Lecture 8
Visual Perception and Colors

[Data Visualization · 1-DAV-105]

Lecture by Broňa Brejová

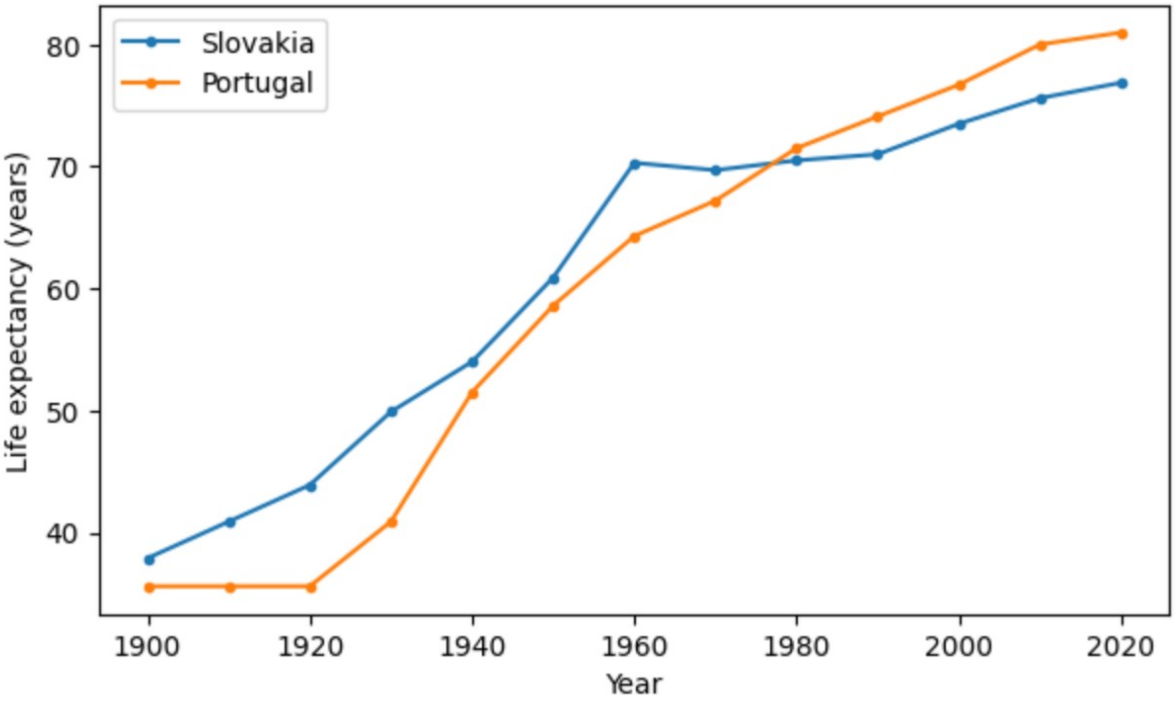
Acknowledgement: materials inspired by lectures from Martina Bátorová in 2021

Why talk about visual perception in visualization?

In which period of time was life expectancy higher in Slovakia than in Portugal?

Country	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020
Slovak Republic	37.9	40.9	43.9	49.9	54.0	60.9	70.3	69.7	70.5	71.0	73.5	75.6	76.9
Portugal	35.6	35.6	35.6	40.9	51.5	58.6	64.3	67.2	71.5	74.1	76.7	80.0	81.0

In which period of time was life expectancy higher in Slovakia than in Portugal?



Visual brain, table vs. graph

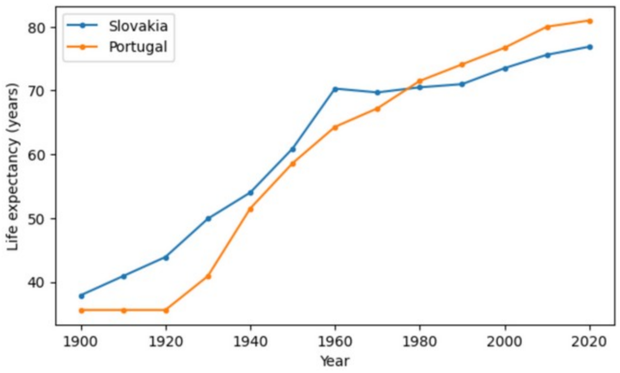
We "read" tables, verbal processing

We "see" plots, visual processing

Visual processing is very parallel and fast, evolved to spot predators

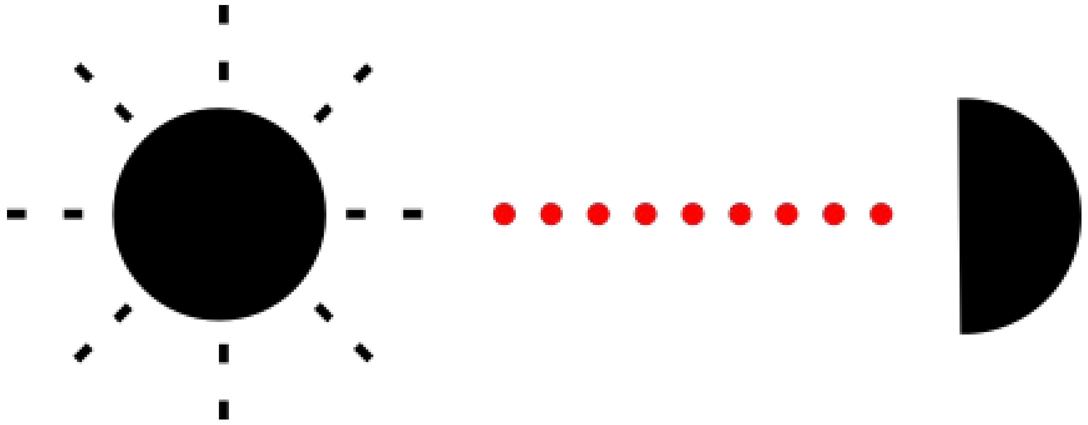
In which situations is a table preferable to a plot?

	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020
Country													
Slovak Republic	37.9	40.9	43.9	49.9	54.0	60.9	70.3	69.7	70.5	71.0	73.5	75.6	76.9
Portugal	35.6	35.6	35.6	40.9	51.5	58.6	64.3	67.2	71.5	74.1	76.7	80.0	81.0



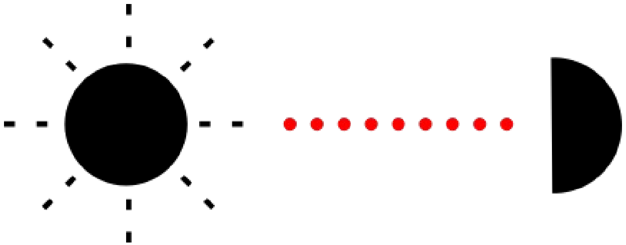
Human visual perception

What happens when we look at the figure below?



