

# Análise Espacial de Dados

Aula teórica 1  
Conceitos de AED  
Main Concepts

23 Fevereiro de 2022

1

## Sumário /Summary

1. Os SIG's e a Análise Espacial	1. GIS and Spatial Analysis
2. Componentes da Análise Espacial	2. Components of Spatial Analysis
2.1. Visualização	2.1. Visualization
2.2. Análise Exploratória de dados Espaciais	2.2. Exploratory Spatial Data Analysis
2.3. Modelação	2.3. Modeling
3. Conceitos Importantes	3. Important Concepts
3.1. Dependência Espacial	3.1. Space Dependence
3.2. Autocorrelação Espacial	3.2. Space Autocorrelation
3.3. Cluster espacial	3.3. Space Cluster
4. Tipos de dados a analisar e exemplos práticos	4. Types of data to analyze and practical examples
4.1. Pontos: análise da distribuição de um padrão pontual	4.1. Points: analysis of the distribution of a point pattern
4.2. Áreas: análise da autocorrelação espacial	4.2. Areas: analysis of spatial autocorrelation
4.3. Superfícies: Métodos de interpolação espacial	4.3. Surfaces: Spatial interpolation methods
4.4. Redes: Análise de caminhos ótimos e áreas de influência	4.4. Networks: Analysis of optimal paths and areas of influence

2

# Dos dados à Informação

From Data to Information

- **Dos dados à Informação**
  - Conjunto de dados organizados de modo a constituírem informação válida.
  - Transformações, manipulações, e aplicação de métodos analíticos aos dados espaciais (geográficos)

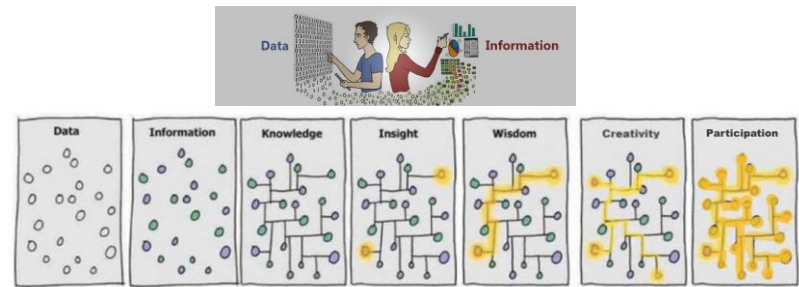


Image Source: <http://orangedotcom.nl/gauc-2016/> and John Hagel's fb page.

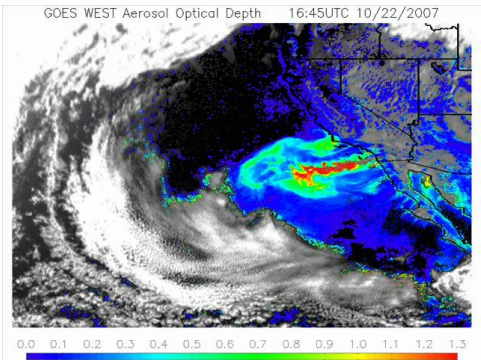
- **From data to information**

Set of data organized to constitute valid information.  
Transformations, manipulations, and application of analytical methods to spatial (geographic) data

# O que estuda a Análise Espacial?

What Spatial Analysis Studies?

- **Análise Espacial estuda a variabilidade espacial de um fenómeno de natureza geográfica**
  - Análise da variabilidade espacial de um fenómeno com a alteração da sua localização;
  - Utiliza medidas de dispersão para avaliar a localização de aglomerados, correlações e autocorrelações espaciais;
  - Procura modelos que expliquem a razão de determinada organização espacial;
  - Foco na questão: **ONDE?**
- **Analysis of spatial variability**
  - Analysis of the spatial variability of a phenomenon with its location;
  - Spatial dispersion measures, clusters, correlation and autocorrelation;
  - Focus on the question: **WHERE?**



# Objectivo da análise espacial ?

Spatial analysis goals?

