

# Métodos de análise espacial aplicados a estudos de população

Diego Rodrigue Macedo



Professor Adjunto  
Departamento de Geografia



INSTITUTO DE GEOCIÊNCIAS  
**PROGRAMA de  
PÓS-GRADUAÇÃO  
em GEOGRAFIA**  
UNIVERSIDADE FEDERAL DE MINAS GERAIS



Programa de Pós-Graduação em  
*Análise e Modelagem de Sistemas Ambientais*  
**UNIVERSIDADE FEDERAL  
DE MINAS GERAIS**



# Análise de dados de área

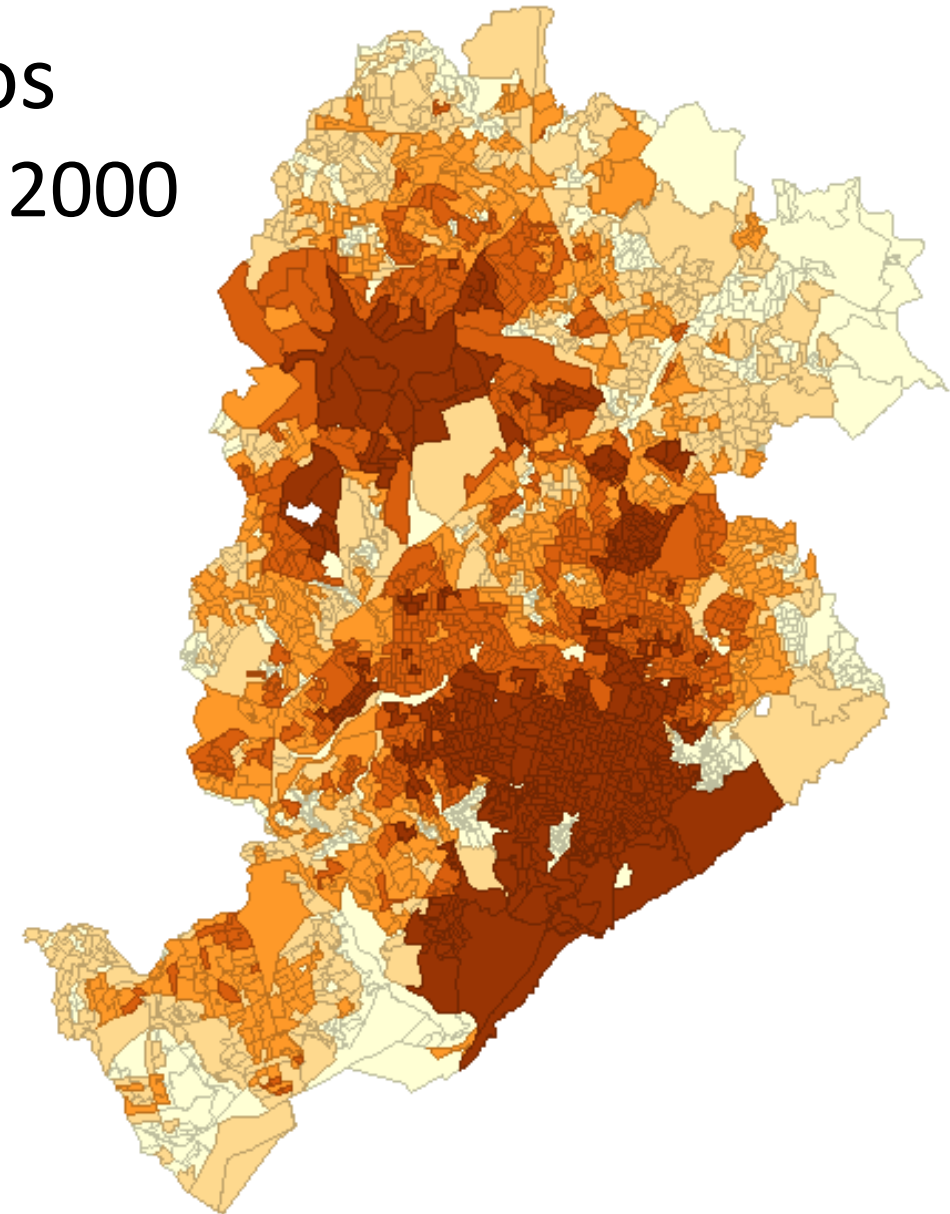
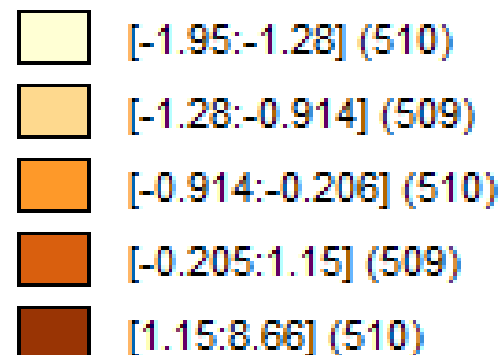
# Análise de dados de área

- Áreas delimitadas por polígonos
- Eventos agregados por alguma área espacial (p.ex. setor censitário, município, mesorregião, etc)
- Não há a localização exata do evento
- Geralmente representa taxas ou eventos binários

# Análise de dados de área

- Representação em mapas temáticos
  - Exemplo: Renda em Belo Horizonte, 2000

Quantile: RENDA

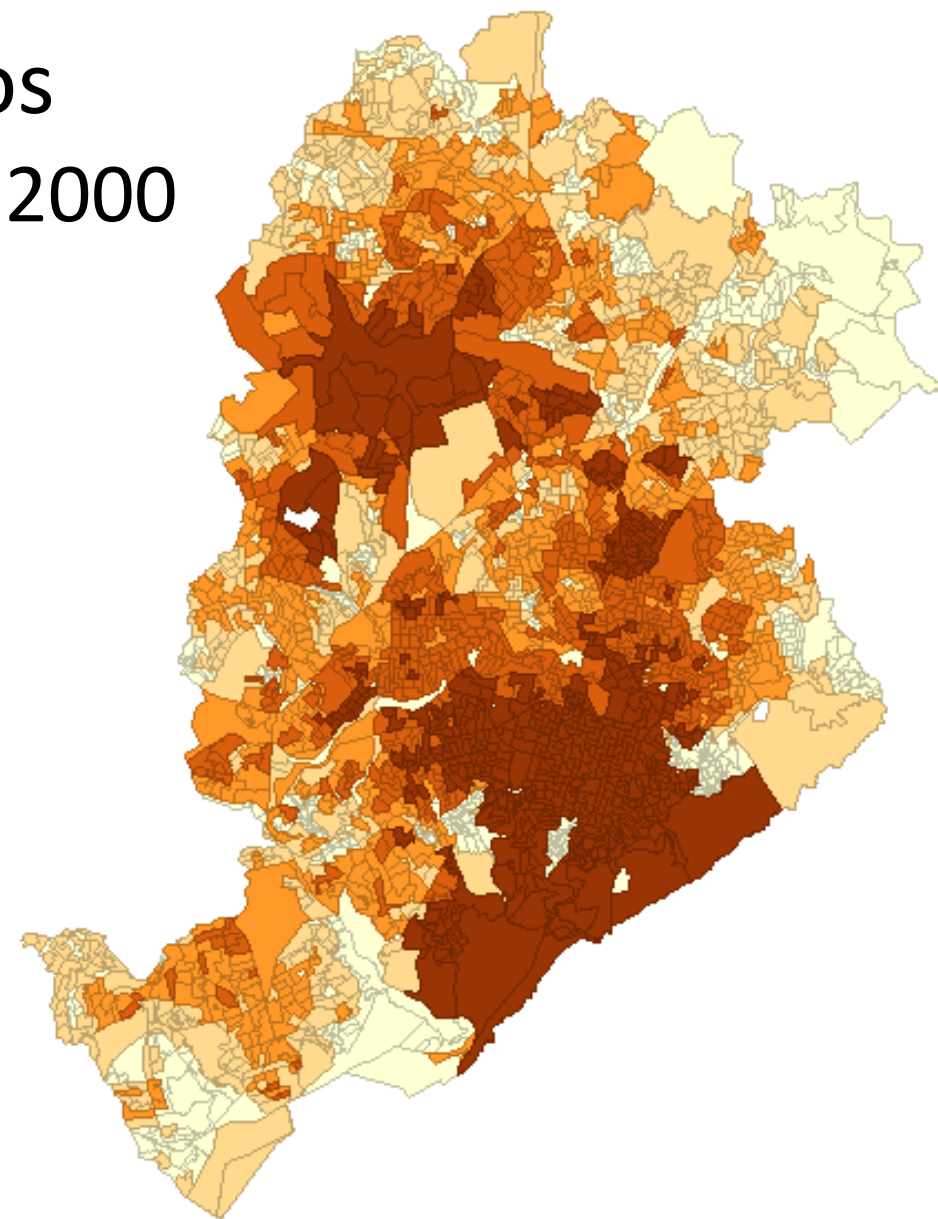
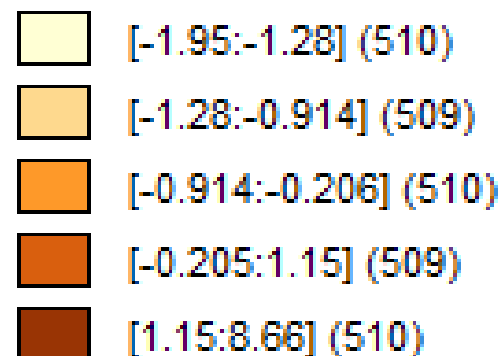


# Análise de dados de área

- Representação em mapas temáticos
  - Exemplo: Renda em Belo Horizonte, 2000

- Abordagem simplista
- Conclusões intuitivas

Quantile: RENDA

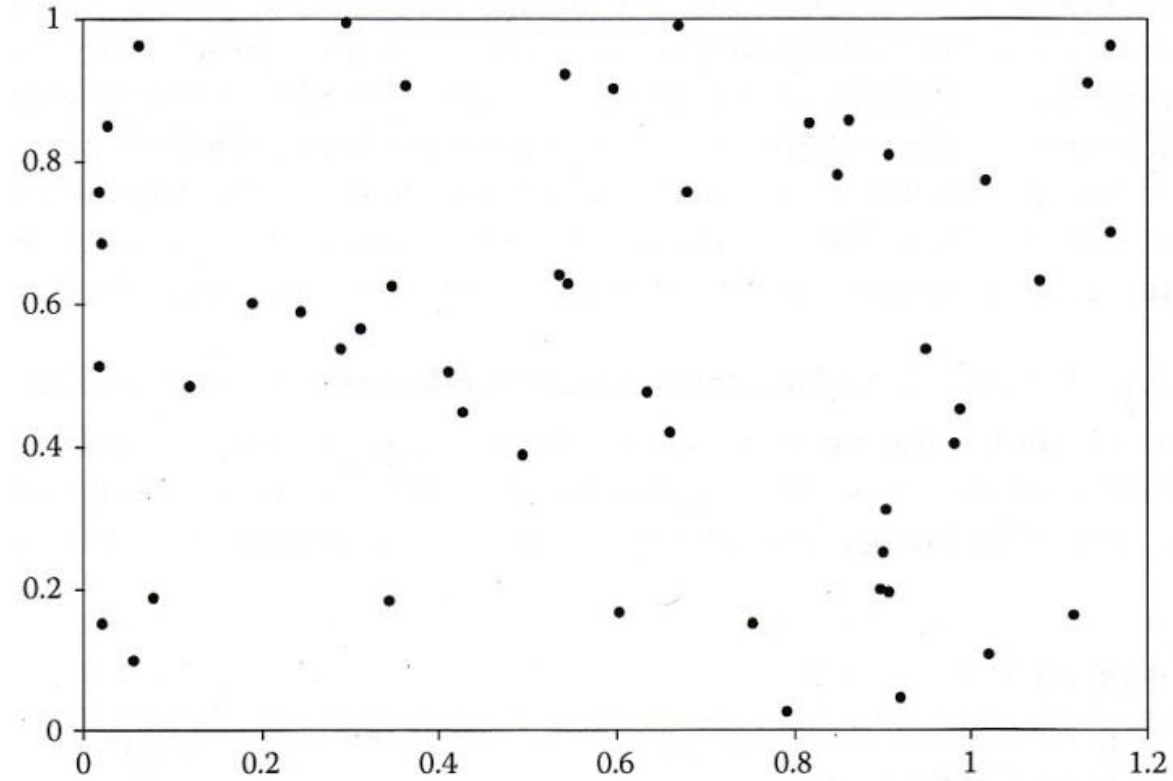
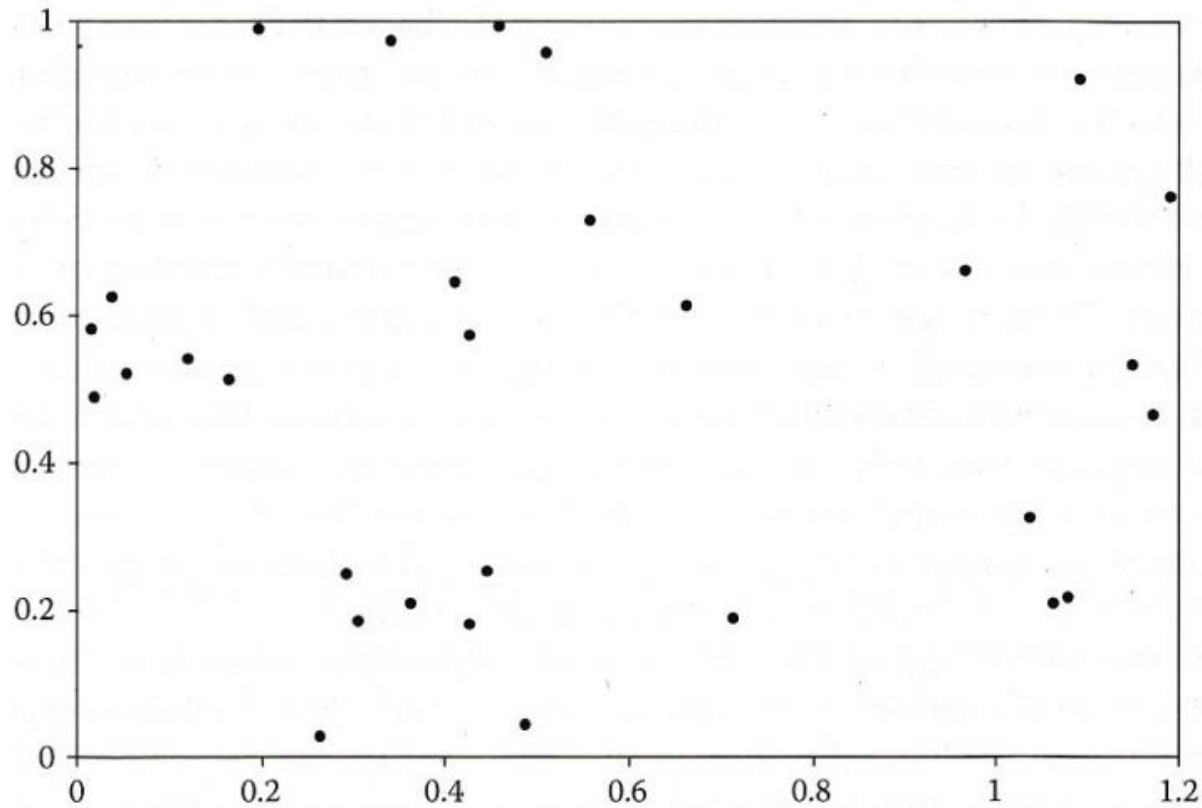


# Análise exploratória de dados espaciais

# Padrão espacial

- Padrão aleatório?
- Padrão agregado?
- Causas mensuráveis?
- Clusters (regiões)?

# Padrão espacial



Padrão aleatório? Ou agregado?

