

RESUMO

Cerca de 90% da variabilidade da densidade de m² de loja de supermercados nas microrregiões do estado de São Paulo pode ser explicada pela densidade de renda ali existente, desde que a autocorrelação espacial presente nas duas variáveis seja levada em consideração. Neste trabalho, apresenta-se o conceito de autocorrelação espacial, sugere-se uma medida para a sua intensidade (Índice de Geary), e estuda-se uma estratégia para sua remoção.

PALAVRAS-CHAVE

Estatística Espacial; Autocorrelação Espacial; Índice de Geary; Matrizes de Conectividade; Vizinhança; Supermercados.

ABSTRACT

Approximately 90% of the variability in the density of supermarket floor space at the micro-regions of São Paulo state can be explained by the income density in the micro-regions, when spatial autocorrelation present in both variables is taken into account. In this report, the concept of spatial autocorrelation is presented, a measurement for its intensity is suggested (Geary's Ratio) and a strategy for its removal is studied.

KEY WORDS

Spatial Statistics; Spatial Autocorrelation; Geary's Ratio; Connectivity Matrices; Neighborhood; Supermarkets.

SUMÁRIO

I.	Introdução	4
1.	Hipótese de trabalho	4
2.	Formulação do problema	5
3.	Justificativa	5
4.	Objetivos	6
II.	Autocorrelação espacial.....	7
1.	Definição de autocorrelação espacial.....	12
1.1.	Correlação	12
1.2.	Autocorrelação	13
1.3.	Autocorrelação espacial.....	13
2.	Uma medida de autocorrelação espacial.....	15
2.1.	Variabilidade de vizinhança.....	17
2.2.	Variabilidade do conjunto total de dados	19
2.3.	Comparando as variabilidades	19
2.4.	Matrizes de conectividade.....	20
2.4.a.	Microrregiões	21
2.4.b.	Faixas de distância.....	22
2.4.c.	Critérios de vizinhança	23
2.4.d.	Matrizes de conectividade	24

2.5. GR da distribuição de metragem de loja	30
3. Significância.....	33
4. Remoção da autocorrelação espacial.....	35
4.1. Variável dependente: densidade de loja na microrregião.....	36
4.2. Variável independente: densidade de renda por microrregião	39
III. Modelo final	41
1. Formulação.....	41
2. Ajuste	45
3. Interpretação.....	47
4. Comparação com o modelo anterior.....	48
IV. Conclusão.....	48
V. Bibliografia	49
VI. Apêndice	50
1. Matrizes de conectividade	50
2. Programa em MatLab para cálculo de GR.....	59
3. Microrregiões do Estado de São Paulo.....	60
4. Valores em m ² de loja e densidade de renda.....	77
5. Programa para simulação da distribuição de GR.....	79
6. Simulação de GRS para distribuição.....	80
7. Resíduos dos modelos de autocorrelação	83

AUTOCORRELAÇÃO ESPACIAL NA DISTRIBUIÇÃO DE ÁREA DE LOJA DE SUPERMERCADOS EM MUNICÍPIOS DO ESTADO DE SÃO PAULO

*Francisco Aranha**

I. INTRODUÇÃO

Em extensivo levantamento de dados patrocinado pelo NPP e concluído em Dezembro de 97¹, amostramos 1.796 estabelecimentos supermercadistas, distribuídos por 288 municípios paulistas. Estes supermercados somaram uma área de loja de 3.288.766 m².

Era nossa intenção elaborar um modelo estatístico capaz de explicar e prever a área de loja de supermercados em cada município. Se, de um lado, pudemos formular um modelo bastante aceitável para a identificação de relações significativas entre as variáveis estudadas, de outro lado temos de reconhecer que sua capacidade explicativa, de pouco mais de 60% da variabilidade total dos dados, é baixa quando se pretende realizar previsões.

1. HIPÓTESE DE TRABALHO

Um das causas identificadas para o desempenho insuficiente do modelo é a presença, nos dados, de autocorrelação espacial. Nossa hipótese de trabalho é de que

* Agradecimentos aos alunos que participaram da pesquisa que originou o presente relatório como auxiliar de pesquisas, Vera Lúcia de Góes, e como monitor, Carlos Augusto Beetz de Souza.

¹ ARANHA, FRANCISCO. **Potencial de Consumo dos Municípios Paulistas**, São Paulo: EAESP/FGVSP/NPP, Relatório de Pesquisa em Editoração.

os fatores determinantes da alta densidade de loja em um município acabam influenciando também os municípios vizinhos. Assim, estes “vazamentos” do efeito das variáveis causais, “transbordando” de um município para o outro, reduzem as diferenças da metragem de loja que seriam observadas entre os municípios, caso o “vazamento” não existisse. A manifestação aparente do fenômeno oculta parcialmente a relevância e capacidade explicativa das variáveis utilizadas.

2. FORMULAÇÃO DO PROBLEMA

Na validade de nossa hipótese de trabalho, uma melhor compreensão dos fatores que influenciam a área de loja de supermercados existente nos municípios depende da confirmação da presença da autocorrelação espacial, sua quantificação e eventual remoção, se possível. É necessário, portanto, pesquisar em fontes especializadas as técnicas adequadas ao trato do problema que temos em mãos e testar sua eficácia no caso particular sob consideração.

Em resumo, o que se pretende neste projeto teórico-prático, é identificar e absorver uma metodologia de análise de dados espacializados, verificando a viabilidade de sua aplicação para a modelagem da área de loja de supermercados nos municípios.

3. JUSTIFICATIVA

O tratamento da autocorrelação é necessário ao aperfeiçoamento da metodologia dos índices de potencial de consumo proposta no trabalho anterior². Estes índices são importantíssimos para a elaboração dos planos de investimento e operação de empresas, principalmente na área de marketing. Além disso, o potencial para

² ARANHA, FRANCISCO. Relatório citado.

supermercados é informação importante em diversos tipos de estudos de mercado e de localização de atividades varejistas.

De outra perspectiva, a previsão do potencial municipal para supermercados é relevante em si, isto é, para os supermercadistas, que poderão avaliar os mercados com “espaço” para abertura de novas lojas e outros com presença excessiva de estabelecimentos.

Finalmente, e talvez mais importante, uma abordagem simples da autocorrelação espacial tem amplo potencial de aplicação empresarial quando os fenômenos de interesse tem um componente espacial proeminente, em áreas tão diversas como logística e administração de risco. Apesar disso, as publicações na área são escassas (principalmente em português) e, em consequência, a identificação, teste e difusão da metodologia é de grande utilidade. Ademais, como pudemos constatar em nossa pesquisa anterior, a maioria das publicações relativas a modelagem de potencial existentes no Brasil acaba sugerindo métodos ingênuos, com graves erros conceituais e estatísticos, que, em nossa opinião, invalidam sua utilização prática já que levam a problemas de especificação e a viés nas estimativas dos parâmetros.

4. OBJETIVOS

Os objetivos deste trabalho são:

4.1. Apresentar o conceito de autocorrelação espacial, em linguagem simples, procurando fazer a ligação entre os textos de estatística básica e textos mais avançados;

4.2. Preparar as bases de dados de potencial dos municípios do estado de São Paulo e realizar os cálculos necessários; estes são extensivos, e envolvem simulações de distribuições de estatísticas³;

4.3. Concluir sobre a existência ou não de autocorrelação espacial na área de supermercados nos municípios;

4.4. Quantificar a intensidade da autocorrelação;

4.5. Se possível, removê-la, para obter dados independentes;

4.6. Recalcular os modelos estudados no projeto anterior.

II. AUTOCORRELAÇÃO ESPACIAL

A grande maioria dos fenômenos não se distribui no espaço casualmente. O relevo de uma região pode ser resultado de falhas geológicas; a presença de uma espécie vegetal pode depender de certas combinações de clima e solo; a concentração de população num país pode refletir o roteiro histórico de sua colonização; a industrialização em determinados estados pode refletir a existência de recursos naturais. Relevo, espécies vegetais, população e indústrias apresentam-se, portanto, organizadas de maneira que pode ser prevista, pelo menos em parte, uma vez conhecidos os seus determinantes.

Assim, quando nos debruçamos sobre um tema e nos interessamos por suas características geográficas, procuramos identificar padrões espaciais e estabelecer

³ Preparo, teste e verificação dos dados consumiram parte significativa do trabalho envolvido neste projeto de pesquisa. Para a realização dos cálculos foram necessários o estudo de álgebra linear e a seleção e aprendizado de software adequado: os livros utilizados nesta etapa foram SEARLE, Shayle R. *Matrix Algebra Useful for Statistics*. New York: Wiley, 1982; SEARLE, Shayle R. *Linear Models*, N.York: Wiley, 1997; e HANSELMAN, Duane and LITTLEFIELD, Bruce. *MatLab: The Language of Technical Computing*. Upper Saddle River: Prentice Hall, 1997.

relações causais com outros fenômenos também distribuídos no espaço. Para isso, geralmente coletamos informação na forma de amostras.

Tomemos o objeto do nosso interesse, a quantidade de supermercados existentes nos municípios do estado de São Paulo, medida em m^2 de loja⁴. Imaginemos que o total de área de loja em cada município tivesse sido determinado única e exclusivamente **por acaso**, através de um sorteio: em uma urna haveria tantas bolas quanto são os municípios, cada uma registrando um valor de m^2 de loja; a metragem construída em cada cidade teria sido definida pela extração de uma bola da urna.

Neste caso, ao fazermos nosso estudo, **a situação de cada município traria informação totalmente nova**: antes de ler a cifra marcada na bola, nada poderíamos afirmar sobre o município.

Imagine agora, uma situação diversa, em que o total de loja de cada cidade tivesse sido determinado em duas etapas. Em cada etapa seria especificada uma, de duas quantidades aditivas. Para definir a primeira, calcular-se-ia a média de m^2 de loja dos municípios vizinhos a um município em foco: este seria o valor-base. Para definir a outra parte, uma bola seria extraída da urna de metragens; se o valor registrado na bola fosse positivo, ele seria somado ao valor-base; se negativo, seria subtraído.

Neste processo em duas etapas, ao fazermos um levantamento sobre um município particular, a informação obtida não seria totalmente nova. Ela seria, em parte, uma repetição do que já sabíamos em função do estudo dos municípios vizinhos.

A autocorrelação espacial refere-se a esta redundância de informação entre duas realizações de um fenômeno quando elas ocorrem próximas uma da outra. Sua presença distorce os resultados obtidos por meio da aplicação de modelos

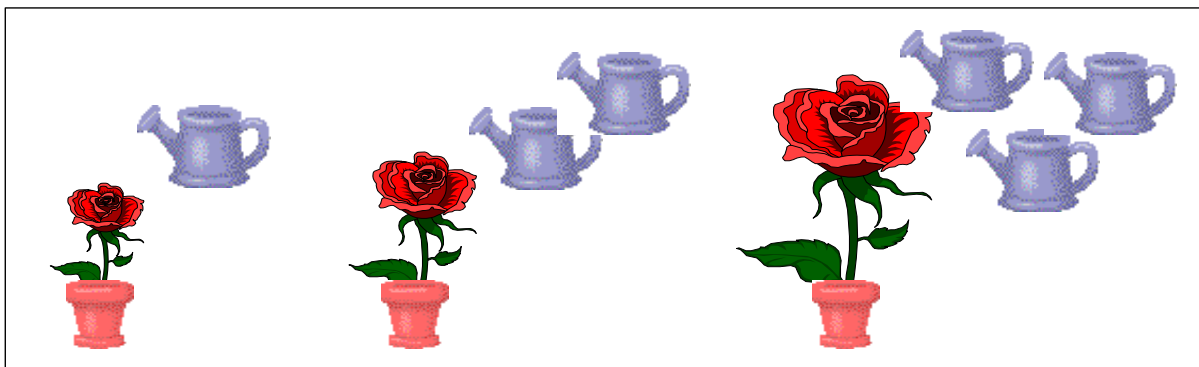
⁴ Esta variável, levantada no trabalho anterior, já citado, foi agregada posteriormente para microrregiões.

estatísticos tradicionais, abafando ou exacerbando relações de causalidade ou associação.

Considere, por exemplo, um cenário em que o tamanho de uma roseira (medido em cm) depende linearmente da quantidade de fertilizante com que é regada durante um determinado período, segundo a relação:

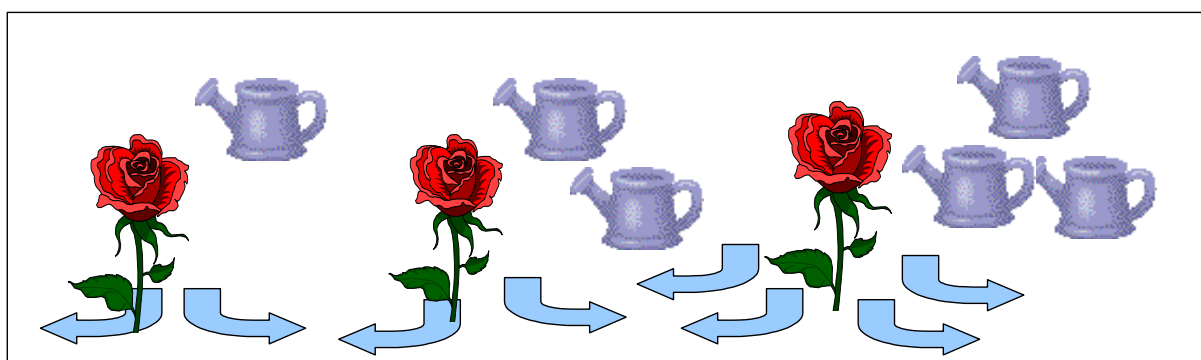
$$\text{tamanho total} = \text{tamanho autônomo} + \alpha (\text{unidades de fertilizante})$$

De acordo com este modelo, uma planta que não recebesse nenhum fertilizante desenvolver-se-ia até o tamanho autônomo (tamanho que não depende do fertilizante). Outra, que tivesse recebido 1 unidade de fertilizante mediria o tamanho autônomo mais α centímetros de altura, já que α corresponde ao número de centímetros de crescimento induzido por uma unidade adicional de fertilizante. Com duas unidades de fertilizante, o tamanho seria o autônomo mais 2 vezes α cm, e assim por diante. Este cenário está representado na ilustração a seguir em que as rosas crescem proporcionalmente mais, conforme são mais fertilizadas quimicamente.



Se, no entanto, nossas rosas não estivessem plantadas em vasos, mas no chão, haveria um extravasamento, de uma planta para a outra, do fertilizante administrado. De tal maneira que uma planta que não tenha sido regada com

fertilizante pode ter se beneficiado de parte do que foi utilizado na rega da planta vizinha – crescendo aparentemente mais do que seria de se esperar se não tivesse havido contaminação. De outro lado, parte do fertilizante administrado em maior concentração numa terceira roseira dissipou-se pelo solo; a planta que recebeu a rega não pôde beneficiar-se integralmente da quantidade de fertilizante aplicada – e assim cresceu aparentemente menos do que se esperaria caso nada tivesse se transferido para uma região fora do seu alcance.



Neste exemplo, se tentássemos estimar α (o efeito no crescimento de uma unidade adicional de fertilizante) utilizando um modelo de regressão linear simples, subestimariamos o efeito do fertilizante sobre o crescimento. A autocorrelação entre os resultados tende a amortecer as diferenças de crescimento observadas entre as plantas. Se uma cresceu muito, é provável que as vizinhas também o tenham feito. De forma análoga, se cresceu pouco, as vizinhas também devem ter pouca altura.

A tendência de valores localizados próximos serem parecidos caracteriza uma autocorrelação positiva. Este tipo de dinâmica deprime a manifestação local de diferenças e reforça as diferenciações regionais. Frequentemente oculta a importância de variáveis explicativas subjacentes ao processo estudado.

Existe também a situação inversa.

Algumas espécies de pássaro, por exemplo, evitam formar grupos muito numerosos, procurando manter uma razoável distância entre os bandos. Assim, sabendo-se que numa determinada área há uma elevada concentração destas aves, podemos prever com segurança que nas regiões vizinhas elas aparecerão em pequenos números.

Este tipo de influência caracteriza uma autocorrelação espacial negativa, em que ocorrências bastante diferentes entre si tendem a estar próximas: muitos pássaros numa área, poucos pássaros na área vizinha. Nesta dinâmica, as diferenças locais do fenômeno são exacerbadas, ao contrário das diferenças regionais, que são enfraquecidas.

