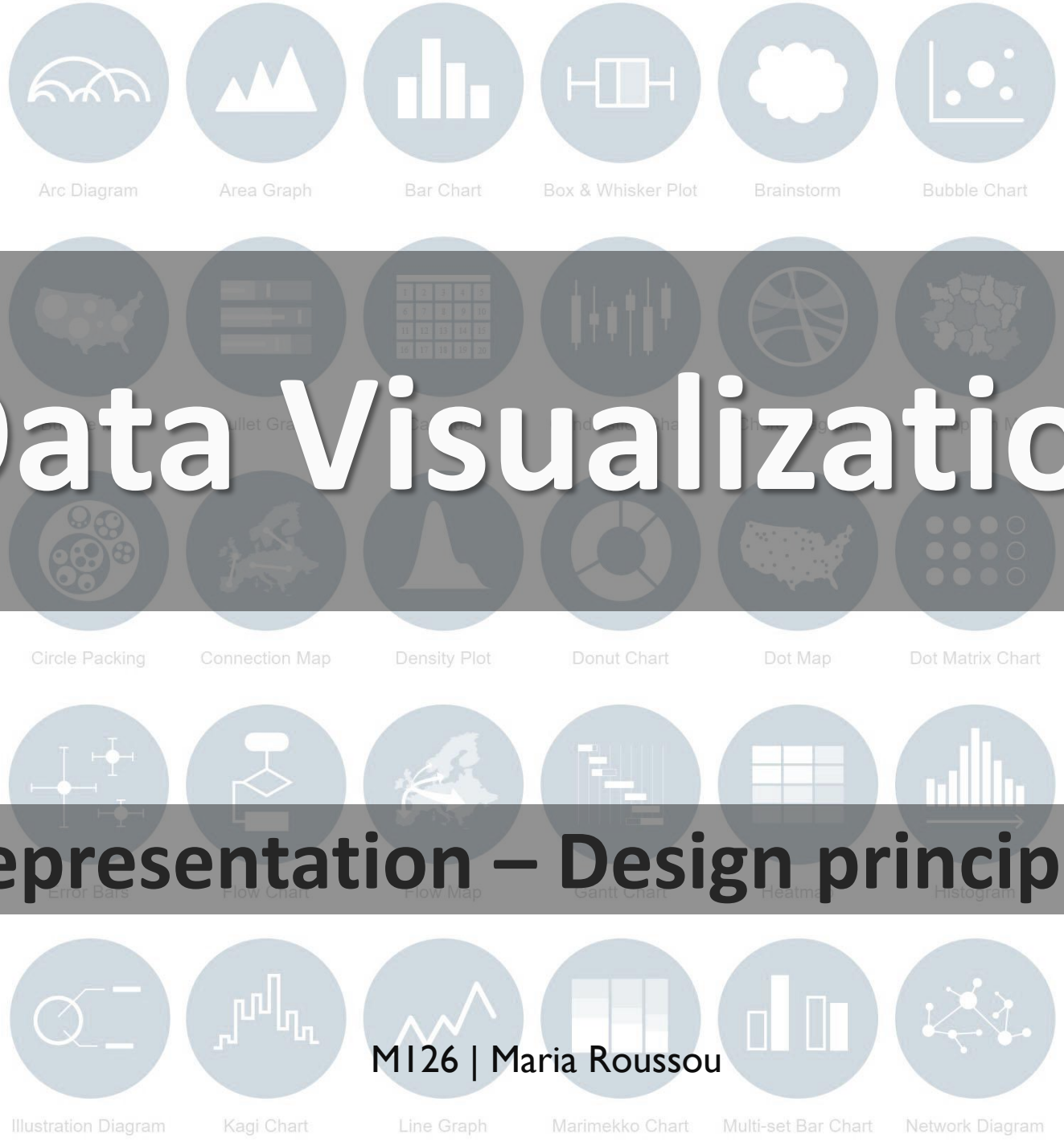


Data Visualization

Representation – Design principles

M126 | Maria Roussou

Representation

- Representation has to do with the visual encoding of our data
- Design decisions affect interpretation

- How one chooses an encoding depends on the type of data to be represented

Representation – visual encoding

City	State	Population
Baton Rouge	Louisiana	191,741
Birmingham	Alabama	220,927
Broken Arrow	Oklahoma	58,018
Eugene	Oregon	115,890
Glendale	Arizona	245,868
Huntsville	Alabama	55,741
Lafayette	Louisiana	87,737
Mobile	Alabama	98,147
Montgomery	Alabama	126,250
New Orleans	Louisiana	322,172
Norman	Oklahoma	101,590
Peoria	Arizona	167,868
Portland	Oregon	514,108
Salem	Oregon	147,631
Scottsdale	Arizona	134,335
Shreveport	Louisiana	68,756
Surprise	Arizona	90,548
Tempe	Arizona	143,369
Tulsa	Oklahoma	392,138

Visual encoding

What kind of data do we have?

How can we represent the data visually?

How can we organize this into a visualization?

Design steps / considerations

1. Problem?
2. Data type?
3. Dimensions?
4. Data structure?
5. Type of interaction?

Design steps

1. Define the problem

Identify users' needs. Why is a representation needed?

- to **communicate** something
- to **explore** for finding new information
- to **confirm** and prove hypotheses

2. Examine the nature of the data to represent

- **quantitative** (e.g., a list of integers or real numbers)
- **ordinal** (data of a non-numeric nature, but which have their own intrinsic order, such as the days of the week)
- **categorical** or **nominal** (data that have no intrinsic order, such as the names of people or cities).

What kind of data do we have? Data types

- Categorical (or nominal):
  - NL, B, F, D, UK, I, ES, …
  - orange, apple, pear, strawberry, …

equal or not equal to other values

non-orderable

- Ordinal:
  - small, medium, large
  - weak, average, strong

obeys ordering relations

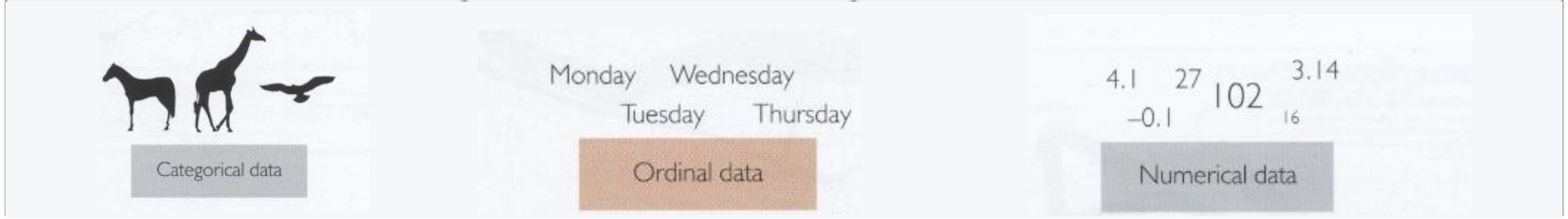
- Numerical (or quantitative):
  - –1, 15, 32, 49, …

can perform arithmetics

orderable

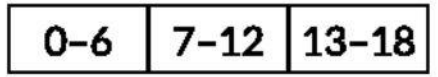
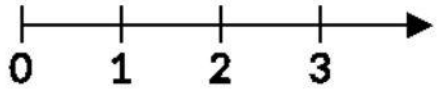
What kind of data do we have? Data types

Categorical	Ordinal	Numerical	
		Interval	Ratio
male / female true / false Asia / Europe / Africa	small, medium, large	latitude, longitude, time (event)	length, count, time (duration)
=	= <>	= <>
-	= <>
+-
* /



What kind of data do we have? Data types

Logical mathematical operations permissible, measure of central tendency, and examples for different data scale types.

Data Scales	Logical Math Operations				Measure of Central Tendency	Examples
	= ≠	< >	+ -	x ÷		
Nominal	✓				mode	
Ordinal	✓	✓			median	
Interval	✓	✓	✓		arithmetic mean	
Ratio	✓	✓	✓	✓	geometric mean	

Katy Börner et al. PNAS 2019;116:6:1857-1864

©2019 by National Academy of Sciences

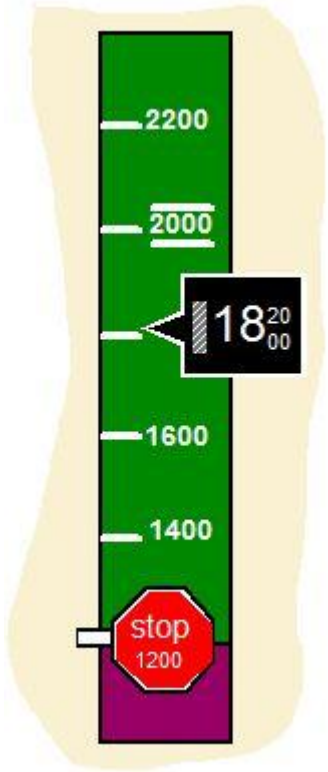
3. Number of dimensions

According to the number of dependent attributes:

- **univariate** (one dimension varies with respect to another)
- **bivariate** (there are two dependent dimensions)
- **trivariate** (three dependent dimensions)
- **multivariate** (four or more dimensions that vary compared to the independent ones)

Number of dimensions: Univariate data – one value

- Visualizing just one number!
- A single value can be displayed
  - as the number itself - a string of digits
  - as a dial (e.g. altimeter, speedometer, …)
  - as a slider or thermometer



Number of dimensions: Univariate data – a collection of values

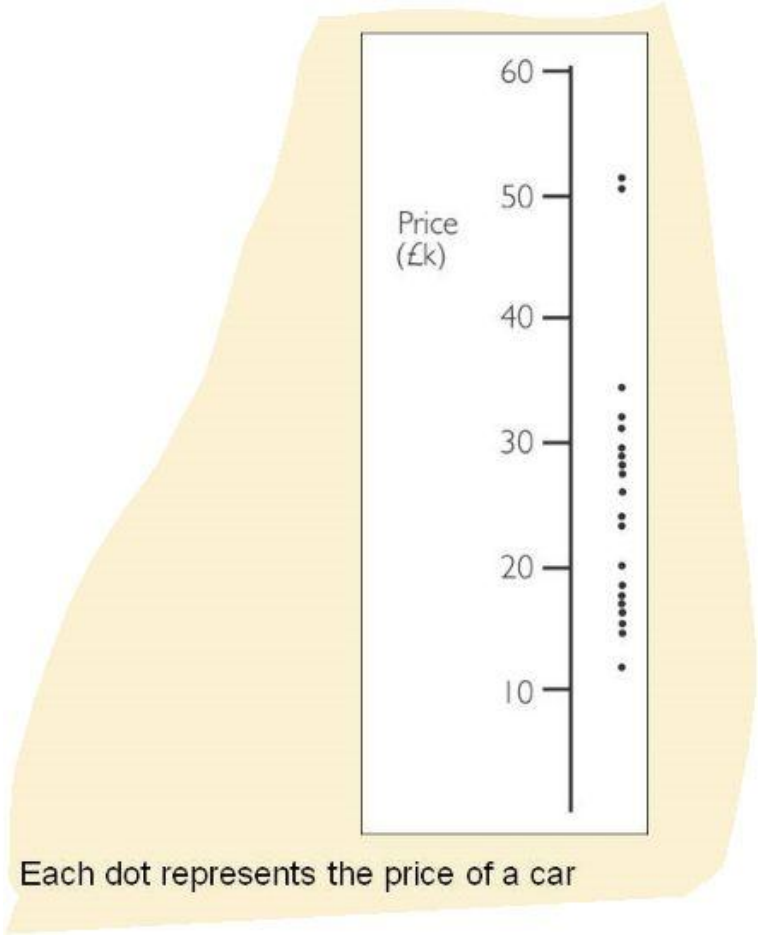
- Often univariate values are a part of a collection, set, or series/sequence of like data shown as:

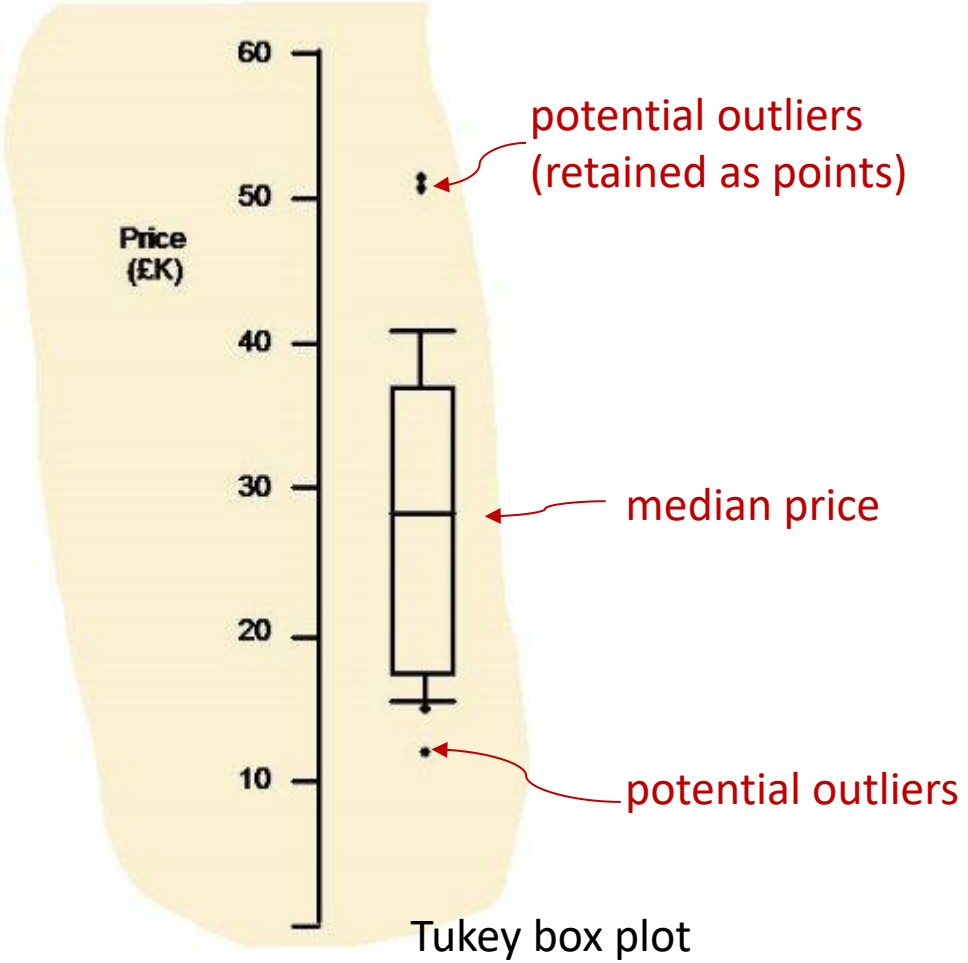
  - characteristic numbers (statistical mean, median or modes), displayed as single numbers
  - a dot plot along an axis
  - a Tukey box plot (mean and standard deviation, or median and quartile range) superimposed on the dot plot
  - bar graphs or histograms of the counts
  - categorized or grouping in a pie chart, stacked graph

Number of dimensions: Univariate data – a collection of values

- Often univariate values are a part of a collection, set, or series/sequence of like data shown as:

- plot





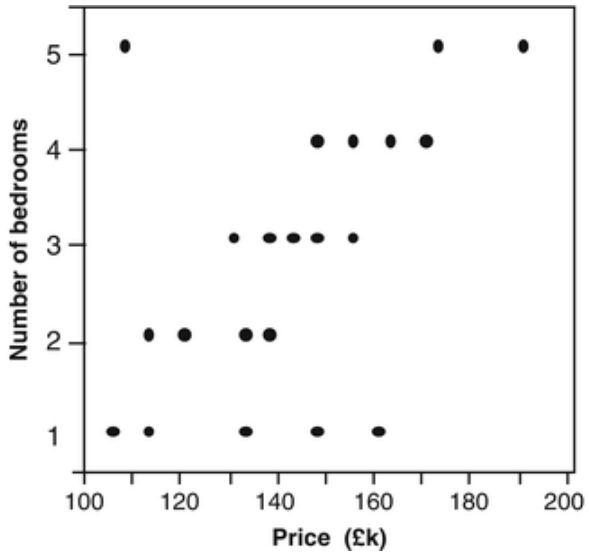
Tukey box plot

Number of dimensions: Bivariate data

- Sets of independent & dependent values

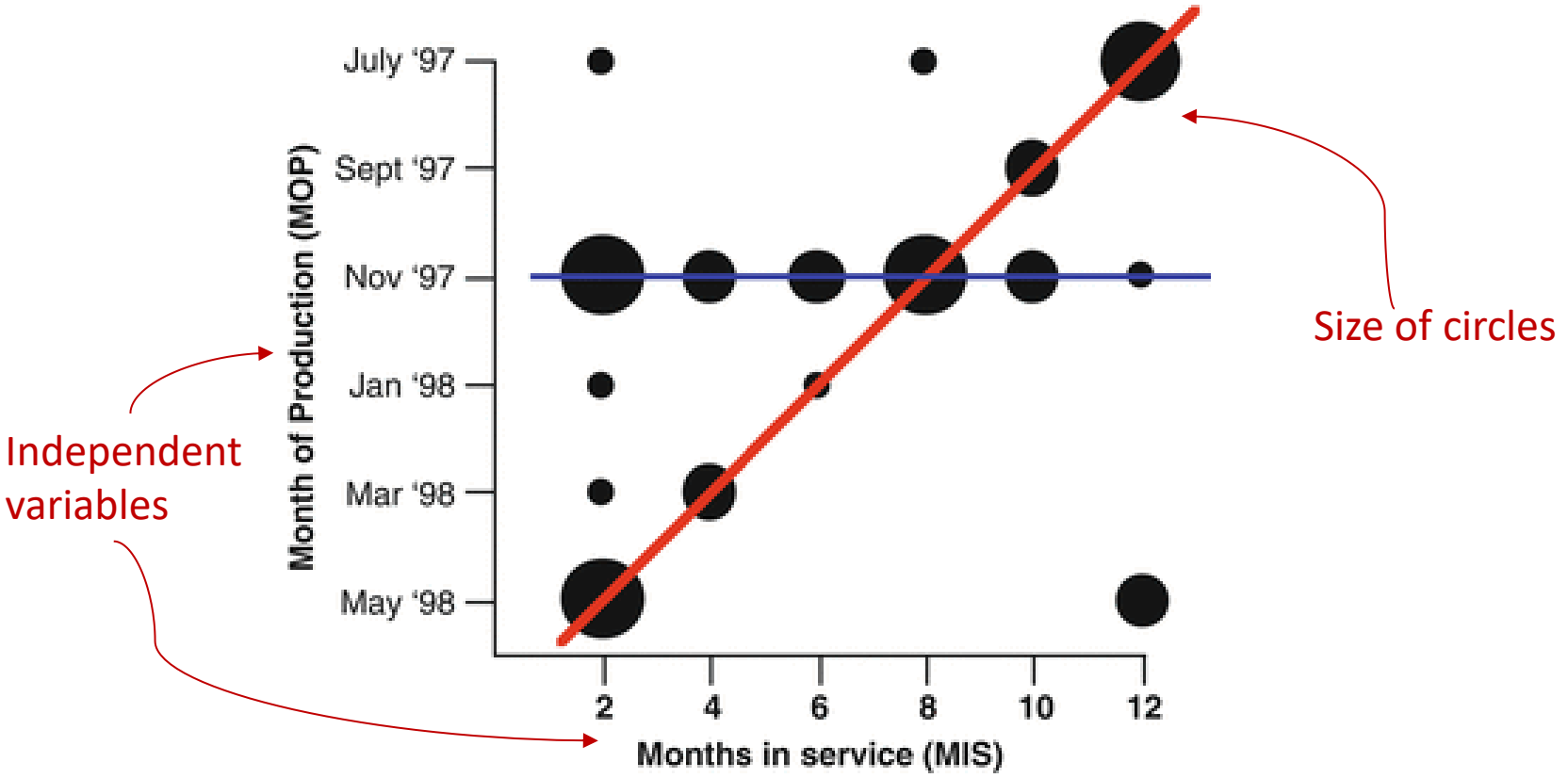
  e.g. independent variables may include:

  - time (by seconds, hourly, daily, monthly, yearly, etc.)
  - discrete values that have no natural numerical relation (colors, days of week, locations)
  - natural sequences



Number of dimensions: Trivariate data

- Objects characterized by 3 attributes

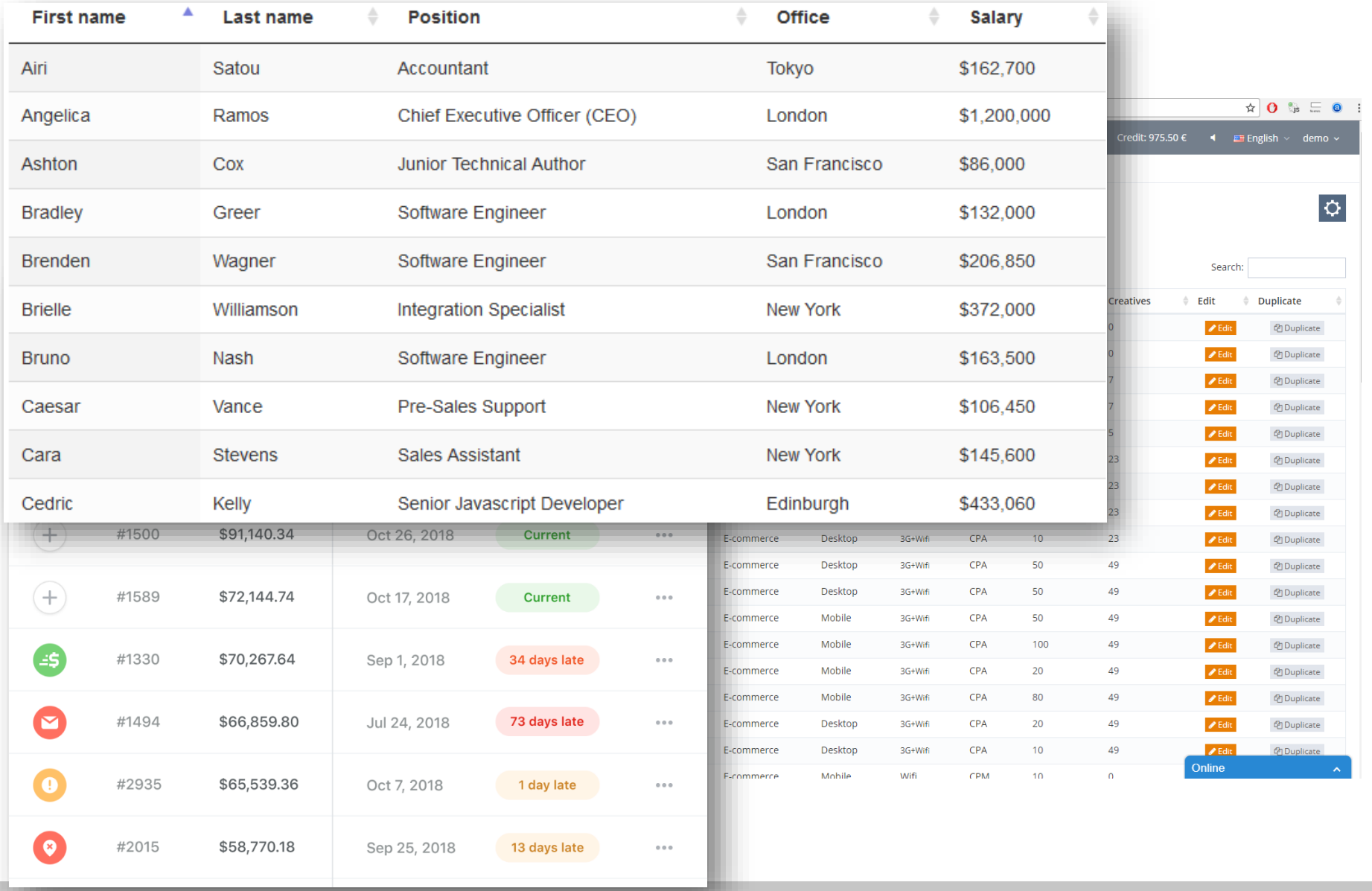


4. Data structures

- **linear** (the data are codified in linear data structures like vectors, tables, collections, etc.)
- **temporal** (data that change in time)
- **spatial** or geographical (data that have a correspondence with something physical, such as a map, floorplan, etc.)
- **hierarchical** (data relative to entities organized on hierarchy, e.g. genealogy, flowcharts, files on a disk, etc.)
- **network** (data that describe relationships between entities)

What kind of data do we have? e.g. in tabular form

First name	Last name	Position	Office	Salary
Airi	Satou	Accountant	Tokyo	$162,700
Angelica	Ramos	Chief Executive Officer (CEO)	London	$1,200,000
Ashton	Cox	Junior Technical Author	San Francisco	$86,000
Bradley	Greer	Software Engineer	London	$132,000
Brenden	Wagner	Software Engineer	San Francisco	$206,850
Brielle	Williamson	Integration Specialist	New York	$372,000
Bruno	Nash	Software Engineer	London	$163,500
Caesar	Vance	Pre-Sales Support	New York	$106,450
Cara	Stevens	Sales Assistant	New York	$145,600
Cedric	Kelly	Senior Javascript Developer	Edinburgh	$433,060



5. Type of interaction

- **static** (not modifiable by the user)
- **transformable** (user can control the process of modification and transformation of data)
- **manipulable** (user can control and modify some parameters that regulate the generation of the views)

Design considerations, in a nutshell

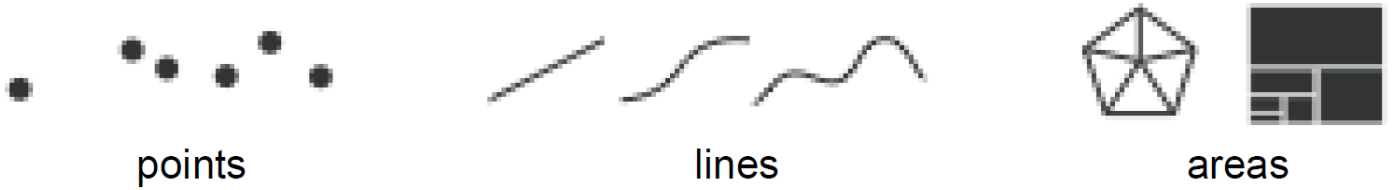
Variables to consider when designing visual representations:

1. Problem
2. Data type
3. Dimensions
4. Data structure
5. Type of interaction

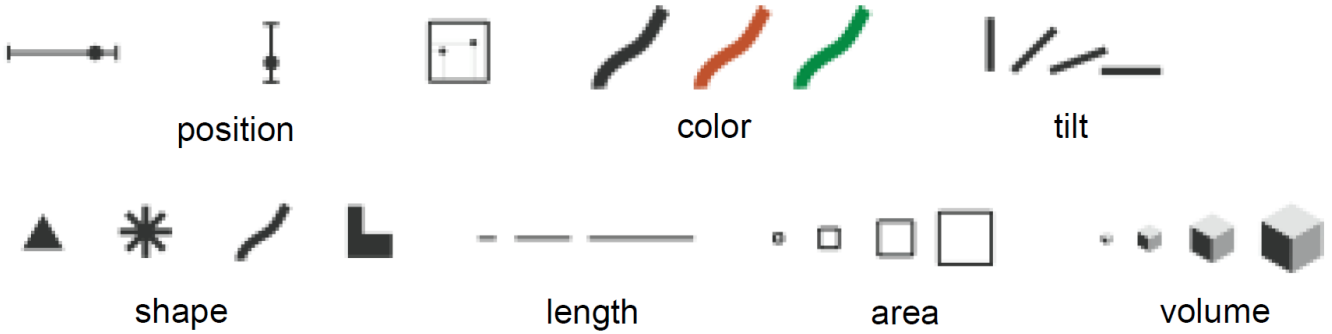
Problem	Data type	Dimensions	Data structure	Type of interaction
Communicate	Categorical	Univariate	Linear	Static
Explore	Ordinal	Bivariate	Temporal	Transformable
Confirm	Numerical	Trivariate	Spatial	Manipulable
		Multivariate	Hierarchical	
			Network	

Visual encodings: mapping data to display variables

- **Marks:** geometric primitives
  - point
  - line
  - area
  - volume



- **Channels:** control appearance of marks
  - position
  - color
  - size
  - shape
  - angle/tilt
  - …



Visual encodings: mapping data to display variables

- Position (2)
- Orientation (1)
- Size (spatial frequency) (1)
- Motion (1)
- Blinking (1)
- Color (3)

~10 degrees of freedom

≈

We can represent 10 different data dimensions “independently” of one another

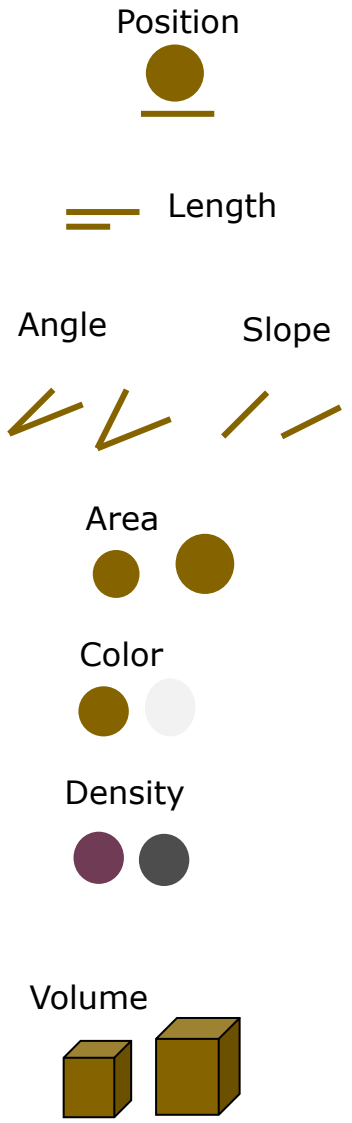
Effectiveness of visual channels

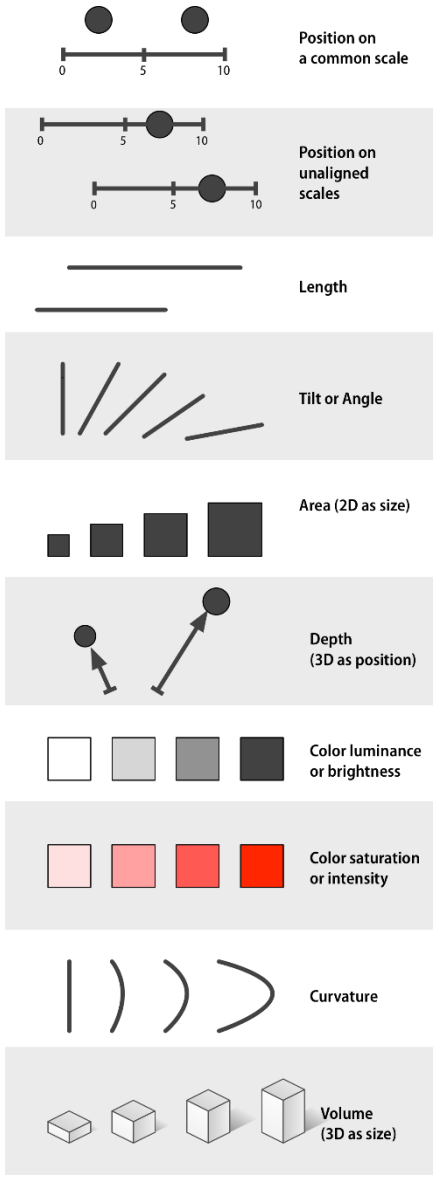
Accuracy ranking for quantitative perceptual tasks

- Position
- Length
- Tilt, angle, slope
- Area
- Hue
- Saturation
- Curvature
- Volume

More Accurate

Less Accurate

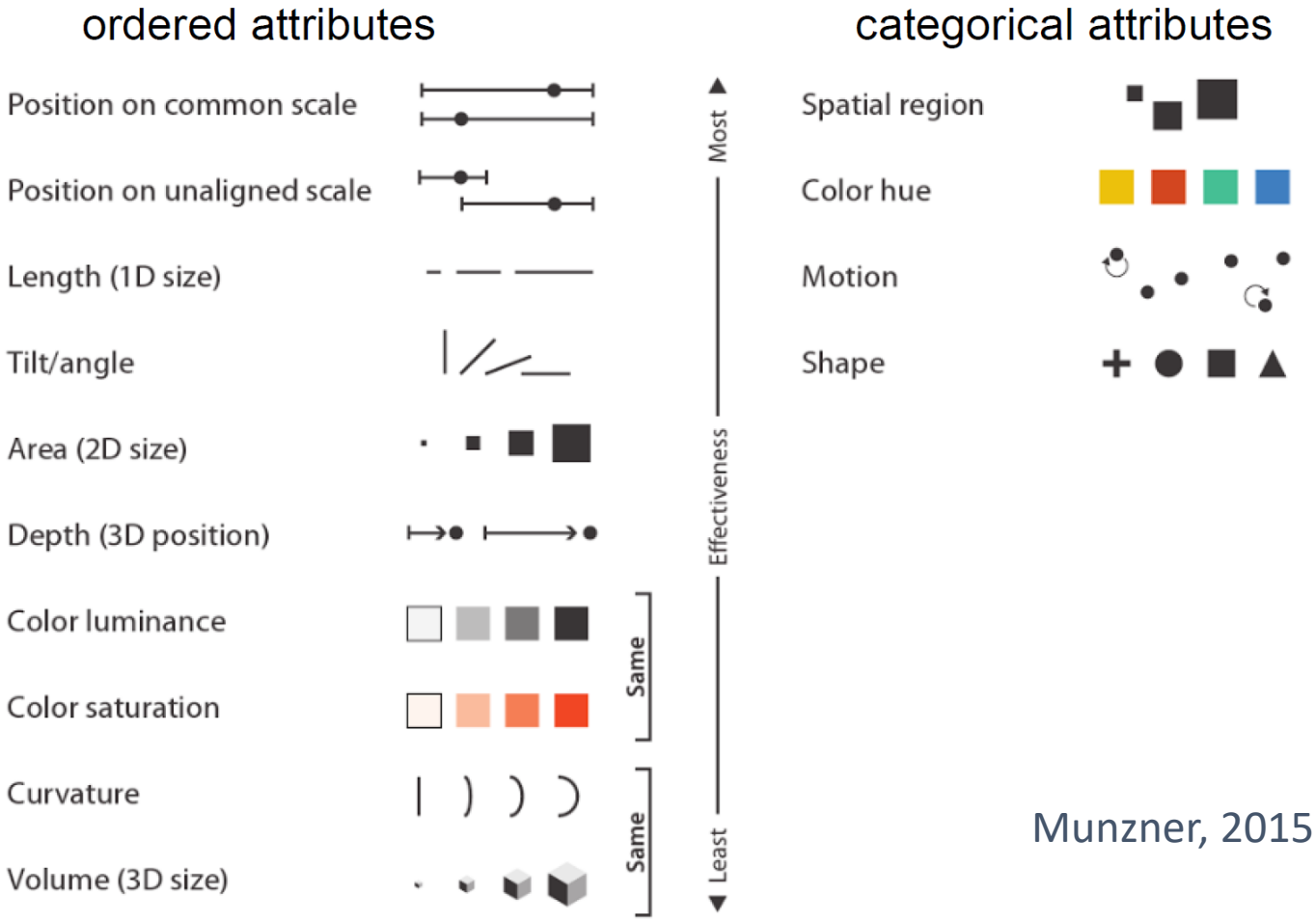




Ranking of Visual Properties for Different Data Types

CATEGORICAL (nominal)	**ORDINAL**	**NUMERICAL** (quantitative)
Position	Position	Position
Color Hue	Density	Length
Texture	Color Saturation	Angle
Connection	Color Hue	Slope
Containment	Texture	Area
Density	Connection	Volume
Color Saturation	Containment	Density
Shape	Length	Color Saturation
Length	Angle	Color Hue

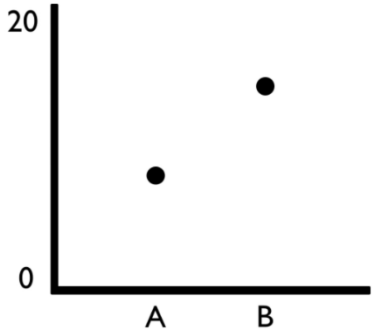
Effectiveness of visual channels



Munzner, 2015

Position: the best channel for all types of data

"Spatial position is such a good visual coding of data that the first decision of visualization design is which variables get spatial encoding at the expense of others"



What does this "tell" us?

- Strongest visual variable
- Suitable for all data types
- Many steps, quantitative judgement

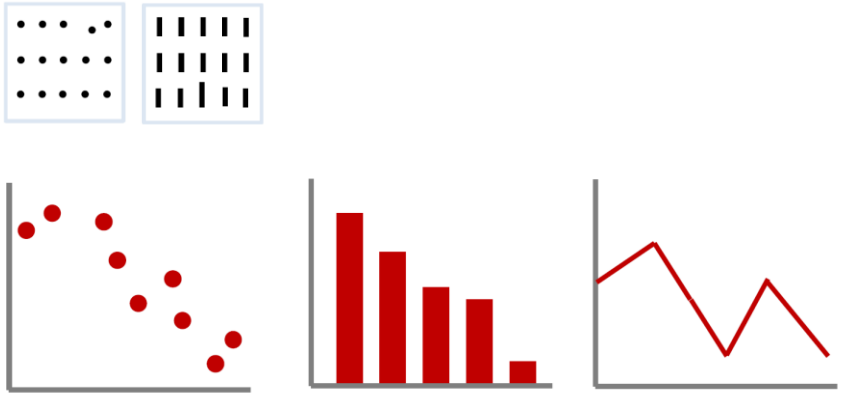
Length: good for numerical (quantitative) data

- Length or height for comparing numerical data

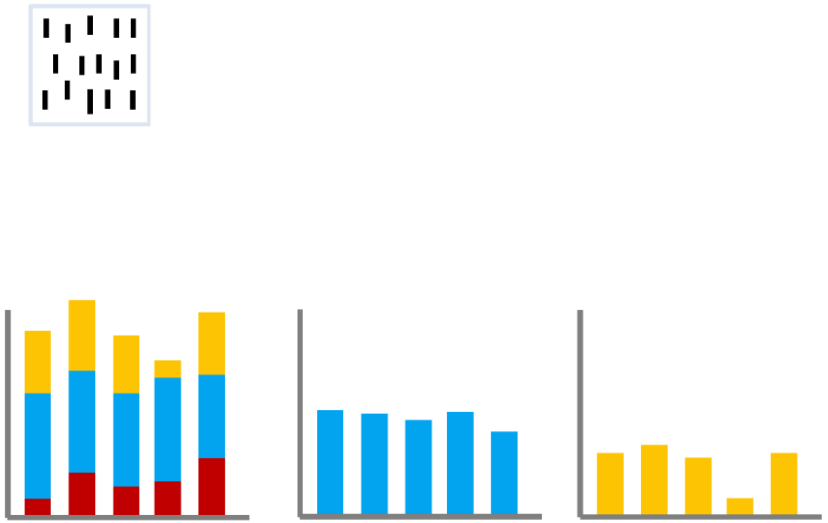
- How much longer?

Channels for numerical (quantitative) data

- Strong:
  - position
  - aligned line length

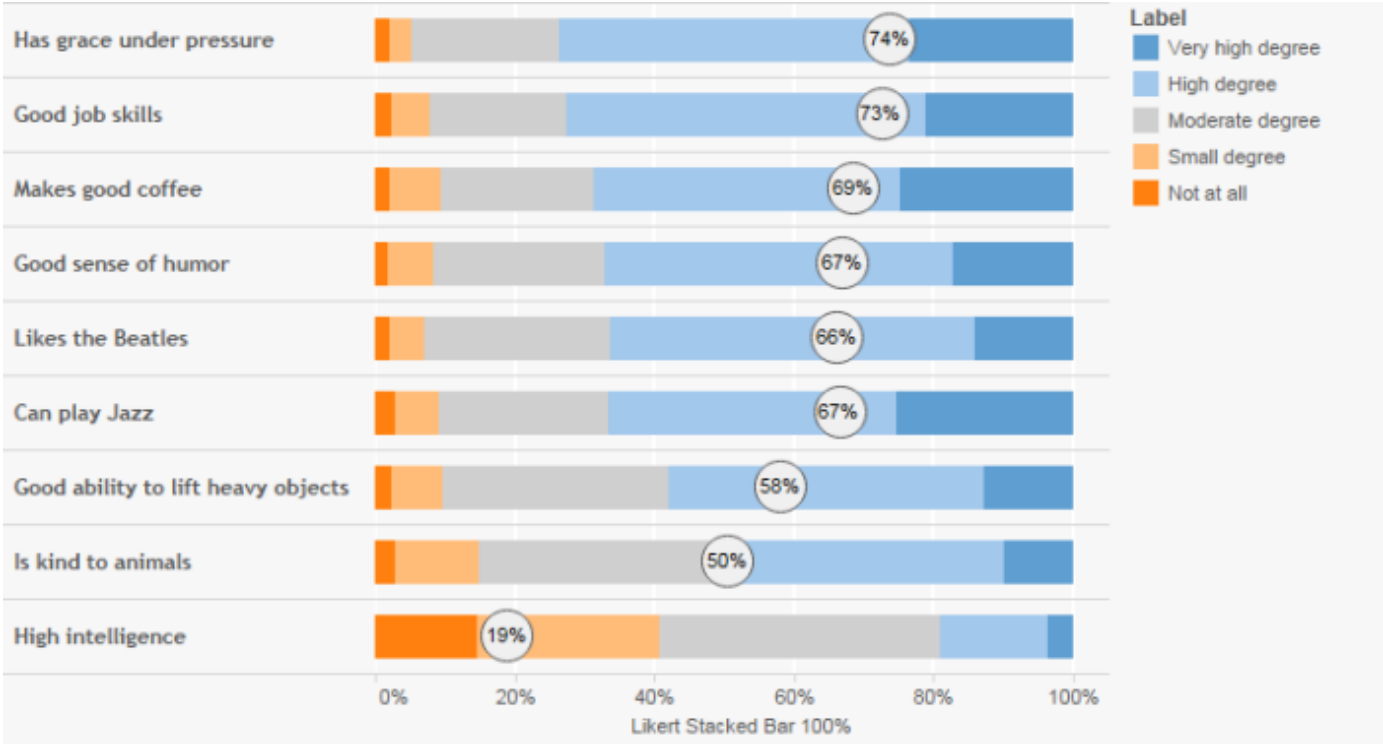


- Much weaker:
  - non-aligned line length



Channels for numerical (quantitative) data

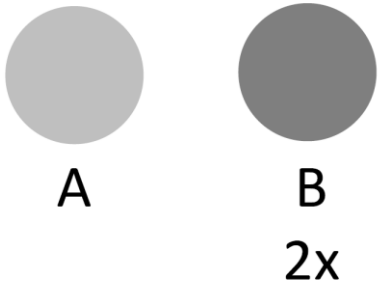
- Much weaker:
  - non-aligned line length: ok for Likert scales



Hue / saturation: good for categorical & ordinal

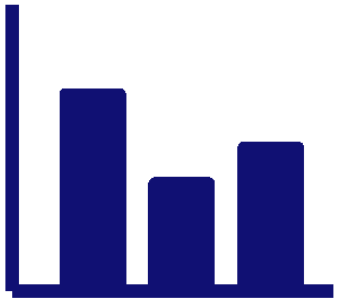
- Color hue, color saturation for distinguishing categorical / ordinal data

- How much darker?

