School of Geography and Geology
McMaster University

Advanced Topics in Spatial Statistics

# Spatial Data Analysis

## This session:

- Spatial Data
  - Classes of problems in spatial data analysis
  - Visualization of Spatial Data
  - Exploration of Spatial Data
  - Modeling of Spatial Data
  - Spatial Correlation Structure

## Spatial vs. Aspatial Data

- How are spatial and aspatial data different?



## Aspatial Data: Examples

- Salary of professional baseball players
- Concrete's strength tests
- Development of new drugs

## Spatial Data: Examples

- Location of bank branches
- Groundwater pressure measurements
- Population density



## Spatial Data

- Point patterns
- Spatially continuous data
- Area data

- Types of data found in the social and natural sciences

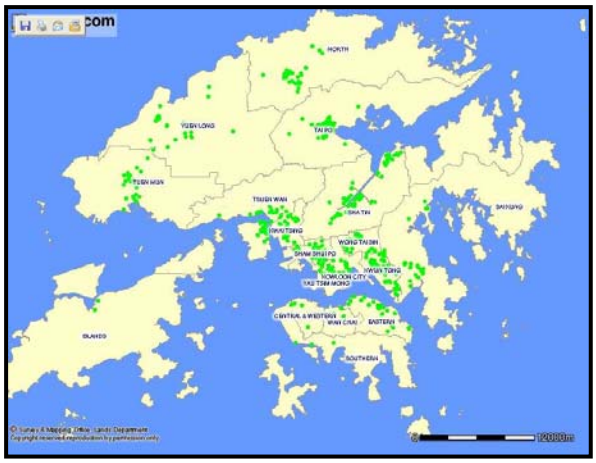
## Point patterns

- Discrete events
- Patterns of human activity
  - Settlements
  - Industry
- Ecology
  - Plant species
  - ...
- Epidemiology
  - Cases
  - ...



## Point Patterns

- SARS–Suspected cases in Hong Kong



## Point Patterns

- Location of events
- Comparing two different patterns
- Space-time clusters

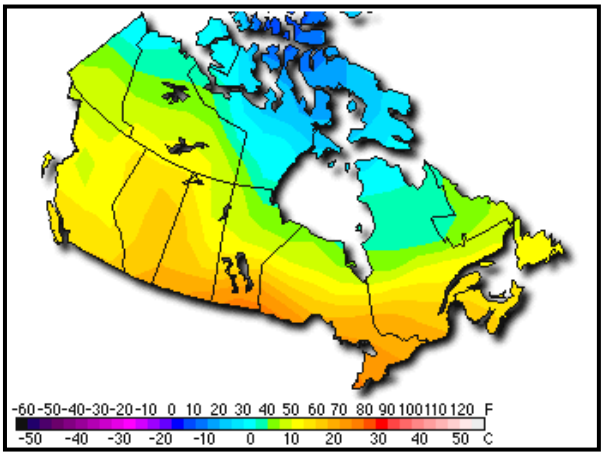


## Spatially Continuous Data

- Process can be measured anywhere
- Natural and environmental sciences
  - Temperature
  - Rain maps
  - Mineral resources

## Spatially Continuous Data

- Temperature map - Canada






## Spatially Continuous Data

- Spatial prediction (Interpolation)

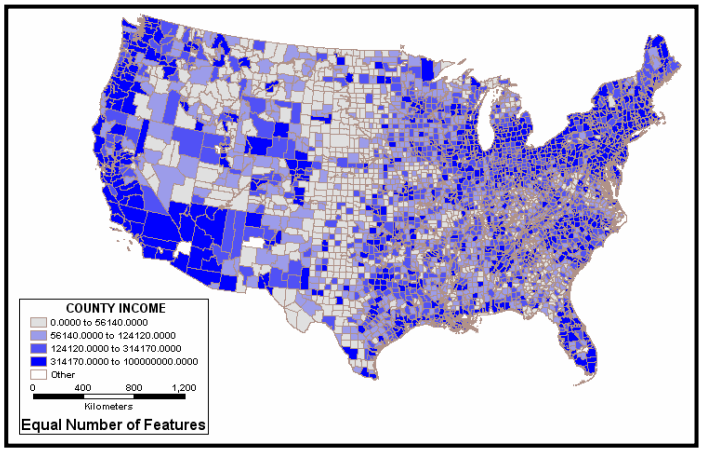
## Area Data

- Discrete spatial units
- Social sciences
  - Income
  - Population
  - Ethnicity
  - ...



