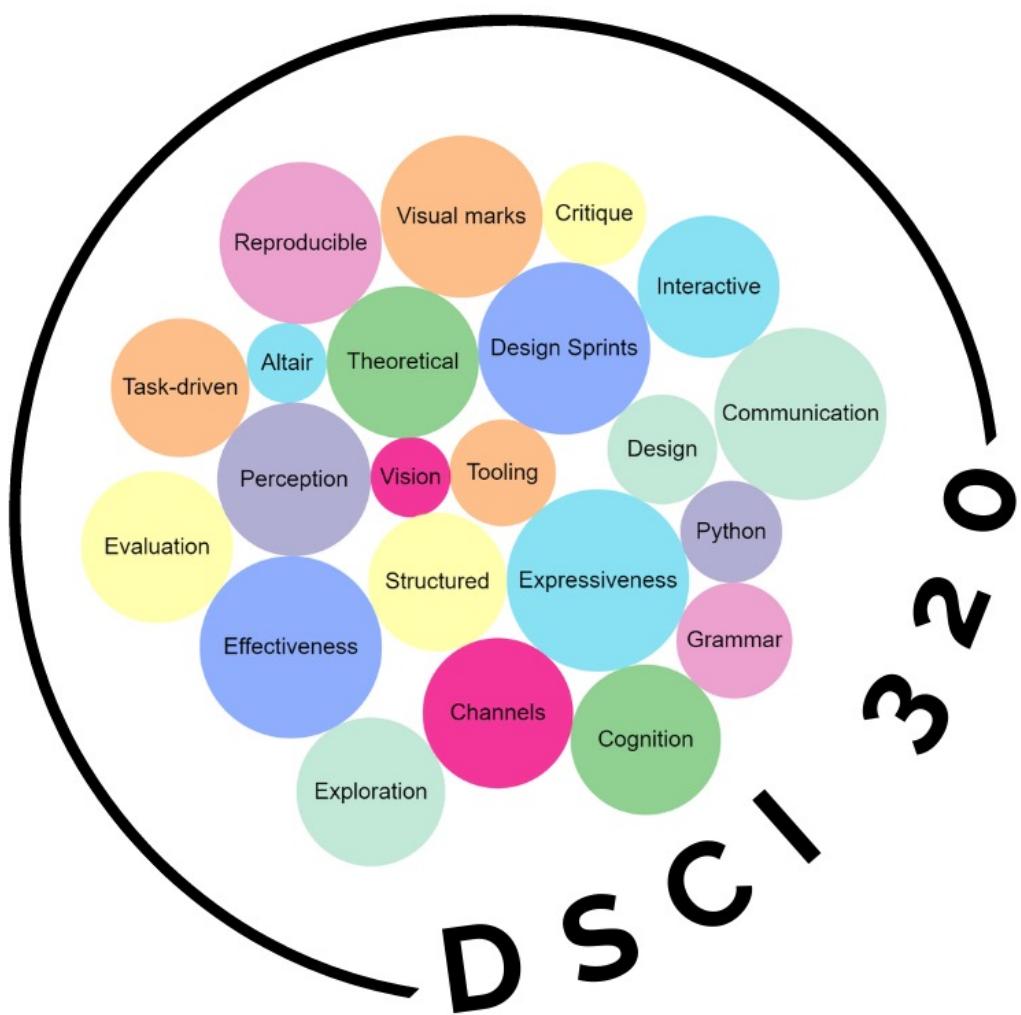


# Visualization for Data Science

## Gestalt Principles

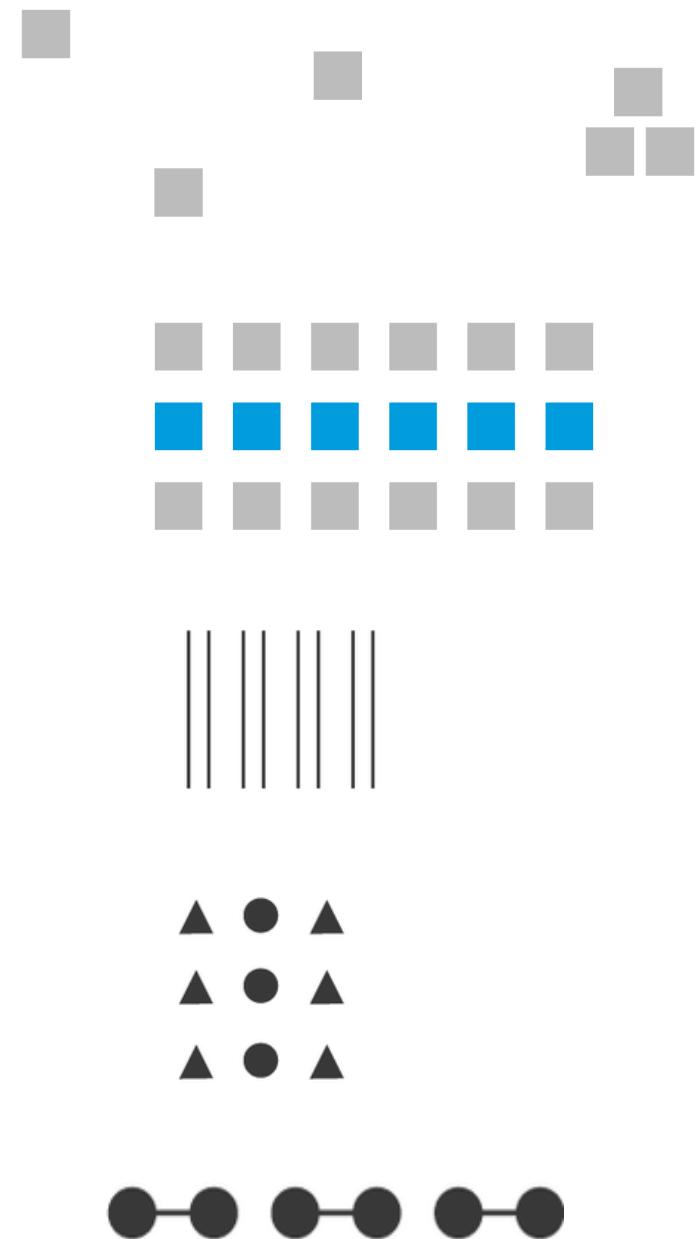
### Task Abstraction



[Gestalt principles \(part 1\)](#). Bang Wong. Nature Methods  
[Gestalt principles \(part 2\)](#). Bang Wong. Nature Methods

# CQ: Which Gestalt Principle(s) are used in the image?

- A. Continuity
- B. Closed Region
- C. Similarity
- D. Proximity
- E. Connection/Connectedness



# Gestalt principles

- Similarity: people tend to see things that physically resemble each other as part of the same object
- Proximity: how close elements are to one another. Similar things should be close to each other
- Connection: grouping effect; we perceive elements as connected to each other thanks to colors, lines, frames, or other shapes
- ~~Enclosure/~~Common Region : we group elements that are in the same closed region
- Continuity: objects that create a continuous pattern or are seen as being connected appear to be grouped together
- Closure: our eyes tend to add any missing pieces of a familiar shape
- Symmetry: elements that are symmetrical tend to be perceived as a unified group
- Figure & Ground: Your brain distinguishes the foreground and the background
- Common Fate: people will group together things that point to or are moving in the same direction

# Gestalt Effects for “Emergence”

“Our visual system attempts to structure what we see into patterns to make sense of information” Bang Wong

Objects emerge and we assign meaning to them, through

- Visually interpolation
- Visually completion

We perceive things that may not exist

# Continuity/Good Continuation

Due to visual interpolation, our internal visual representation tends to be smooth and continuous.

Objects that form a continuous pattern or seem connected are perceived as grouped together.

Our gaze naturally transitions from one object to another.

# Continuity

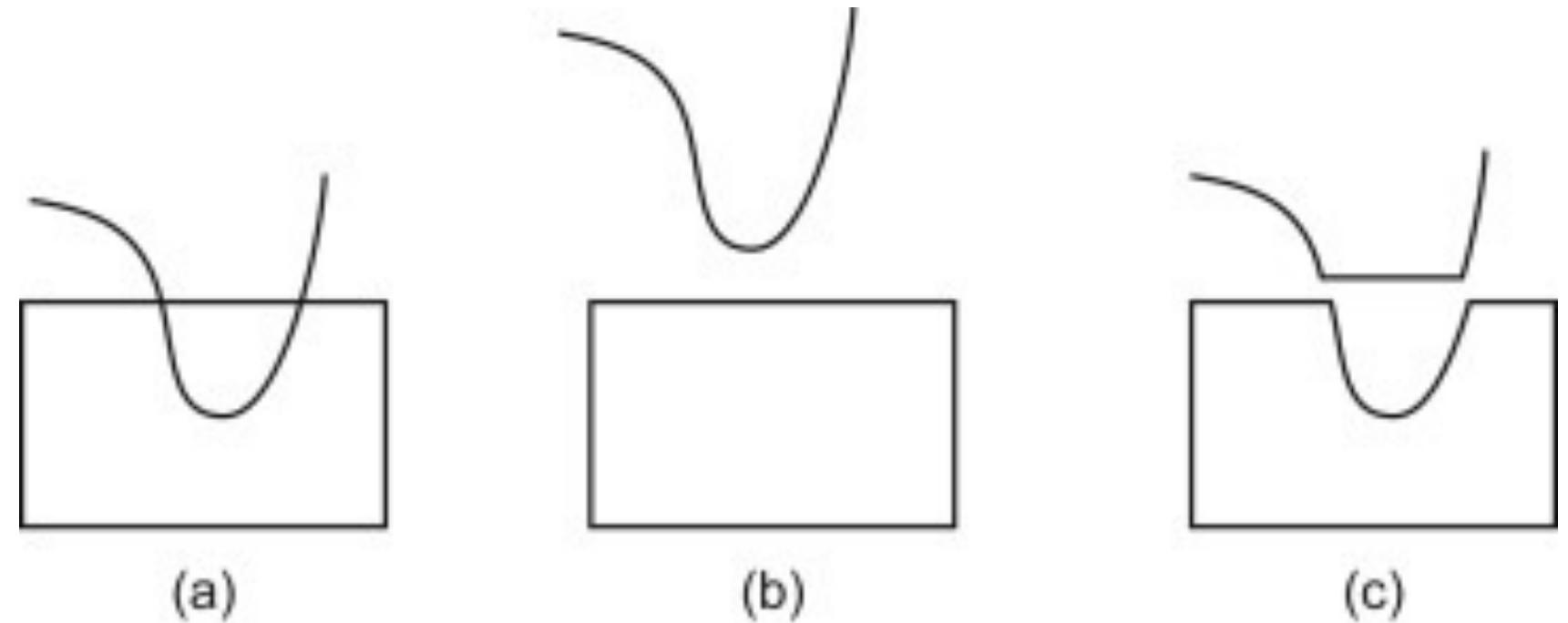
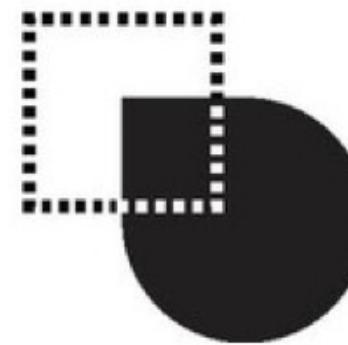
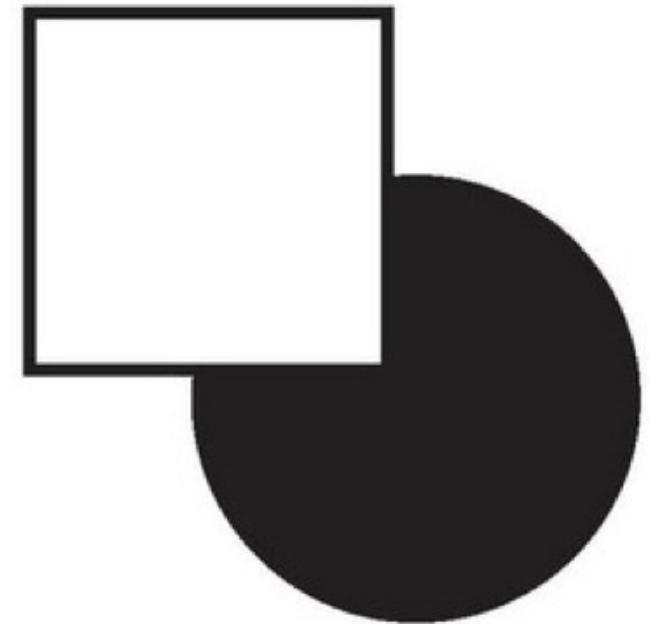


Figure 6.6. Information Visualization. Perception for Design Colin Ware

# Closure

Because of visual completion we have a strong tendency to seen shapes as continuous, our eyes (remember the 3 stages of visual processing) tends to add any missing pieces and forms familiar shapes

A closed contour tends to be seen as an object.



# Closure / Completion



# CQ: Which Gestalt Principle(s) are used in the image?

- A. Figure-Ground
- B. Closure
- C. Continuity
- D. Proximity
- E. Connectedness



# Symmetry

Elements that are symmetrical tend to be perceived as a unified group

*“The pairs of lines shown in Fig. 6.9(b) are perceived more strongly as forming a visual whole than the lines with parallel symmetry (Fig. 6.9a). Also, when edges instead of lines are used, symmetry is more difficult to perceive if the polarity is reversed on the edges (Fig. 6.9(c)).”*

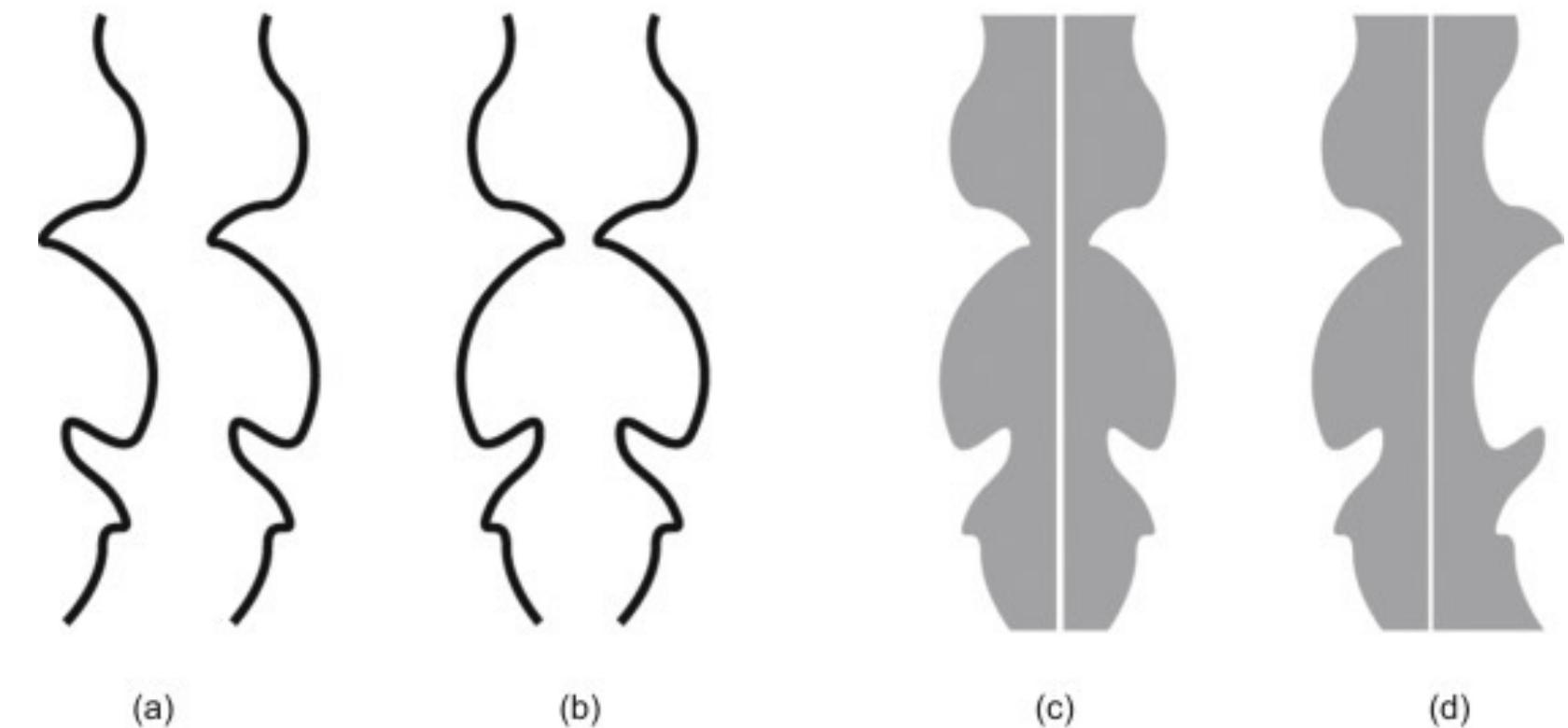
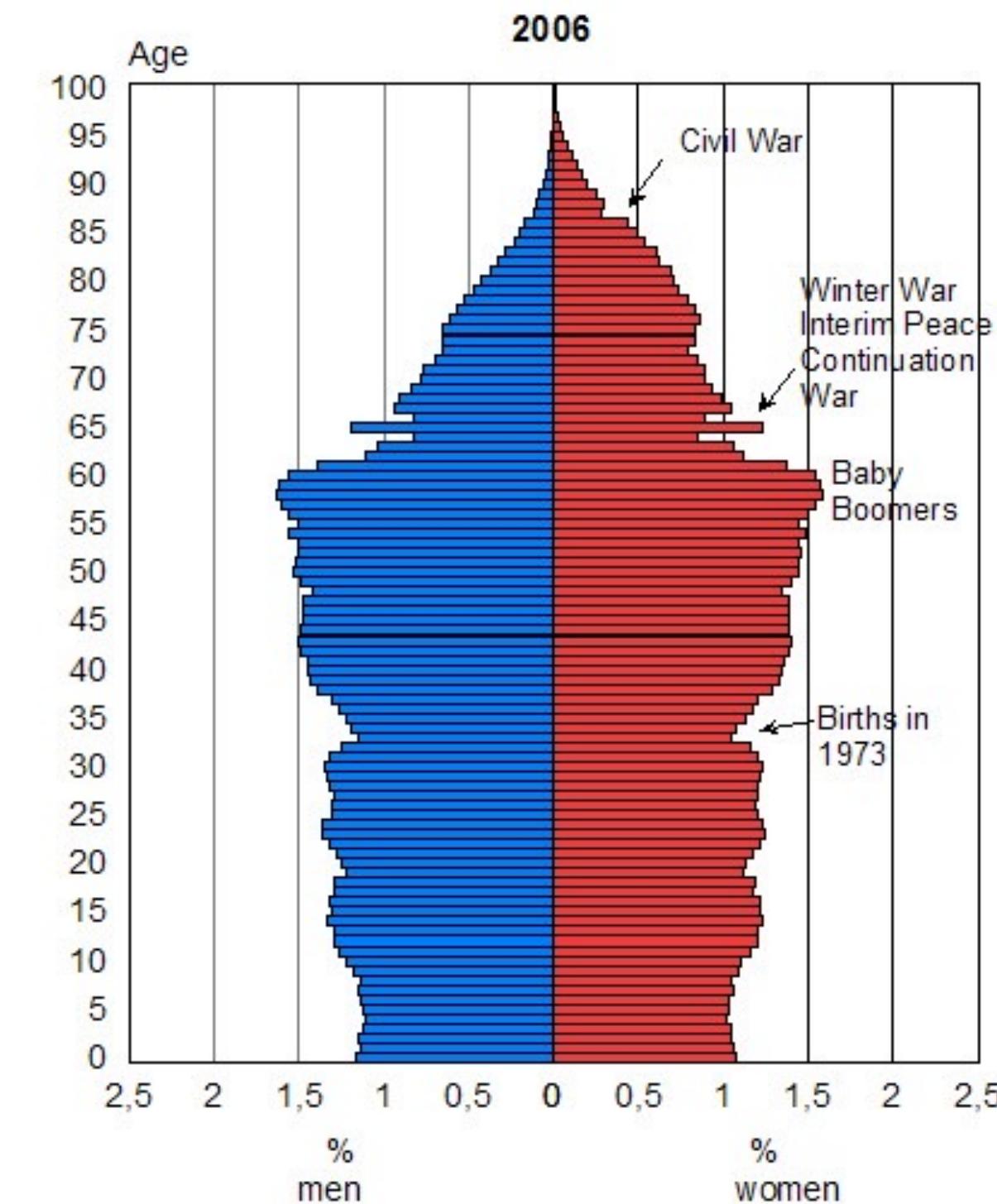
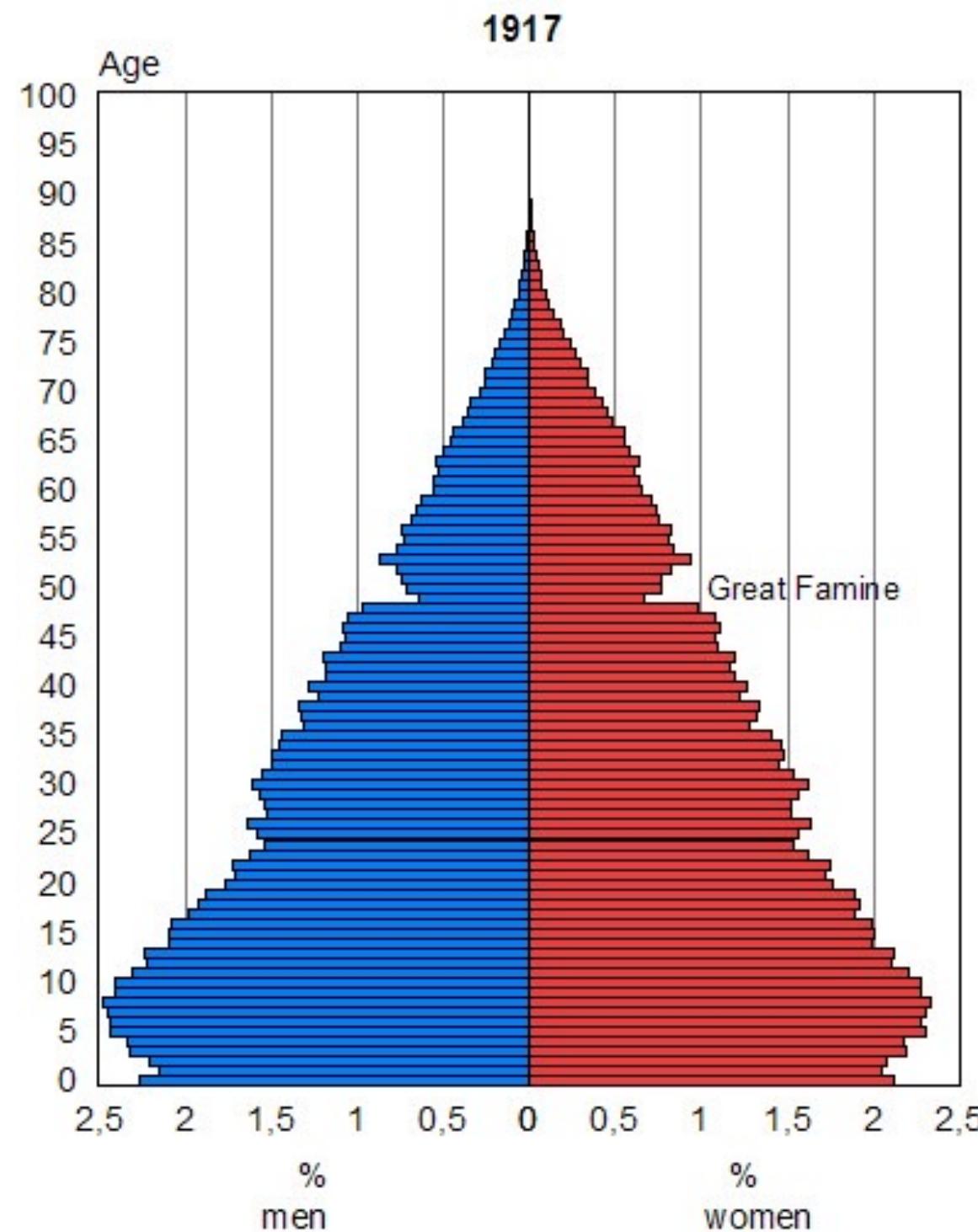


Figure 6.9. Information Visualization. Perception for Design Colin Ware

# Symmetry Example: Population Pyramid



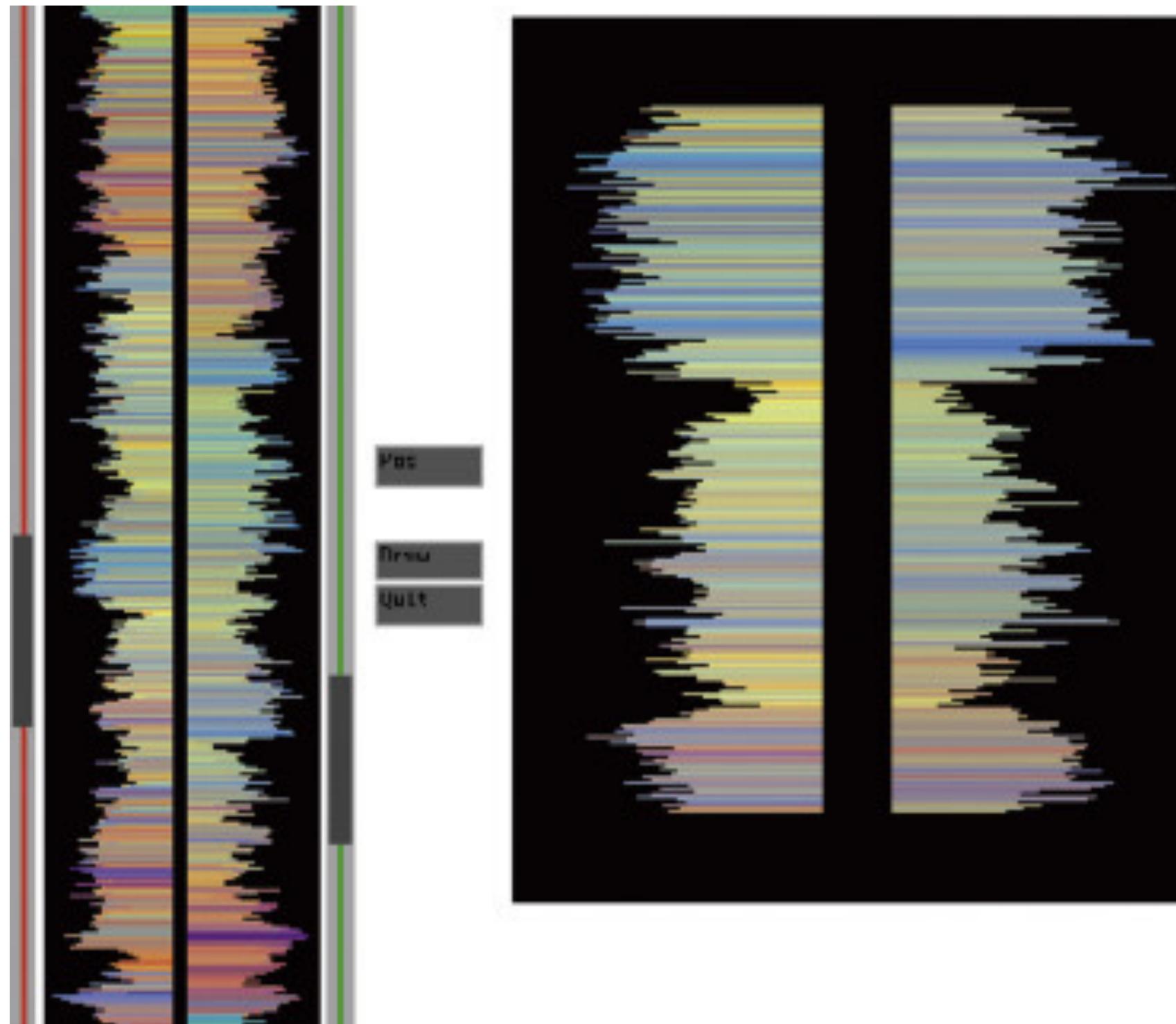


Figure 6.10. Information Visualization. Perception for Design Colin Ware

# Figure / Ground

Your brain distinguishes the foreground and the background



CQ: Which Gestalt Principle(s) are used in the image

- A. Figure-Ground
- B. Closure
- C. Similarity
- D. Proximity
- E. Connection



# CQ: Which Gestalt Principle(s) are used in the image

- A. Figure-Ground
- B. Closure
- C. Similarity
- D. Good Continuation
- E. Connectedness



CQ: Which Gestalt Principle(s) are used in the image

- A. Figure-Ground
- B. Closure
- C. Similarity
- D. Proximity
- E. Connection



When observing a series of dots, if the dots are arranged in a way that they form a recognizable shape, like a circle or a square, individuals tend to perceive the entire shape instead of the individual dots. Which Gestalt principle does this scenario illustrate?

