

**Ministério da Saúde
Fundação Oswaldo Cruz
Escola Nacional de Saúde Pública
Mestrado em Saúde Pública – Epidemiologia Geral**

**"MODELOS DE REGRESSÃO ECOLÓGICA:
UMA APLICAÇÃO EM DOENÇA ISQUÊMICA DO CORAÇÃO
RIO DE JANEIRO – 1991"**

Mirian Carvalho de Souza

**Dissertação apresentada à
Escola Nacional de Saúde
Pública, ao departamento de
Epidemiologia e Métodos
Quantitativos em Saúde, para
obtenção do título de mestre.**

Orientadores:

Marilia Sá Carvalho

Iuri da Costa Leite

Rio de Janeiro, novembro de 2004

*Para Leandro e Maria das Neves, meus pais
Marília e Oswaldo, meus maiores colaboradores
sem vocês eu não teria conseguido chegar aqui.*

AGRADECIMENTOS

Agradeço a Deus por me iluminar e guiar até aqui.

À meus pais pela minha formação, pela força e compreensão. Só quem convive com alguém que está desenvolvendo uma dissertação sabe como é...

À meu irmão pelo companheirismo e amizade.

À Marília pela adoção, orientação, incentivo, carinho, dedicação e tudo mais...

À Oswaldo pelas idéias, pelo socorro, incentivo e dedicação...

À Iuri Leite pela orientação, dedicação, disponibilidade e amizade

À Angela, Dora, Flávio e Rejane pelo empenho em ler este trabalho e pelas sugestões muito bem vindas.

À meu amigo Kaizô Beltrão que me conduziu até esta instituição.

À meu amigo Francisco Oliveira pelo apoio e incentivo.

À meus amigos Angela, Debora, Inês, Luís Carlos, Marise, Sônia, Thuller e Victor Hugo pelo carinho e incentivo.

Aos amigos de minha turma que individualmente, ou em conjunto sempre estiveram presentes no desenvolvimento deste trabalho.

À meus novos companheiros da CONPREV/INCA, minha nova casa que me ofereceu as condições que precisei nos últimos meses para o desenvolvimento deste trabalho.

RESUMO

Análise espacial é o estudo quantitativo de fenômenos localizados no espaço geográfico. A unidade de coleta e de análise da informação define a capacidade de diferenciar áreas, por isto a escolha da unidade afeta a variabilidade dos indicadores em estudo podendo alterar de forma substancial os resultados. Além disso a existência de relações espaciais — tendência e autocorrelação — afeta as associações buscadas em modelos ecológicos de regressão.

Neste estudo analisamos a relação entre a taxa de mortalidade por doença isquêmica do coração (DIC) na população de 30 a 70 anos e indicadores socioeconômicos, nos 153 bairros do Rio de Janeiro em 1991. Os indicadores utilizados foram: proporção de casas (construções unifamiliares), proporção de domicílios alugados, proporção de domicílios ligados à rede geral de água, proporção de domicílios ligados à rede de esgoto, proporção de domicílios com coleta regular de lixo, proporção de chefes de domicílio com 1º grau completo ou mais, proporção de população alfabetizada, proporção de chefes de domicílio com renda inferior a dois salários mínimos, proporção de chefes de domicílio com renda superior a quinze salários mínimos, proporção setores de favelados em cada bairro e proporção de população entre 60 e 70 anos no grupo populacional de 30 a 70 anos.

Utilizaram-se os seguintes métodos estatísticos: classificação multivariada (*K-means*), teste de autocorrelação espacial pelo Índice de Moran, regressão linear múltipla, modelo de regressão espacial com erros correlacionados e modelo aditivo generalizado para detecção de tendência espacial.

Para criar perfis socioeconômicos foi feita classificação multivariada utilizando como unidade de análise os 6.259 setores censitários, que gerou quatro grupos. Foram usadas as variáveis: proporção de casas, de domicílios alugados, ligados à rede de esgoto, proporção de chefes de domicílio com 1º grau completo ou mais, com renda inferior a dois salários mínimos e com renda superior a quinze salários mínimos.

O Índice de Moran indica presença de autocorrelação e/ou tendência espacial nos indicadores, tanto na escala de bairros como para setores censitários, usualmente maior nestes.

Nos modelos de regressão ajustados (linear simples, espacial com e sem tendência), as variáveis significativas foram: proporção de população idosa (positivamente correlacionada), proporção de chefes de domicílio com renda inferior a dois salários mínimos (positivamente correlacionada), proporção de chefes de domicílio com 1º grau completo ou mais (positivamente correlacionada) e proporção de chefes de domicílio com renda superior a quinze salários mínimos (negativamente correlacionada). Como quase todas as variáveis apresentam não estacionariedade de primeira ordem, a regressão espacial foi refeita tendo por base os resíduos do modelo aditivo com alisamento por regressão linear local ponderada (*Loess*).

Os modelos estimados, controlando pela proporção de população idosa, apontam todos na mesma direção, embora com diferenças nos parâmetros estimados. A taxa mortalidade por doença isquêmica do coração, na população de 30 a 70 anos assume valores mais elevados nos bairros onde é menor a presença do estrato social mais rico e maior a do estrato social mais pobre, indicando o efeito protetor da renda sobre a mortalidade por este agravo na população estudada. A correlação também positiva com alta escolaridade, aliada à distribuição espacial observada, sugere uma relação com os estratos sociais médios, possivelmente caracterizados nos grupos "B" e "C" da classificação multivariada. O uso dos modelos espaciais neste caso, controlando associações espúrias devidas à estrutura espacial, confirmou as correlações observadas, trazendo importante contribuição para o conhecimento do perfil coletivo da mortalidade por doença isquêmica do coração na população de 30 a 70 anos, nos bairros do município do Rio de Janeiro.

Palavras-Chave: Análise Espacial, Modelo Autoregressivo, Estudo Ecológico e Doença Isquêmica do Coração.

ABSTRACT

Space analysis is the quantitative study of phenomena localized in geographical space. The basic unit for collecting and analyzing information defines the possibility of differentiating areas. The choice of analysis unit affects the variability of indicators in study and can substantially alter the results. Besides, the presence of spatial patterns — trend and autocorrelation — affects the associations that are being studied in ecological regression models.

In this study, the relationship between the Ischemic Heart Disease mortality rate in the population aged 30 to 70 years and socioeconomic indicators, in 153 neighborhoods of Rio de Janeiro city, in 1991, was analyzed. The selected indicators were: proportion of standalone houses (one family per household), proportion of rented homes, proportion of households served by the public system of water distribution, proportion of households linked to the public sewerage system, proportion of households with regular collection of garbage, heads of the household proportion with at least the basic cycle of education, proportion of alphabetized population, proportion of heads of household with income below two times the national minimum wage, proportion of heads of household with income above fifteen times the national minimum wage, proportion census tracts which are characterized as "favelas" (slums) in each neighborhood and proportion of elderly population.

The statistical methods used were: multivariate classification (K-means), Moran I test of space autocorrelation, multiple linear regression, spatial regression with correlated errors and generalized additive model for detection of space trend.

To create socioeconomic profiles, the spatial unite census tract, with 6.259 areas, was used. The multivariate classification of census tracts generated four groups. The variables used were: proportion of standalone houses, of rented homes, linked to the sewerage system, proportion of heads of household with basic education cycle completed, with income below two minimum wages and with income above fifteen minimum wages.

The Moran index indicates the presence of autocorrelation or space trend, in both spatial unites: neighborhoods and census tracts, usually larger in the last ones.

In the linear and spatial regression models the significant variables were: proportion of elderly population (positively correlated), proportion of heads of household with income below two times the national minimum wage (positively correlated), heads of the household proportion with at least the basic cycle of education (positively correlated) and heads of household proportion with income above fifteen minimum wages (negatively correlated). As almost all the variables presented did not presented first order stationary, the spatial regression was re-done on the residuals of a generalized additive model with a local linear weighted regression (*Loess*) as smoothing function.

The estimated models, controlling for the proportion of elderly population, although presenting differences in parameters, point all in the same direction. The Ischemic Heart Disease mortality rate in the population aged 30 to 70 years is higher in neighborhoods with less people of the richer strata, and more people of the lower income groups, indicating the protective effect of income on Ischemic Heart Disease mortality rate. The positive correlation with high scholarship, together with observed aspects of spatial distribution suggests a relationship between the disease and the medium social strata, possibly characterized in groups "B" and "C" of the multivariate classification. The use of spatial models in this case, controlling for spurious association due to spatial structure, confirmed the observed correlation, bringing important contribution to understand the risk profile of Ischemic Heart Disease mortality in groups.

Key Words: Spatial Analysis, Autoregressive Model, Ecological Study, Isquemic Heart Disease

ÍNDICE ANALÍTICO

1 - INTRODUÇÃO	14
2 - REVISÃO BIBLIOGRÁFICA.....	17
2.1 - Doença Isquêmica do Coração	17
2.2 - Análise Espacial	21
2.3 - Análise Exploratória Espacial	23
2.4 - Métodos de Classificação Multivariada.....	25
2.5 - Modelos de Regressão.....	28
3 - DADOS E METODOLOGIA	32
3.1 - Indicadores Socioeconômicos	32
3.2 - Localização Geográfica	34
3.3 - Indicadores de Mortalidade	37
3.4 - Métodos de Análise	38
4 - RESULTADOS E DISCUSSÃO	40
4.1 - Análise Exploratória dos Indicadores por Bairro	40
4.2 - Classificação Multivariada	69
4.3 - Índice de Moran.....	74
4.4 - Regressão Linear Simples.....	75
4.5 - Regressão Espacial.....	82
4.6 - Detecção de Tendência Espacial	87
4.7 - Regressão Espacial Sem Tendência.....	91
5 - CONSIDERAÇÕES FINAIS	96
6 - BIBLIOGRAFIA	100
ANEXO I - ANÁLISE EXPLORATÓRIA UNIVARIADA DOS INDICADORES NA UNIDADE GEOGRÁFICA SETOR CENSITÁRIO	104

ÍNDICE DE TABELAS

TABELA 3.1 - DIAGNÓSTICO DA LOCALIZAÇÃO DOS ÓBITOS POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS	36
TABELA 3.2 - DIAGNÓSTICO DA LOCALIZAÇÃO DOS ÓBITOS POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS – APÓS O GEOREFERENCIAMENTO MANUAL.....	37
TABELA 4.1 - ÍNDICE DE MORAN PARA SETORES CENSITÁRIOS E BAIROS	75
TABELA 4.2 - COEFICIENTES DE REGRESSÃO DO MODELO DE REGRESSÃO LINEAR SIMPLES	76
TABELA 4.3 - ANÁLISE DA VARIÂNCIA DO MODELO DE REGRESSÃO LINEAR SIMPLES.....	77
TABELA 4.4 - COEFICIENTES DE REGRESSÃO DO MODELO CAR.....	83
TABELA 4.5 - ÍNDICE DE MORAN DOS INDICADORES SÓCIO ECONÔMICOS COM E SEM TENDÊNCIA - BAIROS DO RIO DE JANEIRO.....	91
TABELA 4.6 - COEFICIENTES DE REGRESSÃO DO MODELO CAR SEM TENDÊNCIA	92

ÍNDICE DE QUADROS

QUADRO 2.1- ESCALA, RESOLUÇÃO, HOMOGENEIDADE, ESTABILIDADE E SUA RELAÇÃO COM A DIMINUIÇÃO DA ÁREA DA UNIDADE GEOGRÁFICA.....	21
QUADRO 2.2 - UNIDADES GEOGRÁFICAS DE ANÁLISE: ASPECTOS POSITIVOS E NEGATIVOS	22
QUADRO 4.1- DIREÇÃO DAS CORRELAÇÕES ENTRE OS INDICADORES DOS BAIROS DO RIO DE JANEIRO, 1991	69
QUADRO 4.2 - RAIOS MÉDIOS DOS SETORES CENSITÁRIOS	74
QUADRO 4.3 - AVALIAÇÃO DA TENDÊNCIA ESPACIAL DAS VARIÁVEIS BAIROS DO RIO DE JANEIRO, 1991	88

ÍNDICE DE GRÁFICOS

GRÁFICO 2.1 - MORTALIDADE PROPORCIONAL POR CAPÍTULOS DA CID - FAIXA ETÁRIA 30-69 ANOS - MUNICÍPIO DO RIO DE JANEIRO - 1980/95.....	20
GRÁFICO 4.1 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CASAS BAIROS DO RIO DE JANEIRO, 1991.....	41
GRÁFICO 4.2 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS BAIROS DO RIO DE JANEIRO, 1991.....	43
GRÁFICO 4.3 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA BAIROS DO RIO DE JANEIRO, 1991.....	45
GRÁFICO 4.4 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE DE ESGOTO BAIROS DO RIO DE JANEIRO, 1991.....	47
GRÁFICO 4.5 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS EM QUE O LIXO É COLETADO BAIROS DO RIO DE JANEIRO, 1991.....	49
GRÁFICO 4.6 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS BAIROS DO RIO DE JANEIRO, 1991.....	51
GRÁFICO 4.7 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA BAIROS DO RIO DE JANEIRO, 1991.....	53
GRÁFICO 4.8 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS BAIROS DO RIO DE JANEIRO, 1991.....	55
GRÁFICO 4.9 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS BAIROS DO RIO DE JANEIRO, 1991.....	57
GRÁFICO 4.10 - ANÁLISE EXPLORATÓRIA DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS BAIROS DO RIO DE JANEIRO, 1991.....	59
GRÁFICO 4.11 - ANÁLISE EXPLORATÓRIA DA POPULAÇÃO DE 30 A 70 ANOS BAIROS DO RIO DE JANEIRO, 1991.....	61
GRÁFICO 4.12 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE SETORES FAVELADOS BAIROS DO RIO DE JANEIRO, 1991.....	63
GRÁFICO 4.13 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE POPULAÇÃO IDOSA BAIROS DO RIO DE JANEIRO, 1991.....	65
GRÁFICO 4.14 - DIAGRAMA DE ESPALHAMENTO DOS INDICADORES DOIS A DOIS BAIROS DO RIO DE JANEIRO, 1991.....	68
GRÁFICO 4.15 - ANÁLISE EXPLORATÓRIA DE INDICADORES UTILIZADOS NA CONSTRUÇÃO DO CLUSTER.....	71
GRÁFICO 4.16 - SILHUETA DOS ESTRATOS OBTIDOS NA CLASSIFICAÇÃO MULTIVARIADA.....	72
GRÁFICO 4.17 - HISTOGRAMA DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR SIMPLES.....	77
GRÁFICO 4.18 - PROBABILIDADE NORMAL DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR SIMPLES.....	78
GRÁFICO 4.19 - DISPERSÃO DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR EM RELAÇÃO AOS VALORES AJUSTADOS BAIROS DO RIO DE JANEIRO.....	79
GRÁFICO 4.20 - DISPERSÃO DOS RESÍDUOS DO MODELO CAR EM RELAÇÃO AOS VALORES AJUSTADOS BAIROS DO RIO DE JANEIRO.....	84
GRÁFICO 4.21 - DISPERSÃO DOS RESÍDUOS DO MODELO CAR SEM TENDÊNCIA EM RELAÇÃO AOS VALORES AJUSTADOS BAIROS DO RIO DE JANEIRO.....	93
GRÁFICO AI.1 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CASAS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	104
GRÁFICO AI.2 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	105
GRÁFICO AI.3 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	105
GRÁFICO AI.4 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE PÚBLICA DE ESGOTO SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	106

GRÁFICO AI.5 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS EM QUE O LIXO É COLETADO SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	106
GRÁFICO AI.6 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	107
GRÁFICO AI.7 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	107
GRÁFICO AI.8 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	108
GRÁFICO AI.9 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	108
GRÁFICO AI.10 - ANÁLISE EXPLORATÓRIA DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	109
GRÁFICO AI.11 - ANÁLISE EXPLORATÓRIA DA POPULAÇÃO DE 30 A 70 ANOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	109

ÍNDICE DE MAPAS

MAPA 3.1 - DIVISÃO OFICIAL DO RIO DE JANEIRO	35
MAPA 4.1 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CASAS BAIRROS DO RIO DE JANEIRO, 1991	42
MAPA 4.2 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS BAIRROS DO RIO DE JANEIRO, 1991	44
MAPA 4.3 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA BAIRROS DO RIO DE JANEIRO, 1991.....	46
MAPA 4.4 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE PÚBLICA DE ESGOTO BAIRROS DO RIO DE JANEIRO, 1991	48
MAPA 4.5 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS EM QUE O LIXO É COLETADO BAIRROS DO RIO DE JANEIRO, 1991.....	50
MAPA 4.6 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS BAIRROS DO RIO DE JANEIRO, 1991	52
MAPA 4.7 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA BAIRROS DO RIO DE JANEIRO, 1991	54
MAPA 4.8 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS BAIRROS DO RIO DE JANEIRO, 1991.....	56
MAPA 4.9 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS BAIRROS DO RIO DE JANEIRO, 1991.....	58
MAPA 4.10 - DISTRIBUIÇÃO ESPACIAL DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS BAIRROS DO RIO DE JANEIRO, 1991	60
MAPA 4.11 - DISTRIBUIÇÃO ESPACIAL DA POPULAÇÃO DE 30 A 70 ANOS BAIRROS DO RIO DE JANEIRO, 1991	62
MAPA 4.12 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE SETORES FAVELADOS BAIRROS DO RIO DE JANEIRO, 1991	64
MAPA 4.13 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE POPULAÇÃO IDOSA BAIRROS DO RIO DE JANEIRO, 1991	66
MAPA 4.14 - DISTRIBUIÇÃO ESPACIAL DAS CLASSES ENCONTRADAS	73
MAPA 4.15 - DISTRIBUIÇÃO ESPACIAL DAS TAXAS AJUSTADAS DE MORTALIDADE POR DIC NO GRUPO ETÁRIO 30 A 70 ANOS - MODELO DE REGRESSÃO LINEAR SIMPLES BAIRROS DO RIO DE JANEIRO, 1991	80
MAPA 4.16 - DISTRIBUIÇÃO ESPACIAL DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR SIMPLES BAIRROS DO RIO DE JANEIRO, 1991.....	81
MAPA 4.17 - DISTRIBUIÇÃO ESPACIAL DAS TAXAS AJUSTADAS DE MORTALIDADE POR DIC NO GRUPO ETÁRIO 30 A 70 ANOS - MODELO CAR BAIRROS DO RIO DE JANEIRO, 1991.....	85
MAPA 4.18 - DISTRIBUIÇÃO ESPACIAL DOS RESÍDUOS DO MODELO CAR AJUSTADO BAIRROS DO RIO DE JANEIRO, 1991	86
MAPA 4.19 - DISTRIBUIÇÃO ESPACIAL DAS TAXAS AJUSTADAS DE MORTALIDADE POR DIC NO GRUPO ETÁRIO 30 A 70 ANOS - MODELO CAR SEM TENDÊNCIA BAIRROS DO RIO DE JANEIRO, 1991	94
MAPA 4.20 - DISTRIBUIÇÃO ESPACIAL DOS RESÍDUOS DO MODELO CAR AJUSTADO SEM TENDÊNCIA BAIRROS DO RIO DE JANEIRO, 1991.....	95
MAPA AI.1 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CASAS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	110
MAPA AI.2 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	111
MAPA AI.3 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	112
MAPA AI.4 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE PÚBLICA DE ESGOTO SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	113
MAPA AI.5 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS EM QUE O LIXO É COLETADO SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	114

MAPA AI.6 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	115
MAPA AI.7 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	116
MAPA AI.8 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	117
MAPA AI.9 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	118
MAPA AI.10 - DISTRIBUIÇÃO ESPACIAL DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991.....	119
MAPA AI.11 - DISTRIBUIÇÃO ESPACIAL DA POPULAÇÃO DE 30 A 70 ANOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991	120

1 - INTRODUÇÃO

Análise espacial é o estudo quantitativo de fenômenos que são localizados no espaço. Neste tipo de análise, focalizam-se as técnicas que enfatizam explicitamente a localização espacial, cujos principais objetivos são:

- Análise ambiental;
- Estudos ecológicos;
- Delimitação de áreas — para fins de intervenção;
- Análise de acesso a serviços.

A utilização de mapas e a análise da distribuição geográfica de diversas doenças é antiga, remontando a John Snow (1854); entretanto, somente com a recente evolução da computação começaram a se tornar mais usuais os estudos em que se modela a dimensão espaço.

Os métodos de análise podem ser agrupados segundo o tipo de dado (Bailey & Gatrell, 1995):

- **Distribuição de Pontos** – quando o objeto de análise é a posição relativa de eventos precisamente localizados (Ex.: casos de doenças ou espécies vegetais em estudos de ecologia ambiental);
- **Geoestatística** – conjunto de técnicas de modelagem de uma variável contínua no espaço (Ex.: temperatura, poluição, etc.);
- **Dados de Áreas** – quando a ocorrência do fenômeno em estudo é mensurada a partir de dados agregados por área (Ex.: taxas de mortalidade por município, proporções por população, etc.); e
- **Deslocamento** – quando o objeto em estudo é o acesso e o fluxo entre regiões, inclusive otimizando trajetórias e estudando a localização de equipamentos urbanos.

A unidade geográfica de coleta e de análise da informação define a resolução do estudo ou seja, a capacidade de distinguir a localização geográfica do fenômeno estudado. A escolha da unidade geográfica afeta a variabilidade dos indicadores em estudo e pode alterar de forma substancial os resultados. Um dos problemas mais freqüentes encontrados em trabalhos que utilizam dados de áreas, é a extensão da base territorial de cada unidade geográfica de coleta e análise da informação, em geral, essa área é definida tomando-se por base divisões político-administrativas usuais tais como: bairros, distritos, municípios, etc. (Carvalho e cols., 1997). A utilização destas unidades geográficas como unidade de análise do estudo implica em alguns inconvenientes entre os quais destacam-se:

- Informações agregadas em diferentes unidades geográficas, conforme a divisão administrativa do órgão — bairros, áreas programáticas, regiões administrativas;
- As unidades políticas não coincidem com as discontinuidades populacionais, podendo englobar territórios heterogêneos;
- Tamanhos distintos das unidades geográficas.

Outra questão importante na seleção adequada da unidade geográfica de análise refere-se aos custos de coleta da informação. Em geral as informações — gráficas ou textuais — já estão disponíveis para algumas unidades geográficas pré-determinadas e geralmente abrangendo grandes áreas.

O objetivo geral deste estudo é analisar a relação entre a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos — no município do Rio de Janeiro em 1991 — e indicadores socioeconômicos, visando discutir as áreas e grupos sociais mais atingidos. E os objetivos específicos são:

- i.* Analisar o padrão da distribuição espacial de indicadores socioeconômicos selecionados e da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos em dois conjuntos de unidades geográficas de análise, setores censitários e bairros do município do Rio de Janeiro;
- ii.* Ajustar um modelo que relacione a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos, nos 153 bairros do município do Rio de Janeiro e indicadores socioeconômicos.

A mortalidade por doença isquêmica do coração foi escolhida pela sua magnitude e precocidade com que estes óbitos ocorrem no Brasil. Enquanto a idade mediana de morte por doença isquêmica do coração no Brasil era de 68 anos em 1987, na Argentina, por exemplo, era de 73 anos e nos Estados Unidos e Inglaterra de 78 anos, em 1989 (Chor, 1997). No Brasil, as doenças do aparelho circulatório constituem uma importante causa de óbito entre adultos, e não apenas entre os idosos. Tabulações de dados do Sistema de Informações de Mortalidade (SIM), indicam que estas doenças foram responsáveis por 29,70% da mortalidade na faixa etária 30-69 anos, em 1995 no Brasil.

2 - REVISÃO BIBLIOGRÁFICA

Neste capítulo uma breve revisão da literatura sobre doença isquêmica do coração é apresentada. As bases teóricas dos métodos utilizados são também discutidas. Ênfase especial é dada às técnicas de análise espacial que constituem ferramenta principal de análise deste estudo.

2.1 - Doença Isquêmica do Coração

As doenças cardiovasculares são, em todo mundo, um grupo de causas com grande peso na estrutura de morbimortalidade. Na maioria dos países desenvolvidos do Ocidente, observou-se um aumento, durante trinta anos, da mortalidade por doenças ateroscleróticas a partir da década de 40, seguida de rápida diminuição. Esta tendência foi particularmente importante nos Estados Unidos, cujas taxas de mortalidade masculinas por doença isquêmica do coração decresceram cerca de 38% desde 1970. A partir do final da década de 70, declínio semelhante passou a ser descrito na Inglaterra, no Japão, na Austrália e na América Latina. Entretanto, este declínio não vem ocorrendo em todos os países como por exemplo nos países do Leste Europeu, onde há evidências de que a mortalidade por doença isquêmica do coração está em ascensão (Beaglehole, 1990).

O aumento nas taxas de mortalidade por doença isquêmica do coração não ocorre, necessariamente, como consequência do desenvolvimento econômico, da industrialização e do declínio das doenças infecciosas. No Japão, as taxas de mortalidade por todas as causas eram consideravelmente maiores do que as do Reino Unido. Entretanto, houve uma queda brusca nestas taxas, e atualmente o

Japão apresenta taxas relativamente menores por quase todas as causas de morte (Uemura & Pisa, 1985).

Um desafio no estudo das doenças cardiovasculares, especificamente da doença isquêmica do coração é relacionar o comportamento de sua incidência e mortalidade com determinantes sociais, econômicos e culturais. As diferenças na ocorrência de doença isquêmica do coração entre países, ou mesmo dentro de países, podem ser melhor compreendidas se relacionadas às características sociais, culturais e econômicas das sociedades (Rose, 1989). São dois grupos básicos de determinantes das taxas de mortalidade por doença isquêmica do coração, aqueles relacionados ao "estilo de vida" — dieta, sedentarismo, tabagismo, etc. — e os relacionados a questões socioeconômicas. Estas em grande parte determinam o "estilo de vida", mas também incluem aspectos como o acesso aos serviços de saúde.

Em torno de 1930, quando a doença isquêmica do coração apresentava-se em ascensão nos países ocidentais desenvolvidos, esta doença era mais comum entre as classes sociais mais privilegiadas. Mais tarde, as taxas de mortalidade começaram a diminuir neste grupo, enquanto nas classes mais pobres o declínio observado é mais tardio, ou não acontece (Schenkman, 1998). É consensual que a relação entre estrato social e taxa de mortalidade por doença isquêmica do coração modificou-se ao longo da história.

Vários tipos de explicações foram sugeridas para a compreensão da relação entre as doenças do coração e os estratos sociais: cultura, "estilo de vida" e condições materiais. Citado por Castelli (1989), Marmot (1989) apresenta vários exemplos de relações entre condições socioeconômicas e incidência/mortalidade por doença isquêmica do coração. As evidências sugeriram que as condições materiais de vida, tais como a moradia, o transporte e o meio ambiente estariam fortemente correlacionadas com a renda e com as condições de saúde.

O conceito *fator de risco* teve sua origem nos estudos epidemiológicos das doenças coronarianas (Chor, 1997). A dieta "rica"¹, o colesterol sérico aumentado, a obesidade, o sedentarismo, o tabagismo e a hipertensão arterial sistêmica (HAS), passaram a ser denominados pelo termo *fatores de risco* para doenças cardiovasculares.

A questão nutricional é apontada como sendo uma das responsáveis pelo aumento das taxas de mortalidade por doença isquêmica do coração, particularmente a alta ingestão de gordura saturada e as altas concentrações de colesterol sérico. A dieta "rica", associada ao sedentarismo, aumenta o nível do colesterol sérico e os níveis da pressão arterial da população desde a infância até a vida adulta. Esta dieta é apontada como fator principal na ocorrência de doença isquêmica do coração por alguns autores (Keys, 1970).

Dentre os fatores de risco mais conhecidos para a doença isquêmica do coração, ressalta-se o consumo de cigarros. A importância deste fator, como causa de doença isquêmica do coração em indivíduos, é bastante clara, mas seu efeito na taxa de mortalidade/incidência nacional, é de difícil mensuração (Rose, 1989).

A dieta "rica" e o consumo de cigarros constituem aspectos do "estilo de vida" deste século, e são fenômenos de massa nos países industrializados do Ocidente. Estes fatores são designados "principais" por apresentarem alto impacto na doença isquêmica do coração e por serem passíveis de prevenção desde a infância e de reversão, no caso de mudanças nos hábitos nutricionais.

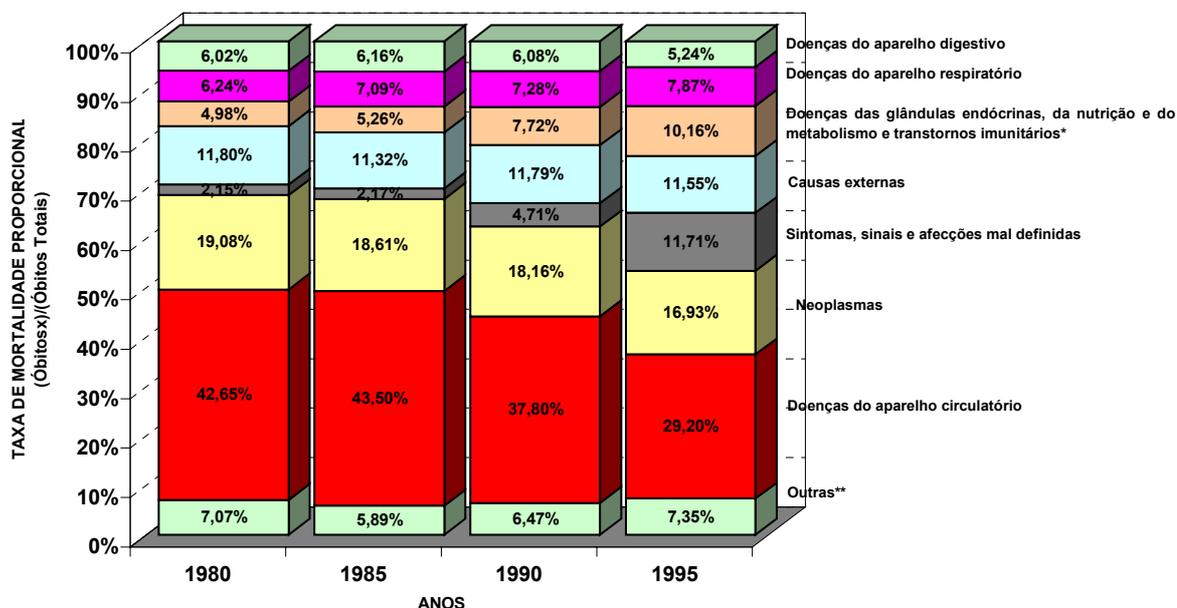
Com a transição demográfica, as doenças crônicas não transmissíveis, passaram a ter um papel fundamental na mortalidade geral dos países desenvolvidos, representando a primeira causa de óbito desde a primeira metade do século. No Brasil, a mortalidade geral vem decrescendo. Com o controle das doenças

¹ A dieta "rica" é caracterizada por altas proporções de produtos de origem animal, de produtos processados comercialmente com alto percentual de gordura total, de gordura hidrogenada, de colesterol e de gordura saturada; altas proporções de açúcar refinado, de sal e de álcool; alta densidade calórica; razão relativamente alta entre calorias e nutrientes essenciais; baixo teor de potássio e de fibras (Stamler, 1992).

infecciosas estas deixaram de ser a causa mais freqüente de óbito, e as de origem aterosclerótica tornaram-se um dos "males do século" (Chor, 1997).

Em 1995, dos óbitos registrados e com causa básica definida, as doenças cardiovasculares determinaram, cerca de um terço do total de mortes, constituindo a primeira causa de óbito na população de 30 a 69 anos de idade. No Gráfico 2.1 temos a mortalidade proporcional por algumas causas de morte no município do Rio de Janeiro em 1980, 1985, 1990 e 1995. Pode-se notar o destaque da mortalidade por doenças do aparelho circulatório em relação às outras causas.

GRÁFICO 2.1 - MORTALIDADE PROPORCIONAL POR CAPÍTULOS DA CID - FAIXA ETÁRIA 30-69 ANOS - MUNICÍPIO DO RIO DE JANEIRO - 1980/95



Fonte: SIM/SUS

* Inclui AIDS

**Outras = Doenças infecciosas e parasitárias, Doenças do aparelho geniturinário, Sistema nervoso e órgãos dos sentidos, Transtornos mentais, Doenças do sangue e dos órgãos hematopoéticos, Doenças do sistema osteomuscular e do tecido conjuntivo, Complicações da gravidez, parto e puerpério, Doenças da pele e tecido celular subcutâneo e Anomalias congênitas.

2.2 - Análise Espacial

A escala, a resolução, a homogeneidade e a estabilidade são atributos associados à unidade geográfica de análise. No Quadro 2.1 temos a relação entre estes atributos e a diminuição da unidade geográfica de análise.

QUADRO 2.1- ESCALA, RESOLUÇÃO, HOMOGENEIDADE, ESTABILIDADE E SUA RELAÇÃO COM A DIMINUIÇÃO DA ÁREA DA UNIDADE GEOGRÁFICA

TERMO	DEFINIÇÃO	
Unidade Geográfica de Análise	Menor área para a qual são analisadas as informações	▲
Escala	Razão entre as medidas do mapa e as distâncias reais	➤
Resolução	Capacidade de distinguir pontos adjacentes	➤
Homogeneidade	Características da distribuição estatística	➤
Estabilidade	Presença de flutuação aleatória	▲

Fonte: Carvalho & Cruz, 1998.

A resolução espacial é definida com base na menor área para a qual estão disponíveis as informações, e quanto maior a escala, maior a resolução, e maior capacidade de diferenciar objetos próximos.

A homogeneidade está relacionada à dispersão do atributo estudado; não é uma medida absoluta, mas um critério que permite avaliar a representatividade dos indicadores adotados para o grupo todo. Sendo assim, região homogênea é uma tipologia de unidades geográficas de informação, cujas semelhanças, em função do problema abordado, são maiores do que suas diferenças (Sanchez & Buzai).

Ao diminuir a área da unidade geográfica de análise, aumentam a escala e a resolução; a homogeneidade dos indicadores, particularmente os sociais, também tende a aumentar. Por outro lado a estabilidade dos indicadores tende a diminuir, uma vez que diminui o número de eventos que ocorrem na área de estudo.

As unidades de coleta e análise da informação devem apresentar a resolução adequada ao fenômeno que se deseja estudar. No Quadro 2.2 temos exemplos de aspectos positivos e negativos de três unidades geográficas de análise.

QUADRO 2.2 - UNIDADES GEOGRÁFICAS DE ANÁLISE: ASPECTOS POSITIVOS E NEGATIVOS

Área	Aspectos Positivos	Aspectos Negativos
Setor Censitário	<ul style="list-style-type: none"> • Menor área para a qual estão disponíveis dados censitários • Grande resolução e homogeneidade socioeconômica 	<ul style="list-style-type: none"> • Flutuação aleatória • Contorno não digitalizado • Difícil georeferenciamento^a
Bairro ou Distrito	<ul style="list-style-type: none"> • Dados censitários disponíveis • Georeferência para algumas bases • Base digital em alguns locais • Digitalização de contornos simples 	<ul style="list-style-type: none"> • Homogeneidade e estabilidade variável • Algumas bases não georeferenciadas • Contornos nem sempre claros
Município	<ul style="list-style-type: none"> • Dados censitários e estimativas intercensitárias disponíveis • Base digital disponibilizada pelo IBGE e DATASUS • Georeferenciamento já definido através do código do município para bases de dados de saúde 	<ul style="list-style-type: none"> • Homogeneidade socioeconômica pequena • Baixa resolução

^a Georeferenciamento – identificador que permite o relacionamento entre bases de dados alfa numéricas (não gráficas) e gráficas.

2.3 - Análise Exploratória Espacial

A análise exploratória de dados espaciais em saúde é feita através de técnicas de mapeamento. Neste estudo o mapeamento será utilizado como instrumento de análise exploratória e de apresentação de resultados por meio de mapas coropléticos, também chamado de temático ou de padrão. Este tipo de mapa é utilizado com frequência na comparação entre áreas, onde a divisão territorial mais usual é a geopolítica (neste estudo setores censitários e bairros). A variável que será disposta no mapa é calculada para cada uma de suas subdivisões, permitindo a divisão da região em classes, segundo algum critério, estatístico ou não. A cada classe será associado um padrão (hachura) ou cor, que preencherá a subdivisão do mapa (Nobre & Carvalho, 1994).

Quando se trabalha com dados de áreas em estudos ecológicos, o objetivo é explicar o atributo em estudo com relação a covariáveis medidas na mesma área ou no mesmo agrupamento de áreas.

Os modelos de regressão linear múltipla, em que a variável resposta (y) depende de $(p-1)$ variáveis explicativas ($x_1, x_2, x_3, \dots, x_{p-1}$), não consideram a estrutura espacial dos dados. Conforme Cressie (1991) "Independência é um pressuposto muito conveniente que faz parte da teoria estatística matemática tratável. Entretanto, modelos que envolvem dependência estatística são frequentemente mais realísticos".

Os principais conceitos estatísticos que definem a estrutura espacial dos dados relacionam-se aos efeitos de 1ª e 2ª ordem. Efeito de 1ª ordem é o valor esperado, isto é, a média do processo no espaço. Efeito de 2ª ordem é a covariância entre as áreas s_i e s_j . Um conceito importante neste tipo de estudo é o de estacionariedade. O processo é considerado estacionário se os efeitos de 1ª e 2ª ordem são constantes, em toda a região estudada, ou seja, não há tendência. Um processo é isotrópico se, além de estacionário, a covariância depende somente da distância entre os pontos e não da direção entre eles.

A estrutura de relacionamento espacial é definida pela matriz de vizinhança (W), onde cada elemento w_{ij} representa uma medida de proximidade espacial entre as áreas s_i e s_j . A escolha de w_{ij} depende do tipo de dado, da região e dos mecanismos da dependência espacial. Bailey & Gatrell (1995), definem alguns critérios usuais para a escolha dos elementos w_{ij} :

- $w_{ij} = \begin{cases} 1, & \text{se o centróide de } s_i \text{ é o mais próximo de } s_j \text{ e} \\ 0, & \text{caso contrário} \end{cases}$;
- $w_{ij} = \begin{cases} 1, & \text{se o centróide de } s_i \text{ está dentro de uma distância específica de } s_j \text{ (buffer) e} \\ 0, & \text{caso contrário} \end{cases}$;
- $w_{ij} = \begin{cases} 1, & \text{se } s_i \text{ tem fronteira comum com } s_j \text{ e} \\ 0, & \text{caso contrário} \end{cases}$;
- $w_{ij} = \frac{l_{ij}}{l_i}$, onde l_{ij} é o comprimento da fronteira comum entre s_i e s_j e l_i é o perímetro de s_i ;
- $w_{ij} = \frac{1}{d}$, onde d é a distância entre os centróides de s_i e s_j .

Diversas medidas são utilizadas na identificação da existência de padrões espaciais — tendência, cluster e/ou regularidade — entre as quais o Índice de Moran.

