Data Visualization and Maps I

HES 505 Fall 2024: Session 24

Carolyn Koehn

Homework Questions / Solutions

Objectives

By the end of today you should be able to:

- Describe some basic principles of data visualization
- Extend principles of data visualization to the development of maps
- Distinguish between several common types of spatial data visualization

Introduction to Data Visualization

Principles vs. Rules

- Lots of examples of *good* and *bad* data visualization
- What makes a graphic good (or bad)?
- Who decides?

- **Rule:** externally compels you, through force, threat or punishment, to do the things someone else has deemed good or right.
- **Principle:** internally motivating because it is a *good practice*; a general statement describing a philosophy that good rules should satisfy
- Rules contribute to the design process, but do not guarantee a satisfactory outcome

> “Graphical excellence is the well-designed presentation of interesting data—a matter of substance, of statistics, and of design … [It] consists of complex ideas communicated with clarity, precision, and efficiency. … [It] is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space … [It] is nearly always multivariate … And graphical excellence requires telling the truth about the data.”
> — Edward Tufte

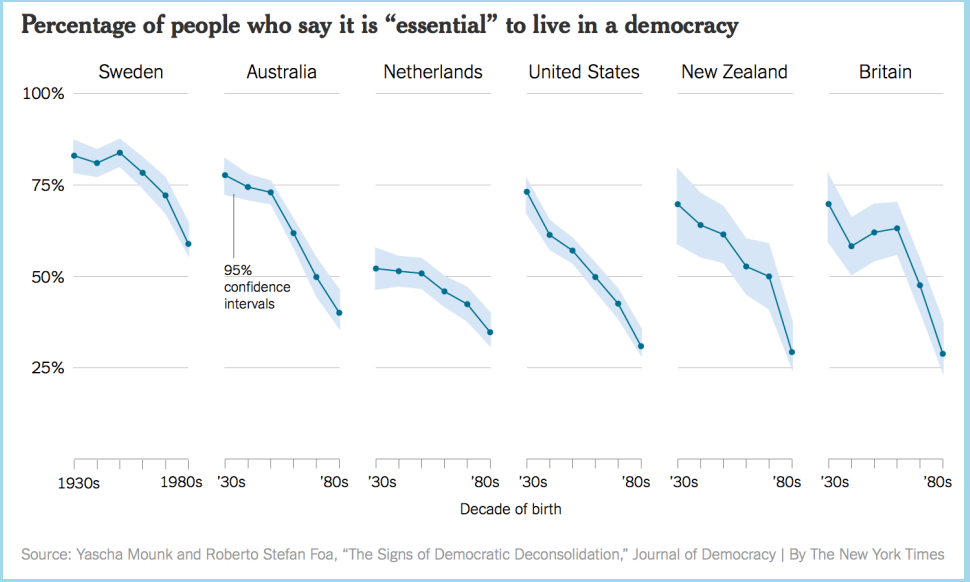
Ugly, Wrong, and Bad

- *Ugly*: graphic is clear and informative, but has aesthetic issues
- *Bad*: graphic is unclear, confusing, or decieving
- *Wrong*: the figure is objectively incorrect

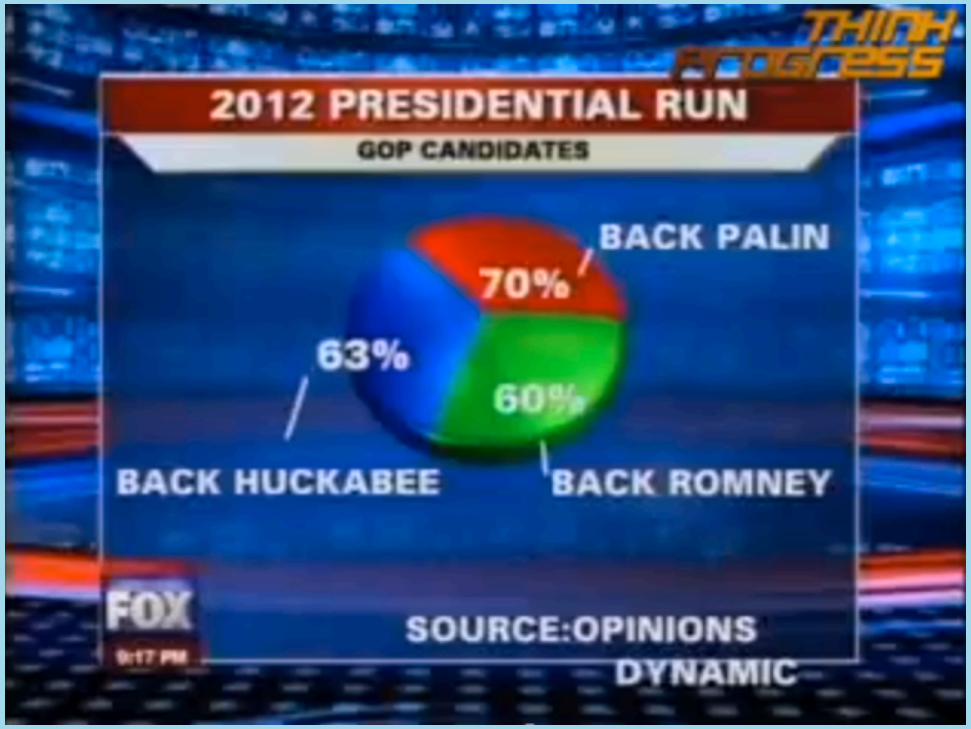
Monstrous Costs' by Nigel Holmes from Healy 2018