# Autocorrelação espacial

- Primeira lei da Geografia: “Todas as coisas se relacionam, porém, coisas mais próximas são mais relacionadas que as distantes” (Tobler, 1979).
- As medidas de autocorrelação espacial procuram lidar simultaneamente com similaridades na localização dos objetos espaciais e de seus atributos

# Autocorrelação espacial

- Três possibilidades:
- Autocorrelação Positiva (Lei de Tobler): Feições similares em localização também são similares em atributos
- Autocorrelação Negativa (oposição à Lei de Tobler): Feições similares em localização tendem a ter atributos menos similares do que feições mais distantes
- Ausência de Autocorrelação: Quando atributos são independentes da localização

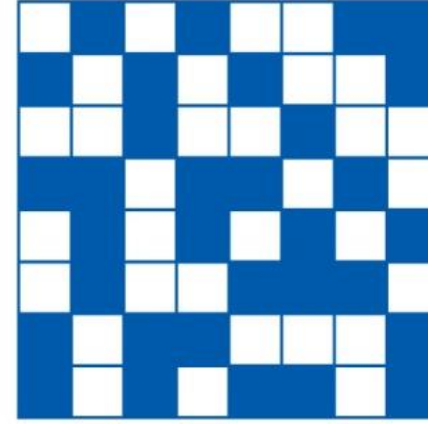
# Field arrangements of blue and white cells exhibiting

(A)



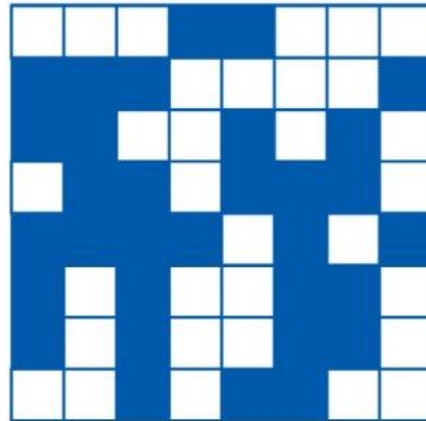
$I = -1.000$   
 $n_{BW} = 112$   
 $n_{BB} = 0$   
 $n_{WW} = 0$

(B)



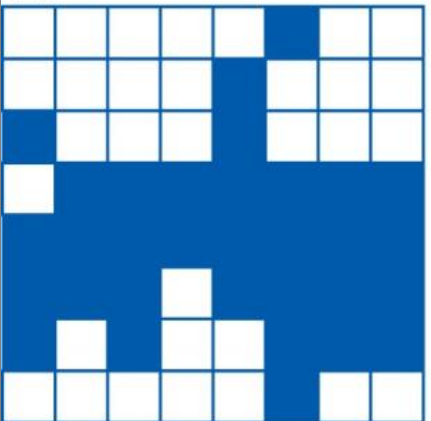
$I = -0.393$   
 $n_{BW} = 78$   
 $n_{BB} = 16$   
 $n_{WW} = 18$

(C)



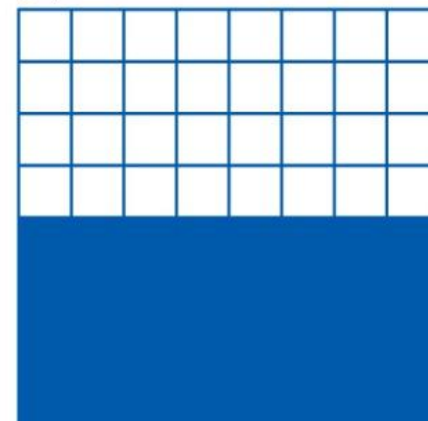
$I = 0.000$   
 $n_{BW} = 56$   
 $n_{BB} = 30$   
 $n_{WW} = 26$

(D)



$I = +0.393$   
 $n_{BW} = 34$   
 $n_{BB} = 42$   
 $n_{WW} = 36$

(E)



$I = +0.857$   
 $n_{BW} = 8$   
 $n_{BB} = 52$   
 $n_{WW} = 52$

(A) extreme negative spatial autocorrelation

(B) a dispersed arrangement

(C) spatial independence

(D) spatial clustering

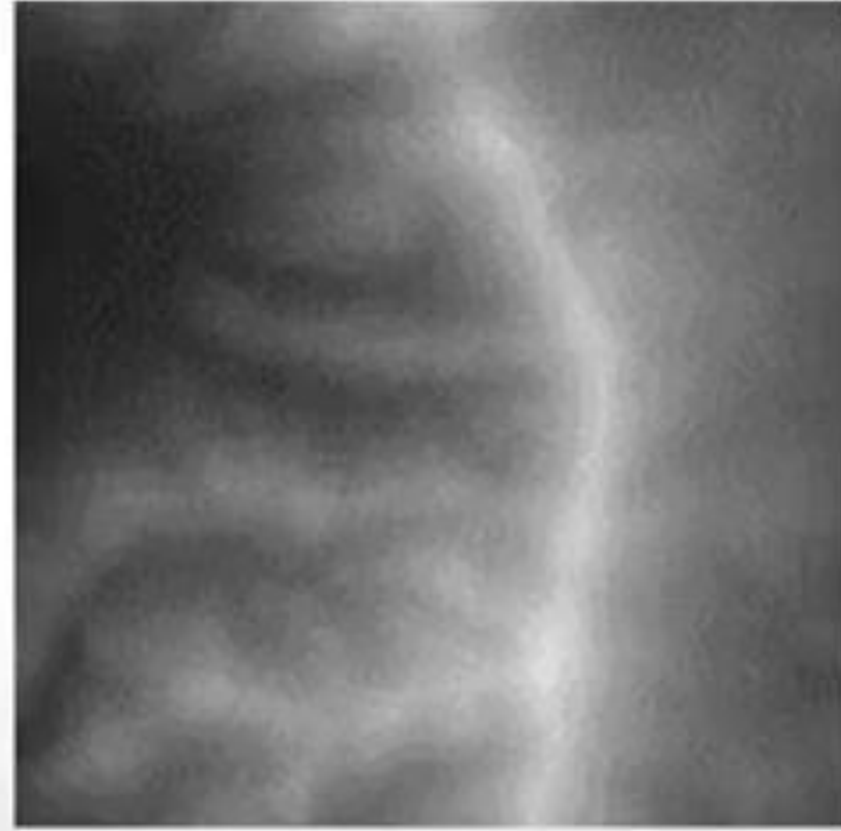
(E) extreme positive spatial autocorrelation

(Goodchild 1986 apud Longley et al, 2013)

# Autocorrelação espacial



$\rho = 0.0$

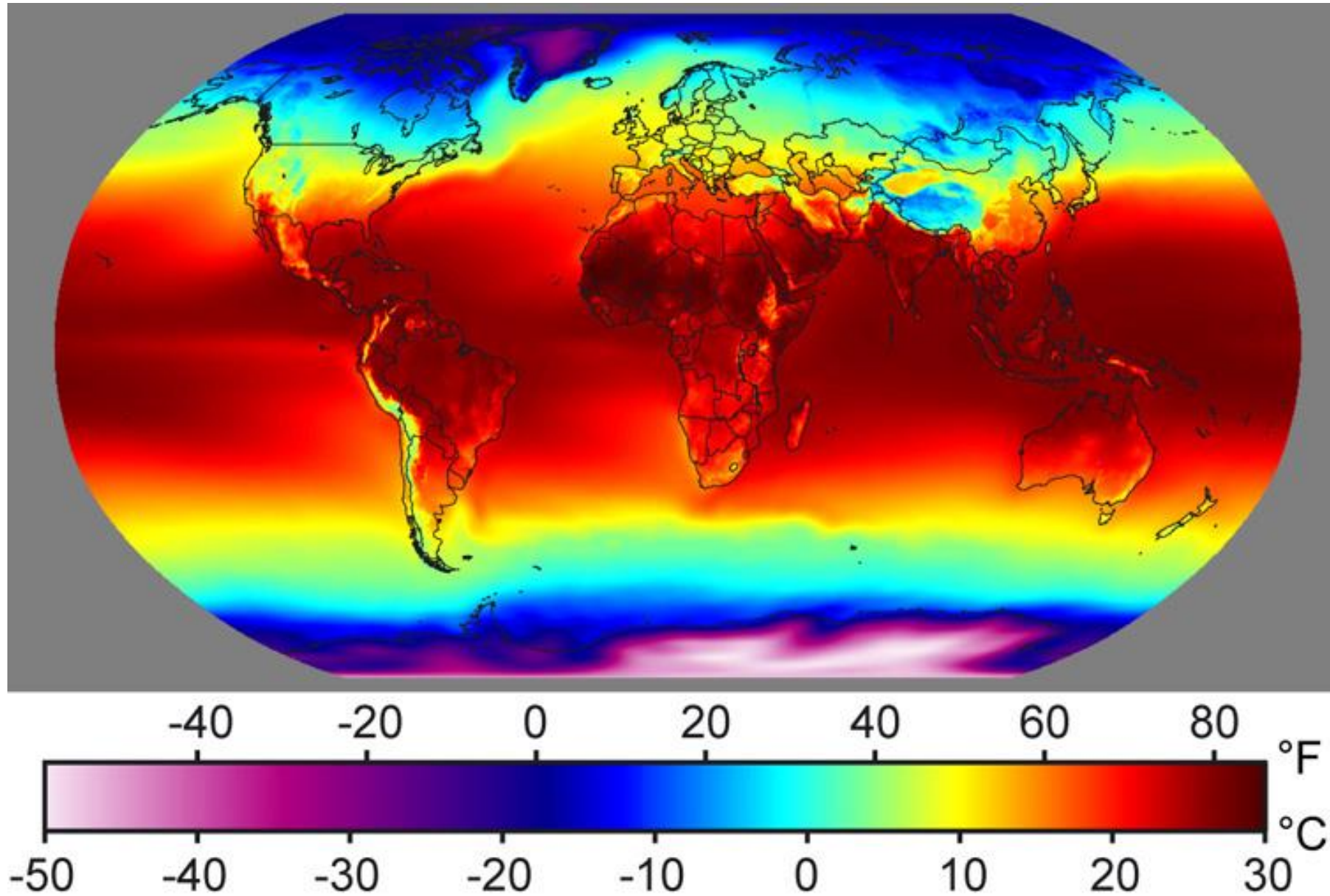


$\rho = 0.8$

# Autocorrelação espacial

- Heterogeneidade espacial: Tendência de lugares de serem diferentes uns dos outros.
- Heterogeneidade maior com o aumento da distância
- Alguns fenômenos geográficos variam de maneira gradual através do espaço (estacionariedade), enquanto outros podem apresentar extrema irregularidade, violando a Lei de Tobler

# Autocorrelação espacial



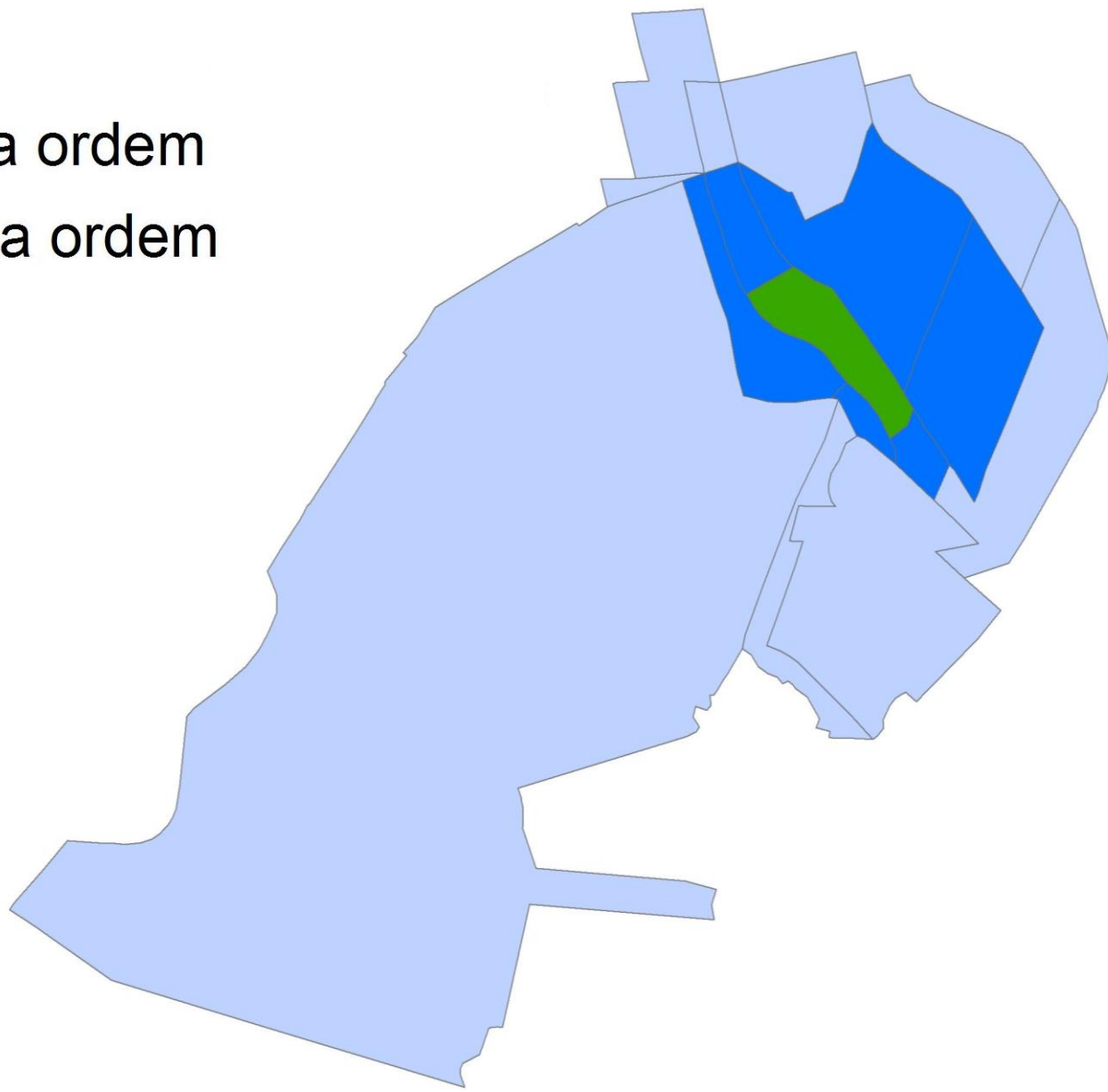
# Autocorrelação espacial



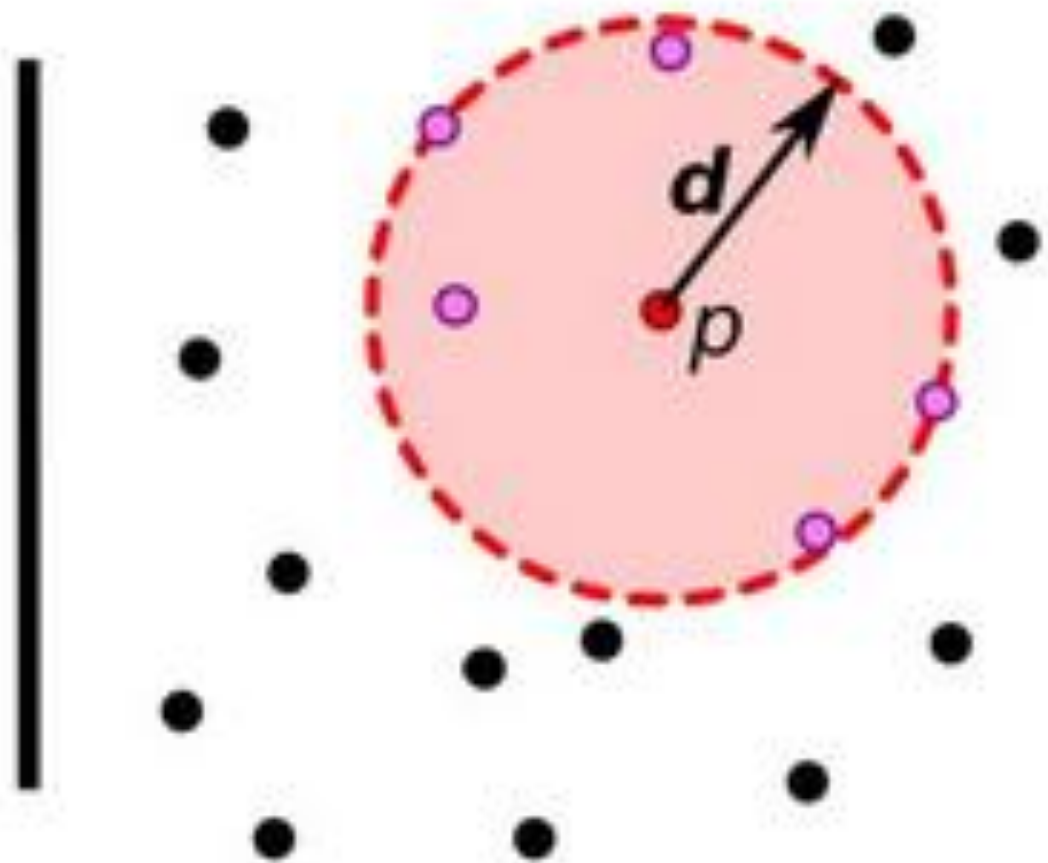
Foto: Tuca Vieira

# Vizinhança

-  Setor
-  Vizinhança de primeira ordem
-  Vizinhança de segunda ordem



# Vizinhança



K-vizinhos

K-vizinhos até um raio “d”

# Indice de Moran

# Índice de Moran

- Dependência espacial entre as unidades de análise em relação à variável de estudo
- A principal medida de correlação espacial é o Índice de Moran ( $I$ )

