



# an introduction to information graphics and data visualisation

max van kleek

University of Southampton  
Open Data Institute Short Course

6179 planes

0:28 EST

part one!

biological basis of information design

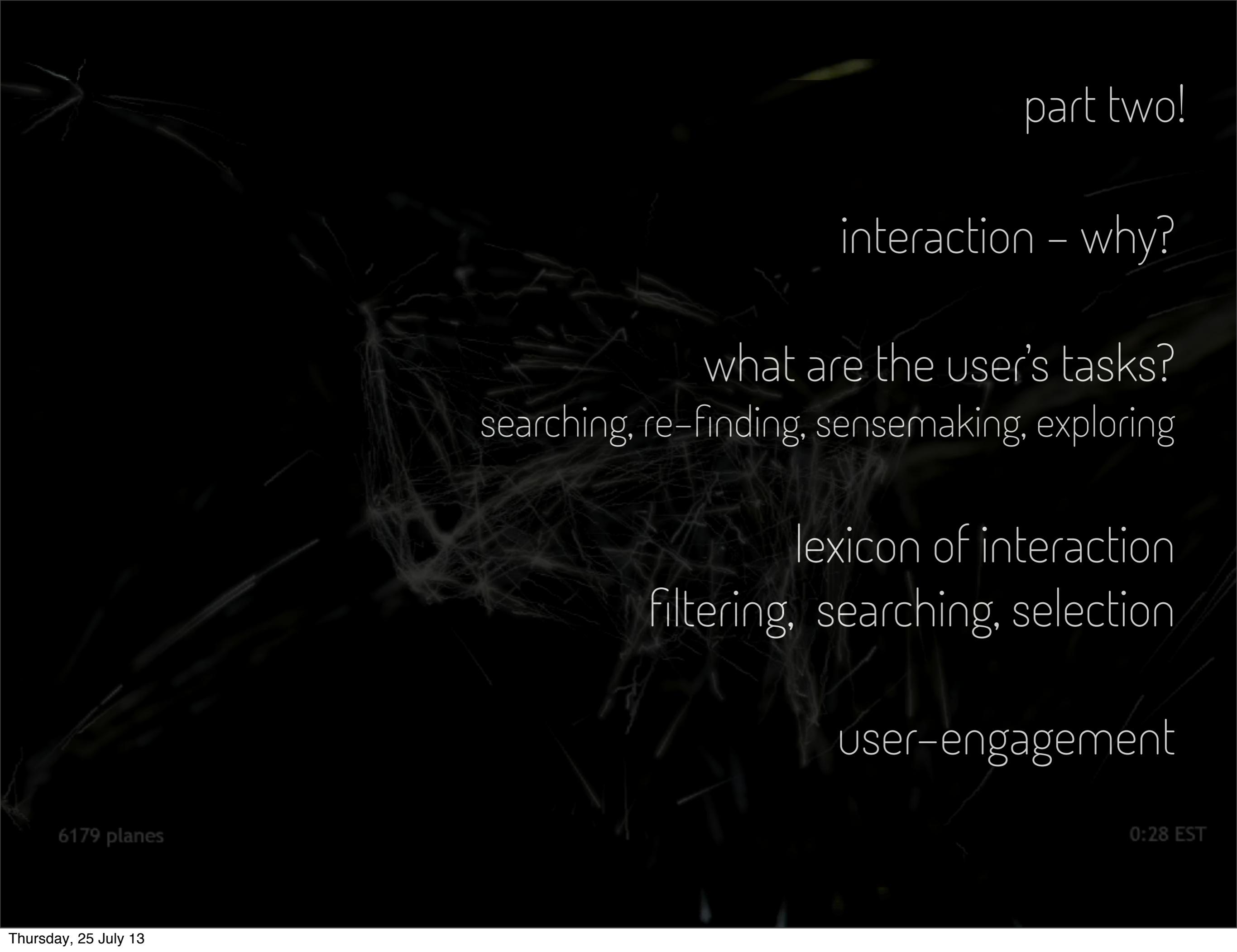
visual dimensions and data dimensions

tasks

deception and bad infographics

6179 planes

0:28 EST



part two!

interaction - why?

what are the user's tasks?  
searching, re-finding, sensemaking, exploring

lexicon of interaction  
filtering, searching, selection

user-engagement

key objectives

what are the **goals** of visualisation?

how do you **choose** a visual representation for data?

how do you **evaluate** a visualisation?

key objectives

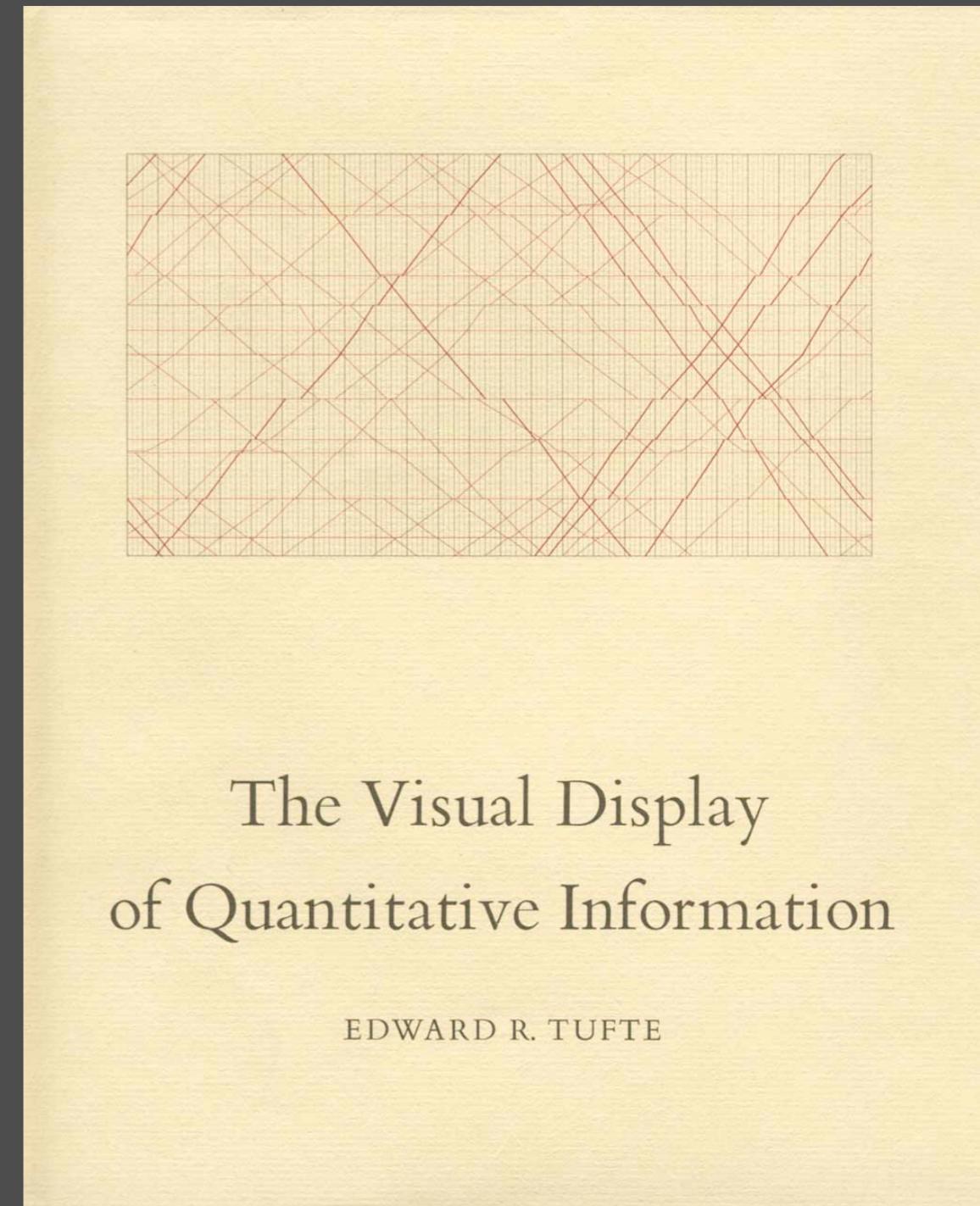
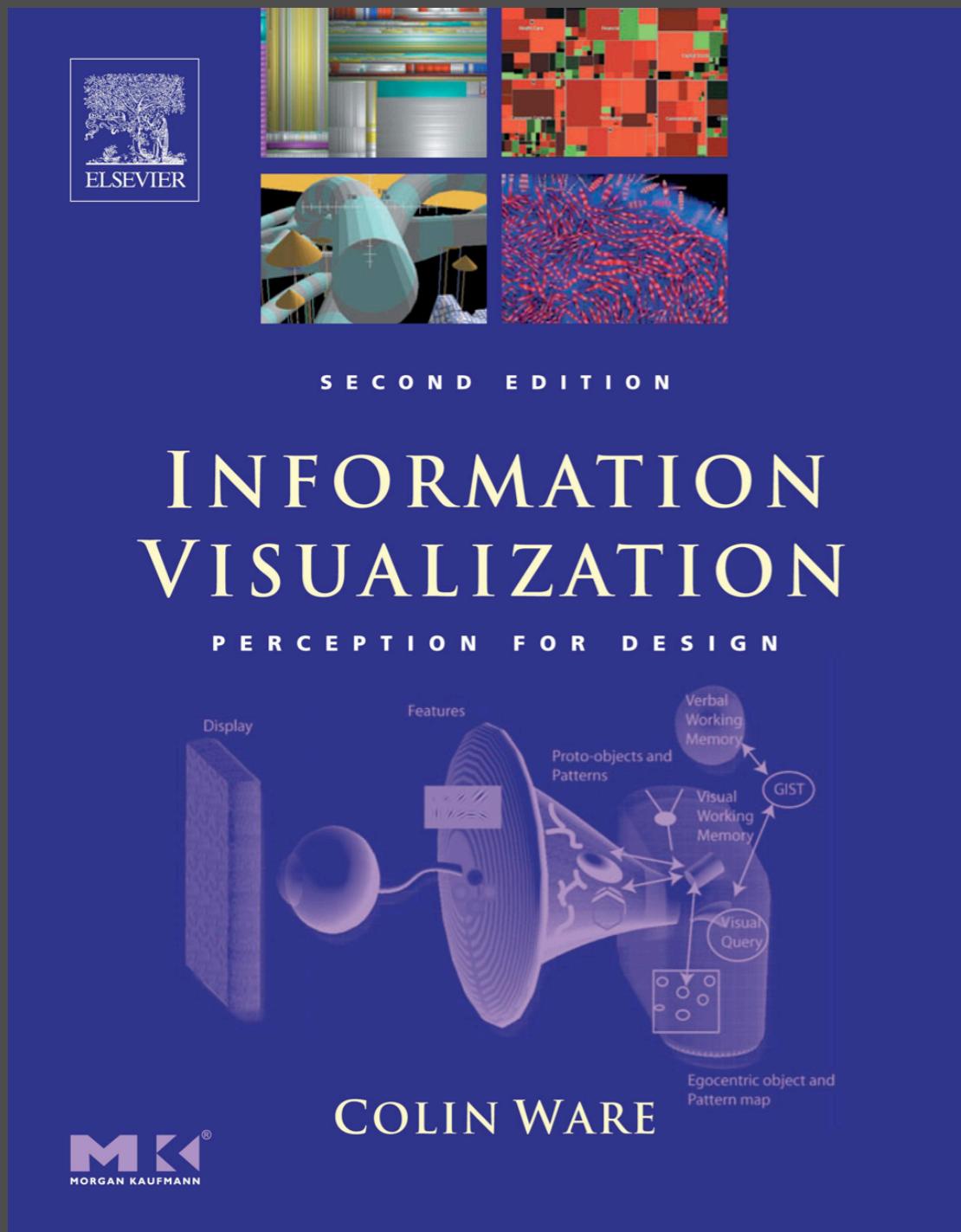
aesthetics + engagement - is ‘pretty’ better?