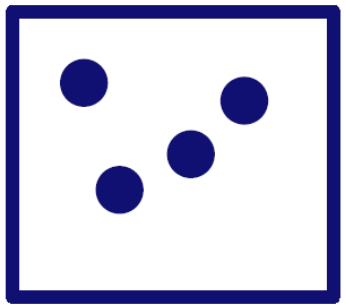


Visual encodings

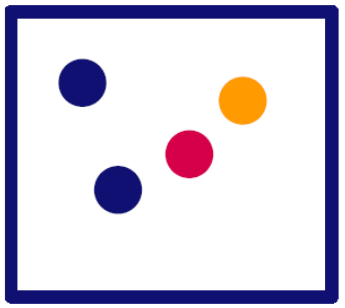
- Combination of marks and channels



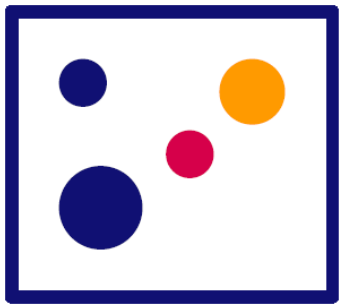
vertical position



vertical
position
+
horizontal
position



vertical
position
+
horizontal
position
+
color



vertical
position
+
horizontal
position
+
color
+
size

Example of marks

Put data in the right order using:

- Shape:

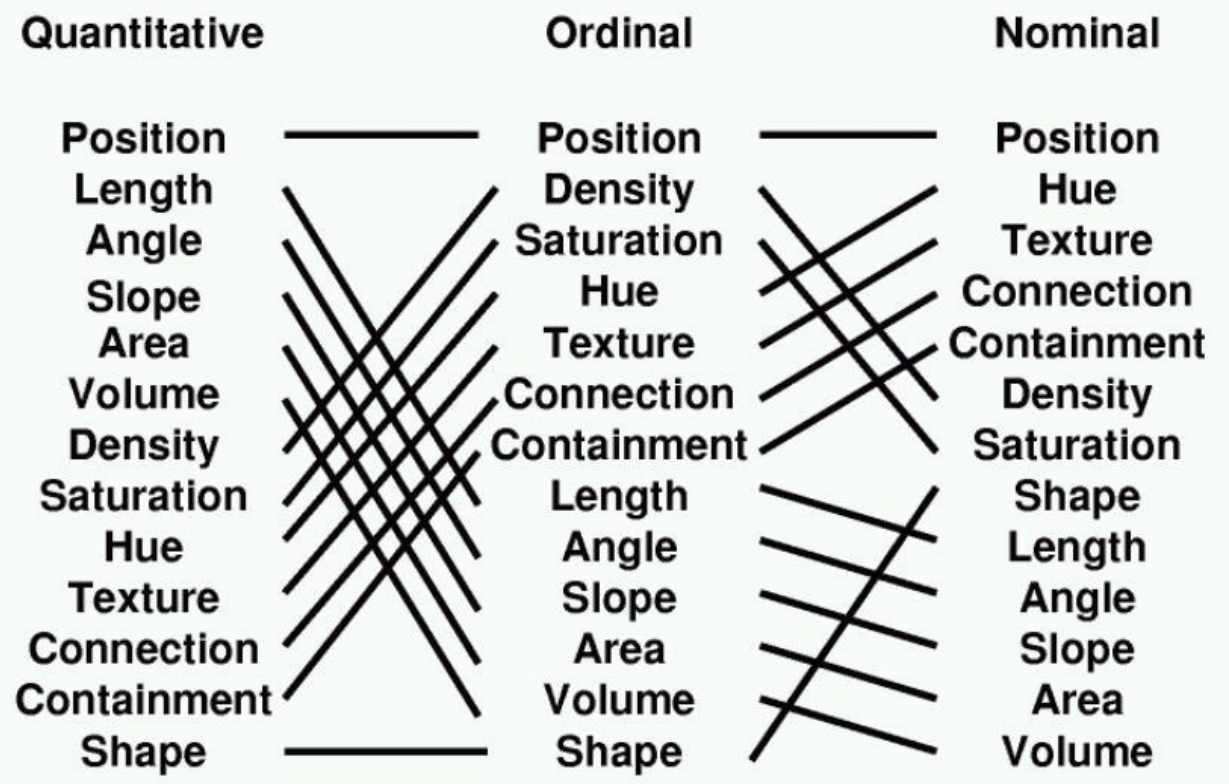
- Orientation:

- Hue:

- Intensity:

- Size:

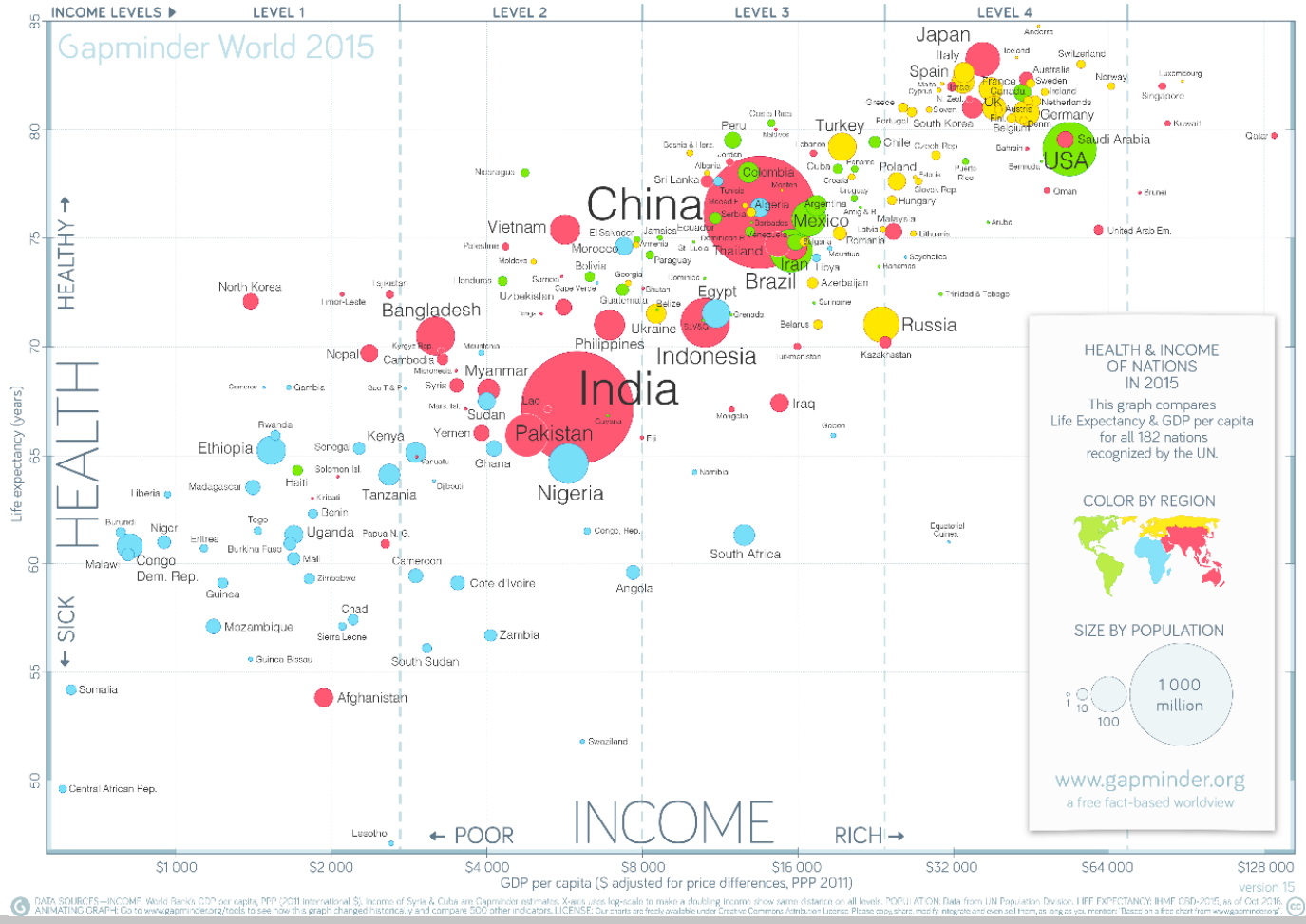
Effectiveness visual channels



Mackinlay, 1986

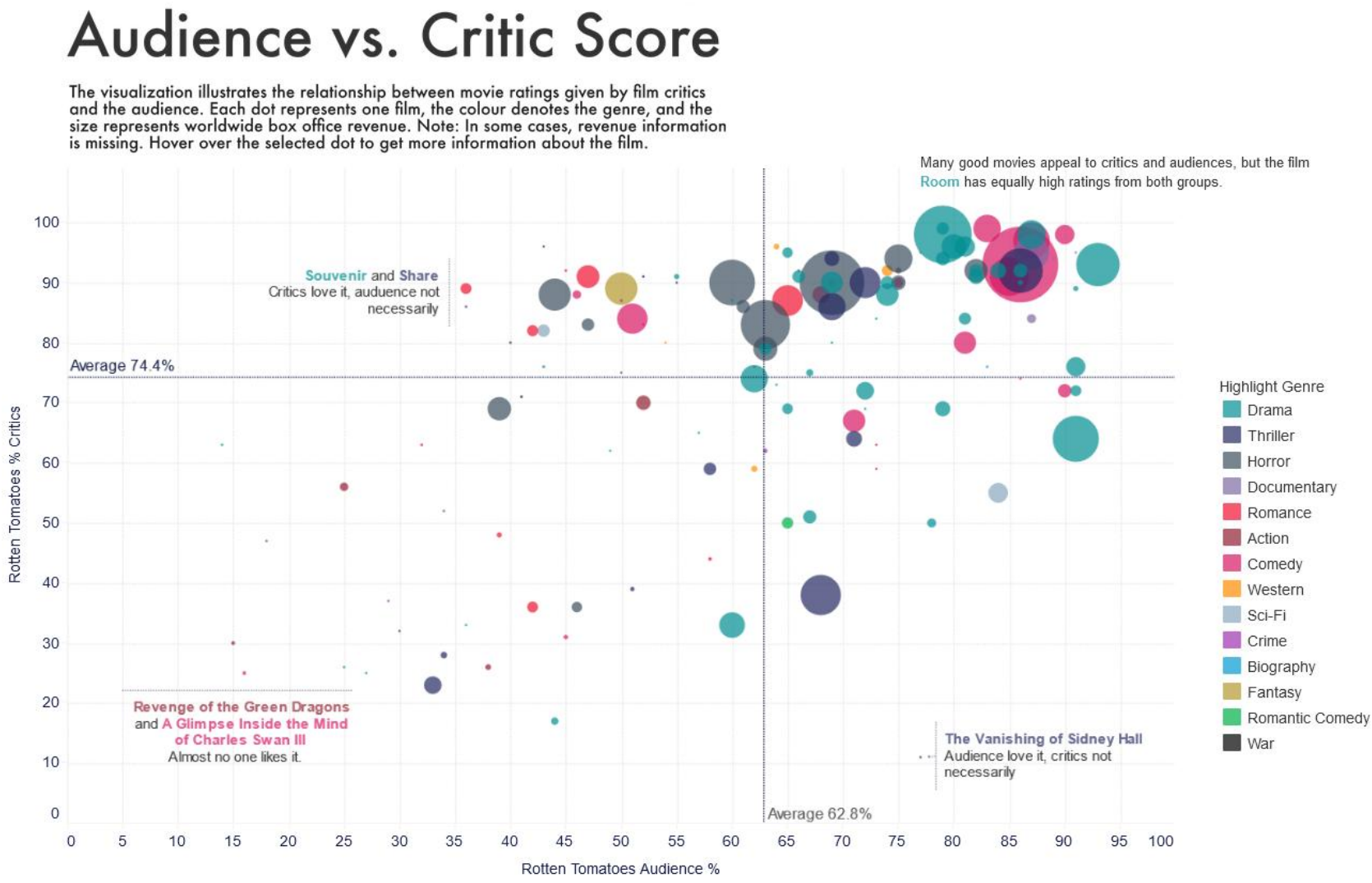
Example: Hans Rosling's mapping

- Most important data (numerical): position
- Population (numerical): size of disks
- Continent (categorical): color
- Time: time
- Two layers

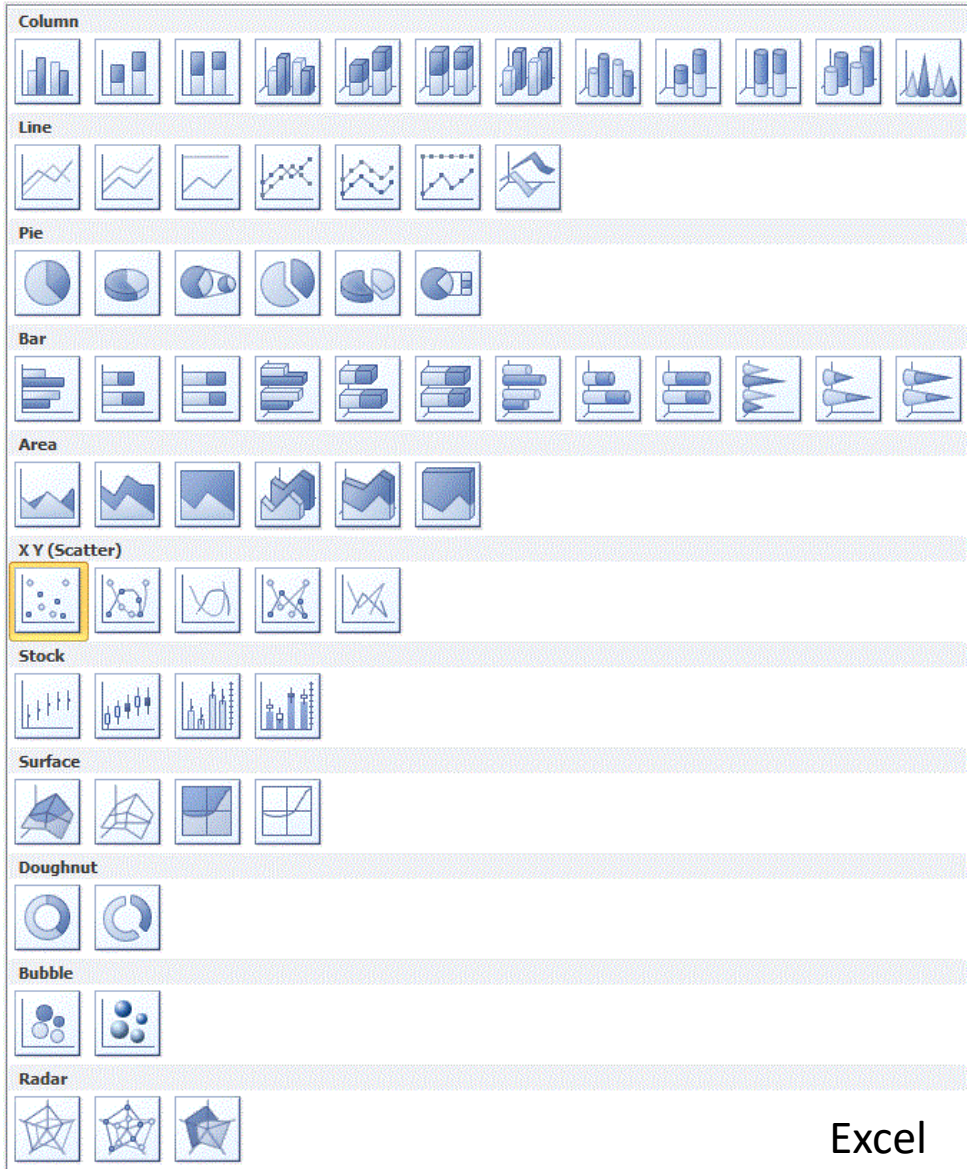


Example: mapping

- Most important data (numerical %): position
- Revenue (numerical): size of disks
- Genre (categorical): color
- Extra layer: details (and interpretation) on hover



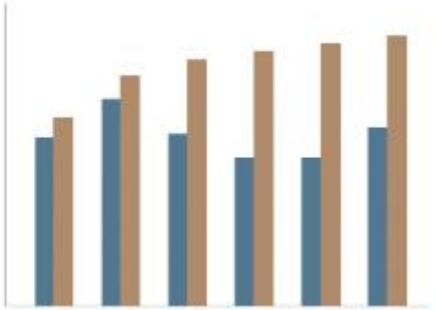
The most common graphs…



Excel

The most common graphs are made of…

Points Lines Bars

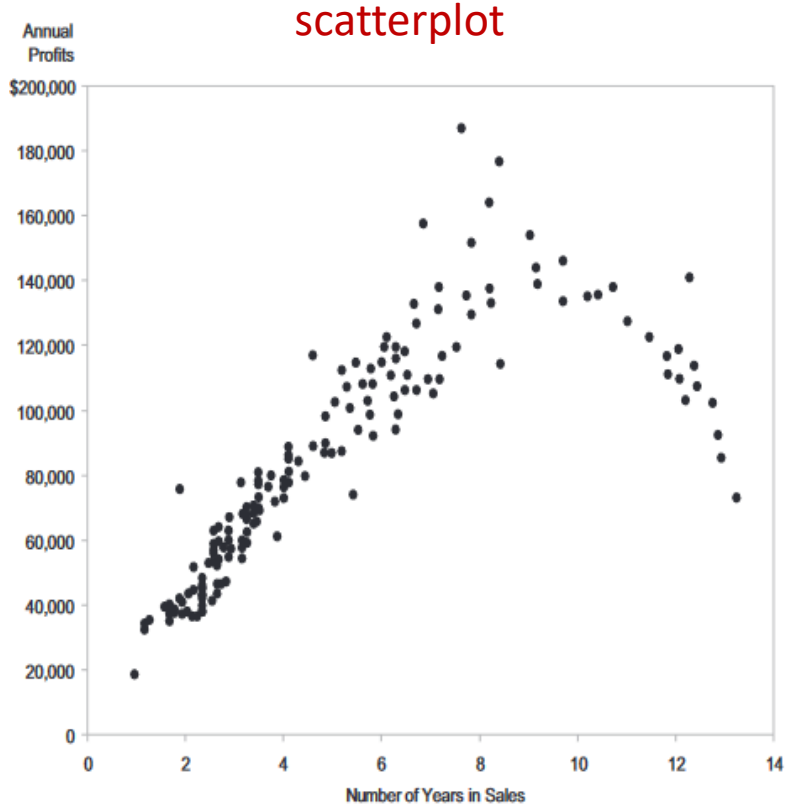


Points

- Small points in the simple shape of circles (dots), squares, triangles, etc., are ideal for pinpointing the precise location of individual values in a graph

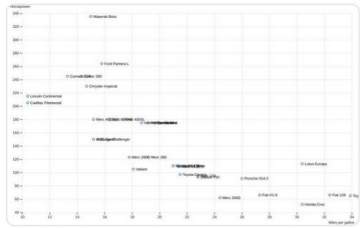
What does this "tell" us?

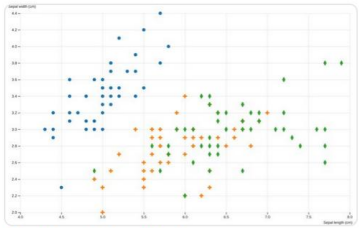
scatterplot

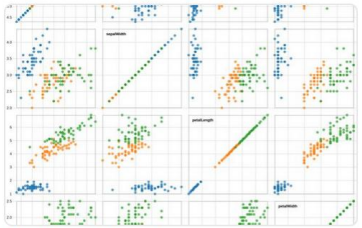


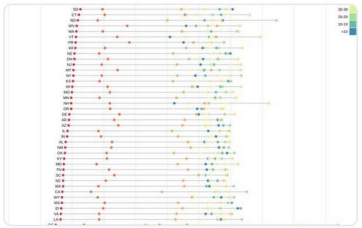
Point-based charts

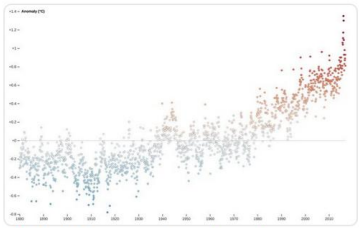
- Scatter plots
- Bubble map / Bubble chart
- Beeswarm (single dimension)
- SPLOM (pairwise dimensional correlations)

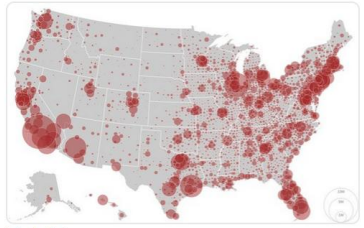

Scatterplot

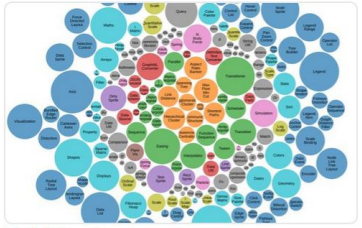

Scatterplot with shapes

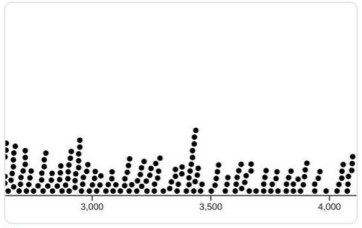

Scatterplot matrix


Dot plot

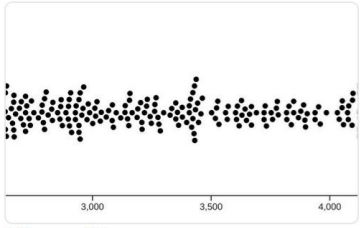

Global temperature trends


Bubble map


Bubble chart



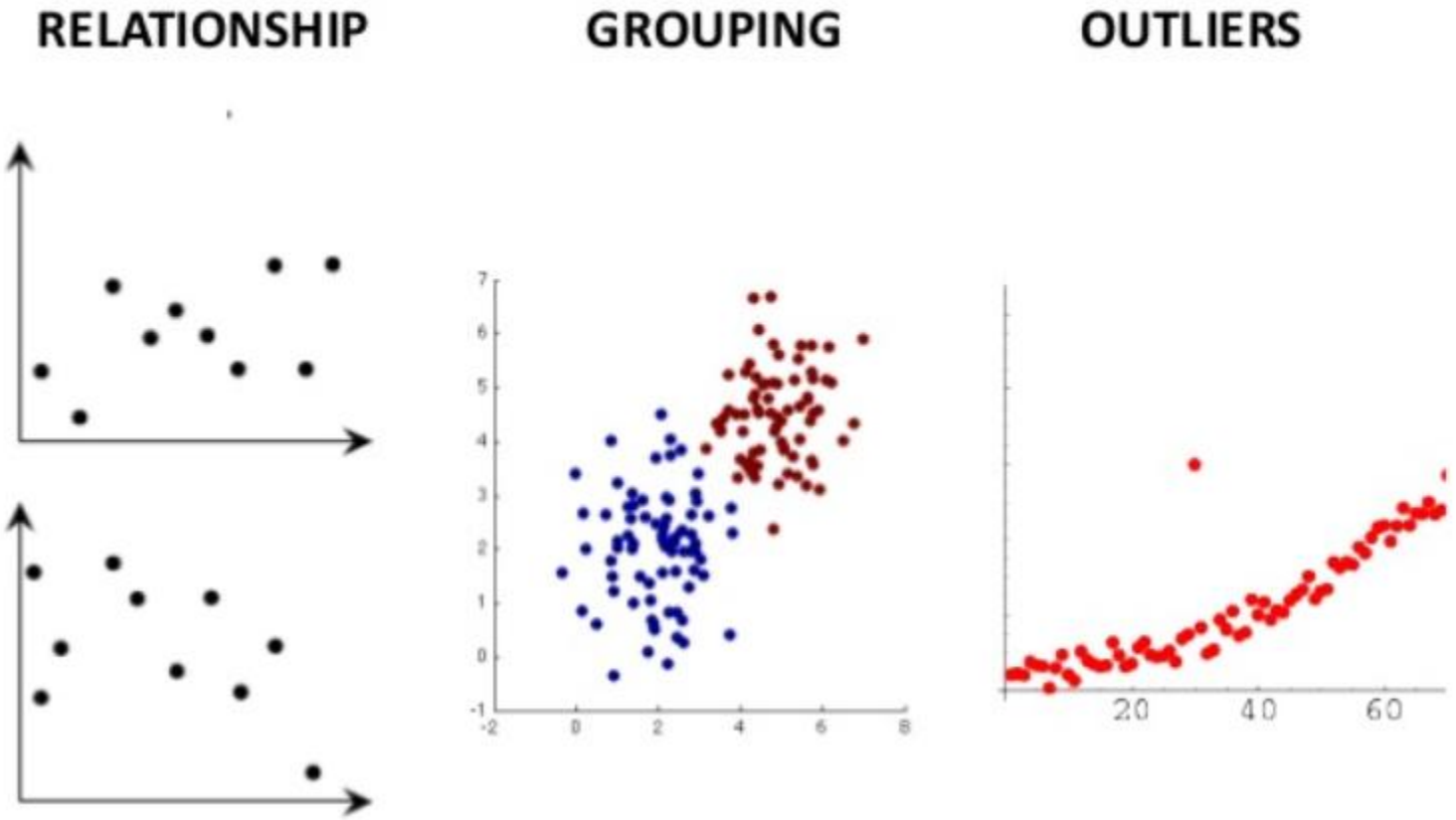
Beeswarm



Mirrored beeswarm

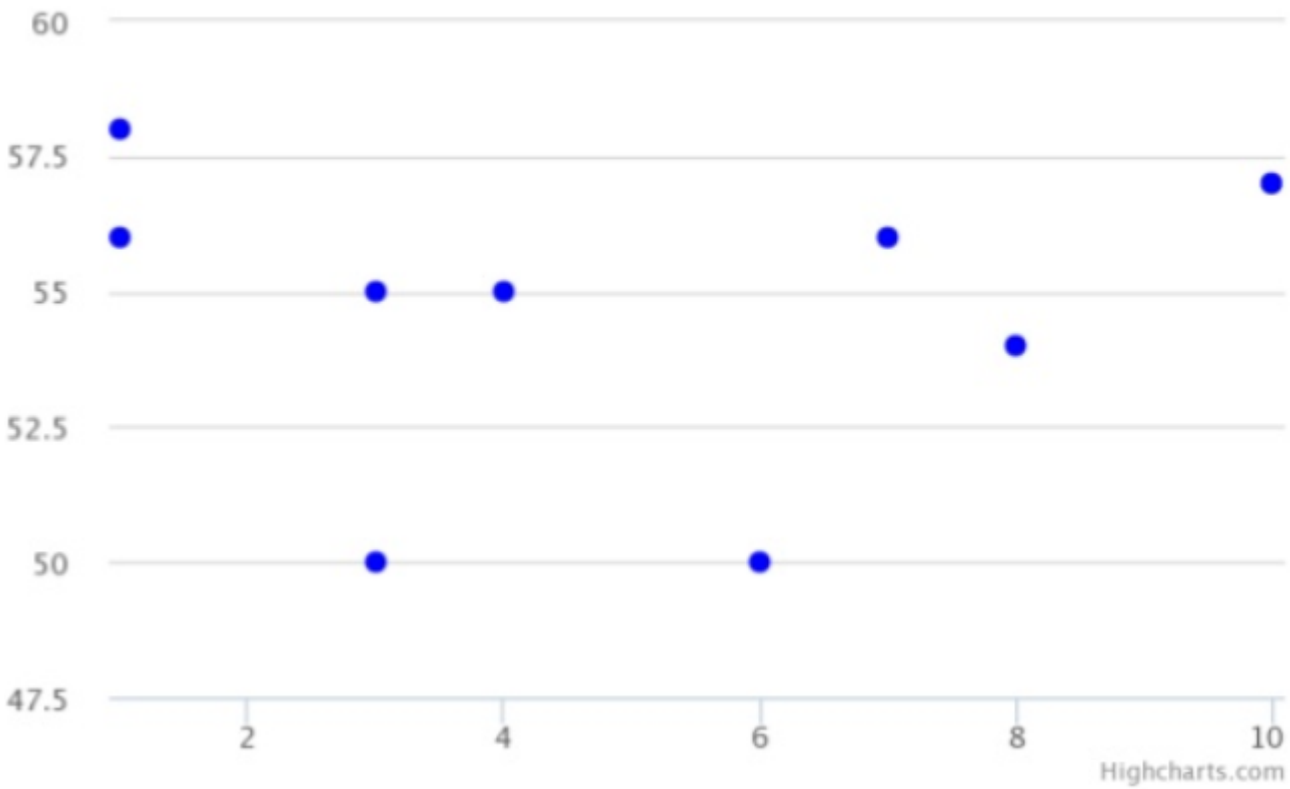
Hertzsprung–Russell diagram

Scatter plot applications



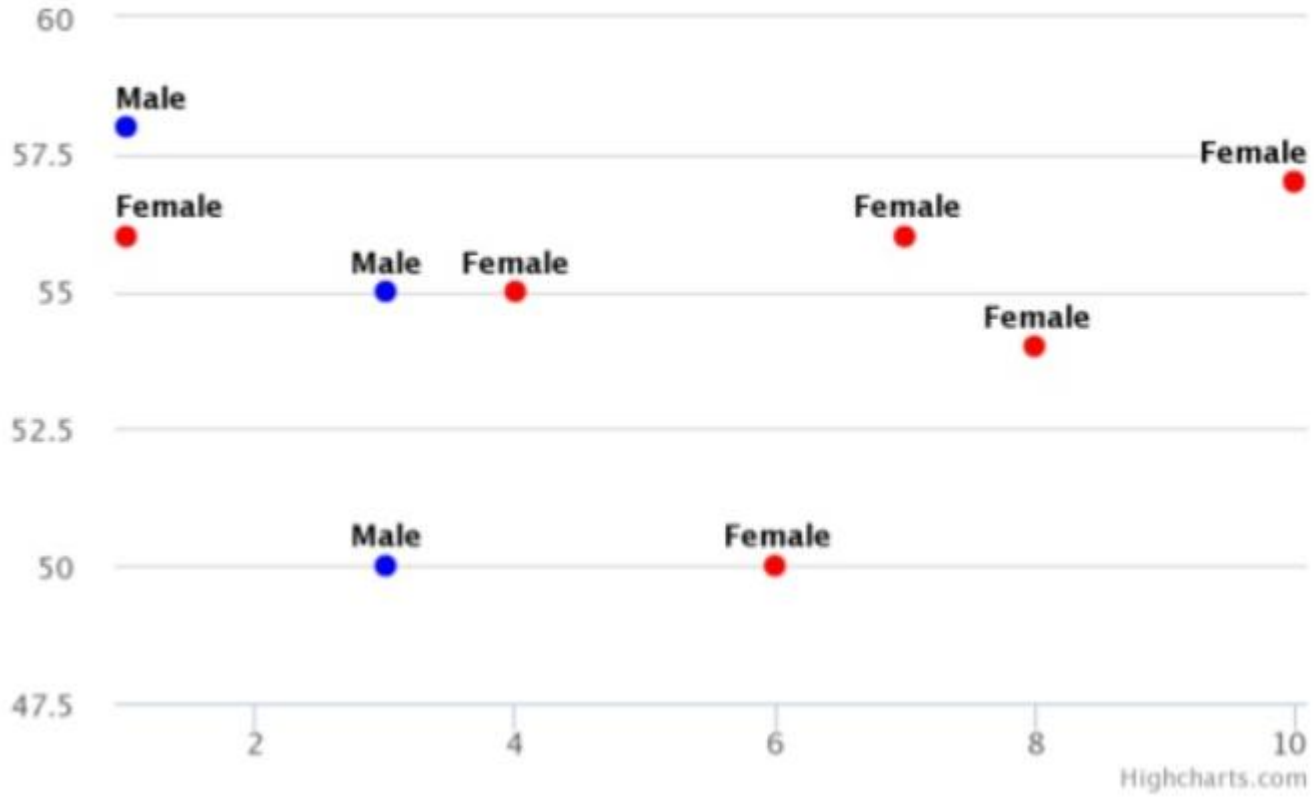
Points

type	mark	channel	data represented
Scatter plot	point	position	2 quantitative



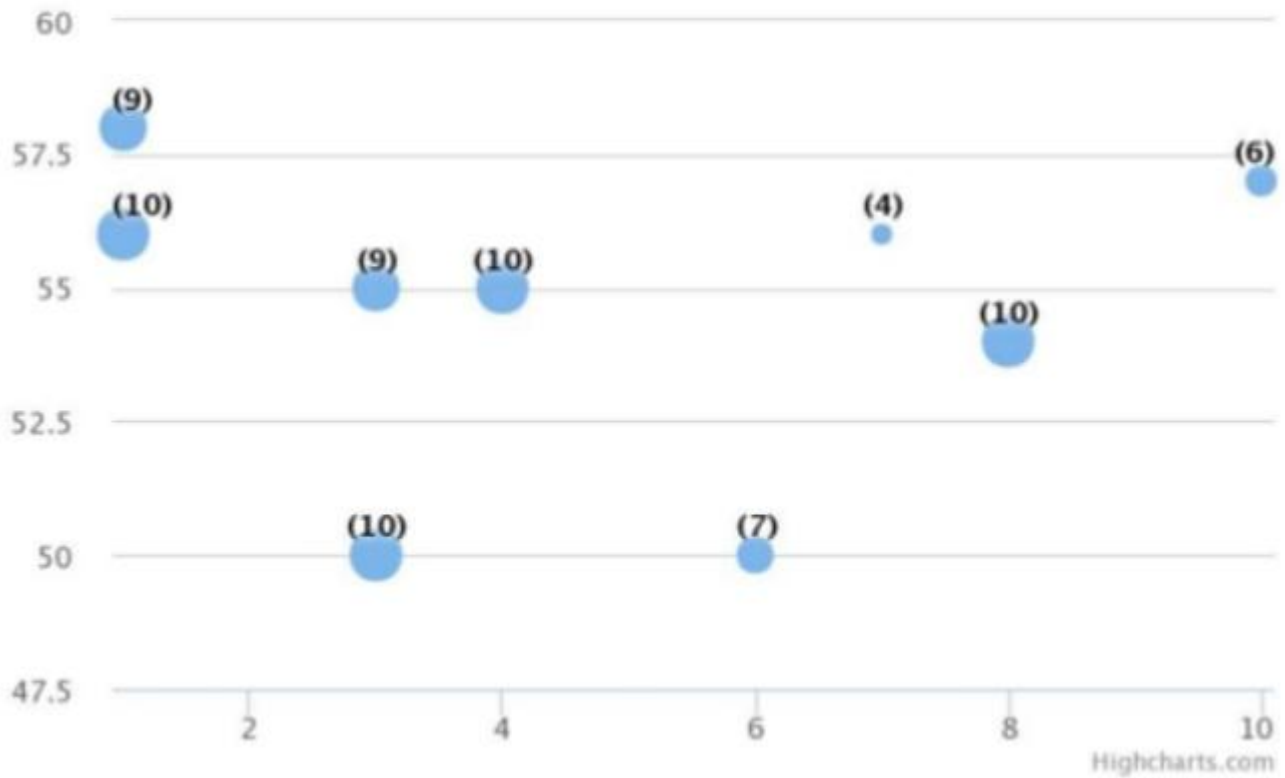
Points

type	mark	channel	data represented
Scatter + hue	point	position, color	2 quantitative, 1 categorical



Points

type	mark	channel	data represented
Scatter + Size	point	position, size	3 quantitative

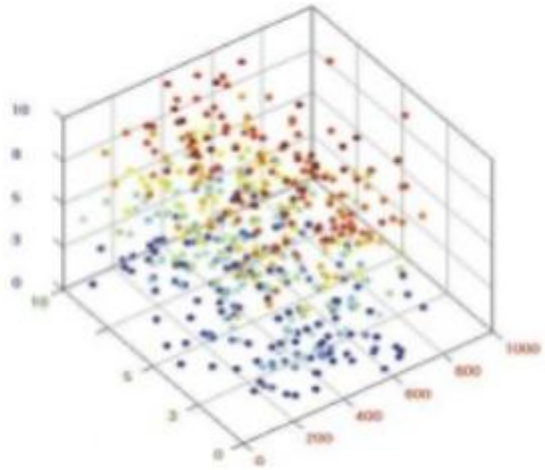


Scatter plot dangers

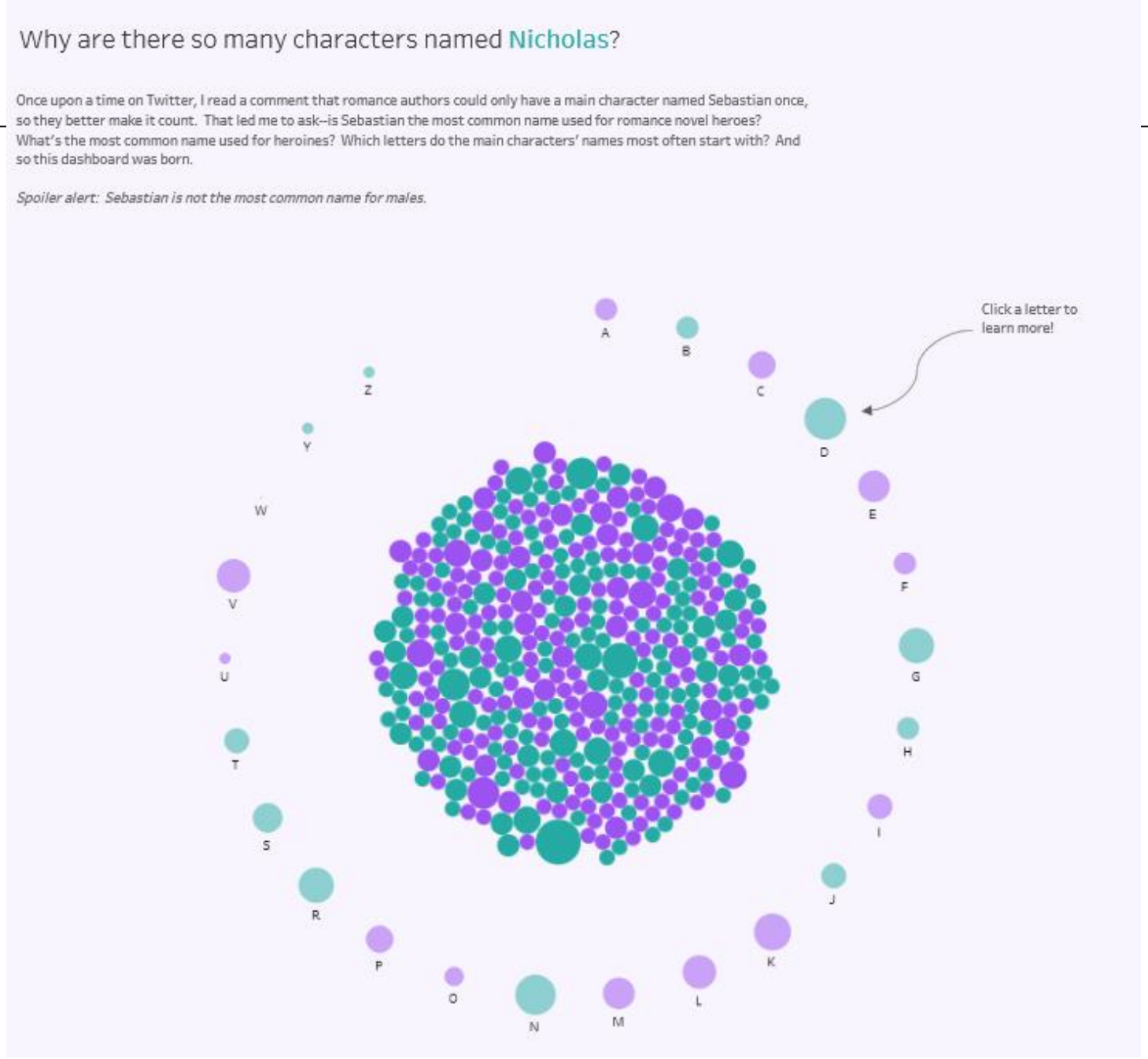
OCCLUSION (DENSITY)

OCCLUSION (OVERLAP)

3-D



Bubble charts



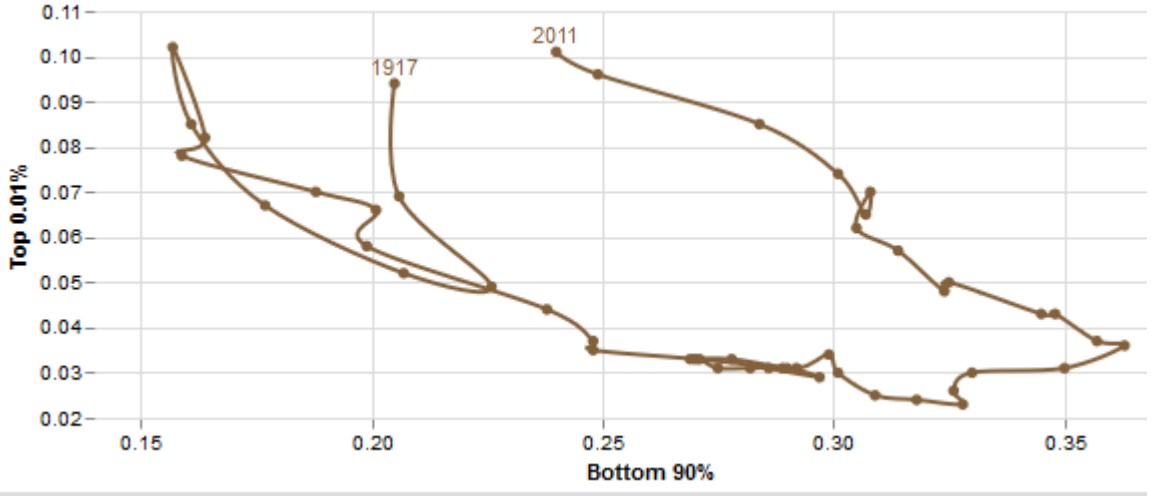
Point patterns and examples of their use

Category	Pattern	Examples
Linear trends	Points are arranged in a way that forms a pattern that looks like a straight line	• Points are going up from left to right, indicating a positive correlation between employee's heights in inches and their salaries in dollars • Points are going down from left to right, indicating a negative correlation between per capita wealth and infant mortality
Non-linear trends	Points are arranged in a way that forms a pattern that looks like a curved line	• Points are going up from left to right at an ever-increasing rate, indicating not just growing sales, but also a growing rate of sales • Points are going down from left to right, indicating fewer device failures as temperature rises until the temperature reaches 100° and failures begin to increase
Concentrations	Dense sets of points appear in particular areas of the graph	• A large concentration of points appears in the upper left corner of the graph, indicating a large number of small orders with large profits • A large collection of points appears in the center of the graph, indicating that most orders are moderate in both size and profits
Clusters	Sets of points appear to be set apart from other points	• A small group of isolated points appears in the lower right corner of the graph, indicating a distinct group of large orders with low profits
Gaps	Area in the graph where no points appear in the midst of surrounding points	• In a scatterplot that correlates home sales amounts and the number of days on the market, there is a glaring gap between 45 and 55 days where no homes appear
Randomness	Points are arranged randomly in the graph, without a discernible pattern	• A random scattering of points throughout the graph indicates an unexpected lack of correlation between marketing expenses and sales
Exceptions	Points that stand out as different from the norm	• A single point appears alone in the lower right corner of the graph indicating a customer who places many orders but receives a low discount

Stephen Few, 2006

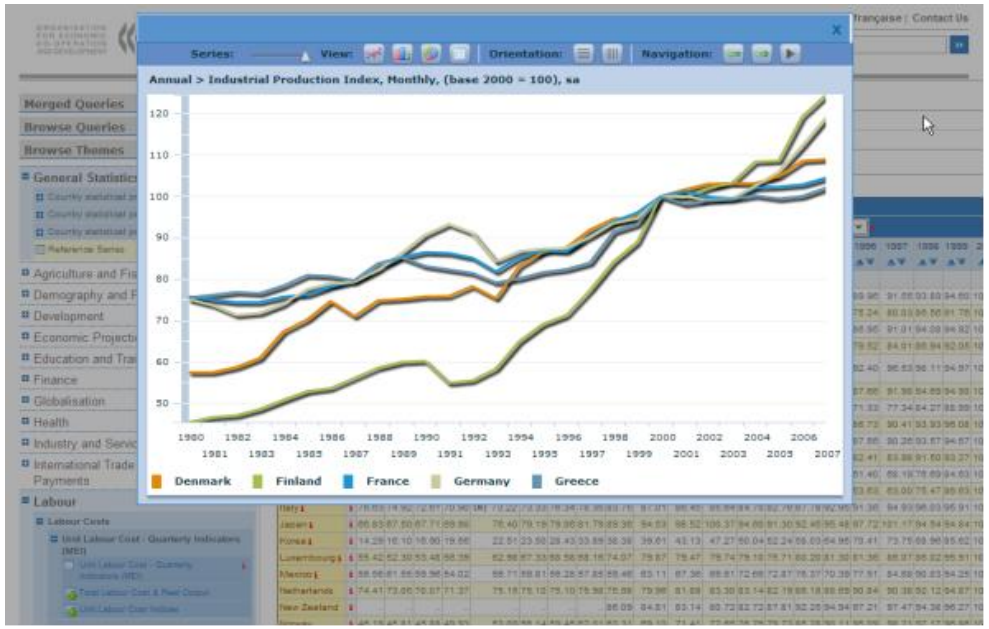
Lines

- By connecting points with a line, we can transform a display of individual values into a **story of transition** from one value to the next
- connects a series of points that represent individual measurements with line segments



