

GIS APPLICATIONS

LECTURE 3

SPATIAL AUTOCORRELATION.

- First law of geography: “everything is related to everything else, but near things are more related than distant things”
- – Waldo Tobler
- Check who is sitting next to you!

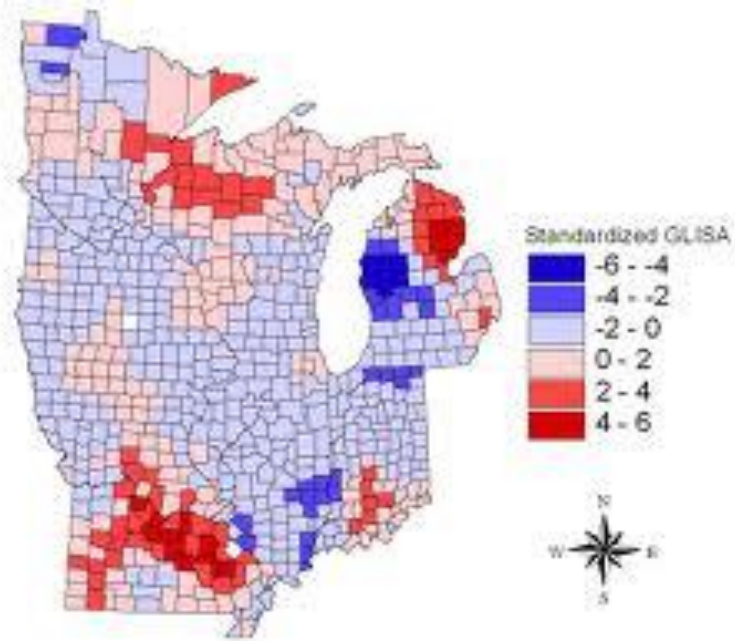


SPATIAL AUTOCORRELATION.

- Spatial autocorrelation (statistics)
- measure and analyze the degree of dependency among observations in a geographic space
- E.g existence of Au in quartz veins!



SPATIAL AUTOCORRELATION



SPATIAL AUTOCORRELATION

- Classic spatial autocorrelation statistics
- Moran's *I* and
- *Geary's C*.
- They require measuring a spatial weights matrix
- reflects the intensity of the geographic relationship between observations in a neighborhood



SPATIAL AUTOCORRELATION.

- Classic spatial autocorrelation statistics compare the **spatial weights** to the **covariance** relationship at pairs of locations.
- Spatial autocorrelation that is more positive than expected from random indicate the clustering of similar values across geographic space
- while significant negative spatial autocorrelation indicates that neighboring values are more dissimilar than expected by chance, suggesting a spatial pattern similar to a chess board.



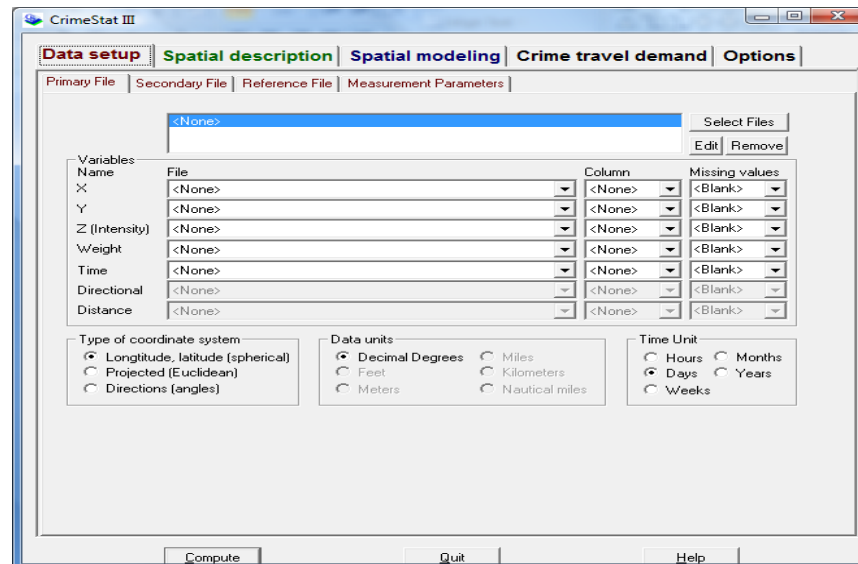
SPATIAL AUTOCORRELATION.

