= 0,95. No processo de escolha *forward stepwise* estes três indicadores permaneceram significantes no modelo (p-valor <0,05), porém o IND_DES teve a maior contribuição, com 0,93 no ganho do R², na explicação do emprego formal, mostrando assim, a importância de políticas voltadas à redução da desigualdade no estado, bem como em todo o Nordeste, na busca de melhores condições de emprego da população.

Tabela 3.2

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Maranhão, 2000 (217 municípios)

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			-0,02	0,01	-1,82	<0,0703
IND_POB	-0,33	0,05	-0,18	0,03	-6,70	<0,0001
IND_DES	1,10	0,05	1,51	0,07	21,46	<0,0001
IND_ESC	0,11	0,06	0,05	0,02	1,92	<0,0567
IND_JUV	0,09	0,03	0,08	0,03	2,84	<0,0050

R= 0,97 R²= 0,95 R² Ajustado = 0,95

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Quanto ao Estado do Piauí (Tabela 3.3), tem-se um comportamento bastante parecido com a do estado do Maranhão, analisado anteriormente, onde os indicadores que contribuíram significativamente para o IND_EMP foram IND_POB, IND_DES, IND_JUV (p-valor<0,5). O coeficiente de correlação múltipla foi de R= 0,98 e o de determinação ajustado de R²= 0,96. Para o procedimento *forward stepwise*, os três indicadores permaneceram significantes no modelo (p-valor <0,05), estando o IND_DES com a maior contribuição (0,95) na explicação do emprego formal, traduzindo o mesmo comportamento verificado nos ajustes anteriores do modelo.

Tabela 3.3

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Piauí, 2000 (n = 220 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,01	0,01	1,41	<0,1590
IND_POB	-0,38	0,05	-0,21	0,03	-7,19	<0,0001
IND_DES	1,28	0,07	1,59	0,08	19,08	<0,0001
IND_ESC	0,04	0,06	0,02	0,03	0,60	<0,5495
IND_JUV	0,04	0,02	0,05	0,02	2,35	<0,0195

R= 0,98 R²= 0,96 R² Ajustado = 0,96

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado do Ceará (Tabela 3.4) observa-se que apenas os indicadores IND_DES e IND_ESC contribuíram significativamente para explicar o IND_EMP (p-valor< 0,01). O coeficiente de correlação múltipla das variáveis envolvidas no modelo foi de R= 0,95 e o coeficiente de determinação ajustado R² igual a 0,90. Na opção *forward stepwise* ambas as variáveis permaneceram significativas no modelo (p-valor <0,05), onde apresentaram uma

contribuição do ganho do R^2 na seguinte ordem: IND_ESC com 0,89 e IND_DES com 0,01; mostrando, que, para os municípios do estado do Ceará deverá haver uma maior preocupação na busca de melhoria na educação de forma a influenciar, positivamente, a taxa de emprego formal sobre a PIA.

Tabela 3.4

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Ceará (181 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			-0,02	0,02	-1,04	<0,3020
IND_POB	-0,05	0,14	-0,03	0,07	-0,37	<0,7112
IND_DES	0,34	0,07	0,32	0,07	4,84	<0,0001
IND_ESC	0,68	0,15	0,31	0,07	4,51	<0,0001
IND_JUV	-0,01	0,04	-0,01	0,05	-0,13	<0,8968

R= 0,95 $R^2= 0,91$ R^2 Ajustado = 0,90

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado do Rio Grande do Norte (Tabela 3.5), pode-se observar que apenas o IND_DES contribuiu significativamente para explicar o IND_EMP (p-valor < 0,01), apresentando um coeficiente de correlação múltipla de $R= 0,94$ e um coeficiente de determinação ajustado de $R^2= 0,89$, mostrando, dessa forma, uma boa representatividade do modelo. No procedimento *forward stepwise* o IND_DES e o IND_ESC permaneceram no modelo, porém apenas o IND_DES contribuiu significativamente (p-valor <0,05) para explicar o IND_EMP; apresentando ainda o IND_DES uma contribuição na variação total de 89%, traduzindo a representatividade total do modelo.

Tabela 3.5

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Rio Grande do Norte (n = 164 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,02	0,03	0,54	<0,5876
IND_POB	-0,01	0,15	-0,01	0,09	-0,06	<0,9557
IND_DES	0,84	0,09	0,80	0,09	9,08	<0,0001
IND_ESC	0,11	0,18	0,06	0,10	0,61	<0,5448
IND_JUV	0,01	0,04	0,01	0,05	0,26	<0,7933

R= 0,94 $R^2= 0,89$ R^2 Ajustado= 0,89

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

O Estado da Paraíba (Tabela 3.6) apresentou um resultado parecido com o estado do Rio Grande do Norte onde apenas o IND_DES apresentou-se significativo para explicar o

IND_EMP (p-valor < 0,01). O coeficiente de correlação múltipla e o coeficiente de determinação ajustado foram respectivamente 0,96 e 0,92, mostrando que o modelo está bem representado pelas variáveis do estudo. Já quando se utiliza o processo de escolha *forward stepwise* o IND_DES e o IND_POB se mostraram significativos (p-valor <0,05) na explicação do IND_EMP. O IND_DES, porém, apresentou sozinho uma contribuição de 92% na variação total dos dados, correspondente ao valor do R².

Tabela 3.6

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado da Paraíba (n = 221 municípios).

Variáveis	β padronizado	Erro padrão de β	<i>B</i>	Erro padrão de <i>B</i>	t(213)	p-valor
Intercept			0,04	0,02	1,55	<0,1217
IND_POB	0,08	0,10	0,05	0,06	0,85	<0,3990
IND_DES	0,81	0,06	0,90	0,07	12,71	<0,0001
IND_ESC	0,09	0,12	0,05	0,07	0,76	<0,4459
IND_JUV	-0,03	0,03	-0,05	0,05	-1,01	<0,3117

R= 0,96 R²= 0,92 R² Ajustado = 0,92

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Quanto ao Estado de Pernambuco (Tabela 3.7), apenas o IND_POB e o IND_DES apresentaram significativos no modelo (p-valor < 0,01). O coeficiente de correlação múltipla das variáveis envolvidas no modelo foi de R= 0,97 e um coeficiente de determinação ajustado de R²= 0,93. No processo de seleção *forward stepwise* os índices de pobreza (IND_POB), desigualdade (IND_DES) e escolaridade (IND_ESC) permaneceram no modelo, todavia apenas os dois primeiros foram significativos (p-valor <0,05). O IND_DES apresentou uma contribuição de ordem de 92% e o IND_POB de 1%, na explicação do IND_EMP.

Tabela 3.7

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado de Pernambuco (n = 178 municípios).

Variáveis	β padronizado	Erro padrão de β	<i>B</i>	Erro padrão de <i>B</i>	t(213)	p-valor
Intercept			0,05	0,02	2,05	<0,0414
IND_POB	-0,31	0,08	-0,24	0,06	-3,79	<0,0002
IND_DES	1,07	0,04	1,15	0,05	25,18	<0,0001
IND_ESC	0,15	0,11	0,09	0,06	1,41	<0,1596
IND_JUV	0,01	0,04	0,02	0,06	0,31	<0,7567

R= 0,97 R²= 0,93 R² Ajustado = 0,93

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

No ajuste do modelo para o Estado de Alagoas (Tabela 3.8), observa-se que apenas o IND_POB e o IND_JUV foram significativos para explicar o IND_EMP (p-valor < 0,01).; estando o modelo com um coeficiente de correlação múltipla de $R= 0,83$ e um coeficiente de determinação ajustado de $R^2= 0,68$. No processo de escolha *forward stepwise* os IND_POB, IND_JUV e o IND_ESC se apresentaram no modelo, permanecendo apenas o IND_POB e o IND_JUV estatisticamente significantes (p-valor <0,05), onde estes tiveram suas contribuições no valor do R^2 de 63% e 5% respectivamente.

Tabela 3.8

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado de Alagoas (n = 101 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,14	0,05	2,68	<0,0088
IND_POB	1,71	0,28	0,89	0,14	6,20	<0,0001
IND_DES	-0,13	0,24	-0,13	0,24	-0,54	<0,5876
IND_ESC	-0,36	0,39	-0,16	0,18	-0,92	<0,3605
IND_JUV	-0,49	0,15	-0,43	0,14	-3,17	<0,0021

$R= 0,83$ $R^2= 0,69$ R^2 Ajustado = 0,68

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado de Sergipe (Tabela 3.9), o IND_DES e o IND_ESC foram os únicos significativos (p-valor <0,05) no ajuste do modelo. O coeficiente de correlação múltipla das variáveis envolvidas foi de $R=0,99$ e o coeficiente de determinação ajustado de $R^2= 0,97$. No procedimento *forward stepwise* os IND_DES e o IND_ESC permaneceram significativos no modelo, estando com suas contribuições no valor do R^2 de 96% e 1% respectivamente.

Tabela 3.9

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Sergipe (n = 72 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,02	0,03	0,85	<0,3988
IND_POB	-0,07	0,09	-0,05	0,07	-0,69	<0,4907
IND_DES	0,78	0,05	0,78	0,05	14,40	<0,0001
IND_ESC	0,31	0,10	0,19	0,06	3,10	<0,0028
IND_JUV	-0,03	0,04	-0,04	0,06	-0,63	<0,5283

$R= 0,99$ $R^2= 0,97$ R^2 Ajustado = 0,97

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Finalmente, no ajuste do modelo para o Estado da Bahia (Tabela 3.10), três indicadores se apresentaram significativos (p-valor <0,05): IND_DES, IND_ESC e o IND_JUV. No processo de escolha forward stepwise estes três indicadores permaneceram significativos, todavia o IND_ESC sozinho contribuiu com 85% e em seguida o IND_DES com 1% da variação total dos dados. O modelo mostra-se com um coeficiente de correlação múltipla de 0,93 e um coeficiente de determinação ajustado de $R^2 = 0,86$, de modo que a educação se apresenta como a variável de maior poder explicativo para o IND_EMP no Estado da Bahia.

Tabela 3.10

Sumário de regressão linear múltipla ponderada para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado da Bahia (n = 415 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,00	0,02	0,08	<0,9332
IND_POB	0,05	0,07	0,03	0,05	0,62	<0,5373
IND_DES	0,40	0,06	0,49	0,08	6,19	<0,0001
IND_ESC	0,57	0,10	0,29	0,05	5,58	<0,0001
IND_JUV	-0,09	0,03	-0,10	0,04	-2,65	<0,0083

R= 0,93 $R^2 = 0,86$ R^2 Ajustado = 0,86

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Em resumo, tem-se que o ajuste do MRLMP para a região Nordeste, tomando IND_EMP como variável resposta, apresenta um coeficiente de determinação ajustado $R^2 = 0,91$, indicando que 91% da variação dos dados é explicada pelas variáveis independentes IND_DES, IND_ESC, IND_POB e IND_JUV, todas elas selecionadas no processo *forward stepwise*, mostraram-se estatisticamente significantes (p-valor <0,05). Para as unidades federativas da região obteve-se os seguintes indicadores com significância estatística (p-valor <0,05) para explicar o IND_EMP: a) Maranhão: IND_DES, IND_POB e IND_JUV; Piauí: IND_DES, IND_POB e IND_JUV; Ceará: IND_ESC e IND_DES; Rio Grande do Norte: IND_DES; Paraíba: IND_DES e IND_POB, Pernambuco: IND_DES e IND_POB; Alagoas: IND_POB e IND_JUV; Sergipe: IND_DES e IND_ESC; Bahia: IND_ESC, IND_DES e IND_JUV. É importante destacar que, com exceção do estado de Alagoas, ao ser excluído o IND_DES do modelo, o IND_ESC passa a ser o fator de maior poder explicativo do IND_EMP, revelando a importância da educação para o emprego formal, tanto nos estados quanto na região Nordeste como um todo.

Dadas as variáveis sob estudo e tendo os resultados da análise de regressão tomando o IND_EMP como variável resposta no modelo, teve-se a curiosidade de investigar como seria

o comportamento dessas variáveis quando analisado o modelo com o IND_POB como variável resposta, pois em todo momento, nos resultados do modelo até agora utilizado, obteve-se o IND_DES e IND_ESC sempre evidentes, como de maior poder explicativo para o IND_EMP, com exceção do estado de Alagoas, em que o IND_POB se apresentou com maior poder explicativo para o IND_EMP, ressaltando ainda que, na maioria dos casos, o IND_POB se apresentou como estatisticamente significativo (p-valor<0,05). Na opção *forward stepwise*, o IND_POB se prestou na maioria dos casos como estatisticamente significativo no modelo (p-valor< 0,05), para os seguintes casos: região Nordeste, estado do Maranhão, Piauí, Paraíba, Pernambuco e Alagoas, ficando de fora do modelo apenas para os estados da Rio Grande do Norte, Ceará, Bahia, e Sergipe, mostrando assim, que essa variável também é bastante importante na explicação do emprego formal. Portanto a seguir, será analisada a relação entre o IND_POB e as demais variáveis sob estudo.

3.3. RELAÇÃO ENTRE A “POBREZA DOS CHEFES DE FAMÍLIA DO MUNICÍPIO (IND_POB)” E OUTRAS VARIÁVEIS SOCIOECONÔMICAS.

Avaliando as possíveis relações de dependência entre as variáveis do modelo de regressão linear múltipla ponderada (MRLMP) tendo a “pobreza dos chefes de família do município (IND_POB)”, definida como variável resposta, em função das demais variáveis explicativas ou independentes (IND_DES, IND_ESC, IND_EMP e IND_JUV), utilizando o município como unidade de análise, teve os seguintes resultados: (Tabelas de 3.1.1 a 3.1.10).

Tabela 3.1.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Região Nordeste, 2000 (n = 1757).

Variáveis	β padronizado	Erro padrão de β	<i>B</i>	Erro padrão de <i>B</i>	t(213)	p-valor
Intercept			-0,01	0,01	-0,96	<0,3353
IND_ESC	0,90	0,02	0,74	0,01	54,67	<0,0001
IND_JUV	0,06	0,01	0,09	0,02	5,96	<0,0001
IND_DES	0,13	0,02	0,22	0,04	5,80	<0,0001
IND_EMP	-0,10	0,02	-0,16	0,03	-5,32	<0,0001

R= 0,97 R²= 0,94 Ajustado R²= 0,94

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para a Região Nordeste (Tabela 3.1.1) o MRLMP apresenta um coeficiente de correlação múltipla entre as variáveis de R= 0,97 e um coeficiente de determinação ajustado

de $R^2 = 0,94$, indicando que 94% da variação dos dados foi explicada pelas variáveis independentes IND_DES, IND_ESC, IND_EMP e IND_JUV (p -valor $< 0,01$). Todas elas, selecionadas no processo de escolha *forward stepwise* mostraram-se estatisticamente significante (p -valor $< 0,01$), porém, o IND_ESC sozinho, apresentou uma contribuição de 94% na variação total, ou seja, apenas ele traduz a representatividade do modelo.

Tabela 3.2.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Maranhão, 2000 (n = 217 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,05	0,02	2,13	$< 0,0346$
IND_EMP	-0,53	0,08	-0,96	0,14	-6,70	$< 0,0001$
IND_DES	0,96	0,09	2,40	0,23	10,21	$< 0,0001$
IND_ESC	0,46	0,07	0,34	0,05	6,80	$< 0,0001$
IND_JUV	0,06	0,04	0,11	0,07	1,58	$< 0,1150$

R= 0,96 R²= 0,91 Ajustado R²= 0,91

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado do Maranhão (Tabelas 3.2.1) três variáveis apresentaram resultados estatisticamente significantes (p -valor $< 0,01$): IND_EMP, IND_DES e IND_ESC. O modelo apresentou o coeficiente de correlação múltipla de $R = 0,96$ e o de determinação ajustado de $R^2 = 0,91$. No procedimento de escolha *forward stepwise*, os três indicadores permaneceram significantes (p -valor $< 0,01$), onde apresentaram uma contribuição no ganho do R^2 na seguinte ordem: IND_ESC com 0,87, IND_DES com 0,03 e o IND_JUV com 0,02 mostrando, que, para os municípios do Estado do Maranhão deverá haver uma maior preocupação na busca de melhoria na educação, de forma a influenciar, positivamente, a pobreza dos chefes de família no município.

Tabela 3.3.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Piauí, 2000 (n = 220 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,04	0,02	2,18	$< 0,0301$
IND_EMP	-0,51	0,07	-0,93	0,13	-7,19	$< 0,0001$
IND_DES	1,12	0,10	2,53	0,23	10,96	$< 0,0001$
IND_ESC	0,29	0,07	0,23	0,06	4,02	$< 0,0001$
IND_JUV	0,10	0,02	0,19	0,04	4,85	$< 0,0001$

R= 0,97 R²= 0,95 Ajustado R²= 0,95

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado Piauí (Tabela 3.3.1) teve-se que 95 % da variação dos dados foi explicada (p -valor < 0,01) pelos quatro indicadores - IND_EMP, IND_DES, IND_ESC, IND_JUV apresentando também um bom coeficiente de correlação múltipla de $R = 0,97$. Na opção *forward stepwise*, todos os indicadores se apresentaram estatisticamente significantes (p -valor < 0,05), estando o IND_DES com maior contribuição (0,92) na explicação da pobreza.

Tabela 3.4.1
Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Ceará, 2000 (n = 181 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			-0,07	0,02	-3,04	<0,0027
IND_EMP	-0,01	0,04	-0,03	0,08	-0,37	<0,7112
IND_DES	-0,02	0,04	-0,03	0,07	-0,40	<0,6931
IND_ESC	0,96	0,04	0,84	0,04	21,40	<0,0001
IND_JUV	0,07	0,02	0,16	0,05	3,37	<0,0009

$R = 0,99$ $R^2 = 0,97$ Ajustado $R^2 = 0,97$

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado do Ceará (Tabela 3.4.1) apenas o IND_ESC e o IND_JUV contribuíram significativamente para explicar o IND_POB (p -valor < 0,01). O coeficiente de correlação múltipla das variáveis envolvidas foi de $R = 0,99$ e o de determinação ajustado de $R^2 = 0,97$. No processo de escolha *forward stepwise*, os dois indicadores permaneceram no modelo, mostrando, todavia, que o IND_ESC apresentou sozinho uma contribuição de 97% na variação total dos dados, correspondente ao valor do R^2 .

Tabela 3.5.1
Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Rio Grande do Norte, 2000 (n = 164 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			-0,04	0,02	-1,77	<0,0793
IND_EMP	0,00	0,04	0,00	0,07	-0,06	<0,9557
IND_DES	-0,05	0,06	-0,08	0,09	-0,89	<0,3767
IND_ESC	1,03	0,05	0,98	0,05	19,61	<0,0001
IND_JUV	0,01	0,02	0,01	0,05	0,30	<0,7640

$R = 0,98$ $R^2 = 0,97$ Ajustado $R^2 = 0,97$

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado do Rio Grande do Norte (Tabela 3.5.1), apenas o IND_ESC foi estatisticamente significativa (p-valor <0,01) para explicar o IND_POB. No processo de seleção *forward stepwise*, o IND_ESC e o IND_DES aparecem no modelo, porém apenas o IND_ESC foi significativo (p-valor <0,01) e explicou, sozinho, 97% na variação total dos dados, mostrado pelo R^2 .

Tabela 3.6.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado da Paraíba, 2000 (n = 221 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,02	0,03	0,92	<0,3602
IND_EMP	0,04	0,05	0,06	0,07	0,85	<0,3990
IND_DES	-0,05	0,06	-0,09	0,10	-0,95	<0,3457
IND_ESC	1,00	0,05	0,82	0,04	19,77	<0,0001
IND_JUV	0,00	0,02	0,00	0,05	0,08	<0,9344

R= 0,98 R²= 0,96 Ajustado R²= 0,96

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado da Paraíba (Tabela 3.6.1), tem-se também o IND_ESC como o único indicador estatisticamente significativo (p-valor <0,01) para explicar o IND_POB. O coeficiente de correlação múltipla foi de R= 0,98 e o de determinação ajustado de R²= 0,96. No processo de escolha *forward stepwise*, o IND_ESC permaneceu sozinho como significativo no modelo (p-valor <0,01), apresentando uma contribuição de 96% na variação total dos dados.