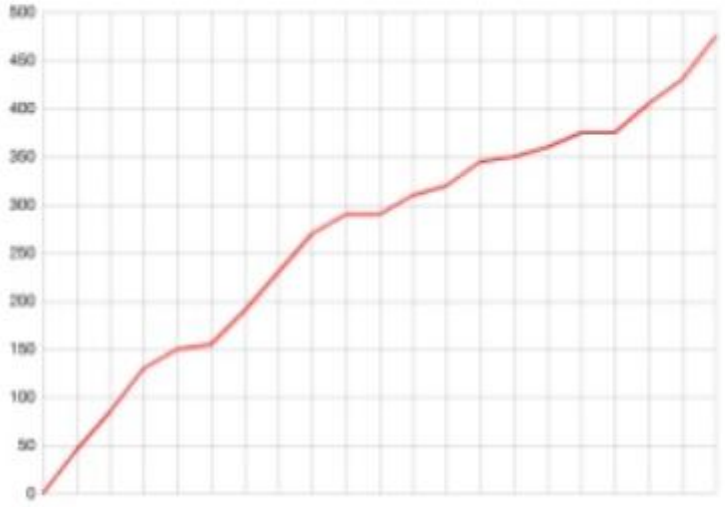
Lines

- Especially useful for **time-series** relationships, i.e. line charts are used to visualise a trend in data over intervals of time – a time series – thus the line is often drawn chronologically

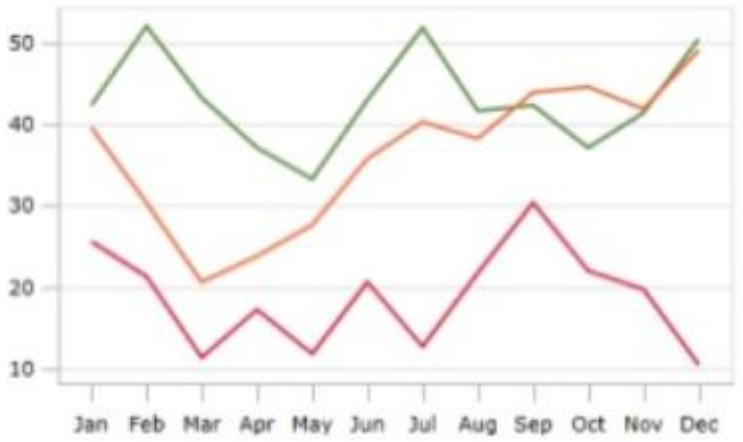


Line chart applications

PATTERN OVER TIME

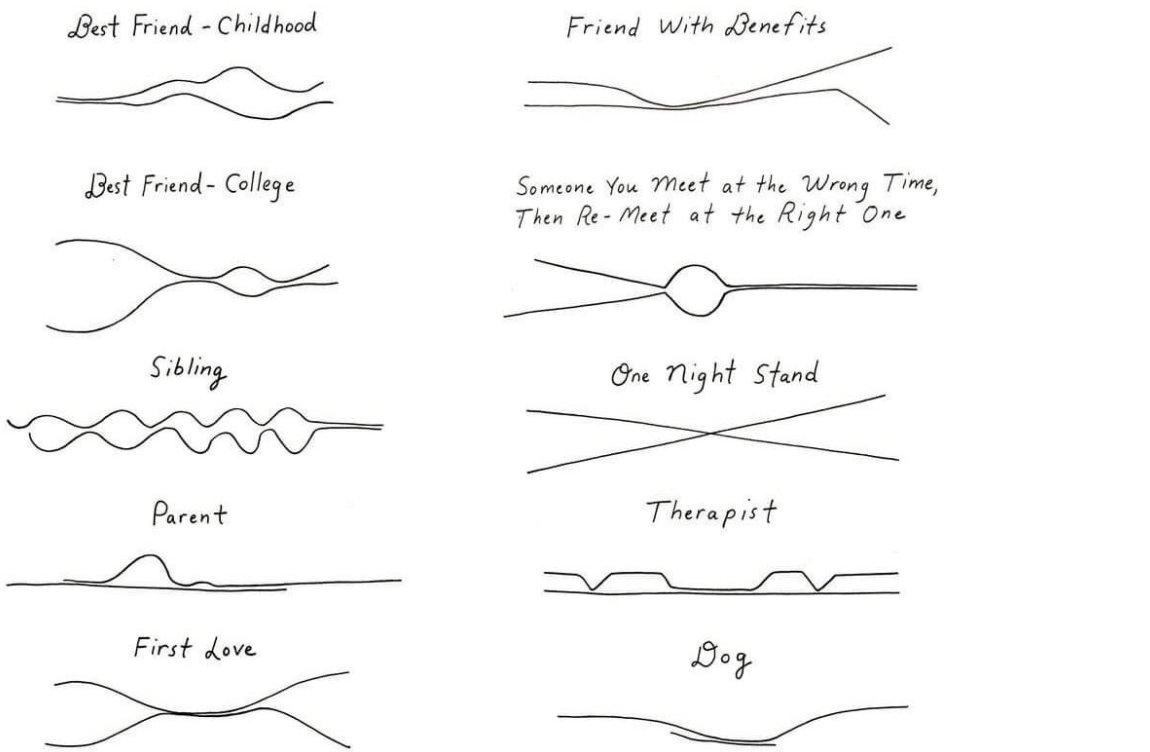


COMPARISON

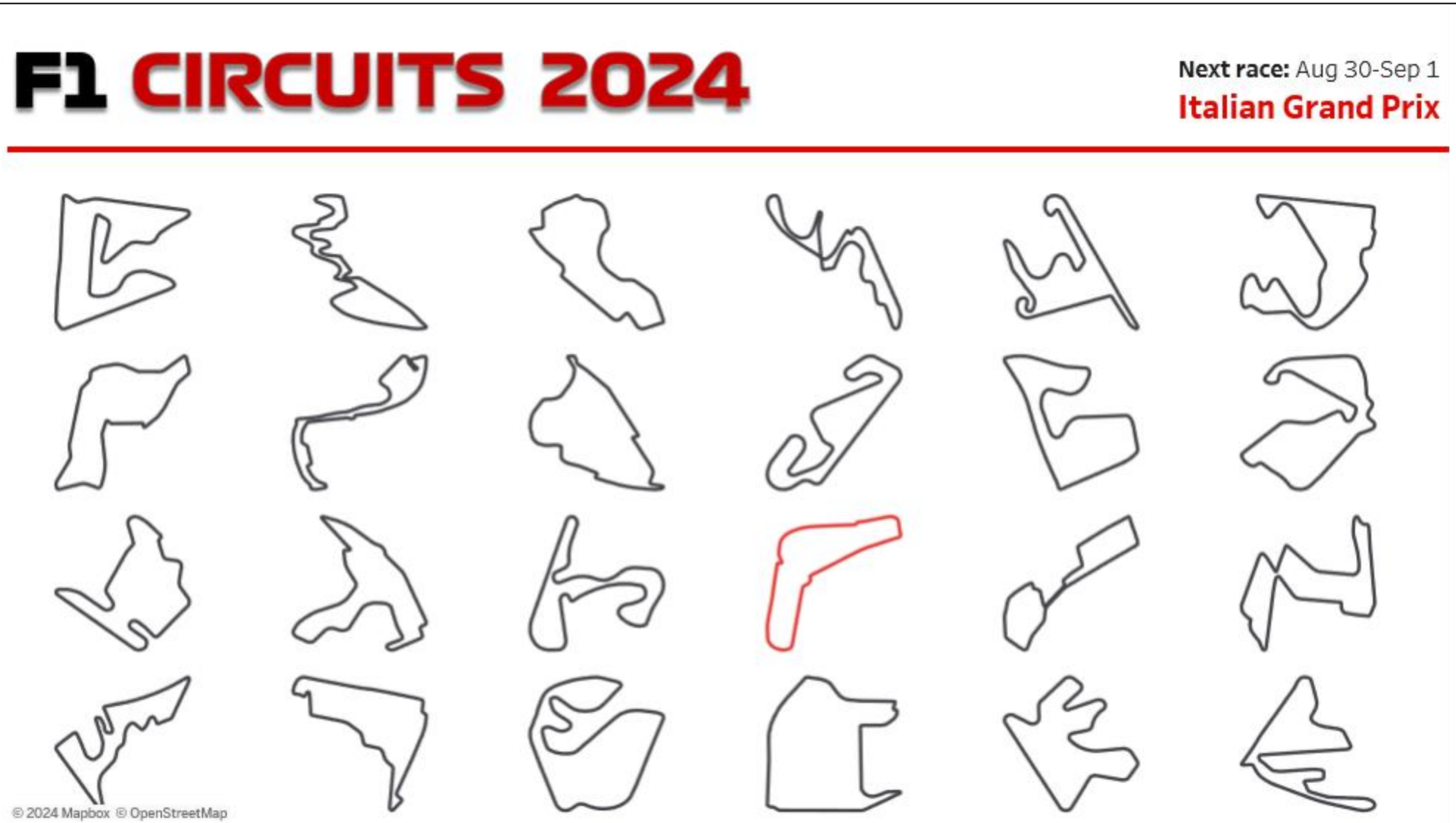


Lines

- Even simple lines can be very effective in telling a story…

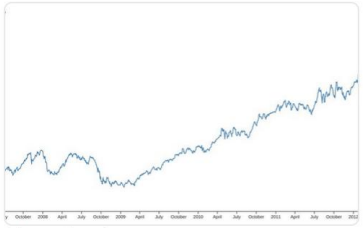


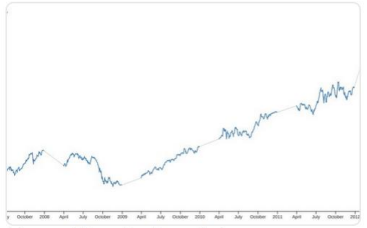
Lines

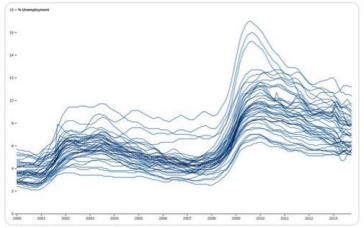

F1 CIRCUITS 2024
Next race: Aug 30-Sep 1
Italian Grand Prix
© 2024 Mapbox © OpenStreetMap

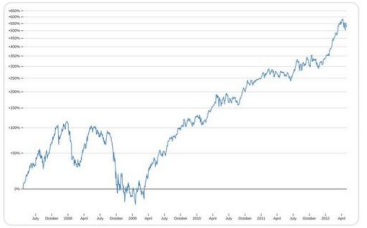
Lines

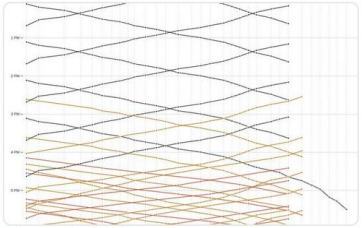
- Line charts
- Candlestick chart, slope charts
- Variable-color line, Gradient, Threshold encoding


Line chart

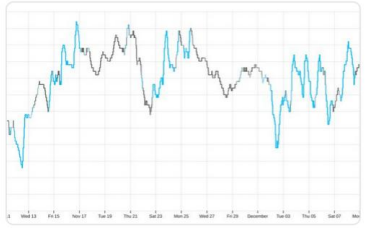

Line with missing data

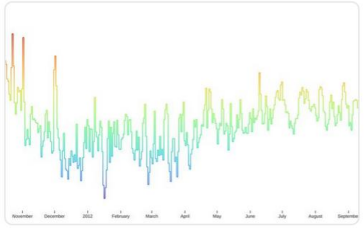

Multi-line chart

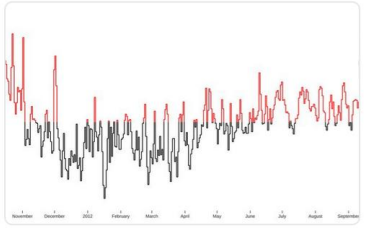

Change line chart

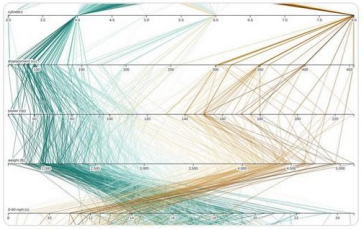

Marey's trains

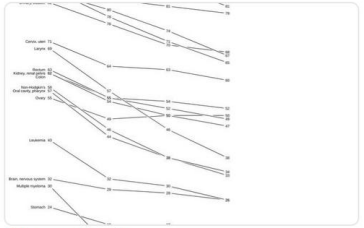
Candlestick chart

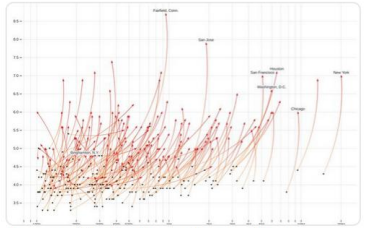

Variable-color line

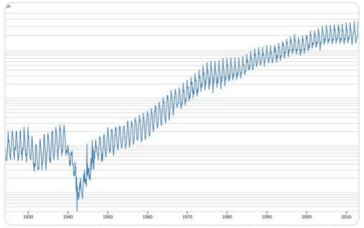

Gradient encoding


Threshold encoding


Parallel coordinates

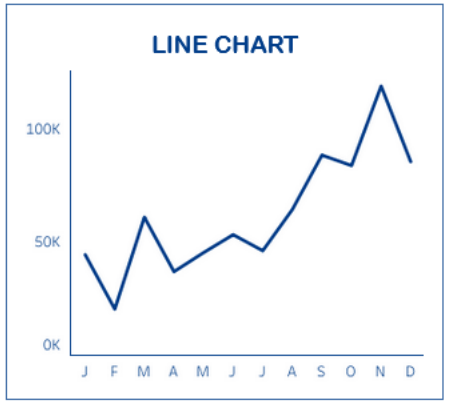

Slope chart

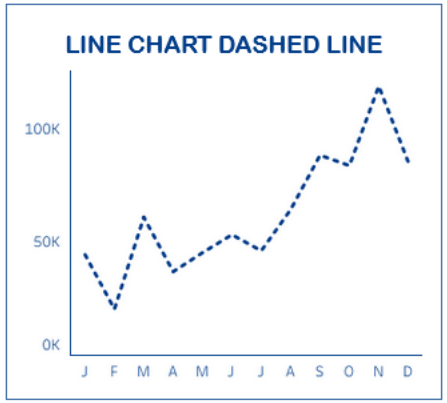

Inequality in American cities

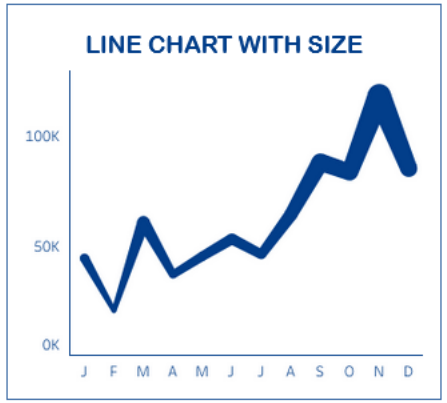

New zealand tourists, 1921–2018

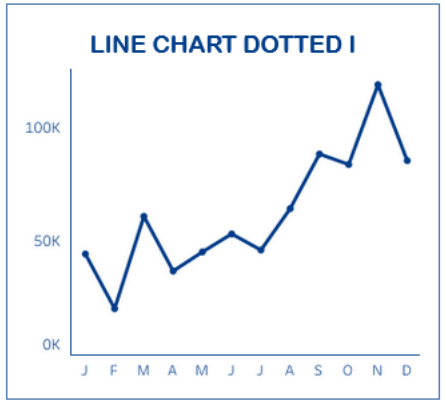
https://observablehq.com/@d3/gallery

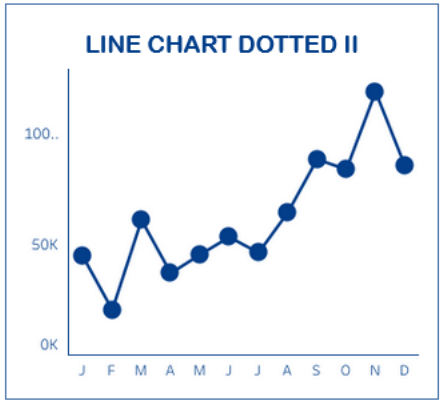
Line & Area chart catalogue

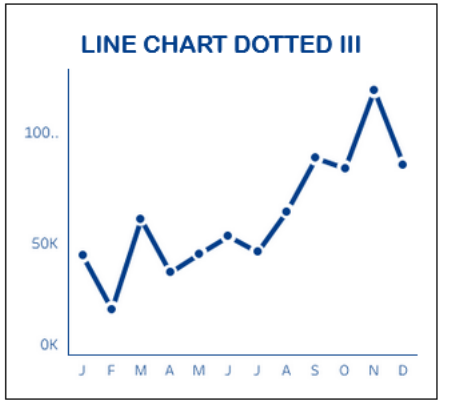






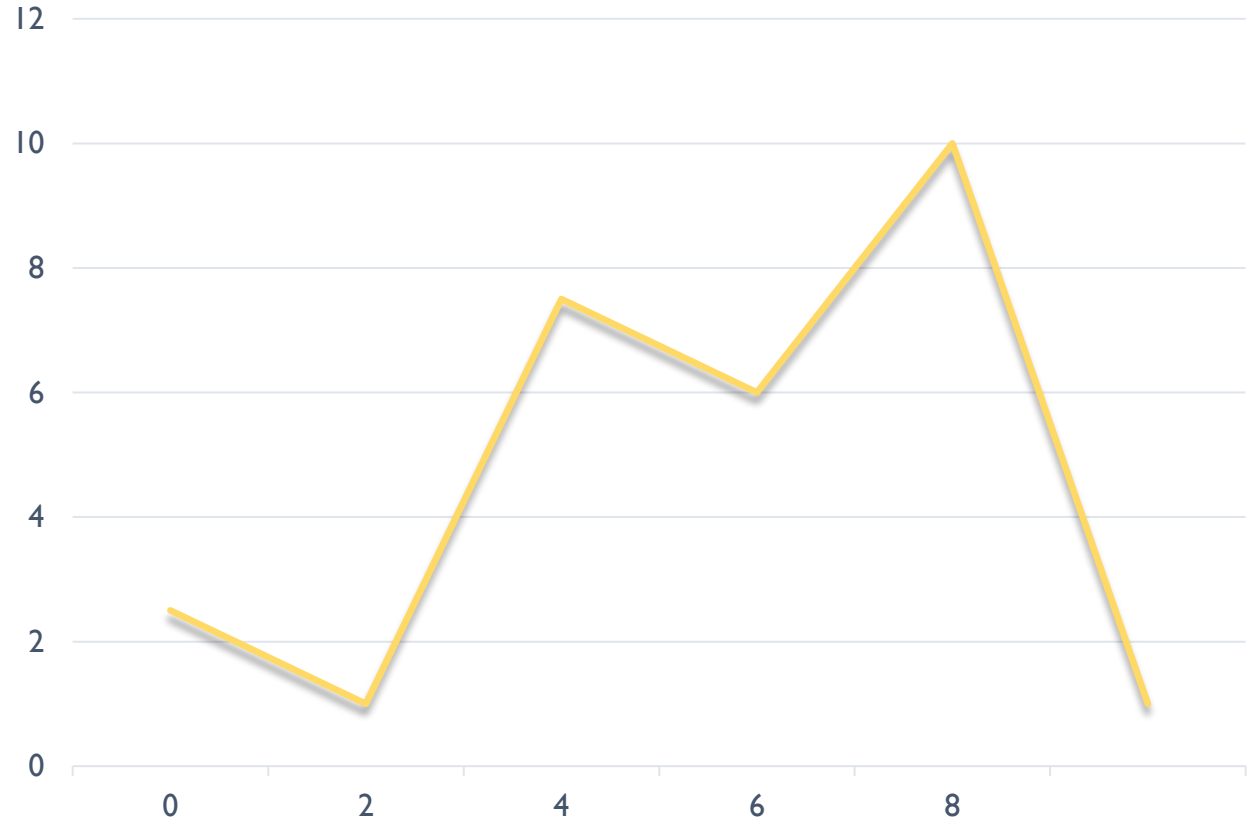






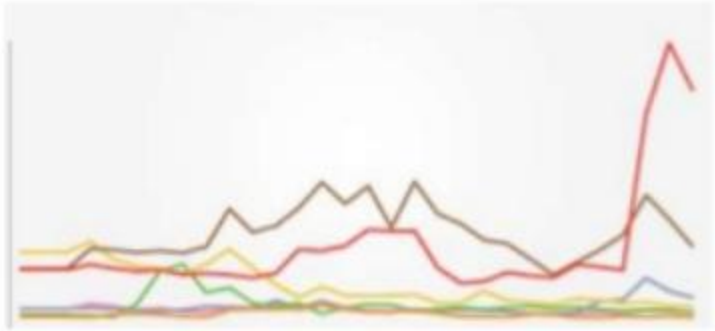
Lines

type	mark	channel	data represented
line chart	line	position (orientation)	2 quantitative

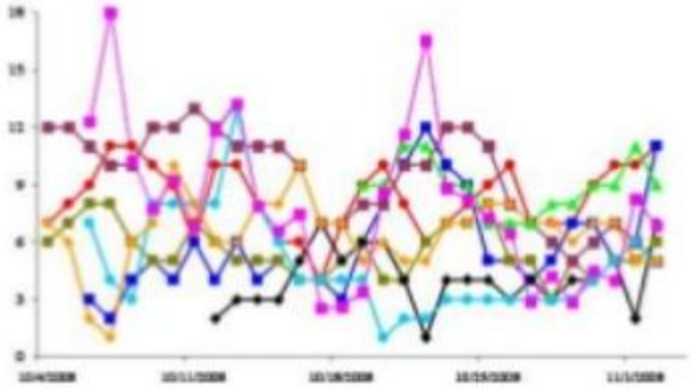


Line chart dangers

X SCALING



Y SCALING



OVERLOAD

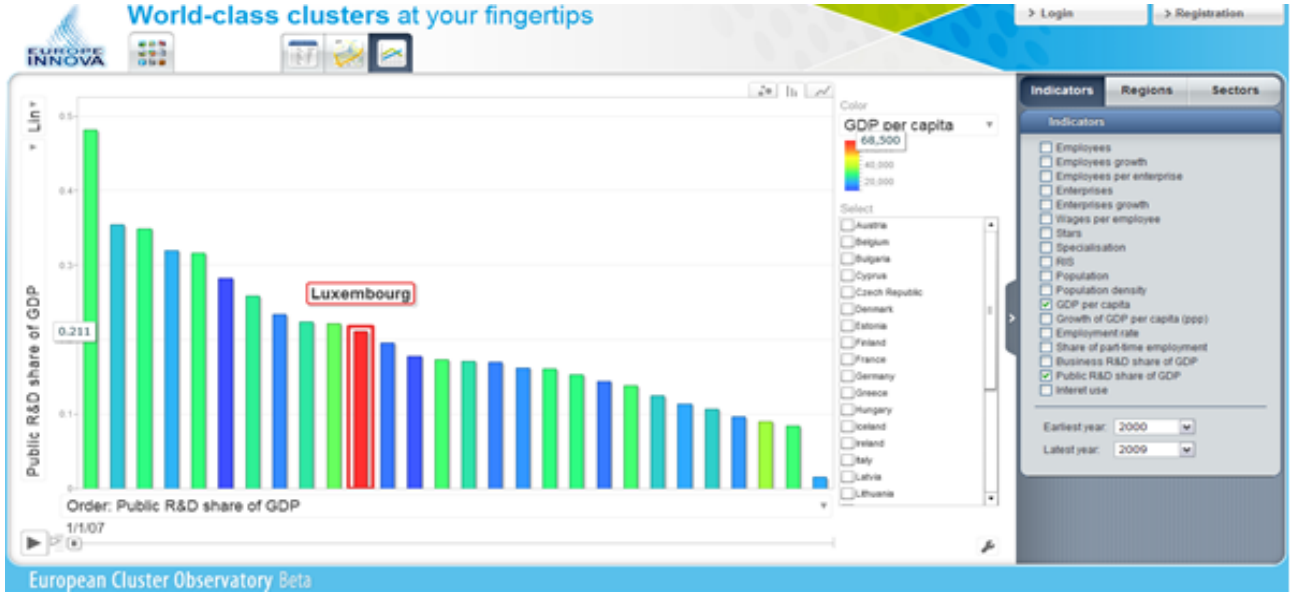
Line patterns and examples of their use

Pattern	Examples
Unidirectional lines	• A line that encodes monthly sales is heading down during the course of the year • A line that encodes a company's total headcount has remained flat for the last five years, despite ups and downs in particular departments
Curved lines	• Sales of a product increased rapidly soon after its introduction but began to increase at a much slower rate after a few months and eventually began to decrease after a few years
Repeating line patterns	• Sales are at their lowest in the first month of each quarter, rise somewhat in the second month, and hit their peak in the final month, then drop again at the beginning of the next quarter to begin the cycle again
Co-variation among multiple lines	• Revenues and profits went up and down together throughout the year • Sales always increased in relation to the number of marketing mailings, but didn't register until five days after the mailings
Intersecting lines	• In June of the year, the decline in coffee sales and the increase in soft drink sales intersected and changed positions in a ranking of sales by product
Smooth vs. jagged lines	• Expenses in the Human Resources department tended to change slowly and smoothly compared to the Information Systems department, which exhibited large and rapid increases in expenses from time to time
Exceptions	• Sales in the month of August decreased far below the norm

Stephen Few, 2006

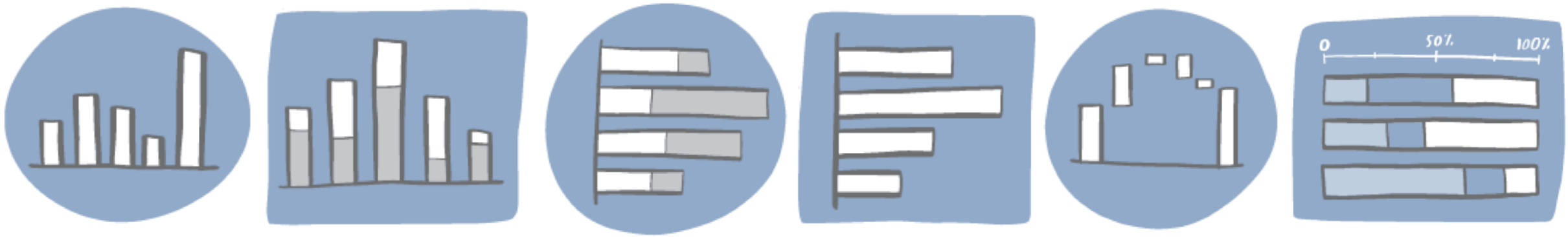
Bar charts / Histograms

- Charts with rectangular bars with **lengths proportional** to the values that they represent
- useful for visualising **summary** data across various **categories** (how a quantity distributes across a set of categories)
- useful for **comparisons of magnitudes**
- used for plotting data with **discrete values** ('discontinuous')



Bars

What is a bar chart?

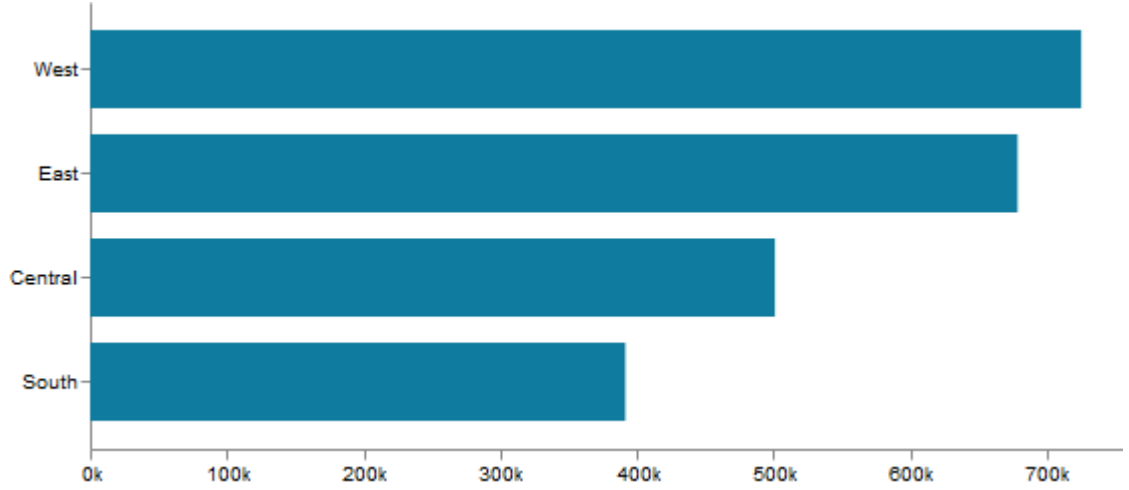


https://www.storytellingwithdata.com/blog/2020/2/19/what-is-a-bar-chart

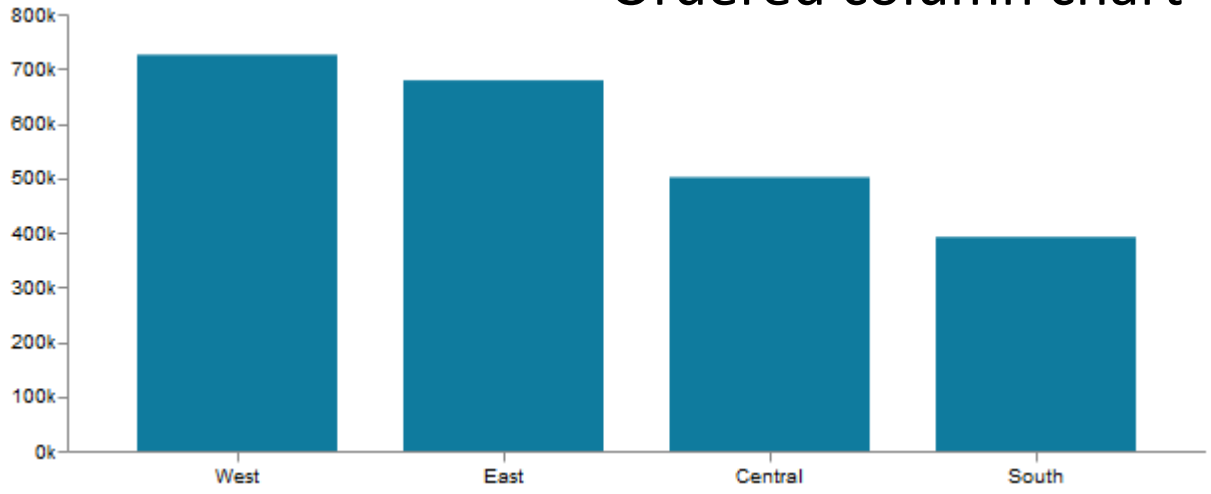
Bars

Bars can be plotted **horizontally** or **vertically** (columns)

- Ordered bar chart

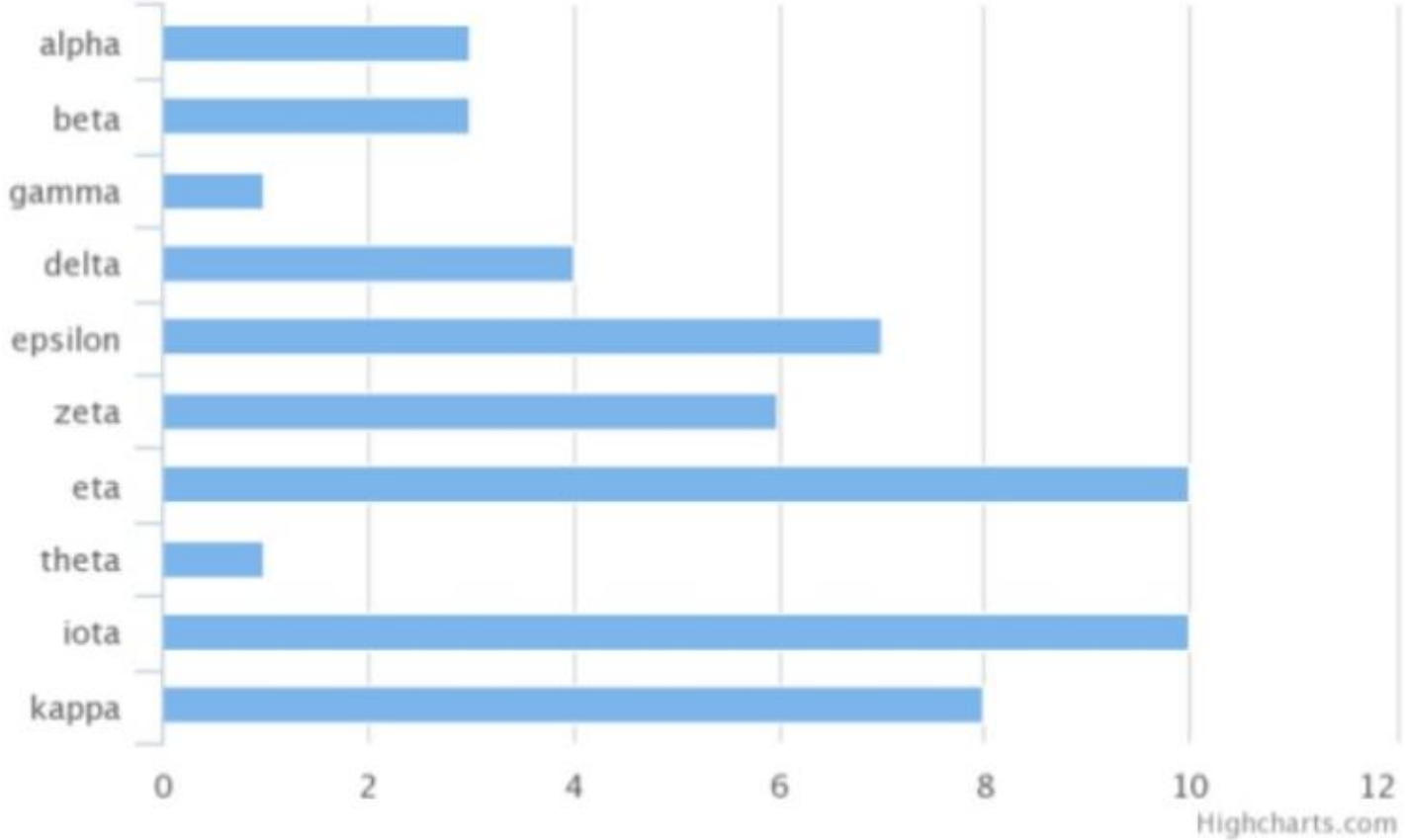


- Ordered column chart



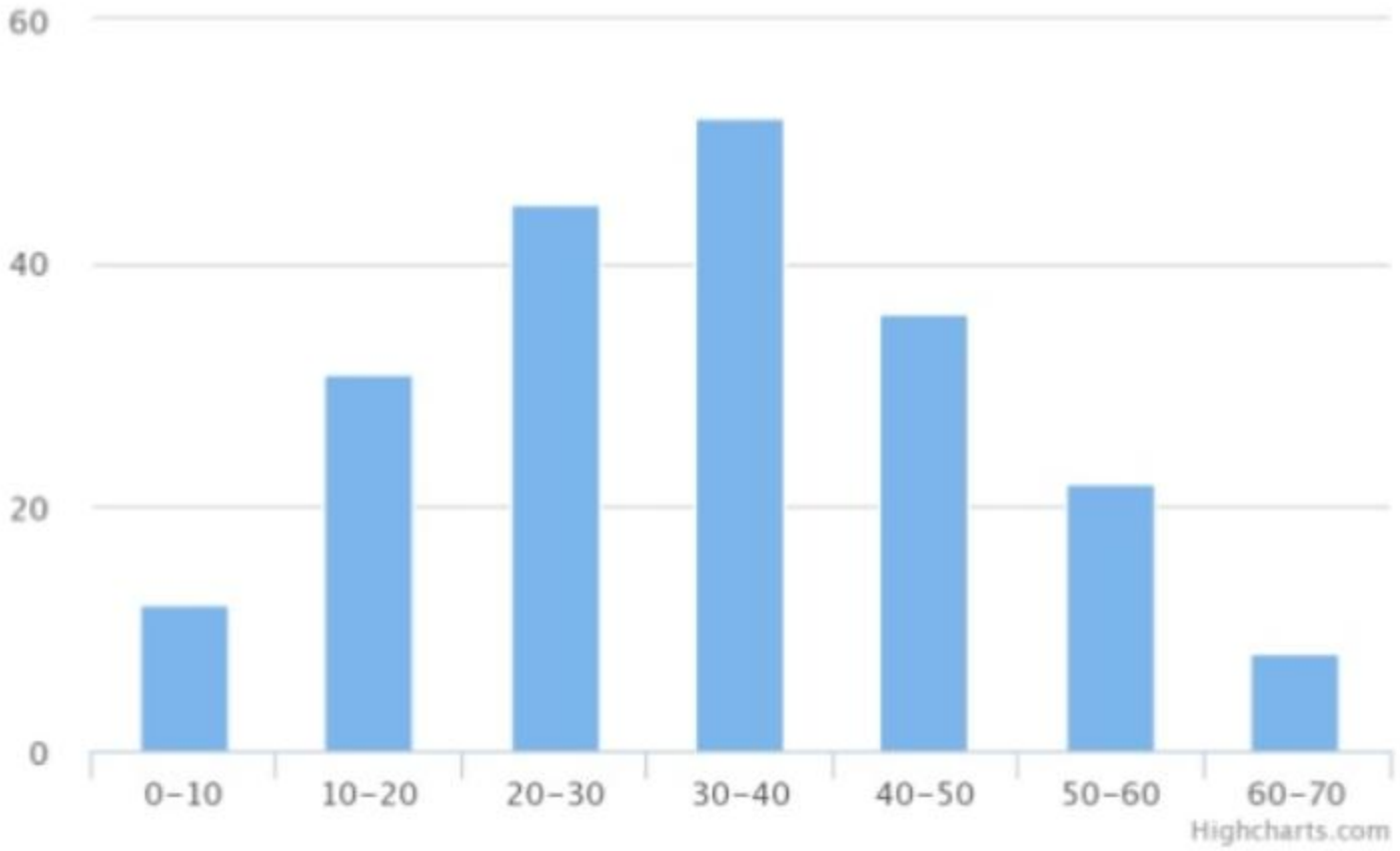
Bars

type	mark	channel	data represented
Bar chart	line	size (length)	1 categorical, 1 quantitative



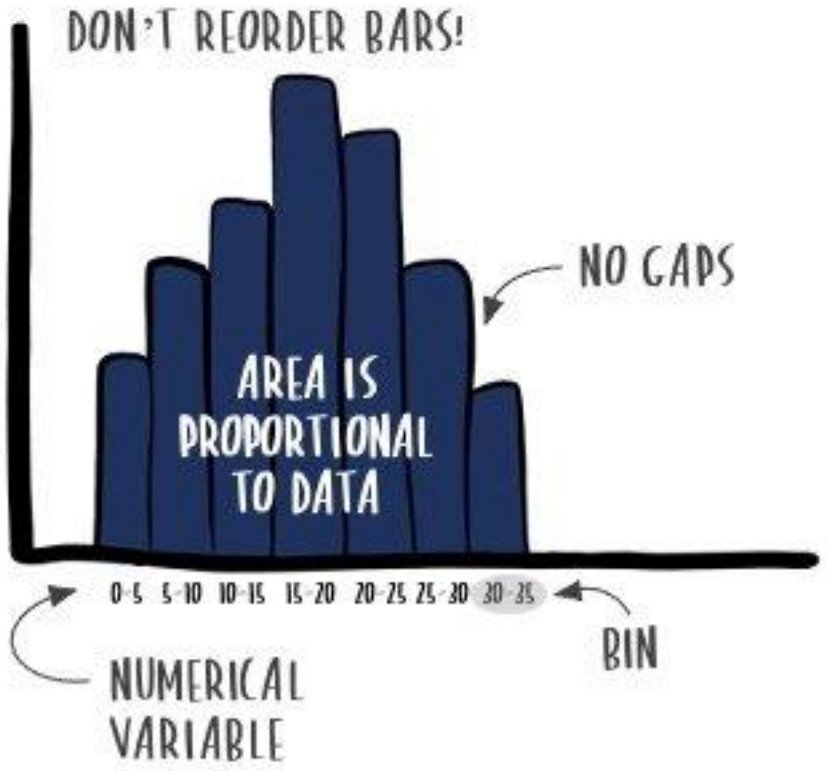
Bars

type	mark	channel	data represented
Histogram	line	size (length)	1 ordinal/quantitative, 1 quantitative (count)



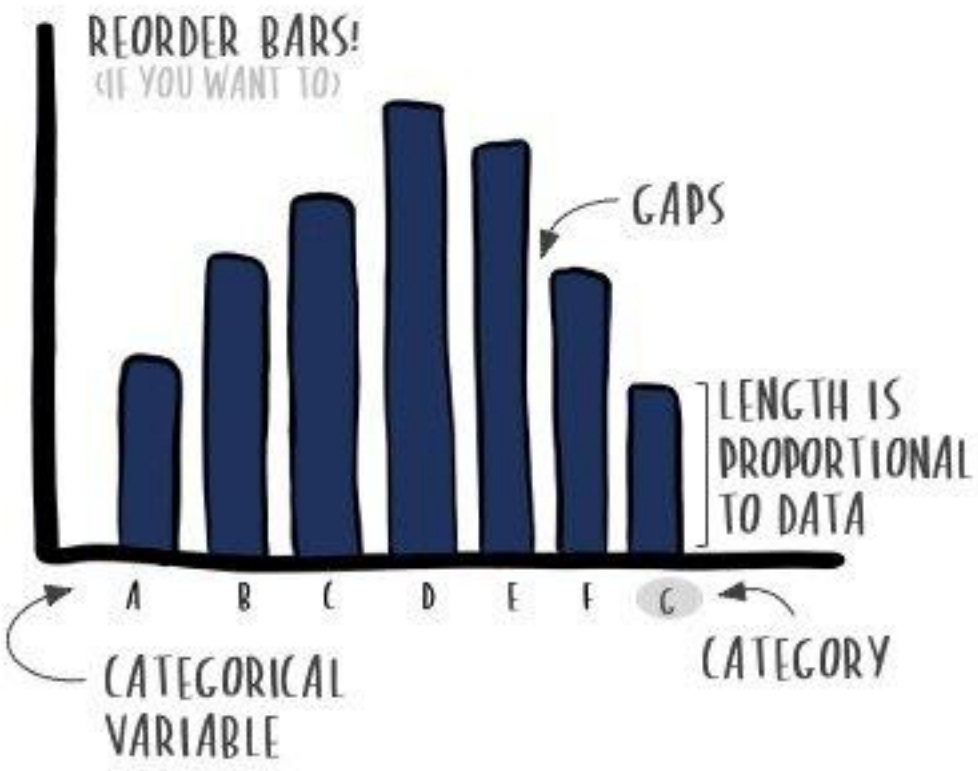
Bars

This is a **histogram**...



