

# INDICADORES DE AUTOCORRELAÇÃO LOCAL

## EXEMPLO DO MUNICÍPIO DE SÃO PAULO

### INTRODUÇÃO

O ferramental de análise espacial em unidades de áreas é muito importante dentro do universo de possibilidades já disponíveis em diversos Sistemas de Informação Geográfica. Divisão geopolítica, setores censitários, unidades de paisagem, bairros são casos típicos de elementos geográficos representados por áreas, ou polígonos, onde para cada polígono há um valor do atributo associado. Este tipo de representação é também conhecida como mapa coroplético. Nestes mapas, a variação do atributo no espaço não é representada por uma superfície contínua, mas apresenta variações abruptas de valor. Apesar de limitadas às características de cada polígono (tamanho, forma, homogeneidade), muitas inferências sobre padrões ou comportamentos espaciais dos diversos atributos podem ser realizadas.

As técnicas de estatística espacial de áreas foram desenvolvidas para tentar identificar regiões onde a distribuição dos valores possa apresentar um padrão específico associado a sua localização geográfica. A informação que se busca é quanto sou parecido com meu vizinho próximo e sou diferente do meu vizinho distante. Neste trabalho, foram utilizadas duas técnicas de estatística espacial que nos fornecem indicadores locais de autocorrelação. Cada polígono, ou unidade de área, terá um determinado valor de dependência espacial. O indicador de associação espacial local Moran (I) (Anselin, 1994) explora o grau de dependência espacial a partir de uma estimativa de segunda ordem, ou seja, um tipo de covariância espacial entre os polígonos. A segunda estatística utilizada neste trabalho é conhecida como indicadores Gi e Gi\* (Getis;Ord, 1992) e se baseia em estimativas de primeira ordem, como um tipo de estimador de médias móveis.

A idéia da utilização destes indicadores de autocorrelação local para realização do trabalho de curso surgiu quando tivemos acesso ao Mapa da Exclusão/Inclusão Social da Cidade de São Paulo, trabalho coordenado pela Prof. Aldaíza Sposati da PUC-SP. Este material possui uma série de indicadores e resultados de pesquisas espacializados nos 96 distritos que compõem o município de São Paulo. Além de ser um banco de dados geográficos extremamente rico, o Mapa da Exclusão/Inclusão trazia em sua gênese a tentativa de captação das discrepâncias e contradições sociais existentes dentro da estrutura urbana de São Paulo. Dessa forma, o trabalho se propôs a avaliar as potencialidades de aplicação destas duas ferramentas estatísticas dentro desta perspectiva.

## OBJETIVOS

Os objetivos principais deste trabalho foram:

Primeiro, estudar os métodos estatísticos Gi e Moran Local para detecção de padrões de forte autocorrelação local e buscar o entendimento do seu significado estatístico e de sua formulação matemática;

Discutir as semelhanças e diferenças entre estes dois indicadores a partir da análise comparativa dos resultados obtidos para os distritos do município de São Paulo sobre os dados presentes no Mapa da Exclusão/Inclusão de São Paulo;

E, por fim, observar e explorar possíveis padrões quanto seu significado na tentativa de mapear a exclusão social no município de São Paulo, uma vez que um dos objetivos do Mapa da Exclusão/Inclusão era captar as diferenças intra-urbanas entre os 96 distritos paulistanos.

## INDICADORES LOCAIS DE ASSOCIAÇÃO ESPACIAL

### MORAN LOCAL

O Índice de Moran Local foi proposto por Luc Anselin (1994) como uma ferramenta estatística para testar a autocorrelação local e para detectar objetos espaciais com influência no indicador Moran Global. Esta família estatística trabalha a partir da estimativa de segunda ordem do comportamento dos seus dados, em outras palavras, a partir da análise das covariâncias entre as diferentes unidades de área. Enquanto o Índice Global de Moran informa o nível de interdependência espacial entre todos os polígonos em estudo, o Índice Local de Moran avalia a covariância entre um determinado polígono e uma certa vizinhança definida em função de uma distância  $d$ .

Em análises de associação espacial, assumir uma estabilidade estrutural ou estacionariedade pode ser complicado, especialmente quando um grande número de observações são utilizadas, (Anselin, 1994), é possível então a utilização do Índice Local de Moran como uma ferramenta estatística que possibilita uma indicação sobre a extensão da significância de um “cluster” de iguais valores.

Anselin define o Índice Local de Moran como o produto do resíduo no polígono de referência com a média local dos resíduos dos seus vizinhos adjacentes. Desta maneira o Índice Local de Moran pode ser escrito como:

$$I_i = z_i \sum_j w_{ij} z_j$$

Onde:

$w_{ij}$  = valor na matriz de vizinhança para a região  $i$  com a região  $j$  em função da distância  $d$ , e onde  $z_i$  e  $z_j$  são os desvios em relação à média.

A matriz  $W_{ij}$  é a que define os vizinhos de um certo polígono. Esta matriz pode ser gerada de diferentes maneiras, ou seja, existem várias possibilidades de

determinação de quem são ou não vizinhos. Neste trabalho adotamos como critério de vizinhança a adjacência, assim, os valores correspondentes aos polígonos com borda em comum assumem valor 1 na matriz de vizinhança, por outro lado, entre os polígonos que não apresentam este tipo de relação o valor então é nulo.

Como por definição os parâmetros  $z_i$  e  $z_j$  representam desvios em relação à média, torna-se clara a semelhança entre o Índice de Moran e a fórmula da correlação entre duas variáveis:

$$Cor = E (x - \bar{x}) (y - \bar{y}) / E (x - \bar{x})^2$$

Dessa forma, o Índice de Moran Local permite duas importantes interpretações combinadas: a indicação de “pockets” de não estacionariedade espacial (indicação de “outliers”) e a possibilidade de teste da hipótese nula sobre a interdependência dos dados. A relação existente entre o Índice de Moran Local e o Global pode ser observada na fórmula que define o I global como a soma dos  $l_i$  locais.

$$\sum l_i = \sum z_i \sum w_{ij} z_j$$

e

$$I = \sum l_i / [S_0 (\sum z_i / n)]$$

A interpretação mais direta do Índice Local de Moran é aquela onde valores significativamente altos e positivos apontam a presença de um “cluster” tanto de valores iguais tanto altos como baixos, já valores significativamente baixos indicam um regime espacial de desigualdade na região, uma espécie de padrão “anti-cluster”, ou melhor, zonas de transição entre um determinado regime espacial e outro.

Anselin expõe em seu trabalho a formulação matemática que define os momentos de  $l_i$  ( $E[l_i]$  e  $Var[l_i]$ ) para a construção de intervalos de confiança para a aceitação da hipótese nula de total aleatoriedade espacial. A partir desta formulação e assumindo-se uma normalidade na distribuição destes índices torna-se possível avaliar a significância da observação de um regime espacial diferenciado.

## ESTATÍSTICA $G_i$ E $G_i^*$

A estatística  $G_i$  e  $G_i^*$  é uma ferramenta estatística apresentada por Getis e Ord como um indicador de associação espacial local. Apesar da estatística  $G$  se aproximar da estatística Moran enquanto objetivo, as informações que ambas apontam são conceitualmente diferentes. Enquanto o Índice Local de Moran é baseado na análise das covariância entre áreas o Índice  $G$  é na realidade uma somatória de valores vizinhos definidos a partir de uma matriz de vizinhança  $W_{ij}$ . (Getis; Ord, 1992)

Desta forma  $G_i$  e  $G_i^*$  podem ser escritos como:

$$G_i = \sum_j w_{ij}(d) x_j / \sum x_j \text{ para } i \neq j$$

e

$$G_i^* = \sum_j w_{ij}(d) x_j / \sum x_j \text{ para } i = j$$

Onde:  $W_{ij}$  é a matriz binária e simétrica que define a vizinhança entre as áreas.

A única diferença existente entre o indicador  $G_i$  e  $G_i^*$  é o fato do valor do polígono referência ser incluído no cálculo do  $G_i^*$  e não no cálculo do  $G_i$ . Neste trabalho, a matriz de vizinhança utilizada para o cálculo dos índices  $G_i$  e  $G_i^*$  foi a mesma utilizada para a definição do Índice Local de Moran, apenas áreas com bordas comuns foram consideradas vizinhas.

Diferentemente do Índice Local de Moran, a estatística  $G_i$  e  $G_i^*$  apresenta uma interpretação mais direta sobre como os dados estão distribuídos no espaço. Uma vez que estes indicadores são compostos por uma somatória de valores de atributos e não desvios em relação à média como no Moran, a observação de valores significativamente altos de  $G_i$  e  $G_i^*$  apontam a existência de altos índices de ocorrência deste atributo, sendo o oposto um indício de agrupamento de valores baixos.

Uma das comparações mais evidentes entre a estatística  $G$  e a Moran é a de que apesar de medirem coisas fundamentalmente diferentes, ambas supõe normalidade em suas distribuições para construir intervalos de confiança para a hipótese nula de total ausência de autocorrelação local.

Em circunstâncias típicas, a hipótese nula é a de que um certo quadro de valores  $x$  dentro de uma distância  $d$  é uma amostra aleatória independente de sua posição geográfica, então assumindo-se uma distribuição normal dos resultados dos  $G_i$  podemos construir intervalos de confiança para um determinado valor  $Z_i$ , onde :

$$Z_i = \{ G_i(d) - E [G_i(d)] \} / (\text{var } G_i(d))^{1/2}$$

A estatística  $G(d)$  mede a concentração total ou falta de concentração de todos os pares de  $(x_i, x_j)$  desde que  $j$  seja vizinho de  $i$ . O índice Moran, por outro lado mede a correlação entre cada  $x_i$  com seus vizinhos  $x_j$ . Desta forma, e como poderá ser observado nos resultados obtidos com estes indicadores no caso de São Paulo, é interessante adotar ambos indicadores na identificação de autocorrelação espacial uma vez que estes trazem informações complementares.

## DADOS E MATERIAIS

Para a realização deste trabalho foram utilizados os seguintes materiais:

- Dados extraídos do Mapa da Exclusão/Inclusão Social da Cidade de São Paulo / Coord. Aldaíza Sposati / 1996 publicado pela educ;
- Excel 97 e Access 97/Windows para a manipulação dos bancos de dados (tabelas);
- O software STATISTICS/Windows para análises estatísticas tradicionais;
- SpaceStat 1.9 /DOS para análises espacial dos dados georeferenciados e Arcview 3.0/ Windows para manipulação e visualização destes dados.

