NYU DATA SCIENCE

Data Visualization using matplotlib

Bruno Gonçalves

www.bgoncalves.com

https://bmtgoncalves.github.io/DataVisualization/

NYU DATA SCIENCE

Data V using matplotlib

Bruno G

JPMorgan Chase & Co.

www.bg

https://bm .io/DataVisualization/

JPMorgan Chase & Co.

Data Visualization using matplotlib

Bruno Gonçalves

www.bgoncalves.com

https://bmtgoncalves.github.io/DataVisualization/

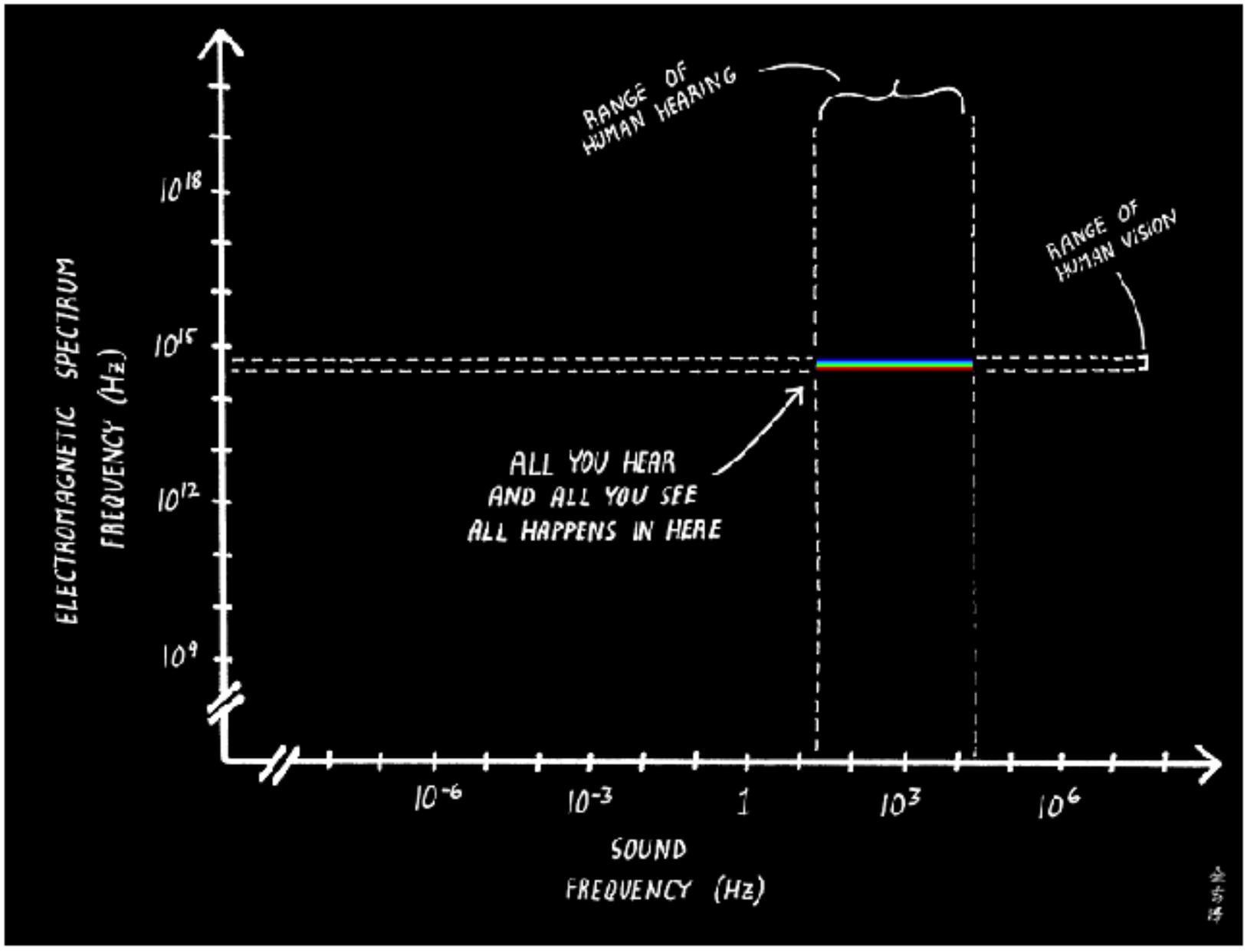
Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of my employer. The examples provided with this tutorial were chosen for their didactic value and are not mean to be representative of my day to day work.

Human Perception

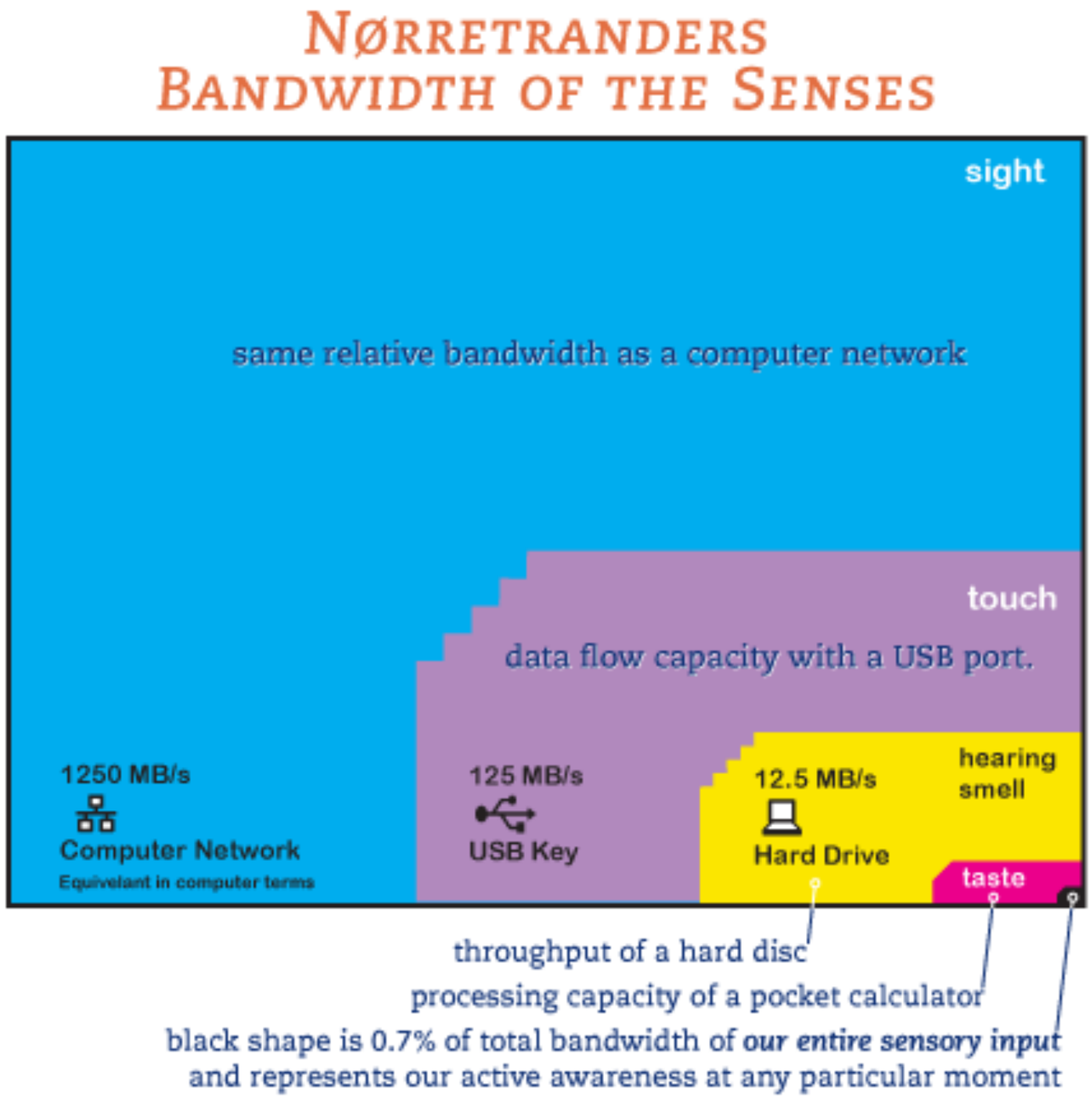
Human Perception

http://abstrusegoose.com/421



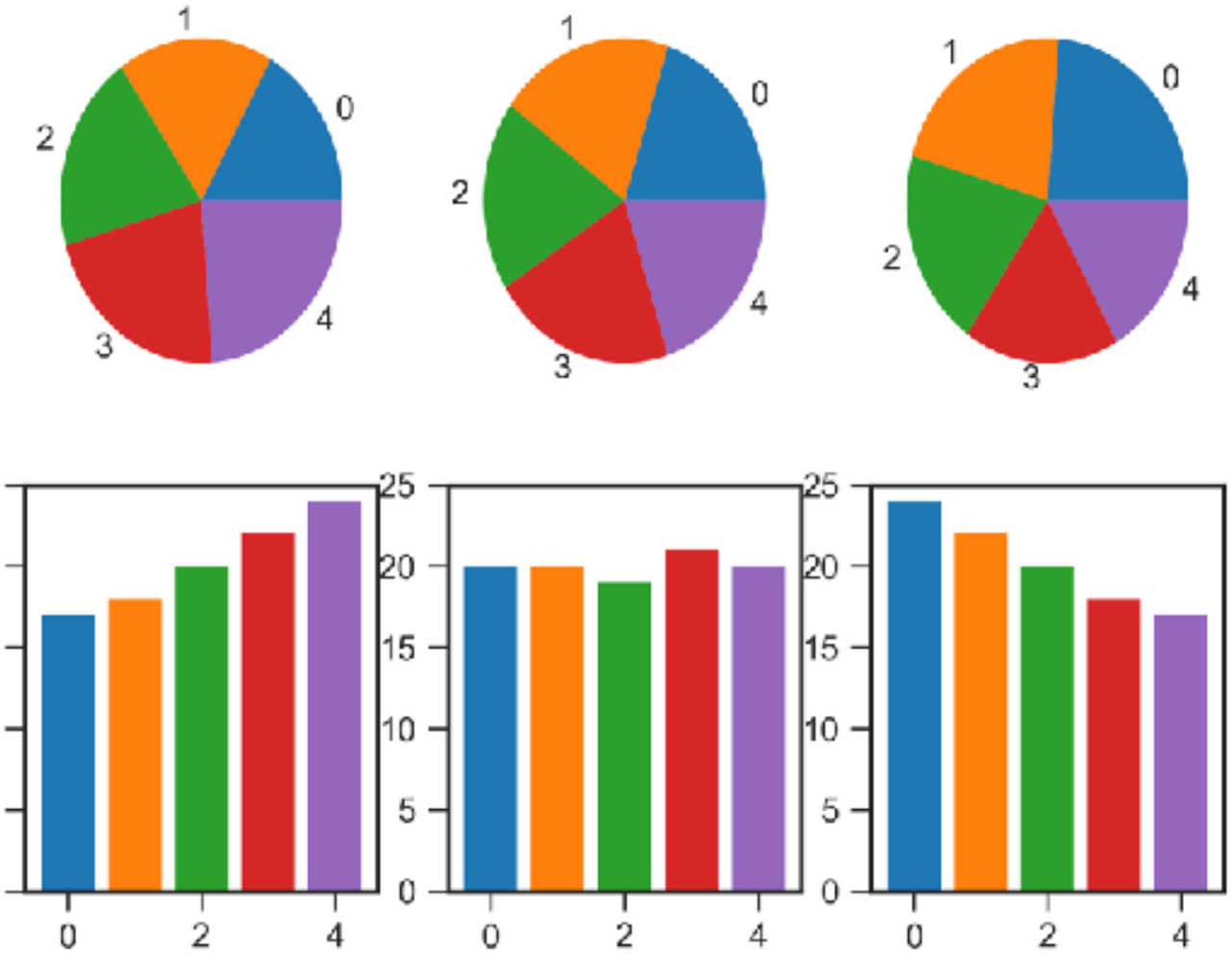
In the grand scheme of things,
we're all pretty much blind and deaf.

Human Senses



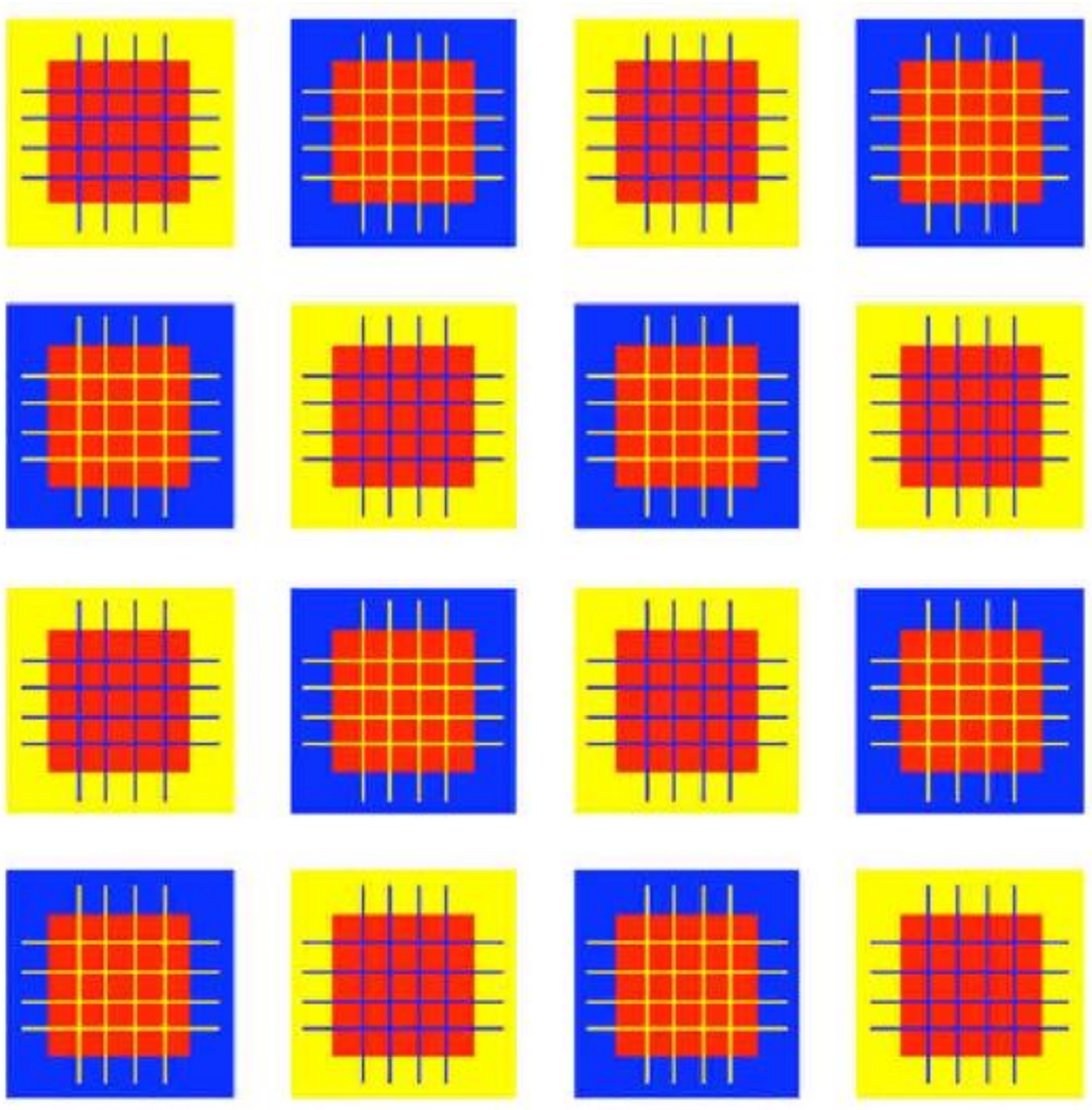
Perception

- Some cognitive tasks are significantly easier than others. In order, we are good a distinguishing:
  - Position, length
  - Direction, Angle, Area
  - Volume, Curvature, Shade



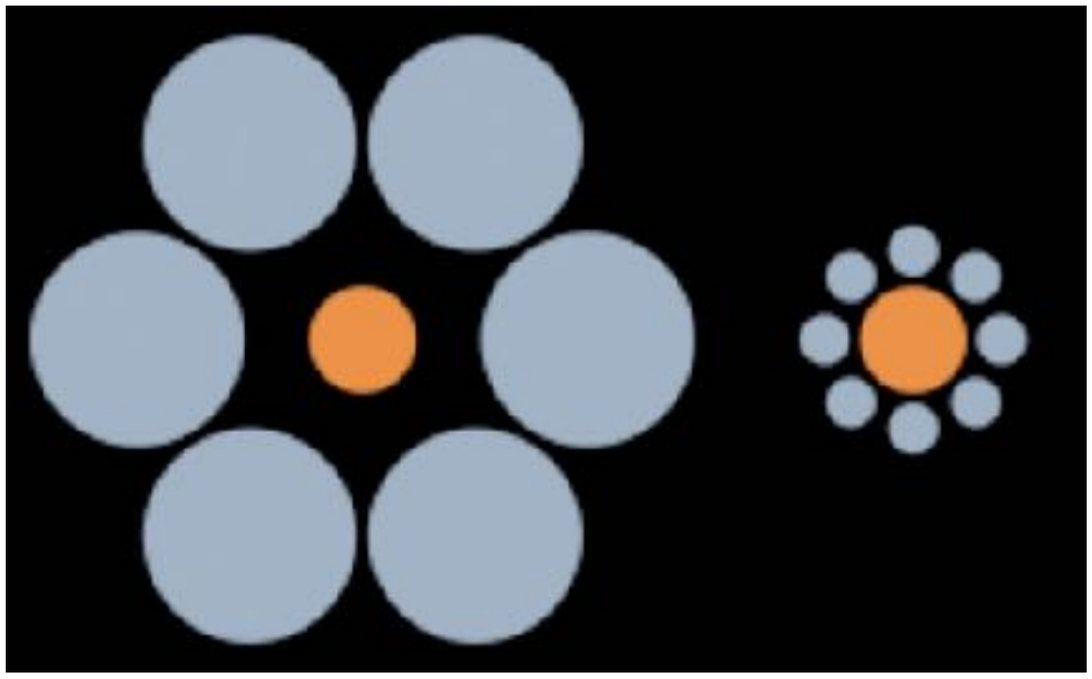
Perception

- Some cognitive tasks are significantly easier than others. In order, we are good a distinguishing:
  - Position, length
  - Direction, Angle, Area
  - Volume, Curvature, Shade
  - Color Saturation.



