06

Visualization Techniques for Geospatial Data

Notice

- **Author**
  - **João Moura Pires (jmp@fct.unl.pt)**

Bibliography

- **Many examples are extracted and adapted from**
  - **Interactive Data Visualization: Foundations, Techniques, and Applications,**

    Matthew O. Ward, Georges Grinstein, Daniel Keim, 2015
  - **Visualization Analysis & Design,**

    Tamara Munzner, 2015

Table of Contents

Visualizing GeoSpatial Data

Geospatial data

- Geospatial data is different from other kinds of data in that spatial data describes objects or phenomena with a specific location in the real world.
- Often, spatial data sets are discrete samples of a continuous phenomenon.
- Because of its special characteristics, the basic visualization strategy for spatial data is straightforward; we map the spatial attributes directly to the two physical screen dimensions, resulting in map visualizations.
- Maps are the world reduced to points, lines, and areas. The visualization parameters, including size, shape, value, texture, color, orientation, and shape, show additional information about the objects under consideration.

Geospatial phenomena data

- **Spatial phenomena according to their spatial dimension:**
  - **point phenomena**
  - **line phenomena: have length, but essentially no width**
  - **area phenomena: have both length and width**
  - **surface phenomena: have length, width, and height**
- **Maps can be subdivided into map types based on properties of the data (qualitative versus quantitative; discrete versus continuous) and the properties of the graphical variables (points, lines, surface, volumes).**

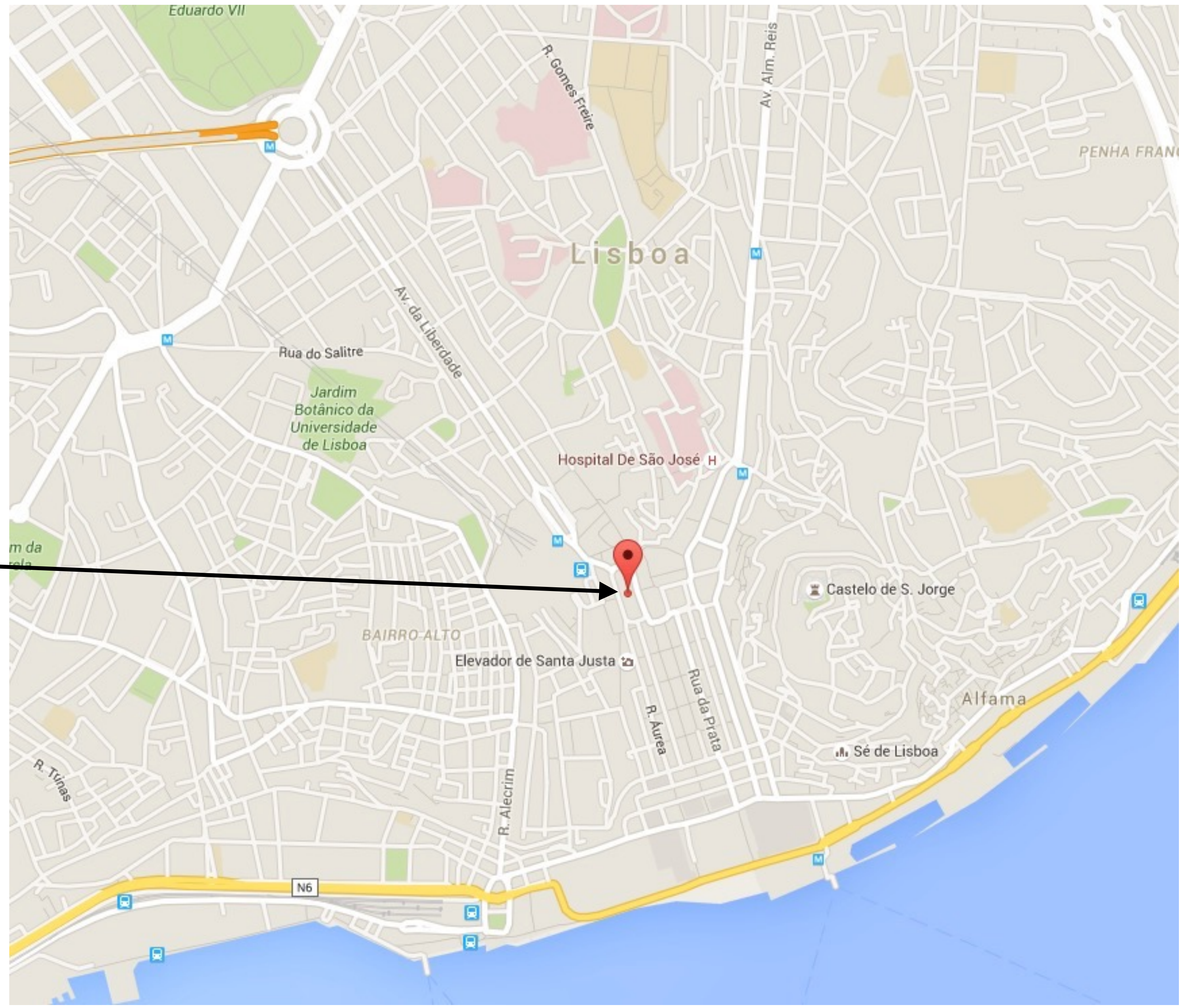
Map Projections

- **Map projections are concerned with mapping the positions on the globe (sphere) to positions on the screen (flat surface).**
- **Π : (λ, φ) → (x, y)**
  - **Degrees of longitude (λ) in [–180, 180], where negative values stand for western degrees and positive values for eastern degrees.**
  - **The degrees of latitude (φ) are defined similarly on the interval [–90, 90], where negative values are used for southern degrees and positive values for northern degrees.**

Map Projections

- **Lisbon is located at 38°42′49.75″N 9°8′21.79″W**

Rossio

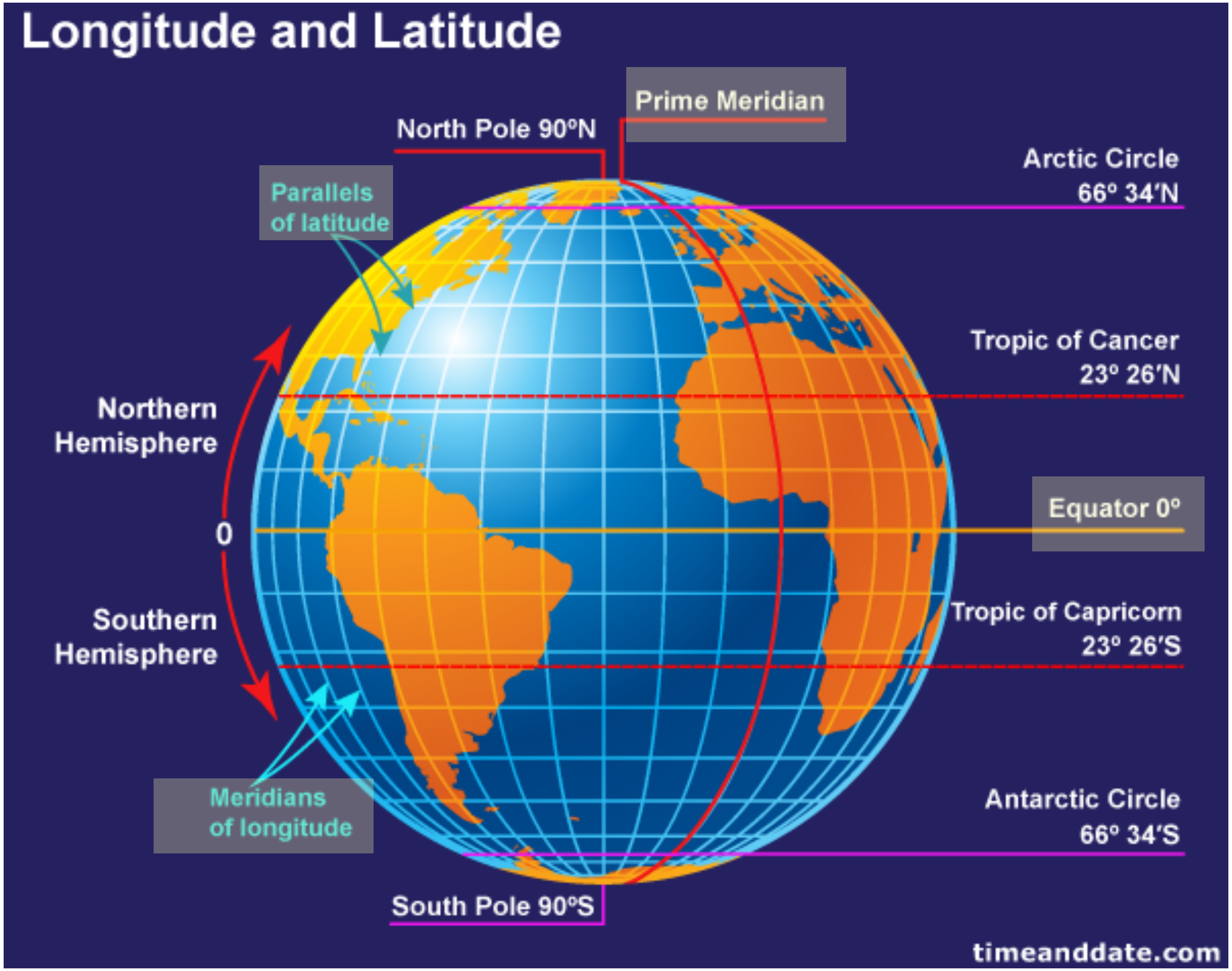


Map Projections

- **Lisbon** is located at **38°42′49.75″N 9°8′21.79″W**

Rossio

Latitude and Longitude



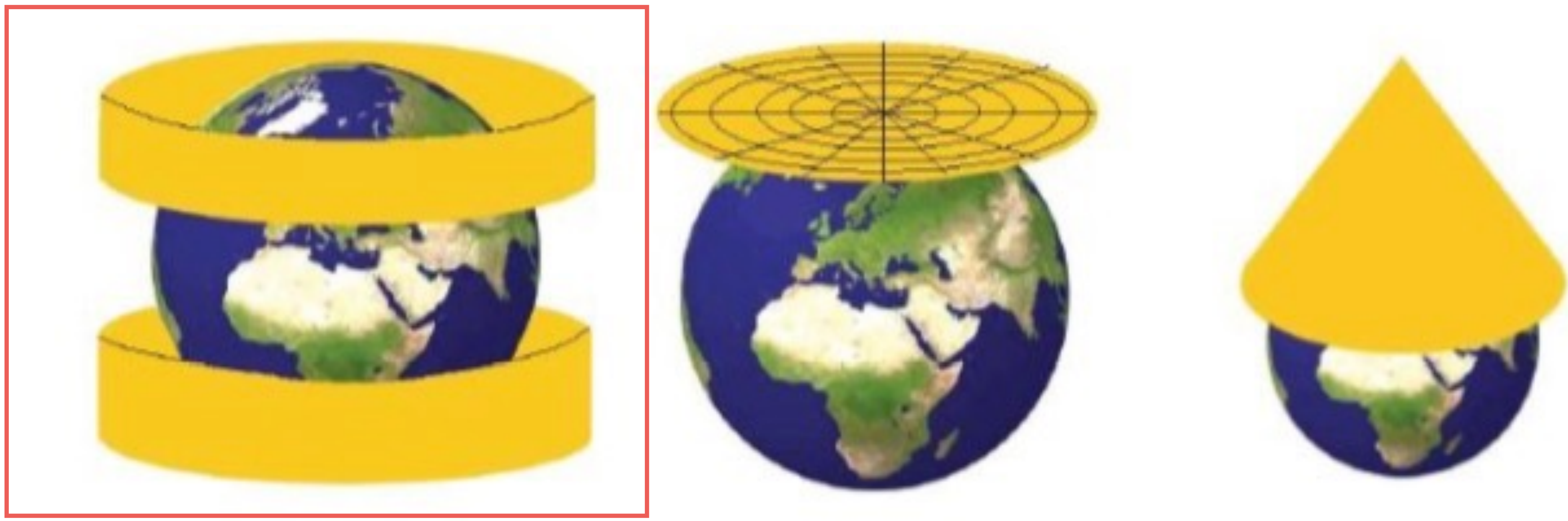
Map Projections

- A conformal projection retains the local angles on each point of a map correctly, which means that they also locally preserve shapes. The area, however, is not preserved.
- Equivalent or equal area if a specific area on one part of the map covers exactly the same surface on the globe, as another part of the map with the same area. Area-accurate projections result in a distortion of form and angles.
- Equidistant if it preserves the distance from some standard point or line.
- Gnomonic projections allow all great circles to be displayed as straight lines. Gnomic projections preserve the shortest route between two points.

Map Projections

- **Gnomonic projections** allow all **great circles to be displayed as straight lines**. They preserve the shortest route between two points.
- **Azimuthal projections** preserve the direction from a central point. Usually these projections also have radial symmetry in the scales, e.g., distances from the central point are independent of the angle and consequently, circles with the central point as center result in circles that have the central point on the map as their center.
- In a **retroazimuthal projection**, the direction from a point S to a fixed location L corresponds to the direction on the map from S to L.

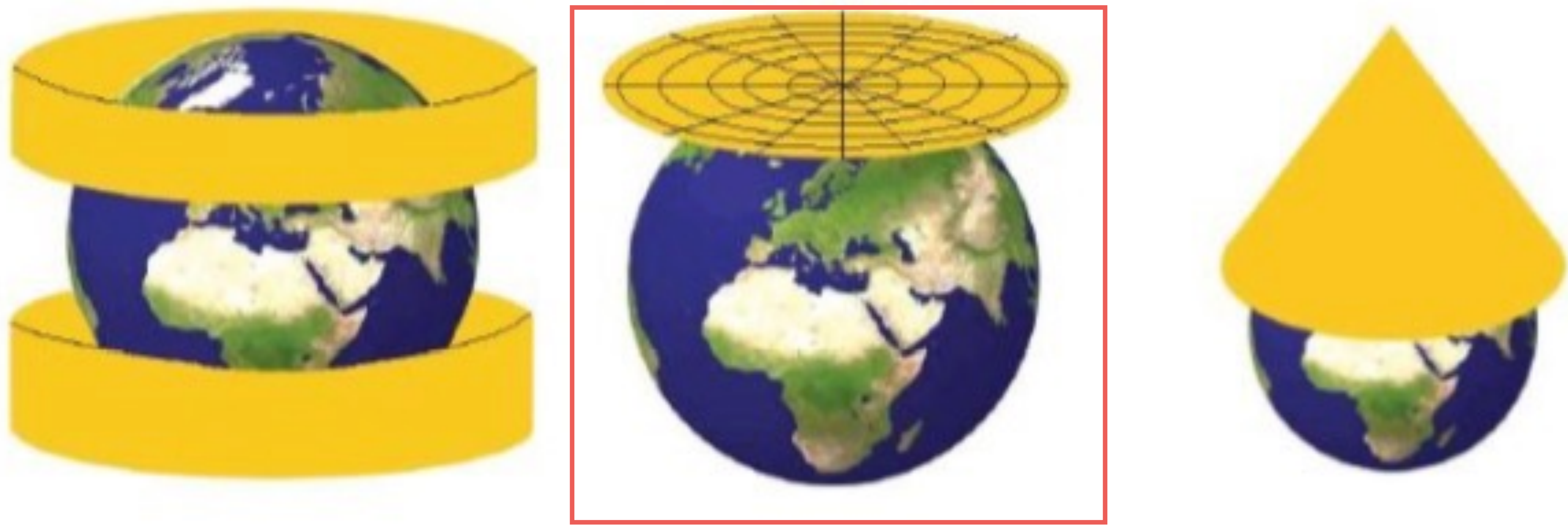
Map Projections: to what kind of surface



Cylinder, plane, and cone projections.

- **Most cylinder projections preserve local angles and are therefore conformal projections. The degrees of longitude and latitude are usually orthogonal to each other**

Map Projections: to what kind of surface



Cylinder, plane, and cone projections.

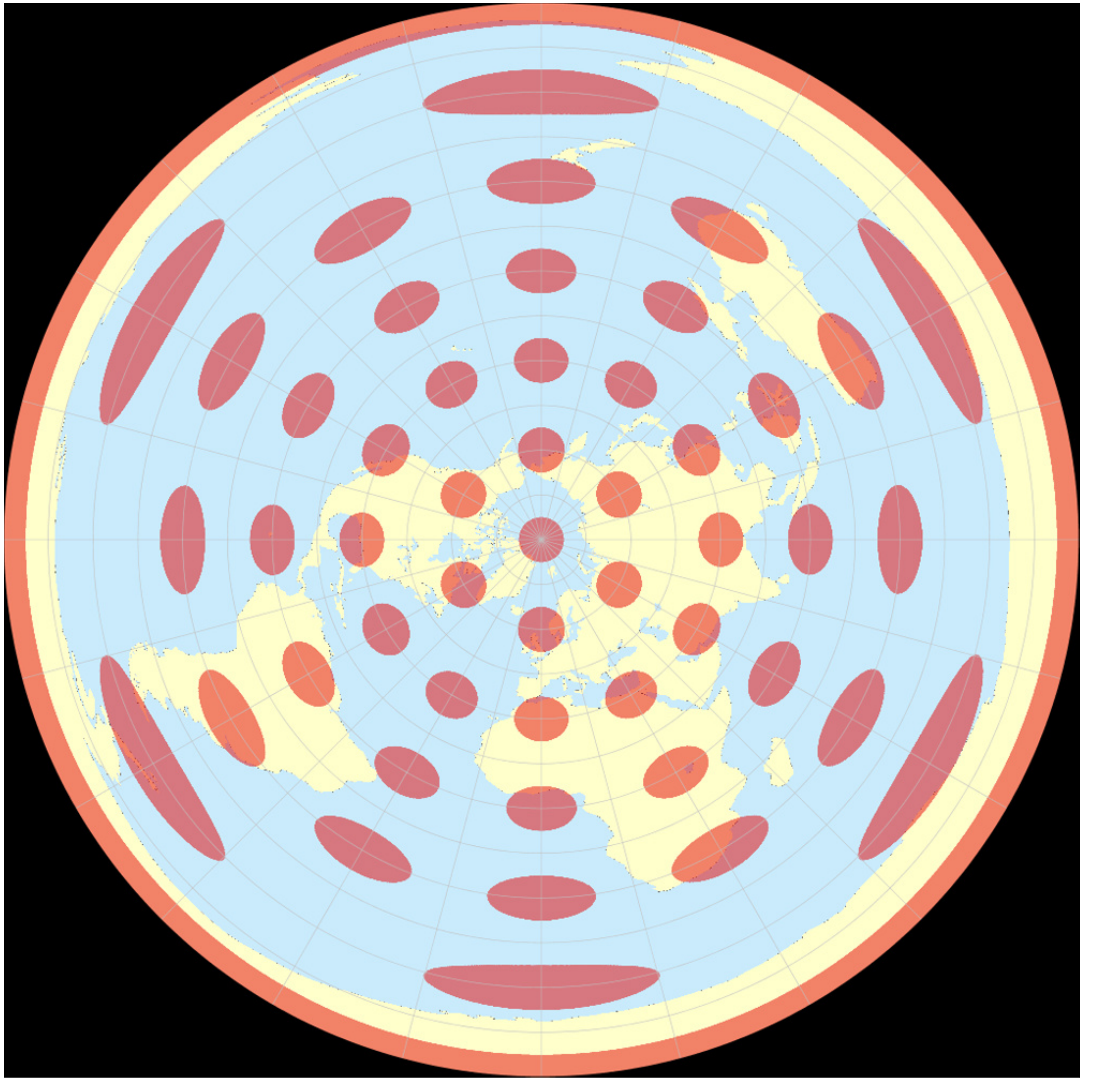
- **Plane projections** are **azimuthal projections** that map the surface of the sphere to a plane that is tangent to the sphere, with the tangent point corresponding to the center point of the projection. Some plane projections are true perspective projections.

Map Projections: to what kind of surface

Cylinder, plane, and cone projections.

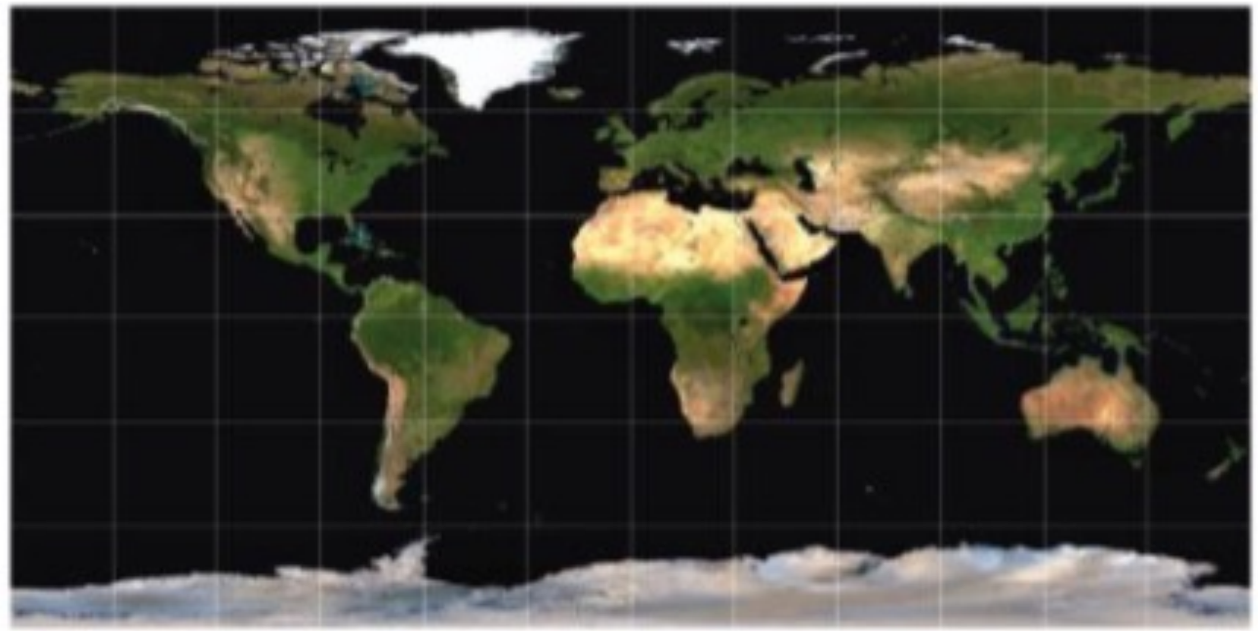
- **Cone projections** map the surface of the sphere to a cone that is tangent to the sphere. **Degrees of latitude are represented as circles around the projection center**, degrees of longitude as straight lines emanating from this center. Cone projections may be designed in a way that preserves the distance from the center of the cone.

Conic Albers Map Projection



Map Projections: Equirectangular Cylindrical Projections

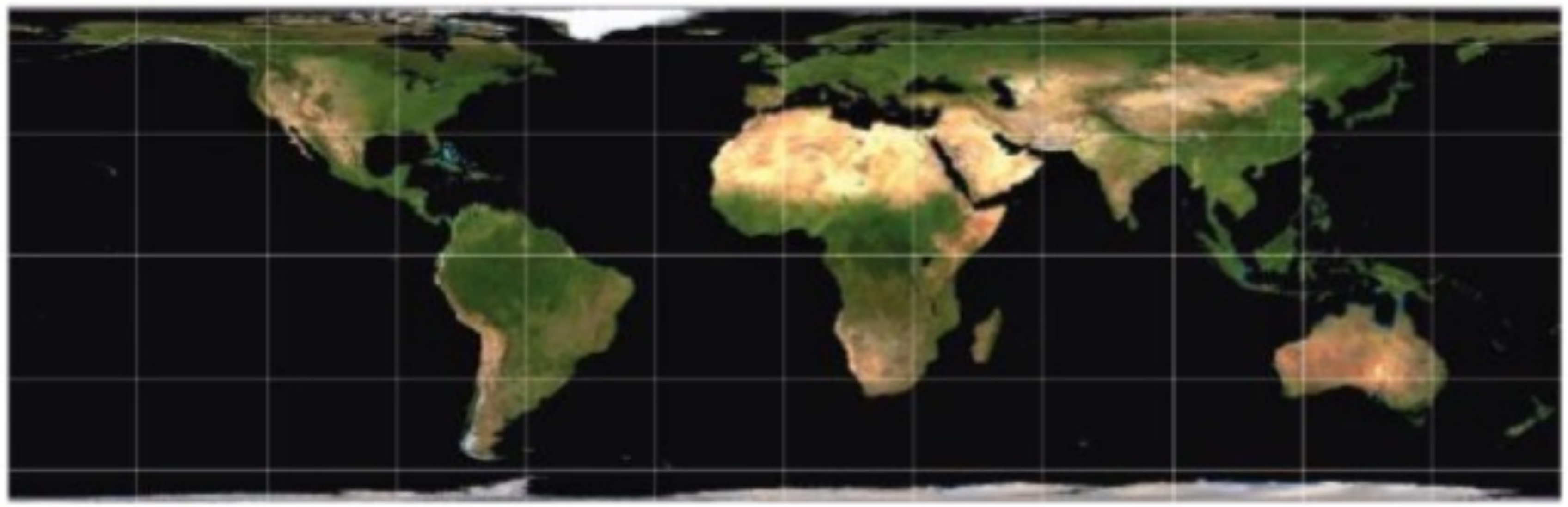
$$x = \lambda,\ y = \varphi.$$



Equirectangular projection.