O Índice de Moran I é uma medida de correlação espacial ponderada por uma matriz de vizinhança, que permite testar se áreas vizinhas apresentam maior semelhança quanto ao indicador estudado do que o esperado, em um padrão de completa aleatoriedade espacial. Os valores obtidos de Moran I encontram-se entre -1 e 1, quantificando a autocorrelação existente. A hipótese² nula (H_0) refere-se à aleatoriedade espacial e ocorre quando o indicador se distribui ao acaso entre as áreas, sem relação com a posição destas. Sendo uma medida de autocorrelação, a existência de tendência nos dados faz com que os valores da estatística se afastem de zero.

² As hipóteses deste teste são: $\begin{cases} H_0) \text{ Não existe autocorrelação espacial} \\ H_1) \text{ Existe autocorrelação espacial} \end{cases}$.

O Índice de Moran é calculado da seguinte forma:

$$\text{Moran } I = \frac{N \sum_{i=1}^N \sum_{j=1}^N w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\left(\sum_{i=1}^N (y_i - \bar{y})^2 \right) \left(\sum_{i \neq j} w_{ij} \right)} \quad (2.1)$$

Onde:

N é o número de áreas;

w_{ij} é o peso atribuído de acordo com a estrutura de vizinhança;

y_i é o valor da estimativa da variável na área s_i ;

y_j é o valor da estimativa da variável na área s_j ;

\bar{y} é o valor médio do processo.

O teste de significância do Índice de Moran é feito por simulação e aproximação Normal, onde os valores do indicador são permutados entre as áreas. A cada permutação é computado o valor da estatística. Todos os índices de Moran assim calculados permitem estimar os parâmetros da distribuição. O cálculo da autocorrelação espacial assume que a variável tem média e variância constantes (estacionariedade). Se estes pressupostos forem violados, os resultados devem ser analisados com cautela.

2.4 - Métodos de Classificação Multivariada

Os métodos de classificação multivariada podem ser definidos como um conjunto de procedimentos estatísticos que visam agrupar e discriminar grupos de indivíduos, regiões ou objetos segundo um critério de semelhança. Os critérios de similaridade mais freqüentes na literatura são os coeficientes de correlação, as medidas de distância e os coeficientes de associação.

Um dos algoritmos de classificação mais utilizados é o "*K-means*". Neste, o número desejado de classes é assumido no início do procedimento, e a base de dados é

fracionada em K grupos, utilizando critérios que diminuam a variância intra-grupos e maximizem a variância inter-grupos. Este método apresenta bom desempenho quando o número de estratos é pequeno em comparação com o número de elementos.

Aldenderfer e Blashfield (1984) afirmam que apesar das diferenças quanto aos objetivos, tipos de dados e métodos usados em cada experiência, os cinco procedimentos básicos que caracterizam todos os algoritmos de classificação são os seguintes:

- Seleção de uma amostra das unidades que deverão ser agrupadas;
- Definição de um conjunto de variáveis que serão medidas, para todas as unidades da amostra;
- Cálculo das similaridades entre as unidades;
- Utilização de um método de análise de classificação para gerar grupos de unidades similares;
- Validação da estrutura de grupos resultante.

O gráfico de perfil de pertinência (Silhouette Average) é um dos instrumentos utilizados para validação da estrutura de grupos resultante do processo de classificação.

Para cada observação i classificada no grupo A, o seu perfil é calculado a partir da dissimilaridade média de i em relação às demais observações do grupo A. No algoritmo *K-means*, a dissimilaridade é calculada em função da distância Euclidiana. Sendo assim a dissimilaridade média $a(i)$ é calculada como:

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} d(i, j) \quad (2.2)$$

Onde:

$|A|$ é o número de observações no grupo A

$d(i,j) = d(j,i)$ é um elemento da matriz de dissimilaridade D ($n \times n$) que mede a distância entre as observações i e j do grupo A.

Logo, $a(i)$ é a distância média entre i e todos os demais elementos de A.

Considerando o grupo B diferente de A, calculamos a dissimilaridade da observação i classificada em A em relação a todas observações de B da seguinte forma.

$$d(i,B) = \frac{1}{|B|} \sum_{j \in B, i \neq j} d(i,j) \quad (2.3)$$

Depois de calcular as dissimilaridades ($d(i,B)$) para todos os grupos $B \neq A$ escolhemos a menor delas, ou seja, aquela referente ao grupo que seja o mais próximo de i , exceto por A. Esta medida, designada por $b(i)$ refere-se a distância média entre a observação i em A, e cada ponto do grupo mais semelhante a i , ou seja:

$$b(i) = \min_{B \neq A} [d(i,B)] \quad (2.4)$$

Com base nestas estatísticas, define-se o perfil médio da observação i como a razão entre a diferença da distância média do grupo ao qual pertence em relação ao grupo mais semelhante, e o valor máximo destas distâncias. Assim, p_i é calculado da seguinte forma:

$$p_i = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (2.5)$$

Observações que assumem p_i próximo a zero encontram-se entre dois grupos, as que assumem valores negativos provavelmente estão mal classificadas, e as que assumem valores próximos a 1 (um) estão muito bem classificadas. Assim, com o valor do perfil calculado de cada observação (p_i) é feito o gráfico denominado de Silhueta. Neste para cada observação i desenha-se uma barra de comprimento igual a p_i . As observações são agrupadas pelo perfil, do grupo 1 ao k . O indicador Perfil Médio é a média dos p_i de todas as observações. Quanto maior o valor do Perfil Médio, melhor o resultado da classificação. A classificação é ótima quando o gráfico

apresenta k grupos muito bem distintos, com poucos valores negativos e próximos de zero (Struyf e cols., 1997).

2.5 - Modelos de Regressão

Para implementar um modelo de regressão espacial faz-se necessário incorporar a estrutura de dependência espacial aos modelos de regressão linear. Considere um modelo de regressão linear múltipla:

$$y_i = b_0 + b_1x_1 + b_2x_2 + \Lambda + b_{p-1}x_{p-1} + \varepsilon_i \quad (2.6)$$

Onde:

y é o valor da variável resposta observada

$b_0, b_1, \Lambda, b_{p-1}$ são os parâmetros do modelo a serem estimados

$x_1, x_2, \Lambda, x_{p-1}$ são vetores de constantes observadas

ε_i é o erro associado ao modelo

Os pressupostos básicos deste modelo são:

$E(\varepsilon_i) = 0$, a média do erro é zero;

$V(\varepsilon_i) = \sigma_{\varepsilon_i}^2$, a variância do erro é constante; e

$Cov(\varepsilon_i, \varepsilon_j) = 0$, as observações não são correlacionadas

Para que testes estatísticos sejam implementados, é comum assumir que os erros são Normalmente distribuídos isto é: $\varepsilon_i \sim N(0, \sigma^2)$.

Assumindo $x_0 = 1$, podemos escrever:

$$y_i = b_0x_0 + b_1x_1 + b_2x_2 + \Lambda + b_{p-1}x_{p-1} + \varepsilon_i \quad (2.7)$$

Em termos matriciais temos:

$$Y = X\beta + \xi$$

Onde:

Y é a Variável resposta;

X é o Vetor de variáveis constantes;

β é o Vetor dos coeficientes de regressão, associados a X ;

ξ é o Vetor de erro, normalmente distribuído com média zero e matriz de covariância

Σ , isto é, $\xi \sim N(0, \Sigma)$, onde $\Sigma = \sigma^2 I$.

No modelo autorregressivo temporal o valor da estimativa da variável no instante t é influenciado pelo valor da variável no instante $t-1$. Já no modelo autoregressivo espacial (Spatial Autoregressive — SAR) o efeito da variável explicativa na área s_i , não pode ser limitado a esta área. Sendo assim, o valor da variável resposta y está sujeita aos valores da variável explicativa x na área s_i e em suas áreas vizinhas. Um modelo com variáveis explicativas espaciais em diferentes intervalos (lags) seria:

$$y_i = b_0 + b_1x_{i,1} + b_2x_{i,2} + \Lambda + b_{p-1}x_{i,p-1} + \tau \sum_{i \in W(i)} x_{i,p} + \varepsilon_i \quad (2.8)$$

Onde:

y_i é o valor da variável resposta na i -ésima área geográfica;

$b_0, b_1, \Lambda, b_{p-1}$ são os parâmetros do modelo;

$x_{i,1}, x_{i,2}, \Lambda, x_{i,p-1}$ são variáveis (constantes) observadas na i -ésima área geográfica;

$x_{i,p}$ é uma variável pertencente ao conjunto $[x_{i,1}, x_{i,2}, \Lambda, x_{i,p-1}]$;

$W(i)$ é um conjunto de áreas adjacentes i (excluindo a área i);

τ é um parâmetro de associação entre o valor da variável explicativa nas áreas vizinhas e a variável resposta;

ε_i é o erro associado a i -ésima área geográfica.

Uma modificação do modelo anterior leva em conta que o valor da variável resposta Y na área i pode ser autocorrelacionado:

$$y_i = b_0 + b_1x_{i,1} + b_2x_{i,2} + \Lambda + b_{p-1}x_{i,p-1} + \rho \sum_{i \in W(i)} y_i + \varepsilon_i \quad (2.9)$$

Onde:

ρ é um parâmetro de associação entre o valor da variável resposta nas áreas vizinhas e a própria variável resposta.

A diferença entre os modelos apresentados nas equações (2.8) e (2.9) é que no modelo (2.8) a variável dependente é correlacionada às variáveis independentes. E no modelo (2.9) a variável dependente é autocorrelacionada.

À medida que a complexidade das relações espaciais aumenta, também cresce o grau de dificuldade para estimar os parâmetros do modelo. Neste caso a alternativa é estimar um modelo que assume erros espacialmente correlacionados (Conditional Autoregressive — CAR).

Assim, a matriz de covariância Σ , assume a seguinte forma: $\Sigma = \sigma^2 \mathbf{V}$, onde $\mathbf{V} = \{v_{ij}\}$ é uma matriz simétrica que descreve a dependência espacial dos erros; v_{ij} denota a covariância entre y_i e y_j e v_{ii} denota a variância de y_i (Haining, 1990).

Para avaliarmos a qualidade do ajuste de um modelo de regressão múltipla utiliza-se em geral as seguintes estatísticas:

- Coeficiente de determinação (R^2) – medida de associação entre a variável dependente e as variáveis explicativas. O R^2 é calculado através da razão entre a Soma de Quadrados atribuída ao modelo de regressão e a Soma de Quadrados Totais da ANOVA $R^2 = SQR/SQTO$. A estatística R^2 assume valores entre [0 e 1], logo quanto mais próximo de 1 o R^2 estiver maior a proporção da variação total explicada pelo modelo;

- Teste de hipótese (estatística t) – testa se o parâmetro ajustado é estatisticamente significativo. As hipóteses do teste são: $\begin{cases} H_0) \beta_k = 0; \\ H_1) \beta_k \neq 0; \end{cases}$
- Teste de hipótese (estatística F) – teste auxiliar da qualidade do ajuste — a estatística F e sua significância são componentes da ANOVA. As hipóteses do teste são: $\begin{cases} H_0) \beta_1 = \Lambda = \beta_{p-1} = 0; \\ H_1) \exists k \text{ tal que } \beta_k \neq 0 \therefore (k = 1, \Lambda, p - 1) \end{cases}$
- Análise de resíduos – realizada através da análise do gráfico de probabilidade Normal, Histograma e gráficos que apresentam a dispersão dos resíduos do ajuste em relação ao valor ajustado da variável dependente. Este último deve ser analisado visando identificar padrões da dispersão dos resíduos (forma de cone, parábola, linha, etc.).

No caso dos modelos de regressão espacial a qualidade do ajuste é verificada através de ferramentas semelhantes, não sendo possível, entretanto, estimar o coeficiente de determinação. A dependência espacial dos resíduos é avaliada com base no Índice de Moran e em mapas coropléticos.

A presença de tendência espacial pode ser estimada através de modelos aditivos. Nestes a relação entre a variável resposta e a variável independente é medida por uma função não linear, usualmente um alisamento não paramétrico como o *Loess* ou *Spline*. Para detectar tendência utiliza-se como variável resposta cada um dos indicadores e como variáveis independentes a localização das duas coordenadas utilizando a função *Loess* como ligação: $y_i = \text{Loess}(\text{Latitude}) + \text{Loess}(\text{Longitude})$.

O *Loess* é uma regressão linear onde o peso das observações diminui à medida em que se afasta do ponto estimado, sendo por isso uma regressão local ponderada (Hastie & Tibshirani, 1990).

3 - DADOS E METODOLOGIA

Este é um estudo ecológico utilizando dois conjuntos de unidades geográficas de análise, bairros e setores censitários. Na primeira parte deste capítulo descreve-se as fontes de dados e as variáveis utilizadas nas análises. Procede-se então com diagnóstico da localização dos óbitos e com uma descrição sucinta da metodologia a ser aplicada na análise dos dados.

3.1 - Indicadores Socioeconômicos

Foram utilizados indicadores socioeconômicos e demográficos, construídos a partir das variáveis do questionário global do censo demográfico de 1991 e consolidados para setores censitários e bairros que compõem o município do Rio de Janeiro. O setor censitário é definido em função das necessidades operacionais do censo, uma vez que corresponde à área que pode ser percorrida por um recenseador durante o período de coleta de dados (FIBGE, 1993). Os setores censitários são classificados segundo a situação em urbanos e rurais. No município do Rio de Janeiro todos os setores censitários são considerados urbanos. Quanto ao tipo, os setores censitários podem ser (FIBGE, 1990):

- **COLETIVOS** – setores censitários exclusivamente de habitações coletivas, como asilos, creches, quartéis, presídios e delegacias, que serão excluídos da análise, pois são em número reduzido (aproximadamente 50 setores censitários) e distorcem os indicadores.
- **NORMAIS** – setores censitários onde há estrutura urbana tradicional, com ruas, endereços, ainda que a propriedade do terreno eventualmente não seja bem definida.

- **SUB-NORMAIS** – A definição destes setores censitários é pouco clara: "(...) conjunto constituído por unidades habitacionais ocupando ou tendo ocupado até período recente terrenos de propriedade alheia, dispostos, em geral, de forma desordenada e densa, e carentes, em sua maioria, de serviços públicos essenciais." Assim, em setores censitários ditos "normais", eventualmente podem ser encontrados domicílios precários, e em setores censitários considerados pelo censo "sub-normais" o tipo de moradia e situação de urbanização pode ser "normal". Para fins deste estudo, consideramos os setores censitários "sub-normais" como sendo favelas.

Quanto à característica do domicílio, a primeira subdivisão separa os domicílios em **particulares** e **coletivos**. Os particulares dividem-se em permanentes e improvisados. A pesquisa das demais características dos domicílios limitou-se aos domicílios particulares permanentes, e pelo mesmo motivo os indicadores criados têm como denominador os domicílios particulares permanentes.

Os indicadores apresentados a seguir foram selecionados a partir de estudos anteriores (Carvalho, 1997), consolidados para setores censitários e bairros:

- **Proporção de casas:** número de casas (construção unifamiliar) em relação ao total de domicílios particulares ou coletivos, permanentes ou improvisados;
- **Proporção de domicílios alugados:** número de domicílios alugados em relação ao total de domicílios particulares permanentes;
- **Proporção de domicílios ligados à rede geral de água:** número de domicílios ligados à rede geral de água em relação ao total de domicílios particulares permanentes;
- **Proporção de domicílios ligados à rede de esgoto:** número de domicílios ligados à rede de esgoto em relação ao total de domicílios particulares permanentes;
- **Proporção de domicílios em que o lixo é coletado:** número de domicílios em que o lixo é coletado regularmente em relação ao total de domicílios particulares permanentes;

- **Proporção de chefes de domicílio com 1º grau completo ou mais:** número de chefes de domicílio com primeiro grau completo ou mais em relação ao total de chefes de domicílio;
- **Proporção de população alfabetizada:** número de residentes acima de 5 anos que sabem ler e escrever em relação à população total acima de 5 anos;
- **Proporção de chefes de domicílio com renda inferior a dois salários mínimos:** número de chefes de domicílio com renda inferior a dois salários mínimos em relação ao total de chefes de domicílio;
- **Proporção de chefes de domicílio com renda superior a quinze salários mínimos:** número de chefes de domicílio com renda superior a quinze salários mínimos em relação ao total de chefes de domicílio;
- **Proporção setores de favelados³:** número de setores censitários sub-normais ou favelados em relação ao total de setores censitários de cada bairro;
- **Proporção de população idosa⁴:** população da faixa etária 60 a 70 anos em relação a população de 30 a 70 anos, dos bairros do município do Rio de Janeiro.

3.2 - Localização Geográfica

A base cartográfica utilizada foi cedida pelo Laboratório de Geoprocessamento da Fundação Oswaldo Cruz – SIG/FIOCRUZ. No Mapa 3.1 estão descritos 153 bairros que compõem a cidade, resultantes da agregação dos 6.259 setores censitários.

³ Este indicador foi calculado apenas para bairros.

⁴ Este indicador foi calculado apenas para bairros.

MAPA 3.1 - DIVISÃO OFICIAL DO RIO DE JANEIRO

Prefeitura da Cidade do Rio de Janeiro
 Secretaria Municipal de Urbanismo - SMU
 Empresa Municipal de Informática e Planejamento S/A - IplanRIO

Município do Rio de Janeiro Bairros

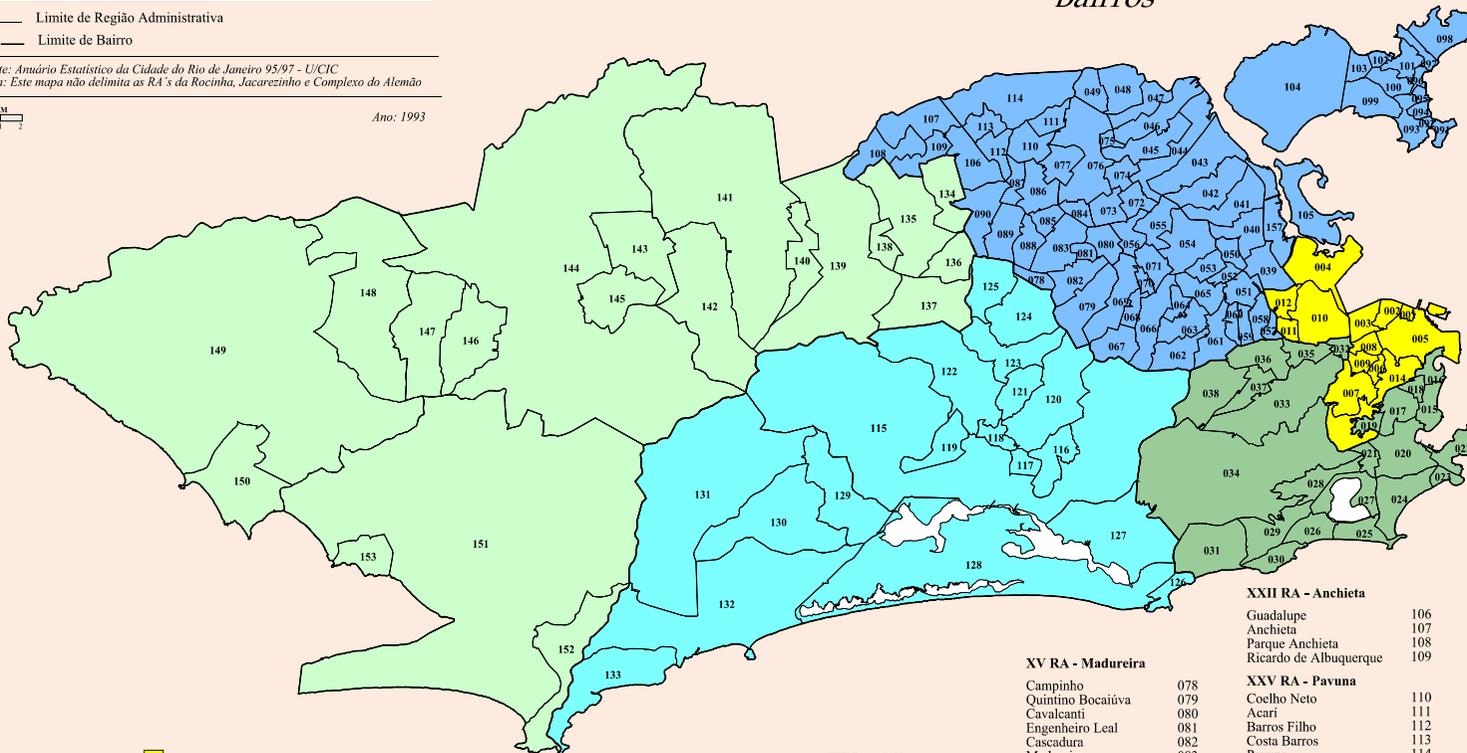


— Limite de Região Administrativa
 — Limite de Bairro

Fonte: Anuário Estatístico da Cidade do Rio de Janeiro 95/97 - U/CIC
 Nota: Este mapa não delimita as RA's da Rocinha, Jacarezinho e Complexo do Alemão



Ano: 1993



AP - 1 ■

I RA - Portuária

Saúde	001
Gamboa	002
Santo Cristo	003
Caju	004

II RA - Centro

Centro	005
--------	-----

III RA - Rio Comprido

Catumbi	006
Rio Comprido	007
Cidade Nova	008
Estácio	009

VII RA - São Cristóvão

São Cristóvão	010
Mangueira	011
Benfica	012

XXI RA - Paquetá

Paquetá	013
---------	-----

XXIII RA - SantaTeresa

Santa Teresa	014
--------------	-----

AP - 2 ■

IV RA - Botafogo

Flamengo	015
Glória	016
Laranjeiras	017
Catete	018
Cosme Velho	019
Centro	020
Humaitá	021
Urca	022

V RA - Copacabana

Leme	023
Copacabana	024

VI RA - Lagoa

Ipanema	025
Leblon	026
Lagoa	027
Jardim Botânico	028
Gávea	029
Vidigal	030
São Conrado	031

VIII RA - Tijuca

Praça da Bandeira	032
Tijuca	033
Alto da Boa Vista	034

IX RA - Vila Isabel

Maracanã	035
Vila Isabel	036
Andaraí	037
Grajaú	038

XXVII RA - Rocinha

Rocinha	154
---------	-----

AP - 3 ■

X RA - Ramos

Manguinhos	039
Bonsucesso	040
Ramos	041
Olaria	042

XI RA - Penha

Penha	043
Penha Circular	044
Brás de Pina	045
Cordovil	046
Parada de Lucas	047
Vigário Geral	048
Jardim América	049

XII RA - Inhaúma

Higienópolis	050
Jacaré	051
Maria da Graça	052
Del Castilho	053
Inhaúma	054
Engenho da Rainha	055
Tomás Coelho	056

XIII RA - Méier

São Francisco Xavier	057
Rocha	058
Riachuelo	059
Sampaio	060
Engenho Novo	061
Lins de Vasconcelos	062
Méier	063
Todos os Santos	064
Cachambi	065
Engenho de Dentro	066
Água Santa	067
Encantado	068
Piedade	069
Abolição	070
Pilares	071

XIV RA - Irajá

Vila Cosmos	072
Vicente de Carvalho	073
Vila da Penha	074
Portuguesa	075
Irajá	076
Colégio	077

XV RA - Madureira

Campinho	078
Quintino Bocaiuva	079
Cavalcanti	080
Engenheiro Leal	081
Cascadura	082
Madureira	083
Vaz Lobo	084
Turiaçu	085
Rocha Miranda	086
Honório Gurgel	087
Oswaldo Cruz	088
Bento Ribeiro	089
Marechal Hermes	090

XX RA - Ilha do Governador

Ribeira	091
Zumbi	092
Cacua	093
Pitangueiras	094
Praia da Bandeira	095
Cocotá	096
Bancários	097
Freguesia	098
Jardim Guanabara	099
Jardim Carioca	100
Taupá	101
Moneró	102
Vista Alegre	103
Galeão	104
Cidade Universitária	105

XXII RA - Anchieta

Guadalupe	106
Anchieta	107
Parque Anchieta	108
Ricardo de Albuquerque	109

XXV RA - Pavuna

Coelho Neto	110
Acari	111
Barros Filho	112
Costa Barros	113
Pavuna	114

XXVIII RA - Jacarezinho

Jacarezinho	155
-------------	-----

XXIX RA - Complexo do Alemão

Complexo do Alemão	156
--------------------	-----

XXX RA - Maré

Maré	157
------	-----

AP 4 ■

XVI RA - Jacarepaguá

Jacarepaguá	115
Anil	116
Gardênia Azul	117
Cidade de Deus	118
Curicica	119
Freguesia	120
Pechincha	121
Taquara	122
Tanque	123
Praça Seca	124
Vila Valqueire	125

XXIV RA - Barra da Tijuca

Joá	126
Itanhangá	127
Barra da Tijuca	128
Camorim	129
Vargem Pequena	130
Vargem Grande	131
Recreio dos Bandeirantes	132
Grumari	133

AP - 5 ■

XVII RA - Bangu

Deodoro	134
Vila Militar	135
Campo dos Afonsos	136
Jardim Sulacap	137
Magalhães Bastos	138
Realengo	139
Padre Miguel	140
Bangu	141
Senador Camará	142

XVIII RA - Campo Grande

Santíssimo	143
Campo Grande	144
Senador Vasconcelos	145
Inhoaíba	146
Cosmos	147

XIX RA - Santa Cruz

Paciência	148
Santa Cruz	149
Sepetiba	150

XXXI RA - Guaratiba

Guaratiba	151
Barra de Guaratiba	152
Pedra de Guaratiba	153

Fonte: Anuário Estatístico da Cidade do Rio de Janeiro, 95-97.

Os dados de endereço das declarações de óbito, apesar de estarem contidos no documento original, não são repassados ao Sistema de Informações de Mortalidade (SIM) na sua face informatizada. Por esse motivo, fez-se um convênio de cooperação técnica entre FIOCRUZ e a Secretaria Estadual de Saúde, através do qual foram recuperados os dados de endereço dos óbitos na região metropolitana do Rio de Janeiro para o ano de 1991. A partir deste banco buscou-se georreferenciar dados de mortalidade para setor censitário utilizando um sistema de localização de endereços — o SISLOC — baseado no cadastro de segmentos de logradouros, por setor censitário, criado a partir das folhas de controle do trabalho do recenseador. Entretanto, conforme pode ser visto na Tabela 3.1, este procedimento permitiu localizar 73,61% dos óbitos por doença isquêmica do coração na população de 30 a 70 anos nos setores censitários do município do Rio de Janeiro, 21,58% foram localizados manualmente e ao final, apenas 4,81% do total de óbitos não foi localizado.

TABELA 3.1 - DIAGNÓSTICO DA LOCALIZAÇÃO DOS ÓBITOS POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS

	Freq.	%
Localizados Automaticamente (SISLOC)	4.257	73,61
Localizados Manualmente	1.248	21,58
Não Localizados	278	4,81
TOTAL	5.783	100,00

Fonte: Pina e cols., 1999.

A Tabela 3.2 mostra os principais erros que impediram a localização dos óbitos por doença isquêmica do coração na população de 30 a 70 anos. Note que foram encontrados 500 erros entre os 278 óbitos não localizados durante o georeferenciamento manual.

TABELA 3.2 - DIAGNÓSTICO DA LOCALIZAÇÃO DOS ÓBITOS POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS – APÓS O GEOREFERENCIAMENTO MANUAL

	Freq.	%
Endereços não localizados	175	35,00
DO's ^a que constam no SIM, mas que não constavam no banco de mortalidade com endereço	166	33,20
Endereço incompleto, sem o número de complemento	50	10,00
DO's da Ilha do Governador e Ramos que já tinham sido localizadas	36	7,20
Logradouro encontrado em mais de um bairro, com o complemento identificado em mais de um setor censitário	30	6,00
Logradouro localizado, mas o complemento não foi localizado	22	4,40
DO's que tinham o campo endereço em branco	20	4,00
Logradouro localizado num só bairro, com o complemento localizado em mais de um setor censitário	1	0,20
TOTAL	500	100,00

Fonte: Pina e cols., 1999.

^a DO's = Declarações de Óbito

3.3 - Indicadores de Mortalidade

As variáveis dos óbitos abaixo relacionadas foram extraídas do Sistema de Informações de Mortalidade (SIM) de 1991.

- Idade;
- Sexo;
- Endereço Residencial — utilizado para georeferenciar os óbitos;
- Causa Básica da Morte — somente os óbitos classificados segundo o capítulo VII da 9ª Classificação Internacional de Doenças, correspondentes aos códigos 410-414.

Duas taxas de mortalidade por doença isquêmica do coração foram calculadas com base nos óbitos ocorridos da população da faixa etária 30 a 70 anos, por setores censitários e bairros do município do Rio de Janeiro, em 1991.

- **Taxa de mortalidade por doença isquêmica do coração nos setores censitários:** óbitos por doença isquêmica do coração ocorridos entre 30-70 anos nos setores censitários do município do Rio de Janeiro em relação a população residente da faixa etária 30 a 70 anos dos setores censitários do município do Rio de Janeiro;

$$Tx_DIC_{\text{setor-RJ}}^{30-70} = \frac{\text{óbitos por DIC}_{\text{setor-RJ}}^{30 \text{ e } 70 \text{ anos}}}{\text{população}_{\text{setor-RJ}}^{30 \text{ e } 70 \text{ anos}}} \cdot 100.000 \quad (3.1)$$

- **Taxa de mortalidade por doença isquêmica do coração nos bairros:** óbitos por doença isquêmica do coração ocorridos entre 30-70 anos nos bairros do município do Rio de Janeiro em relação a população residente da faixa etária 30 a 70 anos dos bairros do município do Rio de Janeiro;

$$Tx_DIC_{\text{bairro-RJ}}^{30-70} = \frac{\text{óbitos por DIC}_{\text{bairro-RJ}}^{30 \text{ e } 70 \text{ anos}}}{\text{população}_{\text{bairro-RJ}}^{30 \text{ e } 70 \text{ anos}}} \cdot 100.000 \quad (3.2)$$

3.4 - Métodos de Análise

Para realizar a análise exploratória univariada dos indicadores em setores censitários e bairros do município do Rio de Janeiro, foram utilizadas as seguintes ferramentas: histogramas, *boxplots*, estatísticas descritivas (máximo, mínimo, média, mediana e quartis) e mapas coropléticos.

Na análise de proporções, utilizou-se escala sempre constante, com cinco faixas: 0,00-0,20, 0,20-0,40, 0,40-0,60, 0,60-0,80 e 0,80-1,00. Além disso, sobrepôs-se a malha⁵ de bairros mesmo quando a unidade de análise era setores censitários para facilitar a descrição do indicador.

Na análise bivariada foi utilizado um diagrama de dispersão entre pares de variáveis, onde podemos observar a direção das correlações entre os indicadores. Para classificar os setores censitários segundo os indicadores socioeconômicos foi

utilizado o método *K-means* obtendo-se quatro grupos distintos entre si. O Índice de Moran foi utilizado para testar a dependência espacial dos indicadores socioeconômicos e dos resíduos resultantes dos modelos de regressão ajustados. Para verificarmos se há associação entre os indicadores socioeconômicos e a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos utilizamos dois tipos de modelos de regressão: a regressão linear simples e a regressão espacial com erros autocorrelacionados (CAR). Para verificar a presença de tendência foram utilizados modelos aditivos com alisamento *Loess*.

Na unidade de análise bairros foram ajustados modelos de regressão relacionando a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos aos indicadores socioeconômicos. Na unidade de análise setores censitários, o ajuste destes modelos não é viável, pois é grande o número de setores com zero óbitos. Não tendo sido possível construir unidades geográficas baseadas em agregados de setores censitários, mais homogêneos que bairros, (Carvalho e cols., 1995) optou-se por trabalhar com setores censitários apenas na classificação multivariada, visando a interpretação dos resultados.

O modelo de regressão linear simples "final" foi encontrado utilizando-se o método de seleção de variáveis Backward. Neste método todas as variáveis independentes entram no primeiro passo, e nos passos seguintes, são removidas, uma a uma, a cada passo, segundo um critério de remoção. Neste estudo foi utilizado o seguinte critério de remoção:

- se o valor de p da estatística F for maior que 0,05 a variável é retirada do modelo;
- caso contrário a variável permanece no modelo.

⁵ Malha é o contorno digitalizado que delimita a área geográfica.

4 - RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados da análise dos dados. Estes foram analisados em duas unidades geográficas, bairros e setores censitários. Os resultados da análise exploratória univariada realizada por setor censitário estão no Anexo I, com exceção da Classificação Multivariada e do Índice de Moran.

A análise dos dados por bairros foi executada de forma que pudéssemos relacionar a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos aos indicadores socioeconômicos, e em seguida, relacionar estes resultados, aos estratos sociais obtidos através da classificação multivariada.

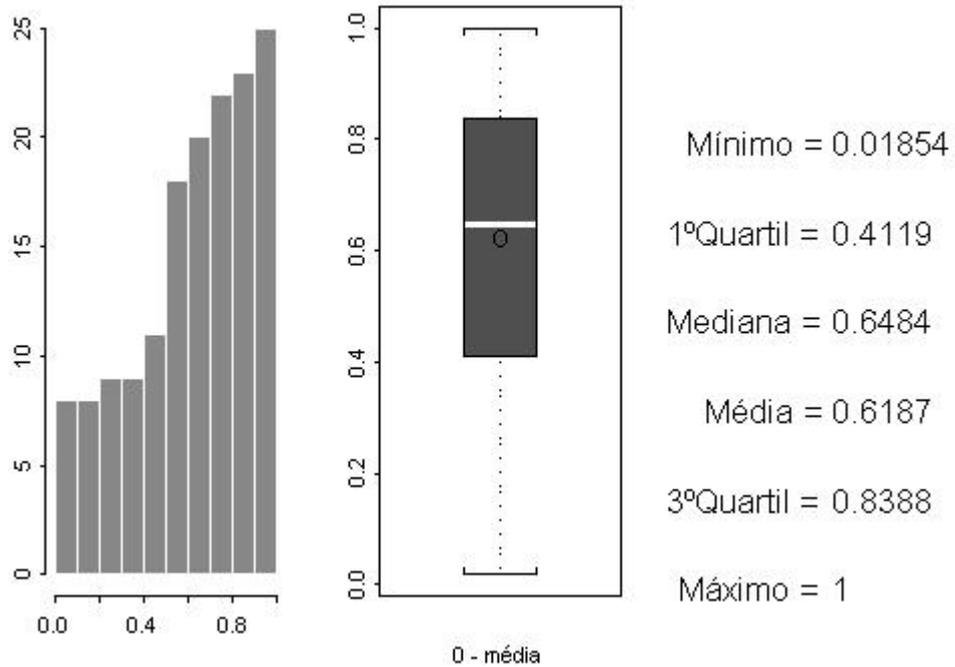
4.1 - Análise Exploratória dos Indicadores por Bairro

Nos Gráficos e Mapas 4.1 a 4.13 estão apresentados os resultados da análise exploratória univariada dos indicadores socioeconômicos e da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros.

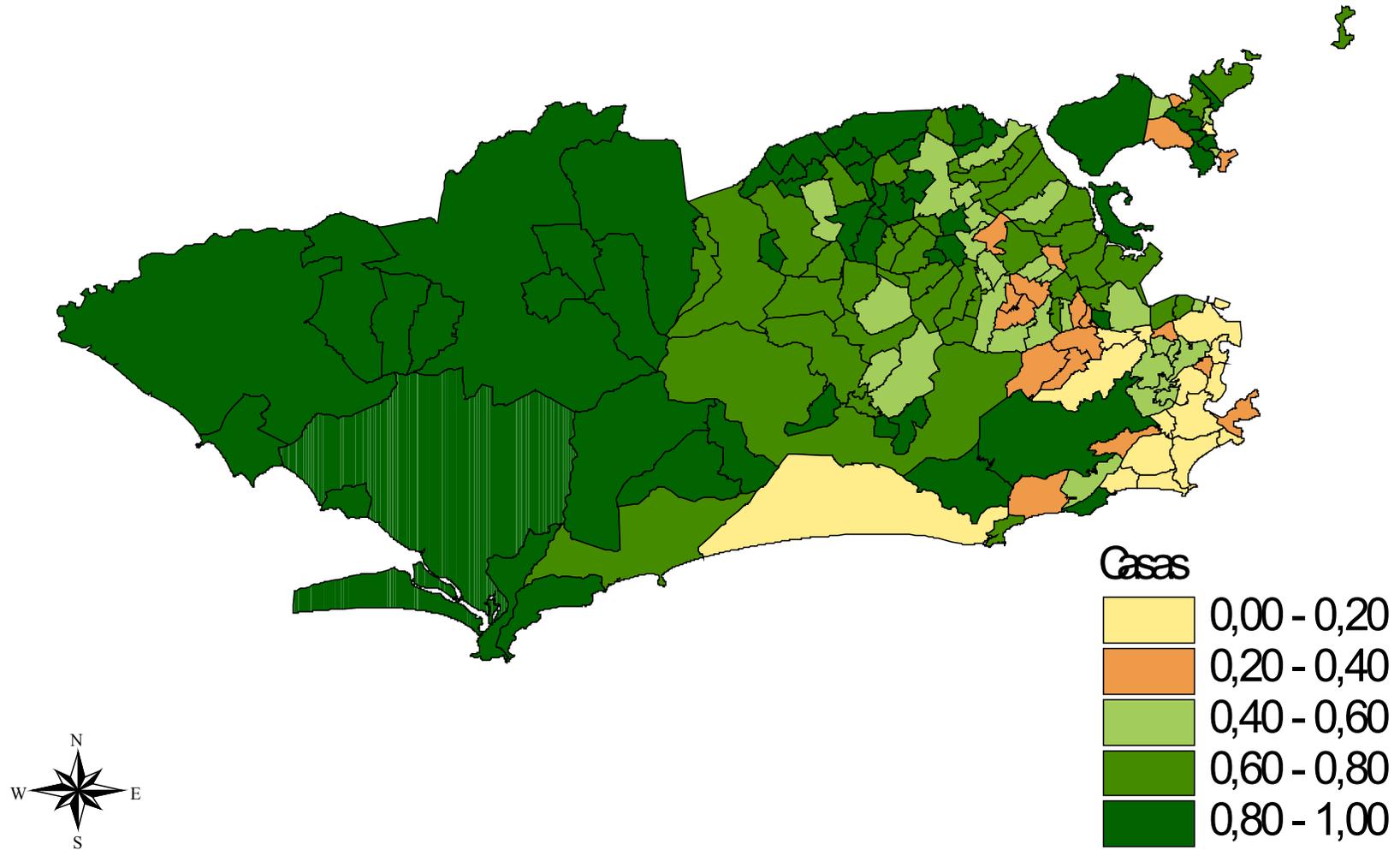
No Gráfico 4.1 observamos que as construções unifamiliares predominam nos bairros. O histograma da proporção de casas indica que a dispersão nos bairros é maior do que nos setores censitários (Gráfico e Mapa AI.1), ou seja, há uma heterogeneidade interna maior da proporção de casas nos bairros — não existem bairros só de apartamentos, mas existem bairros somente de casas (Ex.: Camorim). Através do *boxplot*, percebemos que não há evidências de valores atípicos. No

Mapa 4.1 podemos verificar que na Zona Oeste predominam as casas em relação aos outros tipos de construções.

