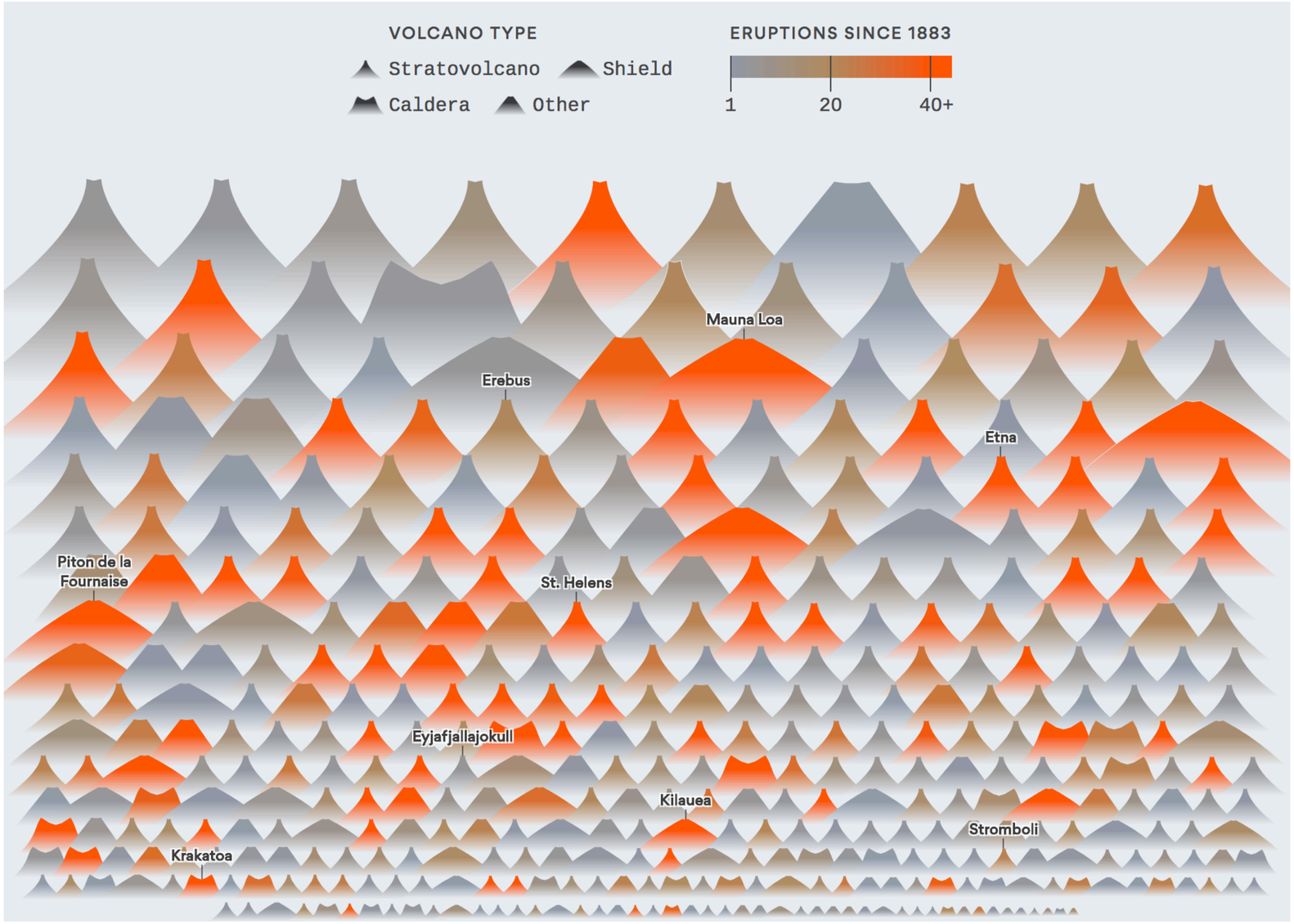
Data Visualization & Design

Week 3

This week in **visualization**…


VOLCANO TYPE
Stratovolcano
Shield
Caldera
Other
ERUPTIONS SINCE 1883
1
20
40+
Mauna Loa
Erebus
Etna
Piton de la Fournaise
St. Helens
Eyjafjallajokull
Kilauea
Stromboli
Krakatoa

Source

Github+Microsoft

1. Color **Perception** & **Representation**
2. **Applying** Color
3. Practical **Tips**
4. Homework Review

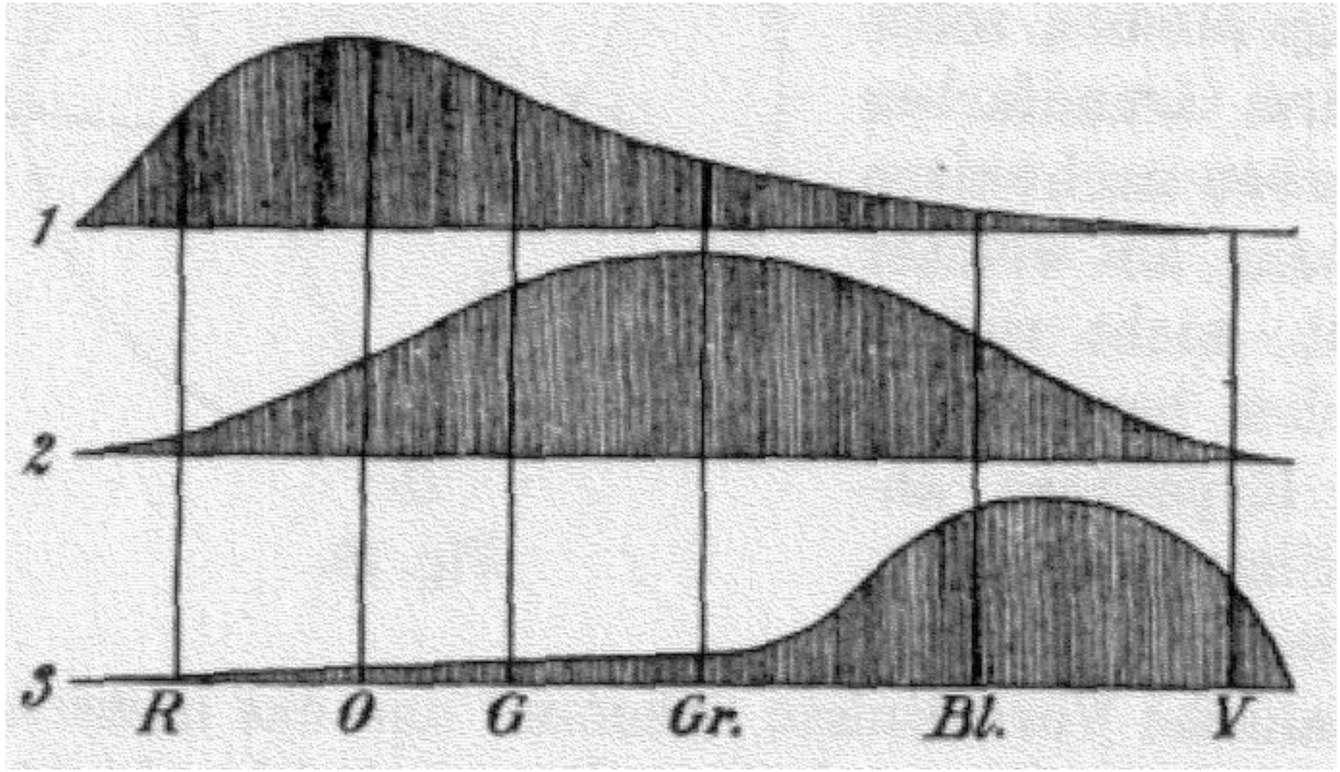
1. Color **Perception** & **Representation**
2. **Applying** Color
3. Practical **Tips**
4. Homework Review

Human Perception of Color

When visualizing data, **color is a versatile tool** for encoding values and communicating difference.

- In 1801, scientist **Thomas Young** proposed that the retina of the human visual system contains three “kinds of fibers,” each sensitive to a different wavelength of light.
- This theory, referred to now as the **“trichromatic theory of color vision,”** informs the way computers display color images, and holds implications for human color perception.

Relative sensitivity



Wavelength

Concept — **Tristimulus theory**

- There are exactly **three different types of color receptors** in the human eye
  - Sensitivity to ***long*** *wavelengths*
  - Sensitivity to ***medium*** *wavelengths*
  - Sensitivity to ***short*** *wavelengths*
- As such, any color can be uniquely represented using **three values**
- Color is an inherently **three-dimensional space**

Representing Color **Digitally**

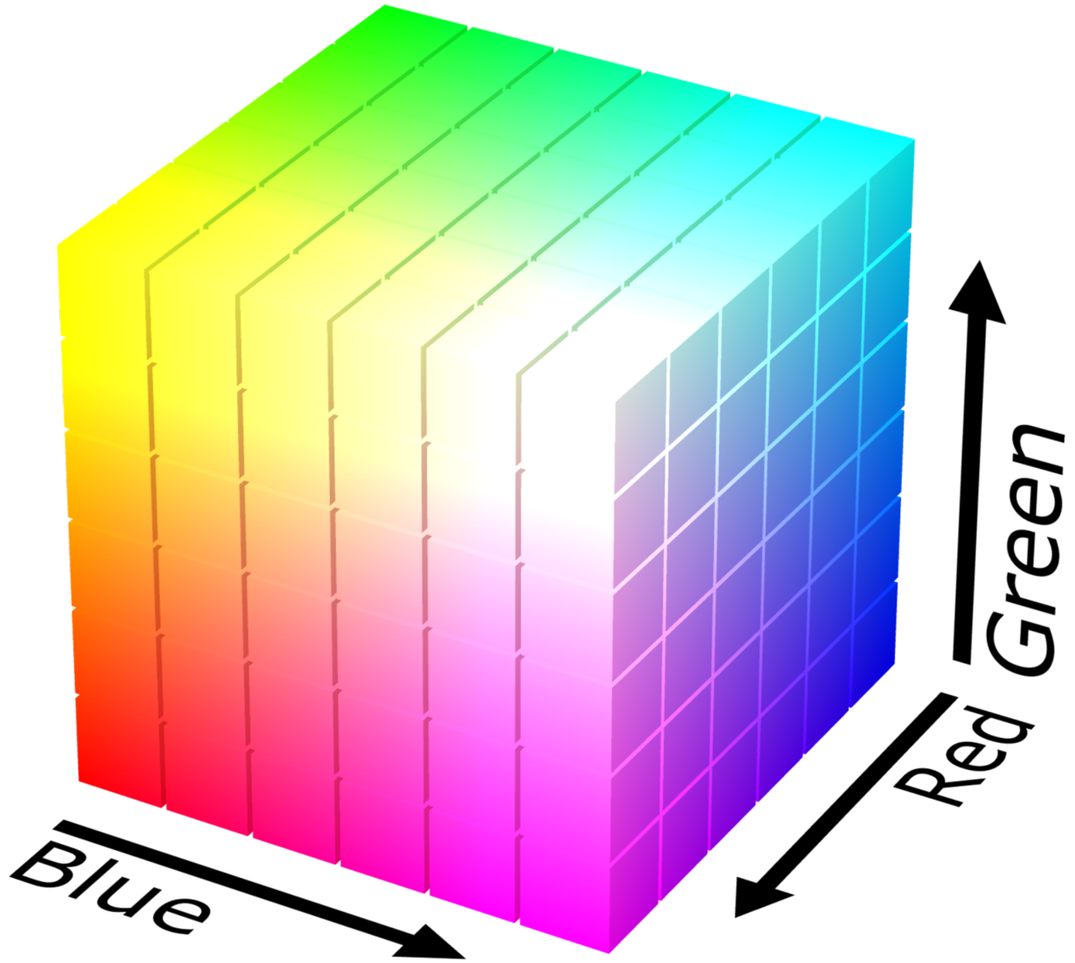
Digital representations of color tie together both human perception and the physical properties of light.