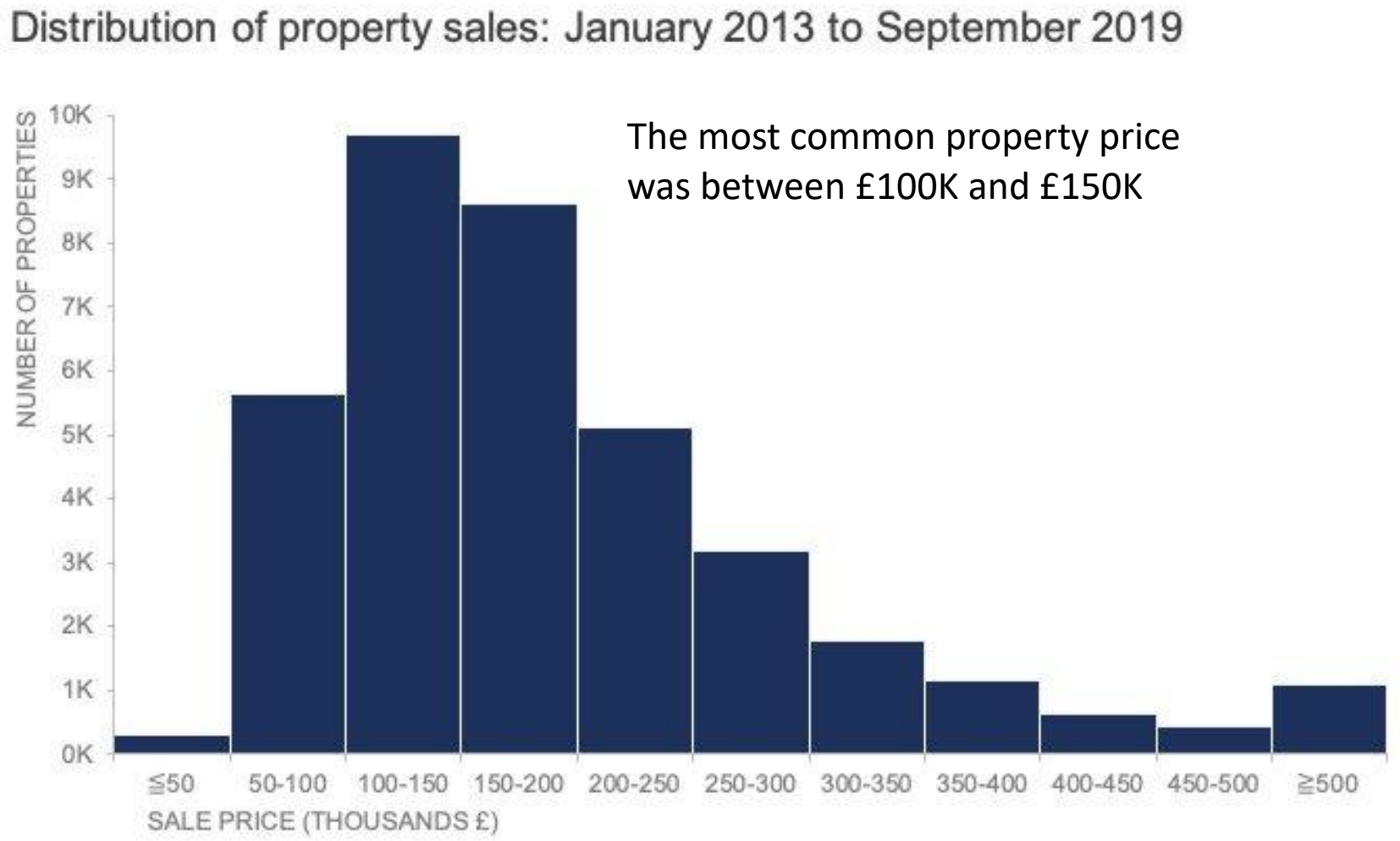
data is grouped into continuous number ranges and each range corresponds to a vertical bar

This is a **bar chart**...



categorical data using rectangular bars, where the length of each bar is proportional to the value they represent

Histogram

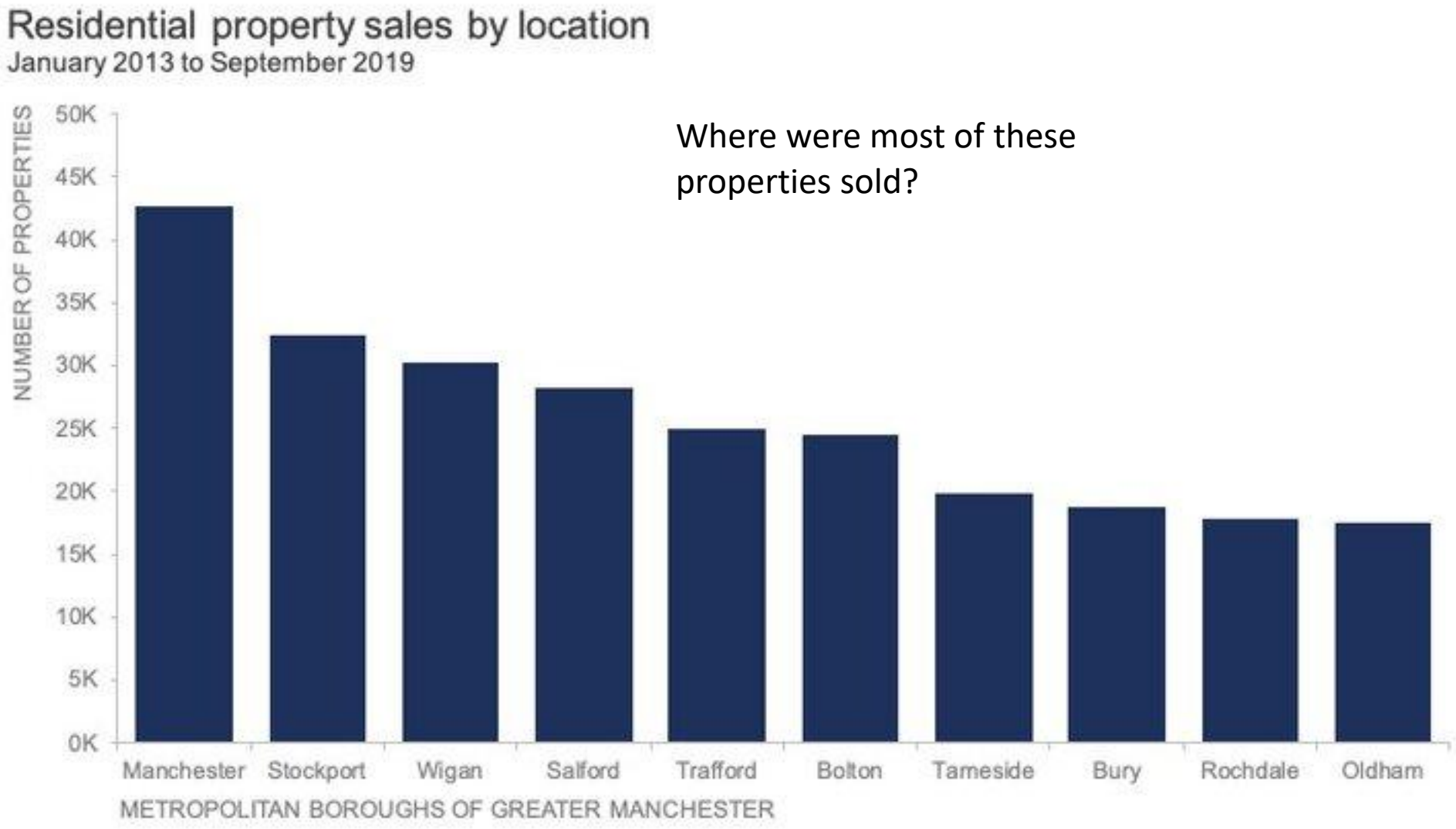


The most common property price was between £100K and £150K

It's rare for a property to sell below £50K

The histogram lets us understand more about the shape or skewness of the numerical variable, sale price. Use when you want to understand the underlying distribution of a numerical data series.

Bar chart

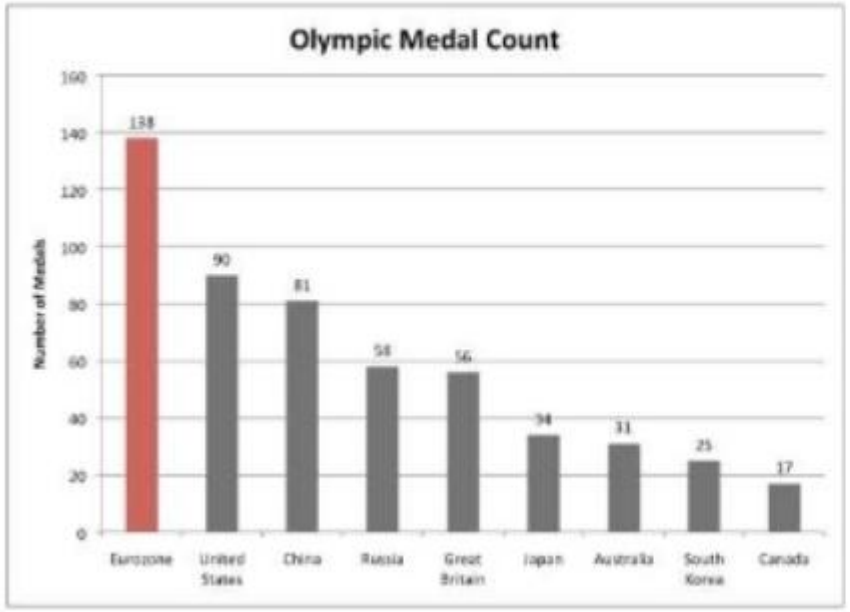


Where were most of these properties sold?

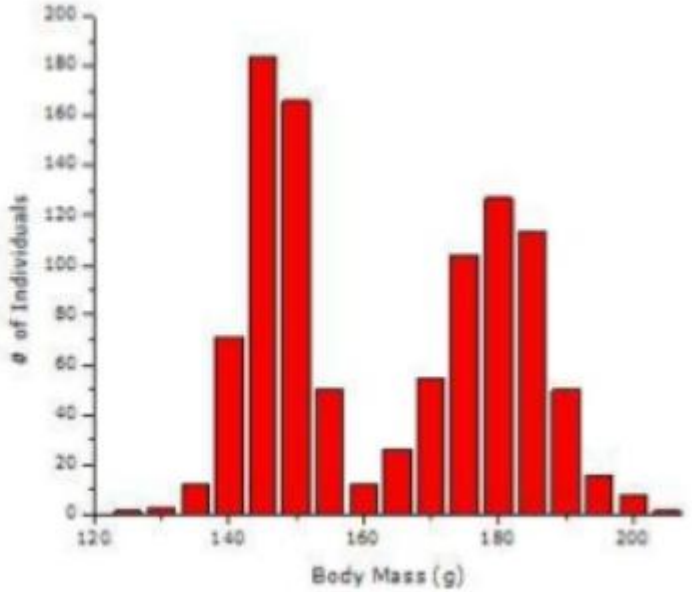
A bar chart works well to make comparisons across categories, not numeric data.

Bar chart applications

COMPARE CATEGORIES

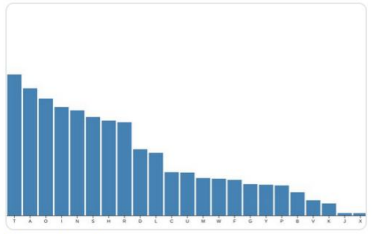


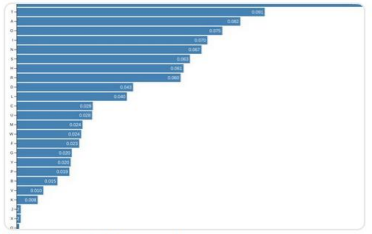
DISTRIBUTION

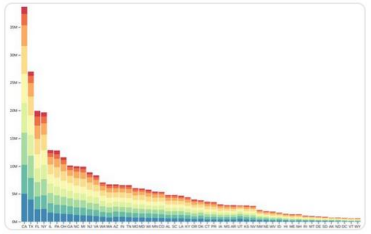


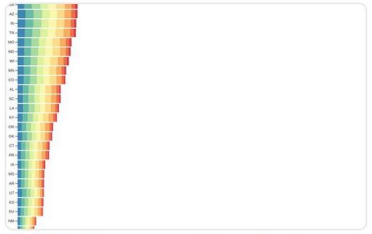
Bars

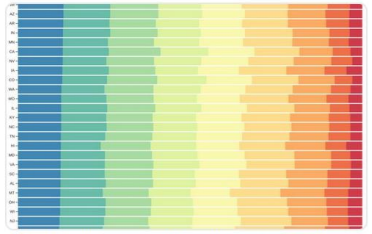
- Scales & Axes
- Bar charts, marimekko chart, timelines, calendar views

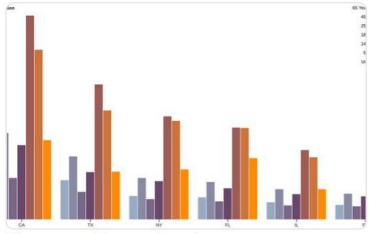

Bar chart

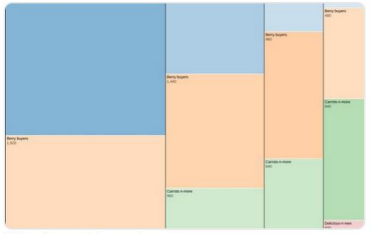

Horizontal bar chart

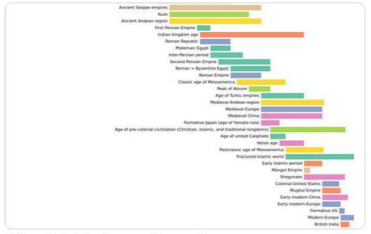

Stacked bar chart

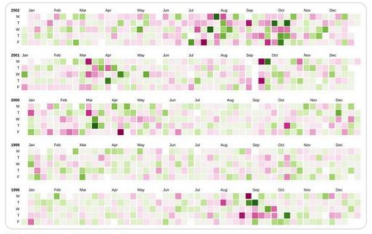

Stacked horizontal bar chart

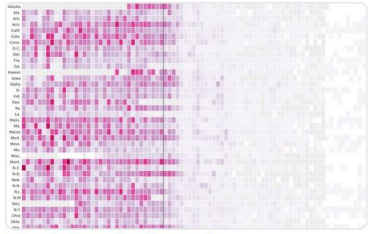

Stacked normalized horizontal …

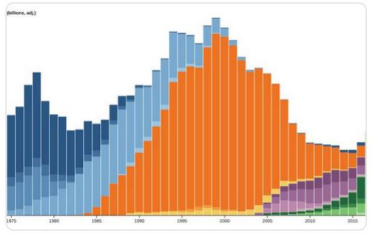

Grouped bar chart


Marimekko chart

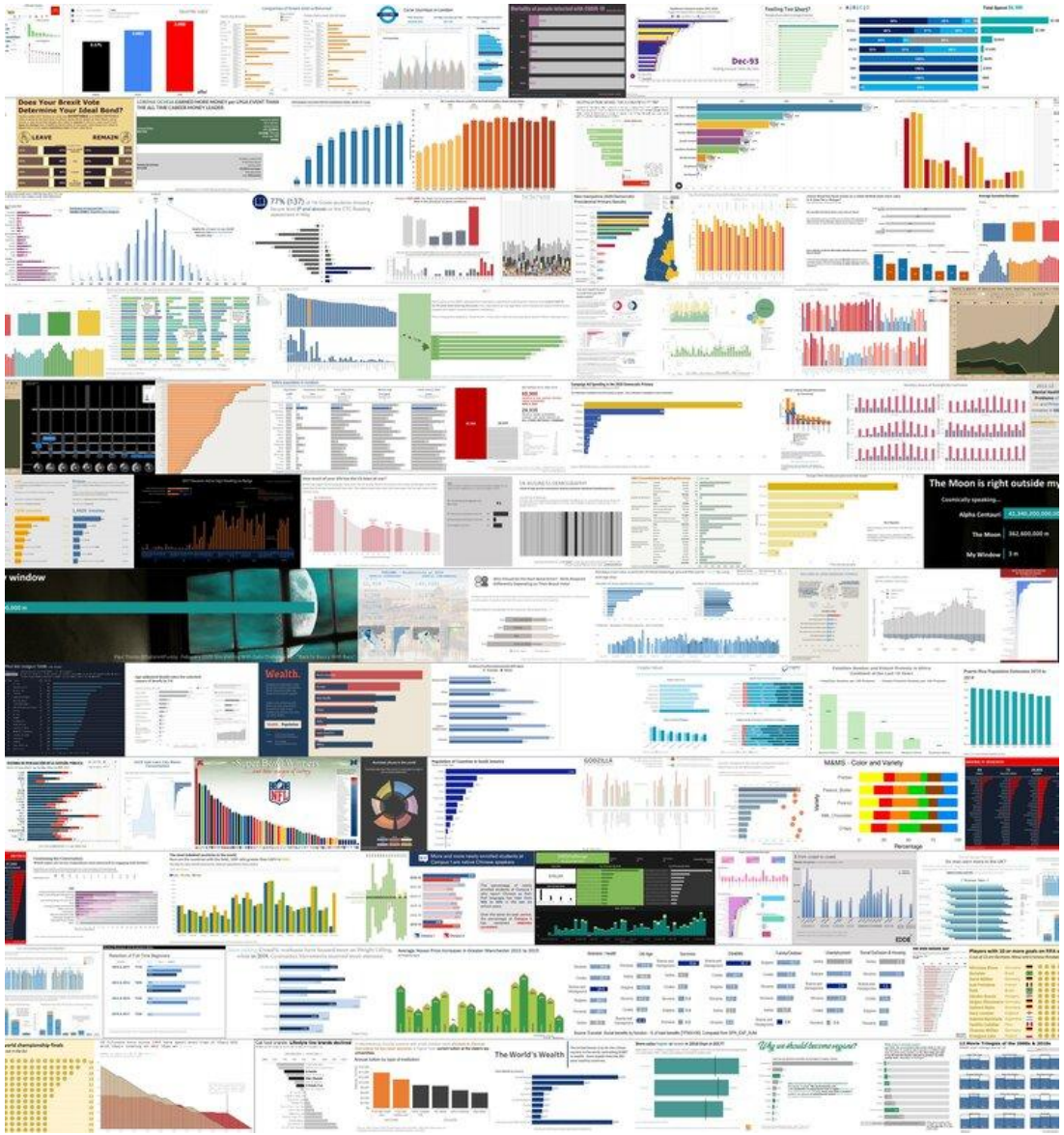

World history timeline


Calendar view

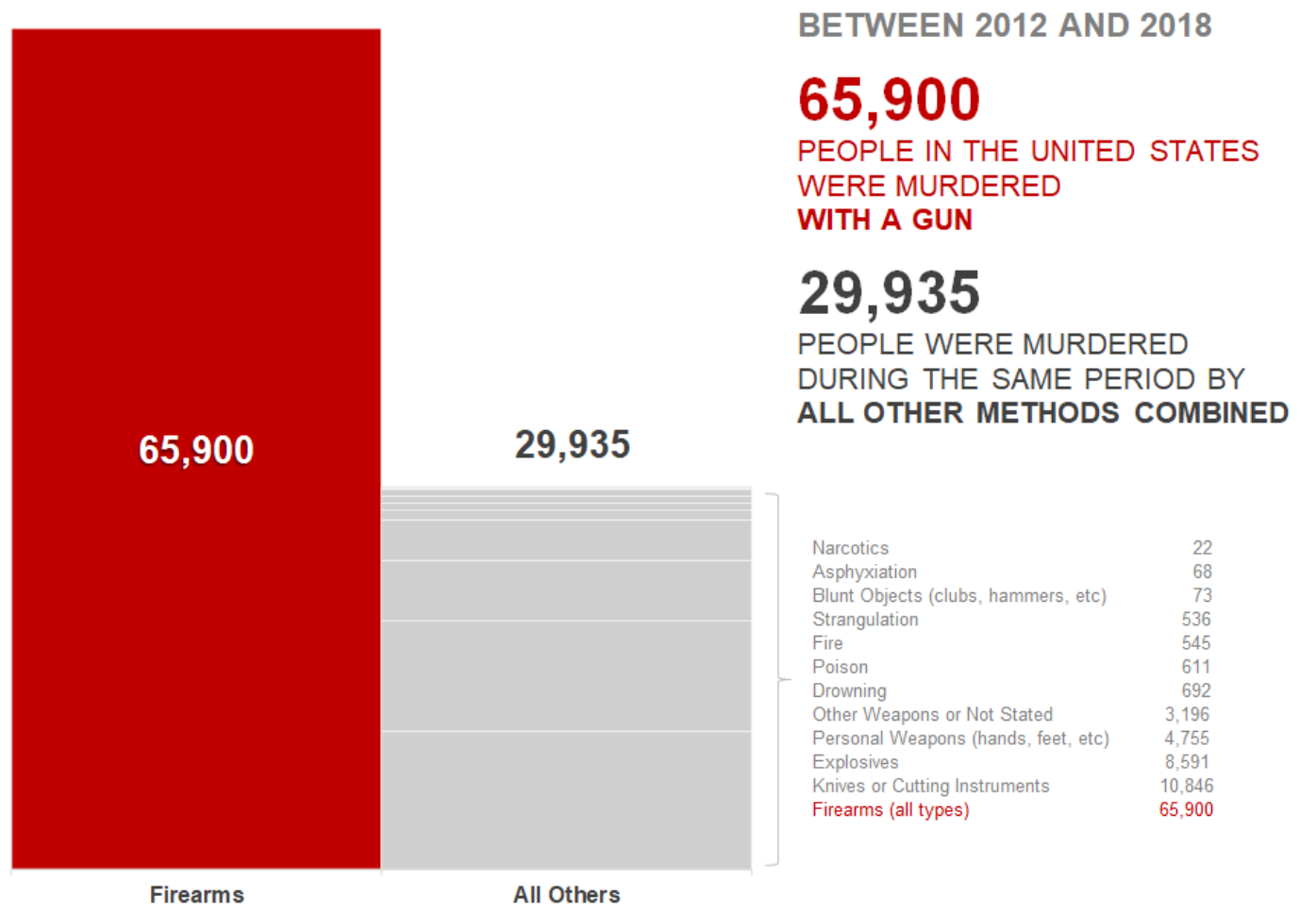

The impact of vaccines


Revenue by music format, 197…

Bars

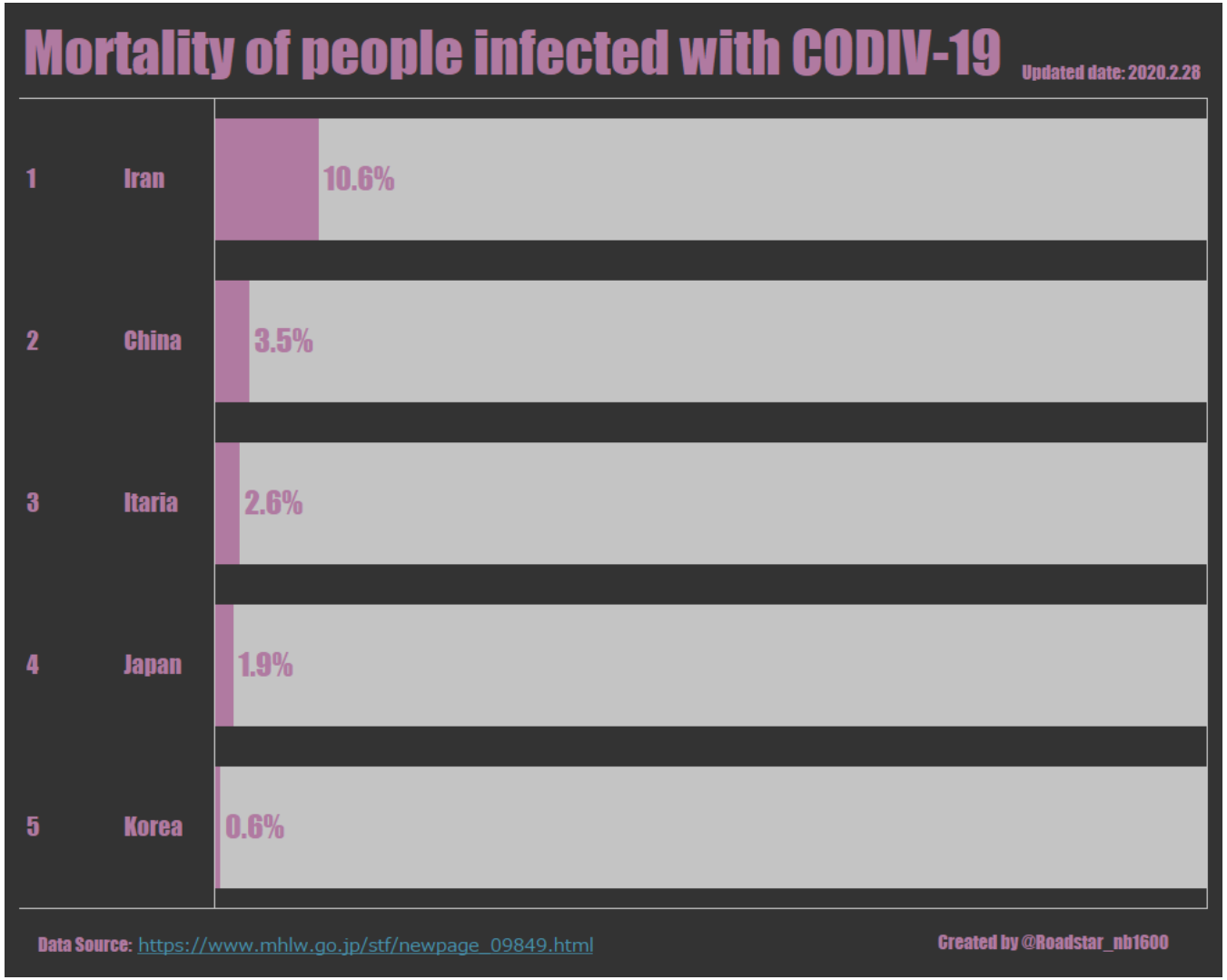

Does Your Brexit Vote Determine Your Ideal Bond?
LEAVE
REMAIN
Dec-93
The Moon is right outside my window
Wealth.
The World's Wealth

Bars

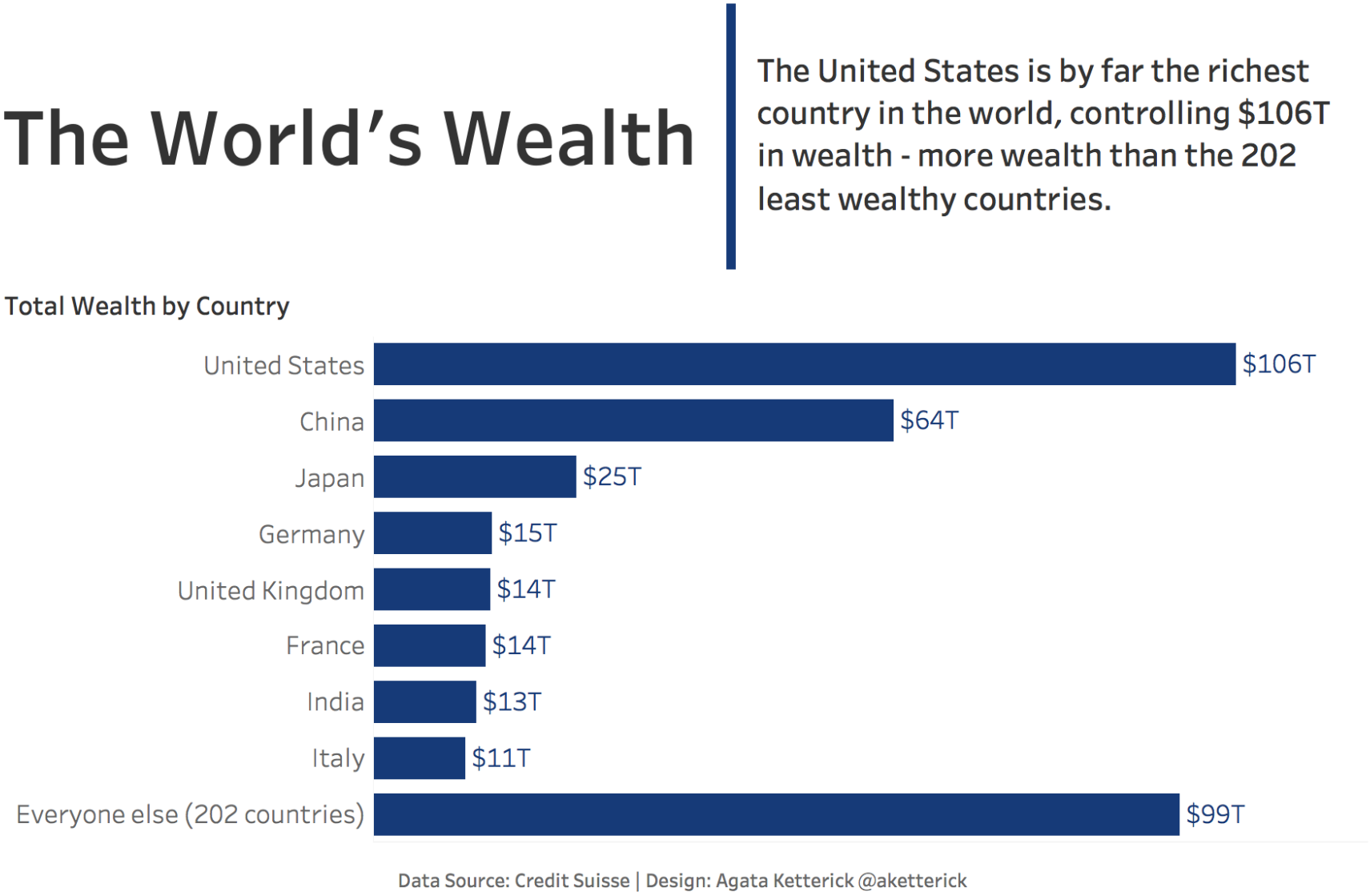


Data Source: FBI • Crime in the U.S. • 2018 • Crime in the U.S. 2018 • Tables • Expanded Homicide Data Table 8
https://ucr.fbi.gov/crime-in-the-u.s/2018/crime-in-the-u.s.-2018/tables/expanded-homicide-data-table-8.xls
Design: Jamie Briggs (@indented)

Bars - horizontal bar chart

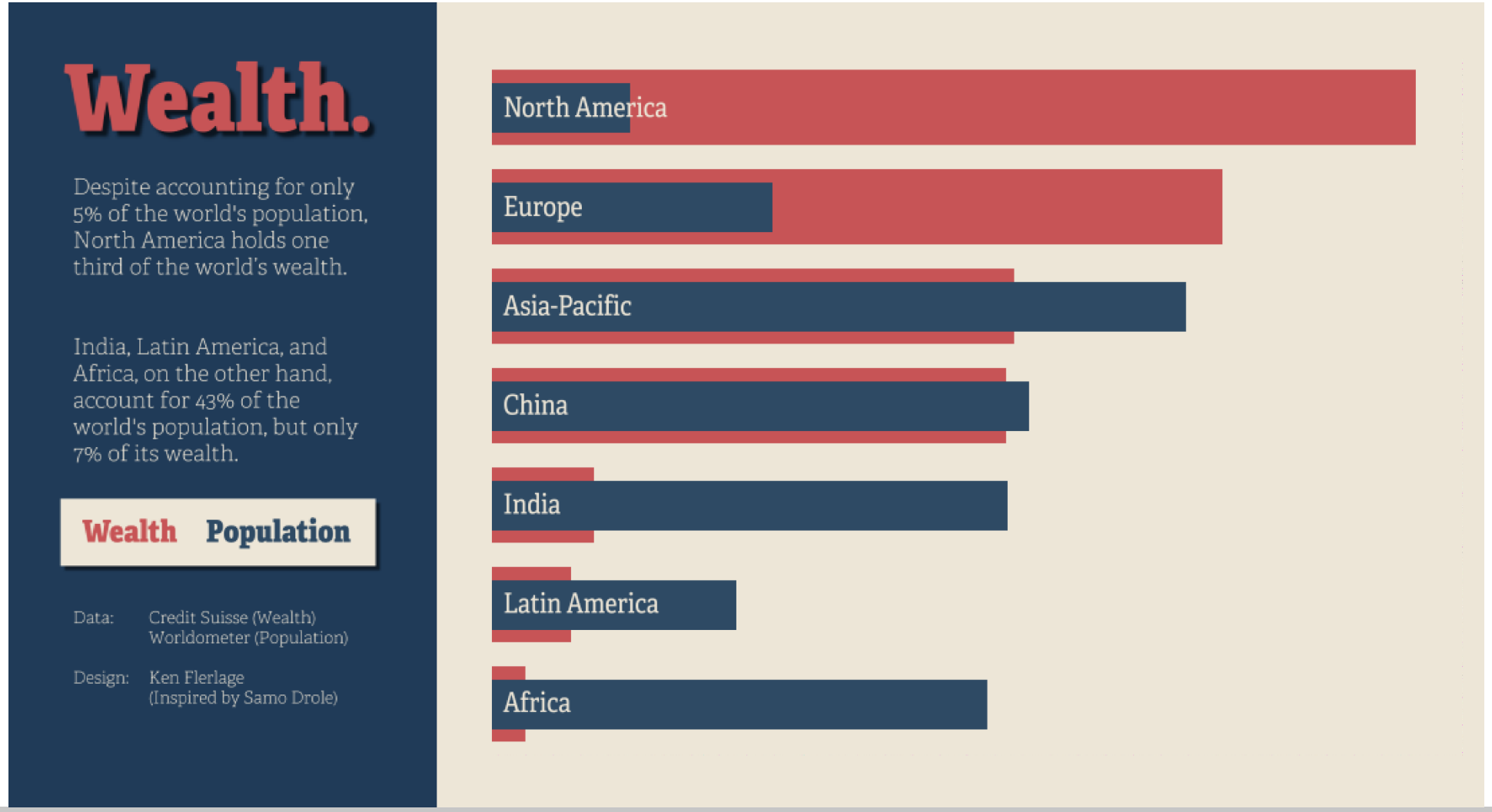


Bars - horizontal bar chart



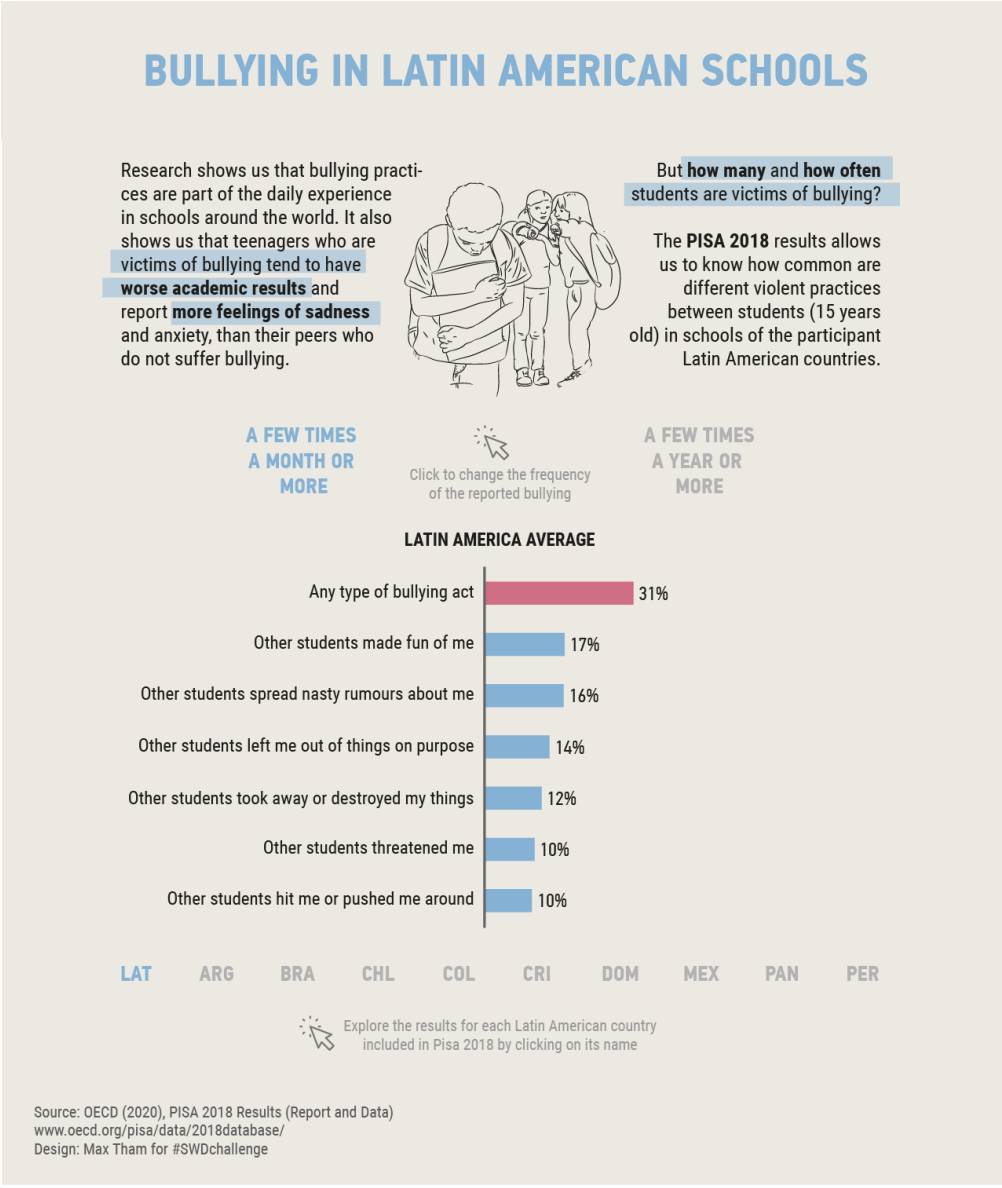
Bars - horizontal bar chart

(same dataset as previous)

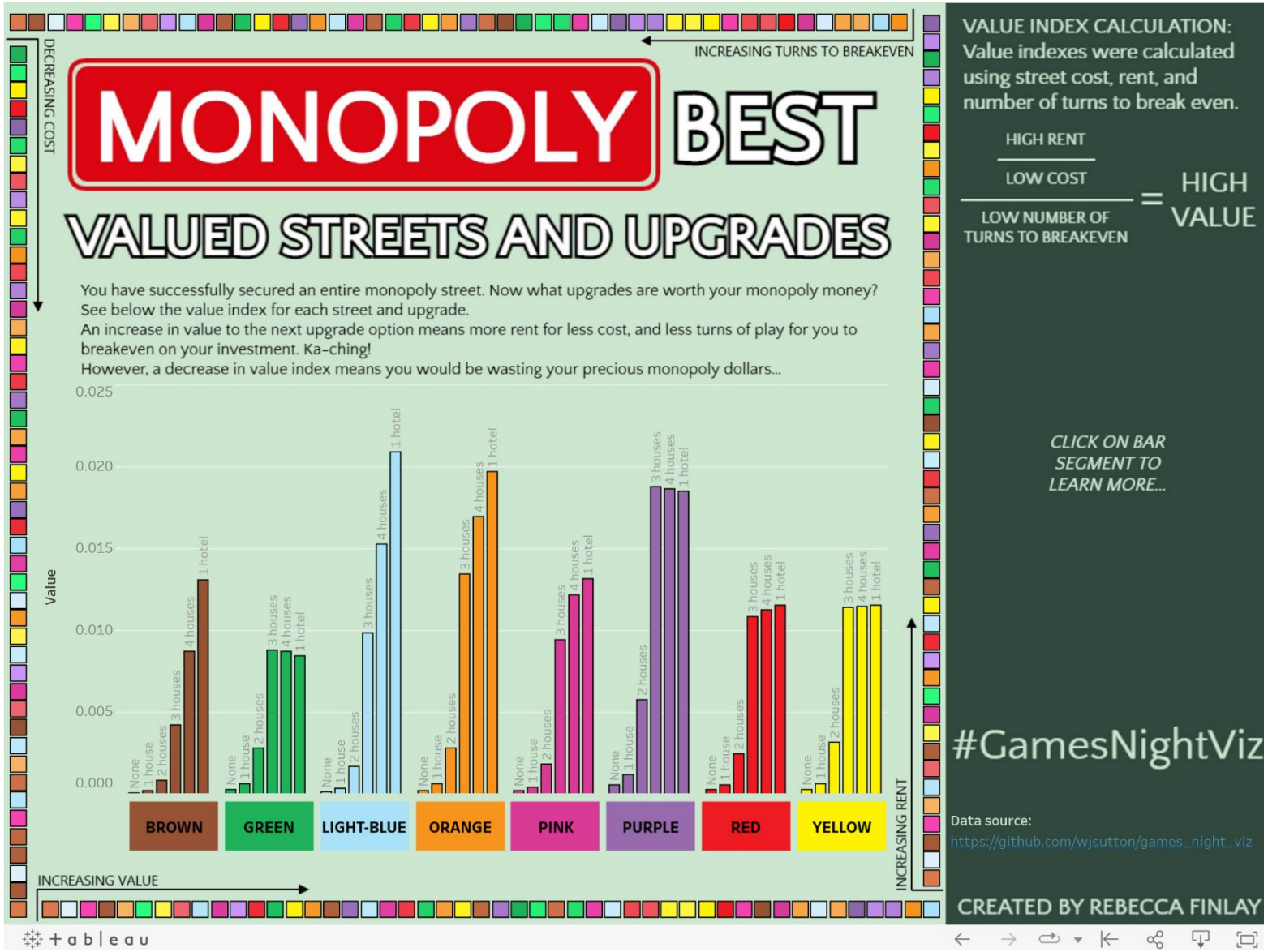


Bars - horizontal bar chart

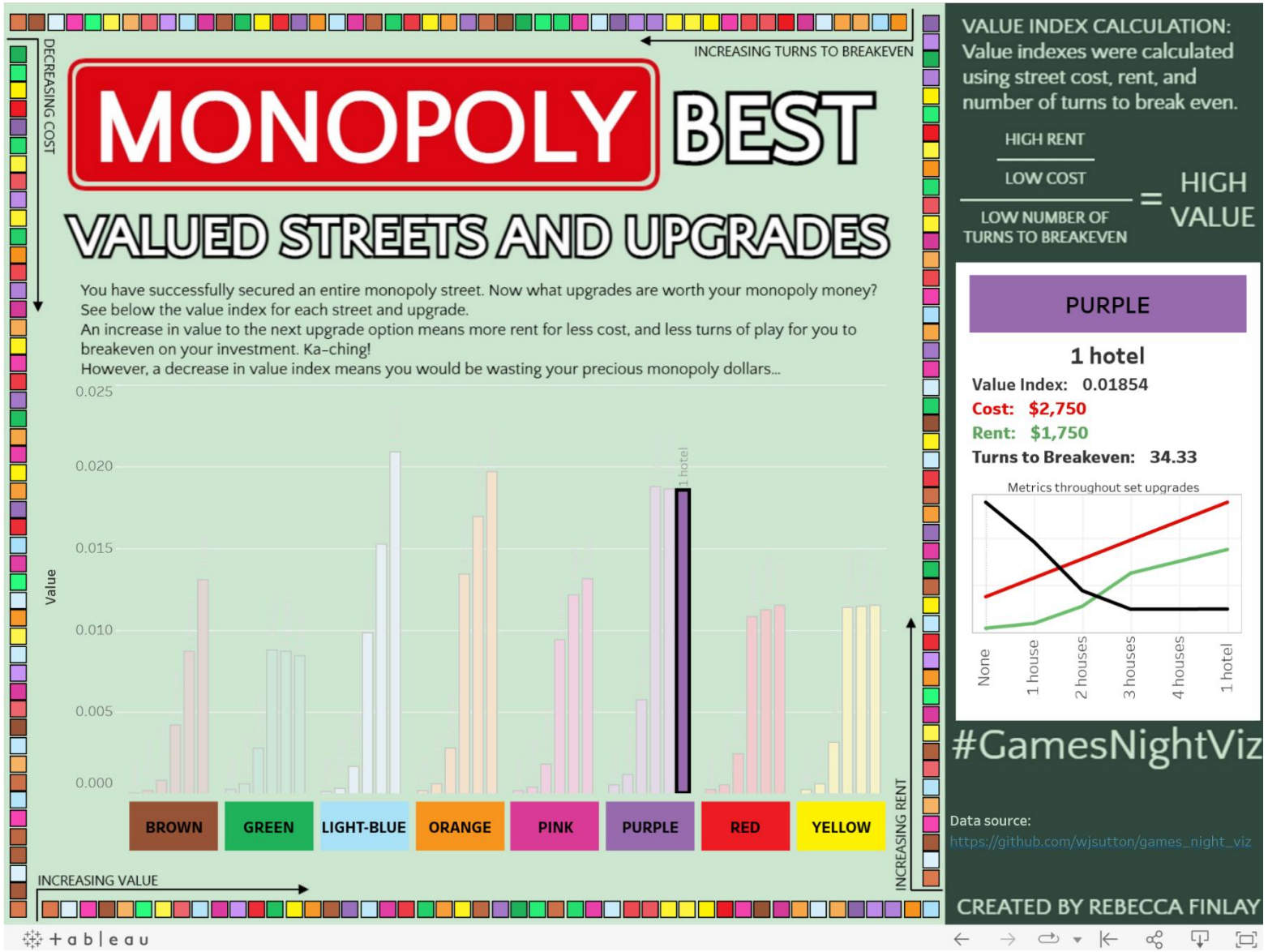
- Ordered bar



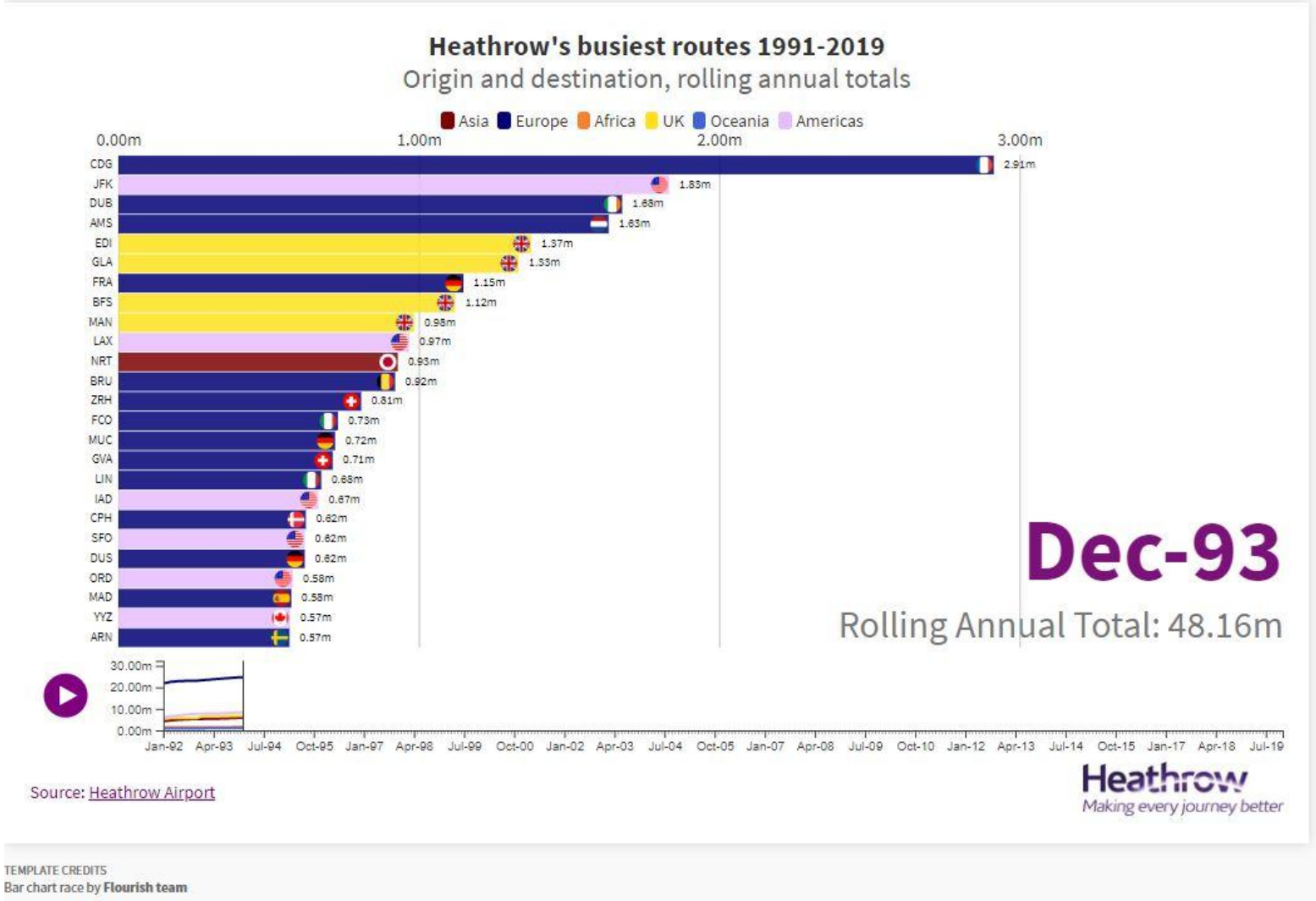
Bars - bar chart



Bars - bar chart

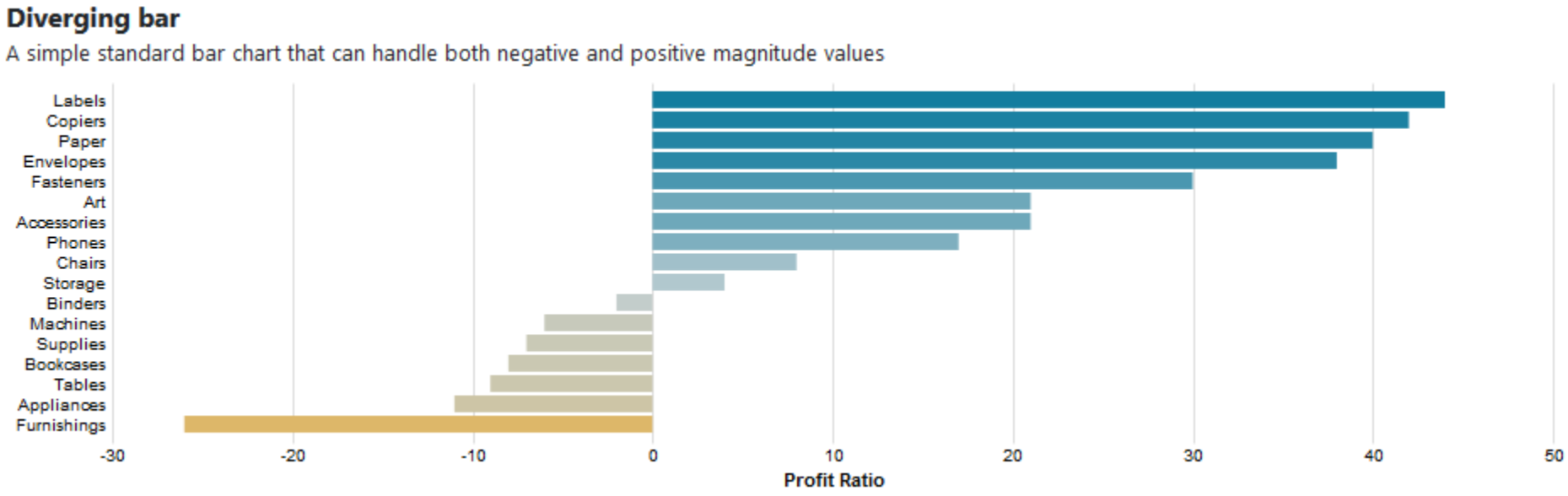


Bars - racing bars (= with animation)