- **It maps meridians to equally spaced vertical straight lines and circles of latitude to evenly spread horizontal straight lines. The projection does not have any of the desirable map properties and is neither conformal nor equal area. It has little use in navigation, but finds its main usage in thematic mapping.**

Map Projections: Lambert cylindrical projection



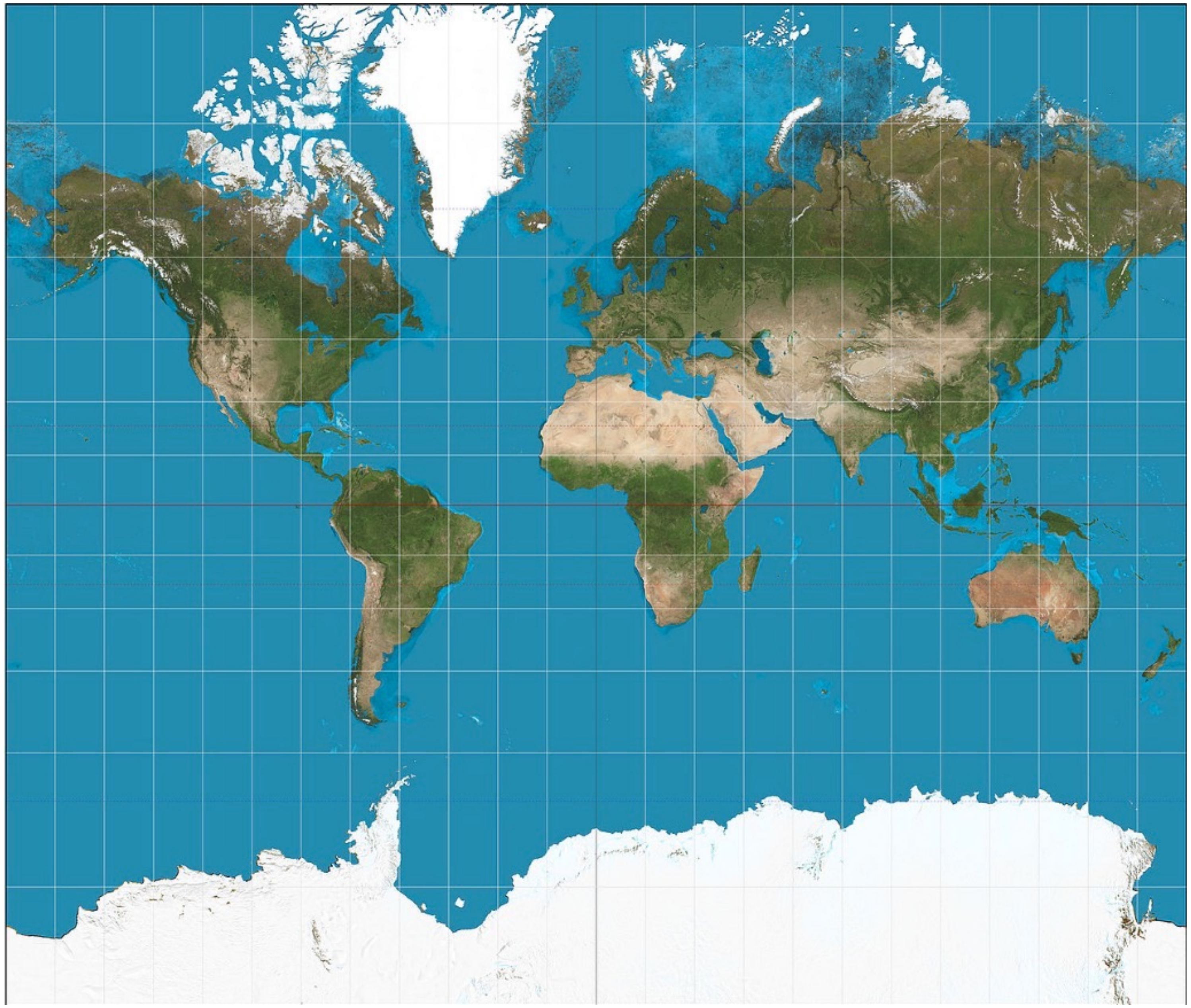
Lambert cylindrical projection.

- **Is an equal area projection that is easy to compute and provides nice world maps.**

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

Map Projections: Mercator projection

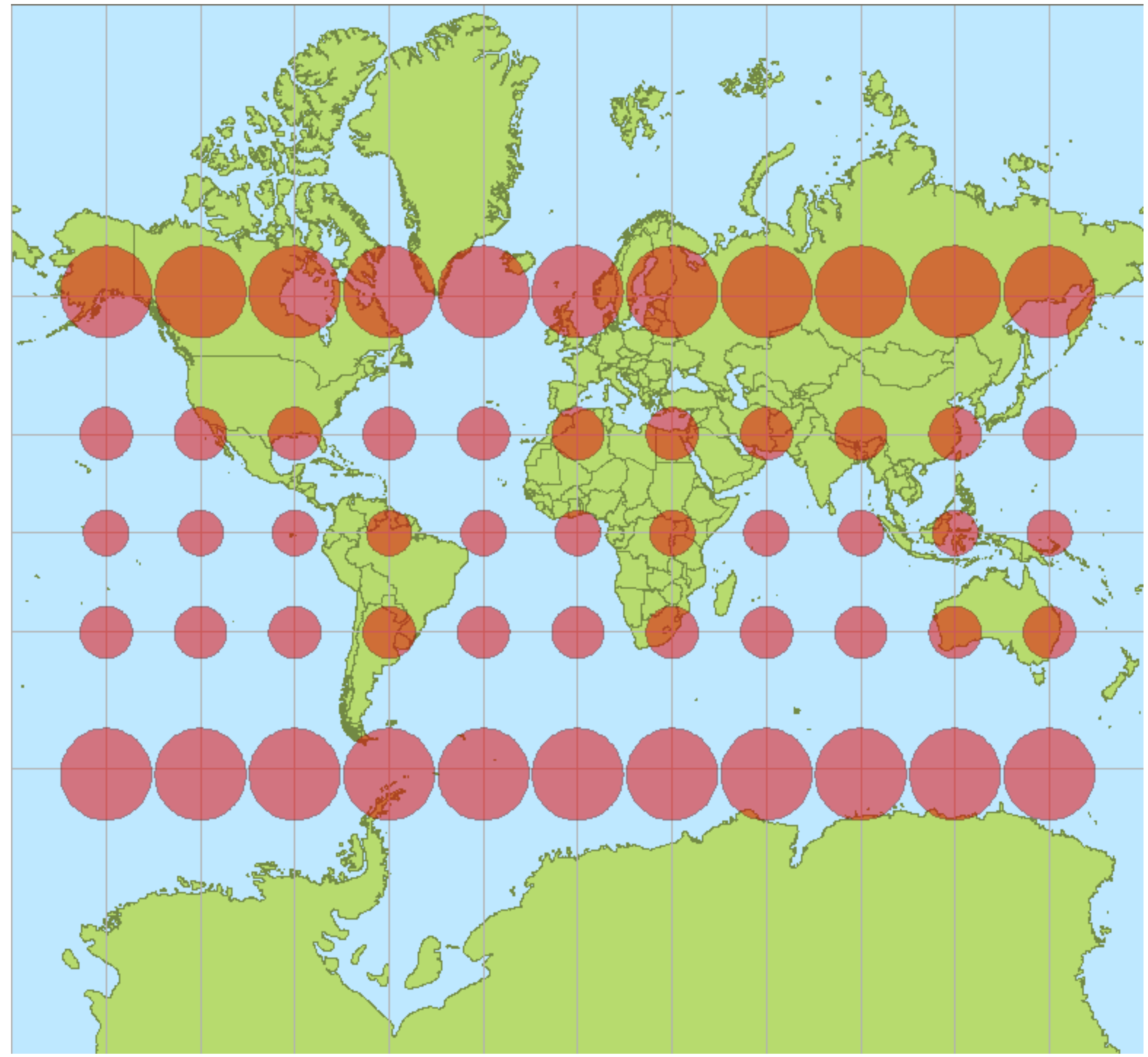
- is a cylindrical map projection.
- It became the standard map projection for nautical purposes because of its ability to represent lines of constant course.



FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE NOVA DE LISBOA

Map Projections: Mercator projection

- is a **cylindrical map projection**.
- It became the standard map projection for **nautical purposes** because of its ability to represent lines of constant course.



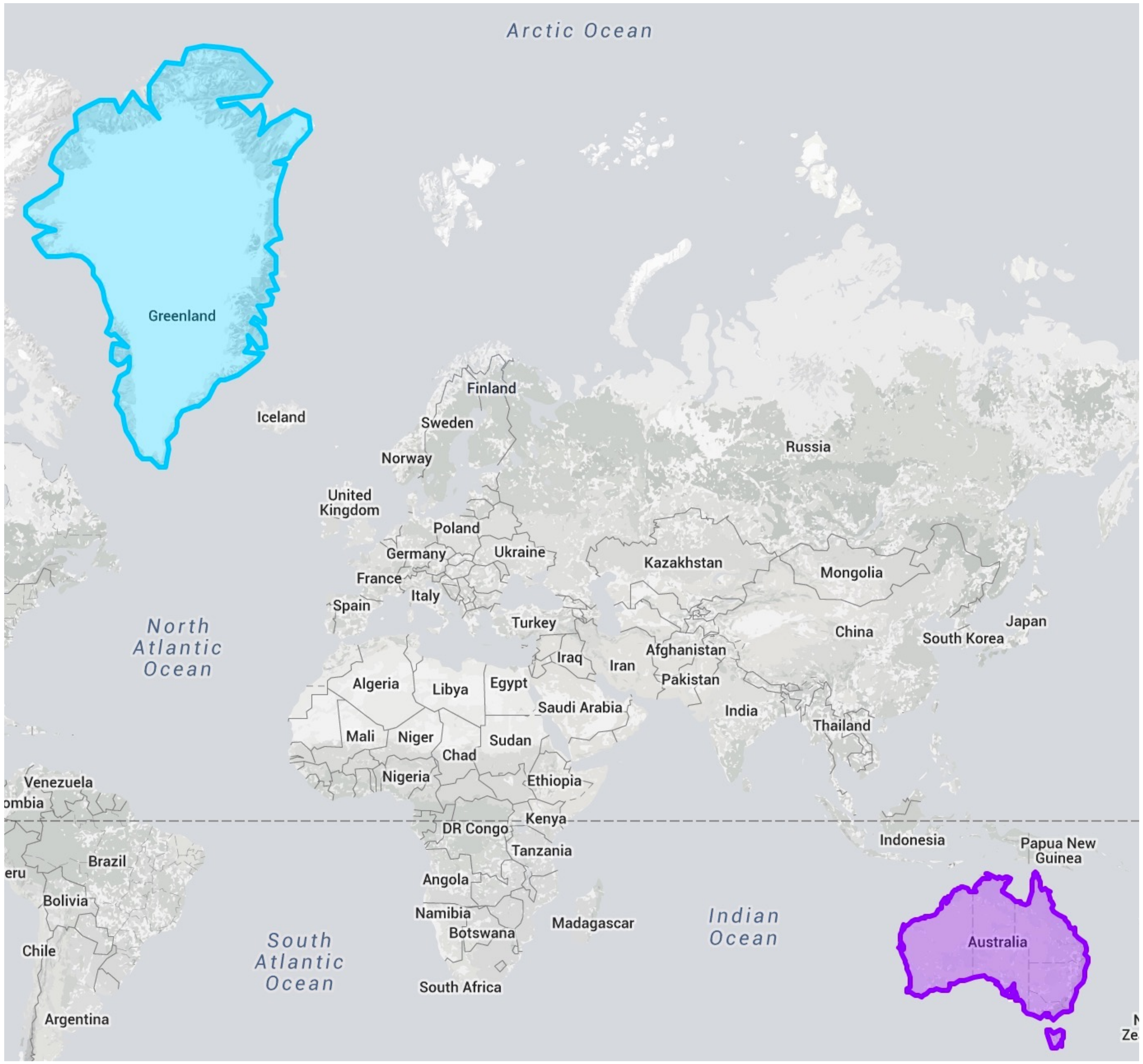
Map Projections: Mercator projection

- is a **cylindrical map projection**.
- **It became the standard map projection for nautical purposes because of its ability to represent lines of constant course.**

Google Maps uses a close variant of the Mercator projection, and therefore cannot accurately show areas around the poles.

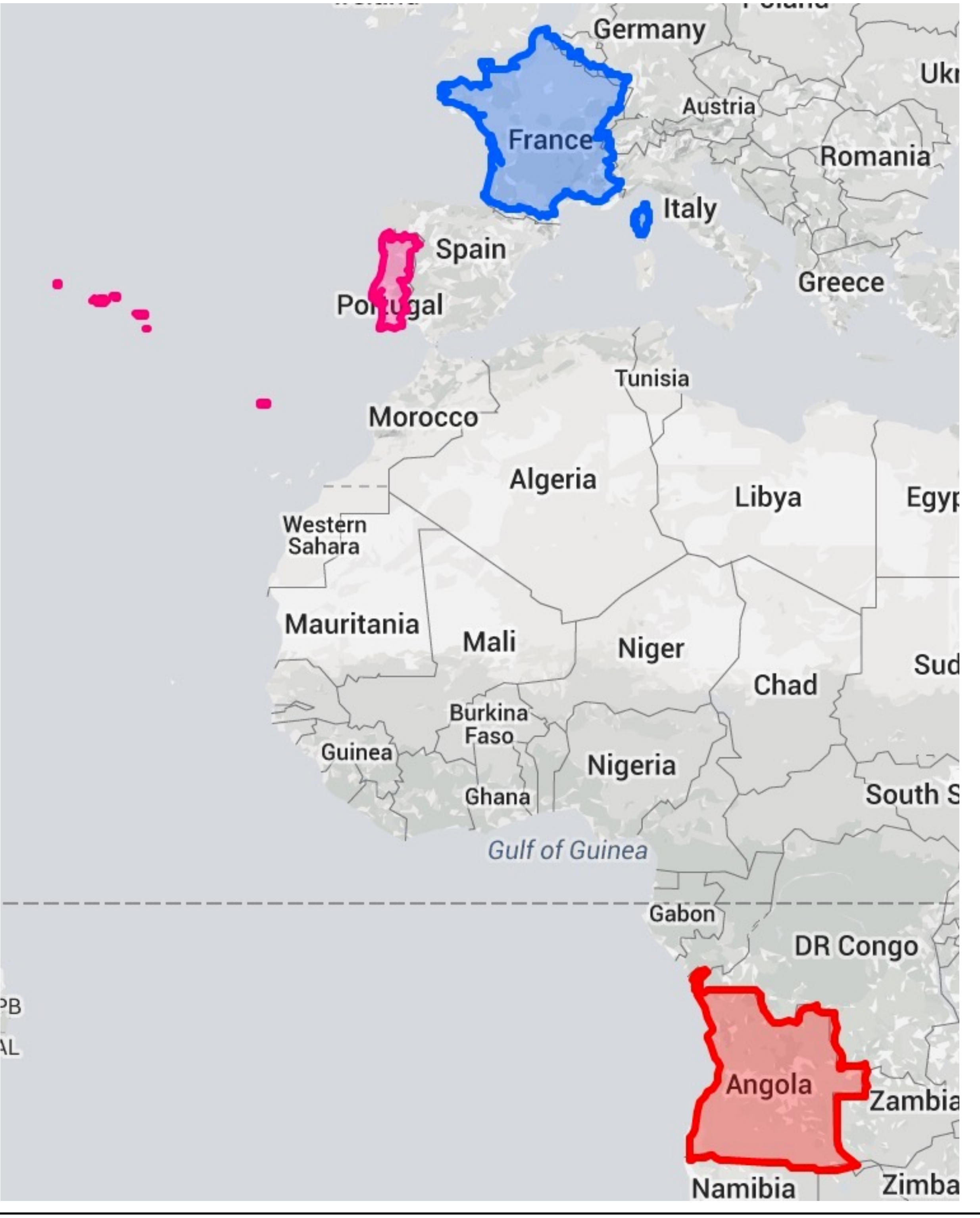
The Mercator projection portrays Greenland as larger than Australia; in actuality, Australia is more than three and a half times larger than Greenland.

