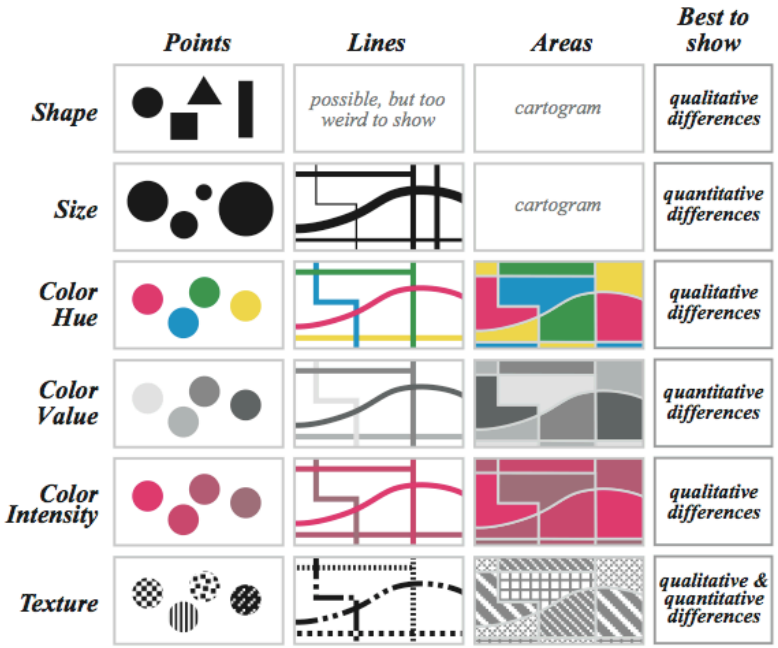
IAT 355



Visual Encoding

Lyn Bartram

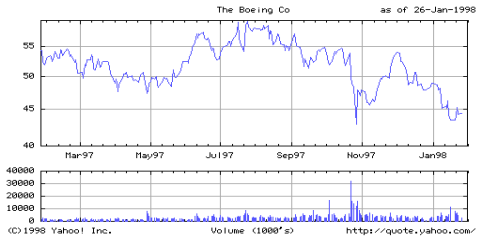
SFU

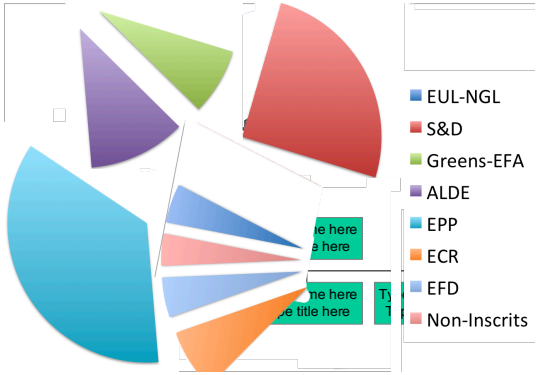
SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY

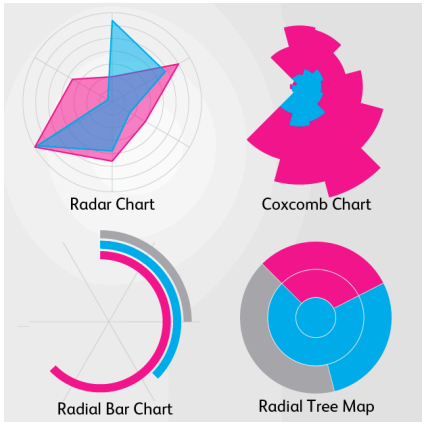
Note: Some of these slides have been borrowed and adapted from T. Munzner and J.Heer

Common Vis Types

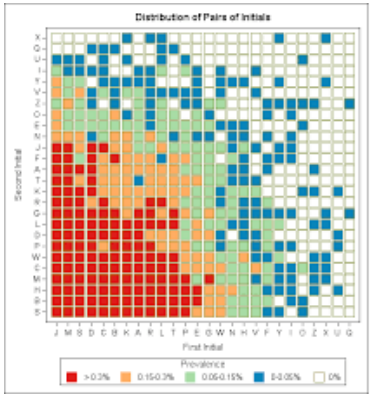
- Graphs



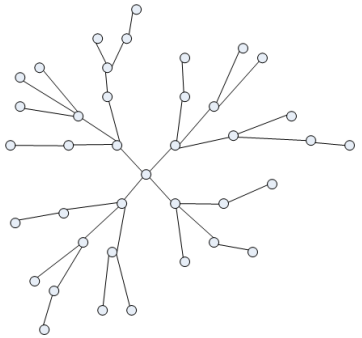




- Maps

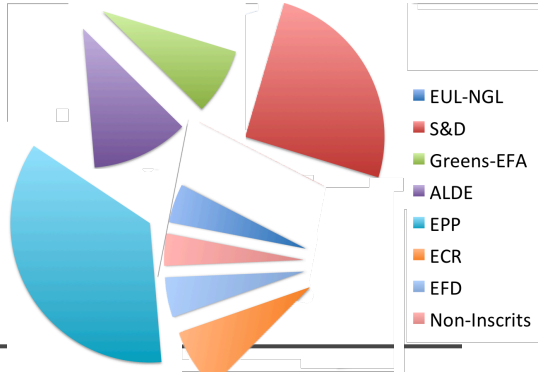


- Networks



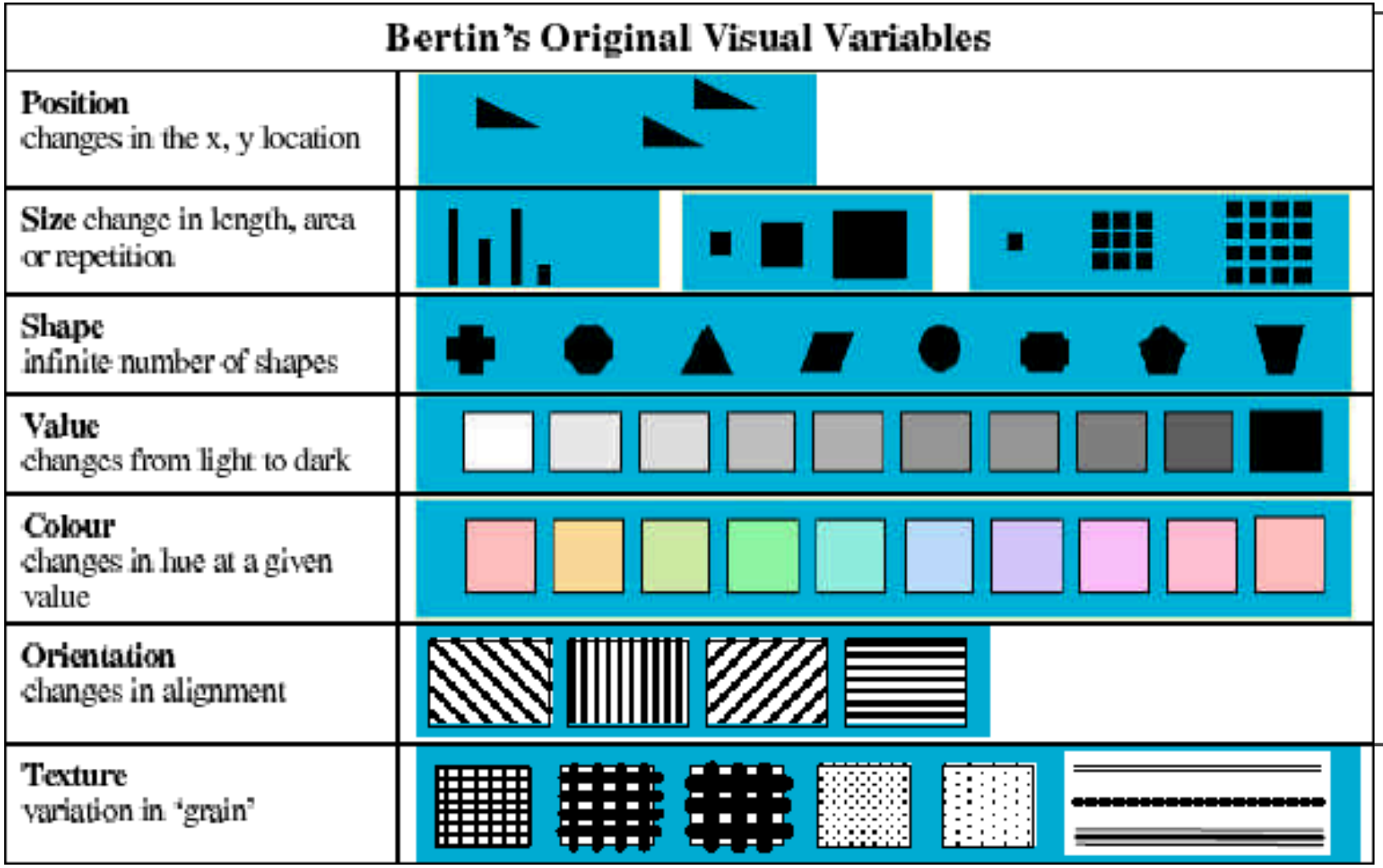
Slide adapted from Marti Hearst

THE BIG QUESTION



- How do you decide how to map data to appearance?
- Image models and visual language
  - Data → Visual feature
- Semiology [Bertin] : an image is perceived as a set of signs and “retinal variables”

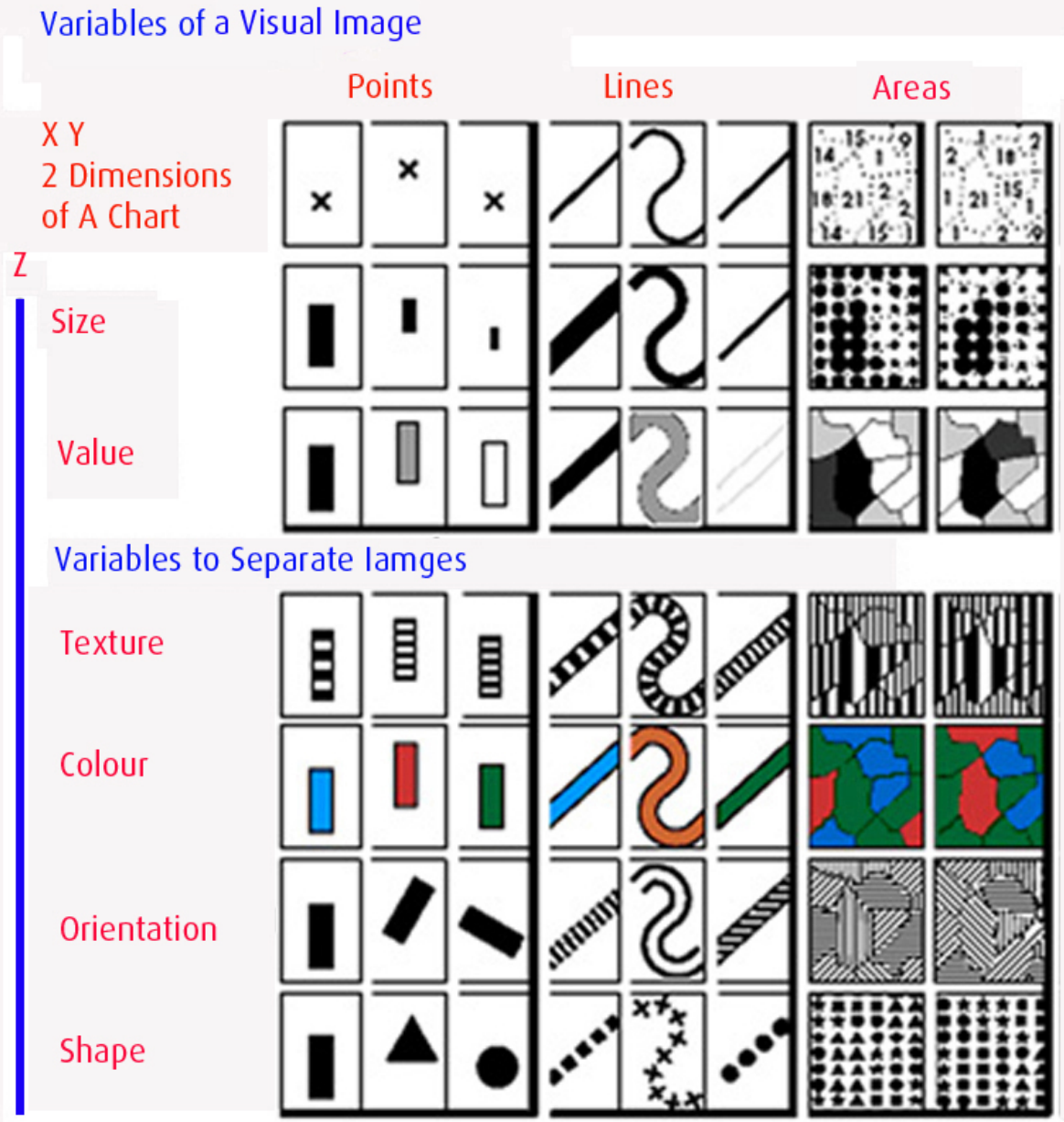
Bertin's graphical language



Visual variables

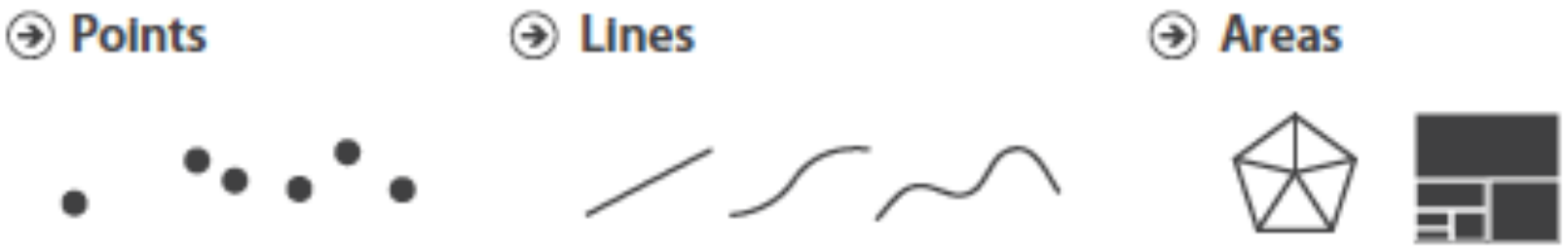
- Position
- Size (length, area volume)

- Value
- Texture
- Color
- Orientation
- Shape
- Transparency
- Blur / Focus …….

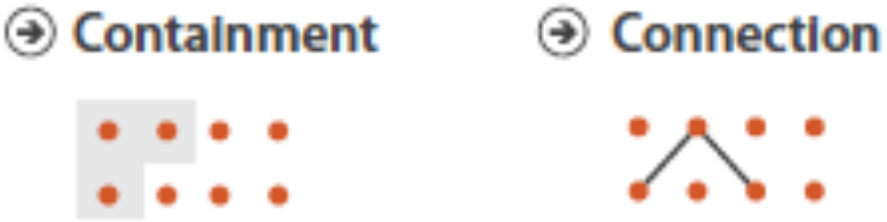


Marks and channels [Munzner]

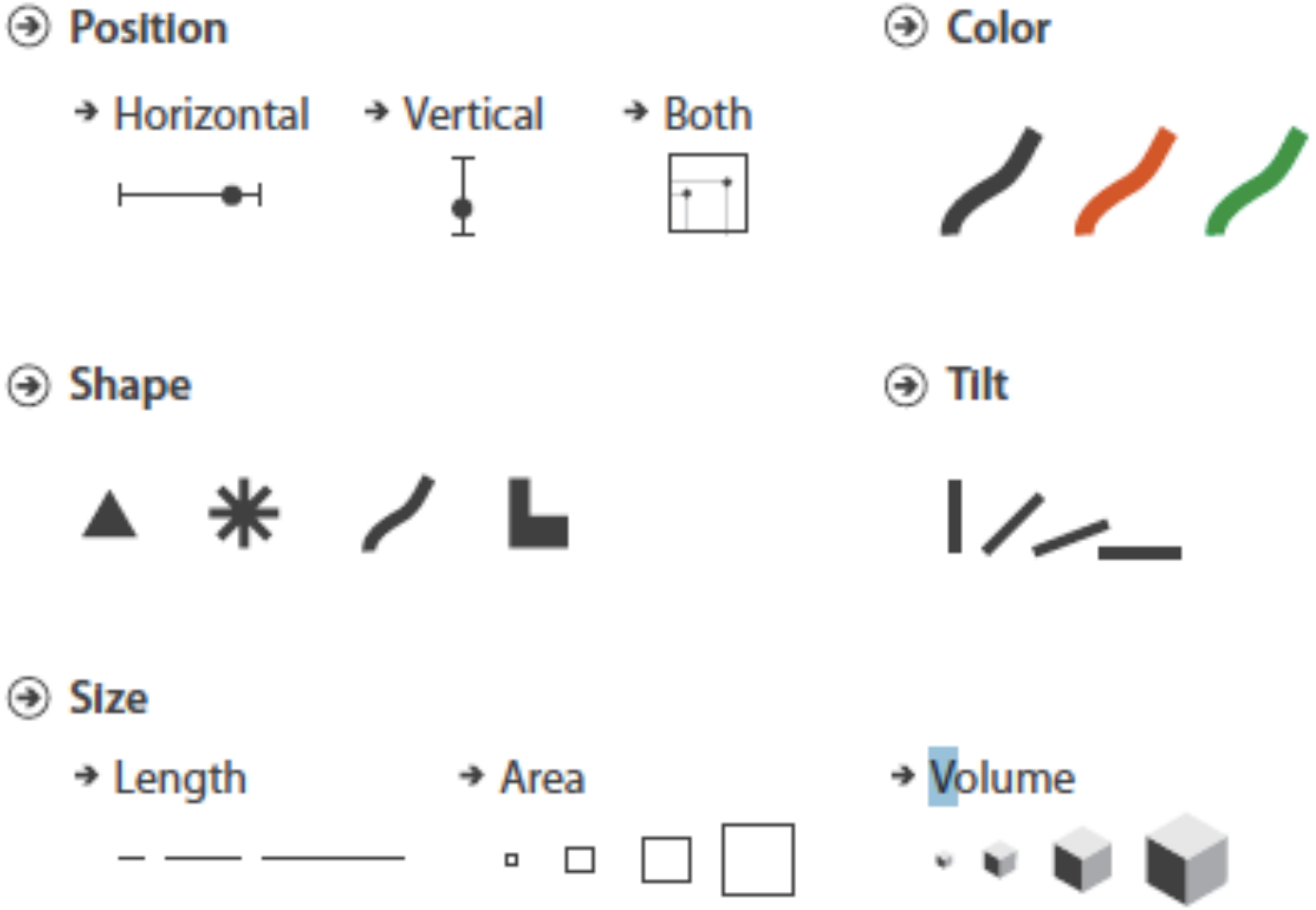
- *Marks* are geometric primitives (items , tabular data)



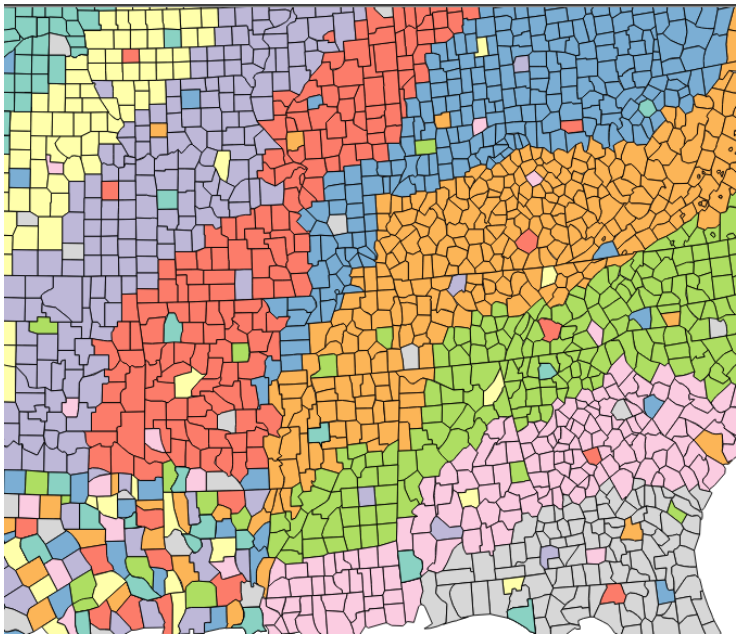
- Or links



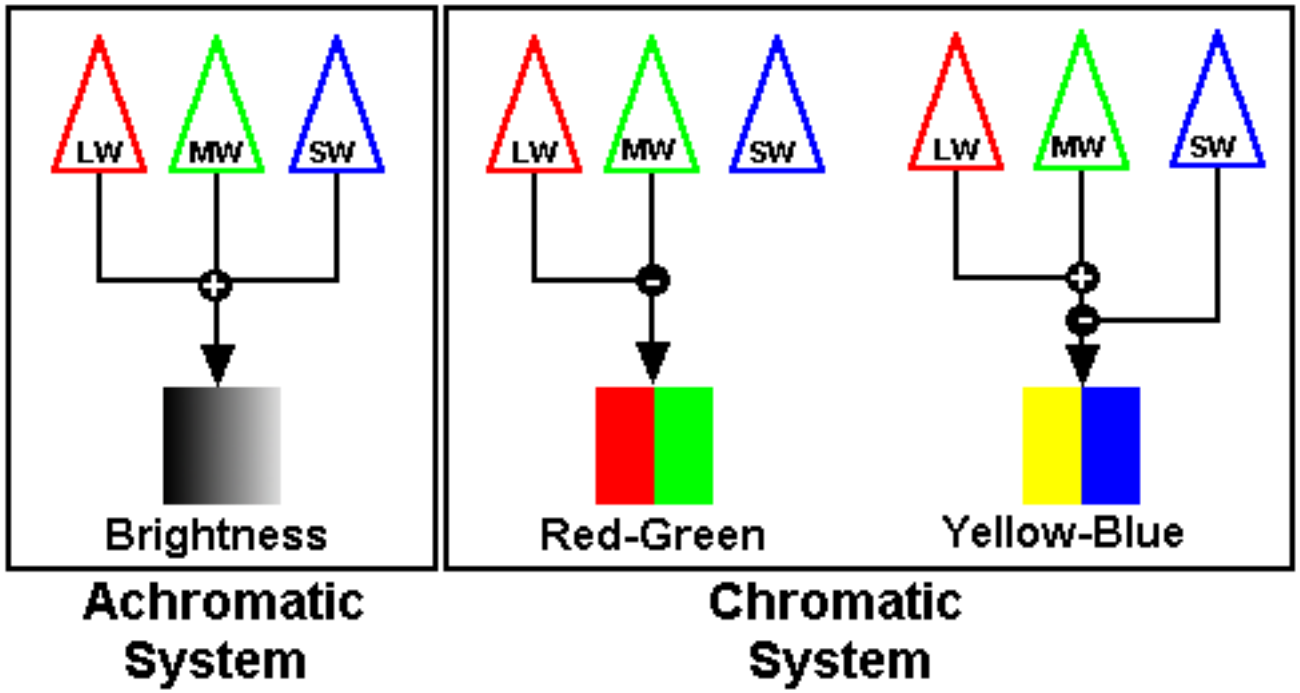
Channels control the appearance of marks



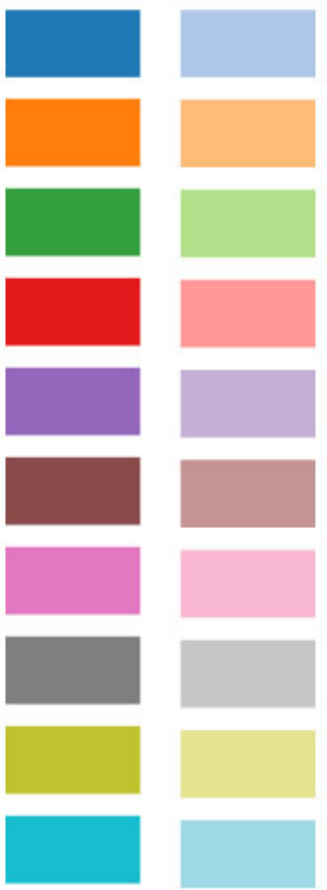
A side trip into COLOUR



How we see colour



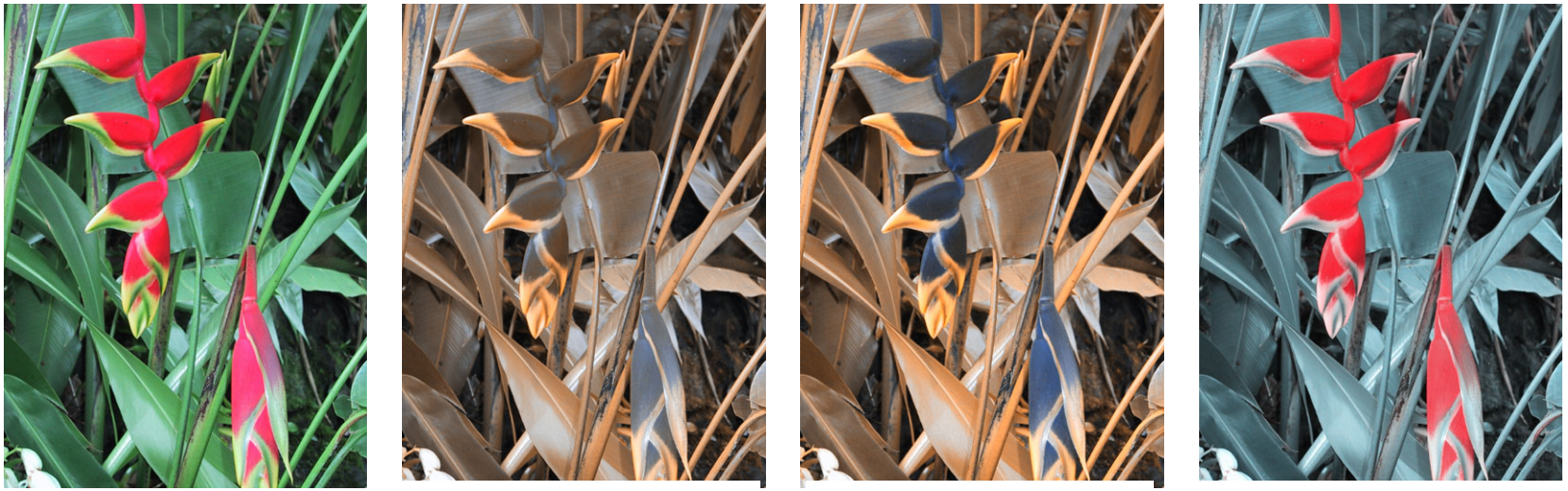
2D Color Space



Normal

Colour deficiency

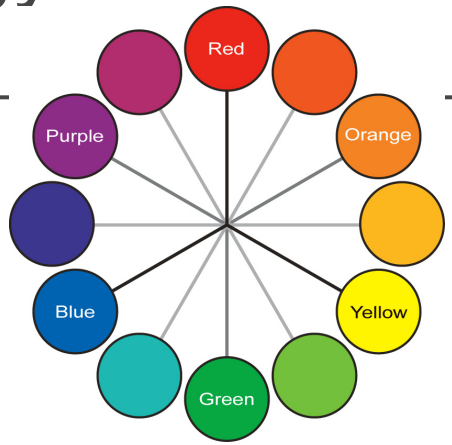
- Simulates color vision deficiencies
  - Photoshop plug-in



Deuteranope — Protanope — Tritanope

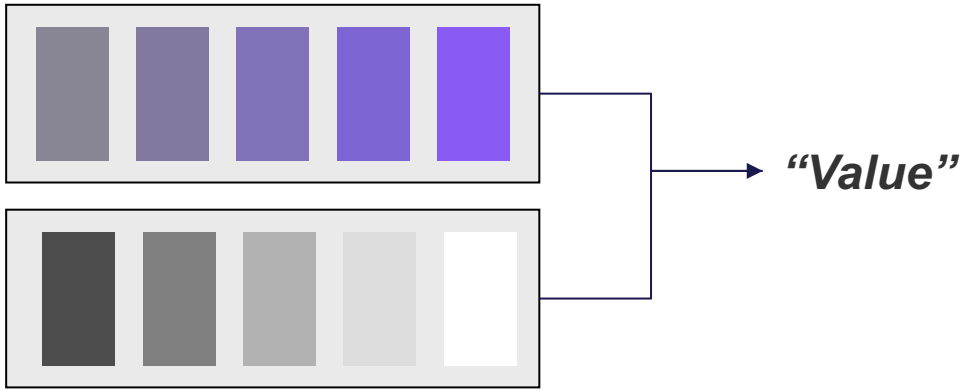
Color Design Terminology

- ***Hue* (color wheel)**
  - Red, yellow, blue (primary)
  - Orange, green, purple (secondary)
  - Opposites complement (contrast)
  - Adjacent are analogous
- *See www.handprint.com for examples

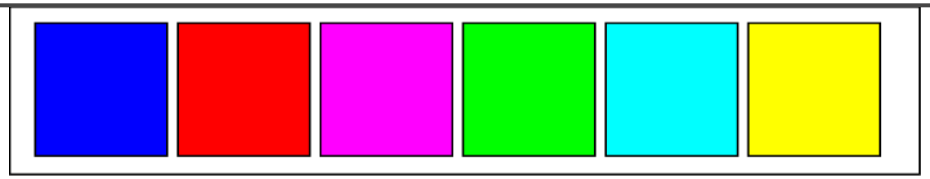


- **Chroma (saturation)**
  - Intensity or purity
  - Distance from gray

- **Lightness**
  - Dark to light
  - Applies to all colors

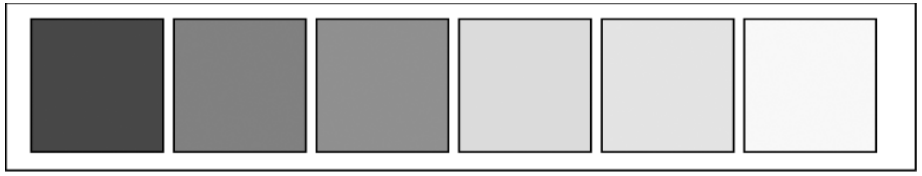


Colour spaces are not perceptually even

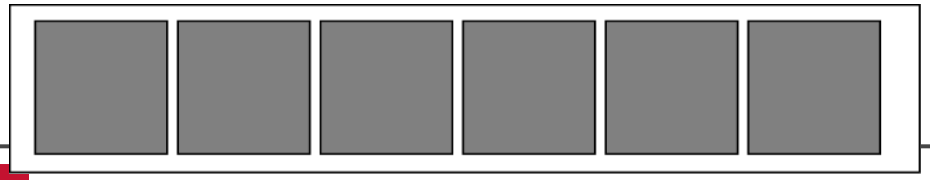


Corners of the RGB color cube

Luminance values (retinal response)



L* values

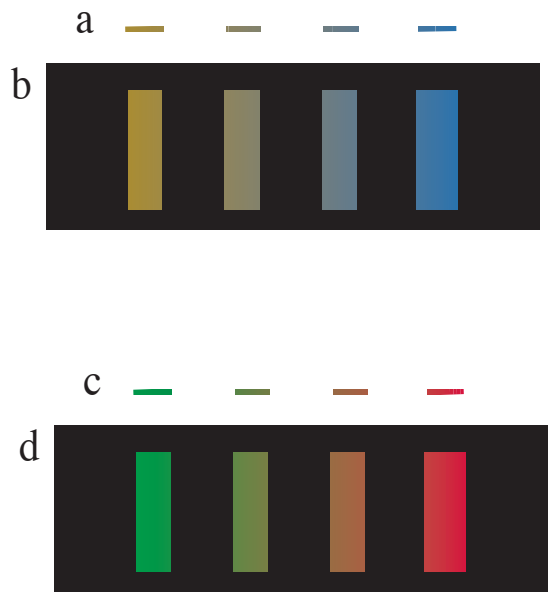


L from HLS
All the same

Size matters

- As the stimulus becomes small, we can't distinguish colour
  - dichromat
- This is more problematic with blue-yellows

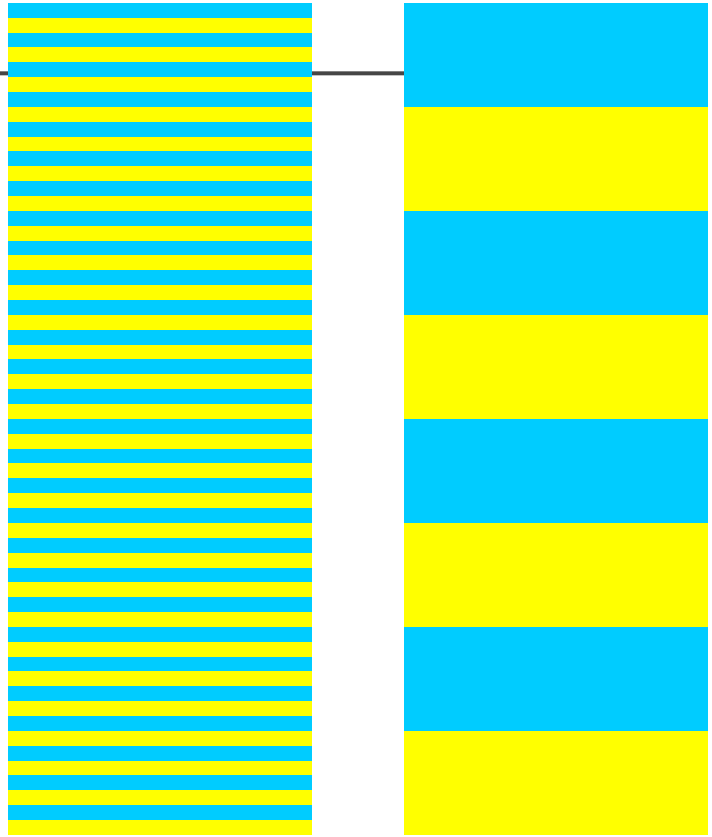
Small field tritanopia



Spreading

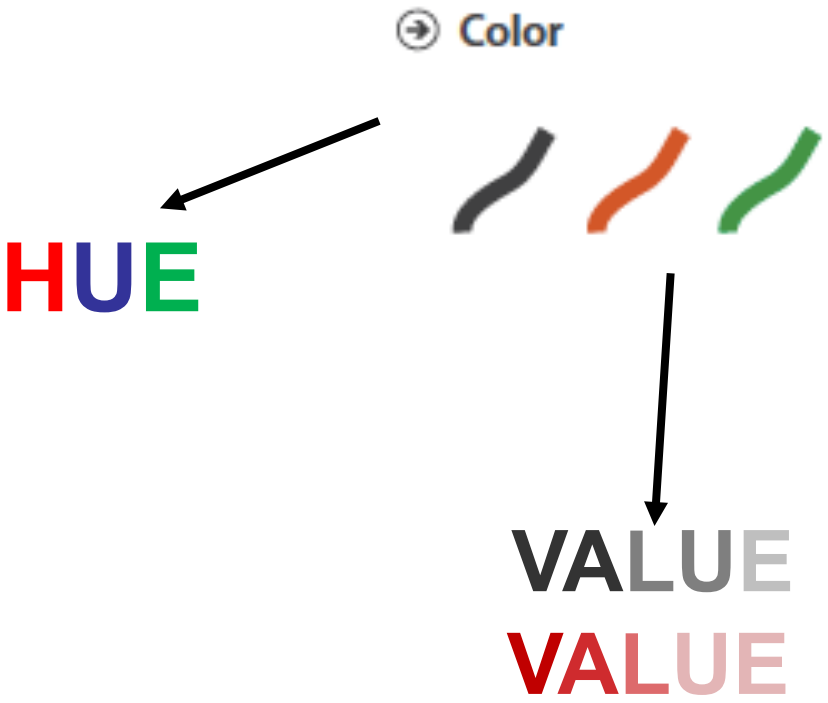
- Spatial frequency
  - The paint chip problem
  - Adjacent colors blend

- The higher the spatial frequency, the less saturated the colour



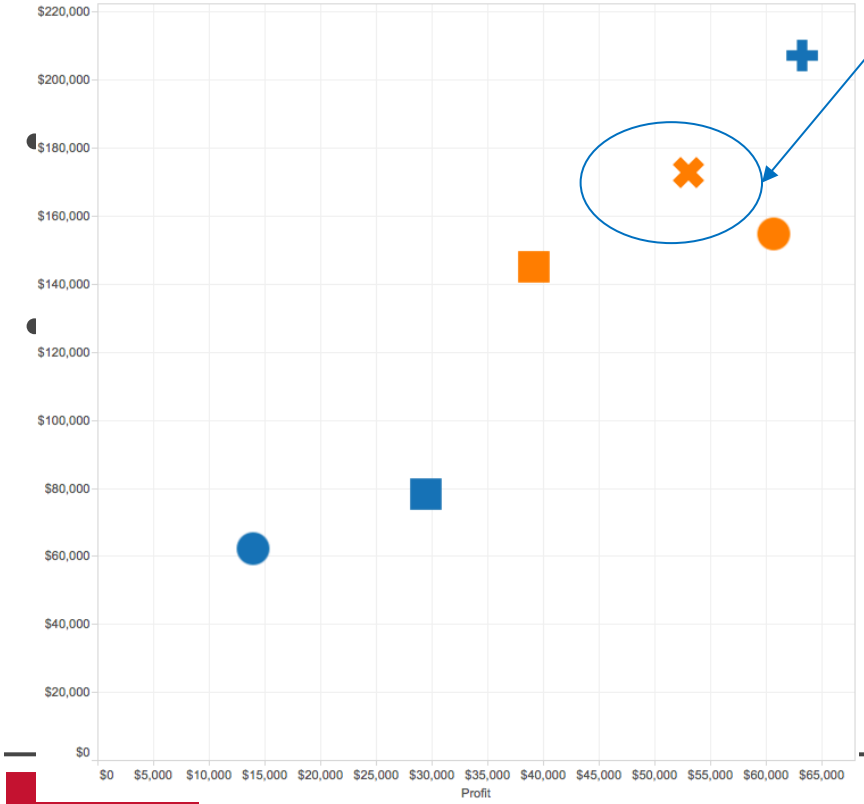
Redrawn from *Foundations of Vision*

Two colour channels

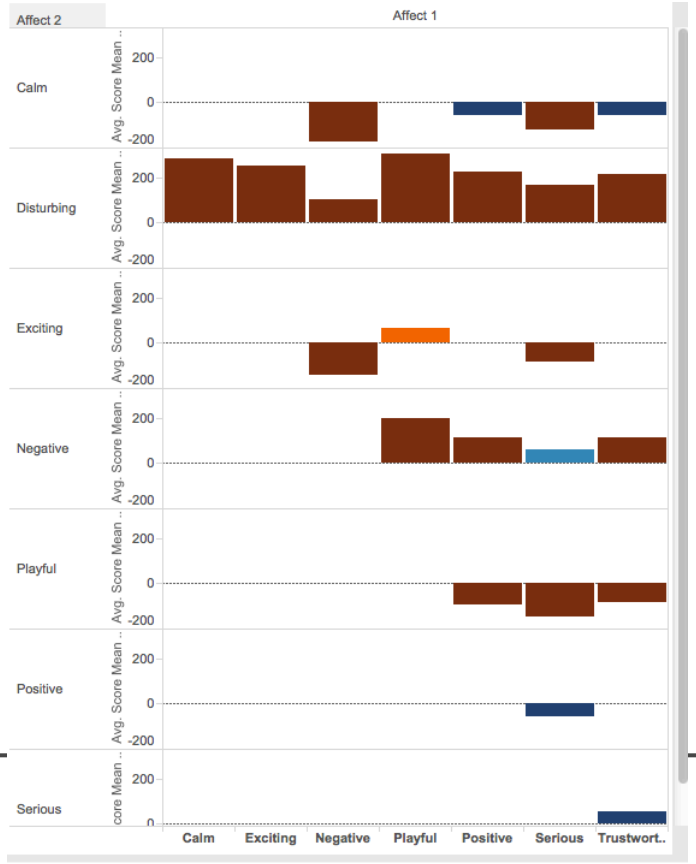


Channel types [Munzner]

