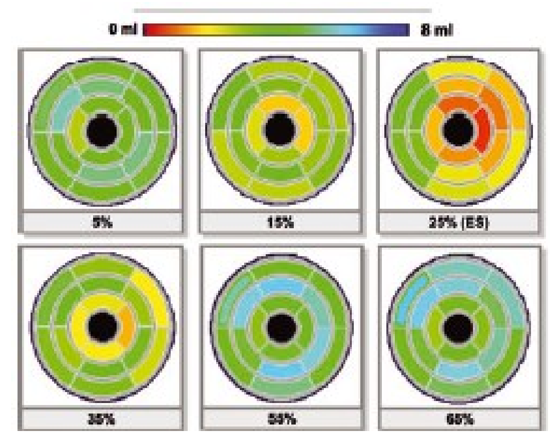
Universidade de Aveiro
Departamento de Electrónica,
Telecomunicações e Informática

48º aniversário
universidade de aveiro
1973 . 2021

An Introduction to Data Visualization



Beatriz Sousa Santos, University of Aveiro, 2022

Definition

Objectives

History

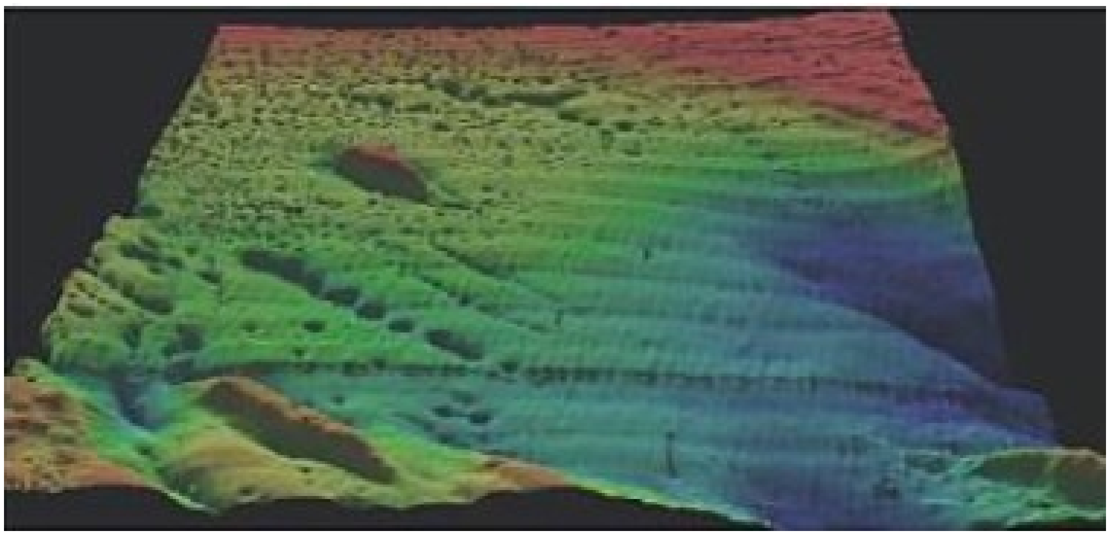
Applications

Model

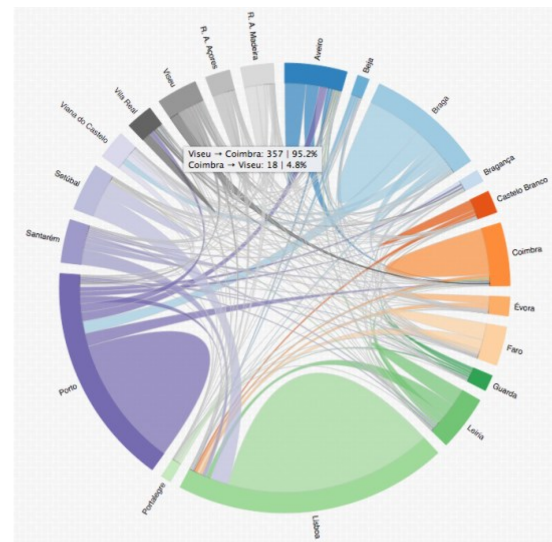
How to obtain and evaluate a Visualization?

What is Visualization?

Is the process of exploring, transforming and **representing data** as images to **gain insight into phenomena**



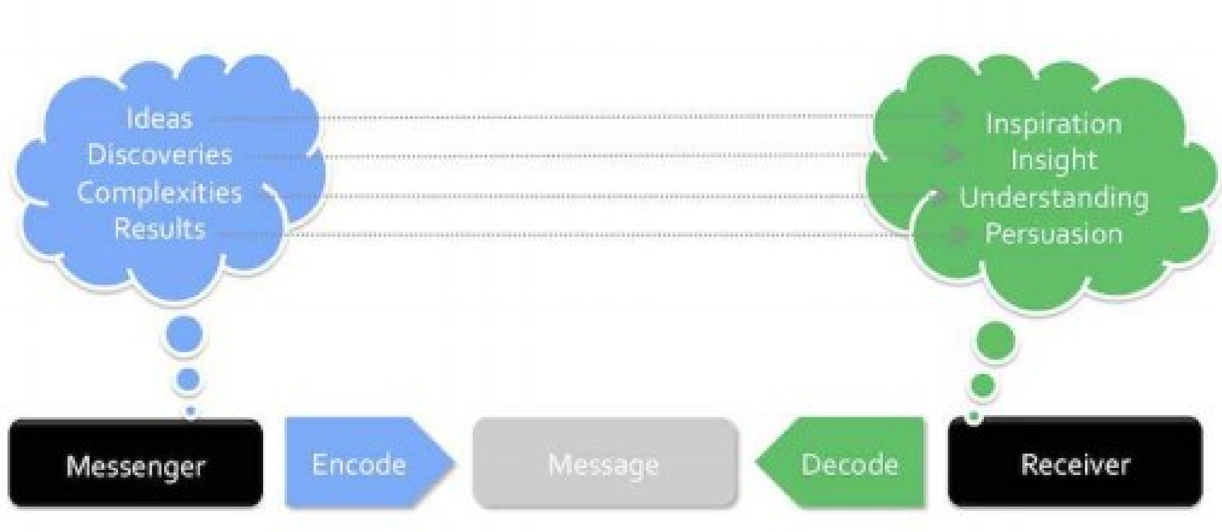
Passamoquoddy Bay
(10^6 measures)
(Ware 2019)



Portuguese Higher Education
(data from 120 000 candidates)

Definition (yet another)

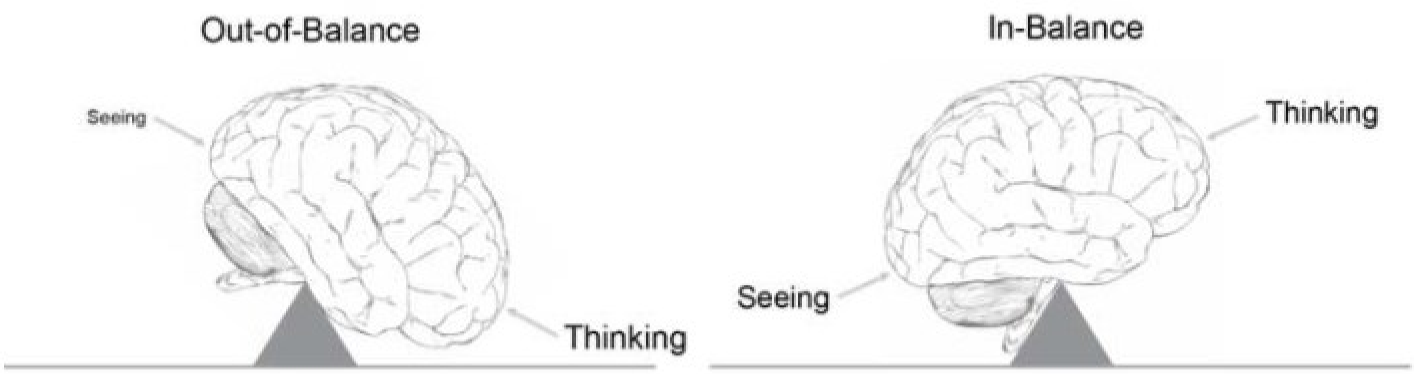
- “The representation and presentation of data that exploits our visual perception abilities in order to amplify cognition” *(Kirk, 2012)*



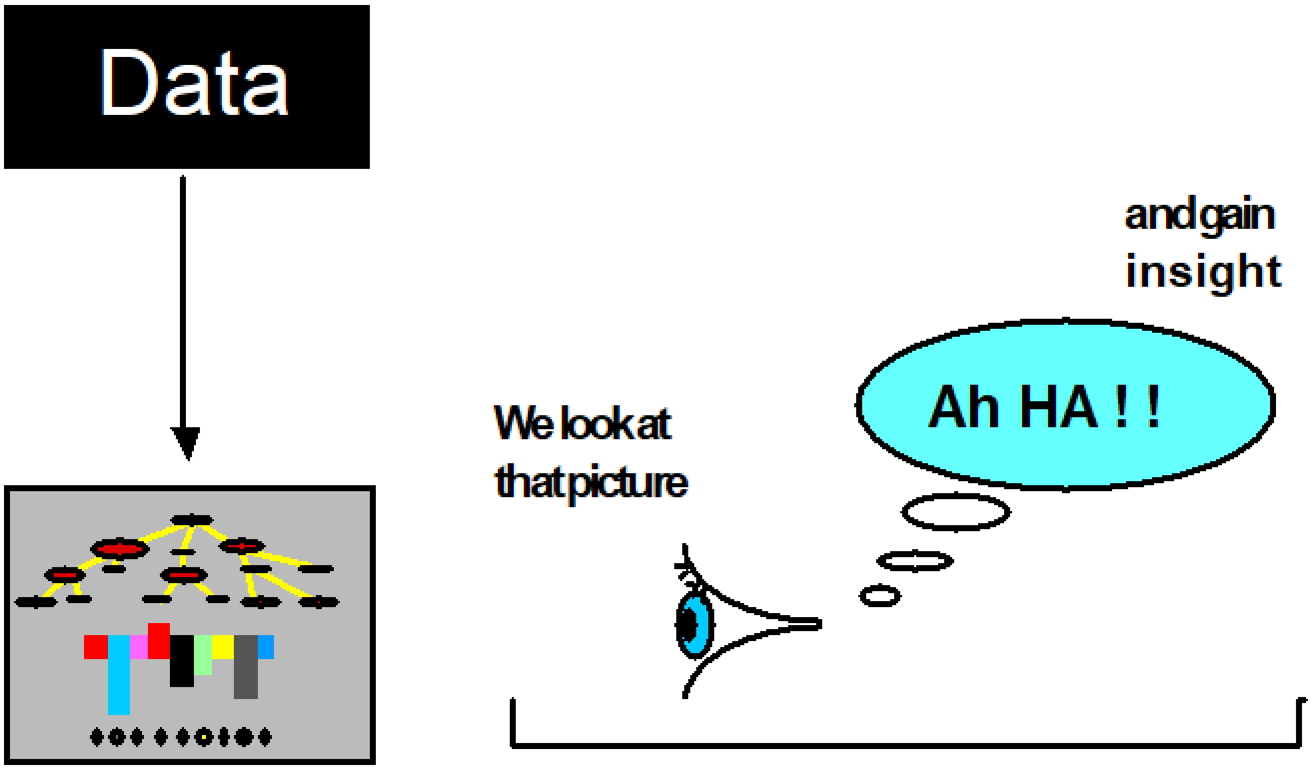
(Kirk, 2012)

Visualization benefits

- Helps us think
- Reduces load on working memory
- Offloads cognition
- Uses the power of human perception



https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/data-visualization-for-human-perception



The process of visualization: graphically encoded data is viewed in order to form a mental model of that data (Spence, 2007)

Human-in-the loop process!

Data and Information Visualization

- In general:

  Data (scientific) Visualization (DV) - Data having an inherent spatial structure

  (e.g., CAT, MR, geophysical, meteorological, fluid dynamics data)

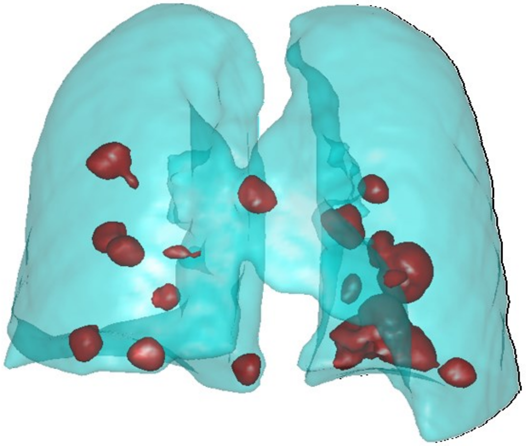
  Information Visualization (IV) – “Abstract” tabular data not having an inherent spatial structure (tabular data)

  (e.g., stock exchange , S/W, Web usage patterns, text)

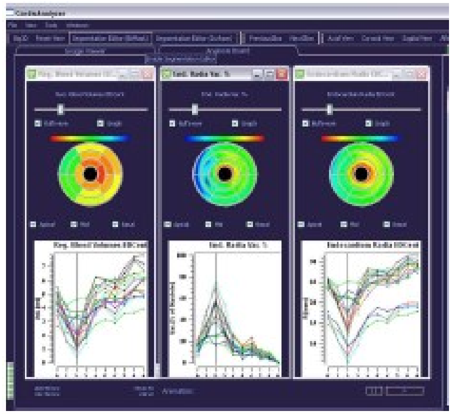
- These designations may be misleading; both DV and IV start with (raw) data and allow to extract information

- Borders between these areas are not well defined …

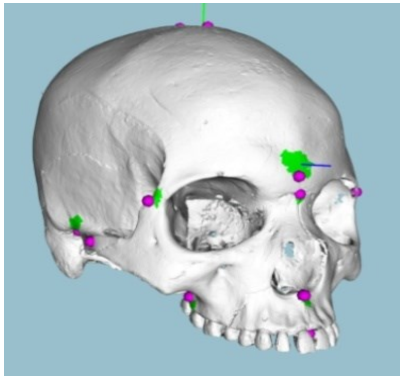
Scientific Visualization (examples "made in UA")