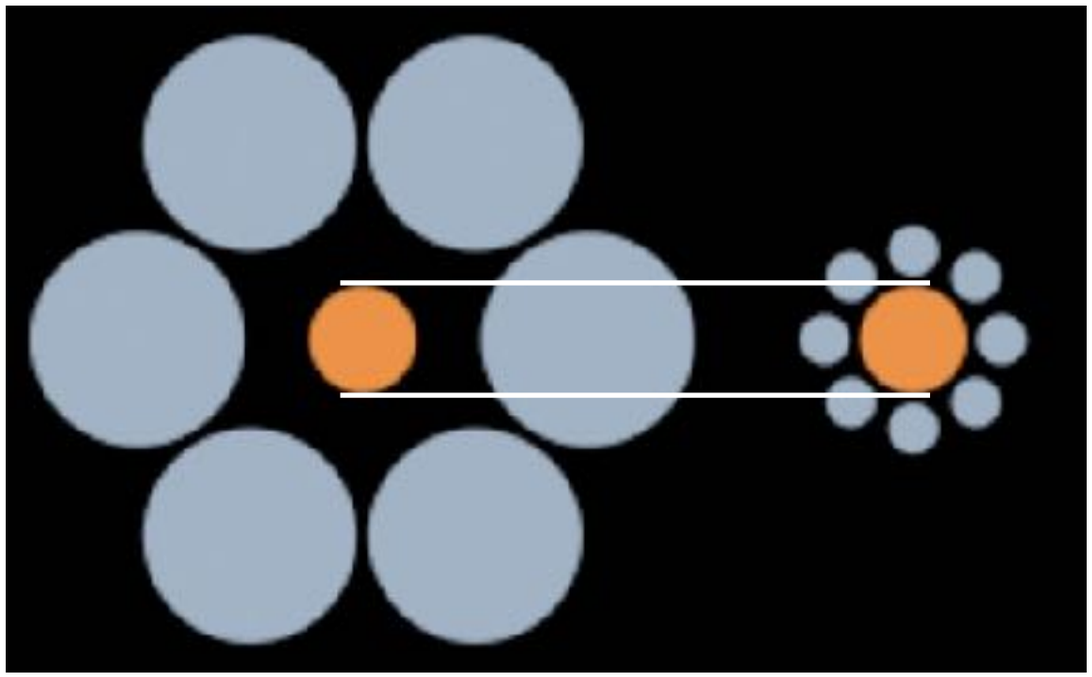
Perception

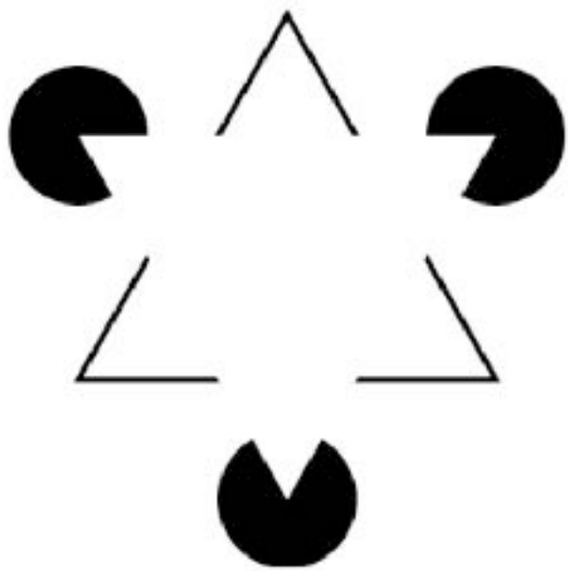
- Some cognitive tasks are significantly easier than others. In order, we are good a distinguishing:
    - Position, length
    - Direction, Angle, Area
    - Volume, Curvature, Shade
    - Color Saturation.
- Context also matters!
    - An object seen in the context of larger objects will appear smaller, while in the content of smaller objects it will appear larger.

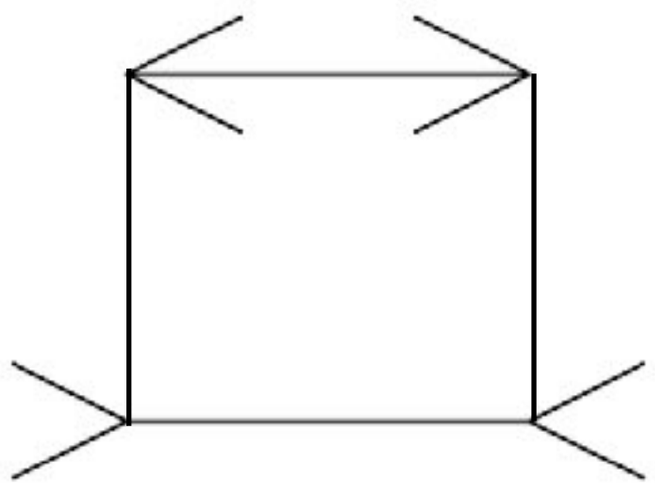


Perception

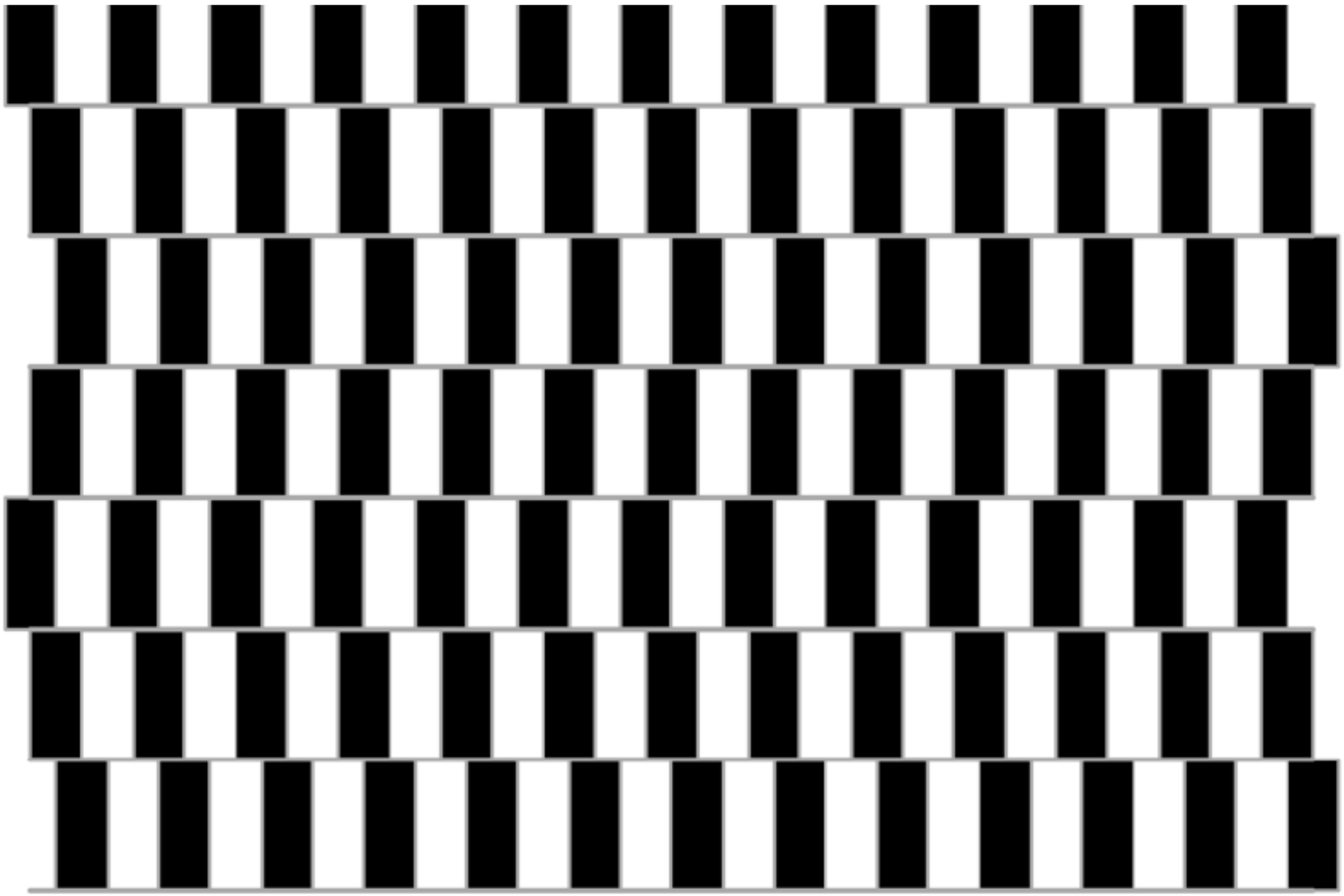
- Some cognitive tasks are significantly easier than others. In order, we are good a distinguishing:
  - Position, length
  - Direction, Angle, Area
  - Volume, Curvature, Shade
  - Color Saturation.
- Context also matters!
  - An object seen in the context of larger objects will appear smaller, while in the content of smaller objects it will appear larger.
  - And we “fill in the gaps”