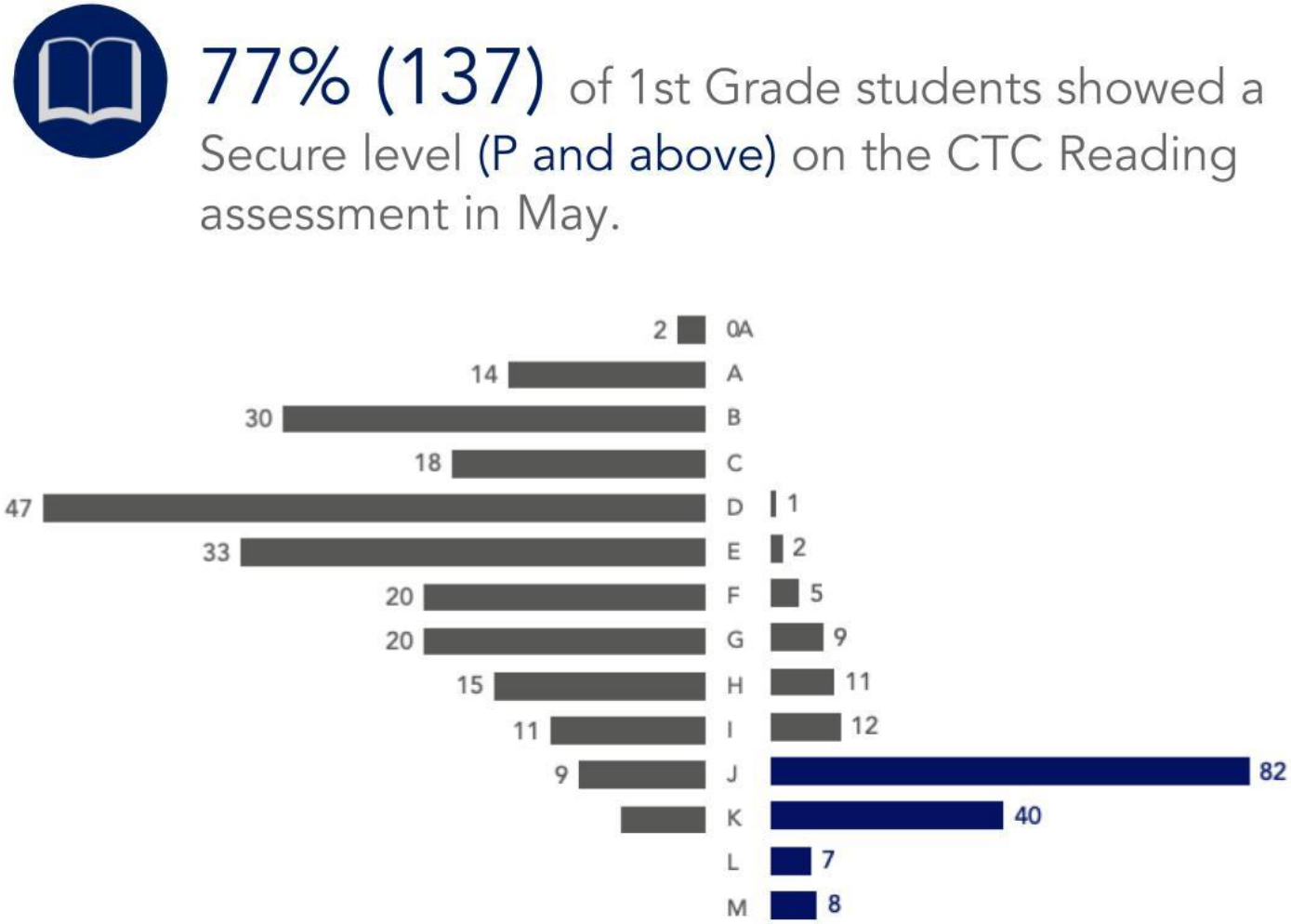


Bars - diverging bars

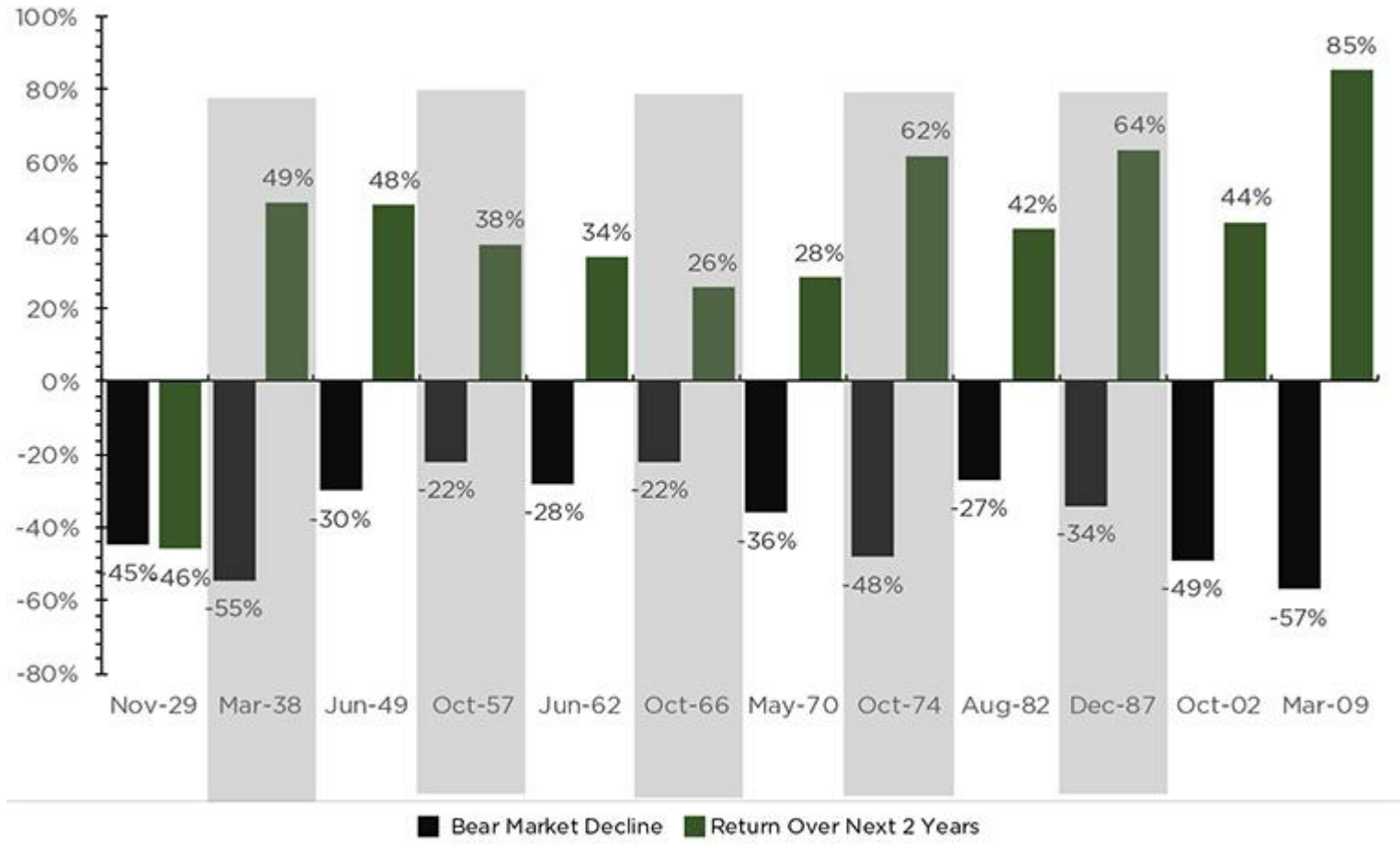
- Emphasise variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a long-term average.
- Can also be used to show sentiment (positive/neutral/negative)



Bars - diverging bars



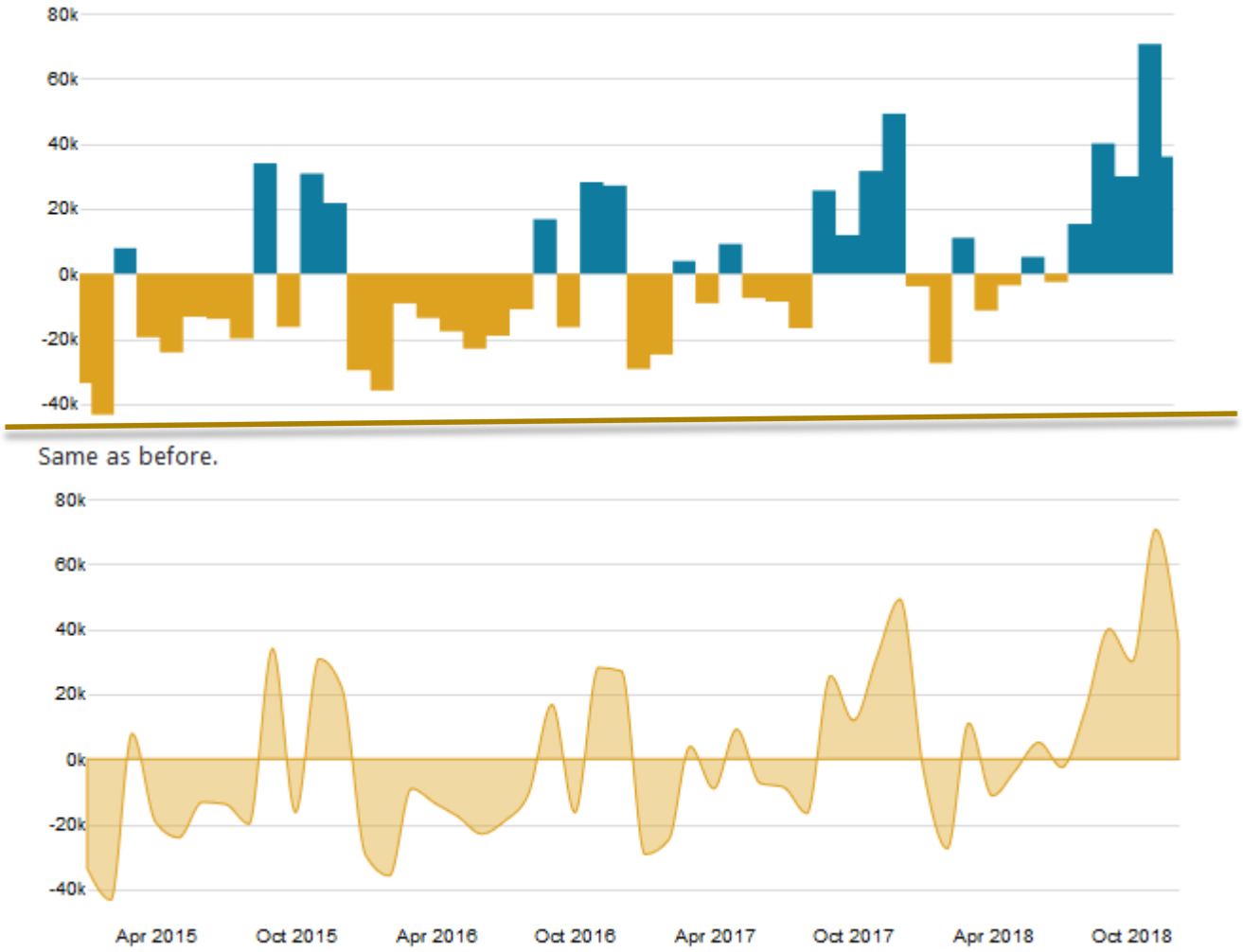
Bars - diverging bars



Edward Jones newsletter 2020

Bars - diverging bars

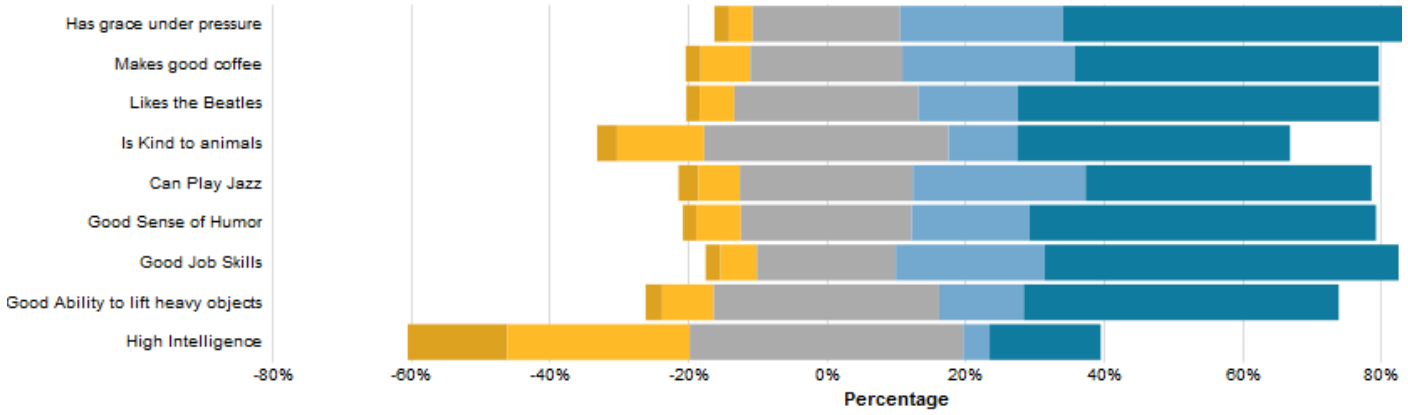
- Surplus/deficit filled line / area



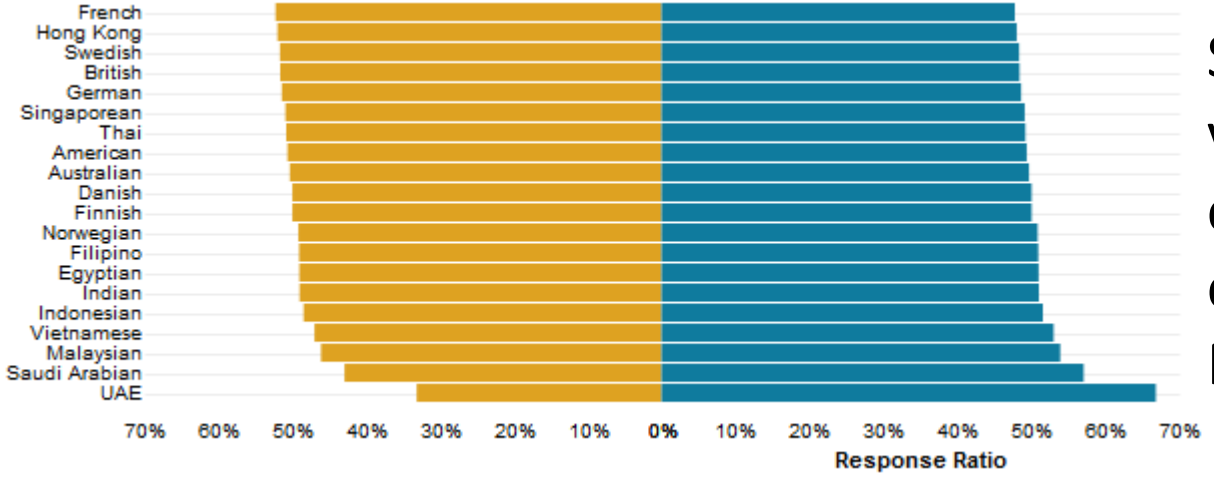
The shaded area of these charts allows a balance to be shown; either against a baseline or between two series

Bars - diverging bars

- Diverging stacked bar / Spine



Perfect for presenting survey results which involve sentiment (e.g. disagree, neutral, agreed)



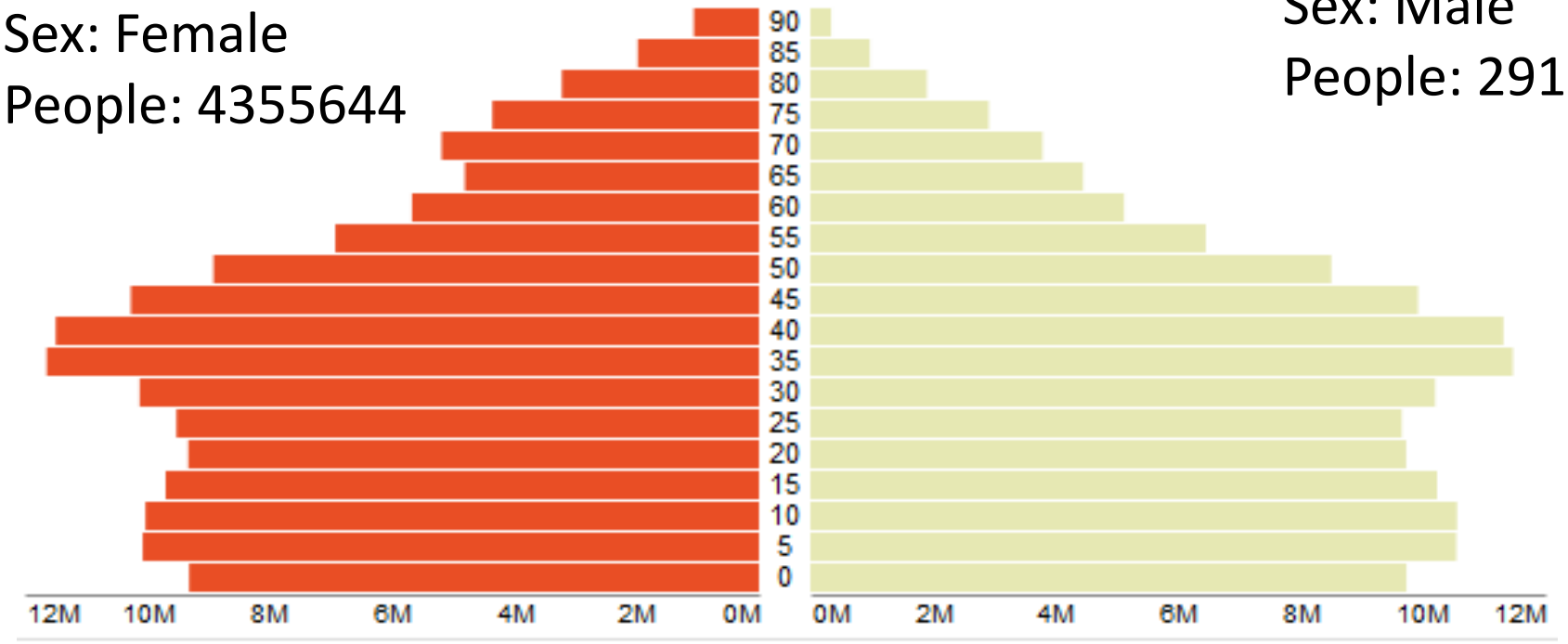
Splits a single value into 2 contrasting components (e.g. Male / Female)

Bars - diverging bars

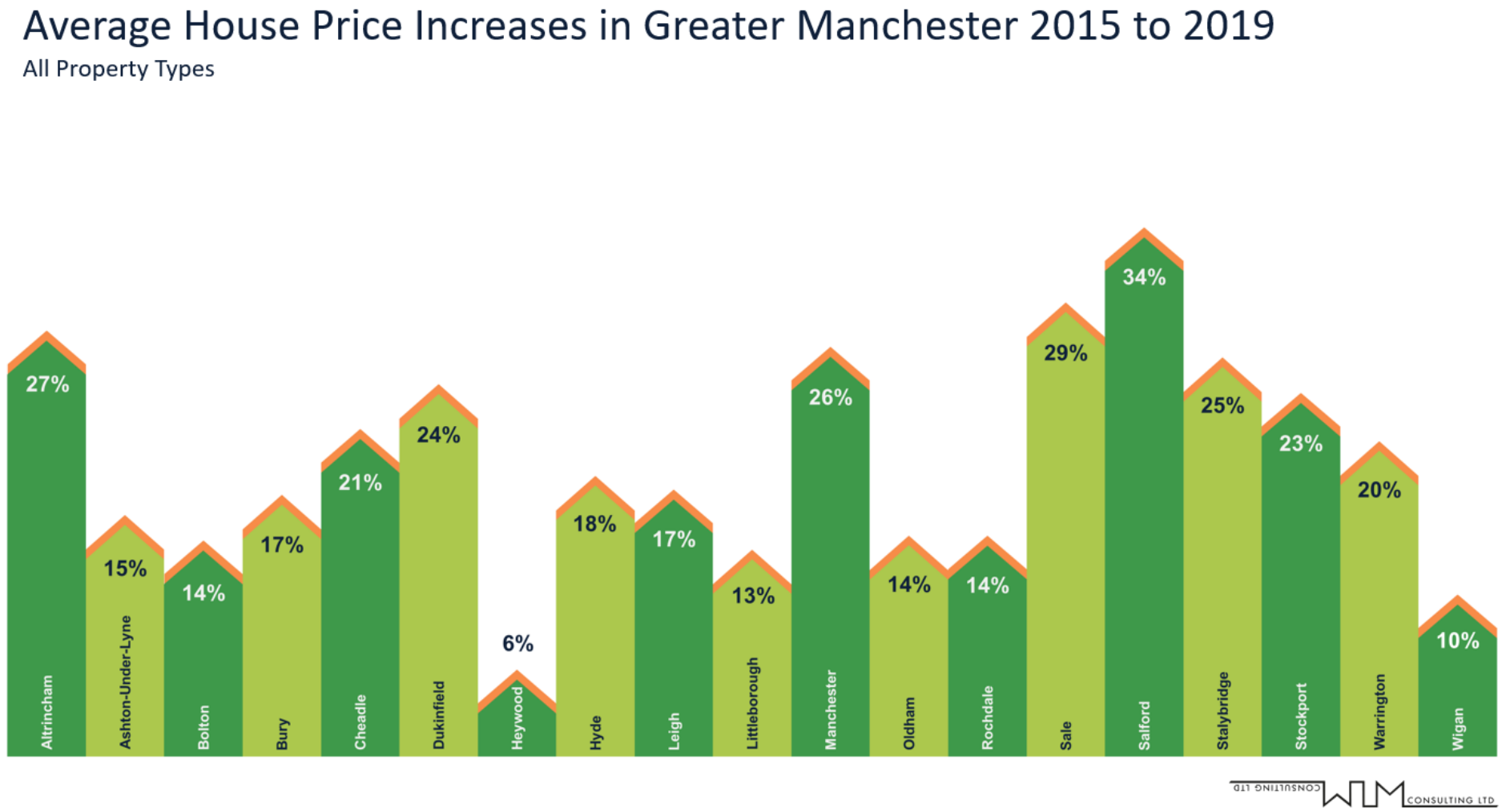
- A standard way for showing the age and sex breakdown of a population distribution; effectively, back to back histograms

Age: 75
Sex: Female
People: 4355644

Age: 75
Sex: Male
People: 2912655

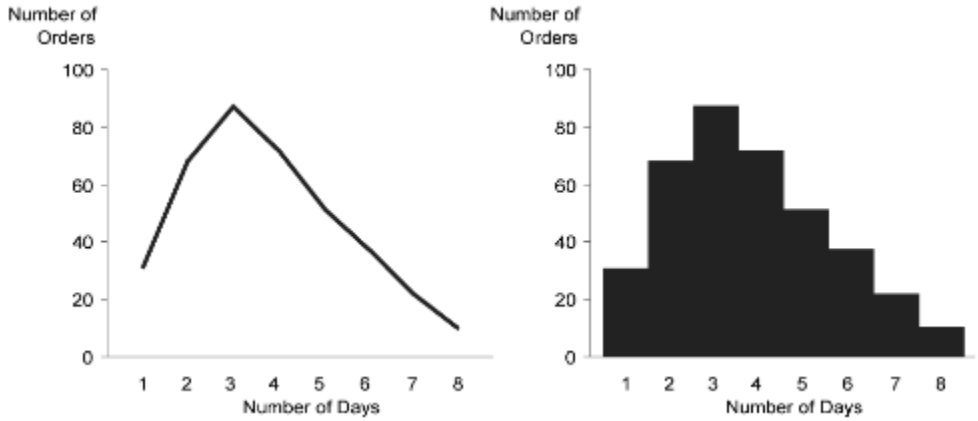


Bars – “themed” bars



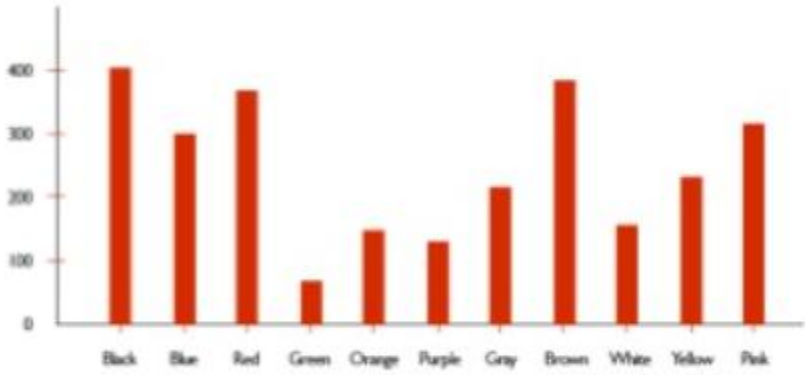
Bar charts instead of line charts

- Bars may be used in place of a line when you wish to emphasize the individuality of the values and to compare one to another, instead of examining the overall shape of the values



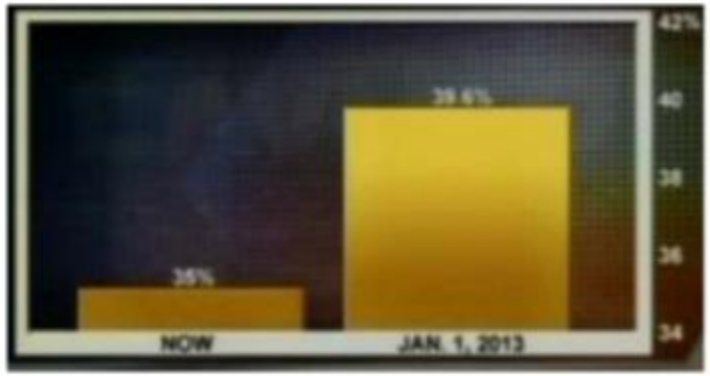
e.g. a distribution relationship involving the number of days that it took to ship the orders that were received in a given month.

Bar chart dangers



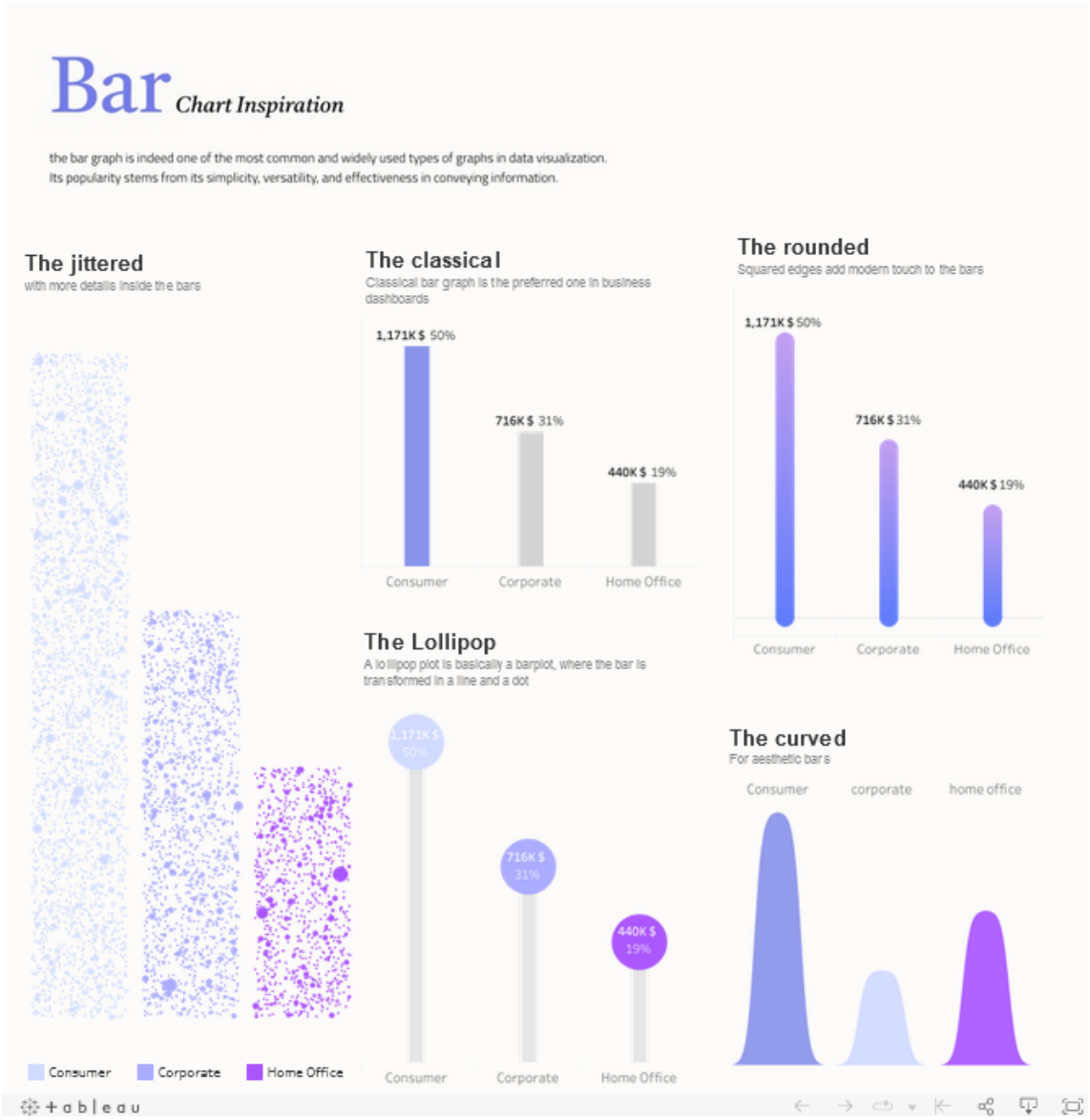
TOO MANY CATEGORIES

POORLY SORTED CATEGORIES



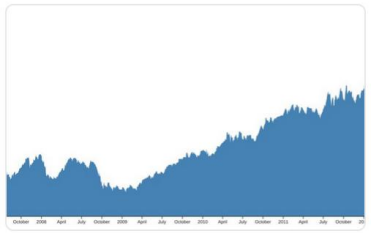
ZERO AXIS

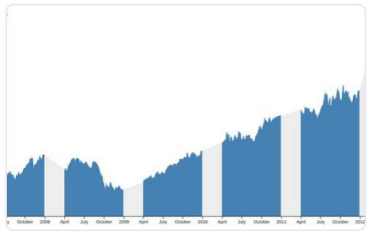
Bar chart styles

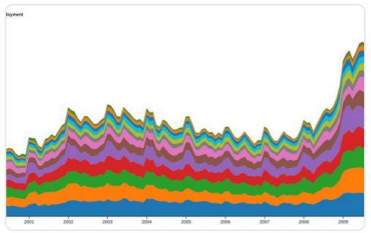


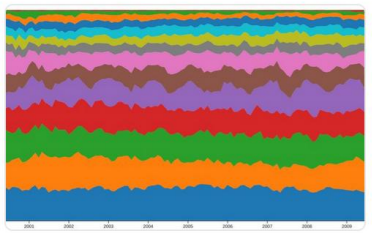
Areas

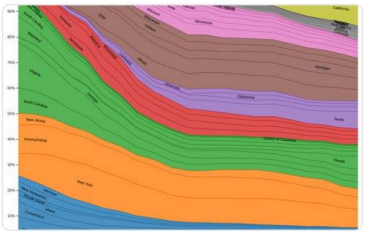
- Difference charts
- Streamgraphs
- Ridgeline plots (to compare many simultaneous time series)

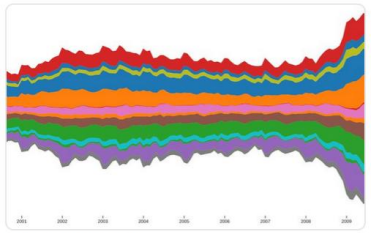

Area chart

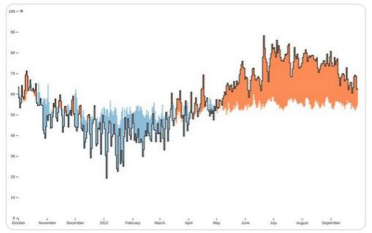

Area with missing data

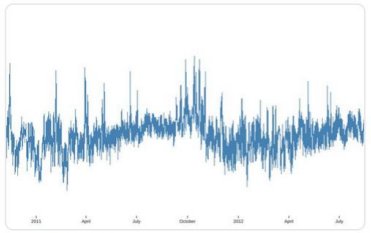

Stacked area chart

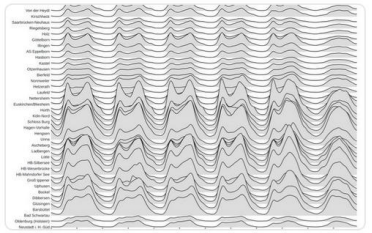

Normalized stacked area chart

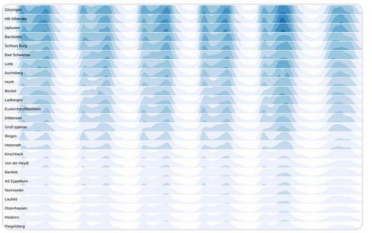

U.S. population by state, 1790–…

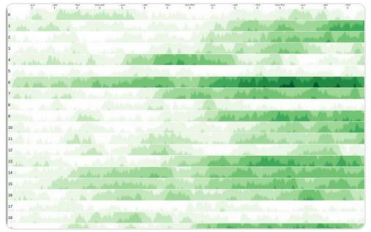

Streamgraph


Difference chart


Band chart

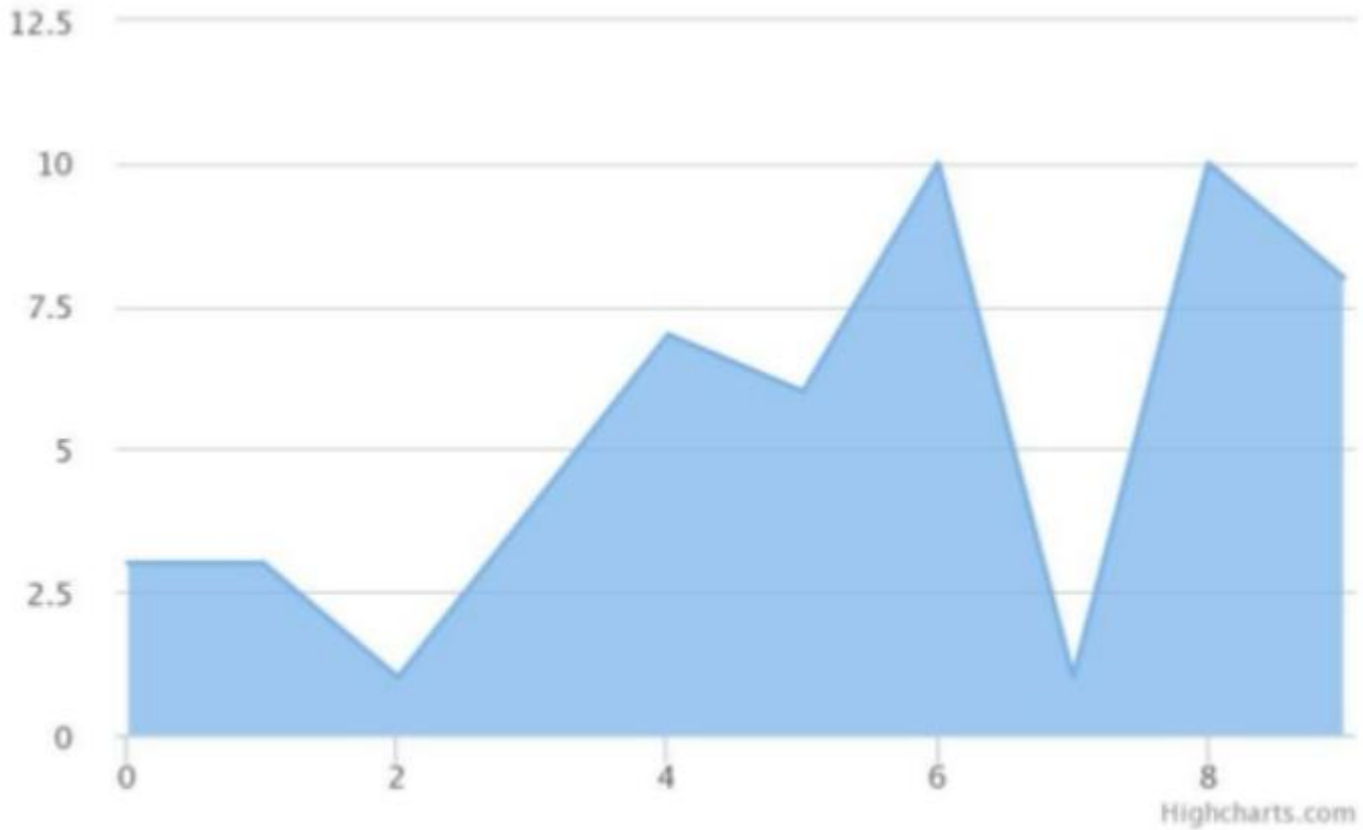

Ridgeline (joy) plot


Horizon chart

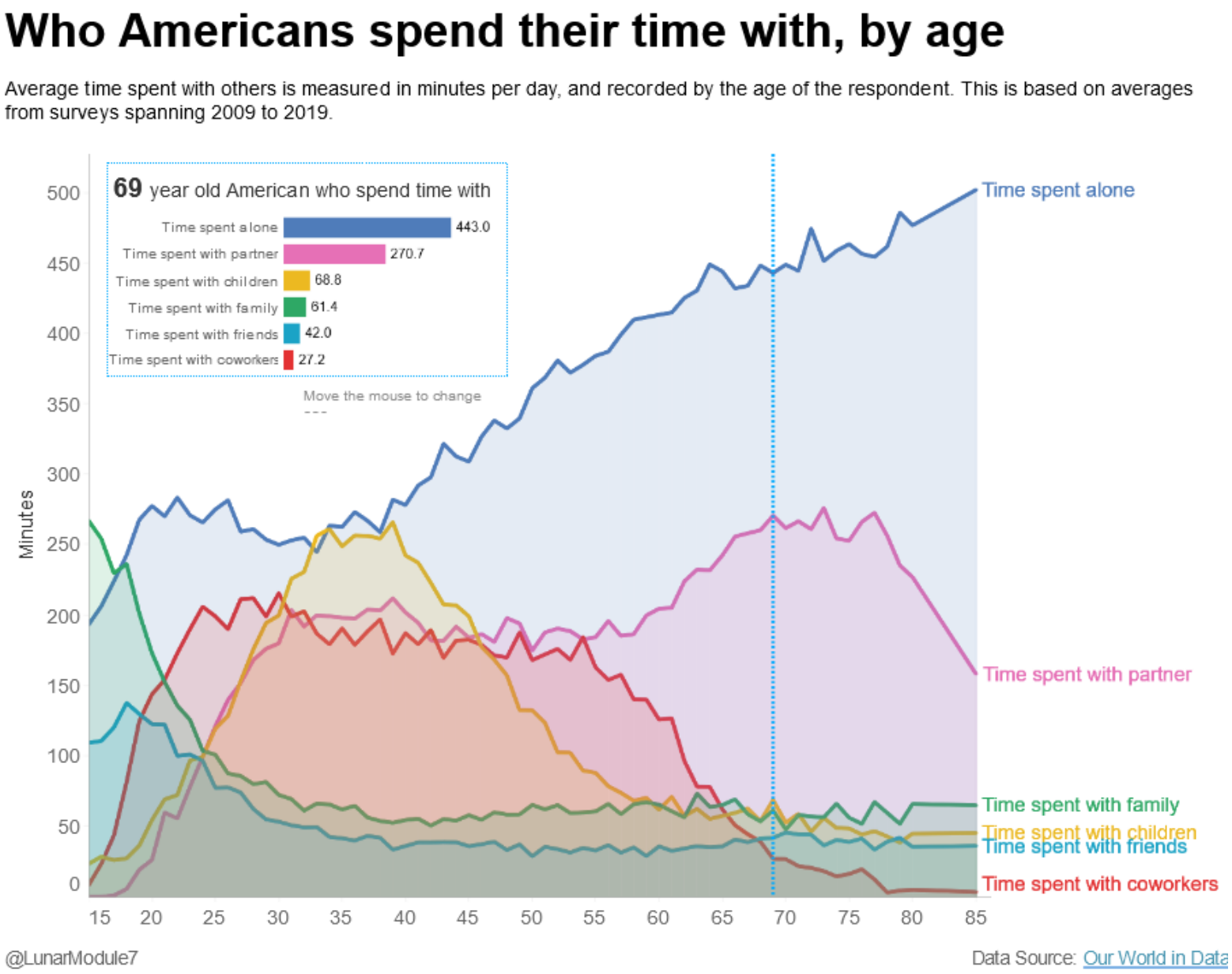

Realtime horizon chart

Areas

type	mark	channel	data represented
Area chart	area	size (length)	2 quantitative

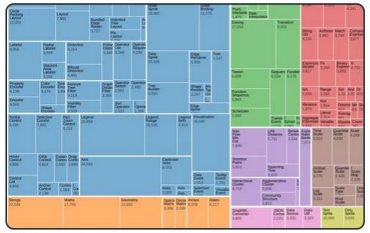


Areas

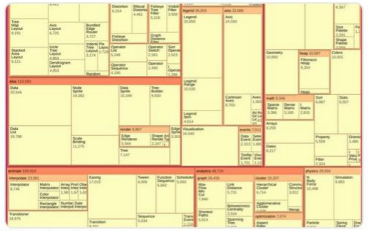

Who Americans spend their time with, by age
Average time spent with others is measured in minutes per day, and recorded by the age of the respondent. This is based on averages from surveys spanning 2009 to 2019.
69 year old American who spend time with
Time spent alone 443.0
Time spent with partner 270.7
Time spent with children 68.8
Time spent with family 61.4
Time spent with friends 42.0
Time spent with coworkers 27.2
Move the mouse to change
Minutes
0
50
100
150
200
250
300
350
400
450
500
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
Time spent alone
Time spent with partner
Time spent with family
Time spent with children
Time spent with friends
Time spent with coworkers
@LunarModule7
Data Source: Our World in Data