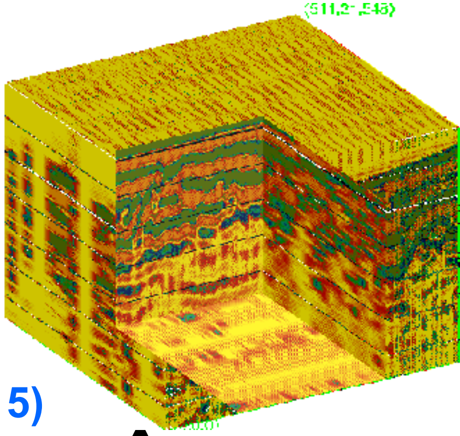
Tomography
(2004)



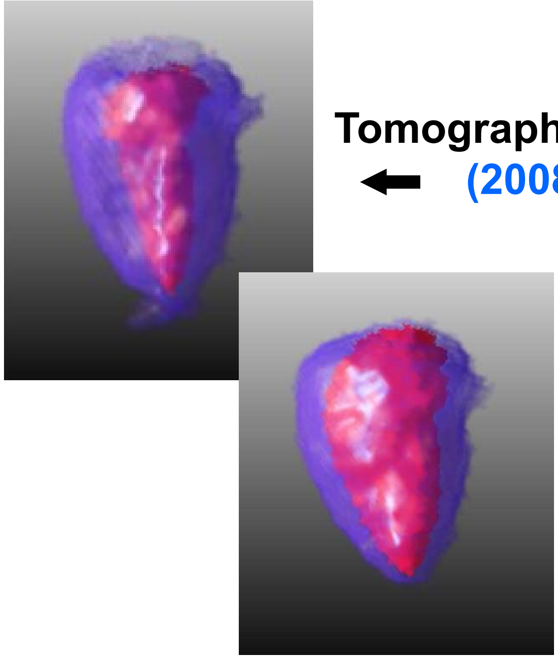
Tomography (2011)



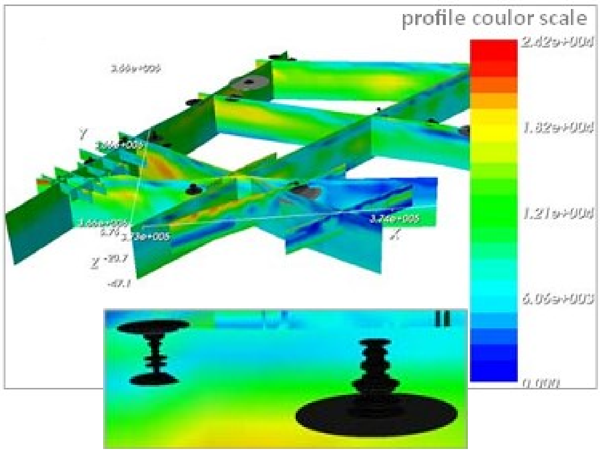
Laser scanner (2015)



Ground
Penetrating Radar (1999)

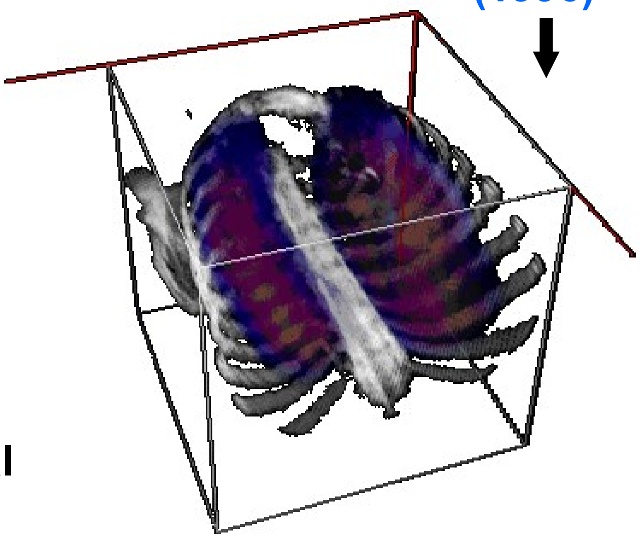


Tomography
(2008)



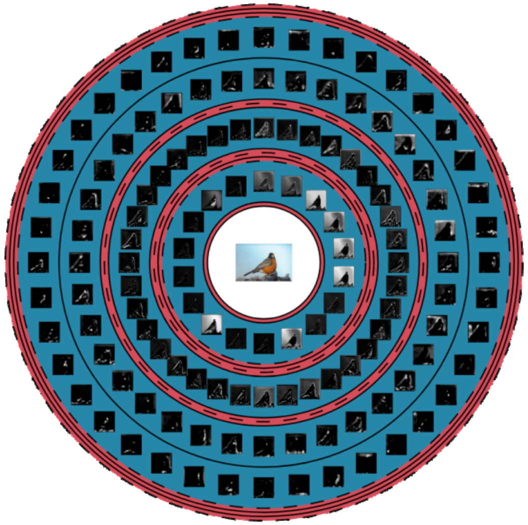
Electrical and mechanical
ground resistivity (2010)

Tomography and SPECT
(1996)

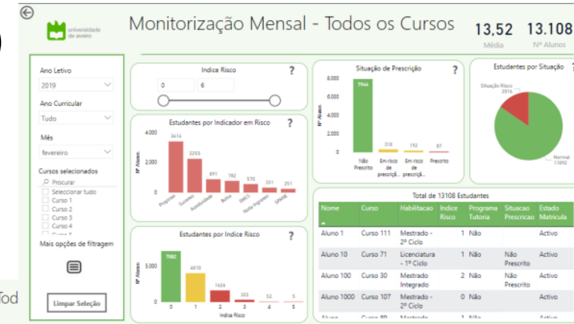


Information Visualization **(examples "made in UA")**

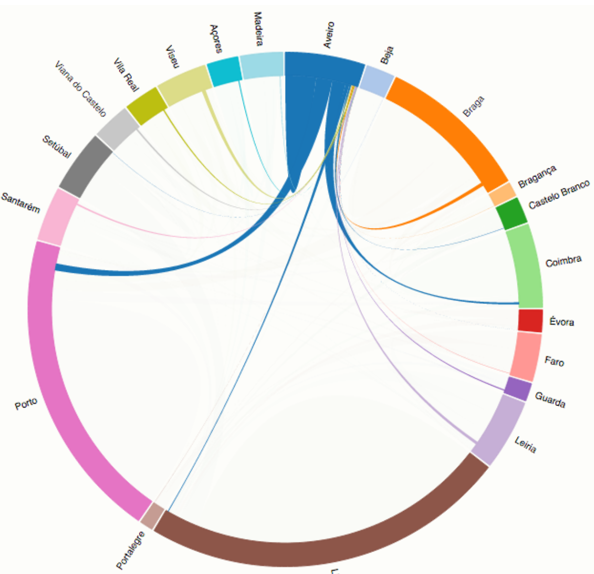
Machine Learning Visualization (XAI)
(UA, 2020)

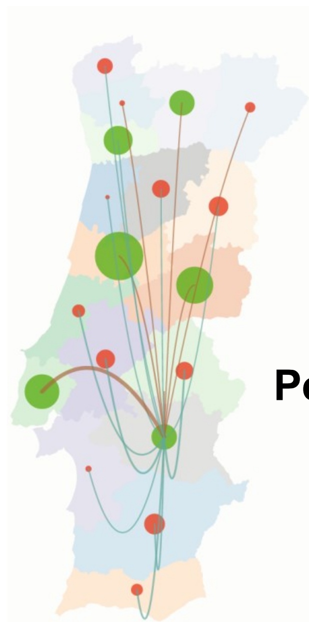


Academic data
(UA, 2020)



Student Migrations
(UA, 2015)

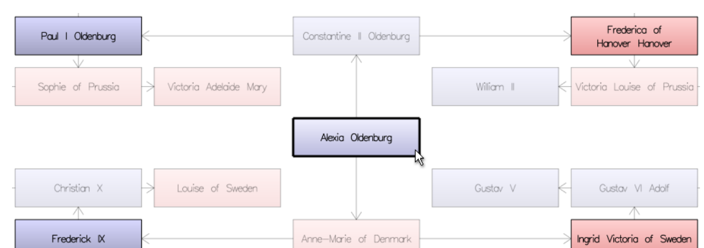




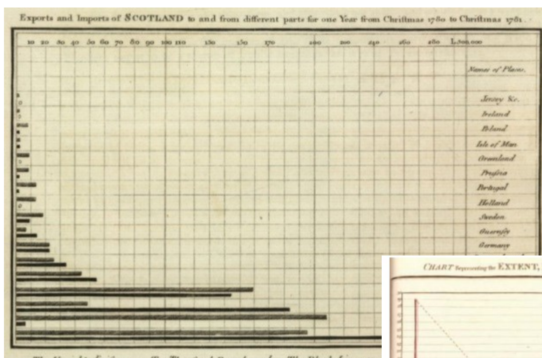
www.portugal-migration.info

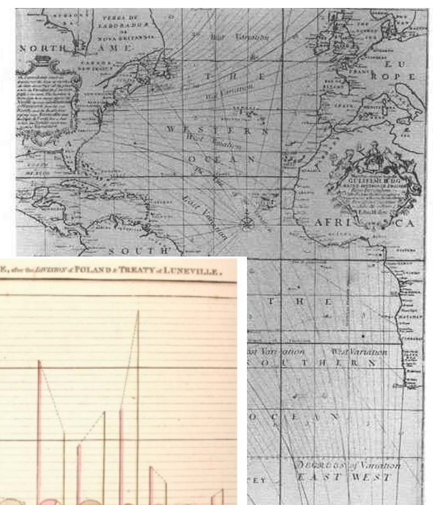
Taxonomy Visualization
(UA, 2021)

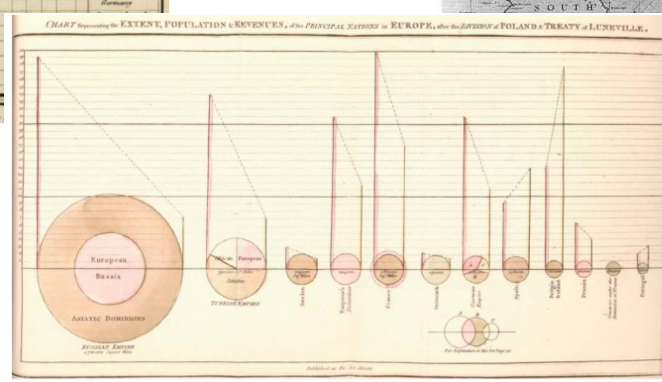
Pedigree trees
(UA, 2011)



Brief History







Brief history

- The usefulness of graphical representations of large amounts of data has been recognized long ago:

  XVIII e XIX centuries- use of graphics in statistics and science:

  W. Playfair, C. J. Minard

  XX century- J. Bertin, E. Tufte

- The use of the computer made Visualization a more practicable discipline:

  1987 - Identification of Visualization as an autonomous discipline

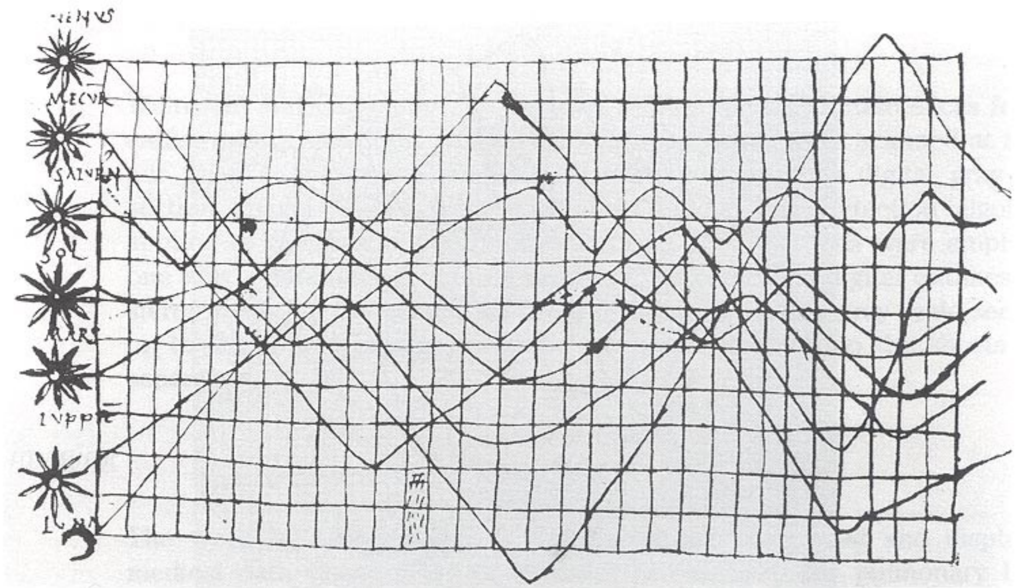
  Visualization in Scientific Computing
  (McCormick, de Fanti and Brown – 1987)

Brief history

- Plenty of Visualization examples of the “pre-computer age”:
  - Inclination of planetary orbits – Xth century
  - Import/ export (Playfair) – XVIIIth century
  - Magnetic declination (Halley) – XVIIIth century
  - Russia campaign of Napoleon (Minard) –XIXth century
  - Cholera out-brake in London (Dr. Snow) – XIXth century