Bad and Wrong

- Presentation of the data is (intentionally?) decieving
- Presentation is just incorrect



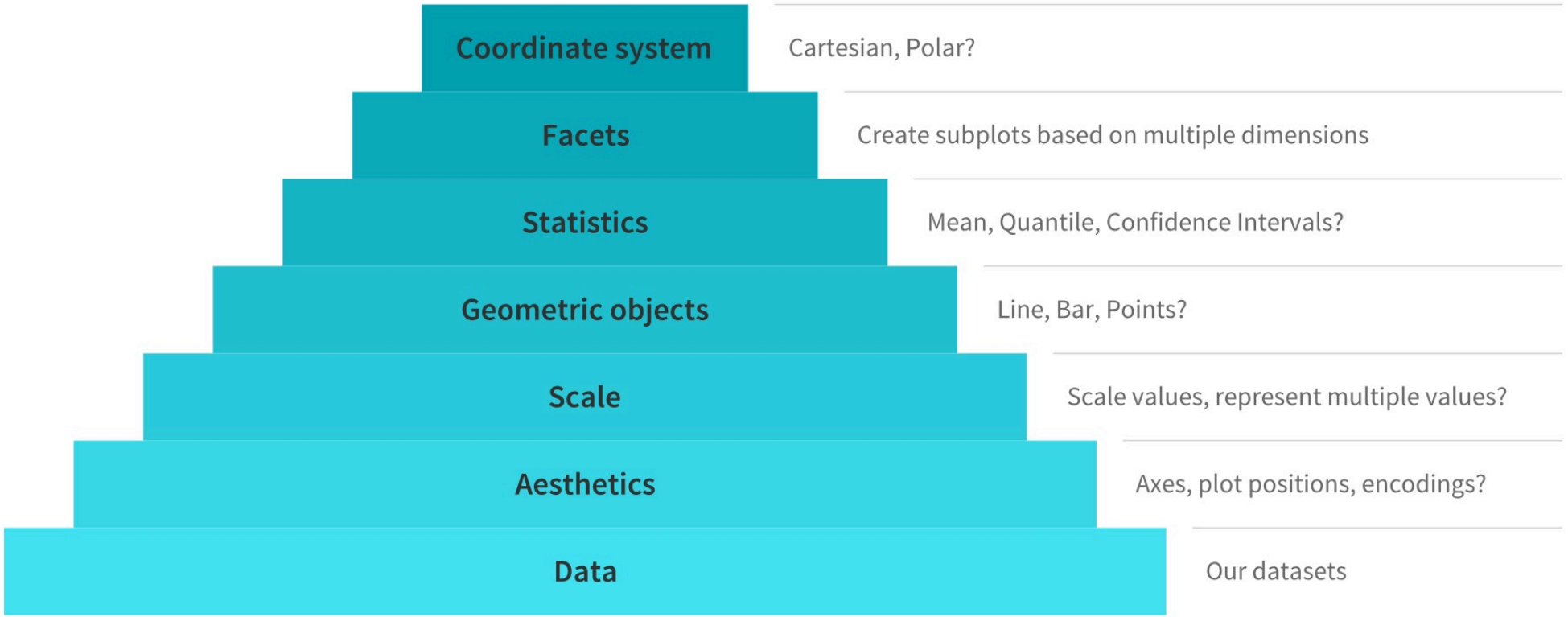
Tricky (from Healy 2018)



Wrong

Grammar of Graphics (Wilkinson 2005)

Major Components of the Grammar of Graphics



Aesthetics: Mapping Data to Visual Elements

- Define the systematic conversion of data into elements of the visualization
- Are either categorical or continuous (exclusively)
- Examples include `x`, `y`, `fill`, `color`, and `alpha`

Table 2.1: Types of variables encountered in typical data visualization scenarios.

