

UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS  
DEPARTAMENTO DE TRANSPORTES

**EXPLORANDO RECURSOS DE ESTATÍSTICA  
ESPACIAL PARA ANÁLISE DA ACESSIBILIDADE  
DA CIDADE DE BAURU**

**Bacharel em Estatística Ana Paula Krempi**

Orientador: Prof. Associado Antônio Néelson Rodrigues da Silva

Dissertação apresentada ao Departamento de Transportes, da Escola de Engenharia de São Carlos (Universidade de São Paulo-USP), como parte dos requisitos para a obtenção do título de Mestre em Engenharia Civil – Área: Transportes.

São Carlos

**2004**

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

K92e Krempi, Ana Paula  
Explorando recursos de estatística espacial para  
análise da acessibilidade na cidade de Bauru / Ana Paula  
Krempi. -- São Carlos, 2004.

Dissertação (Mestrado) -- Escola de Engenharia de São  
Carlos-Universidade de São Paulo, 2004.

Área: Transportes.

Orientador: Prof. Dr. Antônio Néelson Rodrigues da  
Silva.

1. Estatística espacial. 2. Análise espacial.
3. Acessibilidade. 4. Autocorrelação espacial.
5. SIG - Sistemas de Informações Geográficas. I. Título.

**Somente os fortes alcançam a vitória,  
porque os fracos logo se deixam vencer pelo desânimo...  
Somente os fortes conquistam os altos cumes, porque sabem  
escalar a montanha passo a passo e lentamente vencer os percalços...**

## **DEDICATÓRIA**

À toda minha família.

Aos meus pais Carlos e Irma e meu esposo Paulo,  
Vocês foram a força que me impulsionou, fazendo  
acreditar que a realização do sonho era possível.  
A vocês pertence grande parte de mais esta vitória.  
Minha profunda gratidão.

## **AGRADECIMENTOS**

---

A Deus, por ter me dado forças nas horas mais difíceis e de dificuldades;

Aos meus irmãos Carlos e Renata, à Rê, e ao Carlinhos, que foi fundamental para a conclusão desta dissertação, pois toda vez que o meu computador quebrava ele me socorria;

Ao Prof. Associado Antônio Néelson Rodrigues da Silva, por ter me orientado, e passado seus conhecimentos com paciência e dedicação;

Aos meus amigos que sempre me apoiaram, em especial a Rosi, que sempre me ajudou nas horas de apuro;

Aos meus familiares;

Aos meus amigos de trabalho, que me apoiaram ao máximo para que eu alcançasse mais esta vitória.

A todos os professores do Departamento;

A todo pessoal da Administração do Departamento de Transportes, em especial à Heloisa e à Beth, obrigada pelo incentivo;

A todos aqueles que colaboraram para a conclusão desta dissertação, mesmo que indiretamente.

## SUMÁRIO

---

LISTA DE FIGURAS.....	iii
RESUMO.....	vi
<i>ABSTRACT</i> .....	vii
1. INTRODUÇÃO.....	1
1.1. Caracterização do Problema.....	1
1.2. Justificativa do Trabalho.....	2
1.3. Objetivo do Trabalho.....	3
1.4. Estrutura do Trabalho.....	3
2. ANÁLISE ESPACIAL.....	4
2.1. Introdução.....	4
2.2. Análise e Estatística Espacial.....	8
2.3. Dados Espaciais.....	9
2.3.1. Dados de Processos Pontuais ou Dados Contínuos no Espaço.....	11
2.3.2. Dados de Área.....	12
2.4. Análise Exploratória de Dados Espaciais (AEDE).....	12
2.5. Acessibilidade.....	14
3. ANÁLISE ESPACIAL DE ÁREAS.....	16
3.1. Matriz de Proximidade Espacial.....	16
3.2. Média Móvel Espacial.....	18
3.3. Indicadores Globais de Autocorrelação Espacial.....	19
3.4. Indicadores Locais de Autocorrelação Espacial.....	21
3.5. Diagrama de Espalhamento de Moran.....	23
3.6. <i>Box Map</i> , <i>LISA Map</i> e <i>Moran Map</i> .....	24
3.7. Aplicação da técnica.....	27
3.8. Considerações Finais.....	30

4. METODOLOGIA .....	31
4.1. Abordagem Adotada .....	31
4.2. Obtenção dos Dados Básicos .....	32
4.3. Digitalização .....	34
4.4. Importação .....	37
4.5. Análise Espacial de Áreas.....	38
4.6. Atividades Desenvolvidas.....	43
5. ANÁLISE DOS RESULTADOS .....	45
5.1. Estudo de Caso.....	45
5.2. Base de Dados.....	46
5.3. Aplicação das Técnicas e Análise dos Resultados.....	48
6. CONCLUSÕES E RECOMENDAÇÕES .....	76
7. REFERÊNCIAS BIBLIOGRÁFICAS .....	80

## LISTA DE FIGURAS

---

---

Figura 2.1: Relação entre mobilidade, acessibilidade e comportamento de viagens, adaptado de Jones (1981) por Raia Jr. (2000) .....	15
Figura 3.1: Áreas e matriz de proximidade espacial.....	17
Figura 3.2: Gráfico de espalhamento de Moran .....	24
Figura 3.3: <i>Box Map</i> do Mapa de Exclusão/Inclusão Social de São Paulo, 2000 .....	25
Figura 3.4: <i>LISA Map</i> do Mapa de Exclusão/Inclusão Social de São Paulo, 2000.....	26
Figura 3.5: <i>Moran Map</i> do Mapa de Exclusão/Inclusão Social de São Paulo, 2000.....	26
Figura 3.6: Gráfico de barras do Mapa de Exclusão/Inclusão Social de São Paulo, 2000 .....	27
Figura 3.7: Gráfico de Moran para a variável densidade de edifícios por freguesia em Portugal nos anos de 1991 e 2001.....	28
Figura 3.8: Mapas temático representando a distribuição espacial dos pontos do gráfico de Moran, para a variável densidade de edifícios em Portugal nos anos de 1991 e 2001.....	30
Figura 4.1: Setores censitários da cidade de Bauru, em 1997 .....	35
Figura 4.2: Distribuição dos domicílios entrevistados através da pesquisa O-D de 1997 .....	35
Figura 4.3: Passos para criação de uma base de dados.....	37
Figura 4.4: Setores selecionados da cidade de Bauru.....	39
Figura 5.1: Localização da cidade de Bauru no estado de São Paulo e sua configuração geral.....	46
Figura 5.2: Divisão da cidade de Bauru em zonas de tráfego para a pesquisa O/D de 1997 (Raia Jr., 2000).....	47
Figura 5.3: Localizações dos diferentes usos do solo na cidade de Bauru .....	50
Figura 5.4: Distribuição dos percentuais de viagens realizadas por automóvel como motorista .....	51



Figura 5.5: Distribuição dos percentuais de viagens realizadas por automóvel como passageiro.....	51
Figura 5.6: Distribuição dos percentuais de viagens realizadas por ônibus .....	52
Figura 5.7: Distribuição dos percentuais de viagens realizadas por modos não motorizados (a pé e bicicleta) .....	52
Figura 5.8: Distribuição da variável renda.....	53
Figura 5.9: Diagrama de Espalhamento de Moran para a variável % <i>auto_moto</i> .....	54
Figura 5.10: Diagrama de Espalhamento de Moran para a variável % <i>auto_pass</i> .....	54
Figura 5.11: Diagrama de Espalhamento de Moran para a variável % <i>de modos não motorizados</i> .....	55
Figura 5.12: Diagrama de Espalhamento de Moran para a variável % <i>ônibus</i> .....	55
Figura 5.13: Diagrama de Espalhamento de Moran para a variável <i>tend a</i> .....	56
Figura 5.14: <i>Box Map</i> da variável % <i>auto_moto</i> e mapa de barras (desvio x média dos desvios).....	57
Figura 5.15: <i>Box Map</i> da variável % <i>auto_pass</i> e mapa de barras (desvio x média dos desvios).....	58
Figura 5.16: <i>Box Map</i> da variável % <i>modos não motorizados</i> .....	59
Figura 5.17: <i>Box Map</i> da variável % <i>ônibus</i> .....	59
Figura 5.18: <i>Box Map</i> da variável <i>renda</i> .....	60
Figura 5.19: <i>LISA Map</i> da variável % <i>auto_moto</i> .....	60
Figura 5.20: <i>LISA Map</i> da variável % <i>auto_pass</i> .....	61
Figura 5.21: <i>LISA Map</i> da variável % <i>modos não motorizados</i> .....	62
Figura 5.22: <i>LISA Map</i> da variável % <i>ônibus</i> .....	62
Figura 5.23: <i>LISA Map</i> da variável <i>renda</i> .....	63
Figura 5.24: <i>Moran Map</i> da variável % <i>auto_moto</i> .....	64
Figura 5.25: <i>Moran Map</i> da variável % <i>auto_pass</i> .....	64
Figura 5.26: <i>Moran Map</i> da variável % <i>modos não motorizados</i> .....	65
Figura 5.27: <i>Moran Map</i> da variável % <i>ônibus</i> .....	65
Figura 5.28: <i>Moran Map</i> da variável <i>renda</i> .....	66
Figura 5.29: Divisão da cidade de Bauru.....	66
Figura 5.30: Representação dos nove conjuntos de zonas resultantes da combinação de anéis de localização (periferia, transição e centro) e níveis de renda (alta, média e baixa).....	67
Figura 5.31: Legenda para os gráficos de setores das Figuras 5.32 a 5.36.....	71

Figura 5.32: Gráficos de setores da variável <i>% modos não motorizados</i> .....	71
Figura 5.33: Gráficos de setores da variável <i>% ônibus</i> .....	71
Figura 5.34: Gráficos de setores da variável <i>% auto_moto</i> .....	72
Figura 5.35: Gráficos de setores da variável <i>% auto_pass</i> .....	72

## RESUMO

---

A acessibilidade está relacionada com a maneira como a disponibilidade de transportes e os usos do solo afetam os indivíduos na realização de viagens para o desenvolvimento de suas atividades habituais. Frequentemente se assume que os moradores de baixa renda da periferia são os mais afetados pela falta de acesso aos meios de transporte. A questão subjacente a esta afirmação, no entanto, permanece sem uma resposta definitiva: o nível de renda, por si só, seria um indicativo do nível de acessibilidade? O objetivo deste estudo é explorar a união de ferramentas de estatística espacial e SIG (Sistema de Informações Geográficas) com um propósito específico, que é o de analisar as relações entre aspectos da distribuição espacial de características da população (como a renda, por exemplo) de uma cidade média brasileira e os diversos níveis de acessibilidade por diferentes modos de transporte nela observados, buscando possíveis respostas para esta pergunta.

Quando se utiliza procedimentos de visualização e classificação de dados espaciais comuns em SIG, nem sempre as informações são diretamente perceptíveis. Logo, deve-se utilizar ferramentas que ampliem as possibilidades de compreensão e análise dos dados. Inicialmente, as ferramentas selecionadas para uso neste trabalho são apresentadas e discutidas quanto à sua aplicação e utilização na análise proposta. Para tal foram utilizados dados coletados em uma pesquisa origem-destino (O-D) realizada na cidade de Bauru - SP, agrupados por setores censitários e adicionados ao SIG, aplicando técnicas de estatística espacial utilizadas para entidades do tipo área. Os resultados obtidos são apresentados na forma de mapas e de índices que medem a associação espacial global e local entre estas zonas.

Uma das conclusões interessantes da aplicação foi a identificação de regiões da cidade com dinâmica particular, que contrariam o padrão global observado nas demais partes da área urbana. Pôde-se constatar ainda particularidades a respeito do uso de cada modo de transportes. O modo automóvel como motorista, por exemplo, possui agrupamento espacial bem definido no nível de renda alta tanto nas regiões de periferia, como nas de transição e central. Já o modo ônibus é predominantemente utilizado nas zonas de renda baixa das regiões de periferia e transição, enquanto que os modos não motorizados possuem uma dinâmica bem diversificada em toda a área urbana. Estes e outros resultados do estudo de caso deixam claro que as análises de estatística espacial em ambiente SIG criam uma ferramenta para ampliar a análise convencional de acessibilidade em transportes.

## ABSTRACT

---

Transportation accessibility is directly related to the level of transportation supply and land uses and the way they affect individuals in their trip desires for accomplishing regular-basis activities. It is often assumed that low-income segments of the population living at the periphery of the cities are those affected the most by poor conditions of transportation accessibility. There is a subjacent question behind this statement, however, which is: can the income level or the location of an individual alone explain his/her accessibility level? In order to look for answers to this question, the aim of this study is to analyze, making use of spatial statistics tools in a GIS (Geographic Information System) environment, the relationships between accessibility and income and their geographical distributions in a medium-sized Brazilian city.

The application of the most commonly used GIS resources, such as visualization and spatial data classification tools, not always assures a full comprehension of the phenomenon under analysis. As a consequence, many problems require tools that enhance the possibilities of observation and analysis. As tools with this characteristic have been used in this work, they were initially introduced. Thereafter, the possibilities of use of these tools in the problem analyzed were also discussed. Data of an origin-destination (O-D) survey carried out in the city of Bauru, located in the state of São Paulo, which brings information about four different transportation modes, were used in this study. Such data, grouped following the census tracts, were carefully examined in a Geographic Information System in order to look for spatial patterns of accessibility that are not visible in the traditional approaches. The results of the analysis are presented in maps and as indices that are able to capture global and local spatial association patterns in areas.

One of the interesting outcomes of the application was the identification of regions with particular dynamics, which go against the pattern found in the overall urban area. Particularities regarding each particular transportation mode have also been noticed. The zones where the automobile is most used (by drivers, not by passengers) are spatially clustered, regardless if the zone is at the periphery, transition zone or central area of the city. The bus trips are predominantly carried out in low-income areas of the periphery and transition rings, while the non-motorized modes (walk and bicycle) have shown a very diversified dynamics in the entire urban area. This and other results of the case study clearly indicate that spatial statistics analyses in a GIS environment create a powerful tool to extend conventional transportation accessibility analysis.

## 1. INTRODUÇÃO

---

Este capítulo descreve, de modo resumido, as características do problema estudado, que se refere ao estudo da acessibilidade na cidade de Bauru, bem como a justificativa, o objetivo e a estrutura deste trabalho.

### 1.1. Caracterização do Problema

Nas últimas décadas as cidades brasileiras apresentaram um crescimento intenso, o que vem causando significativos impactos nos sistemas de transportes, acarretando problemas como, por exemplo, uma redução generalizada da acessibilidade. Os moradores da periferia, em particular, foram os mais afetados, pois as cidades expandiram-se em grandes áreas, dificultando o deslocamento destes habitantes por meios não motorizados. Com grandes distâncias a percorrer e com serviços de transporte precários, a população da periferia consome a maior parte de seu tempo com viagens, na sua maioria, por motivo de trabalho.

Embora inúmeros trabalhos, tais como Silva *et al.* (1998) e Goto *et al.* (2001), tenham buscado esclarecer como se dá à distribuição da acessibilidade para diferentes regiões das zonas urbanas e para segmentos diversos da população, muitas dúvidas ainda permanecem. Seriam mesmo os habitantes da periferia os mais afetados pela baixa acessibilidade? Não estaria a baixa acessibilidade mais relacionada à renda do que à localização geográfica, uma vez que em muitas cidades a periferia abriga bairros luxuosos, que dispõem de toda a infra-estrutura de transporte e muitas vezes apresentam acessibilidade maior que outros mais bem localizados geograficamente (mais “centrais”)? Como uma alternativa para responder a perguntas como estas, uma possibilidade é fazer um estudo do comportamento da acessibilidade a partir de suas características de distribuição espacial. As técnicas de estatística espacial foram desenvolvidas para tentar identificar regiões onde a distribuição dos valores possa

apresentar um padrão específico associado a sua localização geográfica. Para tal, ferramentas de estatística espacial podem ser utilizadas, tendo um Sistema de Informações Geográficas como suporte.

Dentre os recursos disponíveis nos SIG para utilização direta em planejamento de transportes se incluem aqueles destinados a análises espaciais sob diversos aspectos, tais como: matriz de distância entre pontos, matriz de origem e destino de viagens, caminho mínimo entre pontos, bem como alguns modelos específicos de transporte (Silva *et al.*, 2003). Além disto, ferramentas estatísticas como a análise de padrões de distribuição de pontos e a análise de atributos por área parecem ter sido feitas sob medida para a estrutura dos SIG, o que torna a união das duas técnicas uma poderosa ferramenta de análise. Isto pode ser comprovado através de trabalhos que utilizaram estas ferramentas, como os de Bullen (1997), onde modelos de regressão foram usados para estimar valores de propriedades no Reino Unido; Câmara *et al.* (2001), que fez um levantamento da longevidade da população de acordo com a localização geográfica dos domicílios; e Ramos e Silva (2003), que usaram técnicas de Análise Exploratória de Dados Espaciais como parte de um método para definição de Regiões Metropolitanas em Portugal.

## 1.2. Justificativa do Trabalho

A estatística espacial é definida como uma coleção de técnicas para análise geográfica onde o resultado da análise depende do arranjo espacial dos eventos. Evento geográfico é entendido como uma coleção de pontos, linhas ou objetos, localizados no espaço e ligados com uma classe de atributos.

Baseado na coleção sistemática de informação quantitativa, os objetivos de estatística espacial são: a) Descrição cuidadosa e precisa de eventos no espaço geográfico (incluindo a descrição de padrões), b) Exploração sistemática do padrão dos eventos e de sua associação no espaço com o objetivo de ganhar um melhor entendimento dos processos que podem ser responsáveis pela distribuição observada, c) Melhora da habilidade de prever e controlar eventos que possam ocorrer no espaço geográfico.

Considerando a escassez de estudos sobre o tema no Brasil, a proposta principal deste trabalho é realizar uma análise utilizando ferramentas de estatística espacial para o caso específico de Bauru. Assim, em um contexto onde a acessibilidade

é extremamente relevante para a comunidade, a principal justificativa para este trabalho está em contribuir para uma melhor compreensão da sua distribuição através de quatro modos de transporte, tendo como referência dados de uma pesquisa O-D realizada na cidade em 1997.

### **1.3. Objetivo do Trabalho**

O presente trabalho procura explorar a união de ferramentas de estatística espacial e SIG com um propósito específico, que é o de analisar as relações entre aspectos da distribuição espacial de características da população de uma cidade média brasileira e os diversos níveis de acessibilidade por diferentes modos de transporte nela observados.

### **1.4. Estrutura do Trabalho**

Para atingir o objetivo proposto, este trabalho está estruturado em sete capítulos, como descrito na seqüência. O segundo capítulo contém uma revisão bibliográfica em que são apresentados os principais conceitos de estatística espacial e uma breve discussão acerca do tópico acessibilidade, enquanto o terceiro capítulo apresenta técnicas de estatística espacial aplicadas especificamente a dados de áreas, que é o caso deste estudo. No quarto capítulo é apresentada a metodologia para o desenvolvimento do trabalho, com detalhes a respeito da obtenção dos dados e um exemplo da aplicação das técnicas propostas.

No quinto capítulo são expostos e discutidos os resultados encontrados no estudo de caso desenvolvido. O capítulo 6 contém as principais conclusões e recomendações para trabalhos futuros, seguido pela referências bibliográficas, no capítulo 7.

## 2. ANÁLISE ESPACIAL

---

Este capítulo, baseado em uma revisão da literatura, apresenta os principais conceitos de Estatística Espacial, que é um desdobramento de uma área mais ampla denominada Análise Espacial, destacando os princípios básicos, os principais efeitos e a classificação dos dados espaciais, com maior ênfase para os dados de áreas, aqui analisados. Apresenta também um destaque para a Análise Exploratória de Dados Espaciais, técnica utilizada para analisar dados de áreas, a fim de verificar o comportamento da distribuição espacial da acessibilidade, objetivo deste trabalho.

### 2.1. Introdução

De acordo com Serrano e Valcarce (2000), nos últimos quarenta anos a economia urbana e regional tem apresentado um grande desenvolvimento metodológico, o que ampliou, entre outras coisas, as possibilidades de utilização de dados de natureza espacial para fins de planejamento. Vários estudos foram realizados até que avanços no campo da estatística começassem a ser transferidos para o que hoje se conhece como estatística espacial, desde o reconhecimento dos problemas causados pela dependência espacial, que já se observava nos trabalhos publicados por Student (1914) sobre métodos de estatística tradicionais, até chegar aos trabalhos escritos por Moran (1948) e Geary (1954), que continham os primeiros índices formais para detectar a presença da autocorrelação espacial em um conjunto de dados.

Este desenvolvimento vem acontecendo de forma mais intensa a partir dos anos oitenta e noventa do século XX, trazendo avanços notáveis para o campo do conhecimento hoje conhecido como econometria espacial. O avanço é devido, entre outros aspectos, ao aumento do interesse na área, à crescente disponibilidade de bases de dados socioeconômicas georeferenciadas e à expansão de tecnologias eficientes e de



baixo custo, onde hoje já se poderia incluir tanto os SIG como os programas específicos para análise de dados espaciais, que permitem tratar dados referenciados espacialmente.

Por outro lado, conforme afirmam Câmara *et al.* (2001), grande parte dos usuários de SIG limita-se ainda hoje ao uso do mesmo apenas para visualização de mapas, ignorando que é possível ir muito além, já que a combinação de técnicas estatísticas adaptadas a dados espaciais com funções de visualização formam um conjunto de ferramentas de análise com grande potencial para estudos urbanos e regionais (como demonstra Paez, 2003, ao apresentar exemplos de utilização de técnicas de estatística espacial aplicadas na análise urbana).

De acordo com Câmara *et al.* (2001), os processos de análise espacial incluem métodos de visualização, métodos que investiguem o padrão dos dados, isto é, se os dados apresentam uma agregação definida ou se a distribuição é aleatória, métodos que auxiliem na escolha de um modelo estatístico e a estimação dos parâmetros deste modelo.

Desta forma as ferramentas de estatística espacial podem ser destinadas a: seleção, manipulação, análise exploratória e análise confirmatória. A seleção é feita através da navegação em bancos de dados, identificando os dados de interesse, realizando consultas e apresentando mapas simples. O processo de manipulação envolve todas as funções que criam dados espaciais. As técnicas de análise exploratória descrevem e visualizam dados espaciais, verificando a existência de padrões de associação espacial. As técnicas de análise confirmatória envolvem o conjunto de modelos de estimação e procedimentos para sua validação.

Segundo Assunção (2001), a visualização é um fator importante da estatística espacial, mas não tão útil se não for possível descrever, comparar e interpretar os dados. A estatística descritiva ou inferencial é, portanto, eficaz e essencial para tornar explícito o que está apenas implícito em mapas, ou para analisá-los chamando a atenção para características que não eram perceptíveis apenas pela visualização.

Isto explica porque o interesse por ferramentas de estatística espacial associadas aos SIG tem crescido gradativamente. Alguns resultados obtidos pelos pesquisadores na área até o presente momento sugerem que os grupos de métodos para os quais a ligação com o SIG vêm produzindo melhores resultados seriam aqueles das estatísticas descritivas simples e os relacionados com a estrutura de análise de covariância. Outras áreas da estatística, tais como métodos de alisamento, funções-K e

regressão espacial podem produzir benefícios equivalentes. Porém, até o presente momento, um progresso moderado tem sido em direção à visualização geográfica. Algum progresso também tem sido observado com relação a métodos estatísticos descritivos, o que conduz à impressão de que muita coisa ainda pode ser feita com relação ao “casamento” *estatística espacial-SIG*. Neste sentido, esforços podem ser destinados não ao desenvolvimento de ferramentas internas ao SIG, mas sim à melhora da interface entre programas estatísticos e SIG, como no caso relatado por Anselin e Bao (1997).

De acordo com Paelinck e Klaassen (1979<sup>1</sup>, *apud* Serrano e Valcarce, 2000) destacam-se cinco princípios básicos no campo da estatística espacial:

- **Interdependência:** Todo modelo espacial tem que se caracterizar por sua interdependência, ou seja, devem ser incorporadas as relações mútuas entre as observações e as variáveis. Um exemplo pode ser verificado em áreas onde residem pessoas com elevado nível de renda, uma vez que seus gastos não se restringem à área em que vivem, mas ocorrem também nas regiões vizinhas, igualmente estimulando o crescimento econômico destas últimas.
- **Assimetria:** As relações espaciais são em princípio assimétricas. Um exemplo ocorre no comércio, onde a probabilidade de um residente da periferia ir fazer compras no centro é maior do que a do residente do centro (ou próximo dele) ir fazer compras na periferia.
- **Alotropia:** As causas de um fenômeno espacial podem não se manifestar diretamente no lugar onde ele ocorre. Assim, por exemplo, os fenômenos migratórios se explicam não só pela comparação de vantagens e inconvenientes nos espaços de origem e destino, mas também devido a causas que ocorrem em outros lugares do espaço (nível dos salários, por exemplo).
- **Não linearidade:** A não linearidade de soluções espaciais ótimas obtidas a priori conduz a modelos que requerem

---

<sup>1</sup> PAELINCK, J.H.P; KLAASSEN, L.H. (1979) *Spatial Econometrics*, Farnborough, Saxon House.

especial atenção para sua especificação, uma vez que estes poderão ter comportamento não linear.

- **Inclusão de variáveis topológicas:** Um modelo espacial deve incorporar variáveis topológicas: coordenadas, distâncias, superfícies, densidades, etc.

Apesar de sua importância, nem sempre será possível observar estes cinco princípios na construção dos modelos espaciais. Poderão, no entanto, surgir outros, além daqueles especificados.

Segundo Serrano e Valcarce (2000), quando se trabalha particularmente com dados de natureza espacial podem aparecer os denominados efeitos espaciais: a **heterogeneidade** e a **autocorrelação espacial**. O primeiro efeito aparece quando os dados utilizados para explicar um mesmo fenômeno são de unidades espaciais muito distintas. Neste caso os problemas mais frequentes são a heterocedasticidade e a instabilidade estrutural. A heterocedasticidade espacial ocorre pela omissão de variáveis ou outras formas de erros de especificação que levam à aparição dos denominados erros de medidas. A dependência ou autocorrelação espacial surge sempre que o valor de uma variável em um lugar do espaço está relacionado com seu valor em outro ou outros lugares do espaço. Este fenômeno pode ser entendido como uma situação em que observações próximas no espaço possuem valores similares (correlação de atributos). Neste caso, o desafio da análise é medir o grau de associação espacial entre observações de uma ou mais variáveis. A autocorrelação espacial pode ser positiva ou negativa, sendo positiva quando a presença de um fenômeno em certa região se espalha para as regiões que a rodeiam, influenciando-a. Dois motivos podem induzir a ocorrência de autocorrelação espacial: a existência de erros de medidas e de fenômenos de interação espacial.

