

Graphical Perception

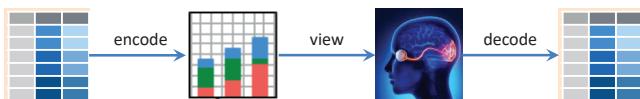
Michael Friendly

Psych 6135

<http://euclid.psych.yorku.ca/www/psy6135/>

Encoding & decoding

- When we construct a graph, we **encode** a numerical or categorical variable as a graphical attribute
- When we view a graph, the goal is to **decode** the graphical attributes and extract information about the data that was encoded

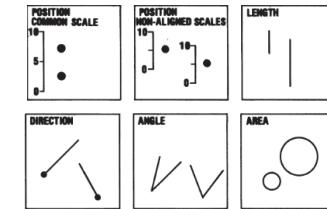


- Encoding should rely on features that can easily be decoded
- Often, easier said than done! The devil is in the details

Graphical Perception

- In constructing a graph, **quantitative** and **categorical** information is encoded by visual attributes:

- Length
- Position along axis
- Angle
- Area
- Color, shape, line style



- What determines the ability of graph viewers to:
 - Make **comparisons** (which is larger?)
 - Estimate a magnitude?
 - See **patterns**, trends, unusual features?

Visual & cognitive systems

- A simplified model: Three stages

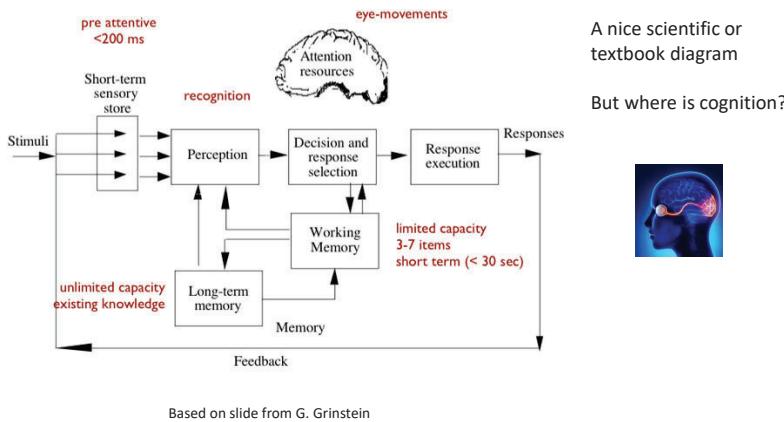
How is info lost?

How to preserve info?
- Sensory (iconic) memory
 - pre-attentive, automatic, feature detection
 - massively parallel, short duration, easily fooled ("thinking fast")
- Working memory
 - requires attention, limited capacity (~ 4-6 "chunks")
- Long-term memory
 - real-world knowledge, ~ unlimited capacity, inference ("thinking slow")

Perception vs. cognition

Another coarse distinction:

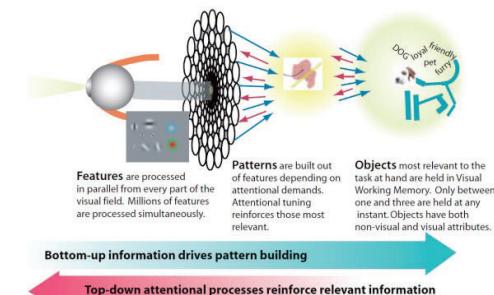
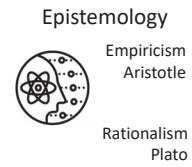
- **Perception:** Processing of the signals coming in: what you “**see**”
- **Cognition:** How you **understand** and **interpret** what you see



5

Perception: Bottom-up & Top-down

- **Bottom-up processing**
 - Low level: features → pattern → object
 - Detect edges, contours, color, motion
- **Top-down processing**
 - Driven by goals, expectations
 - Uses prior knowledge, experience, filters what we “see”



6

Perception: Bottom-up

How many 5s in this display?

1561321203658413076510374627
4173127527327592732990709742
1703707774179527931749270973
4019743217909370945179279417

How many 5s in this display?

1561321203658413076510374627
4173127527327592732990709742
1703707774179527931749270973
4019743217909370945179279417

Numerals differ only in **shape**, and are high-level symbols

You have to literally scan them **all** & count the 5s.

The distinction of **color** is immediate & **pre-attentive**

You only have to scan & count the 5s.

This is why **color** is an important visual attribute for a **categorical** variable in graphs

7

Perception: Top-down

What is in this scene?



What is the middle character?

A B C
I2 B I4

What here?

B

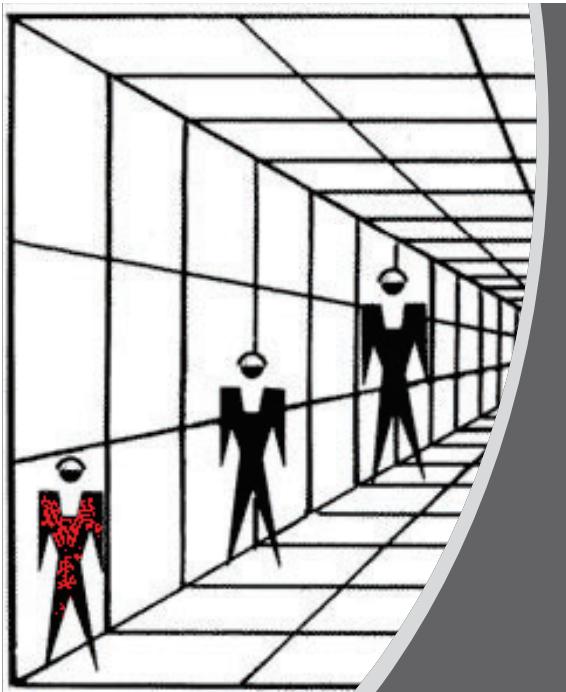
An ambiguous figure!

What is the middle letter in each word?

T A E C A T

All of these are demonstrations of the role of **expectations** (top-down) in determining what we “see”

8



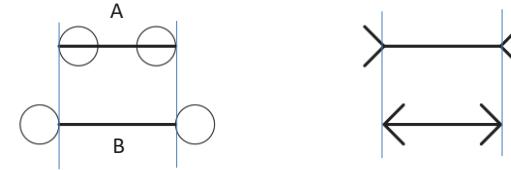
9 Illusions: The Eye-Brain Barrier

Perceptual illusions give some guidance on what **not to do** in data graphics

10 Illusions: Length

Surrounding **context** matters in judging the **length** of objects.

Which **line** is longer? Or are they the same?



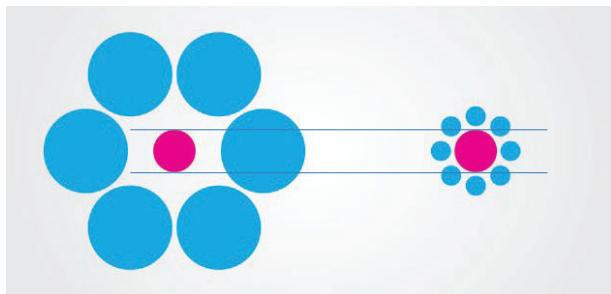
Surrounding context pulls perception of length in its direction
This is the famous **Müller-Lyre** illusion

10

11 Illusions: Area

Surrounding context matters in judging the **area** of objects.

Which **red** circle is larger? Or are they the same?

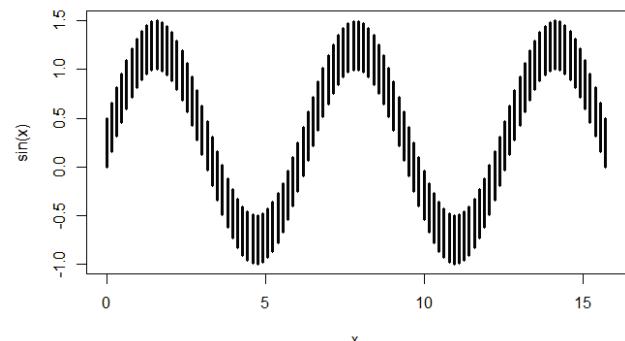


Surrounding context pulls perception of area against the background
This is often called the **Ebbinghaus** illusion or the **Tichener** illusion

11

12 Illusions: Length

Which of the bars are longer? Or, are they all the same length?



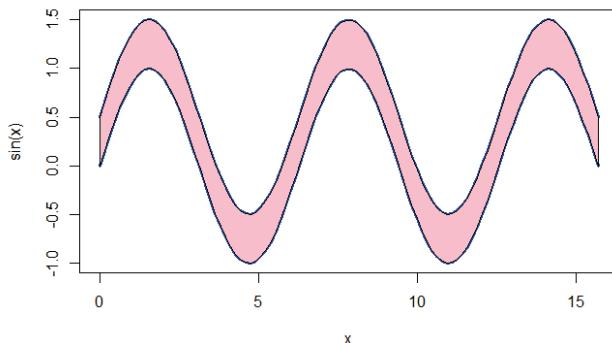
R code:

```
x <- seq(0, 5 * pi, length.out = 100)
w <- 0.5
plot(x, sin(x), ylim = c(-1, 1 + w), type = "n")
segments(x0 = x, y0 = sin(x), y1 = sin(x) + w, lwd = 3)
```

12

Illusions: Difference

Where are **differences** between curves larger? Or, are they all the same?

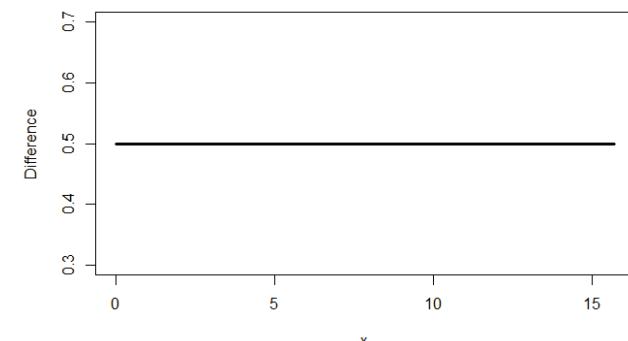
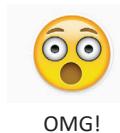


This is sometimes called the “sine illusion”

13

Illusions: Difference

Plotting the difference directly gives the answer.