- **Objectivo da AED**
  - **Compreender a distribuição** de dados com origem nos fenómenos ocorridos no espaço;
  - **Medir as propriedades e as relações entre os objetos** espaciais tendo em conta a sua localização;
  - **Criar modelos que expliquem** a organização espacial de um fenómeno

Para além da percepção visual da distribuição espacial do problema é importante:

- detectar padrões ou anomalias nos dados
- encontrar respostas às questões colocadas
- testar ou confirmar teorias/hipóteses
- gerar novas teorias e generalizações

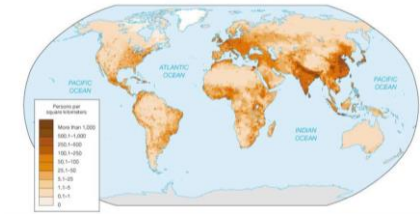
Objective:

- Understand the distribution of data from phenomena occurring in space;
- Measure the properties and relationships between phenomena taking into account their spatial location;

In addition to the visual perception of the spatial distribution of the problem, it is important to:

- detect patterns or anomalies in the data
- find answers to the questions asked
- test or confirm theories / hypotheses
- generate new theories and generalizations

## Exemplo Population Density & Distribution

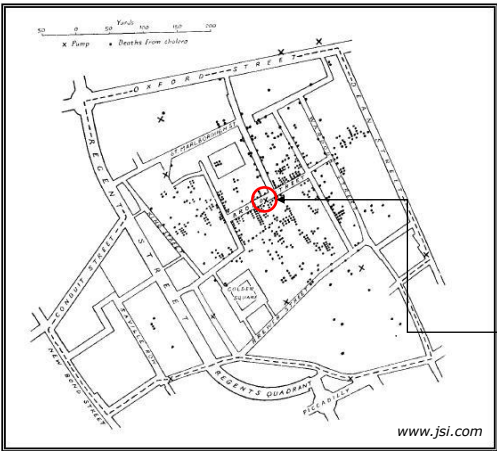


5

# Exemplo de aplicação da AED

- John Snow, 1854
- via polluted water, not air
- “John Snow’s pump”

## 1854 (Londres) : Mapas das áreas contaminadas por cólera



Mapa de Londres com óbitos por cólera identificados por localização do óbito (pontos) e poços de água (cruzes).

Map of London with cholera deaths identified by death location (dots) and water wells (crosses).

Confirmou-se que o ponto de captação de água desta bomba, a jusante da cidade, era um local de concentração de dejectos de pacientes com cólera.

It was confirmed that the water collection point of this pump, downstream of the city, was a concentration site of cholera patients' waste.

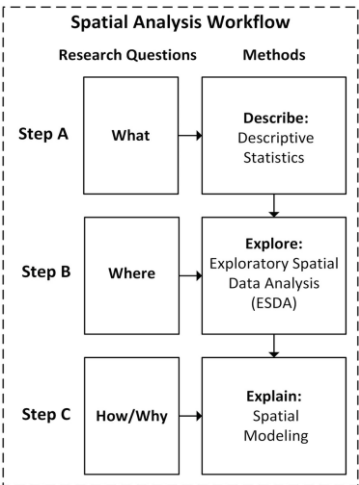
6

# Componentes da análise espacial/

*Components of Spatial Data Analysis*

- 1. **Visualização (Descritivo)**
  - **Identifica e localiza** padrões espaciais de interesse.
- 2. **Análise exploratória de dados (Exploratório)**
  - **Quantifica, com indicadores estatísticos,** padrões espaciais de interesse.
- 3. **Modelação (Explicativo)**
  - **Explica** padrões espaciais de interesse.

- 1. Visualization: Identifies and locates spatial patterns of interest.
- 2. Exploratory data analysis: It quantifies, with statistical indicators, spatial patterns of interest.
- 3. Modeling: Explain spatial patterns of interest.