Tabela 3.7.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado de Pernambuco, 2000 (n = 178 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,03	0,03	1,04	<0,2991
IND_EMP	-0,25	0,07	-0,33	0,09	-3,79	<0,0002
IND_DES	0,13	0,08	0,18	0,11	1,58	<0,1153
IND_ESC	1,05	0,05	0,83	0,04	19,94	<0,0001
IND_JUV	0,00	0,04	0,00	0,07	0,05	<0,9574

R= 0,97 R²= 0,95 Ajustado R²= 0,94

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado de Pernambuco (Tabela 3.7.1), o IND_EMP e o IND_ESC foram os únicos que contribuíram significativamente (p-valor <0,01) para explicar o IND_POB. O coeficiente de correlação múltipla foi de R= 0,97 e o de determinação ajustado de R²= 0,94,

indicando que 94% da variação dos dados foi explicada por essas variáveis. No processo de escolha *forward stepwise*, o IND_ESC e o IND_EMP permanecem significativos (p-valor <0,01) e os mesmo contribuem para o ganho do R² com 94% e 1% respectivamente.

Tabela 3.8.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado de Alagoas, 2000 (n = 101 municípios).

Variáveis	β padronizado	Erro padrão de β	<i>B</i>	Erro padrão de <i>B</i>	t(213)	p-valor
Intercept			-0,10	0,03	-3,35	<0,0012
IND_EMP	0,17	0,03	0,32	0,05	6,20	<0,0001
IND_DES	-0,04	0,07	-0,08	0,15	-0,58	<0,5642
IND_ESC	0,74	0,10	0,65	0,08	7,80	<0,0001
IND_JUV	0,17	0,05	0,28	0,08	3,47	<0,0008

R= 0,98 R²= 0,97 Ajustado R²= 0,97

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado de Alagoas (Tabela 3.8.1), observa-se que três indicadores, IND_EMP, IND_ESC e o IND_JUV, foram estatisticamente significantes (p-valor <0,01) para explicar o IND_POB. O coeficiente de correlação múltipla foi de R= 0,98 e o de determinação ajustado de R²= 0,97. No processo de escolha *forward stepwise*, os três indicadores permanecem significativos no modelo, porém, o IND_ESC apresenta sozinho uma contribuição de 95% na variação total dos dados.

Tabela 3.9.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Sergipe, 2000 (n = 72 municípios).

Variáveis	β padronizado	Erro padrão de β	<i>B</i>	Erro padrão de <i>B</i>	t(213)	p-valor
Intercept			-0,08	0,05	-1,59	<0,1159
IND_EMP	-0,11	0,16	-0,14	0,21	-0,69	<0,4907
IND_DES	0,03	0,14	0,04	0,19	0,21	<0,8316
IND_ESC	0,94	0,07	0,75	0,06	12,74	<0,0001
IND_JUV	0,13	0,05	0,24	0,09	2,56	<0,0126

R= 0,98 R²= 0,95 Ajustado R²= 0,95

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Quanto ao Estado de Sergipe (Tabela 3.9.1), o IND_ESC e o IND_JUV foram os únicos estatisticamente significativos (p-valor <0,05) no modelo. O coeficiente de correlação múltipla foi de R= 0,98 e o de determinação ajustado de R²= 0,95. No procedimento *forward*

stepwise, os dois permanecem significativos no modelo (p-valor <0,05), estando apenas o IND_ESC contribuindo 95% na variação total dos dados.

Tabela 3.10.1

Sumário de regressão linear múltipla ponderada para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado da Bahia, 2000 (415 municípios).

Variáveis	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor
Intercept			0,03	0,02	1,74	<0,0829
IND_EMP	0,02	0,03	0,03	0,05	0,62	<0,5373
IND_DES	-0,05	0,04	-0,09	0,09	-1,10	<0,2709
IND_ESC	1,01	0,05	0,81	0,04	20,83	<0,0001
IND_JUV	-0,02	0,02	-0,03	0,04	-0,71	<0,4791

R= 0,97 R²= 0,94 Ajustado R²= 0,94

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Para o Estado da Bahia (Tabela 3.10.1), tem-se que IND_ESC foi o único indicador estatisticamente significativo (p-valor <0,01) para explicar o IND_POB. O coeficiente de correlação múltipla foi de R= 0,97 e o de determinação ajustado de R²= 0,94. No processo de escolha *forward stepwise*, apenas o IND_ESC foi significativo, apesar do IND_DES aparecer no modelo sem está, contudo, significativo (p-valor <0,05). O IND_ESC contribui sozinho no valor do R² com 0,94.

Portanto, cabe enfatizar que, para os dez modelos envolvidos no estudo (Região Nordeste e suas respectivas Unidades Federativas), tomando o IND_POB como variável resposta, em todos os casos, com exceção apenas para o Estado do Piauí, utilizando a opção *forward stepwise*, apresentaram o IND_ESC como primeiro índice estatisticamente significativo (p-valor <0,05) com a maior contribuição para o R². Com isso, fica evidenciado como são necessários as políticas voltadas especialmente para investimentos que busquem um melhor nível de escolaridade da população, afetando melhoras na pobreza e um aumento do emprego formal (que é bastante influenciado pela educação), de forma que, assim, a desigualdade de renda seja diminuída em cada município da região Nordeste.

Contudo, cabe destacar que, nesta seção as análises dos MRLMP que tem como variável dependente, num primeiro momento, o IND_EMP e, em seguida, o IND_POB, e como variáveis independentes o IND_POB, IND_DES, IND_ESC e IND_JUV. Pode-se observar, no geral, que, para a região Nordeste e seus respectivos estados, os modelos foram muito bem representados pelas variáveis explicativas, estando sempre com um bom coeficiente de correlação e de determinação ajustado, indicando que a variação dos dados

foram bem representados pelas variáveis envolvidas em cada modelo. O relacionamento entre a variável dependente e as variáveis independentes, determinadas no processo de escolha *forward stepwise* (Anexo), evidencia sempre a grande importância de um melhor nível de escolaridade para a população Nordestina, sendo uma grande saída para um crescimento na quantidade de ocupações formais da região e não esquecendo buscar uma diminuição da desigualdade de renda, que é o problema que permeia todo o país e em especial a região Nordeste. Quando observado a pobreza dos chefes do domicílio como variável explicativa no modelo, que é também bastante problemática na região, a variável IND_ESC aparece em todos os casos com maior poder explicativo, ou seja, a busca de melhorias na educação irá remeter a melhorias na pobreza e conseqüentemente, uma diminuição da desigualdade de renda, afetando ainda uma melhoria no emprego formal, que é tão necessária para a população, devido aos seus benefícios de bem-estar para os trabalhadores.

A constatação de que um aumento no nível de escolaridade acarreta melhoras nos níveis de pobreza, emprego, desigualdade e ainda na concentração de jovens (longo prazo), vem reforçar a afirmação de Buarque (1991, p.53), que comenta: “é pelo problema da educação do Brasil que se iniciam as prioridades de sua modernização”.

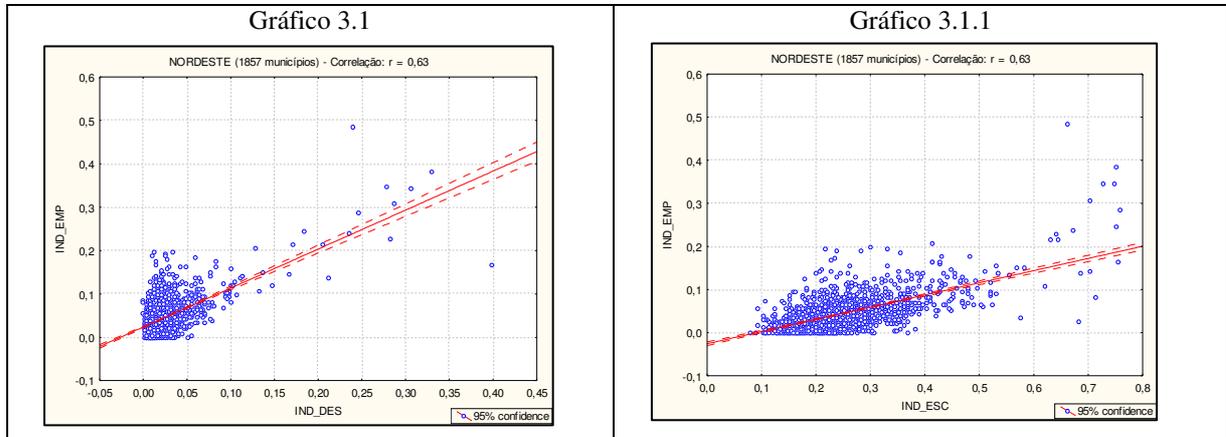
3.4. DISTRIBUIÇÃO DO INDICADOR “TAXA DE EMPREGO FORMAL SOBRE A PIA” COM A VARIÁVEL DE MAIOR PODER EXPLICATIVO.

De acordo com a associação da variável IND_EMP com as demais variáveis estudadas, tem-se que os mais fortes níveis de associação para com esta variável, conforme mostrados anteriormente foram com o IND_DES e com o IND_ESC, para praticamente todos os estados e, inclusive, para a Região Nordeste como um todo. Com isso, se fez necessário analisar graficamente como as devidas observações estão distribuídas no contexto da Região Nordeste como um todo e para cada Unidade Federativa.

▪ **Análise bivariada do IND_EMP relacionada com a Variável de Maior Poder Explicativo dentro de cada Unidade Federativa de interesse.**

A análise bivariada do emprego formal com a variável desigualdade, no caso da Região Nordeste (Gráfico 3.1) como um todo, mostra uma concentração bastante elevada em valores baixos. Mesmo assim, à medida que se reduz a desigualdade aumenta o índice de emprego formal. Já quando analisada a variável emprego com o índice de escolaridade

(Gráfico 3.2), há uma menor concentração nos valores baixos. Na medida em que cresce o índice de escolaridade, também tende a crescer o índice de emprego formal.



Gráficos 3.1 a 3.1.1: Correlação bivariada entre: IND_EMP e a Variável de Maior Poder Explicativo (VMPE) para a região Nordeste (1757 municípios).

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Dando continuidade à análise para os respectivos estados da região Nordeste (Gráficos 3.2 a 3.10) e olhando primeiramente para os casos entre o IND_EMP e o IND_DES, que foi realizada para quase todos os estados, devido ao maior poder explicativo da variável IND_DES para o IND_EMP, ficou claro que, em todos os casos, a concentração em valores baixos é altíssima estando na seguinte ordem para os eixos das variáveis (IND_DES; IND_EMP): SE (0,00 a 0,05; 0,00 a 0,15), PB (0,00 a 0,07; 0,00 a 0,15), PI (0,00 a 0,10; 0,00 a 0,12), RN (0,00 a 0,10; 0,00 a 0,20), MA (0,00 a 0,11; 0,00 a 0,12) e PE (0,00 a 0,15; 0,00 a 0,17). Ou seja, é muito grande a concentração dos dados em valores baixos, tanto para a desigualdade de renda quanto para o nível de emprego formal, que apesar de, em alguns casos, ter apresentado bem mais dispersões que a desigualdade (SE, PB, RN) a concentração é enorme para os dois lados, na qual a maioria dos municípios estão enquadrados em valores de, no máximo, 0,20. Portanto, em todos os estados do Nordeste, há um caso de calamidade, sendo elevada a exclusão social, formando uma mancha bastante homogênea de exclusão dentro no Nordeste.

Para os Estados do Ceará e Bahia (Gráficos 3.4 e 3.10), em que foi mostrada a associação das variáveis IND_EMP com o IND_ESC, tem-se um maior espalhamento desses dados, devido à variável IND_ESC apresentar valores maiores. Observando que, também, na medida em que aumenta o nível de escolaridade, o nível de emprego formal também melhora.

Gráfico 3.2

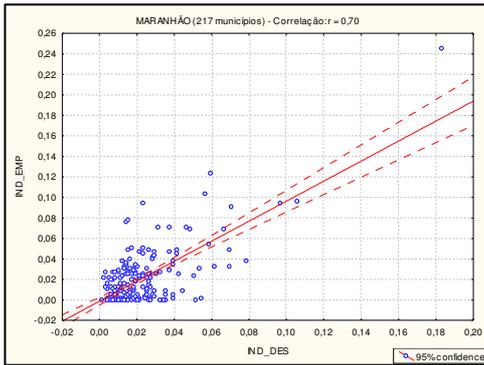


Gráfico 3.3

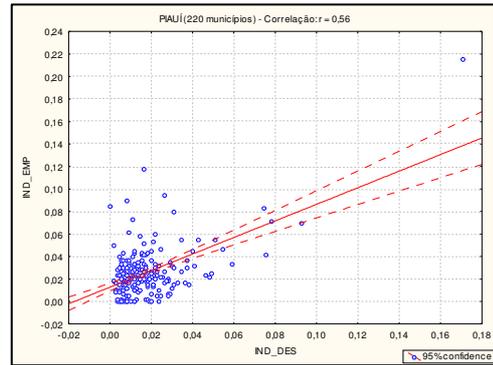


Gráfico 3.4

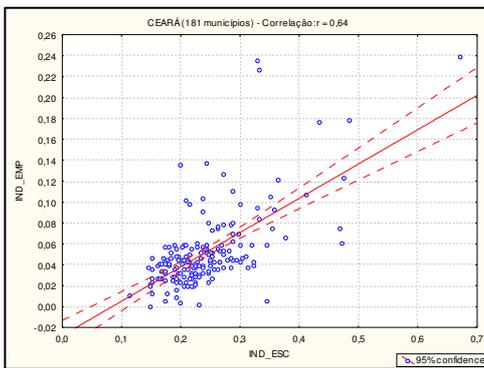


Gráfico 3.5

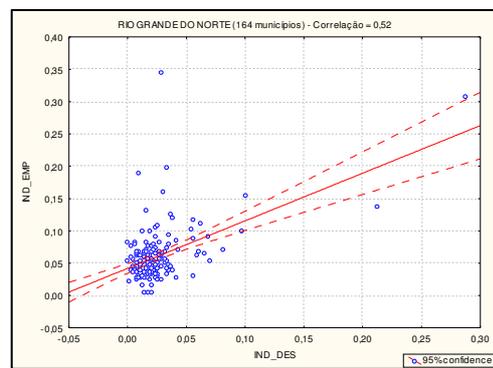


Gráfico 3.6

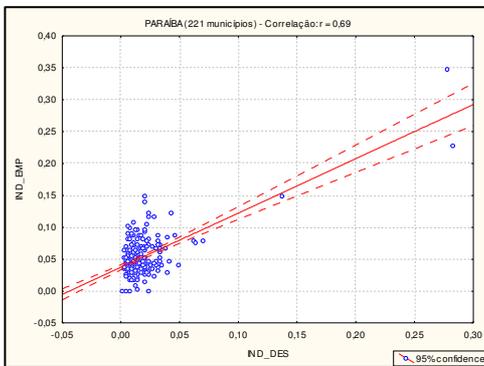
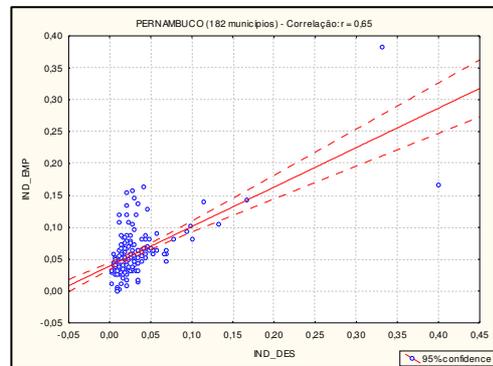


Gráfico 3.7



Continua...

Continuação...

Gráfico 3.8

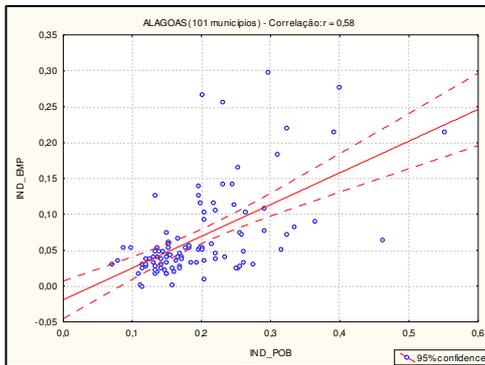


Gráfico 3.9

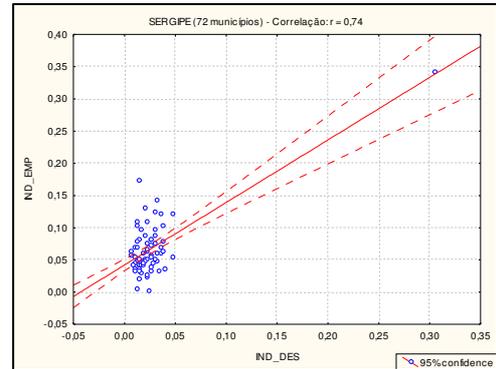
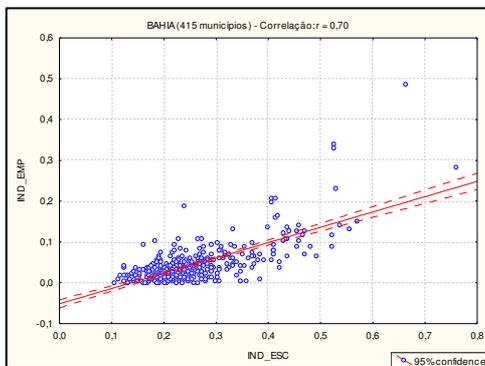


Gráfico 3.10



Gráficos 3.2 a 3.10: Correlação bivariada entre: IND_EMP e VMPE para os nove Estados que compõem a região Nordeste.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

CAPÍTULO 4: ASSOCIAÇÃO DO EMPREGO FORMAL COM A INCLUSÃO SOCIAL NO CONTEXTO DAS UNIDADES FEDERATIVAS NORDESTINAS.

Inicialmente, tem-se que o Índice de Exclusão Social (IES), obtido do AESB (POCHMANN; AMORIM, 2003) é composto pela combinação de sete indicadores/ índices (emprego formal – IND_EMP, pobreza – IND_POB, desigualdade – IND_DES, escolaridade – IND_ESC, alfabetização – IND_ALF, juventude – IND_JUV e violência – IND_VIO) que recebem pesos específicos associados a cada um deles, tendo sido transformados em valores que variam no intervalo de 0 a 1, significando maior inclusão, quanto mais próximo da unidade. A partir do IES foi criado um novo índice de inclusão social (IIS), sem o indicador IND_EMP. Isso foi feito para permitir investigar a associação do emprego formal com a inclusão/ exclusão social, evitando a ocorrência de associações espúrias. Desse modo, o IND_EMP, que entrava na composição do IES foi dele retirado e realizada uma nova combinação linear que determinou o Índice de Inclusão Social (IIS). O procedimento consiste em distribuir, proporcionalmente, o peso referente ao IND_EMP (que corresponde a 17%), para os demais indicadores que irão compor o novo índice (IIS), como pode ser visto no Quadro 4.1. Da mesma forma que os demais índices, o IIS assume valores no intervalo de 0 a 1, representado maior inclusão quanto mais próximo da unidade.

Quadro 4.1
Composição do novo Índice de Inclusão Social (IIS).

Índice de Exclusão Social (IES)	Pesos que compõem o IES	Pesos que compõem o IIS
Pobreza	17 %	20,48 %
Emprego Formal em idade ativa	17 %	Foi retirado
Desigualdade de renda pelos chefes de família	17 %	20,48 %
Anos de estudos do chefe da família	11,3 %	13,61 %
Alfabetização das pessoas maiores de 5 anos	5,7 %	6,87 %
Concentração de Jovens	17 %	20,48 %
Violência (Homicídios por 100 mil habitantes)	15 %	18,07 %
Total	100 %	100 %

Fonte: Pochmann e Amorim (2003) (Dados censitários).