- A. Figure-Ground
- B. Closure
- C. Similarity
- D. Proximity
- E. Connection

CQ

Jake is creating a bar chart to display sales data from various regions. He decides to use different shades of the same color to represent sales from each region. Users can quickly identify and group bars corresponding to each region. Which Gestalt principle is Jake leveraging in this scenario?

- A. Closure
- B. Continuity
- C. Proximity
- D. Similarity
- E. Figure-Ground

# CQ: Which Gestalt Principle is Used in the Following Image

- A. Figure-Ground
- B. Closure
- C. Similarity
- D. Proximity
- E. Connection



CQ: Which Gestalt Principle(s) are used in the image?

- A. Figure-Ground
- B. Closure
- C. Similarity
- D. Proximity
- E. Connection



# Project Update

Scope will be reduced for the following reasons

- Pedagogical efficiency
- LLM-free realistic project scope
- Assessment balance

The image on the next page is the most up-to-date.

It supersedes the text on the Overview page on the website

**All requirements listed below are per group member !**  
**Each analytic question should be answered by one "view" !**

	C	B	A	A+
Report	<ul style="list-style-type: none"> <li>• Markdown file</li> <li>• Contribution summary</li> </ul>	<ul style="list-style-type: none"> <li>• Markdown file</li> <li>• Contribution summary</li> <li>• Significant insights &amp; reflections</li> <li>• Reasonable workload distribution</li> </ul>	<ul style="list-style-type: none"> <li>• Markdown file</li> <li>• Contribution summary</li> <li>• Deep &amp; thoughtful insights &amp; reflections</li> <li>• Equitable workload distribution</li> </ul>	<ul style="list-style-type: none"> <li>• Publish viz on a website (dashboard)</li> <li>• Contribution summary</li> <li>• Deep &amp; thoughtful insights &amp; reflections</li> <li>• Equitable workload distribution</li> </ul>
Visualizations	<ul style="list-style-type: none"> <li>• 1 view</li> <li>• View has at least 2 altair visualizations (vizzes)</li> </ul>	<ul style="list-style-type: none"> <li>• 1 view</li> <li>• The view must have 3 vizzes</li> </ul>	<ul style="list-style-type: none"> <li>• 1 view (with 3 vizzes)</li> <li>• An additional view (with at least 2 vizzes)</li> <li>• Thoughtful layout in views</li> <li>• At least 1 advanced, high-quality, high-effort visualizations.</li> <li>• &gt;=5 vizzes in total</li> </ul>	<ul style="list-style-type: none"> <li>• 2 views (3 viz on 1, 2 viz on other)</li> <li>• An additional view (with at least 2 vizzes)</li> <li>• An additional advanced viz</li> <li>• &gt;=7 vizzes in total</li> </ul>
Interactivity	<ul style="list-style-type: none"> <li>• 1 meaningful interaction (that is more than zoom, pan, tooltip, hover with no propagation)</li> </ul>	<ul style="list-style-type: none"> <li>• 1 meaningful interaction</li> <li>• 1 indirect manipulation interaction (IMI) that is a UI widget (e.g., drop down, check box, radio button, slider, search box)</li> <li>• 1 advanced direct manipulation interaction (DMI) (e.g., brushing, details on demand)</li> </ul>	<ul style="list-style-type: none"> <li>• 1 meaningful interaction</li> <li>• 2 IMI/UI widgets (unique)</li> <li>• 2 advanced direct manipulation interactions</li> <li>• 1 bi-directional interaction</li> </ul>	<ul style="list-style-type: none"> <li>• 1 meaningful interaction</li> <li>• 3 IMI/UI widgets (unique)</li> <li>• 3 advanced DMI</li> <li>• 2 bi-directional interaction</li> <li>• Beyond bean counting - extra flair and effort that supports sensemaking.</li> </ul>
Complexity	<ul style="list-style-type: none"> <li>• Simple analytic tasks</li> <li>• Data wrangling</li> </ul>	<ul style="list-style-type: none"> <li>• Sophisticated analytic tasks</li> <li>• Additional wrangling and processing</li> </ul>	<ul style="list-style-type: none"> <li>• Deep and complex analytic tasks</li> <li>• Deep analysis and insightful commentary</li> </ul>	<ul style="list-style-type: none"> <li>• Deep and complex analytic tasks</li> <li>• Deep analysis and insightful commentary</li> <li>• Extra flair and effort</li> </ul>
Cohesiveness	<ul style="list-style-type: none"> <li>• Separate reports</li> <li>• Mini projects</li> </ul>	<ul style="list-style-type: none"> <li>• Single report</li> <li>• Coordinated styling</li> </ul>	<ul style="list-style-type: none"> <li>• Single report</li> <li>• Coordinated styling</li> </ul>	<ul style="list-style-type: none"> <li>• Single published website</li> <li>• Coordinated styling</li> </ul>

# Project Todos

- Right now, you should be working on your vizzes and your views.
  - A view is a composite of vizzes.
  - Everyone should make sure they have sketched out their views and pushed them to their repo by the end of this week.
  - The TAs will not be reviewing it; you are welcome to update it, but make sure it is done to clarify what you are programming.
- PM2 Feedback and Updates
  - Each group will receive a spreadsheet with both things you can improve for PM2 and advice on what to think about for PM3. The PM3 aspect is most important. Focus on that over the weekend.
  - Information on how we will do regrades will be discussed on Monday.

WHEN? WHERE? WHY?  
WHAT? HOW? WHY? WHEN? WHERE?  
WHO? WHERE? WHAT? HOW?  
HOW? WHO? WHERE? HOW?  
WHERE? WHO? WHERE? HOW?  
WHICH? WHO? WHERE? HOW?  
WHAT? HOW? WHY?  
WHO? WHERE? WHAT? HOW?  
WHY? HOW? WHERE?  
WHO? WHOSE?  
WHERE? WHAT? HOW?  
WHY? WHO?  
WHERE? WHAT? HOW?  
WHERE? WHICH? WHO? WHEN? WHY?  
WHAT? HOW? WHY?  
WHO? WHERE? WHAT? HOW?  
HOW? WHO?  
WHERE? HOW?  
WHAT? HOW? WHY?  
WHERE? WHAT? HOW?  
WHO? WHERE? HOW?



# Visualization for Data Science

## *Color*



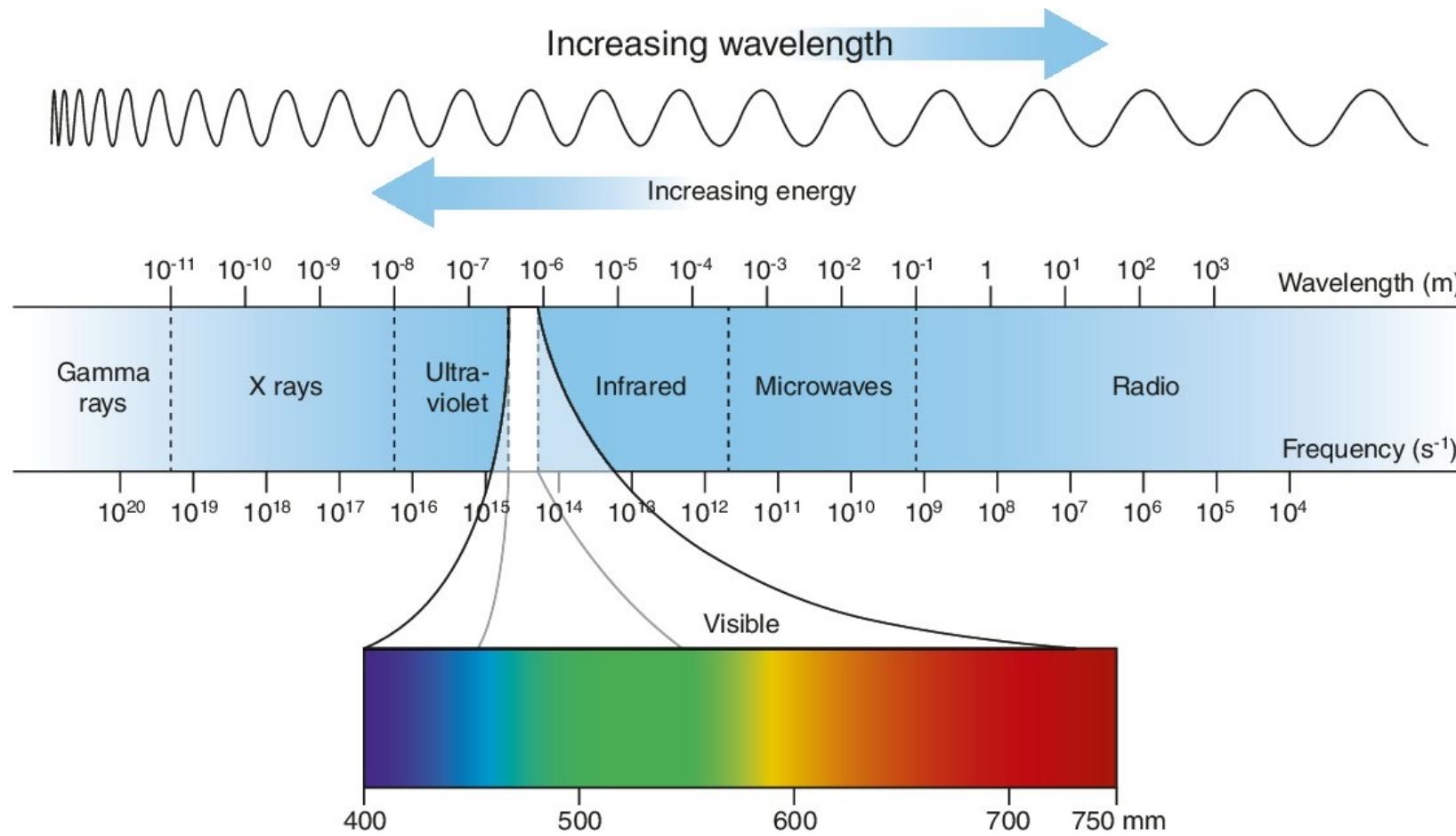
# Learning Outcomes

- List the main three color channels used in visualization
- Explain how color differences affect the user's perception of the visualization
- Detail strategies for making visualizations color accessible
- Describe the various types of color palettes
- Use examples to explain how light and background influence our perceptions
- Discuss how we can use color to effectively represent data attribute values

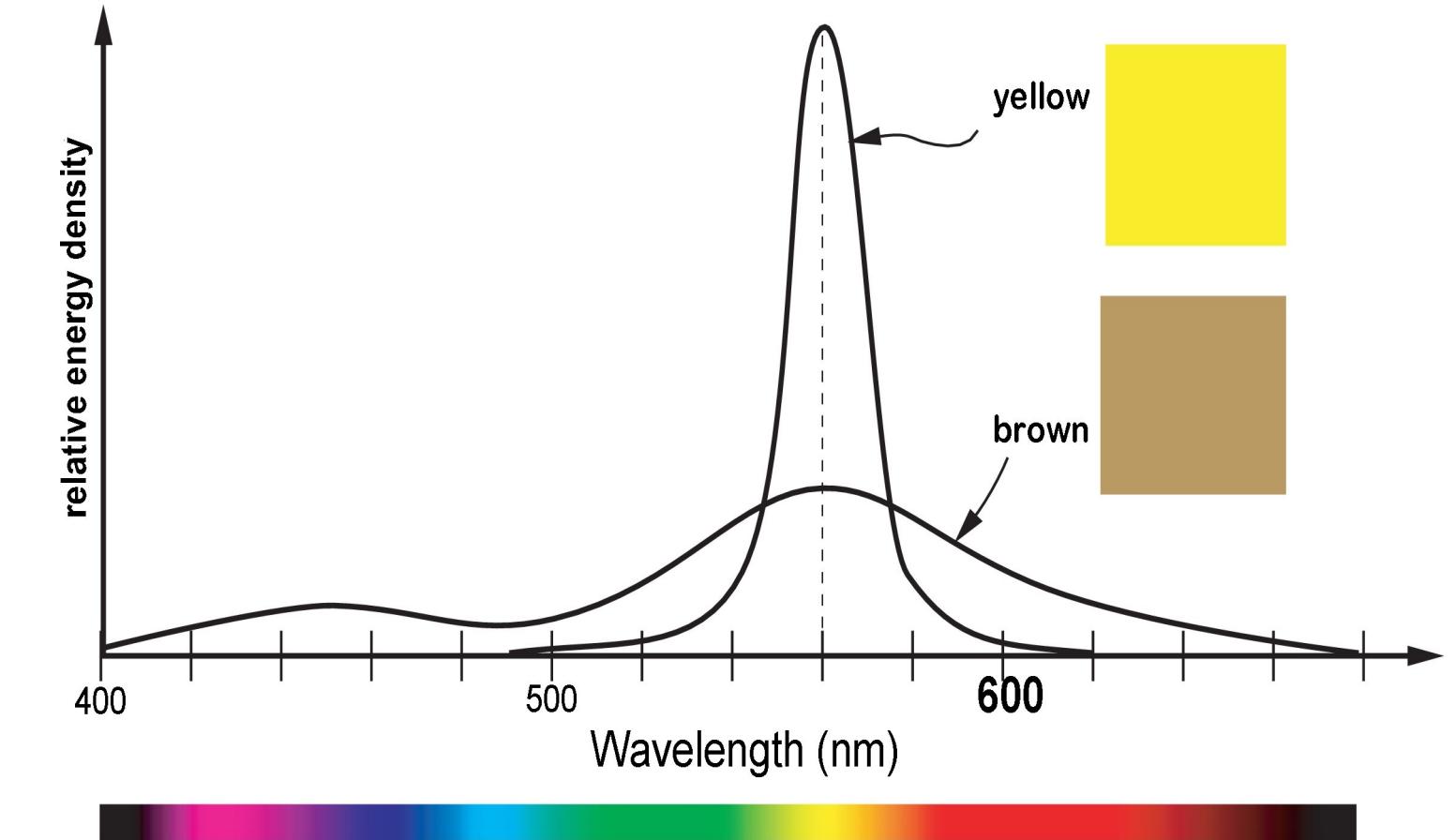
# Color Basics

[Paschos & Bentley, 2019](#)

10



Color is not wavelength, instead it is a combination of wavelengths and energy



# Color is just part of vision system

Color vision is not crucial for many aspects of normal vision

Color does not aid in determining spatial layout, movement, or shape of objects

What are the circular objects? 

CQ: How many circular objects are in the figure?

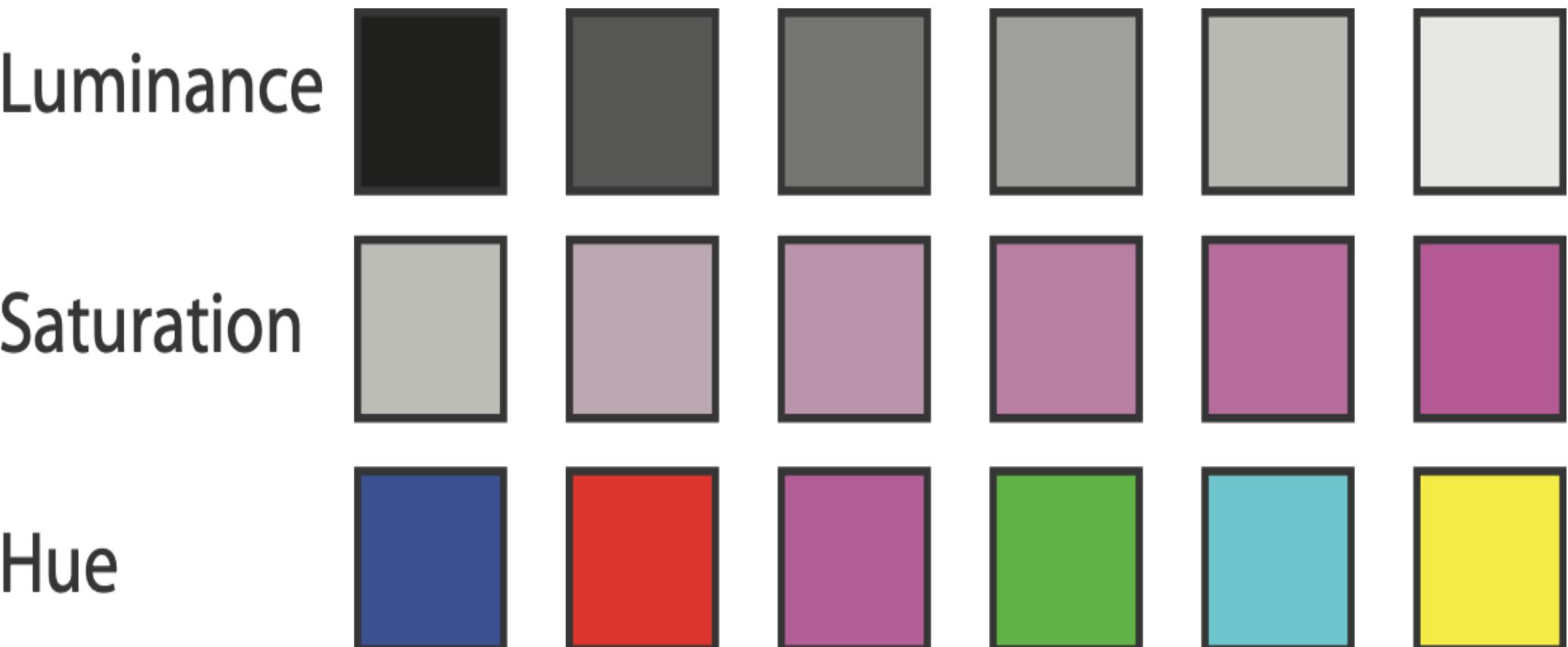


Information Visualization Perception for Design. Colin Ware  
Figure 4.1. Finding the [REDACTED] is much easier with color vision.

# Decomposing color

decompose into three channels (*or perceptual attributes*)

- **luminance:** brightness or darkness value
  - how bright is it?
- **saturation:** intensity of color
  - how colorful is it?
- **hue:** the type of color
  - what color is it?



# Channels: What's up with color?

## → **Magnitude Channels: Ordered Attributes**

Position on common scale



Position on unaligned scale



Length (1D size)



Tilt angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



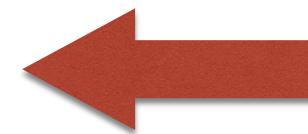
Effectiveness  
Best ↑  
Same  
Least ↓

## → **Identity Channels: Categorical Attributes**

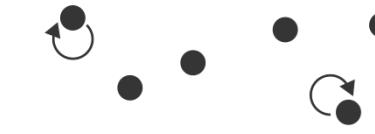
Spatial region



Color hue



Motion



Shape



Effectiveness  
Best ↑  
Same  
Least ↓

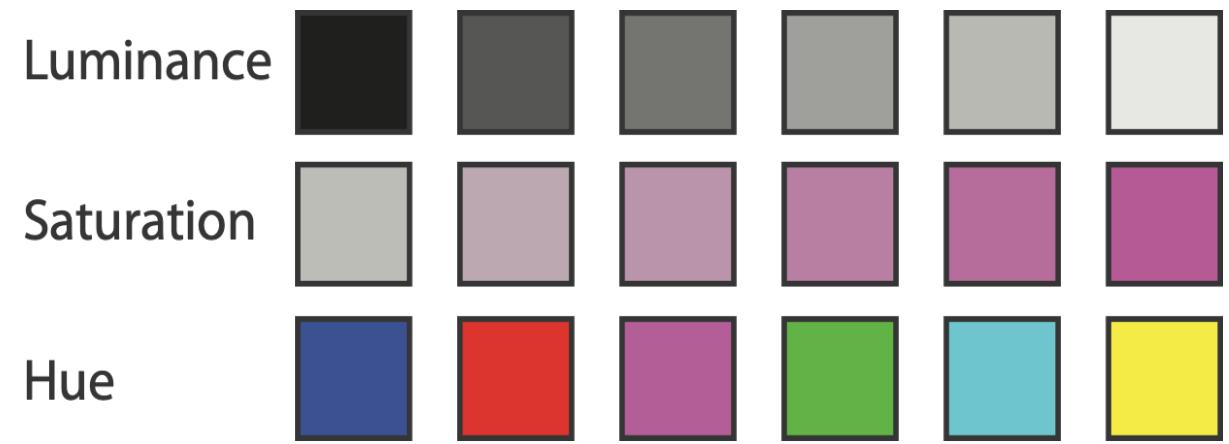
# Decomposing color

Ordered – can show magnitude

- **luminance**: brightness or darkness value
- **saturation**: intensity of color

