

CSSS 569 · Visualizing Data

COGNITIVE ISSUES IN DATA VISUALIZATION

Christopher Adolph

Department of Political Science
and
Center for Statistics and the Social Sciences
University of Washington, Seattle



The Cognitive Science of Visual Displays of Information

Suppose we design the most beautiful,
data rich display we can

But we use elements that humans can't
perceive: Ultraviolet light.

The limits of human vision render our
display useless

The Cognitive Science of Visual Displays of Information

Suppose we design the most beautiful,
data rich display we can

But we use elements that humans can't
perceive: Ultraviolet light.

The limits of human vision render our
display useless

Now suppose we design the most
beautiful, data rich display we can

We use elements humans *can* perceive,
but get systematically wrong

No better? Even worse?

The Cognitive Science of Visual Displays of Information

Suppose we design the most beautiful,
data rich display we can

But we use elements that humans can't
perceive: Ultraviolet light.

The limits of human vision render our
display useless

Now suppose we design the most
beautiful, data rich display we can

We use elements humans can perceive,
but get systematically wrong

No better? Even worse?

Unfortunately, cognitive errors are everywhere

But few designers of scientific visuals pay close attention to them

The Cognitive Science of Visual Displays of Information

Unfortunately, cognitive errors are everywhere

Ideally, we would have an algorithm to accomplish the following:

```
cognitivelyAdjustedGraphic <- correctForErrors(InitialGraphic)
```

The Cognitive Science of Visual Displays of Information

Unfortunately, cognitive errors are everywhere

Ideally, we would have an algorithm to accomplish the following:

```
cognitivelyAdjustedGraphic <- correctForErrors(InitialGraphic)
```

Alas, this does not exist

The cognitive study of graphics is difficult

Hard to systematically understanding how graphical elements combine & interact

Instead, many specific errors known from experiments

These experiments provide warnings about dangerous techniques

To minimize error, we can try to use more reliable graphical elements

The Cognitive Hierarchy of Graphical Elements

Graphical elements used to encode data:

More accurate Position on a plane

Line length

Angle & slope



Area

Volume

Less accurate

Color

Graphical elements are *not* all equal in clarity

People are much better at judging line length than angle or grayscale

Source:

Cleveland & McGill, JRSS, 1987

The Cognitive Hierarchy of Graphical Elements

Graphical elements used to encode data:

More accurate Position on a plane

Line length

Angle & slope



Area

Volume

Less accurate

Color

To show and correlate
multivariate data, we'd like to
use multiple or multifunctional
elements

Color and size and shape, for
example

Source:

Cleveland & McGill, JRSS, 1987

The Cognitive Hierarchy of Graphical Elements

Graphical elements used to encode data:

More accurate Position on a plane

Line length

Angle & slope



Area

Volume

Less accurate

Color

Will they all be processed
equally...

accurately?
quickly?
with similar intensity?
separately?
at highest available level of
measurement?

Unfortunately, NO.

Source:

Cleveland & McGill, JRSS, 1987

The Cognitive Hierarchy of Graphical Elements

Graphical elements used to encode data:

More accurate Position on a plane

Line length

Angle & slope



Area

Volume

Less accurate

Color

Simple advice:

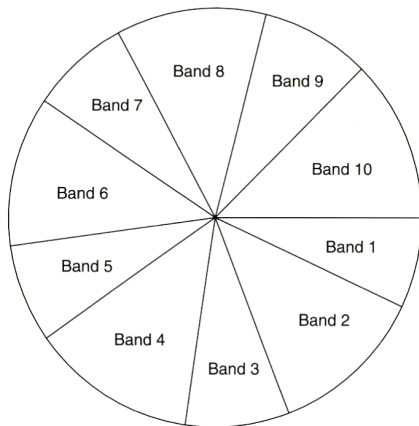
Reserve elements at the top of the list for important variables

Try to avoid using the elements at the bottom of the list to encode *quantitative* data (but redundant usage is fine)

Exception: color can effectively encode qualitative data

Source:

Cleveland & McGill, JRSS, 1987

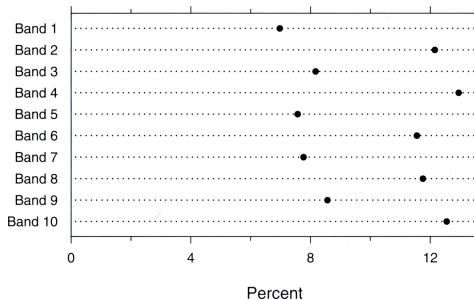


Cognitive failures:
Angular data encoding

Can you describe these data?

The exact sizes of the pies?

Source:
Cleveland, *The Elements of Graphing Data*



Cognitive solution:
Location data encoding

Did you notice that half of the slices
are exactly 50% larger than the
others?

Did you guess the exact sizes
correctly?

Source:

Cleveland, *The Elements of Graphing Data*

Major Categories of Federal Income and Outlays for Fiscal Year 2001

Income and Outlays. These pie charts show the relative sizes of the major categories of Federal income and outlays for fiscal year 2001.

Income

Social security, Medicare, and unemployment and other retirement taxes

35%

Personal income taxes

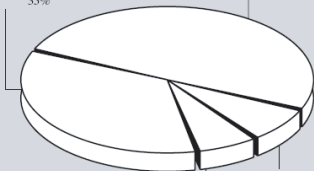
50%

Corporate income taxes

7%

Excise, customs, estate, gift, and miscellaneous taxes

8%



Outlays

Law enforcement and general government

2%

Surplus to pay down the debt

6%

Social security, Medicare, and other retirement ¹

36%

Social programs ⁴

18%

Physical, human, and community development ³

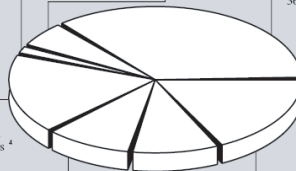
10%

Net interest on the debt

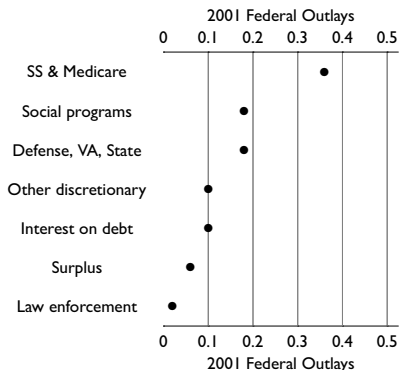
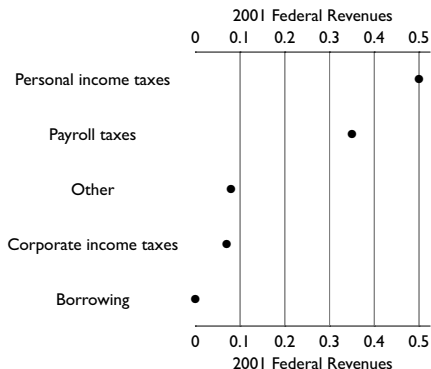
10%

National defense, veterans, and foreign affairs ²

18%



My favorite pie chart. Budget data printed on the back of the US tax forms.
Would a dot chart be as good? Better?



I think this is at least as good as the pie.

Unlike the pie, could be expanded to, say, 10 categories without much fuss.

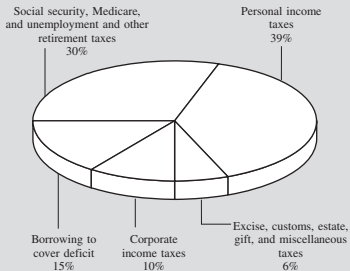
Note that I have **diagonalized** the dot plot by sorting.

This is always helpful for reading and comparing data correctly.

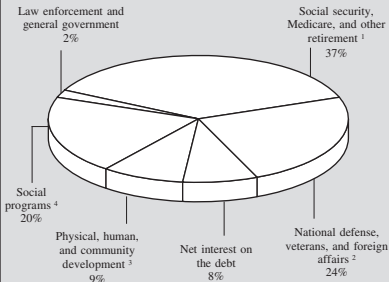
Major Categories of Federal Income and Outlays for Fiscal Year 2008

Income and Outlays. These pie charts show the relative sizes of the major categories of federal income and outlays for fiscal year 2008.

