Warm-Up Activity

http://old.vijayp.ca/blog/2012/06/colours-in-movie-posters-since-1914/

What is the visualization trying to show?

Colours in movie posters since 1914. The width of each hue represents the amount of that hue across all images for that year, and the saturation and lighting were the weighted average for all matching pixels.

What are its methods?

The author downloaded images of movie posters online, grouped the movie posters by the year, and then counted the total number of pixels for each colour in the year. After normalizing and converting to HSL coordinates, the visualization was generated.

What are the strengths / weaknesses?

Strengths: Intuitive

Weakness: Possible bias in the data

Transformations

Affine transformations satisfy:

$$\vec{y} = A\vec{x} + \vec{b}$$

"affine" transformations maintain parallel lines. As compared to projection, euclidian, or similar transformations.

We can use these to accomplish:

Shifts: from one location to another

Rotations: degrees

Scaling: stretching / shrinking

Choosing a Visualization

When we are examining data, what can we look for?

Does this data describe a geometric object?

Are the data points connected to each other?

Can we describe data points with a fixed set of categories?

Is there a quantity associated with the data?

Are the datapoints continuous along one or more dimensions?

Categories and Continuity
Today we'll talk about representing things based on categories and continue discussing continuities.

How Do Colors Work?
Rods (low-light) and cones (color) mediate vision. Humans have about 20 times as many rods (120 million) as cones (6 million).
Rods: help us see light / activated or not / much more than rods: we are better at perceiving brightness than color
Cones: help us to see colors / determine whether we are able to see particular colors

Color Matching Function
How different cones see different colors; wavelength
The graph on the slide shows how strongly the cones are receiving lights of certain wavelength (how activated the cones are)
Ex. green cones see green frequency lights; human eyes are sensitive to red

"Naming" Colors
RGB triplets, sometimes compressed into hexadecimel ("#00FFAA", etc)
Hexadecimal only shows 16 million colors, RGB as uncompressed floats can theoretically represent quite a bit more (but there's a limitation of what monitors can display)
R = 2^8 G = 2^8 B = 2^8
RGB = 2^24 = 16 million
hexadecimal = 16^6 = 16 million

The mandrill image was an early color standardization image established by the Dept of Defense

Color spaces
- HSV (Hue, saturation, value): typically a color space used by color designers.
- CIELAB: the color space that covers the average of human vision.
- sRGB, Adobe sRGB: "standard RGB" is a color space that was standardized to unify different monitors and printers.

Examine how they work in Notebook COLOR provided on course page

Back to the the mandrill image:
This video scans through different hues, which are the part of the rainbow you want
Then different saturations, which is how vibrant or gray the colors are
Then different values, which is how bright or dark the colors are.

This unified space of colors that works on most displays and printers is NOT the full range of human perception.
The images shows that CIELAB > sRGB

List of colors by name
[Web]
[matplotlib]
coding environments will often provide "named" colors if you're more interested in simplicity than flexible design

Color Palettes
Resources:
[colorbrewer]
palettable (package)

Sequential Colormaps
these are also known as monochromatic
the hue doesn't change, but the value and the saturation do

Diverging Colormaps
this shows you that different colors in the rainbow have different perceived brightness.
The outsides of this color map get darker while the center is brighter.

Qualitative Colormaps
The blended areas of the continuous map here are kind of gross looking. This kind of random color map is best for categorical data that is NOT continuous.

Color Meaning
Which end of this is hot and which end is cold? What do red and blue mean?
Answer: misleading in this case (blue means higher temp)
red in the USA is bravery, but red in China is revolution
green in the western world is Christmas, but green in the eastern world is Islam
You're not always picking colors relative to each other, but relative to culture as well.

It's worth noting that "color words" are not consistent across languages or cultures. Color is a product of culture.
Human skin color is a particularly SENSITIVE area of color psychology. Designers are becoming more aware of this. Western culture often treats dark colors as "mysterious" or "evil" and this is largely rooted in a history of racial bias.

It's full of colors
https://commons.wikimedia.org/wiki/File:16777216colors.png
This grid is one way to show all the 16 million colors. They kind of get blurred out when looking at the image in this size, but if you zoom in at the link provided, you can see every pixel from red to green to blue, and from black to white.

HSV Wheel
https://commons.wikimedia.org/wiki/File:HSV_color_solid_cylinder.png
you can see that saturation and value are linear axes while hue is cyclical.

Magenta
just a point of interest, what wavelength would you say magenta (or purple) is?
Answer: Not exist on the wavelength
Invented by our brain: mixture of blue wavelength and red wavelength as we see them at the same time (blue and red cones are activated)

Palette Mapping: Assign each value to a specific color or element.

Color Mapping

$f(v) \rightarrow (R,G,B)$
We can also re-map:
$f(v') \rightarrow (R,G,B)$
$v' = f(v)$
For instance, with logs or squares.

We can fit our continuous distribution linearly to a color map - so a low data value is at the beginning of the color map, a high value is at the top, and a middle value is halfway up the map.
Or we can remap it with exponents, or other mathematical operations.

Color Mapping: Linear Mapping

We map from a range of values to (0, 1):
$v' = (v - v0)/(v1 - v0)$

in order to get the data values to line up with the color bar - which goes from 0 to 1 - we have to transform our data into a range of 0 to 1. This can be as easy as dividing all value by the largest value, or we can do fancier things.

RGB Components
Picture 1: How we get certain colors (ex. gray)
Picture 2: b&w: blue line overlapped with green and red line; same amount of blue, green, and red lead to no hue
Picture 3: saturation/hue, no brightness changes
Picture 4: nice color map to use for quantitative analysis

Color Blindness
these are not the only types of color blindness, just the types where one of the three types of cone are missing. Sometimes that third cone is just "anomalous", and sometimes more than one cone is missing, but it's much rarer.
You can see that designing for a colorblind person can be difficult if you put all the information into one cone. You catch a wider audience by using the colors between the cones, like yellow, cyan, and magenta.
ex. do not use pure red and pure green as it is a common color-blindness
http://enchroma.com/test/instructions/

Coding Demo:
https://uiuc-ischool-dataviz.github.io/spring2019/nbv.html?notebook_name=%2Fspring2019%2Fweek05%2FDEMO_week05.ipynb