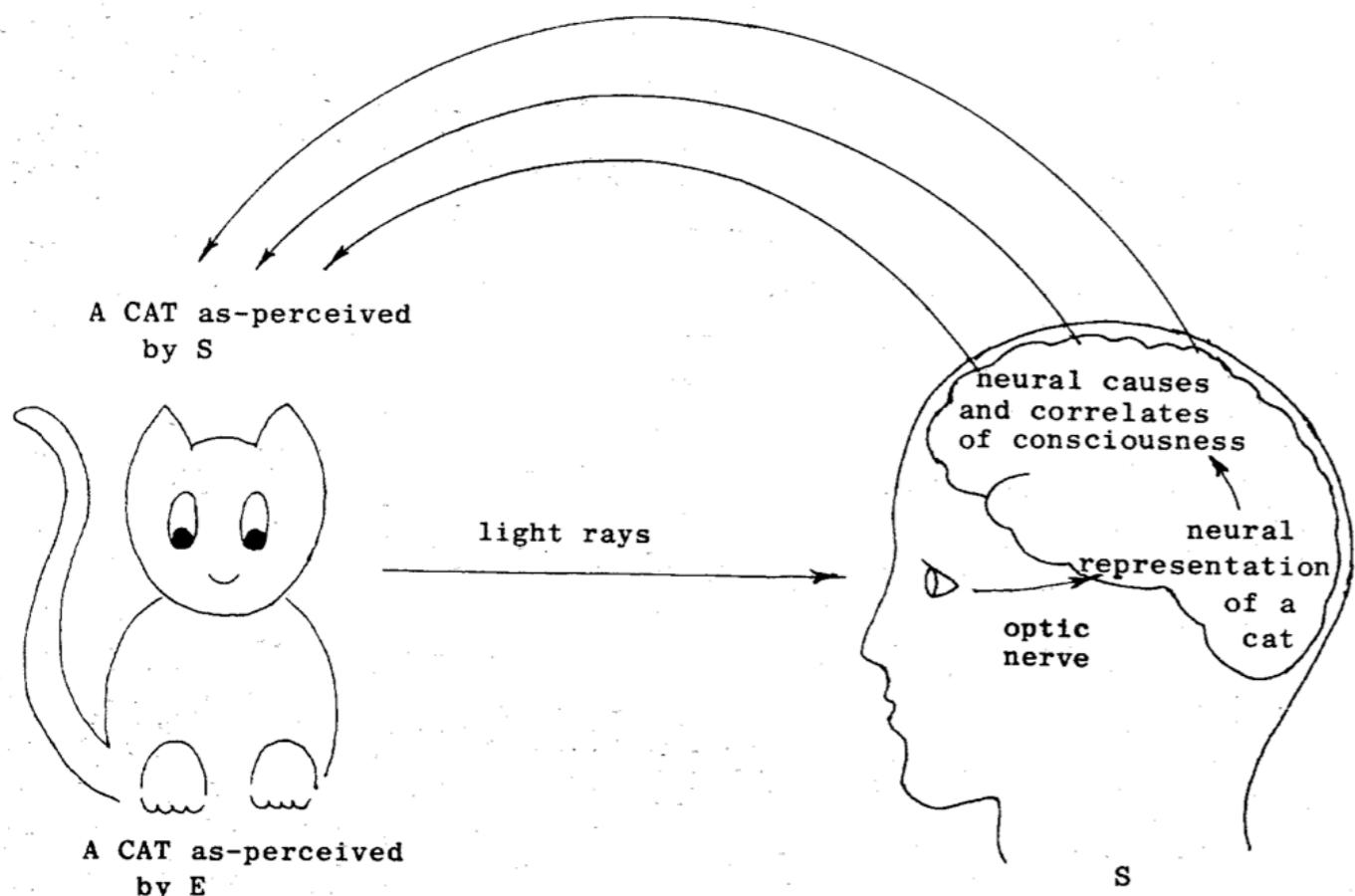
identifying distortion + deception

wielding power tools (excel / matlab / etc )  
vs hacking bespoke approaches

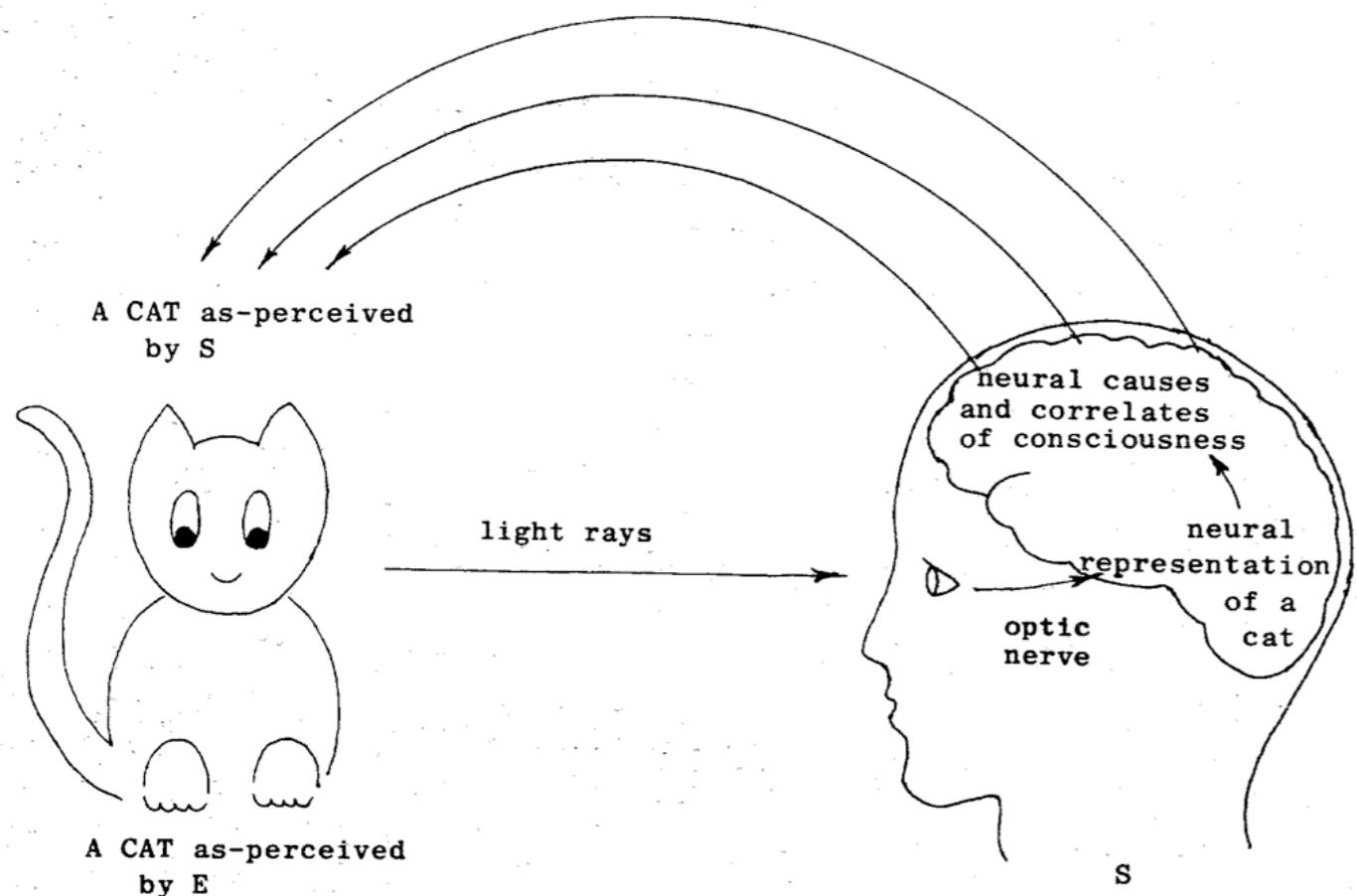
minor objectives

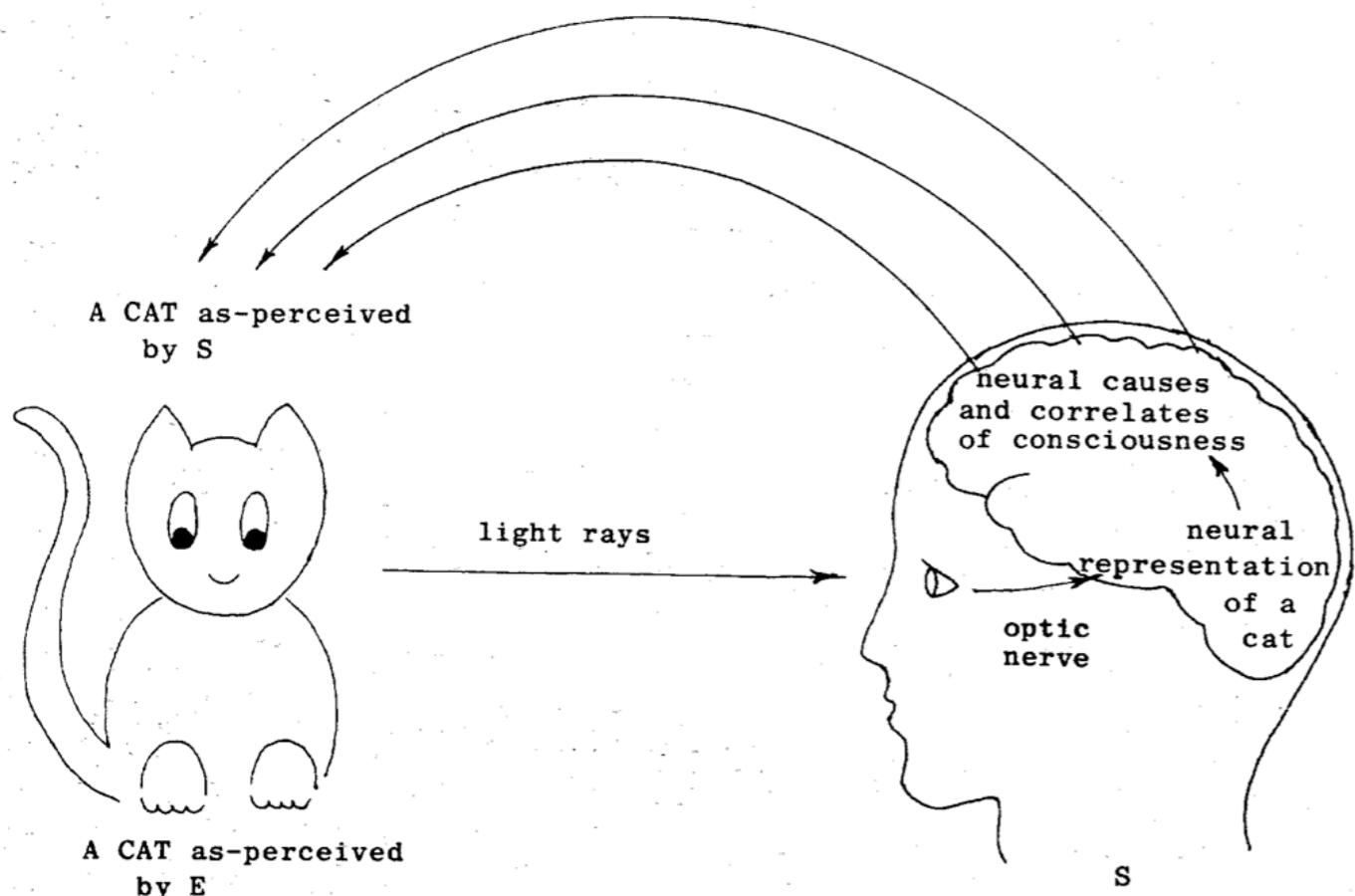
# recommended texts





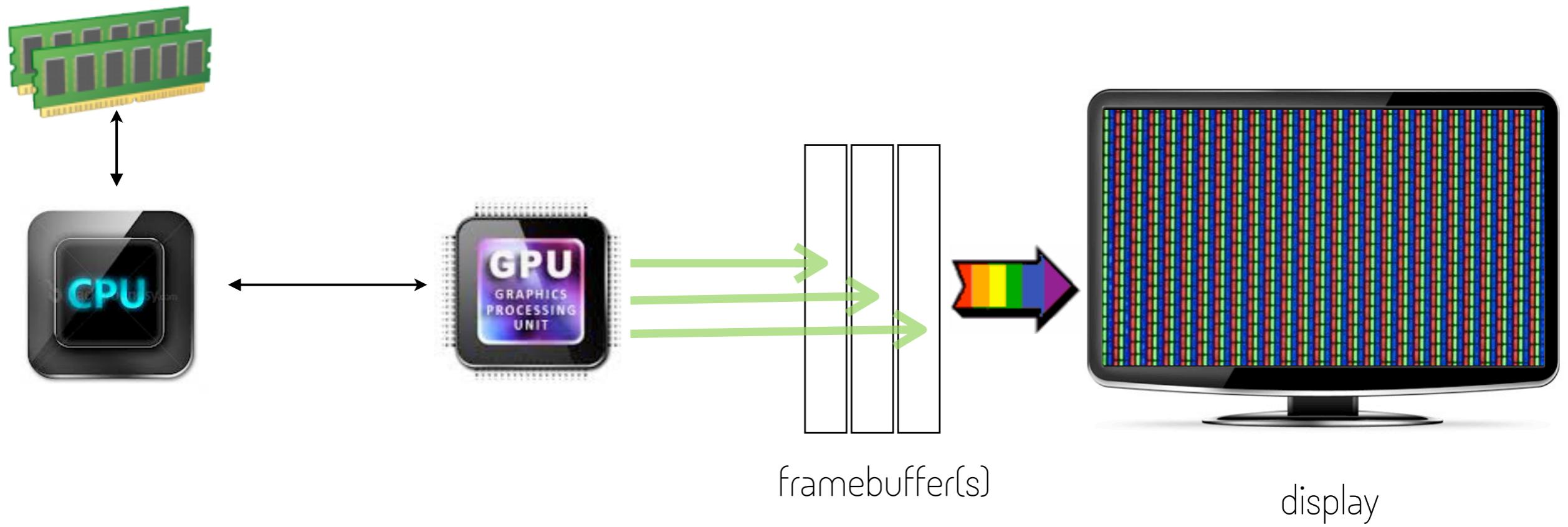
what is the goal  
of information design?

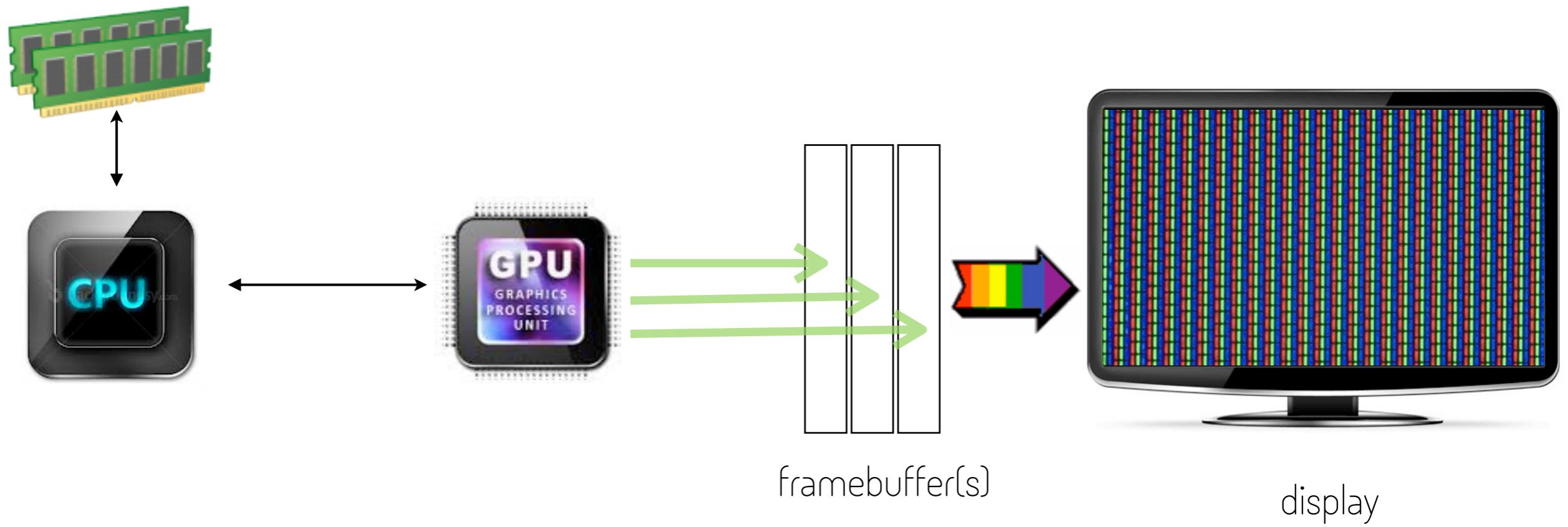




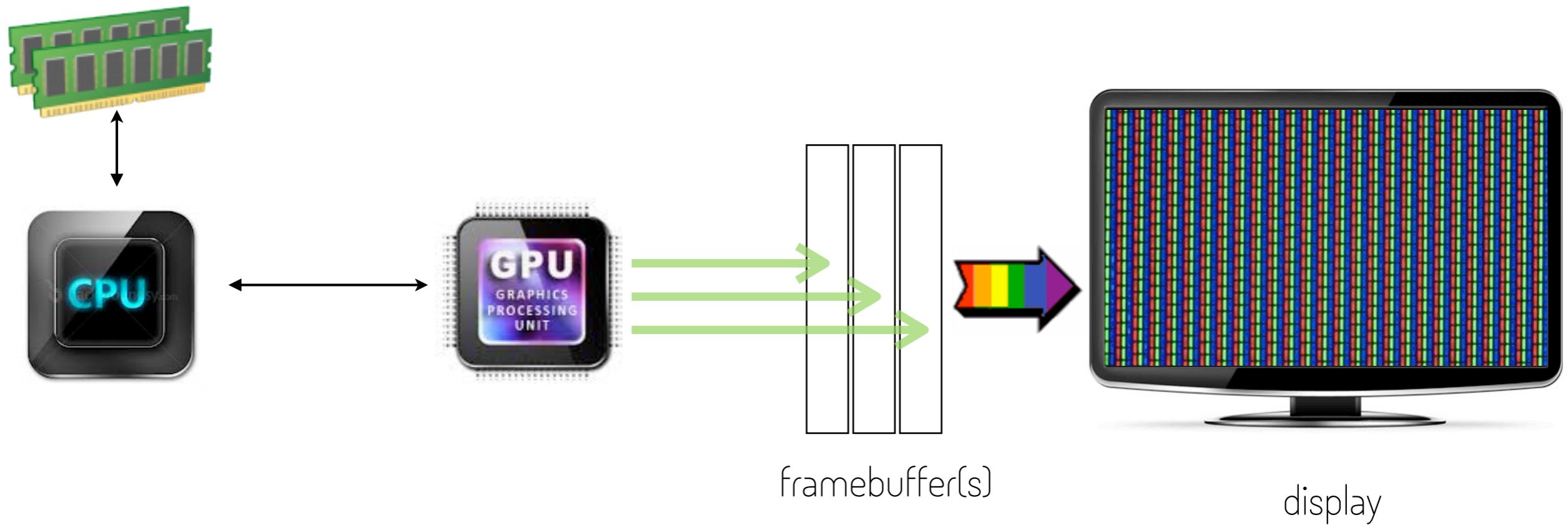
what is the goal  
of information design?

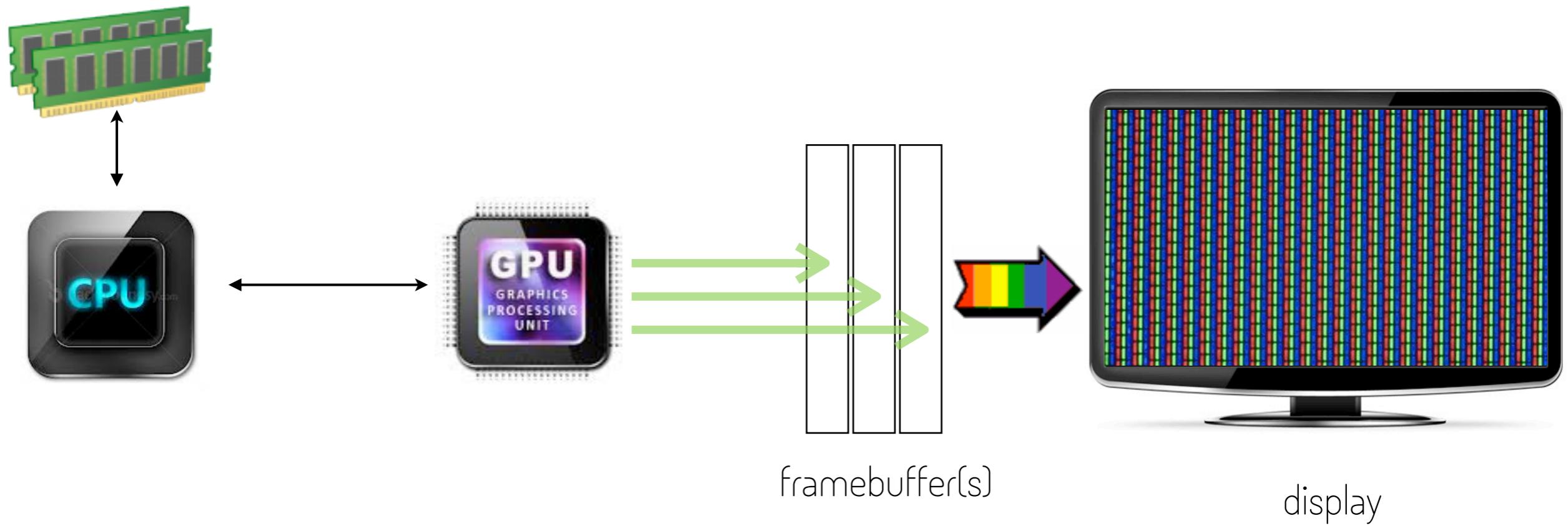
to let us better  
**think** and communicate  
about data





typical computer architecture

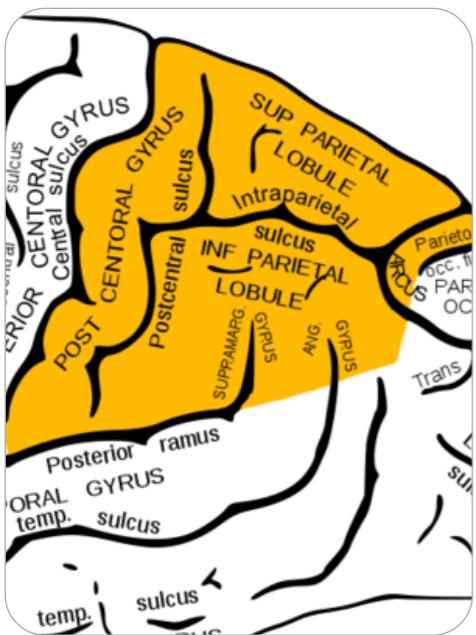




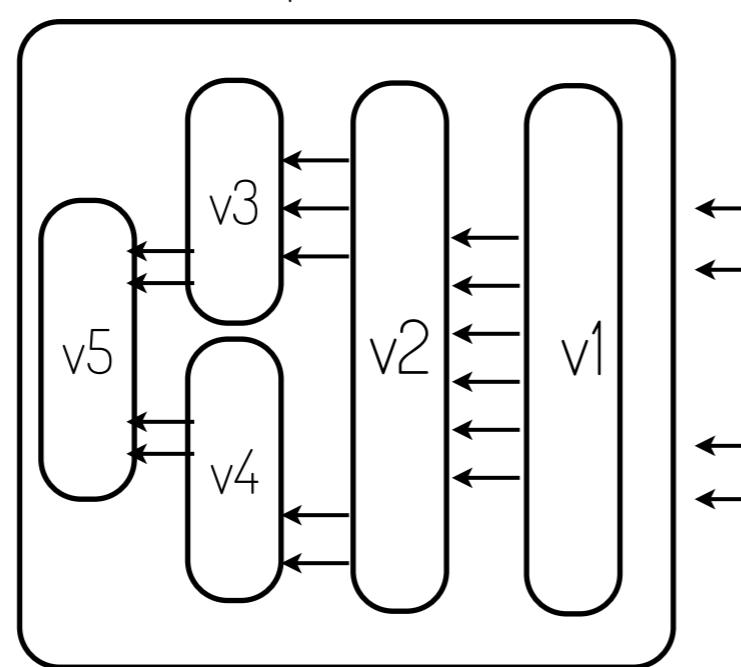
framebuffer(s)

display

parietal lobe + frontal cortex

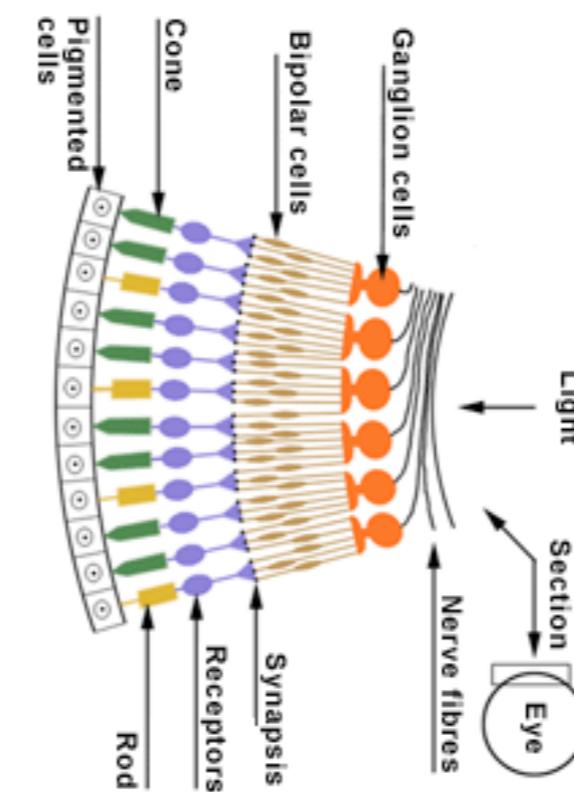


occipital lobe



spatial orientation  
focus of attention  
eye control,  
perceptual fusion

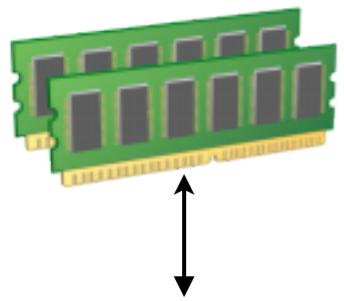
visual cortex  
(pattern detection)



eye / iris / fovea

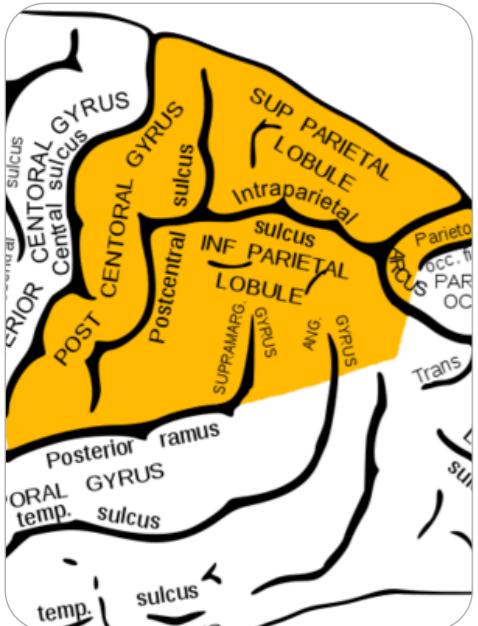
retina  
(sensing)





serial /  
deliberative  
processing  
“attention-focused”