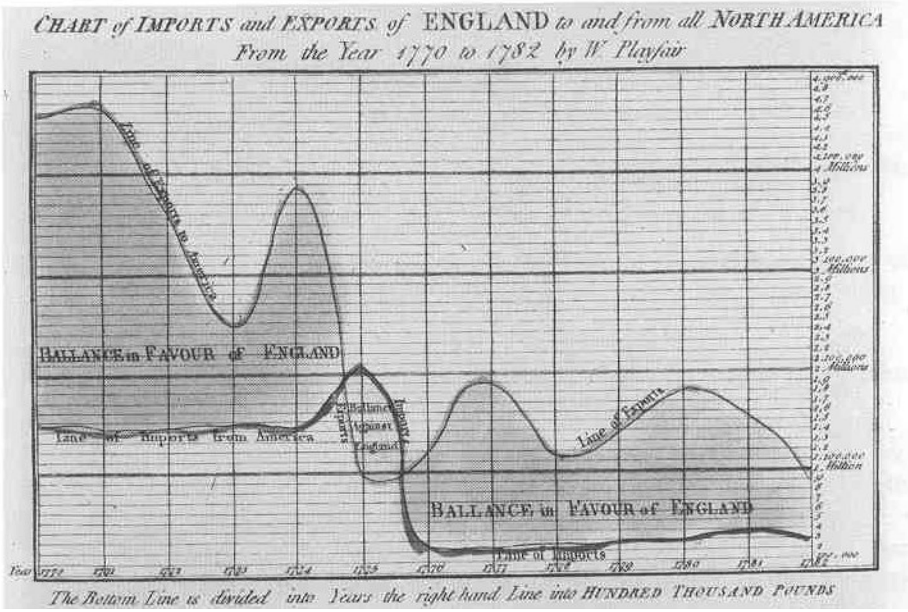
“Pre-computer” Visualization:

One of the oldest known Visualizations



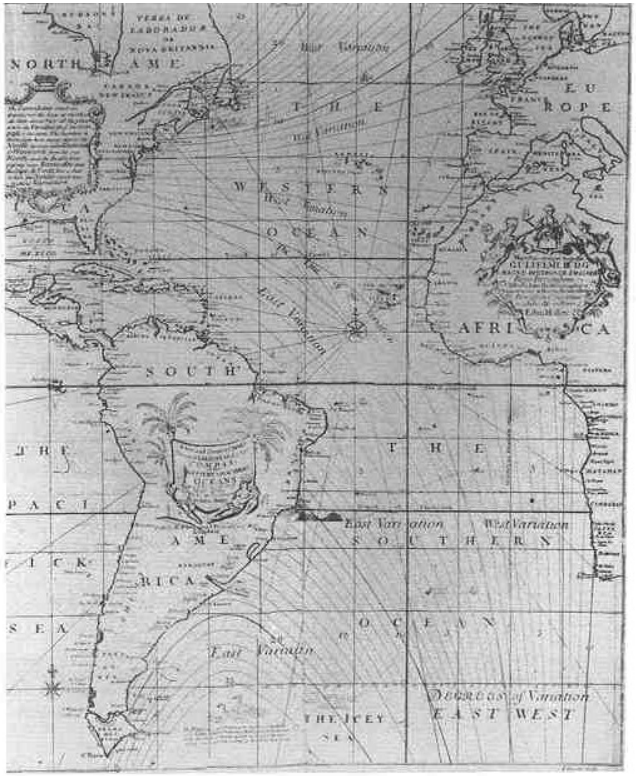
Inclination of orbits along the time - Xth century (Tufte,1983)

One of the first Visualizations used in “business”



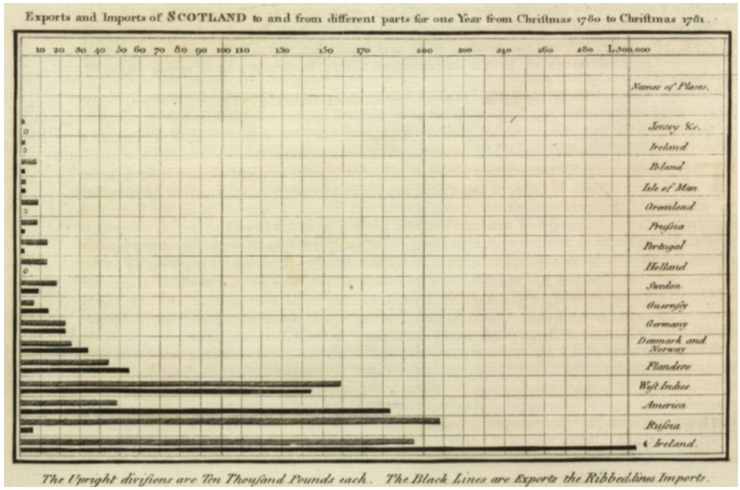
Import/export during the period from1770 to1782 by William Playfair (Tufte, 1983)

One of the first visualizations using contours (isolines)



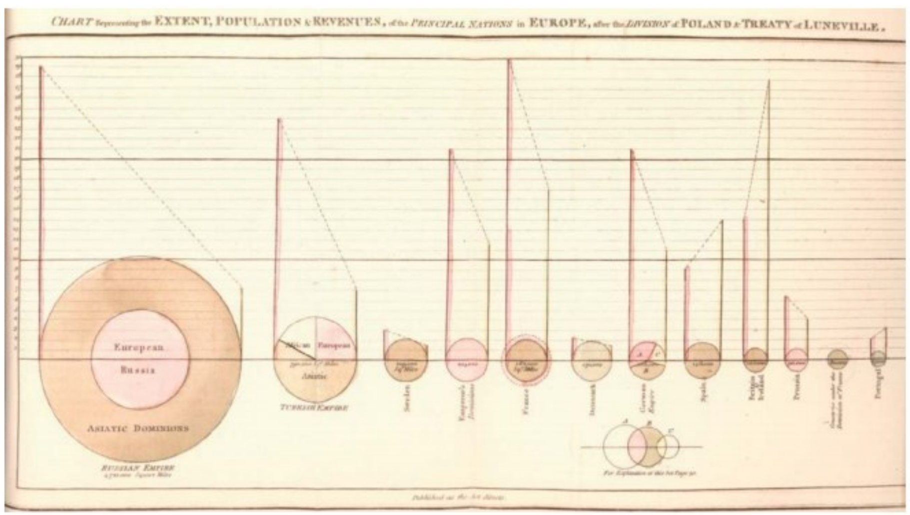
Magnetic declination 1701 Edmund Halley (Tufte,1983)

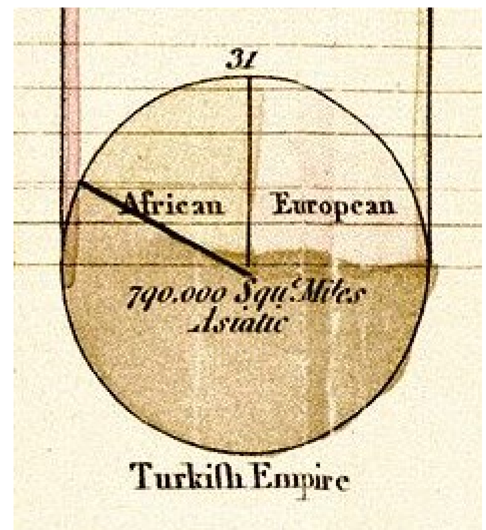
“Ancestors” of simple representations of univariate data



Exports and Imports of Scotland to and from different parts for one
W. Playfair’s *The Commercial and Political Atlas, 1871*

https://en.wikipedia.org/wiki/William_Playfair

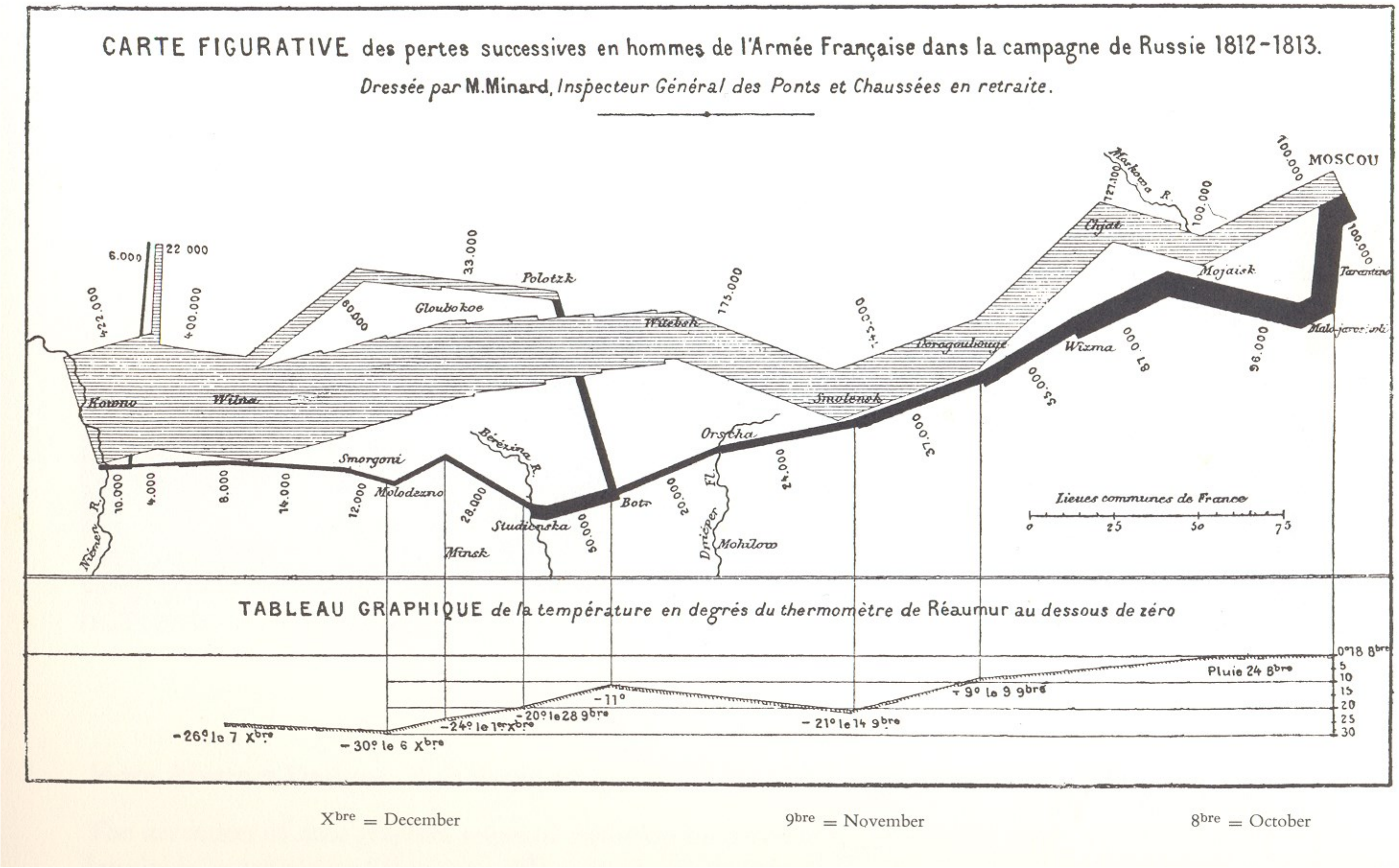




W. Playfair, *Statistical Breviary*, 1801

Multidimensional Visualization

6 dimensions: place (2), n. of men and direction of the army, date, temperature



Russia campaign of Napoleon 1861 by Charles Minard (Tufte,1983)

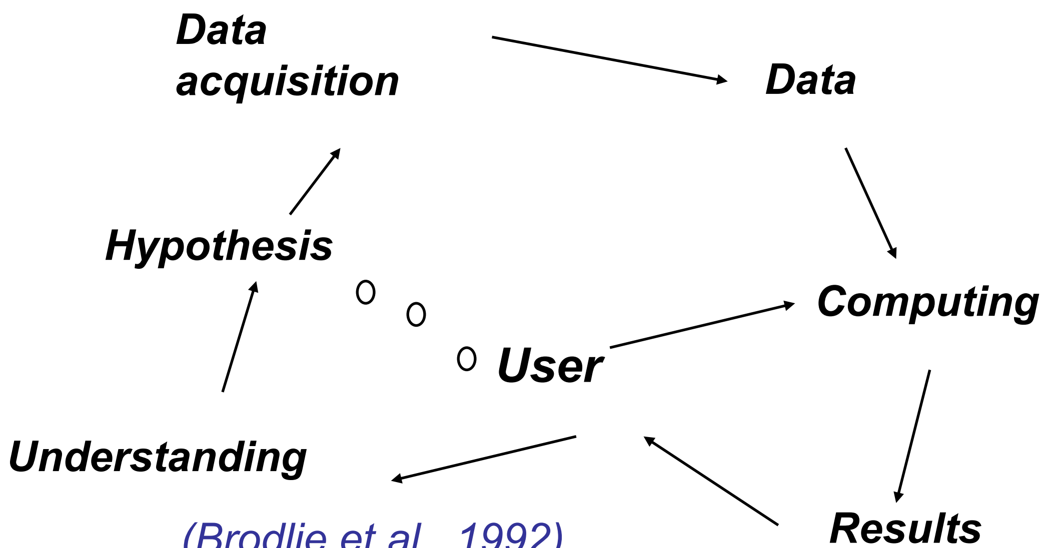
Visualization in scientific discovery

Dr. John Snow

Discovering the cause of the London cholera out brake, 1853-54
(Wikipedia)

Framework

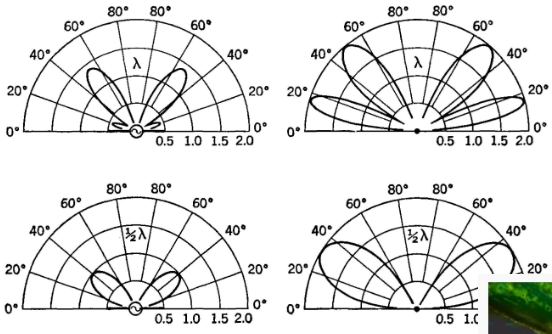
- Visualization includes not only images/charts production from the data, but also their **transformation** and **manipulation** (if possible their acquisition)

