



Ministério da Ciência e Tecnologia
Instituto Nacional de Pesquisas Espaciais

DIVISÃO DE PROCESSAMENTO DE IMAGENS - DPI

ANÁLISE ESPACIAL DE DADOS GEOGRÁFICOS

LABORATÓRIO I

Roteiro prático

ANÁLISE DE PADRÕES DE ÁREAS

Dr. Eduardo Camargo e Dr. Carlos Felgueiras

Atualizado em 05/05/2010

RESUMO

Este laboratório aplica métodos estatísticos espaciais para investigar a existência de padrões espaciais em fenômenos a partir de informações associadas por unidades de área. Este caso ocorre com muita frequência quando lidamos com dados agregados por municípios, distritos, bairros e setores censitários, onde não se dispõe da localização exata dos eventos, mas de valores delimitados por polígonos. Alguns desses indicadores são contagens, como é o caso da maior parte das variáveis coletadas no censo: por exemplo, o IBGE fornece, para cada setor censitário, o número de chefes de família em cada uma das faixas de renda consideradas. Diversos indicadores de saúde também são deste tipo: o Ministério e Secretarias de Saúde organizam e disponibilizam dados de óbitos, partos, doenças transmissíveis e outros por município.

A forma usual de apresentação de dados agregados por áreas é através de mapas temáticos ou coloridos, exibindo o padrão espacial do fenômeno em estudo.

Quando visualizamos um padrão espacial, é muito útil traduzi-lo em considerações objetivas:

- a) o padrão que observamos é aleatório ou apresenta uma agregação definida?
- b) pode ser associado a causas mensuráveis?
- c) os valores observados são suficientes para analisar o fenômeno espacial a ser estudado?
- d) existem agrupamentos de áreas com padrões diferenciados dentro da região de estudo?

Para abordar estas questões é preciso empregar técnicas de análise espacial adequadas. Neste caso, a hipótese mais comum é supor que as áreas são diferenciadas, e que cada uma delas possui uma “identidade” própria. Do ponto de vista estatístico, isto implica em que cada área apresenta uma distribuição de probabilidade distinta das demais, o chamado *modelo espacial discreto*. Para maiores informações sugere-se a leitura do Capítulo 5 do livro *Análise Espacial de Dados Geográficos*, disponível em <http://www.dpi.inpe.br/gilberto/livro/analise>.

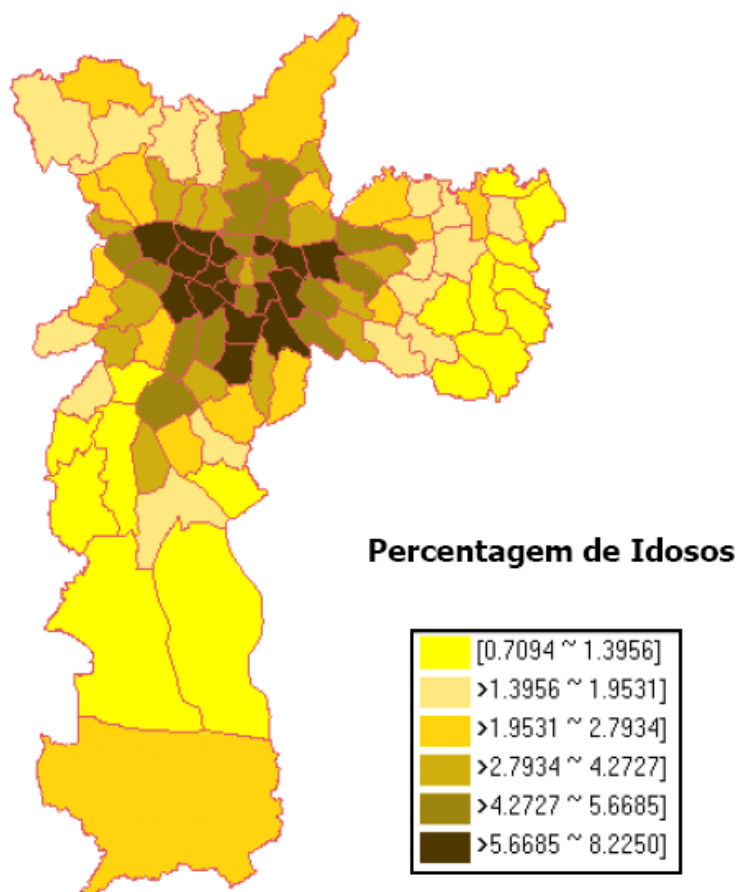
SUMÁRIO

	pg.
1. Introdução	04
2. Carregar os Dados no Sistema SPRING	05
3. Visualizar e Explorar a Tabela de Atributos	07
4. Técnicas de Agrupamento	10
4.1 Agrupamento por Passos Iguais	11
4.2 Agrupamento por Quantil	12
4.3 Agrupamento Estatístico	13
5. Técnicas de Autocorrelação Espacial	14
5.1 Índice Global de Moran	14
5.2 Diagrama de Espalhamento de Moran	16
5.3 Box Map	17
5.4 Média Espacial Móvel	18
5.5 Gráfico de Barras da Média Espacial Móvel versus Z	19
5.6 Índice Local de Associação Espacial (LISA)	20
5.7 LISA Map	21
5.8 Moran Map	22
5.9 Estatísticas Gi e Gi*	23
6. BIBLIOGRAFIAS DE APOIO	31

1. Introdução

Os dados utilizados neste roteiro referem-se ao índice de exclusão social para os distritos da cidade de São Paulo, extraídos do trabalho “Mapa de Exclusão/Inclusão Social na Cidade de São Paulo”, resultado de uma pesquisa multi-institucional liderada pela Profa. Aldaiza Sposati da PUC/SP.

O mapa abaixo ilustra um exemplo no qual se aplicou a técnica de agrupamento por quantil para a percentagem de idosos agregados por distritos.



Neste exemplo o agrupamento por quantil em seis partes é denominado de sextil, ou seja, o atributo de interesse (percentagem de idosos) é dividido em seis classes, tal que, cada classe tenha aproximadamente o mesmo número de distritos.

2. Carregar os Dados no Sistema SPRING

Inicialize o programa SPRING. Depois ative o banco de dados São Paulo, conforme os passos de 1 a 4 indicados na Figura 1.

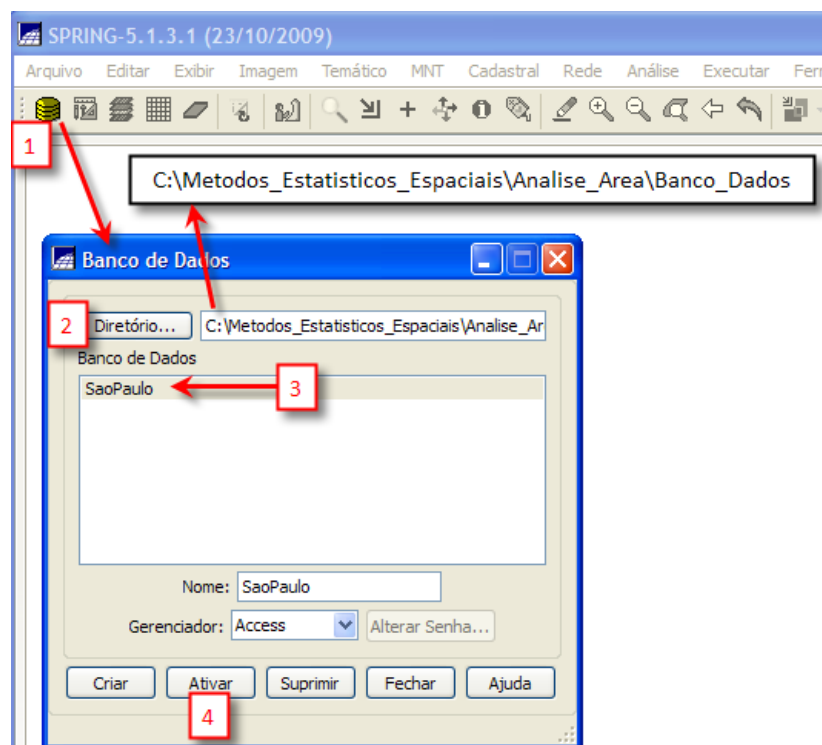


Figura 1

Ativar o projeto contido no banco de dados. Siga os passos de 1 a 3 conforme Figura 2.

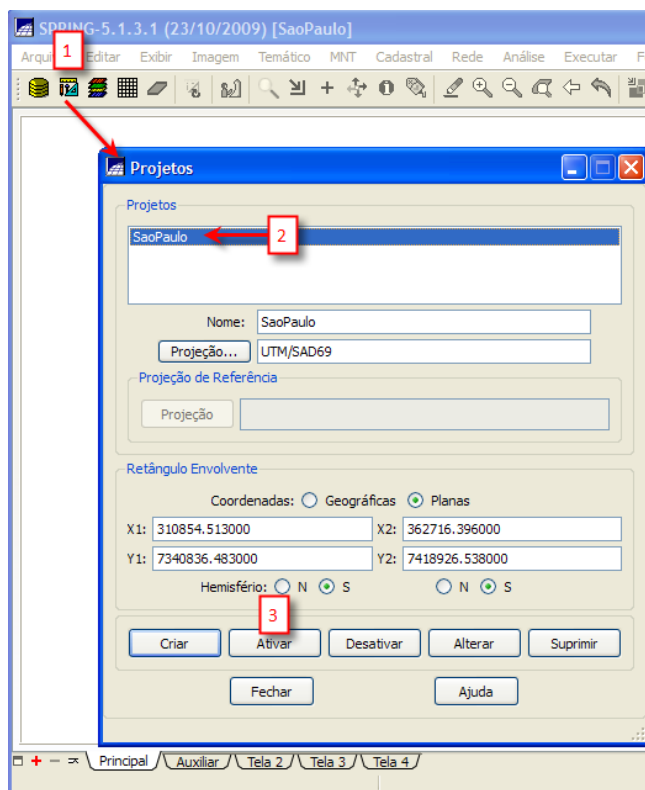


Figura 2

Visualizar os dados contidos no projeto ativo. Siga os passos 1 e 2 conforme indicados Figura 3.

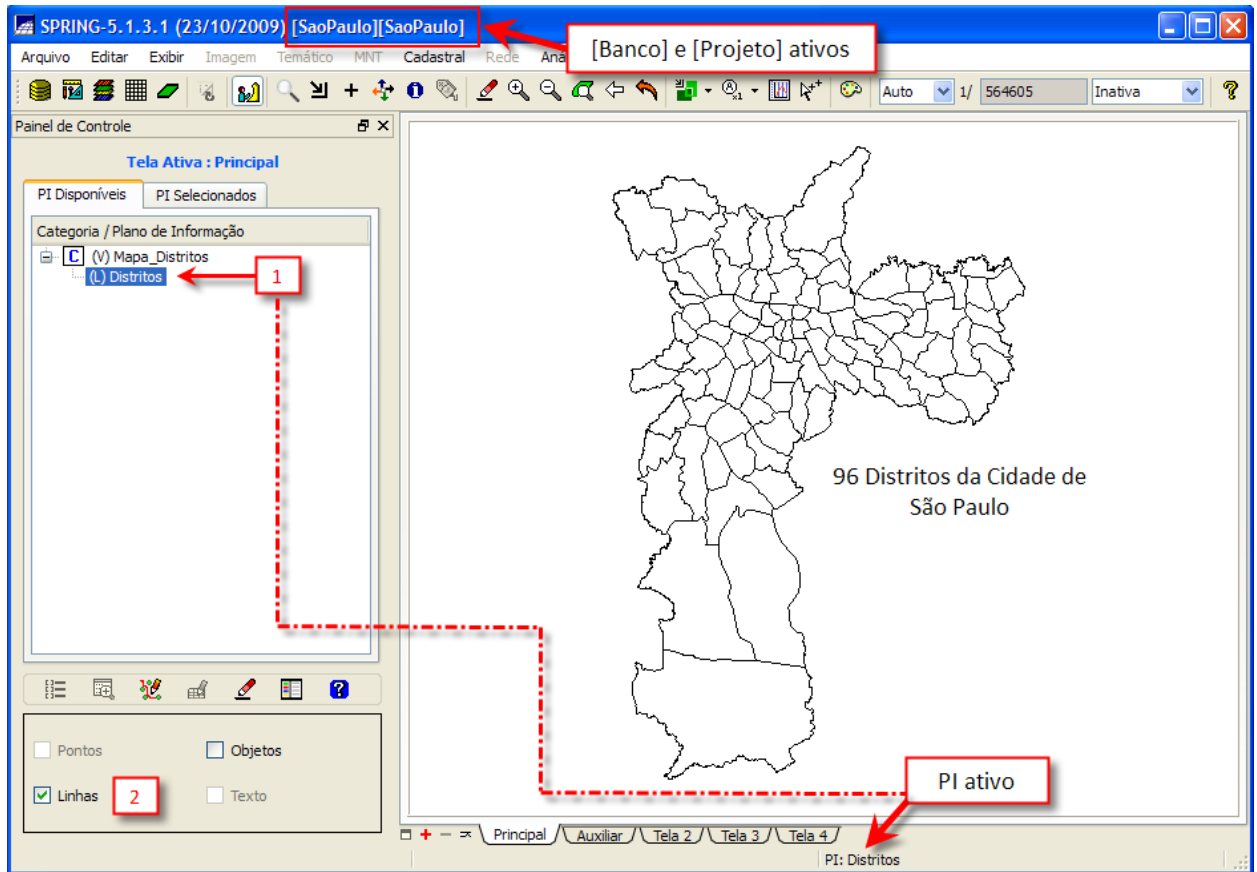



Figura 3

Algumas observações:

- 1) O banco de dados *SaoPaulo* possui duas categorias: *Mapa_Distritos* e *Distritos*, respectivamente. Isto pode ser constatado via *menu principal -> Arquivo -> Modelo de Dados...* ou pressionando o botão esquerdo do mouse sobre o ícone .
- 2) A categoria *Mapa_Distritos* é do tipo *Cadastral* e a categoria *Distritos* é do tipo *Objeto*. Na concepção do SPRING o tipo *Cadastral* deve ser entendido como “**o suporte para conter os Objetos**”. Assim, *Cadastral / Objeto* estão sempre interligados.
- 3) O modelo *Cadastral / Objeto* possui duas componentes:
 - i) *uma espacial*. Neste caso, refere-se ao mapa de distritos composto por 96 entidades do mundo real (os distritos);
 - ii) *outra não espacial*, isto é, uma tabela de atributos que contém informações armazenadas sobre cada um dos 96 distritos que compõem o mapa.

3. Visualizar e Explorar a Tabela de Atributos

Para visualizar a tabela de atributos associada aos *geo-objetos* (Distritos) siga os passos de 1 a 7 conforme indicados na Figura 4.