Identity



Magnitude

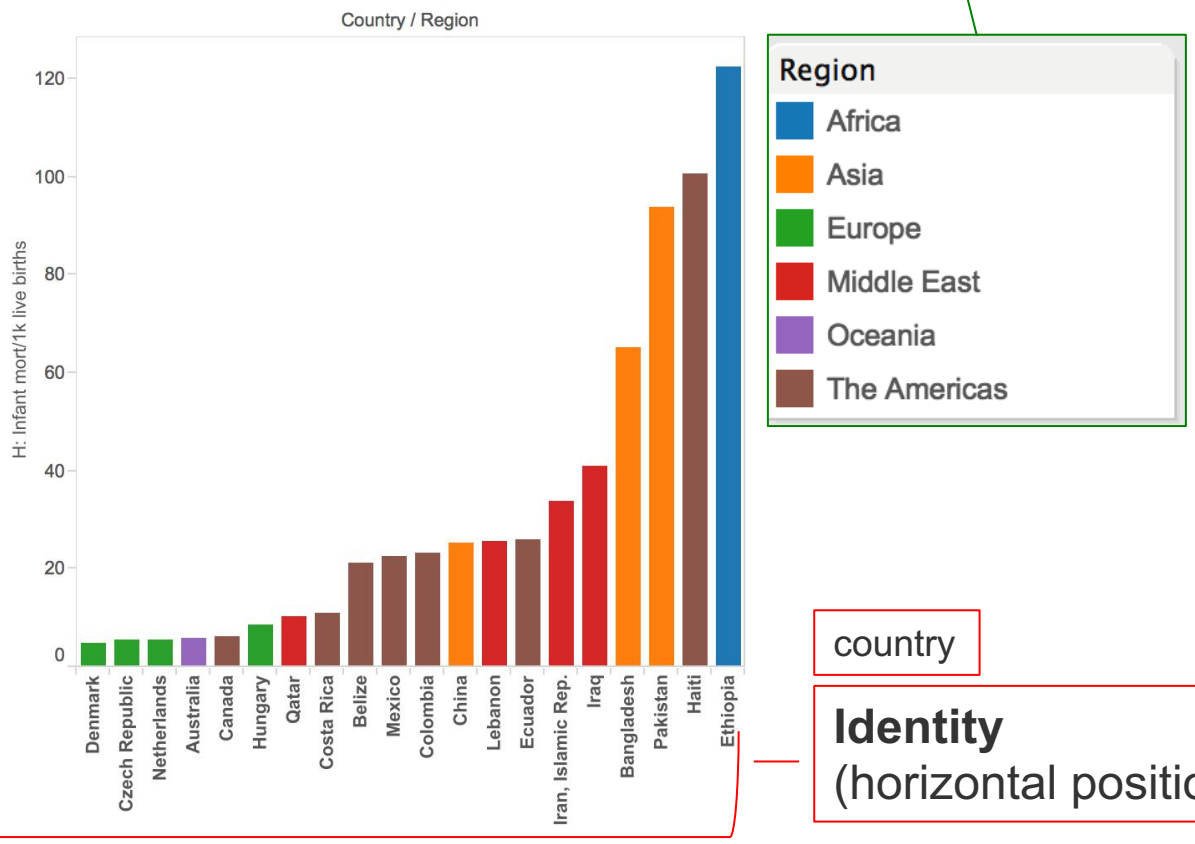


Channels and marks

Identity
Line mark, colour (hue)

Infant mortality

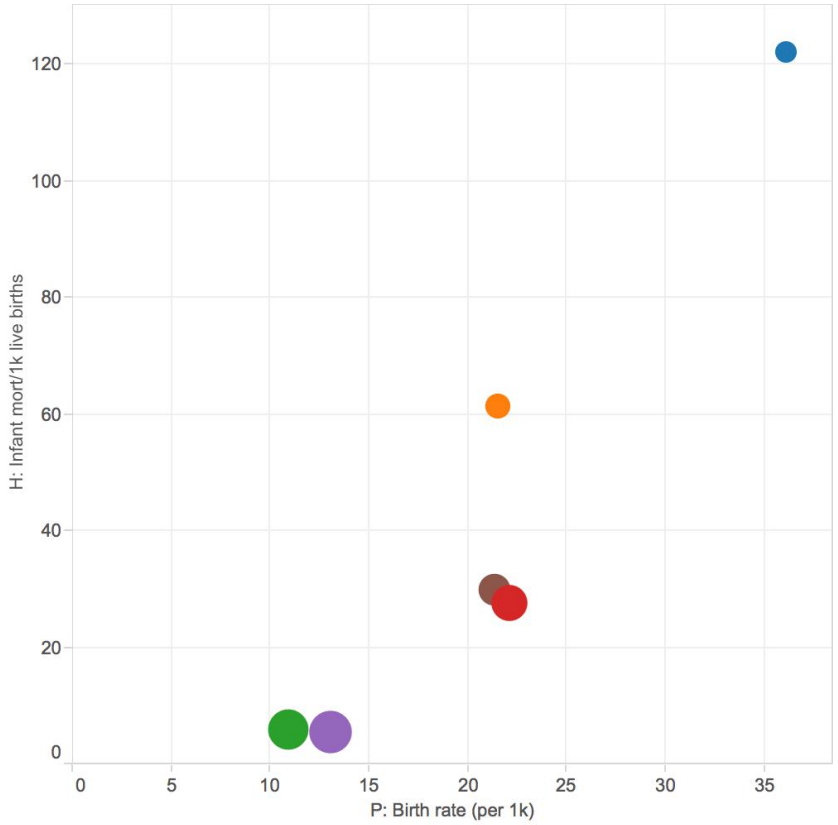
Magnitude
Line mark, Height (spatial position)

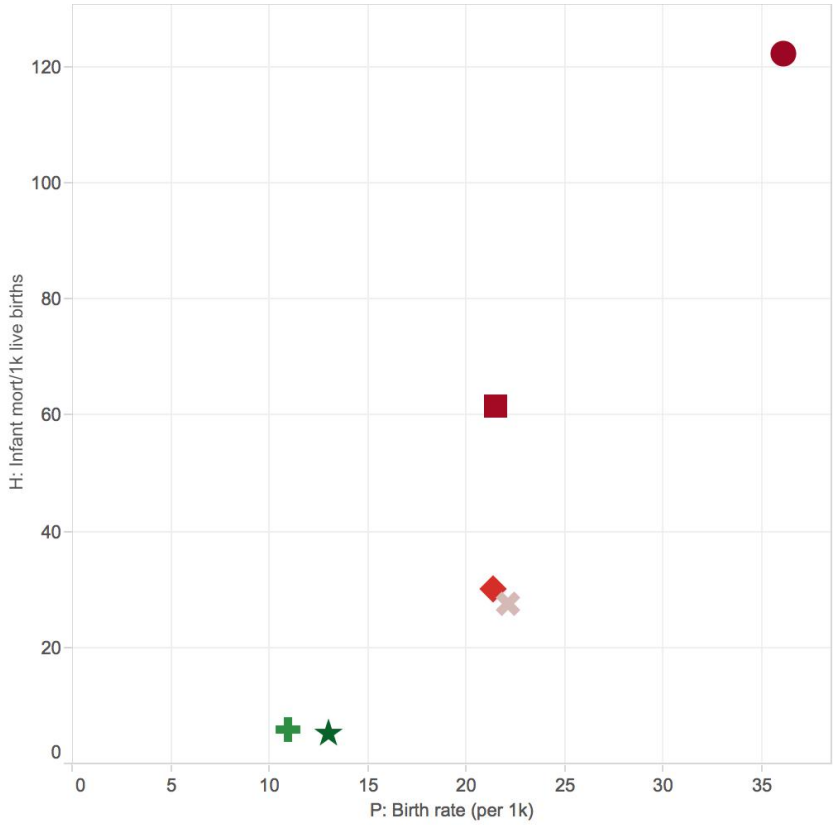


country

Identity
(horizontal position)

Channels are limited





Visual decoding tasks

Identity

1. Selection/Discrimination:
   - Is A different from B?

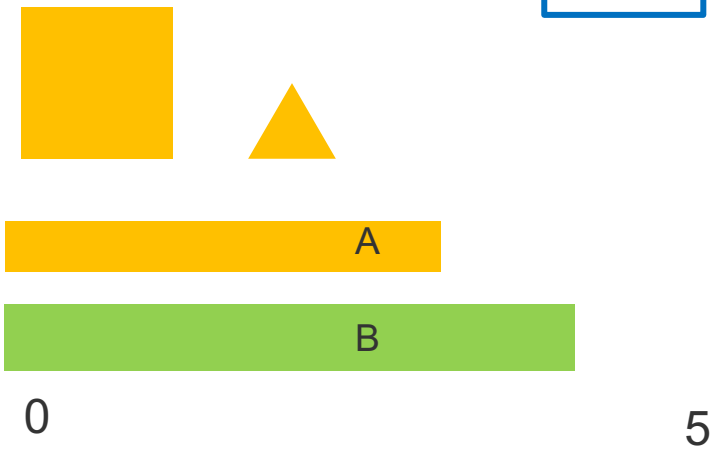
2. Association:
   - Are A and B similar in some way?

3. *Order*
   - *Is A before B?*

Visual decoding tasks

3. *Order*
   - *Is A before B?*

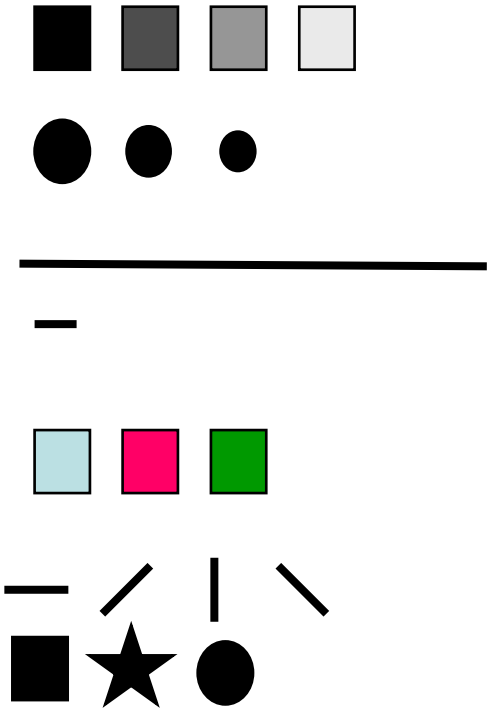
4. Quantification: value
   - How much is A?
   - How much smaller is A than B?



5. Capacity [Carpendale]
   - The number of distinctions possible using the variable
   - How many different things can we represent?

Perceptual semantics

- Density (Greyscale)
- Size / Length / Area
- Position
  - Leftmost -> first,
  - Topmost -> first
- Hue
  - no intrinsic meaning
- Slope / Shape
  - no intrinsic meaning;

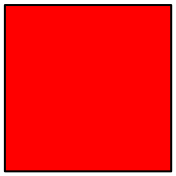


Visual variables: selectivity

Selectivity: Different values are easily seen as different

“Is A different from B?”

preattentive principles



Visual variables: Associativity

- **Associativity:** Similar values can easily be grouped together

  "Is A similar to B?"

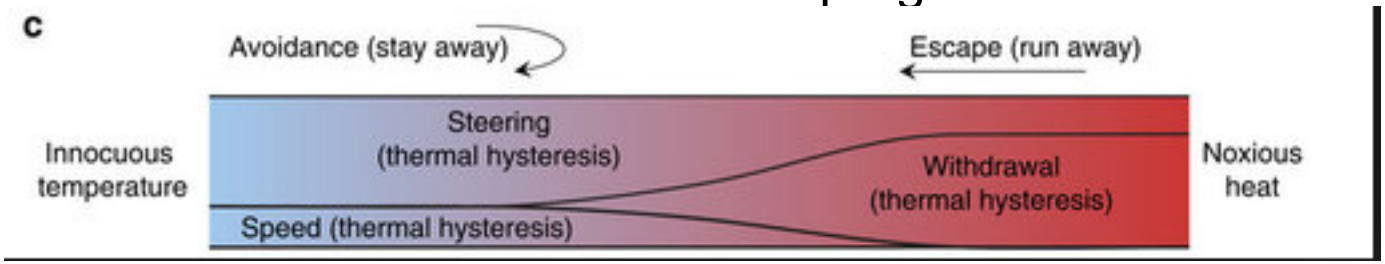
Position > {size, brightness} > {color, orientation } > texture > shape

Visual variables: Order

Order: Different values are perceived as ordered

"Is A more/greater/bigger than B?"

- Size and brightness are ordered
- Orientation, shape, texture are not ordered
- Hue is "not really" ordered
  - Some visual culture of progression



Visual variables: quantity

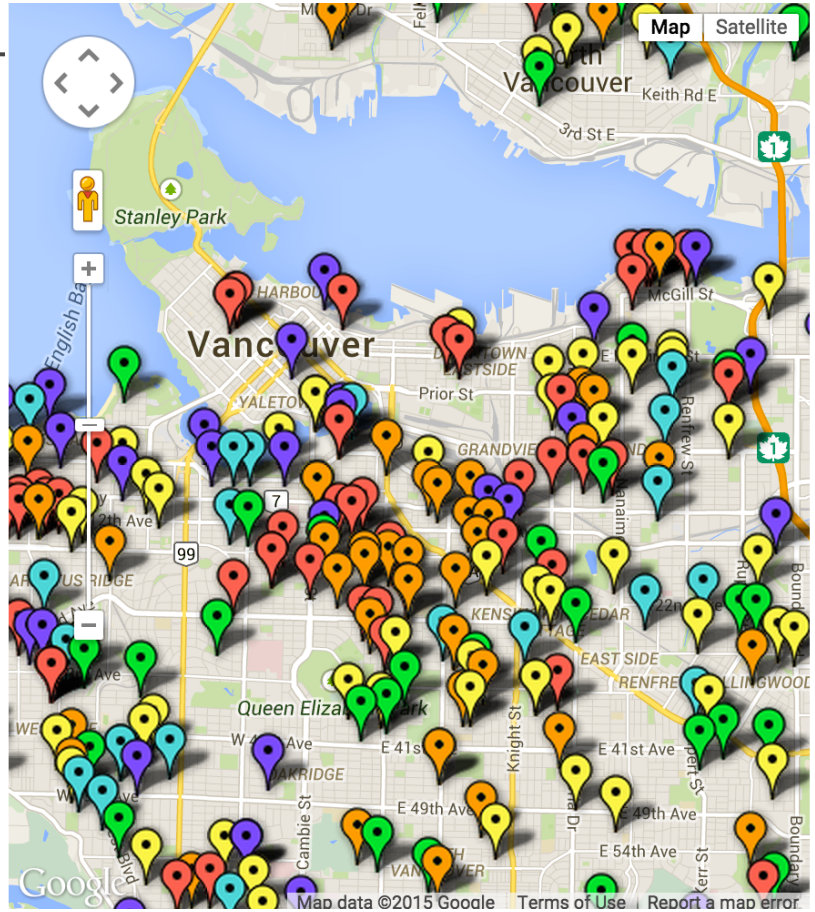
Quantity: A number can be deduced from differences

- "How much is the difference between A and B?"
  - **Position** is quantitative, **size** is somewhat quantitative
  - Other channels are **not** quantitative

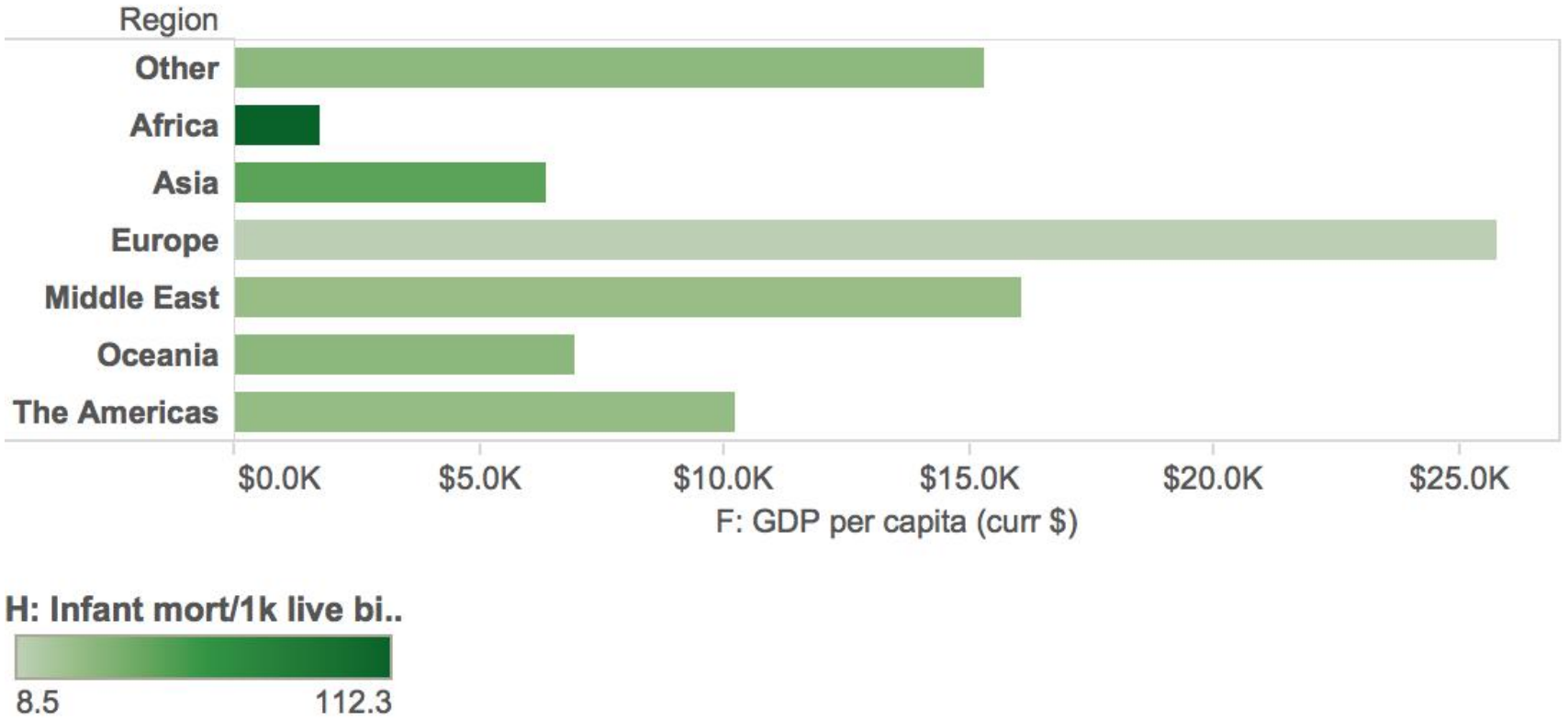
Example: Quantitative values and colour

The 2014 Hallowe'en trick-or-treat counter

HUE is bad for magnitude !!



Relative judgments can be supported by large differences in colour **value**

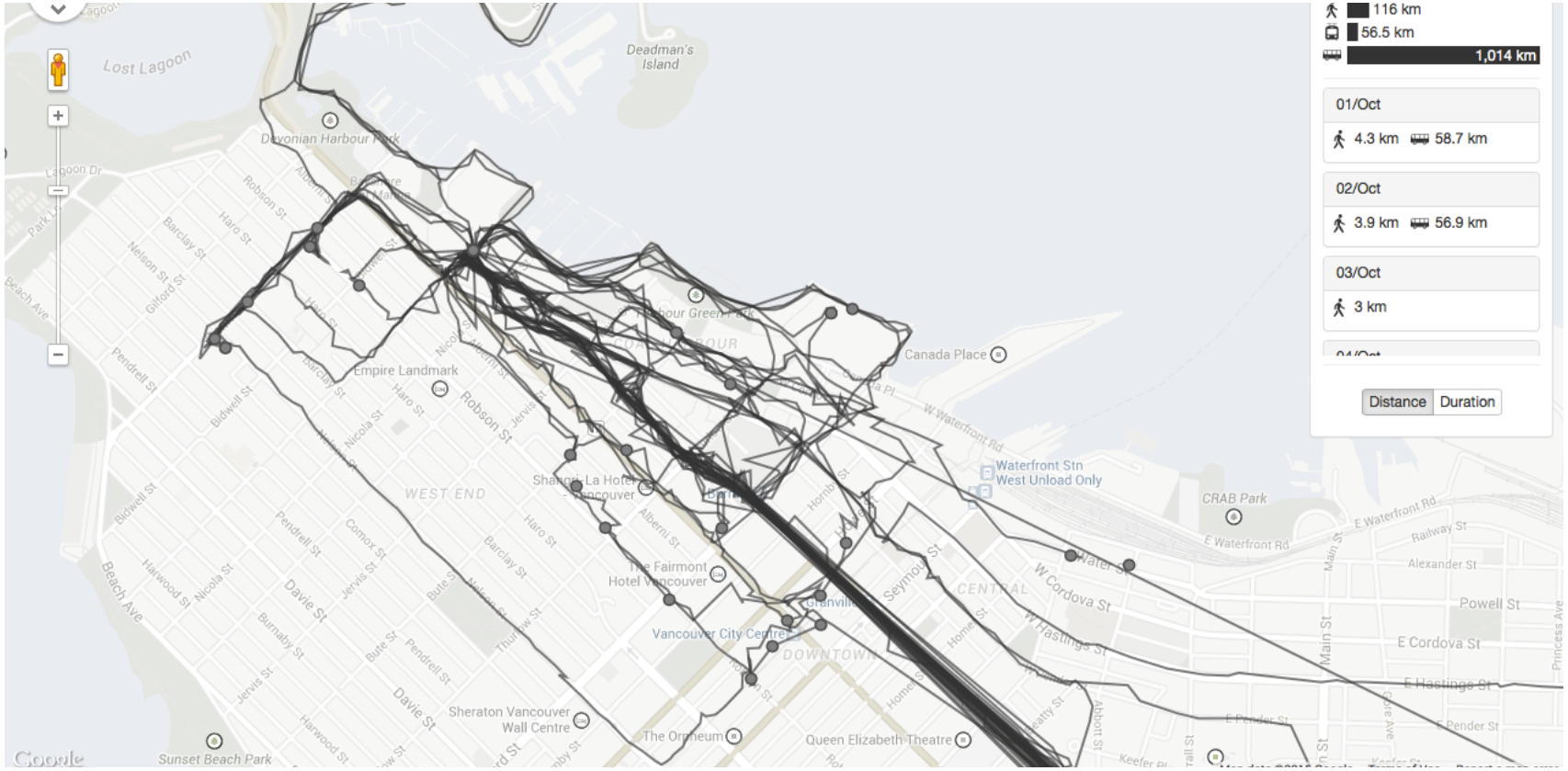


Capacity (discriminability)

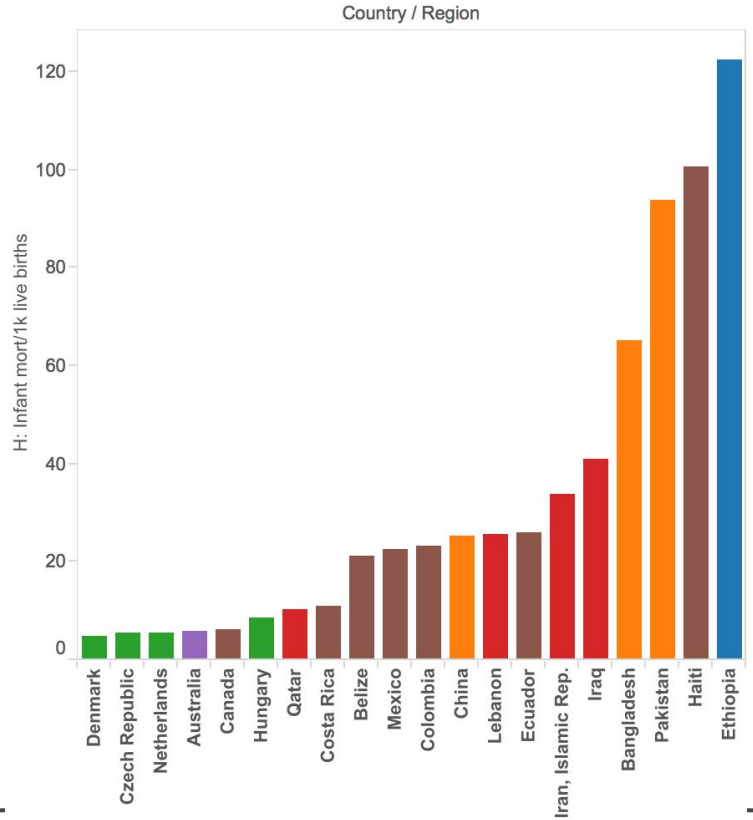
•The number of distinctions (“bins”) possible using the variable

- “How many different things can we represent with this variable?”
  - Shape, Texture: infinite, but …
  - Brightness, hue: 7 (Association) ~ 30 (Distinction)
  - Size: 5 (Association) -20 (Distinction)
  - Orientation: 4

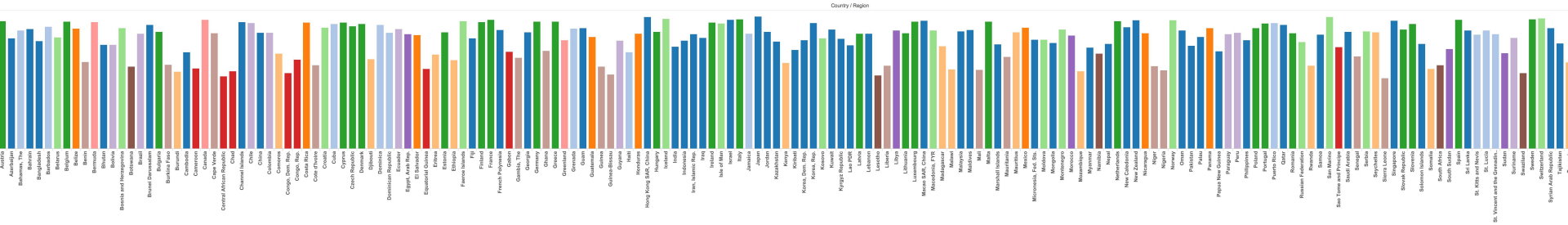
Quantity: line mark and width channel



Hue capacity for identity



Hue capacity for identity



Fundamental principles

Expressiveness:

- the visual encoding should express all of, and **only**, the information in the chosen dataset attributes

Consistency

- The properties of the image should match the properties of the data

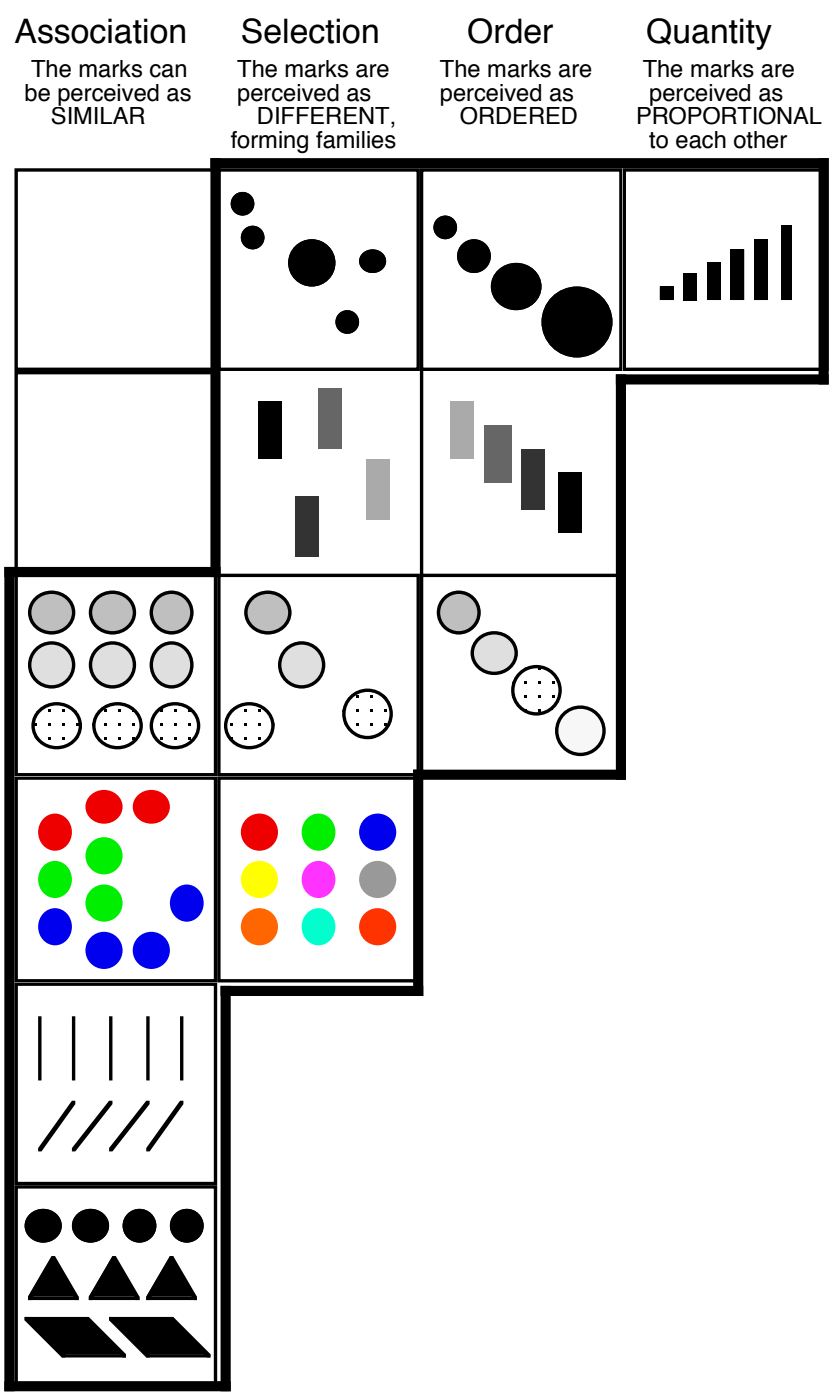
Effectiveness:

- the **importance** of the attribute should match the **salience** of the channel
- Emphasize the critical. Most important gets most perceptually powerful channel.

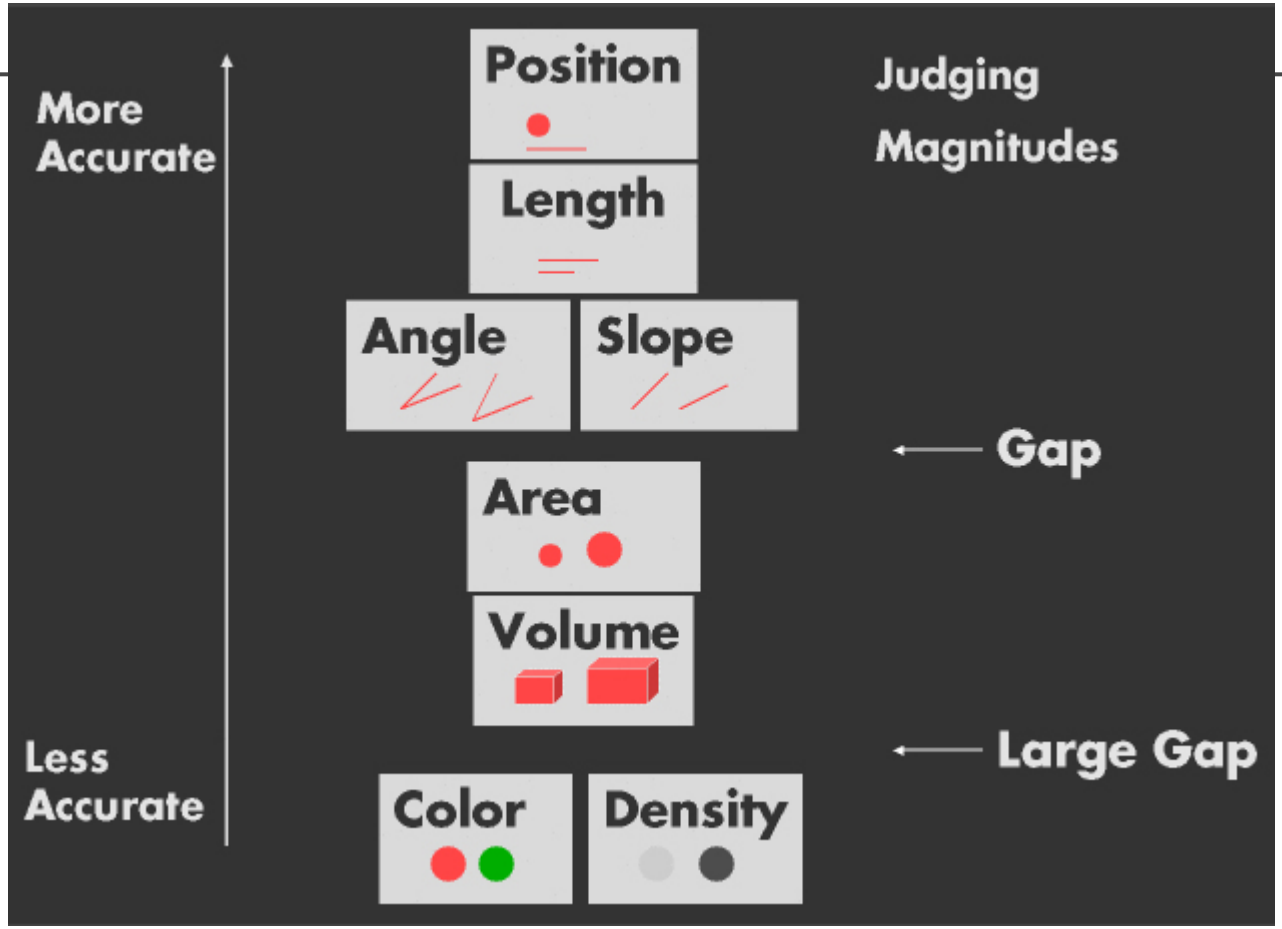
Hierarchy

Bertin's principles:

How encoding methods support common tasks



Visual channels are not created equal



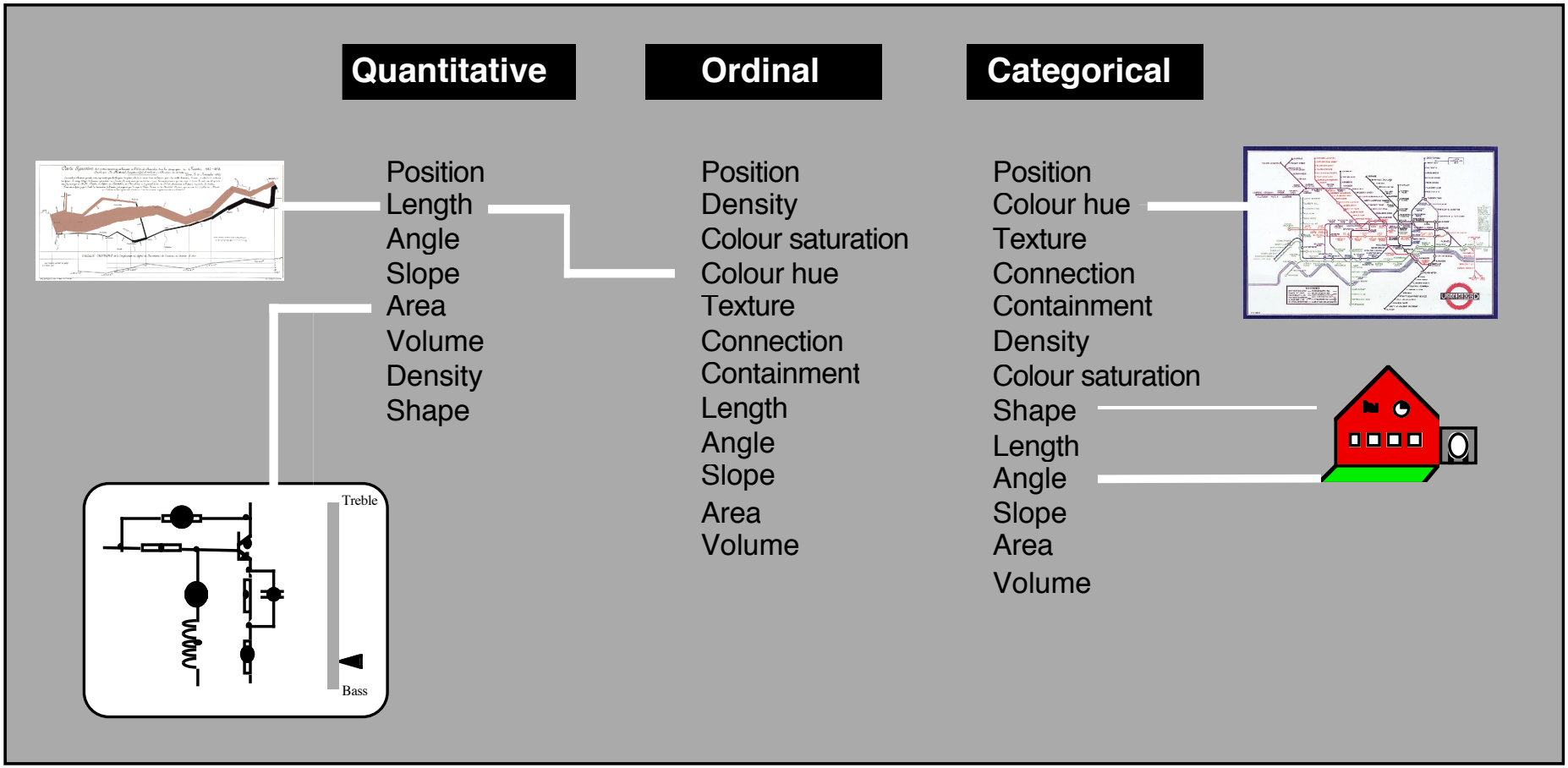
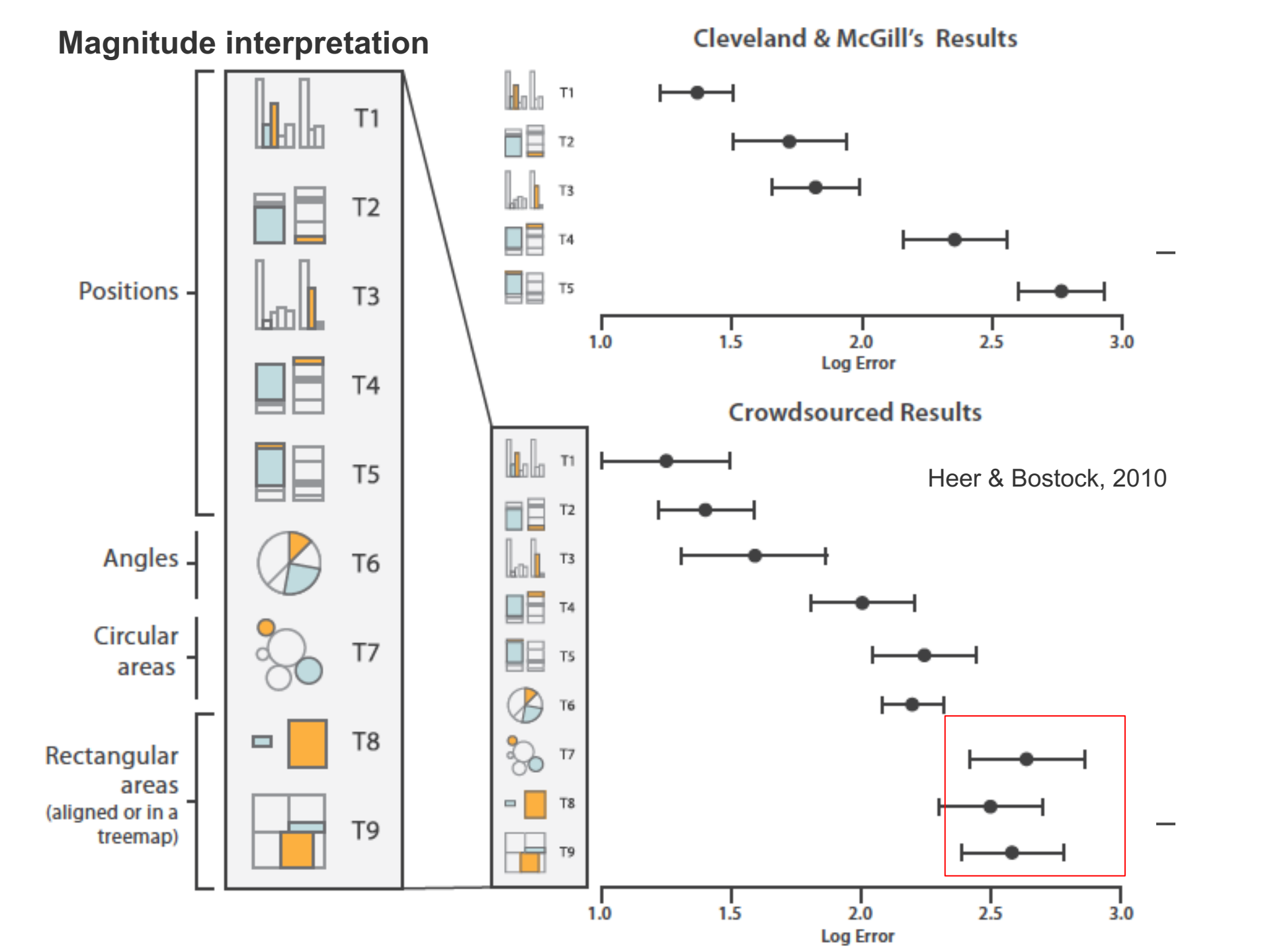


