



UNIVERSITY
OF SKÖVDE

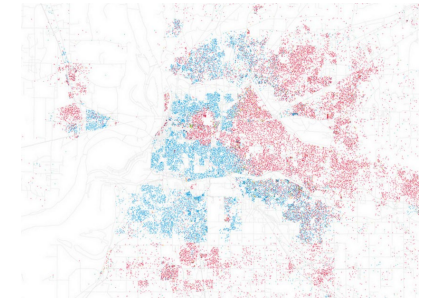
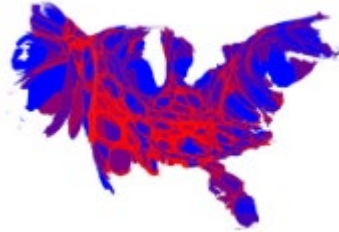
VISUAL DATA ANALYSIS

LECTURE 4
Information
Visualization (2)

The Most Common Type of Data Visualizations & Examples (referred to datalabs)

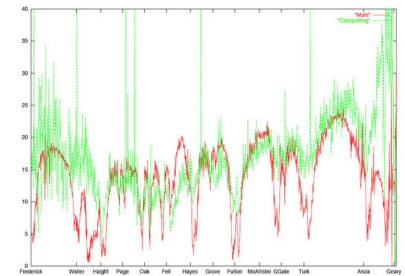
✓ 2D Area

- ✓ Cartogram
- ✓ Choropleth
- ✓ Dot Distribution Map



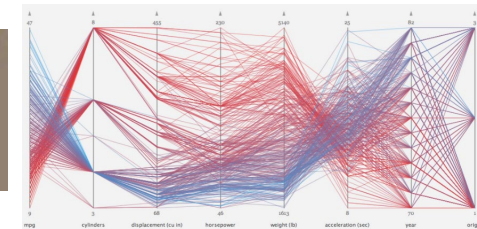
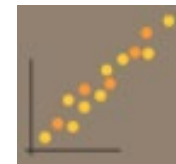
✓ Temporal

- ✓ Line chart
- ✓ Polar Area Diagram
- ✓ Time Series



✓ Multi-dimensional

- ✓ Pie Chart
- ✓ Histogram
- ✓ Scatter Plot
- ✓ Parallel Coordinates



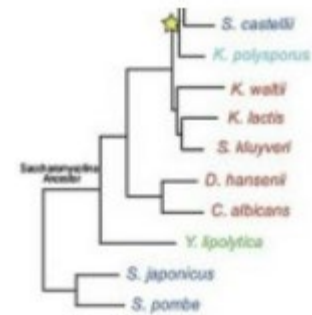
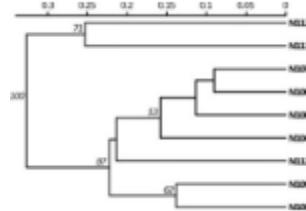
The Most Common Type of Data Visualizations & Examples (referred to datalabs)



UNIVERSITY
OF SKÖVDE

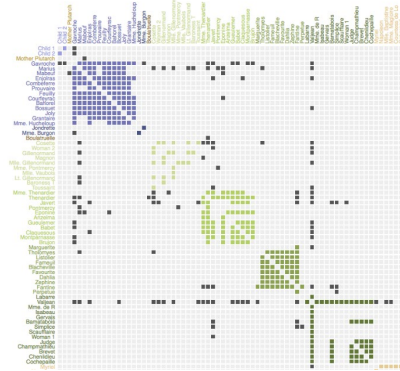
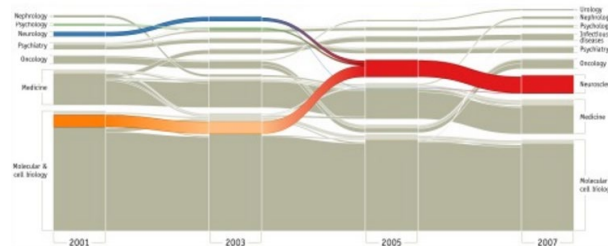
✓ Hierarchical

- ✓ Dendrogram
- ✓ Ring Chart
- ✓ Tree Diagram



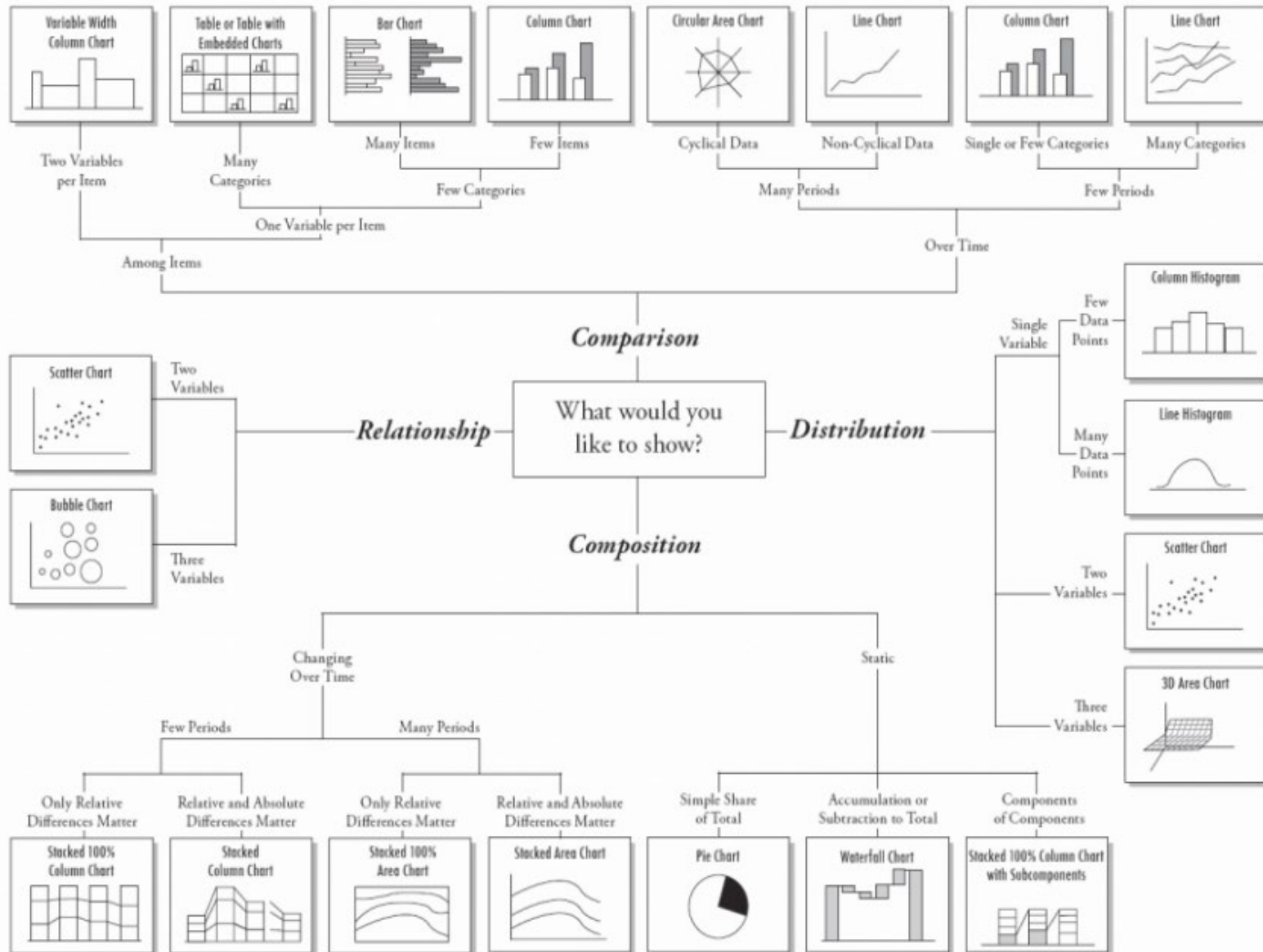
✓ Network

- ✓ Alluvial Diagram (sankey)
- ✓ Node-link Diagram
- ✓ Matrix

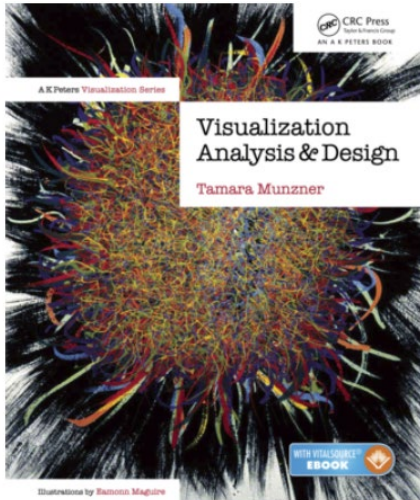


<http://www.datalabsagency.com/data-visualization-news/15-most-common-types-of-data-visualisation/>

The Most Common Type of Data Visualizations & Examples (referred to datalabs)

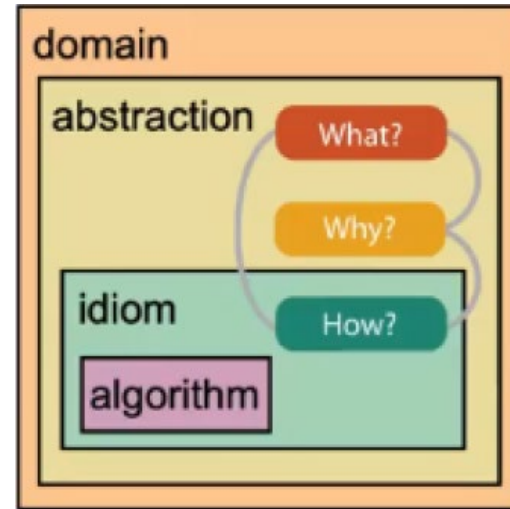


Reference book



@tamaramunzner

Tamara Munzner
Department of Computer Science
University of British Columbia



What:Data

Why:Task

How:Encoding

Domain situation

Who are the target users?

Abstraction

What is shown? **Data** abstraction

Why is the user looking at it? **Task** abstraction

Idiom

How is it shown?