Type of variable	Examples	Appropriate scale	Description
quantitative/numerical continuous	1.3, 5.7, 83, $1.5x10^{-2}$	continuous	Arbitrary numerical values. These can be integers, rational numbers, or real numbers.
quantitative/numerical discrete	1, 2, 3, 4	discrete	Numbers in discrete units. These are most commonly but not necessarily integers. For example, the numbers 0.5, 1.0, 1.5 could also be treated as discrete if intermediate values cannot exist in the given dataset.
qualitative/categorical unordered	dog, cat, fish	discrete	Categories without order. These are discrete and unique categories that have no inherent order. These variables are also called *factors*.
qualitative/categorical ordered	good, fair, poor	discrete	Categories with order. These are discrete and unique categories with an order. For example, "fair" always lies between "good" and "poor". These variables are also called *ordered factors*.
date or time	Jan. 5 2018, 8:03am	continuous or discrete	Specific days and/or times. Also generic dates, such as July 4 or Dec. 25 (without year).
text	The quick brown fox jumps over the lazy dog.	none, or discrete	Free-form text. Can be treated as categorical if needed.

From Wilke 2019

Scales

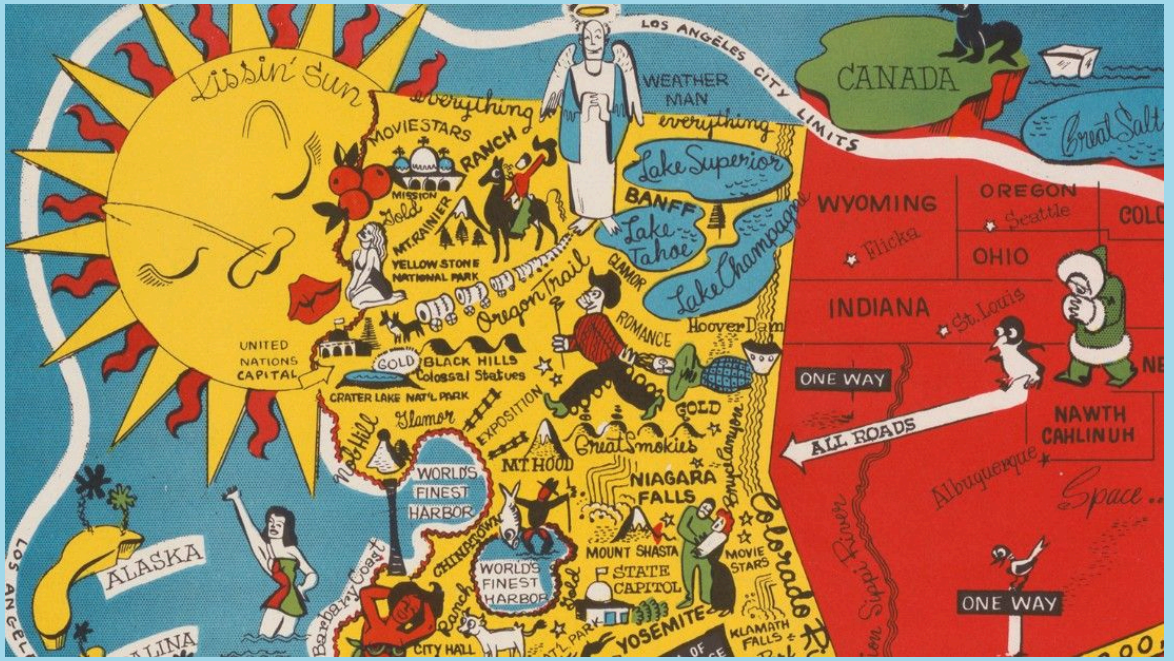
- Scales map data values to their aesthetics
- Must be a one-to-one relationship; each specific data value should map to only one aesthetic

Principles of Data Visualization

- Be Honest
- Principle of proportional ink
- Avoid unnecessary ‘chart junk’
- Use color judiciously
- Balance data and context

Extending Data Viz to Maps

Telling stories with maps



- Maps organize a lot of information in a coherent way
- They invite critique and inspection
- They are also aesthetic objects that can engage broader audiences

