SWE 432 -Web Application Development

Fall 2021

George Mason University

Dr. Kevin Moran

Week 14:

Information Visualization

Overview of Information Visualization

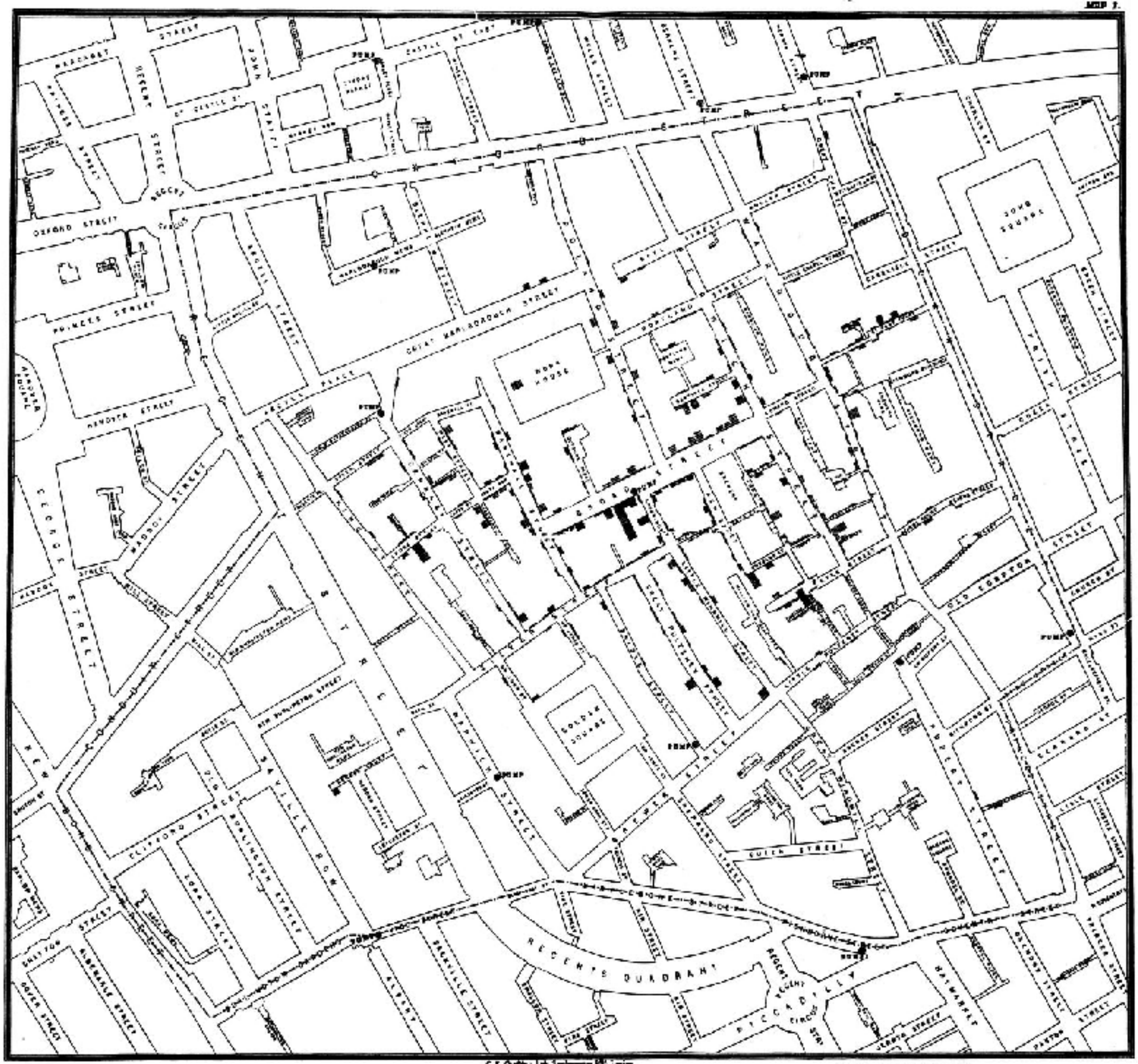
Today

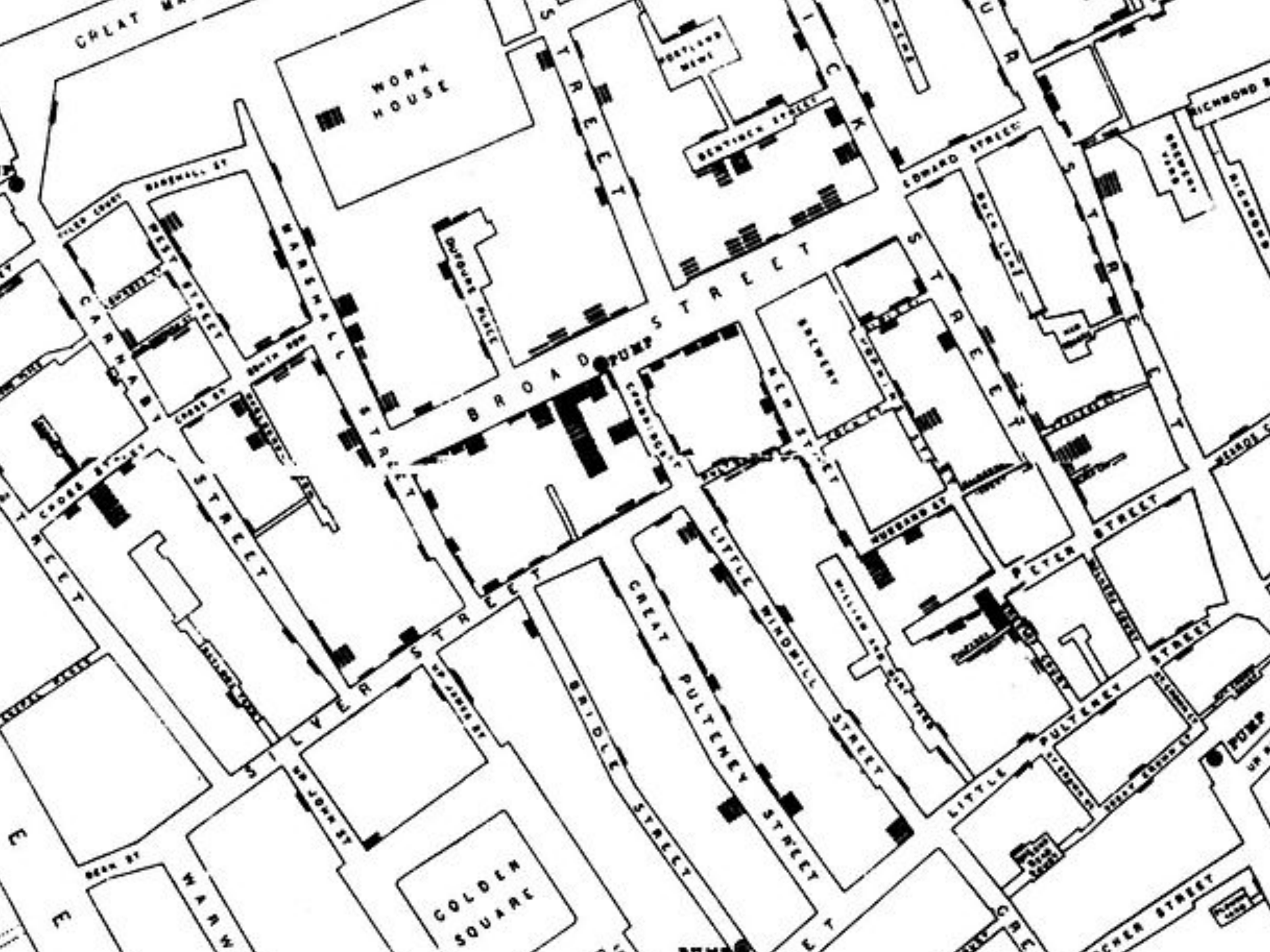
- What types of information visualization are there?
  - Which one should you choose?
- What principles and guidelines inform the design of information visualizations?
- How can interactivity be used to design better information visualizations?

Cholera Epidemic in London, 1854

- >500 fatal attacks of cholera in 10 days
  - Concentrated in Broad Street area of London
  - Many died in a few hours
- Dominant theory of disease: caused by noxious odors
- Afflicted streets deserted by >75% inhabitants

MAP 1.


OXFORD STREET
REGENT STREET
PRINCES STREET
GREAT MARLBOROUGH STREET
HANOVER STREET
BROAD STREET
POLAND STREET
BERWICK STREET
GOLDEN SQUARE
SOHO SQUARE
REGENT'S QUADRANT
PICCADILLY CIRCUS
PICCADILLY
HAYMARKET
CLIFFORD STREET
NEW BURLINGTON STREET
SAVILE ROW
OLD BURLINGTON STREET
BONVIGEDN STREET
CORK STREET
ALBEMARLE STREET
DOVER STREET
BRUTON STREET
BURLINGTON GARDENS
ARGYLL PLACE
MADDOX STREET
GEORGE STREET
HANOVER SQUARE
BROADWICK STREET
GOLDEN SQUARE
QUICK STREET
BEAK STREET
GREAT PULTENEY STREET
LITTLE PULTENEY STREET
CHARLES ST
CRANBOURN STREET
OLD COMPTON STREET
GERRARD STREET
SHAFTESBURY AVENUE
D'ARBLAY STREET
LEXINGTON STREET
BROWER STREET
RUPERT STREET
WARWICK STREET
GLASSHOUSE STREET
REGENT STREET
CHURCH ST
CHURCH ST
CHURCH ST



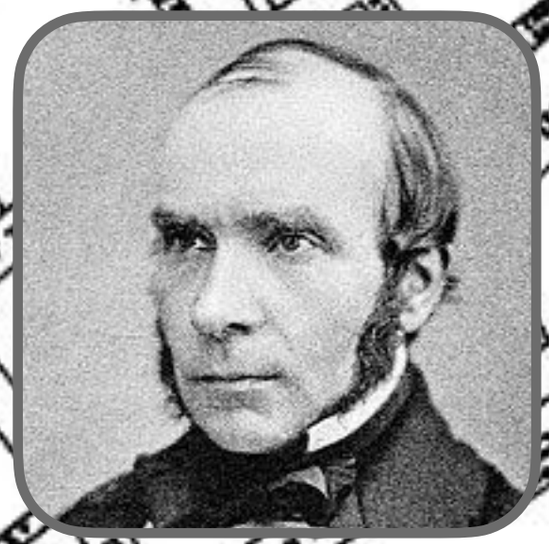
WORK HOUSE
BROAD STREET
PUMP
GREAT PULTENEY STREET
LITTLE WINDMILL STREET
BRIDLE STREET
GOLDEN SQUARE
SILVER STREET
PETER STREET
LITTLE PULTENEY STREET
NEW STREET
BREWERY
EDWARD STREET
RICHMOND
MARSHALL STREET
WEST STREET
CARNABY
CROSS STREET
UP JOHN ST
UP JAMES ST
DEAN ST

WORK HOUSE

BROAD STREET

PUMP

GOLDEN SQUARE

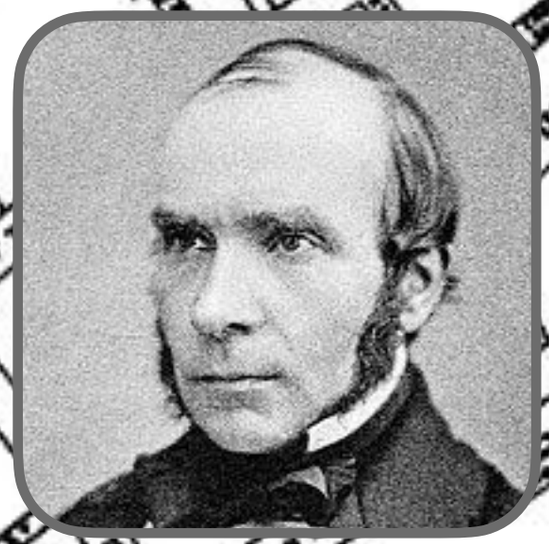


WORK HOUSE

BROAD STREET

PUMP

GOLDEN SQUARE



Investigation and Aftermath

- Based on ***visualization***, did case by case investigation
- Found that ***61 / 83*** positive identified as using well water from Broad Street pump
- Board ordered pump-handle to be removed from well
- Epidemic soon ***ended***
- Solved centuries old question of how cholera spread

Methods used by Snow

- Placed data in appropriate ***context*** for assessing cause & effect
  - Plotted on map, included well location
  - Reveals proximity as cause
- Made quantitative ***comparisons***
  - Fewer deaths closer to brewery, could investigate cause
- Considered ***alternative*** explanations & contrary cases
  - Investigated cases not close to pump, often found connection to pump
- Assessment of possible ***errors*** in numbers

Amplifying Cognition

Amplifying Cognition

- Information Visualization can amplify cognition by:

Amplifying Cognition

- Information Visualization can amplify cognition by:

  1. Increasing the memory and processing resources available to users

Amplifying Cognition

- Information Visualization can amplify cognition by:

  1. Increasing the memory and processing resources available to users

  2. Reducing the search for information

Amplifying Cognition

- Information Visualization can amplify cognition by:

  1. Increasing the memory and processing resources available to users

  2. Reducing the search for information

  3. Using visual representations to enhance the detection of patterns

Amplifying Cognition

- Information Visualization can amplify cognition by:

  1. Increasing the memory and processing resources available to users

  2. Reducing the search for information

  3. Using visual representations to enhance the detection of patterns

  4. Enabling perceptual inference

Amplifying Cognition

- Information Visualization can amplify cognition by:

  1. Increasing the memory and processing resources available to users

  2. Reducing the search for information

  3. Using visual representations to enhance the detection of patterns

  4. Enabling perceptual inference

  5. Using perceptual attention mechanisms for monitoring

Amplifying Cognition

- Information Visualization can amplify cognition by:

  1. Increasing the memory and processing resources available to users

  2. Reducing the search for information

  3. Using visual representations to enhance the detection of patterns

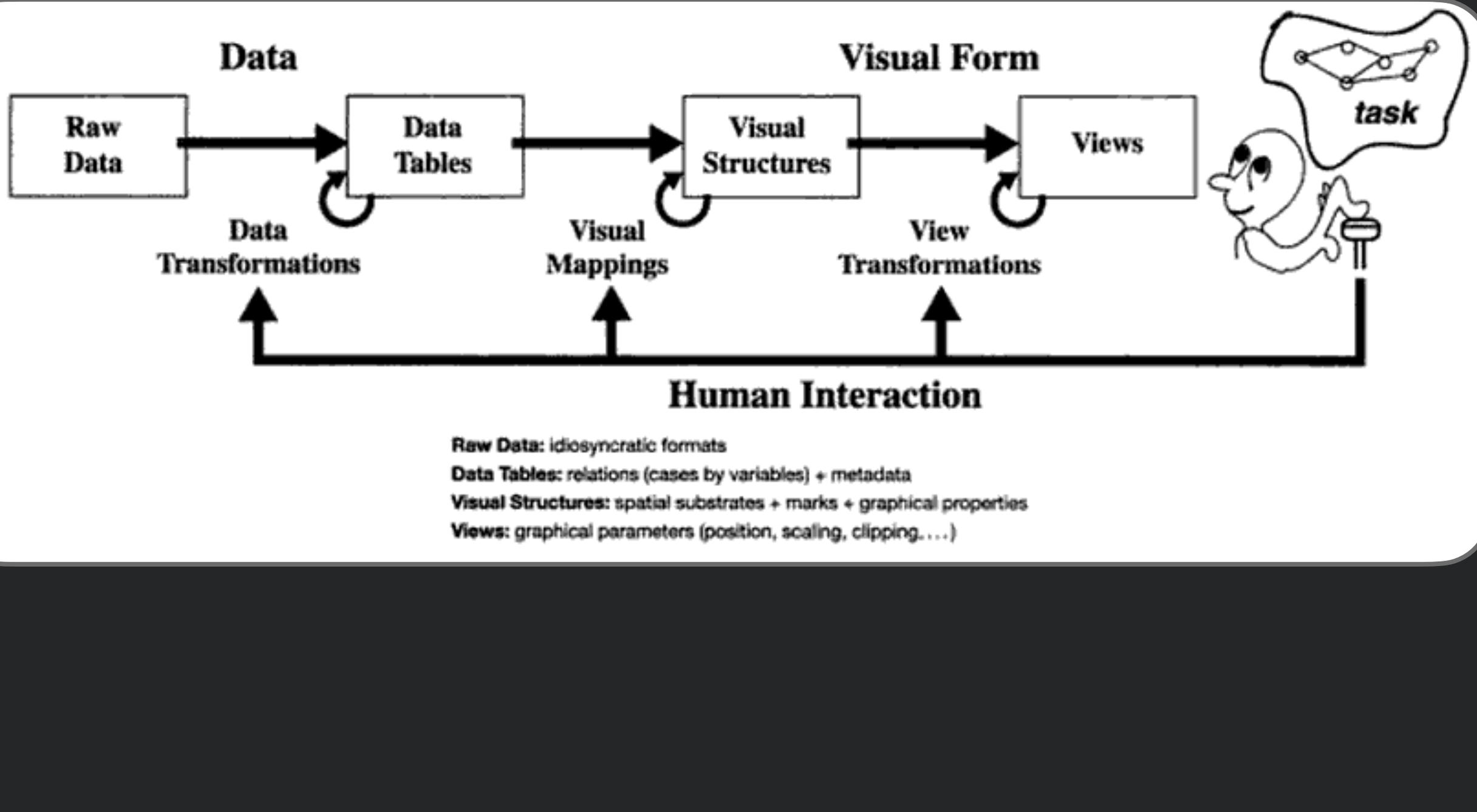
  4. Enabling perceptual inference

  5. Using perceptual attention mechanisms for monitoring

  6. Encoding Information in a manipulable medium

Mapping Data to Visual Form

Designing an Information Visualization



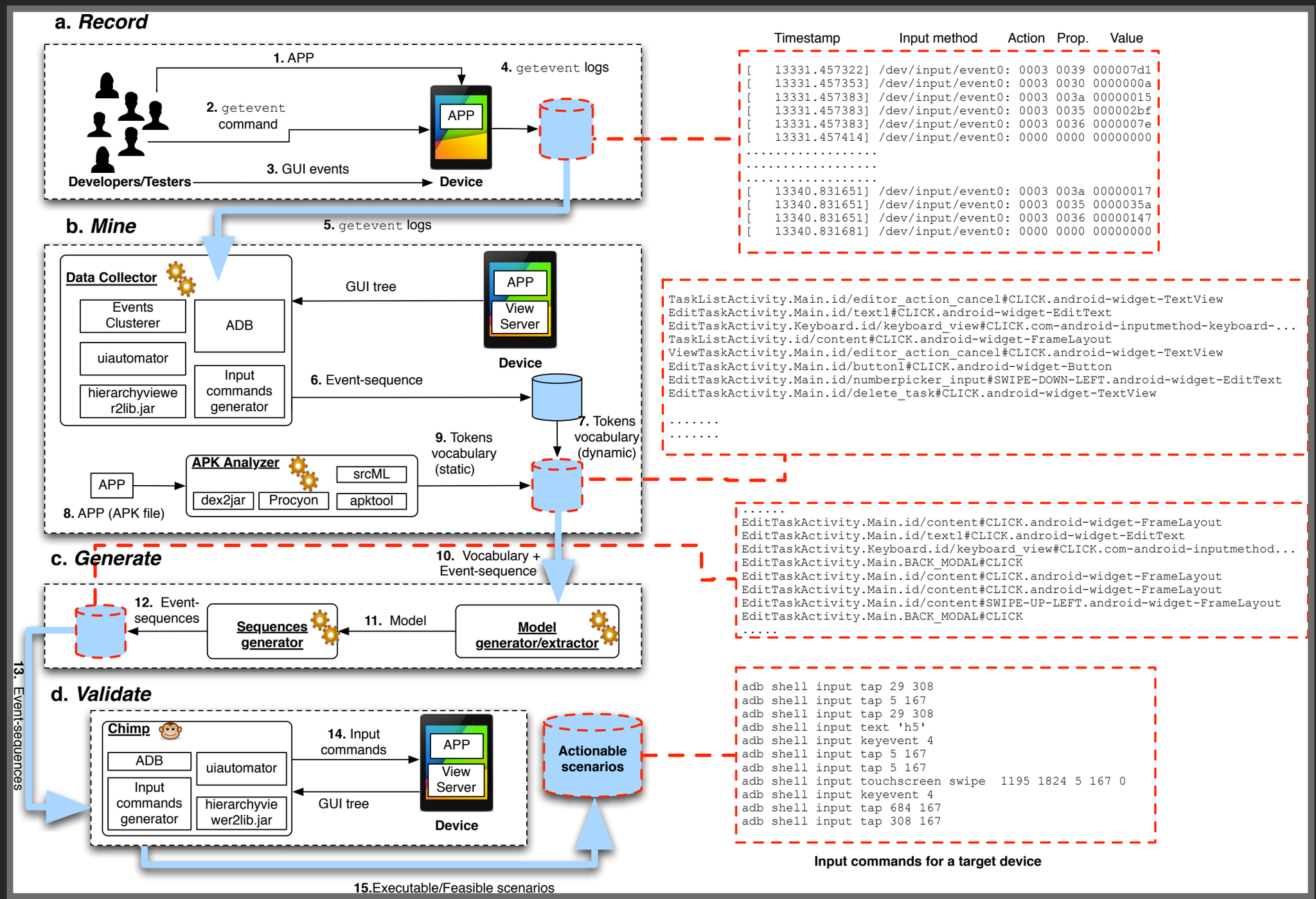
Types of Raw Data

- Nominal - unordered set ***without*** a quantitative value
  - Gender: male, female
  - Hair color: brown, black, blonde, gray, orange, ...
- Ordinal - ***ordered*** set, with no meaning assigned to differences
  - How do you feel today: very unhappy, unhappy, ok, happy, very happy
  - Undefined how much better happy is than ok
- Quantitative - ***numeric*** value
  - Height, weight, distance, ...

Data Transformations

- Classing / binning: Quantitative —> ordinal
  - Maps ranges onto ***classes*** of variables
    - Can also count # of items in each class w/ histogram
- Sorting: Nominal —> ordinal
  - Add order between items in sets
- Descriptive statistics: mean, average, median, max, min, …

Example Uses of a Data Transformation

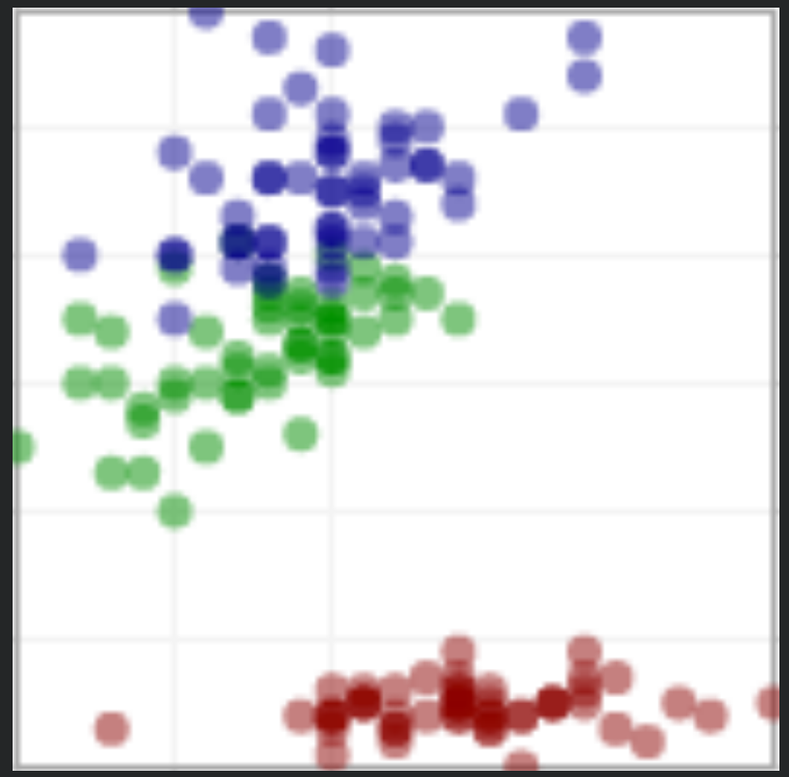


Visual Structures

- 3 components
  - spatial substrate
  - marks
  - marks’ graphical properties

Spatial Substrate

- Axes that divide space
- Types of axes - unstructured, nominal, ordinal, quantitative
- Composition - use of multiple orthogonal axes (e.g., 2D scatterplot, 3D)



Marks

- Points (0D)
- Lines (1D)
- Areas (2D)
- Volumes (3D)

Marks' Graphical Properties

- Quantitative (Q), Ordinal (O), Nominal (N)
- Filled circle - good; open circle - bad

	Spatial	Object
Extent	(Position) —|—|—|—| Size ● ● • ·	Gray Scale ■ ■ ■ □
Dif-ferential	Orientation – / | \	Color ■ ■ ■ ■ Texture ■ ■ ■ ■ Shape ■ ★ ● ◆

Effectiveness of Graphical Properties

- Quantitative (Q), Ordinal (O), Nominal (N)
- Filled circle - good; open circle - bad

	Spatial	Q	O	N	Object	Q	O	N
Extent	(Position)	●	●	●	Grayscale	◑	●	○
	Size	●	●	●				
Differential	Orientation	◑	◑	●	Color	◑	◑	●
					Texture	◑	◑	●
					Shape	○	○	●

Animation

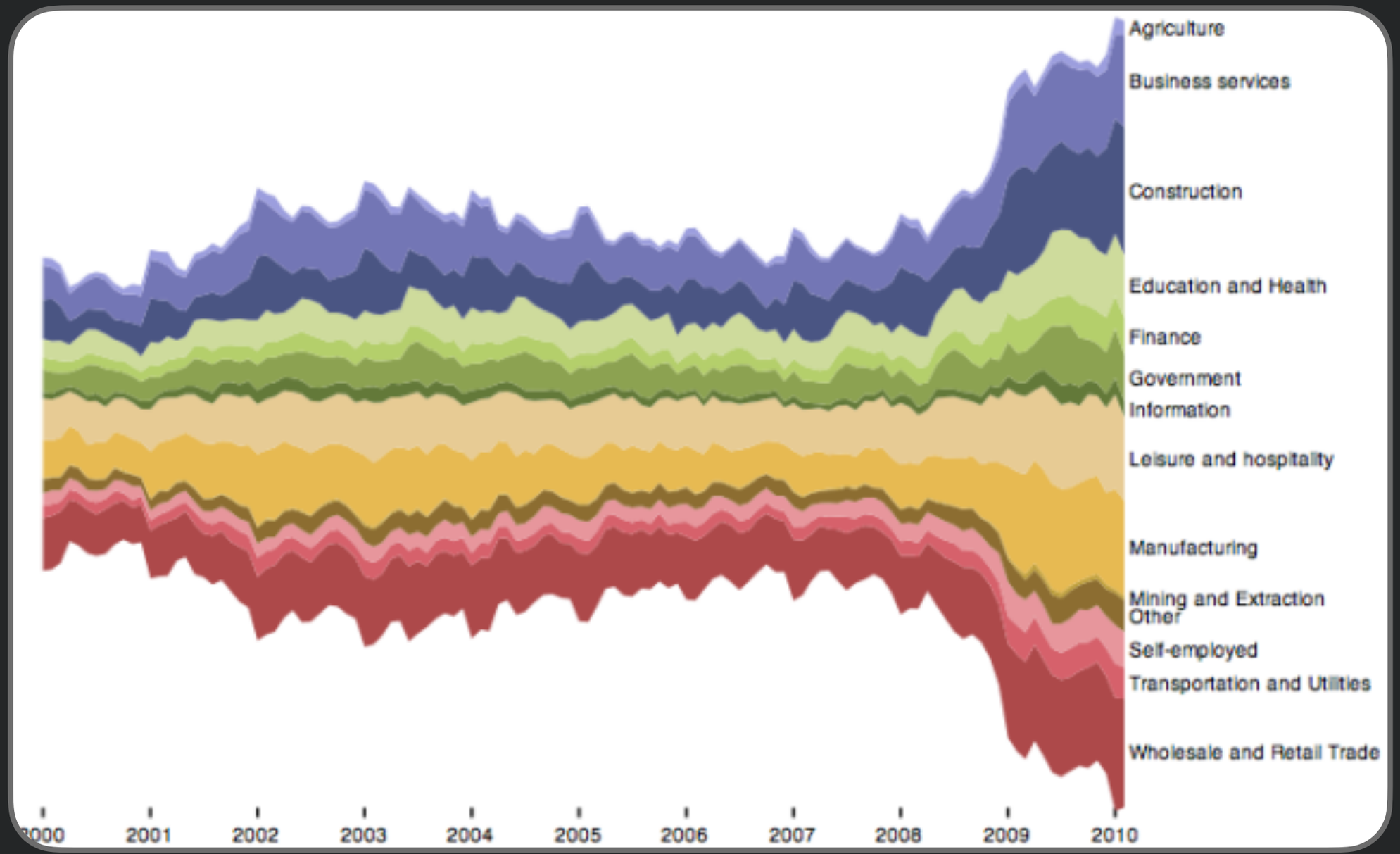
- Visualization can change over time
- Could be used to encode data as a function of time
  - But often not effective as makes direct comparisons hard
- Can be more effective to animate transition from before to after as user configures visualization