Visual encoding idiom: **how** to **draw**

Interaction idiom: **how** to manipulate

Algorithm

Efficient computation

Outline

TYPES

- Data types
- Dataset types
- Attribute types
- Marks and Channels
- Actions

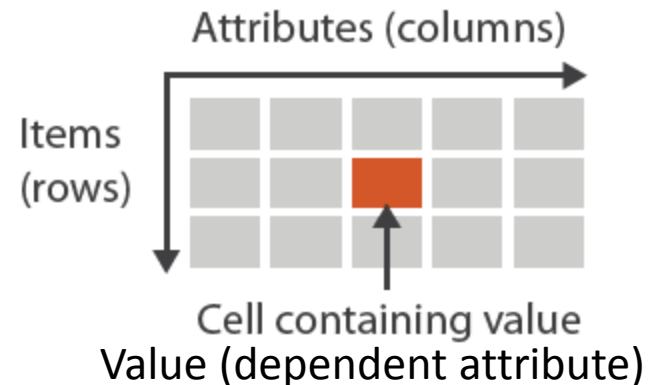
VISUAL ENCODING DESIGNS

- Bar chart
- Dot and line charts
- Stream graph
- Pie chart
- Scatter plot
- Parallel coordinates
- Heat map
- Choropleth map

Data types

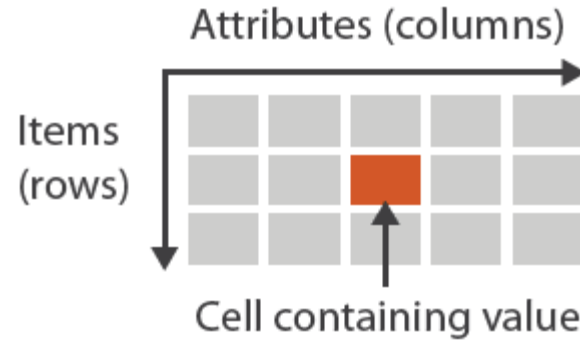
What:Data

- Attributes
 - A property measured, observed, and logged
 - columns, features, variables, key (independent attribute)
 - E.g.) salary, price, number of sales, temperature
- Items
 - An individual and discrete entity
 - rows, instances, observations, examples
 - E.g.) People, stocks, coffee shops, cities
- Link
 - A relationship between items
- Position
 - Spatial data or latitude-longitude pair
 - Location in a region

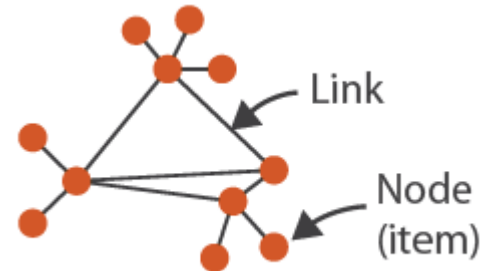


Three major dataset types

- Tables



- Networks and trees



- Geometry
 - Spatial data



Attribute types

- Categorical



- Ordered

- Ordering types
 - Ordinal



- Quantitative



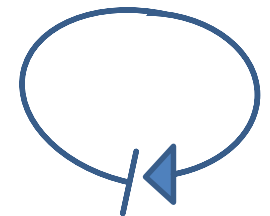
- Ordering direction
 - Sequential



- Diverging



- Cyclic



MARKS AND CHANNELS

How:Encoding

Marks: geometric primitives

- Points



- Lines



- Areas



Channels: control the mark's appearances

Magnitude channels, **ordered** attributes

Identity channels, **categorical** attributes

Position



Position, unaligned



Length



Tilt



Area



Depth, 3D



Color luminance



Color saturation



Curve



Volume, 3D



Best
Effectiveness
Least

Spatial region



Color hue



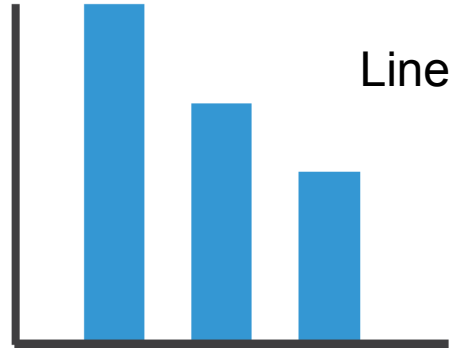
Motion



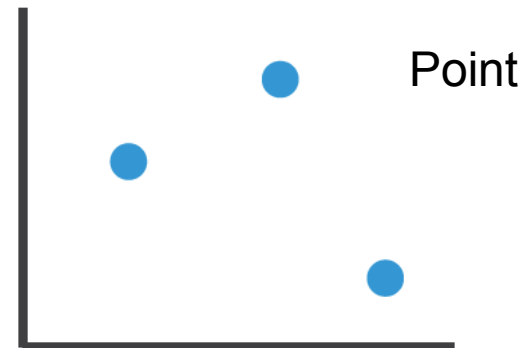
Shape



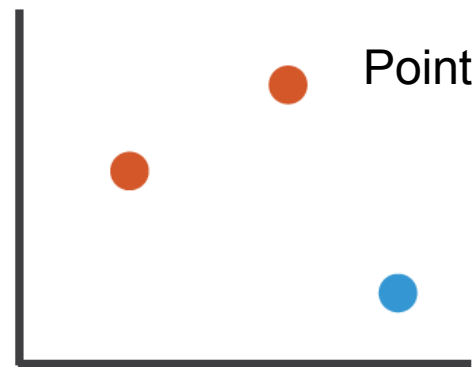
Channels and Marks: example



(a)
Vertical position



(b)
Vertical position
Horizontal position



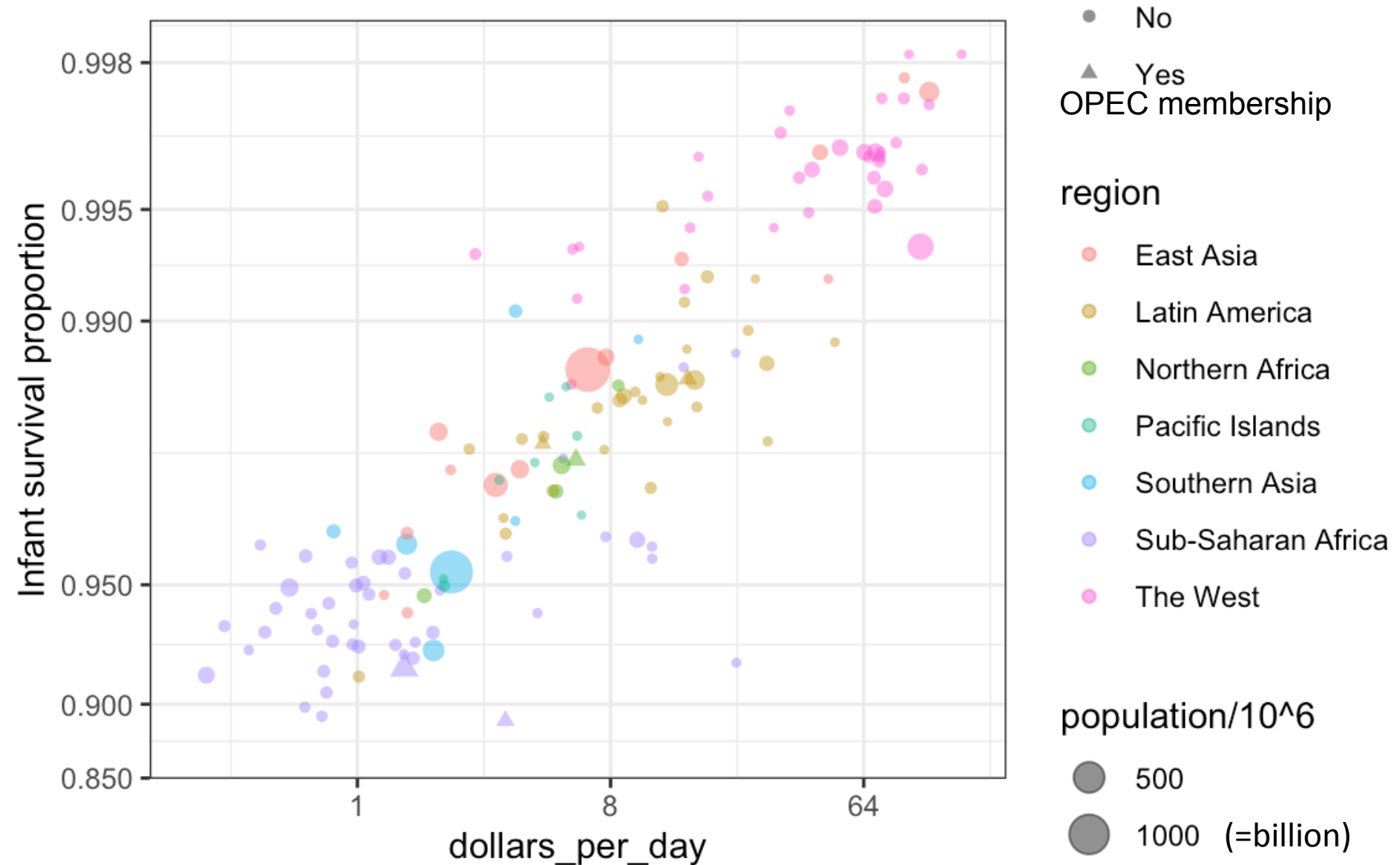
(c)
Vertical position
Horizontal position
Color hue



(d)
Vertical position
Horizontal position
Color hue, area (size)

Channels: example

- Relationship between infant survival and average income



Multiple marks and channels for grouping

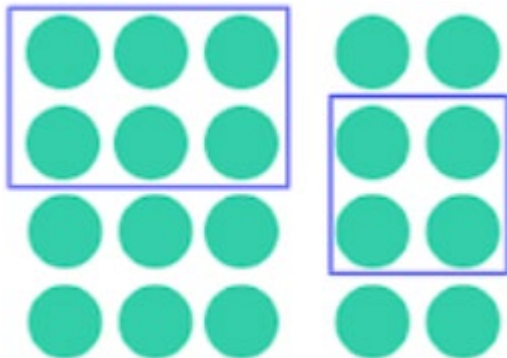
- Grouping the elements based on ...