Categorical – can show identity

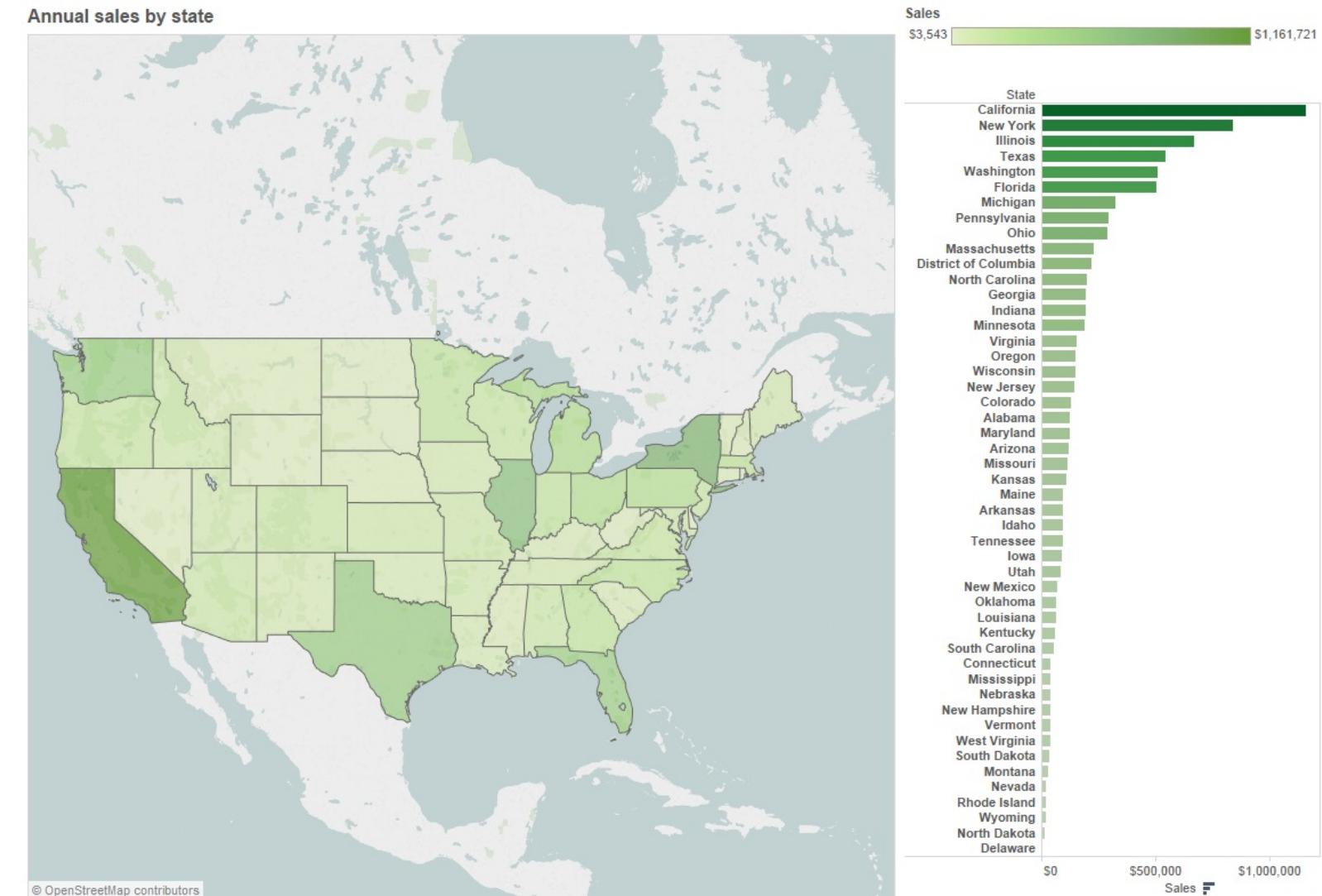
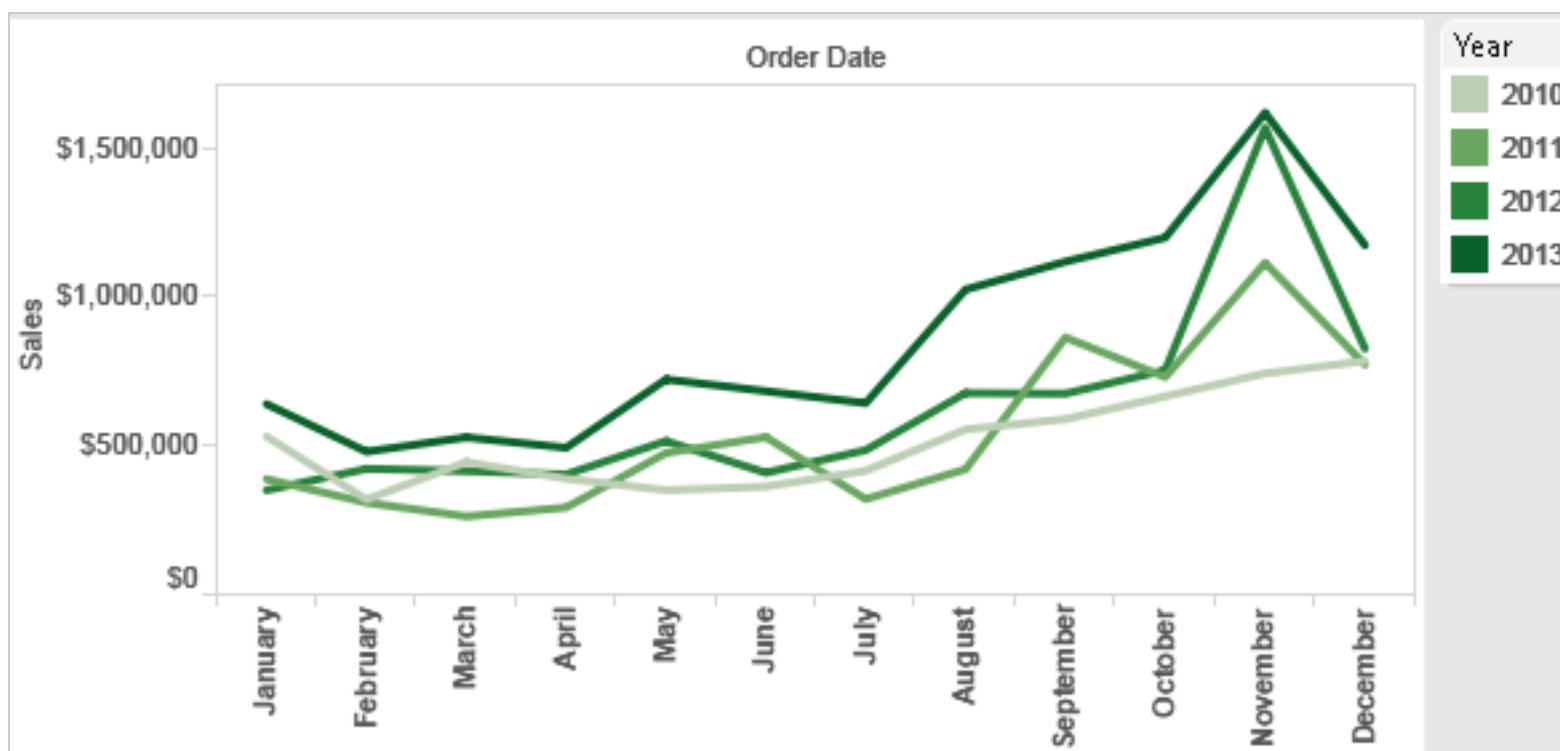
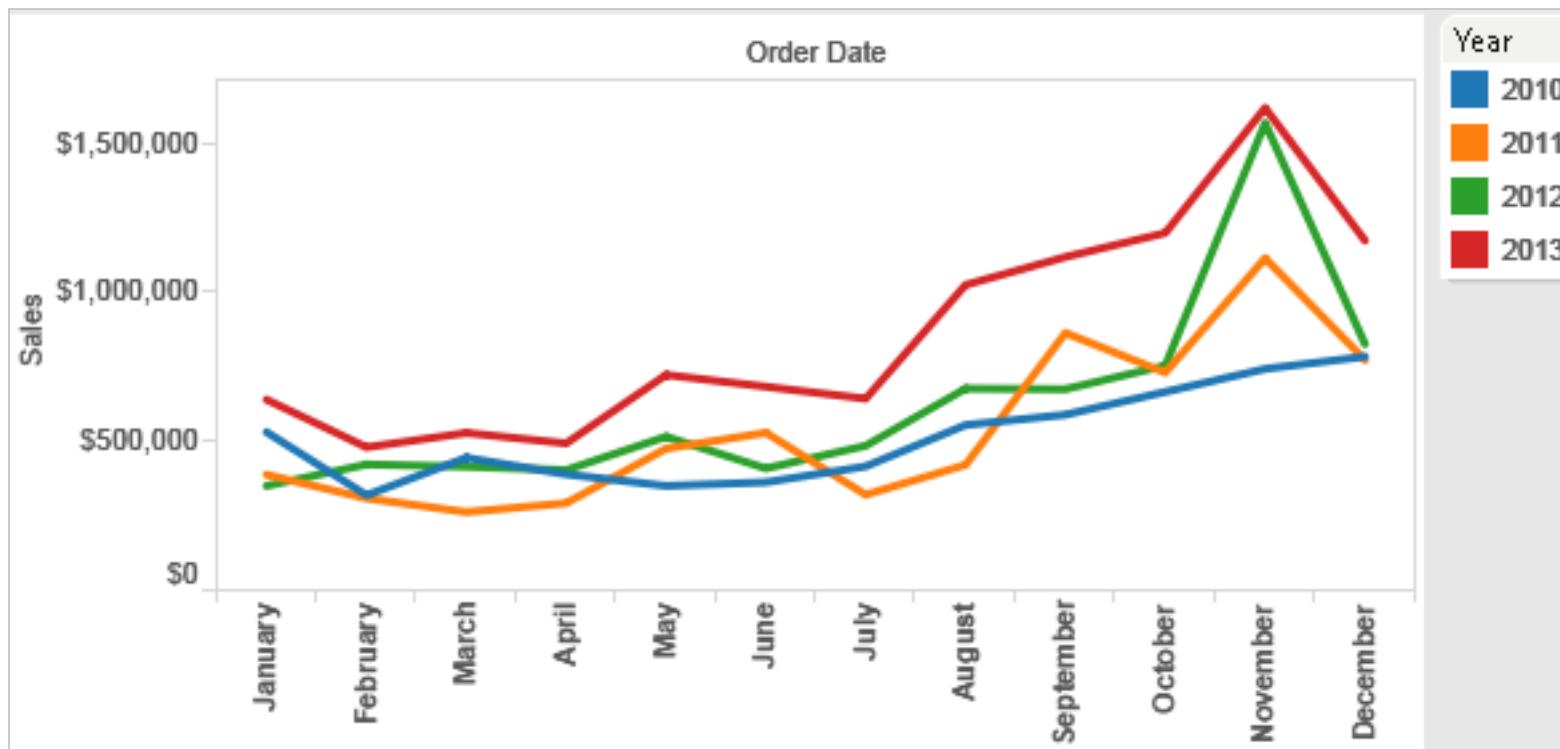
- **hue**: the type of color



Channels have different properties

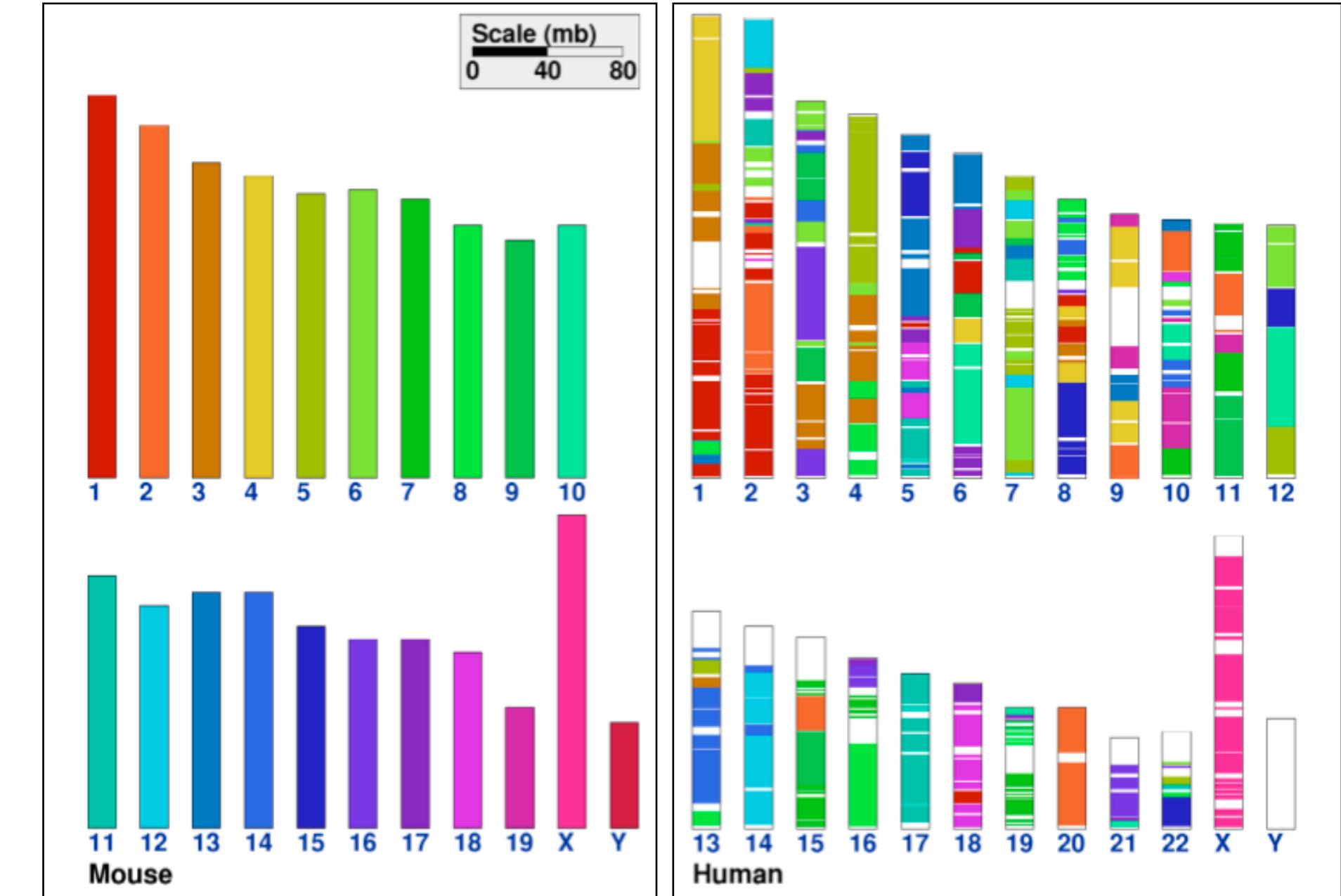
- what they convey directly to perceptual system
- how much they can convey
  - how many discriminable bins can we use?

# Categorical vs ordered color



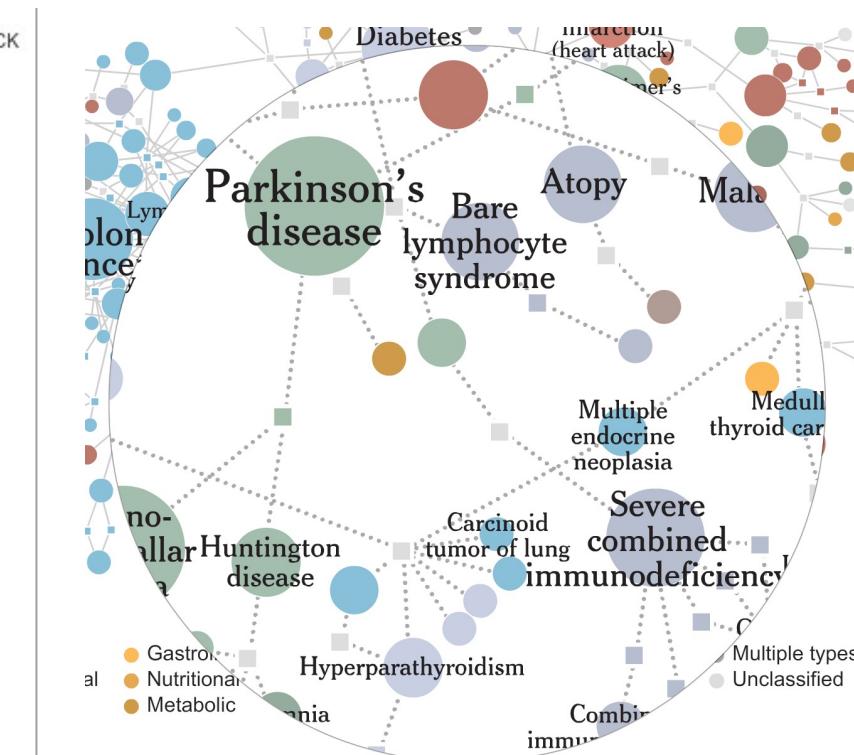
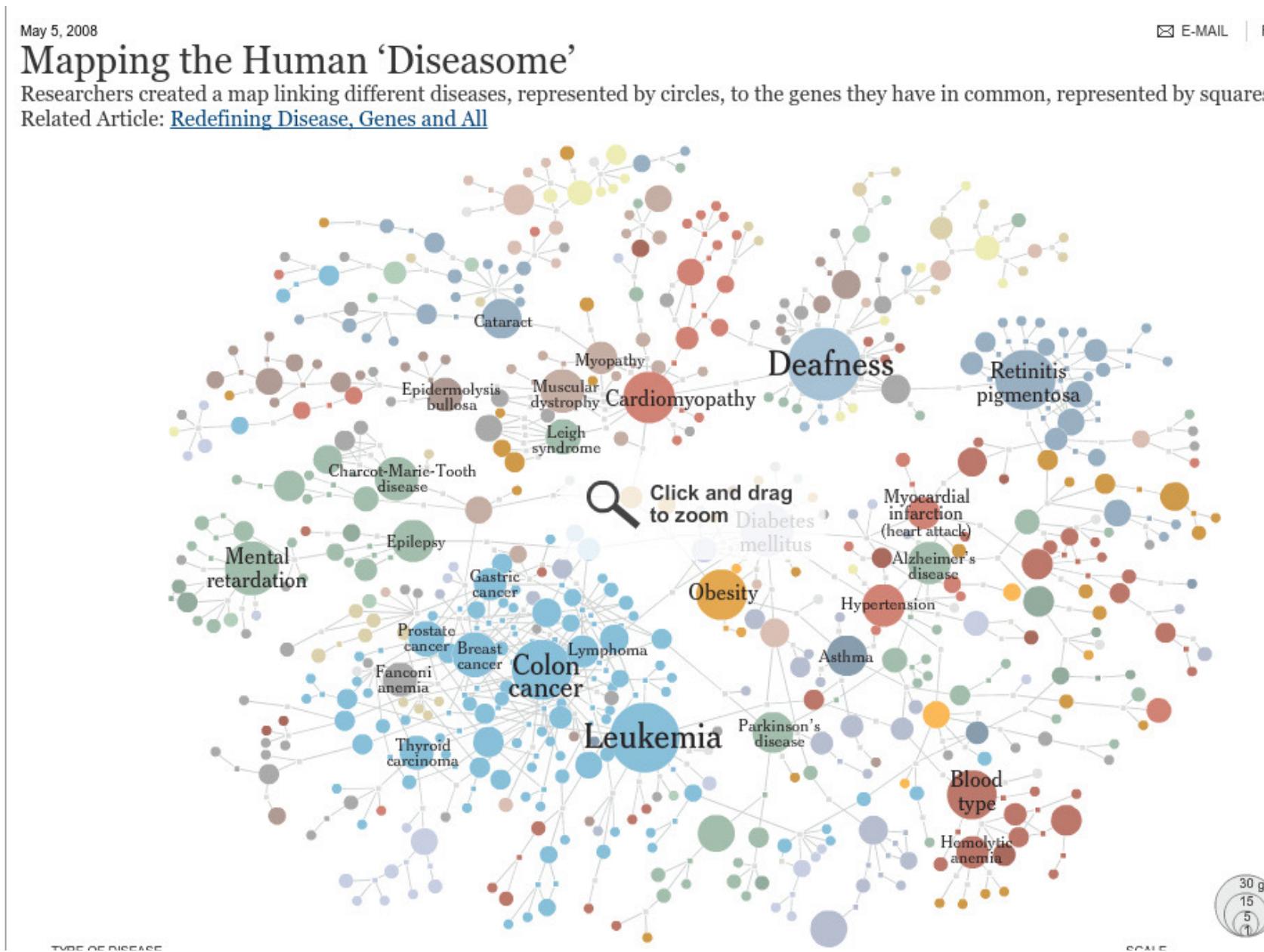
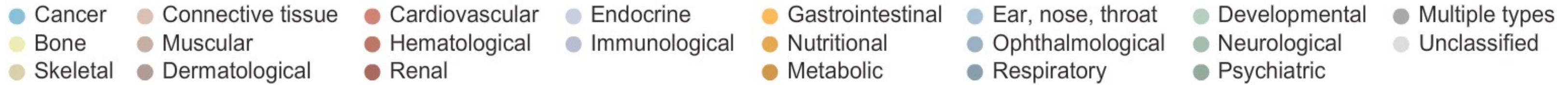
# Categorical color: limited number of discriminable bins

- human perception built on relative comparisons
  - great if color contiguous
  - surprisingly bad for absolute comparisons
- noncontiguous small regions of color
  - fewer bins than you want
  - rule of thumb: 6-12 bins, including background and highlights



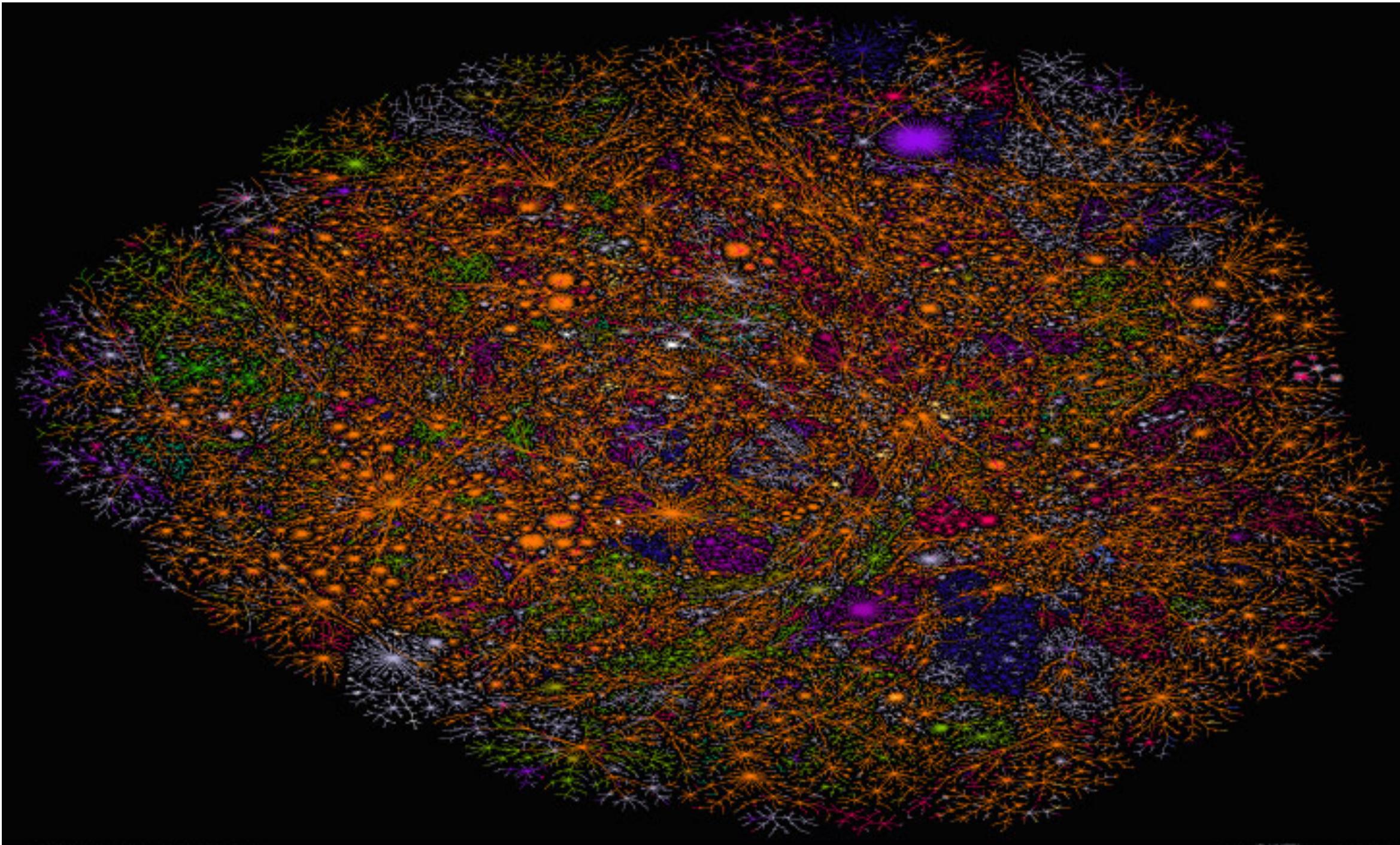
[Cinteny: flexible analysis and visualization of synteny and genome rearrangements in multiple organisms. Sinha and Meller. BMC Bioinformatics, 8:82, 2007.]

