

CSE 564

VISUALIZATION AND VISUAL ANALYTICS

VISUAL DESIGN & AESTHETICS

**KLAUS MUELLER**

COMPUTER SCIENCE DEPARTMENT  
STONY BROOK UNIVERSITY

Lecture	Topic	Projects
1	Intro, schedule, and logistics	
2	Applications of visual analytics, basic tasks, data types	
3	Data sources and introduction to D3	Project #1 out
4	Data wrangling and preparation	
5	Bias in visualization	
6	Data reduction and dimension reduction	
7	Visual perception and cognition	
8	Visual design and aesthetics	Project #1 due, Project #2 out
9	Python/Flask hands-on	
10	Cluster analysis: numerical data	
11	Cluster analysis: categorical data	
12	Foundations of scientific and medical visualization	
13	Computer graphics and volume rendering	Project #2 due / Project #3 out
14	Scientific and medical visualization	
15	Illustrative rendering	Project #3 due
16	High-dimensional data, dimensionality reduction	Final project proposal call out
17	Correlation visualization	
18	Principles of interaction	
19	Midterm #1	
20	Visual analytics and the visual sense making process	Final project proposal due
21	Evaluation and user studies	
22	Visualization of time-varying and time-series data	
23	Visualization of streaming data	
24	Visualization of graph data	Final Project preliminary report due
25	Visualization of text data	
26	Midterm #2	
27	Data journalism	
	Final project presentations	Final Project slides and final report due

# THREE KEY VISUAL REPRESENTATIONS

## Gestalt Principles:

- the tendency to perceive elements as belonging to a group, based on certain visual properties

## Pre-attentiveness:

- certain low level visual aspects are recognized before conscious awareness

## Visual variables:

- the different visual aspects that can be used to encode information

# GESTALT

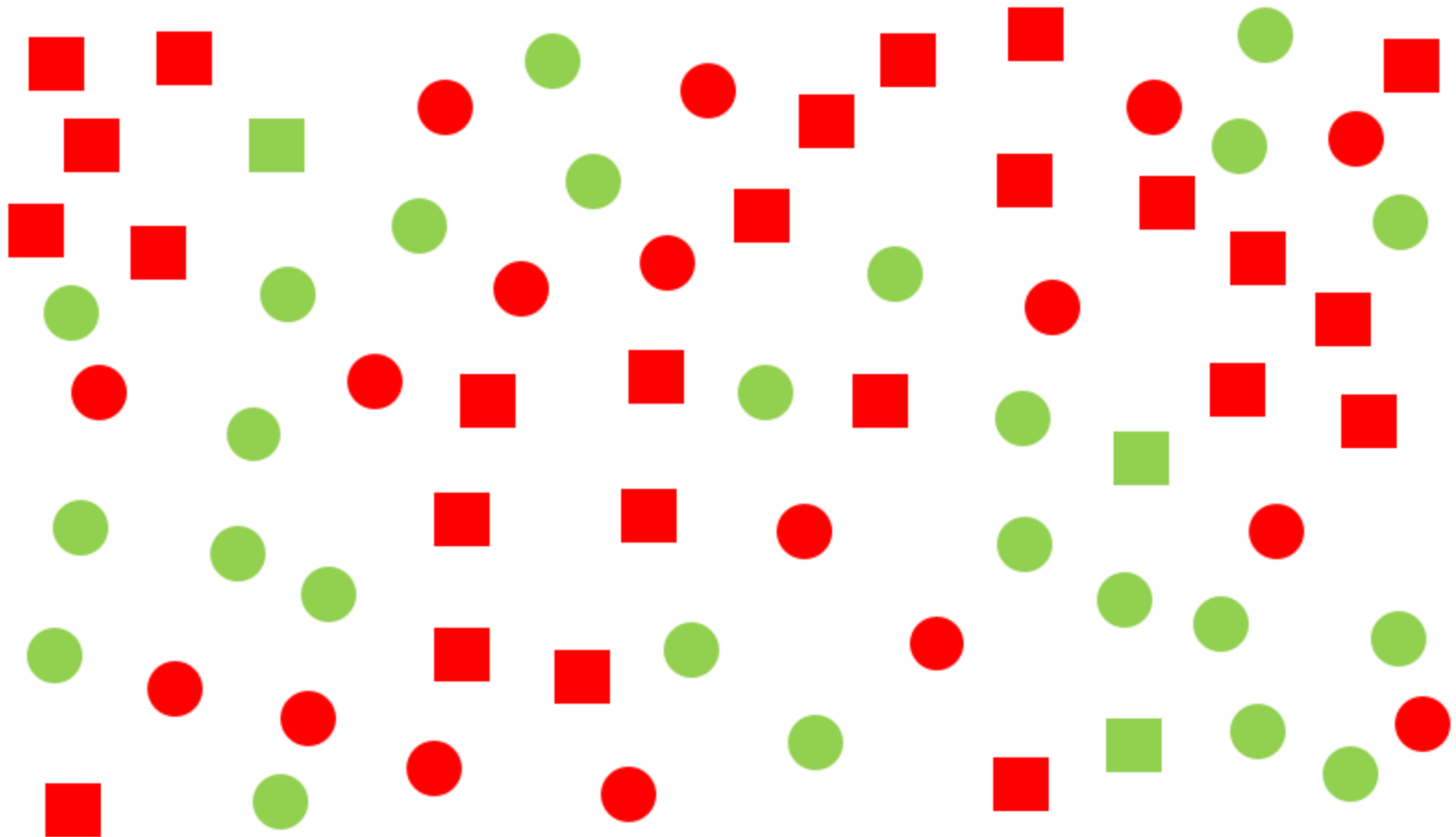
## Concept of totality

- you grasp the "totality" of something before worrying about the details



# PRE-ATTENTIVENESS

Also called pop-out (multiple conjunctions shown here):



# WHICH POPPED-OUT FASTER

Color (red vs. green)

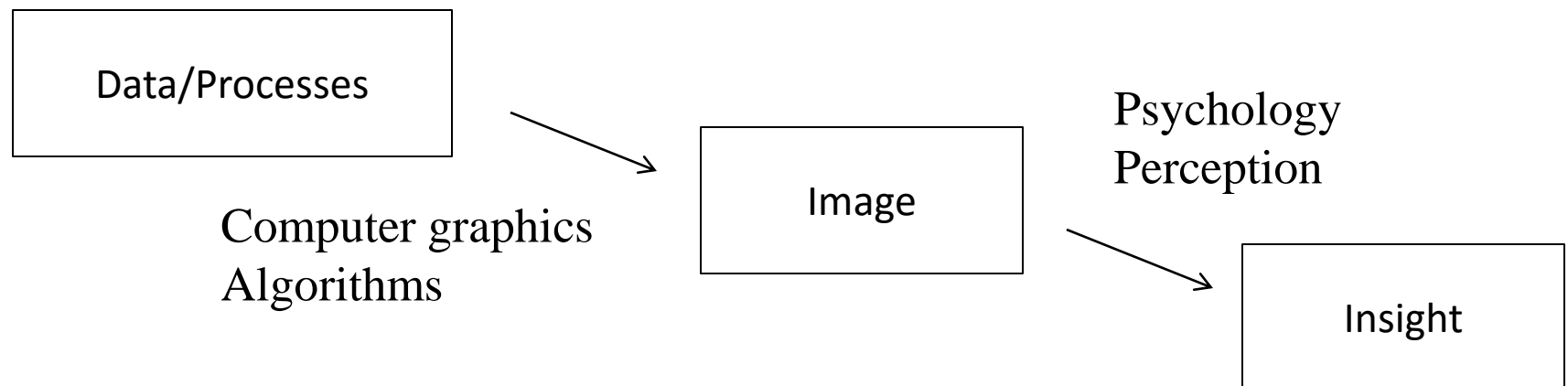
Shape (circle vs. square)

# VISUAL VARIABLES

Formal theory linking perception to visualization

Established by Jacques Bertin (1967)

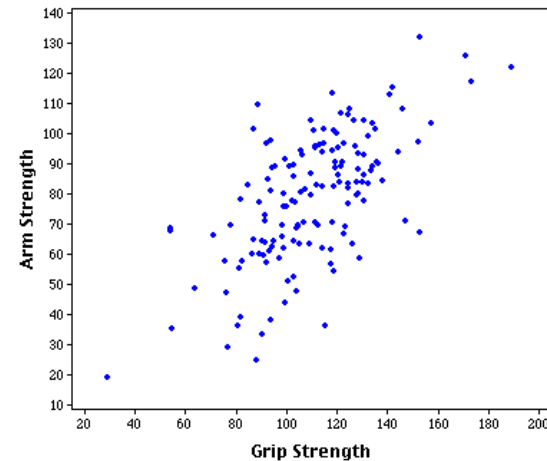
- he called it 'Image Theory'
- original book in French (*Sémiologie Graphique*) translated into English by W. Berg (1983)
- not formally linked to vision research more based on intuition
- but has been shown later by M. Green to be quite accurate



# VISUAL VARIABLES

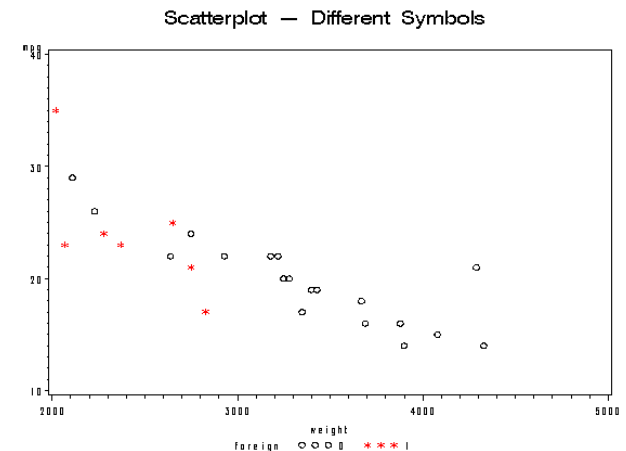
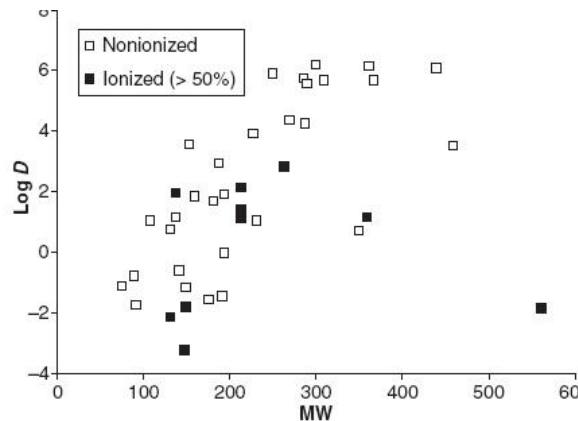
## Two planar variables

- spatial dimensions
- map (arm, grip) to (x,y)



## Six retinal variables

- size
- color
- shape
- orientation
- texture
- brightness



## Retinal variables allow for one more variable to be encoded

- more than three variables will hamper efficient visual search
- recall low decoding speed of conjunctions

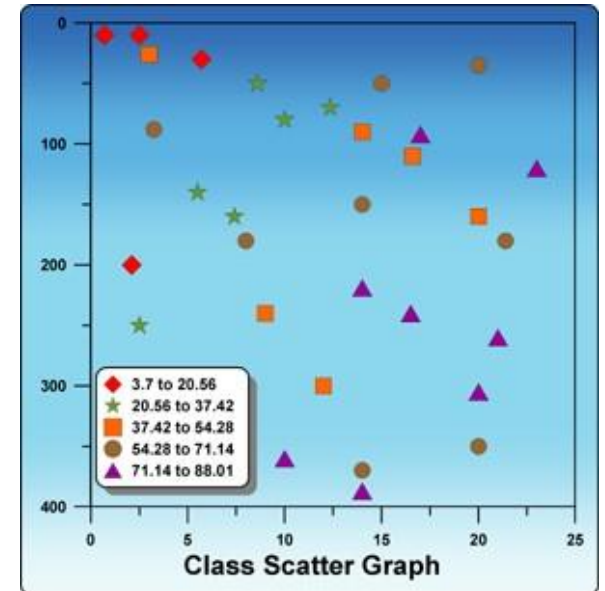


# ASSOCIATIVE VS. SELECTIVE

Both are nominal qualities

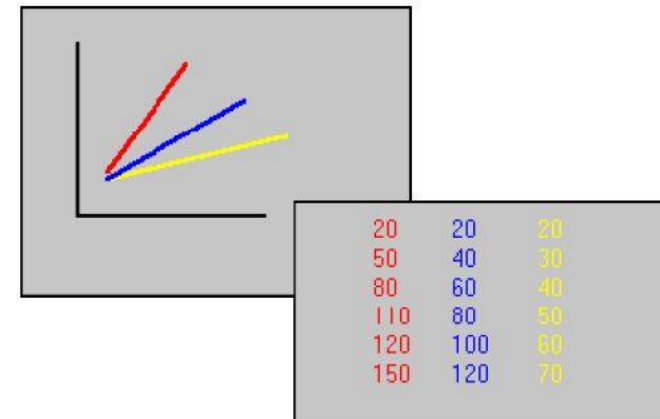
## Associative

- lowest organizational level
- enables **visual grouping** of all elements of a variable