Human visual perception

Human visual perception



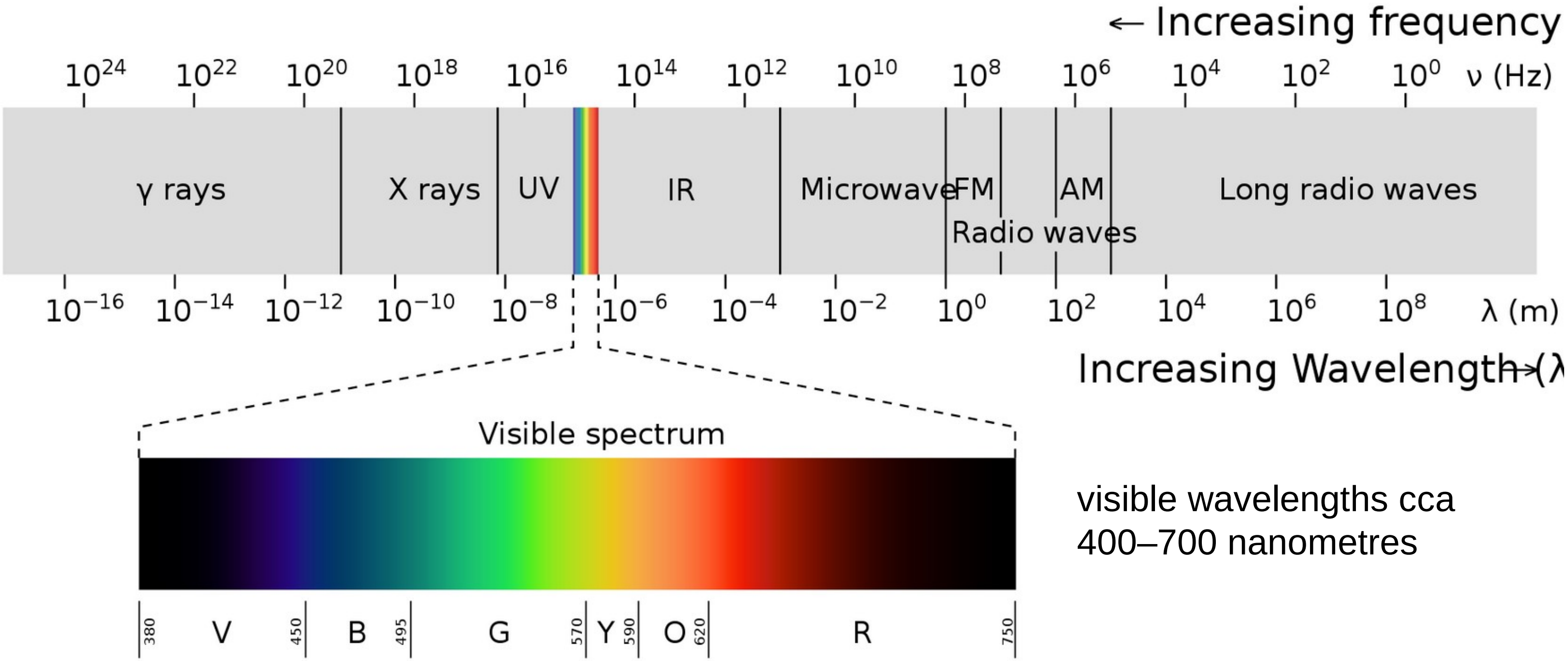
What happens when we look at the figure?

- The **light** from the screen / projector hits the retinas of our **eyes**
- Photoreceptor cells **transduce** (convert) this signal into nerve impulses
- In the brain:
  - detection of **basic features**
  - recognition of **patterns**
  - interpretation, assignment of **meaning**

Today: Focus on the light, eyes and colors, later stages next week

Light

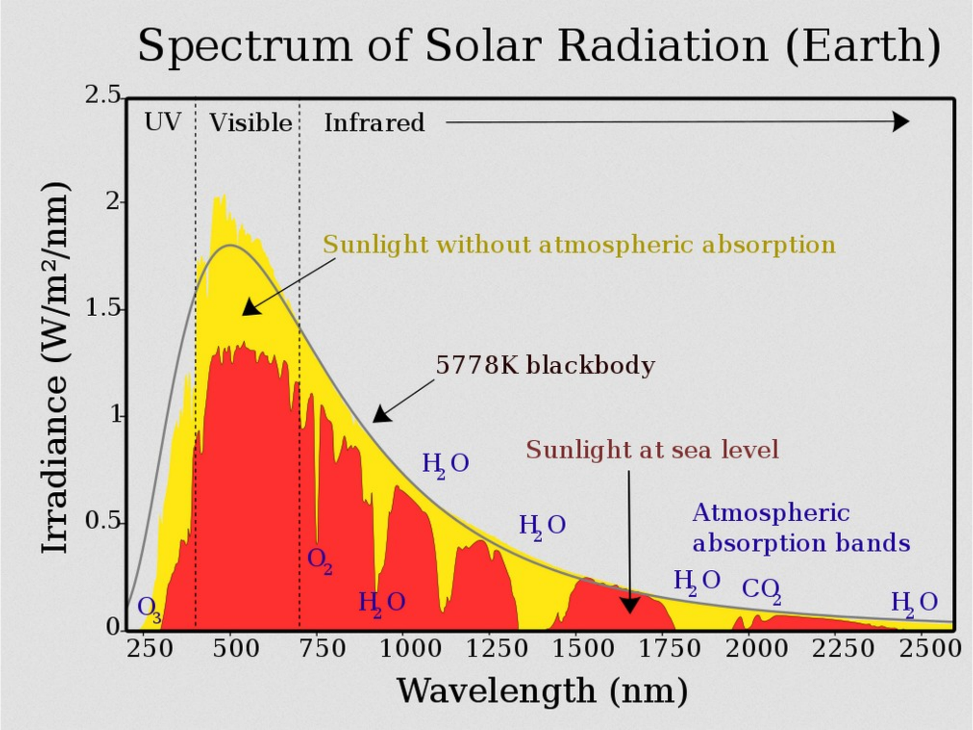
Visible light as a part of electromagnetic spectrum