GRÁFICO 4.1 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CASAS BAIROS DO RIO DE JANEIRO, 1991

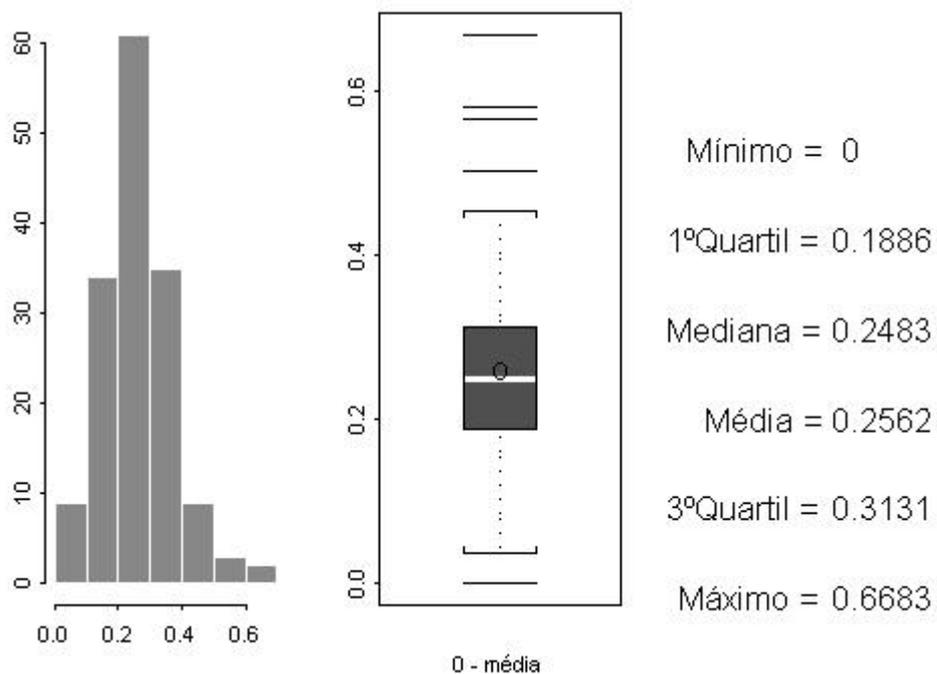


MAPA 4.1 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CASAS
BAIRROS DO RIO DE JANEIRO, 1991

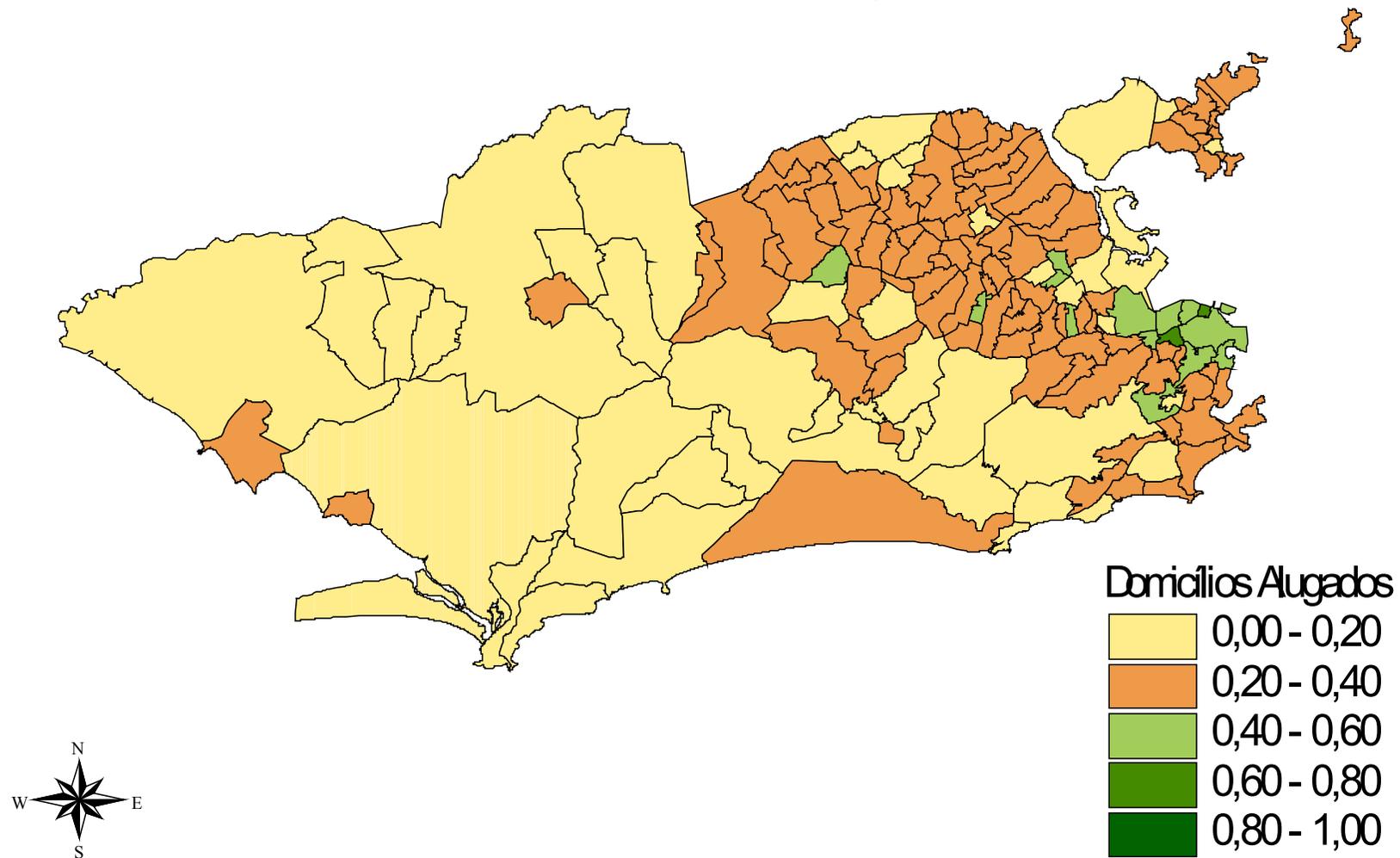


De acordo com o Gráfico e o Mapa 4.2, nota-se que a proporção de domicílios alugados em geral é baixa com média de 25,62%. Observa-se também valores altos atípicos em alguns bairros como por exemplo Saúde e Cidade Nova. Os setores censitários em que esta proporção é extremamente elevada, quando agregados com outros setores para formação dos bairros, deixam de assumir valores extremos (ver Gráfico e Mapa A1.2).

GRÁFICO 4.2 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS BAIROS DO RIO DE JANEIRO, 1991

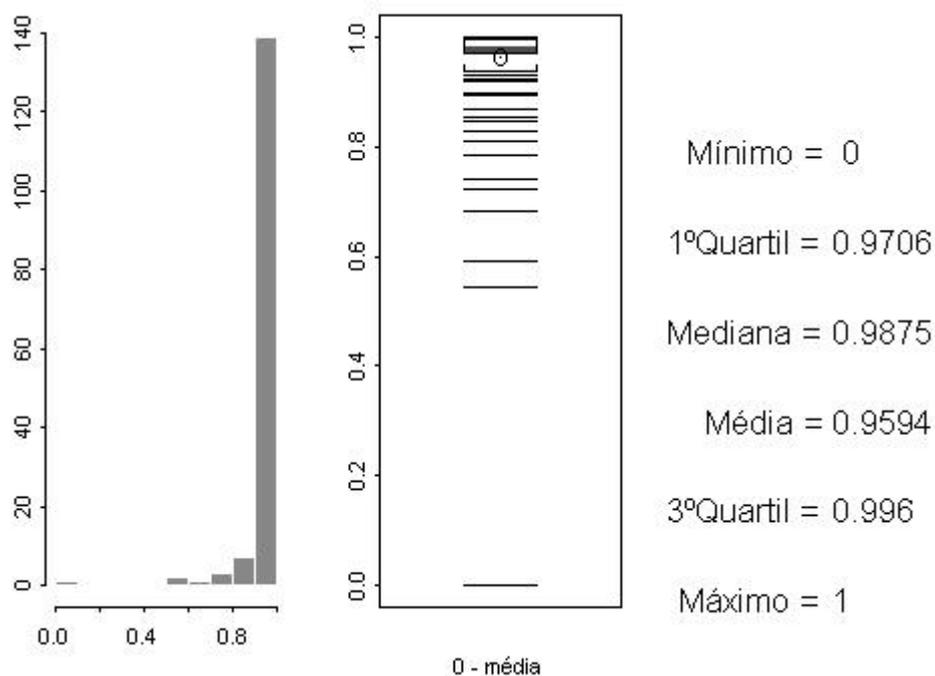


**MAPA 4.2 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS
BAIRROS DO RIO DE JANEIRO, 1991**

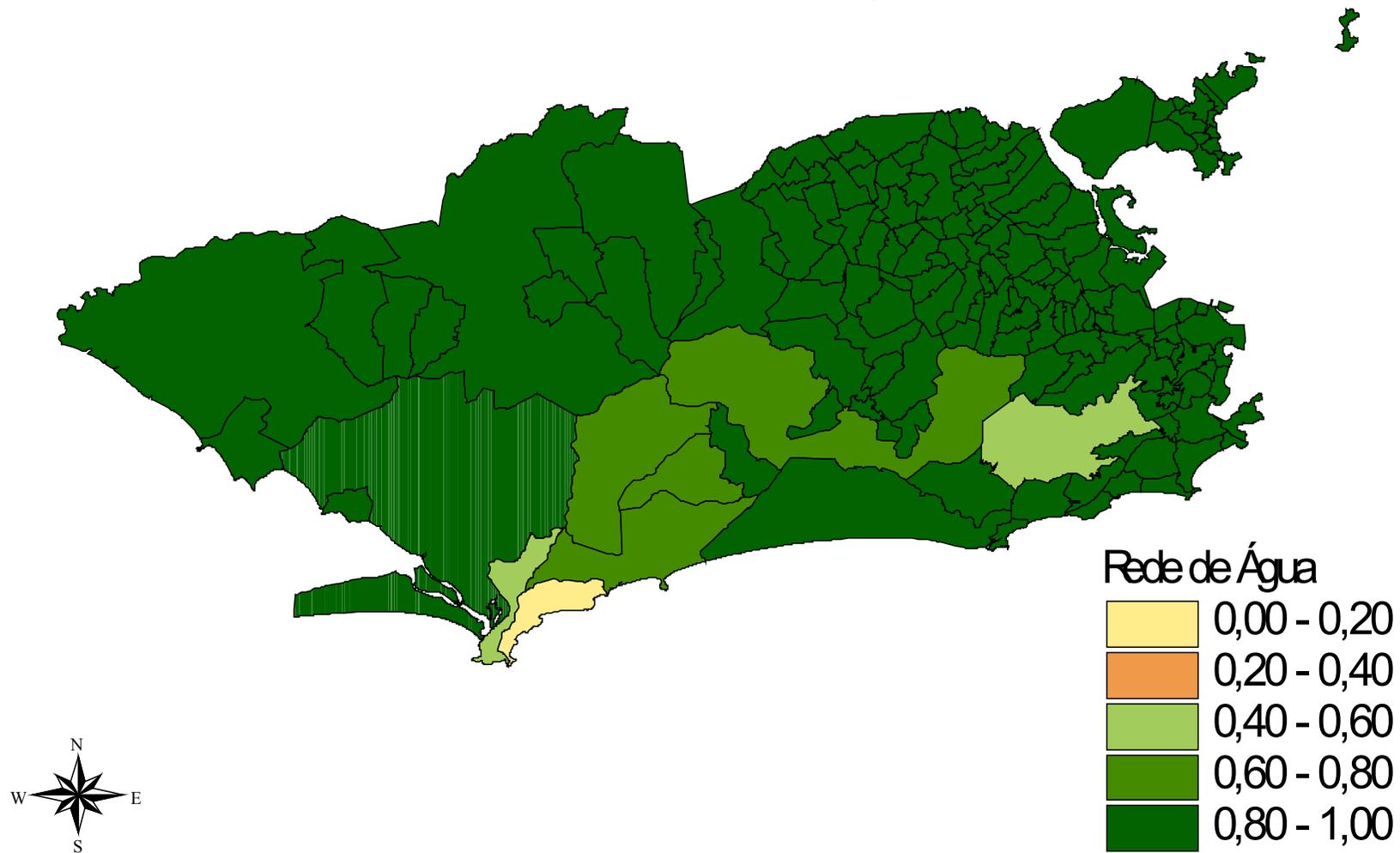


A proporção de domicílios ligados à rede geral de água tem uma distribuição concentrada nos valores altos, em média 95,94%, mas podemos perceber também alguns valores baixos indicando carência de implantação deste serviço básico de infra-estrutura em alguns bairros (Gráfico 4.3). No Mapa 4.3 percebemos que em Grumari e no Alto da Boa Vista esta proporção é baixa.

GRÁFICO 4.3 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA BAIROS DO RIO DE JANEIRO, 1991

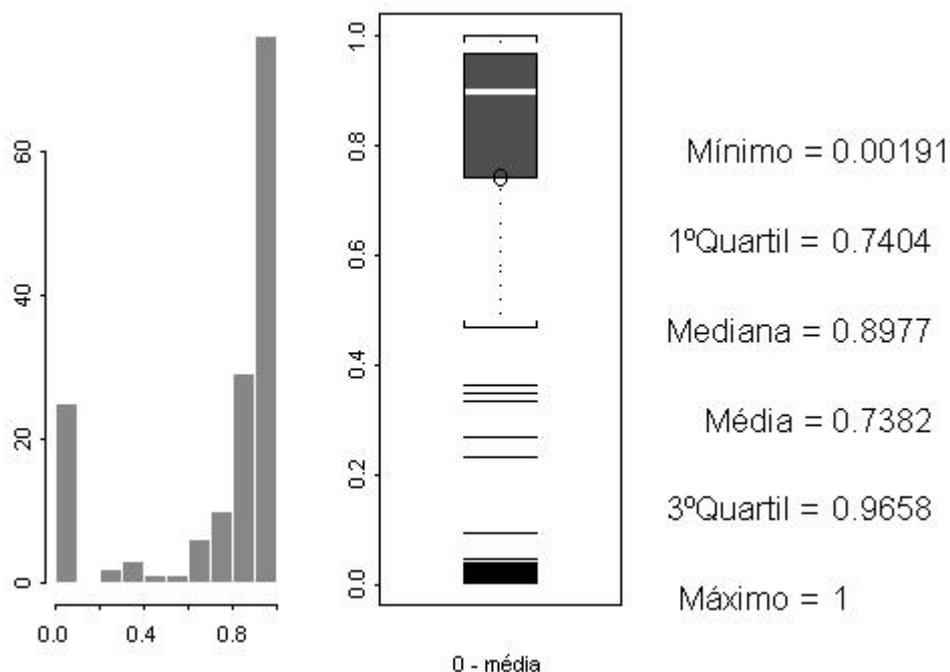


**MAPA 4.3 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA
BAIRROS DO RIO DE JANEIRO, 1991**

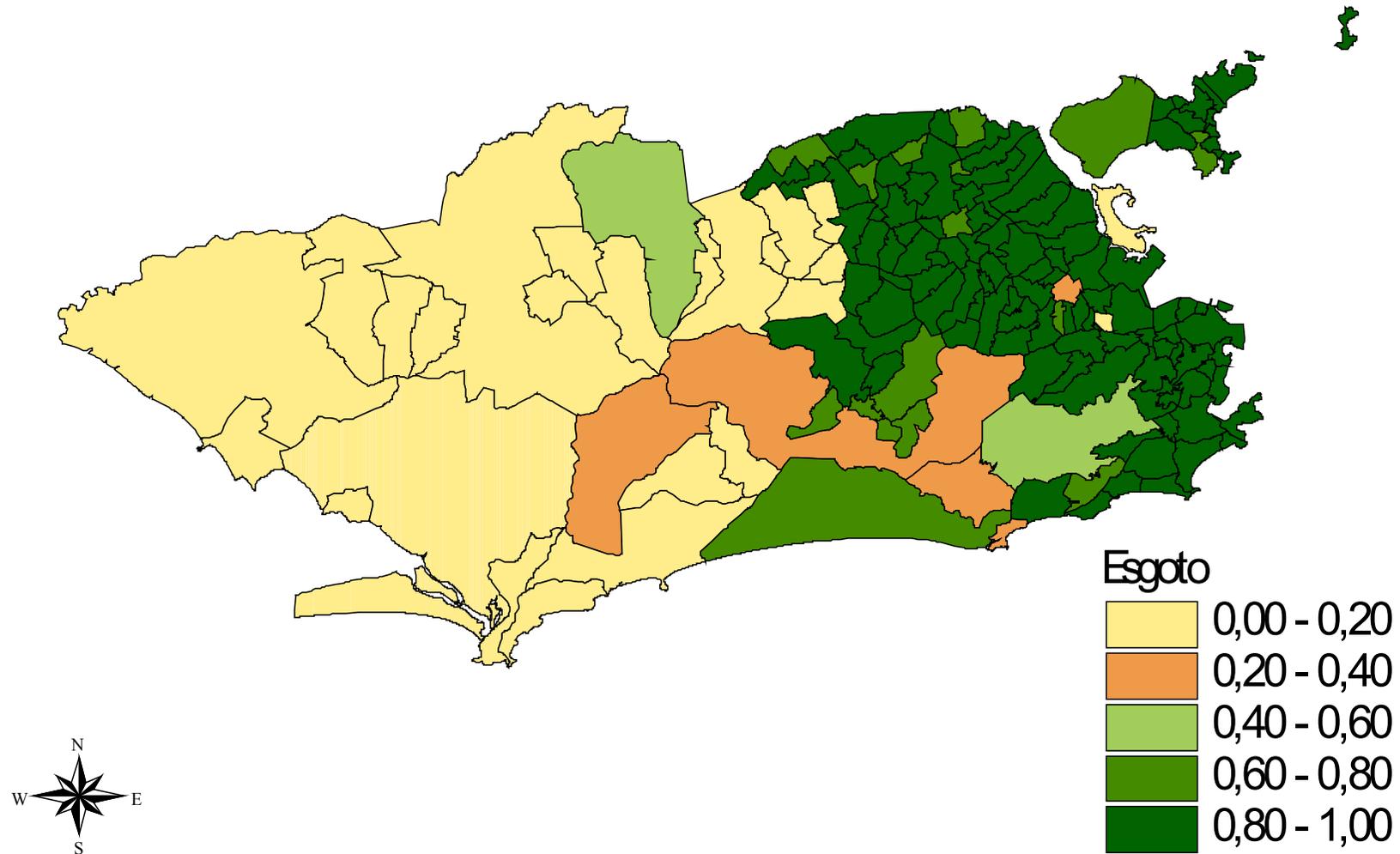


A distribuição da proporção de domicílios ligados à rede de esgoto está concentrada nos extremos da distribuição (Gráfico e Mapa 4.4). Os valores mais altos estão associados aos bairros em que grande parte dos domicílios são ligados à rede pública de esgoto; em geral concentrados nas Zonas Norte e Sul do município. Os valores mais baixos são os bairros em que poucos domicílios são ligados à rede pública de esgoto (concentrado na Zona Oeste). Podemos observar que esta proporção é alta tanto nos bairros quanto nos setores censitários (ver Gráfico e Mapa A1.4).

GRÁFICO 4.4 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE DE ESGOTO BAIROS DO RIO DE JANEIRO, 1991

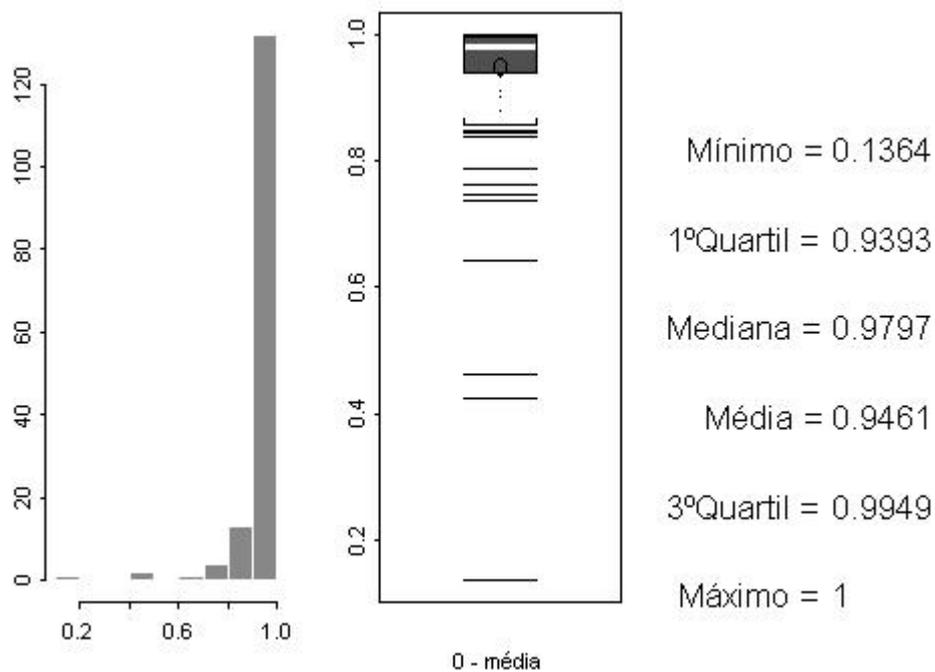


**MAPA 4.4 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE PÚBLICA DE ESGOTO
BAIRROS DO RIO DE JANEIRO, 1991**

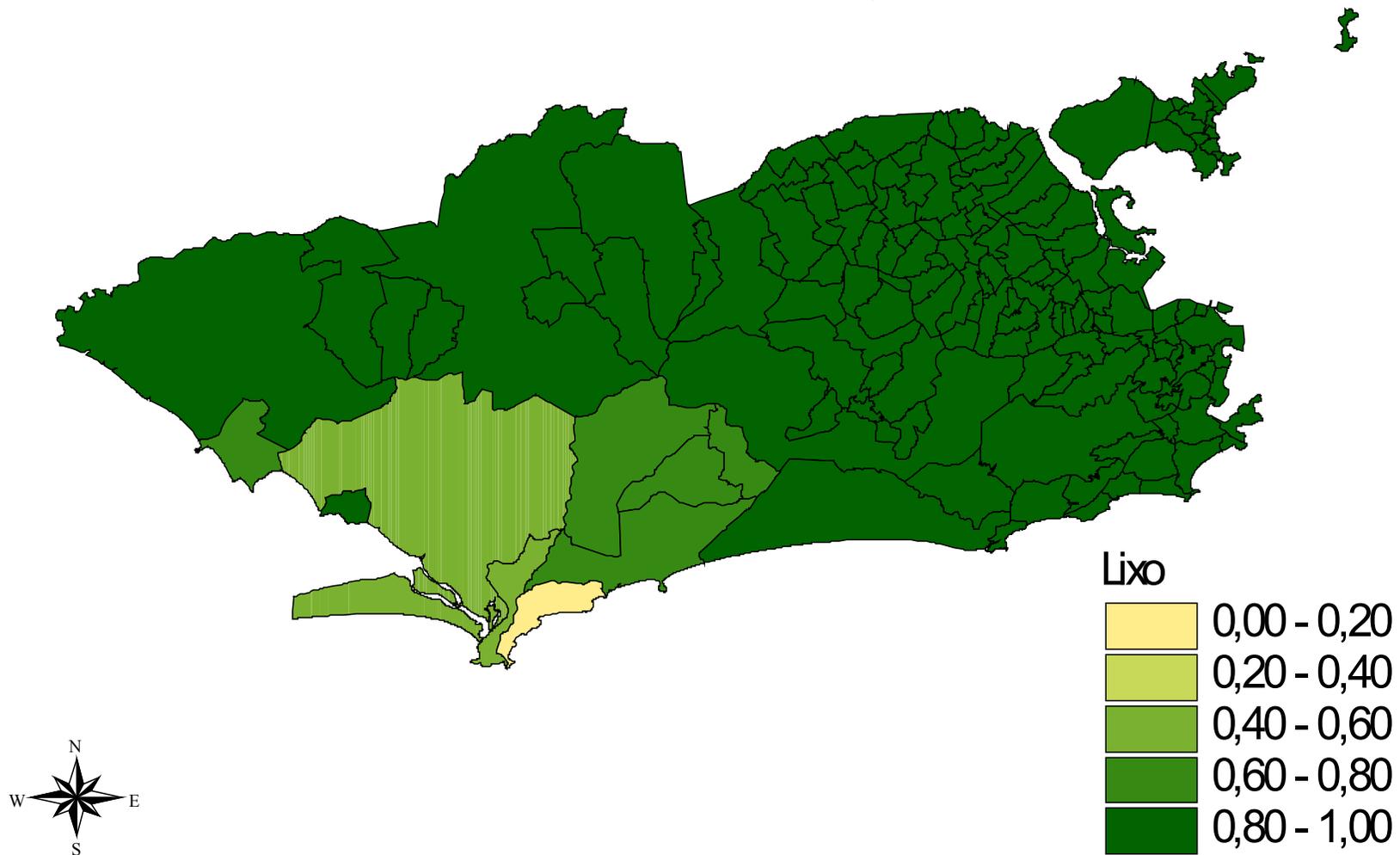


Assim como nos setores censitários, a distribuição da proporção de domicílios com coleta regular de lixo é semelhante à distribuição da variável proporção de domicílios ligados à rede geral de água, ou seja, sua distribuição é concentrada nos valores extremos altos com média de 94,61% (Gráfico 4.5). A presença de valores extremos baixos, indica que este serviço não está implantado em todos os bairros de forma homogênea. No Mapa 4.5 verificamos que em Grumari esta proporção assume o valor mais baixo de todo o município, na unidade de análise bairro.

GRÁFICO 4.5 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS EM QUE O LIXO É COLETADO BAIROS DO RIO DE JANEIRO, 1991

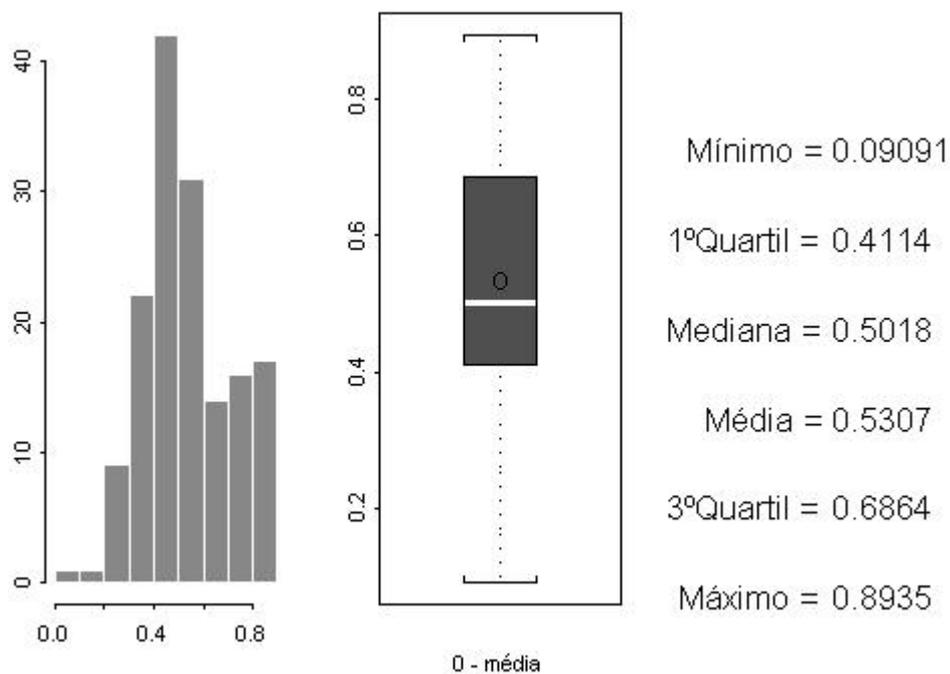


**MAPA 4.5 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS EM QUE O LIXO É COLETADO
BAIRROS DO RIO DE JANEIRO, 1991**

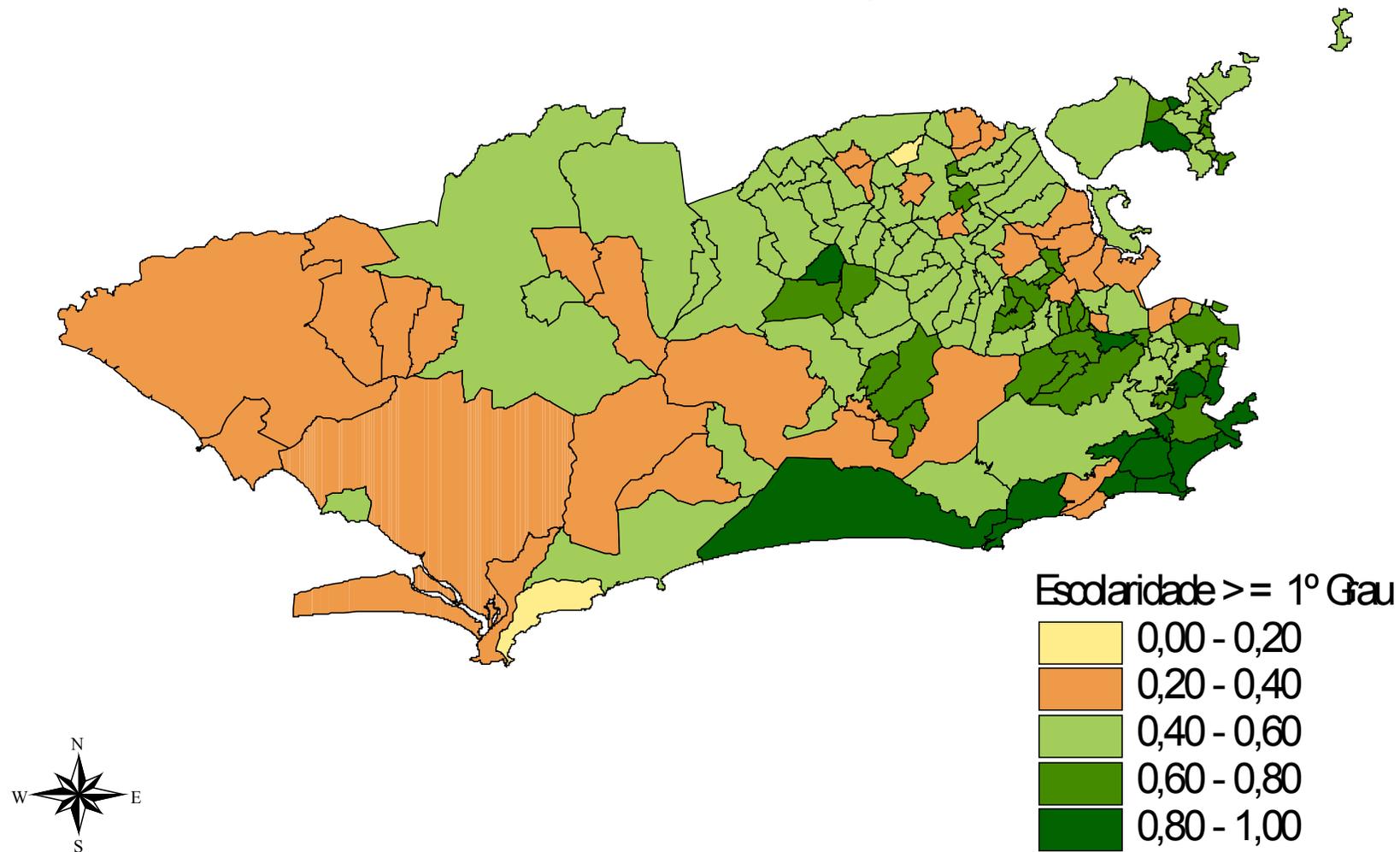


A distribuição da proporção de chefes de domicílio com pelo menos o 1º grau completo varia entre os bairros e não apresenta valores extremos atípicos (Gráfico 4.6), o mesmo ocorre nos setores censitários (Gráfico e Mapa AI.6). Através do Mapa 4.6 verificamos que em geral os valores mais elevados desta proporção estão concentrados nos bairros da Zona Sul, Barra e Ilha do Governador.

GRÁFICO 4.6 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS BAIROS DO RIO DE JANEIRO, 1991

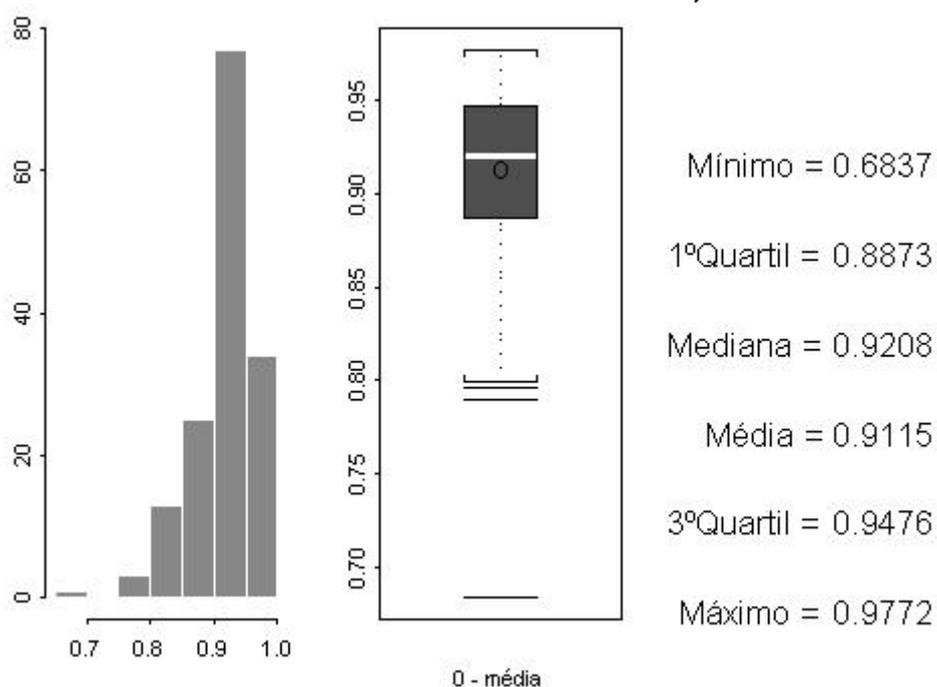


**MAPA 4.6 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS
BAIRROS DO RIO DE JANEIRO, 1991**

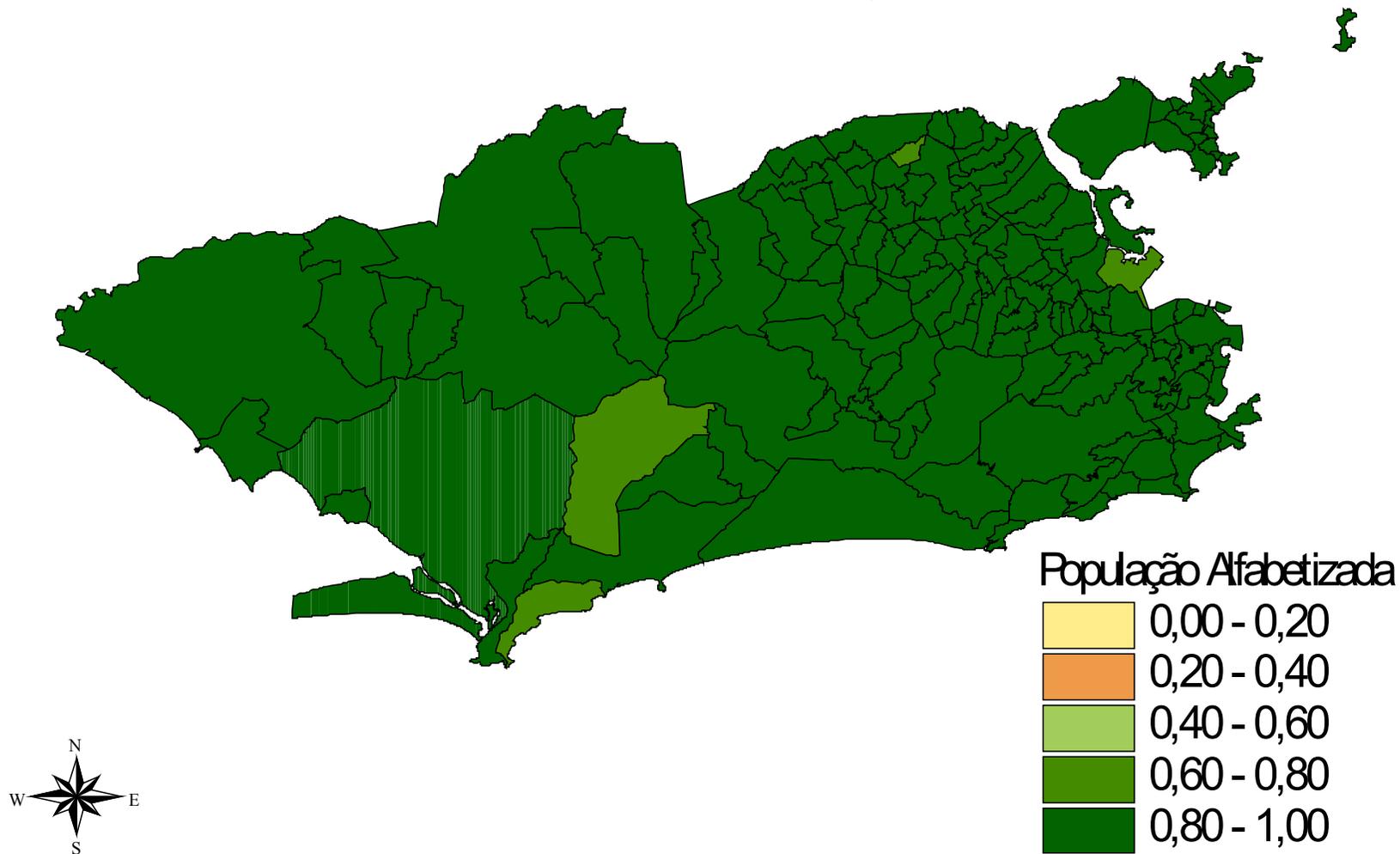


A proporção de população alfabetizada é elevada tanto em bairros quanto em setores censitários (Gráfico e Mapa A1.7). A média (91,15%) e a mediana (92,08%) indicam que apenas uma pequena parcela da população não sabe ler e escrever (Gráfico e Mapa 4.7) e os valores mais baixos correspondem aos bairros de Acari, Caju, Grumari e Vargem Grande.