Key Issues

- Thinking about projections
- Scale of the map
- Errors of Omission

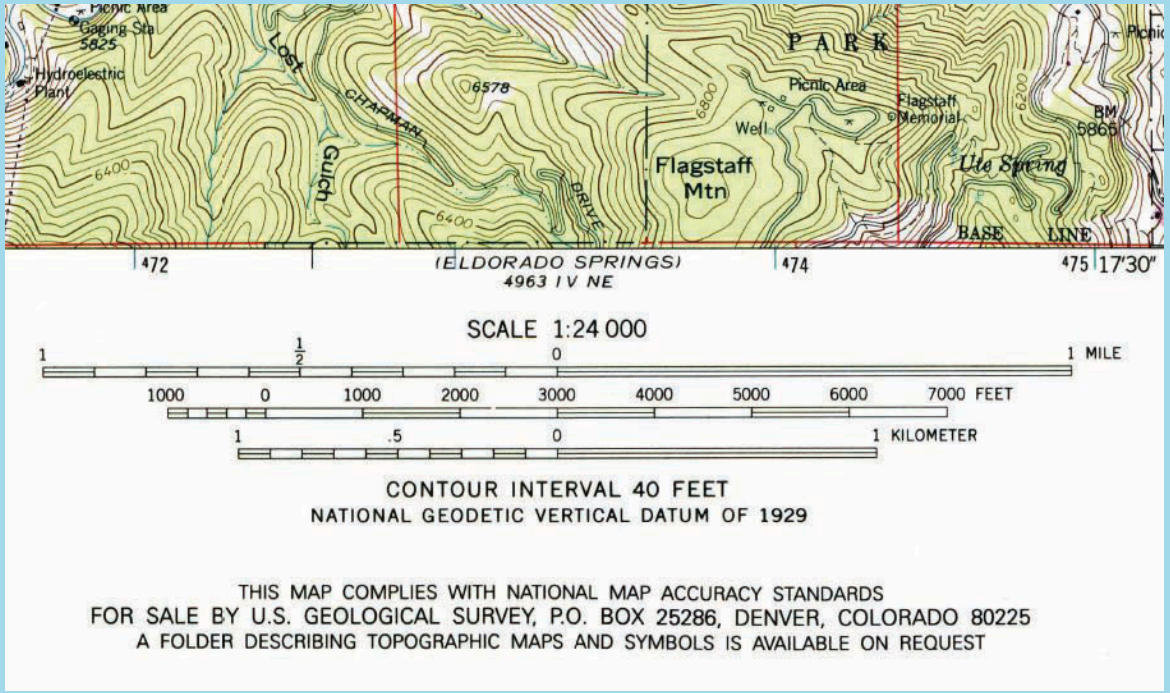
Cartographic Principles

1. Concept before compilation
2. Hierarchy with harmony (Important things should look important)
3. Simplicity from sacrifice
4. Maximum information at minimum cost
5. Engage emotion to enhance understanding

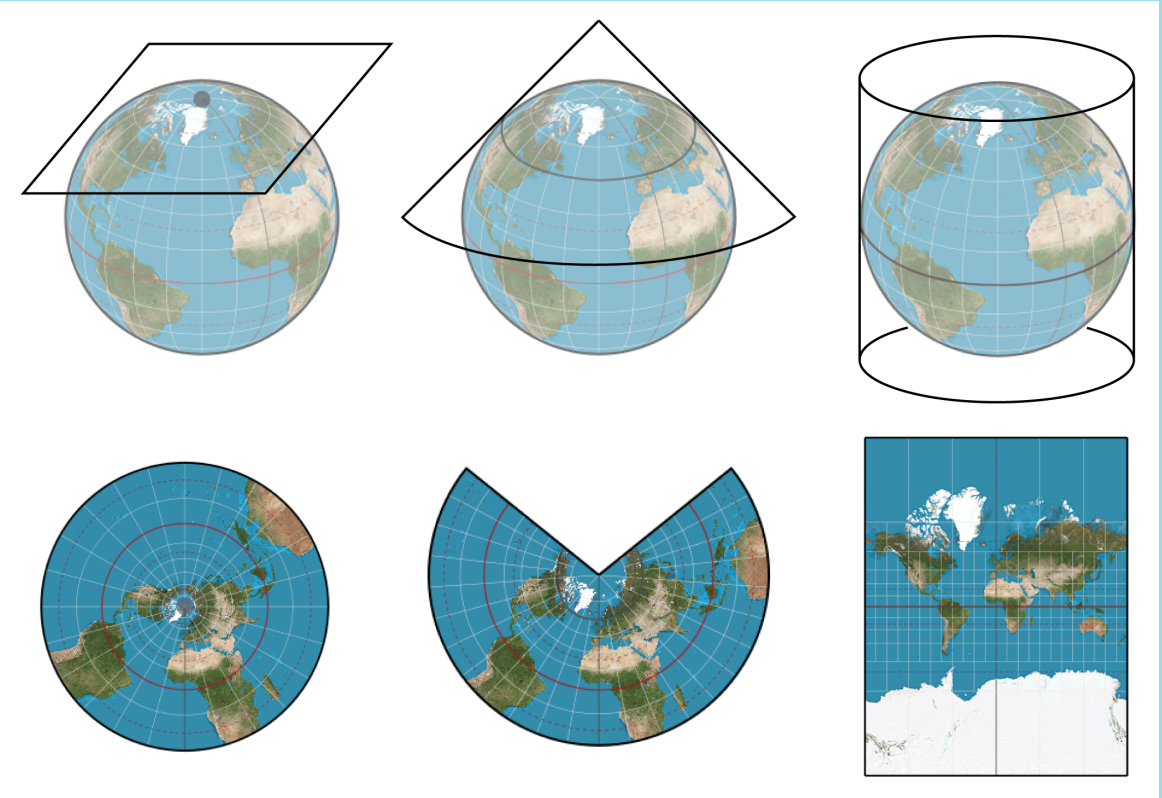
Map Elements

Scale

- Relates map distance to distance on the ground
- Ratio scales (1:24,000 or 1/24,000)
- Graphic scales
- Large vs. small-scale?