O desenvolvimento computacional recente se traduziu também em avanços na área de estatística espacial, proporcionando o crescimento de vários campos de atuação e abrindo inúmeras possibilidades de aplicações. Apesar de existirem no Brasil pesquisas empregando técnicas de estatística espacial, o número de trabalhos dedicados ao tema, particularmente quando se trata de aplicações em transportes, ainda é muito pequeno (por exemplo, Krempf *et al.*, 2002; Queiroz, 2003; Teixeira, 2003). Visando contribuir com as discussões acerca do tema, a proposta deste trabalho é realizar novos estudos de estatística espacial na área de transportes. No caso, serão

estudadas viagens por quatro modos de transporte, agrupadas segundo setores censitários da cidade de Bauru, conforme será detalhado na metodologia.

## 2.2. Análise e Estatística Espacial

Análise espacial é um estudo quantitativo de fenômenos que são localizados no espaço. Assim, um dos índices dos dados utilizados em estatística espacial refere-se sempre a uma localização geográfica, ou seja, o fenômeno estudado possui alguma forma de localização. Muitos dados de uso comum possuem alguma referência espacial como, por exemplo, dados censitários, sempre relacionados ao local de residência do indivíduo. Sob esta ótica, muitos dados que podem ser analisados estatisticamente possuem referência espacial. Mas a característica fundamental da estatística espacial que a diferencia da estatística clássica é o uso explícito da referência geográfica no modelo, isto é, o uso explícito das coordenadas espaciais no processo de coleta, descrição e análise dos dados. Assim sendo, o interesse está centrado nos processos que ocorrem no espaço e os métodos empregados buscam descrever e analisar o comportamento destes processos. Esta característica faz com que estudos sobre o assunto exibam comportamento complexo para serem analisados por métodos tradicionais de estatística (Assunção, 2001).

Segundo Landim (1998), quando se utiliza a estatística clássica para representar as propriedades dos valores amostrais, presume-se que estes sejam realizações de uma variável casual, as posições relativas da amostra são ignoradas e presume-se que todos os valores amostrais tenham a mesma probabilidade de serem escolhidos. Assim, muitos modelos de estatística exigem a independência das variáveis.

A estatística espacial traz resultados diferentes daqueles obtidos pela estatística clássica, sendo os primeiros geralmente mais robustos por incorporarem a dimensão espacial. Para sua análise são necessárias pelo menos as informações sobre a localização e os atributos, que são valores associados aos dados independentemente da forma como sejam medidos, e parte-se do pressuposto que os dados são espacialmente dependentes.

Um exemplo para explicar os conceitos acima se refere ao caso estudado em Londres, em 1854, e citado por Câmara *et al.* (2001). O exemplo refere-se a uma grave epidemia de cólera, doença para a qual na época não era conhecida a forma de contaminação. Como uma das estratégias para lidar com o problema, foi identificada no

mapa da cidade a localização exata dos doentes de cólera e dos poços de água. Com a espacialização dos dados, percebeu-se que a maioria dos casos estava concentrada em torno dos poços. Como consequência, estes foram lacrados, desta forma contribuindo para acabar com a epidemia.

Para Assunção (2001), considerando que a estatística espacial é um ramo da estatística que estuda métodos científicos para a coleta, descrição, visualização e análise de dados que possuem coordenadas geográficas, os dados espaciais podem ser classificados em quatro categorias:

- Dados de Processos Pontuais;
- Dados de Superfícies Aleatórias;
- Dados de Área;
- Dados de Interação Espacial.

Para cada um dos itens acima existem diferentes métodos estatísticos para descrever e analisar os dados. Segundo Levine (1996), as estatísticas usadas para descrever tanto pontos quanto áreas podem ser subdivididas em três categorias gerais:

- **Medidas de distribuição espacial:** descrevem o centro, a dispersão, direção e forma da distribuição de uma variável;
- **Medidas de autocorrelação espacial:** descrevem a relação entre as diferentes localizações para uma variável simples, indicando o grau de concentração ou dispersão (por exemplo, análise de agrupamentos);
- **Medidas de associação espacial entre duas ou mais variáveis:** descrevem a correlação ou associação entre variáveis distribuídas no espaço, por exemplo, a correlação entre a localização de lojas de bebidas com pontos onde ocorrem muitos acidentes de trânsito.

### 2.3. Dados Espaciais

Para Câmara *et al.* (2001), uma das formas de divisão dos dados com os quais a estatística espacial lida seria classificá-los como ambientais ou

socioeconômicos. Dados ambientais seriam todos os tipos de dados obtidos por coleta e que descrevem características físicas do meio (mapas geológicos, topográficos, ecológicos, fitogeográficos e pedológicos). Dados deste tipo podem ser analisados como se constituíssem uma amostra de uma superfície contínua.

Os dados socioeconômicos tratam de aspectos associados a levantamentos populacionais (como os censos demográficos), imóveis, rotas de transporte, saúde etc. Estes dados apresentam a localização específica do indivíduo no espaço, mas por questão de confidencialidade são agrupados em áreas delimitadas por polígonos fechados (como, por exemplo, os setores censitários). Os trabalhos realizados com dados socioeconômicos são em geral mais complexos, pois dentro destes polígonos existe uma variedade de informações, que muitas vezes não são na realidade distribuídas no espaço de forma homogênea.

De acordo com Câmara *et al.* (2001), a escolha das unidades de coleta é parte crucial da análise de dados socioeconômicos em geoprocessamento. Estas devem apresentar resolução definida a partir da menor área para a qual estão disponíveis informações sobre o estudo. Assim quanto menor a escala, maior a população e a área da unidade de estudo, menor a resolução e, portanto, menor a homogeneidade interna e a capacidade de distinguir diferenças. Caso aumente a escala e a resolução, outros problemas surgirão, pois à medida que diminui a área e a população, diminui também a ocorrência do evento estudado. Este problema relacionado à dimensão da área de estudo e aos dados ali contidos é chamado de MAUP (“*Modifiable Area Unit Problem*”, ou Problema da Unidade de Área Modificável), que faz com que mudanças nas dimensões da unidade e na escala possam conduzir a conclusões diversas.

Ainda tratando-se de estatística espacial pode-se classificar os dados seguindo outra denominação: dados de processos pontuais e dados de áreas. Estes dados guardam, respectivamente, forte relação com os dados ambientais e socioeconômicos. Para Cressie (1993), estes dois tipos de dados espaciais seriam ainda denominados, respectivamente, de *Geostatistical data* (dados contínuos no espaço) ou *Lattice data* (dados agrupados em áreas). O primeiro grupo se refere a dados pontuais, como uma amostra de uma distribuição contínua. O segundo grupo consiste em uma coleção fixa de localizações espaciais discretas (pontos ou polígonos).

### 2.3.1. Dados de Processos Pontuais ou Dados Contínuos no Espaço

Segundo Câmara *et al.* (2001), um padrão pontual representa um conjunto de dados consistindo de uma série de localizações pontuais ( $p_1, p_2, \dots$ ) que indicam a ocorrência de eventos de interesse dentro da área de estudo. A palavra “evento” é utilizada para expressar qualquer tipo de fenômeno localizado no espaço. Estes fenômenos podem representar, por exemplo, crimes, ocorrências de doenças e localizações de espécies vegetais.

Neste tipo de análise as observações estão associadas a valores, mas apenas a ocorrência dos eventos é considerada. As áreas das ocorrências dos eventos não são uma medida válida. Os dados não são compostos exclusivamente pela localização dos eventos, podendo, no entanto, também ter outros atributos que podem ser incorporados na análise, a eles associados.

O objetivo da análise espacial de pontos é examinar se o conjunto de eventos apresenta algum tipo de padrão sistemático ou aleatório, buscando entender em qual escala este padrão ocorre. Para tal, alguns métodos, como o estimador de intensidade (*Kernel Estimation*), modelagem de distribuições de pontos e o método do vizinho mais próximo, permitem verificar o padrão dos dados analisados. O estimador de intensidade (*Kernel Estimation*) tem como objetivo analisar o comportamento de padrões de pontos, ajustando uma função bi-dimensional sobre os eventos considerados, e é muito útil para mostrar uma visão geral da distribuição de primeira ordem dos eventos. O modelo mais simples e mais utilizado na prática para analisar estatisticamente a distribuição de pontos é conhecido como CSR (do inglês “*Complete Spatial Randomness*”) ou Aleatoriedade Espacial Completa. Este modelo divide a região de estudo em subáreas e modela a distribuição de eventos pontuais como um processo aleatório. O método do vizinho mais próximo estima a função de distribuição cumulativa baseado nas distâncias entre eventos em uma região de análise, e compara graficamente os padrões de distribuição espacial dos eventos (pontos) observados com aqueles esperados a partir das funções de probabilidade conhecidas, verificando a relação entre cada evento e aqueles mais próximos a ele.

### 2.3.2. Dados de Área

Segundo Assunção (2001), dados de área envolvem elementos que associam o mapa geográfico a uma base de dados. Este mapa geográfico é dividido por áreas, onde cada uma delas possui uma ou mais variáveis aleatórias ( $y_i, \dots, y_n$ ) que representam um valor para toda área e não apenas o valor de um ponto específico (este ponto seria, por exemplo, o centróide da área). Neste caso não se conhece a localização exata do evento, mas sim um valor agregado por área. Este tipo de dado é em geral fornecido em forma agregada como, por exemplo, os dados do censo, que são agrupados em setores censitários a fim de manter o sigilo dos domicílios entrevistados.

O objetivo da análise é verificar a existência de um padrão espacial nos valores observados. A forma usual de apresentação dos dados agregados por área é através de mapas coloridos com o padrão espacial do fenômeno. Quando este padrão espacial é observado, se faz necessário verificar se é aleatório ou apresenta uma agregação definida, ou ainda se esta distribuição pode ser associada a causas mensuráveis.

Quando são utilizadas técnicas de estatística espacial para análise de dados de áreas, estas são apresentadas em forma de índices que medem a associação espacial (Índice de Moran), gráfico de espalhamento (*Box Map*) e mapas (Distribuição espacial). Estes índices auxiliam na identificação de agrupamentos de objetos cujos atributos possuem valores altos e baixos, áreas de transição e casos atípicos.

De particular interesse para os propósitos deste trabalho, que trata de dados associados a áreas adjacentes e independentes, são as técnicas de Análise Exploratória de Dados Espaciais (AEDE), extensão das técnicas de Análise Exploratória de Dados, e que serão tratadas com mais detalhes no próximo item.

### 2.4. Análise Exploratória de Dados Espaciais (AEDE)

Análise Exploratória de Dados Espaciais é feita por um conjunto de ferramentas gráficas e descritivas cujo objetivo é identificar propriedades espaciais dos dados para detectar padrões, formular hipóteses a partir da localização dos dados e avaliar aspectos de modelagem espacial. Esta análise é baseada tanto em métodos gráficos e visuais como em técnicas numéricas, normalmente estatísticas. Tem sido



utilizada nos últimos anos como metodologia para estudar padrões e associações de dados, especialmente tratando-se com grandes bases de dados.

Praticamente nenhuma das ferramentas tradicionais existentes em pacotes estatísticos está preparada para tratar com dados de natureza espacial, pelo fato de ignorarem os efeitos da localização (uma das variáveis mais importantes para o tratamento de dados espaciais), dependência e heterogeneidade espacial. Pode-se inclusive observar que, para observações iniciais de relações bivariadas e multivariadas, muitas técnicas de análise exploratória levam à conclusões incorretas quando se verifica a presença de autocorrelação espacial.

Segundo Anselin e Bao (1997), esta coleção de técnicas permite descrever e visualizar distribuições espaciais, identificar situações atípicas (*outliers*), descobrir padrões de associação espacial, agrupamentos de valores semelhantes (*clusters*) e sugerir regimes espaciais ou outras formas de heterogeneidade espacial. Entende-se por *outliers* espaciais, dados cuja localização podem exercer uma forte influência na hora de realizar estimações. Para que sua presença não atrapalhe na análise é necessário verificar quais dados são considerados *outliers* a fim de distribuir seus efeitos. Já o *cluster* ou agrupamento espacial pode ser definido como um agregado de ocorrências no espaço ou a ocorrência de valores semelhantes em áreas próximas.

Neves *et al.* (2001), em seu artigo sobre análise exploratória de dados socioeconômicos de São Paulo, apresentam um conjunto de técnicas de estatística espacial que permite uma melhor compreensão da dinâmica espacial do fenômeno estudado. Estas técnicas são baseadas no conceito de autocorrelação espacial, gerando como resultados índices globais e locais.

Anselin (1996) apresenta uma classificação das diferentes perspectivas e técnicas de Análise Exploratória de Dados Espaciais. A classificação segue duas dimensões: a primeira refere-se à distinção entre indicadores globais e locais de associação espacial e a segunda centra-se na distinção entre as estatísticas baseadas na vizinhança e distância. Cada uma requer um grupo de métodos específicos e distintos para a exploração de dados espaciais.

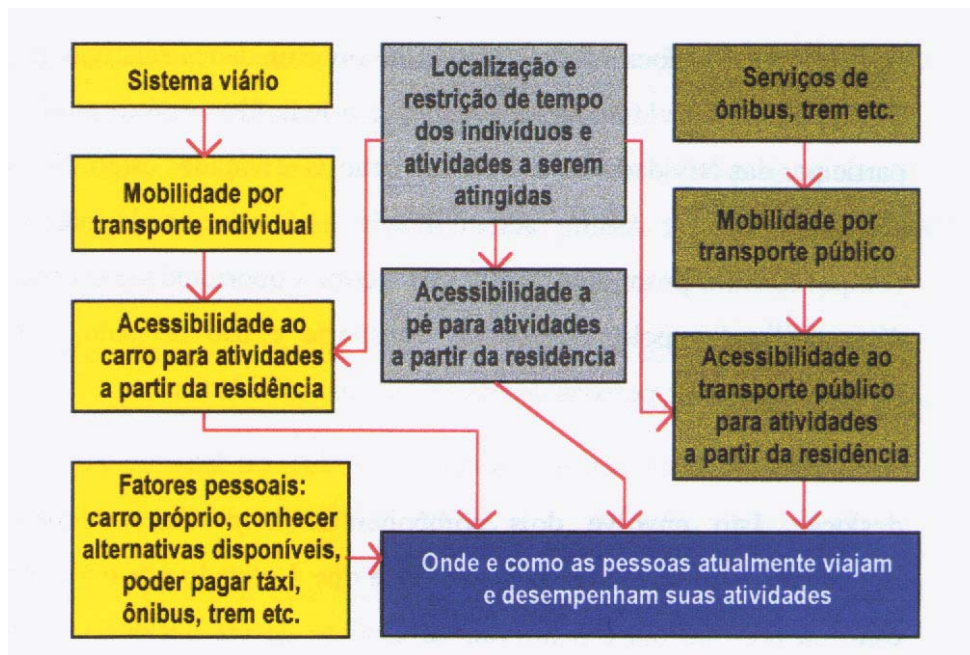
Com relação aos indicadores, os globais constituem uma aproximação mais tradicional do efeito da dependência espacial, em que o esquema geral de dependência pode ser resumido em um único valor. Pode ser obtido, por exemplo, através do *I* de Moran e do *G* de Geary, ou ainda, através do variograma (para dados pontuais). São utilizados para conhecer o grau de interação do conjunto de dados

espaciais. Os indicadores locais, também conhecidos como LISA (*Local Indicator of Spatial Association*, ou Indicador Local de Associação Espacial), por outro lado, são valores associados a cada área específica. Fornecem assim informação acerca da relevância de um agrupamento espacial de valores ao redor de cada uma destas áreas, além da soma destes valores ser proporcional ao indicador global de associação espacial. Estes indicadores são de fácil interpretação mediante a possibilidade de representação dos valores individuais de cada área através de mapas.

Já a segunda dimensão relatada em Anselin (1996), faz a distinção entre os dois modelos de dados nos quais a autocorrelação pode ser analisada, um baseado em dados pontuais e o outro em uma coleção fixa de localizações espaciais discretas (pontos ou polígonos). No primeiro caso se assume que a interação espacial é uma função contínua da distância entre pares de observações. No segundo caso a interação espacial é entendida como uma função degrau (*step function*), segundo a qual uma localização interage com um dado grupo de vizinhos. Esta segunda aproximação, segundo Serrano e Valcarce (2000) a predominante em estatística espacial, assim como a mais usada em estatística aplicada a ciências sociais, requer uma formalização da estrutura de vizinhança para cada observação, isto é, a topologia e a ordenação espacial dos dados na forma de uma matriz de pesos espaciais. Assim, em toda técnica de AEDE são encontrados três elementos básicos: matriz de proximidade espacial ( $W$ ), vetor dos desvios ( $Z$ ) e vetor de médias ponderadas ( $W_z$ ), cujo cálculo será detalhado no capítulo 3.

## 2.5. Acessibilidade

Neste trabalho de dissertação não será realizada uma revisão detalhada da literatura sobre o tema, uma vez que existem inúmeros trabalhos recentes dedicados ao assunto que fazem isto de forma bastante abrangente (tais como Bocanegra, 2002; Goto, 2000; e, em particular, Raia Jr., 2000 e Silva, 1998) e aos quais este estudo dá continuidade. Serão apenas apresentados alguns dos conceitos mais usuais de acessibilidade e mobilidade, já que ambas devem ser consideradas em conjunto, como mostra a Figura 2.1.



**Figura 2.1:** Relação entre mobilidade, acessibilidade e comportamento de viagens, adaptado de Jones (1981) por Raia Jr. (2000).

- A *Acessibilidade* está relacionada não com o comportamento propriamente dito, mas com a oportunidade ou potencial, disponibilizados pelo sistema de transporte e uso do solo, para que diferentes tipos de pessoas desenvolvam suas atividades.
- A *Mobilidade* está relacionada com a capacidade de um indivíduo se deslocar de um lugar para o outro. Isto envolve basicamente dois componentes, que são: efetividade do sistema de transporte em conectar localidades espacialmente separadas, e o limite da capacidade (inclusive financeira) de um determinado indivíduo para fazer uso do sistema de transporte.

Além disso, Davidson (1995), por exemplo, definiu a acessibilidade como a facilidade com que as pessoas em um ponto podem adquirir acesso, por meio de um sistema de transporte (ou qualquer modo ou subsistema), para todas as outras localidades de uma determinada área, levando em conta sua atividade e o custo percebido para obtê-lo. Assim, uma localidade com alta acessibilidade tenderá, para a maioria dos fins, a ser mais atraente do que uma com baixa acessibilidade e deste modo tenderá a ser mais valorizada.

### 3. ANÁLISE ESPACIAL DE ÁREAS (*LATTICE DATA*)

---

Este capítulo apresenta as técnicas de análise espacial aplicadas a dados de área, cujo objetivo é detectar e explicar a presença de padrões e tendências observadas. Assim, são aqui descritas as principais ferramentas utilizadas para analisar este tipo de variável, através de um exemplo prático.

#### 3.1. Matriz de Proximidade Espacial

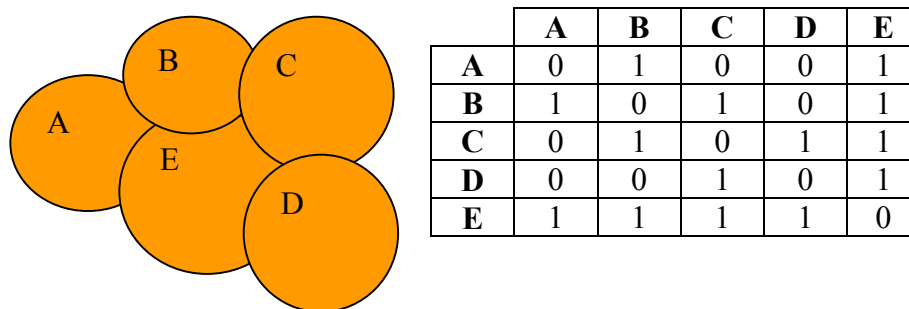
Segundo Serrano e Valcarce (2000), é possível detectar uma certa semelhança entre os conceitos de autocorrelação espacial e temporal, na medida que em ambos os casos se produz um não cumprimento das hipóteses de independência entre as observações amostrais. Contudo existe uma importante diferença entre esses casos, pois a dependência temporal é unidirecional (o passado explica o presente). Já a dependência espacial é multidirecional, isto é, uma região pode estar afetada não somente por outra região adjacente a ela, mas sim por todas as outras que a rodeiam. Esta característica faz com que técnicas para análise de séries temporais não possam ser aplicadas a dados espaciais.

Quando se trata do estudo da dependência espacial de dados de áreas, o elemento chave desta análise é o conceito da matriz de proximidade espacial. A matriz de proximidade espacial ( $W$ ) estima a variabilidade espacial de dados de áreas, sendo uma ferramenta muito útil para descrever o arranjo espacial dos objetos. É uma matriz quadrada, não estocástica, cujos elementos  $w_{ij}$  refletem a intensidade da interdependência existente entre as regiões  $i$  e  $j$ . Na matriz  $W$  ( $n \times n$ ) cada elemento  $w_{ij}$  representa uma medida de proximidade entre  $A_i$  e  $A_j$ , ou seja,  $w_{ij}$ , representa uma medida de proximidade espacial entre o polígono  $i$  e o polígono  $j$ , sendo  $A_i$  e  $A_j$  as zonas que estão sendo analisadas, e onde  $w_{ii} = 0$ .

O valor  $W$  pode ser baseado em algumas medidas de adjacência, tais

como:

- $w_{ij} = 1$ , se o centróide de  $A_i$  está a uma determinada distância de  $A_j$  e  $w_{ij} = 0$ , caso contrário (como no exemplo da Figura 3.1);
- $w_{ij} = 1$ , se  $A_i$  compartilha um lado comum com  $A_j$  e  $w_{ij} = 0$ , caso contrário;
- $w_{ij} = l_{ij}/l_i$ , onde  $l_{ij}$  é o comprimento da fronteira entre  $A_i$  e  $A_j$  e  $l_i$  é o perímetro de  $A_i$ .



**Figura 3.1:** Áreas e matriz de proximidade espacial.

Muitas vezes torna-se importante especificar o grau de proximidade de diferentes ordens, denominadas por faixas de distâncias (ou “lags”), onde a proximidade de primeira ordem é obtida pela conexão de duas zonas, a de segunda ordem por zonas que não se conectam diretamente mas são contíguas, e assim por diante.

Em alguns casos, costuma-se normalizar as linhas da matriz  $W$ , dividindo cada elemento  $w_{ij}$  pela soma total da linha a que pertence, de tal modo que a soma de cada linha seja igual a um, isto é, fazendo-se  $w_i = \sum_j . w_{ij} = 1$ . Assim, os pesos  $w_{ij}$  associados à área  $i$  somam 1.

A matriz de proximidade espacial é utilizada pela sua simplicidade, embora apresente algumas restrições. Por ser uma matriz simétrica, não permite incorporar influências não recíprocas, violando portanto o segundo dos cinco princípios básicos da estatística espacial (descritos no capítulo 2), o princípio da Assimetria.

### 3.2. Média Móvel Espacial

Para o cálculo da média móvel espacial é necessário conhecer o vetor dos desvios. O vetor de desvios ( $Z$ ) é obtido a partir da média geral ( $\mu$ ). Cada elemento deste vetor é calculado subtraindo-se o valor de  $\mu$  do valor do atributo para cada região, ou seja,

$$z_i = y_i - \mu \quad (1)$$

onde

$z_i$ : vetor dos desvios;

$y_i$ : vetor com os valores do atributo analisado;

$\mu$ : média geral.

O vetor das médias ponderadas ( $W_z$ ), também chamado de média dos valores dos vizinhos ou média móvel espacial, é uma medida útil para o cálculo da variação da tendência espacial. A estimativa da média móvel espacial é obtida pela seguinte expressão (2):

$$\hat{\mu}_i = \frac{\sum_{j=1}^n w_{ij} y_j}{\sum_{j=1}^n w_{ij}} \quad (2)$$

onde

$\hat{\mu}_i$ : médias ponderadas;

$\sum_{j=1}^n w_{ij}$ : matriz de proximidade espacial;

$y_i$ : valor do atributo.

A utilização da média móvel espacial ( $W_z$ ) permite a identificação de padrões e tendências espaciais. Pode ser empregada para estudar a dependência espacial do valor de determinado atributo através da verificação da análise da autocorrelação dos valores de acordo com sua localização no espaço, advindo daí o conceito de autocorrelação espacial. Segundo Câmara *et al.* (2001), a autocorrelação espacial mede o quanto o valor observado de um atributo numa região é independente dos valores desta mesma variável nas localizações vizinhas.

### 3.3. Indicadores Globais de Autocorrelação Espacial

O conceito de “correlação”, no contexto da estatística clássica, diz respeito ao relacionamento entre duas variáveis: se são intervalares usa-se como medida o coeficiente de Pearson ( $\rho$ ), que mede o grau de correlação linear entre as variáveis. A informação que se busca através do cálculo da autocorrelação espacial é de quanto o valor de uma variável em uma área é parecido com o do vizinho mais próximo e quão diferente do vizinho mais distante. A autocorrelação espacial ocorre quando observações organizadas no espaço influenciam-se mutuamente. Segundo Câmara *et al.* (2001), a autocorrelação espacial mede o quanto o valor observado de um atributo numa região é independente dos valores desta mesma variável nas localizações vizinhas.

De acordo com Serrano e Valcarce (2000), o valor da autocorrelação espacial pode ser positivo ou negativo. Diante de um caso de elevado nível de autocorrelação espacial positiva, os valores observados em uma localização tenderão a ser similares aos valores de seus vizinhos, isto é, a presença de um determinado fenômeno em uma certa região influencia as regiões próximas com o mesmo fenômeno. Em casos extremos, o valor em uma localização permitirá prever exatamente os valores observados nas localizações vizinhas.

Quando existe um nível elevado de autocorrelação espacial negativa, os valores baixos em uma localização estão rodeados de valores altos, quer dizer, a média ponderada é maior para os vizinhos do que o valor observado na própria localização, ou vice-versa, isto é a presença deste fenômeno dificulta a sua aparição em regiões vizinhas. Esta diferença pode ser mostrada por gráficos de setores e gráficos de barras.