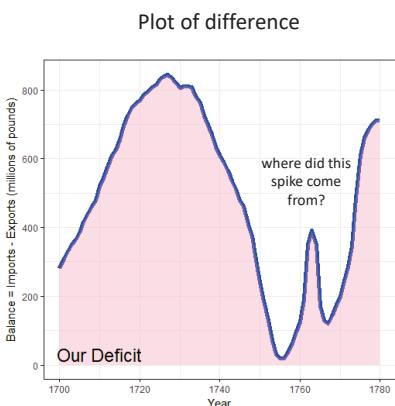
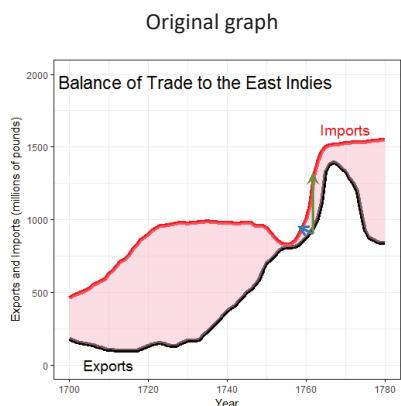


Why does this matter?

14

Distances between curves

Playfair didn't know that judgments of distance between curves are **biased**. We tend to see the **perpendicular** distance rather than the **vertical** distance

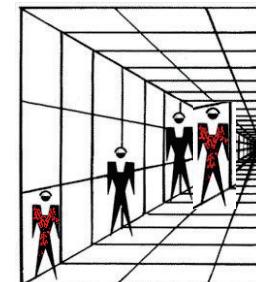
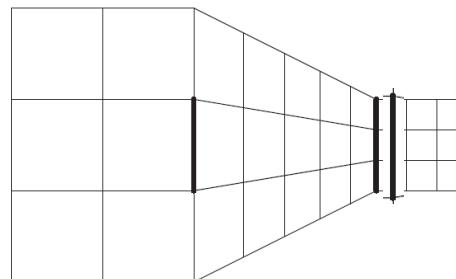


15

Illusions: Perspective

Which **thick** line is longer? Or, both the same?

Which figure is tallest?
Or, all the same?



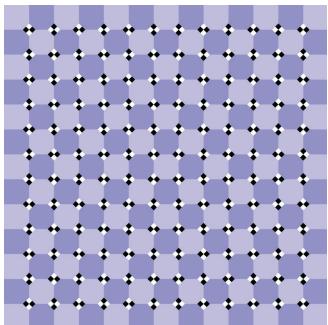
This is often called the **Ponzo** illusion: We judge the **size** of real-world objects relative to their background and perspective.

16

Context illusions

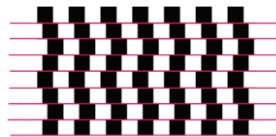
Perception of elements of a scene is affected by context, background, etc.

Are the squares straight or tilted?



It is hard not to be fooled by these!

Are the pink lines straight or curved?



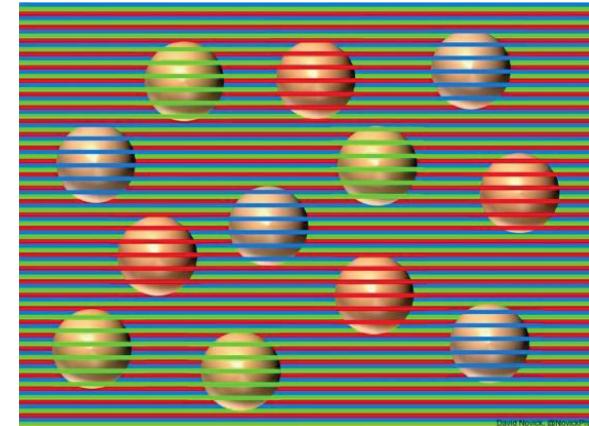
Does the rectangle change in darkness?



17

Context illusions: Color

Are the balls different colors or are they all the same color?



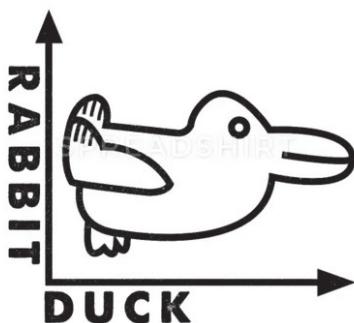
David Novick (@davidnovick)

18

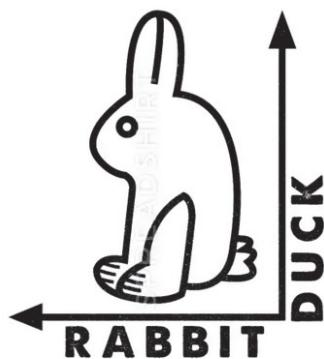
Illusions: Semantic/cognitive

Perception of object figures often shows a preference for **orientation** in nature

Is this image a duck, or a rabbit?



Duck or rabbit?



19

Catalog of visual/auditory illusions

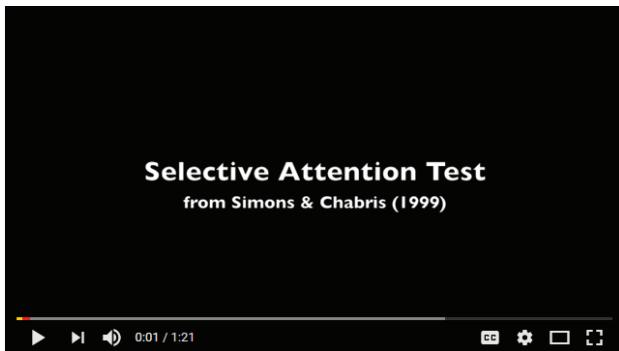
<https://www.illusionsindex.org/illusions>

A large collection of illusions with references to research studies.

The screenshot shows the homepage of The Illusions Index. At the top, there's a navigation bar with links for HOME, ABOUT, EXPLORE, SUGGEST, DONATE, FEEDBACK, and QUIZ. Below the navigation is a search bar labeled "Filter by Tags" and a dropdown menu for "Sort By" with "Most Visited" selected. A message says "Found 63 matching results". On the left, there's a thumbnail for the "YOUNG WOMAN OR OLD WOMAN" illusion. In the center, there's a box for "EXPLORE ILLUSIONS" with text about the site being a searchable curated collection of illusions. To the right, there are thumbnails for "ADELSON'S CHECKER-SHADOW ILLUSION", "TROXLER EFFECT", and "NEGATIVE AFTERIMAGES". At the bottom right, there's a grid of small squares representing a "WATERFALL ILLUSION". The page number "20" is at the bottom right corner.

20

Selective attention



<https://www.youtube.com/watch?v=vJG698U2Mvo>

Attention strongly focused on some feature(s) steals attention from others

21

Magnitude estimation

How large are transport accidents?

How much bigger than non-transport accidents?



Estimation of **length** or ratios of length are more accurate than the same judgments of **area**.

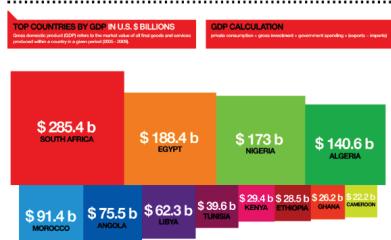
22

Area vs. length judgments

Easy: Which is larger— South Africa or Egypt?

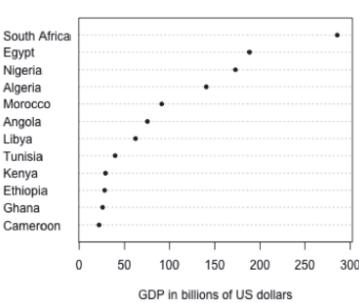
Harder: How much larger is South Africa than Egypt? (% or ratio)

African Countries by GDP



Judgments here based on area

African Countries by GDP



Judgment here based on position along a scale

23

Stevens' Power Law

- How does perceived magnitude of a sensation relate to stimulus intensity?
- S. S. Stevens (1957) showed that, for many domains

$$\text{Sensation} \propto \text{Intensity}^p$$

- These provide ways to assess the **accuracy of magnitude estimation** for visual encodings
 - length judgments most accurate
- But: graph perception is not always a matter of estimating magnitudes.

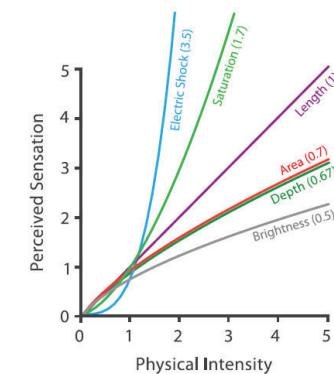


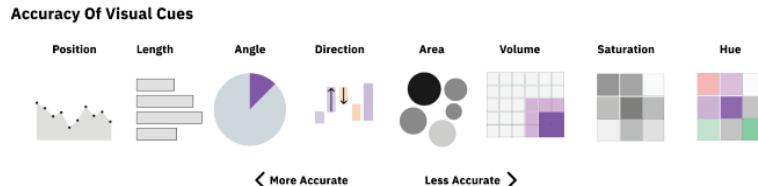
Fig. 5.7 from: Munzner, *Visualization Analysis & Design*

24

Scale of accuracy

The commonly used “scale” of accuracy of magnitude judgments of relative size

- How much smaller/larger is A compared to B?



⚠ Not necessarily the same for other tasks (Part-whole: What % is A of total?)