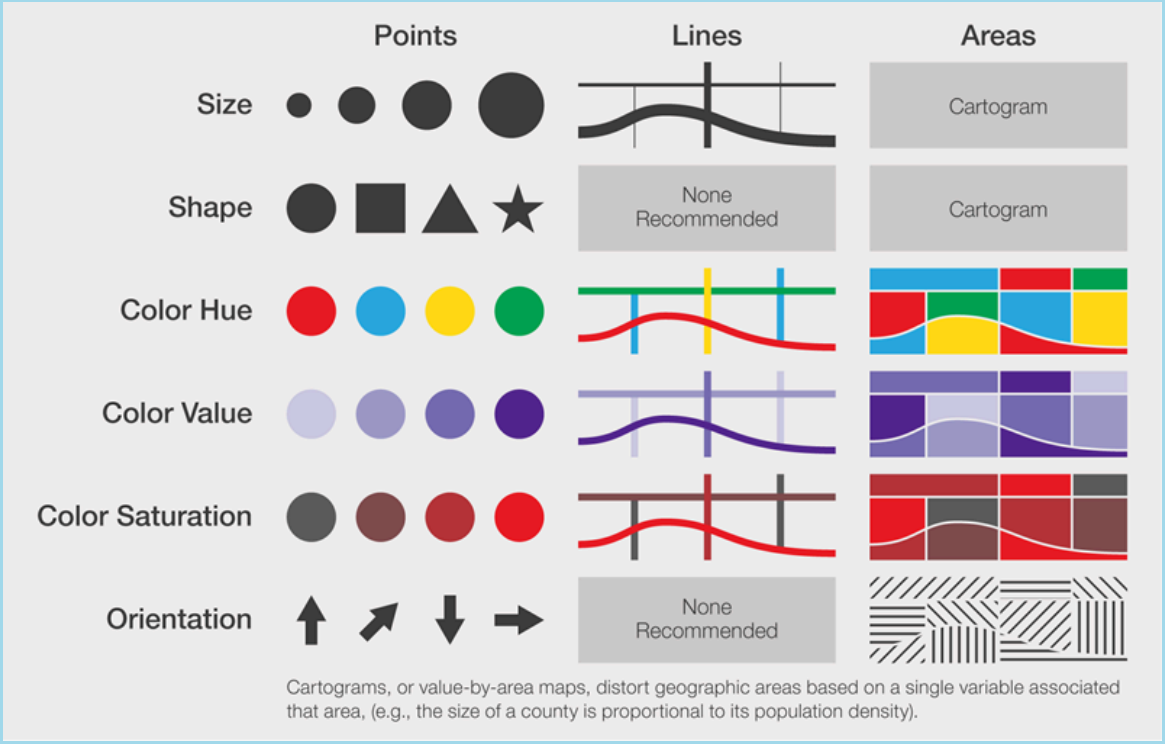
Projection



Developable Surfaces

- Distortion makes scale invalid across large areas
- Distortion increases with distance from standard line
- Five distortions: areas, angles, shapes, distances, and direction

Map Symbols

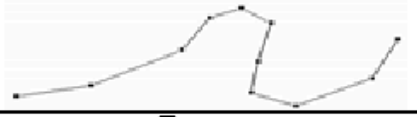
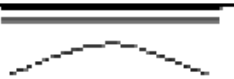
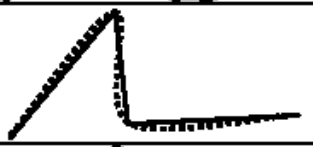


- Graphic code for retrieving information
- (De-)emphasize (un)important information
- Contrast and the role of colors

Generalization

> A good map tells a multitude of little white lies: it supresses truth to help the user see what needs to be seen…
> — Mark Monmonier