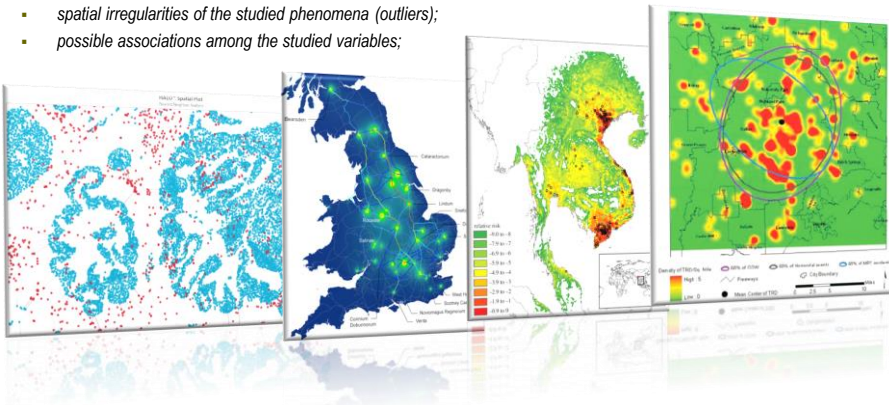


Source:  
[https://static.cambridge.org/binary/version/id/urn:cambridge.org:id/binary:20200421144010724-0711:9781108614528:49898fg1\\_1.png?pub-status=live](https://static.cambridge.org/binary/version/id/urn:cambridge.org:id/binary:20200421144010724-0711:9781108614528:49898fg1_1.png?pub-status=live)

# 1. Visualização

*Visualization*

- representação espacial dos dados (devidamente georreferenciados) para evidenciar:
  - (ir)regularidades espaciais dos fenômenos estudados (outliers);
  - eventuais associações entre as variáveis estudadas;
- Spatial representation of the georeferenced data to show:
  - spatial irregularities of the studied phenomena (outliers);
  - possible associations among the studied variables;

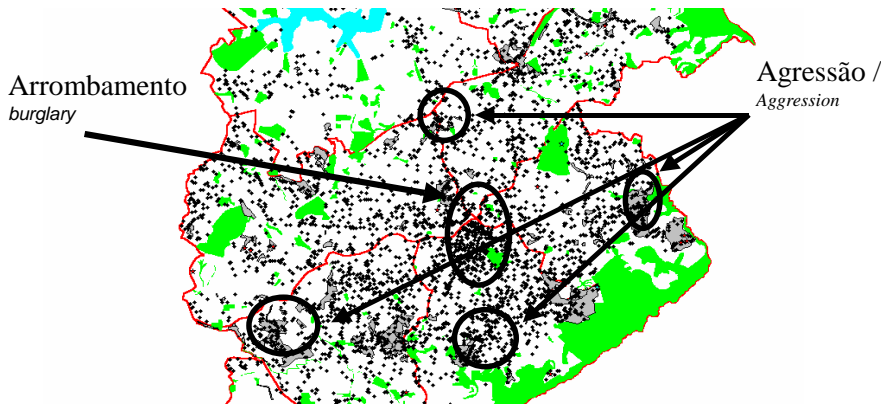


# Visualização de padrões de pontos

Visualization point pattern

## Crimes na cidade de Belo Horizonte

Crimes in the city of Belo Horizonte



Source: Renato Assunção (UFMG/Brasil)

9

# Visualização de padrões de áreas

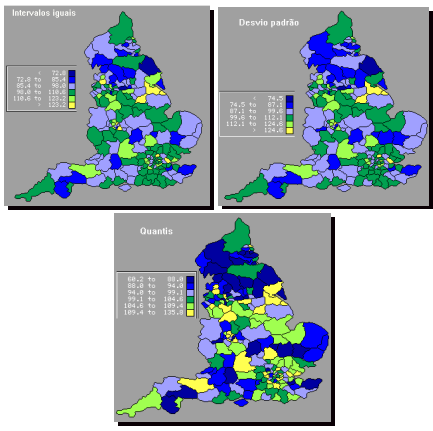
Visualization areal pattern

- Agrupamento por:
  - Intervalos Iguais
  - Quantis
  - Desvio Padrão

- Atenção!
  - Mapas temáticos com cores erradas podem conduzir a interpretações erradas

- Group by:
- Equal Ranges
  - Quantiles
  - Standard deviation

- Attention!
- Thematic maps with the wrong colors can lead to misinterpretations



Cancro da mama em Inglaterra  
Breast Cancer in England  
(1985-1989)

Source: Bailey and Gattrel, 1995

10

## 2. Análise exploratória de dados (AED) /

### Exploratory data Analysis

- A AED tem por objetivo efetuar uma prospecção sistemática e rigorosa de eventuais **padrões espaciais**, bem como detetar **associações espaciais** entre as variáveis que caracterizam os fenómenos estudados.
- “A exploração de dados funciona como um detetive numérico, agrupando e peneirando através dos dados, como se estivesse a produzir prova para o procedimento judicial da análise confirmatória” (K. Jones, 1984)

#### Exploratory data Analysis

The purpose of the AED is to conduct a systematic and rigorous exploration of possible spatial patterns, as well as to detect and verify spatial associations between the variables that characterize the studied phenomena.