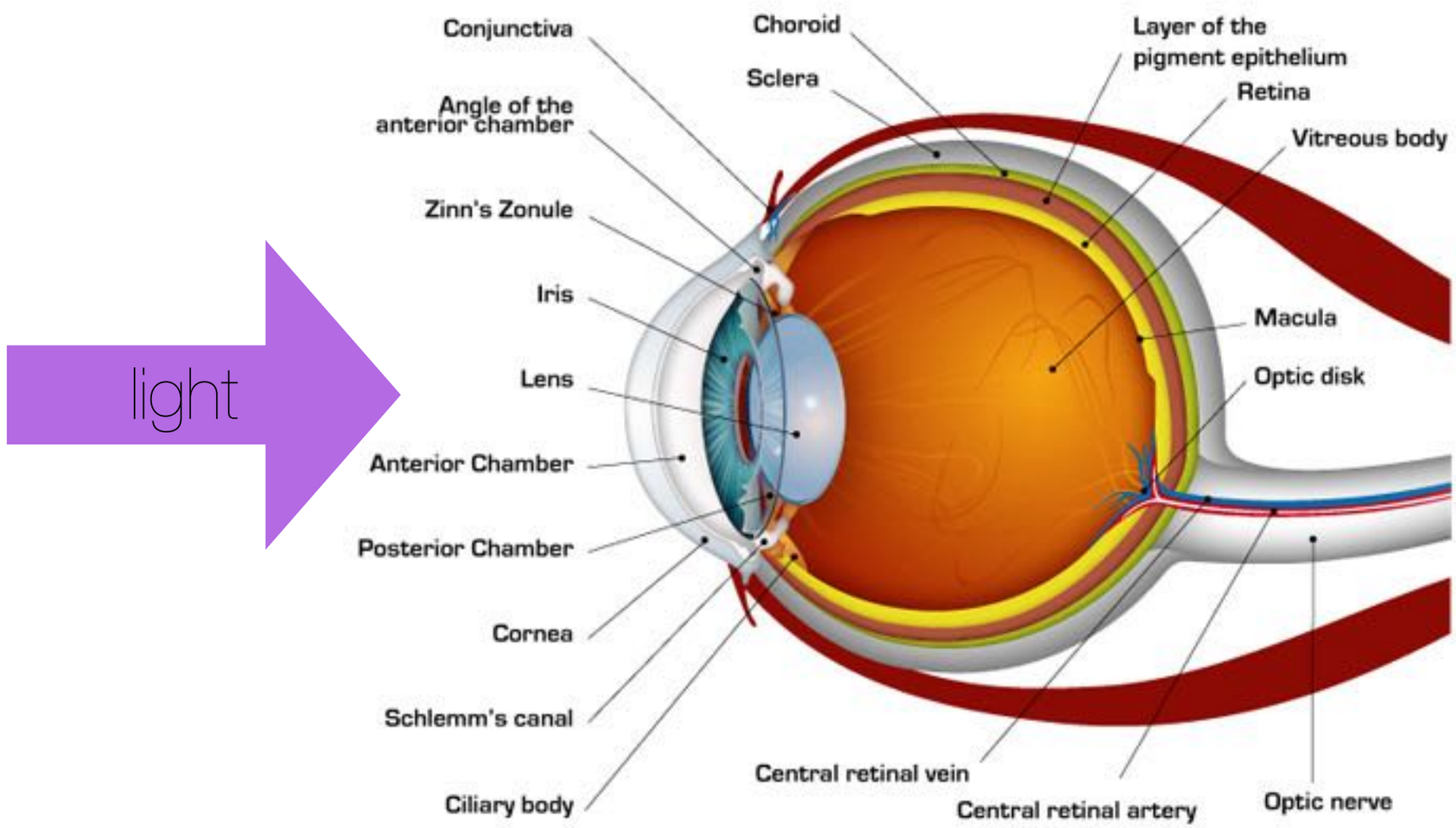


Perception Biases



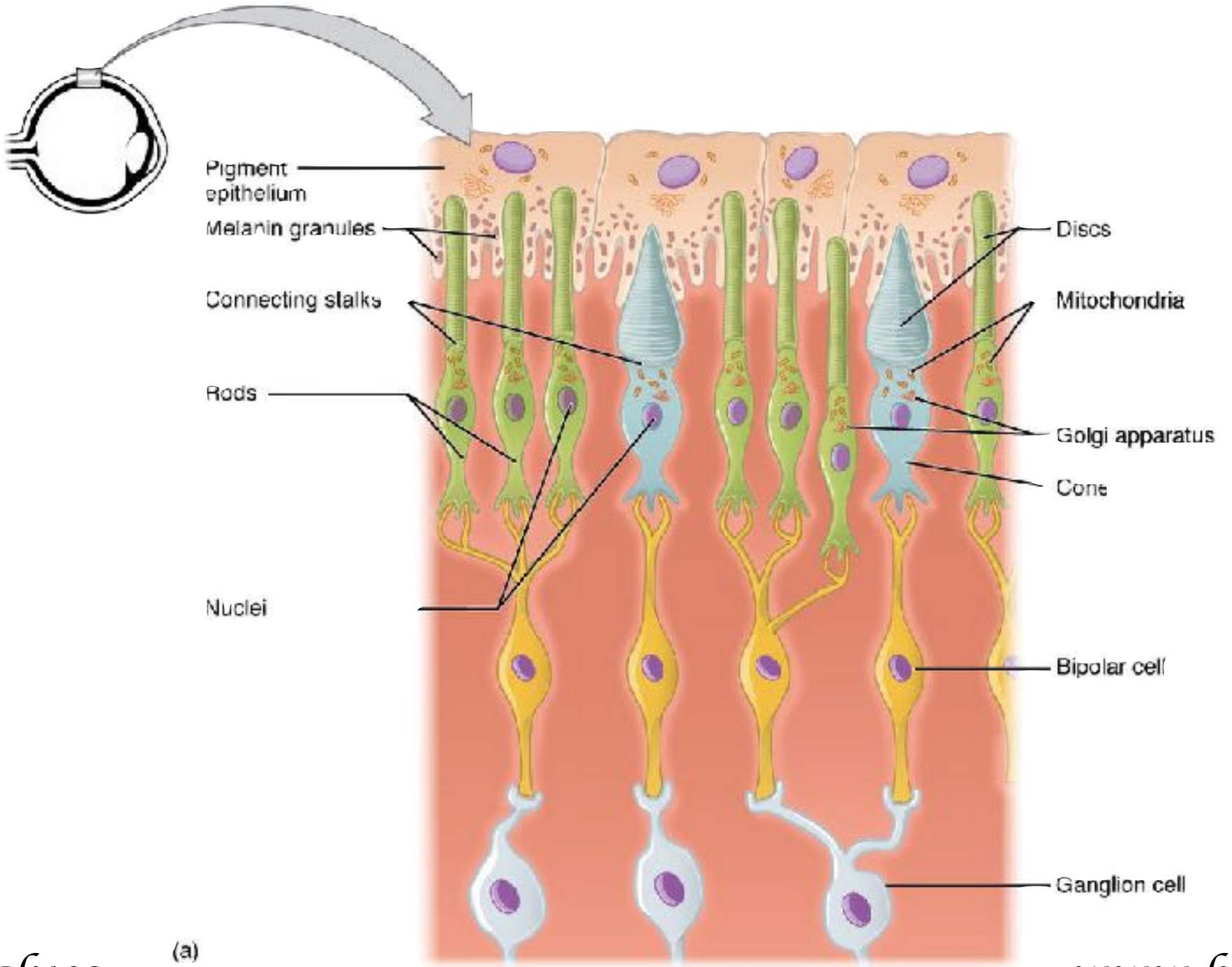
Human Vision

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



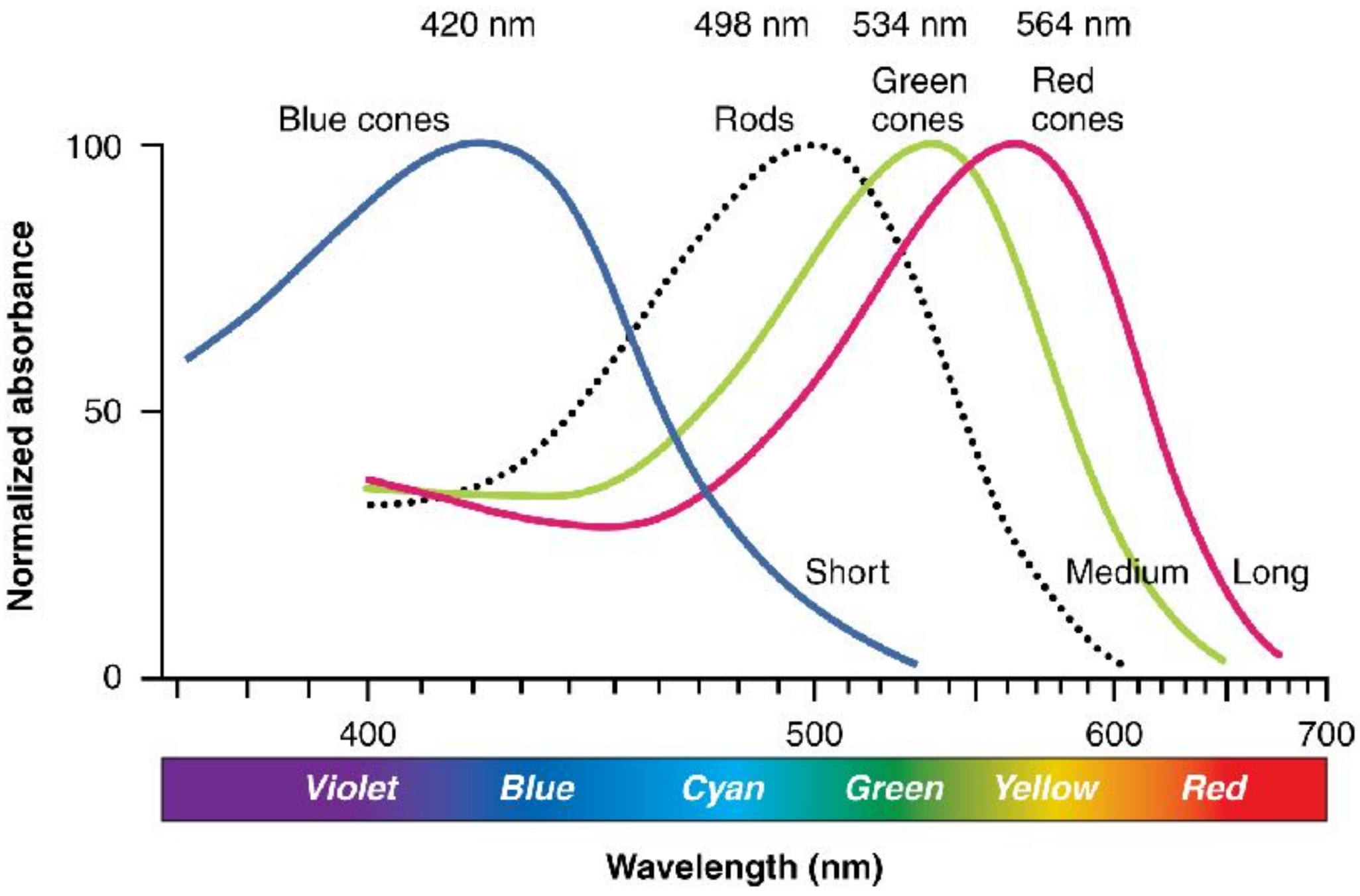
Human Vision

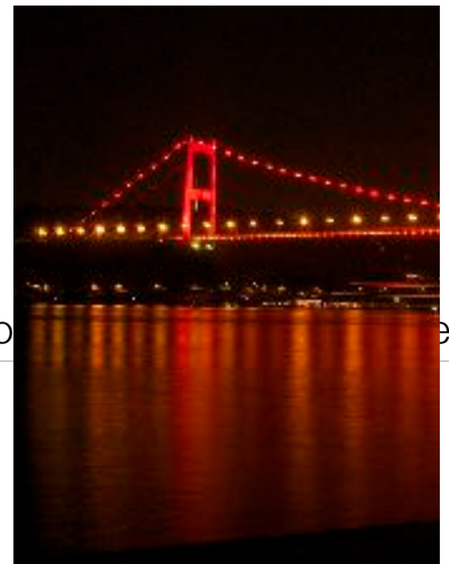
https://en.wikipedia.org/wiki/Photoreceptor_cell



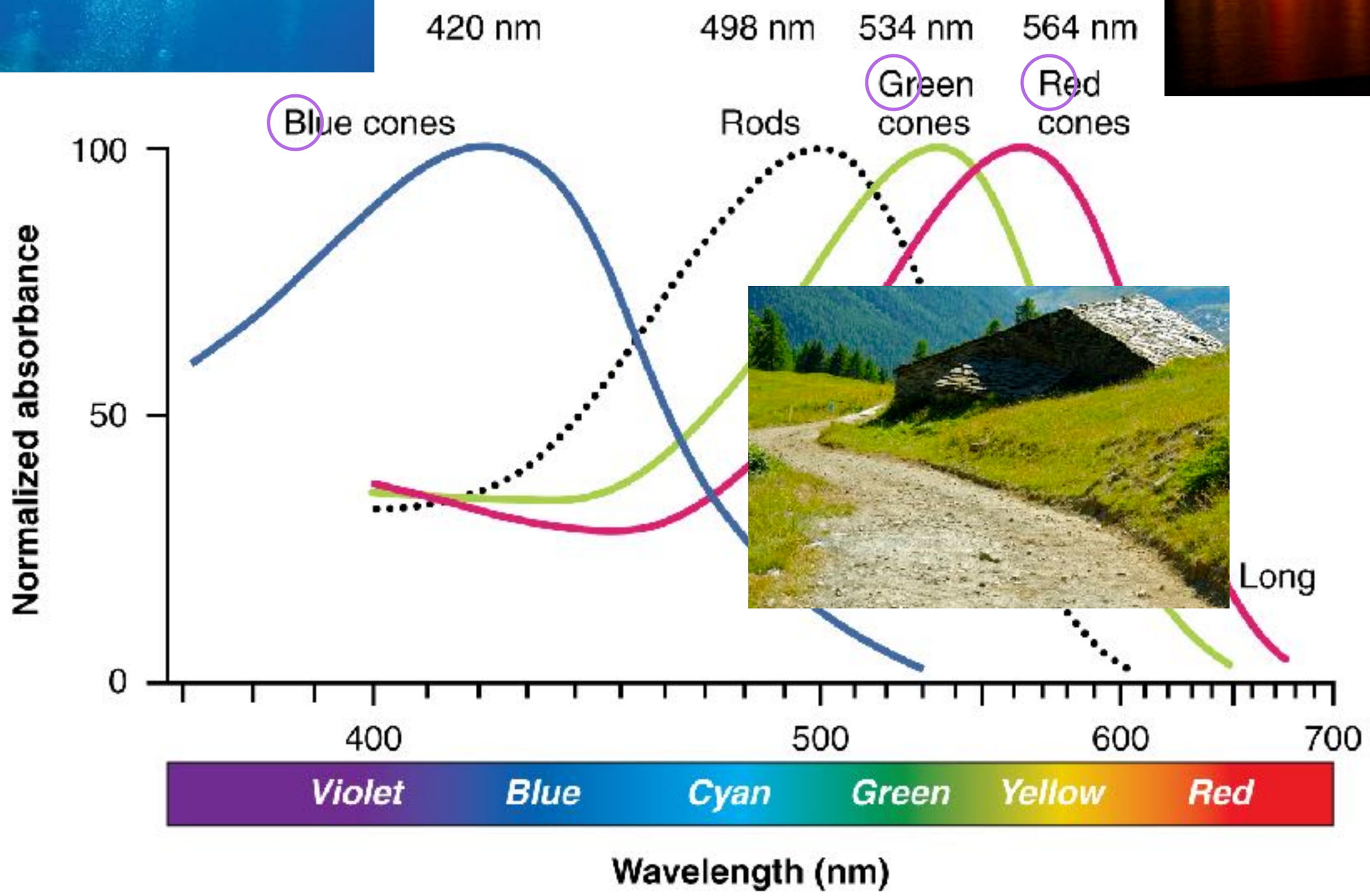
Human Vision

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

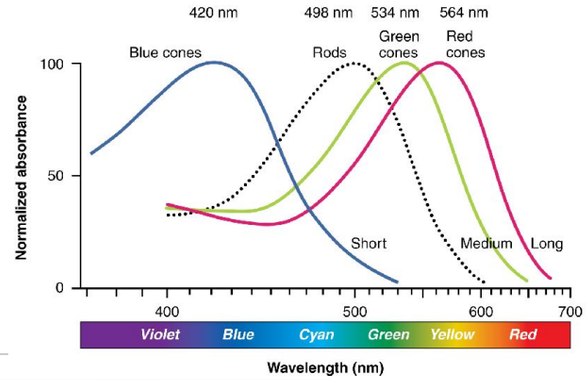




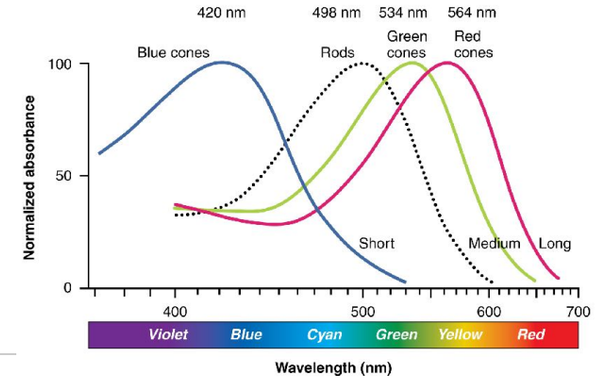
https://en.wikipedia.o ell

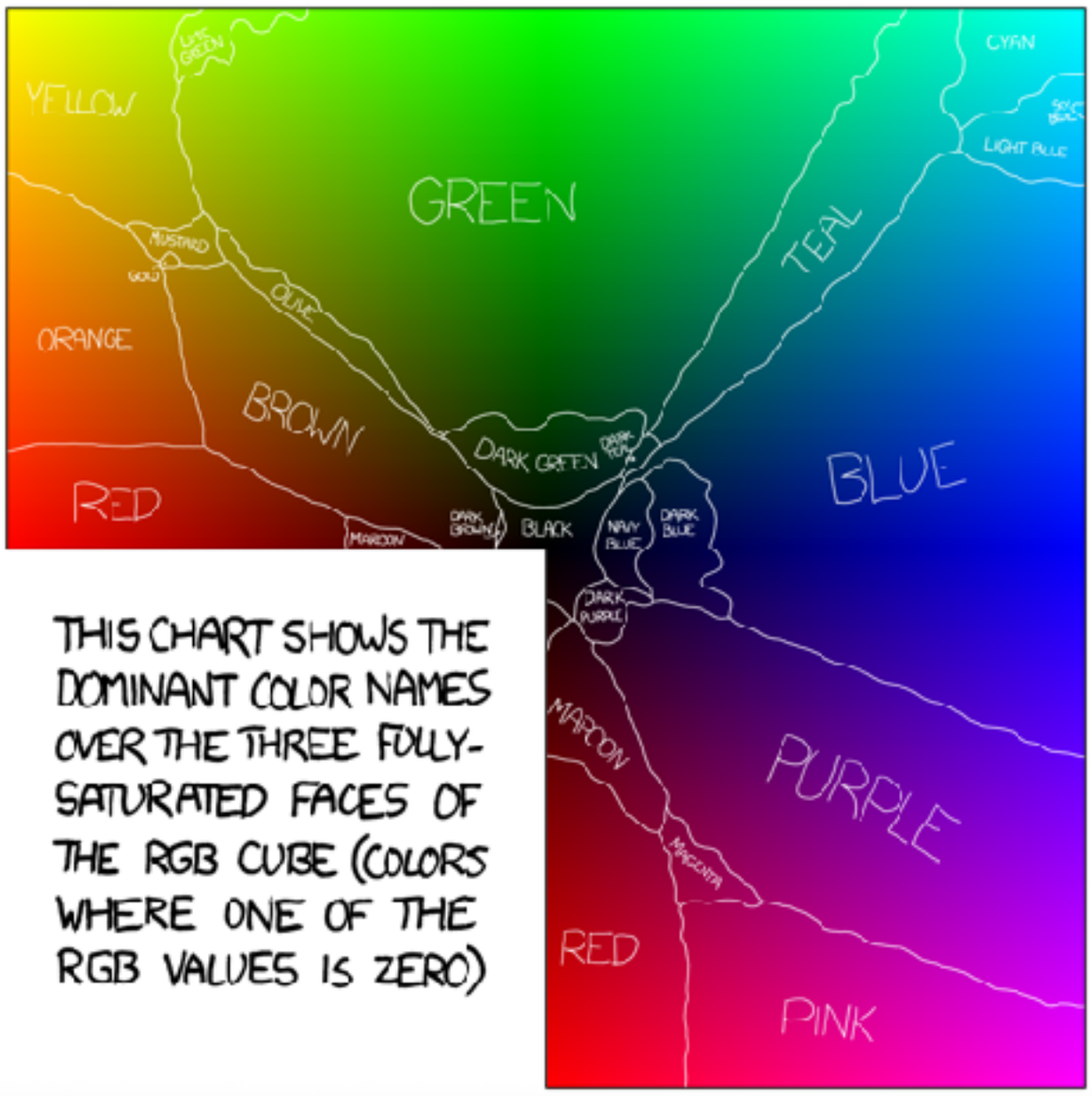


Colors galore!

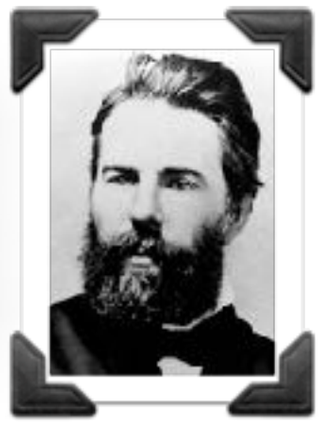


Color Perception



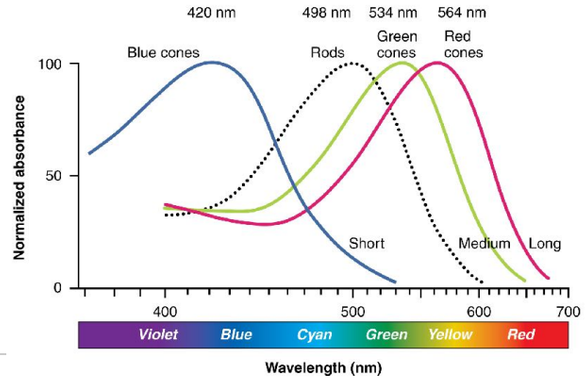


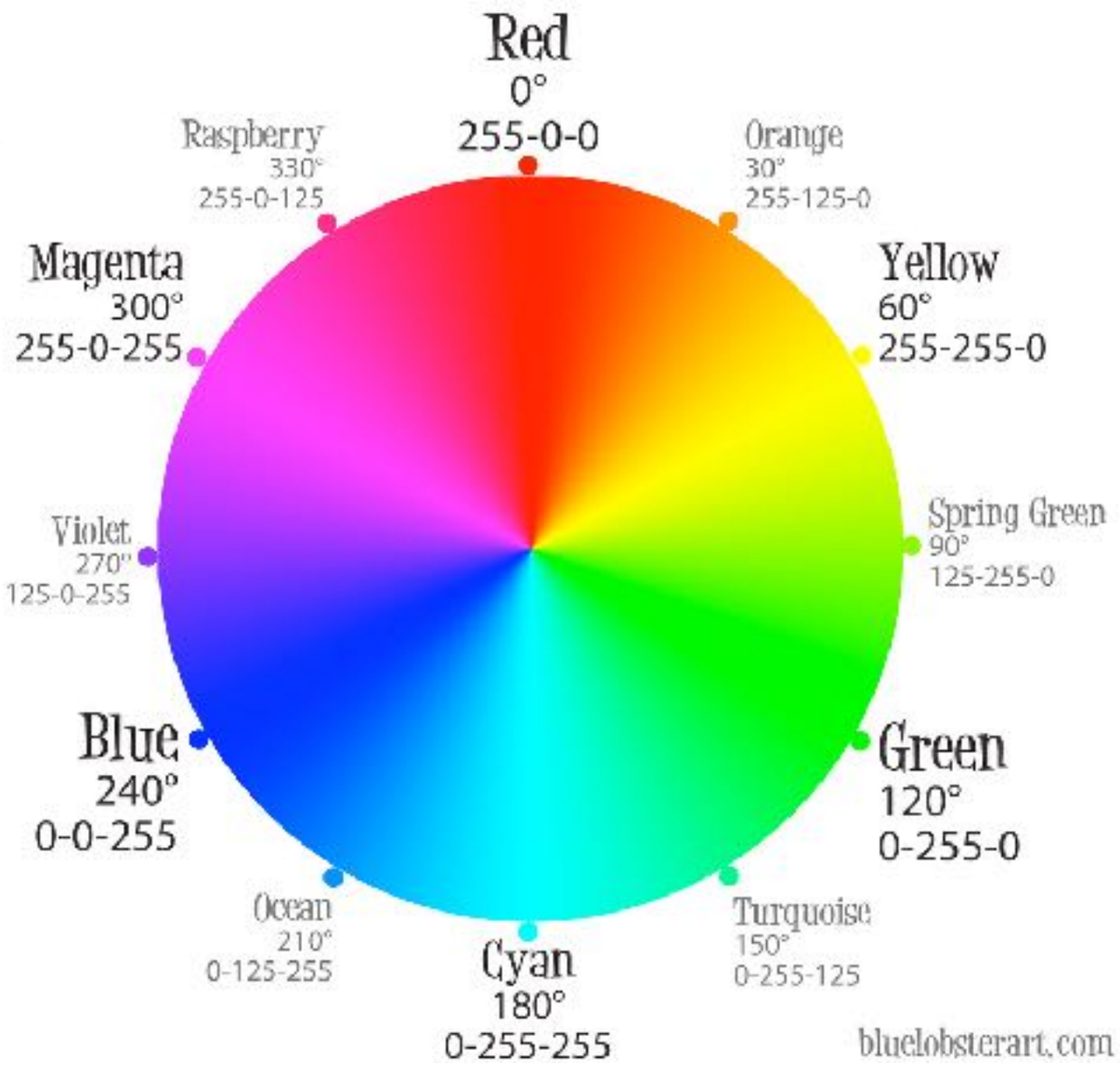
"Who in the rainbow can draw the line where the violet tint ends and the orange tint begins? Distinctly we see the difference of the colors, but where exactly does the one first blendingly enter into the other? So with sanity and insanity."
(H. Melville)