parietal lobe + frontal cortex

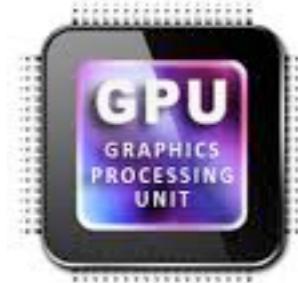


spatial orientation  
focus of attention  
eye control,  
perceptual fusion

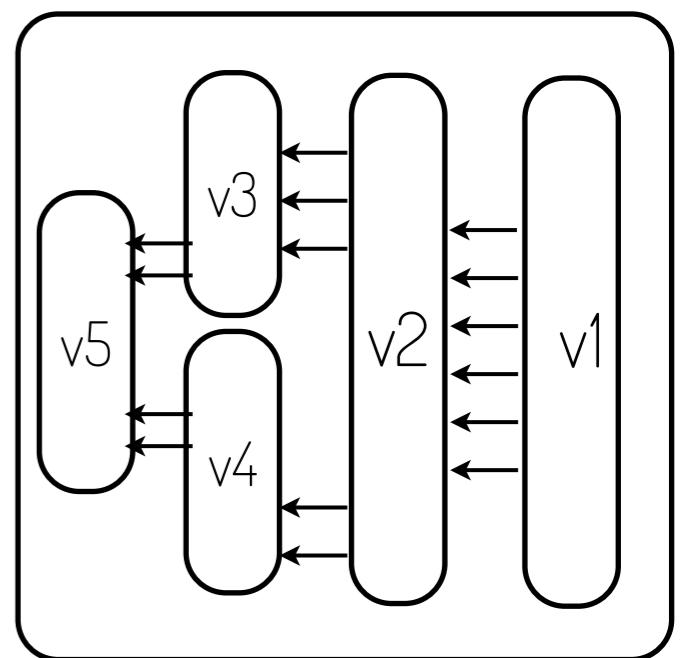
access to  
long term memory

highly parallel

visual processing  
routines  
optimised for  
purpose



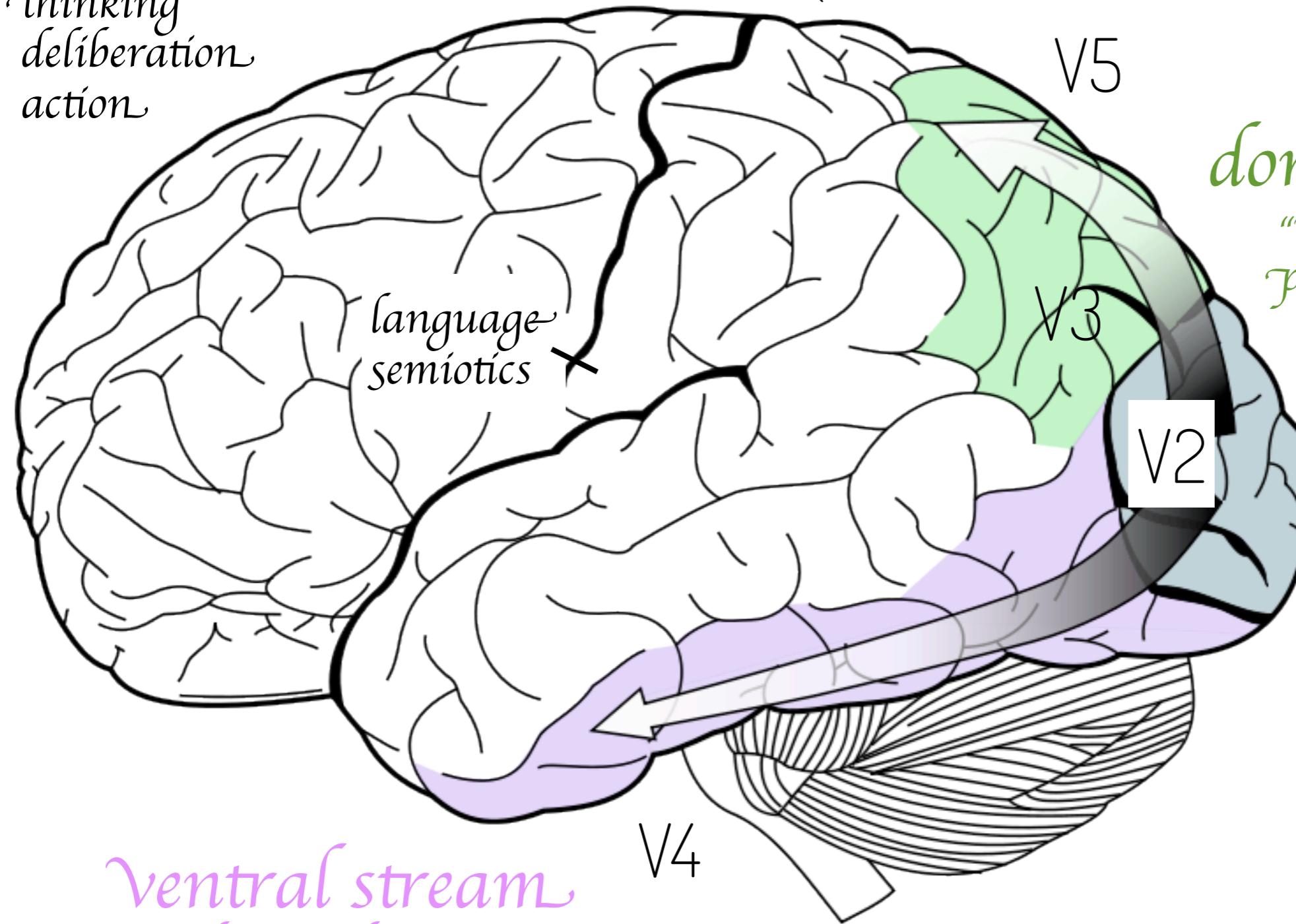
occipital lobe



visual cortex  
(pattern detection)

frontal lobe

*planning  
thinking  
deliberation  
action*



parietal lobe

*spatial reasoning  
perceptual fusion*

V5

*dorsal stream  
"where/how"  
pathways*

V2

V1

occipital lobe

*ventral stream  
"what" pathway*

V4

a

b

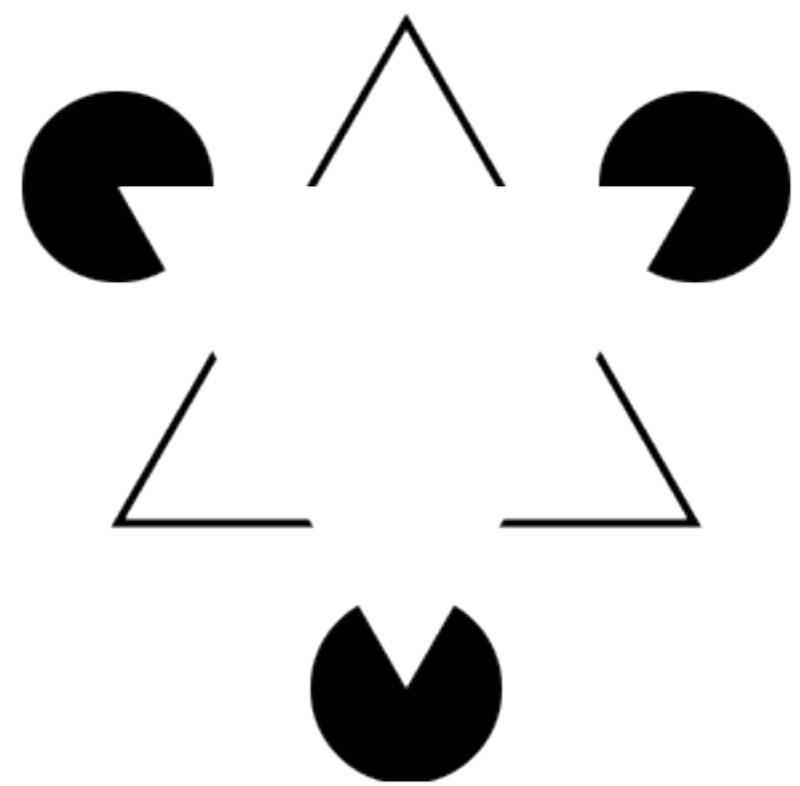
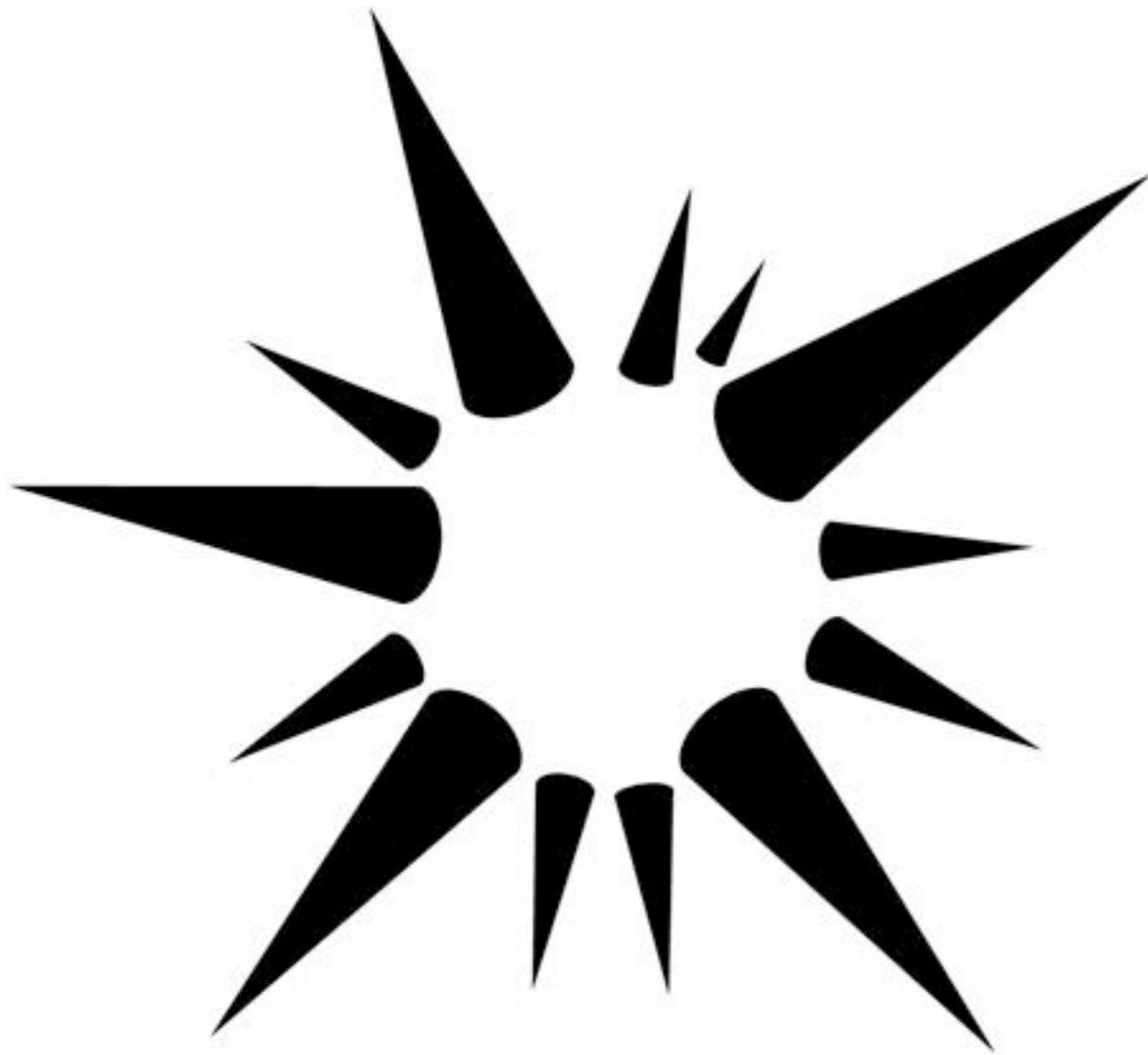
A large grid of 100 identical Persian characters 'ب' (Be) arranged in 10 rows and 10 columns. The characters are black and have a slightly italicized script style.

C

A large grid of handwritten cursive 's' characters, arranged in 15 rows and 15 columns. The characters are written in a black ink-like color on a light gray background. Each character is a single, fluid stroke, showing the characteristic curve and loop of the cursive letter 's'. The grid is centered on the page.

d

A large grid of 100 small white 'Y' symbols on a black background. The grid is composed of 10 rows and 10 columns of these symbols.



50 0 50 100 150 200

X Pump • Deaths from cholera



John Snow, 1854  
London Cholera Outbreak

50 0 50 100 150 200

X Pump • Deaths from cholera

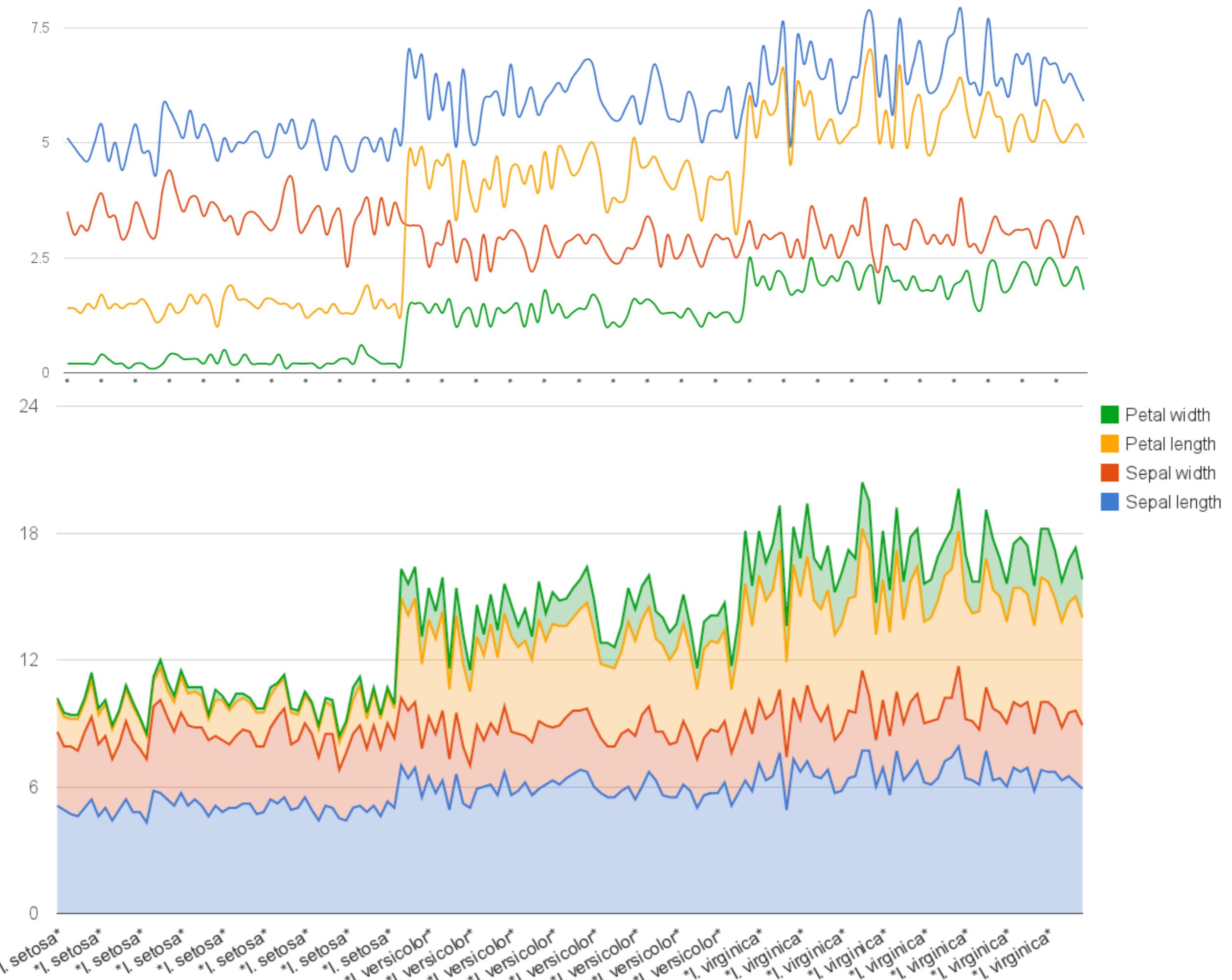


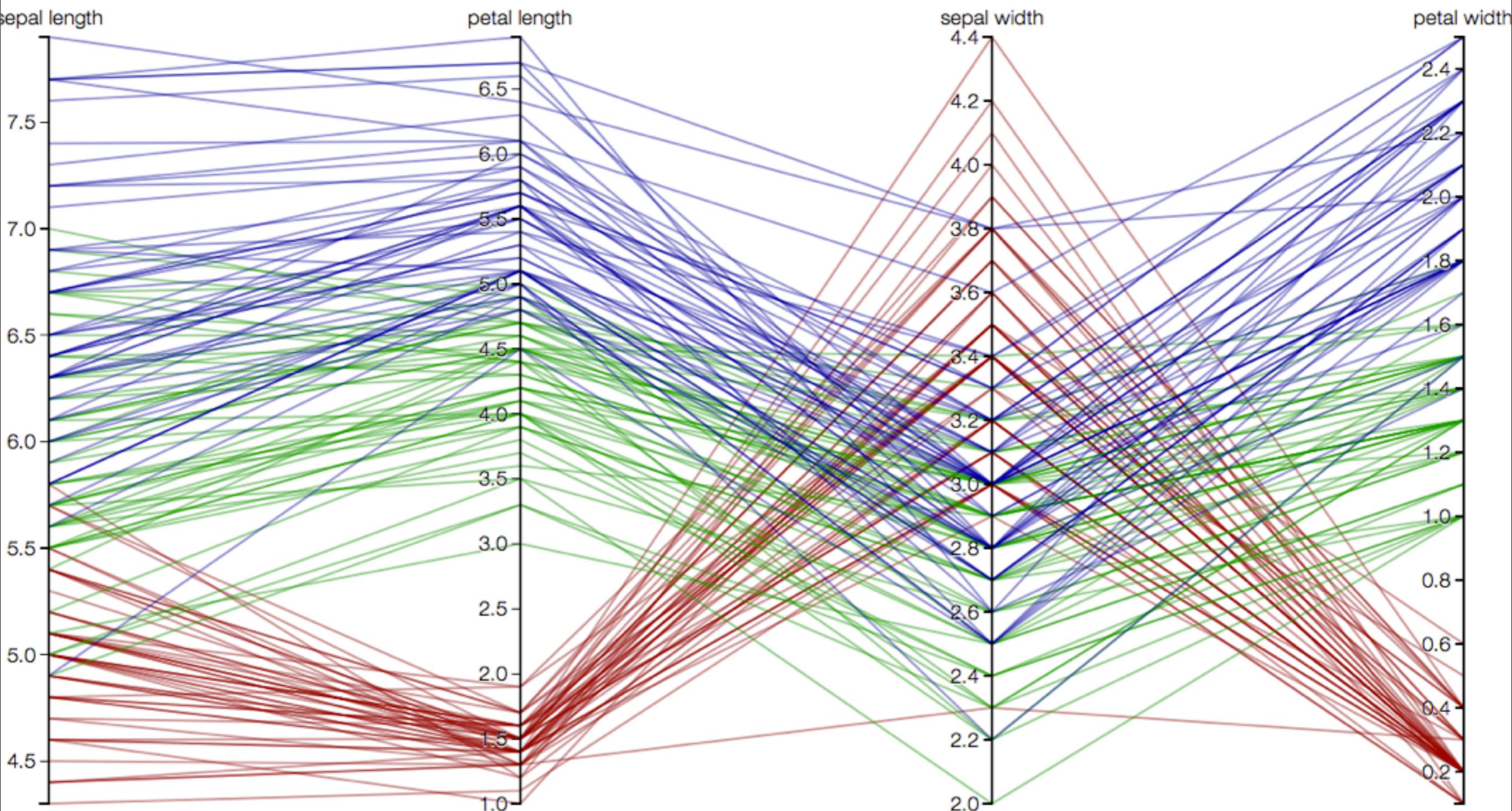
John Snow, 1854  
London Cholera Outbreak

"There was one significant anomaly - none of the monks in the adjacent monastery contracted cholera. Investigation showed that this was not an anomaly, but further evidence, for they drank only beer, which they brewed themselves."