# Dependência espacial

- Aleatoriedade dos dados espaciais:
  - Se **SÃO** aleatórios, o valor de uma variável não depende do valor desta variável nos locais vizinhos
- Autocorrelação espacial global
- Autocorrelação espacial local

# Autocorrelação espacial global

- Índice de Moran (I) Global

$$I = \left( \frac{N}{\sum_{ij} w_{ij}} \right) \frac{\sum_{ij} w_{ij} (x_i - \mu)(x_j - \mu)}{\sum_i (x_i - \mu)^2} \quad (1)$$

Onde:

$I$  = Moran Global,

$N$  = número de observações,

$x$  = variável analisada,

$\mu$  = média de  $x$ ,

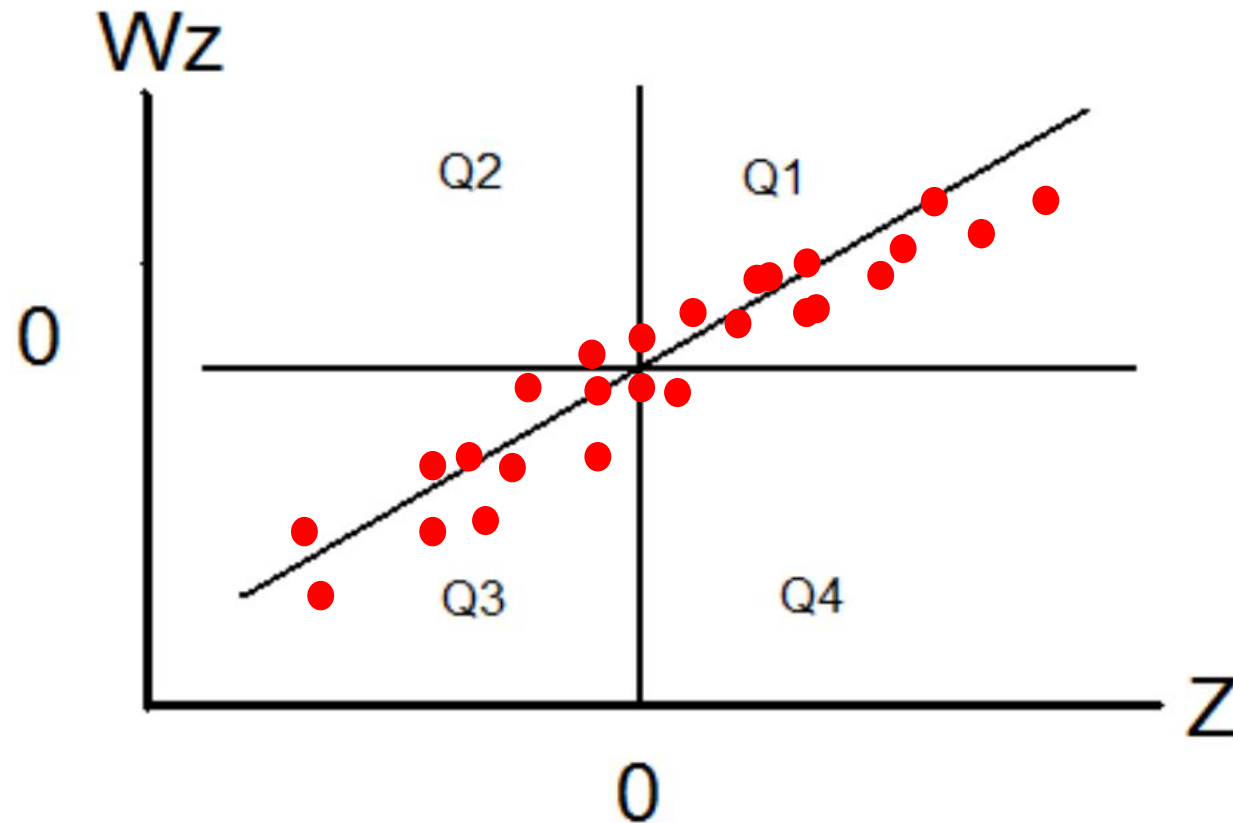
$w_{ij}$  = elementos da matriz de pesos espaciais, nos quais  $i$  e  $j$  referem-se a localização geográfica.

# Autocorrelação espacial global

- Índice de Moran (I) Global
- Medida de associação linear entre os valores de uma determinada área ( $Z$ ) e a média ponderada dos valores da vizinhança ( $Wz$ )
- O índice de Moran varia entre -1 e 1
- Interpretação similar à análise de correlação linear