Examples of Visualizations

Time-series Data

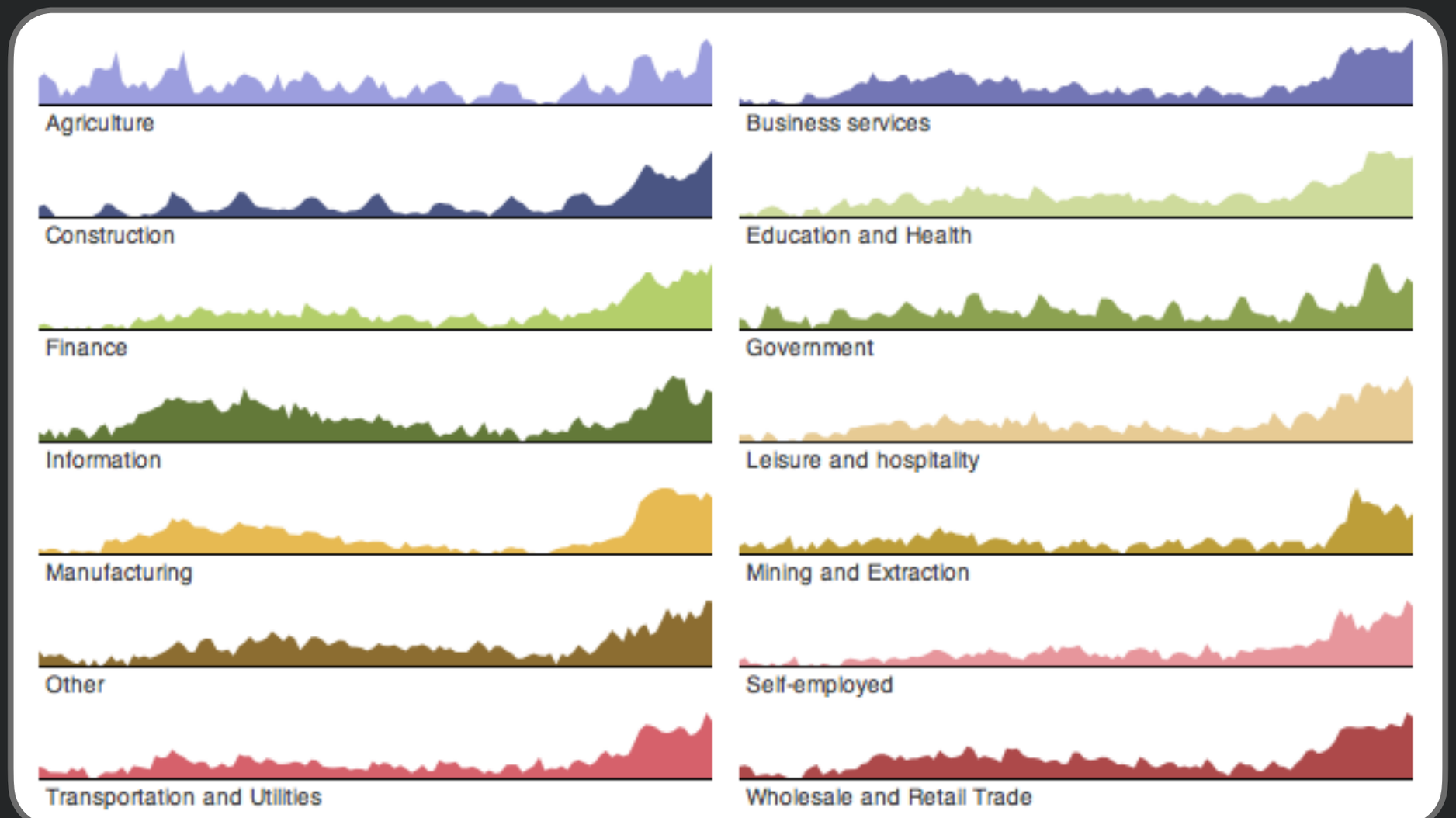
Stacked Graph

- Supports visual summation of multiple components



Small Multiples

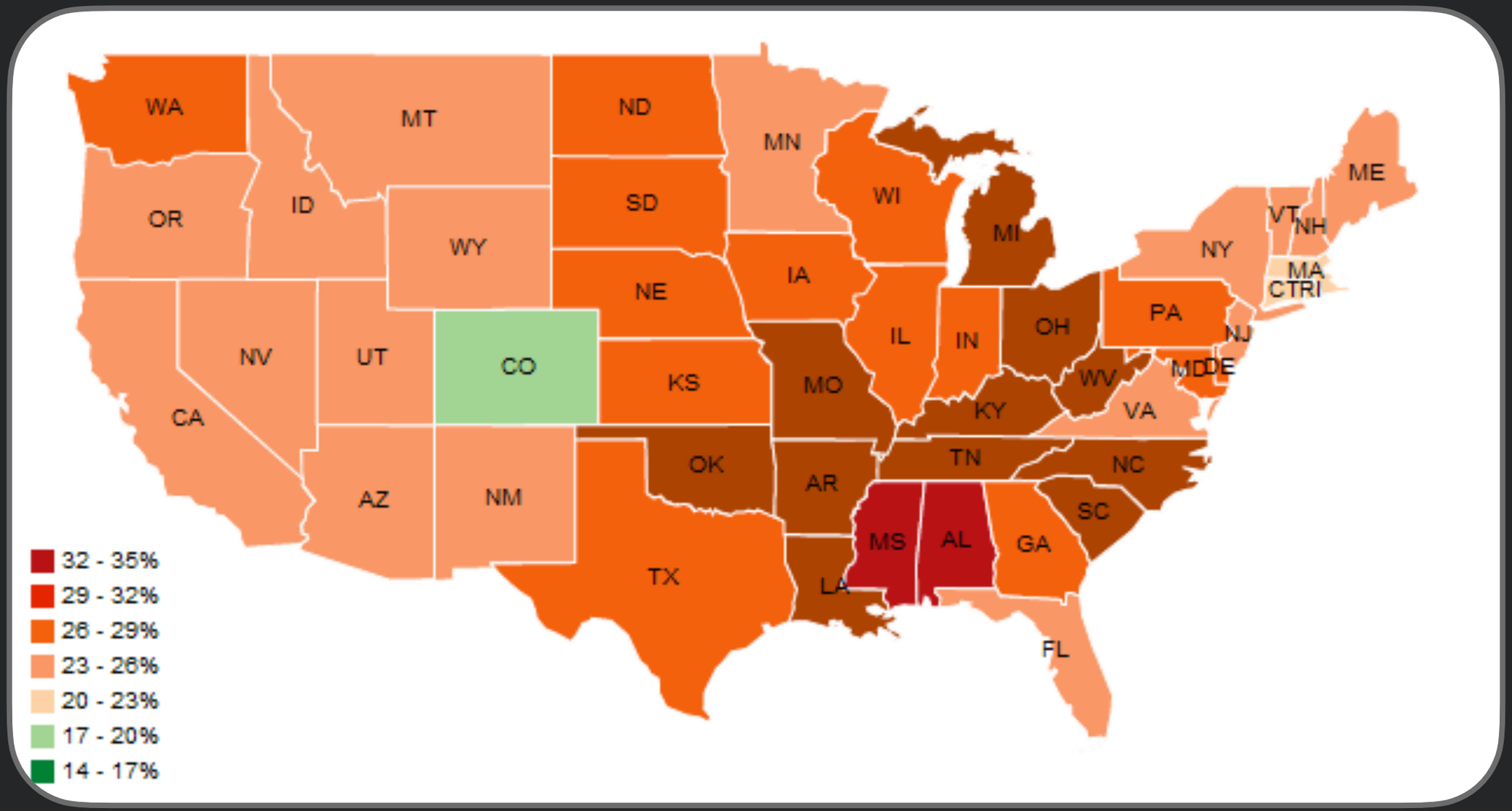
- Supports separate comparison of data series
- May have better legibility than placing all in single plot