GRÁFICO 4.7 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA BAIROS DO RIO DE JANEIRO, 1991

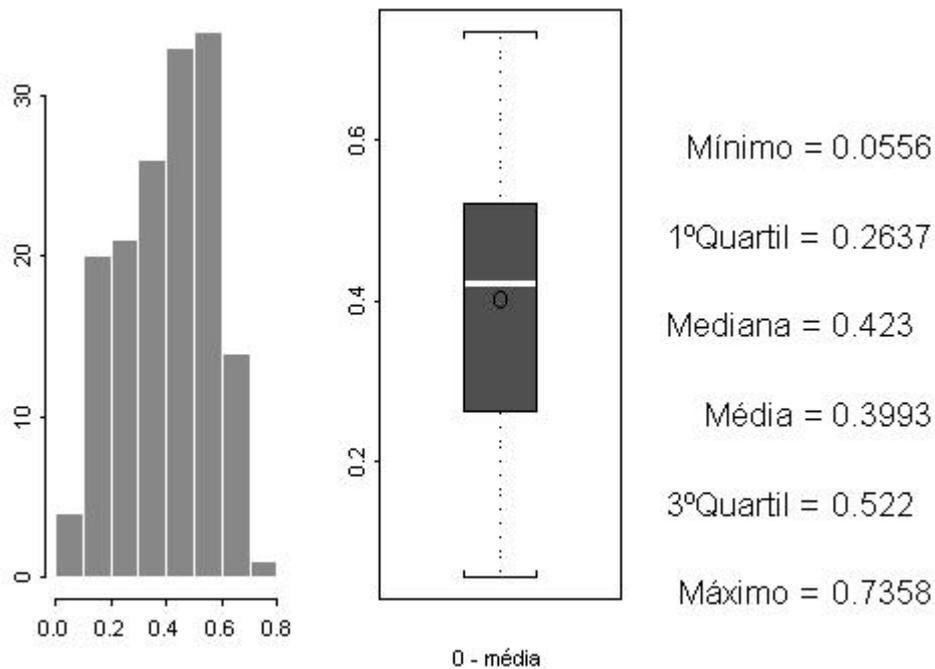


**MAPA 4.7 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA
BAIRROS DO RIO DE JANEIRO, 1991**

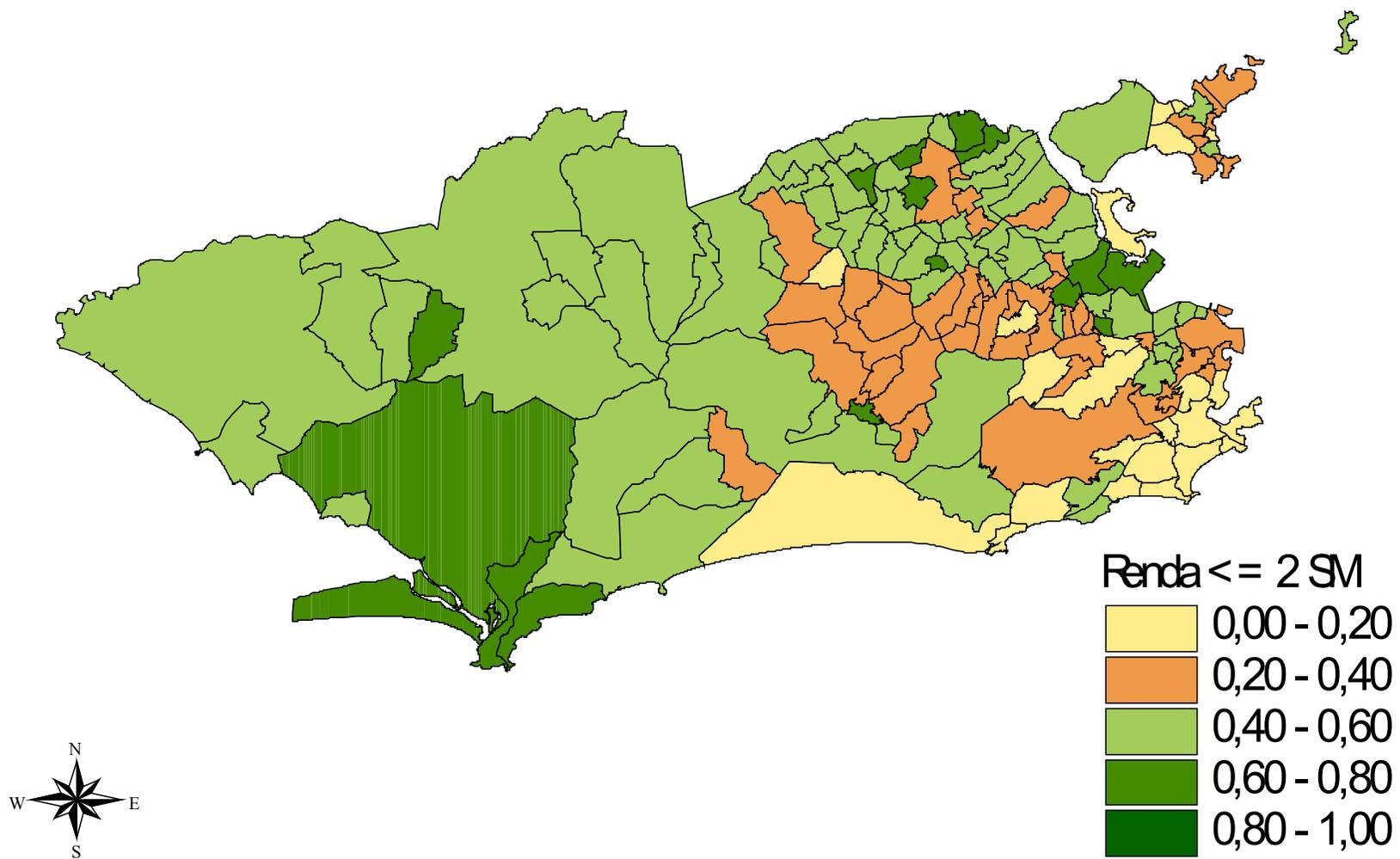


A proporção de chefes de domicílio com renda inferior a dois salários mínimos (Gráfico 4.8) varia muito entre bairros e não possui valores atípicos. Comportamento semelhante pode ser observado nesta variável nos setores censitários (Gráfico e Mapa A1.8). No Mapa 4.8 observamos que dentre os bairros da Zona Sul, o Vidigal e a Gávea são aqueles que assumem os valores mais elevados desta proporção. Vale lembrar que a favela da Rocinha, a maior da América Latina, está situada na Gávea.

GRÁFICO 4.8 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS BAIROS DO RIO DE JANEIRO, 1991

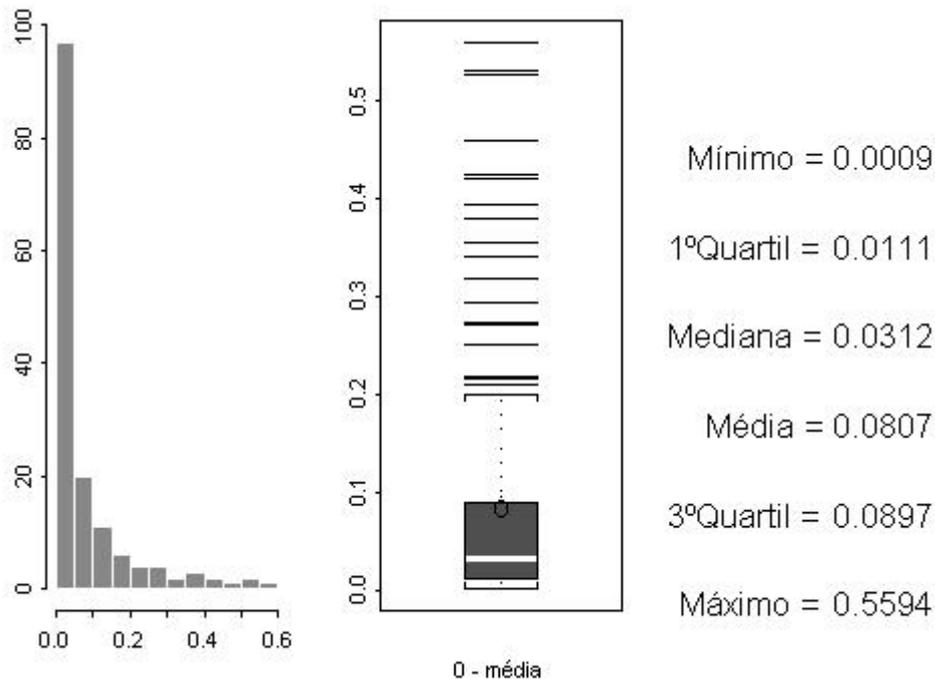


**MAPA 4.8 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS
BAIRROS DO RIO DE JANEIRO, 1991**

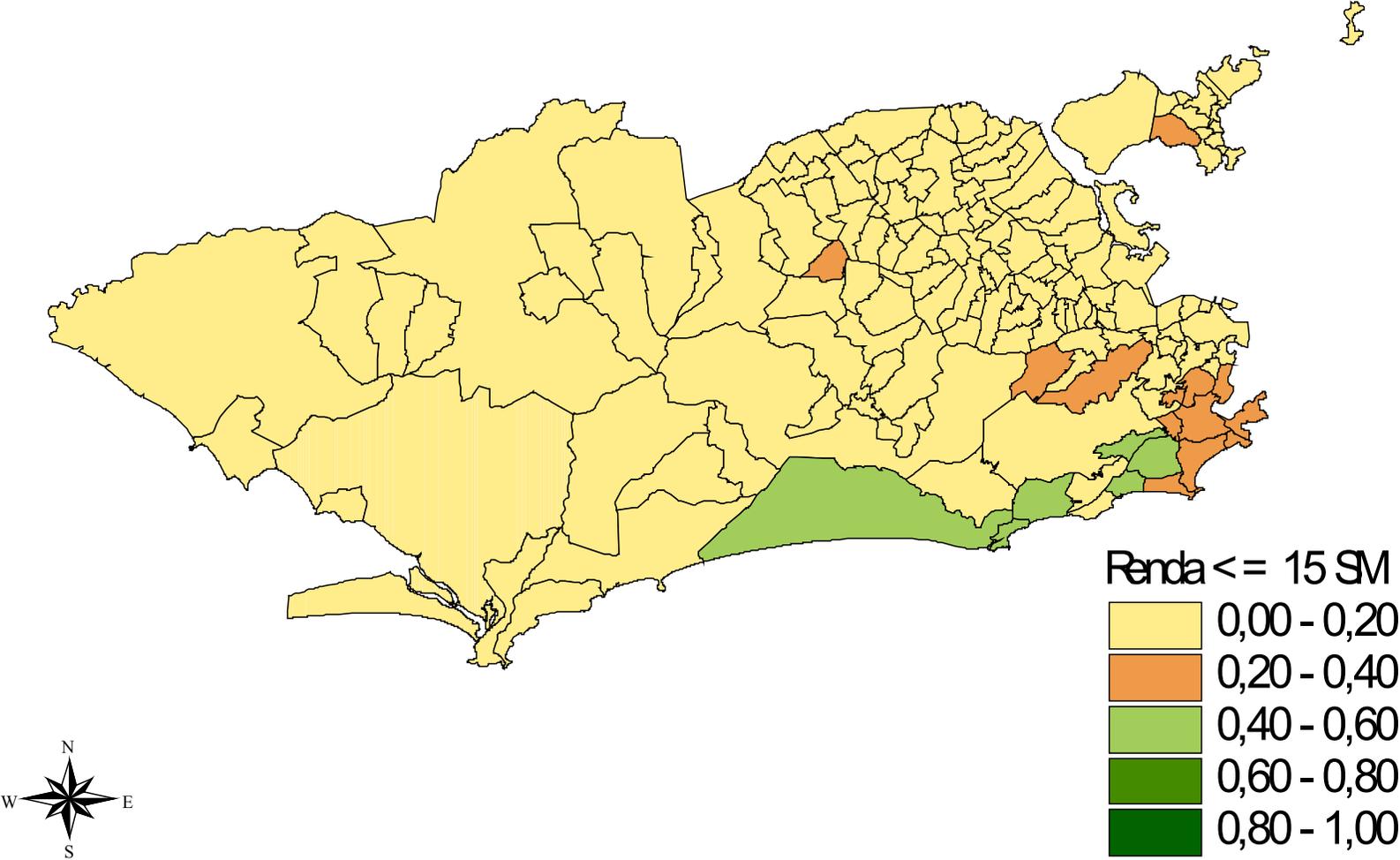


A proporção de chefes de domicílio com renda superior a quinze salários mínimos tem distribuição semelhante nos bairros e setores censitários do município do Rio de Janeiro, concentrada nos valores mais baixos (Gráfico 4.9). Em $\frac{1}{4}$ dos bairros apenas 1,11% dos chefes de domicílios têm renda superior a quinze salários mínimos e em nenhum dos bairros esta proporção é maior que 56%. Há também valores altos atípicos (Ex.: Cosme Velho e Joá) o que indica a heterogeneidade da distribuição desta proporção nos bairros. No Mapa 4.9 podemos ver que os chefes de domicílio com renda superior a quinze salários mínimos estão concentrados nos bairros Barra da Tijuca, Joá, São Conrado, Leblon, Lagoa e Jardim Botânico.

GRÁFICO 4.9 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS BAIROS DO RIO DE JANEIRO, 1991

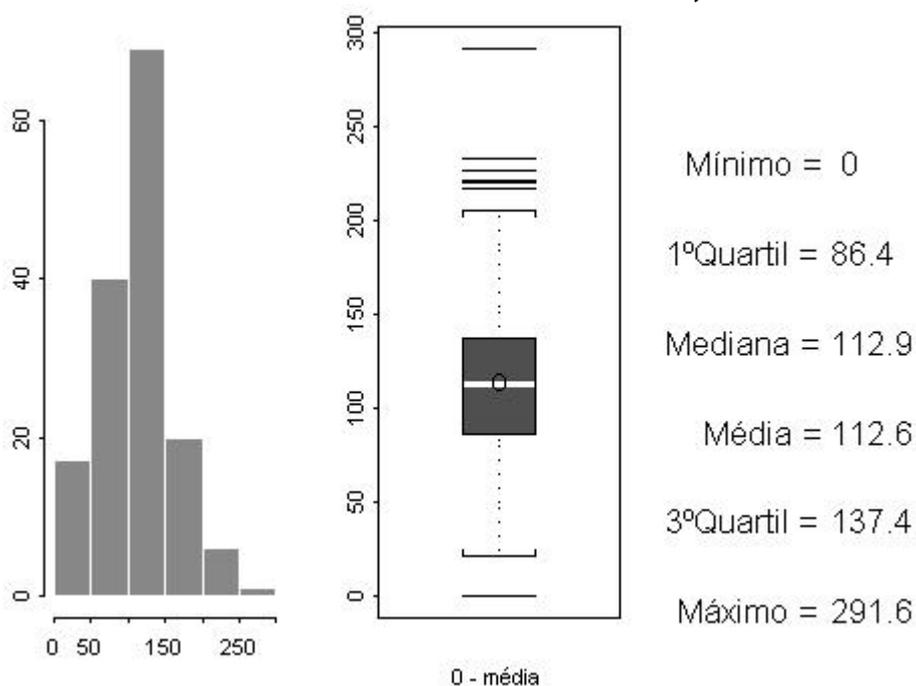


**MAPA 4.9 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS
BAIRROS DO RIO DE JANEIRO, 1991**

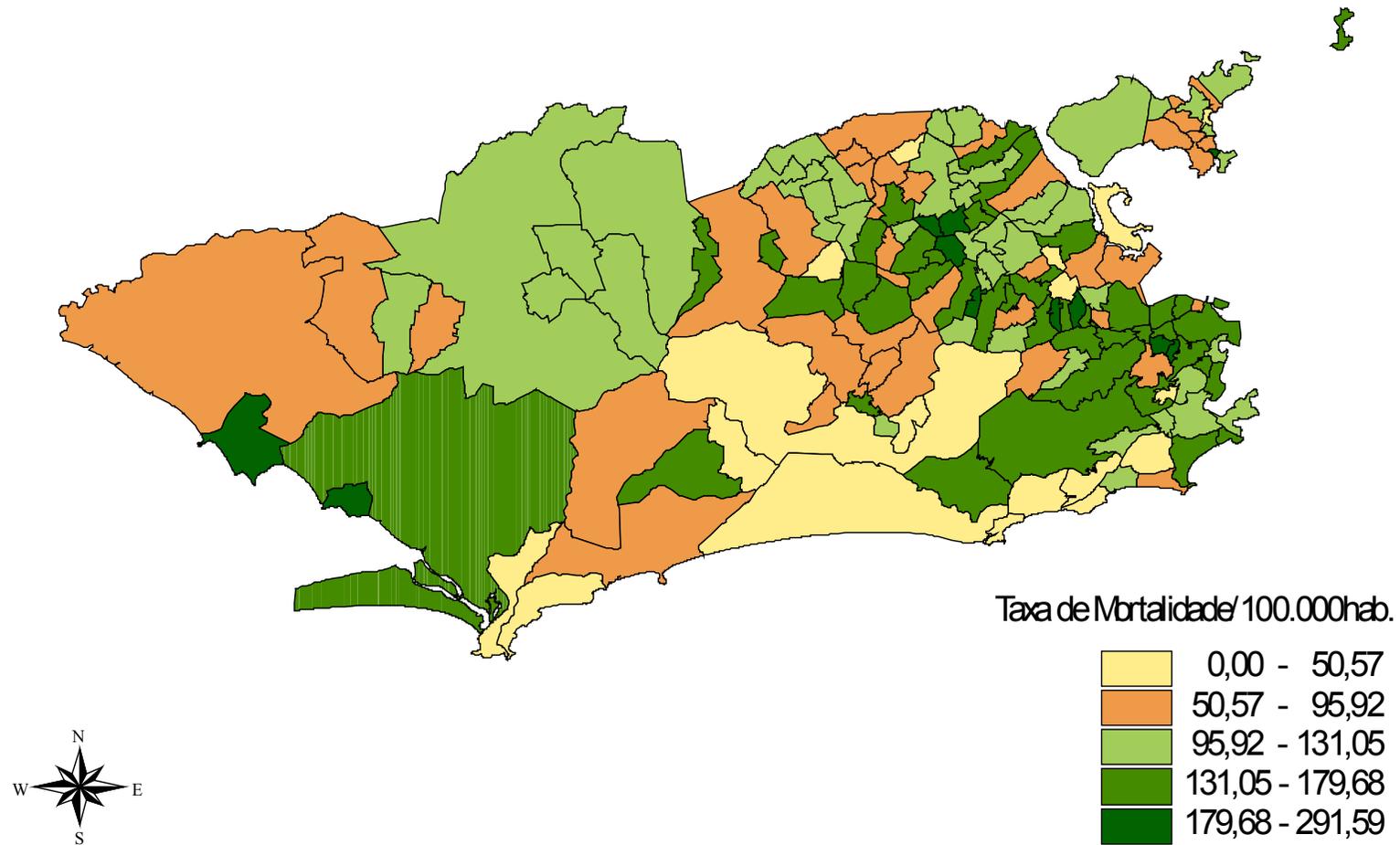


A distribuição da taxa de mortalidade por doença isquêmica do coração, na população de 30 a 70 anos nos bairros (Gráfico 4.10) é bastante distinta da distribuição nos setores censitários. A mediana da taxa nos bairros é 112,90 óbitos por 100.000 habitantes, e nos setores esta mesma medida apresenta valor nulo, já que é grande o número de setores censitários com zero óbitos (ver Gráfico e Mapa A1.10). No Mapa 4.10 podemos visualizar que alguns dos bairros que apresentam maior taxa de mortalidade por doença isquêmica do coração, na população de 30 a 70 anos são Vaz Lobo, Catumbi, Pedra de Guaratiba e Zumbi.

GRÁFICO 4.10 - ANÁLISE EXPLORATÓRIA DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS BAIROS DO RIO DE JANEIRO, 1991

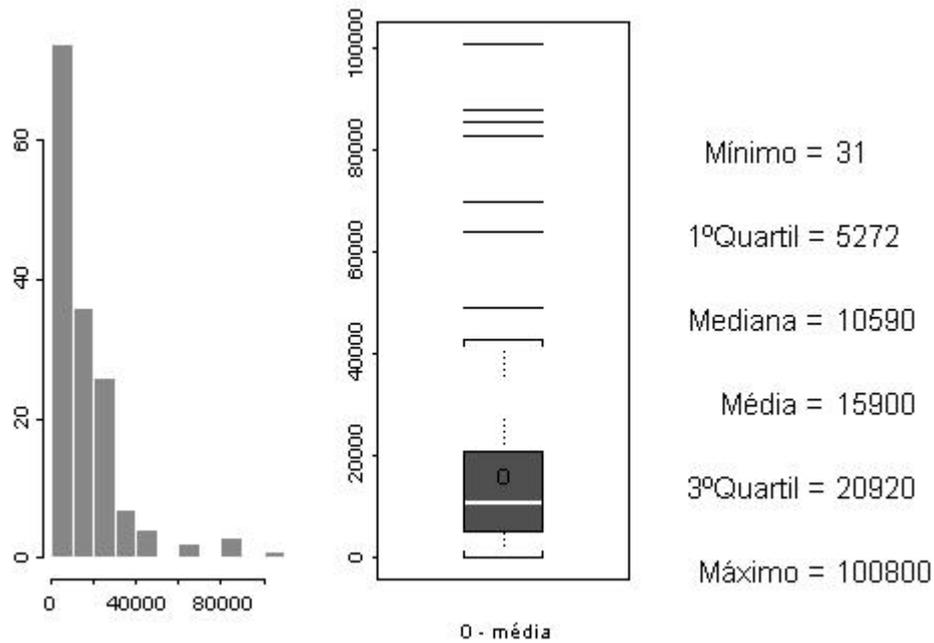


MAPA 4.10 - DISTRIBUIÇÃO ESPACIAL DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS BAIROS DO RIO DE JANEIRO, 1991

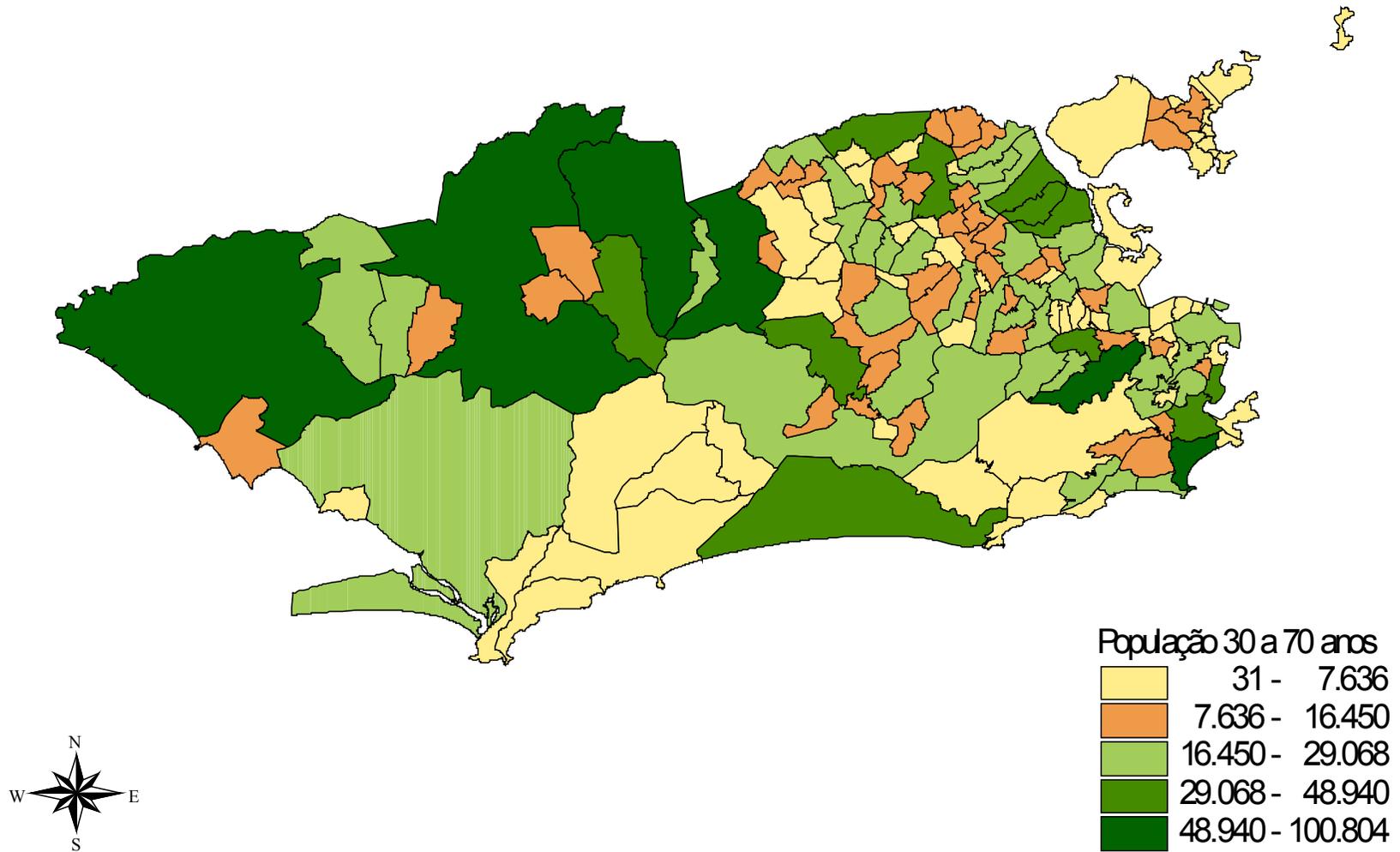


No Gráfico 4.11 podemos visualizar a população de 30 a 70 anos nos bairros do município do Rio de Janeiro. Há bairros com apenas 31 habitantes nesta faixa etária, o que compromete a estabilidade da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos. No Mapa 4.11 observamos que os bairros com menor população dessa faixa etária em geral estão localizados na Zona Norte da cidade.

GRÁFICO 4.11 - ANÁLISE EXPLORATÓRIA DA POPULAÇÃO DE 30 A 70 ANOS BAIROS DO RIO DE JANEIRO, 1991

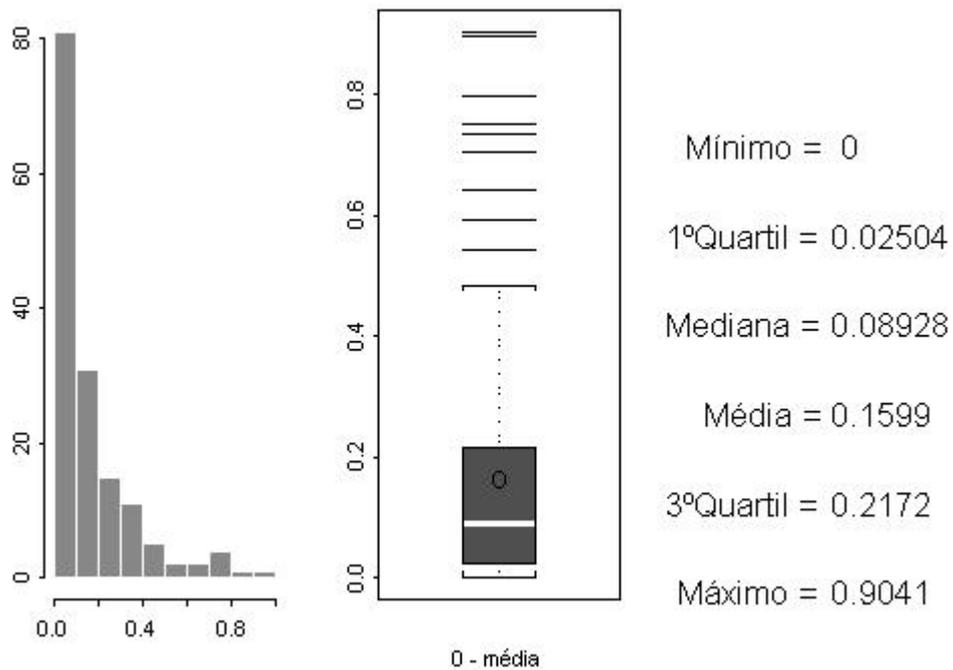


**MAPA 4.11 - DISTRIBUIÇÃO ESPACIAL DA POPULAÇÃO DE 30 A 70 ANOS
BAIRROS DO RIO DE JANEIRO, 1991**

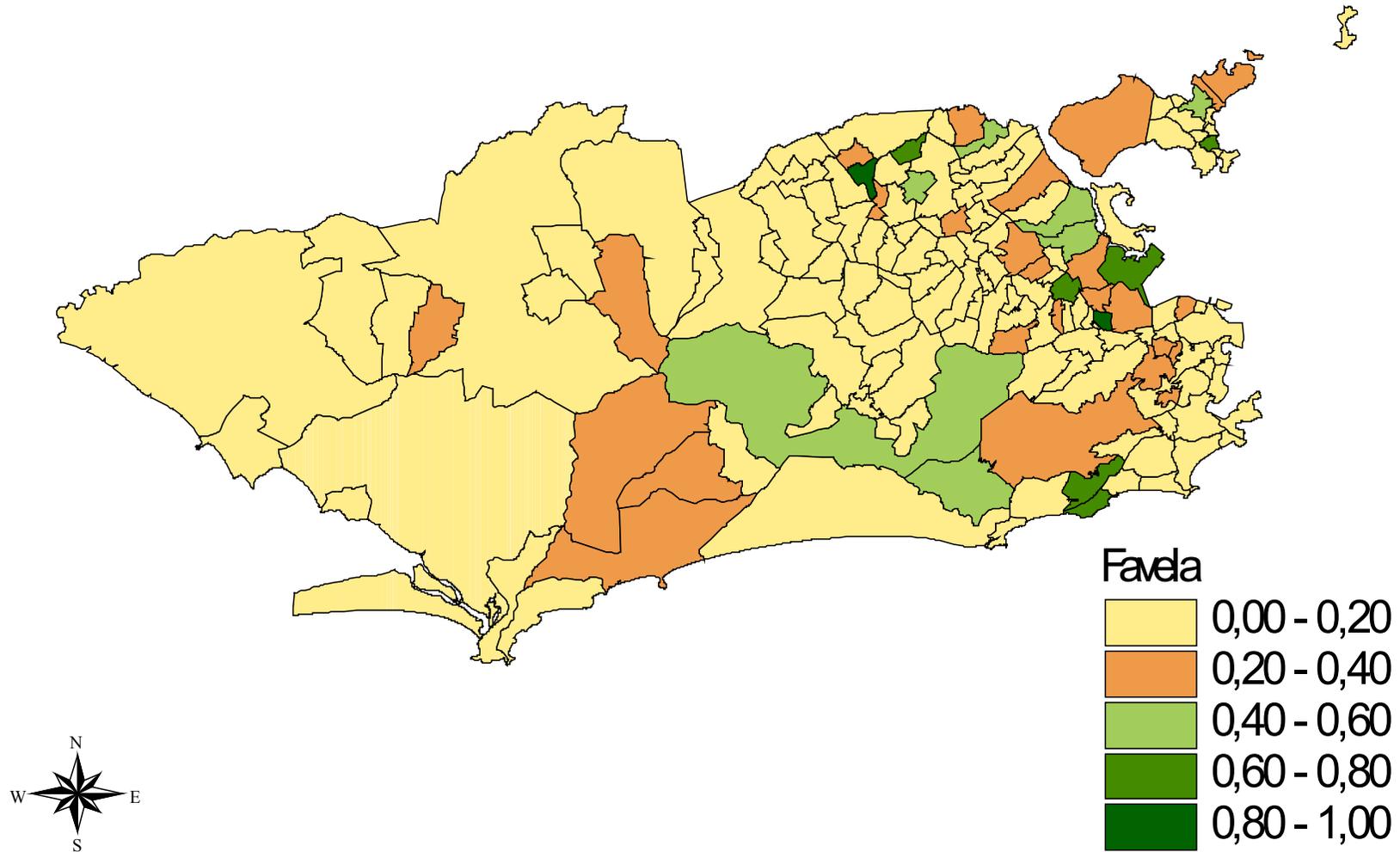


A proporção de setores favelados tem distribuição concentrada nos valores mais baixos (Gráfico e Mapa 4.12). Em $\frac{1}{4}$ dos bairros apenas 2,50% dos domicílios são sub-normais ou favelados. Há também valores altos atípicos o que indica que esta proporção é elevada em determinados bairros (Ex.: Acari, Barros Filho, Caju, Gávea, Jacaré, Mangueira e Vidigal).

**GRÁFICO 4.12 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE SETORES FAVELADOS
BAIRROS DO RIO DE JANEIRO, 1991**

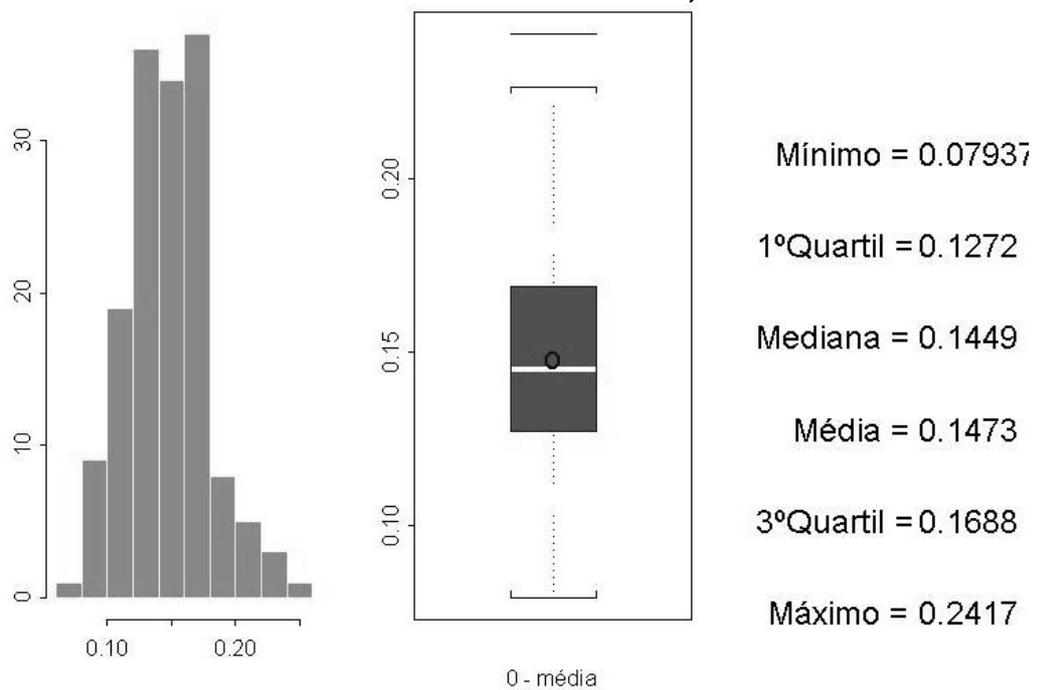


**MAPA 4.12 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE SETORES FAVELADOS
BAIRROS DO RIO DE JANEIRO, 1991**

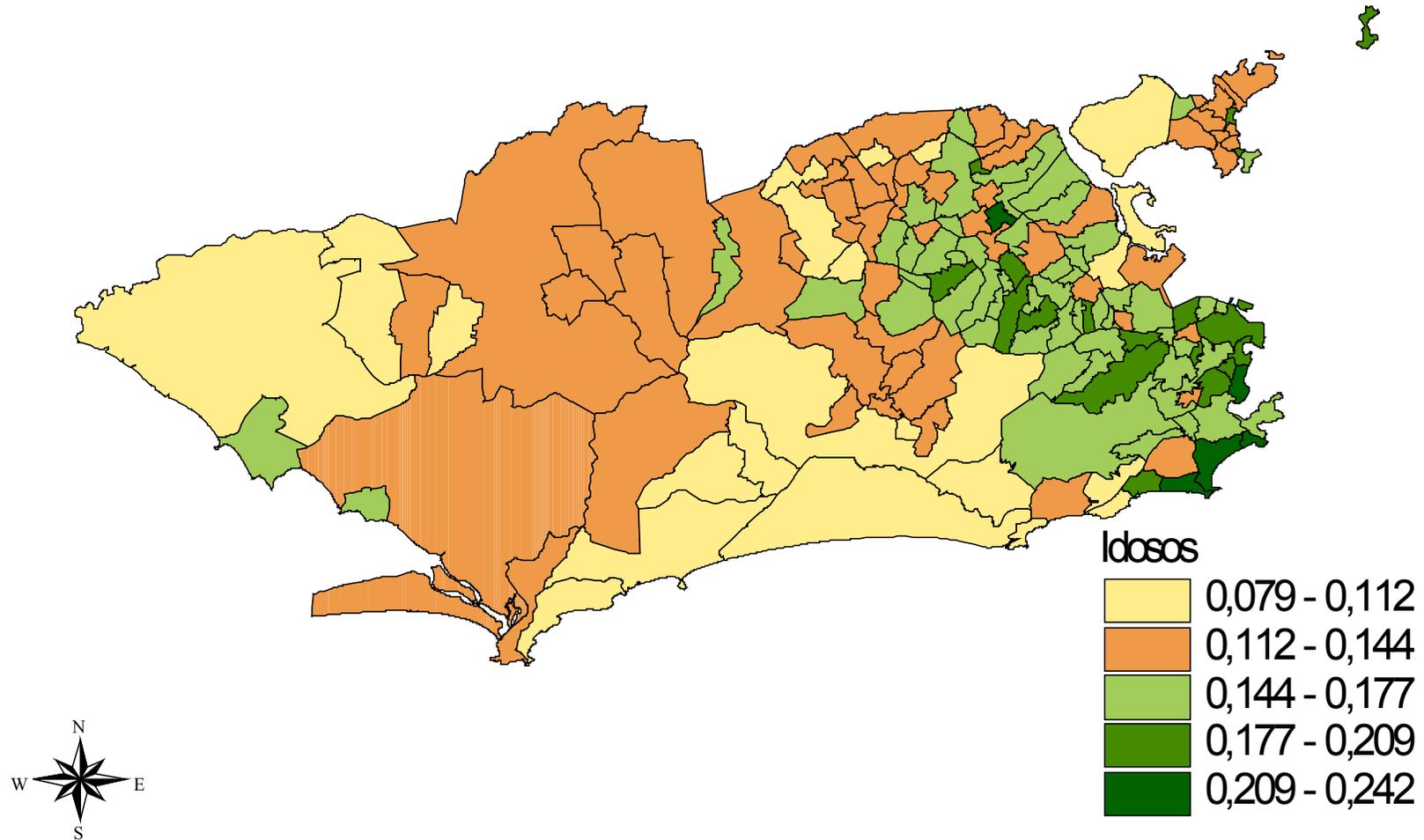


Sem valores atípicos, a proporção da população entre 60 e 70 anos de idade, varia entre 8% e 24% nos bairros (Gráfico e Mapa 4.13). Esta proporção é mais elevada no Centro, na Zona Sul e em Copacabana, bairro reconhecido pela grande população de idosos residentes. Os bairros de ocupação mais recente, localizados na Zona Oeste, apresentam uma baixa proporção de população idosa.