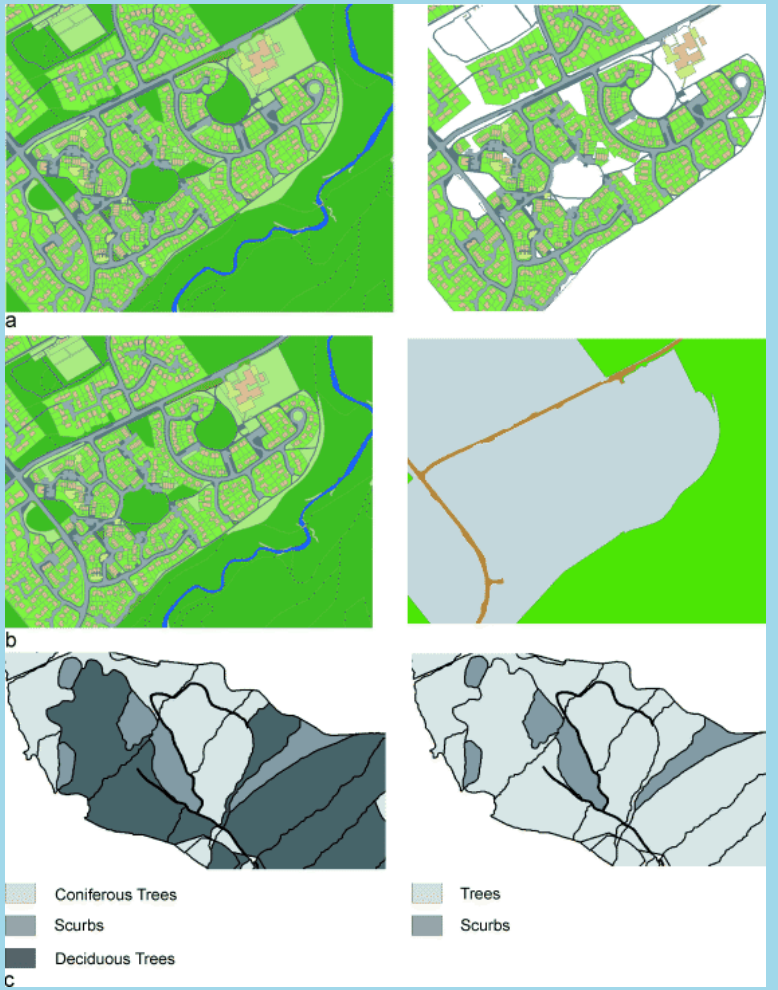
Geometry

Operations		Large-scale	Photo-reduced	Small-scale
Displacement				
Elimination				
(Scale-driven) generalisation				
Partial modification				
Point-reduction				
Smoothing	Curve-fitting			
	Filtering			
Typification				

Zhilin et al. 2008

Context

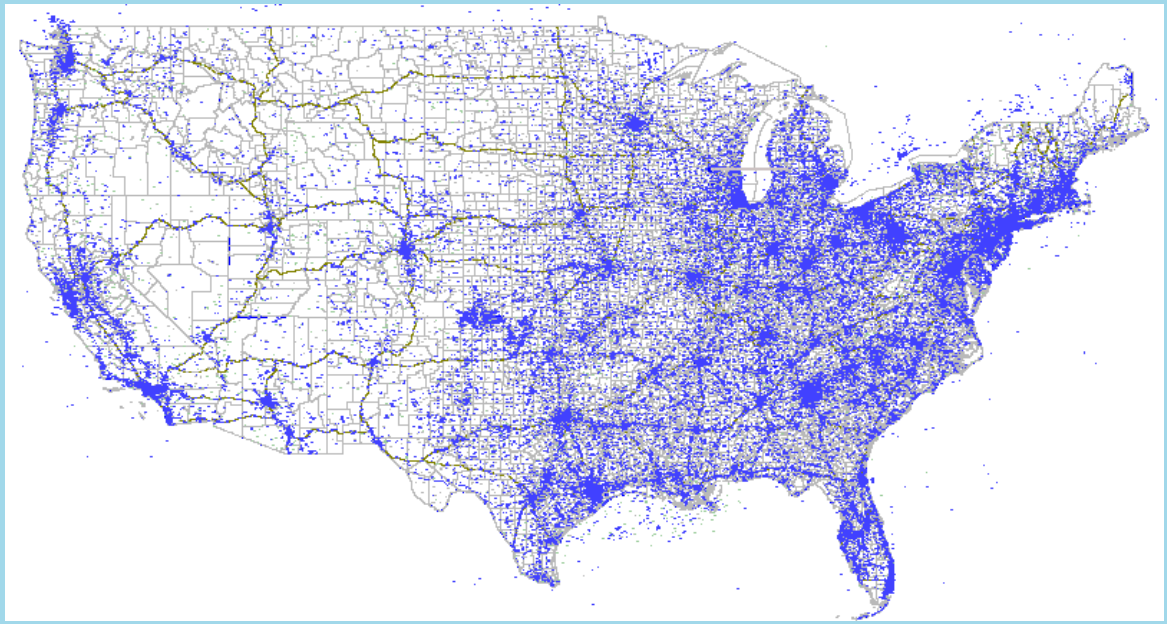
- Filter out irrelevant details
- Two elements: selection and classification
- Reflect interpretations of the relative importance of different features