25

Accuracy: Experimental evidence

Cleveland & McGill (1984) and later Heer & Bostock (2010) carried out experiments to assess the relative accuracy of magnitude judgments for different visual encodings

The task here is to estimate the %age of the smaller highlighted portion.

The details of these studies are interesting & important – more next week

The graph of these results is a great model for data display

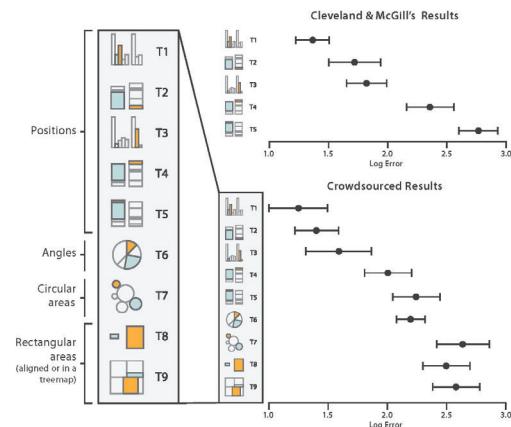


Fig. 5.8 from: Munzner, *Visualization Analysis & Design*

26

Encodings: Types & ranks

Based on this, Munzner (2015) proposes a ranking of visual attributes for **ordered** & **categorical** variables in data displays

These hold when the task is to estimate a **magnitude**.

A different ranking may occur for other graph-based tasks.

angle (pie charts) – good for % of total judgments

color (mosaic plots) – good for pattern perception

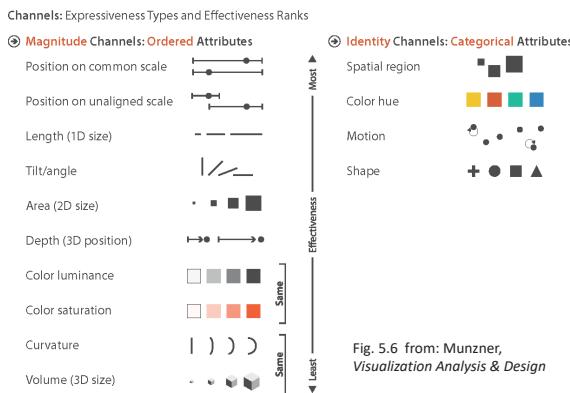


Fig. 5.6 from: Munzner, *Visualization Analysis & Design*

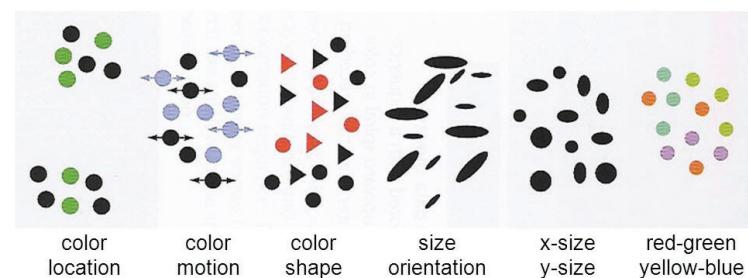
27

Integral & separable encodings

- Some encodings can be viewed **independently**
 - two different variables **can** be decoded separately
- Some **combine** with each other to some degree.
 - different variables **cannot** be easily decoded separately

← Separable

Integral →



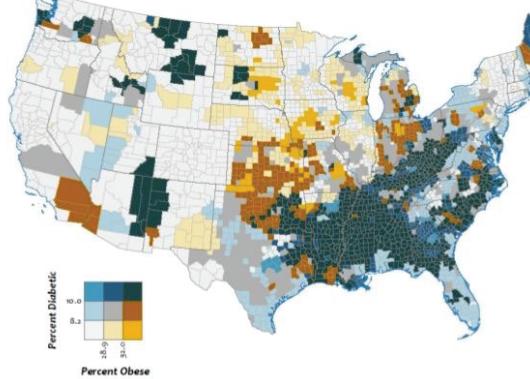
From: Ware, *Information visualization: Perception for Design*

28

Integral dimensions

A bivariate U.S. county-level map showing:
% diabetic (**saturation**) and % obese (**hue**)

Diabetes and Obesity, 2009



It is difficult to see variations in diabetes separately from obesity

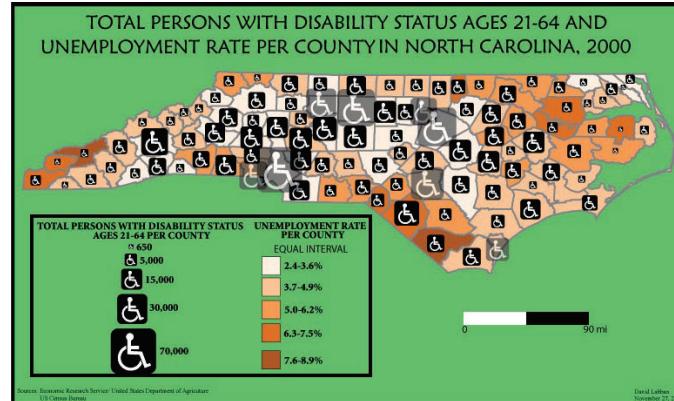
The eye is attracted to the positive correlation between these dark (blue, red) vs. light color

29

Separable dimensions

Bivariate map of N. C.: disability (**size**) and unemployment rate (**saturation**)

- These can be seen separately
- (However, TOTAL disability is confounded with population density)



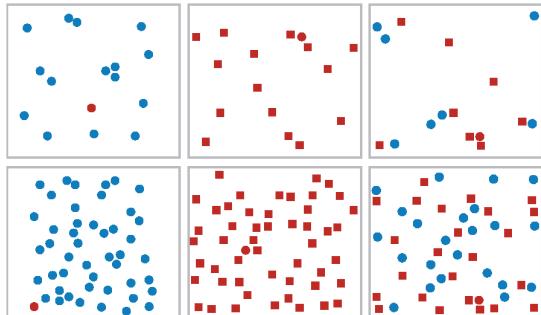
30

Anomaly detection

Find the red dot ● in each of the following displays

- This task is easiest when all the rest are blue dots ●
- Next easiest when **only shape** distinguishes the red dot ■
- Hardest when both **color and shape vary** ● ■

Sometimes called “popout” effect.
Not a good term.



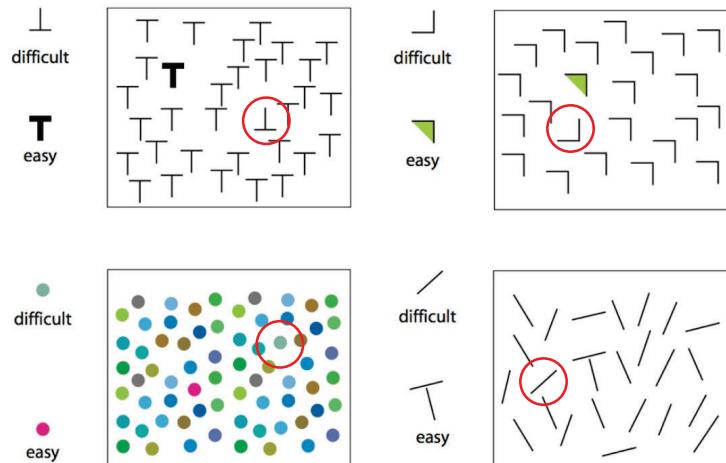
This is important in designing graphs to **highlight** some points.

31

Anomaly detection

For each display, find the anomaly shown at the left

Color and shape: What is easy or hard depends on the background



32

Encodings: Lessons

- Ordered variables
 - Prefer encodings at the top of the hierarchy (position along a scale) to those at the bottom (color saturation, curvature)
- Favor separable encodings
 - Use color and another attribute--- shape, size, orientation
 - Don't overload symbols--- probably two at most
 - Avoid mixing two aspects of color or two aspects of shape
- Small multiples
 - Reduces the need for multiple encodings within a panel
 - But, makes direct comparison more difficult
- Highlighting: to draw attention to one group, use a pre-attentive attribute

33

Encodings: Lessons

- Best to show quantitative variables with **position** or **length**
- Bar charts:
 - Best encoding via length → start at 0
 - Avoid stacked bars (not aligned), where possible
- Dot charts:
 - Best encoding via position along a scale → start at 0
- Frequency data:
 - area/color encoding to show patterns
 - sqrt or log scale often useful to show magnitude
- Color: choose sensibly ordered hues or saturation
- Arrangement
 - make comparisons easier by placing things to be compared nearby

34

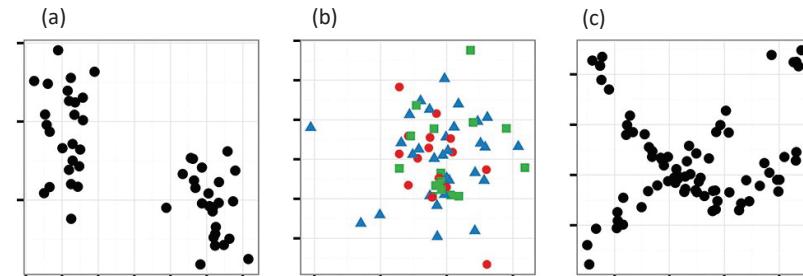
Gestalt principles

- Perception as top-down process governed by holistic principles. “Gestalt” = “form”
 - **proximity**: elements close together likely to belong to the same unit
 - **similarity**: more common visual elements increases belonging together
 - **good continuation**: elements that blend together are likely in the same unit
 - **common region**: elements in the same region likely belong together

35

Gestalt principles

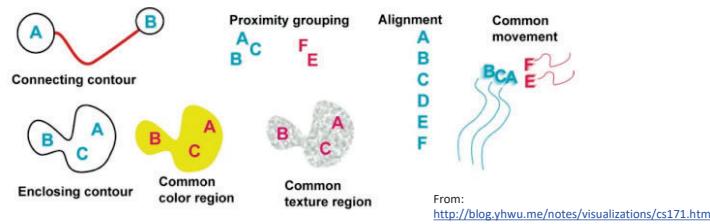
- (a) **proximity** creates impression of 2 groups
- (b) **similarity**: 3 groups via color & shape
- (c) **good continuation** gives impression of 2 groups