Em qualquer dos dois casos, a presença de autocorrelação influencia a quantidade e qualidade da informação contida em dados espacializados, afetando as interpretações possíveis. Os testes estatísticos tradicionais, baseados no pressuposto de independência entre as observações, resultam distorcidos quando aplicados a dados com estas características. Da mesma forma, a distribuição das estimativas de parâmetros populacionais baseadas em amostras não correspondem às formas funcionais tradicionais.

Como a autocorrelação surge pela influência das realizações vizinhas sobre a realização do fenômeno na área de interesse, sua intensidade geralmente decai com a distância, tornando-se irrelevante quando as observações estão suficientemente afastadas entre si. Em outras palavras, a partir de uma certa distância, a rega de uma roseira não influencia mais o crescimento de outra planta. Um dos aspectos que interessa conhecer sobre a autocorrelação, portanto, é a relação de sua intensidade com a distância entre observações.

Nas explicações acima, e nas que se seguirão, estamos nos baseando nos trabalhos de Ripley⁵, Fotheringham⁶ e Griffith⁷, especialmente no trabalho deste último, inteiramente dedicado ao tema de autocorrelação espacial⁸.

1. DEFINIÇÃO DE AUTOCORRELAÇÃO ESPACIAL

1.1. Correlação

O conceito de “correlação” no contexto estatístico diz respeito ao relacionamento entre duas variáveis.

Quando ambas são intervalares, geralmente utiliza-se como medida da intensidade de sua relação o coeficiente de Pearson (ρ), que mede o grau de correlação linear entre elas. Neste caso, o coeficiente de correlação varia de -1 a 1 . Quando $|\rho|$ tende a zero, a relação entre as variáveis é fraca. Quando $|\rho|$ tende a 1 , a relação é forte. O sinal do coeficiente indica a natureza da relação: se negativo, mostra que ao crescimento de uma variável corresponde uma diminuição da outra; se positivo, sinaliza que as duas variáveis caminham na mesma direção, isto é, crescem ou decrescem juntas⁹.

⁵ RIPLEY, Brian D. *Spatial Statistics*, NY: Wiley, Wiley Series in Probability and Statistics, 1981.

⁶ FOTHERINGHAM, Steward and ROGERSON, Peter (eds). *Spatial Analysis and GIS*. London: Taylor & Francis, 1994.

⁷ GRIFFITH, Daniel A. *Spatial Autocorrelation*, Washington: Association of American Geographers, 1987.

⁸ Um dos desafios para a leitura da literatura sobre autocorrelação espacial é a exigência de uma matemática avançada, superior àquela de que dispõem, em geral, os profissionais com formação em Administração de Empresas. Este trabalho de Griffith tem o grande mérito de tentar fazer uma ponte entre as abordagens de estatística básica e as mais avançadas.

⁹ Para maiores detalhes sobre o coeficiente de correlação de Pearson, por favor refira-se a qualquer texto de estatística básica. Entre eles, sugerimos, por exemplo, STEVENSON, William J., *Estatística Aplicada à Administração*, São Paulo: Harbra, 1981, pp. 370-5.

1.2. Autocorrelação

O conceito de autocorrelação surge quando se verificam duas condições:

- as observações de uma variável estão ordenadas de alguma maneira;
- cada observação (observação em foco) da variável relaciona-se com outra ou outras observações da **mesma variável** que ocupam uma determinada posição em relação à observação em foco.

Assim, por exemplo, as cotações do preço de uma *commodity* estão ordenadas no tempo, formando uma série temporal. A cada dia, a cotação não varia livremente: em geral, ela será muito parecida com a cotação da véspera, já que as oscilações diárias, pelo menos em períodos de tranquilidade econômica, costumam ser pequenas, e partem do ponto em que o pregão se encerrou na véspera. Em outras palavras, elas não são independentes.

Em geral, não só a ocorrência imediatamente anterior é relevante. No acompanhamento das vendas do comércio, por exemplo, costuma-se comparar os resultados obtidos pelo varejo no mês precedente e no mesmo mês de anos anteriores.

1.3. Autocorrelação espacial

Dados organizados no tempo são relativamente mais fáceis de analisar do que dados organizados geograficamente. No tempo, os fenômenos organizam-se apenas em uma direção, e em um único sentido: as cotações do ano que vem não afetam as cotações de hoje (ou de amanhã), simplesmente porque ainda não aconteceram. Já no espaço, todas as direções, e todos os sentidos, podem ser relevantes. O potencial

de consumo de uma cidade pode ser afetado por seus vizinhos a Norte, Sul, Leste, Oeste, etc.

Em resumo, a autocorrelação espacial ocorre quando observações (de uma única variável) organizadas no espaço influenciam-se mutuamente.

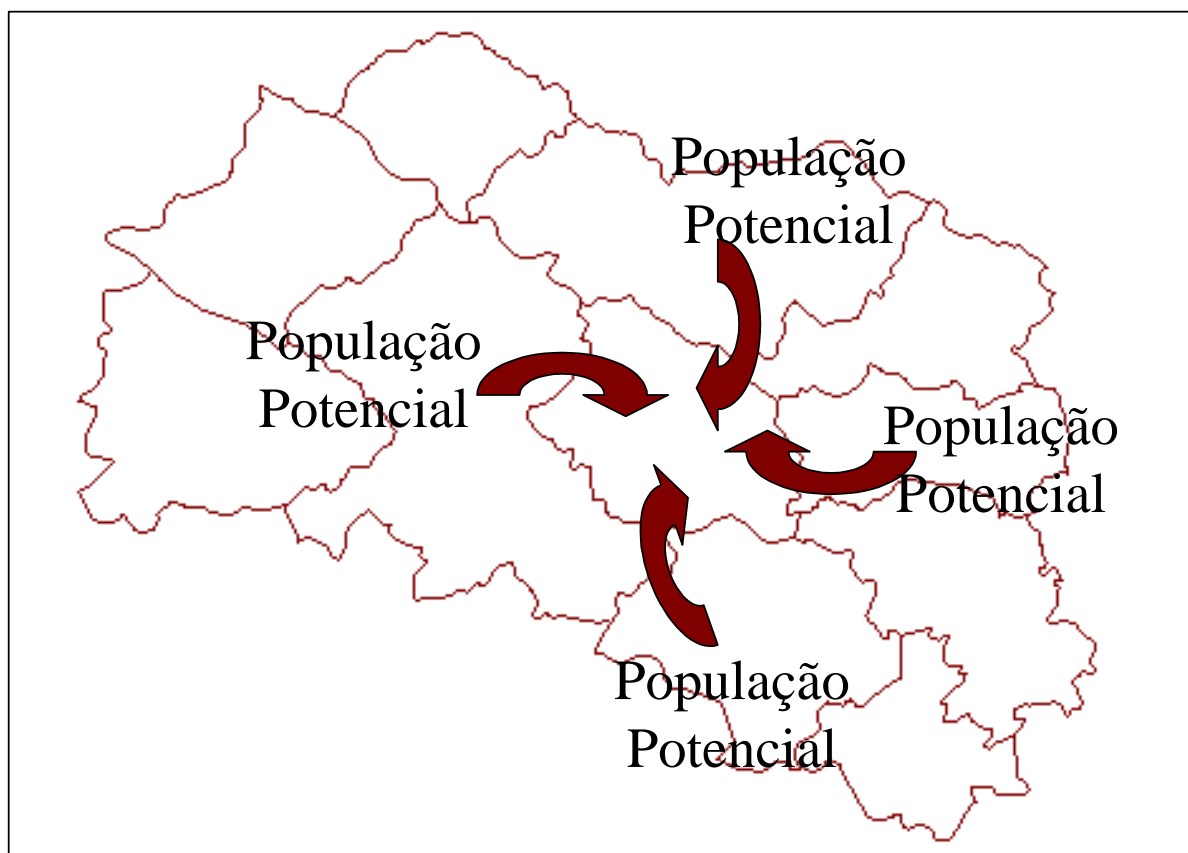
É o caso da concentração de supermercados.

Conforme discutimos em nosso trabalho anterior¹⁰, o potencial de mercado de uma região depende da existência de pessoas, com renda e disposição para comprar. Nosso conhecimento prático já é suficiente para constatar que cidades com alta concentração populacional e alta concentração de renda (portanto, com muitos supermercados) tendem a estar próximas de outras cidades de alta densidade (populacional, de renda e de supermercados).

Até porque é razoável imaginar que o potencial econômico “extravasa” de um município para o outro, como na nossa ilustração do fertilizante, seria raro, como regra geral, encontrar um município com densidade de loja extremamente alta ao lado de outro, com densidade extremamente baixa.

A figura seguinte representa a possibilidade da população de municípios vizinhos gerarem demanda no município de interesse.

¹⁰ ARANHA, Francisco. Relatório citado.



Em resumo, poderíamos detalhar nossa hipótese de trabalho postulando que há, entre as cidades, uma autocorrelação positiva no que diz respeito à concentração de supermercados.

2. UMA MEDIDA DE AUTOCORRELAÇÃO ESPACIAL

O primeiro passo na verificação da hipótese acima é medir a intensidade da autocorrelação nos dados de supermercados.

Como as variáveis de interesse são intervalares, uma medida da autocorrelação espacial pode ser o Índice de Geary (GR = Geary Ratio), dado pela seguinte fórmula¹¹:

$$GR = [(n-1) \sum_{i=1}^n \sum_{j=1}^n c_{ij} (x_i - x_j)^2] / [2(\sum_{i=1}^n \sum_{i=1}^n c_{ij}) \sum_{i=1}^n (x_i - \bar{x})^2]$$


onde:

- x_i e x_j são os valores da variável de interesse nas cidades i e j ;
- \bar{x} é a média da variável de interesse em todas as cidades da amostra;
- n é o número de cidades na amostra;
- c_{ij} assume o valor “1” se x_i e x_j são vizinhos e “0”, caso contrário.


Podemos re-escrever a expressão apresentada por Griffith de forma mais conveniente para uma interpretação:

¹¹ GRIFFITH, Daniel A, obra citada, pp. 36-40.

$$GR = \frac{\sum_{i=1}^n \sum_{j=1}^n c_{ij} (x_i - x_j)^2}{2 \sum_{i=1}^n \sum_{i=1}^n c_{ij}} \bigg/ \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)} \quad (1)$$



Variabilidade de
Vizinhança



Variabilidade
Geral

2.1. Variabilidade de vizinhança

Consideremos o primeiro bloco (numerador do traço de fração representado na vertical na fórmula (1) acima), que estou chamando de “Variabilidade de Vizinhança”.

Neste quociente, não fosse por c_{ij} , o numerador representaria a soma das diferenças ao quadrado entre cada particular “x” e todos os demais. No entanto, como c_{ij} assume o valor “0” quando os elementos comparados não são vizinhos, cancelando o resultado da comparação, isto é, cancelando aquele particular desvio, acabam só entrando na soma os desvios entre os vizinhos, já que estes são multiplicados por “1” e, portanto, não são afetados. No denominador, $\sum_{i=1}^n \sum_{i=1}^n c_{ij}$ corresponde ao número total de parcelas somadas, e o “2” ajusta o fato de que a medida de dispersão de todos os elementos de um conjunto em relação a todos os demais resulta no dobro da dispersão de todos os elementos em relação ao meio do conjunto¹², isto é, resulta

¹² Considere uma situação em que todos os “x” são vizinhos entre si, e, portanto, todos os desvios $(x_i - x_j)$ são incluídos na soma. Então:

no dobro da variância¹³.

$$\frac{\sum_{i=1}^n \sum_{j=1}^n c_{ij} (x_i - x_j)^2}{2 \sum_{i=1}^n \sum_{i=1}^n c_{ij}} = \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2}{2n^2}$$

pois c_{ij} , sendo sempre "1" não influencia o numerador, e, se n é o número total de elementos no conjunto estudado, o número total de comparações possíveis, denotadas por $\sum_{i=1}^n \sum_{i=1}^n c_{ij}$, é n^2 . Adicionando e subtraindo \bar{x} dentro do parêntesis do numerador e expandindo o quadrado, temos:

$$\begin{aligned} \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - x_j)^2}{2n^2} &= \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x} - x_j + \bar{x})^2}{2n^2} = \frac{\sum_{i=1}^n \sum_{j=1}^n [(x_i - \bar{x}) - (x_j - \bar{x})]^2}{2n^2} = \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n [(x_i - \bar{x})^2 + (x_j - \bar{x})^2 + 2(x_i - \bar{x})(x_j - \bar{x})]}{2n^2} = \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n \sum_{j=1}^n (x_j - \bar{x})^2 + 2 \sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})(x_j - \bar{x})}{2n^2} \end{aligned}$$

Como $2 \sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})(x_j - \bar{x}) = 0$, e como $\sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n \sum_{j=1}^n (x_j - \bar{x})^2$ vem

$$\begin{aligned} &= \frac{\sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n \sum_{j=1}^n (x_j - \bar{x})^2 + 2 \sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})(x_j - \bar{x})}{2n^2} = \frac{2 \sum_{i=1}^n \sum_{j=1}^n (x_i - \bar{x})^2}{2n^2} = \\ &= \frac{2 \sum_{i=1}^n n(x_i - \bar{x})^2}{2n^2} = \frac{2n \sum_{i=1}^n (x_i - \bar{x})^2}{2n^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \end{aligned}$$

que é a variância de x_i , como queríamos demonstrar.

¹³ Para maiores esclarecimentos sobre a variância, por favor consulte um texto de estatística básica.

Em outras palavras o valor calculado neste bloco corresponde à variância (espalhamento) das observações definidas por c_{ij} como vizinhas entre si.

2.2. Variabilidade do conjunto total de dados

O segundo bloco, denominador do traço vertical de fração em (1) é, obviamente, a variância - uma medida do espalhamento quadrático médio - das observações do conjunto de dados.

2.3. Comparando as variabilidades

Se quantidade de supermercados for definida espacialmente por acaso em cada município, isto é, se a metragem de loja em uma cidade não tiver nenhuma relação com a metragem das cidades vizinhas, a variabilidade de metragem no conjunto nos dados e no subconjunto dos vizinhos são duas estimativas da mesma grandeza, a variabilidade de metragem entre as cidades em geral. Assim, seu quociente deve ser próximo de 1.

Se, contudo, cidades próximas tiverem uma tendência a serem parecidas, a variância de vizinhança será menor do que a variância geral, e o GR será menor que 1.

Ao contrário, se cidades com grandes metragens ficarem circundadas por cidades com pouca metragem (e vice-versa), a variabilidade entre os vizinhos será maior que a variabilidade geral, e o GR será maior do que 1.

Tipicamente, GR assume valores entre 0 e 2.

2.4. Matrizes de conectividade

Como vimos no item anterior, o coeficiente de Geary detecta e mensura a presença de autocorrelação espacial por meio da comparação entre a variabilidade dos vizinhos e a variabilidade geral dos dados. Assim, precisamos definir o que são “vizinhos” e registrar a ocorrência de vizinhança entre as observações.

A definição de vizinhança pode ser feita de maneira contínua ou discreta.

Na estratégia contínua, duas observações são tão mais fortemente vizinhas quanto mais próximas estiverem. Assim, todas as observações são vizinhas entre si, porém com diferentes intensidades de vizinhança. E, como já vimos, esta intensidade decai com a distância.

Na estratégia discreta, que adotamos ao utilizarmos o coeficiente de Geary definido em (1), ou duas observações são vizinhas – e neste caso o correspondente c_{ij} da fórmula assume o valor “1” -, ou não são – e neste caso c_{ij} assume o valor “0”. A definição de vizinhança, na estratégia discreta, depende da definição de faixas de distância. As observações que distam entre si valores dentro de uma certa faixa, podem ser definidos como vizinhos segundo aquele tipo de faixa. Se distam valores fora daquela faixa, não são vizinhos segundo aquele critério.

As relações de vizinhança segundo os diversos critérios de faixas de distância podem ser representadas por matrizes que organizam os diversos valores de c_{ij} na forma de uma tabela. Este tipo de representação matricial, chamada de matriz de conectividade, facilita o cálculo do índice de Geary.