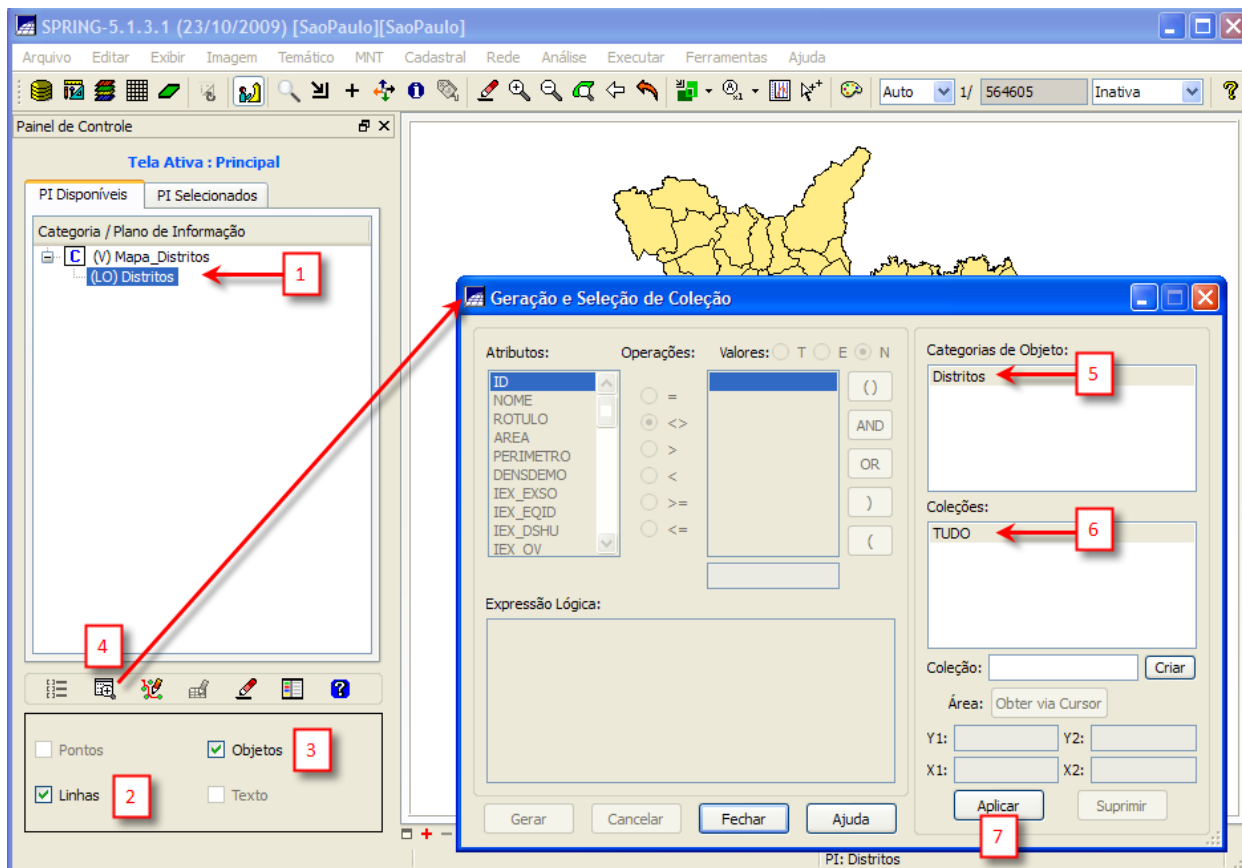


Figura 4

Após a execução do passo 7 leva-se à abertura da tabela de atributos, conforme mostra Figura 5 a seguir.

A Figura 5 ilustra a tela principal do SPRING com a tabela de atributos aberta (*Modo Consulta*).

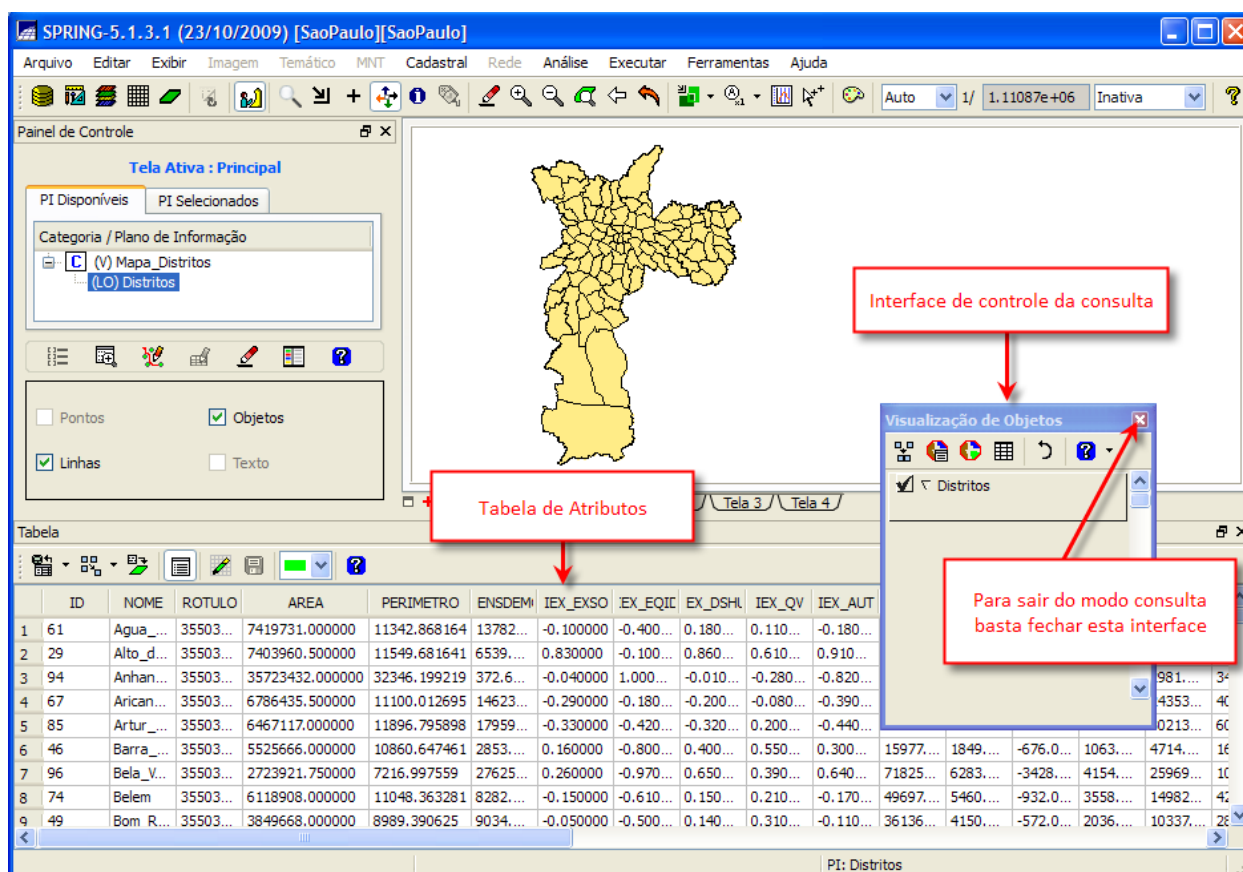


Figura 5

Notas:

- 1) A interface de *Visualização de Objetos*, mostrada na Figura 5, deve ser entendida como o **centro de controle quando o SPRING se encontra no modo consulta**. Esta interface possibilita várias operações:
 - a. agrupamento de atributo;
 - b. consulta por atributo;
 - c. consulta espacial;
 - d. exibir tabela de atributos caso a mesma esteja fechada;
 - e. desfazer consultas.
- 2) Quando a interface *Visualização de Objetos* é fechada o SPRING deixa o modo consulta, preservando o estado das últimas operações realizadas.
- 3) Sobre a tabela de atributos é possível realizar vários tipos de operações, conforme visto a seguir.

Explorando os recursos de seleção e operações sobre a tabela de atributos, conforme Figura 6.

Seleção Via Mapa
BE sobre geo-objeto -> seleciona linha da tabela.
Duplo Click -> mostra a tela de atributos do geo-objeto

Seleção de cor

BE -> seleciona linha da tabela

BE -> seleciona coluna

BD

BD

BE: pressionar botão esquerdo do mouse

BD: pressionar botão direito do mouse

ID	NOME	ROTULO	AREA	PERIMETRO	DENSDEMO	IEX_EXSO	IEX_EQID	IEX_DSHU	IEX_QV	IEX_AUT	POPULAC	DEMVAG
35	Brasilândia	35503011	20477584.000000	24871.447266	9599.571400	-0.830000	-0.620000	-0.760000	-0.270000	-0.790000	201591.000000	36854.000000
57	Ipiranga	35503033	10855735.000000	16771.753906	9669.809500	-0.020000	-0.460000	0.060000	0.410000	-0.080000	101533.000000	12352.000000
79	Penha	35503060	11502566.000000	18286.923828	11770.442500	-0.100000	-0.410000	0.390000	-0.020000	-0.240000	133006.000000	17970.000000
40	Perdizes	35503061	6373885.500000	10118.314453	1784.262300	0.410000	-0.380000	0.570000	0.780000	0.150000	108840.000000	13022.000000
48	Santana	35503071	13249961.000000	16725.835938	10926.904800	0.370000	-0.260000	0.350000	0.450000	0.470000	137679.000000	16033.000000
61	Água Rasa	35503001	7419731.000000	11342.868164	13782.463800	-0.100000	-0.400000	0.180000	0.110000	-0.180000	95099.000000	12214.000000
29	Alto de Pinheiros	35503002	7403960.500000	11549.681641	6539.090900	0.830000	-0.100000	0.860000	0.610000	0.910000	50351.000000	5820.000000
94	Anhanguera	35503003	35723432.000000	32346.199219	372.612600	-0.040000	1.000000	-0.010000	-0.280000	-0.820000	12408.000000	2295.000000
67	Aricanduva	35503004	6786435.500000	11100.012695	14623.030300	-0.290000	-0.180000	-0.200000	-0.080000	-0.390000	96512.000000	15226.000000

PI: Distritos

Marcas
 Remover
 Mover
 Zoom
 Atributos...

Operações
 Estatísticas...
 Ordenar
 Gráfico...
 Ocultar Coluna...
 Exibir Colunas...

Figura 6

4. Técnicas de Agrupamento

As técnicas de agrupamento controlam a forma de visualizar os objetos geográficos baseado nos valores de seus atributos. O SPRING possibilita 4 tipos de agrupamento: i) passo igual, ii) quantil, iii) estatístico e iv) valor único.

Antes de iniciar um agrupamento sugere-se executar a seguinte ação: *menu principal -> Executar -> Limpar -> Tudo*. Esta ação garante limpar qualquer tipo de apontamento e/ou consulta realizada anteriormente. Continuando, para executar um dos tipos de agrupamento pressione o botão esquerdo do mouse sobre o primeiro ícone, conforme indicado na Figura 7. Esta ação leva à abertura da interface *Agrupar Objetos*, conforme a Figura 8.

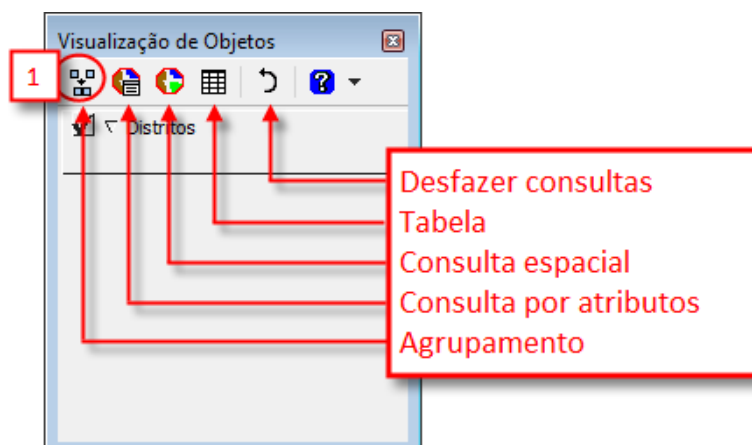


Figura 7

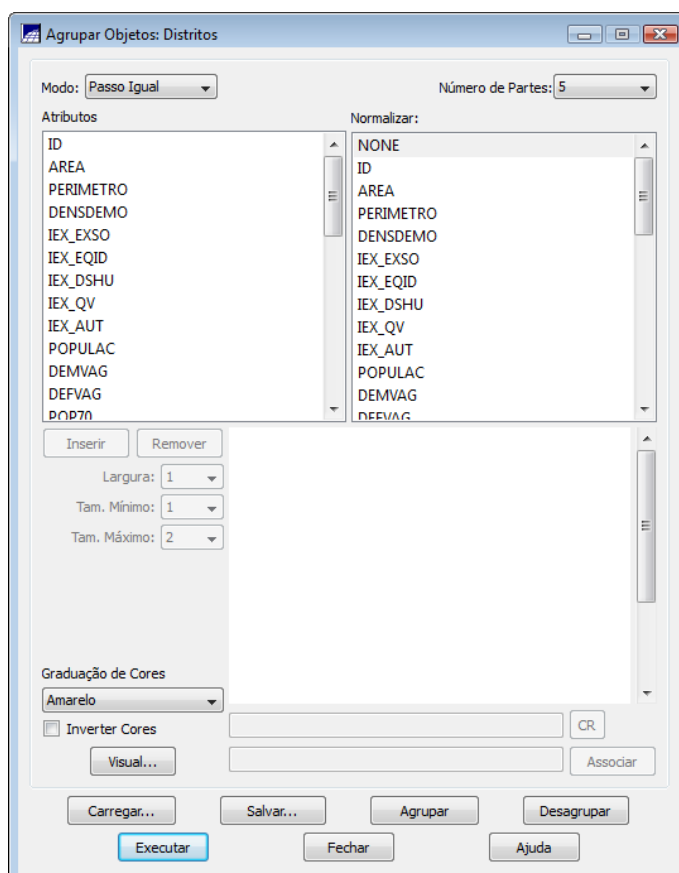


Figura 8

4.1 Agrupamento por Passos Iguais

Para realizar o agrupamento por passos iguais sobre o atributo PERIDOSO (percentagem de idosos) siga os passos de 1 a 6 conforme indicados na Figura 9.