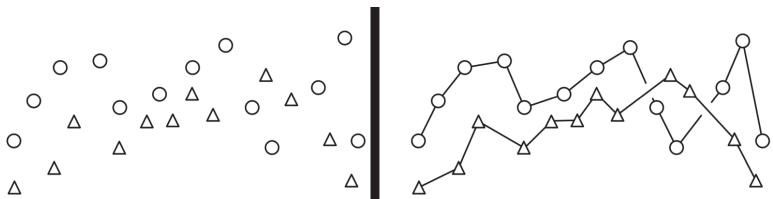
36

Gestalt principles

More gestalt ideas

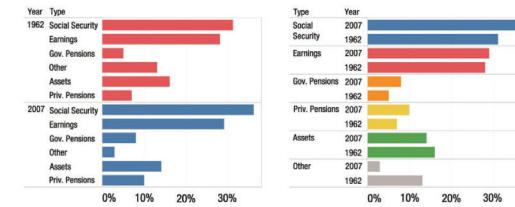


Why lines are good in time series graphs



37

Visual grouping & comparisons



combination of color & proximity grouping leads to different visual comparisons across the two bar graphs



comparisons in a word cloud are weakly controlled by color grouping; more strongly controlled with proximity grouping

From: Franconeri et al (2021) Source DOI: [10.1177/15291006211051956](https://doi.org/10.1177/15291006211051956).

38

Visual grouping & comparisons

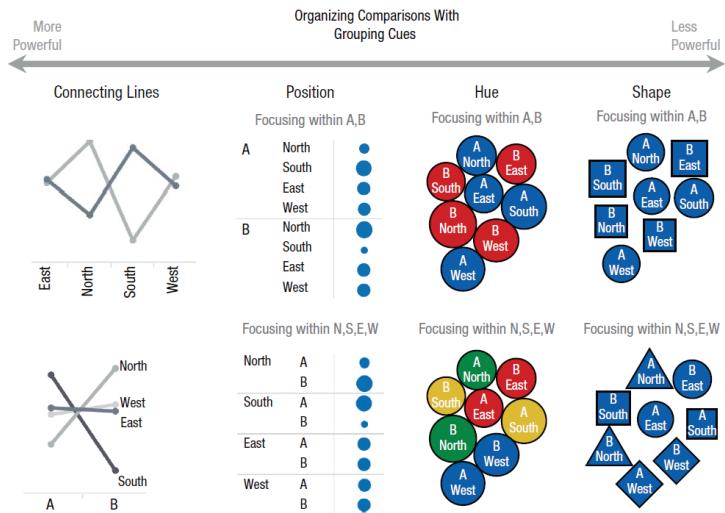
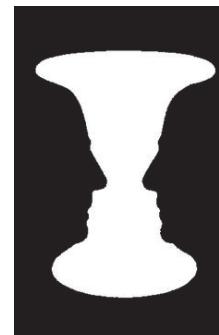


Fig. 8. Several grouping cues that can control how data values are compared. Connecting lines are particularly powerful cues, followed by proximity, color, and shape (Brooks, 2015).

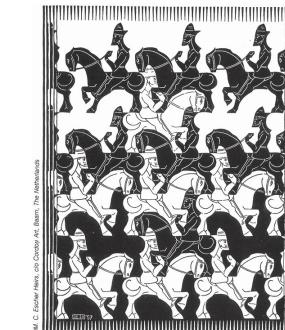
Figure - Ground

What is the figure? What is the background?

Face or vase?



Black or white soldiers?



Face or park?



These examples all use different techniques to create ambiguous figures

40

Ambiguous figures: Priming

Can you see the poodle in this scene?

What about the man?

Semantic priming: Suggestion increases likelihood of perception



41

Discussion

What perceptual features or principles are involved in your reading or understanding of these figures?



What about this?



Top-down?
Bottom-up?
Gestalt?



or this?



or this?

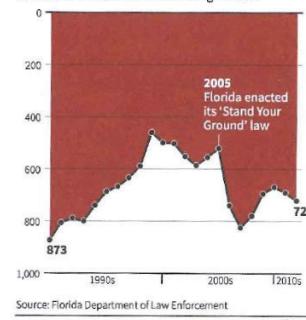
43

Figure - Ground

This graph inverts the y-axis, and shades the area above the curve

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

REUTERS

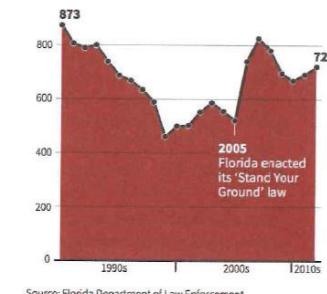
We tend to see 1999 & 2005 as high points

From: Andy Kirk, *Data Visualization: A Handbook for Data Driven Design*

A more conventional version of the same graph

Gun deaths in Florida

Number of murders committed using firearms



Source: Florida Department of Law Enforcement

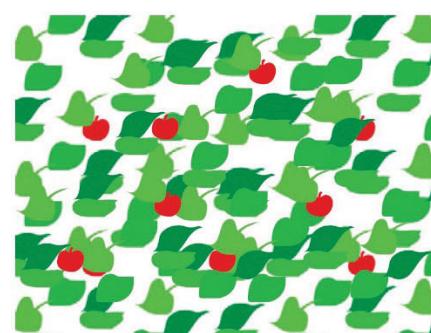
Gun deaths increased after the 'Stand your ground' law

42

Color: Functions in data graphics

Color serves to: **highlight, identify, and group** elements in a visual display

Find the cherries in this display:



Color acts as a preattentive attribute here

From: Colin Ware, *Information Visualization: Perception for Design*

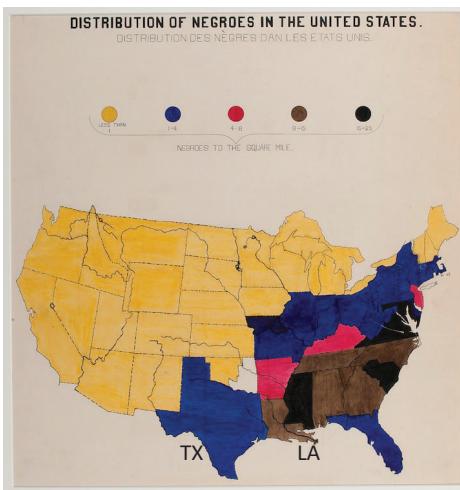
44

Nice graphic, naïve about color

W.E.B. Du Bois presented this as part of an exhibition on The American Negro at the 1900 Paris Exposition.

It is a landmark graphic, but shows no understanding of the use of color for a **quantitative** variable.

Q: Are there more Negroes per sq. mile in Texas (TX) or Louisiana (LA)?

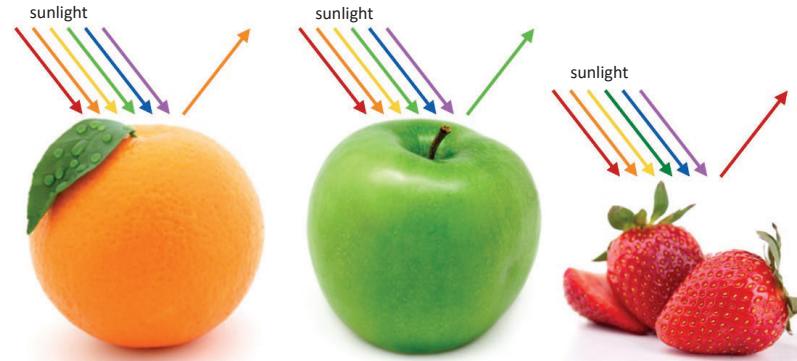


45

Object color

What makes an orange look **orange**, a green apple look **green**, or a strawberry **red**?

Objects absorb colors from the rainbow, but **reflect** their own

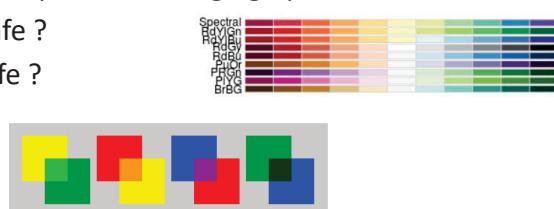
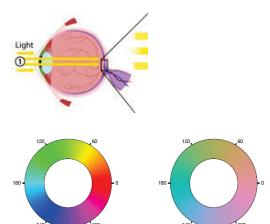


From: Miriah Meyer, lecture notes, cs6330

46

Color: Aspects in data graphics

- Perception: trichromatic theory
 - How the eye sees **color**
- Color spaces:
 - RGB (additive), CMYK (subtractive)
 - HSV, HCL: perceptually based
- Color palettes for computer graphics
 - ColorBrewer: sequential, diverging, qualitative
 - Color-blind safe ?
 - Photocopy safe ?
- Transparency

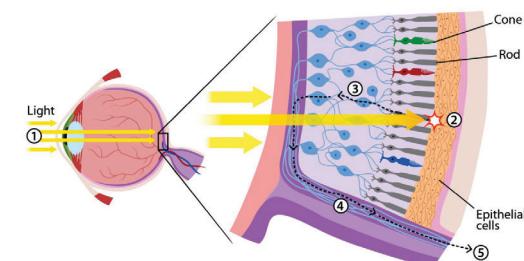
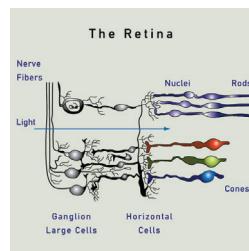


47

Perception: The human eye

- Retina:
 - rods (monochromatic),
 - cones (R, G, B)