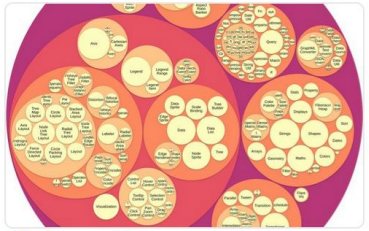
Hierarchies

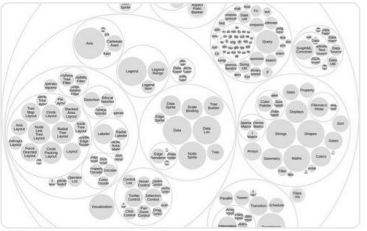
- Treemaps
- tidy trees
- packed circles

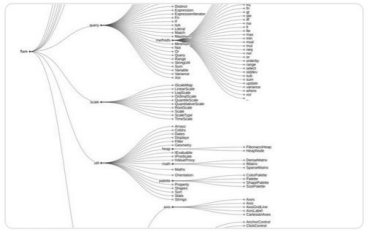

Treemap

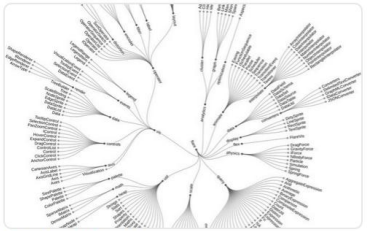
Cascaded treemap

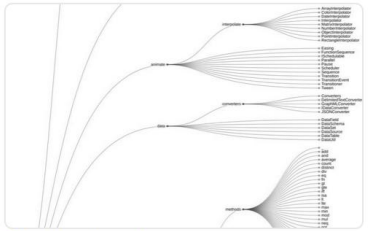

Nested treemap

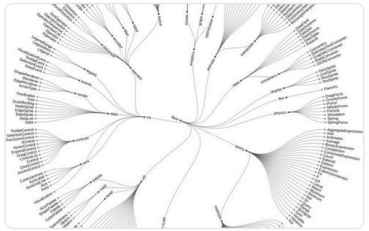

Circle packing

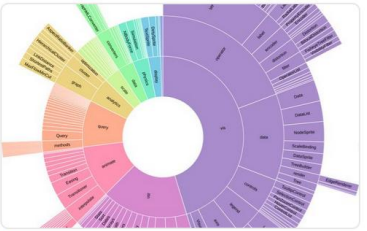

Circle packing (monochrome)


Tidy tree

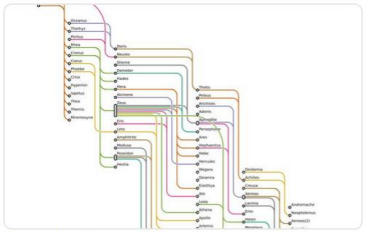

Radial tidy tree

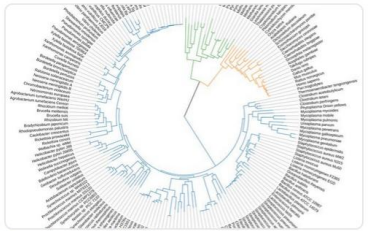

Cluster dendrogram

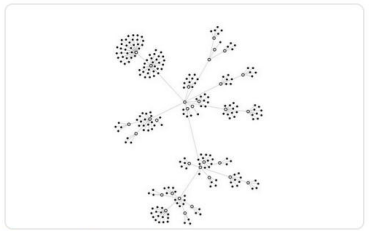

Radial dendrogram


Sunburst

Icicle

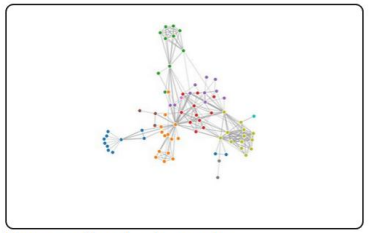

Tangled tree visualization

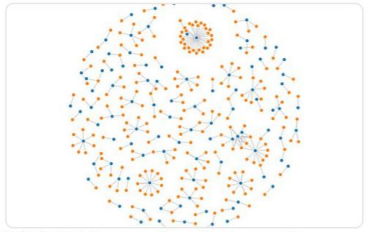

Phylogenetic tree


Force-directed tree

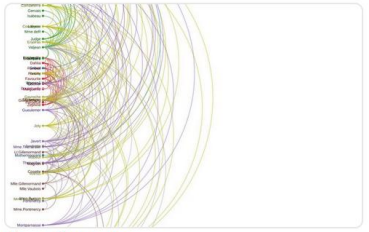
Networks

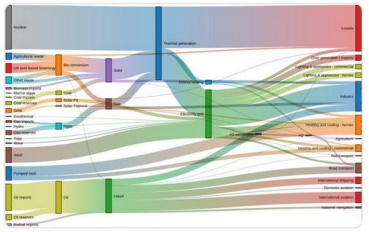
- Matrix chart
- Node-link diagrams
- Word clouds
- Alluvial diagrams

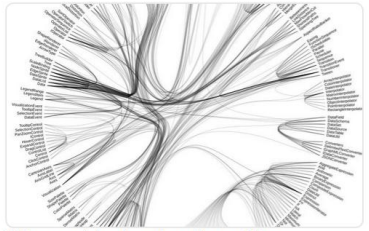

Force-directed graph

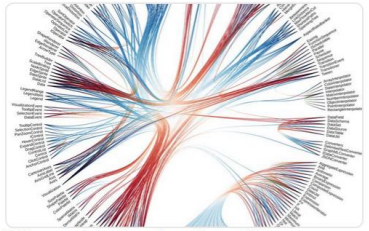

Disjoint force-directed graph

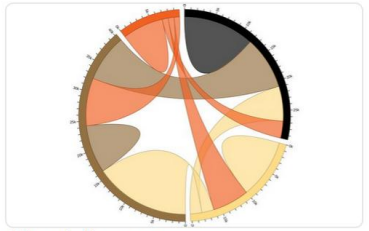
Mobile patent suits

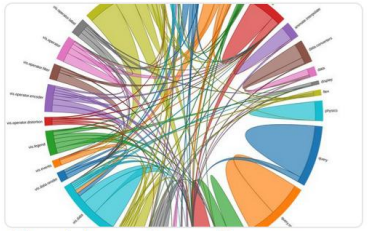

Arc diagram


Sankey diagram


Hierarchical edge bundling

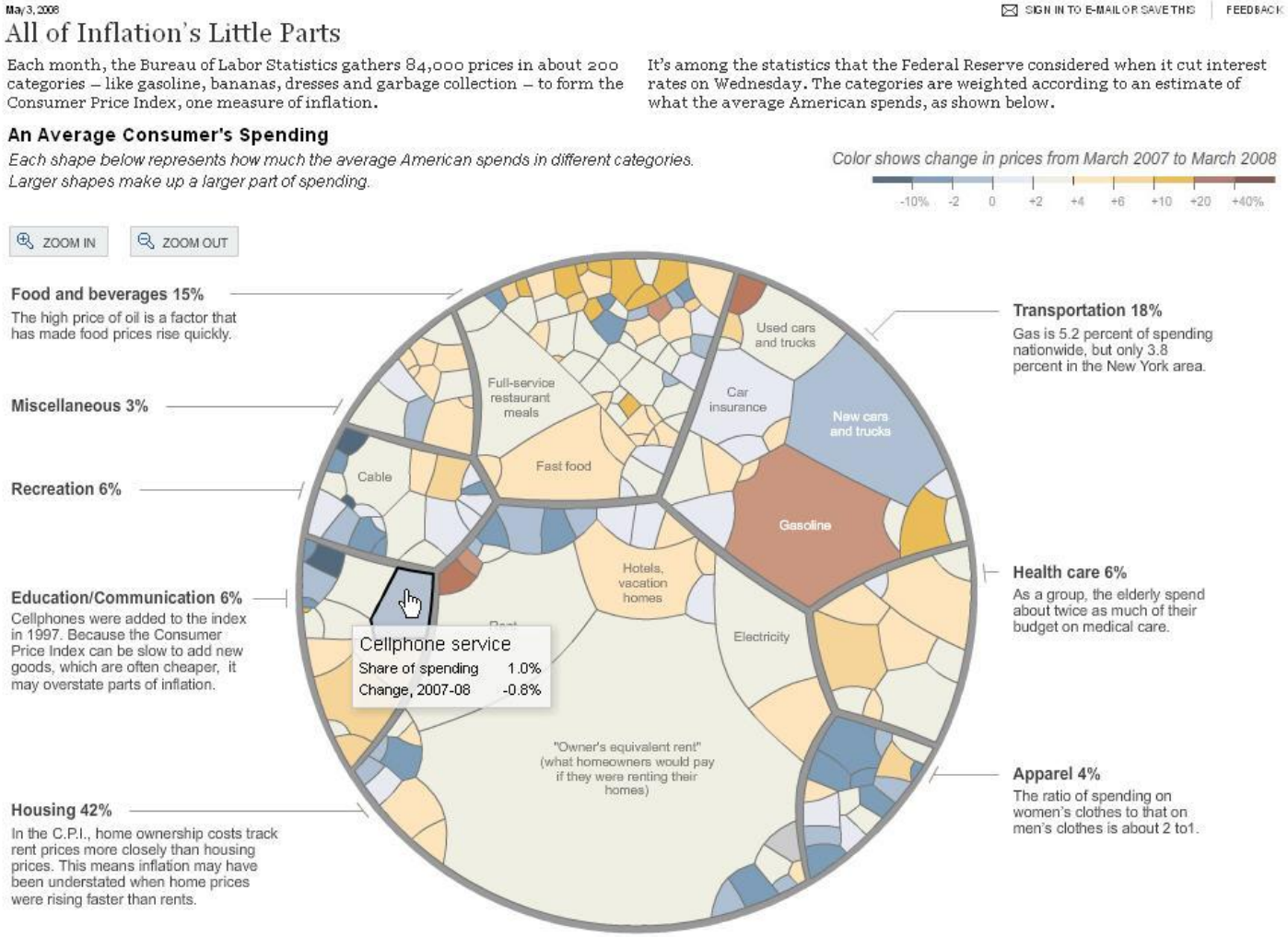

Hierarchical edge bundling


Chord diagram


Chord dependency diagram

Networks

- Voronoi Treemaps

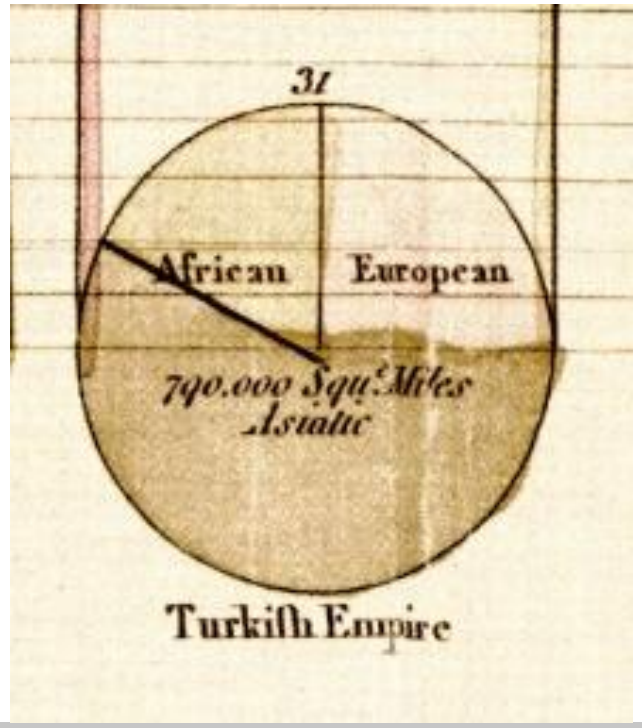


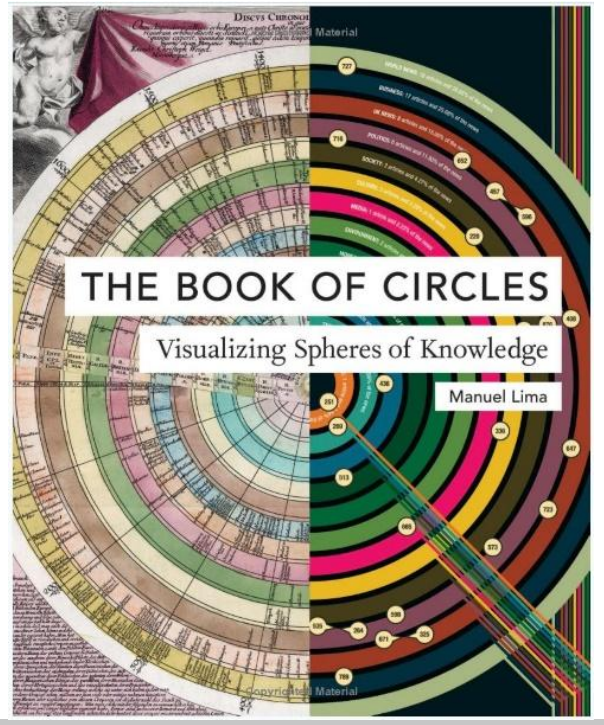
All of Inflation's Little Parts, NYTimes 2008

Radial - Pie chart

Use a pie or donut chart to:

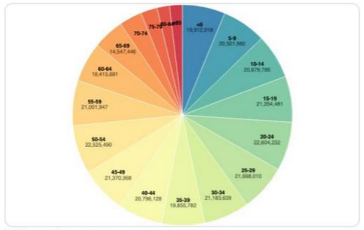
- Compare relative values
- Compare parts of a whole
- Rapidly scan your data

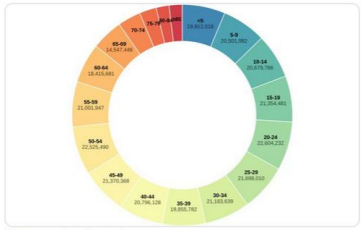


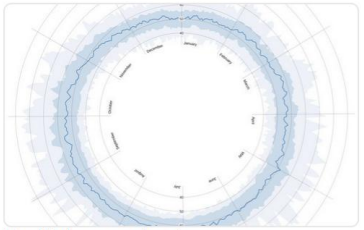


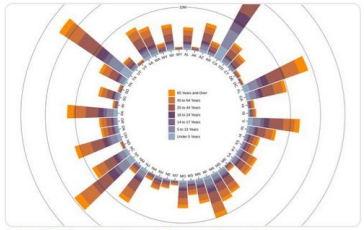
Radial - Pie chart

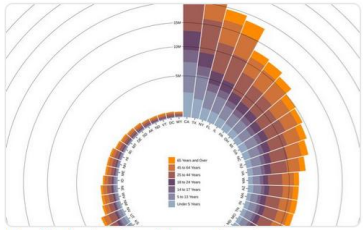
- Pie chart
- Donut chart
- Radial area chart
- Radial stacked bar chart
- Radial stacked bar chart
- Multi level (or multi layer) pie chart


Pie chart

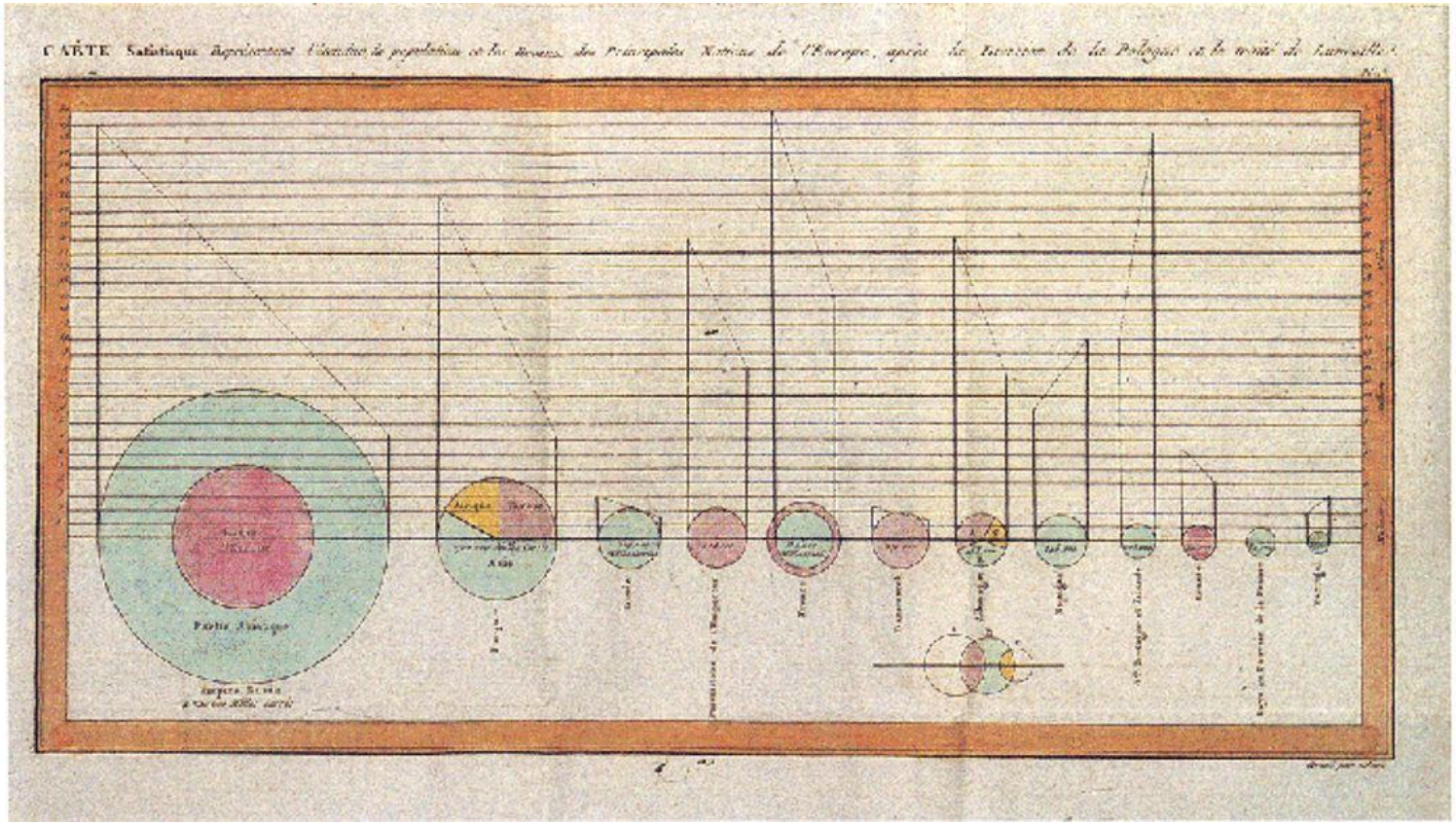

Donut chart


Radial area chart


Radial stacked bar chart


Radial stacked bar chart

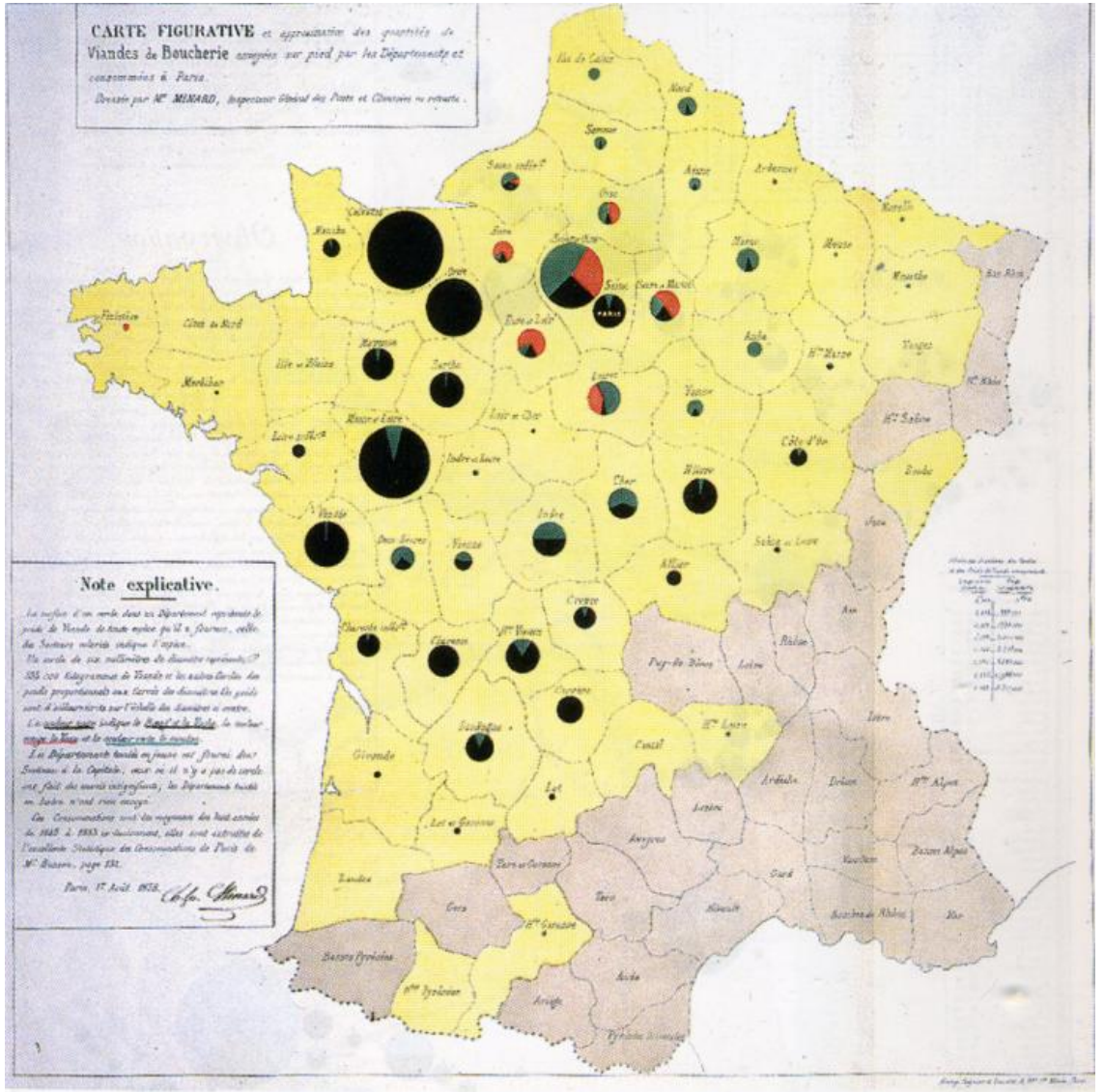
Radial - Pie chart



William Playfair, "Statistical Breviary", 1801

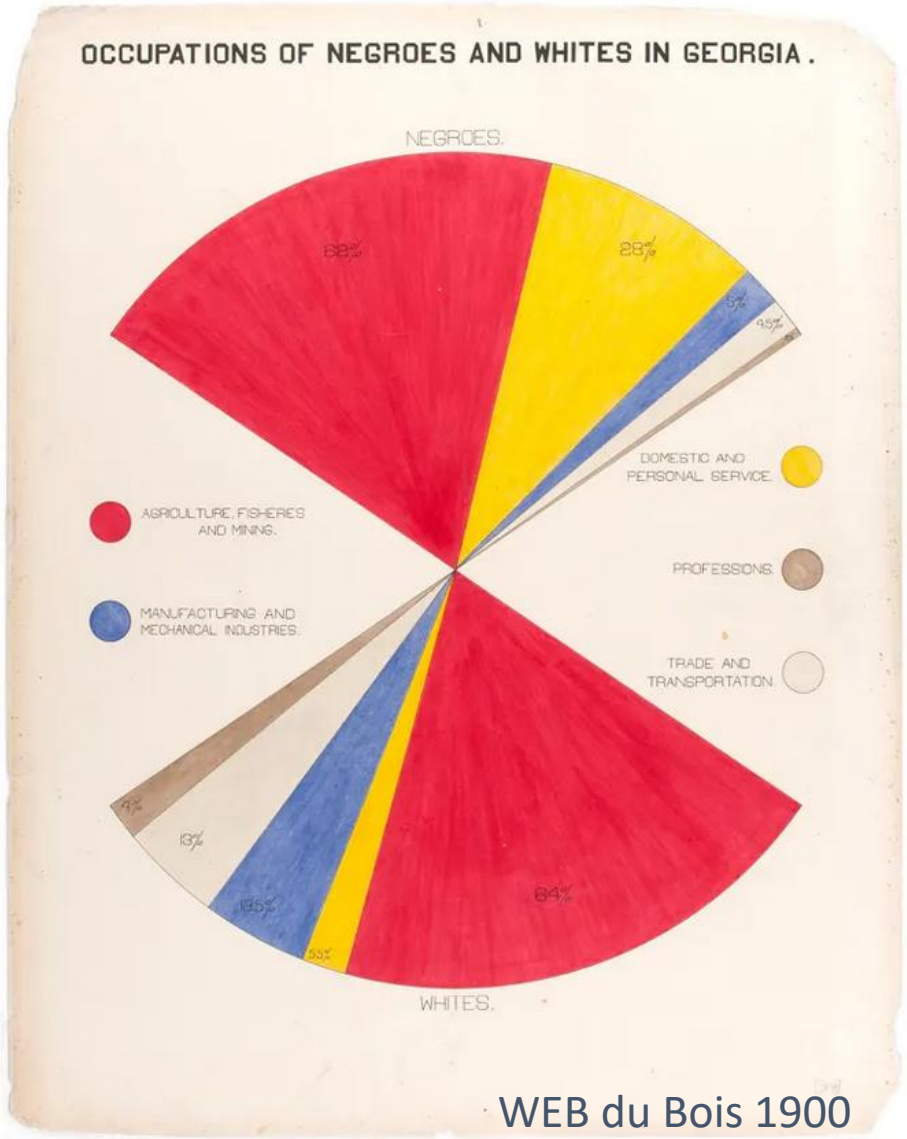
Radial - Pie chart

Minard's map using pie charts to represent the cattle sent from France’s regions to Paris (1858).

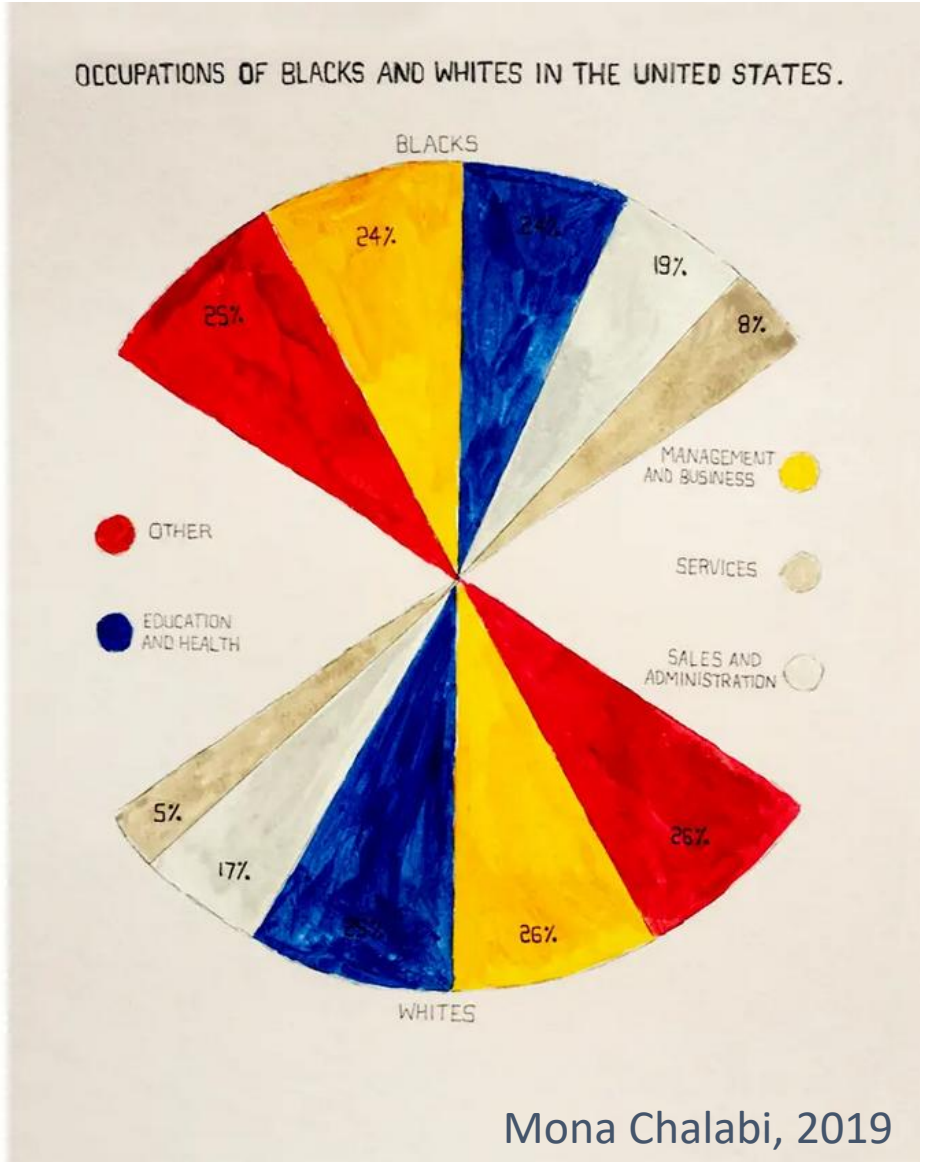


Charles Joseph Minard, 1858

Radial - Pie chart



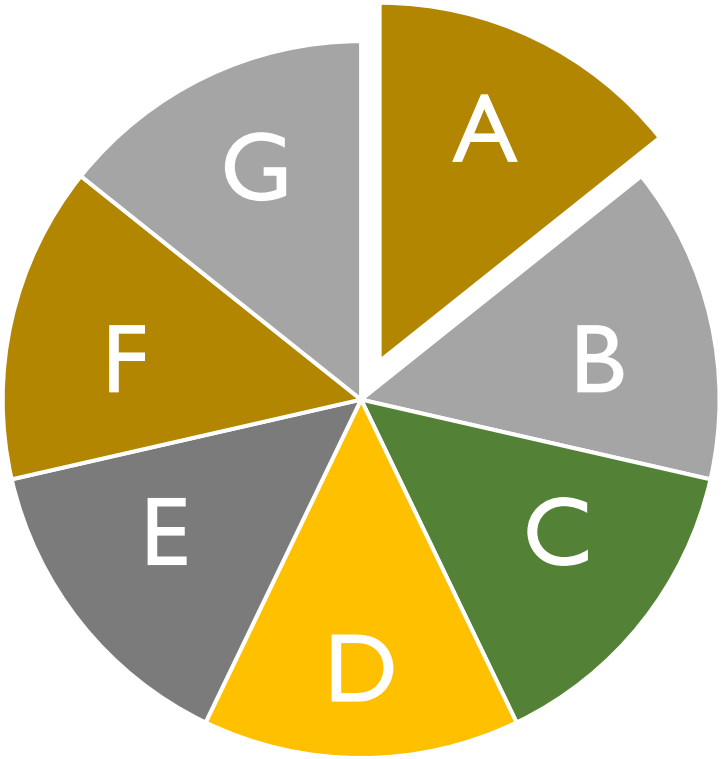
WEB du Bois 1900



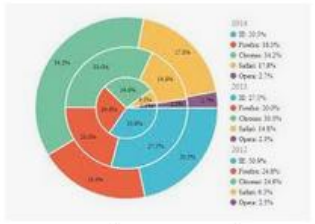
Mona Chalabi, 2019

Radial - Pie chart

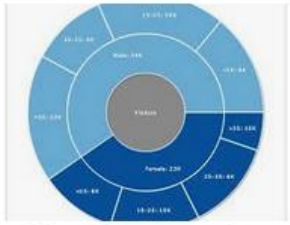
type	mark	channel	data represented
Pie chart	area	size (angle)	1 quantitative



Radial - Multi-level pie chart

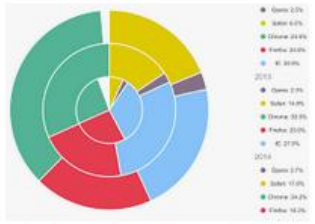


concentric pie chart - Vatan.vtngcf.org
vatan.vtngcf.org

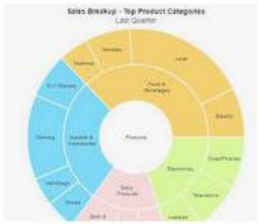


multi layer pie chart - Vatan.vtngcf.org
vatan.vtngcf.org

multi level pie chart online - Vatan ...
vatan.vtngcf.org



multi layer pie chart - Vatan.vtngcf.org
vatan.vtngcf.org

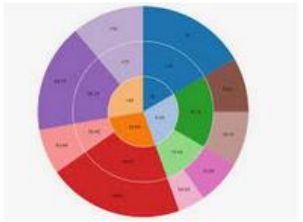


multi layer pie chart - Vatan.vtngcf.o...
vatan.vtngcf.org

multi-level pie chart ...
carrotsearch.com

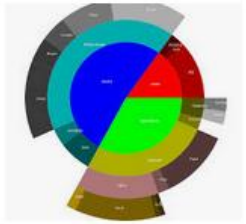
Multi-level pie chart illustrating...
researchgate.net

Multilevel #Pie chart in #powerpoin...
pinterest.com

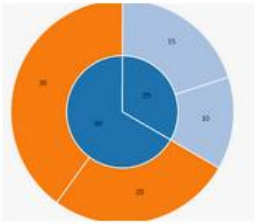


Multi Level Pie Chart Using D3 | know ...
pankajkumar005.wordpress.com

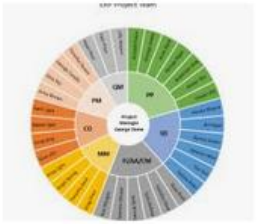
HowTo: Multilevel Pie in Excel - YouTube
youtube.com



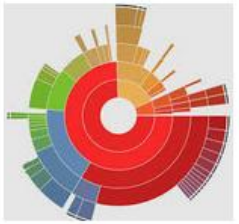
Multi Level Pie Chart in Java
blue-walrus.com



Javascript Pie Chart Library based ...
pshivale.github.io

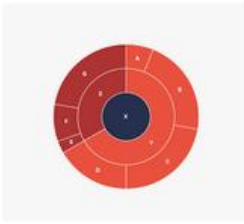


multi layer pie chart - Vatan.vtngcf...
vatan.vtngcf.org

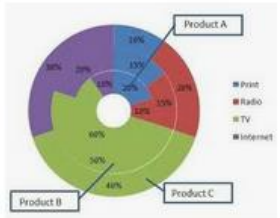


Multi-level pie chart, JavaScript...
softwarerecs.stackexchange.com

Multi-level Pie Charts
neoformix.com



Multi-level Pie Chart | Data Viz Pr...
datavizproject.com



Best Excel Tutorial - Multi Level Pie C...
best-excel-tutorial.com

Donut Chart Template for PowerPoint ...
slidebazaar.com

Nested donut chart - amCharts
amcharts.com

pie chart using html5 canvas ...
stackoverflow.com

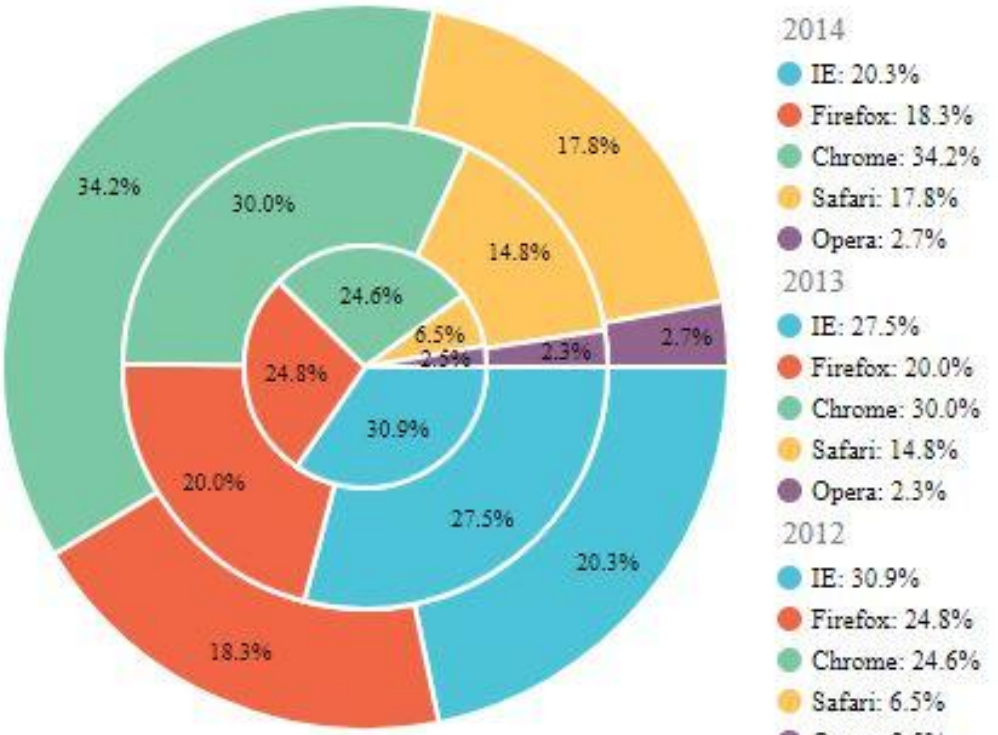
multi layer pie chart - Vatan.vtngcf.org
vatan.vtngcf.org