4.1. AS UNIDADES FEDERATIVAS (UF'S) NO CONTEXTO DAS GRANDES REGIÕES

Para investigar a associação entre os índices de emprego formal (IND_EMP) e inclusão social (IIS) no contexto de suas grandes regiões, com ênfase na região Nordeste

brasileira, utilizou-se o método da análise de *cluster* (ou de agrupamento), que tem por finalidade agrupar as unidades investigadas segundo algumas características consideradas, com base em similaridades ou distâncias. Conforme descrito na metodologia, o critério usado foi o de ligação simples (*single linkage*) e a distância Euclidiana (DE). Replicou-se esta metodologia ao conjunto das 9 (nove) UF's da região Nordeste para identificar o contexto particular desse universo geográfico. Na Figura 4.1, pelo corte na DE = 0,10 os resultados do Diagrama em Árvore (*Tree Diagram*) mostram a formação de agrupamentos, nos quais o emprego formal e a inclusão social variam na mesma direção, bastante consistentes, nos quais as UF's das regiões Sudeste (SE'), Sul (SU), Centro-Oeste (CO), formaram um grande grupo, exceto pelo isolamento do Distrito Federal (DF), enquanto as UF's das regiões Norte (NO) e Nordeste (NE) um outro grande agrupamento, exceto pelas UF's de Piauí (PI) e Maranhão (MA), que formaram um grupo isolado. Conforme mencionado no Quadro 4.2.

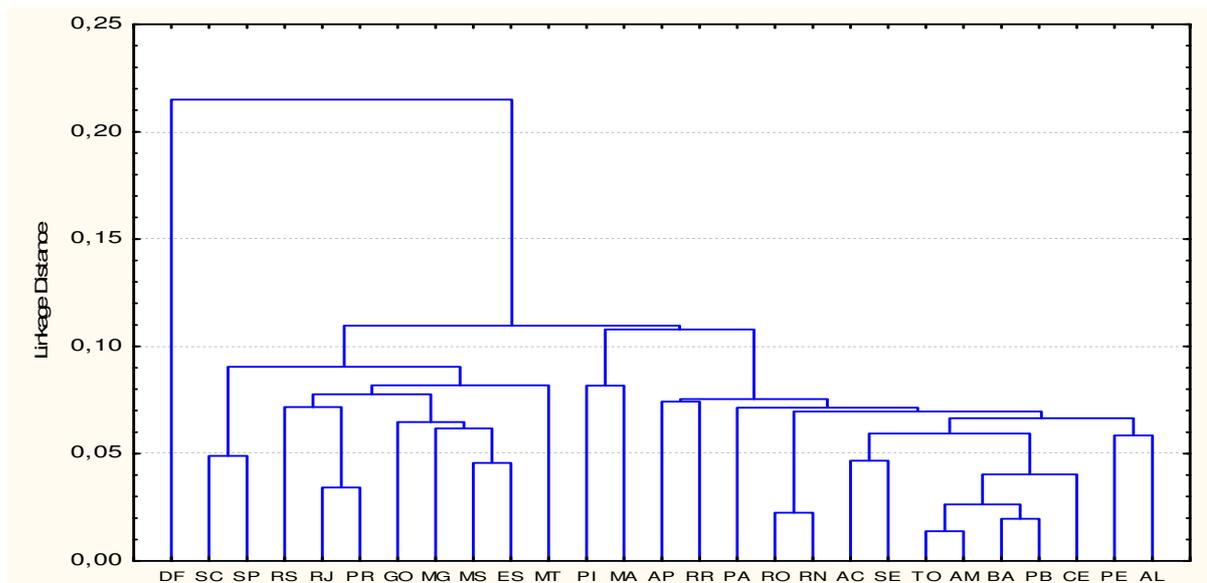


Figura 4.1 – Diagrama em (*Tree Diagram*) segundo as variáveis emprego formal (IND_EMP) e IIS, Brasil (n=27), 2000.

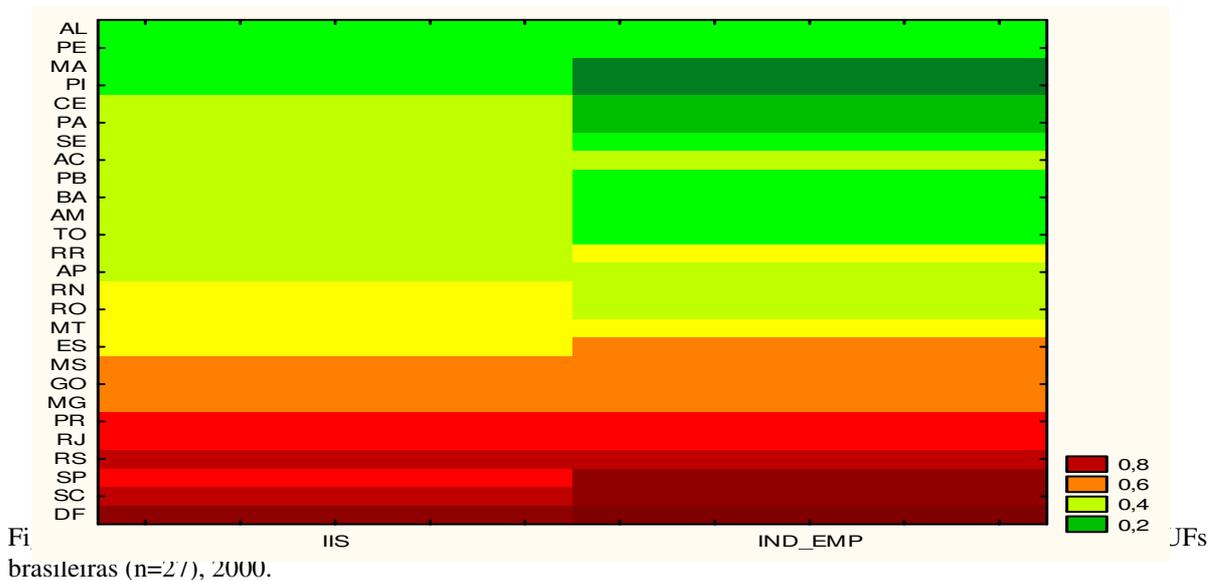
Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Quadro 4.2

Distribuição dos grupos formados pela Análise de cluster, segundo as variáveis IND_EMP e IIS para as UF's brasileiras (n=27), 2000.

<p>Grupo 1 – UFs das Regiões SU, SE' e CO (n=10) – com melhores valores do IND_EMP e IIS: SC, SP, RS, RJ, PR, GO, MG, MS, ES, MT Exceção – DF (Região CO) - isolado</p>	<p>Grupo 3 – UFs das Regiões NO e NE (n=14) – com valores mais baixos do IND_EMP e IIS: AP, RR, PA, RO, RN, AC, SE, TO, AM, BA, PB, CE, PE, AL</p>
<p>Grupo 2 – UFs das Regiões NE (n=2) – com valores mais baixos do IND_EMP e IIS: PI e MA</p>	

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).



Fi
brasileiras (n=27), 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

O comportamento conjunto do IIS e o IND_EMP no contexto das 27 Unidades Federativas, pode ser observado pela Figura 4.2, a seguir, onde os dados apresentaram ordenados crescentemente pelo IIS, ficou claro que ambas as variáveis têm a mesma tendência, variando num mesmo sentido. Nota-se que as UF's das regiões Nordeste e Norte apresentam valores baixos para ambos indicadores, enquanto o inverso pode ser dito para as UF's das regiões Sul, Sudeste e Centro-Oeste.

Em outras palavras, fica reafirmado que o acesso ao emprego formal, (IND_EMP) em todas as UF's brasileiras, apresenta-se tanto maior quanto maior for o nível de inclusão social (IIS) da mesma. Fica claro que, melhores valores do IIS traduzem menor pobreza, menor desigualdade, menor contingente de jovens e melhor escolaridade.

4.2. O CASO DAS UF'S DO NORDESTE BRASILEIRO.

Na análise das variáveis estudadas para a região Nordeste brasileira observou-se que alguns estados apresentaram situações muito semelhantes. Para estabelecer estas semelhanças, procedeu-se uma análise de agrupamento, segundo as duas variáveis de interesse (IND_EMP e IIS), e tomando-se as Unidades Federativas (UF's) da região Nordeste como unidade de análise, onde foi possível identificar, graficamente (Figuras 4.3 e 4.4), grupos de Estados que mais se assemelham quanto a esses indicadores.

Pela análise do Diagrama de árvore (Figura 4.3) e adotando um corte na Distância Euclidiana no ponto 0,085, obtém-se a definição de dois agrupamentos, envolvendo as seguintes UF's: PI e MA, com mais baixos valores dos indicadores e BA, PB, CE, SE e AL, com valores relativamente maiores, ficando isolado o RN. Esses resultados mostram os problemas de inclusão social e emprego, apresentados pelas Unidades Federativas do PI e MA, com os piores indicadores da região, enquanto a UF do RN destaca-se por melhores níveis de ambos os indicadores no contexto regional. Conforme mencionado no Quadro 4.3. Esse comportamento pode ser também visualizado pela Figura 4.4, com o adicional de revelar se os indicadores envolvidos (IND_EMP e IIS) se agrupam por maior grau de semelhança ou de contraste. A apresentação dos diferenciais sócio-espaciais das desigualdades apontou a necessidade de adoção de medidas inclusivas que venham a reduzir o desemprego e melhorar a qualidade de vida das populações.

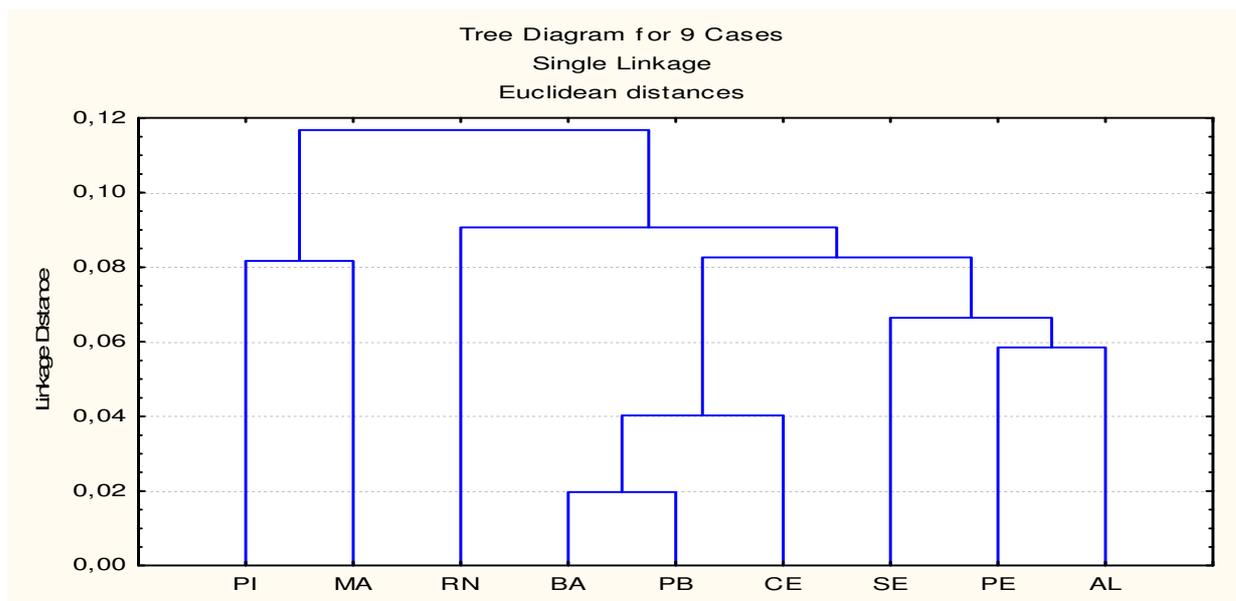


Figura 4.3 - Diagrama em Árvore (Tree Diagram), segundo as variáveis IND_EMP e IIS para as UF's da região Nordeste do Brasil (n=9), 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Quadro 4.3

Distribuição dos grupos formados pela Análise de cluster, segundo as variáveis IND_EMP e IIS para as UF's da região Nordeste do Brasil (n=9), 2000

Grupo 1 – UF isolada – com melhores valores do IND_EMP e IIS: RN	Grupo 3 – UF's com valores intermediários do IND_EMP e IIS: BA, PB, CE, SE, PE, AL
Grupo 2 – UF's com valores mais baixos do IND_EMP e IIS – maior contraste entre os índices - (n=2): PI e MA	

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

A Figura 4.4, dá ênfase aos resultados alcançados, onde mostra o Rio Grande do Norte com o melhor índice, tanto para o IND_EMP, como para o IIS, apresentando uma cor vermelha bastante intensa para os dois indicadores. O Piauí e o Maranhão apresentaram piores índices na associação dos dois indicadores, para ambos os estados, estando inclusive o IND_EMP com a cor verde (que é a pior classificação) nos dois estados e o IIS também com as piores cores apresentadas neste indicador. Os demais estados apresentaram situações intermediárias ao associar os indicadores (IND_EMP e o IIS).

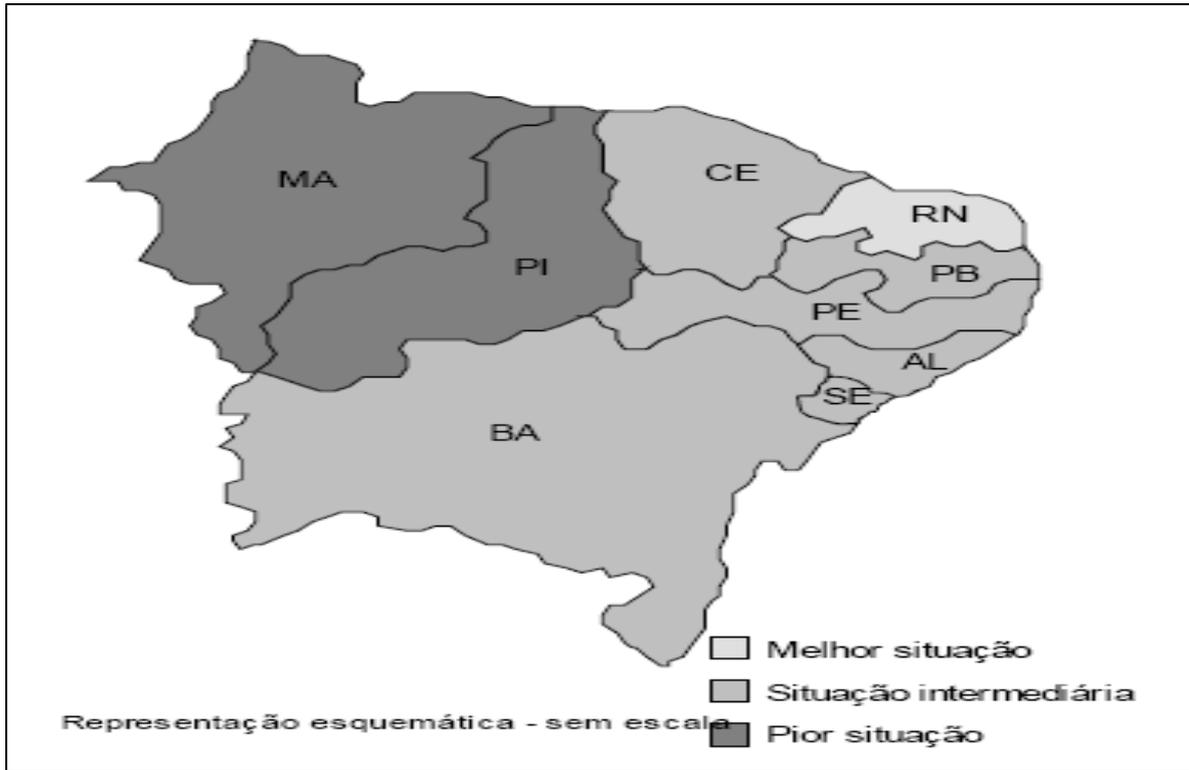


Figura 4.4 – Resultado gráfico da junção (Two-Way joining Results) das variáveis IND_EMP e IIS para as UFs nordestinas (n= 9), 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Pode-se observar ainda, pelo método *Two-Way Joining* (Figura 4.4), que ambos os indicadores traduziram um baixo índice de inclusão social na região, os mesmos variaram entre 0,0 e 0,4. Nos casos em que o IND_EMP foi melhor, o IIS se apresentou com bons índices. Logo, esses dois indicadores ficaram inter-ligados, na medida em que um melhora o outro também se apresenta melhor.

A afirmação de Araújo (1995, p.192) vem a reforçar os resultados encontrados quando comenta: “A realidade social divulga um Nordeste heterogêneo e complexo, múltiplo, diferenciado, apesar de alguns traços comuns marcantes”.



Mapa 4.1: *Clusters* dos Estados do Nordeste agrupados segundo as condições de emprego e do Novo índice de Inclusão Social, 2000.

Fonte: Barbosa e Ribeiro (2006).

De acordo com Barbosa e Ribeiro (2006), os resultados evidenciam que as condições de vida para a população residente nos estados do Nordeste são bem mais precárias, vis-a-vis os residentes do País como um todo, fato que retrata a desigualdade regional existente no Brasil. Entre os estados do Nordeste, o Rio Grande do Norte se destaca como aquele detentor das melhores condições de vida, notadamente no que diz respeito aos serviços básicos e à infra-estrutura existente na moradia. Em contraposição, os Estados do Maranhão e do Piauí apresentaram as piores situações, na maioria das variáveis analisadas.

4.3. ASSOCIAÇÕES ENTRE INDICADORES

Nessa análise dos indicadores que influenciam e são influenciados pelo emprego formal, torna-se indispensável avaliar a magnitude das correlações entre o novo índice de inclusão social (IIS), associados aos indicadores “Taxa de emprego formal sobre a PIA” e o “Percentual de pessoas com mais de 50% da renda provenientes de transferência

governamental”, os quais, como se pode ver pela matriz de correlação (Quadro 4.4), apresentaram significância estatística para todas as associações, ao nível de 5%.

Quadro 4.4

Correlações referente as variáveis: IND_EMP, IIS e PP_50RTG para a Região Nordeste (1757 municípios).

Variáveis	IND_EMP	IIS	PP_50RTG
IND_EMP	1	0,58	-0,15
IIS	0,58	1	-0,15
PP_50RTG	-0,15	-0,15	1

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

A mais forte correlação foi encontrada entre o IND_EMP e o IIS (0,58). Quando se observou o IIS com o PP_50RTG, tem-se que a correlação foi bastante baixa e negativa apesar de ser significativa (p-valor <0,05), ou seja, neste caso, há uma associação inversa onde na medida que aumentam as transferências de renda, o IIS diminui. Mais adiante, foi mostrada graficamente, pela análise bivariada, como se comporta a distribuição desses dados, na qual em alguns casos, esse crescimento inverso foi bem razoável, sendo quase imperceptível, como foi o caso, por exemplo, para a região Nordeste como um todo, que ao aumentar o PP_50RTG não mudou muito o comportamento da variável IIS.

Observando as correlações da variável IND_EMP com as variáveis IIS e PP_50RTG, bem como o grau de associação do IIS com PP_50RTG sob estudo, para cada Unidade Federativa da Região Nordeste (Quadro 4.5), tem-se que, para o primeiro tipo de correlação (IND_EMP com o IIS) em todas as Unidades Federativas mostraram bem correlacionadas e estatisticamente significantes (p-valor <0,05), com valores sempre positivos, ou seja, na medida em que uma variável cresce a outra também aumenta. No segundo tipo de correlação, para o IND_EMP com o PP_50RTG, evidencia-se que, em todos os casos, os valores foram negativos. Logo, na medida em que uma variável cresce, a outra diminui. As associações na maioria dos estados (CE, RN, PB, PE, AL e BA) foram estatisticamente significantes (p-valor <0,05), exceto para três estados (MA, PI e SE) nos quais os valores não foram significantes. Por último, na correlação IIS com o PP_50RTG, todas as associações se mostraram também negativas e estatisticamente significantes (p-valor <0,05), exceto para o estado de Sergipe, que apresentou um valor negativo e não significativo.