visible wavelengths cca
400–700 nanometres

https://commons.wikimedia.org/wiki/File:EM_spectrum.svg

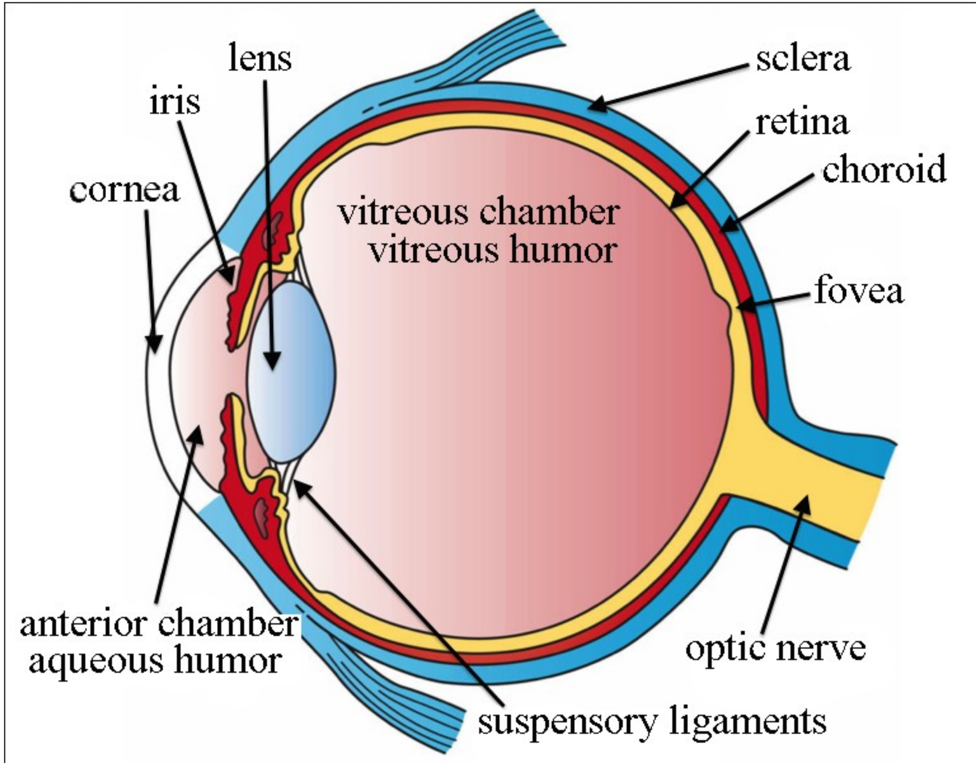
Sunlight is a mixture of different wavelengths



https://commons.wikimedia.org/wiki/File:Solar_spectrum_en.svg
https://commons.wikimedia.org/wiki/File:WhereRainbowRises.jpg

Human eye

Human eye



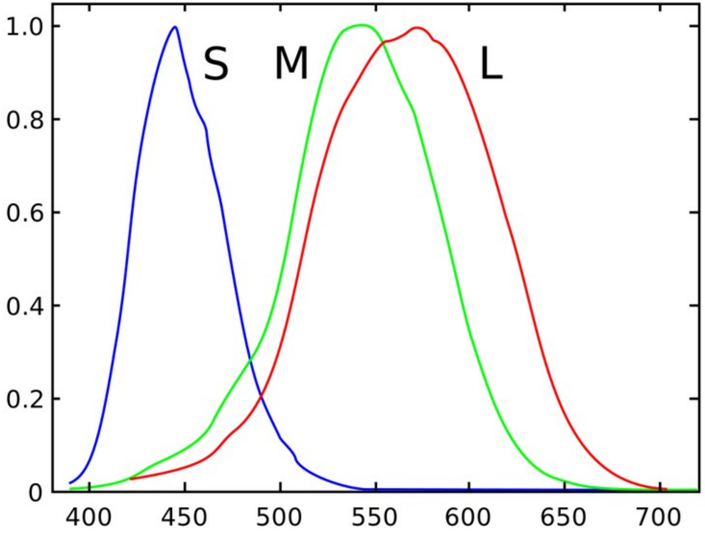
https://commons.wikimedia.org/wiki/File:Three_Main_Layers_of_the_Eye.png

Retina (sietnica): light-sensitive layer

Lens (šošovka): focus light to retina

Pupil (zrenica): hole in iris (dúhovka) where light enters the eye, its size regulated by the amount of light

Photoreceptor cells in the retina

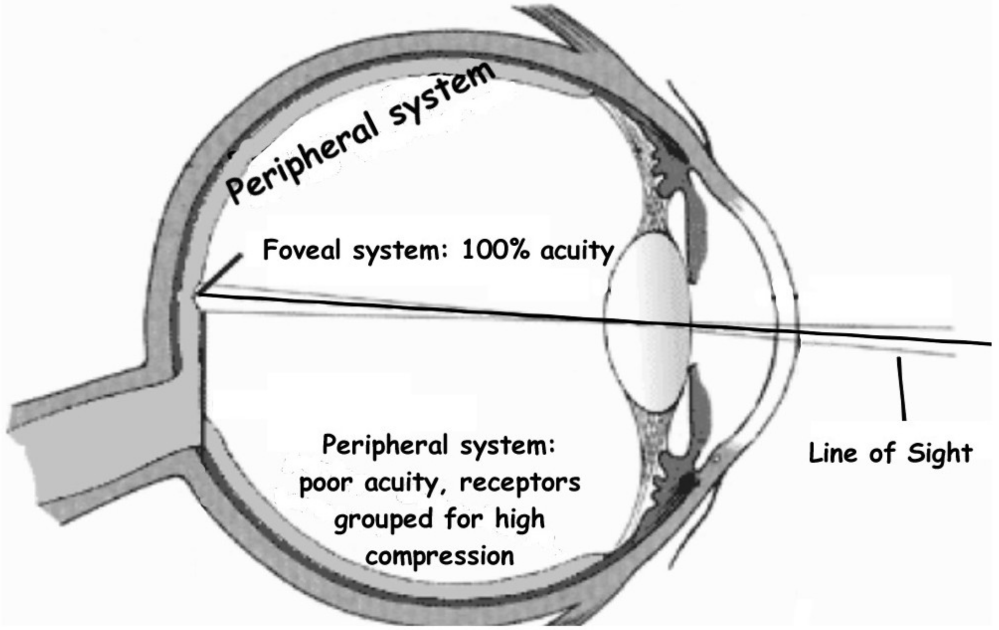


Rods (tyčinky): more sensitive to low light, not used for color vision

Cones (čapíky): color vision, three different types sensitive to different wavelengths (blue, green, red)

https://commons.wikimedia.org/wiki/File:Cones_SMJ2_E.svg

Foveal vs peripheral vision



Peripheral system for general holistic impression of the situation.
Foveal system for high quality details.