11

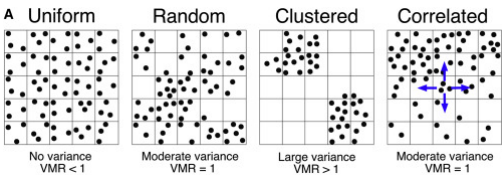
## 2. Análise exploratória de dados (AED) /

### Exploratory data Analysis

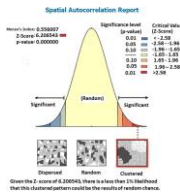
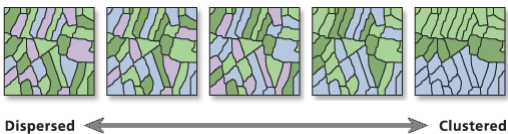
#### Como quantificamos as associações espaciais?

Através de indicadores específicos de associação espacial, como por exemplo:

- Indicadores de associação espacial (Variance to mean Ratio, VMR)



- Indicadores de Autocorrelação Espacial



12

# Conceito de dependência espacial

*Spatial dependence*

❑ Lei de **Waldo Tobler** (*Waldo Tobler Law*)

"Todas as coisas são parecidas, mas as coisas mais próximas parecem-se mais do que coisas mais distantes."

*"All things are alike, but things that are nearer appear to be more than things farther apart."*

❑ Segundo **Noel**

"a dependência (espacial) está presente em todas as direcções e fica mais fraca à medida que aumenta a dispersão na localização dos dados."

*"(Spatial) dependency is present in all directions and weaker as the dispersion in the data location increases."*

❑ **Generalizando / Generalizing**

A maioria das ocorrências, sejam naturais ou sociais, apresentam tipicamente uma relação entre elas que depende da distância.

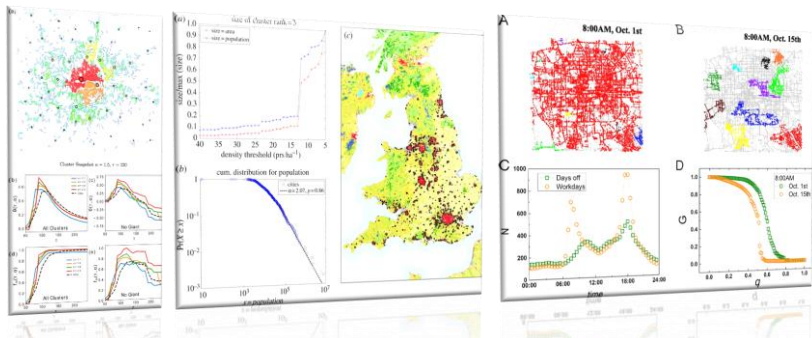
*Most occurrences, whether natural or social, have a relationship that depends on distance.*

13

## 3. Modelação

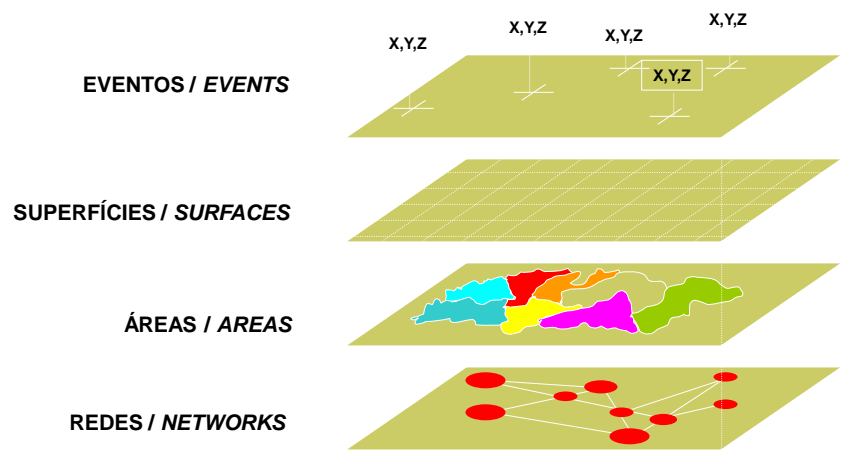
❑ A modelação em análise espacial consiste na aplicação da estatística (e.g. Testes de Hipoteses) para validar os índices usados na fase da análise exploratória dos dados e encontrar os modelos que melhor explicam a distribuição espacial do fenómeno em estudo.