Quadro 4.5

Correlações para as nove Unidades Federativas da região Nordeste referente as variáveis IND_EMP, IIS e PP_50RTG.

Variáveis	MA (n=217)	PI (n=220)	CE (n=181)	RN (n=164)	PB (n=221)	PE (n=178)	AL (n=101)	SE (n=72)	BA (n=415)
IND_EMP com IIS	0,64	0,48	0,55	0,53	0,58	0,61	0,50	0,64	0,67
IND_EMP com PP_50RTG	-0,07	-0,10	-0,45	-0,39	-0,31	-0,37	-0,28	-0,07	-0,37
IIS com PP_50RTG	-0,19	-0,17	-0,38	-0,33	-0,34	-0,47	-0,26	-0,06	-0,21

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Contudo, observou-se que as correlações do IND_EMP com o IIS, além de terem sido mais fortes, foram positivas. Logo, o grau de importância do emprego formal na inclusão foi bastante evidenciado, ou seja, apesar dos níveis de ocupação formal para a região terem sido baixíssimos, há um efeito bastante positivo para inclusão social. Já a variável transferência de renda, foi significativa quando associado à inclusão, mas se apresentou negativa. Logo, não melhorou o grau de inclusão social com o aumento dessa variável, fazendo, em muitos casos, apenas matar a fome dessas pessoas que recebem esse tipo de benefício, não remetendo, dessa forma, a um maior grau de inclusão social. Como citam Resende e Oliveira (2008, p.1), “ é provável que o aumento de consumo destas famílias pobres eleva o nível de bem estar, representando um ‘alívio’ imediato sobre a pobreza “.

Para apresentar uma análise ampla desses três indicadores estudados (IND_EMP, PP_50RTG e IIS), será mostrada a aplicação do modelo de regressão linear múltipla ponderada, que permitirá identificar os indicadores socioeconômicos que são influenciados e influenciam o IIS, tomando o município como unidade de análise e observando esse comportamento, tanto na Região Nordeste como um todo, quanto para cada uma de suas nove Unidades Federativas.

4.4 INVESTIGANDO RELAÇÃO ENTRE VARIÁVEIS ATRAVÉS DA APLICAÇÃO DO MODELO DE REGRESSÃO LINEAR MÚLTIPLA PONDERADA

▪ **Relação entre a variável dependente “Novo Índice de Inclusão Social” e como variáveis independentes “Taxa de emprego formal sobre a PIA” e “Percentual de pessoas com mais de 50% da renda provenientes de transferência governamental”.**

Para avaliar as possíveis relações de dependência entre as variáveis pesquisadas, empregou-se a técnica da Análise de Regressão Linear Múltipla Ponderada, na qual o “novo Índice de Inclusão Social” (IIS), foi definida como variável dependente ou variável resposta, escrita em função das variáveis “Taxa de emprego formal sobre a PIA” (IND_EMP) e “Percentual de pessoas com mais de 50% da renda provenientes de transferência governamental” (PP_50RTG), que são denominadas variáveis independentes ou explicativas sob estudo.

Os coeficientes de correlação e de determinação ajustado, em todos os casos, apresentaram valores acima de 0,70, mostrando que a variação total dos dados foi bem representada por essas variáveis. Na opção *forward stepwise*, IND_EMP foi o maior fator de explicação total dos dados, dado por R^2 . Logo, apesar do nível de emprego formal ser baixíssimo, ela tem um forte poder de explicação do IIS. Ver na Tabela 4.1.

Tabela 4.1

Sumário de regressão linear múltipla ponderada para a variável resposta IIS, ajustado para as variáveis independentes IND_EMP e PP_50RTG, Área Geográfica – Região Nordeste (NE) e Unidades Federativas do NE, 2000.

NORDESTE							R= 0,91 R ² = 0,82 Adjusted R ² = 0,82
	β padronizado	Erro padrão de β	B	Erro padrão de B	t(213)	p-valor	
Intercept			0,43	0,00	96,79	<0,0001	
IND_EMP	0,83	0,01	0,80	0,01	73,12	<0,0001	
PP_50RTG	-0,14	0,01	-0,00	0,00	-12,74	<0,0001	
MARANHÃO							R= 0,93 R ² = 0,86 Adjusted R ² = 0,86
Intercept			0,42	0,01	43,61	<0,0001	
IND_EMP	0,84	0,03	0,92	0,03	29,20	<0,0001	
PP_50RTG	-0,18	0,03	0,00	0,00	-6,21	<0,0001	
PIAUI							R= 0,93 R ² = 0,87 Adjusted R ² = 0,87
Intercept			0,43	0,01	28,86	<0,0001	
IND_EMP	0,82	0,04	0,86	0,04	21,61	<0,0001	
PP_50RTG	-0,14	0,04	0,00	0,00	-3,56	<0,0005	
CEARÁ							R= 0,94 R ² = 0,89 Adjusted R ² = 0,89
Intercept			0,40	0,02	26,25	<0,0001	
IND_EMP	0,86	0,04	0,95	0,04	22,58	<0,0001	
PP_50RTG	-0,11	0,04	0,00	0,00	-2,79	<0,0058	
RIO GRANDE DO NORTE							R= 0,92 R ² = 0,85 Adjusted R ² = 0,85
Intercept			0,46	0,02	23,70	<0,0001	
IND_EMP	0,82	0,04	0,79	0,04	20,74	<0,0001	
PP_50RTG	-0,15	0,04	0,00	0,00	-3,74	<0,0003	
PARAÍBA							R= 0,93 R ² = 0,87 Adjusted R ² = 0,87
Intercept			0,50	0,01	35,52	<0,0001	
IND_EMP	0,74	0,03	0,64	0,03	23,29	<0,0001	
PP_50RTG	-0,26	0,03	0,00	0,00	-8,12	<0,0001	
PERNAMBUCO							R= 0,88 R ² = 0,78 Adjusted R ² = 0,77
Intercept			0,52	0,02	31,12	<0,0001	
IND_EMP	0,72	0,04	0,58	0,03	18,85	<0,0001	
PP_50RTG	-0,30	0,04	-0,01	0,00	-7,86	<0,0001	
ALAGOAS							R= 0,75 R ² = 0,57 Adjusted R ² = 0,56
Intercept			0,37	0,04	9,84	<0,0001	
IND_EMP	0,72	0,07	0,85	0,09	9,90	<0,0001	
PP_50RTG	-0,07	0,07	0,00	0,00	-0,94	<0,3521	
SERGIPE							R= 0,96 R ² = 0,93 Adjusted R ² = 0,92
Intercept			0,41	0,02	24,65	<0,0001	
IND_EMP	0,96	0,03	0,81	0,03	29,28	<0,0001	
PP_50RTG	-0,08	0,03	0,00	0,00	-2,59	0,0118	
BAHIA							R= 0,91 R ² = 0,83 Adjusted R ² = 0,83
Intercept			0,40	0,01	41,35	<0,0001	
IND_EMP	0,89	0,02	0,86	0,02	36,16	<0,0001	
PP_50RTG	-0,03	0,02	0,00	0,00	-1,30	<0,1944	

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

4.5. DISTRIBUIÇÃO DO INDICADOR “NOVO ÍNDICE DE INCLUSÃO SOCIAL” COM AS DEMAIS VARIÁVEIS DE INTERESSE.

Diante da curiosidade de se observar como se comporta graficamente a associação entre o “Novo Índice de Inclusão Social”(IIS) e as demais variáveis de interesse: IND_EMP e PP_50RTG, foi utilizada a técnica de análise bivariada exposta abaixo.

- **Análise bivariada do IIS relacionada com IND_EMP e com o PP_50RTG para cada Unidade Federativa de interesse.**

Observa-se nos gráficos 4.1 a 4.10,1, que há uma grande concentração em valores baixos e, enquanto no primeiro tipo de gráfico se tem um crescimento crescente das variáveis, ou seja, na medida que cresce o IND_EMP, o IIS também aumenta. Para o segundo tipo de gráfico, na medida em que cresce o PP_50RTG, o IIS diminui. Ressaltando ainda que, na primeira situação, o crescimento apresentou bem mais acentuado do que o declínio do segundo caso, mostrando como o IND_EMP afeta muito fortemente na melhora da inclusão social.

Gráfico 4.1

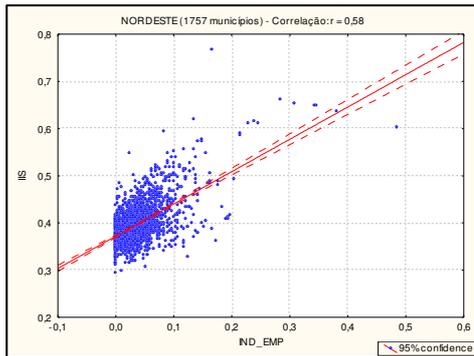


Gráfico 4.1.1

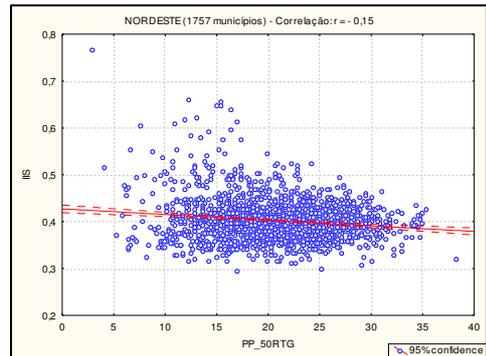


Gráfico 4.2

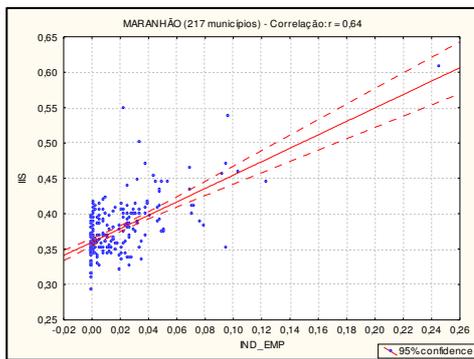


Gráfico 4.2.1

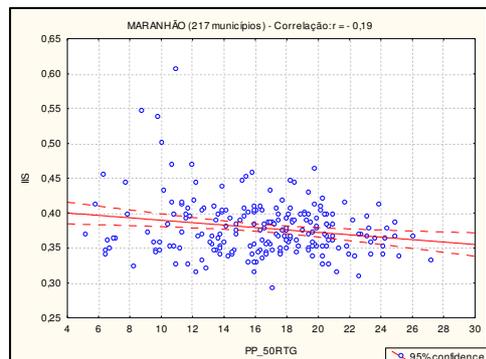


Gráfico 4.3

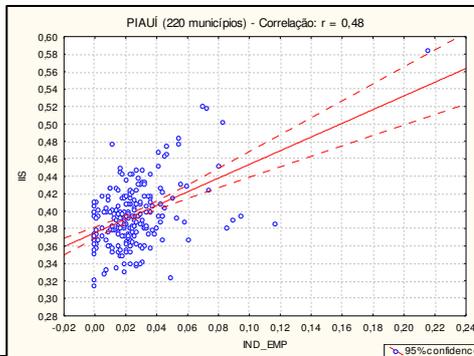
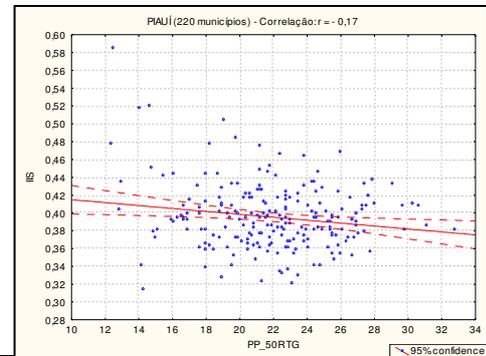


Gráfico 4.3.1



Continua...

Continuação...

Gráfico 4.4

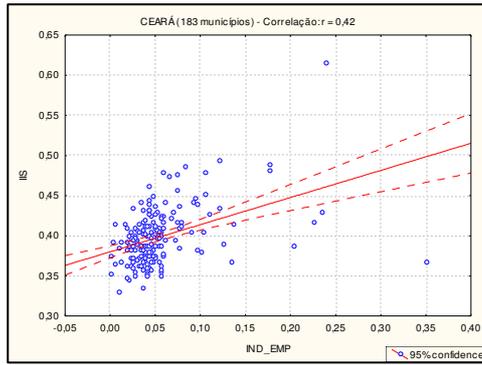


Gráfico 4.4.1

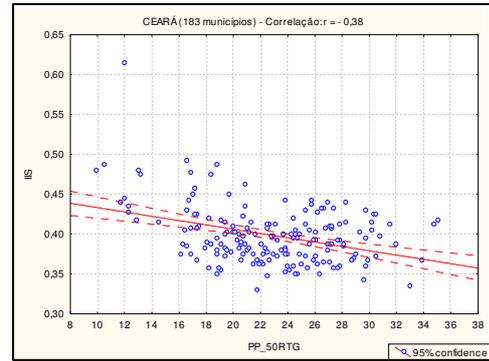


Gráfico 4.5

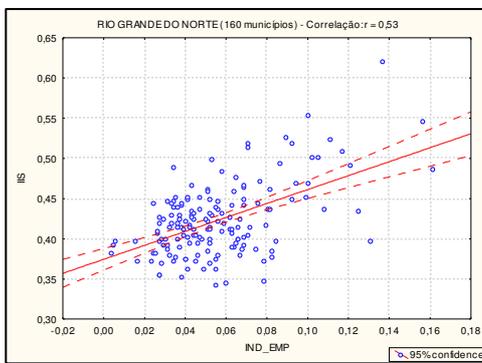


Gráfico 4.5.1

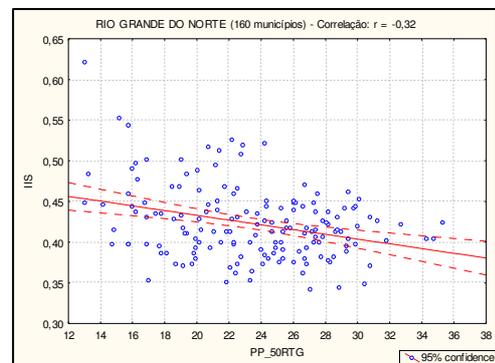


Gráfico 4.6

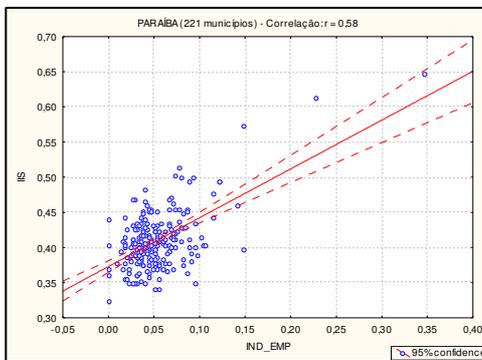
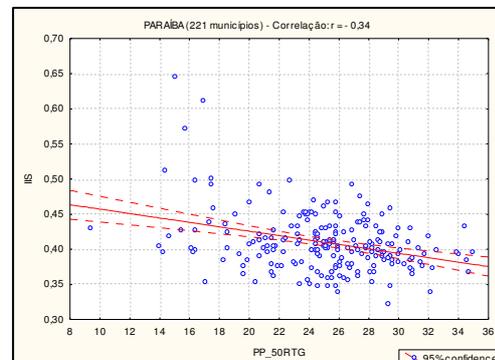
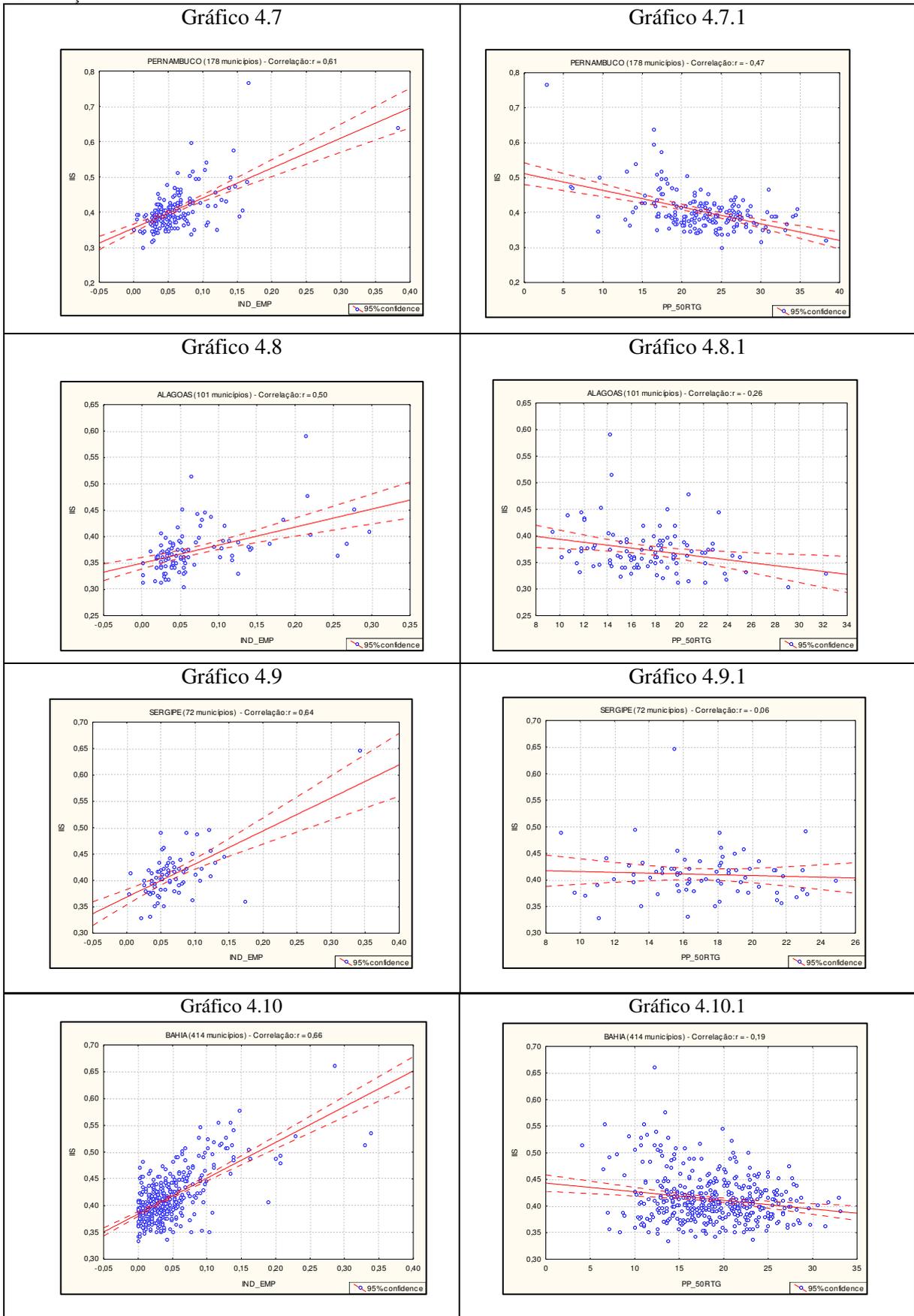


Gráfico 4.6.1



Continua...

Continuação...



Gráficos 4.1 a 4.10.1: Correlação bivariada do IIS com o IND_EMP e com o PP_50RTG para a região Nordeste e para cada uma de suas respectivas Unidades Federativas.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

CAPÍTULO 5 - ANÁLISE EXPLORATÓRIA ESPACIAL DO EMPREGO FORMAL E INCLUSÃO SOCIAL/ EXCLUSÃO SOCIAL NA REGIÃO NORDESTE DO BRASIL.