https://commons.wikimedia.org/wiki/File:Double_system_e.jpg

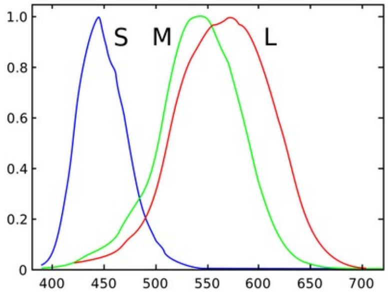
Fovea: central zone with many cones, sharp color vision, only about 1-2°

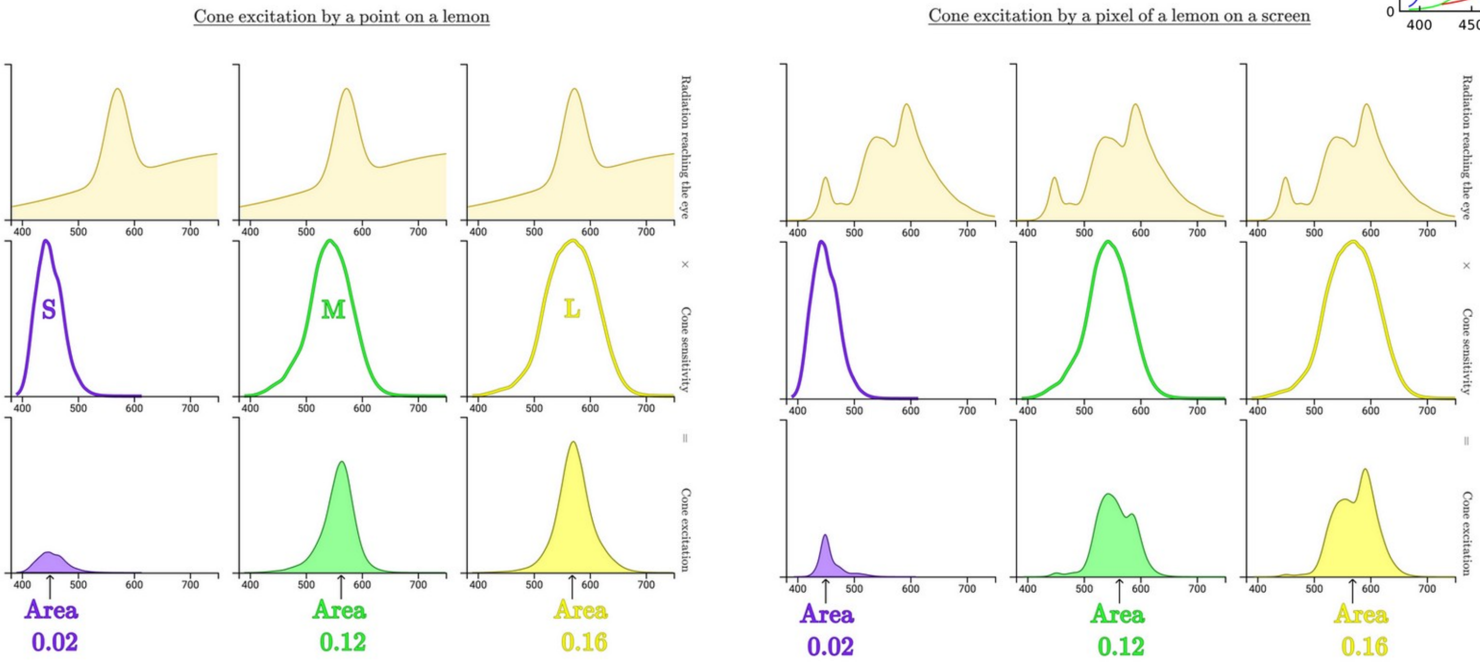
Peripheral vision: mostly rods, fast monochrome vision

The eyes make fast **movements** (saccades) between fixations on different points of interest to create a composite image

Colors and color spaces

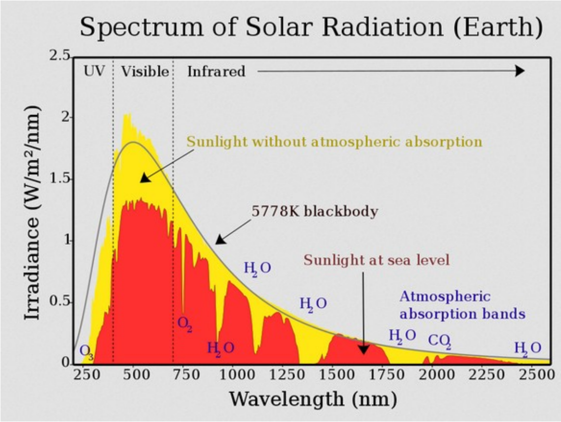
Metamers: light with different spectra that appear the same (to typical humans)

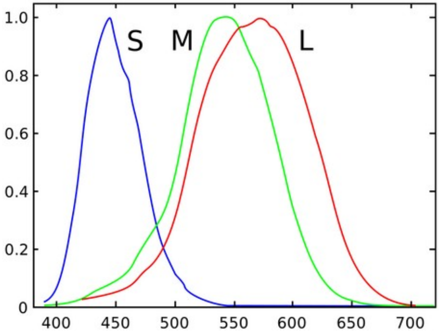




https://jamie-wong.com/post/color/

Color spaces, LMS





A **color space** is an organization of colors.

Our eye projects a full light spectrum into three values: response of the three types of cones.

S (short), M (medium), L (large) wavelength

LMS color space uses these three values to represent a color.

Metameric colors have the same values.

Derived models, e.g. CIE with better properties.

https://commons.wikimedia.org/wiki/File:Solar_spectrum_en.svg
https://commons.wikimedia.org/wiki/File:Cones_SMJ2_E.svg

Do you know some other color spaces?

Additive color models, RGB

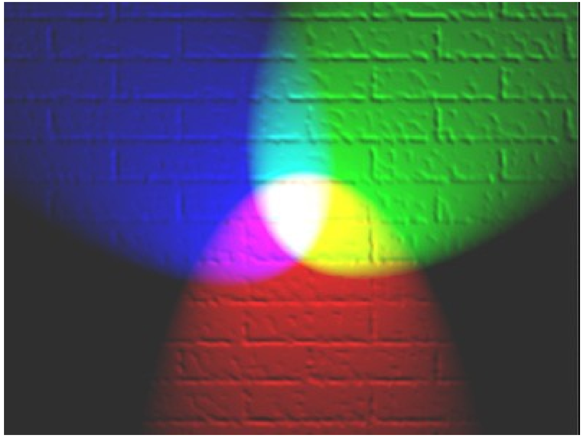
Monitors, projectors etc.