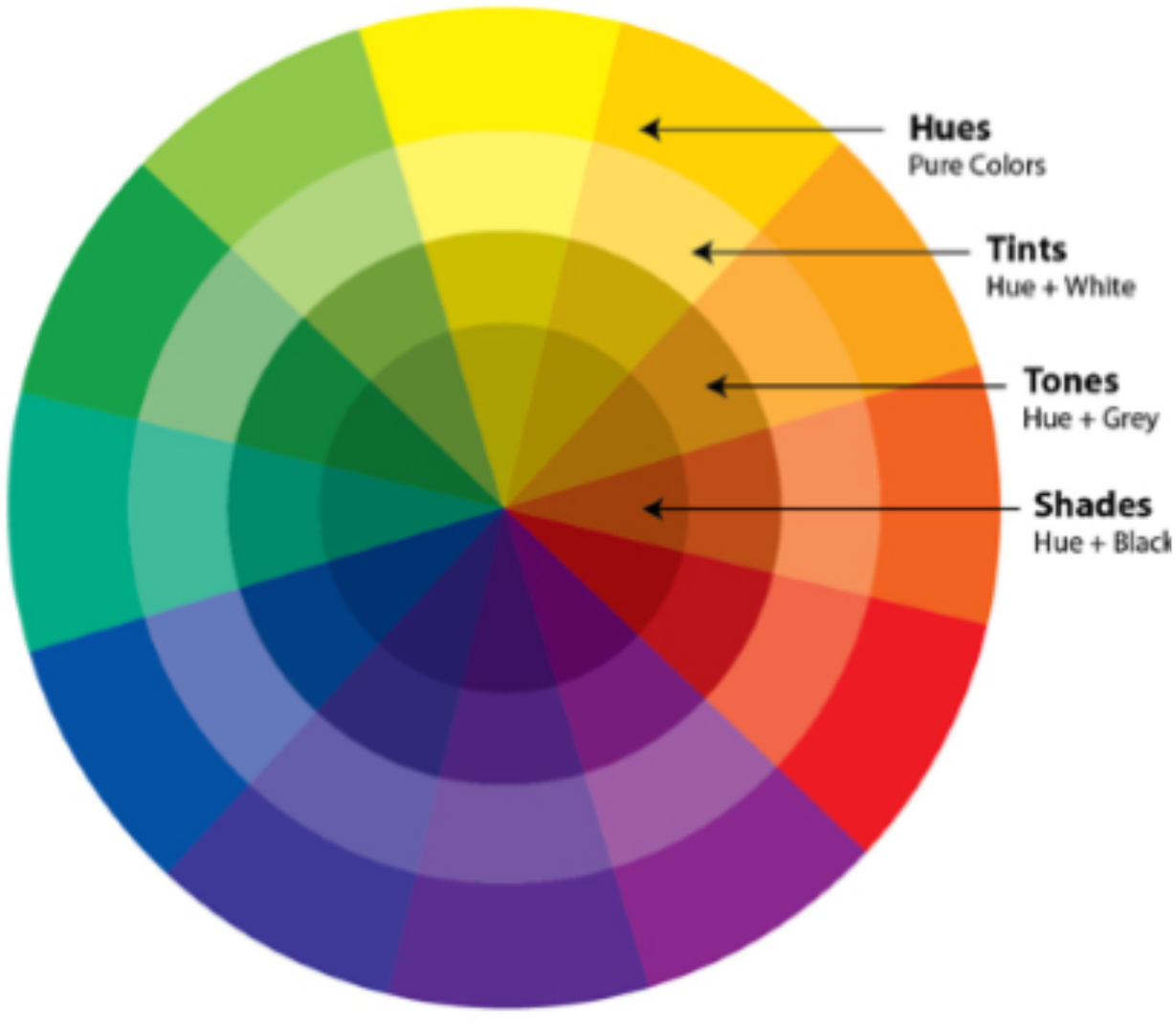


@bgoncalves

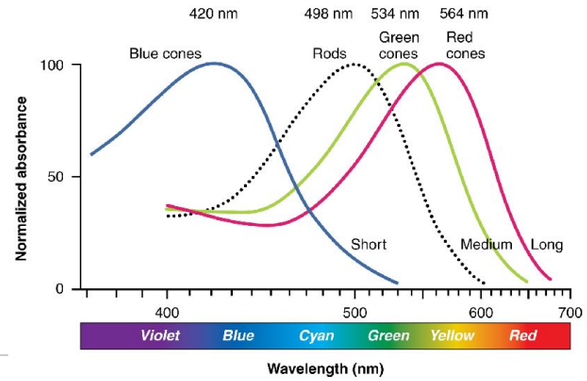
Color Wheel


420 nm
498 nm
534 nm
564 nm
Blue cones
Rods
Green cones
Red cones
100
50
0
Normalized absorbance
Short
Medium
Long
400
500
600
700
Violet
Blue
Cyan
Green
Yellow
Red
Wavelength (nm)


Red
0°
255-0-0
Raspberry
330°
255-0-125
Orange
30°
255-125-0
Magenta
300°
255-0-255
Yellow
60°
255-255-0
Violet
270°
125-0-255
Spring Green
90°
125-255-0
Blue
240°
0-0-255
Green
120°
0-255-0
Ocean
210°
0-125-255
Cyan
180°
0-255-255
Turquoise
150°
0-255-125
bluelobsterart.com


Hues
Pure Colors
Tints
Hue + White
Tones
Hue + Grey
Shades
Hue + Black

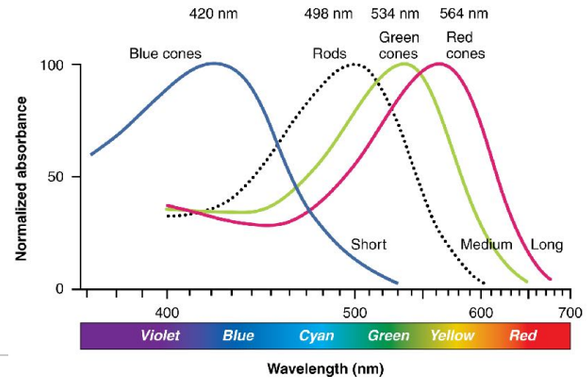
Color Schemes



Warm Colors

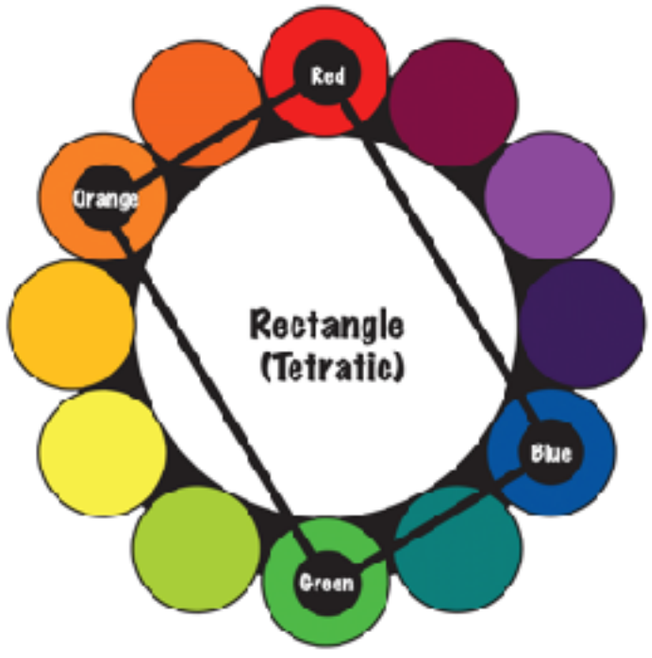
Cold Colors

Color Schemes



Complementary color scheme
Colors that are opposite each other on the color wheel are considered to be complementary colors

(example: Orange and Blue).



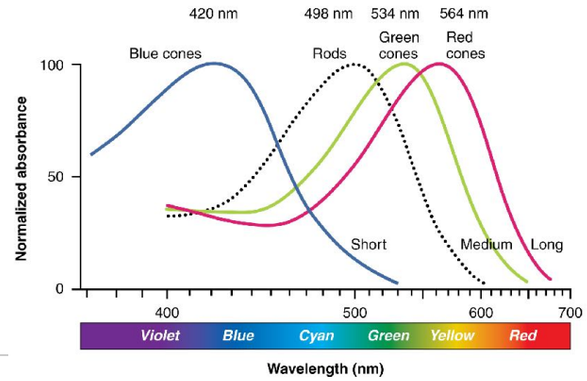
Rectangle (tetradic) color scheme
The rectangle or tetradic color scheme uses four colors arranged into two complementary pairs.

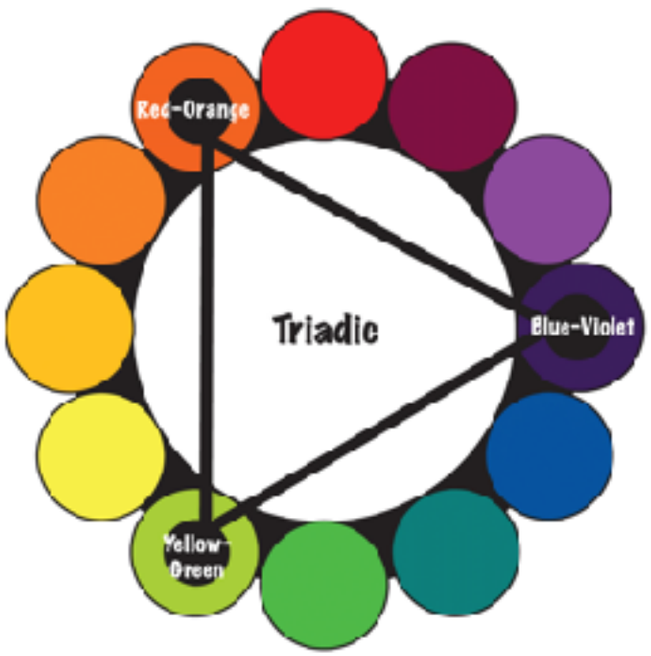
(example: Orange, Red, Blue and Green)

Analogous color scheme
Analogous color schemes use colors that are next to each other on the color wheel.

(example: Green, Blue-Green and Blue)

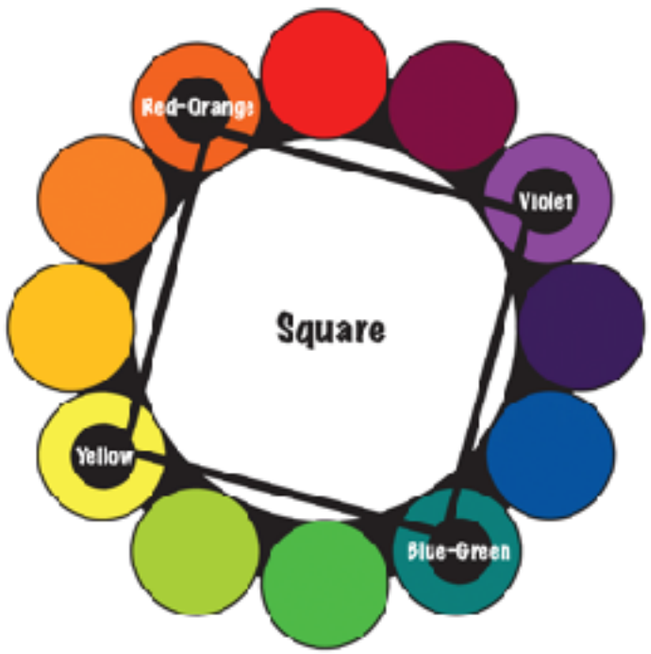
Color Schemes





Triadic color scheme
A triadic color scheme uses colors that are evenly spaced around the color wheel.

(example: Yellow-Green, Red-Orange and Blue-Violet)



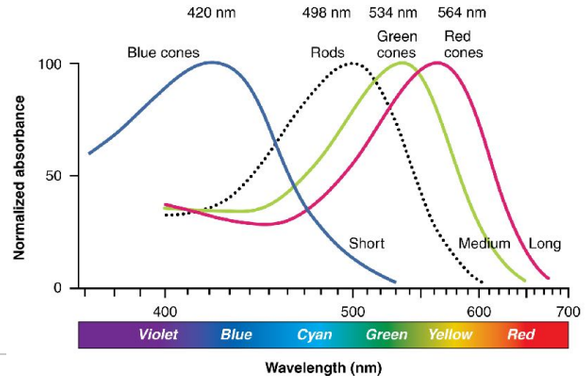
Square color scheme
The square color scheme is similar to the rectangle, but with all four colors spaced evenly around the color circle.

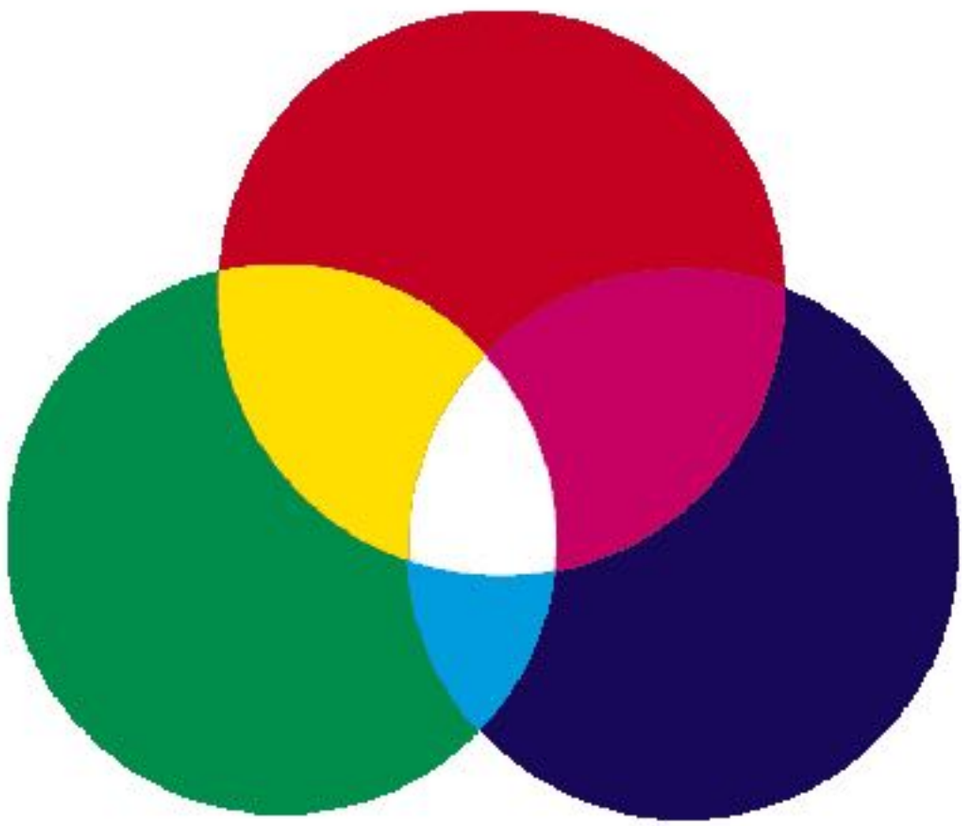
(example: Yellow, Red-Orange, Violet and Blue-Green)

Split-Complementary color scheme
The split complementary color scheme is a variation of the complementary color scheme. In addition to the base color, it uses the two colors adjacent to its complement.