- Moran's I and Geary's C are global
- autocorrelation may vary significantly across geographic space- possibility of heterogeneity
- [Homogeneity and heterogeneity are concepts relating to the uniformity or lack thereof in a substance.]
- Moran's I Statistic



SPATIAL AUTOCORRELATION- *CRIMESTAT* SOFTWARE

- **CrimeStat**
- Windows-based
- spatial statistic software program
- 4 spatial and statistical analysis and is designed to interface with a GI System.



SPATIAL AUTOCORRELATION-MORAN'S I STATISTIC

- where N is the number of cases,
- X_i is the variable value at a particular location, i ,
 X_j is the variable value at another location
(where $i \neq j$),
- \bar{X} is the mean of the variable and W_{ij} is a weight applied to the comparison between location i and location j .

$$I = \frac{N \sum_i \sum_j w_{ij} (X_i - \bar{X})(X_j - \bar{X})}{(\sum_i \sum_j w_{ij}) \sum_i (X_i - \bar{X})^2}$$



SPATIAL AUTOCORRELATION-MORAN'S I STATISTIC

- W_{ij} is a distance-based weight which is the inverse distance between locations i and j ($1/d_{ij}$).
- The weighted Moran's I is similar to a correlation coefficient
- cross-products of values at different locations ,
- two at a time weighted by the inverse of the distance between the locations,
- with the variance of the variable



SPATIAL AUTOCORRELATION-EXAMPLE

- motor vehicle thefts in both Baltimore County, 1996 and Baltimore City by individual blocks.
- 14,853 vehicle theft locations
- 13,101 census blocks
- (*aggregating incident locations to zones*)
- *CrimeStat* calculated “I” as 0.012464 and the theoretical value of “I” as -0.000076.
- The test of significance using the normality assumption gave a Z-value of 125.13



SPATIAL AUTOCORRELATION-MORAN'S I

$$Z(I) = \frac{I - E(I)}{S_{E(I)}} = \frac{0.012464 - (-0.00076)}{0.000100} = 125.13 \text{ (} p \leq .001 \text{)}$$

- Motor thefts are highly and positively spatially autocorrelated



SPATIAL AUTOCORRELATION-MORAN'S I

- Blocks with many incidents tend to be located close to blocks which also have many incidents and, conversely,
- Blocks with few or no incidents tend to be located close to blocks which also have few or no incidents.



SPATIAL AUTOCORRELATION-GEARY'S C STATISTIC

- Similar to Moran's I

$$C = \frac{(N-1) [\sum_i \sum_j w_{ij} (X_i - X_j)^2]}{2(\sum_i \sum_j w_{ij}) \sum_i (X_i - \bar{X})^2}$$

- In this case, however, the interaction is not the cross-product of the deviations from the mean,
- but the deviations in intensities of each observation location with one another. (X_i and X_j)