Sepal length	Sepal width	Petal length	Petal width	Species	Sepal length	Sepal width	Petal length	Petal width	Species	Sepal length	Sepal width	Petal length	Petal width	Species
5.1	3.5	1.4	0.2	*I. setosa*	5.4	3.4	1.5	0.4	*I. setosa*	5.4	3.4	1.5	0.4	*I. setosa*
4.9	3	1.4	0.2	*I. setosa*	5.2	4.1	1.5	0.1	*I. setosa*	5.2	4.1	1.5	0.1	*I. setosa*
4.7	3.2	1.3	0.2	*I. setosa*	5.5	4.2	1.4	0.2	*I. setosa*	5.5	4.2	1.4	0.2	*I. setosa*
4.6	3.1	1.5	0.2	*I. setosa*	4.9	3.1	1.5	0.2	*I. setosa*	4.9	3.1	1.5	0.2	*I. setosa*
5	3.6	1.4	0.2	*I. setosa*	5	3.2	1.2	0.2	*I. setosa*	5	3.2	1.2	0.2	*I. setosa*
5.4	3.9	1.7	0.4	*I. setosa*	5.5	3.5	1.3	0.2	*I. setosa*	5.5	3.5	1.3	0.2	*I. setosa*
4.6	3.4	1.4	0.3	*I. setosa*	4.9	3.6	1.4	0.1	*I. setosa*	4.9	3.6	1.4	0.1	*I. setosa*
5	3.4	1.5	0.2	*I. setosa*	4.4	3	1.3	0.2	*I. setosa*	4.4	3	1.3	0.2	*I. setosa*
4.4	2.9	1.4	0.2	*I. setosa*	5.1	3.4	1.5	0.2	*I. setosa*	5.1	3.4	1.5	0.2	*I. setosa*
4.9	3.1	1.5	0.1	*I. setosa*	5	3.5	1.3	0.3	*I. setosa*	5	3.5	1.3	0.3	*I. setosa*
5.4	3.7	1.5	0.2	*I. setosa*	4.5	2.3	1.3	0.3	*I. setosa*	4.5	2.3	1.3	0.3	*I. setosa*
4.8	3.4	1.6	0.2	*I. setosa*	4.4	3.2	1.3	0.2	*I. setosa*	4.4	3.2	1.3	0.2	*I. setosa*
4.8	3	1.4	0.1	*I. setosa*	5	3.5	1.6	0.6	*I. setosa*	5	3.5	1.6	0.6	*I. setosa*
4.3	3	1.1	0.1	*I. setosa*	5.1	3.8	1.9	0.4	*I. setosa*	5.1	3.8	1.9	0.4	*I. setosa*
5.8	4	1.2	0.2	*I. setosa*	4.8	3	1.4	0.3	*I. setosa*	4.8	3	1.4	0.3	*I. setosa*
5.7	4.4	1.5	0.4	*I. setosa*	5.1	3.8	1.6	0.2	*I. setosa*	5.1	3.8	1.6	0.2	*I. setosa*
5.4	3.9	1.3	0.4	*I. setosa*	4.6	3.2	1.4	0.2	*I. setosa*	4.6	3.2	1.4	0.2	*I. setosa*
5.1	3.5	1.4	0.3	*I. setosa*	5.3	3.7	1.5	0.2	*I. setosa*	5.3	3.7	1.5	0.2	*I. setosa*
5.7	3.8	1.7	0.3	*I. setosa*	5	3.3	1.4	0.2	*I. setosa*	5	3.3	1.4	0.2	*I. setosa*
5.1	3.8	1.5	0.3	*I. setosa*	7	3.2	4.7	1.4	*I. versicolor*	7	3.2	4.7	1.4	*I. versicolor*
5.4	3.4	1.7	0.2	*I. setosa*	6.4	3.2	4.5	1.5	*I. versicolor*	6.4	3.2	4.5	1.5	*I. versicolor*
5.1	3.7	1.5	0.4	*I. setosa*	6.9	3.1	4.9	1.5	*I. versicolor*	6.9	3.1	4.9	1.5	*I. versicolor*
4.6	3.6	1	0.2	*I. setosa*	5.5	2.3	4	1.3	*I. versicolor*	5.5	2.3	4	1.3	*I. versicolor*
5.1	3.3	1.7	0.5	*I. setosa*	6.5	2.8	4.6	1.5	*I. versicolor*	6.5	2.8	4.6	1.5	*I. versicolor*
4.8	3.4	1.9	0.2	*I. setosa*	5.7	2.8	4.5	1.3	*I. versicolor*	5.7	2.8	4.5	1.3	*I. versicolor*
5	3	1.6	0.2	*I. setosa*	6.3	3.3	4.7	1.6	*I. versicolor*	6.3	3.3	4.7	1.6	*I. versicolor*
5	3.4	1.6	0.4	*I. setosa*	4.9	2.4	3.3	1	*I. versicolor*	4.9	2.4	3.3	1	*I. versicolor*
5.2	3.5	1.5	0.2	*I. setosa*	6.6	2.9	4.6	1.3	*I. versicolor*	6.6	2.9	4.6	1.3	*I. versicolor*
5.2	3.4	1.4	0.2	*I. setosa*	5.2	2.7	3.9	1.4	*I. versicolor*	5.2	2.7	3.9	1.4	*I. versicolor*
4.7	3.2	1.6	0.2	*I. setosa*	5	2	3.5	1	*I. versicolor*	5	2	3.5	1	*I. versicolor*
4.8	3.1	1.6	0.2	*I. setosa*	5.9	3	4.2	1.5	*I. versicolor*	5.9	3	4.2	1.5	*I. versicolor*

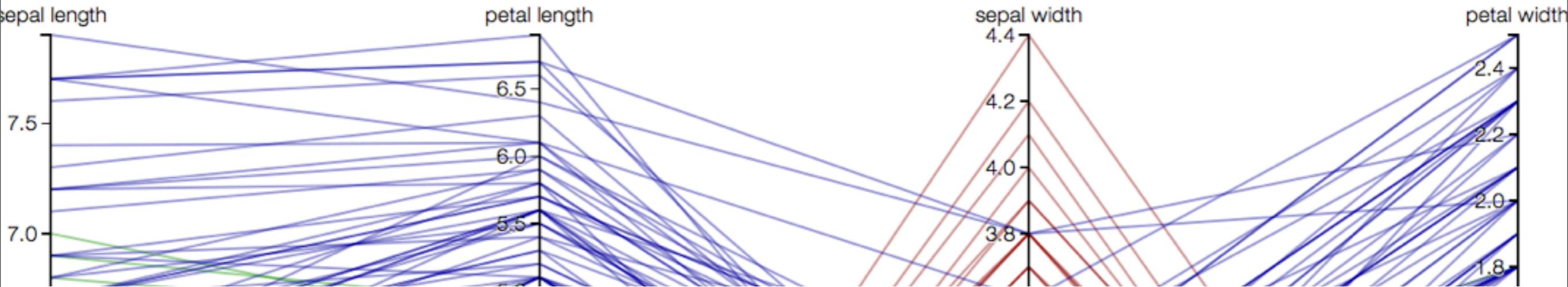
Sepal length	Sepal width	Petal length	Petal width	Species	Sepal length	Sepal width	Petal length	Petal width	Species	Sepal length	Sepal width	Petal length	Petal width	Species
5.1	3.5	1.4	0.2	*I. setosa*	5.4	3.4	1.5	0.4	*I. setosa*	5.4	3.4	1.5	0.4	*I. setosa*
4.9	3	1.4	0.2	*I. setosa*	5.2	4.1	1.5	0.1	*I. setosa*	5.2	4.1	1.5	0.1	*I. setosa*
4.7	3.2	1.3	0.2	*I. setosa*	5.5	4.2	1.4	0.2	*I. setosa*	5.5	4.2	1.4	0.2	*I. setosa*
4.6	3.1	1.5	0.2	*I. setosa*	4.9	3.1	1.5	0.2	*I. setosa*	4.9	3.1	1.5	0.2	*I. setosa*
5	3.6	1.4	0.2	*I. setosa*	5	3.2	1.2	0.2	*I. setosa*	5	3.2	1.2	0.2	*I. setosa*
5.4	3.9	1.7	0.4	*I. setosa*	5.5	3.5	1.3	0.2	*I. setosa*	5.5	3.5	1.3	0.2	*I. setosa*
4.6	3.4	1.4	0.3	*I. setosa*	4.9	3.6	1.4	0.1	*I. setosa*	4.9	3.6	1.4	0.1	*I. setosa*
5	3.4	1.5	0.2	*I. setosa*	4.4	3	1.3	0.2	*I. setosa*	4.4	3	1.3	0.2	*I. setosa*
4.4	2.9	1.4	0.2	*I. setosa*	5.1	3.4	1.5	0.2	*I. setosa*	5.1	3.4	1.5	0.2	*I. setosa*
4.9	3.1	1.5	0.1	*I. setosa*	5	3.5	1.3	0.3	*I. setosa*	5	3.5	1.3	0.3	*I. setosa*
5.4	3.7	1.5	0.2	*I. setosa*	4.5	2.3	1.3	0.3	*I. setosa*	4.5	2.3	1.3	0.3	*I. setosa*
4.8	3.4	1.6	0.2	*I. setosa*	4.4	3.2	1.3	0.2	*I. setosa*	4.4	3.2	1.3	0.2	*I. setosa*
4.8	3	1.4	0.1	*I. setosa*	5	3.5	1.6	0.6	*I. setosa*	5	3.5	1.6	0.6	*I. setosa*
4.3	3	1.1	0.1	*I. setosa*	5.1	3.8	1.9	0.4	*I. setosa*	5.1	3.8	1.9	0.4	*I. setosa*
5.8	4	1.2	0.2	*I. setosa*	4.8	3	1.4	0.3	*I. setosa*	4.8	3	1.4	0.3	*I. setosa*
5.7	4.4	1.5	0.4	*I. setosa*	5.1	3.8	1.6	0.2	*I. setosa*	5.1	3.8	1.6	0.2	*I. setosa*
5.4	3.9	1.3	0.4	*I. setosa*	4.6	3.2	1.4	0.2	*I. setosa*	4.6	3.2	1.4	0.2	*I. setosa*
5.1	3.5	1.4	0.3	*I. setosa*	5.3	3.7	1.5	0.2	*I. setosa*	5.3	3.7	1.5	0.2	*I. setosa*
5.7	3.8	1.7	0.3	*I. setosa*	5	3.3	1.4	0.2	*I. setosa*	5	3.3	1.4	0.2	*I. setosa*
5.1	3.8	1.5	0.3	*I. setosa*	7	3.2	4.7	1.4	*I. versicolor*	7	3.2	4.7	1.4	*I. versicolor*
5.4	3.4	1.7	0.2	*I. setosa*	6.4	3.2	4.5	1.5	*I. versicolor*	6.4	3.2	4.5	1.5	*I. versicolor*
5.1	3.7	1.5	0.4	*I. setosa*	6.9	3.1	4.9	1.5	*I. versicolor*	6.9	3.1	4.9	1.5	*I. versicolor*
4.6	3.6	1	0.2	*I. setosa*	5.5	2.3	4	1.3	*I. versicolor*	5.5	2.3	4	1.3	*I. versicolor*
5.1	3.3	1.7	0.5	*I. setosa*	6.5	2.8	4.6	1.5	*I. versicolor*	6.5	2.8	4.6	1.5	*I. versicolor*
4.8	3.4	1.9	0.2	*I. setosa*	5.7	2.8	4.5	1.3	*I. versicolor*	5.7	2.8	4.5	1.3	*I. versicolor*
5	3	1.6	0.2	*I. setosa*	6.3	3.3	4.7	1.6	*I. versicolor*	6.3	3.3	4.7	1.6	*I. versicolor*
5	3.4	1.6	0.4	*I. setosa*	4.9	2.4	3.3	1	*I. versicolor*	4.9	2.4	3.3	1	*I. versicolor*
5.2	3.5	1.5	0.2	*I. setosa*	6.6	2.9	4.6	1.3	*I. versicolor*	6.6	2.9	4.6	1.3	*I. versicolor*
5.2	3.4	1.4	0.2	*I. setosa*	5.2	2.7	3.9	1.4	*I. versicolor*	5.2	2.7	3.9	1.4	*I. versicolor*
4.7	3.2	1.6	0.2	*I. setosa*	5	2	3.5	1	*I. versicolor*	5	2	3.5	1	*I. versicolor*
4.8	3.1	1.6	0.2	*I. setosa*	5.9	3	4.2	1.5	*I. versicolor*	5.9	3	4.2	1.5	*I. versicolor*



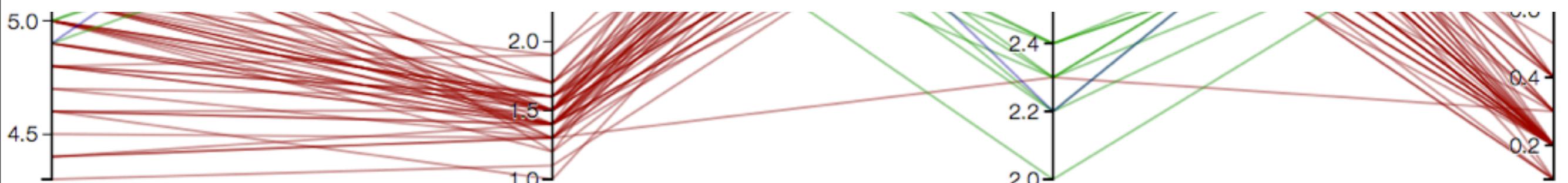


- *Iris setosa*  
- *Iris versicolor*  
- *Iris virginica*

Edgar Anderson's *Iris* data set  
parallel coordinates

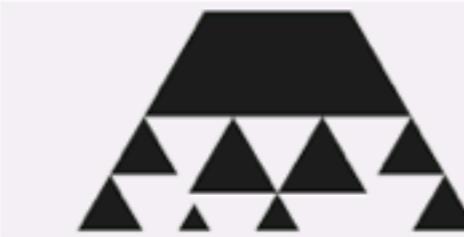
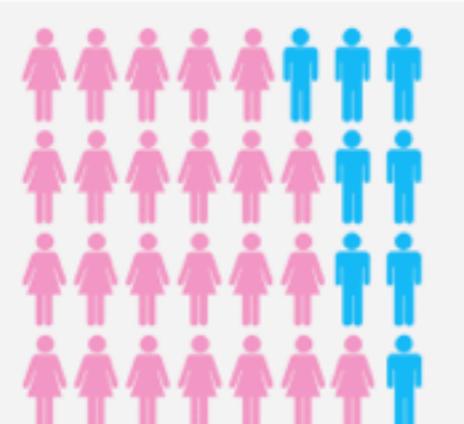
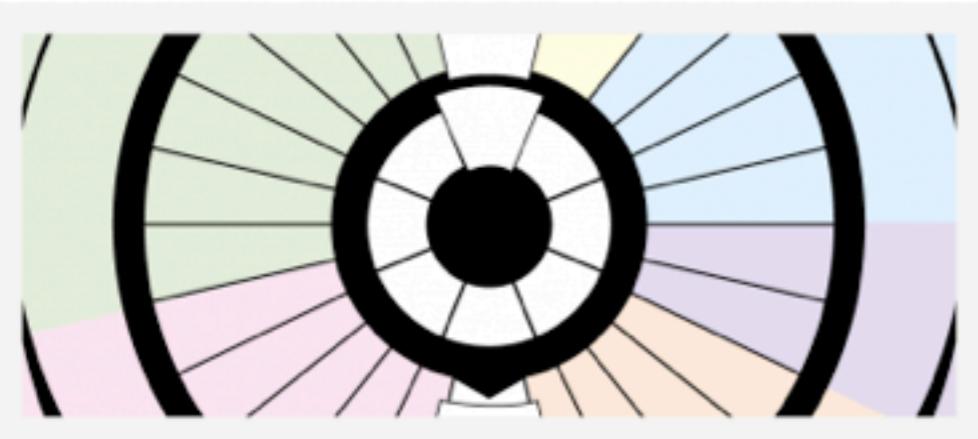
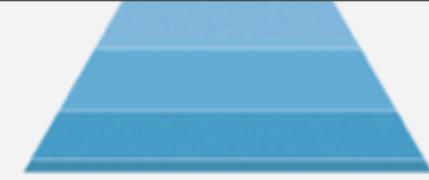


so how do we come up with these  
visual representations and  
which do we choose for a dataset?



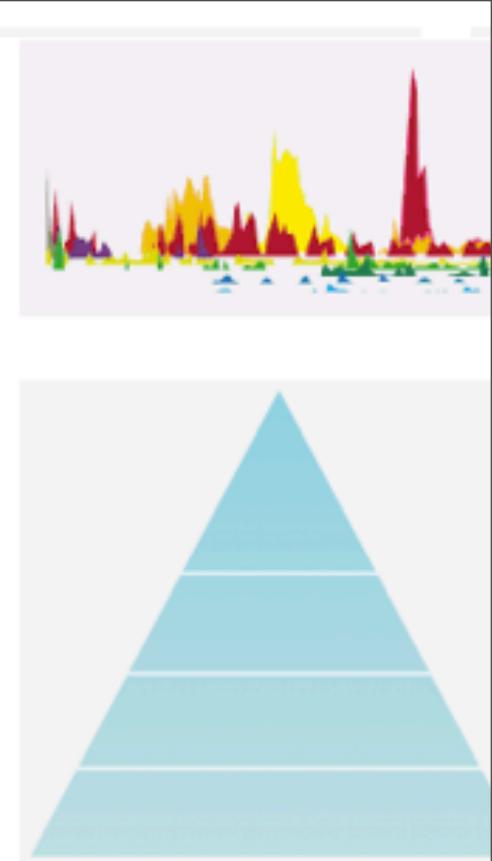
- *Iris setosa*
- *Iris versicolor*
- *Iris virginica*

Edgar Anderson's *Iris* data set  
parallel coordinates



The Stranger Remembrance of Things Pa  
*Lord of the Flies* Or  
*One Hundred*  
*His Dark Materials*  
*To Kill a Mockingbird*  
*The Handmaid's Tale*  
*The Sound and the Fury*  
*The Name of the Rose*  
*One Flew Over the Cuckoo's Nest*  
*The Adventures of Huckleberry Finn*  
*The Rings of Power*  
*To Kill a Mockingbird*

# Visual and Data Dimensions



so you have a dataset...

$\{x_1, x_2, x_3, x_4, \dots\}$        $x_i$  is...

so you have a dataset...

$\{x_1, x_2, x_3, x_4, \dots\}$   $x_i$  is...

$\{1, 200, 5, 6, \dots\}$  integral

$\{1.0, 2.0, 1.2, 4, \dots\}$  fixed point

$\{'a', 'b', '12c', 'd', \dots\}$  alpha(-numeric)

$\{20\%, 30\%, 1\%, 5\% \dots\}$  fractions of a population

$\{\text{pear}, \text{apple}, \text{kiwi}, \text{pineapple}, \dots\}$  categorical

$\{f(\text{pear}, \text{apple}), g(\text{apple}, \text{kiwi}), q(\text{kiwi}, \text{pineapple}) \dots\}$  relational

so you have a dataset...