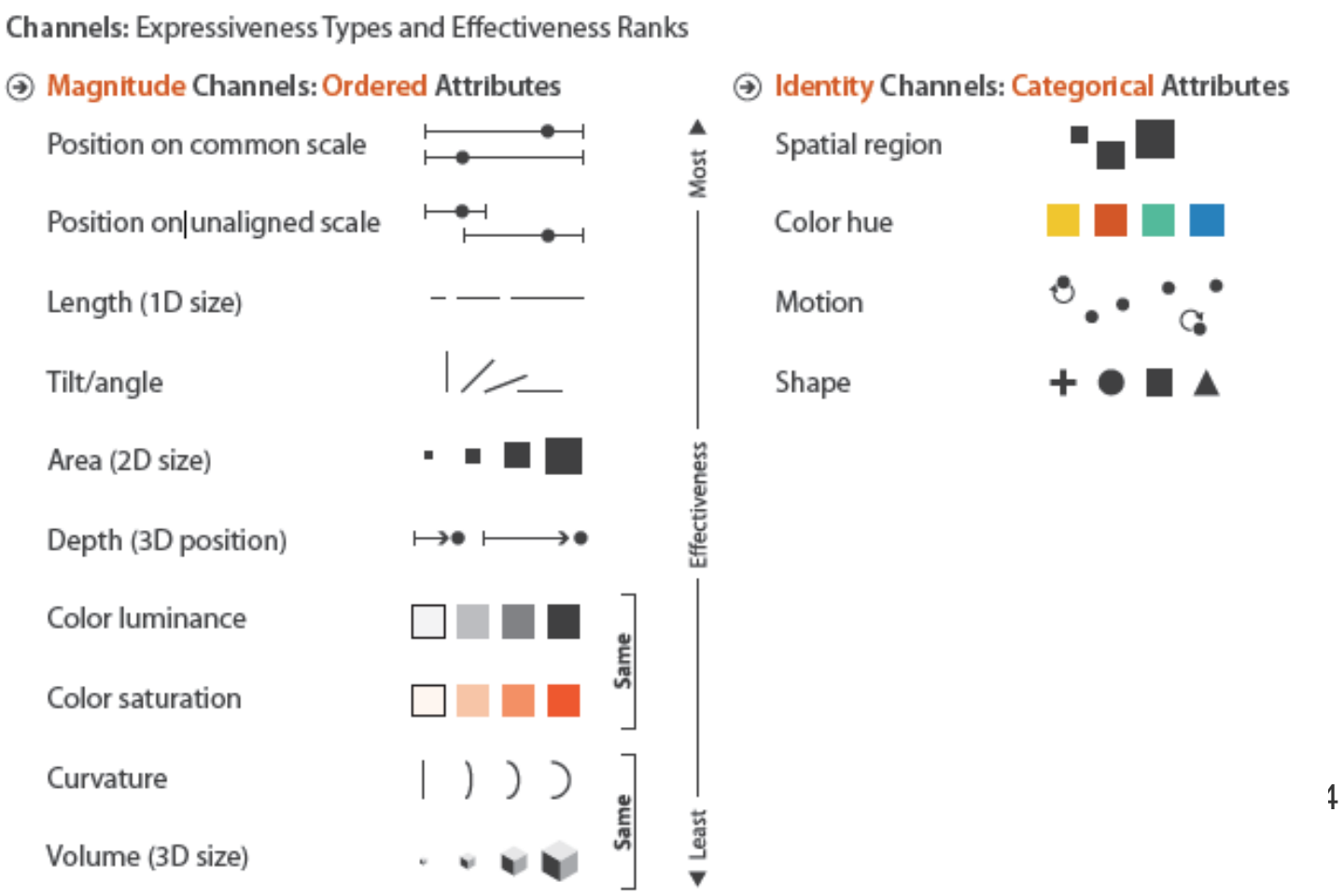
Figure 3.45 Mackinlay's guidance for the encoding of quantitative, ordinal and categorical data

Magnitude interpretation



Heer & Bostock, 2010

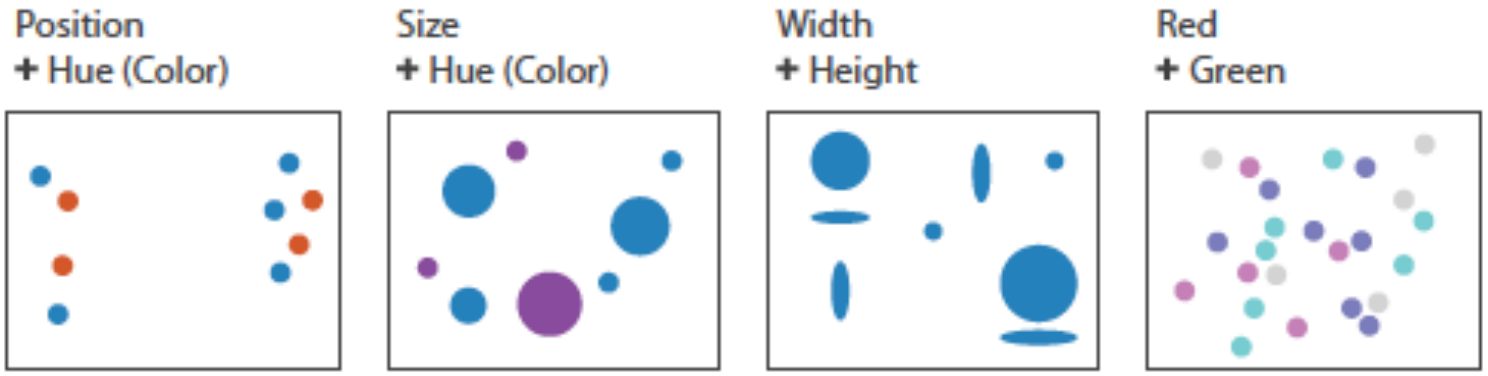
Channels: Expressiveness and Effectiveness



Combining visual channels

- Channels are not all independent from each other
- Our decoding capacity is limited
  - We cannot distinguish some channels
  - Emergent effect

Seperable vs integral [Munzner 2015]



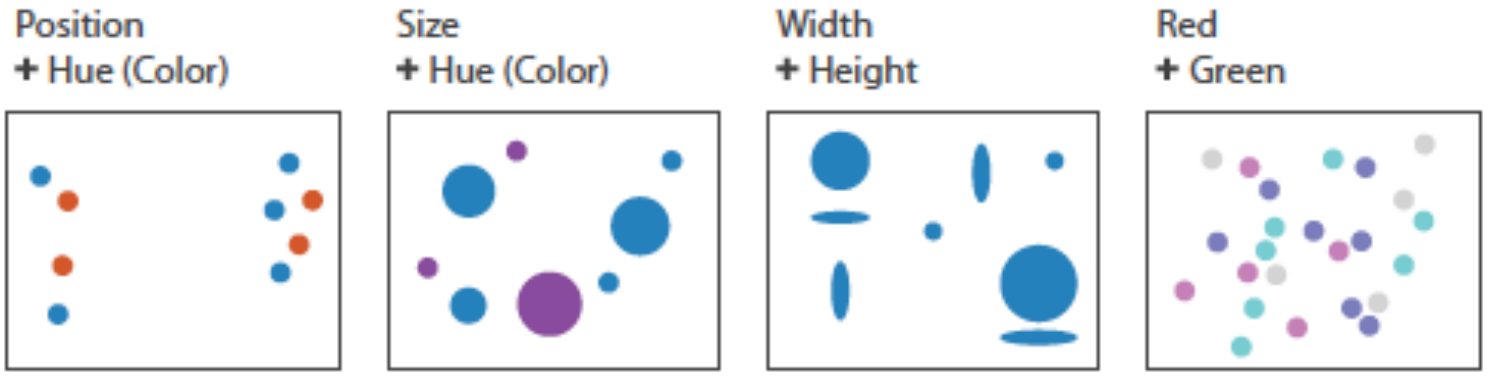
Fully separable

Mostly separable
Small items are hard to see

Some Interference:
Size is **integral** percept

Full Interference:
Hue is **integral** percept

Seperable vs integral [Munzner 2015]



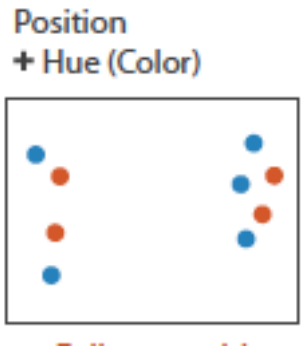
Two distinct groups

Three groups

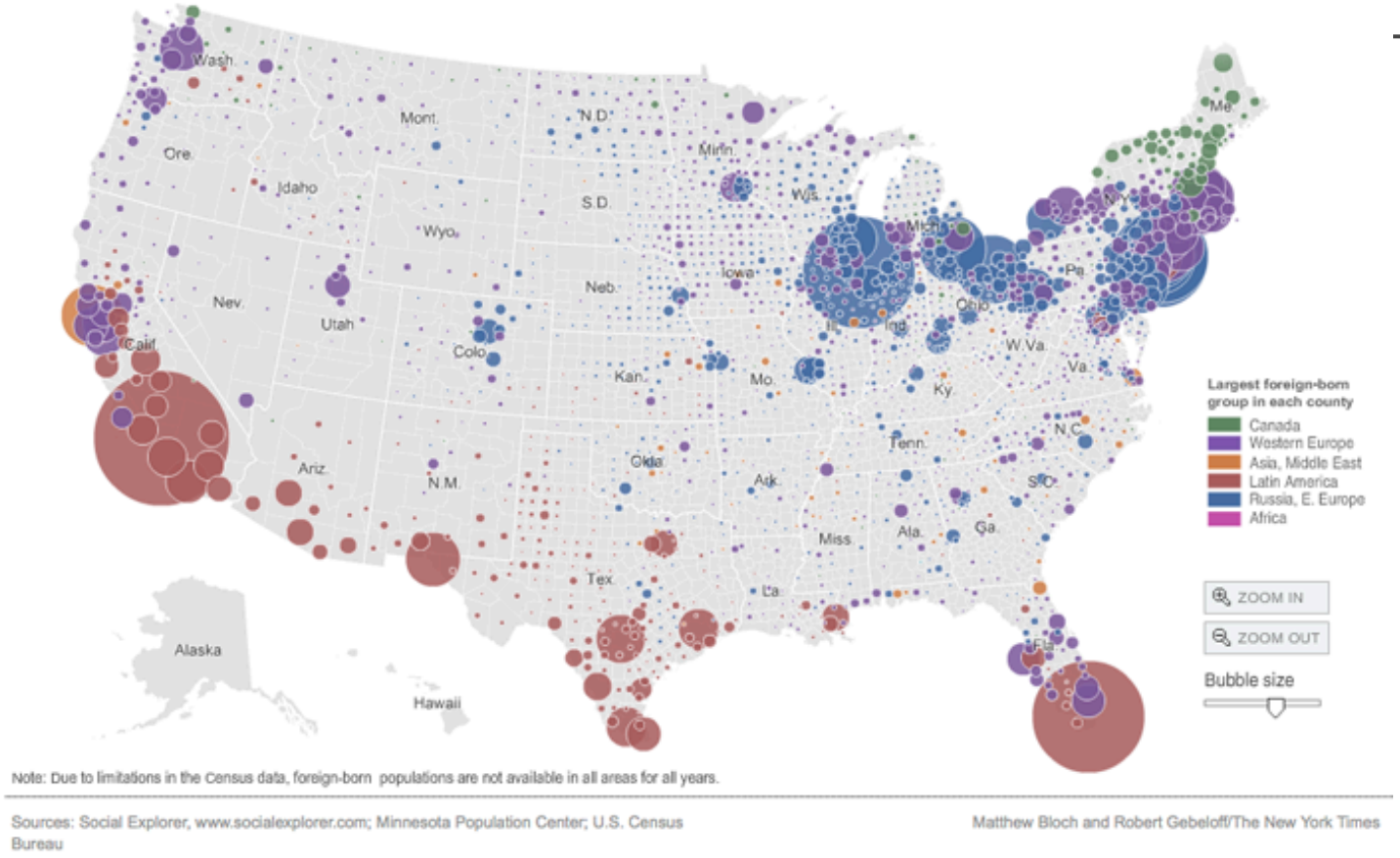
Four groups on hue similarity

Guidelines

- Use separable channels if you want the user to be able to attend to and decode attributes separately
  - Size, position, colour



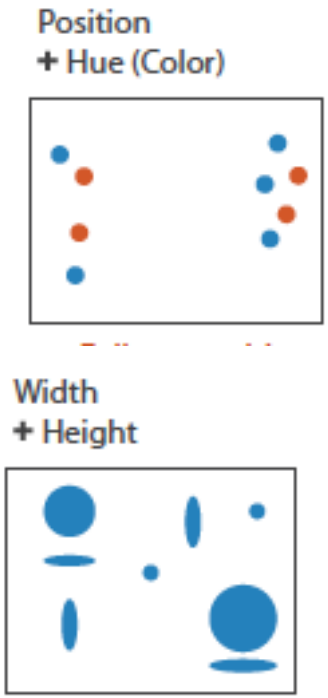
Hue-Size are separable



Guidelines

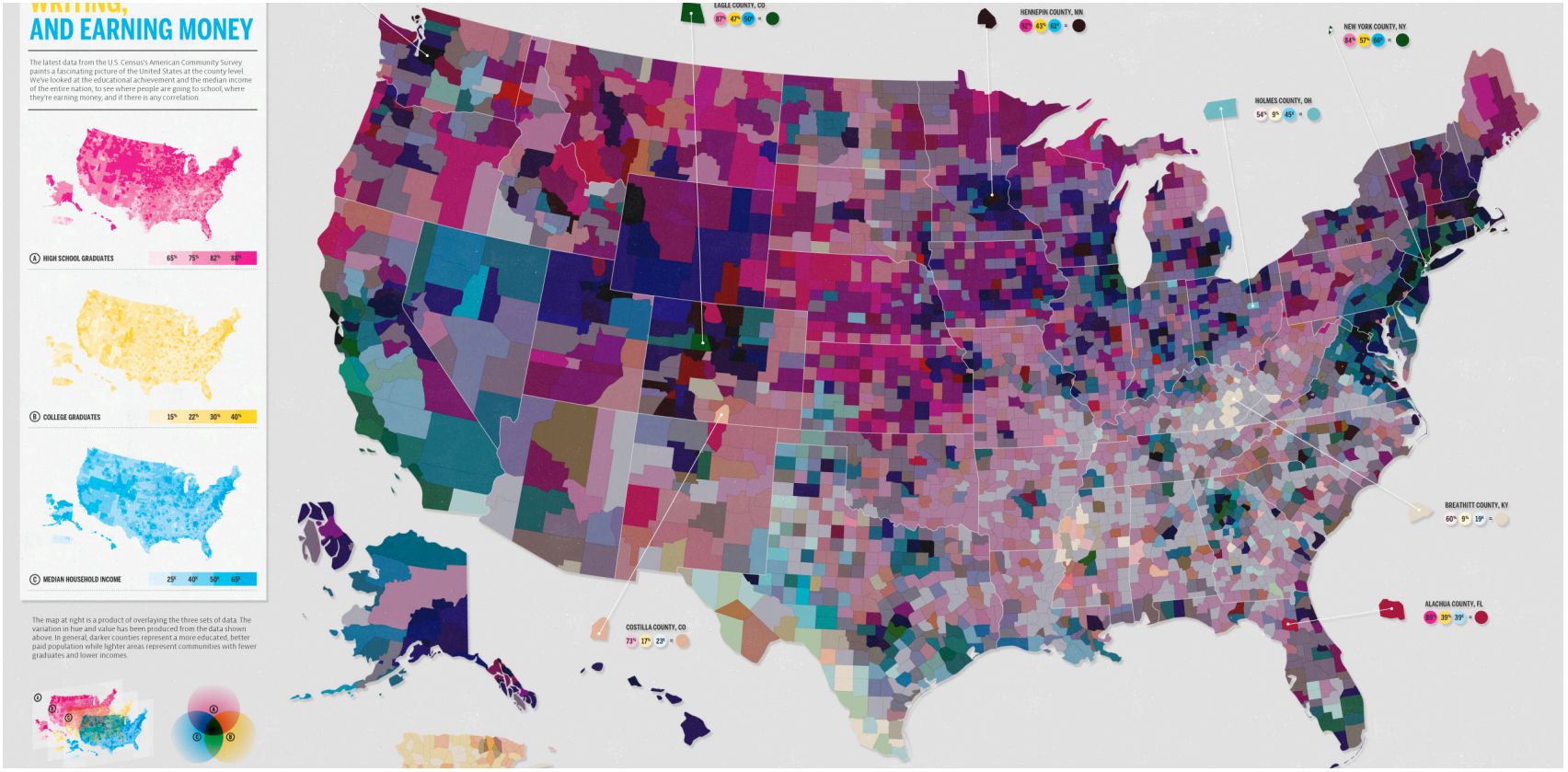
- Use separable channels if you want the user to be able to attend to and decode attributes separately.
  - Size, position, colour

- Use integral channels if you want to show a single data attribute with distinct categories or levels
  - We see three groups distinguished by the shape and size defined by x-size, y-size



And … ?

http://www.visualisingdata.com/index.php/2011/01/discussing-are-the-richest-americans-also-the-best-educated/



Few's Table:

Attribute	*Quantitative*	*Qualitative*
Line length	●	
2-D position	●	
Orientation		●
Line width		●
Size		●
Shape		●
Curvature		●
Added marks		●
Enclosure		●
Hue		●
Intensity		●

What about Colour?

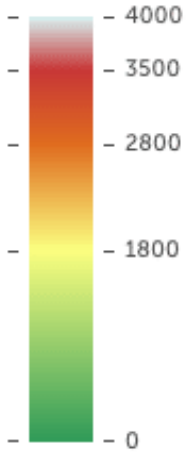
Colour in Information Display

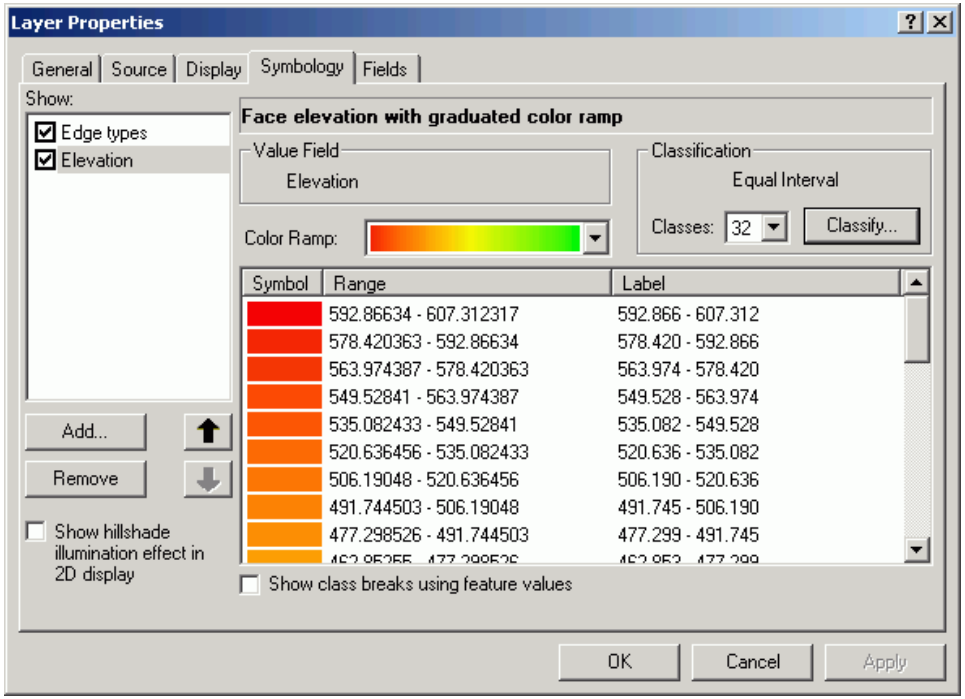
Lyn Bartram

SFU

SCHOOL OF INTERACTIVE ARTS + TECHNOLOGY

What makes color effective?

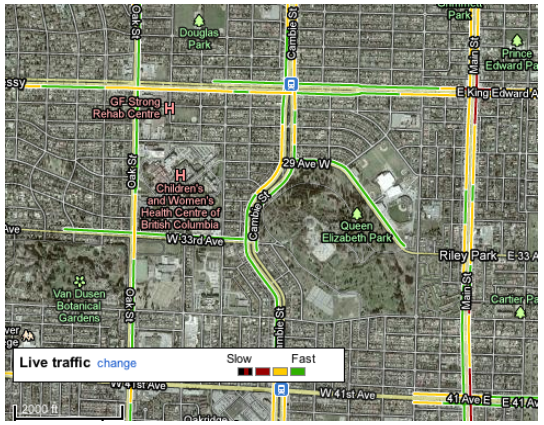


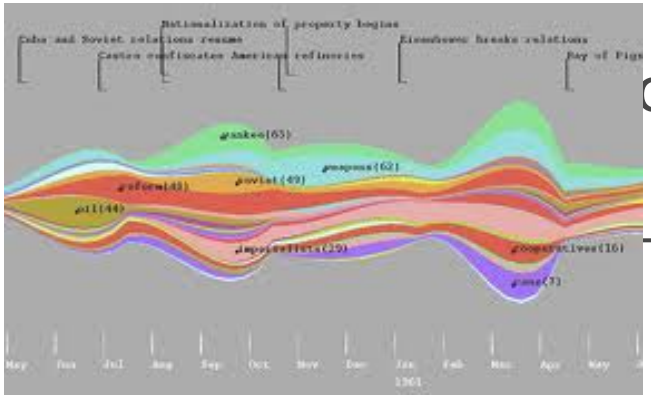


"Good ideas executed with superb craft"
—*E.R. Tufte*

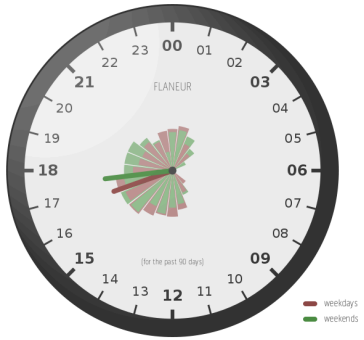
A quick revisit of data characteristics

- What kind of task are you hoping to support?
- What kind of data do you have?
- No single approach



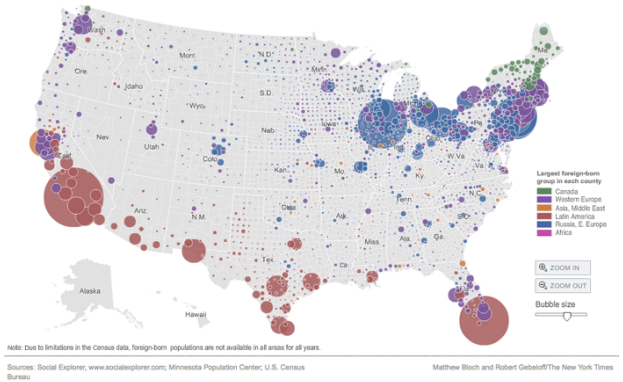


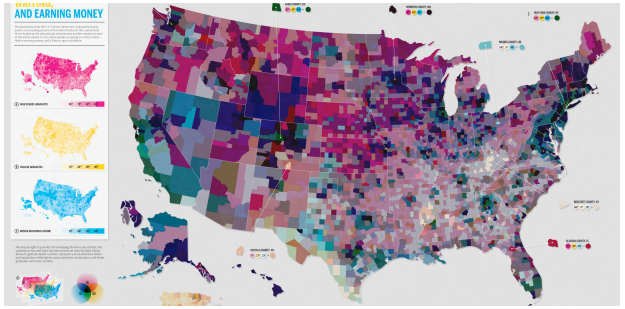
ols



Color Design Principles

- Control value (lightness)
  - Ensure legibility
  - Avoid unwanted emphasis
- Use a limited hue palette
  - Control color “pop out”
  - Define color grouping
  - Avoid clutter (too many)
- Use neutral backgrounds
  - Control impact of color
  - Minimize simultaneous contrast



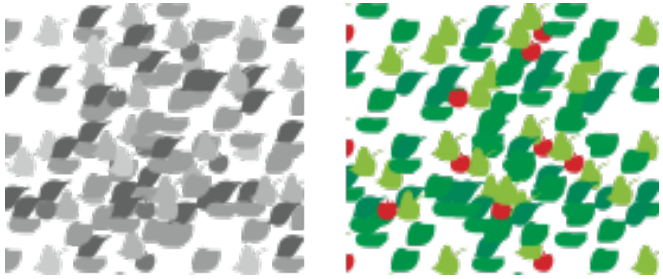


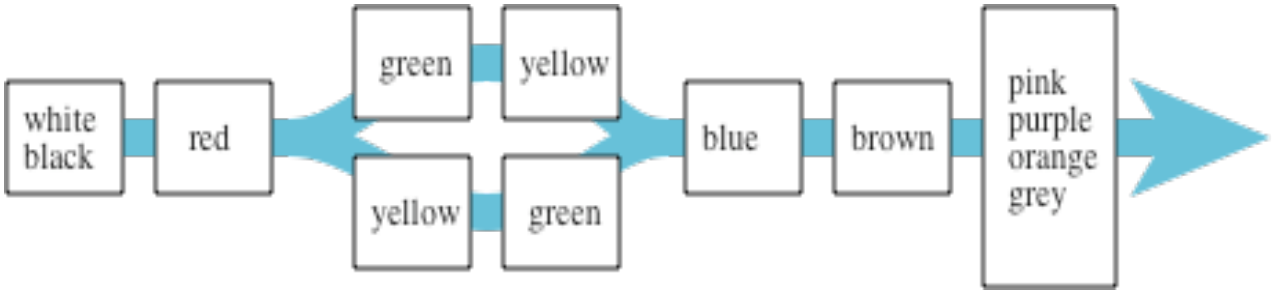
Fundamental Uses

- To label (colour as **noun**)
- To measure (colour as **quantity/value**)
- To represent (colour as **representation**)
  - to imitate reality
- To enliven or decorate (colour as **beauty**)

Hue is great for nominal data (“labels”)

- Pop out
-
- Colour helps us determine type
- Only about six categories

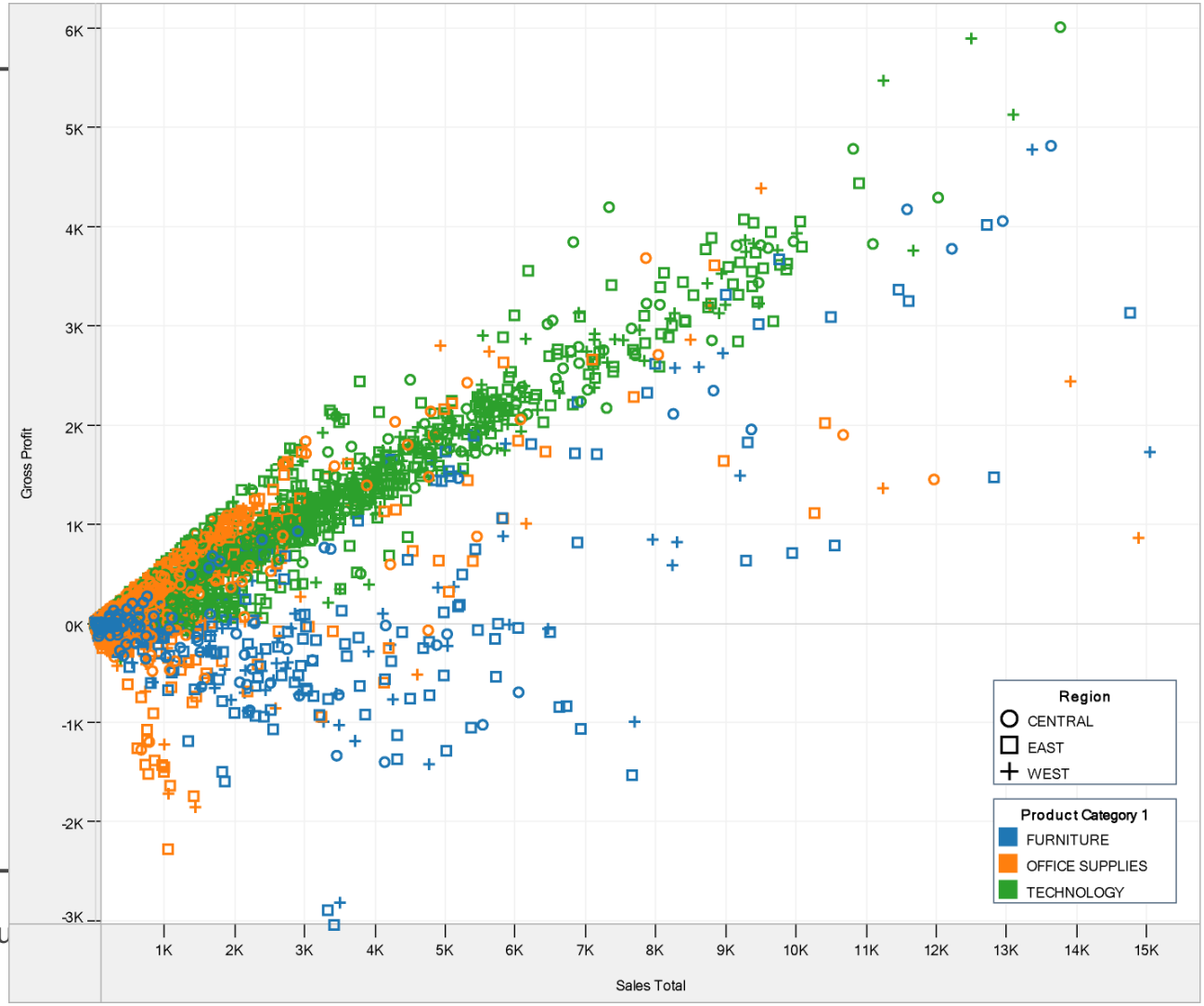




Visual Thinking, Colin Ware

Product Categories

Created by Tableau - Visual Analysis for Databases™



Grouping, Highlighting

	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
red	25.37	13.70	0.05	26.27	14.13	0.04	18.41	10.16	0.05	17.43	9.30	0.00
green	22.14	51.24	0.35	20.68	49.17	0.44	21.11	46.00	0.20	16.36	37.95	0.12
blue	13.17	3.71	74.89	15.38	5.20	86.83	11.55	3.37	65.53	9.96	3.44	56.14
gray	63.46	73.30	78.05	64.66	71.99	90.08	52.96	62.49	67.99	45.54	53.65	58.14
black	0.66	0.70	0.77	0.63	0.66	1.09	0.47	0.58	0.70	0.44	0.54	0.71

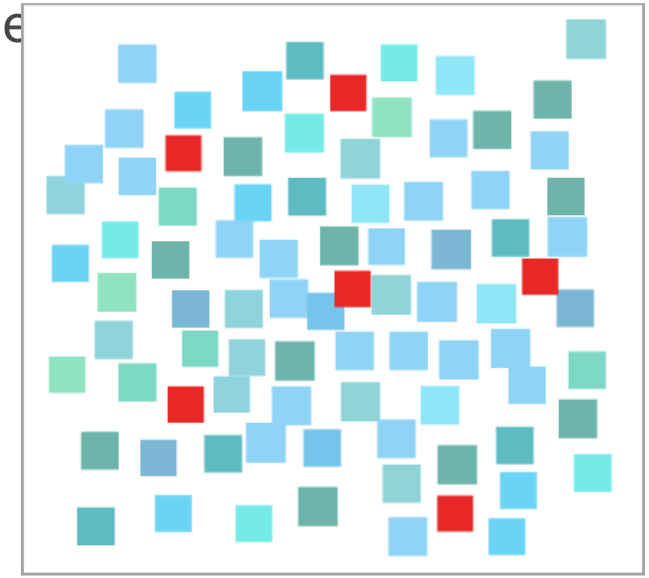
	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
red	25.37	13.70	0.05	26.27	14.13	0.04	18.41	10.16	0.05	17.43	9.30	0.00
green	22.14	51.24	0.35	20.68	49.17	0.44	21.11	46.00	0.20	16.36	37.95	0.12
blue	13.17	3.71	74.89	15.38	5.20	86.83	11.55	3.37	65.53	9.96	3.44	56.14
gray	63.46	73.30	78.05	64.66	71.99	90.08	52.96	62.49	67.99	45.54	53.65	58.14
black	0.66	0.70	0.77	0.63	0.66	1.09	0.47	0.58	0.70	0.44	0.54	0.71

Considerations for Colour Labels

- How critical? Unique specification or is it a “hint”?
  - Quick response, or time for inspection?
  - Is there a legend, or need it be memorized?
- Context
  - Are there established **semantics**?
  - Grouping or ordering relationships?
  - Surrounding shapes and colors?
- Shape and structural issues
  - How big are the objects?
  - How many objects, and could they overlap?
  - Need they be **readable**, or only visible (discernible)?

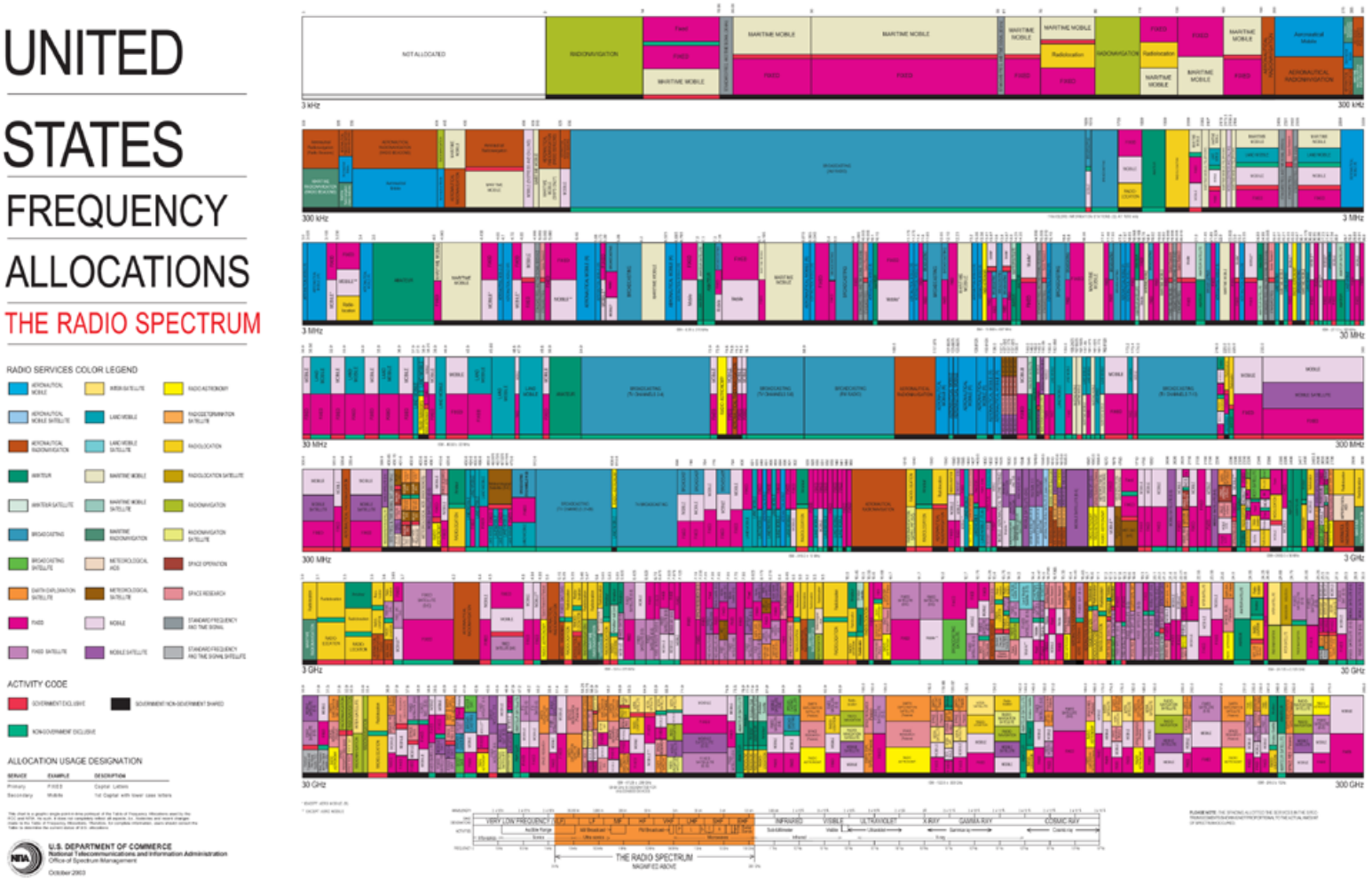
Pop-out vs. Distinguishable

- Pop-out
  - Typically, 5-6 distinct values simultane
  - Up to 9 under controlled conditions



- Distinguishable
  - 20 easily for reasonable sized stimuli
  - More if in a controlled context
  - Usually need a legend

Radio Spectrum Map (33 colors)



http://www.cybergeography.org/atlas/us_spectrum_map.pdf

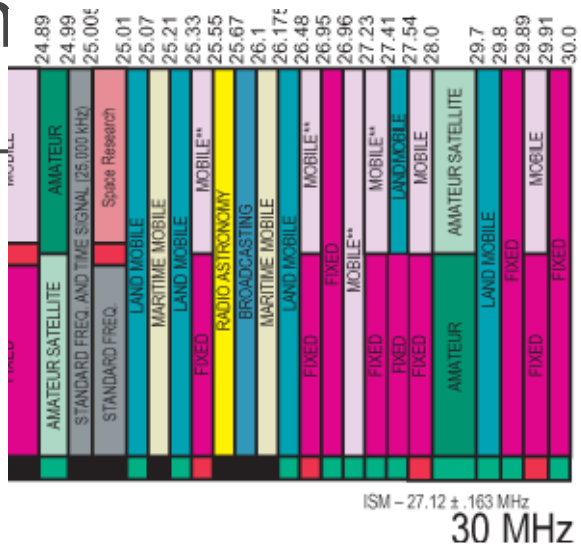
Distinguishable on Inspection

RADIO SERVICES COLOR LEGEND

- AERONAUTICAL MOBILE
- AERONAUTICAL MOBILE SATELLITE
- AERONAUTICAL RADIONAVIGATION
- AMATEUR
- AMATEUR SATELLITE
- BROADCASTING
- BROADCASTING SATELLITE
- EARTH EXPLORATION SATELLITE
- FIXED
- FIXED SATELLITE
- INTER-SATELLITE
- LAND MOBILE
- LAND MOBILE SATELLITE
- MARITIME MOBILE
- MARITIME MOBILE SATELLITE
- MARITIME RADIONAVIGATION
- METEOROLOGICAL AIDS
- METEOROLOGICAL SATELLITE
- MOBILE
- MOBILE SATELLITE
- RADIO ASTRONOMY
- RADIODETERMINATION SATELLITE
- RADIOLOCATION
- RADIOLOCATION SATELLITE
- RADIONAVIGATION
- RADIONAVIGATION SATELLITE
- SPACE OPERATION
- SPACE RESEARCH
- STANDARD FREQUENCY AND TIME SIGNAL
- STANDARD FREQUENCY AND TIME SIGNAL SATELLITE