Com o propósito de identificar as posições relativas dos municípios da Região Nordeste quanto ao nível de inclusão social, realizou-se uma análise exploratória espacial para a região Nordeste como um todo (1790 municípios¹), utilizando-se o software GeoDa (ANSELIN, 2004). Para melhor visualizar a distribuição espacial da variável IND_EMP, bem como seu comportamento quando analisada simultaneamente com cada uma das variáveis IND_POB, IND_DES, IND_ESC, IND_JUV, novo Índice de Inclusão Social (IIS) e PP_50RTG. Os dados utilizados são do censo 2000.

O índice de Moran é uma estatística de autocorrelação espacial que torna possível testar a existência de valores de uma variável com outros valores parecidos que estejam próximos ou vizinhos (ASSUNÇÃO apud RODRIGUES, 2007, p. 25). Relembrando o descrito na metodologia, a autocorrelação poderá ser positiva, se os valores semelhantes se aproximarem no espaço, ou negativa, se tais valores se distanciarem espacialmente. O Índice Local de Associação Espacial (LISA) é uma decomposição do índice global de associação espacial e permite a identificação de padrões de associações espaciais significativos (ANSELIN, 2004), ou seja, permite identificar agrupamentos espaciais (*clusters* espaciais).

Os Índices de Moran Global, ou univariado, Figura 1 (X) e multivariado, Figura 5.1 de (A) a (E), mostram a existência de autocorrelação espacial em todos os casos. O univariado, para o IND_EMP, individualmente, indica que há autocorrelação espacial, com valores positivos, expressando a formação de clusters ou agrupamentos reforçando a existência de vizinhos semelhantes cujo valor é $I = 0,2518$.

No caso do Moran multivariado (Figura 5.1 de (A) a (E)), em que se utiliza cada um dos demais indicadores sob estudo versus IND_EMP (Figuras 5.1.(A) a 5.1.(F)) observa-se que todas as correlações espaciais foram significativas (p-valor $< 0,05$) apresentando-se na seguinte ordem crescente: IND_DES (A) Moran's $I = 0,1539$, IND_ESC (B) Moran's $I = 0,2530$, IND_POB (C) Moran's $I = 0,2262$ e IND_JUV (D) Moran's $I = 0,1367$, indicando a existência de autocorrelação espacial positiva. Deste modo, pode-se concluir que, à medida que os índices aumentam, há uma tendência de melhoria na taxa de emprego formal sobre a PIA, como também no último caso (Figura 1(E)), em que verifica o IND_EMP versus Novo Índice de Inclusão (IIS) mostra o $I = 0,1953$, indicando que na medida que ocorrem melhorias no IND_EMP observa-se melhores níveis de inclusão social para o município.

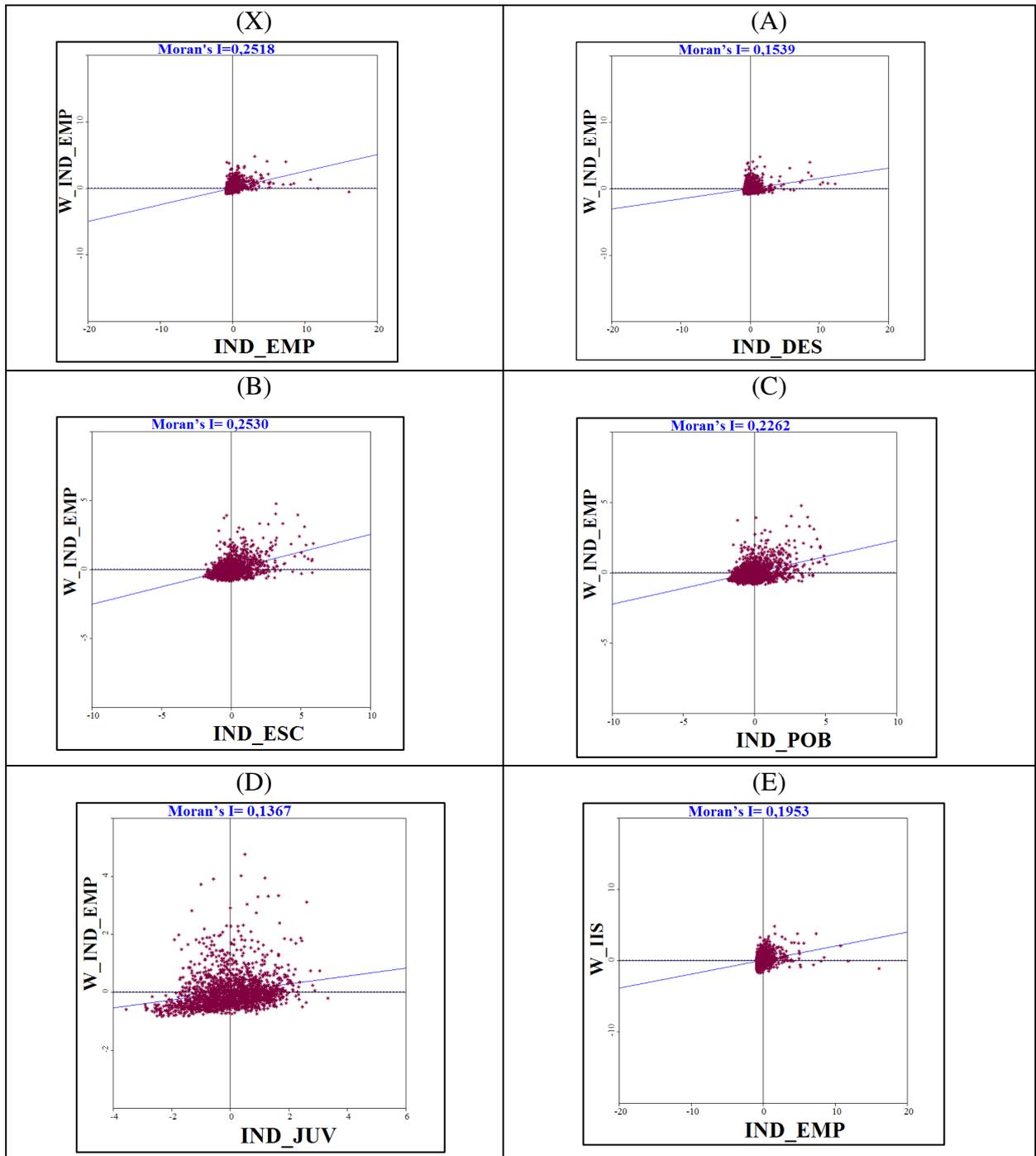
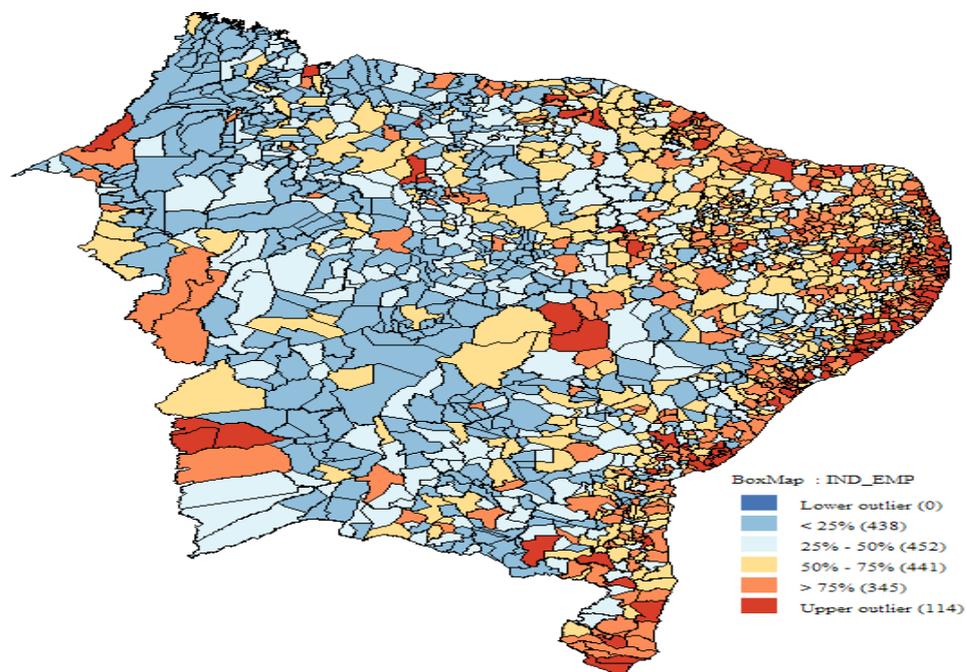


Figura 5.1: Índice de Moran multivariado - versus IND_EMP - para o próprio IND_EMP (X) e os demais índices IND_DES (A), IND_ESC (B), IND_POB (C), IND_JUV (D) e W_IIS com IND_EMP (E), Região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

A distribuição espacial da taxa de emprego formal sobre a PIA representada através do *BoxMap* (Mapa 5.1), equivalente à distribuição dos dados em *Boxplot*, utilizado na análise exploratória convencional, só que espacialmente referido, mostra a existência de 114 *outliers* (correspondente a 6,4% dos municípios) e alguns agrupamentos espaciais bem definidos. Os

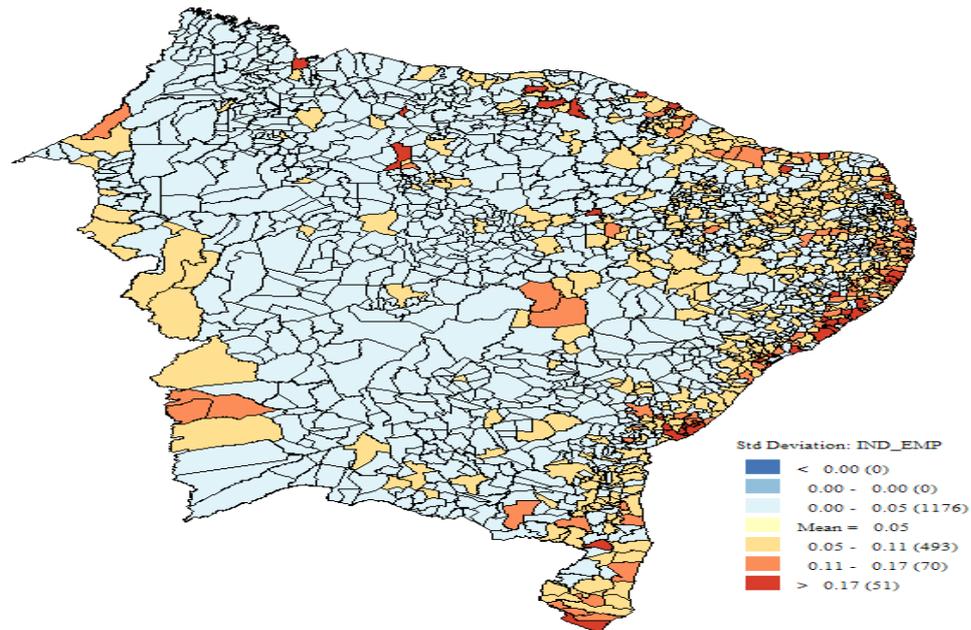
outliers estão enquadrados nos seguintes estados com a correspondente quantidade de municípios: Maranhão (2 municípios - Valores 0,12 e 0,25), Piauí (3 municípios – Valores 0,12, 0,21 e 1), Sergipe (10 municípios - Valores de 0,12 a 0,34), Paraíba (11 municípios – Valores de 0,12 a 0,35), Ceará (13 municípios – Valores de 0,12 a 0,68), Rio Grande do Norte (13 municípios – Valores de 0,12 a 0,75), Alagoas (17 municípios - Valores de 0,11 a 0,30), Pernambuco (21 municípios - Valores de 0,12 a 0,55) e Bahia (24 municípios - Valores de 0,12 a 0,49), mostrando, contudo, que apesar destes municípios apresentarem valores discrepantes para o IND_EMP, em todos os estados esses índices iniciam com valores bastante baixos.



Mapa 5.1: BoxMap da taxa de emprego formal sobre a PIA (IND_EMP) para o ano 2000, Região Nordeste.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Observando o mapa do desvio padrão para o IND_EMP (taxa de empregos formais sobre a PIA), Mapa 5.2, percebe-se que a maior parte dos municípios 65,7% (1176) apresenta IND_EMP inferior à média da região correspondente a 0,05 e localizam-se, em sua maioria, nos municípios centrais até proximidades do litoral da região Nordeste (com cor azul clara no mapa). Logo, a situação do Nordeste é inexpressiva, no que diz respeito ao emprego formal.



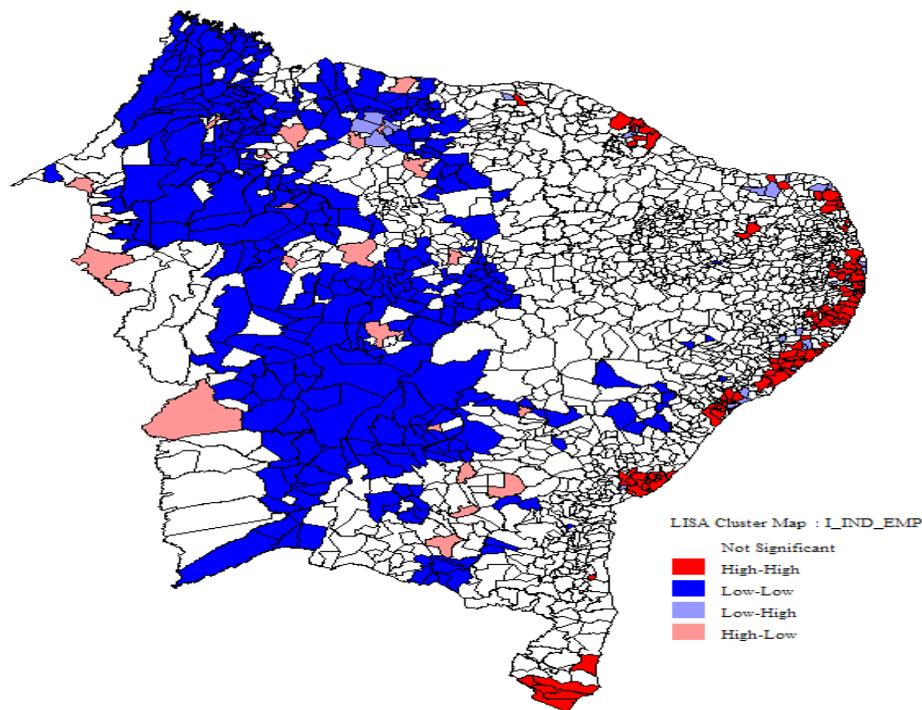
Mapa 5.2: Mapa do desvio padrão (*Standard Deviation*) para o IND_EMP, Região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Os Mapas de 5.3 a 5.8, *Clusters Maps*, retratam o comportamento dos municípios quanto à formação de agrupamentos (clusters) de acordo com o IND_EMP, isoladamente, e quando associado às demais variáveis de interesse sob estudo (IND_POB, IND_DES, IND_ESC, IND_JUV e IIS). A distribuição espacial dos índices através desses mapas são reveladores da autocorrelação espacial apresentada pelos dados em sua forma univariada e bivariada, esta última revelando o efeito de um indicador sobre o outro (Índice Local de Associação Espacial ou LISA, do inglês *Local Indice Spatial Association*), razão pela qual estão sendo analisados.

Nos mapas, os grupos de municípios em vermelho, apresentam a condição *High-High* (alto-alto) que se caracterizam por serem os municípios com alto índice e que apresentam vizinhos próximos também com índices altos; o contrário acontece com a condição *Low-Low* (baixo-baixo), azul escuro, em que um município e seus vizinhos apresentam baixos índices. Observa-se, também, as condições *High-Low*, em que um município com valor alto, e ao seu redor, apresenta vizinhos com valores baixos; e *Low-High*, onde ocorre o contrário, um município com valor baixo apresenta vizinhos com valores altos. Note-se que a formação de clusters *High-High* (alto alto) só se fazem presentes ao longo do litoral da região Nordeste, áreas onde encontram-se encravadas as capitais das UF's nordestinas, nas quais as condições do emprego formal são relativamente mais favoráveis que nos restantes dos municípios. No sentido oposto, a condição *Low-Low* (baixo baixo), é típica de grupos de municípios do interior da região encravadas no semi-árido nordestino, onde a condição do emprego formal é praticamente inexistente.

Ao analisar a distribuição da taxa de emprego formal sobre a PIA através dos mapas indicadores locais de associação espacial (*Lisa map clusters*), Mapa 5.3, percebe-se a existência de agrupamentos de municípios com taxas igualmente baixas (*Low-Low*) com associações espaciais estatisticamente significantes ($p\text{-valor} < 0,05$), nos municípios que se encontram nas áreas, em sua maioria, do lado oeste da região. Agrupamentos, estatisticamente significantes ($p\text{-valor} < 0,05$), reunindo municípios com elevadas taxas (*High-High*) ficaram os municípios localizados nas margens litorâneas da região Nordeste. As taxas contrárias (*Low-High* ou *High-low*), foram distribuídas em poucas localidades ressaltando que os municípios com as cores rosa (*High-low*) possuem taxas elevadas e têm ao seu redor municípios com taxas baixas. Já os que estão com as cores azul claro (*Low-High*), a situação se inverte: eles possuem taxas baixas e estão próximos dos municípios com áreas altas, ou seja, os municípios têm baixo índice e seus vizinhos altos.



Mapa 5.3: Cluster Map, representativo da distribuição espacial do IND_EMP, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.

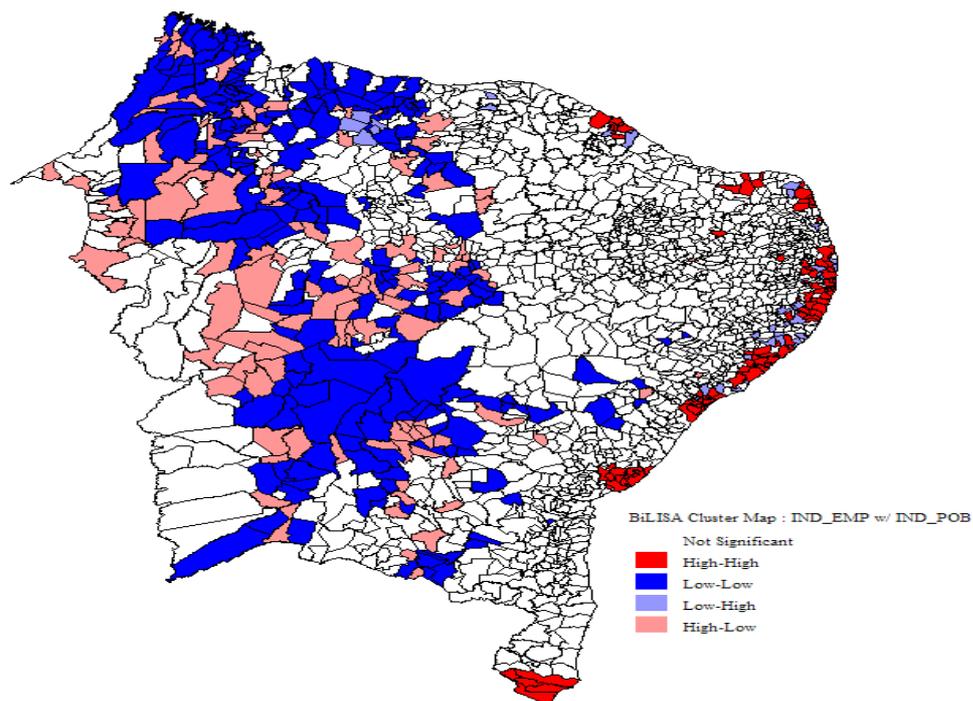
Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Por fim, nas análises das distribuições espaciais através dos mapas indicadores locais de associações espacial, *BiLisa map clusters*, tem-se a análise conjunta do IND_EMP com cada uma das demais variáveis de interesse (Mapas 5.4 a 5.8).

Os *clusters* apresentados pelo Mapa 5.4, do IND_POB versus o IND_EMP, mostram forte correlação na distribuição espacial, ou seja, os clusters espaciais determinados revelam a

forte associação desses indicadores, ou seja municípios com elevados bons níveis de pobreza tendem a apresentar melhores níveis de emprego formal.