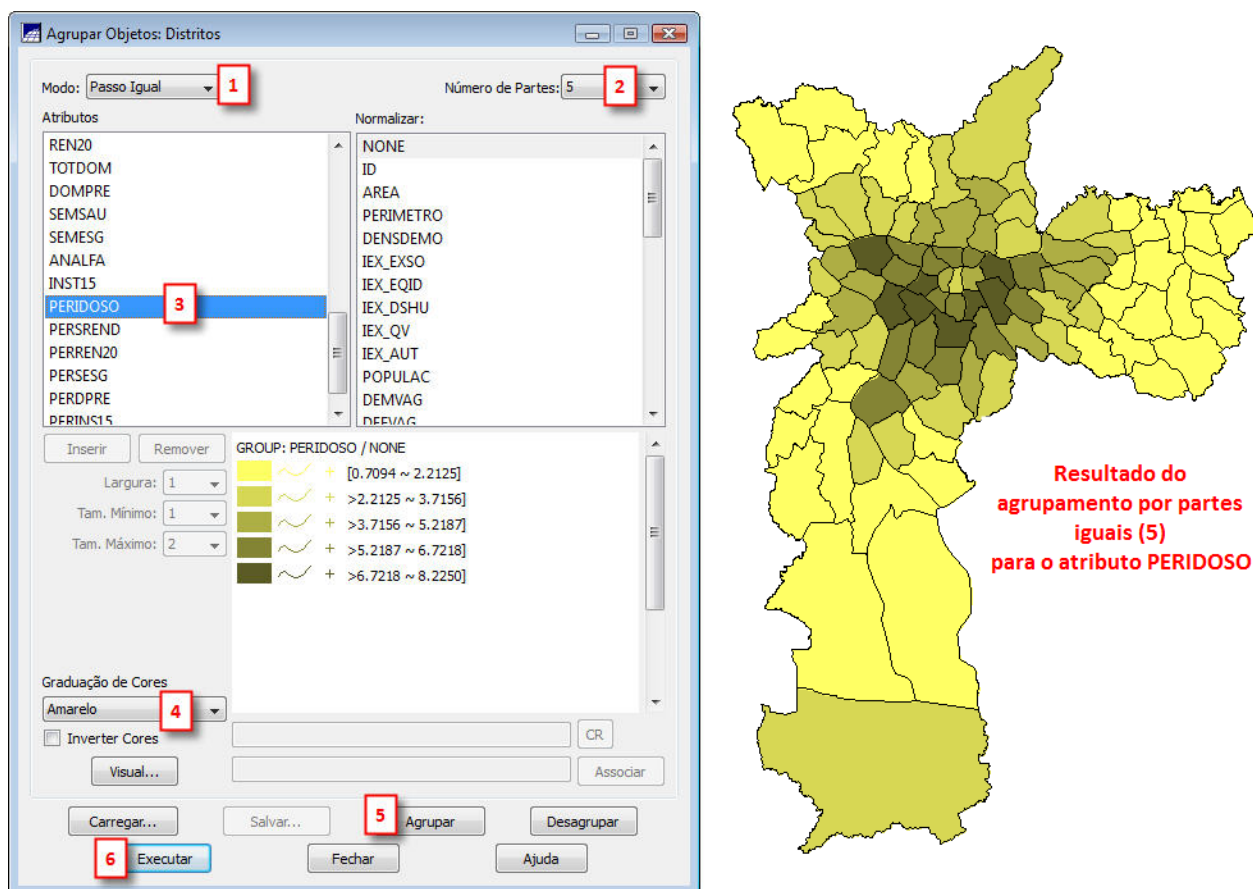


Figura 9

Nota:

Esta forma de agrupamento gera n grupos contendo o mesmo *intervalo* de valores em cada grupo. Seu cálculo é feito da seguinte maneira:

$$\text{Intervalo} = (V_{\max} - V_{\min}) / n$$

em que:

- *Intervalo*: refere-se ao intervalo de valores de cada grupo. No exemplo acima o intervalo é da ordem de 1.5 para cada grupo (vide a escala de cores da Figura 9);
- V_{\max} : é o valor máximo do atributo (neste exemplo o atributo é o PERIDOSO);
- V_{\min} : é o valor mínimo do atributo (idem ao anterior);
- n : é o número de grupos (equivalente ao *Número de Partes* – vide passo 2 da Figura 9).

4.2 Agrupamento por Quantil

Para realizar o agrupamento por quantil sobre o atributo PERIDOSO (percentagem de idosos) siga os passos de 1 a 6 conforme indicados na Figura 10.

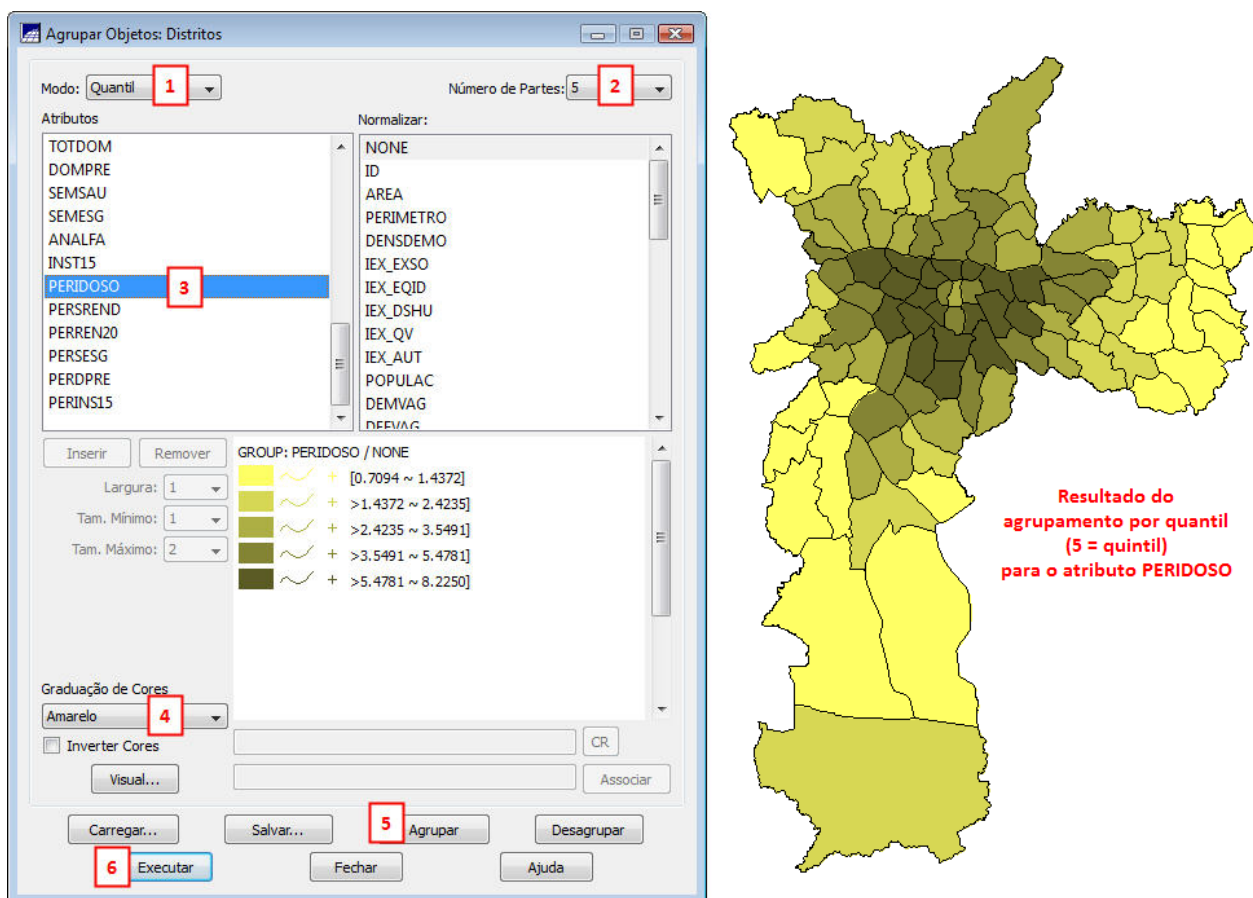


Figura 10

Nota:

Esta técnica de agrupamento gera n grupos contendo o mesmo número de objetos em cada grupo. Primeiro ordena-se os valores do atributo selecionado. Depois, calcula-se o número de objetos (N_{objs}) em cada grupo, que é igual ao número total de objetos (N_{total}) dividido pelo número de grupos (n) que se deseja. Assim: $N_{objs} = N_{total} / n$, em que:

- N_{objs} : refere-se ao número de objetos de cada grupo. No exemplo, $96/5 \cong 19$.
- N_{total} : é o número total de objetos. No exemplo é igual a 96 (distritos).
- n : é o número de grupos. No exemplo é igual a 5.

O $Range$ de cada grupo é o intervalo de valores encontrados entre os objetos localizados na disposição de ordenamento. Assim: $sort(n-1) \geq Range(n) < sort(n)$, em que:

- $Range(n)$ = intervalo do n -ésimo grupo.
- $sort(n-1)$ = valor do atributo relativo à ordem $n-1$ do ordenamento.
- $sort(n)$ = valor do atributo relativo à n -ésima ordem do ordenamento.

4.3 Agrupamento Estatístico

Para realizar o agrupamento estatístico sobre o atributo PERIDOSO (percentagem de idosos) siga os passos de 1 a 6 conforme indicados na Figura 11.

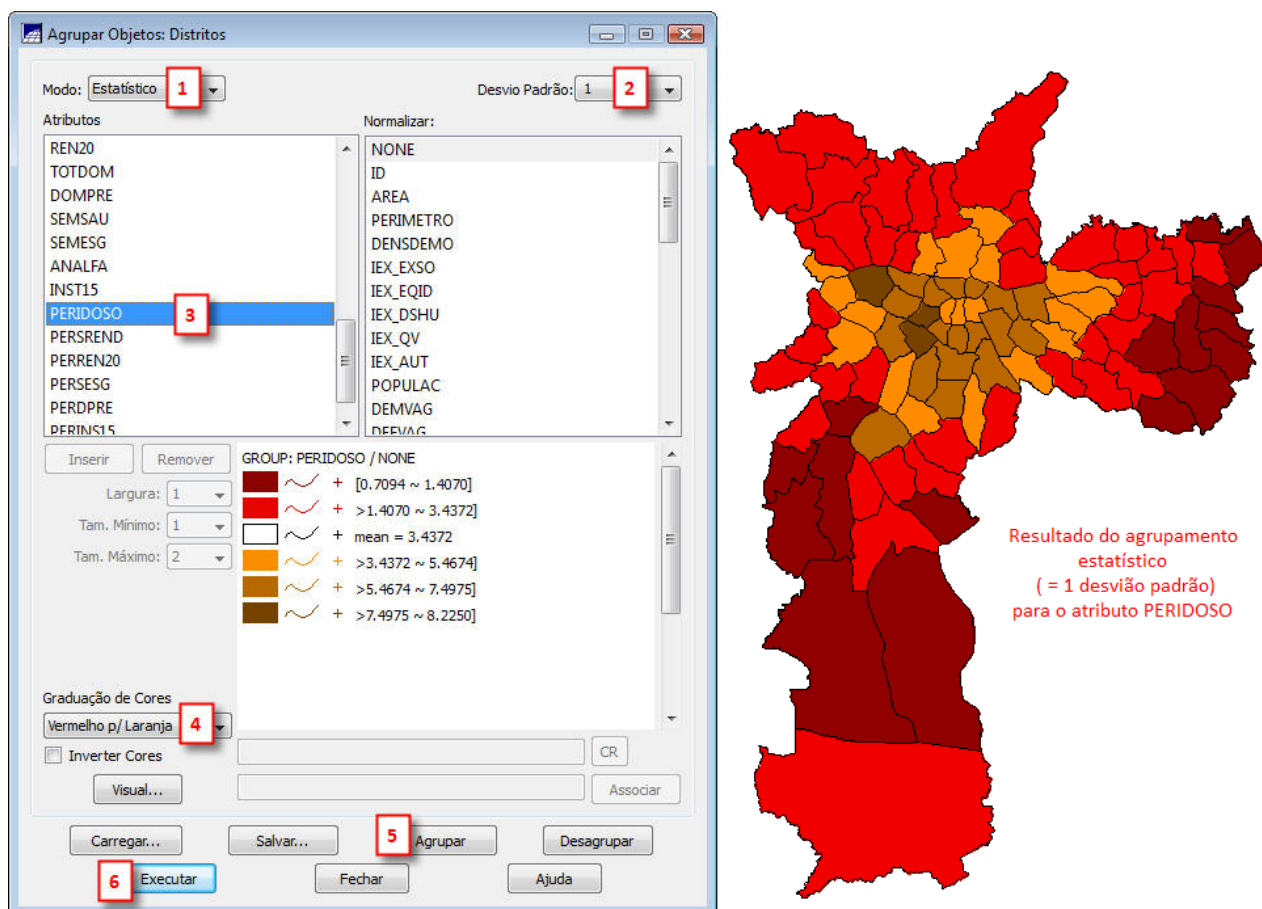
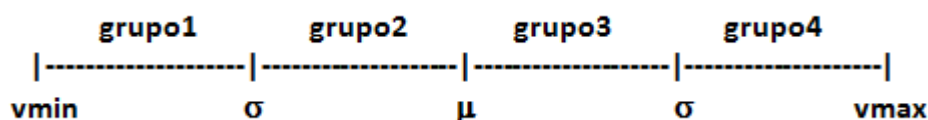


Figura 11

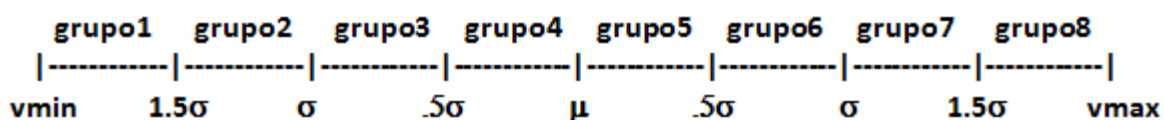
Nota:

Esta forma de agrupamento gera grupos que são separados em intervalos de valores que podem ser: um desvio padrão (σ), 0.5σ ou 0.25σ . Neste caso, o principal divisor é a média (μ). A partir da média é que os grupos são separados, acima e abaixo da média. O número de grupos gerados acima e abaixo da média dependerá do valor da quantidade de desvio padrão escolhido.

Por exemplo, para intervalo igual a um desvio padrão (σ), tem-se:



Outro exemplo, para intervalo igual a meio desvio padrão ($.5\sigma$), tem-se:



5. Técnicas de Autocorrelação Espacial

5.1 Índice Global de Moran

De uma forma geral, o índice global de Moran (I) presta-se a um teste cuja hipótese nula é de independência espacial. Neste caso, seu valor seria zero. Por outro lado, valores positivos (entre 0 e +1) indicam para correlação direta e negativos (entre 0 e -1) indicam para correlação inversa.

Uma vez calculado I , é importante estabelecer sua validade estatística. Em outras palavras, será que o valor calculado representa correlação espacial significativa? Uma forma de avaliação de I é através do *teste de pseudo-significância*. Neste caso, são geradas diferentes permutações dos valores de atributos associados às regiões; cada permutação produz um novo arranjo espacial, onde os valores estão redistribuídos entre as áreas. Como apenas um dos arranjos corresponde à situação observada, pode-se construir uma distribuição empírica de I . Se o valor do índice I medido originalmente corresponder a um “extremo” da distribuição simulada, então se trata de valor com significância estatística.