## MANIPULAÇÃO DOS INDICADORES

Os dados geográficos quando estão representados por unidades de área poligonais apresentam uma dificuldade de interpretação básica que é a falta de informação sobre o comportamento do atributo dentro da unidade espacial básica de análise. A inacessibilidade a esta informação aliada à metodologia de pesquisa e aos critérios de divisão do espaço em unidades torna a análise espacial destes dados um pouco mais específica, na medida que uma se torna necessária uma avaliação sobre o que realmente representa este dado espacializado dentro de um determinado polígono de área igual a  $x$ .

A dificuldade de manipulação destes dados aumenta conforme a heterogeneidade de polígonos. Isto se torna especialmente crítico em áreas urbanas, onde existem áreas altamente povoadas (*ti* alto) e áreas pouco povoadas, como áreas rurais (*ti* baixo) por exemplo. Este tipo de problema surgiu logo nas primeiras análises dos dados, uma vez que São Paulo, como caso típico, apresenta alta diversidade de tamanho, forma e densidade de Distritos (polígonos), podendo ser identificado um padrão de área urbana com polígonos menores e áreas na região sul que apresentavam características de zonas rurais.

A fim de minimizar o problema heterogeneidade nos polígonos foi aplicado aos dados brutos um ponderador de intensidade, desta forma, os dados foram ponderados pelo número de habitantes da área e pelo número relativo aos outros polígonos. Esta técnica de manipulação foi proposta Besag e Newell em 1991 através da seguinte formulação:

$$q_i = (n_i - p t_i) / [t_i p(1 - p)]$$

onde:  $p = \sum n_i / \sum t_i$

$n_i$  = número de casos observado

$t_i$  = população total do polígono I

Os valores (“scores”) que são obtidos a partir desta técnica variam de -100 a 100 e apresentaram uma classificação interessante dos indicadores escolhidos para a análise. Como uma primeira avaliação, estão organizados a seguir em ordem decrescente os onze primeiros distritos em função de a) número total de pessoas idosas; b) porcentagem de pessoas idosas e c) o  $q_i$  de pessoas idosas.

POPULAÇÃO >70	% DE IDOSOS	$q_i$ DE IDOSOS
V. MARIANA	CONSOLAÇÃO	JD. PAULISTA
JD. PAULISTA	JD. PAULISTA	V. MARIANA
SAÚDE	LAPA	CONSOLAÇÃO
SANTANA	MOÓCA	LAPA
PERDIZES	PINHEIROS	PINHEIROS
IPIRANGA	BELÉM	MOÓCA
PENHA	V. MARIANA	STA. CECÍLIA
JABAQUARA	CAMBUCI	SAÚDE
ITAIM BIBI	BARRA FUNDA	IPIRANGA
PINHEIROS	STA. CECÍLIA	PERDIZES
STA. CECÍLIA	IPIRANGA	TATUAPÉ

Neste trabalho foram utilizados três variáveis fornecidas pelo Mapa da Exclusão/Inclusão de São Paulo. Foram elas:



- número de domicílios alugados;
- número de chefes de família com renda até dois salários mínimos ;
- número de pessoas maiores de 70 anos.

A escolha destas variáveis não procurou a princípio ilustrar de maneira completa o quadro da exclusão social dentro do território urbano, mesmo porque, o significado da exclusão social passa pela avaliação e ponderação de inúmeras variáveis. Esta escolha partiu antes da facilidade de interpretação que representaria a aplicação das estatísticas Moran e Gi.

## ÁREA DE ESTUDO

O Município de São Paulo é a maior cidade do Brasil com cerca de 10 milhões de habitantes, juntamente com mais 36 municípios a Região Metropolitana de São Paulo representa uma das maiores aglomerações humanas do planeta. Apesar destes números, todo o processo de expansão ocorrido neste século e, particularmente nas quatro últimas décadas, foi extremamente rápido e caótico culminando num quadro social gravíssimo na cidade atual.

O Mapa da Exclusão / Inclusão Social trabalha com o recorte distrital do espaço urbano. São 96 distritos que correspondem a aglomerações significativamente maiores que muitas cidade brasileiras.

A heterogeneidade das unidades básicas de análise, os distritos, é mais óbvia quando são comparados os tamanhos e formas dos polígonos centrais aos polígonos das regiões de borda, especialmente no quadrante sul. Devido a esta heterogeneidade partiu-se para uma análise, ou que chamei de visão geral sobre a área de estudo, onde são avaliadas as características de cada distrito individualmente e dentro do conjunto total de polígonos. Basicamente busca-se identificar a significância de cada distrito dentro do município.

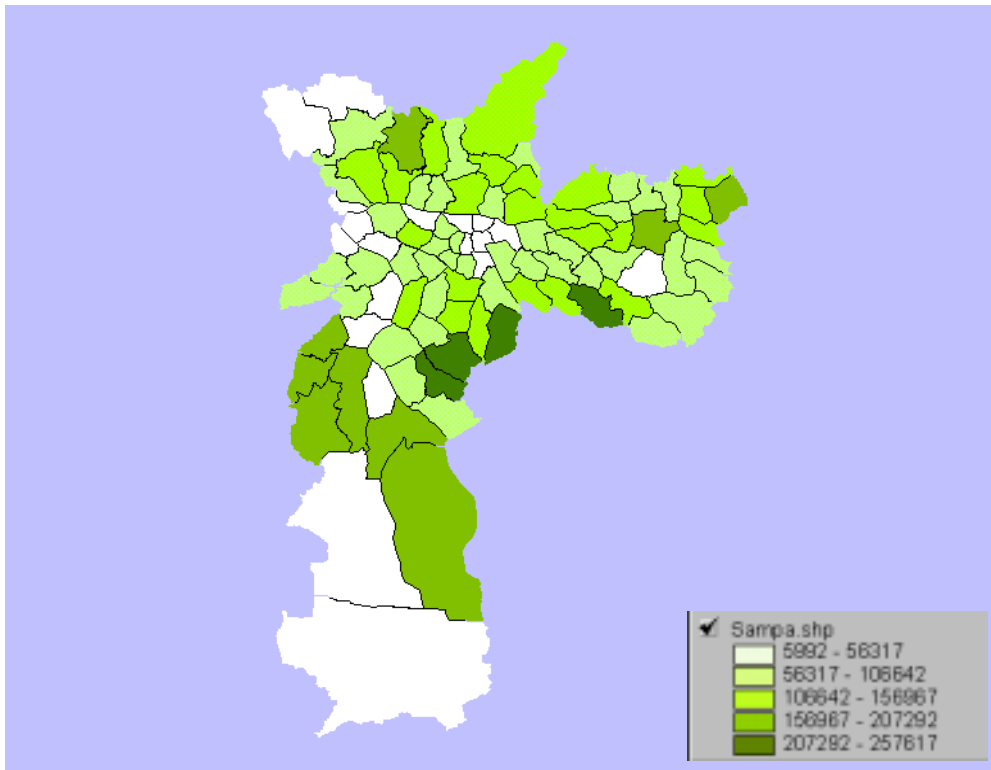


Fig. 1 . Divisão distrital do município de São Paulo classificado por população total por distrito.

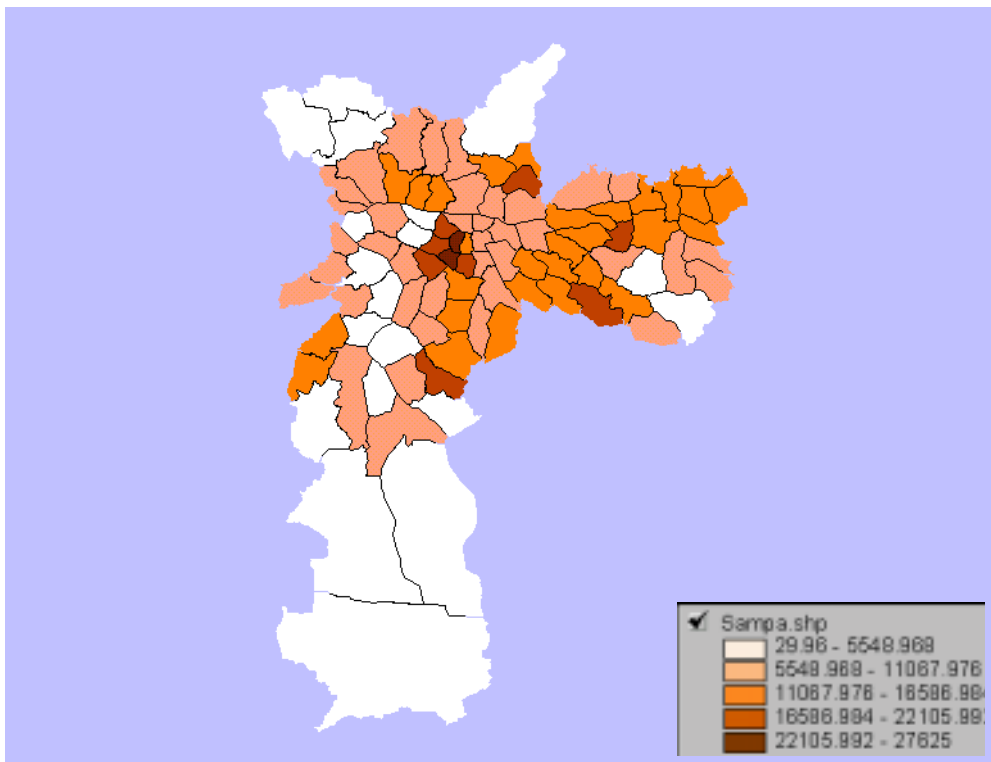


Fig.2. Distritos do município de São Paulo classificados por densidade demográfica.

## APLICAÇÃO DOS MÉTODOS ESTATÍSTICOS AOS DISTRITOS DO MUNICÍPIO DE SÃO PAULO

A partir das três variáveis  $q_i$  construídas (idosos, alugueis e sem renda) foram realizadas basicamente quatro tipos de análises a partir dos indicadores de autocorrelação local Moran e  $G_i$ . Uma refere-se ao cálculo do Índice de Moran Local tradicional e outra ao cálculo do Índice de Moran Local a partir de 99 permutações. A aplicação de permutações permite em tese uma melhor aproximação a uma distribuição normal, dessa forma garantindo melhores inferências aos dados. Os outros dois mapas gerados dizem respeito às estatísticas  $G_i$  e  $G_i^*$ . É importante lembrar que para ambas as estatísticas foi adotada a mesma matriz de vizinhança  $W_{ij}$ .