Similarity



Similarity



Common region, proximity

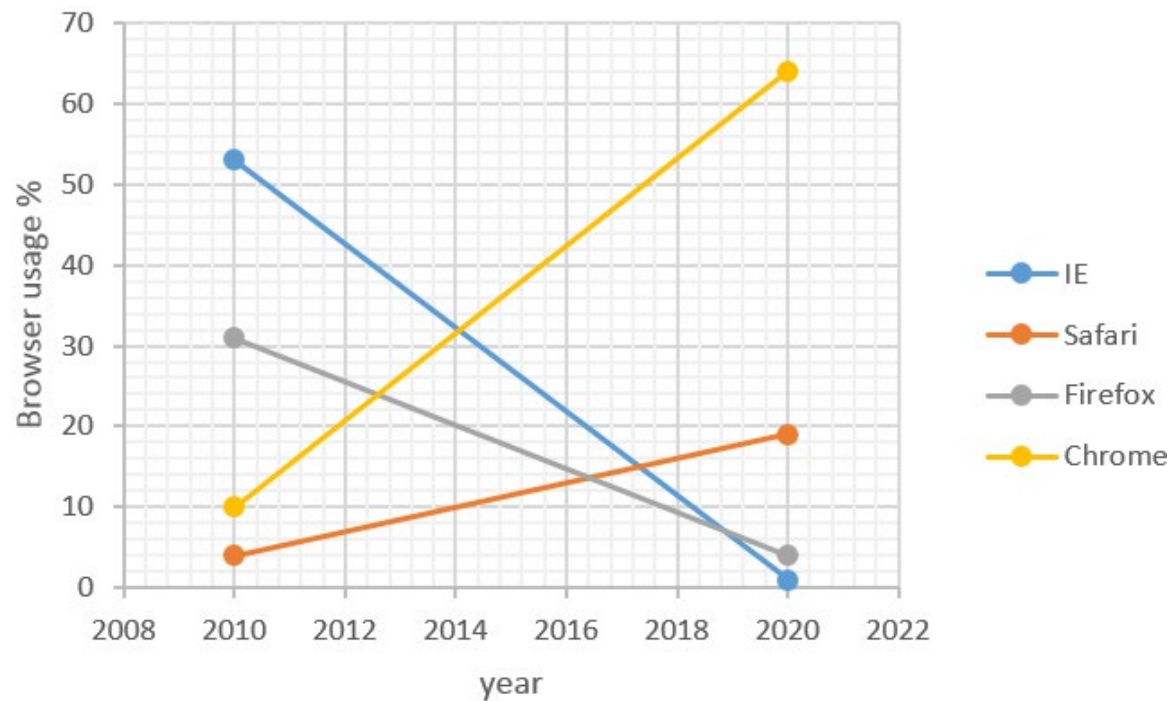
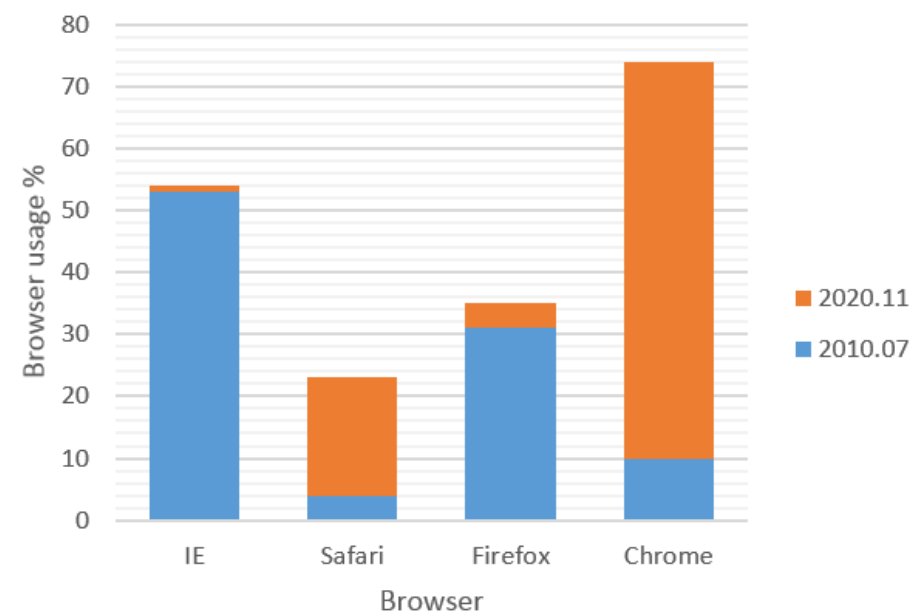
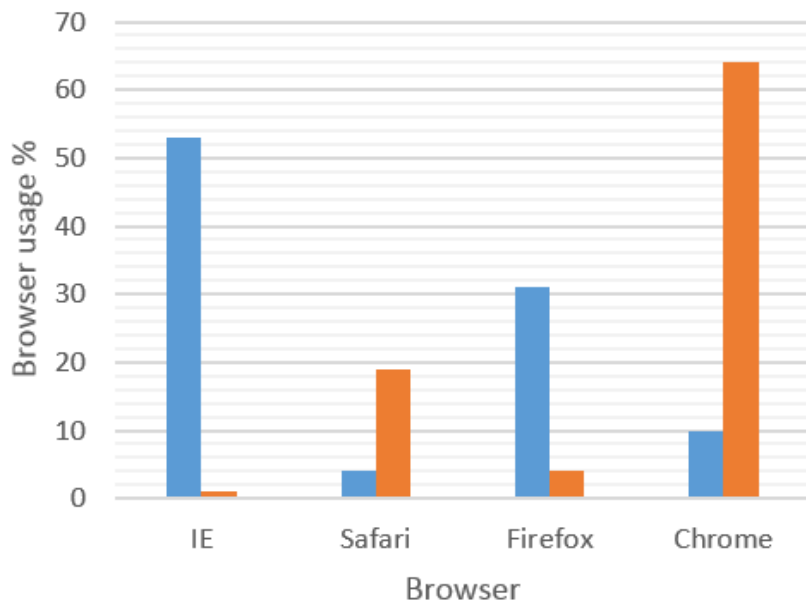


Proximity, color hue


Marks and Channels: Exercise


	2010 July	2020 Nov
Internet Explorer	53	1
Safari	4	19
Firefox	31	4
Chrome	10	64


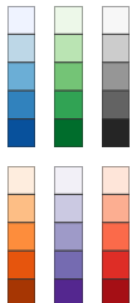
Browser usage in % in 2010 and 2020 (StatCounter)





Selecting a proper color: ColorBrewer 2.0



Number of data classes:  [how to use](#) [updates](#) [downloads](#) [credits](#)










Nature of your data: 
☒ sequential ☐ diverging ☐ qualitative

Pick a color scheme:
Multi-hue:  Single hue: 

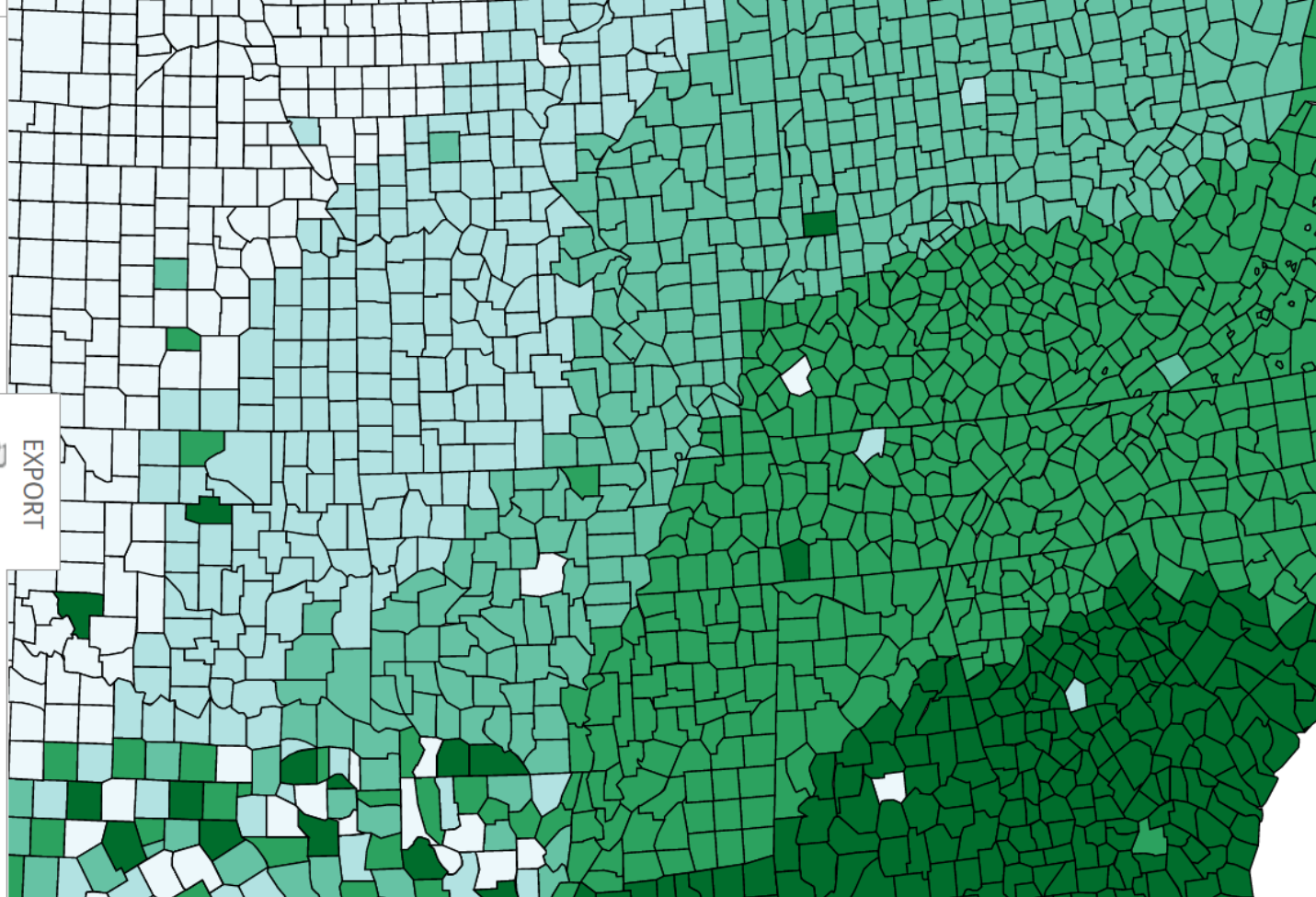
Only show: 
☒ colorblind safe
☐ print friendly
☐ photocopy safe