$\{x_1, x_2, x_3, x_4, \dots\}$	$x_i$ is...
$\{1, 200, 5, 6, \dots\}$	integral
$\{1.0, 2.0, 1.2, 4, \dots\}$	fixed point
$\{'a', 'b', '12c', 'd', \dots\}$	alpha(-numeric)
$\{20\%, 30\%, 1\%, 5\%, \dots\}$	fractions of a population
$\{\text{apple}, \text{apple}, \text{kiwi}, \text{pineapple}, \text{pineapple}, \dots\}$	categorical
$\{f(\text{apple}, \text{apple}), g(\text{apple}, \text{kiwi}), q(\text{kiwi}, \text{pineapple}), \dots\}$	relational

objective - help the user to understand :  
relationships **among the elements of the set**

4

4

9

7

4

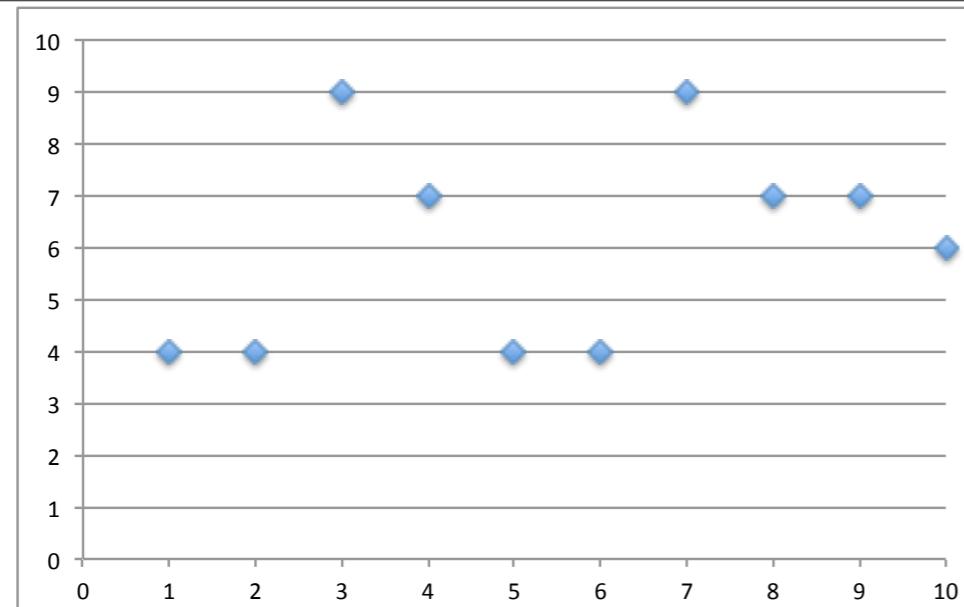
4

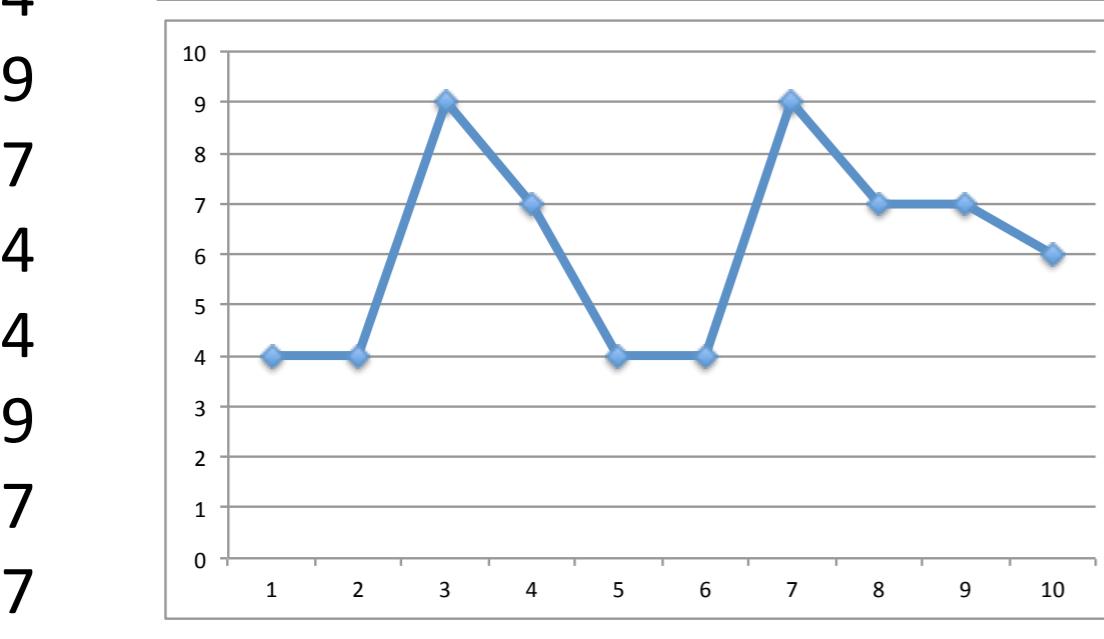
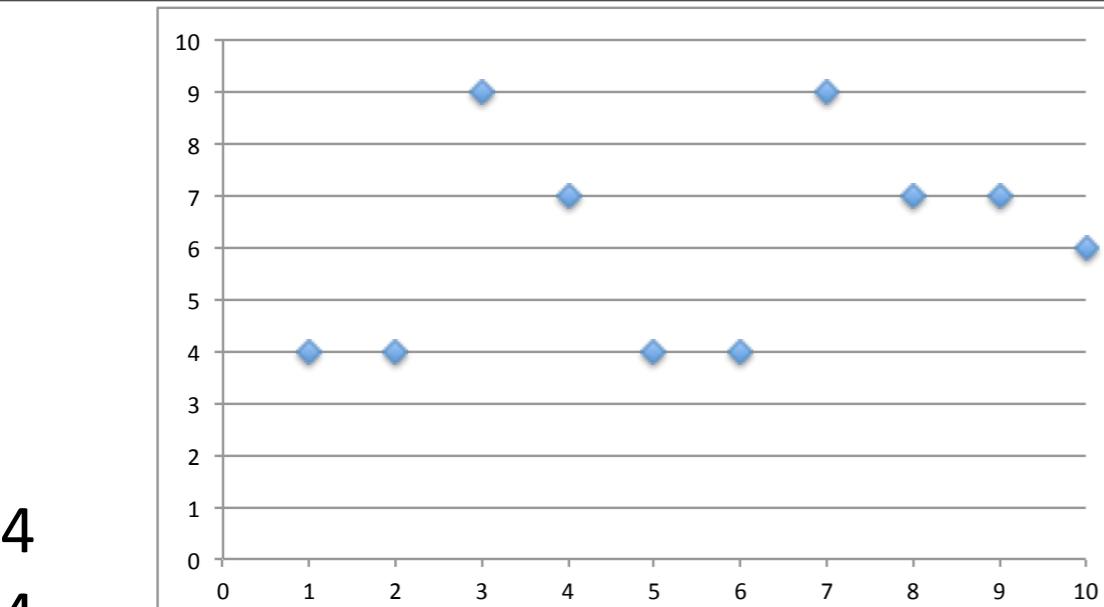
9

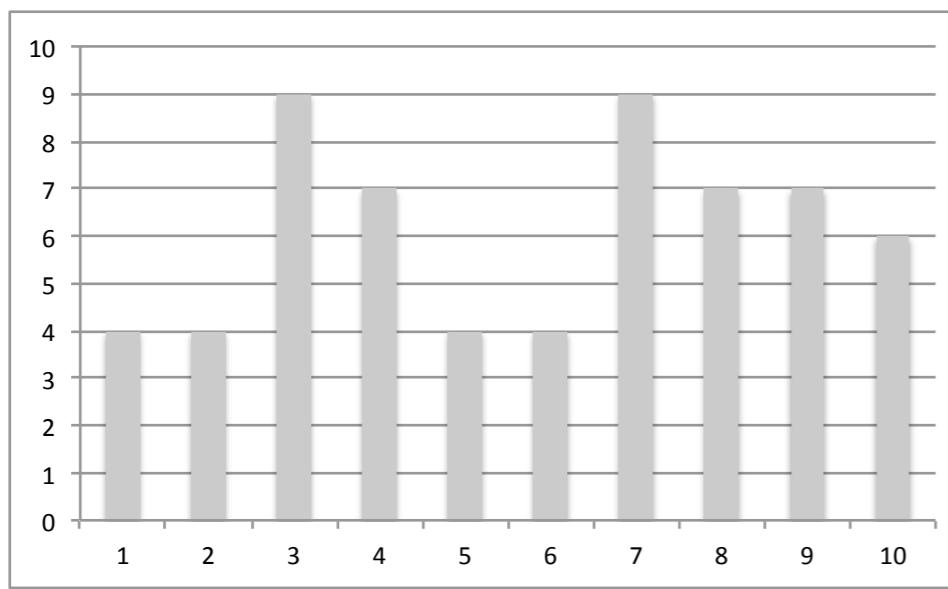
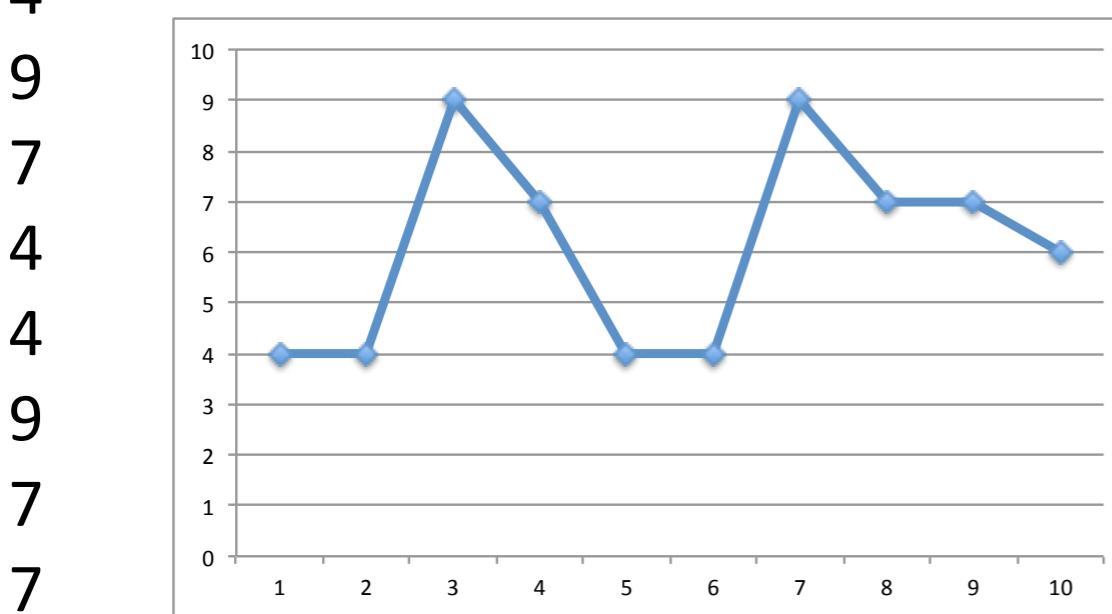
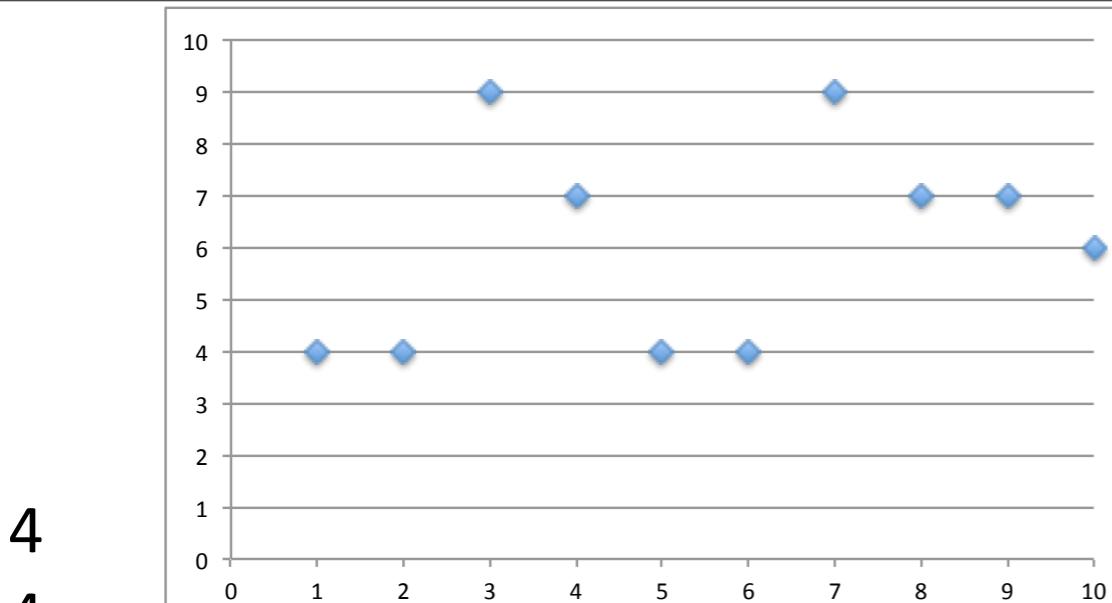
7

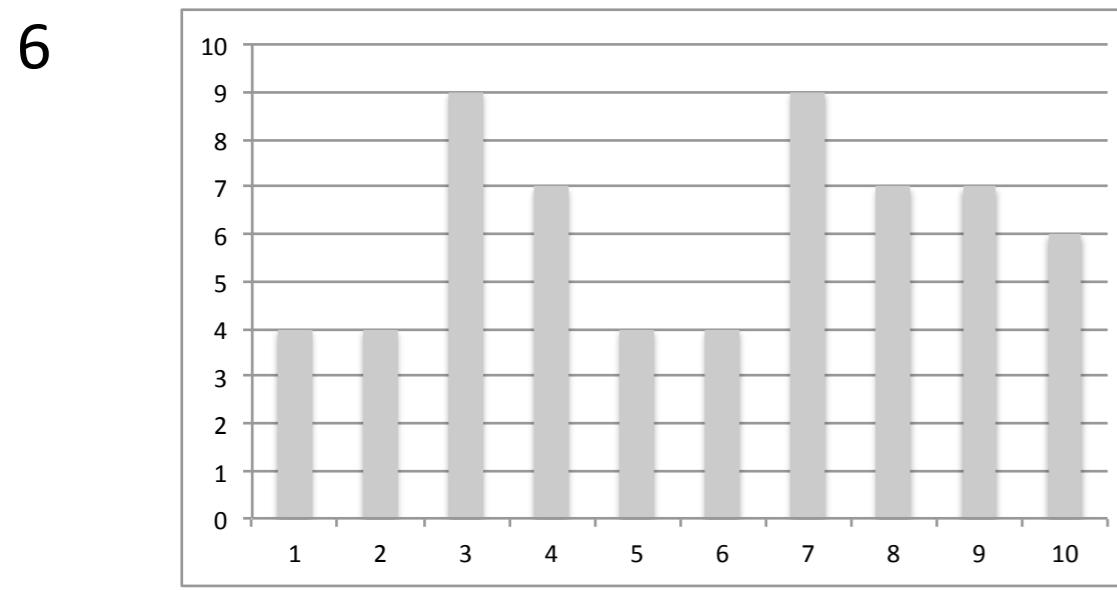
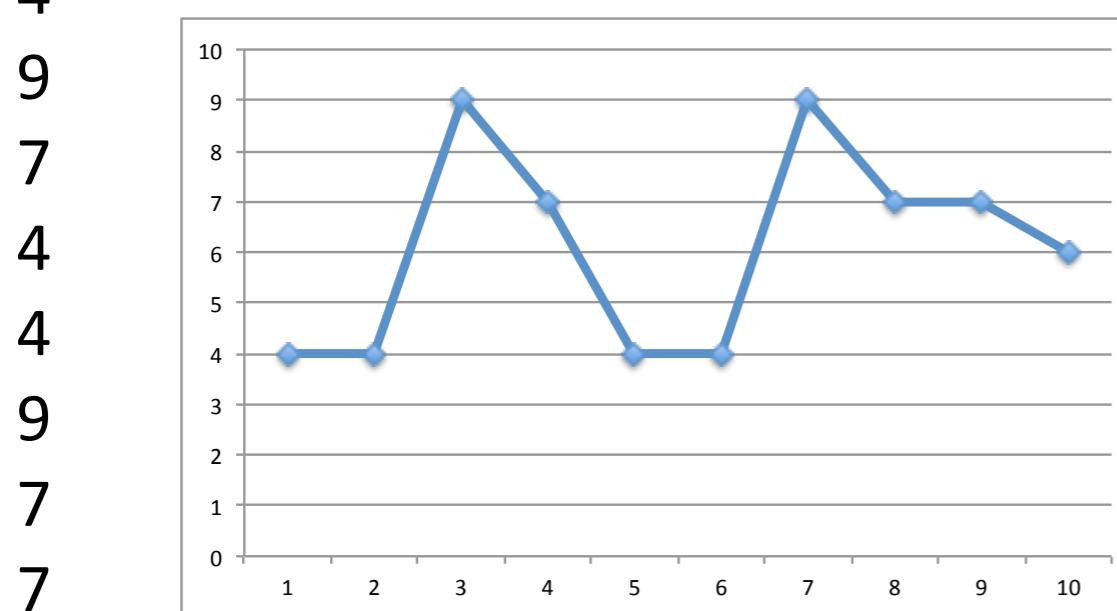
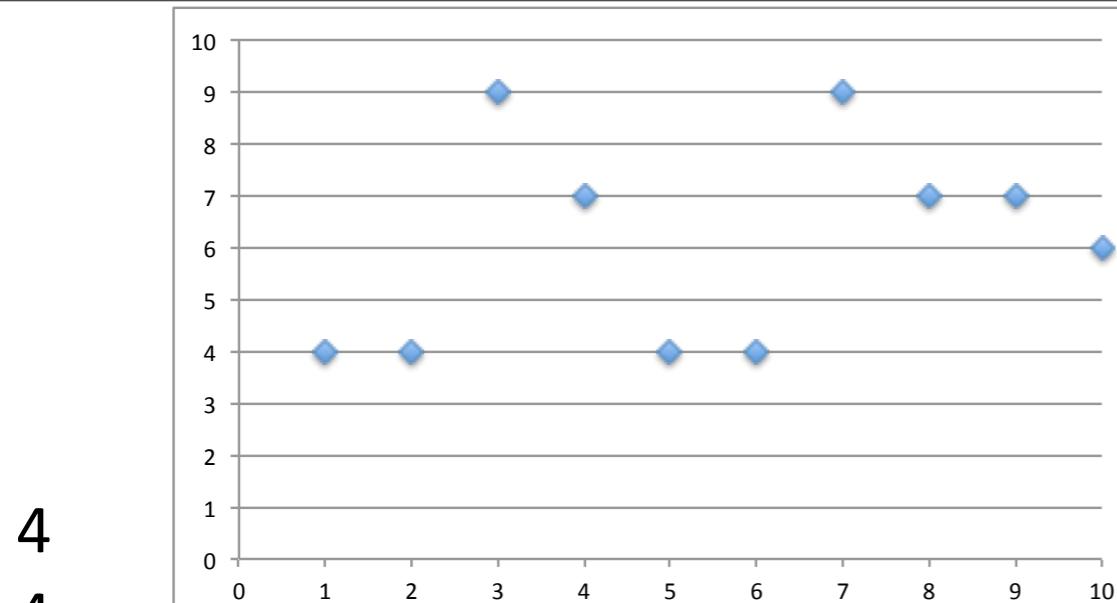
7

