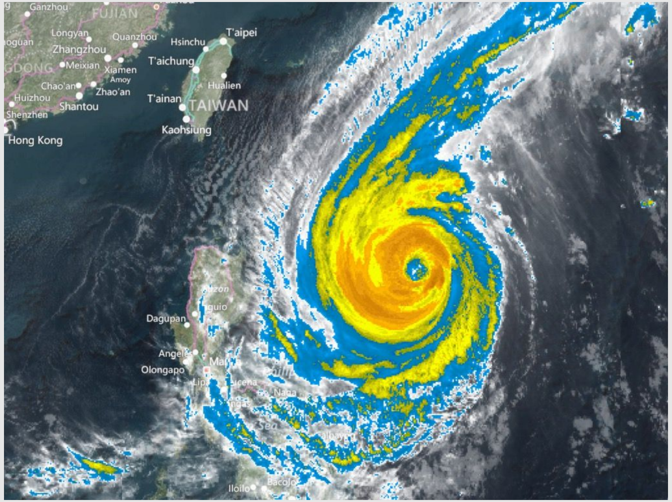
Visualizing Spatial Data

Agenda

- Spatial data
  - 1-D, 2-D, and 3-D data
  - Dynamic data
  - Combining techniques
- Geospatial data
  - Visualizing spatial data
  - Visualization of Point, Line, and Area data

Spatial data

- It has an (several) implicit or explicit spatial or spatio-temporal attribute(s).
- Scientific data, often describing some phenomenon
  - e.g., a typhoon/hurricane (insert an image here)
  - data values are recorded along a path in space



- Spatial attributes can be 1-D, 2-D, and 3-D.

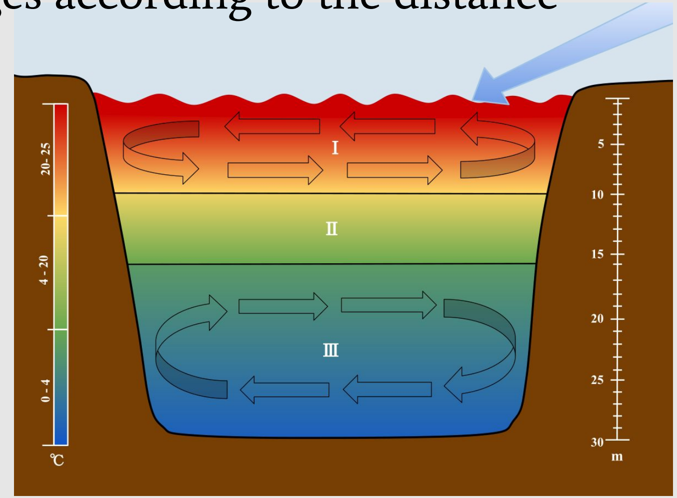
Visualization of 1-D spatial data

Examples:

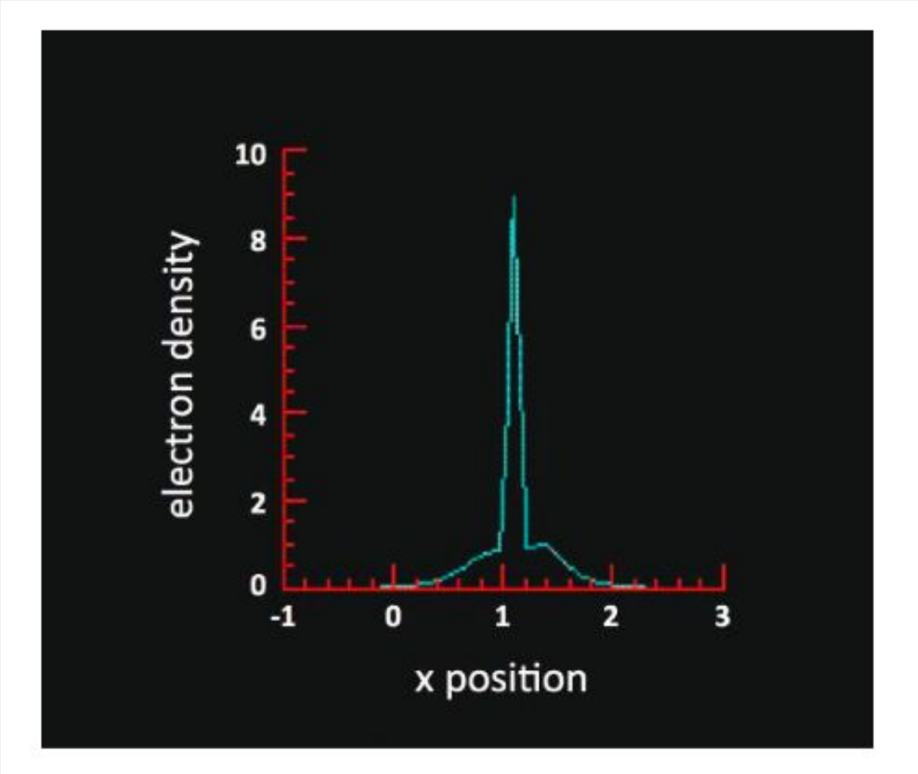
- A drill-hole sample will contain mineral content and ore grade information based on the distance from the top of the drill-hole.
- Sea water temperature data, the temperature changes according to the distance from the surface.

Solution:

- Mapping the spatial data to one of the screen dimensions
- The data values are mapped to either the other screen dimension or the colors of marks or region along the spatial axis.



Visualization of 1-D spatial data



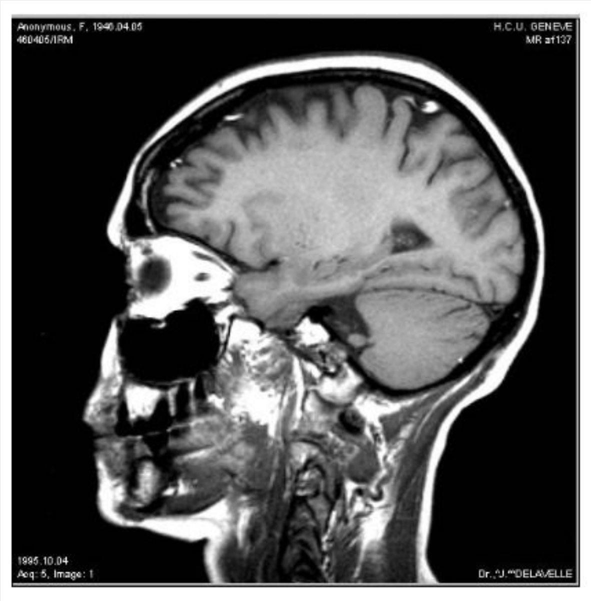
A line graph of a 1-D sequence of data values.

- Scaling is needed to fit the data to the screen.

Visualization of 2-D spatial data

Solution:

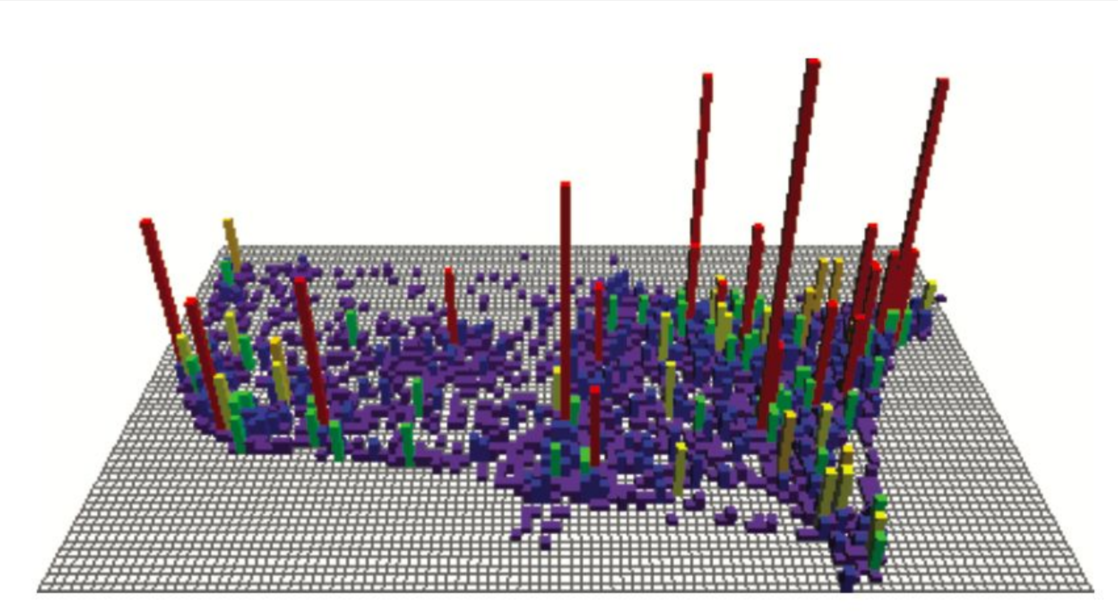
- Mapping the two spatial dimensions to the two axes of the screen.