Map Projections: Mercator projection

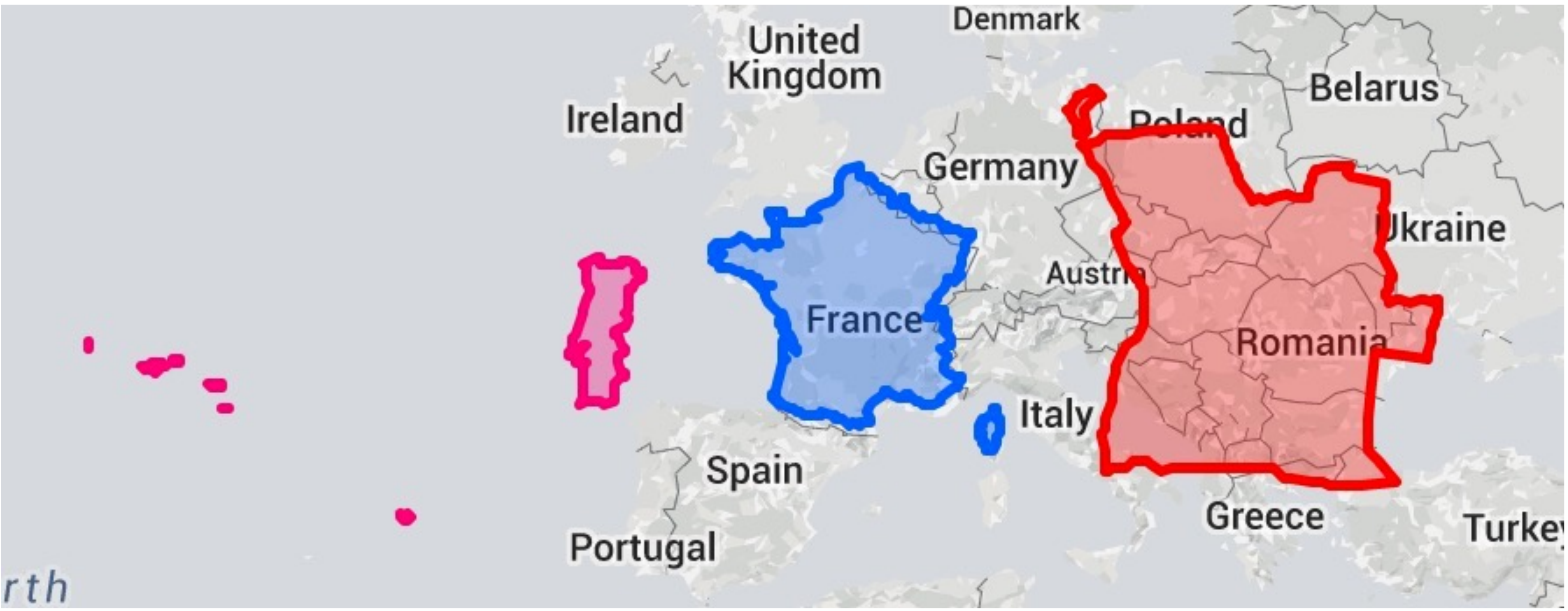


Map Projections: Mercator projection

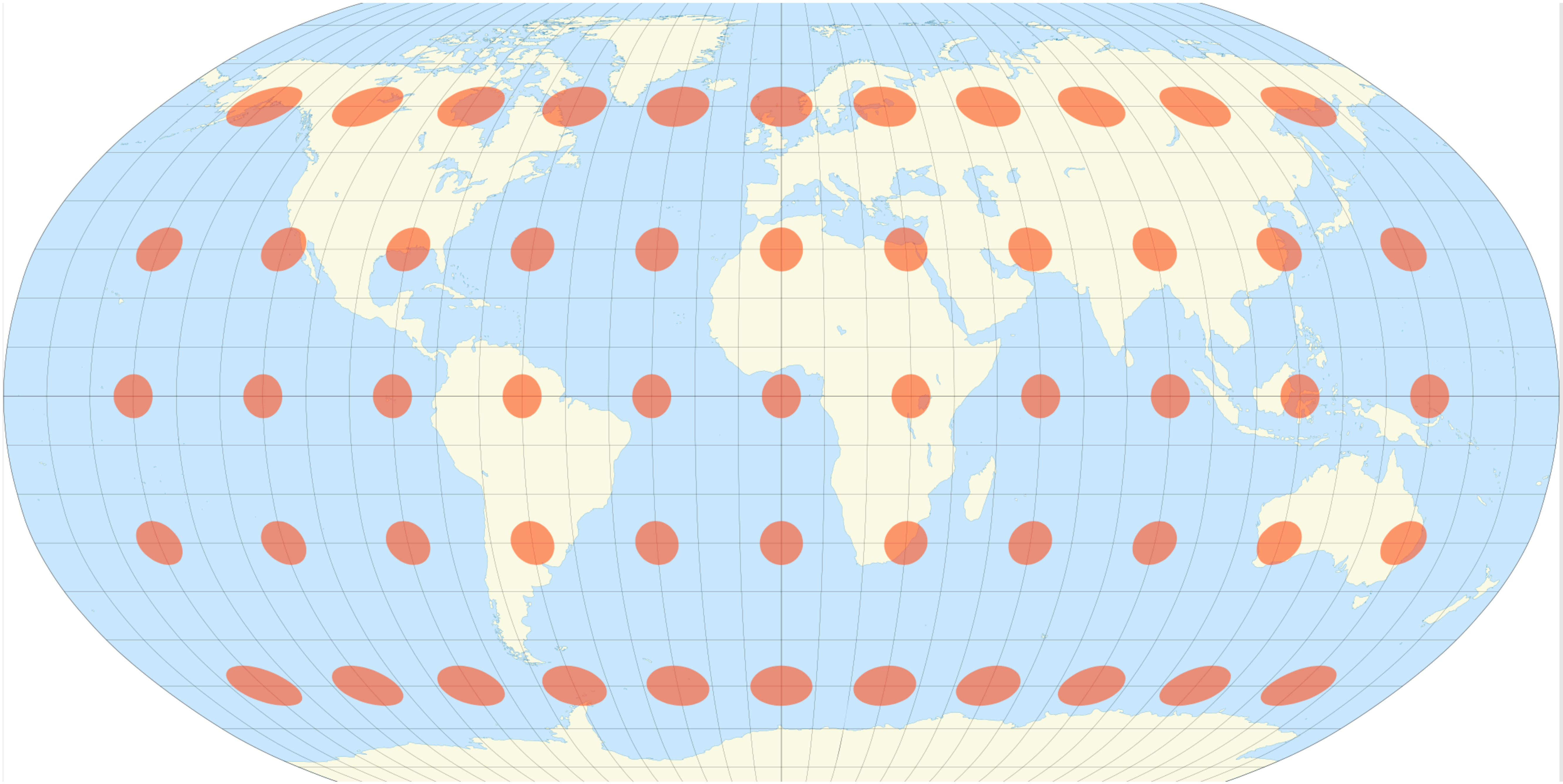
Map Projections: Mercator projection



Map Projections: Mercator projection

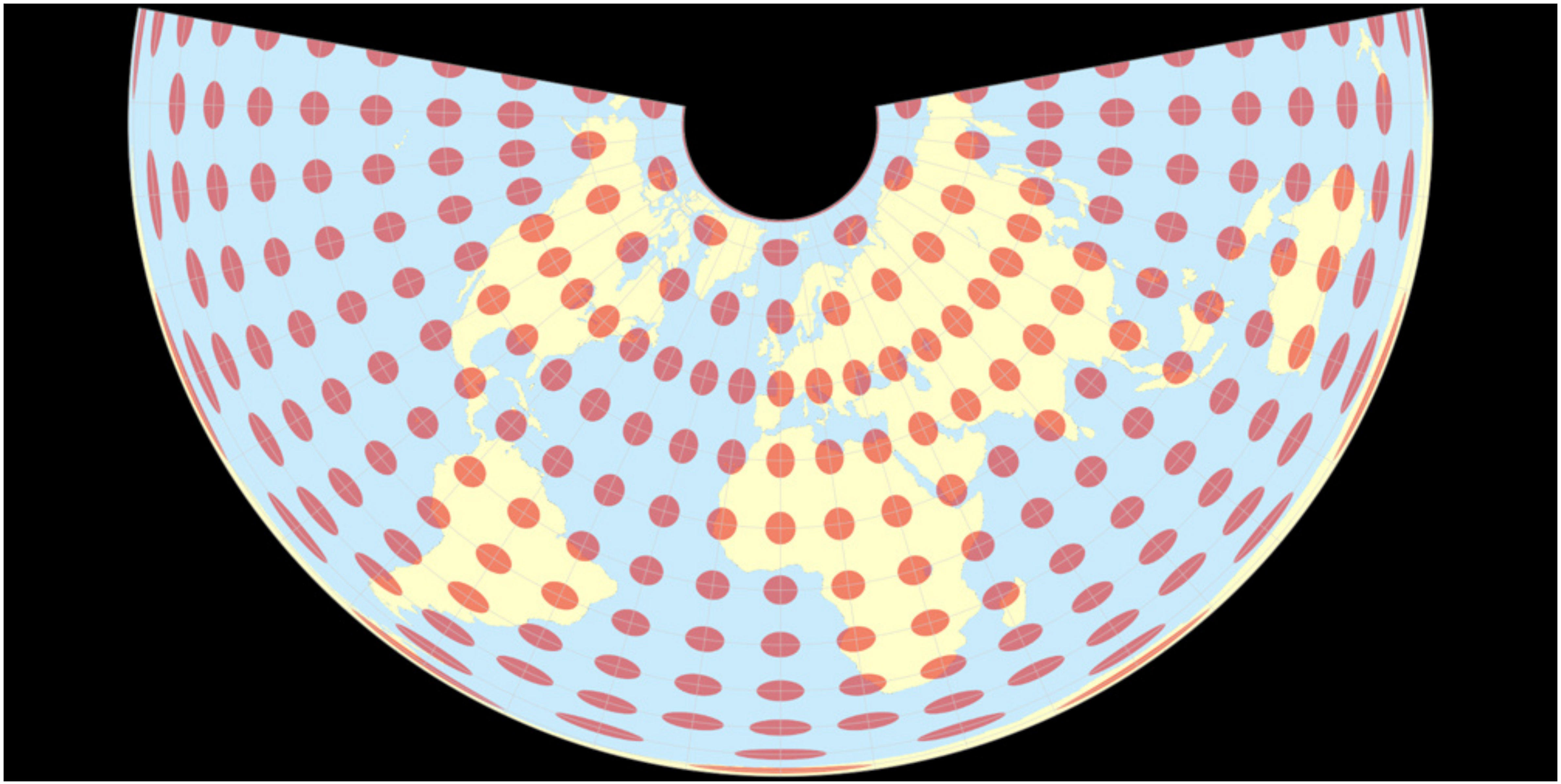


Robinson Projection

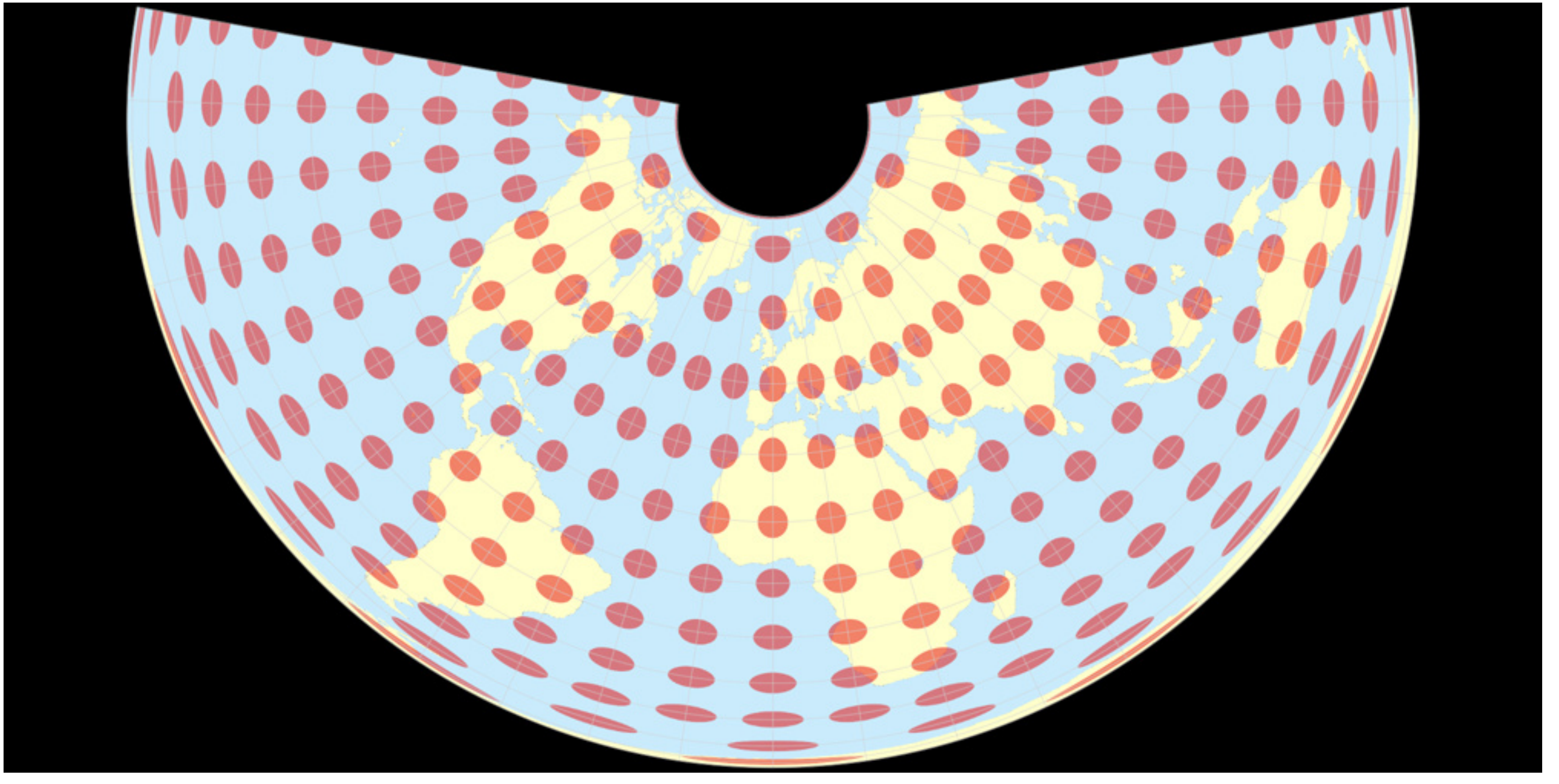


FACULTADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Albers Map Projection



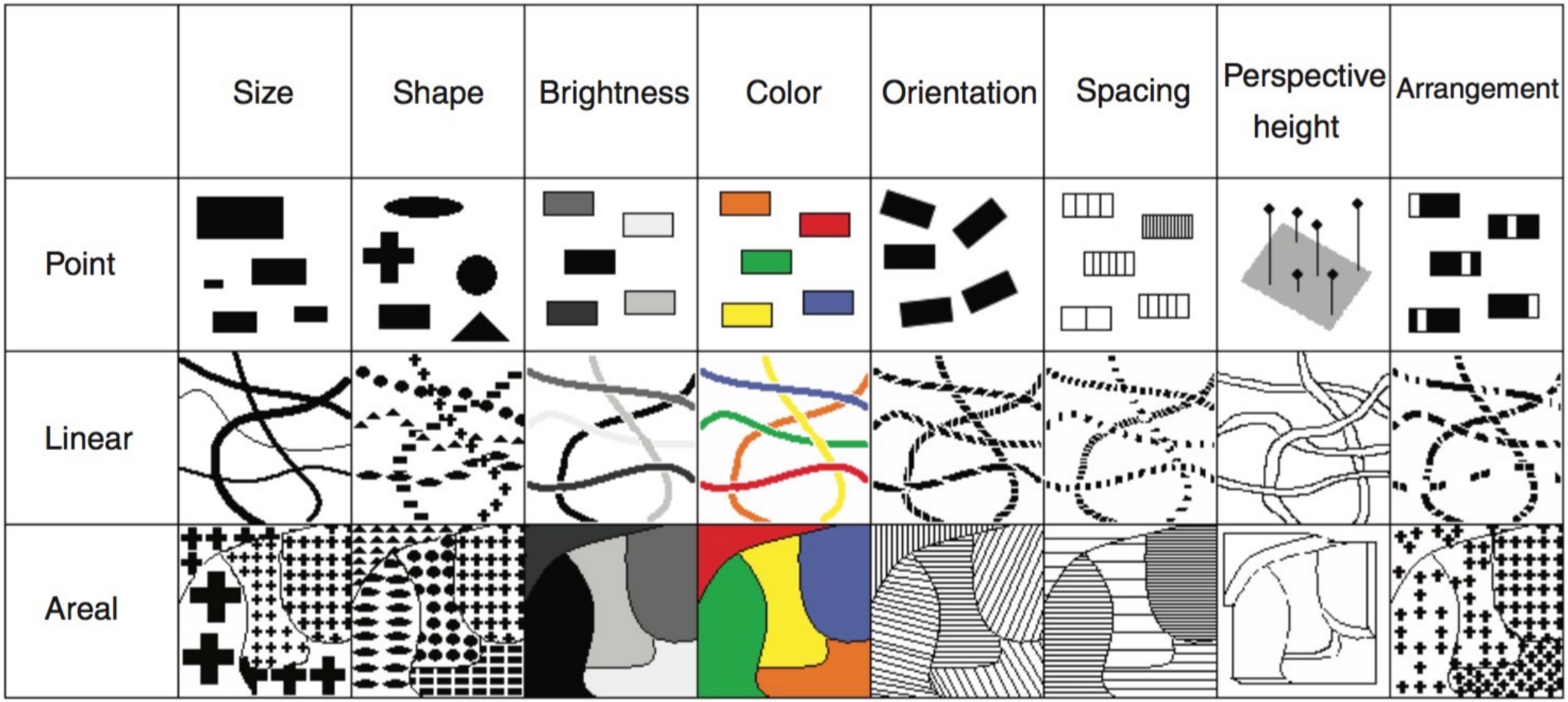
Azimutal Equidistant Polar Projection



Visual Variables for **Spatial Data**

- **size**: size of individual symbols, width of lines, or size of symbols in areas;
- **shape**: shape of individual symbols or pattern symbols in lines and areas;
- **brightness**: brightness of symbols, lines, or areas;
- **color**: color of symbols, lines, or areas;
- **orientation**: orientation of individual symbols or patterns in lines and areas;
- **spacing** (texture): spacing of patterns in symbols, lines, or areas;
- **perspective height**: perspective three-dimensional view of the phenomena with the data value mapped to the perspective height of points, lines, or areas;
- **arrangement**: arrangement of patterns within the individual symbols (for point phenomena), patterns of dots and dashes (for line phenomena), or regular versus random distribution of symbols (for area phenomena).

Visual Variables for Spatial Data

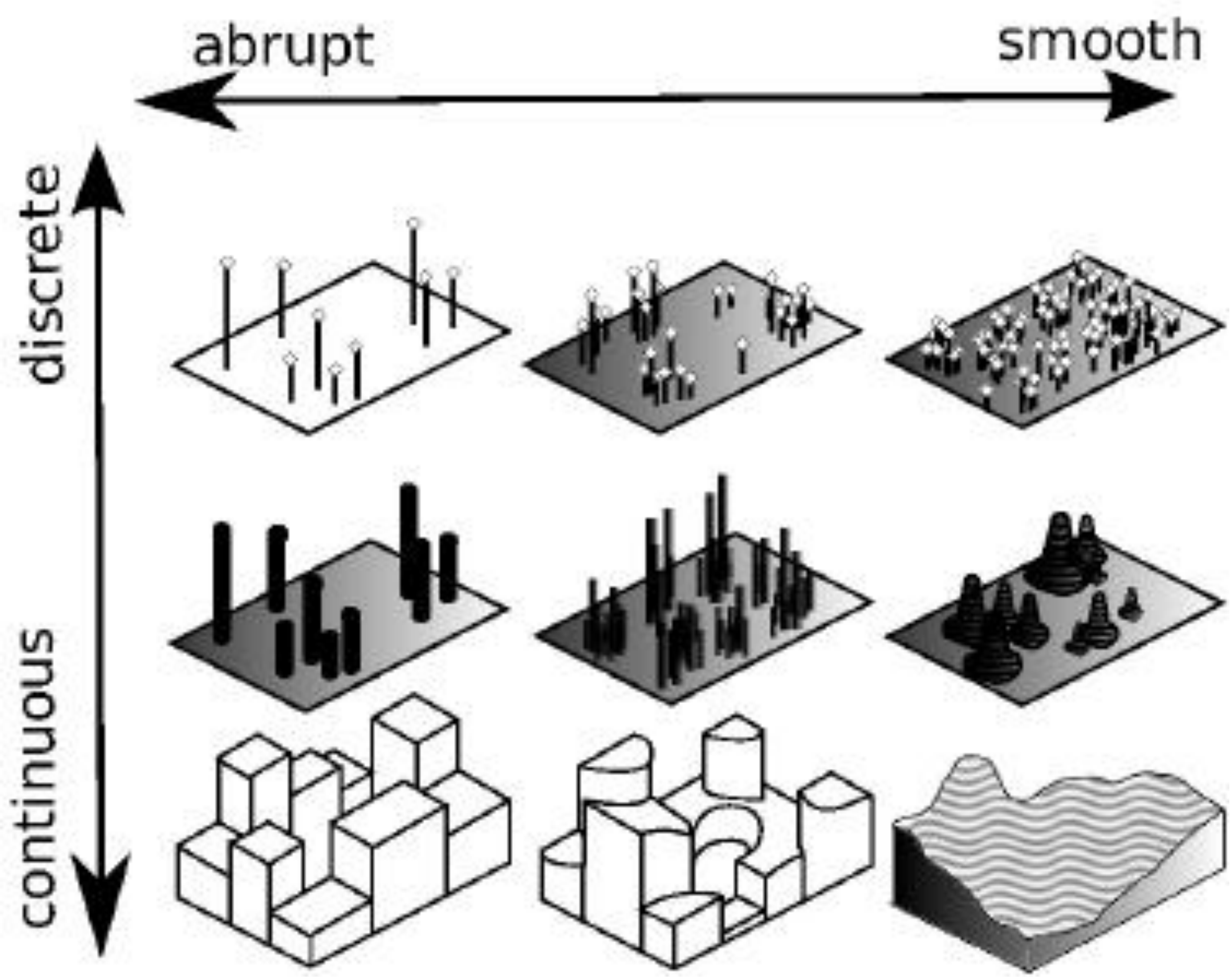


Visual variables for spatial data (Image based on [384].)

Visualization of Point Data

Visualization of Point Data

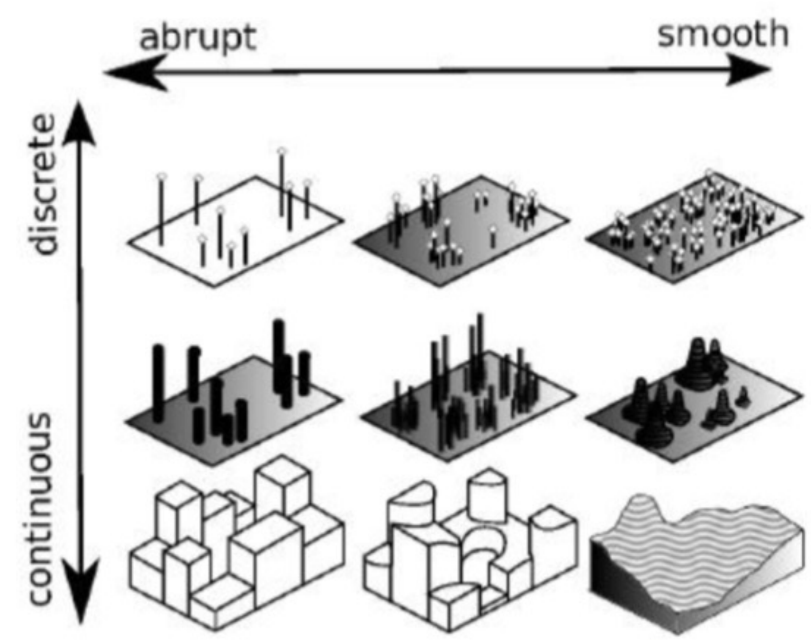
- **Point data are discrete in nature, but they may describe a continuous phenomenon.**



Discrete versus continuous and smooth versus abrupt (based on [281]).

Visualization of Point Data

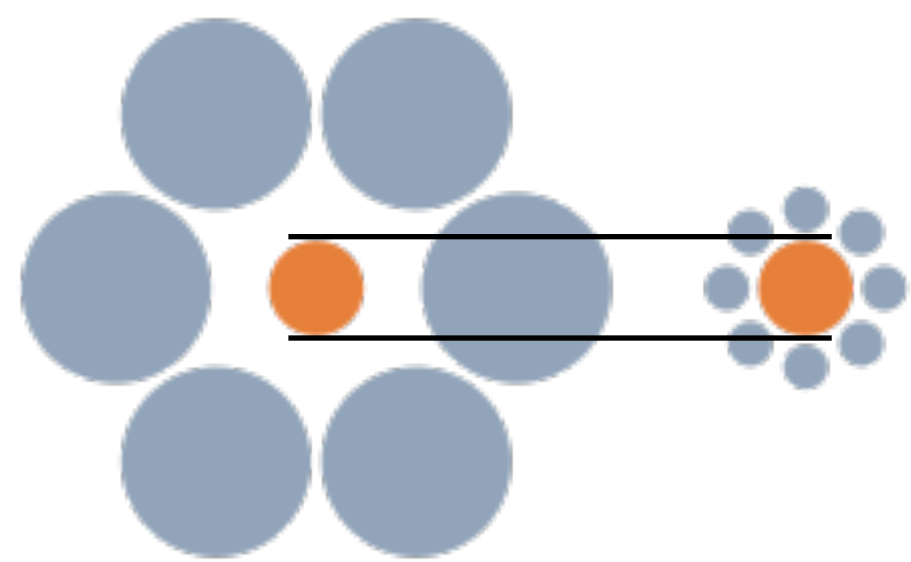
- **Point data are discrete in nature, but they may describe a continuous phenomenon.**
  - **Discrete data are presumed to occur at distinct locations, while continuous data are defined at all locations**
  - **Smooth data refers to data that change in a gradual fashion, while abrupt data change suddenly.**



Discrete versus continuous and smooth versus abrupt (based on [281]).

Dot Maps

- Point phenomena can be visualized by placing a symbol or pixel at the location where that phenomenon occurs.
- A quantitative parameter may be mapped to the size or the color of the symbol or pixel.
  - Symbols: circles are the most widely used symbol in dot maps, but squares, bars, or any other symbol can be used as well
  - Size: calculating the correct size of the symbols does not necessarily mean that the symbols will be perceived correctly.
    The perceived size of the symbols depends on their local neighborhood
    (e.g., the Ebbinghaus illusion)



  - Color: take into account the problems of color

Dot Maps

- **When large data sets are drawn on a map, the problem of overlap or overplotting of data points arises in highly populated areas, while low-population areas are virtually empty.**

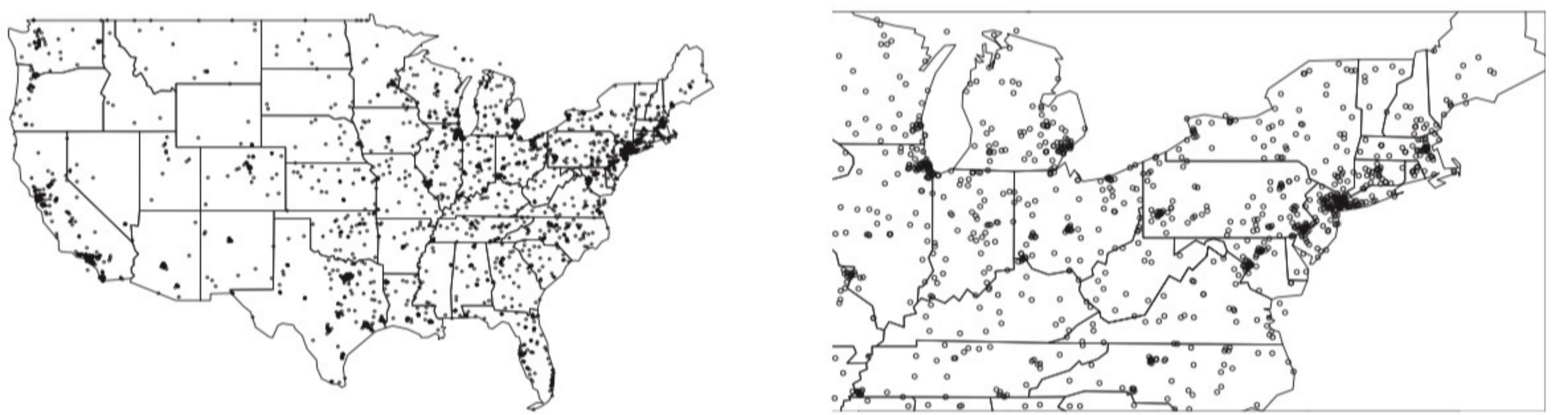
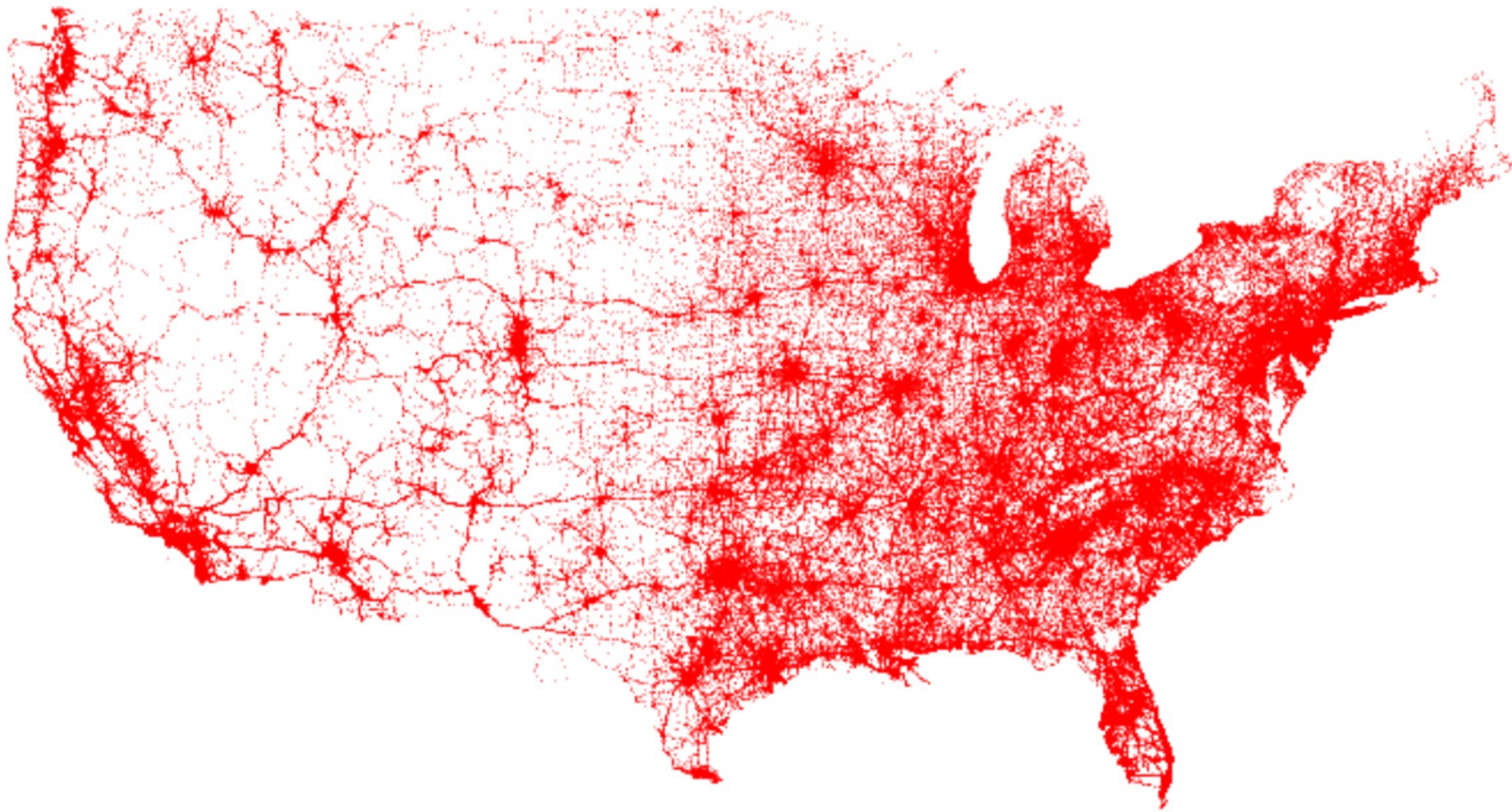


Figure 6.12. USA dot map: every circle represents the spatial location of an event. Even in the zoomed-in version there is a large degree of overlap. (Image reprinted from [227] with permission of Springer Science and Business Media.)