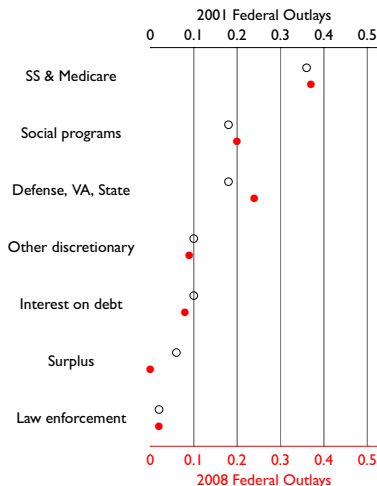
Income



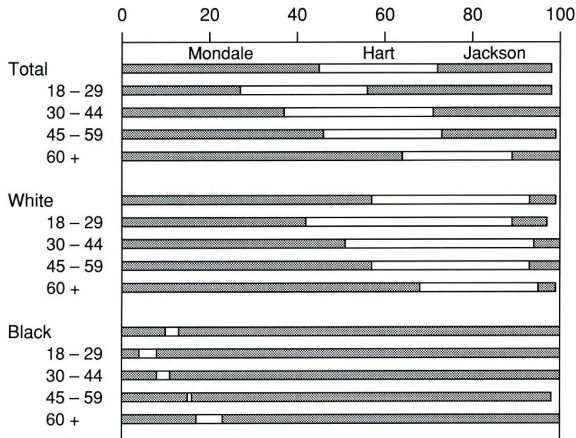
Outlays



Budget pies went missing for most of the decade
They returned in 2009 – right after a presidential election



Another advantage of dotplots: easy comparison across plots through integration
 This is also why dotplots are more useful than barplots



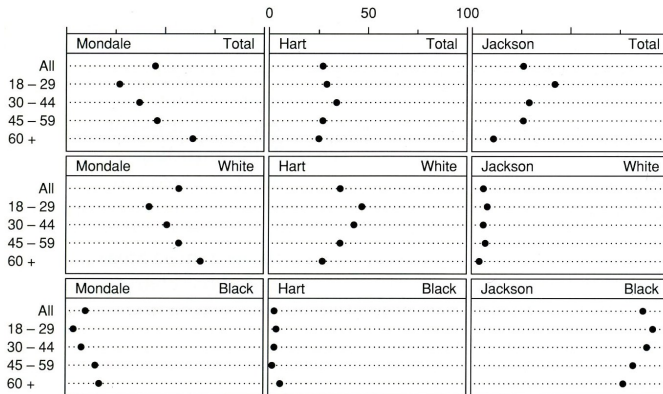
Another reason to prefer dotplots to barplots

Consider this stacked bar plot

How do Hart and Mondale's support depend on age of voter?

How do these patterns vary by race of voter?

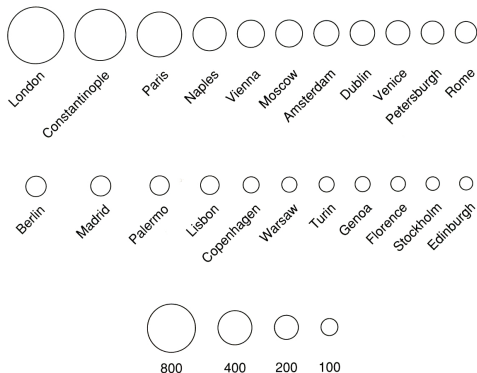
Source:
Cleveland, *The Elements of Graphing Data*



Now consider an array of dotplots of the same data

How do Hart and Mondale's support depend on age of voter?

How do these patterns vary by race of voter? By race \times age?



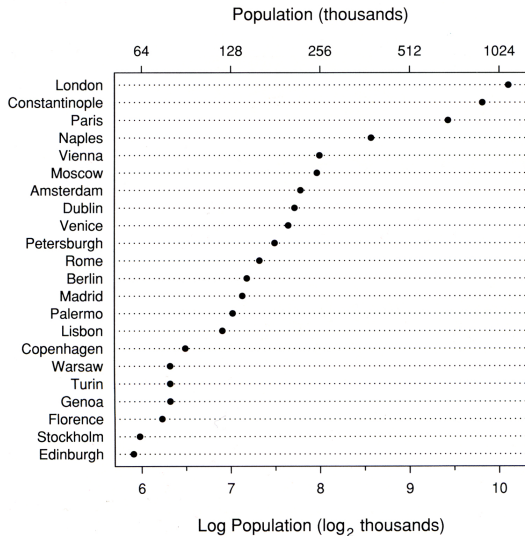
Cognitive failures:
Area data encoding

Is there a smooth increase in city size across these data?

How much bigger was Lisbon than Copenhagen?

Can we even be sure the areas represent population? What if the diameter is what matters?

Source:
Cleveland, *The Elements of Graphing Data*



Cognitive solution:
Location data encoding

Lisbon was about fifty percent bigger than Copenhagen

The progression is far from smooth, even on a log scale

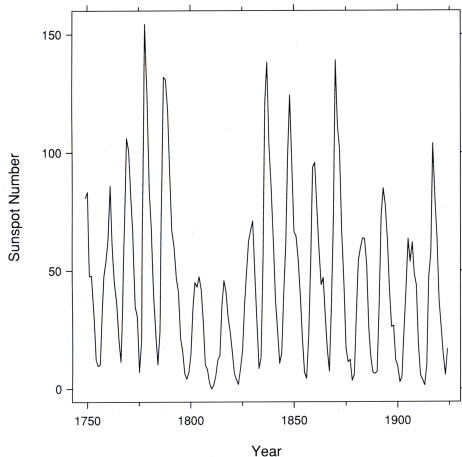


Source: Ware, *Information Visualization*

Perception of volume is even worse than area!

Many viewers read the volumes as if they were areas:
never use volume to represent quantities

Use measures based on length or location, rather than area or volume

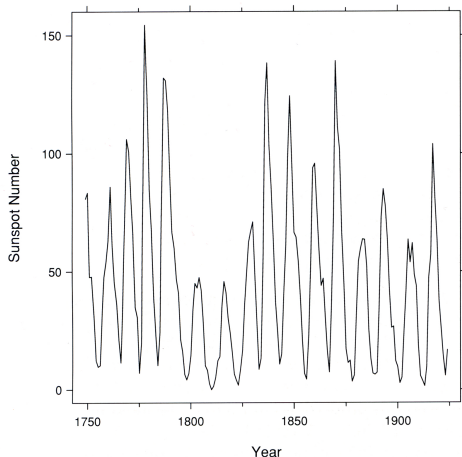


*Cognitive failures:
Hard-to-read lineplots*

Poor angular perception can make
some lineplots hard to read

Are these spikes symmetrical?

Source:
Cleveland, *The Elements of Graphing
Data*

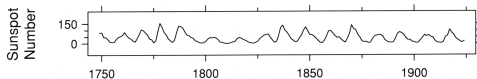


*Cognitive solution:
better aspect ratio*

Choosing a better aspect ratio can
make lineplots easier to read

The goal is to make slopes close to
1.0 or -1.0

Cleveland calls this “banking to 45
degrees,” because angles near 45
are easier for humans to distinguish



Ah-ha! vs Hmm...

Sometimes differences take conscious effort to distinguish. Find the 3's:

85689726984689762689764358922659865986554897689269898
02462996874026557627986789045679232769285460986772098
90834579802790759047098279085790847729087590827908754
98709856749068975786259845690243790472190790709811450
85689726984689762689764458922659865986554897689269898

Ah-ha! vs Hmm...

Sometimes differences take conscious effort to distinguish. Find the 3's:

85689726984689762689764358922659865986554897689269898
02462996874026557627986789045679232769285460986772098
90834579802790759047098279085790847729087590827908754
98709856749068975786259845690243790472190790709811450
85689726984689762689764458922659865986554897689269898

Sometimes encoded data pop right out. Find the 3's:

85689726984689762689764**3**58922659865986554897689269898
024629968740265576279867890456792**3**2769285460986772098
908**3**4579802790759047098279085790847729087590827908754
9870985674906897578625984569024**3**790472190790709811450
85689726984689762689764458922659865986554897689269898

Same information in both examples.

But our brains process color differences “pre-attentively” – fast & effortlessly

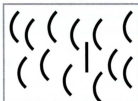
Source: Ware, *Information Visualization*

Ah-ha! vs Hmm...

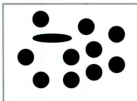
Orientation



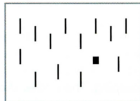
Curved straight



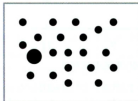
Shape



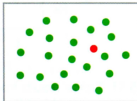
Shape



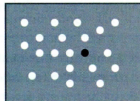
Size



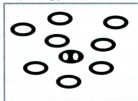
Color



Light/dark



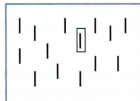
Topology (or count)



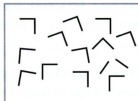
Convex/concave



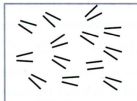
Addition



Juncture (not pre-att)



Parallelism (not pre-att)



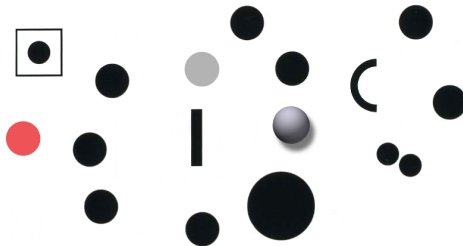
Where possible, **pre-attentive differences** should be exploited

The essence of graphics that “hit you between the eyes”

Tables of numbers seldom if ever achieve this

Source:
Ware, *Information Visualization*

Ah-ha! vs Hmm...



