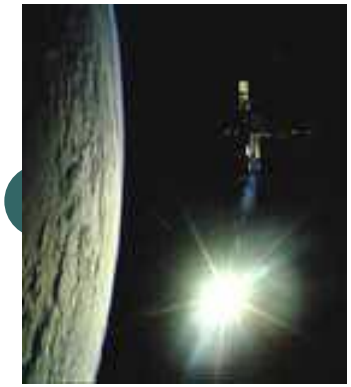


# Spatial Data Analysis: Intro to Spatial Statistical Concepts

Scott Bell
GIS Institute

# Spatial Stats rely on Spatial Data

- Traditional statistics are based on distributions of data along a single axis
- Spatial data by its nature exists on two axes (X and Y)
- I. E. the median in traditional statistics in the sum of all values divided by the number of observations
- Spatial mean is the X, Y coordinate result from calculating the means of X and Y

$$(\bar{x}, \bar{y}) = \left(\frac{\Sigma f_i x_i}{\Sigma f_i}, \frac{\Sigma f_i y_i}{\Sigma f_i}\right)$$

# Exploratory Spatial Data Analysis

- Used like descriptive statistics
- Potentially more options
- Related to Thematic Mapping and Geo-visualization
- Pattern identification/Hypothesis generation

# Traditional vs Spatial

- “Independence of observations” Assumption
- Spatial Statistics operate on data that are assumed to be spatially dependent
- Spatial statistics (Spatial autocorrelation(SA)) have been developed to account for SA so distribution theory can be applied

# Traditional vs Spatial

- "Replication" Assumption
  - Spatial (and other systems) are complex and hard to replicate
- Precise Data
- Samples drawn from hypothetical universe
  - Inability to replicate (due to size and complexity of system) usually means our sample spatial data is the universe
- Distribution under null can be obtained by creating an experiment (environment) in which the null is true
  - due to sample being universe it is virtually impossible to obtain the distribution under null hypothesis conditions

# Spatial Autocorrelation

- What is it?
- Uses of spatial autocorrelation
- Types of spatial dependence
  - Distance
  - K-nearest neighbors
  - Contiguity
    - Rooks, bishops, and Kings cases

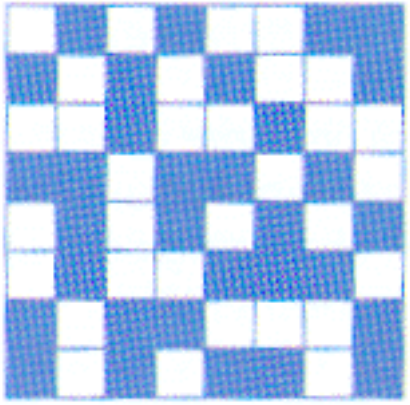
# Spatial Autocorrelation

- Deal simultaneously with similarities in the location (space) of objects and their (non-spatial) attributes. (Goodchild, et. al. 2001)
- Similar location/Similar attribute = high spatial autocorrelation
- Similar location/dissimilar attributes = negative spatial autocorrelation
- Attributes are independent of location = zero/low correlation

(A)

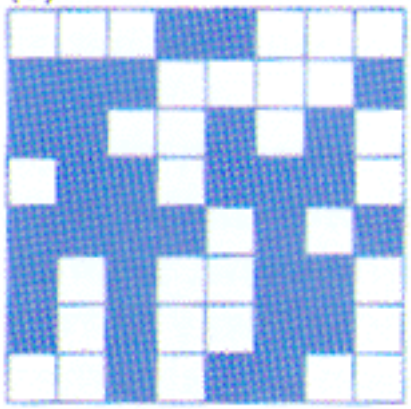
**Correlation= -1.00**

(B)



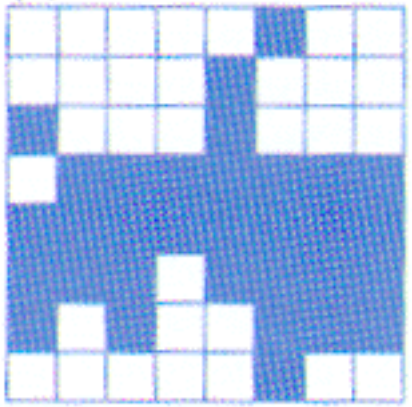
**Correlation= -.393**

(C)



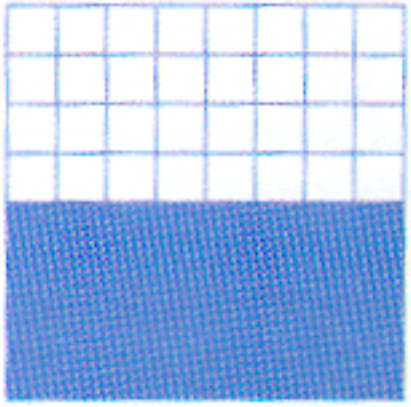
**Correlation= 0**

(D)



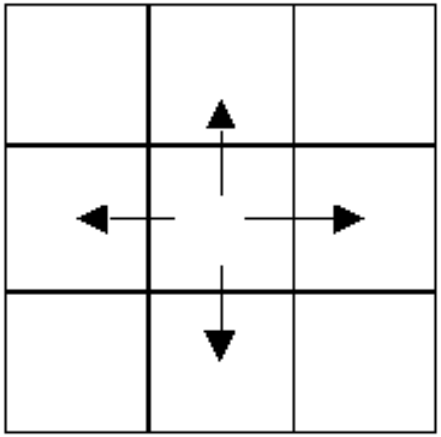
**Correlation= +.393**

(E)

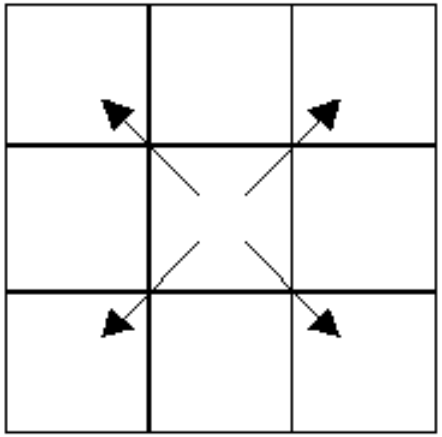


**Correlation= +.857**

Rooks Case



Bishops Case



Queen's (Kings) Case