An tomographic image.

- Each pixel contains (x, y, intensity).

Visualization of 2-D spatial data



A cityscape showing the density of air traffic over US at a particular time period.

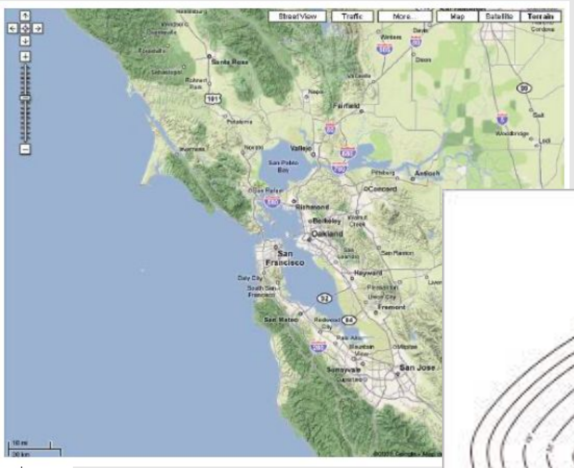
Cityscape

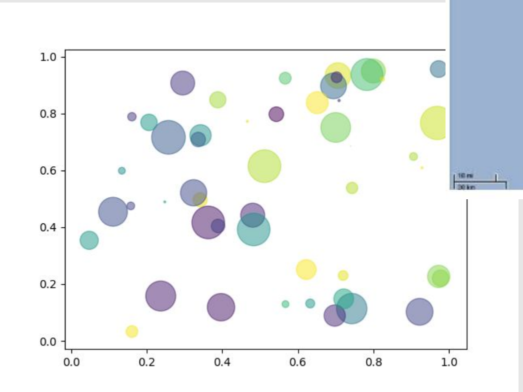
- Each data item contains several attributes
- Spatial attributes are used to locate the data item in the representation

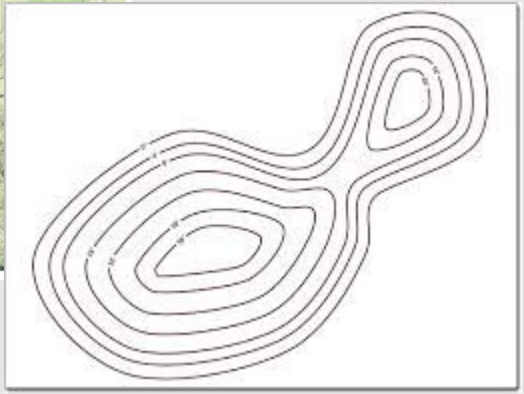
Visualization of 2-D spatial data

Other ways:

- Scatter plot
- Map
- Contour





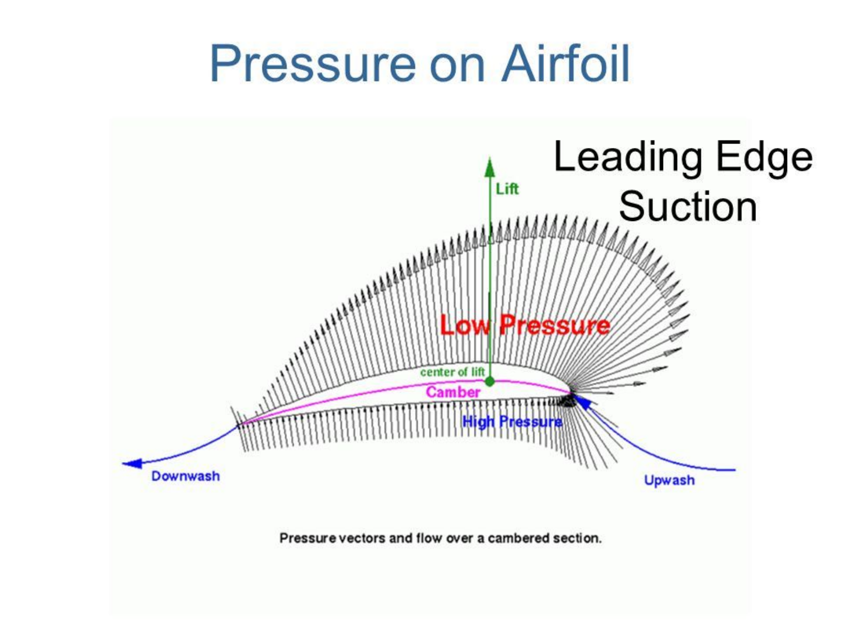


Visualization of 3-D spatial data

Values associate to some surface, for example:

- Air flow around a wing
- Stress attributes of a mechanical parts

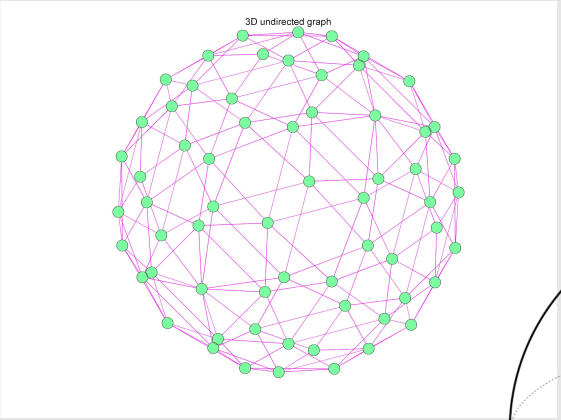
As with the 2-D cases, the spatial data attributes are used to orientate the data values

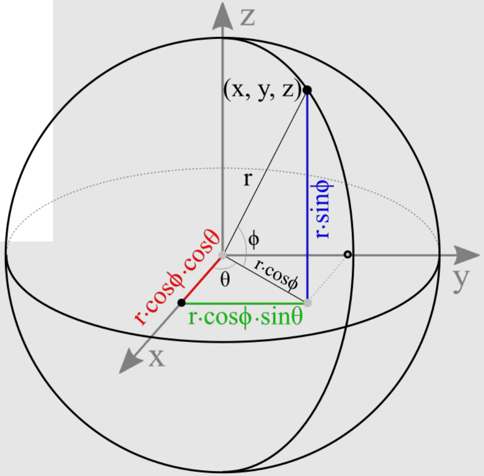


Visualization of 3-D spatial data

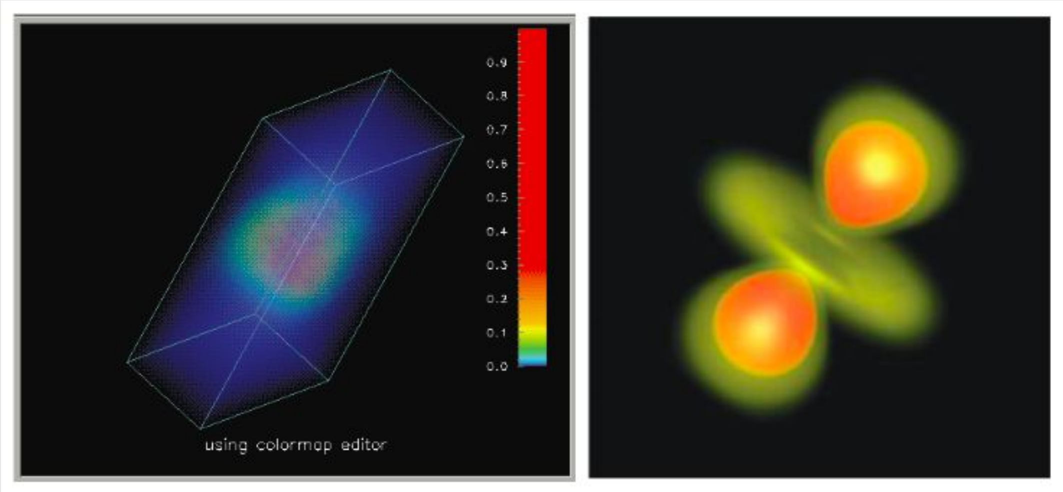
We can define the surface from two ways:

- a set of 3-D vertices and a list of edges, connecting vertices by the edges
- a set of equations





Visualization of 3-D spatial data



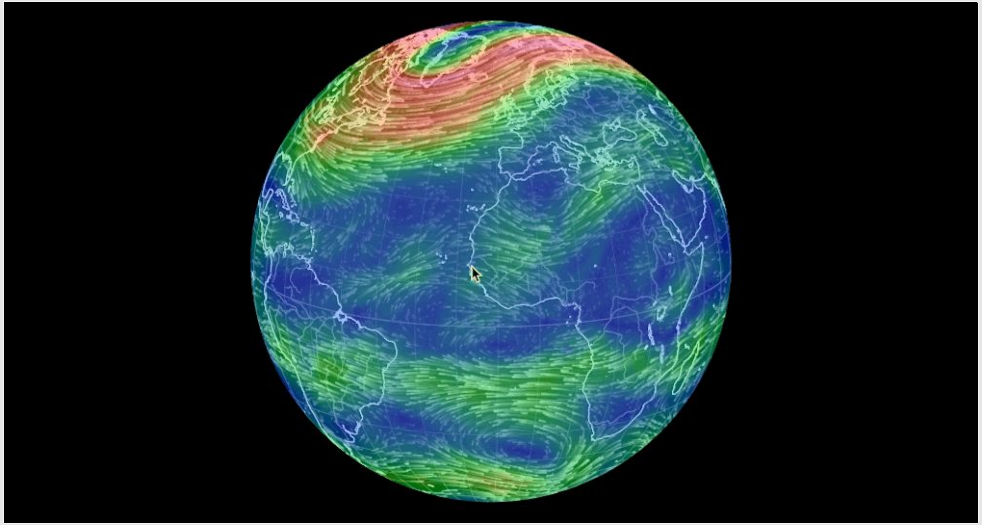
- The left image is generated using an emissive model. All points light in a level proportional to their value.
- The right image uses a texture-based approach and includes diffuse and specular lighting components.