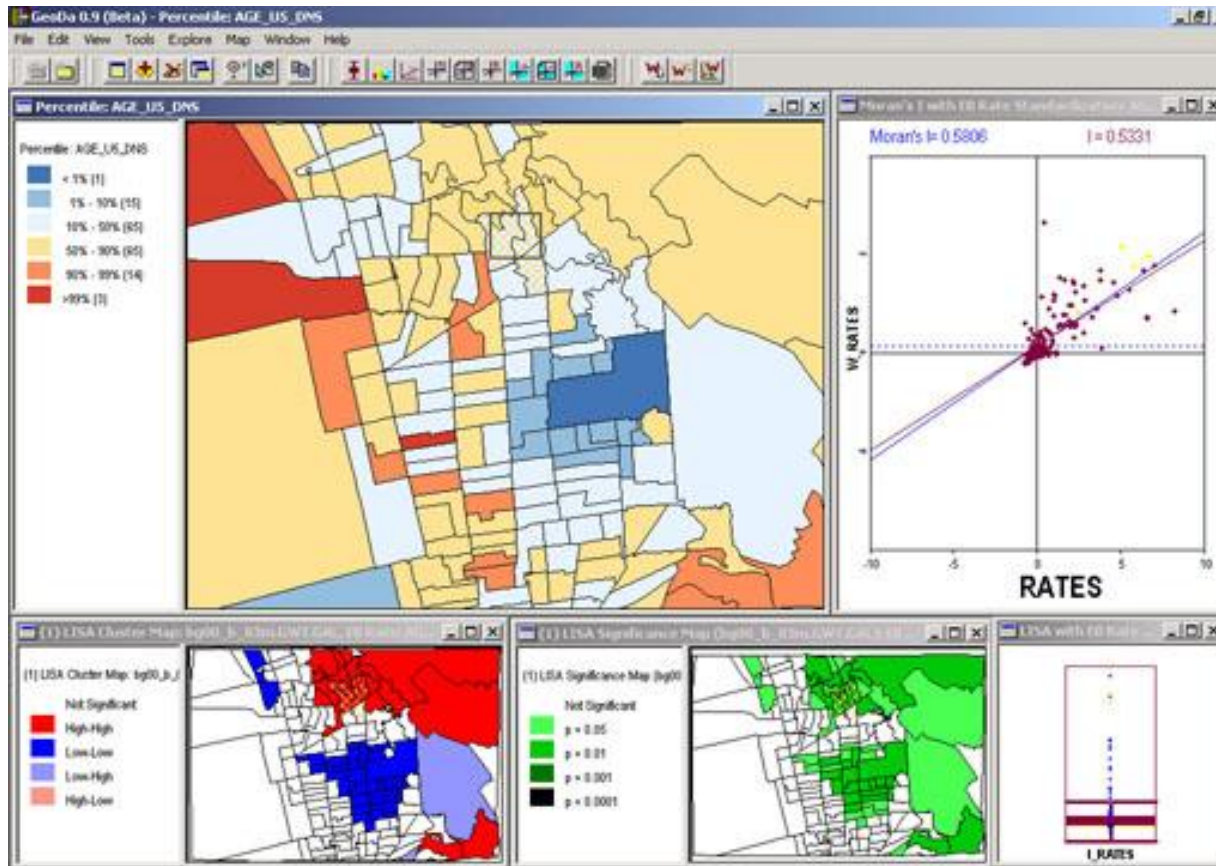
SPATIAL AUTOCORRELATION-GEARY'S C STATISTIC

- The values of C typically vary between 0 and 2
- Values less than 1 (i.e., between 0 and 1) typically indicate positive spatial autocorrelation
- $C < 1$
- while values greater than 1 indicate negative spatial autocorrelation.
- $C > 1$





EXPLORATORY SPATIAL DATA ANALYSIS



EDA

- **First Exploratory data analysis!**



EXPLORATORY DATA ANALYSIS

- The aim is of exploratory data analysis is to identify data properties for purposes of
- pattern detection in data,
- hypothesis formulation from data and
- some aspects of model assessment (e.g. Goodness of fit, identifying data effects on model fit).



EXPLORATORY DATA ANALYSIS

- EDA is based on the use of
- graphical methods, and
- visual methods
- and the use of numerical techniques that are statistically robust. i.e. not much affected by extreme values
- EDA emphasises on descriptive methods
- rather than formal hypothesis testing and the importance of staying close to the original data

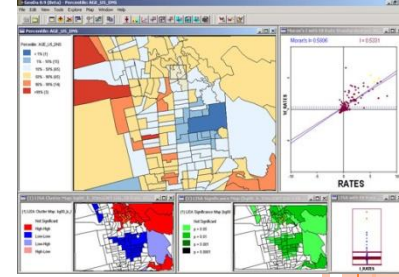


EXPLORATORY SPATIAL DATA ANALYSIS?

- The extension of EDA to detect **spatial** properties of data
- need for additional techniques –EDA+
- ?
 - detecting spatial patterns in data
 - formulating hypotheses based on the geography of the data
 - assessing spatial models.



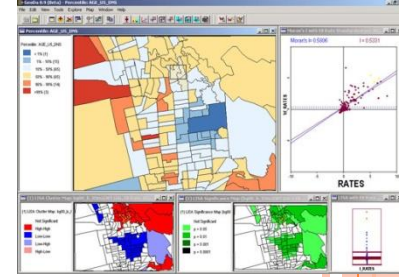
EXPLORATORY SPATIAL DATA ANALYSIS?



- In ESDA we should be able to link numerical and graphical procedures with the map
- We need to be able to answer the question, ‘where are those cases on the map?’
- With modern graphical interfaces this is often done by '**brushing**' - for example cases are identified by brushing the relevant part of a boxplot, and the related regions are identified on the map.