Radial - Multi-level pie chart

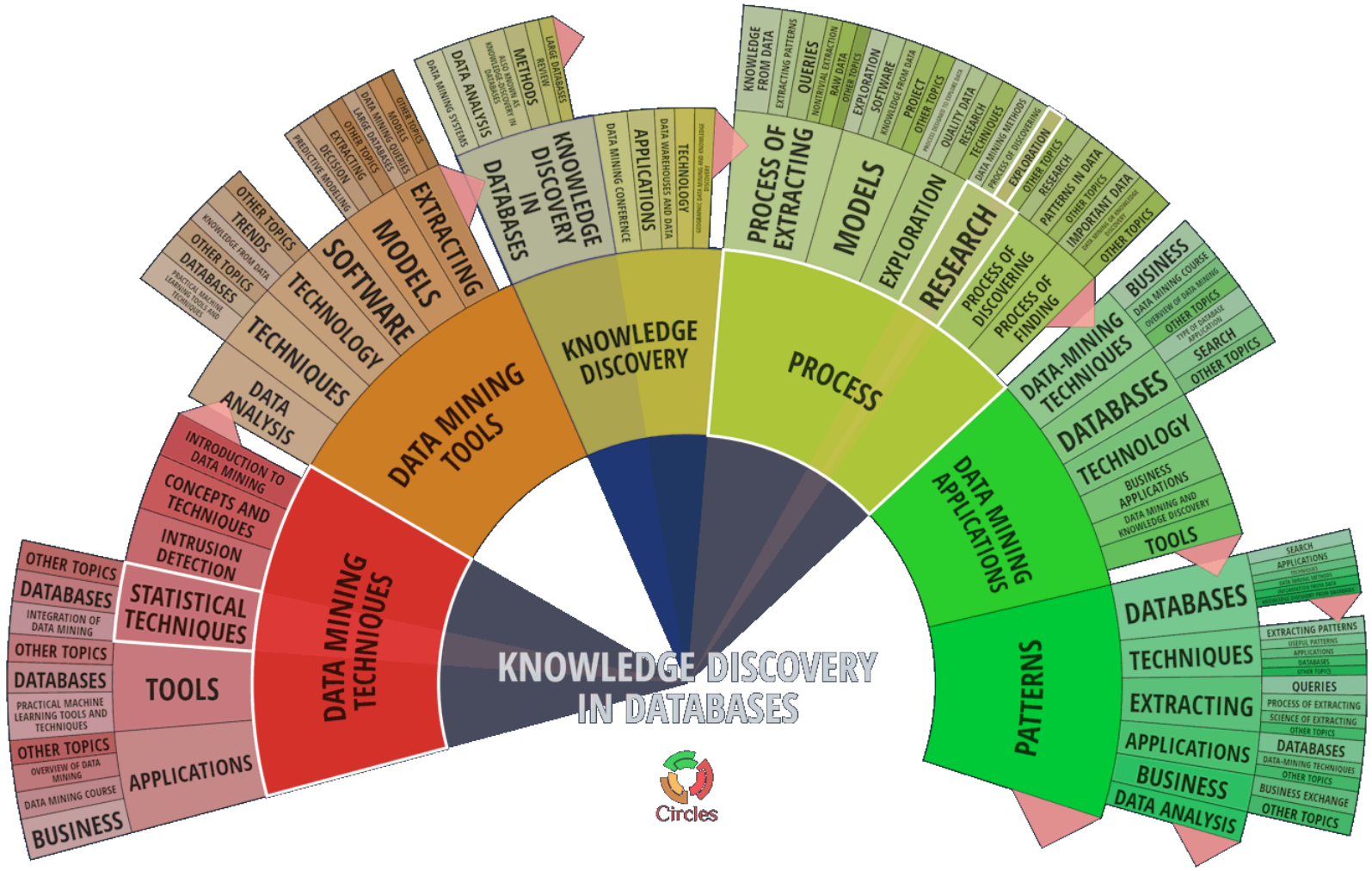
- good for nested relationships

but…

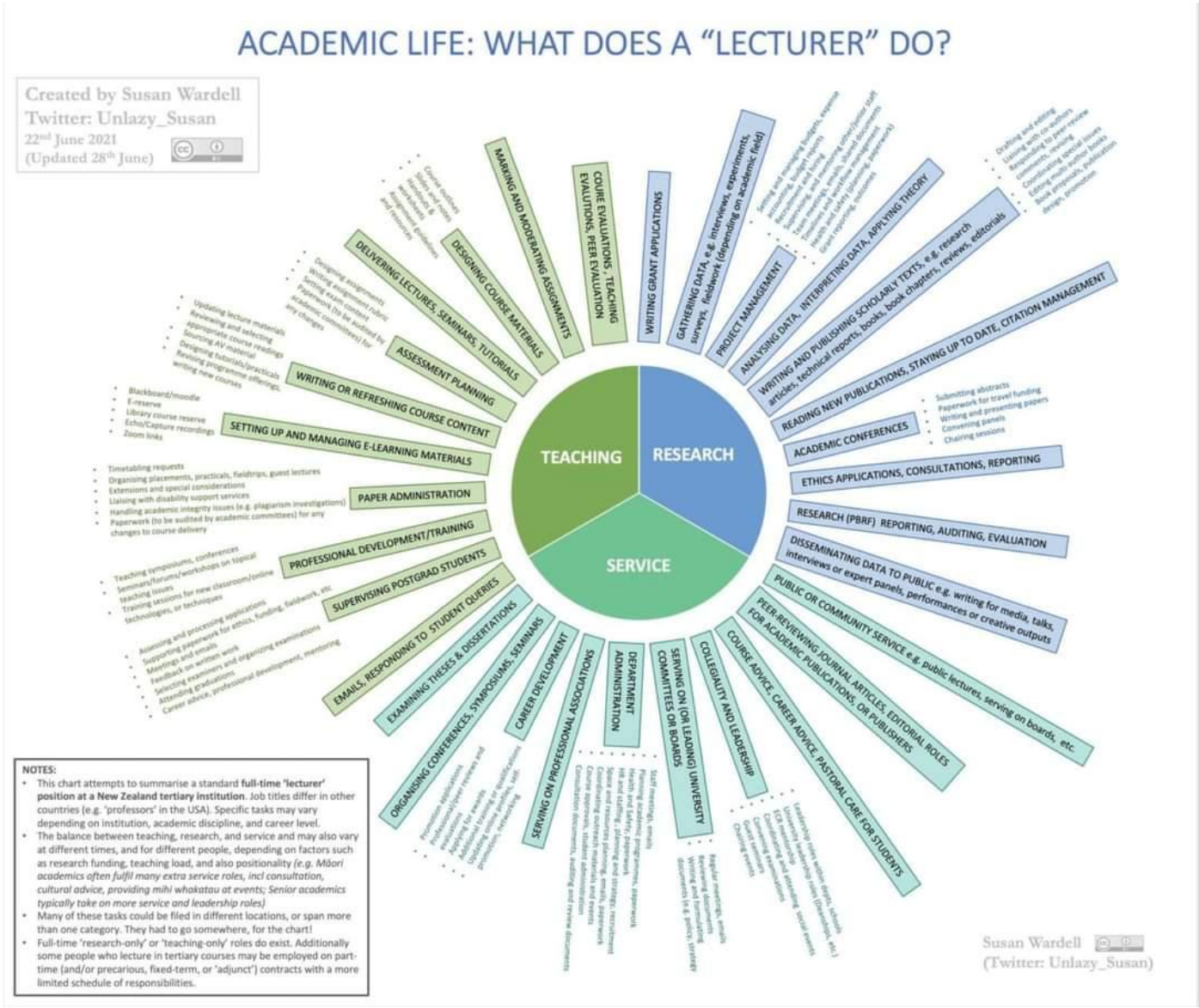
- difficult to read



Radial - Multi-level pie chart



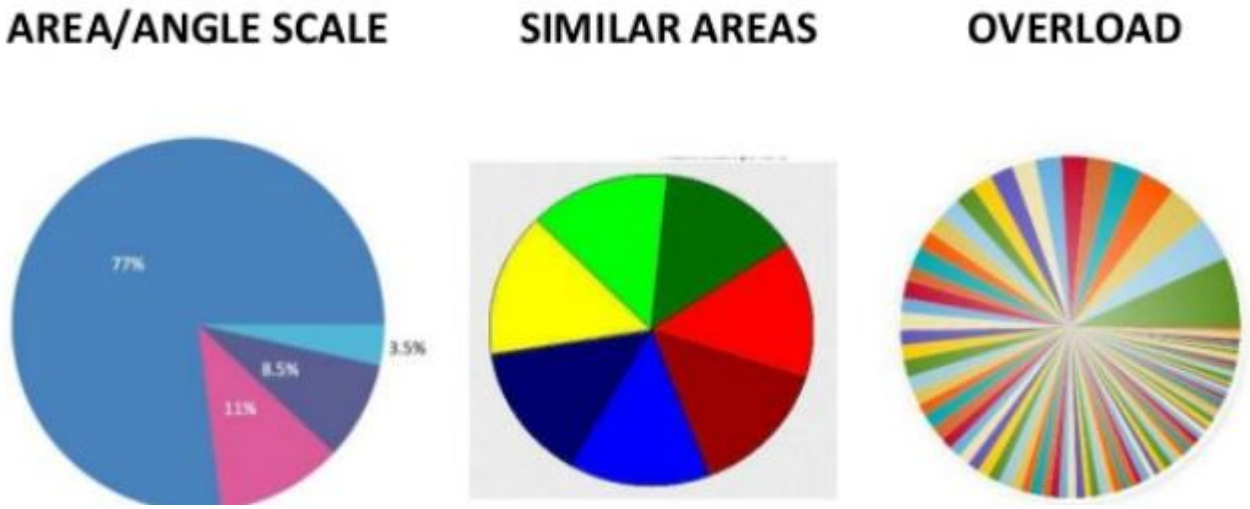
Radial - Multi-level pie chart



Radial - Pie chart dangers

Pie chart flaws:

- that they cannot show more than a few values without separating the visual encoding (the “slices”) from the data they represent (typically %)
- difficult for comparisons to be made between the size of items in a chart when area is used instead of length and when different items are shown as different shapes



When the pie chart is more complex than the data

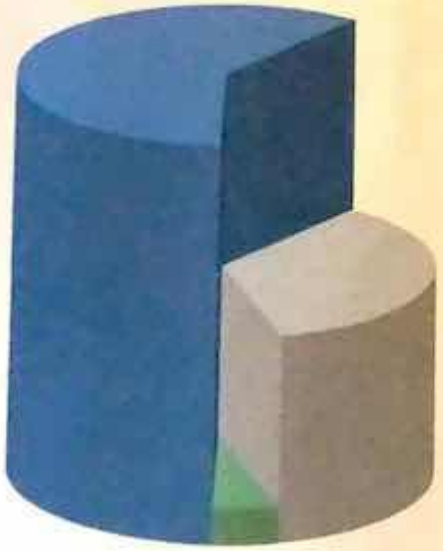
A

Imagine an investor with a $500,000 portfolio and a target allocation of 60% stocks, 35% bonds, and 5% cash investments. After a rough few months, the stock portion of his portfolio falls by $75,000:

Starting portfolio

- Stocks
  $300,000 | 60%
- Bonds
  $175,000 | 35%
- Cash investments
  $25,000 | 5%

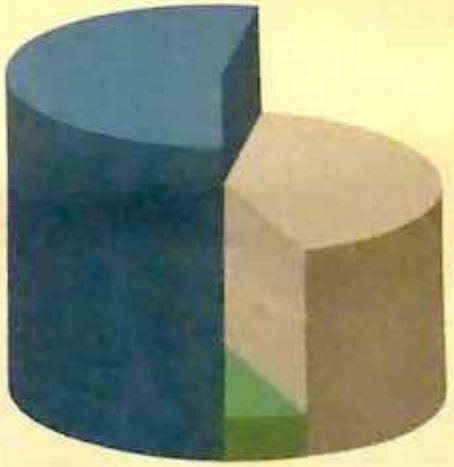
Total: $500,000



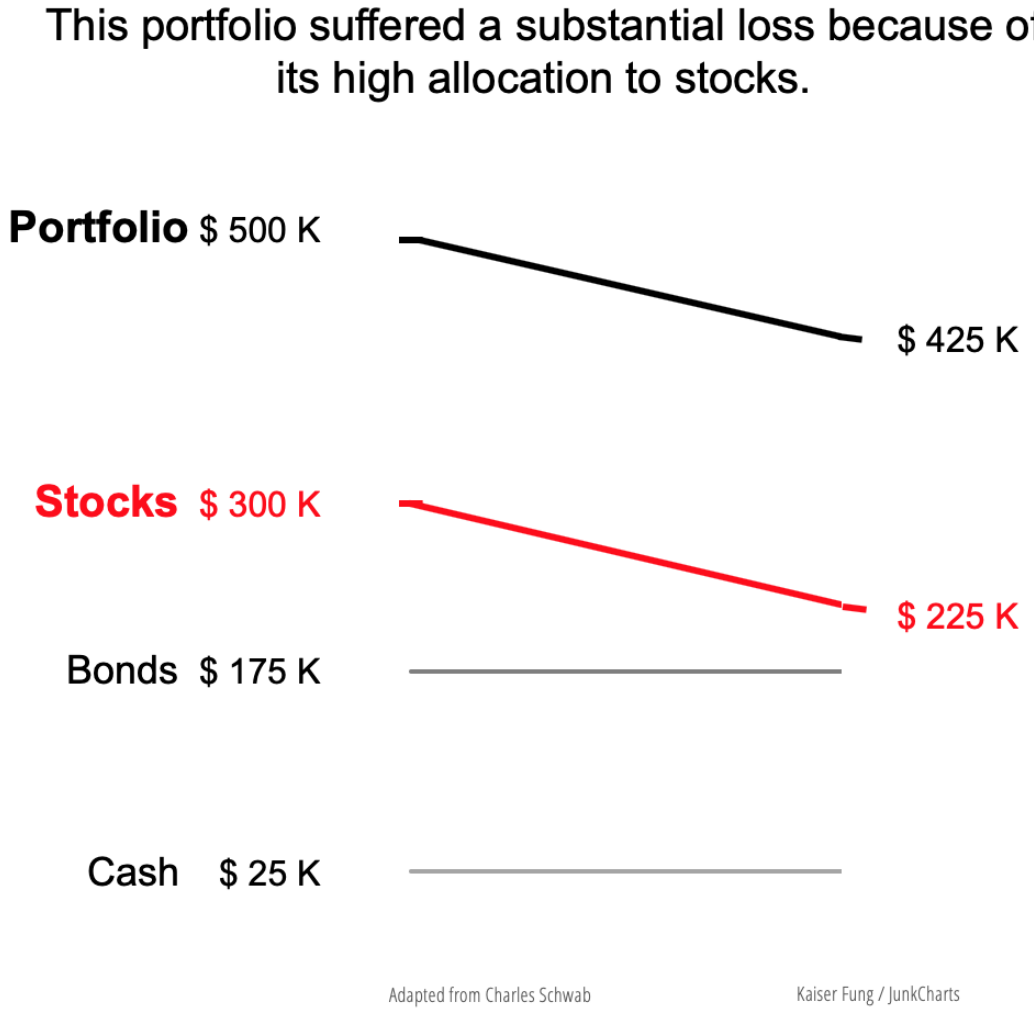
Portfolio following market correction

- Stocks
  $225,000 | 53%
- Bonds
  $175,000 | 41%
- Cash investments
  $25,000 | 6%

Total: $425,000

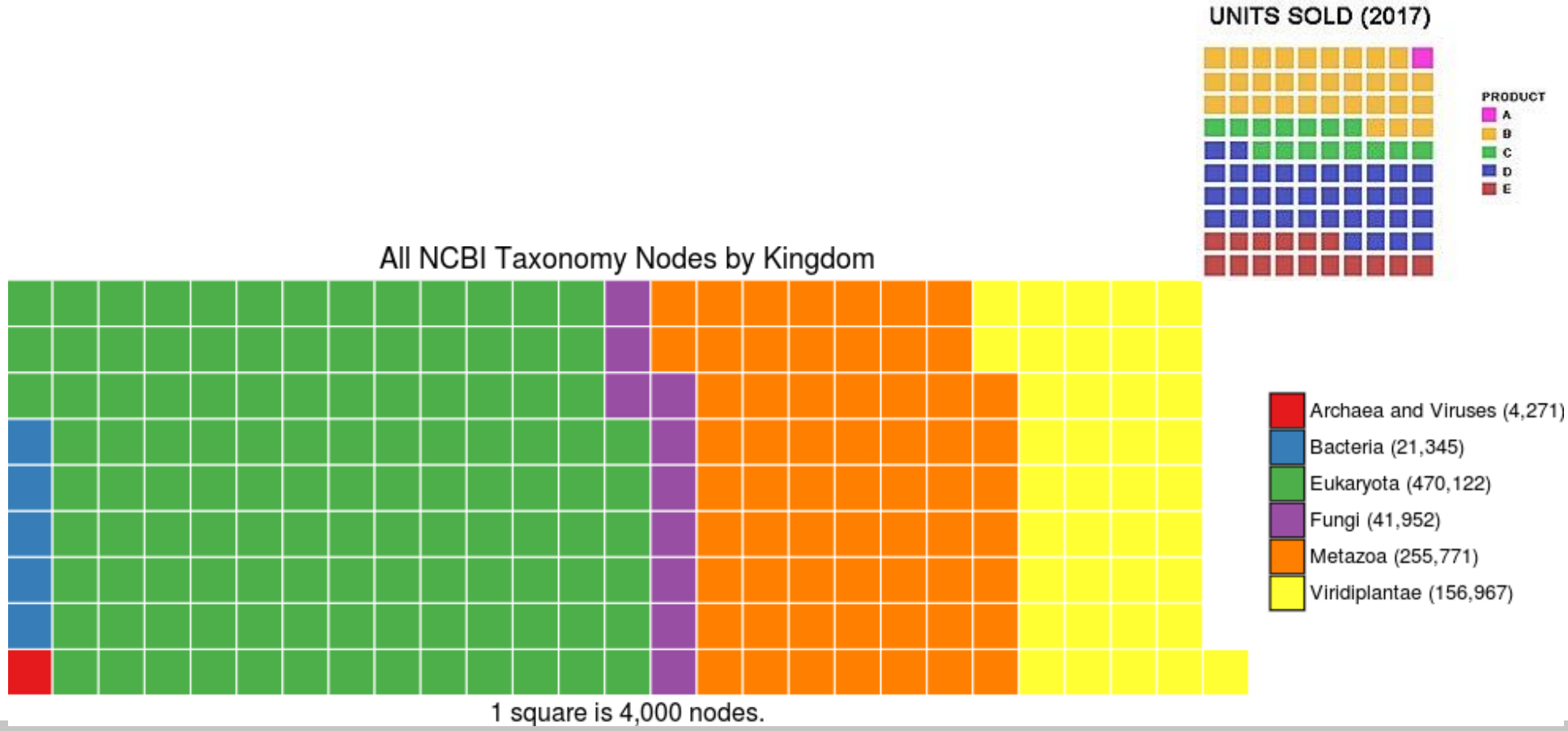


The same data but with a "bumps chart" (slopegraph)



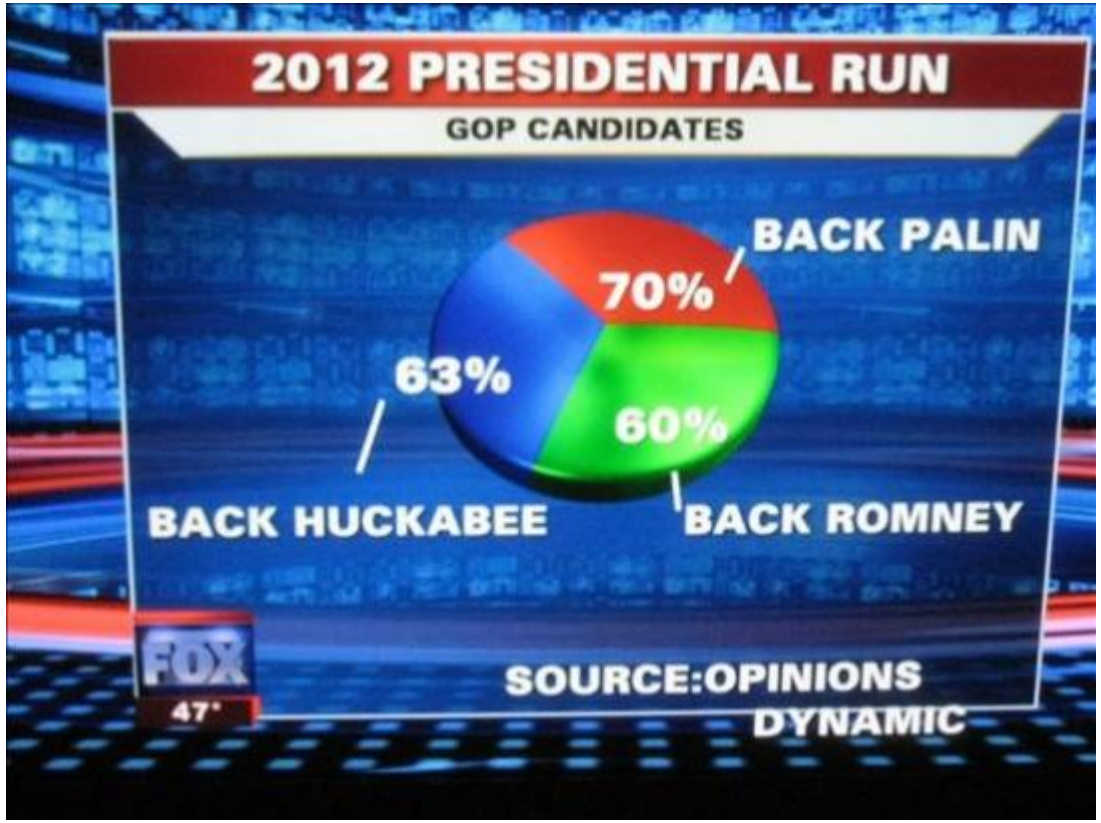
Radial - Pie chart

- Square pie chart (Waffle chart), showing how smaller percentages are more easily shown than on circular charts. On the 10x10 grid, each cell represents 1%.



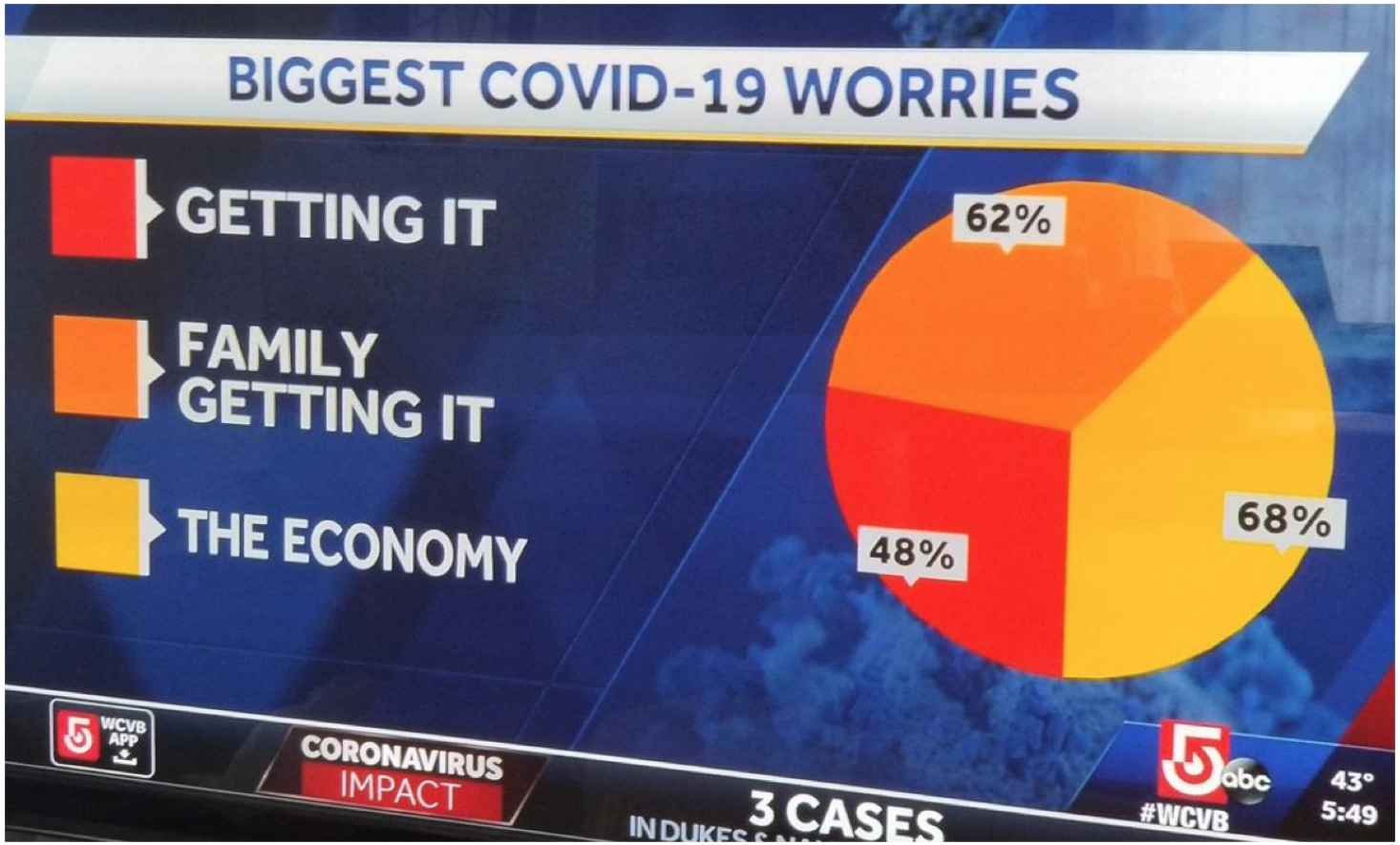
Radial - Pie chart

Beware of the numbers!



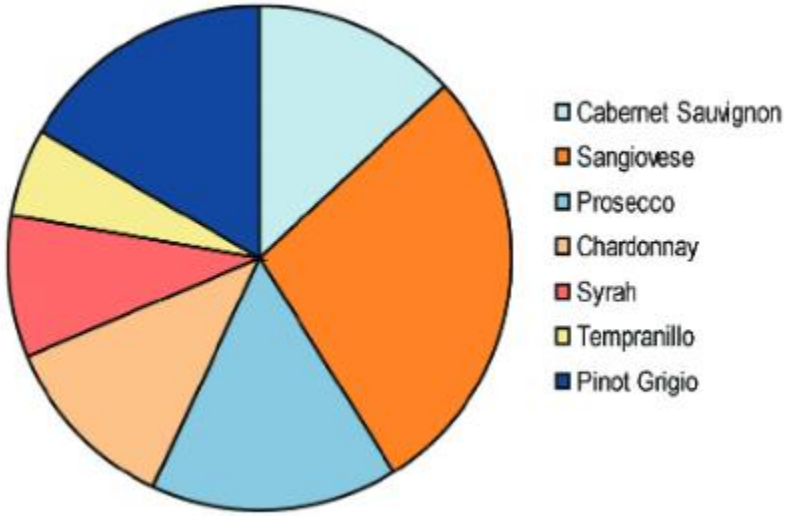
Radial - Pie chart

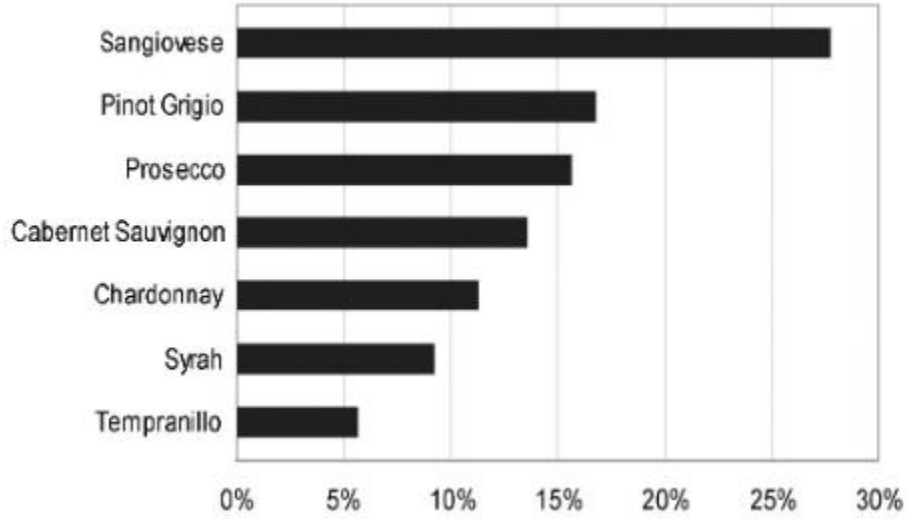
Beware of the numbers!



Found on Kaiser Fung's Junk Charts site

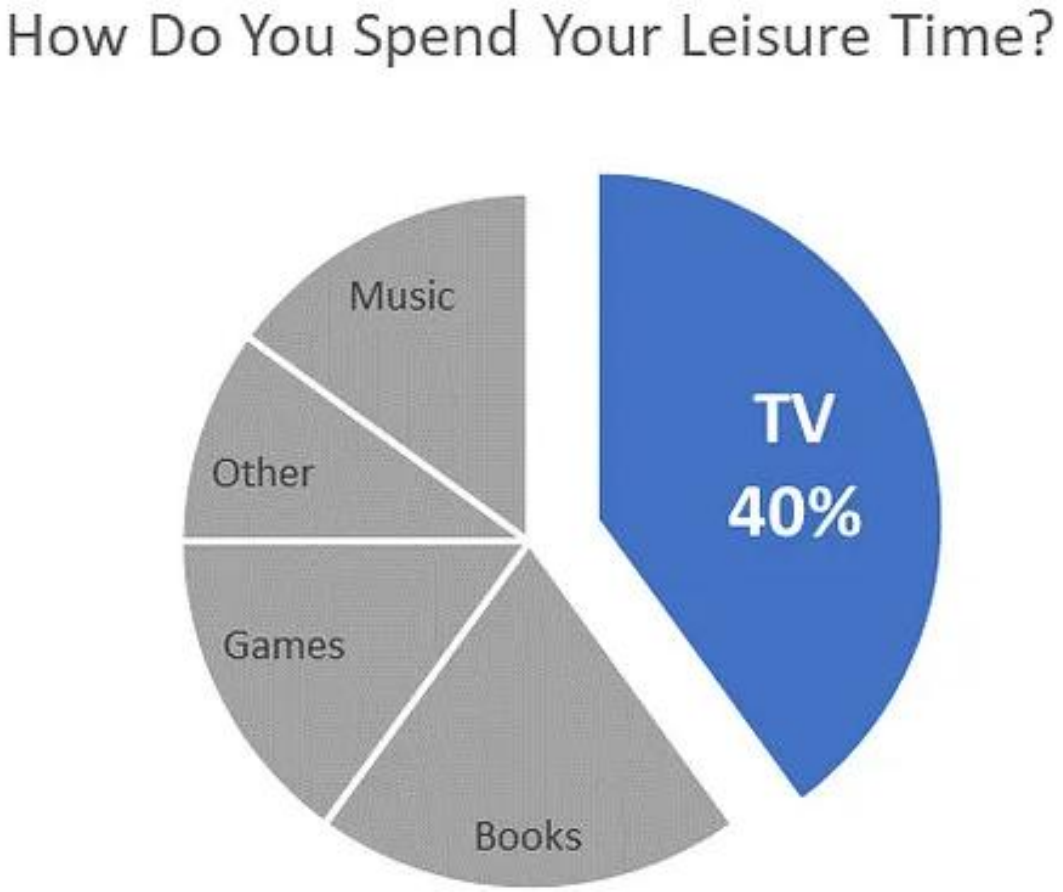
Bar charts instead of pie charts





Making the Case for Pie Charts

- A simple way to help focus attention



Radial - Pie chart

Don’t use a pie chart for:

- precise comparisons of data (not because it’s not capable of it, but there are better visualizations to choose from if precision is a requirement.

Pie or donut chart best practices:

- Make sure the slices of your pie (or donut) equal 100% (add numerical values and % to your visualization)
- Order the pieces of your pie according to size
- Don’t compare more than 5 categories in a pie chart, otherwise you run the risk of unclear differentiation between slices

The world’s most perfect pie chart machines – fun quote

A **watch** (for visualizing time):

Douglas Adams (The Hitchhiker's Guide to the Galaxy) on Digital Watches:

“…the world’s most perfect pie chart machines strapped to our wrists, which we could read at a glance…”

“…trying to find ways of translating purely numeric data into graphic form so that the information leapt easily to the eye.”

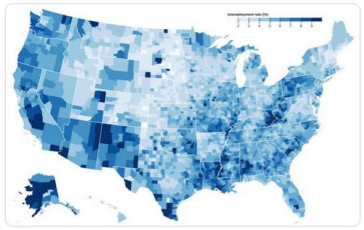
The world's most perfect pie chart machines – fun quote

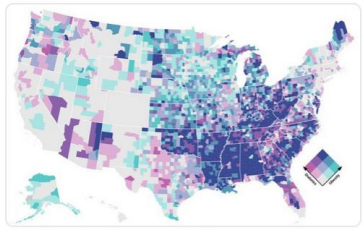
"Digital watches came along at a time that, in other areas, we were trying to find ways of translating purely numeric data into graphic form so that the information leapt easily to the eye. For instance, we noticed that pie charts and bar graphs often told us more about the relationships between things than tables of numbers did.

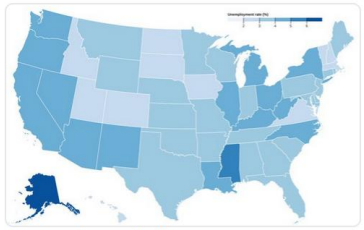
So we worked hard to make our computers capable of translating numbers into graphic displays. At the same time, we each had **the world's most perfect pie chart machines strapped to our wrists**, which we could read at a glance, and we suddenly got terribly excited at the idea of translating them back into numeric data, simply because we suddenly had the technology to do it… so digital watches were mere technological toys rather than significant improvements on anything that went before."

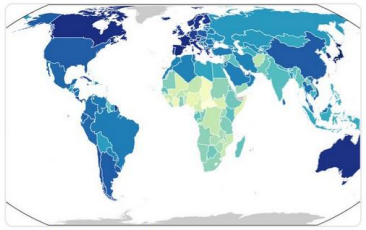
Douglas Adams (The Hitchhiker's Guide to the Galaxy) on Digital Watches: a Fax to US editor, January 1992

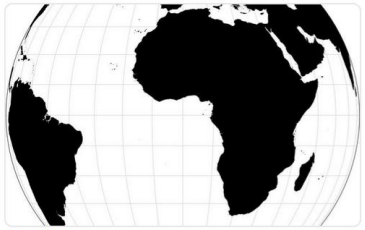
Maps

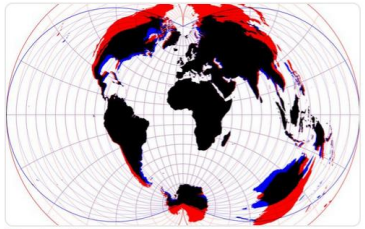

Choropleth


Bivariate choropleth

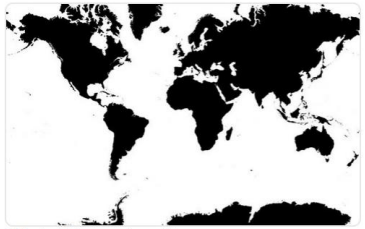

State choropleth


World choropleth


World map


Projection comparison

Tissot's indicatrix


Web Mercator tiles

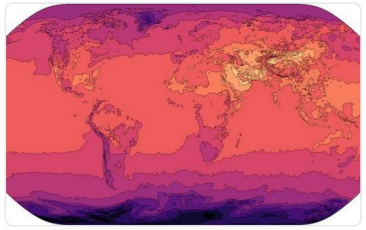
Raster tiles

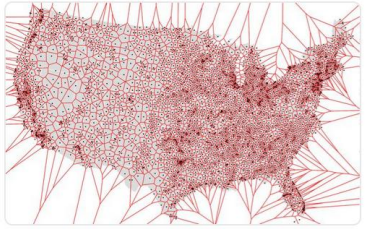
Vector tiles

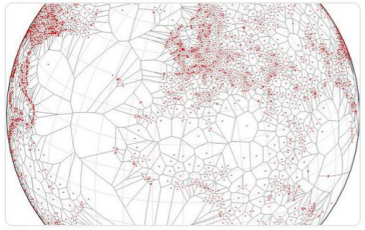
Clipped map tiles

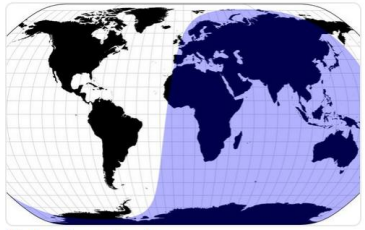
Raster & vector

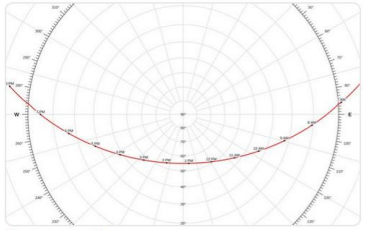
Vector field

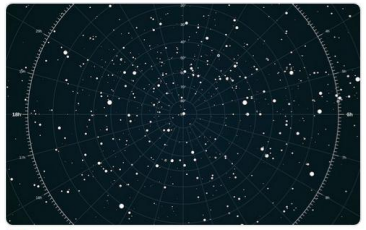

GeoTIFF contours

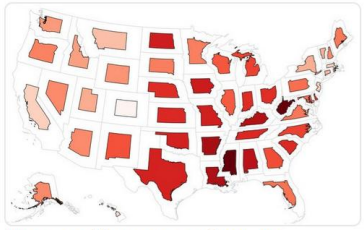

U.S. airports voronoi


World airports voronoi


Solar terminator


Solar path


Star map


Non-contiguous cartogram

Maps

- symbol maps - nominal point data (maybe shapes or colors vary)
- dot maps - ordinal point data (size of brightness of dot vary)
- land use maps - nominal area data (areas are colored or textured)
- choropleth maps - ordinal area data (size of glyph or brightness of area vary)
- line diagrams - nominal or ordinal line data (vary in thickness, color, etc)
- isoline maps - ordinal surface data (heat map)
- surface maps - ordinal volume data

Nature of Geographic Phenomena

- Spatial Dimension
  - 0d - point phenomena located in 2d or 3d space (e.g. data collected at weather monitoring stations)
  - 1d - linear phenomena (e.g. the path an AUV or a drone takes while taking measurements)
  - 2d - areal phenomena (e.g. data collected on the surface of a lake)
  - 2.5d - volumetric phenomena - each x, y position has a single z value associated with it (e.g. the maximum depth at any point in the lake)
  - 3d - volumetric phenomena - each x, y, z position has a value associated with it (e.g. the ph values collected at various points and depths in the lake)

Principles of Symbolization chapter in Thematic Cartography and Geovisualization, 3rd ed. by Slocum, McMaster, Kessler, & Howard

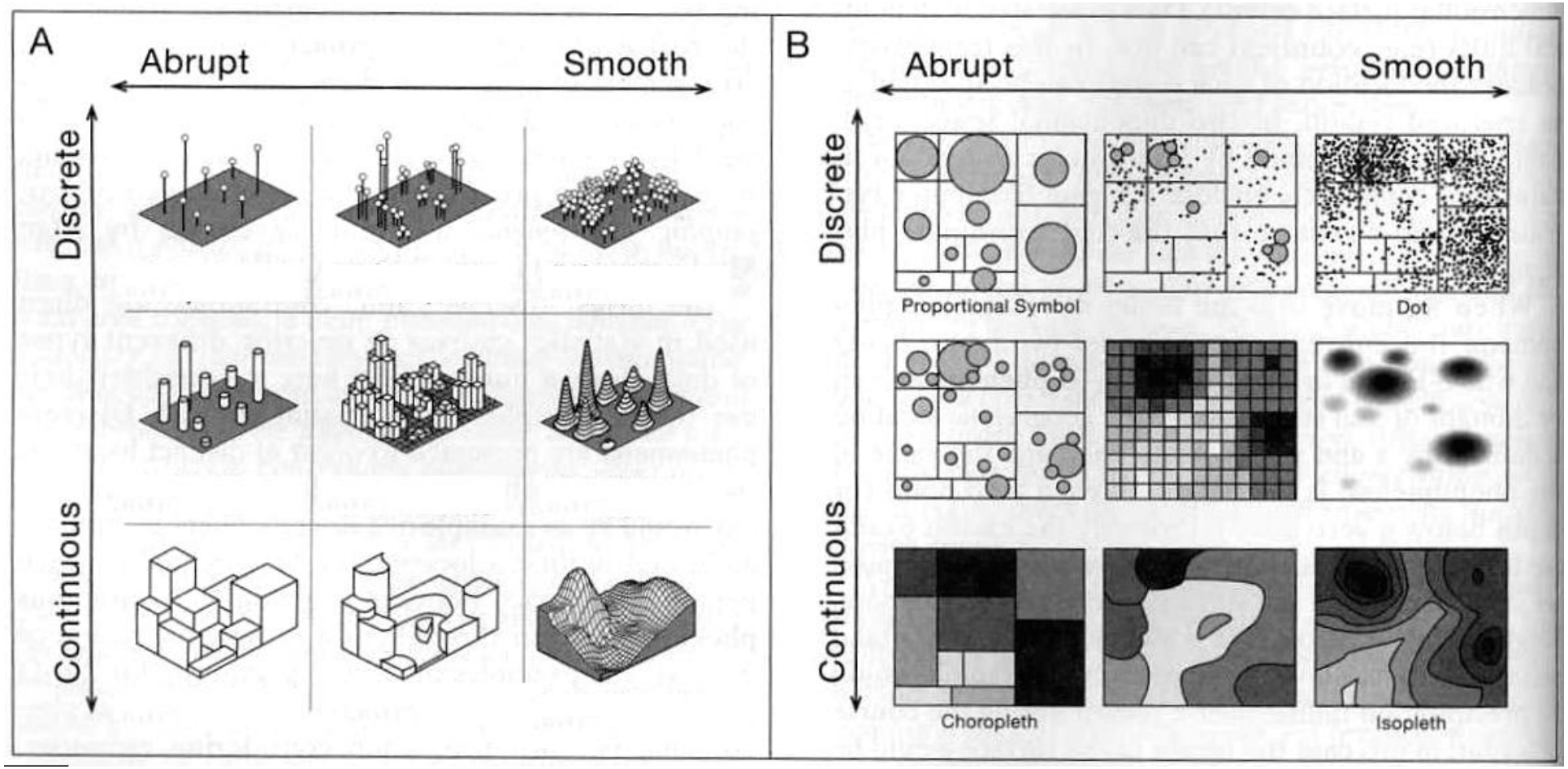
Nature of Geographic Phenomena

- Discrete vs Continuous Phenomena
  - discrete - occur at distinct locations (and have a space between them)
  - continuous - occur throughout a region of interest

- Abrupt vs Smooth Phenomena
  - abrupt - can change suddenly
  - smooth - change gradually

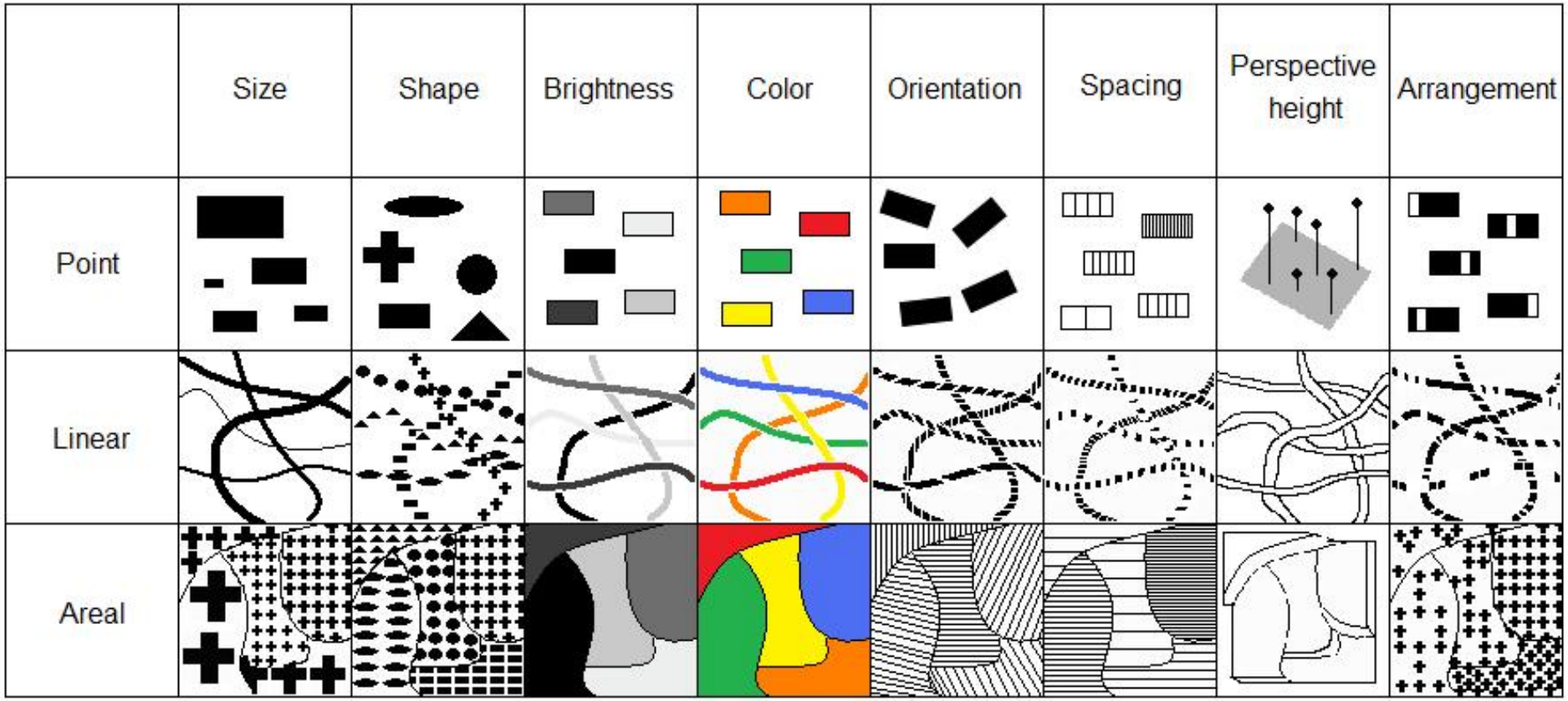
Principles of Symbolization chapter in Thematic Cartography and Geovisualization, 3rd ed. by Slocum, McMaster, Kessler, & Howard

Nature of Geographic Phenomena



Principles of Symbolization chapter in Thematic Cartography and Geovisualization, 3rd ed. by Slocum, McMaster, Kessler, & Howard

Visual variables for spatial data



Nature of Geographic Phenomena

- Distinction between data that has been collected to represent a phenomenon and the phenomenon being mapped
  - i.e. we are typically collecting data at discrete sites (weather stations, well sites) or aggregating over small regions (counties, states) where the actual phenomena being modeled is continuous. Other times we are collecting discrete data on a discrete phenomena.