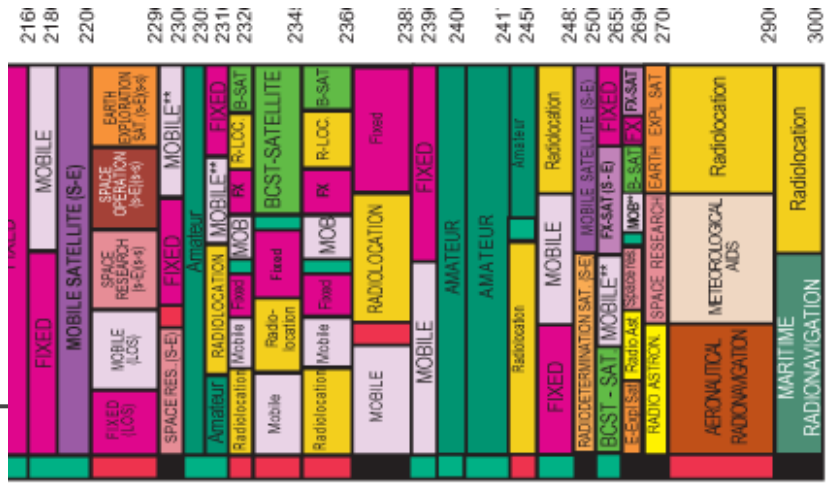
ACTIVITY CODE

- GOVERNMENT EXCLUSIVE
- GOVERNMENT/ NON-GOVERNMENT SHARED
- NON-GOVERNMENT EXCLUSIVE



ISM – 27.12 ± .163 MHz

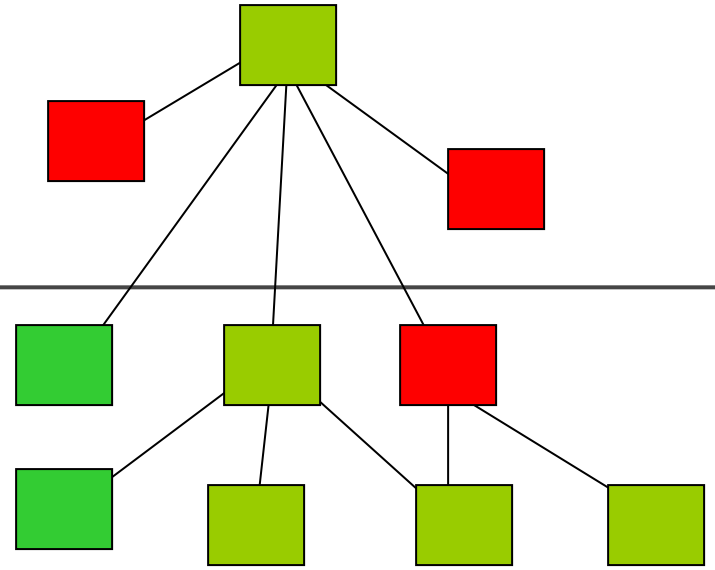
30 MHz



ISM – 2450.0 ± 50 MHz

3 GHz

Cultural semantics

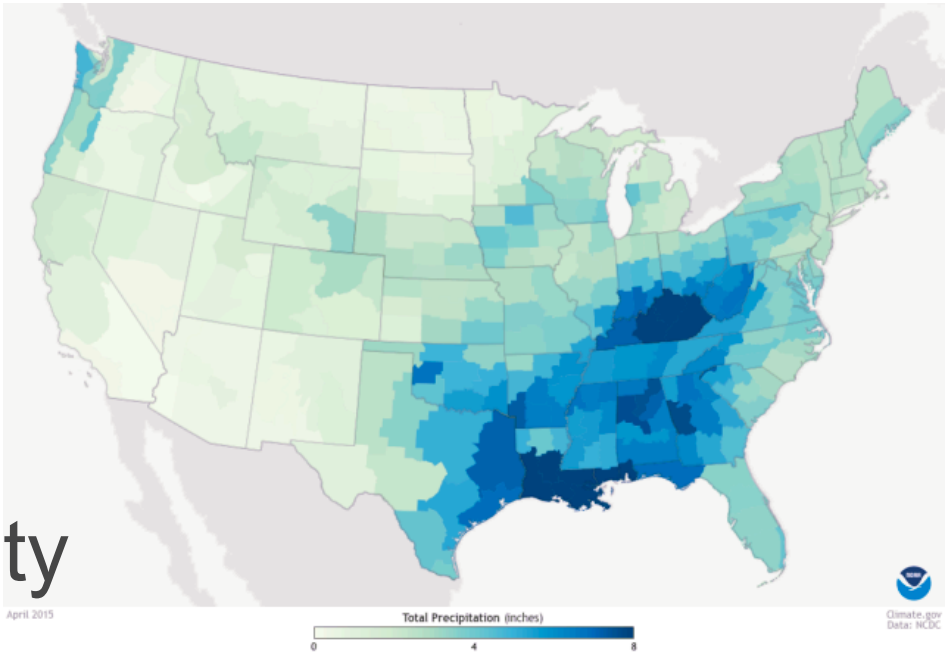


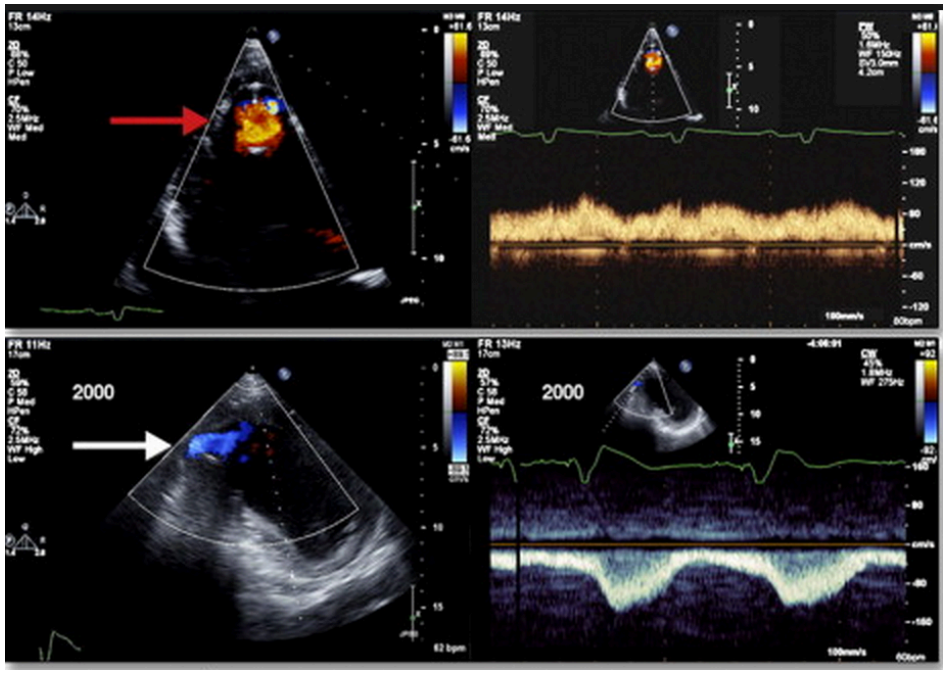
- The fact that certain colours are special is because they are hard wired
- The meaning of those colours is culturally determined

Summary: labeling

- Distinctiveness (perceptual distance)
- Unique hues
  - Don’t choose colours from the same category!
- Contrast with background
  - Luminance
- Compensate for colour deficiency
- Limited number of codes
- Size
  - Objects should not be too small
- Cultural conventions

Colour for quantity





SFU

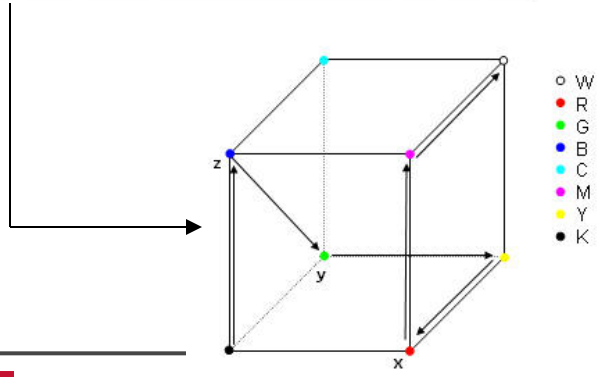
Colour scales

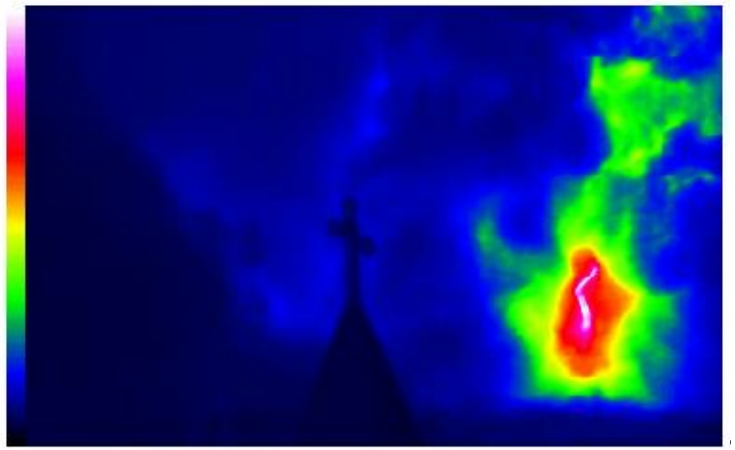
- **Hue scale**
  - Cyclic (learned order)

- **Lightness or saturation scales**
  - Ordered scales
  - Lightness best for high frequency
  - More = darker (or more saturated)
  - Most accurate if *quantized (approximate with discrete values)*

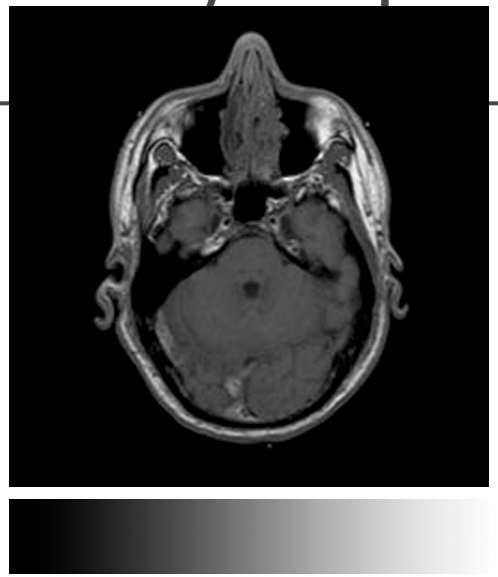
Pseudocoloring

- represent continuously varying values with colour sequence
- Sometimes overlaid on luminosity information
  - Need to use an **isoluminant** color map to avoid distortion

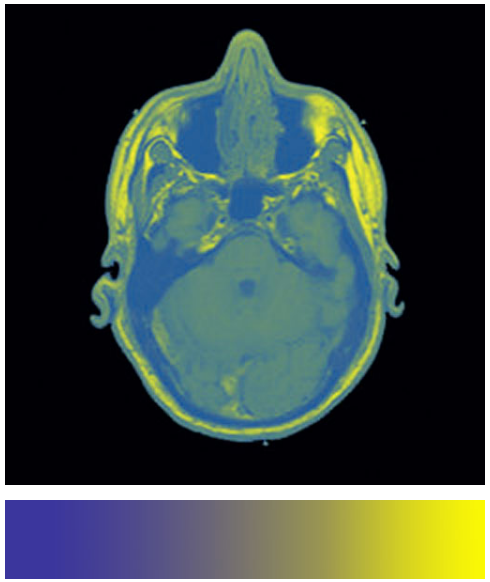




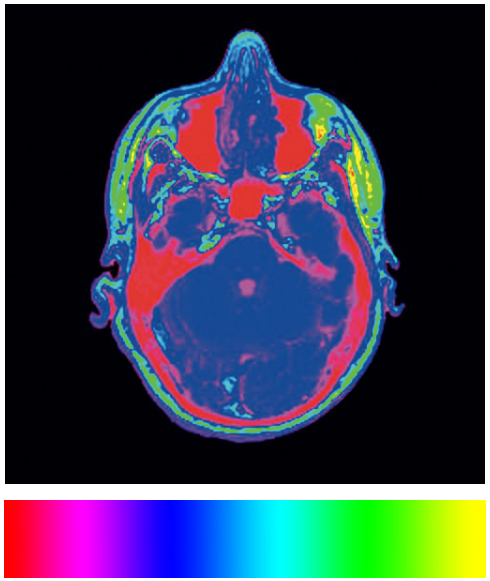
Density Map



Lightness scale

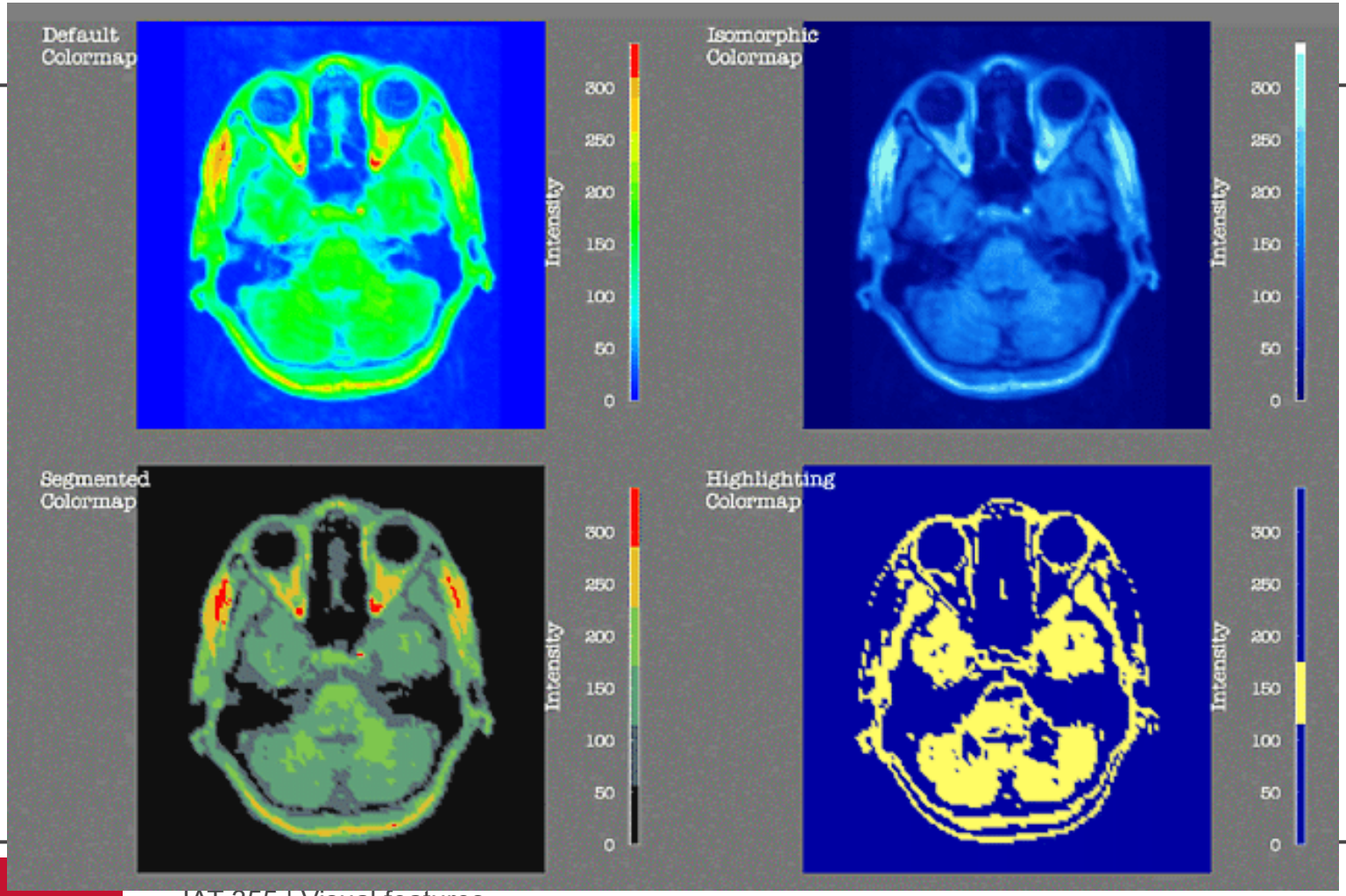


Lightness scale
with hue and
chroma variation



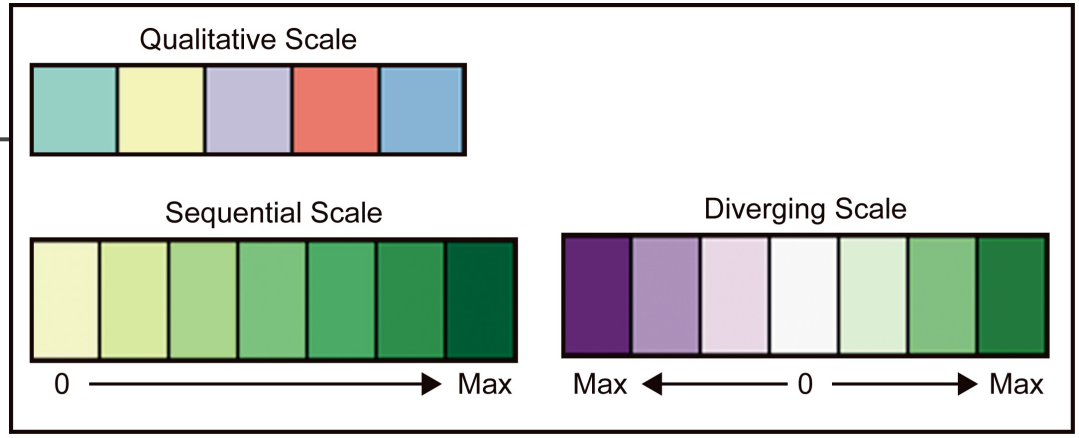
“Rainbow” hue scale
Unordered hue, lightness

Different Scales



Rogowitz & Treinish, “How not to lie with visualization”

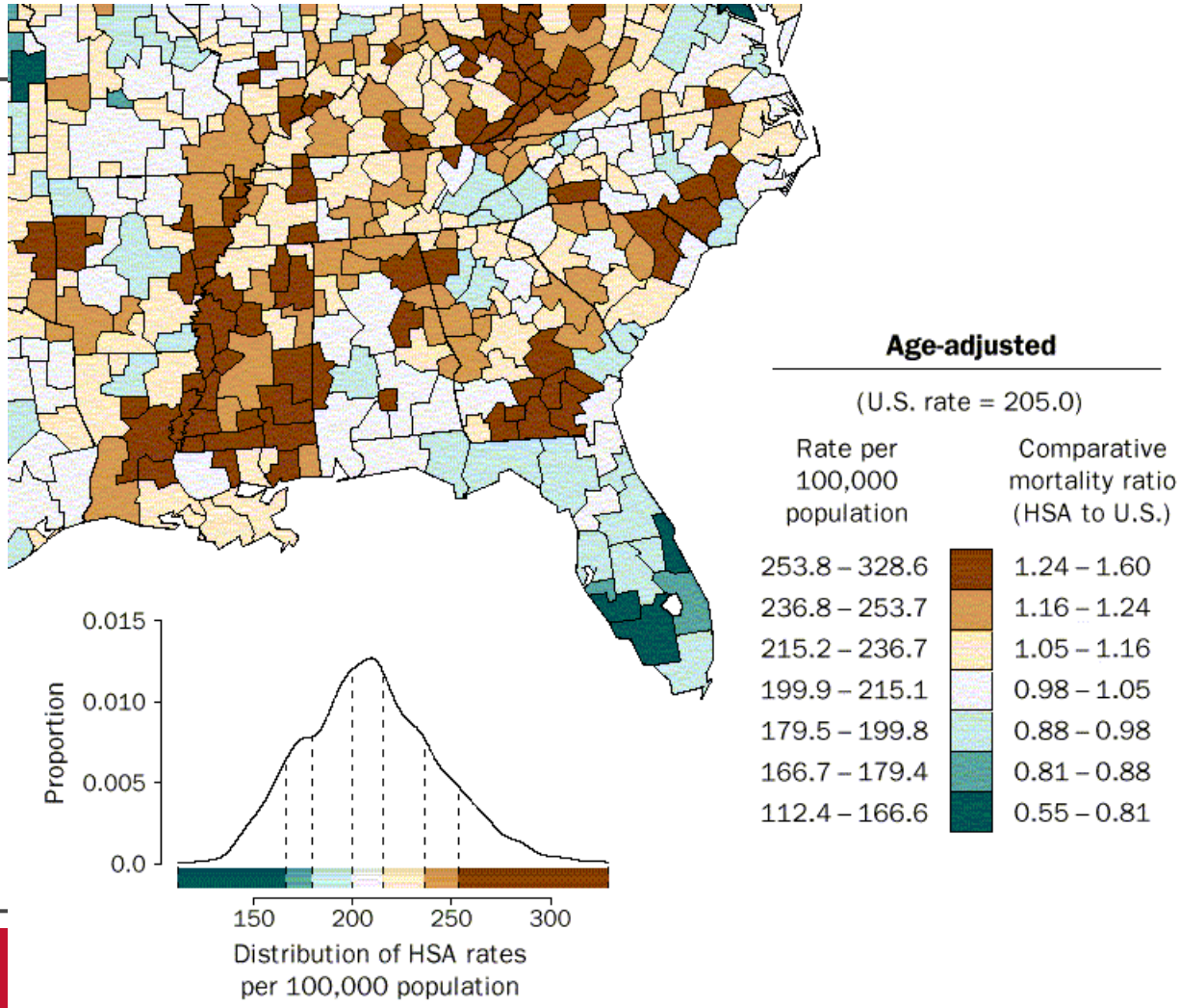
Brewer Scales



- Qualitative scales
  - nominal
  - Distinct hues, similar emphasis
- Sequential scale
  - Vary in lightness and saturation
  - Vary slightly in hue
- Diverging scale
  - Complementary sequential scales
  - Neutral at “zero”
  - Cross-fade through a neutral color

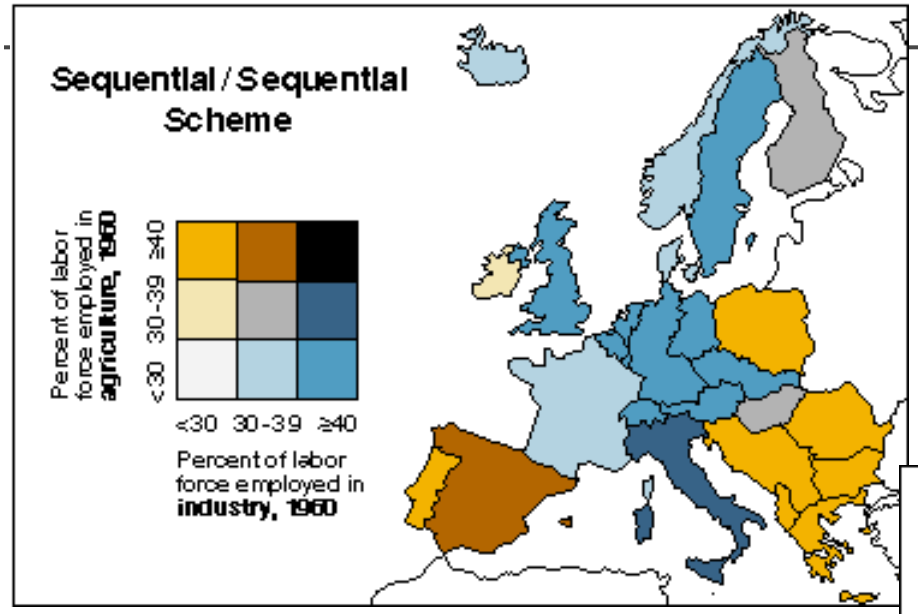
Thematic Maps

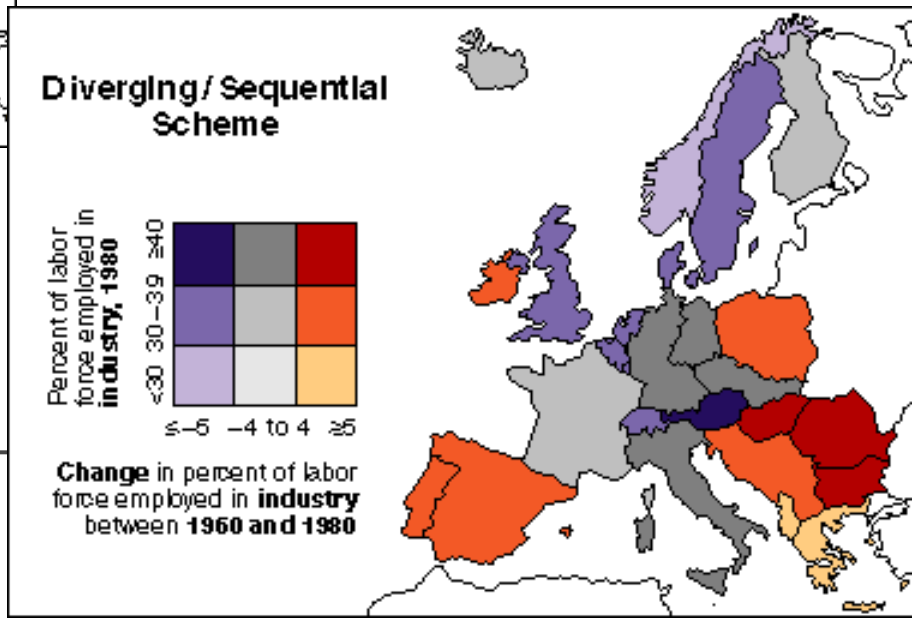
US Census Map



[Mapping Census 2000: The Geography of U.S. Diversity]

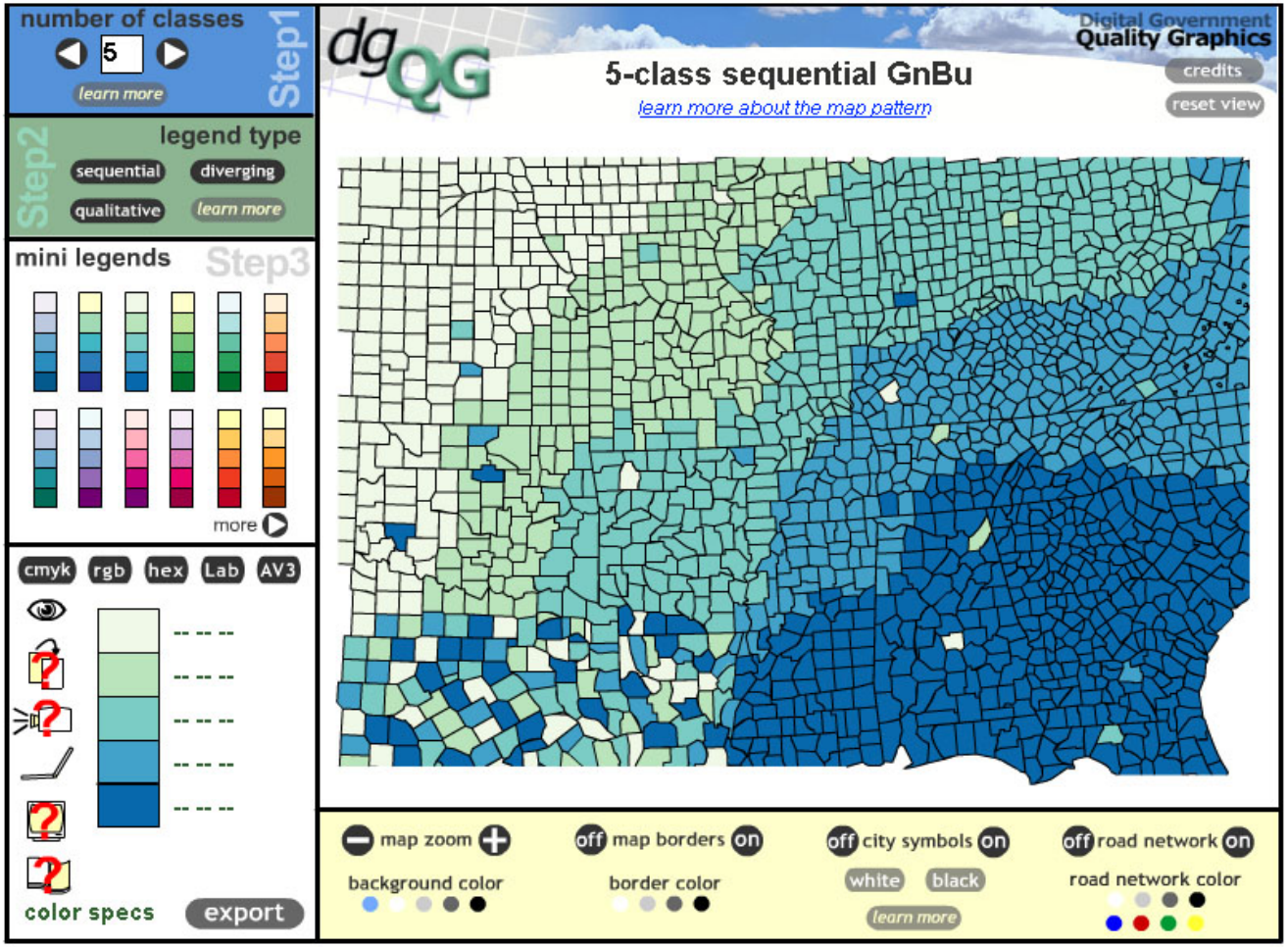
Brewer Examples 2 dimensions





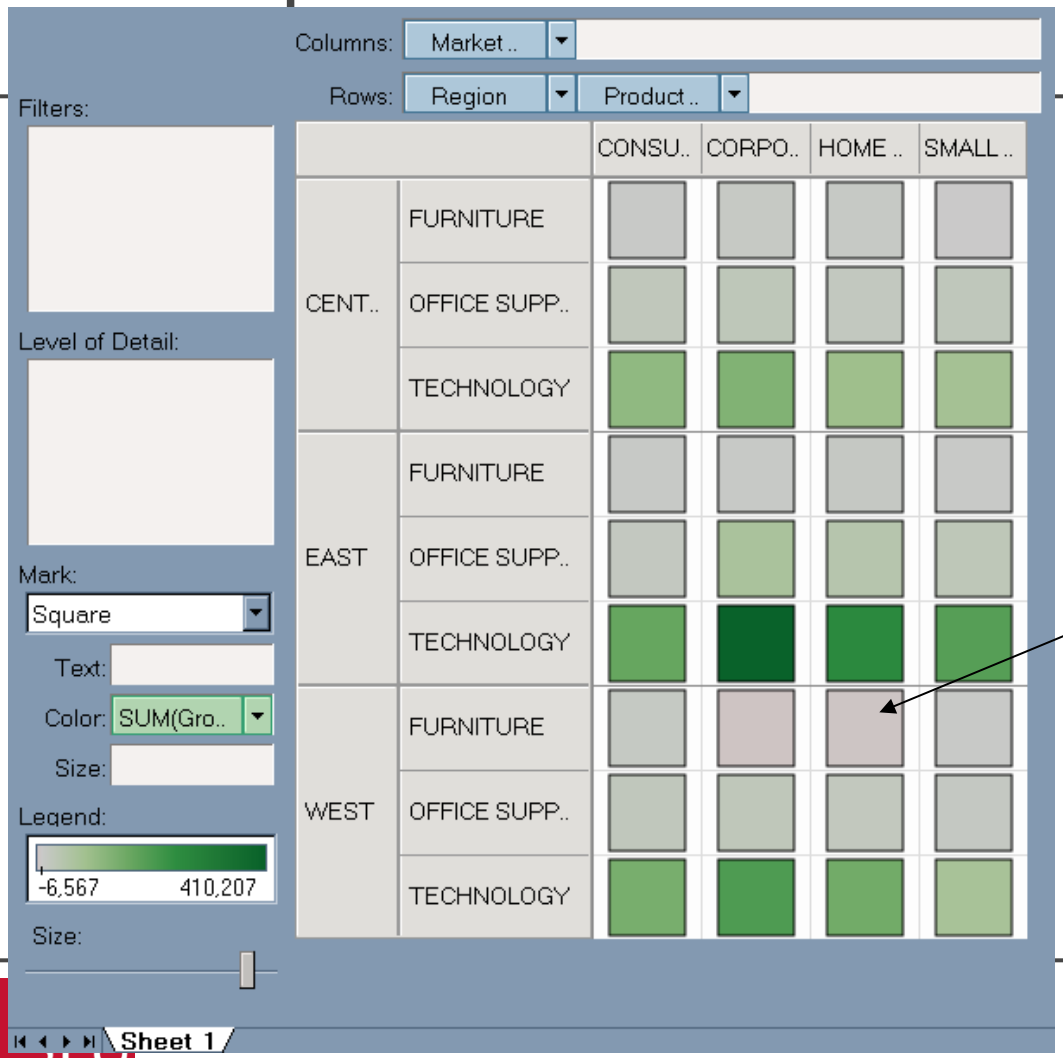
Color Brewer

www.colorbrewer.org



This material is based upon work supported by the National Science Foundation under Grant No. 9983451, 9983459, 9983461