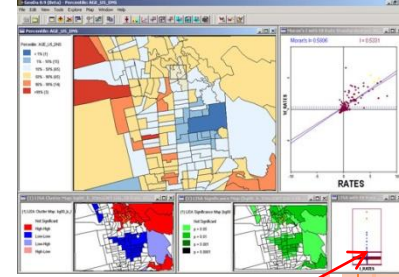
EXPLORATORY SPATIAL DATA ANALYSIS?



- **Data Model for ESDA**
- A set of data can be thought of as having
- general trends (e.g. average values, relationships) and
- local variations from those trends
- AKA- the **smooth** and **rough** properties of the data

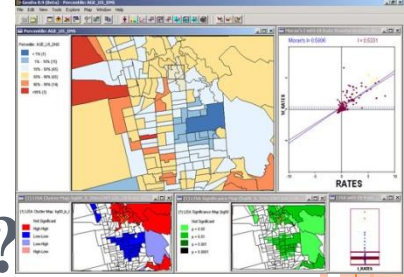


EXPLORATORY SPATIAL DATA ANALYSIS?



- Non-spatial properties
- Smooth
 - ‘Location’ (median)
 - Spread (interquartile range) and
 - Shape (histograms, boxplots)
- rough
 - Outliers e.g. values more than a certain distance above or below the upper or lower quartile of the distribution



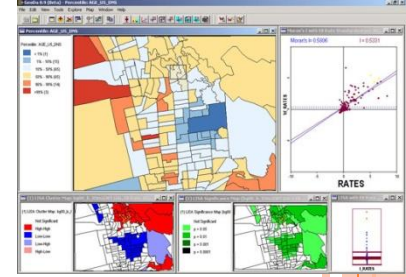


EXPLORATORY SPATIAL DATA ANALYSIS?

- Spatial properties
- smooth
 - spatial trends (or gradients)
 - spatial autocorrelation
- rough
 - difference between data value and smooth value.
- spatial outliers. These are individual attribute values that are very different in magnitude from their neighbouring values



EXPLORATORY SPATIAL DATA ANALYSIS?

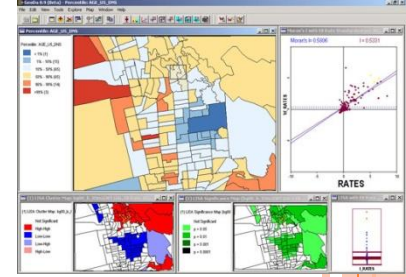


- **Data properties for two variables identified through ESDA**
- In the case of two variables, the scatterplot is used to visualise the relationship between them.
- The best fit line through the scatter plot identifies the **smooth** element of the relationship, and
- the residuals from the best fit line the **rough** element.
- An outlier is a data value more than a certain vertical distance from the best fit line.





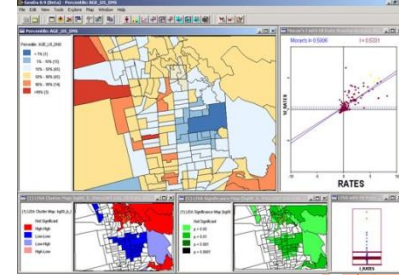
EXPLORATORY SPATIAL DATA ANALYSIS?



- **Classification of ESDA Methods**
- **Global or whole map** statistics process **all** the cases for one (or more) attributes
- or **local** statistics which process **subsets** of the data one at time
 - applying global or focused statistics only to cases in the window
 - executing a spatial query: e.g. identify all the areas within the window possessing attribute property x.



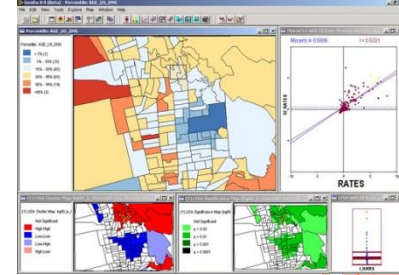
EXPLORATORY SPATIAL DATA ANALYSIS?