Concept: **Color Space**

- A way of organizing colors that takes into account both **human perception** and **digital representation**
- Color spaces enable reproducible representations of color
- May be **arbitrary** (ex. Pantone) or **mathematical** (ex. RGB)

Example: ***Arbitrary*** Color Space (Pantone system)

Example: ***Mathematical*** Color Space (RGB)

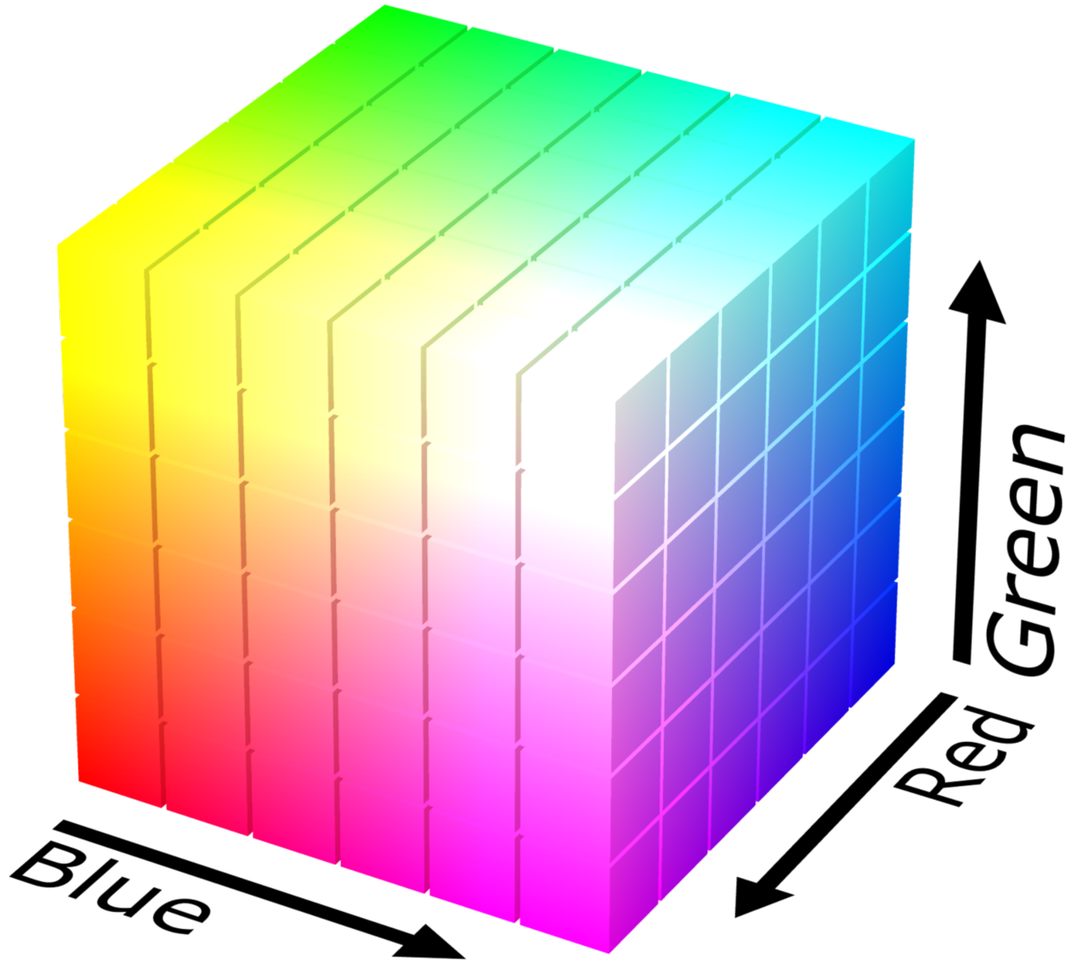


Common color spaces are widely implemented in visualization software and design tools.

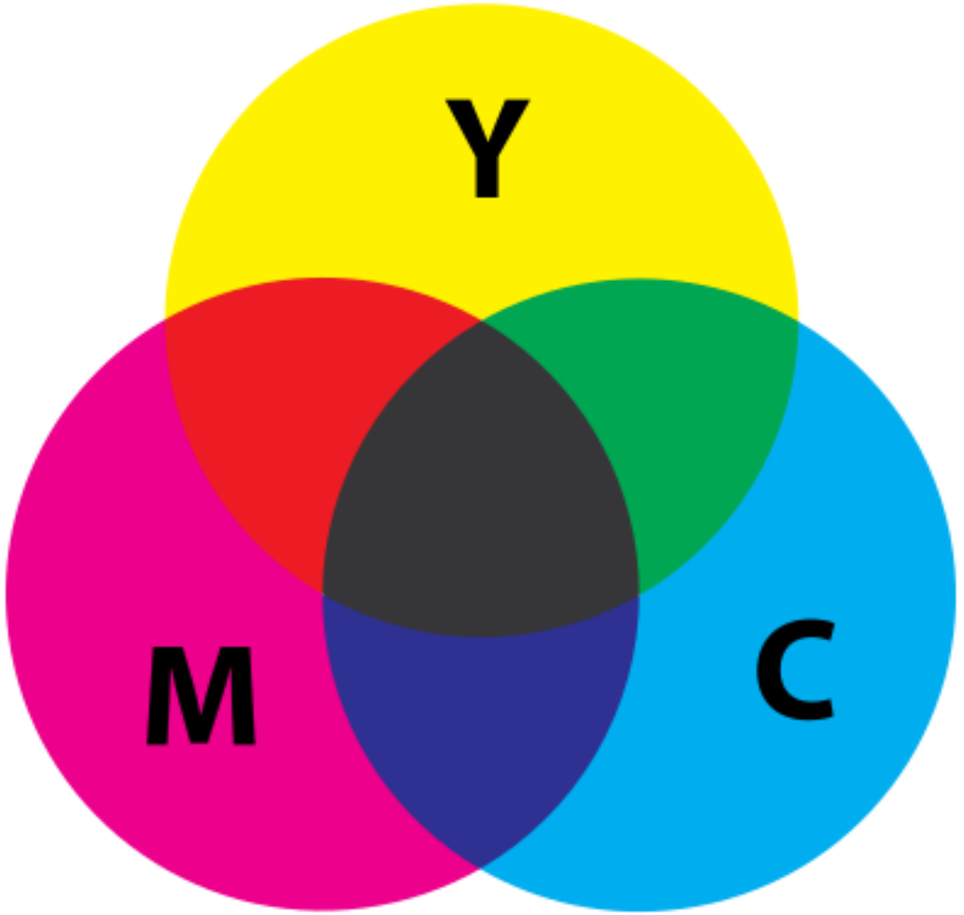
List of common color spaces

- **RGB** (Red, Green, Blue)
- **CMYK** (Cyan, Magenta, Yellow, Black)
- **HSL** (Hue, Saturation, Lightness/Luminance)
- **CIELAB** (or L*a*b*: Lightness, Green–Red and Blue–Yellow)

RGB (Red, Green, Blue)

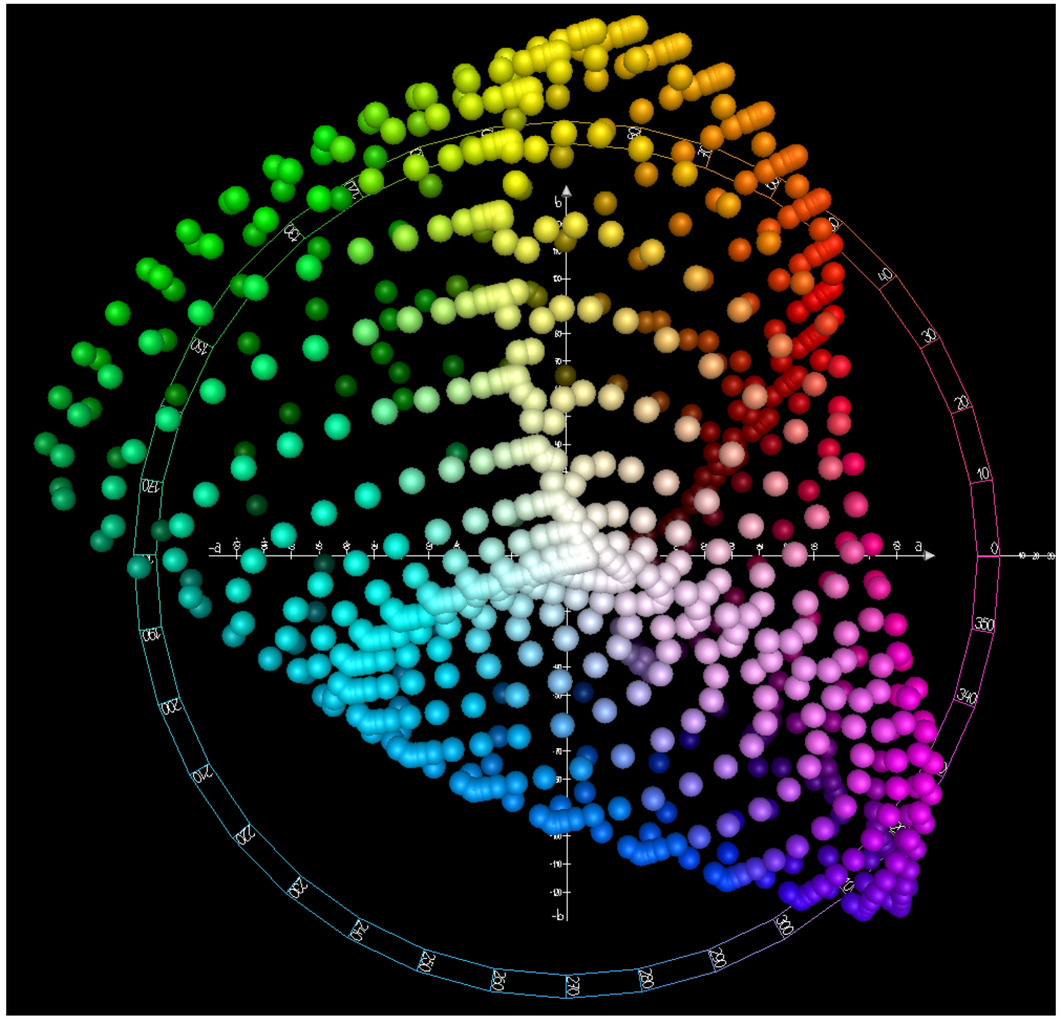


CMYK (Cyan, Magenta, Yellow, Black)



HSL (Hue, Saturation, Lightness/Luminance)

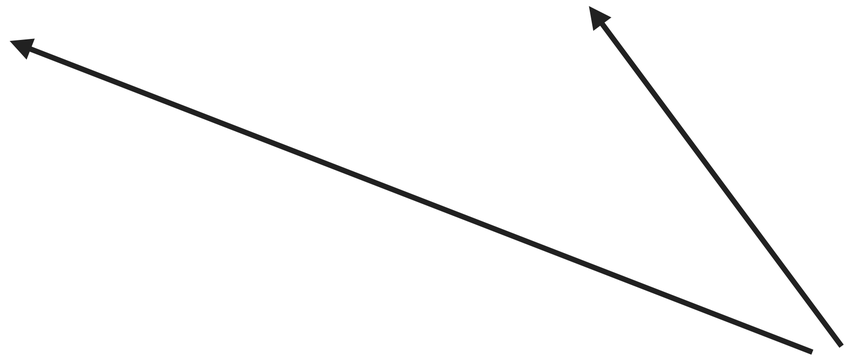
CIELAB (or L*a*b*: Lightness, Green–Red and Blue–Yellow)†



†**International Commission on Illumination**

List of common color spaces

- **RGB** (Red, Green, Blue)
- **CMYK** (Cyan, Magenta, Yellow, Black)
- **HSL** (Hue, Saturation, Lightness/Luminance)
- **CIELAB** (or L*a*b*)



Especially useful for data visualization

In certain color spaces, **color mixing** informs the combined behavior of different hues.

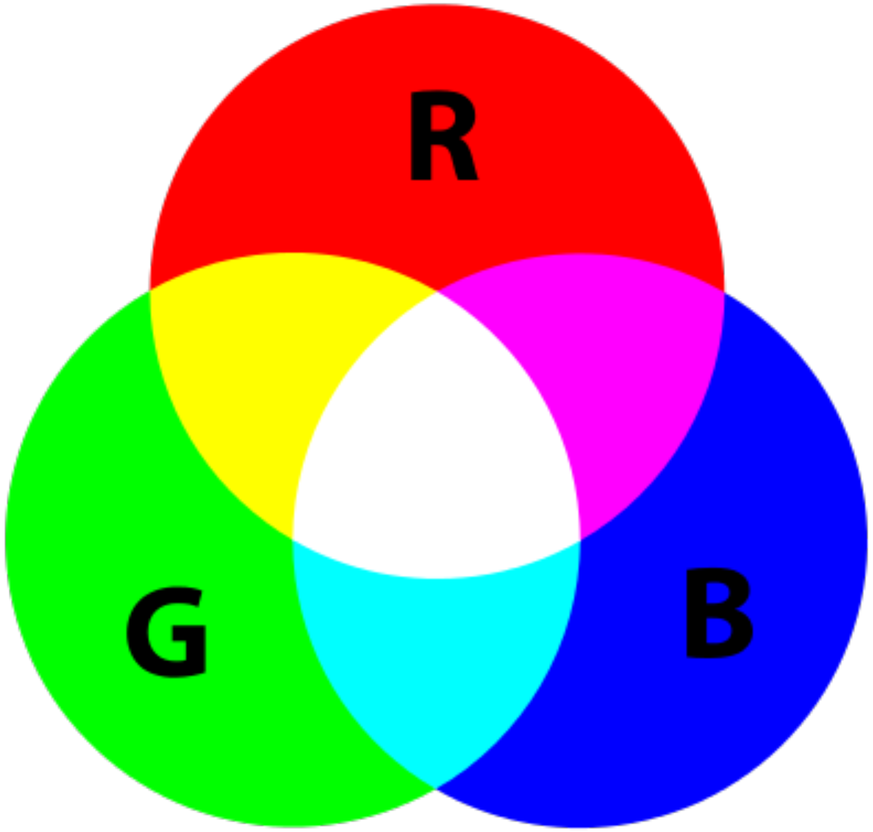
There are two types of color mixing:

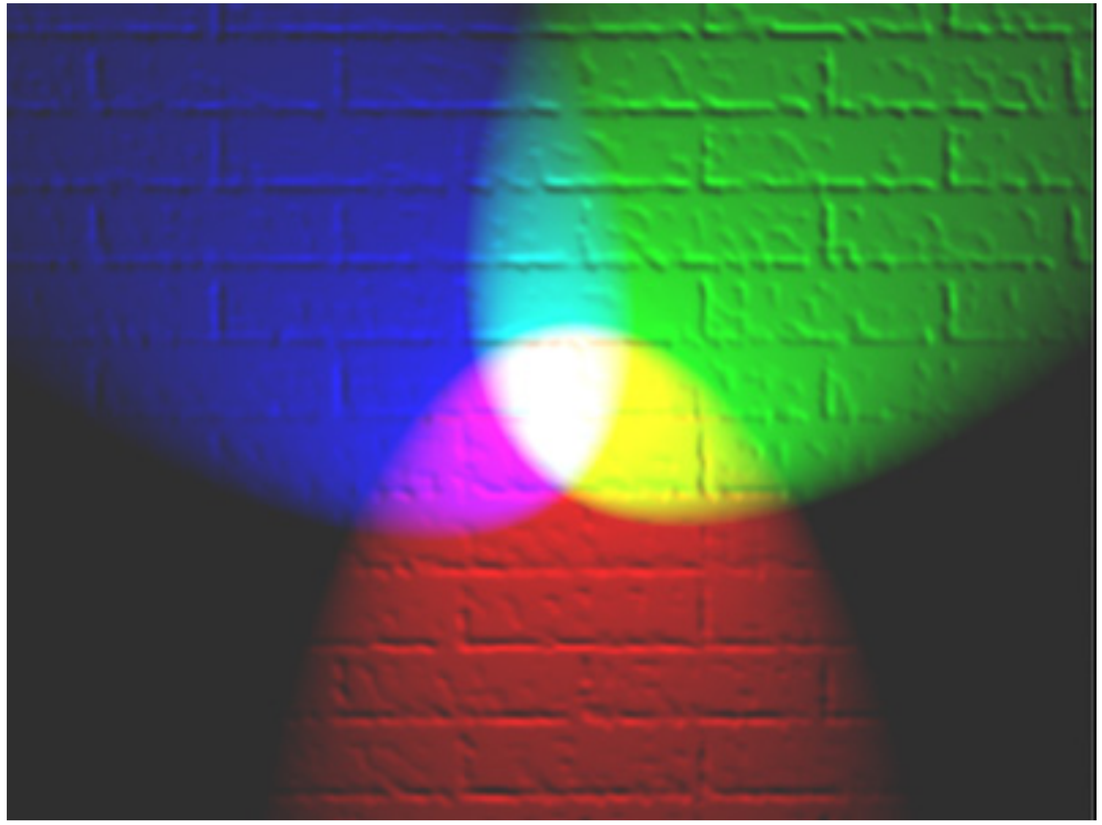
1. **Additive**
2. **Subtractive**

Color Mixing: **Additive**

- Additive color is created by **mixing different light colors**
- **RGB (most common)**: the three values in this space refer to the intensity output of the three light colors used in a monitor, television, or projector
    - **R** — Red
    - **G** — Green
    - **B** — Blue
- Does not correspond to the mixing of physical colors

Color Mixing: **Additive**

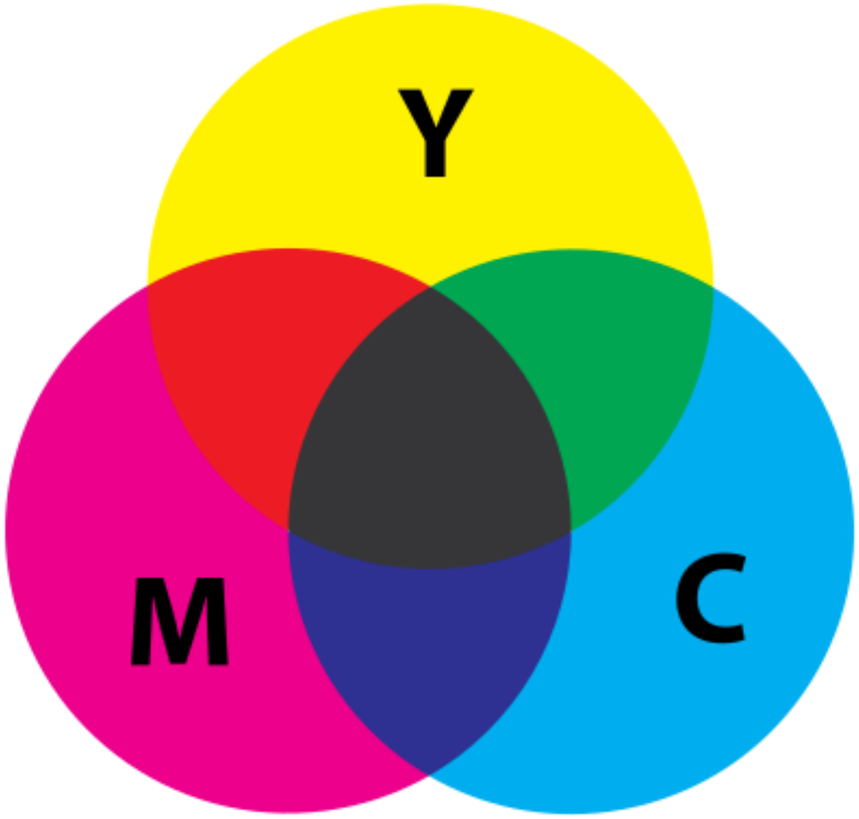




Color Mixing: **Subtractive *(Absorbs Light)***

- Additive color is created by **mixing different ‘paint’ colors**
- **CMYK (widely used)**: the three primary hues in this space are:
    - **C** — Cyan
    - **M** — Magenta
    - **Y** — Yellow
- Does correspond to the mixing of physical colors
- Black can be approximated, but pure black is nearly impossible to achieve

Color Mixing: **Subtractive *(Absorbs Light)***



Some color spaces **do not rely on color mixing** to create additional hues.

HSL (Hue, Saturation, Lightness/Luminance)

- **Hue —** The actual color
- Saturation
- Luminance

PAUSE:

Group and order your swatches by **Hue**.

- Hue
- **Saturation —** The amount of grey in a color
- Luminance

PAUSE:

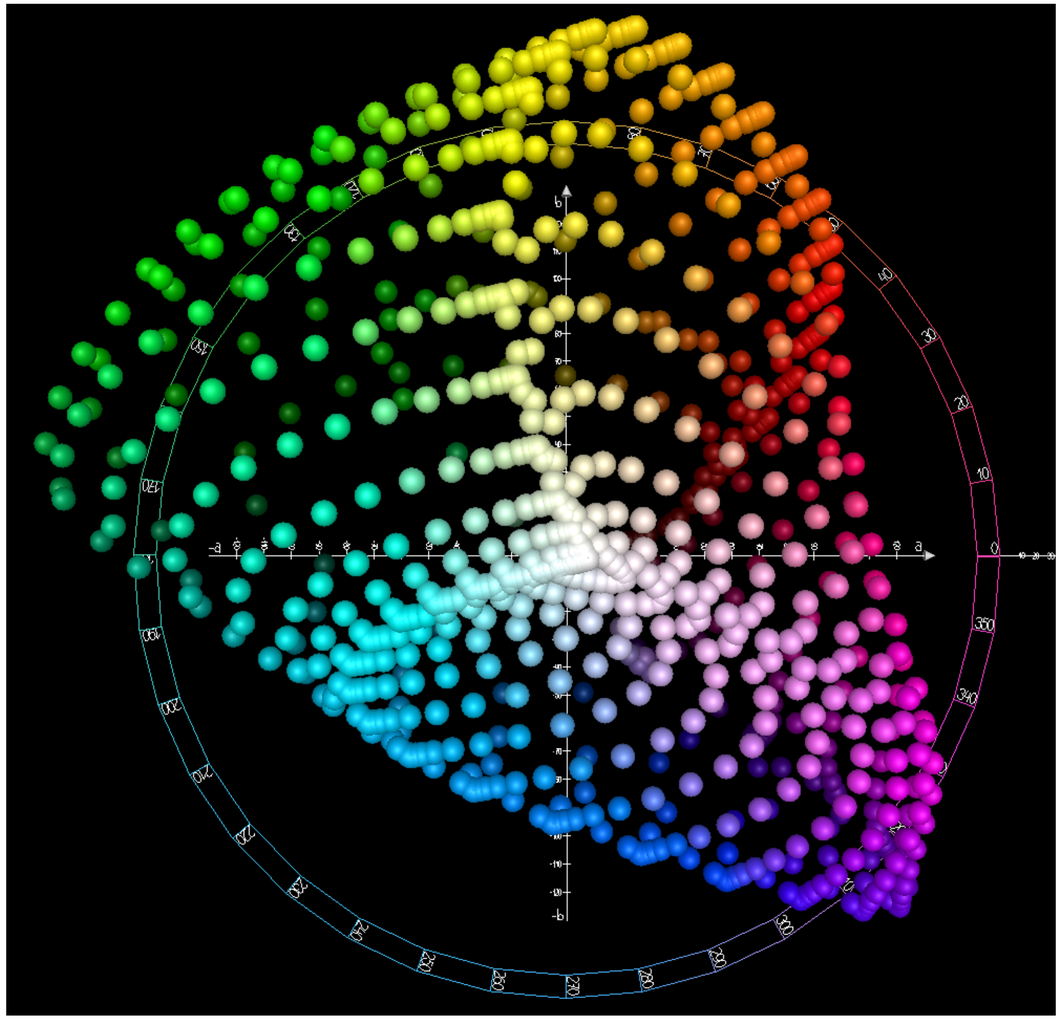
Group and order your swatches by **Saturation**.

- Hue
- Saturation
- **Luminance —** The amount of white or black in a color

PAUSE:

Group and order your swatches by **Luminance**.

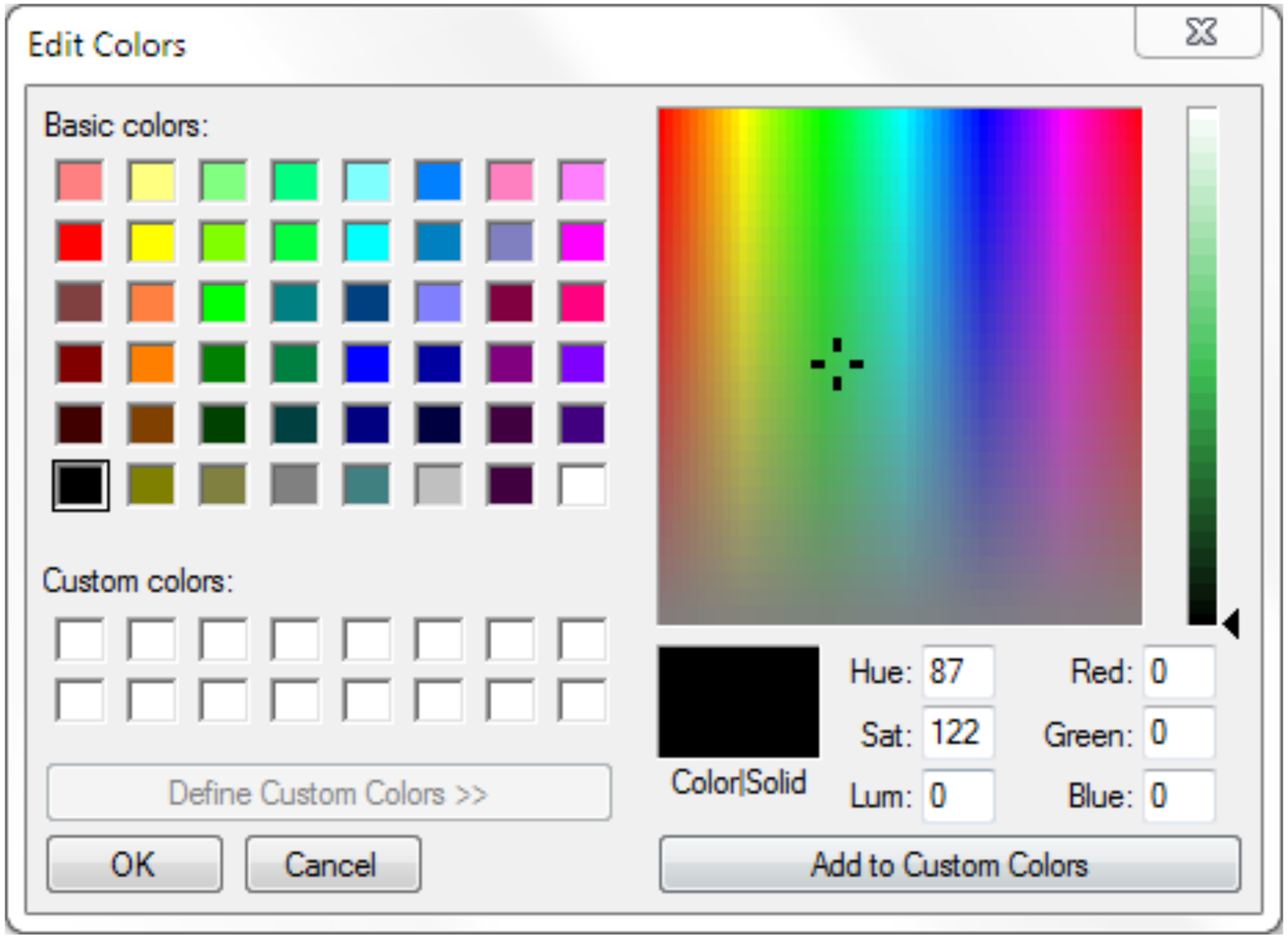
CIELAB (or L*a*b*: Lightness, Green–Red and Blue–Yellow)†