But there's only so much pre-attention to go around

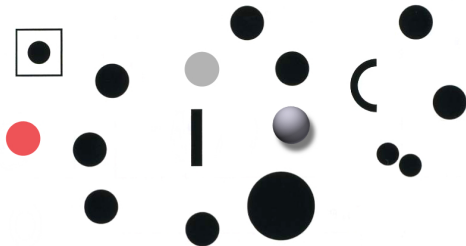
As you add pre-attentive differences, the effect of each diminishes

(though not necessarily equally: some are stronger than others)

Source:

Ware, Information Visualization

Glyphs and Data Display



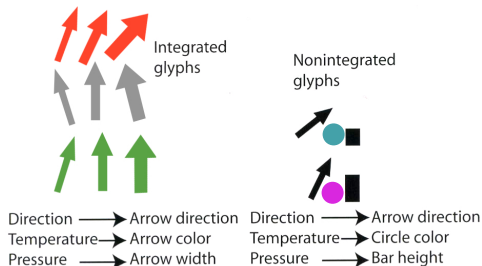
The symbols plotted at the left are *glyphs*

Each might represent a single case in a dataset

Source:

Ware, Information Visualization

Glyphs and Data Display



Source:

Ware, *Information Visualization*

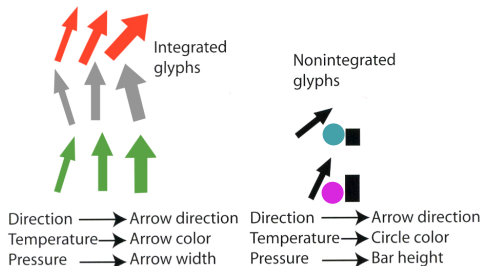
The symbols plotted at the left are *glyphs*

Each might represent a single case in a dataset

But each glyph can carry multiple dimensions

This is best achieved using an *integrated* set of glyph characteristics, one per dimension

Glyphs and Data Display



Number Dimensions of a glyph

\geq

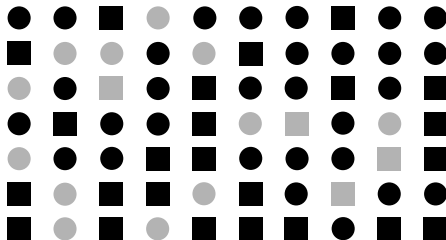
Number of variables encoded

The more variables you encode to dimensions of glyphs, the harder it is to pre-attentively separate the dimensions

Source:

Ware, *Information Visualization*

Glyphs and Data Display



Number Dimensions of a glyph

\geq

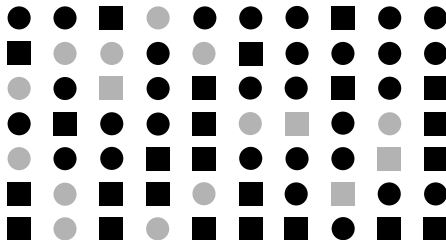
Number of variables encoded

The more variables you encode to dimensions of glyphs, the harder it is to pre-attentively separate the dimensions

Quick! Pick out the **gray squares**

Source: Ware, *Information Visualization*

Glyphs and Data Display



Number Dimensions of a glyph

\geq

Number of variables encoded

The more variables you encode to dimensions of glyphs, the harder it is to pre-attentively separate the dimensions

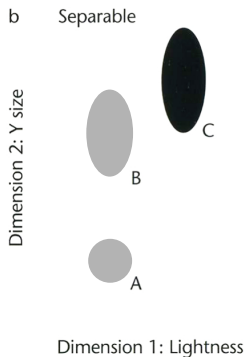
Quick! Pick out the **gray squares**

Source: Ware, *Information Visualization*

This may be an acceptable price for structured comparison across many dimensions

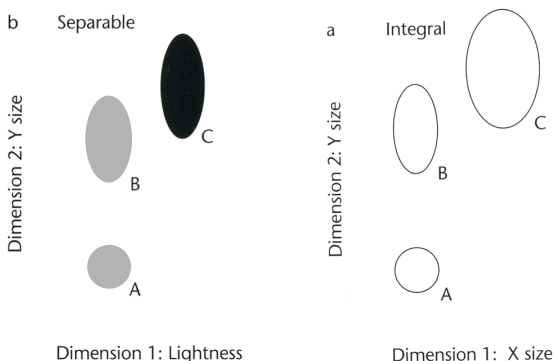
Sometimes, the best graphic – which is the simplest one that makes the desired point – still takes a bit of study to fully comprehend

Glyphs and Data Display



Take care to choose glyph dimensions that can be cleanly separated

Glyphs and Data Display



Take care to choose glyph dimensions that can be cleanly separated

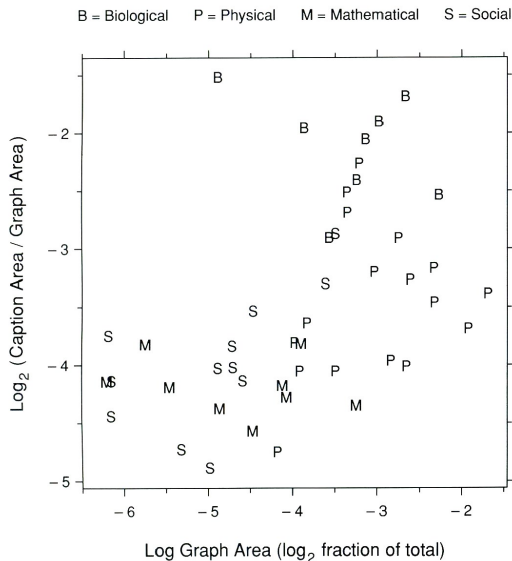
Sometimes, dimensions are reinforcing – and tend to blur together.
This makes it harder to extract information from the plot

Glyphs and Data Display

Glyphs most effective for encoding *discrete* data

Which glyphs should we use?

Glyphs and Data Display



Glyphs most effective for encoding *discrete* data

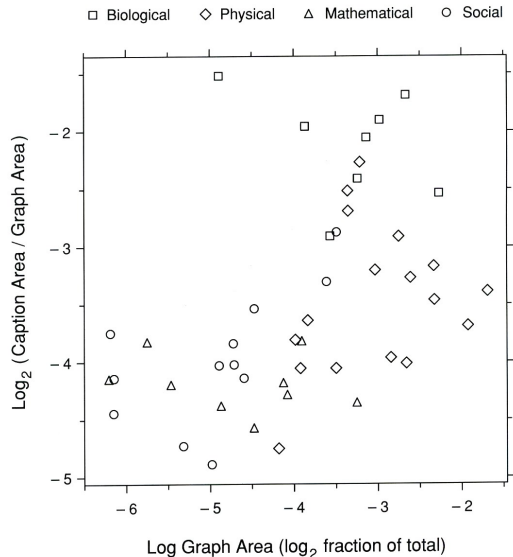
Which glyphs should we use?

First letter of each category?

Easiest to remember; hardest to preattentively distinguish

Source:
Cleveland, *Elements of Graphing Data*

Glyphs and Data Display



Glyphs most effective for encoding *discrete* data

Which glyphs should we use?

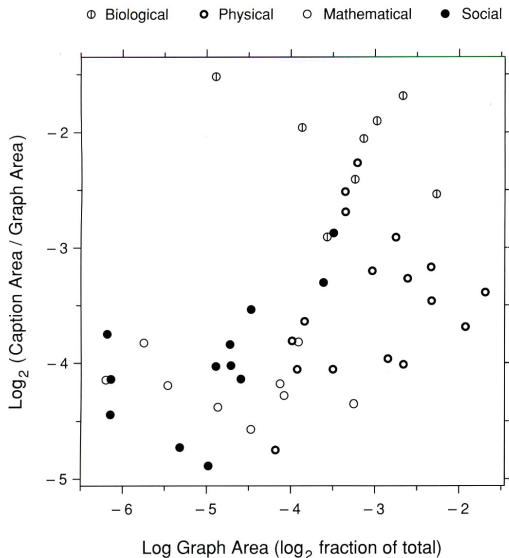
Random open symbols?

Reveal overlaps; still hard to preattentively distinguish

Source:

Cleveland, *Elements of Graphing Data*

Glyphs and Data Display



Not too many overlapping datapoints?

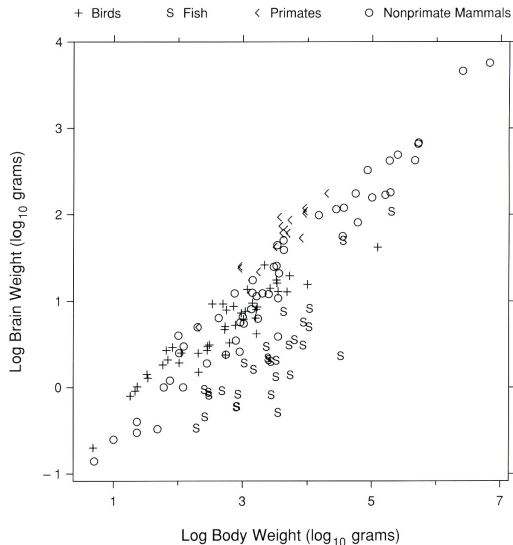
Then Cleveland recommends these glyphs:

○ ● ○ ⊖ ⊕

Hard to remember;
Easy to preattentively
distinguish and mentally
group

Source:
Cleveland, *Elements of
Graphing Data*

Glyphs and Data Display



Lots of overlapping
datapoints?