Antes de prepararmos as matrizes de conectividade de que necessitamos, gostaríamos de detalhar os conceitos de microrregião, faixas de distância e de critérios de vizinhança.

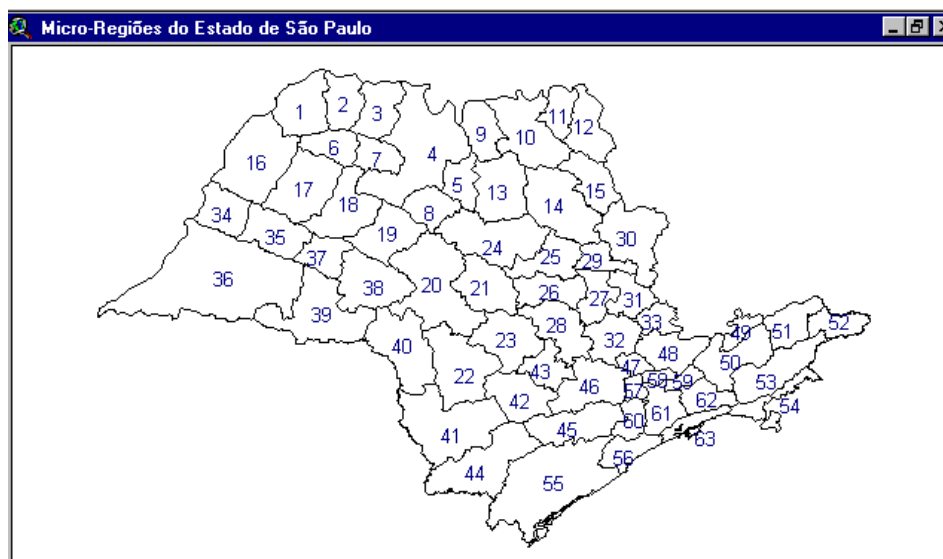
2.4.a. Microrregiões

Embora nossa intenção inicial fosse trabalhar com os municípios do estado de São Paulo individualmente, neste trabalho vamos agrupá-los em microrregiões¹⁴, segundo a definição do IBGE, representada no mapa a seguir. Esta decisão deveu-se a uma restrição de software e à disponibilidade de mapas e bancos de dados relativos a microrregiões:

- o aplicativo que tínhamos à disposição não suportaria manipular matrizes tão grandes quanto as necessárias para tratar as cidades individualmente. Ao contrário, para trabalhar com as microrregiões, as matrizes de conectividade resultaram de dimensão 63 x 63;
- entre as diversas alternativas de agregação, a utilização das microrregiões, além de ser teoricamente coerente, nos facilitava o trabalho, uma vez que, em decorrência de pesquisas anteriores, dispúnhamos de amplo material sobre as microrregiões, inclusive mapas digitais para uso com o software de geoprocessamento¹⁵.

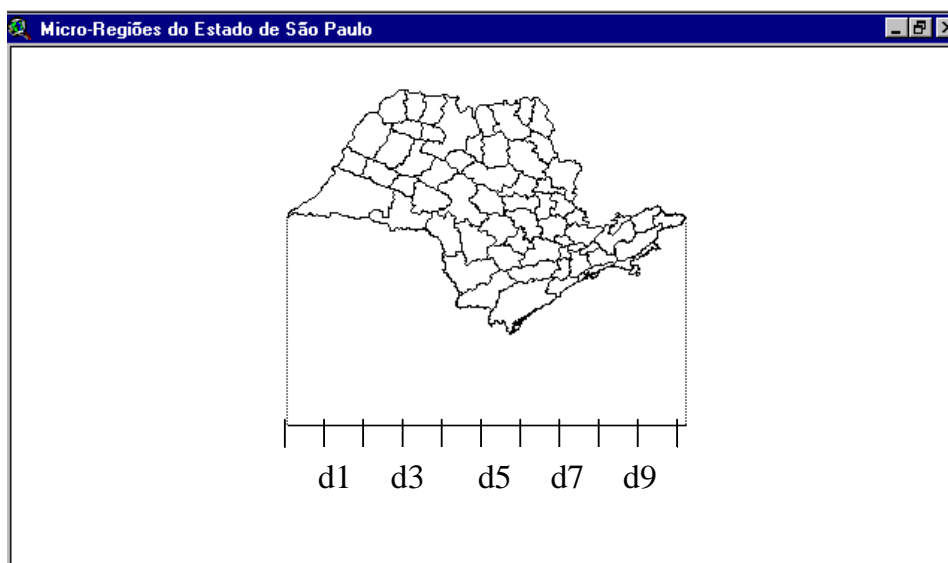
¹⁴ Conforme definição do IBGE, são 63 as microrregiões do estado de São Paulo. Veja no Apêndice VI.3 a lista completa.

¹⁵ Para manipulação e consolidação dos bancos de dados de municípios e elaboração dos mapas, utilizamos o software ARC View, da ESRI. Os interessados em maiores detalhes sobre o software podem consultar a página www.esri.com, extremamente rica em informações sobre Sistemas Geográficos de Informação e suas aplicações.



2.4.b. Faixas de distância

Utilizando a distância dos pontos extremos do Estado como referência e dividindo-a em dez segmentos iguais, definimos 10 faixas de distância, separadas pelos pontos d1 a d10 conforme ilustração a seguir.



2.4.c. Critérios de vizinhança

De posse da escala de distâncias apresentada em 2.4.b acima, definimos os seguintes critérios de vizinhança:

c0: são vizinhas duas microrregiões que partilham uma fronteira;

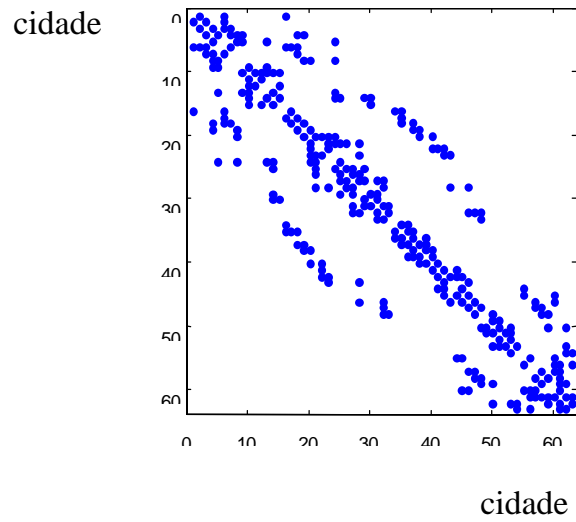
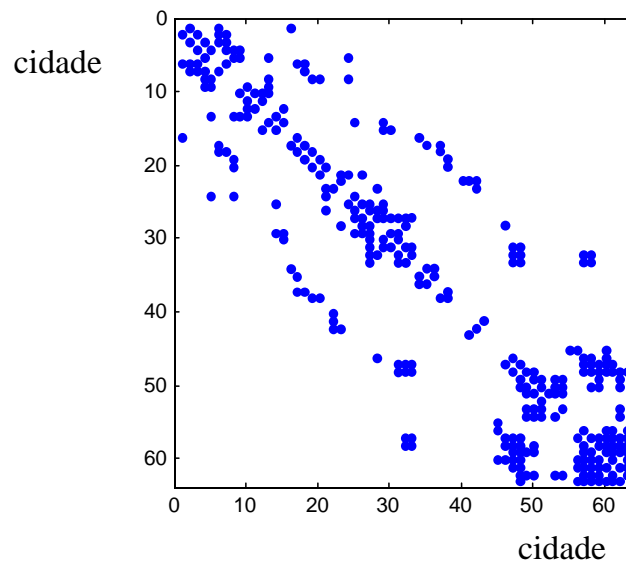
c1: são vizinhas duas microrregiões cujos centróides¹⁶ distam entre 0 e d1;

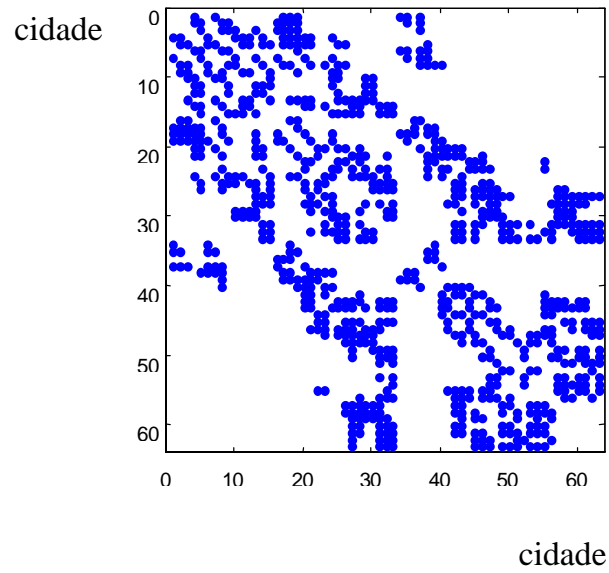
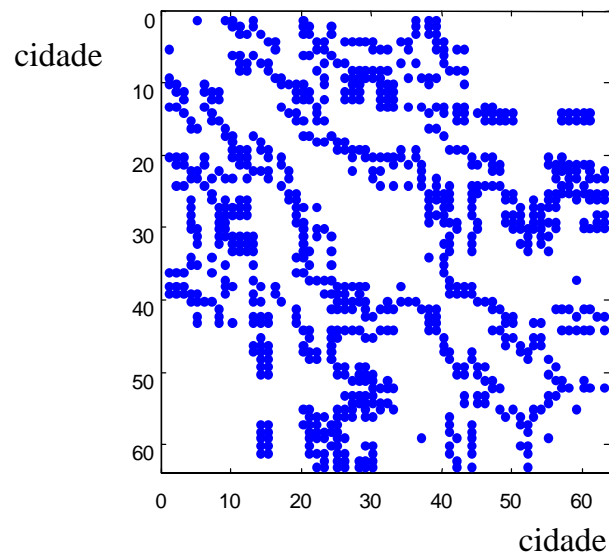
c2: são vizinhas duas microrregiões cujos centróides distam entre d1 e d2;

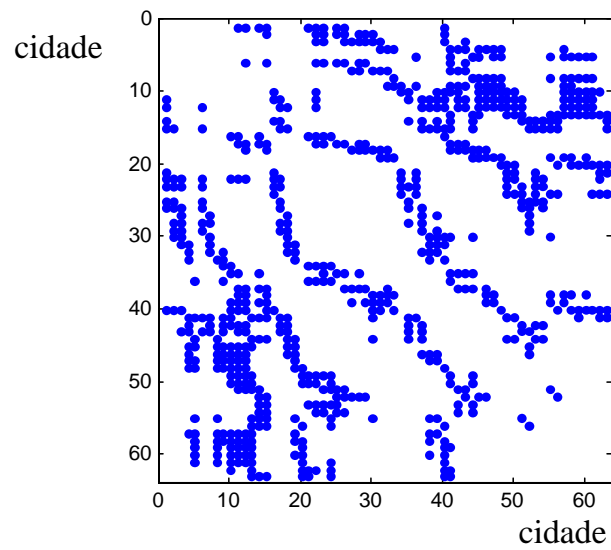
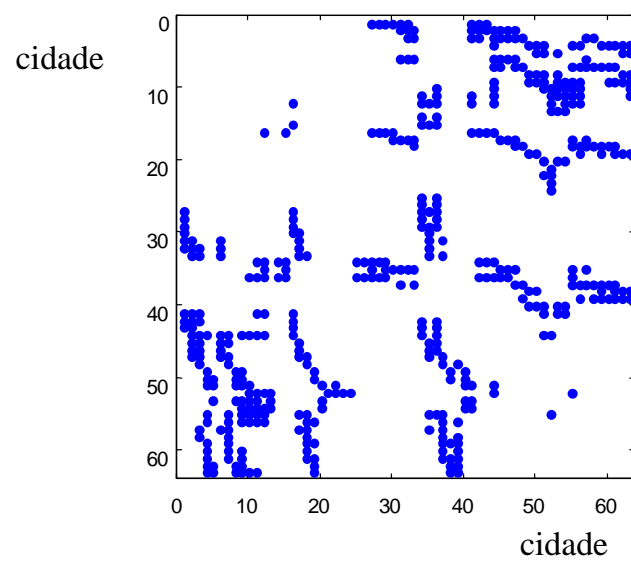
c3: são vizinhas duas microrregiões cujos centróides distam entre d2 e d3;

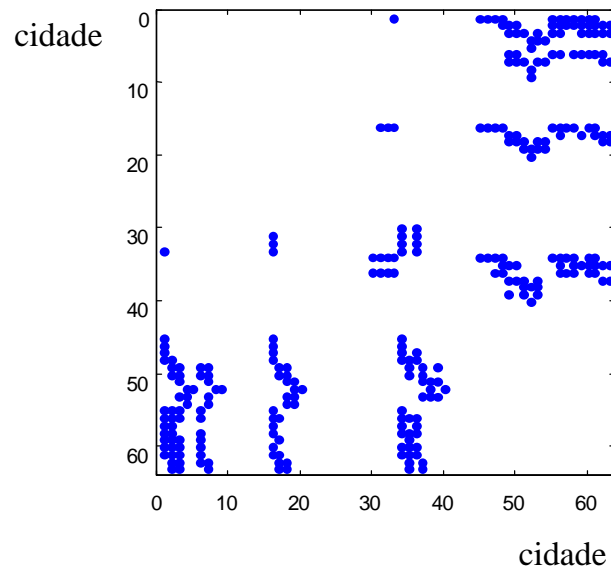
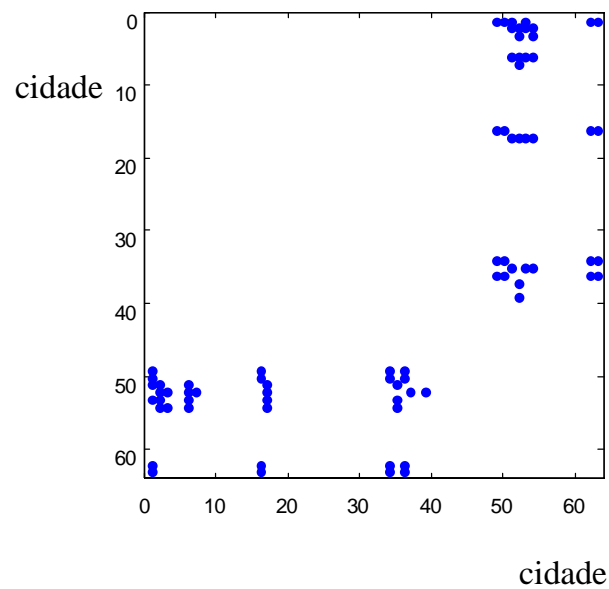
e assim, sucessivamente, até c8.

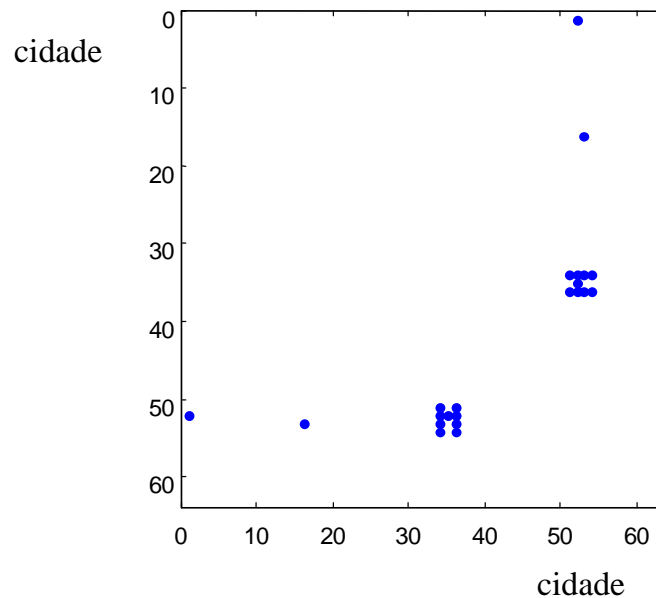
¹⁶ O centróide é o centro geográfico da microrregião.