A aparição da autocorrelação espacial pode ocorrer por erros de medidas ou por fenômenos de interação espacial. Uma das formas de medir a autocorrelação espacial é através do índice global de Moran. Segundo Neves *et al.* (2001), este índice, cujo valor varia de  $-1$  a  $+1$ , fornece uma medida geral da associação espacial. Valores próximos de zero indicam a inexistência de autocorrelação espacial significativa entre os valores dos objetos e seus vizinhos. Valores positivos para o índice indicam autocorrelação espacial positiva, ou seja, o valor do atributo de um objeto tende a ser semelhante aos valores dos seus vizinhos. Valores negativos para o índice indicam autocorrelação espacial negativa. Este índice é uma das formas de detecção de similaridade entre as áreas, pode ser descrito em função dos elementos básicos vistos

anteriormente, e é obtido através da expressão (3):

$$I = \frac{Z'W_z}{Z'Z} \quad (3)$$

onde

I: índice de autocorrelação espacial;

$Z^t$ : vetor transposto dos desvios;

$W_z$ : média móvel espacial.

Segundo Câmara *et al.* (2001), este indicador é uma medida de correlação espacial usada para detectar afastamentos de uma distribuição espacial aleatória, o que permite observar a existência de padrões espaciais, isto é permite detectar a similaridade de áreas de primeira ordem. Os desvios com relação à média são multiplicados pelos desvios da vizinhança. O índice de Moran testa se as áreas vizinhas apresentam maior semelhança quanto ao indicador estudado do que o esperado num padrão aleatório. Tendo em mãos estes resultados, a maior preocupação deve ser com sua validade estatística. Para verificar se a correlação obtida é significativa, duas abordagens são possíveis:

- **Teste de pseudo-significância:** são geradas diferentes permutações dos valores de atributos associados às regiões, onde cada permutação gera um novo arranjo espacial, pois os valores são redistribuídos entre as áreas. Como apenas um dos arranjos corresponde à situação observada, é construída uma situação empírica de I. Caso o valor de I corresponda ao extremo da distribuição simulada, tratar-se-á de um evento com significância.
- **Distribuição Aproximada:** para um número suficiente de sub-regiões, e supondo que as variáveis aleatórias associadas a cada localização do atributo são independentes e normalmente distribuídas, assume-se que o índice I possui distribuição aproximadamente normal.



O índice global de associação espacial também pode ser obtido pela estatística  $C$  de Geary e  $G(d)$  de Getis e Ord, obtidas pelas expressões (4) e (5), respectivamente. Para o cálculo dessas equações, pode ser utilizada qualquer definição da matriz de proximidade espacial, sendo habitual normalizá-la.

$$C = \frac{N-1 \sum_{ij} w_{ij} (x_i - x_j)}{2S_0 \sum_{i=1}^N (x_i - \bar{x})} \quad i \neq j \quad (4)$$

onde

$x_i$ : valor da variável  $x$  na região  $i$ ;

$x_j$ : valor da variável  $x$  na região  $j$ ;

$w_{ij}$ : elementos da matriz de proximidade espacial;

$\bar{x}$ : média amostral;

$$S_0 = \sum_i \sum_j w_{ij}$$

$$G(d) = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij}(d) x_i x_j}{\sum_{i=1}^N \sum_{j=1}^N x_i x_j} \quad i \neq j \quad (5)$$

onde

$x_i$ : valor da variável  $x$  na região  $i$ ;

$x_j$ : valor da variável  $x$  na região  $j$ ;

$w_{ij}$ : elementos da matriz de proximidade espacial, onde  $i$  e  $j$  são considerados vizinhos sempre que se encontram dentro de uma determinada distância  $d$  (tomando neste caso  $w_{ij}(d)$  um valor igual a 1, ou 0 caso contrário).

Outro instrumento útil para analisar o grau de dependência espacial de uma variável é através do *Diagrama de Espalhamento* de Moran, um dos indicadores locais descritos mais adiante.

### 3.4. Indicadores Locais de Autocorrelação Espacial

Segundo Câmara *et al.* (2001), os indicadores globais de autocorrelação espacial fornecem um único valor como medida de associação espacial para todo o conjunto de dados, que é útil para caracterização de toda região de estudo. Muitas vezes

é necessário, no entanto, examinar os padrões com maior nível de detalhes. Para tal, é necessária a utilização de indicadores que possam ser associados a diferentes localizações de uma variável distribuída espacialmente. Quando a área de estudo possui várias divisões, provavelmente aparecerão diferentes regimes de associações espaciais.

Os indicadores de associações espaciais locais produzem um valor específico para cada objeto (área), permitindo assim, a identificação de agrupamentos de objetos com valores de atributos semelhantes (*clusters*), objetos anômalos (*outliers*) e de regimes espaciais, que não são identificados através do cálculo do indicador de associação espacial global.

Segundo Anselin (1996), o indicador local de associação espacial é qualquer estatística que atenda as seguintes condições:

- Permite a identificação de padrões de associação espacial significativa para cada área da região de estudo;
- Constitui uma decomposição do índice global de associação espacial, isto é, a soma dos indicadores de associação espacial local para todas as observações é proporcional ao indicador de associação espacial global.

O  $I_i$  de Moran local, se for calculado com desvios em relação à média, pode ser obtido através da expressão (6). Esta forma computa o índice local do objeto  $I_i$ , através da multiplicação do desvio correspondente,  $z_i$ , pela média dos desvios de seus vizinhos.

$$I_i = \frac{z_i \sum_j w_{ij} z_j}{\sum_{i=1}^n z_i^2} \quad (6)$$

onde

$z_i$ : valor da variável normalizada na região  $i$ ;

$w_{ij}$ : elementos da matriz de proximidade espacial.

Outra forma de obter os indicadores de associação espacial local é através do índice de  $G_i(d)$  de Getis e Ord, obtidos pela expressão (7).

$$G_i(d) = \frac{w_{ij}(d)x_j}{\sum_{j=1}^N x_j} \quad j \neq i \quad (7)$$

onde

$x_j$ : valor da variável não normalizada;

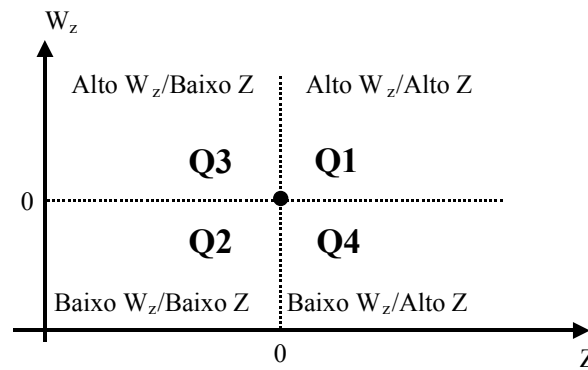
$w_{ij}$ : elementos da matriz de proximidade espacial para região  $i$  com a região  $j$  em função da distância  $d$  (matriz não normalizada de 0 e 1).

Os valores  $y$  usados para calcular os indicadores podem ser os dados originais ou, mais apropriadamente, valores normalizados para evitar a dependência escalar dos indicadores locais. No cálculo do  $I$  de Moran, por exemplo, as observações são normalizadas, subtraindo-se cada observação do valor da média. Esta normalização também deve ser realizada na matriz de proximidade espacial, como já explicado anteriormente.

Uma vez determinada a significância estatística do índice local é muito útil criar mapas indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados.

### 3.5. Diagrama de Espalhamento de Moran

Para observar graficamente as relações entre  $W_z$  x  $Z$  podem ser construídos gráficos, denominados de Diagramas de espalhamento de Moran, o que permite uma comparação visual entre os valores do atributo e de seus vizinhos de maneira quase imediata. A divisão deste tipo de gráfico em quadrantes (Q1, Q2, Q3 e Q4), que são delimitados por retas que se interceptam no ponto zero (Figura 3.2), permite observar quatro combinações possíveis de valores para  $W_z$  e  $Z$ .



**Figura 3.2:** Gráfico de espalhamento de Moran.

Os pontos localizados nos quadrantes Q1 e Q2 indicam as zonas onde o valor medido do atributo se assemelha à média dos vizinhos (valor positivo e média positiva e valor negativo e média negativa para Q1 e Q2, respectivamente), o que é um indicativo de autocorrelação espacial positiva. Pontos localizados nos quadrantes Q3 e Q4, indicam que o valor do atributo medido para determinada zona não se assemelha à média dos seus vizinhos (valor negativo e média positiva e valor positivo e média negativa para os quadrantes Q3 e Q4, respectivamente), indicativo de autocorrelação espacial negativa. As regiões localizadas nestes últimos dois quadrantes podem ser vistas como extremos, uma vez que não obedecem ao padrão observado para seus vizinhos.

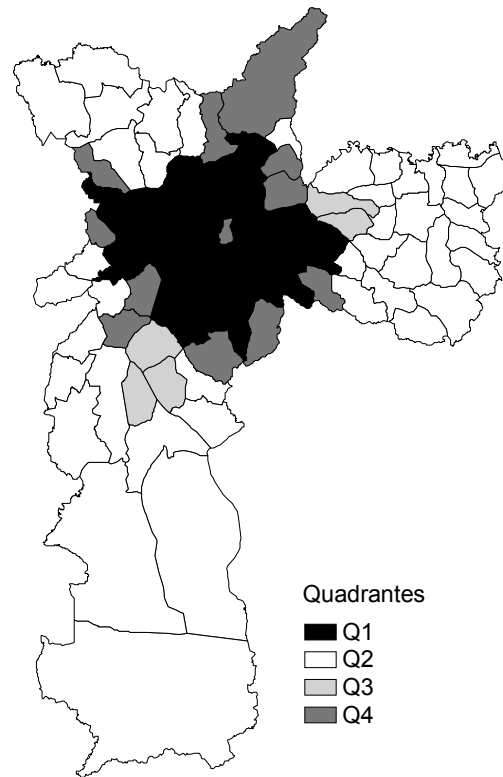
A partir dos gráficos de espalhamento de Moran pode-se construir o *Box Map*. Este tipo de mapa permite a visualização de cada elemento de acordo com sua classificação (níveis de  $W_z$  de acordo com os níveis de  $Z_i$ ). Desta forma, é possível verificar visualmente a correlação entre o atributo medido para determinada zona e o mesmo atributo medido para seus vizinhos.

### 3.6. *Box Map*, *LISA Map* e *Moran Map*

Segundo Neves *et al.* (2001) estes três dispositivos gráficos são baseados nos resultados obtidos para os indicadores locais e do gráfico de espalhamento de Moran.

Para Serrano e Valcarce (2000), uma das formas de identificar *outliers* espaciais, para dados de área, é através do *Box Map*. Este mapa é uma extensão do Diagrama de Espalhamento de Moran, onde os elementos de cada quadrante do gráfico

de espalhamento de Moran são representados por uma cor específica com seus respectivos polígonos (como mostra a Figura 3.3, que indica a distribuição do percentual de idosos dos bairros de São Paulo).

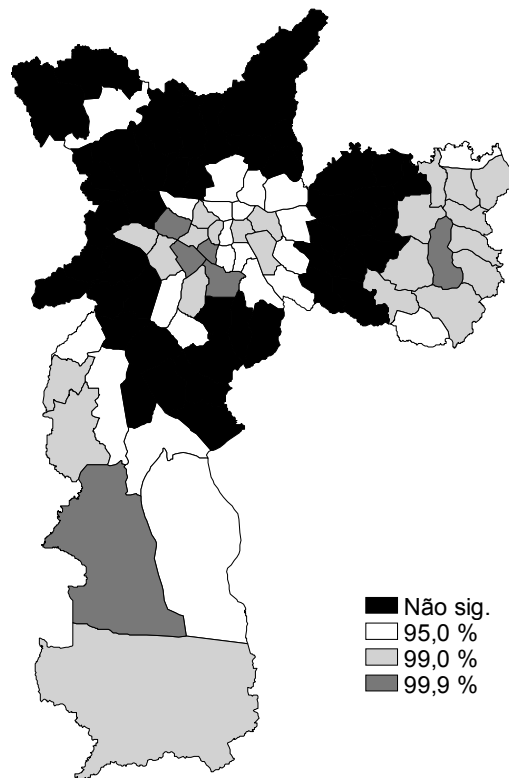


**Figura 3.3:** *Box Map* do Mapa de Exclusão/Inclusão Social de São Paulo, 2000.

Fonte: <http://www.dpi.inpe.br/geopro/exclusao/mapas.html>.

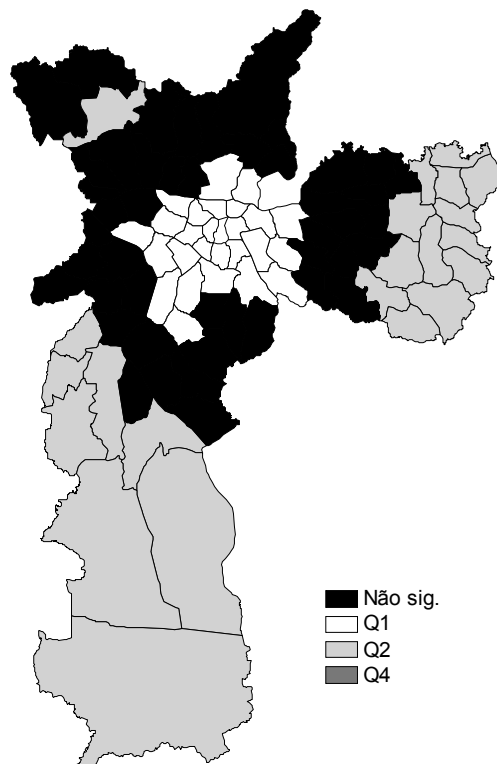
Para análise dos Indicadores Locais de Associação Espacial o mapa denominado *LISA Map* indica as regiões que apresentam correlação local significativamente diferente do resto dos dados. Estas áreas possuem dinâmica espacial própria e merecem análise detalhada. Na geração deste mapa os índices são classificados nos seguintes grupos, como mostra a Figura 3.4: não significantes (valor zero na legenda), com significância de 95 % (valor um na legenda), 99 % (valor dois na legenda) e 99,9 % (valor três na legenda).

Por fim, outra visualização que também pode ser executada é o *Moran Map* (Figura 3.5). De forma semelhante ao *LISA Map*, somente os valores significantes de Indicadores Locais de Associação Espacial são apresentados, porém classificados em quatro grupos, conforme os quadrantes a que pertencem no gráfico de espalhamento de Moran, além dos não significantes (que recebem o valor zero na legenda).



**Figura 3.4:** *LISA Map* do Mapa de Exclusão/Inclusão Social de São Paulo, 2000.

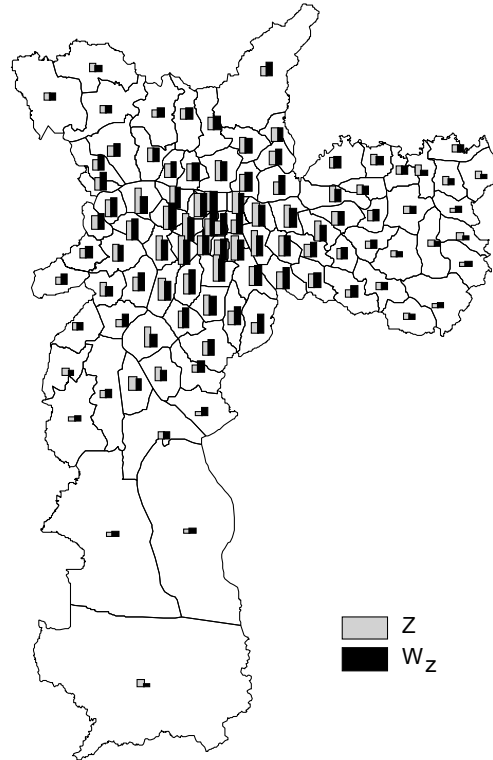
Fonte: <http://www.dpi.inpe.br/geopro/exclusao/mapas.html>.



**Figura 3.5:** *Moran Map* do Mapa de Exclusão/Inclusão Social de São Paulo, 2000.

Fonte: <http://www.dpi.inpe.br/geopro/exclusao/mapas.html>.

O gráfico de barras  $Z \times W_z$  (Figura 3.6) permite ainda visualizar simultaneamente o valor do atributo ( $Z$ ) e do valor correspondente à sua vizinhança ( $W_z$ ). Estes valores são expressos por barras sobre às áreas correspondentes aos objetos e a altura destas barras corresponde aos valores de  $Z$  e  $W_z$ .



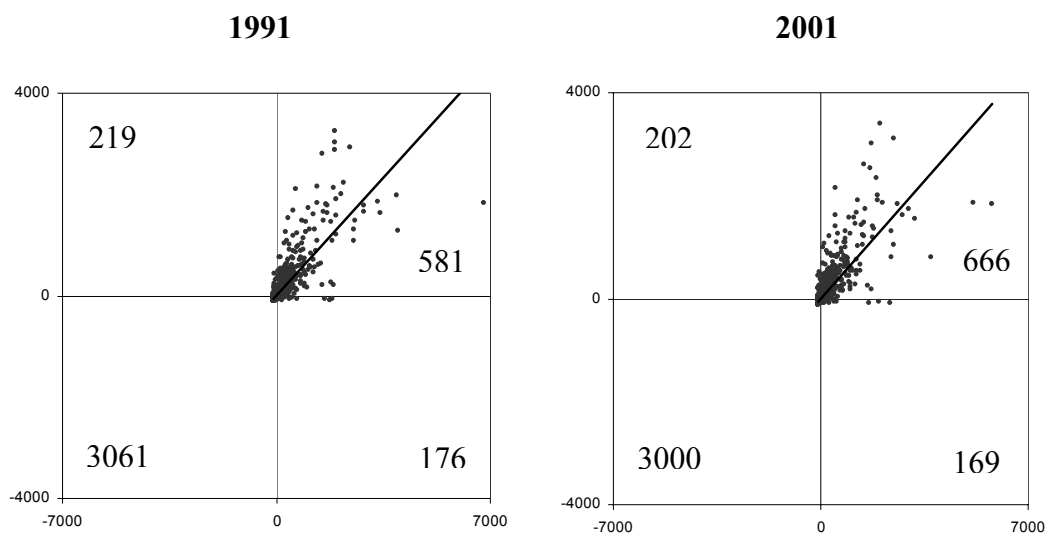
**Figura 3.6:** Gráfico de barras do Mapa de Exclusão/Inclusão Social de São Paulo, 2000. Fonte: <http://www.dpi.inpe.br/geopro/exclusao/mapas.html>.

### 3.7. Aplicação da técnica

Para melhor ilustrar os conceitos de estatística espacial foi aqui empregado um exemplo prático aplicado na região noroeste de Portugal, como uma proposta para delimitação de região metropolitana. Este exemplo foi retirado do artigo escrito por Ramos e Silva (2003), onde a ênfase é dada à avaliação de autocorrelação espacial. No caso particular do presente estudo não se pretende apenas estabelecer um valor global para essa avaliação de autocorrelação espacial, mas realizar uma análise local dessa avaliação. Neste processo de associação devem ser aplicadas técnicas que considerem explicitamente a presença de autocorrelação espacial, tais como: dispositivos de visualização de distribuições e relações espaciais, inclusive associações espaciais locais.

A aplicação da metodologia proposta é realizada em três etapas. Na primeira etapa desenvolve-se a análise espacial, através da representação de Gráfico de Moran, relativamente às condições existentes e definidas de acordo com os dados dos dois últimos Censos gerais de Portugal. Do resultado da análise, que identifica fundamentalmente a distribuição espacial dos quadrantes pelo território em diferentes períodos de tempo, desenvolve-se a regra de transição que irá constituir o modelo de previsão que será implementado para um período futuro. Por fim, é identificada qual a distribuição das zonas pelos quadrantes para a atual delimitação da Área Metropolitana do Porto e para as delimitações propostas de acordo com a evolução identificada no futuro.

Numa primeira fase os dados relativos a 1991 foram reorganizados para corresponderem às 4037 freguesias (divisão administrativa em Portugal) registradas em 2001, de modo a que a divisão espacial fosse coincidente para os dois períodos. A análise foi desenvolvida recorrendo a ferramentas disponíveis no *software* de Sistemas de Informação Geográfica *ArcView* (ESRI, 1996), em conjunto com uma extensão denominada *Spacestat* (Anselin e Bao, 1997; Anselin e Smirnov, 1998). Através dessa extensão foi possível obter a matriz de proximidade espacial entre freguesias, a qual foi posteriormente utilizada para cálculos efetuados recorrendo a uma planilha eletrônica. Após a conversão dos dados absolutos dos edifícios por freguesia em densidades (edifícios por unidade de área), foi possível sintetizar os resultados obtidos nos dois gráficos e nos dois mapas apresentados nas Figuras 3.7 e 3.8, respectivamente.

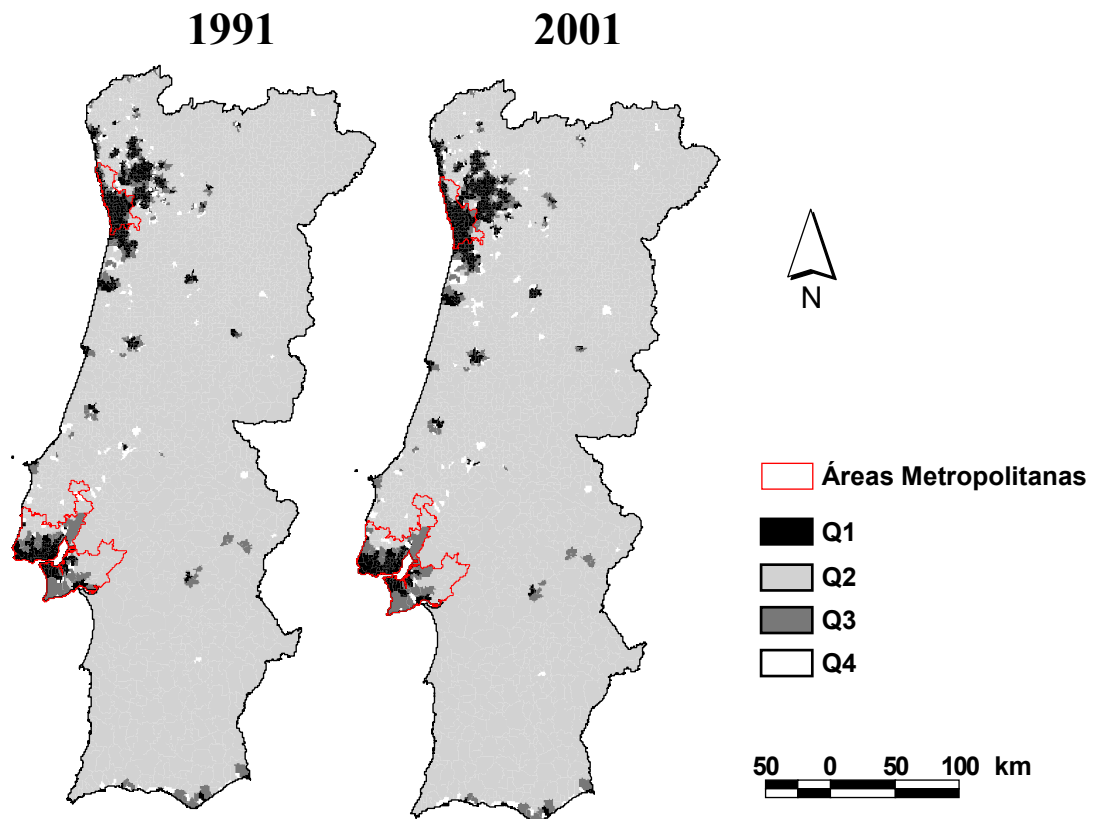


**Figura 3.7:** Gráficos de Moran para a variável densidade de edifícios por freguesia em Portugal nos anos de 1991 e 2001.



Os valores obtidos para o índice I de Moran para a variável densidade de edifícios em Portugal nos anos de 1991 e 2001 de acordo com a equação (3) foram, respectivamente, 0,6561 e 0,6558. Estes valores indicam uma elevada correlação espacial para a variável nos dois períodos. A Figura 3.7 apresenta a distribuição dos valores de Z e Wz nos dois anos, podendo-se detectar pontos nos quatro quadrantes (os valores numéricos inseridos em cada um dos quadrantes identificam o número de pontos aí localizados). Pontos situados nos quadrantes Q1 e Q2 identificam zonas em que o atributo é semelhante à média dos atributos das zonas vizinhas. É nestes dois quadrantes que se situa a maioria dos pontos, particularmente no quadrante 2, em que mais de 3000 pontos se situam na proximidade da intersecção dos eixos  $x$  e  $y$ . A distribuição territorial das freguesias, a que os pontos representados nos gráficos correspondem, pode ser vista nos mapas temáticos da Figura 3.8.

Pela análise dos mapas apresentados na Figura 3.8 identifica-se que a maioria das freguesias correspondentes aos quadrantes 1 dos gráficos da Figura 3.7 se situa em duas áreas bem identificadas, dentro ou nas proximidades das atuais áreas metropolitanas de Lisboa e do Porto. Nessas freguesias existe uma correlação positiva entre o atributo da freguesia e a média dos atributos das freguesias vizinhas, em ambos os casos com valores de densidade de edifícios superiores à média obtida para todo o território. Na verdade não constitui uma surpresa que a maioria dos pontos do quadrante 1 se situe dentro das fronteiras das atuais regiões metropolitanas, cujos limites atuais também estão representados nos mapas da Figura 3.8. Os pontos pertencentes ao quadrante 2 constituem a maioria das situações e distribuem-se por todo o território de Portugal continental, apesar de altamente concentrados no gráfico. Novamente existe uma similaridade entre o valor da freguesia e a média das freguesias vizinhas. Neste caso ambos os valores estão abaixo da média de todo o território. É também interessante analisar, na Figura 3.8, a distribuição espacial das freguesias que se situam nos quadrantes 3 e 4, por constituírem zonas de transição.



**Figura 3.8:** Mapas temáticos representando a distribuição espacial dos pontos do gráfico de Moran para a variável densidade de edifícios em Portugal nos anos de 1991 e 2001.

Este capítulo abrangeu as técnicas de Estatística Espacial com ênfase em dados de áreas (dados estes que serão utilizados nesta pesquisa), tentando explicar de maneira clara e simples as técnicas mais aplicadas neste caso.

No próximo capítulo será apresentada uma aplicação destas técnicas na cidade de Bauru, com ênfase na topologia dos dados e obtenção dos mapas, já que atualmente as técnicas de estatística podem ser empregadas de forma mais fácil e direta, devido à disponibilidade de *software* gratuito na Internet.

### 3.8. Considerações Finais

Muitos dos indicadores aqui apresentados produzem um valor específico para cada zona e permitem, entre outras coisas, a identificação de agrupamentos de objetos com valores de atributos semelhantes. No caso particular deste estudo, espera-se que seja possível identificar bolsões cujo valor de acessibilidade seja alto ou baixo, o que permitirá caracterizar esta variável do ponto de vista de sua distribuição espacial.