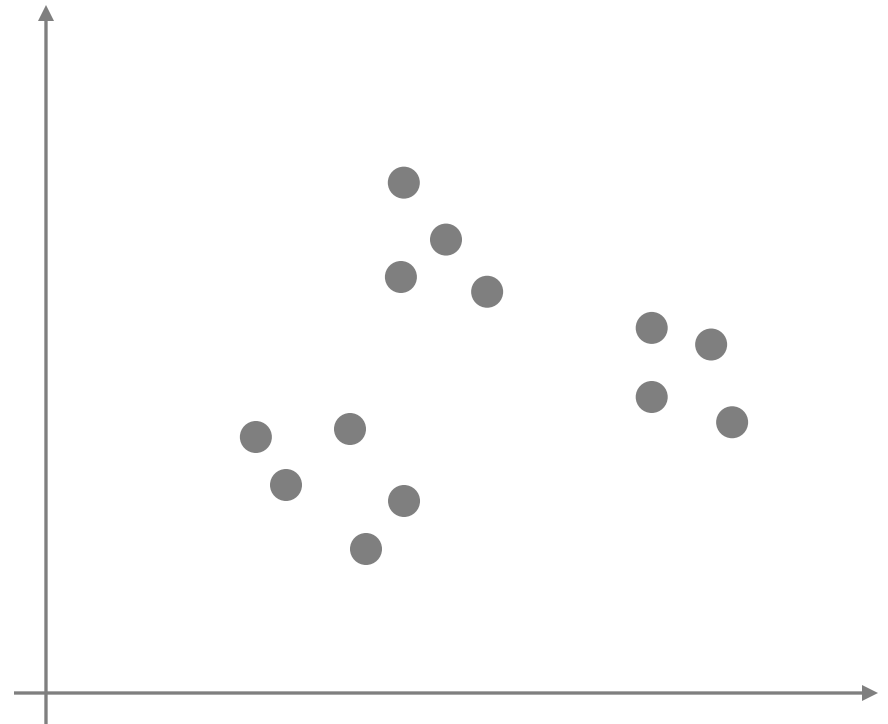
## Selective

- next lowest level
- enables viewer to **isolate encoded data** and ignore others



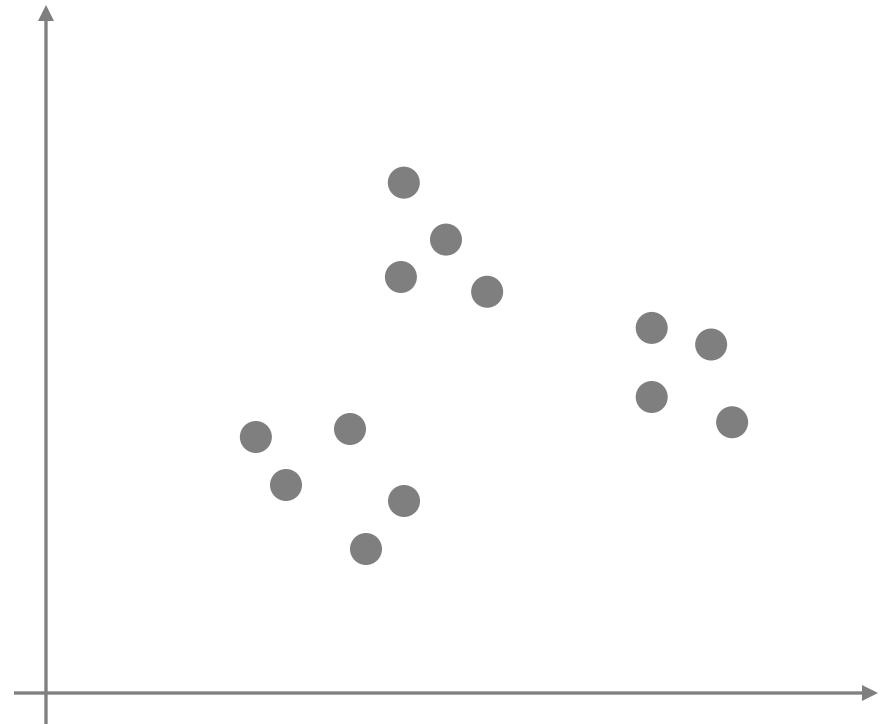
# VISUAL VARIABLE #1 – PLANAR

Visual property	Can convey
Associative	
Selective	
Ordered	
Quantitative	



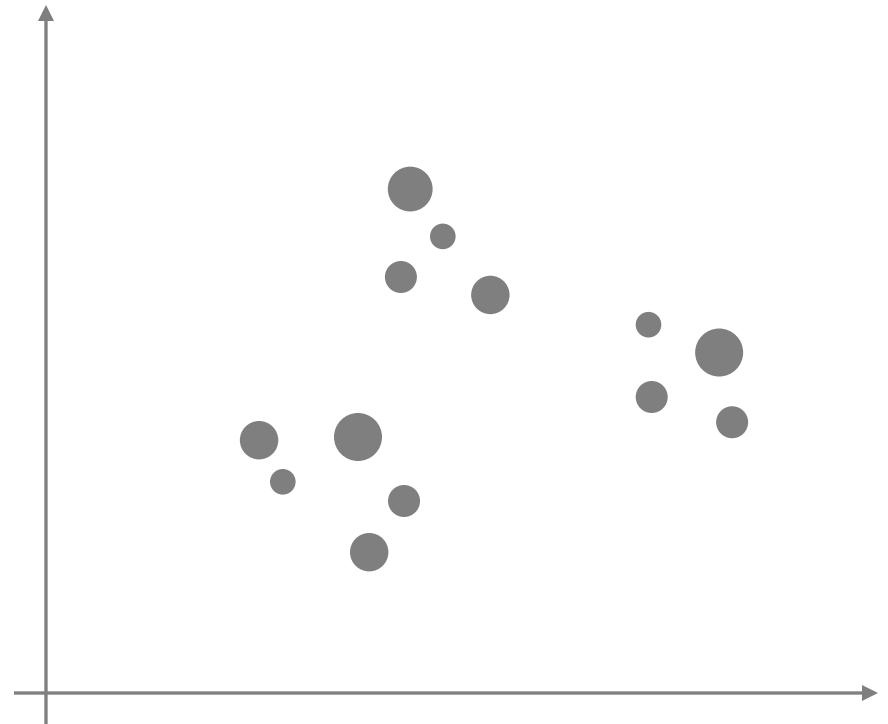
# VISUAL VARIABLE #1 – PLANAR

Visual property	Can convey
Associative	Y
Selective	Y
Ordered	Y
Quantitative	Y



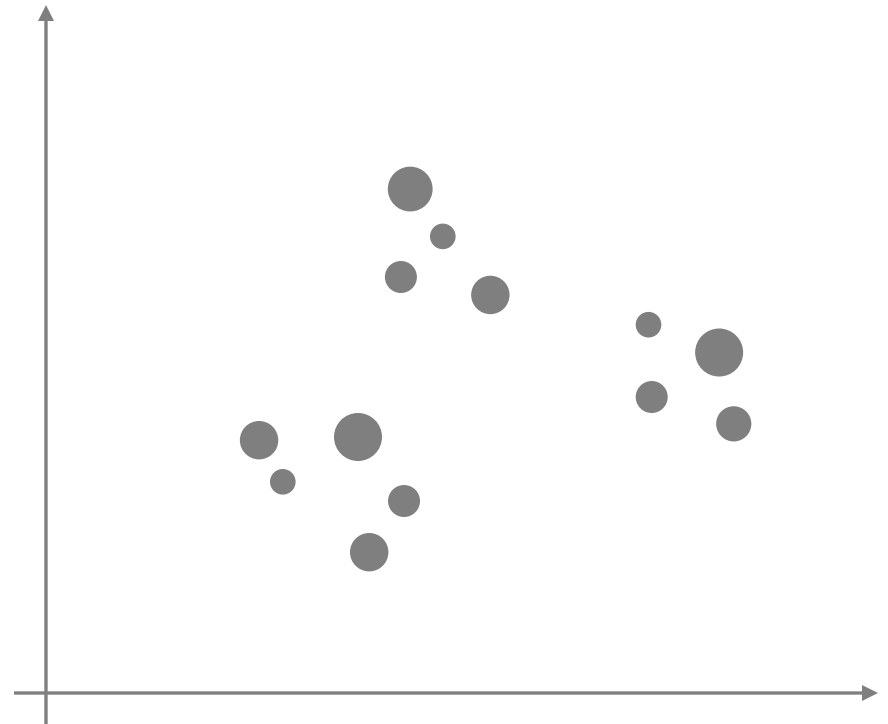
# VISUAL VARIABLE #2 – SIZE

Visual property	Can convey
Associative	
Selective	
Ordered	
Quantitative	



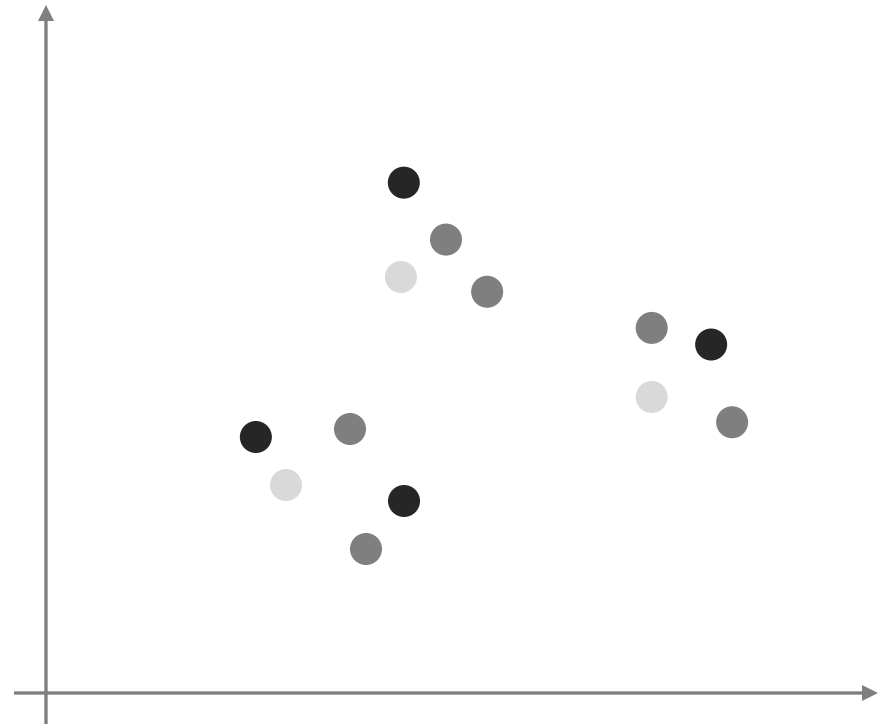
# VISUAL VARIABLE #2 – SIZE

Visual property	Can convey
Associative	Y
Selective	Y
Ordered	Y
Quantitative	(Y)