# Categorical color: limited number of discriminable bins



[https://archive.nytimes.com/www.nytimes.com/interactive/2008/05/05/science/20080506\\_DISEASE.html](https://archive.nytimes.com/www.nytimes.com/interactive/2008/05/05/science/20080506_DISEASE.html)

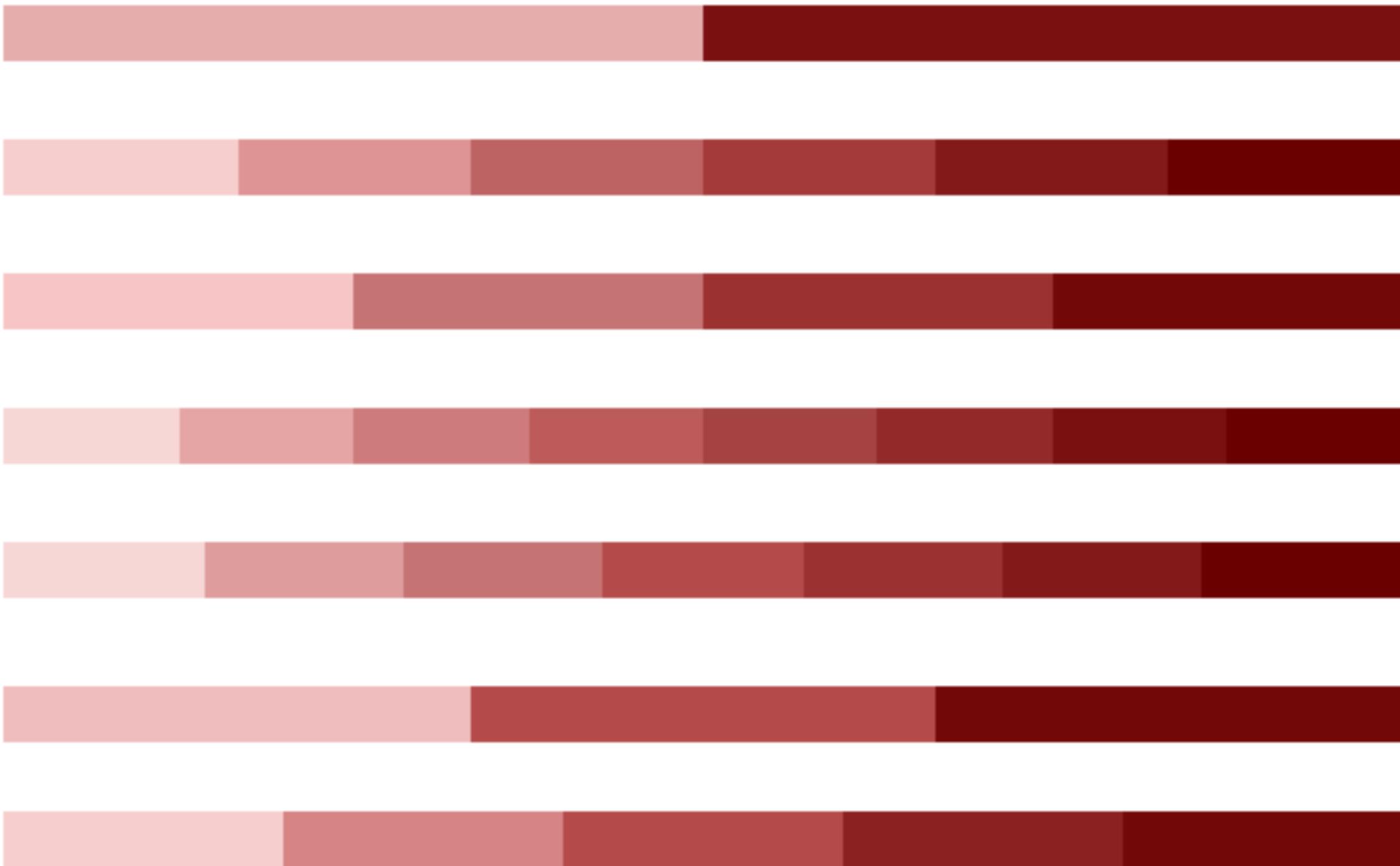
# Categorical color



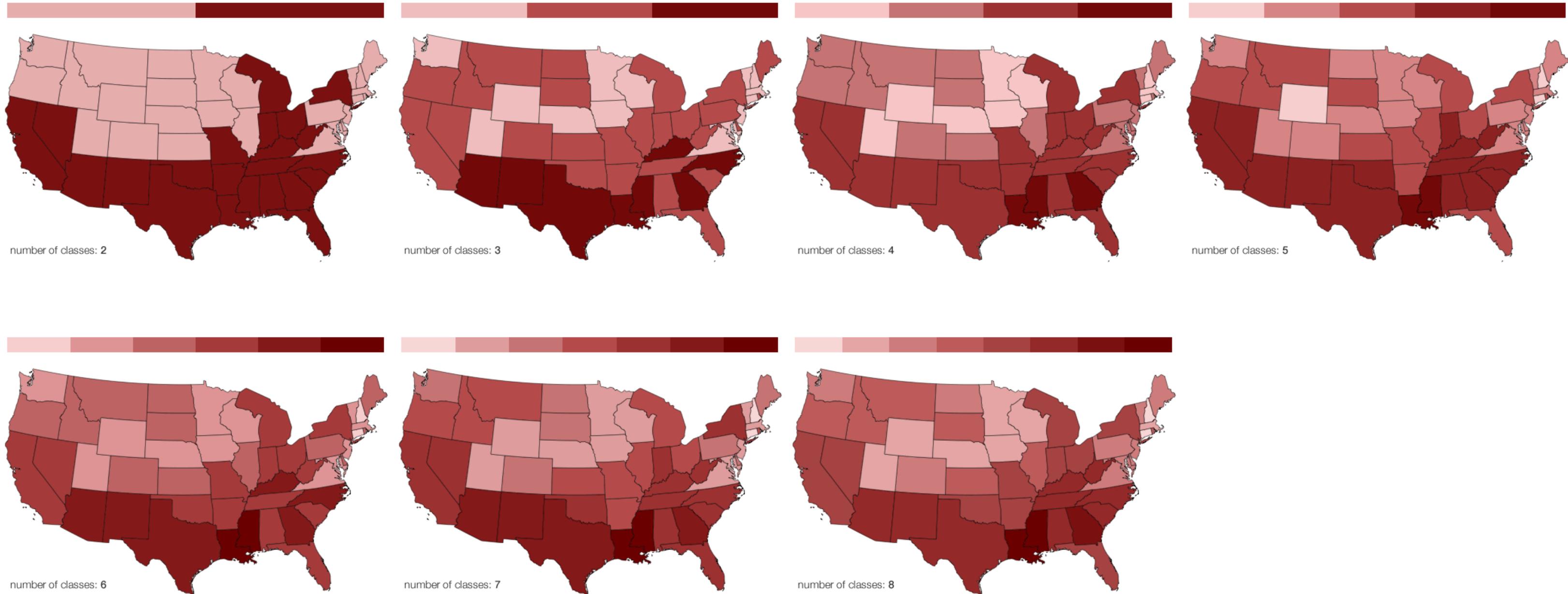
How many colors  
can you  
distinguish?

- A. <=6
- B. 7 - 8
- C. 9 – 10
- D. 11 - 13
- E. 14 - 18

For each bar – how many distinct bins are there?

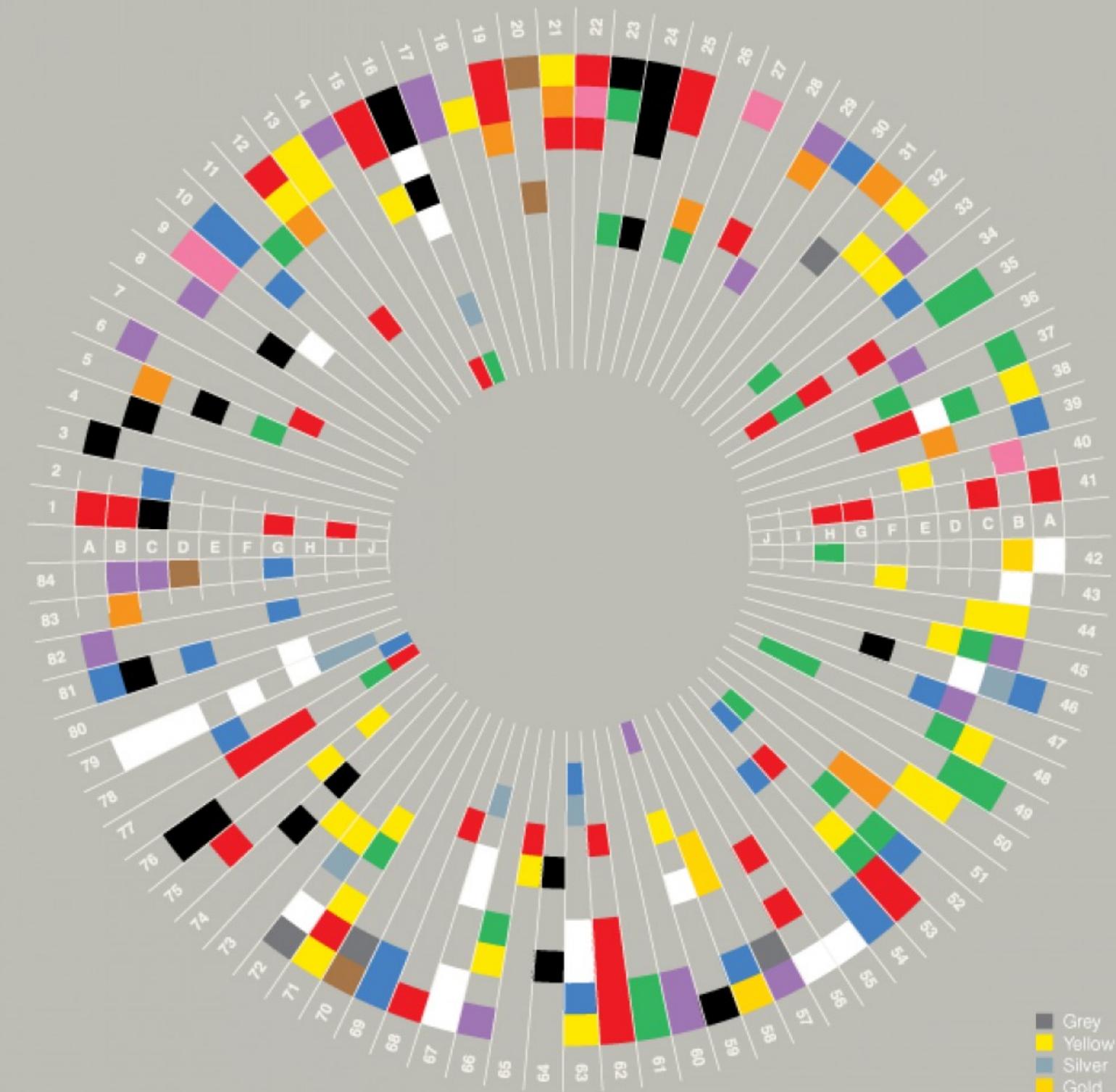


# Ordered color: limited number of discriminable bins



# What does the color mean to you or to them?

## Colours In Cultures



A	Western / American	F	Asian
B	Japanese	G	Eastern European
C	Hindu	H	Muslim
D	Native American	I	African
E	Chinese	J	South American

1	Anger	43	Holiness
2	Art / Creativity	44	Illness
3	Authority	45	Insight
4	Bad Luck	46	Intelligence
5	Balance	47	Intuition
6	Beauty	48	Religion
7	Calm	49	Jealousy
8	Celebration	50	Joy
9	Children	51	Learning
10	Cold	52	Life
11	Compassion	53	Love
12	Courage	54	Loyalty
13	Cowardice	55	Luxury
14	Cruelty	56	Marriage
15	Danger	57	Modesty
16	Death	58	Money
17	Decadence	59	Mourning
18	Deceit	60	Mystery
19	Desire	61	Nature
20	Earthy	62	Passion
21	Energy	63	Peace
22	Erotic	64	Penance
23	Eternity	65	Power
24	Evil	66	Personal power
25	Excitement	67	Purity
26	Family	68	Radicalism
27	Femininity	69	Rational
28	Fertility	70	Reliable
29	Flamboyance	71	Repels Evil
30	Freedom	72	Respect
31	Friendly	73	Royalty
32	Fun	74	Self-cultivation
33	God	75	Strength
34	Gods	76	Style
35	Good Luck	77	Success
36	Gratitude	78	Trouble
37	Growth	79	Truce
38	Happiness	80	Trust
39	Healing	81	Unhappiness
40	Healthy	82	Virtue
41	Heat	83	Warmth
42	Heaven	84	Wisdom

## P: Ordering Colors?

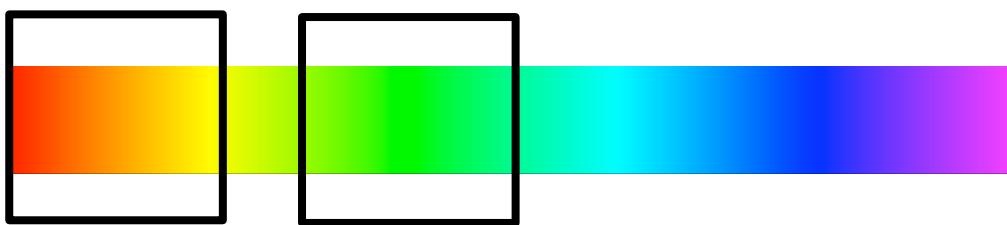
Which ordering of colors is best?

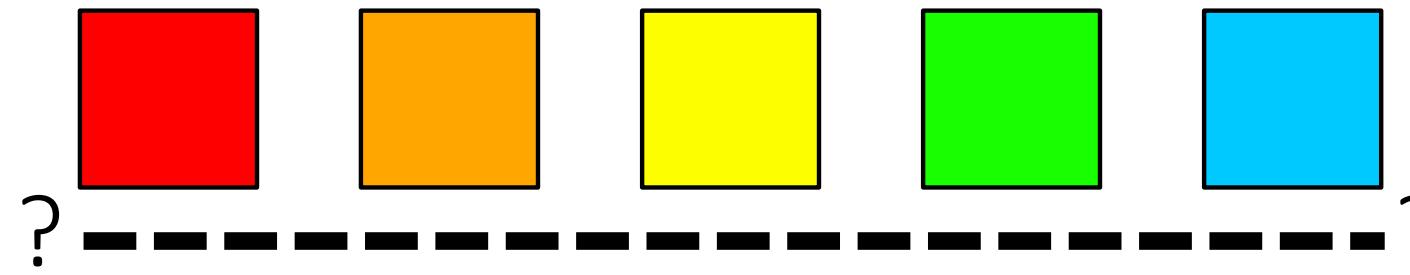
- A. Red Blue Green Yellow Purple
- B. Red Yellow Green Purple Blue
- C. Purple Blue Green Yellow Red
- D. Green Red Blue Purple Yellow
- E. Yellow Red Green Blue Purple

# Ordered color: Rainbow is poor default

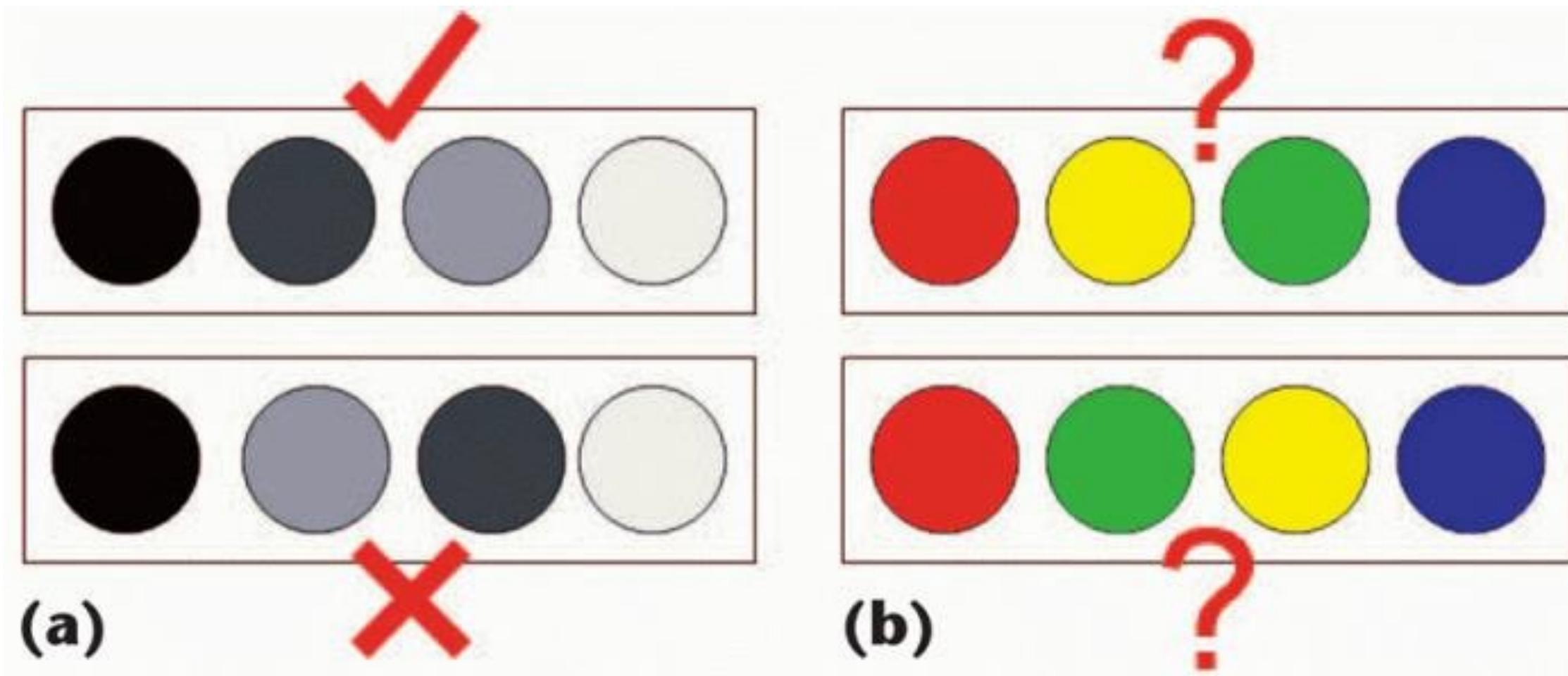
## Problems

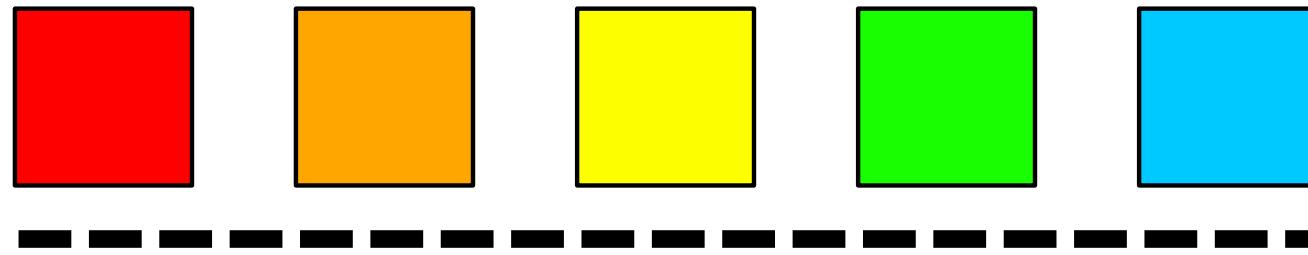
- perceptually unordered
  - Is red bigger than purple, is green bigger than blue, is blue smaller than yellow.
  - There is no common agreed upon (intrinsic conceptualization) of each hue's value
- perceptually nonlinear
  - The region of red is not the same length as the region of green, in the same space that we have red, orange yellow, we have just green green green.





No perceptual ordering (confusing)





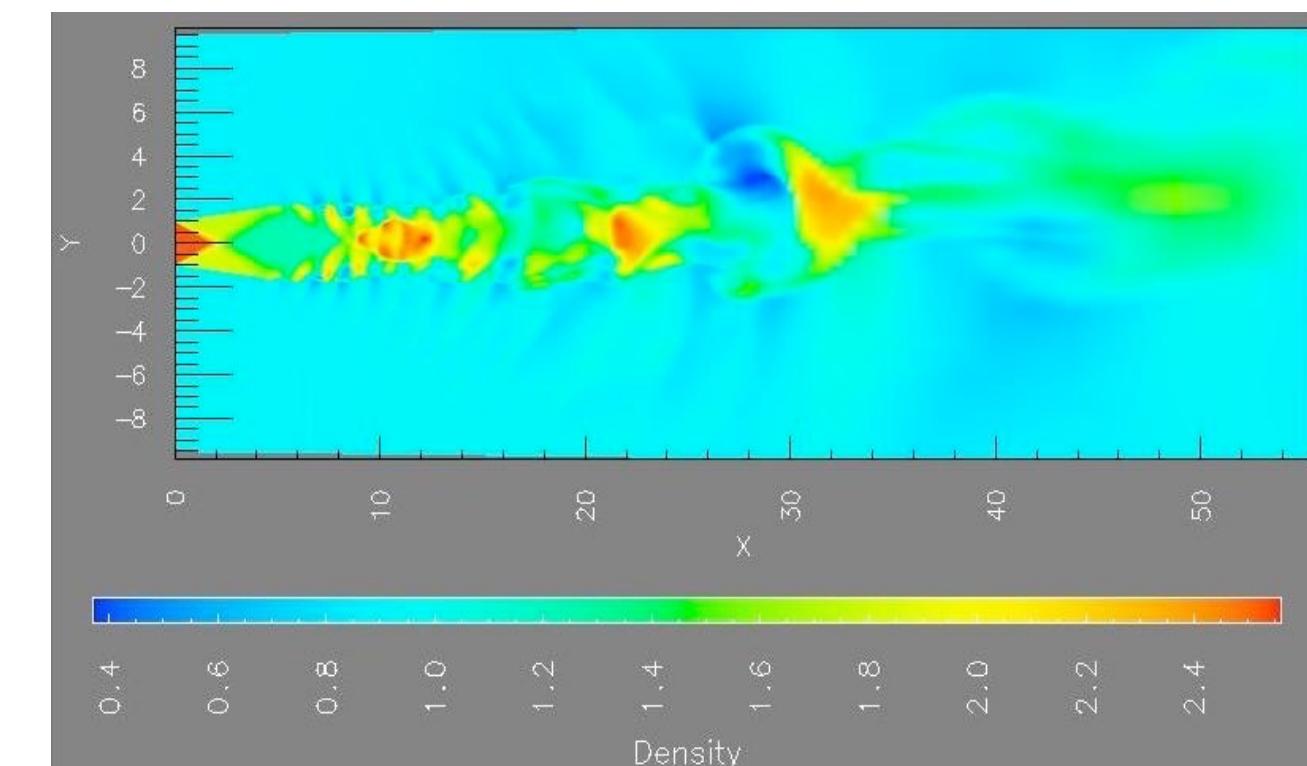
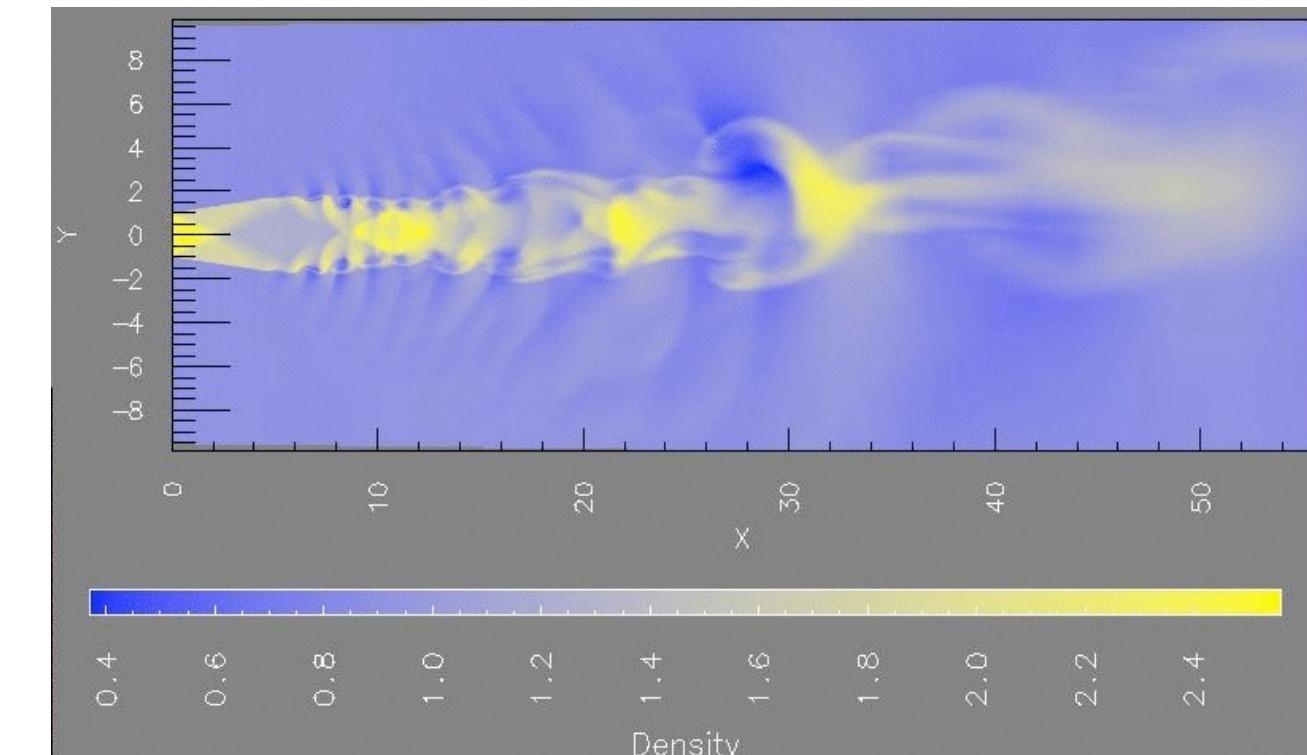
Why this color map is a poor choice for quantitative data...