Dynamic data

Example:

- Flow visualization; displaying the dynamic behavior in liquids and gases.

It uses a mathematical model of particle advection to visualize the flows.



Geospatial data

A kind of spatial data whose spatial attributes represent a specific location in the real world.

- Geospatial data arise in many applications.
  - Environment records
  - Census demographics
  - Credit card payments
  - Online social network
  - Telephone calls

- geographic location is important for many applications:
  - Climate modeling, economic and social measures, customer analysis

Visualizing Geographical data

Solution: Maps

- Map visualizations are a set of **points**, **lines** and **areas.**
  - Spatial contexts are visible
  - Additional information can be shown by size, shape, texture, color, and orientation, etc.

Map projections

Map projection: mapping the position on the globe (sphere) to positions on the screen (flat surface)

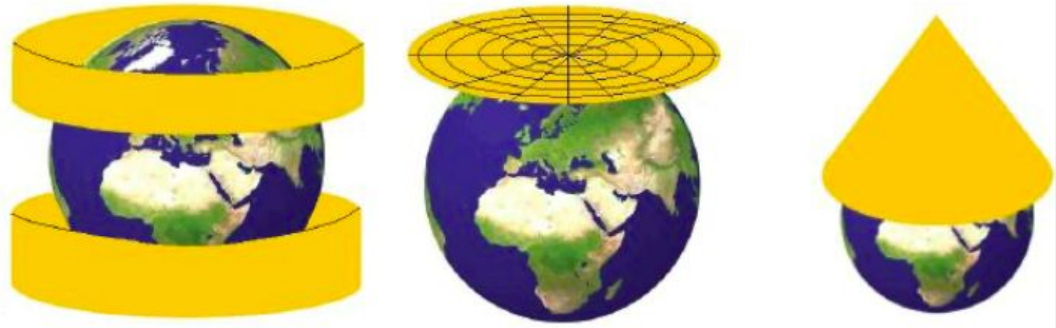
- Longitude: [-180, 180], ("-" stands for western, "+" stands for eastern)
- Latitude: [-90, 90], ("-" uses for southern, "+" uses for northern)

Map properties

- Conformal: retains the local angles; but the area is not preserved
- Equivalent (or, equal area): a specific area on a map is exactly the same to the surface on the globe; area-accurate; result in distortion of angles.
- Gnomonic: all great circles to be displayed as straight lines, i.e., preserve the shortest route between two points.
- Azimuthal: preserve the direction from a central point. (e.g., plane projection)
- Retroazimuthal: the direction from a point S to a fixed location L corresponds to the direction on the map from S to L.

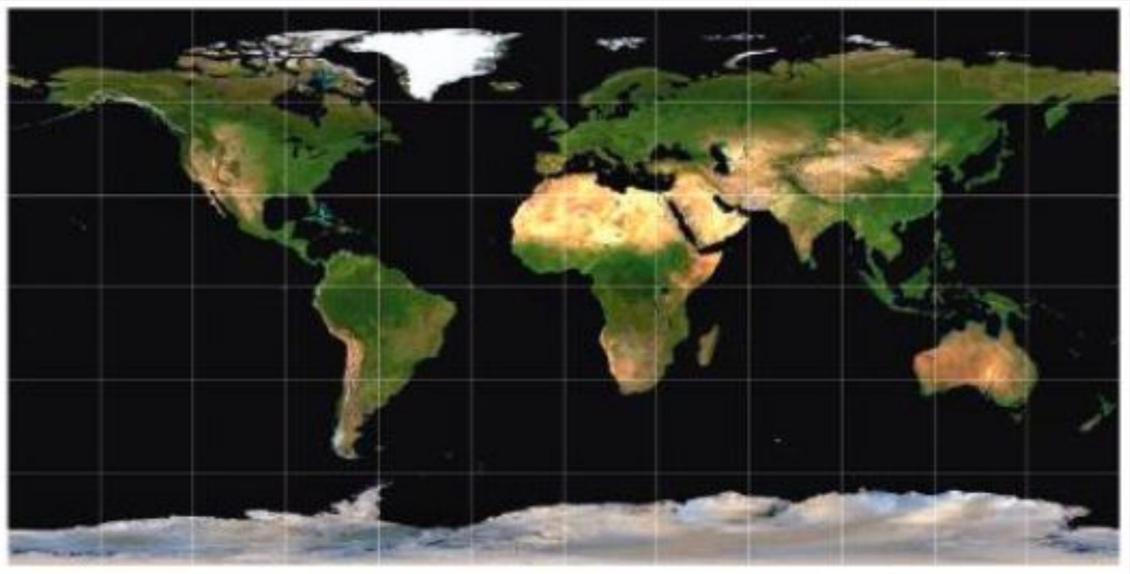
Projection types

- Cylinder projections: project the surface of the sphere on a cylinder that is put around the sphere.
- Plane projections: map the surface of the sphere to a plane that is tangent to the sphere, with the tangent point corresponding to the center point of the projection.
- Cone projections: map the surface to the cone that is tangent to the sphere



Cylinder, plane, and cone projections.

Equirectangular projection



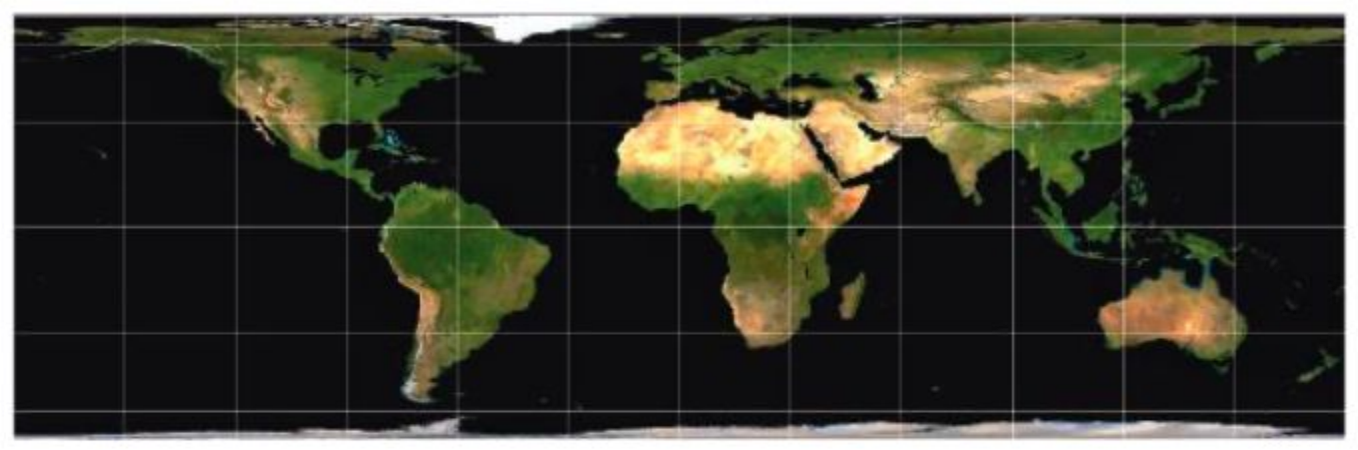
x = longitude

y = latitude

Neither conformal nor equal area;