Matriz de Conectividade **c0**Matriz de Conectividade **c1**

Matriz de Conectividade **c2**Matriz de Conectividade **c3**

Matriz de Conectividade **c4**Matriz de Conectividade **c5**

Matriz de Conectividade **c6**Matriz de Conectividade **c7**

Matriz de Conectividade c_8 

Observe como os vizinhos são mais numerosos nas faixas intermediárias de distância. Obviamente, se a condição de vizinhança é a de uma microrregião estar a uma distância muito pequena de outra, estamos tratando de um território restrito, onde não há espaço para muitos vizinhos. Se a condição de vizinhança é de estar a uma distância muito grande, os candidatos começam a “cair” fora do estado de São Paulo, e, portanto, saem do escopo do trabalho.

Observe também que cada relação de vizinhança é registrada duas vezes pois, se A é vizinho de B, B é vizinho de A. Desta maneira, a matriz de conectividade sempre é simétrica.

2.5. GR da distribuição de metragem de loja

Com o auxílio de um programa escrito em Matlab¹⁹ calculamos o GR para cada faixa de distância, conforme resumo a seguir:

	c0	c1	c2	c3	c4	c5	c6	c7	c8
Vizinhança	0.30	0.34	0.47	0.61	0.53	0.56	0.79	0.66	0.70
Geral	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.53
GR	0.57	0.64	0.89	1.15	1.00	1.06	1.49	1.25	1.32
Vizinhos	156	188	465	459	397	244	145	41	14

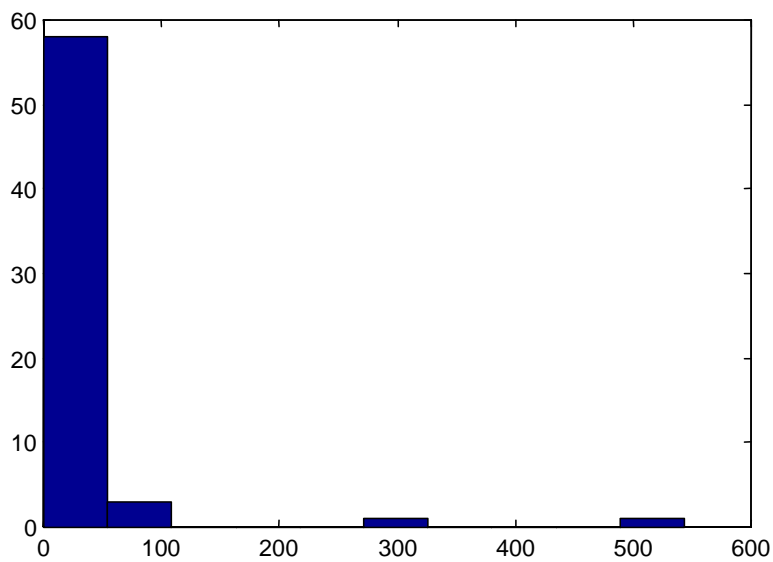
Onde:

- “Vizinhança” é a variância dos m^2 de loja entre os vizinhos;
- “Geral” é a variância dos m^2 de loja entre todas as microrregiões;
- “GR” é o Geary Ratio (Índice de Geary); e
- “Vizinhos” é o número de vizinhos existente naquela faixa de distância.

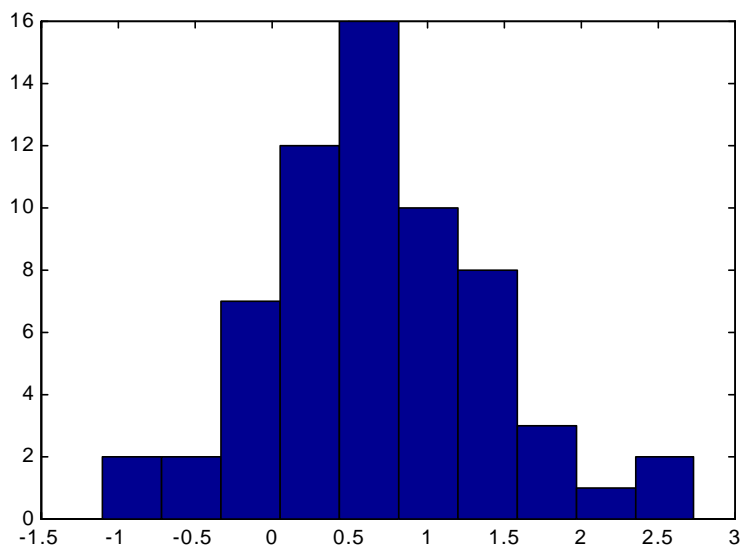
Esta tabela nos mostra, por exemplo, que os vizinhos que fazem fronteira (critério 0) apresentam uma variabilidade de 0,30 em sua metragem de loja²⁰. A variabilidade

¹⁹ O software utilizado foi o MatLab, de The MathWorks Inc. O programa foi reproduzido no apêndice VI.2.

²⁰ Os valores originais da metragem de loja incluídos no apêndice VI (m^2 de loja por km) têm a seguinte distribuição:



que justifica a transformação para escala logarítmica adotada neste relatório de pesquisa, tornando-os mais comportados conforme histograma a seguir (em que os valores apresentados são os logaritmos na base 10 dos dados originais):



dos m² de loja entre todas as microrregiões, é de 0,53. Assim, podemos perceber que os vizinhos segundo o critério 0 são bem mais parecidos entre si do que o conjunto de todas as microrregiões. A variabilidade dos vizinhos é quase metade da variabilidade geral (GR de 0,57).

No outro extremo, conforme nossa previsão inicial de que, na presença de autocorrelação espacial positiva, as diferenças locais são atenuadas e as diferenças regionais exacerbadas (vide item II.4), os vizinhos segundo o critério 8 apresentaram uma variabilidade de vizinhança de 0,70, portanto 60% maior que a variabilidade geral dos dados. Isto significa que os vizinhos segundo o critério c8 são mais dissemelhantes entre si do que as microrregiões em geral.

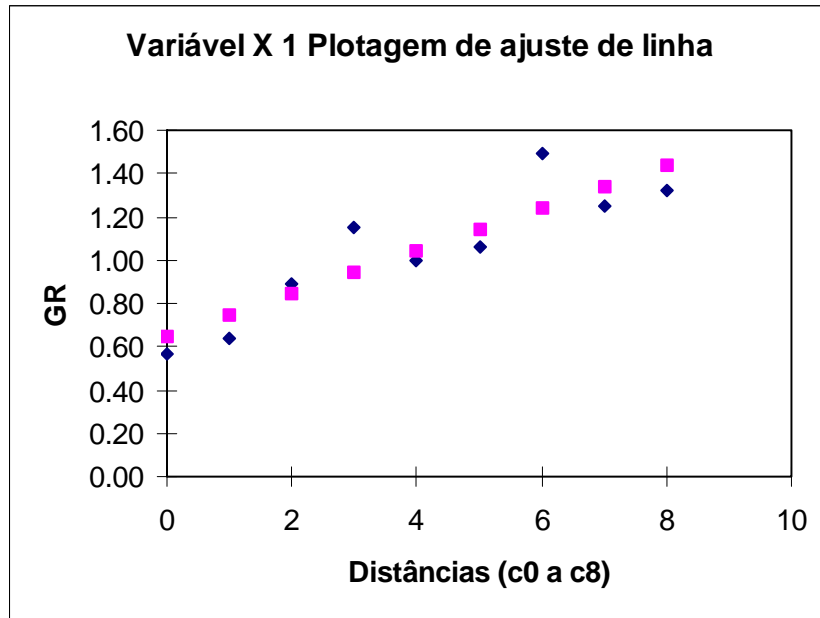
No caso das microrregiões de São Paulo, a relação entre distância e Índice de Geary pode ser representada de maneira geral pela função²¹:

$$GR = 0,64 + 0,10 \text{ distância}$$

onde a distância foi medida como 0, 1, 2, ... , 8 conforme o critério utilizado para definir vizinhança.

Esta relação pode ser melhor visualizada no gráfico a seguir, em que os pontos marcados por losangos são os valores observados e os pontos marcados por quadrados são previstos segundo a função proposta no parágrafo anterior:

²¹ Ajustada por regressão linear simples, com utilização do software Excel. Salvo menção explícita em contrário, os gráficos e regressões apresentados neste trabalho foram preparados com a utilização deste software da família Office.



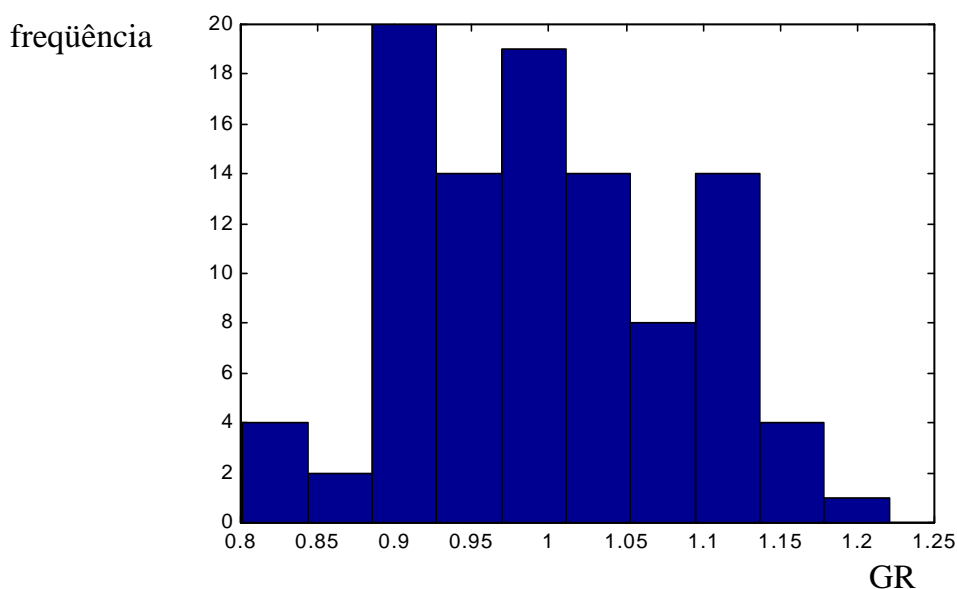
3. SIGNIFICÂNCIA

Mesmo que a localização não tenha nenhuma influência sobre uma variável (o que significa que GR deveria ser próximo a 1), poderíamos obter valores que se afastam de 1 puramente por acaso.

Imagine, por exemplo, que os 63 valores de m^2 de loja com que estamos trabalhando sejam fixos e estejam inscritos cada um numa bola contida numa urna. Teríamos 63! (isto é, aproximadamente 2 seguido de 67 zeros) formas diferentes de dispor estas bolas nas 63 regiões. Cada uma destas disposições teria associado a si um valor de Índice de Geary, formando uma população de medidas de GR. A média desta população deve ser 1, mas quando, por acaso, valores parecidos ocorreram geograficamente próximos, o valor de GR será menor que 1; e inversamente, quando, por acaso, valores muito diferentes ocorrem juntos, GR será maior que 1.

Estudando a distribuição dos valores dos índices associados às possíveis diferentes configurações espaciais dos valores de metragem de loja, poderíamos saber a chance de valores diferentes de 1 aparecerem quando a alocação das bolas é feita por acaso.

Tomemos o caso do critério 0. Utilizando um programa desenvolvido em MatLab²², fizemos a simulação de 100 sorteios de diferentes configurações de m² de loja e calculamos os correspondentes GR segundo este critério de vizinhança. Os resultados obtidos tiveram a seguinte distribuição:



O histograma nos mostra que valores abaixo de 0,80 ou acima de 1,225 são extremamente raros quando a metragem de loja é organizada espacialmente de forma aleatória (nenhum valor foi inferior a 0,80, ou superior a 1,225, em 100 tentativas). Portanto, a chance do valor de $GR = 0,57$ observado para os vizinhos segundo c0 (no item II.2.) ter ocorrido por acaso é muito pequena. O que torna este valor extremamente significativo, isto é, admite-se que ele ocorreu porque, de fato, existe uma autocorrelação positiva entre as microrregiões.

²² O programa citado está reproduzido no apêndice VI.5.

Sempre por simulação²³, montamos a tabela a seguir, indicando os percentis 2,5 e 97,5 dos GRs para cada categoria de distância. A probabilidade de observar um valor inferior ao percentil 2,5 ou superior ao percentil 97,5 é de aproximadamente 5% (não é “exatamente” 5% porque estamos nos baseando numa amostra de 100 simulações para estimar a distribuição na população). Assim, um valor de GR será considerado significativo se ficar fora do intervalo determinado por estes dois pontos de corte.

	c0	c1	c2	c3	c4	c5	c6	c7	c8
P(2,5)	0.85	0.77	0.85	0.84	0.88	0.78	0.57	0.48	0.35
P(97,5)	1.16	1.26	1.14	1.16	1.16	1.23	1.45	1.81	2.26
GRo	0.57	0.64	0.89	1.15	1.00	1.06	1.49	1.25	1.32
Significativo	Sim	Sim	Não	Não	Não	Não	Sim	Não	Não
No. de Vizinhos	156	188	465	459	397	244	145	41	14

Observe que os resultados das categorias de distância c7 e c8 apresentam uma faixa de resultados não significativos muito ampla o que torna muito difícil obter resultados significativos. Isto se deve ao pequeno número de vizinhos nestas categorias, o que expõe as amostras a variações casuais mais fortes.

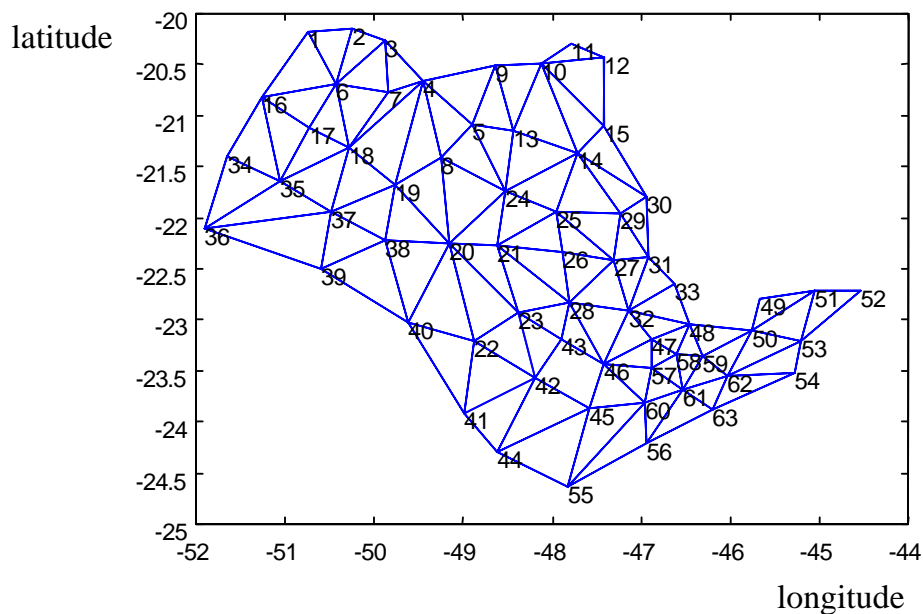
Os valores observados de GR nas categorias c7 e c8, embora não significativos, estão de acordo com o esperado pelos motivos teóricos discutidos no item II.

4. REMOÇÃO DA AUTOCORRELAÇÃO ESPACIAL

Considerando-se que apenas a autocorrelação a pequenas distâncias mostrou-se significativa (a categoria c6 ficou praticamente no limite entre significância e não significância), e que os vizinhos segundo c0 e c1 se sobrepõem parcialmente (veja os gráficos da página 26, [verificar se não mudou de página](#)), só vamos caracterizar como vizinhas as microrregiões que partilham fronteiras (critério 0). Nestes termos,

²³ Uma tabela contendo o conjunto dos valores simulados foi incluída no apêndice VI.6.

a relação de vizinhança entre as microrregiões pode ser representada pelo diagrama a seguir, em que os eixos horizontal e vertical correspondem às coordenadas geográficas (latitude, longitude) em graus decimais; os nós da figura correspondem aos centróides das microrregiões, rotulados segundo os números da tabela de identificação do apêndice VI.3.; e as ligações por meio de linhas representam as relações de vizinhança.



