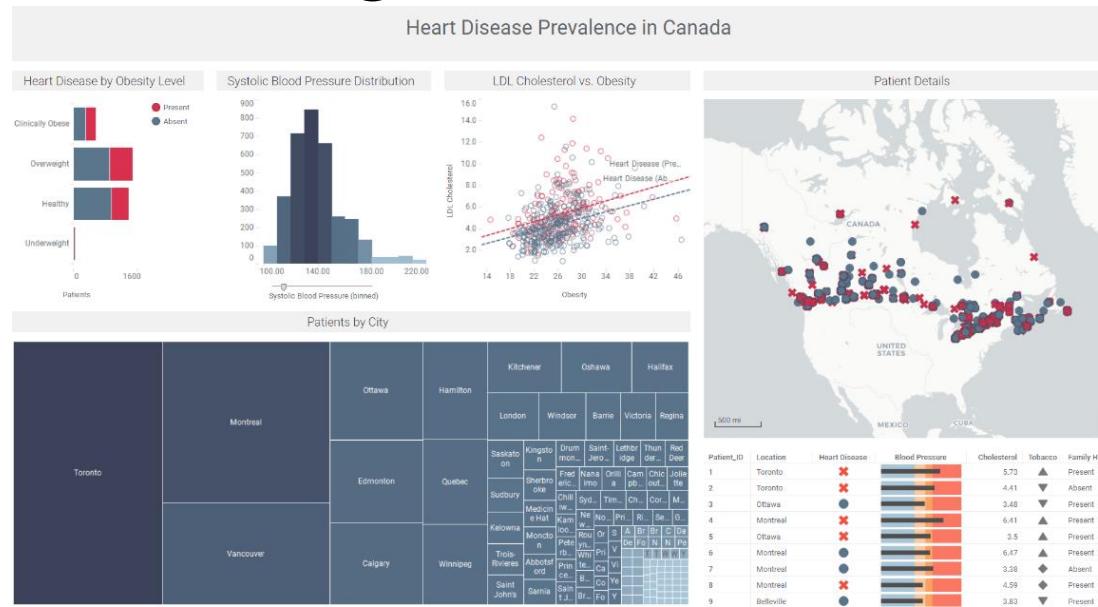




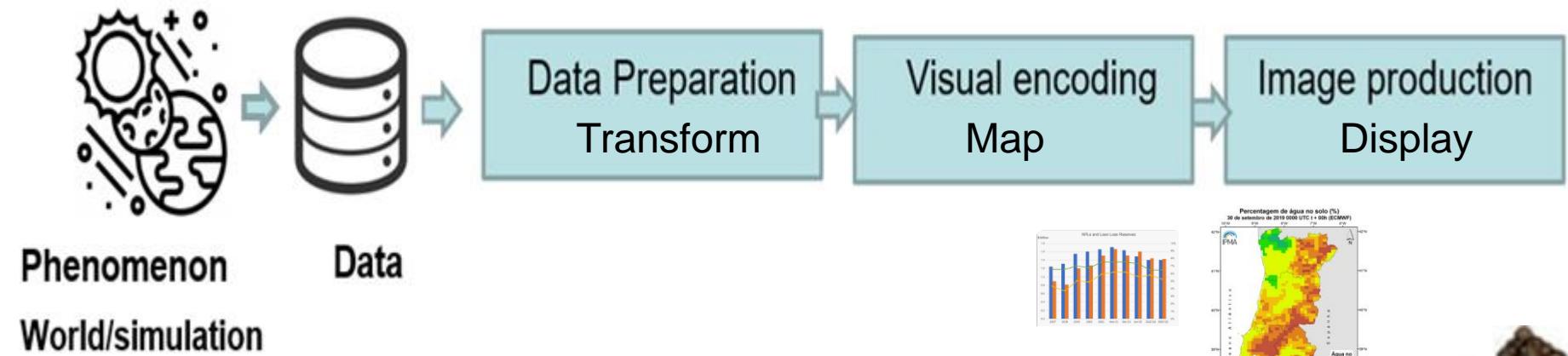
Creating a Visualization



[heart-disease-prevalence](#)

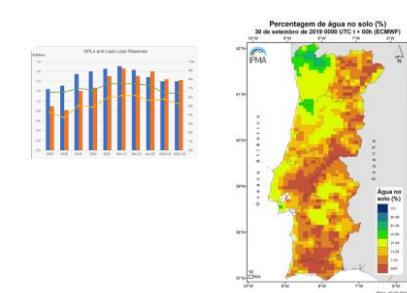
The process of creating a (simple) Visualization

The data may be **acquired from the world** (e.g. sensors, questionnaires)
or **simulated** (e.g. Finite Element Analysis, weather models)



The users should get **insights** from analyzing the visualization
and be supported in their **tasks** (answer their questions)

It is a **human-in-the-loop** process

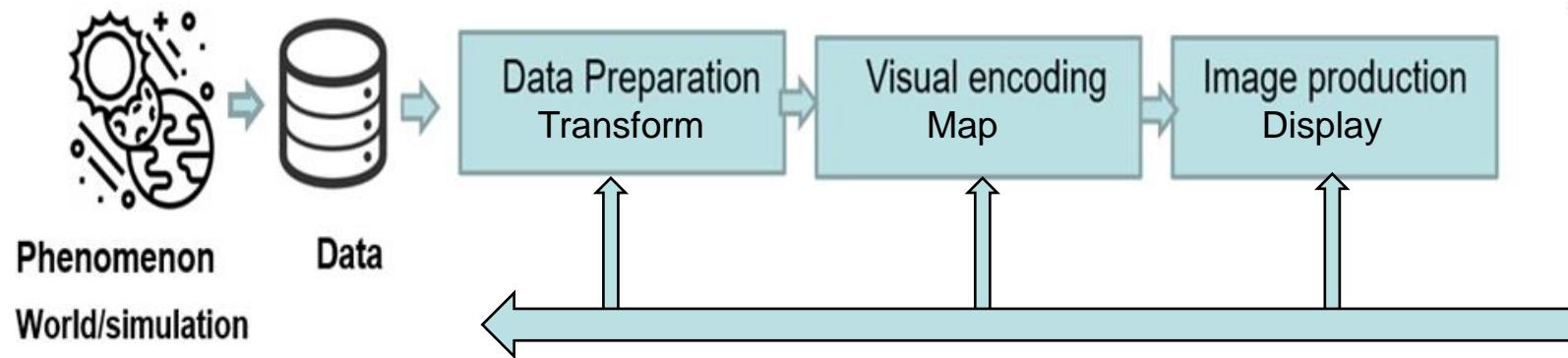
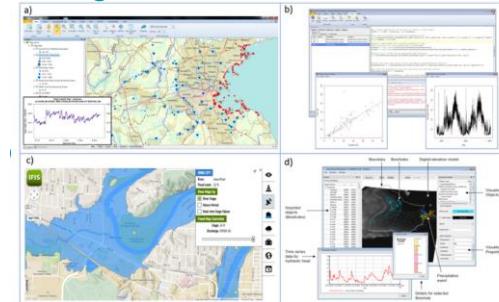


- “***human-in-the-loop***” problems involve the **user as a part of the system**
- They are **very complex** due to the facts that:
 - humans are very complex systems
 - not well known
 - in general we cannot change them
- Target users’ profile, needs, and context of use must be carefully considered whenever designing a visualization

Visual Data Exploration application

In this case it is necessary to provide ways to **let users interact** along the process

[An overview of visualization and visual analytics applications in water resources management - ScienceDirect](#)



Developing this type of application implies a specific approach:
Human-centred design

Questions to guide the Visualization creation process

Who: Users' profile and context of usage

Why: Tasks and questions the visualizations must support

What: Data to visualize

How: Visual representations (visualization techniques) to use

Who

Why

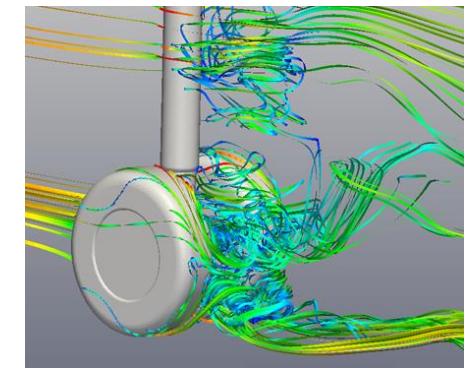
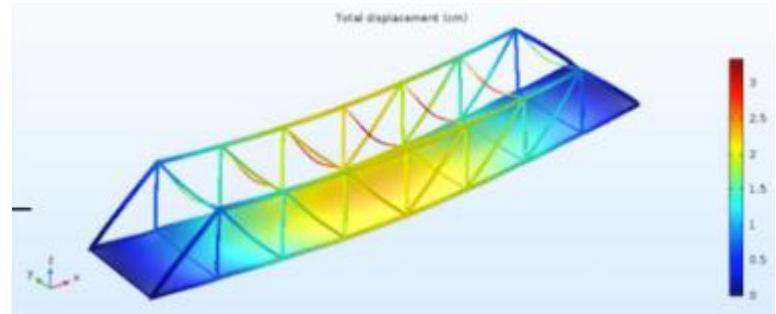
What

How

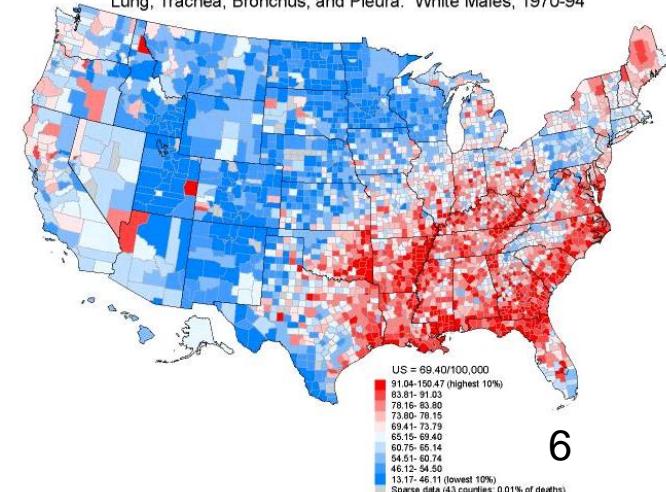
After answering the questions
It is easier to select the techniques

Visual mapping

- It is necessary to decide:
 - which visual structures use to **represent** the data
 - their location in the display
- Some types of data can be easily mapped to a spatial location:
e.g. data with a topological/geographical structure
- Abstract data don't have an easy correspondence with the dimensions of the physical space around us

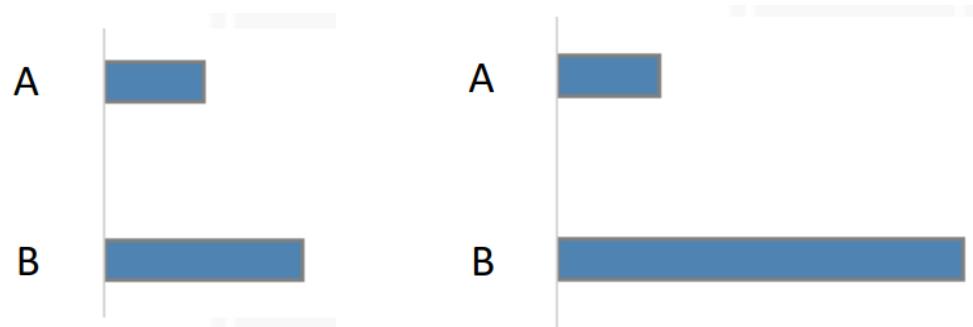


Cancer Mortality Rates by County (Age-adjusted 1970 US Population)
Lung, Trachea, Bronchus, and Pleura: White Males, 1970-94

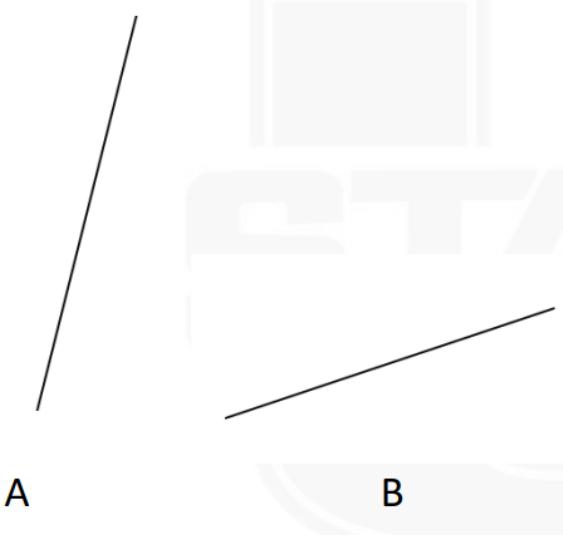


- Some ways are better to visually represent values than others

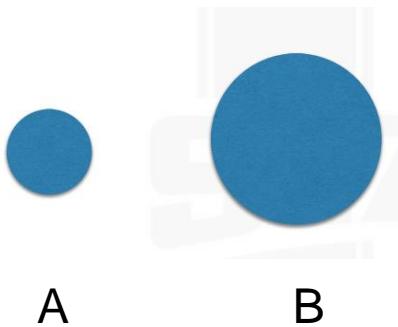
How much longer?



How much steeper?



How much larger?



Three **structures** must be defined in the **visual mapping/encoding**:

- spatial substrate
- graphical elements
- graphical properties

- **Spatial substrate** - dimensions in physical space where the visual representation is created (can be defined in terms of axes and type of data)
- **Graphical elements** - anything visible appearing in the space
points, lines, surfaces, volumes
- **Graphical properties** – properties of the graphical elements to which the human retina is very sensitive - **retinal variables**:
size, orientation, color, texture, and shape

- **Spatial substrate** axes (x, y, ...)
type of data (quantitative, ordinal, categorical)

- **Graphical elements** points
lines
surfaces
volumes

- **Graphical properties** retinal variables:

Size	Color Hue
○ ○ ○ ○ ○	Blue Red Black Orange Dark Blue
Orientation	Shape
/ / - -	□ × ☆ ○ △
Color Saturation	Texture
● ● ● ● ●	●●●●●

size,
orientation
color (depends on physiology and culture)
texture
shape

Retinal variables – Graph workflow

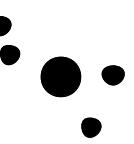
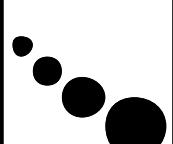
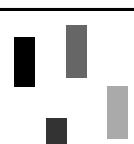
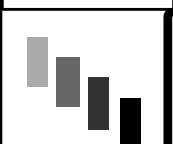
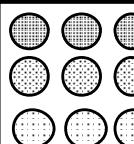
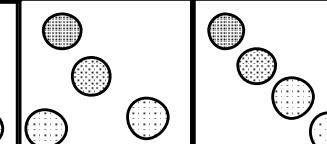
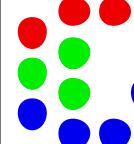
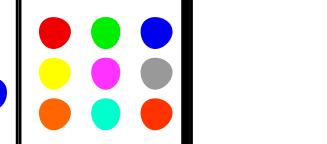
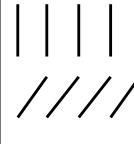
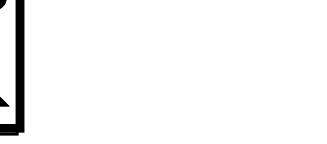
- Some visual attributes as **size**, **proximity** are quickly processed by visual perception, **before the cognitive processes** come into play

Example:
mapping numerical values
to the length of bars: bar chart



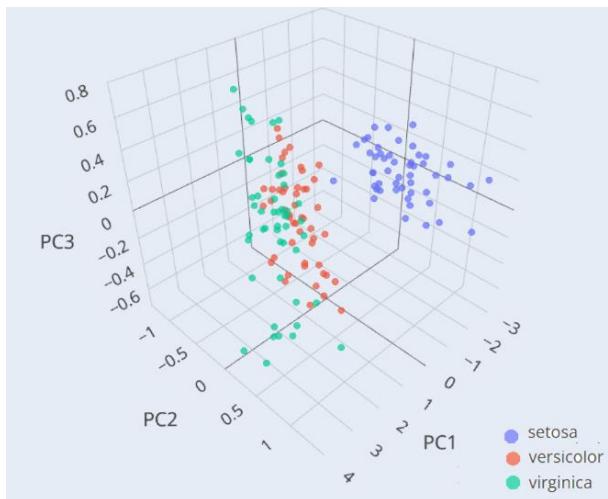
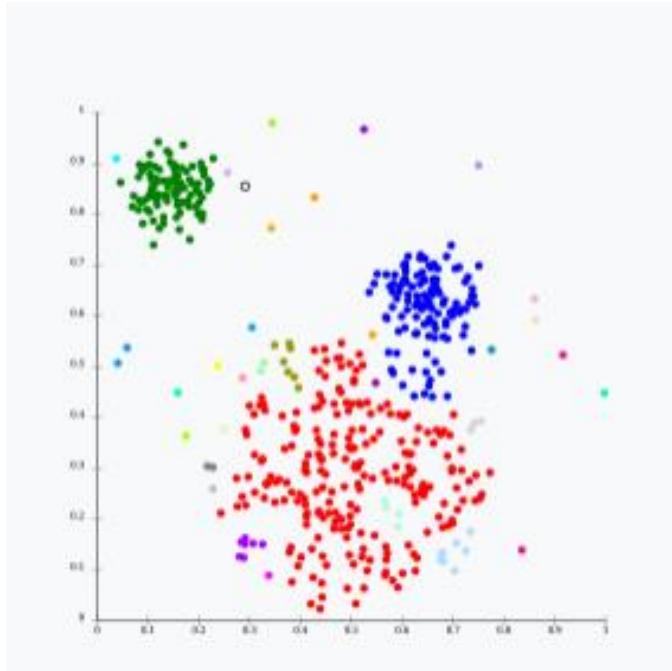
(Mazza, 2009)

How to select visual encodings?

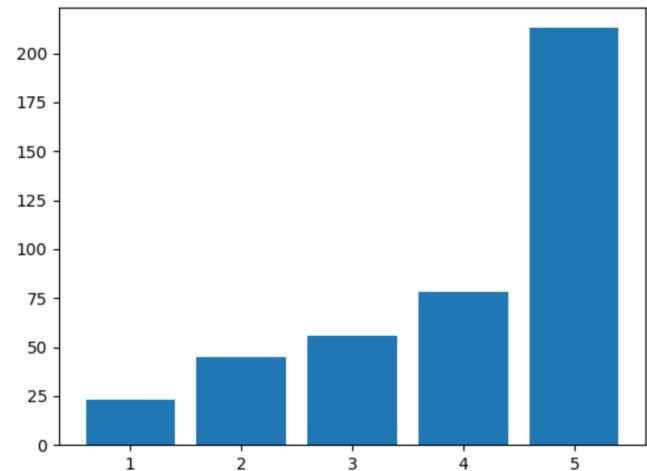
	Association	Selection	Order	Quantity
Size				
Value				
Texture				
Colour				
Orientation				
Shape				

Interpretation of Bertin's guidance
regarding the suitability of various
**encoding methods to support
common tasks** ([Spence, 2007](#))

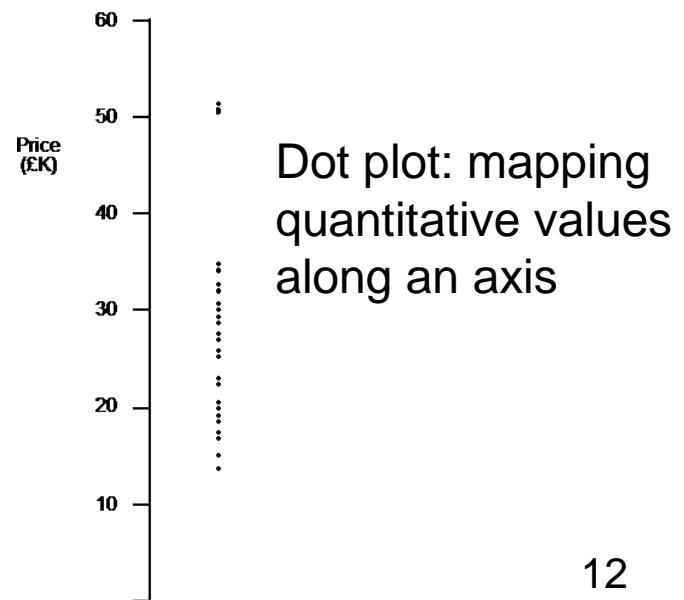
Note that **only size is adequate to
represent quantity accurately!**



Scatter plots with different clusters:
colour used for association

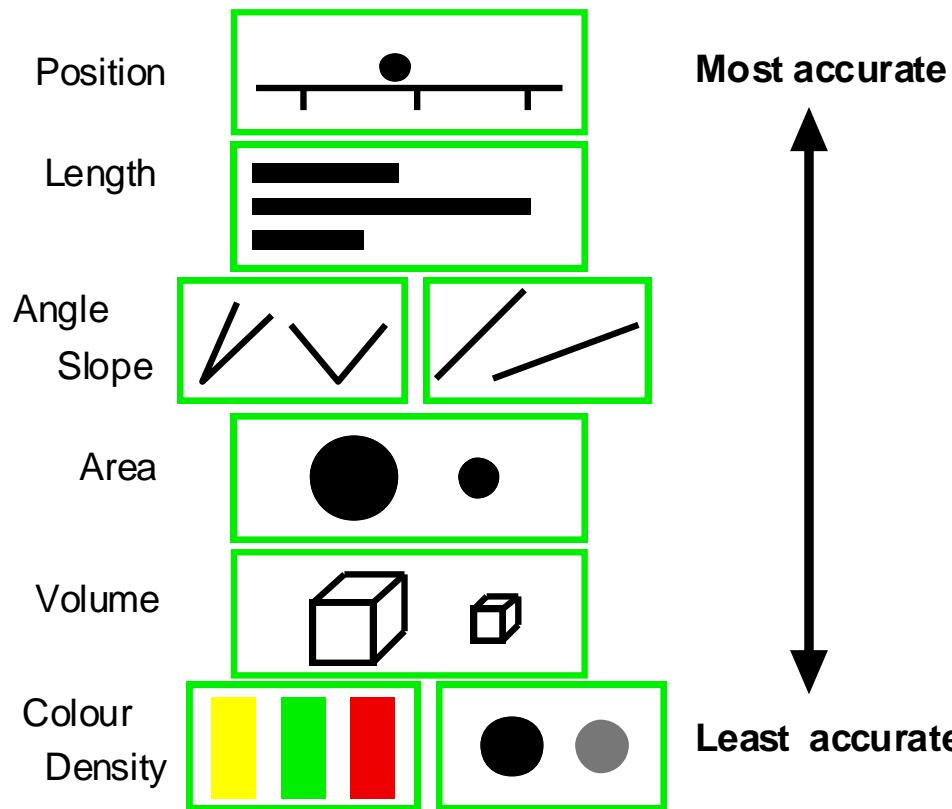


Bar chart: mapping quantitative
values to the length of bars
[Matplotlib Bar Chart - Python Tutorial](#)



Dot plot: mapping
quantitative values
along an axis

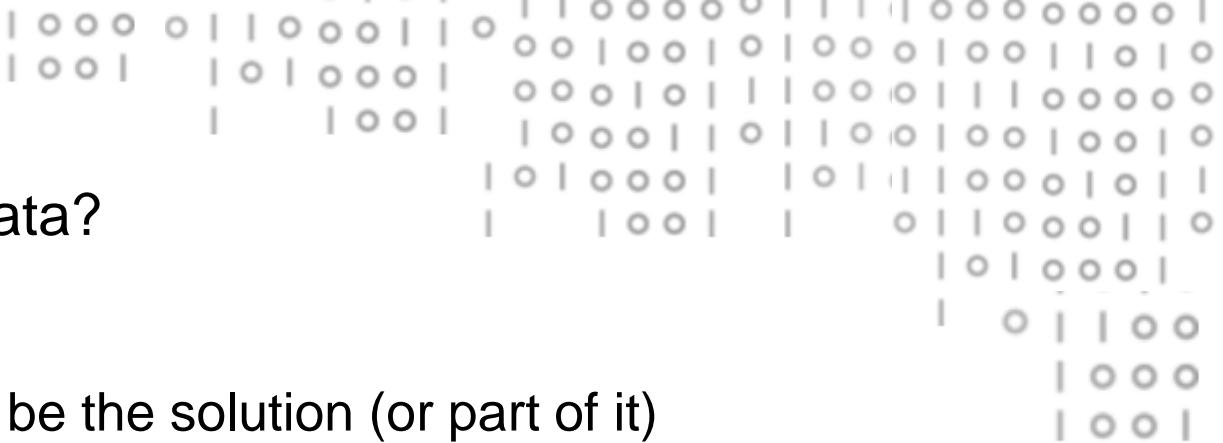
How to select visual encodings to accurately represent quantity?



The relative difficulty of **assessing quantitative value** as a function of encoding mechanism, as established by Cleveland and McGill (Spence, 2007)

In a nut shell:

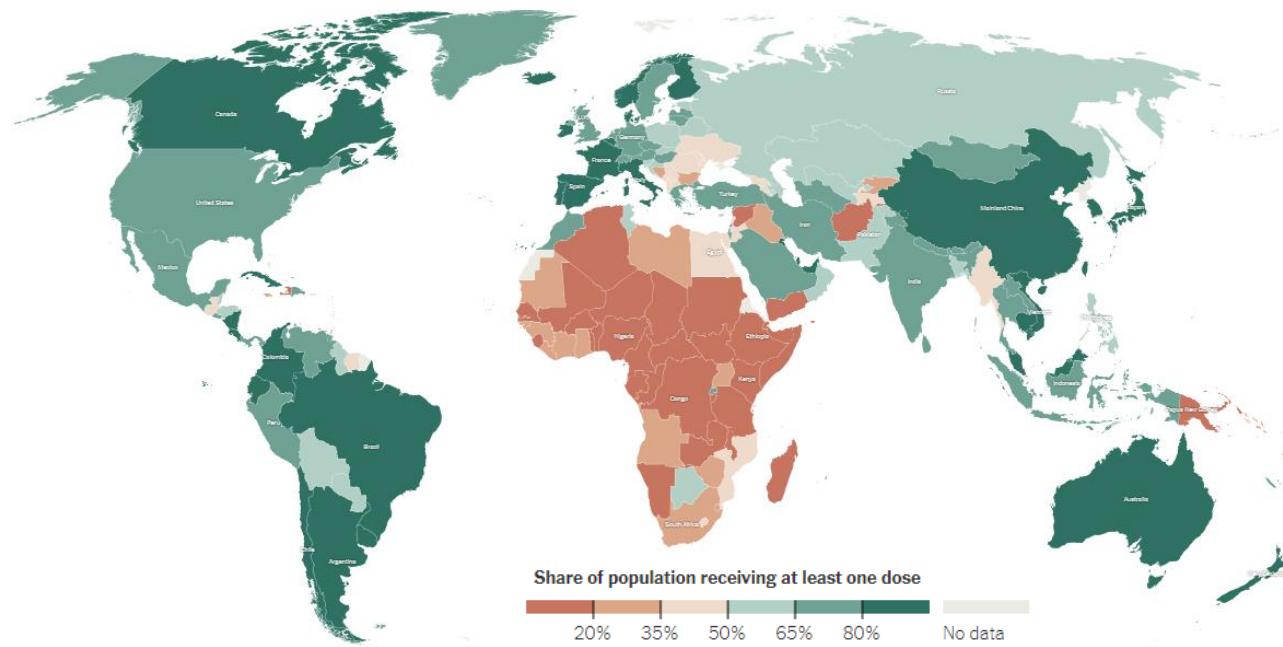
Do you have a lot of data?



