Points to Cover

- Conceptualizing spatial patterns
- Spatial statistics
  - Definition
  - Typology
- Spatial statistics in ArcGIS
Conceptualizing Spatial Patterns

- Geography is concerned with the study of spatial patterns produced by physical and cultural processes
  - Description of the patterns
  - Analysis of the patterns

- The spatial patterns can be conceptualized as
  - Distribution of the features of phenomena across the study area
    - Distribution of cities in a region
  - Distribution of a certain characteristic (variables) associated with those features or phenomena
    - Distribution of average income across the cities
Spatial Statistics

- Spatial statistical analytical methods are tools used to describe and analyze spatial patterns
  - Allow ideas about spatial processes and patterns to be tested
  - Incorporate space (proximity, area, connectivity, and/or other spatial relationships) directly into their mathematics
  - Are based on the assumption of spatial autocorrelation
    - Tobler’s First Law of Geography
Spatial Statistics

- Types of quantitative spatial statistical methods
  - Exploratory and descriptive statistics
    - Describe distribution of the spatial features or phenomena
  - Inferential statistics
    - Compare sample pattern to a theoretical pattern of distribution of features or phenomena
  - Predictive statistics
    - Analyse relationships between features or phenomena
  - Prescriptive statistics
    - Modeling and optimization methods
Spatial Statistics in GIS

- Exploratory and descriptive spatial statistics
  - Simple summary statistics
  - Histograms
  - Measuring Geographic Distribution Toolset
    - Spatial measures of central tendency
    - Spatial measures of dispersion
Spatial Statistics in GIS

- Directional Distribution (Standard Deviational Ellipse)
  - Measures a trend for a set of points or areas by calculating the standard distance separately in the x- and y-directions

\[
SDE_y = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \bar{Y})^2}{n}}
\]

\[
SDE_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{X})^2}{n}}
\]

Where
- \(X_i\) and \(Y_i\) are coordinates of a feature \(l\)
- \(\bar{X}, \bar{Y}\) are coordinates of the mean centre for the set of features
- \(n\) is the number of features in the set

Source: ArcGIS 10 Help files
Spatial Statistics in GIS

- Directional Distribution (Standard Deviational Ellipse)
  - Can be calculated using
    - The locations of the features or
    - The locations influenced by an attribute value associated with the features
  - Application
    - Mapping the distributional trend for a set of crimes to explore a potential relationship between incidents and location of features
    - Assessing the spread of contaminant based on the contaminant concentrations collected at sample sites
    - Comparing patterns over time

Source: ArcGIS 10 Help files
Spatial Statistics in GIS

- **Inferential statistics**
  - Identifying geographic patterns is important for understanding how geographic phenomena behave.
  - **Analysing Patterns Toolset**
    - Based on the null hypothesis that the features (or the values associated with the features) exhibit a spatially random pattern.
    - Compute the probability that the null hypothesis is correct.
Spatial Statistics in GIS

- Average Nearest Neighbour Index
  - Is expressed as the ratio of the **observed mean distance** to the **expected mean distance** between the features
  - The expected distance is the average distance between neighbours in a hypothetical random distribution

Where
- \( D_o \) is the observed mean distance between a feature and its nearest neighbour
- \( D_E \) is the expected mean distance between the features given the random pattern
- \( A \) is the study area

\[
\begin{align*}
ANN &= \frac{\bar{D}_O}{\bar{D}_E} \\
\bar{D}_O &= \frac{\sum_{i=1}^{n} d_i}{n} \\
\bar{D}_E &= \frac{0.5}{\sqrt{n/A}}
\end{align*}
\]
Spatial Statistics in GIS

- Average Nearest Neighbour Index
  - Based on the assumption that the points are free to locate anywhere within the study area
    - There are no barriers
    - Features are located independently of one another
  - Interpretation

![Image showing different levels of clustering with ANN values: ANN > 1, ANN = 1, ANN < 1]
Spatial Statistics in GIS

- Average Nearest Neighbour Index

  **Application**
  
  - Quantify and compare the spatial distribution of a variety of plant or animal species within a fixed study area
  
  - Monitor changes over time in the patterns of spatial distribution of plant and animal species
  
  - Compare an observed distribution to a control distribution
Spatial Statistics in GIS

- Predictive Statistics
  - Beyond analyzing spatial patterns, GIS analysis can be used to examine or quantify relationships among features.
  - The **Modeling Spatial Relationships Toolset**
    - Construct spatial weights matrices
    - Model spatial relationships using regression analyses
Spatial Statistics in GIS

- Ordinary Least Squares Regression
  - Regression is used to evaluate relationships between two or more feature attributes
    - Helps to better understand processes active in an area
    - Helps to predict where something is likely to occur
    - Serves a start for exploring causal relationships between location and characteristics of features
  - OLS regression provides a global model of the variable or process
    - A good starting point for more sophisticated regression analyses
Spatial Statistics in GIS

- **Ordinary Least Squares Regression**
  - Regression is used to evaluate relationships between two or more feature attributes
    - Helps to better understand processes active in an area
    - Helps to predict where something is likely to occur
    - Serves as a start for exploring causal relationships between location and characteristics of features
  - OLS regression provides a global model of the variable or process
    - Creates a single regression model describing a process
    - A good starting point for more sophisticated regression analyses
Spatial Statistics in GIS

- Ordinary Least Squares Regression
  - Issues with applying OLS to spatial data
    - Spatial autocorrelation creates an over count type of bias
    - Nonstationary nature of many geographic processes
      - Regional variation
      - Processes may behave differently in different parts of the area
  - Geographically weighted regression (GWR)
    - Provides a local model of the variable or process by fitting a regression equation to every feature in the dataset
    - Deal effectively with the spatial autocorrelation issue