Context: 
☐ roads
☐ cities
☒ borders

Background:
☒ solid color 
☐ terrain

color transparency

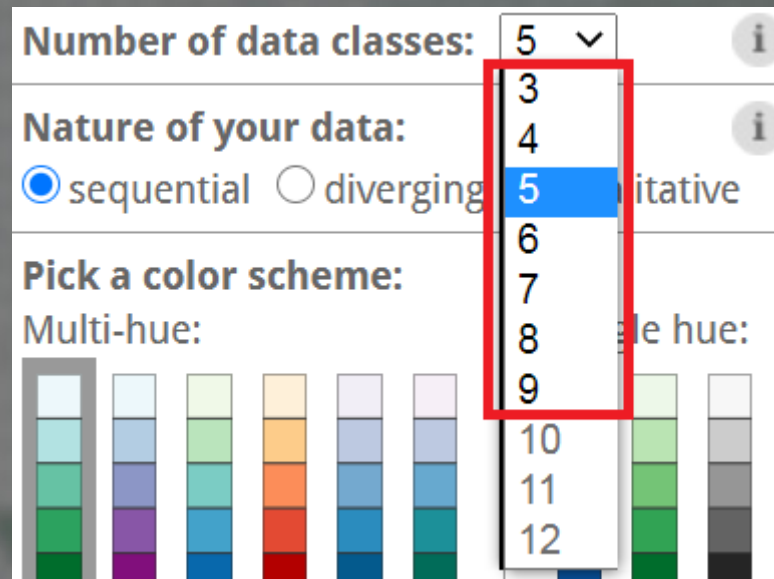
5-class BuGn    
HEX
 #edf8fb
 #b2e2e2
 #66c2a4
 #2ca25f
 #006d2c

EXPORT



COLORBREWER 2.0
color advice for cartography

Selecting a proper color: ColorBrewer 2.0



The screenshot shows the ColorBrewer 2.0 web interface. The 'Number of data classes' dropdown menu is open, displaying a list of numbers from 3 to 12. The number 5 is highlighted in blue. A red rectangle is drawn around the dropdown menu. The background shows various color scheme options and a map of the United States.

Number of data classes: 5 ▼

Nature of your data:
☒ sequential ☐ diverging ☐ qualitative

Pick a color scheme:
Multi-hue:

3
4
5
6
7
8
9
10
11
12

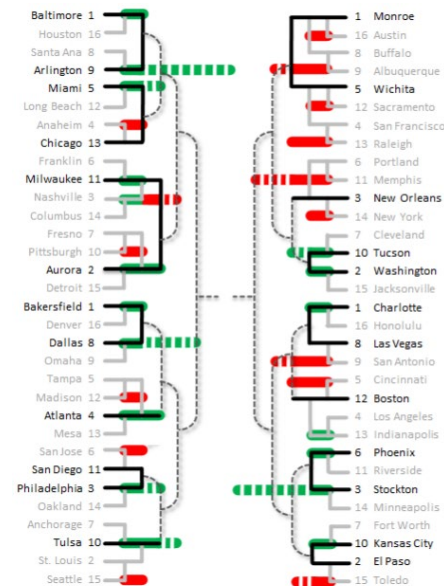
ACTIONS

Why:Task

Action: Analyze

Action

- Consume
 - Discover (explore)
 - Present (explain)
 - Enjoy (social)
- Produce



Tan D, Smith G, Lee B, Robertson G. AdaptiviTree: adaptive tree visualization for tournament-style brackets. IEEE Trans Vis Comput Graph. 2007 Nov-Dec;13(6):1113-20. doi: 10.1109/TVCG.2007.70537. PMID: 17968054




Images from 'Visualization Analysis and Design', Tamara Munzner, 2014

Action: Analyze

- | | Action |
|-----------|----------------------|
| • Consume | • Discover (explore) |
| | • Present (explain) |
| | • Enjoy (social) |
| • Produce | • Annotate |
| | • Record |
| | • Derive |



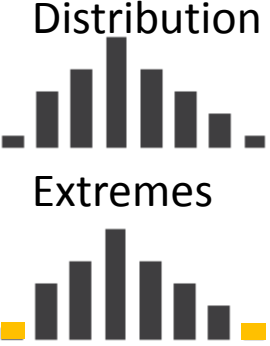
Action: Analyze

	Action	Target
• Consume	• Discover (explore)	• Trends (patterns) 
	• Present (explain)	• Outliers 
	• Enjoy (social)	• Features (structure) 
• Produce	• Annotate	• Attributes
	• Record	• Network, topology
	• Derive	• Spatial data

Action: Analyze

	Action	Target
• Consume	• Discover (explore)	• Trends
	• Present (explain)	• Outliers
	• Enjoy (social)	• Features
• Produce	• Annotate	• Attributes
	• Record	• Network, topology
	• Derive	• Spatial data

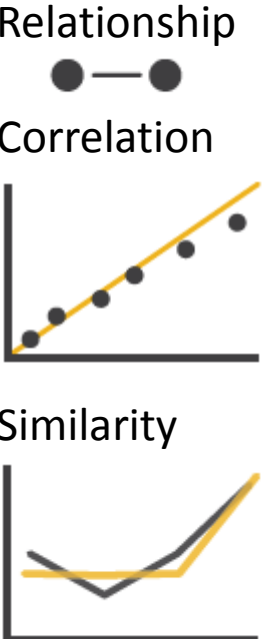
One



Distribution

Extremes

Many




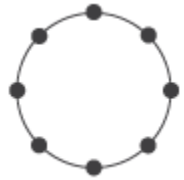




Relationship

Correlation

Similarity

Action: Analyze

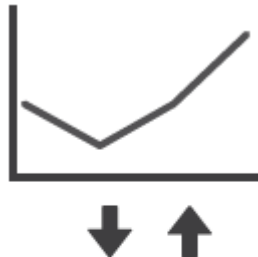
	Action	Target
• Consume	• Discover (explore)	• Trends
	• Present (explain)	• Outliers
	• Enjoy (social)	• Features
• Produce	• Annotate	• Attributes <ul style="list-style-type: none"> • One • Many
		• Network <ul style="list-style-type: none"> • Topology
	• Record	   
	• Derive	• Paths 
		• Spatial data 

Action: Query

- Identify



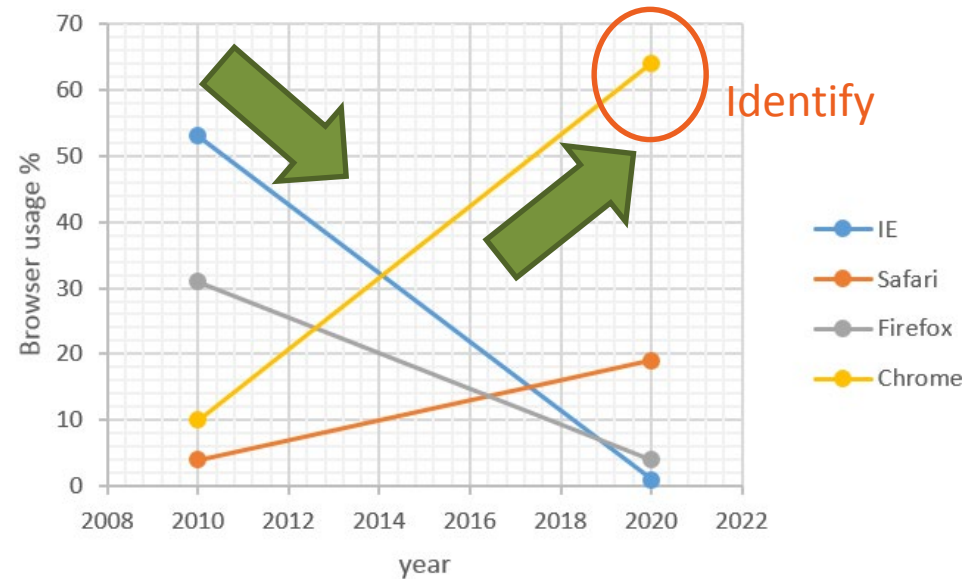
- Compare



- Summarize



Compare decrease vs. increase



TARGET

- Trends
- Outliers
- Feature or structure
- Distribution or range
- Dependency, correlation, similarity
- Topology, paths of network data
- Entire shape of spatial data

SUMMARY

MARKS

- points
- lines
- areas

CHANNELS

- position
- length
- tilt
- area
- depth
- color luminance (brightness or darkness of color), saturation (intensity of color)
- curve
- 3D
- color hue, shape (categorical)

SUMMARY

ACTION

consume - discover (explore)

- present (explain)