A associação espacial, *Bilisa map* (Mapa 5.4), para o IND_EMP com o IND_POB, aponta a existência de agrupamentos de municípios com taxas igualmente baixas (*Low-Low*) com associações espaciais estatisticamente significantes ($p\text{-valor} < 0,05$), nos municípios situados nas áreas, em sua maioria, à oeste da região (do sertão e semi-árido Nordeste). As taxas em que os agrupamentos de municípios se apresentam com altos valores das taxas (*High-High*) estão localizados em áreas litorâneas ou bem próximas ao litoral, no entorno das capitais, conforme já comentado. Quanto aos agrupamentos com taxas contrárias, a maioria é do tipo *High-Low*, apresentam-se próximos aos municípios com áreas baixas, enquanto os agrupamentos do tipo *Low-High*, se apresentam, em sua maioria, próximos aos municípios com áreas altas.



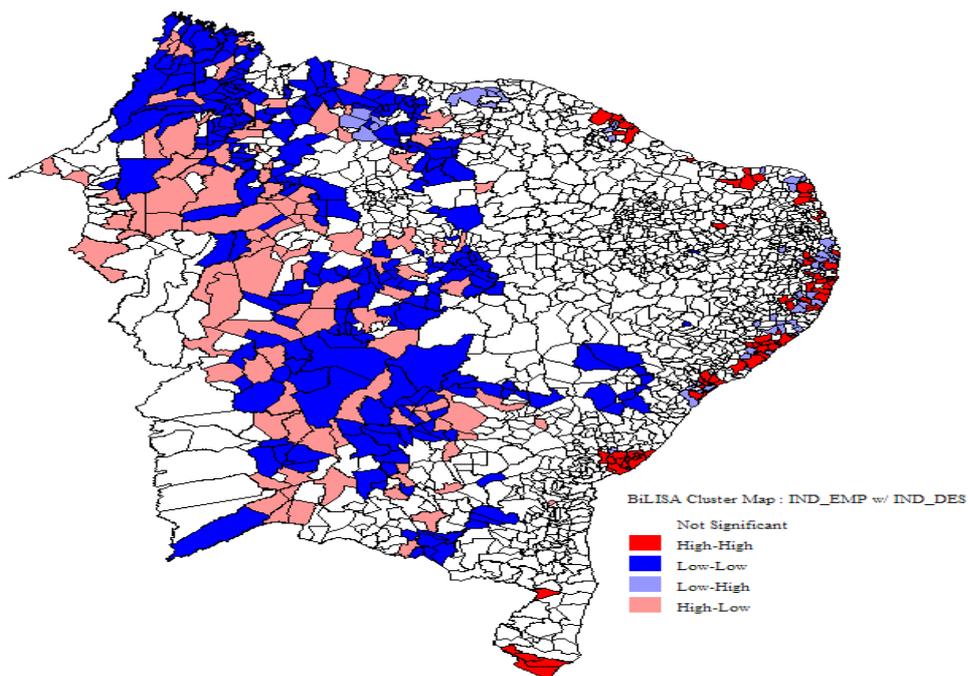
Mapa 5.4: Cluster Map, representativo da distribuição espacial do IND_EMP e IND_POB, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Os *clusters* apresentados pelo Mapa 5.5, do IND_DES versus o IND_EMP, mostram também a forte correlação na distribuição espacial, ou seja, com os clusters espaciais, revelando a forte associação desses indicadores, ou seja municípios com melhores índices de desigualdade de renda, tendem a apresentar bons níveis de emprego formal. Essa análise se

repete para os mapas 5.6 e 5.7, pois há também uma forte correlação espacial do IND_ESC e IND_JUV com IND_EMP.

O Mapa 5.5, mostra a associação espacial *Bilisa map* para o IND_EMP com o IND_DES, apontam a existência de grupamentos de municípios com taxas igualmente baixas (*Low-Low*) com associações espaciais estatisticamente significantes, nos municípios situados nessas áreas, em sua maioria, à oeste da região. Os agrupamentos de municípios onde as taxas se apresentam com altos valores (*High-High*) estão localizados em áreas litorâneas ou bem próximas ao litoral. Quanto aos agrupamentos com taxas contrárias, a maioria é do tipo *High-Low*, pois são municípios com valores altos que estão próximos a valores baixos que neste caso se apresenta nas áreas centrais da região. Já os agrupamentos do tipo *Low-High*, foram mais comuns próximos às áreas litorâneas, revelando o mesmo padrão anteriormente comentado da associação IND_EMP com IND_POB.

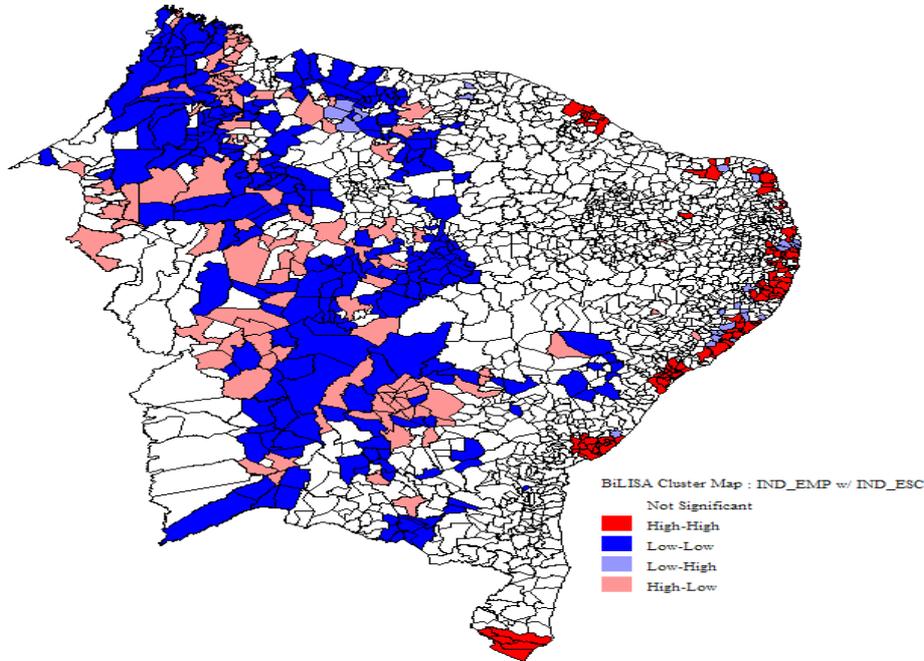


Mapa 5.5: Cluster Map, representativo da distribuição espacial do IND_EMP e IND_DES, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

O Mapa 5.6, expressa a associação espacial *Bilisa map* para o IND_EMP com o IND_DES a análise é bastante parecida com a figura anterior, na qual apontam a existência de agrupamentos de municípios com taxas igualmente baixas (*Low-Low*), com associações espaciais estatisticamente significantes, nos municípios situados nas áreas, em sua maioria, à oeste da região. Os agrupamentos de municípios se apresentam com altos valores das taxas (*High-High*), estão localizados em áreas litorâneas ou bem próximas ao litoral. Quanto aos

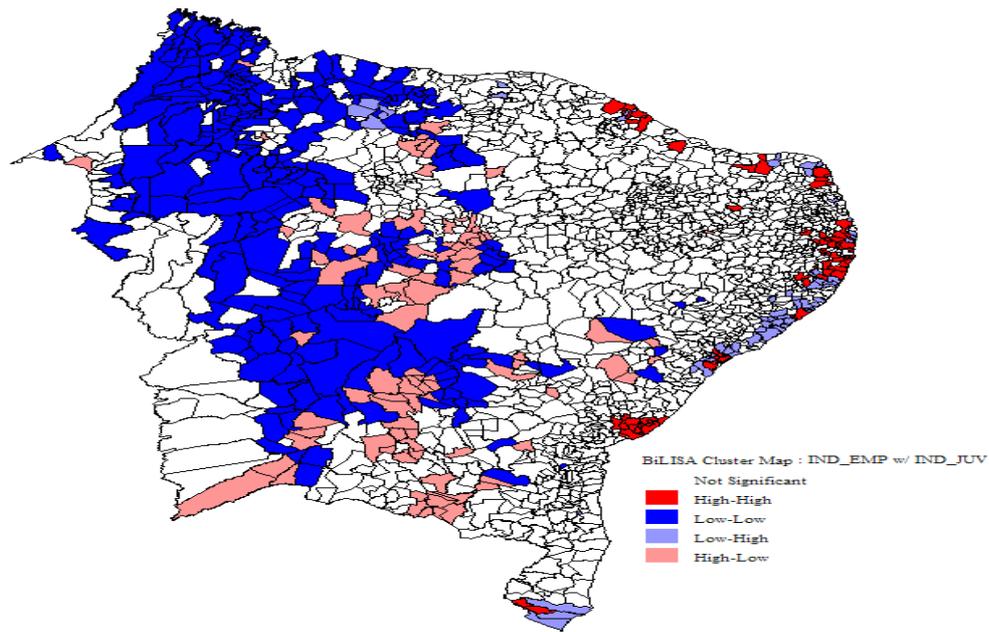
agrupamentos com taxas contrárias, a maioria é do tipo *High-Low*, pois são municípios com valores altos que estão próximos a outros com valores baixos que neste caso se apresentam nas áreas centrais da região. Já os agrupamentos do tipo *Low-High*, foram mais comuns próximos as áreas litorâneas.



Mapa 5.6: Cluster Map, representativo da distribuição espacial do IND_EMP e IND_ESC, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

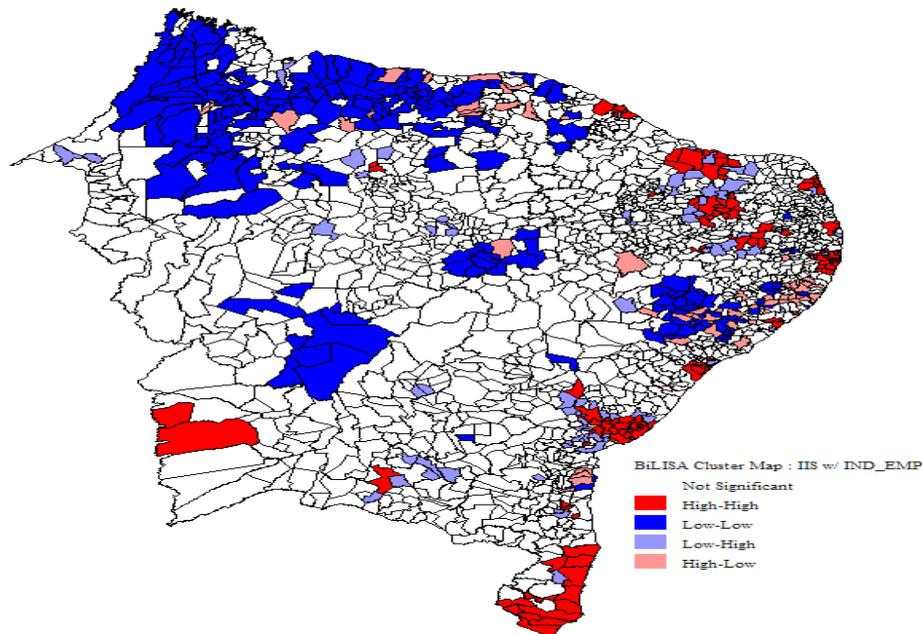
Observa-se no Mapa 5.7, a distribuição da taxa de emprego formal sobre a PIA através do mapa de associação espacial, *BiLisa map clusters*, que a existência de grupamentos de municípios com taxas igualmente baixas (*Low-Low*) com associações espaciais estatisticamente significantes ($p\text{-valor} < 0,05$), nos municípios que se encontram nas áreas, em sua maioria, à esquerda da região. Agrupamentos, estatisticamente significantes ($p\text{-valor} < 0,05$), reunindo municípios com elevadas taxas (*High-High*) ficaram os municípios localizados nas margens litorâneas da região Nordeste. Já os municípios com taxas contrárias, *Low-High* ou *High-low*, foram distribuídos em menos localidades.



Mapa 5.7: Cluster Map, representativo da distribuição espacial do IND_EMP e IND_JUV, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Por último, observa-se no Mapa 5.8, a distribuição da taxa de emprego formal sobre a PIA com o IND_JUV revelado pela associação espacial, *BiLisa map clusters*, que a existência de grupamentos de municípios com taxas igualmente baixas (*Low-Low*) com associações espaciais estatisticamente significantes ($p\text{-valor} < 0,05$), nos municípios que se encontram nas áreas, em sua maioria, na parte noroeste da região, sendo o maior cluster que abrange alguns municípios dos estados do Maranhão e Piauí. Agrupamentos estatisticamente significantes ($p\text{-valor} < 0,05$), reunindo municípios com elevadas taxas (*High-High*), ficaram os municípios localizados, em sua maioria nas margens litorâneas ou próximas ao litoral da região Nordeste, e com taxas contrárias (*Low-High* ou *High-low*), foram distribuídas localidades mais para o interior da região, onde se situa o sertão e semi-árido da região.



Mapa 5.8: Cluster Map, representativo da distribuição espacial do IIS e IND_EMP, segundo a formação de agrupamentos espacialmente definidos, Região Nordeste, 2000.

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

No geral, os mapas mostram que a parte azul quase sempre se apresenta bem espalhado nos estados do Maranhão e Piauí, descendo na parte oeste do mapa incluindo uma parte do estado da Bahia. Esses municípios, em especial, apresentam índices ainda mais precários, no tocante aos 1790 municípios presentes na região Nordeste. Contudo, sabe-se que a região como um todo não se encontra em uma situação confortável, mostrado pelos indicadores analisados, precisando de uma melhora urgente na inclusão social de toda a população.

CAPÍTULO 6: CONCLUSÕES E RECOMENDAÇÕES

Este trabalho buscou empreender uma análise detalhada do emprego formal na região Nordeste brasileira e de seus efeitos na inclusão social, tomando por base indicadores do Atlas da Exclusão Social no Brasil (AESB) (POCHMANN; AMORIM, 2003). Foi mostrado que esses efeitos são de “mão dupla”: tanto o emprego formal pode se traduzir em exclusão social quanto a exclusão social, traduzida por baixa escolaridade, grande contingente de jovens, elevada pobreza e desigualdade, revela-se como perpetuadora dos baixos níveis do emprego formal na Região.

Sendo esse um tema de bastante relevância na sociedade brasileira, em particular na região Nordeste, marcada por uma condição social das mais graves do país, torna-se fundamental o conhecimento detalhado dessa temática, tendo como determinantes os diversos indicadores de inclusão social, entre eles o emprego formal. Assim, a problemática do emprego formal foi analisada de maneira bastante cautelosa, de modo a fornecer alguns subsídios úteis, para que as decisões de políticas e de investimentos sociais possam ser melhor formuladas. Nesse contexto, as conclusões e recomendações do presente trabalho são, a seguir, colocadas.

Um resgate dos resultados da análise exploratória, no contexto das Unidades Federativas da Região Nordeste, mostra que o emprego formal para o estado do Rio Grande do Norte, apresentou-se com o melhor índice dentre as demais UF's, apesar de ainda ser bastante baixo o número de pessoas empregadas com carteira assinada. Quanto aos demais indicadores sob estudo (índices de pobreza, desigualdade, escolaridade e juventude), constata-se, também, um destaque para o estado do Rio Grande do Norte, onde todos eles se apresentaram com melhores índices, no sentido contrário ao observado para os estados do Maranhão e Piauí, onde esses indicadores exibem os piores índices.

Novamente, um resgate dos resultados, agora referentes às capitais de cada estado, permite observar comportamentos semelhantes aos das Unidades Federativas, ficando o diferencial por conta dos níveis mais elevados dos indicadores das capitais, quando comparados com os das UF's. Destacam-se, nesse resgate, os seguintes municípios das capitais: Recife, Salvador, Aracaju, Natal e João Pessoa. Constatam-se, nas análises, que as capitais onde, sabe-se, existem maiores investimentos socioeconômicos, exibem melhorias bem evidenciadas no nível de pobreza, escolaridade e juventude, quando comparadas às das respectivas UF's. Por outro lado, a desigualdade de renda e a taxa de emprego formal pouco

diferem das observadas nas UF's, apresentando, igualmente a estas últimas, baixíssimos valores. Assim, torna-se imperativo que órgãos governamentais busquem estratégias para o aumento do emprego formal, que tem uma forte associação com a desigualdade de renda, de forma que este aumento venha a acarretar, também, melhorias na inclusão social da população.

Descendo a desagregação espacial ao nível municipal, constata-se que, dos 1787 municípios da região Nordeste, 93,6% deles apresentaram uma exacerbada precarização do trabalho na região, refletido pelos baixíssimos níveis do emprego formal, com índices oscilando entre 0,00 e 0,111. Os municípios que apresentaram melhores índices, são todos considerados muito afastados da distribuição geral dos dados, mesmo variando entre 0,111 e 0,747 que representam 6,4% dos municípios da região (valores discrepantes ou *outliers*), padrão que se repete em cada Unidade Federativa, indicando, assim, que há uma grande concentração em baixos níveis de emprego formal em todas as UF's. Ressalta-se, ainda, que existem 40 municípios em que, praticamente, nenhum indivíduo trabalha com carteira assinada, vez que o índice de emprego formal assume o valor zero.

Da análise das associações, pode-se observar que as correlações do emprego formal com as variáveis referentes à pobreza, desigualdade, escolaridade e juventude, apresentaram-se estatisticamente significantes ($p\text{-valor} < 0,05$), tanto para a região Nordeste, quanto para os respectivos estados, observando-se especial destaque para os índices de escolaridade e desigualdade, para todos os estados da região Nordeste, pondo em evidência, mais uma vez, a força da educação e de uma sociedade mais igualitária nos níveis de emprego formal da Região.

Ainda no nível de desagregação municipal, o ajuste dos modelos de regressão para a região Nordeste, bem como para cada Unidade Federativa da região, tomando o IND_EMP como variável resposta e o município como unidade de análise, mostra que, quase sempre, são as variáveis independentes traduzidas pela desigualdade e escolaridade, que contribuem de forma significativa para explicar o emprego formal ($p\text{-valor} < 0,05$). Constata-se, porém, que o IND_DES é a variável de maior poder explicativo para o emprego formal em quase todos os ajustes.

Quando se busca as variáveis mais significativas na explicação da pobreza, pelo ajuste do modelo de regressão, tomando como variável resposta o IND_POB, identifica-se, como principais fatores de explicação, o IND_ESC e o IND_EMP (ao nível de significância de 5%), tanto para a região Nordeste como para as UF's. Logo, é importante que haja investimentos

na busca de melhorias, tanto na educação quanto na oportunidade de gerar empregos formais, acarretando, dessa forma, redução no nível de pobreza da população.

As vantagens de uma sociedade mais educada são notórias e indiscutíveis, apesar do país viver uma “nova exclusão” (que se complementa à exclusão tradicional), constituída por famílias pequenas, nascidas nas metrópoles, com melhor escolaridade e também desempregadas, tidas como excluídas. Portanto, apesar desse indicador ser fundamental para uma melhora em vários aspectos da sociedade, são necessárias mudanças positivas nos demais indicadores estudados, em especial, o emprego formal, que acarretará melhorias da desigualdade de renda e na pobreza acarretando, em consequência, um melhor índice de inclusão social.

Quando se estabeleceu a associação do emprego formal com o novo índice de inclusão social (IIS), tomando-se as UF's brasileiras como unidades de análise (pela aplicação da análise de *cluster* - critério de *linkage*), observou-se que os estados das regiões Nordeste e Norte ficaram, naturalmente, separados dos estados das demais regiões, por estarem entre aqueles de mais baixos índices de emprego e inclusão social. No contexto dos estados da região Nordeste, constatou-se a formação de agrupamentos bastante consistentes, nos quais os estados do Piauí e Maranhão ficaram em um grupo, representando os piores índices e os estados da Bahia, Paraíba, Ceará, Sergipe e Alagoas formaram um outro grande grupo que tiveram índices intermediários, enquanto o estado do Rio Grande do Norte ficou em um grupo isolado, com os melhores índices. Esses resultados mostram os problemas de inclusão social e emprego formal entre as UF's da região Nordeste, especialmente nos estados do Piauí e Maranhão, que estão com os piores indicadores da região. A apresentação dos diferenciais sócio-espaciais das desigualdades apontam a necessidade de adoção de medidas inclusivas, que venham a reduzir o desemprego, aumentando o emprego formal e melhorando a qualidade de vida da população.