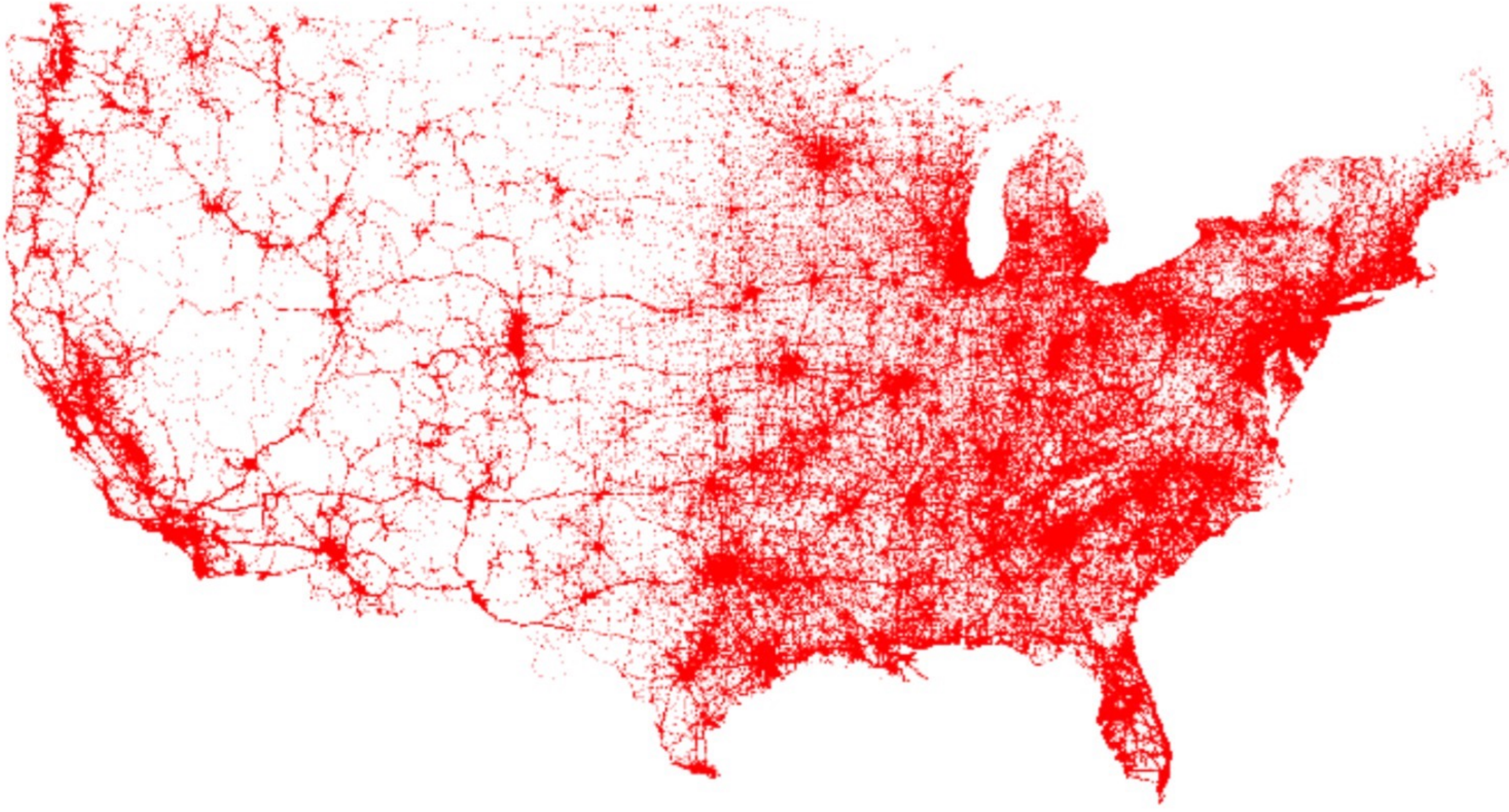
Eight visual variables: Screen resolution

- **450.710 geo-referenced accidents between 2001 and 2013 in US**



Eight visual variables: Screen resolution

- **Preprocessed data: 53% of items from original data set**



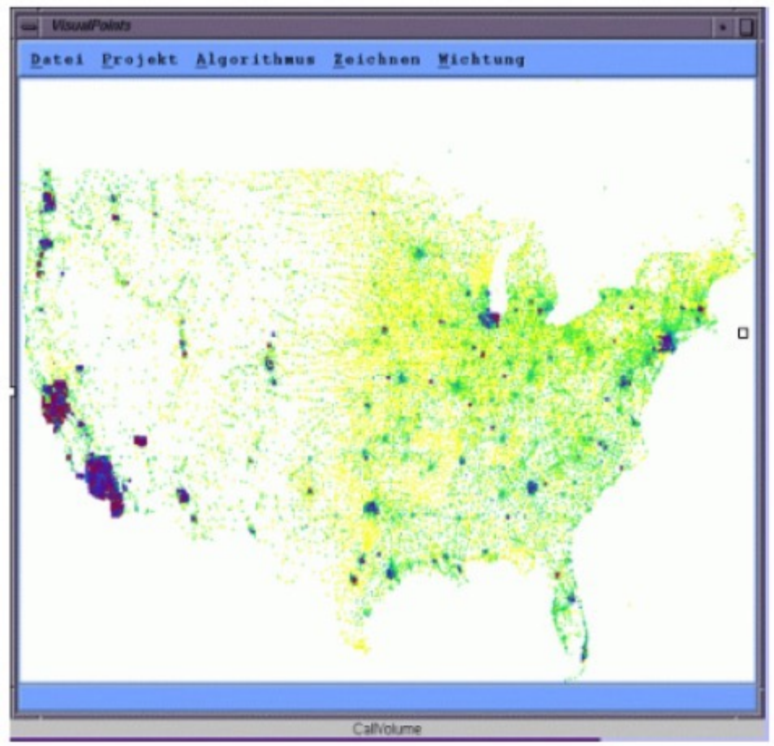
Dot Maps

- **When large data sets are drawn on a map, the problem of overlap or overplotting of data points arises in highly populated areas, while low-population areas are virtually empty**
  - **Spatial data are highly non-uniformly distributed in real-world data sets.**
    - **Credit card payments, telephone calls, health statistics, environmental records, crime data, and census demographics, …, etc..**
- **Approaches for coping with dense spatial data**
  - **2.5D visualization showing data points aggregated up to map regions**
  - **Individual data points as bars, according to their statistical value on a map**

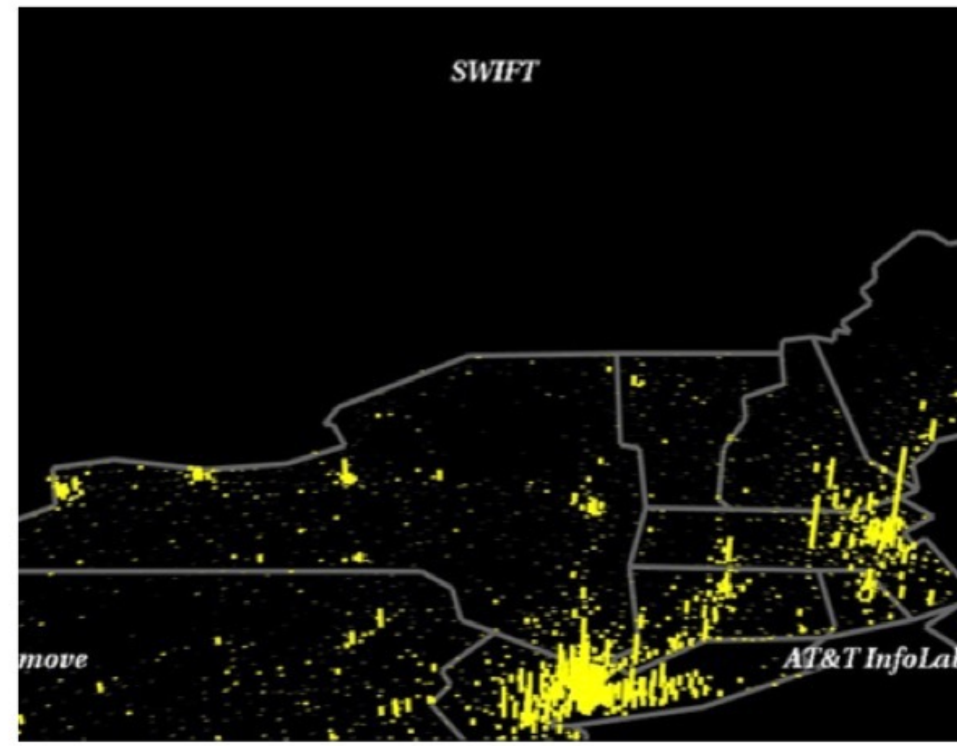
Dot Maps

- **Approaches for coping with dense spatial data**

(a) *Traditional 2D Map* - with overlap



(b) *Non-overlap 2D Map (Gridfit)* - repositioning depends on the ordering of the points in the database

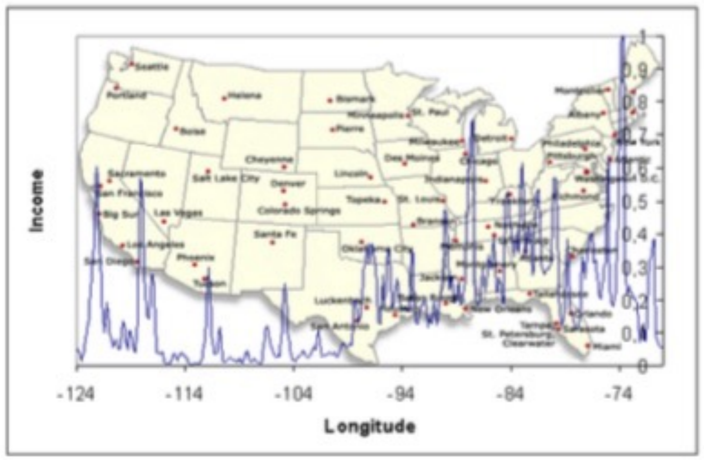


(c) *2.5D Bar Map (Swift)* - too many data points are plotted at the same position, and therefore only a small portion of the data is actually displayed

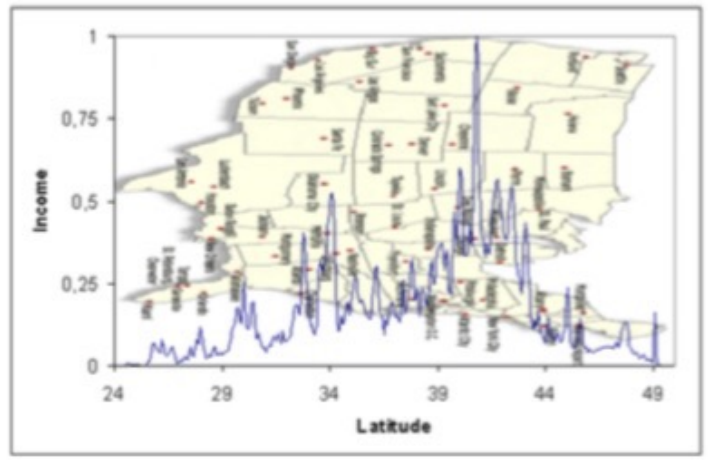
Geo-Spatial Data Viewer: From Familiar Land-covering to Arbitrary Distorted Geo-Spatial Quadtree Maps
Daniel A. Keim, Christian Panse, Jorn Schneidewind, Mike Sips

Dot Maps

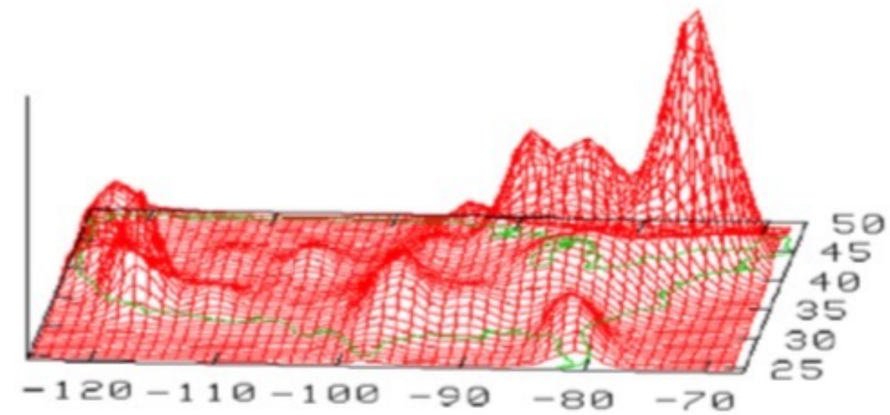
- **Approaches for coping with dense spatial data**



(a) *2D Average Household Income Plot (longitude, median household income)* - The two highest average household income areas (Atlantic Coast and Pacific Coast) regions have up to $100.000 U.S. median household income; the two lowest average household income regions are the New England and Rocky Mountain regions



(b) *2D Average Household Income Plot (latitude, median household income)* - The only significant household income for the United States is in the middle latitude region

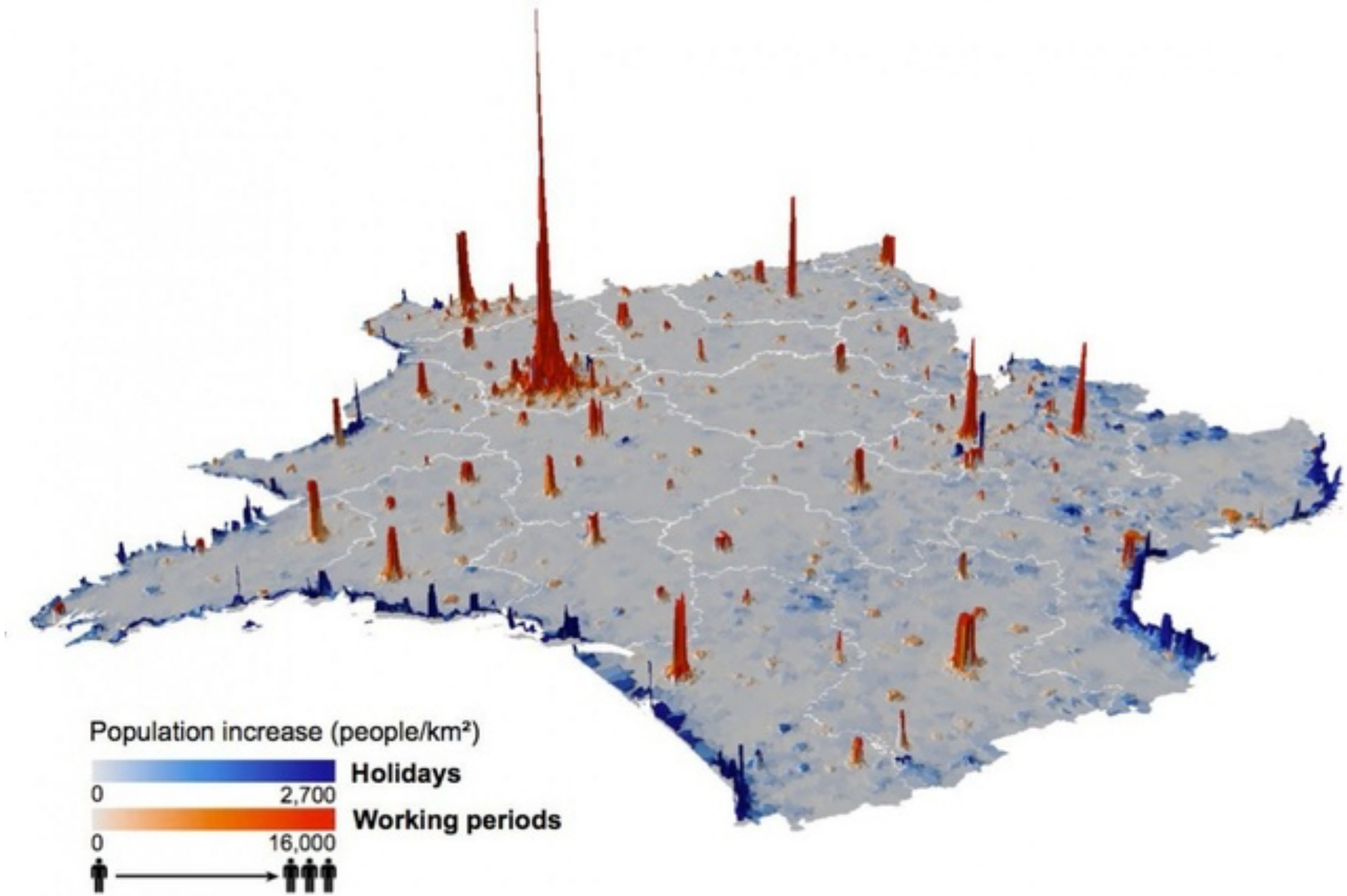


(c) *3D Median Average Income Plot (longitude, latitude, median household income)* - Yields a good separation of household income with respect to six cities that are identified

Geo-Spatial Data Viewer: From Familiar Land-covering to Arbitrary Distorted Geo-Spatial Quadtree Maps
Daniel A. Keim, Christian Panse, Jorn Schneidewind, Mike Sips

Dot Maps

- **Approaches for coping with dense spatial data**



Dot Maps

- **Approaches for coping with dense spatial data**

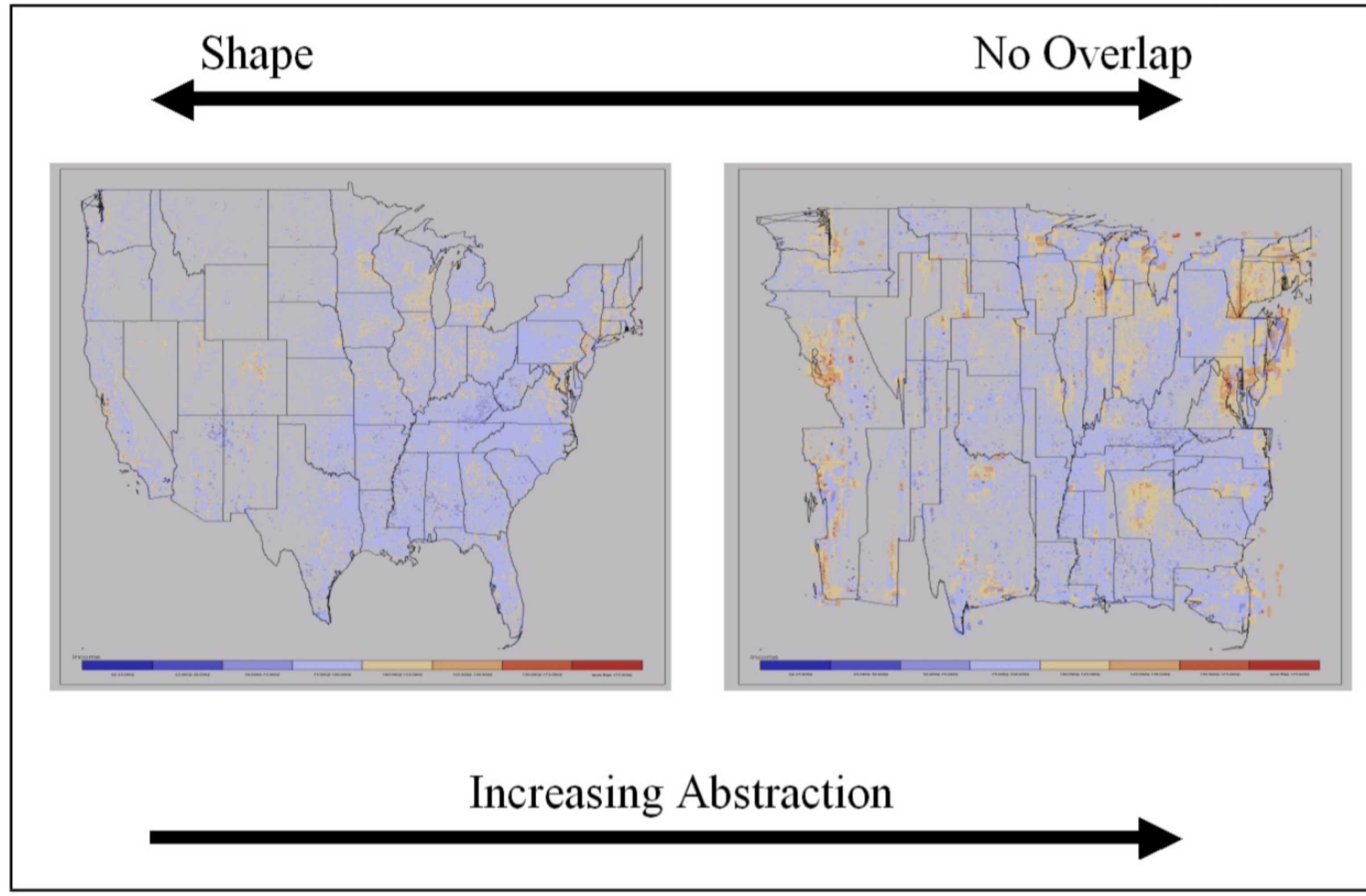


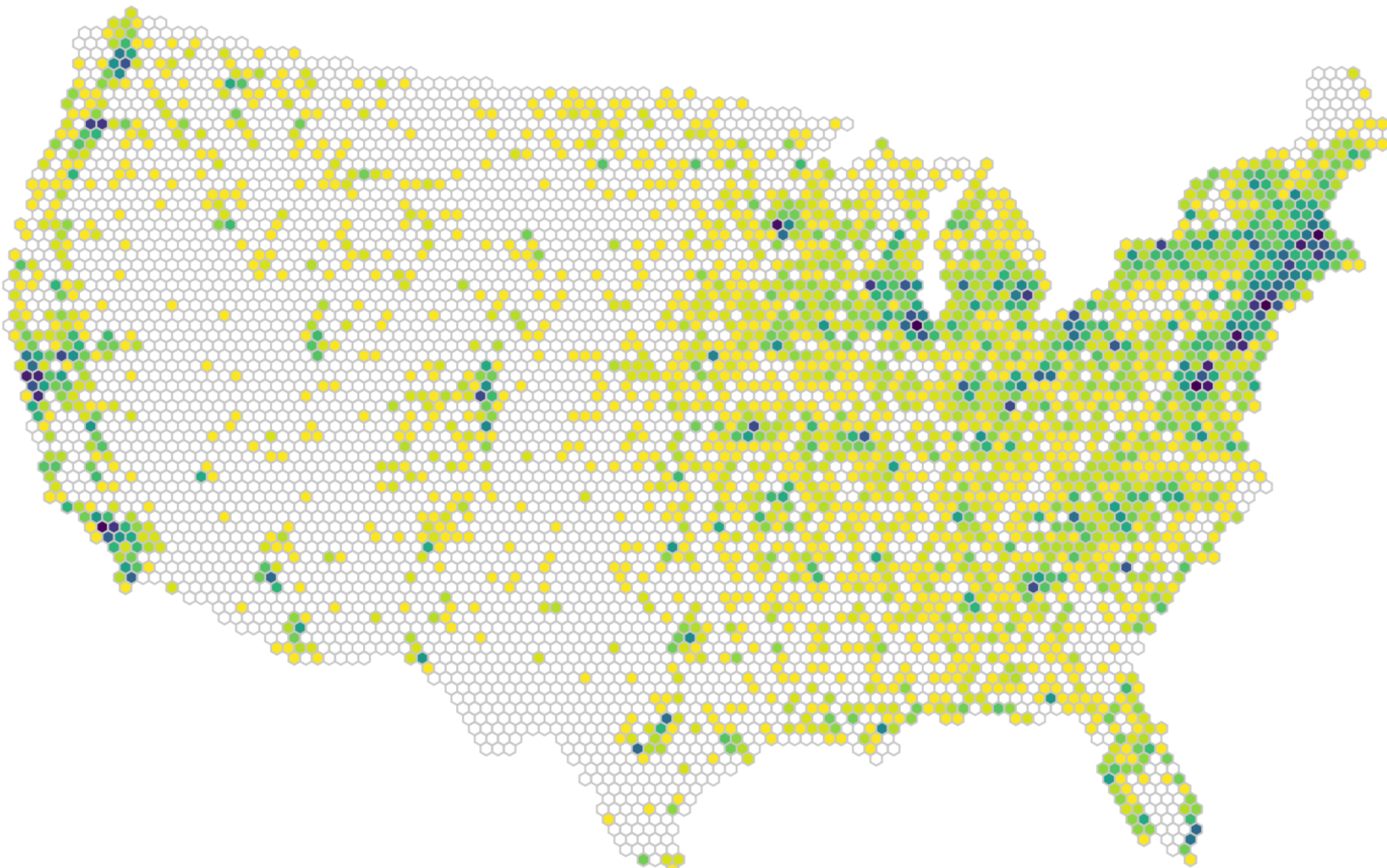
Figure 1: *Tradeoff between Shape and Overlap Factor* – US-Year 2000 Census Median Household Income.