6



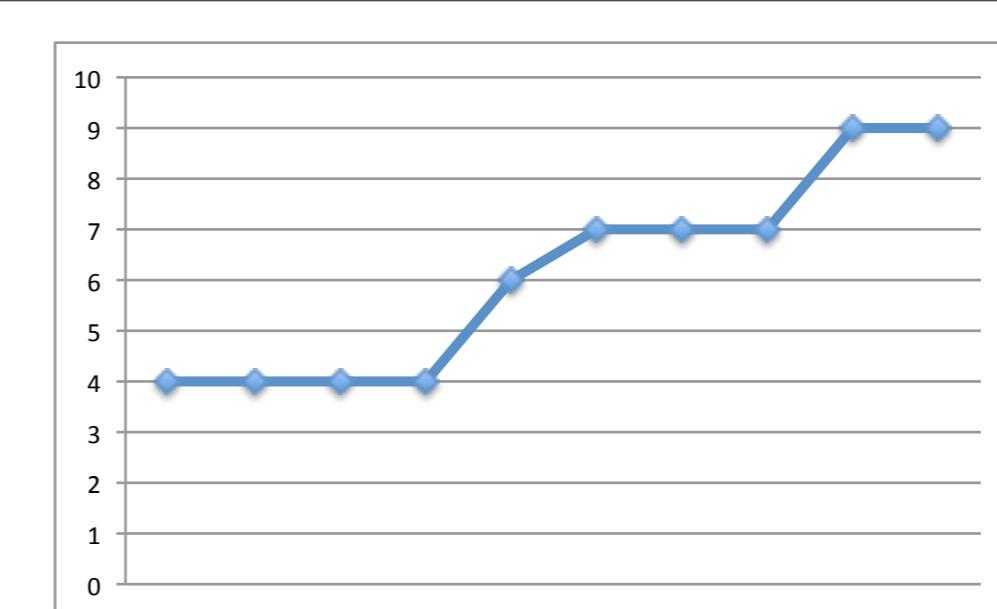
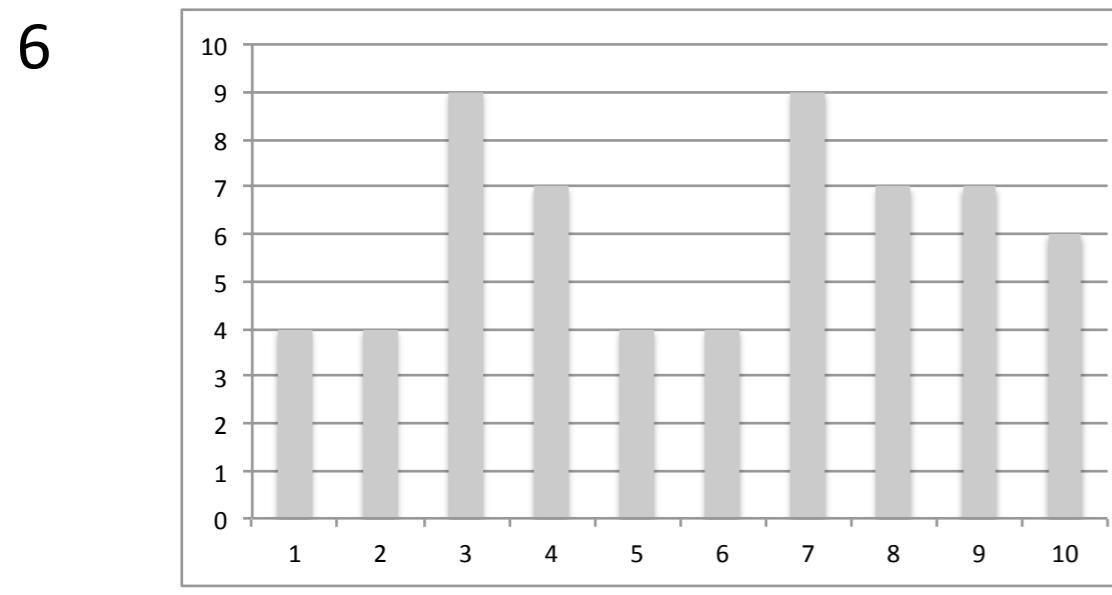
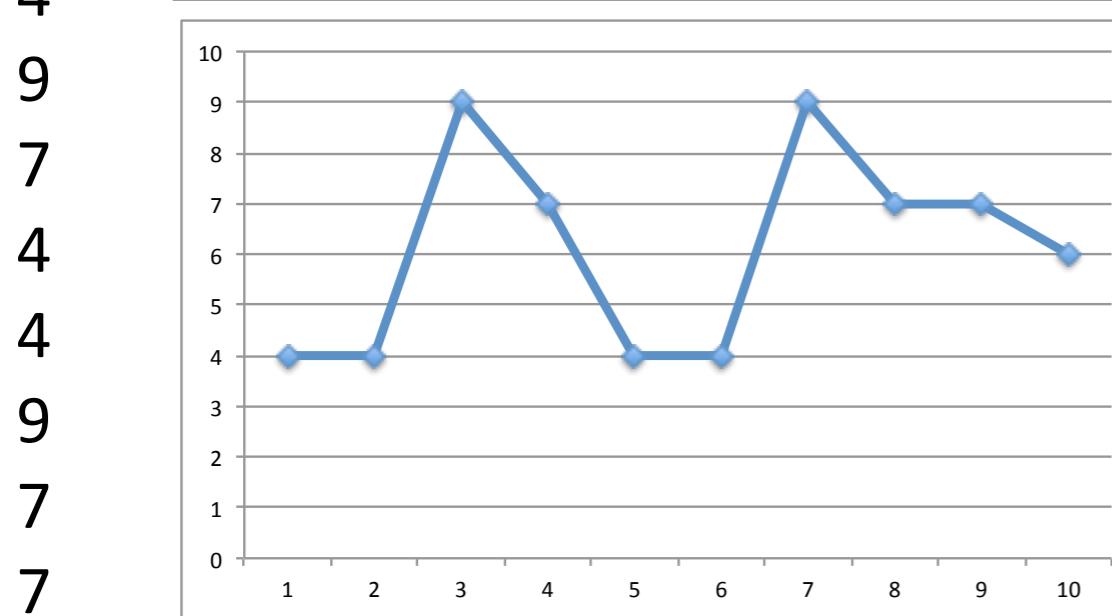
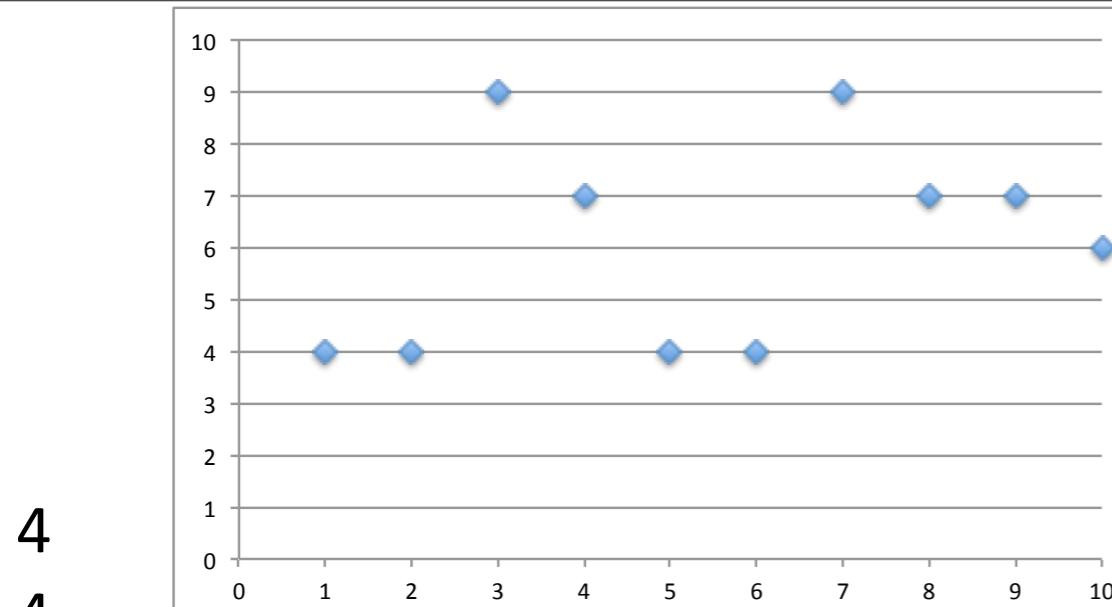






ordering significant

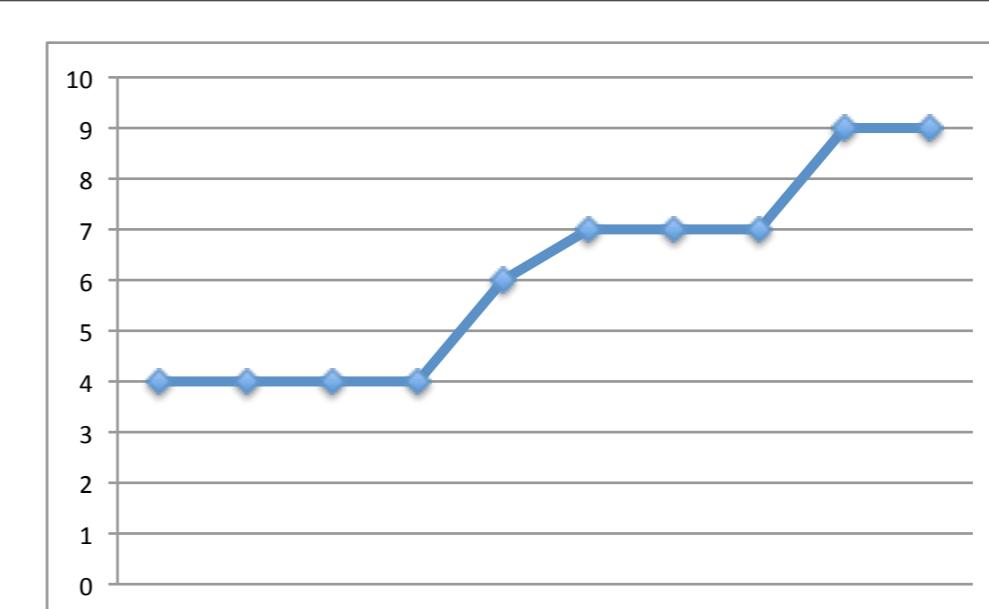
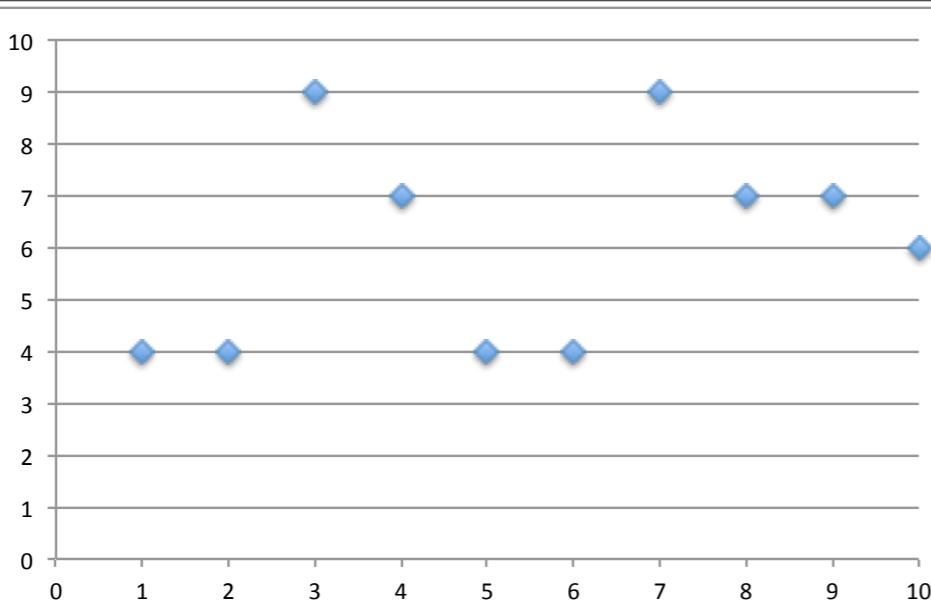
order insignificant



ordering significant

order insignificant

# histogram



4

4

9

7

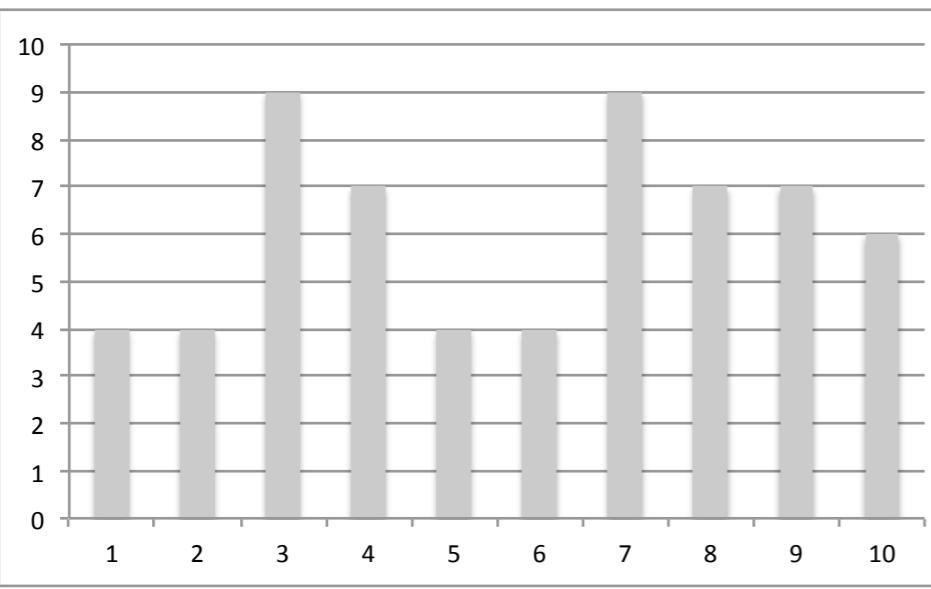
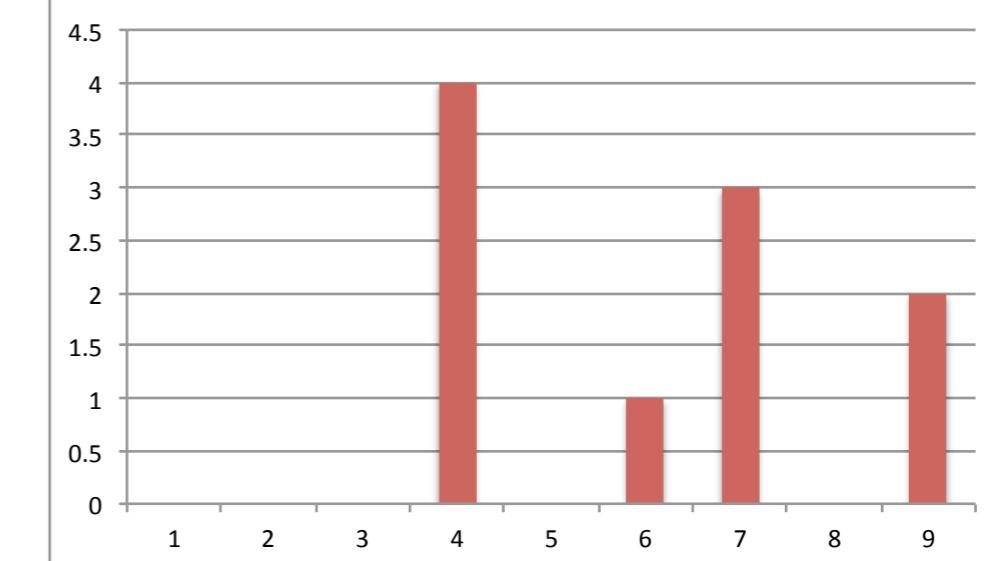
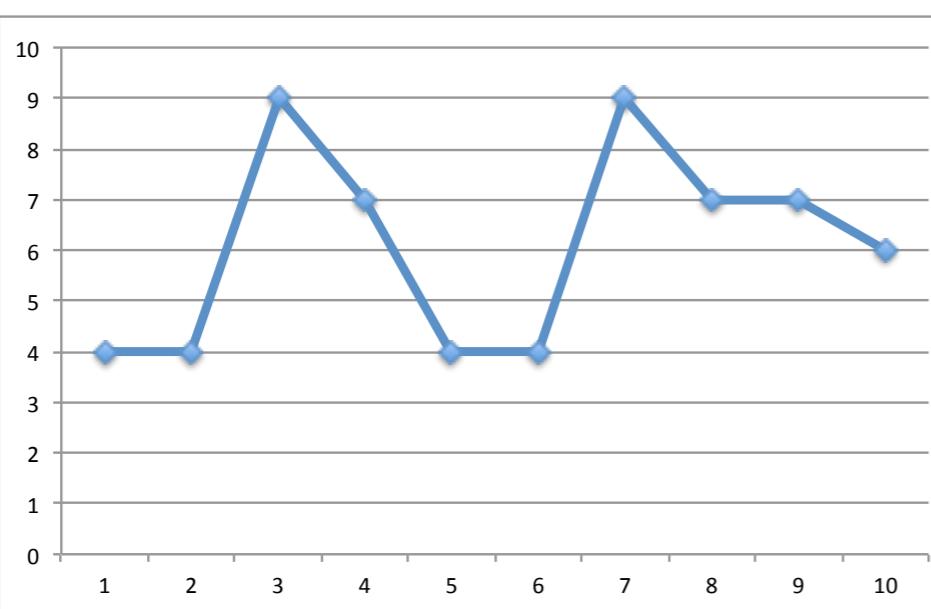
4

9

7

7

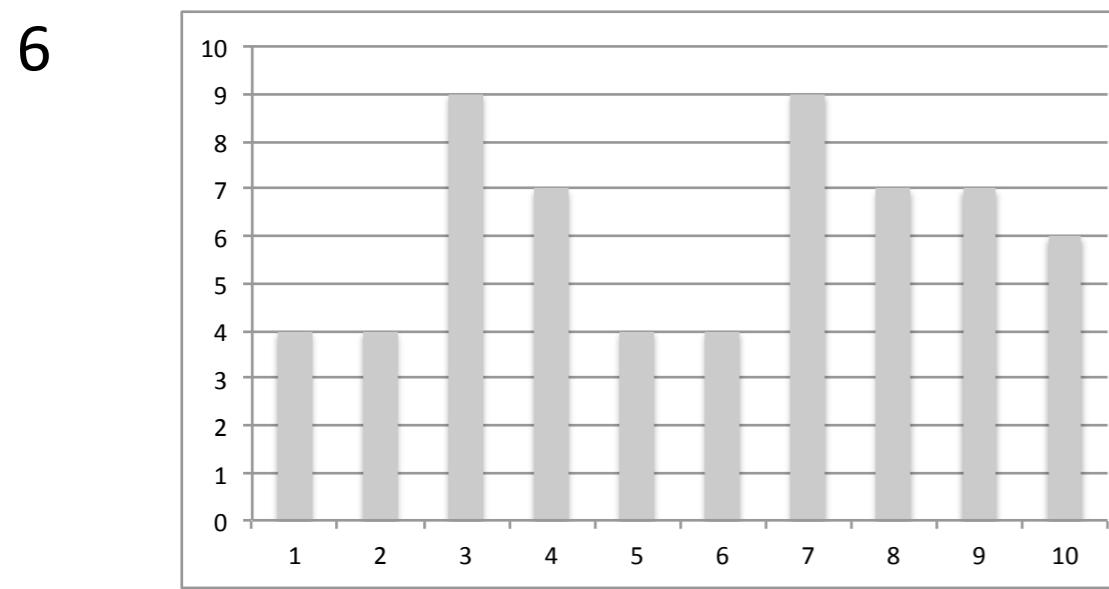
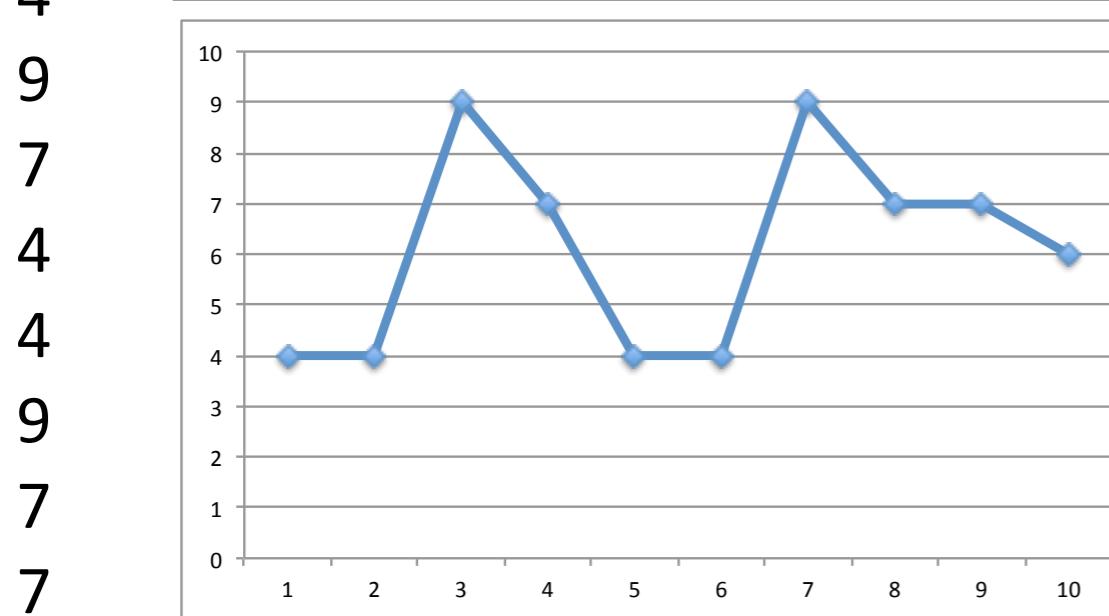
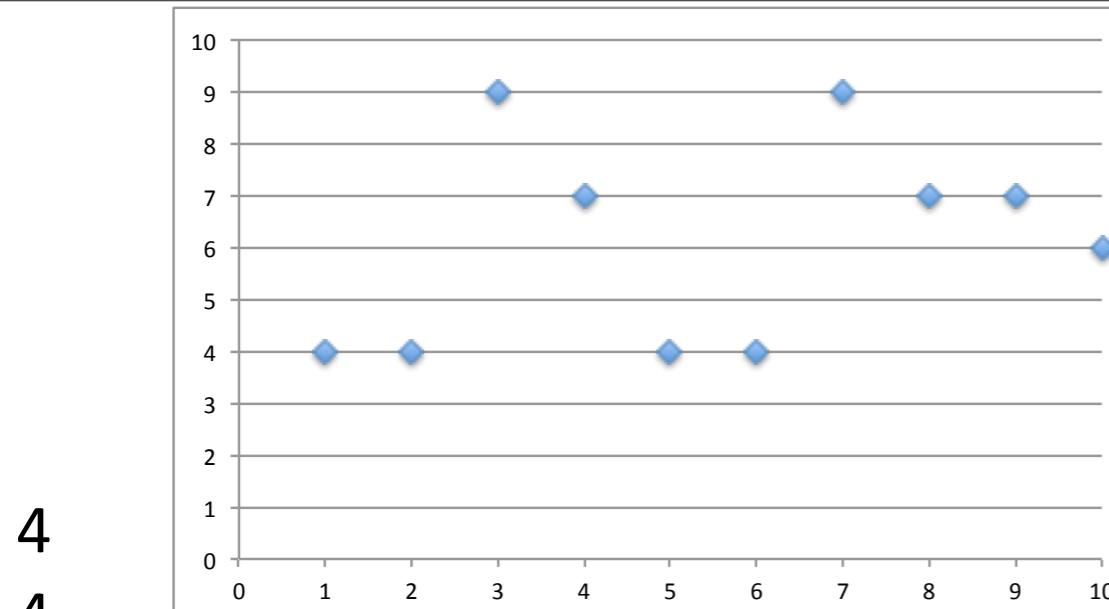
6