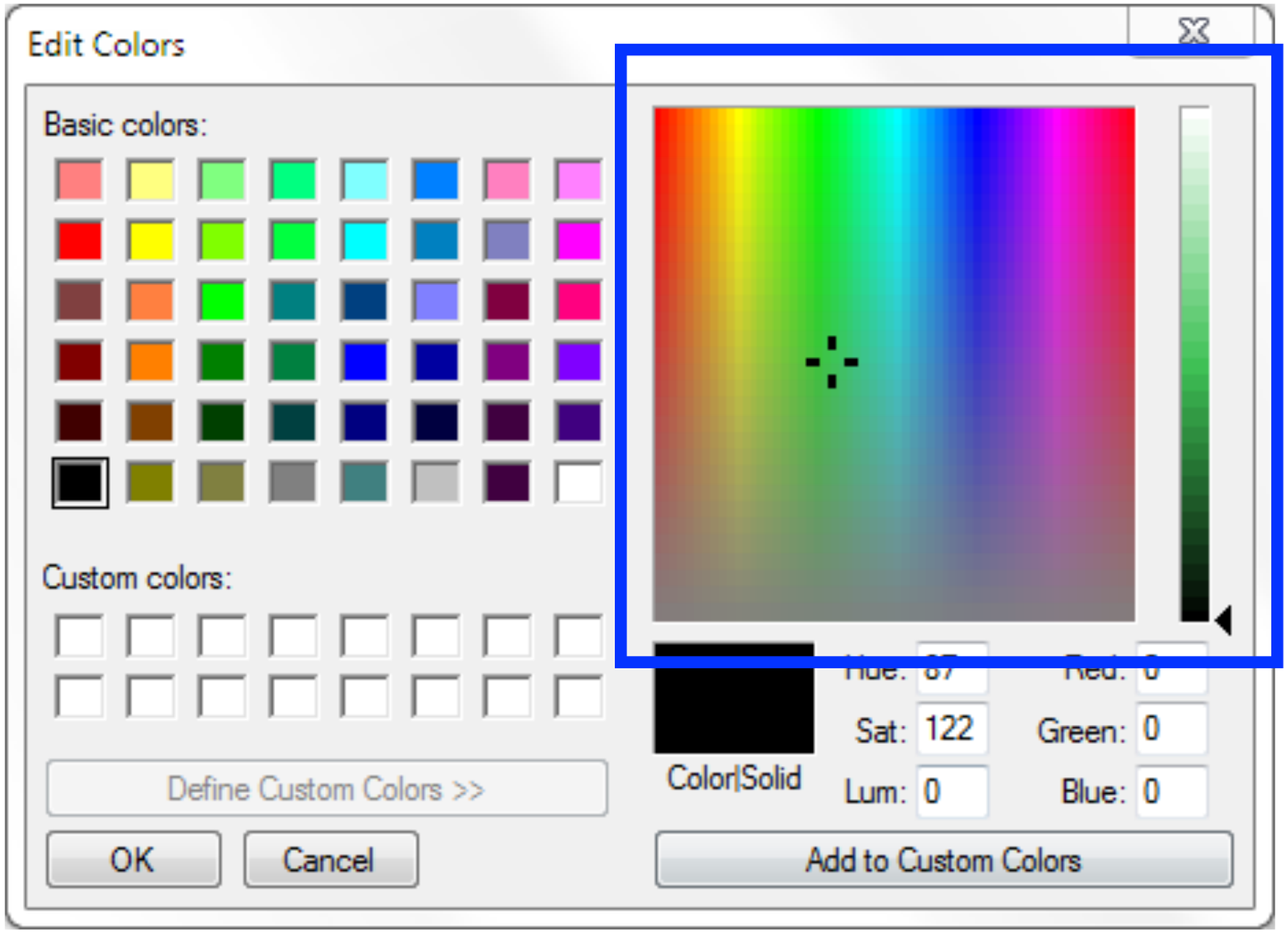


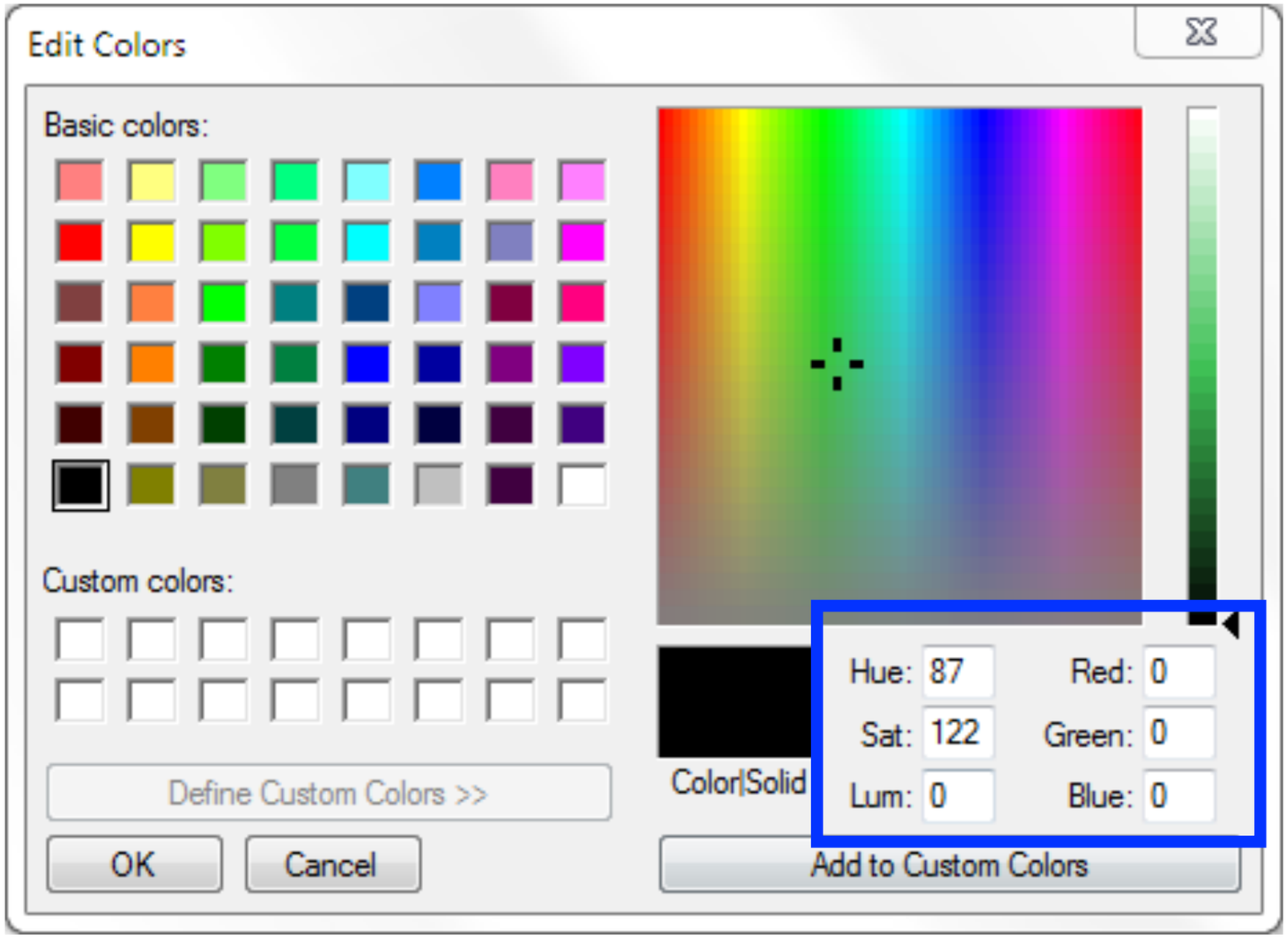
†**International Commission on Illumination**

- Concerned with being **perceptually uniform** to human color vision
- Numerical change = visually perceived change

Typical color controls surface RGB and HSL (and sometimes CMYK) to the user.


Edit Colors
Basic colors:
Custom colors:
Define Custom Colors >>
OK
Cancel
Color|Solid
Hue: 87
Sat: 122
Lum: 0
Red: 0
Green: 0
Blue: 0
Add to Custom Colors


Edit Colors
Basic colors:
Custom colors:
Define Custom Colors >>
OK
Cancel
Color|Solid
Sat: 122
Lum: 0
Green: 0
Blue: 0
Add to Custom Colors


Edit Colors
Basic colors:
Custom colors:
Define Custom Colors >>
OK
Cancel
Color|Solid
Add to Custom Colors
Hue: 87
Sat: 122
Lum: 0
Red: 0
Green: 0
Blue: 0

For the purposes of visualization, we are more concerned with **how color is perceived** than how it is formed…

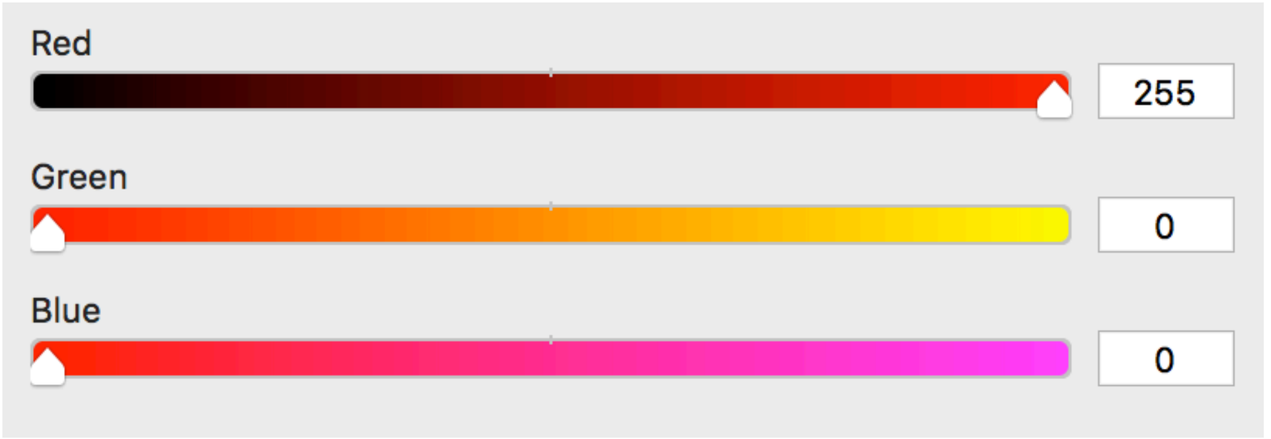
…which means that **using tools that incorporate HSL** (or CIELAB!) will produce the most visually successful results.

Color **syntax**

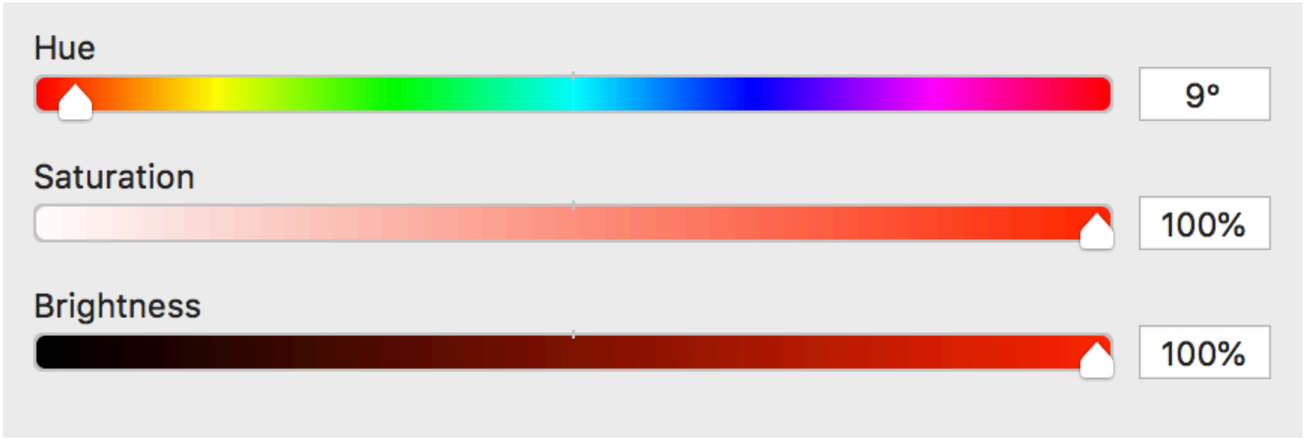
When working with color in a **digital context**, most color pickers allow you to specify values in RGB, CMYK, or HSL/HSB (+ more!).

This color may be specified in any of **these** ways.