Geo-Spatial Data Viewer: From Familiar Land-covering to Arbitrary Distorted Geo-Spatial Quadtree Maps
Daniel A. Keim, Christian Panse, Jorn Schneidewind, Mike Sips

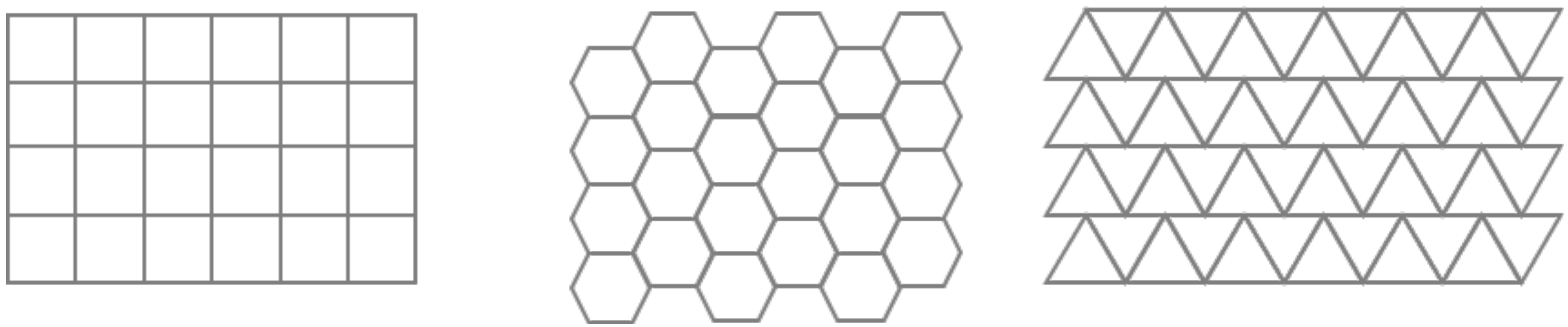
Density Maps and Hexabin Maps

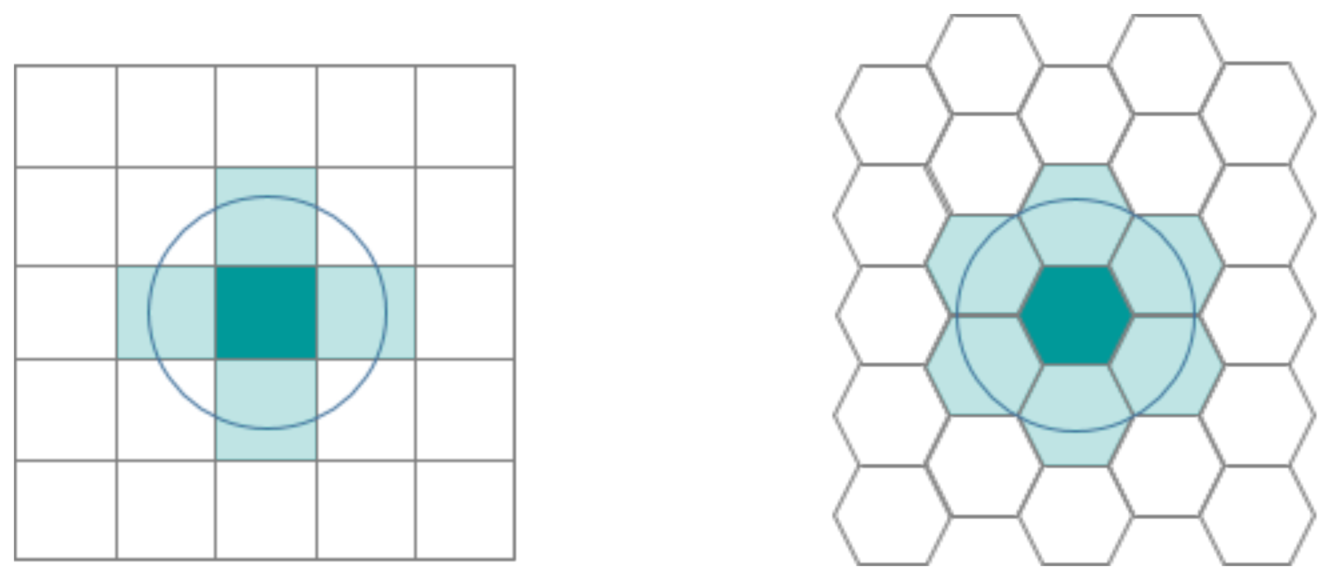
https://www.tableau.com/about/blog/2016/7/how-create-density-maps-using-hexbins-tableau-56511

Density Maps and Hexabin Maps



Why hexagons?





https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-whyhexagons.htm

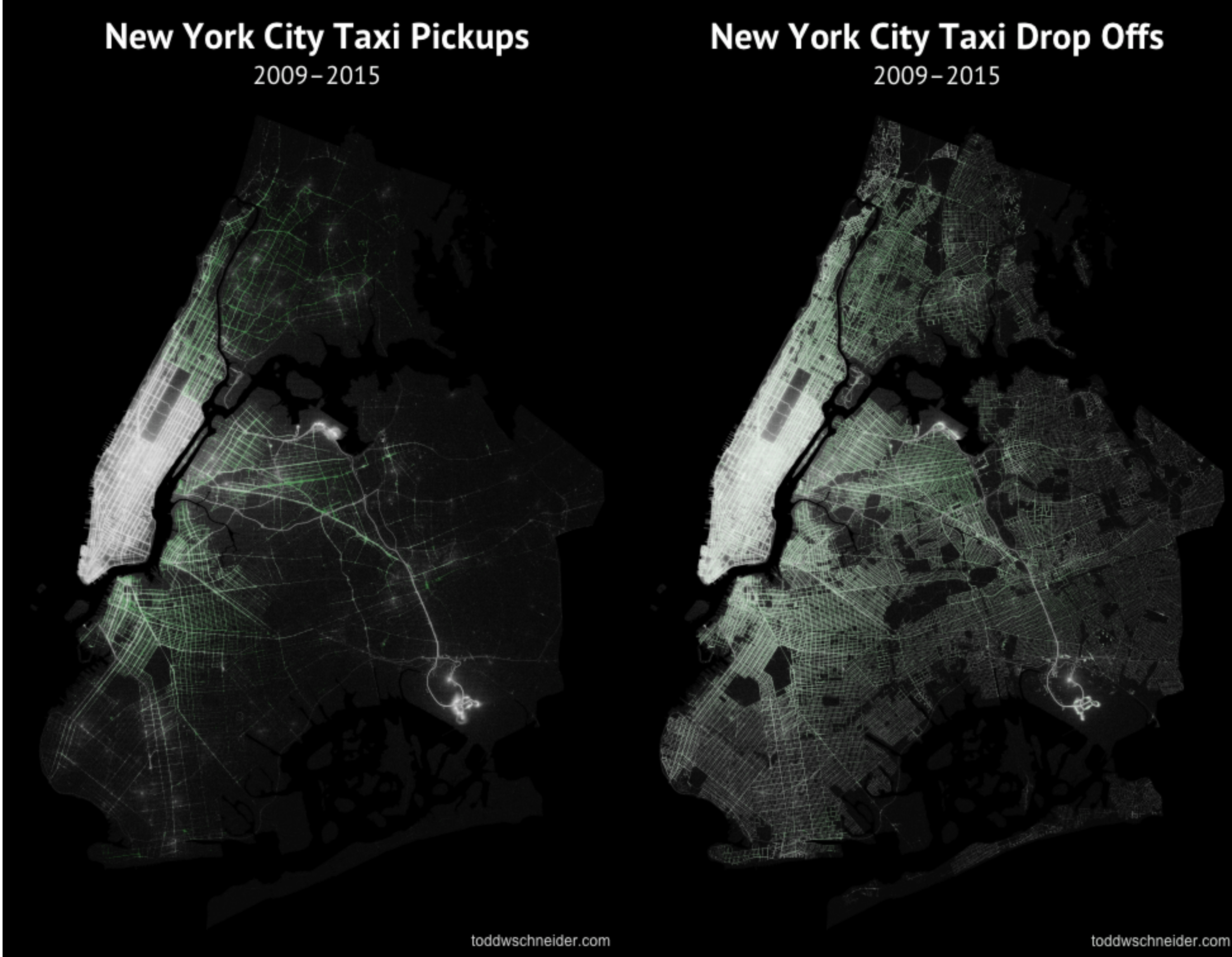
Why hexagons?

- Hexagons reduce sampling bias due to edge effects of the grid shape, this is related to the low perimeter-to-area ratio of the shape of the hexagon. A circle has the lowest ratio but cannot tessellate to form a continuous grid. **Hexagons are the most circular-shaped polygon that can tessellate to form an evenly spaced grid**.
- Due to the linear nature of rectangles, fishnet grids can draw our eyes to the straight, unbroken, parallel lines which may inhibit the underlying patterns in the data. Hexagons tend to break up the lines and allow any curvature of the patterns in the data to be seen more clearly and easily. This breakup of artificial linear patterns also diminishes any orientation bias that can be perceived in fishnet grids.

https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/h-whyhexagons.htm

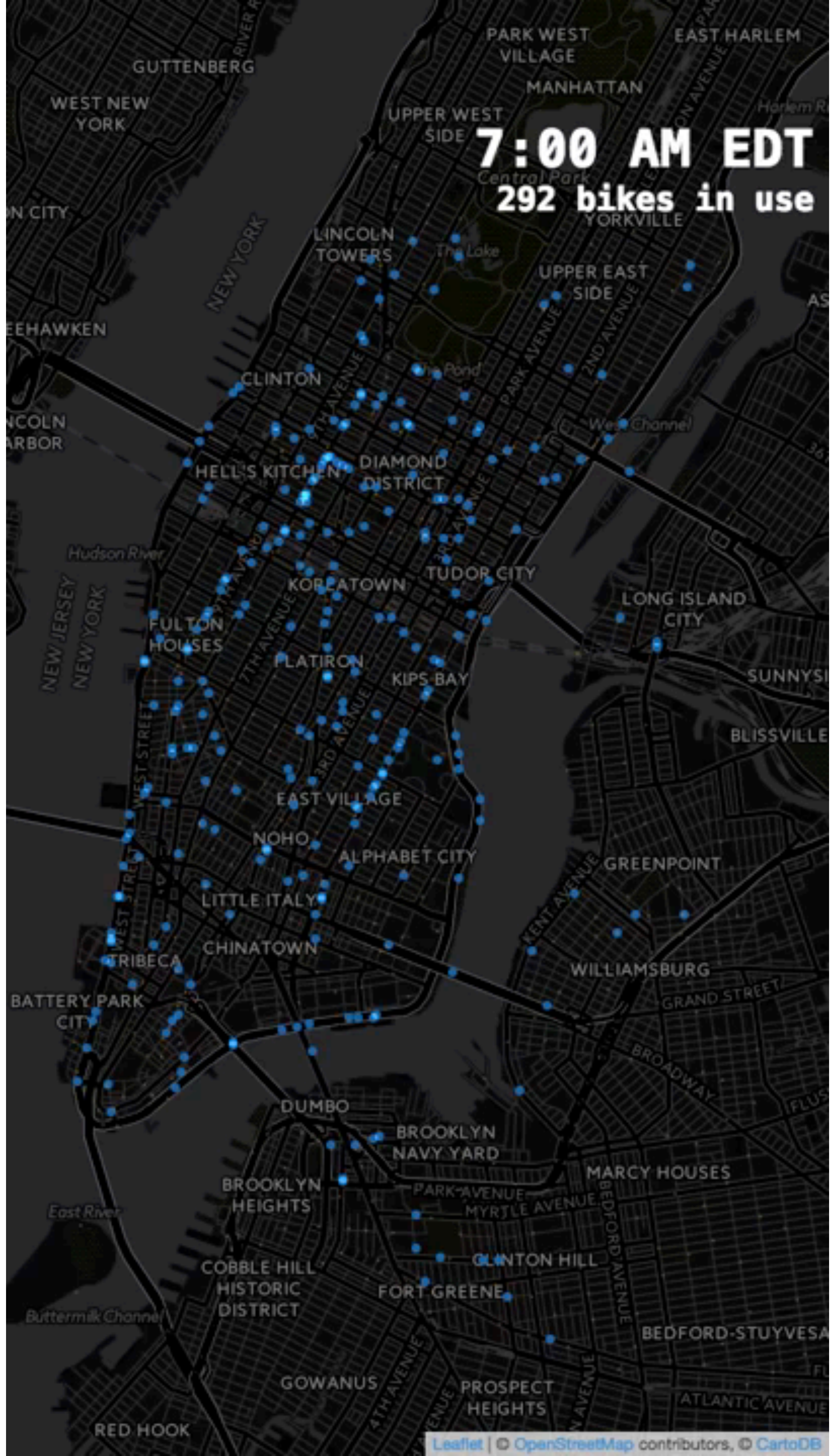
NYC Taxi and Uber Trips

- Analyzing 1.1 Billion NYC Taxi and Uber Trips, with a Vengeance (LINK)


New York City Taxi Pickups
2009–2015
toddwschneider.com
New York City Taxi Drop Offs
2009–2015
toddwschneider.com

NYC Bike Share System

- A Tale of Twenty-Two Million Citi Bikes: Analyzing the NYC Bike Share System (LINK)


7:00 AM EDT
292 bikes in use
PARK WEST VILLAGE
EAST HARLEM
GUTTENBERG
MANHATTAN
WEST NEW YORK
UPPER WEST SIDE
Central Park
YORKVILLE
LINCOLN TOWERS
UPPER EAST SIDE
WEEHAWKEN
CLINTON
HELL'S KITCHEN
DIAMOND DISTRICT
Hudson River
KOREATOWN
TUDOR CITY
LONG ISLAND CITY
FULTON HOUSES
FLATIRON
KIPS BAY
SUNNYSIDE
BLISSVILLE
EAST VILLAGE
NOHO
ALPHABET CITY
GREENPOINT
LITTLE ITALY
CHINATOWN
TRIBECA
WILLIAMSBURG
BATTERY PARK CITY
GRAND STREET
BROADWAY
DUMBO
BROOKLYN NAVY YARD
MARCY HOUSES
BROOKLYN HEIGHTS
PARK AVENUE
MYRTLE AVENUE
East River
CLINTON HILL
COBBLE HILL HISTORIC DISTRICT
FORT GREENE
BEDFORD-STUYVESANT
Buttermilk Channel
GOWANUS
PROSPECT HEIGHTS
ATLANTIC AVENUE
RED HOOK
NEW JERSEY
NEW YORK
Leaflet | © OpenStreetMap contributors, © CartoDB

Visualization of Line Data

Visualization of Line Data

- **The basic idea for visualizing spatial data describing linear phenomena is to represent them as line segments between pairs of endpoints specified by longitude and latitude.**
- **A standard mapping of line data allows data parameters to be mapped to line width, line pattern, line color, and line labeling.**
- **In addition, data properties of the starting and ending points, as well as intersection points, may also be mapped to the visual parameters of the nodes, such as size, shape, color, and labeling.**
- **Lines do not need to be straight, but may be polylines or splines**

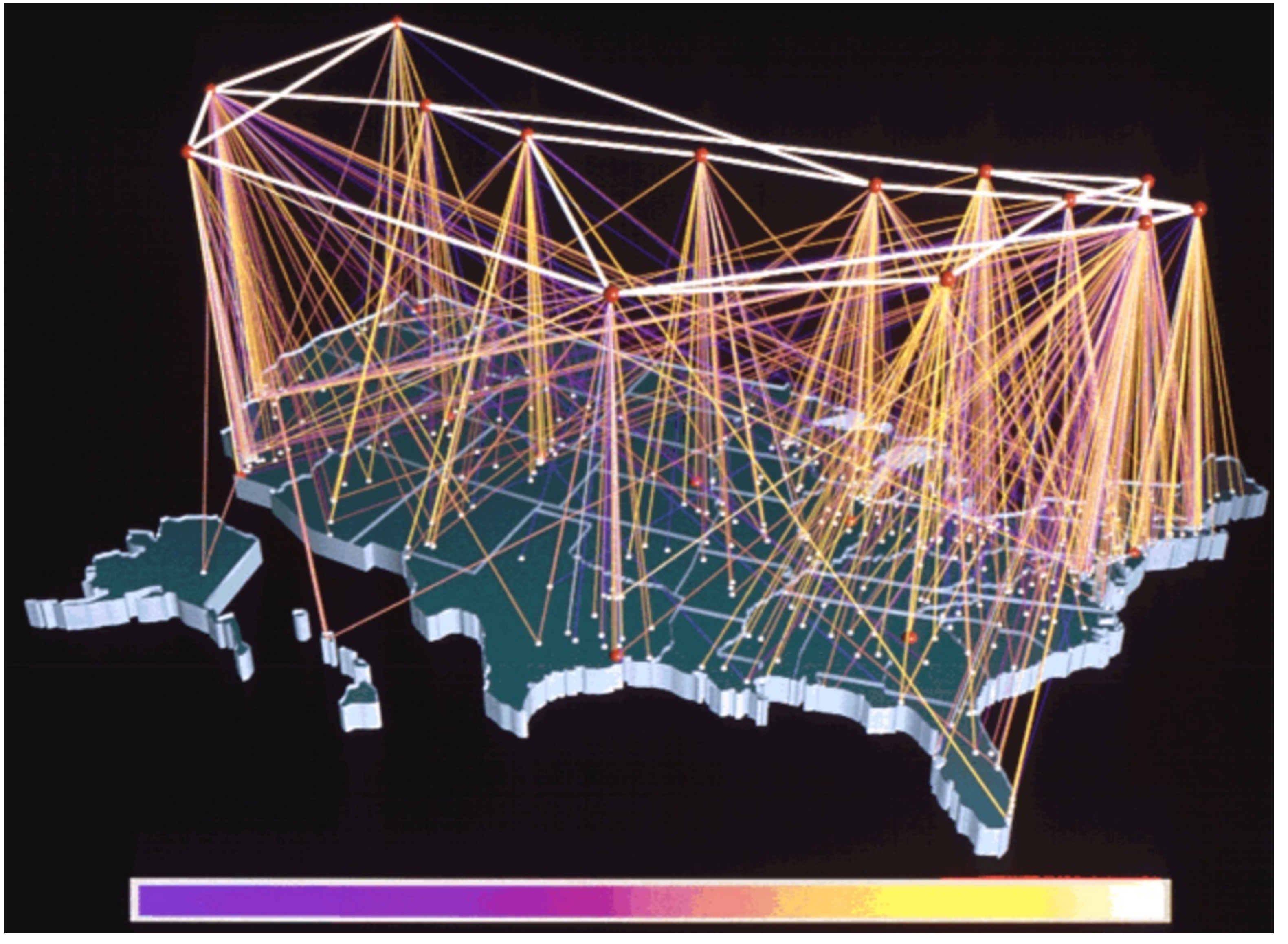
Visualization of Line Data

- **Network Maps**

Swift-3D. (Image from [254].)

Visualization of Line Data

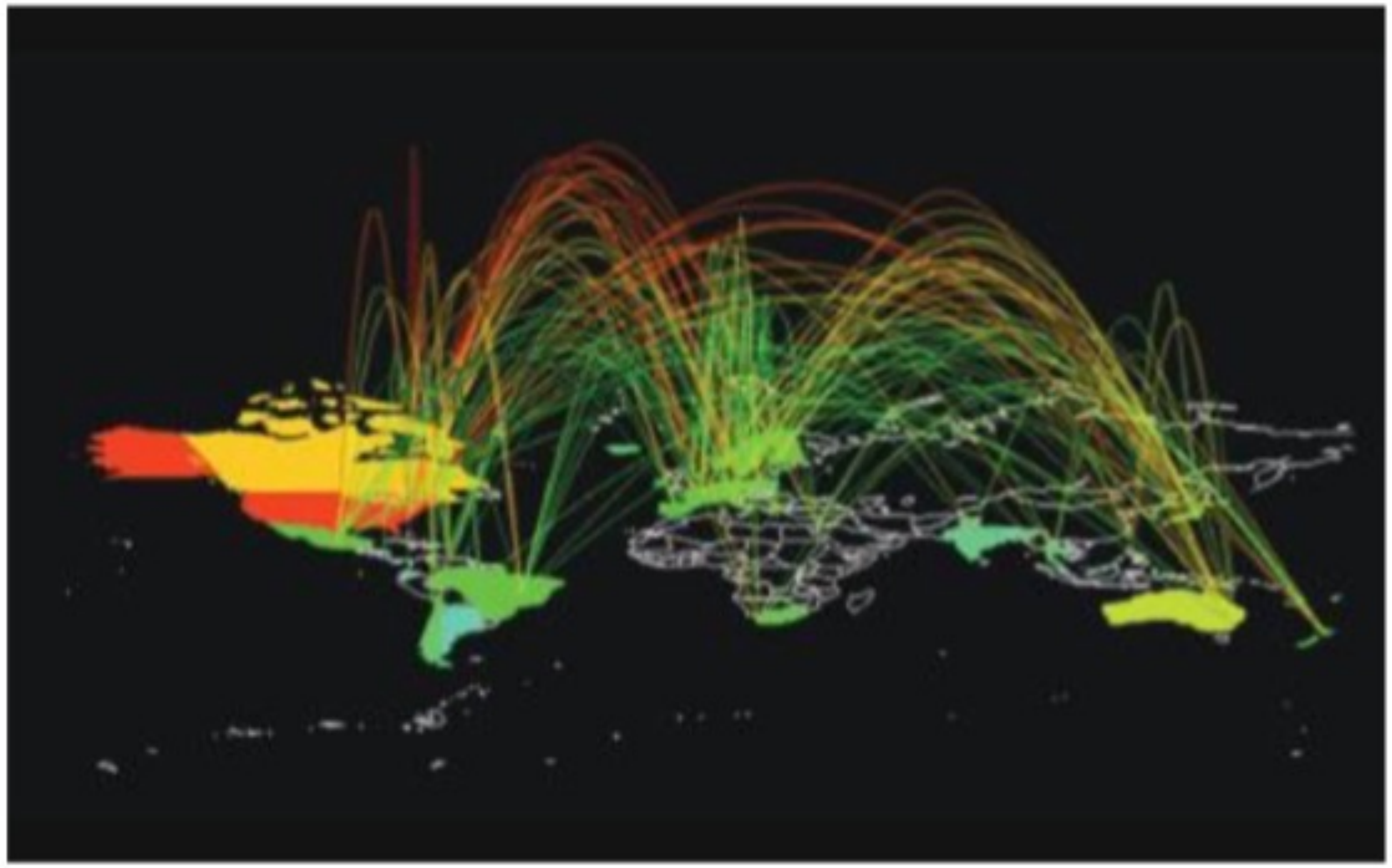
- **Network Maps**



Visualization study of inbound traffic measured in billions of bytes on the NSFNET T1 backbone for September 1991

Visualization of Line Data

- **Flow Maps and Edge Bundling**

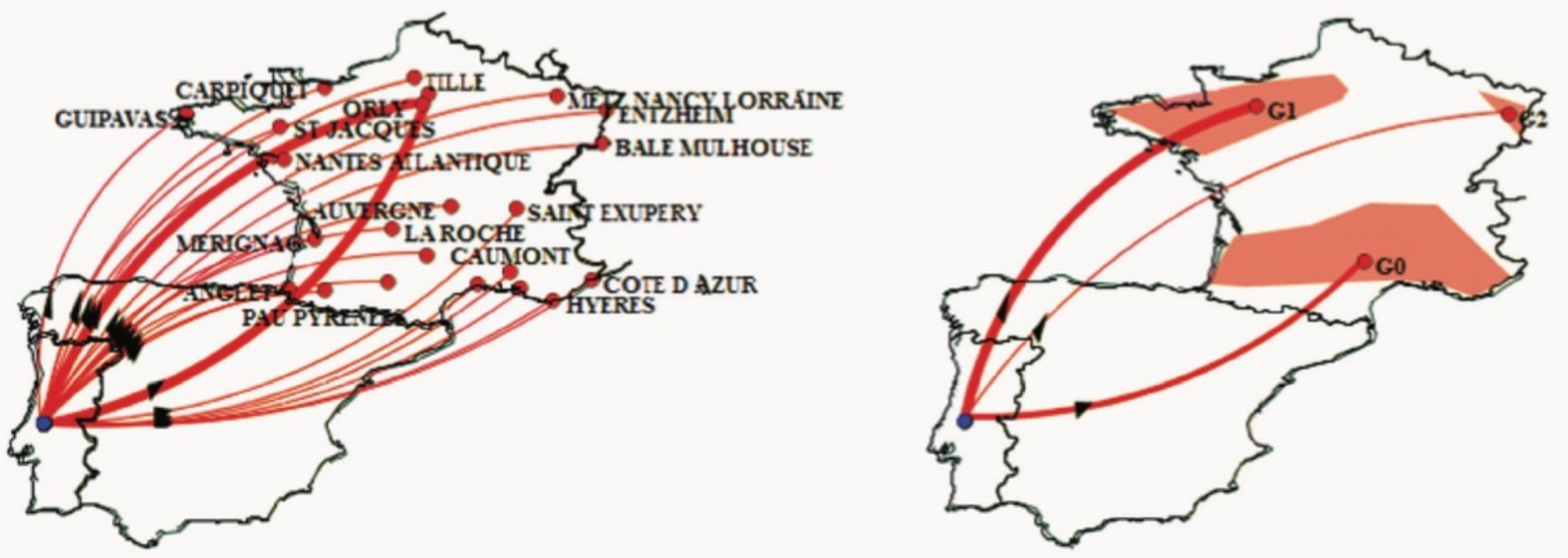


ArcMap. (Image from [87], © 1996 IEEE.)

Visualization of Line Data

- **Flow Maps and Edge Bundling**

Figure 15. The usage of the clustering ad–hoc approach with two spAs from different spDs



a) Without Clustering b) With Clustering

Spatial Clustering in SOLAP Systems to Enhance Map Visualization, Ricardo Silva, João Moura-Pires, Maribel Yasmina Santos

Visualization of Line Data

- **Flow Maps and Edge Bundling**

(a) (b)

Flow maps: (a) flows of tourists in Berlin; (b) produced by the Stanford system showing the migration from California (image from [318], © 2005 IEEE).