- enjoy

produce - annotate

- record

- derive

query - identify

- compare

- summarize

TARGET

- trends

- outliers

- feature or structure

- distribution or range

- dependency, correlation, similarity

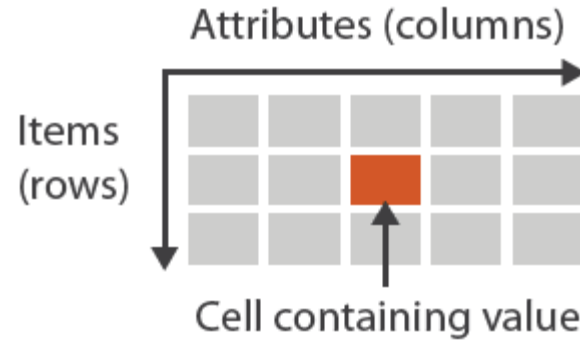
- topology, paths of network data

- entire shape of spatial data

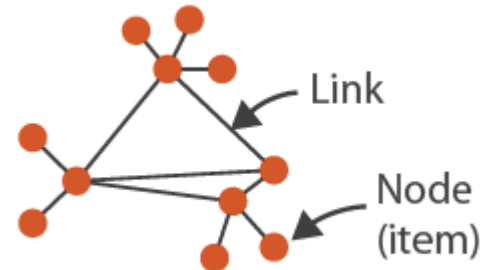
BREAK

Reminder: Three major dataset types

- Tables



- Networks and trees



- Geometry, Spatial



Reminder: Marks and channels

MARKS

- points
- lines
- areas

CHANNELS

Position



Position, unaligned



Length



Tilt



Area



Depth, 3D



Color luminance



Color saturation



Curve



Volume, 3D



Best

Effectiveness

Least

Same

Same

VISUAL ENCODING DESIGNS

- Tables
 - BAR CHART
 - DOT AND LINE CHARTS
 - STREAM GRAPH
 - PIE CHART
 - SCATTERPLOT
 - PARALLEL COORDINATES
 - HEATMAP
- Networks and trees
- Spatial
 - CHOROPLETH MAP

Bar Chart

Search and compare

Why:Task

- Key – value What:Data

- Mark: line

- Channel: position or length

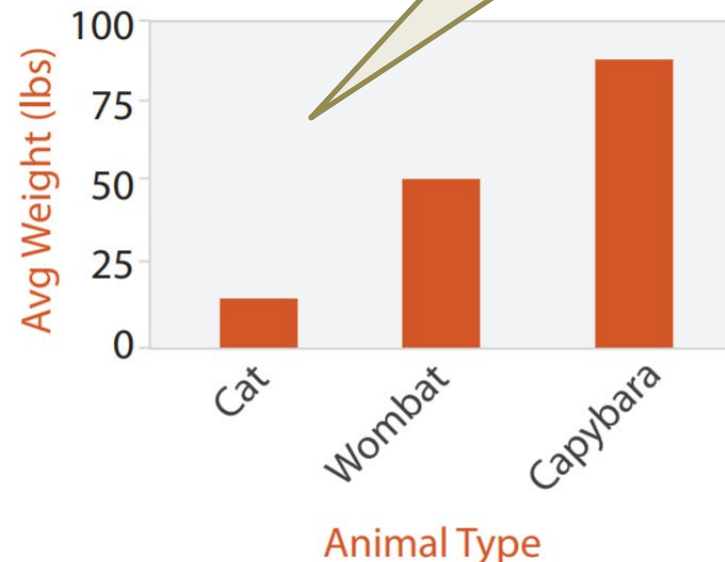
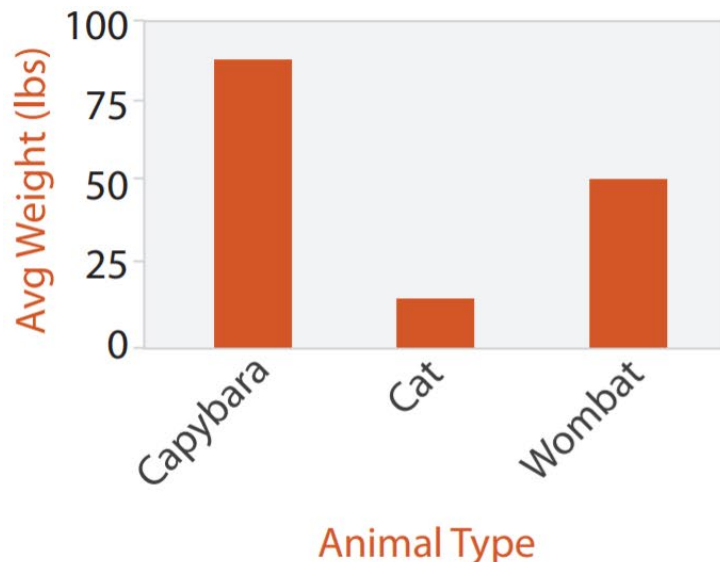
How:Encoding

+) easy to read (compare endpoints of the bar)

*) consider the scale, have a zero baseline

*) for categorical key attributes, not numerical

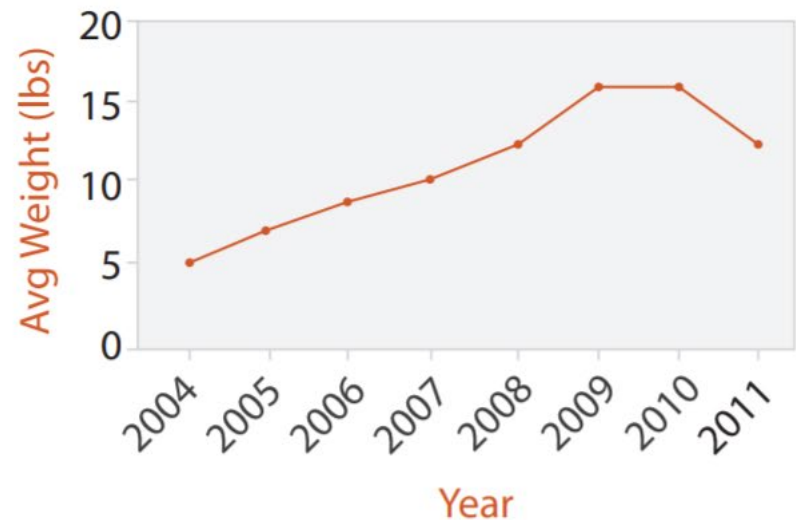
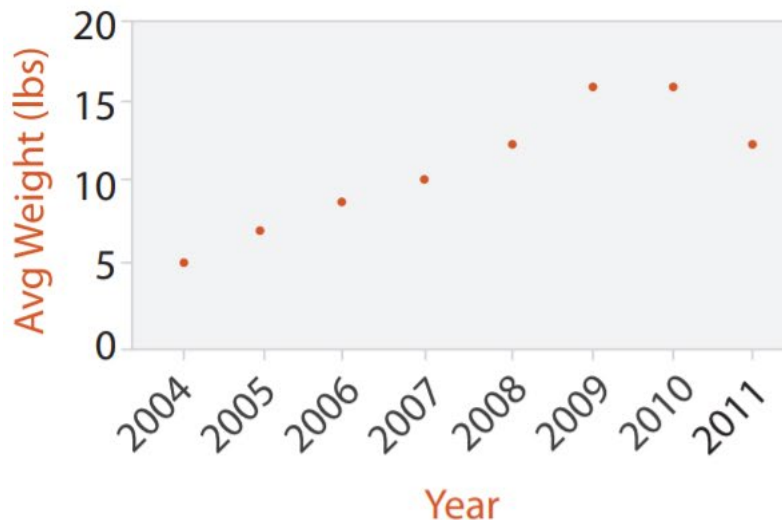
*) show legend, unit of value (labels if needed)



Dot and Line Charts

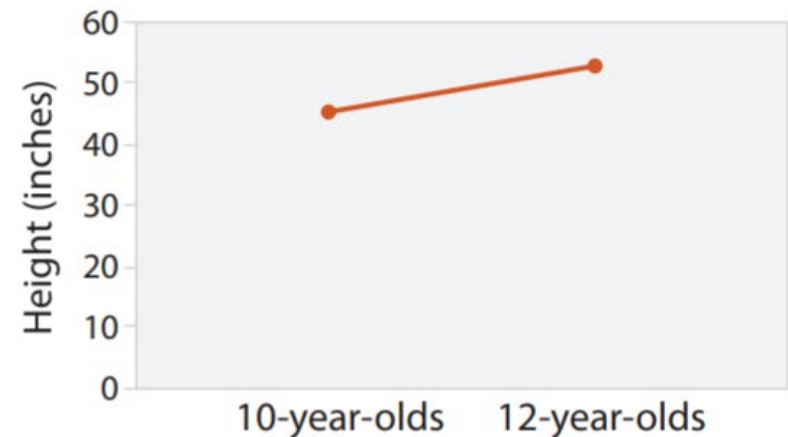
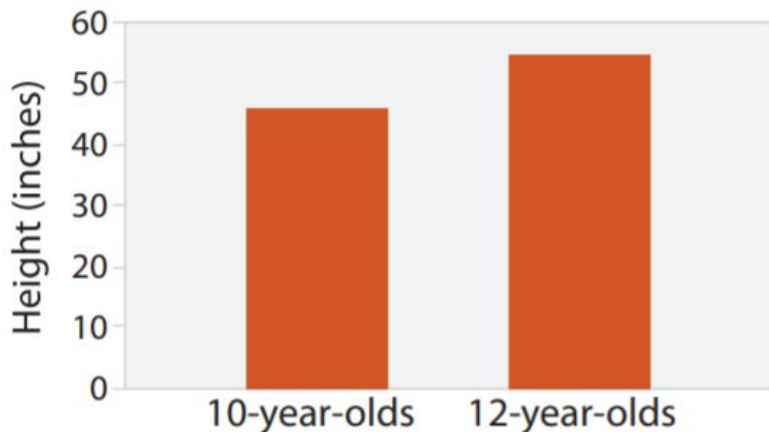
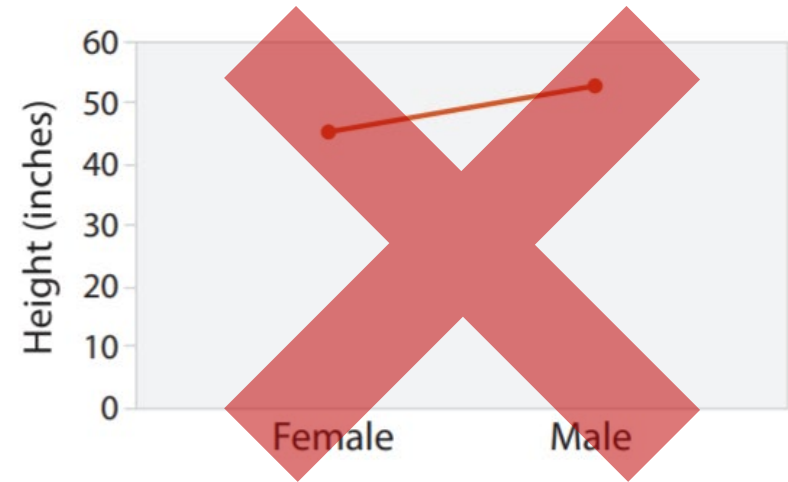
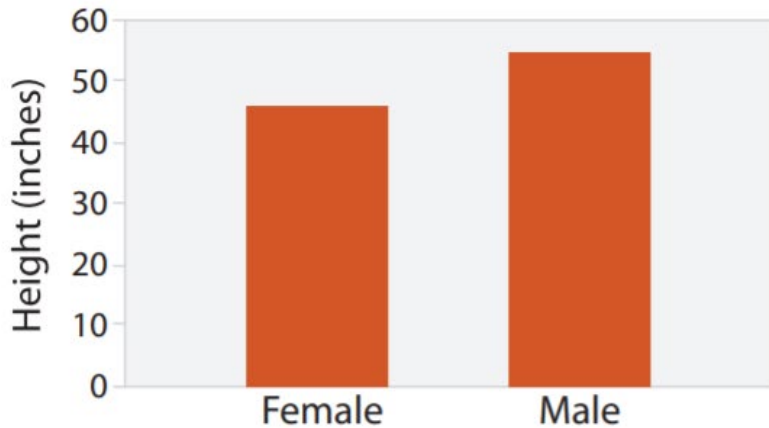
Compare, show trends **Why:Task**