Ainda com relação ao IND_EMP e o IIS, agora em nível de desagregação municipal, constatou-se uma concentração de municípios em valores baixos, retratando, porém, o fato já demonstrado de melhorias no emprego formal se fizerem acompanhar de uma maior inclusão social dos municípios. Relação inversa foi observada, quando se investigou a associação da variável referente à transferência governamental com o IIS, levando a crer que as transferências governamentais não estão acarretando melhorias na inclusão social dos municípios. Tal constatação pode ser explicada pelo fato de a variável de transferência governamental representar a proporção de pessoas no município vivendo em famílias onde 50% ou mais da renda familiar, ser proveniente de transferências governamentais

(PP_50RTG), e assim, quanto maior essa proporção no município, menor o índice de inclusão social (IIS) a ele associado. O que se pode concluir desse fato é que, provavelmente, as transferências governamentais estão sendo direcionadas corretamente, ou seja, para aqueles municípios carentes de inclusão social, porém, apesar desses investimentos o quadro de exclusão social continua presente no município. Certamente, sem maiores investimentos na melhoria da escolaridade da população e em mecanismos de geração de emprego, os municípios da região Nordeste permanecerão exibindo deploráveis níveis de inclusão social.

A distribuição espacial da variável IND_EMP para a região Nordeste vem reforçar ainda mais as constatações até agora apresentadas. A estatística I de Moran, mostrou uma significativa autocorrelação desta variável com formação de *clusters* espaciais, com uma grande concentração em valores baixos. Da análise conjunta do IND_EMP com o IND_DES e com o IIS, deparou-se com um padrão espacial bem semelhante, mostrando, também, uma significativa associação espacial e uma grande concentração em valores baixos desses índices. As associações espaciais do IND_EMP com o IND_POB e com o IND_ESC reforçaram todas as conclusões anteriores, nas quais municípios com baixos níveis de emprego formal encontram-se em piores condições de pobreza e escolaridade.

A distribuição espacial do IND_EMP através do *BoxMap*, Mapa do desvio padrão, e os *Clusters Maps*, apresentaram agrupamentos espaciais bem definidos, onde a grande maioria dos municípios se apresentaram com baixos índices, reforçando o panorama geral do deficiente emprego formal na região Nordeste. Nesse contexto, mesmo os destaques apresentados pelos *outliers*, que correspondem a 6,4% dos municípios, mostram-se, ainda assim, muito baixos, evidenciando, dessa forma, como são insatisfatórios estes índices na Região.

Concluem-se, pelas diversas formas de associações espaciais, estatisticamente significantes ($p\text{-valor} < 0,05$), encontradas, que, mesmo no universo quase uniforme da exclusão social dos municípios, foi possível identificar agrupamentos espaciais reveladores de uma melhor inclusão social, em municípios localizados nas margens litorâneas, onde situam-se as capitais das UF's da região Nordeste, fato já esperado pela maior concentração de investimentos econômicos nessas áreas, conforme já referido. Um padrão oposto, de mais baixa inclusão, muito mais recorrente ao longo de toda a Região, pode ser percebido nos agrupamentos de municípios localizados, em sua maioria, do lado oeste da região Nordeste, mais especificamente nos estados do Maranhão e Piauí e parte do estado da Bahia, adentrando para áreas do semi-árido da Região.

Por fim, de um modo geral, esses resultados ilustram um cenário que apresenta debilidades no âmbito socioeconômico, em toda a região Nordeste, ficando evidente a importância de investimentos que propiciem um aumento na geração do emprego formal e, de modo especial melhorias na escolaridade da população, de forma a acarretar melhores níveis de inclusão social em todos os municípios da Região, os quais se farão visíveis pela redução da pobreza e das desigualdades sociais.

Algumas recomendações para a realização de futuros trabalhos são:

- Investigações detalhadas e mais específicas, especialmente para aqueles municípios nos quais a taxa de emprego formal sobre a PIA é zero (classificado como “exclusão total”), bem como no município de Porto, no estado do Piauí, que apresentou índice igual à unidade (classificado como “inclusão total”), quando se sabe que este estado ficou enquadrado, junto com o estado do Maranhão, nas piores situações;
- Aprofundar os estudos referentes à associação do emprego formal e transferências governamentais na melhoria da inclusão social;
- Investigações focadas nos poucos municípios que se mostraram na análise espacial dos *Clusters Maps*, com taxas igualmente altas (*High-High*) mesmo estando distantes das margens litorâneas onde se situam as capitais das UF's.

REFERÊNCIAS

ANSELIN, L. (Ed.). *SpaceStat Tutorial: A Workbook for Using SpaceStat in the Analysis of Spatial Data*. Morgantown: Regional Research Institute, West Virginia University, 1992.

_____. **GeoDa 0.95i Release Notes**. Spatial Analysis Laboratory. Department of Agricultural and Consumer Economics / University of Illinois, Urbana Champaign/Center for Spatially Integrated Social Science, Illinois, 2004.

ARAÚJO, T. B. Nordeste, Nordestes: Que Nordeste? In: AFFONSO, R.; SILVA, P. L. **Federalismo no Brasil: Desigualdades regionais e desenvolvimento**. São Paulo: FUNDAP/UNESP, 1995.

ARAÚJO, T. P.; LIMA, R. A. Estrutura ocupacional e renda do trabalho no Brasil: traços recentes de desigualdades regionais e sociais. In: JUNIOR, M.; SANTOS, S. M. (orgs). **Brasil e Nordeste: Ocupação, Desemprego e Desigualdade**. Fortaleza: IDT: Banco do Nordeste, 2007.

BARBOSA, L. M.; RIBEIRO, M. N. O. Avaliação das condições de vida da população do Brasil, Nordeste e Rio Grande do Norte. In: **XV Encontro Nacional de Estudos Populacionais**. Caxambu - MG: ABEP, 18-22 set. 2006. Disponível em: <http://www.abep.nepo.unicamp.br/encontro2006/docspdf/ABEP2006_515.pdf>. Acesso em 10 nov. 2008.

CHAHAD, J. P. Z.. A trajetória recente do emprego formal: uma comparação entre o nordeste e o Brasil. In: JUNIOR, M.; SANTOS, S. M. (orgs). **Brasil e Nordeste: Ocupação, Desemprego e Desigualdade**. Fortaleza: IDT: Banco do Nordeste, 2007.

DRAPER, N.R.; SMITH, H. *Applied Regression Analysis*. 3. ed. New York: John Wiley & Sons, 1998.

FALVO, J. F. A dimensão sócio-ocupacional do mercado de trabalho no nordeste. In: JUNIOR, M.; SANTOS, S. M. (orgs.). **Brasil e Nordeste: Ocupação, Desemprego e Desigualdade**. Fortaleza: IDT: Banco do Nordeste, 2007.

HAIR JUNIOR, J. et al. **Análise Multivariada de Dados**. 5. ed. São Paulo: Bookman, 2006 Reimpressão.

MODELO DE CALIBRAÇÃO - **Regressão Linear vs. Regressão Linear ponderada**. p.107-125. Disponível em: <https://estudogeral.sib.uc.pt:8443/dspace/bitstream/10316/5519/7/Capítulo_3.pdf>. Acesso em 01 nov. 2008.

MONTGOMERY, D. C. *Design and analysis of experiments*. 5. ed. New York: John Wiley & Sons, INC, 2001.

MONTGOMERY, D. C.; PECK, E. A. *Introduction to Linear Regression Analysis*. 2. ed. New York: John Wiley & Sons, 1992.

NAGAMENI, R. **Emprego no Nordeste cresce acima da média nacional**. Empresa Brasil de Comunicação (EBC), 2007. Disponível em: <<http://www.agenciabrasil.gov.br/noticias/2007/03/15/materia.2007-03-15.8905657634/view>>. Acesso em 05 nov. 2008.

OLIVEIRA, M. R. **Abertura Econômica e Desemprego**. Novembro de 2002. Disponível em: <www.economiabr.net>. Acesso em 05 jun. 2007.

PIRES, I. J. B. **Mercado de trabalho: uma leitura a partir dos níveis de emprego e rendimento**. In: MACAMBIRA, J.; SANTOS, S. M. (orgs). **O Mercado de Trabalho Formal no Brasil**. Fortaleza: Imprensa Universitária, 2006.

POCHMANN, M.; AMORIM, R. (orgs). **Atlas da exclusão social no Brasil**. 2. ed. São Paulo: Cortez, 2003.

POCHMANN, M. **A exclusão social no Brasil**. Disponível em: <http://www.unioeste.br/cursos/cascavel/economia/xv_cbe/001.pdf> Acesso em 20 nov. 2008.

_____. **O emprego na globalização - a nova divisão internacional do trabalho e os caminhos que o Brasil escolheu**. 1 ed. São Paulo: Bomtempo Editorial, 2001.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO. **Atlas do desenvolvimento humano no Brasil**. Rio de Janeiro: PNUD/IPEA/FJP/IBGE. 1998. CD-ROM.

RESENDE, A. C. C.; OLIVEIRA, A. M. H. C. **Avaliando resultados de um programa de transferência de renda: o impacto do Bolsa - Escola sobre os gastos das famílias brasileiras**. Estudos Econômicos, 2008. Disponível em: <<http://www.scielo.br/scielo.php>>. Acesso em 10 nov. 2008.

RODRIGUES, I. M.. **Inclusão Social da População Feminina do Estado do Rio Grande do Norte, segundo Mesorregiões, e sua Influência na Mortalidade de Crianças Menores de 5 Anos de Idade, Revelada Através de Análise Estatística Multivariada e Espacial**. Natal, 2007. Monografia (Graduação em Estatística). UFRN / CCET / Departamento de Estatística.

SILVA, I. G. Entre o local e o nacional: a "reforma" do aparelho de Estado no Maranhão. **CADERNO CRH**, Salvador, n. 35, p. 199-226, jul./dez. 2001. Disponível em: <www.cadernocrh.ufba.br/include/getdoc.php?id=920&article=124&mode=pdf>. Acesso em 08 out. 2008.

SELLTIZ, C. et al. **Métodos de Pesquisa em Relação Social**. São Paulo: EDUSP, 1967.

ANEXOS

❖ Estatísticas Descritivas:

Tabela 1

Análise descritiva dos indicadores utilizados no estado, Região Nordeste, 2000 (n = 1757 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,21	0,20	0,07	0,95
IND_EMP	0,04	0,04	0,00	0,48
IND_DES	0,02	0,02	0,00	0,40
IND_ESC	0,25	0,23	0,08	0,76
IND_JUV	0,51	0,52	0,21	0,80

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 2

Análise descritiva dos indicadores utilizados no estado do Maranhão, 2000 (n = 217 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,20	0,18	0,07	0,55
IND_EMP	0,02	0,01	0,00	0,25
IND_DES	0,02	0,02	0,00	0,18
IND_ESC	0,24	0,23	0,10	0,75
IND_JUV	0,40	0,40	0,21	0,60

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 3

Análise descritiva dos indicadores utilizados no estado, estado do Piauí, 2000 (n = 220 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,21	0,20	0,07	0,54
IND_EMP	0,03	0,02	0,00	0,21
IND_DES	0,02	0,01	0,00	0,17
IND_ESC	0,21	0,20	0,08	0,63
IND_JUV	0,51	0,51	0,27	0,72

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 4

Análise descritiva dos indicadores utilizados no estado do Ceará, 2000 (n = 181 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,21	0,19	0,10	0,58
IND_EMP	0,05	0,04	0,00	0,24
IND_DES	0,02	0,02	0,00	0,24
IND_ESC	0,24	0,23	0,12	0,67
IND_JUV	0,52	0,52	0,38	0,76

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 5

Análise descritiva dos indicadores utilizados no estado, estado do Rio Grande do Norte, 2000 (n = 164 municípios)

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,23	0,21	0,09	0,62
IND_EMP	0,06	0,05	0,00	0,35
IND_DES	0,03	0,02	0,00	0,29
IND_ESC	0,29	0,27	0,17	0,70
IND_JUV	0,57	0,56	0,38	0,77

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 6

Análise descritiva dos indicadores utilizados no estado da Paraíba, 2000 (n = 221 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,21	0,19	0,09	0,63
IND_EMP	0,05	0,05	0,00	0,35
IND_DES	0,02	0,01	0,00	0,28
IND_ESC	0,23	0,22	0,09	0,75
IND_JUV	0,56	0,57	0,39	0,75

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 7

Análise descritiva dos indicadores utilizados no estado, estado de Pernambuco, 2000 (n = 178 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,25	0,22	0,08	0,95
IND_EMP	0,06	0,05	0,00	0,38
IND_DES	0,03	0,02	0,00	0,40
IND_ESC	0,28	0,25	0,12	0,76
IND_JUV	0,54	0,54	0,33	0,75

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 8

Análise descritiva dos indicadores utilizados no estado de Alagoas, 2000 (n = 101 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,20	0,18	0,07	0,55
IND_EMP	0,07	0,05	0,00	0,30
IND_DES	0,02	0,01	0,00	0,21
IND_ESC	0,24	0,22	0,13	0,64
IND_JUV	0,45	0,45	0,33	0,64

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 9

Análise descritiva dos indicadores utilizados no estado, estado de Sergipe, 2000 (n = 72 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,24	0,22	0,12	0,60
IND_EMP	0,07	0,06	0,00	0,34
IND_DES	0,03	0,02	0,01	0,31
IND_ESC	0,29	0,26	0,17	0,73
IND_JUV	0,51	0,52	0,37	0,68

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

Tabela 10

Análise descritiva dos indicadores utilizados no estado da Bahia, 2000 (n = 415 municípios).

Variáveis	Média	Mediana	Mínimo	Máximo
IND_POB	0,22	0,20	0,08	0,60
IND_EMP	0,04	0,03	0,00	0,48
IND_DES	0,03	0,02	0,00	0,25
IND_ESC	0,25	0,23	0,11	0,76
IND_JUV	0,53	0,53	0,34	0,80

Fonte: Elaborado pela autora / Pochmann e Amorim (2003).

❖ MRLMP pelo método *stepwise* :**Tabela 3.1**

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Região Nordeste (n = 1757 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,95	0,90	0,90	16041,54	<0,0001	1
IND_ESC	2	0,95	0,90	0,00	39,12	<0,0001	2
IND_POB	3	0,95	0,90	0,00	23,69	<0,0001	3
IND_JUV	4	0,95	0,91	0,00	13,21	<0,0003	4

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.2

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Maranhão (n = 217 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. De Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,97	0,93	0,93	2999,70	<0,0001	1
IND_POB	2	0,97	0,94	0,01	28,76	<0,0001	2
IND_JUV	3	0,97	0,95	0,00	16,83	<0,0001	3
IND_ESC	4	0,97	0,95	0,00	3,67	<0,0567	4

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.3

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Piauí (n = 220 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,98	0,95	0,95	4495,98	<0,0001	1
IND_POB	2	0,98	0,96	0,01	46,35	<0,0001	2
IND_JUV	3	0,98	0,96	0,00	6,13	<0,0141	3

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.4

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Ceará (n = 181 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. De Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,94	0,89	0,89	1493,45	<0,0001	1
IND_DES	2	0,95	0,91	0,01	25,24	<0,0001	2

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.5

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Rio Grande do Norte (n = 164 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,94	0,89	0,89	1293,95	<0,0001	1
IND_ESC	2	0,94	0,89	0,00	1,53	<0,2181	2

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.6

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado da Paraíba (n = 221 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,96	0,92	0,92	2586,28	<0,0001	1
IND_POB	2	0,96	0,92	0,00	5,57	<0,0191	2

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.7

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado de Pernambuco (n = 178 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,96	0,92	0,92	2012,76	<0,0001	1
IND_POB	2	0,97	0,93	0,01	30,24	<0,0001	2
IND_ESC	3	0,97	0,93	0,00	2,75	<0,0992	3

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.8

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado de Alagoas (n = 101 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_POB	1	0,79	0,63	0,63	167,78	<0,0001	1
IND_JUV	2	0,82	0,68	0,05	14,60	<0,0002	2
IND_ESC	3	0,83	0,69	0,01	3,21	<0,0763	3

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.9

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado do Sergipe (n = 72 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,98	0,96	0,96	1703,97	<0,0001	1
IND_ESC	2	0,99	0,97	0,01	25,23	<0,0001	2

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.10

Sumário da regressão *stepwise* para a variável resposta IND_EMP, ajustado para as variáveis independentes básicas (IND_POB, IND_ESC, IND_JUV, IND_DES), Estado da Bahia (n = 415 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,92	0,85	0,85	2286,15	<0,0001	1
IND_DES	2	0,93	0,86	0,01	34,51	<0,0001	2
IND_JUV	3	0,93	0,86	0,00	7,19	<0,0076	3

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.1.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Região Nordeste (n = 1757 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,97	0,94	0,94	26080,66	<0,0001	1
IND_JUV	2	0,97	0,94	0,00	33,48	<0,0001	2
IND_DES	3	0,97	0,94	0,00	7,93	<0,0049	3
IND_EMP	4	0,97	0,94	0,00	28,30	<0,0001	4

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.2.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Maranhão (n = 217 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. De Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,93	0,87	0,87	1444,04	<0,0001	1
IND_DES	2	0,95	0,90	0,03	53,56	<0,0001	2
IND_EMP	3	0,96	0,91	0,02	42,28	<0,0001	3
IND_JUV	4	0,96	0,91	0,00	2,50	<0,1150	4

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.3.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Piauí (n = 220 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_DES	1	0,96	0,92	0,92	2688,38	<0,0001	1
IND_EMP	2	0,97	0,94	0,01	46,35	<0,0001	2
IND_JUV	3	0,97	0,95	0,01	33,89	<0,0001	3
IND_ESC	4	0,97	0,95	0,00	16,20	<0,0001	4

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.4.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Ceará (n = 181 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. De Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,99	0,97	0,97	6072,13	<0,0001	1
IND_JUV	2	0,99	0,97	0,00	11,36	<0,0009	2

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.5.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Rio Grande do Norte (n = 164 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,98	0,97	0,97	4981,72	<0,0001	1
IND_DES	2	0,98	0,97	0,00	1,41	<0,2376	2

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.6.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado da Paraíba (n = 221 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,98	0,96	0,96	5959,69	<0,0001	1

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.7.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado de Pernambuco (n = 178 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,97	0,94	0,94	2554,96	<0,0001	1
IND_EMP	2	0,97	0,94	0,01	28,19	<0,0001	2
IND_DES	3	0,97	0,95	0,00	2,52	<0,1145	3

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.8.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado de Alagoas (n = 101 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,98	0,95	0,95	2093,06	<0,0001	1
IND_EMP	2	0,98	0,97	0,01	30,51	<0,0001	2
IND_JUV	3	0,98	0,97	0,00	12,67	<0,0006	3

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.9.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado do Sergipe (n = 72 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,97	0,95	0,95	1263,85	<0,0001	1
IND_JUV	2	0,98	0,95	0,00	6,77	<0,0113	2
IND_EMP	3	0,98	0,95	0,00	1,08	<0,3029	3

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 3.10.1

Sumário da regressão *stepwise* para a variável resposta IND_POB, ajustado para as variáveis independentes básicas (IND_EMP, IND_ESC, IND_JUV, IND_DES), Estado da Bahia (n = 415 municípios)

Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. de Determinação Múltiplo R ²	Ganho do R ²	Estatística F	p-valor	Variáveis incluídas
IND_ESC	1	0,97	0,94	0,94	6371,92	<0,0001	1
IND_DES	2	0,97	0,94	0,00	1,12	<0,2900	2

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

Tabela 4.1.1

Sumário da regressão *stepwise* para a variável resposta IIS, ajustado para as variáveis independentes básicas (IND_EMP e PP_50RTG), Área Geográfica – Região Nordeste (NE) e Unidades Federativas do NE, 2000.