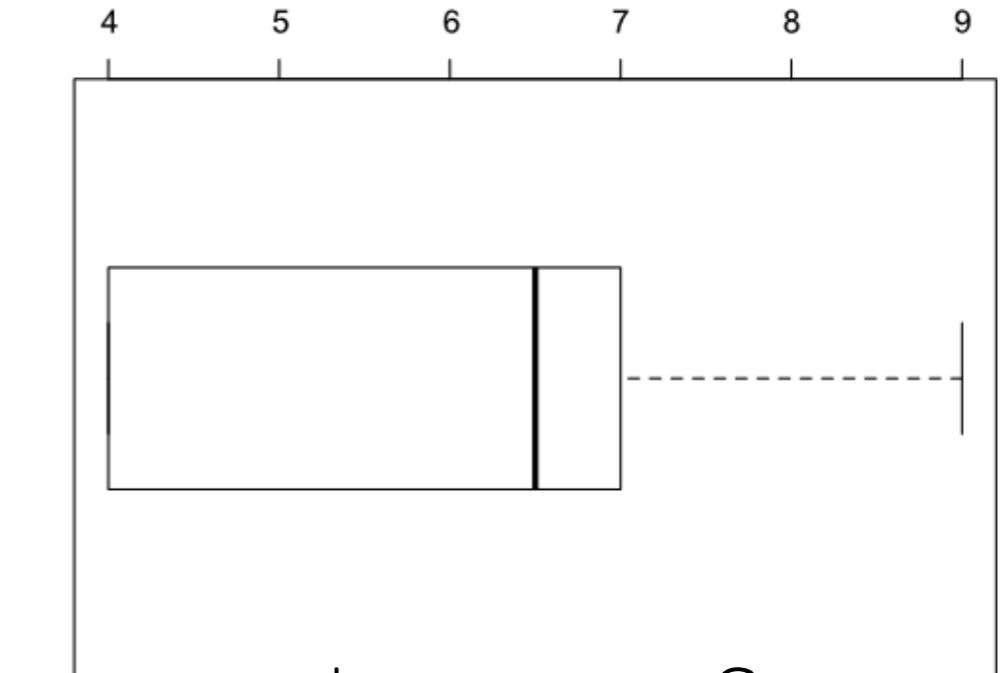
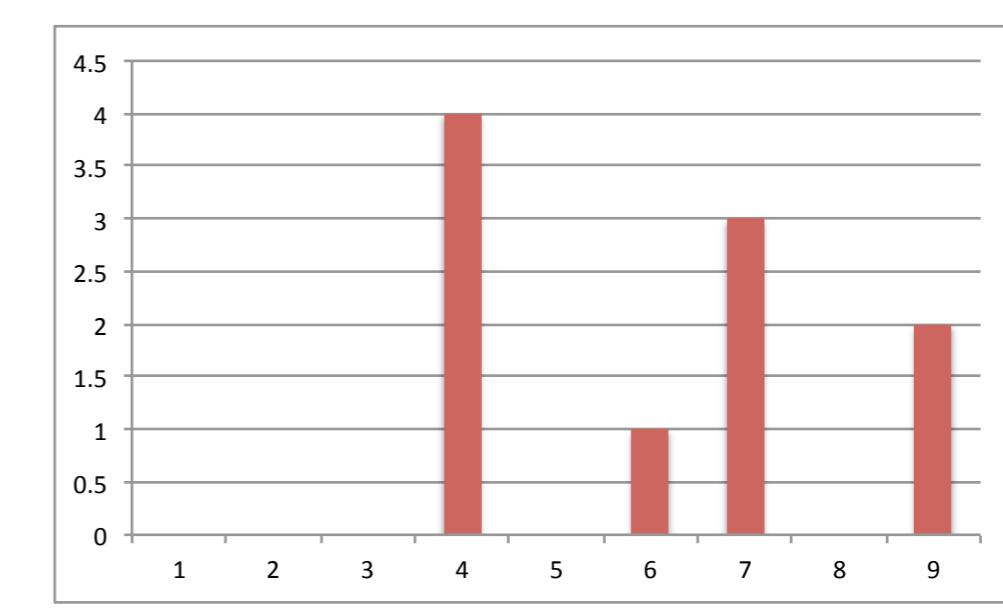
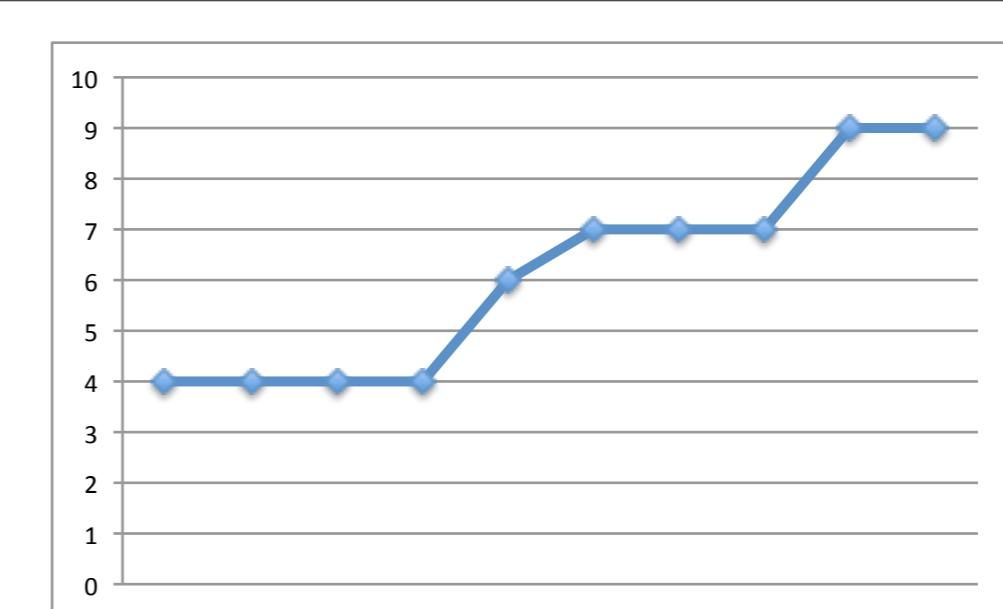
ordering significant

order insignificant

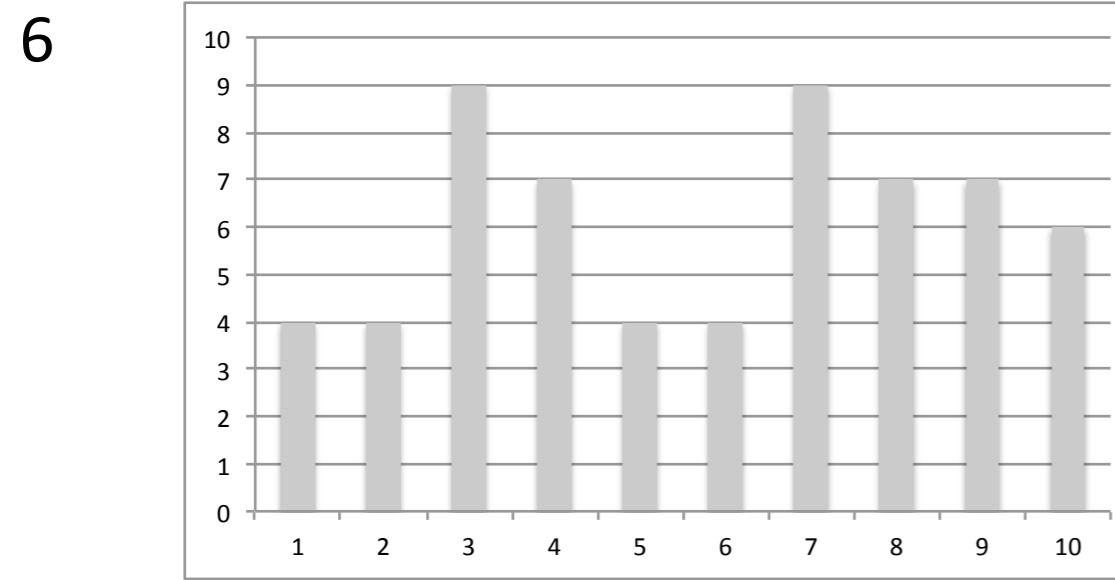
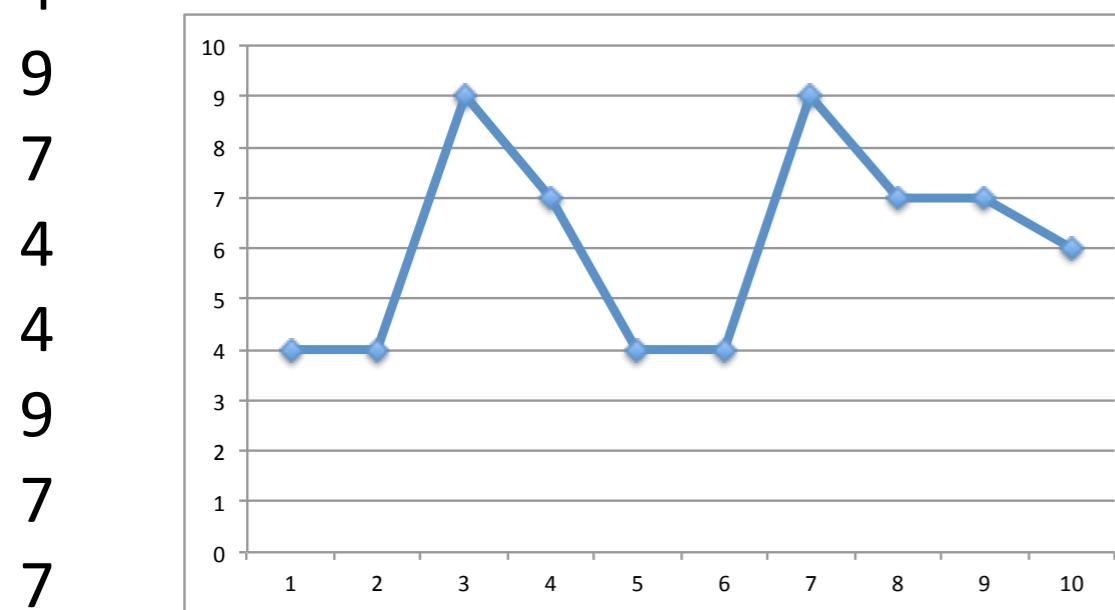
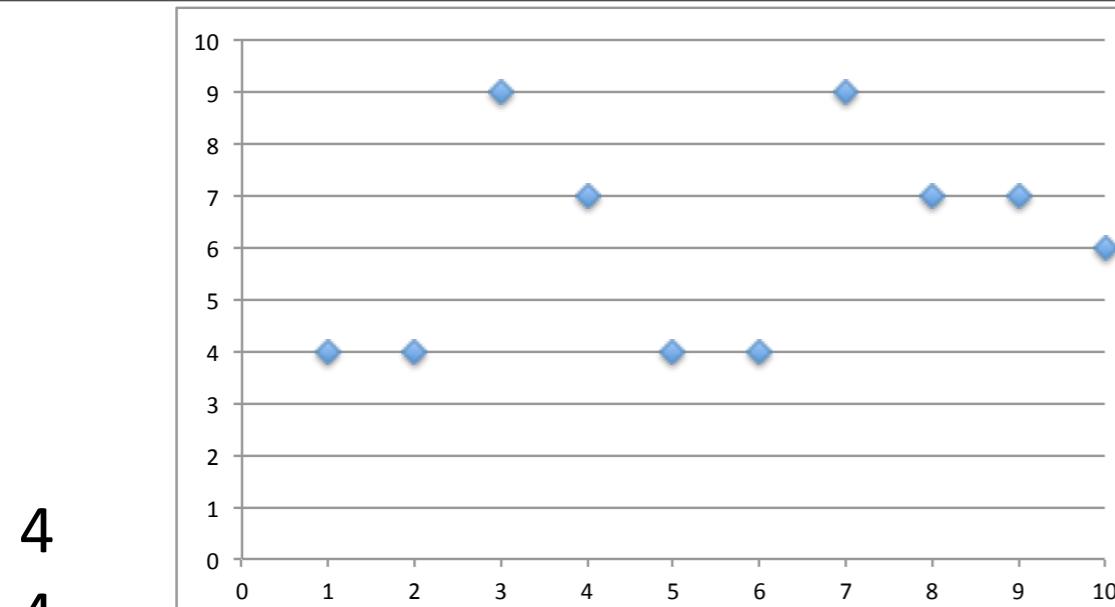
# histogram



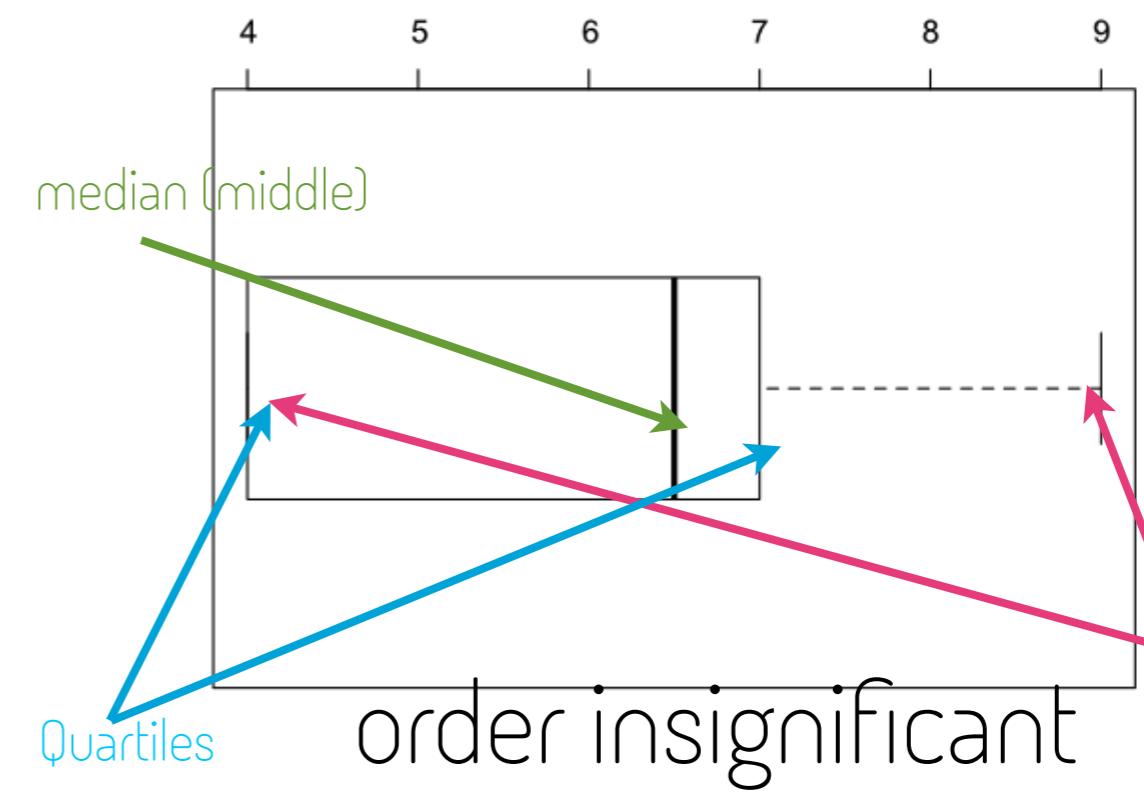
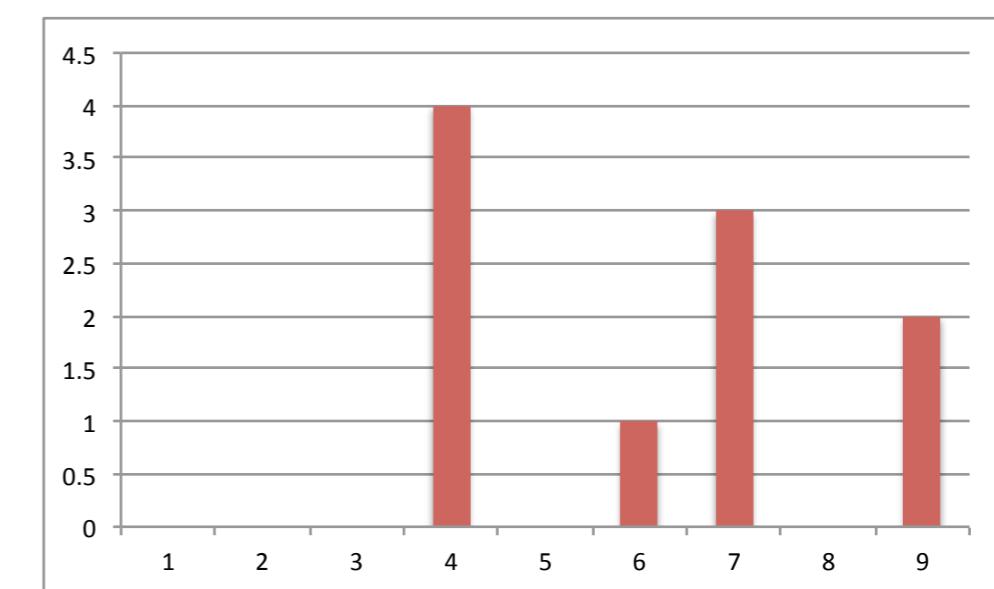
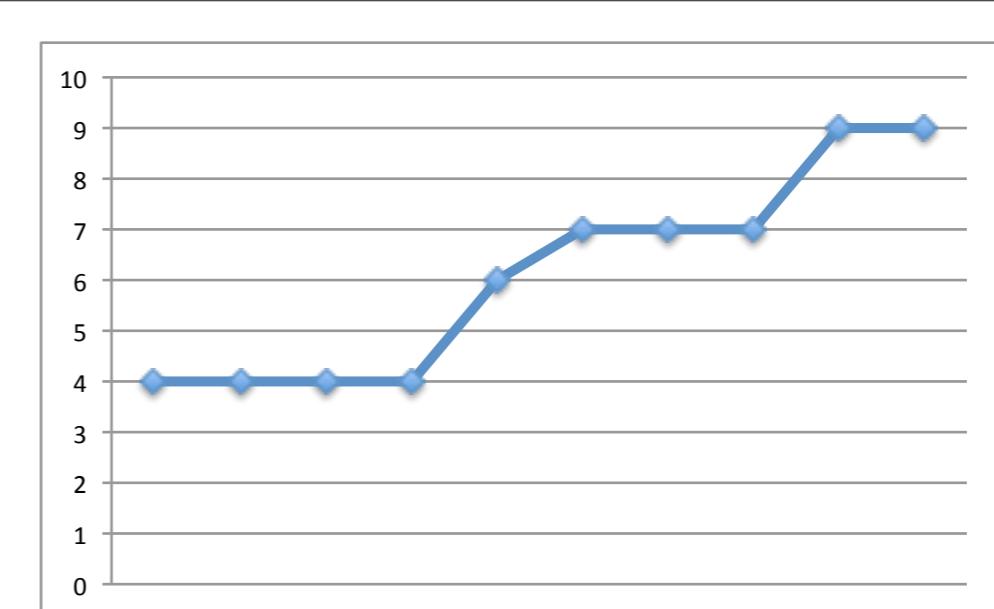
ordering significant



order insignificant



ordering significant



sorted

histogram

box & whisker

extrema  
(whiskers)