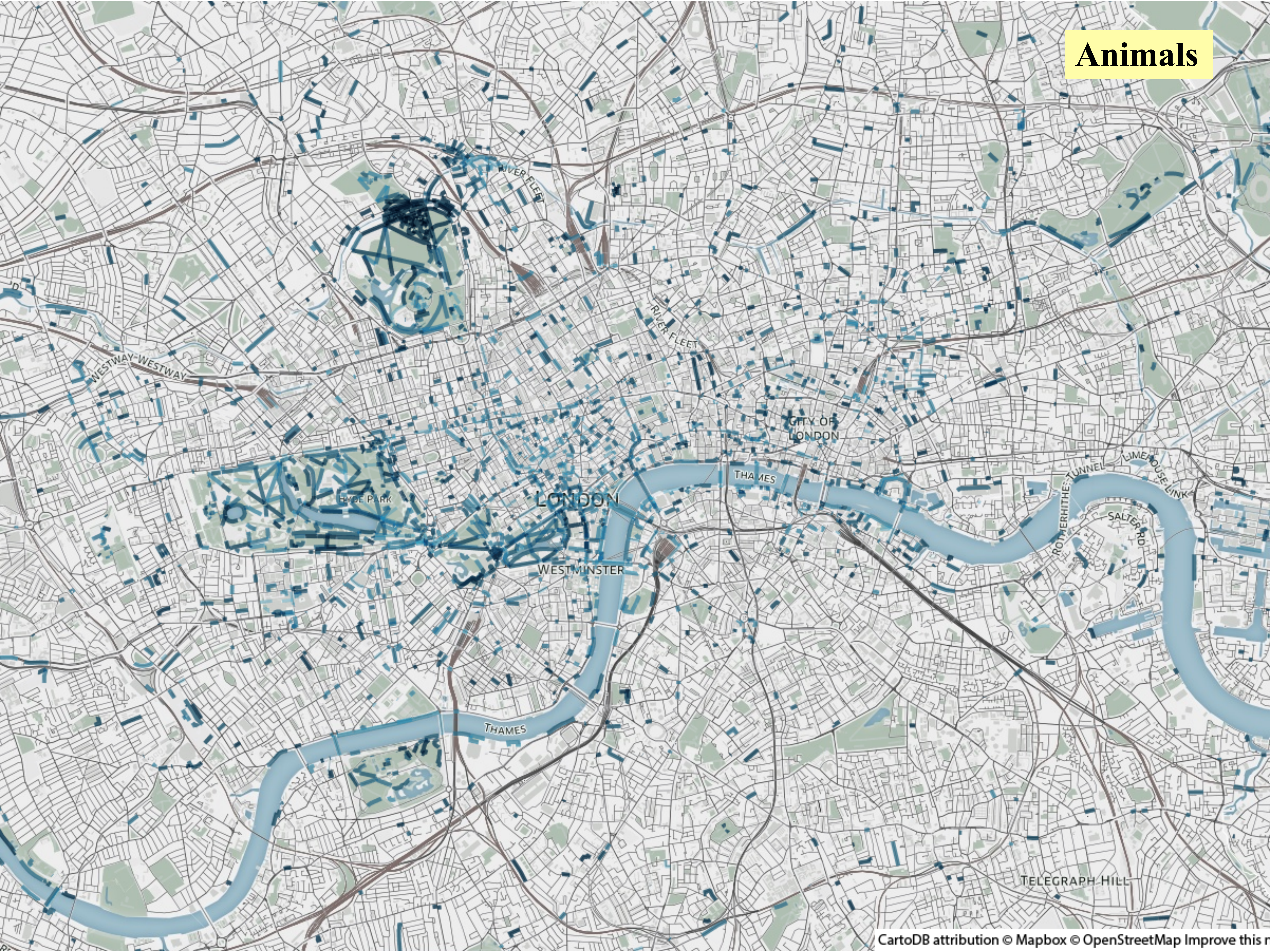
Visualization of Line Data

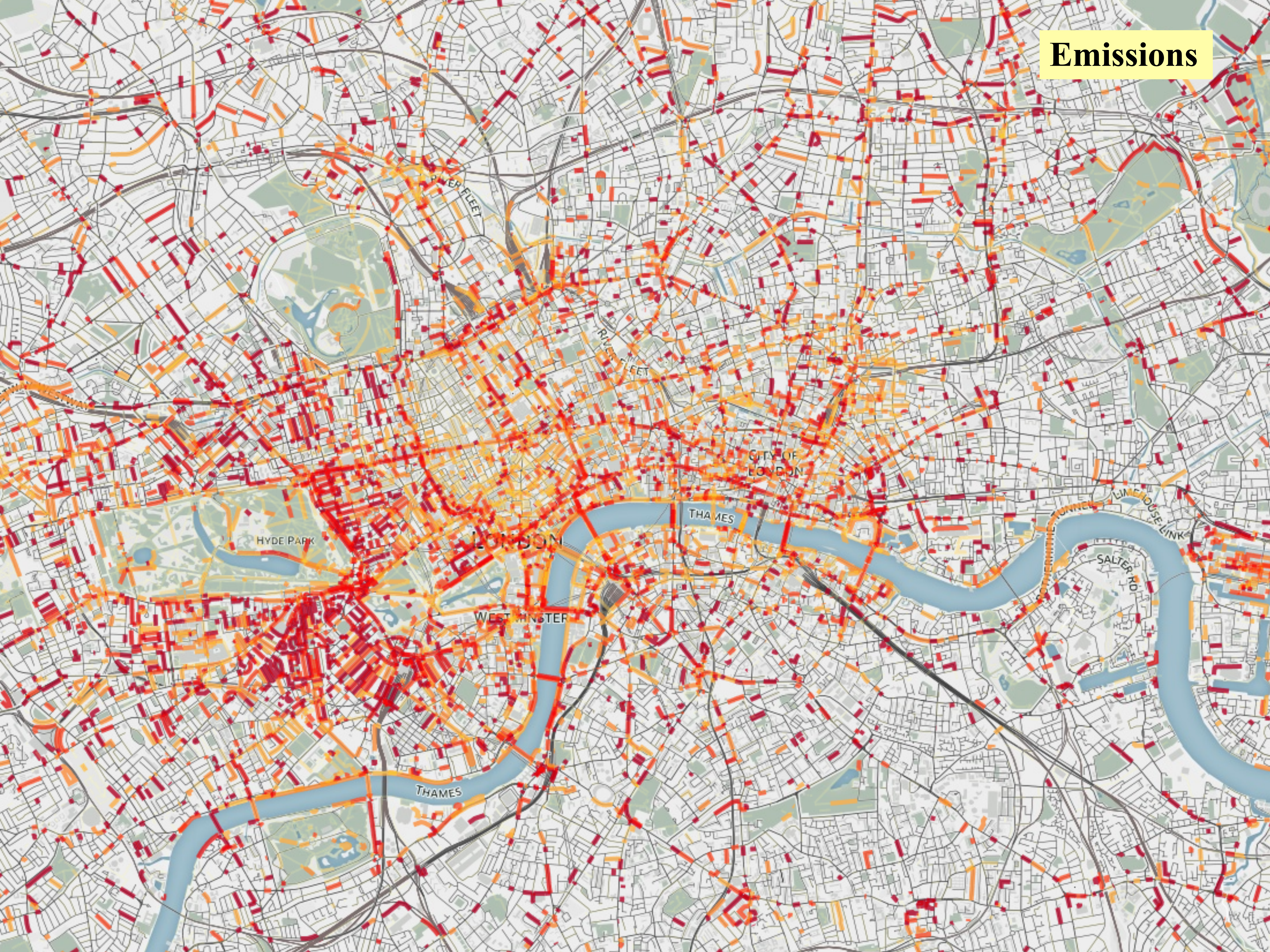
- **Geo-espacial lines**

Mapping London's smells: 'smellscapes' show which streets stink

http://www.telegraph.co.uk/news/earth/environment/11653964/Mapping-Londons-smells-smellscapes-show-which-streets-stink.html

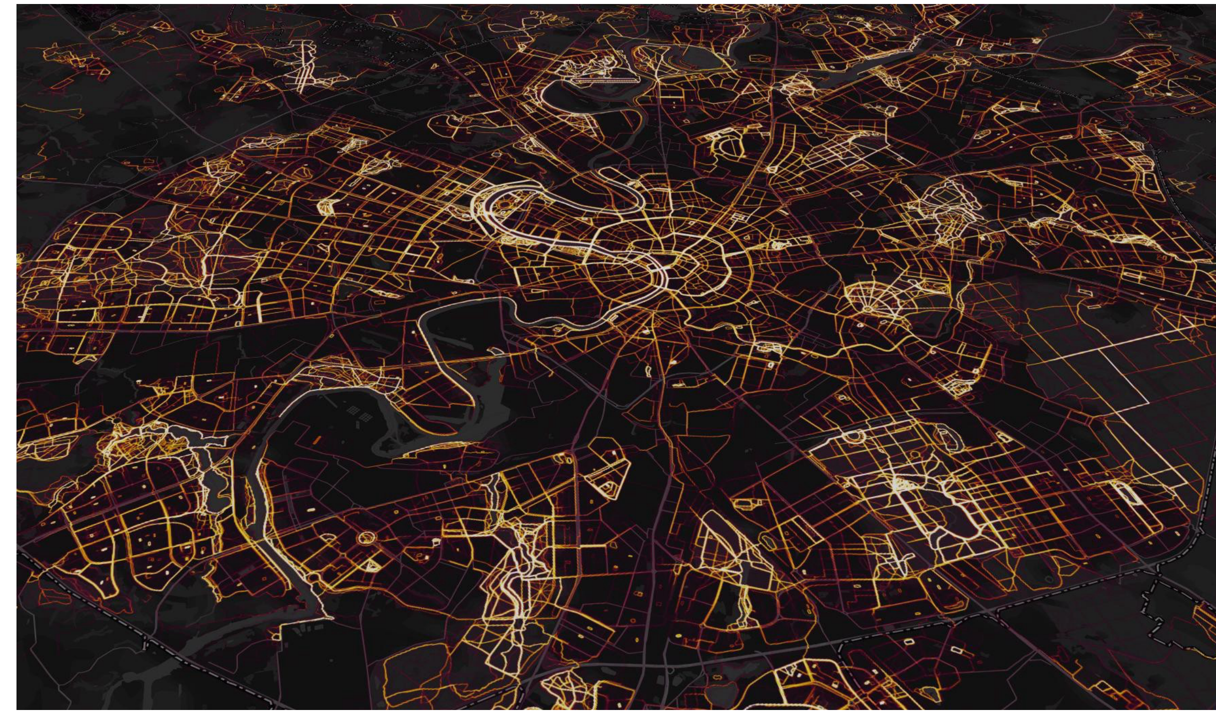

EMISSIONS
NATURE
FOOD
ANIMALS


Animals
RIVER FLEET
RIVER FLEET
WESTWAY-WESTWAY
HYDE PARK
LONDON
CITY OF LONDON
THAMES
WESTMINSTER
ROTHERHITHE TUNNEL
LIMEHOUSE LINK
SALTER RD
THAMES
TELEGRAPH HILL
CartoDB attribution © Mapbox © OpenStreetMap Improve this m


Emissions
HYDE PARK
LONDON
WESTMINSTER
CITY OF LONDON
THAMES
THAMES
RIVER FLEET
RIVER FLEET
LIMEHOUSE LINK
SALTER RD

STRAVA Global Heatmap

- **The heatmap consists of:**
  - **700 million activities**
  - **1.4 trillion latitude/longitude points**
  - **7.7 trillion pixels rasterized**
  - **5 terabytes of raw input data**
  - **A total distance of 16 billion km (10 billion miles)**
  - **A total recorded activity duration of 100 thousand years**



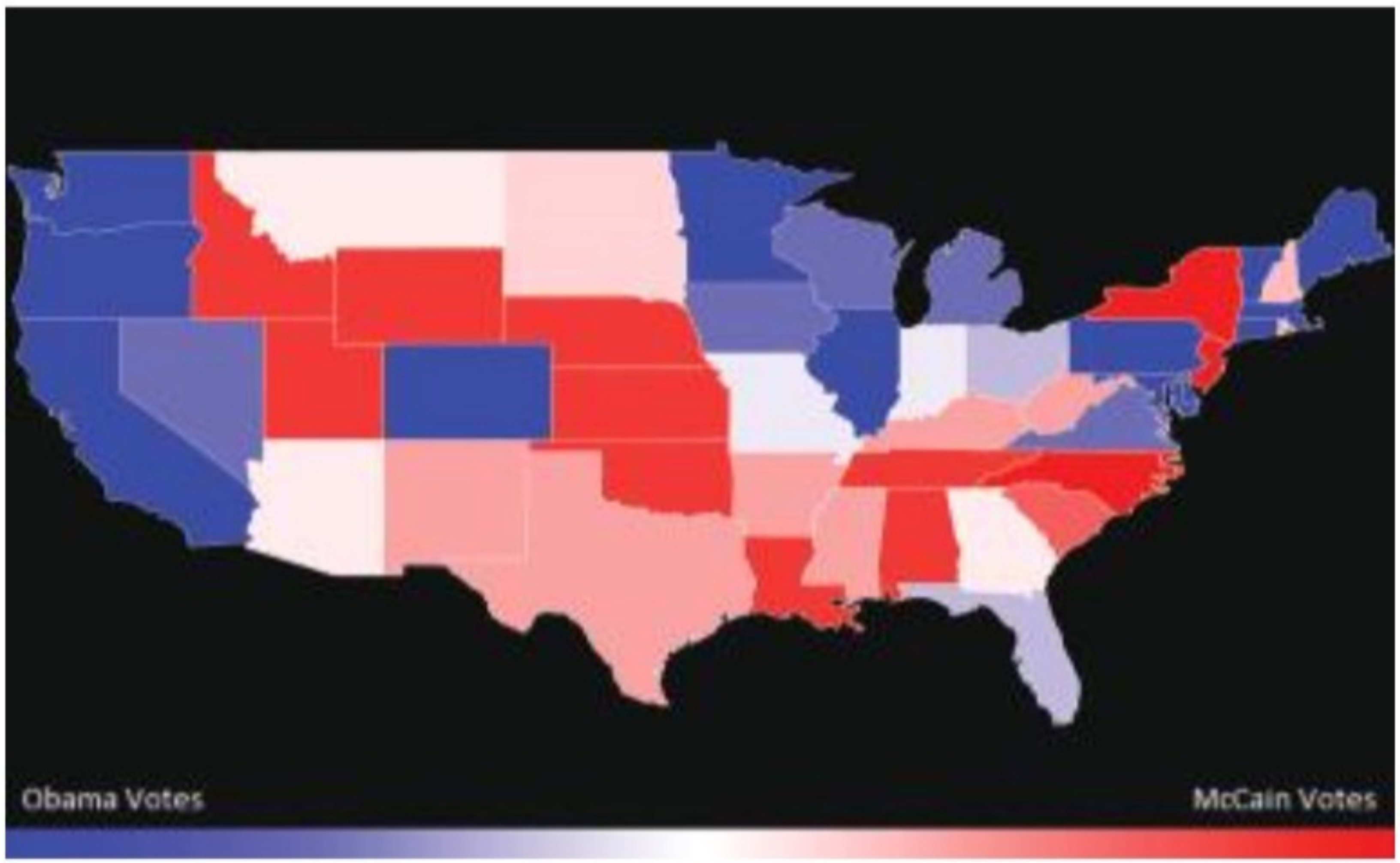
- **Check:**
  - **https://medium.com/strava-engineering/the-global-heatmap-now-6x-hotter-23fc01d301de**

The Global Heatmap in Moscow, Russia

Visualization of Area Data

Visualization of Area Data: Thematic maps

- **Choropleth maps: values of an attribute or statistical variable are encoded as colored or shaded regions on the map**
  - **Assume that the mapped attribute is uniformly distributed in the regions**



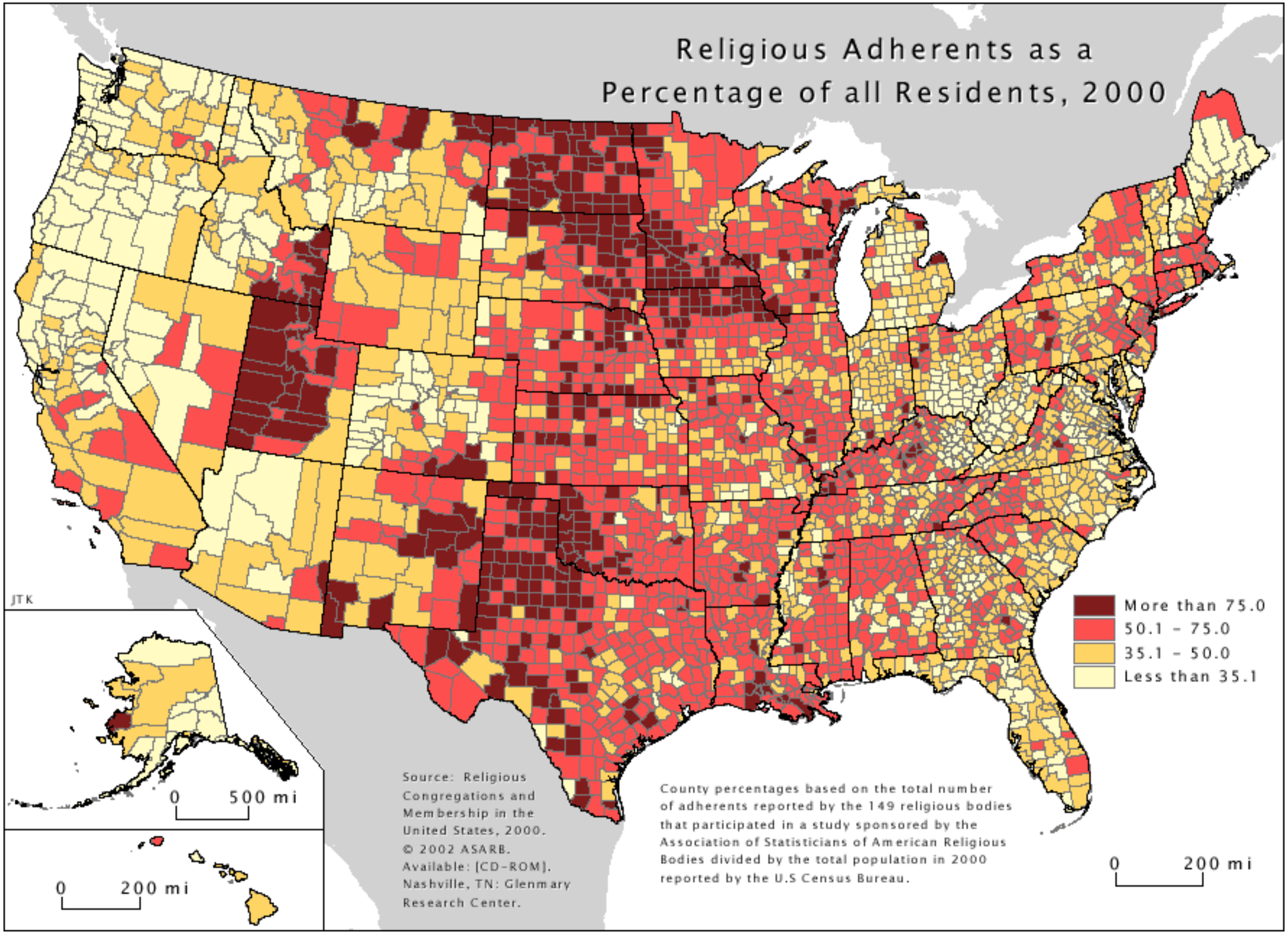
U.S. election results of the 2008 Obama versus McCain presidential election.

Visualization of Area Data: Thematic maps

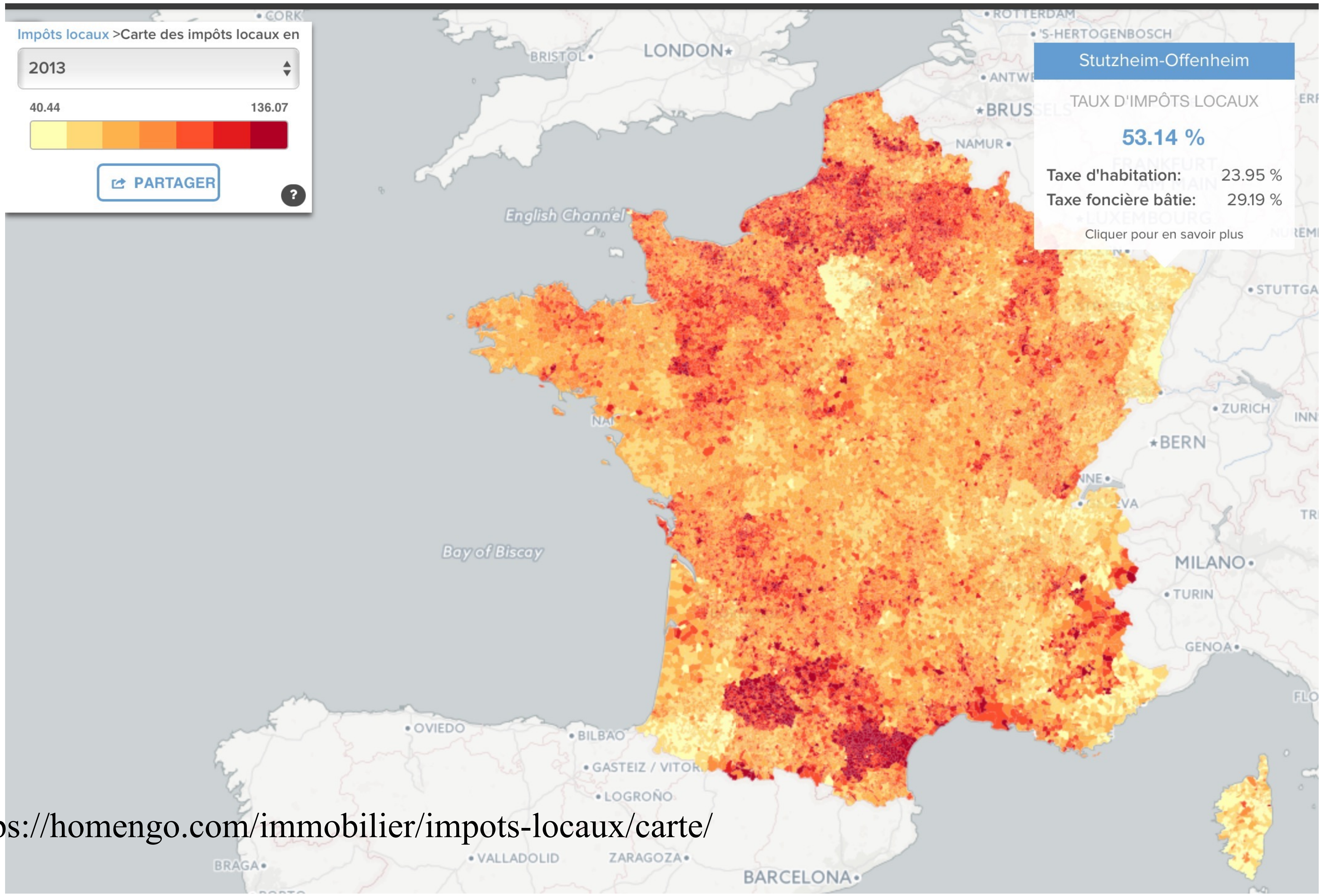
- **Choropleth maps:** values of an attribute or statistical variable are encoded as colored or shaded regions on the map
  - Assume that the mapped attribute is uniformly distributed in the regions

- A problem of choropleth maps is that the **most interesting values are often concentrated in densely populated areas with small and barely visible polygons**, and less interesting values are spread out over sparsely populated areas with large and visually dominating polygons.

Visualization of Area Data: Thematic maps

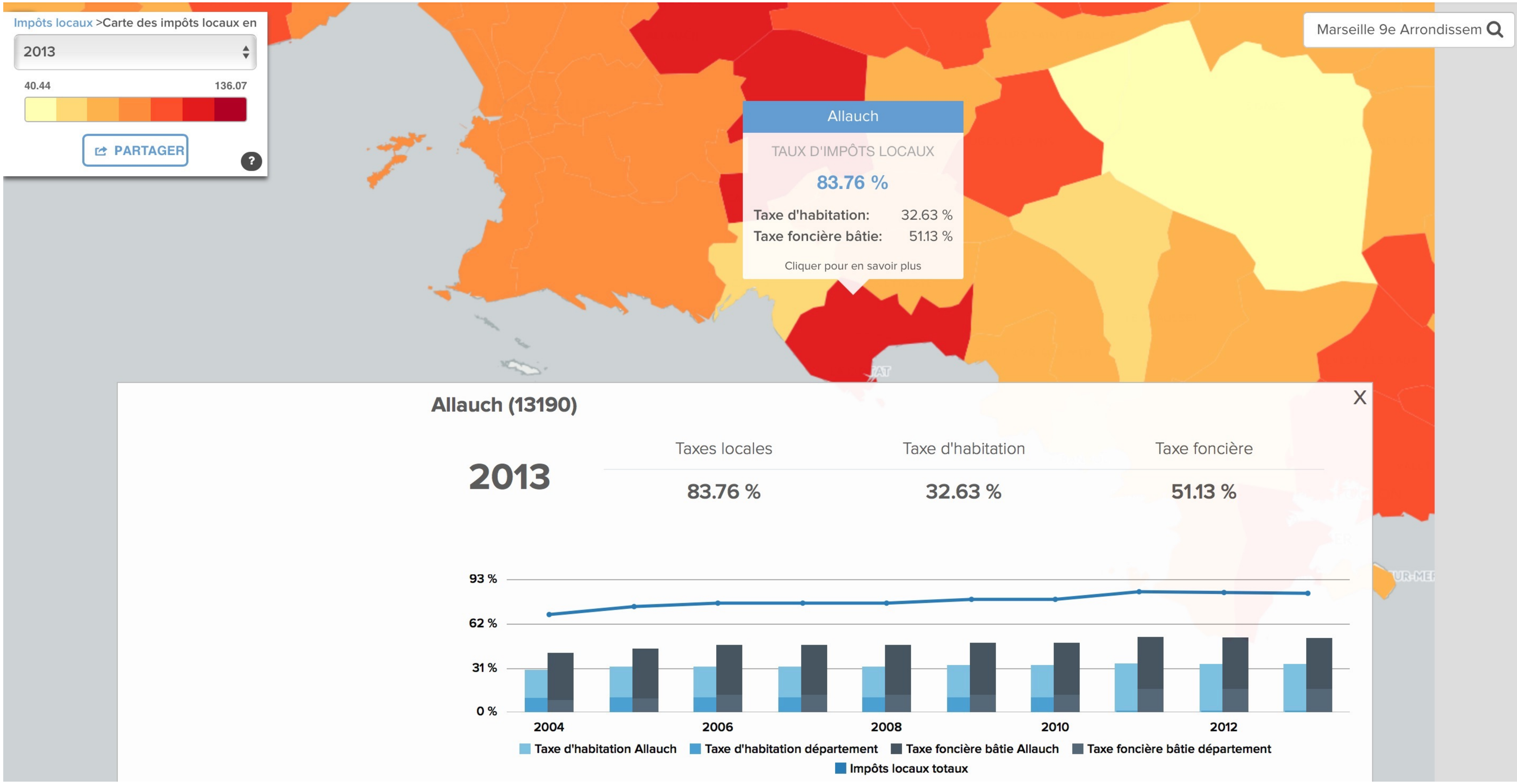


Visualization of Area Data: Thematic maps



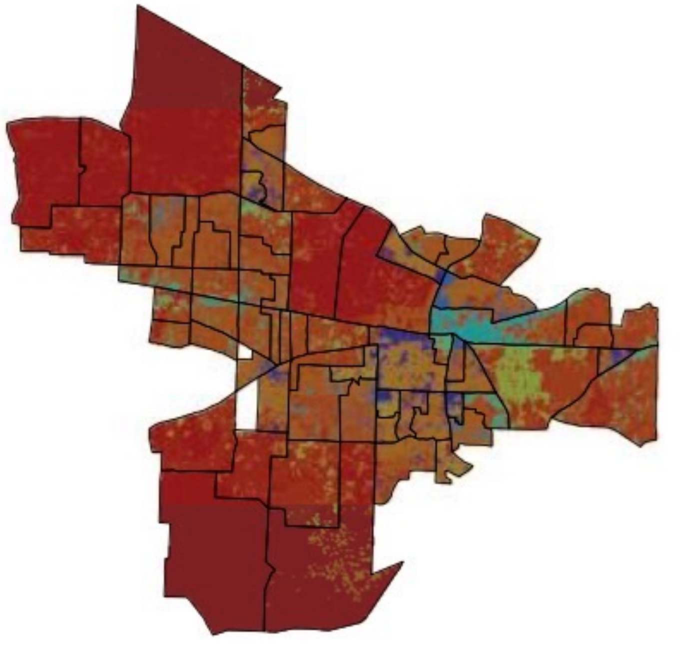
https://homengo.com/immobilier/impots-locaux/carte/