(example: Yellow, Red-Violet and Blue-Violet)

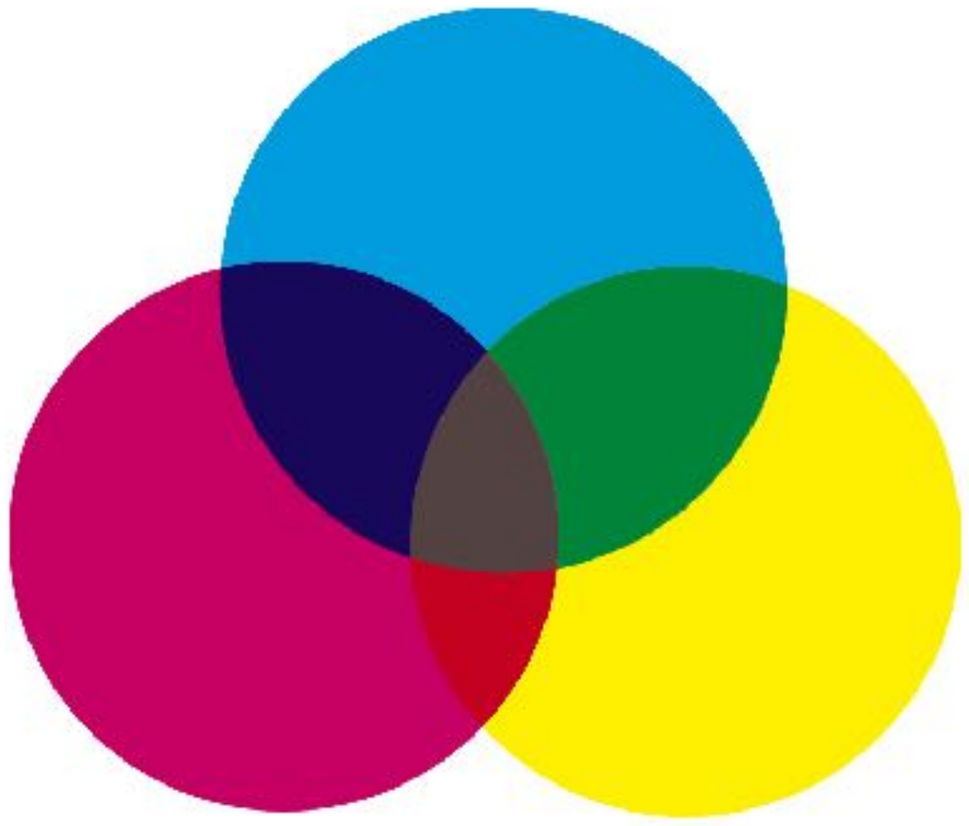
Color Systems





Additive Color (RGB)

Light

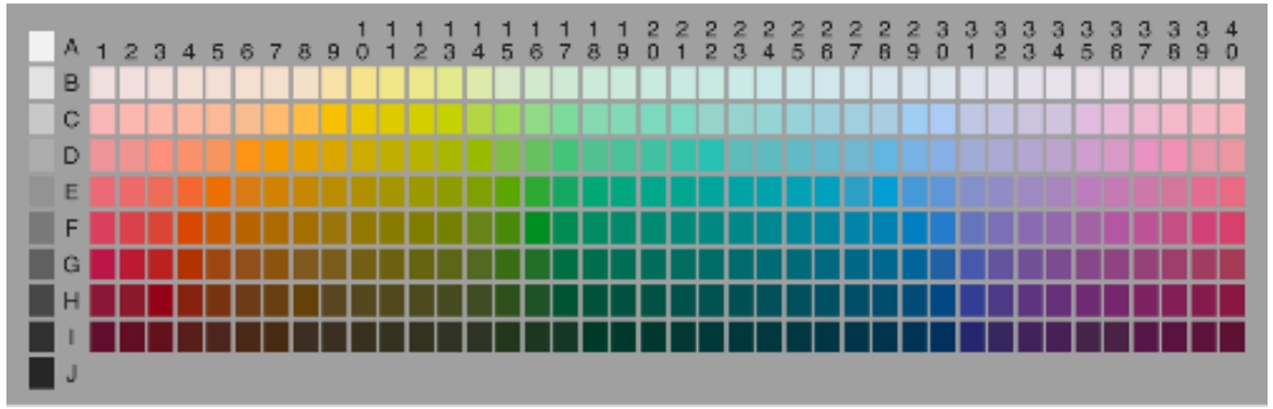


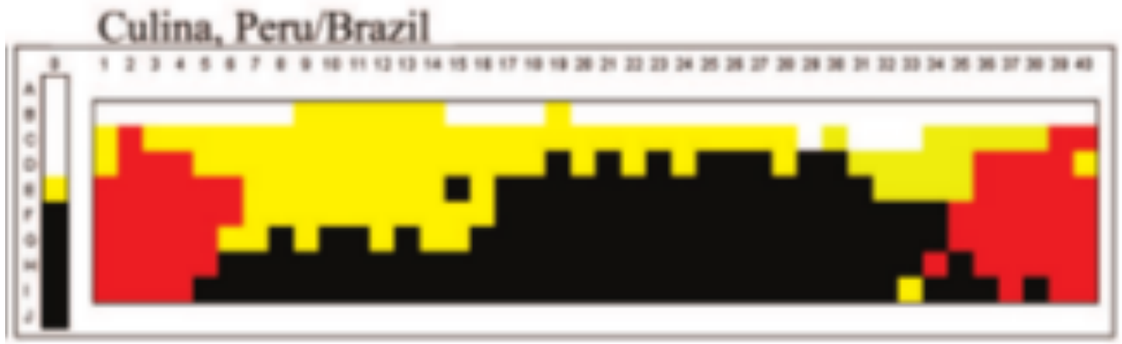
Subtractive color (CMYK)

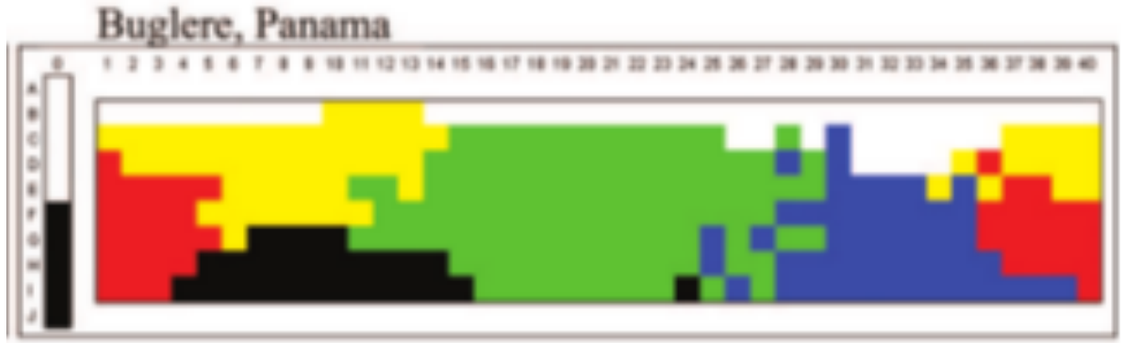
Ink

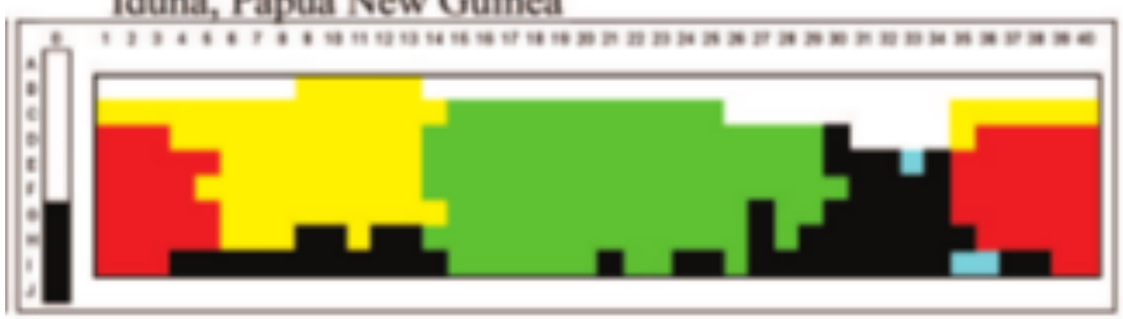
Colors and Culture

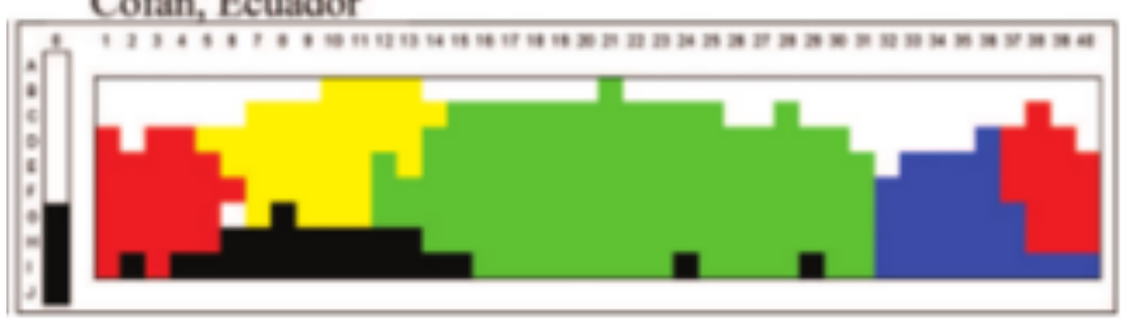
http://www1.icsi.berkeley.edu/wcs/

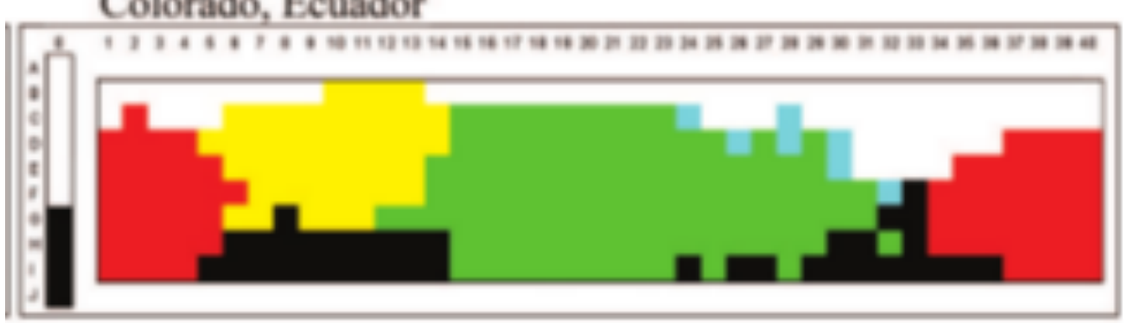










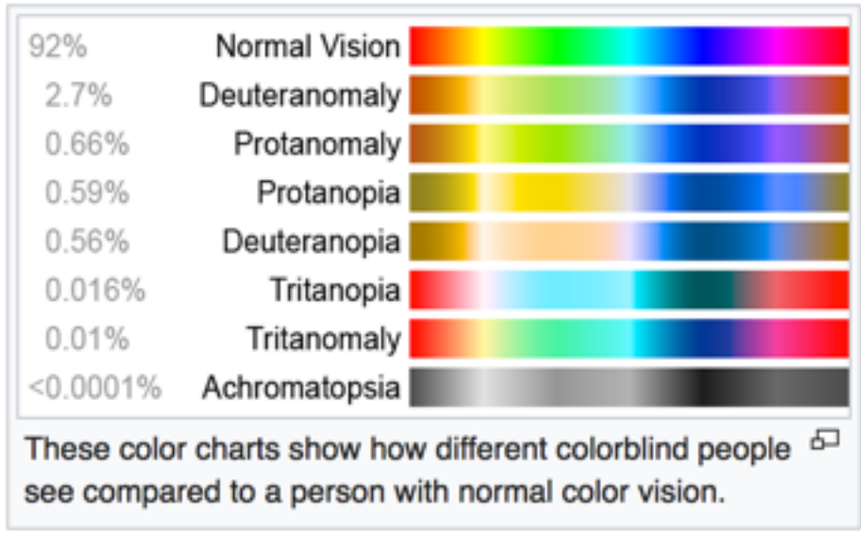


Color Blindness

Color Blindness

https://en.wikipedia.org/wiki/Color_blindness
https://github.com/MaPePeR/jsColorblindSimulator/blob/master/colorblind.js

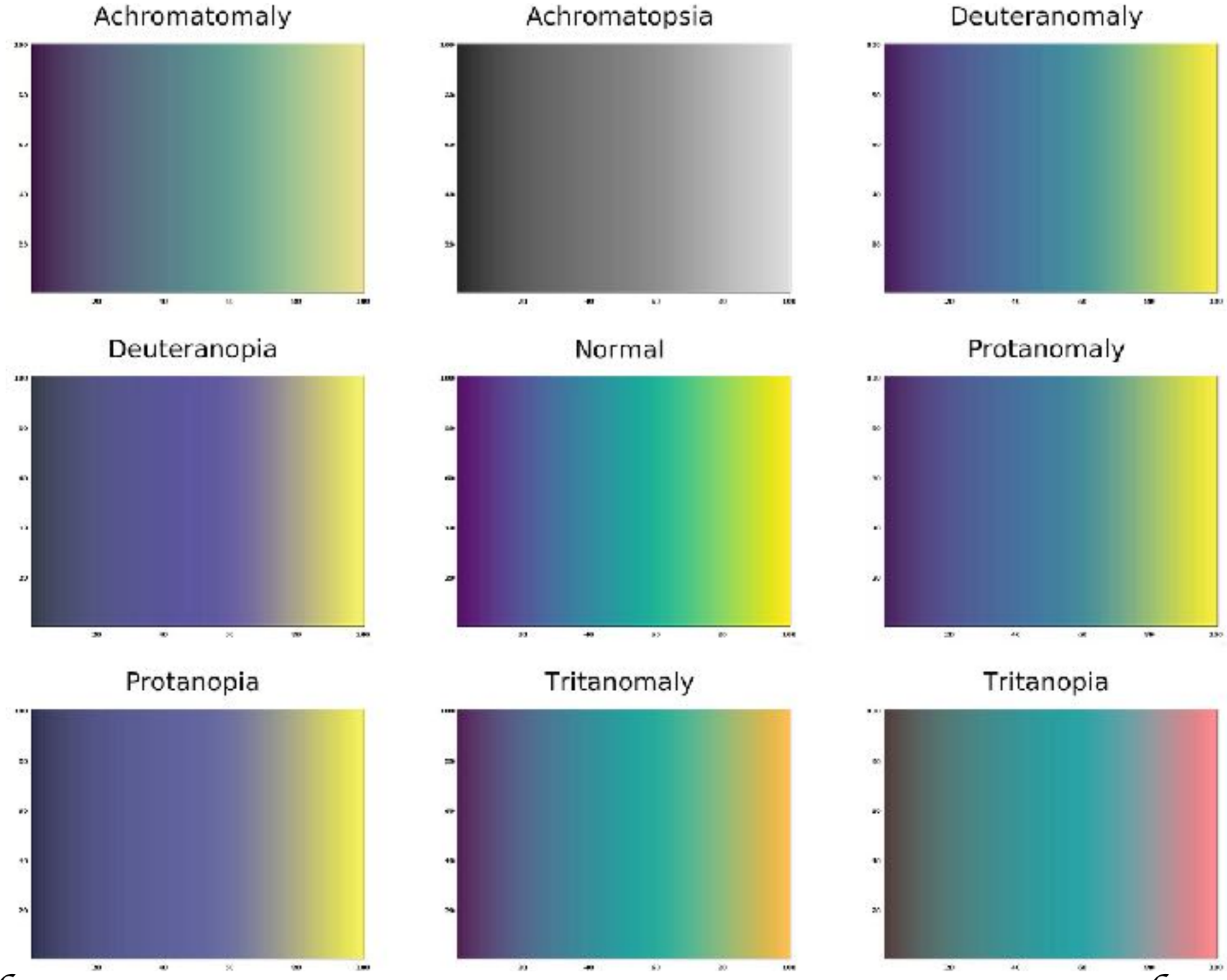
	Cone system	Red		Green		Blue	
	N=normal **A**=anomalous	**N**	**A**	**N**	**A**	**N**	**A**
1	Normal vision						
2	Protanomaly						
3	Protanopia						
4	Deuteranomaly						
5	Deuteranopia						
6	Tritanomaly						
7	Tritanopia						
8	Achromatopsia						
9	Tetrachromat						
10							



These color charts show how different colorblind people see compared to a person with normal color vision.