- Visualization may be the solution (or part of it)
- Creating a Visualization has several phases
- Visual mapping is core
- There are several possible visual encodings/
visualization techniques
- But,
How to select techniques? → next topic

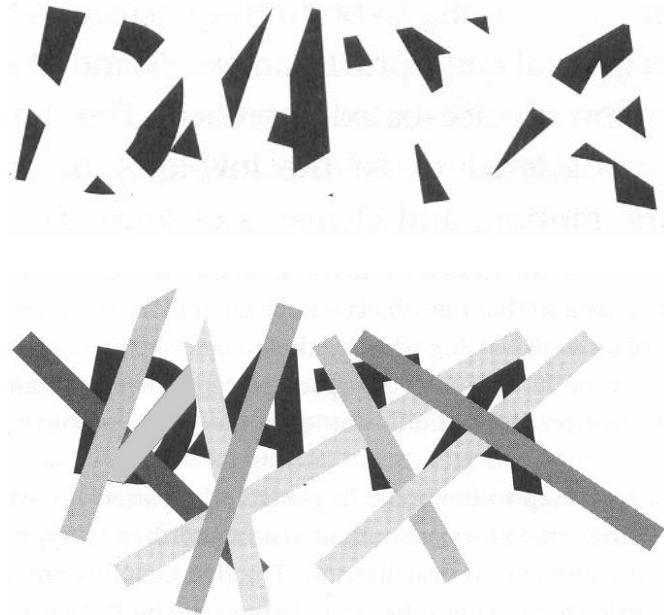


Visually encoding value



Remember:

- The **Human Visual system** is the product of millions of years of evolution
- Although very flexible, **it is tuned to data represented in specific ways**
- If we understand how its mechanisms work we will be able to produce better results



Pre-attentive attributes can help observers to see before though

6970425934749

3587282949546

4244396854634

2356658789376

Example: Count the number of 7s

6970425934749

3587282949546

4244396854634

2356658789376

Procedure to follow to create visual representations of data

1. Understand the **problem**, and the **users**

2. Understand the users' **tasks/ questions**

3. Examine the **nature of the data** to represent and **pre-process** the data
(determine the **number of variables**/ attributes/dimensions)

4. Choose the **visual structures to map**
test several ideas ...

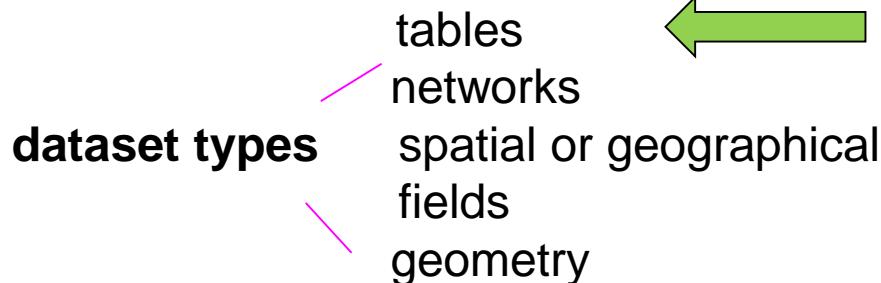
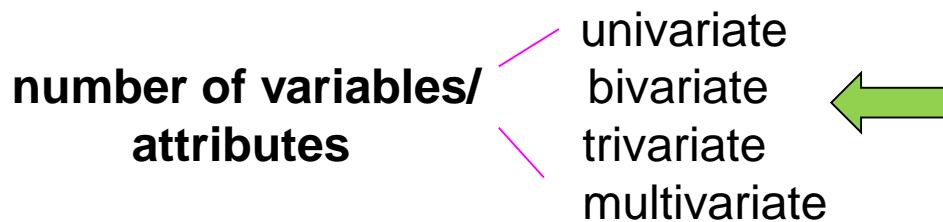
Who

Why

What

How

Important aspects to consider to select a Visualization technique:



Next: visualization techniques
organized according the n. of
variables/attributes (1, 2, 3)

of tabular data

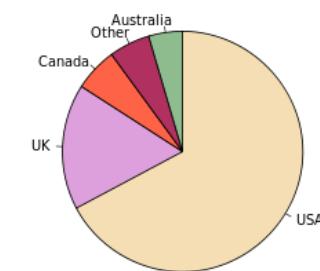
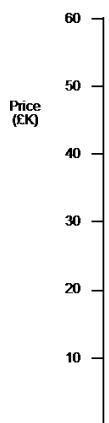
63	1106	2004	1803	2311
140	845	1235	898	3138
1245	1170	287	2430	1095
2350	2481	1325	393	1462
5	1136	376	759	2887
583	2703	1751	2298	1660
1569	861	1465	1467	1180
955	706	1007	2348	564
2429	1486	2514	2771	887
881	1467	1007	2348	564
776	2514	2514	2771	887

Common Visualization Techniques to visually represent univariate, bivariate tabular data

Univariate data

- dot plot
- box plot
- bar chart
- histogram
- pie chart

...

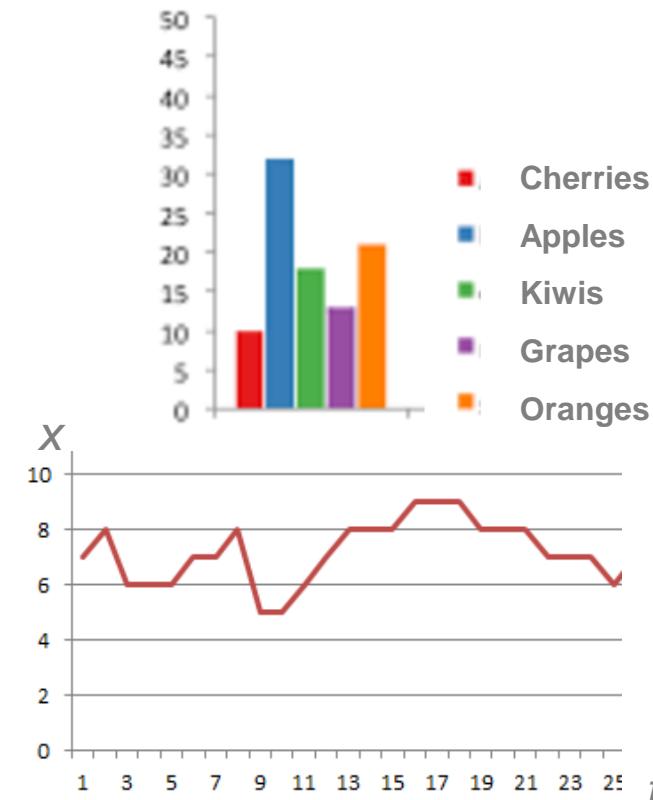
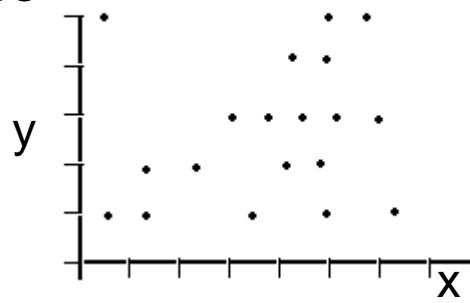


Bivariate data

- scatter plot
- line plot
- time series

...

...



Representing univariate data

- A common situation consists in representing a **set of values**

- Well established techniques exist

- But new ones can be invented!

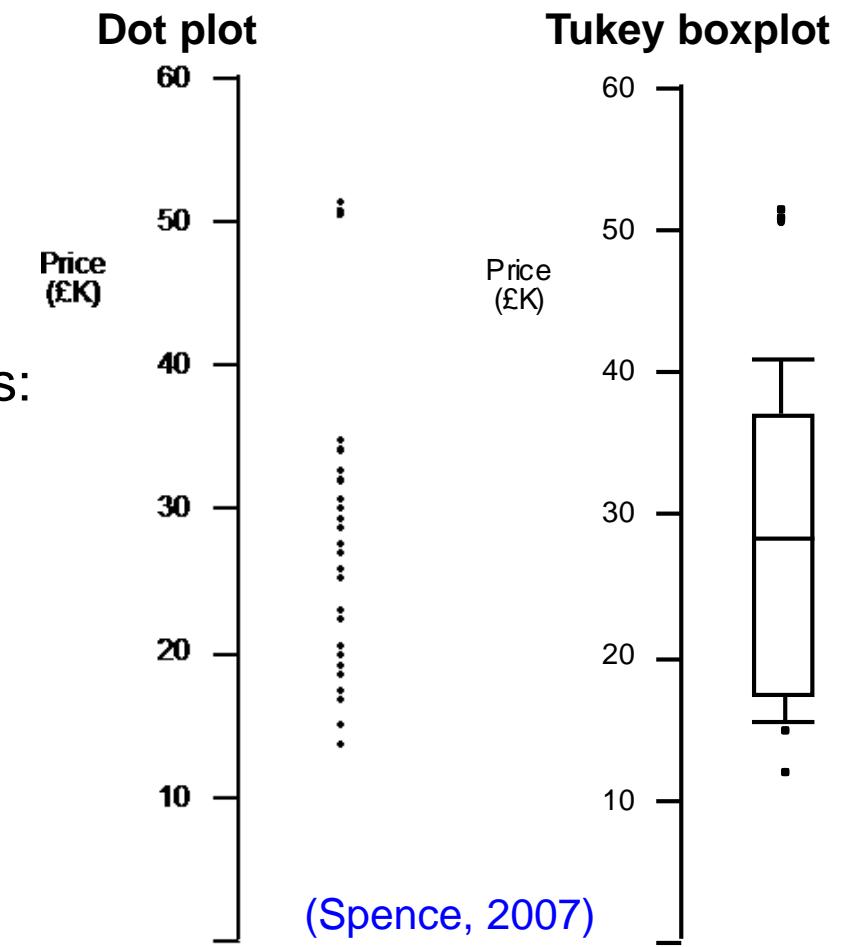
Example: Price for a number of cars:

- dots on a linear scale

- box plot

(that will answer many questions:
median value, outliers,...)

- Including several in the same
Visualization eases comparison

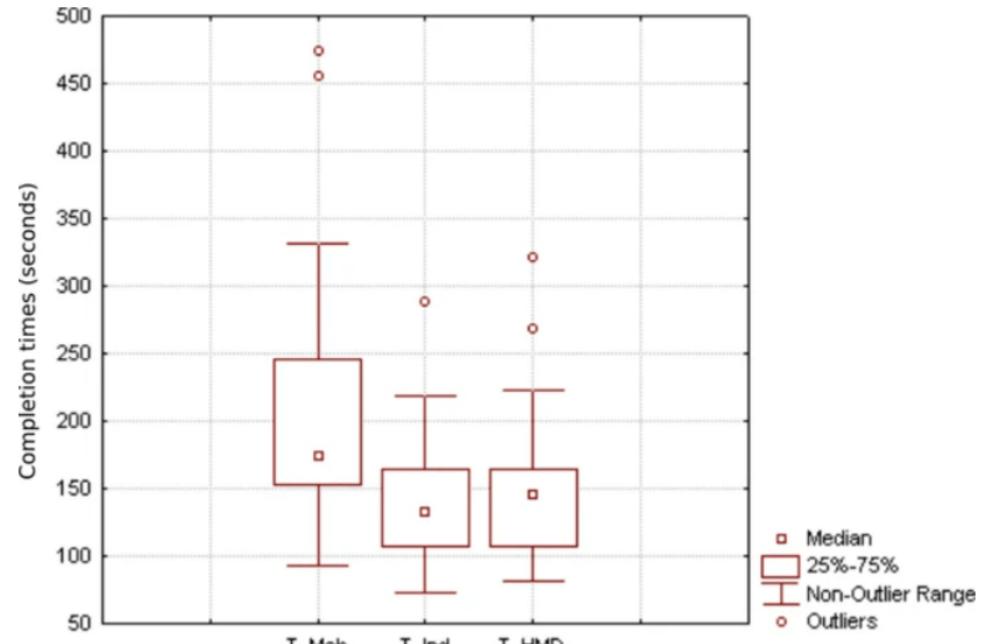


Example:

- Including several box plots in the same Visualization allows comparing:

Comparing 3 augmented reality visualization methods for assembly procedures (times)

Comparing augmented reality visualization methods for assembly procedures | Virtual Reality



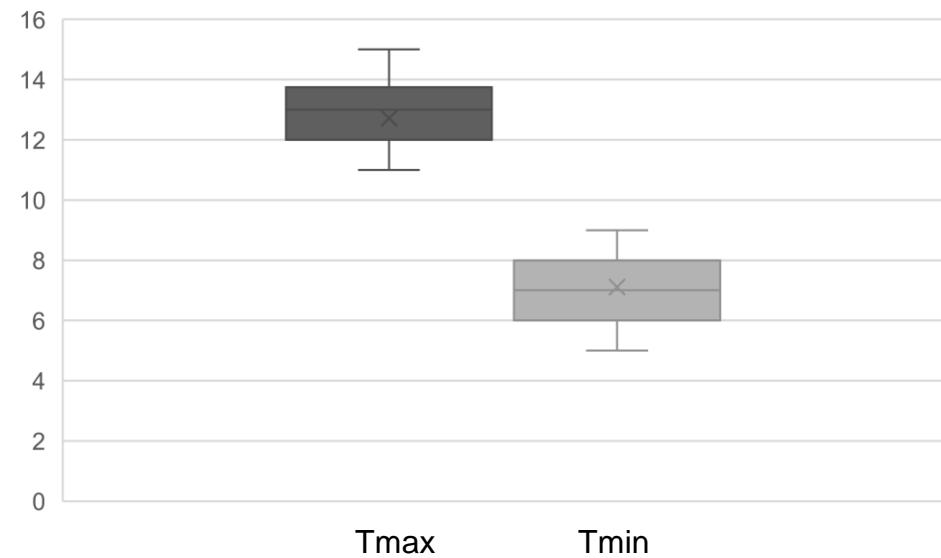
Task performance from the three methods based on time of completion

Exercise

day	Max T	Min. T
1	15	7
2	14	8
3	13	6
4	13	6
5	12	6
6	13	7
7	13	7
8	14	8
9	15	5
10	12	5
11	13	6
12	12	7
13	11	8
14	11	8
15	12	8
16	12	9
17	13	9
18	14	9
19	14	8
20	13	8
21	13	8
22	12	7
23	12	7
24	11	7
25	11	6
26	11	7
27	13	6
28	14	6

- Create boxplots using Excel

Max and min temperatures in February (°C)

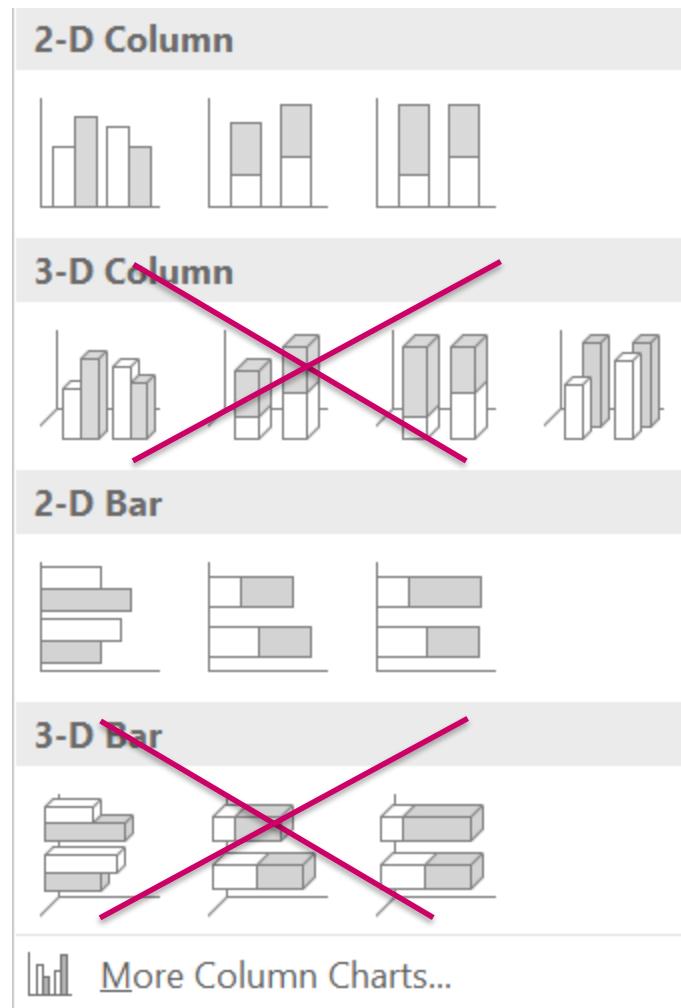


Exercise

- Create a bar chart of these data using Excel (be careful with the options!!)

Fruits/Veg	Quant (t)
oranges	100
lemons	20
mandarines	30
limes	10
hazelnuts	40
nuts	60
peanuts	40
kiwis	70
pineples	50
mangos	20

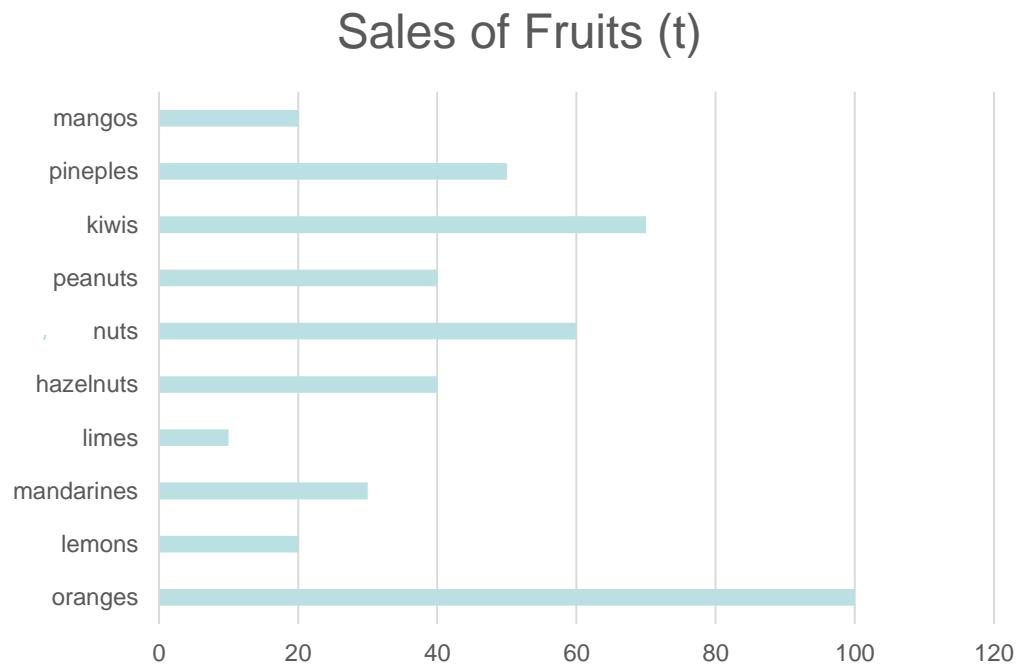
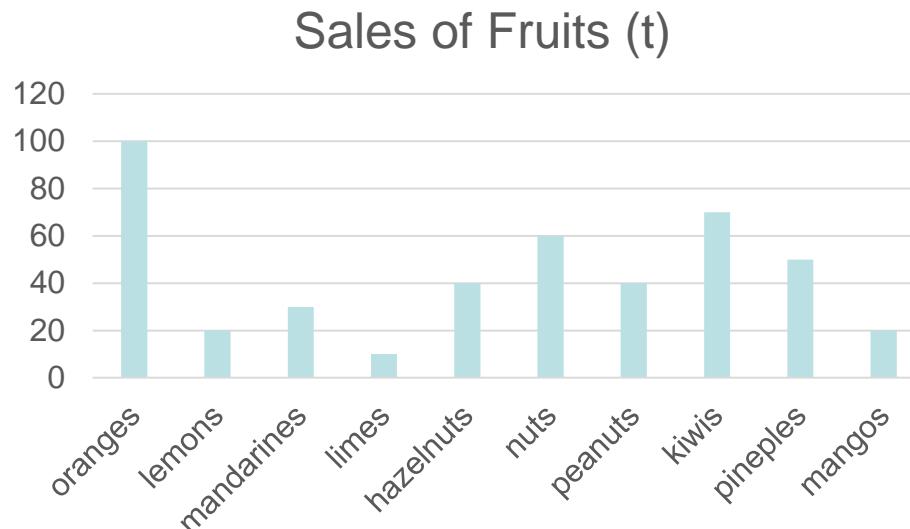
**Remember, 3D
bar charts are
not an option!!!**



Exercise

- Acceptable solutions

Fruits	Quant (t)
oranges	100
lemons	20
mandarines	30
limes	10
hazelnuts	40
nuts	60
peanuts	40
kiwis	70
pineples	50
mangos	20

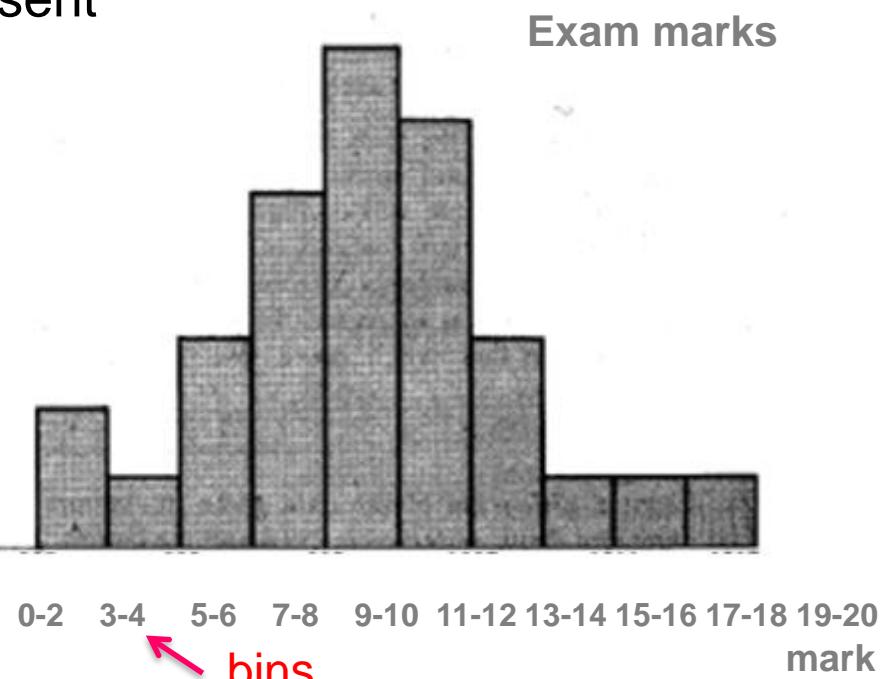
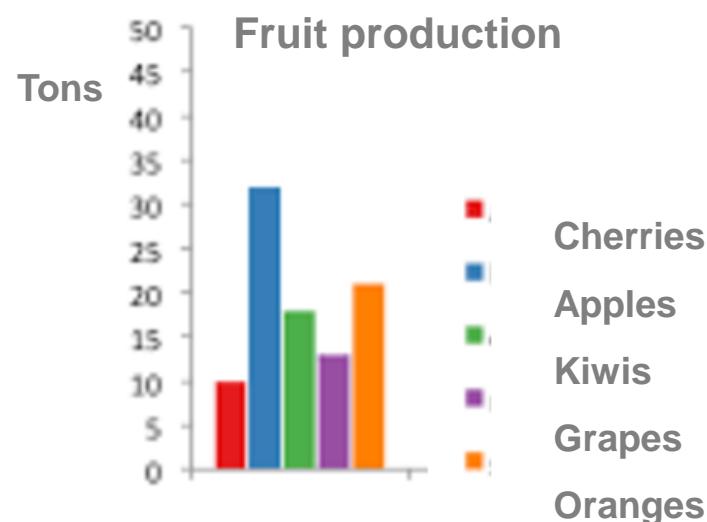


Two common techniques not to be confounded !

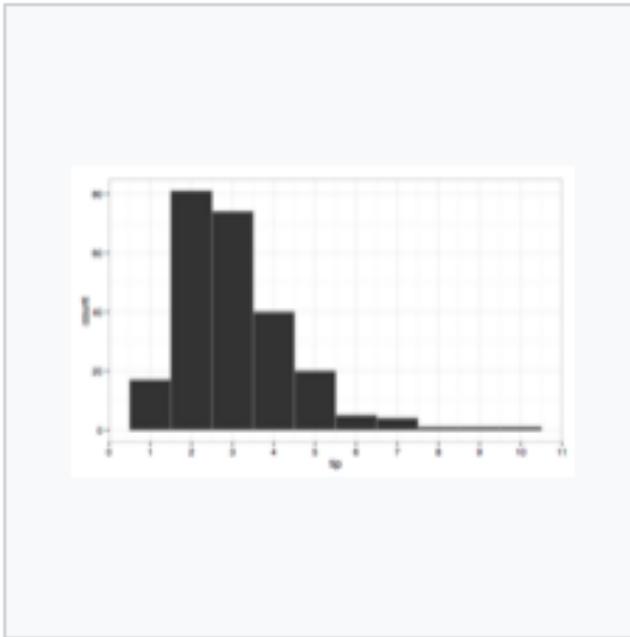
Histogram → represents a distribution of **numerical data**

Bar chart → represents the number of occurrences of a **categorical/ordinal data**

Both represent data by rectangular bars (vertical or horizontal) with length proportional to the values they represent

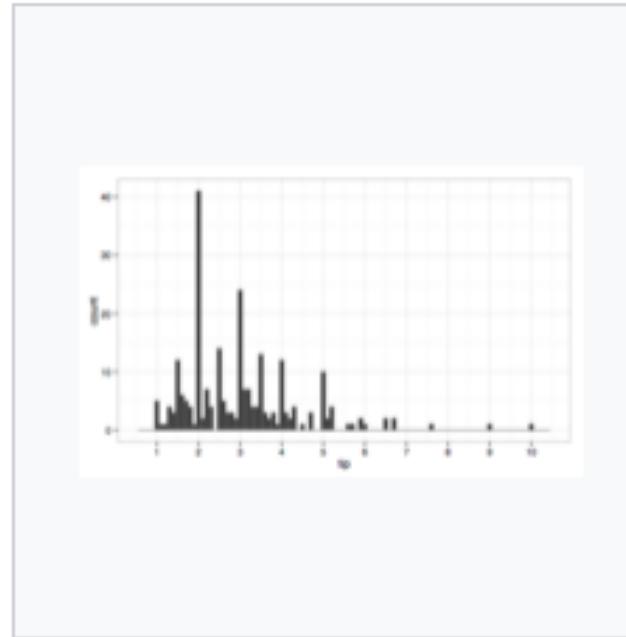


- It is a good idea to plot the data using several different bin widths to learn more about it.
- Example about tips given in a restaurant in the USA:



\$1 bin width, skewed right,
unimodal

[Histogram - Wikipedia](#)

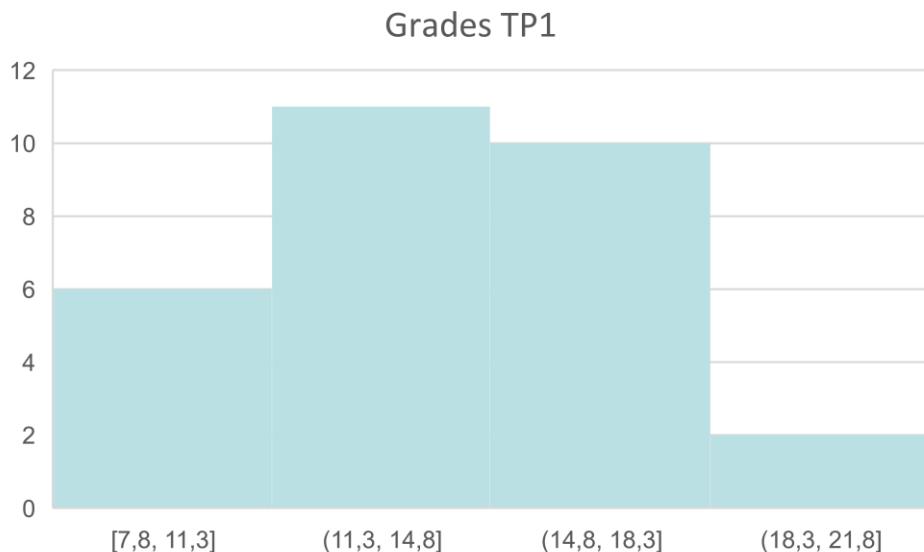


10c bin width, still skewed right,
modes at \$ and 50c, indicates
rounding, also some outliers

Exercise

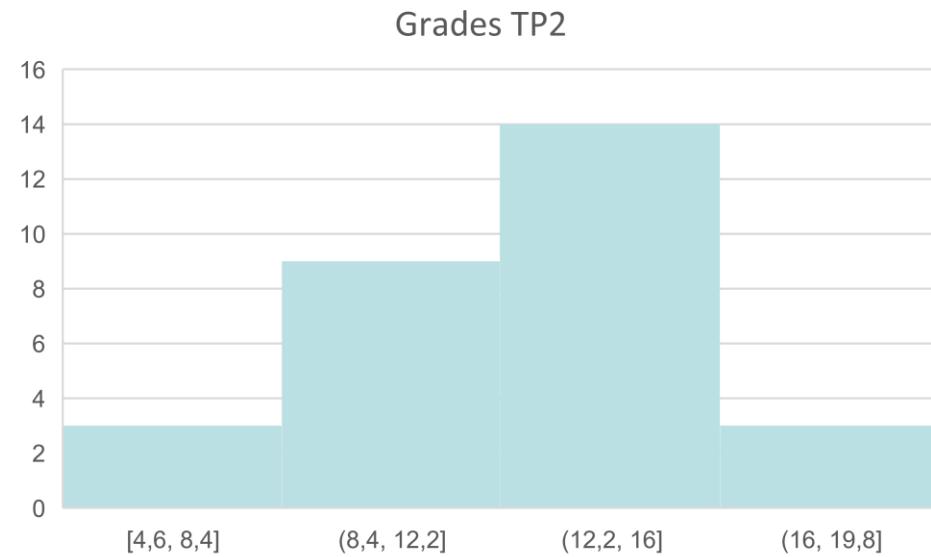
- Create histograms for the grades of T1 and T2 using Excel

TP1	TP2
15,6	9
14,2	15,8
9,8	7,6
10	4,6
15,2	15
10	10
14,2	10,2
9,6	12,8
7,8	15
8	10,4
13,4	15,2
18,4	13,2
14,6	11,2
17,4	15,8
15,2	17,4
16	12,6
11,6	9
14,4	14,8
16,8	7,2
14,2	16,6
15,6	15
15	9,4
13,2	12,8
17,4	14,2
12,8	12
14,4	18,4
13,6	12,6
17,6	9,2
20	13



Bin width 3,5
↑

Bin width 3,8 →

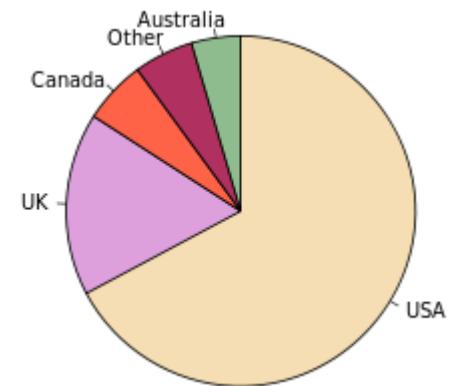


Another simple (and too common) technique

- Pie Chart

Represents numerical proportion, **parts of an whole**

The arc length of each slice (its central angle and area), is proportional to the quantity it represents