❑ *Modeling in spatial analysis consists of the application of statistics (e.g. Hypothesis Tests) to validate the indices used in the exploratory analysis phase of the data.*



14

## 4. Geometrias usadas em análise espacial



15

## 4. Geometrias usadas em análise espacial

- ❑ **Eventos ou padrões pontuais:** fenômenos expressos através de ocorrências identificadas por pontos localizados no espaço, denominados processos pontuais (Exemplos: localização de crimes, ocorrência de doenças, localização de espécies vegetais);
- ❑ **Áreas** (com contagens ou taxas agregadas): trata-se de dados associados a levantamentos populacionais (exemplo: census, estatísticas de saúde, cadastro de imóveis) e que originalmente se referem-se a indivíduos localizados em pontos específicos do espaço. **A agregação em unidades de análise (polígonos) efectua-se ou para simplificar a análise e a representação dos dados ou para efeitos de confidencialidade;**
- ❑ **Superfícies contínuas:** estimadas a partir de um conjunto de amostras de campo, que podem ser regulares ou irregularmente distribuídas (Exemplo: mapas geológicos, mapas topográficos, mapas ecológicos, etc..)
- ❑ *Events or punctual patterns: phenomena expressed through occurrences identified by points located in space, called punctual processes (Examples: location of crimes, occurrence of diseases, location of plant species);*
- ❑ *Areas (with aggregated counts or rates): this is data associated with population surveys (example: census, health statistics, land registry) and which originally refer to individuals located at specific points in space. Aggregation into units of analysis (polygons) is carried out either to simplify the analysis and representation of data or for confidentiality purposes;*
- ❑ *Continuous surfaces: estimated from a set of field samples, which can be regularly or irregularly distributed (Example: geological maps, topographical maps, ecological maps, etc.)*

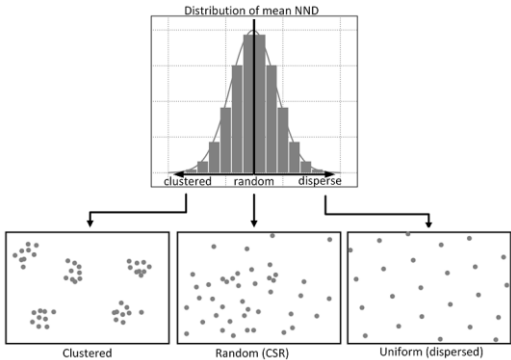
16



# 4.1. Análise de padrões de pontos

- O objecto de interesse é a localização espacial dos eventos em estudo. Pretende-se estudar a distribuição espacial destes pontos, testando hipóteses sobre o padrão observado: se é aleatório, se apresenta aglomerados, ou se é regularmente distribuído.

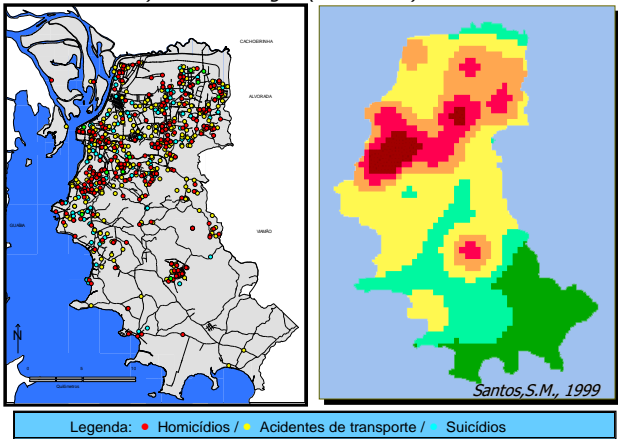
The object of interest is the spatial location of the events being studied. We intend to study the spatial distribution of these points, testing hypotheses about the pattern observed: if it is random, presents clusters, or is regularly distributed.