## 4. METODOLOGIA

---

Inicialmente serão aqui reafirmados os objetivos específicos da pesquisa, os mesmos que foram formulados na introdução desta dissertação. Em seguida serão discutidos detalhes dos procedimentos utilizados na obtenção dos dados, na construção da base de dados, na correta caracterização da sua topologia, e na aplicação das técnicas de estatística espacial para o cálculo dos índices globais e locais de autocorrelação e para a construção dos gráficos de espalhamento e mapas deles derivados.

### 4.1. Abordagem Adotada

A metodologia utilizada nesta pesquisa seguiu, em linhas gerais, os passos do livro de Serrano e Valcarce (2000) e do artigo publicado por Ramos e Silva. (2003). Ambas as obras tentam explorar as técnicas de estatística espacial, a primeira com conteúdo teórico expressivo e a segunda através de uma aplicação prática.

Como descrito no capítulo 1, a presente pesquisa visa explorar ferramentas de estatística espacial para analisar as relações entre aspectos da distribuição espacial de características da população (neste caso, a renda) de uma cidade média brasileira e os diversos níveis de acessibilidade/mobilidade por diferentes modos de transporte nela observados (aqui entendida como a proporção de viagens realizadas em cada um dos modos de transporte considerados). Para tal, procurou-se utilizar apenas pacotes computacionais disponíveis (seja por já terem sido adquiridos pelo Departamento de Transportes da EESC/USP, seja por se tratarem de *software* livre) para o seu desenvolvimento, a fim de viabilizar um projeto de baixo custo e com forte cunho didático, já que esta é uma das primeiras pesquisas sobre esta temática desenvolvida no país na área de transportes. Assim, a análise foi conduzida nos *software* SPRING e TransCAD, utilizando dados de uma pesquisa O/D já existente.

## 4.2. Obtenção dos Dados Básicos

Um dos problemas freqüentemente observados em pesquisas empíricas encontra-se na fase de levantamento dos dados, já que após a definição dos objetivos do trabalho e das variáveis exigidas para o desenvolvimento da análise é necessário proceder à coleta dos dados correspondentes. Embora muitos dados úteis para o planejamento urbano e de transportes já sejam atualmente disponibilizados pelo Instituto Brasileiro de Geografia e Estatística ou por outras fontes (públicas e privadas), o que reduz custos e economiza tempo para projetos de natureza semelhante a este, muitos outros dados igualmente importantes não estão disponíveis. Isto pode demandar uma série de procedimentos para sua obtenção, envolvendo a formulação de questionários, o treinamento de equipes para sua aplicação, e a tabulação dos resultados. Somente depois de todos estes procedimentos e de uma verificação da qualidade dos dados obtidos no campo se pode então executar as análises pretendidas.

Por outro lado, quando os dados estão disponíveis *a priori*, várias etapas podem ser evitadas, o que pode conduzir o pesquisador diretamente à fase de tabulação. É nesta fase que se pode, como forma de reduzir a duração da etapa, selecionar e trabalhar apenas com os dados necessários para a pesquisa em questão. Uma grande variedade de dados disponíveis, no entanto, nem sempre elimina totalmente a necessidade de submetê-los a transformações para adequação das variáveis ao modelo a ser utilizado.

Como neste trabalho a fonte principal de dados foi uma pesquisa Origem/Destino (O/D), este e os próximos parágrafos serão dedicados a este tema. Com o objetivo de caracterizar os deslocamentos efetuados pela população em suas atividades diárias, tais como viagens por motivo de trabalho, estudo, compras, etc., são realizadas as pesquisas de origem e destino, que constituem levantamentos de dados de viagens dos mais abrangentes. Um exemplo de sua grande importância é o caso da cidade de São Paulo, onde desde 1967 foram realizadas quatro pesquisas O/D envolvendo não só a capital, mas também sua região Metropolitana. A primeira delas abrangiu quinze municípios, que correspondiam à mancha urbanizada contínua da Região Metropolitana de São Paulo (RMSP), e subsidiou os estudos para implantação da primeira linha de metrô na capital paulista.

Várias pesquisas O/D foram realizadas em outras cidades brasileiras para fundamentar estudos de melhorias no sistema de transportes, entre os quais estão: Belém, Belo Horizonte, Brasília, Fortaleza, Goiânia, Natal, Porto Alegre, Recife, Salvador, Rio de Janeiro e Campinas (esta última de 2003, portanto uma das mais recentes pesquisas O/D que se tem notícia no Brasil).

No presente trabalho utilizou-se como fonte de dados uma pesquisa O/D realizada em 1997 na cidade de Bauru, através da qual foram levantados os dados sobre as viagens internas, já que a mesma possui duas pesquisas distintas e complementares que são: pesquisa domiciliar e de linha de contorno, esta última para registrar viagens externas.

A pesquisa domiciliar deve ser realizada em uma amostra representativa de domicílios, selecionada em cada uma das zonas de tráfego. Depois de realizados os treinamentos adequados para aplicação dos questionários, os moradores sorteados para responder a pesquisa são então entrevistados, devendo informar a respeito das características de todas as viagens realizadas no dia anterior, que devem corresponder a um dia típico da semana. São coletadas informações como endereço de origem e destino das viagens, horário de início e fim, motivos, modos de transporte utilizados (ônibus, trem, motocicleta, bicicleta, etc.), tempos de acesso aos modos de transporte, além de outras informações relativas a características socioeconômicas (como renda, sexo, idade, posse de automóveis, etc.)

Para obtenção dos dados, a área da pesquisa deve ser dividida em grandes bolsões, definidos geralmente por barreiras físicas, tais como componentes do sistema viário ou linhas de trem. A pesquisa levanta o fluxo de pessoas ou veículos que passam por esses pontos de cruzamento, classificando-os por modo, horário ou motivo. Neste caso é então necessário entrevistar os ocupantes para verificar o motivo da viagem. A quantidade de informações coletadas neste tipo de pesquisa permite que sejam utilizadas em outras áreas, não limitando-se apenas ao planejamento de transportes.

Uma vez estabelecido o objetivo do trabalho, a maior preocupação passa a ser a de assegurar dados que atendam às necessidades do estudo dentro dos recursos disponíveis para a sua obtenção. Isto faz com que, de maneira geral, apesar da pesquisa O/D ser uma das mais ricas fontes de dados para estudos de transportes, raramente seja realizada apenas para fins de pesquisa acadêmica, em virtude de seu alto custo. Assim sendo, somente se já estiver disponível este tipo de pesquisa poderá ser utilizado em trabalhos como o que aqui se propõe.

Adicionalmente, com a pesquisa em mãos, uma das maiores preocupações das técnicas de estatística é a verificação do tamanho da amostra e a validação dos resultados, isto é, a verificação de sua adequação para a aplicação das técnicas necessárias ao estudo, bem como a correção de possíveis erros nos dados, de tal modo que a pesquisa torne-se confiável e fidedigna.

Posteriormente, pode-se proceder à agregação dos dados individuais por domicílios ou por zonas, como no caso específico deste estudo, em que foram somadas as viagens realizadas por cada um dos modos considerados de acordo com os limites dos setores censitários.

Em síntese, o material básico necessário para dar início ao desenvolvimento deste trabalho resumiu-se aos limites dos setores censitários da Contagem Populacional de 1996 e diversos dados de uma pesquisa O/D realizada em 1997. Os dados relativos aos limites dos setores censitários foram adquiridos junto ao Instituto Brasileiro de Geografia e Estatística (IBGE) e foram também utilizados por Bocanegra (2002). Todos os demais dados foram os mesmos empregados nos trabalhos de López-Reyes (1999), Raia Jr. (2000) e Bocanegra (2002).

### 4.3. Digitalização

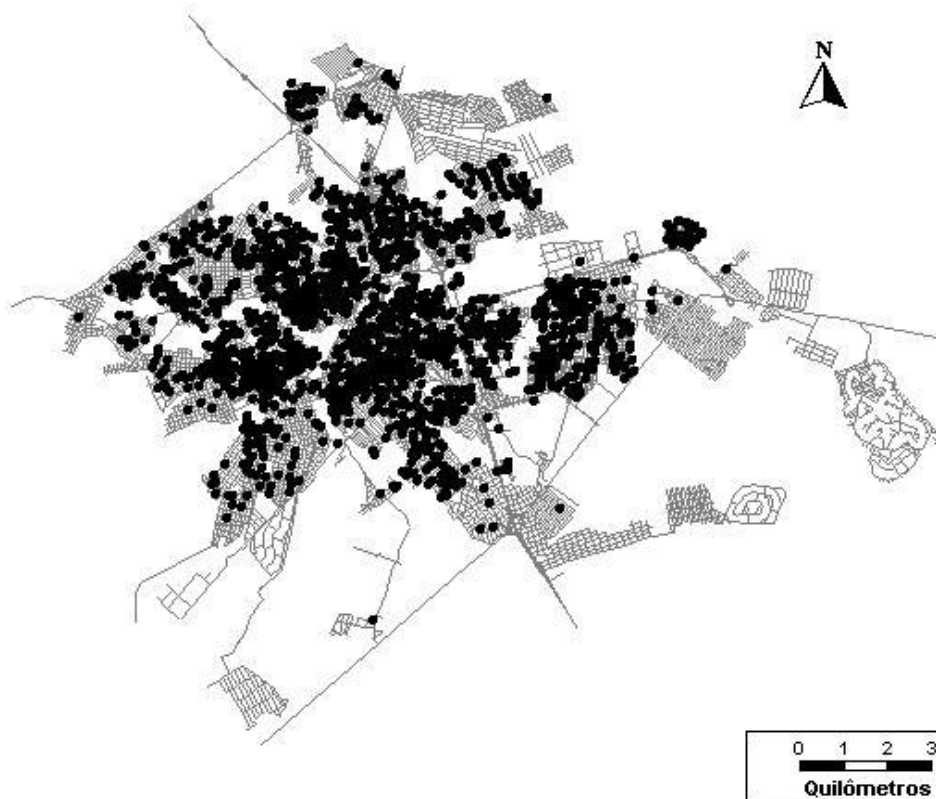
Os limites dos setores censitários foram obtidos em um mapa impresso em papel, o que exigiu a conversão de cada um desses setores. Isto foi feito diretamente no *software* TransCAD<sup>2</sup>, sendo os setores censitários tratados como áreas (Figura 4.1). Foi também necessário georreferenciar os domicílios entrevistados na pesquisa O-D de 1997, como mostrado na Figura 4.2.

---

<sup>2</sup> *Software* desenvolvido pela empresa Caliper Corporation. 1172 Beacon Street, Newton, Massachusetts, USA. Página na *internet* em [www.caliper.com](http://www.caliper.com).



**Figura 4.1:** Setores censitários da cidade de Bauru, em 1996.



**Figura 4.2:** Distribuição dos domicílios entrevistados através da pesquisa O-D de 1997.

O uso de bases de dados digitais georreferenciadas é fundamental para os cálculos relacionados a análise de dados espaciais. Isto é, os dados devem possuir localização geográfica expressa através de coordenadas no mapa e seus atributos devem ser armazenados em bancos de dados digitais. Estes atributos referem-se a informações sobre as ruas, as vias, os setores, etc.

A obtenção de dados gráficos para a elaboração de um mapa pode ser efetuada através da importação ou da digitalização dos vetores, mas sempre tomando-se o devido cuidado com a topologia. Na criação dos vetores no SPRING, especialmente de mapas cadastrais, temáticos e redes, é necessário realizar as etapas de digitalização, ajustes e poligonização. A primeira etapa de digitalização é o processo que transforma os dados do meio analógico para o digital, para que se possa futuramente aplicar técnicas espaciais.

De acordo com Antenucci *at al.* (1991, *apud* Raia Jr, 2000) existem algumas formas de se obter estas bases de dados: aquisição através de mesa digitalizadora, aquisição através de digitalização tomando-se como base imagens de satélites, aquisição por meio de cópia de imagens *raster* ou matriciais (“*scanning*”) servindo como “pano de fundo” para posterior digitalização via *mouse* e tela do computador, aquisição por meio da transferência de arquivos digitais existentes, etc.

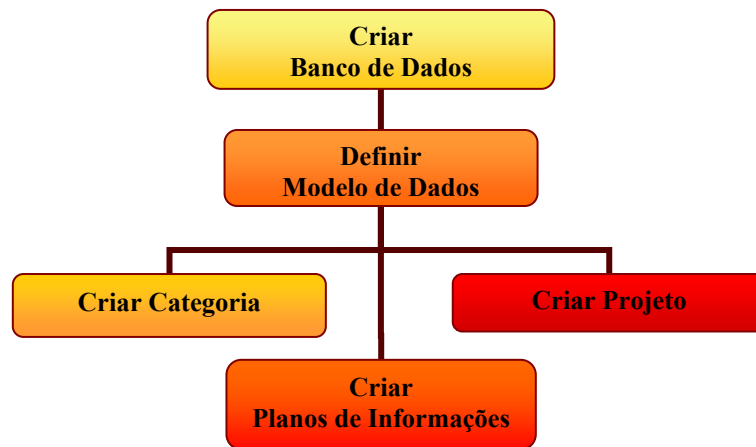
Como as análises espaciais realizadas neste trabalho foram todas executadas em ambiente SPRING (Sistema de Processamento de Informações Georreferenciadas), alguns conceitos básicos para tratar a base de dados serão descritos neste item. Para começar a trabalhar neste ambiente torna-se necessária a criação de um banco de dados, onde serão armazenadas todas as entradas dos dados e a definição do modelo. Este modelo no banco de dados especifica as categorias e classes dos dados que serão armazenados, pois antes de começar a trabalhar estas categorias já deverão estar especificadas. O conceito de categoria agrupa dados de mesma natureza no banco, definindo uma classe de dados. Cada categoria é sempre associada a um único modelo de dados e poderá conter inúmeros Planos de Informações em diferentes projetos no mesmo Banco de Dados. Apenas após a criação das Categorias será possível a criação de Planos de Informações.

Nesta análise o modelo utilizado foi do tipo cadastral, pois este pode ser usado para representar regiões, tais como setores censitários, divisões político-administrativas, etc. As características de apresentação gráfica dos dados também são definidas e armazenadas no modelo de dados do banco, assim como as tabelas que são



utilizadas para inserir, armazenar e recuperar os atributos descritivos dos dados. Para cada categoria existe uma tabela de atributos única, onde são definidos seus campos de acordo com os atributos fornecidos.

Para criação do projeto, subdiretório dentro do banco de dados, é necessário conhecer o tipo de projeção a ser utilizada e estabelecer o limite da área de estudo. O próximo passo é a criação de um plano de informação, o qual armazena um conjunto de características básicas em comum. Em um projeto cada PI está associado a uma única categoria e, conseqüentemente, a um único modelo. Quando ativado um projeto, surge a janela Painel de Controle. É através dela que os Planos de Informações e suas diferentes representações são selecionados, tanto para a visualização quanto para qualquer operação desejada. A Figura 4.3 contém, de forma resumida, os passos para criação de uma base de dados no *software* SPRING.



**Figura 4.3:** Passos para criação de uma base de dados.

#### 4.4. Importação

Para importar os dados não é necessário ter um projeto ativo, mas neste caso deve-se informar nome, projeção e retângulo envolvente do mesmo, na própria janela de importação. O SPRING aceita a importação dos dados em qualquer um destes formatos; *ASCII-SPRING*, *ARCINFO-UNG*, *DXF-R12*, *SHAPEFILE*, *JPEG*, *RAW* ou *SITIM*, sendo necessária apenas a definição das categorias adequadas a cada dado a ser importado.

O que acontece na maioria das vezes é que os mapas (Planos de Informações) importados de outros sistemas, tais como *AutoCAD*, *ArcView* etc., não

consideram a questão da topologia. Tendo o mapa em formato digital e representado por pontos, linhas e polígonos, esta relação espacial deverá ser definida explicitamente para que se possa proceder às operações de análise espacial dos dados. Assim, uma primeira preocupação no SPRING, após a importação das linhas que compõem um mapa, é a eliminação das linhas duplicadas (no caso das áreas, são as linhas de fronteiras ou divisas entre uma região e outra). Arquivos que são oriundos de outros programas devem ter estas linhas de fronteiras ou divisas eliminadas. Isto é fundamental no caso do SPRING, no qual a definição da topologia para um modelo cadastral resulta na criação dos polígonos, isto é, no armazenamento das informações referentes às linhas, nós e identificadores que compõem cada polígono, assim como as linhas que são compartilhadas por polígonos adjacentes.

Após a eliminação das linhas duplicadas procede-se a um ajuste, isto é, a criação dos nós. Estando as linhas ajustadas, realiza-se a poligonização (criação de regiões). Neste ponto a topologia deve estar criada, e o número de regiões ou áreas ou polígonos deve ser exatamente o esperado (igual ao do arquivo original importado), nem mesmo um a mais ou a menos. Uma vez criada a topologia é necessário identificar cada uma das regiões, somente após o que se poderá ligar tabelas de atributos ao mapa e por fim escolher a estatística a ser utilizada dentre as opções oferecidas pelo *software*.

#### **4.5. Análise Espacial de Áreas**

Como apresentado no capítulo 3, as técnicas de Análise Espacial de Áreas são aplicadas em eventos cuja localização está associada a áreas, o que ocorre ao se trabalhar com fenômenos agregados por bairros, municípios, setores censitários, etc. O que a diferencia das outras técnicas, como a de análise de dados pontuais, é que não se considera neste caso a localização exata dos eventos individuais, mas sim de valores agregados por área.

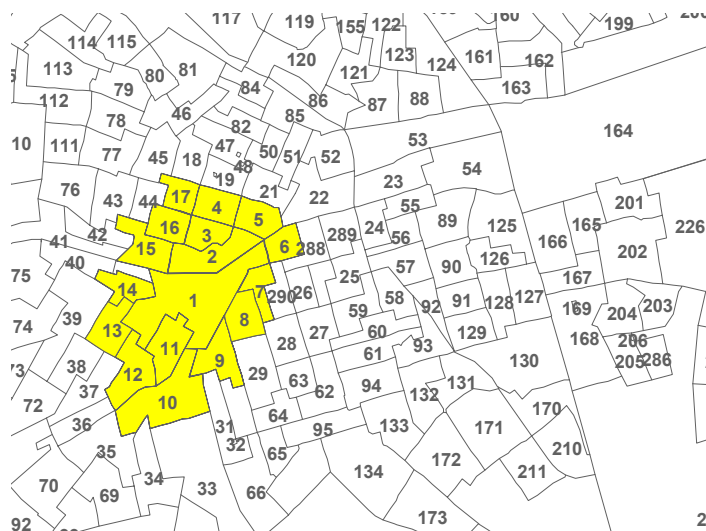
Nesta seção serão descritos inicialmente os métodos para obtenção dos resultados destinados a verificar a existência de padrões e tendências espaciais, com o objetivo de obter uma primeira aproximação da distribuição espacial da variável selecionada para análise. Como primeira aproximação desta variabilidade estimar-se-á a média móvel espacial. Assim, o cálculo da média móvel espacial levará em consideração a variabilidade de cada setor e de seus vizinhos imediatos. Como resultado, espera-se que estas primeiras aproximações sejam capazes de fornecer as

grandes tendências do fenômeno em estudo.

Para melhor compreensão da técnica foi aplicada a teoria descrita no capítulo 3 em uma amostra retirada dos dados originais. Inicialmente serão calculados os três elementos básicos e necessários para as aplicações das técnicas de Estatística Espacial aqui exploradas, que são: média móvel espacial, denominada por ( $W_z$ ), matriz de proximidade espacial ( $W$ ) e vetor dos desvios ( $Z$ ).

O primeiro passo para a análise dos dados de áreas é o cálculo da matriz de proximidade espacial (como descrito no item 3.1), pois este é o ponto de partida para os demais cálculos. O cálculo desta matriz estima a variabilidade espacial dos dados. Alguns programas possuem rotinas pré-programadas para o cálculo desta matriz, bastando apenas selecioná-las. O SPRING (programa utilizado neste trabalho) não exige que o cálculo seja feito separadamente, pois o próprio programa entende as relações de vizinhança entre áreas. Para isto basta apenas a topologia estar correta, de tal modo que o programa “entenda” os polígonos e a conectividade entre as áreas, para proceder às análises.

Um dos modos de calcular a matriz de proximidade espacial é através do critério de lado comum, onde  $w_{ij} = 1$ , se  $A_i$  compartilha um lado comum com  $A_j$ , caso contrário  $w_{ij} = 0$ , como ilustram a Figura 4.4 e a Tabela 4.1, que contém um detalhe de 17 áreas da base de dados de setores censitários empregada neste estudo. Estes setores serão utilizados na seqüência para exemplificar alguns dos procedimentos de cálculo apresentados no capítulo 3.



**Figura 4.4:** Setores selecionados da cidade de Bauru.

Uma vez calculada a matriz de proximidade espacial, uma forma simples e útil de avaliar a variação da tendência espacial é através do cálculo das médias dos valores dos vizinhos. Para o cálculo deste valor, denominado de média móvel espacial, é necessário primeiramente calcular o vetor dos desvios para depois realizar o cálculo da média móvel espacial. As formulações para o cálculo destas estatísticas podem ser vistas no item 3.2, equações (1) e (2).

**Tabela 4.1:** Matriz de proximidade espacial.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0
2	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	1	0
3	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
4	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	1
5	0	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
6	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
7	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
8	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
9	1	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
10	1	0	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0
11	1	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
12	1	0	0	0	0	0	0	0	0	1	1	0	1	0	0	0	0
13	1	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
14	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
15	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1	0	1
17	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0

A próxima etapa é o cálculo do vetor dos desvios ( $z_i = y_i - \mu$ ), que exige o cálculo preliminar do valor médio do atributo considerando todas as zonas que são objeto da análise. Neste exemplo, a média das viagens realizadas pelo modo automóvel como motorista para os 17 setores foi de 0,269.

**Tabela 4.2:** Vetor dos desvios.

<i>%auto_moto</i>	<b>Desvio</b>
0,0976	-0,1714
0,4286	0,1596
0,2000	-0,0690
0,3125	0,0435
0,3611	0,0921
0,0000	-0,2690
0,0909	-0,1781
0,1538	-0,1152
0,3929	0,1239
0,1429	-0,1261
0,1786	-0,0904
0,2442	-0,0248
0,2273	-0,0417
0,6250	0,3560
0,2593	-0,0097
0,1667	-0,1023
0,7000	0,4310

Com os valores da matriz de proximidade espacial normalizada (Tabela 4.3 – obtida pela divisão do valor de cada célula pela soma da linha a que pertence na Tabela 4.1), e do vetor dos desvios (Tabela 4.2), pode-se calcular o terceiro elemento básico, o vetor das médias ponderadas (Tabela 4.4). Este é obtido através da multiplicação da matriz de proximidade espacial normalizada (Tabela 4.3) pelo vetor dos desvios (Tabela 4.2). Assim, o primeiro valor (-0,0185) é obtido da seguinte forma:

$$\begin{aligned}
 & ((0 \cdot -0,1714) + (0,09 \cdot 0,1596) + (0 \cdot -0,0690) + (0 \cdot 0,0435) + (0 \cdot 0,0921) + \\
 & + (0,09 \cdot -0,2690) + (0,090 \cdot -0,1781) + (0,09 \cdot -0,1152) + (0,09 \cdot 0,1239) + \\
 & + (0,09 \cdot -0,1261) + (0,09 \cdot -0,0904) + (0,09 \cdot -0,0248) + (0,09 \cdot -0,0417) + \\
 & + (0,09 \cdot 0,3560) + (0 \cdot -0,0097) + (0 \cdot -0,1023) + (0 \cdot 0,4310)) = -0,0185
 \end{aligned}$$

**Tabela 4.3:** Matriz de adjacência normalizada.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
1	0	0,09	0	0	0	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0,09	0	0	
2	0,2	0	0,2	0	0,2	0	0	0	0	0	0	0	0	0	0,2	0,2	0	
3	0	0,33	0	0,33	0	0	0	0	0	0	0	0	0	0	0	0,33	0	
4	0	0	0,33	0	0,33	0	0	0	0	0	0	0	0	0	0	0	0,33	
5	0	0,33	0	0,33	0	0,33	0	0	0	0	0	0	0	0	0	0	0	
6	0,5	0	0	0	0,5	0	0	0	0	0	0	0	0	0	0	0	0	
7	0,5	0	0	0	0	0	0	0,5	0	0	0	0	0	0	0	0	0	
8	0,33	0	0	0	0	0	0,33	0	0,33	0	0	0	0	0	0	0	0	
9	0,33	0	0	0	0	0	0	0,33	0	0,33	0	0	0	0	0	0	0	
10	0,25	0	0	0	0	0	0	0	0,25	0	0,25	0,25	0	0	0	0	0	
11	0,33	0	0	0	0	0	0	0	0	0,33	0	0,33	0	0	0	0	0	
12	0,25	0	0	0	0	0	0	0	0	0,25	0,25	0	0,25	0	0	0	0	
13	0,33	0	0	0	0	0	0	0	0	0	0	0,33	0	0,33	0	0	0	
14	0,50	0	0	0	0	0	0	0	0	0	0	0	0,50	0	0	0	0	
15	0,33	0,33	0	0	0	0	0	0	0	0	0	0	0	0	0	0,33	0	
16	0	0,25	0,25	0	0	0	0	0	0	0	0	0	0	0	0	0,25	0	0,25
17	0	0	0	0,50	0	0	0	0	0	0	0	0	0	0	0	0	0,50	0

**Tabela 4.4:** Vetor das médias ponderadas.

Média Móvel
-0,0185
-0,0521
0,0332
0,1498
-0,0217
0,0396
-0,0258
-0,0740
-0,1362
-0,0407
-0,1063
-0,1074
0,0527
-0,1065
-0,0376
0,1280
-0,0294

A partir destes elementos básicos pode-se analisar a autocorrelação espacial da variável que é o objeto de estudo desta pesquisa. Este índice é calculado através da equação (3) do item 3.3. Numa definição bem geral pode-se dizer que um índice de autocorrelação espacial é capaz de medir o quanto o valor observado de um atributo em dada região é independente desta mesma variável para as regiões vizinhas.

Para o presente trabalho, o índice a ser utilizado será o de *Moran*, que serve para testar se as áreas conectadas apresentam maior semelhança quanto ao indicador estudado do que seria esperado num padrão aleatório. Em particular, este tipo de indicador fornece um único valor como medida de associação espacial para todo o conjunto de dados (índice global). Uma das maiores preocupações com relação ao índice de *Moran* é estabelecer a significância estatística, o que o *software* SPRING avalia através do teste de pseudo-significância. Este teste gera diferentes permutações, sendo possível escolher o número desejado no *software*, entre as opções 99, 999 e 9999. Cada permutação produz um novo arranjo espacial dos valores do atributo. Se o valor do índice *I* encontrado nos dados originais corresponder ao extremo da distribuição simulada, trata-se de um evento com significância estatística.