- Ordered Key – value **What:Data**
- Mark: point and line **How:Encoding**
- Channel: position
- Shows the changes over time by connecting each point in the series
- +) easy to read
- *) use for continuous data
- *) show a limited number of datapoints
- *) show unit, legend or label



Dot and Line Charts

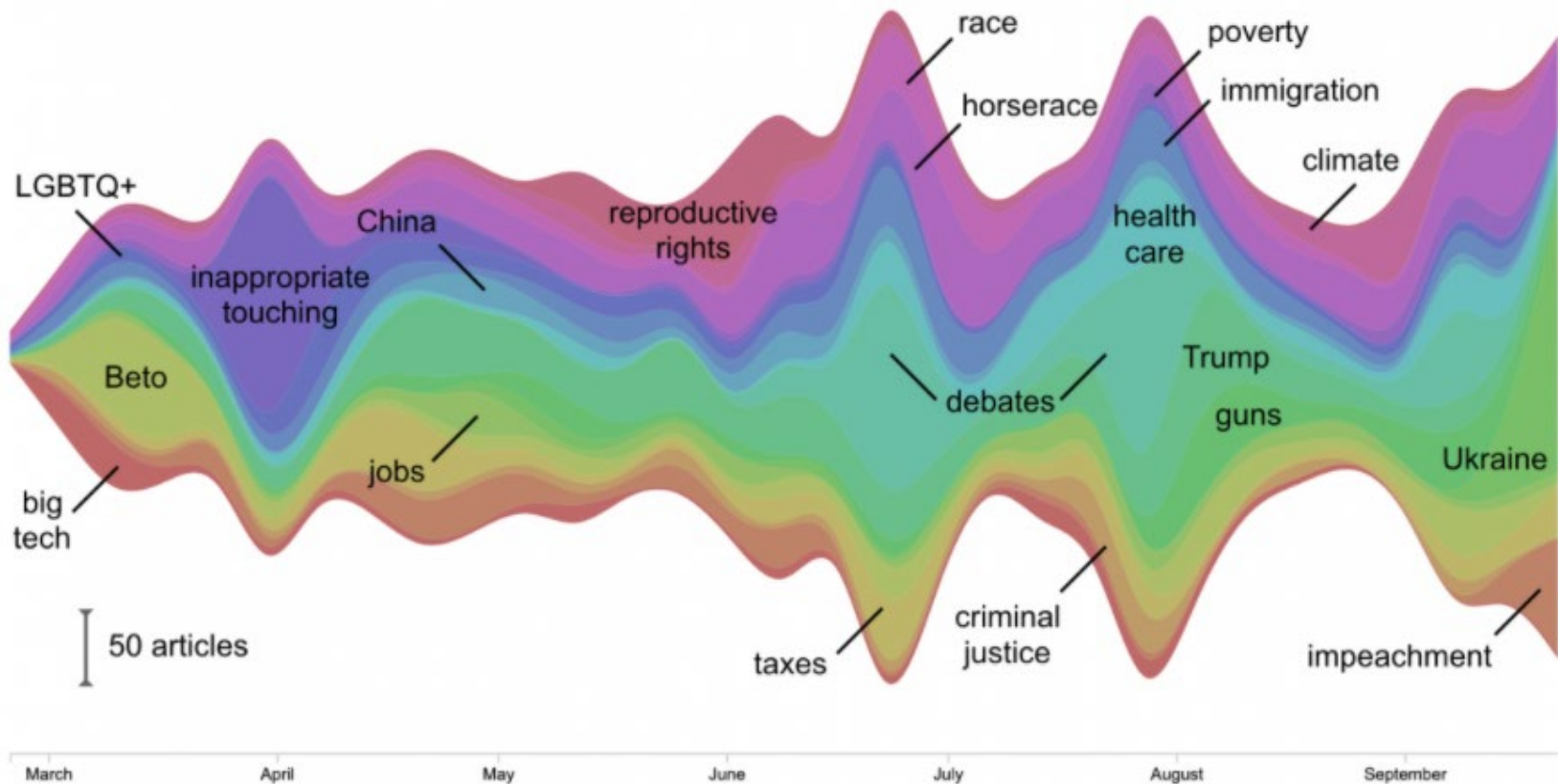
- Misleading implication, line charts are not for categorical data



Stream graph

How media are setting the 2020 agenda

A topic analysis of news articles published by 28 outlets since March 2019 mentioning Joe Biden, Bernie Sanders, Elizabeth Warren, Kamala Harris, Pete Buttigieg, Beto O'Rourke, Cory Booker, Kirsten Gillibrand, Amy Klobuchar, or Tulsi Gabbard



$n = 5,850$

Data: Media Cloud

<https://www.storybench.org/how-news-media-are-setting-the-2020-election-agenda-chasing-daily-controversies-often-burying-policy/>

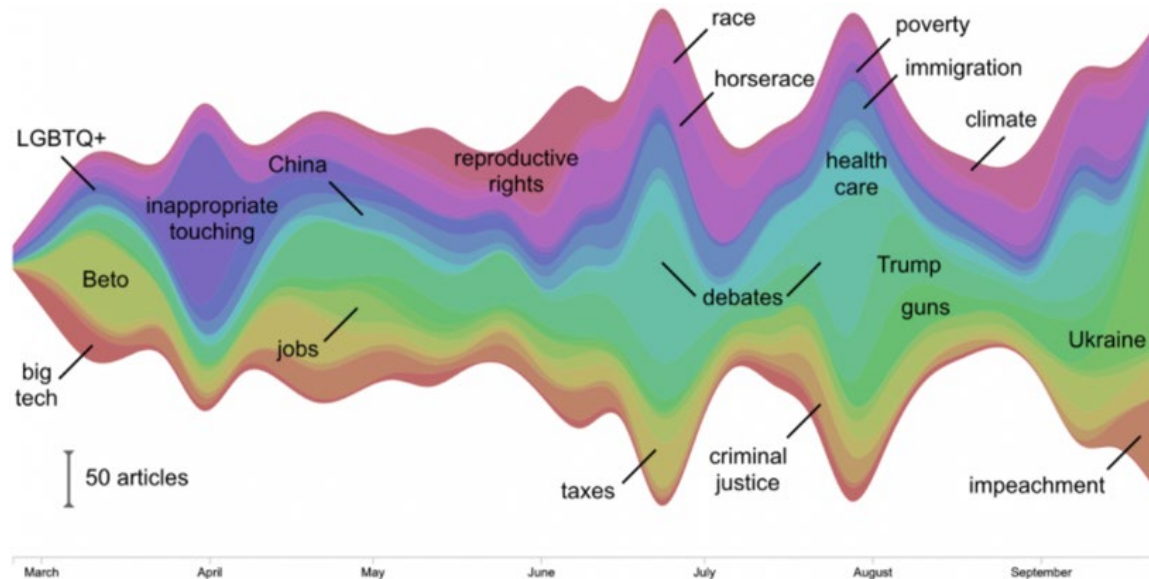
Stream graph

Trends, compare, derive

- Ordered key (time) – categorical key (topic)- value (counts)
- Mark: layers across time (or area)
- Channel: height (of layers)
- Shows the changes over time using color for different categories
- Each stream for proportional change over time

*) show a limited number of datapoints

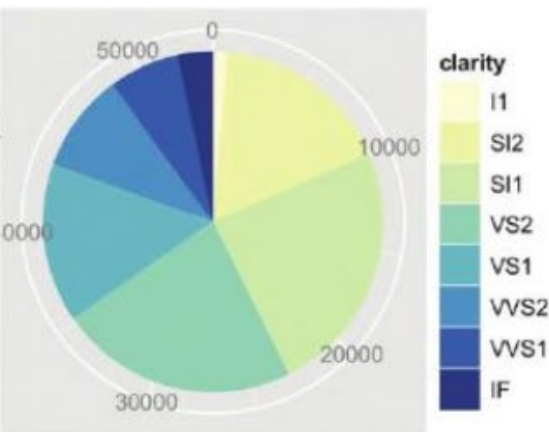
*) show legend or label



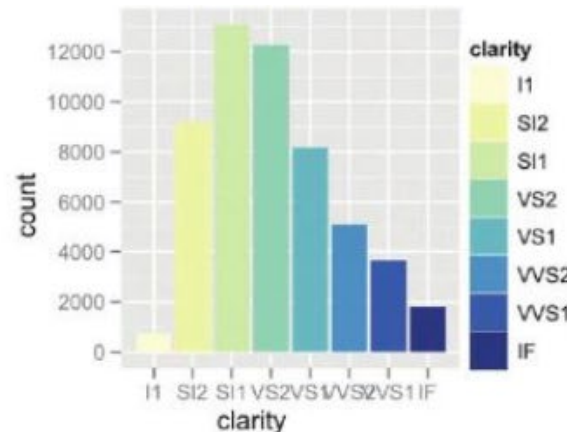
Pie Chart

Proportion (part-whole relationship)