Para computar o índice global de Moran siga os passos de 1 a 10 conforme ilustrados na Figura 12.

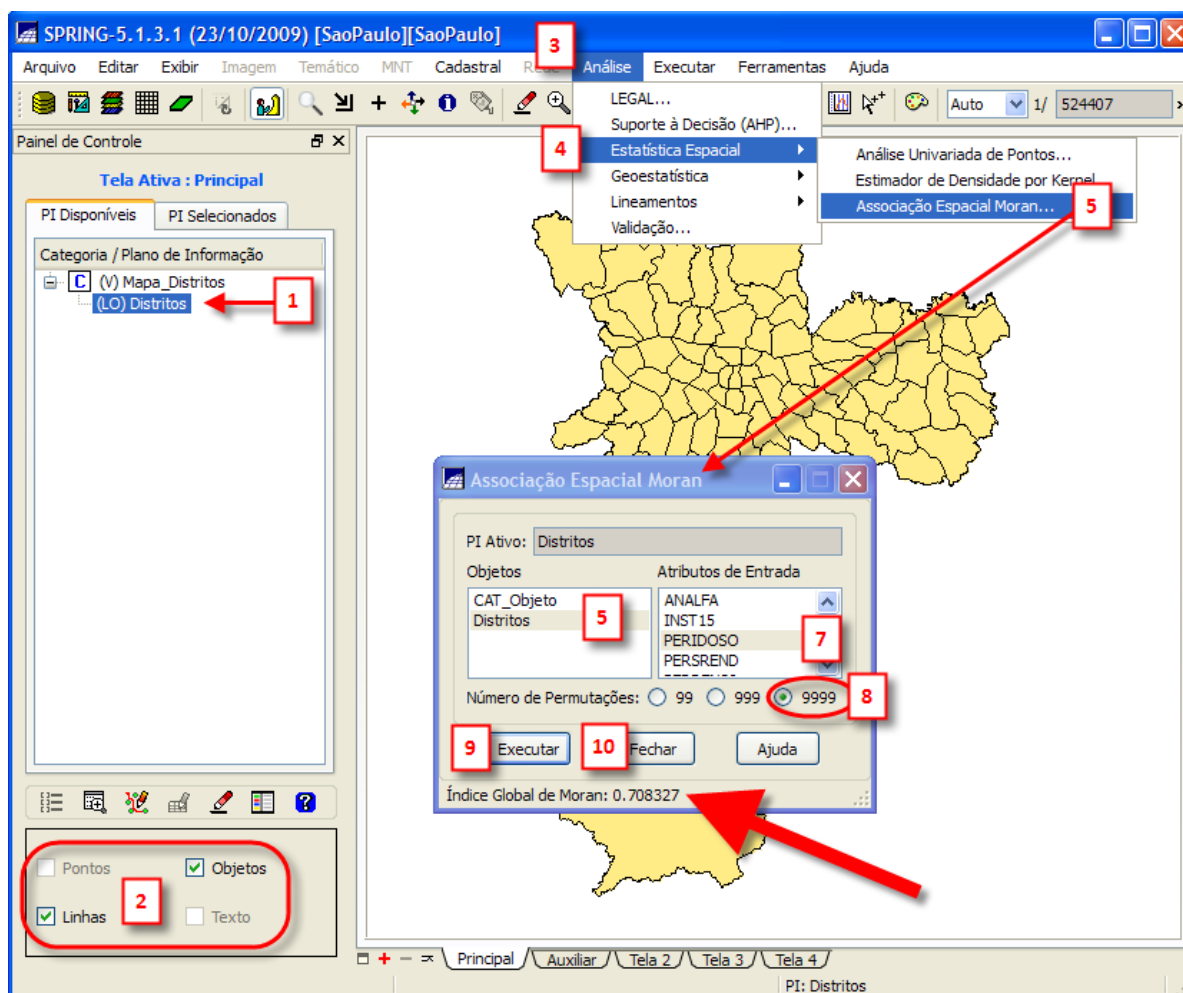


Figura 12

O valor do índice global de Moran é reportado no rodapé da tela *Associação Espacial Moran*, conforme mostra a Figura 12. Além deste valor outros índices são calculados e armazenados na tabela de atributos, conforme apresentados a seguir.

Para visualizar os indicadores armazenados na tabela de atributos siga os passos de 1 a 3, conforme indicados na Figura 13.

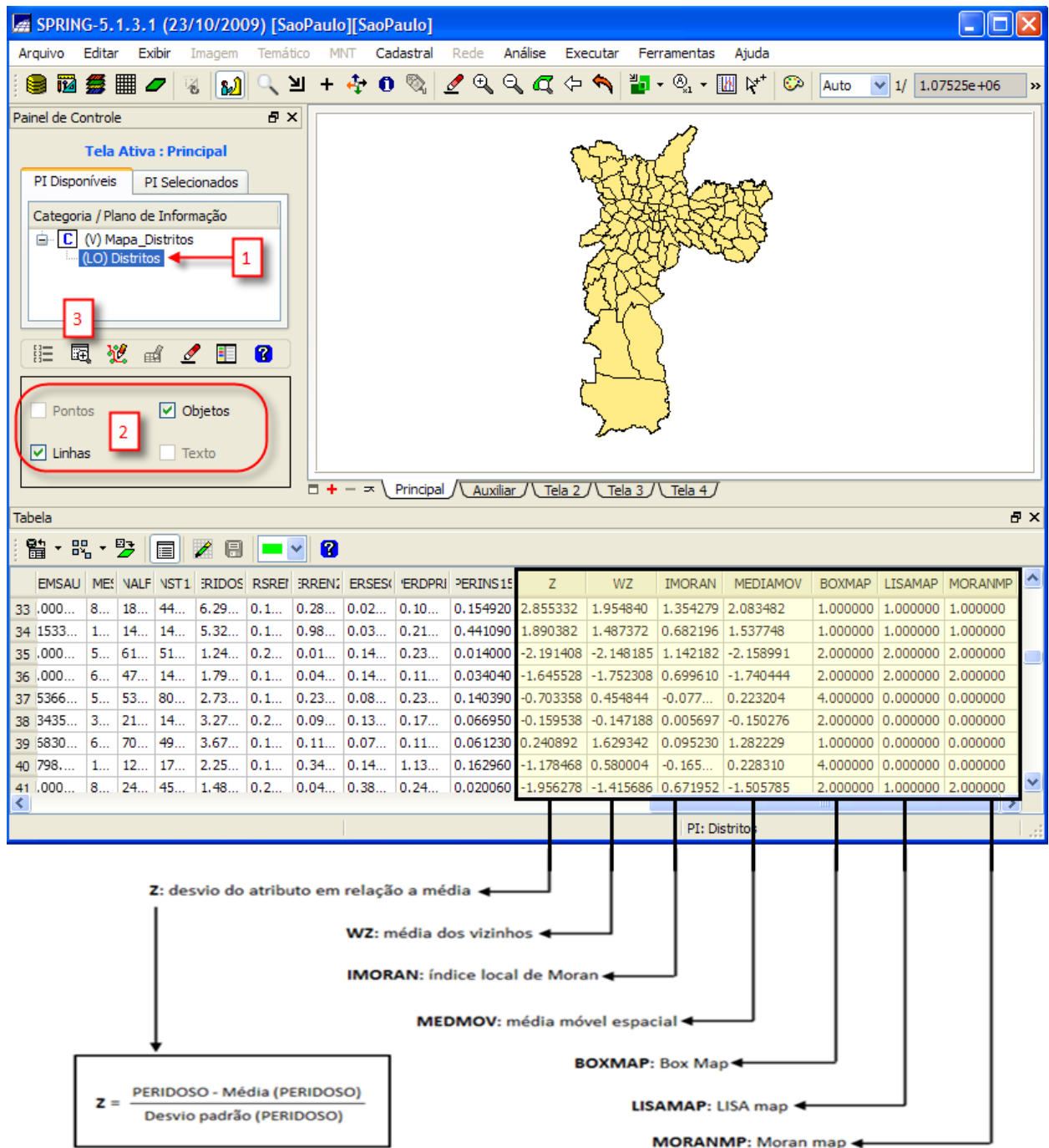


Figura 13.

5.2 Diagrama de Espalhamento de Moran

O diagrama de espalhamento de Moran é uma maneira adicional de visualizar a dependência espacial. A ideia é comparar os valores normalizados do atributo numa área com a média dos seus vizinhos, construindo um gráfico bidimensional de **Z** (valores normalizados) por **WZ** (média dos vizinhos), que é dividido em quatro quadrantes:

- **Q1** (valores positivos, médias positivas) e **Q2** (valores negativos, médias negativas): indicam pontos de associação espacial positiva, no sentido que uma localização possui vizinhos com valores semelhantes.
- **Q3** (valores positivos, médias negativas) e **Q4** (valores negativos, médias positivas): indicam pontos de associação espacial negativa, no sentido que uma localização possui vizinhos com valores distintos.

Para visualizar o diagrama de espalhamento de Moran siga os passos de 1 a 3, conforme indicados na Figura 14.

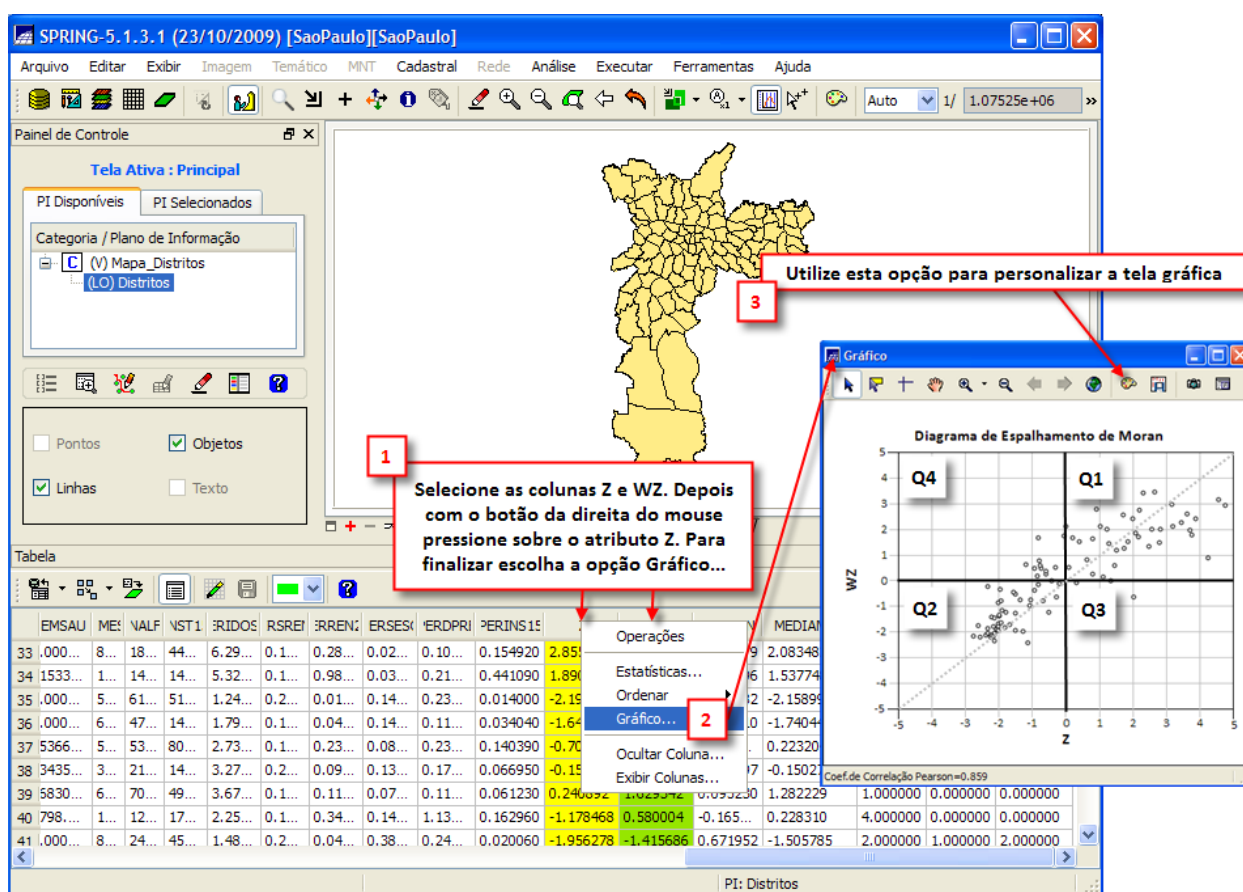


Figura 14

5.3 Box Map

O box map é uma forma alternativa de apresentar o diagrama de espalhamento de Moran. Neste caso, os valores são exibidos na forma de mapa, no qual cada polígono é classificado conforme a posição de seu valor em relação aos quadrantes Q_1 , Q_2 , Q_3 e Q_4 . Para cada quadrante há uma cor correspondente no mapa a ser gerado.

Para visualizar o resultado do box map siga os passos de 1 a 5 conforme indicados na Figura 16.