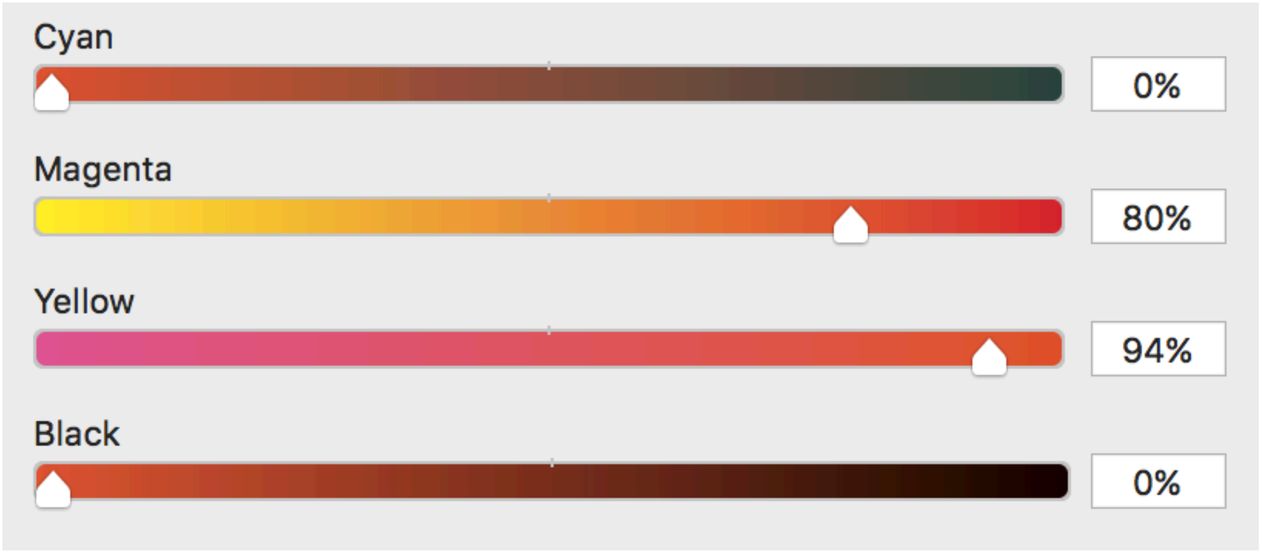
- RGB(255, 0, 0)
- HSL(9, 100, 100)
- CMYK(0, 80, 94, 0)



- **RGB(255, 0, 0)**
- HSL(9, 100, 100)
- CMYK(0, 80, 94, 0)



- RGB(255, 0, 0)
- **HSL(9, 100, 100)**
- CMYK(0, 80, 94, 0)



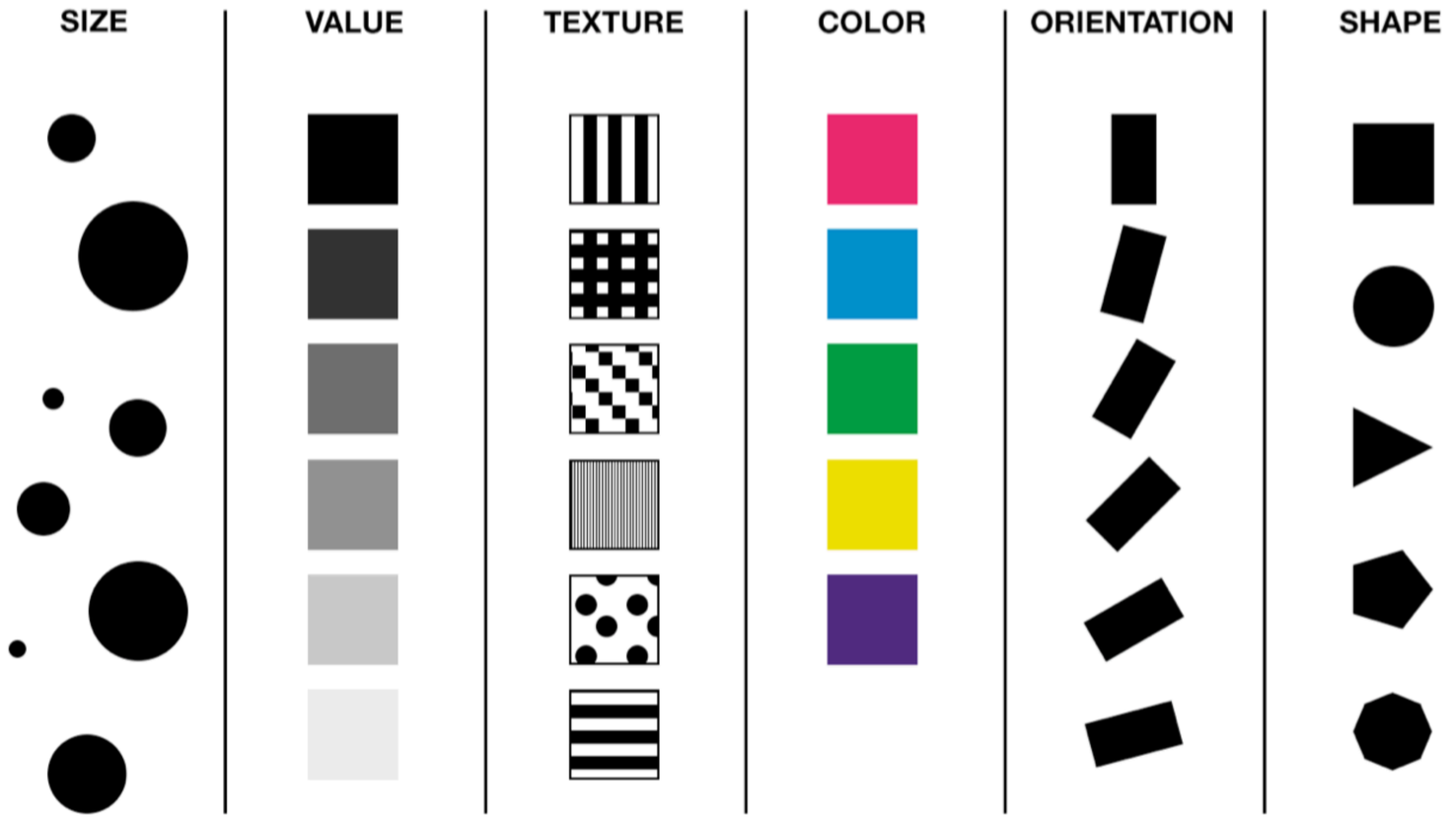
- RGB(255, 0, 0)
- HSL(9, 100, 100)
- **CMYK(0, 80, 94, 0)**

1. Color **Perception** & **Representation**
2. **Applying** Color
3. Practical **Tips**
4. Homework Review

Recap: **Visual Variables**

Jacques Bertin

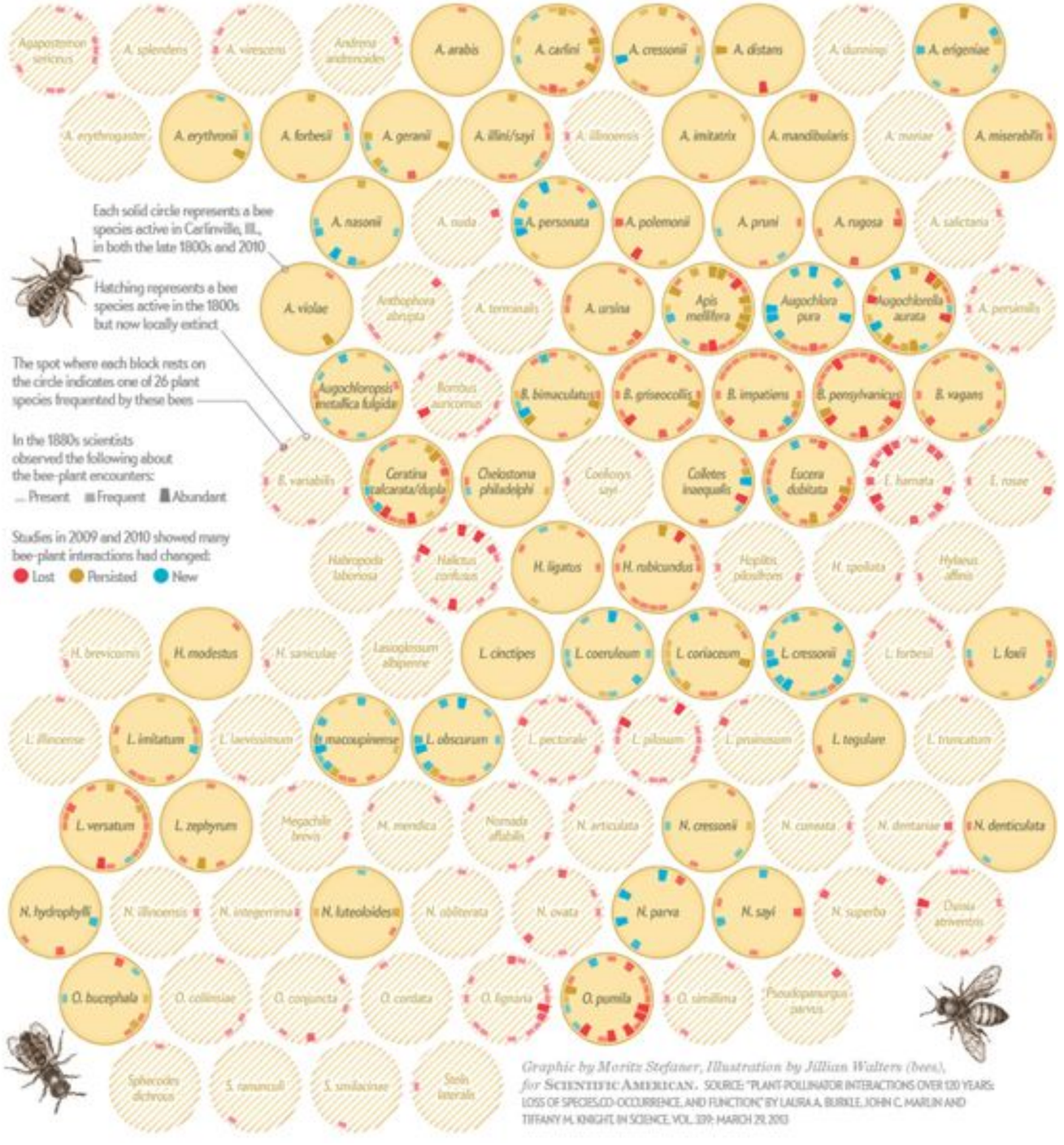
The Semiology of Graphics (1967)


SIZE
VALUE
TEXTURE
COLOR
ORIENTATION
SHAPE

The process of mapping data to visual variables is called **visual mapping**.

Selecting visual variables to represent different aspects of the same information can greatly influence the **perception and understanding** of the presented information.

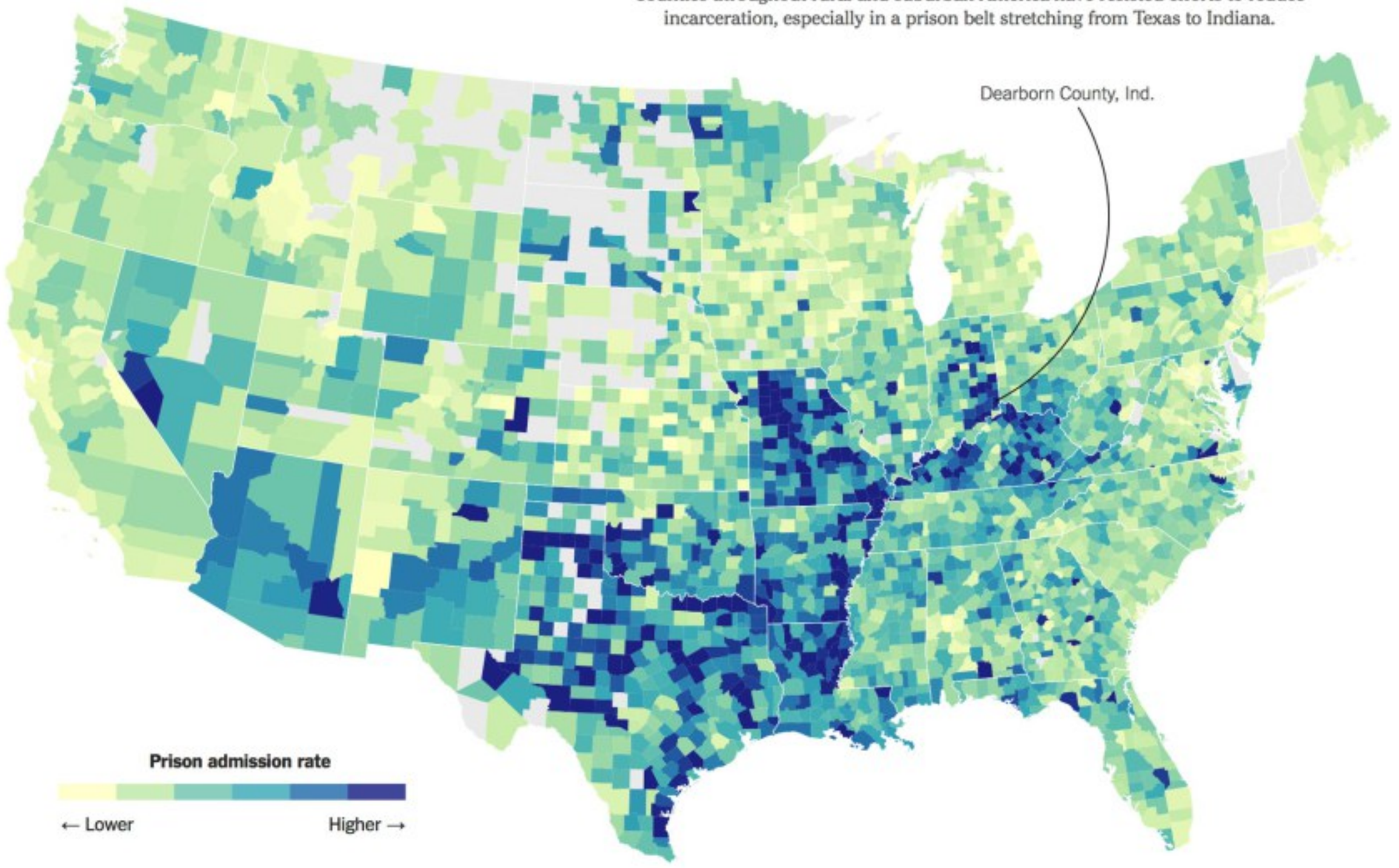
Color is an especially useful visual variable, because it is flexible enough to represent both qualitative and quantitative data.