4.1. Variável dependente: densidade de loja na microrregião

Consideremos que a metragem de loja de cada microrregião é definida pela média da metragem de loja das microrregiões vizinhas mais um valor explicado por outros motivos, por hora chamados, em conjunto de “acaso”, conforme discussão realizada no item II. Este modelo pode ser enunciado algebricamente como:

$$x_i = \rho \left(\sum_{j=1}^n c_{ij} x_j / \sum_{j=1}^n c_{ij} \right) + u_i \quad (2)$$

onde:

- x_i representa os m^2 de loja da cidade i ,
- $\rho \left(\sum_{j=1}^n c_{ij} x_j / \sum_{j=1}^n c_{ij} \right)$ representa a média de m^2 dos vizinhos da cidade i , ponderada pelo coeficiente de correlação, e
- u_i representa a parcela de m^2 da cidade i não explicada pelos m^2 dos seus vizinhos.

Calculando o valor de ρ por regressão linear simples, chegamos aos seguintes resultados:

RESUMO DOS RESULTADOS

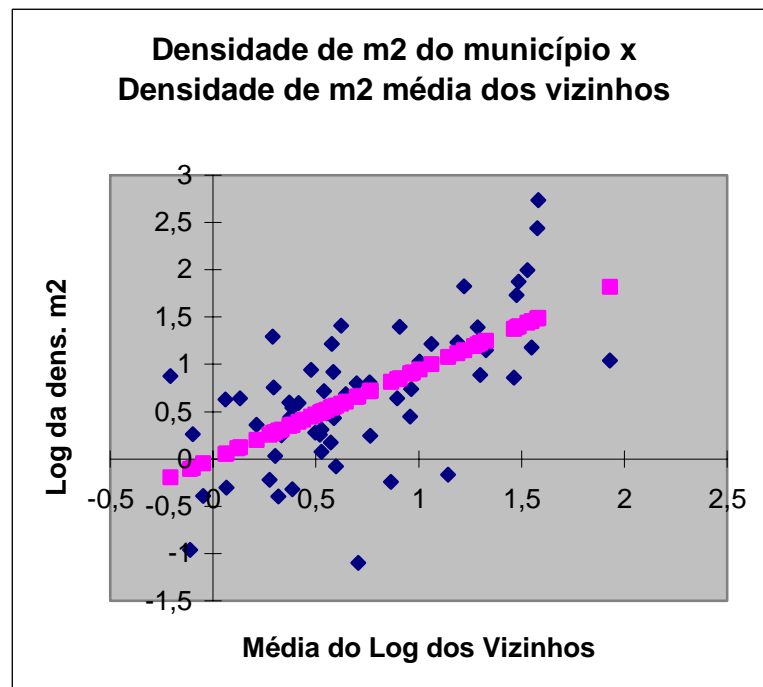
<i>Estatística de regressão</i>	
R múltiplo	0.64
R-Quadrado	0.41
R-quadrado ajustado	0.40
Erro padrão	0.56
Observações	63

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	13.65	13.65	43.80	1.04616E-08
Resíduo	62	19.33	0.31		
Total	63	32.98			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
mediavizinhos	0.94	0.08	11.57	3.98E-17

Este resultado mostra que, com um nível altíssimo de significância, a autocorrelação espacial explica cerca de 40% da variabilidade dos dados. Além disso, mostra que o valor esperado da densidade de m^2 de loja de um município é igual a 94% do valor da densidade média de seus vizinhos. A diferença entre o valor observado em cada município e o esperado segundo o modelo foi registrado no apêndice VI. Esta diferença, chamada de resíduo do modelo de autocorrelação, é o valor que precisa ainda ser explicado por outras variáveis independentes. Corresponde aos dados originais sem a parcela que pode ser explicada por autocorrelação.



No gráfico acima, novamente, os valores representados por losangos são os valores originais (log da densidade de m^2 da microrregião); e os representados por quadrados são os previstos pelo modelo (média do log da densidade das microrregiões vizinhas, vezes 0,94).

O Índice de Geary para os resíduos é de 1,17, não significativamente diferente de 1 ao nível de significância de 5% (portanto, não há evidência de autocorrelação espacial remanescente entre os resíduos).

4.2. Variável independente: densidade de renda por microrregião

No trabalho sobre potencial para supermercados, já citado, procuramos explicar a densidade de supermercados num município em função da densidade de renda ali existente. Aplicaremos o mesmo conceito no item III, a seguir, porém antes agregando os dados para microrregiões e eliminando o efeito da autocorrelação espacial, tanto na variável dependente quanto na variável explicativa, já que esta também sofre do problema (GR de 0,45, significativo ao nível de 5%).

Utilizando a mesma estratégia empregada com o log da densidade de loja, obtivemos os seguintes resultados para o log da densidade de renda:

RESUMO DOS RESULTADOS

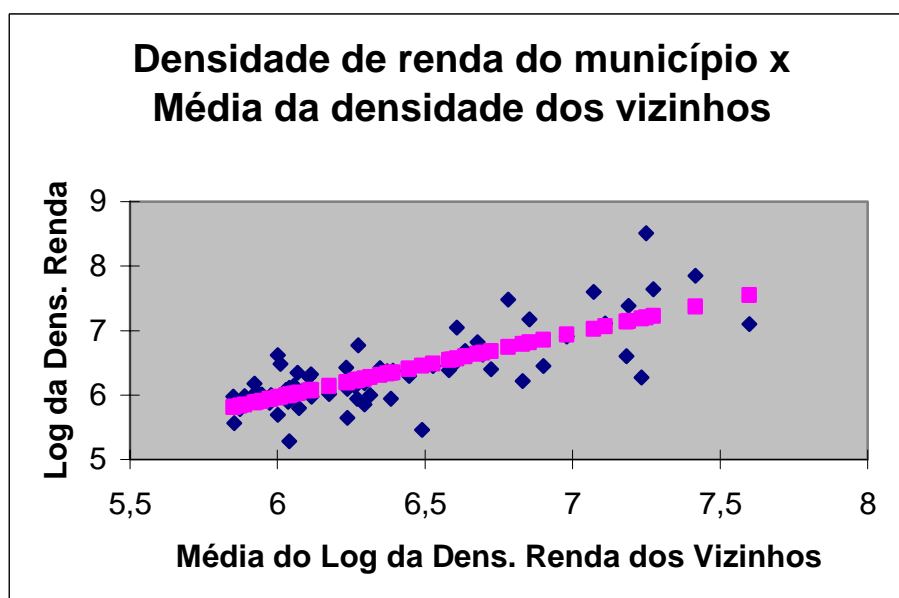
<i>Estatística de regressão</i>	
R múltiplo	0.76
R-Quadrado	0.58
R-quadrado ajustado	0.57
Erro padrão	0.39
Observações	63

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	13.21	13.21	87.20	2.2899E-13
Resíduo	62	9.40	0.15		
Total	63	22.61			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Média Vizinhos	0.99	0.01	130.37	2.4E-77

Estes resultados indicam um modelo altamente significativo, capaz de explicar cerca de 60% da variabilidade total do log da densidade de renda com base na média do log dos vizinhos, com um coeficiente de correlação de quase 1.



Os resíduos deste modelo de autocorrelação para o log da densidade de renda²⁴ (parcela da densidade de renda de uma microrregião que não depende da densidade de renda das microrregiões vizinhas) serão utilizados para explicar os resíduos do modelo de autocorrelação para a densidade de loja (parcela da densidade de loja de uma microrregião que não depende da média da densidade de loja das microrregiões vizinhas).

²⁴ Também registrados na tabela do apêndice VI.6.

III. MODELO FINAL

1. FORMULAÇÃO

O modelo final de explicação da densidade de supermercado nas microrregiões postula o seguinte:

Parte 1: A parcela de densidade de renda livre de autocorrelação espacial determina a parcela de densidade de supermercado livre de autocorrelação espacial, conforme resultados abaixo (dados sempre em escala logarítmica):

RESUMO DOS RESULTADOS

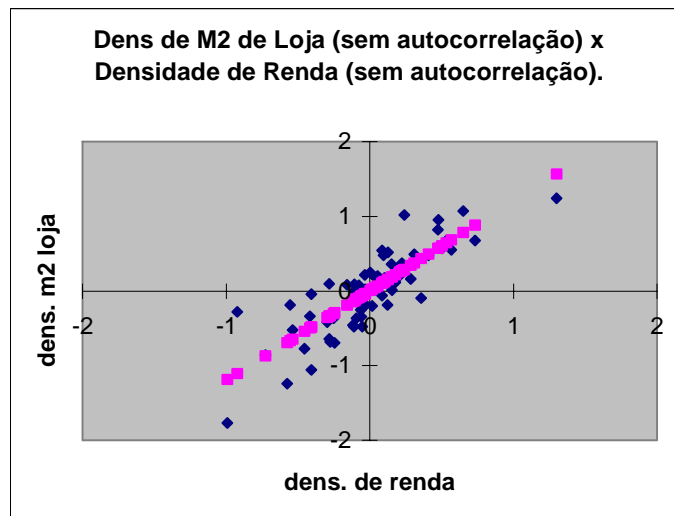
<i>Estatística de regressão</i>	
R múltiplo	0.84
R-Quadrado	0.70
R-quadrado ajustado	0.70
Erro padrão	0.31
Observações	63

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	13.56	13.56	143.57	1.13147E-17
Resíduo	61	5.76	0.09		
Total	62	19.33			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	(0.00)	0.04	(0.03)	0.974894
dens. de renda	1.20	0.10	11.98	1.13E-17

A qualidade do ajuste pode ser observada no diagrama de dispersão a seguir:



Parte 2: A média da densidade de renda dos vizinhos de uma microrregião determina a média da densidade de m^2 de supermercados, conforme resultados abaixo (dados sempre em escala logarítmica):

RESUMO DOS RESULTADOS

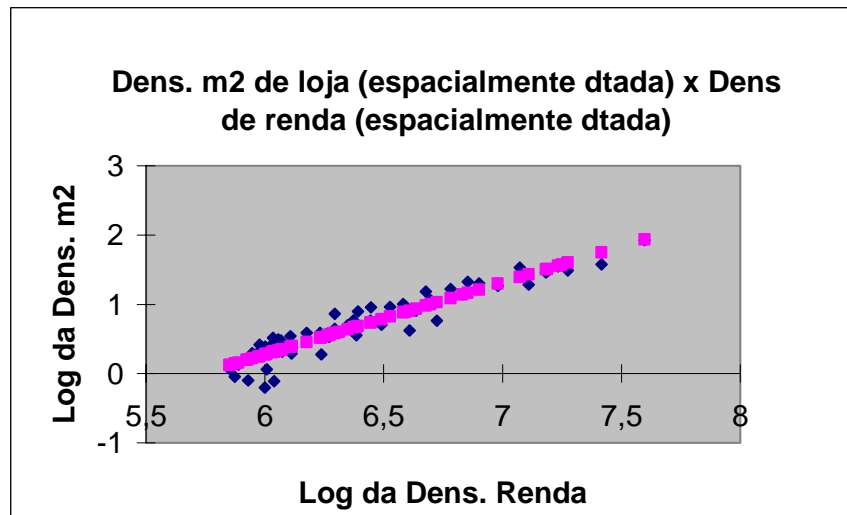
<i>Estatística de regressão</i>	
R múltiplo	0.95
R-Quadrado	0.90
R-quadrado ajustado	0.90
Erro padrão	0.16
Observações	63

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	13.68	13.68	569.14	1.2547E-32
Resíduo	61	1.47	0.02		
Total	62	15.15			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	(5.96)	0.28	(21.27)	6.65E-30
dens renda	1.04	0.04	23.86	1.25E-32

Estes resultados mostram que, com altíssima confiança, o componente espacialmente determinado da renda define o componente espacialmente determinado da densidade de loja, explicando 90% da variabilidade deste.



Combinando as duas partes, temos:

$$\log(\text{dens m}^2 \text{ loja}) = -5,96 + 1,04 (\log \text{ dens. de renda sem autocorrelação espacial}) + \\ +1,20 (\text{resíduo do log dens. de renda após remoção da autocorr.})$$

Com a seguinte capacidade de explicação:

RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0.91
R-Quadrado	0.84
Observações	63

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	1	27.55	27.55	309.35	1.42974E-25
Resíduo	61	5.43	0.09		
Total	62	32.98			

Este resultado é completamente compatível com um modelo alternativo em que a densidade de m² de loja é explicada pela densidade da renda desmembrada em duas partes funcionando como variáveis independentes e ajustada por regressão múltipla (dados em escala logarítmica):

RESUMO DOS RESULTADOS

<i>Estatística de regressão</i>	
R múltiplo	0.91
R-Quadrado	0.84
R-quadrado ajustado	0.83
Erro padrão	0.30
Observações	63

ANOVA

	<i>gl</i>	<i>SQ</i>	<i>MQ</i>	<i>F</i>	<i>F de significação</i>
Regressão	2	27.55	13.77	152.20	3.14049E-24
Resíduo	60	5.43	0.09		
Total	62	32.98			

	<i>Coefficientes</i>	<i>Erro padrão</i>	<i>Stat t</i>	<i>valor-P</i>
Interseção	(5.88)	0.54	(10.80)	0.00
dens y auto	1.02	0.08	12.06	0.00
dens y res	1.20	0.10	12.20	0.00

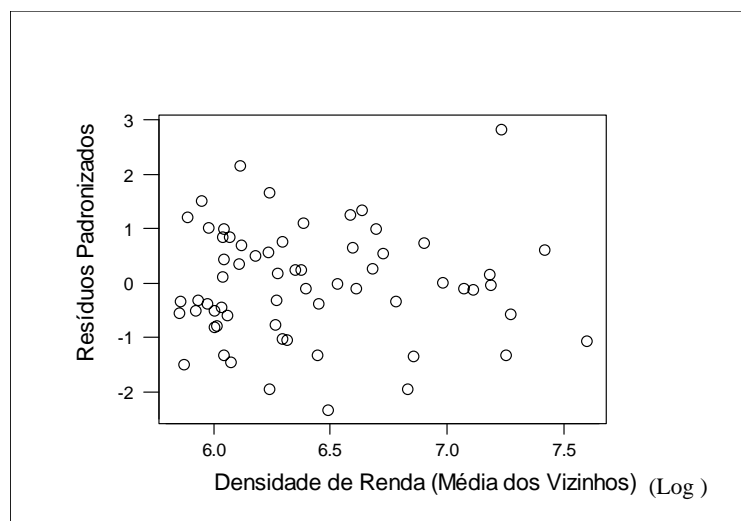
Neste caso a equação ficaria:

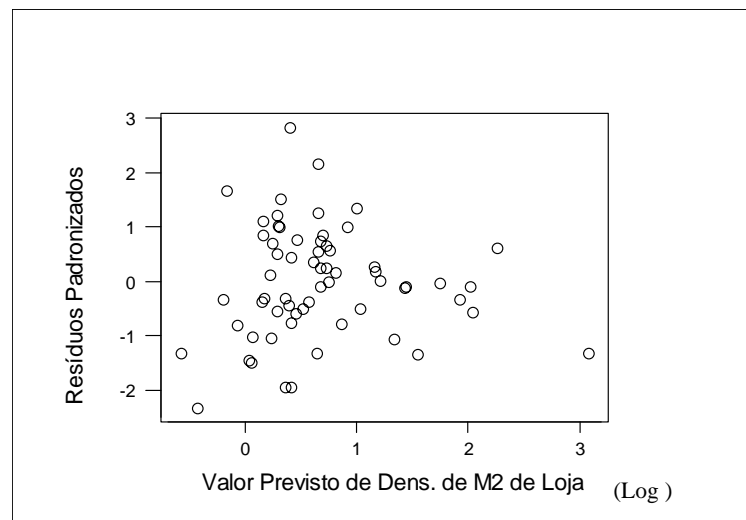
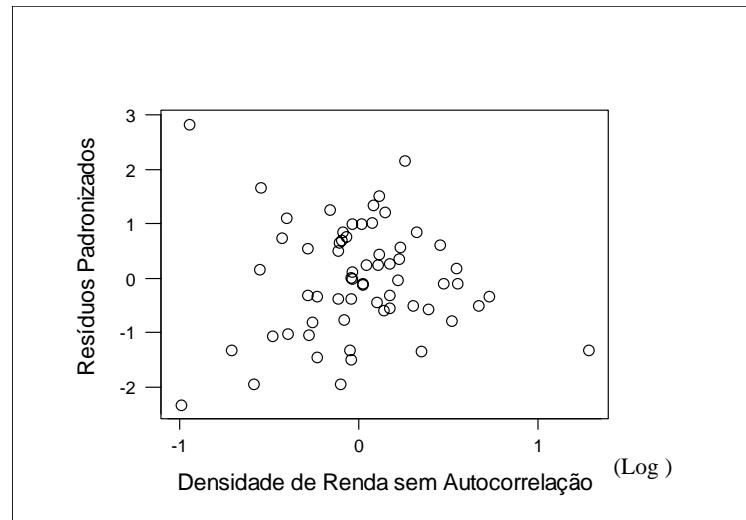
$$\log(\text{dens m}^2 \text{ loja}) = -5,88 + 1,02 (\log \text{ dens. de renda sem autocorrelação espacial}) + \\ +1,20 (\text{resíduo do log da dens. de renda após remoção da autocorr.})$$

2. AJUSTE

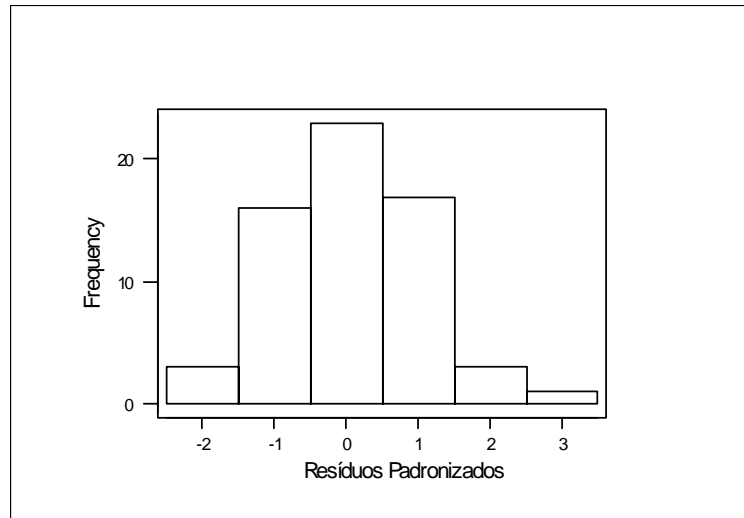
Como as duas alternativas resultam muito próximas, optaremos pelo modelo tradicional, obtido por regressão múltipla, cujas propriedades são bem conhecidas. Tanto a equação como um todo, quanto os coeficientes individuais das variáveis explicativas são significantes a um nível inferior a 0,1%.