Mainly use in thematic mapping

Lambert cylindrical projection

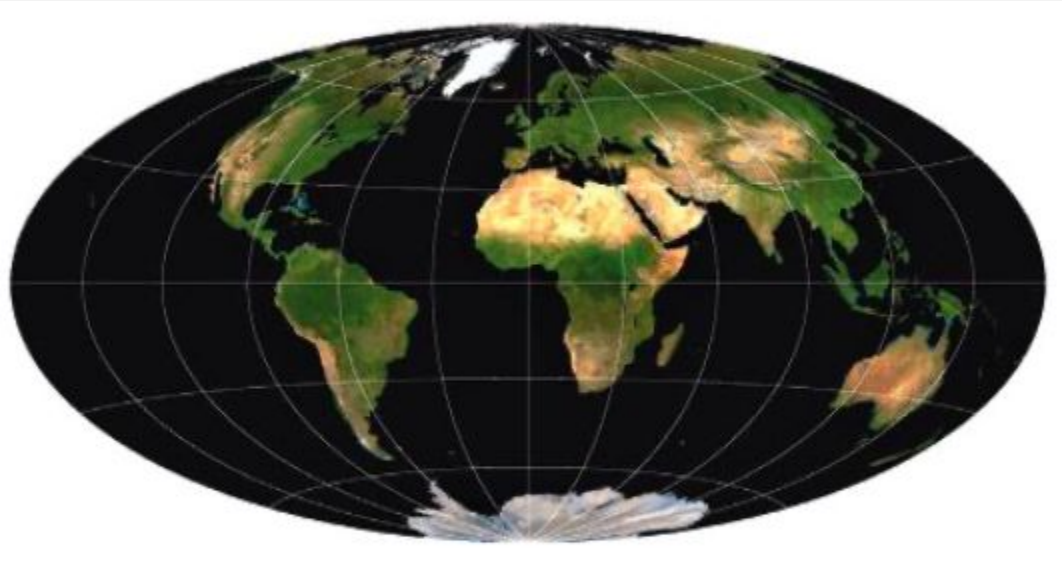


It is an equal area projection.

$$x = (\lambda - \lambda_0) * \cos\varphi_0, \qquad y = \frac{\sin\varphi}{\cos\varphi_0}.$$

λ -- the longitude; λ_0-- meridian; φ -- latitude; φ_0 -- equator;

Hammer-Aitoff projection

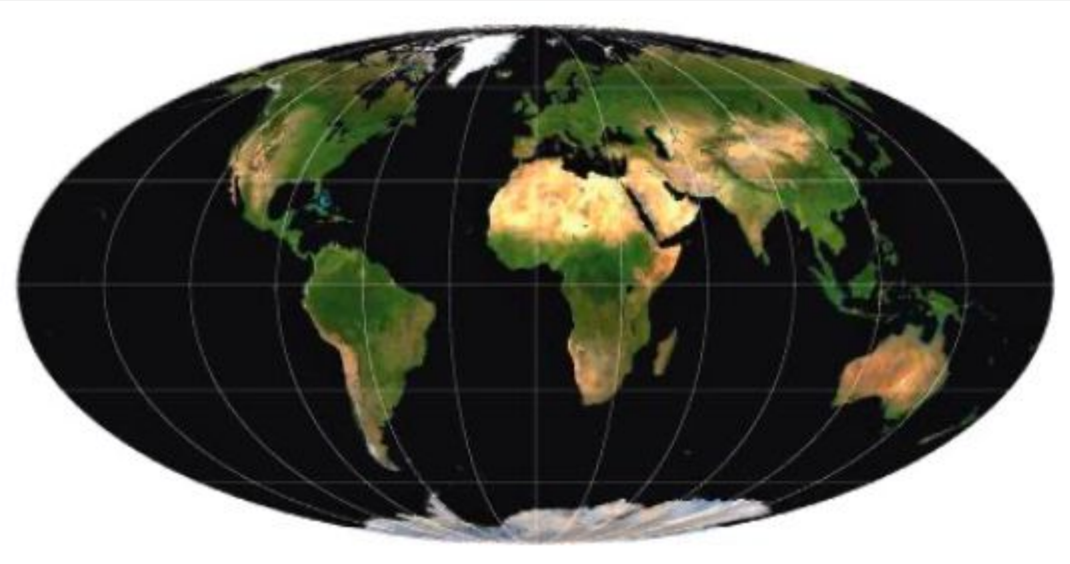


The central meridian and the equator are straight lines.

- an equal area projection.
- Mainly used in thematic world mapping

$$x = \frac{2\sqrt{2}\cos\varphi\sin\frac{\lambda}{2}}{\left(1+\cos\varphi\cos\frac{\lambda}{2}\right)^{\frac{1}{2}}}, \qquad y = \frac{\sqrt{2}\sin\varphi}{\left(1+\cos\varphi\cos\frac{\lambda}{2}\right)^{\frac{1}{2}}}.$$

Mollweide projection

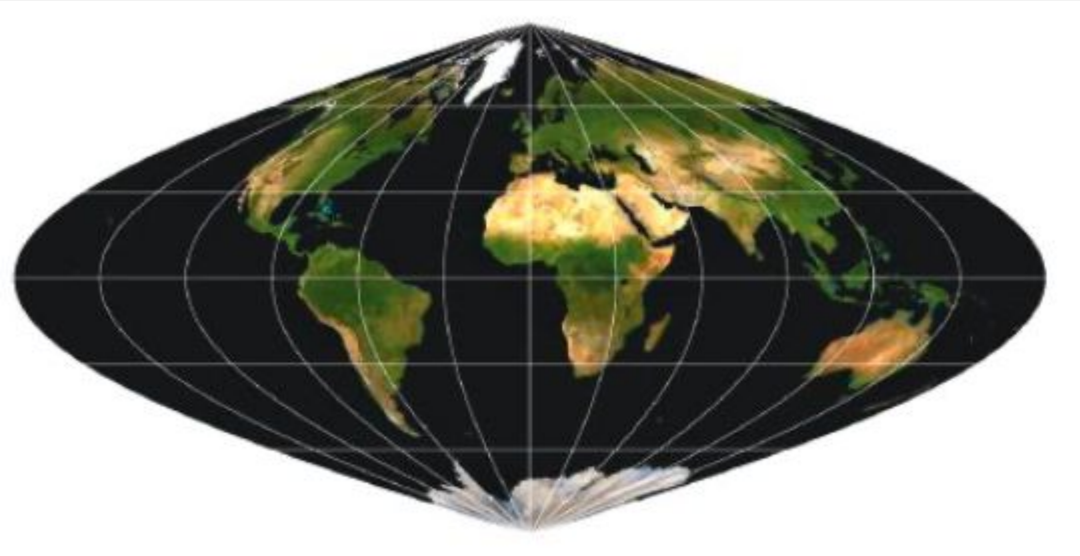


It is an equal area projection, mainly used for thematic maps of the entire world.

$$x = \frac{2\sqrt{2}(\lambda - \lambda_0)\cos\theta}{\pi}, \qquad y = 2^{\frac{1}{2}}\sin\theta, \qquad 2\theta + \sin(2\theta) = \pi\sin\varphi.$$

θ is a auxiliary angle defined by $2\theta + \sin 2\theta = \pi \sin\phi$

Cosinusoidal projection



$$x = (\lambda - \lambda_0) * \cos\varphi, \qquad y = \varphi.$$

It is an equal-area projection, which can be quickly computed.

Albers Equal-area Conic Projection

Shape and distance are both incorrected, but the distortion of these properties is minimized in the region between two standard parallels.

Mostly used for small regions with east-west orientation location in the middle latitude since the distances are most accurate there.

$$n = \frac{\cos\beta_1 + \cos\beta_2}{2}, \quad p = \sqrt{\frac{4}{n} * \sin\left(\frac{\frac{\pi}{2} - \varphi}{2}\right) + \frac{4}{n^2} * \left(\sin\frac{\beta_1}{2}\right)^2 * \left(\sin\frac{\beta_2}{2}\right)^2},$$