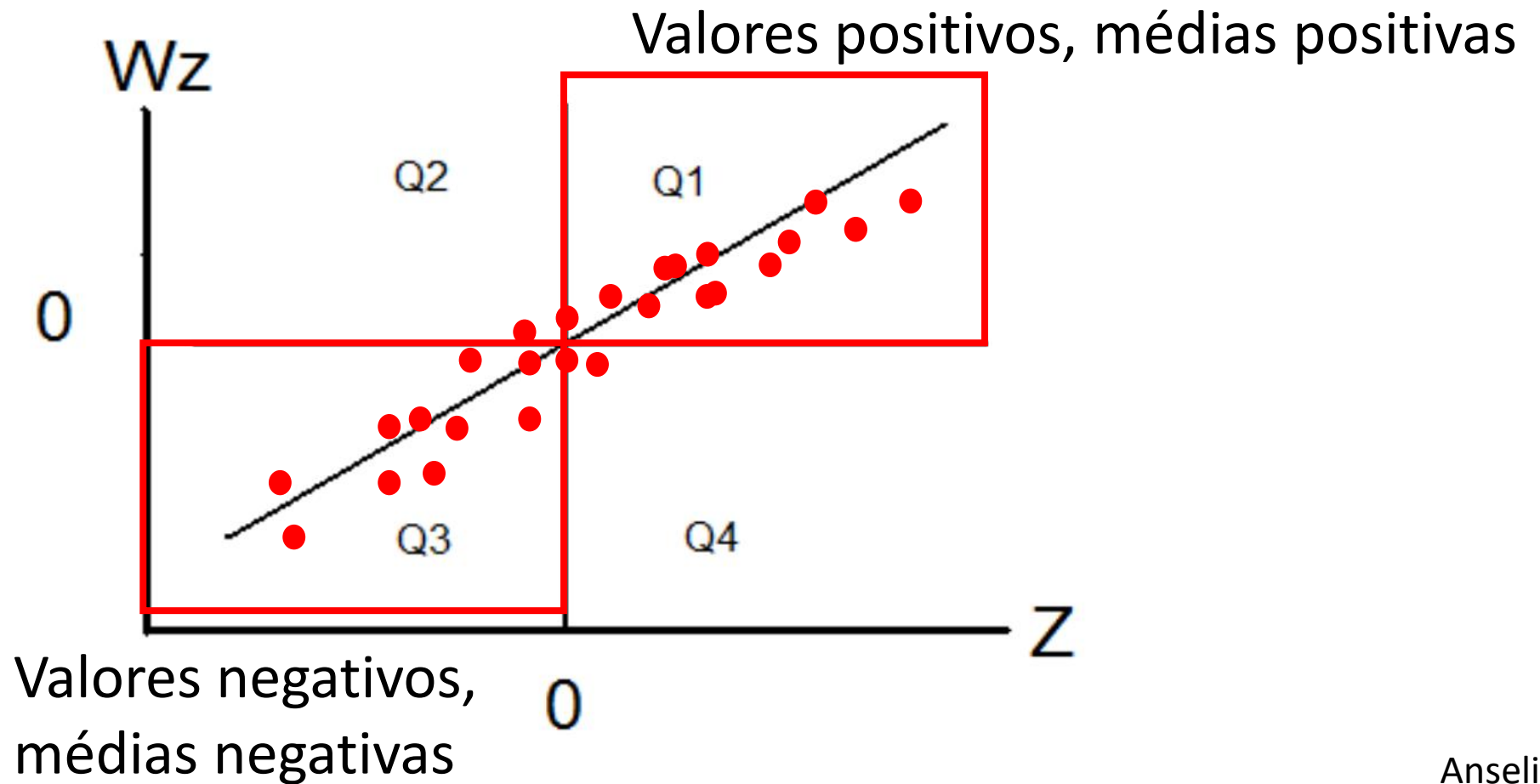
# Diagrama de dispersão

- Diagrama de dispersão



# Diagrama de dispersão

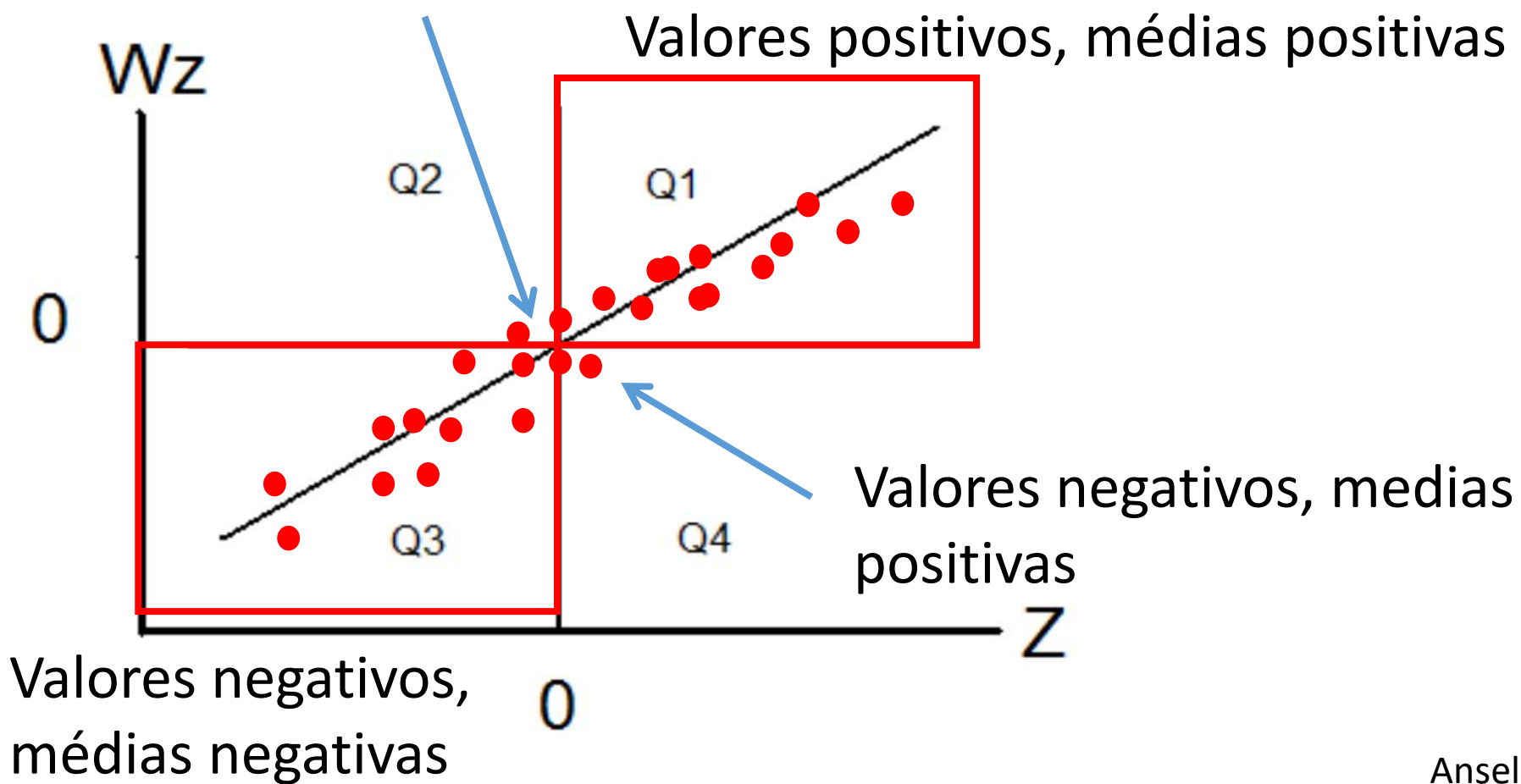
- Diagrama de dispersão



# Diagrama de dispersão

- Diagrama de dispersão

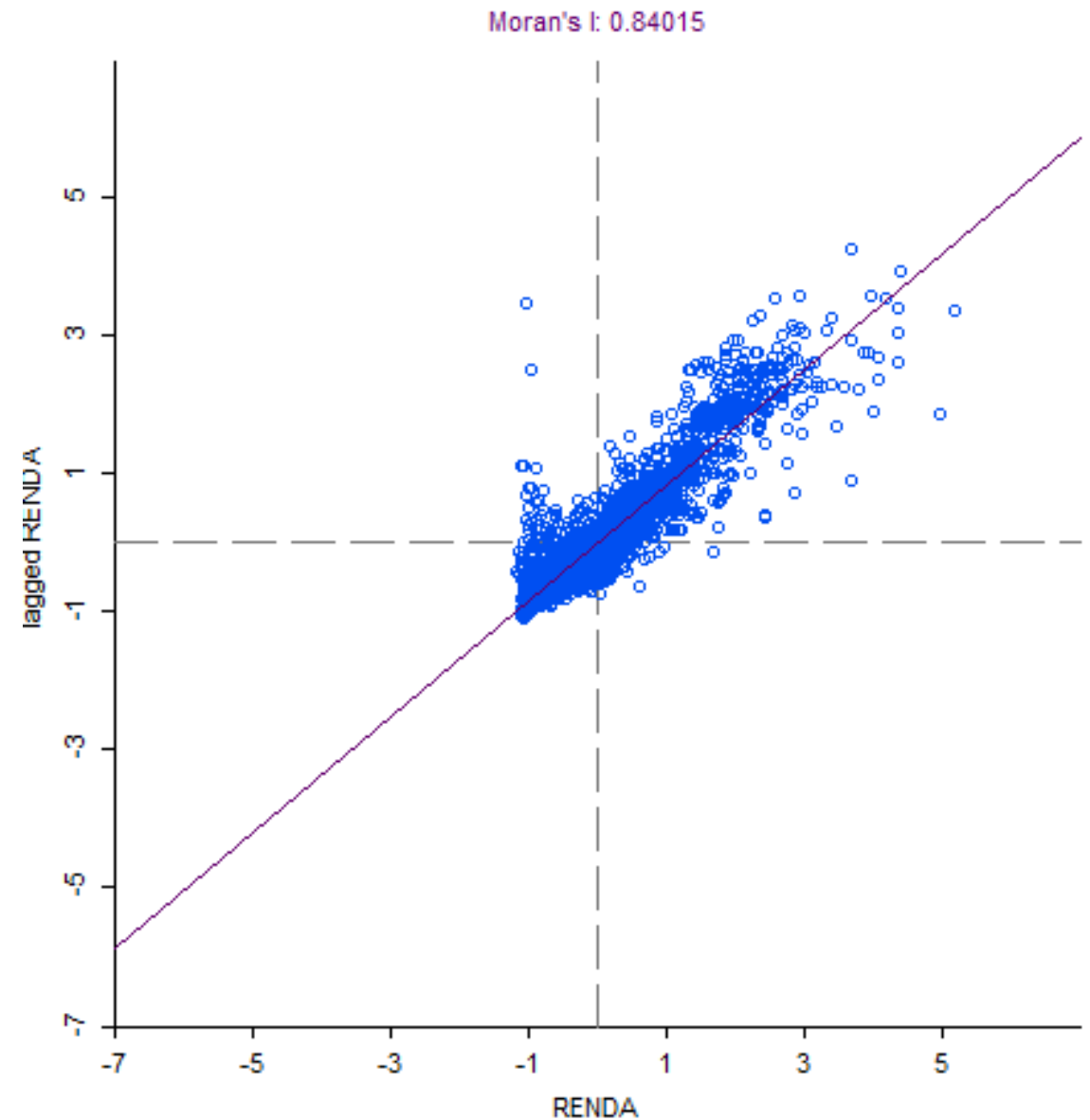
Valores positivos, médias negativas



# Autocorrelação espacial global

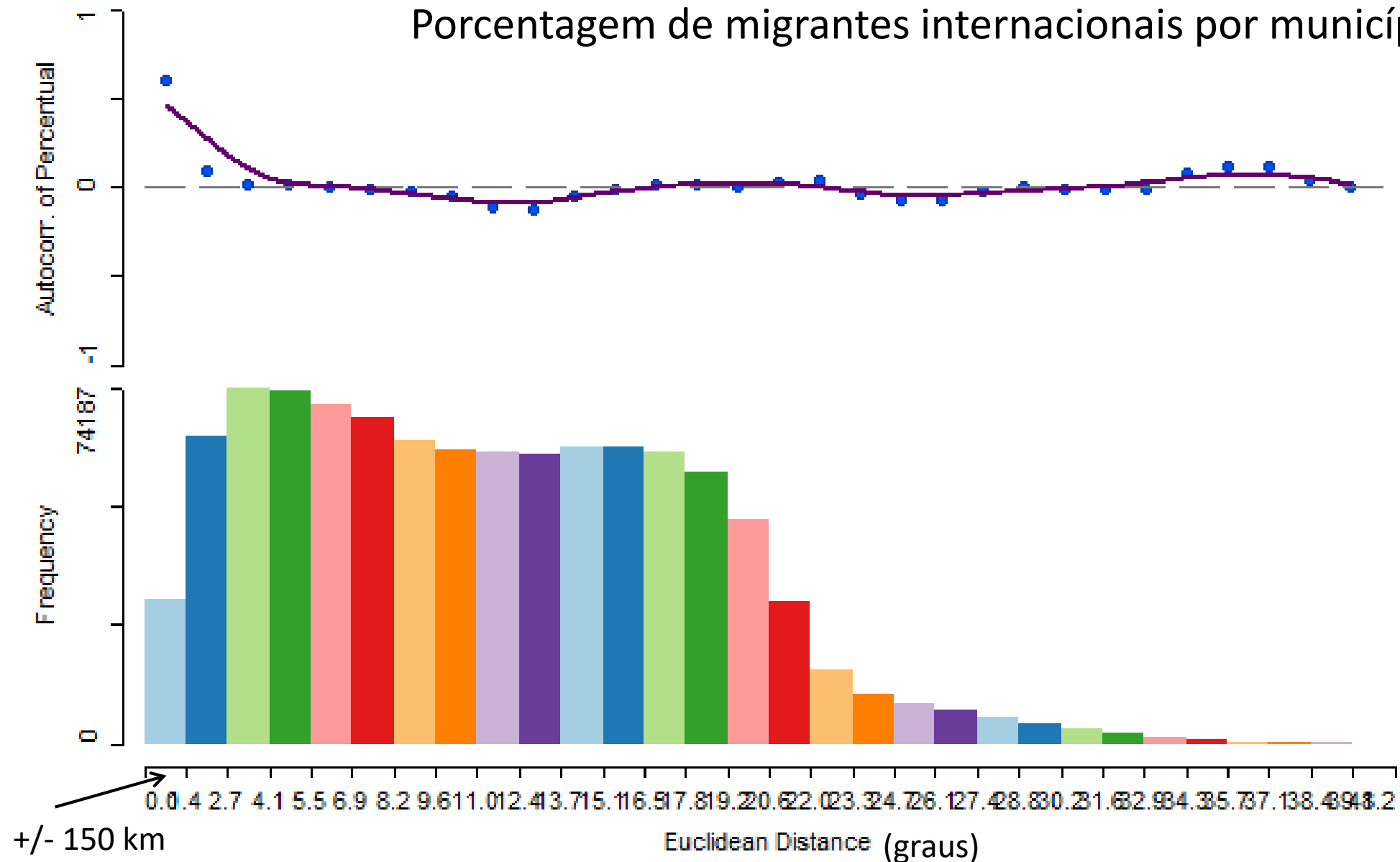
- Renda em Belo Horizonte (2000)

$$I = 0,84$$



# Correlograma

## Porcentagem de migrantes internacionais por município (Brasil, 2010)



# Z-Score e P-Value

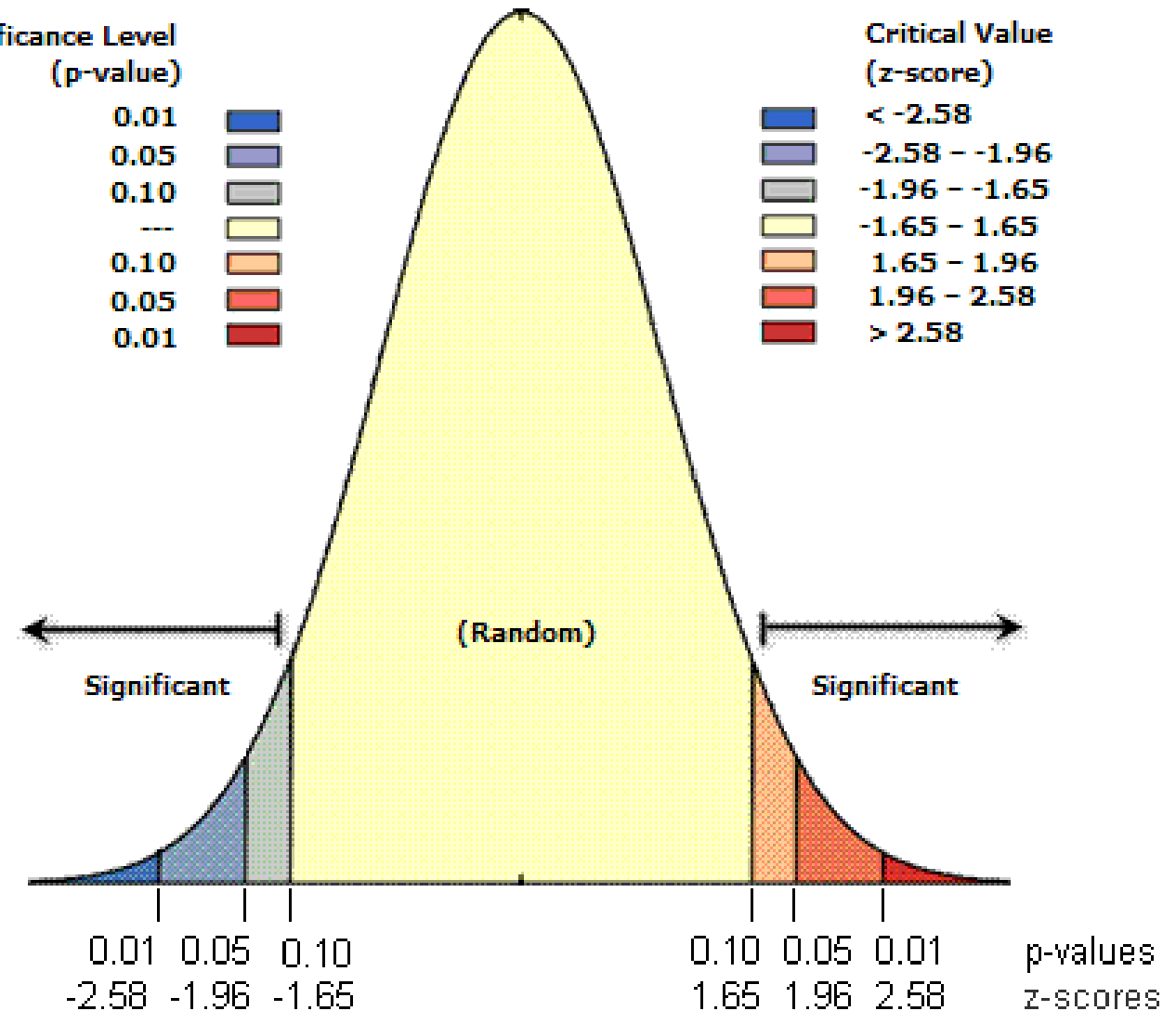
Hipótese nula: dados são randomicamente distribuídos na área de estudo

Significance Level  
(p-value)

|      |  |
|------|--|
| 0.01 |  |
| 0.05 |  |
| 0.10 |  |
| ---  |  |
| 0.10 |  |
| 0.05 |  |
| 0.01 |  |

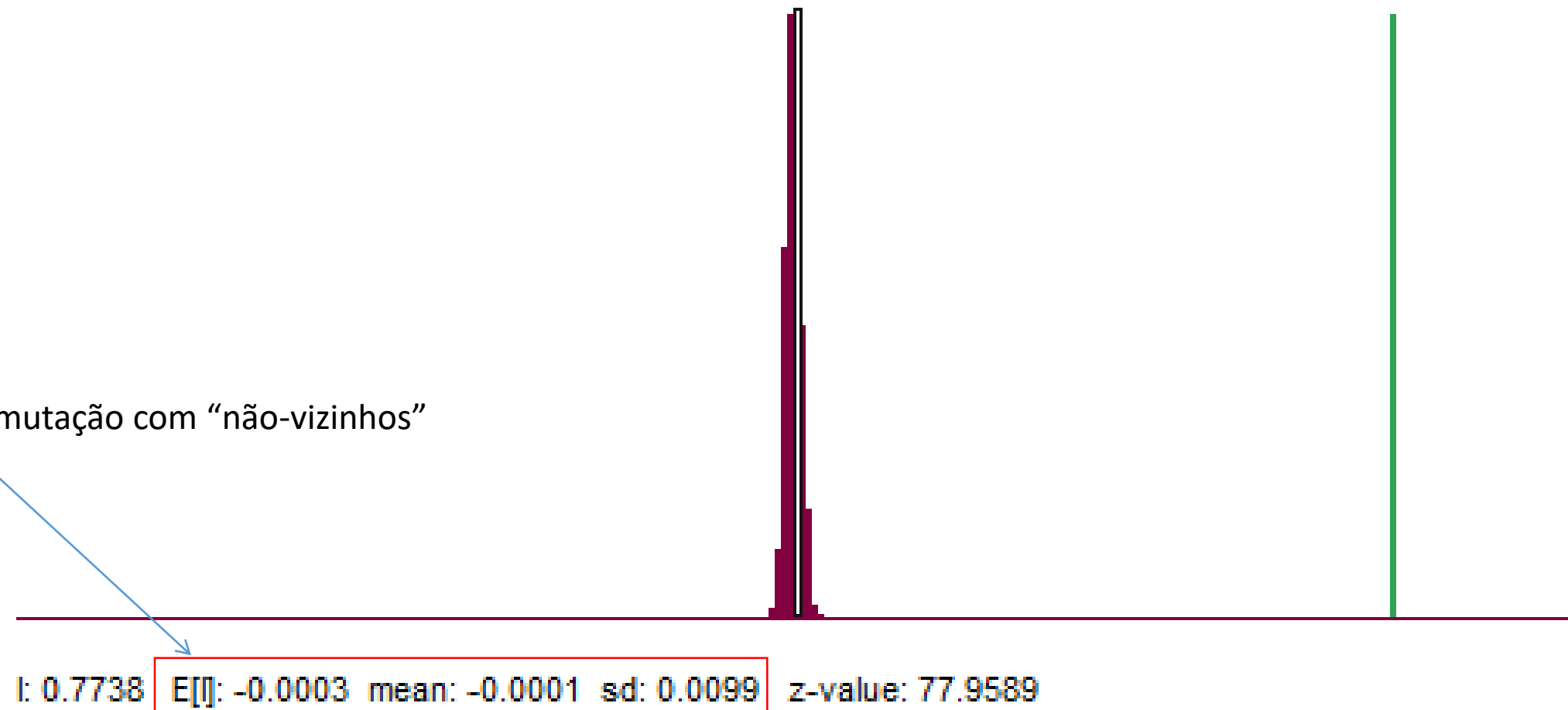
Critical Value  
(z-score)

|               |
|---------------|
| < -2.58       |
| -2.58 - -1.96 |
| -1.96 - -1.65 |
| -1.65 - 1.65  |
| 1.65 - 1.96   |
| 1.96 - 2.58   |
| > 2.58        |



permutations: 999  
pseudo p-value: 0.001000

Permutação com “não-vizinhos”



# Autocorrelação espacial local

- Índice de Moran (I) Local
- Detecta regimes de associação espacial em nível local dentro da área de estudo
- Permite construir clusters estatisticamente significantes
- A utilização destes indicadores em conjunto com os indicadores globais, refinam nosso conhecimento sobre o processos que dão origem a dependência espacial

# Autocorrelação espacial local

- Os indicadores locais de associação espacial, produzem um valor específico para cada objeto.
- Isto acarreta a identificação de:
  - “Clusters”: objetos com valores de atributos semelhantes,
  - “Outliers”: objetos anômalos,
  - A presença de mais de um regime espacial

# Autocorrelação espacial local

- Índice de Moran (I) Local

$$I_i = \left( \frac{x_i - \mu}{\left( \frac{\sum (x_i - \mu)^2}{N} \right)} \right) \sum_j w_{ij} (x_j - \mu) \quad (2)$$

Onde:

$I_i$  = Moran Local para a observação  $i$ ;

$N$  = número de observações;

$x$  = variável analisada;

$\mu$  = média de  $x$

$w_{ij}$  = elementos da matriz de pesos espaciais, nos quais  $i$  e  $j$  referem-se a localização geográfica.

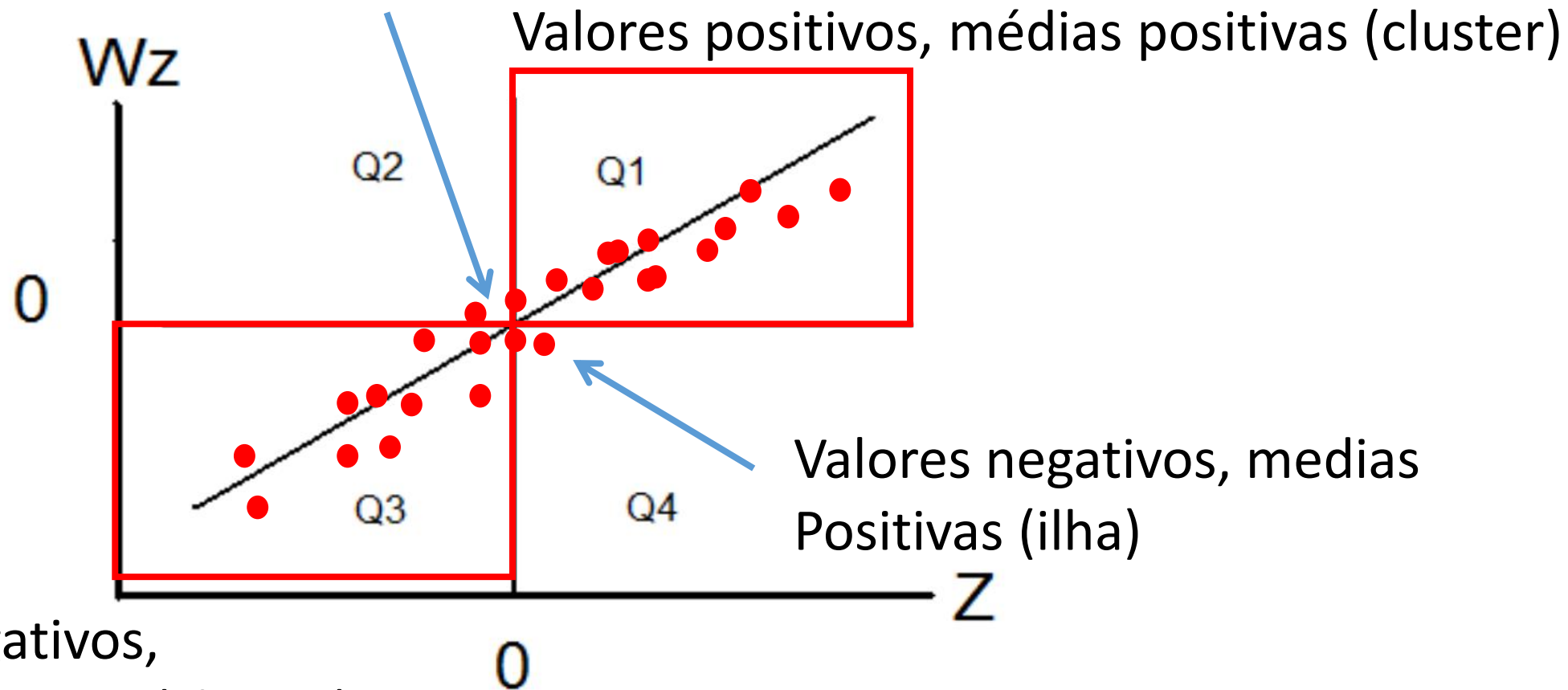
# Mapa Lisa

- Uma vez determinada a significância estatística de Moran ( $I$ ) é muito útil gerar um mapa indicando as regiões que apresentam correlação local significativamente diferente do resto dos dados.
- Este mapa é denominado por Anselin (1995) de “LISA MAP”.