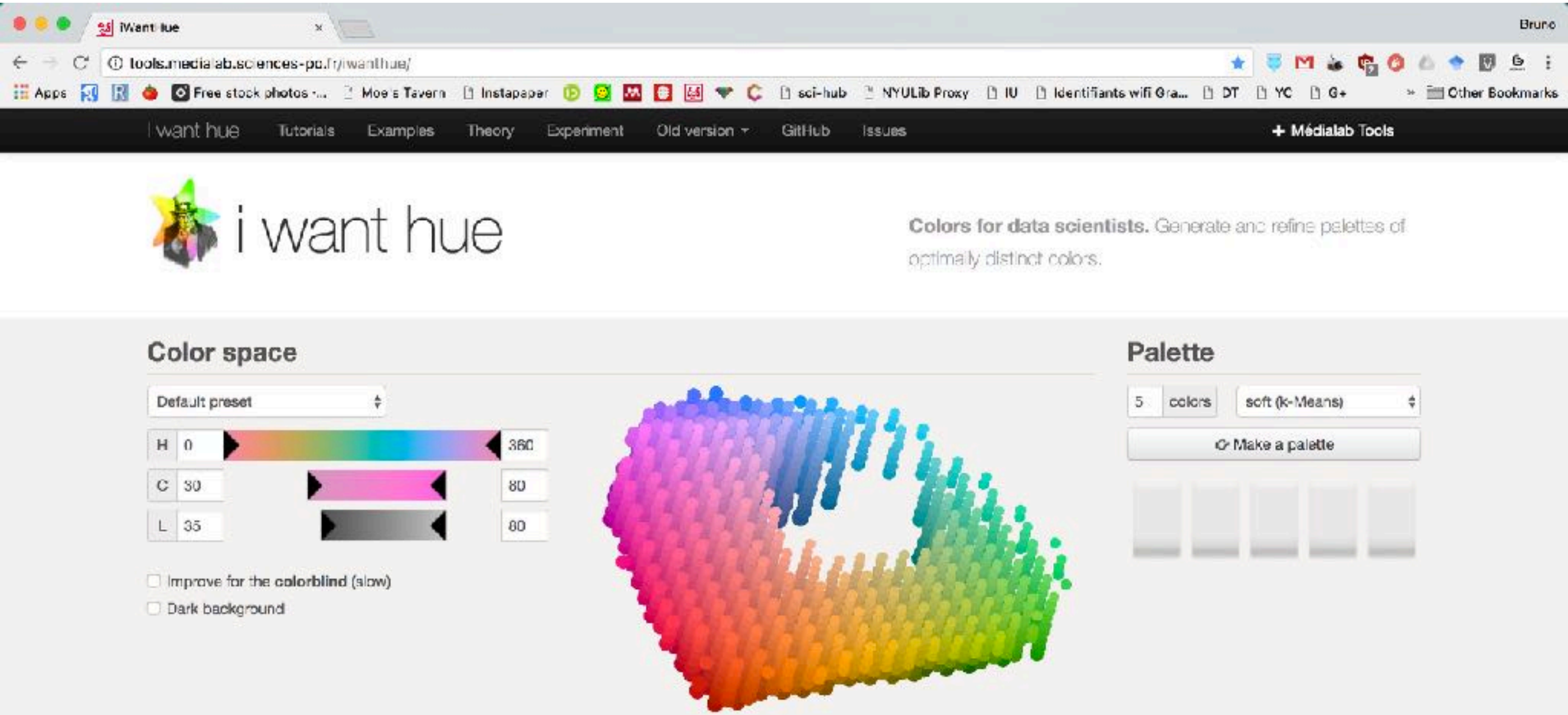
@bgoncalves

Viridis Color Scheme



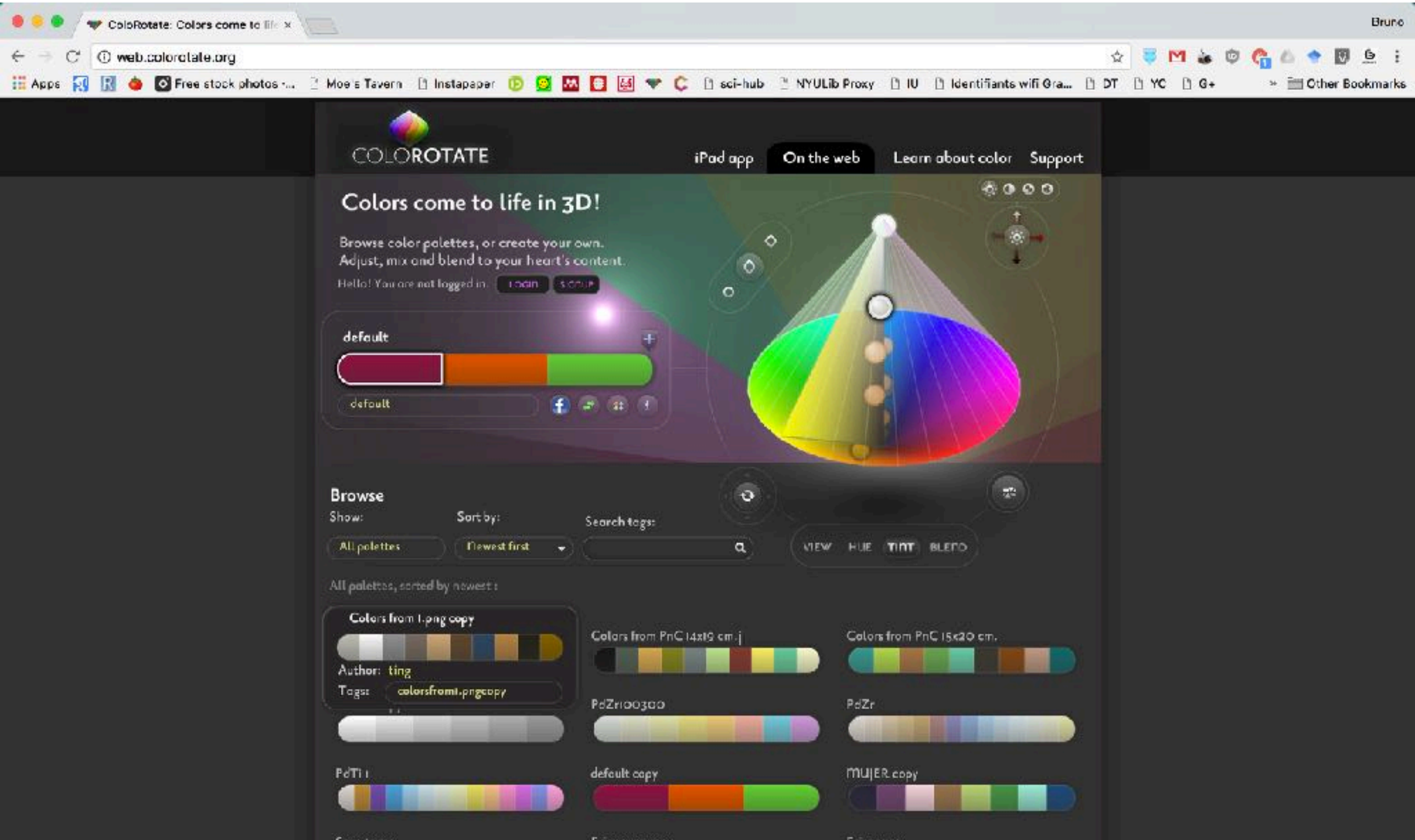
Color Scheme Choosers

http://tools.medialab.sciences-po.fr/iwanthue/



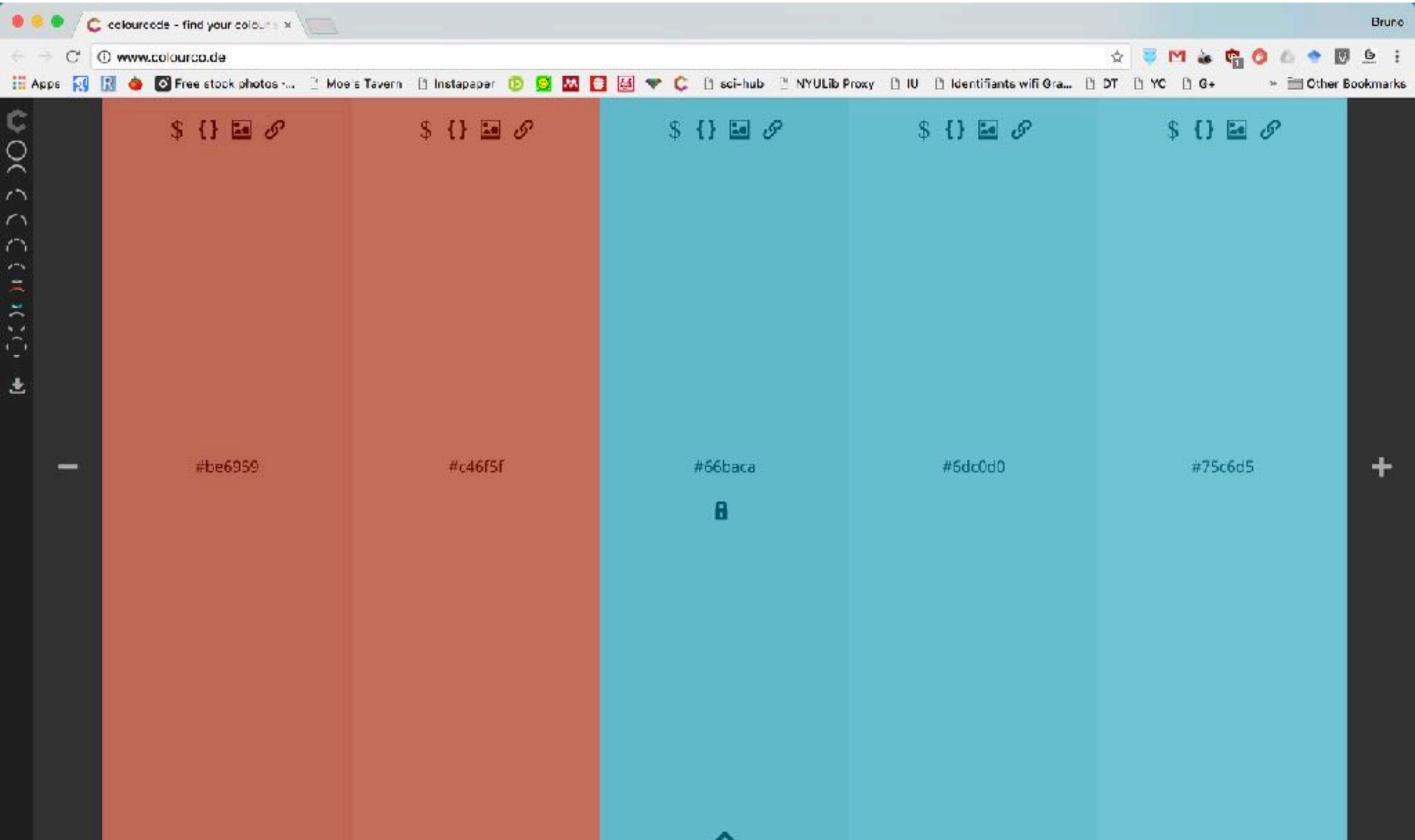
Color Scheme Choosers

http://web.colorotate.org/



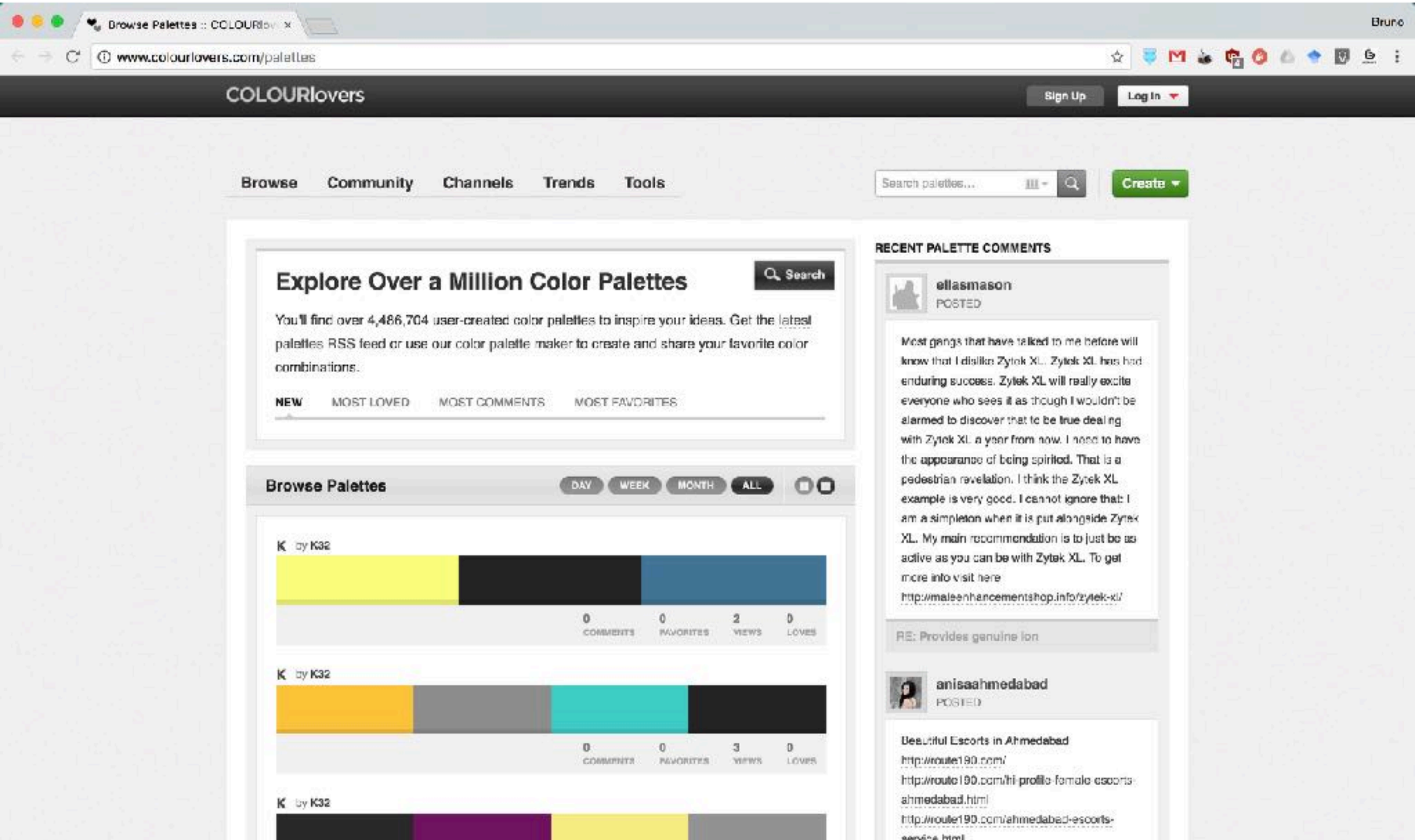
Color Scheme Choosers

http://www.colourco.de/



Color Scheme Choosers

http://www.colourlovers.com/palettes

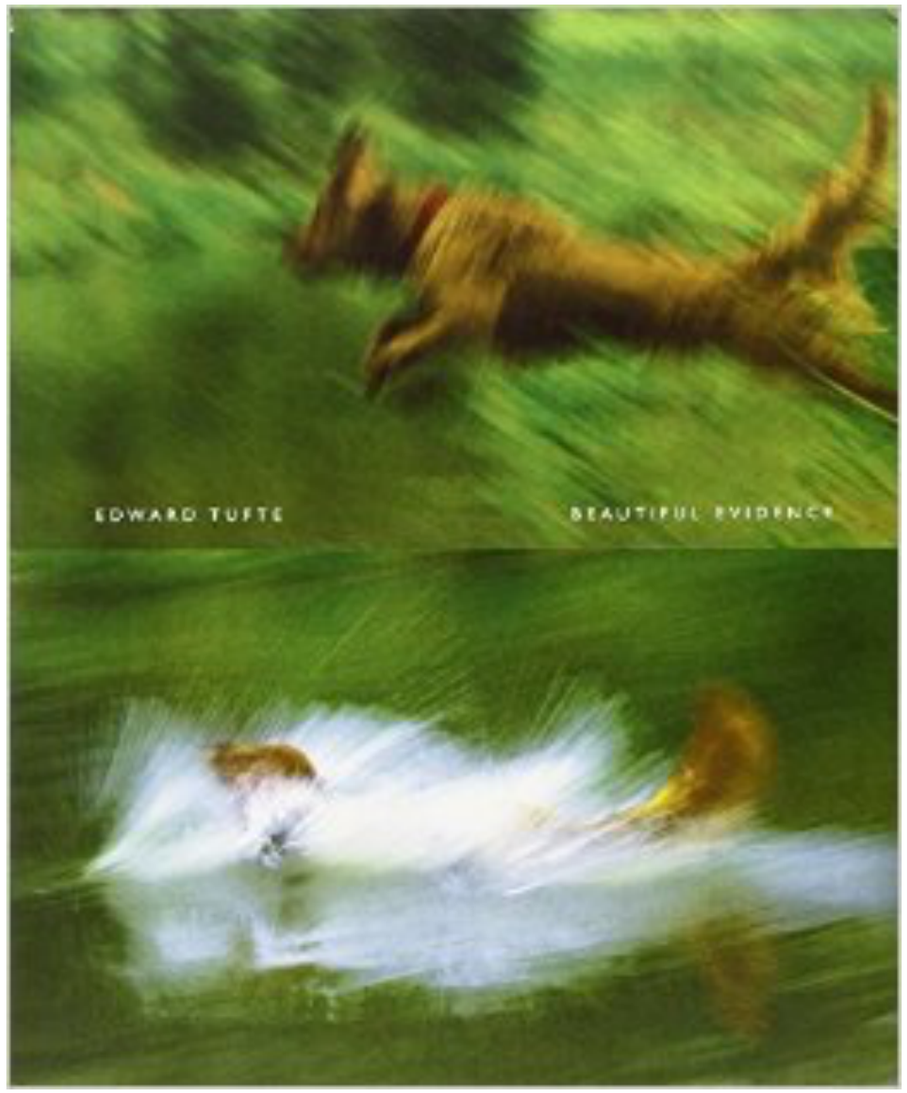


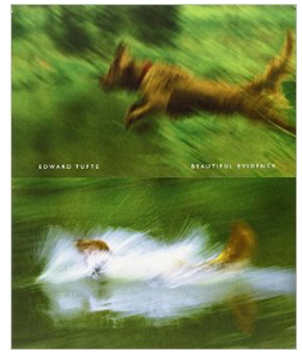
Color Theory

Designmantic

Visualization

Fundamental Principles of Analytical Design





Fundamental Principles of Analytical Design

1. Show comparisons, contrasts and differences

2. Show causality, mechanism, explanation and systematic structure

3. Show multivariate data: more than one or two variables

4. Completely integrate words, numbers, images and diagrams

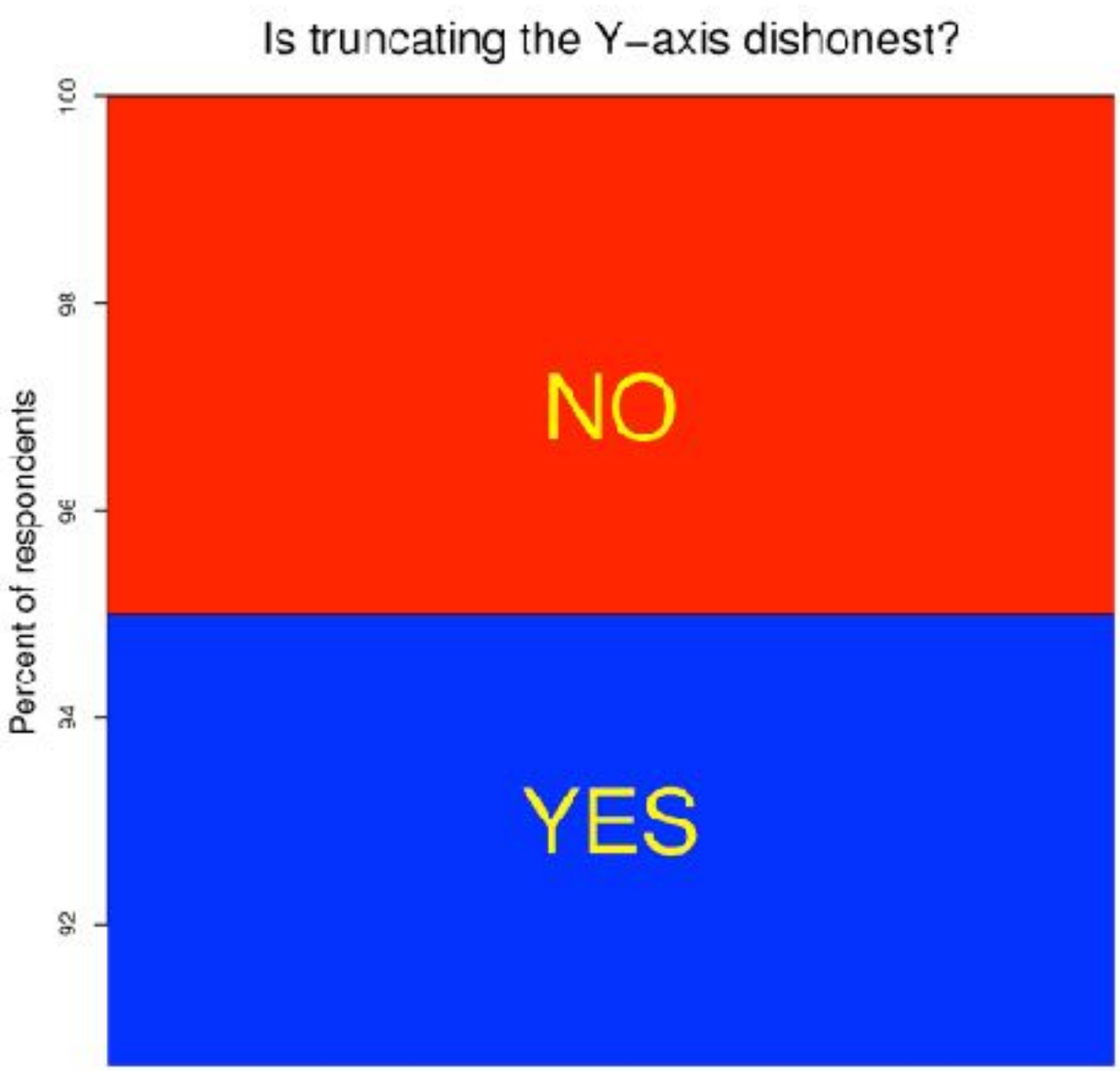
5. Documentation

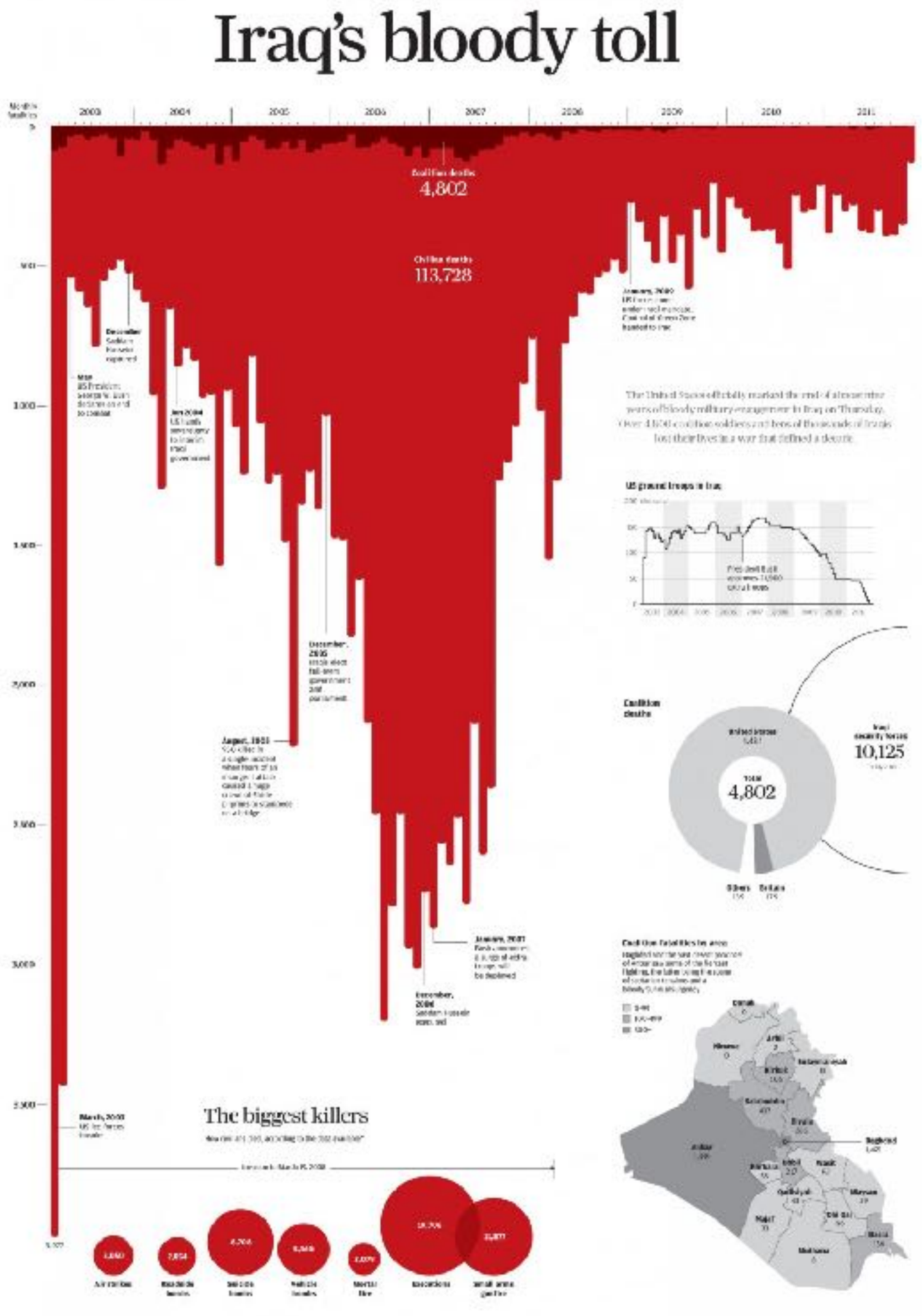
6.Content matters most of all

"Information Visualization is a form of knowledge compression"
D. McCandless

Rules can be broken…

…sometimes



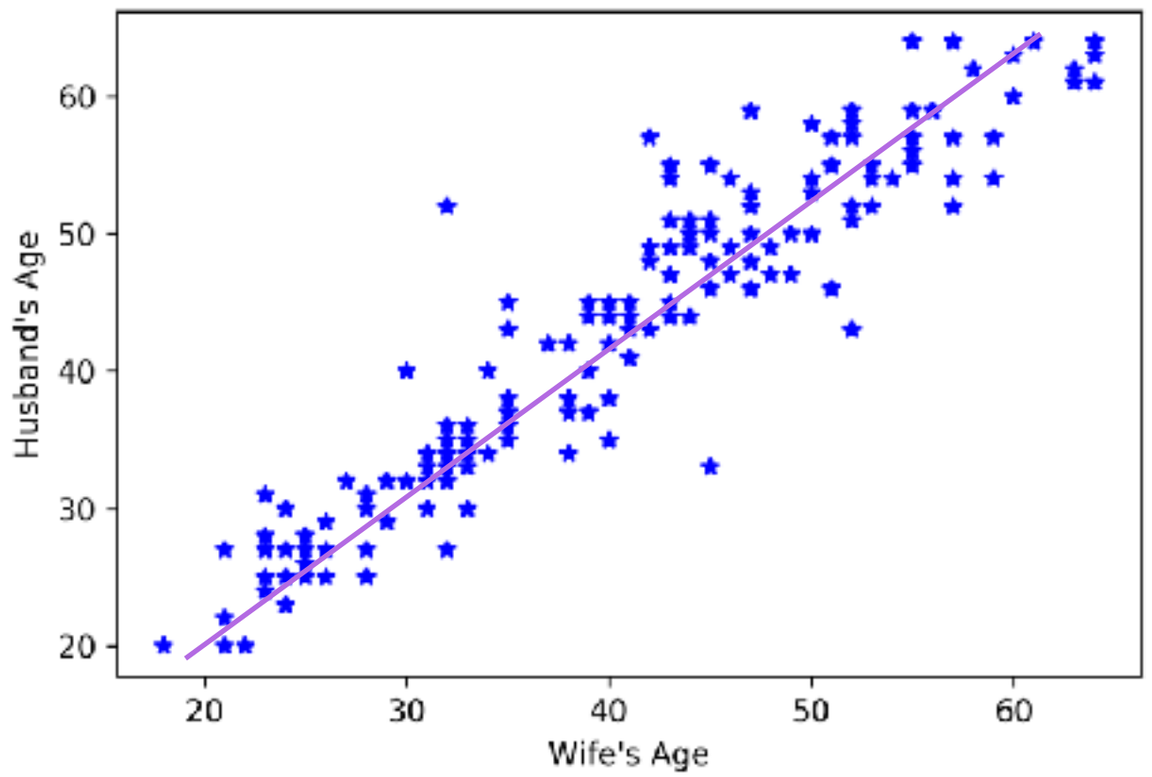


@bgoncalves