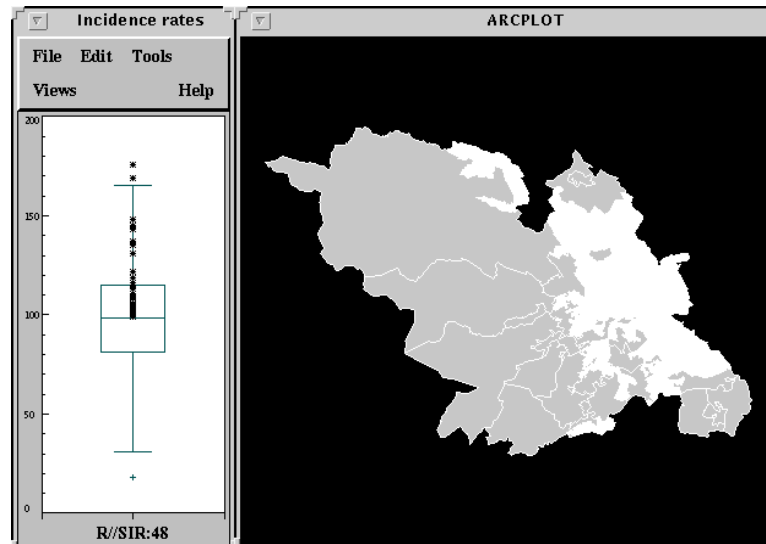
- **Describing Non-Spatial Properties of Attribute**
- Median-the location (or centre) of the distribution
- **ESDA query**- values below median?
- Quartiles and Inter-quartile spread.
 - The measure of spread of values about the median.
 - **ESDA query**: which are the areas that lie in the upper (lower) quartile?



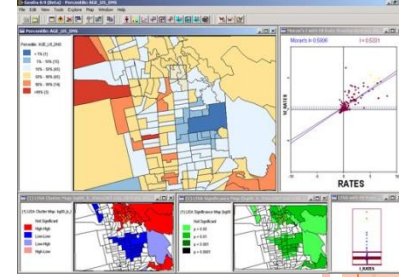
EXPLORATORY SPATIAL DATA ANALYSIS?



- Box plots.
 - Graphical summary of the distribution of attribute values.
 - **ESDA query:** where do cases that lie in specific parts of the boxplot occur on the map?; where are the outlier cases located on the map?

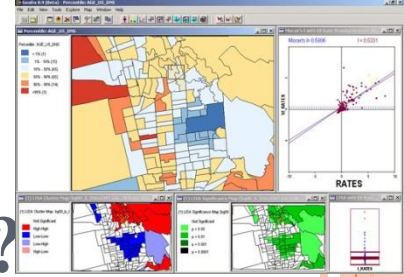


EXPLORATORY SPATIAL DATA ANALYSIS?



- **ESDA for Describing Spatial Properties of an Attribute**
- Smoothing.
- Where the map consists of many small areas- apply simple smoothing methods
- depends on the scale of the smoother- reveal the presence of general patterns that are
- **Technique:** simply take the attribute value of an area and its neighbours and average them. – Median can also be used

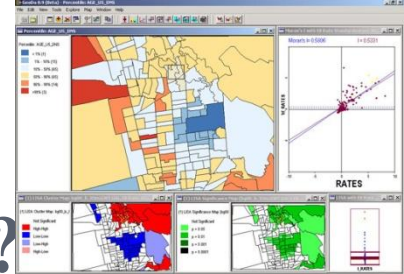




EXPLORATORY SPATIAL DATA ANALYSIS?

- E.g the existence of a general increase in heart disease incidence from Southern to Northern England



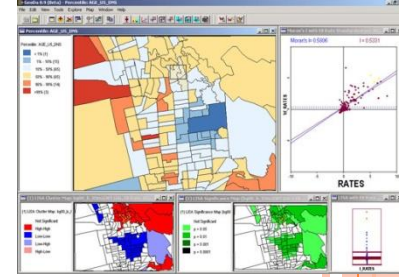


EXPLORATORY SPATIAL DATA ANALYSIS?

- Identifying trends and gradients on the map.
- Are there any general trends or gradients in the map distribution of values () **ESDA techniques:**
 - kernel estimation
 - taking transects through the data and plotting with attribute value on vertical axis and spatial location on horizontal axis
 - spatially lagged boxplots with lag order specified with respect to a particular area or zone (Haining 1993 p224)
 - two way median polish adapted to non-regular lattice data suggested by Cressie and described in Haining 1993 p215-220.