# Diagrama de dispersão

- Diagrama de dispersão

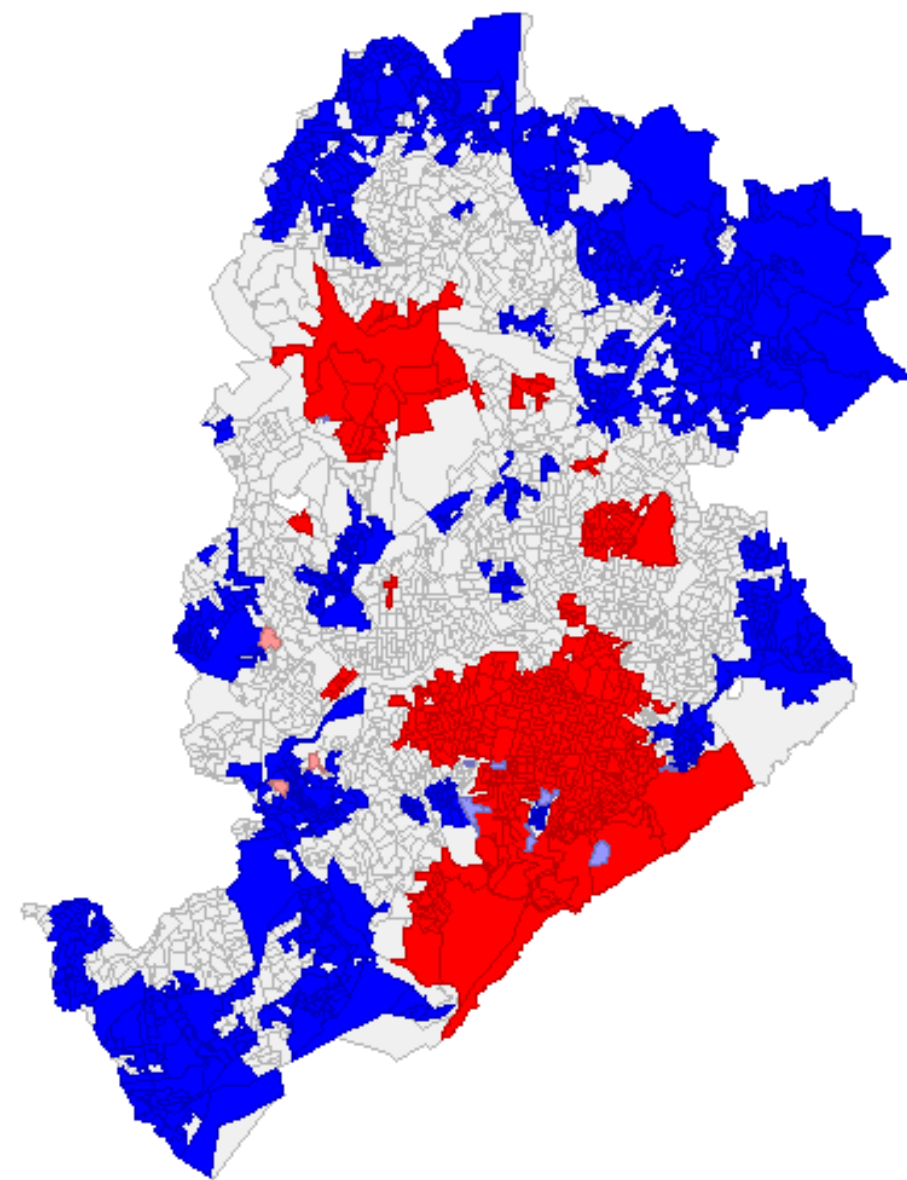
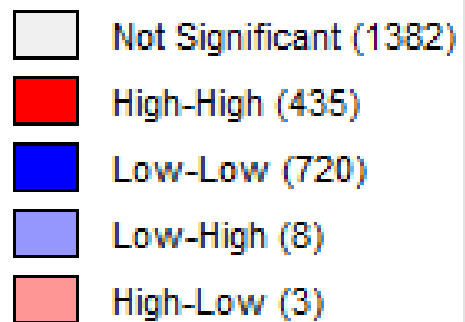
Valores positivos, médias negativas (ilha)



Valores negativos,  
médias negativas (cluster)

# Renda em Belo Horizonte, 2000

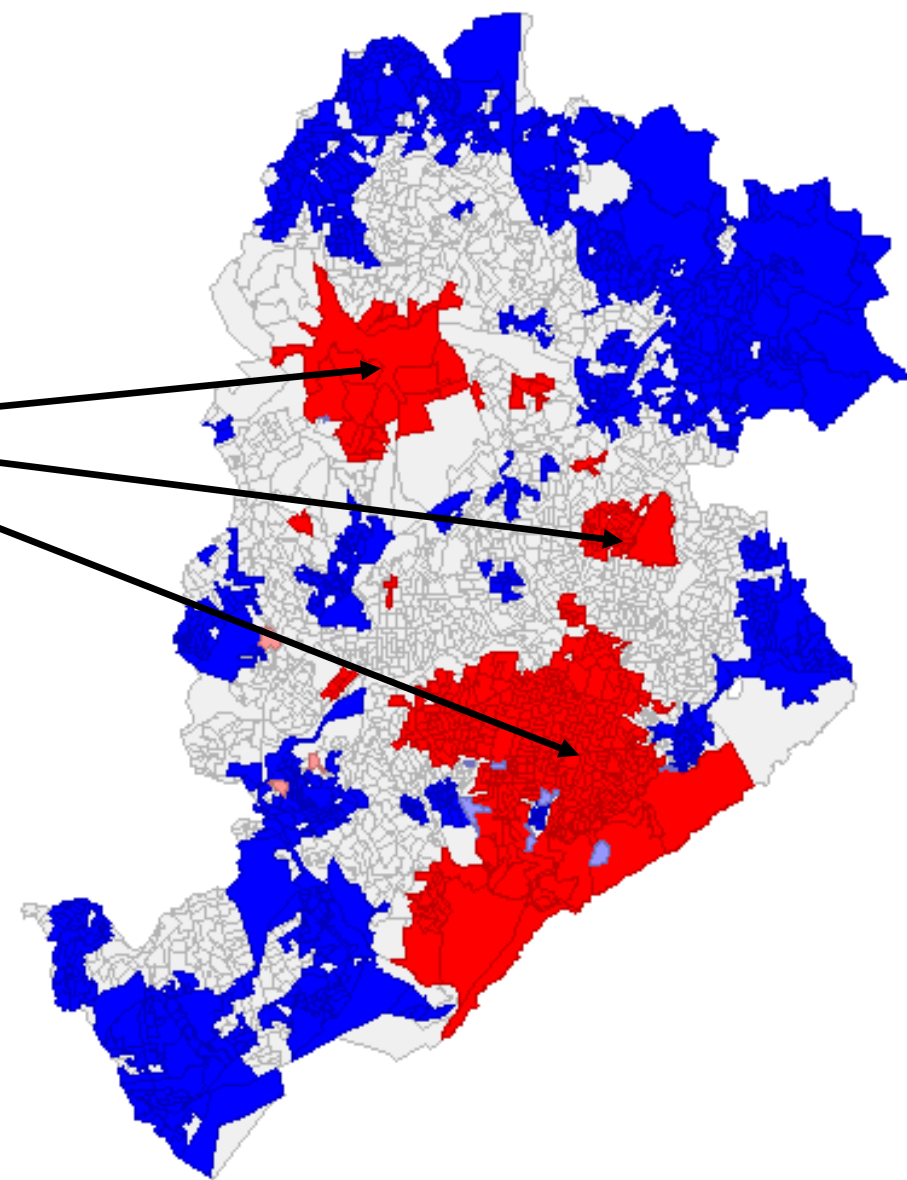
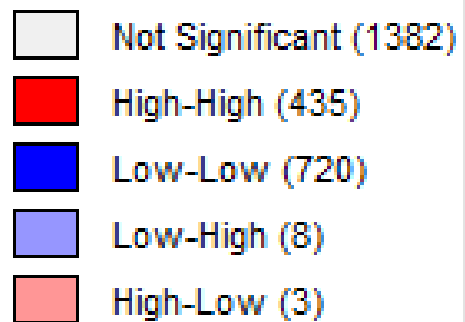
LISA Cluster Map: final, I\_RE



# Renda em Belo Horizonte, 2000

Clusters de alta renda

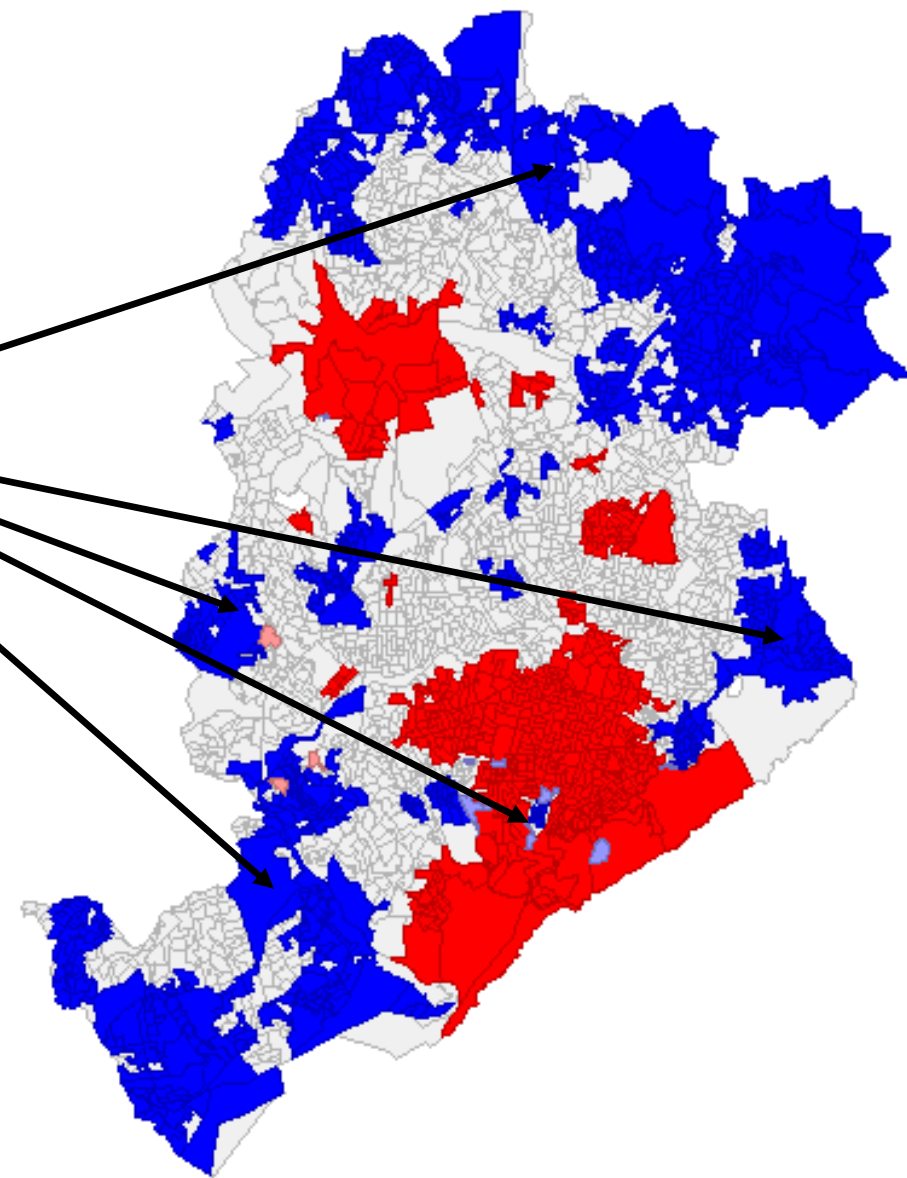
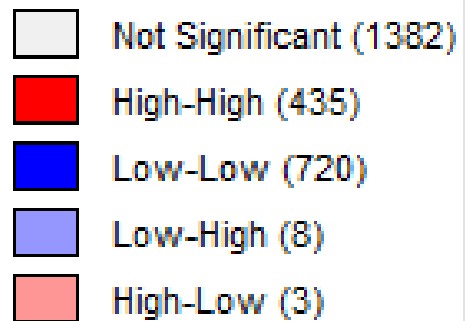
LISA Cluster Map: final, I\_RE



# Renda em Belo Horizonte, 2000

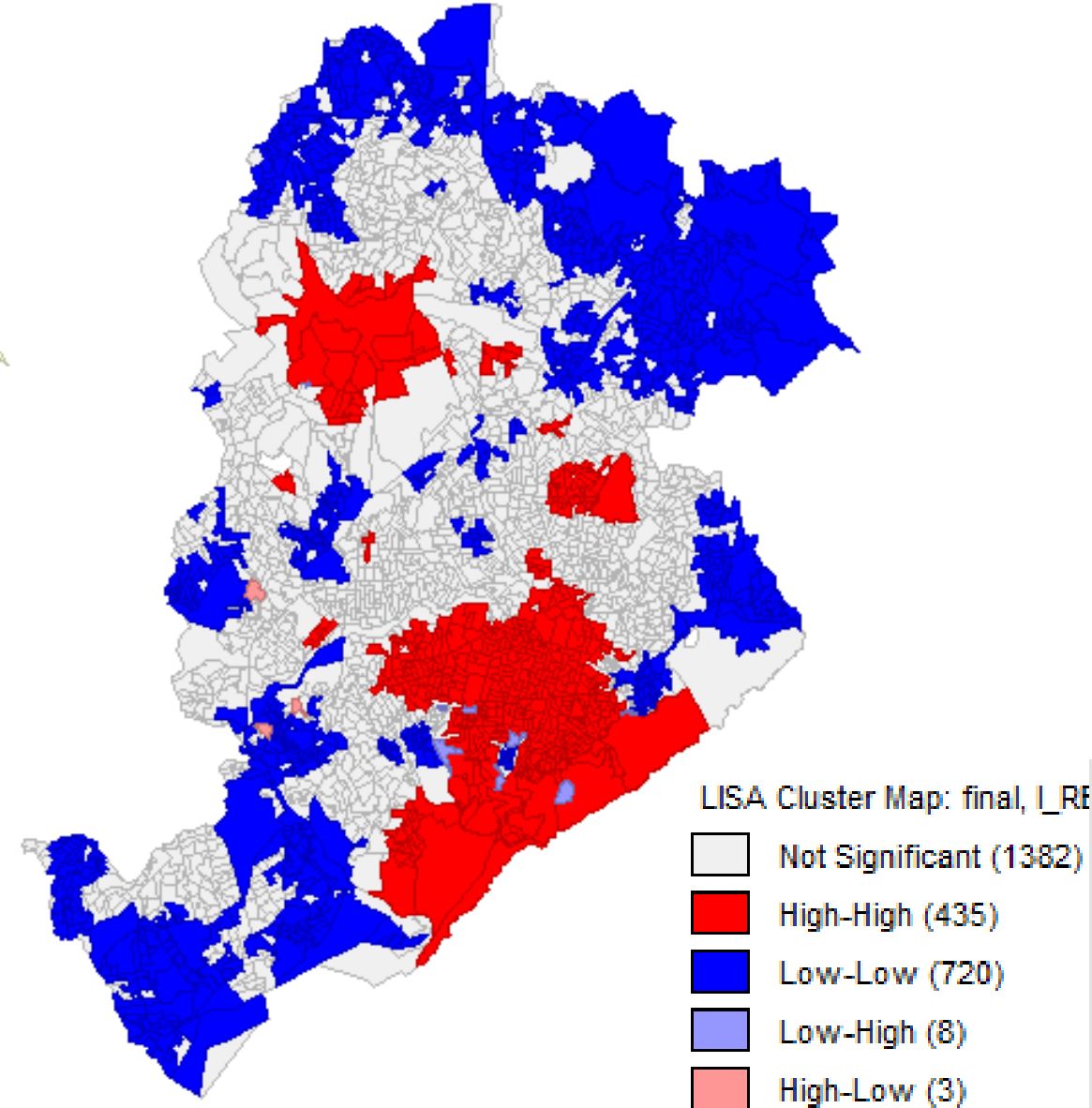
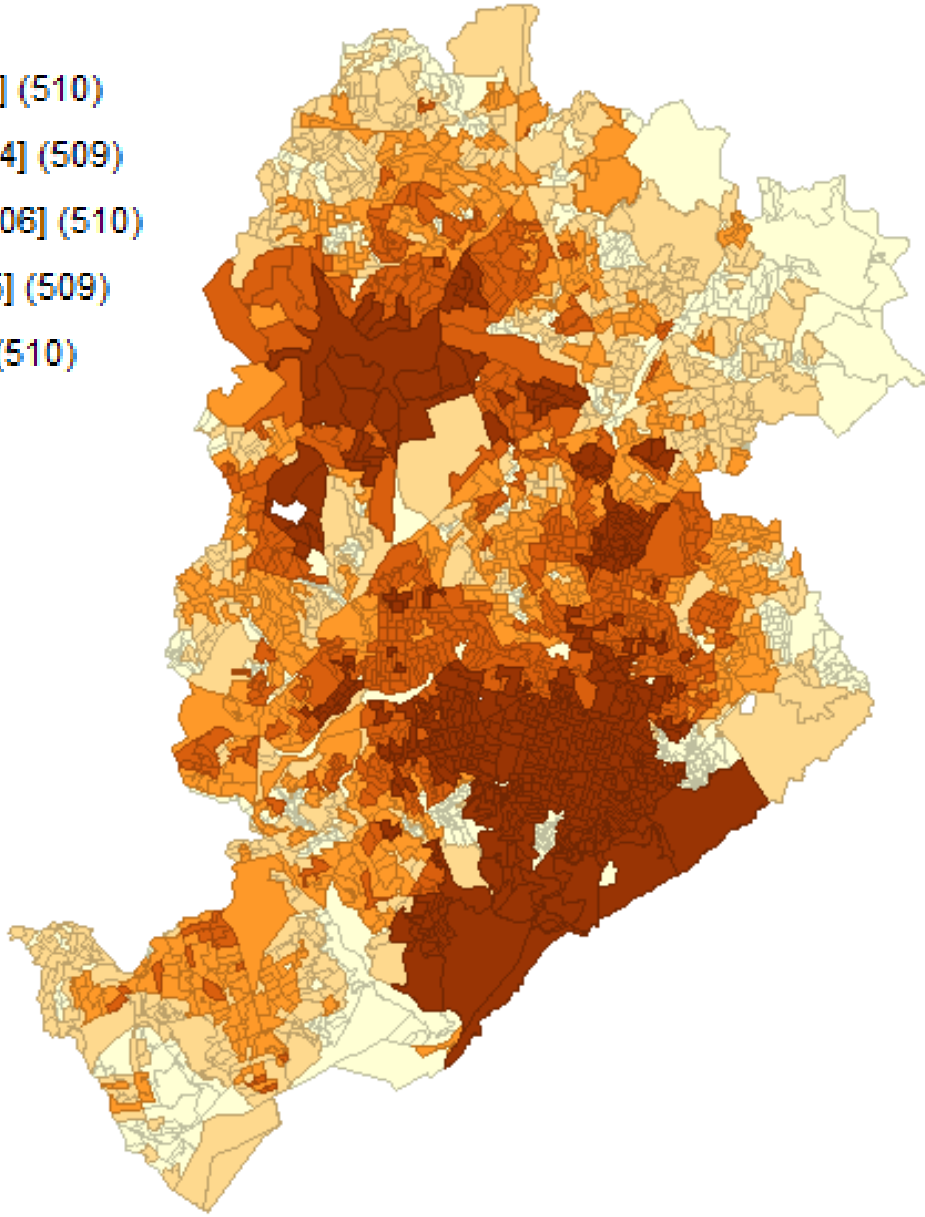
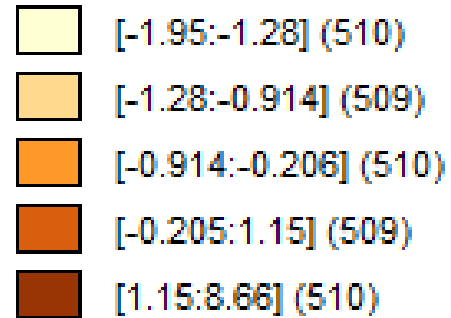
Clusters de baixa renda