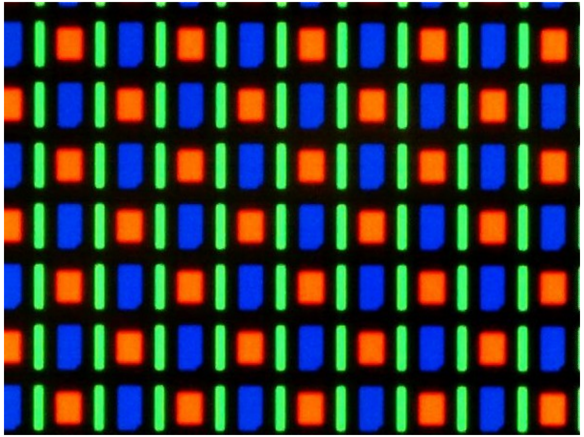
Component lights in **primary colors**,
other colors mixtures of these (adding up light).

White can be achieved by combining colors.

RGB uses **red, green, blue** as primary
(corresponds to LMS peaks).

The **gamut** is the set of colors representable by a device, usually a subset of the visible spectrum.



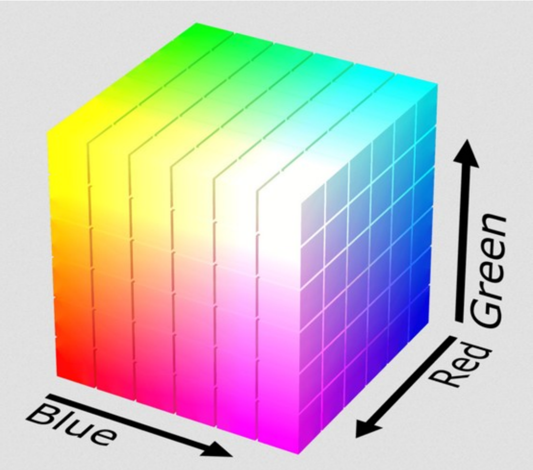


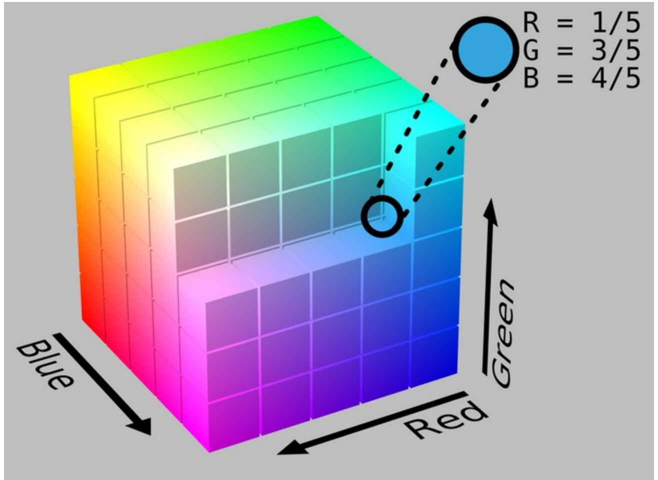
https://commons.wikimedia.org/wiki/File:RGB_illumination.jpg
https://commons.wikimedia.org/wiki/File:Nexus_one_screen_microscope.jpg

Colors in RGB space (RGB cube)

RGB is often used to specify colors.
Each coordinate e.g. a real number between 0 and 1 or integer between 0 and 255.
Also hexadecimal notation, e.g. #ff0000 is pure red.

Greytones on the main diagonal (x,x,x).





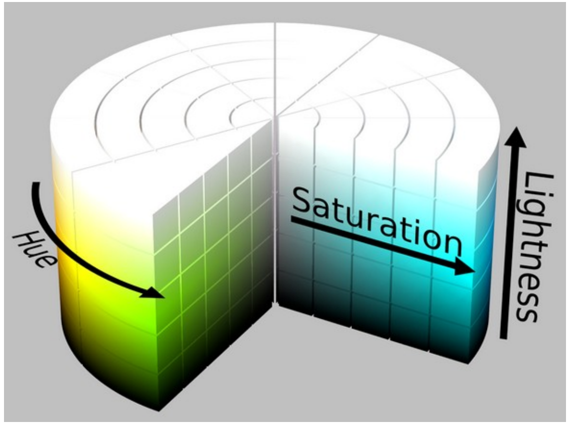
https://commons.wikimedia.org/wiki/File:RGB_color_solid_cube.png
https://commons.wikimedia.org/wiki/File:RGB_Cube_Show_lowgamma_cutout_b.png

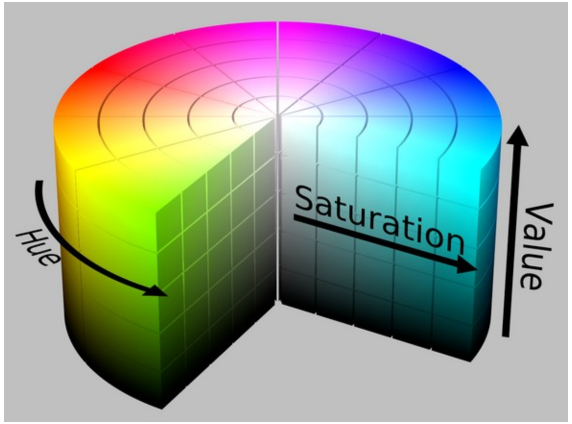
HSL and HSV color models

Transformations of RGB model with more intuitive coordinates.

Useful for color pickers, color palettes, image transformations etc.

Hue, saturation, lightness / hue, saturation, value.





https://commons.wikimedia.org/wiki/File:HSL_color_solid_cylinder_saturation_gray.png
https://commons.wikimedia.org/wiki/File:HSV_color_solid_cylinder_saturation_gray.png

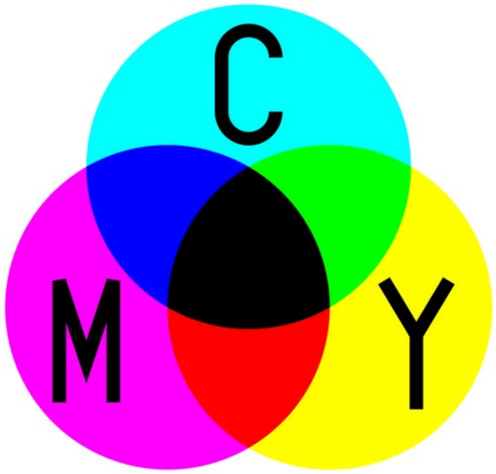
Subtractive models, pigments

Pigments block part of the light spectrum.