- Key - value
- Mark: area
- Channel: angle (& color for easier legibility)
-) avoid if slices are similar size
-) limit to lesser than 8 slices
- *) use whitespace between slices
- *) angle on area is less accurate than length on lines



(a) Pie chart



(b) Bar chart



(c) Polar area chart

REMINDER: Marks and channels

MARKS

- points
- lines
- areas

CHANNELS

Position



Position, unaligned



Length



Tilt



Area



Depth, 3D



Color luminance



Color saturation



Curve



Volume, 3D



Best

Effectiveness

Least

Same

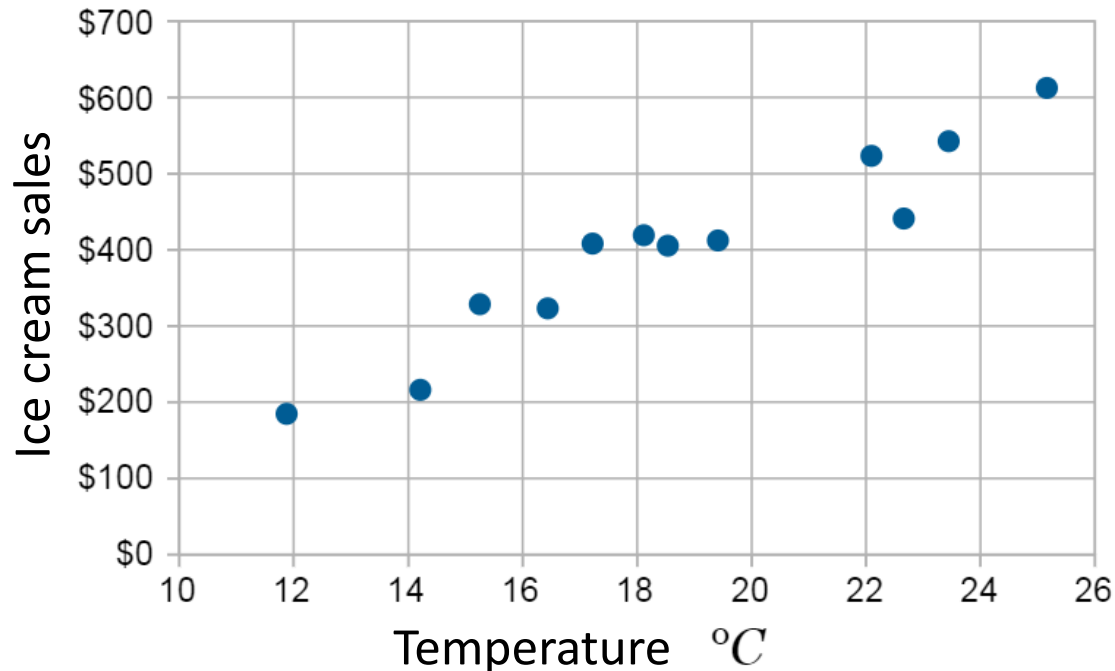
Same

Scatterplot

Find trends, outliers, distribution, correlation, locate clusters

- Value - value
- Mark: point
- Channel: vertical and horizontal positions

*) stronger correlation fall along a perfect diagonal line

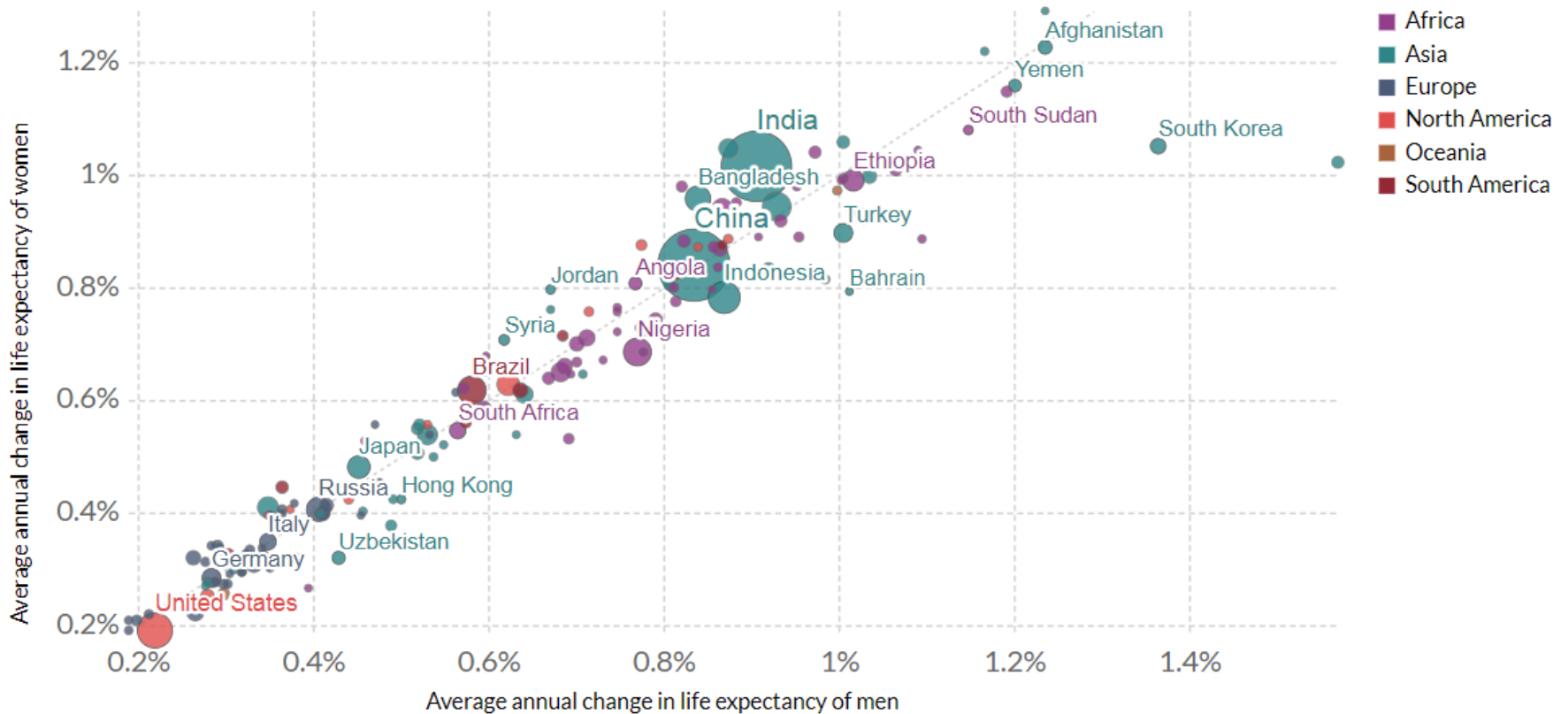


Bubble plot

Life expectancy of women vs life expectancy of men, 1950 to 2020

In countries that lie above the grey line the life expectancy of women is higher than for men.

 Select countries ☒ Average annual change ☒ Hide countries < 1 million people



Source: UN Population Division (2019 Revision)
Note: Shown is the period life expectancy at birth measured in years.

CC BY

▶ 1950  2020

Parallel coordinates

Find trends and clusters, outliers, extremes, relationship, correlation

- Categorical keys - many values
- Mark: lines crossing through the axes (e.g., each line is a student)
- Channel: vertical position for multiple values and horizontal position for separate axes

-) learning time

-) how to order the axes

+) scalable

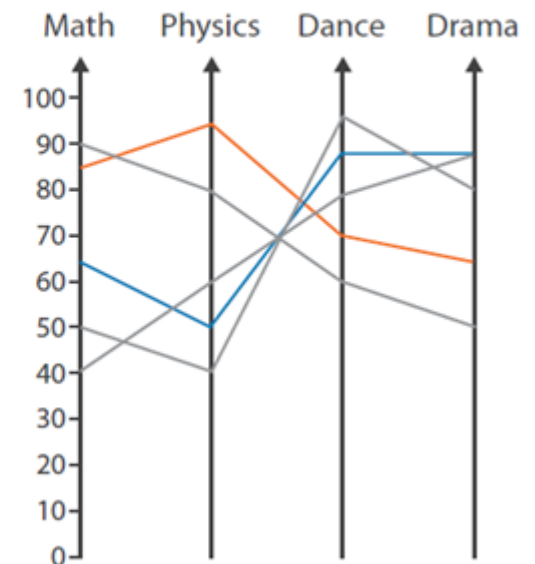
*) normalized values

*) better with interaction,
highlight the lines with
brushing to focus on
selected series