Native English speaking population

Are **much controversial**:

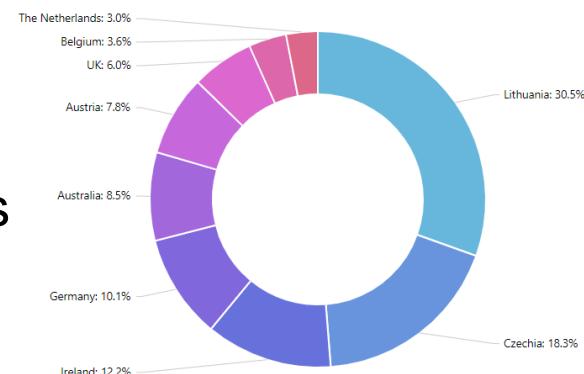
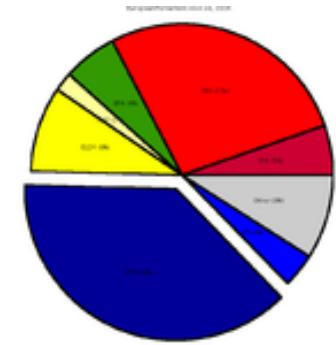
many experts recommend avoiding them

Save the Pies for Dessert



ficult to compare different sections of a pie chart, or to compare data across different pie charts

Variations of pie charts:



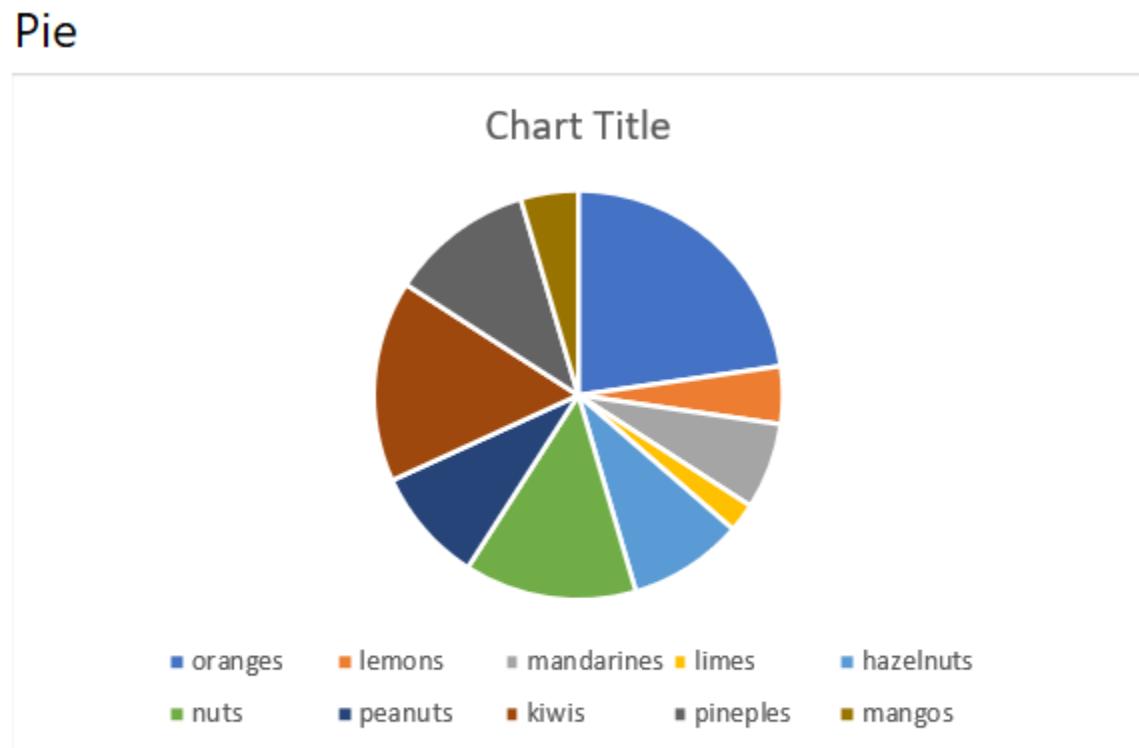
- Simple criteria to determine whether a pie chart is acceptable
- Consider it **only if:**
- **The parts make up a meaningful whole**
- + **The parts are mutually exclusive**
- + **There are <6 parts and slices have not very different sizes**

**If the main purpose is to compare between the parts,
use a different chart!**

<https://eagereyes.org/techniques/pie-charts>

- Excel (good) advise:

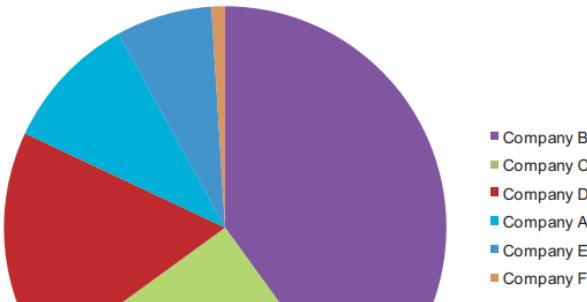
Do not use this chart if it contains many slices
as angles are hard to estimate!



A pie chart is used to show proportions of a whole. Use it to show numbers that relate to a larger sum and always equal 100%. Do not use this chart if it contains many slices as angles are hard to estimate.

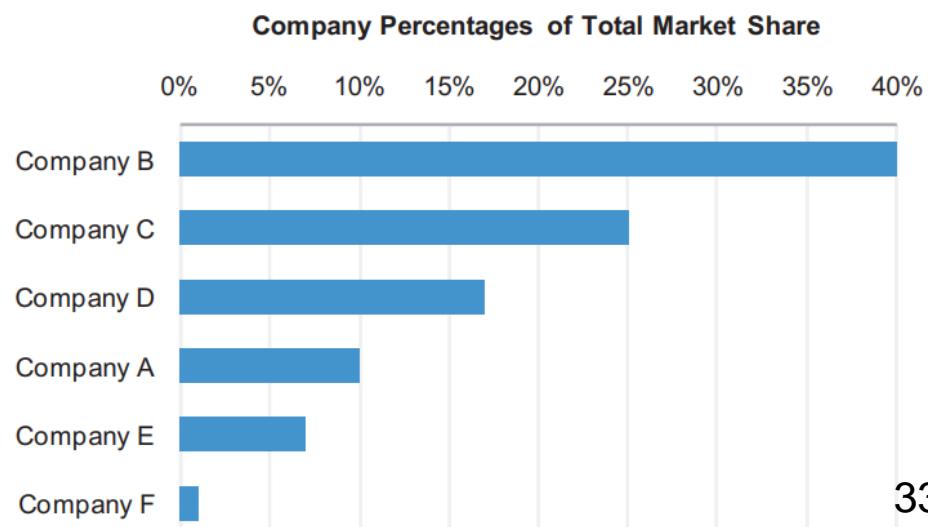
Exercise

- What are the issues with this pie chart?



- What if the user needs to assess and compare the companies market share?

Save the Pies for Dessert



Representing bivariate data

- The **scatter plot** is the conventional representation for two **quantitative** variables

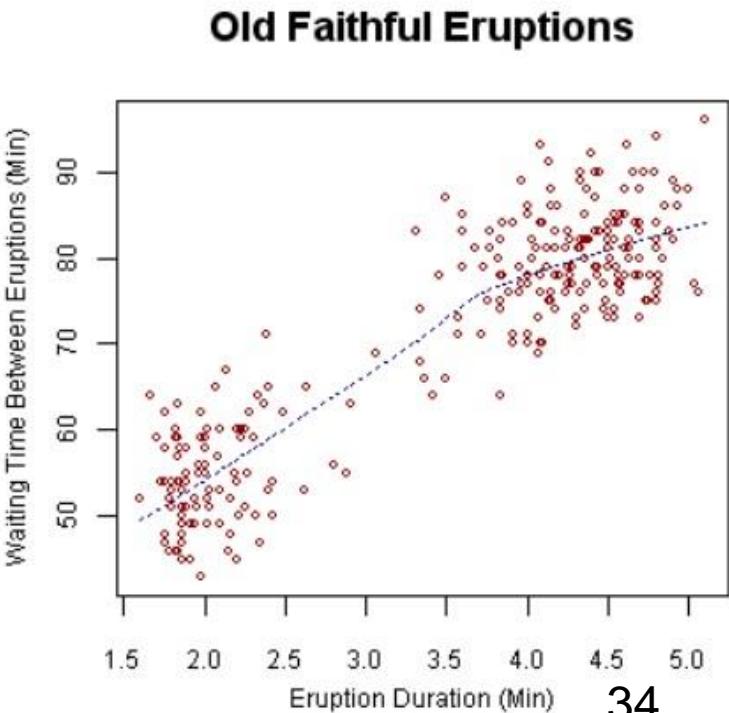
Each observation is represented by a point on a two dimensional space

The axes are associated with these two quantitative attributes

This representation affords awareness of:

- **general trends**
- **local trade-offs**
- **outliers**

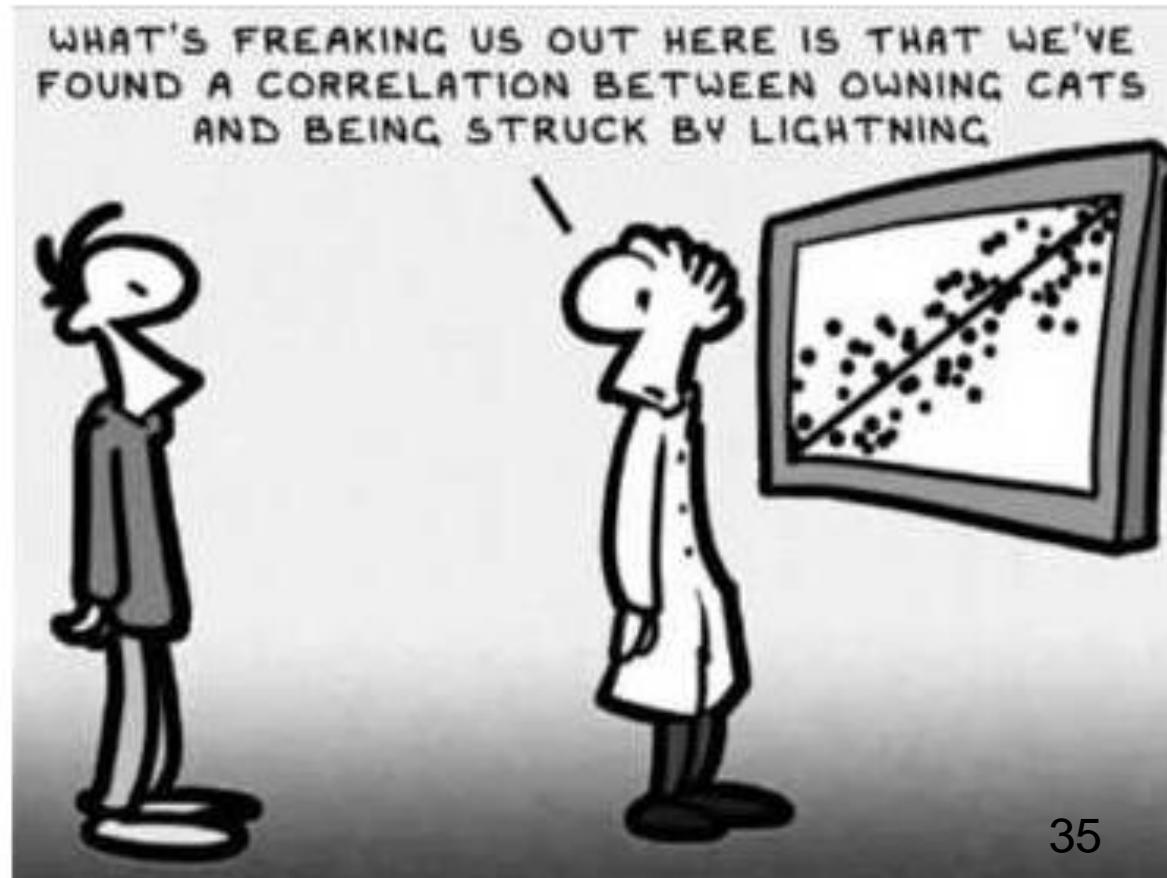
[Scatter plot - Wikipedia](#)



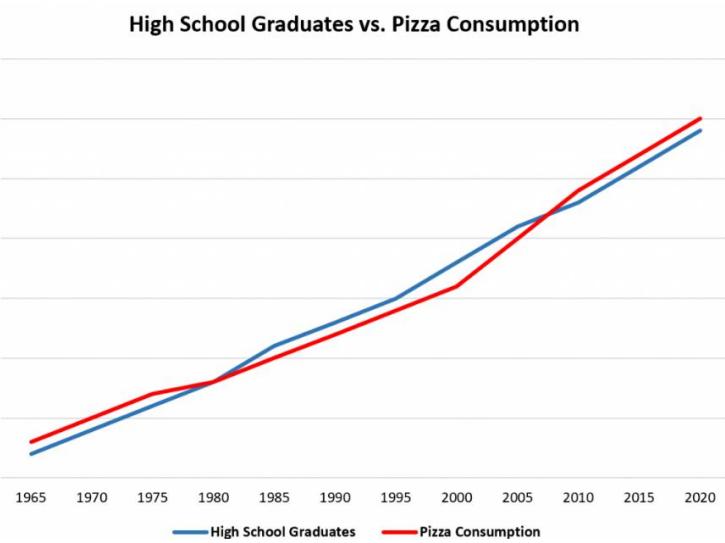
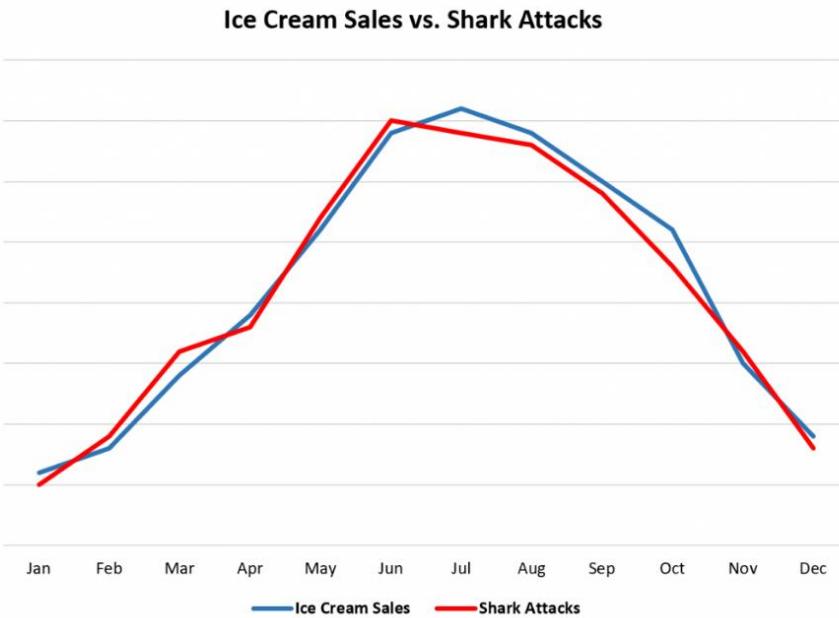
Representing bivariate data

- Be careful that to establish causation you need to use other methods:

Correlation is not causation!



Correlation is not causation:



Correlation Does Not Imply Causation:
5 Real-World Examples

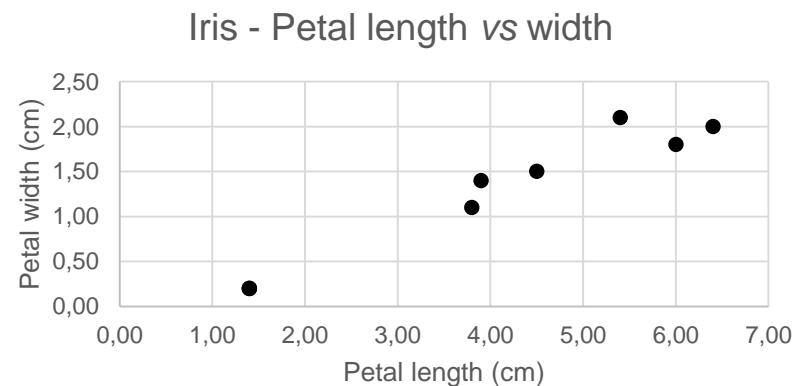
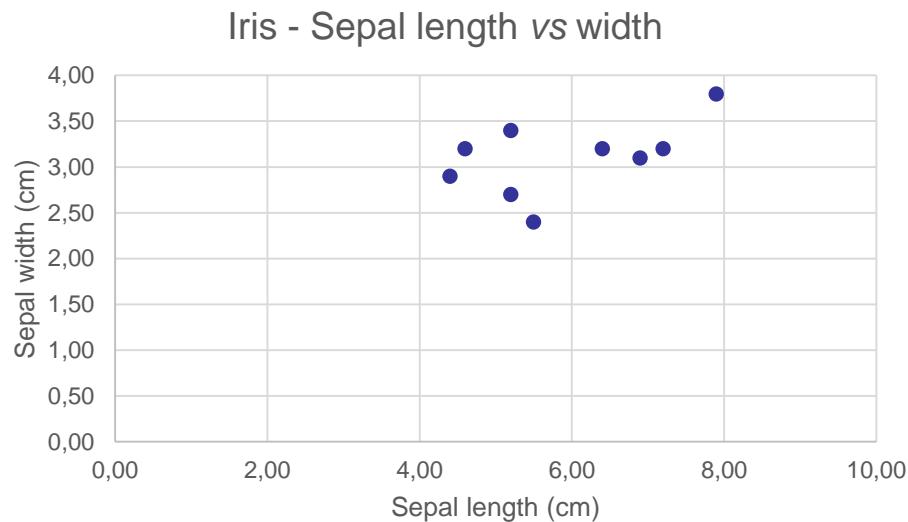
Exercise

- Create scatter plots using Excel
Iris dataset (a few observations)

	sepal length	sepal width	petal length	petal width	class
1	4,40	2,90	1,40	0,20	setosa
2	5,20	3,40	1,40	0,20	setosa
3	4,60	3,20	1,40	0,20	setosa
4	6,40	3,20	4,50	1,50	versicolor
5	5,20	2,70	3,90	1,40	versicolor
6	5,50	2,40	3,80	1,10	versicolor
7	7,20	3,20	6,00	1,80	virginica
8	7,90	3,80	6,40	2,00	virginica
9	6,90	3,10	5,40	2,10	virginica



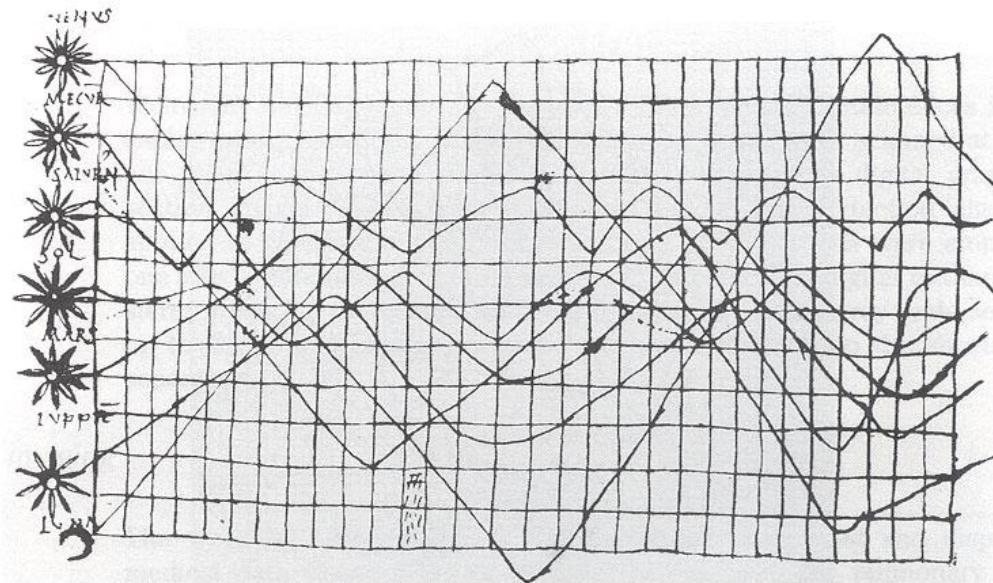
Iris Species



Representing bivariate data

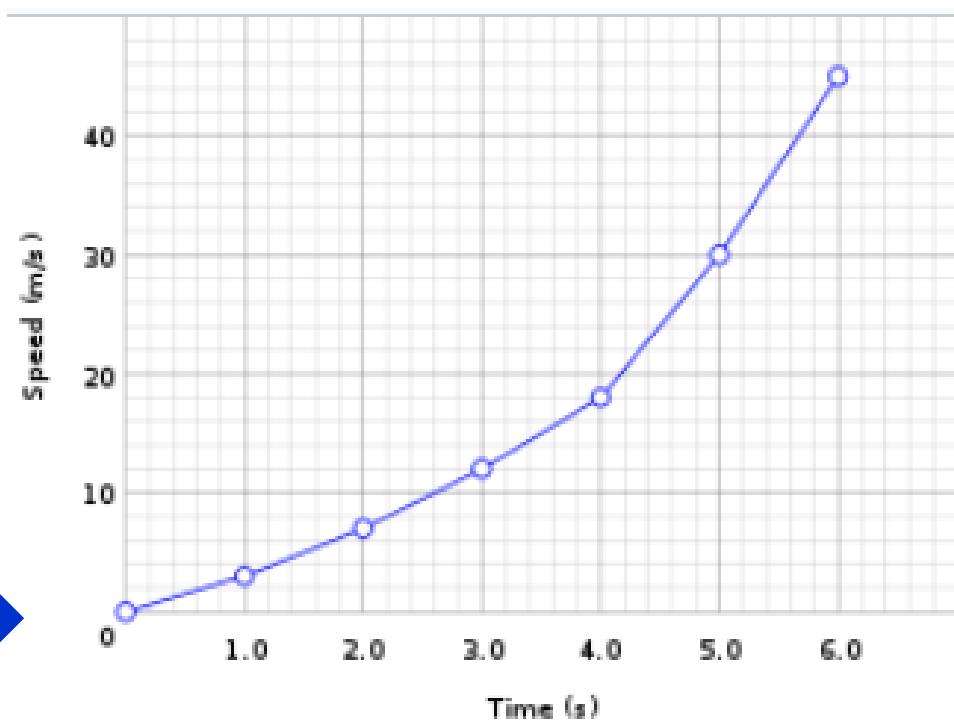
The line chart

One of the oldest known and ubiquitous Visualization techniques



Inclination of orbits along the time - Xth century (Tufte, 1983)

- A **line chart** or **line plot** or **line graph** or **curve chart** displays information as a series of data points called 'markers' connected by straight line segments
- Basic type of chart common in many fields
- Often used to visualize a trend in data over intervals of time (**time series**)

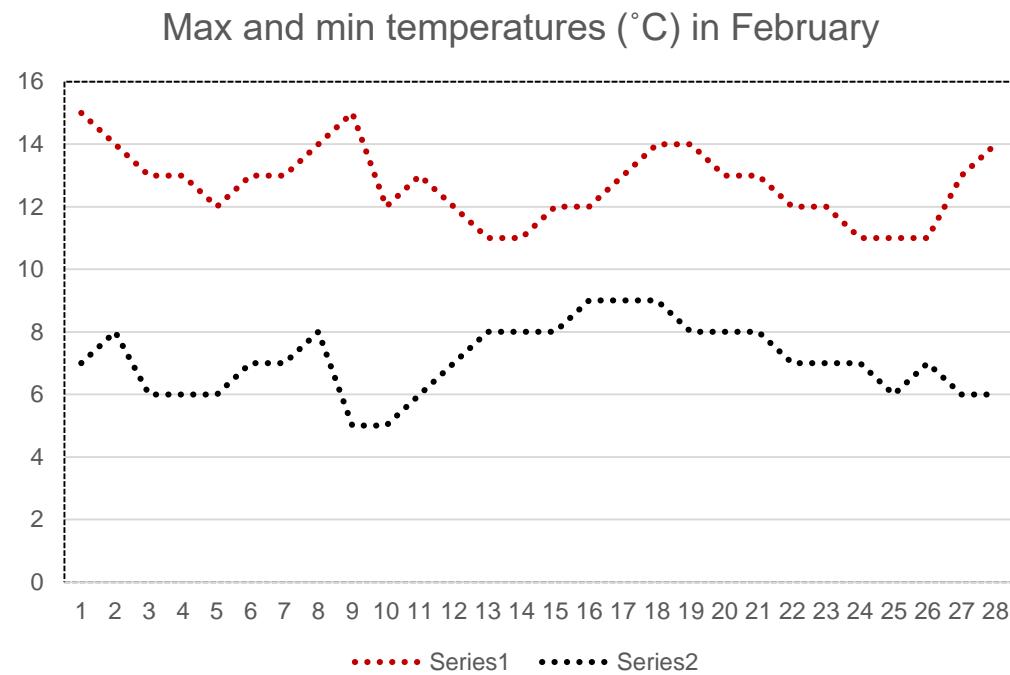


[Scatter plot - Wikipedia](#)

Exercise

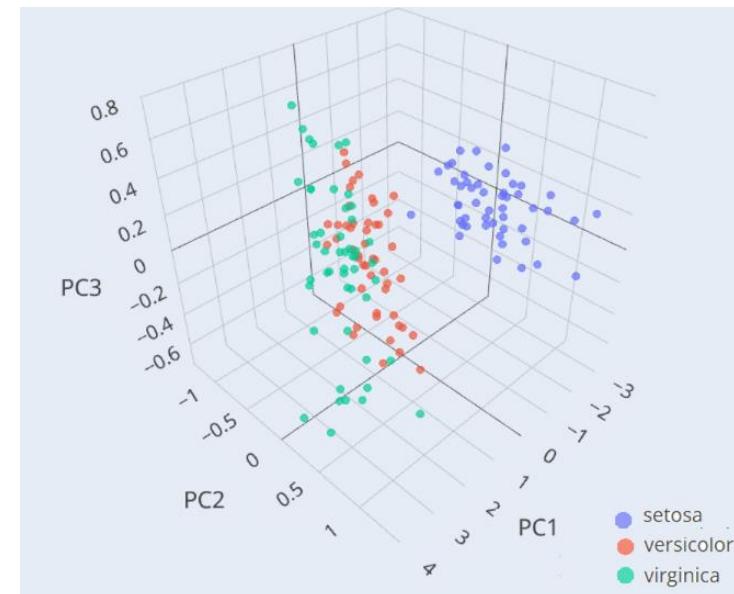
day	Tmax	Tmin
1	15	7
2	14	8
3	13	6
4	13	6
5	12	6
6	13	7
7	13	7
8	14	8
9	15	5
10	12	5
11	13	6
12	12	7
13	11	8
14	11	8
15	12	8
16	12	9
17	13	9
18	14	9
19	14	8
20	13	8
21	13	8
22	12	7
23	12	7
24	11	7
25	11	6
26	11	7
27	13	6
28	14	6

- Create line charts using Excel

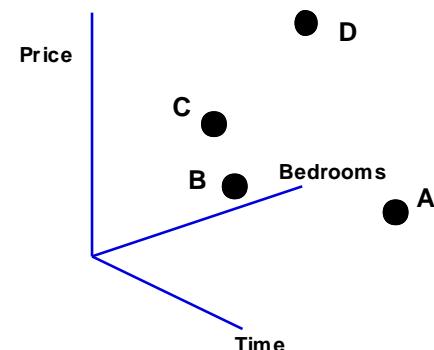


Representing trivariate data

- Since we live in a 3D world, representing trivariate data as points in a 3D space and displaying a 2D view seems natural
- However, these representations of **abstract data can be ambiguous**
- This can be mitigated by interaction, allowing the user to reorient the representation



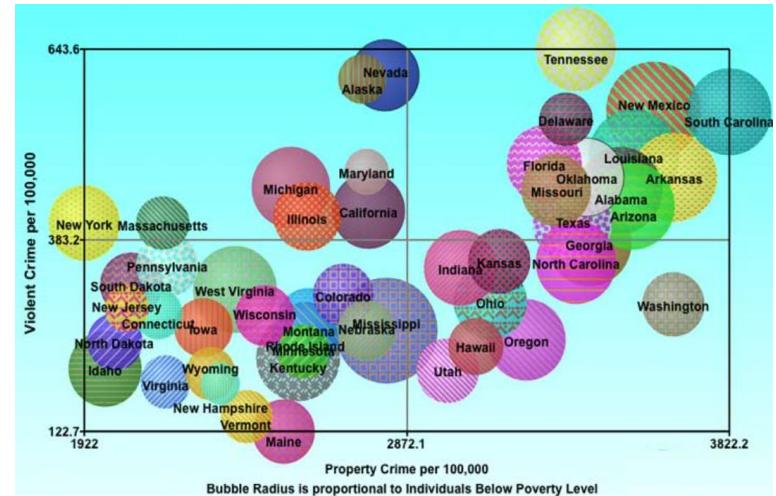
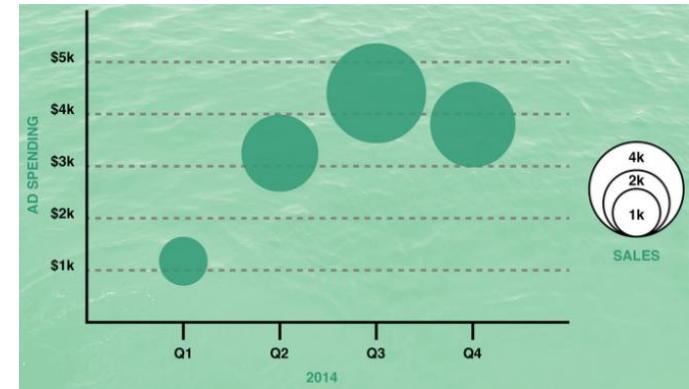
“for 3D to be useful, you’ ve got to be able to move it” (Spence, 2007)