Mackaness and Chaudry

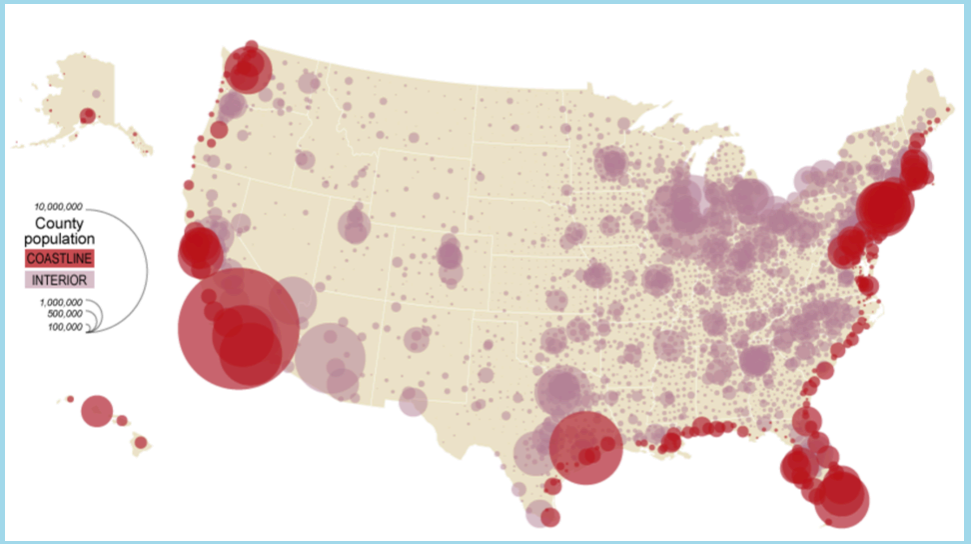
Data Maps

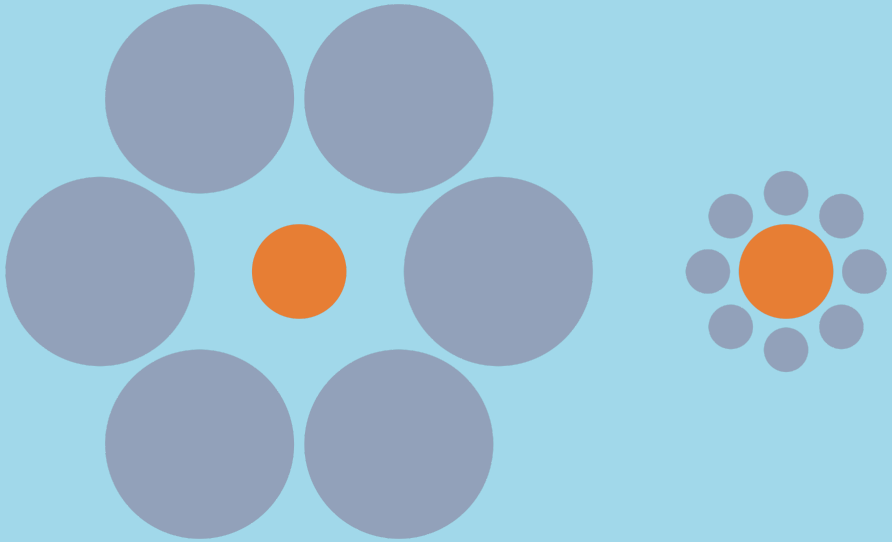
Point Maps



- Dot Maps: quantity represented by amount and concentration of dots
- Proportional Symbol Map: Geometric symbols scaled in proportion to a quantity