Heat Maps



Skewed Data

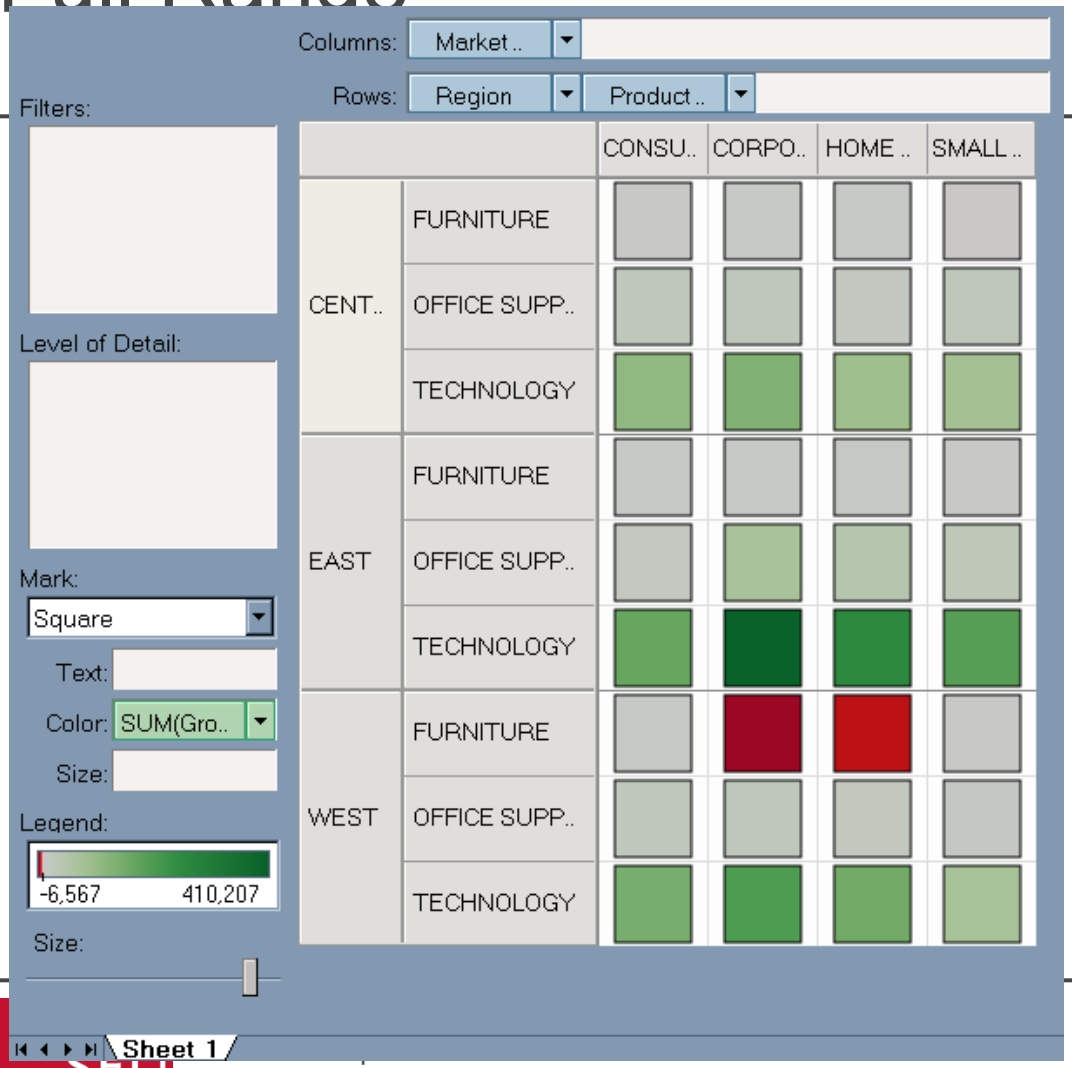
Default Tableau colour scale

Slightly negative

www.tableausoftware.com

Slide courtesy of Maureen Stone, Tableau Software

Full Range



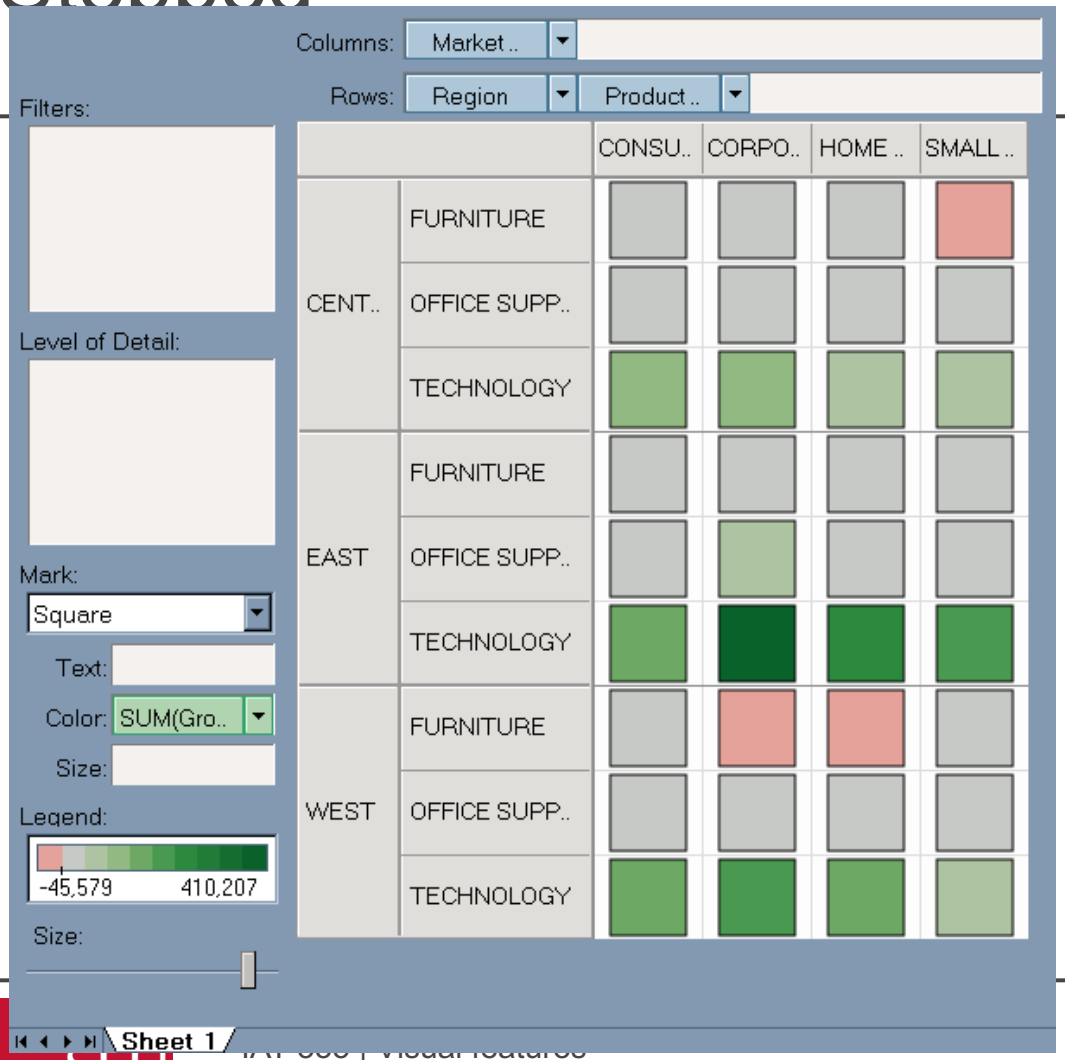
Skewed Data

Divergent Tableau colour scale

www.tableausoftware.com

Slide courtesy of Maureen Stone, Tableau Software

Stepped



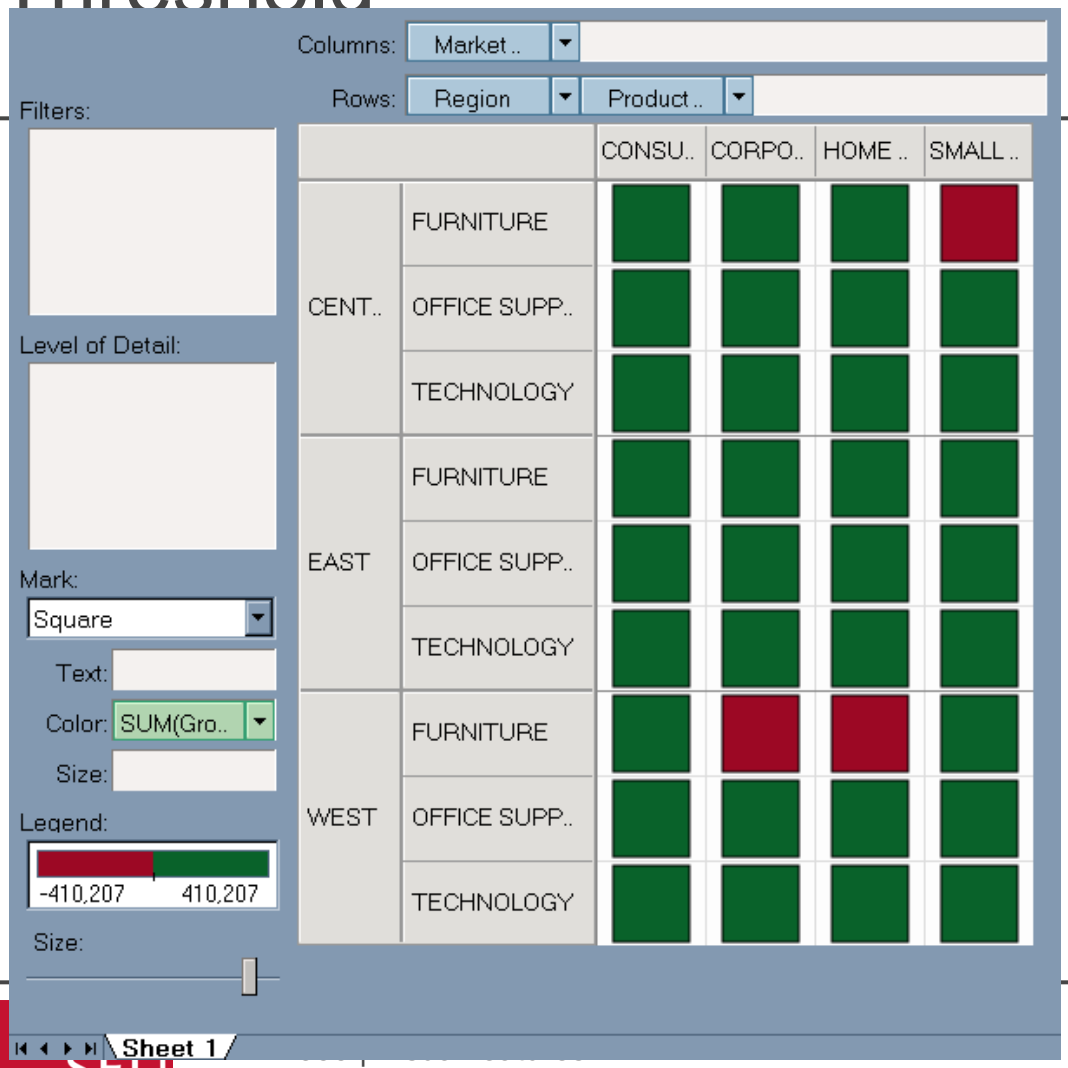
Skewed Data

Stepped Tableau colour scale

www.tableausoftware.com

Slide courtesy of Maureen Stone, Tableau Software

Threshold



Skewed Data

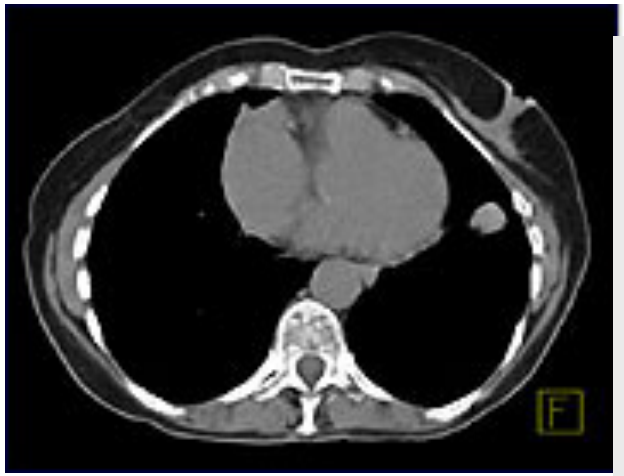
Binary Tableau colour scale

www.tableausoftware.com

Slide courtesy of Maureen Stone, Tableau Software

Color and Shading

- Shape is defined by lightness (shading)
- “Color” (hue, saturation) labels



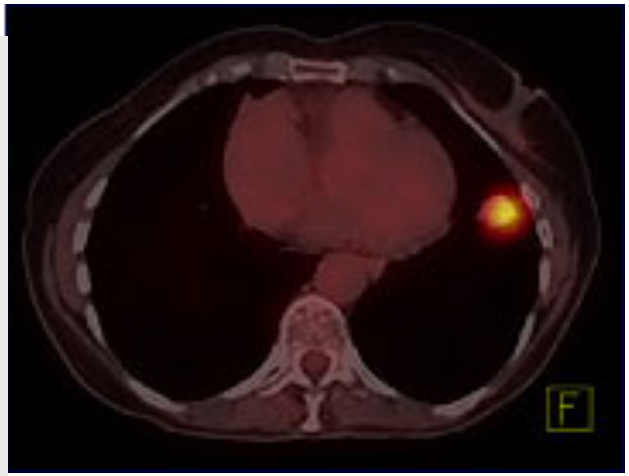


Image co[illegible] of Siemens

CT image (defines shape)

PET color highlights tumor

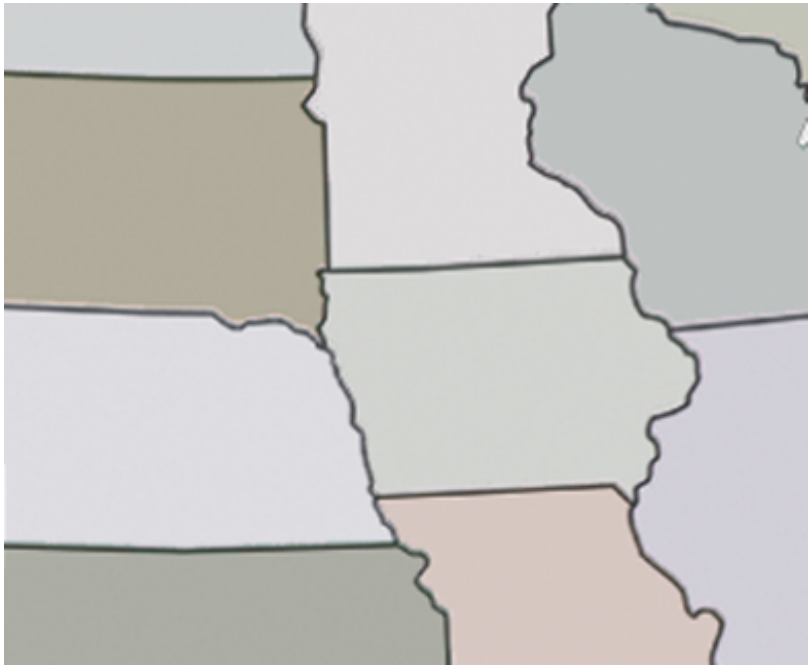
Can colour map multiple dimensions?

Do people interpret color blends as combinations of variables?

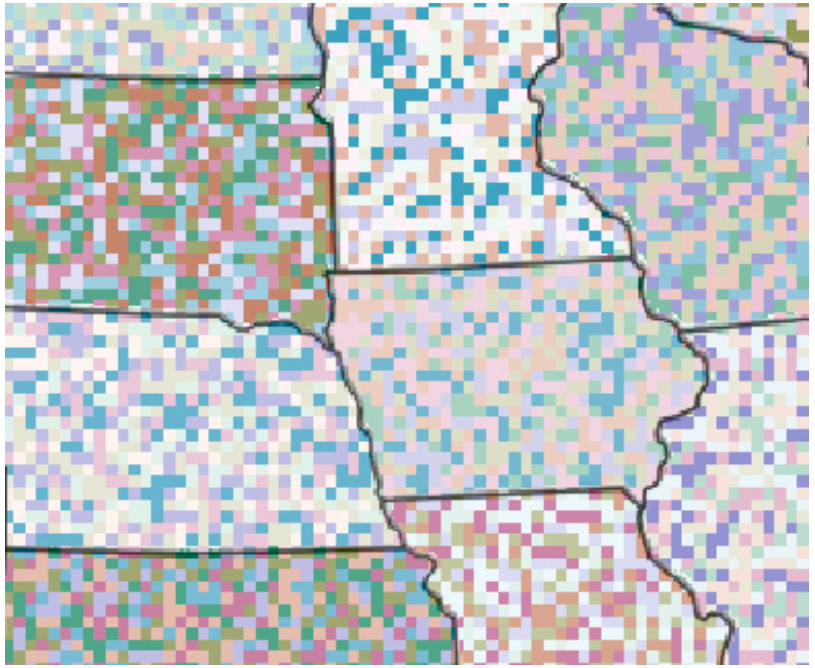
Using Color Dimensions to Display Data Dimensions Beatty and Ware

Color Weaves

6 variables = 6 hues, which vary in brightness



Additive mixture (blend)

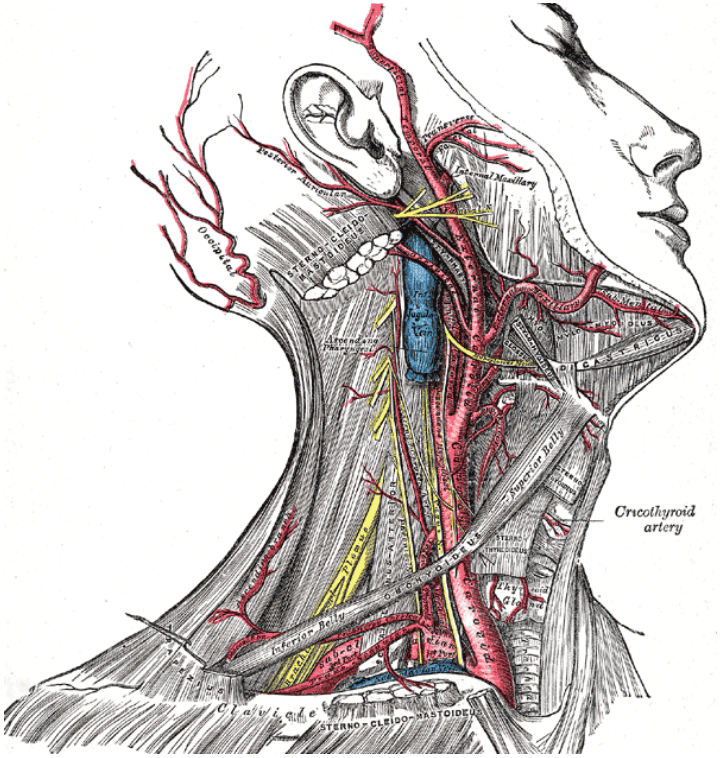


Spatial texture (weave)

Weaving versus Blending (APGV06 and SIGGRAPH poster)
Haleh Hagh-Shenas, Victoria Interrante, Christopher Healey and Sunghee Kim

Illustrative Color: to represent or mimic reality

Gray's *Anatomy of the Human Body*



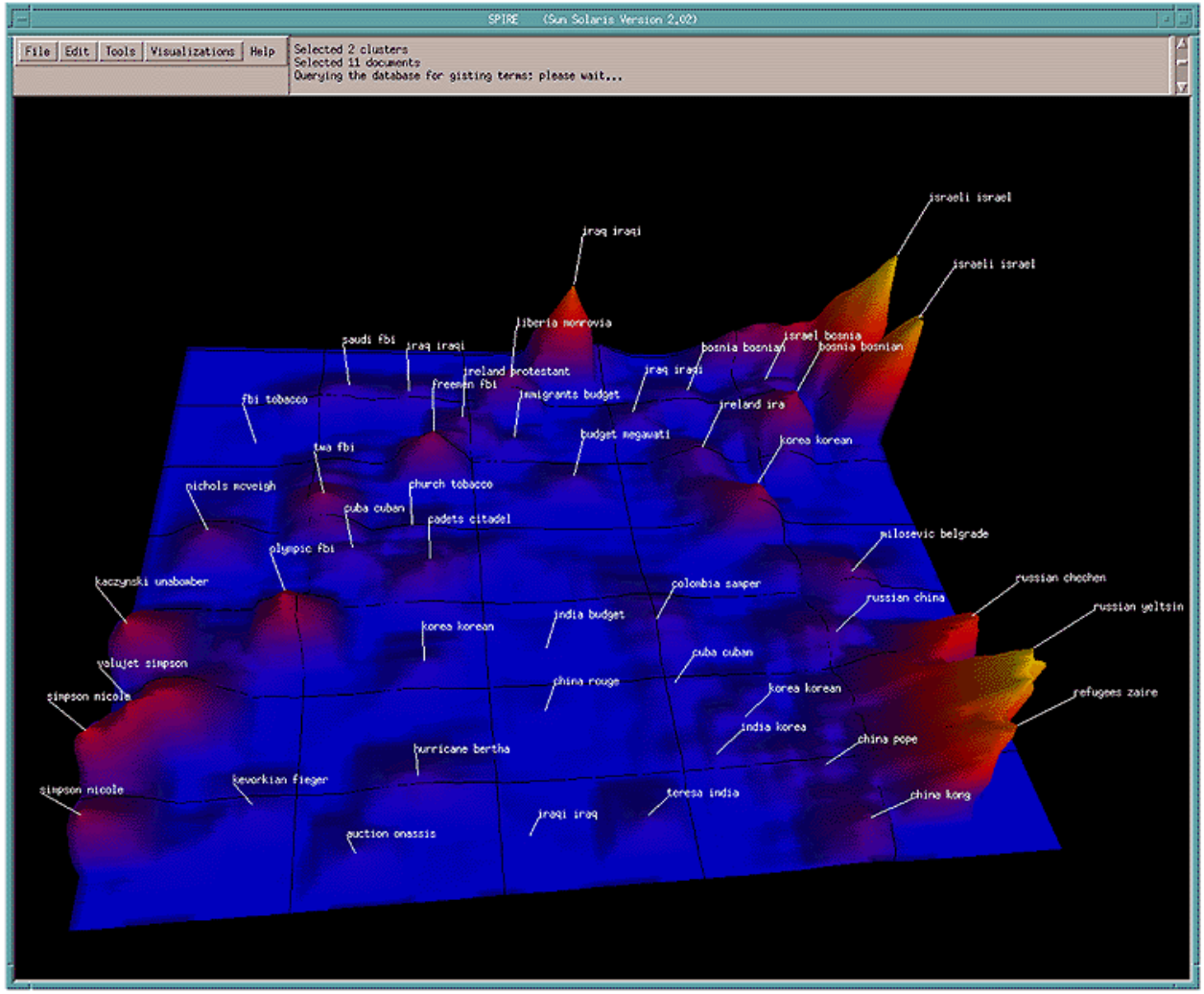
www.bartleby.com/107/illus520.html

Map of Point Reyes

www.nps.gov

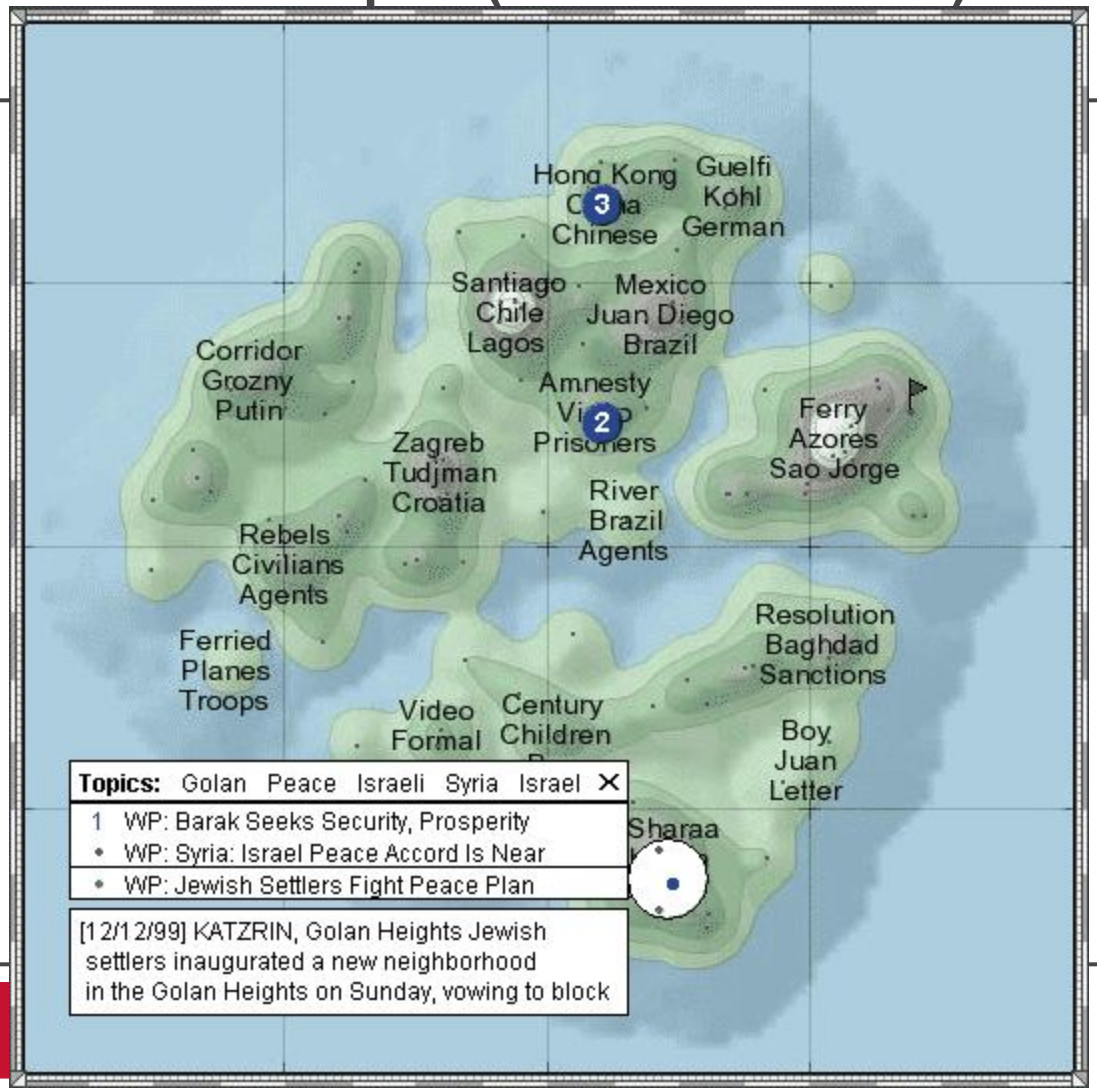
Slide courtesy of Maureen Stone, Tableau Software

ThemeView (original)

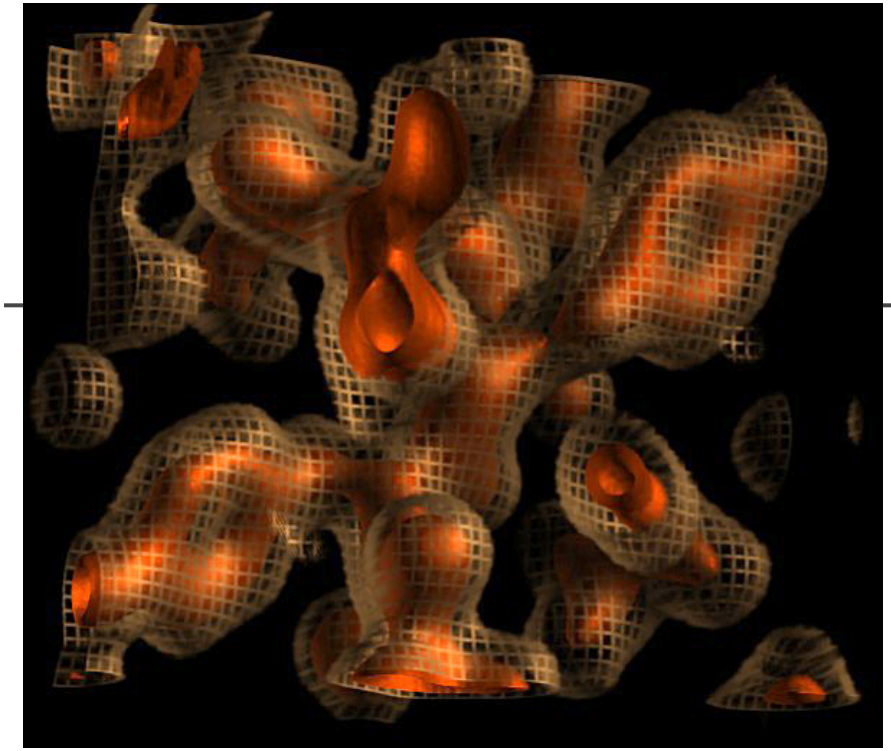


Courtesy of Pacific Northwest National Laboratories

ThemeScape (commercial)

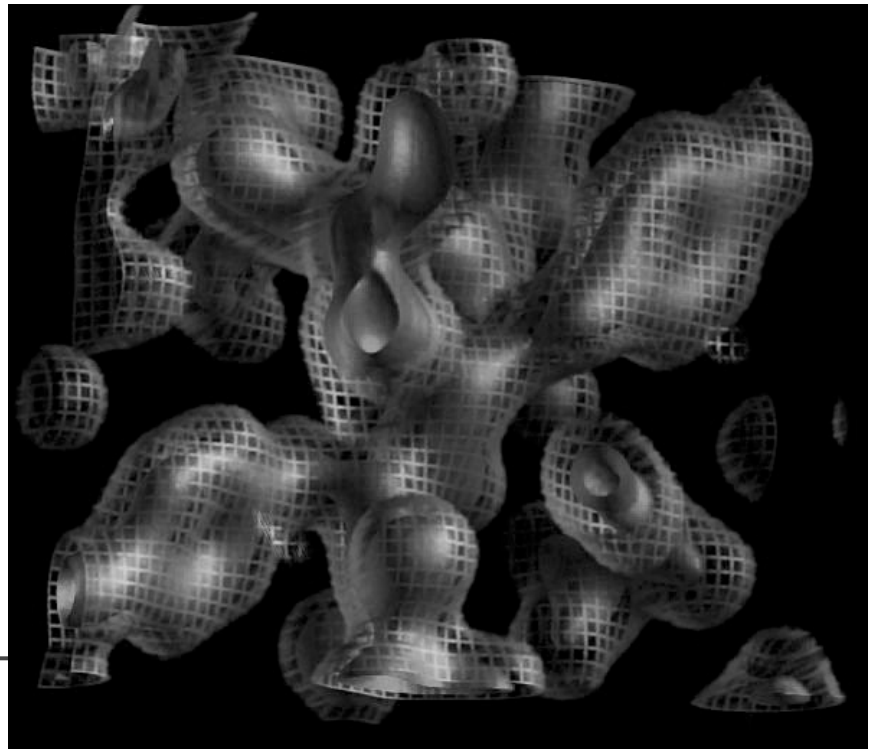


Courtesy of Cartia



Which has more information?

Which would you rather look at?



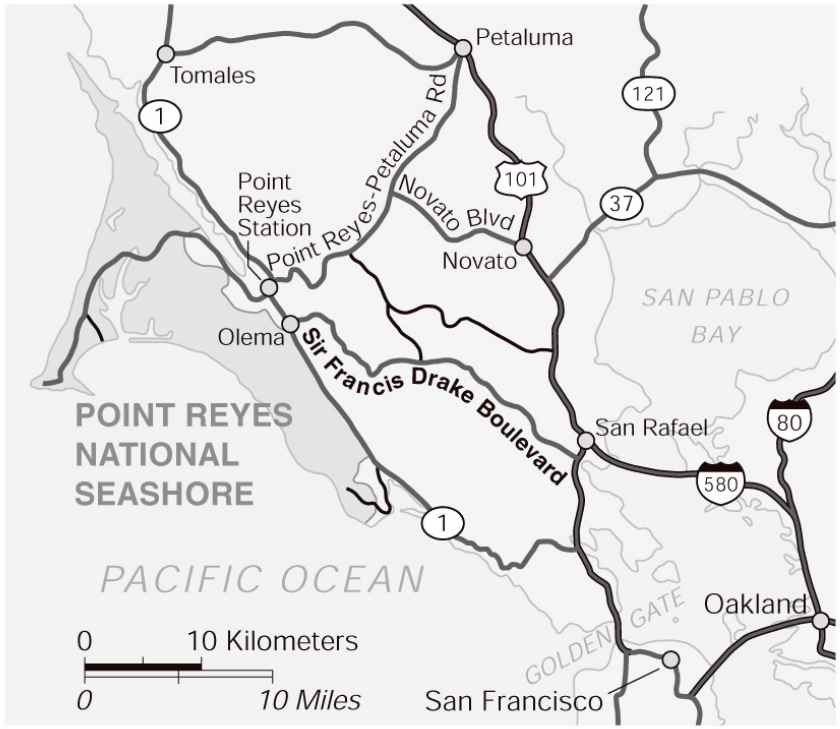
Visualization of isoelectron density surfaces around molecules

Marc Levoy (1988)

Get it Right in Black and White

- Value
  - Perceived lightness/darkness
  - Controlling value primary rule for design
- Value defines shape
  - No edge without lightness difference
  - No shading without lightness variation
- Value difference (contrast)
  - Defines legibility
  - Controls attention
  - Creates layering

Get it Right in Black and White



Colour includes Grey

www.nps.gov

Why does the logo work?

Value control

Contrast defines Legibility

Drop Shadows

Drop Shadow

Drop shadow adds edge

Primary colors on white

Primary colors on white

Primary colors on white

Primary colors on white

Primary colors on white

Primary colors on white

Primary colors on black

Primary colors on black

Primary colors on black

Primary colors on black

Primary colors on black

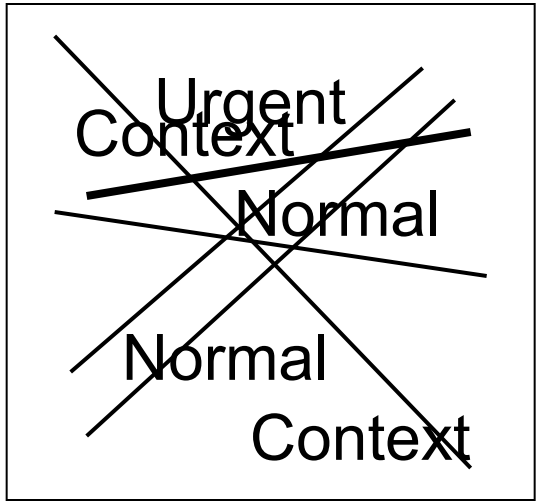
Primary colors on black

Readability

If you can't use color wisely,
it is best to avoid it entirely
Above all, do no harm

If you can't use color wisely,
it is best to avoid it entirely
Above all, do no harm.

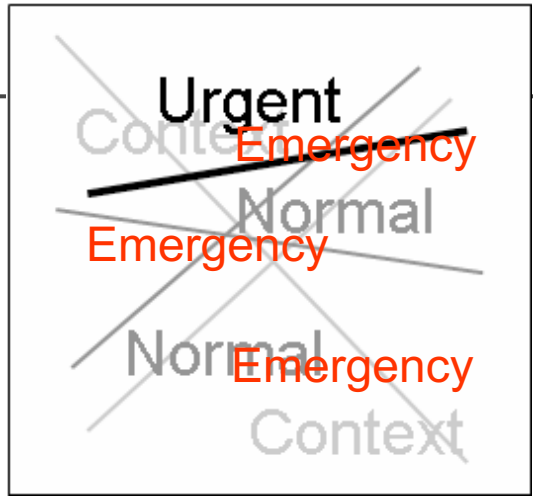
Value Contrast and Layering

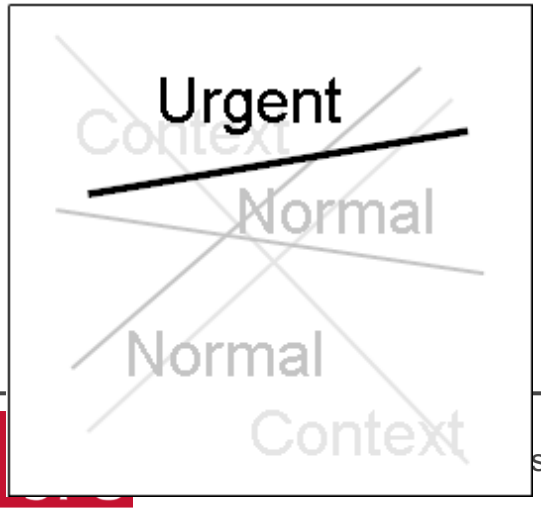


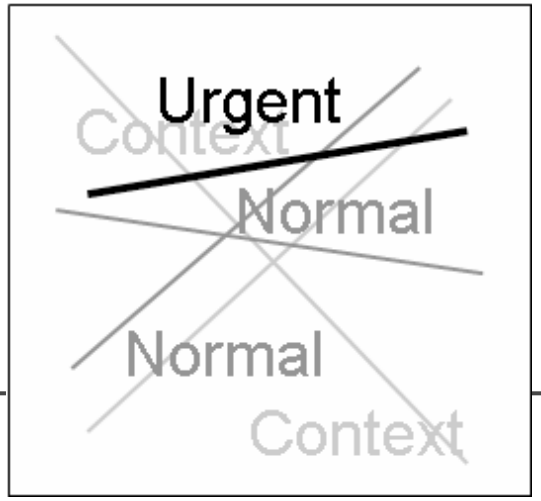
colorusage.arc.nasa.gov

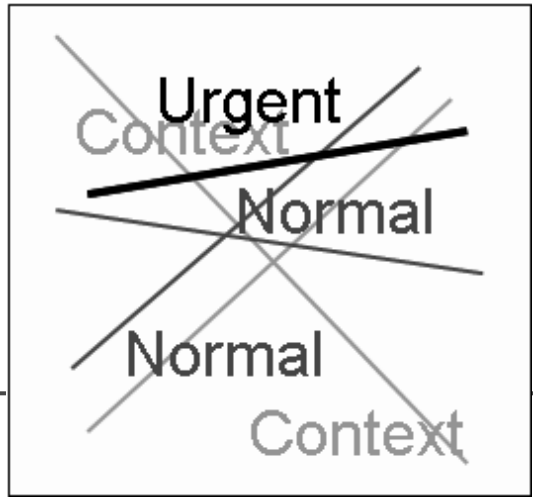
What Defines Layering?

- enhance with hue
- (popout effect essentially pulls forward)





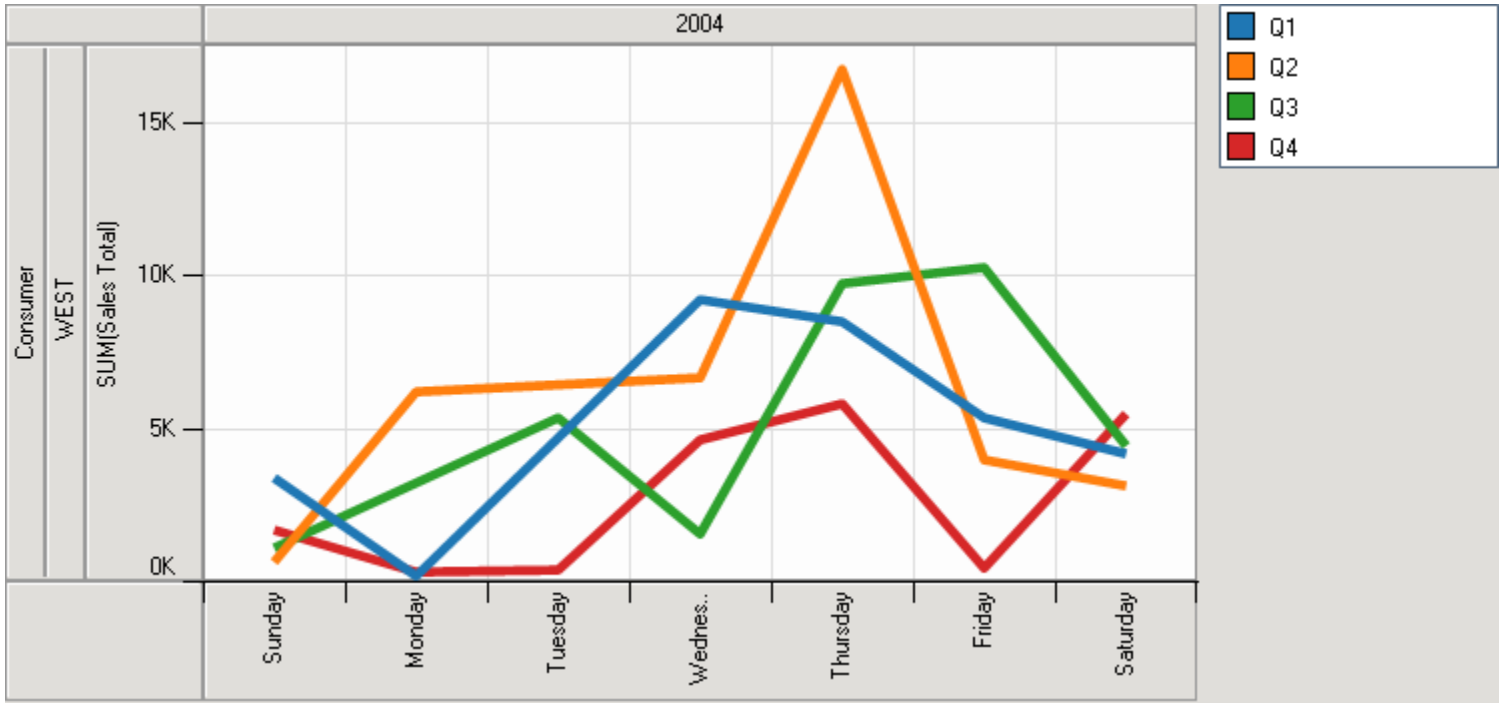




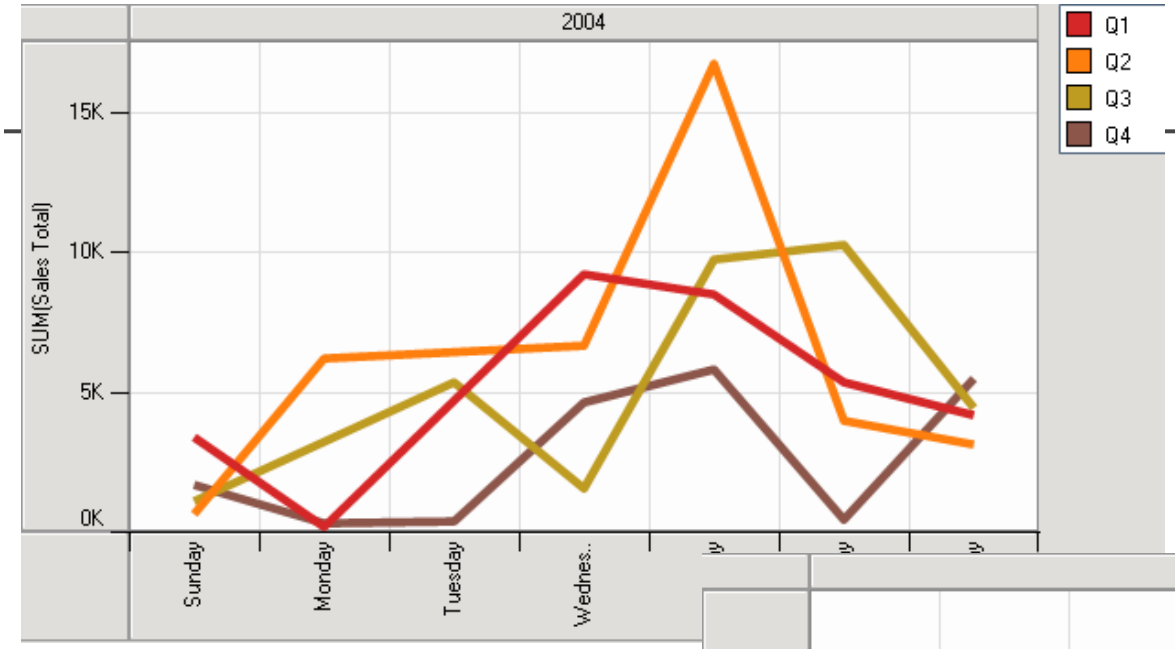
General guidelines … or from Tufte to practice [Stone, Ware]

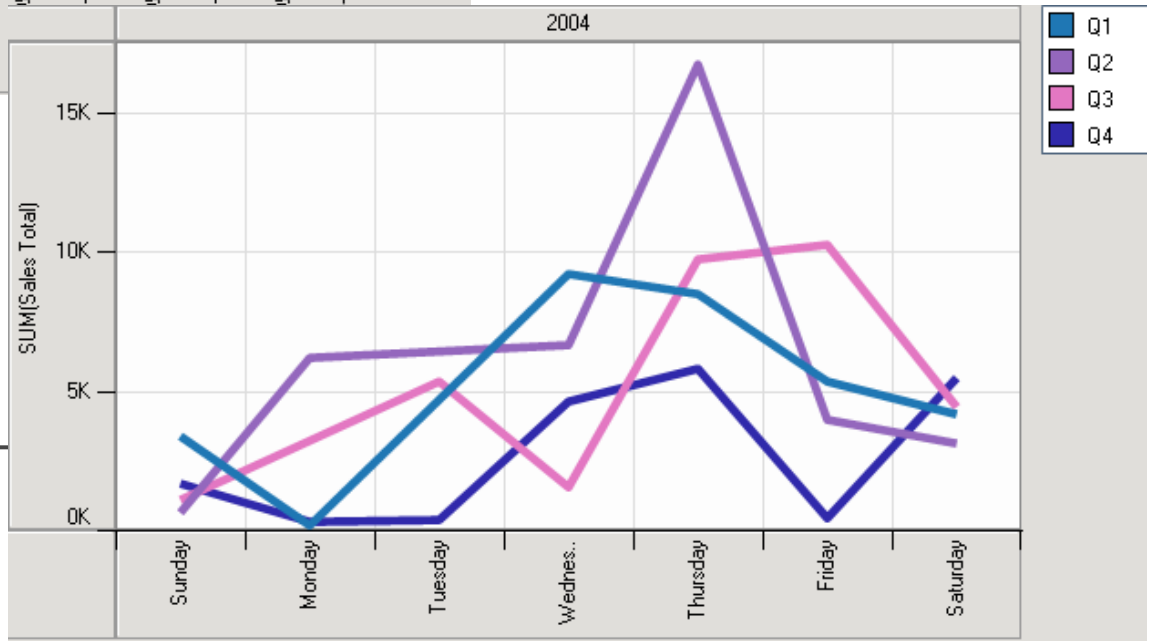
- Assign colour according to function
- Use **contrast** to highlight
- Use analogy to group
- Control value contrast for legibility
- Break isoluminance with borders

