

Análise espacial

Patrícia de Siqueira Ramos

UNIFAL-MG, *campus* Varginha

15 de Agosto de 2019

Conteúdo

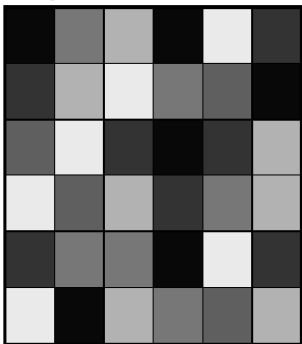
- Introdução à análise espacial
- Dados espaciais
 - Fonte de dados espaciais
 - Processo estocástico espacial
 - Dados espaciais e inferência estatística
 - Problemas com dados espaciais
- Matrizes de ponderação espacial
 - Normalização da matriz de pesos espaciais
 - Operador de defasagem espacial
 - Propriedades das matrizes de ponderação espacial
 - Escolha da matriz de ponderação espacial
- Análise exploratória de dados espaciais
 - Autocorrelação espacial global e local
 - *Clusters* espaciais
 - *Outliers* espaciais

Metodologia das aulas

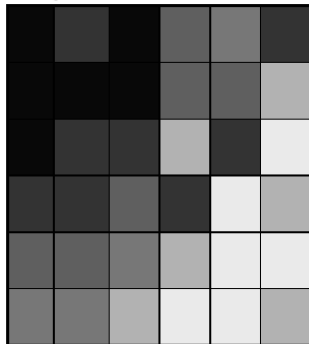
- Livro texto principal:
 - ALMEIDA, E. *Econometria espacial*. Campinas-SP. Alínea (2012).
- Aulas:
 - Expositivas
 - Leitura e discussão de artigos
 - Análise de dados espaciais usando a linguagem Python
- Avaliação:
 - Prova: 04/10/2019 - peso 30%
 - Atividades para entrega/apresentação - peso 20%
 - Trabalho - peso 50%
 - Prova Especial: 02/12/2019

Motivação

Sem levar em conta o padrão espacial, uma análise tradicional dos dois mapas levaria aos mesmos resultados



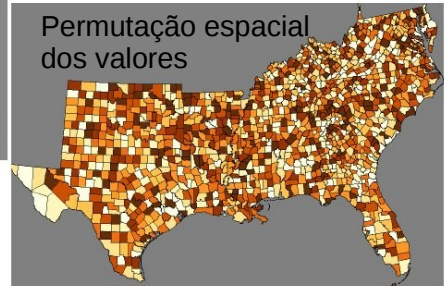
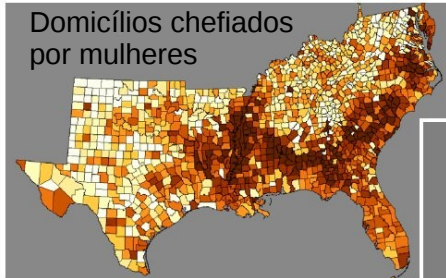
Mapa 1



Mapa 2

Motivação

O mesmo ocorreria aqui (EUA - do Texas à Flórida):



Componentes da análise espacial

- Visualização:
Mostrar padrões interessantes
- Análise Exploratória de Dados Espaciais (AEDE):
Encontrar padrões interessantes
- Modelagem espacial:
Explicar padrões interessantes
(Não veremos)

Correlação espacial

- H_0 : aleatoriedade espacial é ausência de padrão espacial
- Se H_0 for rejeitada há evidência de estrutura espacial (queremos rejeitá-la!)
- Aleatoriedade espacial: padrão espacial observado tem a mesma probabilidade de ocorrer que qualquer outro padrão, assim, o valor de uma variável em um local não depende do valor nos outros locais

Aleatoriedade espacial

LOCAL 1	VALOR 1
LOCAL 2	VALOR 2
LOCAL 3	VALOR 3
⋮	⋮

Quando há aleatoriedade espacial, qualquer valor pode estar em qualquer local.

Difícil: obter a distribuição de valores nos locais sob H_0 , mas é fácil simular aleatoriedade

1ª Lei da Geografia de Tobler

Tudo depende de tudo, mas coisas mais próximas dependem mais.

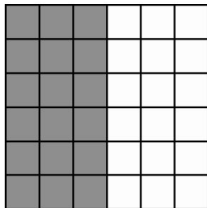
- À medida que a distância aumenta, a autocorrelação tende a desaparecer
- Se H_0 for rejeitada, ou seja, se houver padrão ou autocorrelação espacial, há duas alternativas:
 - a autocorrelação pode ser positiva (*clusters* espaciais): valores parecidos próximos ocorrem mais frequentemente do que sob H_0 (e há menos variabilidade)
 - a autocorrelação pode ser negativa (*outliers* espaciais): valores diferentes próximos ocorrem mais frequentemente do que sob H_0 (e há mais variabilidade)

Autocorrelação positiva e negativa

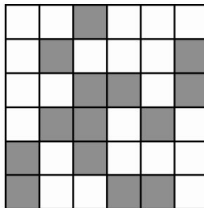
- Atenção! Diferentemente de séries temporais e análise multivariada, autocorrelação positiva e negativa não dão ideias opostas

Autocorrelação positiva e negativa

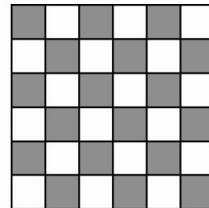
- Atenção! Diferentemente de séries temporais e análise multivariada, autocorrelação positiva e negativa não dão ideias opostas



autocorrelação espacial
positiva
(*clusters* espaciais)