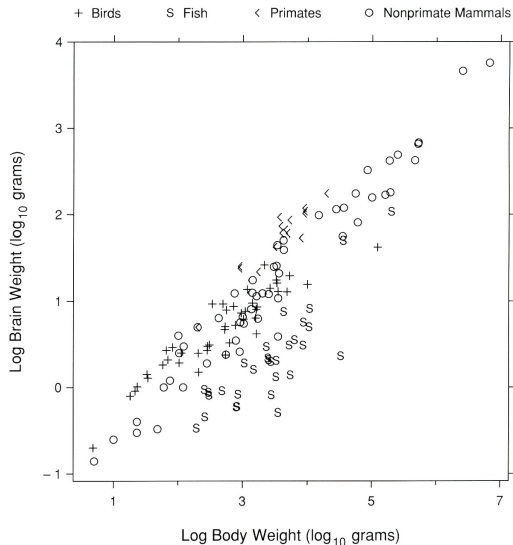
Then Cleveland
recommends these glyphs:

O + < S w

Hard to remember;
Easy to preattentively
distinguish, even in pieces

Source:
Cleveland, *Elements of
Graphing Data*

Glyphs and Data Display



Glyphs most effective for encoding *discrete* data

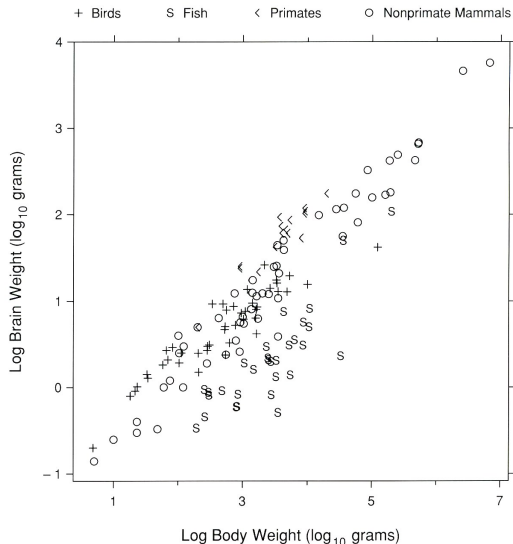
Which glyphs should we use? *It depends*

Always tradeoffs between preattentive distinctions and category recall

What matters will vary by display

Source:
Cleveland, *Elements of Graphing Data*

Glyphs and Data Display



Glyphs most effective for encoding *discrete* data

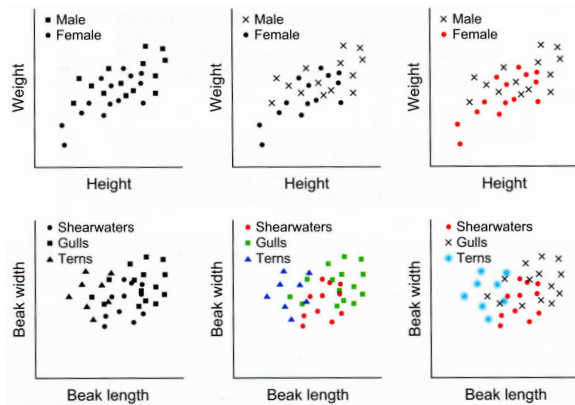
Which glyphs should we use? *It depends*

Note we haven't used color or transparency.

If we add these features, we can differentiate even identical, overlapping glyphs.

Source:
Cleveland, *Elements of Graphing Data*

Glyphs and Data Display



Source: Ware, *Information Visualization*

The dimensions of a glyph need not always code different variables

Coding the same variables to multiple pre-attentive features can help emphasize distinctions and pull apart distributions, even when points are strongly overlapping

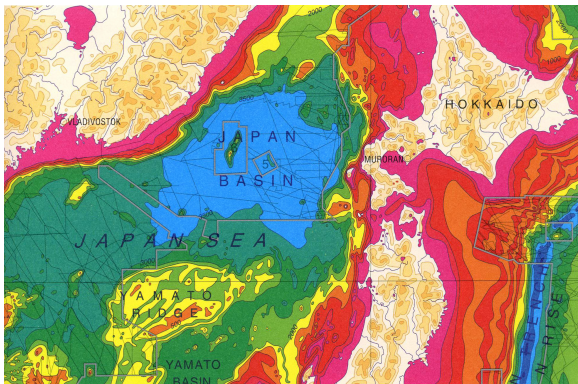
Using Color (In)effectively

While striking and colorful, this map is not clear or useful

The color scale is so ineffectively chosen that we likely wouldn't know this was a map if the place names were missing

Source:

Tufte, *Information Visualization*



Using Color (In)effectively

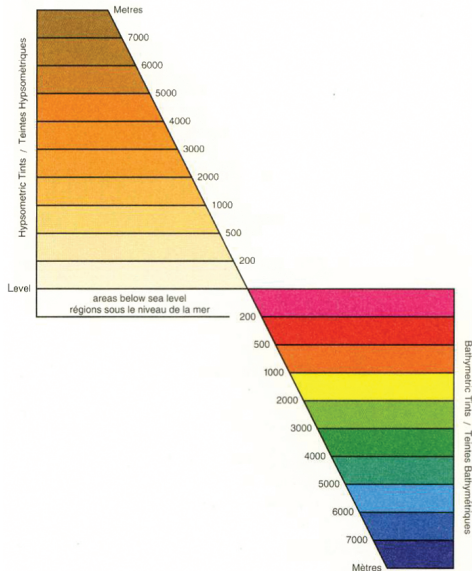
The mapmaker used a rainbow scale for underwater depth

Normally, mapmakers maintain a constant hue for a terrain type, and vary its brightness and saturation

Our eyes can order brightness, but not the rainbow

Source:

Tufte, *Information Visualization*



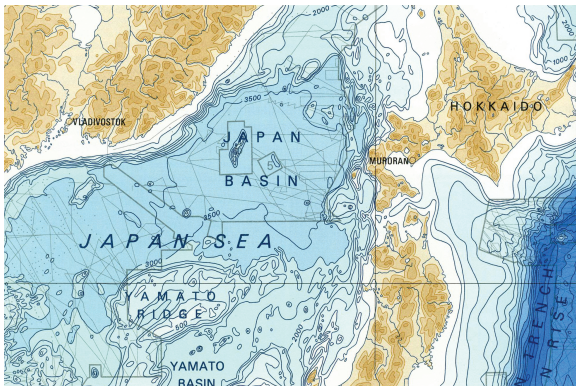
Using Color Effectively

For centuries, cartographers have effectively used color on maps

While not as flashy as a rainbow scale, this map is far more effective for both lookup and comparison

Source:

Tufte, *Information Visualization*



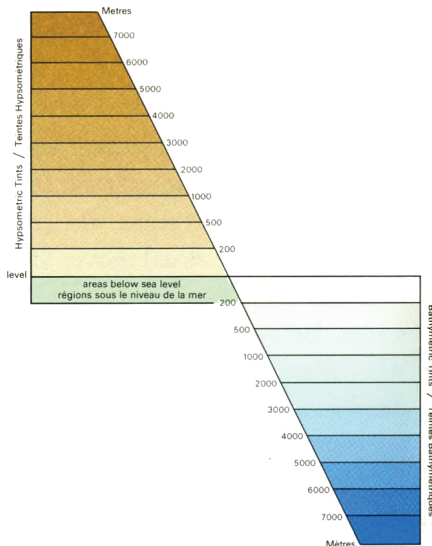
Using Color Effectively

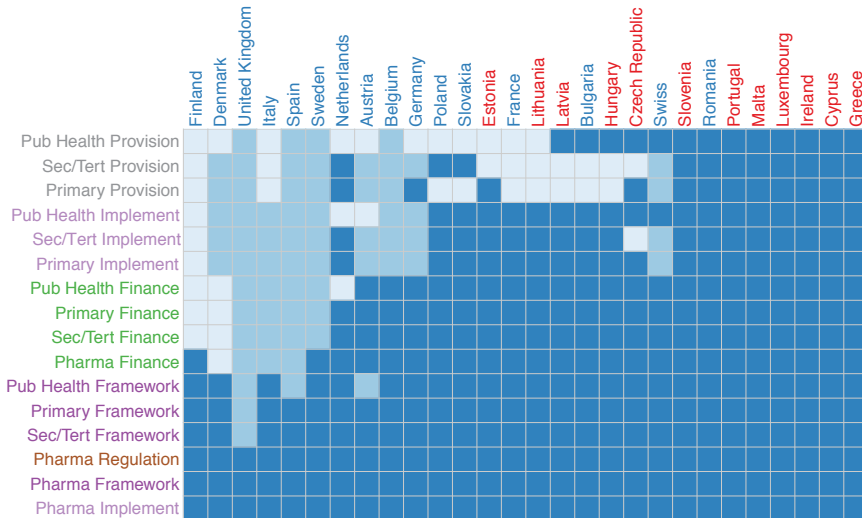
Humans are bad at precise reading of color

When plotting a quantitative variable on a color scale, care should be taken to find pre-attentively smooth gradients

Source:

Tufte, *Information Visualization*





Available Levels of Government

Two Tiers

Three Tiers

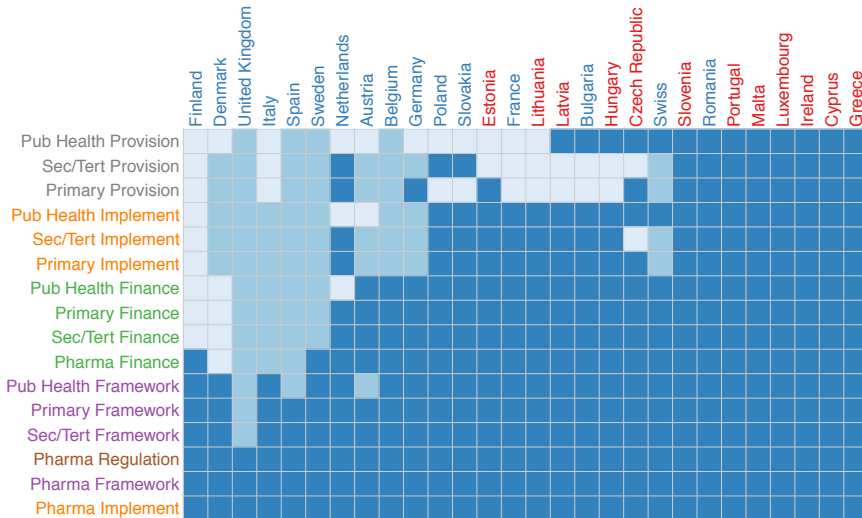
Observed Allocation of Authority

Local

Regional

State

What if I'd drawn my heatmap with *these* colors



Available Levels of Government

Two Tiers

Three Tiers

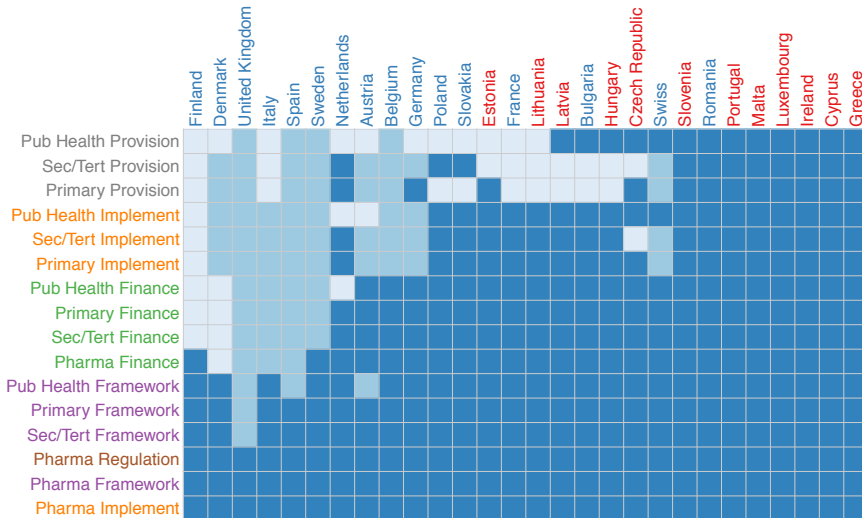
Observed Allocation of Authority

Local

Regional

State

Poor choice of colors for categorical data can distort the data



Available Levels of Government

Two Tiers

Three Tiers

Observed Allocation of Authority

Local

Regional

State

Choose colors for categories to achieve equal pairwise distinctions

If you only learn four things about using color in graphs, it should be to:

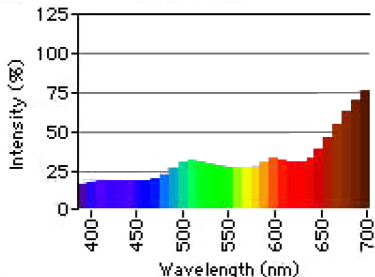
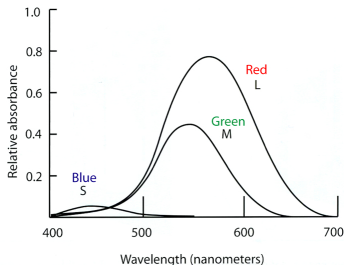
- 1 Choose colors for quantities using pre-attentively smooth gradients
- 2 Choose colors for categories to achieve equal pairwise distinctions
- 3 Avoid overlapping colors with similar brightness (value)
- 4 Use pastels for large-area colors and saturated colors for small points

To understand and implement the above,
we need to know something about the science of color

Color Science Vastly Oversimplified by an Outsider

Color is

- a measure of the wavelength of light
- something we perceive
- a mixture of red, blue, and green
- a mixture of hue, luminosity, and saturation
- broadly categorical; also continuous
- hard to perceive accurately
- hard to reproduce accurately
- an element of many visuals



Color Science Vastly Oversimplified by an Outsider

Human eyes contain two kinds of photo-receptive elements:

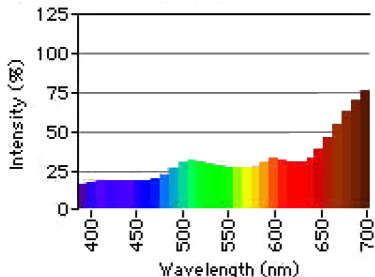
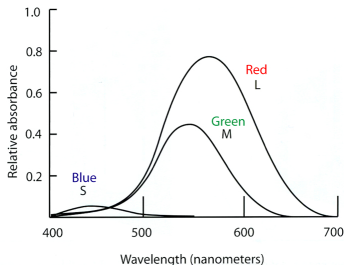
Rods. Sensitive to brightness

Single photon receptors

Little use in sunlight

Cones. Come in three varieties...

- 1 Short wavelength (**red**);
most sensitive
- 2 Medium wavelength (**green**);
moderately sensitive
- 3 Long wavelength (**blue**); weak



Color Science Vastly Oversimplified by an Outsider

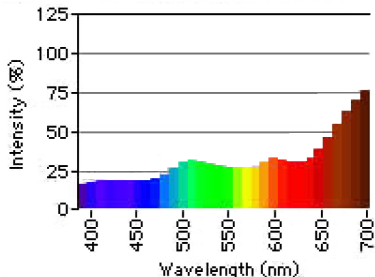
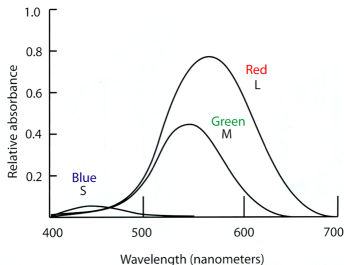
Humans are best at seeing red;
worst at seeing blue

Species vary in color vision ability:

- dogs have only two cones, are red-green colorblind, and see less detail in daylight
- birds have more cones than humans – chickens have 12!

Number of cones = number of primary colors a species perceives

Mixing the three (human) primaries in different amounts makes any color humans can see.

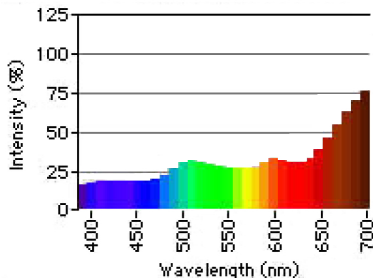
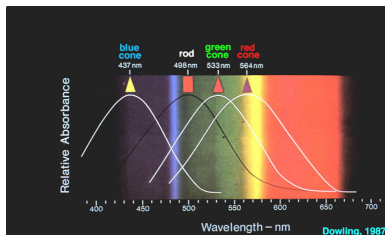


Color Science Vastly Oversimplified by an Outsider

Suppose we used scientific equipment to measure the number of photons received at each of N wavelengths throughout the visible spectrum

Clearly, we could choose N to be arbitrarily large, and still learn more about the distribution of photon wavelengths by sampling in still more places

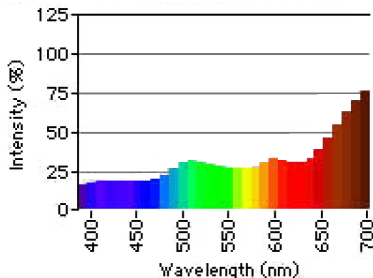
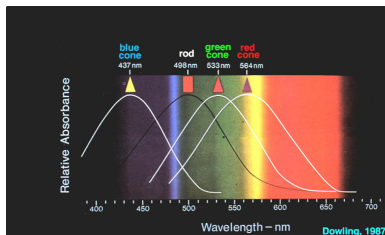
Just like adding bins to a histogram, this gives us a finer view of the distribution



Color Science Vastly Oversimplified by an Outsider

The human eye has cones tuned to just three wavelengths, so our eyes approximate complex histogram as mixture of three densities

Many different colors (that is, distinct histograms on the visual spectrum) that look the “same” to us would look different to a chicken, which samples the distribution in 12 places!



Color Spaces

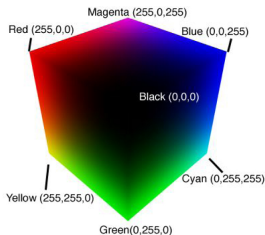


Figure 3a: The black corner (0,0,0) of the RGB Cube

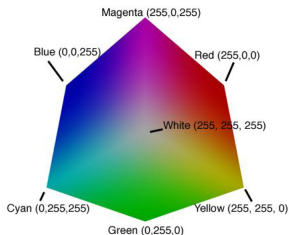


Figure 3b: The white corner (255,255,255) of the RGB Cube

Primaries and color can be expressed in many equivalent ways.