- No perceptual ordering (confusing)
- No darkness variation (obscures details)
- Viewers perceive sharp transitions in color as sharp transitions in the data, even when this is not the case (misleading)

# Ordered color: Rainbow is poor default

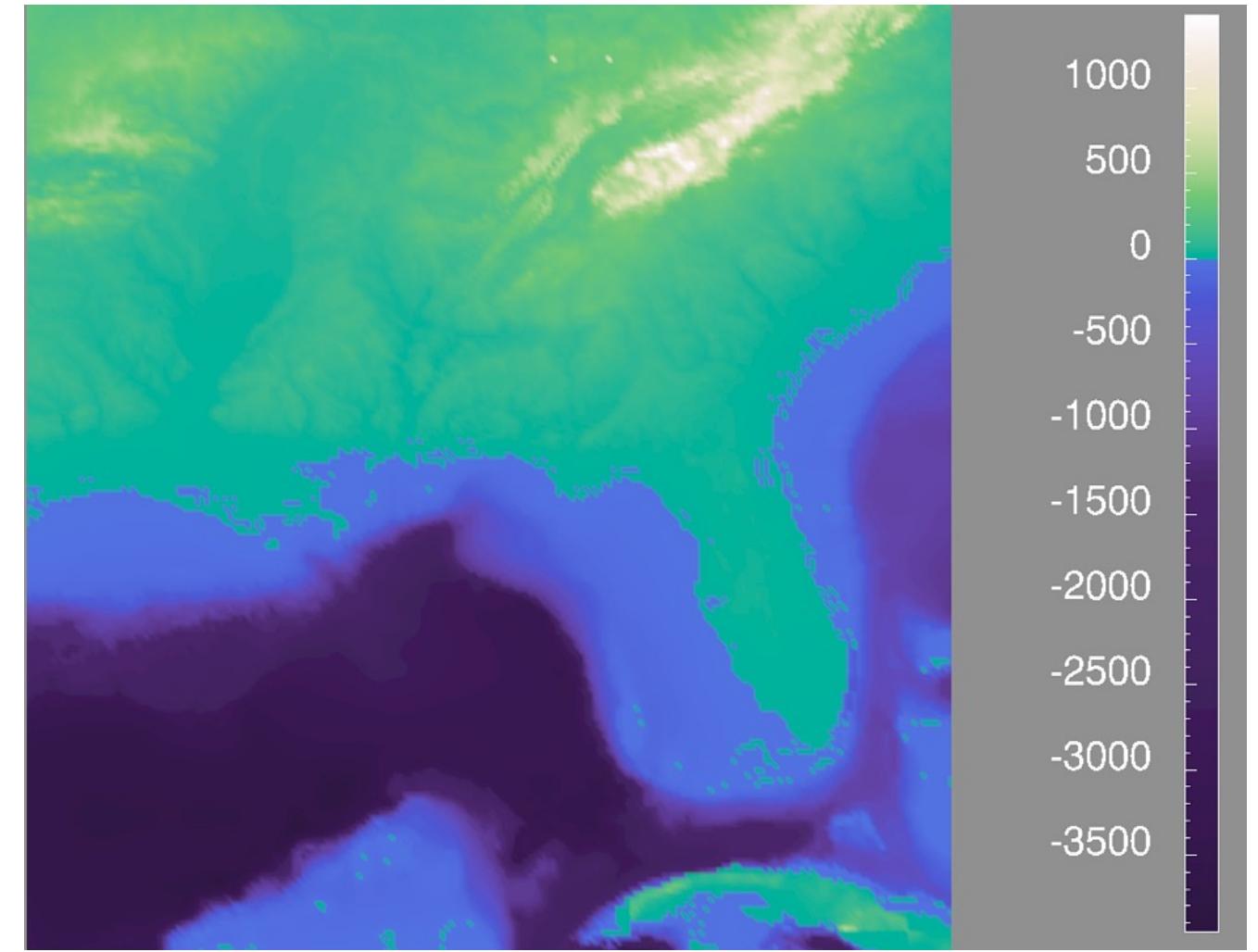
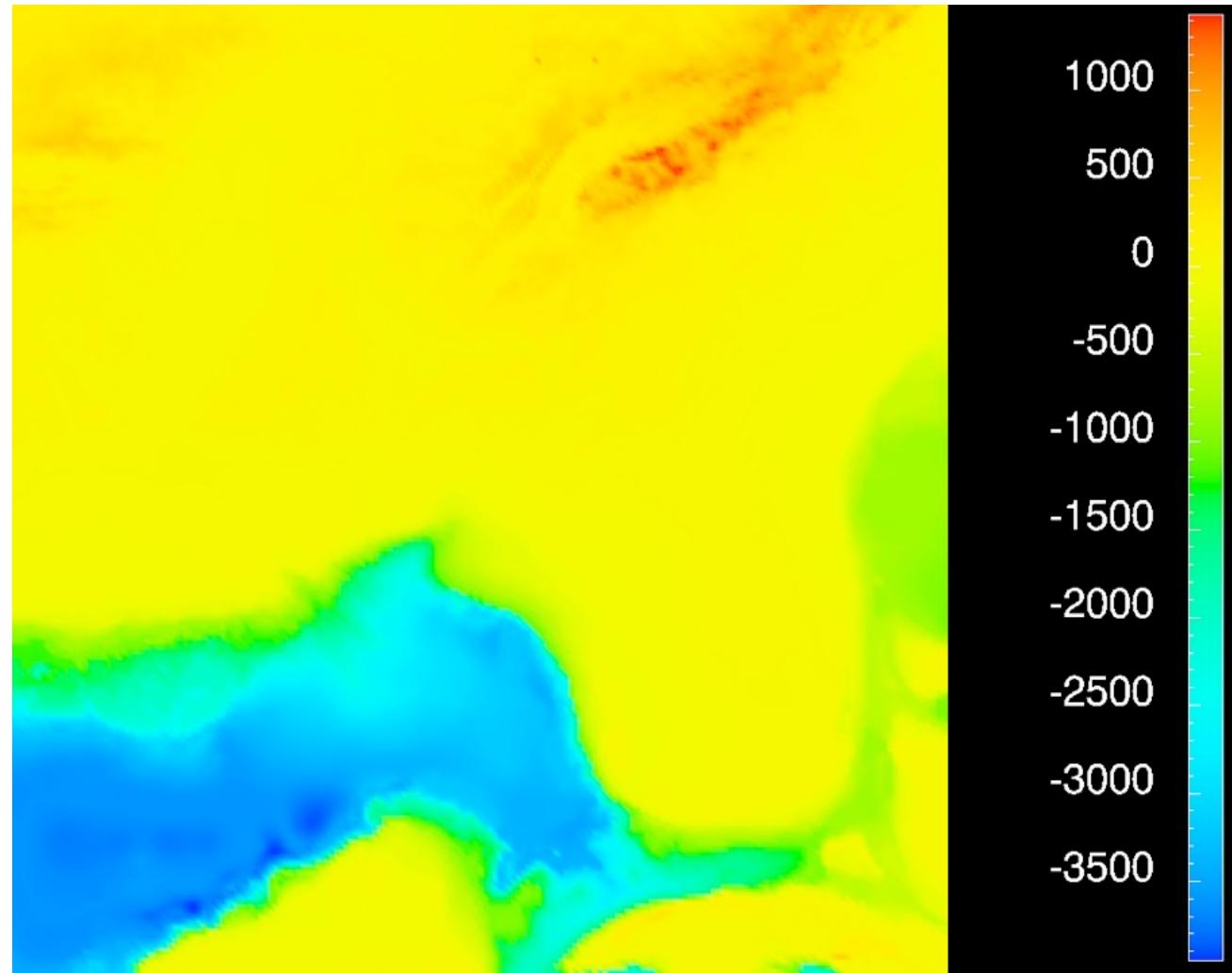
## Benefits

- Having a wide range of colors allows us to emphasize different aspects of a structure
- For instance, in the top image, we use 3 hues and can get a sense of the large-scale structure
- In the lower image, we use more hues and fine-grained structure are visible and nameable



# Ordered color: Rainbow is poor default

Effective at the discretion of the designer

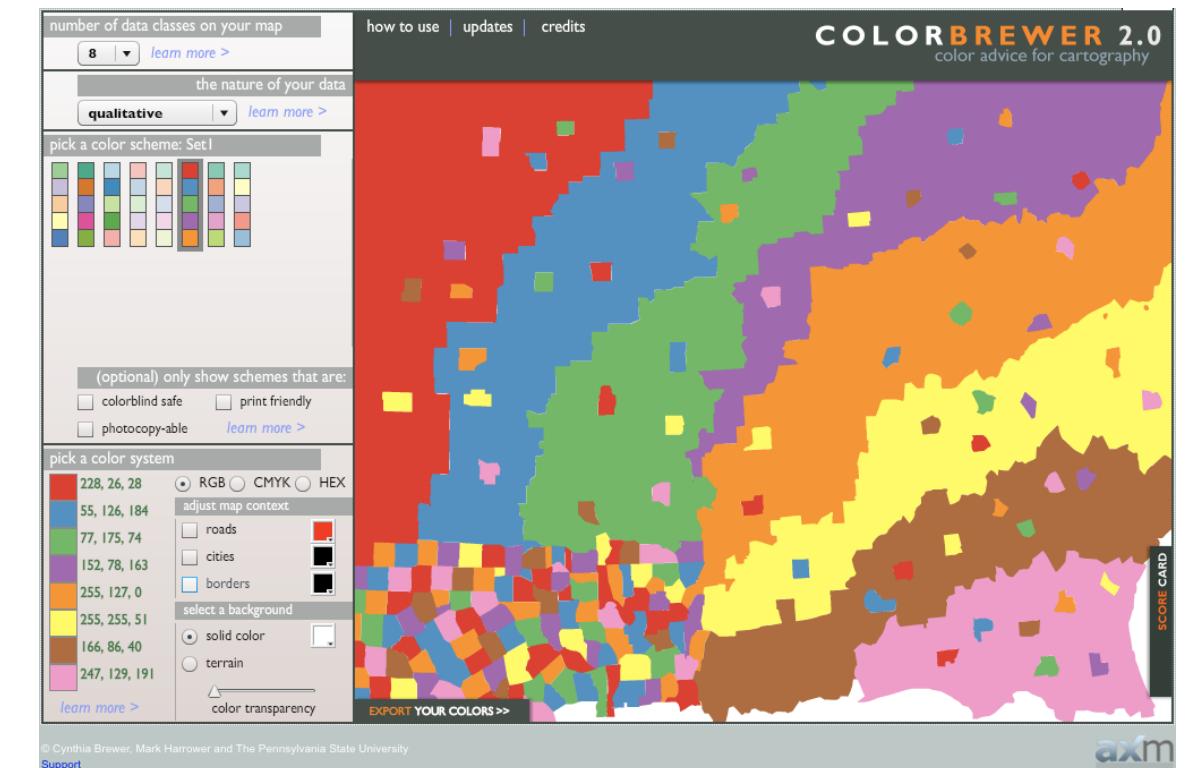
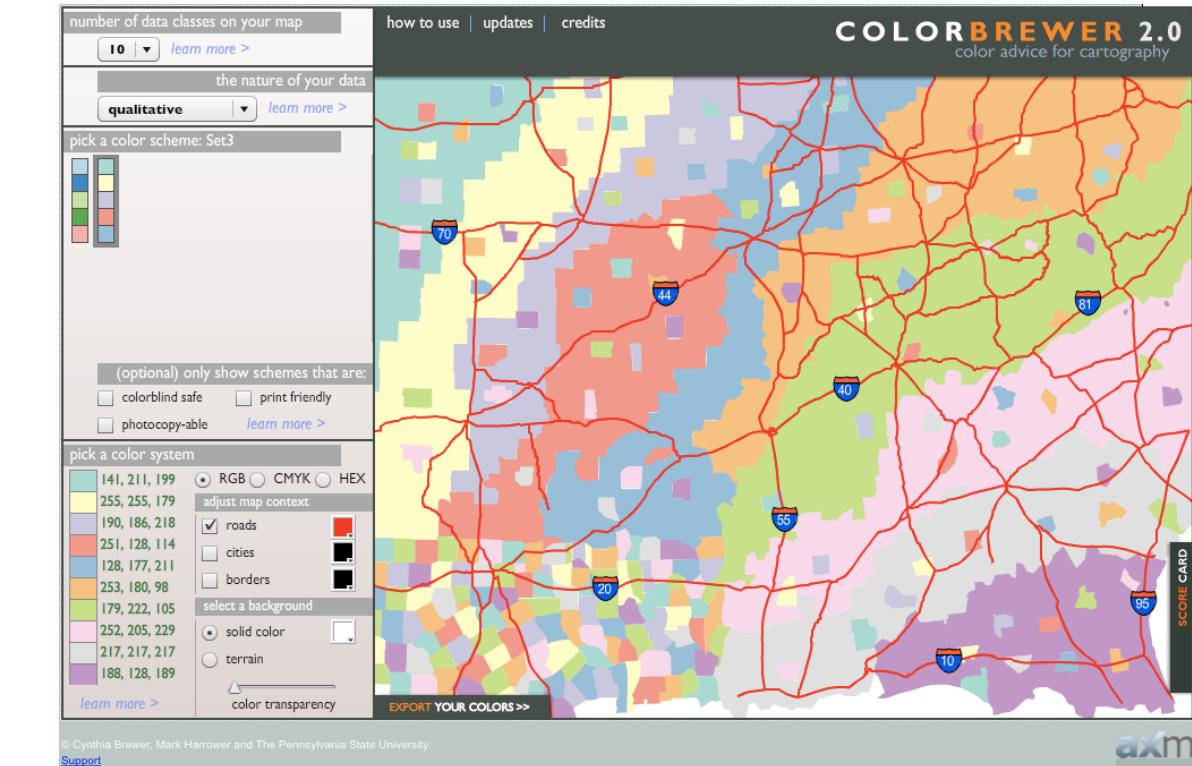


[\[Why Should Engineers Be Worried About Color? Treinish and Rogowitz 1998.\]](#)  
<http://www.research.ibm.com/people/l/lloydt/color/color.HTM>

# Interaction between channels: Not fully separable

## color channel interactions

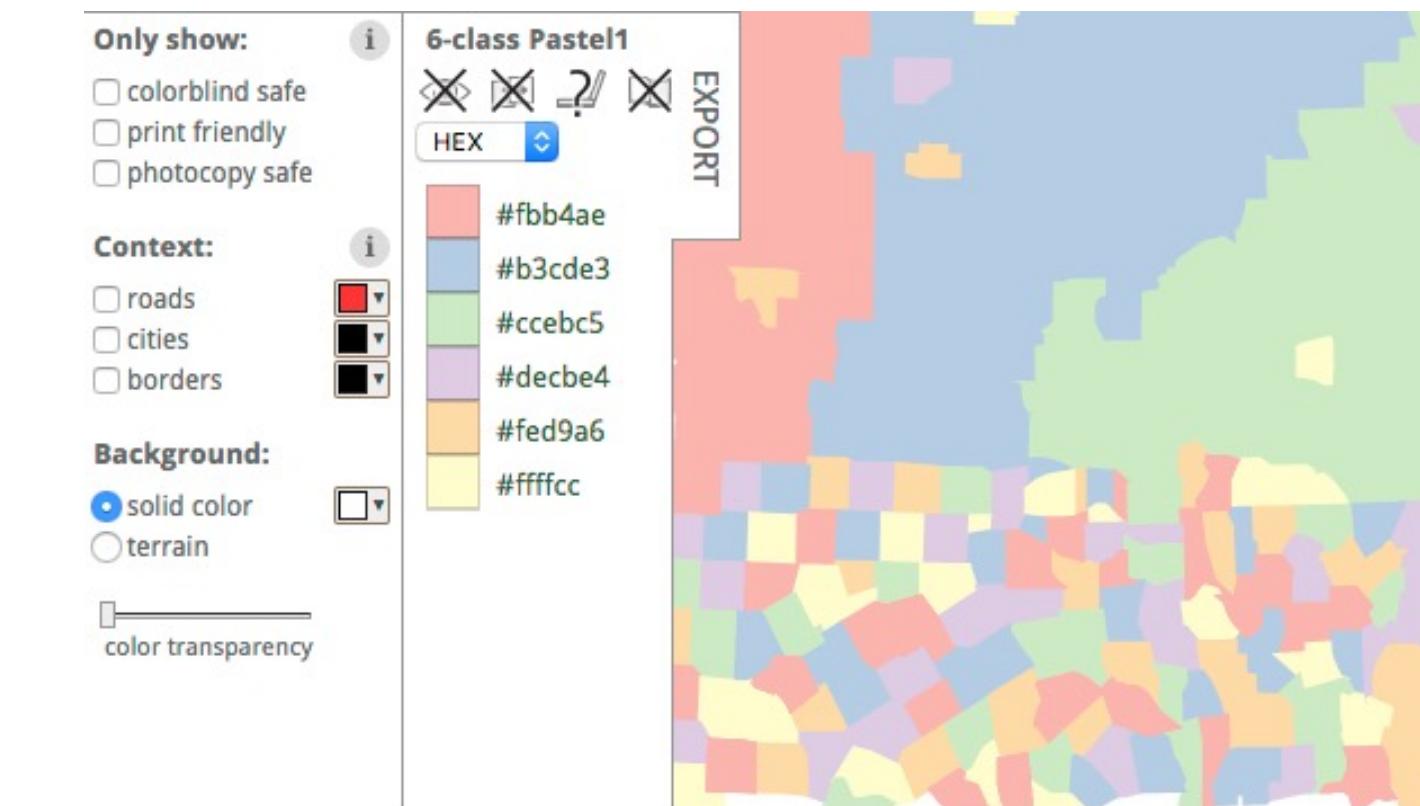
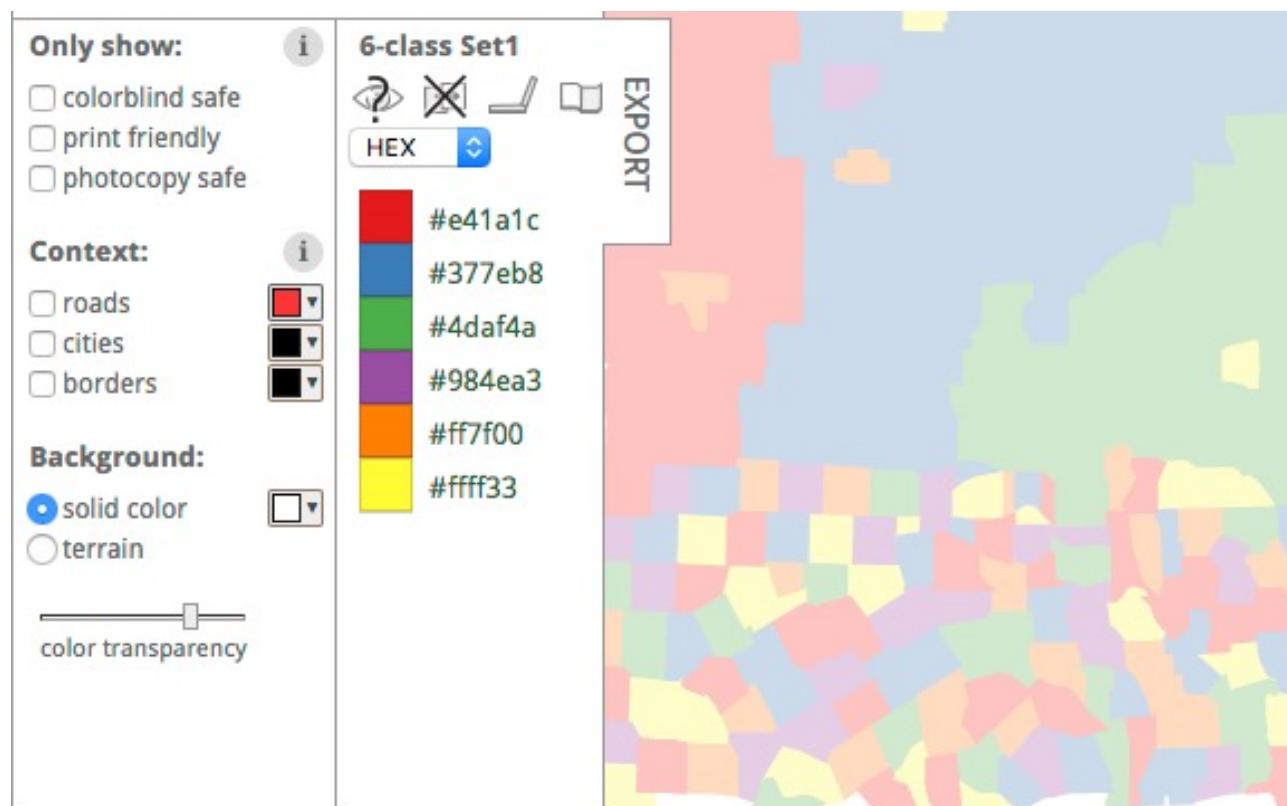
- size heavily affects salience
- small regions need high saturation
- large regions need low saturation



# Interaction between channels: Not fully separable

saturation & luminance:

- not separable from each other!
- also not separable from transparency
- small separated regions: 2 bins safest (use only one of these channels), 3-4 bins max
- contiguous regions: many bins (use only one of these channels)



# What is a colormap/ color palette?

specifies a mapping between color and values

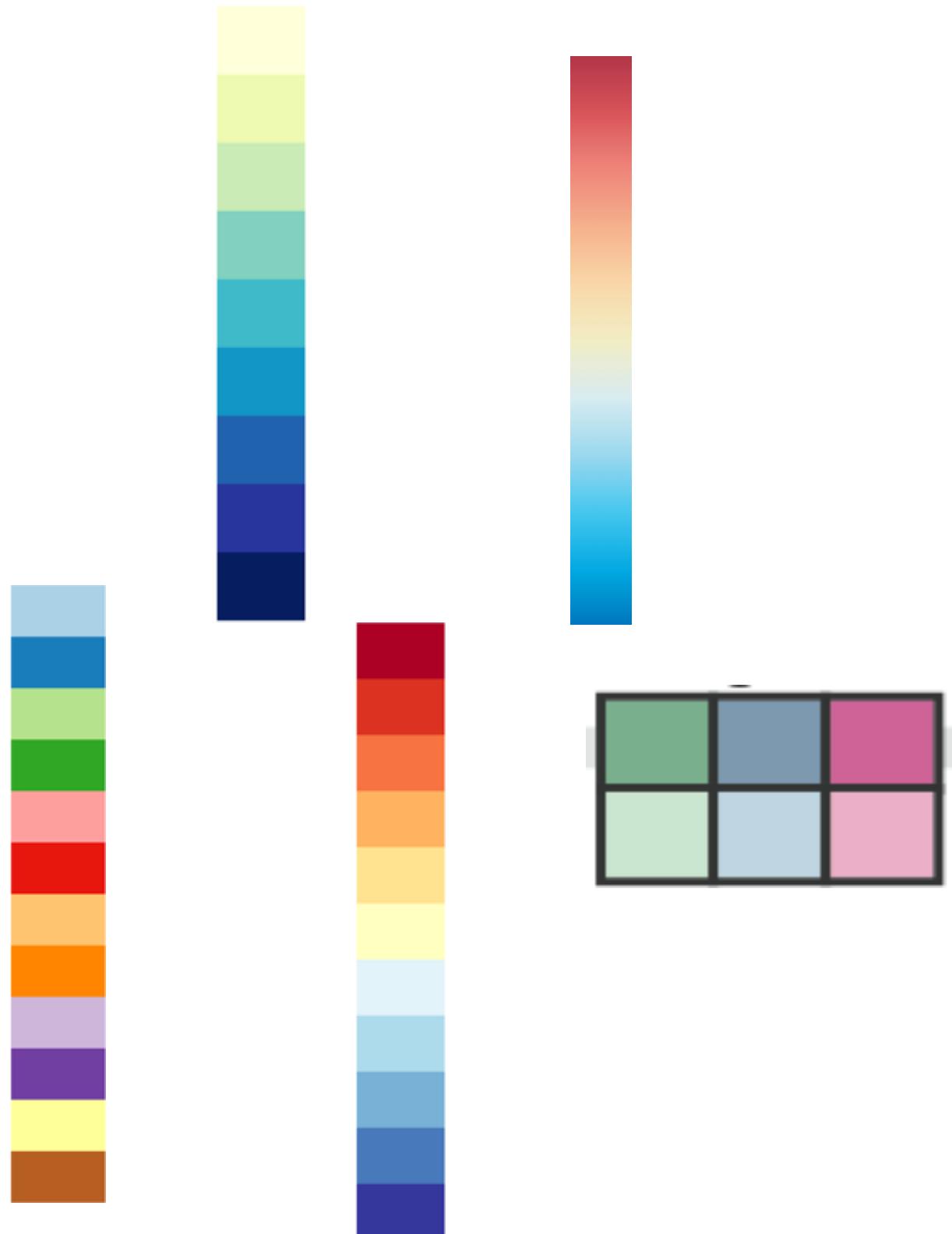
categorical vs ordered

sequential vs diverging

segmented vs continuous

univariate vs bivariate

Design Tip: Match colormap to attribute characteristics!