A interpretação dos diferentes mapas gerados, parte primeiro de uma análise estatística básica sobre os dados. Todos os mapas estão acompanhados do histograma e do gráfico de probabilidade de normalidade dos indicadores de autocorrelação espacial, dessa forma torna-se mais direta a interpretação dos valores e a coerência dos intervalos de confiança construídos.

Ao lado de cada mapa, ou resultados obtidos, o SpaceStat 1.9 gera uma lista apontando os comportamentos mais extremos e os respectivos distritos relacionados a estes comportamentos. Para esta classificação de “outliers”, o programa se vale da regra dos desvios padrões. Dessa forma, pode-se esperar que se encontre uma estatística amostral real,  $S$ , situada nos intervalos de  $\mu_s - \sigma_s$  a  $\mu_s + \sigma_s$ , de  $\mu_s - 2\sigma_s$  a  $\mu_s + 2\sigma_s$ , ou de  $\mu_s - 3\sigma_s$  a  $\mu_s + 3\sigma_s$ , aproximadamente em 68,27%, 95,45% e 99,73% de vezes respectivamente. Como a identificação de comportamentos extremos se baseia na construção destes intervalos de confiança, todos os mapas foram classificados a partir dos desvios em relação à média. Assim a visualização e identificação destes casos extremos se dá de maneira direta e intuitiva.

## ÍNDICE LOCAL DE MORAN PARA $q_i$ DE DOMICÍLIOS ALUGADOS

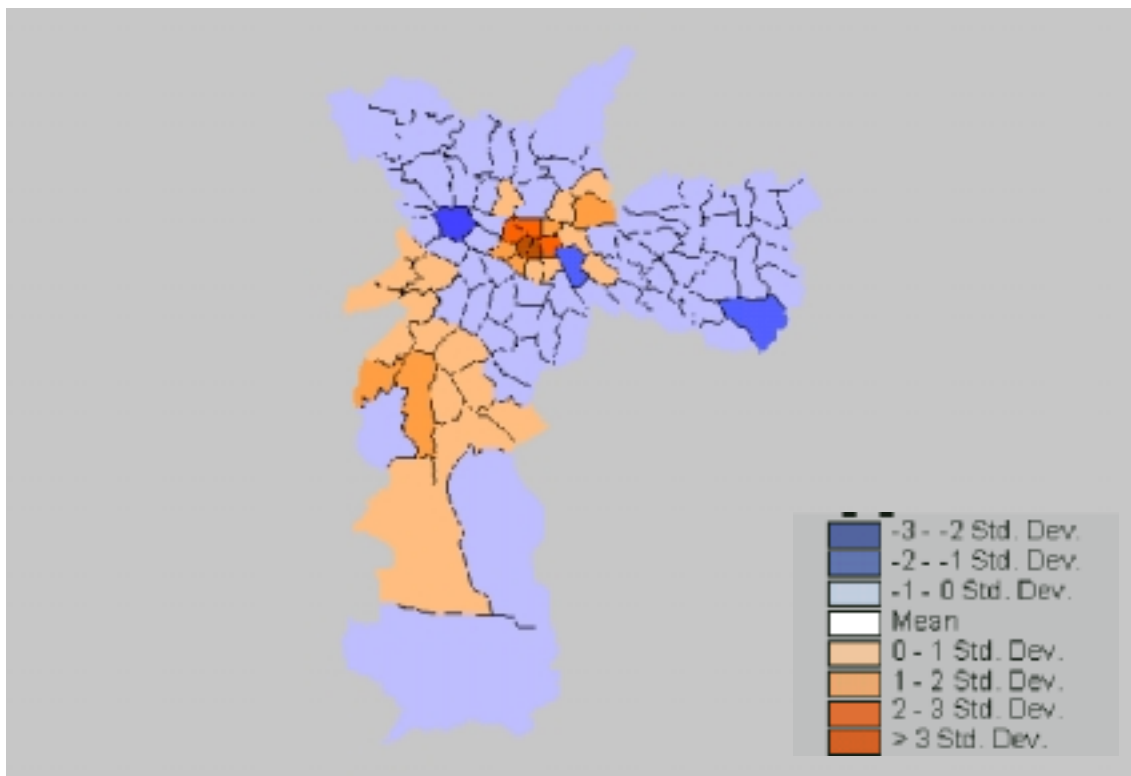


Fig. 3. Mapa dos índices de Moran Local gerados a partir da variável  $q_i$  de domicílios alugados e classificados em função da regra da média mais desvios padrões.

Histograma do LM\_  $q_i$  aluguel

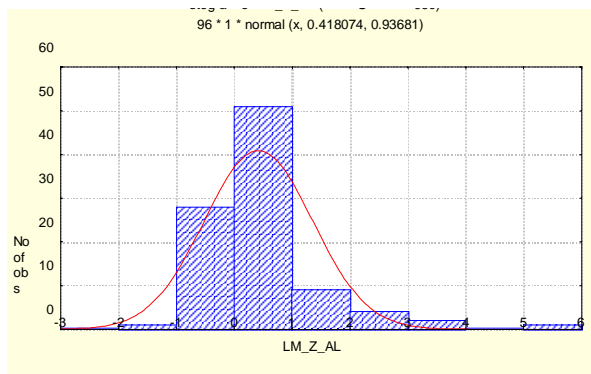
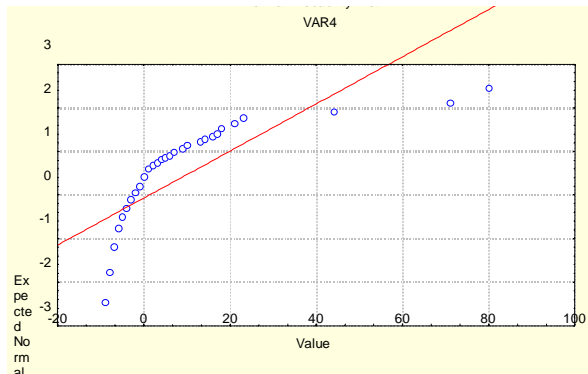


Gráfico de probabilidade normal



Distritos identificados como “outliers”:

IBGE	LM
SÉ	3.579176
LAPA	-1.55504
SANTA CECÍLIA	3.128312
<b>REPÚBLICA</b>	<b>5.159141</b>
BOM RETIRO	2.788280
BRÁS	2.360710

## ÍNDICE LOCAL DE MORAN SOBRE PERMUTAÇÕES

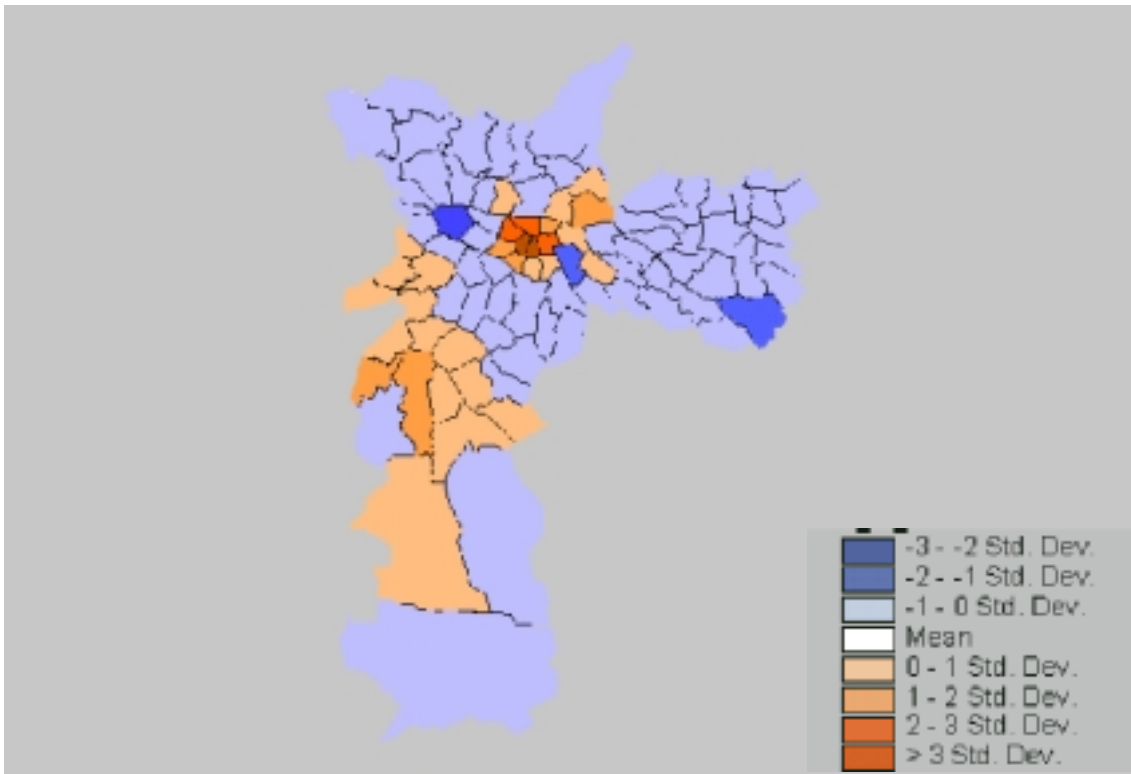


Fig. 4. Distritos do município de São Paulo classificados de acordo com índice de moran local sobre permutações para qi de alugueis.