Table

Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

Parallel Coordinates

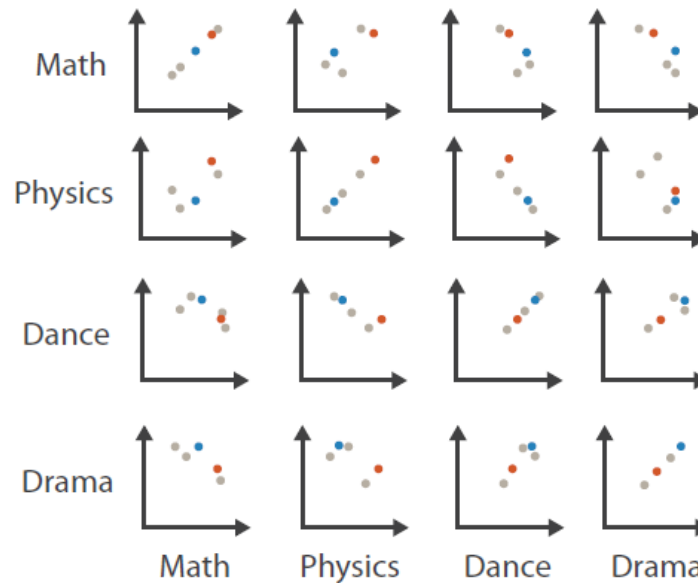


Scatterplot Matrix vs. Parallel coordinates

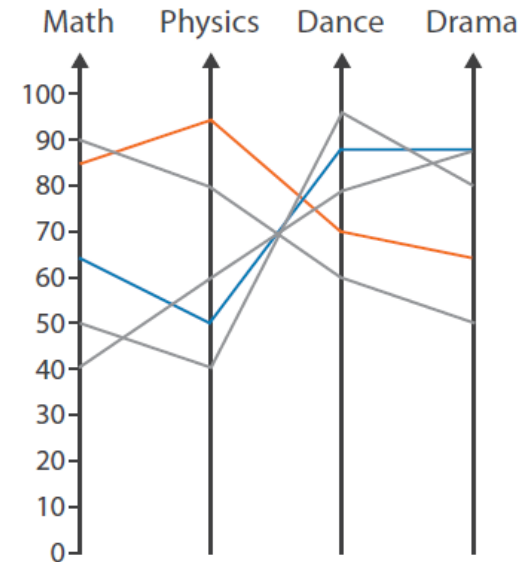
Table

Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90

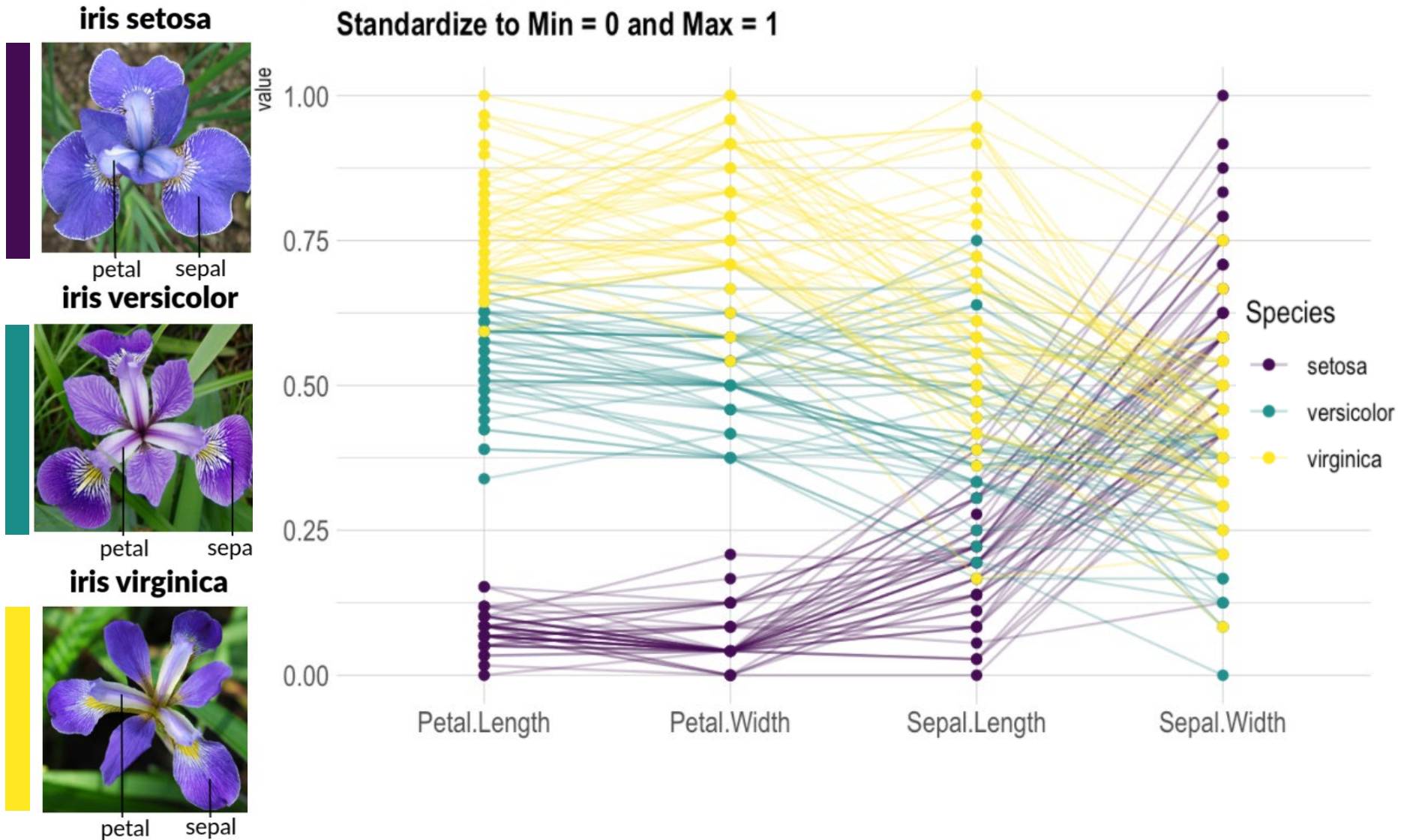
Scatterplot Matrix



Parallel Coordinates



Parallel coordinates



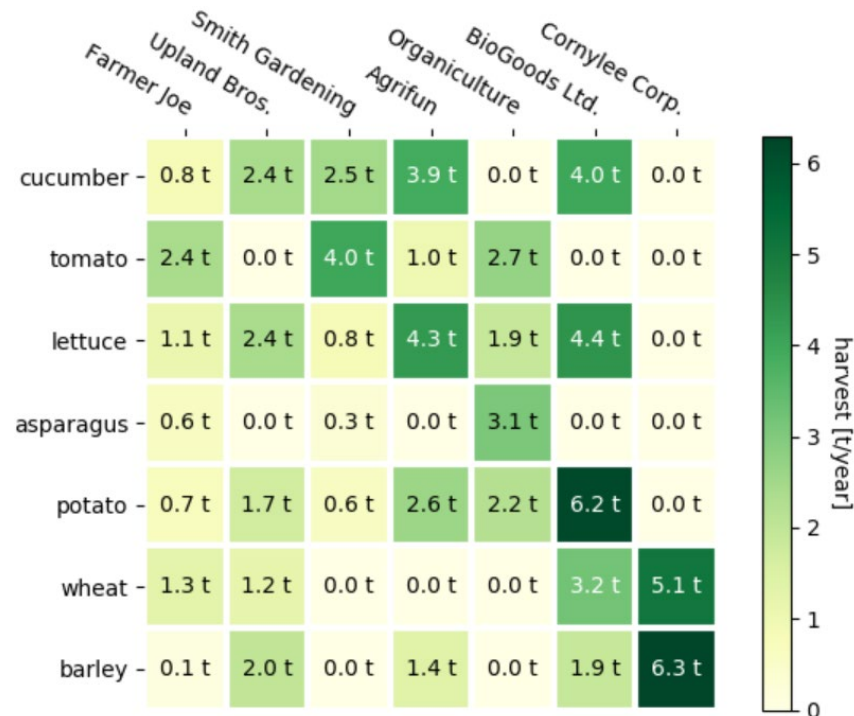
Heat map

Find clusters (patterns), outliers, summarize

- Two keys - value
- Mark: area in 2D matrix alignment
- Channel: color

+) scalable

*) normalized values or same scale



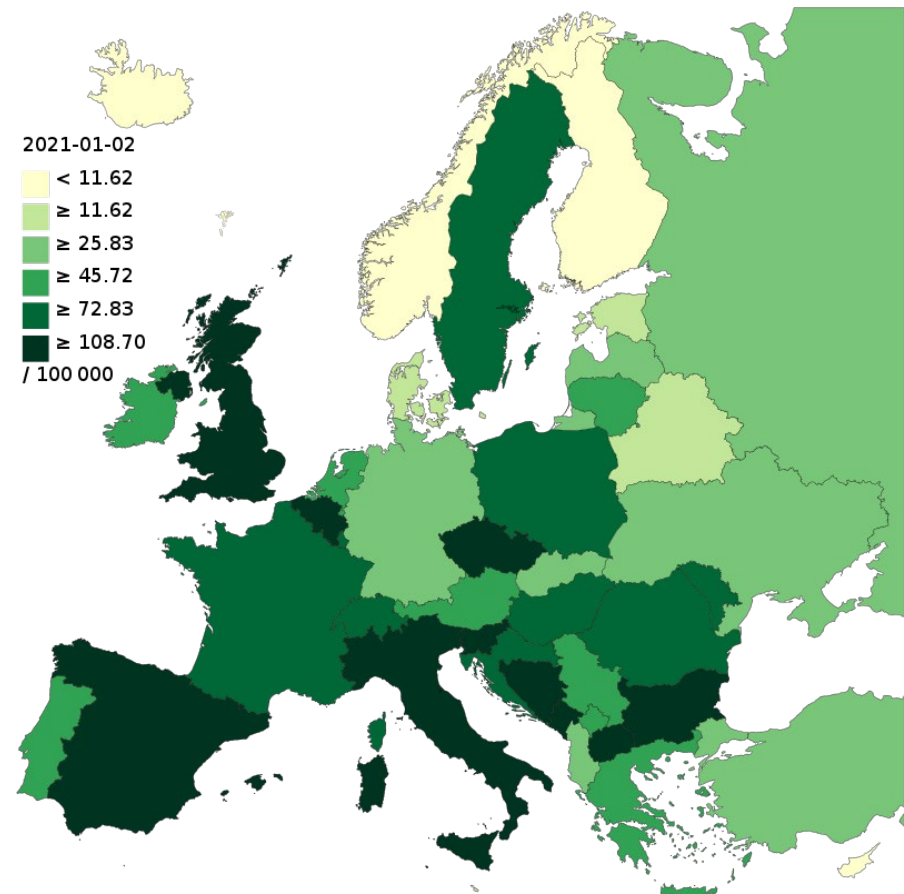
Choropleth map

Find patterns, outliers, summarize

- Geographic data. A value for each region.
- Mark: given geometry for area
- Channel: color

*) Consider how to construct the color map, which region boundaries to use

*) Can compare different regions
e.g.) continent, country, state,
territory, zip code



Deaths per 100,000 residents as of 2 January 2021

Summary

TYPES

- Data types
 - Dataset types
 - Attribute types
 - Marks and Channels
 - Actions
- What:Data
- How:Encoding
- Why:Task

VISUAL ENCODING DESIGNS

- Bar chart
- Dot and line charts
- Stream graph
- Pie chart
- Scatter plot
- Parallel coordinates
- Heat map
- Choropleth map