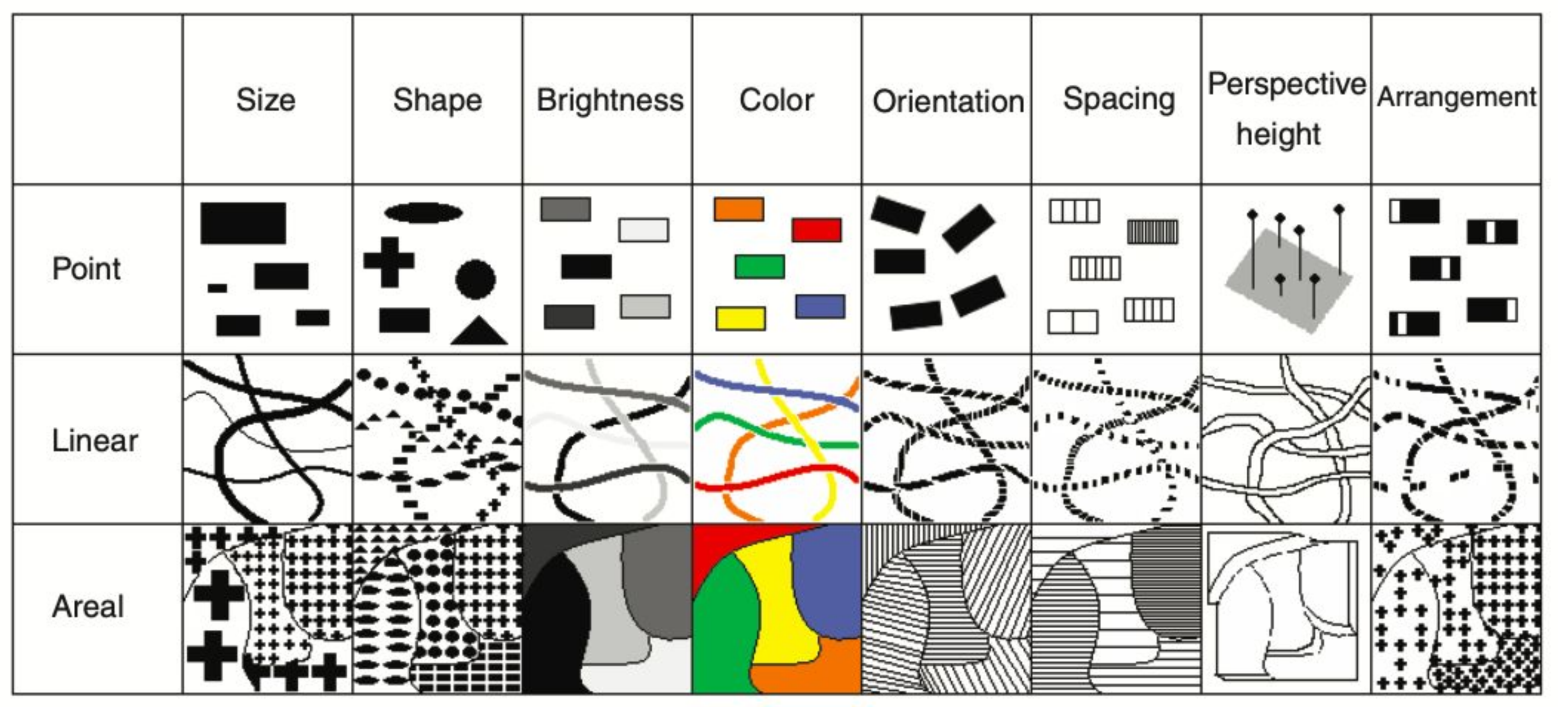
$$x = \frac{p}{\sin(n * \lambda)}, \quad y = -\frac{p}{\cos(n * \lambda)}.$$

β_1 and β_2 are two standard parallels which can be any of the latitudes.

Visual codes for spatial data

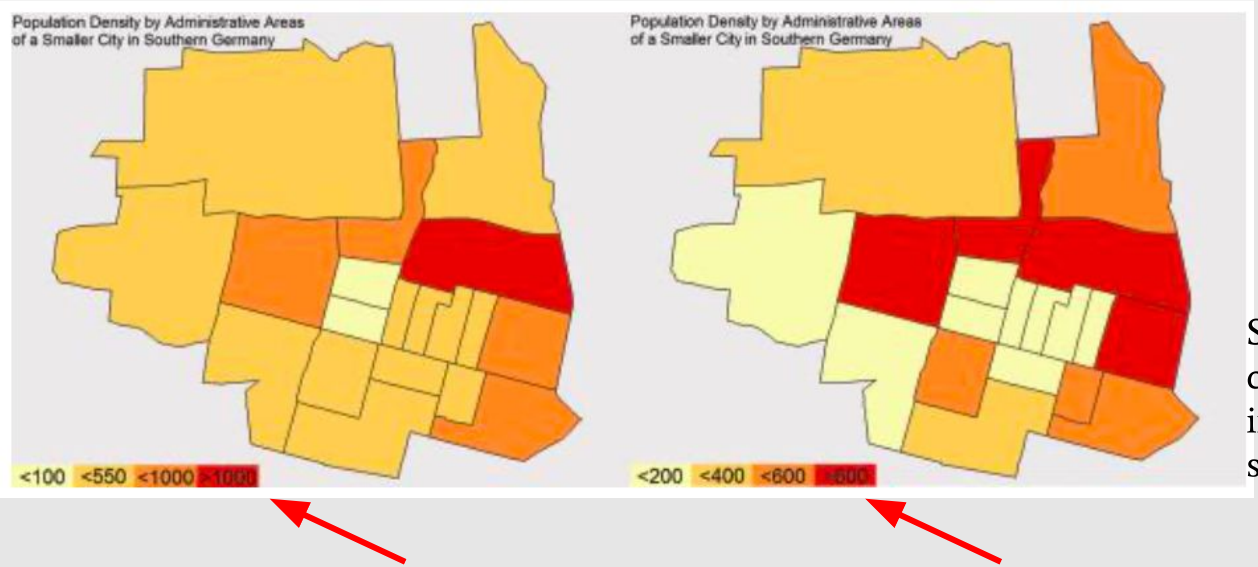
- Size -- size of symbols, width of lines
- Shape -- shape of symbols, line patterns
- Brightness
- Color
- Orientation
- Texture (spacing) -- spacing of patterns in symbols, lines, or areas;
- Perspective height -- perspective 3-D view of phenomena with the data value mapped to the perspective height of points, lines, or areas.
- Arrangement -- arrangement of patterns within the individual symbols, lines, or areas.

Visual codes for spatial data



Impact of class separation, normalization, and spatial aggregation

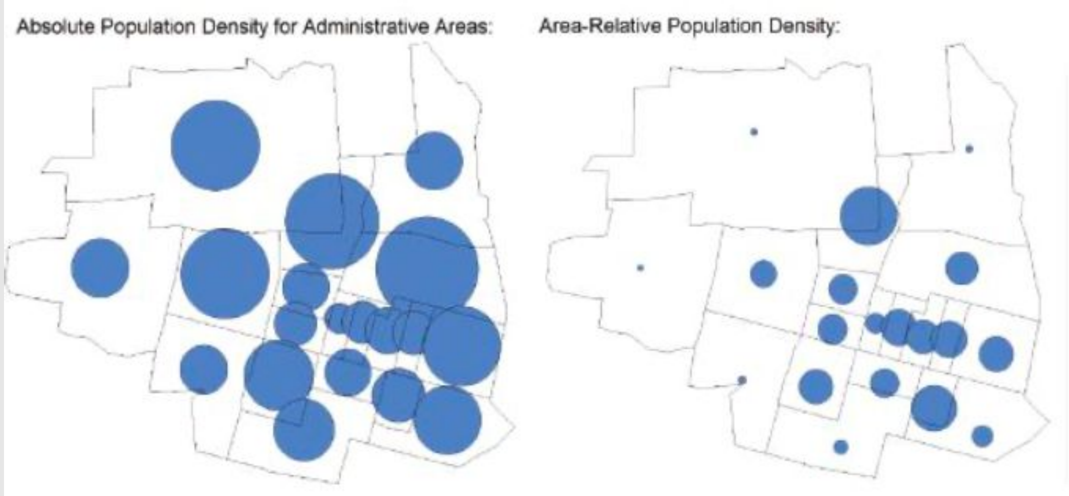
- In spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization.



Same data set with different class separation, the changes in the visualization are significant.

Impact of class separation, normalization, and spatial aggregation

- In spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization.



The left one uses absolute number of the population; the right one shows the relative population.

The absolute population difference is larger between some areas, which causes this inverted effect result.

Impact of class separation, normalization, and spatial aggregation

- In spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization.

A well-known example about London cholera with different area aggregation. Different spatial aggregation yields different choropleth maps.

Visualizing geographical data

According their spatial dimension or extent, spatial phenomena have four types:

- Point phenomena -- have no spatial extent; can be specified by a longitude and latitude pair.
    - e.g., Buildings, oil wells, cities
- Line phenomena -- have length, no width; can be specified by an unclosed series of longitude and latitude pairs
    - e.g., telecommunication networks, roads, boundaries between countries

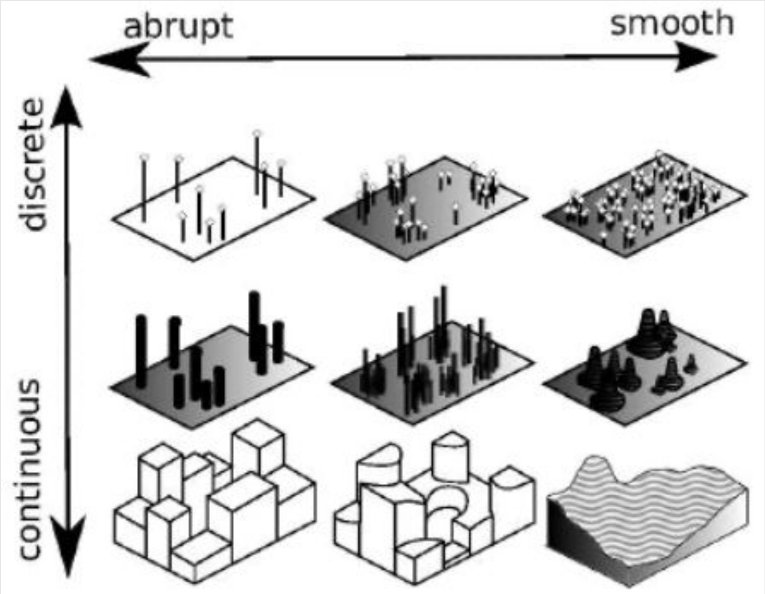
Visualizing geographical data

- Area phenomena -- have both length and width; can be specified by a set of longitude and latitude pairs that completely enclose a region.
  - e.g., lakes, parks, and states or countries
- Surface phenomena -- have length, width, and height; can be specified by a series of longitude, latitude and height vectors.
  - e.g., mountains