## Area Data

- County income in the US



## Spatial Data Analysis

- Visualizing spatial data
- Exploring spatial data
- Modeling spatial data

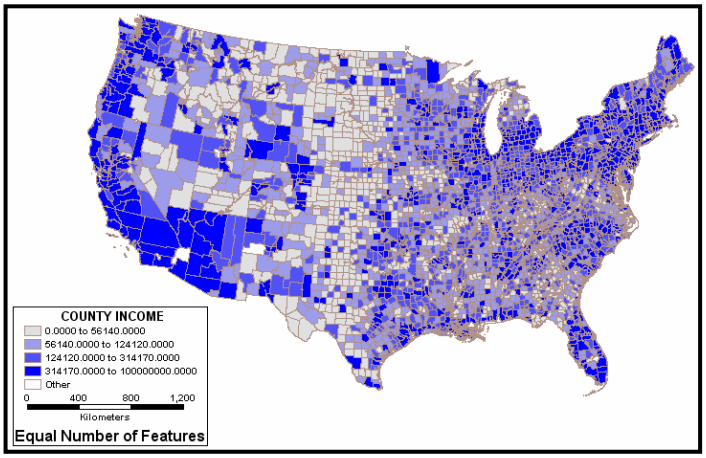


## Visualizing Spatial Data

- Dot maps (Point patterns)
- Contours (Spatially continuous data)
- Choropleth maps (Area data)
- Interactive mapping (GIS)
- ...