These are different **colorspaces**: mappings from 3 variables to a color

Computer space · **RGB**

Red, Green, Blue

Printer space · **CMYK**

Cyan, Magenta, Yellow, Black

Artist space · **HSV**

Hue, Saturation, Value

Brain space · **CIElab**

Lightness, blue/yellow, red/green

Color Spaces: RGB

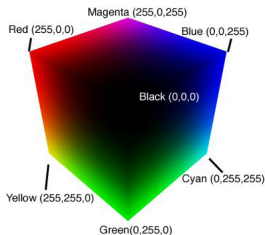


Figure 3a: The black corner (0,0,0) of the RGB Cube

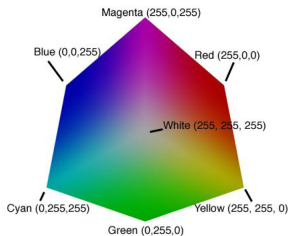


Figure 3b: The white corner (255,255,255) of the RGB Cube

RGB is mainly useful for telling a computer what color of light to display

If you want to tell a printer what color of ink to print, it is easier to use CMYK: cyan, magenta, yellow, and black

But neither color space is useful for choosing colors, either aesthetically or scientifically

Color Spaces: HSV

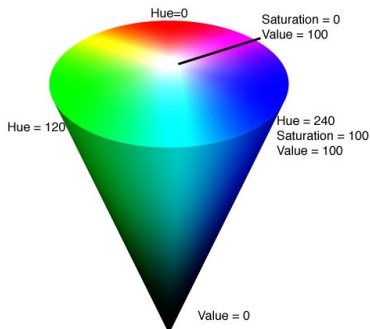


Figure 2: The HSV Cone

Source: Darrin Cardani, "Adventures in HSV Space"

Artists find RGB an inconvenient space to think about color

Instead, they often use HSV:
hue is the "name" of the color

think rainbow

saturation is the richness of the color;
desaturated colors have been mixed with gray

think solids vs pastels

value is the brightness of the color;
how much white or black is mixed in

think stop sign vs. red wine

Color Spaces: HSV

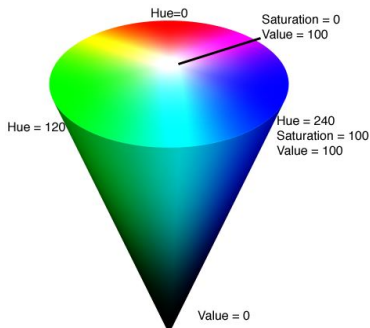


Figure 2: The HSV Cone

HSV is useful for constructing beautiful complementary colors for artistic palettes

Colors on opposite sides of the cone are aesthetically harmonious contrasting colors

Colors 120 degrees apart (color triads) are too

Any HSV color can also be represented in RGB and CMYK, and vice versa

Source: Darrin Cardani, "Adventures in HSV Space"

Color Spaces: Opponent Color Theory

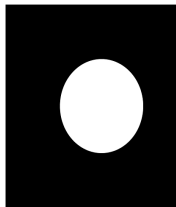
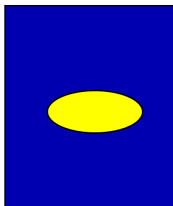
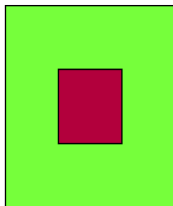
Does your brain read off RGB values from your cones? Or maybe HSV values?

Probably not.

Opponent color theory:

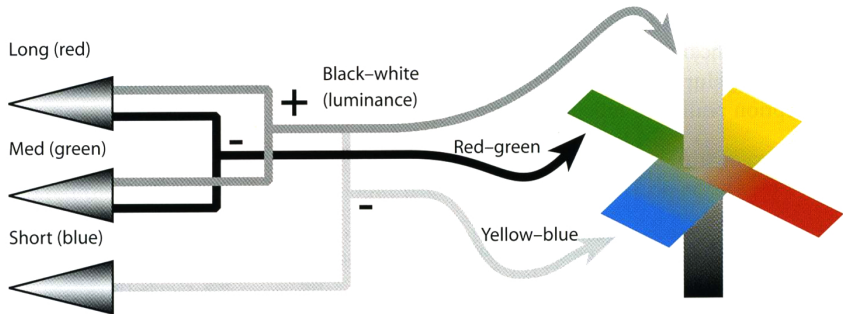
Human optical system converts $\{S,M,L\}$ cone readouts to three channels

- Redness vs. greenness
- Blueness vs. yellowness
- Brightness



Source: Maya Gupta, Electrical Engineering, UW

Color Spaces: Opponent Color Theory



Source: Colin Ware, *Information Visualization*

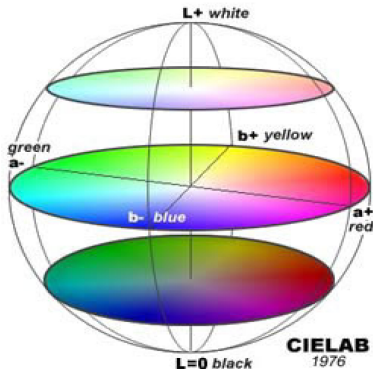
In other words, red/green and blue/yellow are “opponent colors”

They appear in zero-sum combinations
(no one ever says “the yellowish-blue sweater”)

Color blindness: absence or weakness of one set of cones

Most common are people who can’t distinguish red & green

Color Spaces: CIElab



Opponent color theory suggests a new color space, CIElab

(CIE stands for *Commission internationale de l'éclairage*, or International Commission on Illumination)

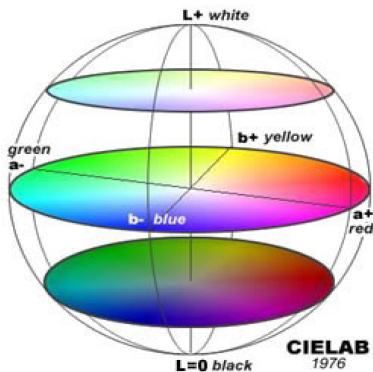
L = luminance (white vs. black)

a = red vs. green

b = blue vs. yellow

Equal Euclidian distances in CIElab space are (approximately) perceptually "equal" to humans

Color Spaces: CIElab



Equal Euclidian distances in CIElab space are (approximately) perceptually “equal” to humans

If CIElab is the brain’s color space, it’s the best one for choosing colors to convey precise scientific information

If you want to convey distinct categories, choose colors that are well separated in CIElab space

If you want to convey precise numerical steps, choose equal steps through CIElab space

The Cognitive Science of Visual Displays of Information



Does this mean you need to learn a lot of cognitive science before you can make a color graphic?

Not really.

Easy shortcuts available: RColorBrewer will choose appropriate colors for you

The Cognitive Science of Visual Displays of Information

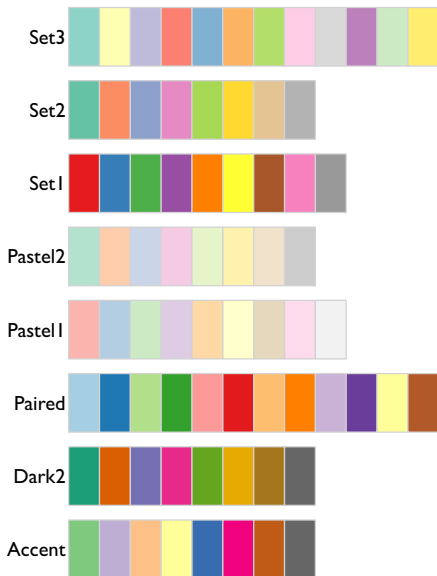


At left are perceptually equal gradients for different color hues, as suggested by RColorBrewer

```
library(RColorBrewer)  
display.brewer.all(type="seq")
```

Pick one horizontal strip for your color scale to plot quantitative data

The Cognitive Science of Visual Displays of Information



RColorBrewer will also suggest colors for qualitative variables

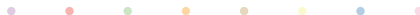
Goal here is to make each category equally distinct from the others

```
library(RColorBrewer)  
display.brewer.all(type="qual")
```

Why so many choices? Not for aesthetics, but because they solve different color cognition problems

Suppose we use Pastel1 to encode categories to glyphs

Can you easily tell which color is which?



Suppose we use Pastel1 to encode categories to glyphs

Can you easily tell which color is which?



Hard to distinguish the hue of small areas of desaturated color

Don't use pastels to color small glyphs

Suppose we use Pastel1 to encode categories to regions

Can you easily tell which color is which?



Suppose we use Pastel1 to encode categories to regions

Can you easily tell which color is which?



Easy to distinguish the hue of large areas of desaturated color

Use pastels to color large regions

Suppose we use Set1 to encode categories to glyphs

Can you easily tell which color is which?



Suppose we use Set1 to encode categories to glyphs

Can you easily tell which color is which?



Easy to distinguish the hue of small areas of saturated color

Use jewel tones to color large regions

Suppose we use Set1 to encode categories to regions

Would a graph with large bright regions be readable?



Suppose we use Set1 to encode categories to regions

Would a graph with large bright regions be readable?



Large areas of saturated color command attention – distract from small details

Avoid jewel tones when coloring large regions

Use jewel tones for glyphs

Use pastels for regions



Use jewel tones for glyphs

Use pastels for regions



Avoid pastel glyphs and saturated regions!



Use jewel tones for glyphs

Use pastels for regions



Avoid pastel glyphs and saturated regions!



Use jewel tones for glyphs

Use pastels for regions



Avoid pastel glyphs and saturated regions!



Color Cognition Problem 2

Background Contrast

Text is only readable when it differs significantly from the background in *value*

Dark text only works on light backgrounds

Legible text requires
value contrast

Legible text requires
value contrast

Legible text requires
value contrast

Color Cognition Problem 2

Background Contrast

Text is only readable when it differs significantly from the background in *value*

Dark text only works on light backgrounds

Light text only works on dark backgrounds

Legible text requires
value contrast

Legible text requires
value contrast

Legible text requires
value contrast

Color Cognition Problem 2

Background Contrast

Text is only readable when it differs significantly from the background in *value*

Dark text only works on light backgrounds

Light text only works on dark backgrounds

Mid-value backgrounds make muddy images: *avoid*

Legible text requires
value contrast

Avoid mid gray
backgrounds

Legible text requires
value contrast

Color Cognition Problem 2

Background Contrast

Text is only readable when it differs significantly from the background in *value*

Mid-value backgrounds make muddy images: *avoid*

Applies to graphs generally: don't use a gray background to your plot

Warning: gray backgrounds are default in ggplot2 and Excel!