Ebbinghaus' illusion



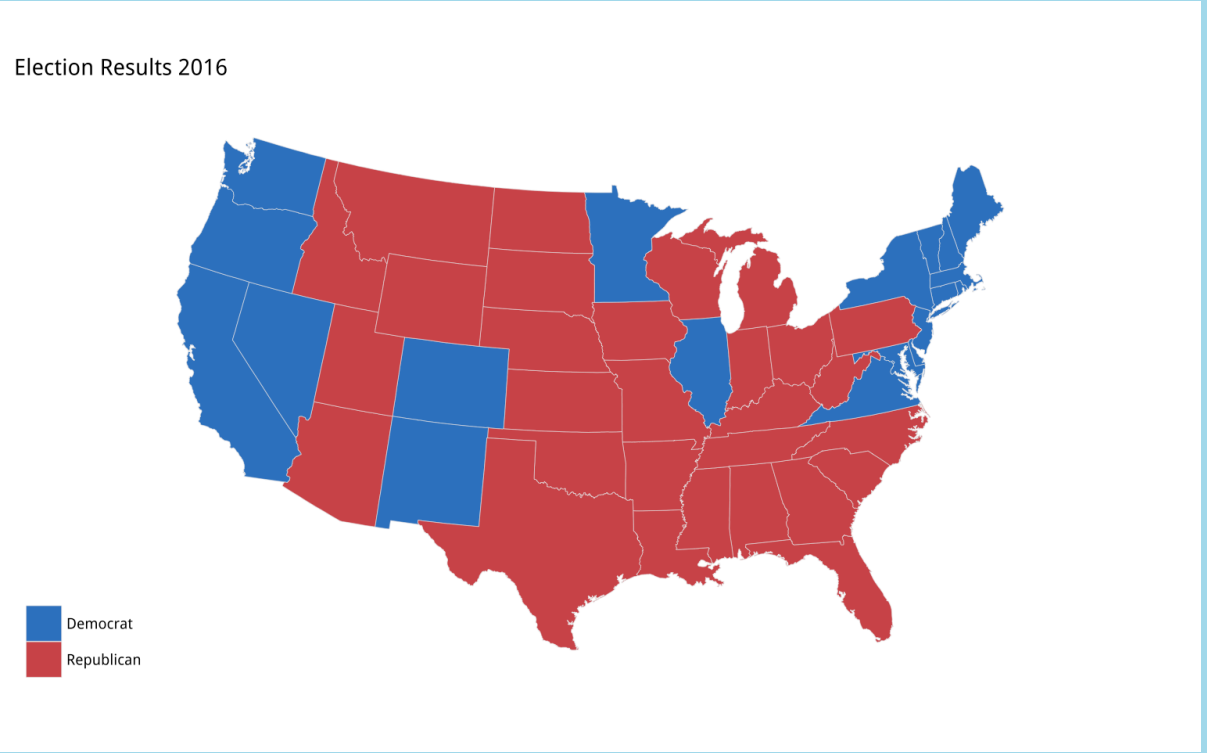


Line Maps

From High Country News

Choropleth

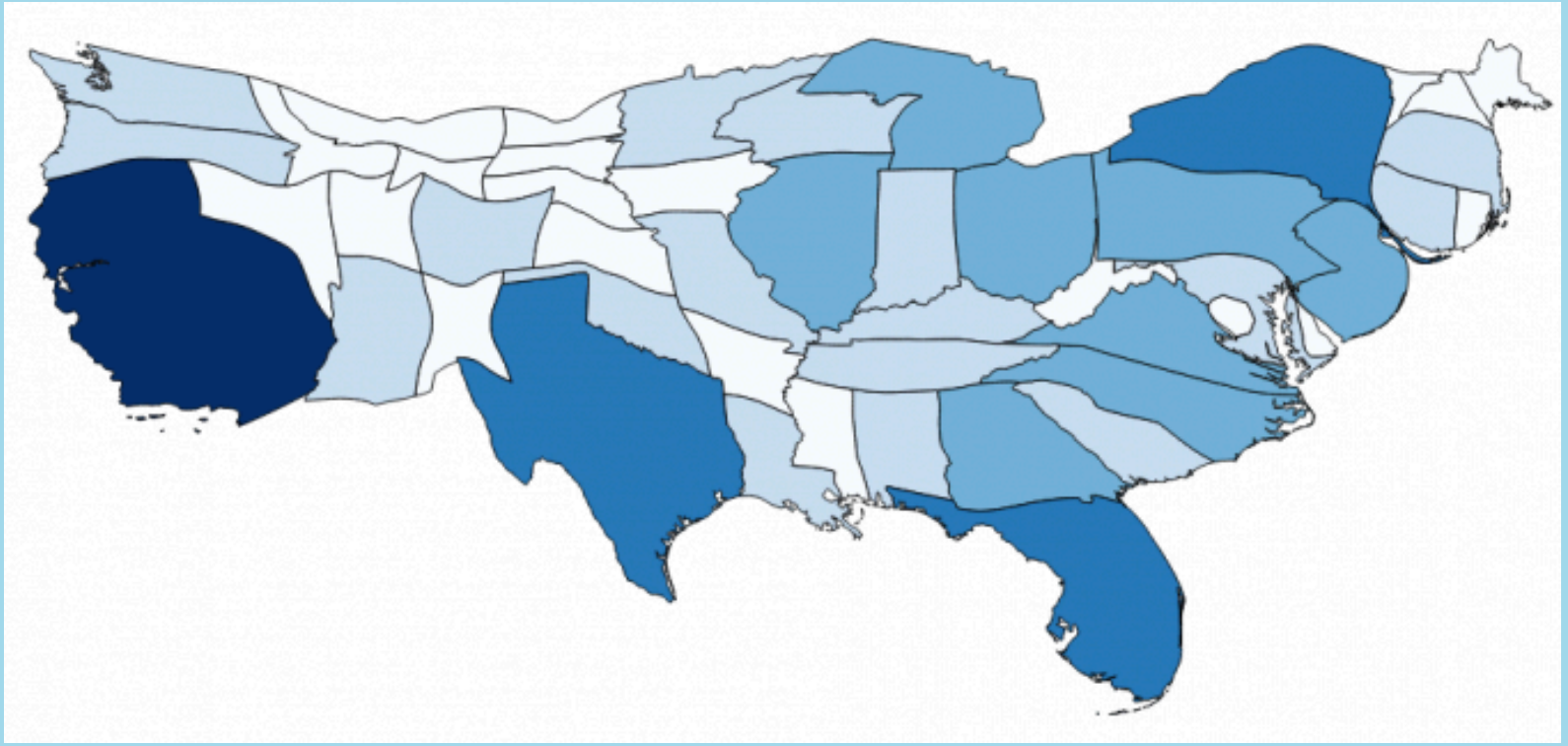
- Mapping color to geographies
- Common problems



From Healy 2019

Cartogram

- Adjusts for differences in area, population, etc
- Common Problems



From Healy 2019