**GRÁFICO 4.13 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE POPULAÇÃO IDOSA
BAIRROS DO RIO DE JANEIRO, 1991**



**MAPA 4.13 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE POPULAÇÃO IDOSA
BAIRROS DO RIO DE JANEIRO, 1991**



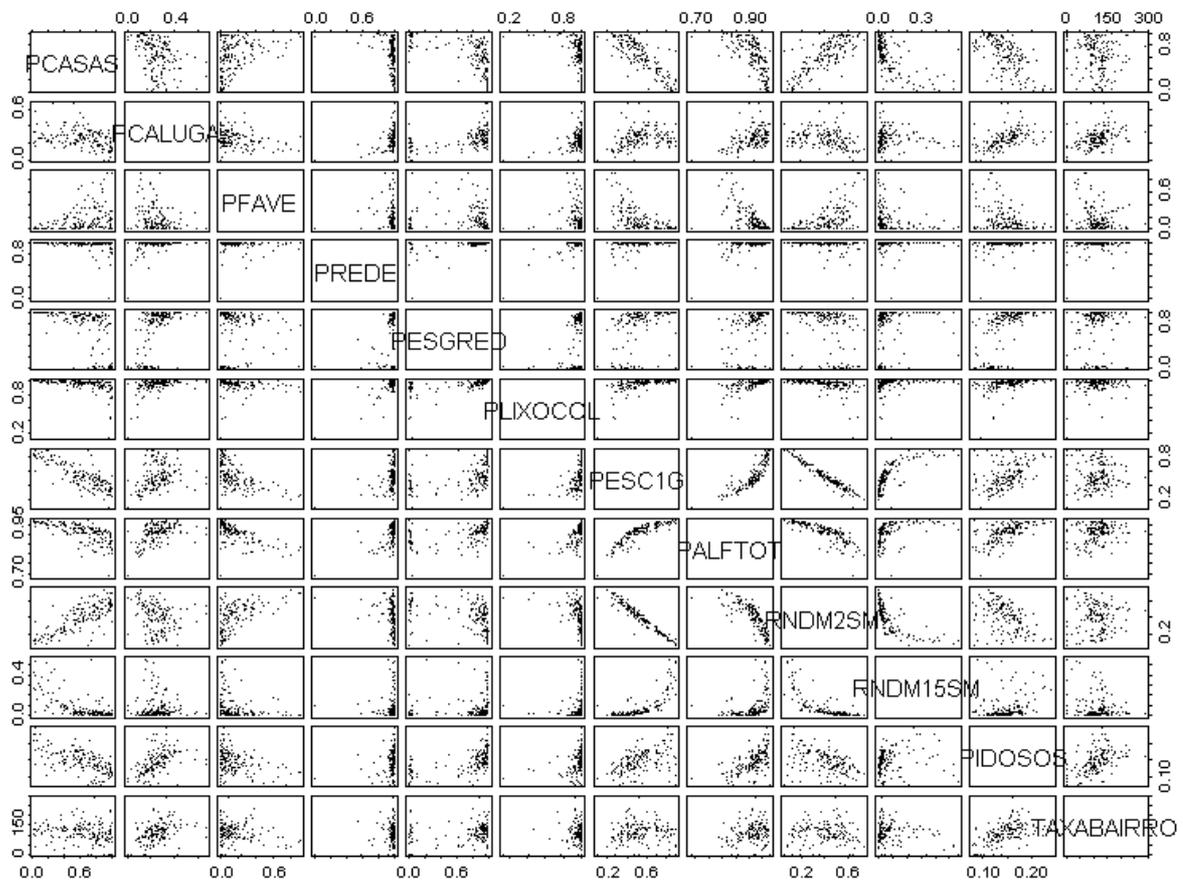
A análise exploratória das correlações entre os indicadores, realizada através do diagrama de espalhamento do Gráfico 4.14, com a direção da correlação resumida no Quadro 4.1, permite as seguintes observações:

- Variáveis indicativas de estratos sociais mais elevados — escolaridade do chefe de domicílio acima do 1º grau, alfabetização, renda acima de quinze salários mínimos, disponibilidade dos serviços de infra-estrutura urbana — esgoto e coleta de lixo — são positivamente correlacionadas. Com este bloco correlacionam-se negativamente os indicadores de carência de serviços públicos, renda baixa, proporção de setores favelados e de casas.
- No outro extremo da escala social encontram-se os bairros onde é alta a proporção de setores favelados, correlacionados positivamente com a presença de domicílios unifamiliares (casas), baixa renda e escolaridade. Entretanto, nestes bairros com alta proporção de setores favelados é grande a proporção de domicílios ligados a rede geral de água e com coleta regular de lixo. Em parte isto se explica pelo grau de urbanização das favelas no município do Rio de Janeiro, particularmente em relação à disponibilidade de água. Em contraste, temos a Zona Oeste, área desprovida de serviços públicos e não favelizada, onde predominam domicílios unifamiliares e de baixa renda (Carvalho e cols., 1997).
- Duas variáveis relacionadas aos serviços públicos essenciais — rede de água e coleta regular de lixo — positivamente correlacionadas à proporção de população alfabetizada e à proporção de domicílios alugados sugere a presença de bairros onde residem estratos sociais médios, com acesso à equipamento urbano básico. Nos bairros onde é grande a proporção de casas alugadas é pequena a proporção de setores favelados e de renda abaixo de dois e acima de quinze salários mínimos.

Nos bairros, a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos apresentou correlação com as variáveis: proporção de setores favelados (negativamente correlacionada), proporção de chefes de domicílio com renda superior a quinze salários mínimos (negativamente correlacionada) e proporção de população idosa (positivamente correlacionada).

Não se procurou testar a significância das correlações, uma vez que algumas são visivelmente não lineares, como por exemplo: escolaridade superior ao 1º grau e alfabetização; escolaridade superior ao 1º grau e renda alta; alfabetização e renda alta. Entretanto, como esta não linearidade não aparece na relação com o indicador resposta, não se fez transformação visando a linearidade destas relações que poderiam, inclusive, introduzir outras não linearidades.

**GRÁFICO 4.14 - DIAGRAMA DE ESPALHAMENTO DOS INDICADORES DOIS A DOIS
BAIRROS DO RIO DE JANEIRO, 1991**



QUADRO 4.1- DIREÇÃO DAS CORRELAÇÕES ENTRE OS INDICADORES DOS BAIROS DO RIO DE JANEIRO, 1991

	PCASAS	PCALUGA	PFAVE	PREDE	PESGRED	PLIXOCOL	PESC1G	PALFTOT	RNDM2SM	RNDM15SM	PIDOSOS	TAXABAIRRO
PCASAS												
PCALUGA												
PFAVE	+	-										
PREDE	-	+	+									
PESGRED												
PLIXOCOL	-	+	+		+							
PESC1G	-		-			+						
PALFTOT	-	+	-		+	+	+					
RNDM2SM	+		+			-	-	-				
RNDM15SM	-	-	-		+	-	+	+	-			
PIDOSOS	-	+						+	-			
TAXABAIRRO			-							-	+	

4.2 - Classificação Multivariada

Os setores censitários foram classificados a partir dos indicadores socioeconômicos em quatro grupos definidos previamente. Esta análise foi realizada somente para setores censitários visando uma classificação baseada em unidades com maior homogeneidade interna, classificação esta que será utilizada posteriormente na discussão dos resultados da modelagem.

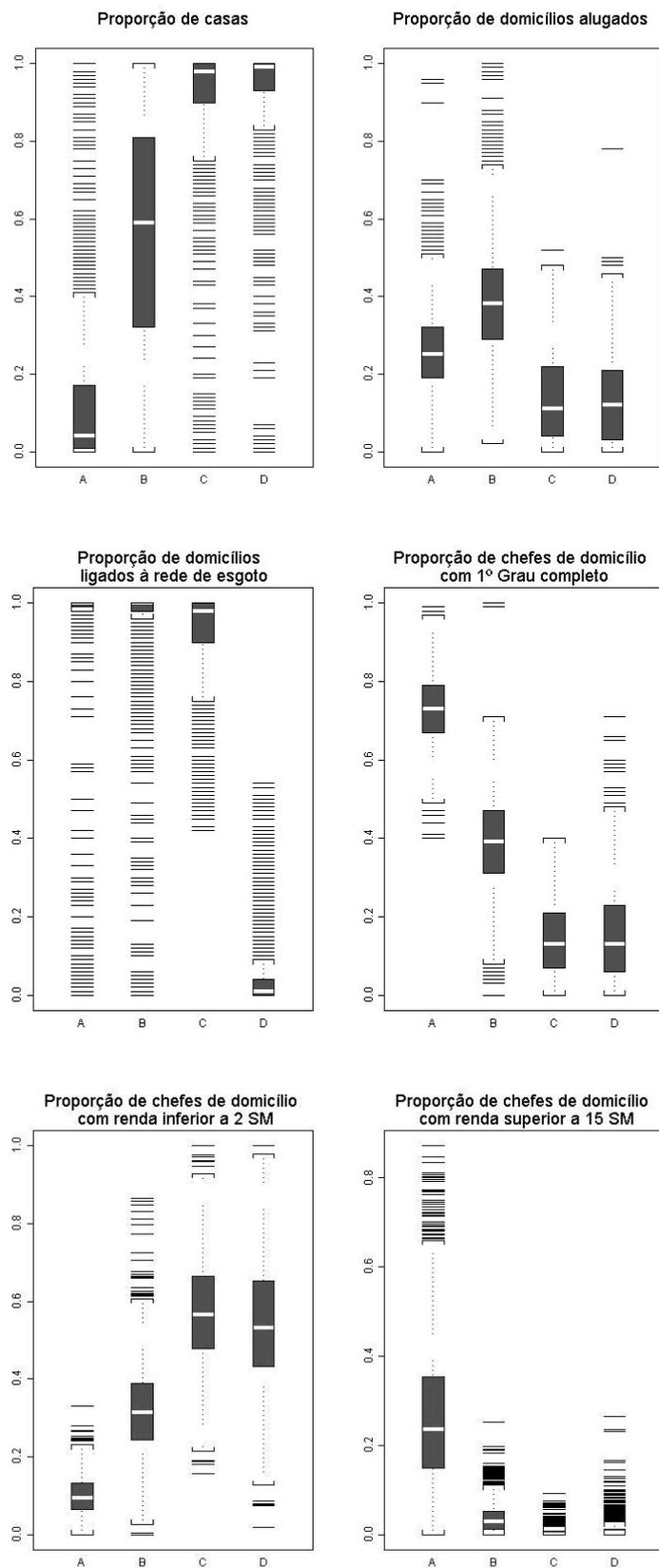
Não se incluiu o indicador da proporção de população idosa na construção do cluster, porque o objetivo desta classificação, é estabelecer um padrão socioeconômico sobre a unidade geográfica setor censitário. Este procedimento foi utilizado, visando discutir os perfis dos grupos sociais, em relação à mortalidade por doença isquêmica do coração na população de 30 a 70 anos.

Para atingirmos estes grupos, observamos dentre as classificações computadas, aquela que apresentou o maior valor do Perfil Médio do Cluster (seção 2.4) e os quatro grupos mais bem definidos graficamente. As variáveis que entraram na classificação final foram: proporção de casas, proporção de domicílios alugados, proporção de domicílios ligados à rede de esgoto, proporção de chefes de domicílio

com 1º grau completo ou mais, proporção de chefes de domicílio com renda inferior a dois salários mínimos e proporção de chefes de domicílio com renda superior a quinze salários mínimos. Os *boxplots* destas variáveis estão dispostos no Gráfico 4.15.

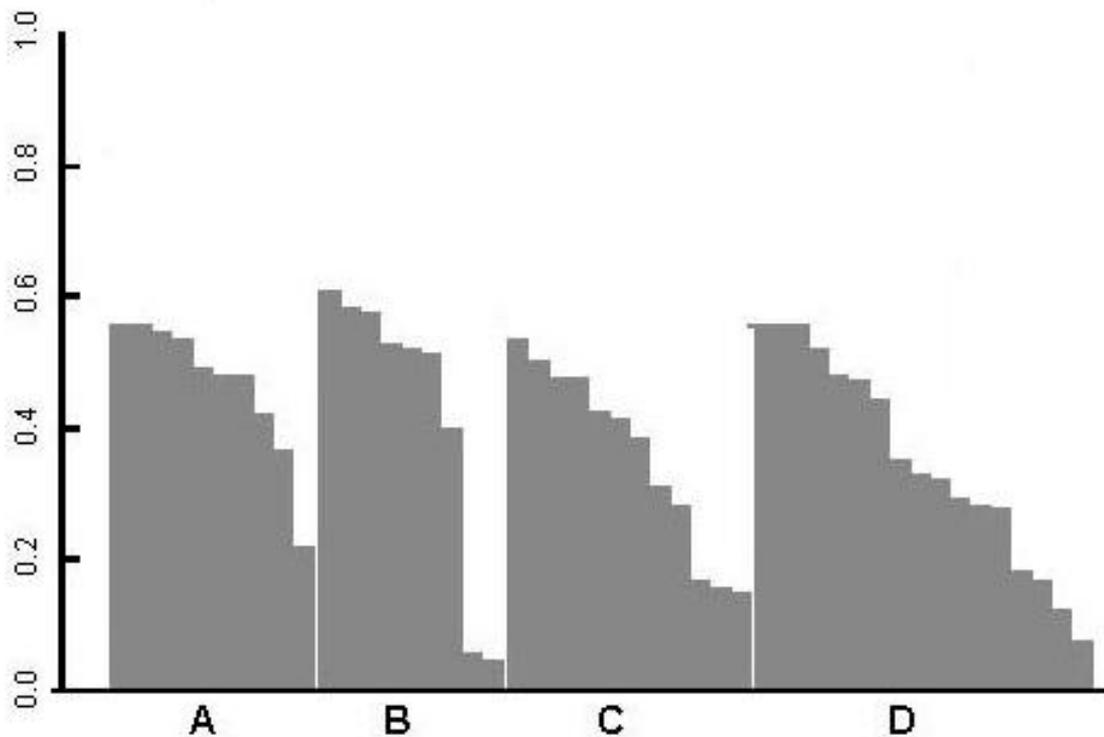
O grupo A refere-se aos setores censitários do estrato social mais alto em que os indicadores de infra-estrutura, escolaridade e renda acima de quinze salários mínimos assumem valores elevados. O grupo B refere-se ao estrato social médio. Neste estrato, os indicadores de alta renda e escolaridade em geral não são elevados e as proporções de casas e de domicílios alugados são altas. Este grupo é distinto do grupo C onde a mediana da proporção de domicílios alugados e da proporção de chefes de domicílio com 1º grau completo ou mais é inferior à do grupo B. O grupo C difere do grupo D apenas em relação à rede de esgoto. O grupo D refere-se aos setores censitários do estrato social menos privilegiado em que os indicadores de escolaridade, esgoto e alta renda assumem valores baixos.

GRÁFICO 4.15 - ANÁLISE EXPLORATÓRIA DE INDICADORES UTILIZADOS NA CONSTRUÇÃO DO CLUSTER

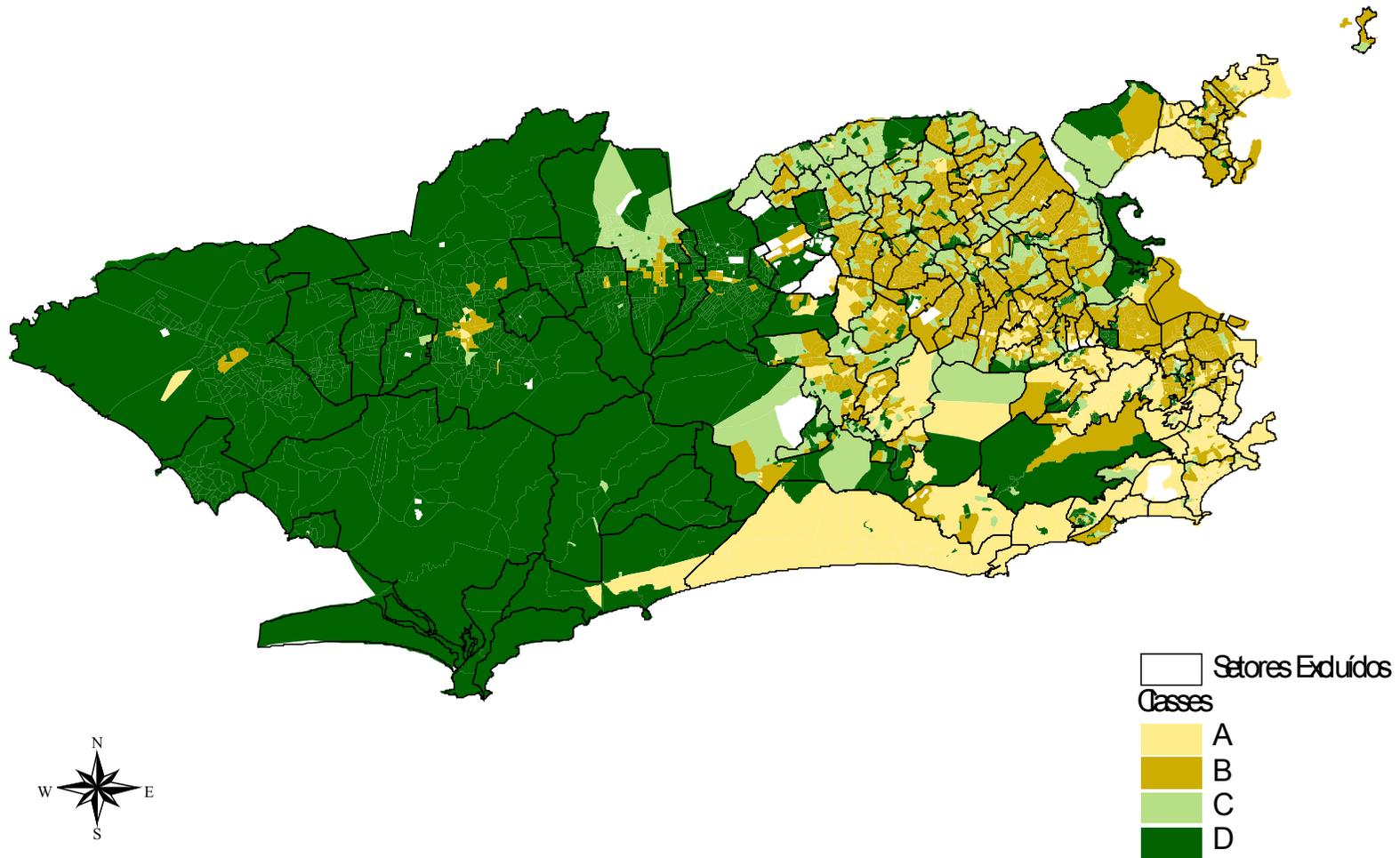


A silhueta dos estratos encontrados pode ser visualizada no Gráfico 4.16. Observe que o gráfico não possui valores negativos e os quatro grupos estão bem definidos indicando que o cluster criado é adequado aos dados. No Mapa 4.14, observamos que os setores censitários classificados no estrato A concentram-se principalmente na Zona Sul e Barra da Tijuca. Os setores dos estratos B e C são predominantemente da Zona Norte e das regiões centrais de Campo Grande e Santa Cruz. Já os setores censitários do estrato D são predominantemente da Zona Oeste.

GRÁFICO 4.16 - SILHUETA DOS ESTRATOS OBTIDOS NA CLASSIFICAÇÃO MULTIVARIADA



MAPA 4.14 - DISTRIBUIÇÃO ESPACIAL DAS CLASSES ENCONTRADAS



4.3 - Índice de Moran

O Índice de Moran dos indicadores socioeconômicos, para cada unidade geográfica de análise — setor censitário e bairro — é apresentado na Tabela 4.1. Nos setores censitários a autocorrelação espacial dos indicadores socioeconômicos é, em geral, maior do que nos bairros. Isto se deve, em primeiro lugar, ao efeito do tamanho dos setores censitários, como estas áreas em geral são pequenas, os vizinhos de 1ª ordem são bem próximos (Quadro 4.2). Além disso, sendo os bairros uma composição de setores censitários, com grande diversidade de perfis, os indicadores são média do processo, cuja distribuição tem, em geral, menor variabilidade. Assim, a semelhança inter-bairros — valores altos próximos a valores altos e valores baixos próximos a valores baixos — detectada através do Índice de Moran, fica menor.

QUADRO 4.2 - RAIO MÉDIO DOS SETORES CENSITÁRIOS

Raio^a Médio dos Setores Censitários	Região
247 metros	Município do Rio de Janeiro
144 metros	Bairros: Ramos, Irajá e Madureira
67 metros	Bairro: Copacabana

^a Raio é distância que vai de um ponto central para a periferia, numa dada área.

As exceções são as variáveis proporção de domicílios em que o lixo é coletado, proporção de domicílios ligados à rede geral de água e taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos. Esta última, devido a grande instabilidade nos setores censitários e ao número acentuado de taxas de mortalidade iguais a zero, apresenta correlação espacial de apenas 0,018.

TABELA 4.1 - ÍNDICE DE MORAN PARA SETORES CENSITÁRIOS E BAIRROS

Indicadores	Setor Censitário		Bairro	
	Moran	p	Moran	p
Proporção de chefes de domicílio com renda superior a 15 SM	0,799	0,000	0,556	0,000
Proporção de chefes de domicílio com 1º grau completo ou mais	0,760	0,000	0,385	0,000
Proporção de domicílios ligados à rede de esgoto	0,735	0,000	0,633	0,000
Proporção de casas	0,735	0,000	0,514	0,000
Proporção de chefes de domicílio com renda inferior a 2 SM	0,650	0,000	0,376	0,000
Proporção de população alfabetizada	0,538	0,000	0,311	0,000
Proporção de domicílios alugados	0,489	0,000	0,484	0,000
Proporção de domicílios em que o lixo é coletado	0,337	0,000	0,445	0,000
Proporção de Domicílios ligados à rede geral de água	0,287	0,000	0,355	0,000
Taxa de mortalidade por DIC na população de 30 a 70 anos	0,018	0,017	0,196	0,000
Proporção de setores favelados	-	-	0,093	0,029
Proporção de população idosa	-	-	0,227	0,000

4.4 - Regressão Linear Simples

A regressão linear simples foi calculada tomando-se como base a seguinte relação:

$$Tx_DIC_{\text{bairro-RJ}}^{30-70} \sim [\text{constante} + (\% \text{ de casas}) + (\% \text{ de domicílios alugados}) + (\% \text{ de domicílios ligados à rede geral de água}) + (\% \text{ de domicílios ligados à rede de esgoto}) + (\% \text{ de domicílios em que o lixo é coletado}) + (\% \text{ de chefes de domicílio com 1º grau completo ou mais}) + (\% \text{ de população alfabetizada}) + (\% \text{ de chefes de domicílio com renda inferior a dois SM}) + (\% \text{ de chefes de domicílio com renda superior a quinze SM}) + (\% \text{ setores de favelados}) + (\% \text{ da população idosa})]$$

Este modelo foi ajustado para que posteriormente seja realizada a comparação dos seus resultados com os resultados do modelo de regressão espacial.

A equação do modelo de regressão linear simples ajustado é:

$$\begin{aligned} \text{Tx_DIC}_{\text{bairro-RJ}}^{30-70} = & -181,603 \\ & + 779,851 * (\% \text{ de população idosa}) \\ & + 205,684 * (\% \text{ chefes domicílio com } 1^\circ \text{ Grau completo ou mais}) \\ & + 206,945 * (\% \text{ chefes domicílio com renda inferior a } 2 \text{ SM}) \\ & - 154,695 * (\% \text{ chefes domicílio com renda superior a } 15 \text{ SM}) \end{aligned}$$

A taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros é inversamente correlacionada à proporção de chefes de domicílio com renda superior a quinze salários mínimos, e é positivamente correlacionada com a proporção de chefes de domicílio com 1º grau completo e proporção de chefes de domicílio com renda inferior a dois salários mínimos, controlados pela proporção da população idosa.

Na Tabela 4.2 temos discriminados os valores dos coeficientes de regressão e o teste de significância — estatística t ($\alpha = 0,050$) — para cada um dos coeficientes. Considerando-se que a associação negativa com a proporção de chefes de domicílio com renda superior a quinze salários mínimos discrimina o estrato social mais elevado, o modelo ajustado sugere que a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos é menor neste estrato. Por outro lado, os valores mais elevados da taxa de mortalidade ajustada estão associados aqueles bairros em que a escolaridade dos chefes de domicílio é alta e a renda é baixa.

TABELA 4.2 - COEFICIENTES DE REGRESSÃO DO MODELO DE REGRESSÃO LINEAR SIMPLES

	Coeficiente	Erro Padrão	Estatística t	p
Intercepto	-181,603	67,543	-2,689	0,008
Proporção da população idosa	779,851	122,445	6,369	0,000
Proporção de chefes de domicílio com 1º grau completo ou mais	205,684	77,606	2,650	0,009
Proporção de chefes de domicílio com renda inferior a 2 SM	206,945	77,906	2,656	0,009
Proporção de chefes de domicílio com renda superior a 15 SM	-154,695	43,323	-3,571	0,000

O R^2 encontrado no modelo ajustado foi de 0,381 o que indica que 38,10% da variabilidade total da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros é explicada pelas variáveis independentes do modelo.

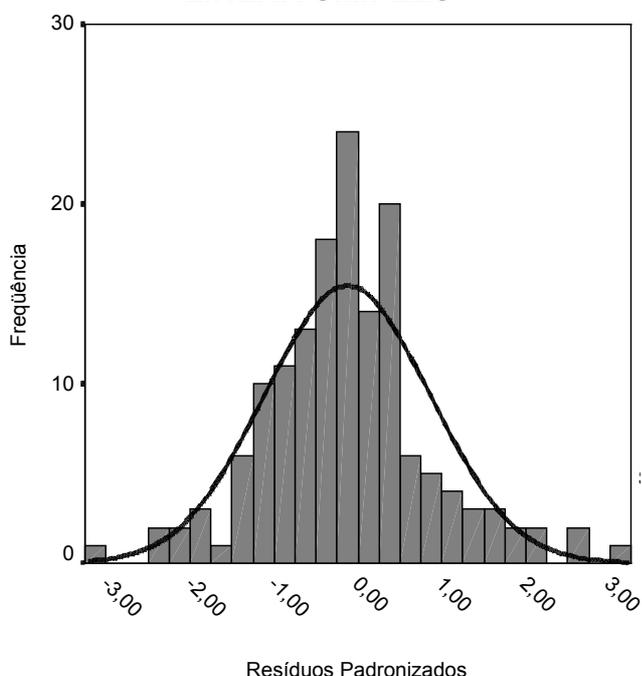
A análise da variância (Tabela 4.3) nos permite verificar que o ajuste é significativo ao nível $\alpha = 0,050$.

TABELA 4.3 - ANÁLISE DA VARIÂNCIA DO MODELO DE REGRESSÃO LINEAR SIMPLES

	Soma de Quadrados	g.l.	Média Quadrática	F	p
Regressão	144.486,732	4	36.121,683	22,752	0,000
Erro	234.971,511	148	1.587,645		
Total	379.458,243	152			

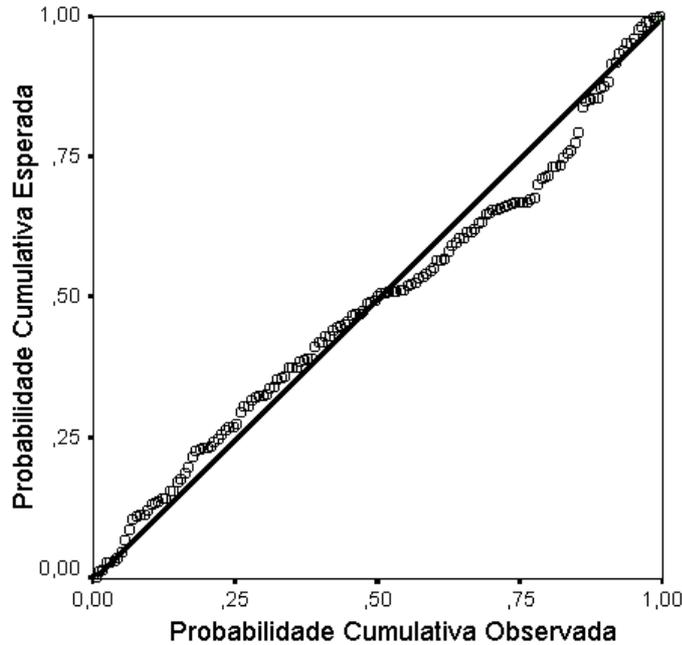
Os Gráficos 4.17 a 4.19 apresentam a análise gráfica dos resíduos. No histograma (Gráfico 4.17), podemos observar que os resíduos do modelo têm uma distribuição próxima da distribuição Normal, com média 0 e desvio padrão 0,99.

GRÁFICO 4.17 - HISTOGRAMA DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR SIMPLES



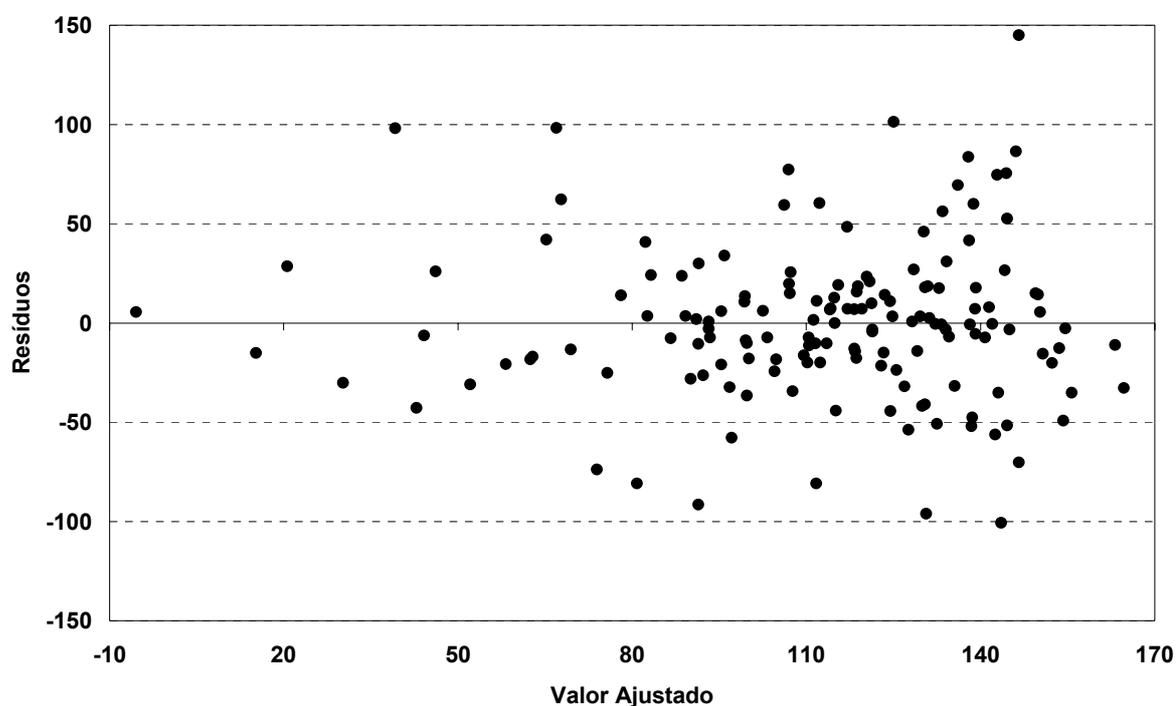
No Gráfico 4.18 temos a representação da probabilidade Normal, onde observamos que há indicativos de afastamento das hipóteses de Normalidade dos resíduos.

GRÁFICO 4.18 - PROBABILIDADE NORMAL DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR SIMPLES



A observação da dispersão dos resíduos do modelo em relação aos valores ajustados da taxa de mortalidade por doença isquêmica do coração, deve ser realizada de forma cuidadosa no Gráfico 4.19. Neste gráfico, percebemos a presença de um padrão na distribuição dos resíduos que lembra a forma de um cone, o que indica que os resíduos do modelos são correlacionados. Foram ajustados modelos de regressão linear com transformações Log, Inversa e Normal mas os resultados pouco se alteraram optando-se por manter as variáveis originais.

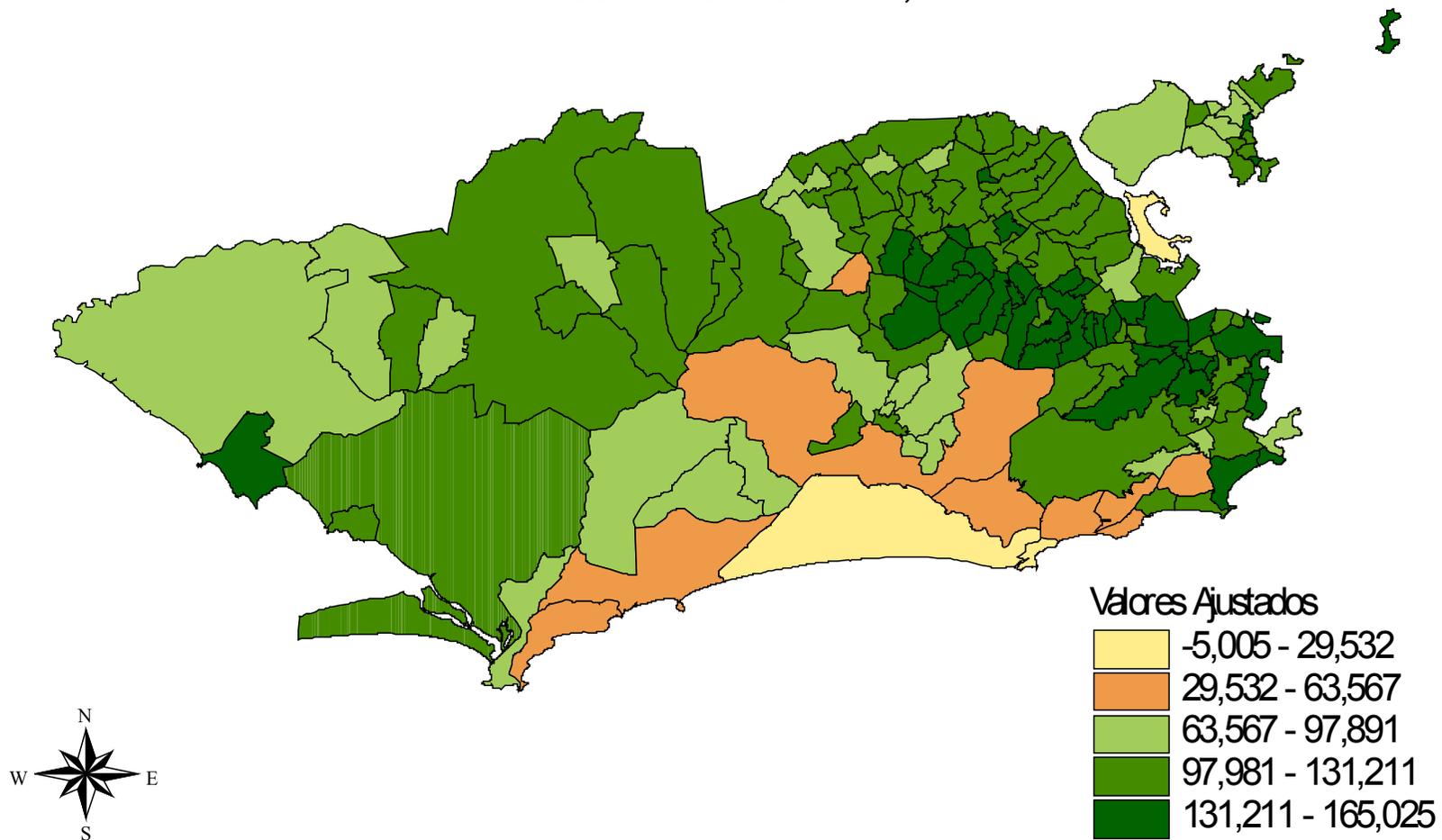
GRÁFICO 4.19 - DISPERSÃO DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR EM RELAÇÃO AOS VALORES AJUSTADOS BAIROS DO RIO DE JANEIRO



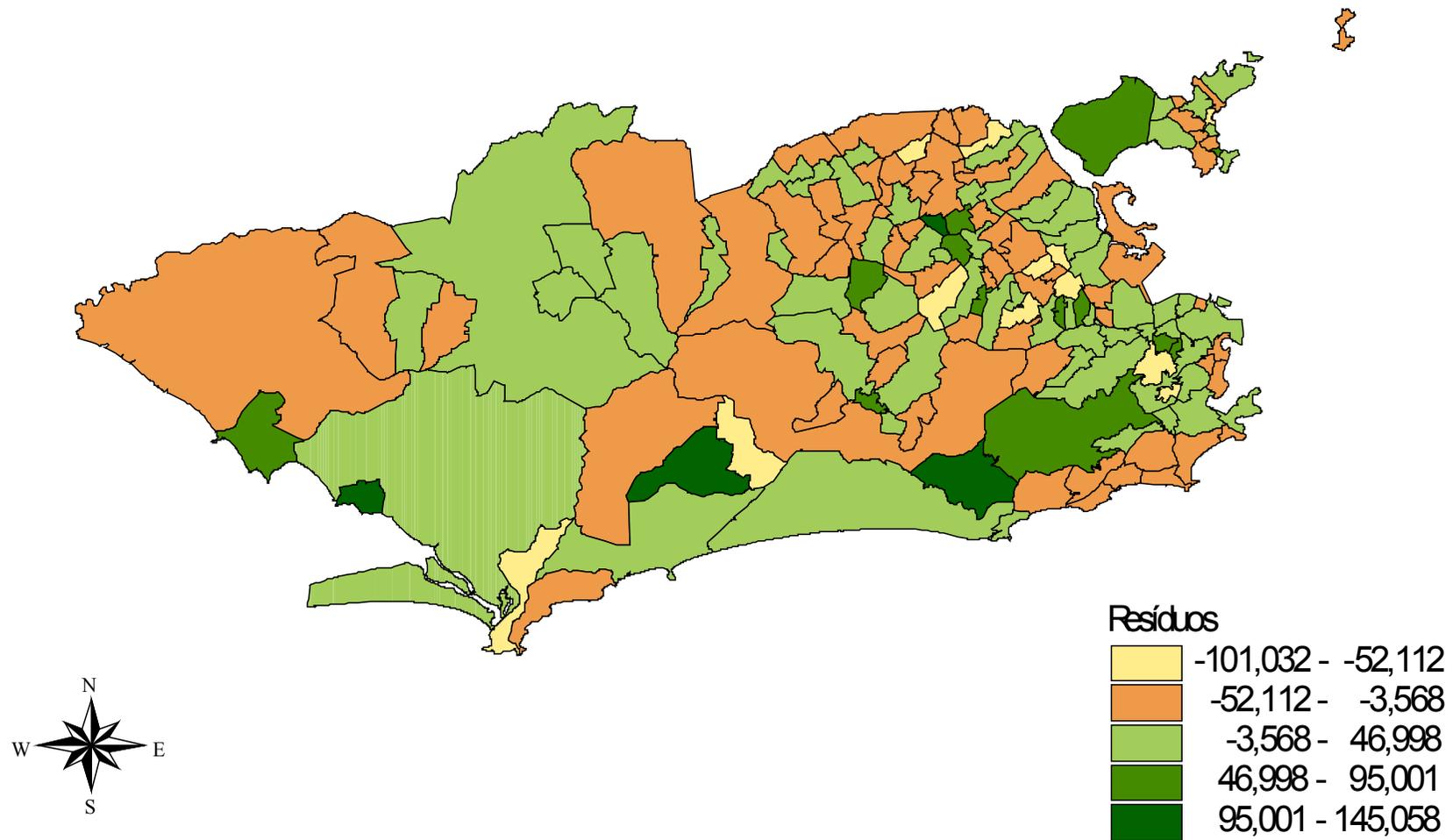
Para avaliar a correlação espacial dos resíduos do modelo de regressão linear, executamos o teste de autocorrelação espacial (Moran I). Neste modelo, há indicação de pequena autocorrelação espacial (Moran I = 0,079, $p = 0,050$). Em relação à distribuição espacial dos resíduos, observamos no Mapa 4.16 que não há indicativos da presença de tendência, ou seja, os erros distribuem-se aleatoriamente.