Histograma do LM\_qi\_aluguel

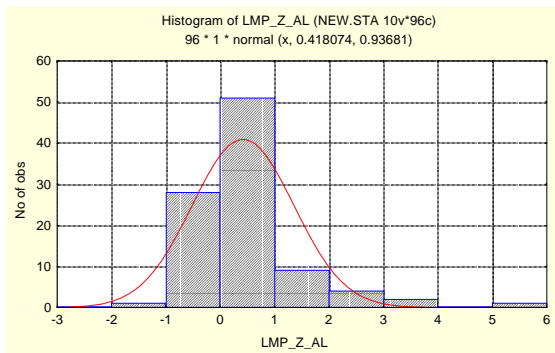
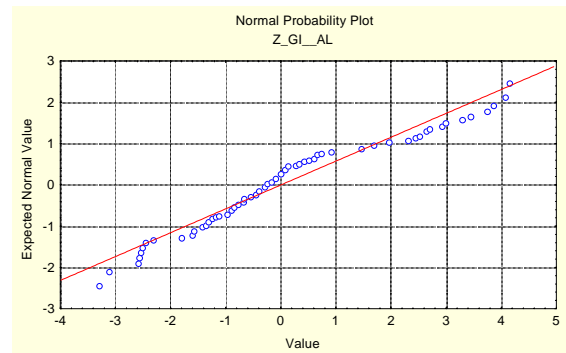


Gráfico de probabilidade normal



Distritos identificados como “outliers”:

IBGE	LMP
SÉ	3.579176
LAPA	-1.55504
SANATA CECÍLIA	3.128312
<b>REPÚBLICA</b>	<b>5.159141</b>

# ESTATÍSTICA Gi PARA qi ALUGUEIS

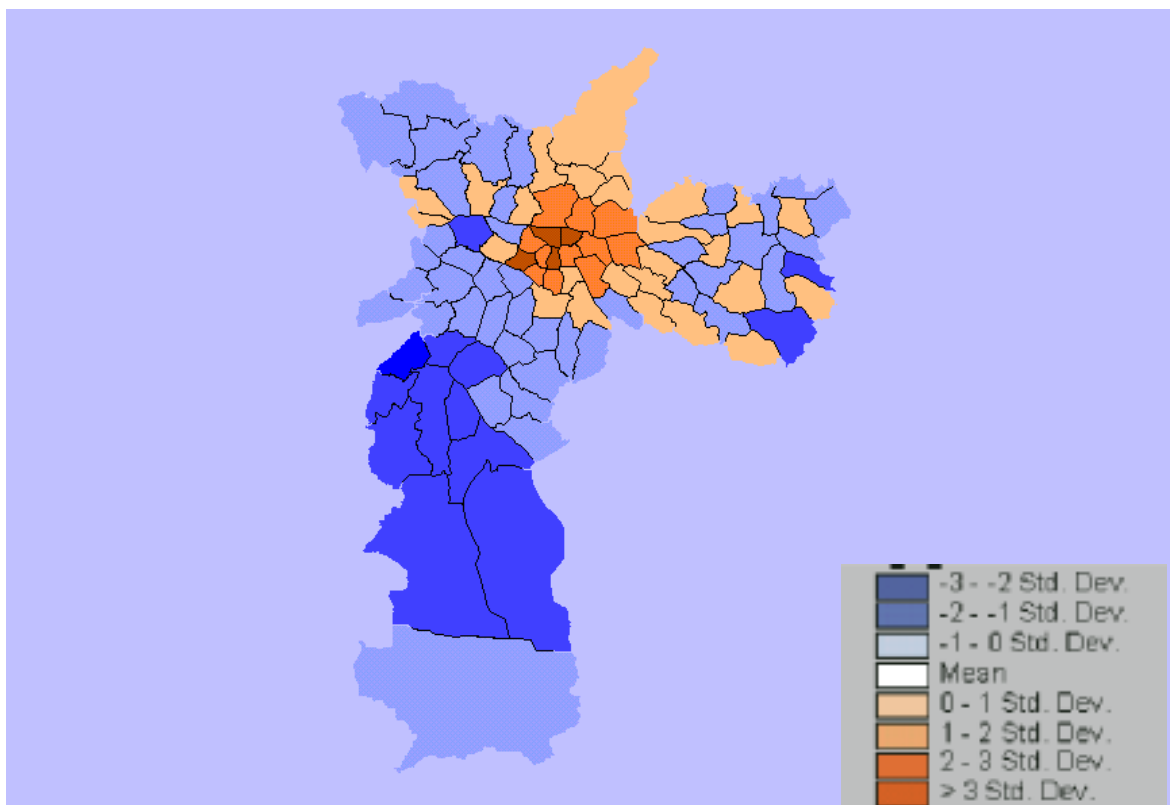
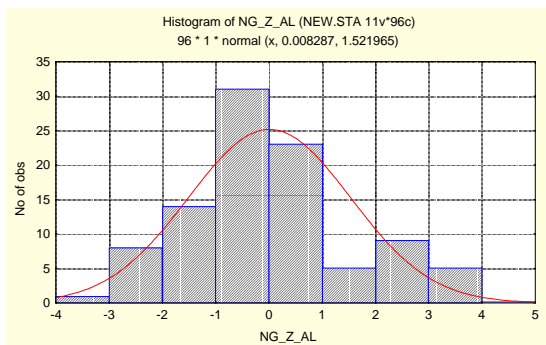
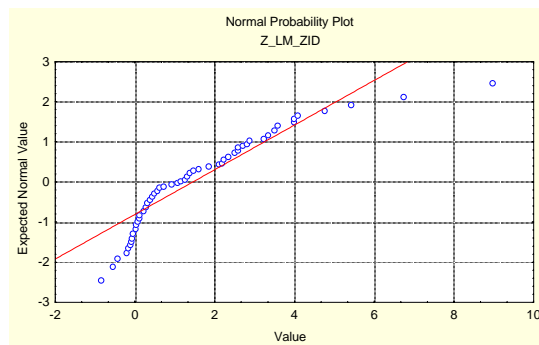


Fig. 5. Distritos de São Paulo classificados pelo valor de Gi.

## Histograma de Gi para qi\_aluguel



## Gráfico de probabilidade normal



Distritos identificados como “outliers” para Gi de qi de aluguel:

IBGE	Z	PROB
<b>CONSOLAÇÃO</b>	<b>3.7602</b>	<b>0.0002</b>
<b>SÉ</b>	<b>3.6893</b>	<b>0.0002</b>
<b>CAMPO LIMPO</b>	<b>-3.0400</b>	<b>0.0024</b>
<b>JD. ÂNGELA</b>	<b>-2.7608</b>	<b>0.0058</b>

## ESTADÍSTICA $G_i^*$ PARA $q_i$ ALUGUÉIS

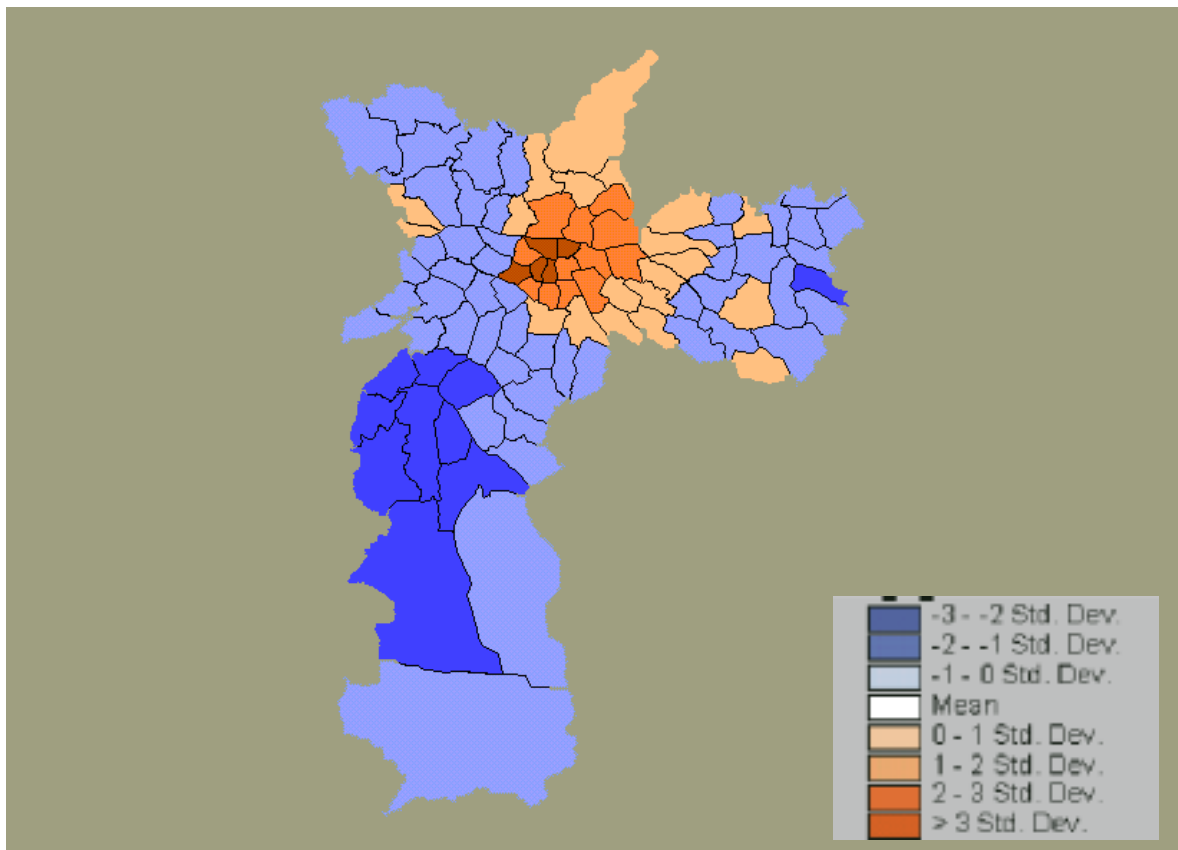


Fig. 6. Distritos de São Paulo classificados pelo valor de  $G_i^*$  de  $q_i$  de aluguel.