Adding more pigments blocks (subtracts) more light.

Black can be achieved by combining colors.

Example: CMY(K) color model used in **printing**.

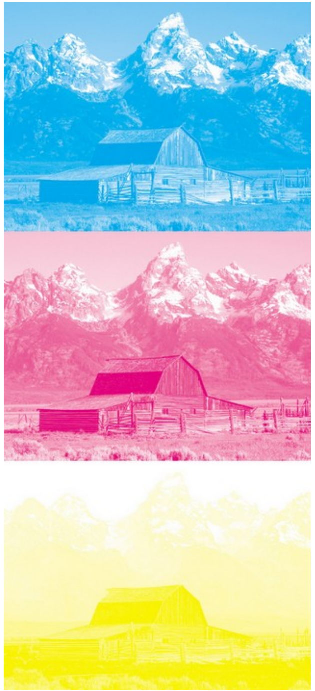


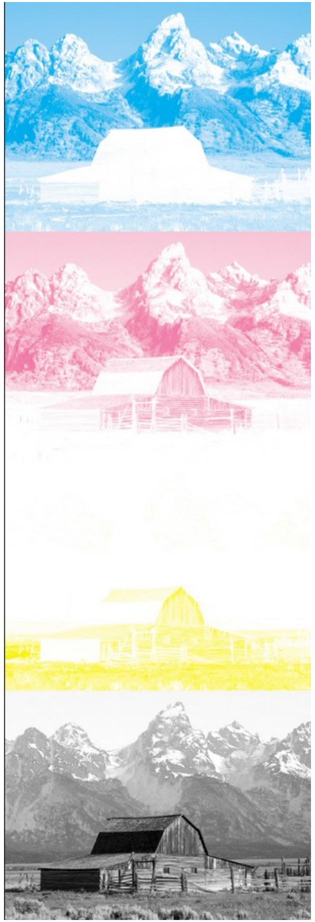
https://commons.wikimedia.org/wiki/File:SubtractiveColor.svg

CMY(K) color model

Primary colors cyan, magenta and yellow
- **cyan** absorbs red
- **magenta** absorbs green
- **yellow** absorbs blue

Black (K) added because
- it is cheaper
- it hides artifacts in dark colors





Conversion from RGB to CMYK is difficult, device-dependent.

https://commons.wikimedia.org/wiki/File:Barns_grand_tetons.jpg
https://commons.wikimedia.org/wiki/File:CMY_separation_%E2%80%93_no_black.jpg
https://commons.wikimedia.org/wiki/File:CMYK_separation_%E2%80%93_maximum_black.jpg

Color wheel, RYB model

Subtractive model developed in art, for mixing pigments

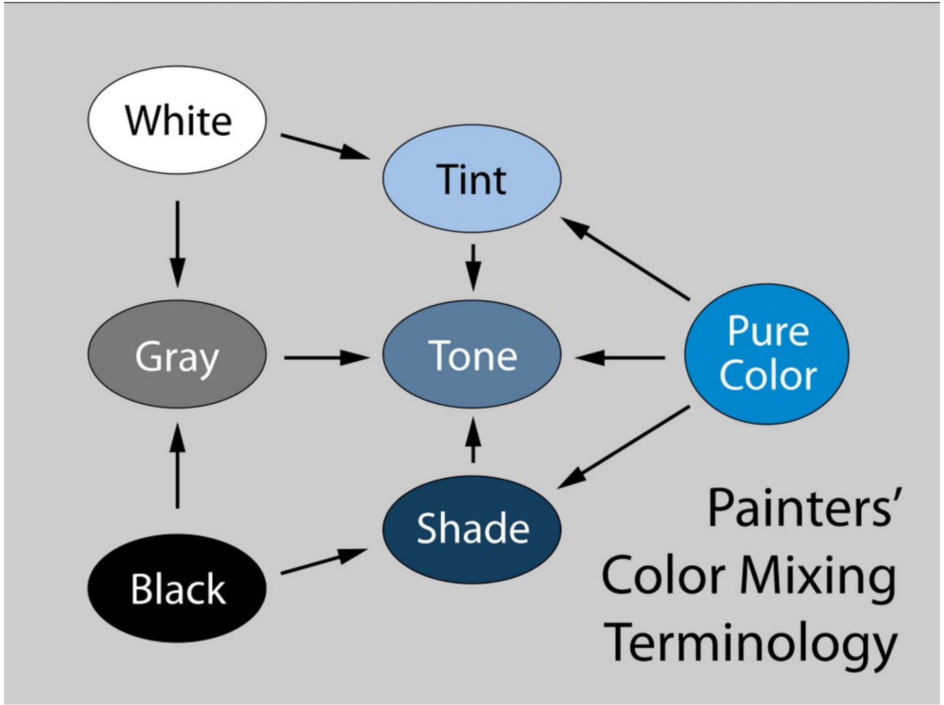
Dates back to 17th century

Primary colors red, yellow, blue

Secondary orange, green, purple (each a mix of two primaries)

https://www.w3schools.com/colors/colors_wheels.asp

Tint, tone, shade - more painter terminology



https://commons.wikimedia.org/wiki/File:Tint-tone-shade.svg

Examples of color schemes

Monochromatic: tints / tones / shades of the same hue

Complementary: 2 colors opposite on the color wheel
(e.g. orange and blue)

Split complementary: color and 2 neighbors
of its complement (e.g. orange and blue-green, blue-purple)

Analogous: 3-5 adjacent colors on the wheel

Each of these can be desaturated (tints / tones / shades)

See also https://color.adobe.com/create/color-wheel

https://www.w3schools.com/colors/colors_wheels.asp

Color and meaning, cultural differences

Colors often symbolize different things both within and between cultures:

- for example red: blood, love, passion, life, anger, violence, danger, emergency, speed, heat ...
  - China: good luck, prosperity vs. Europe: warning
  - "in the red" mean losses in English, what about China?
- mourning color is black in Europe, white in the East