No Mapa 4.15, visualizamos que os valores mais elevados da taxa ajustada estão concentrados nos bairros da Zona Leste do município. Na Zona Oeste, em especial na Barra da Tijuca e no Joá, onde é grande a proporção de chefes de domicílio com renda superior a quinze salários mínimos, a taxa ajustada aproxima-se de zero.

MAPA 4.15 - DISTRIBUIÇÃO ESPACIAL DAS TAXAS AJUSTADAS DE MORTALIDADE POR DIC NO GRUPO ETÁRIO 30 A 70 ANOS - MODELO DE REGRESSÃO LINEAR SIMPLES BAIROS DO RIO DE JANEIRO, 1991



**MAPA 4.16 - DISTRIBUIÇÃO ESPACIAL DOS RESÍDUOS DO MODELO DE REGRESSÃO LINEAR SIMPLES
BAIRROS DO RIO DE JANEIRO, 1991**



4.5 - Regressão Espacial

A regressão espacial foi calculada tomando-se como base a seguinte relação:

$$\text{Tx_DIC}_{\text{bairro-RJ}}^{30-70} \sim [\text{constante} + (\% \text{ de casas}) + (\% \text{ de domicílios alugados}) + (\% \text{ de domicílios ligados à rede geral de água}) + (\% \text{ de domicílios ligados à rede de esgoto}) + (\% \text{ de domicílios em que o lixo é coletado}) + (\% \text{ de chefes de domicílio com 1º grau completo ou mais}) + (\% \text{ de população alfabetizada}) + (\% \text{ de chefes de domicílio com renda inferior a dois SM}) + (\% \text{ de chefes de domicílio com renda superior a quinze SM}) + (\% \text{ setores de favelados}) + (\% \text{ da população idosa}) + (\text{erros autocorrelacionados})]$$

A equação do modelo CAR ajustado é:

$$\begin{aligned} \text{Tx_DIC}_{\text{bairro-RJ}}^{30-70} = & -187,316 \\ & + 821,259 * (\% \text{ de população idosa}) \\ & + 205,052 * (\% \text{ chefes domicílio com 1º Grau completo ou mais}) \\ & + 204,614 * (\% \text{ chefes domicílio com renda inferior a 2 SM}) \\ & - 154,173 * (\% \text{ chefes domicílio com renda superior a 15 SM}) \end{aligned}$$

Na Tabela 4.4 temos os coeficientes de regressão do modelo espacial ajustado. Os resultados do modelo de regressão espacial são semelhantes aos do modelo de regressão linear simples, onde o ajuste sugere que a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos aproxima-se de zero no estrato social mais elevado, discriminado pela variável proporção de chefes de domicílio com renda superior a quinze salários mínimos.

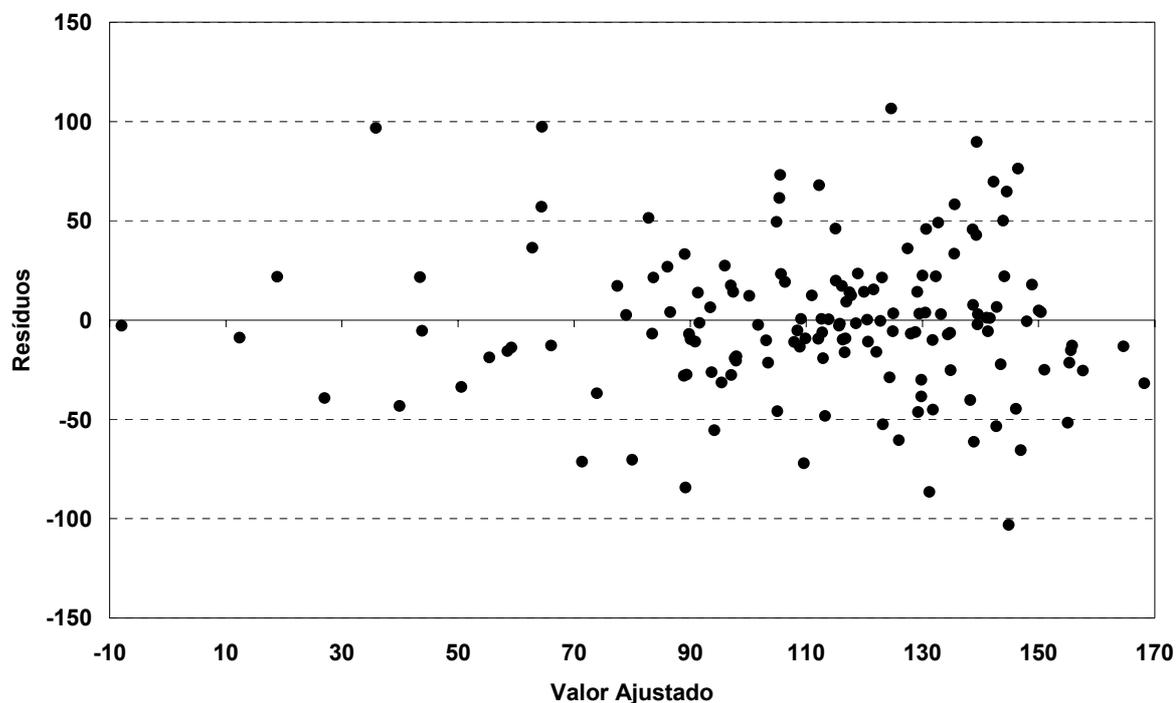
TABELA 4.4 - COEFICIENTES DE REGRESSÃO DO MODELO CAR

	Coefficiente	Erro padrão	Estatística t	p
Intercepto	-187,316	66,645	-2,810	0,005
Proporção da população idosa	821,259	125,934	6,521	0,000
Proporção de chefes de domicílio com 1º grau completo ou mais	205,052	76,704	2,673	0,008
Proporção de chefes de domicílio com renda inferior a 2 SM	204,614	76,671	2,668	0,008
Proporção de chefes de domicílio com renda superior a 15 SM	-154,173	45,541	-3,385	0,000

Para avaliar os resíduos do modelo espacial, é necessário fazer o teste de autocorrelação espacial (Moran I). No caso do modelo selecionado, não há indicação de correlação espacial (Moran I = -0,068, p = 0,901). No Gráfico 4.20 temos a dispersão dos resíduos do modelo CAR ajustado. Segundo a inspeção gráfica, a dispersão dos resíduos ainda lembra a forma de um cone.

Mesmo com o Índice de Moran indicando que não há autocorrelação espacial dos resíduos do modelo, ajustamos um novo modelo com a retirada da tendência espacial dos dados. É possível que com a retirada da tendência, este padrão de concentração dos valores extremos numa determinada região se modifique.

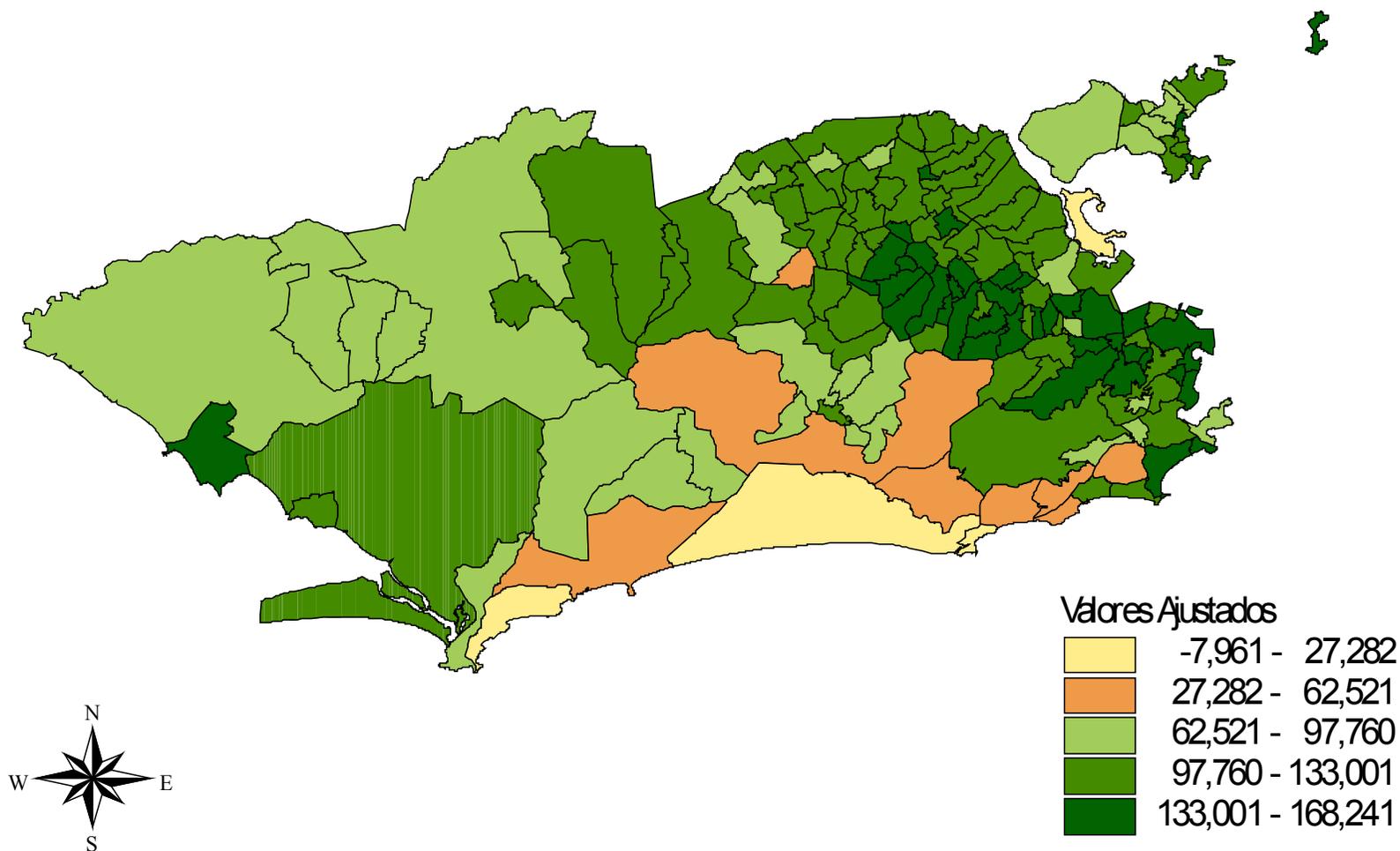
GRÁFICO 4.20 - DISPERSÃO DOS RESÍDUOS DO MODELO CAR EM RELAÇÃO AOS VALORES AJUSTADOS BAIROS DO RIO DE JANEIRO



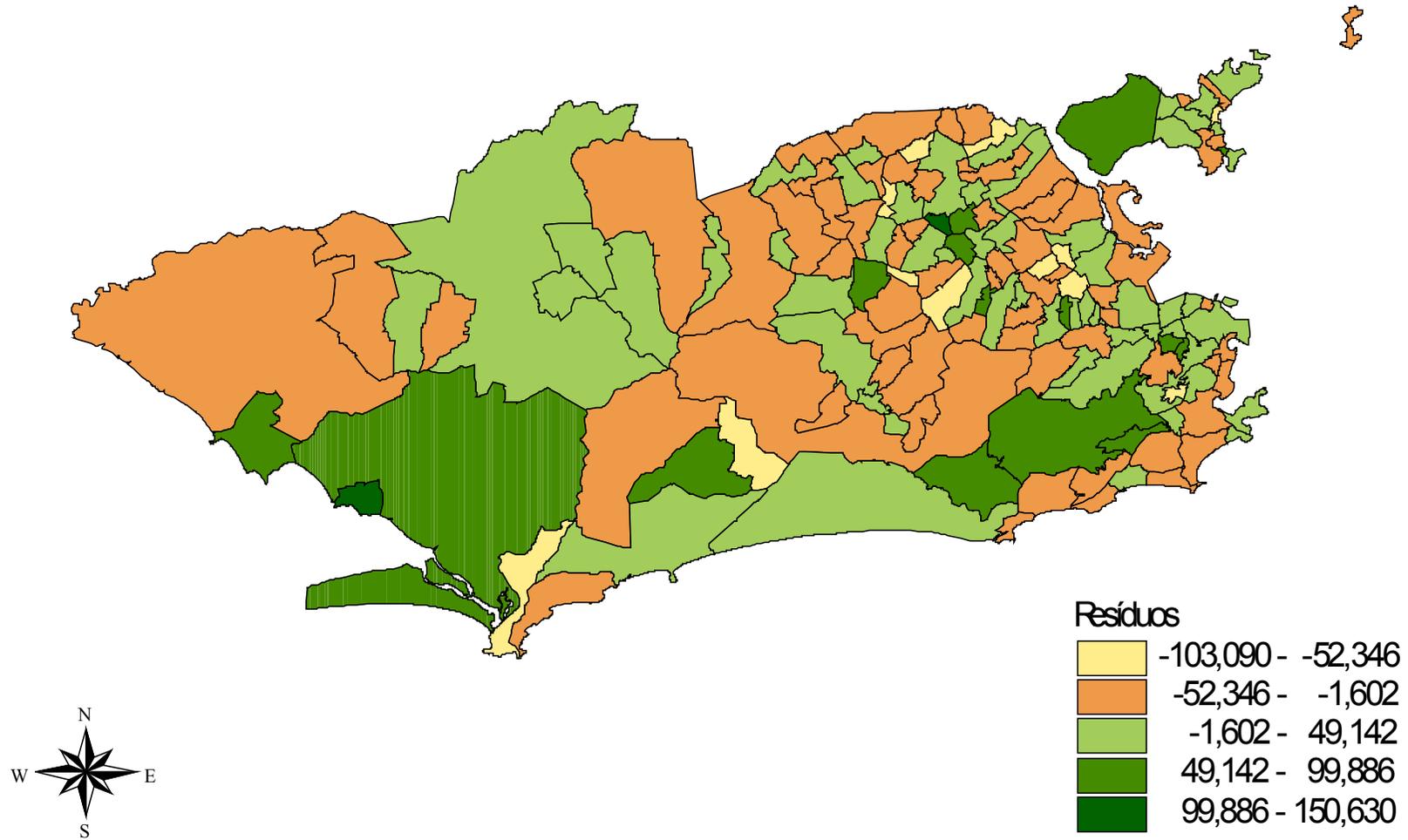
Assim como no modelo de regressão linear, no Mapa 4.17 podemos observar que os valores mais elevados da taxa de mortalidade ajustada pelo modelo CAR estão associados aos bairros da Zona Leste do município; e os valores próximos de zero ocorrem em bairros onde é grande a proporção de chefes de domicílio com renda superior a quinze salários mínimos (Ex.: Barra da Tijuca e Joá).

No Mapa 4.18 observamos que a distribuição espacial dos resíduos do modelo CAR ajustado não é muito diferente da distribuição dos resíduos do modelo de regressão linear simples, onde não há indicativos da presença de tendência, ou seja, os resíduos distribuem-se aleatoriamente.

**MAPA 4.17 - DISTRIBUIÇÃO ESPACIAL DAS TAXAS AJUSTADAS DE MORTALIDADE POR DIC NO GRUPO ETÁRIO 30 A 70 ANOS - MODELO CAR
BAIRROS DO RIO DE JANEIRO, 1991**



**MAPA 4.18 - DISTRIBUIÇÃO ESPACIAL DOS RESÍDUOS DO MODELO CAR AJUSTADO
BAIRROS DO RIO DE JANEIRO, 1991**



4.6 - Detecção de Tendência Espacial

A avaliação da presença de tendência espacial nos indicadores utilizados gerou o Quadro 4.3, onde estão representados os resíduos e a linha de tendência estimados no modelo em relação à posição geográfica. O gráfico indicativo de estacionariedade da média é aquele que apresenta a linha de tendência constante paralela ao eixo das abscissas com os resíduos aleatoriamente distribuídos em volta desta reta. A densidade de pontos em determinada região do gráfico é expressa pela quantidade de pequenos traços verticais (|) que cortam o eixo das abscissas.

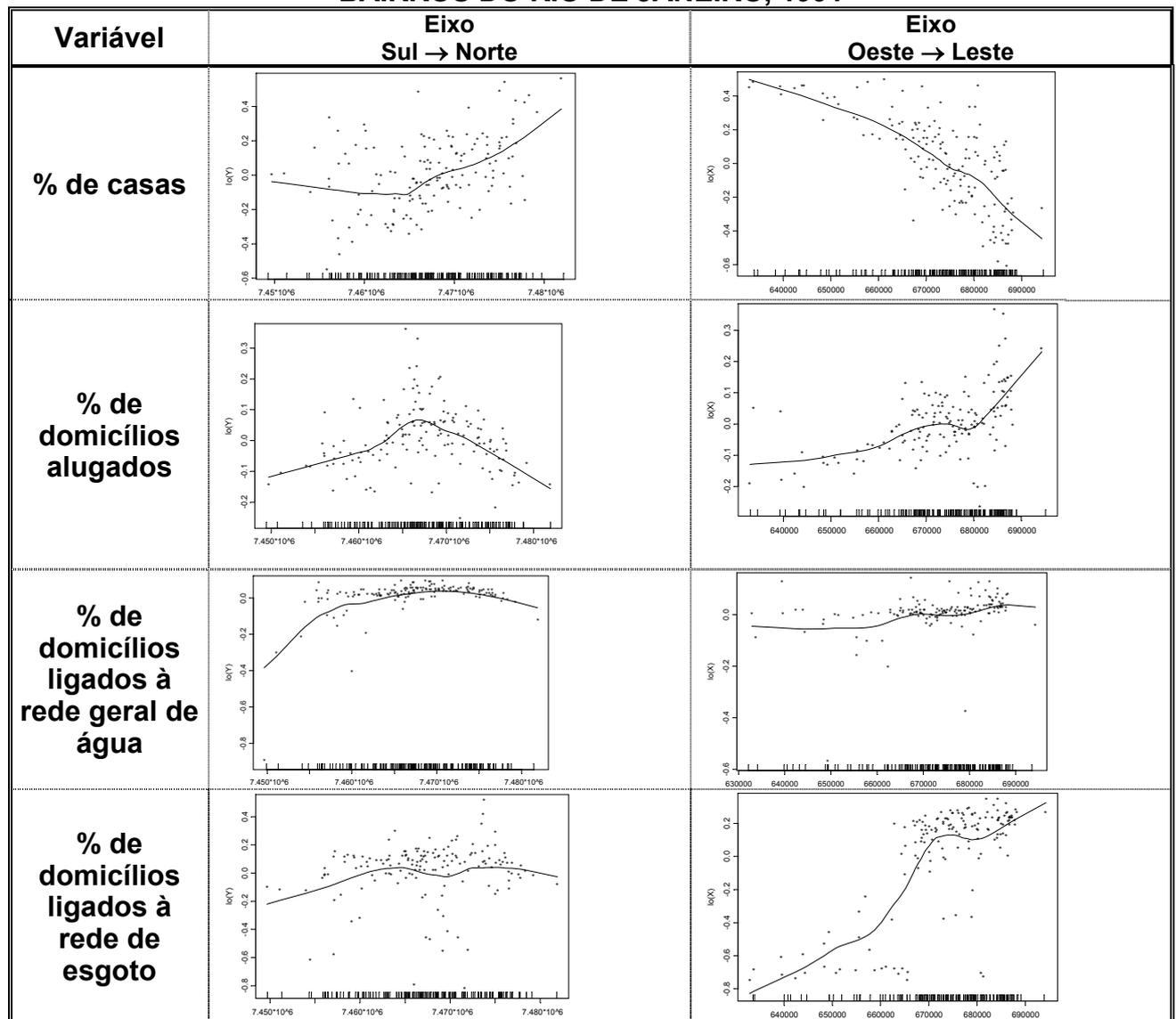
Observou-se para todos os indicadores que a média do processo não era constante, isto é, não havia estacionariedade de 1ª ordem. Os indicadores proporção de domicílios alugados, domicílios ligados à rede de esgotos, população alfabetizada, população idosa, chefes de domicílios com escolaridade acima do 1º Grau e renda superior a quinze salários mínimos apresentam forte tendência ascendente na direção Oeste → Leste: os bairros da Zona Leste, mais antigos da cidade (Ilha do Governador toda e Centro antigo) tem melhores indicadores.

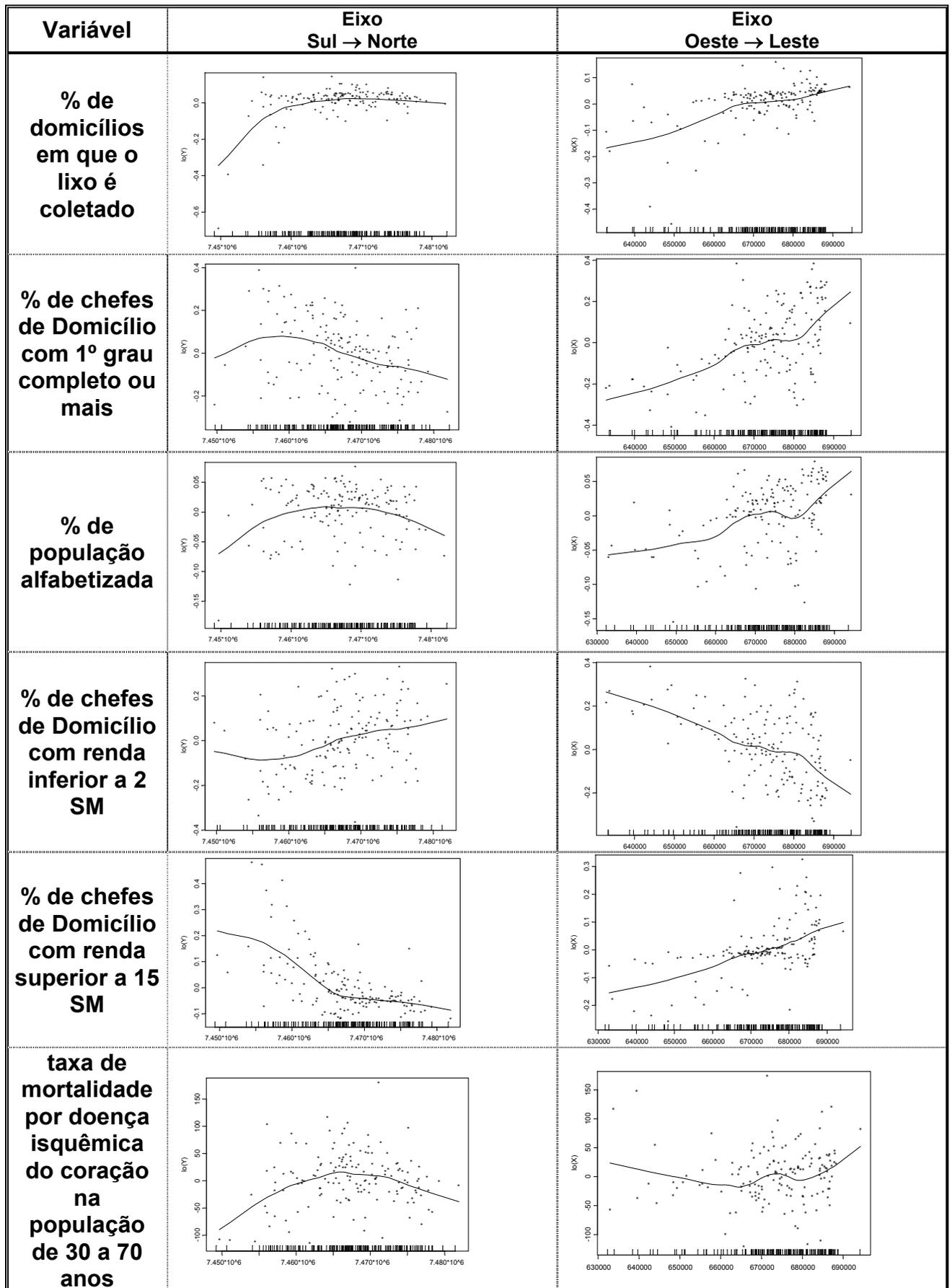
O comportamento no eixo Norte → Sul não é tão homogêneo: os domicílios alugados e a população idosa são mais freqüentes na região central do município. Por outro lado, a escolaridade, medida pelos dois indicadores tem maior dispersão, com maiores valores no Centro. A rede de esgotos não apresenta tendência neste eixo e a renda alta assume valores elevados na Zona Sul.

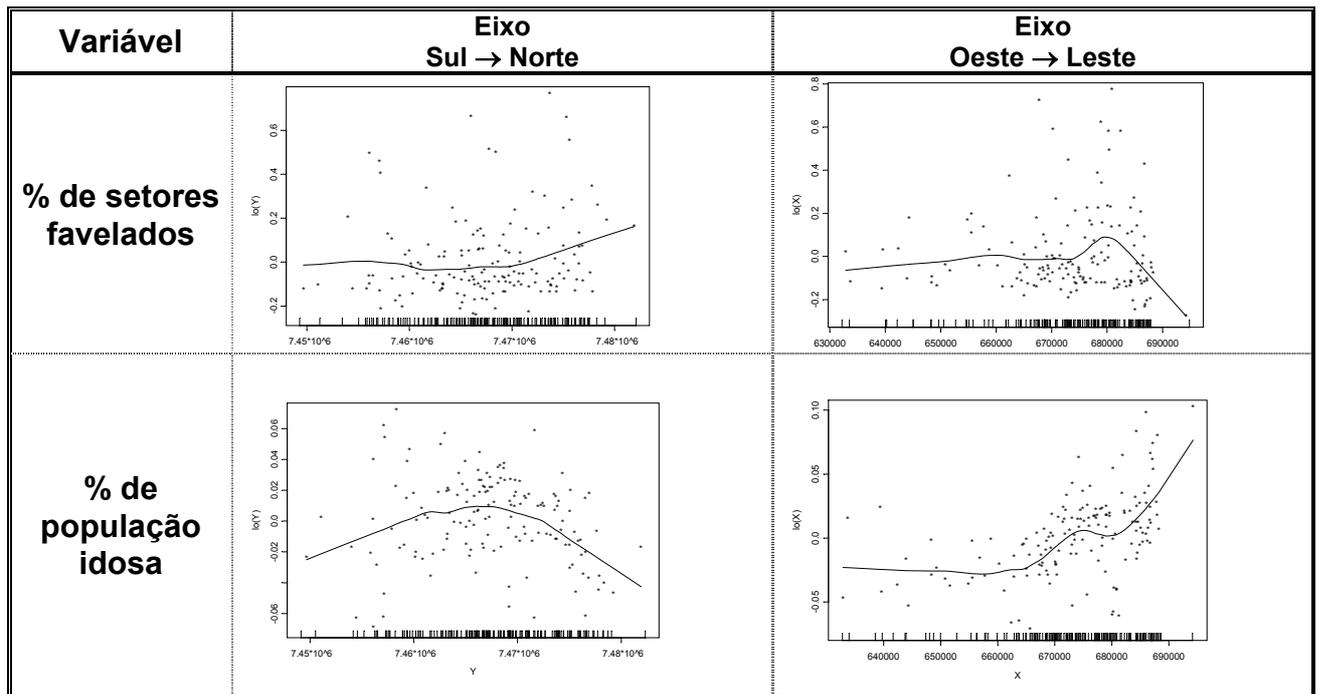
A rede de água e a coleta regular de lixo destacam-se por não apresentarem tendência no eixo Oeste → Leste, porém existem valores extremos (baixos) na Zona Oeste da cidade. A presença de favelas é maior na Região de Longitude 680.000, que inclui os bairros da Leopoldina e a favela da Rocinha. Os indicadores proporção de casas e proporção de chefes de domicílio com baixa renda apresentam valores mais elevados na Zona Oeste, diminuindo em direção ao Leste (Centro) e ao Sul.

A taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos não tem padrão claro, porém, considerando a região com maior número de bairros (entre longitude 670.000 e 690.000), pode-se observar uma tendência ascendente no eixo Oeste → Leste, isto é, as taxas aumentam em direção aos bairros do Centro, Ilha do Governador e ao longo da Avenida Brasil até Irajá.

QUADRO 4.3 - AVALIAÇÃO DA TENDÊNCIA ESPACIAL DAS VARIÁVEIS BAIROS DO RIO DE JANEIRO, 1991







O Índice de Moran foi recalculado para cada uma das variáveis utilizando-se os resíduos do modelo aditivo (Tabela 4.5). Conforme discutido anteriormente, o Índice de Moran se afasta de zero na presença de tendência. Mesmo retirando a tendência, este índice permanece significativo, indicando presença de autocorrelação espacial, para todas as variáveis exceto para a proporção de casas, proporção de setores favelados, proporção de população idosa e taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos.

TABELA 4.5 - ÍNDICE DE MORAN DOS INDICADORES SÓCIO ECONÔMICOS COM E SEM TENDÊNCIA - BAIRROS DO RIO DE JANEIRO

Indicadores	Bairro com Tendência		Bairro sem Tendência	
	Moran	p	Moran	p
Proporção de chefes de Domicílio com renda superior a 15 SM	0,556	0,000	0,218	0,000
Proporção de chefes de domicílio com 1º grau completo ou mais	0,385	0,000	0,190	0,000
Proporção de chefes de domicílio com renda inferior a 2 SM	0,376	0,000	0,189	0,000
Proporção de Domicílios em que o lixo é coletado	0,445	0,000	0,124	0,006
Proporção de Domicílios ligados à rede de esgoto	0,633	0,000	0,121	0,007
Proporção de população alfabetizada	0,311	0,000	0,116	0,010
Proporção de Domicílios alugados	0,484	0,000	0,108	0,011
Proporção de domicílios ligados à rede geral de água	0,355	0,000	0,090	0,019
Proporção de casas	0,514	0,000	0,051	0,125
Taxa de mortalidade por DIC na população de 30 a 70 anos	0,196	0,196	0,043	0,169
Proporção de setores favelados	0,093	0,029	0,020	0,286
Proporção de população idosa	0,227	0,000	0,038	0,184

4.7 - Regressão Espacial Sem Tendência

Como foi detectada a tendência espacial dos indicadores socioeconômicos, ajustou-se um novo modelo de regressão espacial. Ainda que, segundo o Índice de Moran, da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos não exista autocorrelação ou tendência espacial, o gráfico de tendência (Quadro 4.3), sugere não estacionariedade (a média da variável não é constante).

Na Tabela 4.6 temos os coeficientes de regressão do modelo CAR sem tendência ajustado. As variáveis significativas neste modelo são as mesmas dos modelos anteriores (regressão linear simples e CAR), onde os coeficientes de regressão indicam que a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros é menor no estrato social mais elevado. Por outro lado, a relação positiva com a escolaridade superior ao 1º grau e baixa renda, indica que

os valores mais altos da taxa de mortalidade estão associados aos bairros em que há uma combinação de chefes de domicílio com renda baixa e escolaridade alta.

A equação do modelo CAR sem tendência ajustado é:

$$Tx_DIC_{\text{bairro-RJ}}^{30-70} = -0,930$$

$$+ 791,267 * (\% \text{ de população idosa})$$

$$+ 214,257 * (\% \text{ chefes domicílio com 1º Grau completo ou mais})$$

$$+ 194,851 * (\% \text{ chefes domicílio com renda inferior a 2 SM})$$

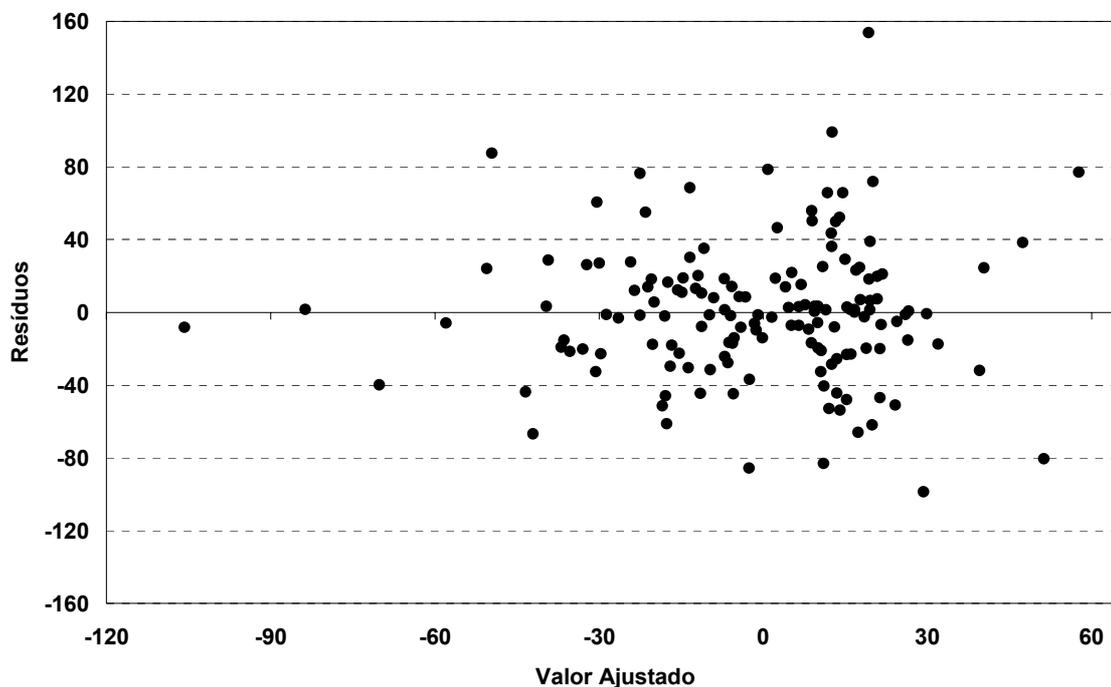
$$- 197,986 * (\% \text{ chefes domicílio com renda superior a 15 SM})$$

TABELA 4.6 - COEFICIENTES DE REGRESSÃO DO MODELO CAR SEM TENDÊNCIA

	Coefficiente	Erro padrão	Estatística t	p
Intercepto	-0,930	3,902	-0,238	0,8111
Proporção da população idosa	791,267	138,368	5,718	0,000
Proporção de chefes de domicílio com 1º grau completo ou mais	214,257	89,583	2,391	0,018
Proporção de chefes de domicílio com renda inferior a 2 SM	194,851	80,327	2,425	0,016
Proporção de chefes de domicílio com renda superior a 15 SM	-197,986	65,756	-3,010	0,003

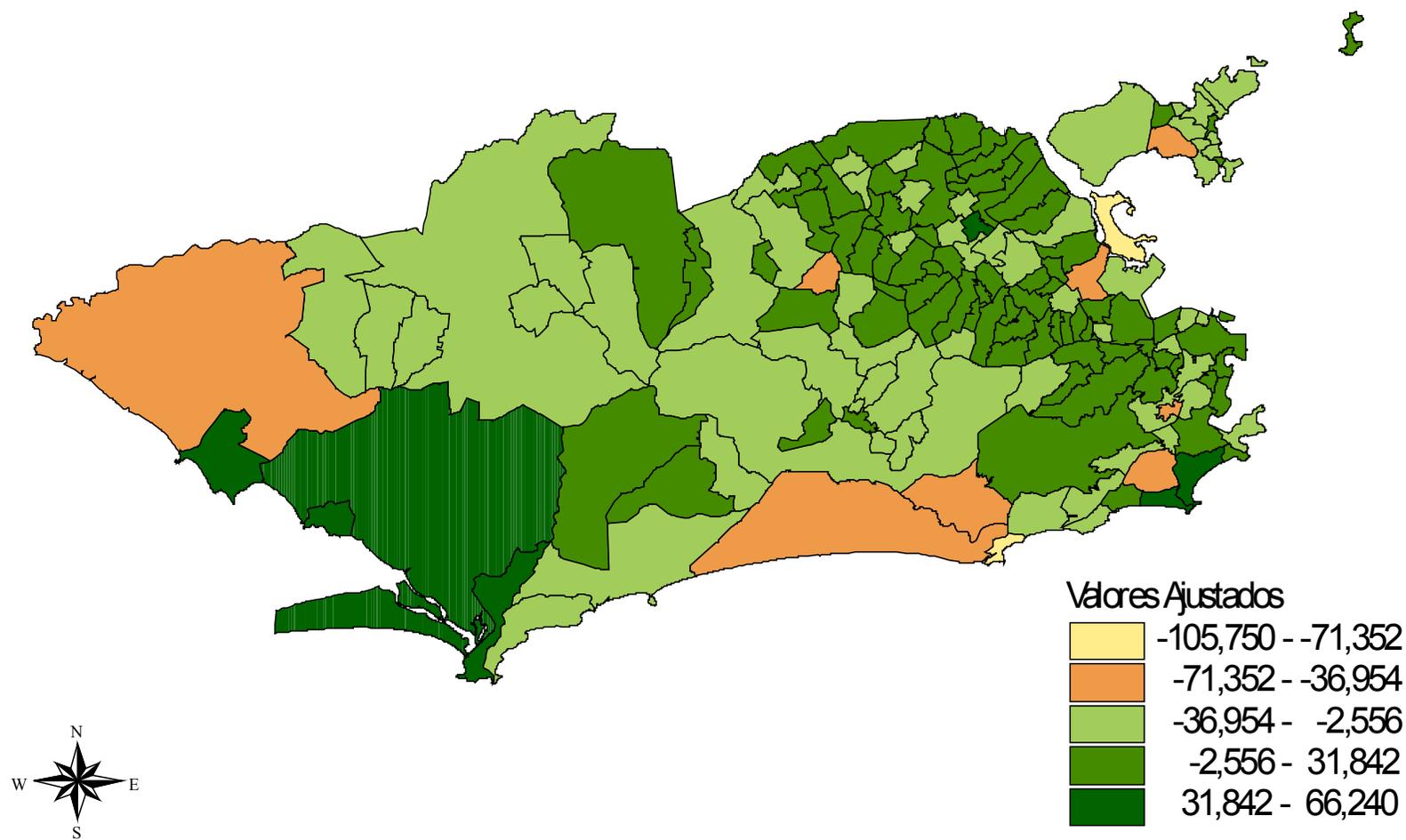
Segundo o teste de correlação dos resíduos, não há evidências de correlação espacial (Moran I = -0,072, p = 0,921). Além deste teste, a análise dos resíduos foi executada através do Gráfico 4.21 e do Mapa 4.20. No Gráfico 4.21, observamos que a dispersão dos resíduos ainda lembra um cone.