Histograma de  $G_i^*$  para  $q_i$  aluguel

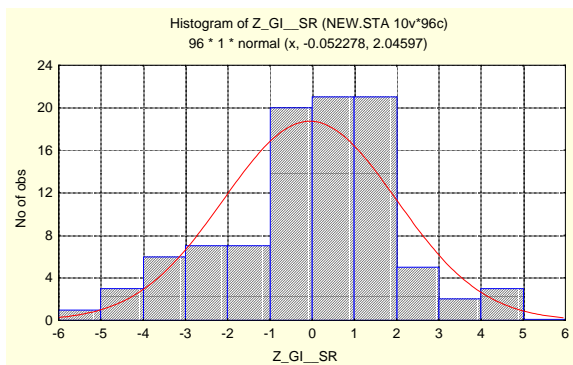
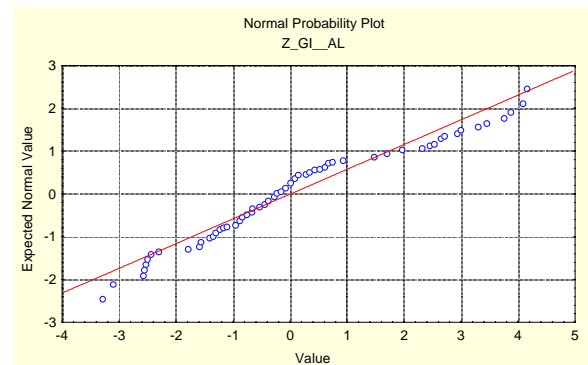


Gráfico de probabilidade normal



Distritos identificados como “outliers” para  $G_i^*$  de  $q_i$  de aluguel:

IBGE	Z	PROB
<b>SÉ</b>	<b>4.1501</b>	<b>0.0000</b>
REPÚBLICA	4.0764	0.0000
<b>JD. SÃO LUIS</b>	<b>-3.2949</b>	<b>0.0010</b>
CAMPO LIMPO	-3.1093	0.0019

## RESULTADOS SOBRE DOMICÍLIOS ALUGADOS

Nas análises realizadas sobre a variável “qi” de domicílios alugados foi possível identificar um comportamento diferenciado para os distritos da região central da cidade. Todos os indicadores de autocorrelação gerados apontaram o centro da cidade como a área da cidade que apresenta maior número de famílias vivendo em imóveis não próprios. O quadrante sudoeste da cidade também foi indicado como uma região com comportamento diferenciado em relação à média da cidade, porém, neste caso os indicadores Gi apontaram como sendo uma alta concentração de valores baixos da variável “qi”.

A indicação da região central da cidade como um “outlier” em relação ao número de imóveis alugados poderia, em uma primeira análise, ser explicado pelo alto número de cortiços, principalmente quando analisamos que este comportamento se dá mais sobre o quadrante nordeste, quadrante que engloba as regiões do Brás, Belém e Bom Retiro, conhecidos distritos com este tipo de ocupação.

O comportamento apontado no quadrante sudoeste, sobretudo nos distritos de Campo Limpo, Jardim Ângela e Jardim São Luís poderia ser explicado pela alta concentração de loteamentos irregulares e invasões presentes nestas regiões. É interessante notar que através de apenas uma rápida especulação sobre o significado destas estatísticas, já pudemos identificar dois padrões distintos de ocupação e uma forte polarização nordeste/sudoeste dentro do espaço urbano.

No caso desta variável é possível observar como a aplicação da técnica de permutações consegue um ajuste significativo em relação à normalidade dos dados (veja Fig.3 e Fig.4). Os indicadores Gi e Gi\* apresentaram distribuições onde a aceitação da normalidade é bem coerente.



## ÍNDICE LOCAL DE MORAN PARA $q_i$ DE CHEFES DE FAMÍLIA SEM RENDA

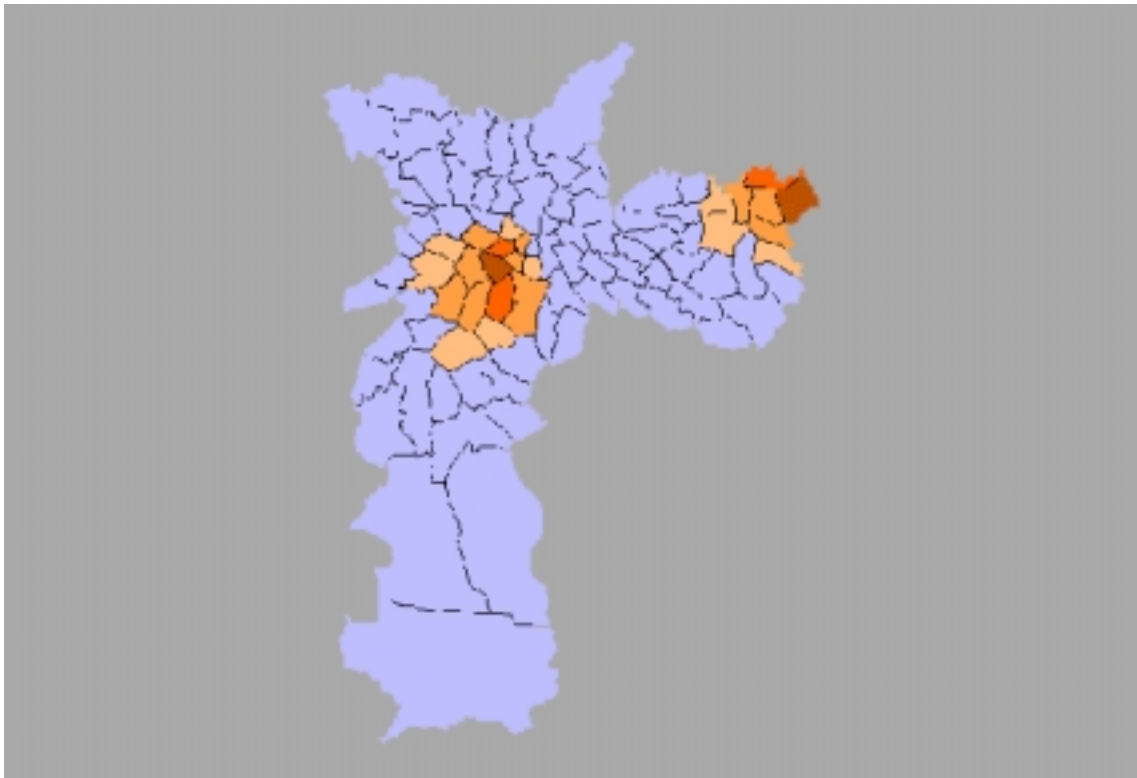


Fig. 7. Mapa dos distritos de São Paulo classificados em relação ao Índice Local de Moran para a variável “chefes de família sem renda”

Histograma de  $l_i$  para  $q_i$ \_sem renda

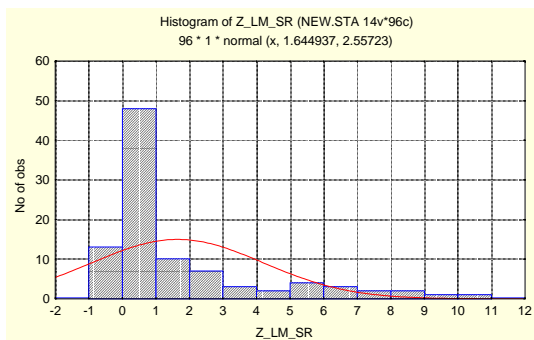
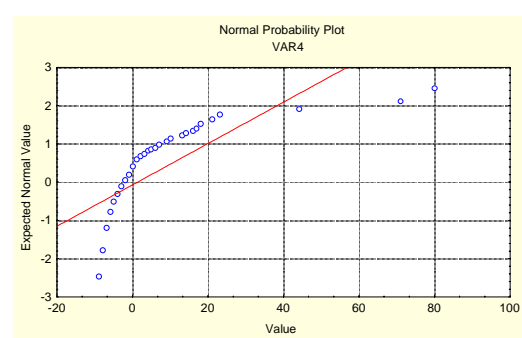


Gráfico de Prob. de Normalidade



“Outliers” apontados pelo SpaceStat:

IBGE	LOCAL MORAN
CONSOLAÇÃO	3.372106
<b>JD. PAULISTA</b>	<b>5.925623</b>
<b>ITAIM PAULISTA</b>	<b>4.440743</b>
JD. HELENA	3.608146
MOEMA	4.258492

## ÍNDICE LOCAL DE MORAN SOBRE PERMUTAÇÕES

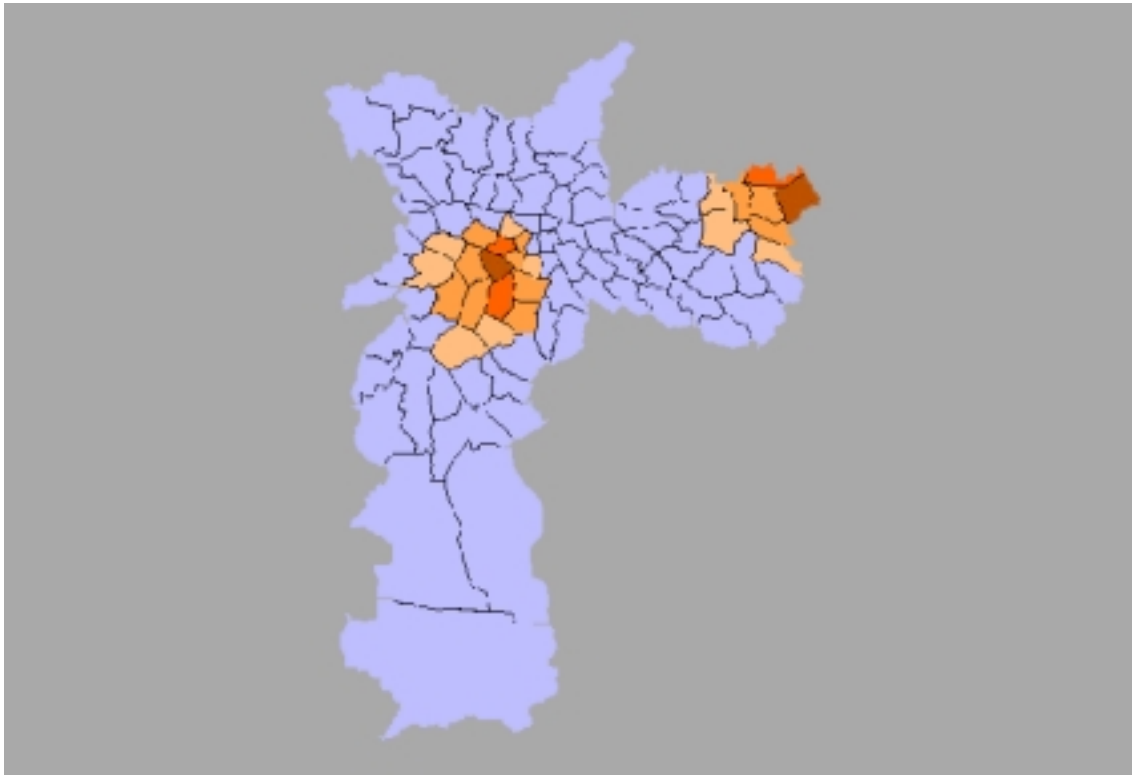


Fig. 8. Índice Local de Moran após 99 permutações para a variável “chefe de família sem renda” para os distritos de São Paulo.