Maps

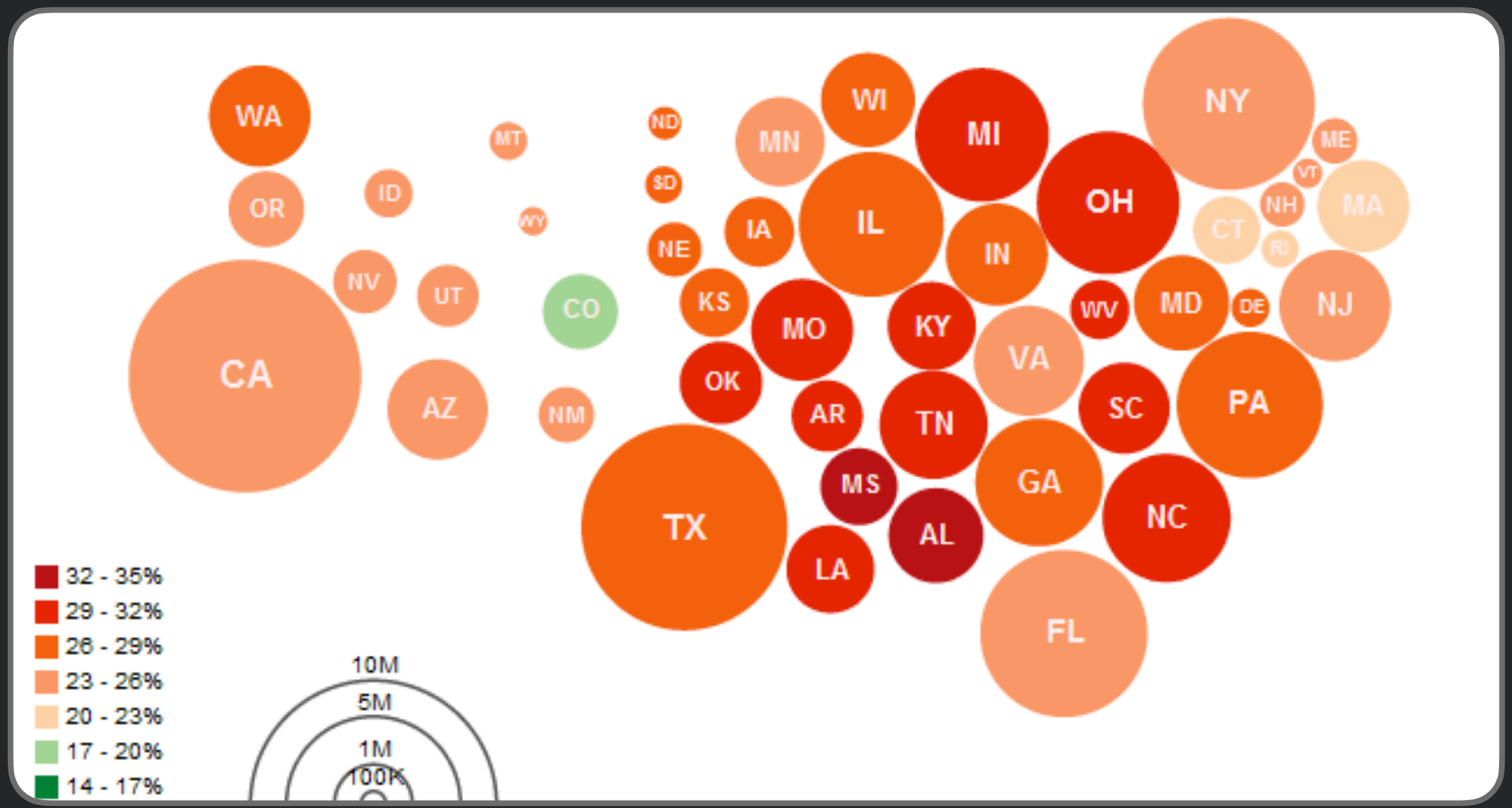
Choropleth Map

- Groups data by area, maps to color



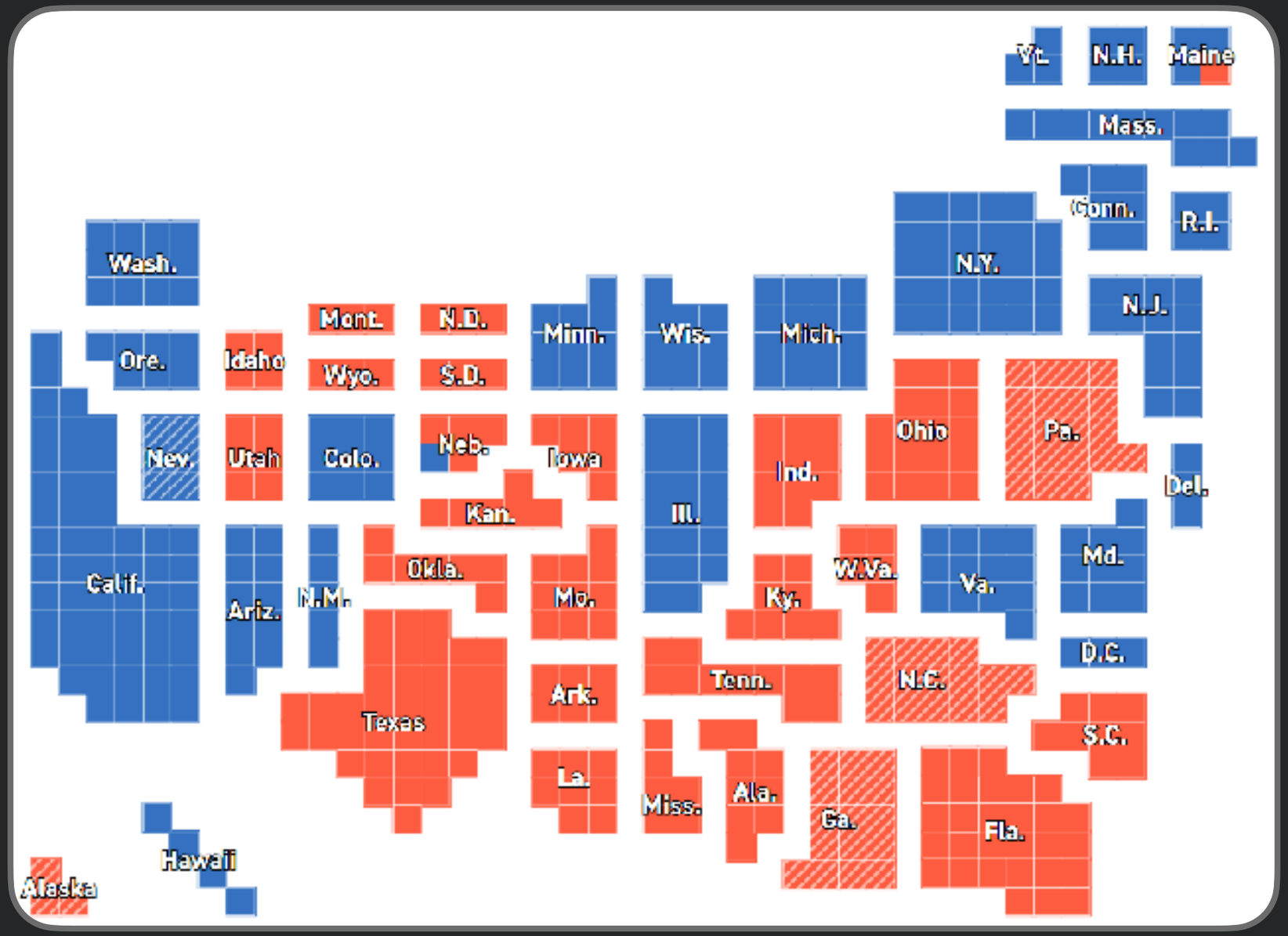
Cartograms

- Encodes two variables w/ size & color



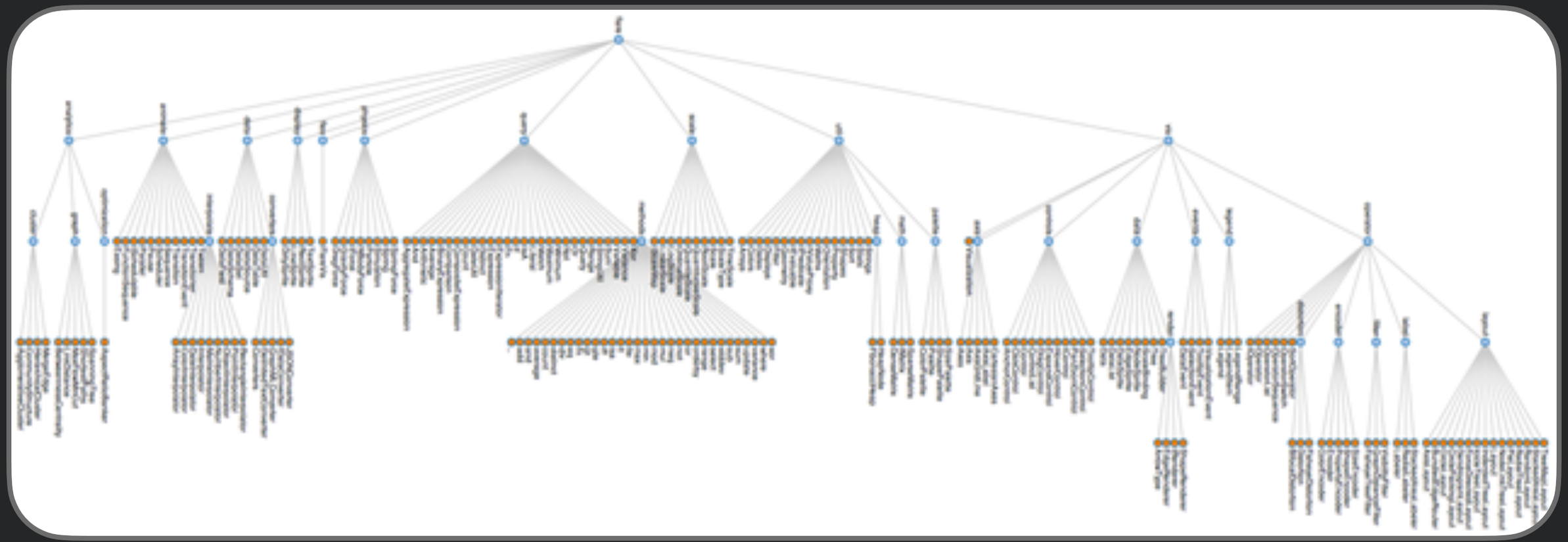
Cartograms

- Encodes two variables w/ size & color



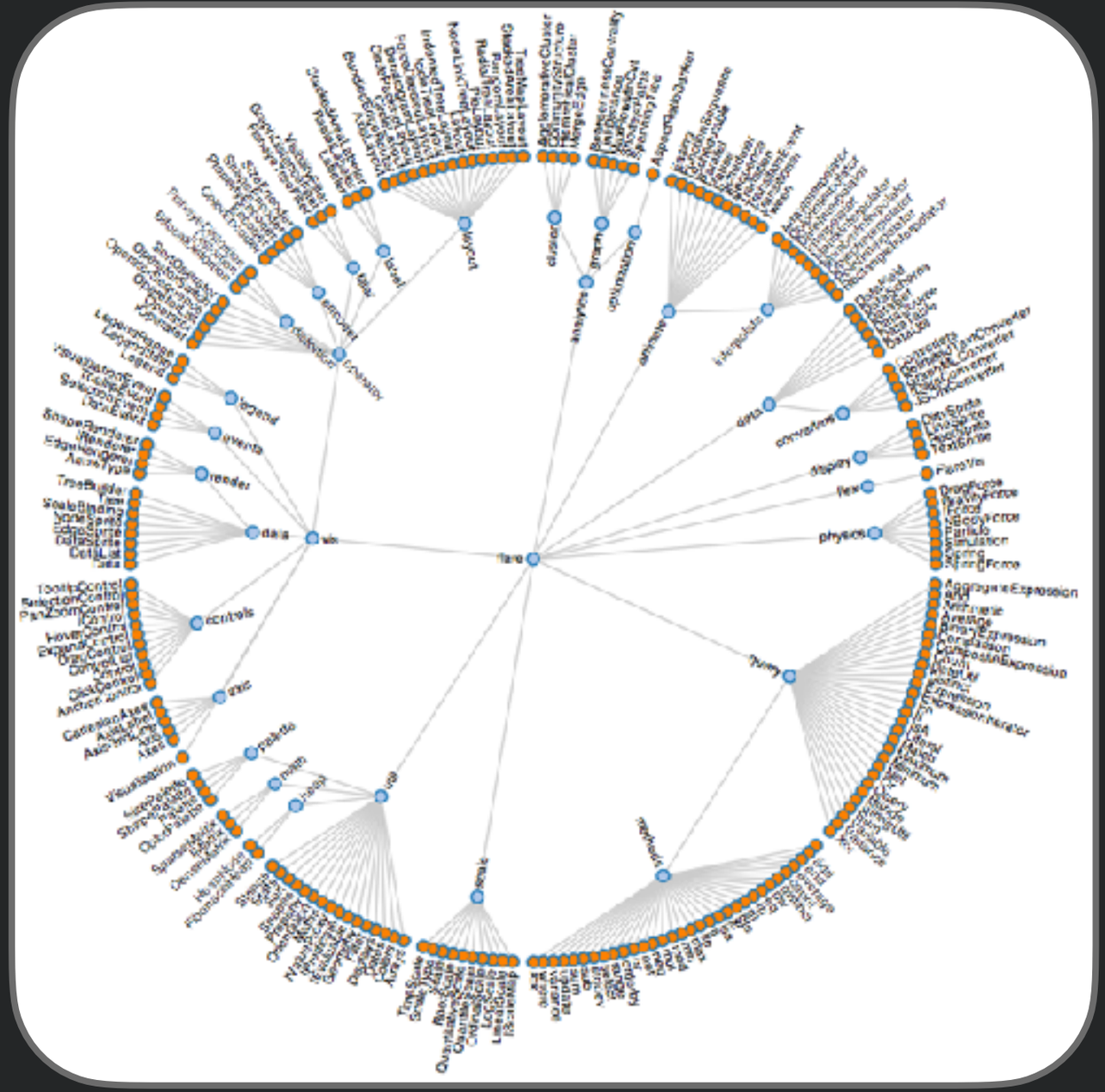
Hierarchies

Node Link Diagram

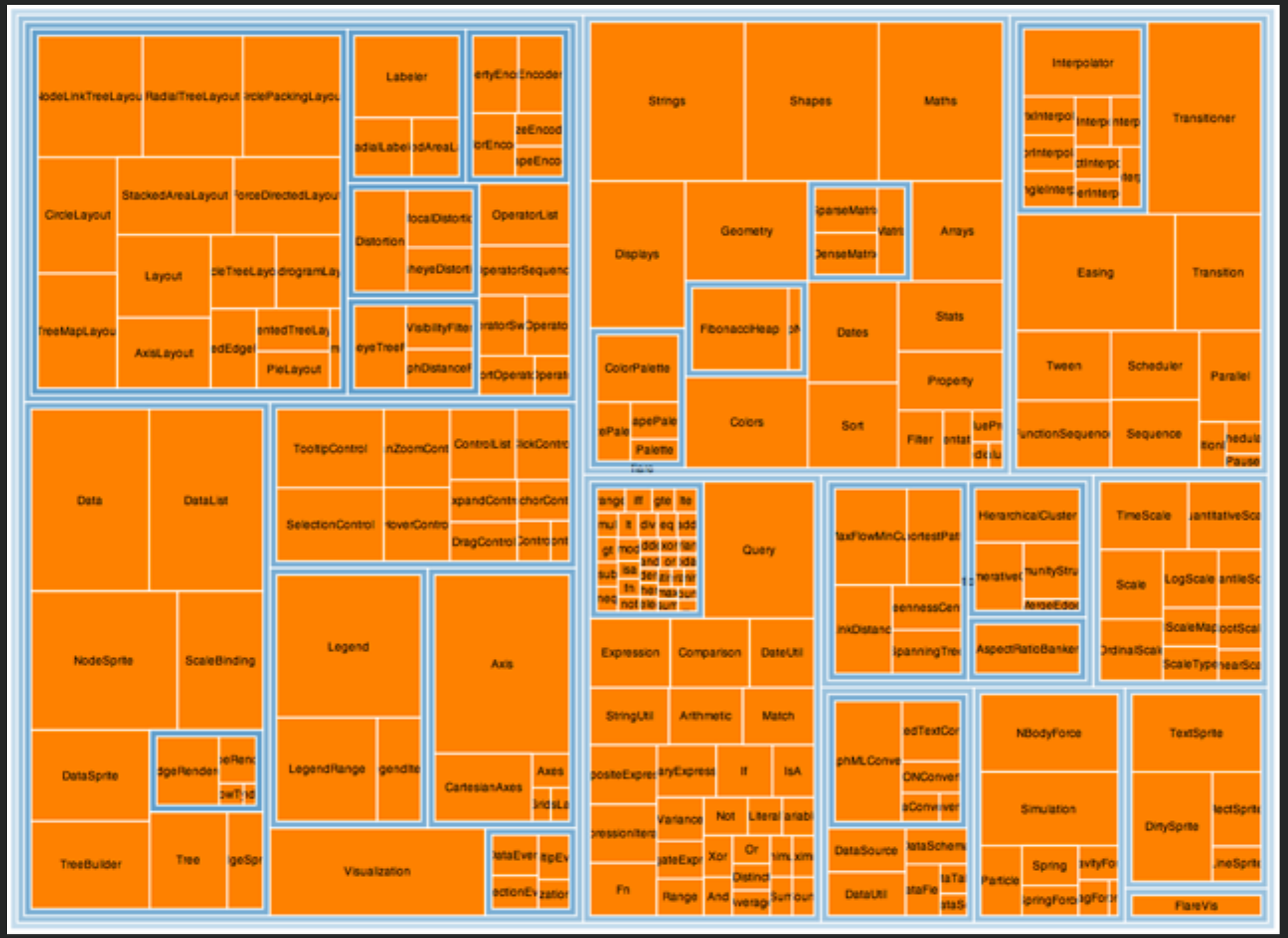


Dendrogram

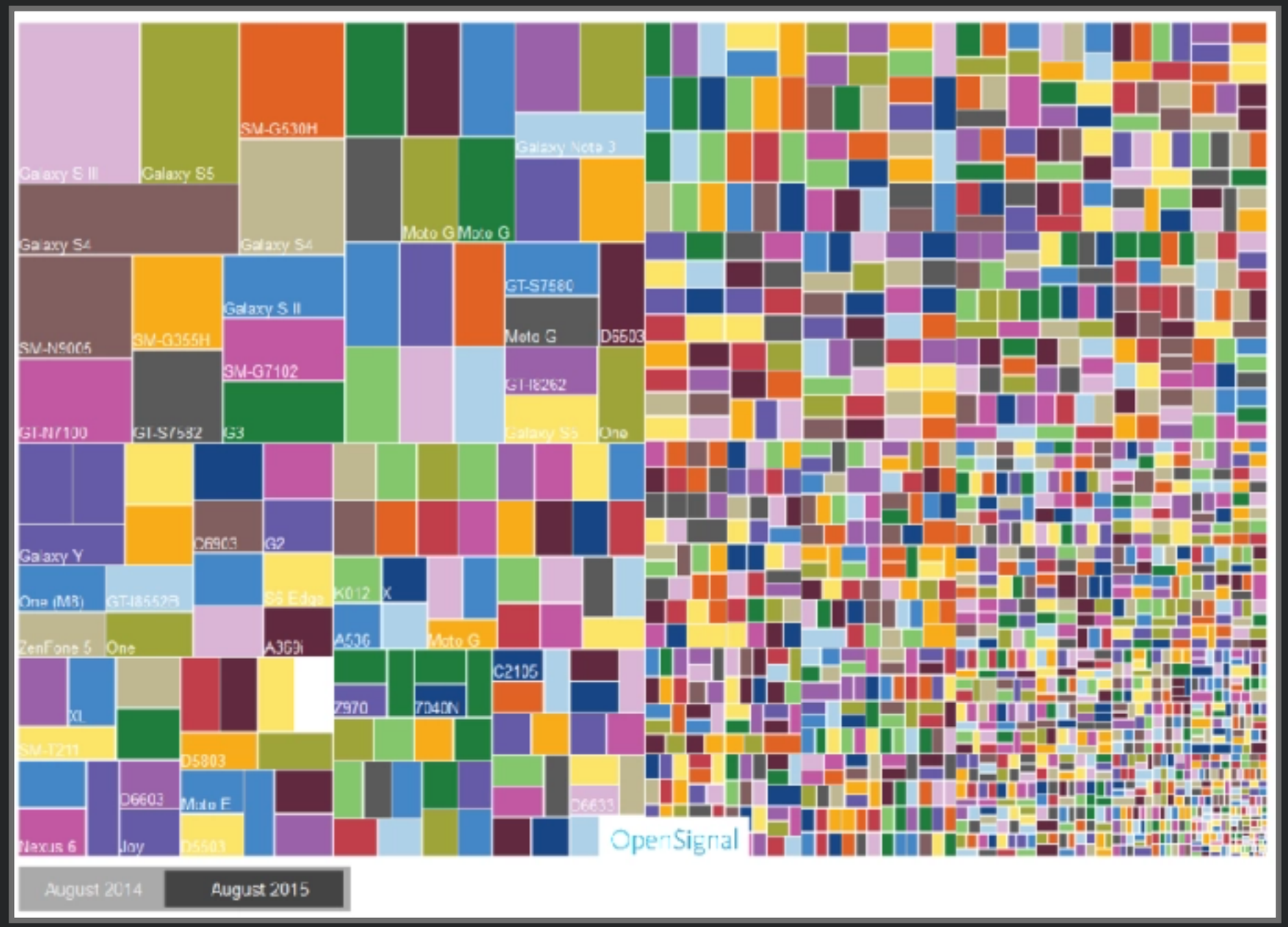
- Leaf nodes of hierarchy on edges of circle



Treemaps


Labeler
Strings
Shapes
Maths
Interpolator
Transitioner
CircleLayout
StackedAreaLayout
Layout
AxisLayout
PieLayout
Distortion
OperatorList
Displays
Geometry
Arrays
Easing
Transition
FibonacciHeap
Dates
Stats
Tween
Scheduler
Parallel
ColorPalette
Palette
Colors
Sort
Property
Filter
Sequence
Data
DataList
NodeSprite
ScaleBinding
DataSprite
TreeBuilder
Tree
Visualization
TooltipControl
ControlList
SelectionControl
DragControl
Legend
LegendRange
Axis
CartesianAxes
Axes
Query
Expression
Comparison
DateUtil
StringUtil
Arithmetic
Match
Variance
Not
Or
Xor
Distinct
Range
And
IsA
Fn
HierarchicalCluster
AspectRatioBanker
TimeScale
Scale
LogScale
ScaleType
NBodyForce
Simulation
Spring
Particle
TextSprite
DirtySprite
FlareVis
DataSource
DataUtil

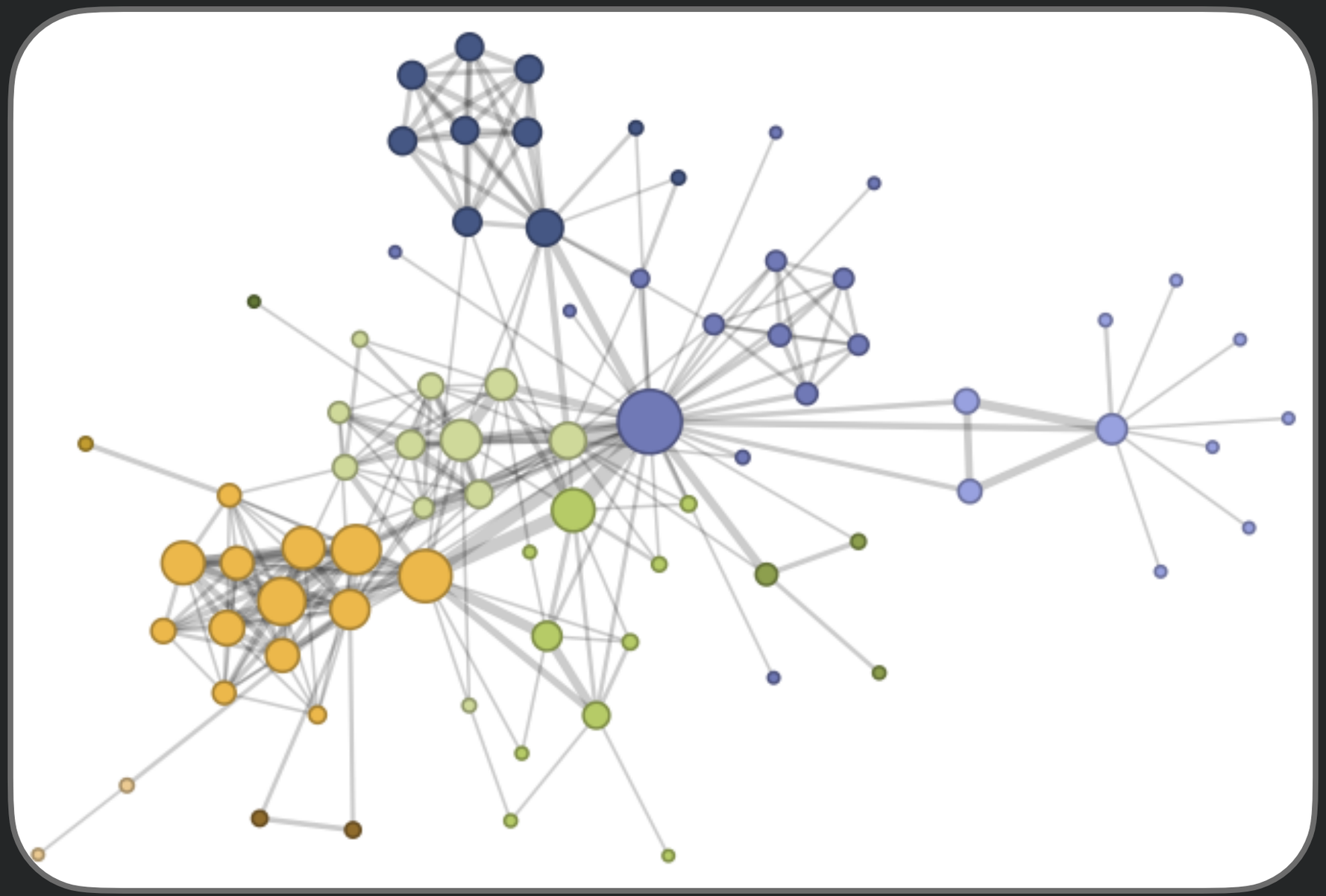
Treemaps


Galaxy S III
Galaxy S5
SM-G530H
Galaxy S4
Galaxy S4
Galaxy Note 3
Moto G
Moto G
GT-S7580
Galaxy S II
SM-N9005
SM-G355H
SM-G7102
Moto G
D5503
GT-I8262
GT-N7100
GT-S7582
G3
Galaxy S5
One
Galaxy Y
D6603
G2
One (M8)
GT-I8552B
S5 Edge
K012
X
ZenFone 5
One
A369i
A536
Moto G
C2105
Z970
T040N
XL
SM-T211
D5803
D6603
Moto E
D6633
Nexus 6
Joy
D5503
OpenSignal
August 2014
August 2015