Legible text requires
value contrast

Avoid mid gray
backgrounds

Legible text requires
value contrast

Common mistaken intuition:

Different hues (“colors”) are sufficient to distinguish background and foreground

It is very difficult to read text that is isoluminant with its background color. If clear text material is to be presented it is essential that there be substantial luminance contrast with the background color. Color contrast is not enough. This particular example is especially difficult because the chromatic difference is in the yellow blue direction. The only exception to the requirement for luminance contrast is when the purpose is artistic effect and not clarity.



Common mistaken intuition:

Different hues (“colors”) are sufficient to distinguish background and foreground

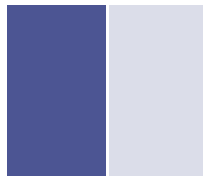
Even two color opposites (blue and yellow) can blend
when they have similar values (brightness)

It is very difficult to read text that is isoluminant with its background color. If clear text material is to be presented it is essential that there be substantial luminance contrast with the background color. Color contrast is not enough. This particular example is especially difficult because the chromatic difference is in the yellow blue direction. The only exception to the requirement for luminance contrast is when the purpose is artistic effect and not clarity.



To avoid unreadable text,
make sure background and foreground have different values

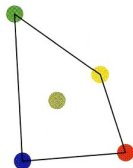
It is very difficult to read text that is isoluminant with its background color. If clear text material is to be presented it is essential that there be substantial luminance contrast with the background color. Color contrast is not enough. This particular example is especially difficult because the chromatic difference is in the yellow blue direction. The only exception to the requirement for luminance contrast is when the purpose is artistic effect and not clarity.



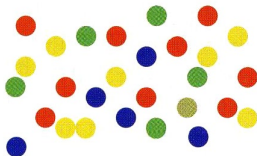
To avoid unreadable text,
make sure background and foreground have different values

With a large value contrast,
even background and foreground of the same hue can be effective

a



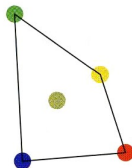
c



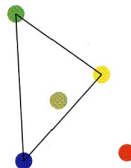
RColorBrewer chooses “equally distinct” colors. How?

Colors in a are plotted in CIElab space: Interior colors blend in

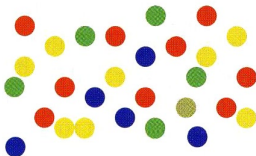
a



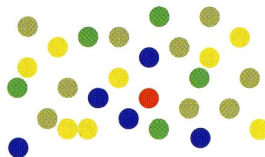
b



c



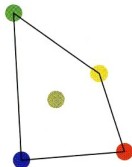
d



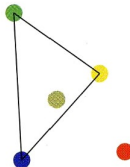
RColorBrewer chooses “equally distinct” colors. How?

Colors outside the convex *hull* of the other colors *stand out*

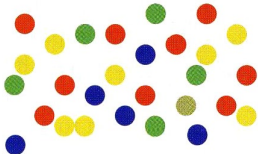
a



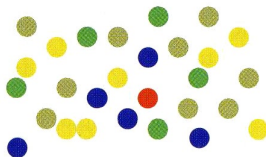
b



c



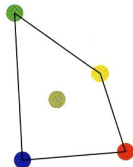
d



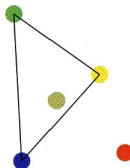
RColorBrewer chooses “equally distinct” colors. How?

RColorBrewer “qual” colors are equidistant from each other on the convex hull

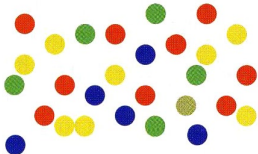
a



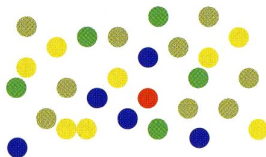
b



c



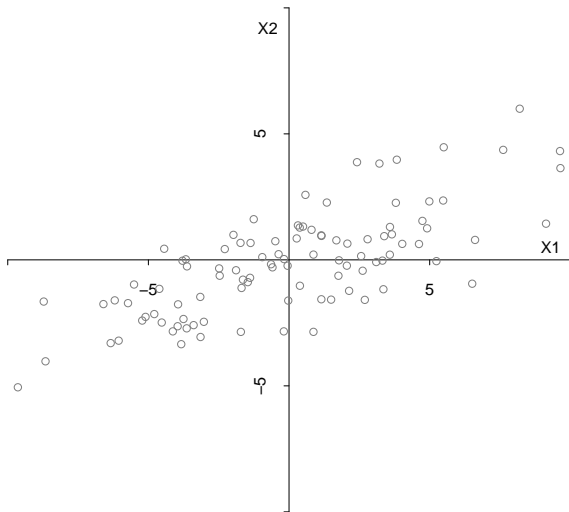
d



If your goal is to *highlight* a point or category,
choose something outside the convex hull of the other colors

Aside: Convex Hulls

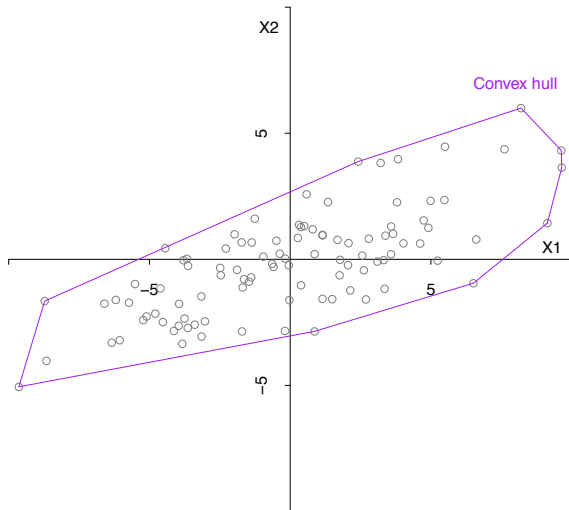
What is a convex hull?



Aside: Convex Hulls

What is a convex hull?

An elastic band wrapped around the cloud of points such that it contains the smallest convex set containing those points

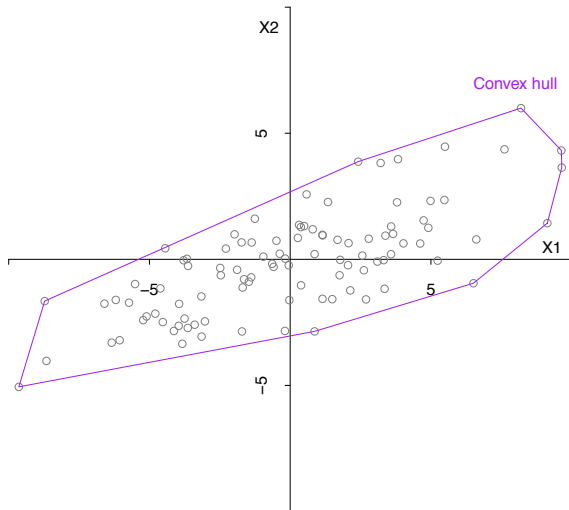


Aside: Convex Hulls

What is a convex hull?

An elastic band wrapped around the cloud of points such that it contains the smallest convex set containing those points

What is a convex set?



Aside: Convex Hulls

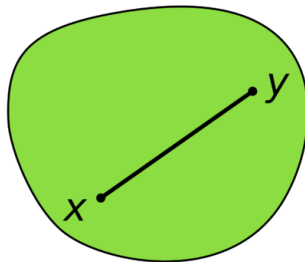
What is a convex hull?

An elastic band wrapped around the cloud of points such that it contains the smallest convex set containing those points

What is a convex set?

If a straight line between any two points in a region remains within that region, that region is a convex set

A Convex Set



Aside: Convex Hulls

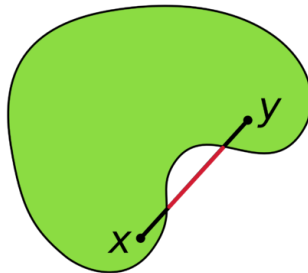
What is a convex hull?

An elastic band wrapped around the cloud of points such that it contains the smallest convex set containing those points

What is a convex set?

If a straight line between any two points in a region remains within that region, that region is a convex set

Not a Convex Set

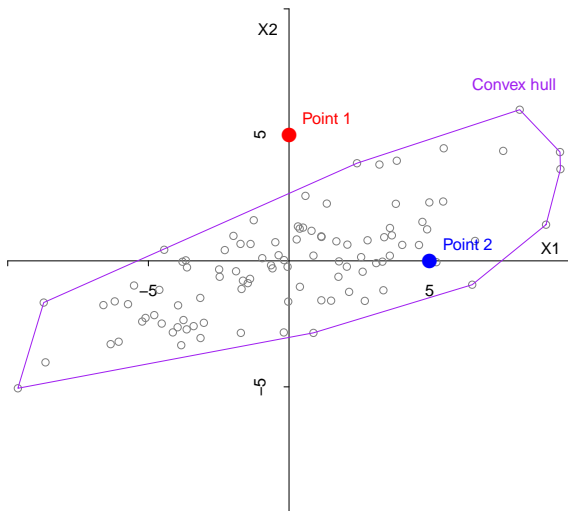


Aside: Convex Hulls

Convex hulls will come up again later when we discuss the difference between extrapolation from a dataset and interpolation from a dataset