Histograma de li para qi\_sem renda

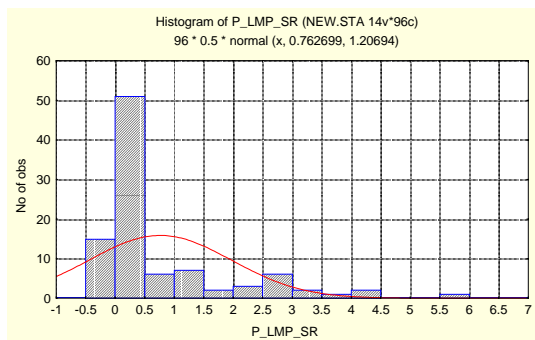
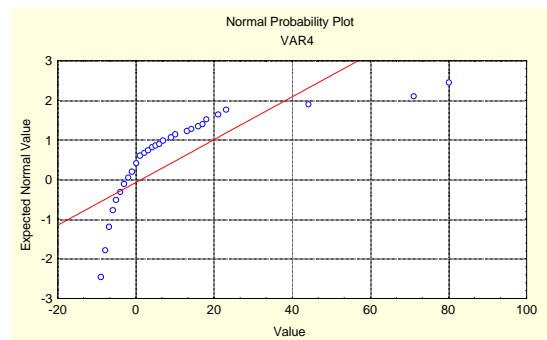


Gráfico de Prob. de Normalidade



“Outliers” apontados pelo SpaceStat:

IBGE	LOCAL MORAN
CONSOLAÇÃO	3.372106
<b>JD. PAULISTA</b>	<b>5.925623</b>
<b>ITAIM PAULISTA</b>	<b>4.440743</b>
JD. HELENA	3.608146
MOEMA	4.258492

# ESTATÍSTICA $G_i$ PARA A VARIÁVEL $q_i$ CHEFES DE FAMÍLIA SEM RENDA

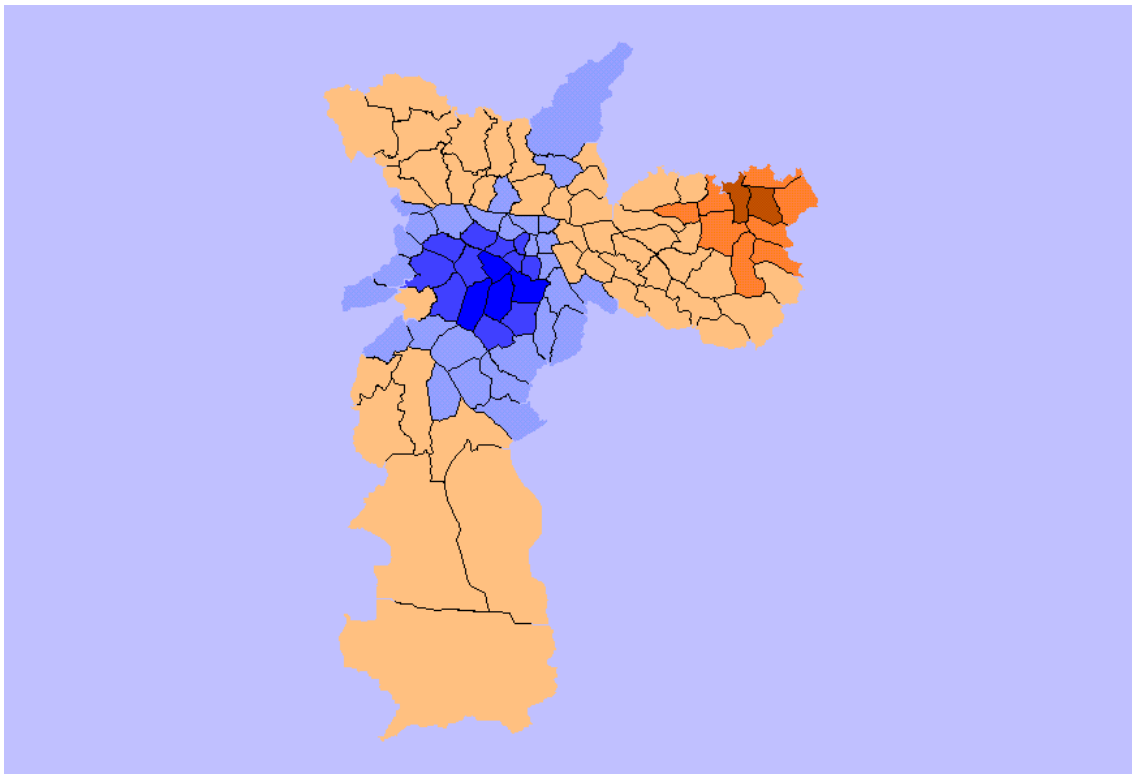
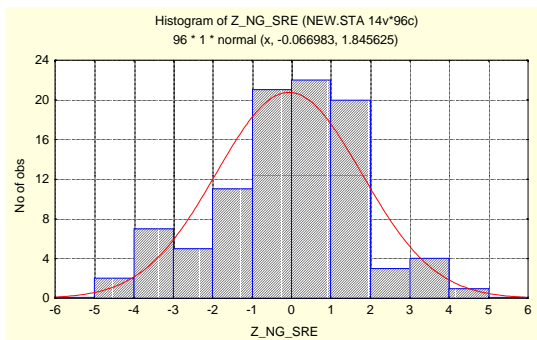
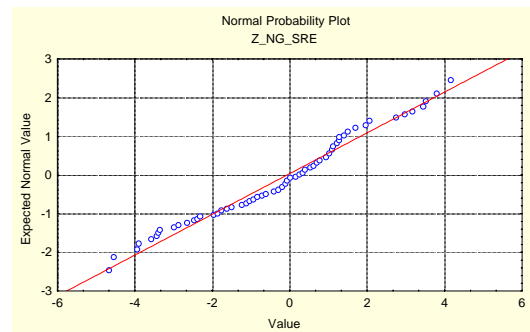


Fig. 9. Distritos de São Paulo classificados de acordo com o valor  $G_i$  calculados para a variável “chefes de família sem renda”.

Histograma de  $G_i$ \_sem renda



Probabilidade de normalidade



“Outliers” apontados para  $q_i$ \_sem renda

IBGE	Z	PROB
VILA CURUÇÁ	4.1568	0.0000
SÃO MIGUEL	3.7919	0.0001
<b>MOEMA</b>	<b>-4.6730</b>	<b>0.0000</b>
<b>ITAIM BIBI</b>	<b>-4.5468</b>	<b>0.0000</b>

## ESTATÍSTICA $G_i^*$ PARA A VARIÁVEL $q_i$ CHEFES DE FAMÍLIA SEM RENDA

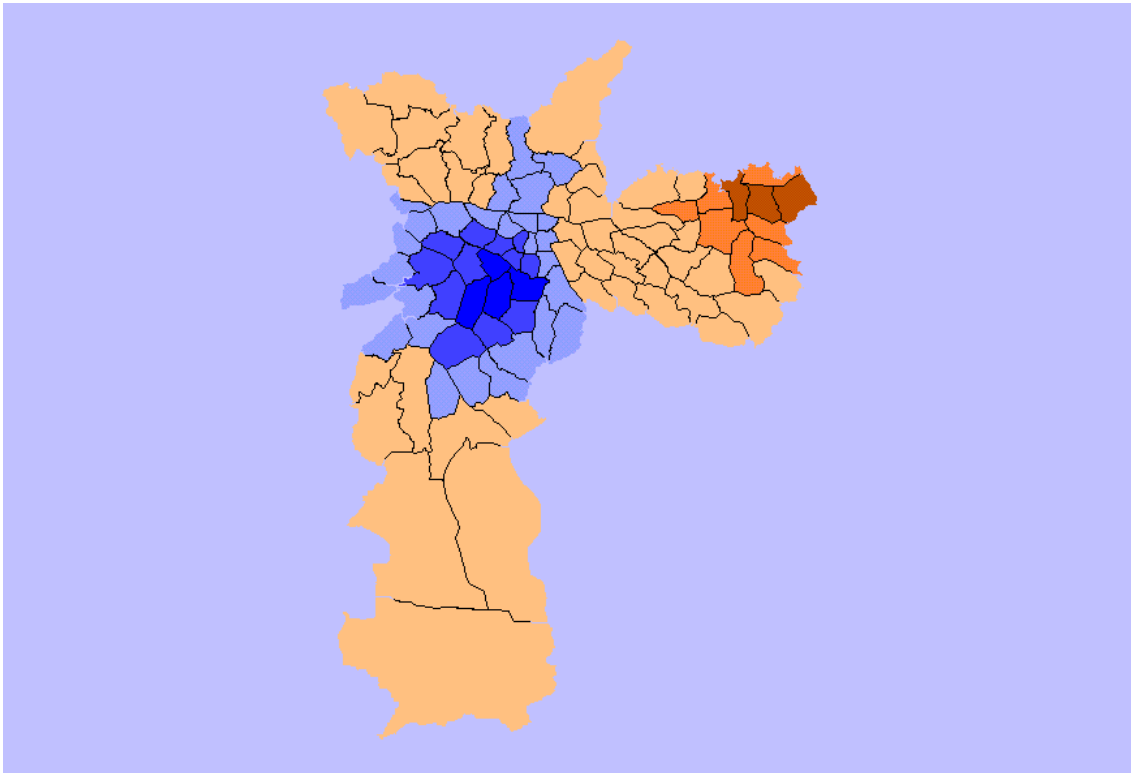
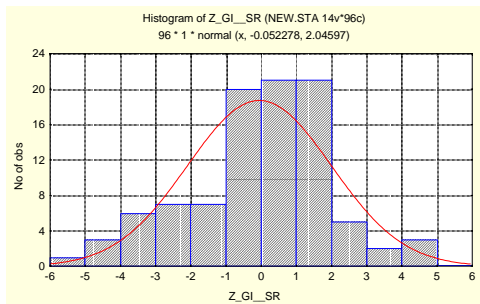
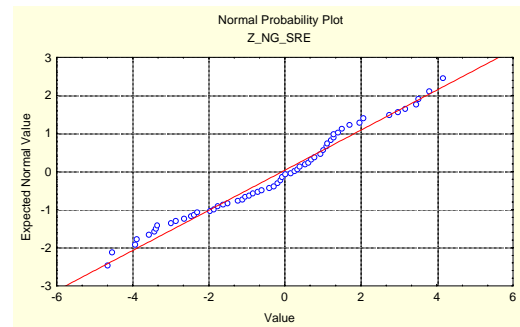


Fig. 10. Classificação dos distritos do município de São Paulo em relação ao valor de  $G_i^*$  calculado para a variável  $q_i$  de chefes de família sem renda.