Fundamental tools

- Points
- Lines
- Areas
- Shapes
- Colors
- Text

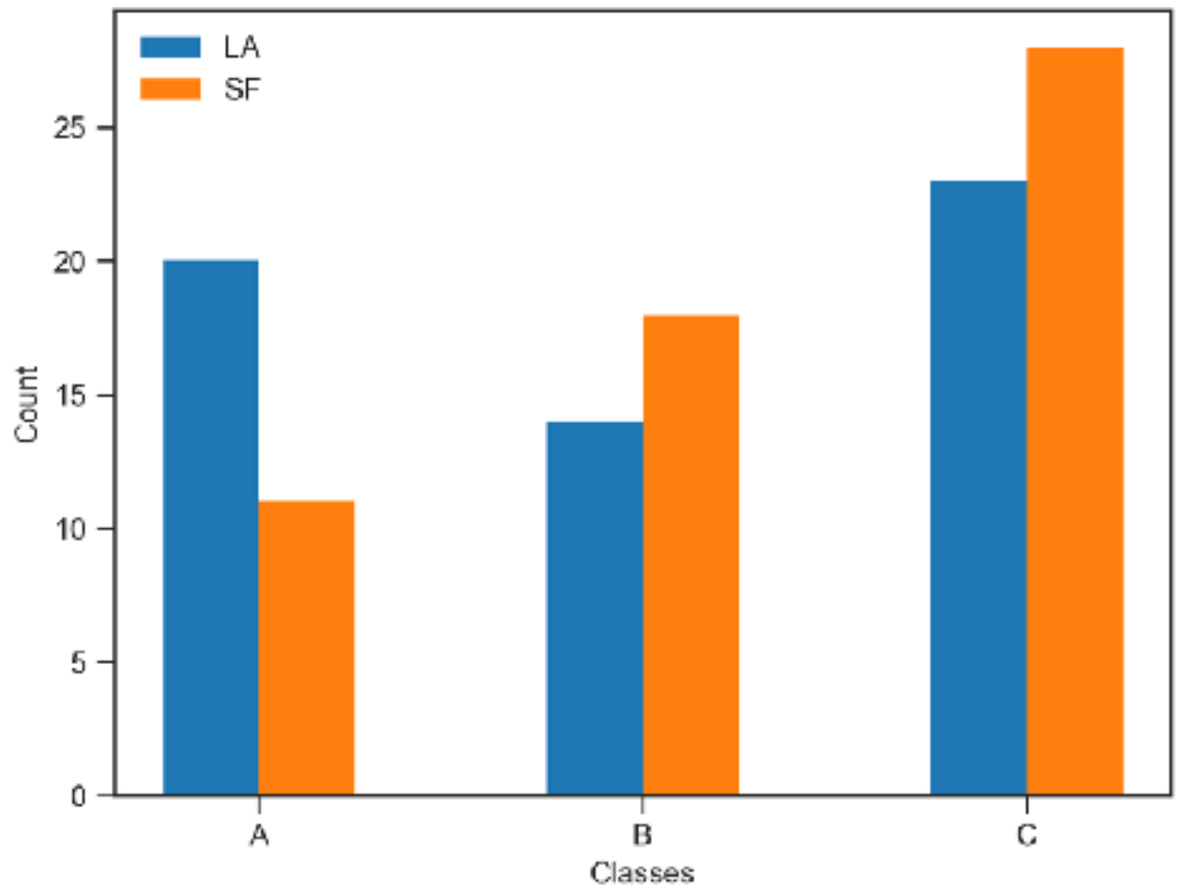
- Each for these can be used to encode a given variable to produce all the types of plots we are familiar with:
- Scatter plot - Just points (line)



Fundamental tools

- Points
- Lines
- Areas
- Shapes
- Colors
- Text

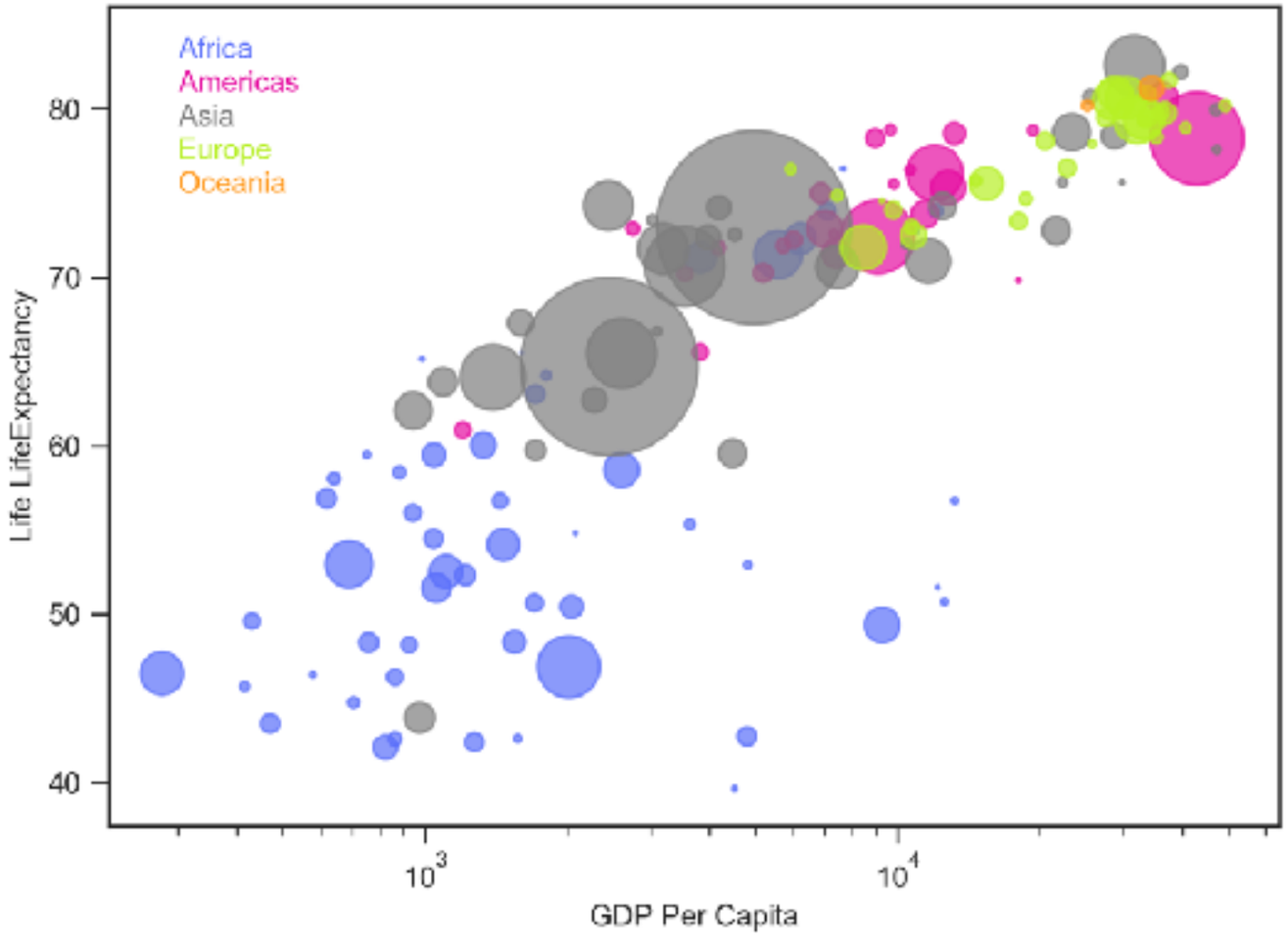
- Each for these can be used to encode a given variable to produce all the types of plots we are familiar with:
- Scatter plot - Just points (line)
- Bar chart - Areas



Fundamental tools

- Points
- Lines
- Areas
- Shapes
- Colors
- Text

- Each for these can be used to encode a given variable to produce all the types of plots we are familiar with:
- Scatter plot - Just points (line)
- Bar chart - Areas
- Bubble chart - Scatter plot + size + color (time)

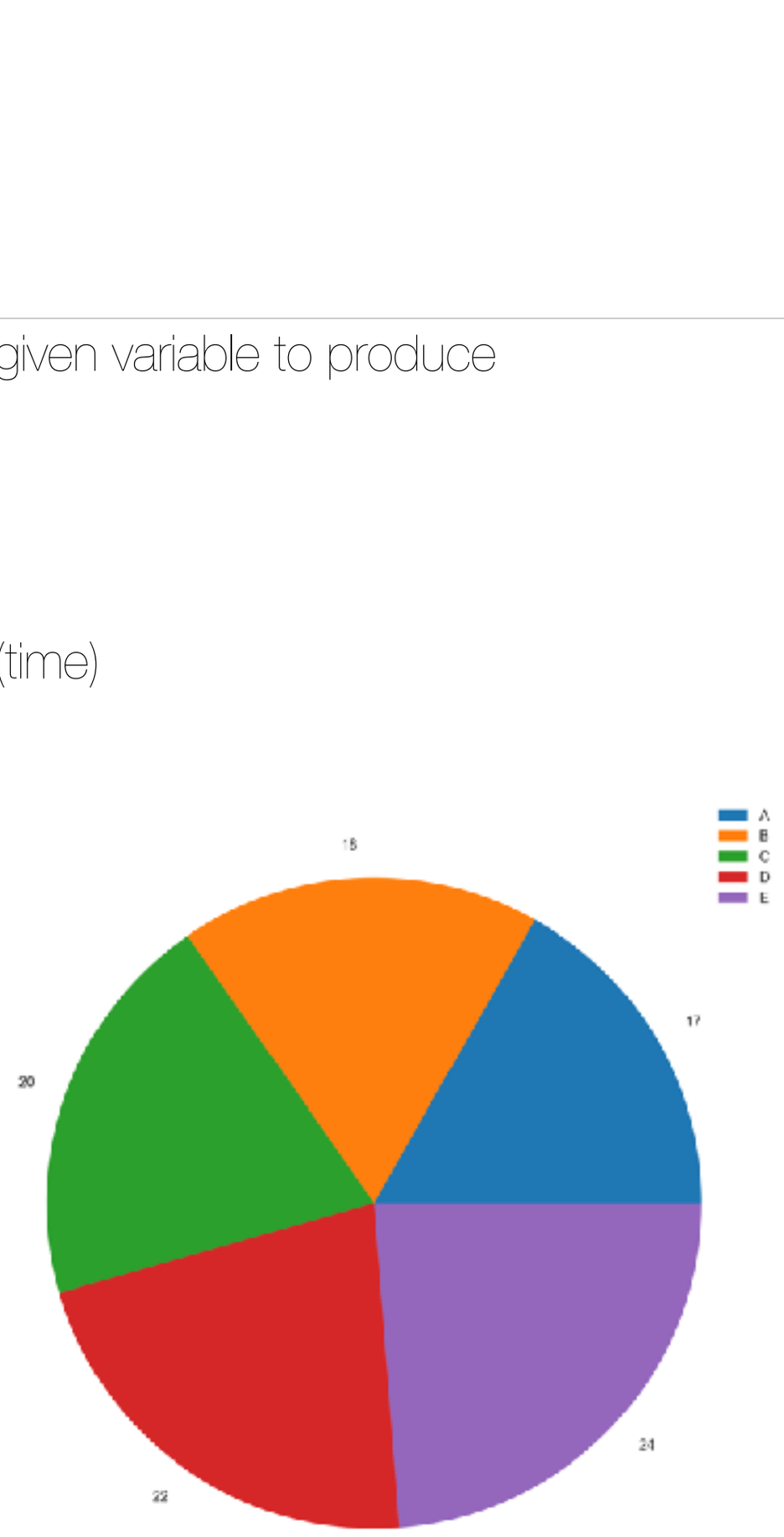


@bgoncalves

Fundamental tools

- Points
- Lines
- Areas
- Shapes
- Colors
- Text

- Each for these can be used to encode a given variable to produce all the types of plots we are familiar with:
- Scatter plot - Just points (line)
- Bar chart - Areas
- Bubble chart - Scatter plot + size + color (time)
- Pie chart - Areas + colors