Visualization of Area Data: Thematic maps

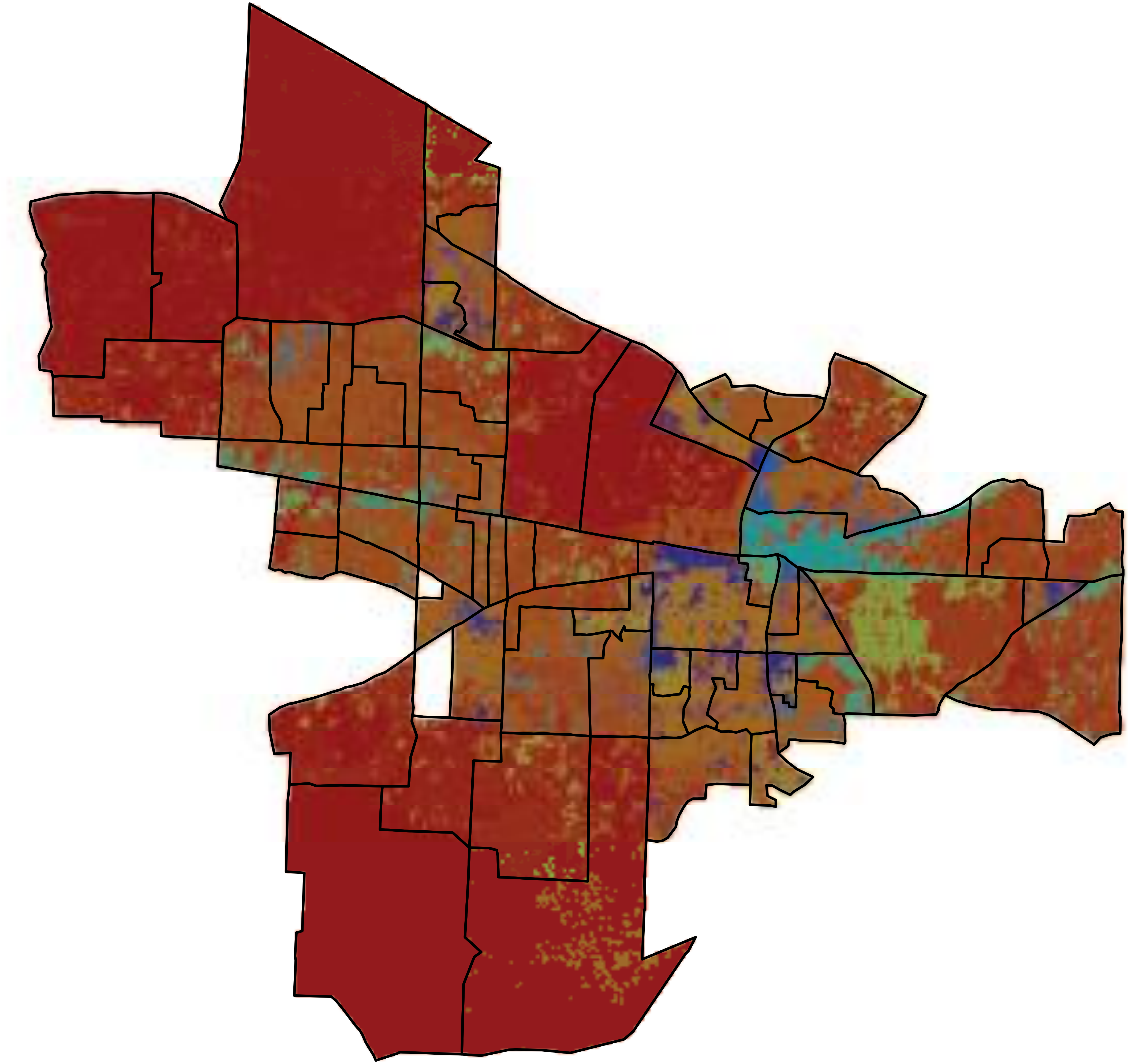


Visualization of Area Data: Thematic maps

- **If the attribute has a different distribution than the partitioning into regions, other techniques, such as dasymetric maps, are used. The variable to be shown forms areas independent of the original regions. To do this, ancillary information is acquired, which means the cartographer steps statistical data according to extra information collected within the boundary**



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



Visualization of Area Data: Thematic maps

- A third important type of map is an **isarithmic map**, which shows the contours of some continuous phenomena.
  - **isometric maps**, if the contours are determined from real data points such as temperatures measured at a specific location.
  - **isopleth maps**, if the data are measured for a certain region (such as a county) and, for example, the centroid is considered as the data point.

- One of the main tasks in generating **isarithmic maps** is the **interpolation of the data points** to **obtain smooth contours**, which is done, for example, by triangulation, or inverse distance mapping.

Visualization of Area Data: Thematic maps

An isarithmic map showing the the average anual temperature

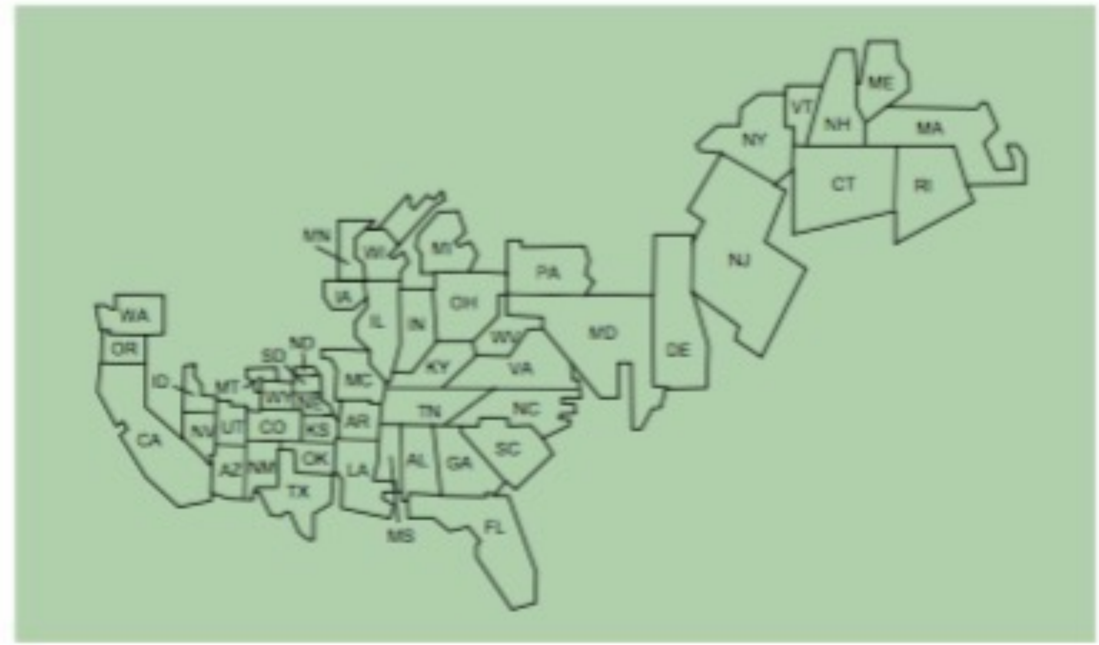
https://en.wikipedia.org/wiki/Thematic_map

Visualization of Area Data: Cartograms

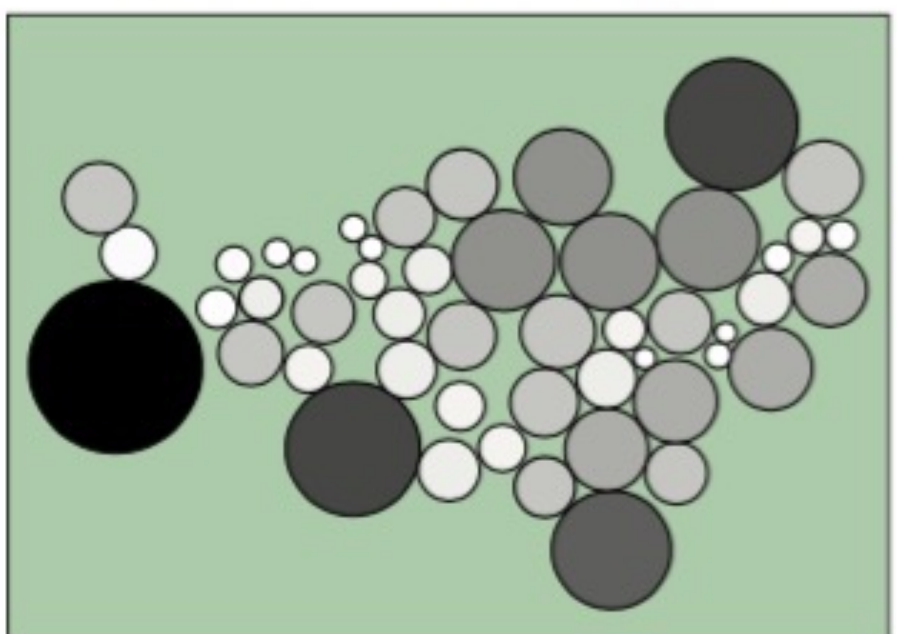
- A problem of choropleth maps is that the most interesting values are often concentrated in densely populated areas with small and barely visible polygons, and less interesting values are spread out over sparsely populated areas with large and visually dominating polygons.
- Cartograms are generalizations of ordinary thematic maps that avoid the problems of choropleth maps by distorting the geography according to the displayed statistical value:
  - The size of regions is scaled to reflect a statistical variable, leading to unique distortions of the map geometry.

Visualization of Area Data: Cartograms

(a) Noncontinuous cartogram.



(b) Noncontiguous cartogram.



(c) Circular cartogram.

(d) Continuous cartogram.

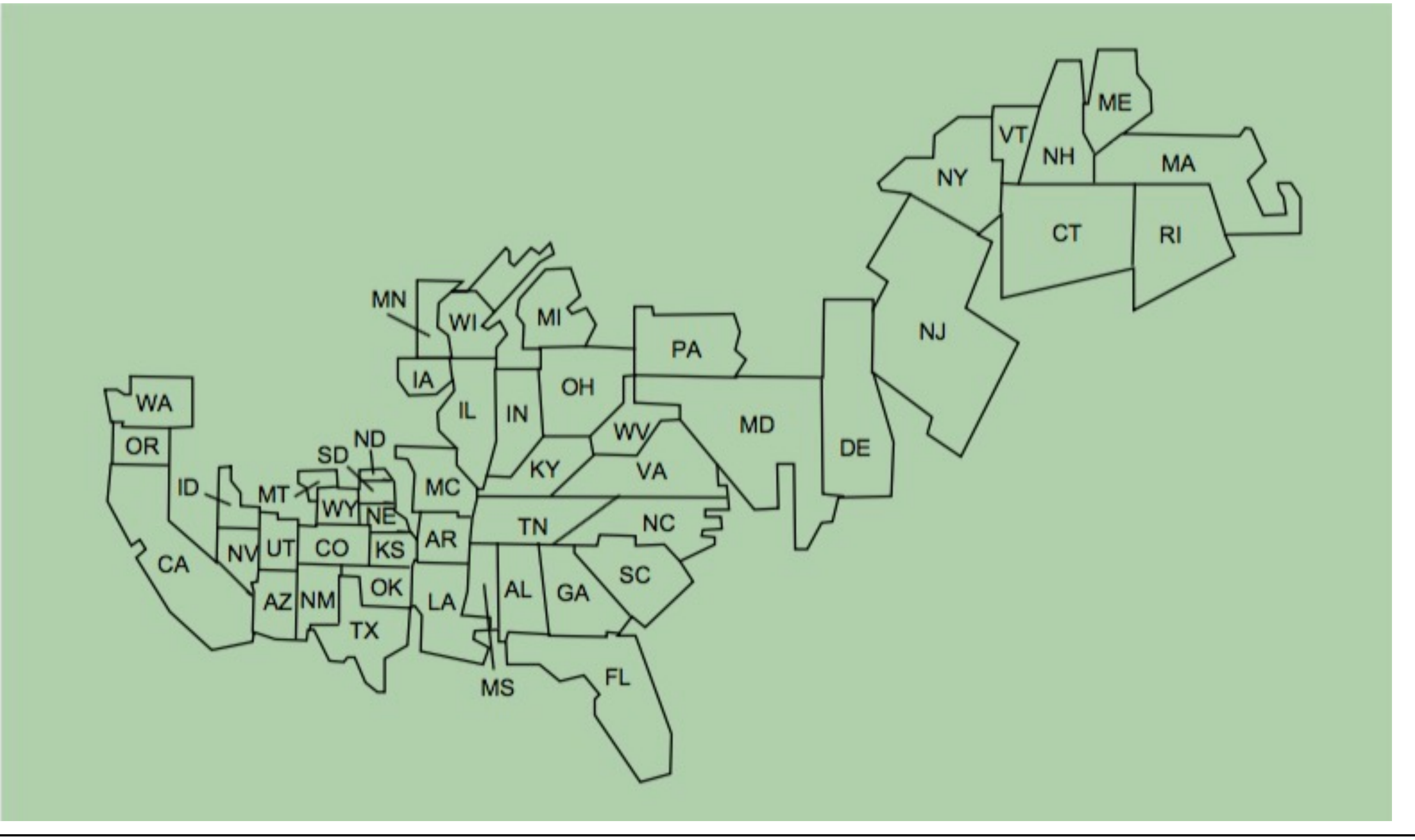
Visualization of Area Data: Cartograms

- **Noncontinuous cartograms can exactly satisfy area and shape constraints, but don't preserve the input map's topology. Because the scaled polygons are drawn inside the original regions, the loss of topology doesn't cause perceptual problems.**
  - **More critical is that the polygon's original size restricts its final size.**

FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE NOVA DE LISBOA

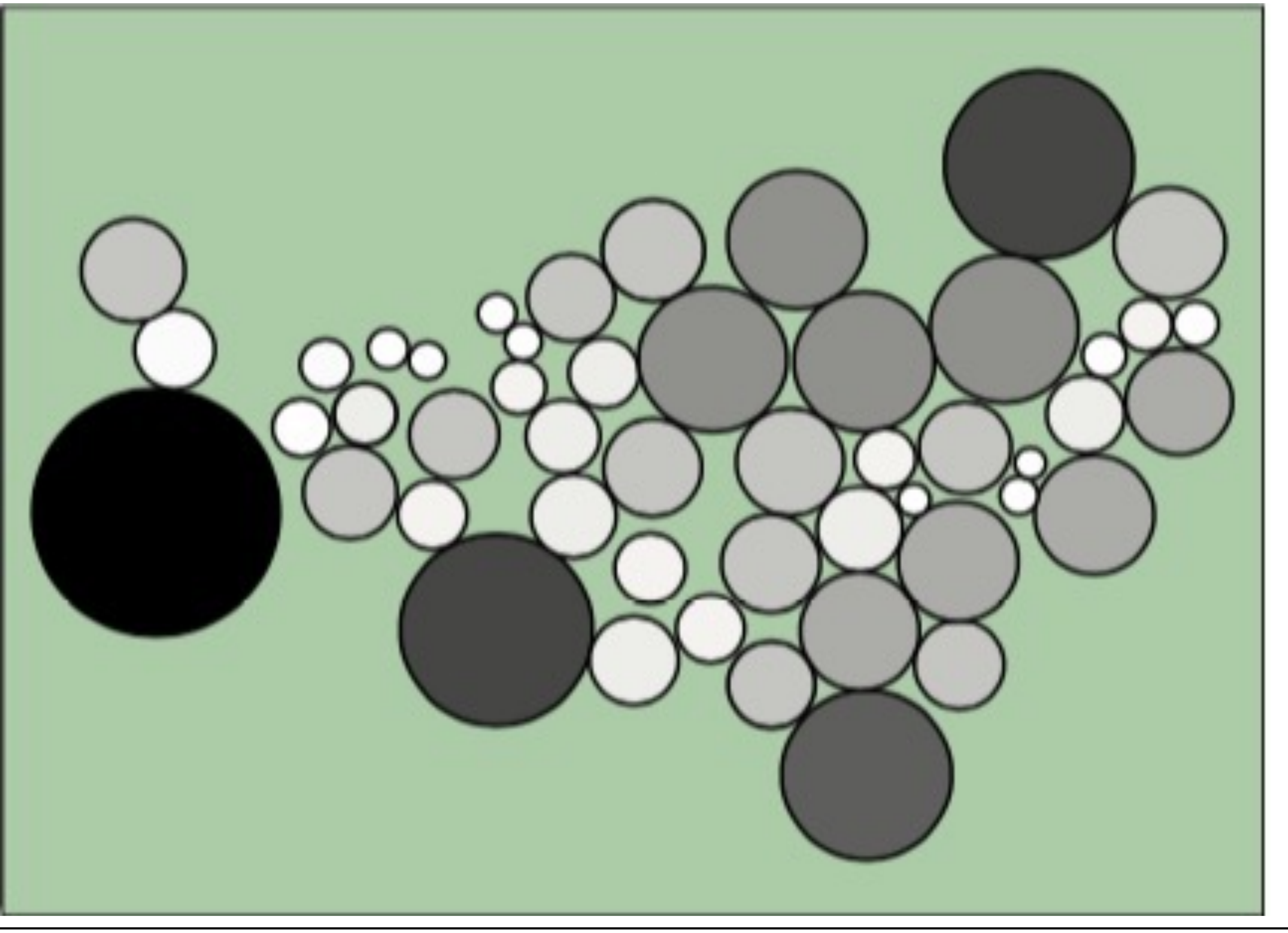
Visualization of Area Data: Cartograms

- **Noncontiguous cartograms,** scale all polygons to their target sizes, perfectly satisfying the area objectives. They provide perfect area adjustment, with good shape preservation. **However, they lose the map's global shape and topology**, which can make perceiving the generated visualization as a map difficult.



Visualization of Area Data: Cartograms

- **Circular cartograms, completely ignore the input polygon's shape, representing each as a circle in the output. Circular cartograms have some of the same problems as noncontiguous cartograms.**



Visualization of Area Data: Cartograms

- **Continuous cartograms** retain a map's topology perfectly, but they relax the given area and shape constraints. In general, cartograms can't fully satisfy shape or area objectives, so cartogram generation involves a complex optimization problem in searching for a good compromise between shape and area preservation

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Visualization of Area Data: Cartograms

- **A U.S. state population cartogram with the presidential election results of 2000. The area of the states in the cartogram corresponds to the population, and the color (shaded and not shaded areas) corresponds to the percentage of the vote. A bipolar color map depicts which candidate has won each state.**

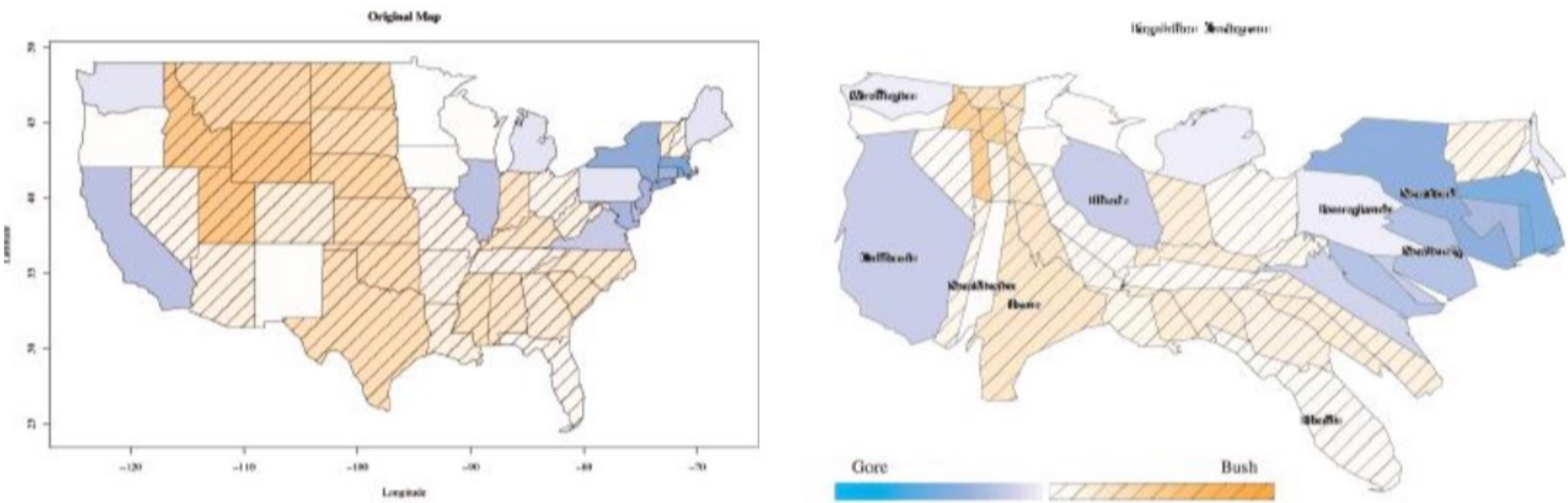


Figure 6.21 - Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward, Georges Grinstein, Daniel Keim, 2015

Visualization of Area Data: Cartograms

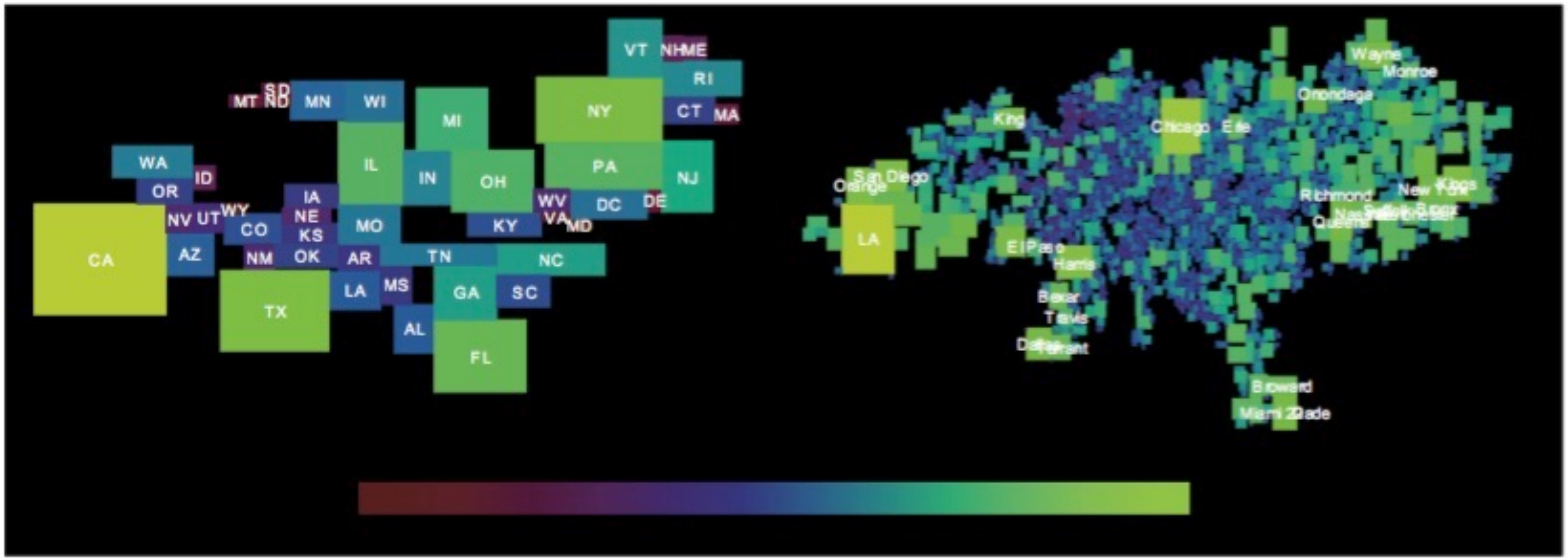
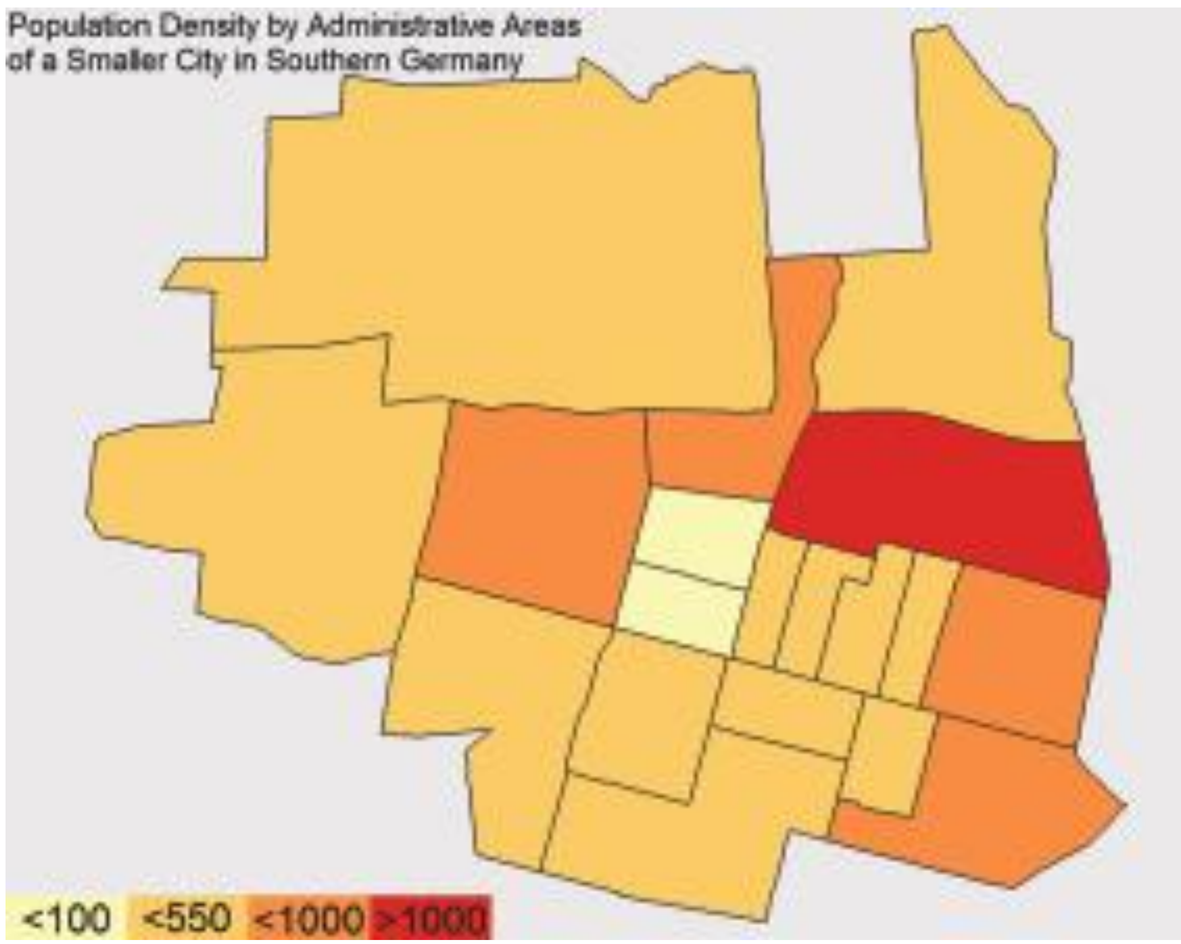


Figure 6.23. A rectangular U.S. population cartogram on the state and county level. The area of the rectangles corresponds to the population and the color redundantly encodes the population numbers. (Image from [184], © 2004 IEEE.)