17

# 4.1. Análise de padrões de pontos (exemplo)

- Análise de padrões pontuais para o caso de mortalidade (homicídios/acidentes de transporte) em Porto Alegre (dados de 1996).
- Analysis of specific patterns for the case of mortality (homicides / transport accidents) in Porto Alegre (1996 data).



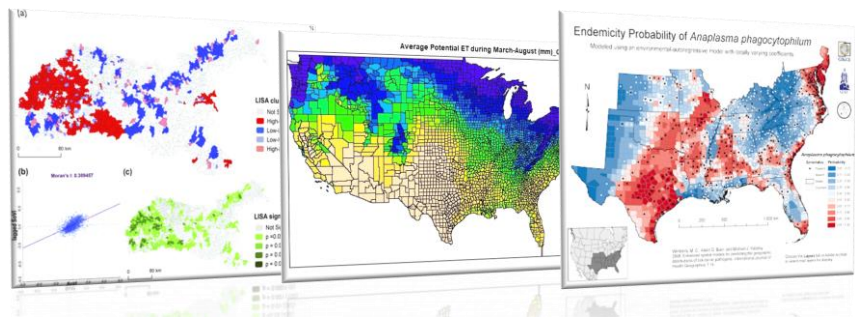
Superfície interpolada que mostra o padrão de distribuição de pontos com uma forte concentração no centro da cidade e decrescendo em direcção aos bairros mais afastados.

Interpolated surface showing the pattern of dot distribution with a strong concentration in the center of the city and decreasing towards the more distant neighborhoods.

18

## 4.2. Análise de áreas

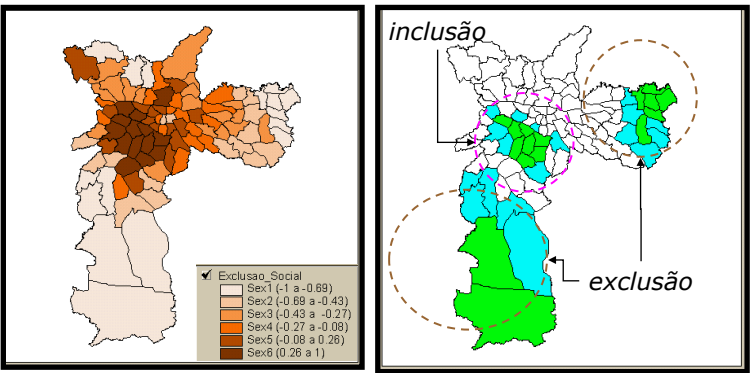
- Derivam geralmente de levantamentos populacionais. As áreas são delimitadas por polígonos fechados onde se supõe (não é garantido) existir homogeneidade interna, ou seja, mudanças importantes ocorrem apenas nos limites.
- They are generally derived from population surveys. The areas are delimited by closed polygons where it is assumed (not guaranteed) to exist internal homogeneity, that is, important changes occur only in the limits.*



19

## Análise de áreas (exemplo)

- Mapa de exclusão/Inclusão Social de São Paulo (1991) e agrupamentos de exclusão social (Zonas Leste e Sul) e inclusão social (centro)



Map of exclusion / social inclusion of São Paulo (1991) and groups of social exclusion (East and South zones) and social inclusion (center)