Os gráficos de resíduos mostram valores bem comportados nas várias dimensões relevantes:





E também, resíduos bem distribuídos:



3. INTERPRETAÇÃO

O modelo nos mostra que a densidade de supermercado de uma microrregião depende:

- em parte, da densidade de renda das microrregiões vizinhas;
- em parte, do componente autônomo da densidade de renda da própria microrregião, isto é, da parcela da sua densidade de renda que é independente da densidade de renda dos vizinhos. Esta última tem um peso maior na determinação da densidade de área de loja na própria microrregião.

Retornando os valores das estimativas para a escala original temos uma correlação de 0,944 entre os valores previstos e observados, portanto uma capacidade de explicação de quase 90% da variabilidade total dos dados.

Observe que o desempenho do modelo é melhor na escala original (R^2 de 90%) do que na escala logarítmica (R^2 de 84%).

4. COMPARAÇÃO COM O MODELO ANTERIOR

No trabalho anterior, o modelo proposto foi:

$$\text{Densidade de Loja} = 0.000\ 022 (\text{Densidade de Renda})^{0.849}$$

com uma capacidade de explicação de 74% da variabilidade total em escala logarítmica, que caía para 49% quando reconvertido para a escala original.

IV. CONCLUSÃO

Com os resultados obtidos no item III, confirmamos nossas hipóteses de que:

- havia autocorrelação espacial nos nossos dados de densidade de loja (GR de 0,57, significativo, quando consideradas vizinhas as microrregiões que partilham fronteiras) e de densidade de renda;
- esta autocorrelação fazia parecer menor a influência da densidade de renda sobre a densidade de loja (R^2 de 49% sem isolamento da autocorrelação espacial, contra R^2 de 90% quando a autocorrelação é explicitada e levada em consideração separadamente);
- a remoção da autocorrelação espacial permitiu elaborar um modelo com alta capacidade de explicação da densidade de loja, agora passível de utilização em exercícios de previsão.

V. BIBLIOGRAFIA

ARANHA FILHO, Francisco José Espósito. *Potencial de Consumo dos Municípios Paulistas*, São Paulo: EAESP/FGV/NPP, relatório de pesquisa em editoração.

FOTHERINGHAM, Steward and ROGERSON, Peter (eds). *Spatial Analysis and GIS*. London: Taylor & Francis, 1994.

GRIFFITH, Daniel A. *Spatial Autocorrelation*, Washington: Association of American Geographers, 1987.

HANSELMAN, Duane and LITTLEFIELD, Bruce. *MatLab: The Language of Technical Computing*. Upper Saddle River: Prentice Hall, 1997.

RIPLEY, Brian D. *Spatial Statistics*, NY: Willey, 1981.

SEARLE, Shayle R. *Linear Models*, N.York: Wiley, 1997.

SEARLE, Shayle R. *Matrix Algebra Useful for Statistics*. New York: Willey, 1982.

STEVENSON, William J., *Estatística Aplicada à Administração*, São Paulo: Harbra, 1981.

2. PROGRAMA EM MATLAB PARA CÁLCULO DE GR

```
function [varger,varviz,gratio]=gr(ax,ac)
%GR Geary Ratio sintaxe: gr(x,c)
%x é um vetor coluna contendo as n observações
%c é uma matriz de conectividade n x n
%The Geary Ratio (GR) indica o quanto um conjunto de
%dados é auto correlacionado. Seu valor esperado é 1.
%Valores próximos de zero indicam forte autocorrelação
%positiva. Valores muito maiores do que 1 indicam
%forte autocorrelação negativa.
[m,n]= size(ax);
if n>1
    error('O primeiro argumento deve ser um vetor coluna.')
end
[o,p]=size(ac);
if o~=p
    error('O segundo argumento deve ser uma matriz quadrada')
end
if o~=m
    error('A matriz de conectividade deve ter uma dimensão compatível com o vetor de
dados')
end
for i=1:m
    for j=1:m
        dif2(i,j)=(ax(i)-ax(j)).^2 ;
    end
end
varger= std(ax).^2;%varger é a variabilidade geral
varviz=(ones(1,m)*(ac.*dif2)*ones(m,1))/(ones(1,m)*ac*ones(m,1)*2);%varviz
%é a variabilidade de vizinhança
nviz=(ones(1,m)*ac*ones(m,1))/2;%nviz é o número de vizinhos
global gratio
gratio=varviz/varger;%gratio é o Índice de Geary
varger
varviz
nviz
gratio
```

3. MICRORREGIÕES DO ESTADO DE SÃO PAULO

Código	Município	Microrregião
		Jales
350260	APARECIDA D'OESTE	1
350395	ASPÁSIA	1
351385	DIRCE REIS	1
351420	DOLCINÓPOLIS	1
352480	JALES	1
352910	MARINÓPOLIS	1
352965	MESÓPOLIS	1
353284	NOVA CANAÃ PAULISTA	1
353520	PALMEIRA D'OESTE	1
353590	PARANAPUÃ	1
354025	PONTALINDA	1
354040	POPULINA	1
354450	RUBINÉIA	1
354570	SANTA ALBERTINA	1
354610	SANTA CLARA D'OESTE	1
354660	SANTA FÉ DO SUL	1
354740	SANTA RITA D'OESTE	1
354720	SANTANA DA PONTE PENSA	1
354900	SÃO FRANCISCO	1
355490	TRÊS FRONTEIRAS	1
355580	URÂNIA	1
		Fernandópolis
351520	ESTRELA D'OESTE	2
351550	FERNANDÓPOLIS	2
351800	GUARANI D'OESTE	2
352070	INDIAPORÃ	2
352820	MACEDÔNIA	2
352960	MERIDIANO	2
353000	MIRA ESTRELA	2
353690	PEDRANÓPOLIS	2
354920	SÃO JOÃO DAS DUAS PONTES	2
355530	TURMALINA	2
		Votuporanga
350120	ÁLVARES FLORENCE	3
350180	AMÉRICO DE CAMPOS	3
351070	CARDOSO	3
351290	COSMORAMA	3
353625	PARISI	3

354030	PONTES GESTAL	3
354420	RIOLÂNDIA	3
355610	VALENTIM GENTIL	3
355710	VOTUPORANGA	3
		São José do Rio Preto
350020	ADOLFO	4
350090	ALTAIR	4
350460	BADY BASSITT	4
350480	BÁLSAMO	4
351130	CEDRAL	4
351750	GUAPIAÇU	4
351790	GUARACI	4
351940	IBIRÁ	4
351980	ICÉM	4
352450	JACI	4
352570	JOSÉ BONIFÁCIO	4
352950	MENDONÇA	4
353030	MIRASSOL	4
353040	MIRASSOLÂNDIA	4
353280	NOVA ALIANÇA	4
353300	NOVA GRANADA	4
353390	OLÍMPIA	4
353400	ONDA VERDE	4
353420	ORINDIÚVA	4
353500	PALESTINA	4
353660	PAULO DE FARIA	4
353960	PLANALTO	4
354080	POTIRENDABA	4
354980	SÃO JOSÉ DO RIO PRETO	4
355340	TANABI	4
355535	UBARANA	4
355560	UCHOA	4
355715	ZACARIAS	4
		Catanduva
350370	ARIRANHA	5
350930	CAJOBI	5
351110	CATANDUVA	5
351120	CATIGUÁ	5
351492	ELISIARIO	5
351495	EMBAÚBA	5
353325	NOVAIS	5
353510	PALMARES PAULISTA	5

353570	PARAÍSO	5
353810	PINDORAMA	5
354560	SANTA ADÉLIA	5
355190	SEVERÍNIA	5
355260	TABAPUÃ	5
		Auriflama
350420	AURIFLAMA	6
351590	FLOREAL	6
351680	GASTÃO VIDIGAL	6
351690	GENERAL SALGADO	6
351890	GUZOLÂNDIA	6
352830	MAGDA	6
353330	NOVA LUZITÂNIA	6
354925	SÃO JOÃO DE IRACEMA	6
		Nhandeara
352810	MACAUBAL	7
353100	MONÇÕES	7
353140	MONTE APRAZÍVEL	7
353250	NEVES PAULISTA	7
353260	NHANDEARA	7
353270	NIPOÃ	7
353990	POLONI	7
355130	SEBASTIANÓPOLIS DO SUL	7
355570	UNIÃO PAULISTA	7
		Novo Horizonte
352150	IRAPUÃ	8
352190	ITAJOBÍ	8
352885	MARAPOANA	8
353350	NOVO HORIZONTE	8
354480	SALES	8
355600	URUPÊS	8
		Barretos
350550	BARRETOS	9
351200	COLINA	9
351210	COLÔMBIA	9
		São Joaquim da Barra
351740	GUAÍRA	10
352130	IPUÃ	10
352420	JABORANDI	10
352970	MIGUELÓPOLIS	10
353190	MORRO AGUDO	10
353360	NUPORANGA	10

353430	ORLÂNDIA	10
354490	SALES OLIVEIRA	10
354940	SÃO JOAQUIM DA BARRA	10
		Ituverava
350300	ARAMINA	11
350820	BURITIZAL	11
351770	GUARÁ	11
352010	IGARAPAVA	11
352410	ITUVERAVA	11
		Franca
351320	CRISTAIS PAULISTA	12
351620	FRANCA	12
352370	ITIRAPUÃ	12
352540	JERIQUARA	12
353630	PATROCÍNIO PAULISTA	12
353700	PEDREGULHO	12
354270	RESTINGA	12
354310	RIBEIRÃO CORRENTE	12
354360	RIFAINA	12
354950	SÃO JOSÉ DA BELA VISTA	12
		Jaboticabal
350610	BEBEDOURO	13
351010	CÂNDIDO RODRIGUES	13
351560	FERNANDO PRESTES	13
351860	GUARIBA	13
352430	JABOTICABAL	13
353130	MONTE ALTO	13
353150	MONTE AZUL PAULISTA	13
353900	PIRANGI	13
353950	PITANGUEIRAS	13
354650	SANTA ERNESTINA	13
355310	TAIAÇU	13
355320	TAIÚVA	13
355370	TAQUARITINGA	13
355440	TERRA ROXA	13
355680	VIRADOURO	13
355690	VISTA ALEGRE DO ALTO	13
		Ribeirão Preto
350560	BARRINHA	14
350780	BRODÓSQUI	14
351310	CRAVINHOS	14
351460	DUMONT	14

351885	GUATAPARÁ	14
352510	JARDINÓPOLIS	14
352760	LUIS ANTÔNIO	14
354020	PONTAL	14
354090	PRADÓPOLIS	14
354340	RIBEIRÃO PRETO	14
354750	SANTA RITA DO PASSA QUATRO	14
354760	SANTA ROSA DE VITERBO	14
355090	SÃO SIMÃO	14
355140	SERRA AZUL	14
355150	SERRANA	14
355170	SERTÃOZINHO	14
		Batatais
350100	ALTINÓPOLIS	15
350590	BATATAIS	15
350940	CAJURU	15
351090	CÁSSIA DOS COQUEIROS	15
354790	SANTO ANTÔNIO DA ALEGRIA	15
		Andradina
350210	ANDRADINA	16
351100	CASTILHO	16
351780	GUARAÇAI	16
352044	ILHA SOLTEIRA	16
352300	ITAPURA	16
353010	MIRANDÓPOLIS	16
353210	MURUTINGA DO SUL	16
353320	NOVA INDEPENDÊNCIA	16
353740	PEREIRA BARRETO	16
355230	SUD MENUCCI	16
355255	SUZANÁPOLIS	16
		Araçatuba
350280	ARAÇATUBA	17
350620	BENTO DE ABREU	17
351820	GUARARAPES	17
352650	LAVÍNIA	17
354440	RUBIÁCEA	17
354805	SANTO ANTÔNIO DO ARACANGUÁ	17
355630	VALPARAÍSO	17
		Birigui
350110	ALTO ALEGRE	18
350440	AVANHANDAVA	18

350510	BARBOSA	18
350640	BILAC	18
350650	BIRIGUI	18
350770	BRAÚNA	18
350810	BURITAMA	18
351190	CLEMENTINA	18
351250	COROADOS	18
351650	GABRIEL MONTEIRO	18
351710	GLICÉRIO	18
352725	LOURDES	18
352770	LUIZIÂNIA	18
353730	PENÁPOLIS	18
353770	PIACATU	18
354840	SANTÓPOLIS DO AGUAPEI	18
355520	TURIÚBA	18
		Lins
350880	CAFELÂNDIA	19
351700	GETULINA	19
351720	GUAÍÇARA	19
351730	GUAIMBÉ	19
352580	JÚLIO MESQUITA	19
352710	LINS	19
354160	PROMISSÃO	19
354460	SABINO	19
		Bauru
350070	AGUDOS	20
350340	AREALVA	20
350360	AREIÓPOLIS	20
350430	AVAÍ	20
350470	BALBINOS	20
350600	BAURU	20
350745	BOREBI	20
350830	CABRÁLIA PAULISTA	20
351450	DUARTINA	20
351810	GUARANTÃ	20
351910	IACANGA	20
352680	LENÇÓIS PAULISTA	20
352750	LUCIANÓPOLIS	20
353890	PIRAJUI	20
353940	PIRATININGA	20
354010	PONGAÍ	20
354110	PRESIDENTE ALVES	20

354250	REGINÓPOLIS	20
355550	UBIRAJARA	20
355590	URU	20
		Jaú
350520	BARIRI	21
350530	BARRA BONITA	21
350680	BOCAINA	21
350730	BORACÉIA	21
351410	DOIS CÓRREGOS	21
352000	IGARAÇU DO TIETÊ	21
352200	ITAJU	21
352290	ITAPUÍ	21
352530	JAÚ	21
352800	MACATUBA	21
352980	MINEIROS DO TIETÊ	21
353670	PEDERNEIRAS	21
		Avaré
350055	ÁGUAS DE SANTA BÁRBARA	22
350310	ARANDU	22
350450	AVARÉ	22
351140	CERQUEIRA CESAR	22
351925	IARAS	22
352180	ITAI	22
352350	ITATINGA	22
353580	PARANAPANEMA	22
		Botucatu
350230	ANHEMBI	23
350690	BOFETE	23
350750	BOTUCATU	23
351230	CONCHAS	23
353610	PARDINHO	23
355010	SÃO MANUEL	23
		Araraquara
350170	AMÉRICO BRASILIENSE	24
350320	ARARAQUARA	24
350670	BOA ESPERANÇA DO SUL	24
350740	BORBOREMA	24
351400	DOBRADA	24
351960	IBITINGA	24
352270	ITÁPOLIS	24
352930	MATÃO	24
353205	MOTUCA	24