Desta forma, se a autocorrelação espacial for verificada, o próximo passo constará da análise de indicadores locais de associação espacial.

O indicador global fornece um único valor como medida de associação espacial para toda região. Para refinar o processo que dá origem a dependência espacial que não é evidenciada pelo indicador global são calculados os indicadores de autocorrelação espacial local. Estes indicadores podem ser calculados através da equação (6), descrita no item 3.4, que fornece como resultado padrões de associação espacial significativos. Novamente aqui podem ser utilizadas técnicas semelhantes às empregadas para testar os índices globais de *Moran*, como a simulação de distribuição por permutação aleatória nos valores dos atributos.

#### **4.6. Atividades Desenvolvidas**

A integração da Estatística Espacial com SIG possibilita a aplicação de inúmeras técnicas de visualização e cálculo que permitem algumas análises interessantes, como será apresentado mais adiante. Para que isto aconteça de forma satisfatória torna-se necessário o planejamento da pesquisa, estipulando uma seqüência de atividades e direcionando o estudo de forma a atingir os objetivos propostos. Abaixo estão descritas as atividades a serem seguidas, com uma prévia do desenvolvimento dos procedimentos realizados no próximo capítulo.

1. Obtenção do mapa da cidade, em formato digital, com as áreas dos setores censitários, para uso no *software* SPRING;
2. Obtenção de dados de renda média domiciliar por setor censitário;
3. Obtenção dos dados da pesquisa O-D em formato digital;
4. Aplicação das técnicas de Estatística Espacial para as variáveis analisadas (no caso em questão, proporção de viagens por modo);
  - a. Obtenção do Índice de Autocorrelação Espacial Global;
  - b. Construção do Diagrama de Espalhamento de Moran e *Box Map*.
  - c. Cálculo dos Indicadores Locais de Autocorrelação Espacial;
  - d. Construção do *LISA Map* e *Moran Map*, a fim de identificar os agrupamentos de associação espacial local significantes;
5. Divisão da cidade em 3 anéis, a fim de analisar como a acessibilidade, pelos diferentes modos de viagens analisados, é afetada pela localização;
6. Divisão da variável renda em três níveis, a fim de analisar como a mobilidade, para os diferentes modos de viagens analisados, é afetada pela renda;
7. Construções de tabelas cruzadas com as variáveis renda e localização;
8. Construção de gráficos de setores para diferentes combinações de renda e localização;
9. Seleção de zonas em diferentes regiões da cidade, e com diferentes níveis de renda, para análise detalhada;
10. Análise dos resultados.

O próximo capítulo traz a aplicação das técnicas apresentadas neste capítulo e nos anteriores, sempre procurando conduzir as análises utilizando dados existentes e *software* de baixo custo ou fácil acesso, como é o caso do SPRING.



## 5. ANÁLISE DOS RESULTADOS

---

A seguir serão apresentados os principais resultados obtidos neste trabalho. Para tal, foram aplicadas as técnicas de estatística espacial descritas na metodologia para o cálculo do índice de autocorrelação global, local e para a construção dos gráficos de espalhamento e mapas temáticos com o propósito de melhor compreender a relação existente entre acessibilidade, renda e a distribuição da população no espaço urbano.

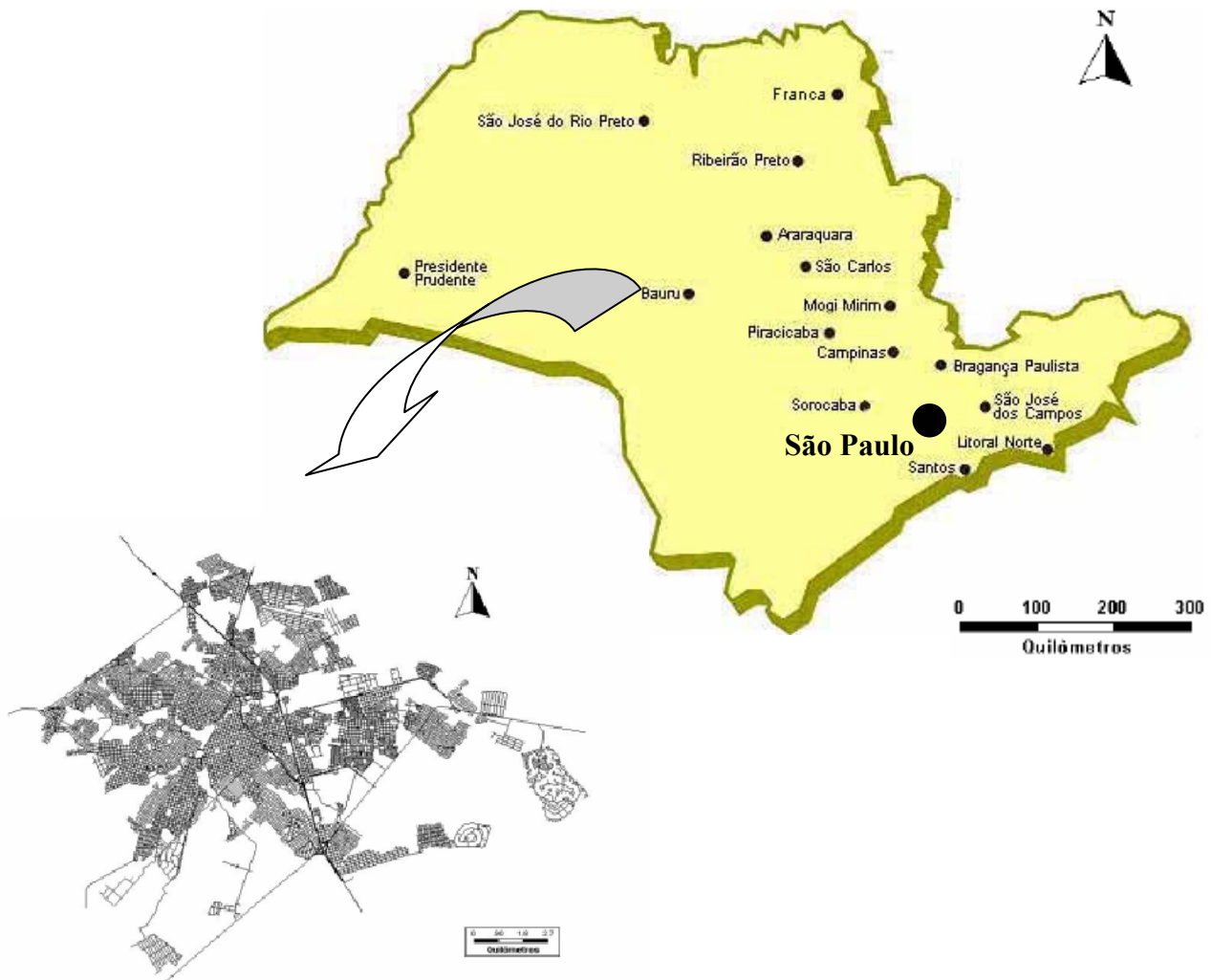
### 5.1. Estudo de Caso

A cidade de Bauru está localizada na região centro-oeste do estado de São Paulo, a 325 km da capital, como mostra a Figura 5.1. É uma cidade de médio porte com população superior a 316 mil habitantes (IBGE, 2000), tendo sido sua área urbana originalmente dividida em 306 setores censitários no levantamento empregado neste estudo. Com relação as suas coordenadas geográficas, localiza-se entre as latitudes 22° 15' S e 22° 25' S e entre as longitudes 49° W e 49° 10' W. É uma região de clima tropical, apresentando altitudes de 500 a 630 m, predominantemente planáltica e definida como uma região pouco acidentada topograficamente.

Uma mistura de agricultura (cafeicultura), ferrovias, confronto com os indígenas, além de fatos políticos e religiosos, envolvem o surgimento e desenvolvimento da cidade de Bauru. A ferrovia, por exemplo, que tanto projetou a cidade no passado hoje não mais é significativa e resta como uma referência apenas na memória da população.

O desenvolvimento e toda transformação de uma cidade é notado no crescimento de sua mancha urbana, no fluxo de capitais, pessoas, veículos, na violência urbana, no aumento da necessidade da instalação de infra-estruturas públicas, entre outras questões que merecem maior atenção. Bauru não foge à regra das demais cidades

de porte médio do interior paulista e apresenta uma série de problemas ocasionados por seu crescimento rápido, sendo hoje o transporte uma das infra-estruturas mais importantes da cidade.



**Figura 5.1:** Localização da cidade de Bauru no estado de São Paulo e sua configuração geral.

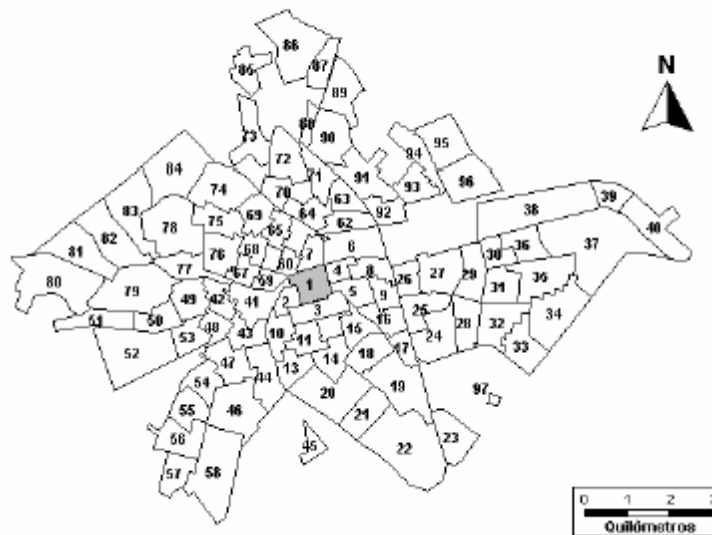
## 5.2. Base de Dados

A pesquisa O/D utilizada como fonte de dados neste estudo foi realizada no período de maio e junho de 1997, abrangendo 24 mil domicílios. Nesta época, Bauru possuía aproximadamente 300 mil habitantes.

Segundo Ortúzar e Willumsen (1994, *apud* Raia Jr., 2000), para uma cidade com número de habitantes entre 150 mil e 300 mil, o número de domicílios

abordados em uma pesquisa deve ser, no mínimo, de 2,9 % do total. Assim sendo, o índice obtido para Bauru é considerado aceitável, pois foram abordados 4,5 % do total dos domicílios.

De acordo com Raia Jr. (2000), a pesquisa catalogou 23.314 deslocamentos, com interesse em dados como modo de transporte utilizado em cada viagem, idade e grau de instrução do viajante, horário e motivo da viagem, localização das origens e destino das viagens, etc. A área urbana foi dividida em 98 zonas de tráfego para esta pesquisa, como mostra a Figura 5.2. A região em destaque, referente ao número 1, representa a zona central de negócios, principal pólo de atração de viagens.



**Figura 5.2:** Divisão da cidade de Bauru em zonas de tráfego para a pesquisa O/D de 1997 (Raia Jr., 2000).

Os dados foram analisados por Raia Jr. (2000) em extensas planilhas eletrônicas, sendo tratados inicialmente de forma individual, isto é, considerando as viagens realizadas por cada pessoa entrevistada e associando a elas a renda efetivamente percebida. Distorções foram encontradas nesta associação e alteradas por Bocanegra (2002), que posteriormente analisou a consistência e a correção dos dados. Dois exemplos de inconsistência são: pessoas com menos de dezoito anos com carteira de habilitação e viagens declaradas por motivo de trabalho por indivíduos com renda igual a zero. Constatadas as inconsistências, foram verificados os dados iniciais e realizadas correções, quando possível. Caso contrário, os dados eram excluídos da planilha.

Foram considerados quatro modos de viagens na pesquisa:

- *Auto\_moto*: viagens realizadas pelo modo automóvel como motorista.
- *Auto\_pass*: viagens realizadas pelo modo automóvel como passageiro.
- *Bike\_pé*: viagens realizadas através de modos não motorizados (bicicleta ou caminhada).
- *Ônibus*: viagens realizadas por ônibus.

Outros erros e inconsistências também foram encontrados e alterados, deixando os dados mais próximos da realidade e mais consistentes.

Trabalhos como o de López-Reyes (1999), Raia Jr. (2000) e Bocanegra (2002) foram realizados com a mesma base de dados, facilitando assim a análise a ser realizada nesta dissertação, pois não foi necessário realizar nenhum tratamento específico nos dados originais para obter as informações necessárias para este estudo.

Para iniciar a pesquisa foi preciso reunir o mapa digitalizado da cidade de Bauru, dados de renda média dos domicílios por setor censitário (1996) e alguns dados da pesquisa O/D. De acordo com Bocanegra (2002), os limites dos setores censitários foram obtidos em um mapa em papel exigindo a digitalização de cada um dos setores. Isto foi realizado diretamente no *software* TransCAD, sendo os setores censitários tratados como áreas.

Como o mapa foi todo digitalizado no *software* TransCAD foram utilizadas ferramentas de sobreposição de camadas para verificar o número de viagens realizadas por setor censitário, já que os dados referentes à pesquisa O/D não estavam originalmente agrupados. A cidade foi então agrupada de acordo com a divisão de setores estipulada para o levantamento censitário de 1991, num total de 306 zonas. Foram calculados os percentuais de viagens dos quatro modos de transporte para cada uma das zonas e adicionado a base de dados digital, que já estava disponível. Excluídas as regiões que não possuíam dados relativos a nenhum tipo de viagem, restaram 272 zonas.

### **5.3. Aplicação das Técnicas e Análise dos Resultados**

As técnicas descritas na metodologia foram aplicadas ao conjunto de dados para extrair informações dos percentuais de viagens, realizados por quatro modos

de transporte. As quais permitiram uma melhor compreensão da dinâmica espacial existente no fenômeno estudado, de forma a identificar padrões de associação espacial, agrupamento de objetos com valores de atributos semelhantes, áreas de transição e situações atípicas. Estas técnicas são aplicadas em conjunto de dados com áreas definidas e atributos associados.

Com os dados digitalizados e a topologia conferida foram calculados os percentuais para os quatro modos de viagens, com o intuito de verificar a existência de regiões com variações extremas nos dados, que podem indicar casos críticos de acessibilidade. Os modos analisados foram:

- % *auto\_moto*: percentagem de viagens realizadas pelo modo automóvel (como motorista).
- % *auto\_pass*: percentagem de viagens realizadas pelo modo automóvel (como passageiro).
- % *bike\_pé*: percentagem de viagens realizadas através de modos não motorizados (bicicleta e caminhada).
- % *ônibus*: percentagem de viagens realizadas por ônibus.

Este conjunto de dados foi, a princípio, analisado através de estatísticas descritivas, para se obter as primeiras idéias sobre o comportamento das variáveis (Tabela 5.1). Para esta identificação vale o princípio da estatística clássica, cujo resumo estatístico fornece informações sobre as medidas de posição, dispersão, entre outras, na qual se estabelece um sistema de classificação que descreve o padrão de variação de um determinado fenômeno estatístico.

A análise conduzida a seguir foi dividida em duas etapas, onde a primeira consiste no cálculo dos elementos básicos da estatística espacial e construção dos mapas, e a segunda etapa identifica, dentro dos agrupamentos espaciais encontrados, quais os valores significantes, procurando avaliar seu impacto no fenômeno estudado.

Na primeira etapa, o principal resultado na fase inicial da análise foi a obtenção dos índices de autocorrelação espacial, da média e do desvio padrão de cada variável (Tabela 5.1). O índice *I* fornece uma medida geral da associação espacial existente nos dados.

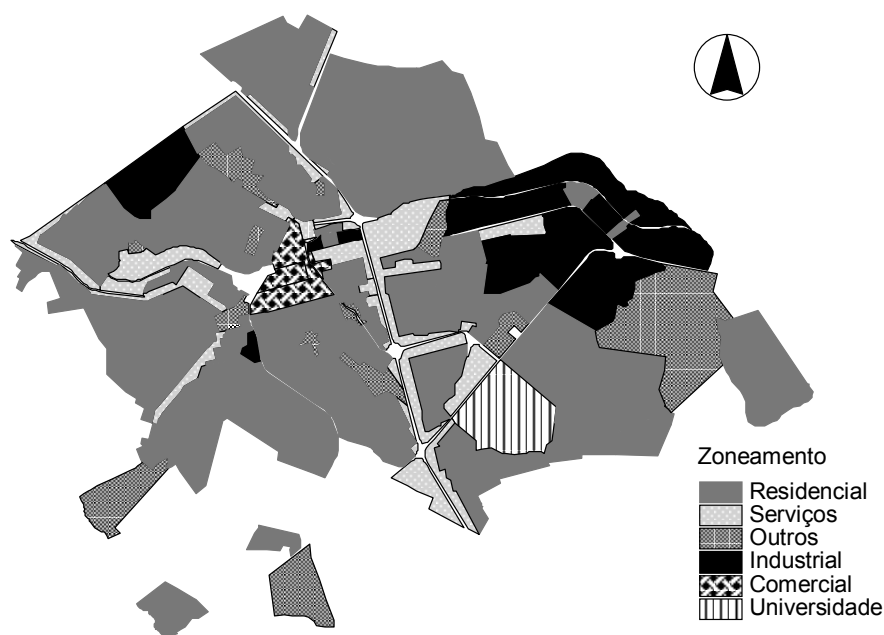
**Tabela 5.1:** Índices dos atributos.

Variáveis	Total de Zonas	Proporção de viagens	Média	Desvio Padrão	<i>I Moran</i>
[Renda]	272	---	1785,04	818,20	0,73
[% <i>auto_moto</i> ]	272	27 %	0,261	0,164	0,39
[% <i>auto_pass</i> ]	272	15 %	0,142	0,108	0,20
[% <i>bike_pe</i> ]	272	30 %	0,304	0,175	0,21
[% <i>onibus</i> ]	272	28 %	0,293	0,189	0,38

Analisando a Tabela 5.1 pode-se observar que todos os índices de Moran são positivos, indicando existir uma autocorrelação espacial positiva, embora com exceção da renda, em nenhum caso muito alta.

Quanto mais baixos os valores do índice de autocorrelação espacial, mais dispersas estão as variáveis, portanto menor o agrupamento e maior dificuldade de analisar este comportamento. O maior índice foi obtido para a variável percentual % *auto\_moto*, que representa as viagens realizadas pelo modo automóvel como motorista, seguido pela variável % *ônibus*, que representa as viagens realizadas pelo modo ônibus. Os menores valores foram encontrados para as variáveis % *bike\_pé*, que representa as viagens realizadas por modos não motorizados, e % *auto\_pass*, que representa o percentual das viagens realizadas pelo modo automóvel como passageiro (“carona”).

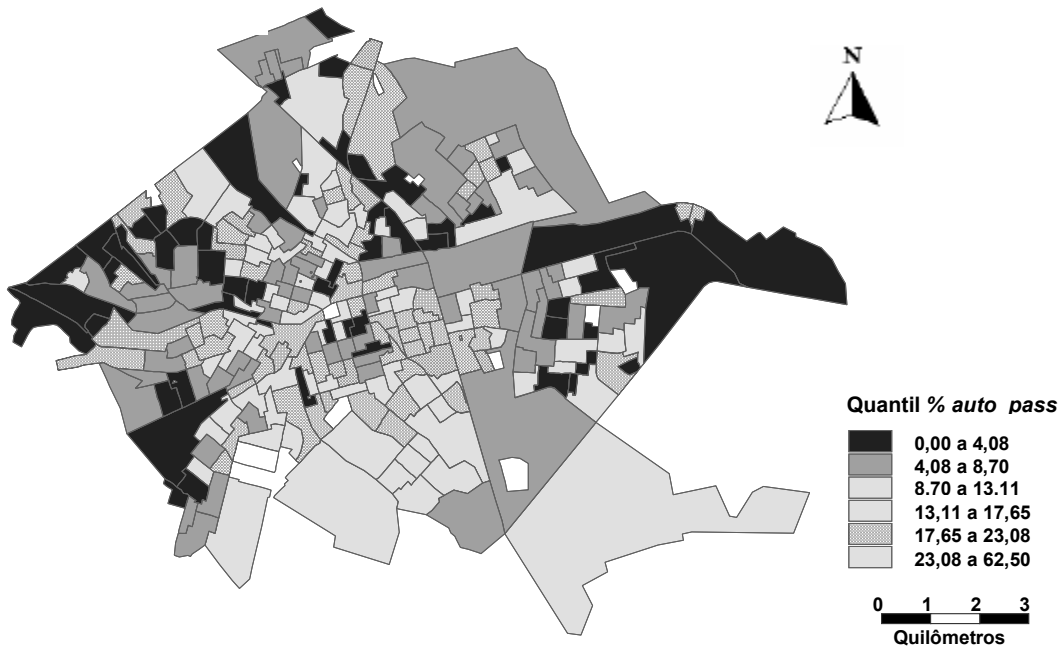
A Figura 5.3 representa a localização dos diferentes usos do solo na cidade de Bauru. Isto irá auxiliar na análise, ao permitir a identificação da localização dos agrupamentos espaciais em relação aos usos do solo existentes.

**Figura 5.3:** Localizações dos diferentes usos do solo na cidade de Bauru.

Para ter uma visão geral da distribuição espacial do uso dos modos de transporte considerados, foram construídos os mapas temáticos classificados por quantil. Estes mapas permitem visualizar os diferentes regimes espaciais existentes em cada modo, como mostram as Figuras 5.4, 5.5, 5.6 e 5.7. A Figura 5.8 refere-se à outra variável considerada no estudo, renda.



**Figura 5.4:** Distribuição dos percentuais de viagens realizadas por automóvel como motorista.



**Figura 5.5:** Distribuição dos percentuais de viagens realizadas por automóvel como passageiro.



**Figura 5.6:** Distribuição dos percentuais de viagens realizadas por ônibus.



**Figura 5.7:** Distribuição dos percentuais de viagens realizadas por modos não motorizados (a pé e bicicleta).



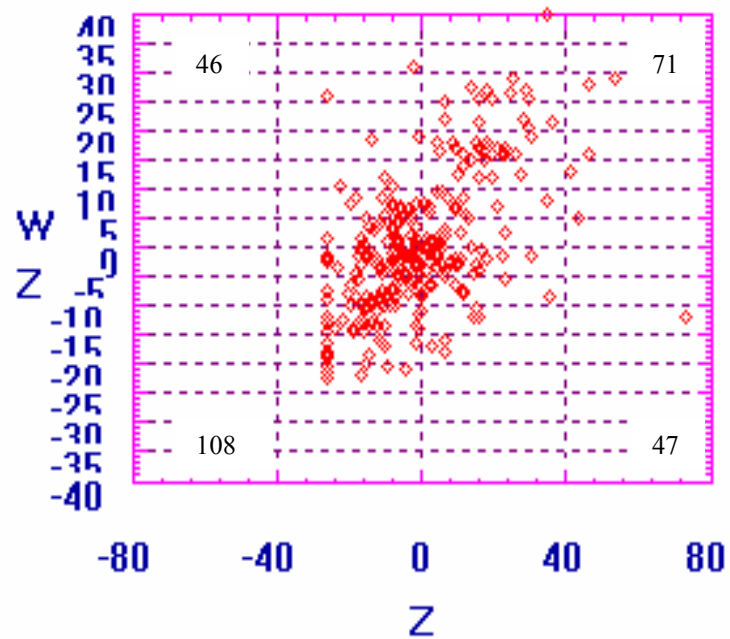


**Figura 5.8:** Distribuição da variável renda.

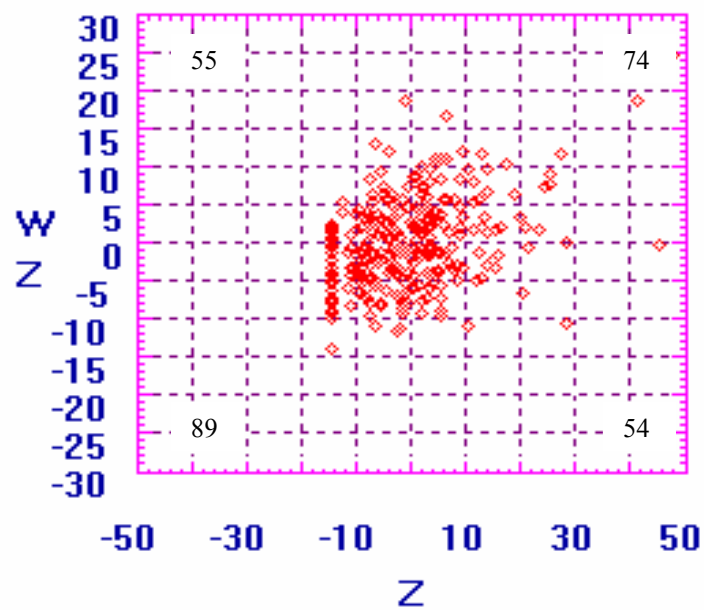
As Figuras 5.4, 5.5, 5.6, 5.7 e 5.8 representam a distribuição das variáveis *% auto\_moto*, *% auto\_pass*, *% ônibus*, *% bike\_pé* e *renda*, classificadas em seis quantis, a fim de visualizar os diferentes regimes espaciais existentes para as mesmas, e como refletem os índices globais de autocorrelação espacial que quantificam o grau da relação existente internamente as variáveis.

Este primeiro resultado mostra os valores extremos, mas é preciso ir muito além, isto é, calcular a média local do percentual de viagens de cada modo para cada zona, e comparar com o valor do atributo (*% de viagens*). Caso a diferença seja pequena, pode-se estar diante de regiões com dinâmica própria, fortemente relacionada, conhecidas como “aglomerados espaciais”. Este resultado é apresentado em duas formas: pelo Diagrama de Espalhamento de Moran e pelo *Box Map*, como descrito no capítulo três.

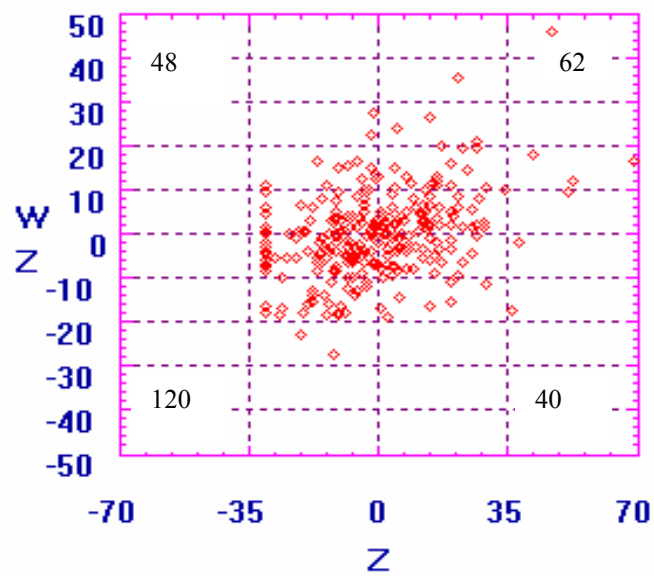
As Figuras 5.9, 5.10, 5.11, 5.12 e 5.13 representam os Diagramas de Espalhamento de Moran para os modos de viagens analisados e para renda.