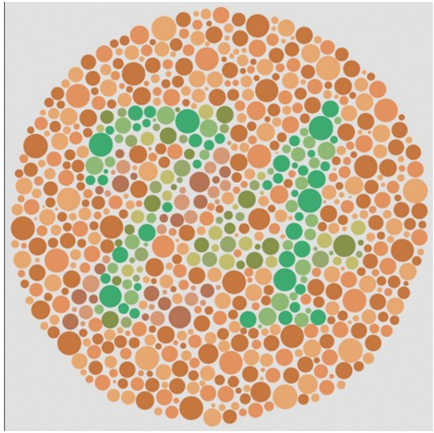
Colors in data visualization

Not everybody will be able to see your colors

Color blindness or color vision deficiency

Various forms and causes

Most often genetic red–green color blindness, where L or M opsin gene is mutated (8% of males)



https://commons.wikimedia.org/wiki/File:Ishihara_9.svg

Technical problems

- Projectors often distort colors

- Black-and-white printing

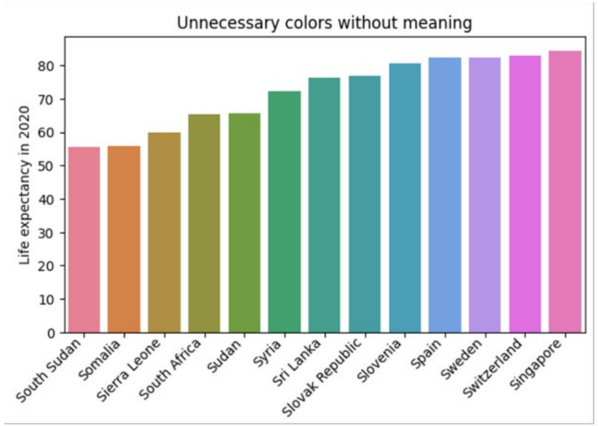
Choose your colors wisely to avoid these problems as much as possible

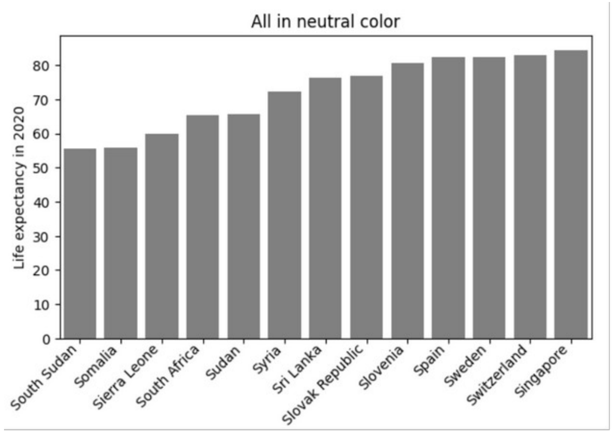
Color draws attention, use it sparingly

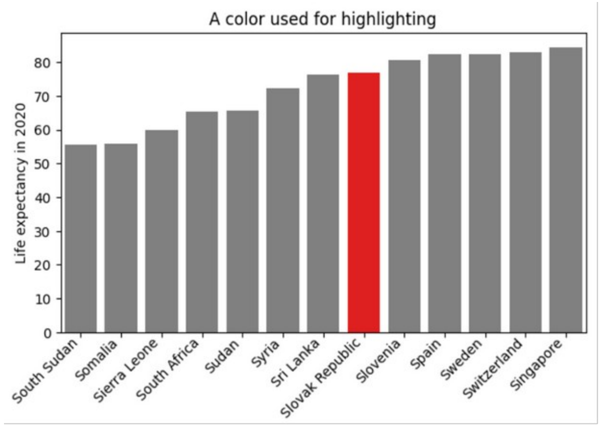
Rely on neutral tones

Use color sparingly, avoid unnecessary colors

Colors are great for highlighting points of interest



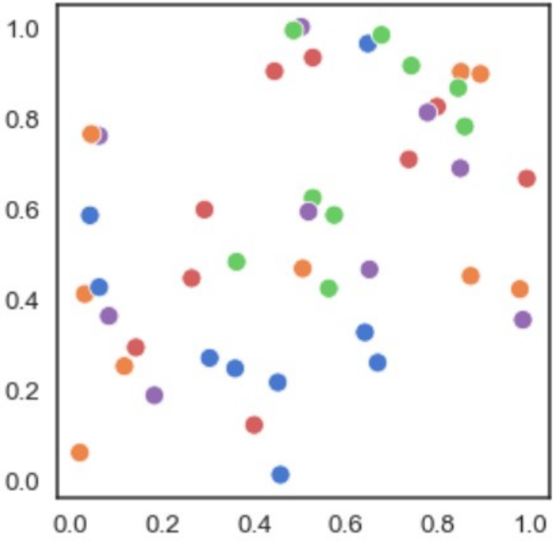


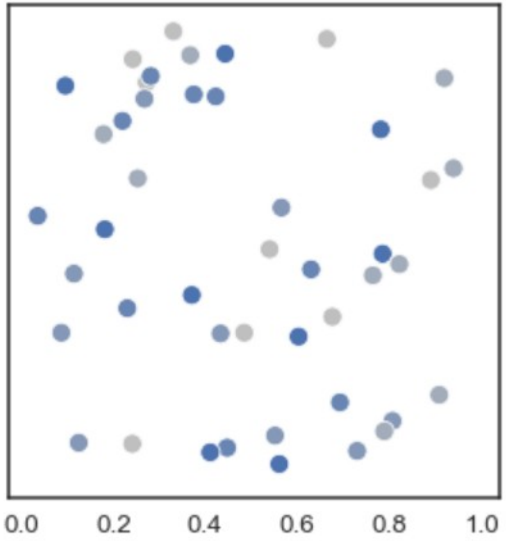


Qualitative palettes (for categorical variables)