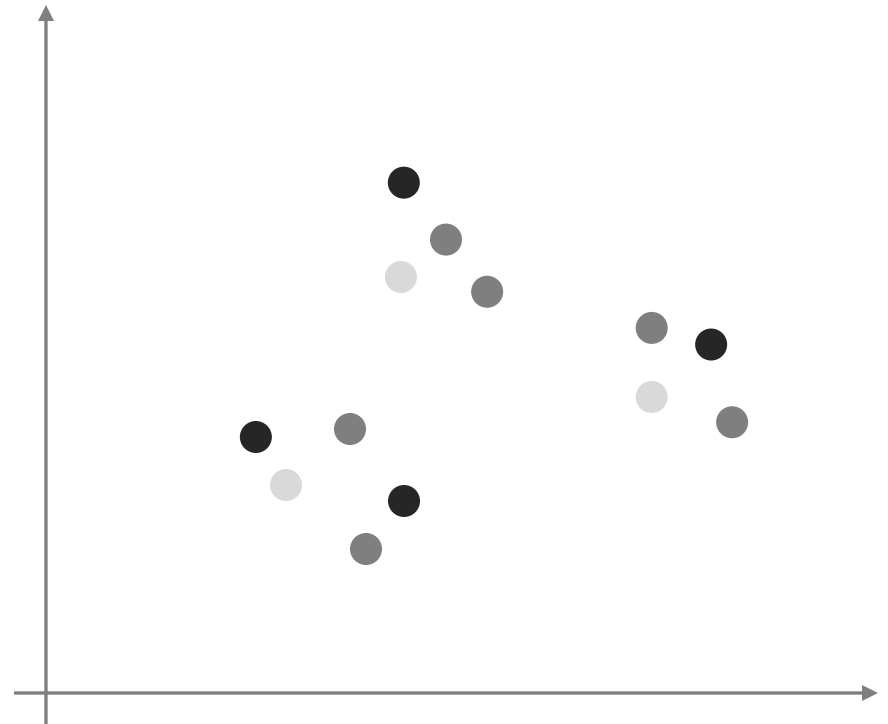
# VISUAL VARIABLE #3 – BRIGHTNESS

Visual property	Can convey
Associative	
Selective	
Ordered	
Quantitative	



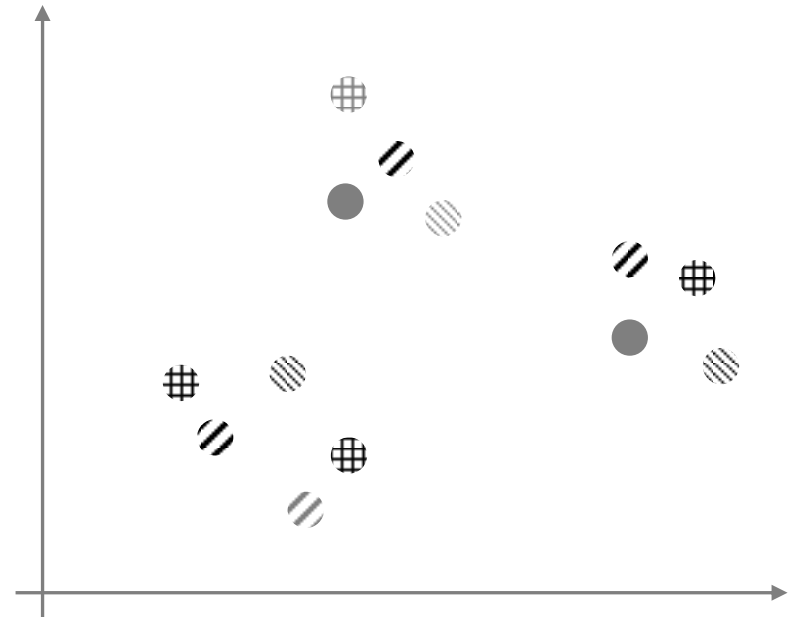
# VISUAL VARIABLE #3 – BRIGHTNESS

Visual property	Can convey
Associative	Y
Selective	Y
Ordered	Y
Quantitative	



# VISUAL VARIABLE #4 – TEXTURE

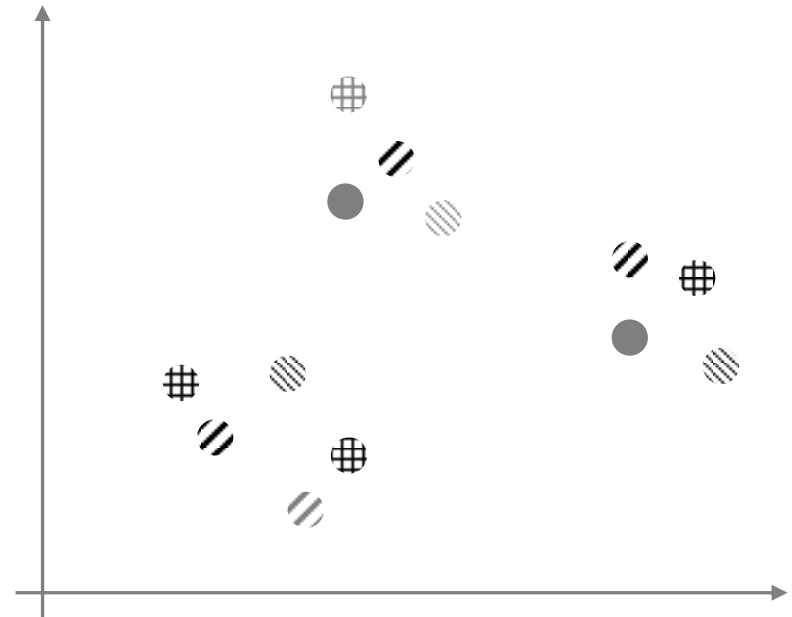
Visual property	Can convey
Associative	
Selective	
Ordered	
Quantitative	