GRÁFICO 4.21 - DISPERSÃO DOS RESÍDUOS DO MODELO CAR SEM TENDÊNCIA EM RELAÇÃO AOS VALORES AJUSTADOS BAIROS DO RIO DE JANEIRO

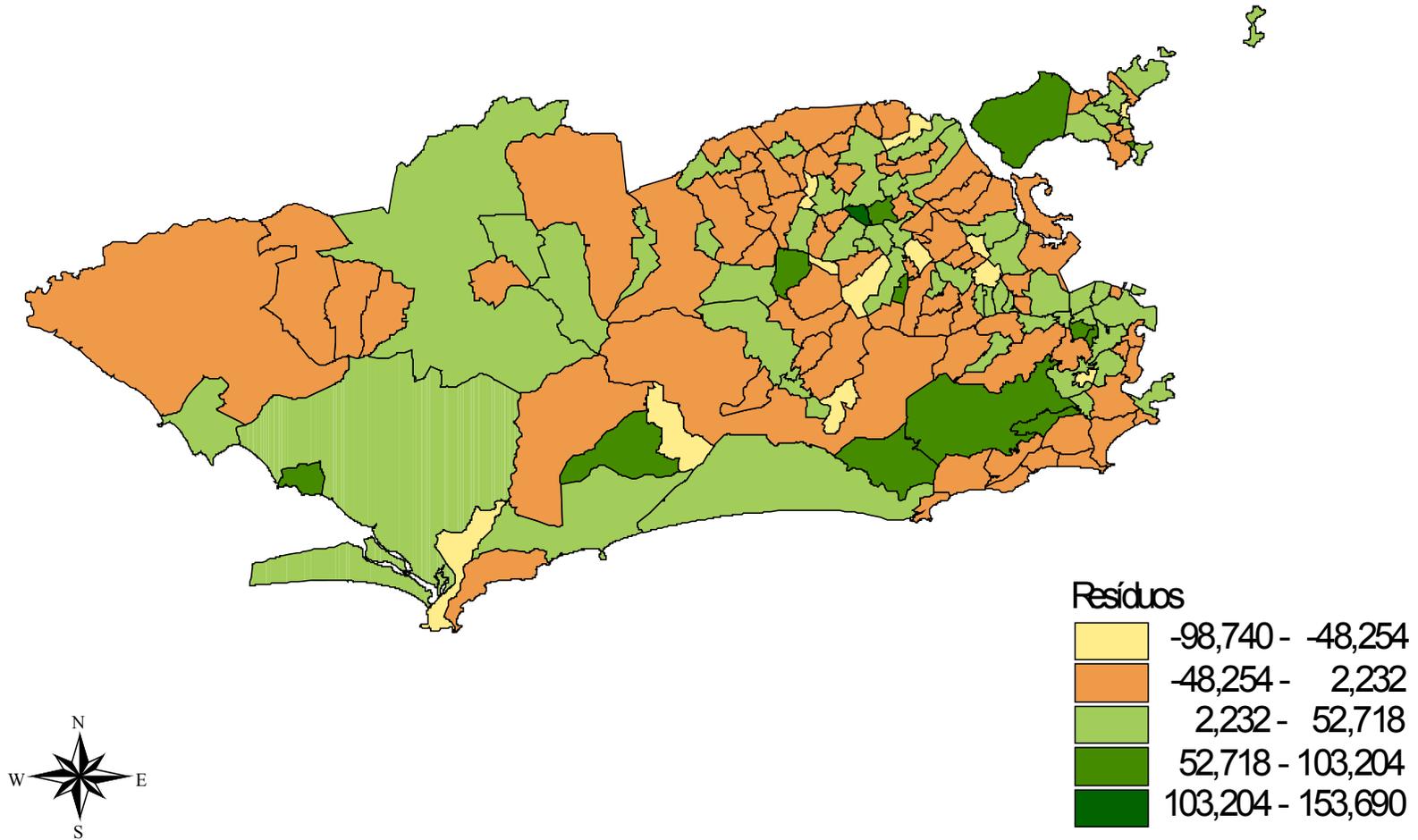


Com a retirada da tendência, a distribuição espacial da taxa ajustada pelo modelo CAR se modifica. Observe que a concentração dos valores elevados não se dá da mesma forma que nos modelos anteriores (ver Mapa 4.19).

MAPA 4.19 - DISTRIBUIÇÃO ESPACIAL DAS TAXAS AJUSTADAS DE MORTALIDADE POR DIC NO GRUPO ETÁRIO 30 A 70 ANOS - MODELO CAR SEM TENDÊNCIA BAIROS DO RIO DE JANEIRO, 1991



**MAPA 4.20 - DISTRIBUIÇÃO ESPACIAL DOS RESÍDUOS DO MODELO CAR AJUSTADO SEM TENDÊNCIA
BAIRROS DO RIO DE JANEIRO, 1991**



5 - CONSIDERAÇÕES FINAIS

Neste capítulo é apresentada a síntese das análises realizadas para setores censitários e bairros. Os métodos utilizados para o estudo da relação entre a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos e indicadores socioeconômicos, no município do Rio de Janeiro em 1991 sugerem que os valores mais elevados da taxa estão associados ao estrato social médio, classificados como B e C na seção 4.2.

A análise univariada permitiu identificar comportamento diferenciado em bairros e setores censitários. A flutuação dos indicadores, típica de pequenas áreas, é maior nos setores censitários. O Índice de Moran também tende a ser maior entre setores censitários, em parte graças às pequenas distâncias entre estes e devido ao efeito da agregação destes em bairros que, no Rio de Janeiro, apresentam um padrão de ocupação do espaço em mosaico. Áreas ricas coexistem ao lado de favelas, são poucos os bairros do Rio de Janeiro que não apresentam setores censitários favelados, que são um indicativo de precárias condições de vida. Por outro lado, mesmo nas regiões mais pobres, por exemplo na Zona Oeste, o fenômeno inverso pode ser observado: regiões de extrema pobreza são intercaladas com setores censitários mais ricos. Assim, os indicadores são médias, e a diminuição do escopo da variação diminui também a capacidade do Índice de detectar cluster espacial. Um desdobramento interessante seria a comparação deste índice com o Moran I_{pop} — Moran ajustado para densidade populacional — que permite também a detecção de cluster intra-áreas (Cruz, 1996). Além disso, em quase todos os mapas pode-se observar tendência nos indicadores, usualmente melhores na Zona Sul e Centro, piorando em direção à Zona Oeste.

No caso da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos, o efeito de pequenas áreas gerando grande quantidade de setores censitários com zero óbito impede a análise. Um desdobramento deste trabalho será agregar setores censitários semelhantes quando a indicadores socioeconômicos, até obter um contingente populacional que permita estimar a taxa de mortalidade, mantendo a homogeneidade interna das áreas criadas (Carvalho e cols., 1995).

A análise da relação entre a taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos e indicadores socioeconômicos selecionados, na unidade geográfica bairros, utilizando diferentes métodos de regressão (simples, espacial com erros autocorrelacionados com e sem tendência espacial), permite validar as associações encontradas, semelhantes nas diferentes técnicas.

Na regressão linear simples, o R^2 obtido, indica que o modelo ajustado explica uma parcela importante da variabilidade da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros. Os fatores de risco associados ao estilo de vida, poderiam influenciar o modelo. Entretanto, estes, por exemplo indicadores relacionados ao consumo de cigarros e hábitos alimentares, são de difícil mensuração a nível ecológico, sendo estimados usualmente apenas para regiões isoladas do Brasil.

Para os modelos de regressão espacial, não existe uma medida resumo como o R^2 . Contudo, como a suposição básica destes modelos é a existência de autocorrelação e/ou tendência, ao se controlar estes elementos da variabilidade, eventualmente a associação detectada desaparece (Huel e cols., 1978). Não foi o caso dos ajustes obtidos neste trabalho, onde se mantiveram as mesmas variáveis em todos os ajustes.

Controlando-se pela proporção de população idosa, as variáveis socioeconômicas significativas nos três modelos de regressão relacionam mortalidade por doença isquêmica do coração na população de 30 a 70 anos com renda e escolaridade.

A proporção de chefes de domicílio com renda superior a quinze salários mínimos é inversamente correlacionada à taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros, sugerindo que a alta renda apresenta um "efeito protetor" sobre a mortalidade por doença isquêmica do coração na população de 30 a 70 anos. A faixa de renda alta adotada abrange uma parcela importante da população de profissionais de nível superior, empresários, administradores, executivos, entre outros. Este grupo usualmente dispõe de assistência médica suplementar, que em 1996 era composta por 1.180 empresas que comercializam um grande número de produtos, com faturamento anual estimado em 14,80 bilhões de dólares e cerca de 42,3 milhões de beneficiários (Almeida, 1998).

A correlação positiva entre baixa renda e mortalidade por doença isquêmica do coração na população de 30 a 70 anos, reforça o achado anterior. Entretanto, a correlação também positiva entre escolaridade superior ao 1º grau — mais freqüente nas populações de maior poder aquisitivo — sugere que além de atingir os setores censitários mais pobres da população, com alta proporção de chefes de domicílio com renda inferior a dois salários mínimos, o indicador de mortalidade aumenta também nos grupos de alta escolaridade. Na população idosa, a educação ginasial (1º Grau completo) ainda era pouco comum. Este nível de escolaridade sugere relação entre mortalidade por doença isquêmica do coração na população de 30 a 70 anos e estratos sociais médios.

Na literatura, é comum encontrarmos referências que indicam que a ocorrência da doença isquêmica do coração na população é influenciada pelo "estilo de vida", sendo assim, a hipótese básica é que ela ocorre de forma diferenciada nos estratos sociais. Se observarmos simultaneamente os mapas referentes à distribuição da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros (Mapa 4.10) e os grupos encontrados na classificação multivariada (Mapa 4.14), percebemos que os bairros que apresentam as maiores taxas, foram classificados nos estratos sociais médios (B e C). Quando adicionamos à este resultado, os resultados do modelo de regressão linear simples (Mapa 4.15) e do modelo CAR (Mapa 4.17), verificamos que os valores ajustados mais elevados

também concentram-se nesta região. Os valores ajustados do modelo CAR sem tendência (Mapa 4.19) distribuem os valores mais altos da taxa de mortalidade pelo município, mas ainda assim há uma concentração na Zona Leste do município.

Este estudo reafirma a importância dos indicadores socioeconômicos de renda e escolaridade na determinação das altas taxas de mortalidade por doença isquêmica do coração nos bairros do Rio de Janeiro. Mesmo ajustando os modelos pela proporção de população idosa e pela estrutura espacial, tanto local — autocorrelação modelada através do CAR — como global — retirada de tendência através do *Loess* — a mortalidade por doença isquêmica do coração na população de 30 a 70 anos é menor nos estratos sociais elevados e influenciada positivamente e com peso semelhante pela concentração de chefes de domicílio com baixa renda e alta escolaridade.

6 - BIBLIOGRAFIA

ALMEIDA, C., 1998. *O Mercado Privado de Serviços de Saúde no Brasil: Panorama Atual e Tendências da Assistência Médica Suplementar*. Brasília, Instituto de Pesquisa Econômica Aplicada.

Anuário Estatístico 95-97 Cidade do Rio de Janeiro. CD-ROM. Rio de Janeiro.

BAILEY, T.C., GATRELL, A.C., 1995. *Interactive Spatial Data Analysis*. Essex: Longman Scientific & Technical.

BEAGLEHOLE, R., 1990. International Trends in Coronary Heart Disease Mortality, Morbidity, and Risk Factors. *Epidemiologie Reviews*, 12:1-15.

CARVALHO, M. S., 1997. *Aplicação de Métodos de Análise Espacial na Caracterização de Áreas de Risco à Saúde*. Dissertação de Doutorado, Rio de Janeiro: Engenharia Biomédica, Universidade Federal do Rio de Janeiro.

CARVALHO, M. S., CRUZ, O. G., 1998. Análise Espacial por Micro-Áreas: Métodos e Experiências. In: *Epidemiologia Contextos e Pluralidade* (Veras, R. P., org.), pp.79-89, Rio de Janeiro: Editora Fiocruz/Abrasco.

CARVALHO, M. S., CRUZ, O. G., NOBRE, F. F., 1997. Perfil de Risco: Método Multivariado de Classificação Socioeconômica de Microáreas Urbanas — Os Setores Censitários da Região Metropolitana do Rio de Janeiro. *Cadernos de Saúde Pública*, 4: 435-445.

CARVALHO, M. S., CRUZ, O. G., NOBRE, F. F., 1995. The use of multivariate cluster analysis and aggregation algorithm to propose a new spatial partitioning of

Rio de Janeiro/Brazil. In: Proceedings of the CDC/ATSDR Symposium on Statistical methods, p.18, Atlanta.

CASTELLI, W., 1989. Determinants of CHD Mortality: Genetic, Socioeconomic, Lifestyle and Risk Factor Influences — an Over View. *International Journal of Epidemiology*, 3:S180-S182.

CHOR, D., 1997. *Perfil de Risco Cardiovascular de Funcionários de Banco Estatal*. Dissertação de Doutorado, São Paulo: Faculdade de Saúde Pública, Universidade de São Paulo.

CLEEK, R. K., 1979. Cancers and Environment: the Effect of Scale. *Soc. Sci. & Med.*, 3D: 241-247.

CRESSIE, N., 1991. *Statistics for spatial data*. New York: Wiley Interscience.

CRUZ, O. G., 1996. *Homicídios no Estado do Rio de Janeiro: Análise da Distribuição Espacial e Sua Evolução*. Dissertação de Mestrado, São Paulo: Faculdade de Saúde Pública, Universidade de São Paulo.

FIBGE (Fundação Instituto Brasileiro de Geografia e Estatística), 1990. *X Recenseamento Geral do Brasil. Questionário Básico*. Rio de Janeiro: FIBGE.

FIBGE (Fundação Instituto Brasileiro de Geografia e Estatística), 1993. *Censo Demográfico 1991 - Resultado do Universo Relativo às Características da População e dos Domicílios*. Rio de Janeiro: FIBGE.

HAINING, R., 1990. *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge: Cambridge University Press.

HASTIE, T. J., TIBSHIRANI R. J., 1990. *Generalized Additive Models*. Chapman and Hall.

HUEL, G., THOMAZEAU, R., DERRIENNIC, F., LAZAR, P., 1978. Dureté de l'eau et mortalité cardiovasculaire. analyse portant sur 947 communes alsaciennes. *Rev. Epidém. et Santé Publ.*, 26: 381 - 390.

KEYS, A., 1970. Coronary Heart Disease in Seven Countries. *Circulation*, suppl. I, 41: 1-211.

MARMOT, M., 1989. Socioeconomic Determinants of CHD Mortality. *International Journal of Epidemiology*, 3:S196-S202.

NOBRE, F. F., CARVALHO, M. S., 1994. Spatial and Temporal Analysis of Epidemiological Data in GIS for Health and the Environment. In: Proceedings of an International Workshop Held in Colombo. Sri Lanka.

PINA, MF., CABRAL, R., ARGENTO, R., CARRIJO, R., LUA, W., 1999. Georeferenciamento do Banco de Mortalidade de 1991 no Município do Rio de Janeiro. In: XIX Congresso Brasileiro de Cartografia, Pernambuco.

ROSE, G., 1989. Causes of Trends and Variations in CHD Mortality in Different Countries. *International Journal of Epidemiology*, supp. S174-S179.

SANCHEZ, D., BUZAI, G. D. Consideraciones de la Forma y el Tamaño de las Unidades Espaciales de Information. Una Aplicación SIG a la Argentina, pp. 465-484.

SCHENKMAN, S., 1999. *Tendência da Mortalidade por Doença Isquêmica do Coração e de seus Determinantes. Estudo Ecológico de Série Temporal, Município de São Paulo, 1941-1996*. Dissertação de Mestrado, São Paulo: Escola Paulista de Medicina, Universidade Federal de São Paulo.

STAMLER, J., 1992. Established Major Coronary Risk Factors. In: *Coronary Heart Disease Epidemiology: From aetiology to public health* (Marmot, M. G. e Elliot, P.). pp. 35-66, Oxford University, Press.

STRUYF A., HUBERT M., ROUSSEEUW, P. J., 1997. Integrating Robust Clustering Techniques in S-PLUS. *Computational Statistics and Data Analysis*, 26:17-37.

UEMURA, K. e PISA, Z., 1985. Recent Trends in Cardiovascular Disease Mortality in 27 Industrialized Countries. *World Health Statistics Quarterly*, 38:142-149.

ANEXO I - ANÁLISE EXPLORATÓRIA UNIVARIADA DOS INDICADORES NA UNIDADE GEOGRÁFICA SETOR CENSITÁRIO

Nos Gráficos AI.1 a AI.11 e Mapas AI.1 a AI.11 estão apresentados os resultados da análise exploratória univariada dos indicadores socioeconômicos e da taxa de mortalidade por doença isquêmica do coração na população de 30 a 70 anos nos bairros nos setores censitários que foram utilizados para comparação com a escala Bairros.

GRÁFICO AI.1 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CASAS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991

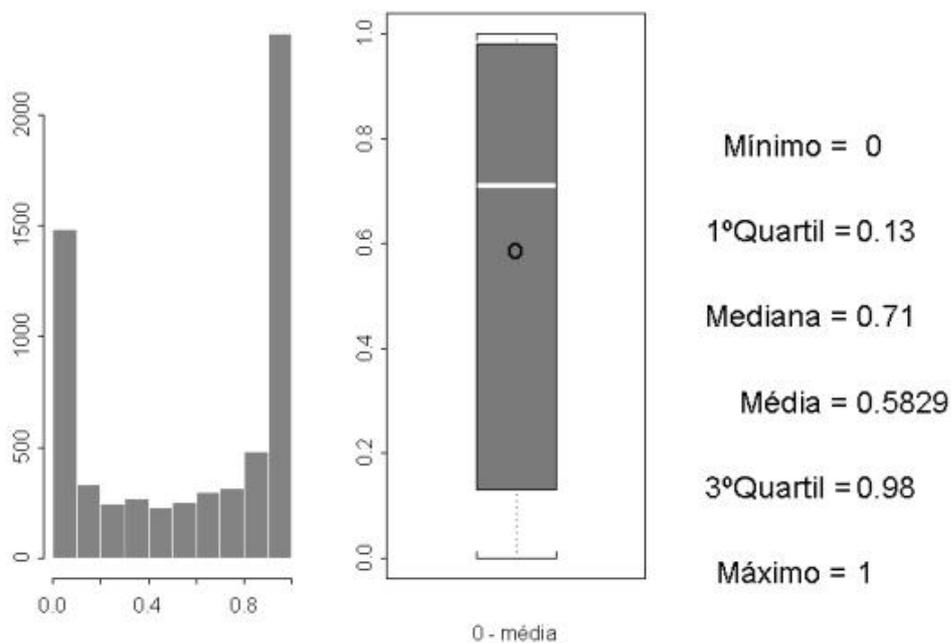


GRÁFICO AI.2 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991

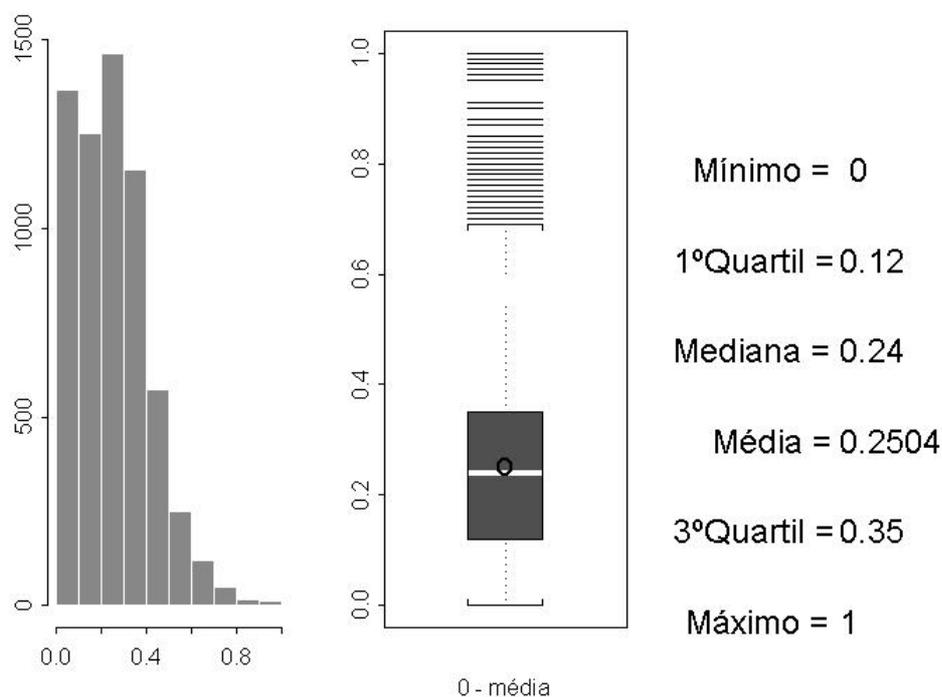
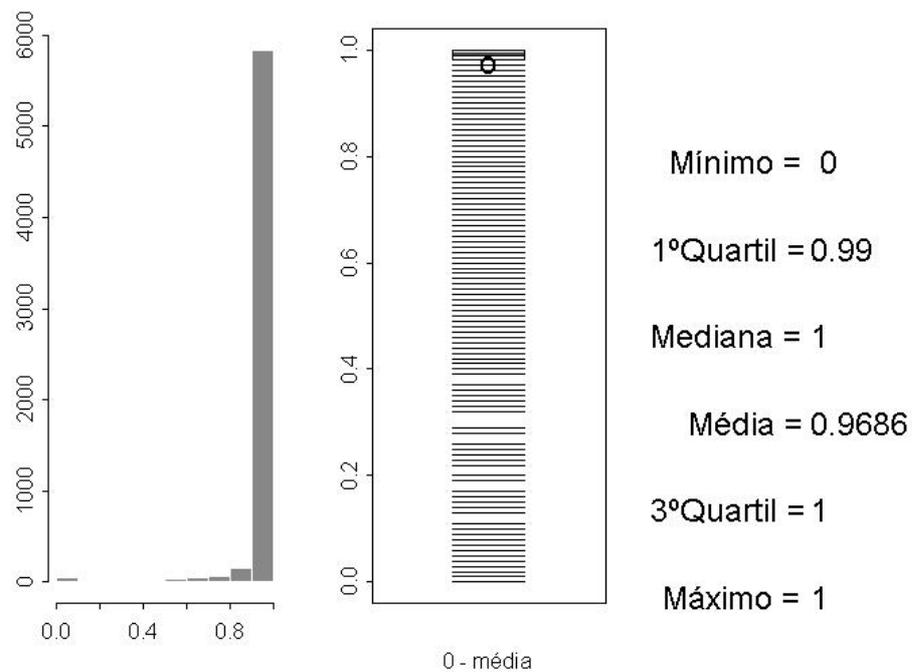
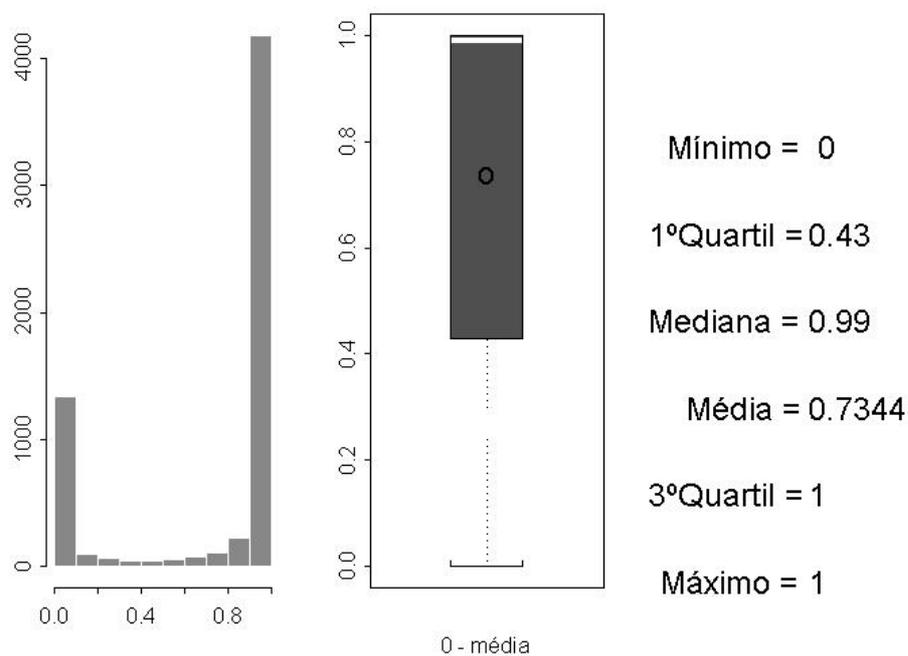


GRÁFICO AI.3 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991



**GRÁFICO AI.4 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS
LIGADOS À REDE PÚBLICA DE ESGOTO
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**



**GRÁFICO AI.5 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE DOMICÍLIOS
EM QUE O LIXO É COLETADO
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**

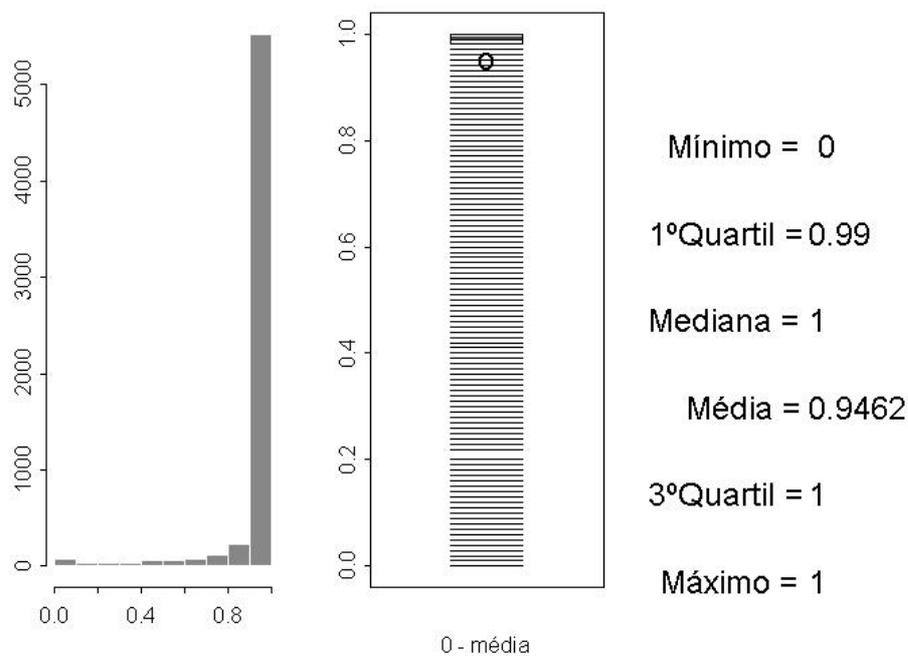


GRÁFICO AI.6 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991

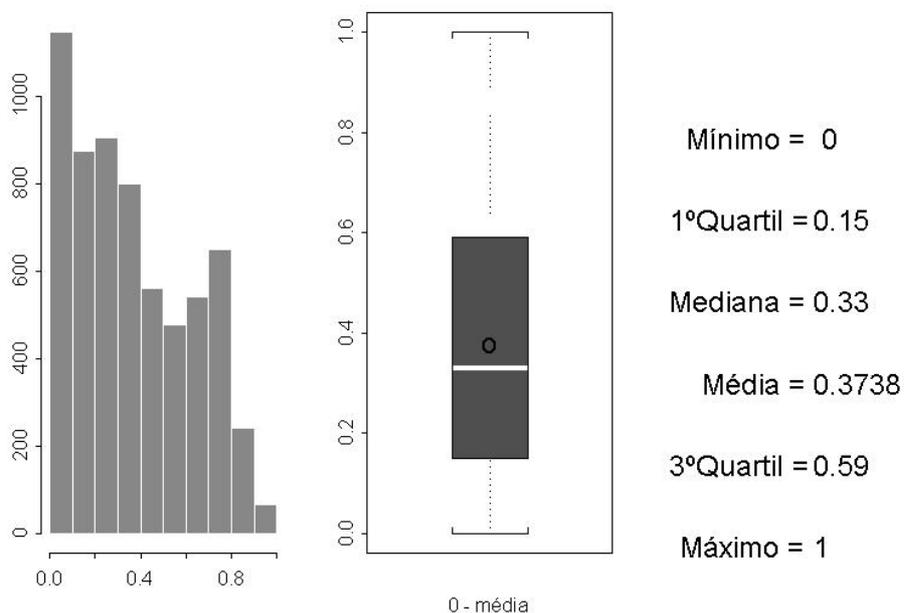


GRÁFICO AI.7 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991

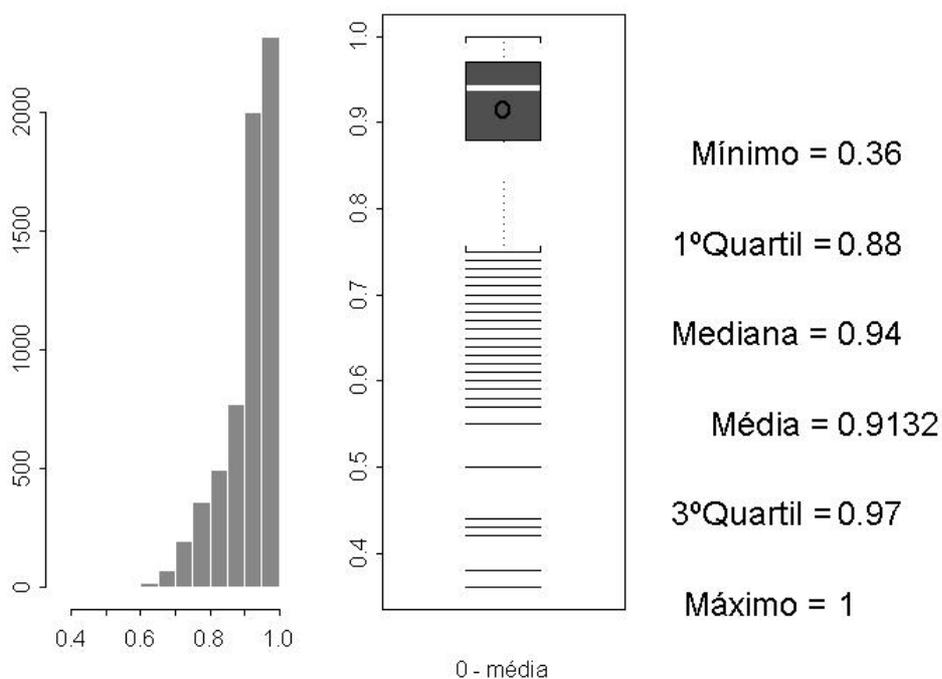


GRÁFICO AI.8 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991

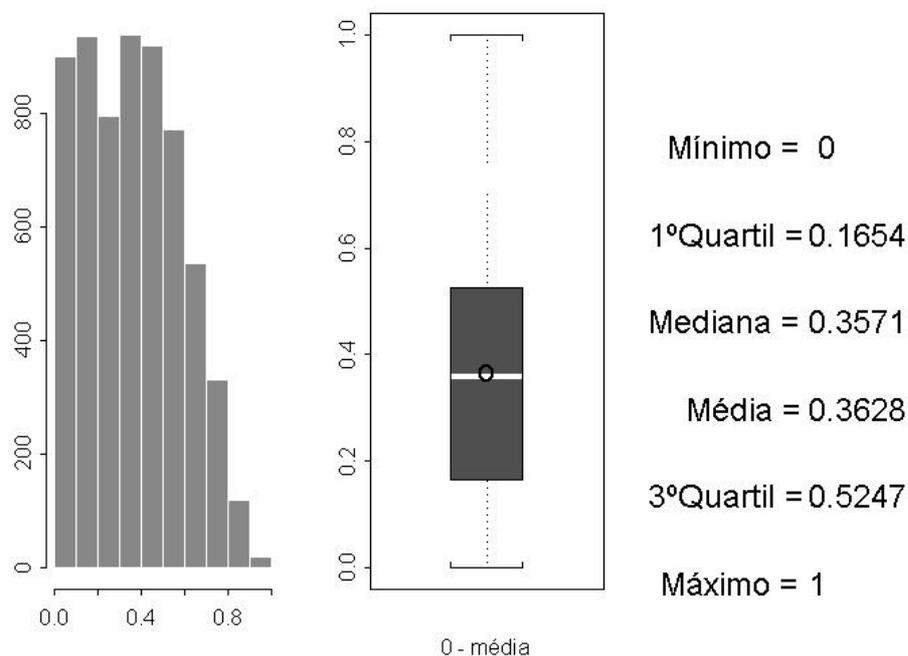


GRÁFICO AI.9 - ANÁLISE EXPLORATÓRIA DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991

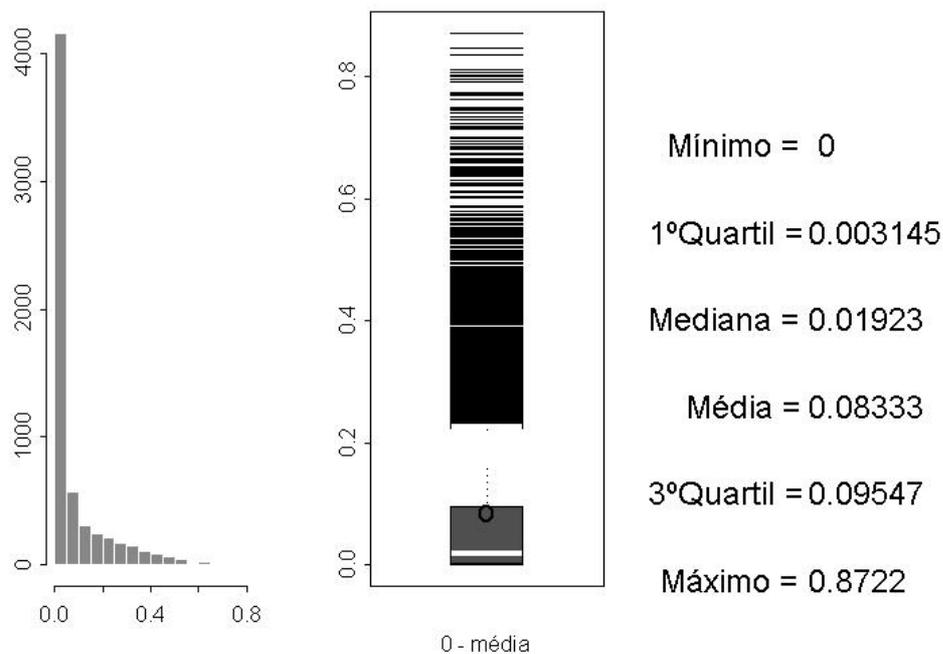


GRÁFICO AI.10 - ANÁLISE EXPLORATÓRIA DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991

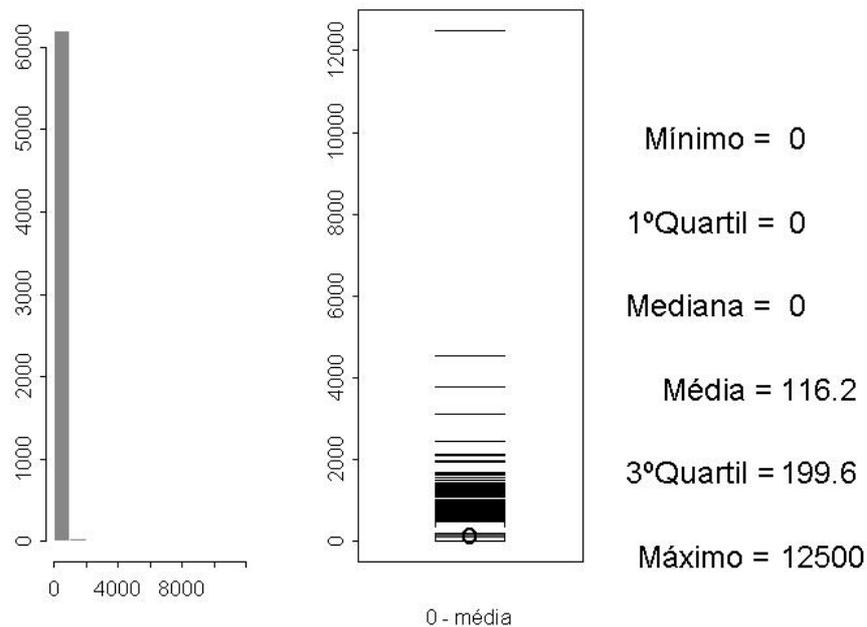
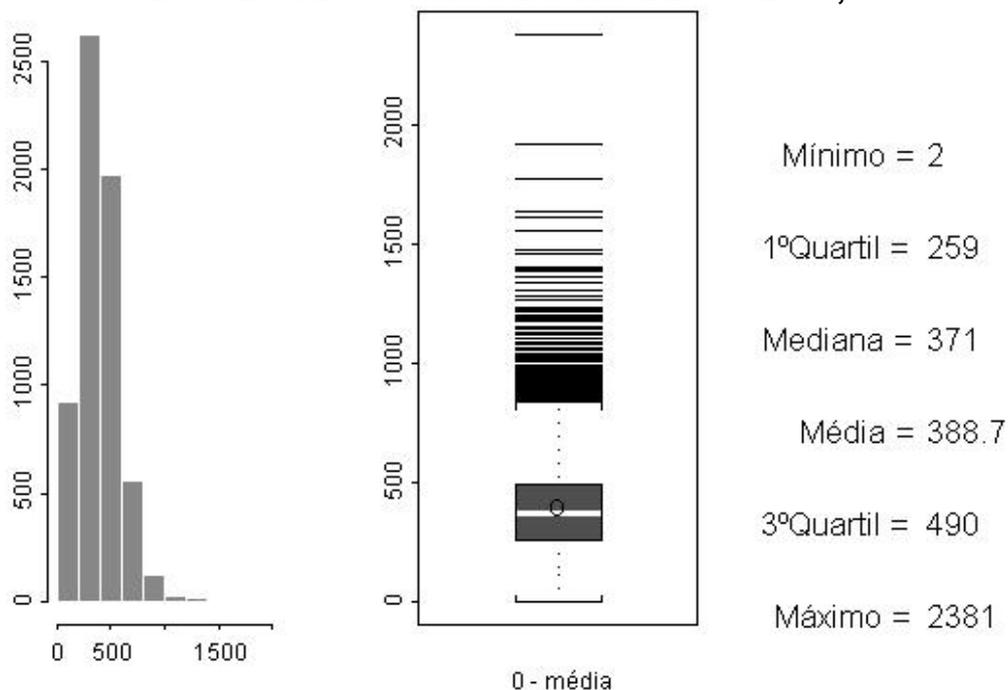
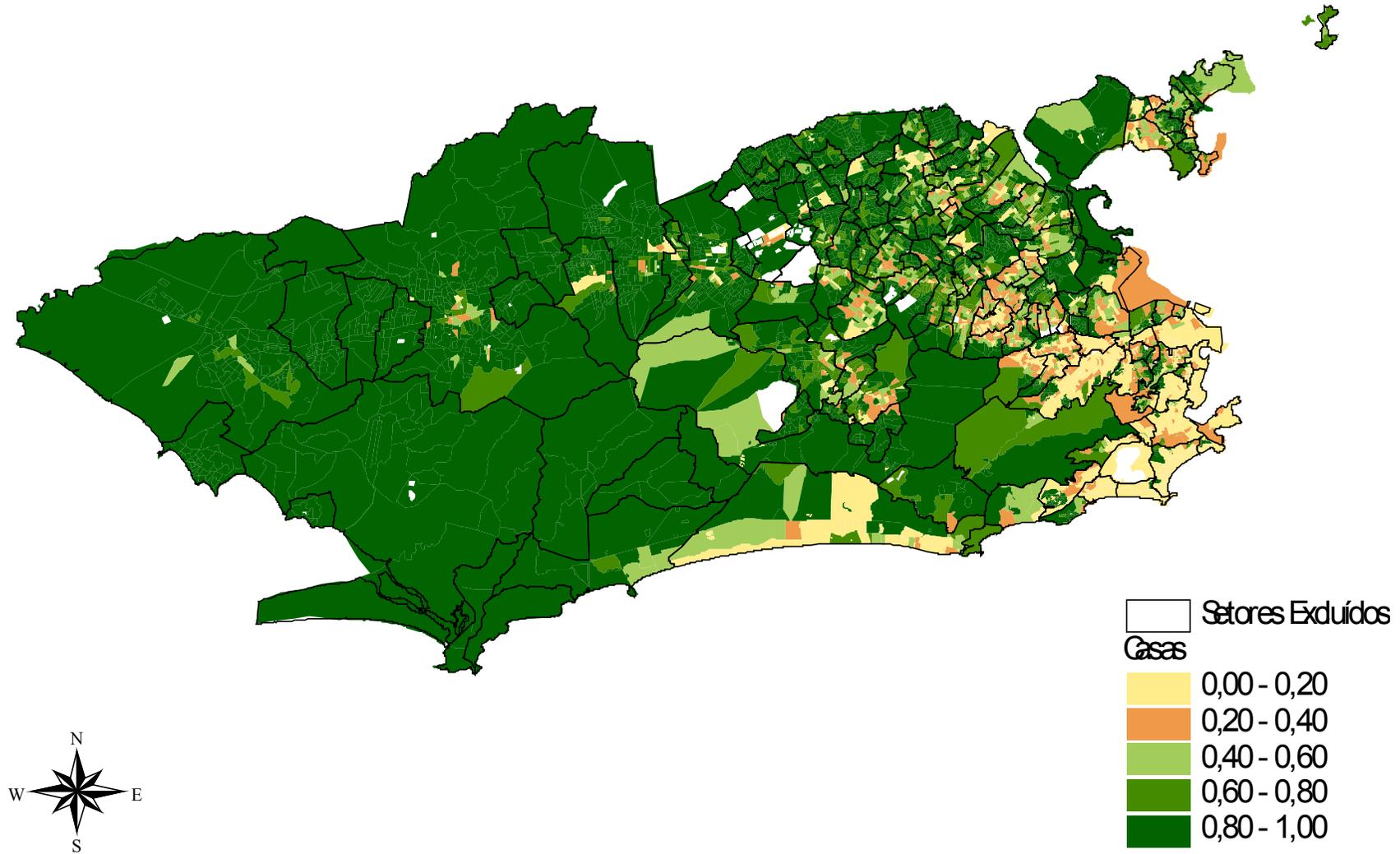