353290	NOVA EUROPA	24
354370	RINCÃO	24
354690	SANTA LÚCIA	24
355270	TABATINGA	24
		São Carlos
350200	ANALÂNDIA	25
351370	DESCALVADO	25
351430	DOURADO	25
351930	IBATÉ	25
354290	RIBEIRÃO BONITO	25
354890	SÃO CARLOS	25
		Rio Claro
350790	BROTAS	26
351270	CORUMBATAÍ	26
352110	IPEÚNA	26
352360	ITIRAPINA	26
354390	RIO CLARO	26
355470	TORRINHA	26
		Limeira
350330	ARARAS	27
351220	CONCHAL	27
351240	CORDEIRÓPOLIS	27
352140	IRACEMÁPOLIS	27
352670	LEME	27
352690	LIMEIRA	27
354620	SANTA CRUZ DA CONCEIÇÃO	27
354670	SANTA GERTRUDES	27
		Piracicaba
350060	ÁGUAS DE SÃO PEDRO	28
351040	CAPIVARI	28
351170	CHARQUEADA	28
353090	MOMBUCA	28
353870	PIRACICABA	28
354210	RAFARD	28
354400	RIO DAS PEDRAS	28
354515	SALTINHO	28
354700	SANTA MARIA DA SERRA	28
355040	SÃO PEDRO	28
355450	TIETÊ	28
		Pirassununga
350030	AGUAÍ	29
353930	PIRASSUNUNGA	29

354070	PORTO FERREIRA	29
354630	SANTA CRUZ DAS PALMEIRAS	29
		São João da Boa Vista
350040	ÁGUAS DA PRATA	30
350870	CACONDE	30
351080	CASA BRANCA	30
351390	DIVINOLÂNDIA	30
351518	ESPÍRITO SANTO DO PINHAL	30
352380	ITOBI	30
353050	MOCOCA	30
354810	SANTO ANTÔNIO DO JARDIM	30
354910	SÃO JOÃO DA BOA VISTA	30
354970	SÃO JOSÉ DO RIO PARDO	30
355080	SÃO SEBASTIÃO DA GRAMA	30
355330	TAMBAÚ	30
355360	TAPIRATIBA	30
355640	VARGEM GRANDE DO SUL	30
		Mogi-Mirim
350380	ARTUR NOGUEIRA	31
351515	ENG. COELHO	31
355730	ESTIVA GERBI	31
352260	ITAPIRA	31
353070	MOGI-GUAÇU	31
353080	MOGI-MIRIM	31
354800	SANTO ANTÔNIO DE POSSE	31
		Campinas
350160	AMERICANA	32
350950	CAMPINAS	32
351280	COSMÓPOLIS	32
351490	ELIAS FAUSTO	32
351905	HOLAMBRA	32
351907	HORTOLÂNDIA	32
352050	INDAIATUBA	32
352470	JAGUARIÚNA	32
353180	MONTE MOR	32
353340	NOVA ODESSA	32
353650	PAULÍNIA	32
353710	PEDREIRA	32
354580	SANTA BARBARA D'OESTE	32
355240	SUMARÉ	32
355620	VALINHOS	32
355670	VINHEDO	32

		Amparo
350050	ÁGUAS DE LINDÓIA	33
350190	AMPARO	33
352700	LINDÓIA	33
353120	MONTE ALEGRE DO SUL	33
353680	PEDRA BELA	33
353820	PINHALZINHO	33
355160	SERRA NEGRA	33
355210	SOCORRO	33
		Dracena
351440	DRACENA	34
352600	JUNQUEIRÓPOLIS	34
353160	MONTE CASTELO	34
353310	NOVA GUATAPORANGA	34
353480	OURO VERDE	34
353540	PANORAMA	34
353640	PAULICÉIA	34
354710	SANTA MERCEDES	34
354930	SÃO JOÃO DO PAU D'ALHO	34
355510	TUPI PAULISTA	34
		Adamantina
350010	ADAMANTINA	35
351580	FLORA RICA	35
351600	FLÓRIDA PAULISTA	35
352080	INÚBIA PAULISTA	35
352160	IRAPURU	35
352740	LUCÉLIA	35
352890	MARIÁPOLIS	35
353460	OSVALDO CRUZ	35
353490	PACAEMBU	35
353600	PARAPUÃ	35
354380	RINÓPOLIS	35
354470	SAGRES	35
354510	SALMOURÃO	35
		Presidente Prudente
350080	ALFREDO MARCONDES	36
350130	ÁLVARES MACHADO	36
350240	ANHUMAS	36
350890	CAIABU	36
350910	CAIUA	36
351512	EMILIANÓPOLIS	36
351530	ESTRELA DO NORTE	36

351535	EUCLIDES DA CUNHA PAULISTA	36
352060	INDIANA	36
352560	JOÃO RAMALHO	36
352870	MARABÁ PAULISTA	36
352920	MARTINÓPOLIS	36
353020	MIRANTE DO PARANAPANEMA	36
353220	NARANDIBA	36
353830	PIQUEROBI	36
353920	PIRAPOZINHO	36
354120	PRESIDENTE BERNARDES	36
354130	PRESIDENTE EPITÁCIO	36
354140	PRESIDENTE PRUDENTE	36
354150	PRESIDENTE VENCESLAU	36
354220	RANCHARIA	36
354240	REGENTE FEIJÓ	36
354425	ROSANA	36
354550	SANDOVALINA	36
354770	SANTO ANASTÁCIO	36
354830	SANTO EXPEDITO	36
355290	TACIBA	36
355390	TARABAI	36
355430	TEODORO SAMPAIO	36
		Tupã
350580	BASTOS	37
351900	HERCULÂNDIA	37
351920	IACRI	37
354180	QUEIRÓZ	37
354200	QUINTANA	37
355500	TUPÃ	37
		Marília
350140	ÁLVARO DE CARVALHO	38
350150	ALVINLÂNDIA	38
351470	ECHAPORÁ	38
351660	GÁLIA	38
351670	GARÇA	38
352780	LUPÉRCIO	38
352900	MARÍLIA	38
353370	OCAUÇU	38
353410	ORIENTE	38
353450	OSCAR BRESSANE	38
354000	POMPÉIA	38
355660	VERA CRUZ	38

		Assis
350400	ASSIS	39
350720	BORÁ	39
350980	CAMPOS NOVOS PAULISTA	39
351000	CÂNDIDO MOTA	39
351330	CRUZÁLIA	39
351610	FLORÍNIA	39
351950	IBIRAREMA	39
351990	IEPÉ	39
352790	LUTÉCIA	39
352880	MARACAÍ	39
353530	PALMITAL	39
353550	PARAGUAÇU PAULISTA	39
353715	PEDRINHAS PAULISTA	39
353970	PLATINA	39
354170	QUATÁ	39
355395	TARUMA	39
		Ourinhos
350630	BERNARDINO DE CAMPOS	40
351015	CANITAR	40
355720	CHAVANTES	40
351519	ESPIRITO SANTO DO TURVO	40
351540	FARTURA	40
352090	IPAUCU	40
352860	MANDURI	40
353380	ÓLEO	40
353470	OURINHOS	40
353880	PIRAJU	40
354320	RIBEIRÃO DO SUL	40
354540	SALTO GRANDE	40
354640	SANTA CRUZ DO RIO PARDO	40
355050	SÃO PEDRO DO TURVO	40
355120	SARUTAIÁ	40
355300	TAGUAI	40
355420	TEJUPÁ	40
355460	TIMBURI	40
		Itapeva
350500	BARÃO DE ANTONINA	41
350715	BOM SUCESSO DE ITARARÉ	41
350800	BURI	41
351260	CORONEL MACEDO	41
352170	ITABERÁ	41

352240	ITAPEVA	41
352280	ITAPORANGA	41
352320	ITARARÉ	41
353282	NOVA CAMPINA	41
354350	RIVERSUL	41
355380	TAQUARITUBA	41
355385	TAQUARIVAÍ	41
		Itapetininga
350075	ALAMBARI	42
350220	ANGATUBA	42
350945	CAMPINA DO MONTE ALEGRE	42
351850	GUAREI	42
352230	ITAPETININGA	42
		Tatui
350700	BOITUVA	43
351150	CERQUILHO	43
351160	CESÁRIO LANGE	43
352640	LARANJAL PAULISTA	43
353750	PEREIRAS	43
354050	PORANGABA	43
355400	TATUI	43
355465	TORRE DE PEDRA	43
		Capão Bonito
350270	APIAI	44
350535	BARRA DO CHAPÉU	44
351020	CAPÃO BONITO	44
351760	GUAPIARA	44
352120	IPORANGA	44
352215	ITAOCA	44
352265	ITAPIRAPUA PAULISTA	44
354280	RIBEIRA	44
354300	RIBEIRÃO BRANCO	44
354325	RIBEIRÃO GRANDE	44
		Piedade
351970	IBIÚNA	45
353780	PIEDADE	45
353790	PILAR DO SUL	45
355020	SÃO MIGUEL ARCANJO	45
355350	TAPIRAI	45
		Sorocaba
350115	ALUMÍNIO	46
350275	ARAÇARIGUAMA	46

350290	ARAÇOIABA DA SERRA	46
350840	CABREÚVA	46
351030	CAPELA DO ALTO	46
352100	IPERÓ	46
352390	ITU	46
352840	MAIRINQUE	46
354060	PORTO FELIZ	46
354520	SALTO	46
354530	SALTO DE PIRAPORA	46
355060	SÃO ROQUE	46
355110	SARAPUI	46
355220	SOROCABA	46
355700	VOTORANTIM	46
		Jundiaí
350960	CAMPO LIMPO PAULISTA	47
352400	ITUPEVA	47
352590	JUNDIAI	47
352730	LOUVEIRA	47
355650	VÁRZEA PAULISTA	47
		Bragança Paulista
350410	ATIBAIA	48
350710	BOM JESUS DOS PERDÕES	48
350760	BRAGANÇA PAULISTA	48
352340	ITATIBA	48
352520	JARINU	48
352550	JOANÓPOLIS	48
353200	MORUNGABA	48
353240	NAZARÉ PAULISTA	48
353860	PIRACAIA	48
355495	TUIUTI	48
355635	VARGEM	48
		Campos do Jordão
350970	CAMPOS DO JORDÃO	49
353170	MONTEIRO LOBATO	49
354820	SANTO ANTÔNIO DO PINHAL	49
354860	SÃO BENTO DO SAPUCAI	49
		São José dos Campos
350850	CAÇAPAVA	50
352020	IGARATA	50
352440	JACAREI	50
353800	PINDAMONHANGABA	50
354600	SANTA BRANCA	50

354990	SÃO JOSÉ DOS CAMPOS	50
355410	TAUBATÉ	50
355480	TREMEMBÉ	50
		Guaratinguetá
350250	APARECIDA	51
350860	CACHOEIRA PAULISTA	51
351340	CRUZEIRO	51
351840	GUARATINGUETÁ	51
352660	LAVRINHAS	51
352720	LORENA	51
353850	PIQUETE	51
354075	POTIM	51
354190	QUELUZ	51
354430	ROSEIRA	51
		Bananal
350315	ARAPEI	52
350350	AREIAS	52
350490	BANANAL	52
354960	SÃO JOSÉ DO BARREIRO	52
355200	SILVEIRAS	52
		Paraibuna/Paratinga
351360	CUNHA	53
352490	JAMBEIRO	53
352630	LAGOINHA	53
353230	NATIVIDADE DA SERRA	53
353560	PARAIBUNA	53
354230	REDENÇÃO DA SERRA	53
355000	SÃO LUIS DO PARAITINGA	53
		Caraguatatuba
351050	CARAGUATATUBA	54
352040	ILHABELA	54
355070	SÃO SEBASTIÃO	54
355540	UBATUBA	54
		Registro
350540	BARRA DO TURVO	55
350925	CAJATI	55
350990	CANANÉIA	55
351480	ELDORADO	55
352030	IGUAPE	55
352042	ILHA COMPRIDA	55
352460	JACUPIRANGA	55
352610	JUQUIÁ	55

352990	MIRACATU	55
353620	PARIQUERA-AÇU	55
354260	REGISTRO	55
355180	SETE BARRAS	55
		Itanhaém
352210	ITANHAÉM	56
352330	ITARIRI	56
353110	MONGAGUÁ	56
353720	PEDRO DE TOLEDO	56
353760	PERUÍBE	56
		Osasco
350570	BARUERI	57
350920	CAJAMAR	57
351060	CARAPICUIBA	57
352250	ITAPEVI	57
352500	JANDIRA	57
353440	OSASCO	57
353910	PIRAPORA DO BOM JESUS	57
354730	SANTANA DE PARNAÍBA	57
		Franco da Rocha
350900	CAIEIRAS	58
351630	FRANCISCO MORATO	58
351640	FRANCO DA ROCHA	58
352850	MAIRIPORÃ	58
		Guarulhos
350390	ARUJÁ	59
351880	GUARULHOS	59
354680	SANTA ISABEL	59
		Itapecerica da Serra
351300	COTIA	60
351500	EMBU	60
351510	EMBU-GUAÇU	60
352220	ITAPECERICA DA SERRA	60
352620	JUQUITIBA	60
354995	SÃO LOURENÇO DA SERRA	60
355280	TABOÃO DA SERRA	60
355645	VARGEM GRANDE PAULISTA	60
		São Paulo
351380	DIADEMA	61
352940	MAUÁ	61
354330	RIBEIRÃO PIRES	61
354410	RIO GRANDE DA SERRA	61

354780	SANTO ANDRÉ	61
354870	SÃO BERNARDO DO CAMPO	61
354880	SÃO CAETANO DO SUL	61
355030	SÃO PAULO	61
		Mogi das Cruzes
350660	BIRITIBA-MIRIM	62
351570	FERRAZ DE VASCONCELOS	62
351830	GUARAREMA	62
352310	ITAQUAQUECETUBA	62
353060	MOGI DAS CRUZES	62
353980	POÁ	62
354500	SALESÓPOLIS	62
355250	SUZANO	62
		Santos
350635	BERTIOGA	63
351350	CUBATÃO	63
351870	GUARUJÁ	63
354100	PRAIA GRANDE	63
354850	SANTOS	63
355100	SÃO VICENTE	63

4. VALORES EM M² DE LOJA E DENSIDADE DE RENDA

	Micro-Região	Dados Originais	
		Dens. de Loja	Dens. de Renda
1	Jales	1.320	937,346
2	Fernandópolis	4.383	957,635
3	Votuporanga	3.911	999,079
4	São José do Rio Preto	8.698	2,202,926
5	Catanduva	8.327	2,639,976
6	Auriflama	0.479	491,700
7	Nhandeara	2.550	787,408
8	Novo Horizonte	0.573	713,346
9	Barretos	4.848	1,517,400
10	São Joaquim da Barra	0.831	989,008
11	Ituverava	0.599	1,233,105
12	Franca	4.248	3,008,154
13	Jaboticabal	6.285	2,610,358
14	Ribeirão Preto	16.497	5,921,067
15	Batatais	3.019	872,569
16	Andradina	1.078	754,050
17	Araçatuba	2.306	1,495,202
18	Birigui	3.465	1,281,445
19	Lins	2.685	1,035,142
20	Bauru	19.656	2,090,740
21	Jaú	4.375	2,361,022
22	Avaré	0.399	623,017
23	Botucatu	1.493	1,365,526
24	Araraquara	5.544	2,346,929
25	São Carlos	5.480	2,832,044
26	Rio Claro	10.664	2,426,610
27	Limeira	17.100	6,575,064
28	Piracicaba	25.061	4,839,494
29	Pirassununga	8.236	2,794,474
30	São João da Boa Vista	2.795	1,957,943
31	Mogi-Mirim	16.451	4,201,458
32	Campinas	67.004	29,888,216
33	Amparo	7.700	2,793,187
34	Dracena	5.728	1,019,530
35	Adamantina	3.938	1,032,048
36	Presidente Prudente	1.795	888,383
37	Tupã	1.898	1,397,655
38	Marília	5.200	1,914,697