Networks

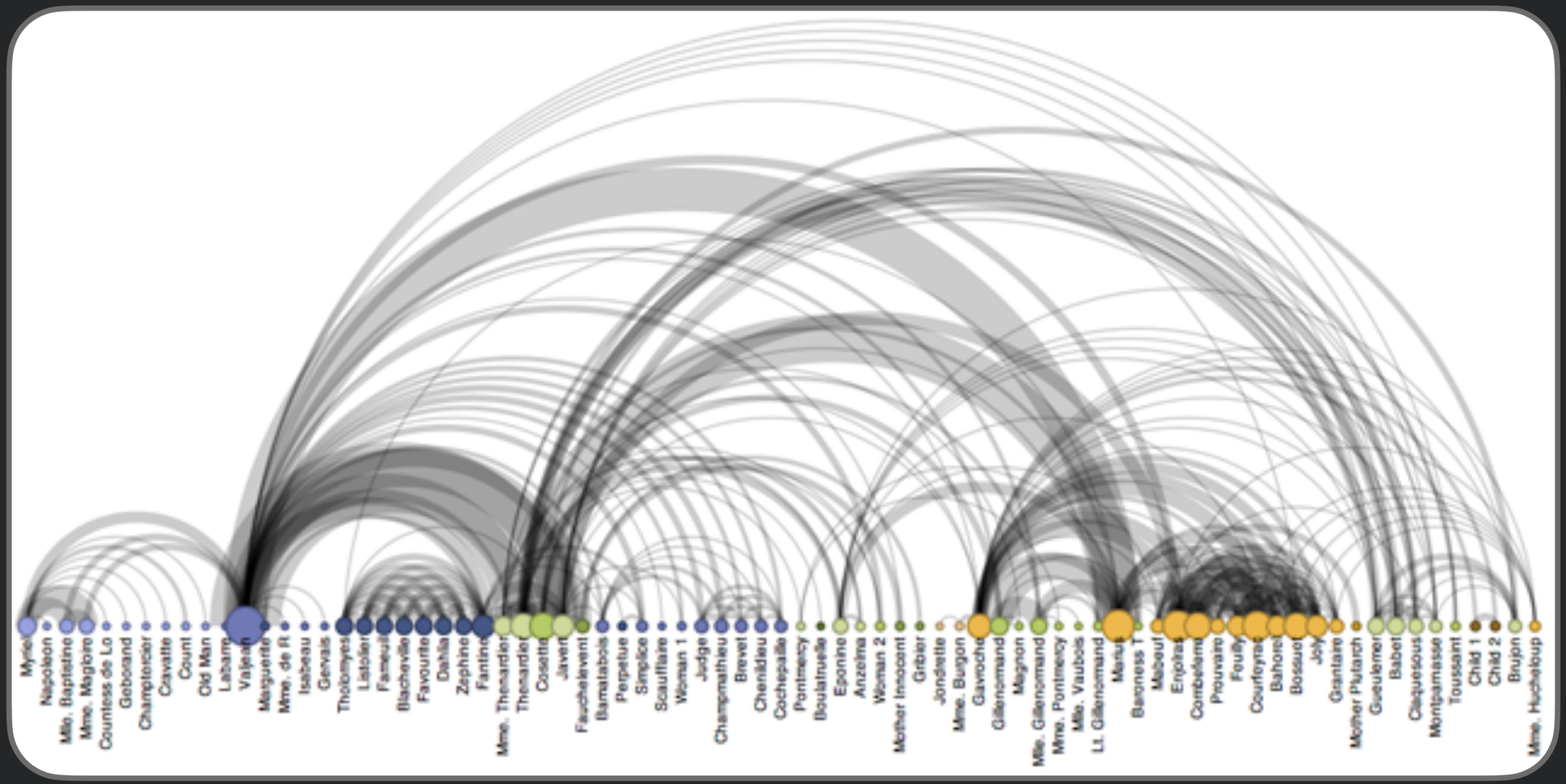
Force-directed Layout

- Edges function as springs, find least energy configuration

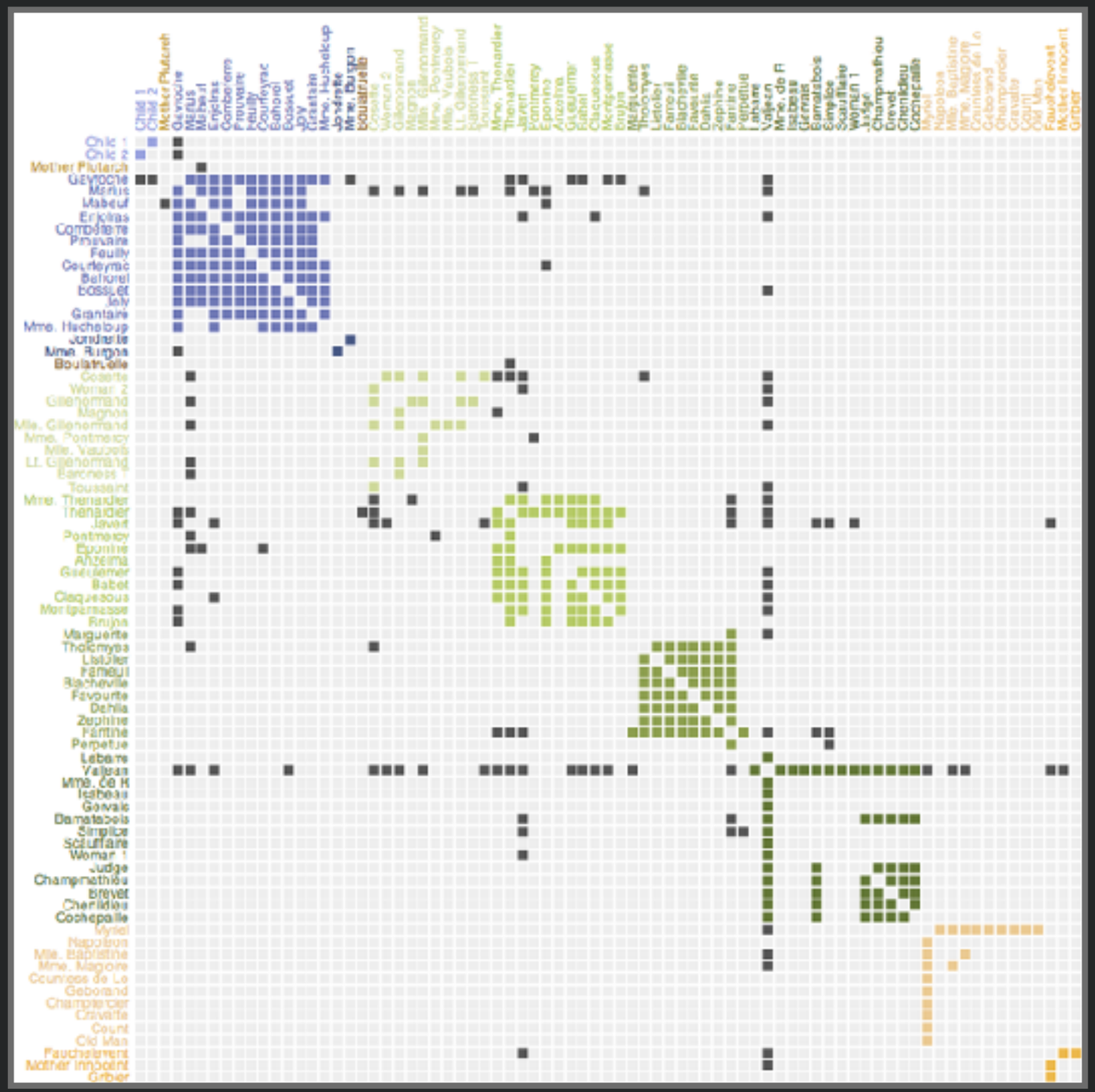


Arc Diagram

- Can support identifying cliques & bridges w/ right order



Adjacency Matrix



Design Considerations

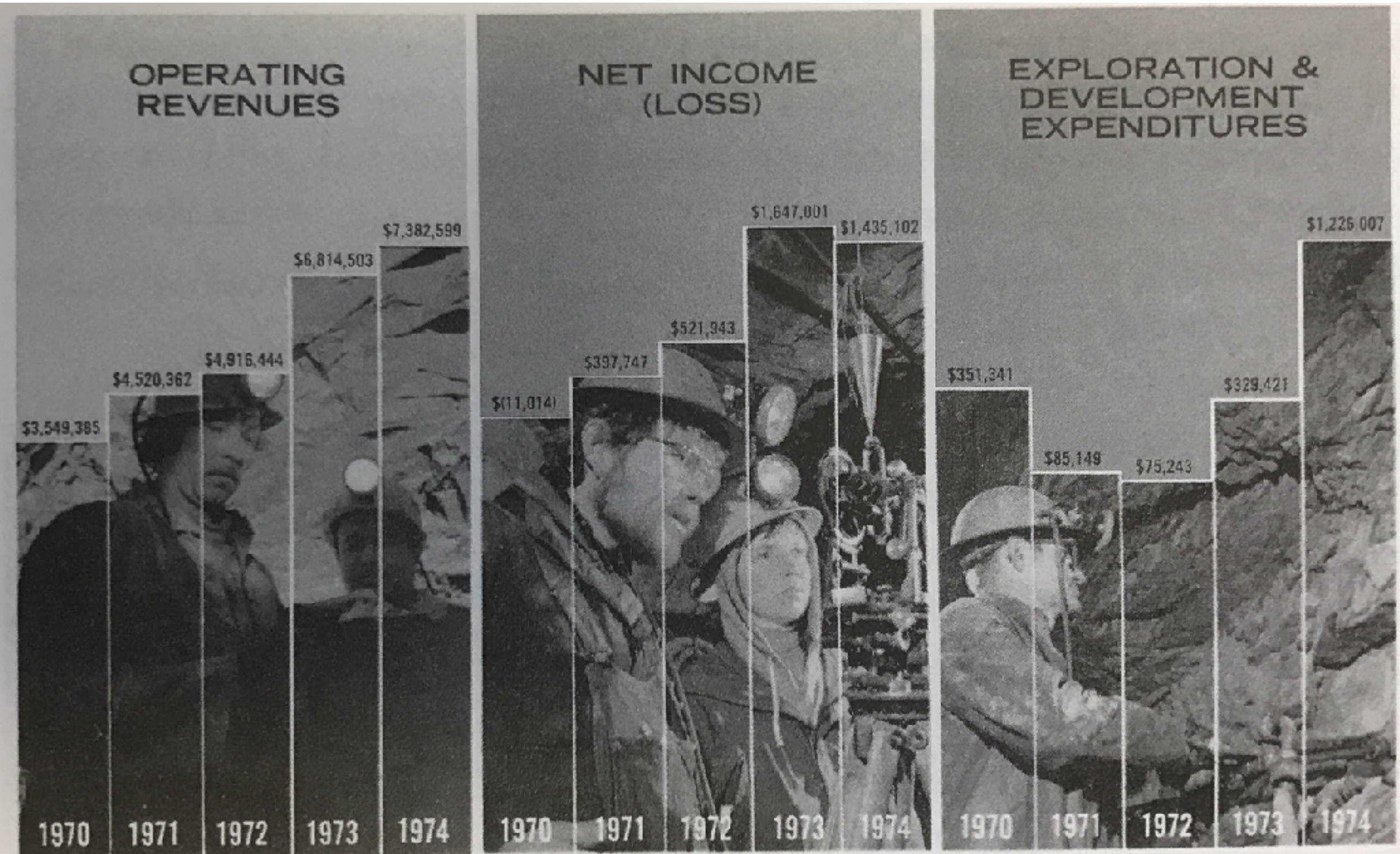
Tufte's principles of graphical excellence

- Show the ***data***
- Induce the viewer to think about the substance rather than the methodology
- Avoid distorting what the data have to say
- Present ***many*** numbers in a small space
- Make large data sets ***coherent***
- Encourage the eye to ***compare*** different pieces of data
- Reveal data at several levels of detail, from overview to fine structure
- Serve reasonable clear ***purpose***: description, exploration, tabulation, decoration

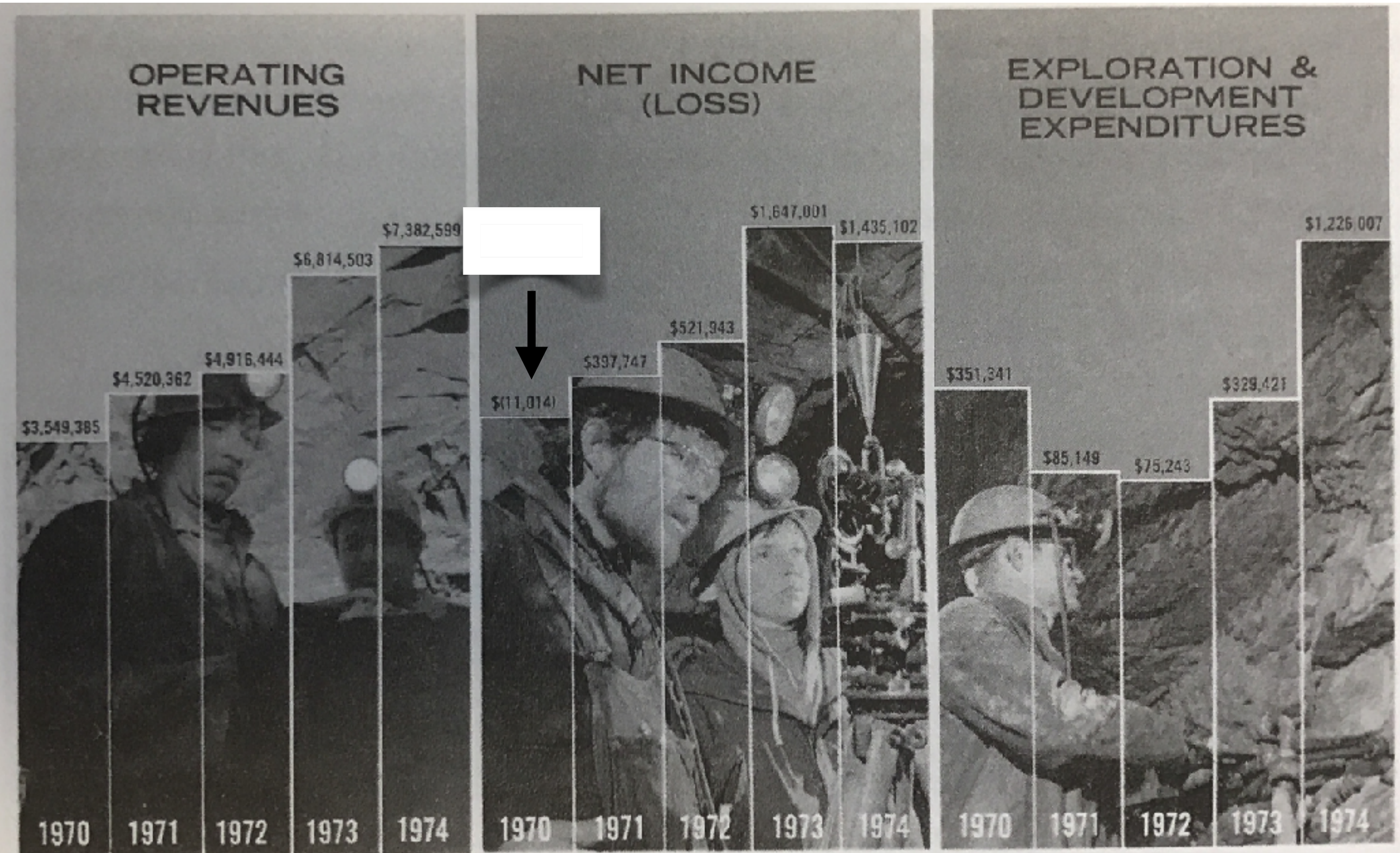
Distortions in Visualizations

- Visualizations may distort the underlying data, making it harder for reader to understand truth
- Use of ***design*** variation to try to falsely communicate ***data*** variation