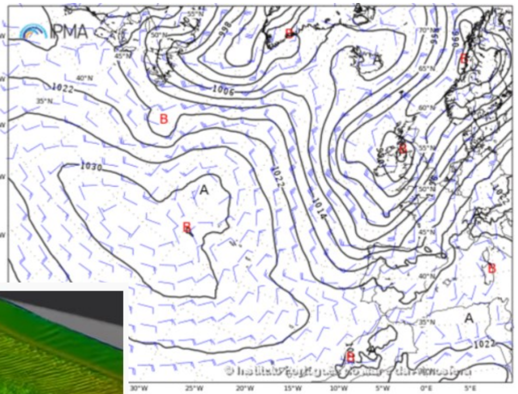


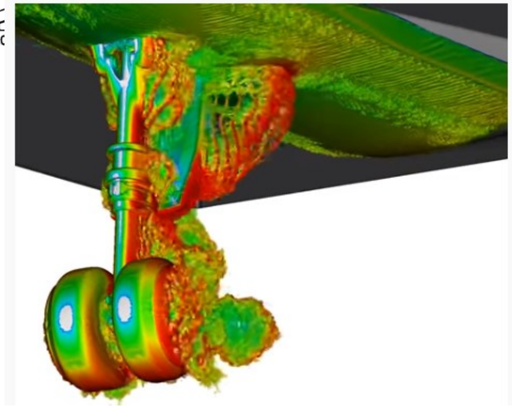
(Brodlie et al., 1992)

- It is a “**human-in-the-loop**” problem

Applications







Applications of Scientific Visualization

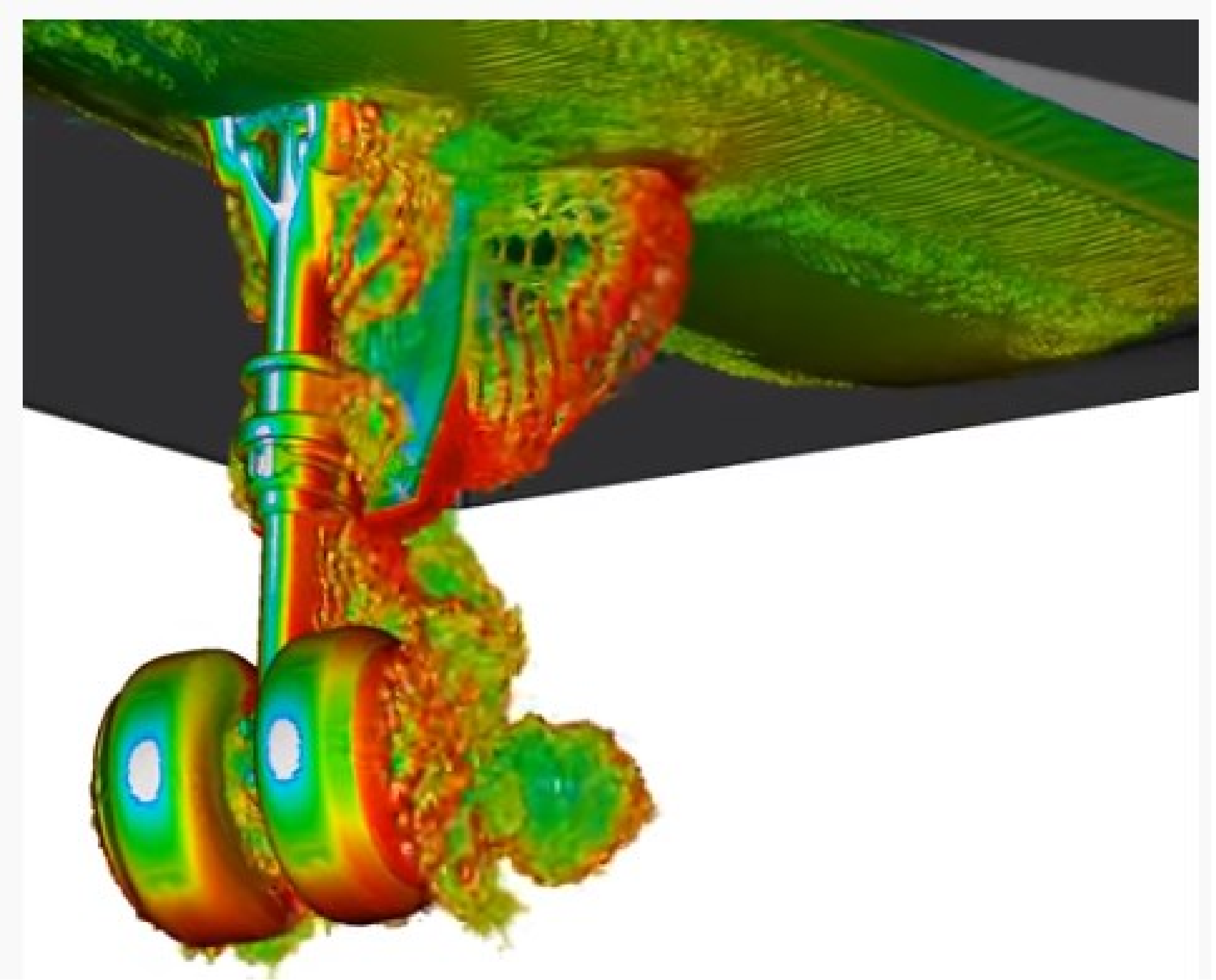
- **Scientific Visualization** is currently used in many scientific areas:
  - All engineering fields …
  - Medicine
  - Meteorology, climatology, oceanography
  - Fluid dynamics
  - Cosmology
  - etc., etc.

- Let us see some examples …

- Can you think of an area where data visualization cannot be applied?

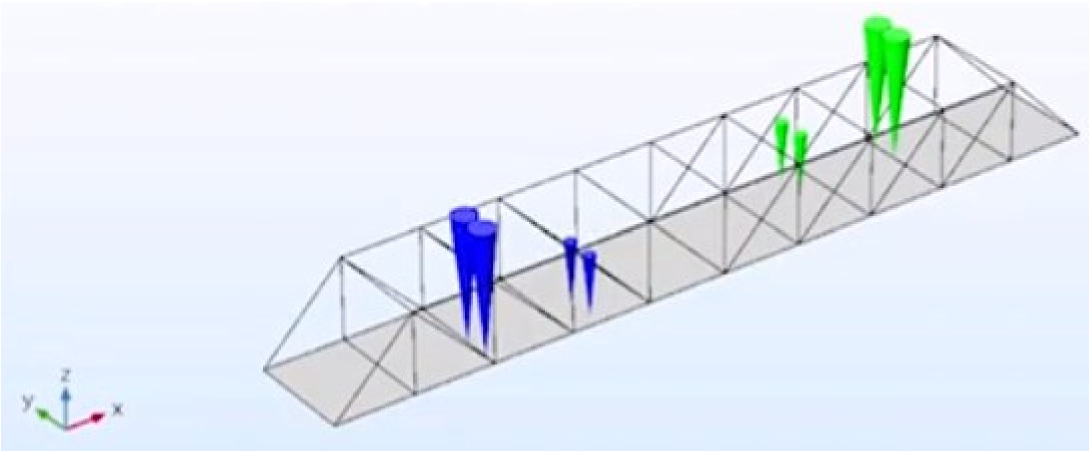
Fluid mechanics visualization

NASA/Boeing CFD visualization of vortices responsible for the noise created by the 777's noise landing

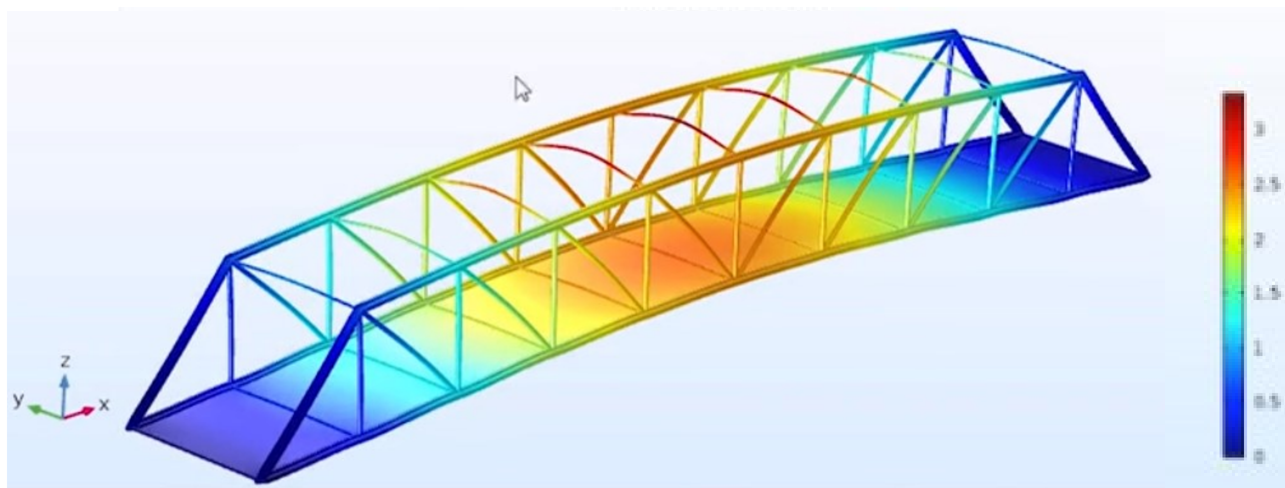


https://www.youtube.com/watch?v=F9EFx7aQuhw

Civil engineering visualization

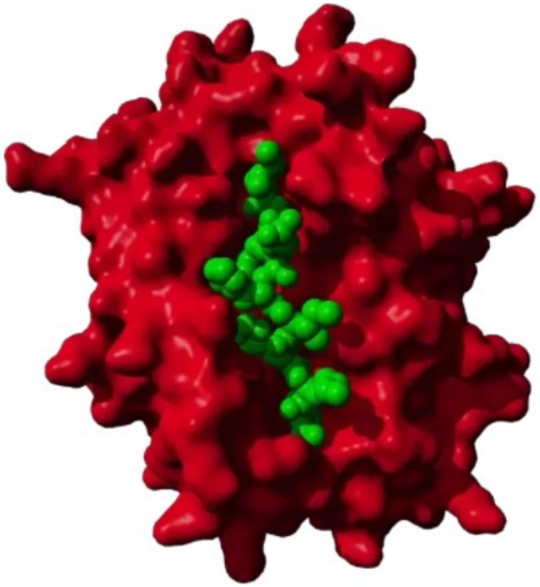


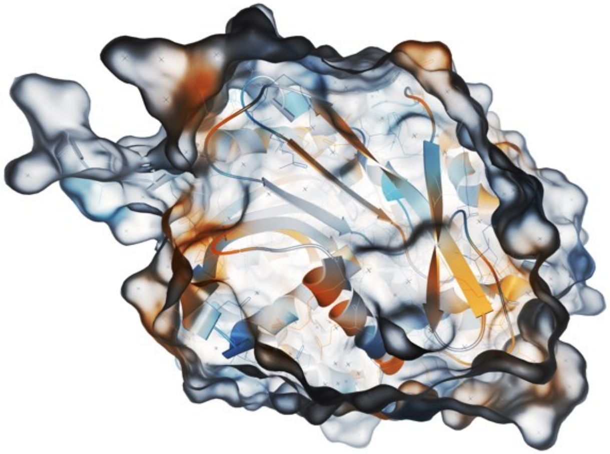
Bridge project:
Visualizing Displacement, Force and Moment in Beams, Stress in Beams, and Stress in Roadway

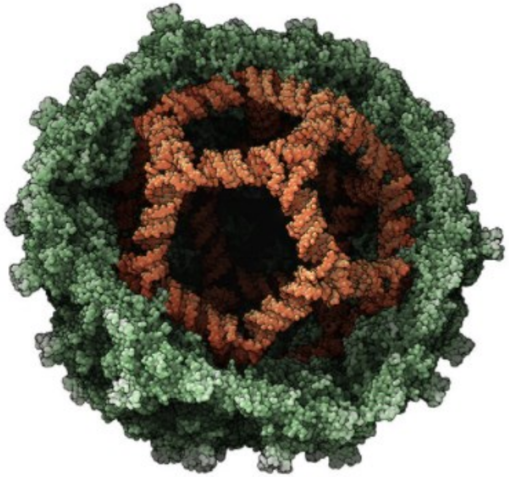


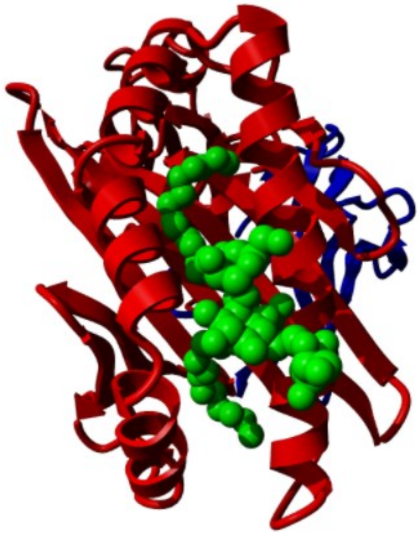
https://www.comsol.com/blogs/efficiently-analyze-civil-engineering-designs-using-an-app