NORDESTE (n = 1757 municípios)							
Variáveis	Ordem de entrada no modelo	Coef. de Correlação Múltiplo R	Coef. De Determinação Múltiplo R2	Ganho do R2	Estatística F	p-valor	Variáveis incluídas
IND_EMP	1	0,90	0,80	0,80	7244,13	<0,0001	1
PP_50RTG	2	0,91	0,82	0,02	162,25	<0,0001	2
MARANHÃO (n = 217 municípios)							
IND_EMP	1	0,91	0,83	0,83	1068,99	<0,0001	1
PP_50RTG	2	0,93	0,86	0,03	38,61	<0,0001	2
PIAUI (n = 220 municípios)							
IND_EMP	1	0,93	0,86	0,86	1355,13	<0,0001	1
PP_50RTG	2	0,93	0,87	0,01	12,64	<0,0001	2
CEARÁ (n = 181 municípios)							
IND_EMP	1	0,94	0,89	0,89	1406,74	<0,0001	1
PP_50RTG	2	0,94	0,89	0,00	7,81	<0,0058	2
RIO GRANDE DO NORTE (n = 164 municípios)							
IND_EMP	1	0,91	0,83	0,83	816,77	<0,0001	1
PP_50RTG	2	0,92	0,85	0,01	14,02	<0,0003	2
PARAIBA (n = 221 municípios)							
IND_EMP	1	0,91	0,83	0,83	1088,84	<0,0001	1
PP_50RTG	2	0,93	0,87	0,04	65,89	<0,0001	2
PERNAMBUCO (n = 178 municípios)							
IND_EMP	1	0,84	0,70	0,70	406,03	<0,0001	1
PP_50RTG	2	0,88	0,78	0,08	61,75	<0,0001	2
ALAGOAS (n = 101 municípios)							
IND_EMP	1	0,75	0,56	0,56	127,76	<0,0001	1
SERGIPE (n = 72 municípios)							
IND_EMP	1	0,96	0,92	0,92	799,62	<0,0001	1
PP_50RTG	2	0,96	0,93	0,01	6,69	<0,0118	2
BAHIA (n = 415 municípios)							
IND_EMP	1	0,93	0,87	0,87	2724,31	<0,0001	1

Fonte: Elaborado pela autora / Pochmann e Amorim (2003) (Dados censitários).

❖ Descrição dos Métodos Estatísticos

a) Coeficiente de Correlação

Sejam X_1, X_2, \dots, X_n e Y_1, Y_2, \dots, Y_n duas amostras das variáveis aleatórias X e Y , respectivamente. O coeficiente de correlação linear de Pearson é definido, do ponto de vista amostral por:

Equação 2.2 Cálculo do coeficiente de correlação

$$r = \frac{\sum_{i=1}^n (x_i \cdot y_i) - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{\sqrt{\left[\sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2 \right] \cdot \left[\sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2 \right]}}, \text{ onde } n \text{ é o número de observações.}$$

Os valores do coeficiente de correlação (r) se encontram no intervalo $[-1, +1]$. Como r é uma estimativa do coeficiente de correlação de Pearson populacional (ρ) tem-se que, quando o valor é igual a $+1$, a correlação entre as variáveis é perfeita e positiva; quando é igual a -1 , a correlação entre as variáveis é perfeita e negativa e quando é igual a zero não há correlação ou a relação que porventura exista não é linear.

Teste para o coeficiente de correlação

As hipóteses científicas de interesse são:

H_0 = Não existe correlação linear entre as variáveis envolvidas (analisando duas a duas).
($\rho = 0$)

H_1 = Existe correlação linear entre as variáveis. ($\rho \neq 0$)

A estatística de teste é a estatística t de Student, dada por:

Equação 2.3

Cálculo do teste t de Student

$$t = \frac{r}{\sqrt{1 - r^2}} \sqrt{n - 2} \sim t_{(n-2)}$$

Rejeita-se H_0 , Se $|t| \leq t(\alpha, n-2)$

b) Análise de Regressão Linear Múltipla.

A interpretação da variável estatística pode se apoiar em qualquer uma das três perspectivas: a importância das variáveis independentes (sua contribuição individual à variável estatística), os tipos de relações encontradas ou as inter-relações entre as variáveis independentes (HAIR JUNIOR et al., 2006).

Em todas as aplicações, a seleção de variáveis independentes deve ser baseada em suas relações teóricas com a variável dependente. A análise de regressão então fornece um meio de avaliar, objetivamente, a magnitude e a direção (positiva ou negativa) da relação de cada variável independente. O caráter da Regressão Múltipla, que diferencia de sua contraparte univariada, é a avaliação simultânea de relações entre cada variável independente e a medida dependente. Ao fazer essa avaliação simultânea, a importância relativa de cada variável independente é determinada. (HAIR JUNIOR et al., 2006).

Além de avaliar a importância de cada variável, a regressão múltipla também dá ao pesquisador um meio de avaliar a natureza das relações entre variáveis independentes e a variável dependente. A relação assumida é uma associação linear baseada nas correlações entre as variáveis independentes e a medida dependente.

A correlação entre variáveis independentes pode tornar algumas variáveis redundantes no esforço preditivo. Significa que essas variáveis não são necessárias, se um outro conjunto de variáveis independentes que explique essa variável for empregado.

Para Montgomery e Peck (1992), uma equação de regressão linear múltipla apresenta o relacionamento entre uma determinada variável dependente, denominada por y , e duas ou mais variáveis independentes ($x_1, x_2, x_3, \dots, x_k$).

O modelo que expressa a análise de regressão múltipla tem a seguinte representação:

Equação 2.4 **Modelo Geral de Regressão Linear Múltipla.**

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_k X_k + \varepsilon$$

Equação 2.4.1 **Equação de estimação da Regressão Linear Múltipla**

$$\hat{Y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_k x_k$$

em que:

k = número de variáveis independentes;

\hat{Y} = valor predito da variável dependente;

$x_1, x_2, x_3, \dots, x_k$ = variáveis independentes;

β_0 = intercepto da variável dependente, ou valor para variável dependente quando todas as variáveis independentes são iguais a zero;

b_0 = estimativa amostral para o intercepto da variável dependente, ou valor para variável dependente quando todas as variáveis independentes são iguais a zero;

$\beta_1, \beta_2, \beta_3, \dots, \beta_k$ = coeficientes das variáveis independentes;

$b_1, b_2, b_3, \dots, b_k$ = estimativas amostrais para os coeficientes das variáveis independentes.

ε = componente de erro aleatório. Assume-se $\varepsilon \sim N(0, \sigma^2)$

Dentre alguns preceitos básicos para o modelo de Regressão Linear Múltipla pode-se citar observância da normalidade dos erros (ε), ou seja, a suposição de que os erros observados tenham Distribuição Normal, de forma que o fenômeno da homocedasticidade, que é definida como ocorrência de que o termo de erro tenha uma variância homogênea (ou seja, variância constante). Quando tal fato não acontece, então ocorre o fenômeno estatístico de heterocedasticidade, ou heterogeneidade da variância do erro observado na variável.

Uma alternativa para checar a suposição de normalidade dos erros assumida no modelo 2.4 é construir o gráfico de probabilidade normal (MONTGOMERY, 2001). Esse gráfico tem como finalidade determinar se uma amostra de “n” observações pode ser assumida como observações independentes e identicamente distribuídas de uma distribuição normal, de mesma média e mesma variância.

Para construir o gráfico de probabilidade normal, os resíduos do modelo (2.4) são ordenados do menor para o maior. Estes valores ordenados são então colocados em um gráfico de dispersão em função dos respectivos valores esperados segundo a hipótese de normalidade. Se os pontos não se alinharem a uma reta, obtém-se, portanto, evidências contra a normalidade dos erros. Caso a disposição dos pontos corresponda a uma reta, não se descarta a normalidade.

Este procedimento é bastante útil em uma análise de regressão, não somente para averiguar a distribuição dos erros, mas também para observar se um resíduo é muito maior que qualquer outro, sendo este chamado de uma observação discrepante (*outlier*). A presença de uma ou mais observação discrepantes pode distorcer o resultado da análise.

De acordo com Hair Jr et al. (2006), a heterocedasticidade é causada quando a resposta em um modelo de Análise de Regressão Linear obedece a uma distribuição que caracteriza a não aleatoriedade dos erros das observações, ou seja, a não normalidade, que é caracterizada quando a variância para a variável estudada se modifica de acordo com a média.

Para detectar esse problema deve-se proceder a uma análise de resíduos, e em caso de ocorrência de tal situação, como forma de reduzir a heterocedasticidade na distribuição de um modelo de regressão, deve-se transformar os dados, com a utilização de transformações do tipo logarítmicas, inversas ou exponenciais, de acordo com as tendências apresentadas nas análises de resíduos.

Em grande parte dos casos, pelo fato de estarmos perante intervalos de concentração amplos – normalmente, mais de uma ordem de grandeza – é de esperar que a variância em cada ponto da reta seja diferente (MODELO DE CALIBRAÇÃO, 2008), pelo que, a condição de homogeneidade de variância, designada por homoscedasticidade, não é cumprida.

Conseqüentemente, as concentrações mais elevadas tendem a influenciar (ou “pesar”) a reta de calibração, resultando na perda da exatidão, especialmente nos seus pontos mais baixos (MODELO DE CALIBRAÇÃO, 2008).

A utilização do modelo de Regressão Linear Ponderada (RLP) é a forma mais simples e efetiva de harmonizar as discrepâncias entre as variâncias dos diferentes pontos que constituem a reta (MODELO DE CALIBRAÇÃO, 2008).

É muito comum encontrar situações em que o desvio-padrão das subpopulações de y se altera com o valor de x , o que se descreve como uma situação de heteroscedasticidade. Em muitos casos, o desvio-padrão aumenta proporcionalmente com a concentração, o que conduz a um coeficiente de variação constante. Não obstante poderem ser observados comportamentos diferentes, tendo em consideração que estes podem ser função do sinal ou da concentração ou ainda de outros fatores, a ocorrência mais comum de heterocedasticidade é um aumento de variância em função da concentração.

Quando as concentrações mais elevadas tendem a influenciar (“pesar”) mais a reta de regressão do que os pequenos desvios associados às concentrações mais baixas, resultando na diminuição da exatidão, deve-se recorrer ao modelo de regressão linear ponderada. A expressão a ser minimizada toma a seguinte forma (MODELO DE CALIBRAÇÃO, 2008):

Para exemplificar, ao se ter um modelo de regressão linear simples ($y = b_0 + b_1X_1$), tem-se os parâmetros do modelo (b_0 e b_1) da equação de regressão linear ponderada podem agora ser determinados, de forma que a conversão das equações da regressão linear simples nas suas contrapartes ponderadas é facilmente realizada por multiplicação de um fator por

cada somatório da equação, e alterando n (total de pares de dados) por $\sum w_i$ (fator de ponderação) (MODELO DE CALIBRAÇÃO, 2008).

c) **Análise de análise de *cluster* (ou de agrupamento)**

As etapas da operacionalização da análise de agrupamento são:

- a) Definição do problema e respectivos objetivos;
- b) Obtenção dos dados;
- c) Tratamento dos dados (onde são feitos basicamente tratamentos descritivos);
- d) Definição do critério de parença;
- e) Definição da Técnica de agrupamento;
- f) Apresentação dos resultados (grupos);
- g) Avaliação e interpretação dos resultados.

Dentre essas etapas, recomenda-se atenção especial para o tratamento descritivo dos dados, pois caso esta análise não seja devidamente realizada, poderão ocorrer distorções nos resultados da análise de agrupamento.

Medidas de Parencça

Um dos passos fundamentais da análise de agrupamento é a escolha de um critério que meça o quanto dois objetos são “parecidos” ou “diferentes”. Isto é feito através dos chamados coeficientes de parencça, que classificam-se:



Com relação à similaridade, quanto maior o valor respectivo, mais os objetos são “parecidos”. Já com relação à dissimilaridade, quanto maior o valor, menos “parecidos” são os objetos. Dois exemplos desta classificação são dados pelo coeficiente de correlação e pela distância Euclidiana, respectivamente, cujas definições são dadas por:

Dados os vetores aleatórios $\mathbf{X}' = (X_1, X_2, \dots, X_p)$ e $\mathbf{Y}' = (Y_1, Y_2, \dots, Y_p)$, então duas medidas de associação por similaridade podem ser:

- 1 – O Coeficiente de correlação de Pearson entre X e Y , já definido anteriormente;

2 – A distância Euclidiana entre X e Y, definida por:

Equação 2.5

Cálculo da Distância Euclidiana entre dois pontos X e Y.

$$d(\mathbf{X}, \mathbf{Y}) = \sqrt{\sum_{i=1}^p (X_i - Y_i)^2}$$

Os objetos a serem agrupados podem estar descritos por variáveis escaladas em intervalos, binárias, nominais, ordinais, escaladas em proporção ou ainda combinações desses tipos de variáveis. Não há uma medida de dissimilaridade ou de similaridade que sirva para todos os tipos de variáveis. Portanto, devem-se ter os devidos cuidados no tratamento das variáveis, para poder utilizar a medida de parecnça que seja mais adequada.

Técnicas de agrupamento

Uma definição formal de agrupamento nem sempre é aceita universalmente. Porém todas as definições que surgem baseiam-se em duas idéias básicas: Coesão interna dos objetos e isolamento externo entre os grupos.

Os métodos utilizados no trabalho foram: de ligação simples (dendogramas) e por Gráficos dos resultados cruzados (no método *two-way joining results*).

Para a formação dos agrupamentos é necessário que se tenha uma medida de similaridade. O método de ligação simples, usou-se a distância euclidiana, segmento de linha reta entre cada par de observações, para medir a proximidade de cada Unidade Federativa no universo da região Nordeste; verifica-se que quanto menor a distância entre os estados, maior a similaridade entre eles. Para a partição dos grupos faz-se necessário um procedimento de agrupamento, neste caso o procedimento escolhido foi o hierárquico, que tem uma estrutura do tipo árvore, sua representação é chamada de dendograma ou gráfico em árvore e na terminologia em inglês *joining (tree clustering)*. Os grupos são formados por um processo aglomerativo ou divisivo, isto é, o agrupamento se dá através de uma série de sucessivas junções ou uma série de sucessivas divisões. O algoritmo aglomerativo usado foi o da ligação completa, em que o critério de agrupamento se baseia na distância máxima e pode ser chamado de abordagem do vizinho mais distante.

O método *Two-Was joining* é representado por um gráfico de escala de cores que expressa na leitura vertical, os diferenciais entre cada índice composto (quão desigual é a

realidade nos vários estados), e na horizontal, uma diferenciação dos vários índices dentro de cada estado. Uma maior variação de cores indica maior heterogeneidade.

d) Análise espacial

Os dados espaciais consistem em mensurações ou observações feitas em locais específicos ou região (dados de área). A principal característica destes dados é algum tipo de estrutura espacial que tem uma tendência de que observações próximas entre si sejam mais parecidas do que observações distantes. Atualmente, a análise espacial tem sido usada em vários campos, bem como na observância de questões socioeconômicas, em saúde ambiental, na modelagem da exposição a campos eletromagnéticos, na análise de políticas de controle das doenças e seu planejamento etc.

As técnicas da estatística espacial se aplicam de acordo com dimensionalidade e a estrutura de dados.

A dimensionalidade usa técnicas univariadas úteis para a ocorrência de um único padrão de distribuição espacial. Pode-se usar também técnicas multivariadas utilizadas quando mais de um padrão são envolvidos, onde se pretende estudar as relações entre estes padrões. Na dimensionalidade a diferença básica é que no caso univariado a única explicação possível de um padrão observado é em termos de sua tendência geral nas localizações ou, possivelmente, através da tendência dos vizinhos em “mover juntos”, isto é, em ser “espacialmente autocorrelacionados”. No caso multivariado, a explicação adicional está disponível por outros atributos “espacialmente correlacionados”, medidos em cada localização.

Na estrutura dos dados, existem três tipos básicos de dados que são descritos: os dados de localização, os dados de atributos e os dados de interação.

Dados de localização são localidades onde um conjunto de eventos ocorreu, como, por exemplo, uma localização de casos de alta taxa de emprego formal em uma área de estudo.

Dados de atributos são valores ou atributos associados a um conjunto de localizações e também é chamado de dados de área, que se constitui no tipo de dados utilizado no presente estudo.

Dados de interação são medidas quantitativas onde cada associação tem uma ligação ou par de localizações, e geralmente, são dois pontos.

A estrutura de dados pode ter uma outra forma de classificar os dados espaciais. São elas: dados de casos (eventos), dados de área e dados de amostra pontuais. Onde os dados de

casos, freqüentemente denominados processos pontuais, referem-se a localização dos eventos. Dados de área referem-se a características de uma população adstrita à região geográfica. Por último, os dados relativos a amostras de um universo que corresponde à uma distribuição de um evento contínuo no espaço, também denominados de dados geoestatísticos.

Conforme referido, para nosso estudo, a ênfase será na análise espacial de dados de área, mais precisamente na análise exploratória espacial de dados de área.

Os dados de área geralmente são agregados numa base municipal, ou por setores censitários, não dispondo da localização exata dos eventos, mas de um valor agregado por área. O objetivo é determinar a existência de um padrão espacial nos valores observados e, partindo destes valores, construir mapas temáticos. São geralmente agregados artificialmente, para fins administrativos. Os tipos de mapas mais utilizados são:

- Mapas topográficos, que mostram contornos de aspectos naturais ou construídos pelo homem, selecionados, freqüentemente como esquema de referência.
- Mapas temáticos, que mostram conceitos geográficos, como densidade populacional e aspectos climáticos, entre outros.

Quanto às visualizações desses mapas existem várias. Contudo, foi utilizada a visualização por intervalos iguais. Esta visualização são aqueles que definimos pelo valor máximo e (ou) mínimo para mostrar a dispersão nos dados.

Índice Global de associação espacial: Índice de Moran

O Índice de Moran fornece uma medida geral da associação espacial existente no conjunto dos dados. Seu valor varia de -1 a 1 . Valores próximos de zero, indicam a inexistência de autocorrelação espacial significativa entre os valores dos objetos e seus vizinhos. Valores positivos para o índice, indicam autocorrelação espacial positiva, ou seja, o valor do atributo de um objeto tende a ser semelhante aos valores dos seus vizinhos. Valores negativos para o índice, por sua vez, indicam autocorrelação negativa.

O índice de Moran pode ser descrito em função dos elementos básicos vistos anteriormente, é dado por:

Equação 2.6

Cálculo do índice de Moran

$$I = \frac{Z^t \cdot W_z}{Z^t \cdot Z}$$

onde: Z^t é o vetor de desvios transposto e W_z representa a ponderação espacial atribuída ao par de vizinhos i e j . I é formalmente equivalente ao coeficiente de regressão linear. Este coeficiente indica a inclinação da reta de regressão (β_0) de W_z em Z . A interpretação do índice de Moran como um coeficiente de regressão, indica o caminho para se construir um dispositivo gráfico para visualizar a associação espacial entre o valor do atributo de cada elemento (z_i) com a média dos valores dos atributos dos seus vizinhos (Wz_i). Este procedimento é denominado de gráfico de espalhamento de Moran (ANSELIN, 1992).

Testes de significância

Sob a premissa de distribuição normal da variável e variância constante, a medida utilizada tem distribuição normal sob H_0 (ausência de autocorrelação espacial) com esperança e variáveis conhecidas.

Portanto, testes de significância para I podem ser calculados computando-se variáveis padronizadas (z -variáveis). Os W_z são usados no teste de significância e representam a proximidade entre áreas geográficas, expressada pela matriz de vizinhança. Essa matriz é definida por meio da contigüidade, da distância ou quaisquer critérios de vizinhança espacial. No presente trabalho a matriz de vizinhança foi definida pelo critério dos vizinhos mais próximos ou contíguos, com pelo menos um ponto de fronteira em comum.

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](#)