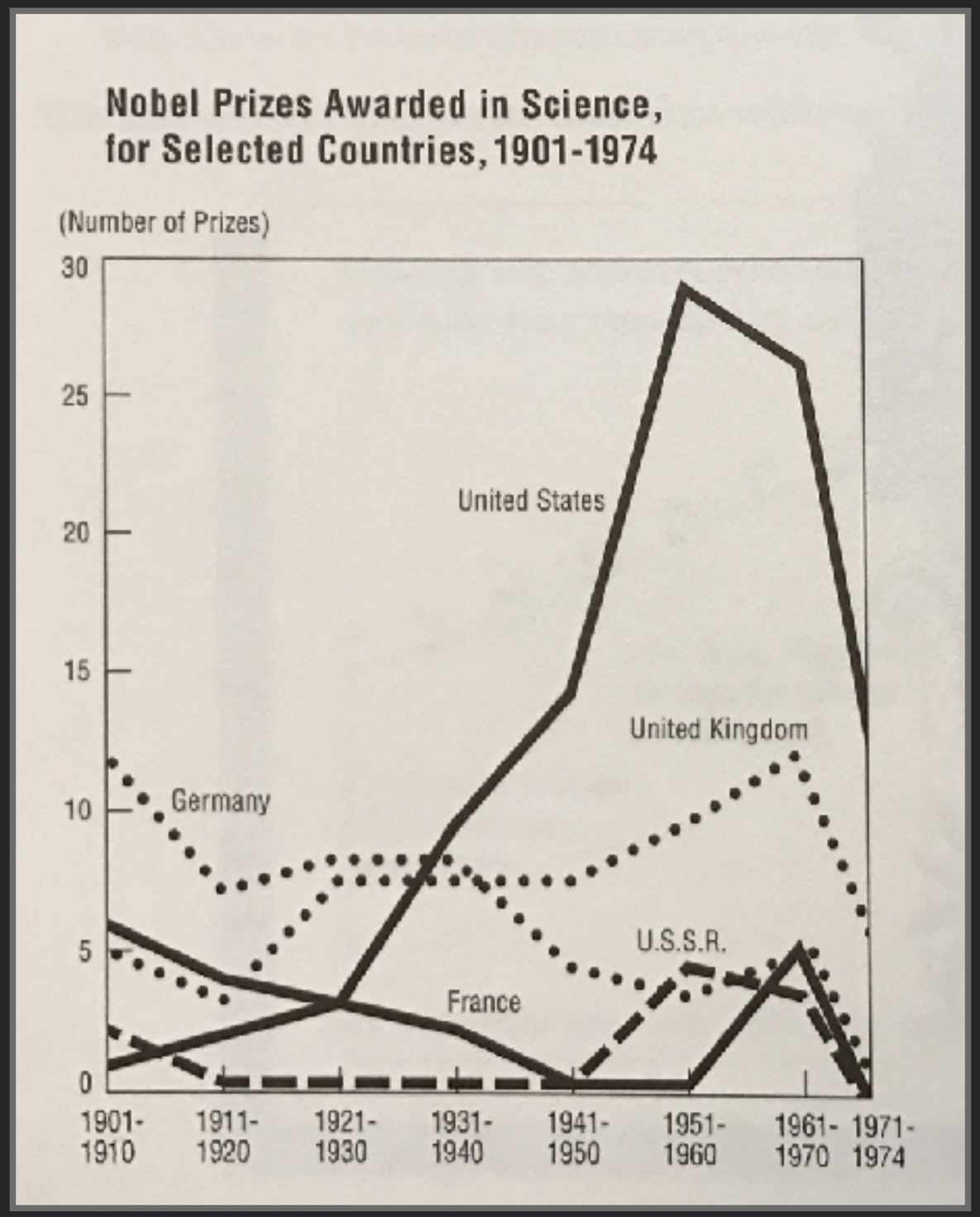
Example


OPERATING REVENUES
$3,549,385
$4,520,362
$4,916,444
$6,814,503
$7,382,599
1970
1971
1972
1973
1974
NET INCOME (LOSS)
$(11,014)
$397,747
$521,943
$1,647,001
$1,435,102
1970
1971
1972
1973
1974
EXPLORATION & DEVELOPMENT EXPENDITURES
$351,341
$85,149
$75,243
$329,421
$1,226,007
1970
1971
1972
1973
1974

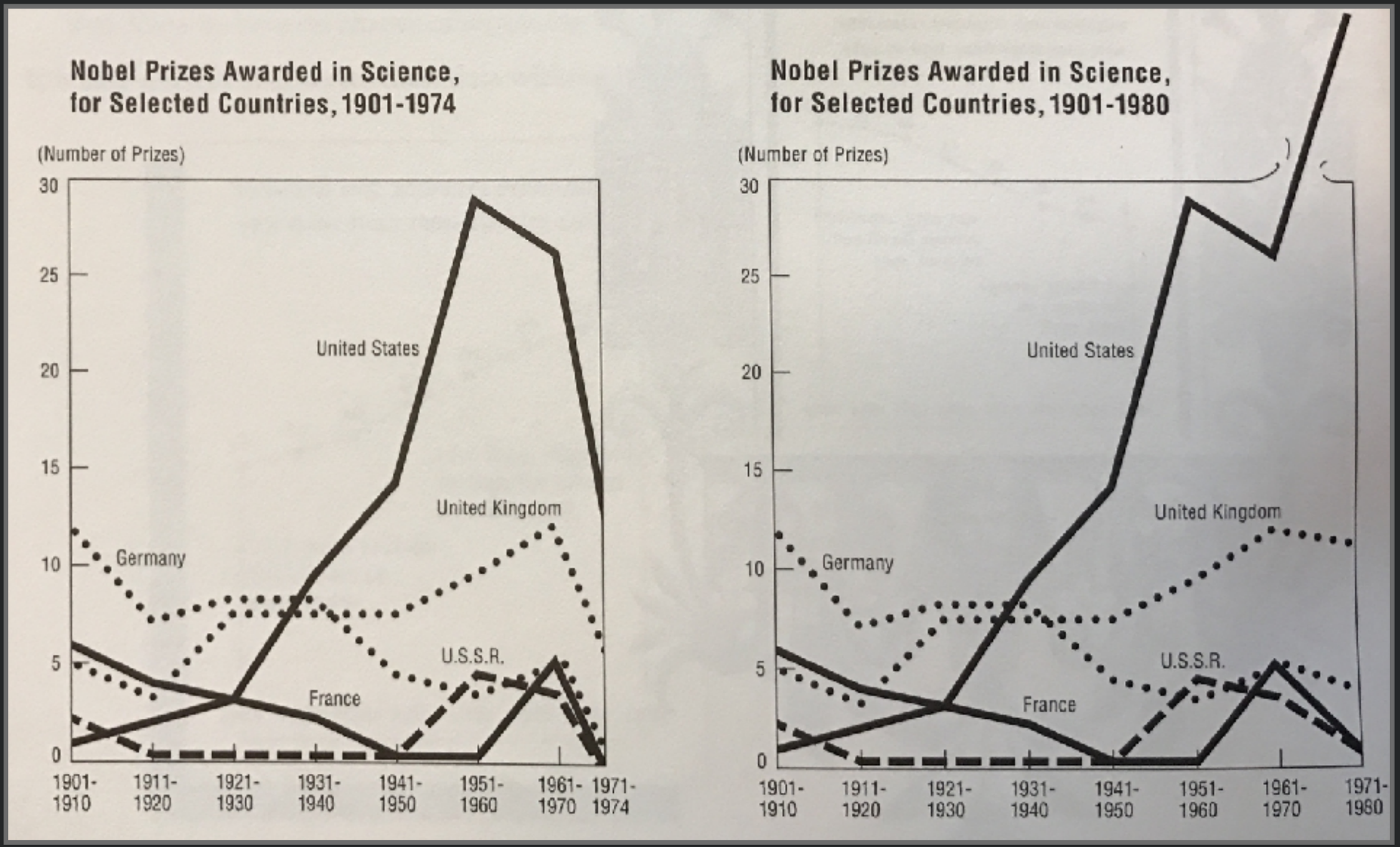
Example


OPERATING REVENUES
$3,549,385
$4,520,362
$4,916,444
$6,814,503
$7,382,599
1970
1971
1972
1973
1974
NET INCOME (LOSS)
$(11,014)
$397,747
$521,943
$1,647,001
$1,435,102
1970
1971
1972
1973
1974
EXPLORATION & DEVELOPMENT EXPENDITURES
$351,341
$85,149
$75,243
$329,421
$1,226,007
1970
1971
1972
1973
1974

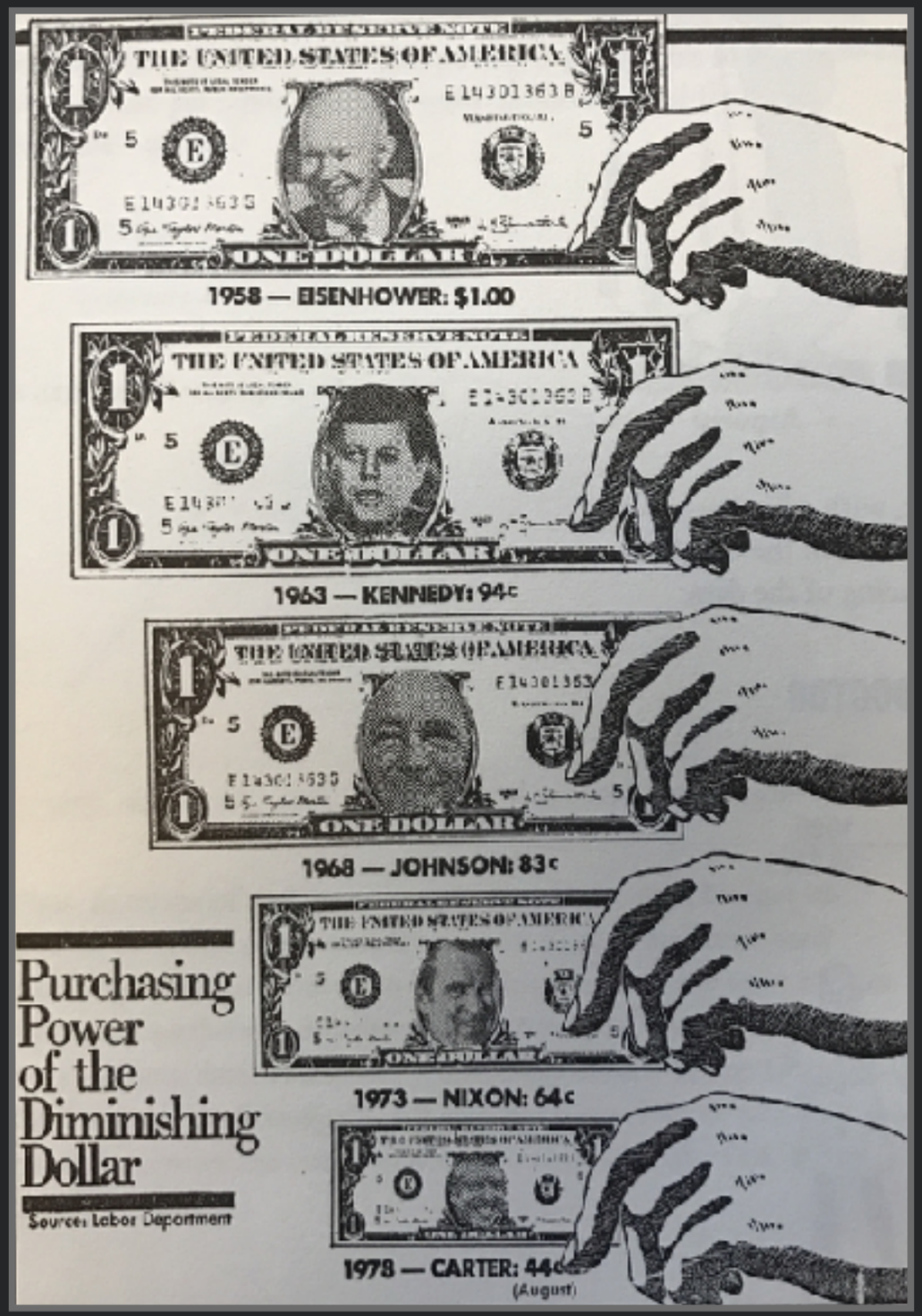
Example


Nobel Prizes Awarded in Science,
for Selected Countries, 1901-1974
(Number of Prizes)
30
25
20
15
10
5
0
United States
United Kingdom
Germany
U.S.S.R.
France
1901-1910
1911-1920
1921-1930
1931-1940
1941-1950
1951-1960
1961-1970
1971-1974

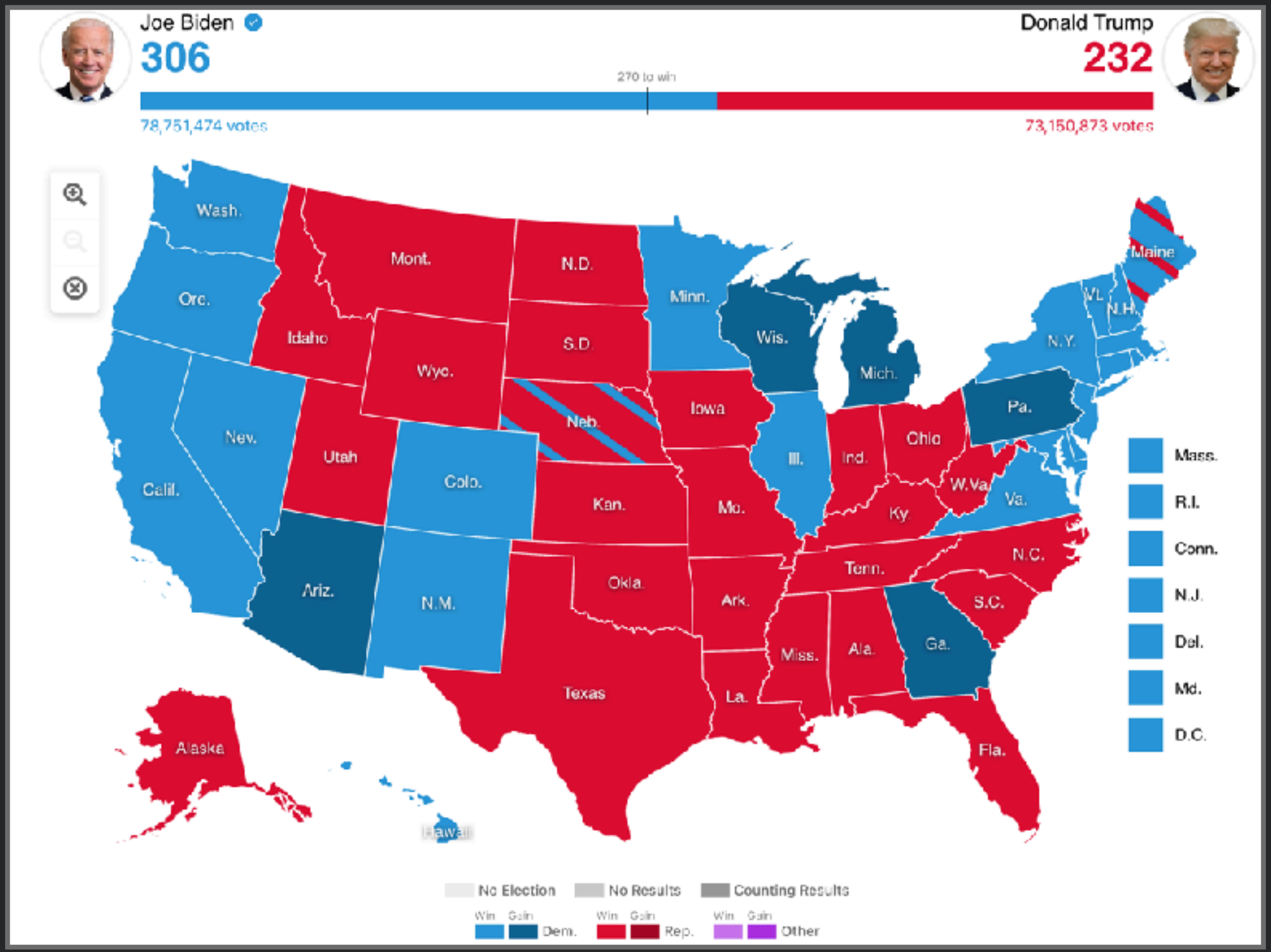
Example (corrected)