Graphic by Moritz Stefaner, Illustration by Jillian Walters (bees), for SCIENTIFIC AMERICAN. SOURCE: "PLANT-POLLINATOR INTERACTIONS OVER 120 YEARS: LOSS OF SPECIES, CO-OCCURRENCE, AND FUNCTION" BY LAURA A. BURKLE, JOHN C. MARLIN AND TIFFANY M. KNIGHT, IN SCIENCE, VOL. 339; MARCH 29, 2013

Where Americans Are Sent to Prison Most

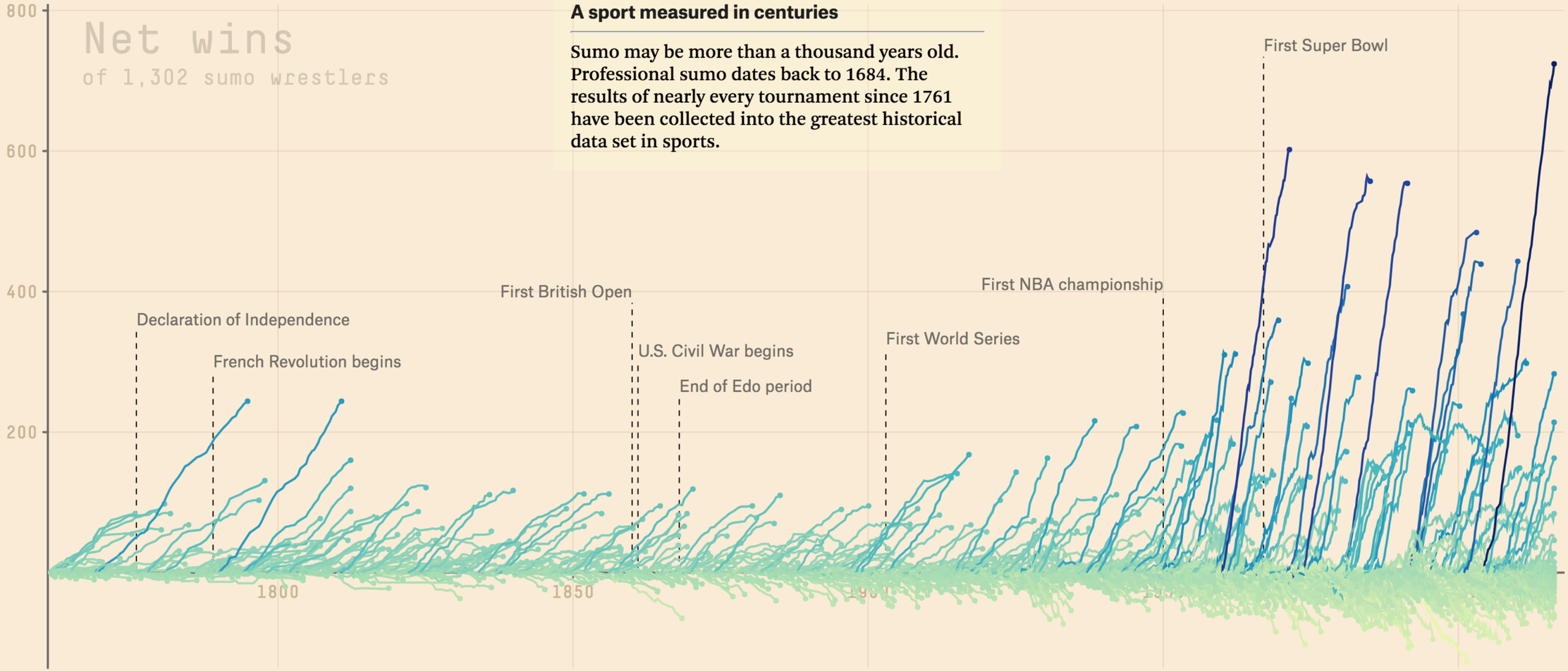
Counties throughout rural and suburban America have resisted efforts to reduce incarceration, especially in a prison belt stretching from Texas to Indiana.



Analysis of the National Corrections Reporting Program by The New York Times and John Pfaff, Fordham University

A History Of Sumo

By Matthew Conlen



A sport measured in centuries

Sumo may be more than a thousand years old. Professional sumo dates back to 1684. The results of nearly every tournament since 1761 have been collected into the greatest historical data set in sports.

Color and **Data**

Not all channels work for all types of data. Think about what best corresponds to what you're working with.

Qualitative

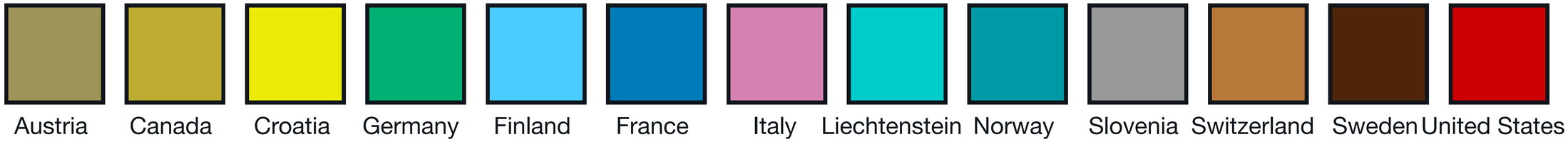
- Nominal
- Ordinal

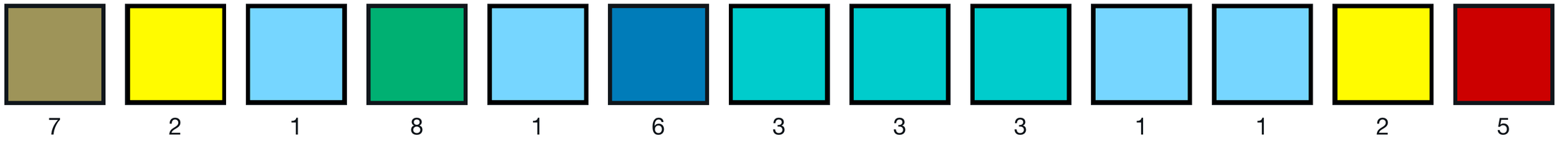
Quantitative

- Interval
- Ratio

Which is more effective?

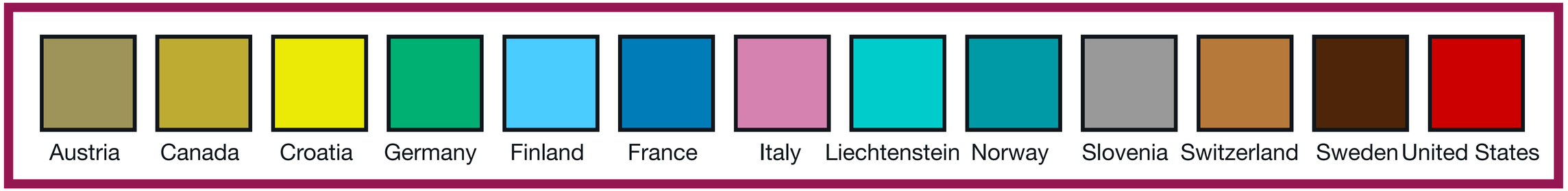
Channel: ***Color***



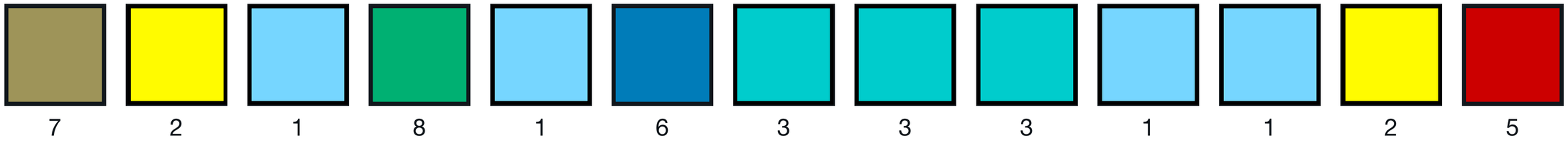


Which is more effective?

Channel: ***Color***

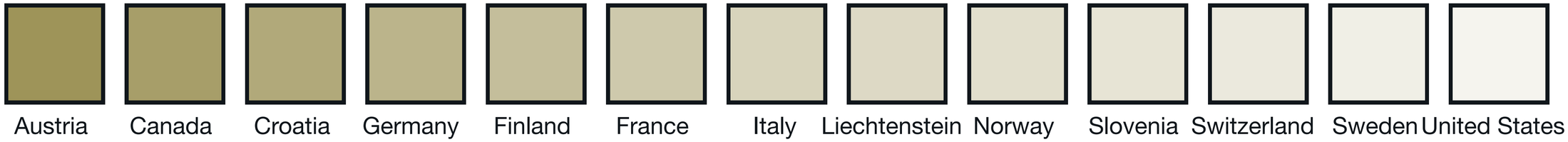


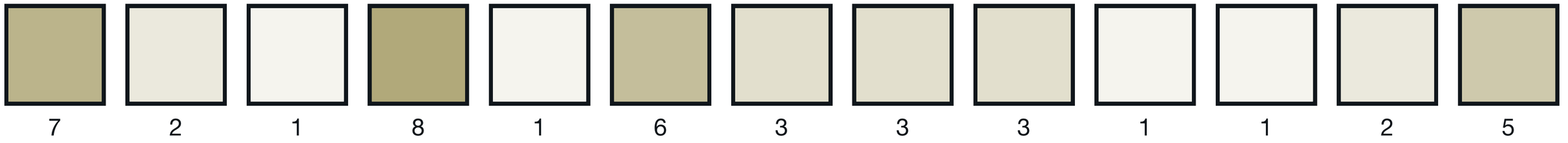
Qualitative data *(nominal)*



Which is more effective?

Channel: ***Value (Luminance, Saturation)***

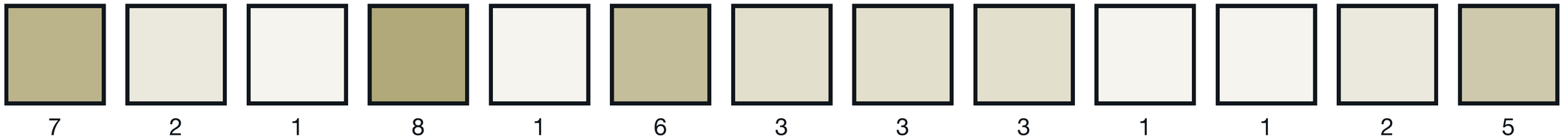




Which is more effective?

Channel: ***Value (Luminance, Saturation)***

Qualitative data *(ratio)*



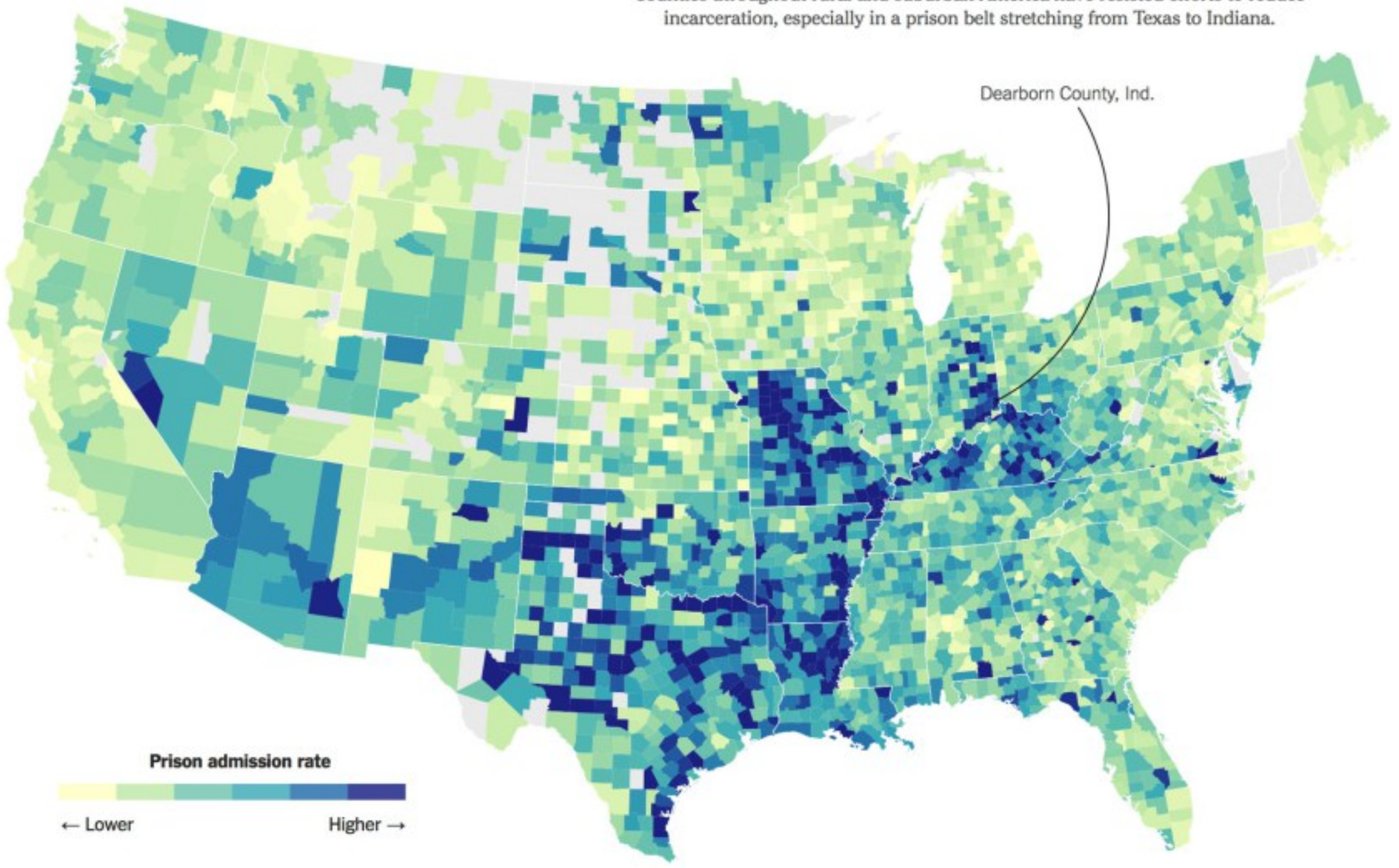
Working with color becomes challenging when trying to create **scales that appear to have even increments** *(perceptual uniformity)*.