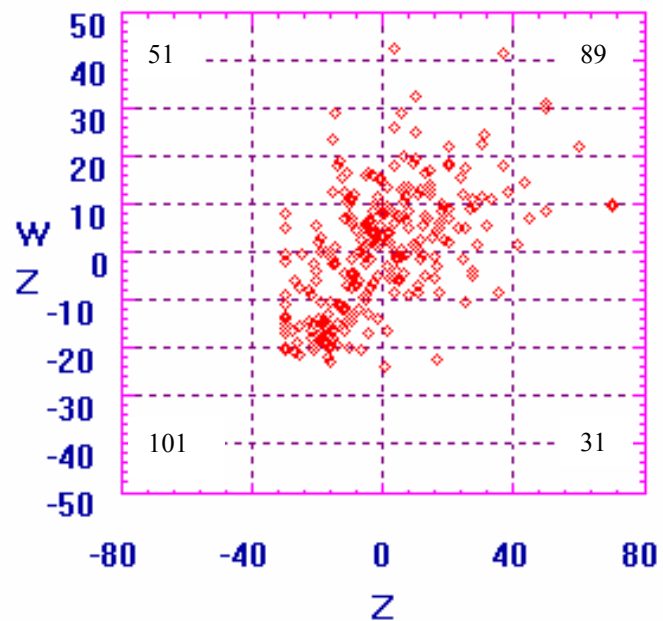
**Figura 5.9:** Diagrama de Espalhamento de Moran para a variável *% auto\_moto*.



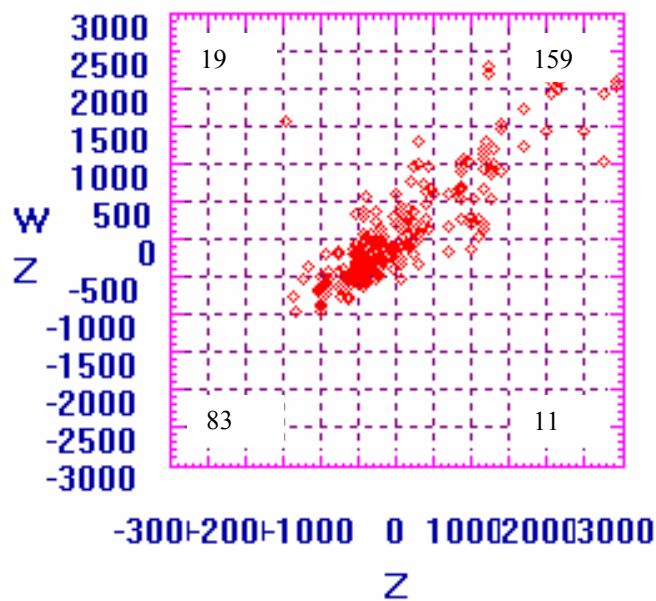
**Figura 5.10:** Diagrama de Espalhamento de Moran para a variável *% auto\_pass*.



**Figura 5.11:** Diagrama de Espalhamento de Moran para a variável *% de modos não motorizados*.



**Figura 5.12:** Diagrama de Espalhamento de Moran para a variável *% ônibus*.



**Figura 5.13:** Diagrama de Espalhamento de Moran para a variável *renda*.

Como pode ser observado para o modo *% auto\_moto* (Figura 5.9), existe uma grande concentração de pontos localizados nos quadrantes Q1 e Q2, indicando a existência de muitas regiões onde o valor do atributo medido assemelha-se aos valores de seus vizinhos, e regiões que podem ser consideradas de transição, representadas nos quadrantes Q3 e Q4.

Para a variável *% auto\_pass* existem mais regiões consideradas de transição e que podem ser analisadas como extremos (cuja localização são os quadrantes Q3 e Q4) do que no caso anterior, ou seja, regiões que não seguem o mesmo processo de dependência espacial das demais observações, do que agrupamentos espaciais localizados em Q1 e Q2. Estes últimos indicam que uma localização possui vizinhos com valores semelhantes. Estas tendências também podem ser observadas no *Box Map* (Figuras 5.14 e 5.15), onde cada polígono está associado ao quadrante a que pertence no gráfico de Espalhamento de Moran.



**Figura 5.14:** *Box Map* da variável *% auto\_moto* e detalhe do mapa de barras (desvio x média dos desvios).

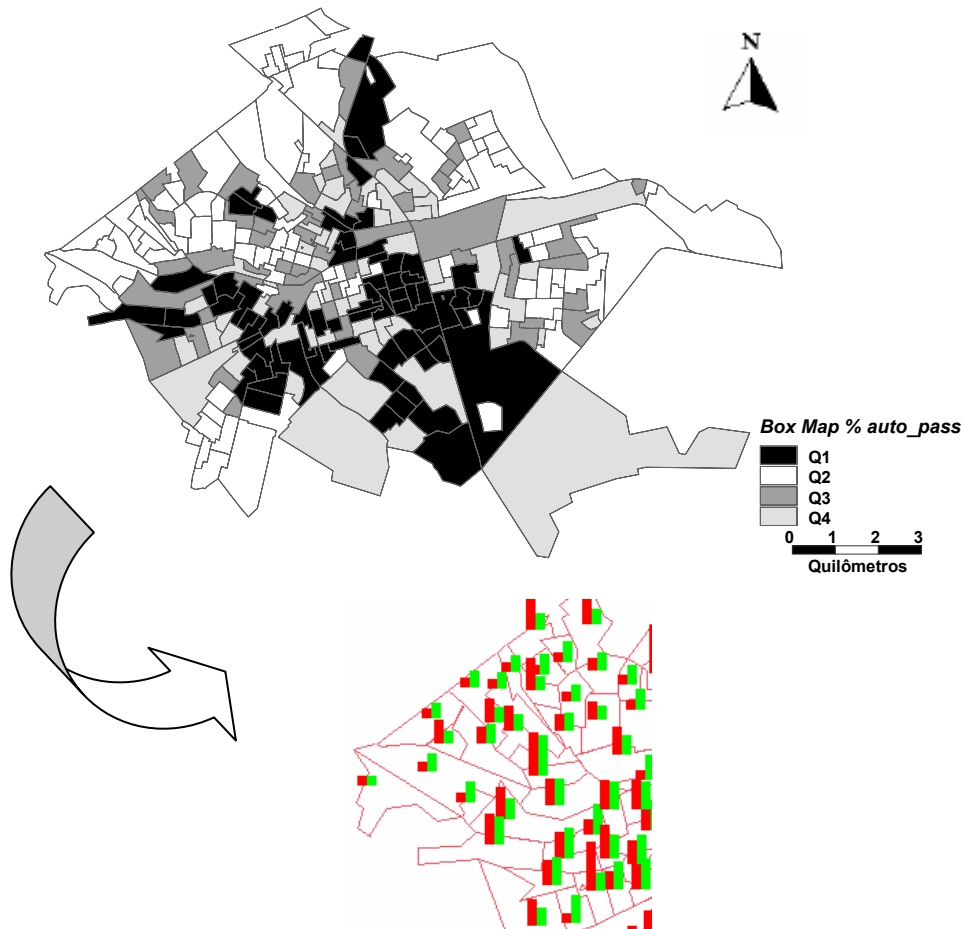
No gráfico de barras da variável *% auto\_moto*, mostrado na Figura 5.14, são apresentadas duas barras sobre cada setor censitário, uma com altura proporcional ao valor do desvio do *% auto-moto*, representada pela barra vermelha (desvio é o valor do atributo menos a média global deste atributo), e outra verde, que representa proporcionalmente o valor médio dos desvios dos percentuais de viagens de automóvel como motorista das zonas vizinhas.

Interpreta-se este gráfico através dos tamanhos das barras, onde barras que aparecem com valores semelhantes indicam agrupamentos de altos (caso do detalhe da Figura 5.14) ou baixos valores e barras com alturas distintas indicam áreas de transição e situações atípicas.

O *Box Map* de cada uma das cinco variáveis estudadas é apresentado nas Figuras 5.14, 5.15, 5.16, 5.17 e 5.18. As zonas da classe 1 possuem os percentuais de viagens acima da média, e percentuais de viagens dos vizinhos com média positiva (quadrante 1 do Diagrama de Espalhamento de Moran). A classe 2 representa o oposto, os percentuais de viagens negativos, e a média dos vizinhos abaixo da média global

(quadrante 2 do Diagrama de Espalhamento de Moran). A classe 3 representa as zonas com valores abaixo da média e o valor médio dos percentuais de viagens dos vizinhos acima da média (quadrante 3 do Diagrama de Espalhamento de Moran). A classe 4 apresenta zonas com valores acima da média e valor médio dos percentuais de viagens menor que a média (quadrante 4 do Diagrama de Espalhamento de Moran).

Como no *Box Map* as representações referentes aos quadrantes Q1 e Q2 indicam uma associação espacial positiva, no sentido que uma localização possui vizinhos com valores semelhantes, a comparação deste gráfico com o da Figura 5.3, permite localizar esta distribuição com relação aos usos do solo na cidade de Bauru. Assim quanto ao modo % *auto\_moto* verifica-se que Q1 e Q2 são áreas predominantemente residenciais, localizadas nas zonas Norte (Q2), Sudeste (Q1) e Oeste (Q2) da cidade. O mesmo procedimento pode ser realizado para o modo % *auto\_pass*, com base no *Box Map* mostrado na Figura 5.15, bem como para os outros dois modos, cujos *Box Map* constam nas Figuras 5.16 e 5.17.



**Figura 5.15:** *Box Map* da variável % *auto\_pass* e mapa de barras (desvio x média dos desvios).



Figura 5.16: Box Map da variável % modos não motorizados.

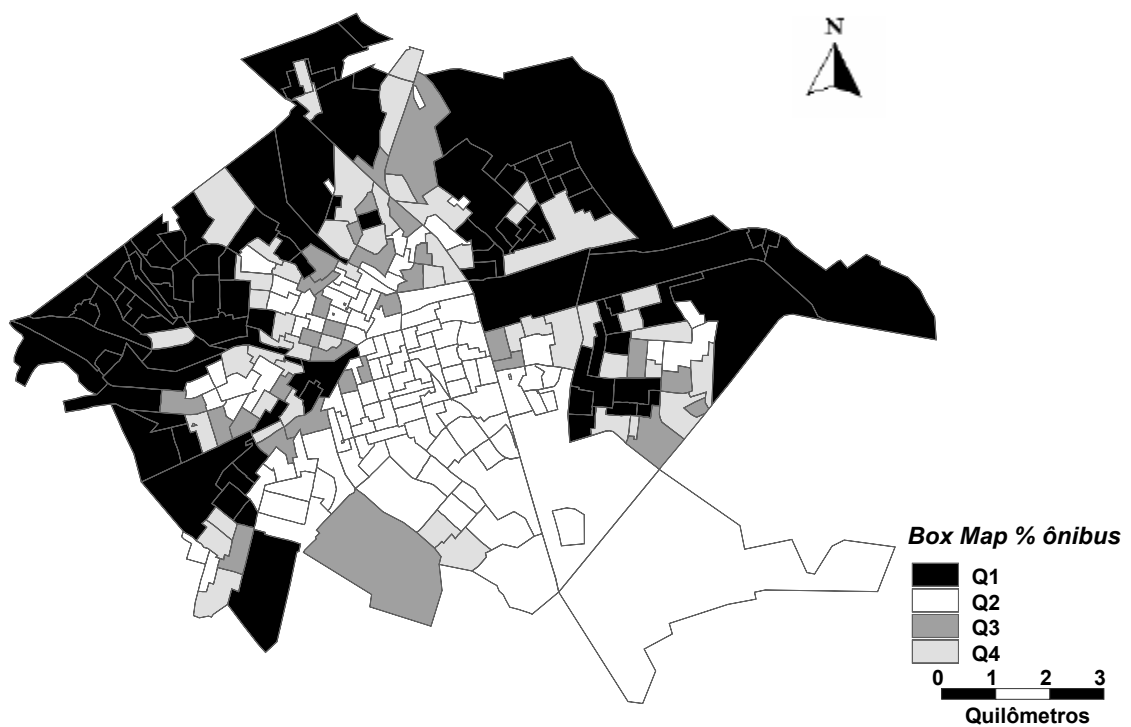
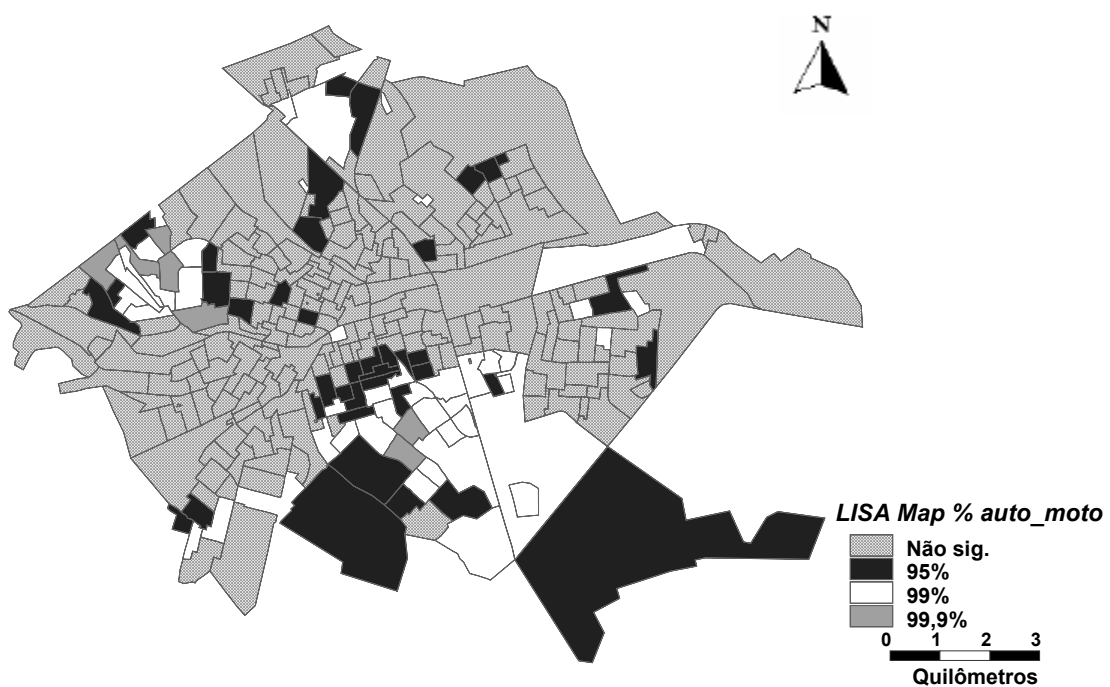


Figura 5.17: Box Map da variável % ônibus.



**Figura 5.18:** *Box Map* da variável *renda*.

A Figura 5.19 apresenta o *LISA Map* da variável *% auto\_moto*, onde os percentuais são classificados em função da significância dos valores de seus índices locais. Esta ferramenta identifica os agrupamentos com valores dos indicadores de associação espacial local significantes, sem contudo diferenciá-los em agrupamentos de altos ou baixos valores.

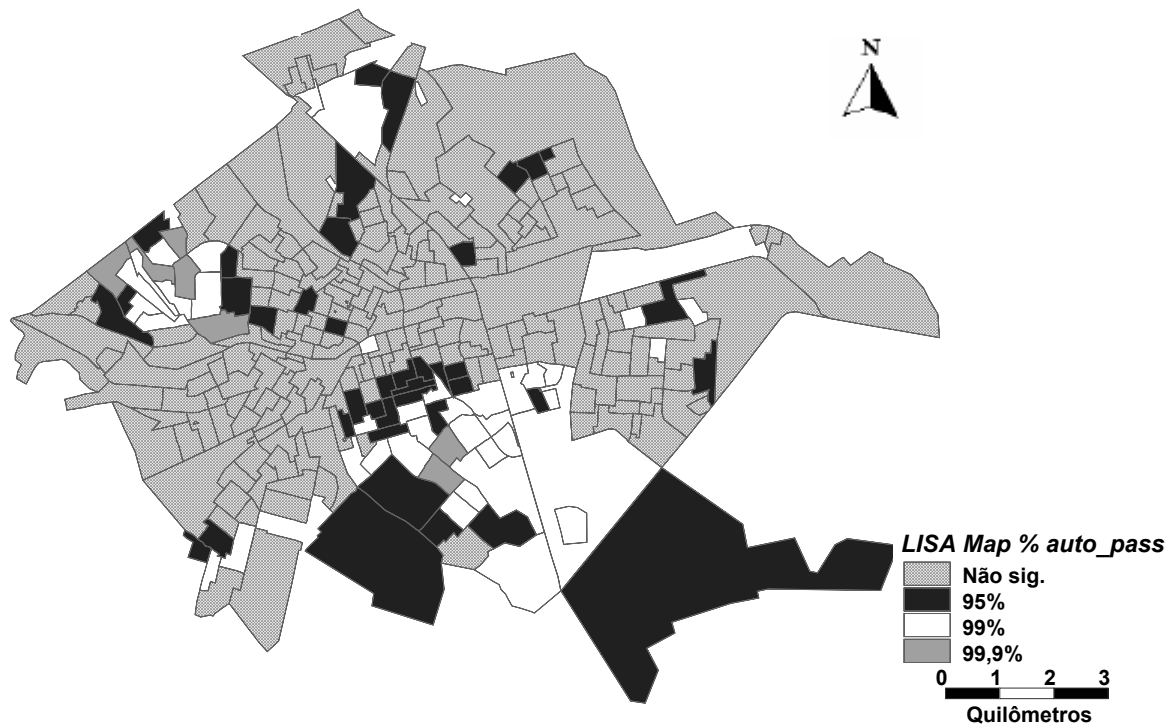


**Figura 5.19:** *LISA Map* da variável *% auto\_moto*.

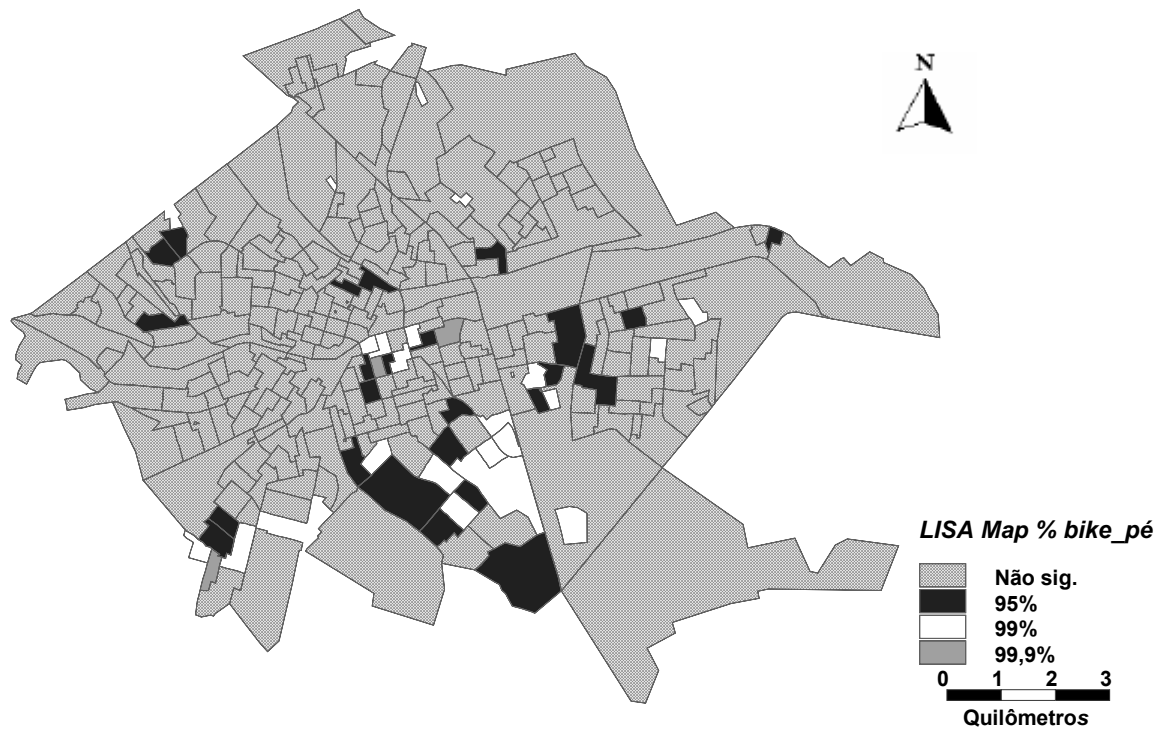


Como pode ser observado na Figura 5.19, as zonas que possuem índice local significativo são representadas no mapa como classe 1 (43 zonas), com 95% de significância, classe 2 (31 zonas), com 99% de significância, e classe 3 (7 zonas), com 99,9% de significância. Encontram-se bem espalhadas pela cidade, estando algumas regiões na zona norte, zona sul e zona oeste. De acordo com a Figura 5.3 estas zonas são de predominância residencial.

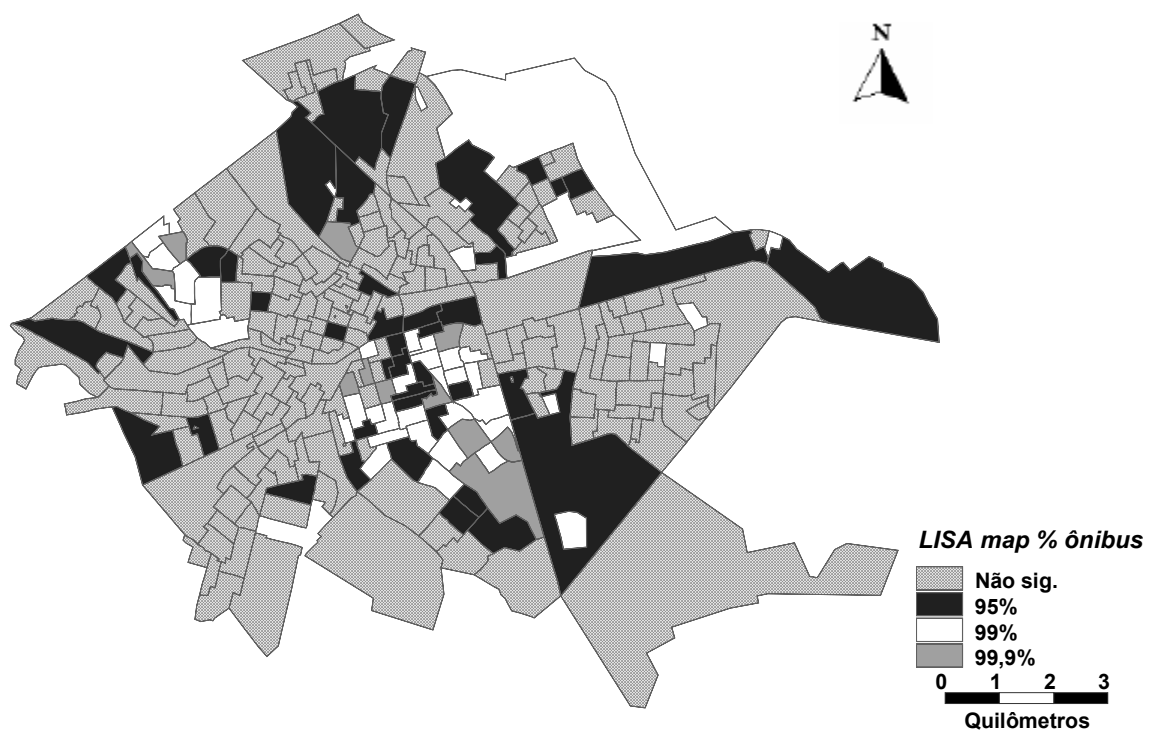
A mesma análise foi realizada para a variável *% auto\_pass*, representada pela Figura 5.20, cujas zonas significativas (46 zonas) são poucas com relação às não significativas (228 zonas do total de 272): 30 zonas possuem índice com significância de 95%, 15 zonas possuem índice com significância de 99% e 1 possui índice com significância de 99,9%. Comparando estas zonas com a Figura 5.3, pode-se observar que estas últimas são áreas pertencentes às regiões norte e oeste, com predominância residencial, mas longe da área universitária e da área comercial. Os *LISA Map* das demais variáveis estudadas aparecem nas Figuras 5.21 a 5.23.



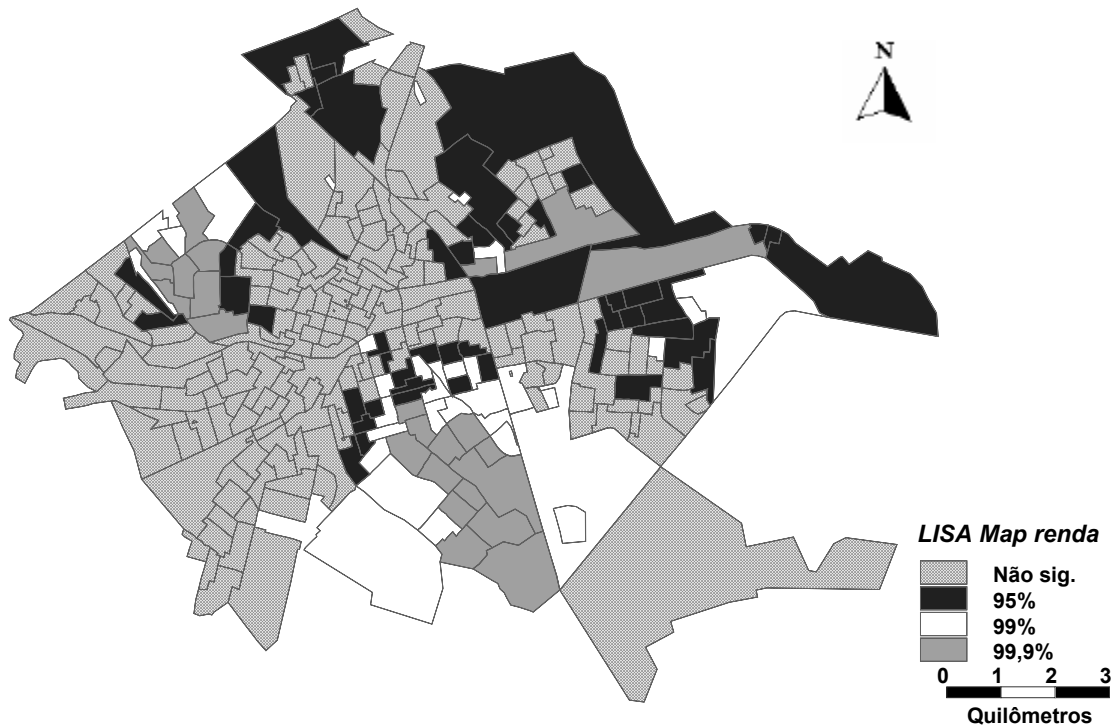
**Figura 5.20:** *LISA Map* da variável *% auto\_pass*.