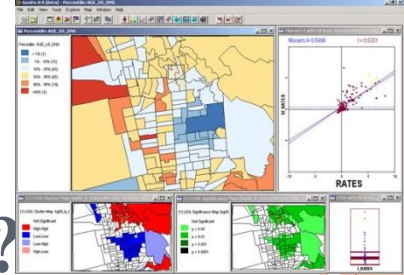


EXPLORATORY SPATIAL DATA ANALYSIS?



- Spatial autocorrelation.
- This is the propensity for attribute values in neighbouring areas to be similar.
- **ESDA technique:** use a scatterplot with area attribute value on the vertical axis plotted against the average of the attribute values in the adjacent areas on the horizontal axis
- A scatterplot upward sloping scatter to the right - -positive spatial autocorrelation (adjacent values tend to be similar).
- Scatter slopes upward to the left this is indicative of negative spatial autocorrelation (adjacent values tend to be dissimilar).



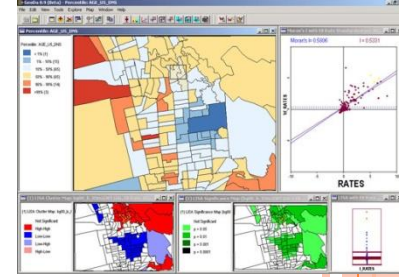


EXPLORATORY SPATIAL DATA ANALYSIS?

- Detecting spatial outliers
- This is the situation where an individual attribute value is not necessarily extreme in the distributional sense!
- but is extreme in terms of the attribute values in adjacent (spatial!) areas.
- **ESDA technique:** Use the scatterplot technique as for spatial autocorrelation and then run a least squares regression on the plot. Cases with standardized residuals greater than 3.0 or less than -3.0 might be flagged as possible outliers although for reasons beyond the level of this course this will tend to overstate the number of outliers (Haining 1993 p214). can be adapted for this purpose but has the same problem of tending to overstate the existence of outliers.



EXPLORATORY SPATIAL DATA ANALYSIS?-REFERENCE

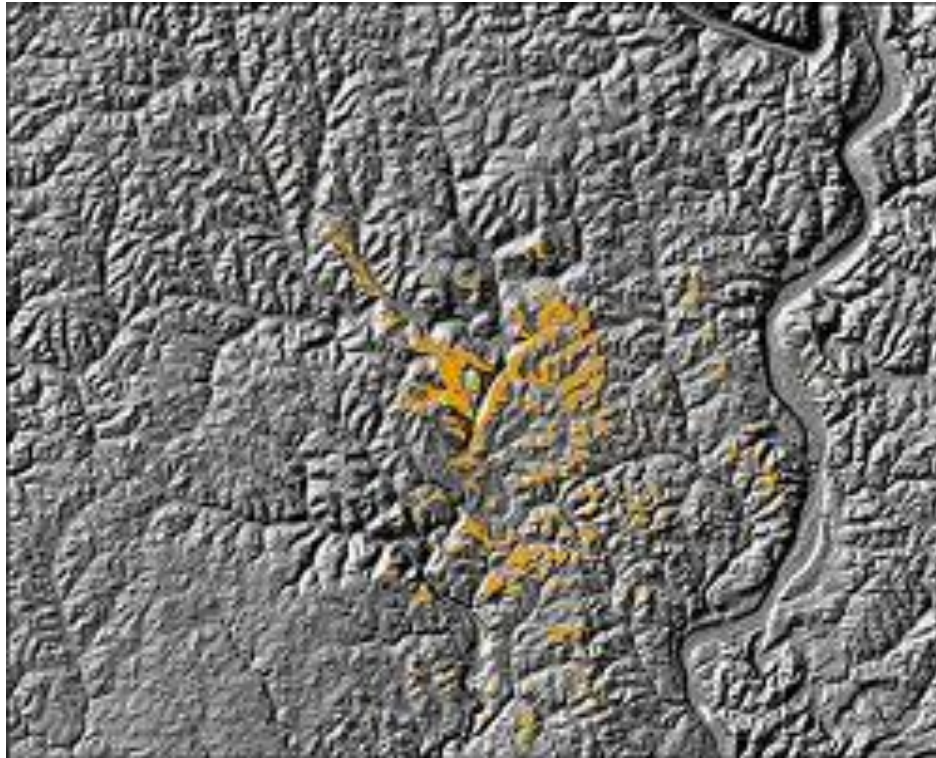


- Haining R.P.(1993) *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press.
- NCGIA- National Centre for Geographical Information Analysis





3D ANALYSIS AND SURFACE RENDERINGS.