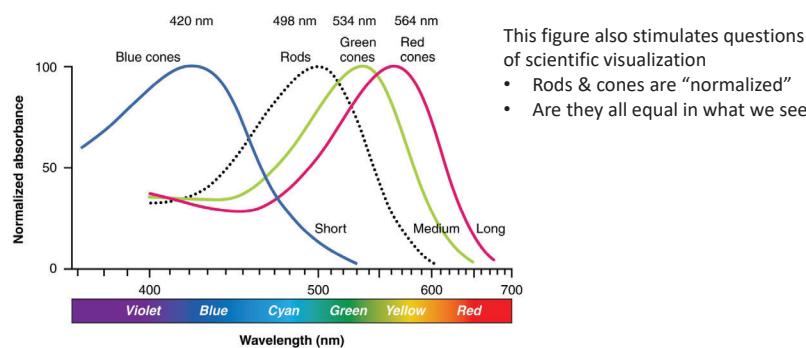
It is of interest to see the wide variety of ways this is conveyed in scientific diagrams:



48

Perception: color sensitivity

- Cells in the retina are differentially sensitive to colors of different wavelength
 - Each have a **distribution** of sensitivity for short, medium & long
 - Their **peaks** are used to name them as **Blue, Green, Red or Rods**

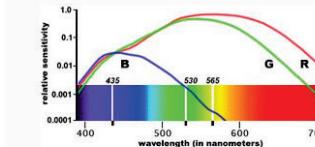


49

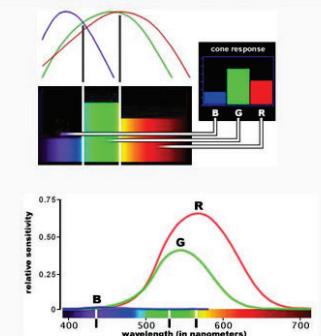
Perception: color sensitivity

This slide, from <http://slideplayer.com/slide/6329532/>, shows color sensitivity on three different scales

- Cone receptors least sensitive to (least output for) to blue



Relative sensitivity curves for the three types of cones, log vertical scale, cone spectral curves from Vos & Walraven, 1974



Relative sensitivity curves for the three types of cones, the Vos & Walraven curves on a normal vertical scale

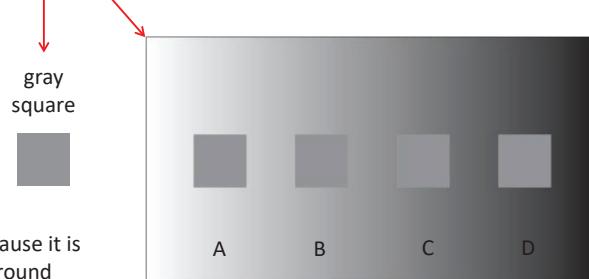
50

Perception: Contrast

Color perception, even of gray, is influenced by **contrast** against a background

Q: Which gray square at right is most similar to that at the left?

A: it is the **same** gray square against a changing background



51

Luminance contrast

Showing blue text on a black background doesn't work very well. There is insufficient luminance contrast.

Showing yellow text on a white background doesn't work very well. There is insufficient luminance contrast.

Showing blue text on a white background works better. There is sufficient luminance contrast.

Showing yellow text on a black background works better. There is sufficient luminance contrast.

TIP: For presentations, light text on a **dark background** is often preferred. I don't do this, because I'm also concerned with printing slides. (With LaTeX Beamer, it is easy to have separate setups for presentation & print)

52

Brightness illusion

Is the white at the center the **same** brightness as the white at the edges?
Or, is it **brighter**?

They are the **same**. We interpret the center as a source of light relative to the surrounding bright yellow.

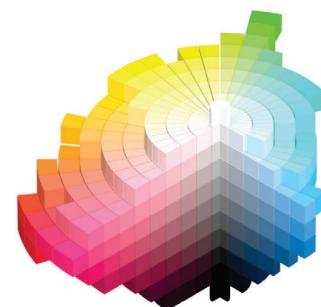
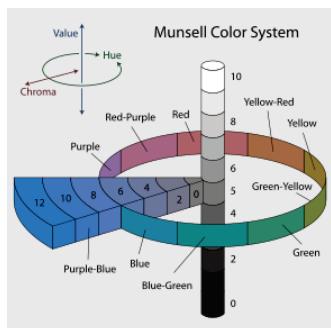


The illusion also affects pupil size! <https://www.pnas.org/content/109/6/2162>

53

Color space: Munsell colors

- Color space is 3D
 - How to specify a given color in **perceptual** terms?
 - Albert Munsell (~1930): hue, chroma, lightness (HCL)
 - These form **perceptually uniform & independent** dimensions



By SharkD - CC BY-SA 3.0,
<https://commons.wikimedia.org/w/index.php?curid=8401562> 55

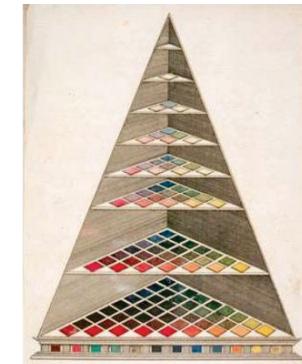
Early color theory

Tobias Mayer (1755) – color theory composed of (blue, red, yellow) as basic colors



Introduces the idea of color "primaries"

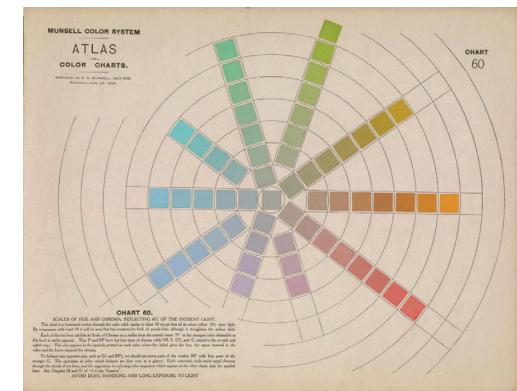
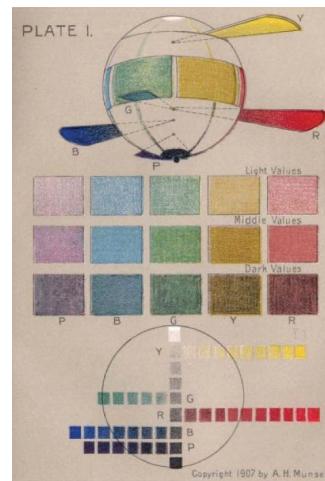
Johannes Lambert (1772) – A color pyramid, composed of 7 layers



Introduces the idea of color saturation

54

Munsell's color scheme was highly influential in Psychology research
Nearly every lab investigating color used standard sets of Munsell color chips



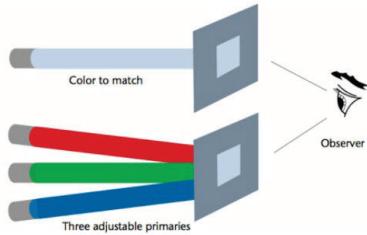
56

CIE color space

- How do we know about the **perceptual** properties of colors, taking spectral sensitivity into account?

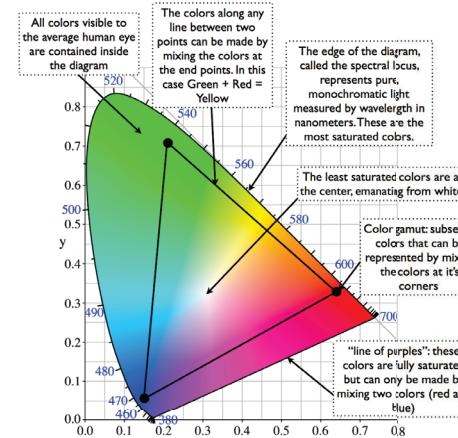
Experiments used a **color-matching** task:

- Adjust the intensity of pure R, G, B lights to match a given color
- This defines a new color theory connecting **physical** properties and human **perception** (spectral sensitivity).
- The CIE (International Commission on Illumination) becomes the standard to calibrate color in scientific instruments and human experiments



57

CIE color space



Anatomy of a CIE Chromaticity Diagram

The International Commission on Illumination (CIE) in 1931 defines a color space of (x,y,z) coordinates based on color-matching experiments combining R, G, B light sources in additive mixtures, and a "standard colorimetric observer"

This defines a new color theory connecting **physical** properties and human **perception** (spectral sensitivity).

There are eventually a variety of CIE color spaces (CIELab, CIEuv, ...) and lots of formulas for converting among them.

58

Color space: RGB & CMYK

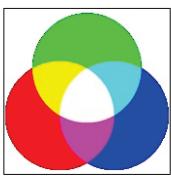
Enter technology: how to produce color?

- RGB:
 - Combine **light**: **R + G + B** = white
 - Used in computer monitors, TV, film
- CMYK:
 - Combine **ink**: **Cyan + Magenta + Yellow** = Black
 - Used in color laser printers, the print industry



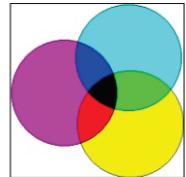
Caution: R, SAS, SPSS use RGB by default

Additive colour system



Mixture of primary light colours
White

Subtractive Colour System



Mixture of primary pigment colours
Black

TIP: for publishing, you may need to prepare or convert graphics from RGB to CMYK.

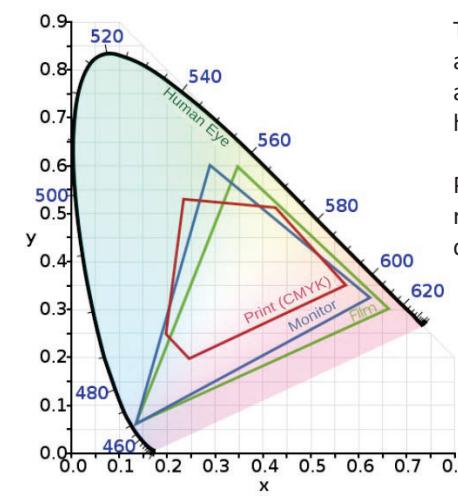
Some software offer useful tools for this:

- Adobe Acrobat Pro
- ImageMagik

Color space: RGB & CMYK

The standard gamut of colors available for different **display media** are a restricted subset of what the human eye can see.