Maximum hue separation

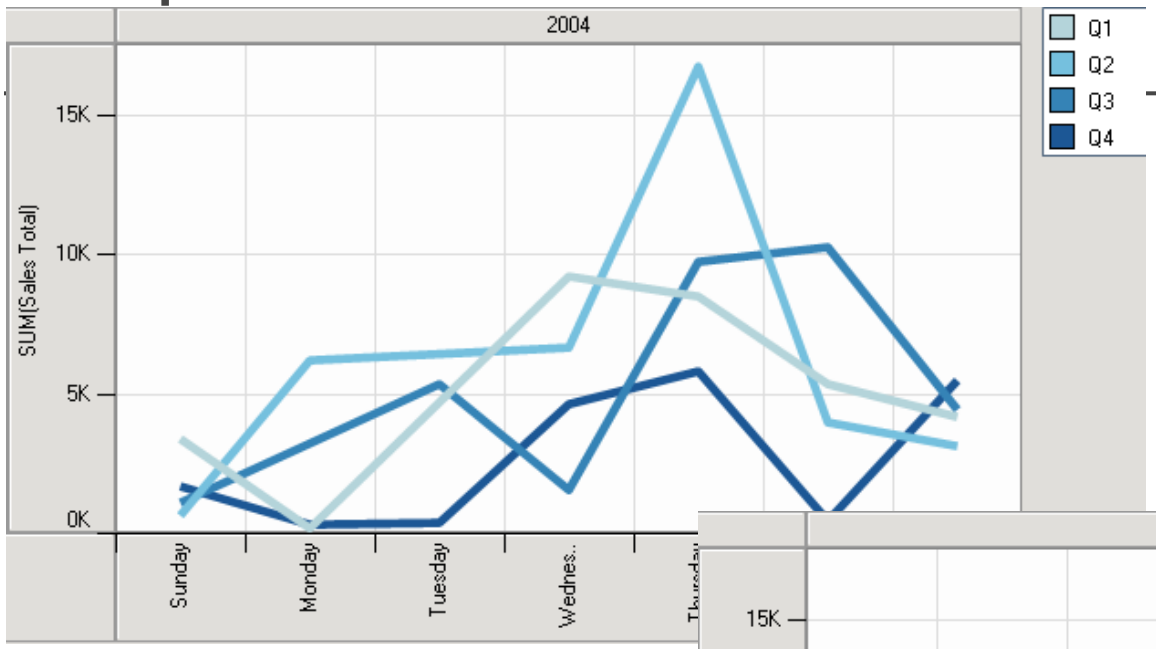


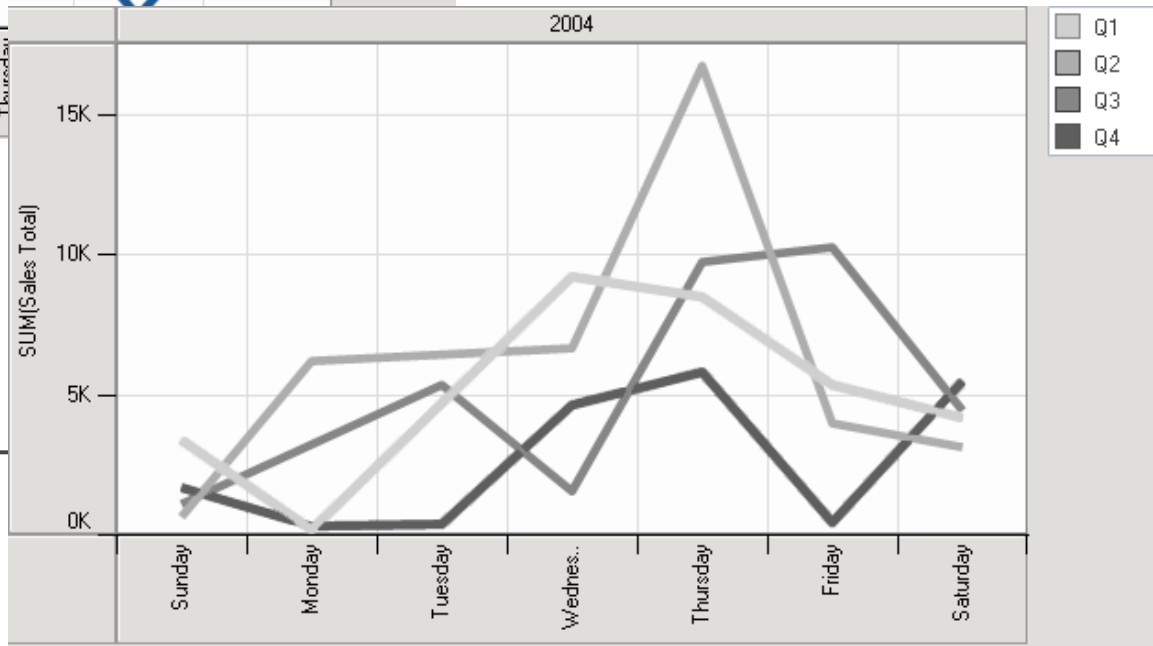
Analogous, yet distinct

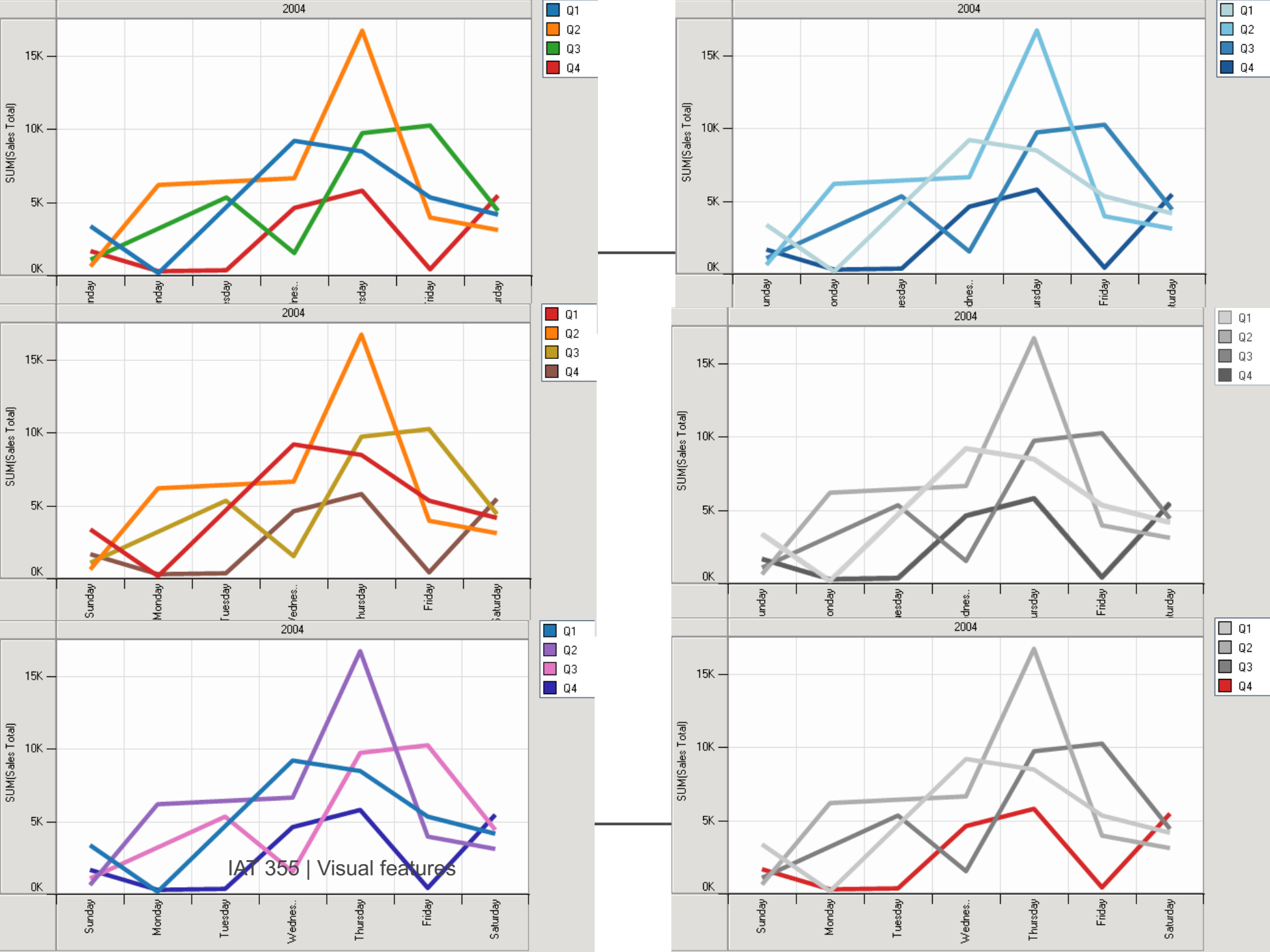




Sequential


2004
Q1
Q2
Q3
Q4
15K
10K
5K
0K
SUM(Sales Total)
Sunday
Monday
Tuesday
Wednes..


2004
Q1
Q2
Q3
Q4
15K
10K
5K
0K
SUM(Sales Total)
Sunday
Monday
Tuesday
Wednes..
Thursday
Friday
Saturday


2004
SUM(Sales Total)
15K
10K
5K
0K
Q1
Q2
Q3
Q4
Sunday
Monday
Tuesday
Thursday
Friday
Saturday

Stephen Few's practical rules on charts

1. If you want different objects of the same color in a table or graph to look the same, make sure that the background—the color that surrounds them—is consistent.
2. If you want objects in a table or graph to be easily seen, use a background color that contrasts sufficiently with the object.

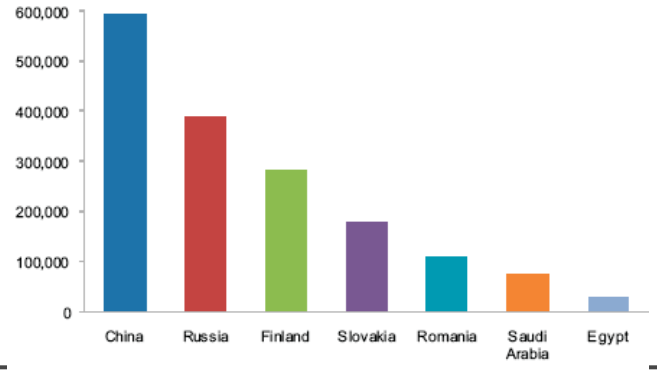
Don't do this!

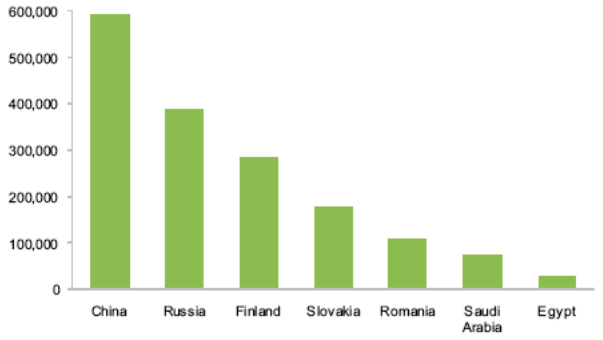
	Laptops	Desktop PCs	Harddisks	Flash Memory	Screens	Keyboards	Printers	Scanners	PDAs	Projectors	Cameras
California	-2,216	4,497	884	3,252	8,564	3,418	6,582	-3,891	2,333	1,356	5,450
Colorado	3,410	0	2,338	2,676	1,567	367	1,361	3,249	828	1,272	-141
Connecticut	0	0	4	0	2,998	1,219	673	0	0	610	989
Florida	0	0	583	765	2,305	940	1,737	2,727	0	1,344	497
Illinois	0	0	9,384	3,405	4,362	2,331	4,495	3,464	1,366	0	1,569
Iowa	377	0	412	6,577	210	4,487	872	189	5,622	0	3,466
Louisiana	0	2	1,455	1,501	1,138	0	920	1,349	0	0	991
Massachusetts	0	0	-297	0	12,489	707	0	0	0	492	367
Missouri	0	0	1,224	-169	986	367	1,011	1,125	-210	-45	-685
Nevada	0	875	374	884	-45	11,934	410	210	4,497	-10,978	4,355
New Hampshire	1,009	0	-166	0	897	376	0	0	0	882	-43
New Mexico	0	-208	-161	377	902	0	369	-683	0	0	202
New York	0	0	-17,034	0	8,564	3,254	0	0	3,405	2,326	7,589
Ohio	820	0	2,671	495	827	2,305	355	-134	2,724	0	705
Oklahoma	0	2,722	492	495	1,274	0	2,305	700	-15,034	0	571
Oregon	-139	357	818	708	395	487	2,671	834	2,310	2,727	1,269
Texas	0	1,356	3,415	808	5,452	0	2,336	1,567	0	0	829
Utah	1,130	1,006	1,501	1,009	678	-156	1,222	979	367	-208	904
Washington	0	768	488	2,998	1,661	1,452	1,418	564	918	-5	1,141
Wisconsin	497	0	1,418	1,450	576	916	767	1,740	2	0	1,337

Few (2)

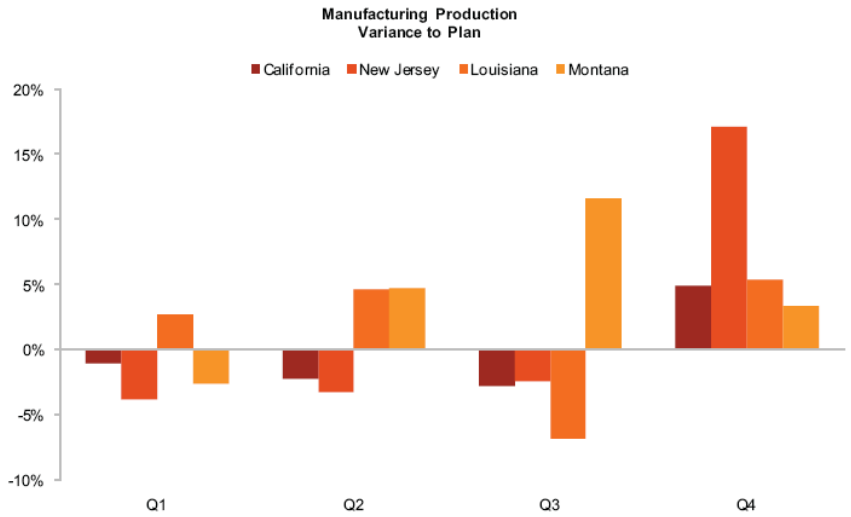
3. Use colour only when needed to serve a particular communication goal

4. Use different colours only when they correspond to differences of meaning in the data



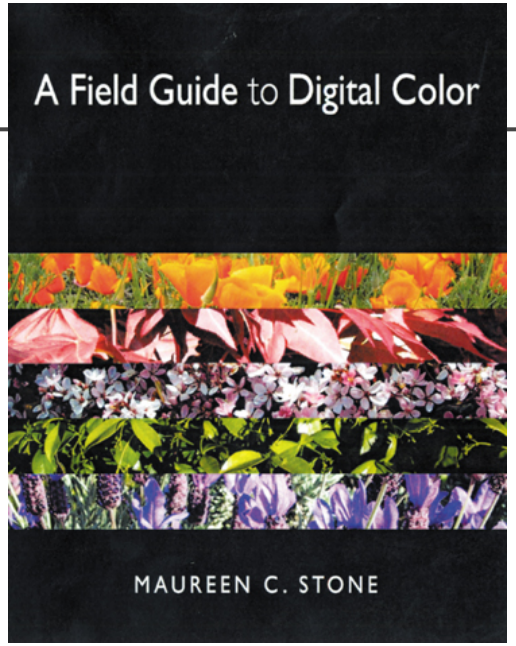


Few (3)



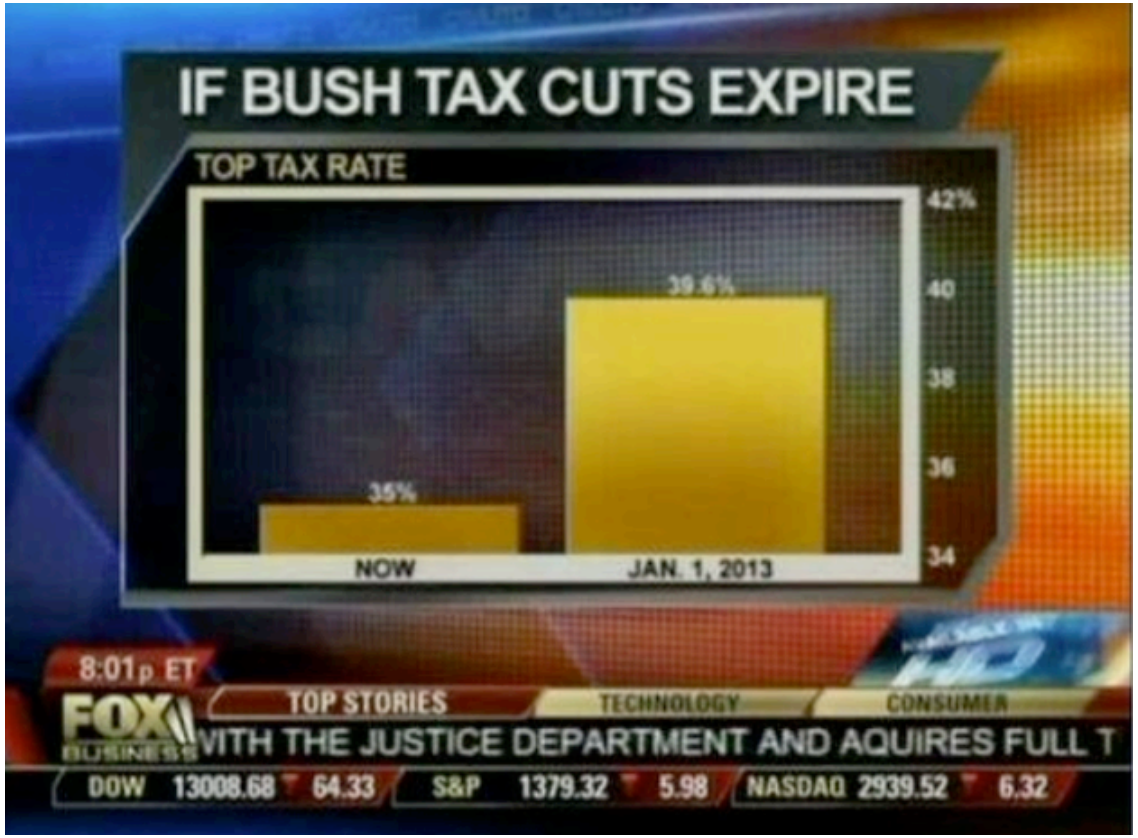
5. Use soft, natural colors for most: bright and/or dark colors to highlight.
6. When using color to encode a sequential range of quantitative values, stick with a single hue (or a small set of closely related hues) and vary intensity from pale colors for low values to increasingly darker and brighter colors for high values.

Additional Resources on Color

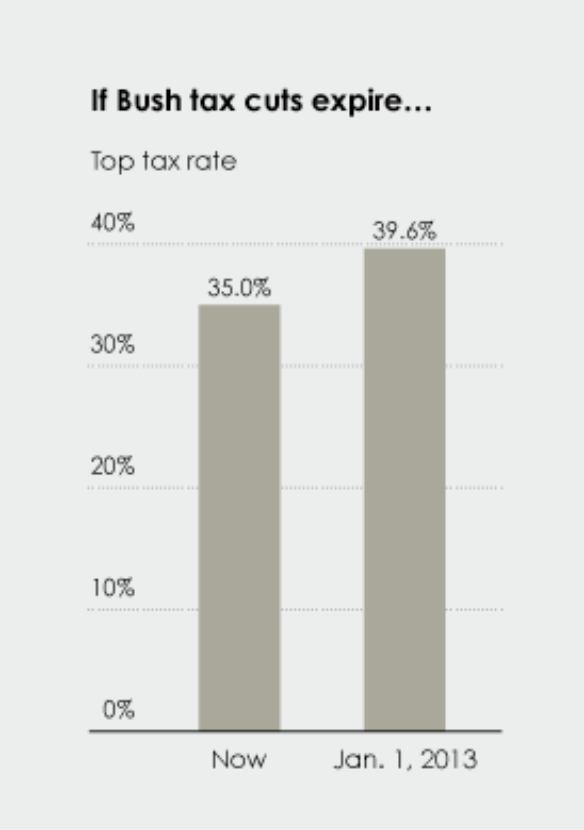


- Even without the "u" ….
- Stone Soup website
  - http://www.stonesc.com/Vis06
  - Final copy of slides, references
- *A Field Guide to Digital Color*
  - Maureen C. Stone
  - Published by A.K. Peters
- Stephen Few's articles on color
  - http://www.perceptualedge.com

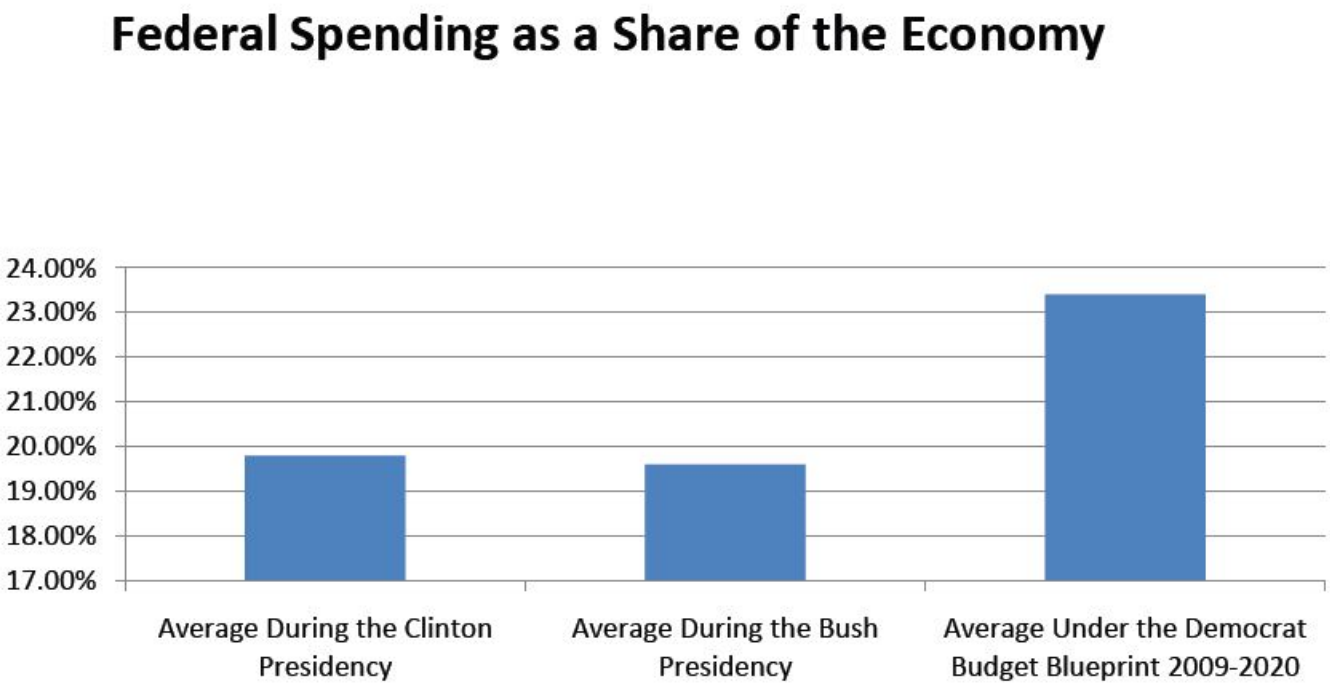
Fox News to the (lie) – what is this??



The real story

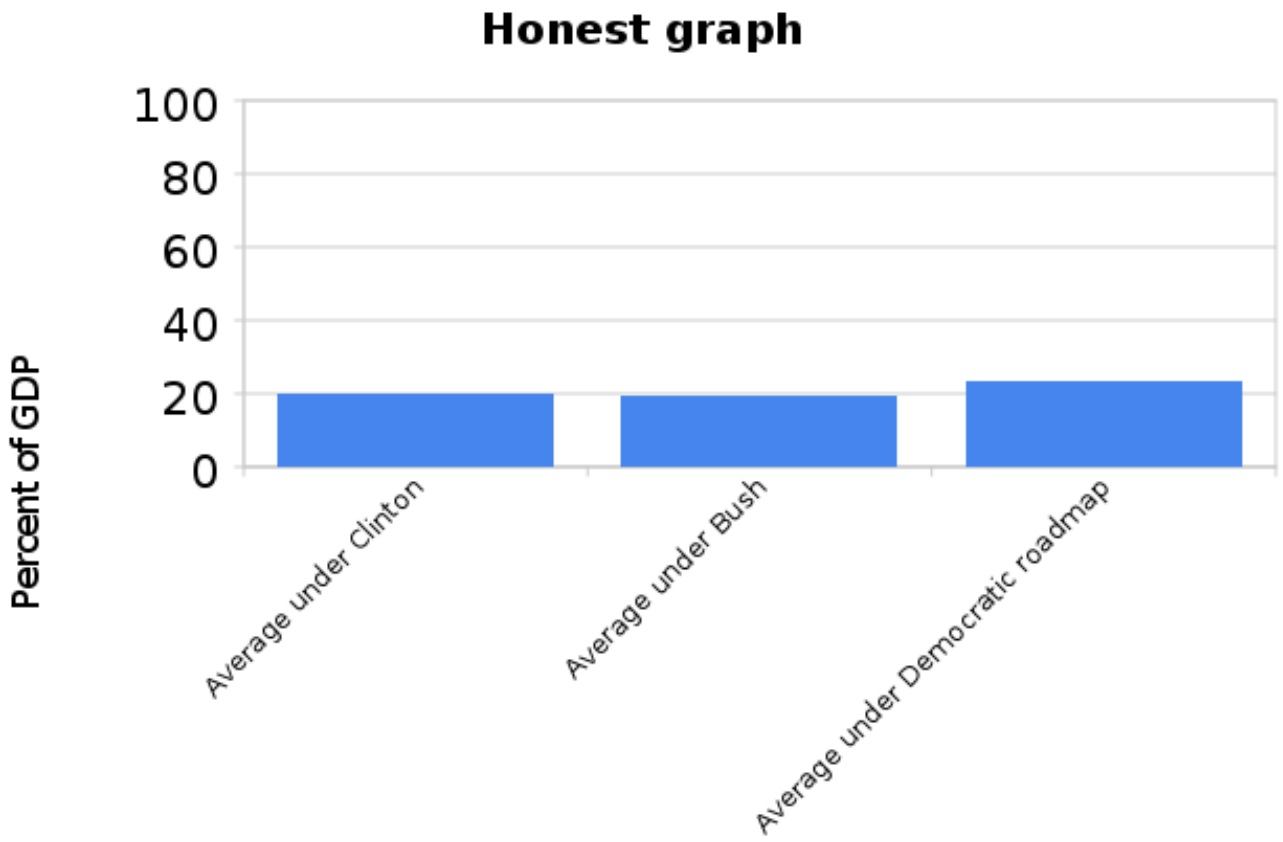


Lies, Damn Lies, and Bad Graphs



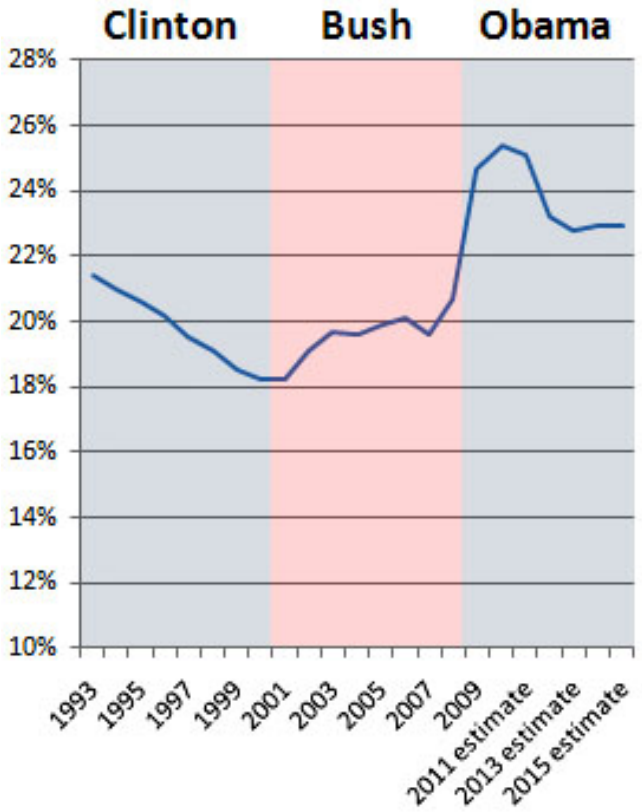
http://www.tnr.com/blog/jonathan-cohn/77893/lying-graphs-republican-style

More honest



http://www.tnr.com/blog/jonathan-cohn/77893/lying-graphs-republican-style

Better



http://www.tnr.com/blog/jonathan-cohn/77893/lying-graphs-republican-style

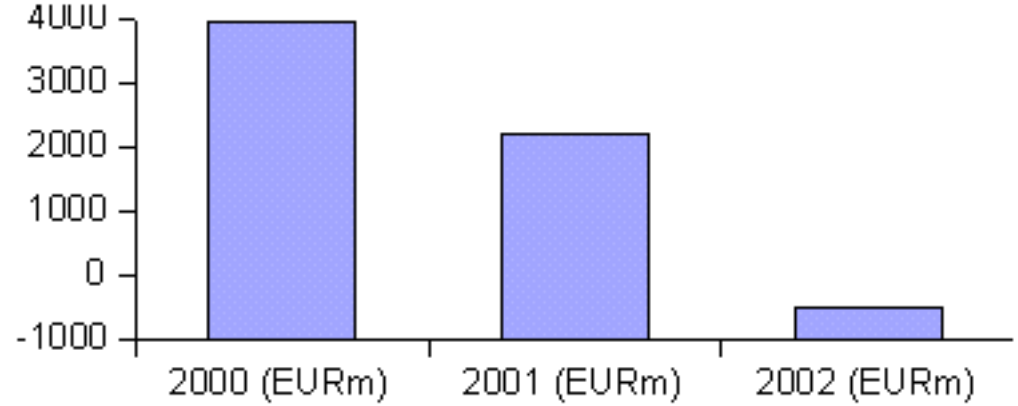


Chart 8: Imaginative net profits for Nokia 2000-2002, with x-axis at -1000

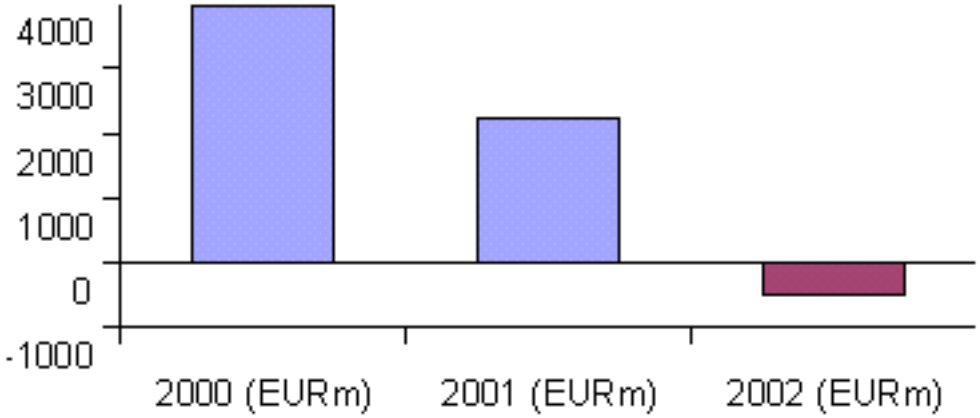


Chart 7: Imaginative net profits for Nokia 2000-2002

Perspective as well as value scales

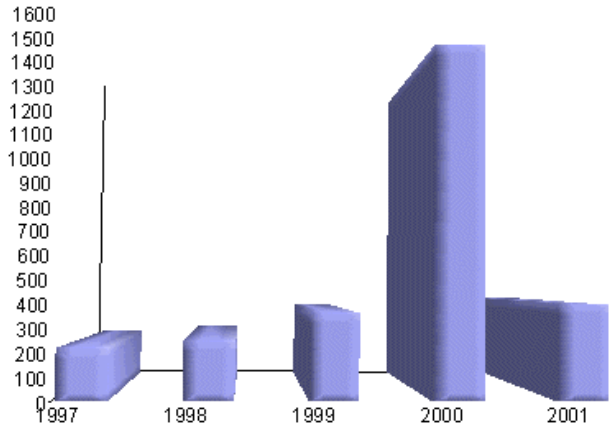


Chart 19: Net profits for Sonera 1997 - 2001 (EURm), nice 3d effect

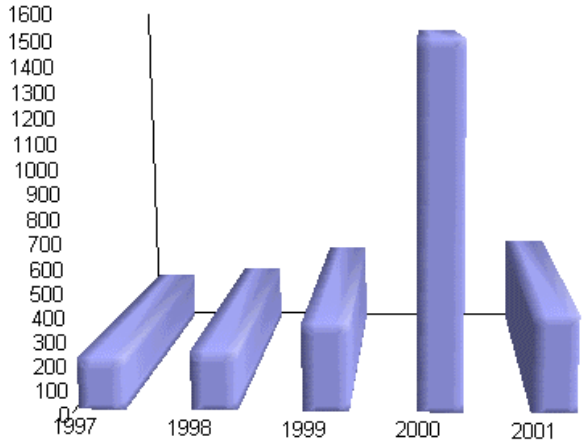
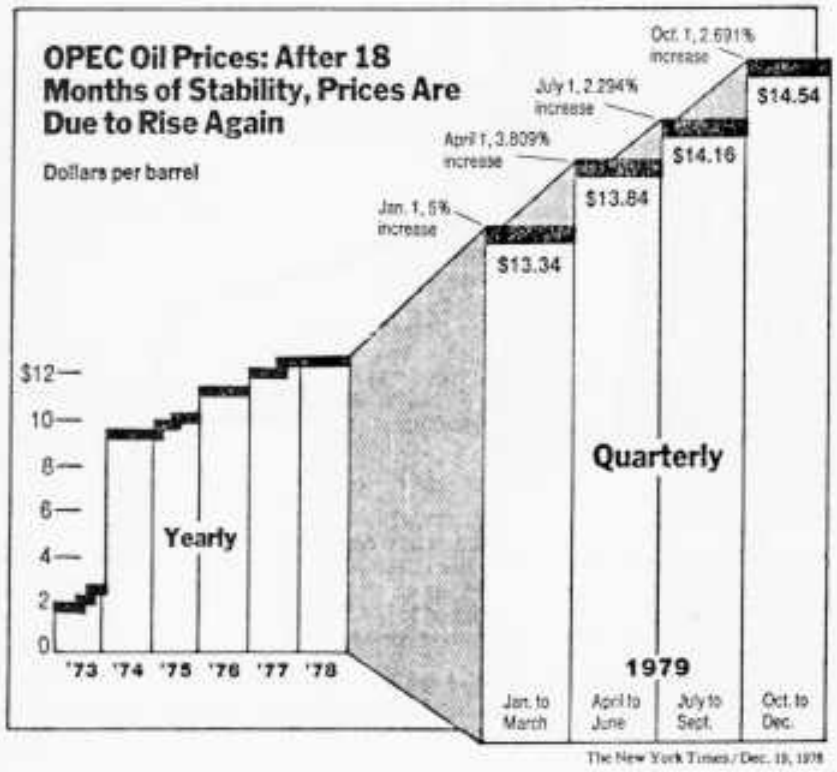


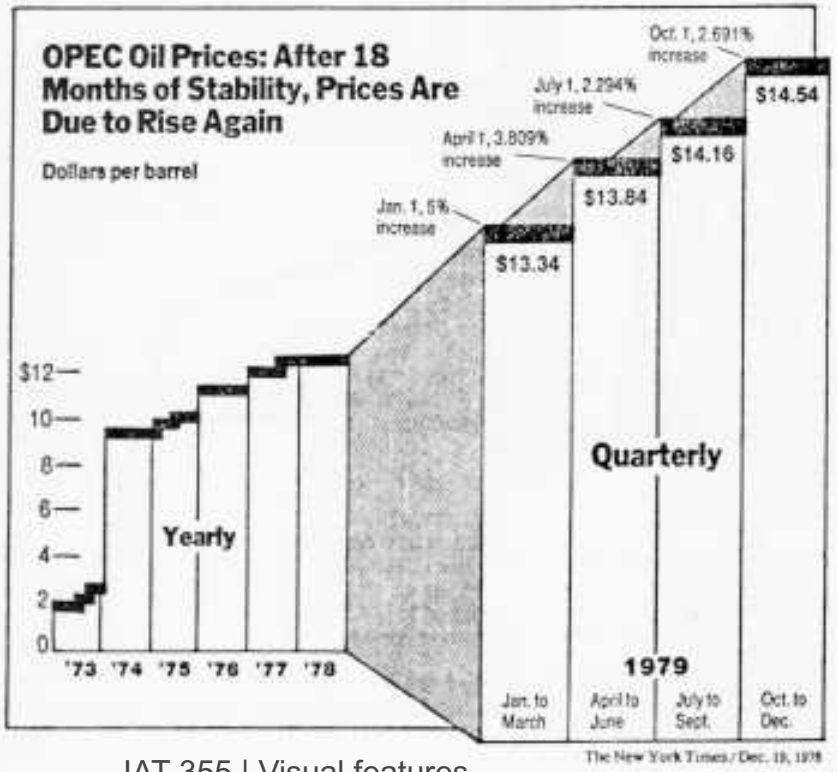
Chart 20: Net profits for Sonera 1997 - 2001 (EURm), another perspective into same 3d effect

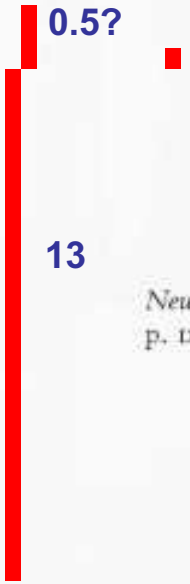


New York Times, December 19, 1978, p. D-7.

Changing Scale

Design variation corrupts this display:





New York Times, December 19, 1978, p. D-7.