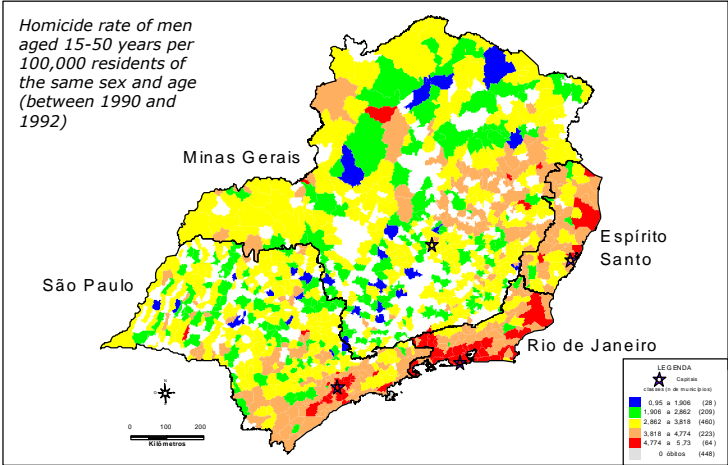
Mapa temático

Agrupamentos ou "Clusters" de exclusão/inclusão social em São Paulo

20

## Análise de áreas (exemplo)

**Taxa de homicídio de homens de idades compreendidas entre 15-50 anos por cada 100.000 residentes do mesmo sexo e idade (entre 1990 e 1992)**

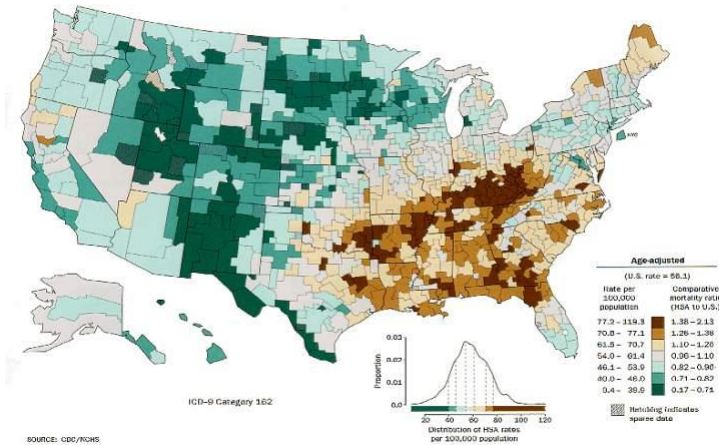


21

## Análise de áreas (exemplo)

**Mortes por cancro de pulmão em homens nos EUA**

*Deaths from lung cancer in men in the US*



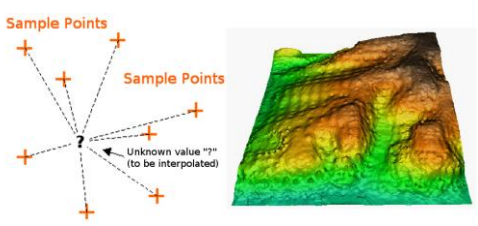
22

### 4.3. Análise de superfícies

- O objectivo é reconstruir a superfície da qual se retirou e mediu as amostras, para o efeito é necessário construir um modelo de dependência espacial (**modelo inferencial**) que permite construir a superfície através de métodos de interpolação.
- O **modelo inferencial** tem por objectivo **quantificar a dependência espacial entre os valores das amostras**. Para gerar este modelo utilizam-se técnicas de **geoestatística**.

The objective is to reconstruct the surface from a set of measured samples. It is necessary to construct a spatial dependence model (inferential model) that allows constructing the surface through interpolation methods.

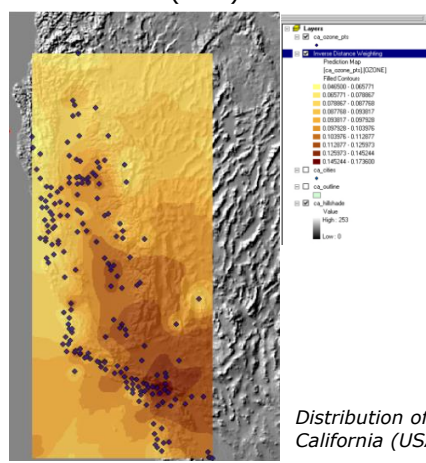
The purpose of the inferential model is to quantify the spatial dependence of sample values. To generate this model geostatistics techniques are used.