### Histograma de $G_i^*$ \_sem renda



### Prob. Normal



“Outliers” apontados para  $G_i^*$  de chefes de família sem renda

IBGE	Z	PROB
VILA CURUÇÁ	4.3464	0.0000
ITAIM PAULISTA	4.0837	0.0000
<b>MOEMA</b>	<b>-5.0586</b>	<b>0.0000</b>
<b>ITAIM BIBI</b>	<b>-4.8552</b>	<b>0.0000</b>

## RESULTADOS SOBRE CHEFES DE FAMÍLIA SEM RENDA

Novamente o padrão altamente polarizado pode ser observado. Neste caso, porém, o contraste ocorreu mais fortemente entre os distritos do chamado centro expandido e os distritos do extremo leste da capital. É possível avaliar as técnicas estatísticas Local Moran e Gi através desta variável e perceber que o Indicador de autocorrelação local é mais restritivo na atribuição de nível de confiança para aceitação da hipótese de ausência de regime espacial extremo.

Apesar da distribuição dos indicadores Moran não apresentarem uma normalidade, nem mesmo após aplicada a técnica de permutações, os resultados encontrados em comparação aos levantados nas estatísticas Gi são muito compatíveis e coerentes.

A zona leste é caracterizada pela ocupação dos grandes projetos habitacionais das décadas de 70 e 80, e são na realidade verdadeiros distritos dormitórios dentro da estrutura urbana da capital paulista.

A polarização é extrema quando estes distritos da zona leste são comparados aos da região do centro expandido onde os distritos de Moema e Itaim Bibi indicados como “outliers” por todas as estatísticas. Num primeiro momento parece óbvio o resultado alcançado nestas estatísticas, uma vez que é bastante conhecido o processo de apropriação do espaço urbano pelas elites paulistanas no decorrer da história. Este deslocamento no sentido do quadrante sudoeste começou já no início do processo de metropolização da cidade com as grandes mansões nos Campos Elísios, posteriormente Av. Paulista, Av. Brasil e atualmente a Av. Berrini. Neste aspecto é muito interessante a visualização do indicador Gi, pois o mapa desenha quase que perfeitamente este processo de apropriação.

## ÍNDICE DE MORAN LOCAL PARA $q_i$ DE POSSOAS IDOSAS

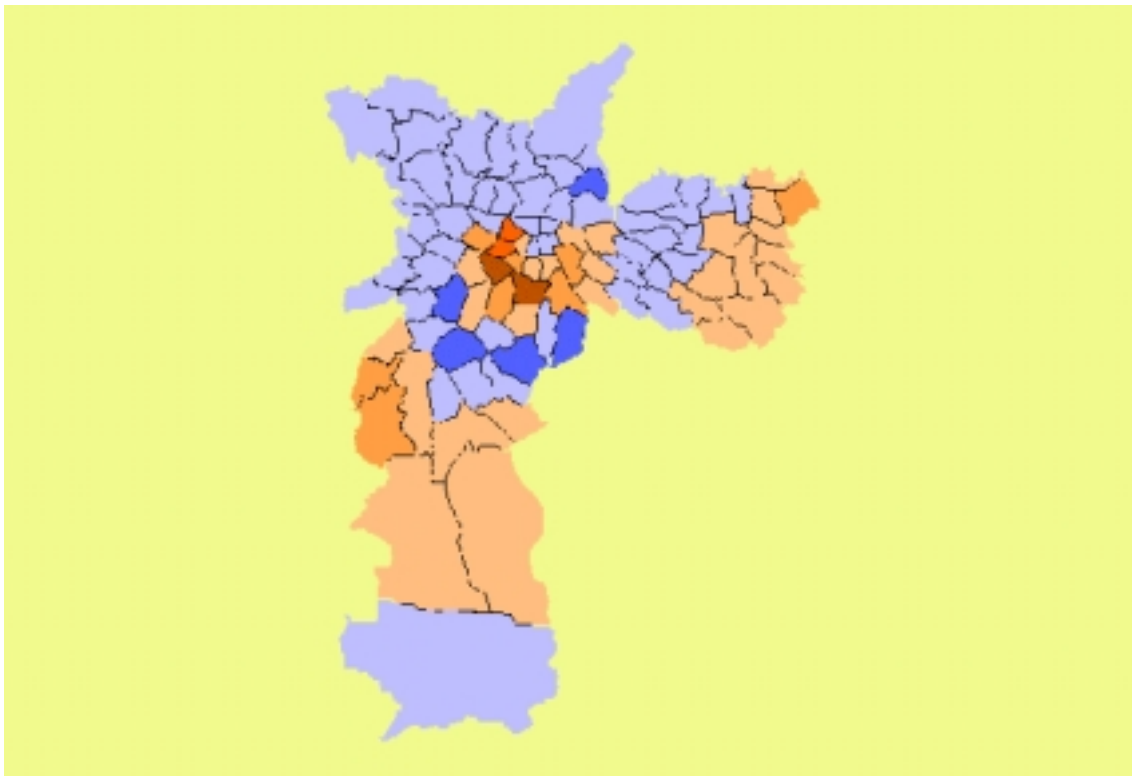
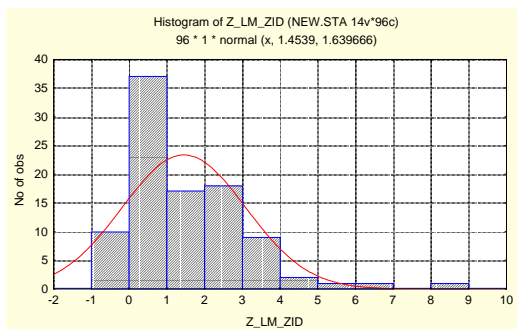
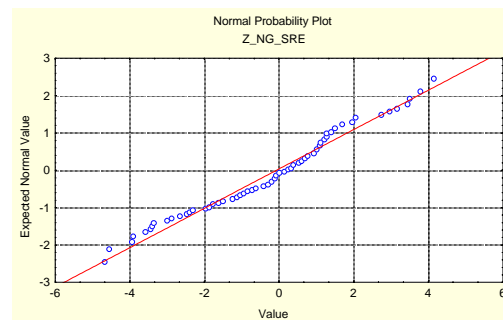


Fig. 11. Distritos de São Paulo classificados pelo índice de Moran Local calculados sobre a variável  $q_i$  de pessoas idosas.

Histograma de  $li_{q_i}$  idosos



Prob. Normal



Outliers identificados para o Moran Local de  $q_i$  de pessoas idosas:

IBGE	LOCAL MORAN
SANTA CECÍLIA	2.273214
CONSOLAÇÃO	2.640032
<b>JD. PAULISTA</b>	<b>3.812878</b>
VILA MARIANA	3.018939

## MORAN LOCAL APÓS 99 PERMUTAÇÕES PARA $q_i$ DE PESSOAS IDOSAS

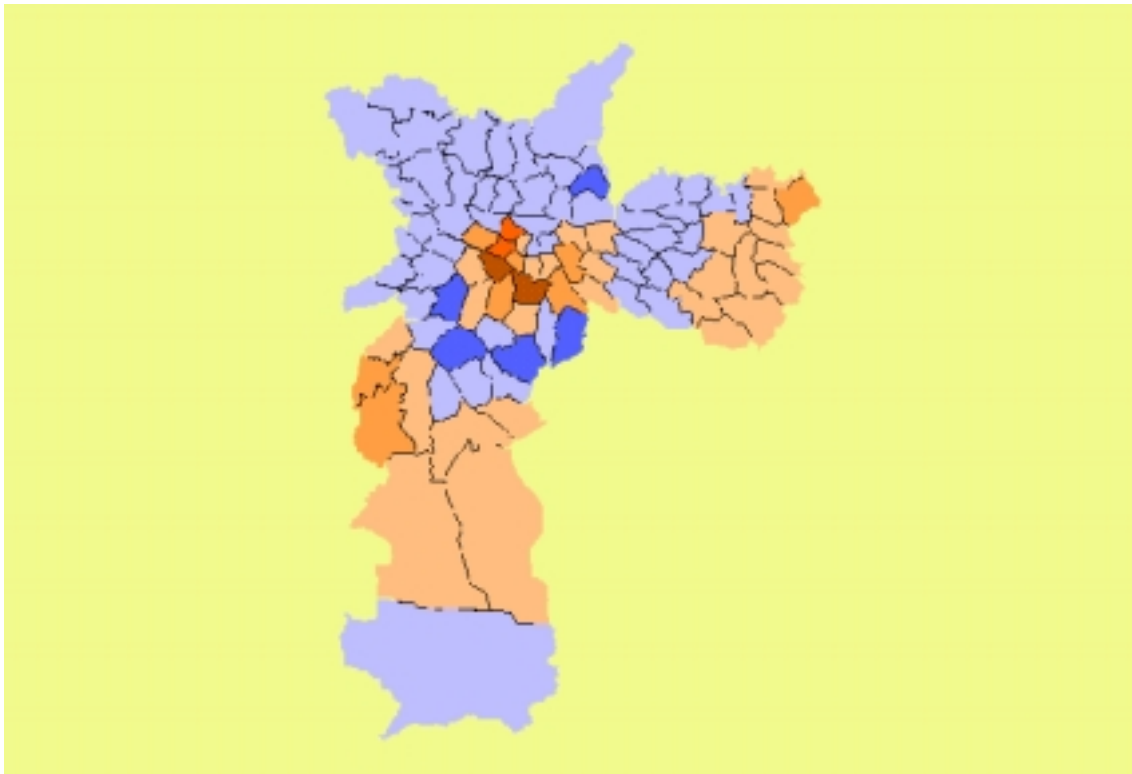
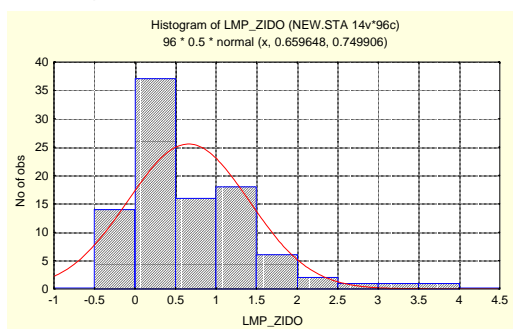
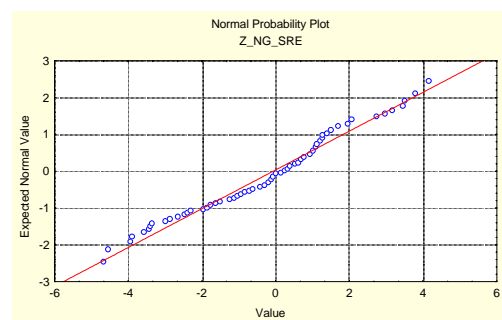


Fig. 12. Distritos classificados de acordo com o índice local de moran após 99 permutações sobre  $q_i$  de idosos.