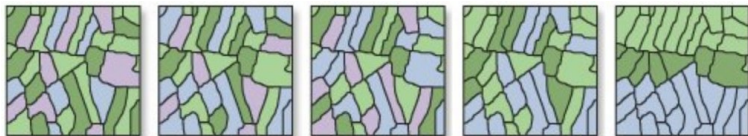


aleatoriedade espacial
(homogeneidade)

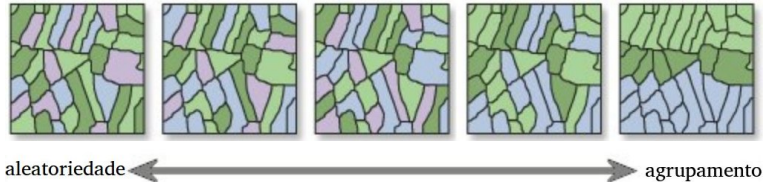


autocorrelação espacial
negativa
(*outliers* espaciais)

Aleatoriedade × agrupamento

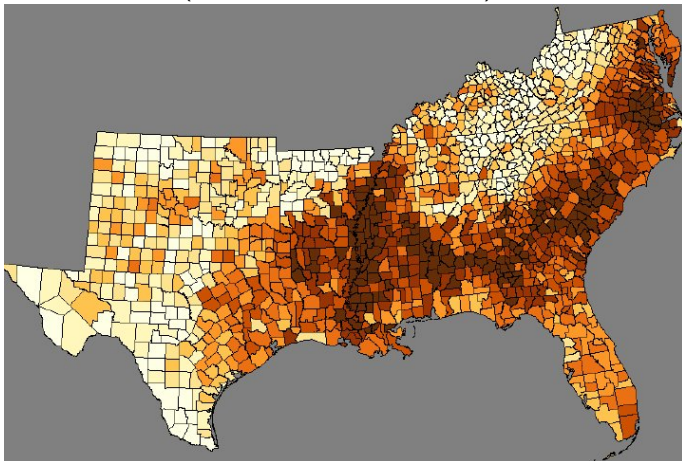


Aleatoriedade × agrupamento



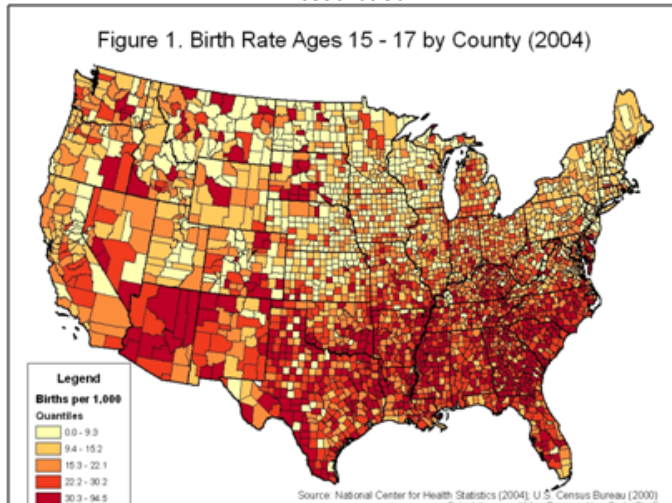
Aleatoriedade × agrupamento

A distribuição espacial de negros no sul dos EUA parece ser aleatória
(dados do censo de 2000)?



Aleatoriedade × agrupamento

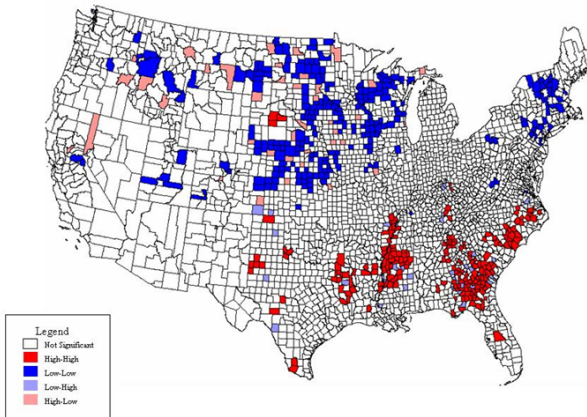
E esse caso?



Por curiosidade

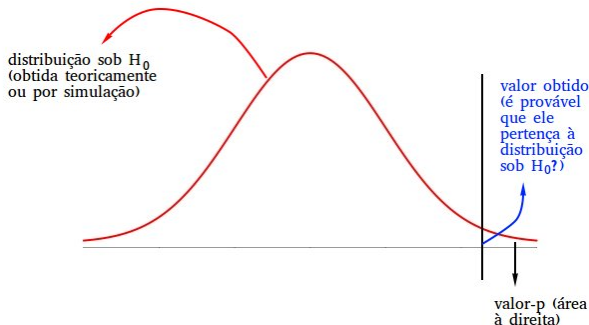
I de Moran global foi significativo e o mapa LISA obtido foi esse

Figure 3. LISA Cluster Map of Birth Rates Ages 15 – 17 Using Queen's 1 Spatial Weights Matrix



Estatísticas de teste

- Assumindo aleatoriedade espacial, qual a probabilidade de um determinado valor ter ocorrido?
- Quando a probabilidade é baixa (valor- p pequeno), H_0 é rejeitada



Estatísticas para medir a autocorrelação

- Devem capturar a similaridade de valor e de localização
- Similaridade de valor: y em i (sua localização) e em j (seus vizinhos)

As medidas de similaridade podem envolver três tipos de cálculo:

- produto cruzado: $y_i \cdot y_j$ (I de Moran é desse tipo)
- diferenças ao quadrado: $(y_i - y_j)^2$
- módulo das diferenças: $|y_i - y_j|$