Point 2 is interpolated

Point 1 is extrapolated



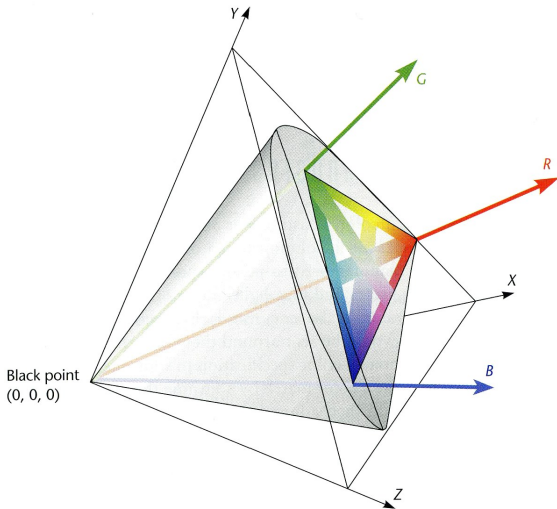
Color Cognition Problem 4 *Color Reproduction & Color Gamuts*

LCD displays can't capture
all the colors humans can
see

Or all the colors a good
printer can print

\$ and fancy chemistry
needed to get all perceived
colors

Fascinating history of
science topic:
discovery of pigments for
various colors took centuries



Color Cognition Problem 4 Color Reproduction & Color Gamuts

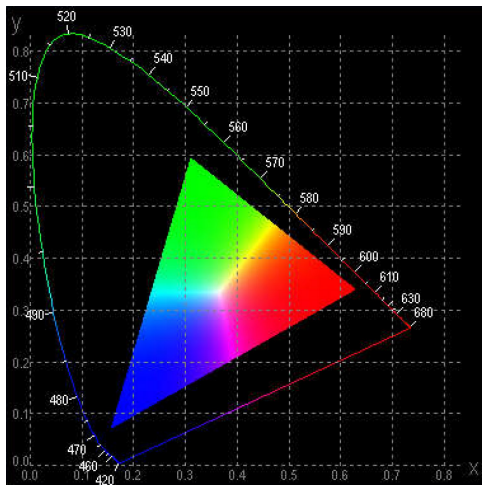
Focus on the LCD display
versus human perception
problem

Most greens and blues are
missing!

OLED helps, but rarely
available

Bigger problem: device
dependence

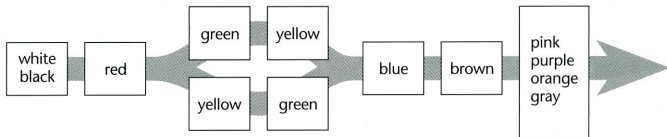
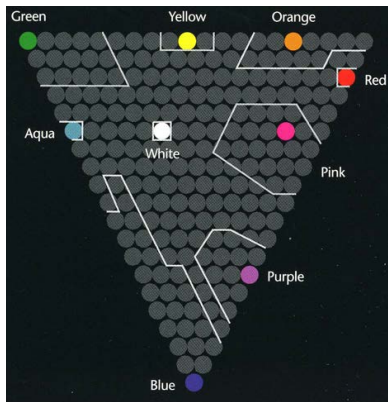
This LCD projector \neq my
computer \neq your computer
 \neq your printer



Why colors work for categories:

Humans recognize certain hues as categorically distinct

Note the misclassification of red and orange in terms of primaries



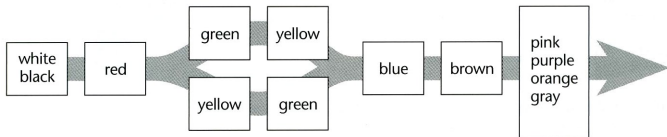
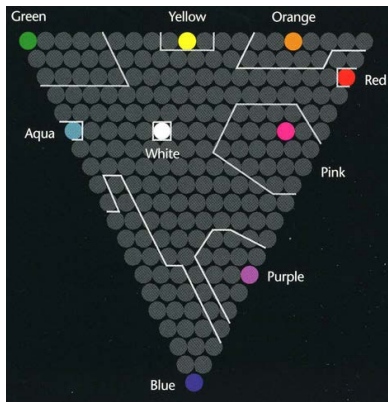
Note:

misclassification of red
and orange in terms of
primaries

narrow range of reds

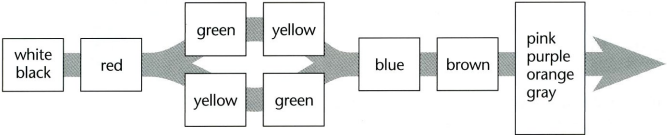
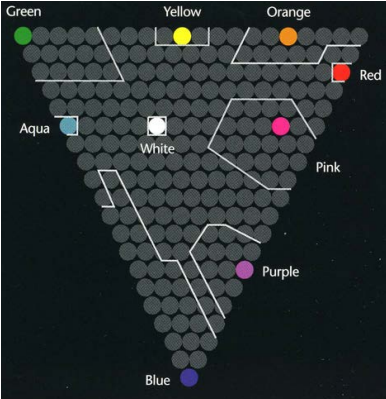
wide range of blues

Result of strong red
cones/weak blue cones



Where is **brown** on this chart?

For that matter, where is **brown** in the color spectrum?

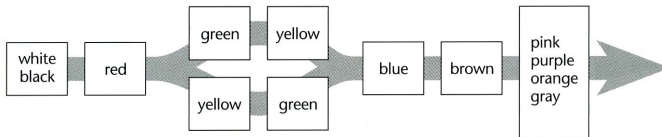
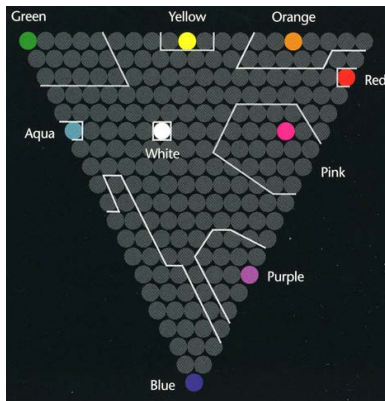


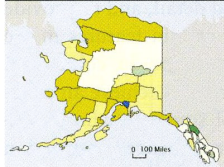
Where is **brown** on this chart?

For that matter, where is **brown** in the color spectrum?

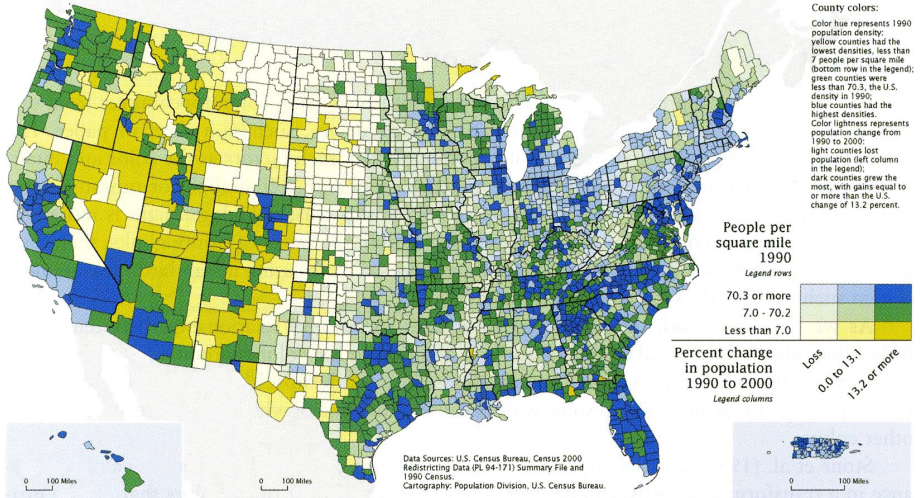
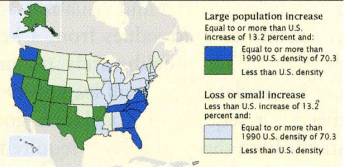
Browns are dark yellows and dark oranges

An imperfection in mapping of hue to human color perception



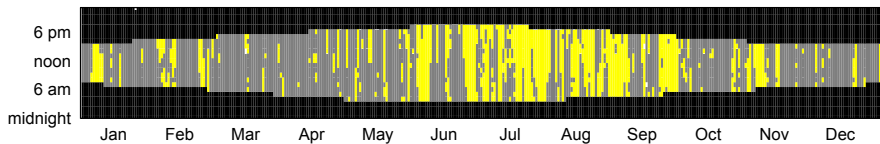


Percent Change, 1990 to 2000 and Population Density, 1990

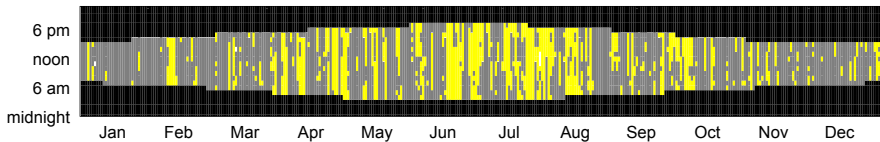


2003 Seattle, WA

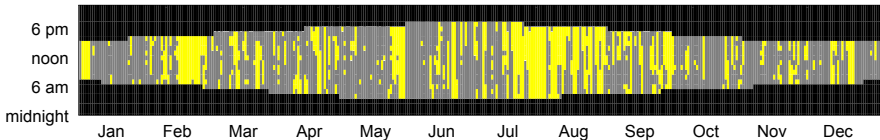
Daylight/Cloudcover/Night Cocoon



2004 Seattle, WA



2005 Seattle, WA



Graphic: Christopher Adolph (www.chrisadolph.com/). Data: NOAA and US Naval Observatory.