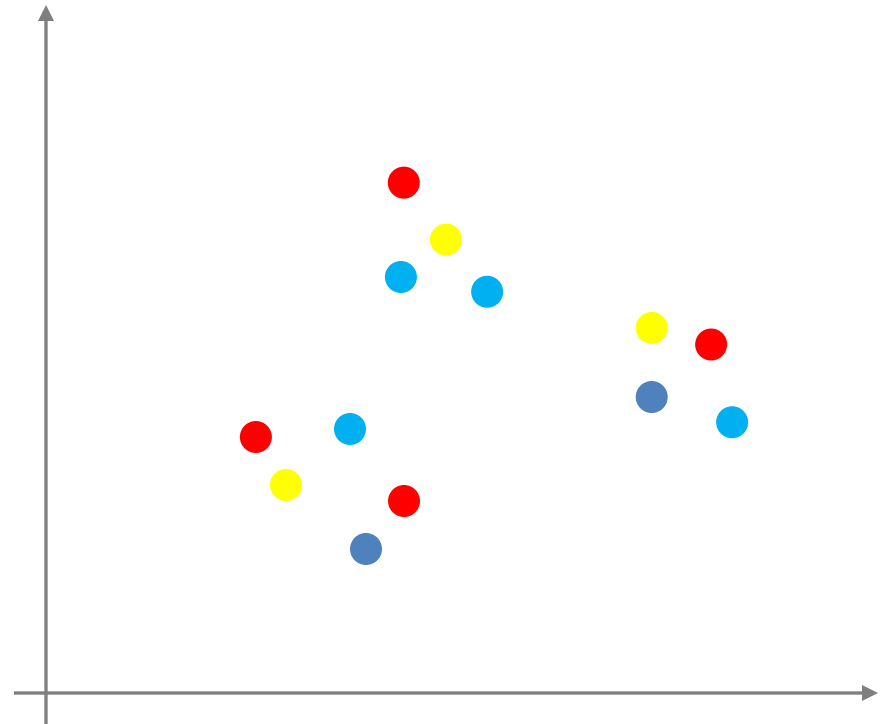
# VISUAL VARIABLE #4 – TEXTURE

Visual property	Can convey
Associative	Y
Selective	Y
Ordered	
Quantitative	



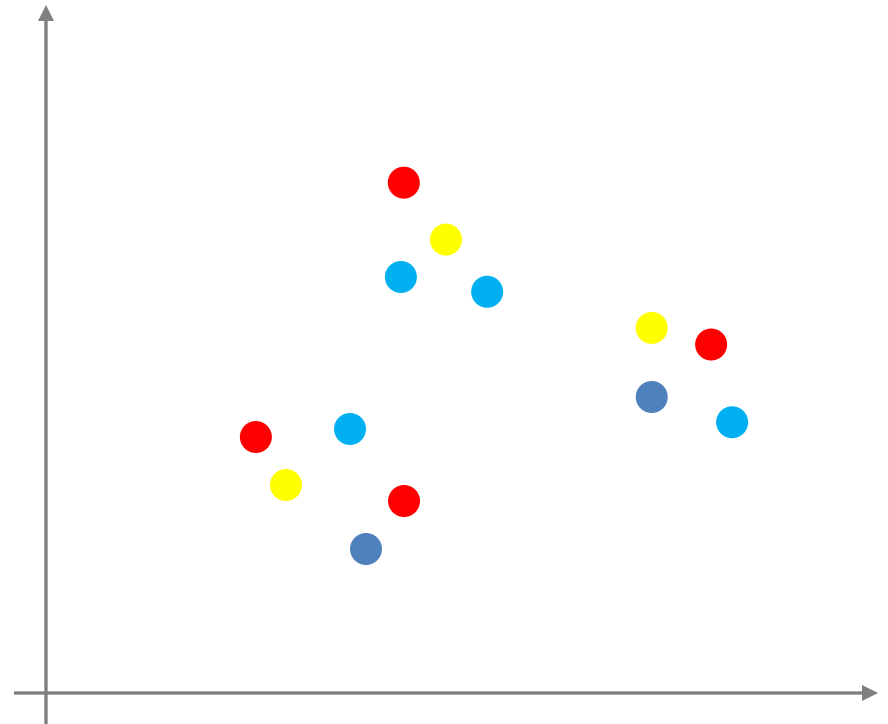
# VISUAL VARIABLE #4 – COLOR

Visual property	Can convey
Associative	
Selective	
Ordered	
Quantitative	



# VISUAL VARIABLE #4 – COLOR

Visual property	Can convey
Associative	Y
Selective	Y
Ordered	
Quantitative	



# VISUAL VARIABLE #5 – ORIENTATION

Visual property	Can convey
Associative	
Selective	
Ordered	
Quantitative	



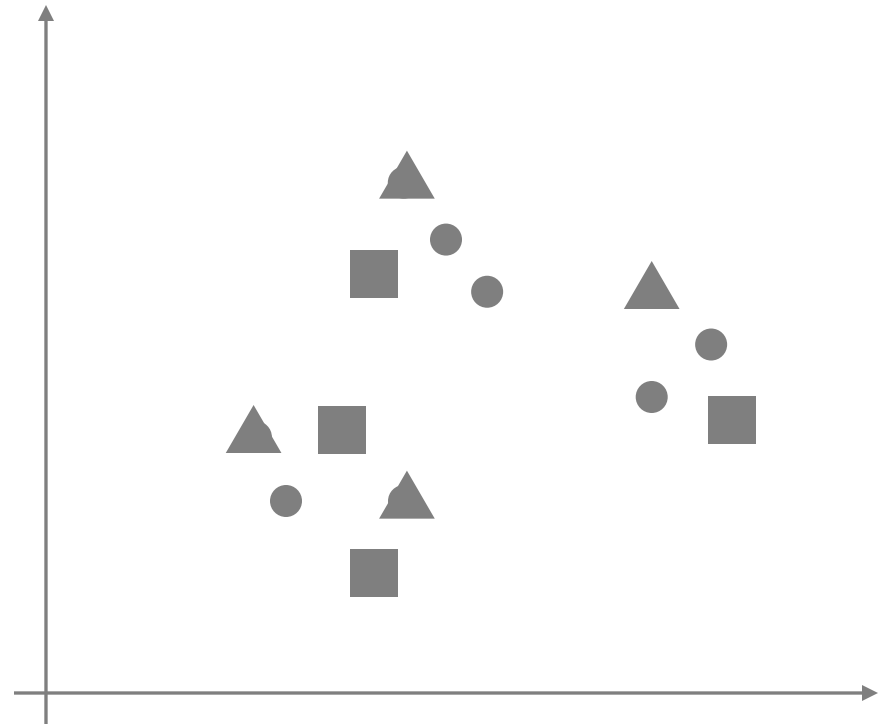
# VISUAL VARIABLE #5 – ORIENTATION

Visual property	Can convey
Associative	(Y)
Selective	(Y)
Ordered	
Quantitative	



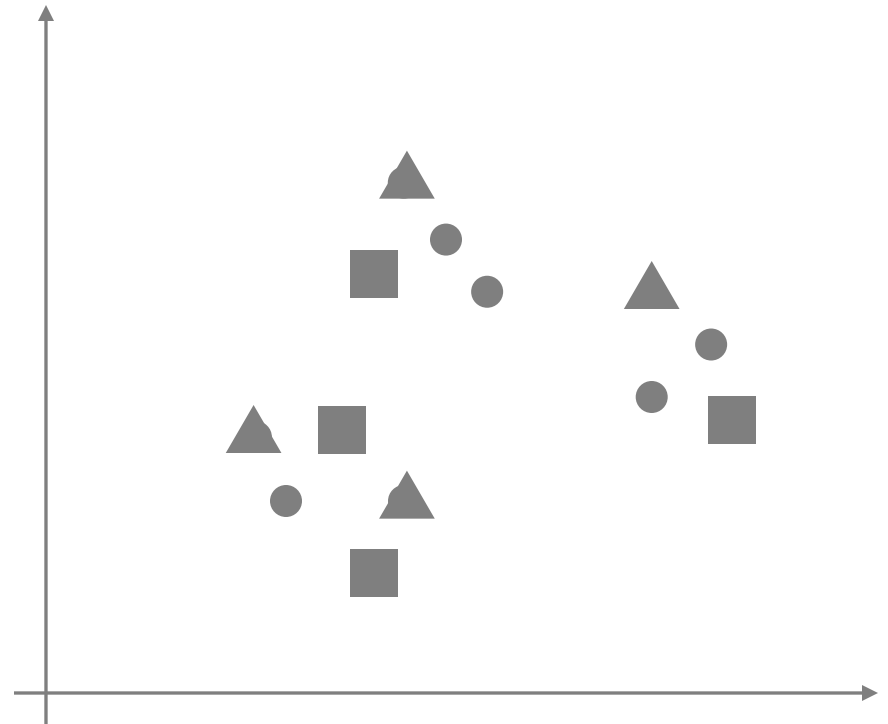
# VISUAL VARIABLE #6 – SHAPE

Visual property	Can convey
Associative	
Selective	
Ordered	
Quantitative	



# VISUAL VARIABLE #6 – SHAPE

Visual property	Can convey
Associative	(Y)
Selective	(Y)
Ordered	
Quantitative	



# LEVELS OF ORGANIZATION

Visual variables differ in what data properties they can convey

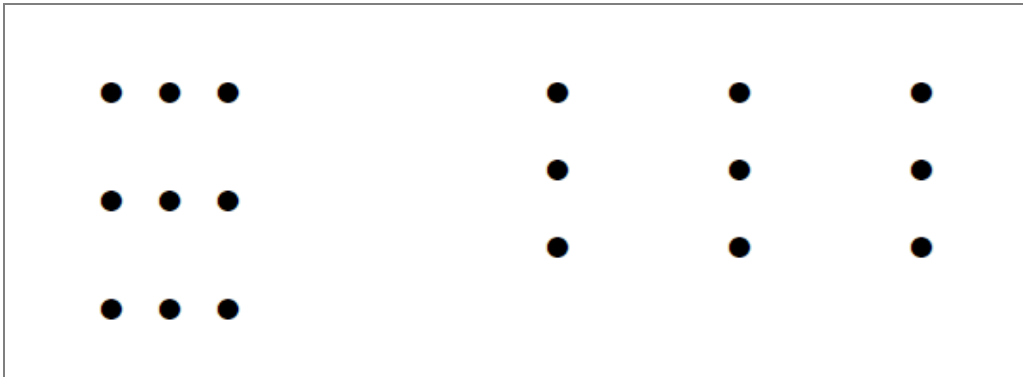
	Associative	Selective	Ordered	Quantitative
Planar	yes	yes	yes	yes
Size	yes	yes	yes	(yes)
Brightness (Value)	yes	yes	yes	
Texture	yes	yes		
Color (Hue)	yes	yes		
Orientation	(yes)	(yes)		
Shape	(yes)	(yes)		



# TAKE-AWAYS (1)

Planar variable is the single most strongest visual variable

- maps to proximity
- provides an intuitive organization of information
- things close together are perceptually grouped together



**TYPICAL WEB FORM**

**Personal Information**

First Name

Last Name

**Contact Information**

Address

City

County

Post Code  Country

| [Cancel](#)

PRIMARY ACTION | SECONDARY ACTION

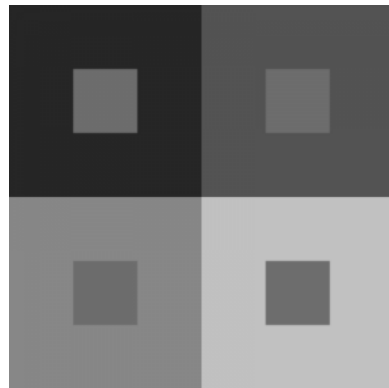
# TAKE-AWAYS (2)

Size and brightness are good secondary visual variables to encode *relative* magnitude

- size appeals to spatial perceptive channels

What are the advantages and disadvantages of brightness

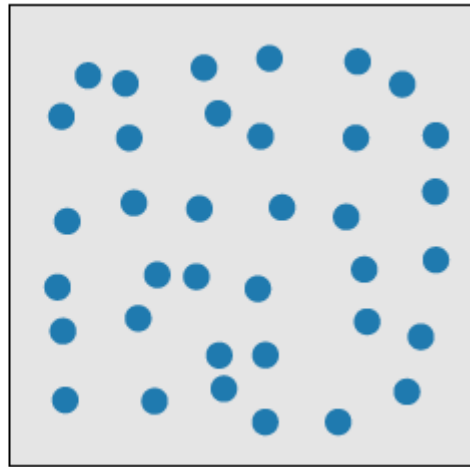
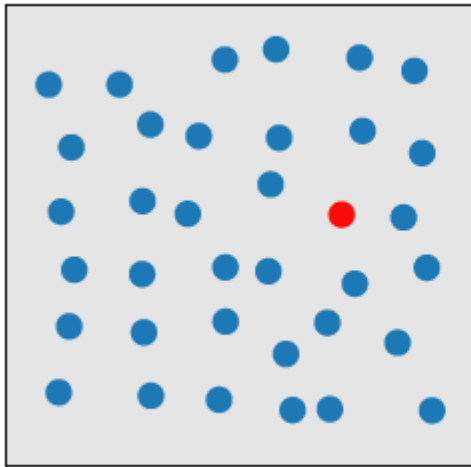
- + brightness does not consume extra space (bigger disks do)
- brightness depends on environmental lighting (size does not)  
where do you view the visualization (office, outdoors, night or day?)



# TAKE-AWAYS (3)

Color is a good visual variable for labeling

- texture can do this as well, but it does not support pop-out much



color pop-out

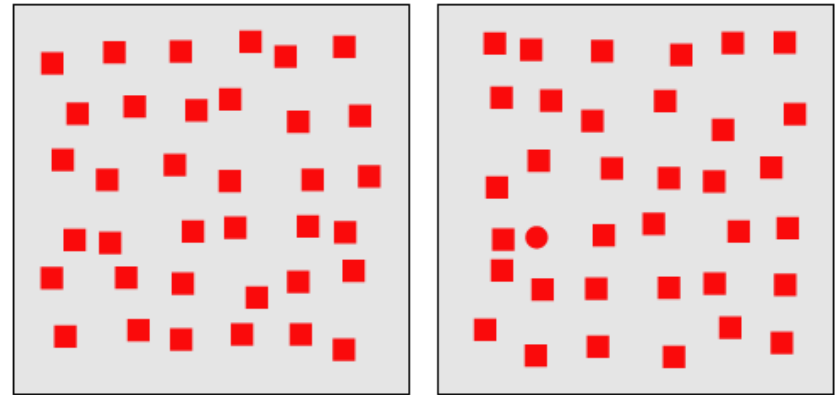


texture pop-out?

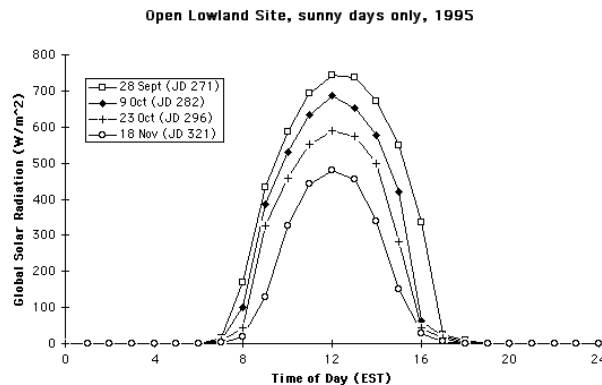


# TAKE AWAYS (4)

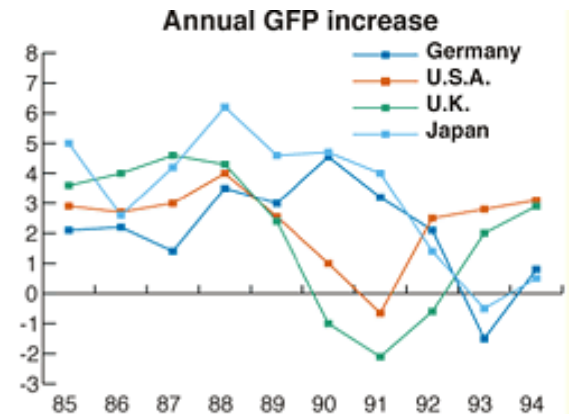
*Shape* provides only limited pop-out



- compare with color pop-out on the previous slide
- another example: coloring of graphs



VS.



# COLOR AND CONTRAST



Background with same-colored object at the same brightness

- can you see the shape?
- can you count the number of gaps?

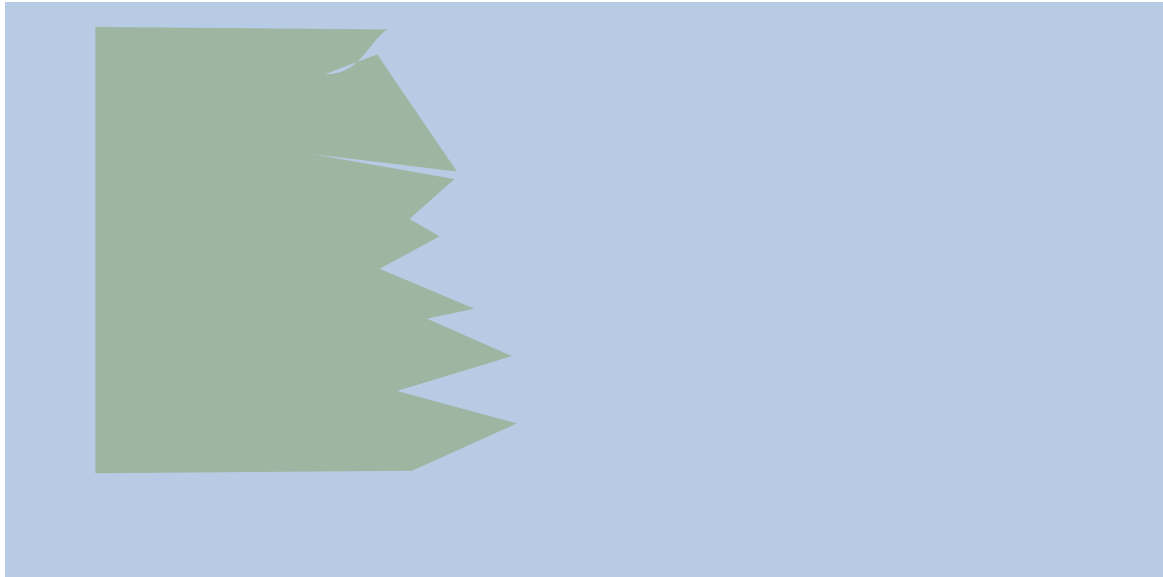
# COLOR AND CONTRAST



Background with different-colored object at similar brightness

- can you see the shape?
- can you count the number of gaps?

# COLOR AND CONTRAST



Background with different-colored object at lower brightness

- can you see the shape?
- can you count the number of gaps?

# COLOR AND CONTRAST



Background with different-colored object at higher brightness

- can you see the shape?
- can you count the number of gaps?