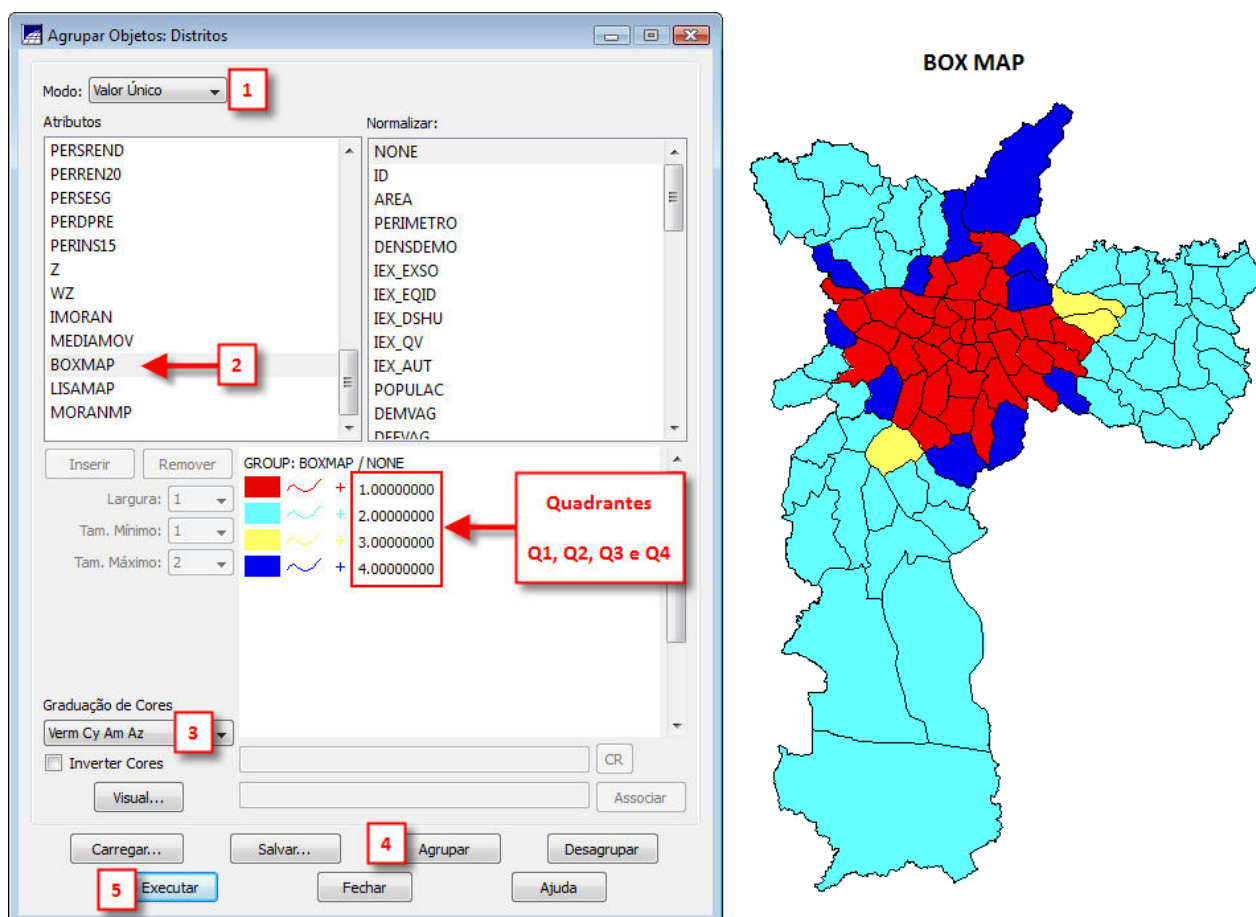


Figura 16

Nota:

Após a visualização do mapa, desfazer o agrupamento. Para isto, pressione o botão *Desagrupar* e depois o botão *Executar*.

5.4 Média Espacial Móvel

Uma forma simples de explorar a variação da tendência espacial dos dados é calcular a média dos valores dos vizinhos. Isto reduz a variabilidade espacial, pois este método tende a produzir uma superfície com menor flutuação que os dados originais. Para visualizar o resultado da média espacial móvel siga os passos de 1 a 6 conforme indicados na Figura 17.

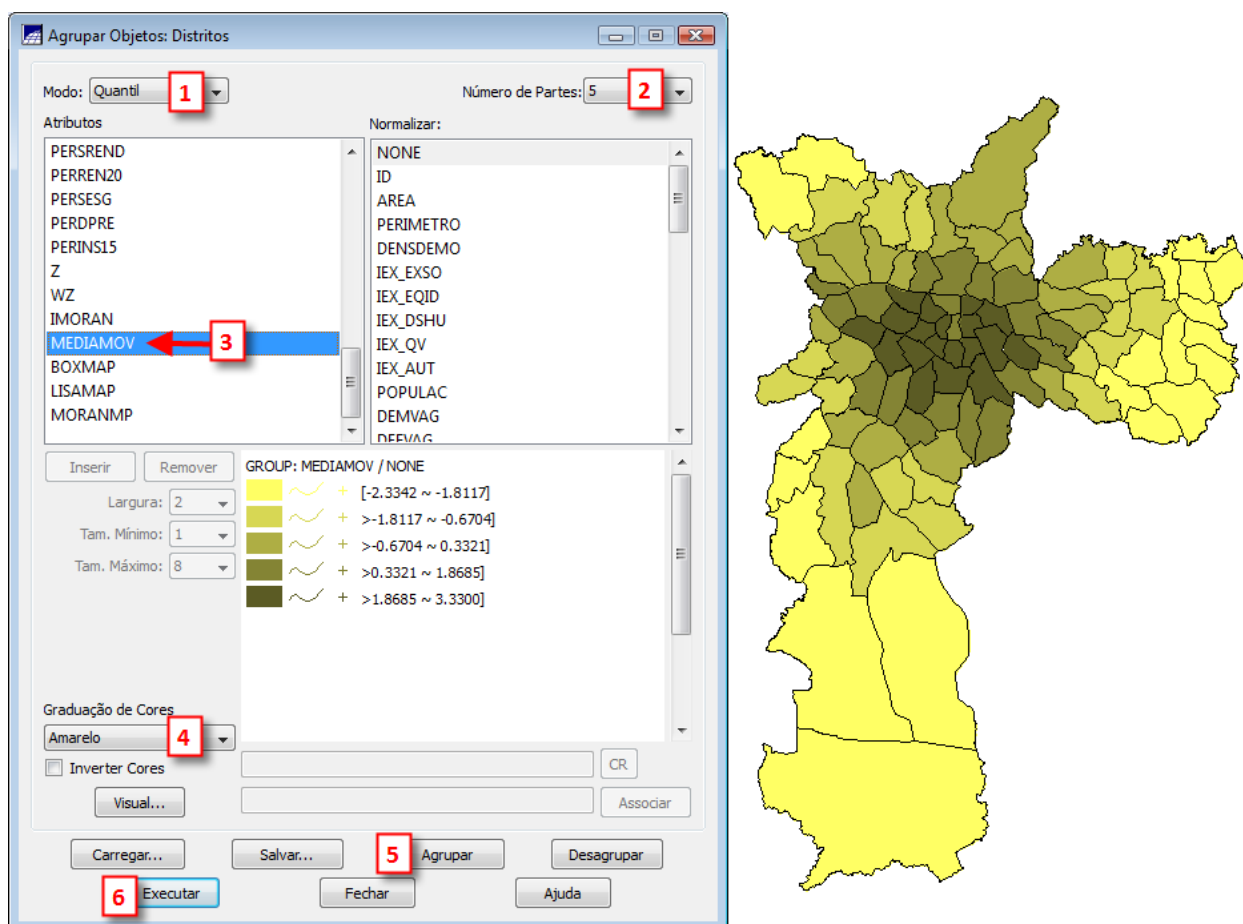


Figura 17

Notas:

- A Figura 16 apresenta o resultado do estimador da média espacial móvel para o percentual de idosos (mais de 70 anos) para os 96 distritos da cidade de São Paulo. Estes dados são indicadores da grande disparidade social da cidade, mostrando uma grande variação entre a região central e as periferias. A região central aponta para um máximo de percentual de idosos da ordem de 6,8% [que é a média do atributo *PERIDOSO* (3.43) mais o último valor da escala de cores (3.33)]. As periferias apontam para um máximo de percentual de idosos por volta de 1% [que é a média do atributo *PERIDOSO* (3.43) menos o primeiro valor da escala de cores (-2.33), o sinal negativo significa abaixo da média].
- Após a visualização do mapa, desfazer o agrupamento. Para isto, pressione o botão *Desagrupar* e depois o botão *Executar*.

5.5 Gráfico de Barras da Média Espacial Móvel versus Z

Outra forma bastante didática de apresentar a média espacial móvel é por meio de um gráfico de barras, em que comparamos o valor do atributo (Z) com sua média local, como mostrado na Figura 15. Neste caso, podemos considerar que regiões onde o valor do atributo é semelhante à sua média local (sejam estes valores altos ou baixos) são pontos de autocorrelação espacial positiva. Já os casos de sub-áreas com grande disparidade entre o valor do atributo e os valores da média local indicam pontos de transição entre regimes espaciais, conforme indicado por círculo sobre o mapa da Figura 15.

Para executar o gráfico de barras Z versus a média espacial móvel siga os passos de 1 a 6 conforme indicados na Figura 15.

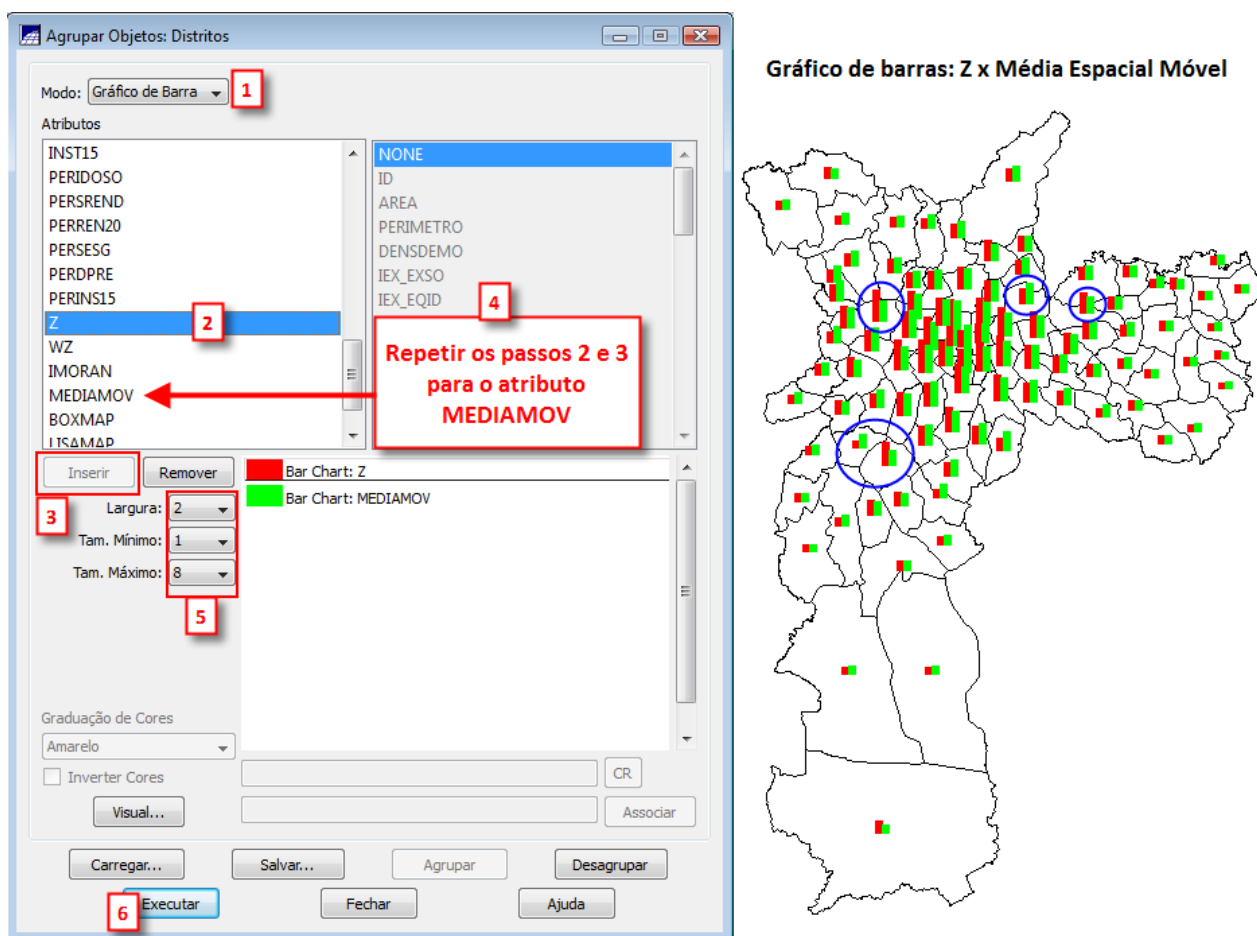


Figura 15

Nota:

Para desfazer o gráfico de barras:

- i) Primeiro selecione com “duplo click” (botão esquerdo do mouse) o elemento *Bar Chart: Z*;
- ii) Depois pressione o botão *Remove*.
- iii) Repita os dois passos anteriores para remover o outro elemento *Bar Chart: MEDIAMOV*.
- iv) Removido os dois elementos pressione o botão *Executar*.

5.6 Índice Local de Associação Espacial (LISA)

Como vimos anteriormente o Estimador de Autocorrelação Espacial Global, Moran (I), fornece um valor único como medida da associação espacial. Por outro lado, muitas vezes é necessário examinar padrões numa escala mais detalhada. Neste caso, é preciso utilizar indicadores locais de associação espacial que possam ser associados a diferentes localizações de uma variável distribuída espacialmente.

Os Indicadores Locais de Associação Espacial produzem um valor específico para cada objeto (polígono). Isto acarreta a identificação de:

- “Clusters”: objetos com valores de atributos semelhantes;
- “Outliers”: objetos anômalos;
- A presença de mais de um regime espacial.

Na geração do LISA é avaliada a significância dos valores do índice de Moran Local obtido para cada objeto (polígono), em relação à hipótese de não existência de autocorrelação espacial (hipótese nula). Na avaliação da significância é utilizada a abordagem de permutação dos atributos dos vizinhos, conforme descrito em Anselin (1995).

Uma vez determinada a significância estatística do Índice Local de Moran, é muito útil gerar um mapa indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados. Este mapa é chamado por Anselin (1995) de **LISA Map**, e na sua geração, os valores do Índice Local de Moran são classificados em quatro grupos:

- Não significantes;
- Com significância entre 0,05 e 0,01;
- Com significância entre 0,01 e 0,001;
- Com significância maior que 0,001.

O emprego de Indicadores Locais em conjunto com Indicadores Globais refina nosso conhecimento sobre os processos que dão origem a dependência espacial (Anselin, 1995).

5.7 LISA Map

Para visualizar o LISA Map siga os passos de 1 a 7 conforme indicados na Figura 18.