Issues for spatial data mapping

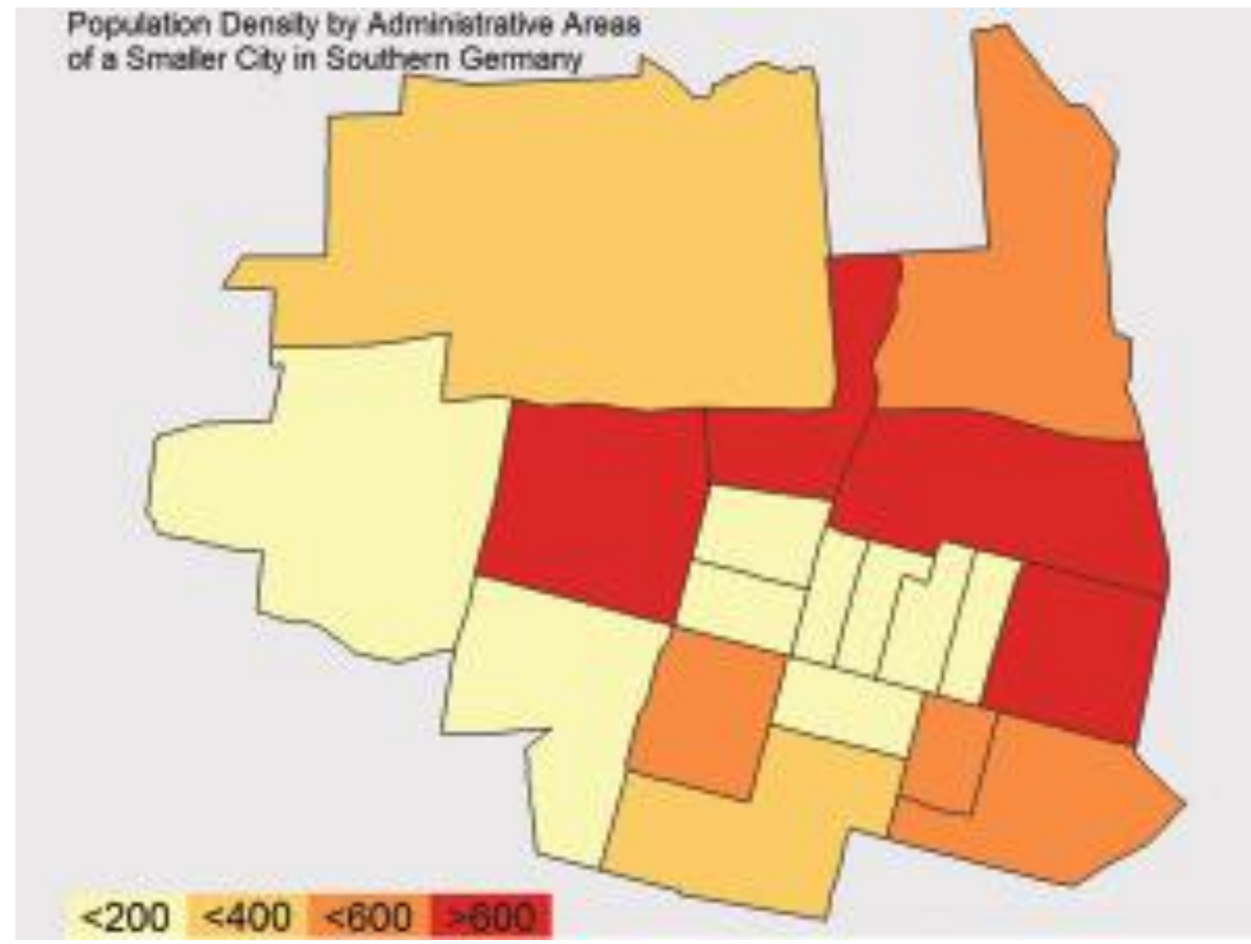
- **Note that in spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization:**



Different class separation with a significant impact on the generated map

Issues for spatial data mapping

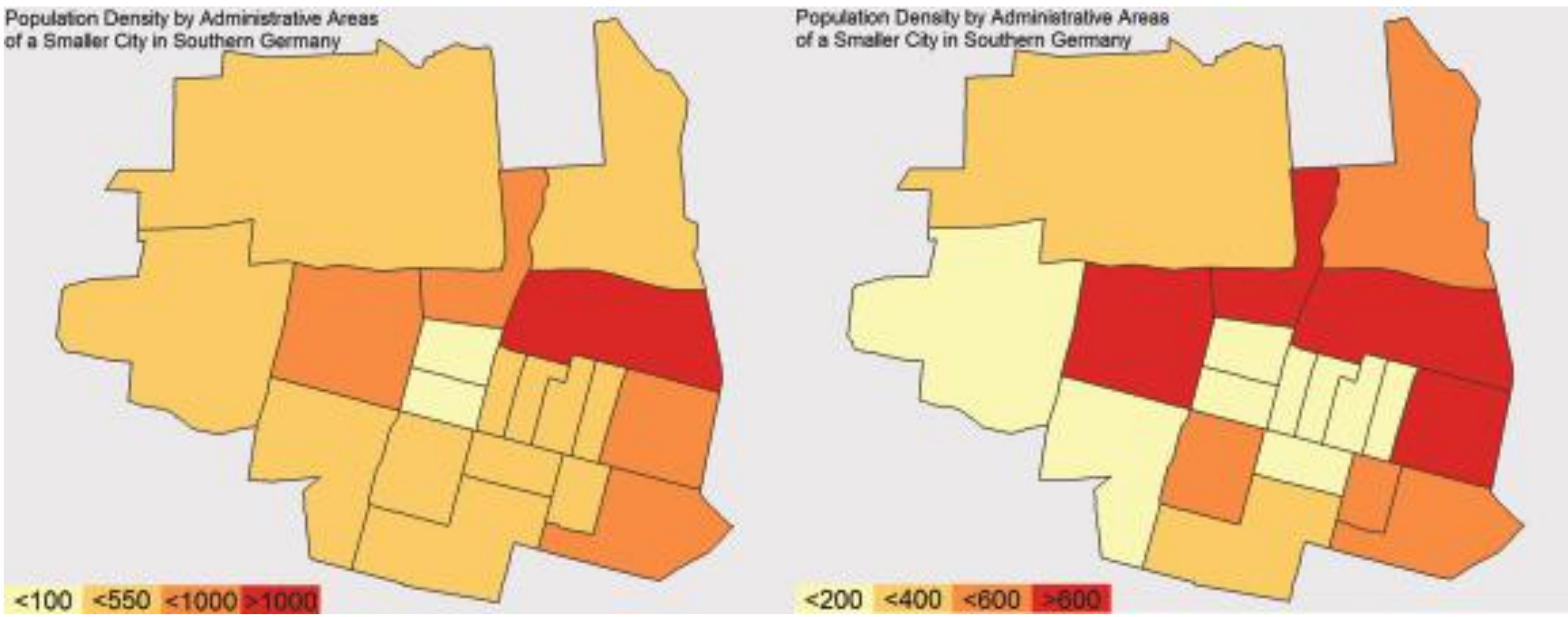
- **Note that in spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization:**



Different class separation with a significant impact on the generated map

Issues for spatial data mapping

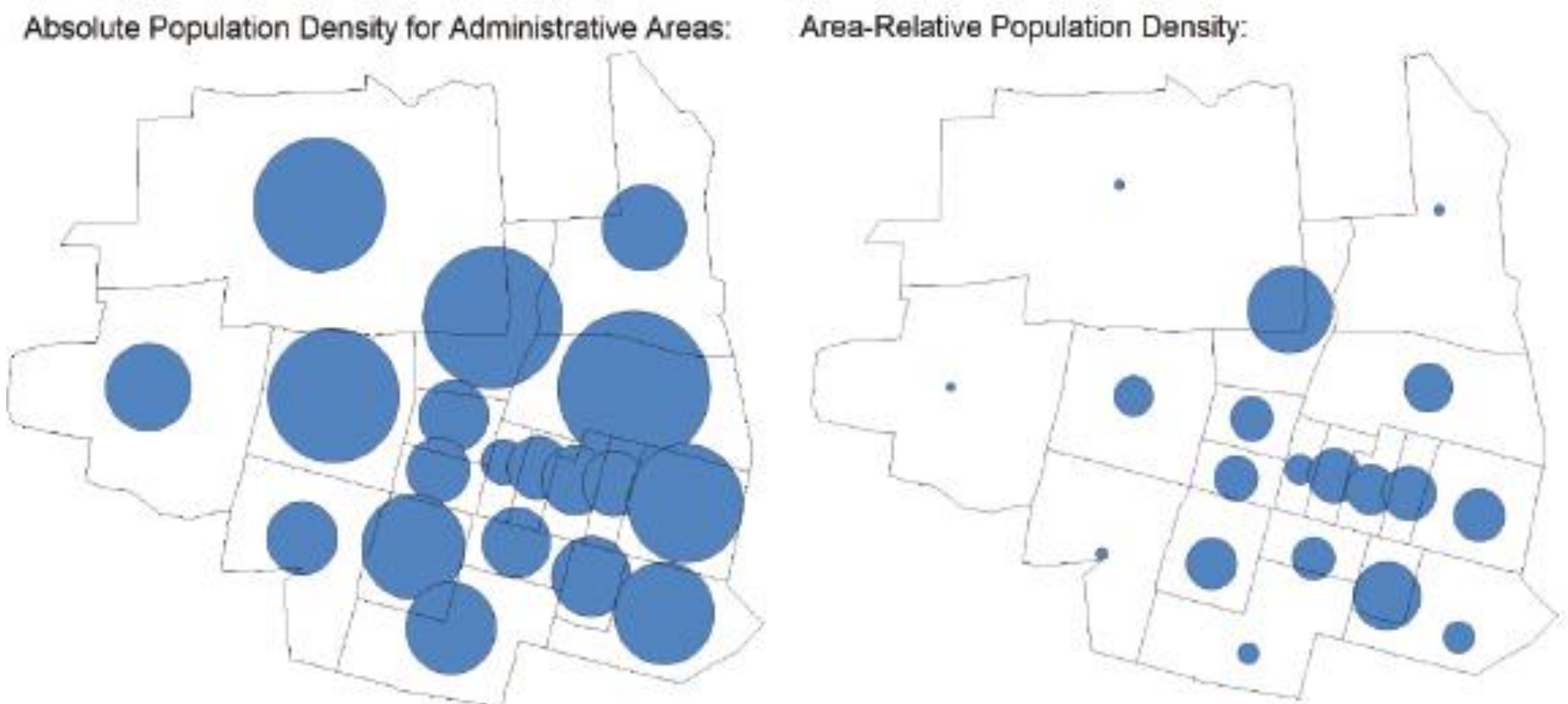
- **Note that in spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization:**



Different class separation with a significant impact on the generated map

Issues for spatial data mapping

- **Note that in spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization:**



Absolute versus relative mapping.
On the right numbers are displayed relative to the population numbers

Issues for spatial data mapping

- **Note that in spatial data mapping, the chosen class separation, normalization, and spatial aggregation may have a severe impact on the resulting visualization:**

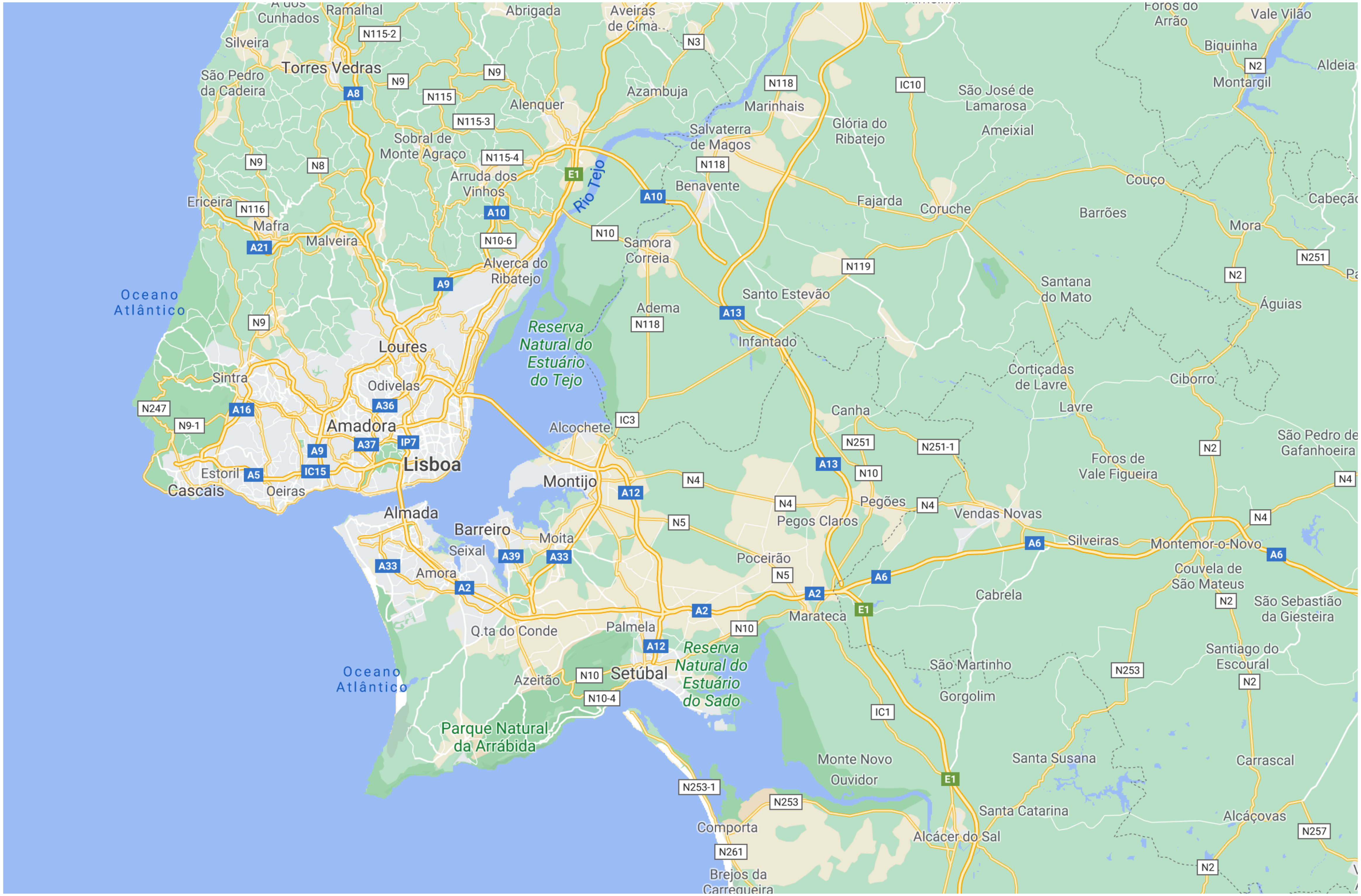
London cholera example with different area aggregations, resulting in quite different maps

Other Issues in Geospatial Data Visualization

Other Issues in Geospatial Data Visualization

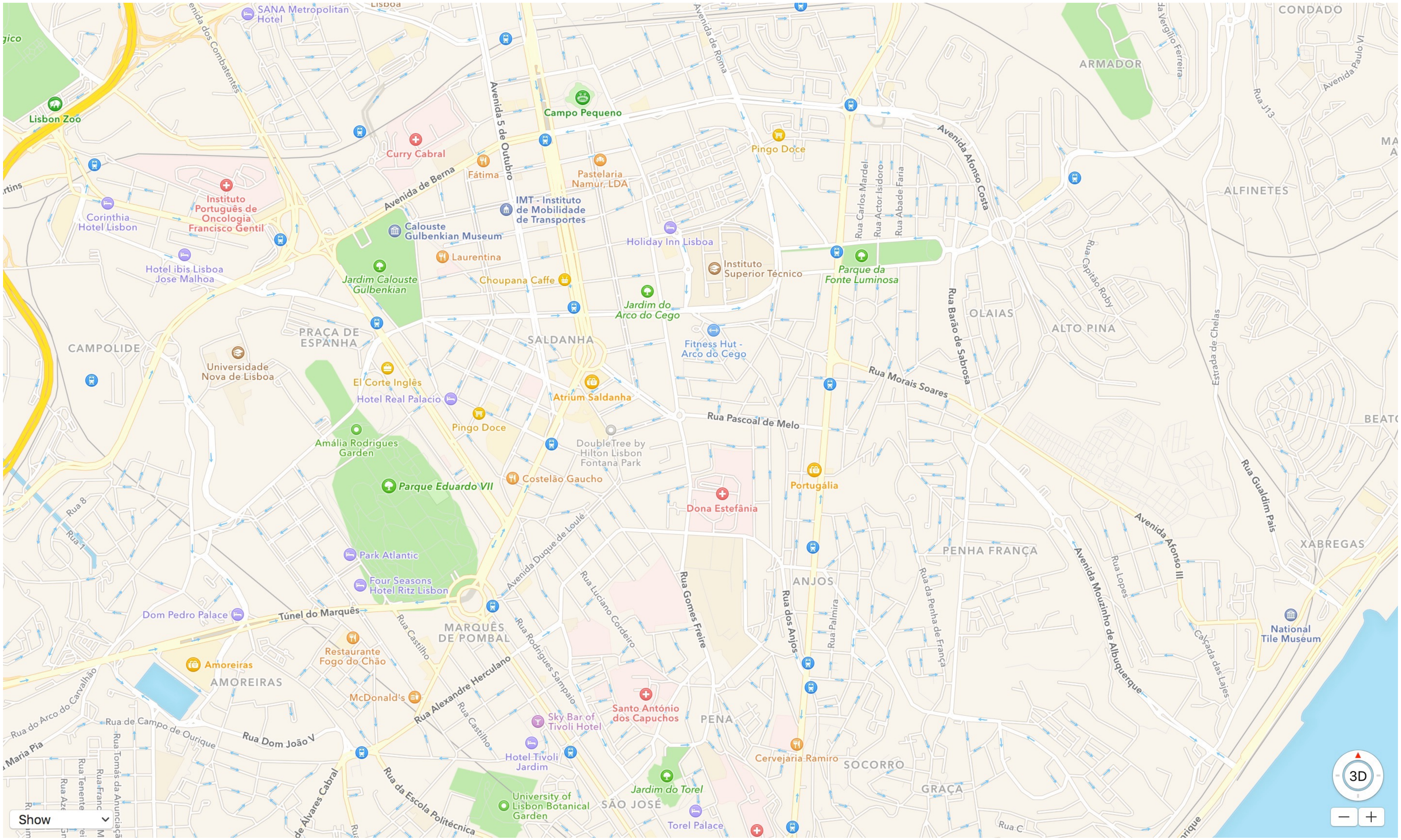
- **Map generalization** is the process of selecting and abstracting information on a map. Generalization is used when a small-scale map is derived from a large-scale map containing detailed information.
  - **map generalizations are application- and task-dependent**, e.g., good map generalizations emphasize the map elements that are most important for the task at hand, while still representing the geography in the most accurate and recognizable way: simplify points, simplify lines, simplify polygons, etc..
  - Map labeling deals with placing textual or figurative labels close to points, lines, and polygons

Other Issues in Geospatial Data Visualization

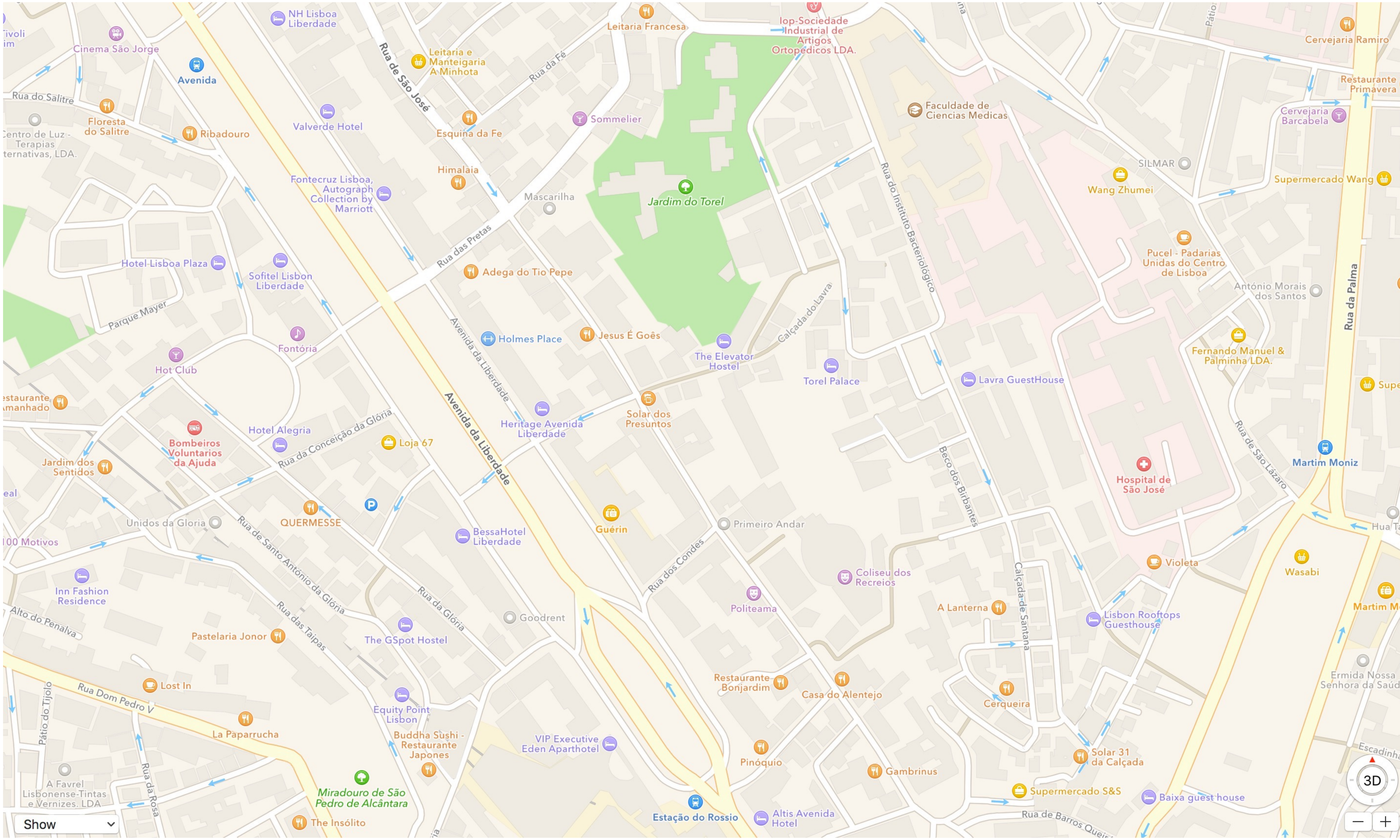


Other Issues in Geospatial Data Visualization

Other Issues in Geospatial Data Visualization



Other Issues in Geospatial Data Visualization



Further Reading and Summary

Further Reading

- **Pag 221 - 253 from Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward, Georges Grinstein, Daniel Keim, 2015**

Tips and Tools

- **MapBox: https://www.mapbox.com**
- **MapMap: https://github.com/floledermann/mapmap-examples**
- **CartoDB: https://carto.com/solutions/web-mobile/**
- **Esri Maps Javascript API: https://developers.arcgis.com/javascript/**

- **Tableau:**
  - https://help.tableau.com/current/pro/desktop/en-us/maps_options.htm
  - **Use Mapbox Maps in Tableau**
    - https://help.tableau.com/current/pro/desktop/en-us/maps_mapsources_mapbox.htm