23

### 4.3. Análise de superfícies (exemplo)

- Distribuição de amostras de concentração de ozono na Califórnia (EUA)



1. Localizar espacialmente as amostras;
2. Estimar o campo de concentração de ozono através de um modelo de dependência espacial (modelo inferencial) para produzir um mapa (superfície - continuidade espacial);
3. Estimar a concentração de ozono em localizações específicas onde não existem amostras (e.g., Los Angeles)

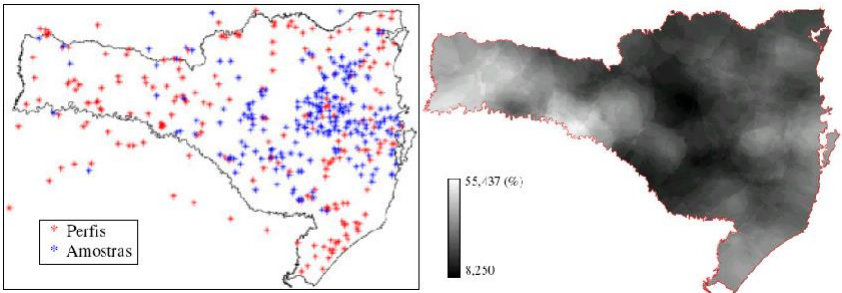
**Geoestatística**

Distribution of ozone concentration samples in California (USA)

24

### 4.3. Análise de superfícies (exemplo)

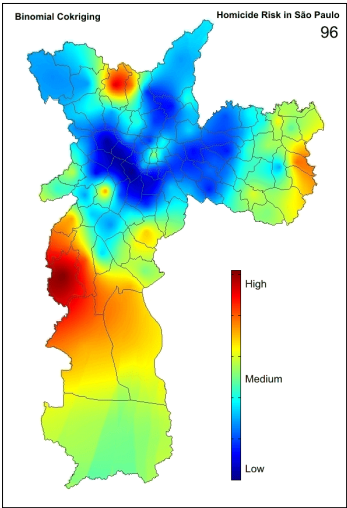
- Distribuição de perfis e amostras de solo para o estado de Santa Catarina e áreas próximas, e mapa de distribuição espacial da variável *saturação por bases*.



Distribution of profiles and soil samples for the state of Santa Catarina and nearby areas, and spatial distribution map of the base saturation variable.

25

### 4.3. Análise de superfícies (exemplo)



Risco de homicídio na cidade de São Paulo – 1996

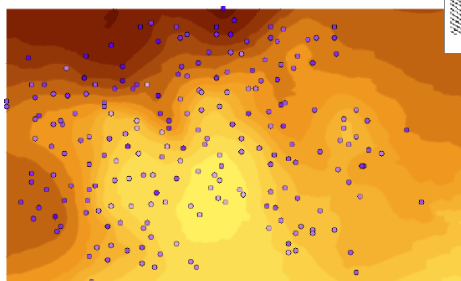
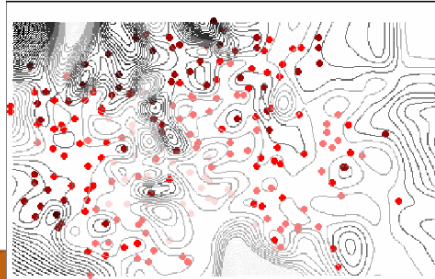
Risk of homicide in the city of São Paulo - 1996

26

## 4.3. Análise de superfícies (exemplo)

- Preço da habitação na cidade de Baltimore

*Price of housing in the city of Baltimore*



27

## 4.4. Análise de redes

### *Network Analysis*

- Procura de caminhos ótimos entre vários destinos sobre uma rede (grafo);
  - Procura de caminho mais rápido/curto sobre uma rede viária;
  - Requer uma representação vectorial;
- Procurar o caminho de menor custo:
  - Entre uma origem e um destino;
  - Para minimizar o custo total;
  - O custo pode combinar várias variáveis: construção, impacto ambiental, aquisição de terras (e.g. expropriações), custo operacional, etc...
  - Usado para construir auto-estradas, linhas energéticas, redes de esgotos, etc...
  - Geralme

*Search for optimal paths between several destinations over a network (graph);*

- Search for the fastest / shortest route on a road network;
- Requires a vector representation;

*Look for the lowest cost path:*

- Between an origin and a destination;
- To minimize total cost;
- The cost can combine several variables: construction, environmental impact, land acquisition (e.g. expropriations), operational cost, etc ...
- Used to build motorways, energy lines, networks of esgotos, etc ...
- Usually requires a raster representation; nte requer uma representação raster;

28