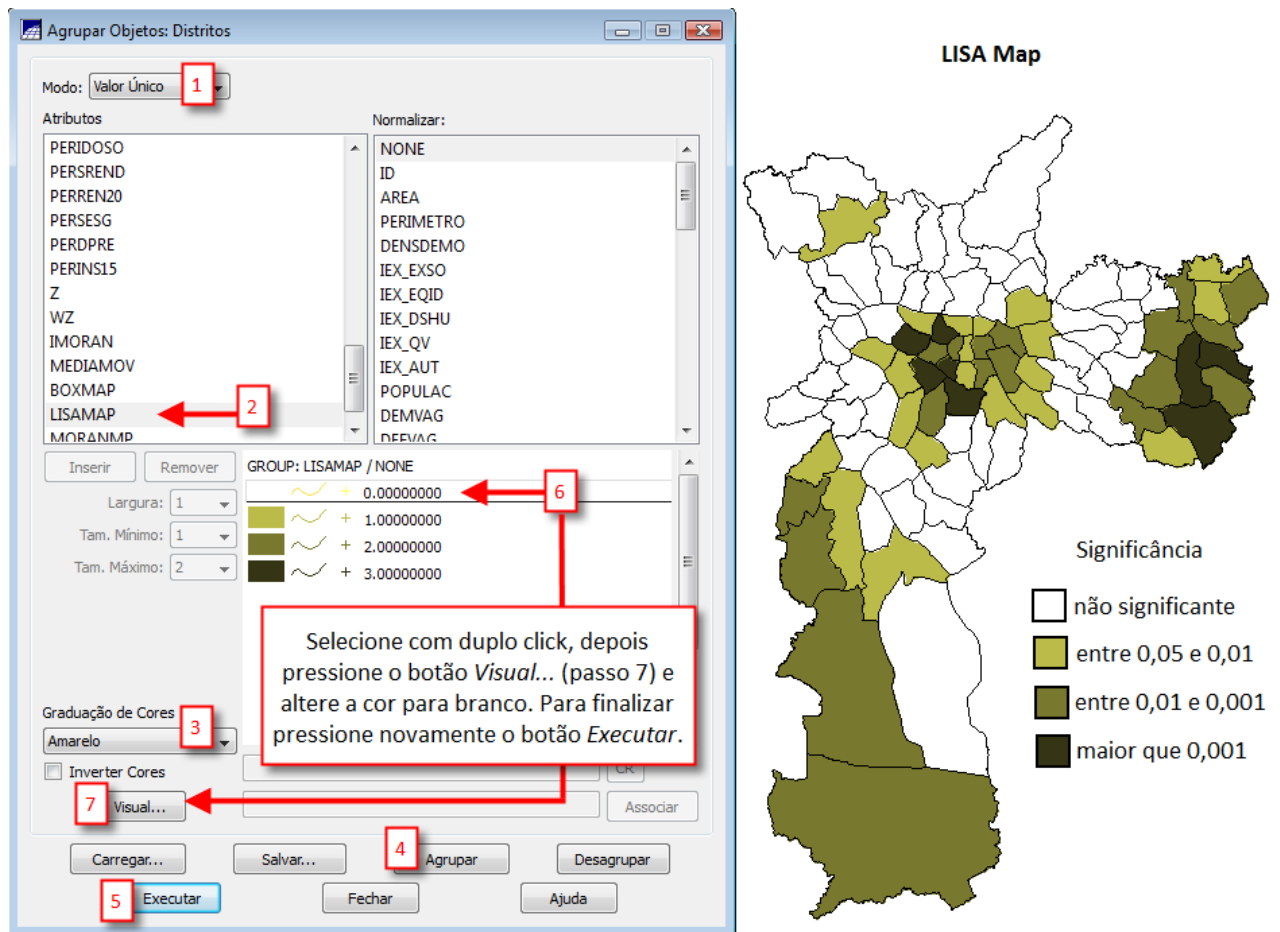


Figura 18

Notas:

- Neste exemplo, o resultado do LISA Map indica claramente uma forte polarização centro-periferia do atributo PERIDOSO, indicando a presença de “bolsões”. São áreas com dinâmica espacial própria e que merecem análise detalhada.
- Após a visualização do mapa, desfazer o agrupamento. Para isto, pressione o botão *Desagrupar* e depois o botão *Executar*.

5.8 Moran Map

No Moran Map somente os objetos para os quais os valores de LISA foram considerados significantes ($p \leq 0,05$), são destacados, porém, aparecem classificados em quatro grupos, conforme sua localização no quadrante do gráfico de espalhamento. Os demais objetos ficam classificados como *sem significância*.

Para visualizar o Moran Map siga os passos de 1 a 7 conforme indicados na Figura 19.

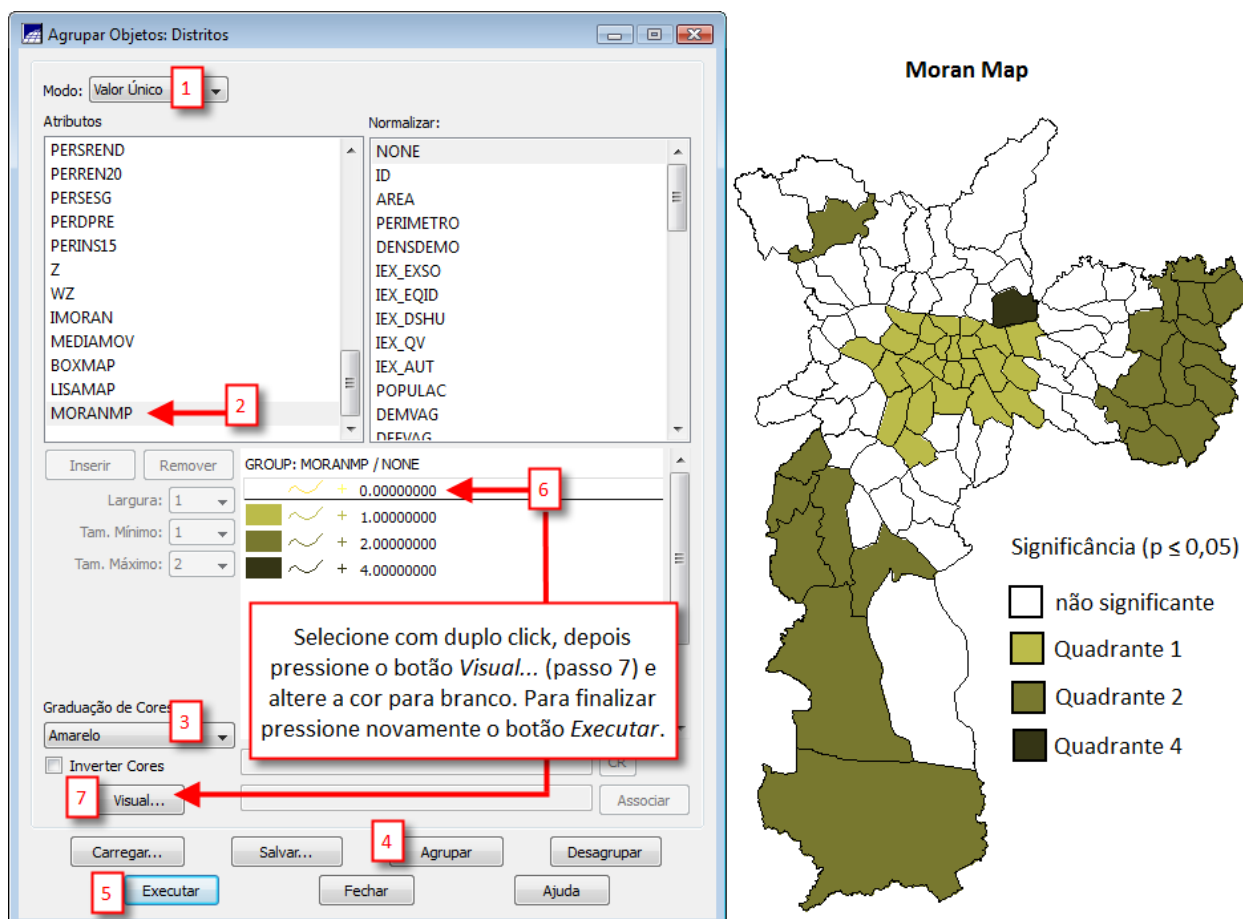


Figura 19

5.9 Estatísticas G_i e G_i^*

As estatísticas G_i e G_i^* foram introduzidas por Getis e Ord (1992) como formas alternativas de medidas de associação espacial. São baseadas em distâncias estatísticas e calculadas a partir de um conjunto de vizinhos para cada localização. Essas medidas indicam o tamanho para o qual uma localização é rodeada por um “cluster” de valores altos ou baixos para a variável em consideração.

Formalmente são definidas como:

$$G_i = \frac{\sum_{j=1}^n w_{ij}(d) x_j}{\sum_{i=1}^n x_i}, j \neq i \qquad G_i^* = \frac{\sum_{j=1}^n w_{ij}(d) x_j}{\sum_{i=1}^n x_i}, j = i$$

em que:

- w_{ij} : é um valor na matriz de proximidade para região i com a região j em função da distância (d);
- x_i e x_j : são os valores dos atributos considerados nas áreas i e j ;
- d : é distância entre pontos;
- n : é o número de áreas (polígonos).

NOTA: a estatística G_i inclui no numerador a soma dos valores de todos vizinhos dentro de uma distância (d) do ponto considerado. G_i^* difere de G_i por incluir a localização visitada.

A partir deste ponto as medidas de associação espacial G_i e G_i^* serão calculadas utilizando o software **SpaceStat**. Depois os resultados são visualizados em forma de mapas no sistema SPRING. Para isto, execute os passos seguintes.

5.9.1 - Ativar Banco SaoPaulo e Carregar Projeto SaoPaulo

5.9.2 – Selecione o plano de informação *Distritos* e depois exportar em formato SpaceStat a Tabela de Atributos e a Matriz de Proximidade Espacial. Isto é realizado conforme as Figuras 20.

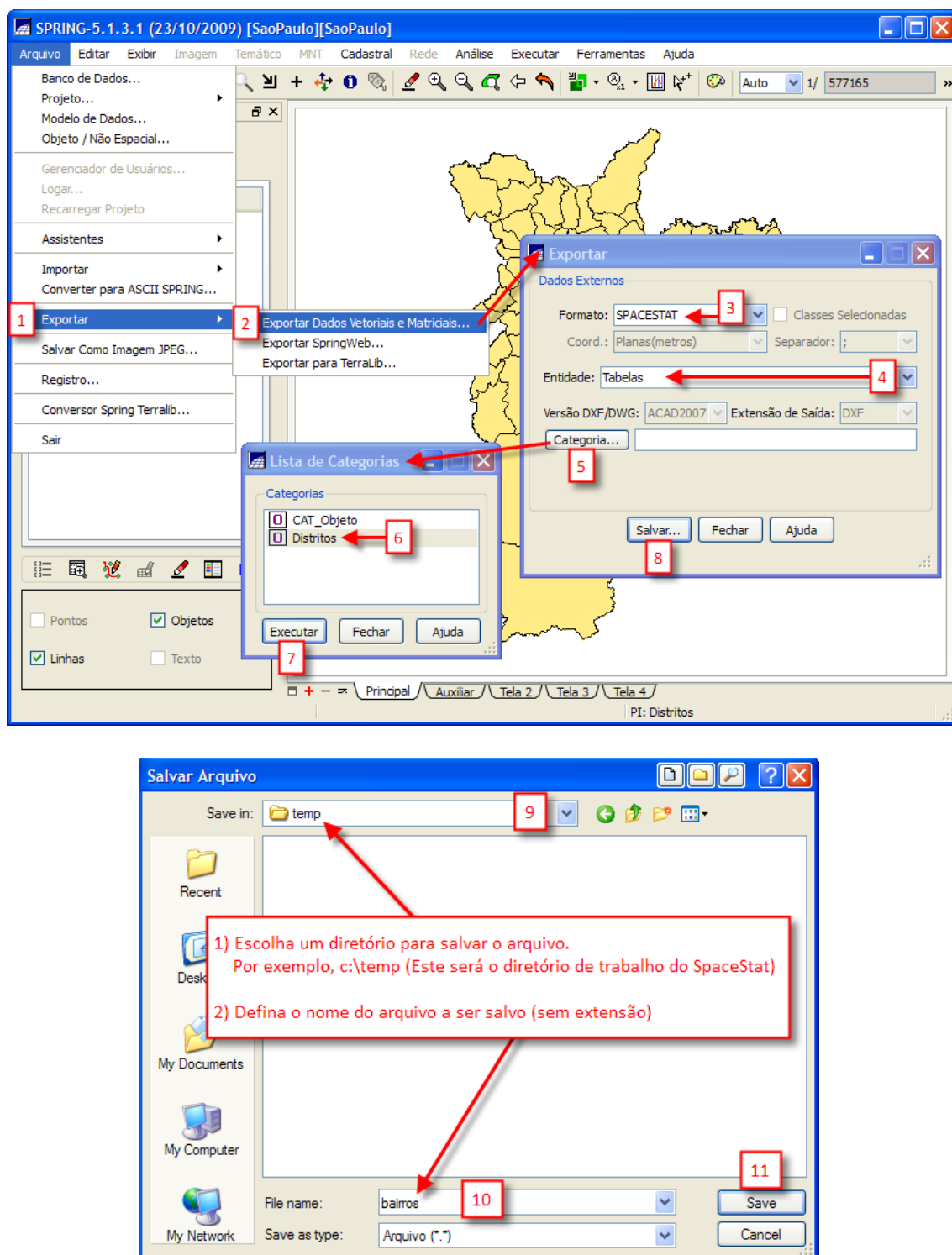


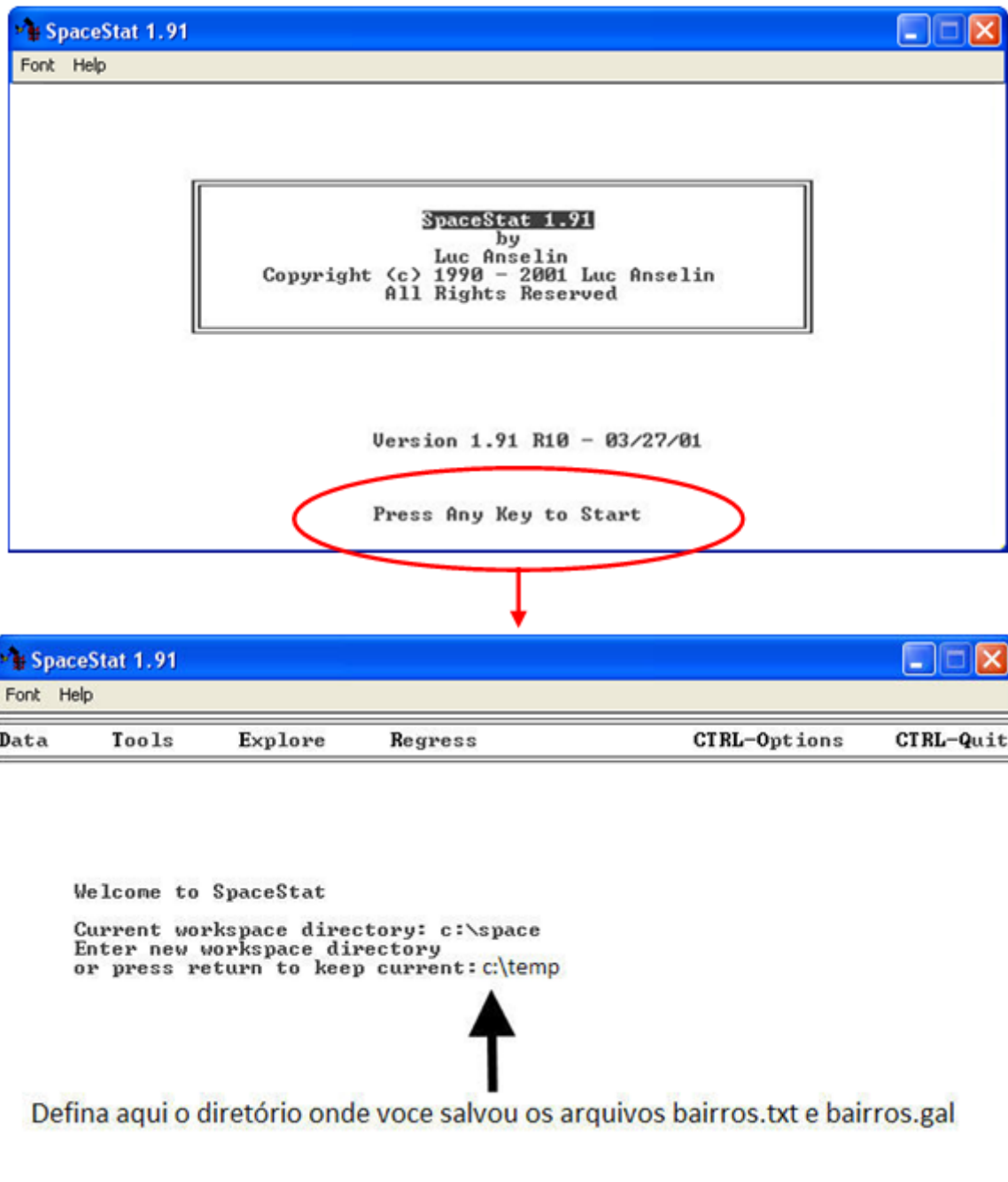
Figura 20.