Molecule visualization





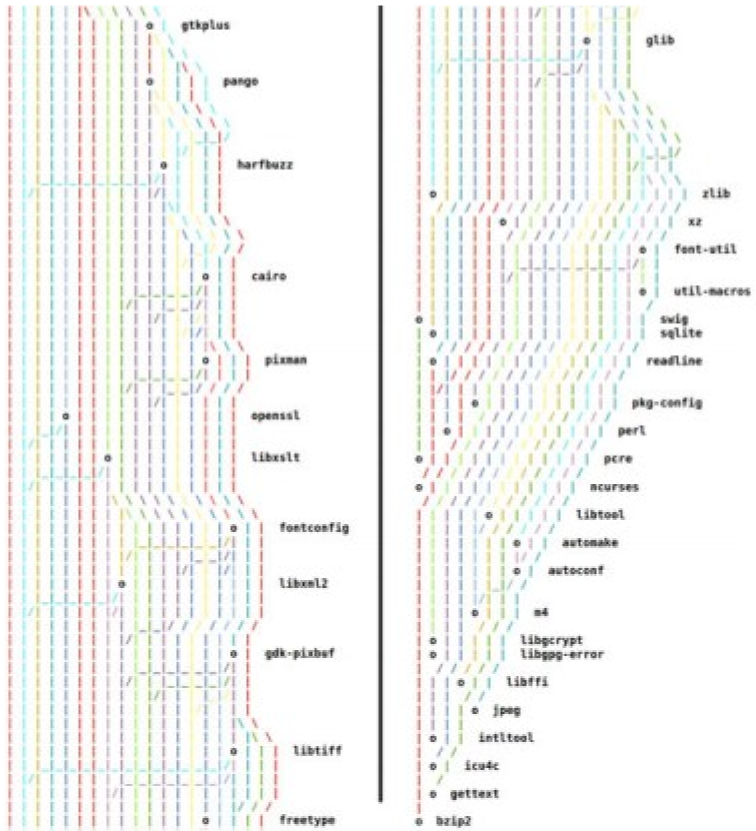




Design:
Drugs
Proteins
Materials
Nanosystems

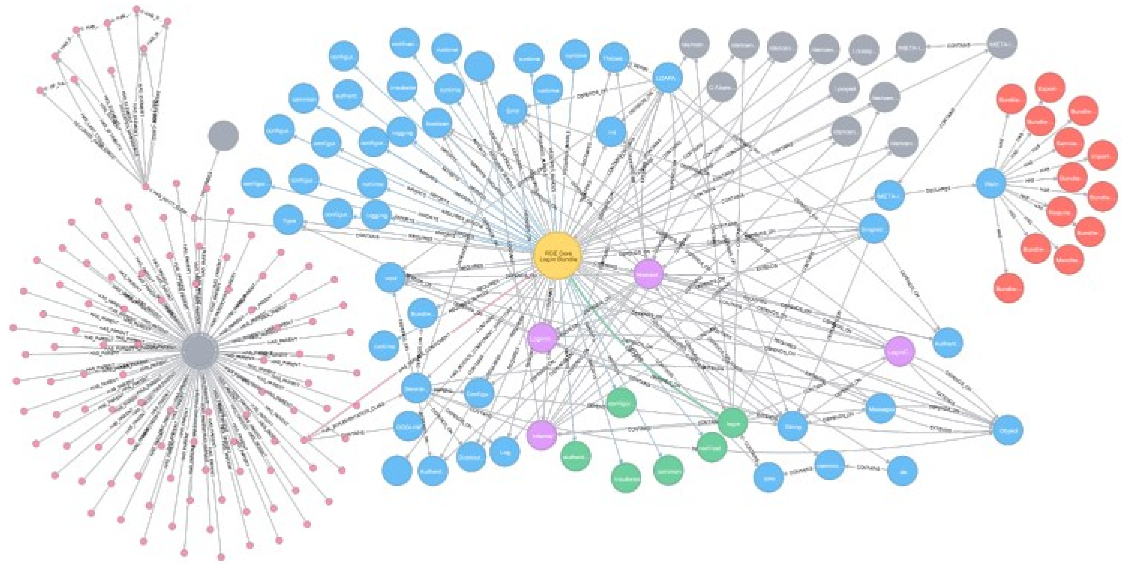
https://www.samson-connect.net/

Software visualization

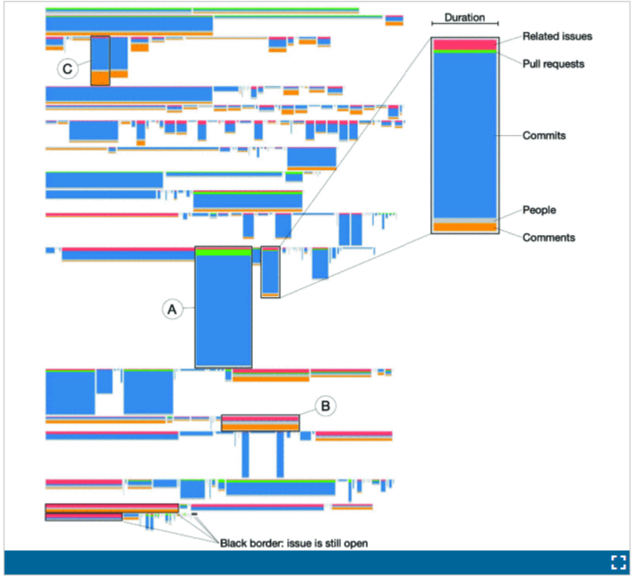


git-style package dependency graph of dia (also shown in Fig. 1). The freetype node has been duplicated to show alignment between the two halves.

https://ieeexplore.ieee.org/document/8419271



https://ieeexplore.ieee.org/document/8742198

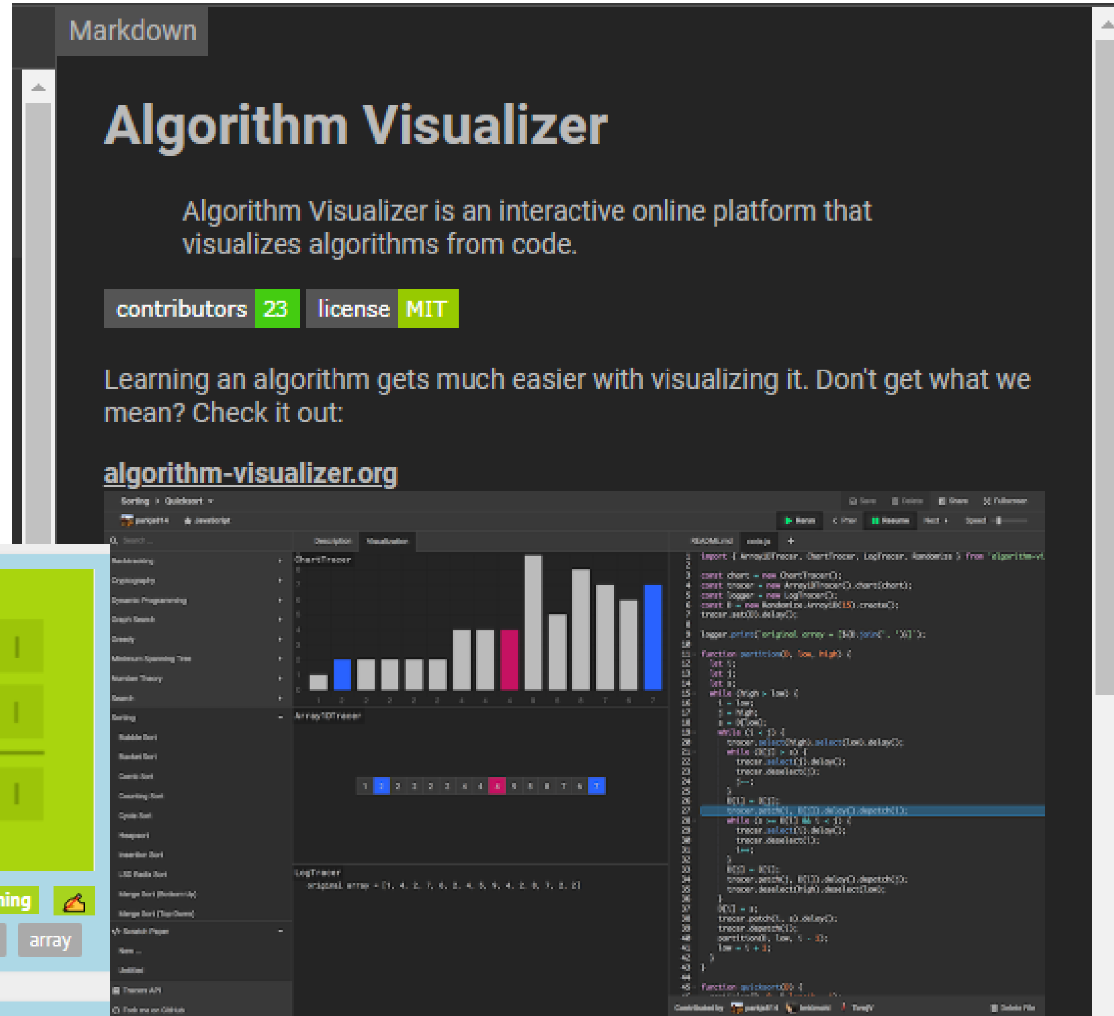


https://ieeexplore.ieee.org/document/9604892

Algorithm visualization

- Beyond mathematical and empirical analyses of algorithms

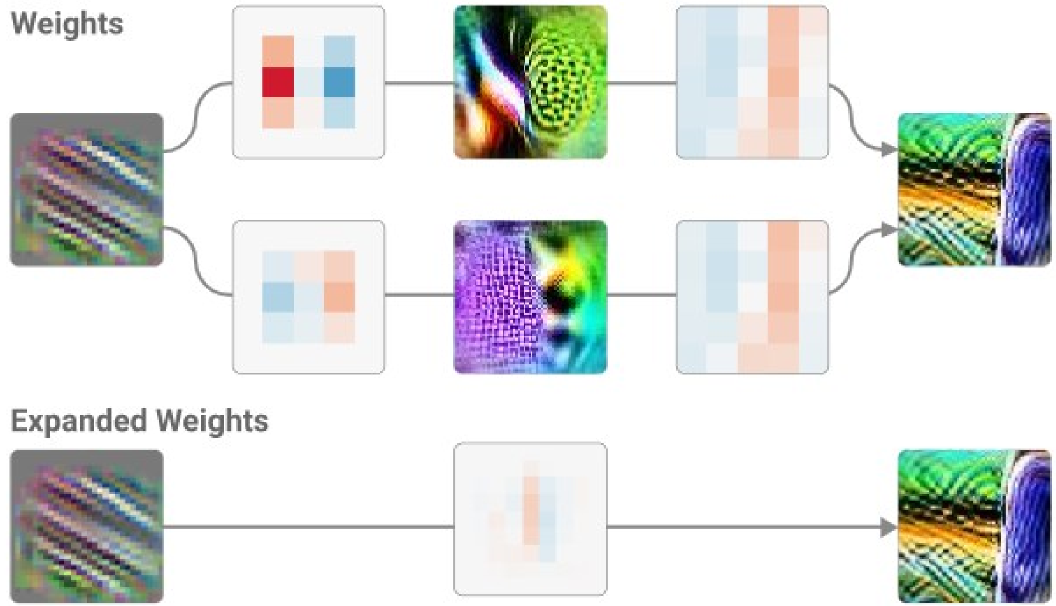
https://visualgo.net/en

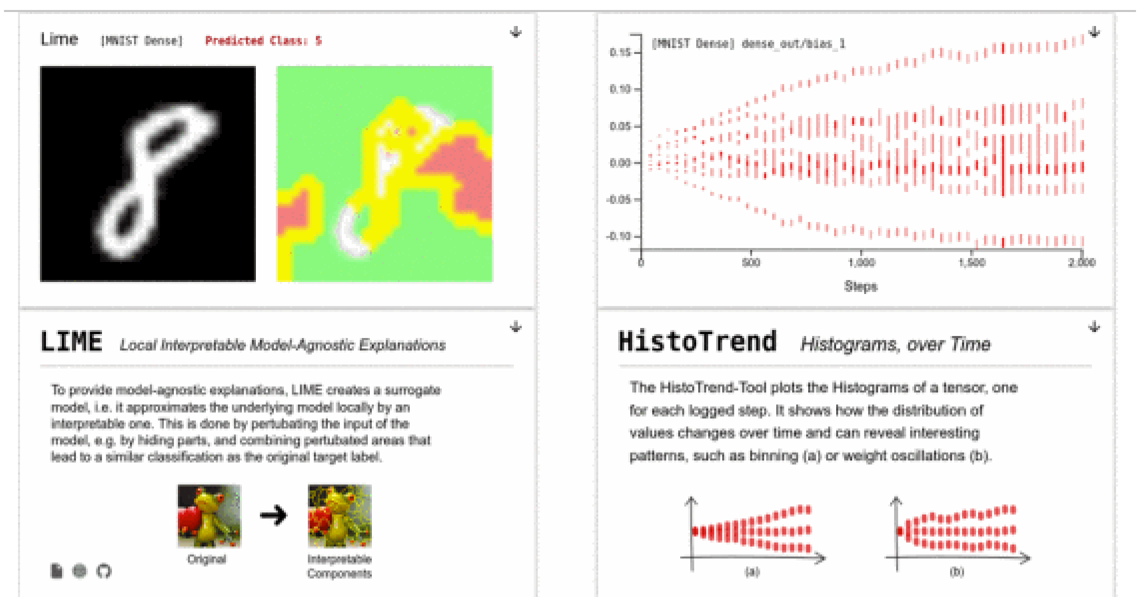


https://algorithm-visualizer.org/

Machine Learning visualization

- To help understand the “inner workings” of neural networks and other AI methods





(a) LIME (high-abstraction explainer) (b) HistoTrend (low-abstraction explainer)

https://distill.pub/2020/circuits/visualizing-weights/

https://ieeexplore.ieee.org/document/8807299

Medicine
(education)

- Human anatomy

- using volume rendering

- VOXELman (University of Hamburg)