Histograma de LMP\_qi\_idosos



Prob. Normal



“Outlier” identificados para local moran após permutações:

IBGE	LOCAL MORAN
SANTA CECÍLIA	2.273214
CONSOLAÇÃO	2.640032
<b>JD. PAULISTA</b>	<b>3.812878</b>
VILA MARIANA	3.018939

## ESTATÍSTICA $G_i$ PARA $q_i$ DE IDOSOS

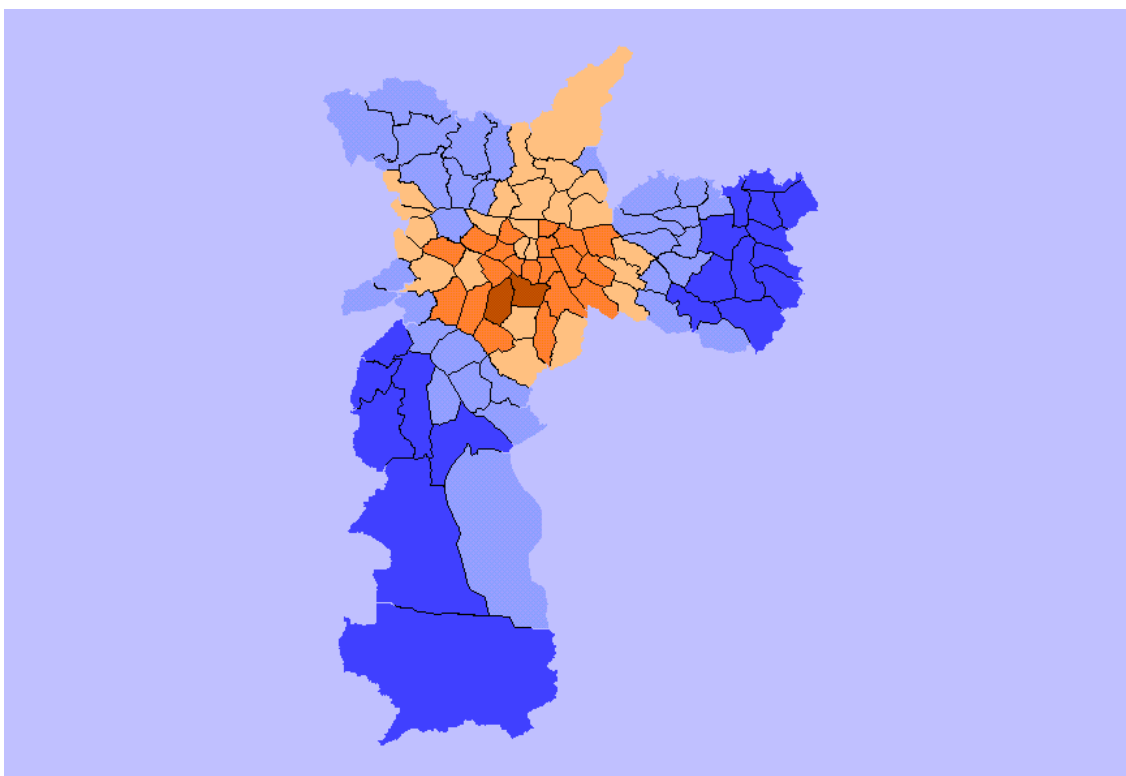
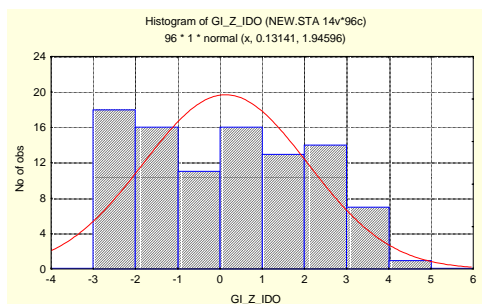
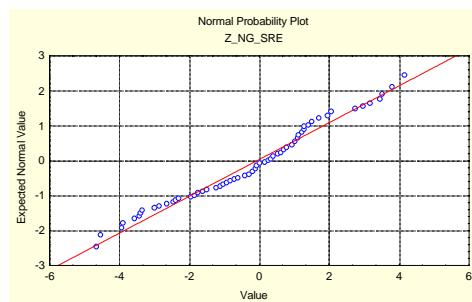


Fig. 13. Distritos paulistanos classificados segundo o  $G_i$  encontrado para a variável  $q_i$  de idosos.

### Histograma de $G_i$ $q_i$ idosos



### Prob. Normal



“Outliers” apontados pelo SpaceStat para os  $G_i$  do  $q_i$  de pessoas idosas:

IBGE	Z	PROB
<b>VILA MARIANA</b>	<b>3.9700</b>	<b>0.0001</b>
<b>MOEMA</b>	<b>3.6893</b>	<b>0.0002</b>
PARELHEIROS	-2.7774	0.0055
JOSÉ BONIFÁCIO	-2.5313	0.0114



## RESULTADOS DAS ANÁLISES SOBRE qi DE IDOSOS

Novamente o padrão apontado por todos os indicadores revelam dois regimes espaciais distintos, um na área mais central e outro antagônico nas periferias. Dentro desta variável é possível observar que as áreas que apresentam maiores qi de idosos coincidem com as áreas mais abastadas do espaço urbano. Não seria equivocado estabelecer esta relação, na medida em que o acesso a serviços de saúde, as condições de habitabilidade, violência e padrões de renda são fatores diretamente relacionados à longevidade da população.

Nesta variável, tanto as periferias leste como sul foram apontadas como tendo um comportamento extremo porém negativo, ou seja, há uma clara necessidade de se investigarem os aspectos listados acima para que se estude até que ponto as condições de vida estão influenciando esta variável.

É interessante notar nos indicadores locais de Moran a presença de uma espécie de anel de polígonos com baixos valores justamente nas áreas de transição entre o regime espacial da região central com o regime espacial das periferias. Este tipo de informação é só é possível no Moran Local, é por este motivo que quando utilizamos o Moran Local em conjunto com as estatísticas Gi, podemos observar o comportamento espacial das variáveis de maneira muito mais global.

A construção dos intervalos de confiança para “outliers” pode ser questionada até certo ponto, uma vez que as distribuições apresentaram um certo grau de assimetria, fator que interfere diretamente nesta inferência.

## CONCLUSÕES

O que se pode concluir ao final deste trabalho é muito mais do que as observações sobre os resultados observados nos mapas de indicadores de autocorrelação. Entende-se que estes tipos de análises espaciais são processos complexos uma vez que toda uma manipulação de indicadores de pesquisas, verificação do suporte gráfico (mapa coroplético) e pesquisa de ferramentas estatísticas são peças fundamentais para a coerência das análises. Como conclusões mais diretas podemos apontar:

- A escala dos dados (tamanho e semelhança entre os polígonos) é um fator decisivo para qualquer tipo de análise espacial. No caso dos distritos de São Paulo a heterogeneidade tanto espacial quanto em relação às características internas de cada distrito limitaram uma visualização mais refinada das análises. Fica agora como sugestão para próximas pesquisas a necessidade de se buscar uma base com áreas mais refinadas (escala mais local) e que englobe toda a região metropolitana, uma vez que alguns desvios observados poderia ser explicado pela localização de borda de um polígono.
- A manipulação dos dados brutos deve sempre estar relacionada a que tipo de informação se quer extrair. As diferenças internas de cada polígono com relação a número de habitantes e densidade de habitantes é um complicador para o entendimento do que estes dados estão nos revelando. A utilização do ponderador de intensidade aparentemente se mostrou uma possibilidade bastante interessante.
- Ambos indicadores de autocorrelação local apresentaram bons resultados e apontaram padrões coerentes à realidade do Município. Como já foi apontado antes, estes indicadores quando utilizados em conjunto nos propiciam uma visão muito mais abrangente da situação em estudo.

- Dentro da perspectiva de se mapear a exclusão social no município de São Paulo estas duas ferramentas estatísticas se mostraram de grande valia. A capacidade de síntese e a visualização cartográfica destes indicadores. Podem ser observados através destes indicadores regimes espaciais extremamente polarizados dentro do ambiente urbano, a caracterização das zonas de centralidade (incluídos) e das periferias (excluídos) ocorre em todos os dados analisados e evidencia o paradoxo que vivem atualmente nossas cidades.

## BIBLIOGRAFIA

Anselin, L., Local Indicators of Spatial Association – LISA. In: **Geographical Analysis**, Vol. 27, No. 2 (April 1995)

Anselin, L., **SpaceStat 1.8 user's guide**, Technical Report S-92, National Center for geographic Information and Analysis, 1992.

Bailey, T.; Gatrell, A.C., **Interactive Spatial Data Analysis.**, Longman Group Limited, England, 1995.

Getis, A.; Ord, J.K., Local Spatial Autocorrelation Statistics: Distributional Issues and an Application., In: **Geographical Analysis**, vol. 27, No.4 (October 1995)

Getis, A.; Ord, J.K., The Analysis of Spatial Association by Use of Distance Statistics. In: **Geographical Analysis**, Vol. 24, No. 3 (July 1992)

Sposati, A., **Mapa da Exclusão/Inclusão da Cidade de São Paulo**, Editora da PUC-SP, São Paulo, 1996.