**Figura 5.21:** *LISA Map da variável % modos não motorizados.*



**Figura 5.22:** *LISA Map da variável % ônibus.*



**Figura 5.23:** *LISA Map* da variável renda.

O gráfico *Moran Map* analisa os valores significantes, isto é, faz uma junção do *Box Map* com *LISA Map*, mostrando apenas as zonas dos respectivos quadrantes que são significantes. Para a variável % *auto\_moto*, por exemplo, pode-se observar, através da Figura 5.24, que do total de 274 zonas analisadas, 194 não são significantes, 39 pertencem a Q1, 32 pertencem a Q2, 5 pertencem a Q3 e 4 pertencem a Q4.

Foram poucas as zonas de transição com significância espacial (Quadrantes 3 e 4). As zonas pertencentes à classe 1 indicam que possuem significância espacial (com 95%, 99% e 99,9%) e estão basicamente localizadas na região sul da cidade, região com predominância de residências. As zonas pertencentes a classe 2 são zonas com relação espacial negativa mas significantes. Estas zonas pertencem às regiões norte e leste e algumas zonas espalhadas pela cidade, como pode ser observado na Figura 5.24.

A mesma análise é feita para o modo *auto\_pass* (Figura 5.25), onde são poucas as zonas de transição com significância espacial (Quadrantes 3 e 4) - apenas 7 pertencentes a Q3 e 4 pertencentes a Q4. São 17 zonas pertencentes a Q1 e que possuem significância espacial (com 95%, 99% e 99,9%), na sua maioria, espalhadas pelo centro da cidade e região sul e estritamente residenciais. As zonas pertencentes à classe 2

(zonas com relação espacial negativa mas significantes) somam 18 casos na região norte, oeste e algumas zonas espalhadas pela cidade, todas regiões consideradas com predominância residencial, como pode ser observado na Figura 5.25. Foram também elaborados *Moran Maps* para as outras variáveis em estudo, como consta nas Figuras 5.26, 5.27 e 5.28.

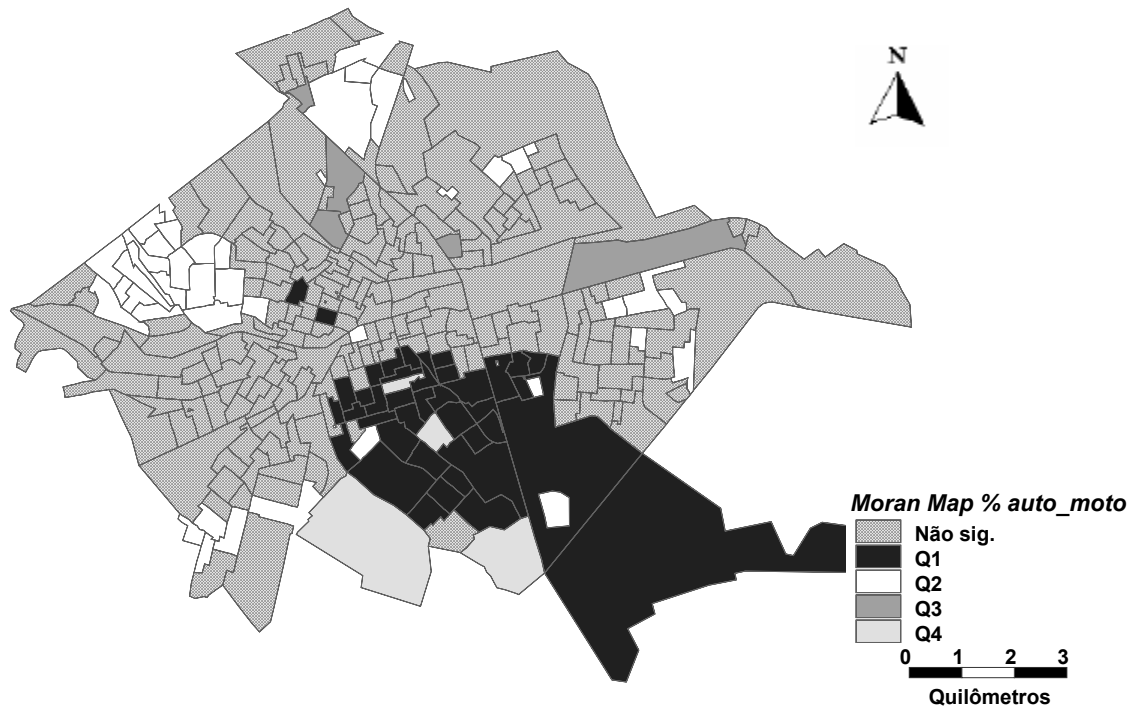


Figura 5.24: *Moran Map* da variável % auto\_moto.



Figura 5.25: *Moran Map* da variável % auto\_pass.

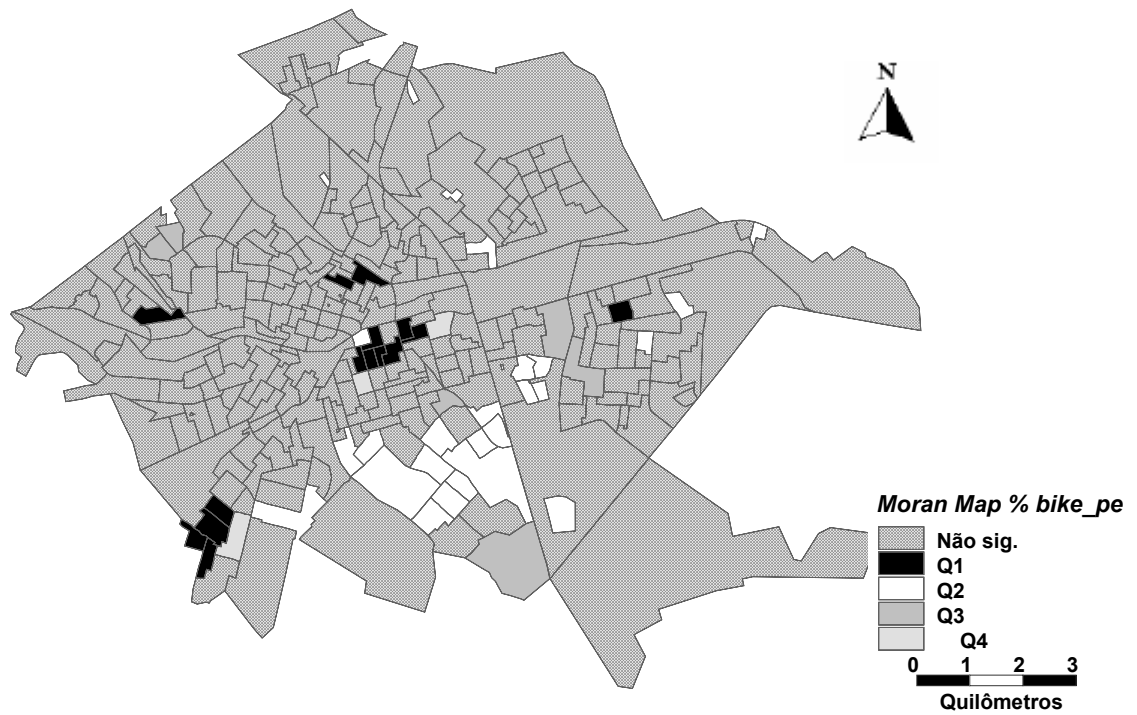


Figura 5.26: Moran Map da variável % modos não motorizados.

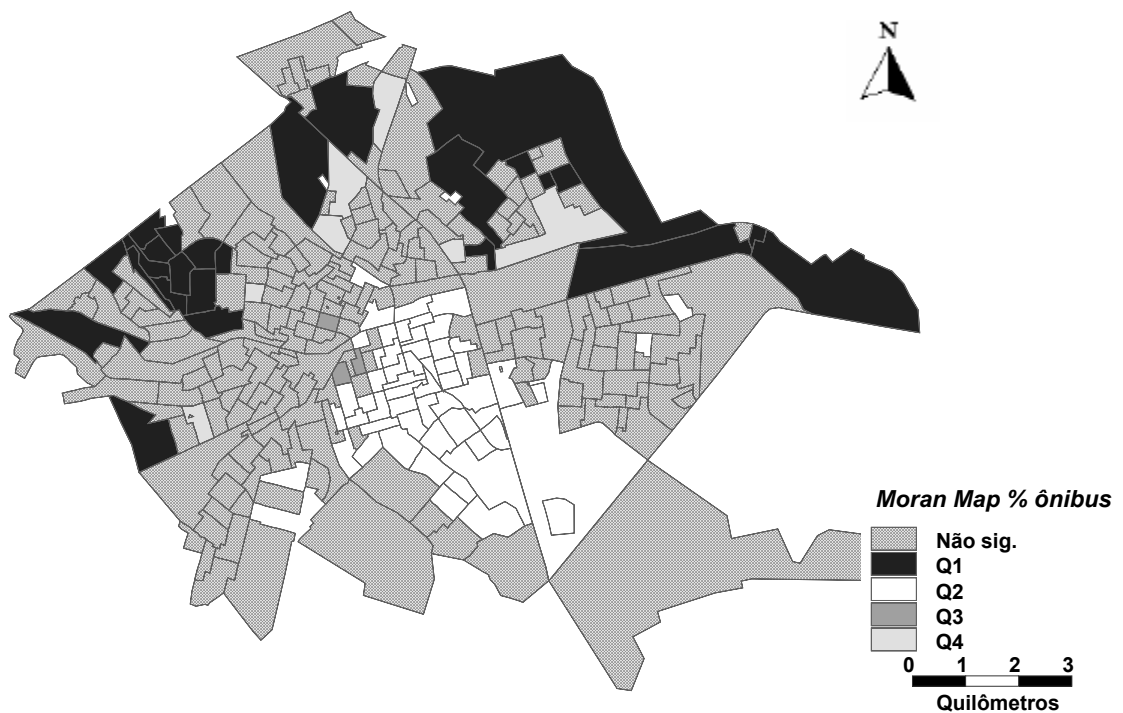
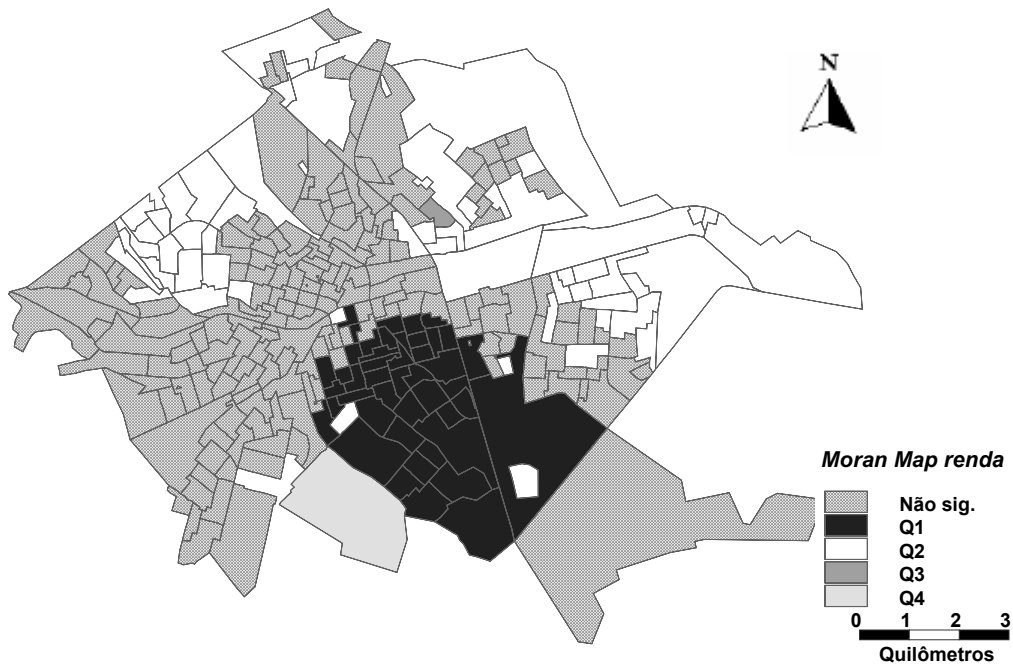
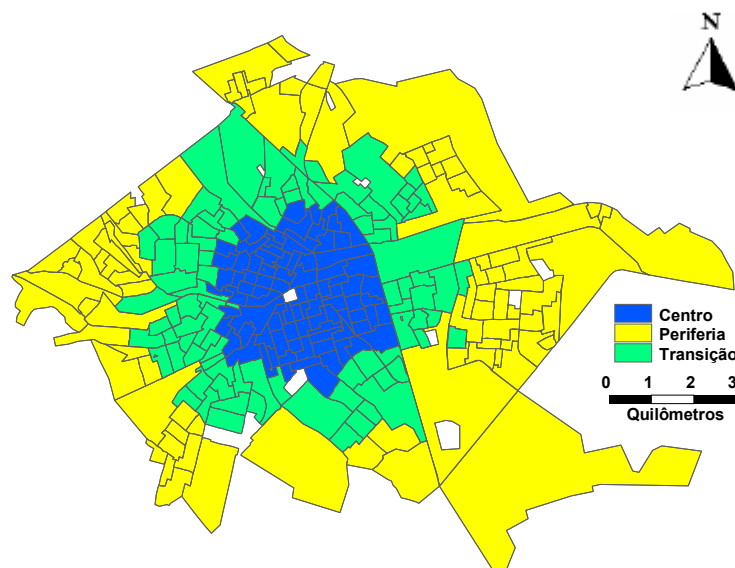


Figura 5.27: Moran Map da variável % ônibus.



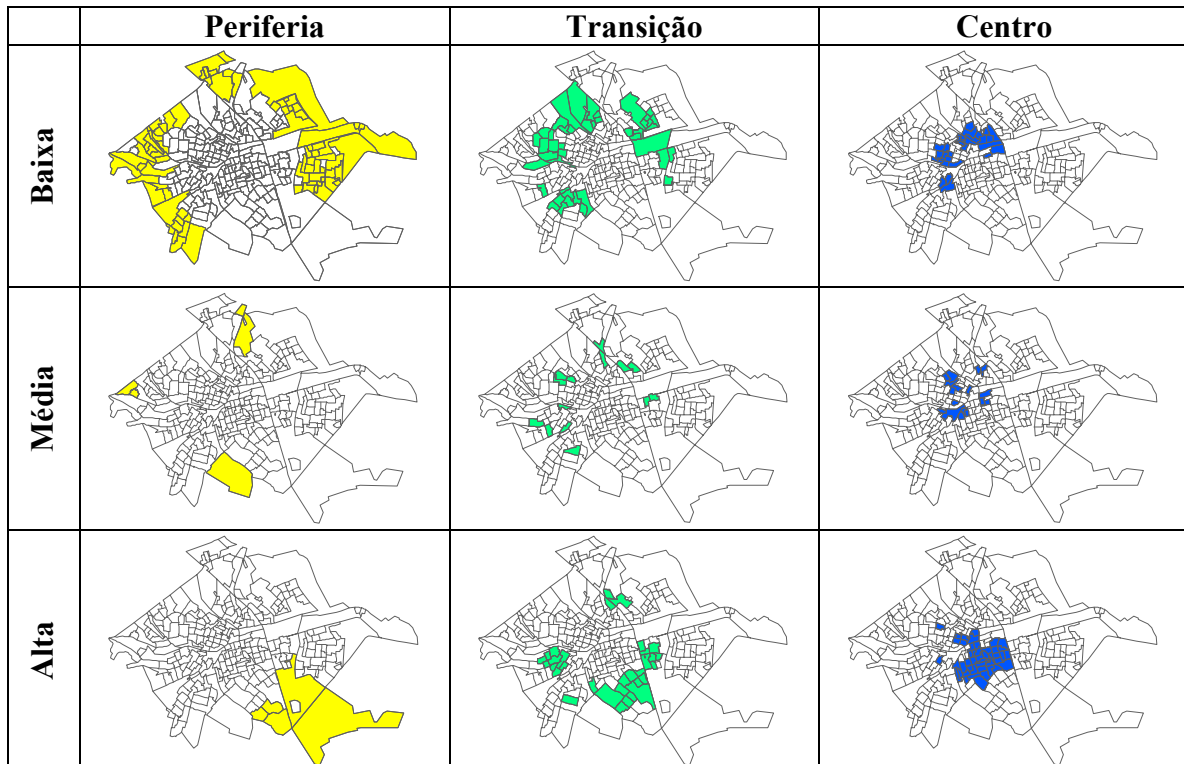
**Figura 5.28:** Moran Map da variável renda.

A segunda parte da análise consiste em uma avaliação detalhada dos gráficos e de tabulações cruzadas realizadas para os modos e para a distribuição da variável renda. Através dos gráficos, pode-se observar que cada modo apresenta uma distribuição específica. Para avaliar como a acessibilidade através dos diferentes modos descritos é afetada pela localização dos usuários pela cidade, as zonas foram agrupadas em 3 anéis: Centro (com 90 zonas), Transição (com 91 zonas) e Periferia (com 91 zonas) (Figura 5.29). Para a variável renda foram também considerados três níveis: baixa, média e alta.



**Figura 5.29:** Divisão da cidade de Bauru.

Para análise das variáveis foram realizadas tabulações cruzadas entre a variável renda e localização (conforme a Figura 5.30), e analisadas por quadrantes, como resumido nas Tabelas 5.2, 5.3 e 5.4. Cada uma das três tabelas representa um anel da cidade, de acordo com a Figura 5.29, onde as linhas indicam os modos de viagens estudados e as colunas os níveis de renda.



**Figura 5.30:** Representação dos nove conjuntos de zonas resultantes da combinação de anéis de localização (periferia, transição e centro) e níveis de renda (alta, média e baixa).

Para o preenchimento da tabela foram analisados os valores Q1, Q2, Q3 e Q4 do Diagrama de Espalhamento de Moran e os *Box Map* para cada modo, considerando a divisão de anéis. Estes valores representam as zonas em que os percentuais de viagens realizados se situam nos respectivos quadrantes. Por exemplo, tomando-se o valor  $Q1 = 11$  da Tabela 5.4, que representa a classificação das zonas do centro, nível de renda médio e viagens realizadas por modos não motorizados, isto significa que, das 17 zonas de renda média e modos não motorizados para o centro, 11 realizam viagens por este modo e se situam no Quadrante 1.

A terceira coluna de cada faixa de níveis de renda indica os valores significativos representados pelo *Moran Map*. Deste modo pode-se destacar as zonas com significativo índice de autocorrelação espacial e, ao mesmo tempo, saber se trata de um agrupamento de alto ou baixo valor.

**Tabela 5.2:** Classificação das zonas da periferia de acordo com os modos de viagem e níveis de renda.

	NÍVEIS DE RENDA								
	BAIXA			MÉDIA			ALTA		
<b>VIAGENS NÃO MOTORIZADAS</b>	Q1	30	6	Q1	2	0	Q1	0	0
	Q2	18	1	Q2	1	0	Q2	4	0
	Q3	14	3	Q3	0	0	Q3	1	1
	Q4	19	1	Q4	2	0	Q4	0	0
	Não Sign.	--	70	Não Sign.	--	5	Não Sign.	--	4
	Total	81	81	Total	5	5	Total	5	5
<b>ÔNIBUS</b>	Q1	55	17	Q1	2	0	Q1	0	0
	Q2	3	0	Q2	0	0	Q2	4	0
	Q3	5	0	Q3	2	0	Q3	1	1
	Q4	18	2	Q4	1	0	Q4	0	0
	Não Sign.	--	62	Não Sign.	--	5	Não Sign.	--	4
	Total	81	81	Total	5	5	Total	5	5
<b>AUTOMÓVEL COMO MOTORISTA</b>	Q1	0	0	Q1	0	0	Q1	3	3
	Q2	57	25	Q2	4	0	Q2	0	0
	Q3	16	2	Q3	0	0	Q3	1	0
	Q4	8	0	Q4	1	1	Q4	1	1
	Não Sign.	--	54	Não Sign.	--	4	Não Sign.	--	1
	Total	81	81	Total	5	5	Total	5	5
<b>AUTOMÓVEL COMO PASSAGEIRO</b>	Q1	3	0	Q1	2	0	Q1	3	0
	Q2	53	14	Q2	2	2	Q2	0	0
	Q3	17	1	Q3	0	0	Q3	0	0
	Q4	8	0	Q4	1	0	Q4	2	0
	Não Sign.	--	66	Não Sign.	--	3	Não Sign.	--	5
	Total	81	81	Total	5	5	Total	5	5



**Tabela 5.3:** Classificação das zonas de transição de acordo com os modos de viagem e níveis de renda.

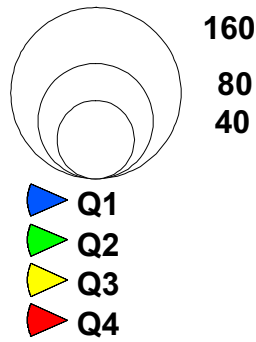
	NÍVEIS DE RENDA								
	BAIXA			MÉDIA			ALTA		
	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3
<b>VIAGENS NÃO MOTORIZADAS</b>	Q1	5	0	Q1	2	0	Q1	1	0
	Q2	18	0	Q2	5	0	Q2	24	14
	Q3	13	1	Q3	6	0	Q3	3	0
	Q4	9	1	Q4	1	0	Q4	4	0
	Não Sign.	--	43	Não Sign.	--	14	Não Sign.	--	18
	Total	45	45	Total	14	14	Total	32	32
<b>ÔNIBUS</b>	Q1	26	8	Q1	2	0	Q1	0	0
	Q2	5	0	Q2	3	1	Q2	23	11
	Q3	3	0	Q3	3	0	Q3	4	0
	Q4	11	5	Q4	6	0	Q4	5	0
	Não Sign.	--	32	Não Sign.	--	13	Não Sign.	--	21
	Total	45	45	Total	14	14	Total	32	32
<b>AUTOMÓVEL COMO MOTORISTA</b>	Q1	3	0	Q1	2	0	Q1	23	15
	Q2	23	7	Q2	5	0	Q2	1	0
	Q3	12	3	Q3	3	0	Q3	3	0
	Q4	7	0	Q4	4	0	Q4	5	1
	Não Sign.	--	35	Não Sign.	--	14	Não Sign.	--	16
	Total	45	45	Total	14	14	Total	32	32
<b>AUTOMÓVEL COMO PASSAGEIRO</b>	Q1	8	3	Q1	5	1	Q1	23	6
	Q2	17	3	Q2	1	0	Q2	0	0
	Q3	12	3	Q3	4	0	Q3	2	0
	Q4	8	1	Q4	4	0	Q4	7	1
	Não Sign.	--	35	Não Sign.	--	13	Não Sign.	--	25
	Total	45	45	Total	14	14	Total	32	32

**Tabela 5.4:** Classificação das zonas do centro de acordo com os modos de viagem e níveis de renda.

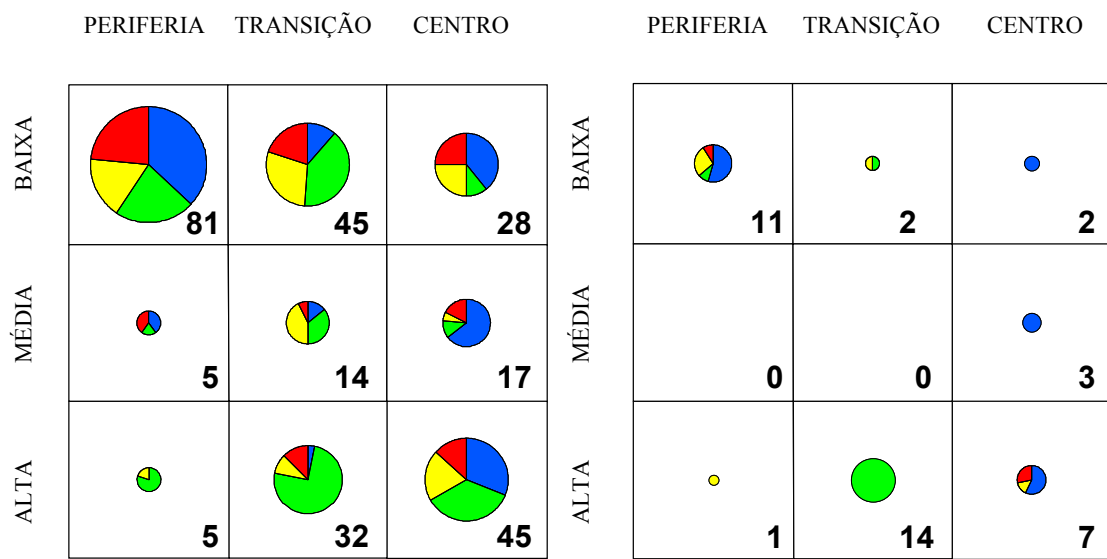
	NÍVEIS DE RENDA								
	BAIXA			MÉDIA			ALTA		
<b>VIAGENS NÃO MOTORIZADAS</b>	Q1	11	2	Q1	11	3	Q1	14	4
	Q2	3	0	Q2	2	0	Q2	16	0
	Q3	7	0	Q3	1	0	Q3	9	1
	Q4	7	0	Q4	3	0	Q4	6	2
	Não Sign.	--	26	Não Sign.	--	14	Não Sign.	--	38
	Total	28	28	Total	17	17	Total	45	45
<b>ÔNIBUS</b>	Q1	1	0	Q1	2	0	Q1	0	0
	Q2	13	3	Q2	8	0	Q2	42	30
	Q3	6	0	Q3	5	3	Q3	3	1
	Q4	8	0	Q4	2	2	Q4	0	0
	Não Sign.	--	25	Não Sign.	--	12	Não Sign.	--	14
	Total	28	28	Total	17	17	Total	45	45
<b>AUTOMÓVEL COMO MOTORISTA</b>	Q1	7	1	Q1	2	0	Q1	31	21
	Q2	10	0	Q2	6	0	Q2	2	0
	Q3	5	0	Q3	3	0	Q3	4	0
	Q4	6	0	Q4	6	0	Q4	8	1
	Não Sign.	--	27	Não Sign.	--	17	Não Sign.	--	23
	Total	28	28	Total	17	17	Total	45	45
<b>AUTOMÓVEL COMO PASSAGEIRO</b>	Q1	6	0	Q1	1	0	Q1	23	8
	Q2	6	1	Q2	6	0	Q2	4	0
	Q3	8	1	Q3	4	0	Q3	7	1
	Q4	8	1	Q4	6	0	Q4	11	1
	Não Sign.	--	25	Não Sign.	--	17	Não Sign.	--	35
	Total	28	28	Total	17	17	Total	45	45

Dado a sua importância no conjunto dos resultados, as células em amarelo foram selecionadas para análises mais detalhadas. Esta seleção se deu através de uma análise preliminar dos gráficos de setores (Figuras 5.32 a 5.35).