## Visualizing Spatial Data: Choropleth Map

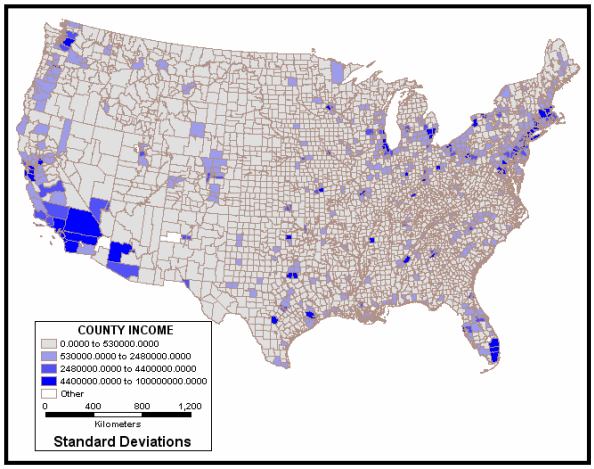
- County income in the US





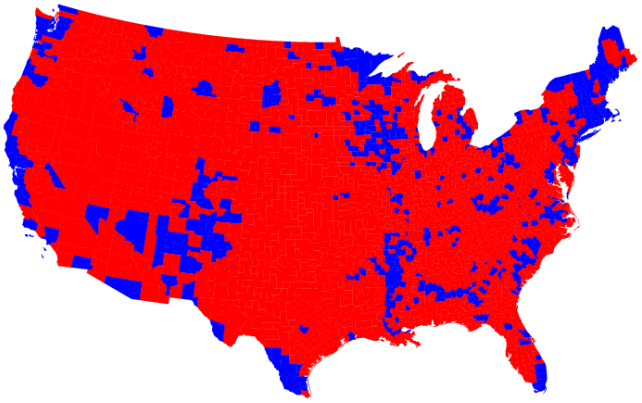

## Visualizing Spatial Data: Choropleth Map

- County income in the US



## Visualizing Spatial Data: Choropleth Map

- US 2004 Election results

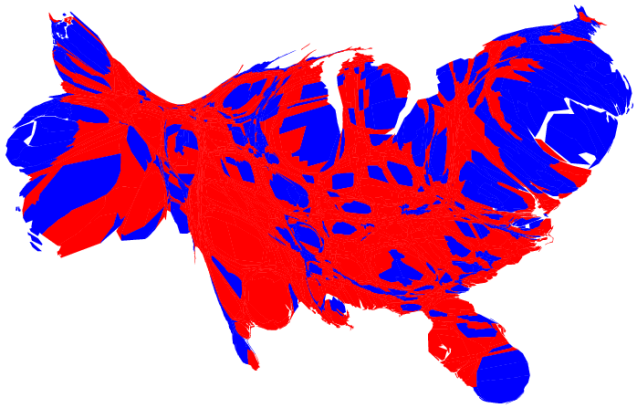






## Visualizing Spatial Data: Population weighted cartogram

- US 2004 Election results



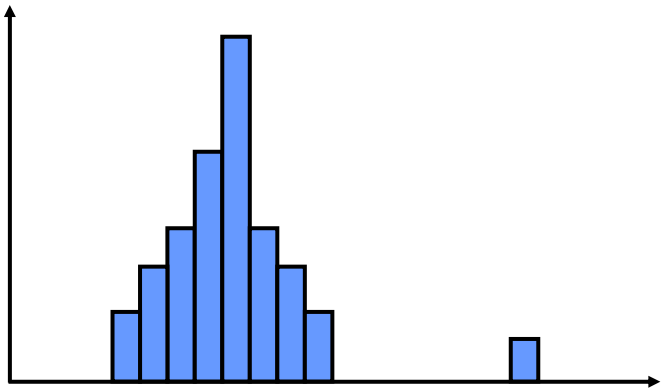
## Exploring Spatial Data

- Good descriptions of the data
- Few a priori assumptions
- Fact finding, development of working hypotheses, etc.
- **Robust**, **Sufficient** methods
- Typically univariate or bivariate methods (involving one or two variables)



## Exploring Spatial Data

- What is a robust method?



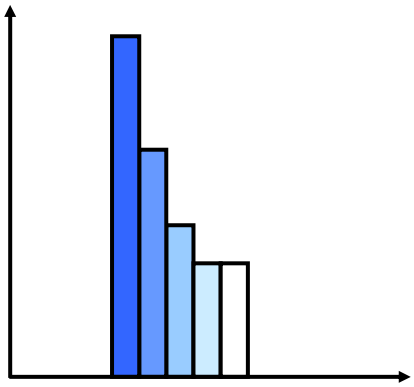
## Exploring Spatial Data

- Robust method are resistant to atypical observations (outliers)



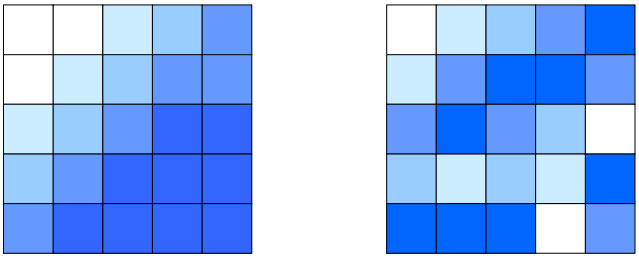
## Exploring Spatial Data

- Sufficiency criterion



## Exploring Spatial Data

- Sufficiency criterion





## Exploring Spatial Data

- A combination of conventional techniques + maps
- Specialized techniques
  - Statistics for detecting spatial patterns

## Modeling Spatial Data

- Explanation
- Relationships
- Stronger assumptions
- Testing of hypotheses
- **Robust**, **Sufficient** methods
- Typically multivariate (involving two or more variables)