Changing Scale

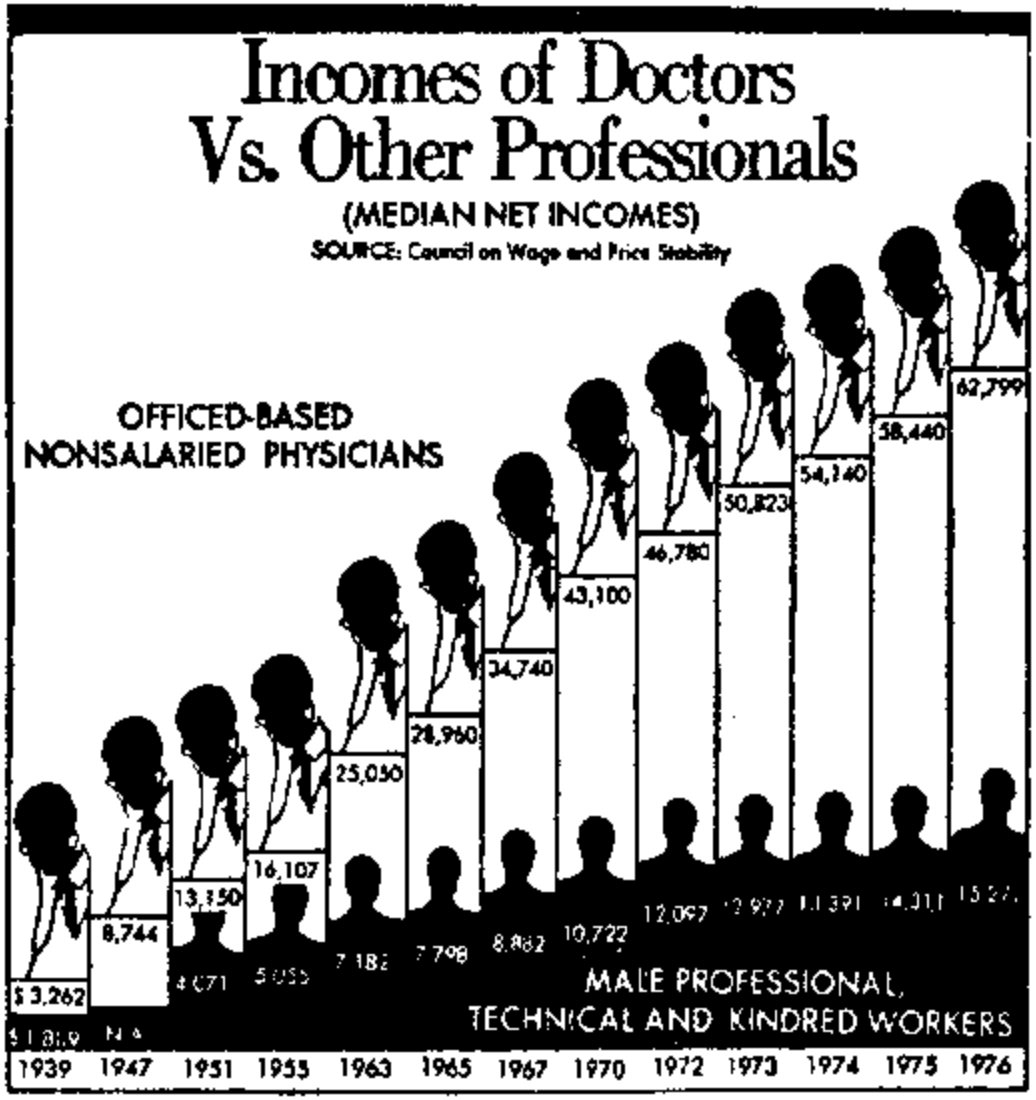
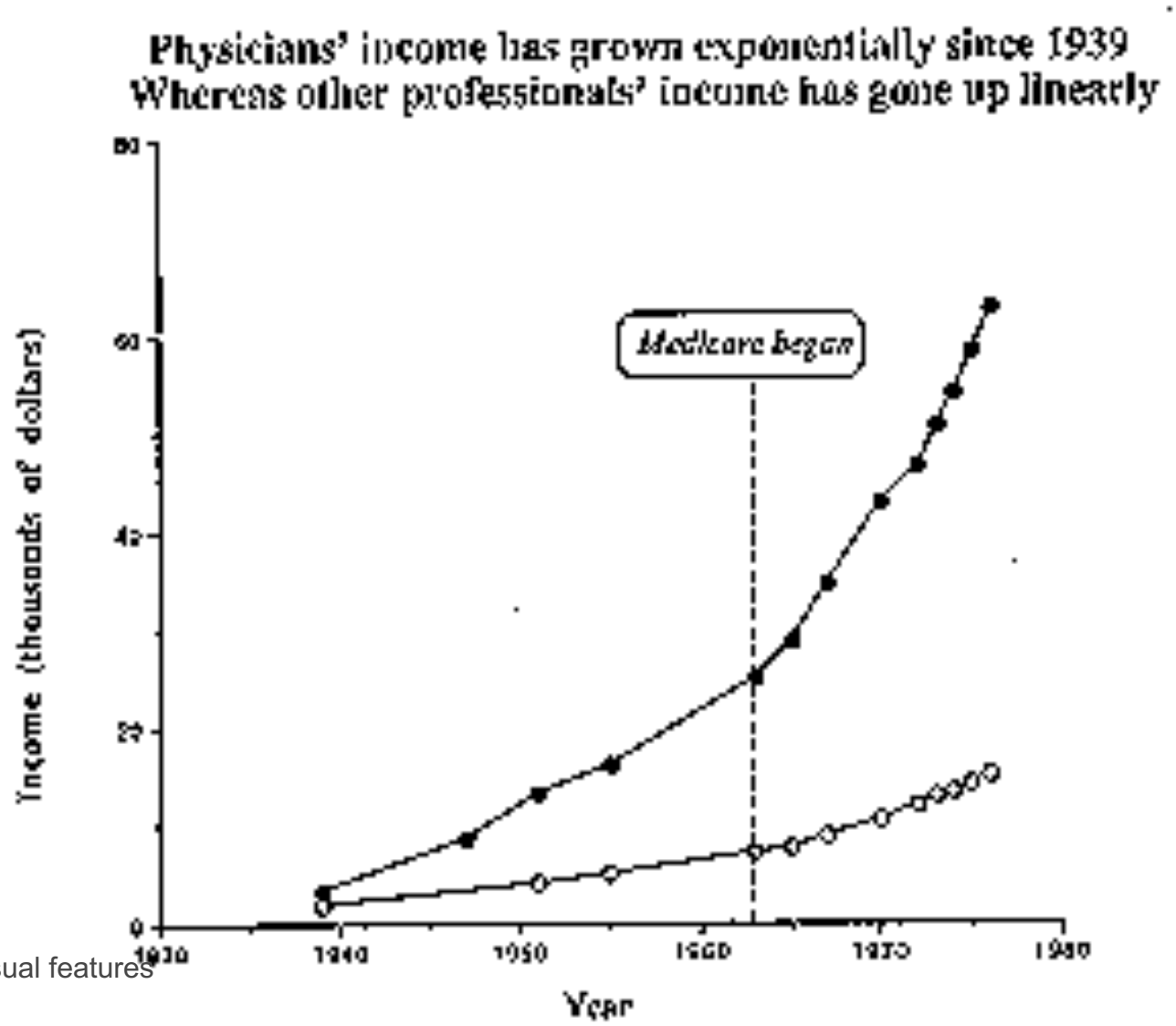


Figure 13. Changing scale in mid-axis to make exponential growth linear (© The Washington Post).

…with linear time scale



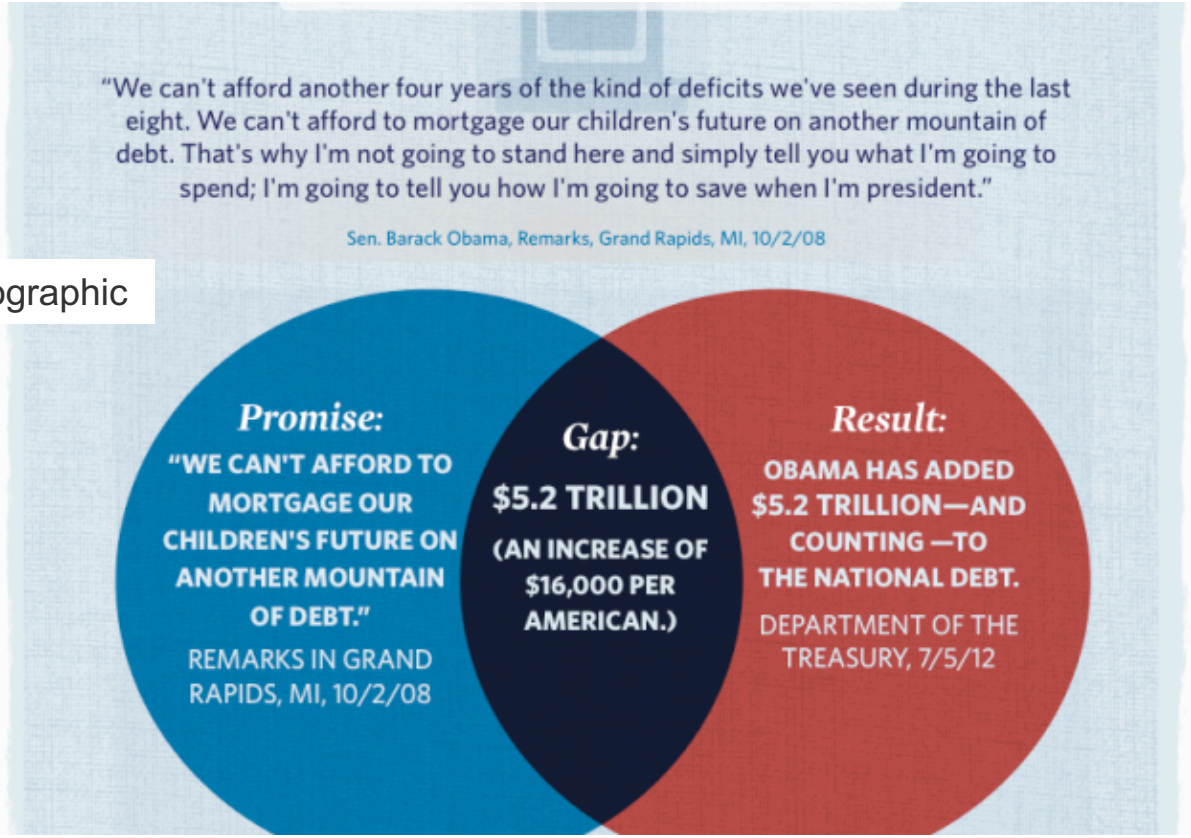
And more

SFU

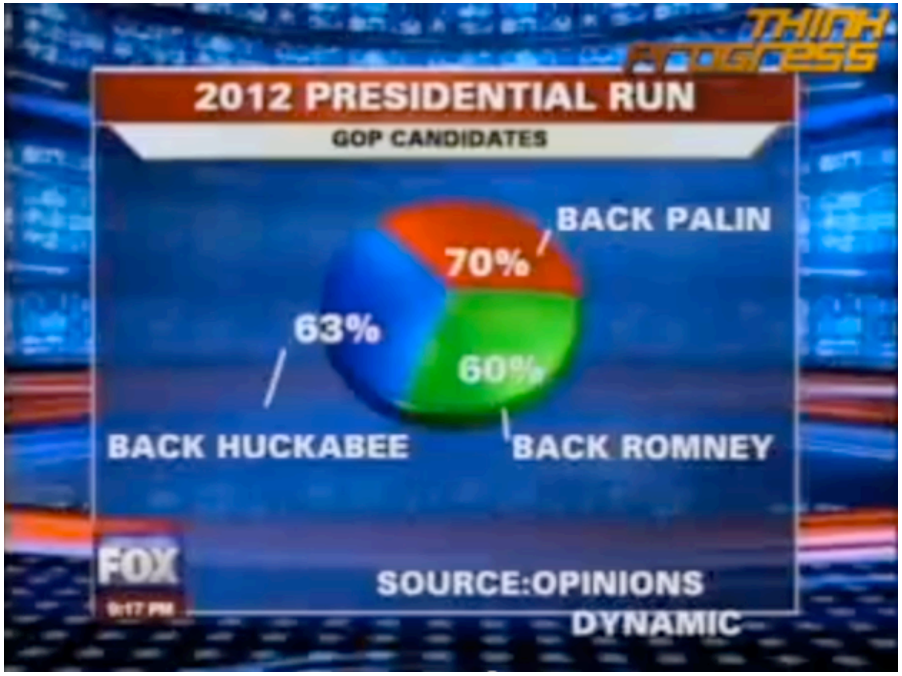
SCHOOL OF INTERACTIVE
ARTS + TECHNOLOGY

Umm, what's wrong here?

Mitt Romney campaign infographic

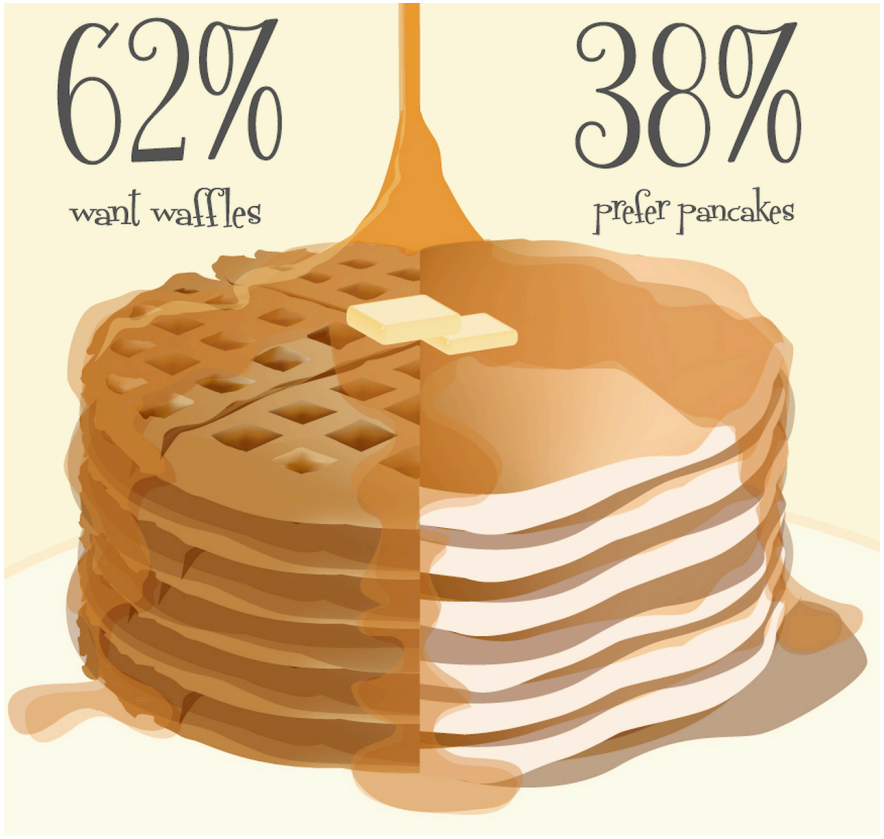


Fox news again: best pie chart ever

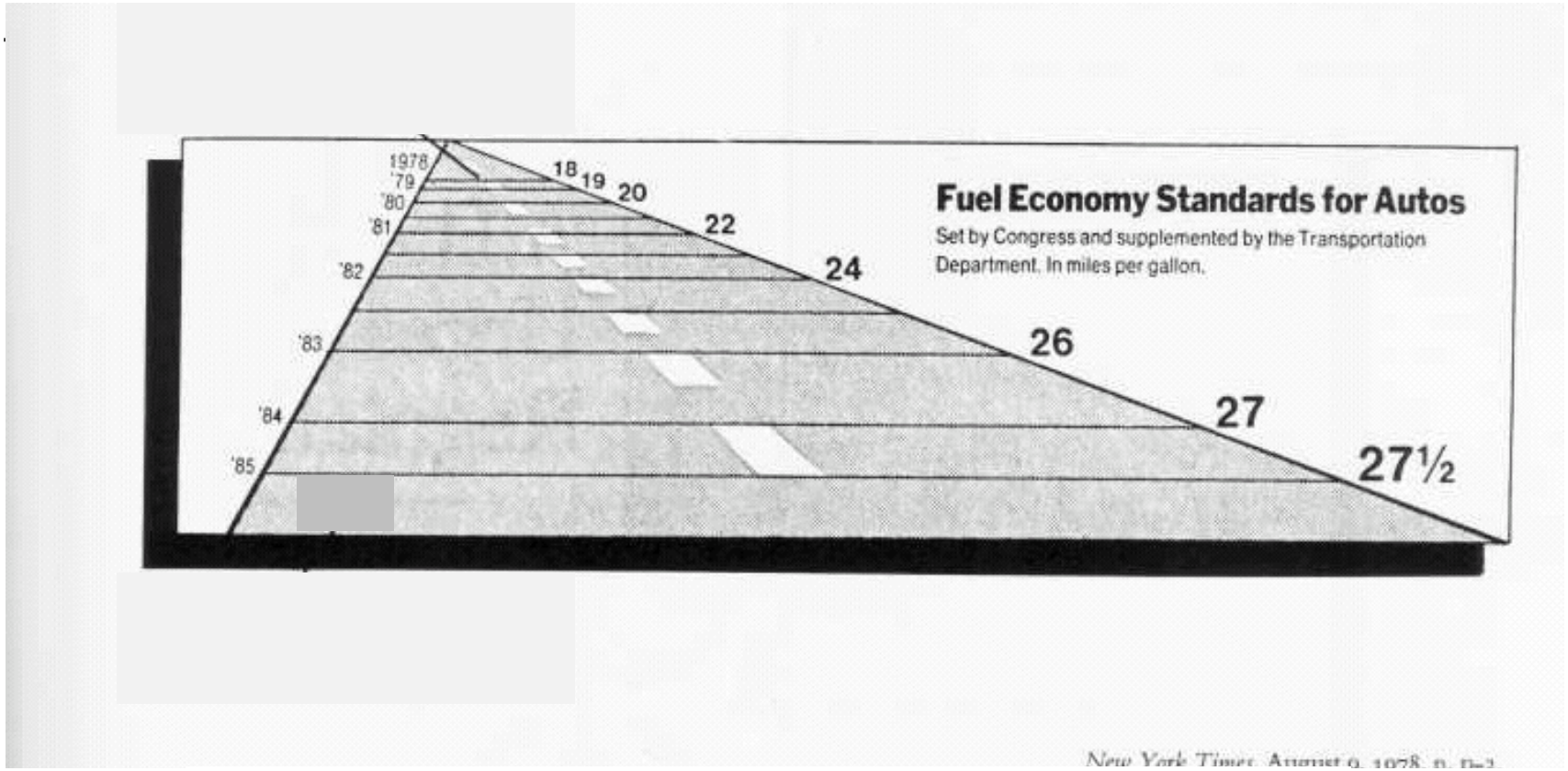


http://flowingdata.com/2009/11/26/fox-news-makes-the-best-pie-chart-ever/

??

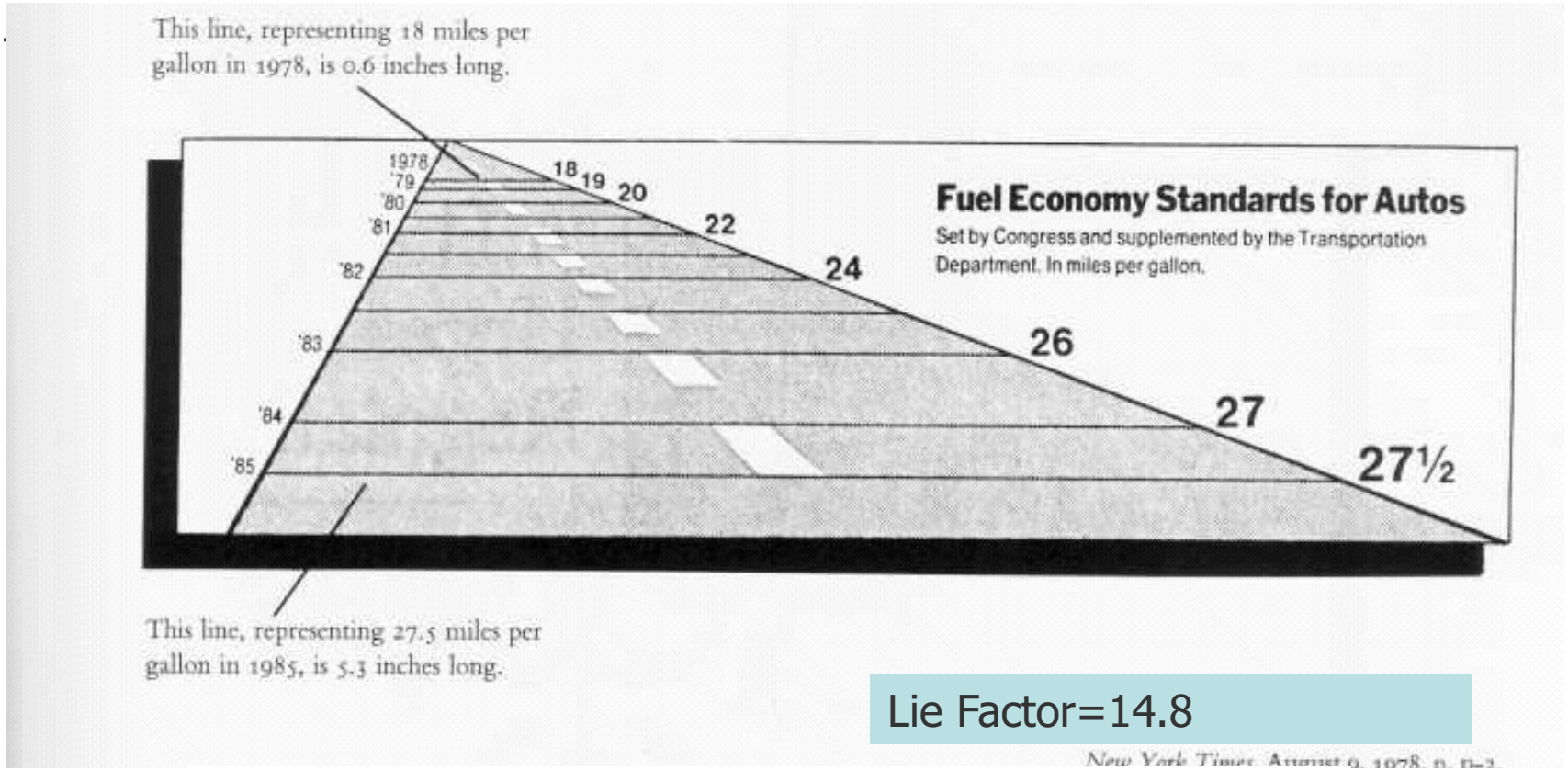

62%
want waffles
38%
prefer pancakes

What about this?



E.R. Tufte, "The Visual Display of Quantitative Information", 2nd edition

Lie factor



E.R. Tufte, "The Visual Display of Quantitative Information", 2nd edition

Lie Factor

$$Lie\ Factor = \frac{size\ of\ effect\ shown\ in\ graphic}{size\ of\ effect\ in\ data} =$$

$$= \frac{\frac{(5.3-0.6)}{0.6}}{\frac{(27.5-18.0)}{18}} = \frac{7.833}{0.528} = 14.8$$

Tufte requirement: 0.95<Lie Factor<1.05

(E.R. Tufte, "The Visual Display of Quantitative Information", 2nd edition)

Tufte's Principles of Graphical Excellence

- Give the viewer
  - the greatest number of ideas
  - in the shortest time
  - with the least ink in the smallest space.

- Tell the truth about the data!

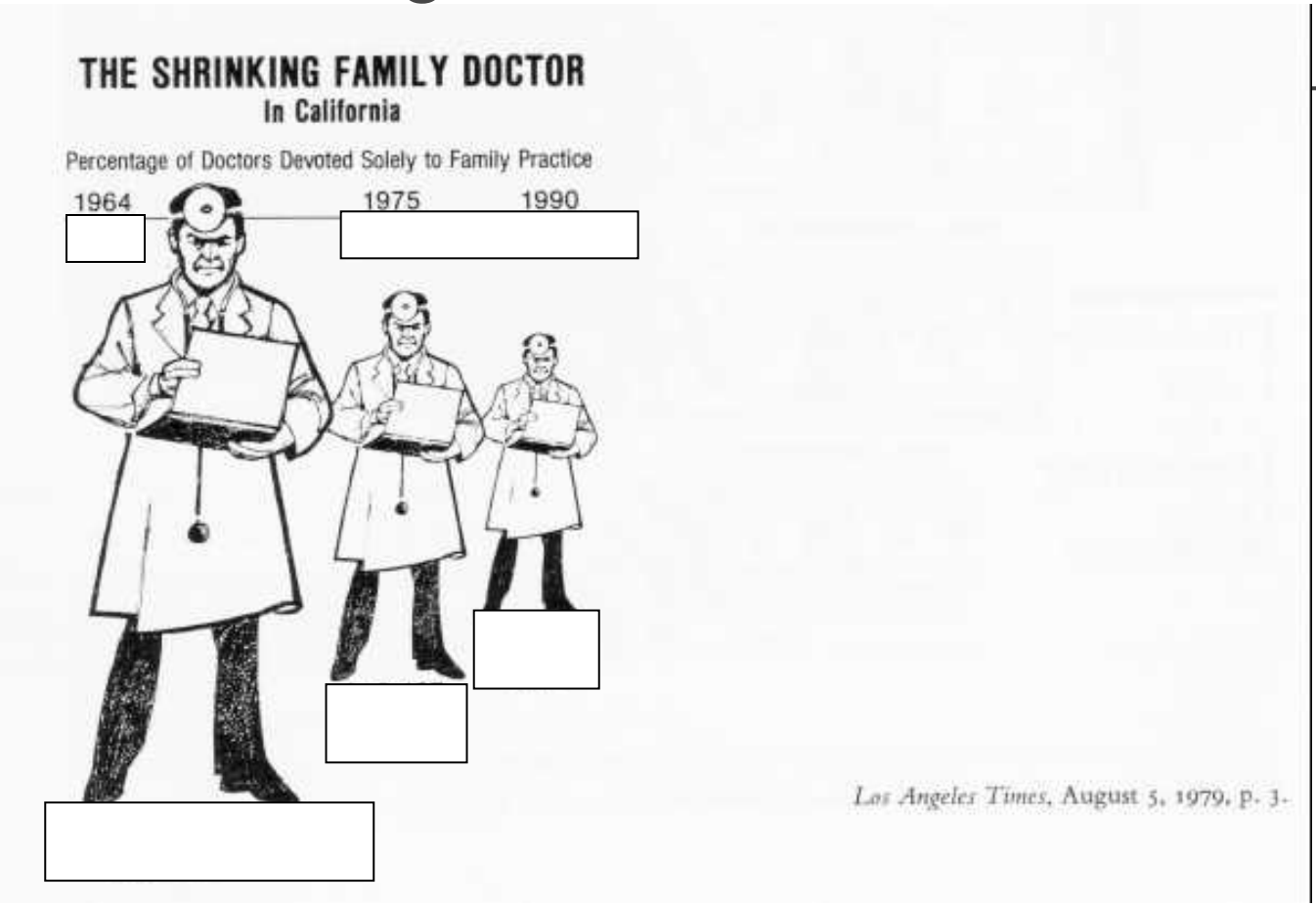
E.R. Tufte, "The Visual Display of Quantitative Information", 2nd edition

Two Principles

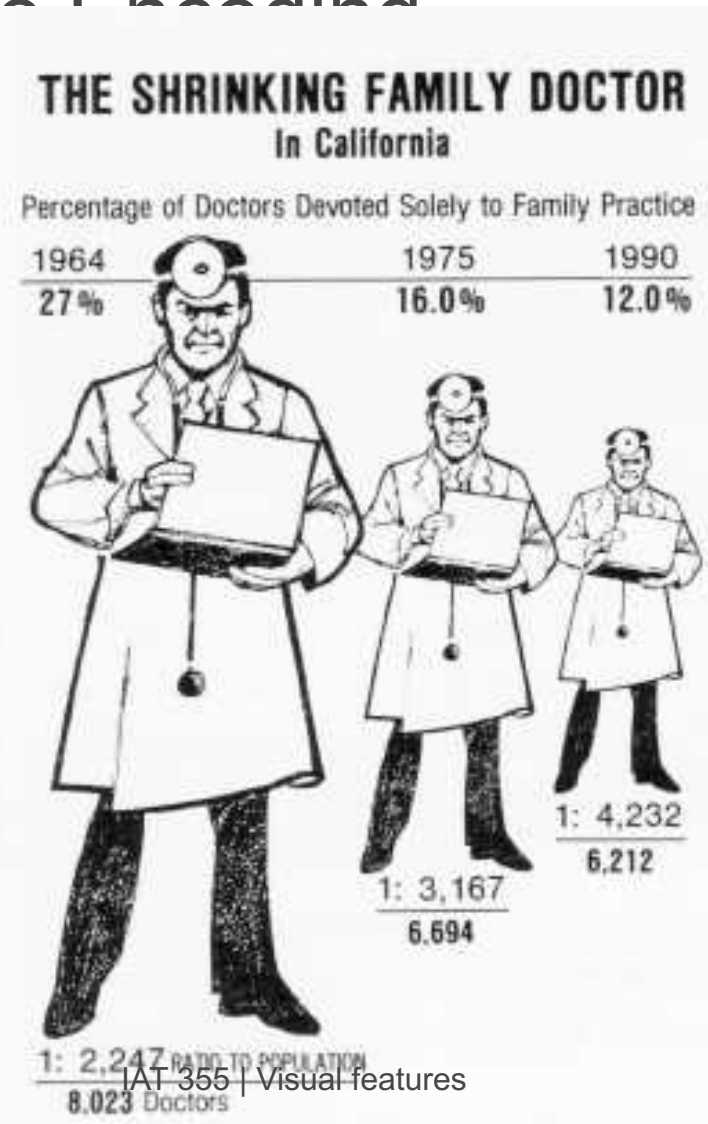
- The representation of numbers, as physically measured on the surface of the graphics, should be directly proportional to the numerical quantities represented

- Clear, detailed and thorough **labeling** should be used to defeat **distortion**

Size Encoding

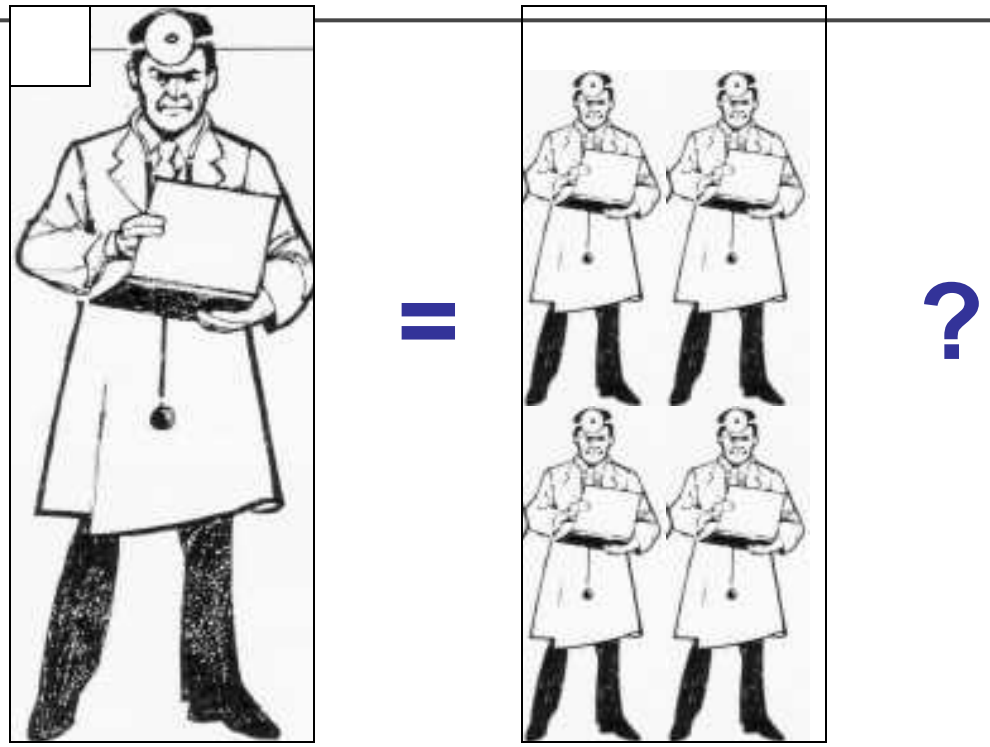


Size Encoding



Los Angeles Times, August 5, 1979, p. 3.

Size Encoding: height or area?



Height or Area

- Height = value
  Width = value
  Area = value^2

  or

- Area = value
  height*width = value
  height = width = value $^{0.5}$

Problem:
Using 2 dimensions to represent 1 dimension.

Height = value

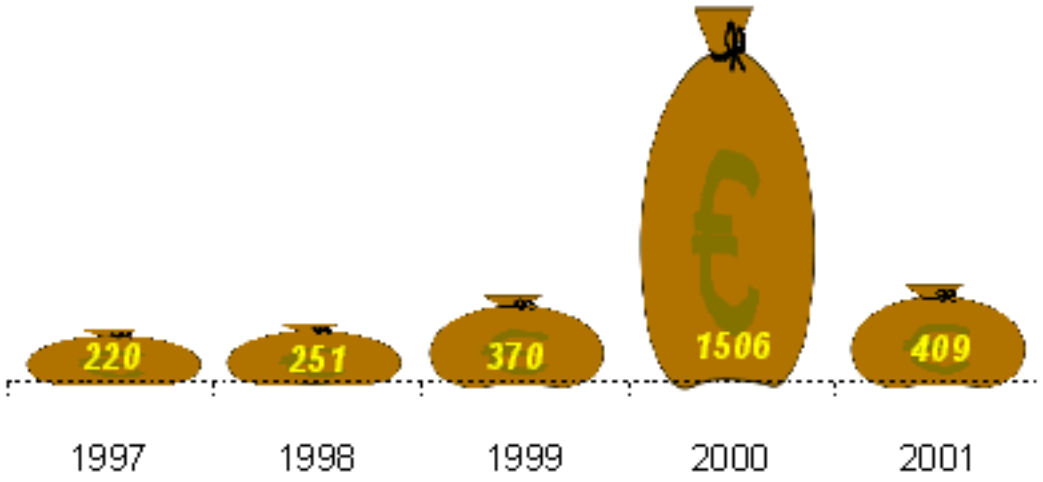


Chart 15: Soneras result drawn with moneybags as bars

With radius=value

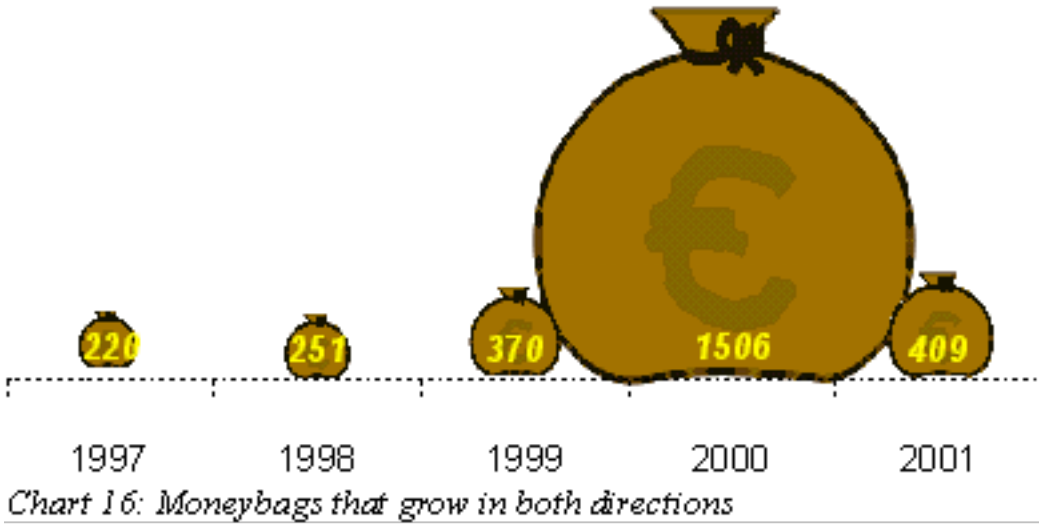


Chart 16: Moneybags that grow in both directions

Area = value

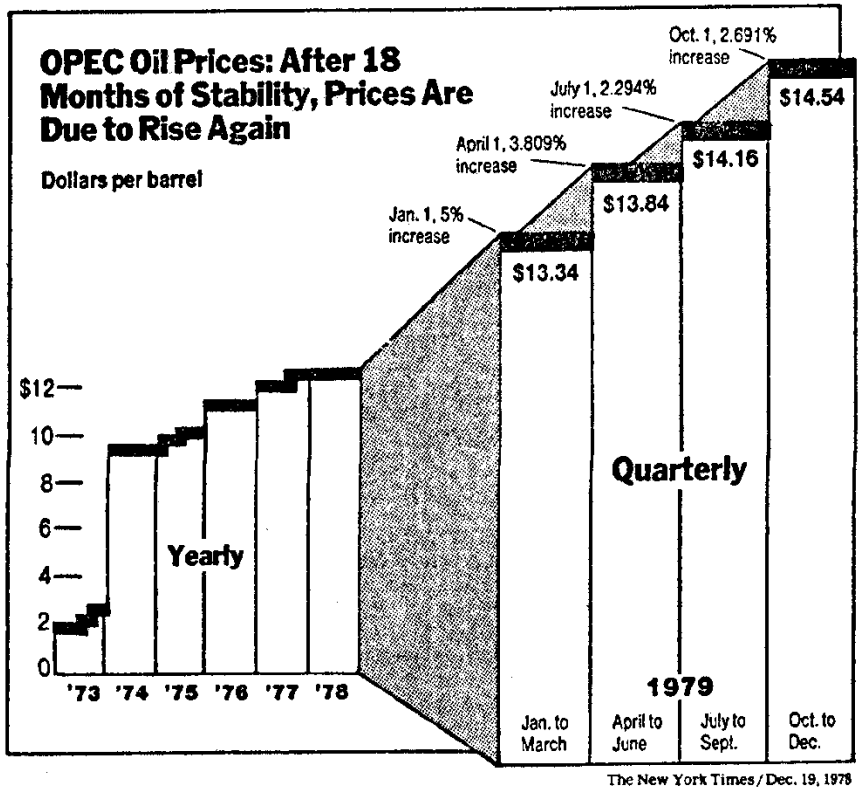
Chart 17: True 2-dimensional moneybags

Design and Data Variation

- Show data variation, not design variation

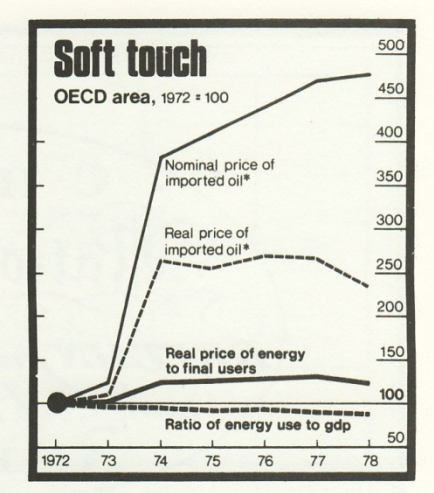
• 1973-1978: one vertical inch equals to $8.00. In 1979, One vertical inch equals $3-4

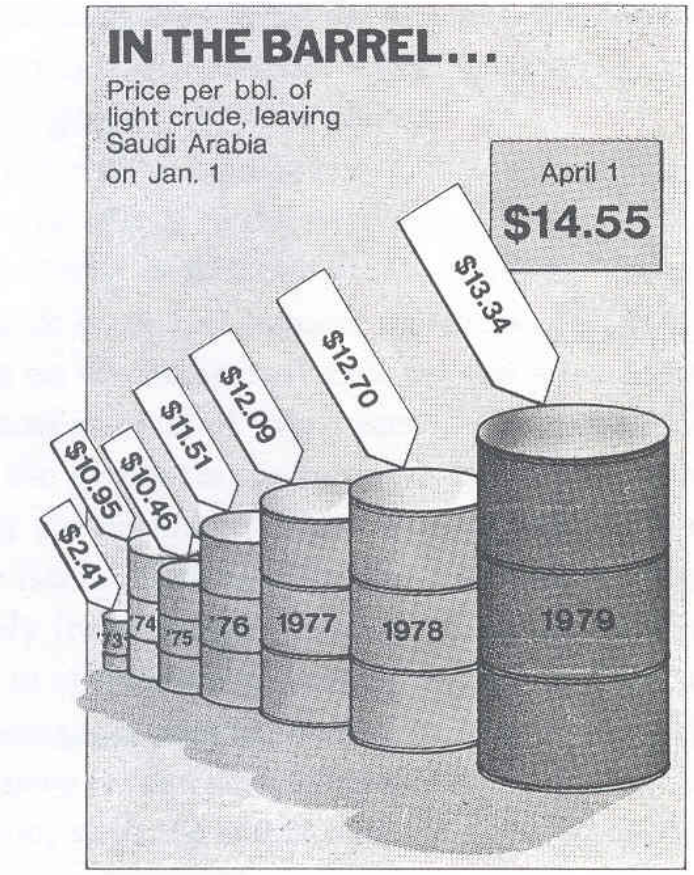
• 1973-1978: one horzontal inch equals 3.7 years, while 1979 equals 0.57 year



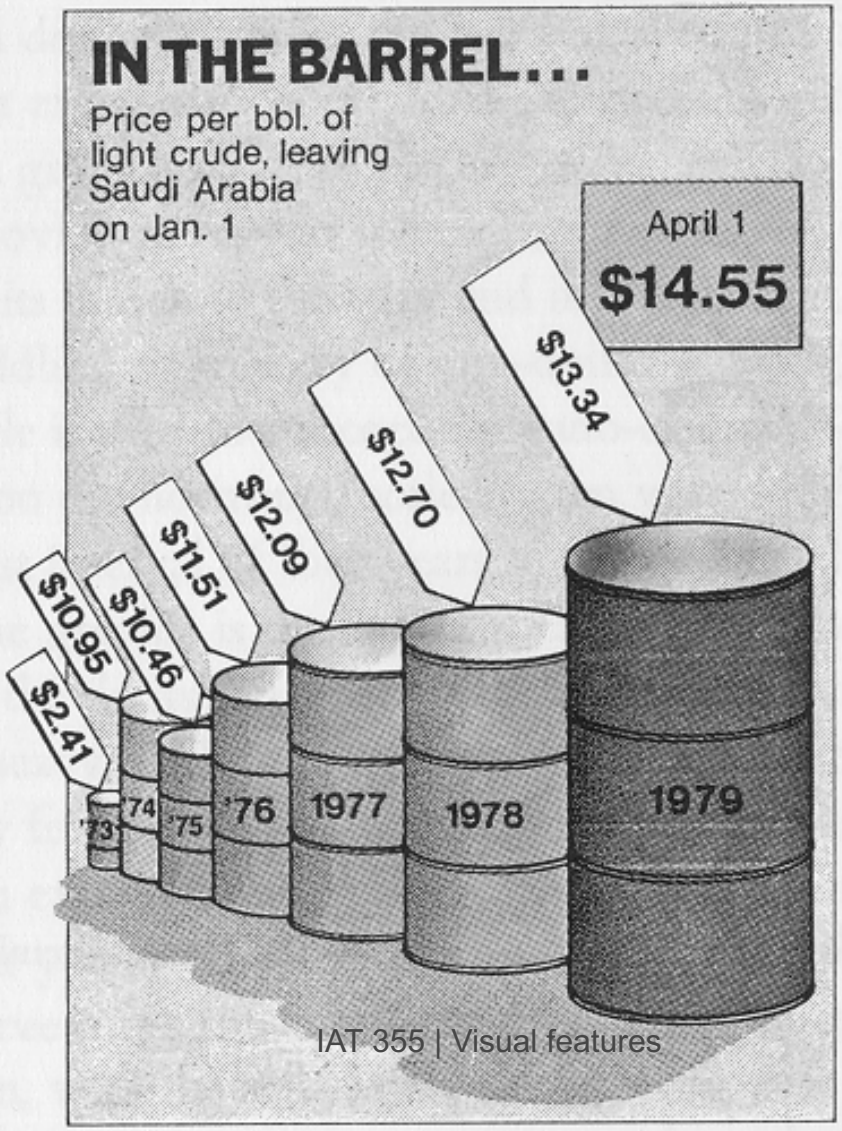
Example

- Lie factor: 9.5
- The price of oil is inflated so needs to be repaired.