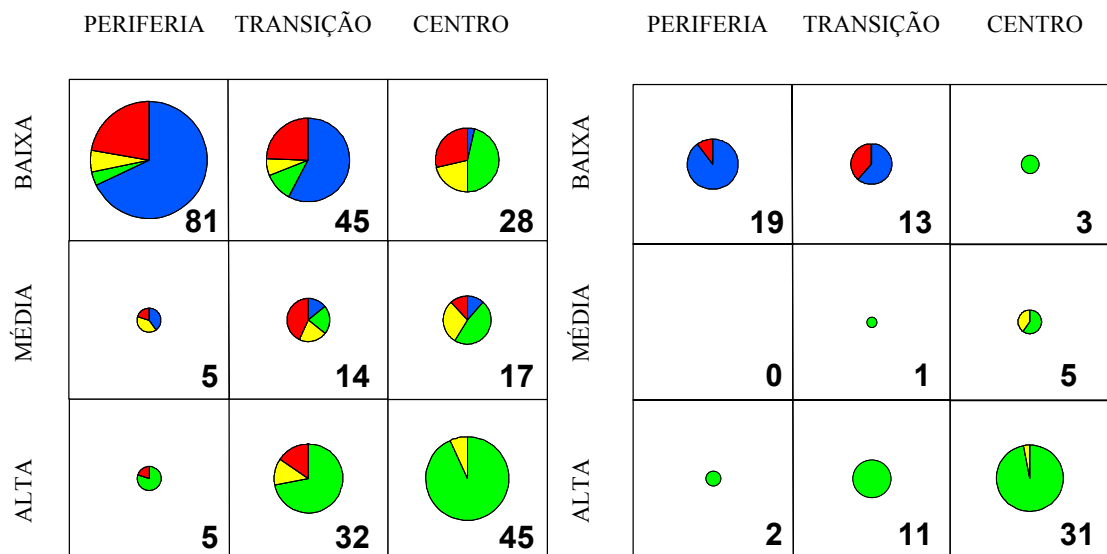
Os gráficos de setores representam os diferentes modos de viagem, onde cada quadrante é representado por uma cor específica (ver legenda na Figura 5.31). Os tamanhos dos gráficos são proporcionais ao número de viagens. Do lado esquerdo das Figuras 5.32 a 5.35 os quadrantes se referem aos *Box Map*, e do lado direito se referem aos *LISA Map* (quadrantes com valores significativos). Em síntese, Q1 implica em grande proporção de viagens realizadas pelo modo, Q2 implica em pequena proporção de viagens realizadas pelo modo, e Q3 e Q4 são valores de transição.



**Figura 5.31:** Legenda para os gráficos de setores das Figuras 5.32 a 5.36.



**Figura 5.32:** Gráficos de setores da variável % *modos não motorizados*.



**Figura 5.33:** Gráficos de setores da variável % *ônibus*.

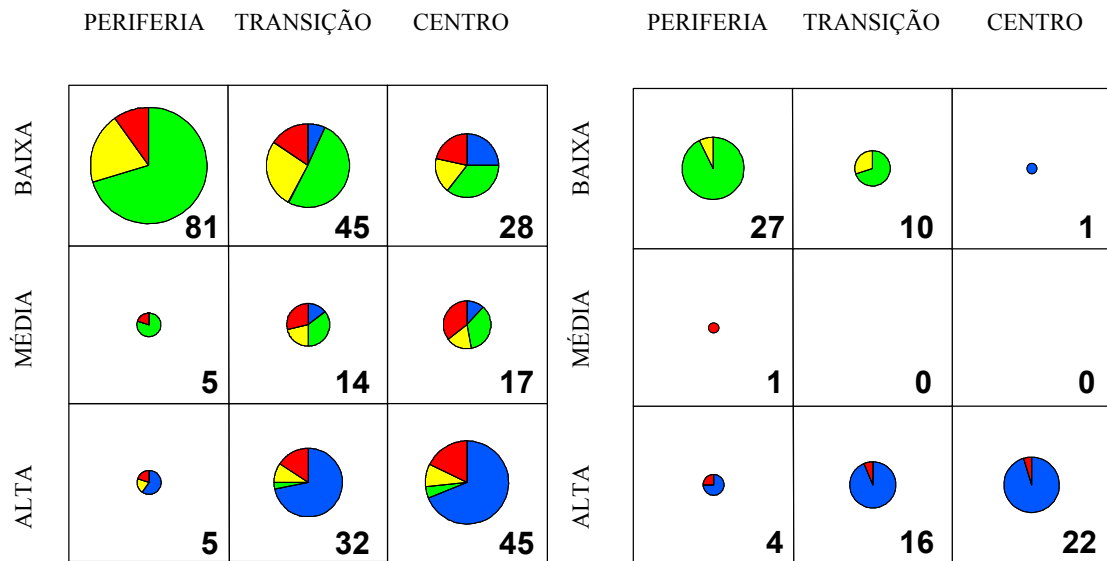


Figura 5.34: Gráficos de setores da variável % *auto\_moto*.

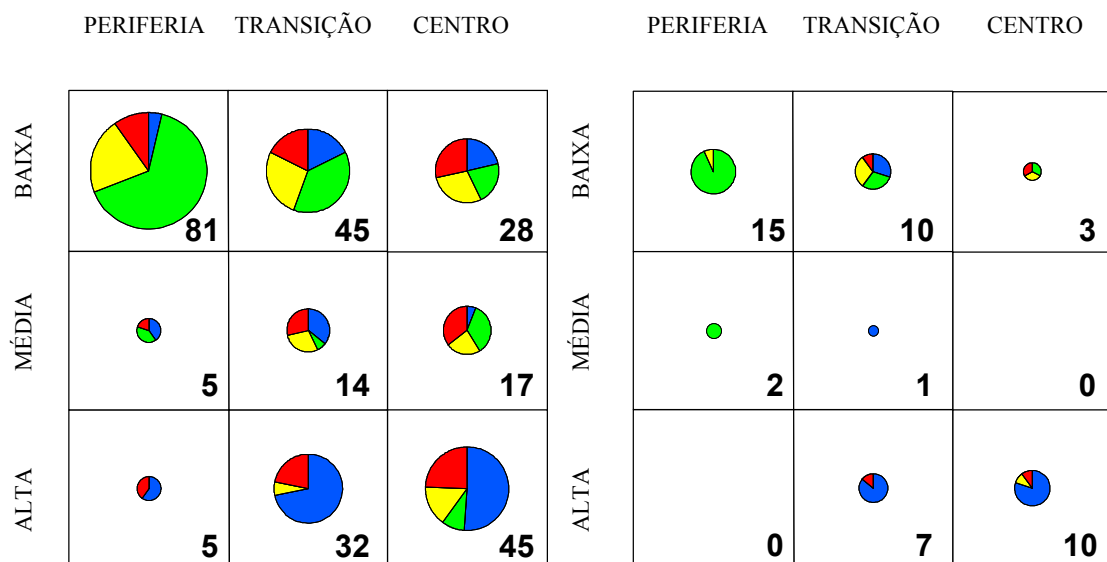


Figura 5.35: Gráficos de setores da variável % *auto\_pass*.

Para as zonas da periferia (Tabela 5.2) são analisadas as zonas de nível de renda baixa que utilizam o modo ônibus, e as zonas de nível de renda baixa que utilizam automóvel como motorista.

Como se pode observar na Tabela 5.2, a primeira célula em destaque indica que das 81 zonas de renda baixa na periferia, em 55 são realizadas muitas viagens pelo modo ônibus, podendo ser classificadas no quadrante Q1. Nestas zonas, tanto os valores dos desvios por setor quanto a média dos desvios são positivos, indicando que o percentual de viagens por ônibus é maior que a média da mesma variável para toda a cidade. Analisando a significância deste quadrante, 17 zonas possuem significativo

índice local de autocorrelação, o que é um agrupamento expressivo. As zonas de Q2 não possuem significância, e das zonas Q3 e Q4, que indicam zonas de transição de padrão espacial, somente 2 apresentam valores significantes. Analisando o gráfico de setores da variável % *ônibus* (Figura 5.33), da esquerda para a direita e no nível de renda baixo, observa-se que o uso deste modo é predominante na periferia, diminuindo progressivamente da periferia para o centro. Isto é indicado nas Tabelas 5.2, 5.3 e 5.4 pelos valores associados ao quadrante Q1, que nos gráficos têm a cor azul.

A outra célula em destaque na Tabela 5.2 representa nível de renda baixa e modo de viagem automóvel como motorista. O valor do Q2 representa que, das 81 zonas, em 57 são realizadas poucas viagens pelo modo automóvel como motorista, podendo ser classificadas no quadrante Q2. Analisando a significância deste quadrante, 25 zonas possuem significativo índice local de autocorrelação. Analisando o gráfico de setores da variável % *auto\_moto* (Figura 5.34), da esquerda para a direita e no nível de renda baixo, observa-se que o uso deste modo é quase inexistente na periferia, aumentando progressivamente da periferia para o centro. Analisando a mesma Figura 5.34, mas no nível de renda alta, as viagens realizadas na periferia são predominantemente feitas pelo modo % *auto\_moto*, o que também ocorre nas zonas de transição e centro.

Para as zonas de transição (Tabela 5.3) são analisadas as zonas de nível de renda alta que utilizam os modos não motorizados, e as zonas de nível de renda alta que utilizam automóvel como motorista.

Como se pode observar na Tabela 5.3, a primeira célula em destaque indica que das 32 zonas de renda alta na região de transição, em 24 são realizadas poucas viagens pelos modos não motorizados, podendo ser classificadas no quadrante Q2. Nestas zonas, tanto os valores dos desvios por setor quanto a média dos desvios são negativos, indicando que o percentual de viagens por modos não motorizados é menor que a média da mesma variável para toda a cidade. Analisando a significância deste quadrante, 14 zonas possuem significativo índice local de autocorrelação, o que é um agrupamento expressivo. As zonas de Q1, Q3 e Q4 não apresentam valores significantes. Analisando o gráfico de setores da variável % *modos não motorizados* (Figura 5.32), da esquerda para a direita e no nível de renda alta, observa-se que o uso deste modo cresce progressivamente da periferia para o centro. Isto é indicado nas Tabelas 5.2, 5.3 e 5.4 pelos valores associados ao quadrante Q1, que nos gráficos têm a cor azul.

A outra célula em destaque na Tabela 5.3 representa nível de renda alta e modo de viagem automóvel como motorista. O valor do Q1 representa que, das 32 zonas, em 23 são realizadas muitas viagens pelo modo automóvel como motorista, podendo ser classificadas no quadrante Q1. Analisando a significância deste quadrante, 15 zonas possuem significativo índice local de autocorrelação. A análise do gráfico de setores da variável % *auto\_moto* (Figura 5.34) não mais se justifica neste caso, uma vez que foi feita no caso anterior.

Para as zonas centrais (Tabela 5.4) são analisadas as zonas de nível de renda alta que utilizam ônibus, e as zonas de nível de renda alta que utilizam automóvel como motorista.

Como se pode observar na Tabela 5.4, a primeira célula em destaque indica que das 45 zonas de renda alta no centro, em 42 são realizadas poucas viagens pelo modo ônibus, podendo ser classificadas no quadrante Q2. Nestas zonas, tanto os valores dos desvios por setor quanto a média dos desvios são negativos, indicando que o percentual de viagens por ônibus é menor que a média da mesma variável para toda a cidade. Analisando a significância deste quadrante, 30 zonas possuem significativo índice local de autocorrelação, o que é um agrupamento expressivo. As zonas de Q1 e Q4 não possuem significância e, das zonas Q3, somente uma é significativa. A análise do gráfico de setores da variável % *ônibus* (Figura 5.33) não mais se justifica neste caso, uma vez que foi feita no caso anterior.

A outra célula em destaque representa nível de renda alta e modo de viagem automóvel como motorista. O valor do Q1 representa que, das 45 zonas, 31 utilizam o automóvel como principal meio de transporte e 21 destas são significantes. Portanto, desta análise pode-se inferir que na periferia o modo ônibus é muito utilizado por indivíduos de renda baixa (cujas viagens provavelmente devem ser motivadas por trabalho e estudo. O modo automóvel é utilizado pelas zonas com predominância de renda alta e independente do motivo, em qualquer ponto da cidade. Já para a renda baixa, na periferia o modo automóvel não possui significância.

A outra célula em destaque na Tabela 5.4 representa nível de renda alta e modo de viagem automóvel como motorista. O valor do Q1 representa que, das 45 zonas, em 31 são realizadas muitas viagens pelo modo automóvel como motorista, podendo ser classificadas no quadrante Q1. Analisando a significância deste quadrante, 21 zonas possuem significativo índice local de autocorrelação. A análise do gráfico de setores da variável % *auto\_moto* (Figura 5.34) já foi feita anteriormente.

Na região de transição e para renda média, o modo mais utilizado é o ônibus, o que pode ser explicado devido ao fato das pessoas estarem longe do trabalho ou escola e não contarem com o automóvel para se locomover.

Na região central as pessoas de renda média utilizam os modos não motorizados para se locomoverem. Este fato pode ser explicado por estarem mais próximas aos pólos de atividade. Isto não afeta, no entanto, as zonas de renda alta, que embora estejam no centro têm, como modo mais utilizado, o automóvel.

## 6. CONCLUSÕES E RECOMENDAÇÕES

---

O presente trabalho procurou explorar a união de ferramentas de estatística espacial e SIG com um propósito específico, que foi o de analisar as relações entre aspectos da distribuição espacial de características da população de uma cidade média brasileira e os diversos níveis de acessibilidade por diferentes modos de transporte nela observados. Isto foi feito com as seguintes preocupações em mente: utilizar recursos disponíveis e de baixo custo; aplicar técnicas de estatística espacial de forma didática e explorar os recursos gráficos existentes em estatística espacial, de modo que os procedimentos aqui utilizados possam ser aplicados em outros trabalhos.

Foram analisados dados de quatro modos de transporte com ferramentas de estatística espacial em um ambiente SIG neste estudo de caso, visando ampliar a compreensão de influências locais na acessibilidade em transporte. A princípio foi calculado o índice de autocorrelação global de Moran para todos os modos. Os resultados obtidos foram: 0,388 para o percentual de viagens realizado pelo modo automóvel como motorista (*% auto\_moto*), 0,199 para o percentual de viagens realizado pelo modo automóvel como passageiro (*% auto\_pass*), 0,209 para o percentual de viagens realizado por modos não motorizados (*% bike\_pe*), 0,375 para o percentual de viagens realizado pelo modo ônibus (*% ônibus*) e 0,730 para a variável *renda*. Os valores encontrados para os modos de viagens foram todos de baixa autocorrelação espacial positiva, e alta autocorrelação espacial positiva para a renda.

Valores de baixa autocorrelação significam que o valor do atributo medido para algumas zonas não é semelhante aos encontrados para os vizinhos imediatos. Porém, esta estatística não faz uma distinção das zonas que têm esta condição particular, uma vez que se refere a um indicador global para a cidade toda.



Foram aplicadas outras técnicas para a identificação de agrupamentos espaciais, que avaliassem os agrupamentos quanto ao nível de significância, o que permite assumir a hipótese de que tende a haver semelhança entre os valores dos atributos das zonas fisicamente mais próximas. Cabe destacar que estas técnicas foram também aplicadas utilizando exclusivamente as rotinas do programa, mantendo a proposta inicial de baixo custo.

As análises gráficas permitem algumas conclusões particulares sobre o uso dos diferentes modos de transporte nas zonas da cidade, como pode ser visto nos gráficos de espalhamento de Moran, que permitem identificar as regiões com significância.

A conclusão deste trabalho é que as viagens consideradas significativas originadas nas zonas periféricas, mostradas no *Moran Map* e analisadas através da Tabela 5.2, são viagens feitas por modos motorizados mesmo no nível de renda baixo, embora predominantemente no modo ônibus. No nível de renda médio são poucos setores com esta situação, portanto nenhum agrupamento espacial possui significância estatística. Logo, cada zona possui uma característica, mesclando o uso dos modos, principalmente ônibus, modos não motorizados e “carona”. Para as zonas de renda alta da periferia o modo predominantemente utilizado é o automóvel como motorista, em regiões preponderantemente residenciais. Estas aglomerações podem ser vistas nas Figuras 5.24, 5.25, 5.26, e 5.27, e de alguma maneira complementam umas às outras, pois são inclusive fisicamente próximas.

Analisando a zona de transição para o nível de renda baixo pode-se perceber que os modos ônibus e automóvel como passageiro são muito utilizados em zonas de renda média, embora o modo automóvel como passageiro seja o único que possui significância neste caso. As zonas de renda alta possuem uma característica particular, onde modos não motorizados e ônibus não são utilizados, sendo apenas utilizado o automóvel como motorista e como passageiro.

Analisando as zonas da região central com nível de renda baixo pode-se observar (Tabela 5.4) que os modos não motorizados são os mais utilizados, provavelmente devido ao fato de estarem em região central e próxima aos pólos de atividades. Com relação ao nível de renda médio, os modos não motorizados também são os mais utilizados. Já as zonas de renda alta utilizam o automóvel como motorista, como passageiro (“carona”) e modos não motorizados. No caso da região central, uma

característica particular é que o modo automóvel como passageiro é mais utilizado do que o modo ônibus.

Estas características podem ser vistas nos gráficos de setores, Figuras 5.32 a 5.35, nas quais se pode perceber que o modo automóvel como motorista possui agrupamento espacial bem definido no nível de renda alta nas regiões de periferia, transição e central. Isto sugere que nestas regiões este modo é predominantemente utilizado em zonas de renda alta. O modo automóvel como passageiro (“carona”) possui uma dinâmica bem diferente da variável automóvel como motorista, pois são as regiões de transição e renda baixa, e as regiões de renda alta e transição do centro que utilizam este modo.

O modo ônibus é predominantemente utilizado pelas regiões de renda baixa da periferia e transição. Já os modos não motorizados possuem uma dinâmica bem diversificada, sendo utilizados nas regiões de renda baixa da periferia e centro, regiões de renda média do centro e regiões de renda alta do centro.

Um dos resultados interessantes da aplicação é a identificação de regiões com dinâmica particular, que vão contra o padrão global observado na área urbana. Esta diferença é bem explícita no índice global de Moran, pois para o modo *% auto\_moto*, onde o valor do índice é 0,38, a associação é mais definida, enquanto que no modo *% auto\_pass*, onde o valor é 0,19, esta associação é bem menor.

Em outras palavras, parece que os residentes da periferia da cidade têm um nível de acessibilidade razoável, aparentemente porque podem dispor dos modos motorizados (no caso, o ônibus).

A aplicação que foi realizada parece sugerir que a combinação dessas ferramentas com SIG é realmente frutífera. Este trabalho também mostrou que, embora existam no mercado programas comerciais com ferramentas de estatística espacial a eles incorporados, isto não é realmente necessário para aplicações simples, pois se o pacote de SIG dispuser de ferramentas de operação com matrizes e se o usuário tiver algum conhecimento de programação básica, os mesmos resultados podem ser obtidos sem muitas dificuldades.

Uma recomendação para trabalhos posteriores seria analisar a base de dados verificando todas as viagens, a fim de visualizar novas divisões da cidade e outras variáveis socioeconômicas para futuras análises, identificando possíveis parâmetros que possam afetar os motivos das viagens, e analisando o relacionamento destas considerando também a sua componente de distribuição espacial.

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

---

ANSELIN, L. (1996) The Moran scatterplot as an ESDA tool to assess local instability in spatial association. *In: Fisher, M., Scholten, H. J. e Unwin, D. (eds.). Spatial Analytical Perspectives on GIS.* p. 111-126, Taylor & Francis, Londres.

ANSELIN, L.; BAO, S. (1997) Exploratory spatial data analysis linking SpaceStat and ArcView. *In: Fisher, M. e Getis, A (eds.), Recent Developments in Spatial Analysis.* Springer-Verlag, Berlim.

ANSELIN, L. ; O. SMIRNOV, O. (1998). The SpaceStat extension for ArcView 3.0. Regional Research Institute, West Virginia University, Morgantown.

ASSUNÇÃO, R. M. (2001) Estatística Espacial com aplicações em Epidemiologia, Economia, Sociologia. *7ª Escola de Modelos de Regressão*, São Carlos, SP.

BOCANEGRA, C. W. R. (2002) *Procedimentos para tornar mais efetivo o uso das redes neurais artificiais em planejamento de transporte.* São Carlos Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

BRONDINO, N. C. M.; SILVA, A. N. R. (1999) Combining artificial neural networks and GIS for land valuation purposes. *Anais do 6<sup>th</sup> International Conference on Computers in Urban Planning and Urban Management*, Veneza, Itália, em CD-ROM.

BULLEN, N. (1997) Linking GIS and multilevel modelling in the spatial analysis of house prices. *In: Massimo Craglia e Helen Couclelis (eds.) Geographical Information Research: Bridging the Atlantic.* p. 412-426, Taylor & Francis, Londres.

CALIPER (1996) *TransCAD User's Guide.* Newton, Massachusetts, Caliper Corporation.

CÂMARA, G.; CORREA, V.; PAIVA, J. A.; MONTEIRO, A. M. V.; CARVALHO, M. S.; FREITAS, C. C.; RAMOS, F. R.; NEVES, M. C. (2001) *Geoprocessamento: Teorias e aplicações.* <http://www.ptr.Usp.br/Cursos/AnaliseEspacial/Livro/cap13-estatespacial.pdf>.

CRESSIE, N. (1993) *Statistics for spatial data.* New York, Wiley.

DAVIDSON, K. B. (1995) *Accessibility and isolation in transportation network evaluation*. Artigo apresentado na 7<sup>th</sup> World Conference on Transport Research. The University of New South Wales, Sydney, Australia.

ESRI (1996) ArcView GIS, The Geographic Information System for everyone, Using ArcView GIS. Environmental Systems Research Institute, Redlands-CA.

GEARY, R. (1954) The contiguity ratio and statistical mapping. *The Incorporated Statistician*. v. 5, p. 115-145.

GOTO, M. (2000) *Uma análise da acessibilidade sob a ótica da equidade – o caso da região metropolitana de Bauru*. São Carlos Dissertação (Mestrado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

GOTO, M.; SILVA, A. N. R.; MENDES, J. F. G. (2001) Uma análise de acessibilidade sob a ótica da equidade - O caso da Região Metropolitana de Belém, Brasil. *Engenharia Civil/Civil Engineering*, Departamento de Engenharia Civil da Universidade do Minho, Guimarães - Portugal, n. 10, p. 55-66.

IBGE (2000) Censo Demográfico 2000 - Malha Municipal Digital do Brasil 1997.

KREMPI, A. P.; BRONDINO, N. C. M.; SILVA, A. N. R. Evaluating transportation accessibility with spatial statistics tools in a GIS environment. In: International Conference on Design and Decision Support Systems in Architecture and Urban Planning, 6, Ellecom - Holanda, 2002. Proceedings... Eindhoven, Eindhoven University of Technology. p. 132-144.

LANDIM, P. M. B. (1998) *Análise Estatística de Dados Geográficos* – Editora Unesp – São Paulo, Brasil.

LEVINE, N. (1996) Spatial statistics and GIS – software tools to quantify spatial patterns. *Journal of the American Planning Association*. v. 62, n. 3, p. 381-391.

LÓPEZ-REYES, D. E. (1999) *Um procedimento de determinação de matriz origem-destino para diferentes modos: método indireto baseado em um modelo de escolha discreta*. São Carlos Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

MORAN, P. (1948) The interpretation of statistical maps. *Journal of the Royal Statistic Society B*. v. 10, p. 243-251.

NEVES, M. C; RAMOS, F. R; CAMARGO, E. C. G; CÂMARA, G; MONTEIRO, A. M (2001). *Análise exploratória espacial de dados sócio-econômicos de São Paulo*, Campinas (SP): Embrapa Meio Ambiente e Instituto Nacional de Pesquisa Espacial - INPE.

PAEZ, A (2003) Spatial statistics for urban analysis: a review of new techniques with examples. *Anais do 8<sup>th</sup> International Conference on Computers in Urban Planning and Urban Management*, Sendai, Japão, em CD-ROM.

QUEIROZ, M. P. (2003) *Análise espacial dos acidentes de trânsito do município de Fortaleza*. Fortaleza Dissertação (Mestrado) – Universidade Federal do Ceará.

RAIA JR., A. A. (2000) *Acessibilidade e mobilidade na estimativa de um índice de potencial de viagens utilizando redes neurais artificiais*. São Carlos. Tese (Doutorado) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

RAMOS, R. A. R.; SILVA, A. N. R. (2003) A data-driven approach for the definition of metropolitan regions. *Anais do 8<sup>th</sup> International Conference on Computers in Urban Planning and Urban Management*, Sendai, Japão, em CD-ROM.

SERRANO, R. M.; VALCARCE, E. V. (2000) *Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial*. Barcelona, Edicions Universita de Barcelona.

SILVA, A. N. R. (1998). *Sistemas de Informações Geográficas para Planejamento de Transporte*. Texto Sistematizado elaborado para concurso de Livre-Docência, Departamento de Transporte, Escola de Engenharia de São Carlos, Universidade de São Paulo. 112p.

SILVA, A. N. R.; LIMA, R. S.; RAIA JR, A. A.; VAN DER WAERDEN, P. (1998) Urban transportation accessibility and social inequity in a developing country. In: Freeman, P. e C. Jamet (eds.) *Urban transport policy - A sustainable development tool*. Rotterdam, Balkema. p.709-714.

SILVA, A. N. R.; ROSE, A.; RODRIGUES, D. S. (2003) An assessment method for GIS-T software. *Anais do 8<sup>th</sup> International Conference on Computers in Urban Planning and Urban Management*, Sendai, Japão, em CD-ROM.

STUDENT. (1914) *The elimination of spurious correlation due to position in time or space*. *Biometrika* v. 10, p. 179-180.

TEIXEIRA, G. L. (2003) *Uso de dados censitários para identificação de zonas homogêneas para planejamento de transportes utilizando técnicas de estatística espacial*. Brasília Dissertação (Mestrado) – Faculdade de Tecnologia da Universidade de Brasília.