Visualization of Point data

Point data are discrete in nature, but they may describe a continuous phenomenon

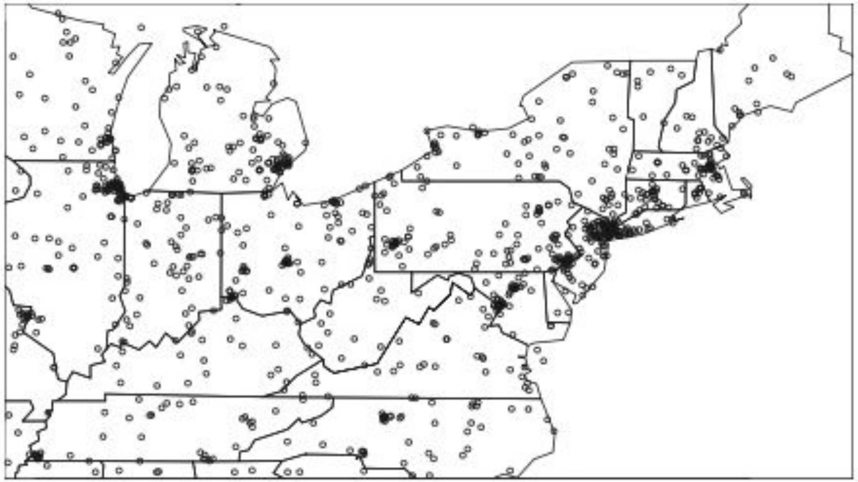
- e.g., temperature measurements at specific locations



The data can be displayed continuously or discrete, and, smooth or abrupt. It all dependents on the nature of the data and the task.

Visualization of Point Data: Dot map

Point phenomena can be visualized by placing a symbol or pixel at the location where that phenomenon occurs; this simple visualization is called dot map.



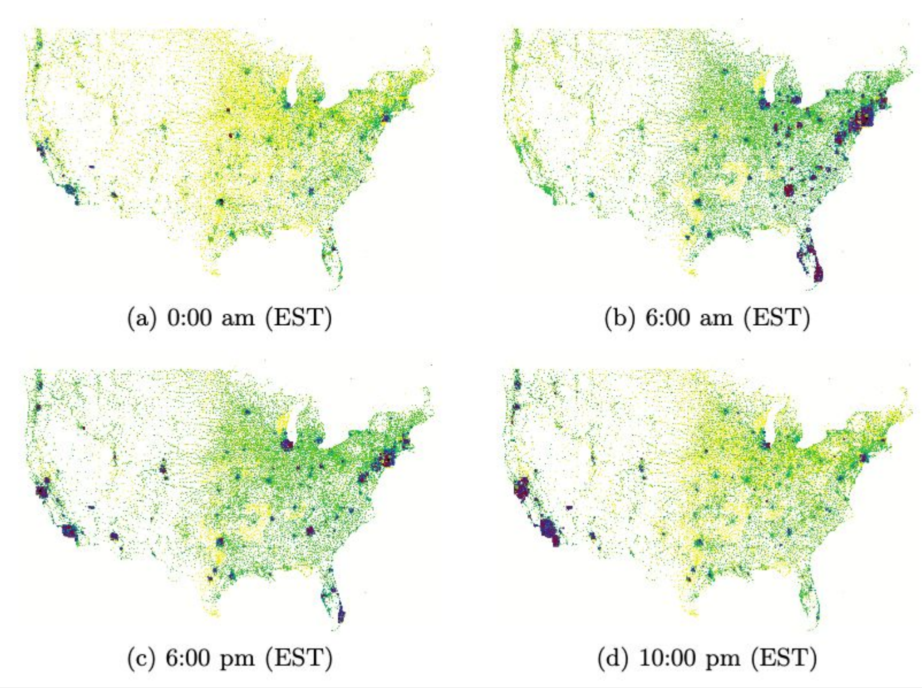
Every circle represents the spatial location of an event.

Visualization of Point Data: Dot map

- The perceived size of the symbols (marks) does not necessarily correspond to the actual size due to the problems in size perception. (perceived size depends on their local neighborhood, accuracy compress)
- If color is used to represent a quantitative parameter, the problem of color perception must be taken into account (popout, preattentive, accuracy compress)
- Overlap occurs when data set is highly dense.

Visualization of Point Data: PixelMaps

- PixelMap does not aggregate the data, but avoids overlap in the 2D display.



The figure display U.S. Telephone Call Volume at four different times during a day. The idea is to place the first data items at their correct position and position overlapping data at nearby unoccupied positions.

Visualization of Linear Data

For linear phenomena, the basic idea is to represent them as line segments between pairs of endpoints specified by longitude and latitude.

- Data pattern can be mapped to line width, line pattern, line color, and line labeling.
- Start, end, and intersection points can be mapped to the visual code of nodes, such as size, shape, color, and labeling.
- Line can be curve to avoid clutter in the display.

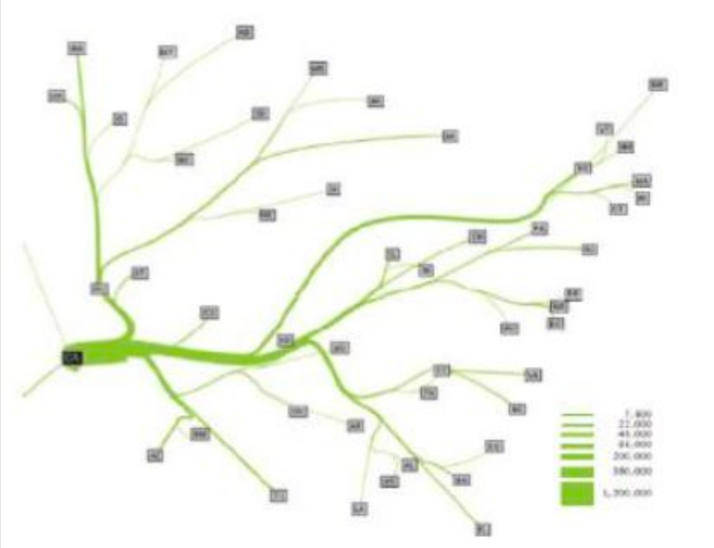
Network maps

Network can be used to display the connectivity.

Lines in dense area may overlap severely.

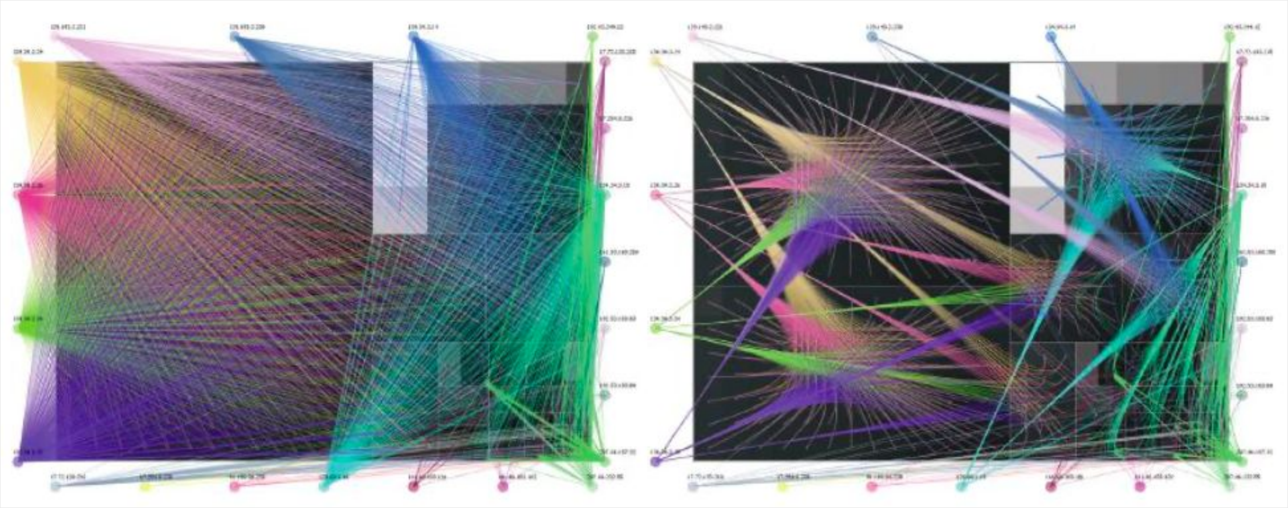
Flow Maps

- Inspired by graph layout algorithms
  - Minimizing the edge crossing
  - Retaining the relative position of nodes



Edge bundling

- The edges can be bundled according to the hierarchy defined on the nodes.
  - Nodes connected through the root of the hierarchy are maximally bent.
  - Nodes within the same sub-hierarchy are only minimally bent.