Other Simple (and common) representations of trivariate data

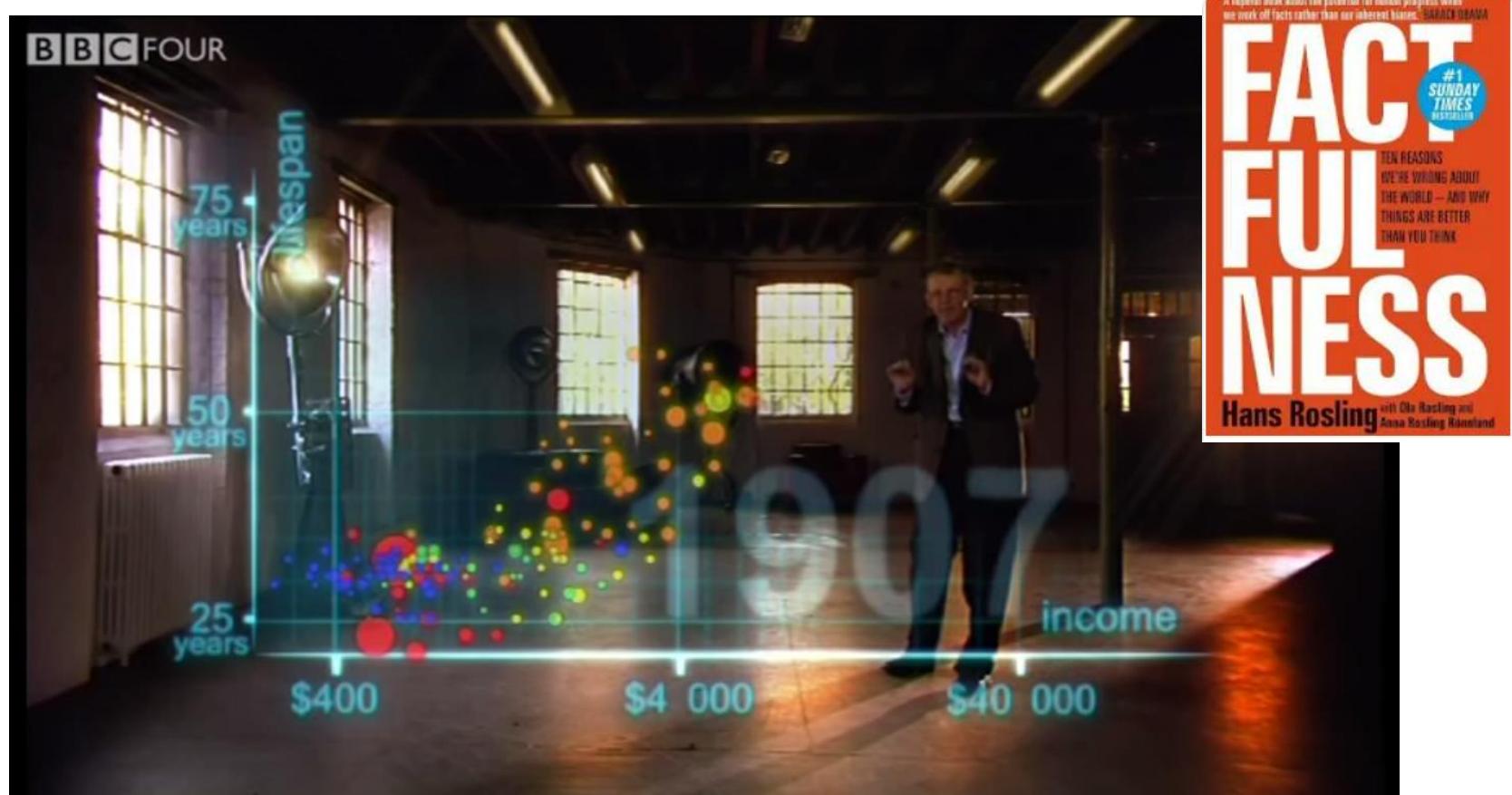
- In a **bubble chart** data are represented as a disk that expresses two of the values through the disk's *xy* location and the third through its size (radius or area?)
- Mapping the variable to circle size must be done carefully. The interpretation may be **ambiguous**
- Represent through size of the circle the variable to be perceived less accurately
- Representing one more dimension through color is possible

[Bubble chart - Wikipedia](#)



Example of scatter plot representing more than two variables: World health by Hans Rosling

life span/income (axes) + population (marker size) + continent (colour)
along the time (animation)



[Hans Rosling's 200 Countries, 200 Years, 4 Minutes - The Joy of Stats – BBC](#)
[Factfulness: Ten Reasons We're Wrong About the World - Rosling - Google Books](#) 43

Marker size represents another variable

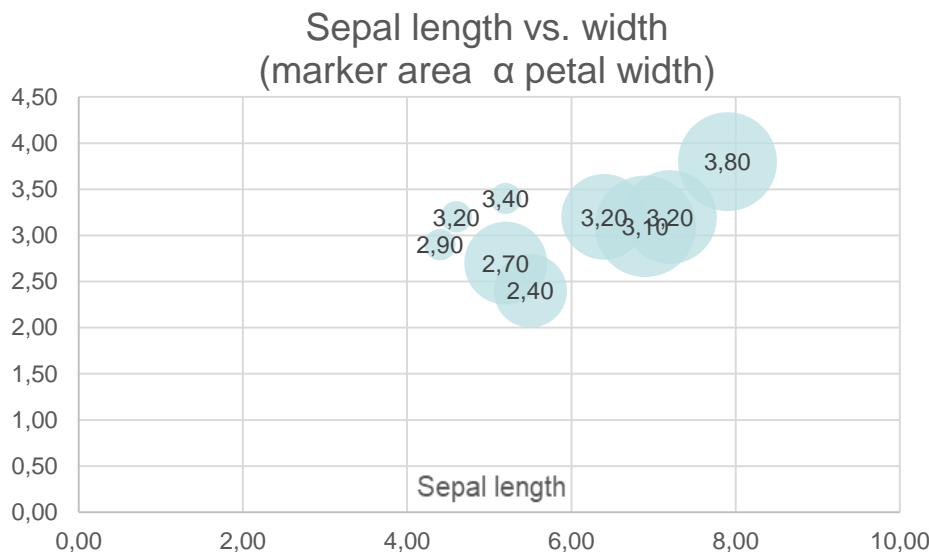
Exercise

- Draw bubble plots using Excel

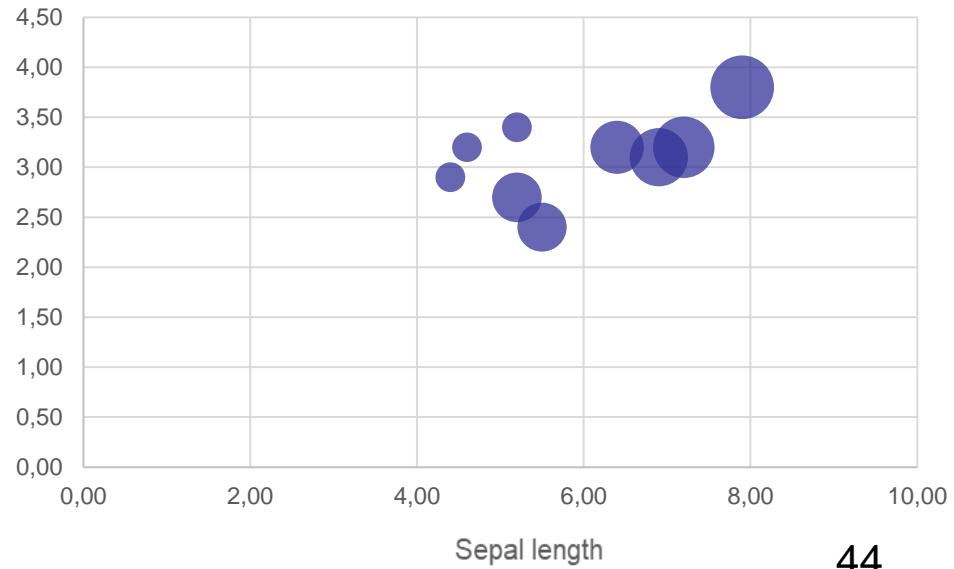
Iris dataset (a few observations):

	sepal length	sepal width	petal length	petal width	class
1	4,40	2,90	1,40	0,20	setosa
2	5,20	3,40	1,40	0,20	setosa
3	4,60	3,20	1,40	0,20	setosa
4	6,40	3,20	4,50	1,50	versicolor
5	5,20	2,70	3,90	1,40	versicolor
6	5,50	2,40	3,80	1,10	versicolor
7	7,20	3,20	6,00	1,80	virginica
8	7,90	3,80	6,40	2,00	virginica
9	6,90	3,10	5,40	2,10	virginica

Marker area proportional to petal length



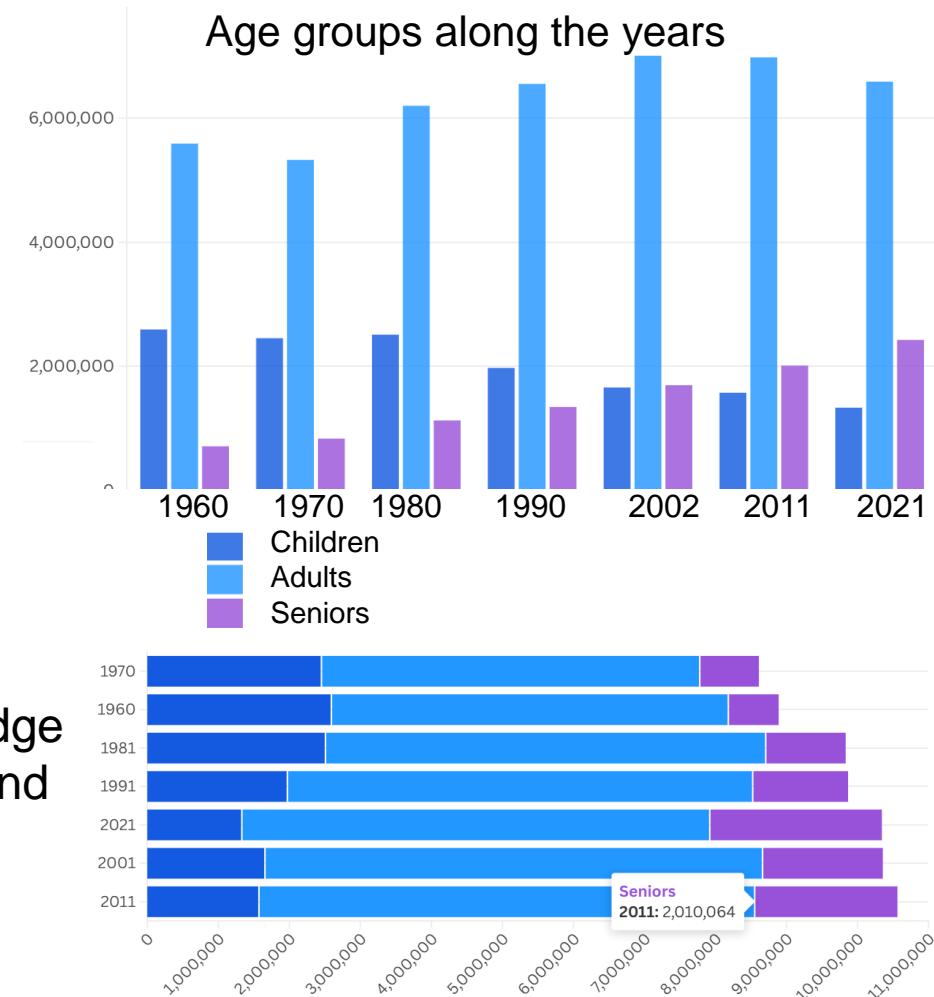
Sepal length vs. width
marker size α petal length



Other simple representations of trivariate data

- In a multiple **bar chart** data regarding all variables are represented side by side with the same axis (and scale) to comparability

Bars of different colour represent different variables/attributes (age groups) for different categories (year of the Census)



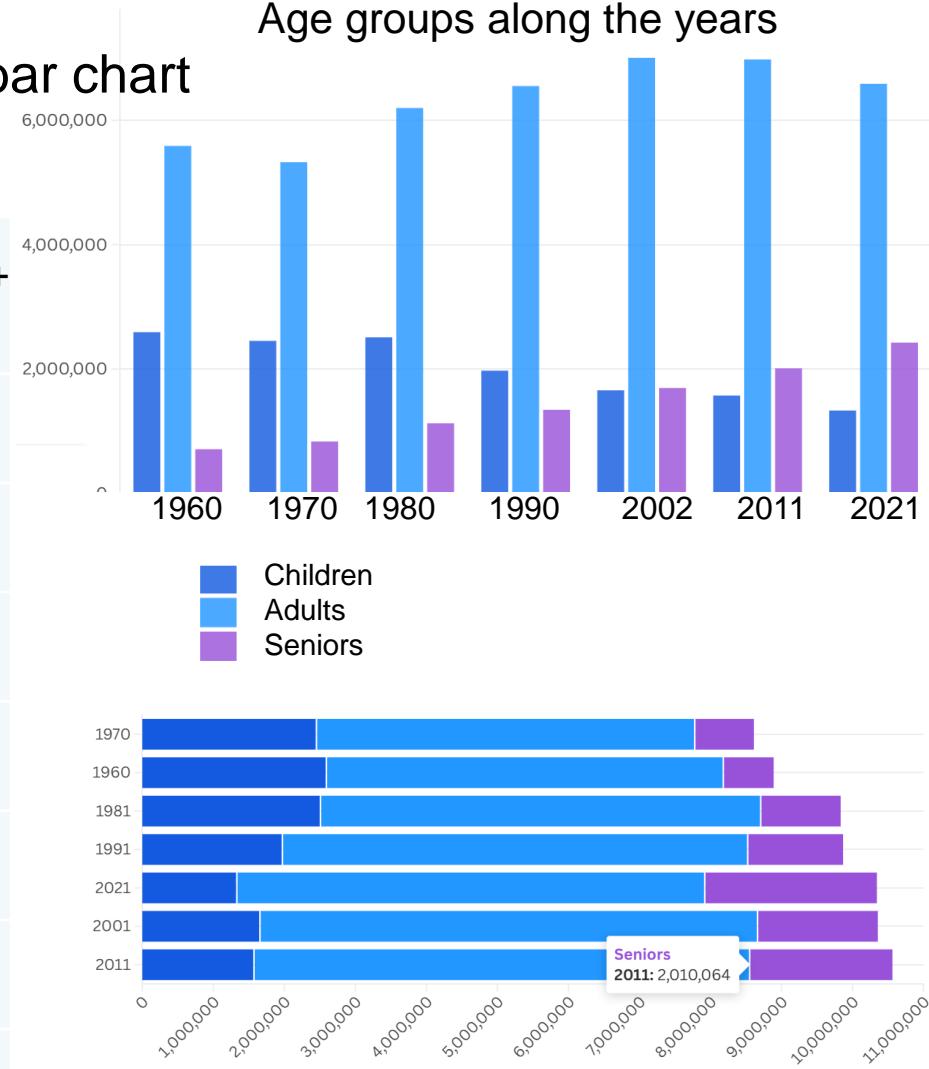
- In a **stacked bar chart** it is easy to judge the total value, but difficult to understand individual values (worse for more variables)

Exercise

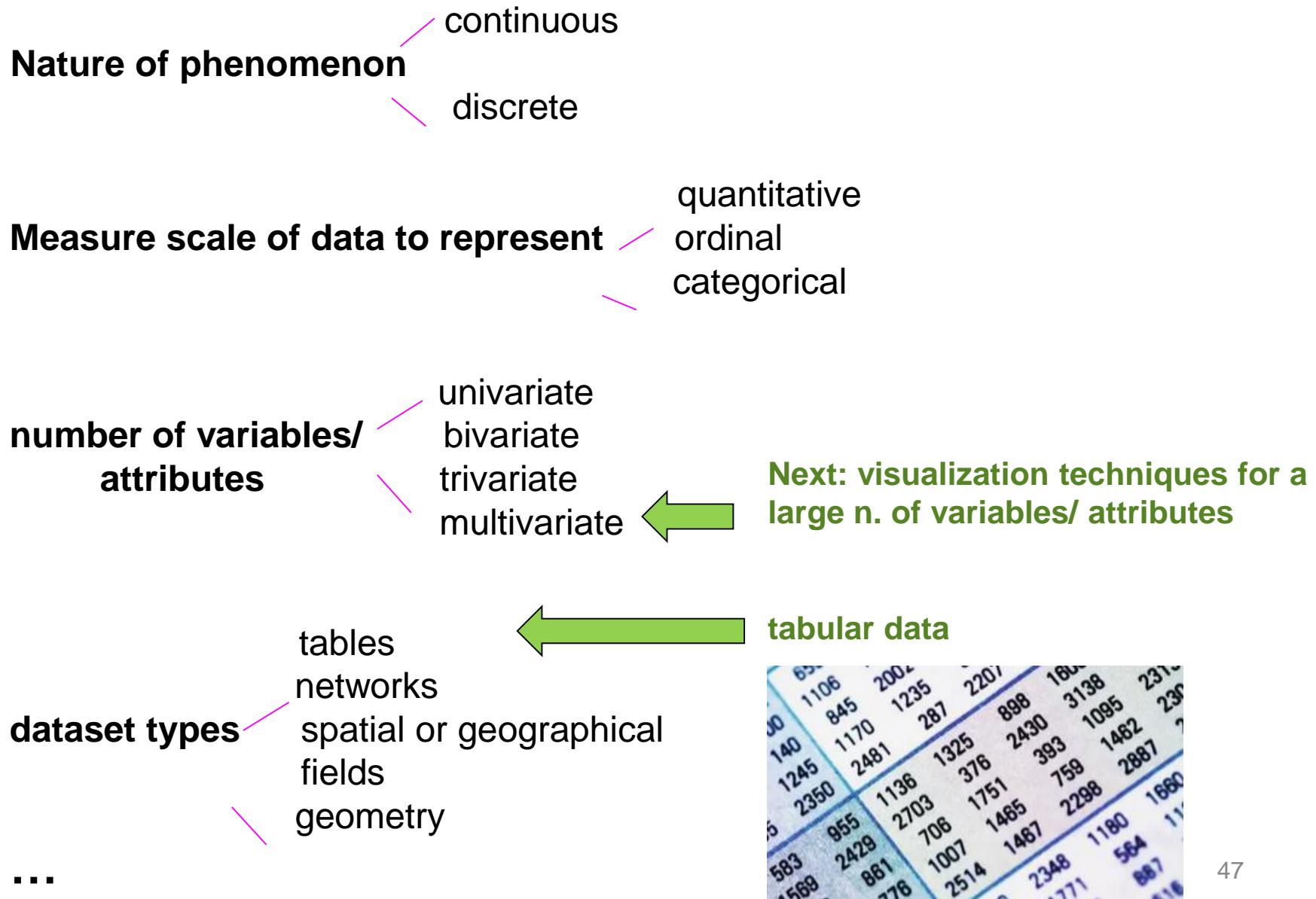
- Create a bar chart and a stacked bar chart of these data using [Flourish](#)

Census Year	Children: 0-14 years	Adults: 15-64 years	Seniors: 65 + years
1960	2591955	5588868	708569
1970	2451850	5326515	832760
1981	2508673	6198883	1125458
1991	1972403	6552000	1342744
2001	1656602	7006022	1693493
2011	1572329	6979785	2010064
2021	1331188	6588239	2423639

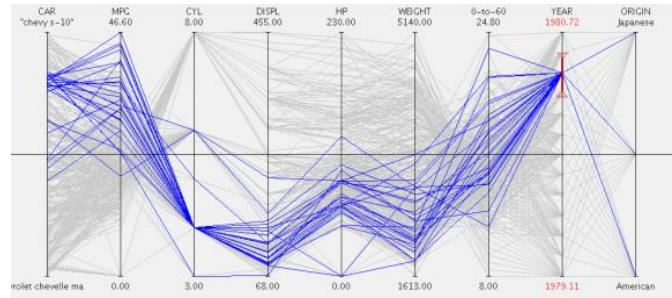
Age groups along the years



Important aspects to consider to select a Visualization technique:

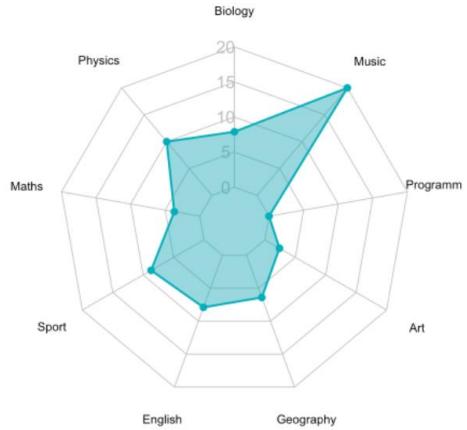


Techniques for Multivariate (or Hypervariate) data

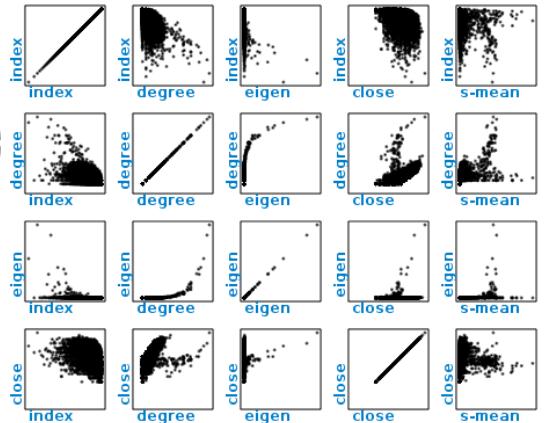


- Coordinate plots — parallel coordinate plots

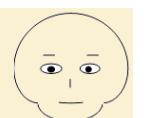
star (radar/spider) plots



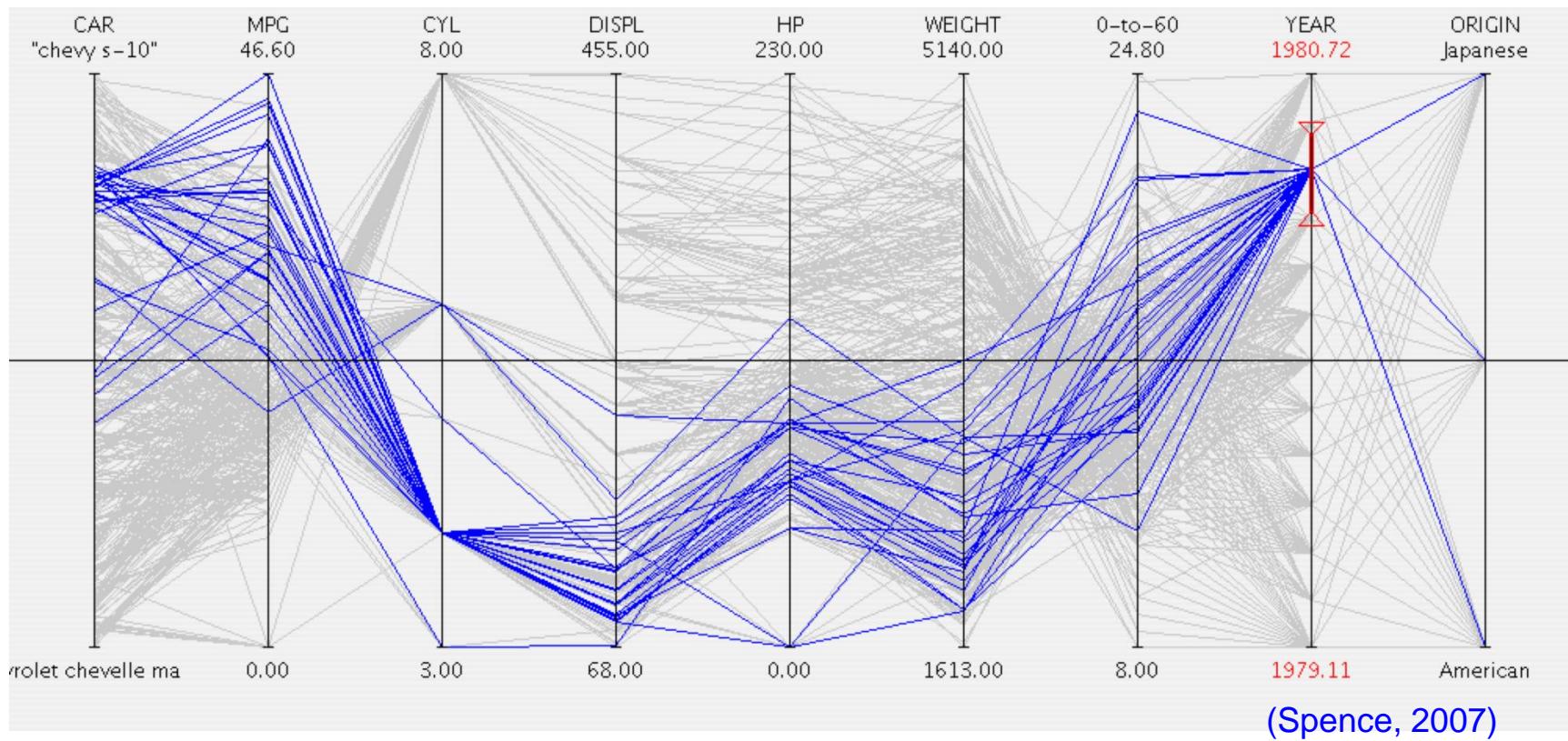
- Scatterplot Matrix



- Icons/glyphs

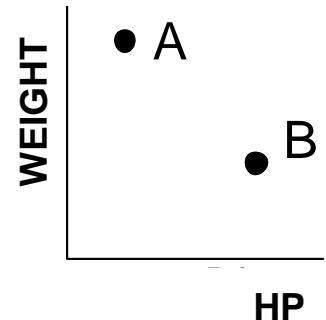


- **Parallel coordinates plots** are one of the most popular techniques for hypervariate data
- They have a very simple basis



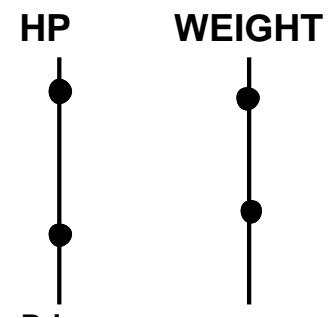
(Dataset with cars to sell with several interesting variables/attributes)

Consider a simple case of bivariate data
(data concerning cars to sell: HP, Weight)



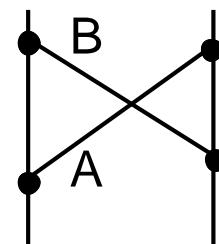
1- A scatterplot represents the horse power and weight associated with two cars

2- the axes are detached and made parallel; each car is represented by a point on each axis



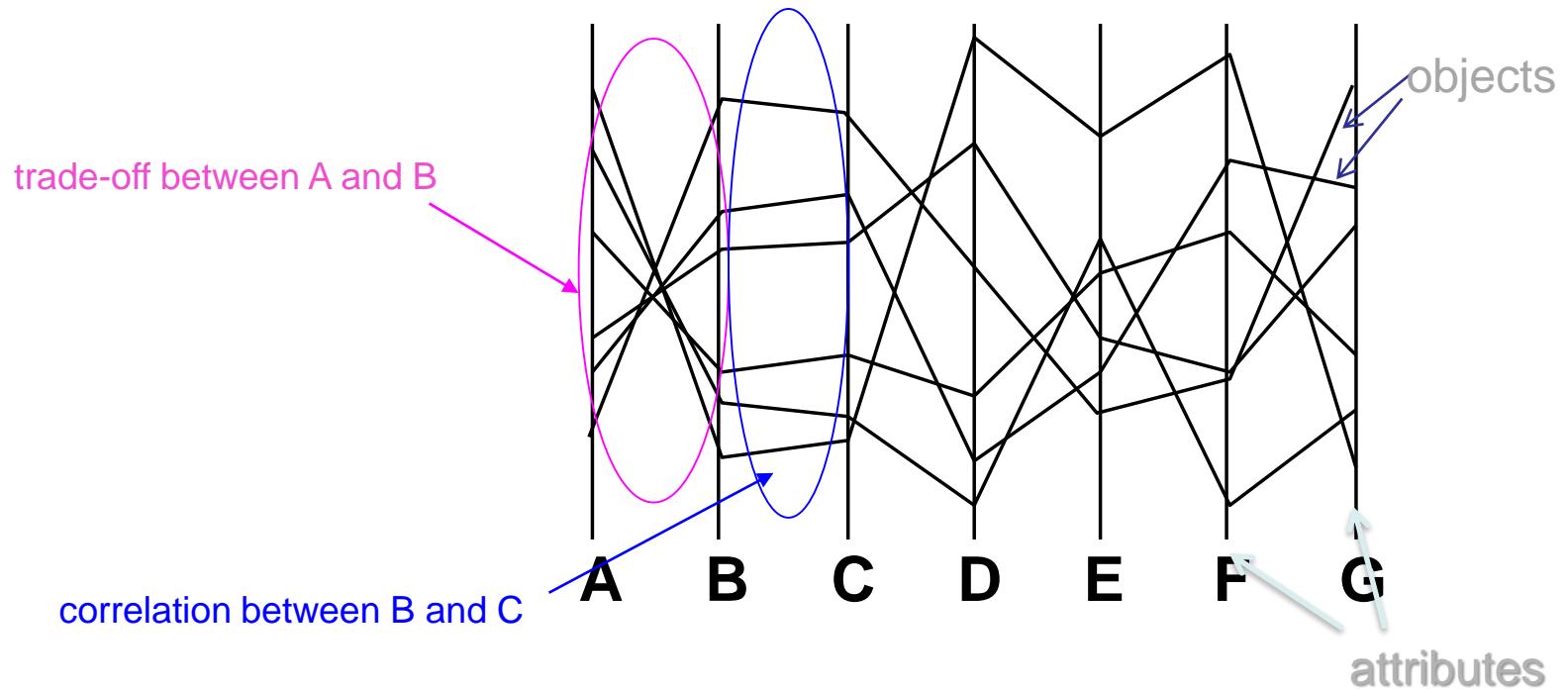
3- To avoid ambiguity the pair of points representing a car are joined and labeled

Each object/item (car) is represented by a polyline



- For objects characterized by many attributes the parallel coordinate plots offer many advantages

An example for six objects, each characterized by seven attributes:



The trade-off between A and B, and the correlation between B and C, are immediately apparent. The trade-off between B and E, and the correlation between C and G, are not.