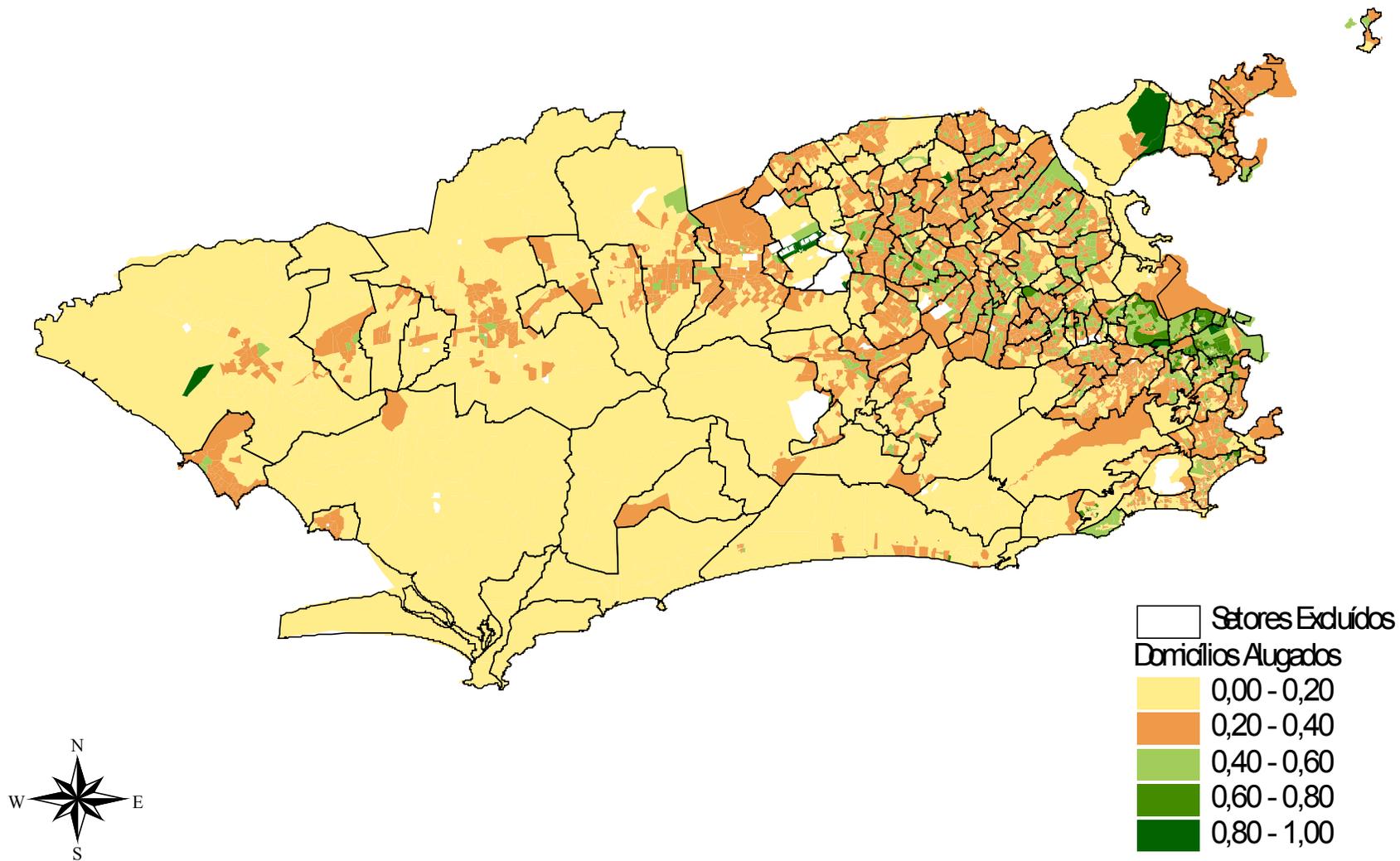
GRÁFICO AI.11 - ANÁLISE EXPLORATÓRIA DA POPULAÇÃO DE 30 A 70 ANOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991



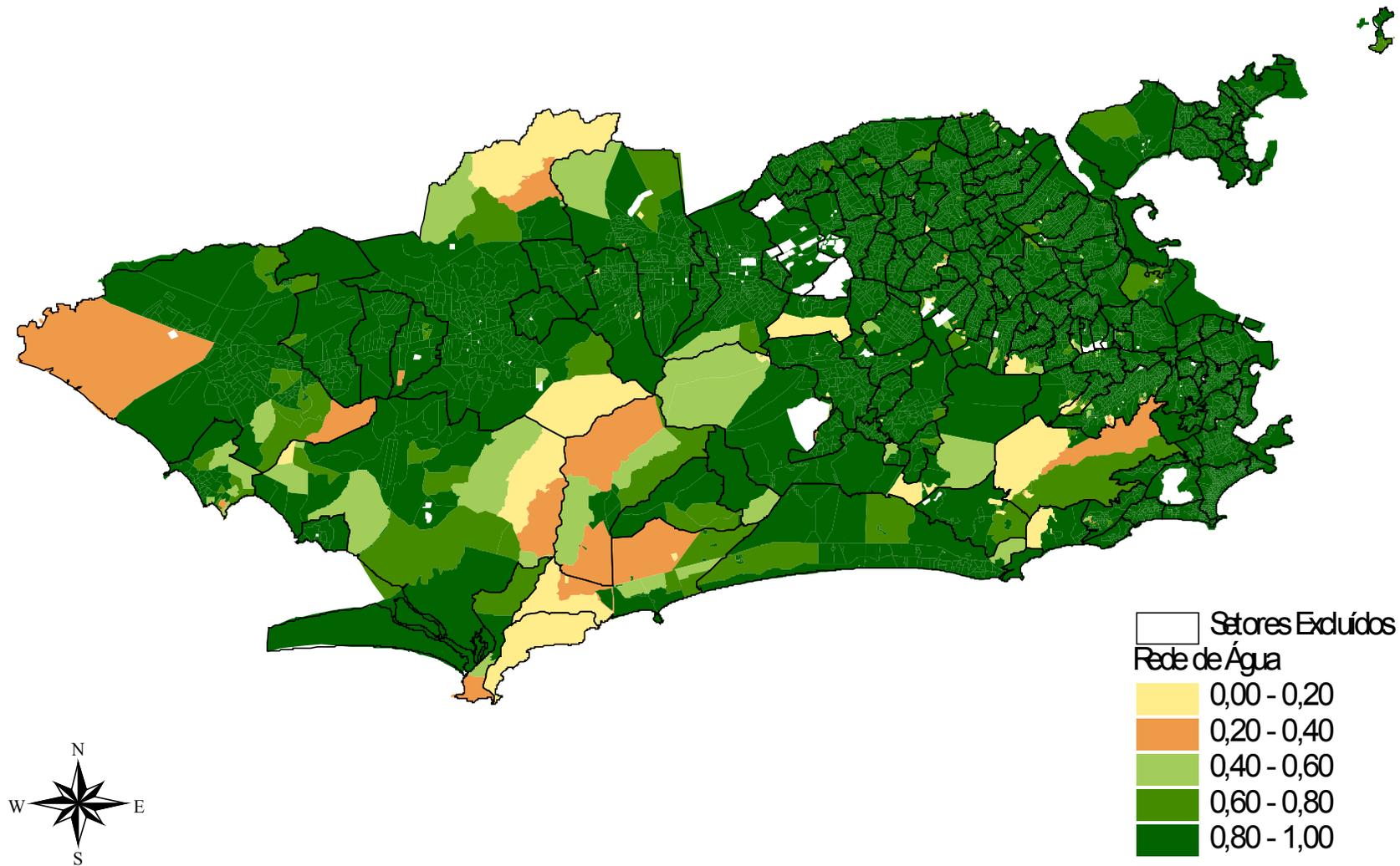
MAPA AI.1 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CASAS
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991



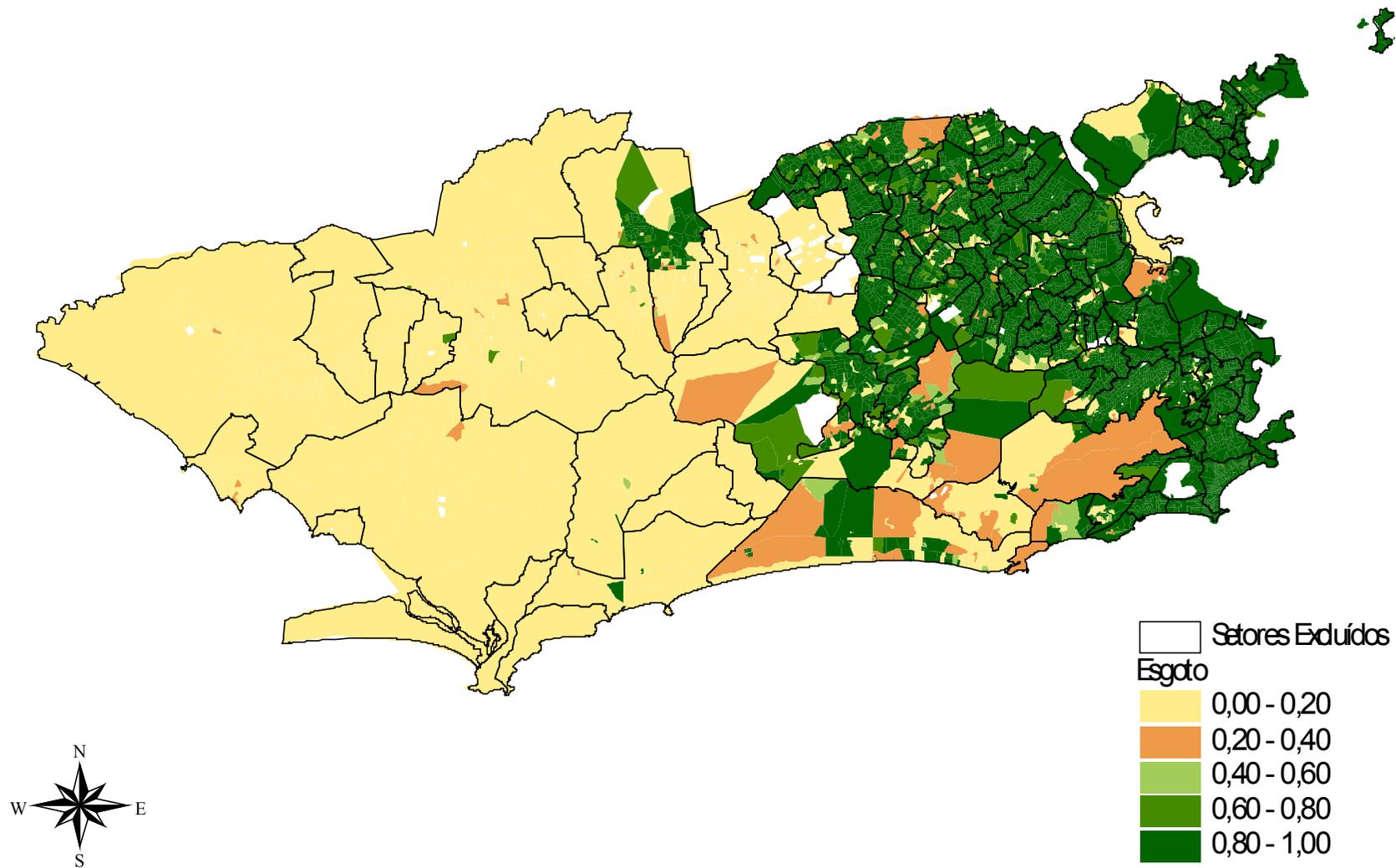
**MAPA AI.2 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS ALUGADOS
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**



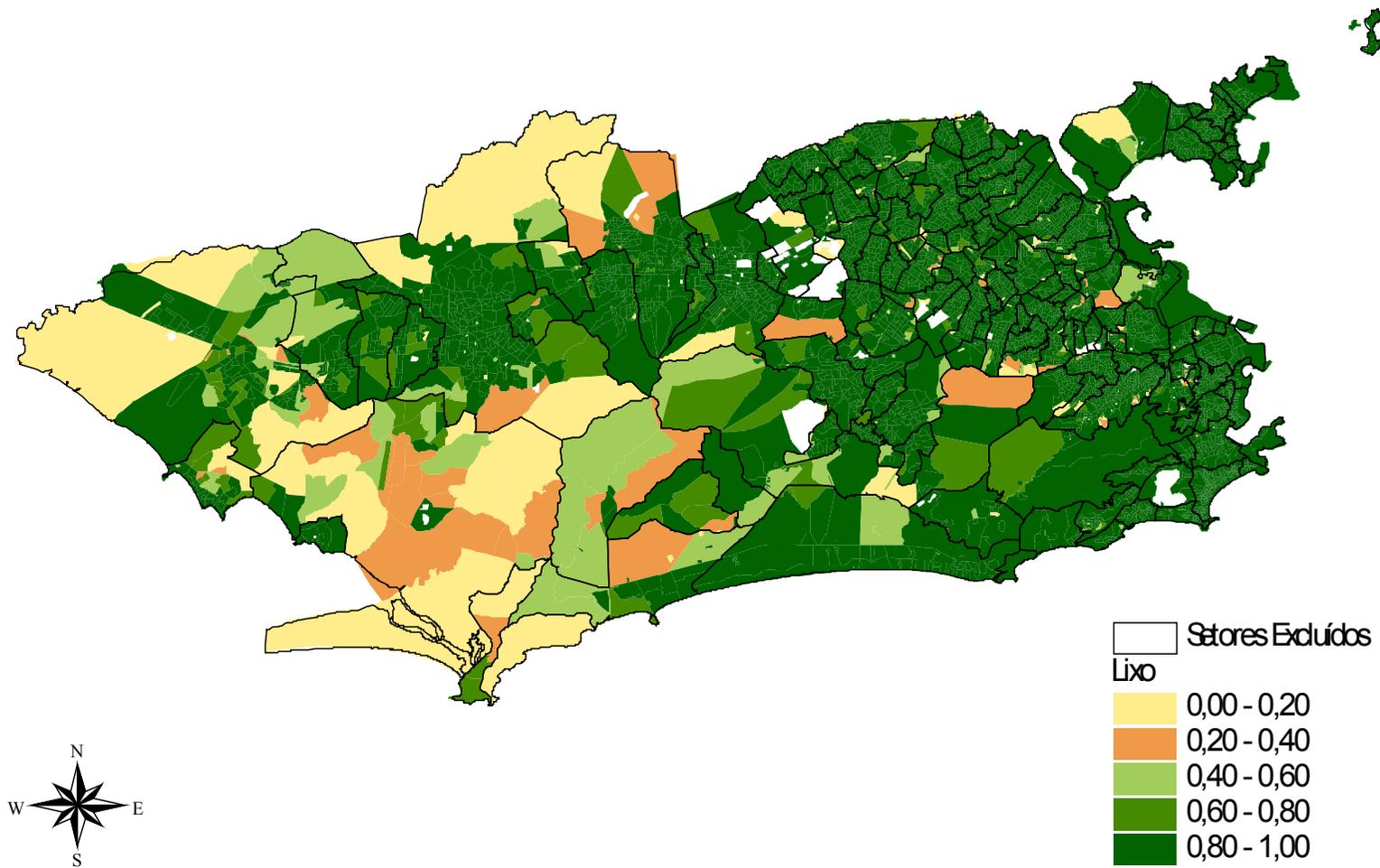
**MAPA AI.3 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE GERAL DE ÁGUA
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**



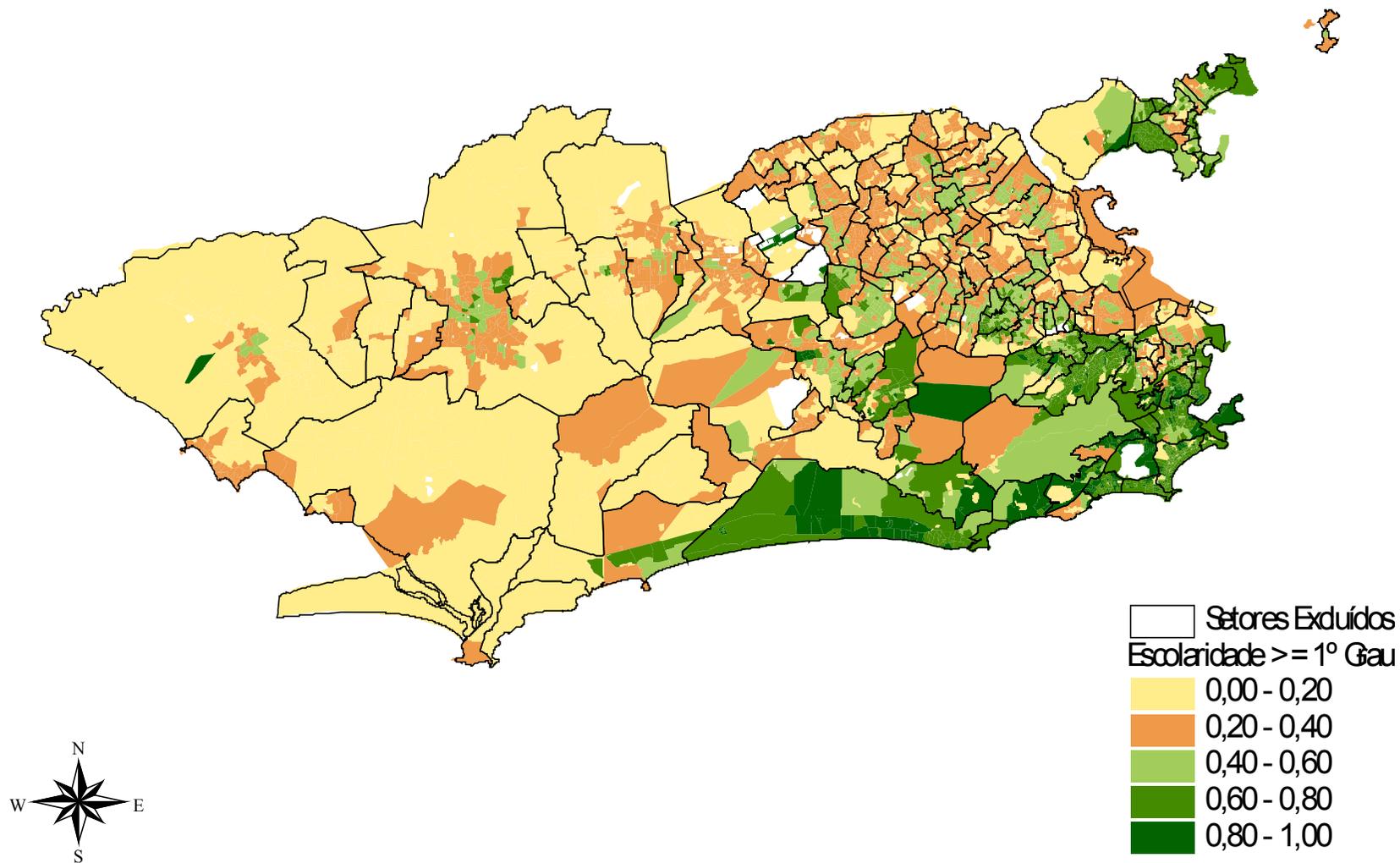
MAPA AI.4 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS LIGADOS À REDE PÚBLICA DE ESGOTO SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991



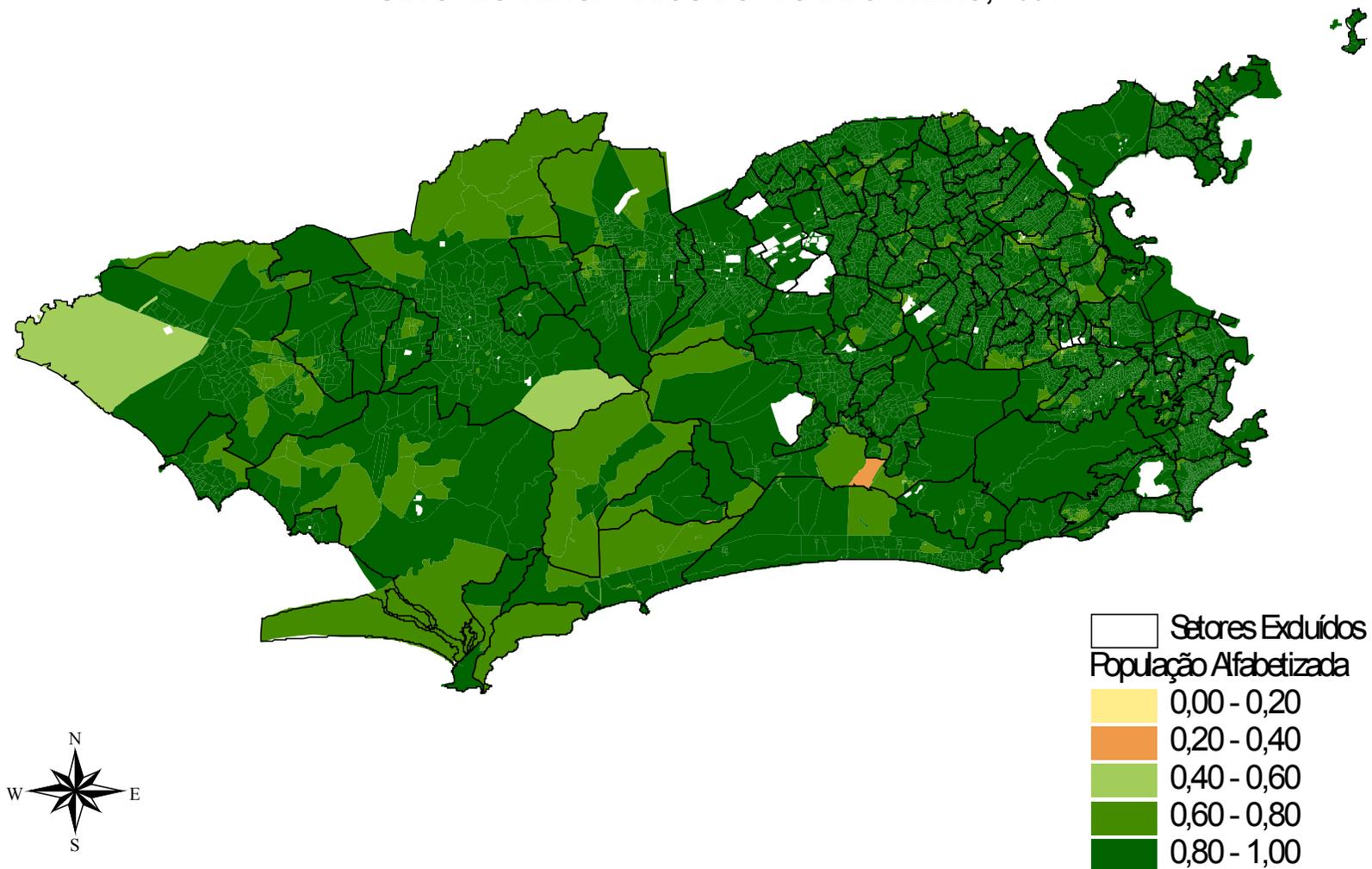
**MAPA A1.5 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE DOMICÍLIOS EM QUE O LIXO É COLETADO
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**



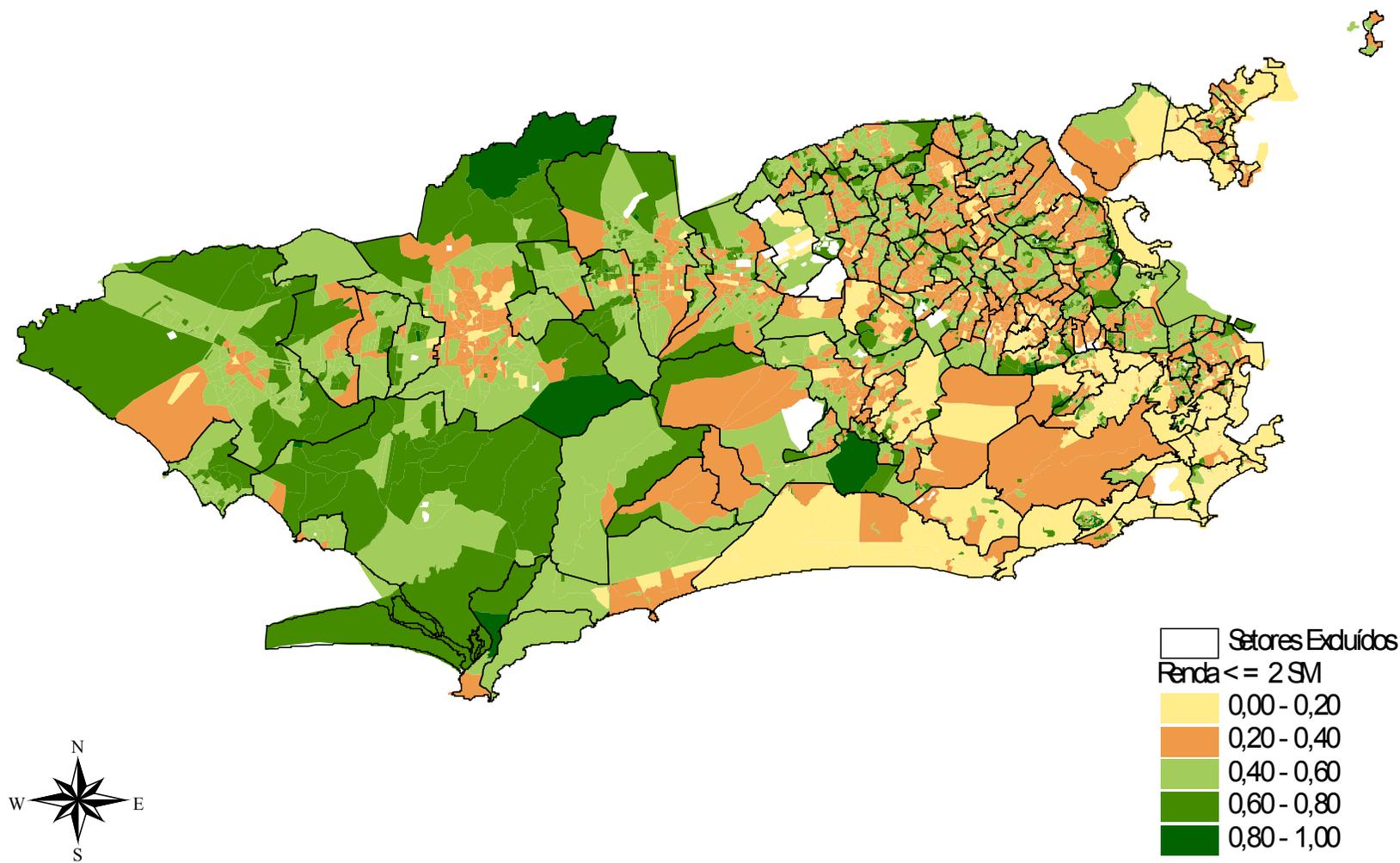
**MAPA AI.6 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM 1º GRAU COMPLETO OU MAIS
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**



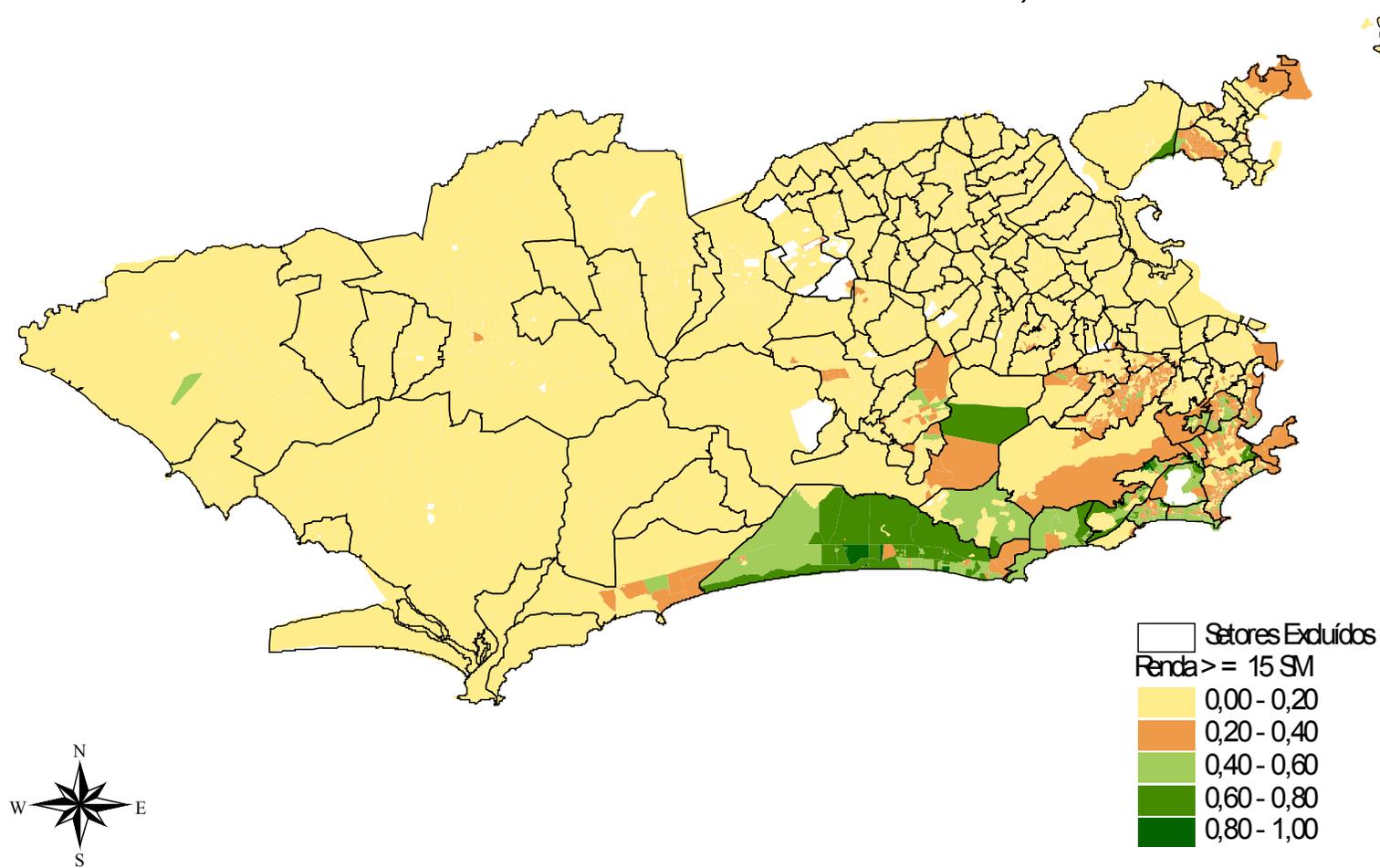
**MAPA AI.7 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE POPULAÇÃO ALFABETIZADA
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**



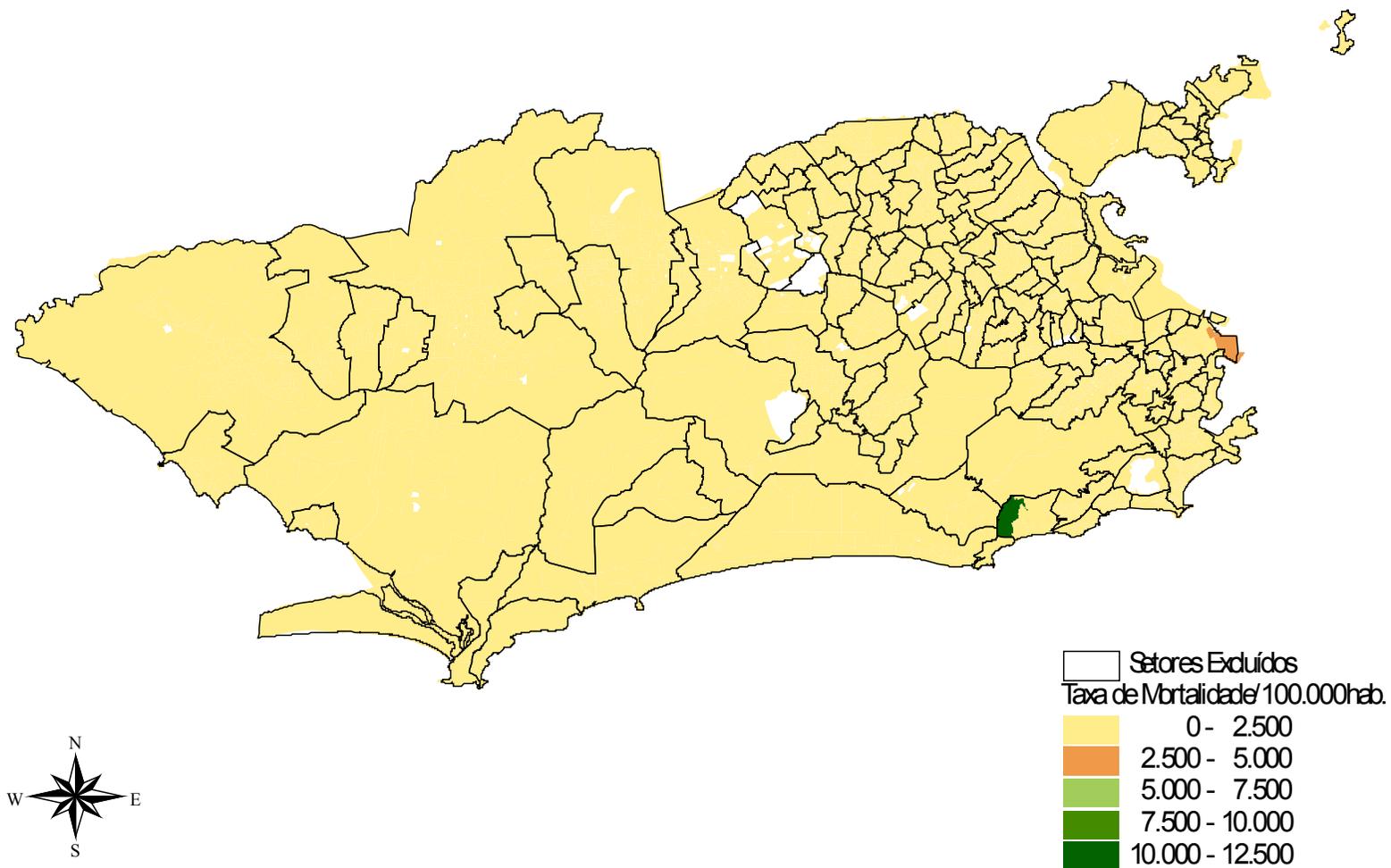
MAPA AI.8 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA INFERIOR A DOIS SALÁRIOS MÍNIMOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991



**MAPA AI.9 - DISTRIBUIÇÃO ESPACIAL DA PROPORÇÃO DE CHEFES DE DOMICÍLIO COM RENDA SUPERIOR A QUINZE SALÁRIOS MÍNIMOS
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**



MAPA AI.10 - DISTRIBUIÇÃO ESPACIAL DA TAXA DE MORTALIDADE POR DOENÇA ISQUÊMICA DO CORAÇÃO NA POPULAÇÃO DE 30 A 70 ANOS SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991



**MAPA AI.11 - DISTRIBUIÇÃO ESPACIAL DA POPULAÇÃO DE 30 A 70 ANOS
SETORES CENSITÁRIOS DO RIO DE JANEIRO, 1991**