# WHAT DID WE LEARN FROM THAT EXPERIMENT?

Color is for ...

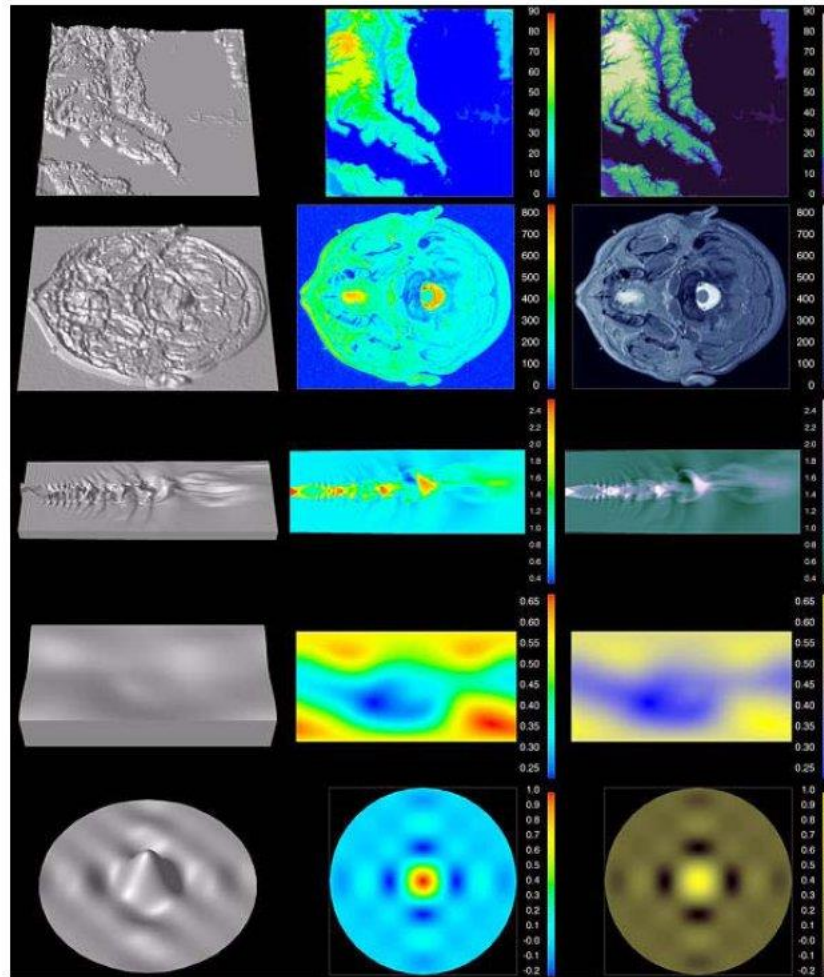
Brightness (intensity, luminance) is for ...

# WHAT DID WE LEARN FROM THAT EXPERIMENT?

Color is for ... labeling

Brightness (intensity, luminance) is for ... fine detail contrast

# LUMINANCE AND HUE



luminance mapped  
to height

just hue

hue and luminance

encode high frequency information by L

# ROLE OF SATURATION

## Art & Money

By: JeanAbbateci



### ART & MONEY

#datavisualisation



This Animated Bubble Chart shows the 270 most expensive artworks sold in auction since 2008 until end 2011

drawing painting  
silkscreen sculpture

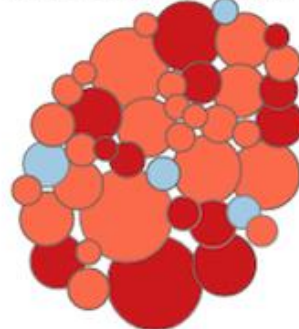
Share 1161 Tweet 558

#### SORTING

- year by year
- top 10 artworks
- men / women
- dead / alive
- by nationality**
- best-selling artists
- auction houses
- size of artworks
- date of creation (all centuries)

#### AMERICAS

United States of America

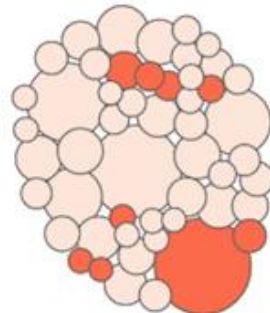


Argentina



#### ASIA

China

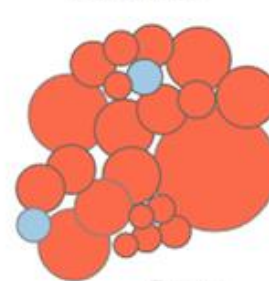


Japan

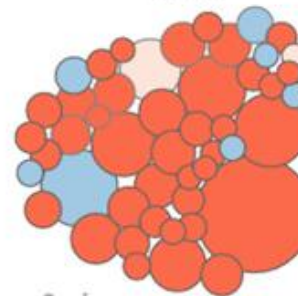


#### EUROPA

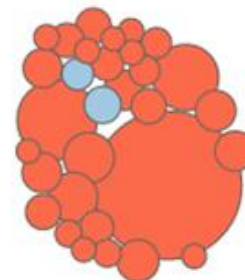
Great Britain



France



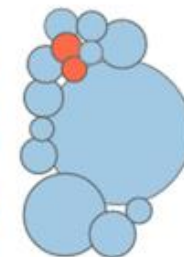
Spain



Netherlands - Belgium



Switzerland



Italy



Russia



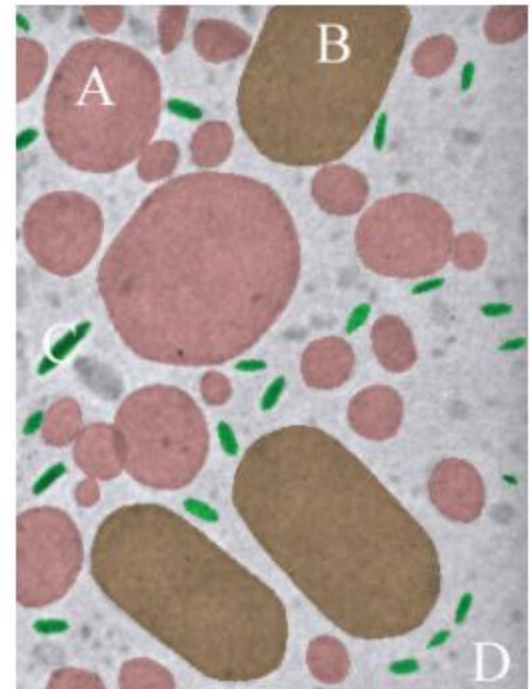
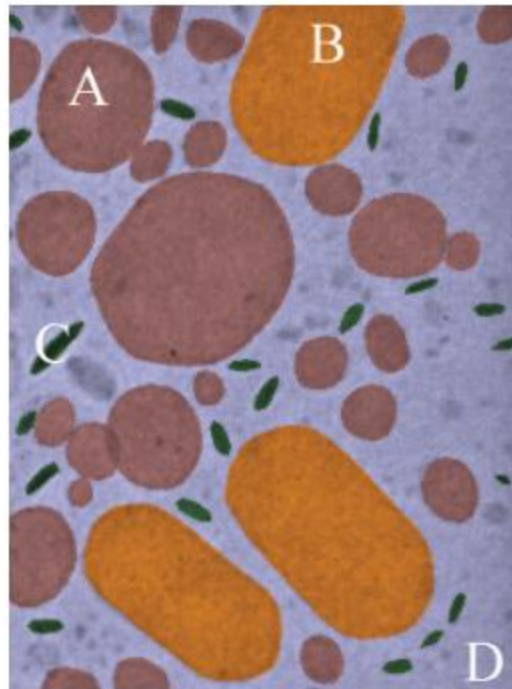
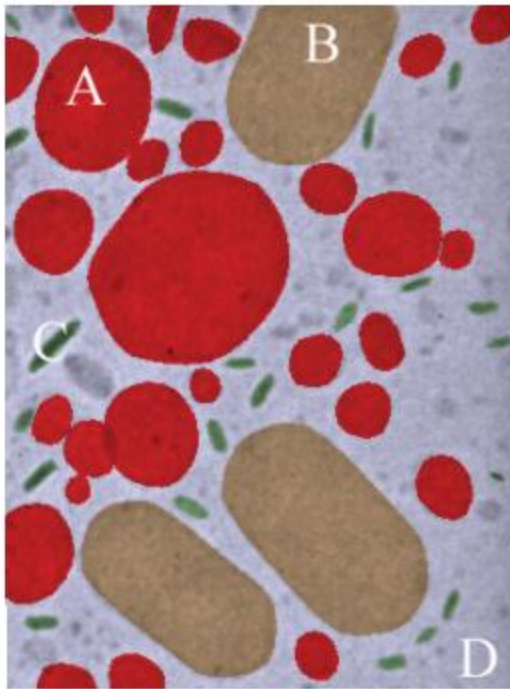
Germany



Others



# COLOR TAGGING FOR IMPORTANCE



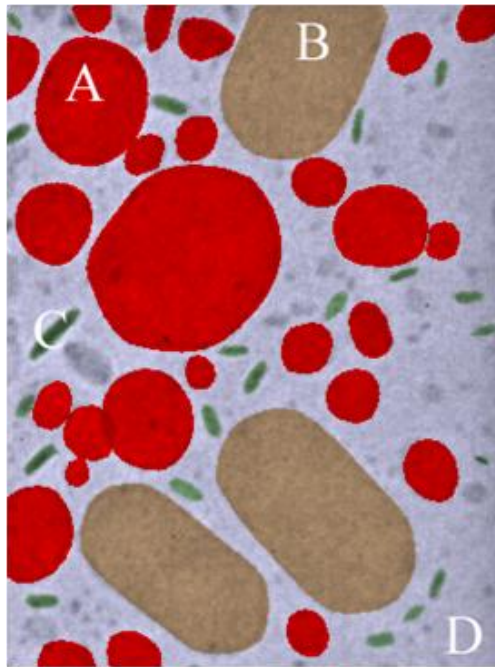
Which is the most important structure in each (as intended by the author)



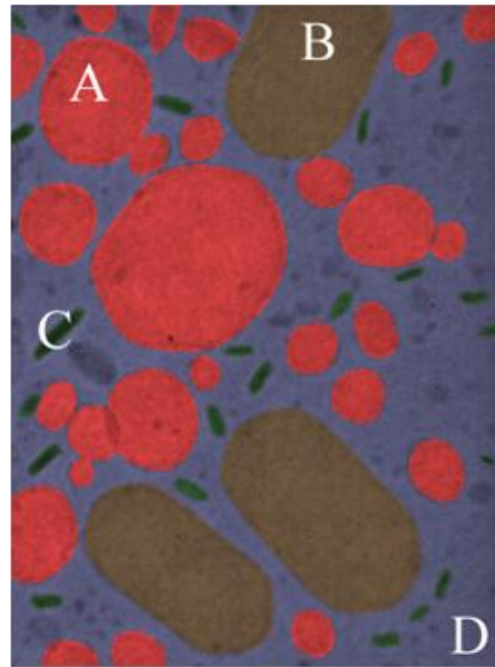
# HOW ABOUT AESTHETICS?

Which one do people like better?