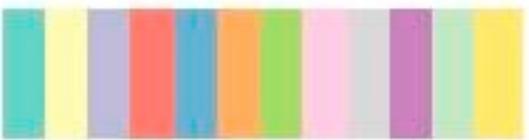


# Color palettes: univariate

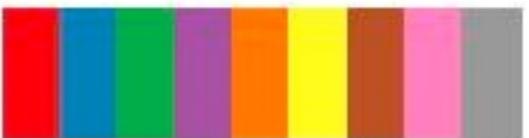
## → Categorical



- aim for maximum distinguishability
- aka *qualitative, nominal*



categorical



# Color palettes: univariate

→ Ordered

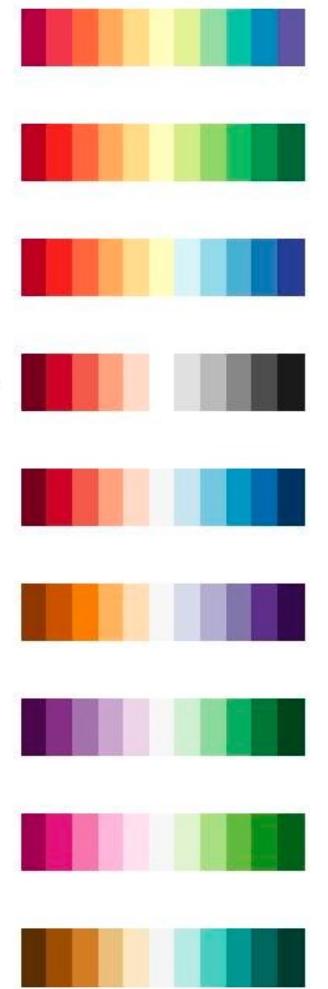
→ *Sequential*

→ *Diverging*



- diverging
  - useful when data has meaningful "midpoint"
  - use **neutral color for midpoint**
    - white, yellow, grey
  - use saturated colors for endpoints
- sequential
  - ramp luminance or saturation
  - if multi-hue, good to order by luminance

diverging



sequential



Cividis

Viridis

Inferno

Magma

Plasma

Warm

Cool

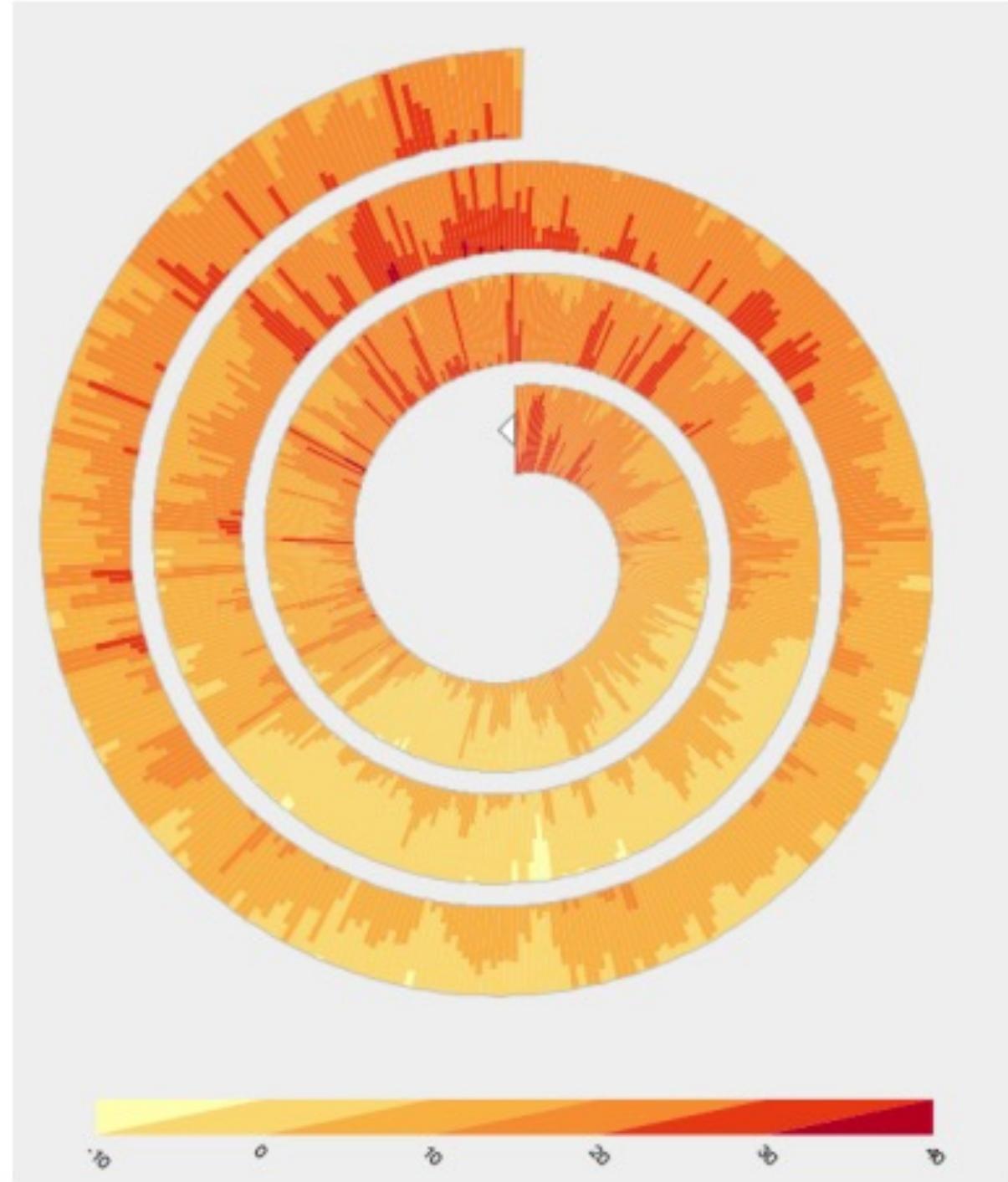
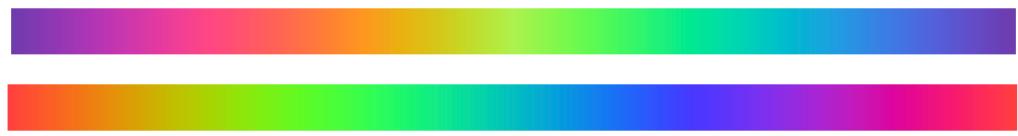
CubehelixDefault

# Color palettes: univariate

→ Cyclic



cyclic multihue



# Color palette design considerations: univariate

segmented

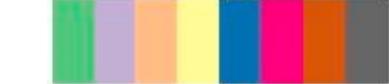
diverging



sequential



categorical



continuous

(a)

(b)

(c)

cyclic multihue



sequential  
single hue

diverging  
two hue

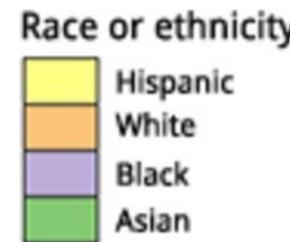
sequential  
multihue

# Color map/palette design considerations

- segmented or continuous?
- diverging or sequential or cyclic?
- single-hue or two-hue or multi-hue?
- perceptually linear?
- ordered by luminance?
- colorblind safe?

# Application: Example

## Categorical

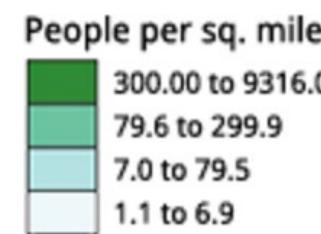


Does not imply magnitude differences  
(categorical/nominal data)

Distinct hues with similar emphasis

---

## Sequential

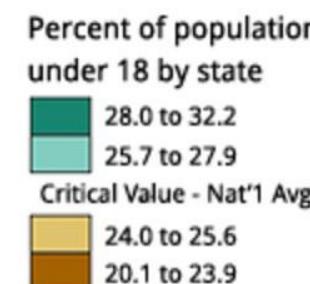


Best for ordered data that progresses from low to high (ordinal, quantitative data)

Darkness (lightness) channel effectively employed

---

## Diverging

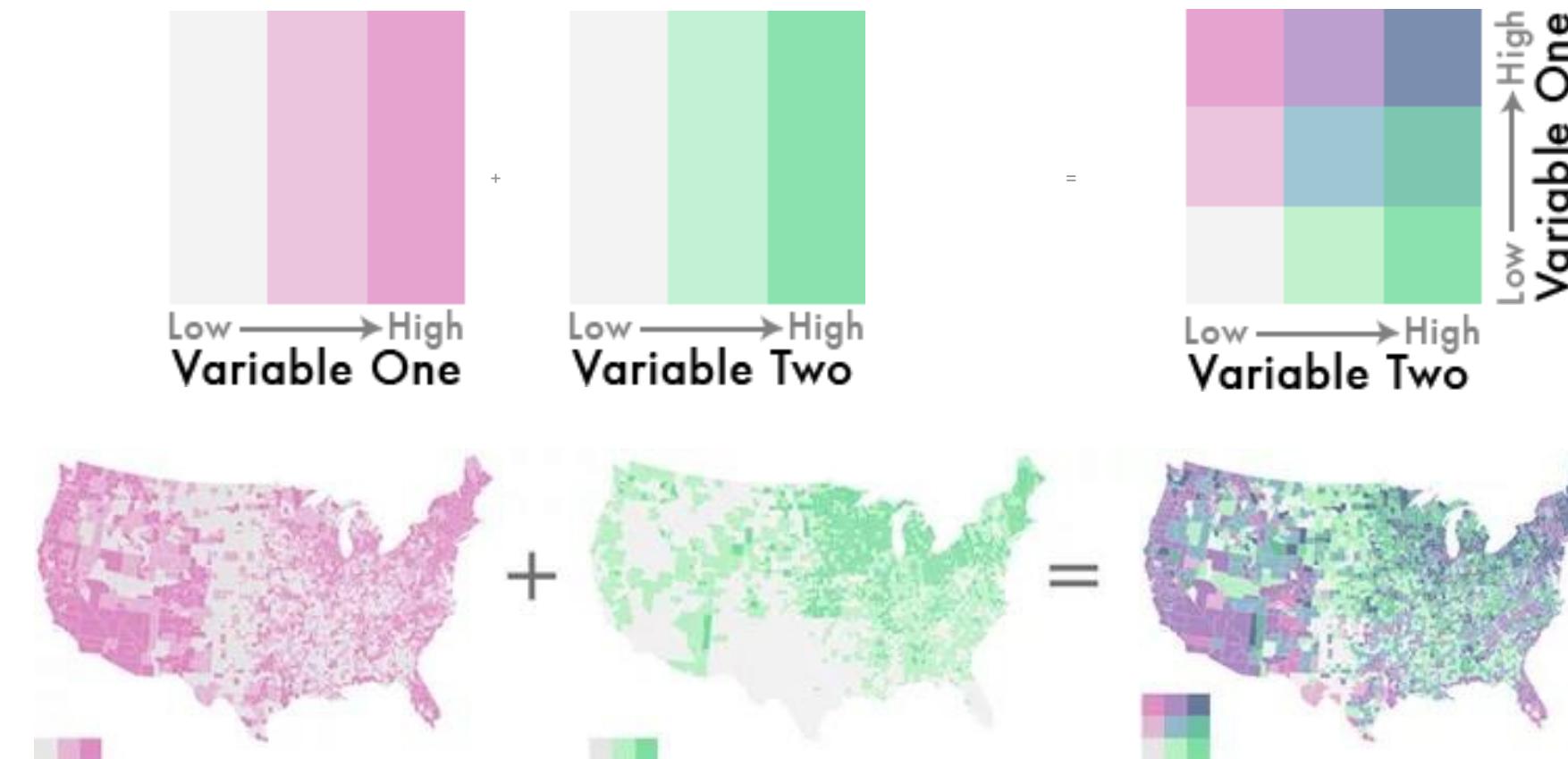


For data with a “diverging” (mid) point (quantitative data)

Equal emphasis on mid-range critical values and extremes at both ends of the data range

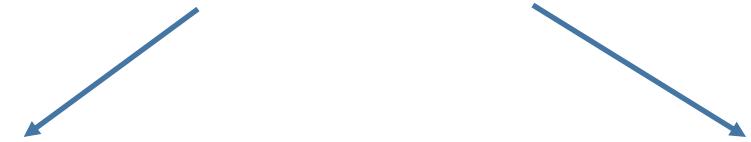
# Application: Example

- Displays two variables
- Combination of two sequential color schemes
- *These are very difficult to design effectively, make intelligible, and be color blind friendly.*



# Luminance

- need luminance for edge detection
  - fine-grained detail only visible through luminance contrast
  - legible text requires luminance contrast!



Luminance information



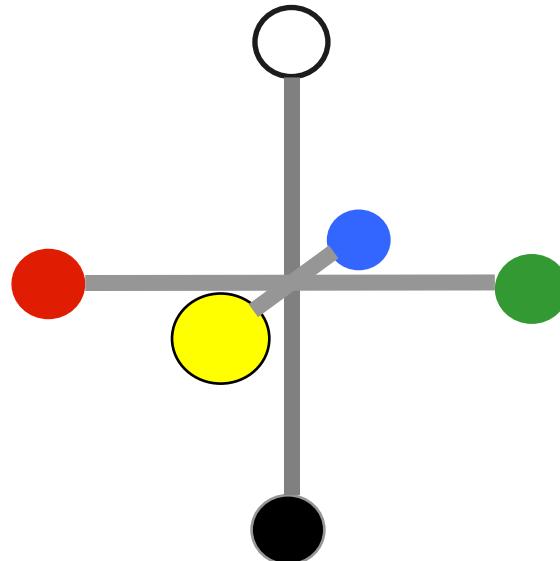
Saturation/hue information



*[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]*

# Opponent color and color deficiency

- perceptual processing before optic nerve
  - one achromatic luminance channel ( $L^*$ )
    - edge detection through luminance contrast
  - 2 chroma channels
    - red-green ( $a^*$ ) & yellow-blue axis ( $b^*$ )
- “colorblind”: degraded acuity, one axis
  - ~10% of men are red/green color deficient
  - blue/yellow is rare



Luminance information

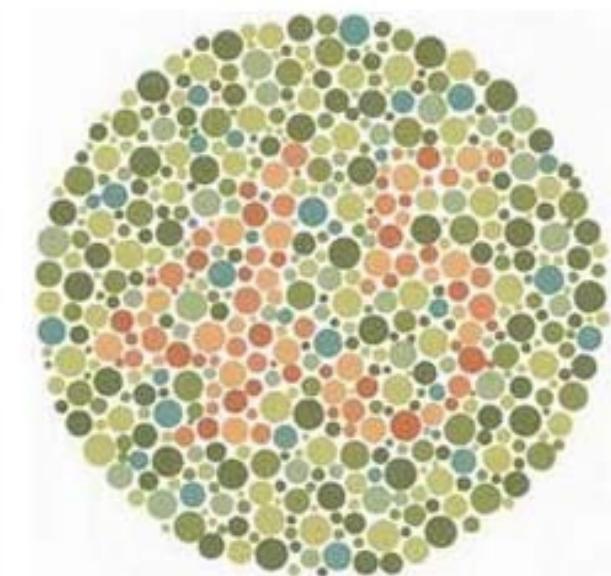
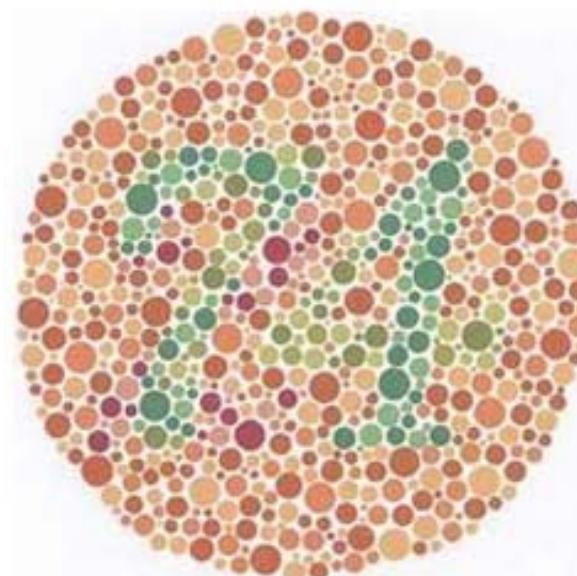
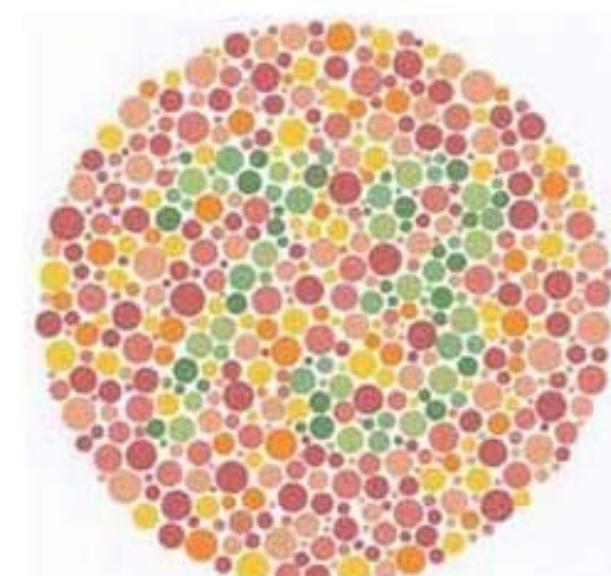
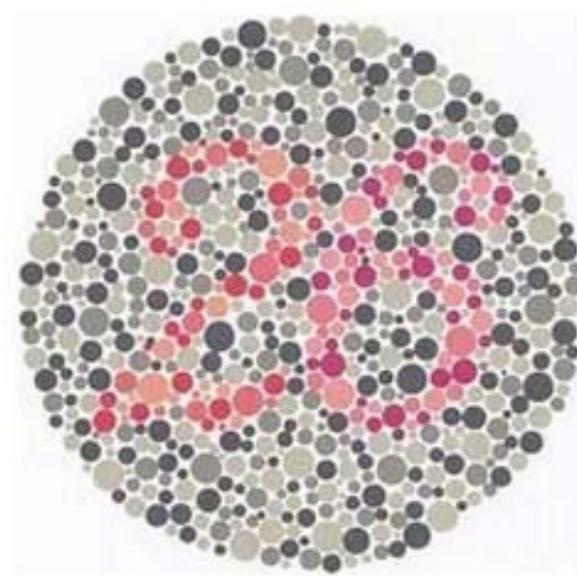
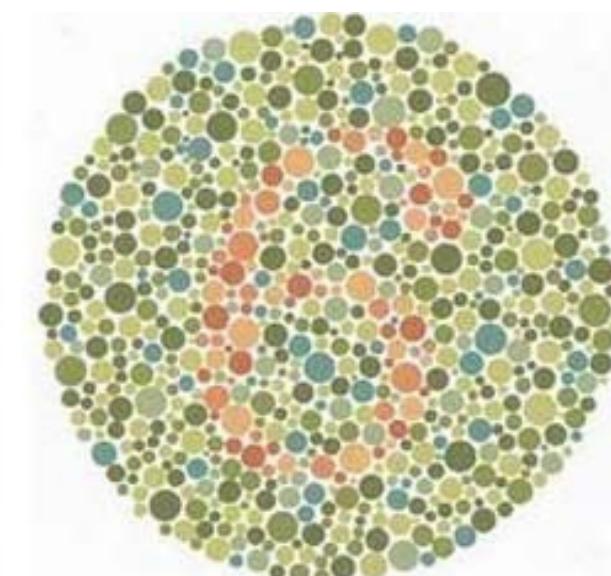
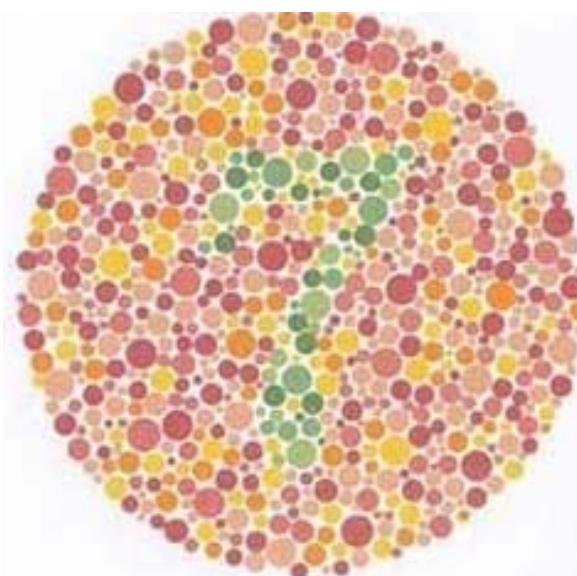
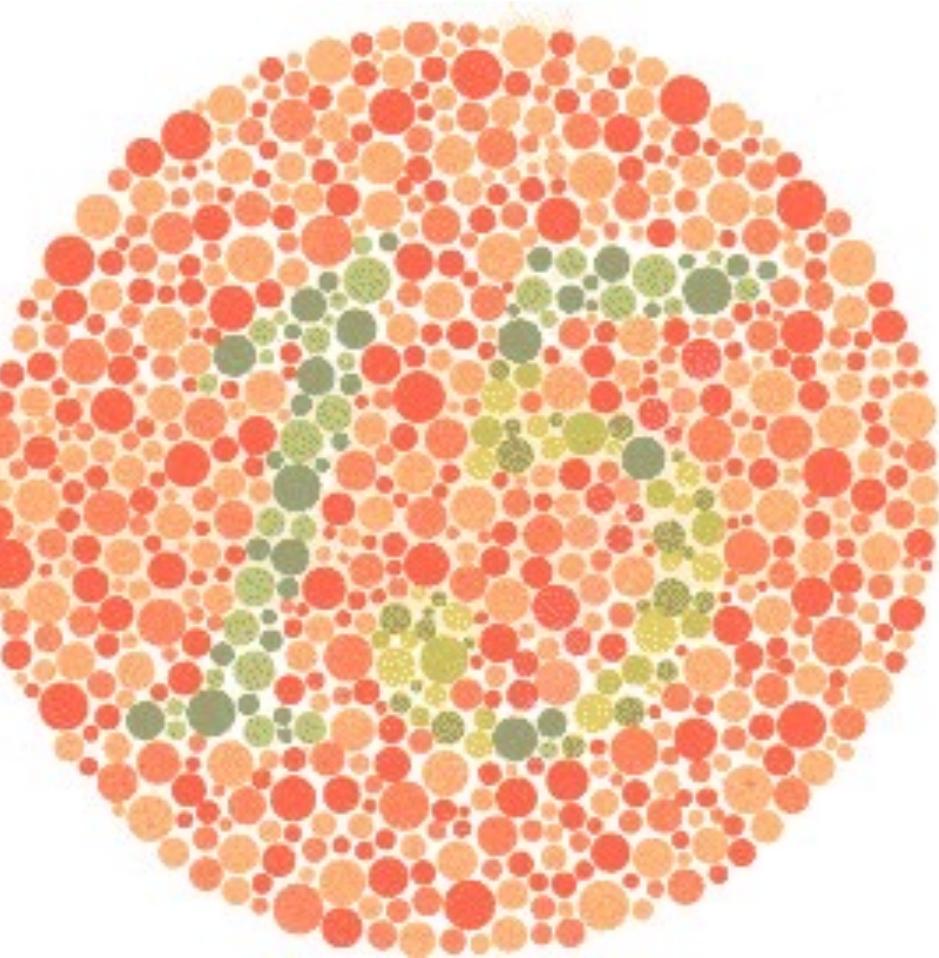
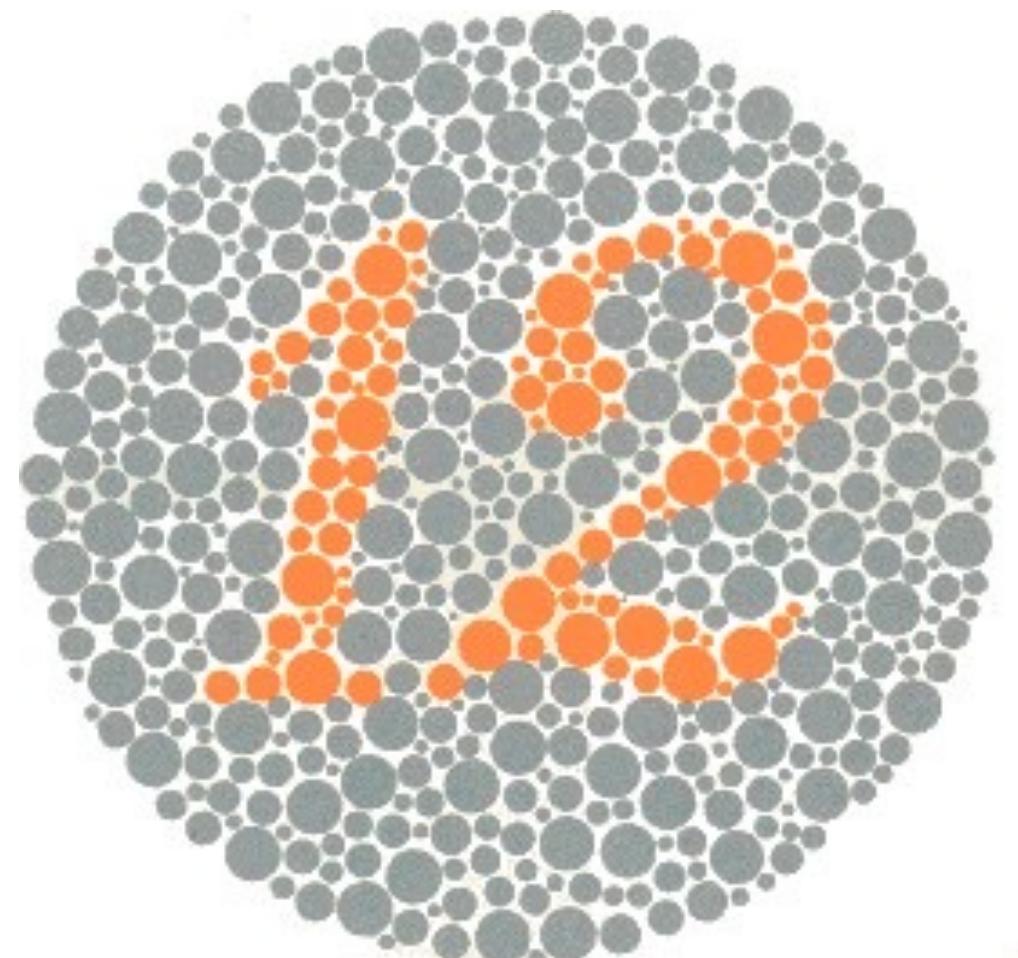