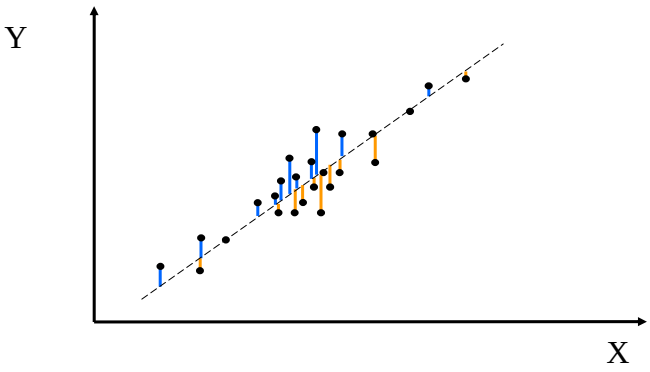


## Modeling Spatial Data

- Specialized techniques
  - Simulation
  - Regression analysis

## Modeling Spatial Data

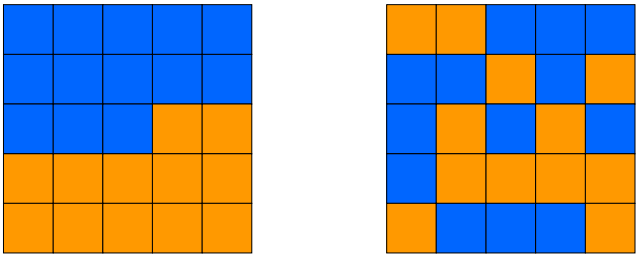
- Sufficiency criterion





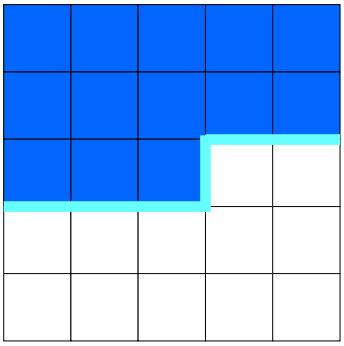

## Modeling Spatial Data

- Sufficiency criterion



## Why are there spatial patterns?

- Identical reaction to environmental factors





## Why are there spatial patterns?

- Identical reaction to environmental factors
  - If we know all the factors, conventional analysis is sometimes sufficient
  - Often we cannot observe all factors: specialized techniques are required

## Why are there spatial patterns?

- Spatial processes
  - Diffusion
  - Transfer and exchange
  - Interaction
  - Dispersal



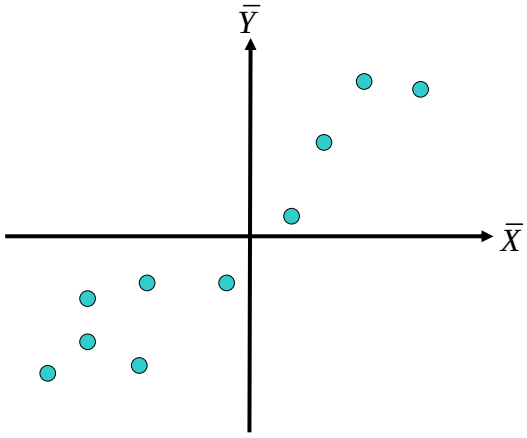
## Why are there spatial patterns?

- Covariance

$$C_{XY} = \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})$$

## Why are there spatial patterns?

- Covariance






## Why are there spatial patterns?

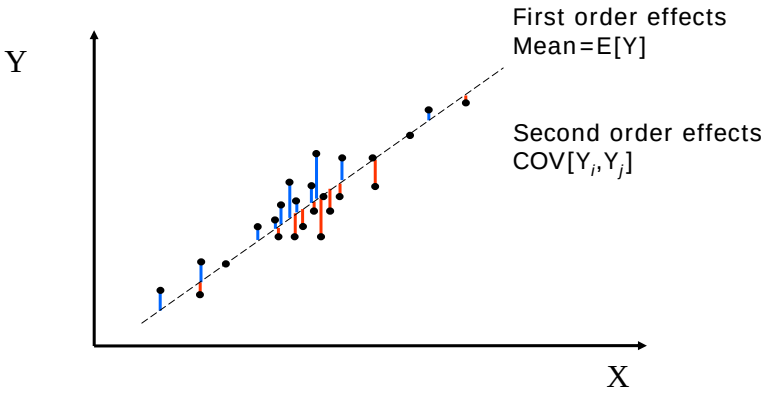
- Autocovariance

$$C_{XX} = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (X_i - \bar{X})(X_j - \bar{X})$$

- Autocorrelation

## Why are there spatial patterns?

- Spatial correlation structure






## Next

- Point pattern analysis I