- perceived importance level of red object is the same



Vis 1

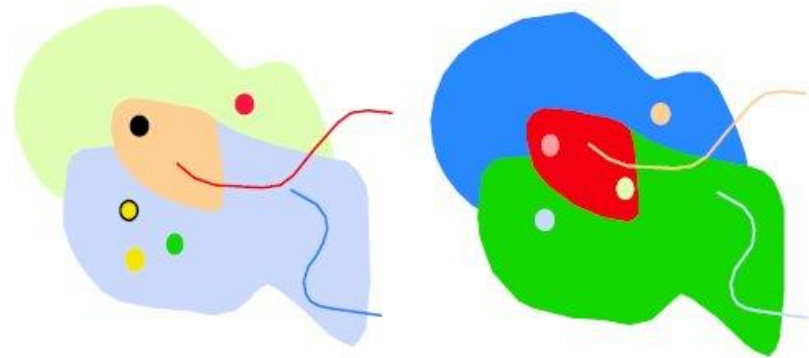


Vis 2



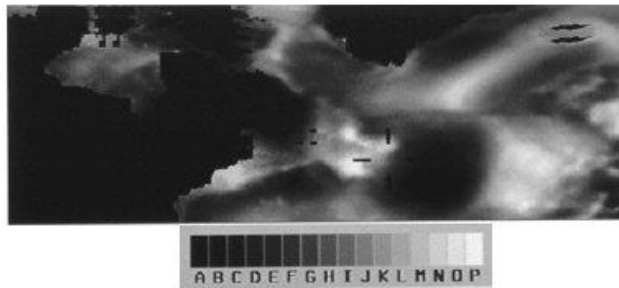
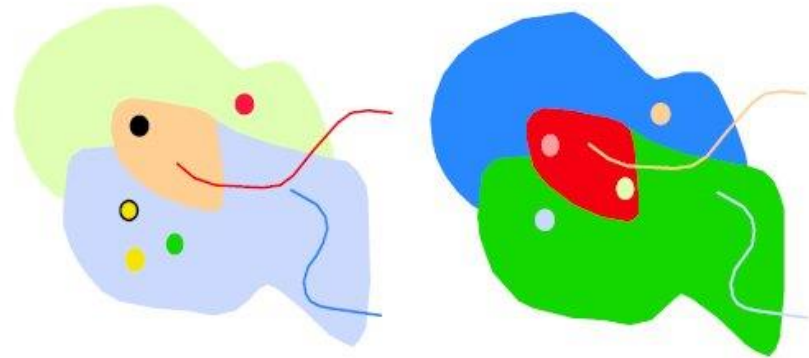
# COLOR CODING AND COLORMAPS

- Color coding
  - large areas: low saturation
  - small areas: high saturation
  - maintain luminance contrast
  - break iso-luminances with borders
- Pseudo-coloring: assign colors to grey levels by indexing the grey levels into a color map

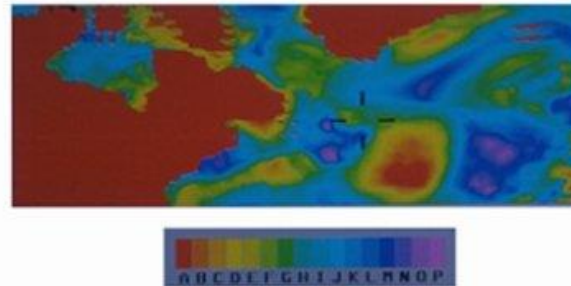


# COLOR CODING AND COLORMAPS

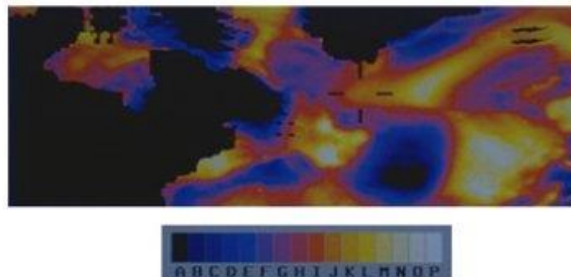
- Color coding
  - large areas: low saturation
  - small areas: high saturation
  - maintain luminance contrast
  - break iso-luminances with borders
- Pseudo-coloring: assign colors to grey levels by indexing the grey levels into a color map



original greylevel map



simple spectrum sequence  
with iso-luminance



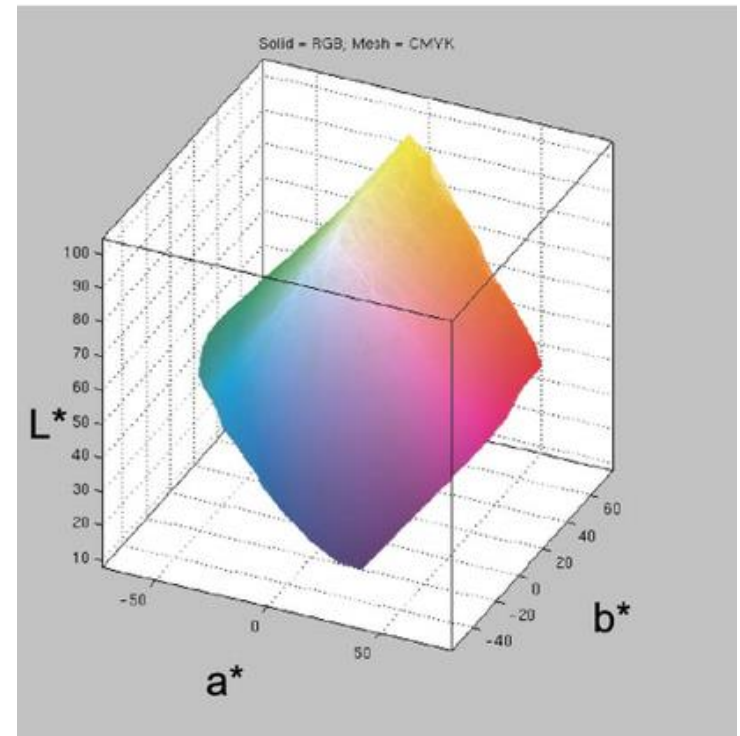
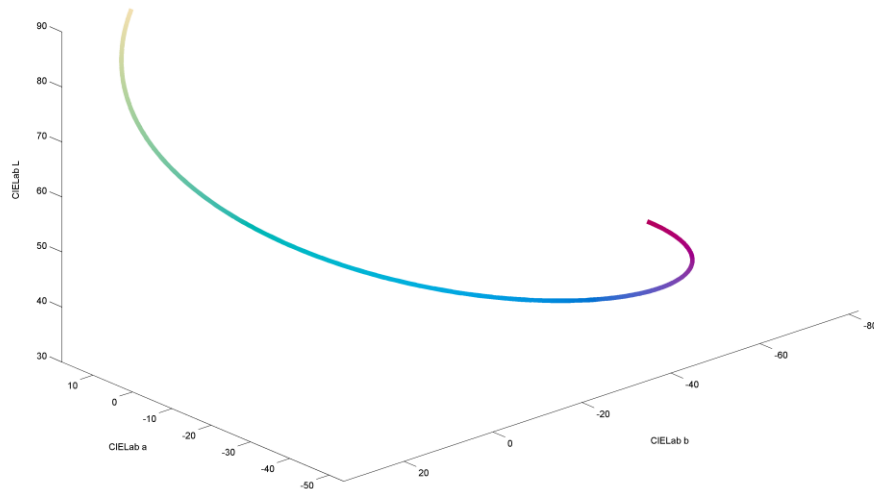
more effective:  
spiral sequence through  
color space  
luminance increases with  
hue



# SPIRAL THROUGH COLOR SPACE

Varies hue and intensity at the same time

- shown here: CIE Lab color space



# THE RAINBOW COLORMAP

As we saw, colors can add detail information to a visualization

- instead of 256 levels get  $256^3 = 16,777,216$

Oftentimes you have a visualization with just one variable

- this would give you a grey level image
- how to turn this into a color image for better detail

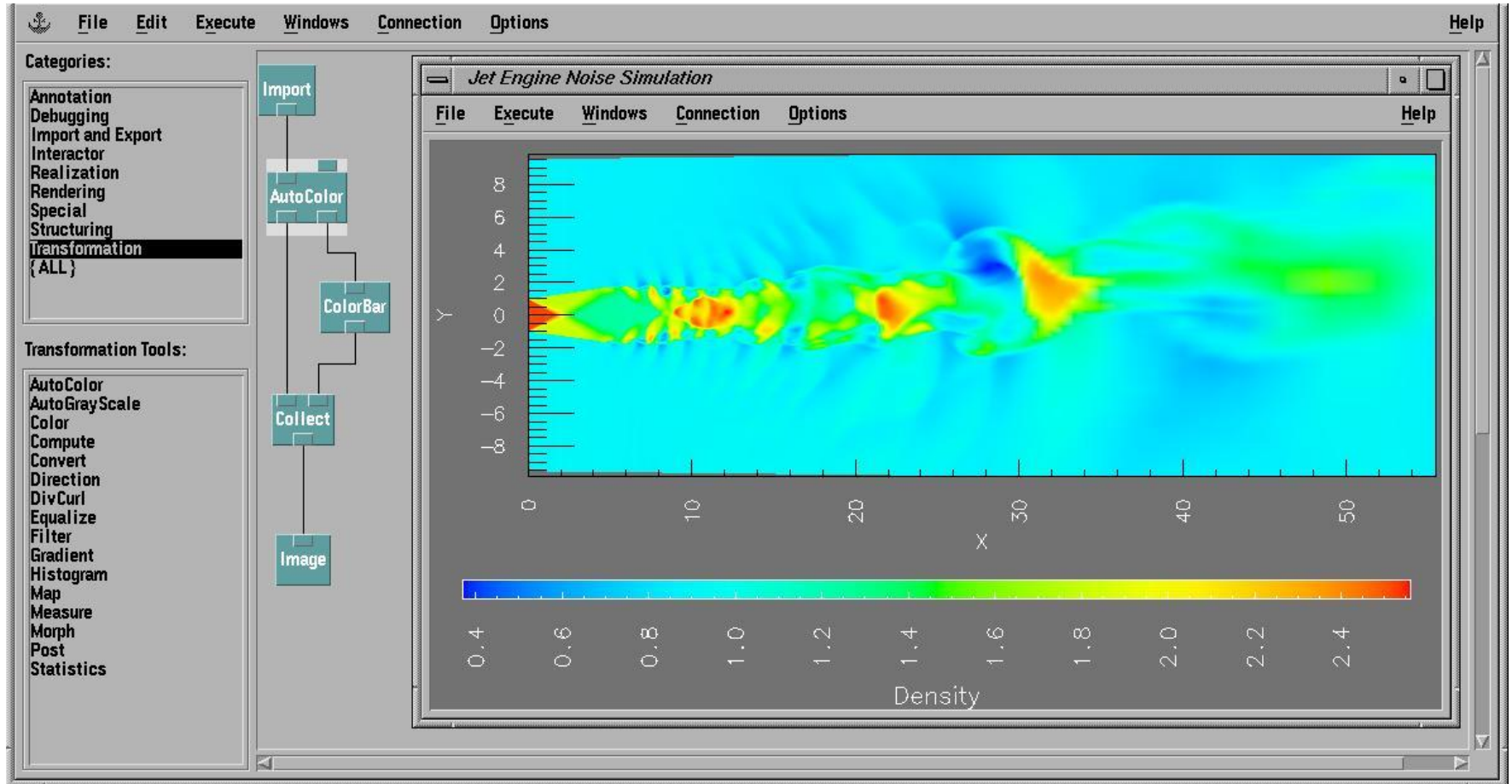
Solution 1:

- map to hue → the rainbow colormap

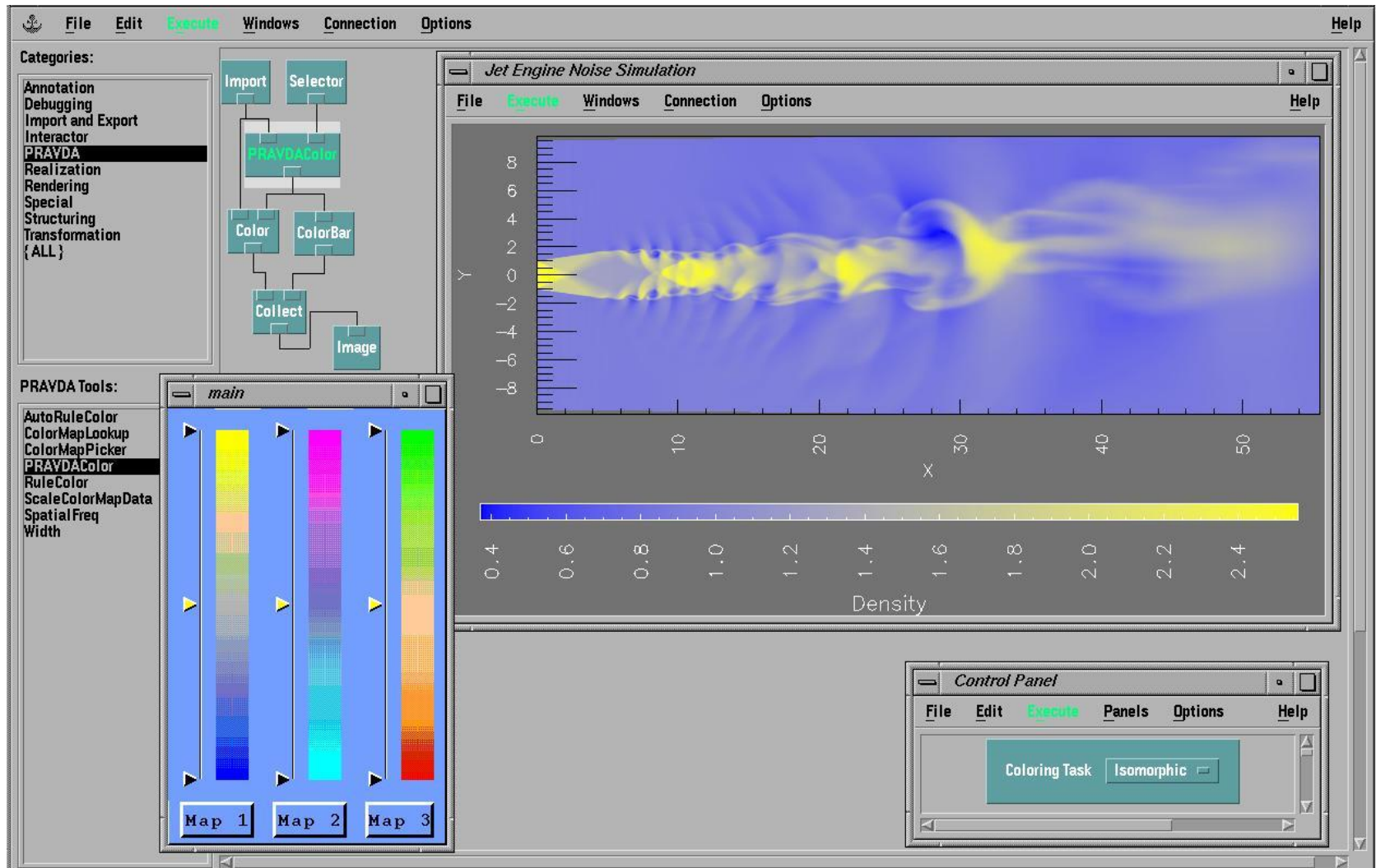


- can you see all adjacent colors at the same contrast?

# AVOID RAINBOW COLORMAPS



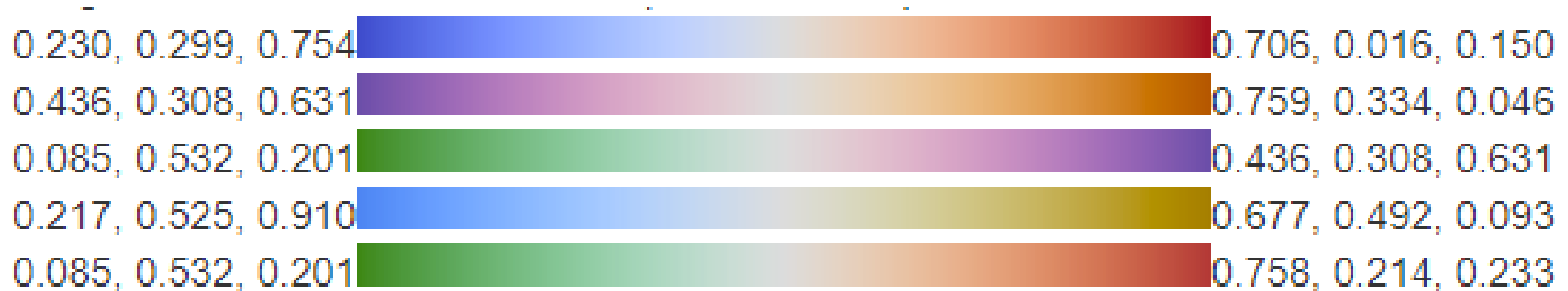
# BETTER: LINEAR HUE



# MORELAND'S DIVERGING COLORMAPS

## Algorithmically generated

- all have the same midpoint value (0.865, 0.865, 0.865)
- begin and end point listed here



<https://www.kennethmoreland.com/color-maps/>

# BREWER SCALES

## Nominal scales

- distinct hues, but similar emphasis

## Sequential scales

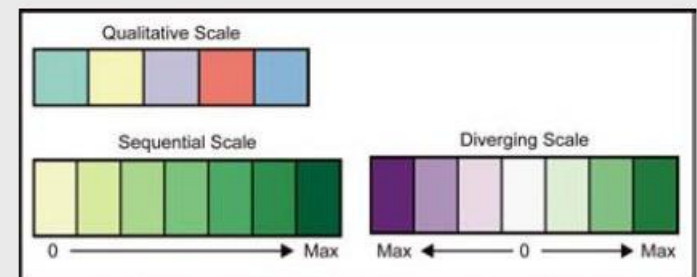
- vary in lightness and saturation
- vary slightly in hue

<http://colorbrewer2.org/>


## Diverging scale


- complementary sequential scales
- neutral at "zero"

### Brewer's Categories

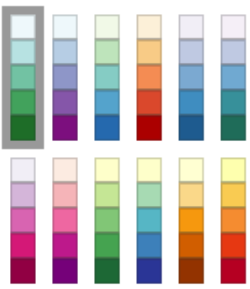


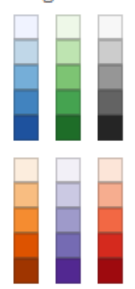
# COLOR BREWER


Number of data classes: 3  [how to use](#) [updates](#) [downloads](#) [credits](#)


**Nature of your data:**   
☒ sequential ☐ diverging ☐ qualitative



**Pick a color scheme:**









Multi-hue: 

Single hue: 

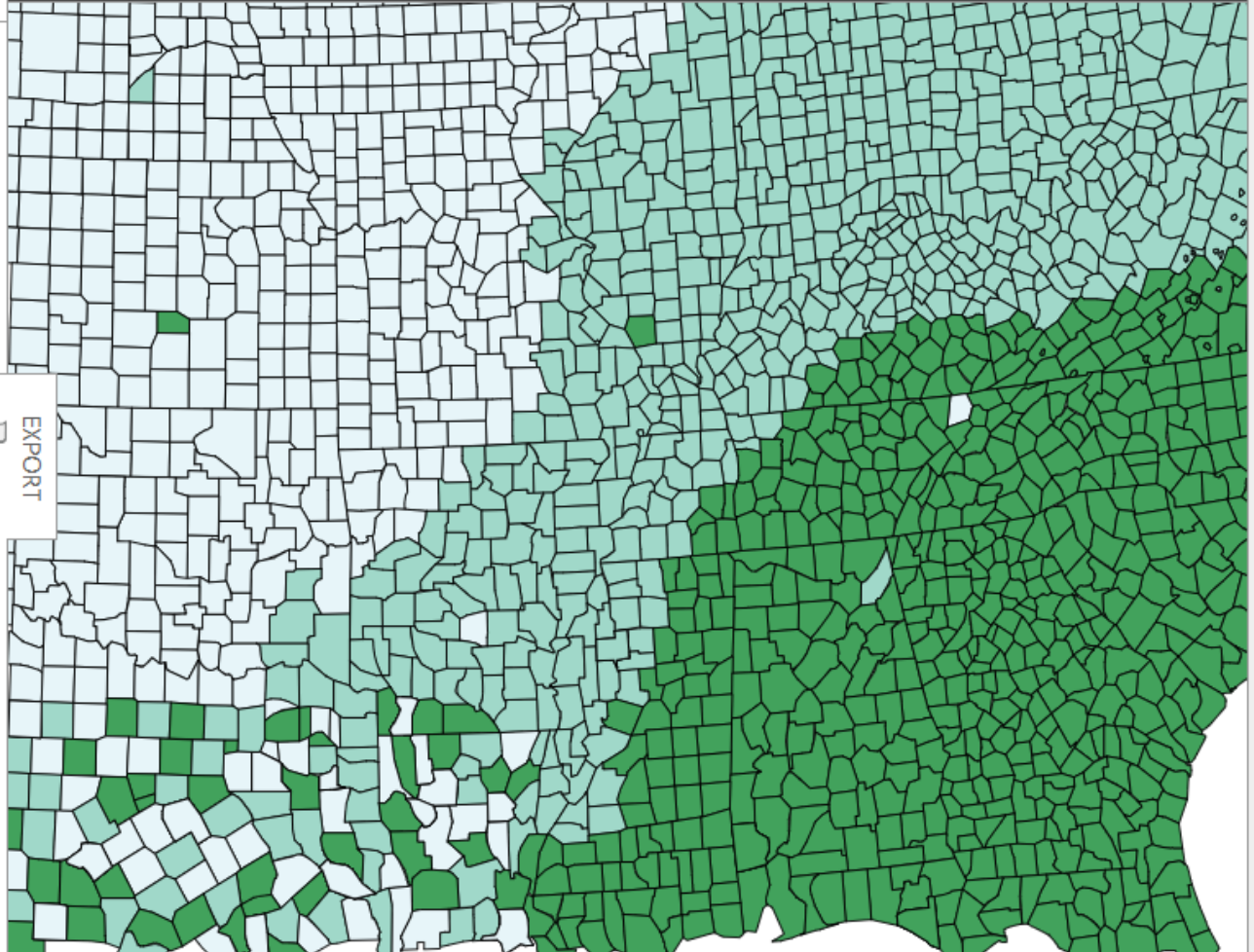
**Only show:**   
☐ colorblind safe  
☐ print friendly  
☐ photocopy safe

**Context:**   
☐ roads  
☐ cities  
☒ borders

**Background:**  
☒ solid color   
☐ terrain  
  
color transparency

**3-class BuGn**  
    EXPORT  
HEX   
 #e5f5f9  
 #99d8c9  
 #2ca25f

**COLORBREWER 2.0**  
color advice for cartography





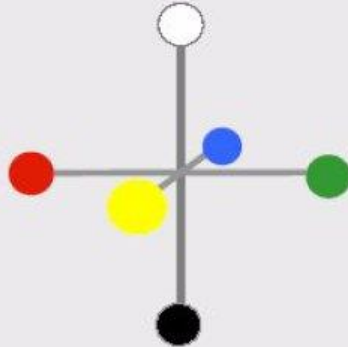
# OPPONENT COLOR

## Definition

- Achromatic axis
- R-G and Y-B axis
- Separate lightness from chroma channels

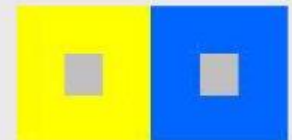
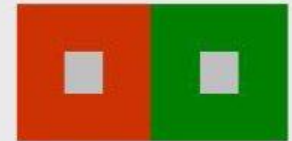
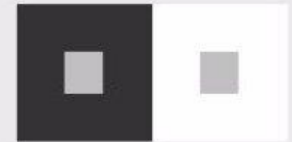
## First level encoding

- Linear combination of LMS
- Before optic nerve
- Basis for perception
- Defines “color blindness”



## Add Opponent Color

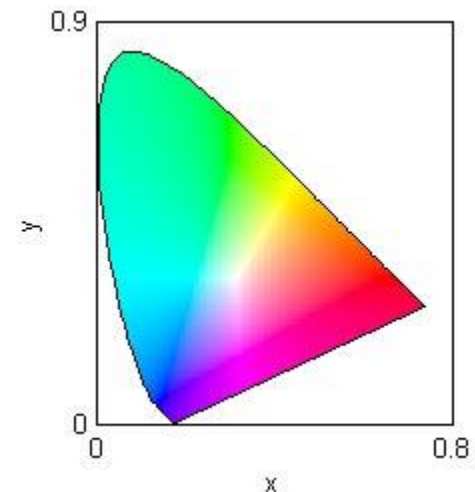
- Dark adds light
- Red adds green
- Blue adds yellow