Volume encoding?

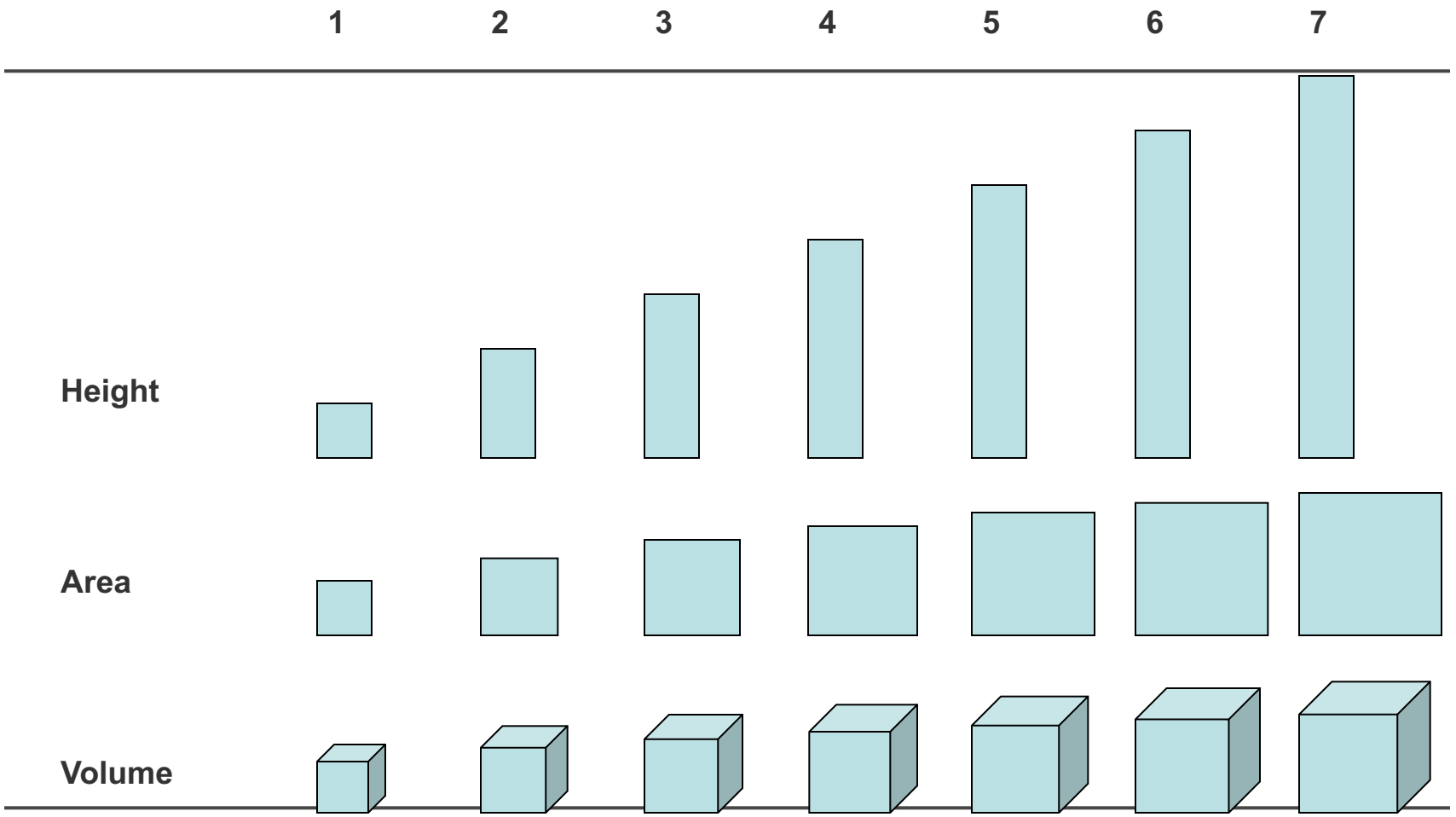


Time, April 9, 1979, p. 57.

Height?
Diameter?
Surface area?
Volume?

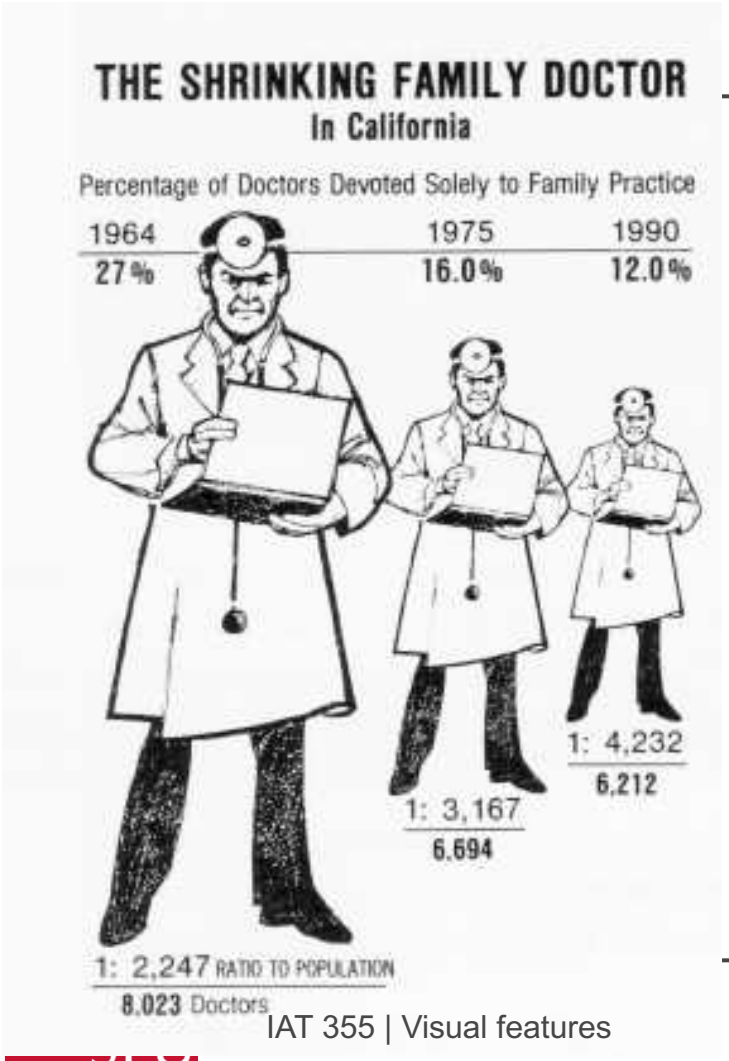
73 – 79 data difference = 5.5x
73 – 79 volume difference = 270x

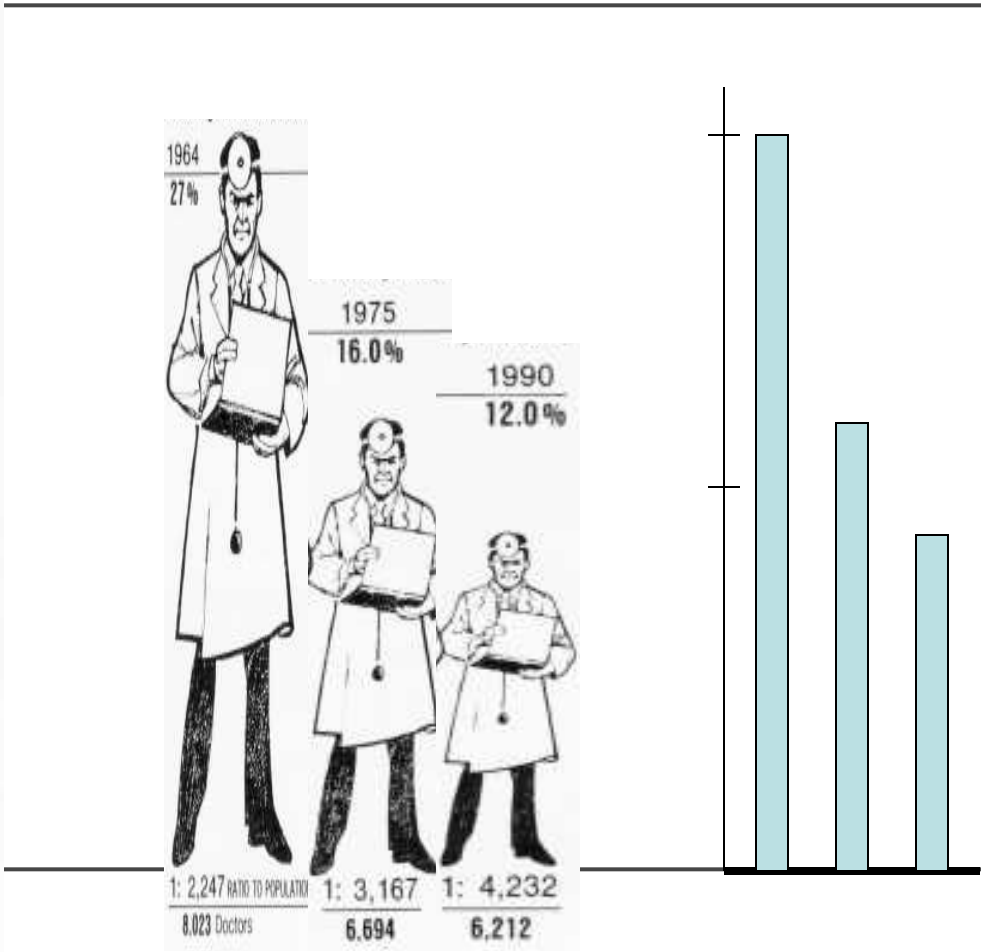
Problem with size encoding



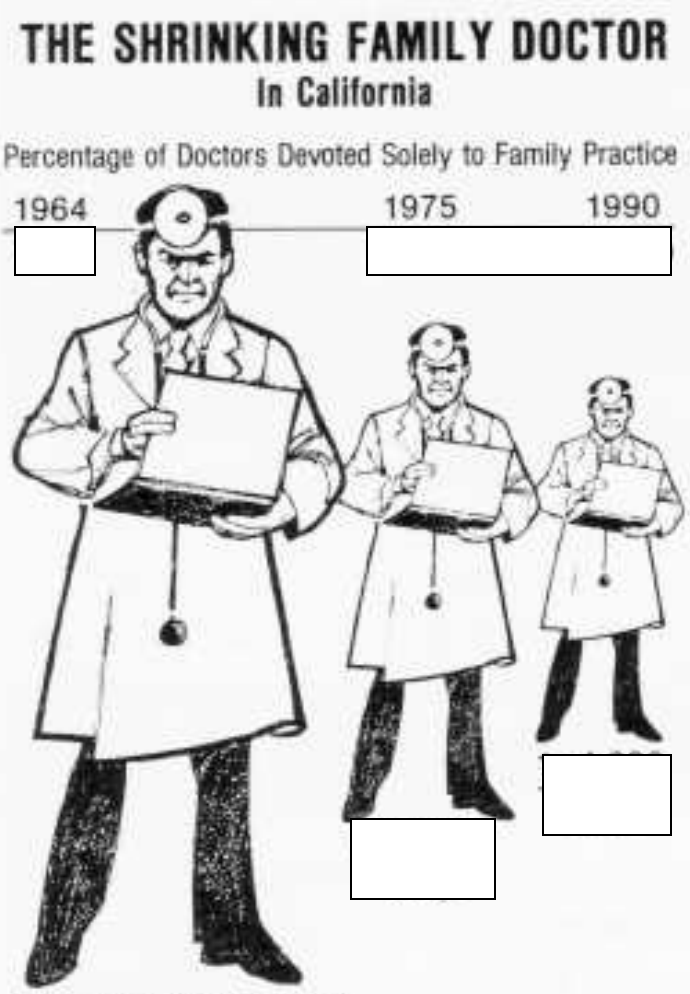
Height & width encoding

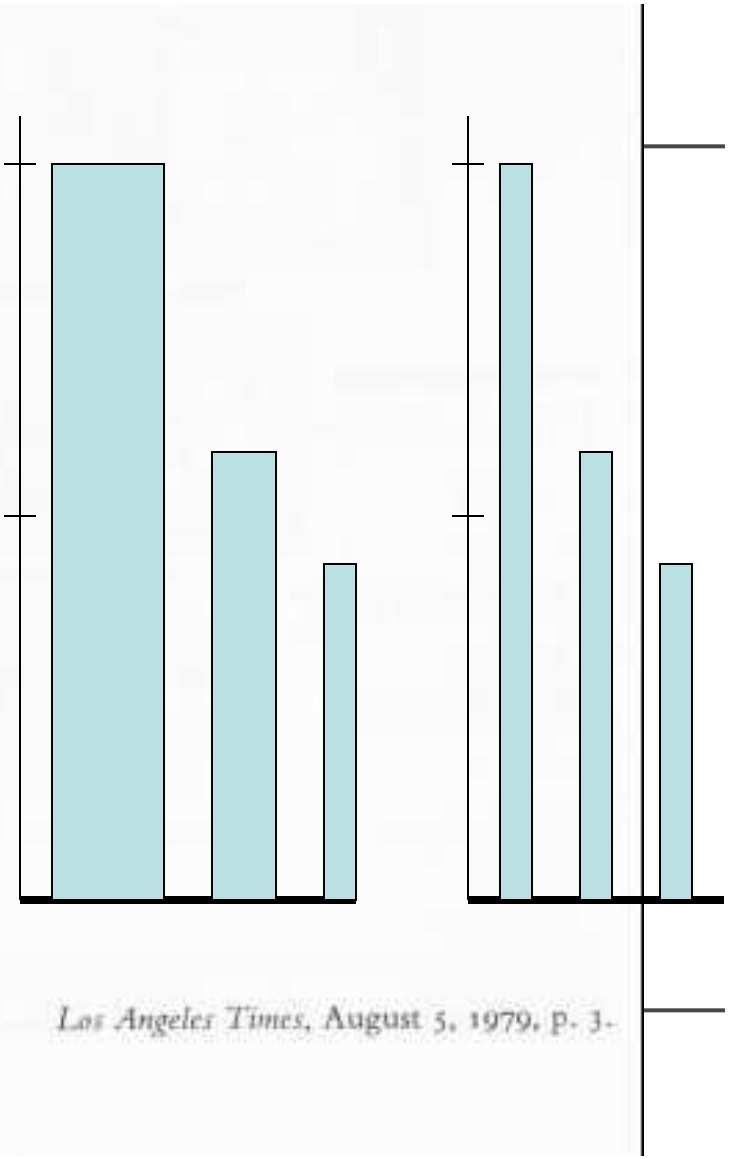
Solution: just use height





Size Encoding

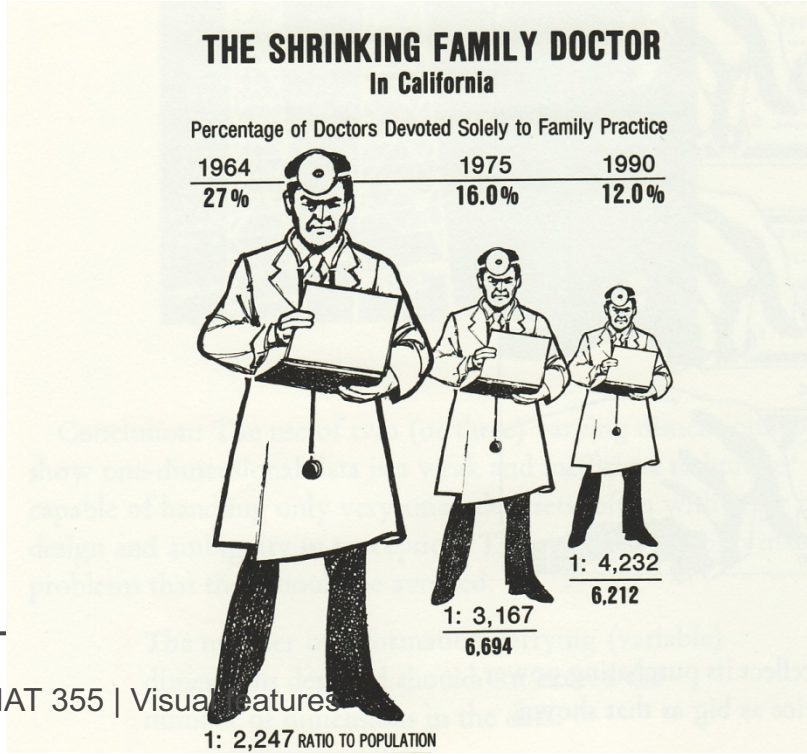




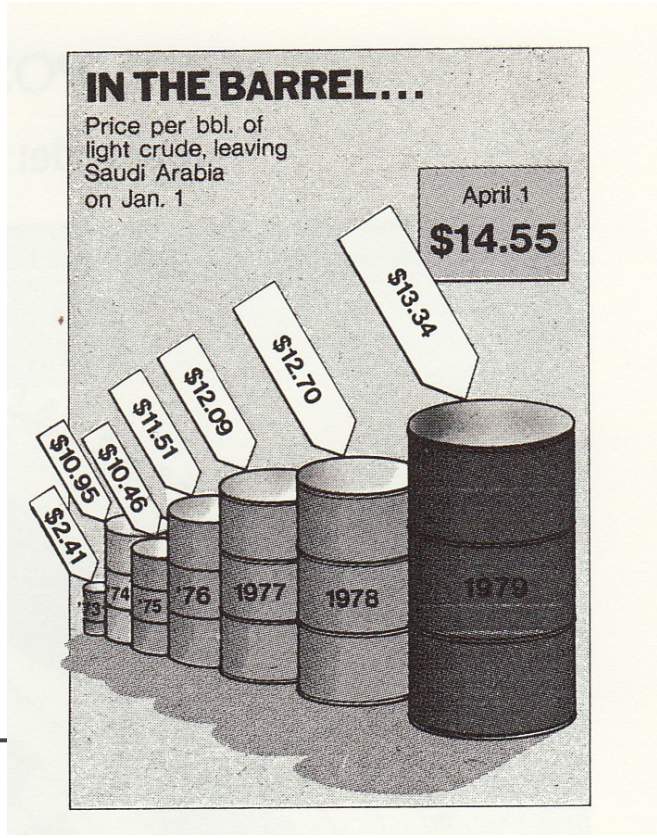
Los Angeles Times, August 5, 1979, p. 3.

Visual Area and Numerical Measure

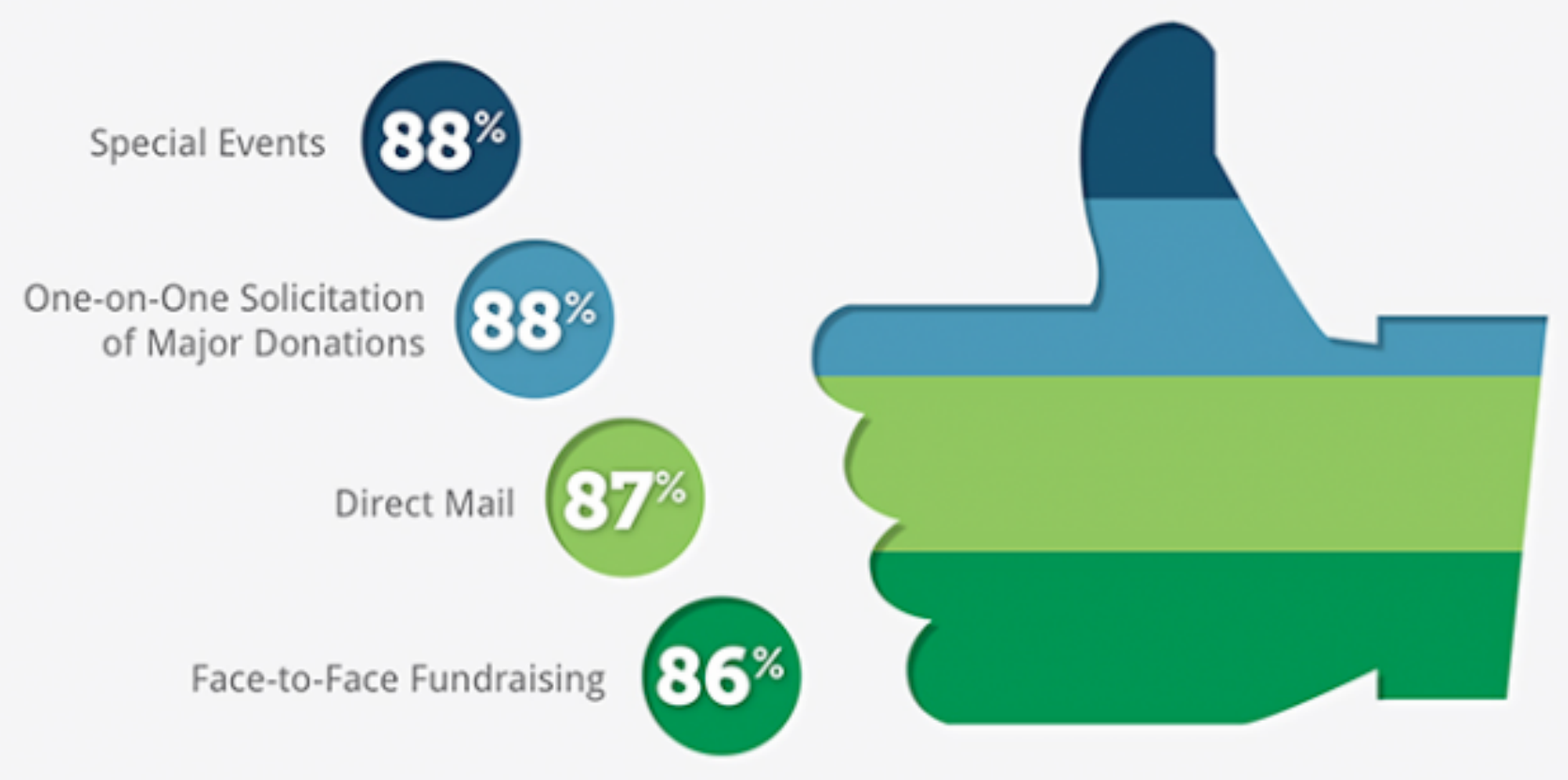
- Tricking the reviewer with design variation is to use areas to show 1D data



Lie factor: 2.8



What's being used here?

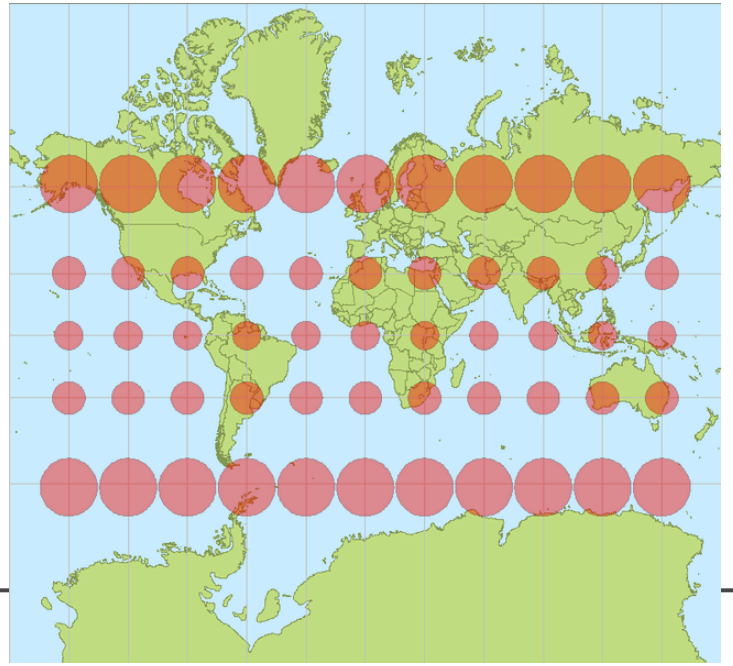


Something more familiar

Mercator projection

- Most maps use a version of the Mercator projection
- Designed for navigators to represent lines of constant course as straight segments

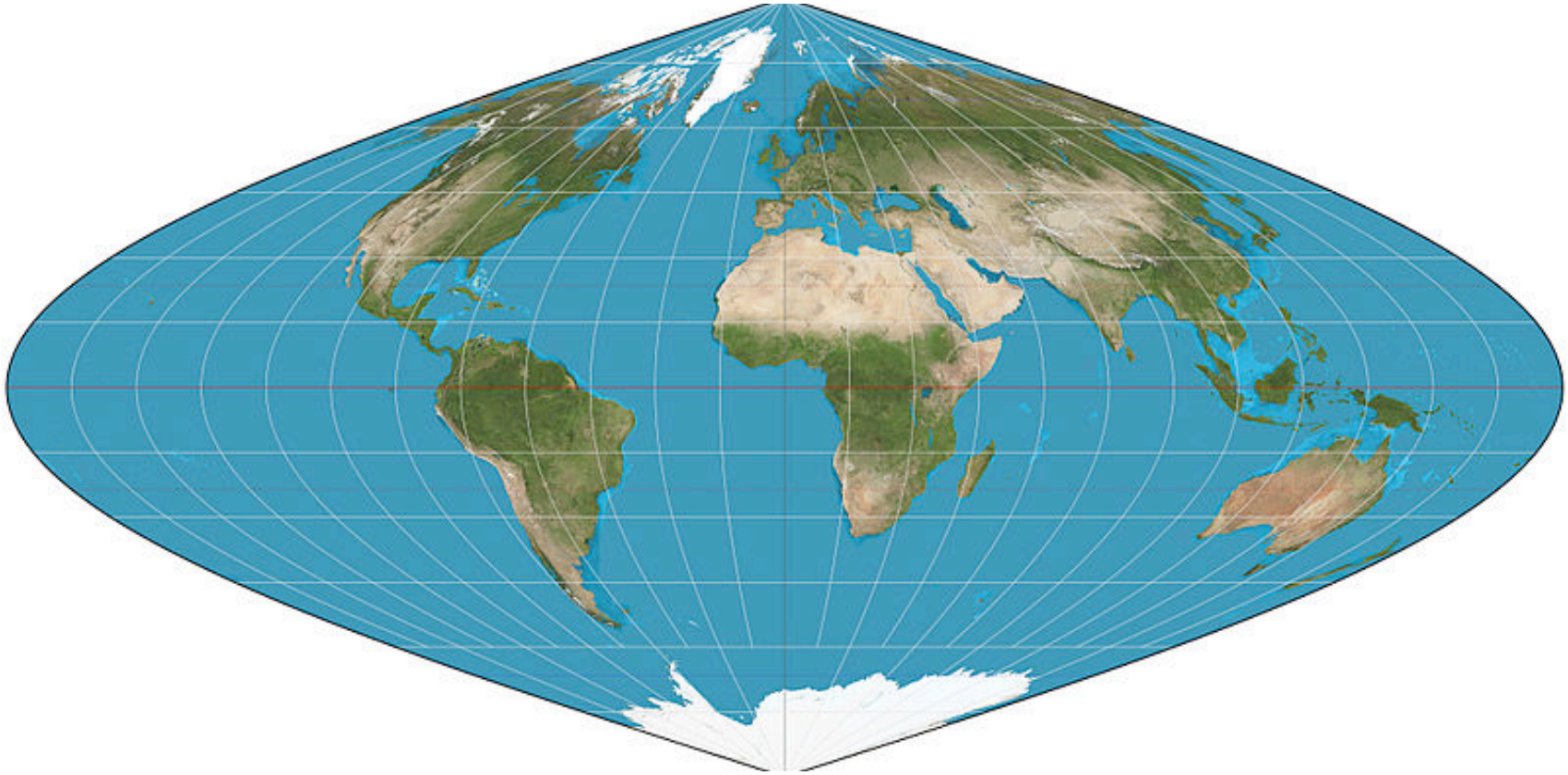
- So what's wrong?



Projections distort

- Greenland takes as much space on the map as Africa, when in reality Africa's area is 14 times greater and Greenland's is comparable to Algeria's alone.
- Alaska takes as much area on the map as Brazil, when Brazil's area is nearly five times that of Alaska.
- Finland appears with a greater north-south extent than India, although India's is greater.
- Antarctica appears as the biggest continent, although it is actually the fifth in terms of area.

More accurate … but unfamiliar!



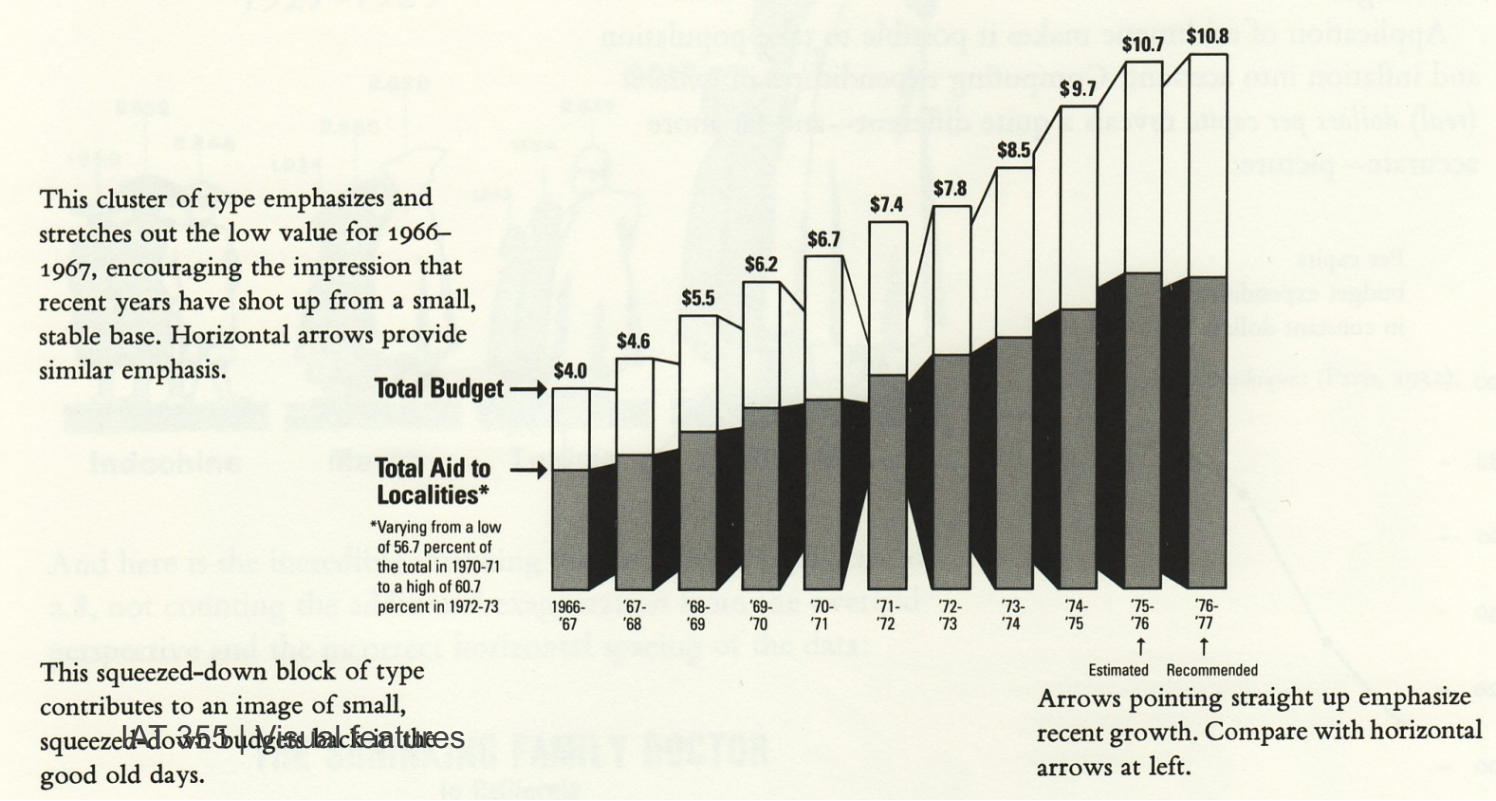
And more

SFU

SCHOOL OF INTERACTIVE
ARTS + TECHNOLOGY

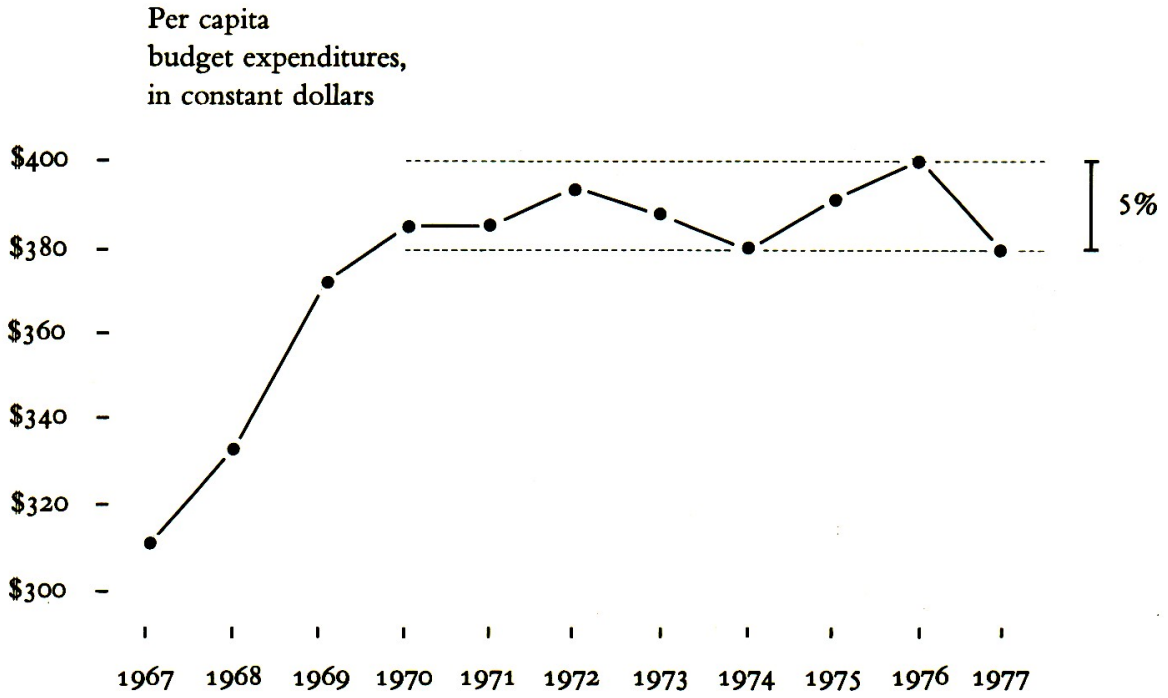
Government Spending

- Tricks to exaggerate the growth of spending

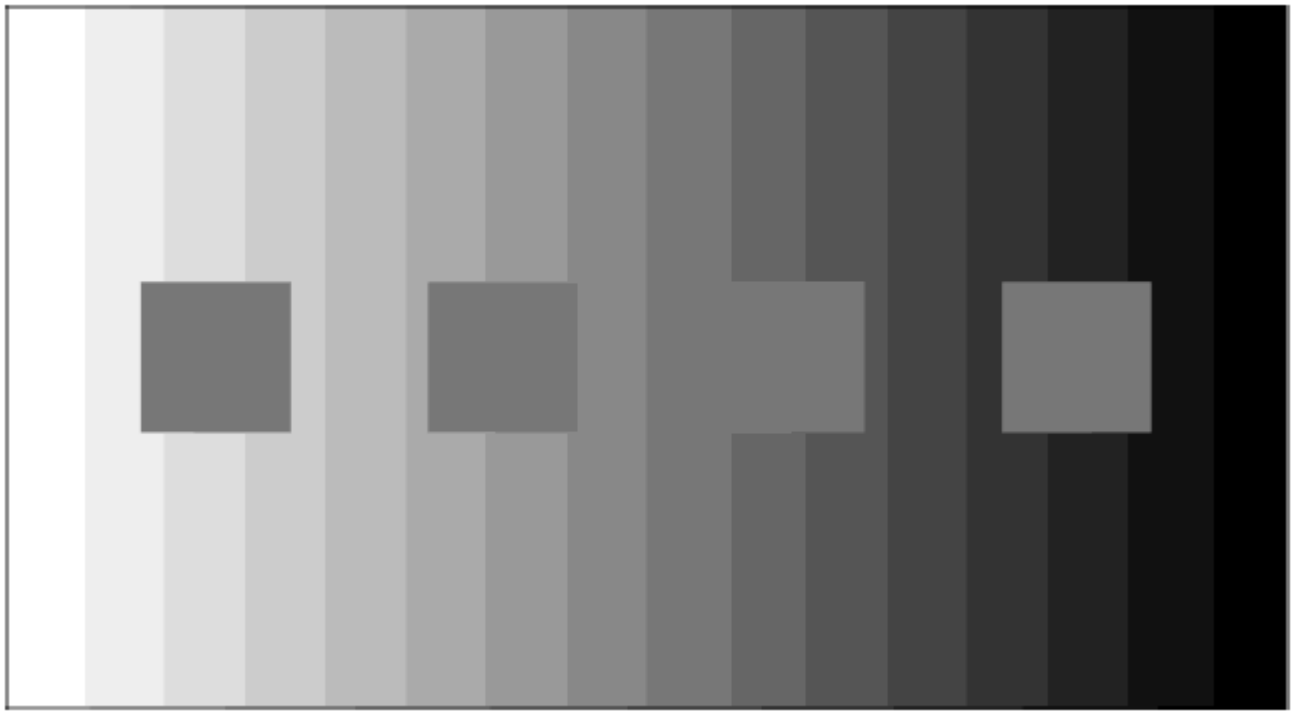


Real Government Spending

- Tricks to exaggerate the growth of spending



Remember the perceptual issues of colour



Contrast rules

- Rule 1: if you want objects of the same colour to look the same, make sure the (surrounding) background is consistent.
- Rule 2: if you want them to be easily see, use a background colour that contrasts sufficiently.

Be careful about implied meaning

- Certain representations impute meanings

- Cultural
- perceptual

Verbal-visual conflict: the Stroop effect

Space

Spatial Channels

Values

- Express

Regions

- Separate
- order
- Align
  - 1D
  - 2D

Given Use

- Geographic
- Fields
  - scalar

Spatial Layouts

- Parallel
- Rectilinear
- Radial
- Spacefilling
- Dense

Space

Spatial Channels

Values

- Express

Regions

- Separate
- order
- Align
  - 1D
  - 2D

Given Use

- Geographic
- Fields
  - scalar

Spatial Layouts

- Parallel
- Rectilinear
- Radial
- Spacefilling
- Dense