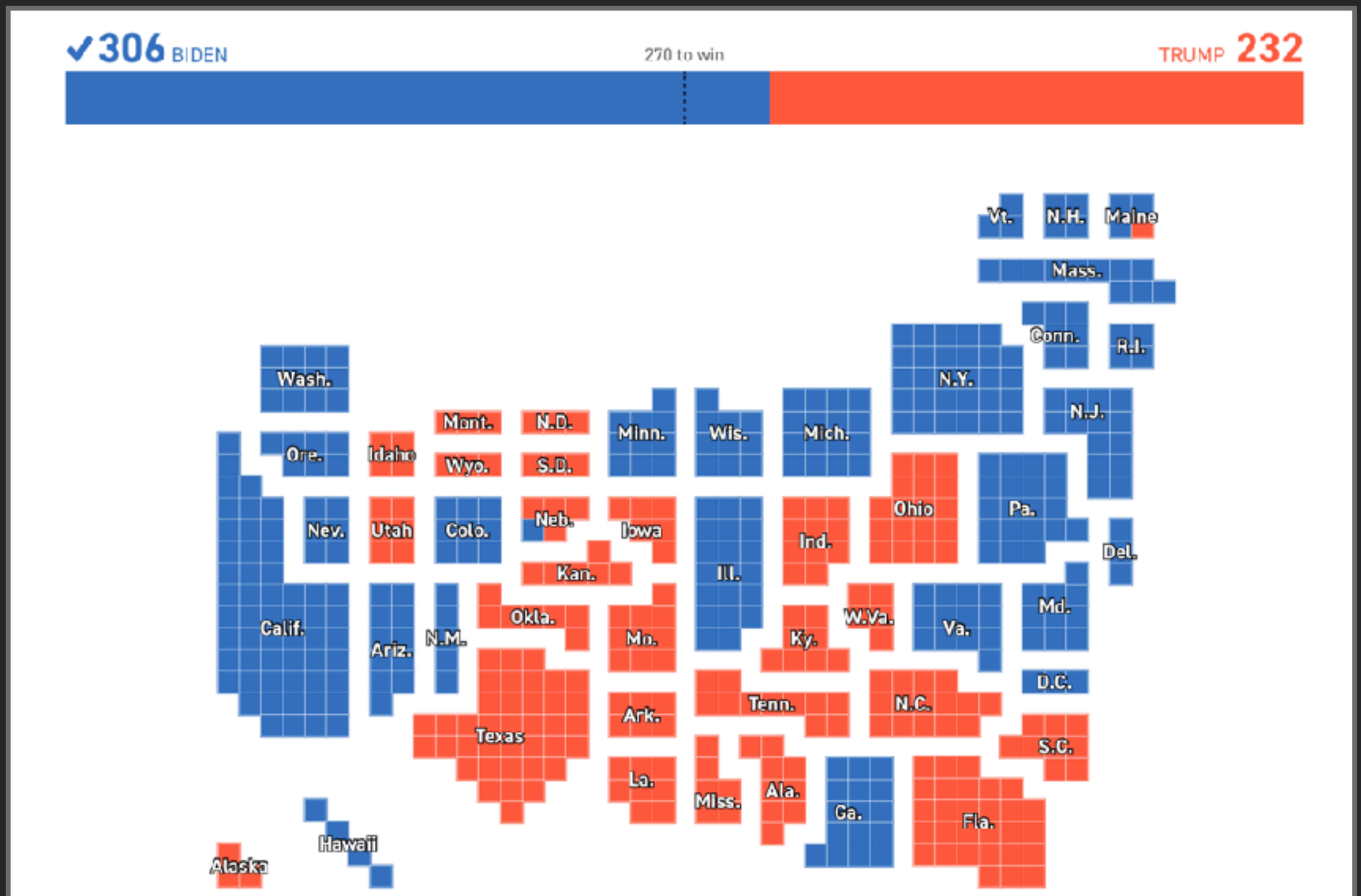
Example


THE UNITED STATES OF AMERICA
ONE DOLLAR
1958 — EISENHOWER: $1.00
1963 — KENNEDY: 94¢
1968 — JOHNSON: 83¢
1973 — NIXON: 64¢
1978 — CARTER: 44¢
(August)
Purchasing Power of the Diminishing Dollar
Source: Labor Department

Traditional Electoral Map



Weighted Electoral Map



Data-ink

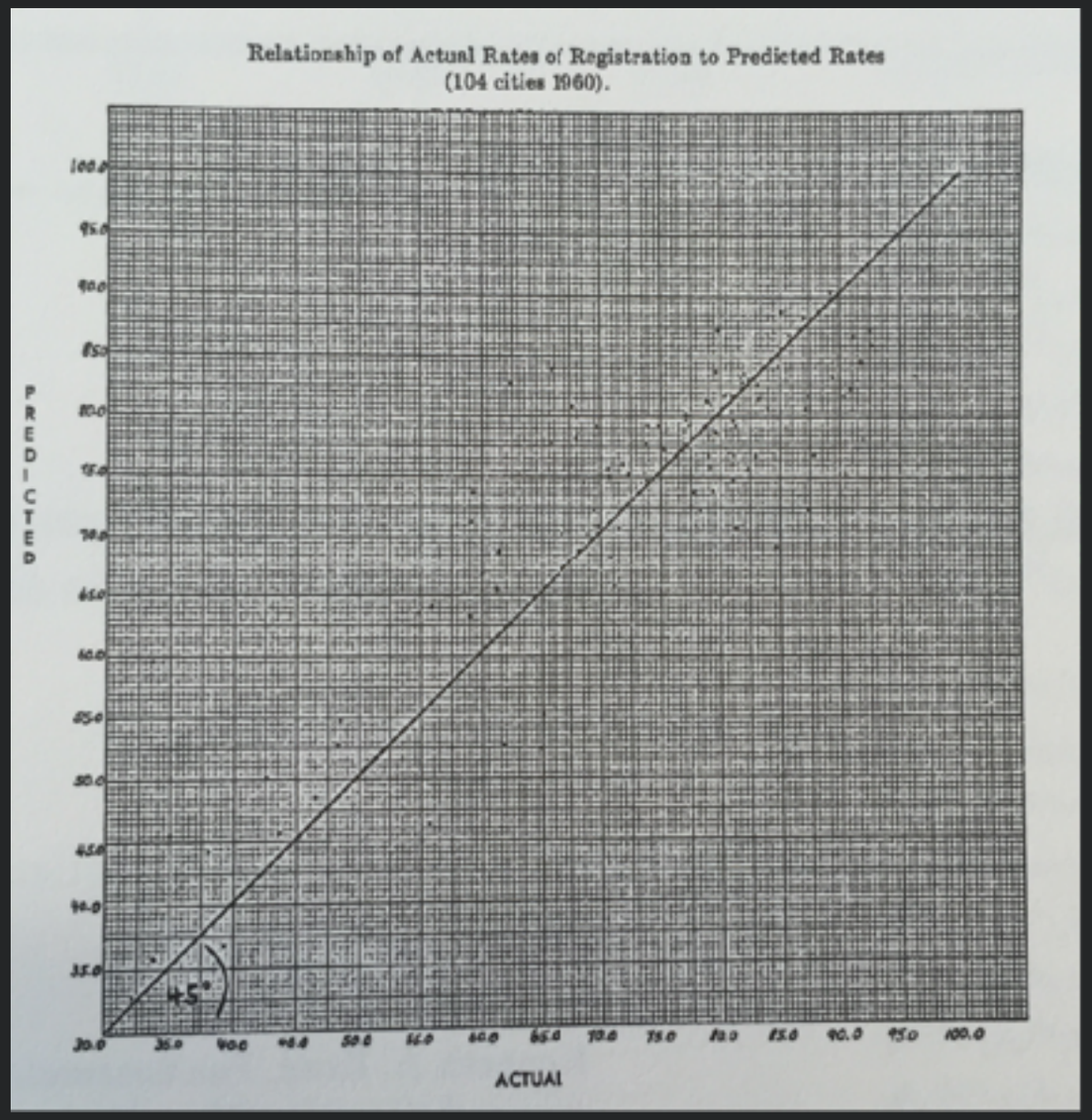
- Data-ink - non-redundant ink encoding data information

$$\text{Data-ink ratio} = \frac{\text{Data-ink}}{\text{Total ink used to print the graphic}}$$

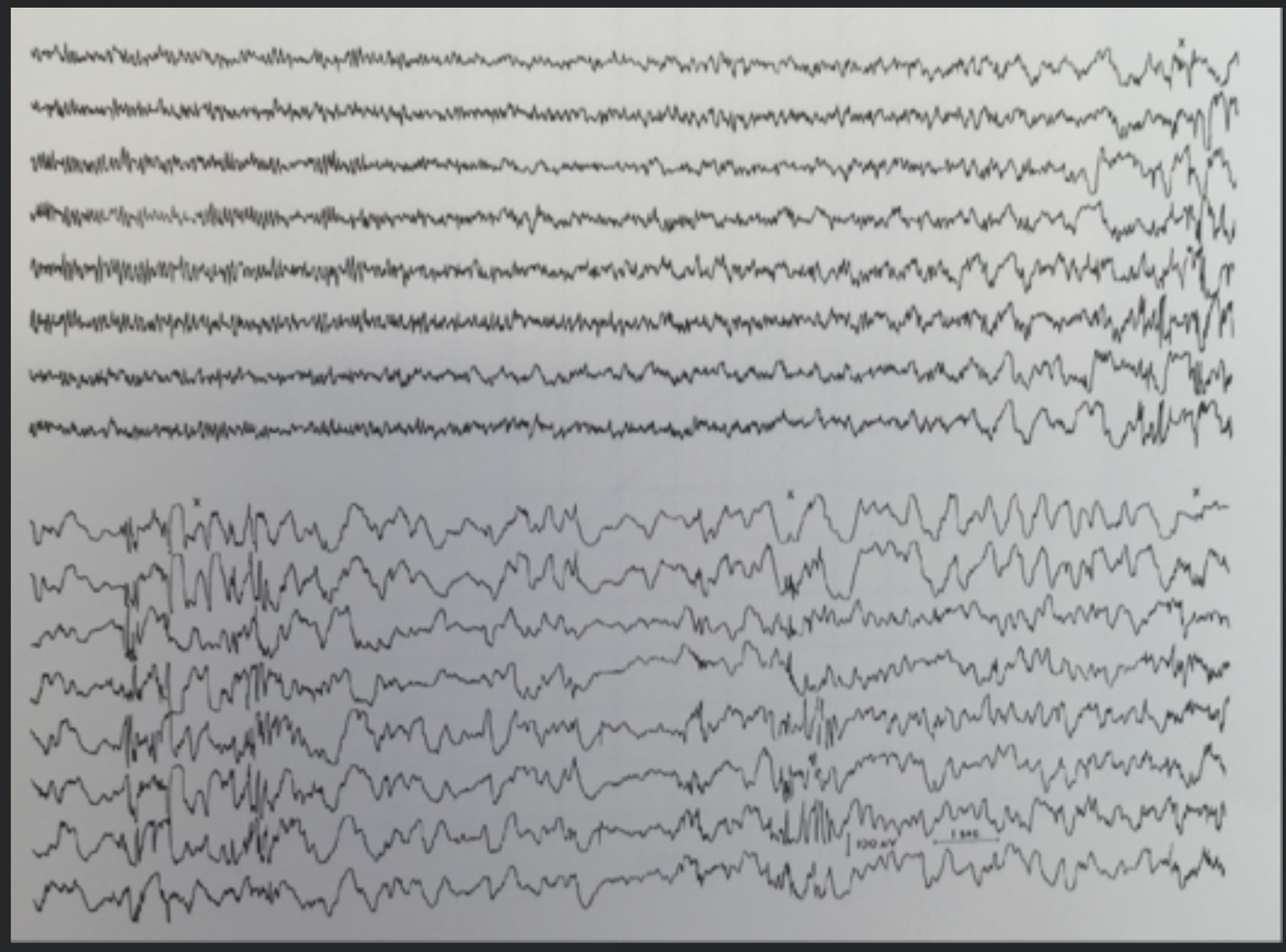
= proportion of a graphic's ink devoted to the non-redundant display of data-information