LISA Cluster Map: final, I\_RE



# Mapa coroplético x Mapa Lisa

Quantile: RENDA



# Autocorrelação espacial: Aplicação

**GEOGRAFIA, Rio Claro, v. 39, n. 2, p. 257-272, mai./ago. 2014.**

**AGRUPAMENTOS DE EMIGRAÇÃO INTERNACIONAL NO BRASIL:  
O PAPEL DAS REDES SOCIAIS NA FORMAÇÃO  
DOS ESPAÇOS DE EMIGRAÇÃO**

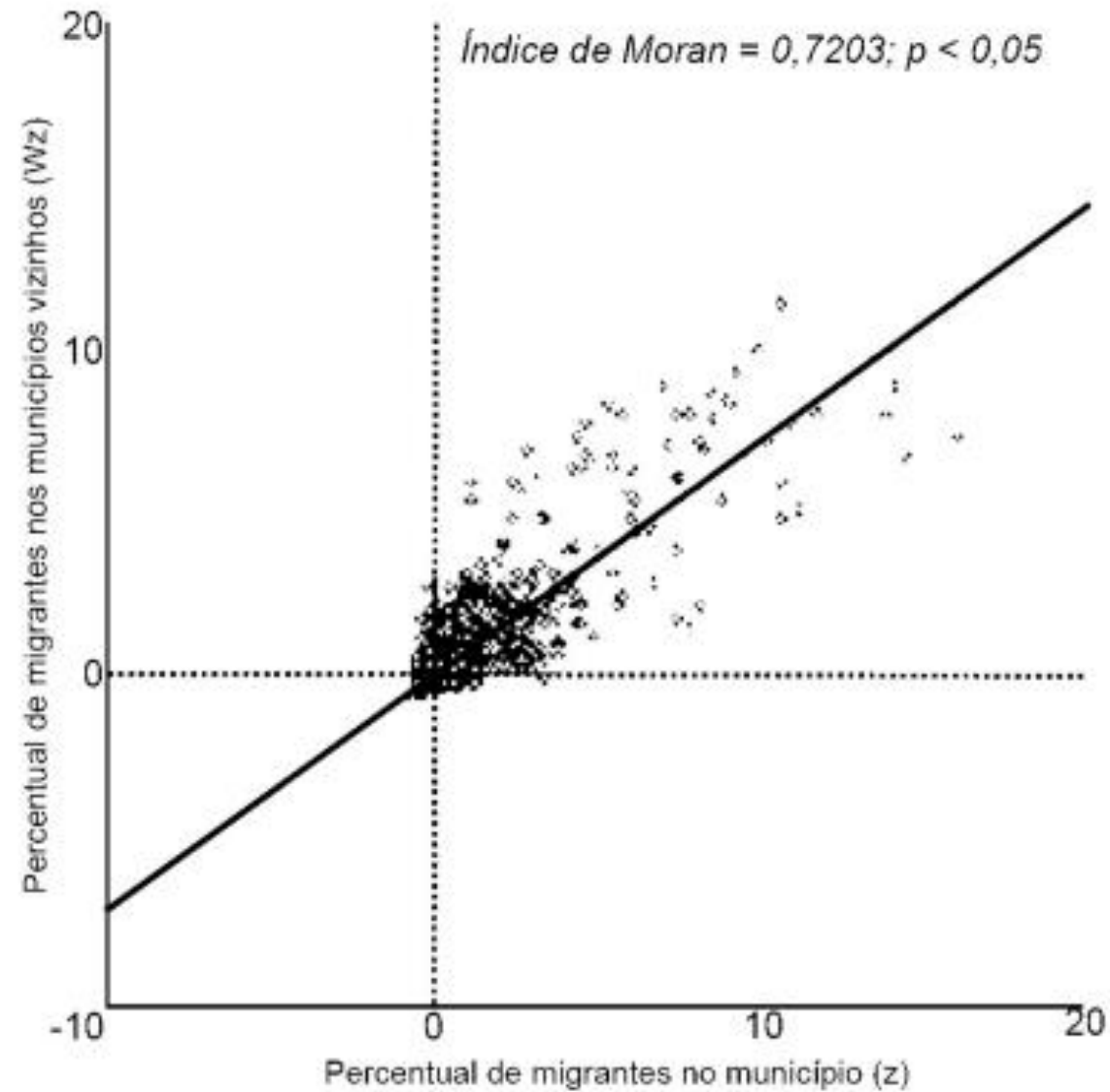
*Marden Barbosa de CAMPOS<sup>1</sup>*

*Diego Rodrigues de MACEDO<sup>2</sup>*

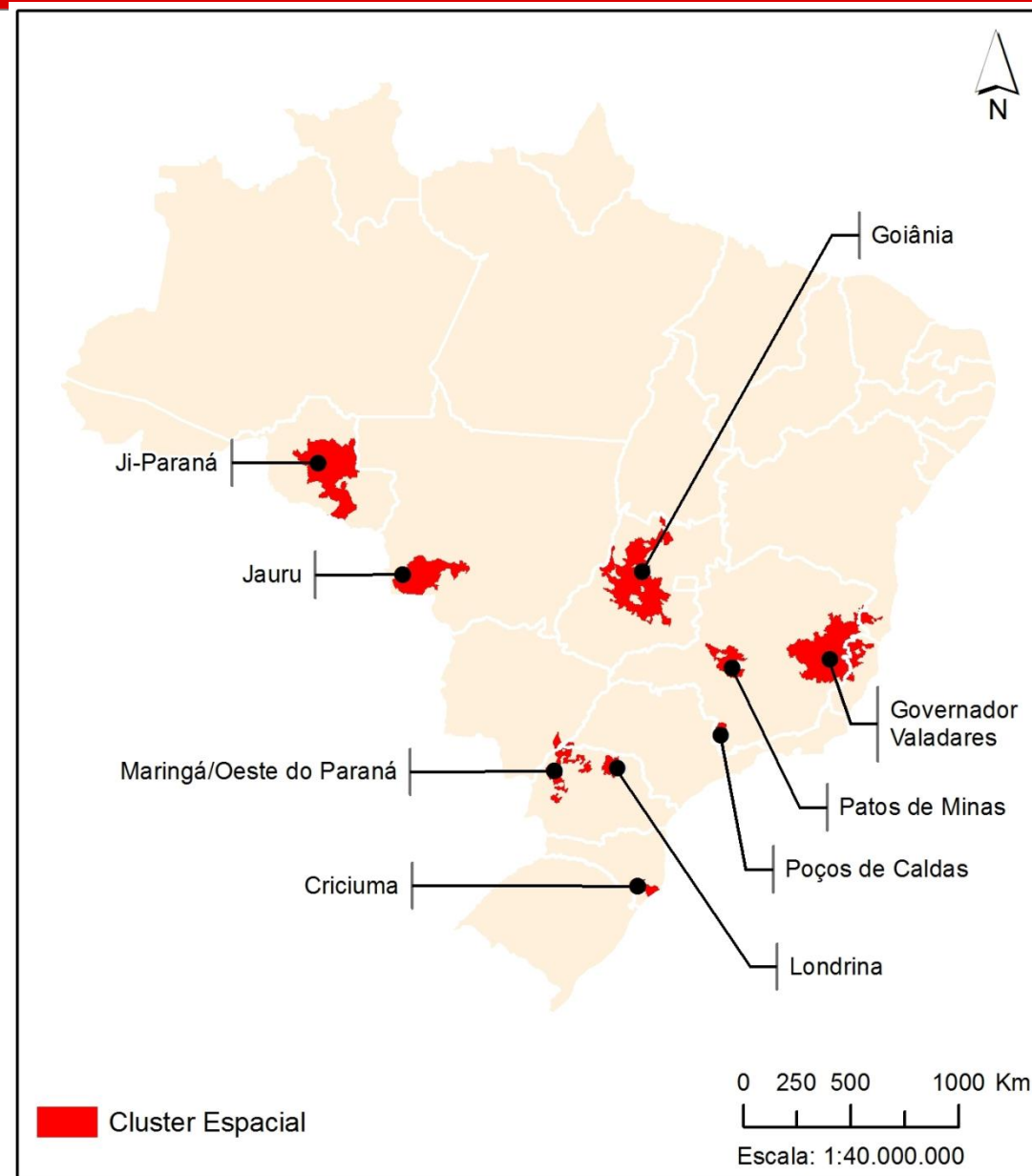
# Aplicação

- Objetivo: identificar as principais regiões de origem dos migrantes internacionais no território brasileiro
- Mapeamento de agrupamentos espaciais de emigração internacional
- Identificação dos clusters espaciais do tipo “alto-alto”, através do mapa de significancia de Moran

# Resultados



# Resultados



# Percentual e total de migrantes nos clusters

## Governador Valadares

|                |                |
|----------------|----------------|
| Estados Unidos | 55.81% (21826) |
| Portugal       | 27.82% (10880) |
| Itália         | 5.11% (2000)   |
| Espanha        | 2.75% (1075)   |
| Reino Unido    | 2.58% (1008)   |
| Outros         | 5.92% (2317)   |

## Goiânia

|                |               |
|----------------|---------------|
| Estados Unidos | 23.91% (6828) |
| Espanha        | 19.92% (5690) |
| Portugal       | 17.62% (5033) |
| Reino Unido    | 9.51% (2717)  |
| Bélgica        | 5.63% (1609)  |
| Outros         | 23.41% (6685) |

## Mariangá/Oeste do Paraná

|                |               |
|----------------|---------------|
| Portugal       | 21.09% (1049) |
| Espanha        | 19.83% (986)  |
| Japão          | 18.88% (939)  |
| Estados Unidos | 12.29% (611)  |
| Reino Unido    | 12.13% (603)  |
| Outros         | 15.79% (785)  |

## Ji Paraná

|                |               |
|----------------|---------------|
| Espanha        | 29.13% (1139) |
| Portugal       | 28.03% (1096) |
| Estados Unidos | 25.22% (986)  |
| Bolívia        | 8.26% (323)   |
| Reino Unido    | 1.89% (74)    |
| Outros         | 7.47% (292)   |

## Patos de Minas

|                |               |
|----------------|---------------|
| Estados Unidos | 65.38% (1590) |
| Reino Unido    | 12.17% (296)  |
| Espanha        | 8.55% (208)   |
| Bélgica        | 3.13% (76)    |
| Portugal       | 2.96% (72)    |
| Outros         | 7.81% (190)   |

## Jauru

|                |              |
|----------------|--------------|
| Estados Unidos | 55.23% (660) |
| Bolívia        | 16.49% (197) |
| Espanha        | 11.63% (139) |
| Portugal       | 4.85% (58)   |
| Japão          | 2.18% (26)   |
| Outros         | 9.62% (115)  |

## Criciúma

|                |              |
|----------------|--------------|
| Estados Unidos | 14.25% (835) |
| Itália         | 13.2% (773)  |
| Alemanha       | 10% (586)    |
| Portugal       | 6.59% (386)  |
| Reino Unido    | 5.96% (349)  |
| Outros         | 50% (2929)   |

## Poços de Caldas

|                |               |
|----------------|---------------|
| Estados Unidos | 61.89% (1163) |
| Portugal       | 8.89% (167)   |
| Itália         | 8.3% (156)    |
| Espanha        | 5.91% (111)   |
| Reino Unido    | 2.93% (55)    |
| Outros         | 12.08% (227)  |

## Londrina

|                |               |
|----------------|---------------|
| Japão          | 28.09% (2081) |
| Portugal       | 17.96% (1331) |
| Reino Unido    | 16.13% (1195) |
| Estados Unidos | 12.47% (924)  |
| Espanha        | 6.69% (496)   |
| Outros         | 18.65% (1382) |

Fonte: IBGE, 2010.

# Aplicação

- Resultados corroboram a hipótese de que a migração é impulsionada por redes já estabelecidas nos países que recebem os migrantes

# I-Moran como diagnóstico em modelos de regressão

# Suposições do modelo de regressão

- 1) **Distribuição Normal** Para um valor fixo da variável aleatória  $X$ ,  $Y$  é uma variável aleatória com distribuição Normal (com média e variâncias finitas);

$$Y_i \sim N(E(y/x); \sigma^2)$$

- 2) **Linearidade**

Todos os valores médios de  $Y$  ( $E(y/x)=\mu_{y/x}$ ) permanecem sobre uma reta, para um particular valor de  $X$ .

$$E(y/x)=\mu_{y/x} = \beta_0 + \beta_1 x$$

- 3) **Independência**

Os valores de  $Y$  são estatisticamente independentes.