[The figures show IP flow traffic](#).

- The left side is the straight line visualization; the right side is the edge bundling visualization which significantly reduces the visual clutter.

Visualization of Area Data

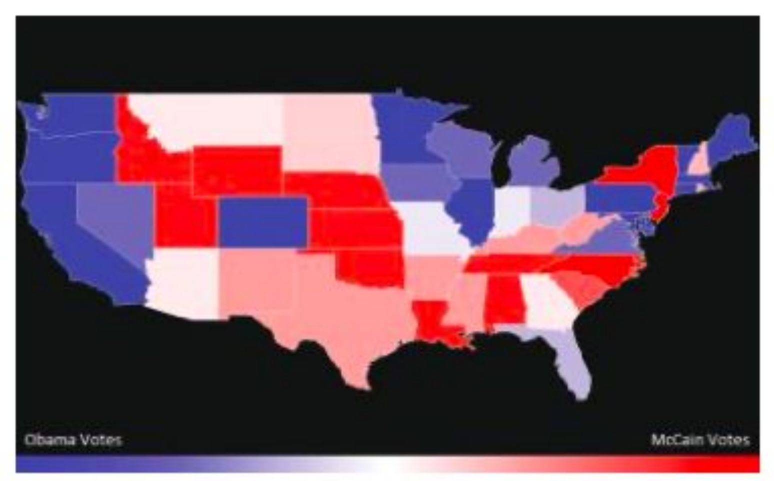
Thematic maps are the main approach to visualizing area phenomena. There are several variants of thematic maps.

- Choropleth maps
- Dasymetric maps
- Isarithmic maps
- Cartograms

Choropleth maps

- The values of an attribute or statistical variable are encoded as colored or shaded regions on the map.

  - The attribute is assumed to be uniformly distributed in the regions.
  - Data normalization and color or grayscale mapping are important issues.



A choropleth map showing 2008 U.S. election results.

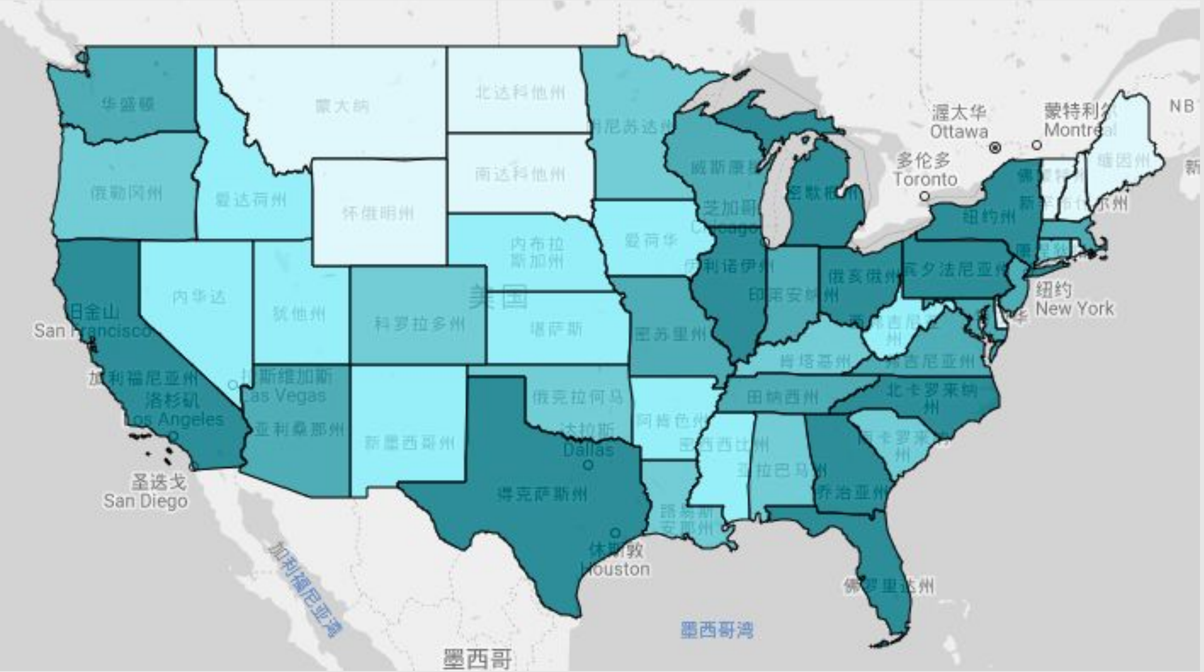
Choropleth maps

Problems:

- Most interesting values are often concentrated in densely populated areas with barely visible polygons
- Less interesting values are spread out over sparsely populated areas with large and visually dominating polygons.

Disadvantage: It tends to highlight pattern in large areas, which often be of lower importance.

Choropleth maps

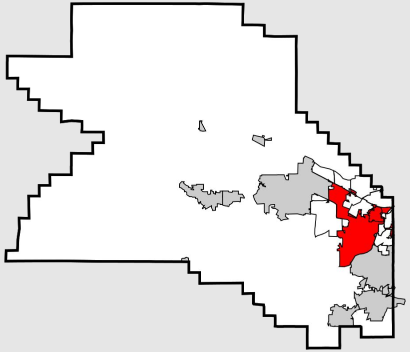


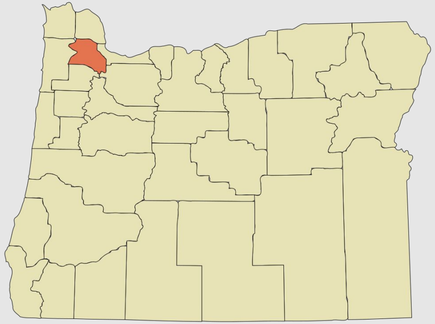
[U.S. Census Demographics]() : Readers may focus on the states with larger area rather than these with higher population.

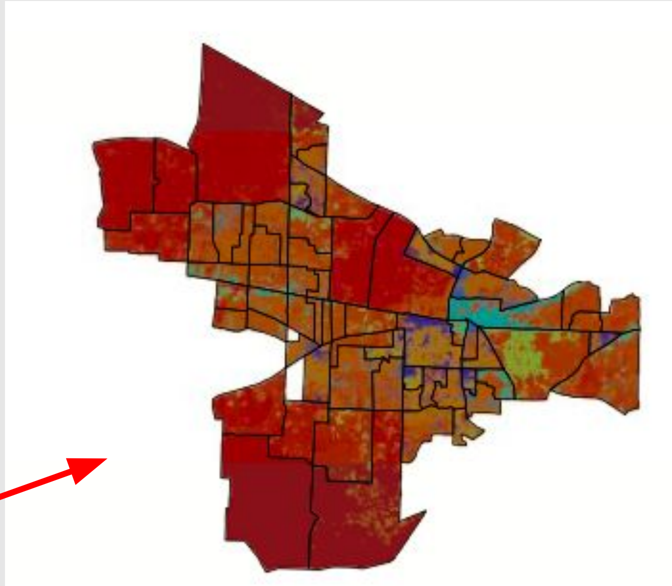
Dasymetric maps

The attribute to be shown forms areas independent of the original regions.

- e.g., boundaries of the areas derived from the attribute do not need to match the given map's regions







A dasymetric map showing the population in Beaverton Creek, Oregon, USA

Isarithmic map

It shows the contours of some continuou phenomena.