Print (CMYK) is most restricted, and requires a more careful choice of color in graphics



Always check the display on different media.

59

60

Software: Color specification

- Color is often hard to use effectively in software, because the ways to specify it are so varied:
 - Color names: "black", "red", "green3", "skyblue", "cyan"
 - RGB: black=(0,0,0); green3=(0, 205, 0), cyan=(0, 255, 255)
 - Hex: black="#000000"; cyan="#00FFFF"

| | | |
|-------------------|---------|-------------|
| 18 beige | #F5F5DC | 245 245 220 |
| 19 bisque | #FFB4C4 | 255 228 196 |
| 20 bisque1 | #FFB4C4 | 255 228 196 |
| 21 bisque2 | #EED5B7 | 238 213 183 |
| 22 bisque3 | #CD979E | 205 182 159 |
| 23 bisque4 | #B87D6B | 139 125 107 |
| 24 black | #000000 | 0 0 0 |
| 25 blanchedalmond | #FFEBCD | 255 235 205 |
| 26 blue | #0000FF | 0 0 255 |
| 27 blue1 | #0000FF | 0 0 255 |
| 28 blue2 | #0000EE | 0 0 238 |
| 29 blue3 | #0000CD | 0 0 205 |
| 30 blue4 | #00008B | 0 0 139 |
| 31 blueviolet | #8A2BE2 | 138 42 224 |
| 32 brown | #8B52AA | 165 42 42 |
| 68 cyan | #00FFFF | 0 255 255 |
| 69 cyan1 | #00FFFF | 0 255 255 |
| 70 cyan2 | #00EEFF | 0 238 255 |
| 71 cyan3 | #00CCFF | 0 205 255 |
| 72 cyan4 | #00BFFF | 0 139 255 |
| 73 darkblue | #00008B | 0 0 139 |
| 74 darkcyan | #00BFFF | 0 139 139 |
| 75 darkgoldenrod | #8B8A0B | 194 134 11 |
| 76 darkgoldenrod1 | #FFB90F | 255 105 15 |
| 77 darkgoldenrod2 | #E8A0D0 | 238 173 14 |
| 78 darkgoldenrod3 | #CD958C | 205 149 12 |
| 79 darkgoldenrod4 | #8B650B | 139 101 9 |
| 80 darkgray | #A9A9A9 | 169 169 169 |
| 81 darkgreen | #006400 | 0 100 0 |
| 82 darkgrey | #A9A9A9 | 169 169 169 |

See: <http://research.stowers.org/mcm/efg/R/Color/Chart/> for R color charts

61

Every time you are forced to say "#008B8B" or "cyan4" a puppy dies somewhere

-- MF, 2018

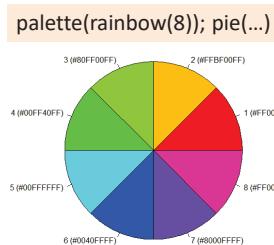
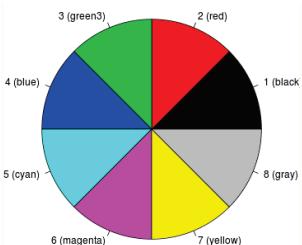


62

Software: palettes

- R (and other software) provide palettes of colors used for **defaults** in graphs
 - Not all are nice— depends on your purpose
 - But, there are lot of choices
 - You can change them **once** for all graphs in a session or paper

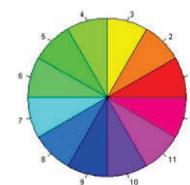
```
> (pal <- palette())
[1] "black" "red"  "green3" "blue"  "cyan"  "magenta" "yellow" "gray"
> pie(rep(1, length(pal)), labels = sprintf("%d (%s)", seq_along(pal), pal), col = pal)
```



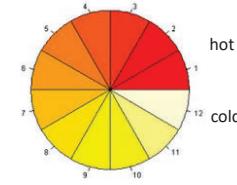
63

R: basic palettes

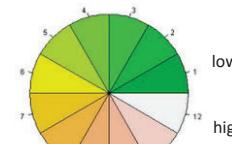
n < 12
pie(rep(1, n), col=rainbow(n))



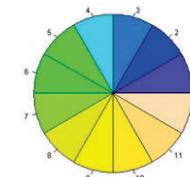
pie(rep(1, n), col=heat.colors(n))



pie(rep(1, n), col=terrain.colors(n))



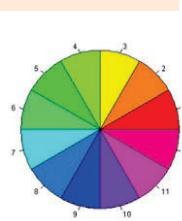
pie(rep(1, n), col=topo.colors(n))



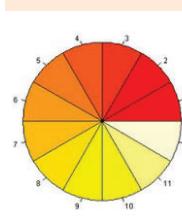
64

R: basic palettes

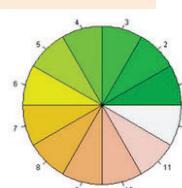
rainbow



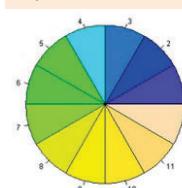
heat



terrain



topo



Discussion Q:

- Which of these are better for **quantitative** variables?
- Which for **categorical**?

These are shown for **area fill**. How effective would they be for:

- **point** colors
- **line** colors

E.g., yellow is bright as an area, but nearly invisible as points (●) or lines (→) or **text on a white background**

65

palettes: ColorBrewer

ColorBrewer, by Cynthia Brewer provides an interactive application for choosing color palettes, <http://colorbrewer2.org>

This is one example of a **multi-hue** scheme for a **quantitative, sequential** variable, shown from low to high with 4 color classes

The screenshot shows the ColorBrewer 2.0 interface. A red arrow points to the 'variable type' dropdown set to 'sequential'. Another red arrow points to the 'choose different versions of the scheme' section, which displays four variations of the BuGn scheme. A third red arrow points to the 'export color specs to HEX, RGB, CMYK' button. On the right, a map of the United States is colored using the 4-class BuGn scheme, transitioning from light blue to dark green across the country's terrain.

This example: <http://colorbrewer2.org/#type=sequential&scheme=BuGn&n=4>

66

palettes: ColorBrewer

Diverging schemes are designed to show a quantitative variable, where we want to see what is low vs. what is high, leaving the middle of less visual impact – difference from average, residuals, ...

there are different schemes within this rubric

The screenshot shows the ColorBrewer 2.0 interface. A red arrow points to the 'choose # classes' dropdown set to 5. Another red arrow points to the 'various schemes' section, which displays five variations of the RdBu scheme. A third red arrow points to the 'see other context' section, which shows a map of the United States colored using the 5-class RdBu scheme, with red in the West and blue in the East. A fourth red arrow points to the 'add transparency' section.

This example: <http://colorbrewer2.org/#type=diverging&scheme=RdBu&n=5>

67

palettes: ColorBrewer

Qualitative schemes are designed to show a **categorical** variable, where we want to see differences among **unordered** categories

The screenshot shows the ColorBrewer 2.0 interface. A red arrow points to the 'choose # classes' dropdown set to 5. Another red arrow points to the 'various schemes' section, which displays five variations of the Accent scheme. A third red arrow points to the 'see other context' section, which shows a map of the United States colored using the 5-class Accent scheme, with various colors (green, orange, blue, etc.) distributed across the country. A fourth red arrow points to the 'add transparency' section.

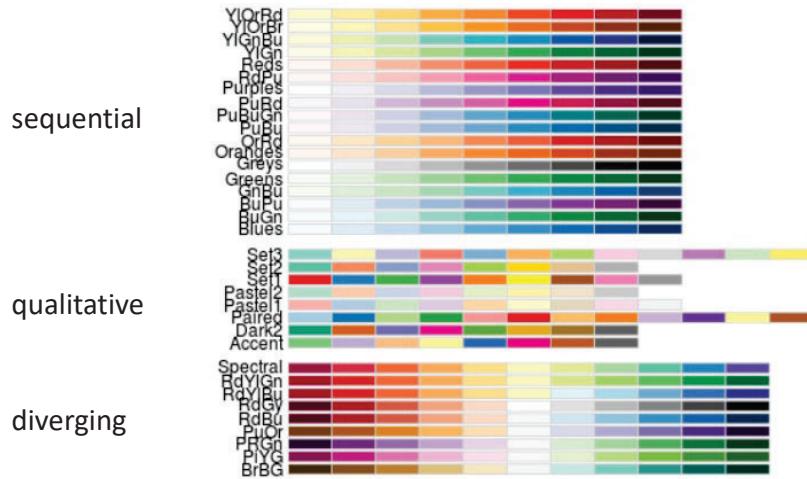
These are all available in the RColorBrewer package

This example: <http://colorbrewer2.org/#type=qualitative&scheme=Accent&n=5>

68

palettes: RColorBrewer

RColorBrewer::display.brewer.all()



R: choose_palette()

The colorspace package in R has an interactive palette widget.

It also provides functions for many kinds of color manipulations.