Repita os passos de 4 a 11. No passo 4 empregue a opção *Entidade: Matriz de proximidade*. Após a execução do passo 11 dois arquivos são criados: **bairros.txt** (a Tabela de Atributos) e **bairros.gal** (a Matriz de Proximidade Espacial). A partir deste ponto têm-se os arquivos necessários para calcular as estatísticas G_i e G_i^* no SpaceStat.

5.9.3 - Minimizar Spring

5.9.4 - Inicializar o SpaceStat

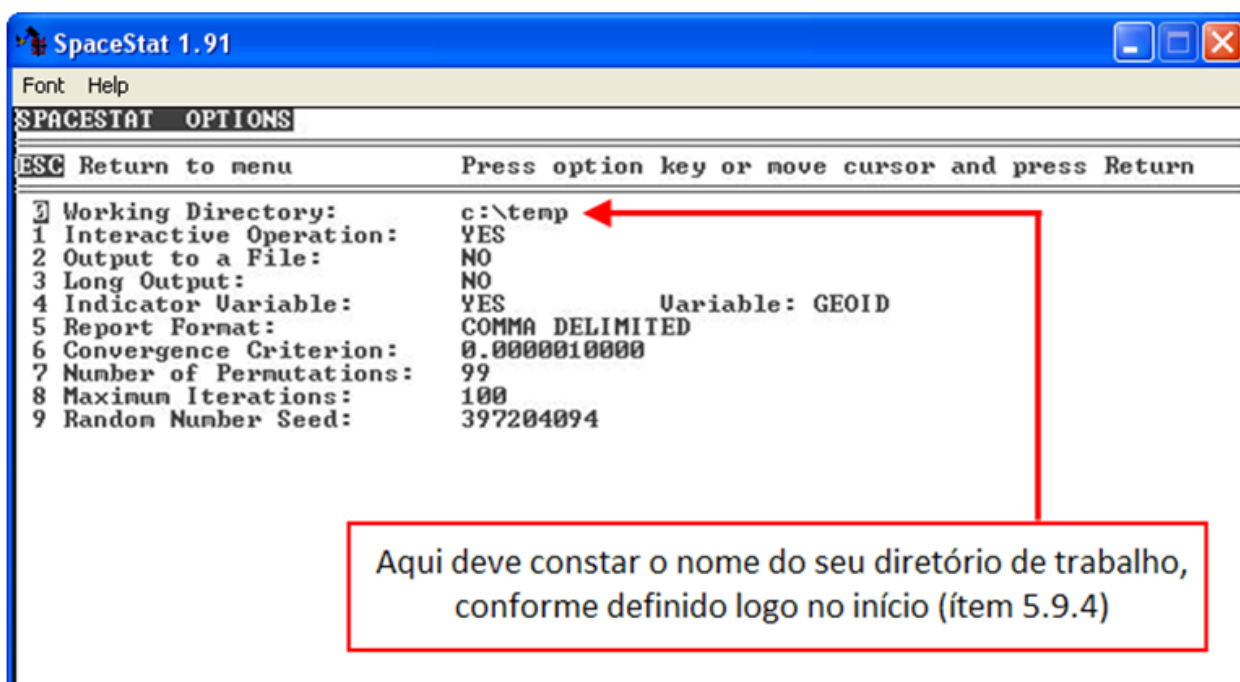
Depois que o SpaceStat é inicializado surgirá a tela abaixo. Pressione qualquer tecla para prosseguir.



No exemplo da Figura acima o diretório de trabalho é C:\temp. O diretório de trabalho pode ser qualquer e deve conter os arquivos bairros.txt e bairros.gal exportados do SPRING. Vale lembrar que o arquivo bairros.txt é a tabela de atributos e o arquivo bairros.gal é a matriz de proximidade espacial.

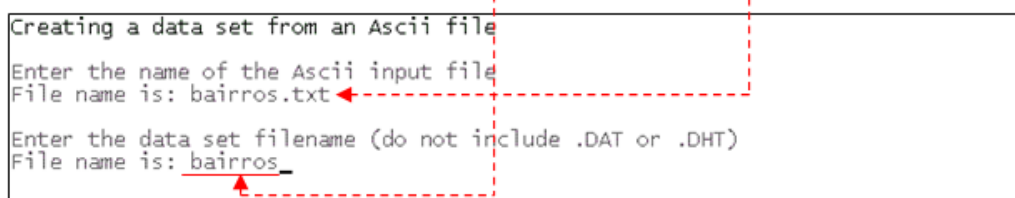
5.9.5 Configurar o SpaceStat

Pressione as teclas Ctrl O (Control mais a tecla O) e configure exatamente conforme a tela abaixo:



5.9.6 - Tecle Esc para retornar ao menu principal.

5.9.7 - Tecle: D 1 1 para visualizar a tela abaixo; defina o arquivo Ascii, no caso bairros.txt; para finalizar tecle bairros e ENTER.



A execução acima converte a tabela de atributos (bairros.txt) em dois novos arquivos: bairros.dat e bairros.dht. São formatos internos e necessários para o SpaceStat trabalhar. Quando a operação acima é executada com êxito, surgirá o "Report" abaixo sobre a tabela de atributos.

```
The GAUSS dataset bairros has been created
It contains 96 observations on 44 variables
The variables contained in the data set are:
  GEOID DENSDEMO IEX_EXSO IEX_EQID IEX_DSHU IEX_QV IEX_AUT POPULAC
  DEMVAG DEFWAG POP70 TOTCH CHMULH S_REND REN1_5 REN1_5_3
  REN3_5 REN5_10 REN10_20 REND20 REN20 TOTDOM DMPRE SEMSAU
  SEMESG ANALFA INST15 PERIDOSO PERSREND PERREN20 PERSESG PERDPRE
  PERINS15 Z WZ IMORAN MEDIAMOV BOXMAP LISAMAP MORANMP
  G_PERIDO Z_PERIDO P_PERIDO S_PERIDO
```

Para retornar ao Menu do SpaceStat tecle ENTER.

5.9.8 - Calculando a estatística G_i . Tecla: E 5 3 2 ENTER para visualizar a tela abaixo; digite bairros e ENTER.

```
INTERACTIVE PROBLEM FILE CREATION
Problem no. 1
Data set for problem
Enter the data set filename (do not include .DAT or .DHT),
or press Return for directory listing: bairros
```

Defina o arquivo .gal (Matriz de Proximidade), no caso bairros. Para finalizar tecla ENTER.

```
Spatial weights for problem
Enter the filenames for the spatial weights from the following list
or press Return to stop

.FMT files in default directory:

.GAL files in default directory:
  BAIRROS  CÓPIAD~1

.GWT files in default directory:

Weights file: bairros
Weights file:
```

Seguindo defina a variável, no caso peridoso. Para finalizar tecla ENTER.

```
Variables for problem
Choose the variable(s) from the following list
(each variable name should correspond exactly to one in the list):
GEOID  DENSDEMO  IEX_EXSO  IEX_EQID  IEX_DSHU  IEX_QV  IEX_AUT
POPULAC  DEMVAG  DEFVAG  POP70  TOTCH  CHMULH  S_REND
REN1_5  REN1_5_3  REN3_5  REN5_10  REN10_20  REN20  REN20
TOTDOM  DOPRE  SEMSAU  SEMESG  ANALFA  INST15  PERIDOSO
PERSREND  PERREN20  PERSESG  PERDPRE  PERINS15  Z  WZ
IMORAN  MEDIAMOV  BOXMAP  LISAMAP  MORANMP  G_PERIDO  Z_PERIDO
P_PERIDO  S_PERIDO

Enter the variable name, or press Return to stop
Variable name: peridoso
Variable name: -
```

```
G-I STATISTICS FOR SPATIAL ASSOCIATION
Weights matrix BAIRROS is not row standardized
```

Tecla ENTER para apresentar o relatório final, conforme ilustrado abaixo.

G-I STATISTICS FOR SPATIAL ASSOCIATION									
DATA SET: BAIRROS				WEIGHTS: BAIRROS		VARIABLE: PERIDOSO			
LARGEST (MOST POSITIVE Z) VALUES				SMALLEST (MOST NEGATIVE Z) VALUES					
GEOID	G-I	Z	PROB	GEOID	G-I	Z	PROB		
40	0.127017	4.2911	0.0000	69	0.0226153	-2.7554	0.0059		
38	0.122179	4.0163	0.0001	86	0.0410726	-2.5582	0.0105		
96	0.104908	3.8224	0.0001	64	0.0194348	-2.4630	0.0138		
42	0.148607	3.8099	0.0001	84	0.028214	-2.3784	0.0174		
37	0.0986282	3.4158	0.0006	4	0.0208781	-2.3545	0.0185		
36	0.0983584	3.3550	0.0008	62	0.0286379	-2.3395	0.0193		
75	0.094934	3.0881	0.0020	90	0.021771	-2.2701	0.0232		
95	0.0945506	3.0640	0.0022	66	0.02277	-2.2092	0.0272		
58	0.107834	3.0225	0.0025	9	0.0487229	-2.1137	0.0345		
59	0.105327	2.8452	0.0044	12	0.0320244	-2.1119	0.0347		

Lado Esquerdo

Lado Direito

A tela de relatório anterior exibe do lado esquerdo apenas os 10 “Z valores” mais positivos, com seus respectivos valores de GEOID, G_i e níveis de significância. Do lado direito os 10 “Z valores” mais negativos, também com seus respectivos valores de GEOID, G_i e níveis de significância.

O “Z valor” é definido como:

$$Z_i = (G_i - E[G_i]) / \sigma(G_i) \quad i=1, \dots, n \quad (n = \text{número de polígonos}).$$

Um “Z valor” positivo e significativo para a estatística G_i indica **Agrupamento Espacial de alto valores**, por outro lado “Z valor” negativo e significativo indica **Agrupamento Espacial de baixo valores**.

Por exemplo, do lado esquerdo, os 3 maiores “Z valores” são 4.2911 para vizinhança 40 (GEOID), 4.0163 para vizinhança 38(GEOID) e 3.8224 para vizinhança 96(GEOID). Todos 3 são altamente significantes, indicando um forte agrupamento de alto valores de Peridoso (a variável em questão) concentrado nestas localizações.

>>> Fechar o SpaceStat.

Além da tela de relatório, o SpaceStat gera também um arquivo ASCII no diretório de trabalho conforme definido no item 5.9.4 (no caso c:\temp). O nome deste arquivo está relacionado com o tipo de operação executada. Como executamos a estatística G_i , então o nome do arquivo a ser gerado segue a regra: G_i <name>.txt, onde <name> refere-se ao nome da Matriz de Proximidade (no caso, bairros). Outro detalhe é que o nome do arquivo é limitado a 8 caracteres (limitações do DOS). Então, neste caso, o SpaceStat gera o arquivo: G_i _bairr.txt.

O arquivo G_i _bairr.txt contém o valor da variável indicadora (GEOID), e para cada variável, seu valor original (PERIDOSO), o valor da estatística G_i ou G_i^* (G_PERIDO), os “Z valores”, (Z_PERIDO) suas probabilidades (P_PERIDO) e um indicador de significância (S_PERIDO) tomando os seguintes valores: 3 se $p < 0.001$; 2 se $p < 0.01$; 1 se $p < 0.05$ e zero não significativa.