- e.g., contour maps or topographic maps
- Sometime, it needs to interpolate the data points to obtain smooth contours

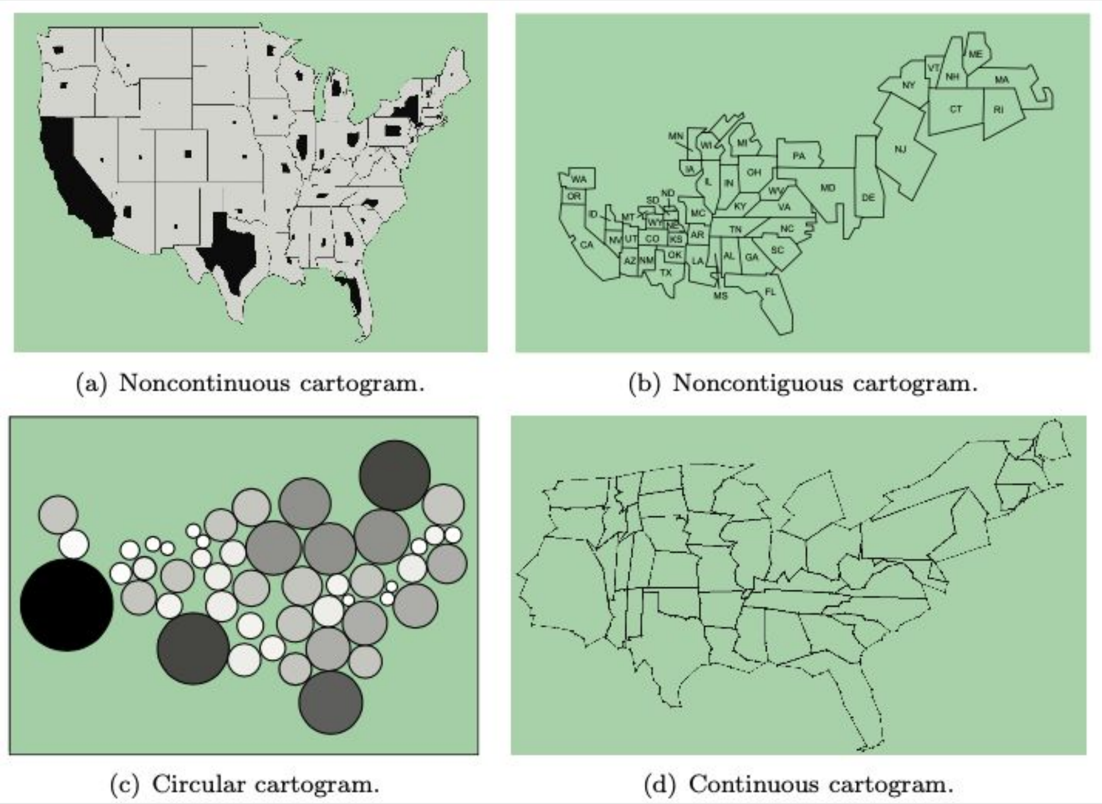
An isarithmic map showing the number of pictures taken on Mainau Island, using a heatmap, where the colors range from black to red to yellow, with yellow representing the most photographs.

Cartograms

Cartograms are generalizations of ordinary thematic maps.

- Avoiding the problems of choropleth maps by distorting the geography according to the displayed statistical value
- Regions are resized according to a geographically related attribute

Types of Cartograms



(a) Noncontinuous cartogram.

(b) Noncontiguous cartogram.

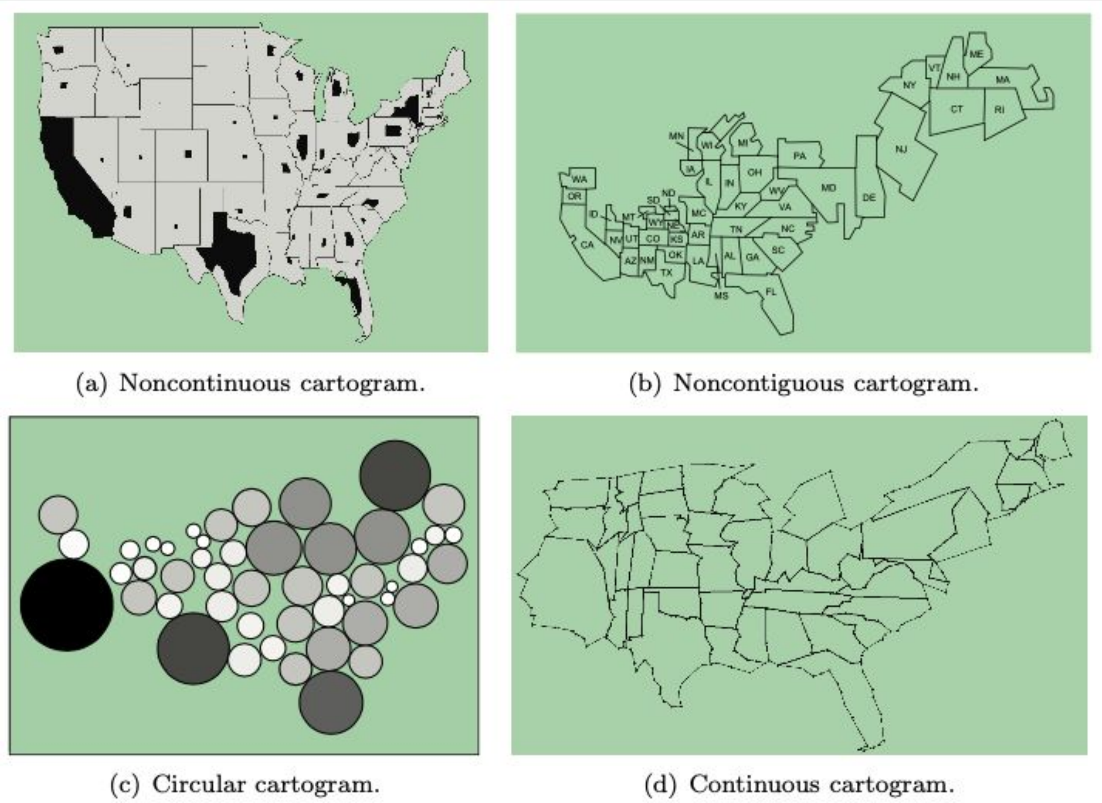
(c) Circular cartogram.

(d) Continuous cartogram.

(a) It exactly satisfies area and shape constraints, but the map's topology is not preserved. Note: the scaled polygons inside the original regions avoid the perceptual problems.

(b) It scales all polygons to their target sizes, perfectly satisfying the area objective; but the map's topology is highly relaxed since the polygons don't retain their adjacency.

Types of Cartograms

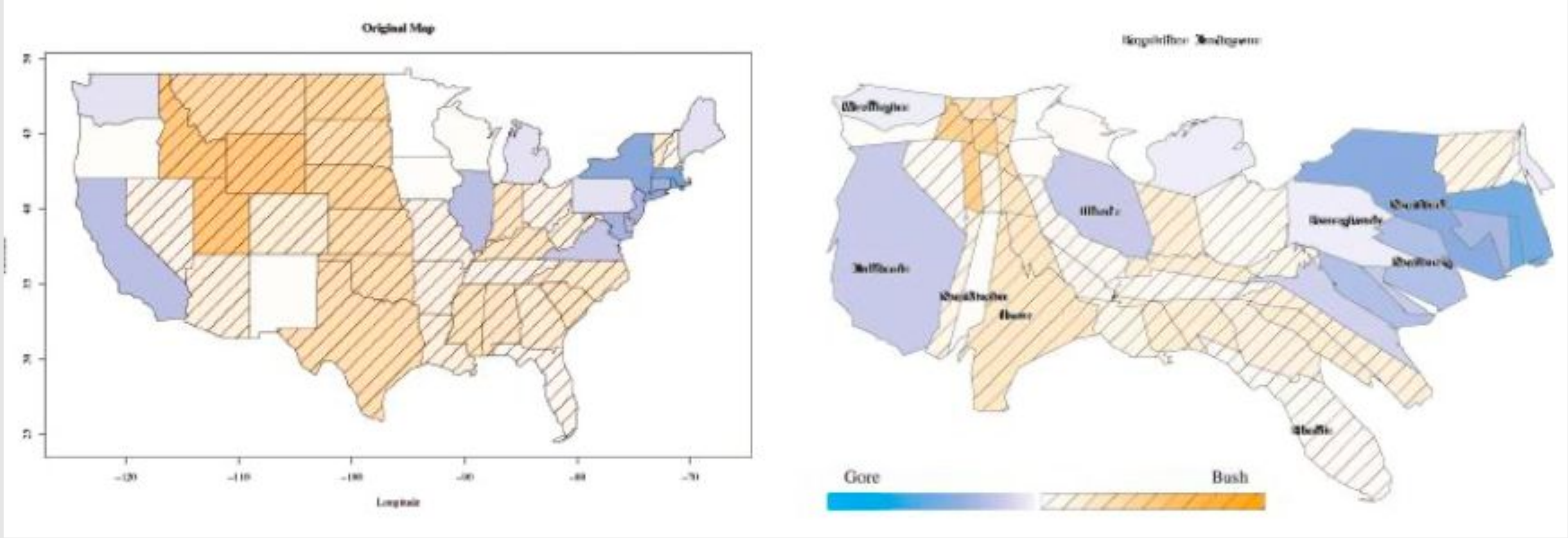


(a) Noncontinuous cartogram.
(b) Noncontiguous cartogram.
(c) Circular cartogram.
(d) Continuous cartogram.

(c) It ignores the polygons' shape, representing each as a circle in the output. The topology constraints is still perceivable.

(d) It retains a map's topology perfectly, but the area and shape constraints are relaxed.

Cartogram Examples



The left is a conventional map (choropleth map) of the 2000 U.S. presidential election; the right is a population-based cartogram of it. In the right one, the states are scaled to their population, so it reveals the close result of the election more effectively than the left one.

Cartogram Examples

A rectangular U.S. population cartogram on the state and country level. The area of the rectangles corresponds to the population and the color redundantly encodes the population numbers.