Why bother with perceptual uniformity?

- Faithful representation of data
- Easy to compare and compute values

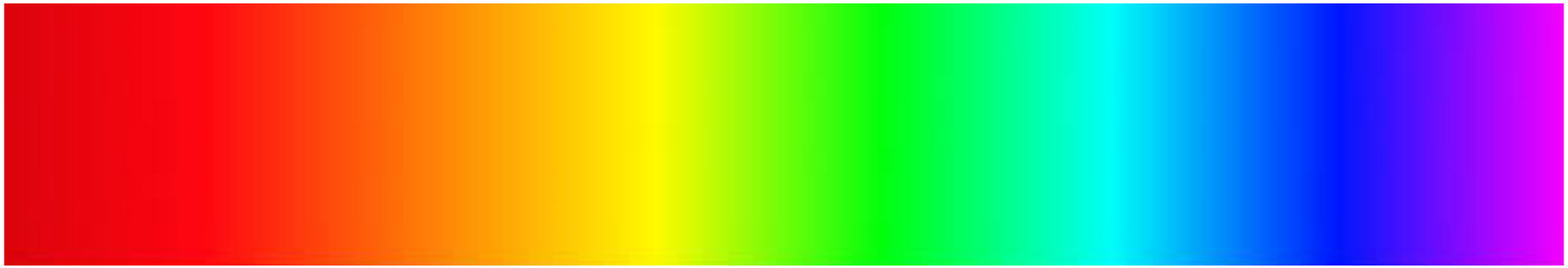
Where Americans Are Sent to Prison Most

Counties throughout rural and suburban America have resisted efforts to reduce incarceration, especially in a prison belt stretching from Texas to Indiana.

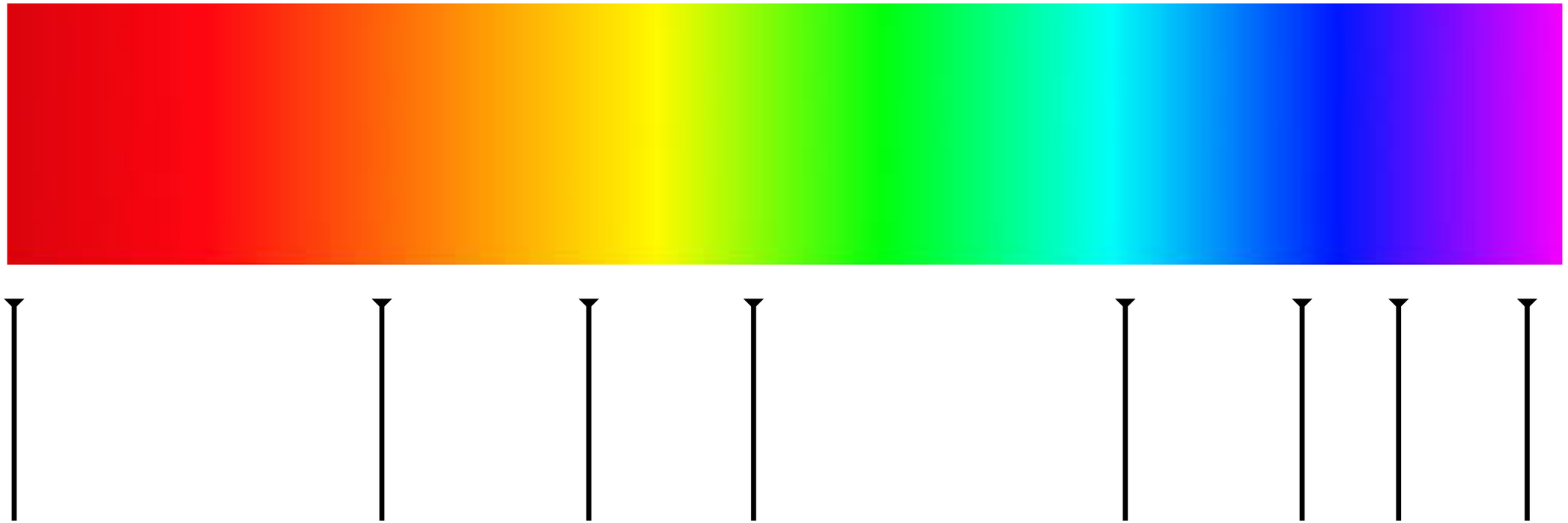


Analysis of the National Corrections Reporting Program by The New York Times and John Pfaff, Fordham University

Why not use a rainbow?

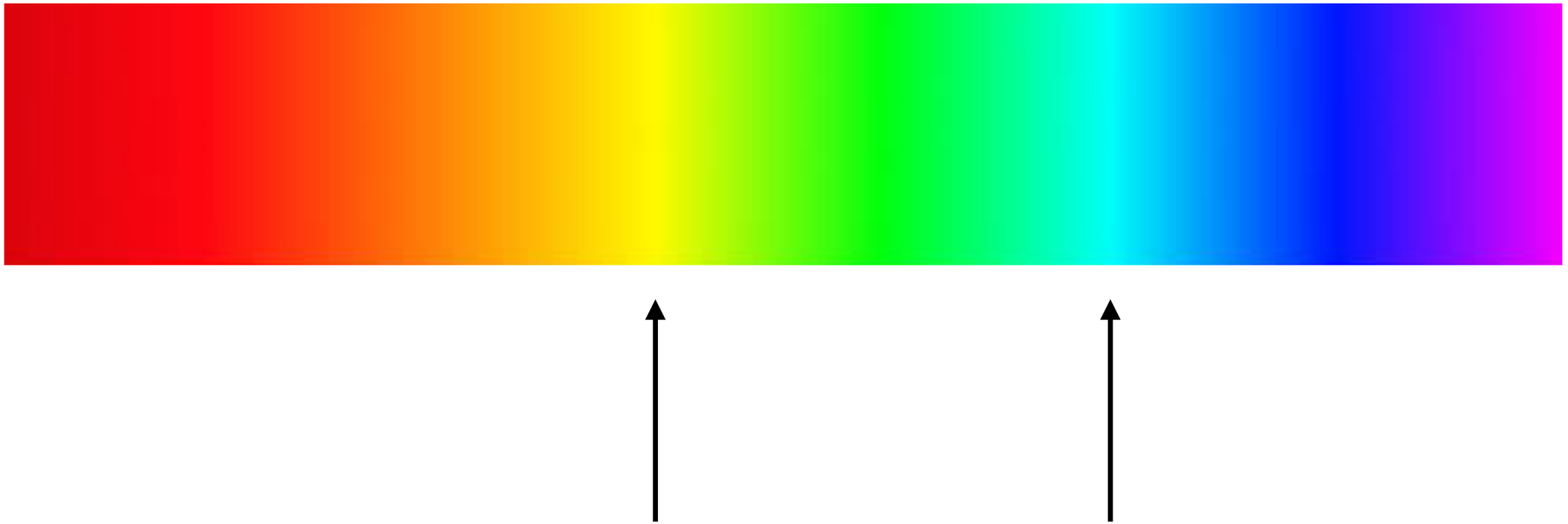


Why not use a rainbow?



Weirdly spaced differences

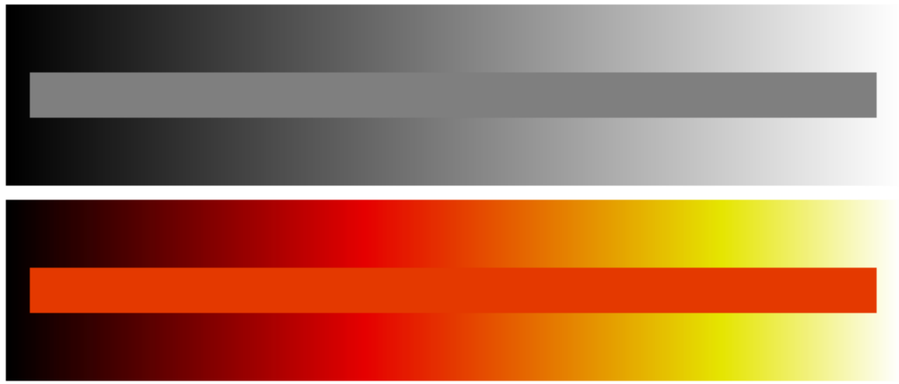
Why not use a rainbow?



Mach color bands

Even though rainbow scales are the default in many visualization applications, they produce undesirable **perceptual effects that distort the underlying data**.

Why not use a **grayscale** (luminance only)?



Pixel appearance is affected
by surrounding pixels

Detail lost

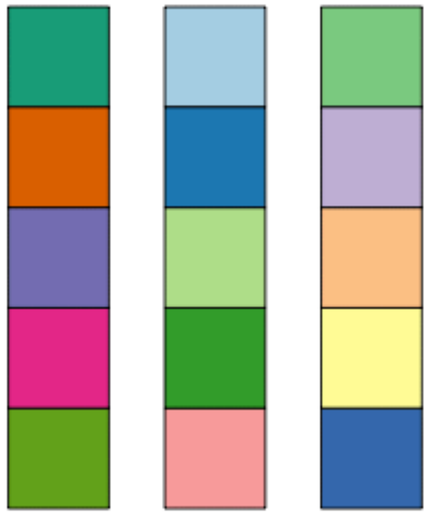
Why not use an **isoluminant map** (chromatic shifts)?

- Humans are more sensitive to changes in luminance
- Limits the number of colors represented, which can lead to lower-fidelity graphics

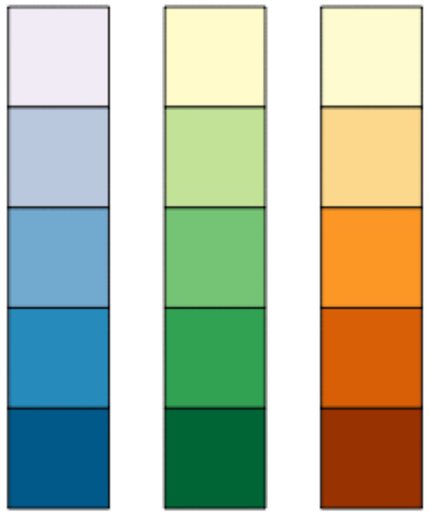
The right color scale **controls changes in hue** with shifts in saturation and luminance.

Many researchers have worked on this problem, including **Cynthia Brewer**…

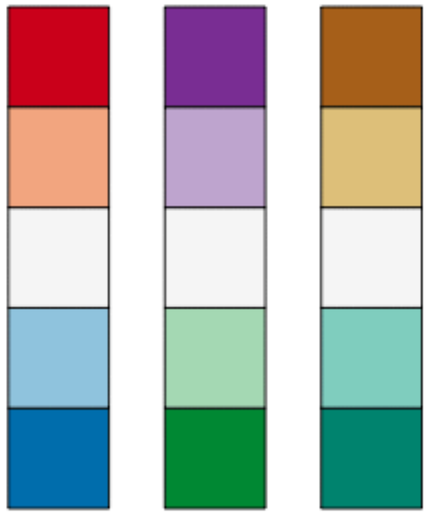
Cynthia Brewer — **Color Scales**



Qualitative

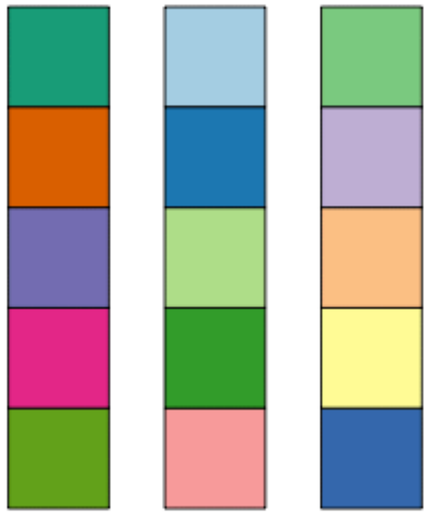


Sequential

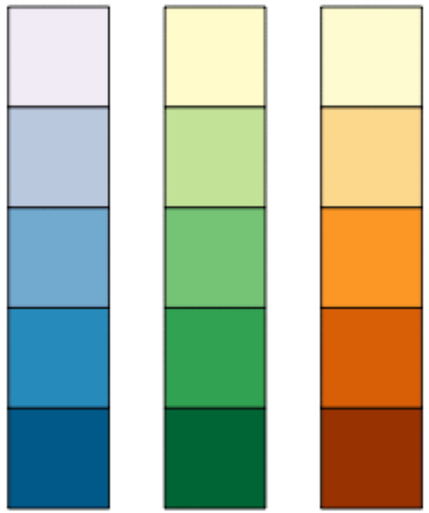


Diverging

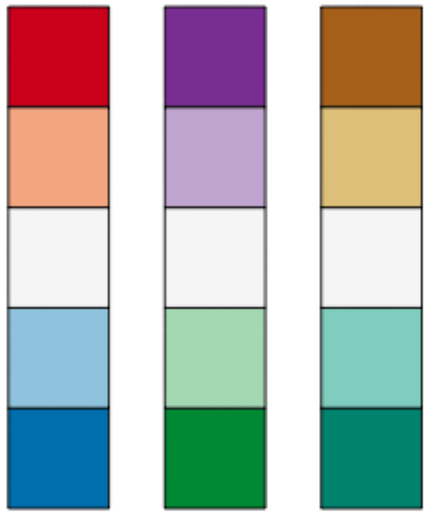
Cynthia Brewer — **Color Scales**



Qualitative
(Nominal)



Sequential
(Ordinal, interval, ratio)



Diverging
(Ordinal, interval, ratio)

…who make their work **easily accessible to all**.