A figura abaixo ilustra um trecho deste arquivo:

```

GEOID,PERIDOSO,G_PERIDO,Z_PERIDO,P_PERIDO,S_PERIDO
1,3.472380,0.084460,2.327314,0.019949,1.000000
2,1.432020,0.015384,-0.652102,0.514335,0.000000
3,1.136790,0.058080,-0.345621,0.729628,0.000000
4,1.129620,0.020878,-2.354543,0.018546,1.000000
5,2.319760,0.006036,-1.716693,0.086035,0.000000
6,1.437250,0.019877,-1.827772,0.067584,0.000000
7,0.909890,0.010879,-1.963878,0.049544,1.000000
8,1.203120,0.010933,-1.953519,0.050758,0.000000
9,1.247280,0.048723,-2.113688,0.034542,1.000000
10,2.943160,0.033466,-1.404123,0.160282,0.000000
11,1.506970,0.037520,-0.376863,0.706275,0.000000
12,1.540690,0.032024,-2.111863,0.034698,1.000000
13,3.181000,0.033842,-0.672431,0.501309,0.000000
14,0.847980,0.015162,-1.558535,0.119107,0.000000
15,2.793350,0.039811,-0.939888,0.347275,0.000000
16,1.318590,0.017773,-1.305051,0.191875,0.000000
17,5.327590,0.106184,2.028852,0.042473,1.000000
18,5.478060,0.059930,-0.858891,0.390401,0.000000
19,2.665430,0.063894,0.826162,0.408712,0.000000
20,4.253900,0.077233,1.796051,0.072486,0.000000
21,2.733850,0.047575,0.445944,0.655638,0.000000
22,5.948060,0.072872,1.480720,0.138681,0.000000
23,3.040510,0.051868,-0.055895,0.955425,0.000000
24,7.699200,0.106515,1.325506,0.185004,0.000000
25,2.258740,0.049033,0.566348,0.571157,0.000000
26,2.111080,0.024289,-0.684985,0.493353,0.000000
27,4.272740,0.067289,0.276860,0.781888,0.000000
28,4.980060,0.056435,1.163197,0.244749,0.000000
29,4.704970,0.100076,2.473779,0.013369,1.000000
30,3.678100,0.046583,1.400061,0.161493,0.000000
31,4.480930,0.030770,-1.616201,0.106051,0.000000
32,2.520390,0.047334,0.426808,0.669519,0.000000
33,1.961080,0.008888,-1.393762,0.163389,0.000000
34,2.641410,0.049770,-0.209950,0.833707,0.000000
35,1.550170,0.027600,-1.191694,0.233381,0.000000
36,6.621310,0.098358,3.355024,0.000794,3.000000
37,8.224960,0.098628,3.415817,0.000636,3.000000
38,8.036810,0.122179,4.016339,0.000059,3.000000
39,7.232590,0.096411,2.248081,0.024571,1.000000
40,6.108050,0.127017,4.291116,0.000018,3.000000
41,5.467900,0.071733,2.406698,0.016097,1.000000
42,7.096720,0.148607,3.809860,0.000139,3.000000
43,5.183660,0.091556,2.842356,0.004478,2.000000
44,4.086270,0.077982,0.993708,0.320365,0.000000
45,3.165160,0.047984,0.478447,0.632332,0.000000
46,6.653310,0.085612,2.420226,0.015511,1.000000
47,1.901440,0.044805,-0.576773,0.564093,0.000000
  
```

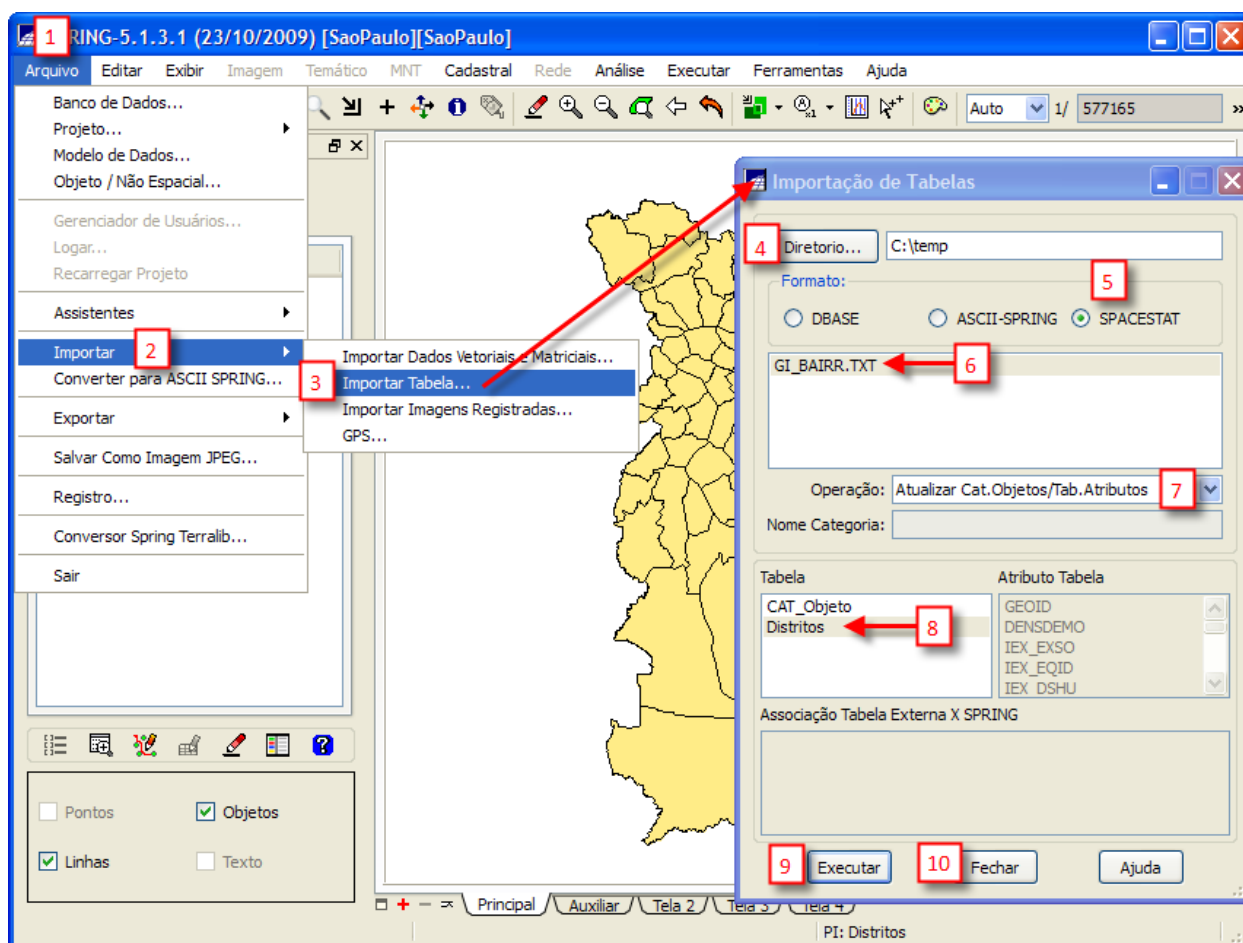
5.9.9 - O próximo passo é importar para o SPRING o Arquivo Gi_bairr.txt. Este arquivo nada mais é do que uma tabela de atributos.

Antes de executar a importação:

a) Se o SPRING estiver no modo consulta, feche a consulta. Isto é necessário porque vamos adicionar novas informações na tabela já existente no banco de dados São Paulo.

b) Sempre execute esta ação, *MENU: Executar -> Limpar -> Tudo*. Isto garante que não haverá nenhum objeto selecionado e a atualização ocorrerá sem problemas.

c) Agora realize a importação da tabela Gi_bairr.txt conforme ilustra a figura abaixo.



O resultado desta operação pode ser visto na tabela de atributos, conforme ilustra a figura seguinte.

Z	WZ	IMORAN	MEDIAMOV	BOXMAP	LISAMAP	MORANMP	G_PERID	Z_PERID	P_PERID	S_PERID
2.111...	1.840...	0.942815	1.878920	1.000...	1.000...	1.000000	0.097601	2.310409	0.020866	1.000000
1.267...	1.988...	0.611504	1.885128	1.000...	1.000...	1.000000	0.100076	2.473779	0.013369	1.000000
-2.002...	-1.716...	0.833900	-1.811685	2.000...	0.000...	0.000000	0.010477	-1.216578	0.223765	0.000000
-1.013...	-0.773...	0.190256	-0.813586	2.000...	0.000...	0.000000	0.040660	-0.879311	0.379233	0.000000
-1.484...	-0.554...	0.199702	-0.709512	2.000...	0.000...	0.000000	0.043940	-0.640300	0.521977	0.000000
3.216...	2.098...	1.637718	2.285008	1.000...	1.000...	1.000000	0.085612	2.420226	0.015511	1.000000
2.346...	3.364...	1.915499	3.195050	1.000...	3.000...	1.000000	0.104908	3.822386	0.000132	3.000000
3.722...	1.947...	1.758585	2.169131	1.000...	2.000...	1.000000	0.116759	2.709992	0.006728	2.000000
2.197...	1.661...	0.885434	1.737593	1.000...	1.000...	1.000000	0.094313	2.090324	0.036589	1.000000
0.928...	2.745...	0.618231	2.442220	1.000...	2.000...	1.000000	0.094934	3.088092	0.002014	2.000000
-1.887...	-1.171...	0.536202	-1.314316	2.000...	0.000...	0.000000	0.027600	-1.191694	0.233381	0.000000
0.835...	0.215...	0.043677	0.304036	1.000...	0.000...	0.000000	0.067289	0.276860	0.781888	0.000000
-1.535...	-0.497...	0.185329	-0.670435	2.000...	0.000...	0.000000	0.044805	-0.576773	0.564093	0.000000
3.487...	2.233...	1.890221	2.412875	1.000...	2.000...	1.000000	0.105327	2.845154	0.004439	2.000000

Observe os quatro últimos atributos em destaque na tabela acima; são oriundos do arquivo Gi_bairr.txt. Relembrando:

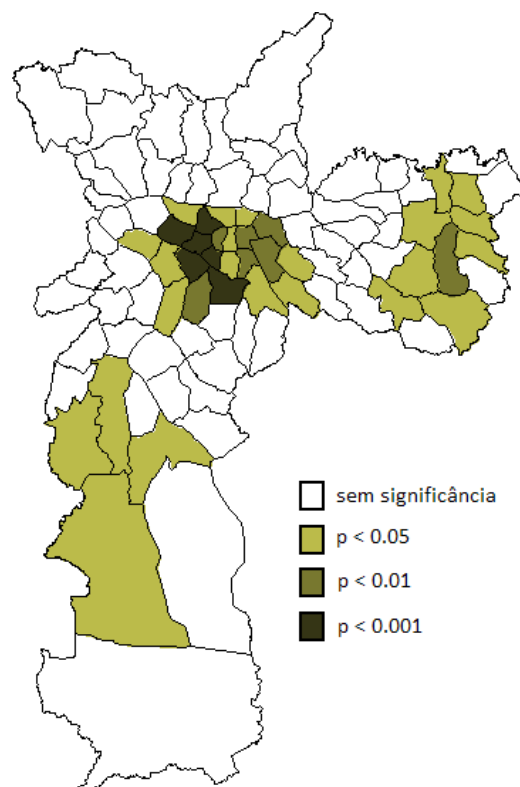
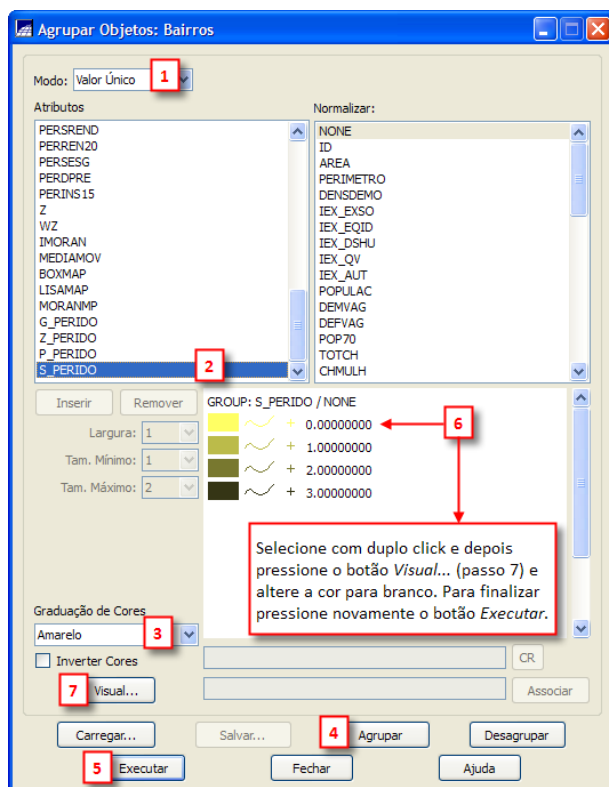
G_PERIDO: refere-se à estatística G_i ;

Z_PERIDO: é o “Z valor” igual a $(G_i - E[G_i]) / \sigma G_i$

P_PERIDO: a probabilidade da estatística G_i

S_PERIDO: é um indicador de significância que leva os seguintes valores: **3** se $p < 0.001$; **2** se $p < 0.01$; **1** se $p < 0.05$ e **0** sem significância.

Um resultado interessante de ser apresentado e analisado refere-se ao mapa de significância do indicador de associação espacial G_i . Isto é realizado conforme ilustra a figura seguinte.



O mesmo procedimento pode ser realizado para a estatística G_i^* . Inicia-se no item 5.9, com a seguinte restrição: no item 5.9.8 TECLE: E 5 4 2 ENTER.

BIBLIOGRAFIAS DE APOIO

- ANSELIN, L. (1995) "Local indicators of spatial association - LISA". **Geographical Analysis**, 27 : 91-115.
- ANSELIN, L. (1996) "The Moran scatterplot as ESDA tool to assess local instability in spatial association". In: Fisher, M.; Scholten, H. J.; Unwin, D. *Spatial Analytical Perspectives on GIS*. London: Taylor & Francis, p. 111-126.
- ANSELIN, L. (1998) "Exploratory Spatial Data Analysis in a Geocomputational Environment". In: LONGLEY, P. A.; BROOKS, S. M.; MCDONNELL, R.; MACMILLIAN, B. *Geocomputation: a primer*. Chichester, John Wiley, p. 77-94.
- ANSELIN, L. & BAO, S. (1997) *Exploratory Spatial Data Analysis Linking SpaceStat and ArcView*. In: Fischer, M. M. & Getis, A. *Recent developments in spatial analysis*. New York, Springer, p 35-59. BAILEY, T.C., GATRELL, A.C., *Interactive spatial data analysis*. Essex, Longman Scientific & Technical, 1995.
- BESAG, C.; NEWELL, J. (1991). The detection of clusters in rare diseases. **Journal of the Royal Statistical Society**, A, 154. p. 143-155.
- CARVALHO, M. S. (1997). **Aplicação de Métodos de Análise Espacial na Caracterização de Áreas de Risco à Saúde**. Tese de Doutorado, Rio de Janeiro: Instituto Alberto Luiz Coimbra de Pós-graduação e Pesquisa de Engenharia, Universidade Federal do Rio de Janeiro.
- CARVALHO, M.S. ; CRUZ, O.G (1998). "Mortalidade por causas externas: análise exploratória espacial, Região Sudeste/Brasil". **Anais do XIº Encontro Nacional de Estudos Populacionais, Caxambu**, 19 a 23 de Outubro de 1998.
- CRESSIE, N. (1991). **Statistics for Spatial Data**. Chichester, John Wiley.
- FISCHER, M.; GETIS, A. (1996). **Recent developments in spatial analysis**. New York, Springer.
- FISCHER, M.; SCHOLTEN, H.; UNWIN, D. (1996) *Spatial Analytical Perspectives on GIS*. London: Taylor & Francis.
- GETIS, A.; ORD, J.K. (1992). The analysis of spatial association by use of distance statistics. **Geographical Analysis**, 24: -206, 1992.
- HERBERT, D. T. (1980). The British experience. In Georges-Abeyie, D. E. e Harries, K. D. (Eds.), **Crime: a Spatial Perspective**, Columbia University Press.
- HAINING, R.P. (1990). **Spatial data analysis in the social and environmental sciences**. Cambridge University Press, Cambridge.
- LONGLEY, P. A.; BROOKS, S. M.; MCDONNELL, R.; MACMILLIAN, B. (1998). **Geocomputation: a primer**. Chichester, John Wiley.
- ORD, J. K.; GETIS, A. (1995). Local spatial autocorrelation statistics: distributional issues and an application. **Geographical Analysis**, 27 : 286-306.
- SPOSATI, A. et al. **Mapa da exclusão/inclusão social da cidade de São Paulo**. São Paulo: EDUC, 1996.