Exercise

- 1- Draw a parallel coordinates representation of the following data:
(candidates to a position)
- 2- Draw a star plot
- 3- Ask a question better supported by the star plot

#	Education	Age	Experience	Salary	English	French
#	(MSc/PhD)	(years)	(years)	(euros)	(Bas/Ada)	(Y/N)
1	MSc	22	0	2500	Advanced	N
2	BSc	23	0	2500	Basic	N
3	MSc	24	1	2500	Advanced	Y
4	PhD	38	7	6000	Intermediate	Y
5	MSc	25	1	2500	Basic	N
6	PhD	35	5	5000	Advanced	Y
7	MSc	31	7	3000	Advanced	N
8	BSc	23	0	2500	Advanced	N
9	MSc	26	2	3000	Basic	Y
10	MSc	28	4	3000	Advanced	N

Exercise

Draw the parallel coordinates representation using [Examples | Flourish](#)

It is necessary to code the ordinal variables:

PhD-4 MSc-2 BSc-1 Advanced-4 Intermed-2 Basic-1 Y-1 N-0

Select the variables units to get similar values in all axes:

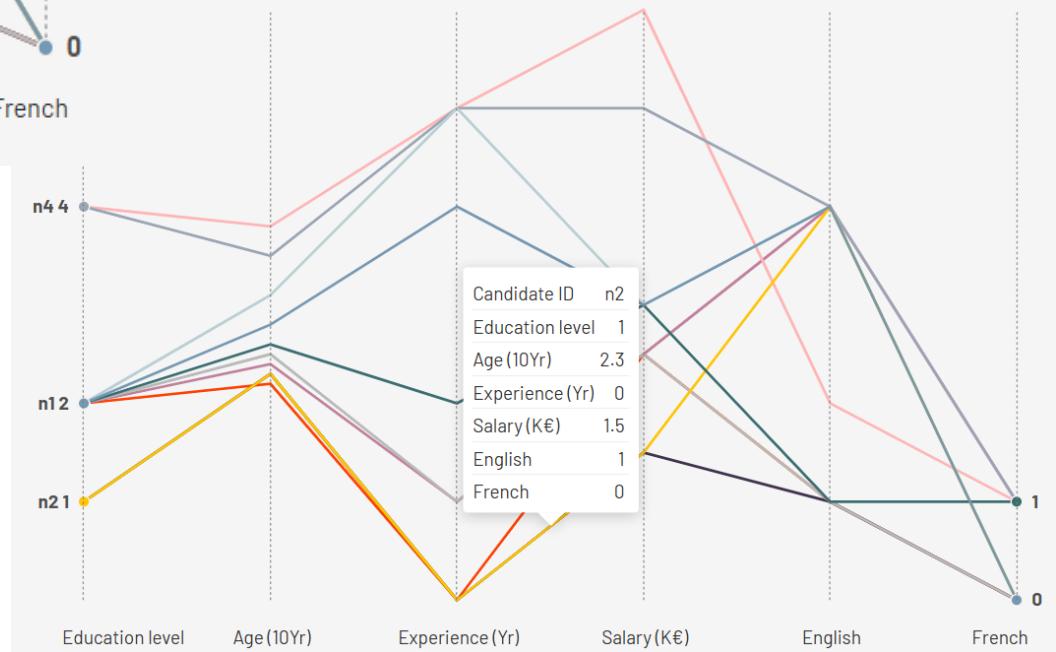
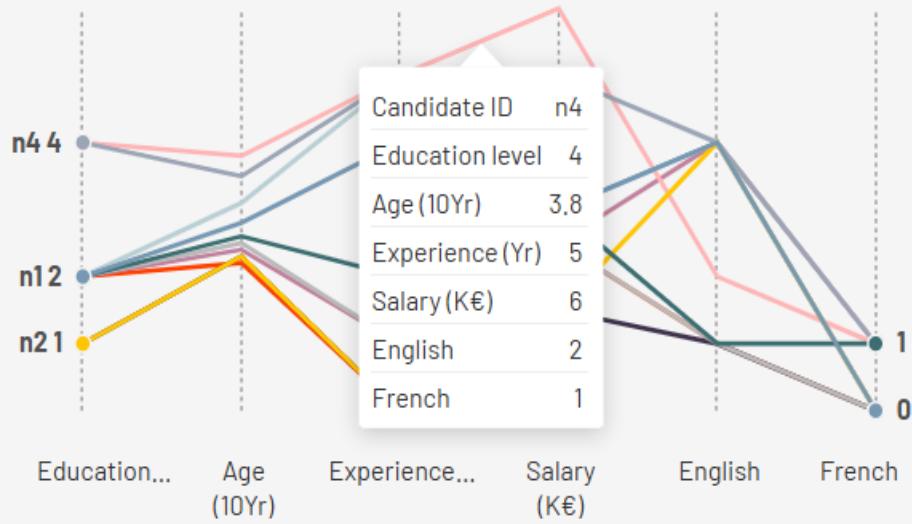
Candidates to a position



ABC	A	123	B	123	C	123	D	123	E	123	F	123	G
Candidate ID	Education level	Age (10Yr)	Experience (Yr)	Salary (KEuros)	English	French							
n1	2	2.2	0	2.5	1	0							
n2	1	2.3	0	1.5	1	0							
n3	2	2.4	1	2.5	4	1							
n4	4	3.8	5	6	2	1							
n5	2	2.5	1	2.5	1	0							
n6	4	3.5	5	5	4	1							
n7	2	3.1	5	3	4	0							
n8	1	2.3	0	1.5	4	0							
n9	2	2.6	2	3	1	1							
n10	2	2.8	4	3	4	0							

It is possible to have “details on demand” by user interaction

Candidates to a position

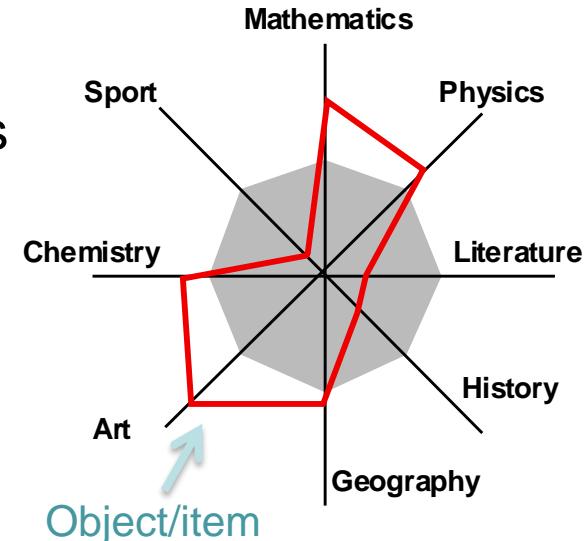


- **Star (aka radar/spider) plots** have many features in common with parallel coordinate plots

- A variable/attribute value is represented by a point on a coordinate axis

- Attribute axes radiate from a common origin

- For a given object/item, points are joined by straight lines (making a closed polygon)
- Other useful information such as average values or thresholds can be visually encoded



(Spence, 2007)

All units in per 90

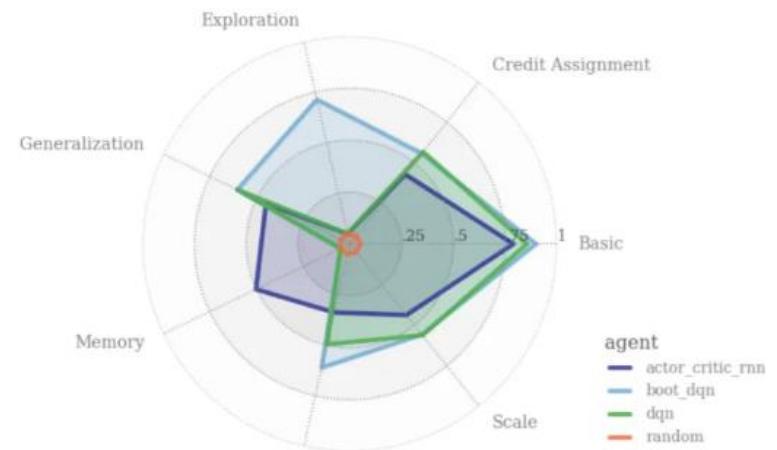
Messi vs. Joe Average

Barcelona

Season: 2012-13

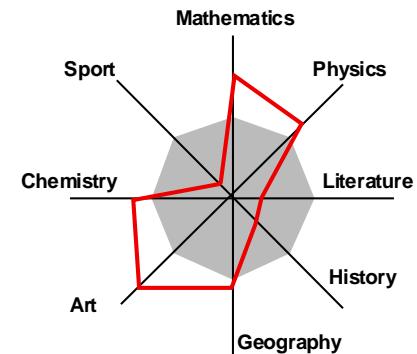
Age: 25

Time played: 29.1 90's



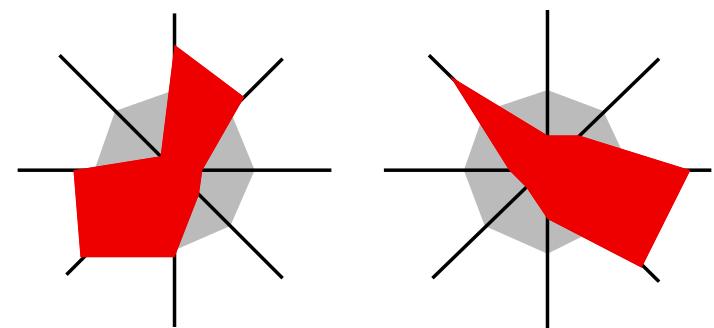
[DeepMind 'Bsuite' Evaluates Reinforcement Learning Agents | Synced](#)

Properties of star plots:



- Their shape can provide a reasonably rapid appreciation of the attributes/variables of the objects/items
- They offer **object visibility** and are suitable to compare objects

(by visibility it is meant the ability to gain insight pre-attentively; without a great cognitive effort)



Bob's performance

Tony's performance

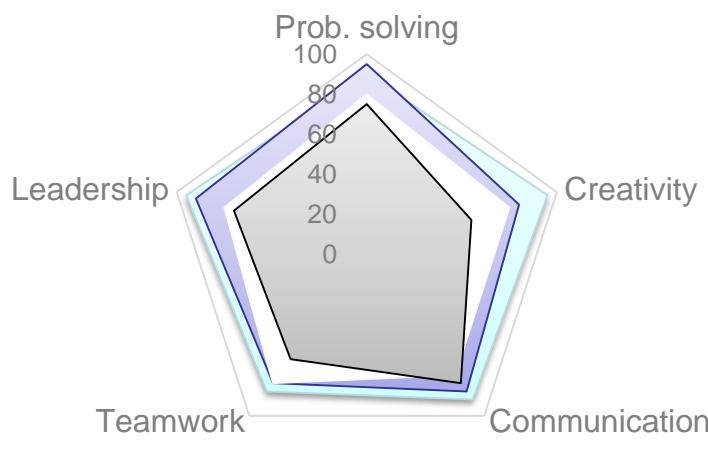
(Spence, 2007)

Exercise

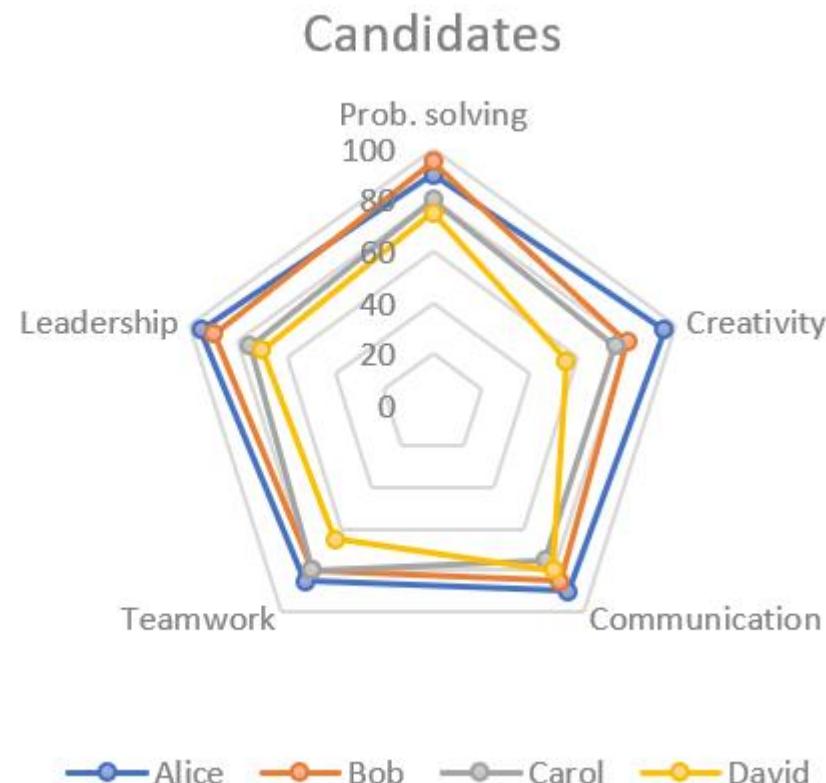
- Create a radar plot for this Human Resources case using Excel

	Prob. solving	Creativity	Communication	Teamwork	Leadership
Alice	90	95	90	85	95
Bob	95	80	85	80	90
Carol	80	75	75	80	75
David	75	55	80		

Candidates



■ Alice ■ Bob ■ Carol ■ David

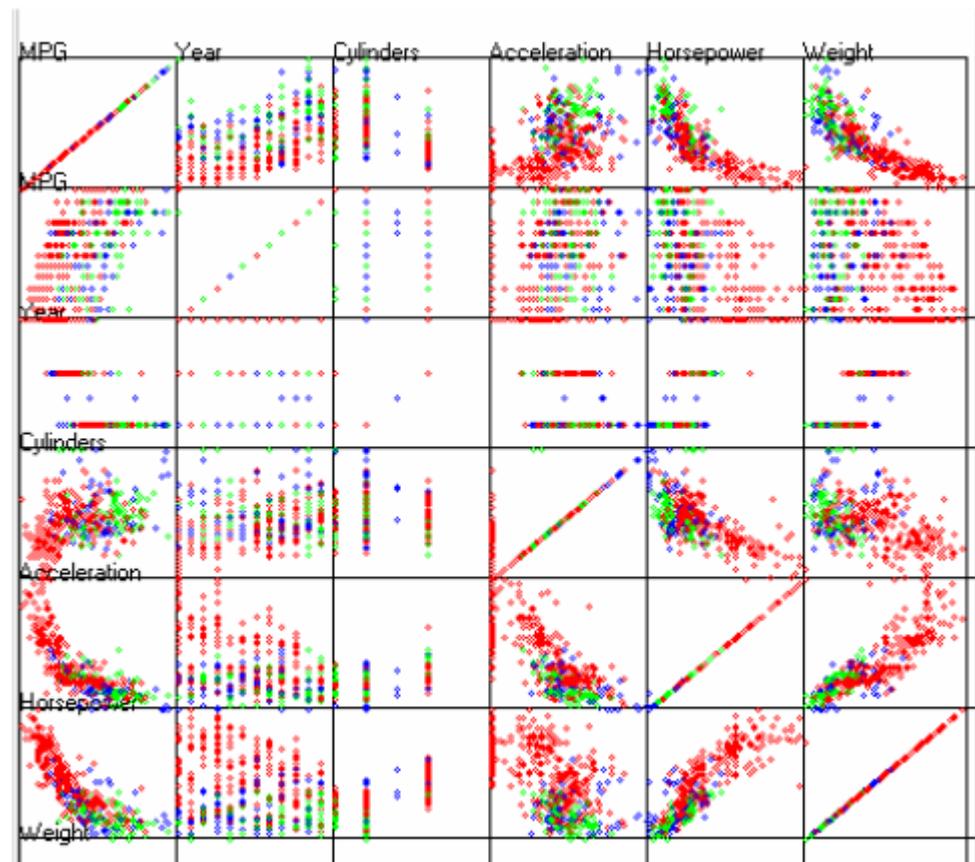


—●— Alice —●— Bob —●— Carol —●— David

- The **scatterplot matrix** (SPLOM) is applicable to higher n. of variables
- However, as the number of attributes increases, the number of different pairs of attributes increases rapidly:

- 2 variables-> 1 scatterplot
- 3 variables -> 3 scatterplots
- 4 variables -> 6 scatterplots

We may try to reduce the number of dimensions keeping the more relevant:
Dimensionality reduction!



Scatterplot matrix for 6 attributes of a car dataset

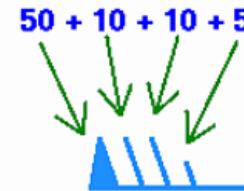
Windy: Wind map & weather forecast

Glyphs

The properties of the shape represent different variables sized according to the associated quantitative value and possibly using color



This example shows Chernoff faces for lawyers' ratings of twelve judges [Chernoff face - Wikipedia](#)



Wind blowing from the west at 75 knots



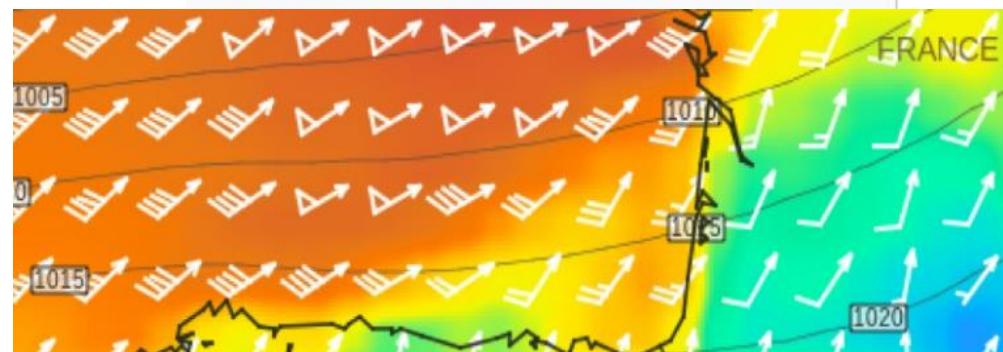
Wind blowing from the northeast at 25 knots



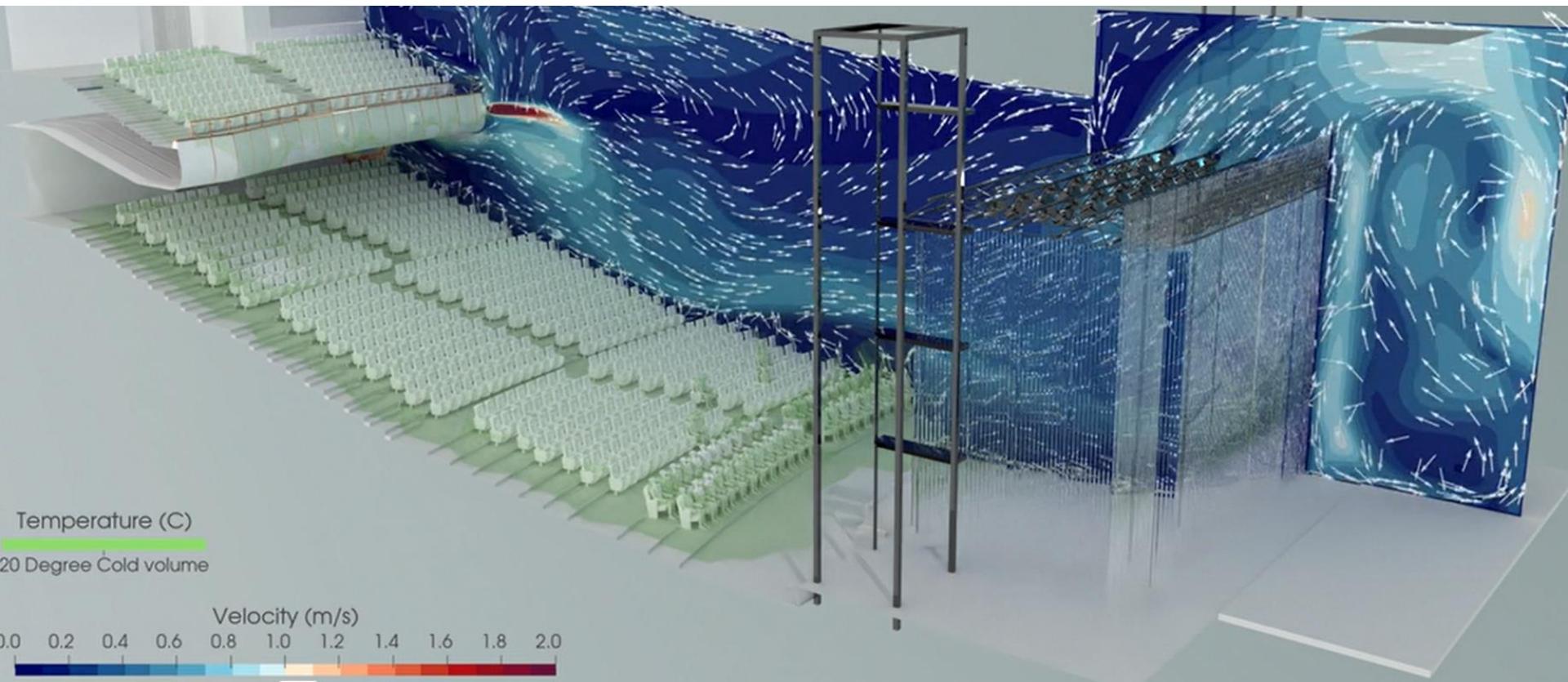
Wind blowing from the south at 5 knots



Calm winds

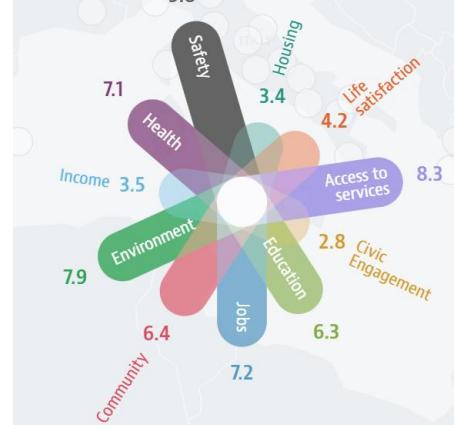


Oriented glyph and streamlines showing the overall direction of flow within a flow dynamics data visualization

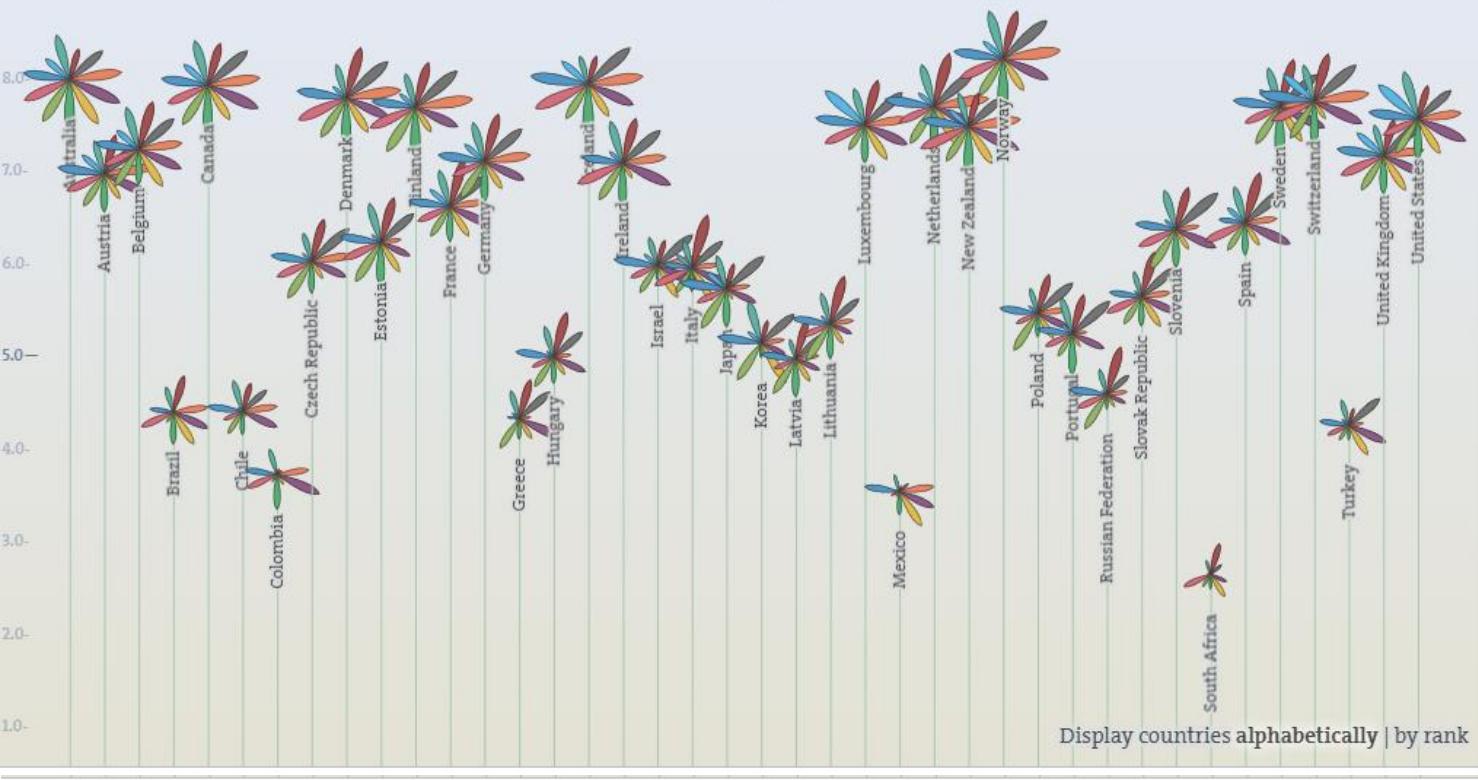


[ParaView for Computational Fluid Dynamics](#)

The physical properties of the shape represent different variables sized according to the associated quantitative value and distinguished through color



In Lisbon



Create Your Better Life Index

Rate the topics according to their importance to you:

Housing	<input type="range"/>
Income	<input type="range"/>
Jobs	<input type="range"/>
Community	<input type="range"/>
Education	<input type="range"/>
Environment	<input type="range"/>
Civic Engagement	<input type="range"/>
Health	<input type="range"/>
Life Satisfaction	<input type="range"/>
Safety	<input type="range"/>
Work-Life Balance	<input type="range"/>

Reset Help

Important aspects to consider to select a Visualization technique:

Nature of phenomenon

- continuous
- discrete

Measure scale of data to represent

- quantitative
- ordinal
- categorical

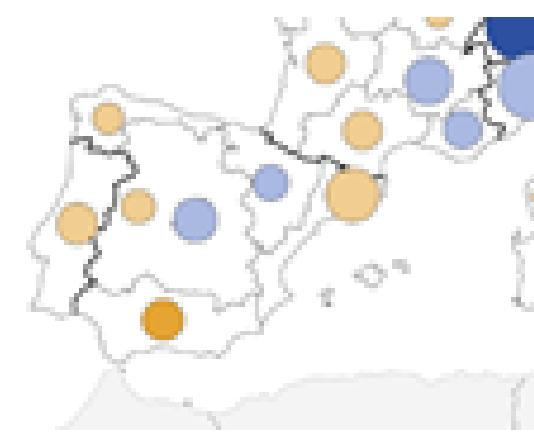
number of attributes

- univariate
- bivariate
- trivariate
- multivariate

dataset types

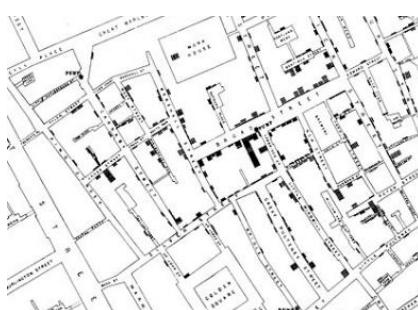
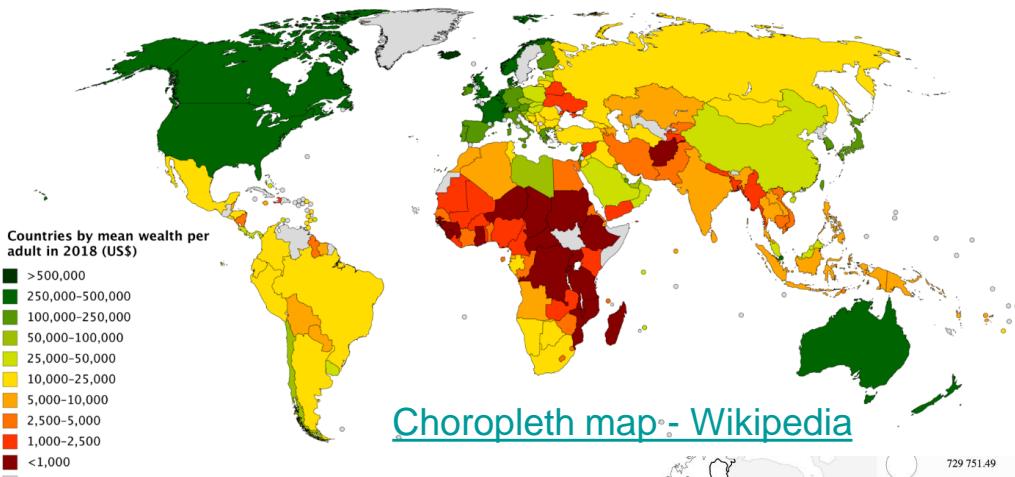
- tables
- networks
- spatial or geographical fields
- geometry

maps

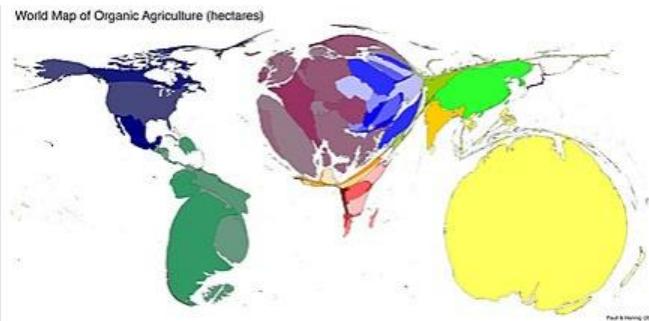


Representing a special type of data: maps

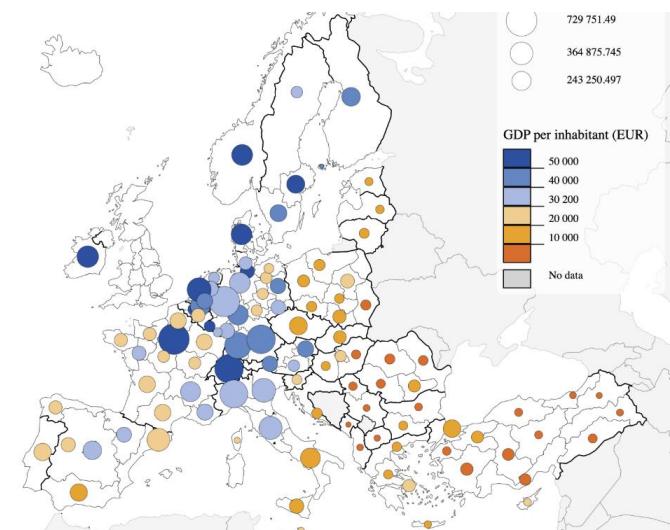
- **Maps** are common visualizations of **georeferenced data** (latitude, longitude, value)
- Several types of maps:
 - Choropleth maps
 - Scaled symbol maps
 - Dot maps
 - Cartograms, ...