- 4) **Homocedasticidade**

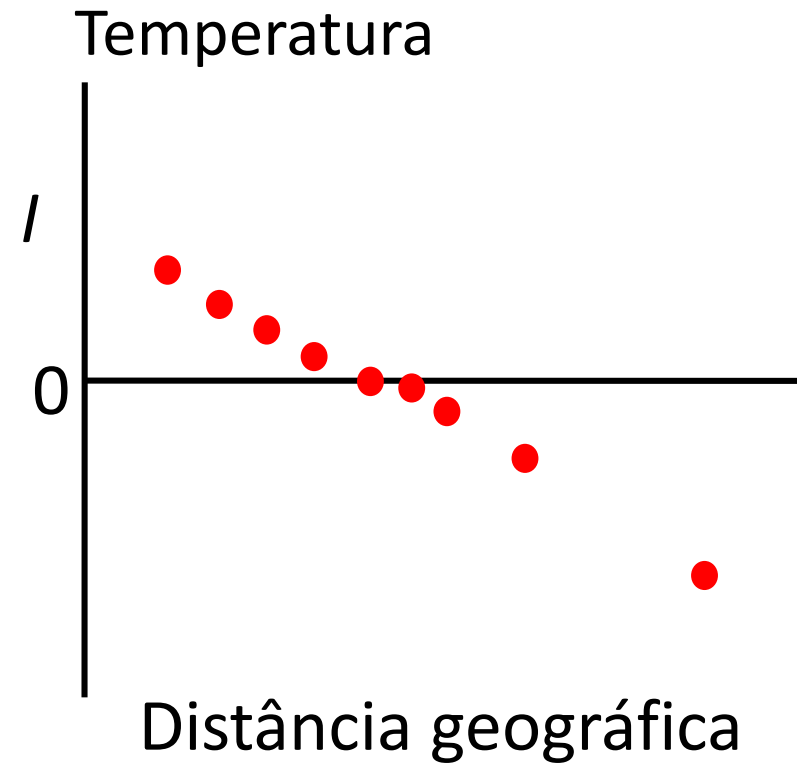
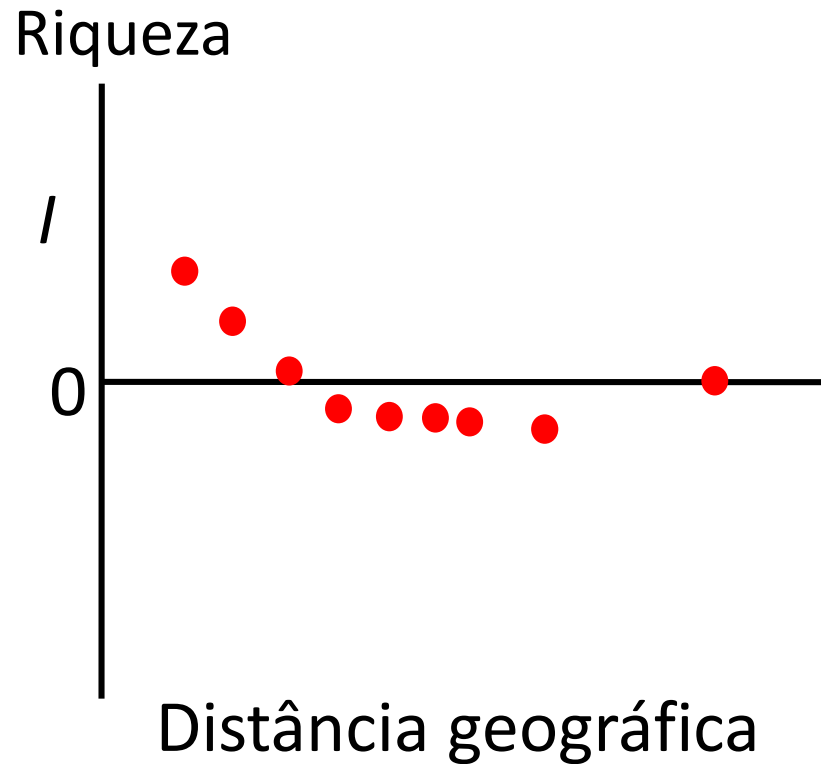
A variância de  $Y$  é igual, qq que seja  $X$ .

# Análise dos Resíduos

- Hipótese de independência das observações em geral é Falsa → **Dependência Espacial**
- Efeitos Espaciais
  - Se existir forte tendência ou correlação espacial, os resultados serão influenciados, apresentando associação estatística onde não existe (e vice-versa) Erro Tipo I e II
- Como verificar?
  - Medir a autocorrelação espacial dos resíduos da regressão (Índice de Moran dos resíduos)

# Resíduos: como NÃO DEVEM ser

## Dependencia espacial



# I-Moran dos resíduos do modelo de regressão

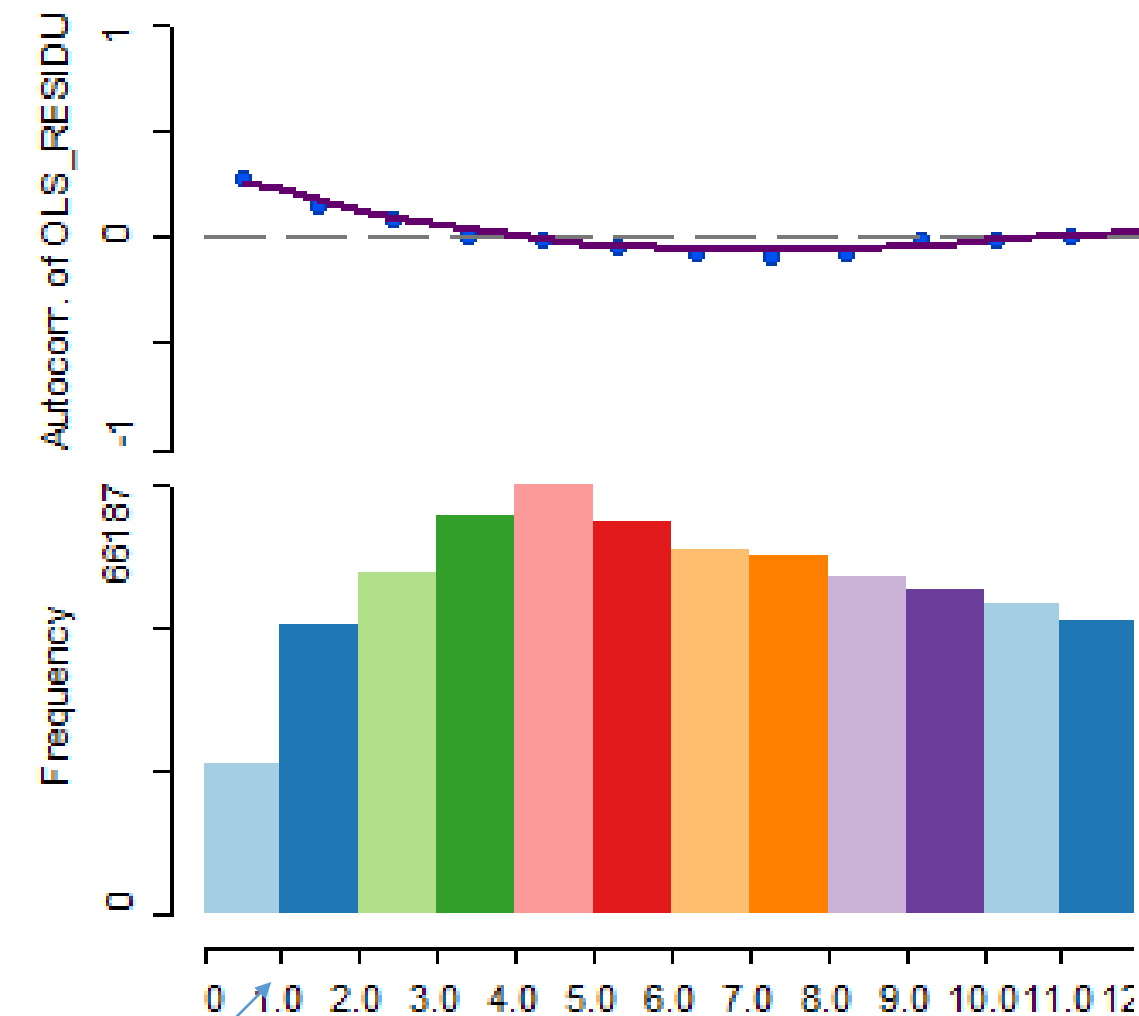


Table 2. Lagrange multiplier diagnostics for spatial dependence

|        | In-migration | Out-migration |
|--------|--------------|---------------|
| LMerr  | 84.94***     | 127.88***     |
| LMlag  | 154.60***    | 136.70***     |
| RLMerr | 4.48*        | 5.35*         |
| RLMlag | 74.14***     | 14.18***      |
| SARMA  | 159.08***    | 142.06***     |

Source: IBGE (2000; 2010). PNUD (2010). INPE (2000-2010).

\*\*\*  $p < 0.001$ ;

\*\*  $p < 0.01$ ;

\*  $p < 0.05$

# Regressão Espacial

- Autocorrelação espacial constatada!

**E agora?**

- Modelos de regressão que incorporam efeitos espaciais:
  - **Globais:** utilizam um único parâmetro para capturar a estrutura de correlação espacial
  - **Locais:** parâmetros variam continuamente no espaço

# *Spatial Lag Model (LAG)*

- Suposição

- a variável  $Y_i$  depende dos valores da variável dependente nas áreas vizinhas a  $i$ :

$$Y = \rho WY + X\beta + \varepsilon$$

- $\rho$  = coeficiente espacial autoregressivo - medida de correlação espacial
- $\rho = 0$ , se autocorrelação é nula (hipótese nula)
- $W$  = matriz de proximidade espacial
- $WY$  expressa a dependência espacial em  $Y$

# Aplicação: Modelos de Regressão Espacial Global

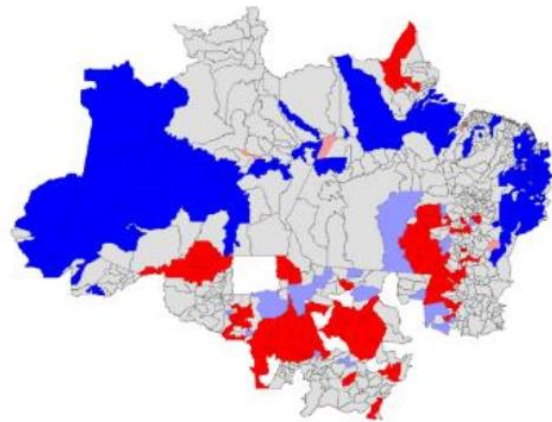
# Assessing the regional context of migration in the Brazilian Amazon through spatial regression modeling

Douglas Sathler; Susana Adamo; Everton E. C. Lima; Diego Rodrigues Macedo; Alex de Sherbinin; Paola Kim-Blanco

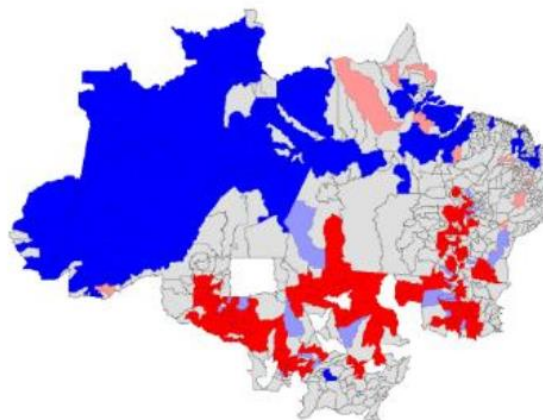
Submetido: Applied Geography

- (1) Existe alguma dependência espacial nas variáveis exploradas e, em caso afirmativo, como essa característica deve ser adequadamente abordada nos modelos de regressão espacial?
- (2) Quais fatores contextuais explicam significativamente as taxas de imigração e emigração (data fixa) na Amazônia brasileira?