Principles of Symbolization chapter in Thematic Cartography and Geovisualization, 3rd ed. by Slocum, McMaster, Kessler, & Howard

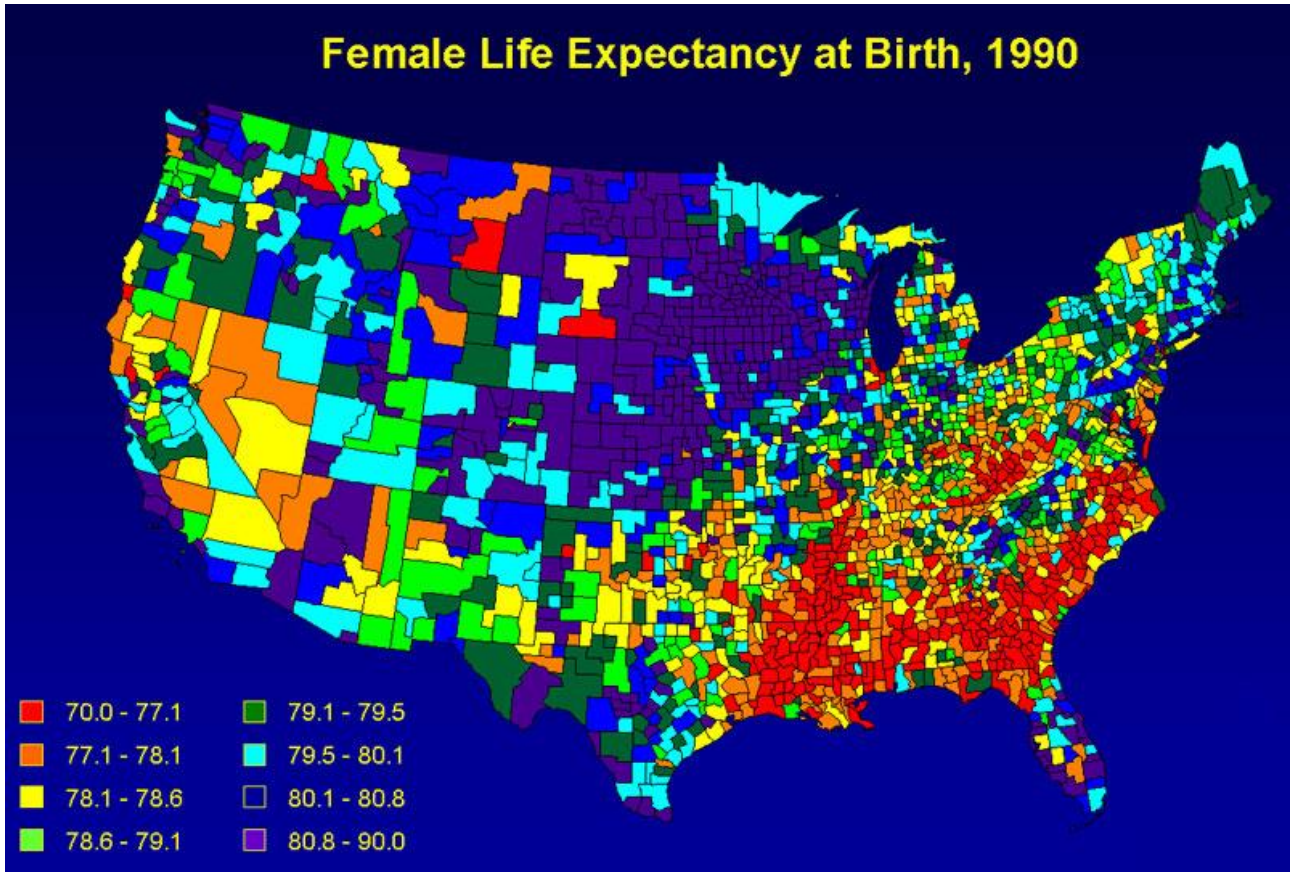
Nature of Geographic Phenomena

- Type of visualization used depends both on the nature of the underlying phenomenon and the purpose of the map

Principles of Symbolization chapter in Thematic Cartography and Geovisualization, 3rd ed. by Slocum, McMaster, Kessler, & Howard

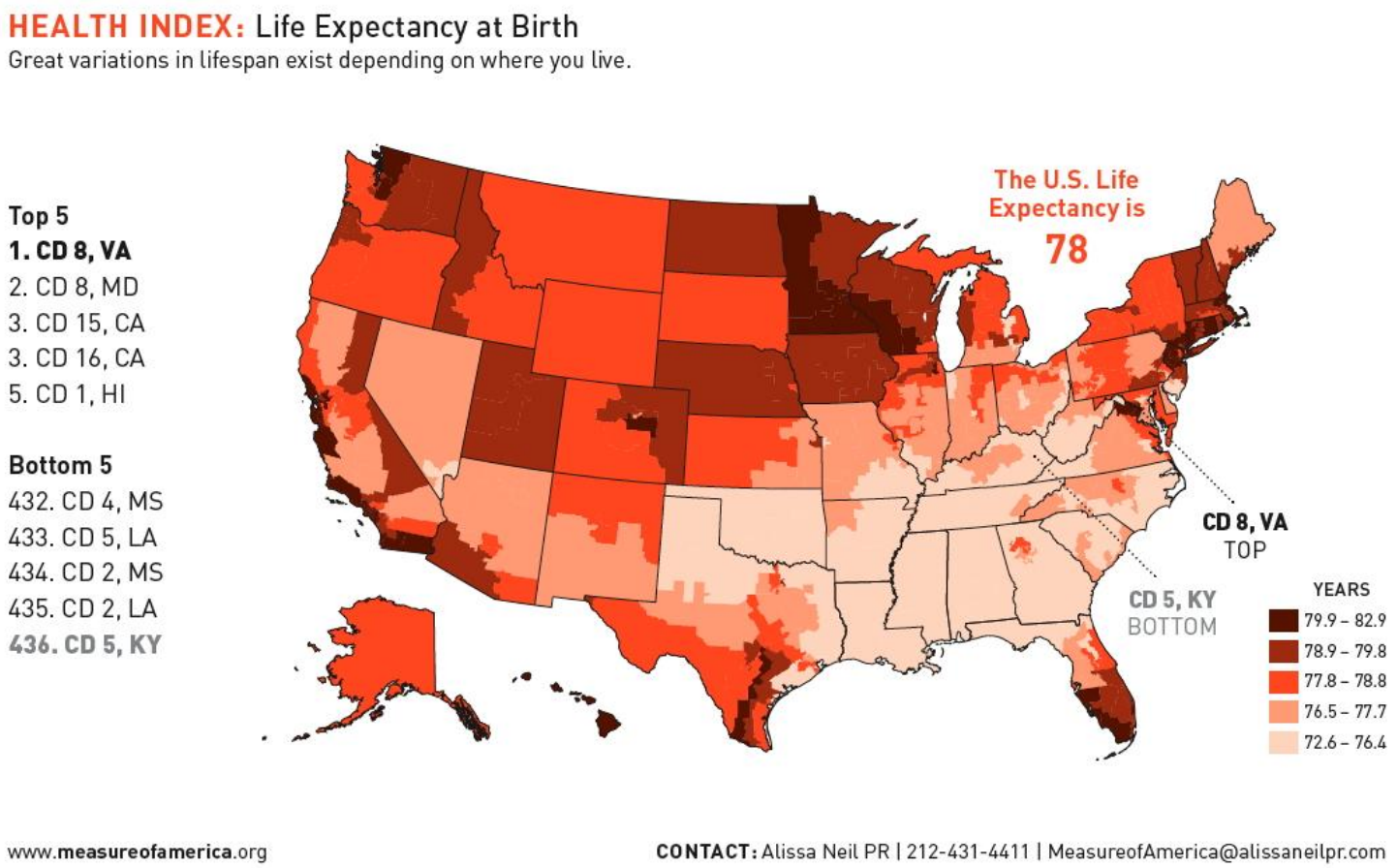
Maps

- Different ways of using colour to map life expectancy in the US (quantitative)

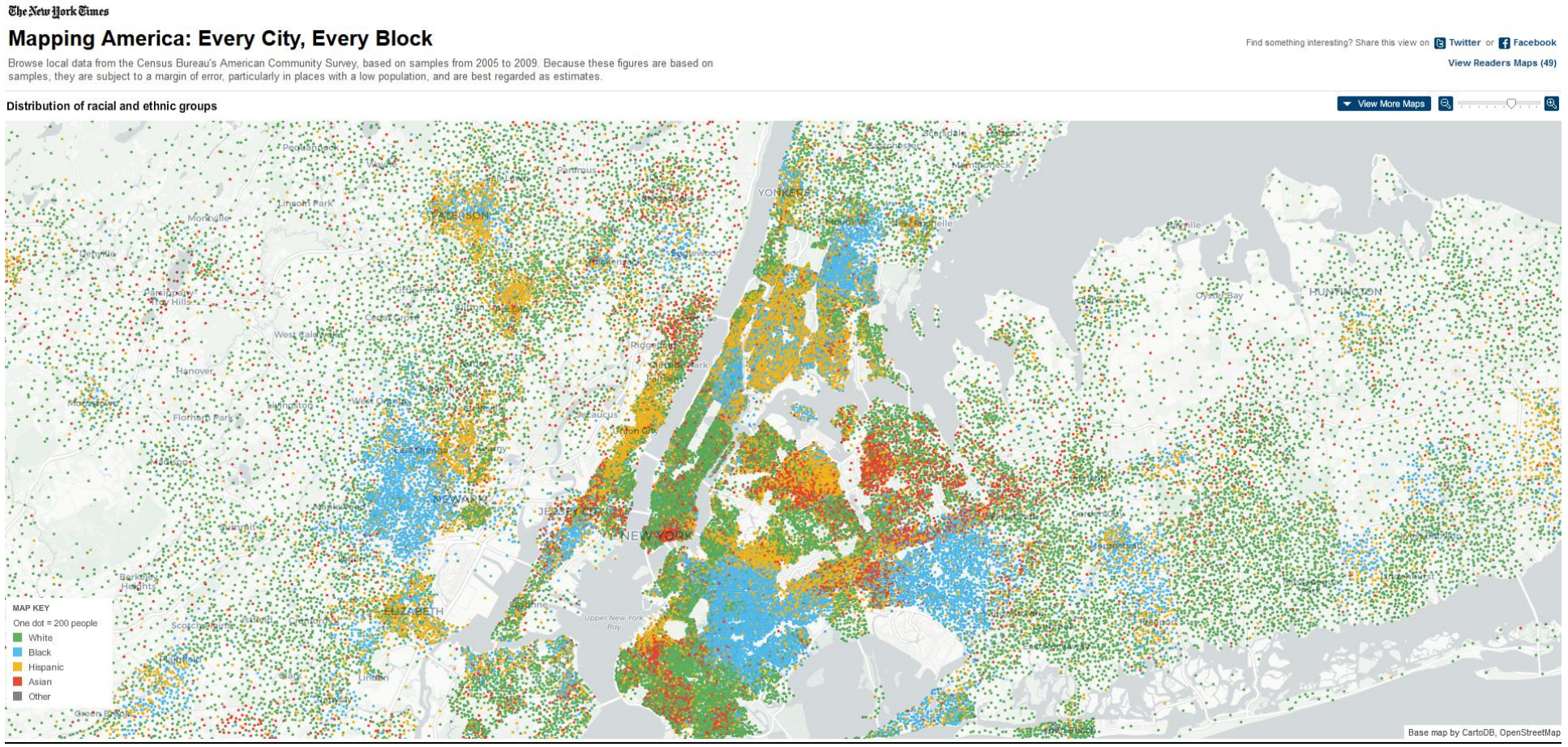


Maps

- Different ways of using colour to map life expectancy in the US (quantitative)

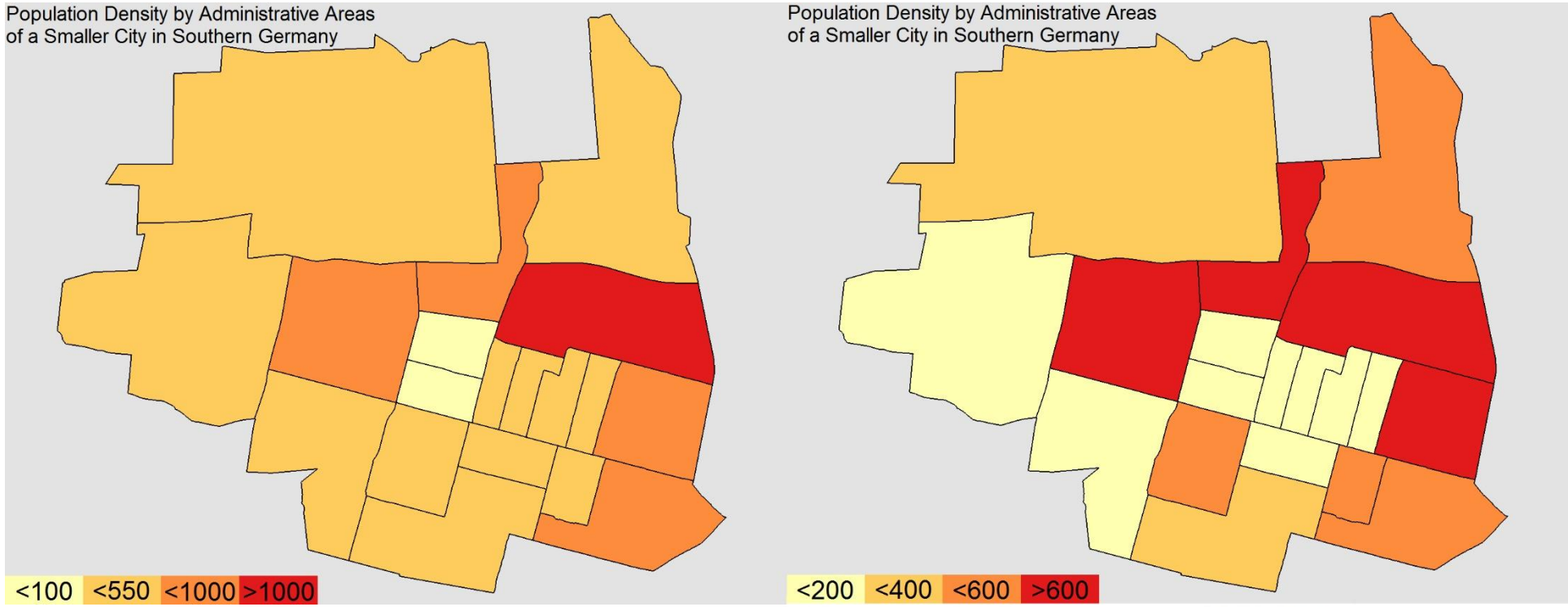


Dot maps

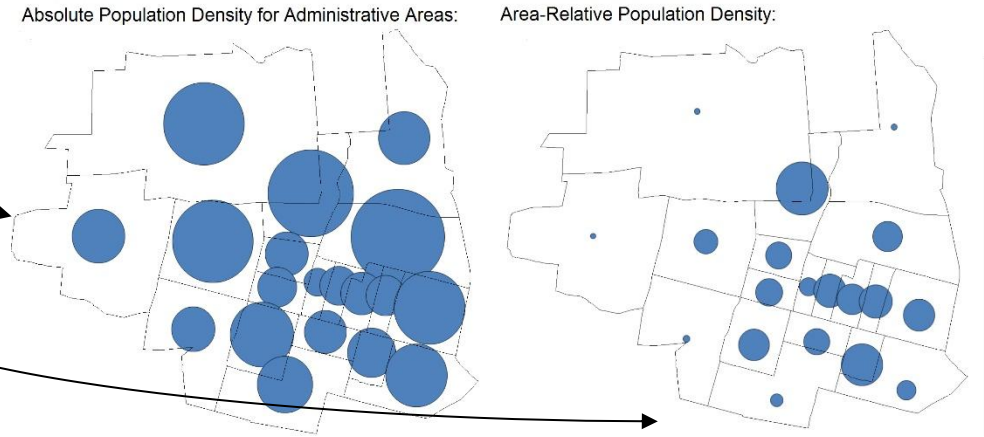


New York Times, 2005-2009

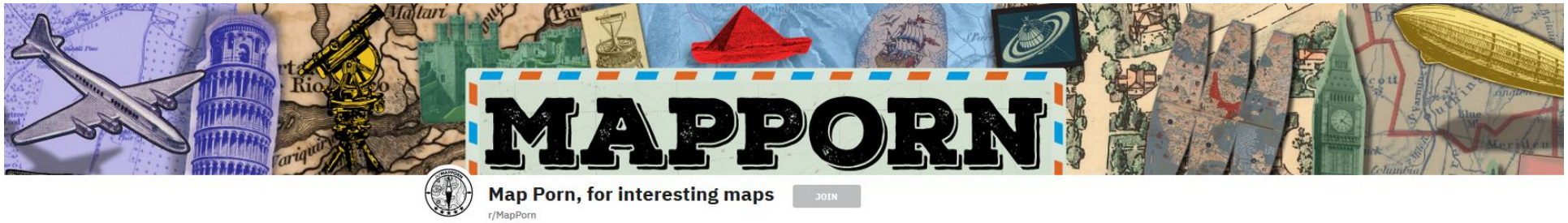
Choropleth maps



Whether you use **actual values** or **ratios**, the results will vary

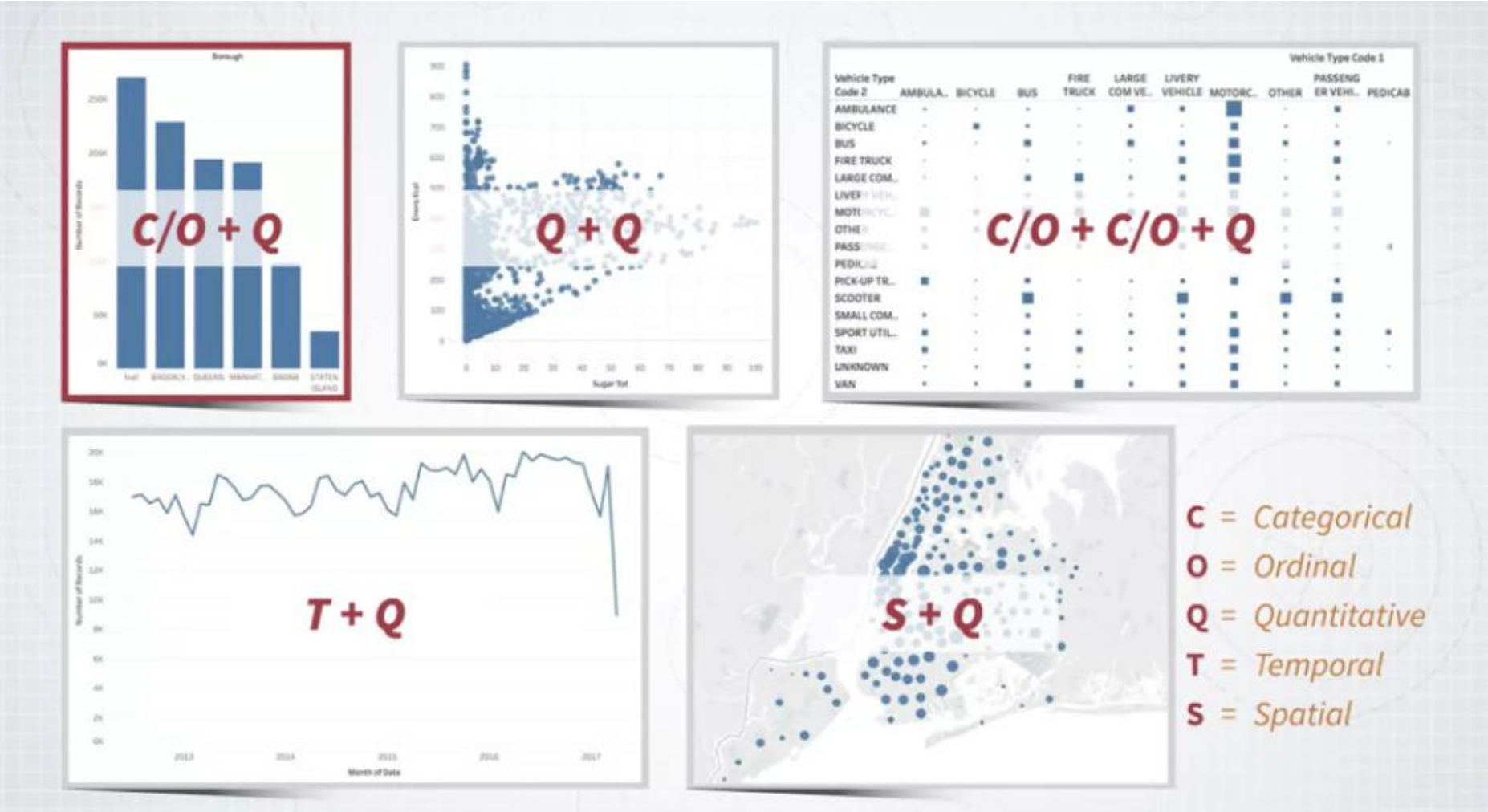


Maps

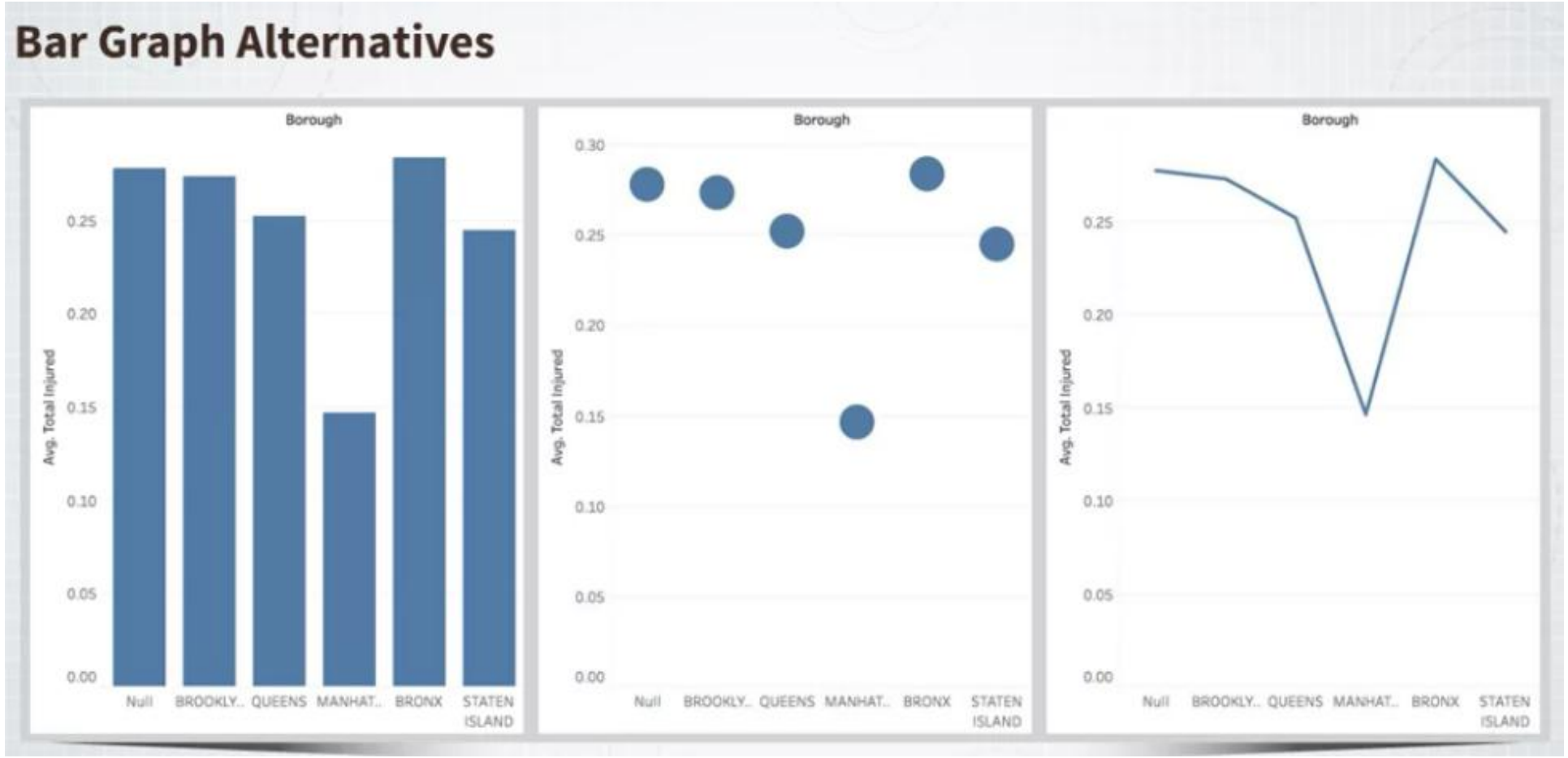

MAPPORN
Map Porn, for interesting maps
JOIN
r/MapPorn

r/MapPorn - Posted by u/Tranquil-Golden-Hour 3 hours ago
4.7k
Average 'First Leaf' Appearance in the contiguous U.S.
Average 'First Leaf' Appearance
March
February
January
April 13
The New York Times | USA National Phenology Network
120 Comments
Share
Save
Hide
Report
97% Upvoted

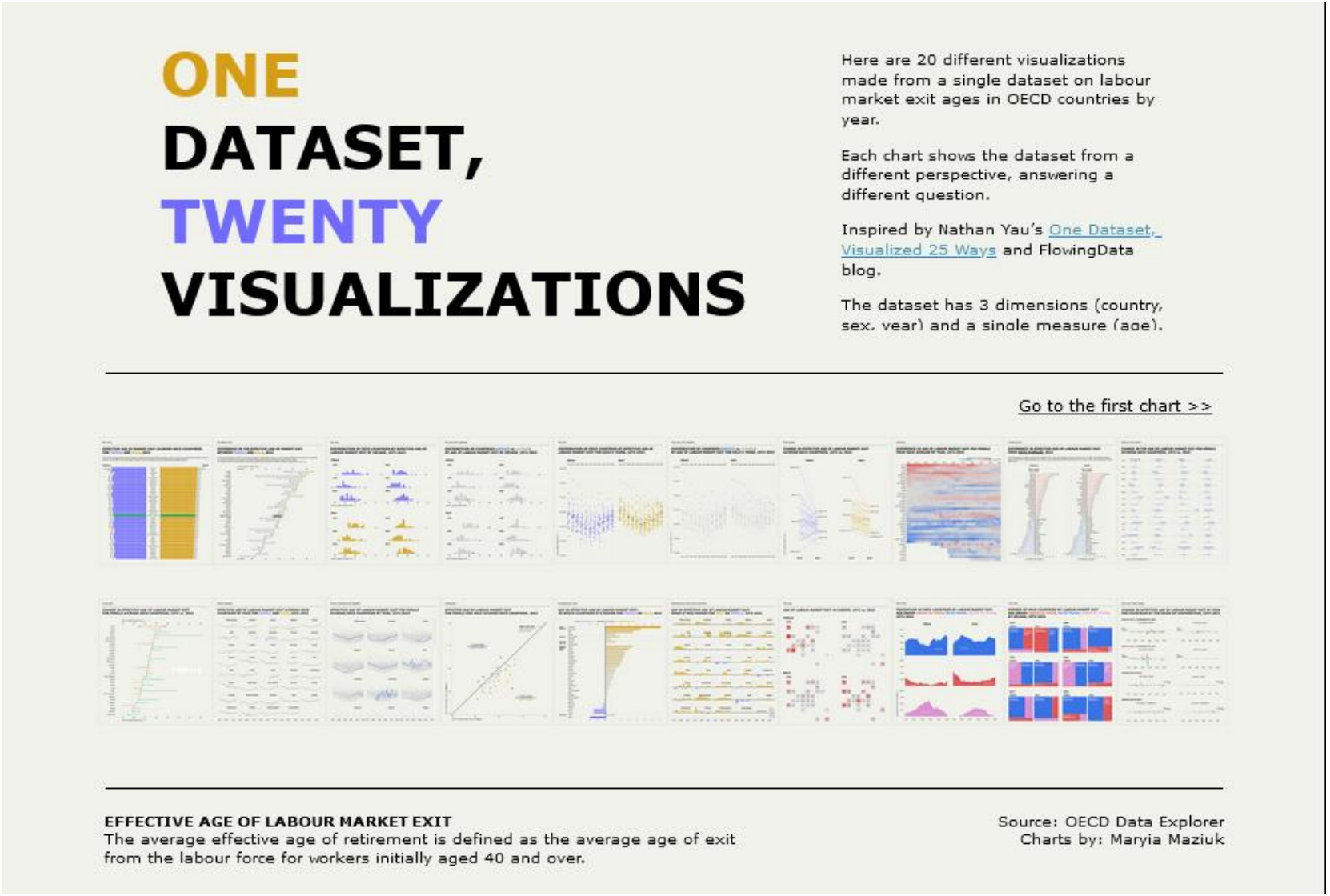
Summary of most common graphs



Summary of most common graphs



One dataset, 20 visualizations



Ways to visualize ‘time’

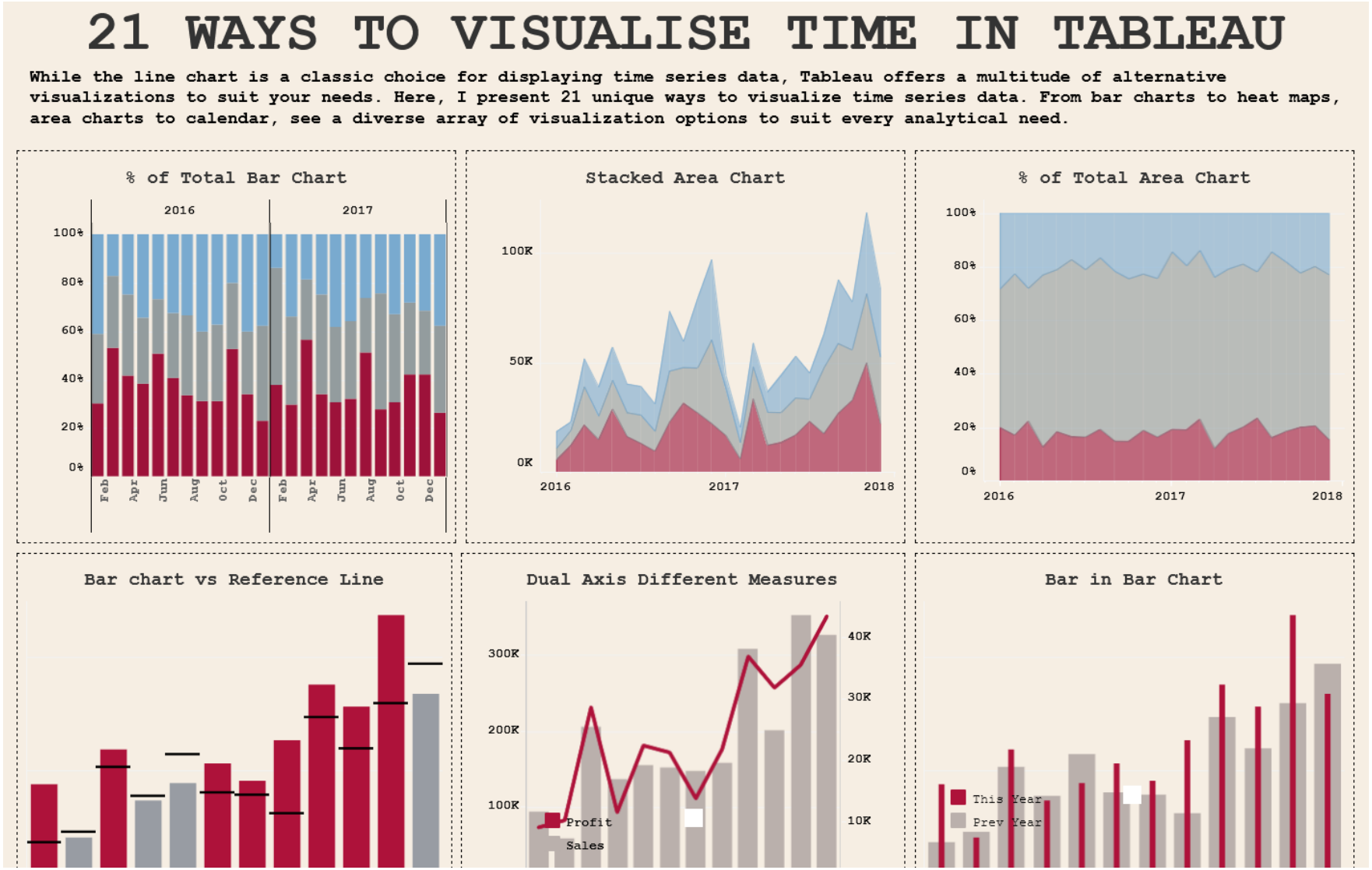
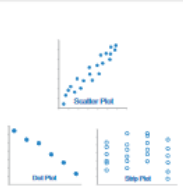
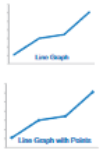
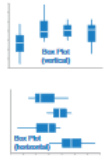


Chart (graph) selection matrix

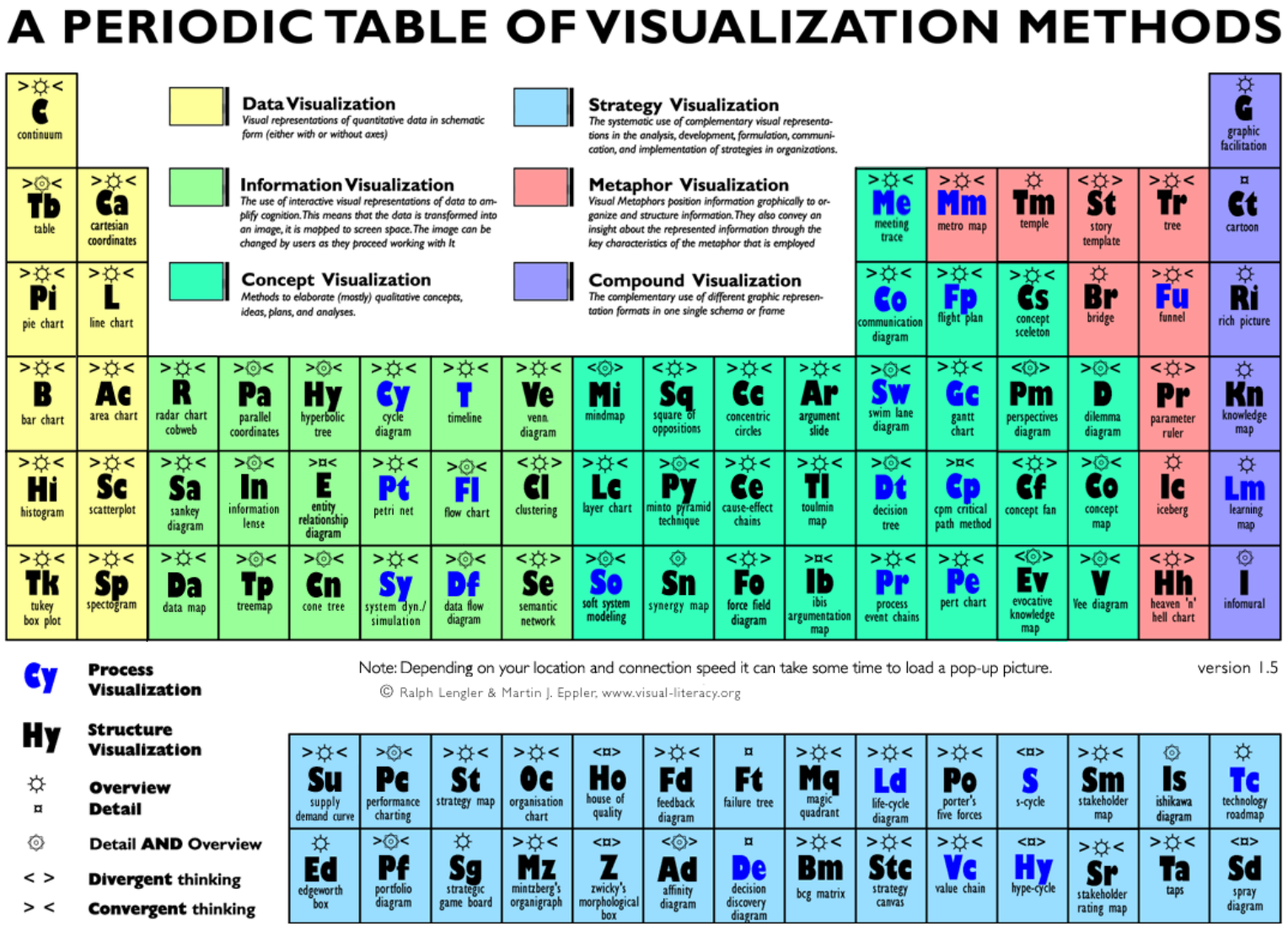
Graph Selection Matrix

Featured Relationships	Value-Encoding Objects: Points	Lines	Bars	Boxes
				
Time Series Values display how something changed through time (yearly, monthly, etc.)	**Sometimes** (As a *dot plot*, if values occur at irregular intervals of time)	**Often** (To feature overall trends and patterns and to support their comparisons)	**Sometimes** (Vertical bars only, to feature individual values and to support their comparisons)	**Sometimes** (Vertical boxes only, to display how a distribution changes through time)
Ranking Values are ordered by size (descending or ascending)	**Sometimes** (As a *dot plot*, especially when the quantitative scale does not begin at zero)	**Sometimes** (As a *bumps chart*, to show how rankings change through time)	**Often**	**Sometimes** (To display a ranked set of distributions)
Part-to-Whole Values represent parts (proportions) of a whole (for example, regional portions of total sales)	**Sometimes** (Especially when the visual weight of bars creates excessive clutter)	**Sometimes** (To display how parts of a whole have changed through time)	**Often**	**Sometimes** (When displaying typical part-to-whole values along with the ranges across which they vary)
Deviation The difference between two sets of values (for example, the variance between actual and budgeted expenses)	**Sometimes** (As a *dot plot*, especially when the quantitative scale does not begin at zero)	**Often** (When also featuring a time series)	**Often**	**Sometimes** (When displaying typical deviation values along with the ranges across which they vary)
Distribution Counts of values per interval from lowest to highest (for example, counts of people by age intervals of 10 years each)	**Often** (As a *strip plot*, to feature individual values in one or more distributions)	**Often** (As a *frequency polygon*, to feature the overall shape of one or more distributions)	**Often** (As a *histogram*, especially when displaying a single distribution)	**Often** (Only when comparing multiple distributions)
Correlation Comparison of two or more sets of values to determine if there is a relationship between them	**Often** (As a *scatter plot*, when displaying the relationship between two sets of values)	**Sometimes** (Only when summarizing correlations as lines of best fit without displaying individual correlation values)	**Sometimes** (As a *table lens*, especially when your audience is not familiar with *scatter plots*)	**Never**
Geospatial Values are located in space (e.g., on a map) to show their location	**Often** (As bubbles of various sizes on a map)	**Sometimes** (To display routes on a map)	**Sometimes** (Only when there is enough space for bars to vary enough in length for easy comparisons)	**Never** (Although you might want to show distributions using boxes on a map, it isn't practical)
Nominal Comparison A simple comparison of unordered discrete values	**Sometimes** (As a *dot plot*, especially when the quantitative scale does not begin at zero)	**Never**	**Often**	**Never** (A nominal comparison consists solely of discrete values along a nominal scale)

Check out

- The Data Visualisation Catalogue: https://datavizcatalogue.com/
- The R-graph gallery: https://www.r-graph-gallery.com/
- D3 examples: https://observablehq.com/@d3/gallery
- Google charts: https://developers.google.com/chart/interactive/docs/gallery
- PolicyViz: https://policyviz.com/
- Visual Vocabulary – Vega: https://gramener.github.io/visual-vocabulary-vega/

Check out

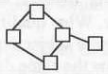

A PERIODIC TABLE OF VISUALIZATION METHODS
Data Visualization
Visual representations of quantitative data in schematic form (either with or without axes)
Strategy Visualization
The systematic use of complementary visual representations in the analysis, development, formulation, communication, and implementation of strategies in organizations.
Information Visualization
The use of interactive visual representations of data to amplify cognition. This means that the data is transformed into an image, it is mapped to screen space. The image can be changed by users as they proceed working with it
Metaphor Visualization
Visual Metaphors position information graphically to organize and structure information. They also convey an insight about the represented information through the key characteristics of the metaphor that is employed
Concept Visualization
Methods to elaborate (mostly) qualitative concepts, ideas, plans, and analyses.
Compound Visualization
The complementary use of different graphic representation formats in one single schema or frame
C continuum
G graphic facilitation
Tb table
Ca cartesian coordinates
Me meeting trace
Mm metro map
Tm temple
St story template
Tr tree
Ct cartoon
Pi pie chart
L line chart
Co communication diagram
Fp flight plan
Cs concept sceleton
Br bridge
Fu funnel
Ri rich picture
B bar chart
Ac area chart
R radar chart cobweb
Pa parallel coordinates
Hy hyperbolic tree
Cy cycle diagram
T timeline
Ve venn diagram
Mi mindmap
Sq square of oppositions
Cc concentric circles
Ar argument slide
Sw swim lane diagram
Gc gantt chart
Pm perspectives diagram
D dilemma diagram
Pr parameter ruler
Kn knowledge map
Hi histogram
Sc scatterplot
Sa sankey diagram
In information lense
E entity relationship diagram
Pt petri net
Fl flow chart
Cl clustering
Le layer chart
Py minto pyramid technique
Ce cause-effect chains
Tl toulmin map
Dt decision tree
Cp cpm critical path method
Cf concept fan
Co concept map
Ic iceberg
Lm learning map
Tk tukey box plot
Sp spectogram
Da data map
Tp treemap
Cn cone tree
Sy system dyn./ simulation
Df data flow diagram
Se semantic network
So soft system modeling
Sn synergy map
Fo force field diagram
Ib ibis argumentation map
Pr process event chains
Pe pert chart
Ev evocative knowledge map
V Vee diagram
Hh heaven 'n' hell chart
I infomural
Cy Process Visualization
Hy Structure Visualization
Overview
Detail
Detail AND Overview
< > Divergent thinking
> < Convergent thinking
Note: Depending on your location and connection speed it can take some time to load a pop-up picture.
version 1.5
© Ralph Lengler & Martin J. Eppler, www.visual-literacy.org
Su supply demand curve
Pe performance charting
St strategy map
Oc organisation chart
Ho house of quality
Fd feedback diagram
Ft failure tree
Mq magic quadrant
Ld life-cycle diagram
Po porter's five forces
S s-cycle
Sm stakeholder map
Is ishikawa diagram
Tc technology roadmap
Ed edgeworth box
Pf portfolio diagram
Sg strategic game board
Mz mintzberg's organigraph
Z zwicky's morphological box
Ad affinity diagram
De decision discovery diagram
Bm bcg matrix
Stc strategy canvas
Vc value chain
Hy hype-cycle
Sr stakeholder rating map
Ta taps
Sd spray diagram

Further reading

- Harris, R. (2000). *Information Graphics - A Comprehensive Illustrated Reference*. Oxford University Press.

Chart[1]

Sometimes referred to as an information graphic. A chart is a vehicle for consolidating and displaying information for purposes such as analysis, planning, monitoring, communicating, etc. Previously, charts were tangible things such as single sheets of paper, display boards, or flip charts. Today, many charts are generated and displayed electronically so that no hard copy ever exists. This book discusses charts without regard to the medium used. Instead, it discusses them in terms of content, function, format, etc. • There are five major categories of charts: graphs, maps, diagrams, tables, and other (those charts that do not fit into one of the other four categories). Each major category is broken into multiple subcategories, some of which are noted in the table below. • All individual information graphics can be included in multiple subcategories depending on the criteria that is used, such as shape, format, function, type of scales, type of data displayed, usage, number of axes, etc. For example, a widely used graph for plotting the distribution of data elements in a data set might be called as a histogram, data distribution graph, joined column graph, column graph, two-axis graph, two-dimensional graph, rectangular graph, quantitative graph, graph, or chart. All 10 terms are perfectly correct.

		Major Categories of Charts				
		Graph (plot)	Map	Diagram	Table	Other Charts
						
Key features		Quantitative patterns and comparisons	Spatial and directional relationships	Nonquantitative interrelationships	Preciseness of information and ease of reference	Differs depending on specific chart
Primary Function		Examples are shown below				
Quantitative	Shows patterns and/or relationships of quantitative data at point in time	Area graph Bar graph Circle graph Column graph Line graph Nomograph Polar graph Radar graph Scatter graph Trilinear graph	Contour map Demographic map Distorted map Elevation map Prism map Shaded map Smooth statistical map Weather map		Analytical table Bidirectional table General table Quantitative table Reference table Spreadsheet	Icon comparison display Pie chart Proportional chart Ranking chart Unit chart Venn diagram
Quantitative	Shows patterns and/or relationships of quantitative data over time	Area graph Candlestick chart Cash flow graph Column graph Control chart Index graph Line graph Run chart			Analytical table General table Quantitative table Reference table Spreadsheet	Comparative chart Icon comparison display Point & figure chart Proportional chart Stair chart Unit chart
Quantitative & Nonquantitative	Shows how/where things are distributed or located	100% graph Border plot Box plot Histogram Pareto graph Population pyramid Probability graph Quantile graph	Block map Blot or patch map Dot density map Geological map Pin map Profile map Topographic map Weather map	Block diagram Network diagram Voroni diagram	Analytical table Frequency table General table Percent table Quantitative table Reference table Spreadsheet	Business matrix Conceptual chart Floor plan
Nonquantitative	Relates time and activities			CPM chart PERT chart Time line	Calendar Time table	Explanatory chart Gantt chart Loading chart Milestone chart Scheduling chart
Nonquantitative	Shows how nonquantitative things are organized, arranged, interconnected or interrelated		Network map Ray map Road map Strip map Thematic map	Schematic: Block diagram Cause & effect diag. Conceptual diagram Flow chart Network diagram Organization chart Relational diagram Venn diagram	General table Pictorial table Reference table	Conceptual chart Dendrogram Distribution channel chart Exploded diagram Gantt & milestone charts Minimum spanning Process chart Structure diagram
Nonquantitative	Shows how nonquantitative things proceed		Flow map Weather map	Schematic: Conceptual diagram Decision chart Flow chart PERT & CPM charts Process chart Tree chart	Analytical table General table Reference table Spreadsheet	Conceptual chart Gantt & milestone charts Illustration chart Process chart Vector chart
Nonquantitative	Shows how nonquantitative things evolve or work			Conceptual diagram Flow chart Process chart	Analysis table General table Pictorial table Reference table	Conceptual chart Cross-section Exploded diagram Illustration chart Pictorial chart
Nonquantitative	Shows how to do things			Calculation chart "How to" diagram Procedural diagram		Calculation chart How-to chart Illustration chart Pictorial instruction

Thank you!

mroussou@di.uoa.gr

http://eclass.uoa.gr/courses/DI411/