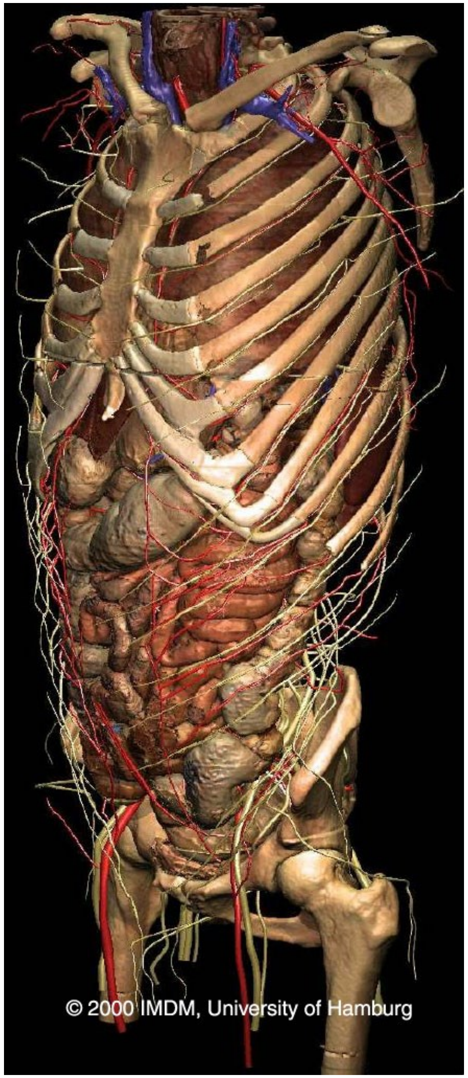
- Visible Human project
  (National Library of Medicine-USA)

https://www.visiblebody.com/

http://www.voxel-man.de/3d-navigator/inner_organs/

http://www.nlm.nih.gov/research/visible/visible_human.html

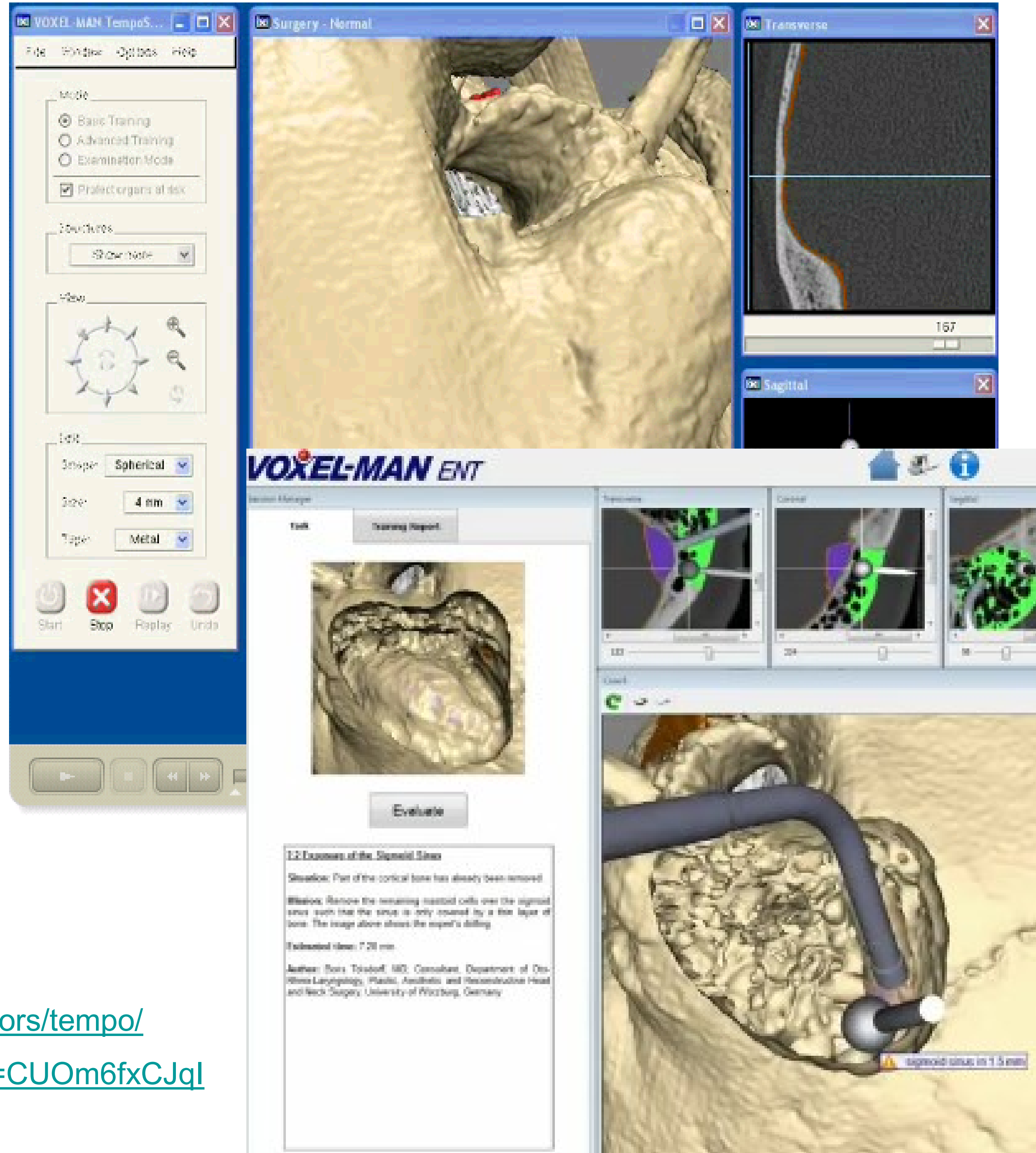
https://www.nlm.nih.gov/research/visible/applications.html



Medicine
(e.g. surgery training)

VOXELman,
University of Hamburg

- Temporal bone surgery
- Movement of the drill is controlled with a force feedback device



https://www.voxel-man.com/simulators/tempo/

https://www.youtube.com/watch?v=CUOm6fxCJqI

Dentistry (e.g. training)

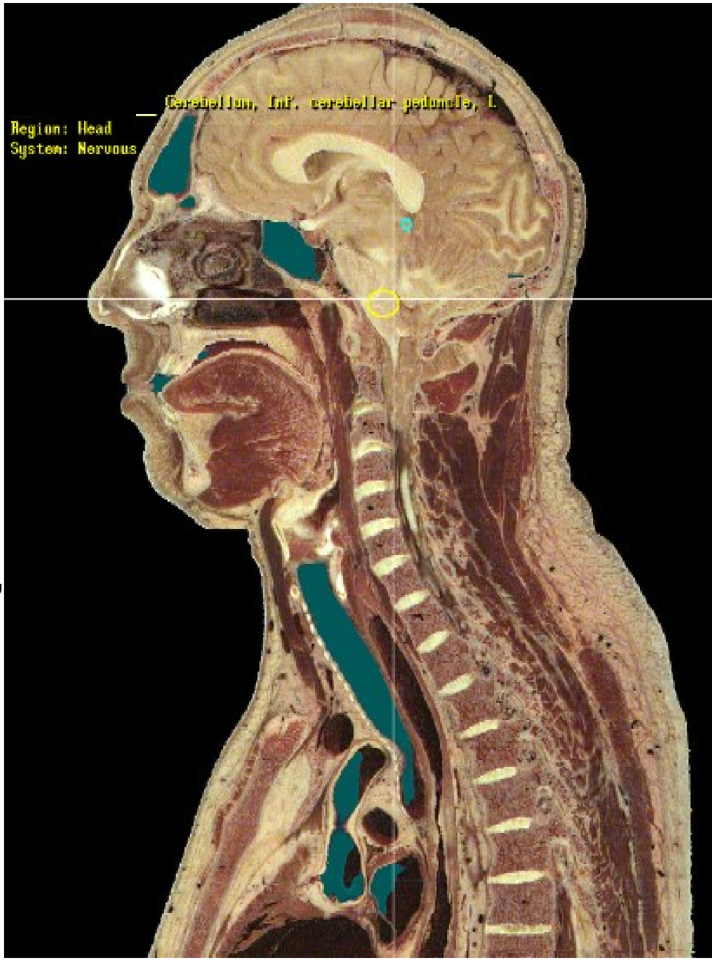
Stereoscopic display + glasses

Interaction devices:
- two force feedback devices
- foot pedal

https://www.voxel-man.com/simulators/dental/

https://www.youtube.com/watch?v=CB_vdW6K42o

An example of Scientific Visualization: The visible Human Project
(1994,1995)



The data sets were designed to serve as
(1) a reference for the study of human anatomy,
(2) public-domain data for testing medical imaging algorithms,
(3) a test bed and model for the construction of network-accessible image libraries.

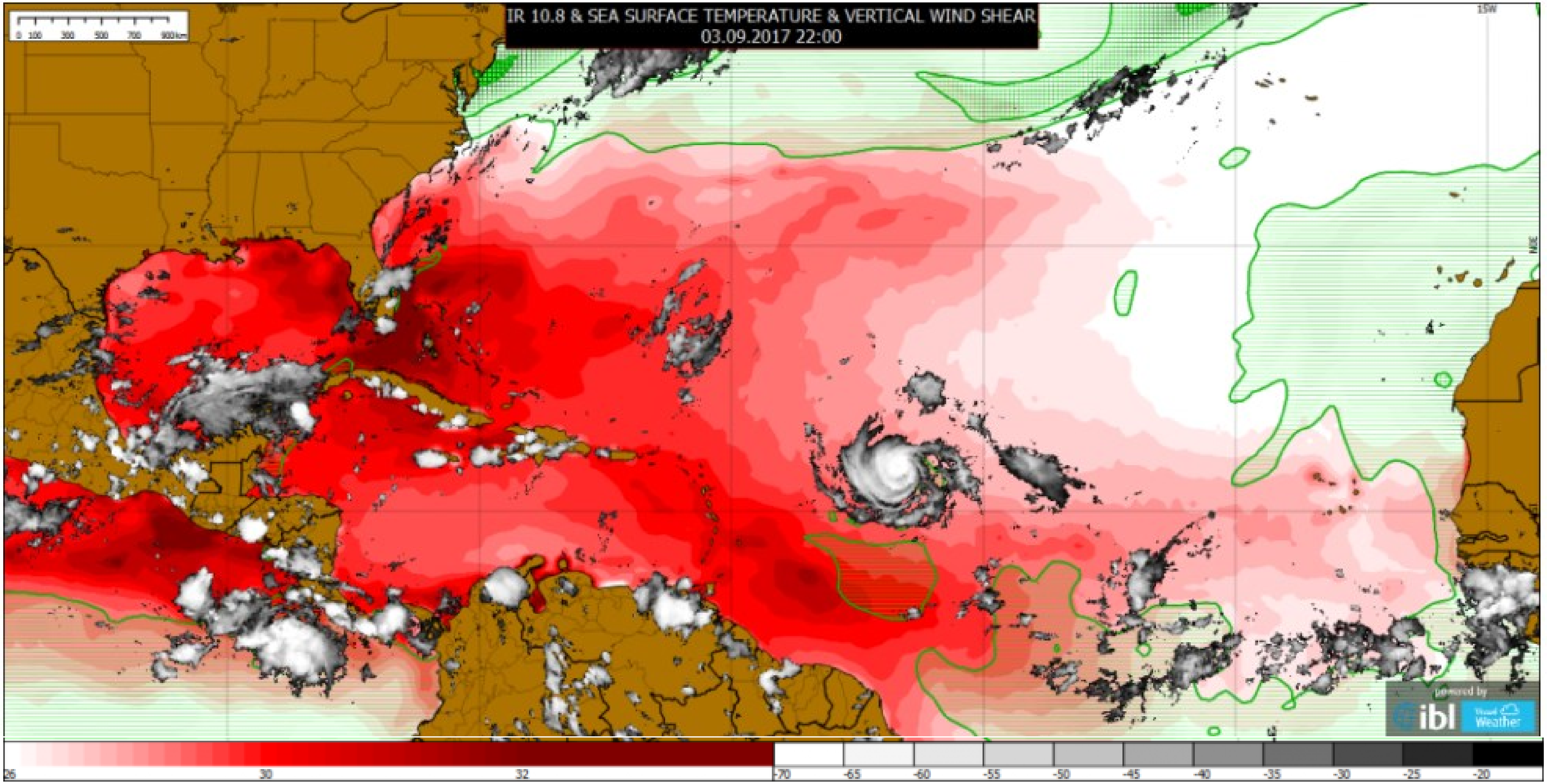
Have been applied to a wide range of educational, diagnostic, treatment planning, virtual reality, artistic, mathematical, and industrial uses.