Chroma information



[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

# Color Deficiency Tests



# Color Deficiency

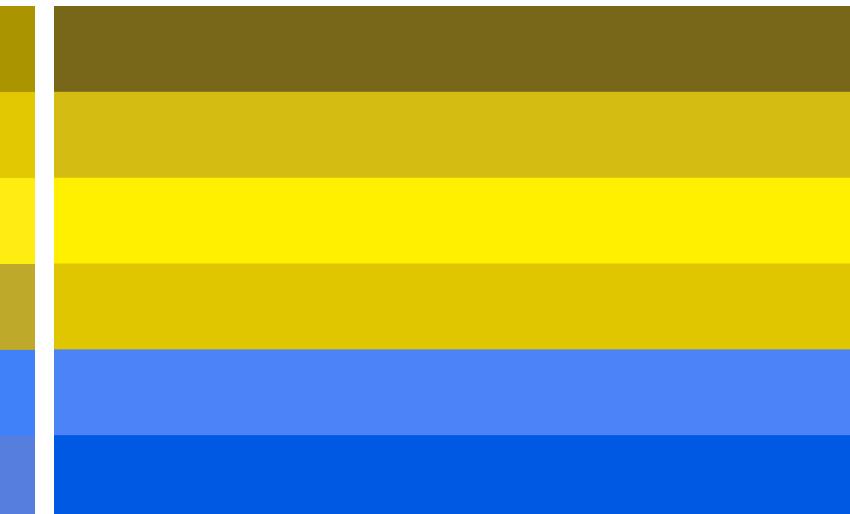
- ~10% of males, 1% of females (probably due to X-chromosomal recessive inheritance)
- Most common: red-green weakness / blindness
- Reason: lack of medium or long wavelength receptors, or altered spectral sensitivity (most common: green shift)



Normal Color Perception

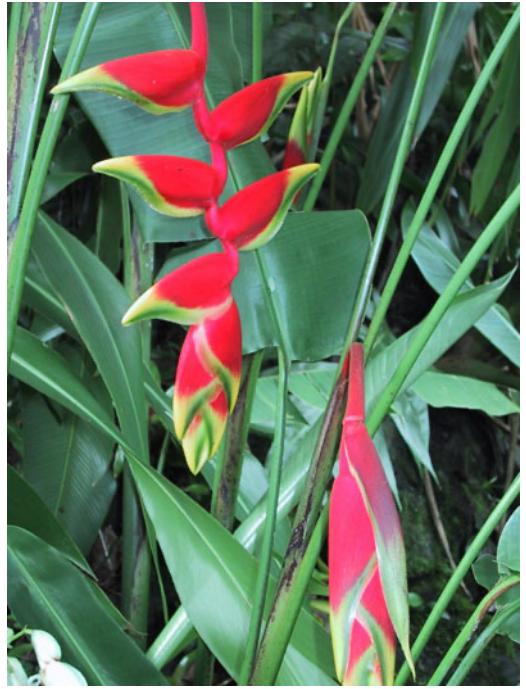


Deutanopia (no green receptors)



Protanopia (no red receptors)

# Designing for color deficiency: Check with simulator



Normal vision



Deutanope  
green-weak



Protanope  
red-weak

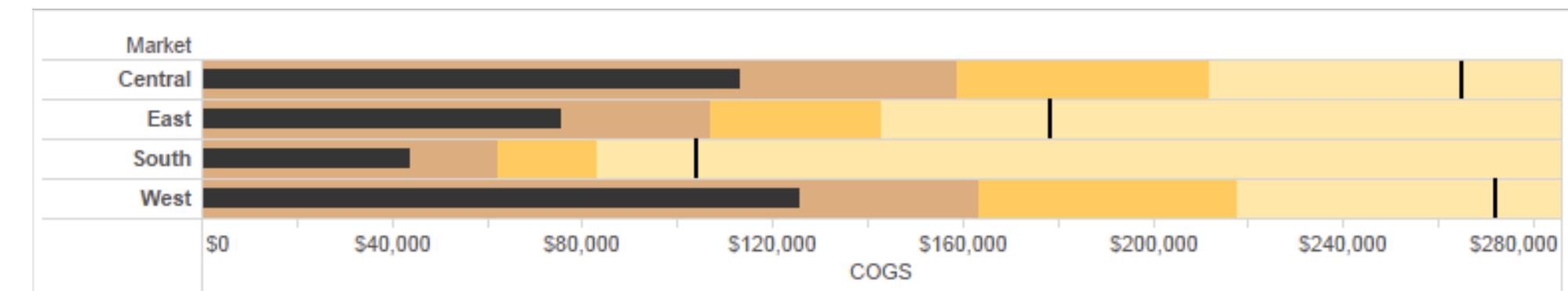
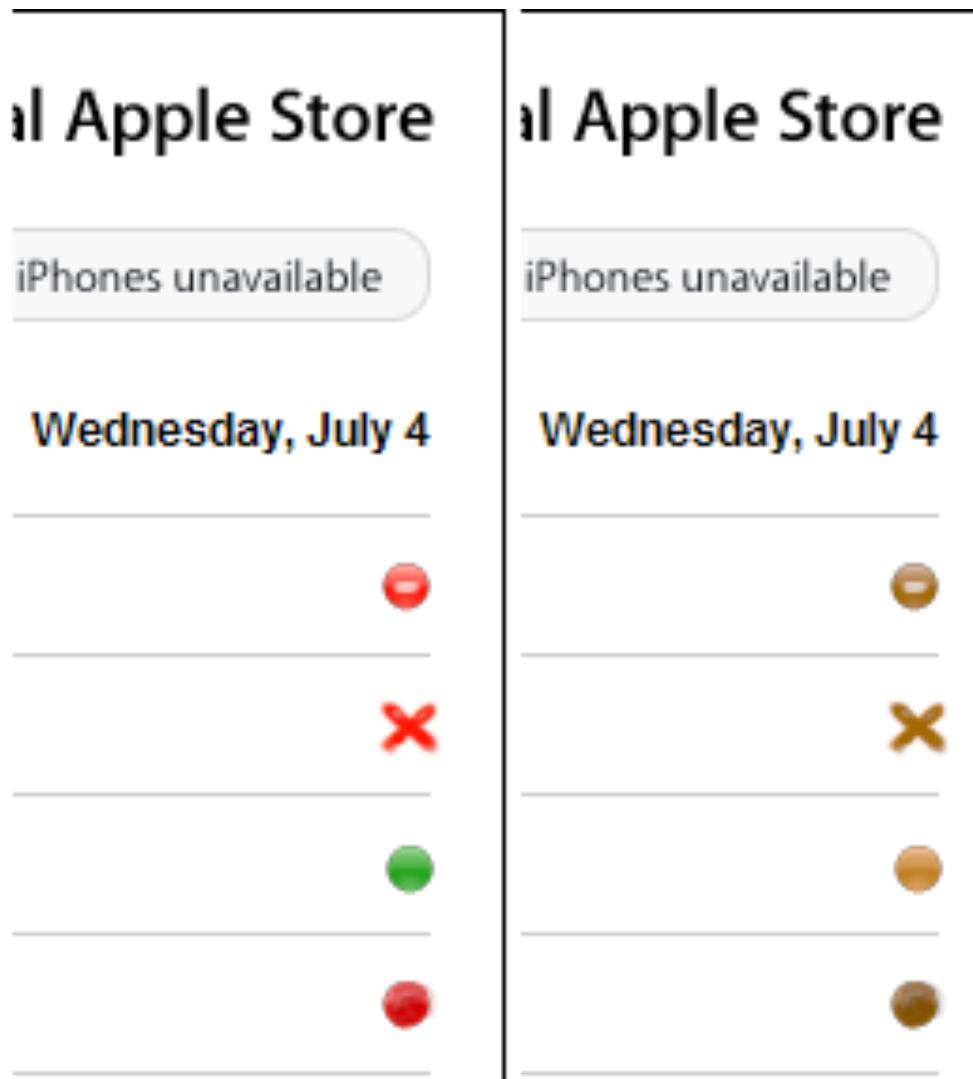
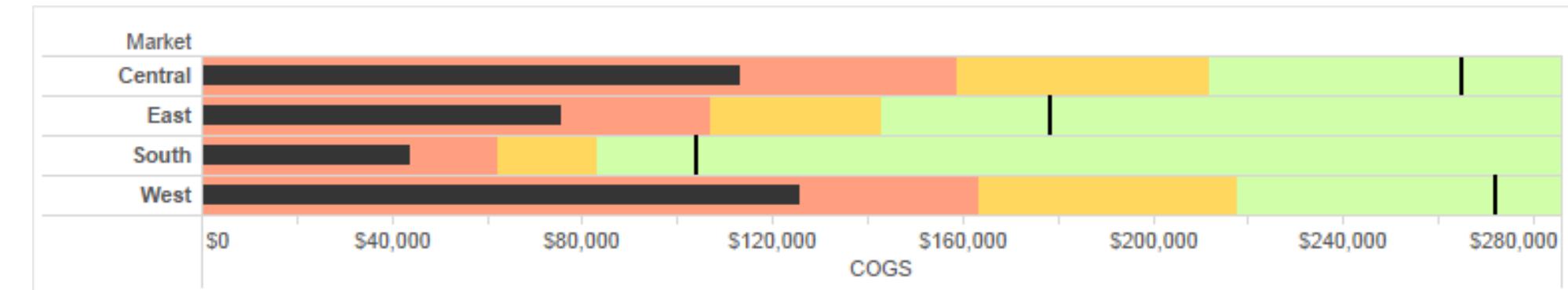


Tritanope  
blue-weak



# Designing for color deficiency: Avoid encoding by hue alone

- redundantly encode
  - vary luminance
  - change shape

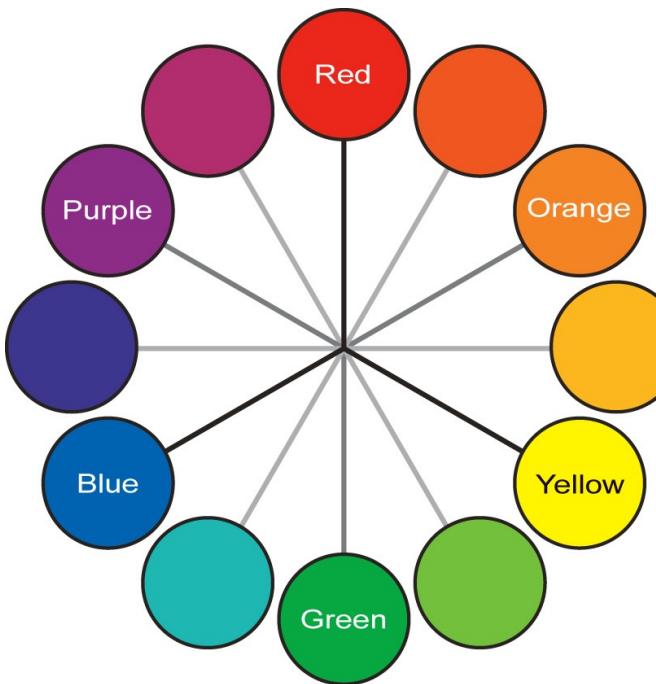


Deuteranope simulation

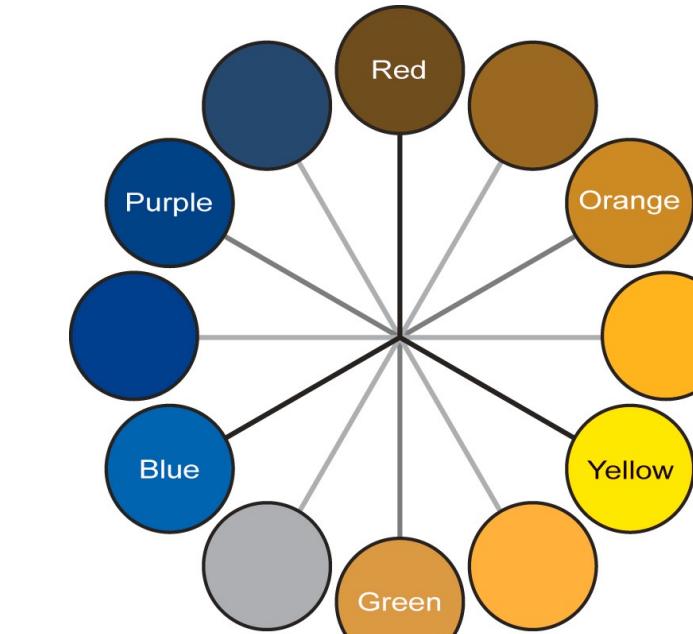
Change the shape

Vary luminance

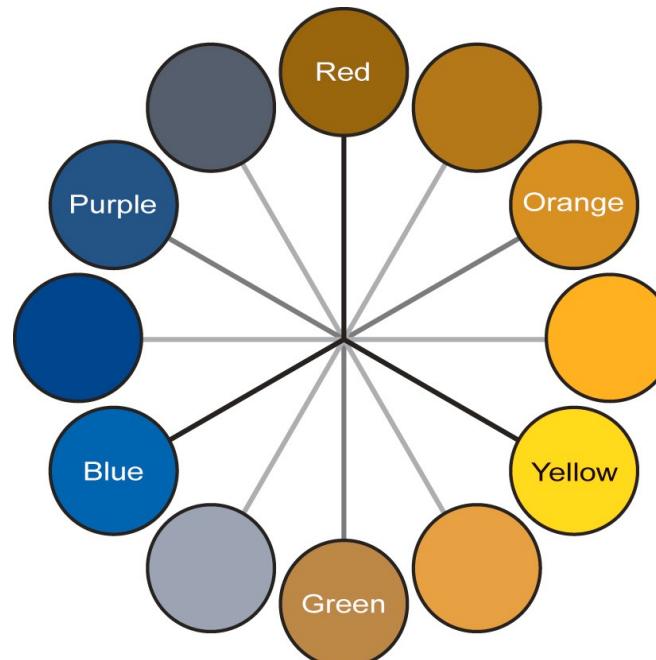
# Color deficiency: Reduces color to 2 dimensions



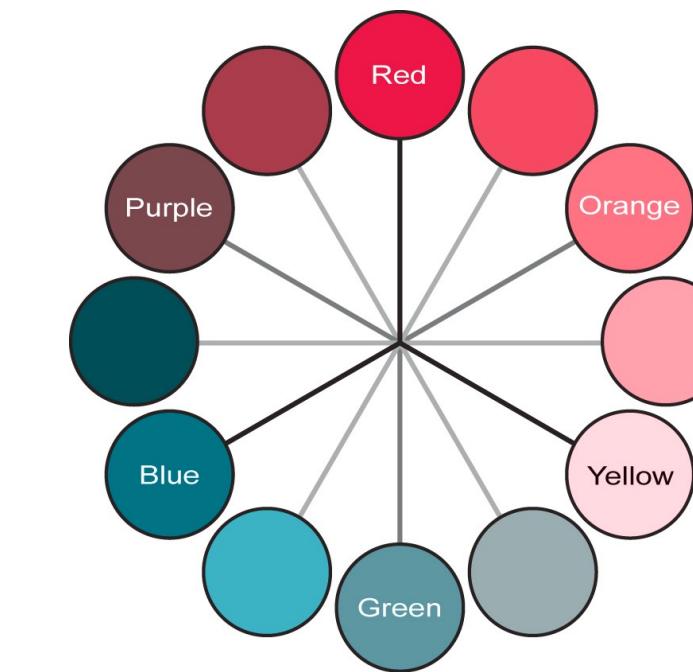
**Normal**



**Protanope**



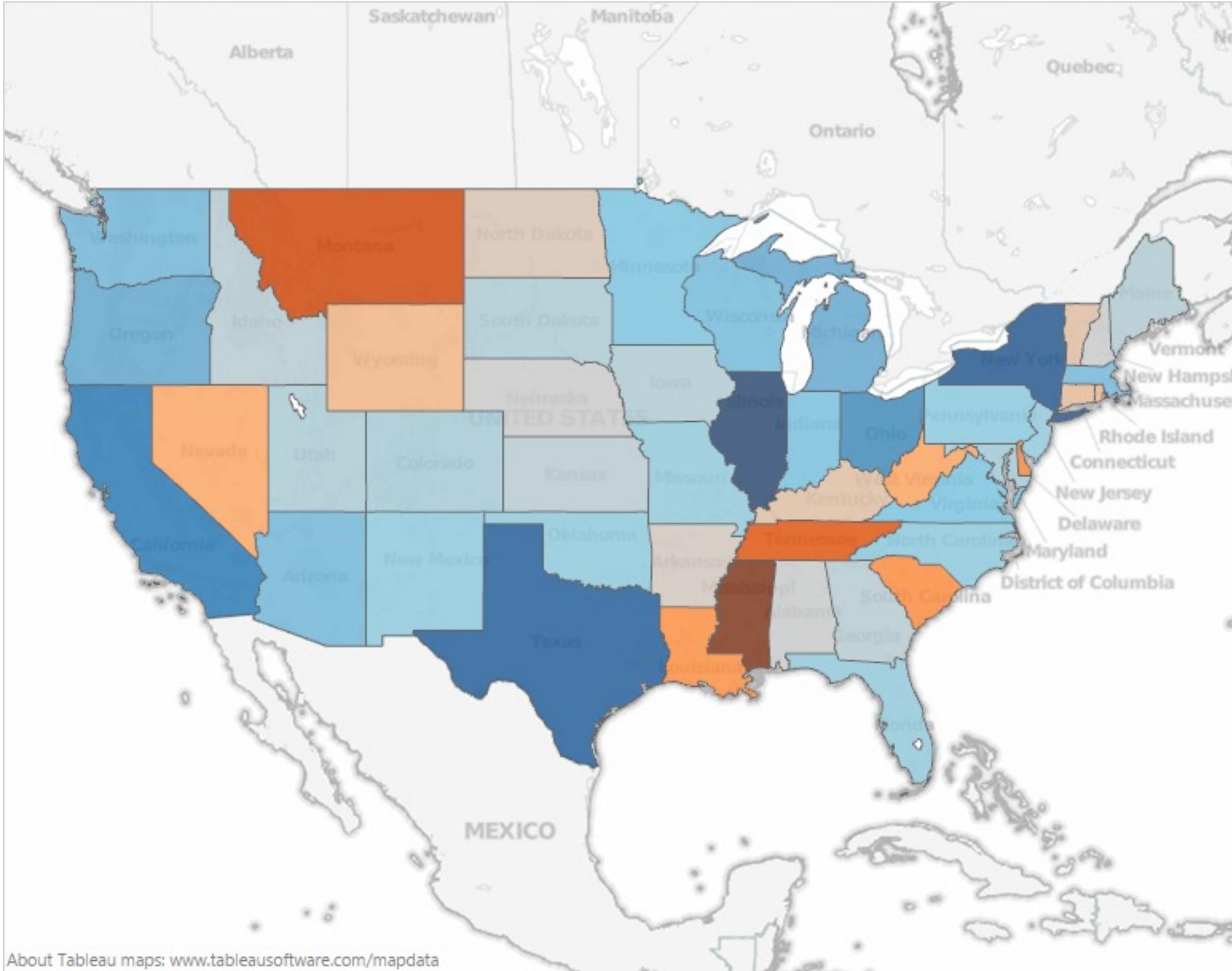
**Deuteranope**



**Tritanope**

[Seriously Colorful: Advanced Color Principles & Practices. Stone.Tableau Customer Conference 2014.]

# Designing for color deficiency: Blue-Orange is safe



# What we see, what others see

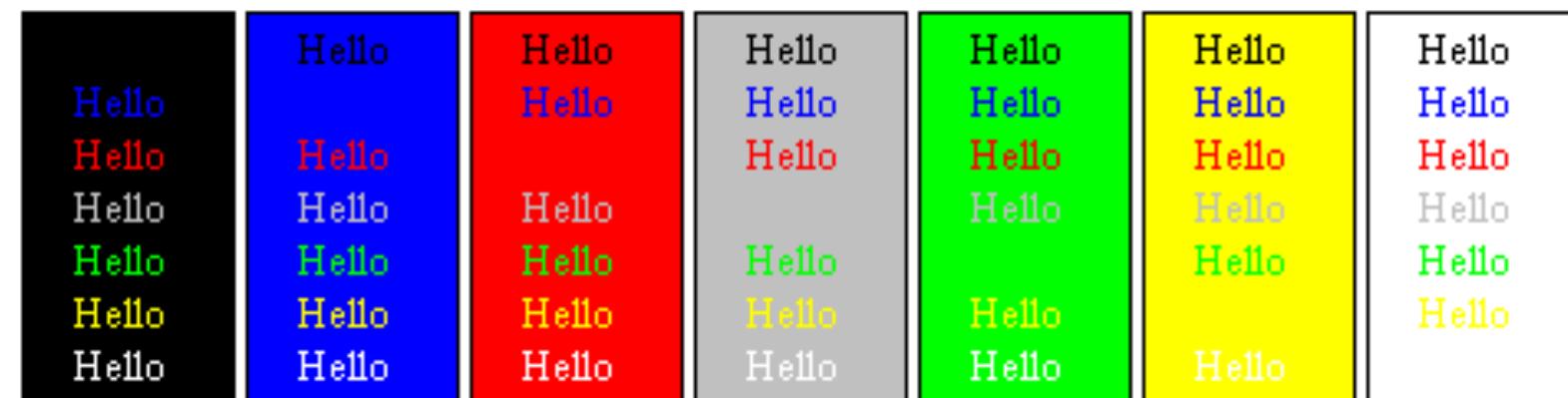


# Luminance Contrast

When adding text to a visualization make sure that there is a distinct difference between the foreground and the background.

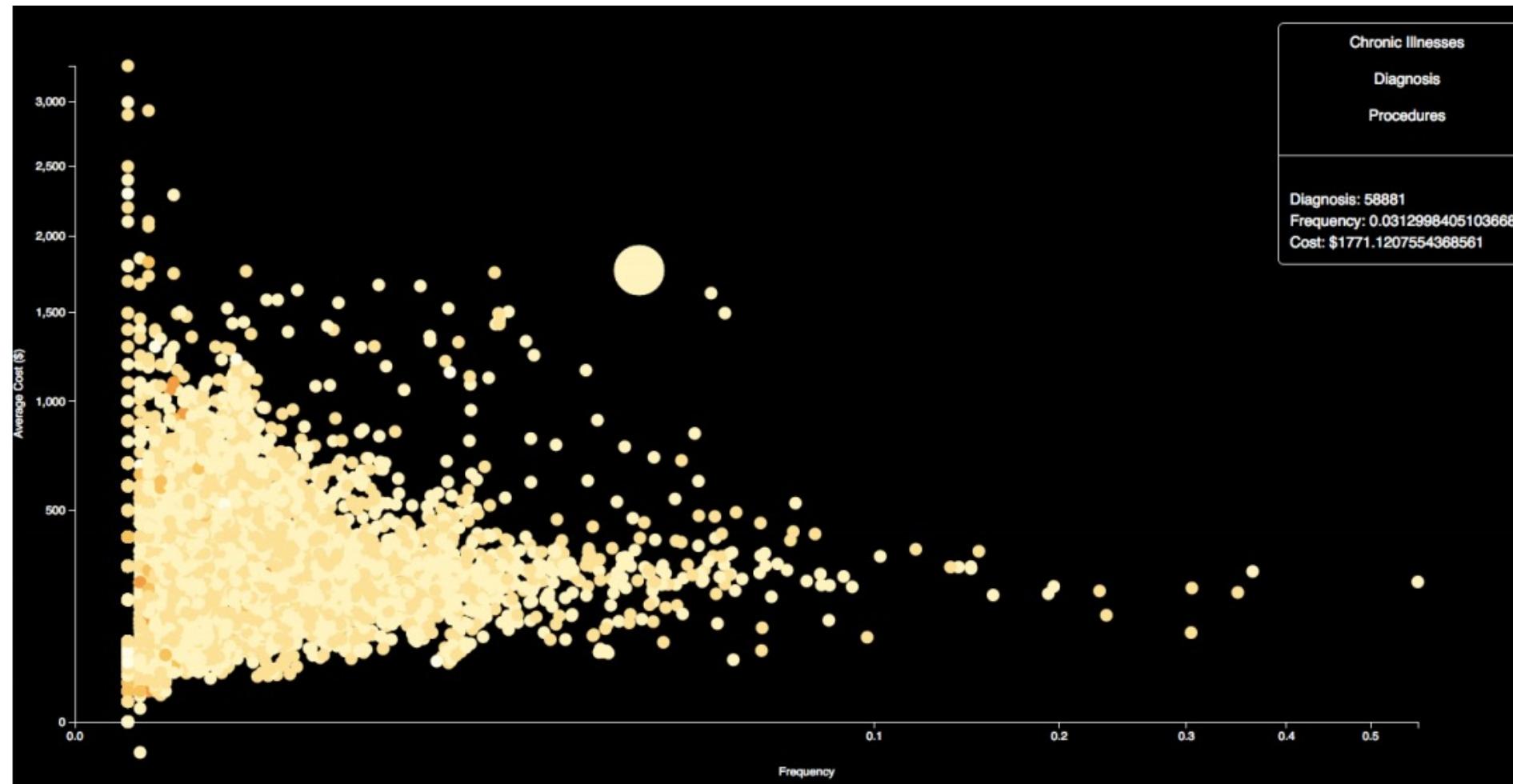
Luminance Contrast is Critical

When you try to use Hue difference for edge detection it is hard to perceive



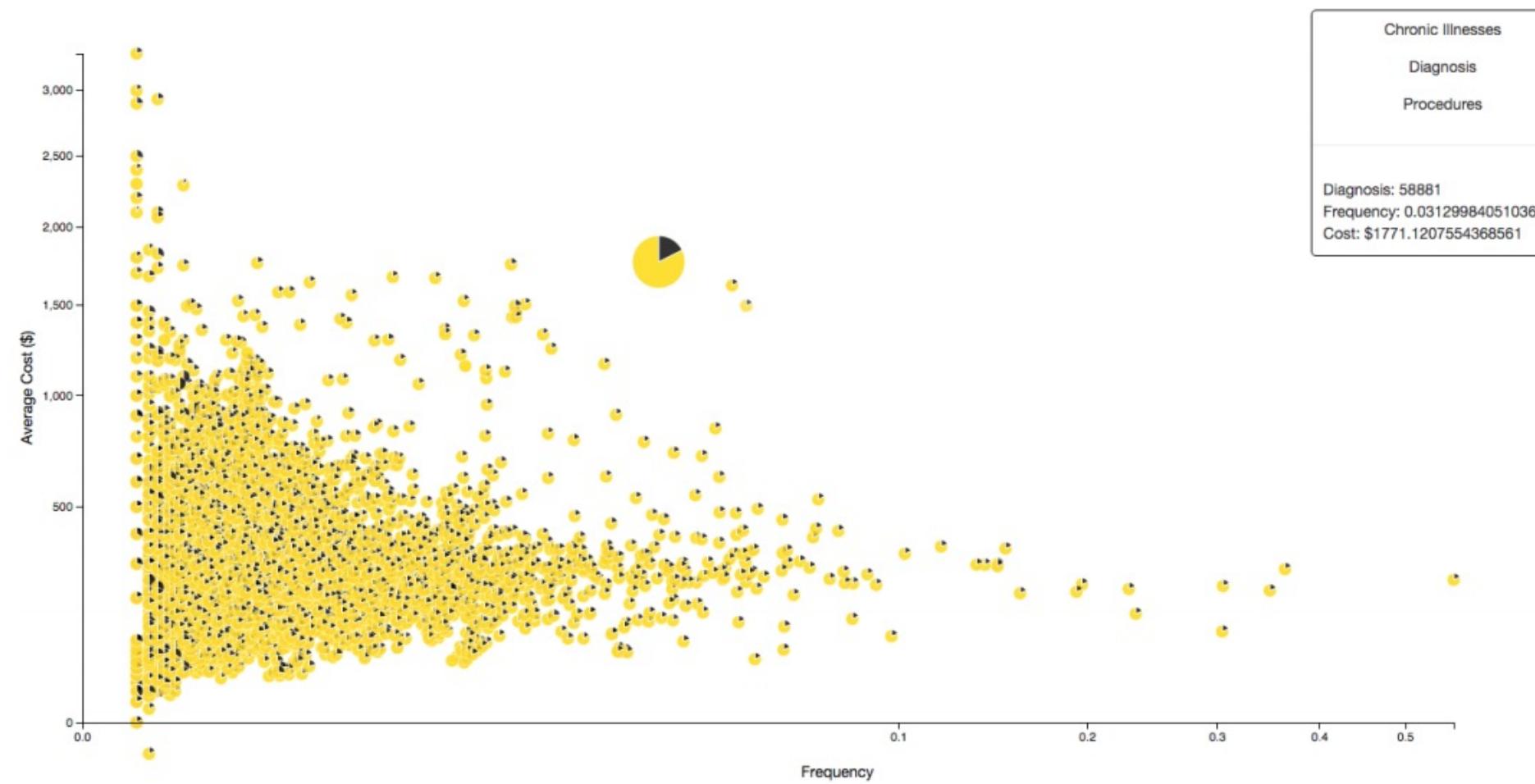
# Interaction with the background: tweaking yellow for visibility