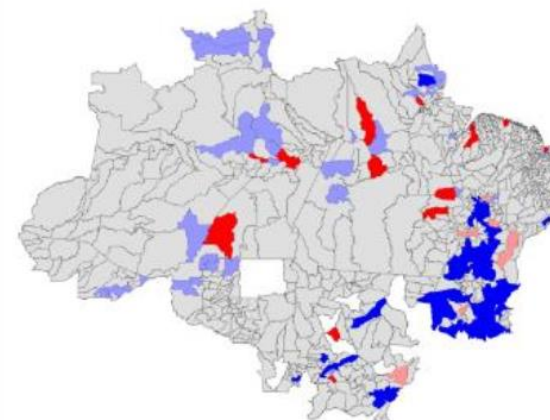
In-migration rate



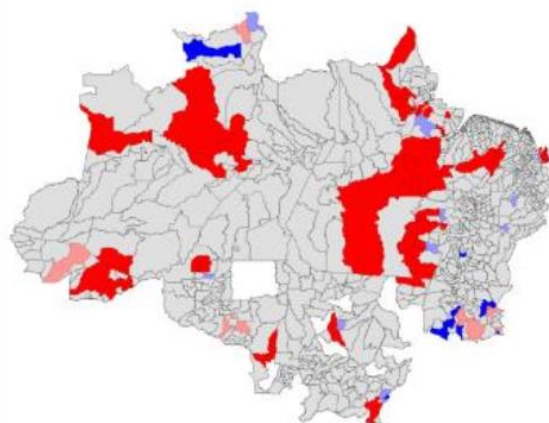
Out-migration rate



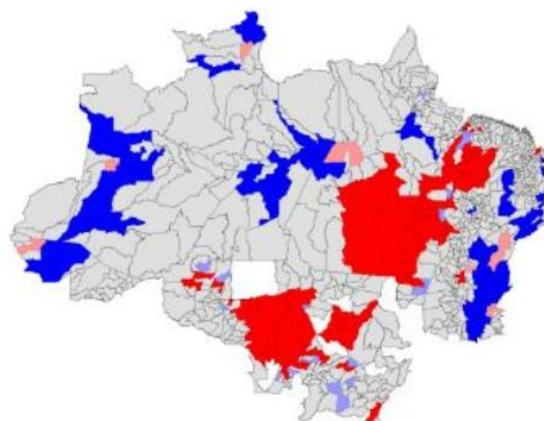
Population



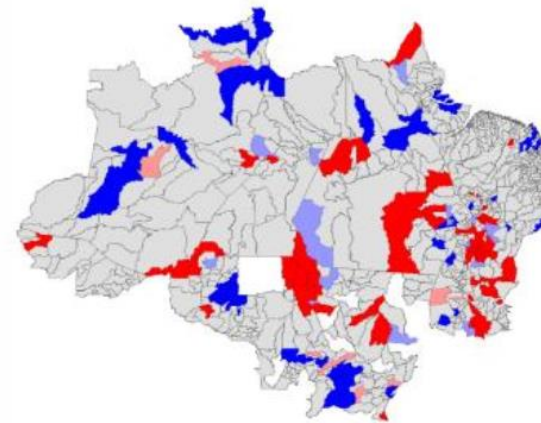
Urban Population Growth



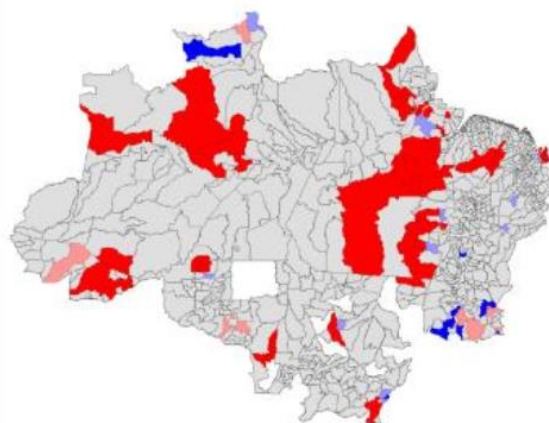
% Population with 20 to 39 years



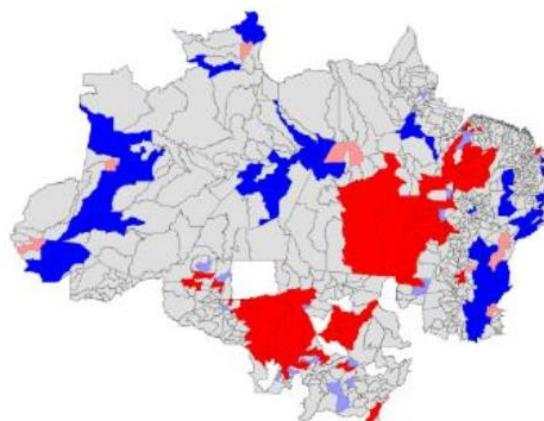
Variation in GDP



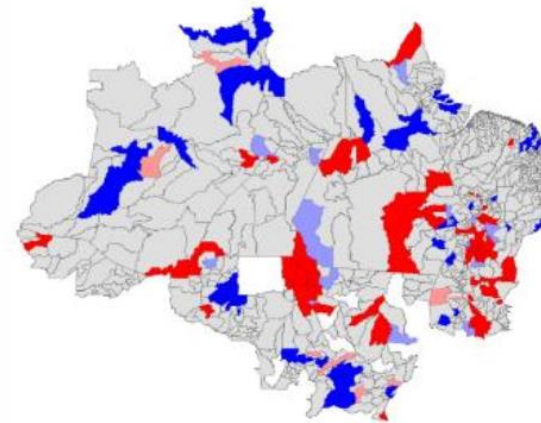
% Services GDP

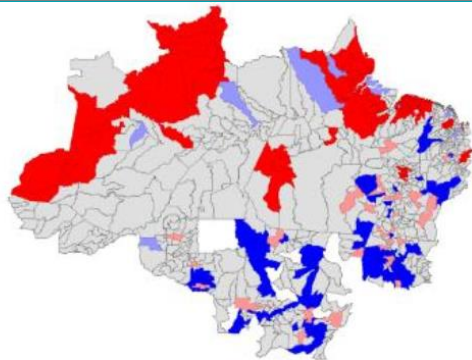


HDI

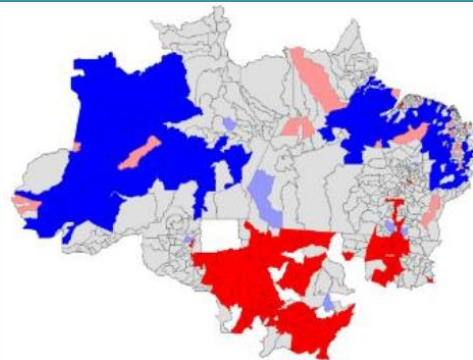


Variation in HDI

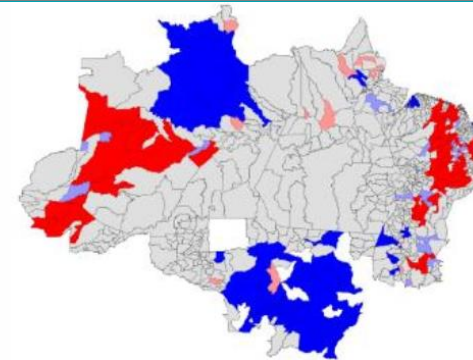




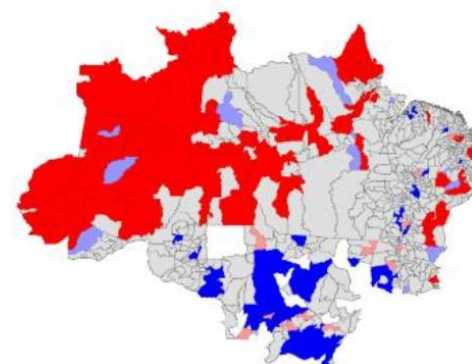
GINI Index



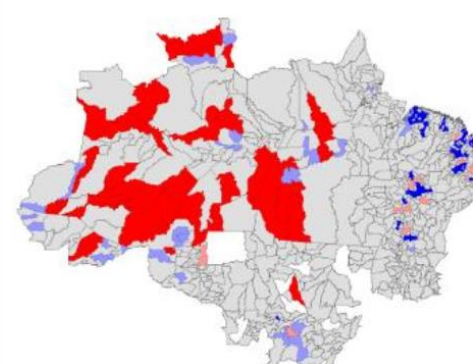
Distance to the closest major center



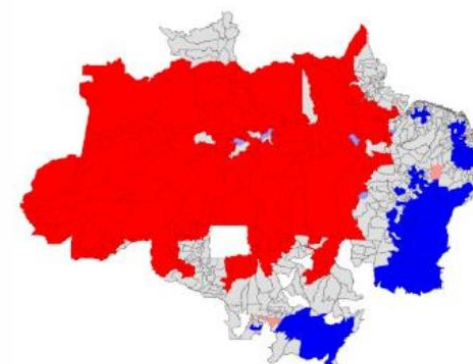
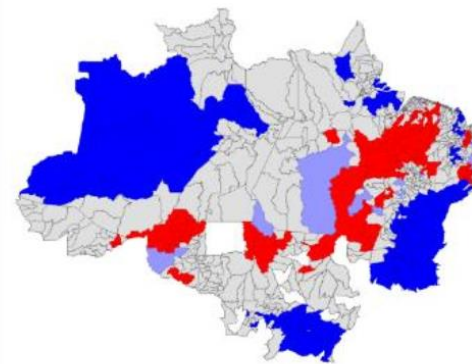
Distance to a paved road



% Deforested area



% Forest area



0 1,000 2,000 Km



Cartographic data: IBGE, 2010.  
Geographic projection  
Datum: Sirgas 2000.  
Date: 06/09/2016.

# Modelos Ajustados

|               | In-migration |         | Out-migration |         |
|---------------|--------------|---------|---------------|---------|
|               | OLS          | SLM     | OLS           | SLM     |
| Adj-R2        | 0.438***     |         | 0.346***      |         |
| Rho           | 0.513***     |         | 0.466***      |         |
| AIC           | 4383.17      | 4236.37 | 4018.7        | 3900.48 |
| *** p < 0.001 |              |         |               |         |

# Resultados

- Os resultados mostraram que as variáveis e os modelos ajustados possuem dependência espacial

# Modelos com Efeitos Espaciais Locais

# Modelos com Efeitos Espaciais Locais

- Modelos de Regressão com Efeitos Espaciais **Contínuos**
  - variações espaciais modeladas de forma contínua, com parâmetros variando no espaço.
  - “*Geographically Weighted Regression*” – GWR.

# Modelos com Efeitos Espaciais Locais

- Uma maneira de incorporar o termo espacial em uma regressão é utilizar a Regressão Geograficamente Ponderada (GWR)

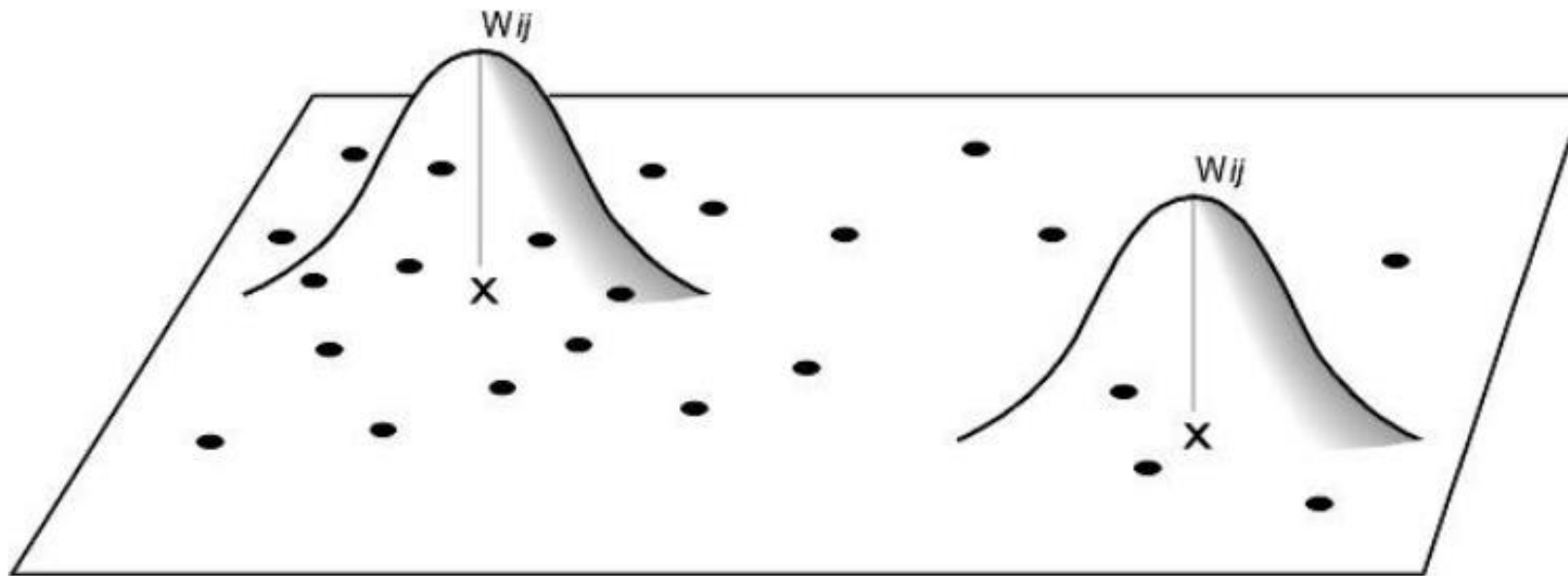
Regressão clássica:  $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

Parâmetros  $\beta_0 + \beta_1$  são iguais para todos os valores estimados

# Modelos com Efeitos Espaciais Locais

- Regressão Geograficamente Ponderada, cada Y possuirá um valor para  $\beta_0 + \beta_1$

$$y(i) = \beta_0(i) + \beta_1(i)x_i + \varepsilon_i$$





***XXI Encontro Nacional de Estudos Populacionais***  
*População, sociedade e políticas: desafios presentes e futuros*  
*ST: Tendências na distribuição espacial da população,*  
*desenvolvimento e urbanização 2*

**Renda e condições de vida nos municípios de Minas Gerais:  
possibilidades a partir da Regressão Geograficamente Ponderada  
(GWR)**

Júlio César Paiva  
Carlos Lobo  
Diego Macedo

Universidade Federal de Minas Gerais  
PPG Análise e Modelagem de Sistemas Ambientais  
Setembro/2018

# Objetivo e Hipótese

- Analisar a relação entre renda e indicadores sociais (mortalidade infantil, analfabetismo e porcentagem de pobres) nos municípios mineiros com base nos dados do último Censo Demográfico (2010).
- A renda não se materializa necessariamente em melhores condições de vida?

Porcentagem de  
Pobres

Mortalidade  
Infantil

Taxa de  
Analfabetismo

Renda Média  
Per Capita

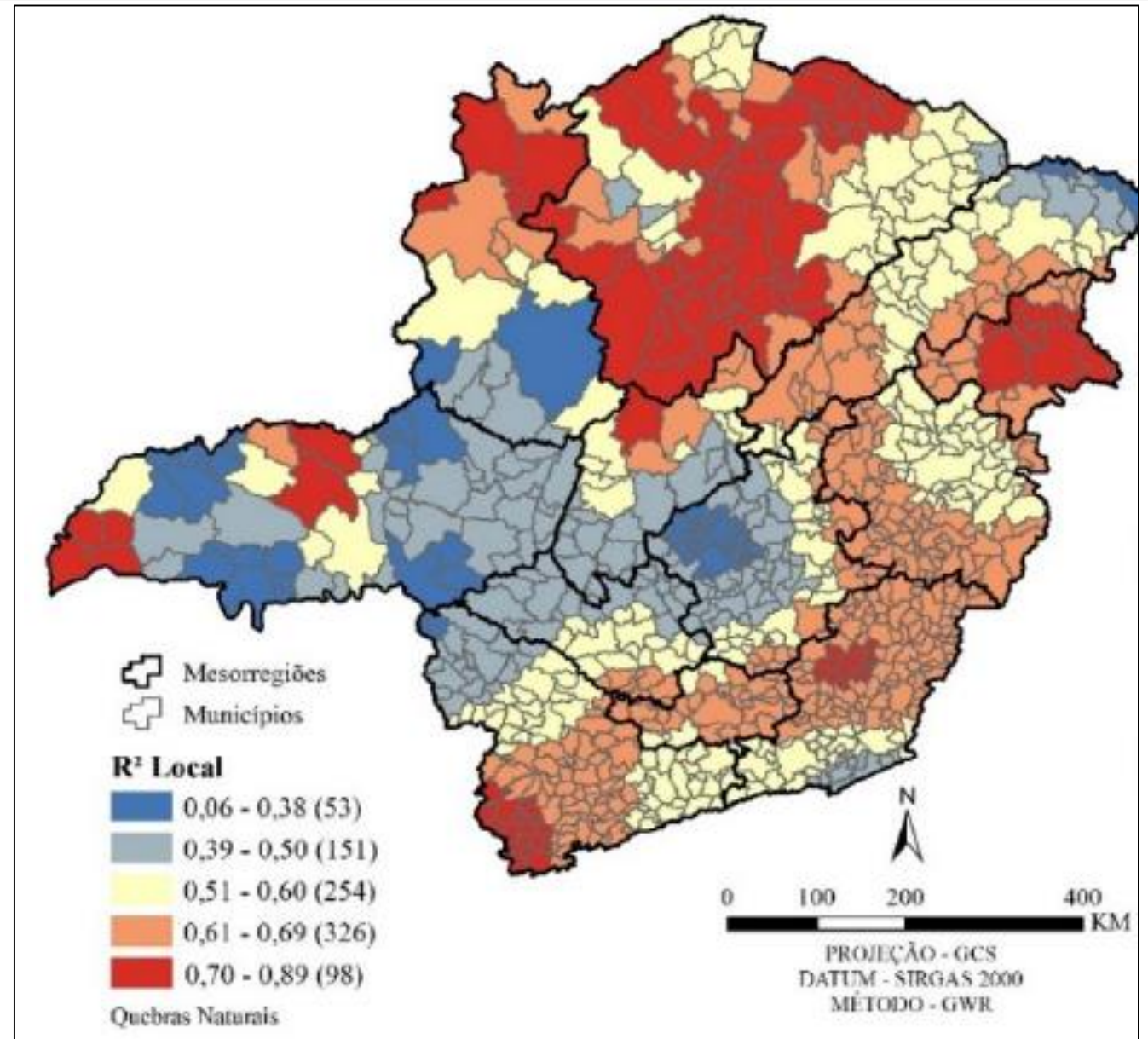


ICV

$R^2=0,69$

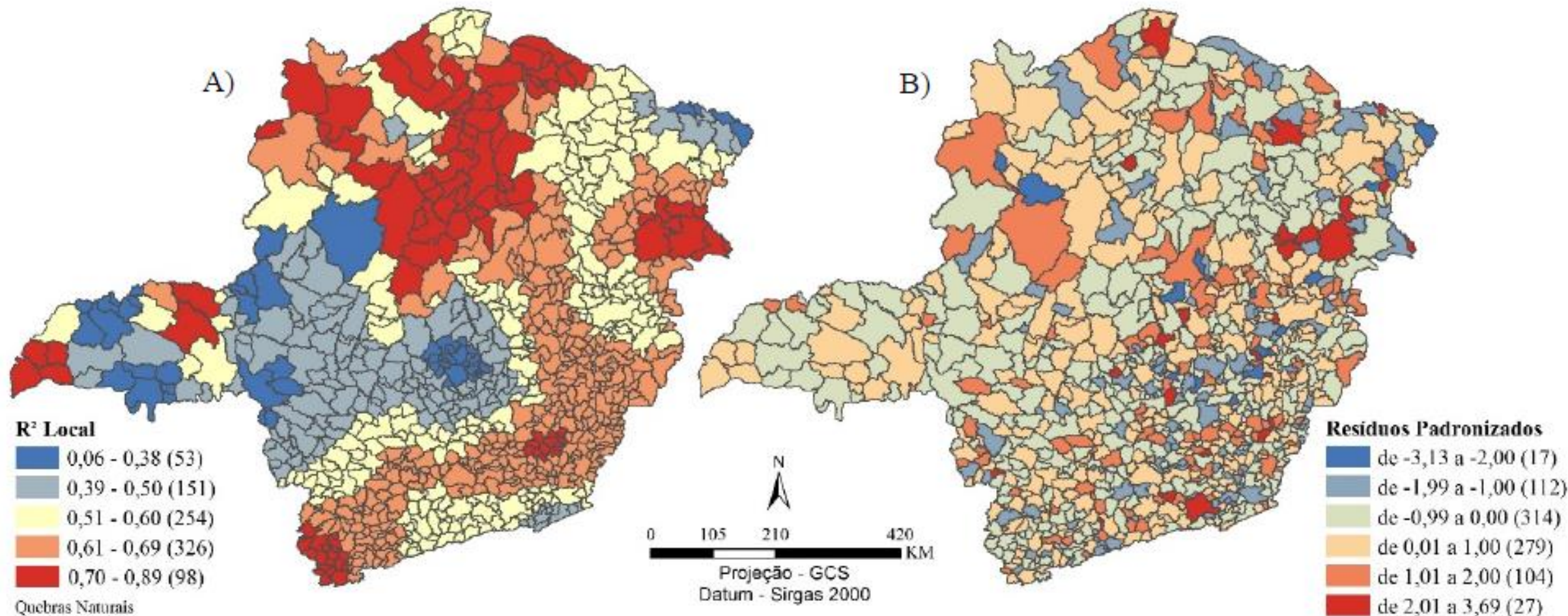
# Resultados

- Relação direta entre ICV e Renda ( $R^2$  Global 0,89);
- Modelo não espacial  $R^2=0,69$
- Escores de  $R^2$  Local altos (678 municípios +0,51 / 175 municípios -0,50);
- (+) Norte de Minas e Mucuri;
- (-) Triângulo/Alto Paranaíba e Metropolitana de BH.



# Resultados

**Figura 3 - ICV e Renda média per capita nos municípios de Minas Gerais, 2010 – Regressão Espacialmente Ponderada (GWR):  $R^2$  e Resíduos Padronizados**



- I de Moran 0,01; P-valor 0,26

- No geral, a renda se materializa em melhores condições de vida.
- No entanto, localmente estas variações são distintas.
- Disparidade entre os municípios e regiões mineiras: Triângulo/Alto Paranaíba se as relações são mais baixas, e Norte de Minas, Vales do Jequitinhonha e Mucuri, as relações são maiores.