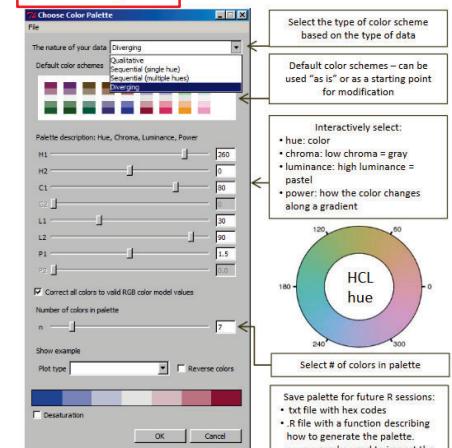
The R Color cheatsheet, by Malcolm Fraser is a goto source for all aspects of color in R:

<https://www.nceas.ucsb.edu/~frazier/RSpatialGuides/colorPaletteCheatsheet.pdf>

R color cheatsheet

Overview of colorspace palette selector

library("colorspace")
pal <- choose_palette()



70

Viridis palettes

Designed by Stéfan van der Walt and Nathaniel Smith for Python; ported to R in the [viridis](#) package.

Goals:

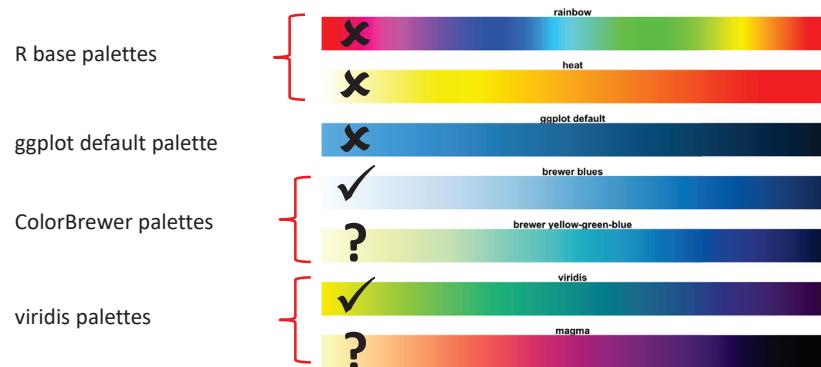
- **Colorful**, spanning as wide a palette as possible so as to make differences easy to see
- **Perceptually uniform**: values close to each other have similar-appearing colors and values far away from each other have more different-appearing colors
- **Robust to colorblindness**: these properties hold true for people with common forms of colorblindness, as well as in grey scale printing
- **Pretty**: much nicer as a defaults in software

These assertions are largely **untested**. Perhaps a good research topic!

71

Comparing palettes

For a **quantitative** variable and a **continuous** color scale, there are many choices. How well do they work?

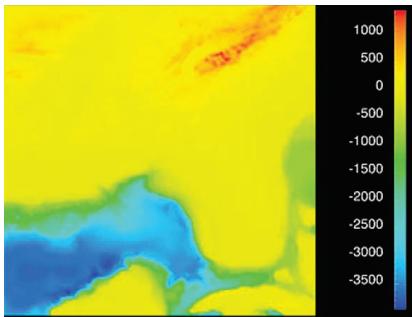


This is a bit tricky: ideally, we want a **wide range** of color

72

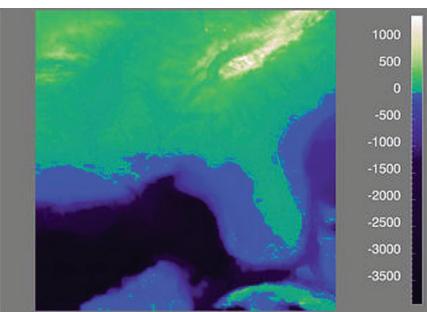
Comparing palettes

What is shown in this map?



The rainbow color scale obscures the main features

Now we can see it—elevation in the Florida coast: above or below 0



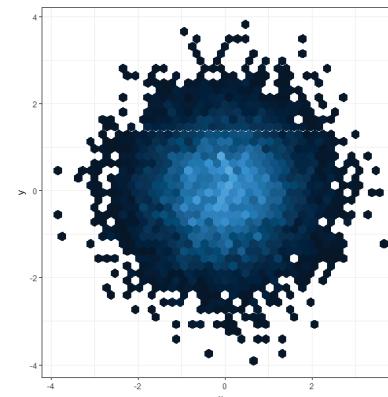
This color scheme was designed to reveal the essential topography of the map & to have perceptually equal elevation steps

From: <http://www.research.ibm.com/people/l/lloyd/color/color.HTM>

73

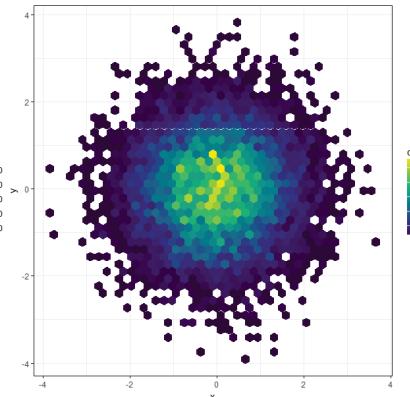
Comparing palettes

ggplot default palette



```
df <- data.frame(x = rnorm(10000), y = rnorm(10000))
g <- ggplot(df, aes(x = x, y = y)) +
  geom_hex(bins=40) + coord_fixed() + theme_bw()
g
```

viridis default palette



```
library(viridis)
g + scale_fill_viridis()
```

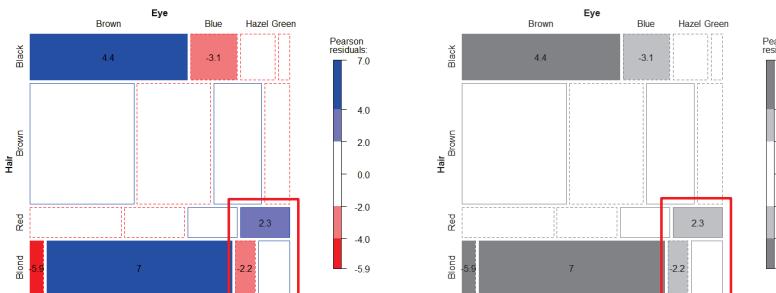
74

Color → B/W ?

Graphics designed in color often have to consider what happens when graphs are reproduced in B/W: grayscale

- This is particularly hard for a **diverging** color scale
- My original design for mosaic plots used solid vs. dashed lines to distinguish + vs. -

`mosaic(haireye, labeling=labeling_residuals, gp=shading_Friendly)`



75

Color → B/W ?

The design of this graphic table was crafted to preserve readability if printed in B/W.
NB: text for numbers changes from black to white depending on background color.

Figure 9: Section 37 benefits by type (1998–2015)

| | 1998–2002 | 2003–2005 | 2006–2009 | 2010–2013 | 2014–2016 | Scale |
|--------------------------------|-----------|-----------|-----------|-----------|-----------|---------|
| Roads, streetscapes | 30 | 35 | 54 | 83 | 15 | 0 – 10 |
| Culture, community, recreation | 26 | 50 | 59 | 47 | 16 | 11 – 20 |
| Parks | 27 | 41 | 41 | 52 | 20 | 21 – 30 |
| Affordable housing | 17 | 26 | 38 | 56 | 11 | 31 – 40 |
| Public art | 26 | 25 | 41 | 32 | 4 | 41 – 50 |
| Heritage | 16 | 13 | 26 | 18 | 3 | 51 – 60 |
| Transit | 11 | 7 | 10 | 20 | 3 | 61 – 70 |
| Libraries | 6 | 2 | 5 | 11 | 1 | 71 – 80 |
| Other | 3 | 6 | 7 | 8 | 3 | 81 – 90 |

Figure 9: Section 37 benefits by type (1998–2015)

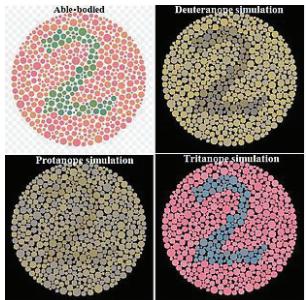
| | 1998–2002 | 2003–2005 | 2006–2009 | 2010–2013 | 2014–2016 | Scale |
|--------------------------------|-----------|-----------|-----------|-----------|-----------|---------|
| Roads, streetscapes | 30 | 35 | 54 | 83 | 15 | 0 – 10 |
| Culture, community, recreation | 26 | 50 | 59 | 47 | 16 | 11 – 20 |
| Parks | 27 | 41 | 41 | 52 | 20 | 21 – 30 |
| Affordable housing | 17 | 26 | 38 | 56 | 11 | 31 – 40 |
| Public art | 26 | 25 | 41 | 32 | 4 | 41 – 50 |
| Heritage | 16 | 13 | 26 | 18 | 3 | 51 – 60 |
| Transit | 11 | 7 | 10 | 20 | 3 | 61 – 70 |
| Libraries | 6 | 2 | 5 | 11 | 1 | 71 – 80 |
| Other | 3 | 6 | 7 | 8 | 3 | 81 – 90 |

Background shading works equally well in color or B/W
A+ for visual design!