About 4,000 licensees from 66 countries

As of 2019, a license is no longer required to access the VHP datasets.

https://www.nlm.nih.gov/research/visible/applications.html

Meteorology and oceanography

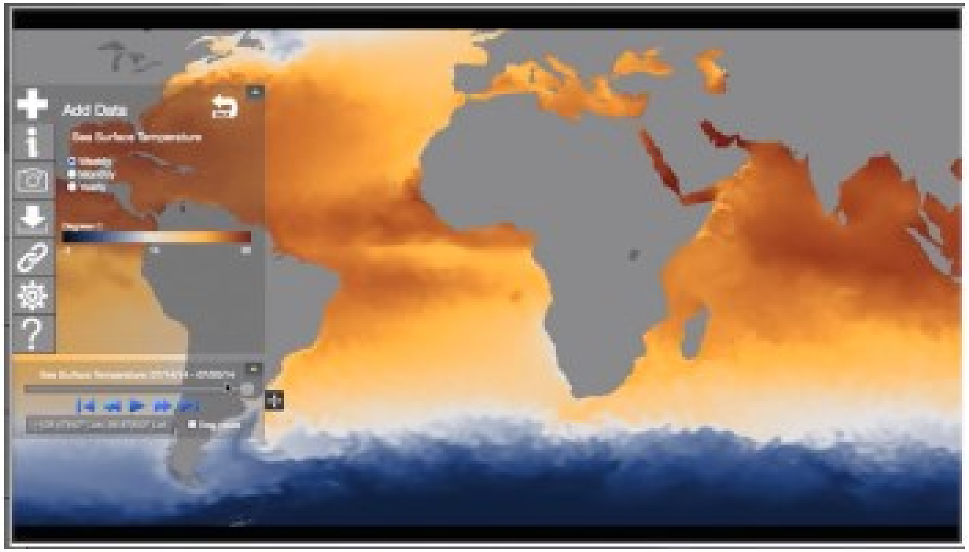


https://www.iblsoft.com/products/visualweather/

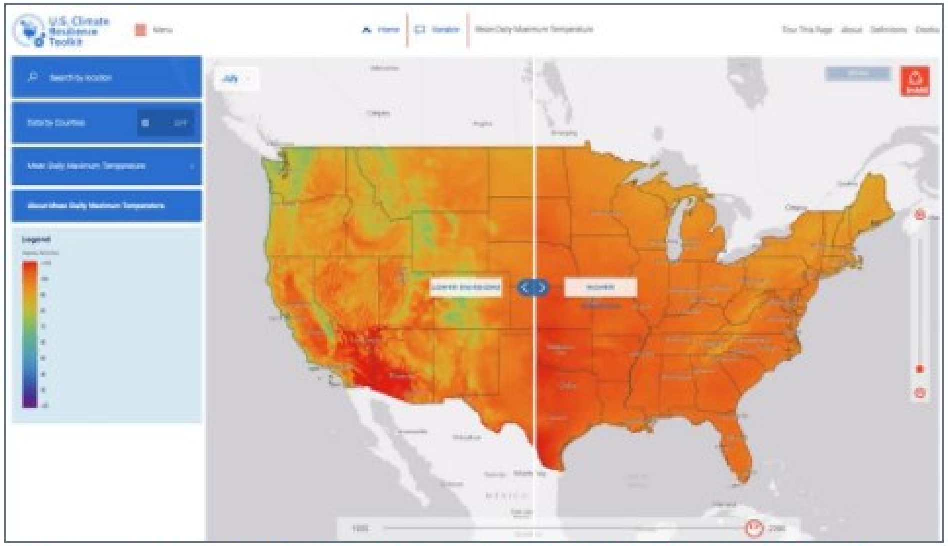
Climate research
(by NOAA)

•The Climate Explorer offers graphs, maps, and data of observed and projected temperature, precipitation, and related climate variables for every county in the contiguous US

• The tool shows projected conditions for two possible futures:
  – one in which humans make a moderate attempt to reduce global emissions of heat-trapping gases,
  – one in which we go on conducting business as usual.



https://www.climate.gov/maps-data/primer/visualizing-climate-data

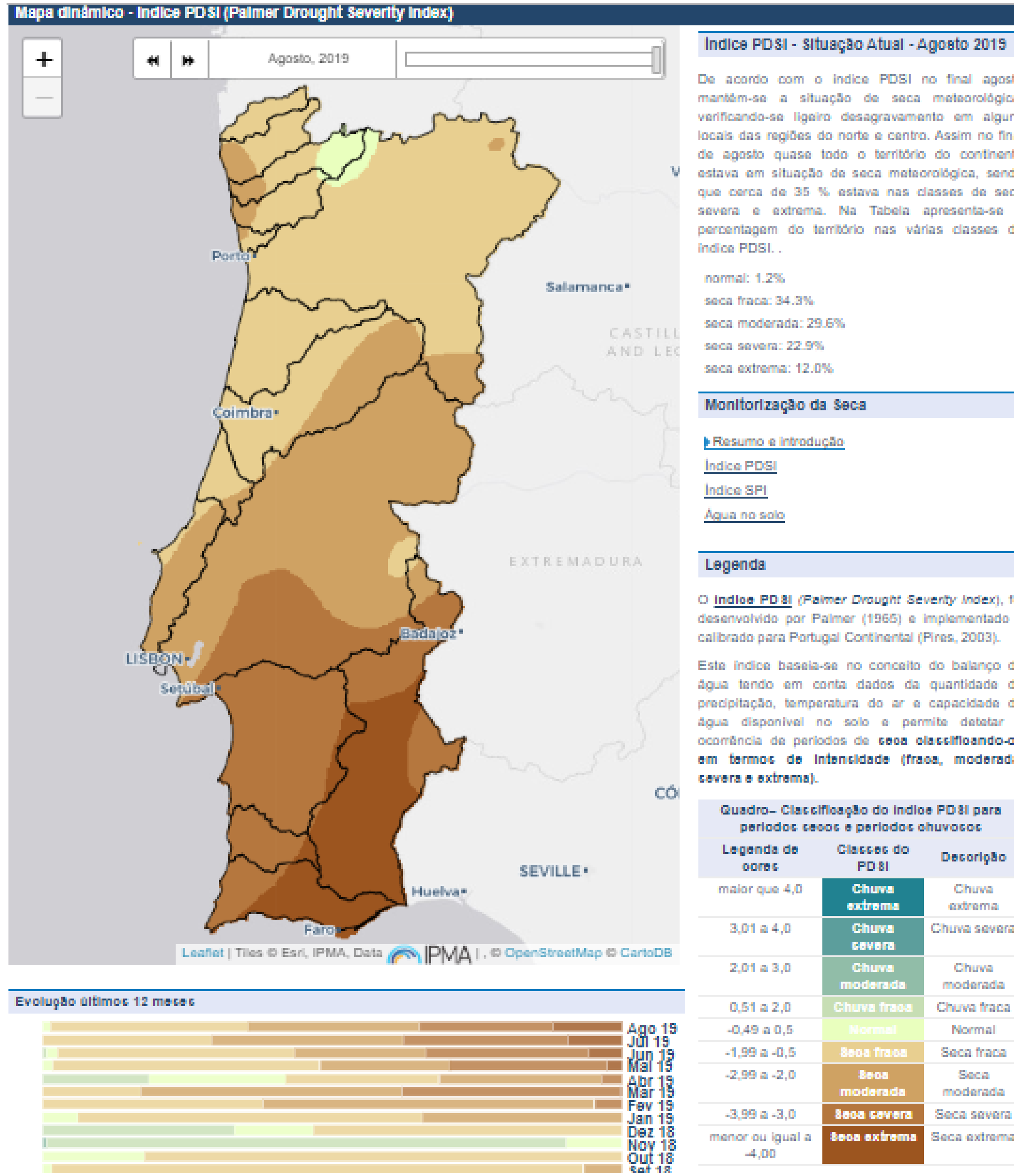


View by Variable interface. View Maximum Daily Temperature variable in Climate Explorer.

https://toolkit.climate.gov/tools/climate-explorer

Example in Climate monitoring:

Drought Severity Index (by IPMA)

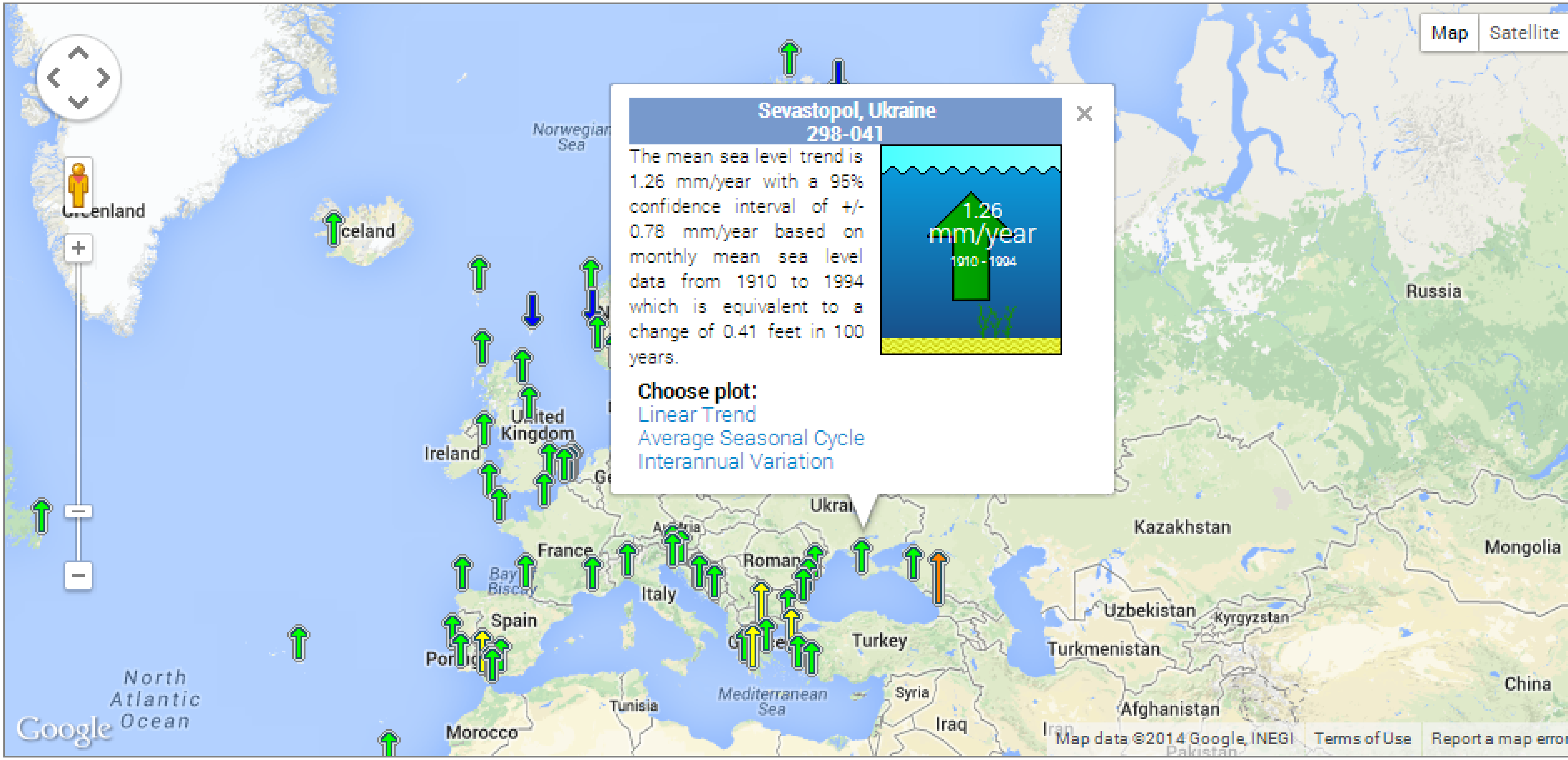


http://www.ipma.pt/pt/oclima/observatorio.secas/

Sea Level Trends

East Coast | West Coast | Gulf Coast | Alaska | Hawaii | Global

View in Google Earth



The map above illustrates regional trends in sea level, with arrows representing the direction and magnitude of change. Click on an arrow to access additional information about that station.

Sea Level Trends
mm/yr (feet/century)

15 to 21 (5 to 7)	6 to 9 (2 to 3)	-3 to 0 (-1 to 0)	-12 to -9 (-4 to -3)
12 to 15 (4 to 5)	3 to 6 (1 to 2)	-6 to -3 (-2 to -1)	-15 to -12 (-5 to -4)
9 to 12 (3 to 4)	0 to 3 (0 to 1)	-9 to -6 (-3 to -2)	-18 to -15 (-6 to -5)

https://tidesandcurrents.noaa.gov/sltrends/regionalcomparison.html?region=GNEATL

- “***human-in-the-loop***” problems involve the user as a part of the system

- They are very complex due to the facts that:

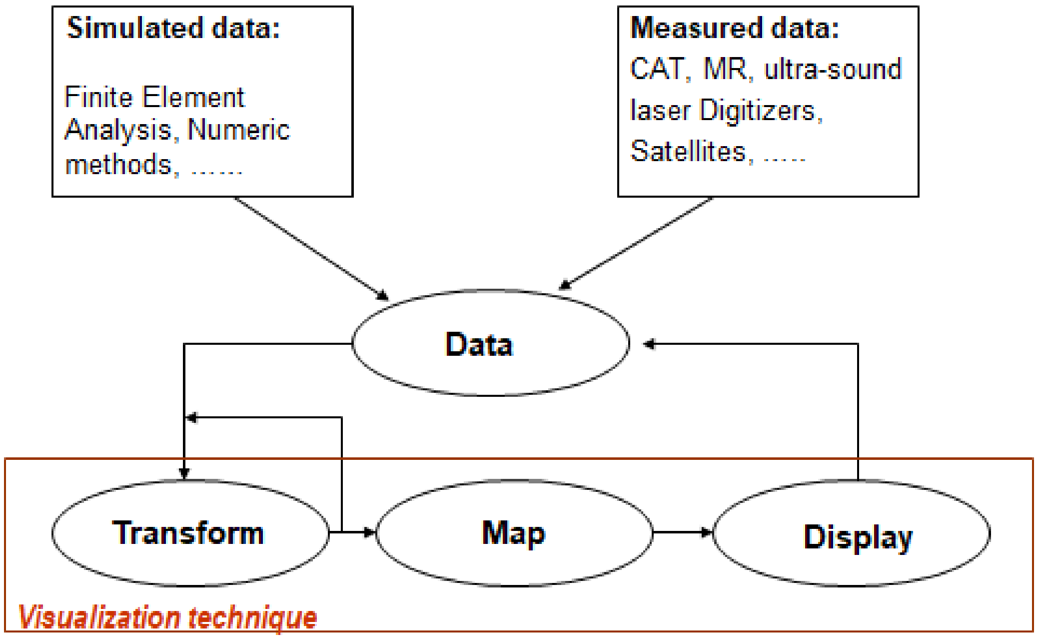
  - humans are very complex systems

  - 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

Scientific Visualization reference model



(adapted from Schroeder et al., 2006)

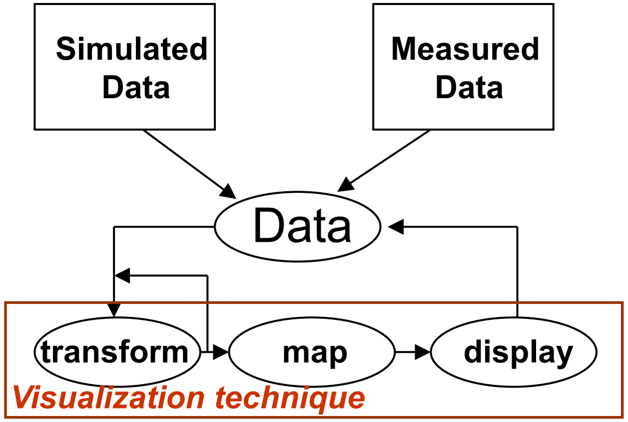
The visualization creator is involved in all the phases after obtaining the data

The user should get insights from the visualization

•:**Data can be**

- simulated
  (e.g. stress of a mechanical part, phantom of the human body, etc.)