[Dot distribution map - Wikipedia](#)



[Cartogram - Wikipedia](#)

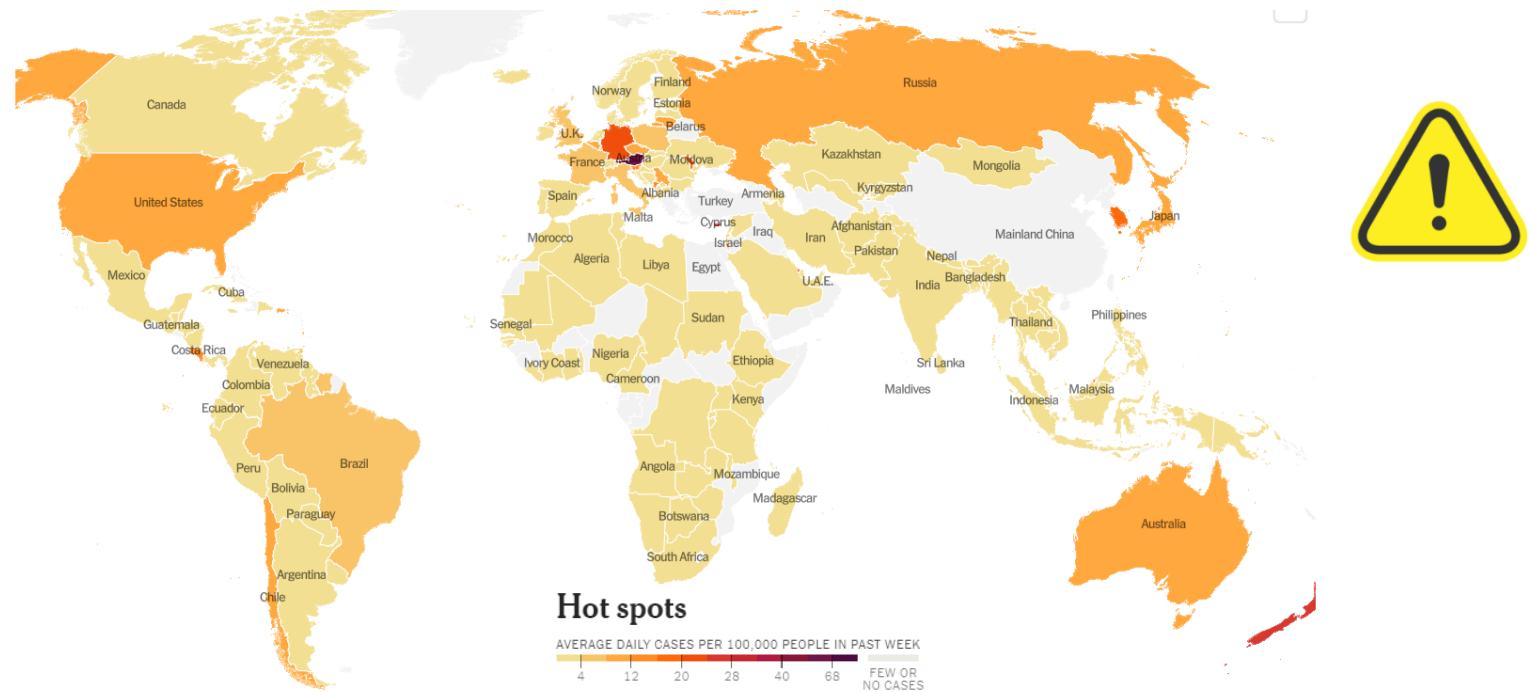


[Scaled symbol maps](#)

Simple representations of attributes on a map

Choropleth maps - A standard approach to communicating aggregated data by geographical areas using color encoding of the geographic area

They require some care: what are the possible issues?

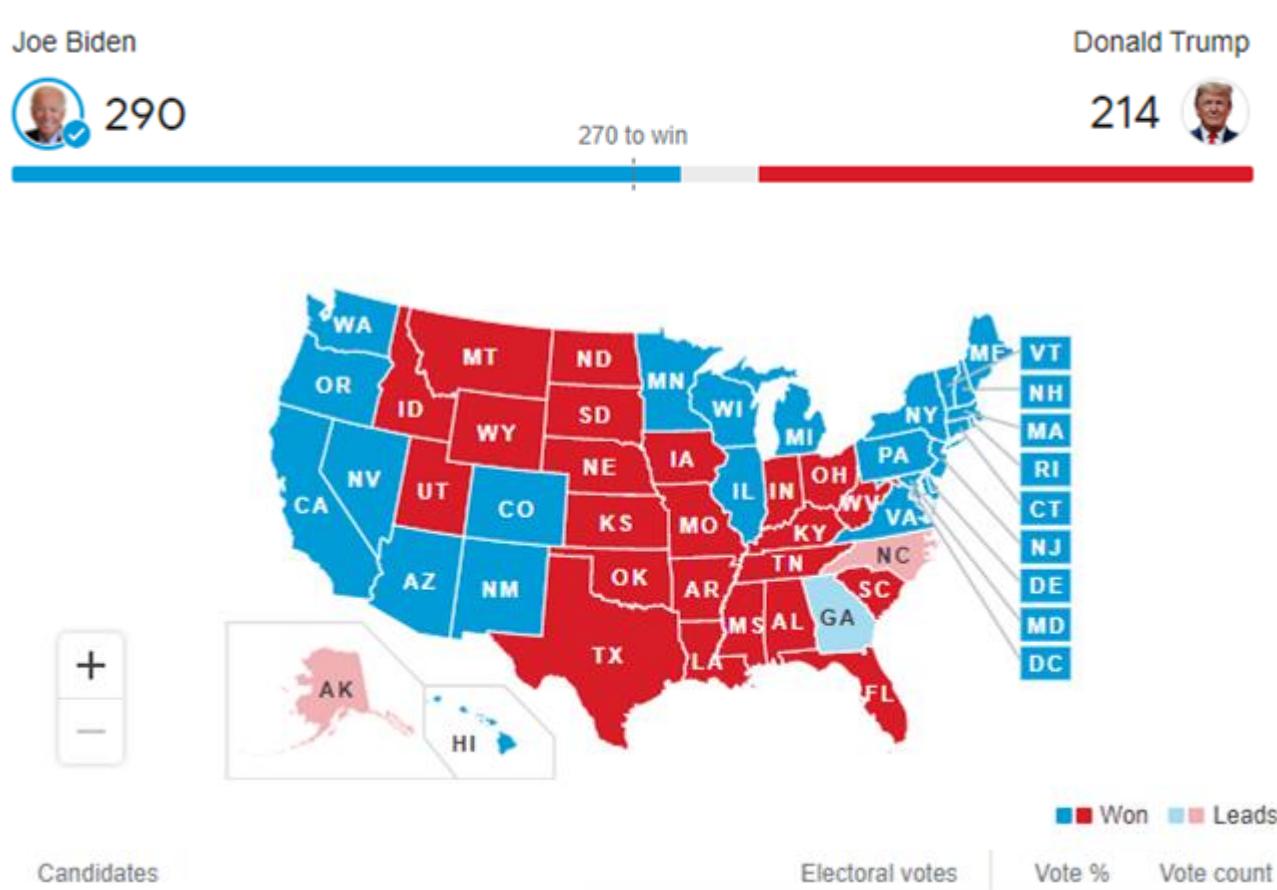


March 1/2023

Visualizations of the US 2020 Election

(choropleth + bar)

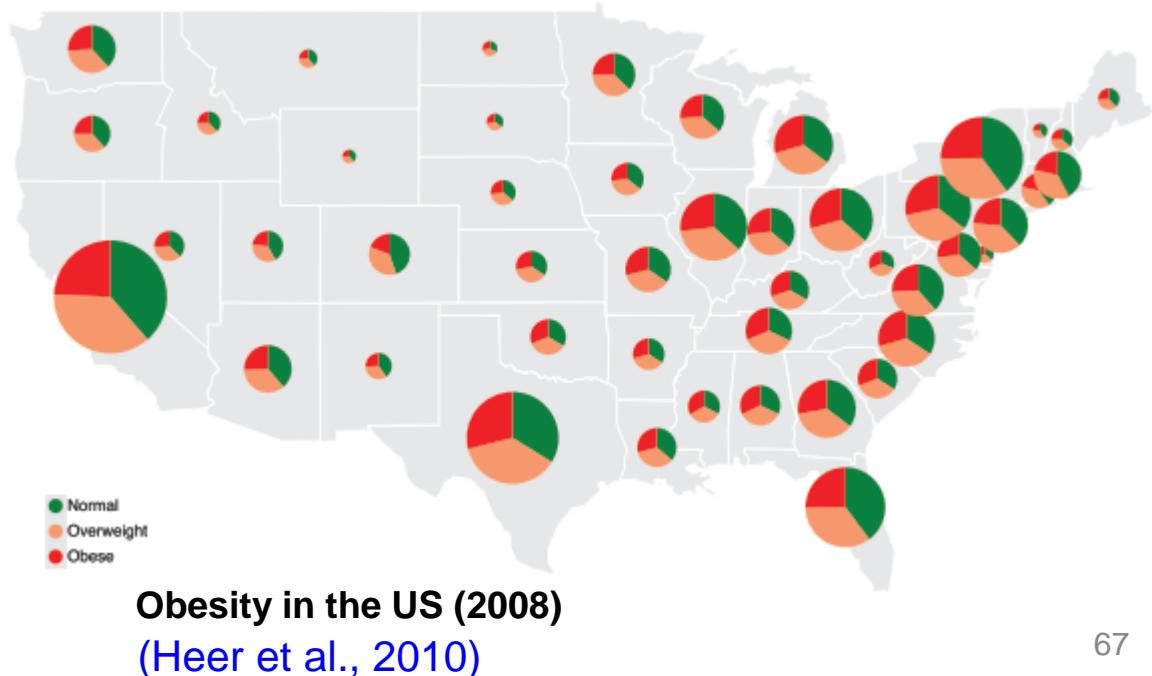
the bar helps better understand the ratio of votes
(and the number provides the maximum accuracy)



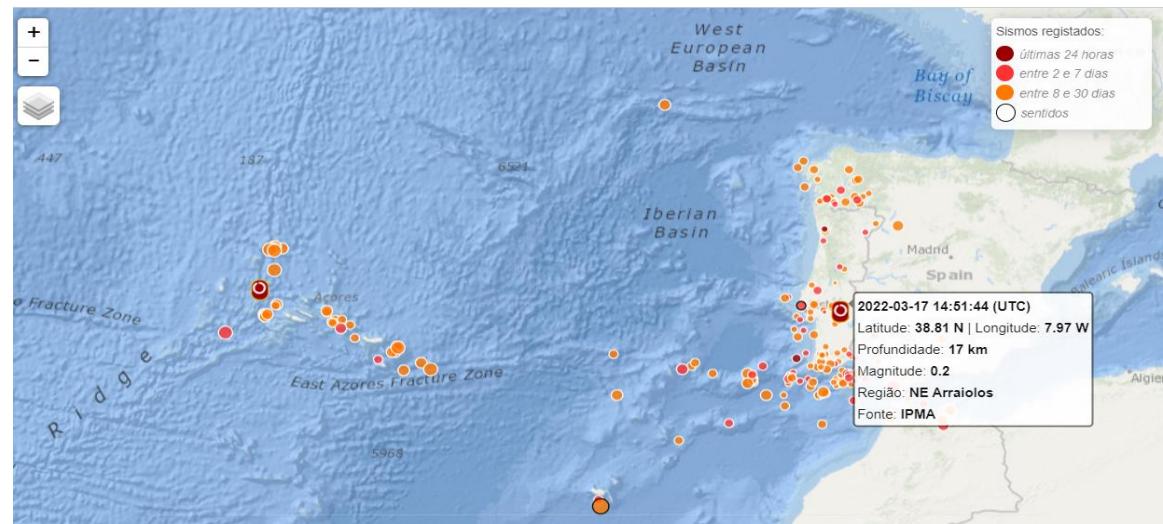
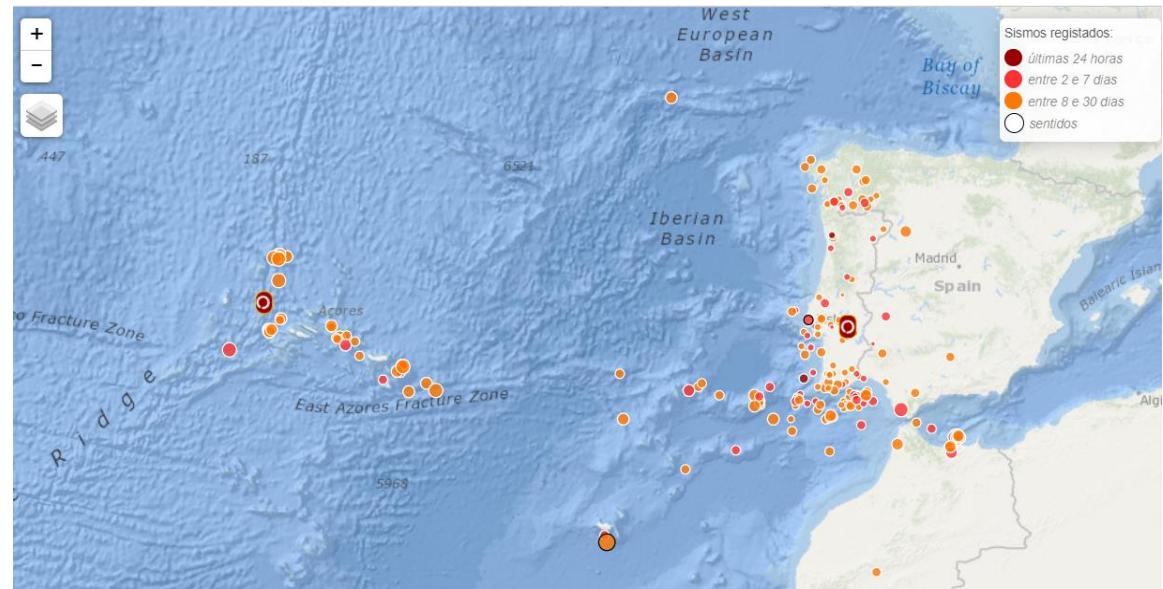
Simple representations of attributes on a map

- **Scaled Symbol Maps** are an alternative to the choropleth map;
- Symbols are placed over an underlying map; may show more attributes
- Avoid confounding geographic area with data values

What is missing in this visualization?



Example:
Seismic activity:
Is something missing in
this visualization?



IPMA - Mapa sismicidade

Cartograms:
Using distortion to represent a variable

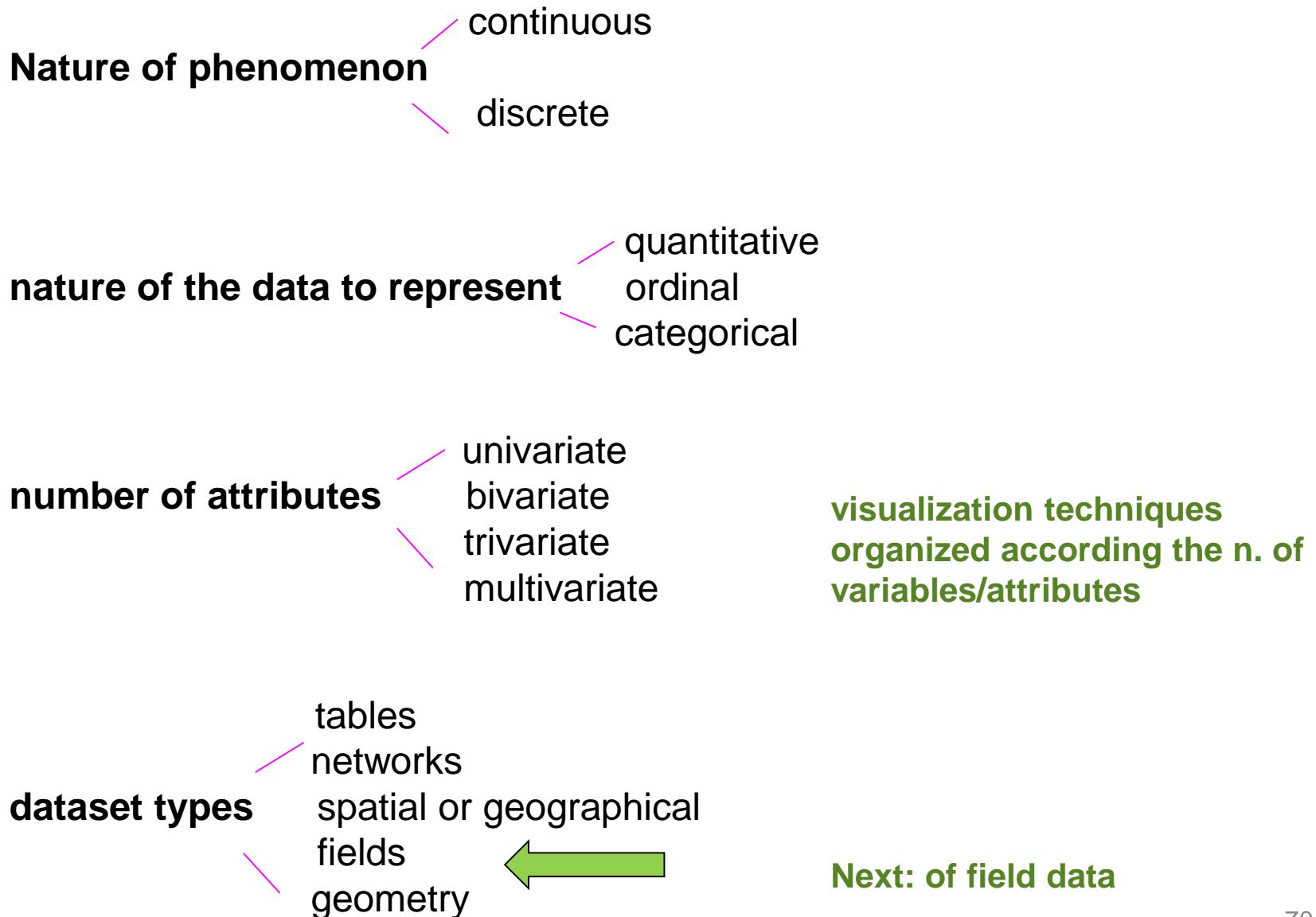
Population Increase 1973 to 2022

WORLD
MAPPER
www.worldmapper.org

- North America
- Central America
- Caribbean
- South America
- Northern Europe
- Central Europe
- Eastern Europe
- Southern Europe
- Southern Africa
- Northern Africa
- Western Africa
- Central Africa
- Eastern Africa
- Russia
- Central Asia
- Western Asia
- Southern Asia
- Eastern Asia
- Australasia
- Southeastern Asia
- Pacific Oceania

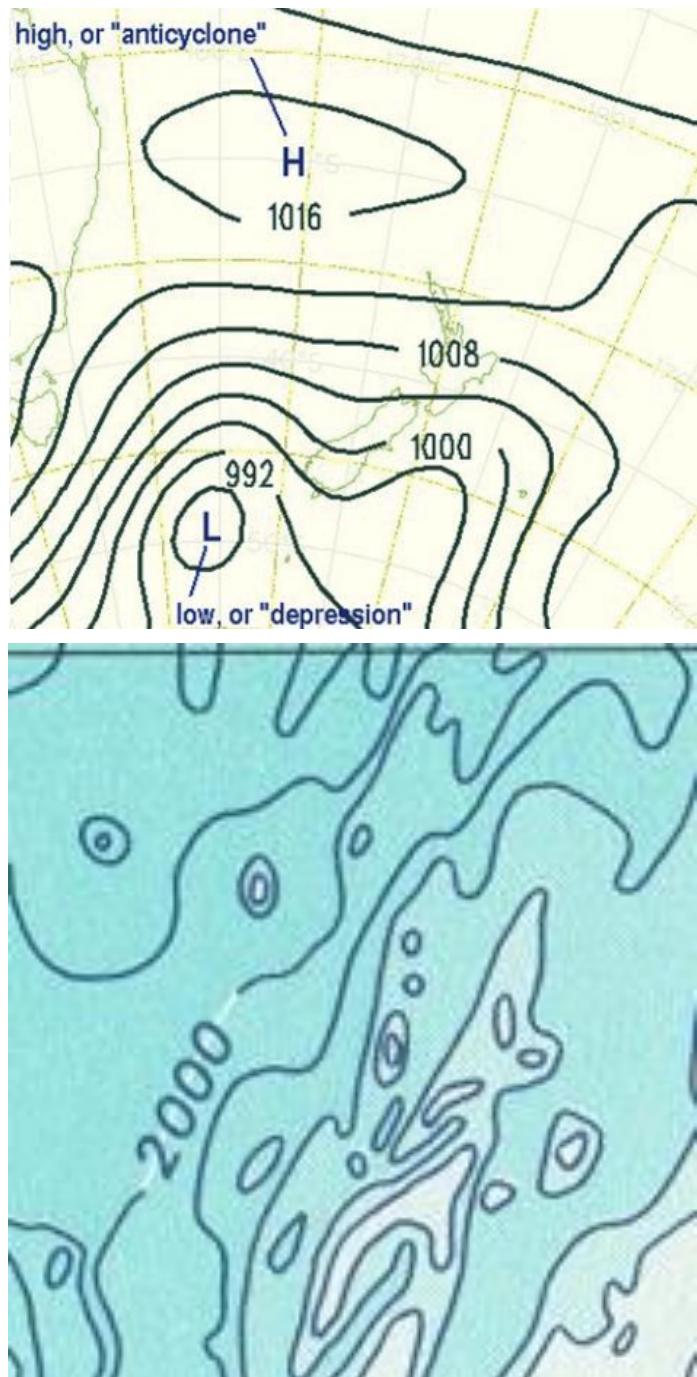


Important aspects to consider to select a Visualization technique:



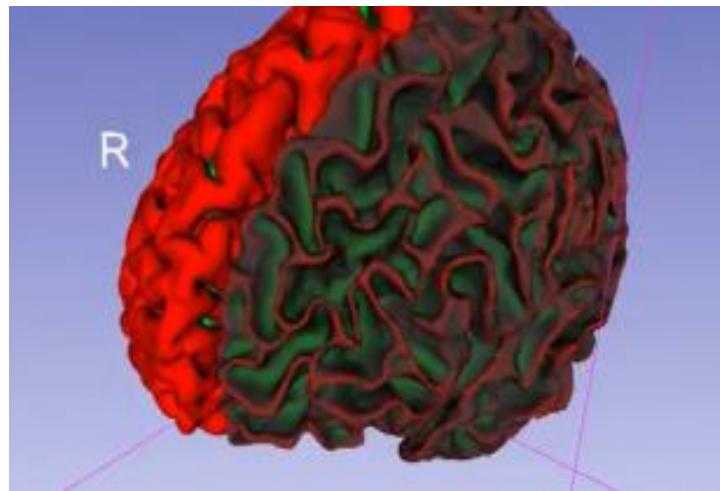
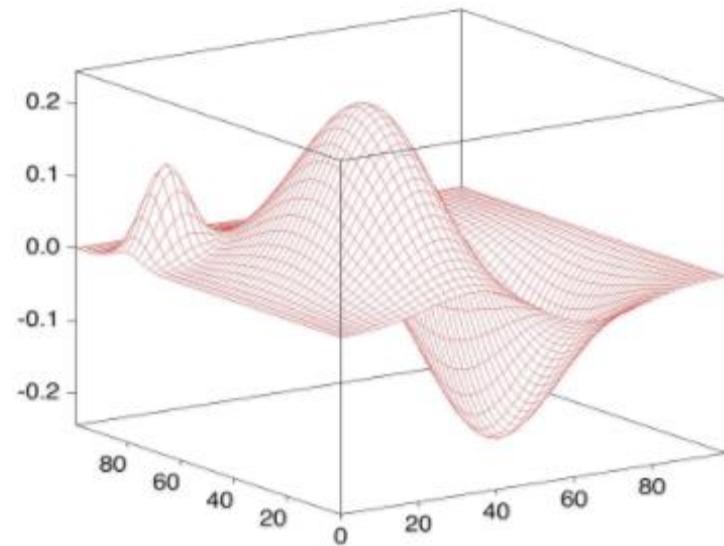
Representing a scalar in a 2D continuous field

- Contour plots
- **contour line** (also **isoline**, **isopleth**, or **equipotential curve**) of a function of two variables is a curve along which the function has a constant value, so that the curve joins points of equal value.
- Typical in meteorological charts (isobars and isothermal curves)
- and maps (to represent altitude or depth)

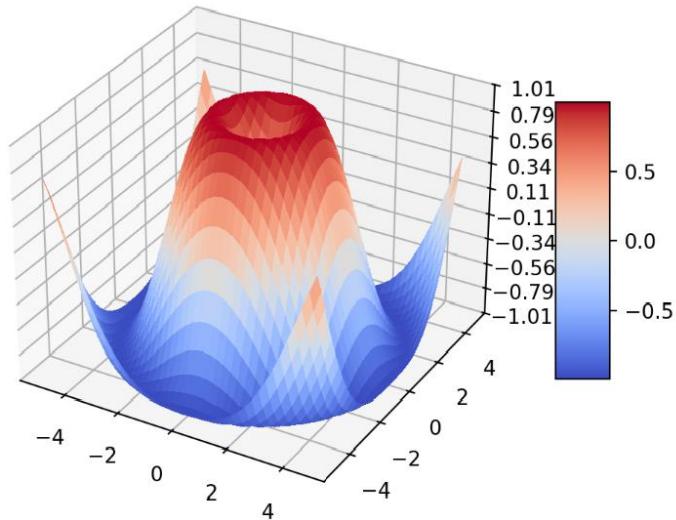


Representing a 3D scalar function and scalar volumetric data

- Surface plots may be combined with color
(preferably in a redundant way and carefully selecting the scale)
- Volumetric data