Source: Friendly, A. R. (2017). *Land Value Capture and Social Benefits: Toronto and São Paulo Compared*. IMF Papers on Municipal Finance and Governance, No 33, University of Toronto, <https://munkschool.utoronto.ca/imf/>

76

Colorblindness



Most common forms are genetic, and involve a deficiency in one of the cone type sensitivities

- Protanopia (red deficient: L cone absent)
- Deutanopia (green deficient: M cone absent)
- Tritanopia (blue deficient: S cone absent)

Some form of red-green insensitivity is most common

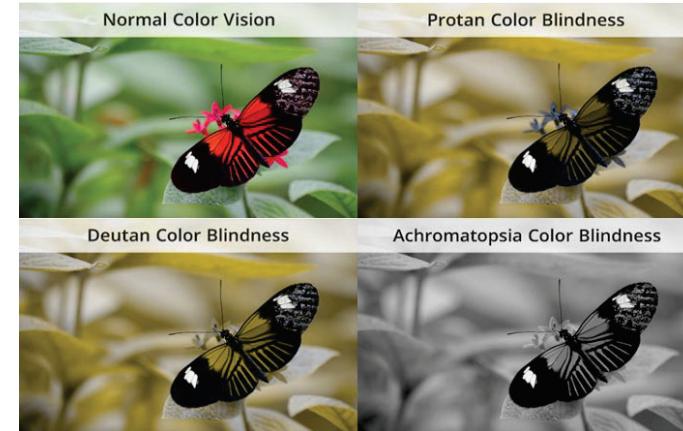
- about 6-8% of population
- more common in males

TIP: Avoid color scales with main variation between red & green

77

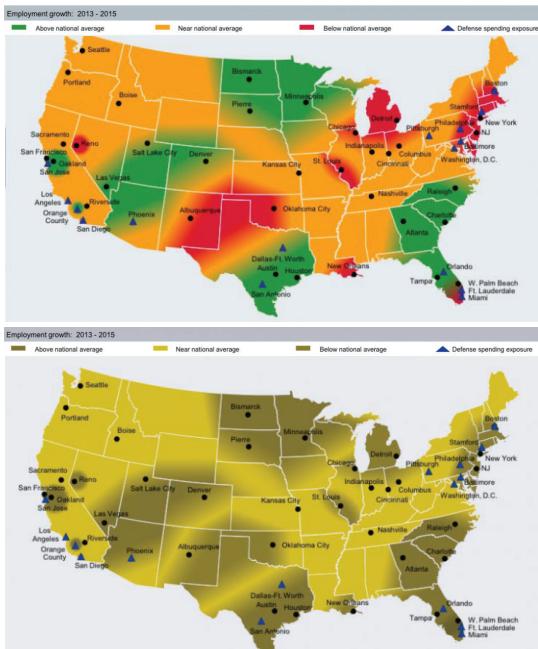
Colorblindness

What an image looks like with various forms of color deficiency



Red-green colorblindness: $-R$ (protan) $\approx -G$ (deutan)

78



Goal: Show employment growth, 2013–2015

Original design, using

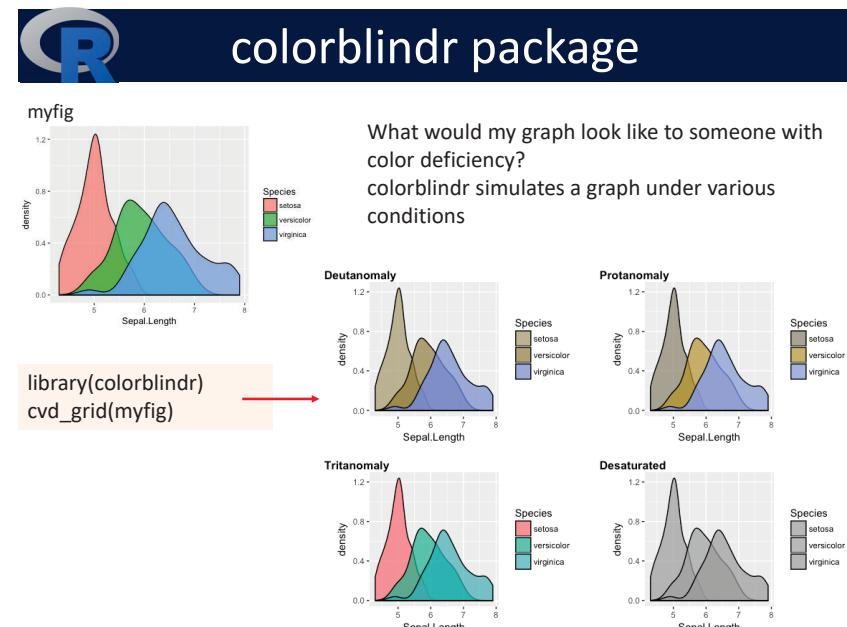
- green: above average
- red: below average

How this looks to someone with red-green colorblindness

red & green become indistinguishable

From: <http://www.mena-forum.com/category/u-s-a/>

79



colorblindr package

What would my graph look like to someone with color deficiency?
colorblindr simulates a graph under various conditions

library(colorblindr)
cvd_grid(myfig)

80



colorspace

2.0-0

A Toolbox for Manipulating and Assessing Colors and Palettes

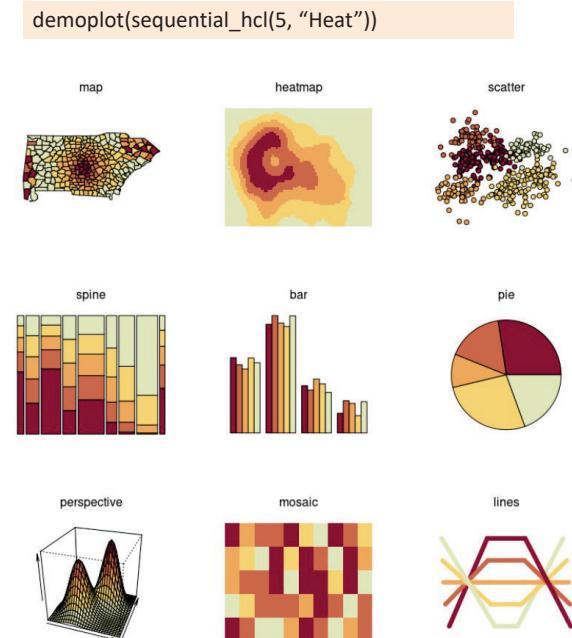


See: <http://colorspace.r-forge.r-project.org/>

81



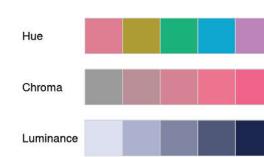
See how color palettes work in different kinds of statistical displays



83

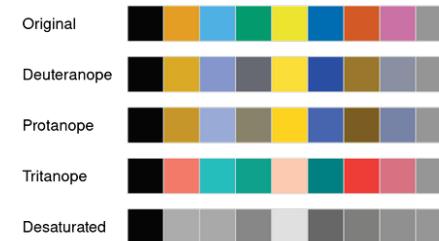
colorspace: palette visualization

swatchplot(): display collections of palettes in flexible ways



Effect of varying hue, chroma and luminance individually

```
swatchplot(
  "Hue"    = sequential_hcl(5, h = c(0, 300), c = c(60, 60), l = 65),
  "Chroma" = sequential_hcl(5, h = 0, c = c(100, 0), l = 65, rev = TRUE, power = 1),
  "Luminance" = sequential_hcl(5, h = 260, c = c(25, 25), l = c(25, 90), rev = TRUE, power = 1),
  off = 0
)
```



Emulate different types of color vision deficiency for one or more palettes

```
swatchplot(palette.colors(), cvd = TRUE)
```

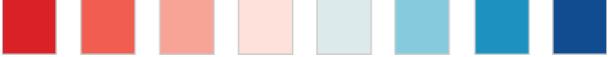
82

Color: Lessons

- Use colors to represent differences in meaning
 - Avoid gratuitous use of multiple colors
 - Use consistent color scheme across multiple graphs of the same data
- Consider presentation goal:
 - Highlight one subset against the rest
 - Group a categorical variable
 - Encode a quantitative variable
- Consider differences in color perception, B/W printing

84

Color: Lessons

- Consider encoding scheme:
 - Categorical: Use a wide range of hues, of ~ same saturation
 - Sequential: use a small range of hues of varying intensity
 - Diverging: Use two sequential schemes, decreasing toward the middle

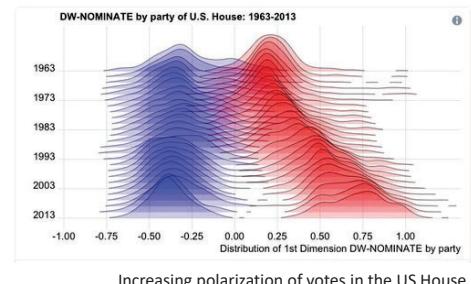
Images from: Stephen Few,
http://www.perceptualedge.com/articles/visual_business_intelligence/rules_for_using_color.pdf

85

Transparency

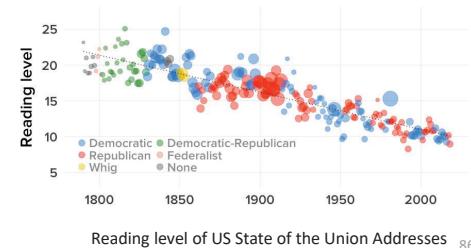
Colors can be made partially **transparent**, by adding an “alpha” channel,
 $0 \leq \alpha \leq 1$ (opaque)

Filled areas combine to look more saturated
What do you see here?



This also works well with filled point symbols, which would otherwise be obscured when they overlap

Different colors “blend”
What do you see here?

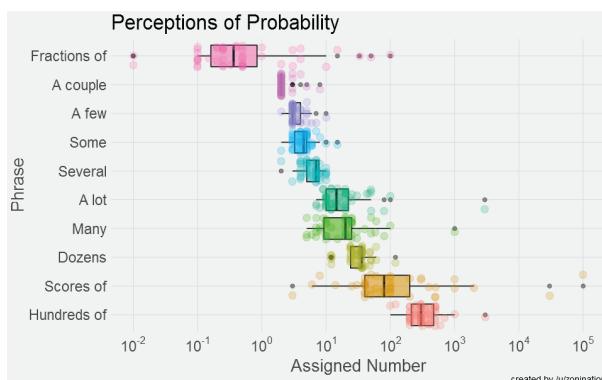


86

Transparency: Adding another layer

Transparency also works well to combine different graphical features in a plot
Here, a filled boxplot and dots representing individual observations

What number would you assign to the following phrases?



From: <https://github.com/zonination/perceptions>

87

Summary

- In designing data graphics, consider the viewer
 - Info → encoding → image → decoding → understanding
- Perception: much is known, with ~ links to graphics
 - Bottom up: perceptual features, what grabs attention
 - Top down: expectations provide a context
 - Encoding attributes must consider what is to be seen
- Color: What is the presentation goal?
 - Color palettes for different purposes
 - Transparency increases the effective use of color

88