## marks with high luminance on a background with low luminance



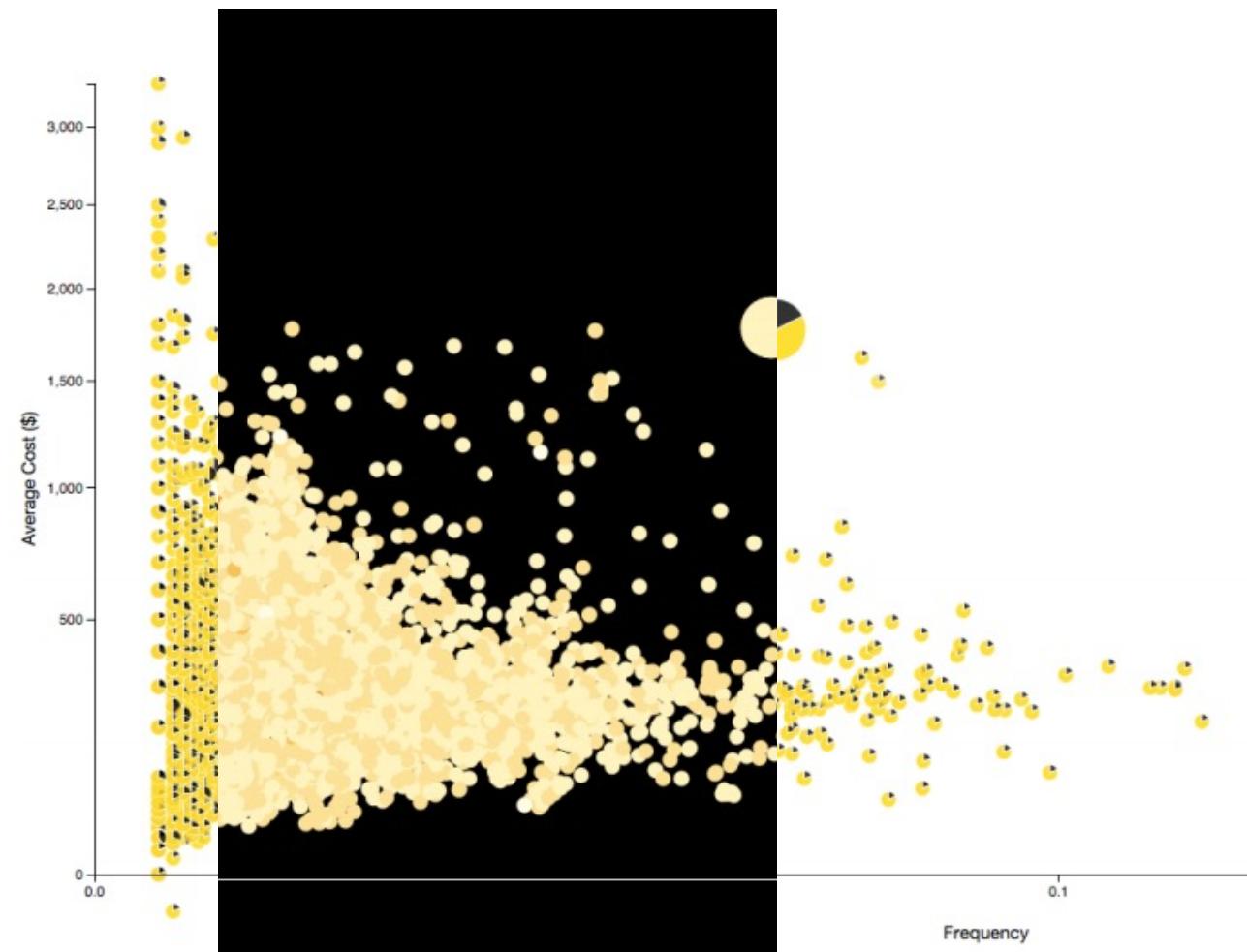
# Interaction with the background: tweaking yellow for visibility

## marks with medium luminance on a background with high luminance

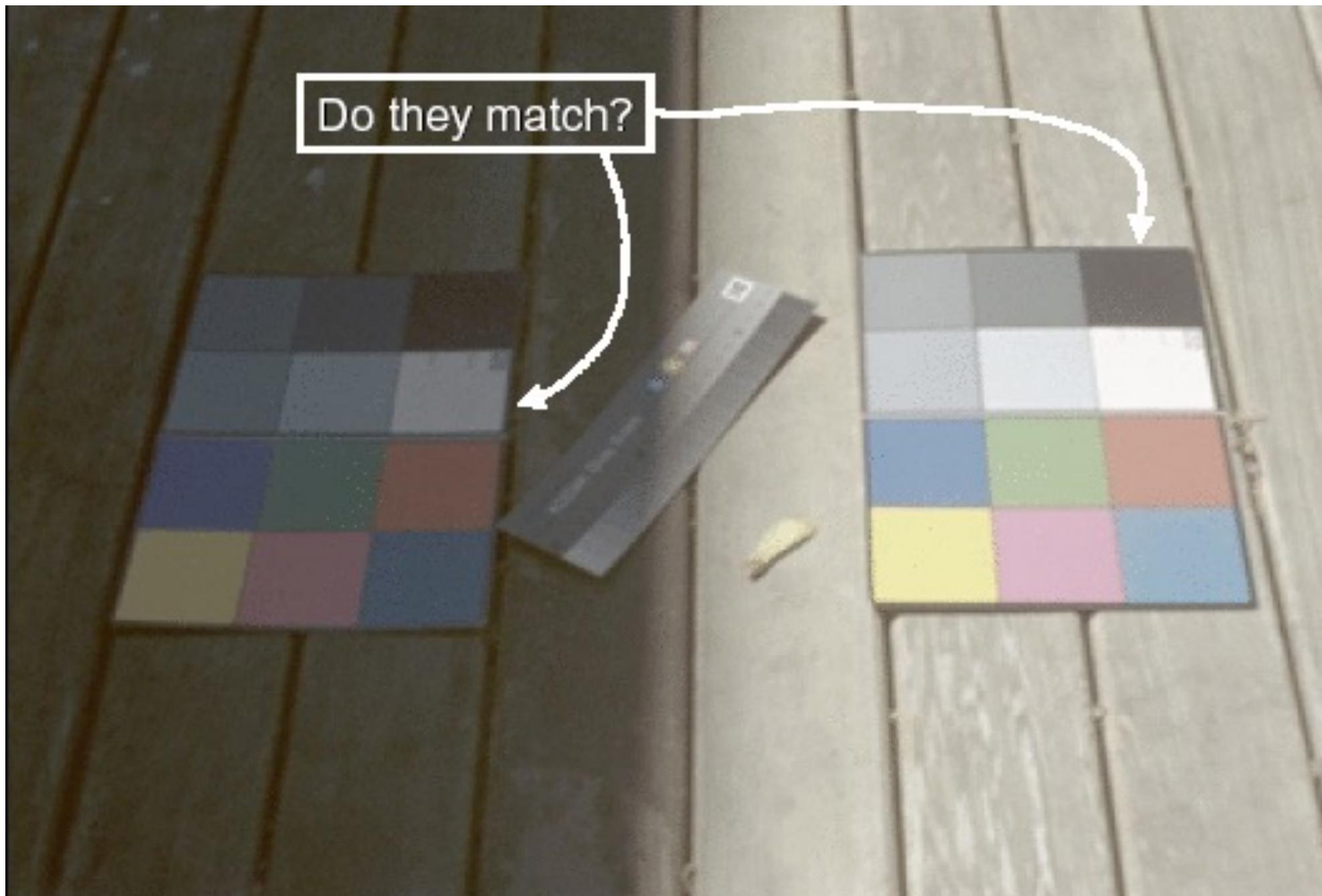


# Interaction with the background: tweaking yellow for visibility

## change luminance of marks depending on background



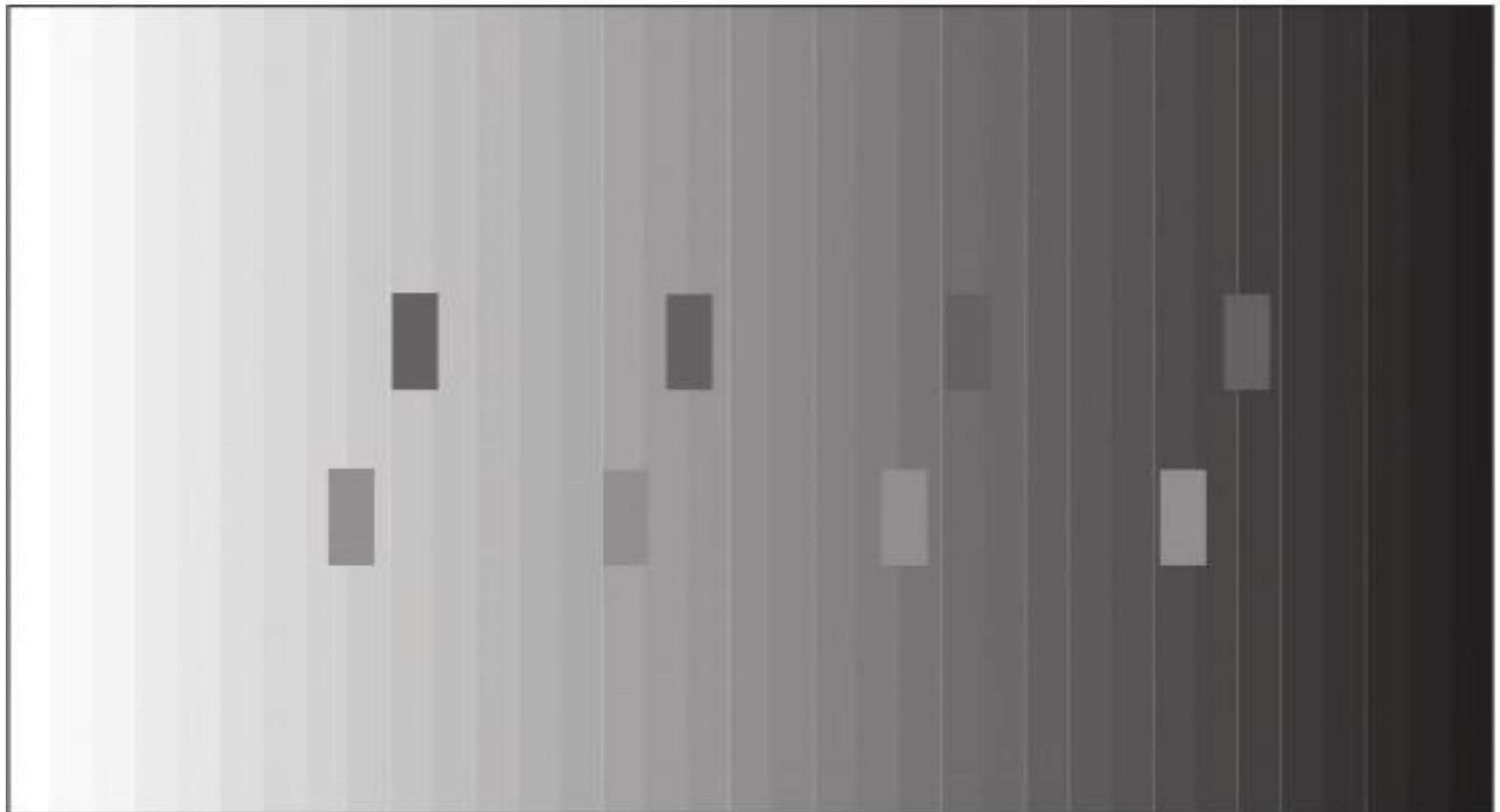
# Color/Lightness constancy: Illumination conditions



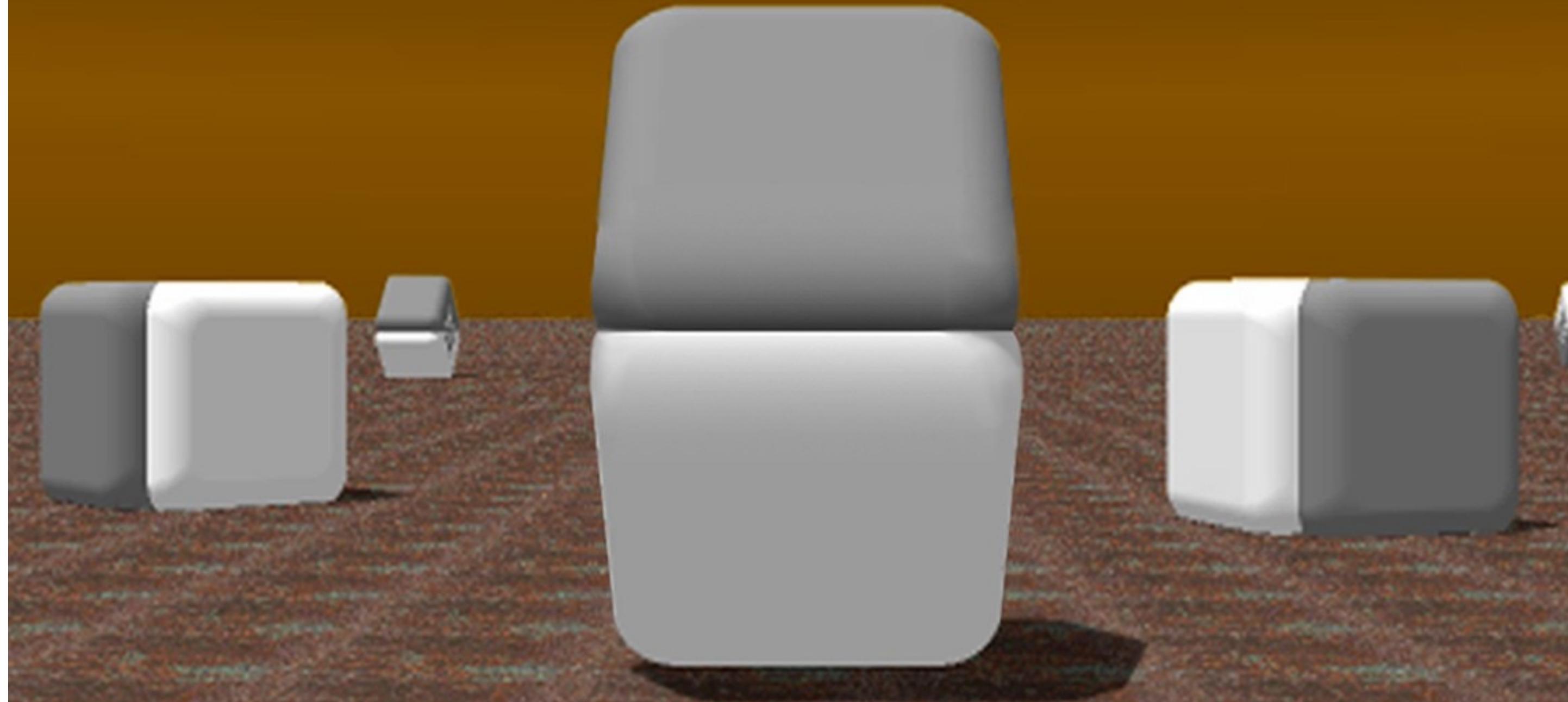
*Image courtesy of John McCann via Maureen Stone*

Which bar is the darkest  
Which bar is the lightest

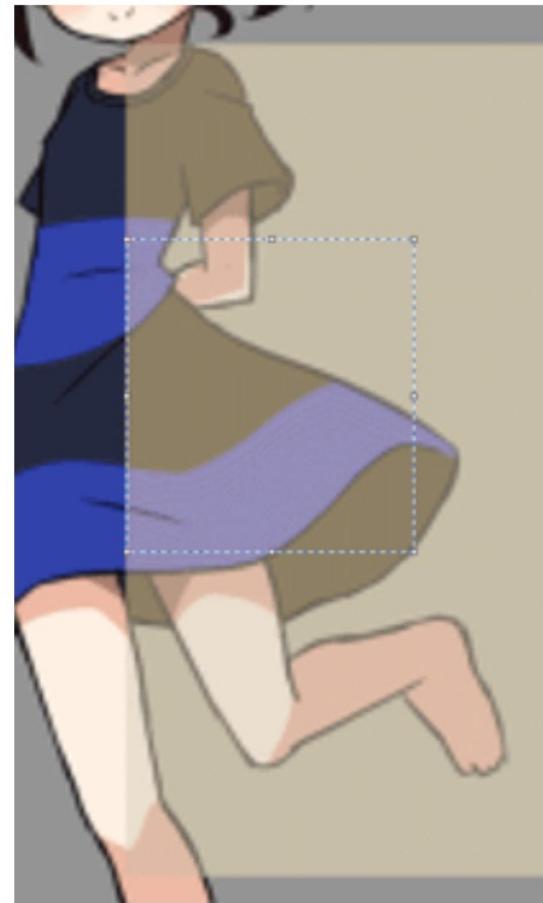
Top Row A B C D  
Bottom Row E F G H



# Cornsweet Illusion



# Contrast with background



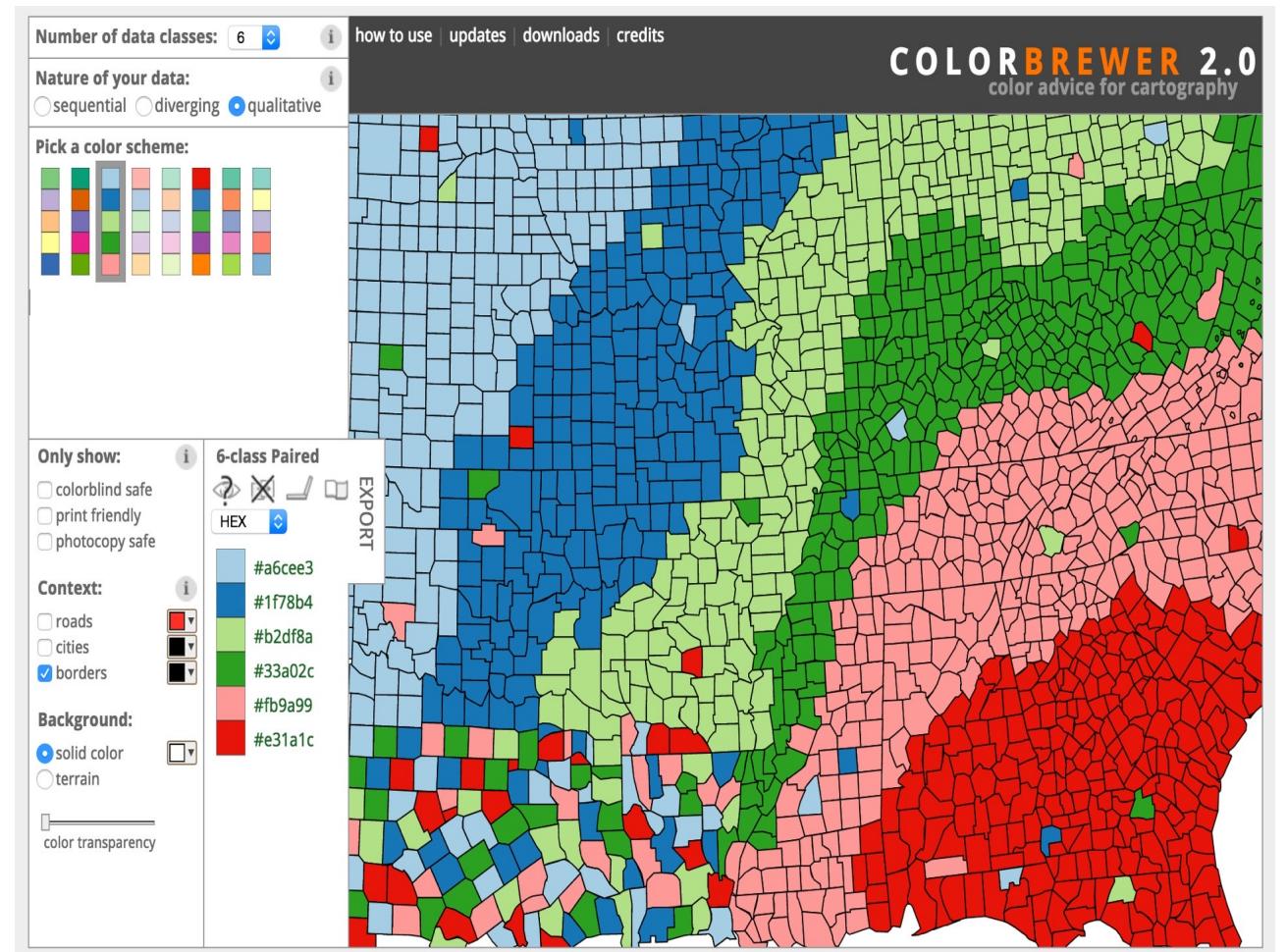
Black and blue? White and gold?

<https://imgur.com/hxJjUQB>

[https://en.wikipedia.org/wiki/The\\_dress](https://en.wikipedia.org/wiki/The_dress)

# Tools for picking Color

- Color Brewer
- Colorgorical
- Programming Libraries



The screenshot shows the Colorgorical interface. On the left, there's a sidebar with a "Generate" button, a color palette icon, and controls for "Number of colors" (set to 5), "Score importance" (with sliders for Perceptual Distance, Name Difference, Pair Preference, and Name Uniqueness), and "Select hue filters" (a circular color wheel with 90°, 180°, and 270° markers). The main area has tabs for "Results: Color space, Hex, RGB, Lab, LCH" and "Array format: No quote". It displays a preview of a 5-color palette with hex codes: "#57146,131", "#148,210,207", "#25,79,70", "#57,238,192", and "#148,210,207". Below this are three charts: a histogram of color distribution, a scatter plot of individual colors, and a bar chart of color scores. The bottom sections include "Instructions" (instructions for generating palettes), "About" (information about the tool's creators), and "Documentation" (links to the paper and GitHub repository).