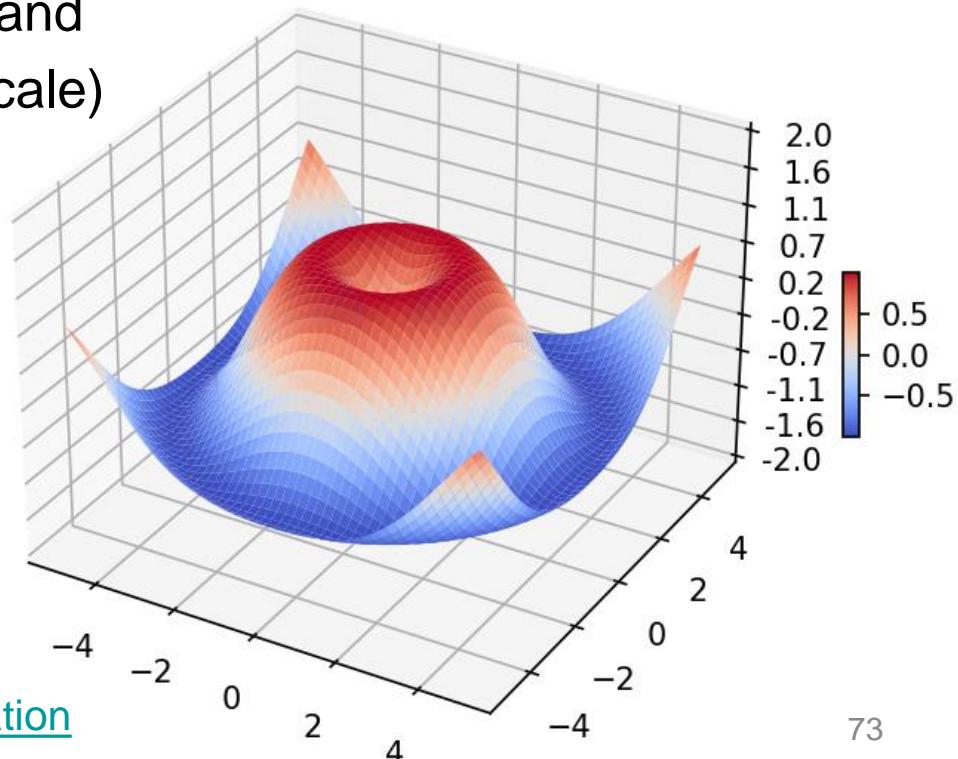
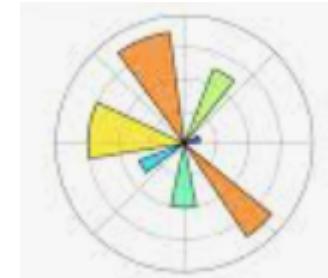


[3D plotting — Matplotlib 3.10.0 documentation](#)
[Plot \(graphics\) – Wikipedia](#)
[3D Slicer image computing platform | 3D Slicer](#)



Exercise

- Represent a 3D scalar function using Matplotlib
- Surface plots may be combined with colour
(preferably in a redundant way and
carefully selecting the colour scale)



```

# plot of 3D function

import matplotlib.pyplot as plt
import numpy as np
from matplotlib import cm
from matplotlib.ticker import LinearLocator

fig, ax = plt.subplots(subplot_kw={"projection": "3d"})


# data.
X = np.arange(-5, 5, 0.01)
Y = np.arange(-5, 5, 0.01)
X, Y = np.meshgrid(X, Y)
R = np.sqrt(X**2 + Y**2)
Z = np.sin(R)

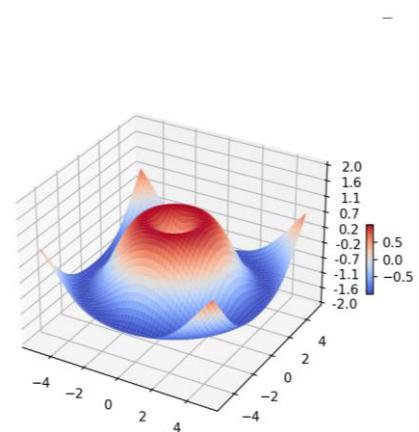
# Plot the surface.
surf = ax.plot_surface(X, Y, Z, cmap=cm.coolwarm,
                       linewidth=0, antialiased=True)

# Customize the z axis.
ax.set_zlim(-2, 2)
ax.zaxis.set_major_locator(LinearLocator(10))
# A StrMethodFormatter is used automatically
ax.zaxis.set_major_formatter('{x:.0lf}'

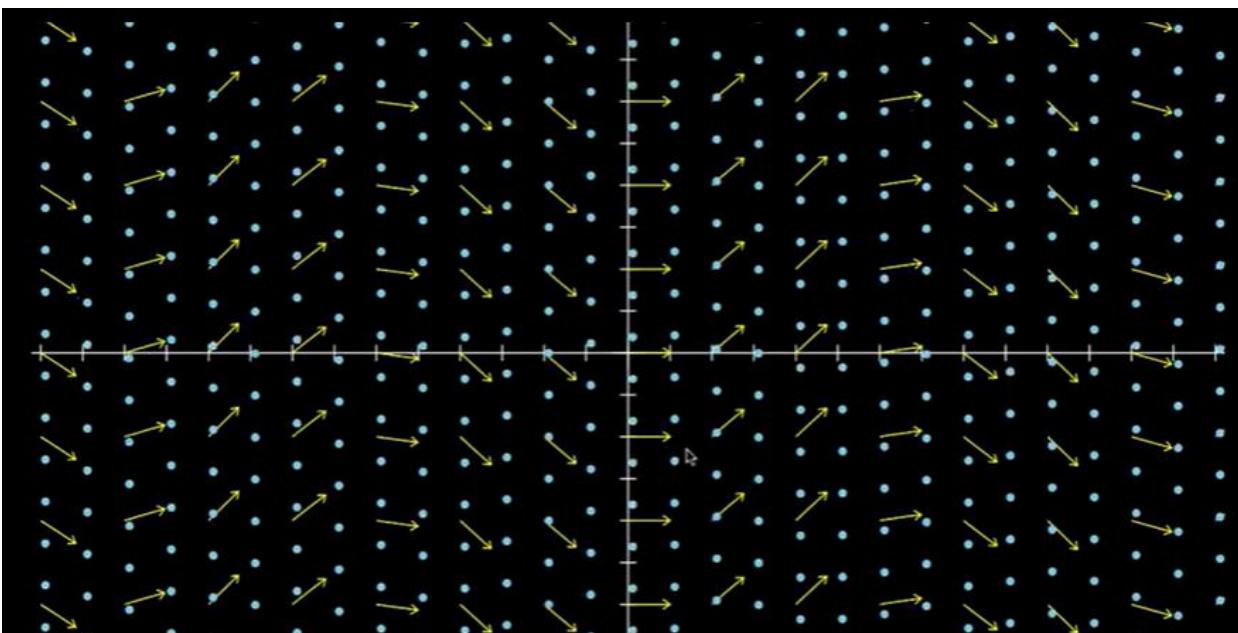
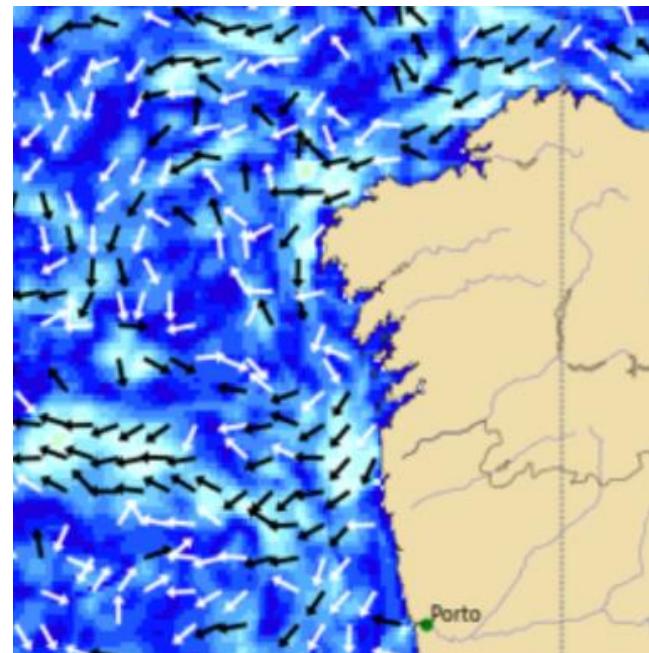
# Add a color bar which maps values to colors.
fig.colorbar(surf, shrink=0.2, aspect=10)

plt.show()

```



Representing vectors in a 2D field



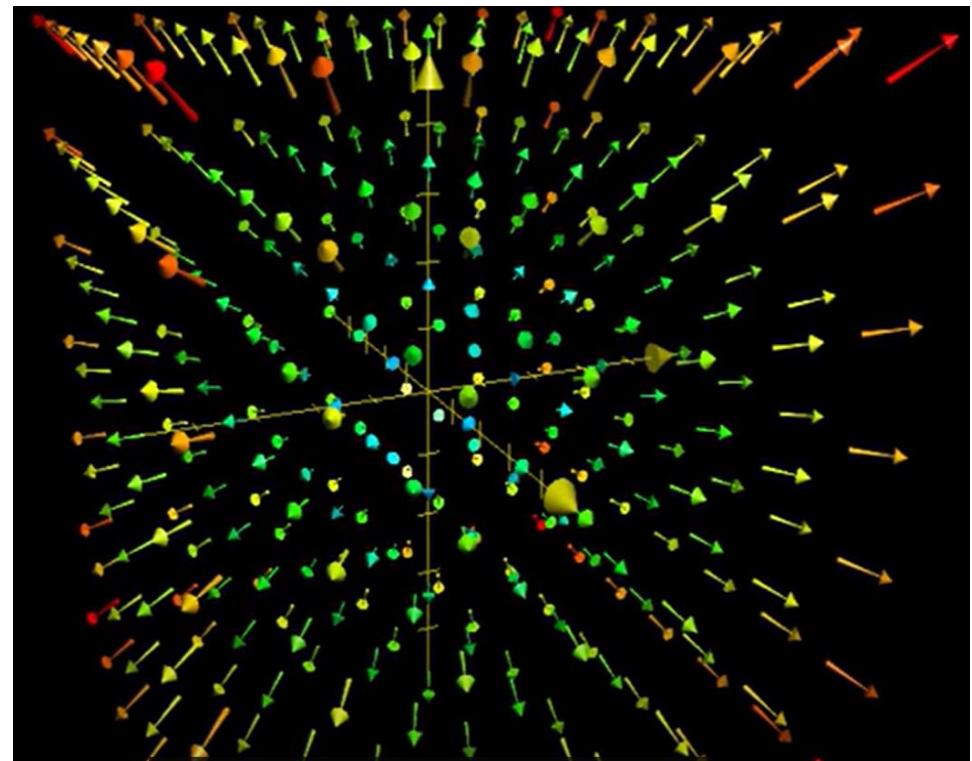
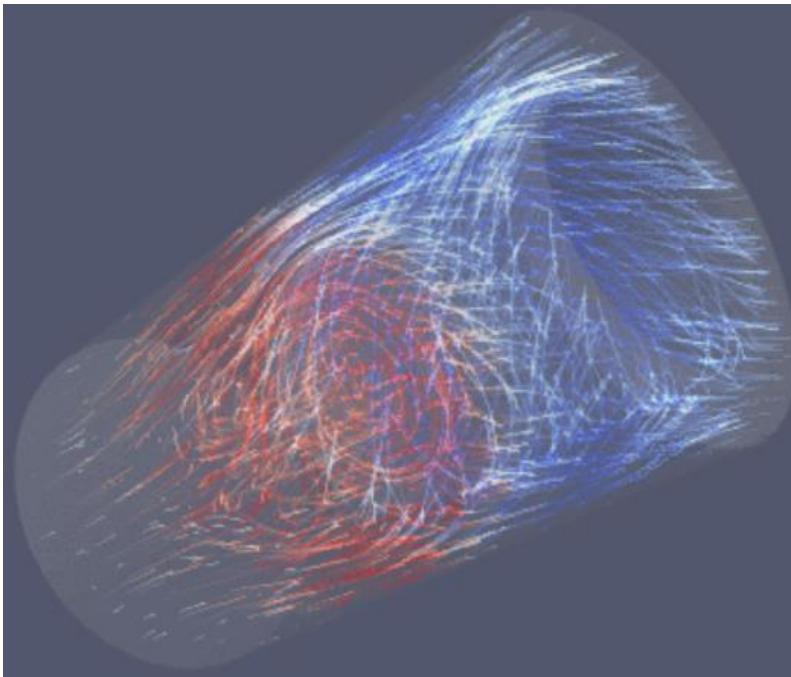
[IPMA - Corrente à superfície](#)

[Fluid flow and vector fields \(video\)](#) | Khan Academy

Representations of vectors in a 3D field

Vector data visualization: in 3D space a vector variable is visualized

Stream Lines are another interesting representation



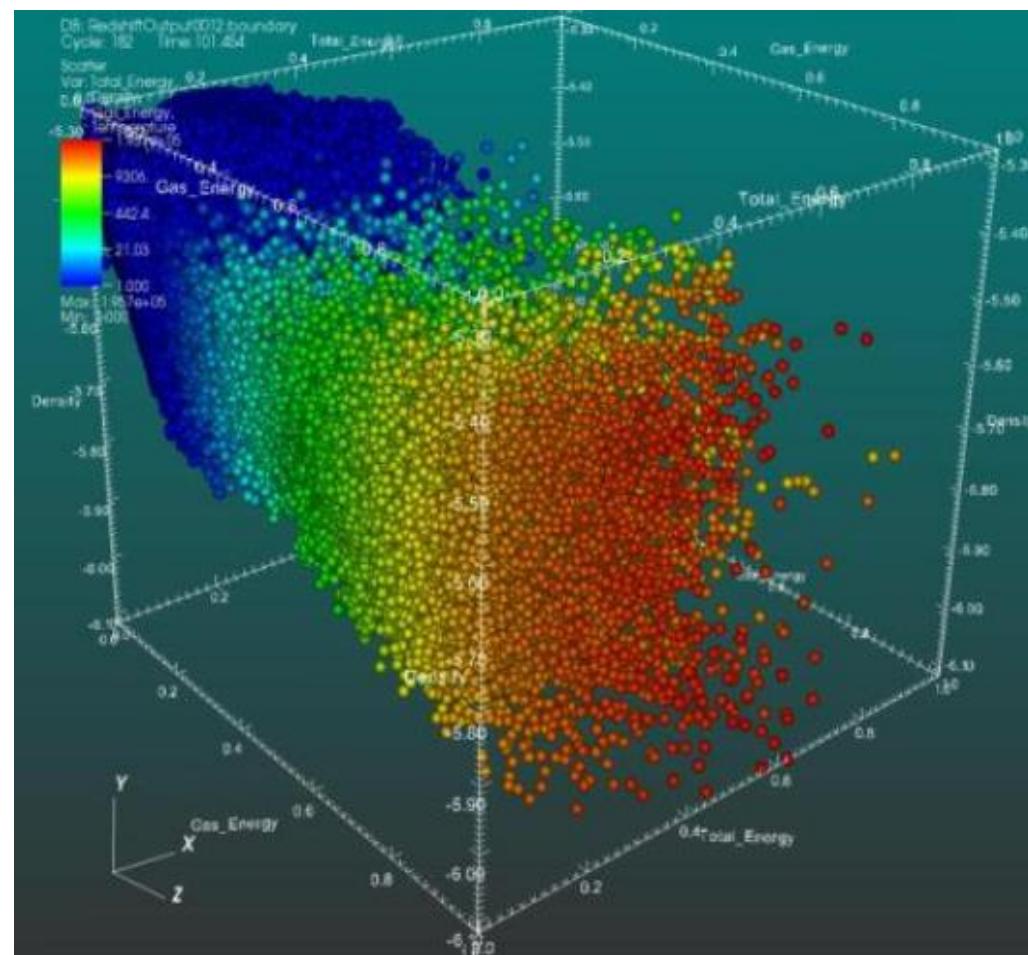
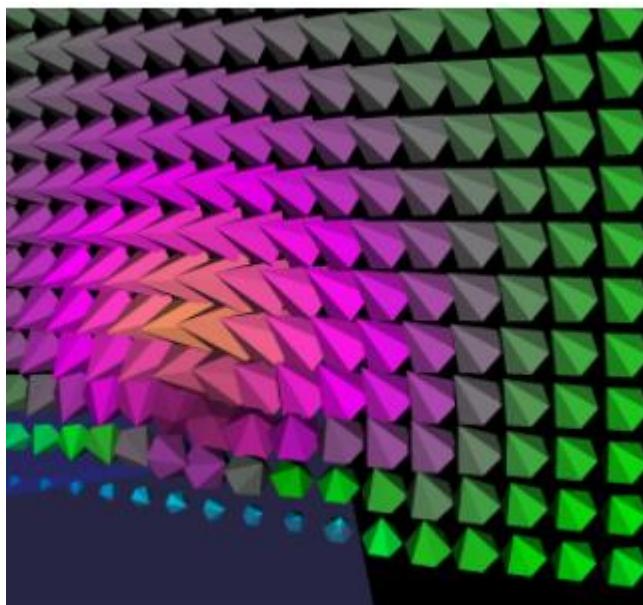
[3d vector field example \(video\) | Khan Academy](#)

[New Animated Stream Lines Representation for ParaView 5.3 \(kitware.com\)](#)

Glyphs

Four-dimensional data visualization: in 3D space a fourth scalar variable is visualized using colored glyphs ➔

Glyphs for Visualizing a 3D Vector Field



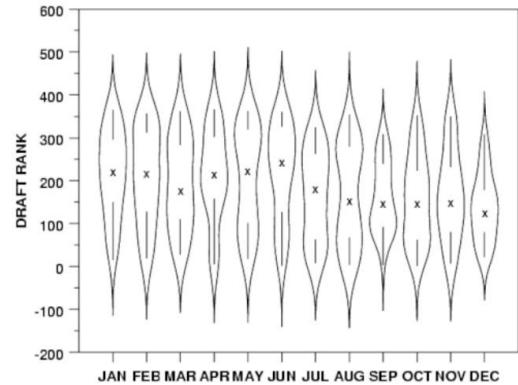
(Wikipedia/Glyph)

Some more visualization techniques:

- These are only some of the visualization techniques to represent value
- There are a lot more ...
- Here are some that you may have seen already

- violin plots

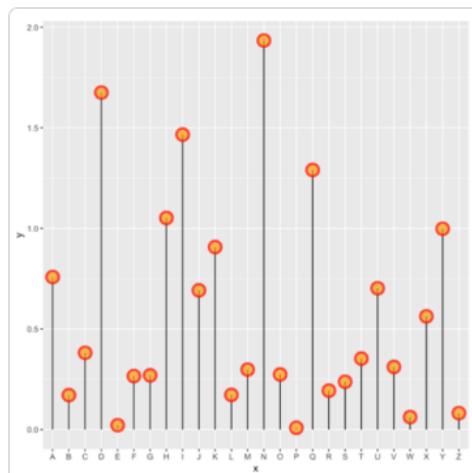
[Violin plot - Wikipedia](#)



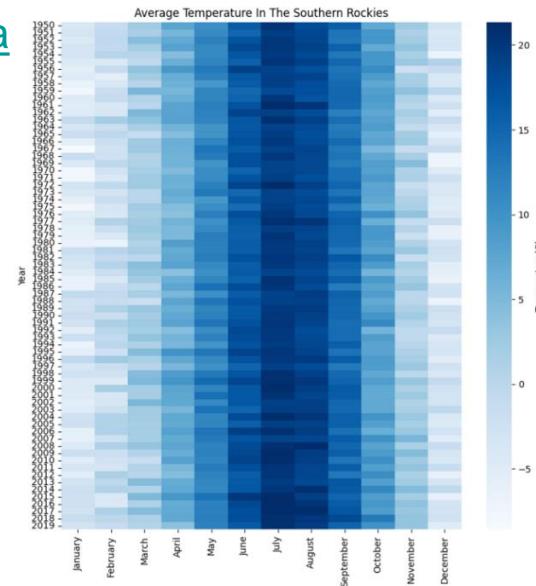
- lollipop charts

[Heat map – Wikipedia](#)

- Heat-maps

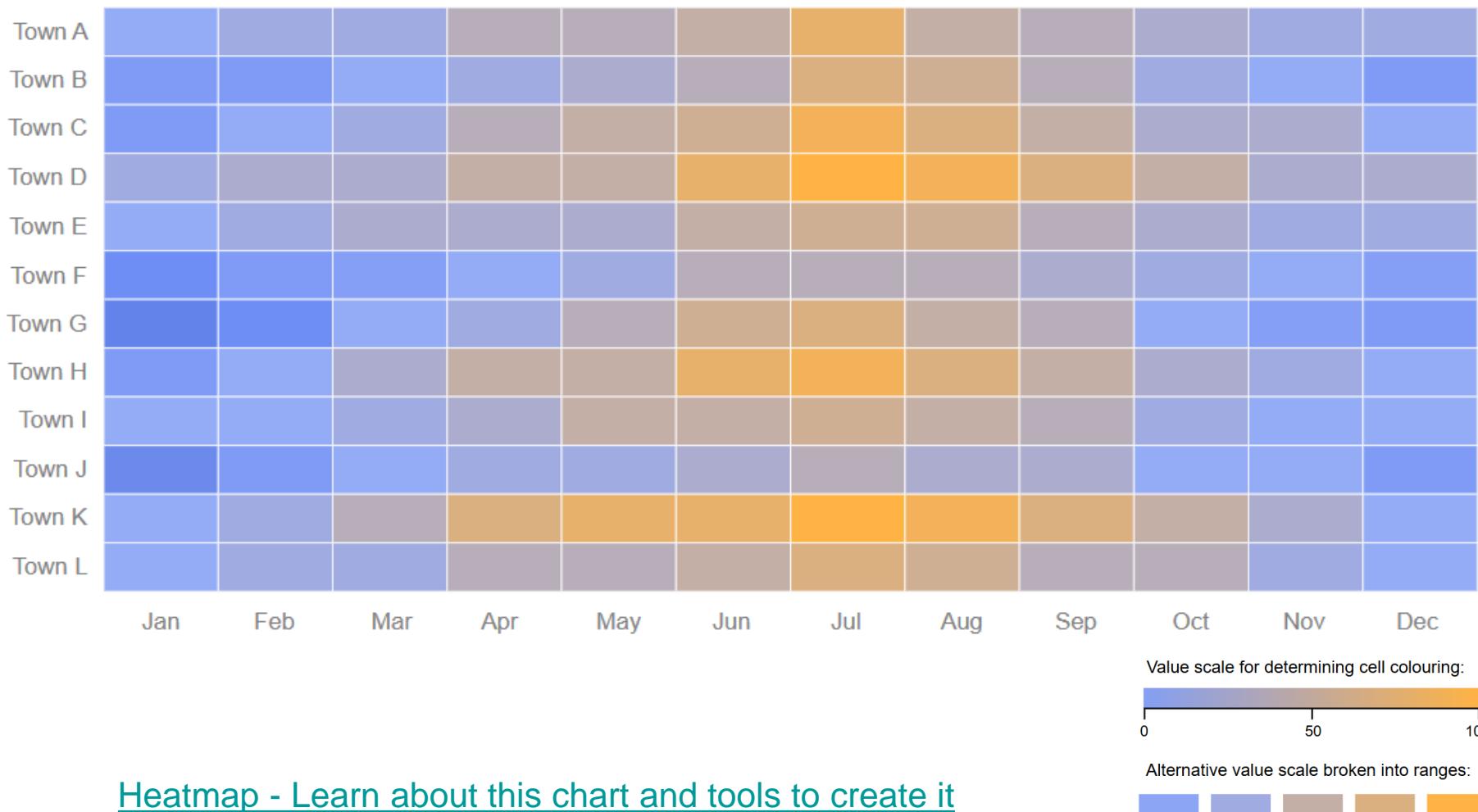


[Lollipop plot | the R Graph Gallery](#)



Heat map

May represent quantitative or ordinal data also in tabular format through colour and help detect patterns



[Heatmap - Learn about this chart and tools to create it](#)

Important aspects to consider to select a Visualization technique:

nature of the problem

- communicate
- explore
- confirm

nature of the data to represent

- quantitative
- ordinal
- categorical

number of attributes

- univariate
- bivariate
- trivariate
- multivariate

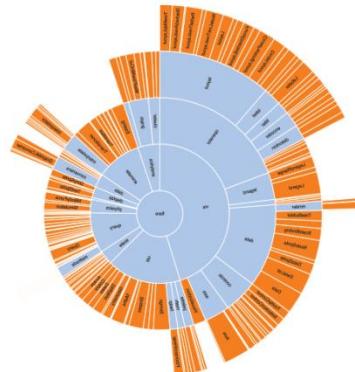
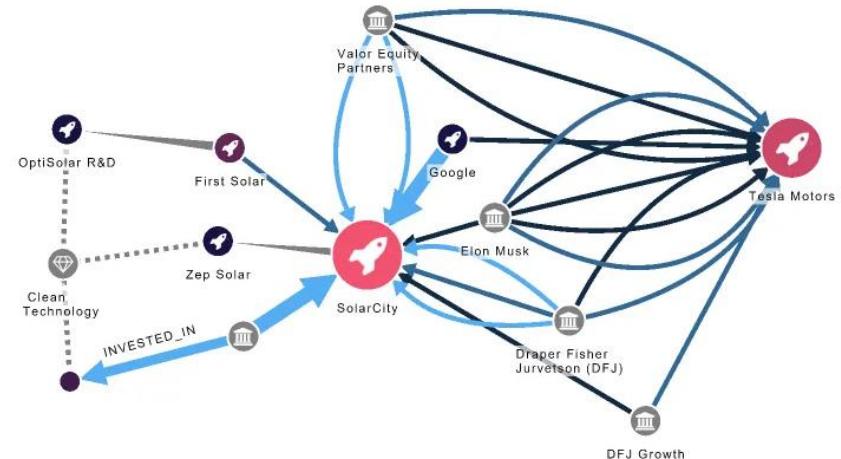
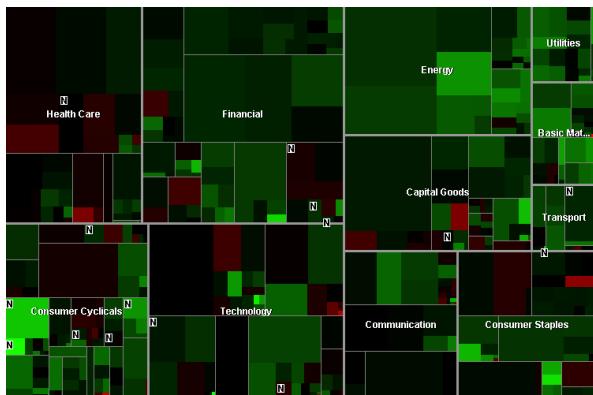
dataset types

- tables
- networks
- spatial or geographical fields
- geometry

Next: networks (relation)

Beyond value: representing relation

- Often we may want to visually represent beyond value: **relation**
 - Networks
 - Hierarchical data (trees)



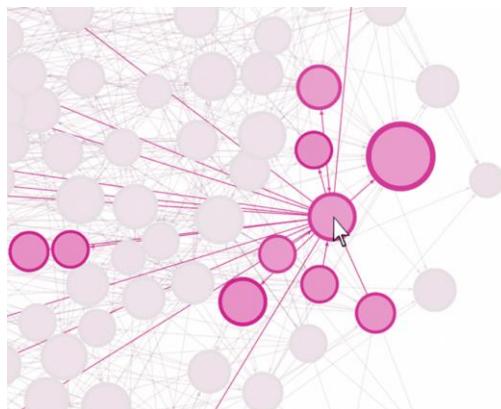
<https://plotly.com/python/sunburst-charts/>

[Understanding and Using Tree Maps | Tableau](#)

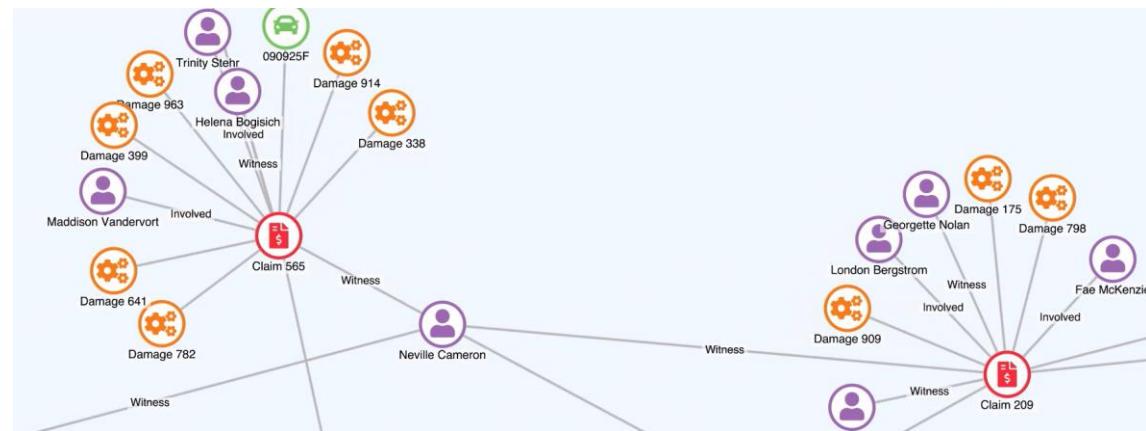
- Often we may want to visually represent networks, e.g.:
 - Transportation networks
 - Social networks
 - Graphs
 - Etc.