Typically vary hues, easier to distinguish than lightness of the same hue

Try to keep the number of colors low



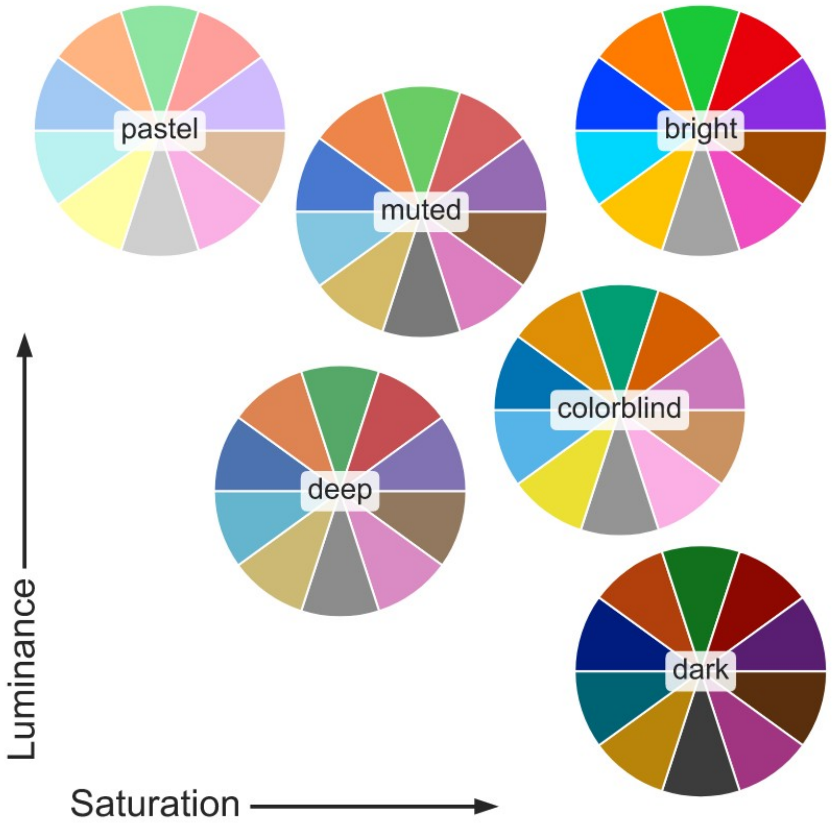


https://seaborn.pydata.org/tutorial/color_palettes.html

Qualitative palettes (for categorical variables)

Examples from Seaborn library

on the right:



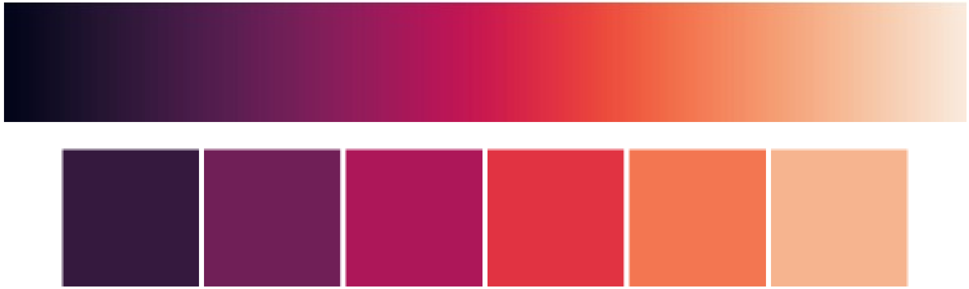
Color Brewer tool:

https://seaborn.pydata.org/tutorial/color_palettes.html

Quantitative sequential palettes
(for numerical variables)

Vary lightness rather than (just) hue (hue is circular, hard to interpret)

Discrete vs continuous:



https://seaborn.pydata.org/tutorial/color_palettes.html

Quantitative diverging palettes

For numerical variables with a special midpoint, often 0 (such as increase / decrease)

Midpoint displayed as a special neutral color, e.g. grey

https://seaborn.pydata.org/tutorial/color_palettes.html

Summary

Visual perception:

- light on the retina, transduction on photoreceptor cells
- feature detection, pattern formation, interpretation

Light and color:

- Visible light as a part of electromagnetic spectrum
- Three types of cones sensitive to different wavelengths of light

Color spaces (LMS, RGB, HSL, HSV, CMYK, RYB)

Colors in visualization (qualitative and quantitative palettes, color blindness, sparing use of colors)

Graphics file formats (for exporting graphs)

Raster formats: store your plot as pixels at some resolution
- resolution leads to tradeoff between size and quality
- prefer lossless compression (PNG rather than JPEG)
- transparent background may be also a good idea

Vector formats: store your plot as geometric objects (SVG, PDF, EPS)
- can be arbitrarily enlarged, editable
- can be large if many points / lines (subsample data?)
- vector formats usually preferable unless not supported by software
- it is a good idea to include fonts in the file
- beware that these file formats may include bitmaps

Data analysis project phases

Data analysis / visualization project phases

- Obtaining data
- Data preprocessing, **checking**, cleaning
- **Exploratory** analysis (many tables and graphs for your use)
- Formation of **hypotheses**
- **Testing** hypotheses (careful reanalysis, new data, other sources)
- **Explanatory** visualizations for the final report / presentation (best views selected for the audience)

Details: obtaining data

- Obtaining data
  - This course: we download whole datasets in a tabular form
  - But often also web scraping, manual collection of data, measurements, surveys,...
  - Requires careful planning
- Data preprocessing, **checking**, cleaning
- **Exploratory** analysis
- Formation of **hypotheses**
- Testing hypotheses
- **Explanatory** visualizations for the final report / presentation

Details: preprocessing data

- Obtaining data
- Data preprocessing, **checking**, cleaning
    - Try to understand how (and why) the data was obtained and processed
    - Convert them to a convenient format
    - Check for missing values and suspicious outliers
    - Very important phrase: "Garbage in, garbage out"

- **Exploratory** analysis
- Formation of **hypotheses**
- Testing hypotheses
- **Explanatory** visualizations for the final report / presentation

Details: exploratory analysis

- Obtaining data
- Data preprocessing, **checking**, cleaning
- **Exploratory** analysis
  - Try many analyses
  - This course: visualizations and simple statistics
  - Later you learn advanced statistical and machine learning models
  - Less successful attempts may suggest new directions

- Formation of **hypotheses**
- Testing hypotheses
- **Explanatory** visualizations for the final report / presentation

Details: Formation of hypotheses

- Obtaining data
- Data preprocessing, **checking**, cleaning
- **Exploratory** analysis
- Formation of **hypotheses**
    - Select visualizations showing interesting trends / exceptions in the data
    - Formulate possible relationships
      (but remember, correlation does not imply causation)

- Testing hypotheses
- **Explanatory** visualizations for the final report / presentation

Details: Testing hypotheses

- Obtaining data
- Data preprocessing, **checking**, cleaning
- **Exploratory** analysis
- Formation of **hypotheses**
- Testing hypotheses
  - Recheck your code and data, try other related analyses
  - Try to find other relevant data or existing analyses by other people
  - If important decisions will be based on your result,
    test it particularly thoroughly
    (what would happen if our plot was all wrong?)

- **Explanatory** visualizations for the final report / presentation

Details: Explanatory visualizations

- Obtaining data
- Data preprocessing, **checking**, cleaning
- **Exploratory** analysis
- Formation of **hypotheses**
- Testing hypotheses
- **Explanatory** visualizations for the final report / presentation
  - Formulate your conclusions
  - Support them with your analysis and visualizations
  - Do not include all exploratory analyses
    (but do not hide data contradicting your conclusion)
  - Polish visualizations that you selected

Data analysis / visualization project phases

- Obtaining data
- Data preprocessing, **checking**, cleaning
- **Exploratory** analysis (many tables and graphs for your use)
- Formation of **hypotheses**
- **Testing** hypotheses (careful reanalysis, new data, other sources)
- **Explanatory** visualizations for the final report / presentation (best views selected for the audience)