= 1.0 – proportion of a graphic that can be erased

Examples of Data-ink Ratio



~0

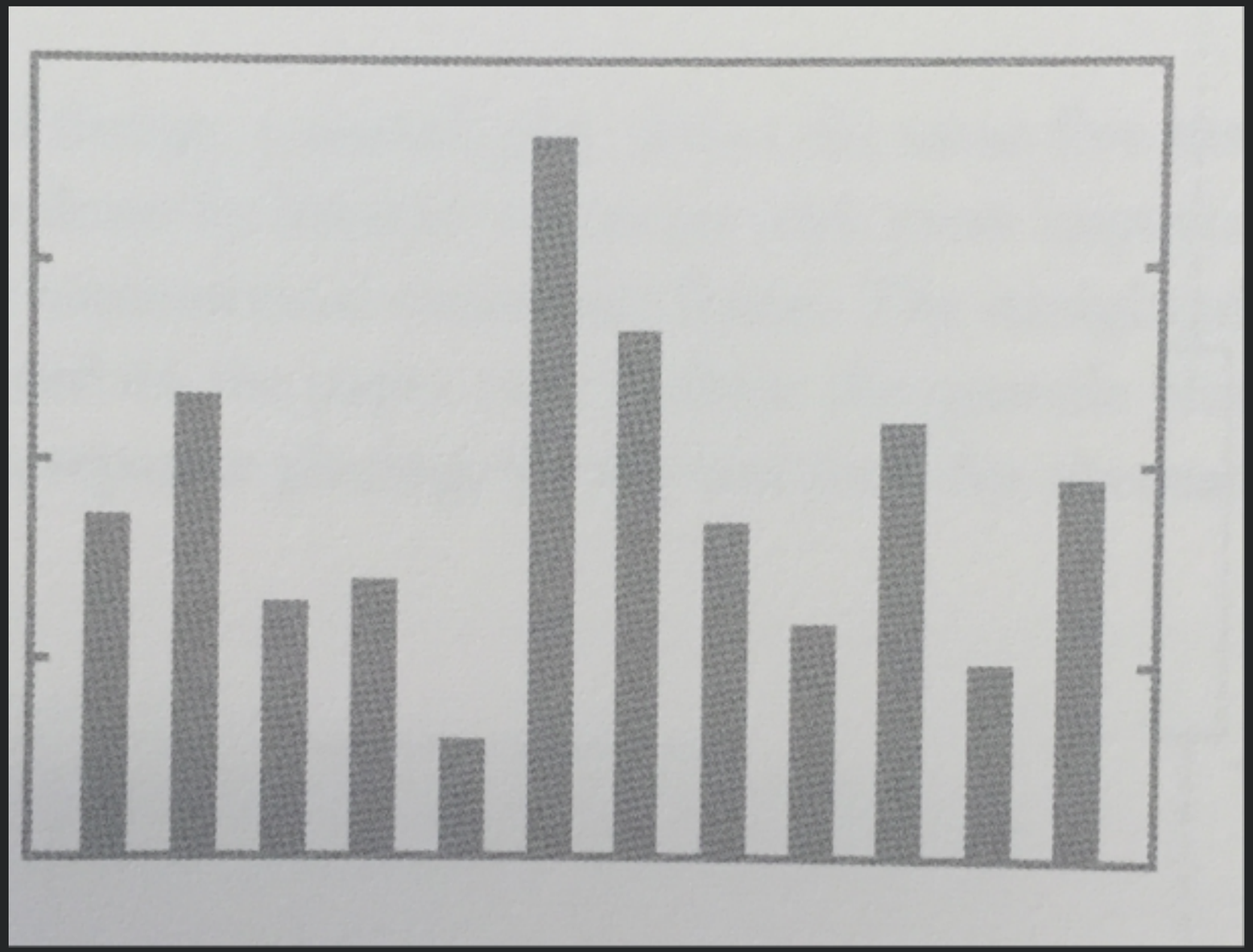


1.0

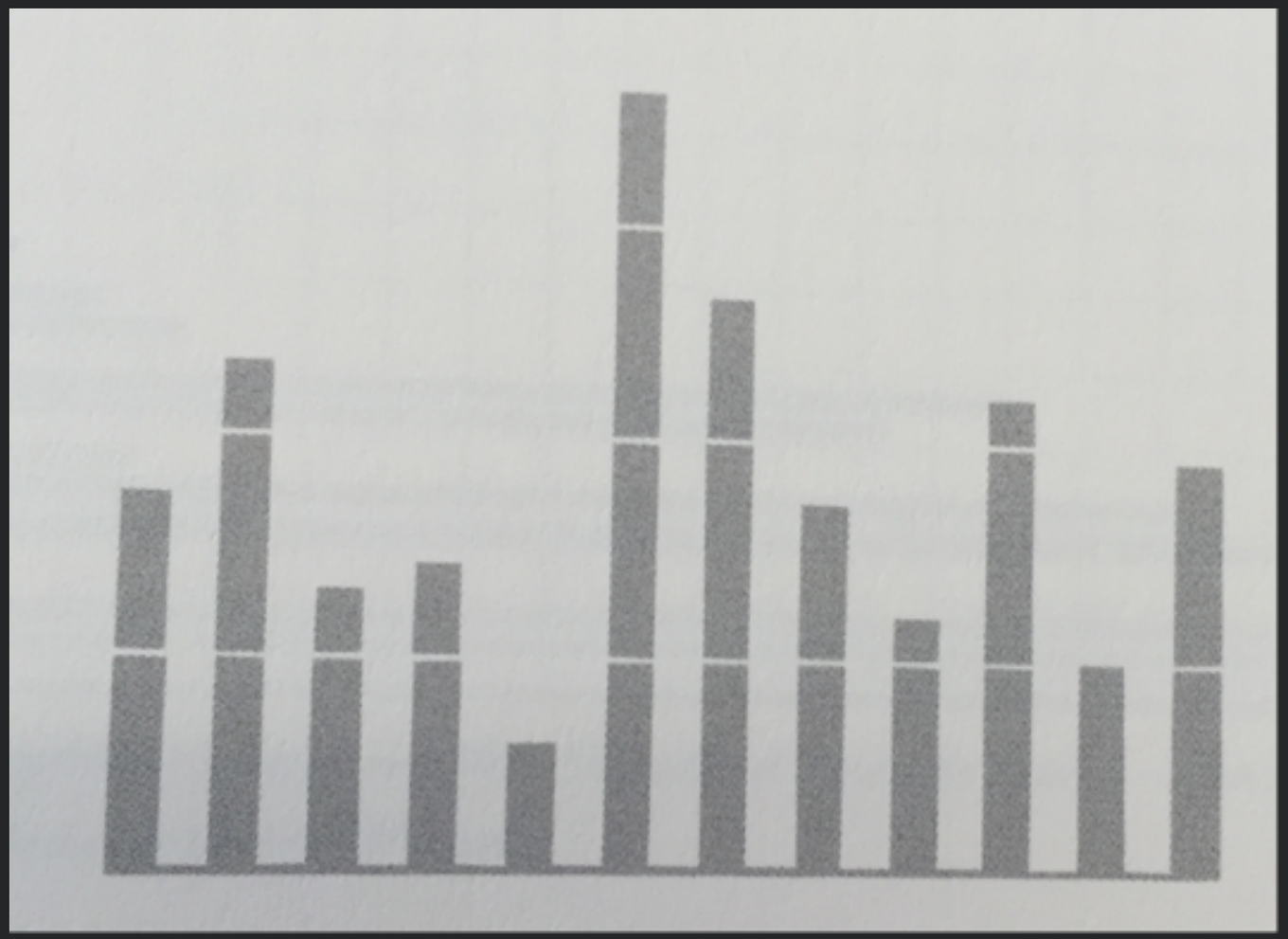
Design Principles for Data-ink

- (a.k.a. aesthetics & minimalism / elegance & simplicity)
- ***Above all else show the data***
- Erase non-data-ink, within reason
  - Often not valuable and distracting
  - Redundancy not usually useful

Example



Example (revised)



Interacting with Visualizations

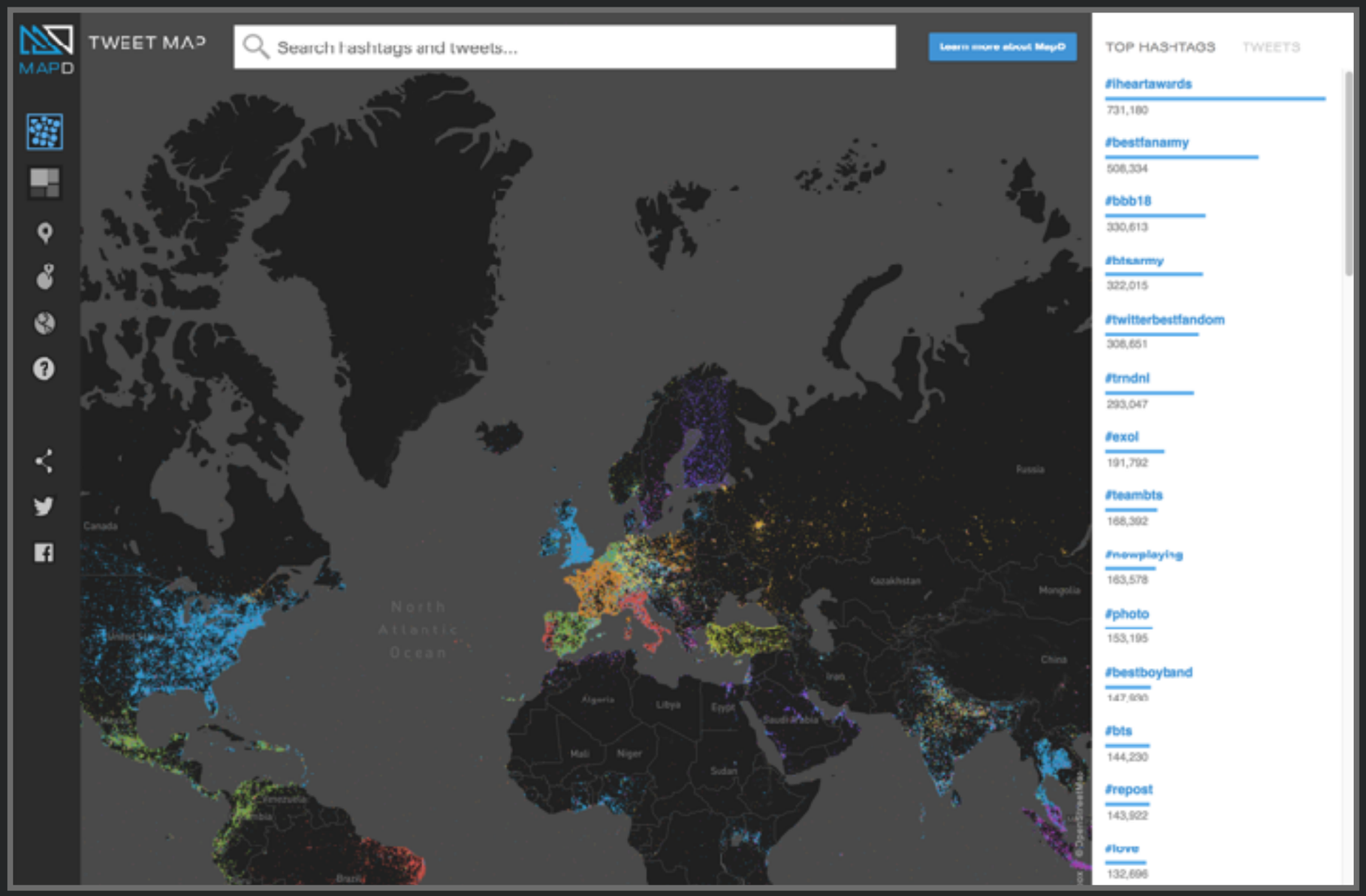
Interactive Visualizations

- Users often use iterative process of making ***sense*** of the data
  - Answers lead to new questions
- Interactivity helps user constantly change display of information to answer new questions
- Should offer visualization that offers best view of data moment to ***moment*** as desired view ***changes***

Information Visualization Tasks

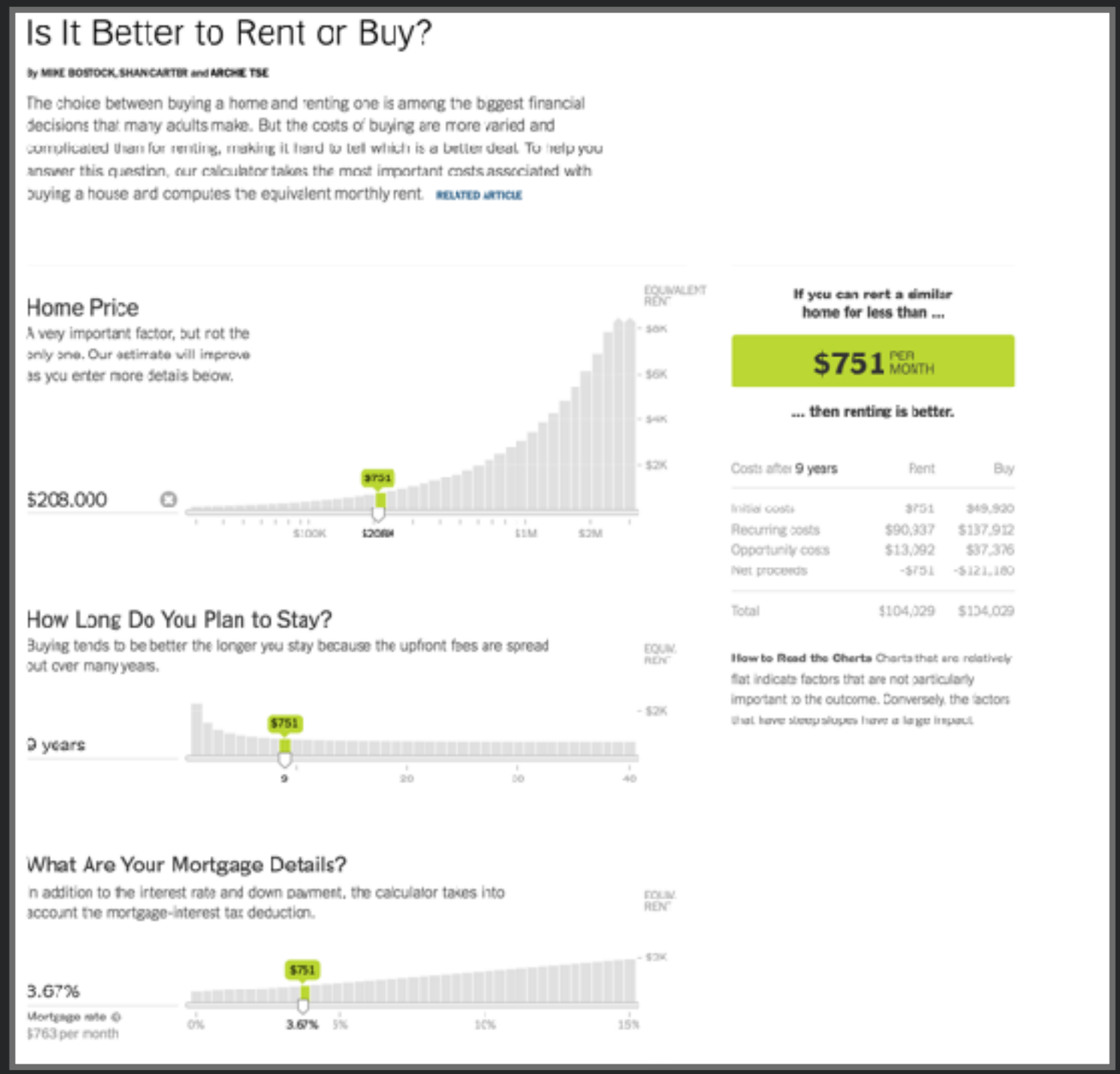
- ***Overview:*** gain an overview of entire collection
- ***Zoom:*** zoom in on items of interest
- ***Filter:*** filter out uninteresting items
- ***Details on Demand:*** select an item or group and get details
- ***Relate:*** view relationships between items
- ***History:*** support undo, replay, progressive refinement
- ***Extract:*** allow extraction of sub-collections through queries

Global Tweet Map



https://www.mapd.com/demos/tweetmap/

Renting vs. Buying Utility



Is It Better to Rent or Buy?

By MIKE BOSTOCK, SHAN CARTER and ARCHIE TSE

The choice between buying a home and renting one is among the biggest financial decisions that many adults make. But the costs of buying are more varied and complicated than for renting, making it hard to tell which is a better deal. To help you answer this question, our calculator takes the most important costs associated with buying a house and computes the equivalent monthly rent. RELATED ARTICLE

Home Price

A very important factor, but not the only one. Our estimate will improve as you enter more details below.

$208,000

EQUIVALENT RENT

$6K
$4K
$2K

$751

$100K 208K $1M $2M

If you can rent a similar home for less than ...

$751 PER MONTH

... then renting is better.

Costs after 9 years	Rent	Buy
Initial costs	$751	$49,920
Recurring costs	$90,337	$137,912
Opportunity costs	$13,092	$37,376
Net proceeds	-$751	-$121,180
Total	$104,029	$104,029

How to Read the Charts Charts that are relatively flat indicate factors that are not particularly important to the outcome. Conversely, the factors that have steep slopes have a large impact.

How Long Do You Plan to Stay?

Buying tends to be better the longer you stay because the upfront fees are spread out over many years.

9 years

EQUIV. RENT

$2K

$751

9 20 30 40

What Are Your Mortgage Details?

In addition to the interest rate and down payment, the calculator takes into account the mortgage-interest tax deduction.

3.67%

Mortgage rate
$763 per month

EQUIV. RENT

$2K

$751

0% 3.67% 5% 10% 15%

https://www.nytimes.com/interactive/2014/upshot/buy-rent-calculator.html?_r=0

Acknowledgements

- Slides adapted from Dr. Thomas Latoza's SWE 432 course