These samples will have both light/dark and hue contrast

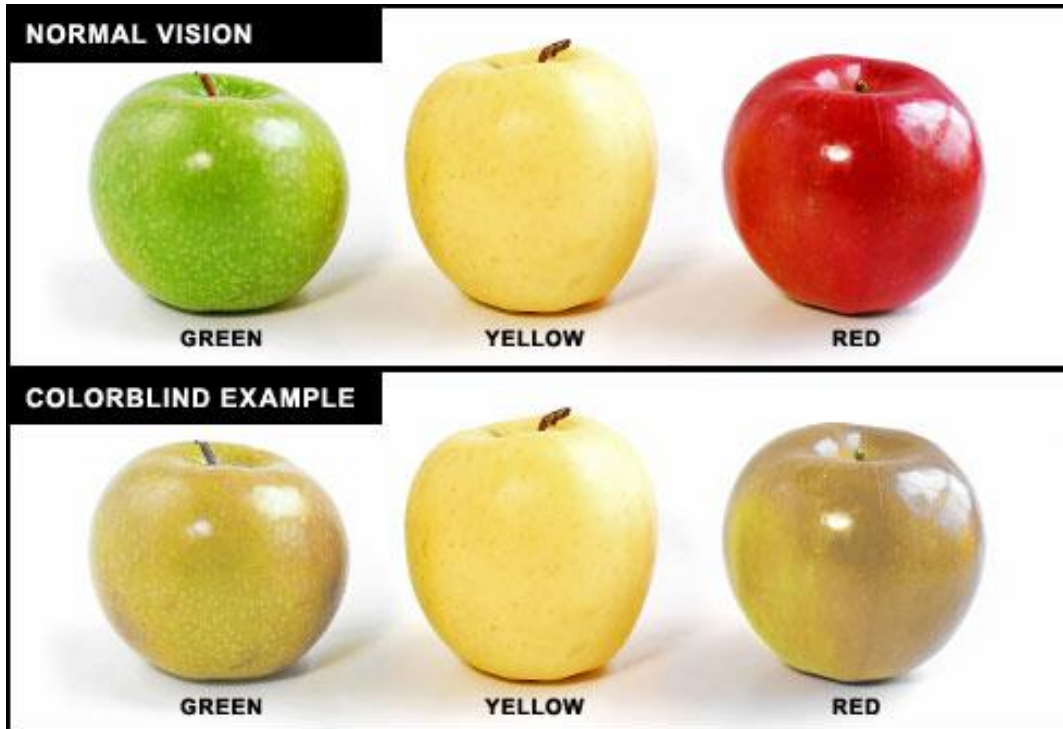
## Opponent colors do not mix

- can only see one of the opponents
- there is no blueish yellow
- there is no reddish green





# COLOR BLINDNESS



Most common is deficiency in distinguishing red and green

# FORMS OF COLOR BLINDNESS

normal



The colors of the rainbow as viewed by a person with no color vision deficiencies.

green missing



The colors of the rainbow as viewed by a person with deuteranopia.

red missing



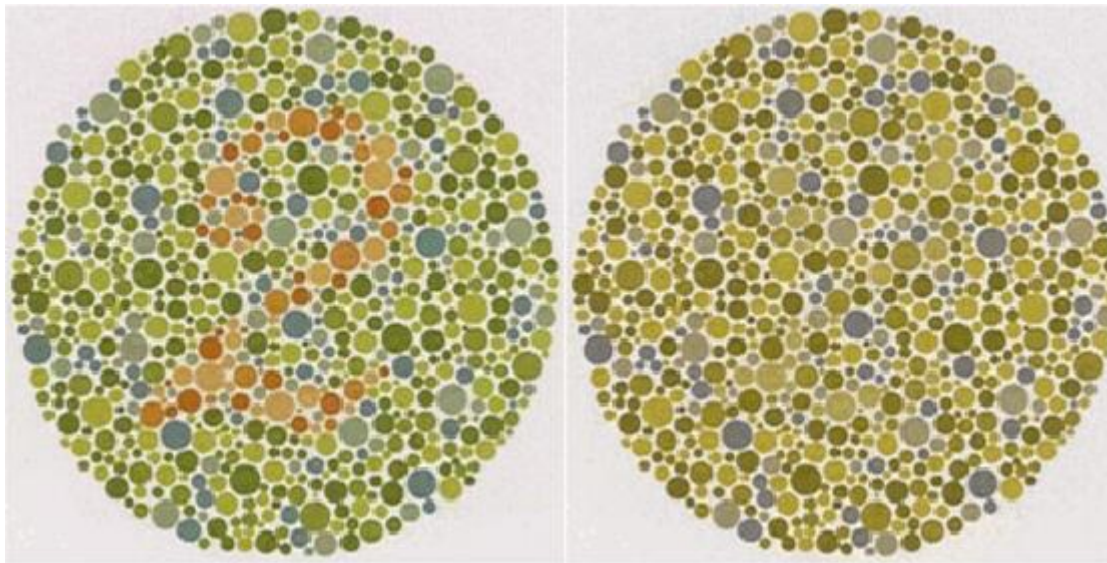
The colors of the rainbow as viewed by a person with protanopia.

blue missing  
(rare)



The colors of the rainbow as viewed by a person with tritanopia.

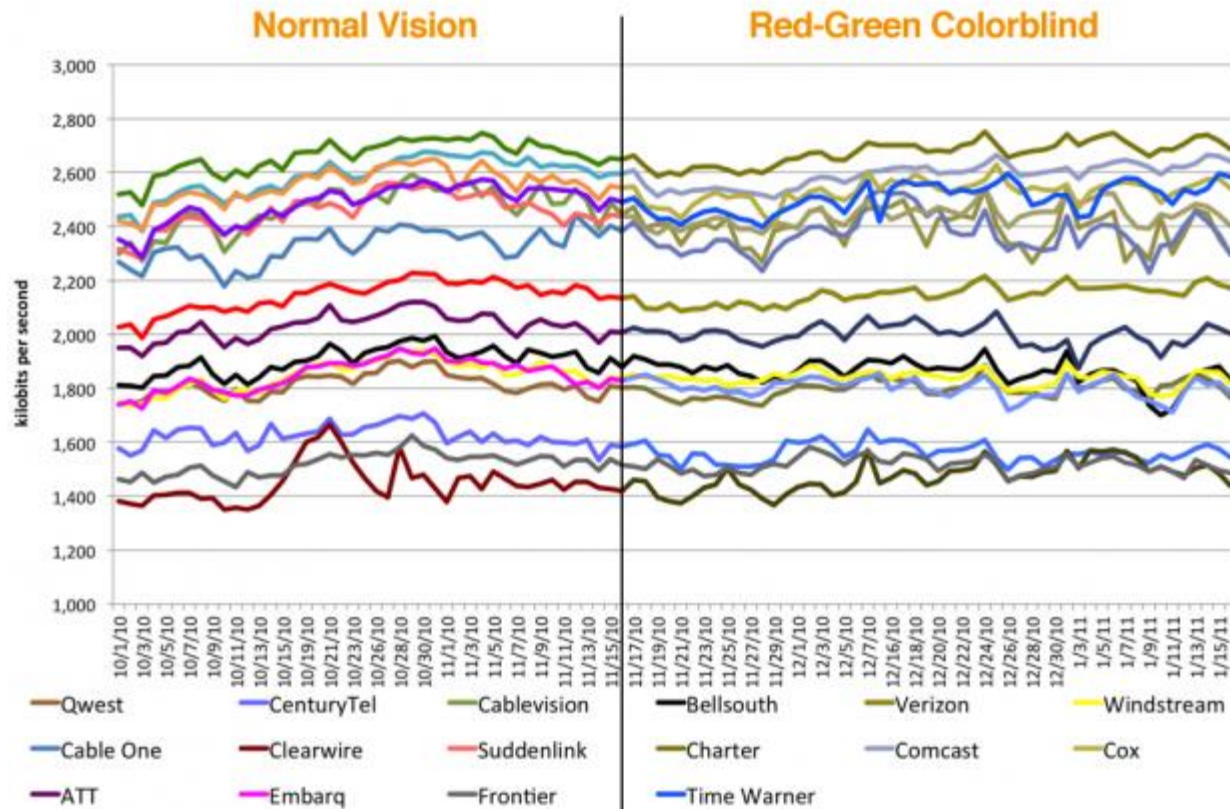
# ISHIHARA TEST



normal

protanopia

# LINE CHARTS



# DESIGNING FOR COLOR DEFICIENT USERS

8% (0.5%) of US males (females) are color deficient

- so be careful when designing visualizations

What to do?

- use different intensities for red-green (e.g. light green, dark red)
- space red and green colored colors dots far apart or make large
- add symbols to line charts or vary line style
- avoid using gradient colors to indicate data value

# SUMMING UP

Use Luminance for detail, shape, and form

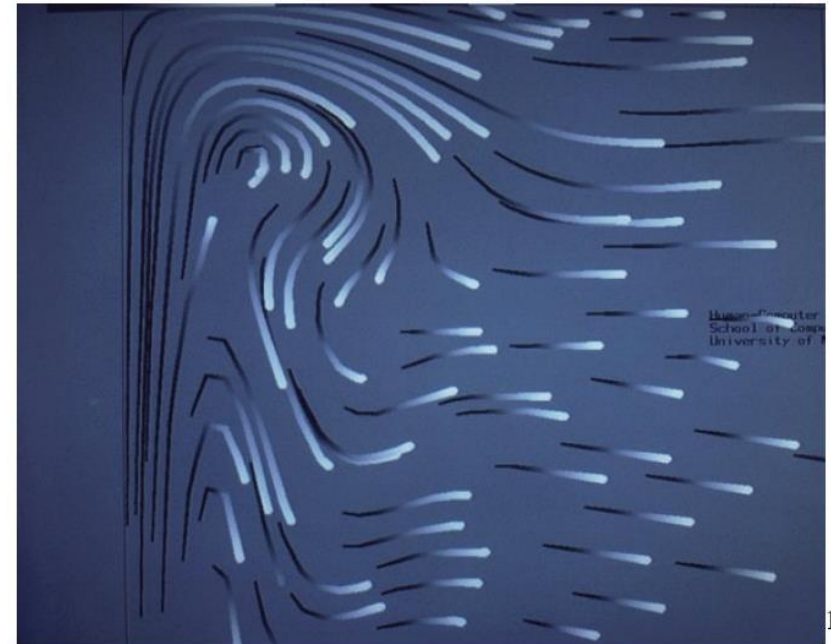
Use color for coding – few colors

Use strong colors for small areas

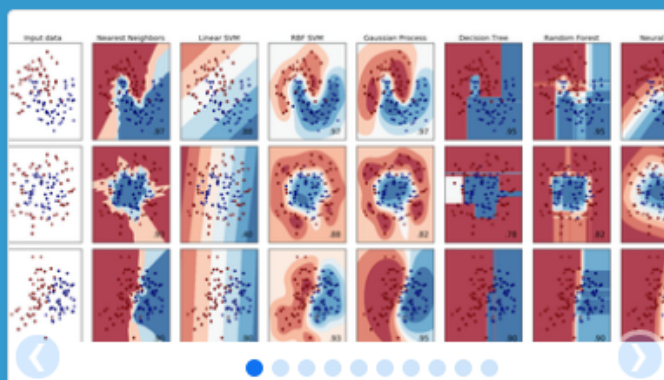
Use subtle colors to code large areas

Visualization artistry:

- Use of luminance to indicate direction







# scikit-learn

*Machine Learning in Python*

- Simple and efficient tools for data mining and data analysis
- Accessible to everybody, and reusable in various contexts
- Built on NumPy, SciPy, and matplotlib
- Open source, commercially usable - BSD license

## Classification

Identifying to which category an object belongs to.

**Applications:** Spam detection, Image recognition.

**Algorithms:** SVM, nearest neighbors, random forest, ... [— Examples](#)

## Regression

Predicting a continuous-valued attribute associated with an object.

**Applications:** Drug response, Stock prices.

**Algorithms:** SVR, ridge regression, Lasso, ... [— Examples](#)

## Clustering

Automatic grouping of similar objects into sets.

**Applications:** Customer segmentation, Grouping experiment outcomes

**Algorithms:** k-Means, spectral clustering, mean-shift, ... [— Examples](#)

## Dimensionality reduction

Reducing the number of random variables to consider.

**Applications:** Visualization, Increased efficiency

**Algorithms:** PCA, feature selection, non-negative matrix factorization. [— Examples](#)

## Model selection

Comparing, validating and choosing parameters and models.

**Goal:** Improved accuracy via parameter tuning

**Modules:** grid search, cross validation, metrics. [— Examples](#)

## Preprocessing

Feature extraction and normalization.

**Application:** Transforming input data such as text for use with machine learning algorithms.

**Modules:** preprocessing, feature extraction. [— Examples](#)