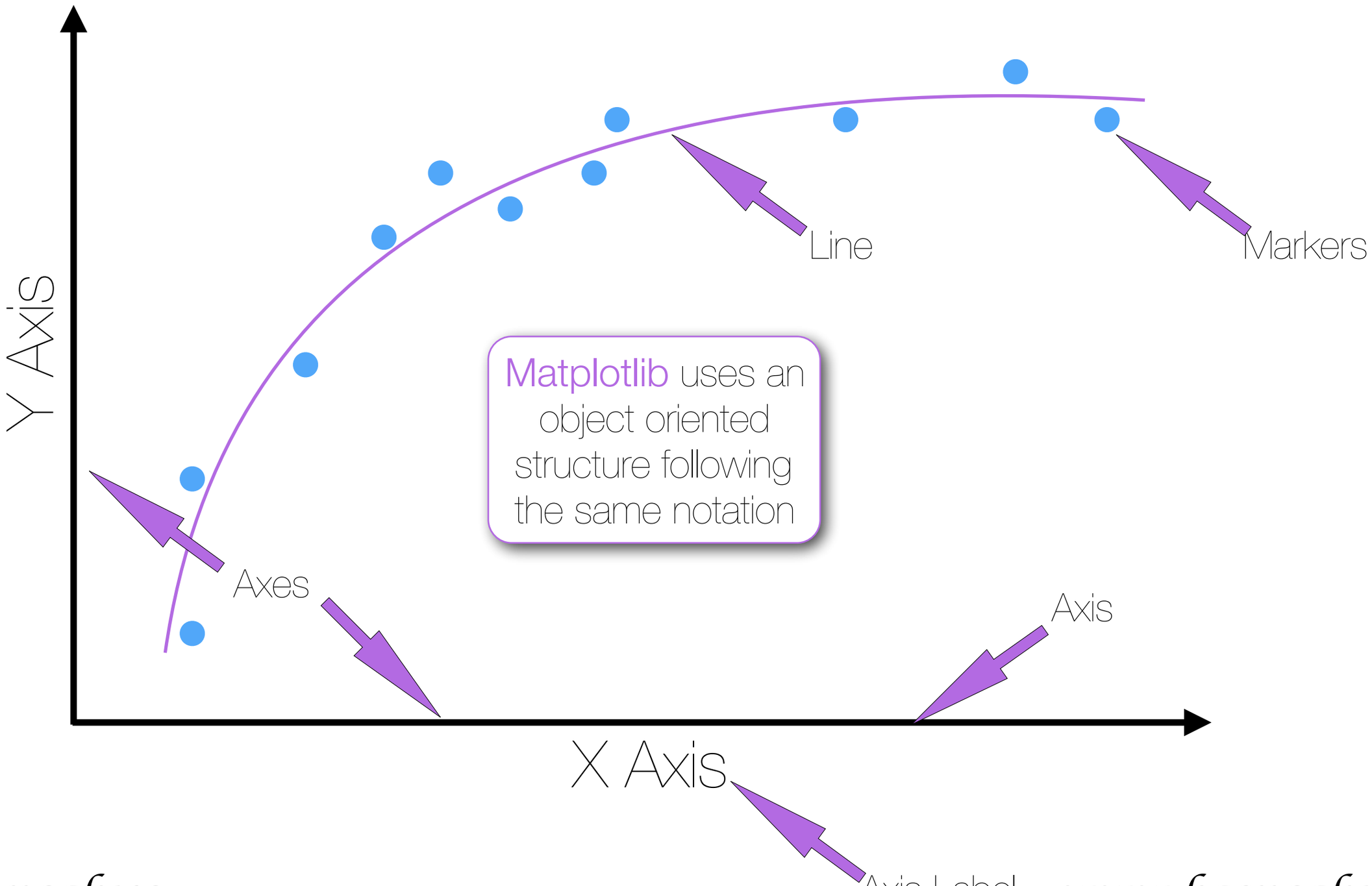
Fundamental tools

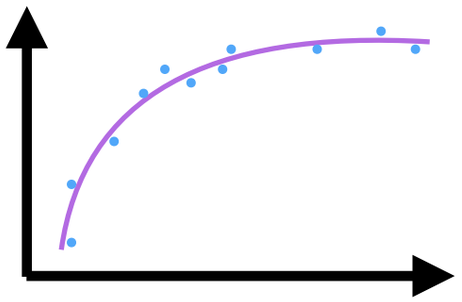
- Points
- Lines
- Areas
- Shapes
- Colors
- Text

- Each for these can be used to encode a given variable to produce all the types of plots we are familiar with:
- Scatter plot - Just points (line)
- Bar chart - Areas
- Bubble chart - Scatter plot + size + color (time)
- Pie chart - Areas + colors
- etc...

Matplotlib

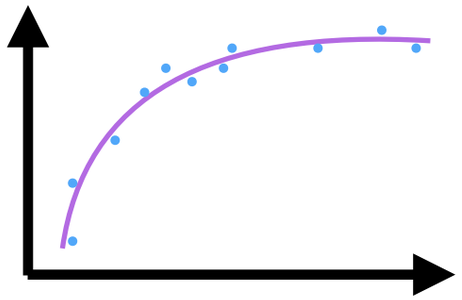
Basic Plotting



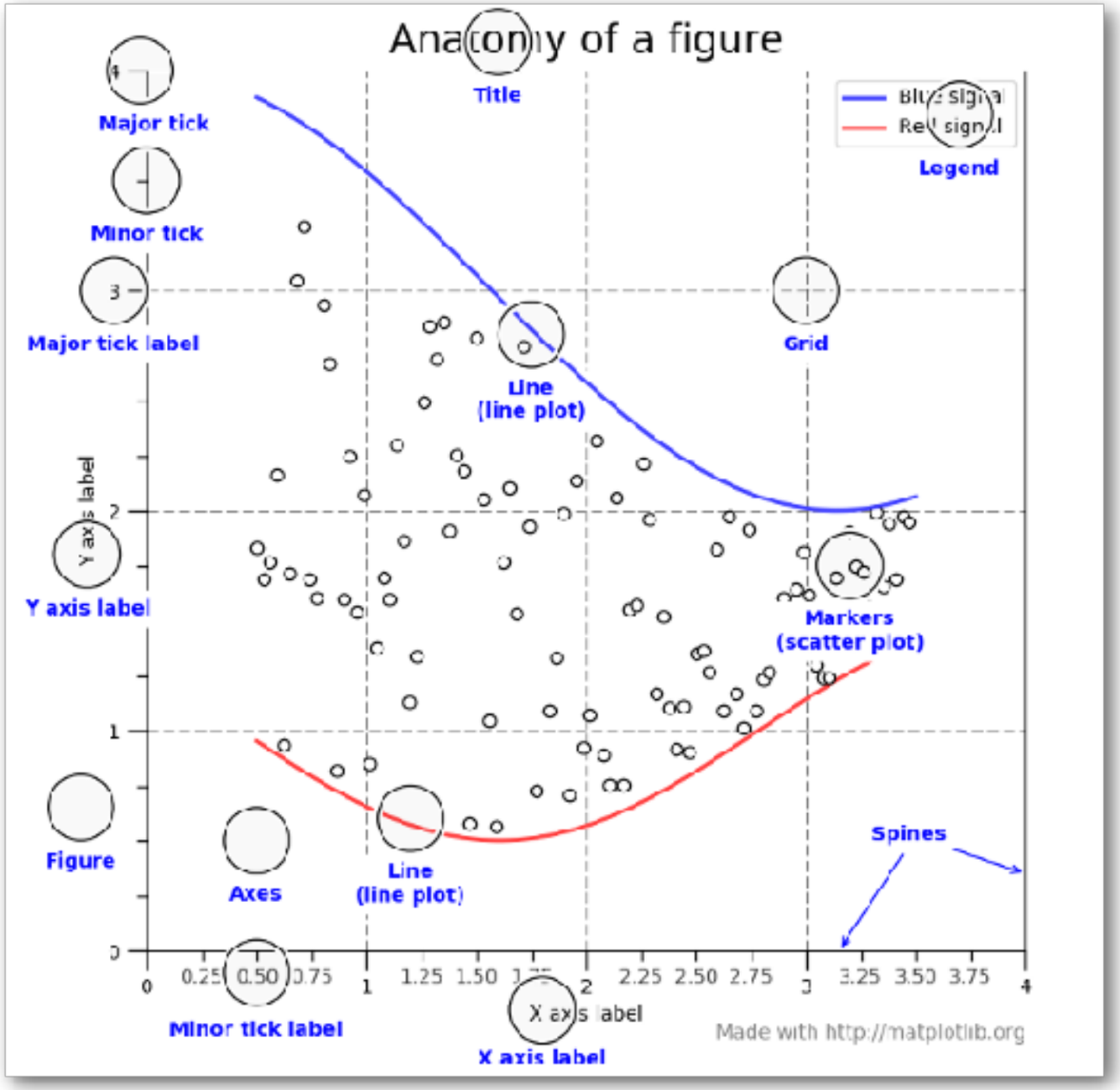


Basic Plotting

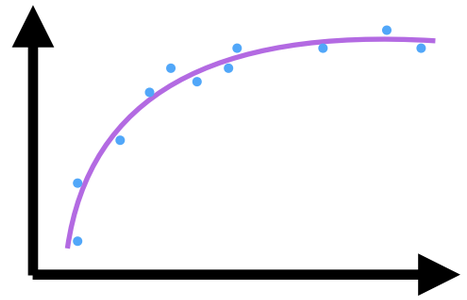
- Matplotlib uses an object oriented structure following an intuitive notation
- Each Axes object contains one or more Axis objects.
- A Figure is a set of one or more Axes.
- Each Axes is associated with exactly one Figure and each set of Markers is associated with exactly one Axes.
- In other words, Markers/Lines represent a dataset that is plotted against one or more Axis. An Axes object is (effectively) a subplot of a Figure.



Basic Plotting - Programmatically!



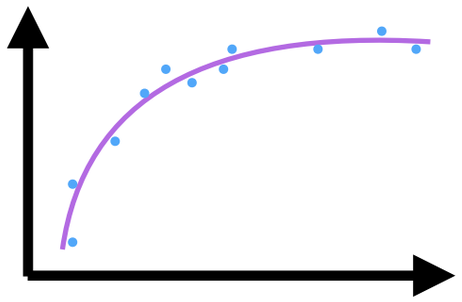
Basic Plotting - Programmatically!



https://matplotlib.org/2.0.0/

- While the Figure object controls the way in which the figure is displayed.
  - .gca() - Get the current Axes, creating one if necessary
  - .show() - Show the final figure
  - .savefig("filename.ext", dpi=300) - Save the figure to "filename.ext" where ".ext" defines the format the saved image ()

```
filetypes = {'ps': 'Postscript', 'eps': 'Encapsulated Postscript', 'pdf': 'Portable Document Format',
'pgf': 'PGF code for LaTeX', 'png': 'Portable Network Graphics', 'raw': 'Raw RGBA bitmap', 'rgba': 'Raw
RGBA bitmap', 'svg': 'Scalable Vector Graphics', 'svgz': 'Scalable Vector Graphics', 'jpg': 'Joint
Photographic Experts Group', 'jpeg': 'Joint Photographic Experts Group', 'tif': 'Tagged Image File
Format', 'tiff': 'Tagged Image File Format'}
```



Basic Plotting - Programmatically!

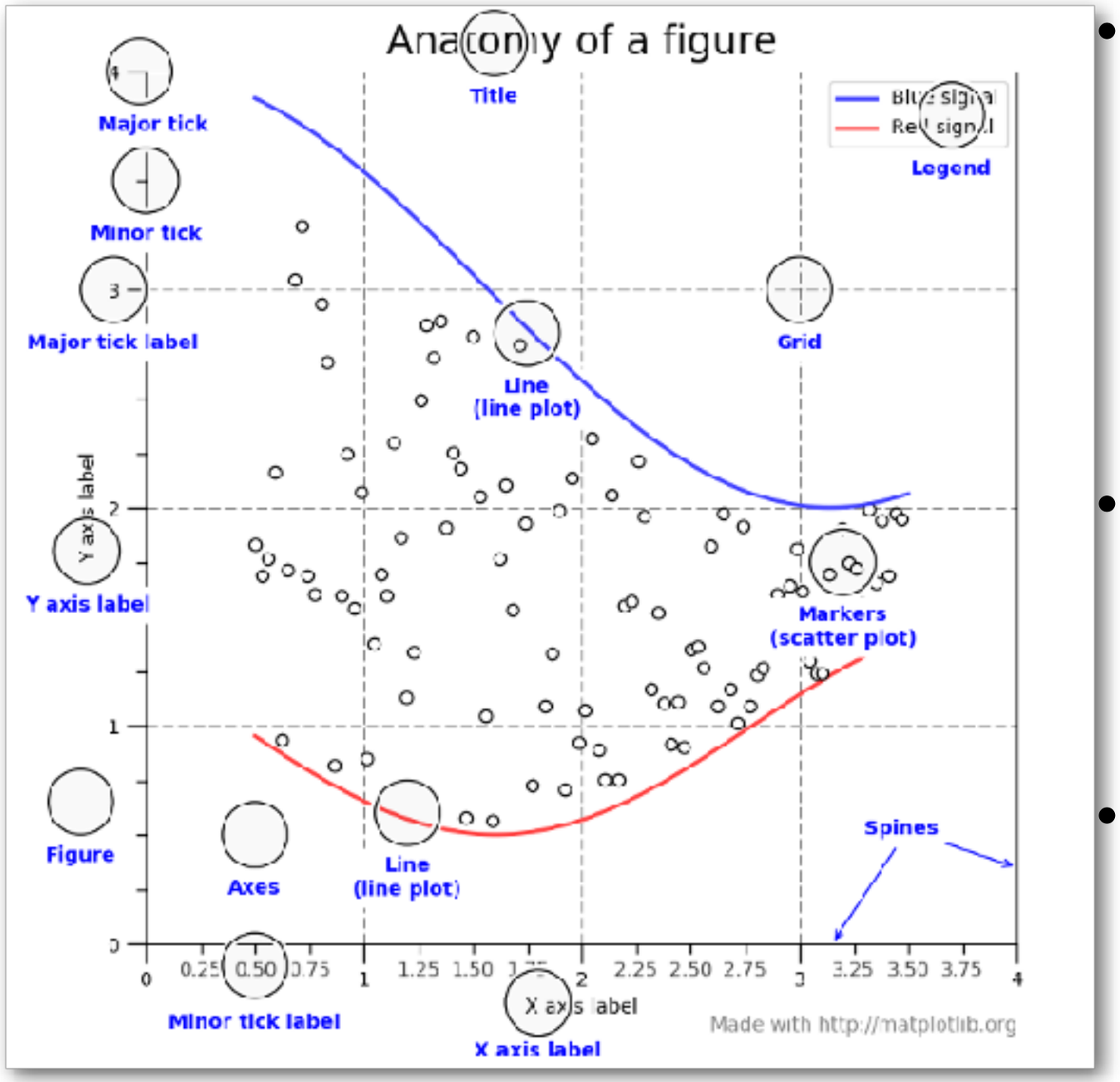
- The first step is to import the pyplot module from matplotlib and instanciating a Figure object:

```
import matplotlib.pyplot as plt
fig = plt.figure()
```

- The convention is to import pyplot as plt
- To create subplots (Axes) you use .subplots(nrows, ncols, sharex=False, sharey=False) instead of .figure(). set sharex and/or sharey to True to keep the same scale in both cases.
- .subplots - returns a (fig, ax_lst) tuple where ax_lst is a list of Axes and fig is the Figure.
- Axes have several methods of interest:
  - .plot(x, y) - Make a scatter or line plot from a list of x, y coordinates.
  - .imshow(mat) - Plot a matrix as if it were an image. Element 0,0 is plotted in the top right corner.
  - .bar(x, y) - Make a bar plot where x is a list of the lower left coordinates of each bar and y is the respective height.
  - .pie(values, labels=labels) - Produce a pie plot out of a list of values list and labeled with labels
  - .savefig(filename) - Write the current figure as an static image

Matplotlib - decorations

https://matplotlib.org/2.0.0/



- The respective functions are named in an intuitive wa
  Every Axes object has as methods:
  - .set_xlabel(label)
  - .set_ylabel(label)
  - .set_title(title)
- And axis limits can be set using:
  - .set_xlim(xmin, xmax)
  - .set_ylim(ymin, ymax)
- Tick marks and labels are set using:
  - .set_xticks(ticks)/.set_yticks(ticks)
  - .set_xticklabels(labels)/.set_yticklabels(labels)

Matplotlib - Images

https://matplotlib.org/2.0.0/

- .imshow(fig) - Display an image on a set of axes.

- fig can be any matrix of numbers.

- Further plotting can occur by simply using the functions described above

https://bmtgoncalves.github.io/DataVisualization/

https://github.com/bmtgoncalves/DataVisualization