[Gephi - The Open
Graph Viz Platform](#)

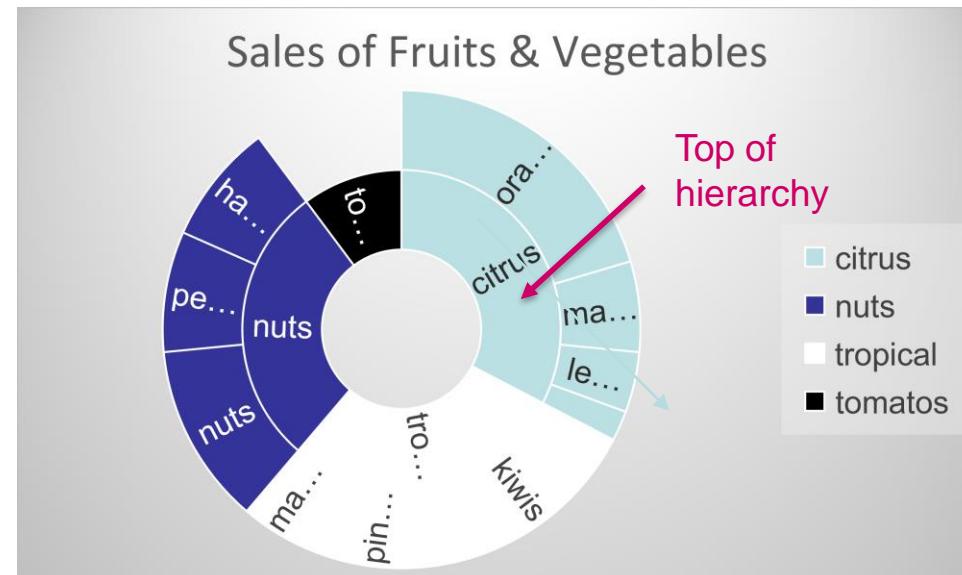


[Mapas e diagramas - Metro de Lisboa
Social Network Visualization Use Cases](#)



- A **sunburst** is adequate to represent hierarchical data.
- Each level of the hierarchy is represented by one ring or circle with the innermost circle as the top of the hierarchy

Types	Fruits/Veg	Quant (t)
citrus	oranges	100
	lemons	20
	mandarines	30
	limes	10
nuts	hazelnuts	40
	nuts	60
	peanuts	40
tropical	kiwis	70
	pineapples	50
	mangos	20
tomatos		50



[Create a sunburst chart in Office - Microsoft Support](#)

Exercise

- Create a Sunburst of these data using Excel

Types	Fruits/Veg	Quant (t)
citrus	oranges	100
	lemons	20
	mandarines	30
	limes	10
nuts	hazelnuts	40
	nuts	60
	peanuts	40
tropical	kiwis	70
	pineapples	50
	mangos	20
tomatos		50

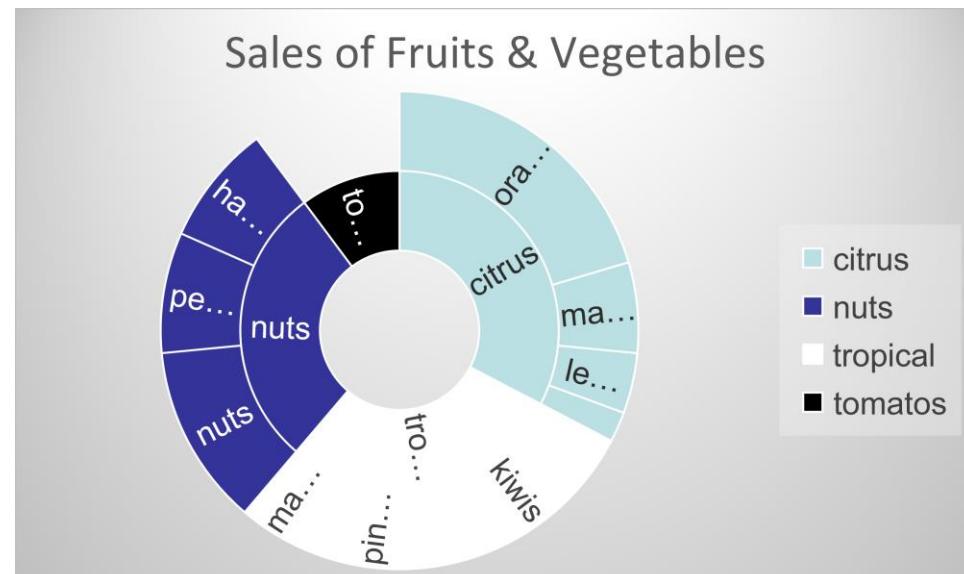
Sunburst

Use this chart type to:

- Compare values across hierarchy levels.
- Show proportions within hierarchical levels as rings.

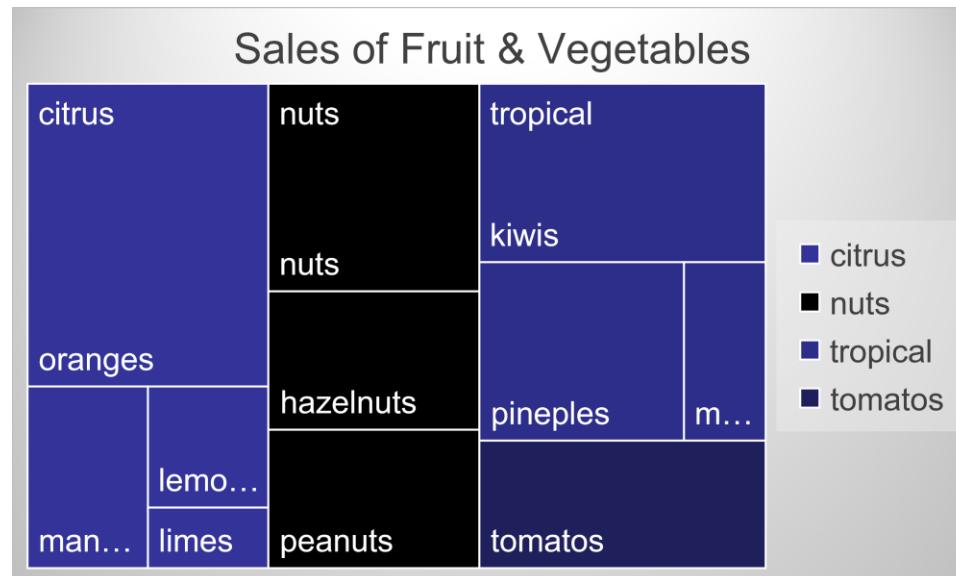
Use it when:

- The data is organized hierarchically and has many categories.



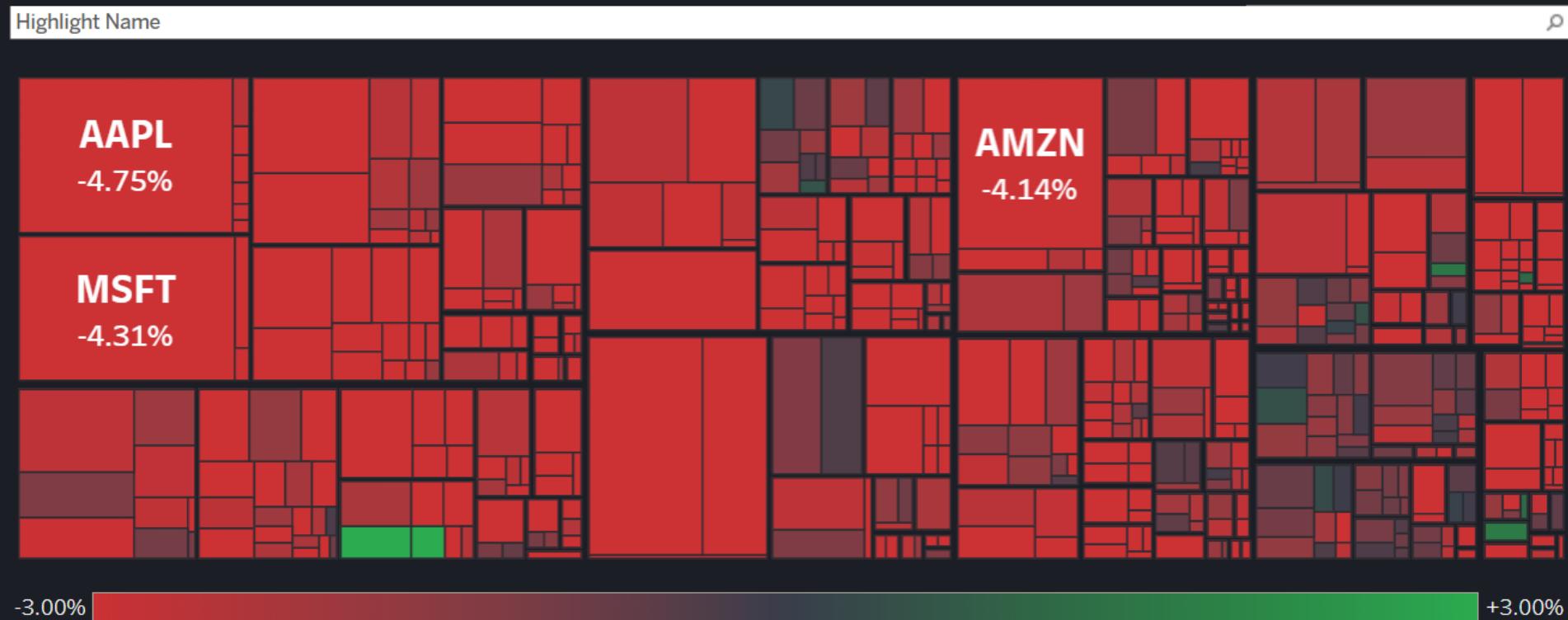
- **Treemaps** represent hierarchical data as a set of nested rectangles
- Each node is given a rectangle, tiled with smaller rectangles

Types	Fruits/Veg	Quant (t)
citrus	oranges	100
	lemons	20
	mandarines	30
	limes	10
nuts	hazelnuts	40
	nuts	60
	peanuts	40
tropical	kiwis	70
	pineples	50
	mangos	20
tomatos		50



- It is possible to use colour (in this case using cultural convention: red negative, green positive)

Data: Yahoo! Finance
 S&P 500 index stocks categorized by sectors and industries. Size represents market cap.

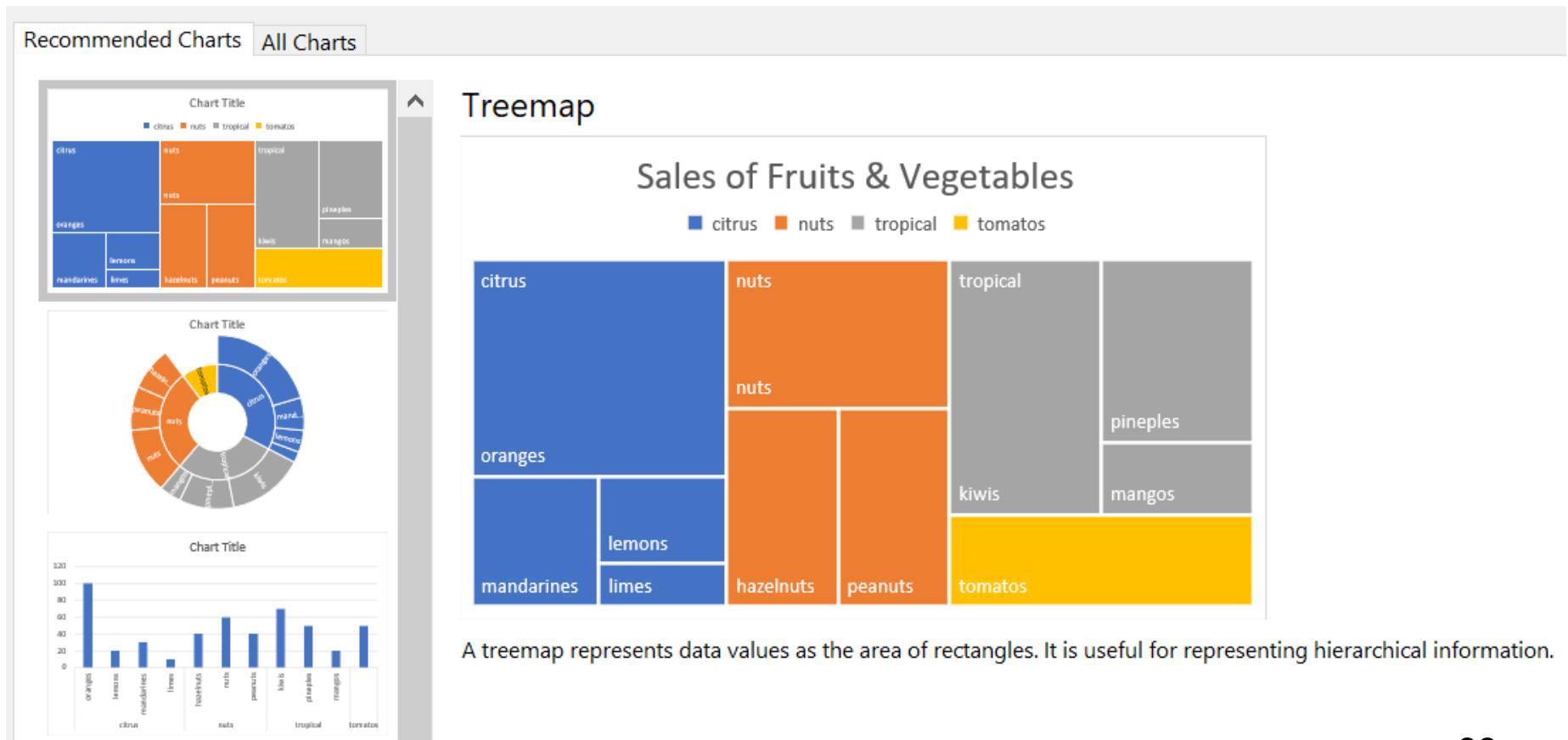


[Stock Market Overview - Nested Treemap | Tableau Public](#)

Exercise

- Create a Treemap of the same data using Excel

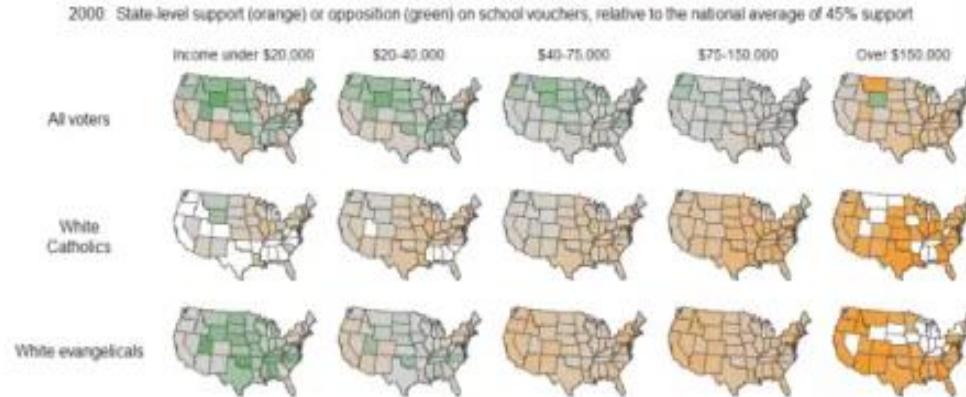
Insert Chart



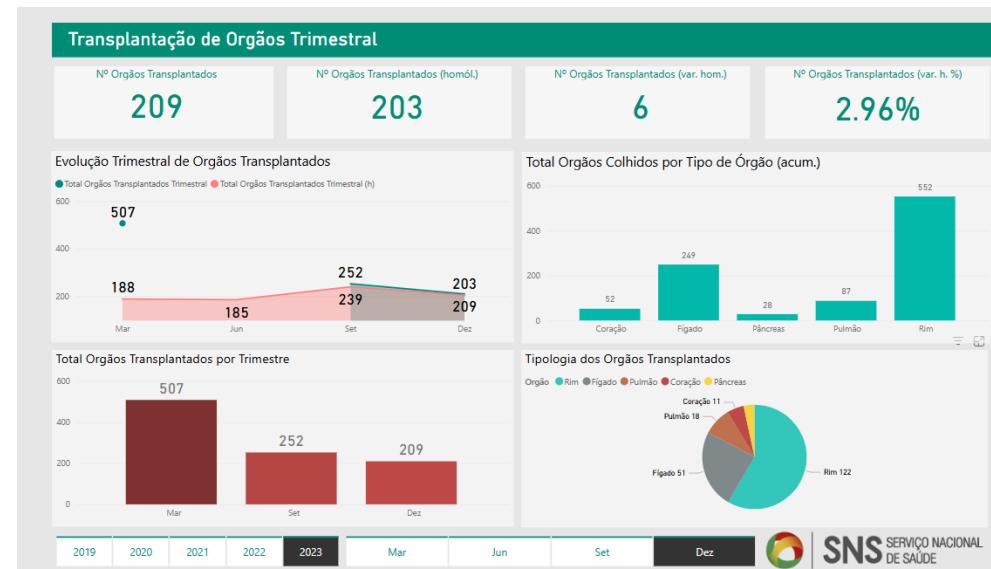
Arranging multiple visualizations:

Often it is interesting to several Visualizations

Small multiples



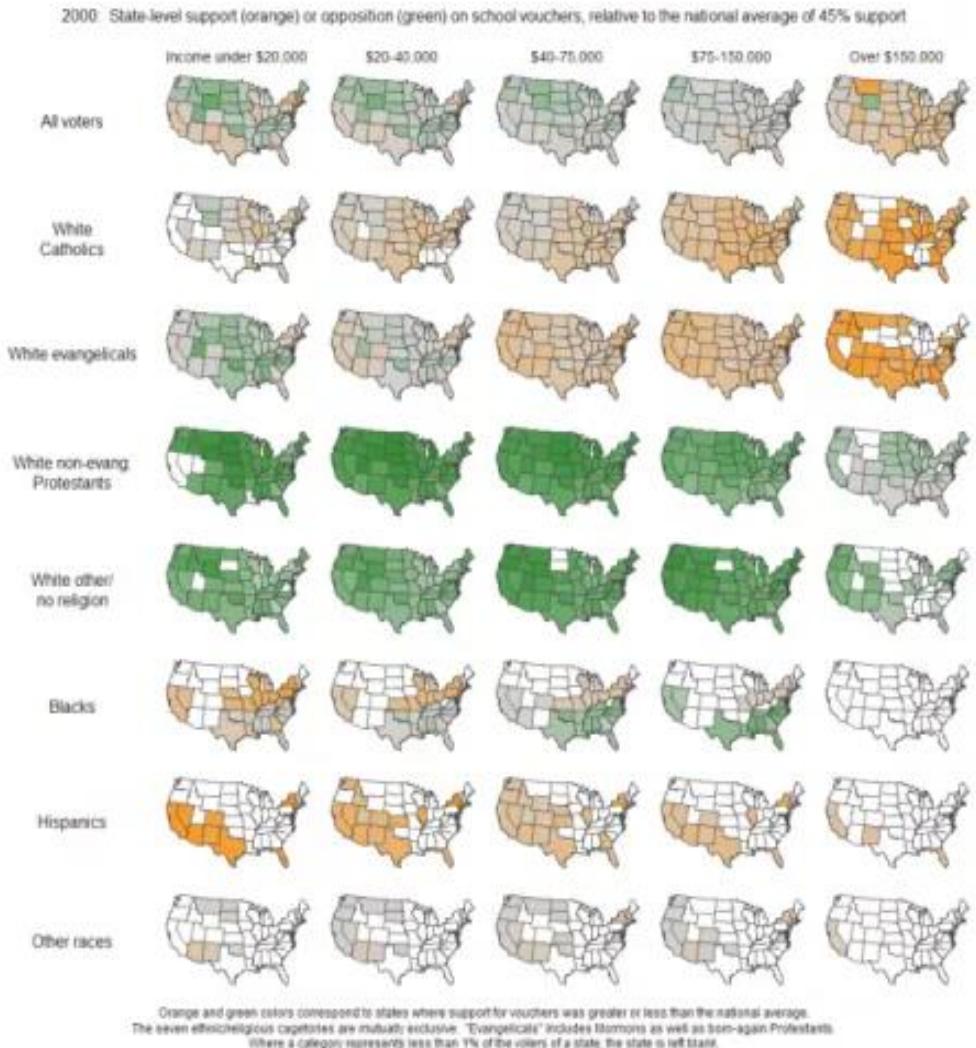
Dashboards



Small multiple

arrangement of similar visualizations
facilitating efficient and effective comparisons

(Kirk, 2012)



Better Know a Visualization: Small Multiples

Dashboard

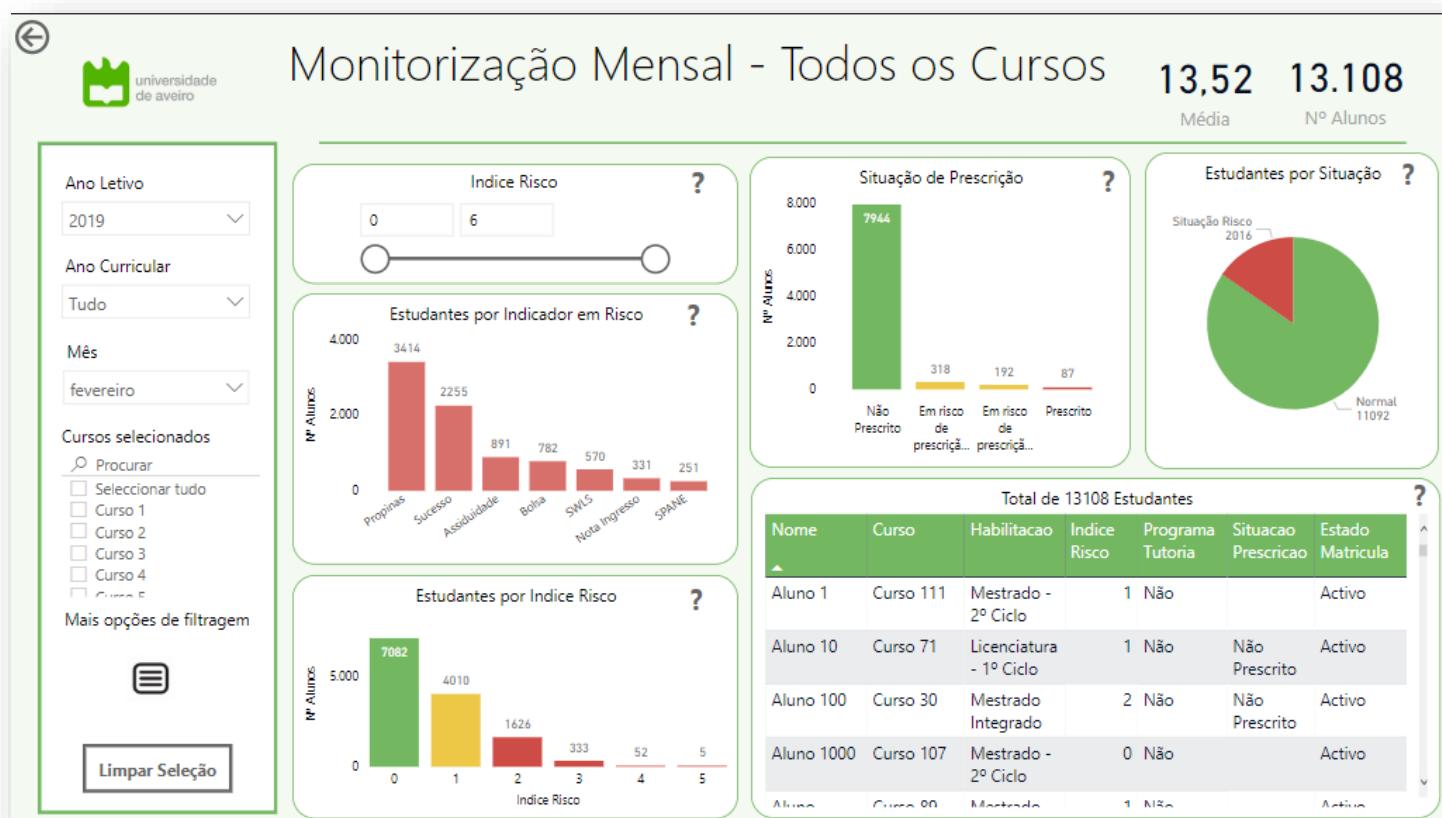
Visual display summarizing a dataset providing information at-a-glance (e.g. KPIs)

" A dashboard is a visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance. " (Few, 2004)

Dashboards: Making Charts and Graphs Easier to Understand

Prototype:
“Portal dos
indicadores,
UA”

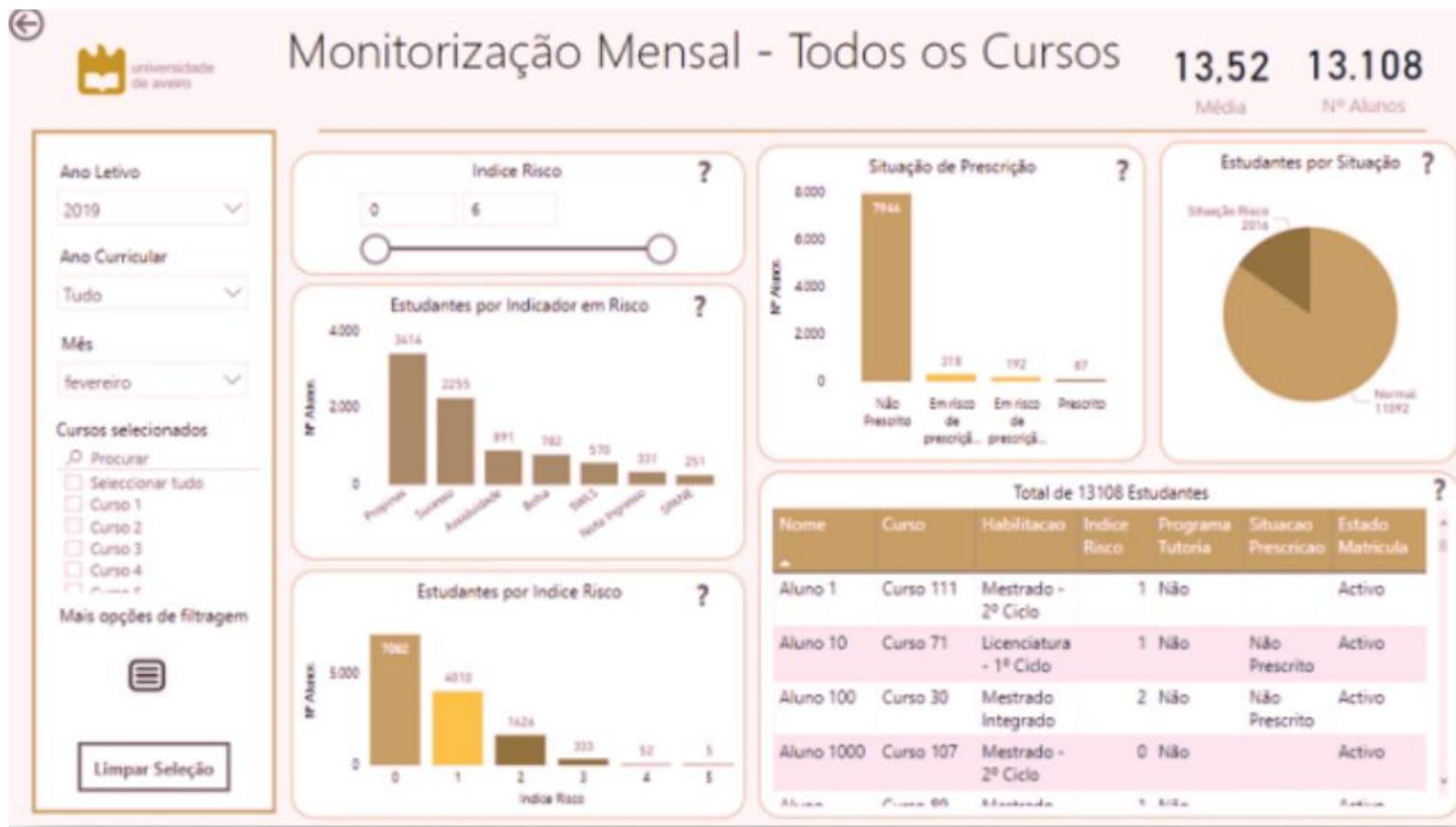
Any issues
about the
color?



Testing readability to color-blind users:

As seen by someone with deutanopia (red-green blindness)

[Coblis — Color Blindness Simulator – Colblindor](#)



Bibliography

- Camões, J., *Data at Work : Best practices for creating effective charts and information graphics in Microsoft Excel*, Pearson Education, 2016
<https://learning.oreilly.com/library/view/data-at-work/9780134268798>
- Heer, J., Bostock, M., & Ogievetsky, V. A tour through the visualization zoo.
Communications of the ACM, vol. 8, n.1, 2010 Kirk, A., *Data Visualization: A successful design process*, Pack Publishing, 2012
- Kirk, A., *Data Visualisation: a successful design process*, Packt publishing, 2012
- Spence, R., *Information Visualization, Design for Interaction*, 2nd ed., Prentice Hall, 2007
- Ware, C., *Information Visualization, Perception for Design*, 3rd ed., Morgan Kaufman, 2012
- Wilke, C., *Fundamentals of Data Visualization: A Primer on Making In informative and Compelling Figures*, O'Reilly, 2019