39	Assis	2.769	935,837
40	Ourinhos	1.779	1,207,814
41	Itapeva	0.404	602,837
42	Itapetininga	1.832	1,116,606
43	Tatui	1.752	2,276,145
44	Capão Bonito	0.499	369,148
45	Piedade	1.195	867,162
46	Sorocaba	16.421	8,137,669
47	Jundiai	54.300	24,149,047
48	Bragança Paulista	7.199	4,021,610
49	Campos do Jordão	0.685	1,642,102
50	São José dos Campos	25.462	11,104,525
51	Guaratinguetá	7.558	4,136,185
52	Bananal	0.109	189,758
53	Paraibuna/Paratinga	0.079	289,601
54	Caraguatatuba	6.473	2,527,551
55	Registro	2.043	442,894
56	Itanhaém	15.004	1,854,812
57	Osasco	275.867	71,088,606
58	Franco da Rocha	11.000	12,641,540
59	Guarulhos	75.353	43,913,697
60	Itapeçerica da Serra	14.194	14,941,700
61	São Paulo	543.000	324,190,288
62	Mogi das Cruzes	24.677	12,841,034
63	Santos	98.673	39,810,717

5. PROGRAMA PARA SIMULAÇÃO DA DISTRIBUIÇÃO DE GR

```
%montec
%Espera encontrar no workspace o vetor x de dados
%Gera uma Matriz 63x100 em que cada
%coluna é uma permutação dos x
%Em seguida, calcula o GR para cada coluna e apresenta os resultados
fim=100 %fim é o # de simulações
n=63 %n é o # de obs. no vetor x
cdavez= input('Informe a matriz de conectividade: ');
for j=1:fim
    temp=randperm(n);
    MC(:,j)=temp;
end
for j=1:fim
    for i=1:n
        MCx(i,j)=x(MC(i,j));
    end
end
for j=1:fim
    global gratio
    gr2(MCx(:,j),cdavez);
    GRMC(j)=gratio;
end
GRMC'
```

6. SIMULAÇÃO DE GRS PARA DISTRIBUIÇÃO

Simulação	c0	c1	c2	c3	c4	c5	c6	c7	c8
1	0.9517	0.9491	1.004	1.0947	1.0182	1.1567	1.1197	0.7669	0.4875
2	0.8670	0.9000	1.029	0.9744	1.0065	1.1261	1.2922	1.2255	1.8622
3	1.0244	1.0848	1.000	1.0068	0.9160	0.9031	1.3200	0.4891	1.1198
4	0.9708	1.1044	1.048	1.0801	1.0269	1.1391	0.9669	1.2313	2.3544
5	0.9607	0.9788	0.996	0.8116	0.9379	1.0718	1.1288	1.0752	0.9966
6	0.9498	0.8096	0.908	1.0142	0.9611	0.9113	0.9917	1.4334	1.4250
7	1.0511	1.0147	0.959	0.8286	1.0781	1.0275	0.9518	0.8012	1.3090
8	0.9930	1.1344	0.935	0.8993	1.0049	1.0070	0.7632	1.4222	1.3891
9	1.0849	1.0538	0.997	1.0185	0.9773	0.9294	1.2585	1.6813	1.3097
10	0.9934	0.7965	1.029	1.0704	0.8773	1.0362	0.8407	0.9174	1.2076
11	1.1647	1.0664	0.962	0.9913	0.9727	1.0409	1.0455	1.0785	1.7340
12	0.7850	0.9434	1.209	1.0154	0.8870	0.8460	0.7470	1.2257	0.5797
13	0.9587	1.1355	1.083	0.8709	1.0744	1.0297	1.0999	1.0565	0.4990
14	0.9101	0.9910	1.001	1.0341	0.9510	1.0216	1.2554	1.5865	0.5720
15	1.0847	1.1457	0.914	1.0074	1.1227	0.8829	0.5668	1.1539	0.5990
16	0.9572	1.0055	0.911	1.0505	0.9262	1.1235	1.2201	1.6293	2.1608
17	1.1289	0.9321	0.999	1.0575	1.0931	0.7095	0.9940	1.6558	0.9516
18	0.9551	0.9576	0.844	0.9199	1.0142	0.9433	0.9550	0.9236	0.7210
19	1.1330	0.9878	1.028	1.0102	1.1136	0.9831	0.8266	0.6099	1.0728
20	1.0496	0.9827	1.073	0.9318	0.9215	0.8747	1.1516	0.9100	1.1041
21	1.1407	1.0012	1.115	0.7867	0.8796	0.9404	0.9346	1.1987	0.3493
22	0.9515	1.1229	0.849	1.1137	0.9994	0.9459	1.1017	0.7154	0.5457
23	0.8870	0.9424	1.027	1.0689	1.0124	0.8913	0.8065	0.9916	1.1143
24	0.9638	1.0358	1.040	0.9122	0.9380	1.0655	0.9901	0.4633	0.4937
25	1.1715	0.8660	1.018	0.8929	0.9962	0.8933	0.9480	0.6484	0.8323
26	0.9294	1.1539	0.929	1.0125	1.1137	0.8426	0.7224	0.9950	1.4429
27	0.8527	1.0647	0.898	1.1035	0.9415	1.0058	0.9737	0.6493	0.8377
28	1.0499	1.0100	0.848	0.9464	1.1551	1.0698	1.2191	1.2940	1.7090
29	0.9472	1.0104	0.989	0.9634	1.0553	0.8924	1.2754	1.2597	1.2464
30	0.9849	1.2399	1.069	0.9092	0.9891	1.0359	1.2856	0.5002	0.5742
31	0.8929	1.2492	1.051	1.1624	0.9011	0.9103	0.8875	1.5525	0.5804
32	1.0651	0.9004	1.075	1.0434	0.8755	1.0078	1.2047	1.4617	0.5388
33	1.1162	1.0838	0.905	1.0949	0.9944	1.0034	1.4554	1.0112	0.6315
34	0.9754	1.1162	1.023	0.9595	0.9417	0.9981	1.4477	0.8023	1.3712
35	1.0402	0.7940	0.978	0.9727	1.0050	1.1038	1.0066	1.0557	0.8847
36	1.0983	0.8959	1.113	0.9415	1.0063	1.0242	1.0949	0.5293	0.6606
37	1.1368	0.8274	0.985	0.9130	1.2388	1.1353	0.9044	0.5917	0.6343
38	1.0028	1.1185	0.947	1.0953	0.9804	0.8731	1.2243	0.8625	1.0796
39	1.0106	1.0116	0.999	0.8755	0.9548	0.9644	0.9079	1.7128	1.4422

40	0.9377	0.9343	1.061	1.0095	1.1782	1.0678	1.3982	0.6636	0.6293
41	1.0504	0.9110	1.007	1.1202	1.0831	0.9255	0.7459	2.1164	1.6700
42	0.9491	1.0599	0.860	0.8704	0.8832	1.1534	1.0857	1.4727	1.4476
43	1.1550	1.0000	1.027	1.1770	1.1226	1.1248	0.8209	1.0389	0.4303
44	1.0569	0.8946	0.895	1.0129	1.0495	0.9483	0.8438	1.0552	1.0357
45	1.0042	1.1058	1.003	1.0198	0.9954	1.1605	0.5406	1.0745	0.1923
46	0.9837	1.0041	1.100	0.8752	0.8725	1.0073	1.2016	1.0055	0.3650
47	0.9330	0.8960	0.990	0.9123	1.0849	0.8442	0.7521	1.0950	0.4299
48	0.9935	0.9106	1.018	0.9639	0.9170	0.9973	0.7808	1.1312	1.9005
49	1.1062	0.9105	1.075	0.8483	0.9919	1.0674	1.0357	0.8421	0.7506
50	0.9487	0.8757	1.006	1.0814	1.0305	1.0195	0.7088	0.9004	0.8012
51	1.0733	1.2575	0.972	0.9992	0.9591	1.2482	1.3350	1.6335	1.1817
52	1.0067	1.0713	0.933	1.1160	1.0162	0.9566	1.1461	1.8086	0.4867
53	1.0605	0.9354	1.023	0.9068	0.9044	1.0881	0.5660	0.9386	0.4284
54	0.9851	0.9482	1.067	1.0125	1.0065	0.8842	1.0464	0.8189	0.8381
55	1.0099	1.1890	1.108	1.0068	0.9857	1.0177	1.0754	0.6620	0.9182
56	0.8694	1.1836	0.987	1.0256	1.0332	0.7605	0.9693	1.3242	0.7699
57	1.1457	0.9085	1.049	0.9702	1.0188	0.8952	0.6943	1.0673	1.1071
58	1.0096	1.3231	1.151	0.9052	0.9890	0.9168	0.9567	1.8151	0.5835
59	0.9301	1.2414	0.862	1.0926	0.9451	0.9782	1.1691	1.0236	1.2398
60	0.9838	0.7717	1.053	0.9473	1.1157	0.9678	0.7985	0.5187	0.9249
61	0.9954	1.0162	0.900	1.0792	0.9667	1.1151	1.1292	1.6321	0.8101
62	1.1000	1.0397	0.996	1.0958	1.0793	0.8568	0.4517	0.6700	1.0076
63	1.1101	1.0363	1.002	0.9395	1.0900	0.9341	1.3519	0.9413	0.8103
64	0.9183	1.1220	0.816	0.9842	1.0836	0.7473	0.8909	0.9365	0.7749
65	0.9691	0.9505	1.061	0.9763	0.9253	0.8978	1.0171	0.7903	0.5927
66	0.9231	0.7495	1.027	0.9278	0.8930	0.9906	0.8915	0.5514	0.7650
67	1.0017	1.1357	1.005	1.0181	0.9882	0.9666	0.6707	0.7139	0.5434
68	1.1150	0.9328	0.948	0.9966	0.9998	1.2684	1.4492	1.1127	0.6768
69	0.9722	0.9765	1.017	1.1538	0.9354	0.9375	1.2283	0.9053	2.4380
70	1.1643	0.8427	1.049	1.0322	1.0435	1.2055	0.9110	0.9845	0.7905
71	0.8494	0.9842	1.133	1.1642	1.1612	0.9279	1.0397	1.0305	0.8063
72	0.8662	1.0232	1.073	0.9797	0.8802	0.8756	1.0245	1.7958	1.1739
73	0.9663	0.8753	1.064	1.0543	0.9817	1.0979	1.0592	0.7568	1.0453
74	1.0908	1.0179	0.983	0.9367	1.0431	1.0197	0.6882	0.8820	0.5203
75	0.8812	1.1464	0.964	0.9597	0.8931	0.8784	1.2438	1.1831	1.2154
76	1.0786	0.9683	1.037	0.9640	0.8919	0.7962	1.4938	1.1072	1.5829
77	0.9524	1.0550	1.005	0.9229	1.0789	0.8643	0.8947	1.2657	1.2773
78	0.9078	1.1179	1.027	0.9983	1.0534	0.9440	1.4762	1.9202	1.8935
79	1.1108	1.1025	0.856	1.0853	1.0121	0.9448	0.7680	0.9711	1.0391
80	0.9241	1.2993	1.052	0.9920	1.0197	1.1645	0.9336	1.0291	0.5050
81	1.0552	0.8284	0.904	1.0990	0.9548	1.0489	1.0421	1.1641	0.8228

82	1.0245	0.7217	1.071	0.9376	0.9125	1.0501	0.6871	1.6060	0.5301
83	0.9778	1.0222	1.078	1.0272	0.9861	1.0011	1.2698	0.2579	0.5929
84	0.9695	0.8722	0.846	0.9827	1.0594	1.0360	1.1684	0.6399	1.5691
85	1.0959	0.9976	1.036	0.9962	1.0760	0.9943	1.3527	1.2000	1.1220
86	1.0102	0.8561	0.915	1.0558	1.0355	0.9731	0.8690	1.4393	1.9123
87	1.2664	0.9868	1.090	1.0124	0.9799	1.0860	0.9428	1.1800	1.1671
88	1.1477	1.2648	0.997	0.9579	0.9810	1.0498	0.9079	1.1087	2.0426
89	1.0838	0.8655	1.016	1.0031	0.9519	0.9730	1.2969	0.8188	0.6877
90	1.0414	0.9964	1.145	0.9215	1.0911	1.1035	1.2250	0.9354	0.6565
91	0.9541	0.7743	0.911	1.0057	0.9757	0.9261	0.9199	0.8927	0.5383
92	0.9370	0.9513	0.901	1.0911	0.9374	1.0537	0.7575	0.6384	0.3606
93	0.9996	0.8529	1.059	1.0380	1.0418	1.1289	0.9534	1.1446	2.4799
94	1.0402	1.0357	1.037	0.9378	0.9122	1.0872	0.9369	0.8235	1.0172
95	1.0081	0.8919	1.038	1.0131	0.9701	0.8163	1.2570	1.1329	0.6349
96	0.8937	0.8874	0.945	1.0365	0.9148	0.9730	0.9719	1.2049	1.5610
97	0.9291	1.0538	0.973	0.8566	0.9980	0.9070	0.8687	1.0023	0.2956
98	0.9599	0.7964	1.137	1.0276	0.8827	0.9165	0.9832	1.2742	0.6968
99	1.1092	0.8954	0.993	1.0094	0.9795	1.2322	1.1272	0.3920	0.9869
100	0.8478	0.8176	0.956	1.0097	0.9964	1.2367	1.0740	0.9298	0.5756
P(2,5)	0.8510	0.7729	0.8466	0.8380	0.8784	0.7775	0.5664	0.4756	0.3547
P(97,5)	1.1645	1.2613	1.1413	1.1583	1.1583	1.2346	1.4525	1.8120	2.2624

7. RESÍDUOS DOS MODELOS DE AUTOCORRELAÇÃO

Microrregião	Resíduo da dens. m2 loja	Resíduo da dens. de renda
1	0.0092	0.1553
2	0.5183	0.1272
3	0.1992	0.0558
4	0.4908	0.3099
5	0.3691	0.2244
6	(0.6830)	(0.2751)
7	(0.0063)	(0.1047)
8	(1.0578)	(0.4048)
9	0.0782	(0.0770)
10	(0.6441)	(0.2821)
11	(0.4805)	(0.1101)
12	0.5727	0.5037
13	0.1410	0.1048
14	0.6730	0.5349
15	(0.0413)	(0.4070)
16	(0.2534)	(0.0605)
17	0.1631	0.2860
18	0.1773	0.1023
19	(0.1247)	(0.1249)
20	1.0210	0.2421
21	(0.2038)	0.0183
22	(0.6985)	(0.2438)
23	(0.3664)	(0.0918)
24	0.0211	0.0342
25	(0.1690)	(0.0380)
26	0.0825	(0.1593)
27	0.1138	0.1782
28	0.5424	0.0864
29	0.0011	(0.1075)
30	(0.4560)	(0.1170)
31	0.2156	(0.0338)
32	0.6758	0.7329
33	(0.3388)	(0.4155)
34	0.4809	0.0958
35	0.2457	0.0046
36	(0.2347)	(0.0516)
37	(0.1894)	0.1237
38	0.2083	0.2103

39	0.0893	(0.1083)
40	(0.0628)	0.0868
41	(0.3489)	(0.0591)
42	0.3571	0.1531
43	(0.4755)	(0.0505)
44	(0.3627)	(0.2524)
45	(0.4199)	(0.2952)
46	0.0204	(0.0302)
47	0.3433	0.2353
48	(0.5208)	(0.5369)
49	(1.2413)	(0.5765)
50	0.8186	0.4751
51	1.0727	0.6508
52	(0.8583)	(0.7264)
53	(1.7663)	(0.9913)
54	0.0933	(0.2823)
55	(0.1856)	(0.5550)
56	(0.2828)	(0.9232)
57	0.9554	0.4782
58	(0.7769)	(0.4532)
59	0.4760	0.4100
60	(0.0976)	0.3593
61	1.2436	1.3027
62	0.1792	0.0387
63	0.5533	0.5679