→ *LINK:* ***http://colorbrewer2.org/***
→ *LINK:* ***https://gka.github.io/palettes/***

Use your intuition. Some combinations of variables are more readable than others.

1. Color **Perception** & **Representation**
2. **Applying** Color
3. Practical **Tips**
4. Homework Review

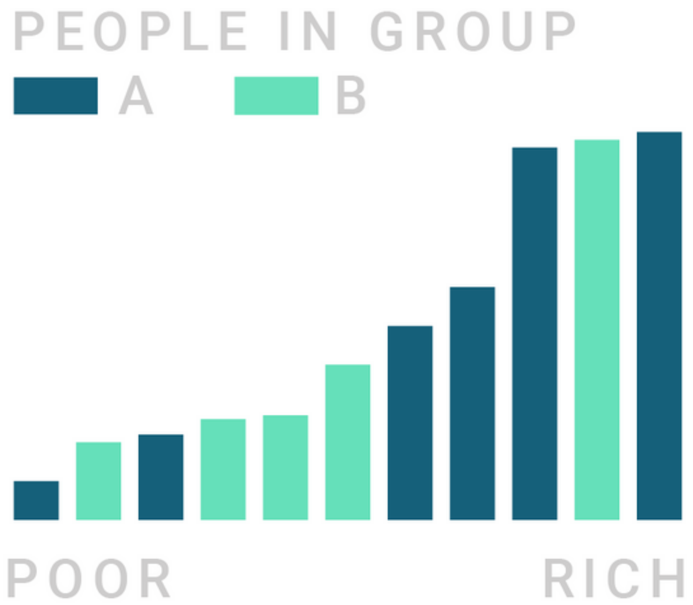
Lisa Charlotte Rost for ***Datawrapper*** put together a handy guide for using color in visualization.

→ *LINK:* **https://blog.datawrapper.de/colors/**

Consider alternatives to **gradients** when encoding important values.

NOT IDEAL

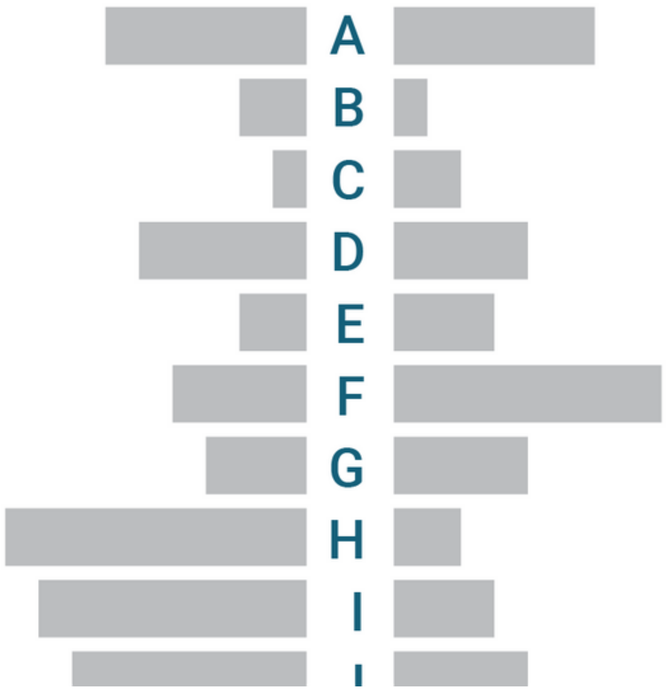
BETTER



If you need more than **seven colors** in a chart, consider using another chart type or to group categories together.

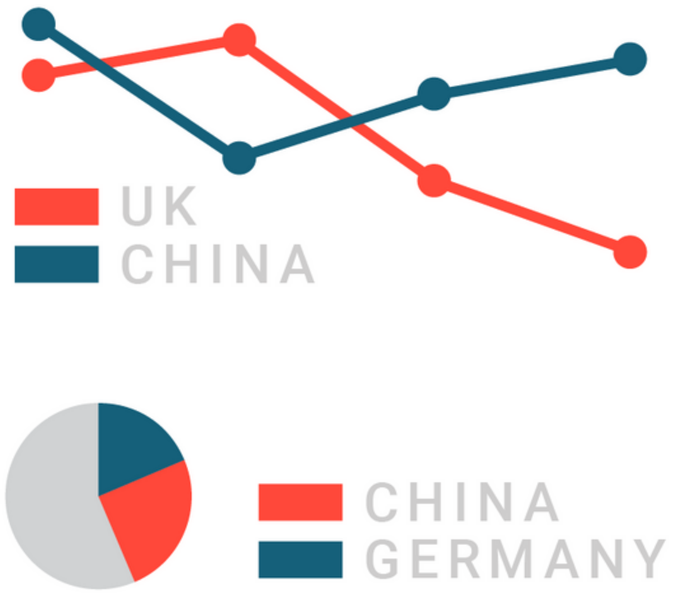
NOT IDEAL

BETTER

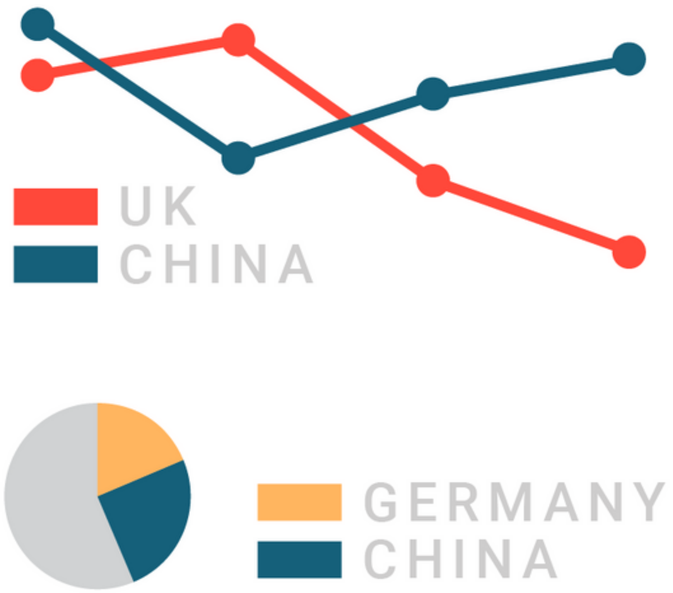


Consider using the **same color** for the **same variables**.

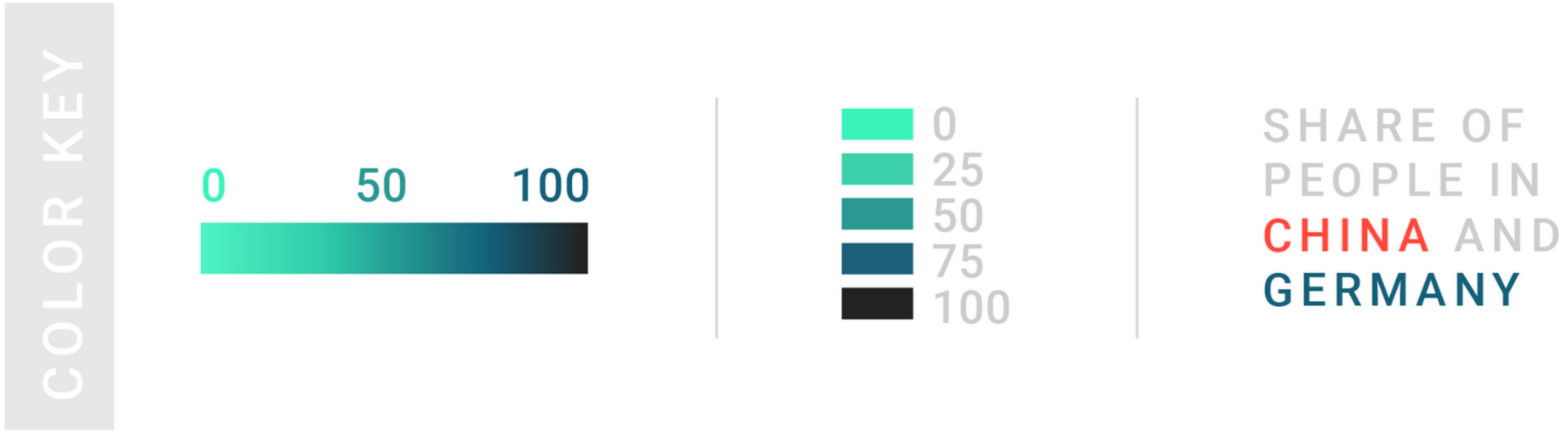
NOT IDEAL



BETTER

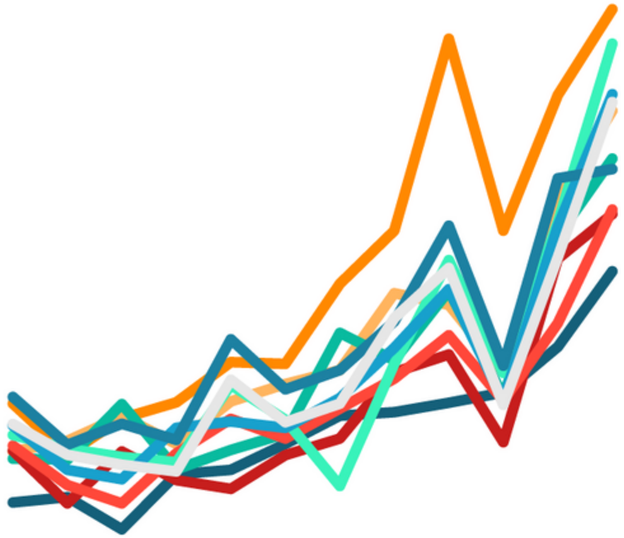


Make sure to **explain to readers** what your colors encode.

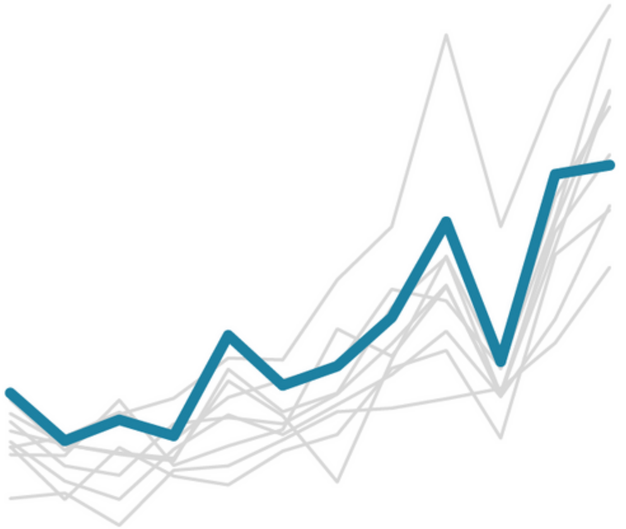


Consider the color grey as the **most important color** in visualization.

NOT IDEAL



BETTER

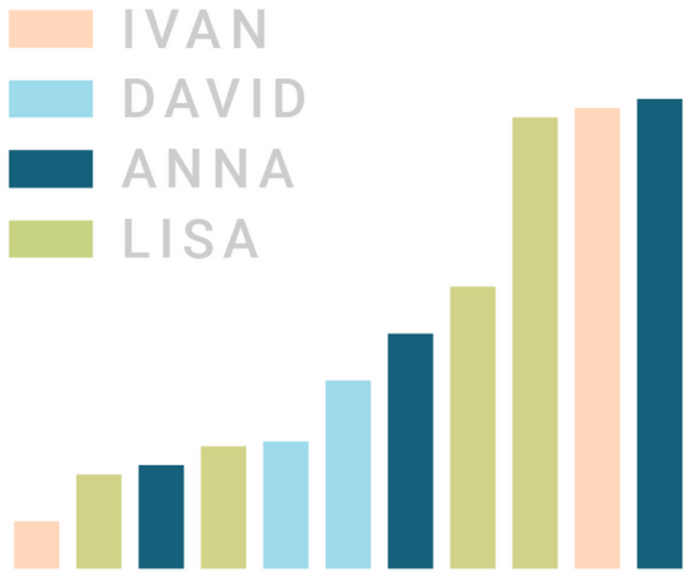


Make sure your **contrasts** are high enough.

Don’t use a **gradient color palette** for categories and the other way around.

NOT IDEAL

BETTER



1. Color **Perception** & **Representation**
2. **Applying** Color
3. Practical **Tips**
4. Homework Review