3D ANALYSIS AND SURFACE RENDERINGS VIEWSHEDS

- A viewshed is an area that is visible from a specific location.
- The analysis uses the elevation value of each cell of the DEM to determine visibility to or from a particular cell(raster).
- E.g.
 - Locate communication towers- masts
 - View from a road



VIEWSHEDS

- Viewsheds can be calculated using an individual point such as a tower or multiple points such as a line representing a road
- **algorithm** that estimates the difference of elevation from one cell (the viewpoint cell) to the next
- each cell between the viewpoint cell and target cell is examined for line of sight
- Software-
- ArcMap, Maptitude or ERDAS IMAGINE

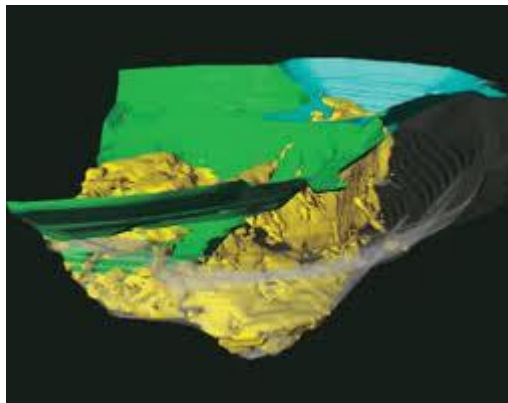
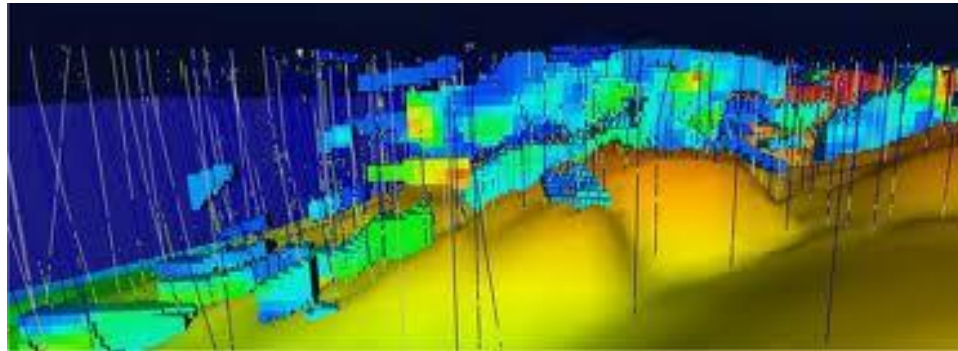
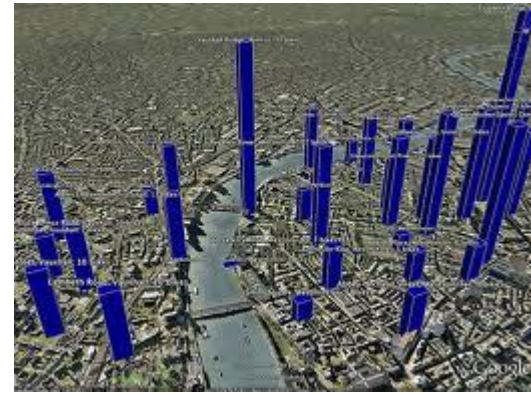


VIEWSHEDS

- cells of higher value are between the viewpoint and target cells the line of sight is blocked
 - not be part of the viewshed
- If not include (Kim et al., 2004)
- several variables can be used to limit or adjust the calculation
- a limited viewing angle. . A vertical angle can be added as well
 - 90 degrees (looking straight up) to -90 degrees (looking straight down)



VISUALISATION



VISUALISATION-3D

- 50 percent of the brain's neurons are involved in vision
- Considering the cartographic complexity of some terrain
- simulates spatial reality
- Geographic Visualization depends on psychological cues to create a natural 3D
- subtle change of color or texture with distance
- Depth-
 - the addition of atmospheric attenuation such as fog or haze
 - feature obstruction and overlap



VISUALISATION-3D

- In ArcGIS Arcscene
- 3D is ignored or disregarded
- But!
- Z coordinates data is now easily recorded or readily available for surface features!
- geological analysis
- 2D → with different layers to identify the relationship between lithology and landforms
- 3D → a geologic map can simply be draped over the topography



VISUALISATION