- measured from real phenomena



- Then a **visualization technique** is applied, involving:

  - data **transformation** through several methods
    (e.g. scale transformation, noise filtering, outlier elimination, changing resolution, etc.)
  - **mapping** to an adequate form to represent data visually
    (e.g. lines, points, color)

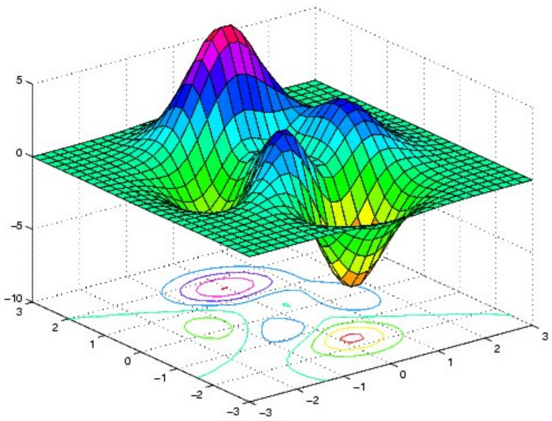
  - producing an image or sequence of images (rendering)

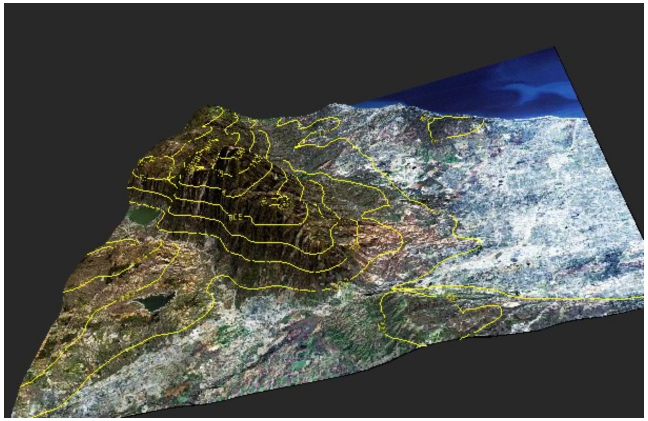
- **This process is repeated as needed to provide insight**

- The choice of the right mapping is fundamental

- Consider the values of a function or terrain altitude data, or sea depth:

  - different mappings or visualization techniques can be used, e.g.
    - three-dimensional surface
    - pseudo-color
    - contours (isolines)





- Visualization may be used with different purposes:

  - personal exploration
  - discussion with colleagues
  - presentation to other people

for

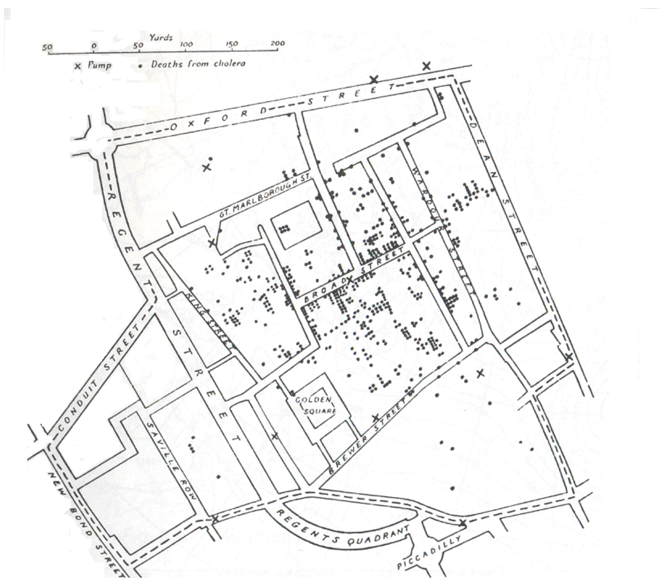
  - explorative analysis
  - confirmative analysis

Classical examples for:

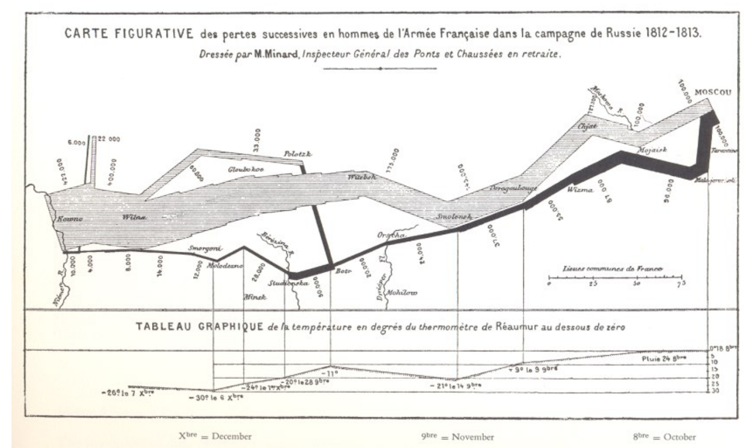
a) exploration

b) presentation

a)



b)



Example of Presentation to other people: World health by Hans Rosling: 200 years of health/income – 120 000 values in 4 min

https://www.youtube.com/watch?v=jbkSRLYSojo

Whatever the purpose, a visualization:

- Should allow **offload internal cognition and memory** usage to the **perceptual system**, using **carefully designed images** as a form of external representations (external memory)

- To **support users' tasks**

To design simple or complex visualizations:

- Need to **find what are the questions** users will ask!

Example: how to select simple charts?

Max and Min temperatures along the month of February (in ºC):

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

Q1- What were the maximum and minimum values of MaxT?

Q2- What was the most frequent MaxT?

Q3- In how many days was that MaxT value attained?

Q4- How were the daily temperature ranges?

Q5 – What was the maximum temperature range?

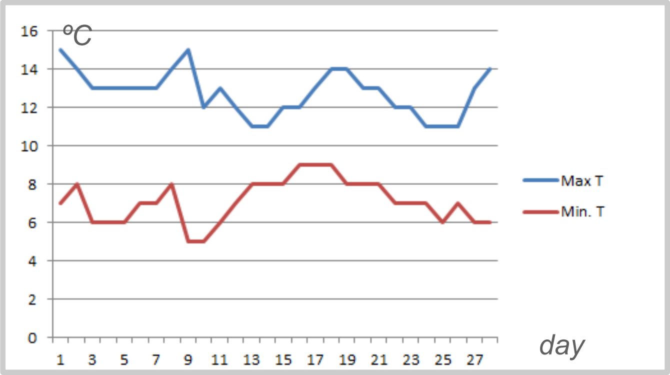
- What type of chart would you use to answer Q1?

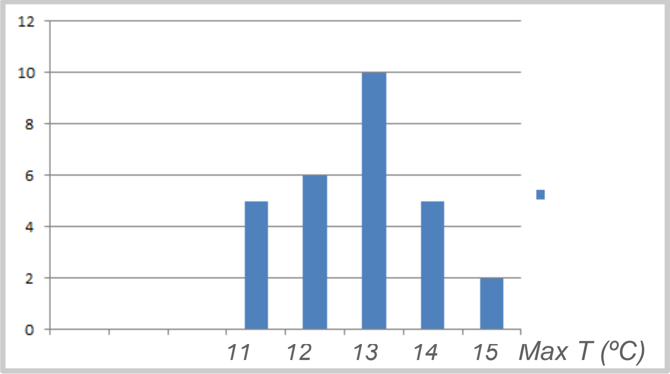
- And the other questions?

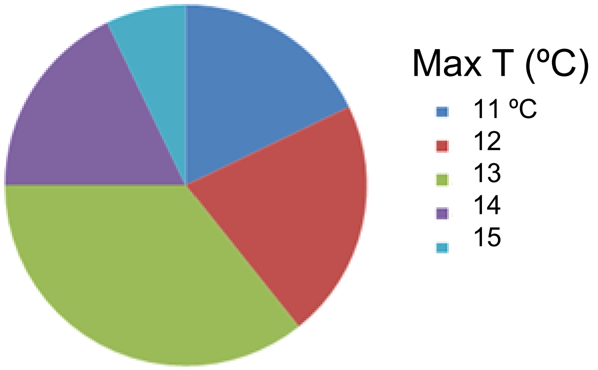
Example: how to select simple charts?

Temperatures along the month of February (in ºC): a few possible charts

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



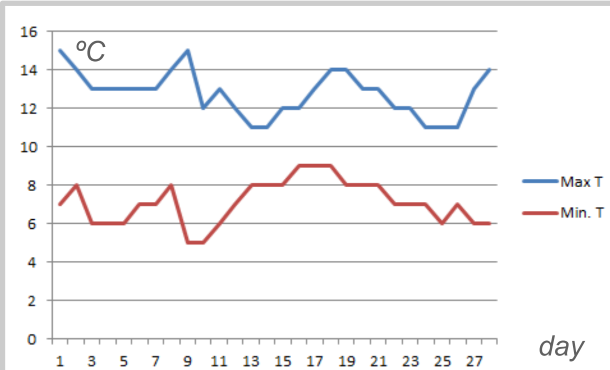




Simple example

Temperatures along the month of February (in ºC):

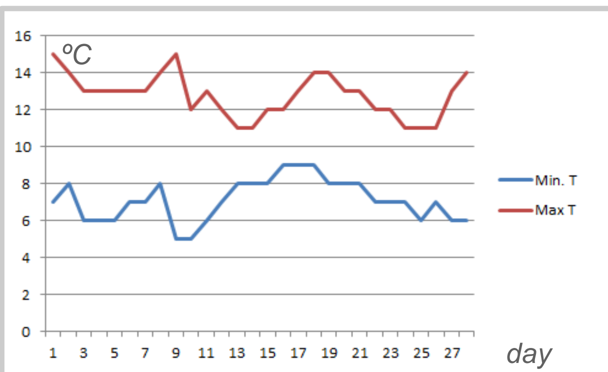
Anything “odd” about this chart?

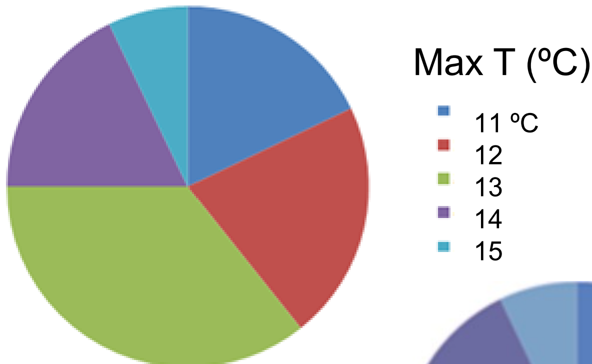


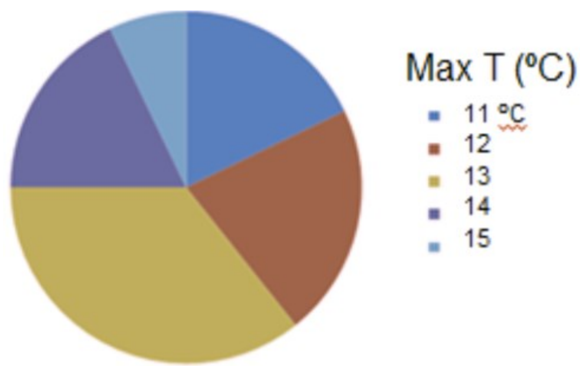
What if the user is color-blind?

Test it using https://www.color-blindness.com/coblis-color-blindness-simulator/

Would you prefer this one?







Do not forget “cultural” aspects,
nor individual differences!

Next sessions:

- Data characteristics, the phenomena they represent and pre-processing
- Human characteristics fundamental for Visualization
- Creating a Visualization: visually representing
  - 1D, 2D, 3D and nD quantitative data
  - Other types of data (maps, networks, hierarchical data, text…)
- Effective Visualization

Bibliography

- Munzner, T., *Visualization Analysis and Design*, A K Peters/CRC Press, 2014
- Kirk, A., *Data Visualisation: A Handbook for Data Driven Design*, 2nd. Ed., Sage, 2019
- Kirk, A., *Data Visualization : a successful design process*. Packt Publishing., 2012
- Spence, R., *Information Visualization, Design for Interaction*, Prentice Hall, 2007
- *Tufte, E., The Visual Display of Quantitative Information*, Graphics Press, 1983
- Tufte, E., *Envisioning Information,* Graphics Press, 1990
- Friendly, M., "Milestones in the history of thematic cartography, statistical graphics, and data visualization“, 2008

Images of the 1rst slide:

B. Marques, S. Silva, J. Alves, T. Araújo, P. Dias, B. Sousa Santos, “A conceptual model and taxonomy for collaborative Augmented Reality”, *IEEE Transactions on Visualization and Computer Graphics*, 2021.

V. Gonçalves, P. Dias, M. J. Fontoura, R. Moura, and B. Sousa Santos, “Investigating landfill contamination by visualizing geophysical data.,” *IEEE Comput. Graph. Appl.*, vol. 34, no. 1, pp. 16–21, 2015.

S. Silva, B. Sousa Santos, J. Madeira, “Exploring different parameters to assess left ventricle global and regional functional analysis from coronary CT angiography”. *Computer Graphics Forum*, vol. *31*, n. 1, 146–159, 2012.