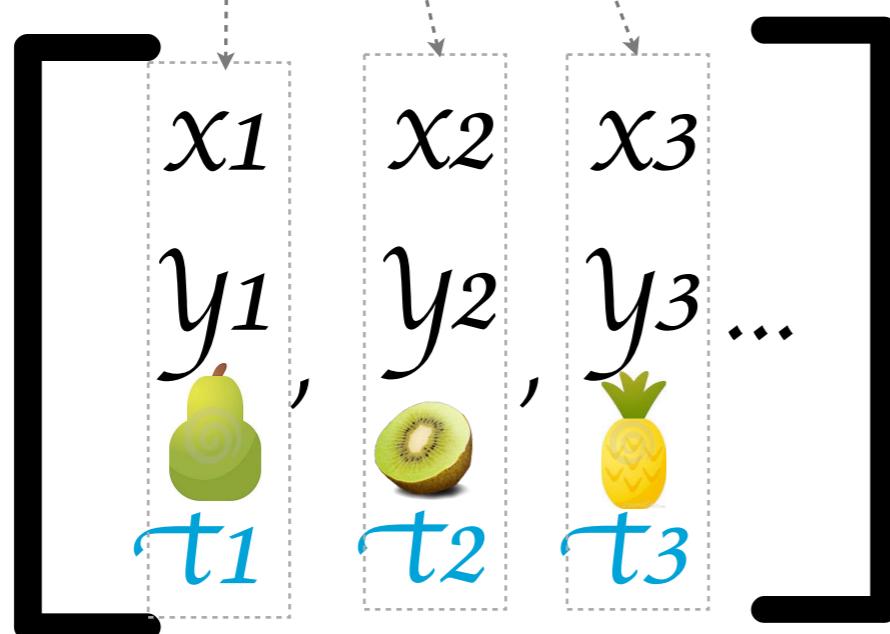
so you have a dataset...

it's probably multivariate

$x =$

$\{\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \dots\}$

$x =$



if these are observations of the  
(same) object(s) over time  
“time series”

if these are observations of different  
things at a single point in time  
“population”

if these are observations of different  
things at different points in time  
“observations”

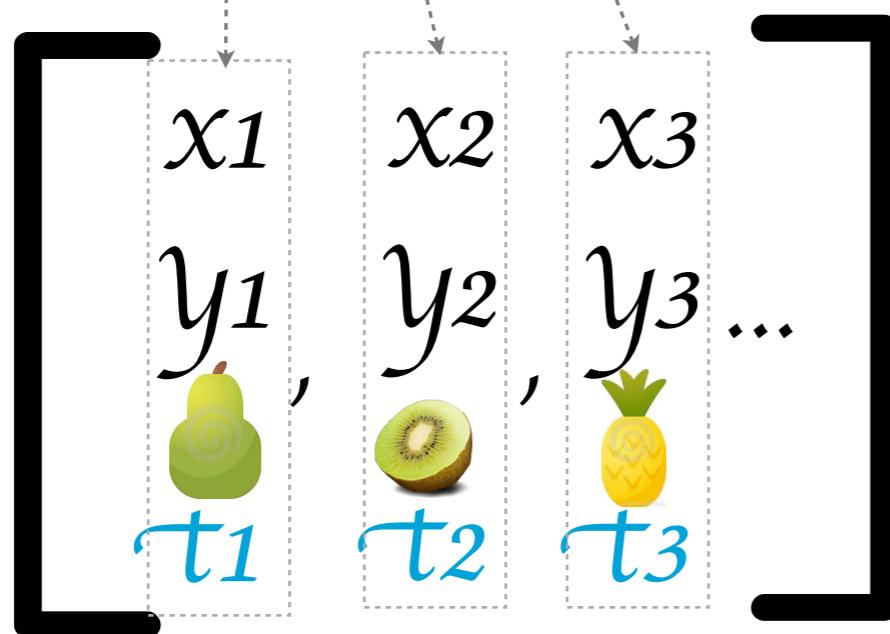
so you have a dataset...

it's probably multivariate

$\mathcal{X} =$

$\{\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4, \dots\}$

$\mathcal{X} =$



if these are observations of the  
(same) of object(s) over time  
“time series”

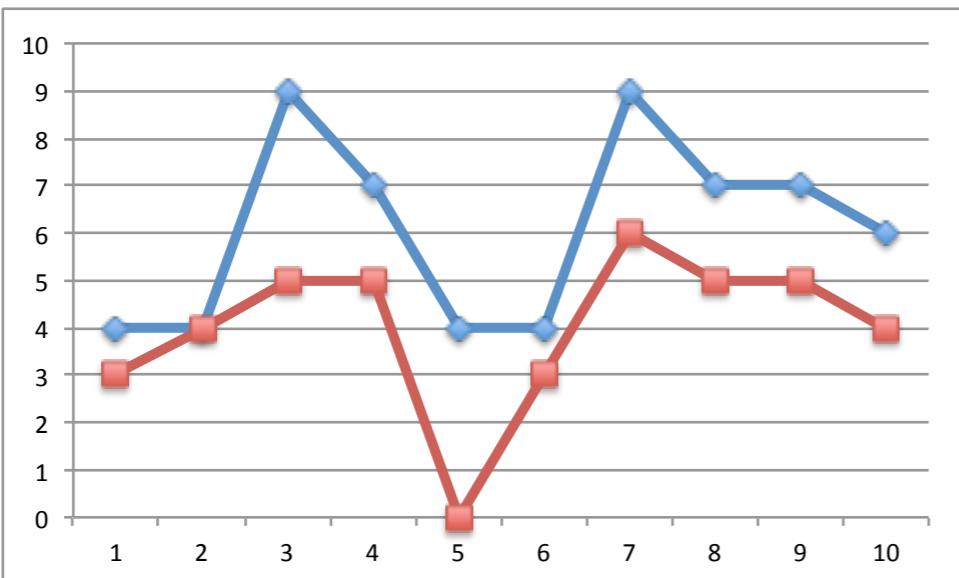
if these are observations of different  
things at a single point in time  
“population”

if these are observations of different  
things at a different points in time  
“observations”

objectives - to show :

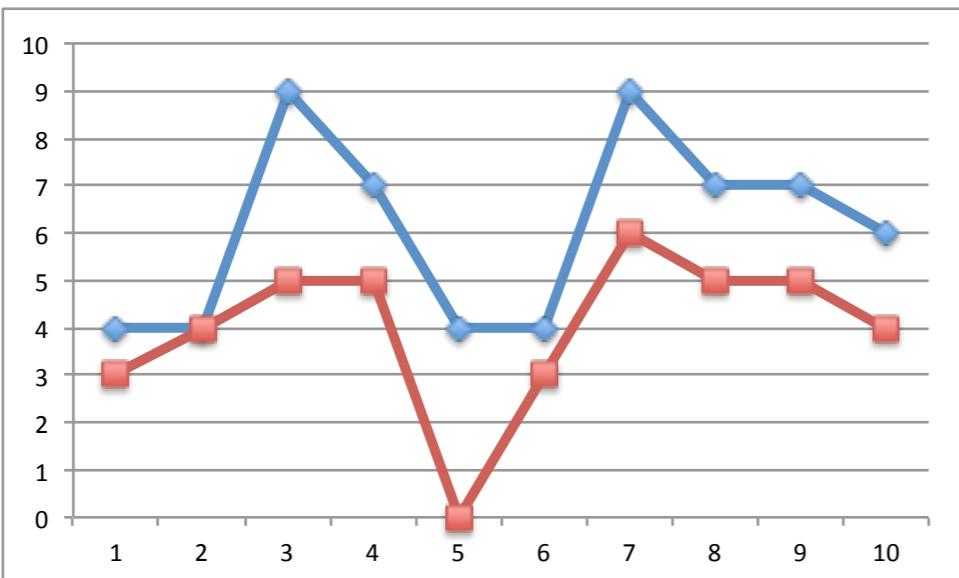
1. relations among dimensions of each example (multivariate)
2. relations among examples/observations (multidimensional)

4	3
4	4
9	5
7	5
4	0
4	3
9	6
7	5
7	5
6	4

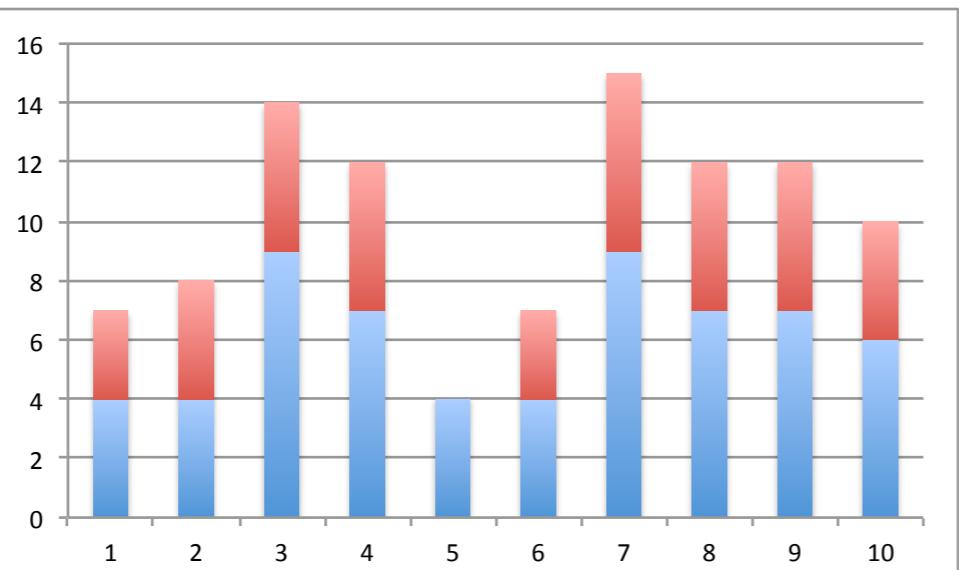


lines

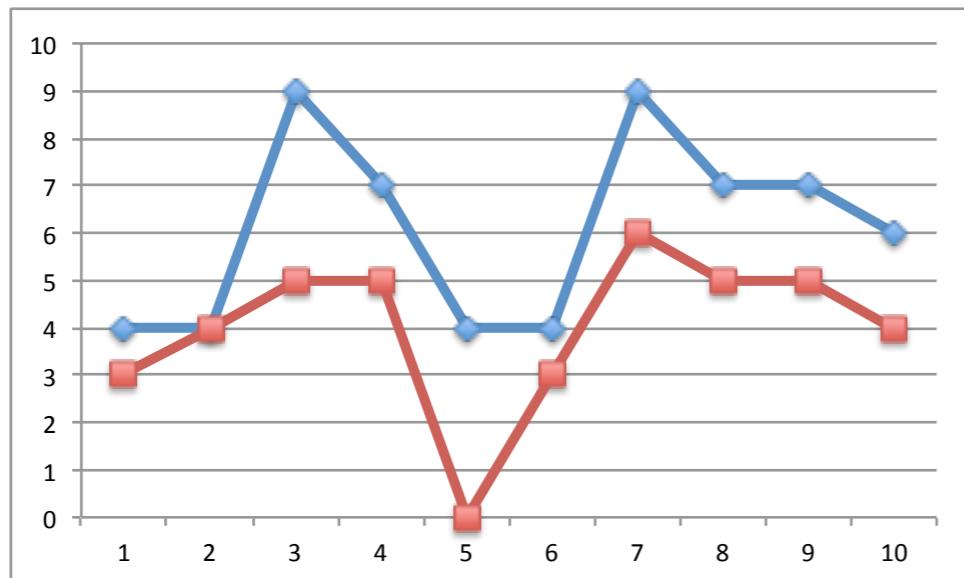
4	3
4	4
9	5
7	5
4	0
4	3
9	6
7	5
7	5
6	4



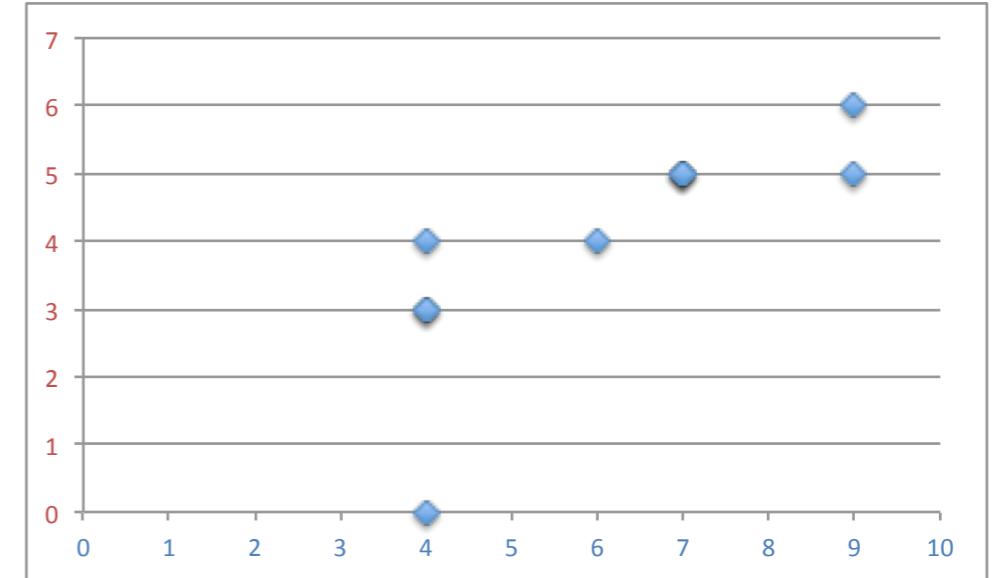
lines



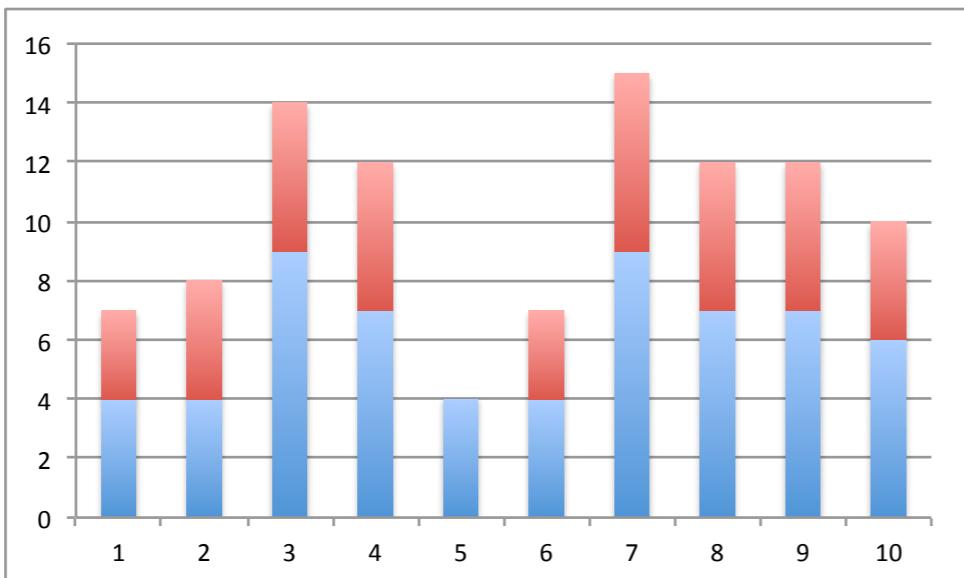
stacked bar



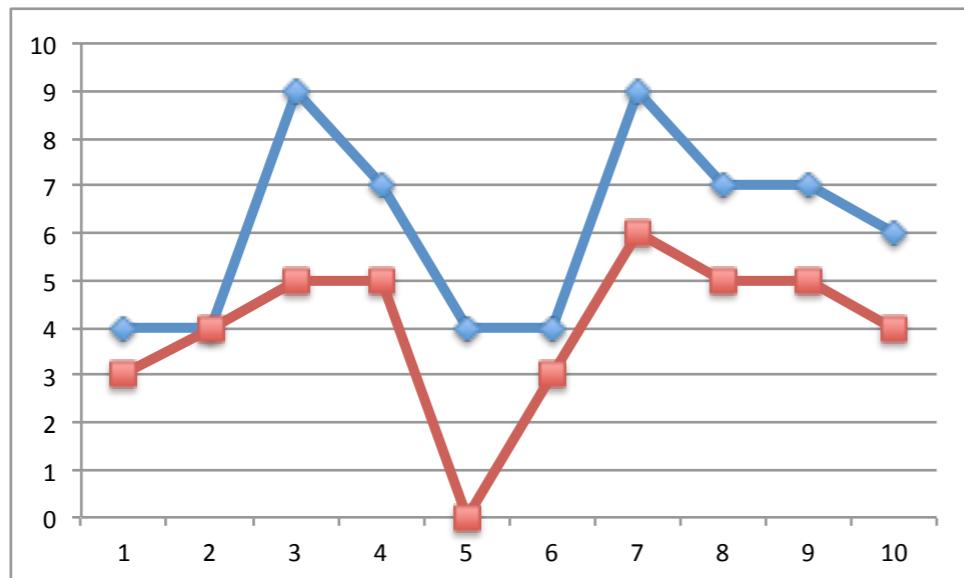
lines



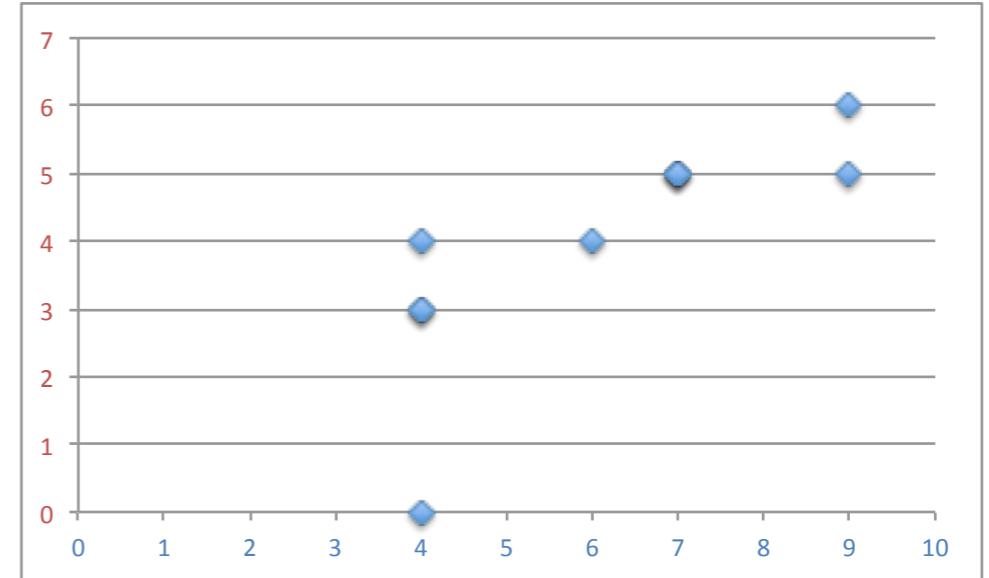
scatter



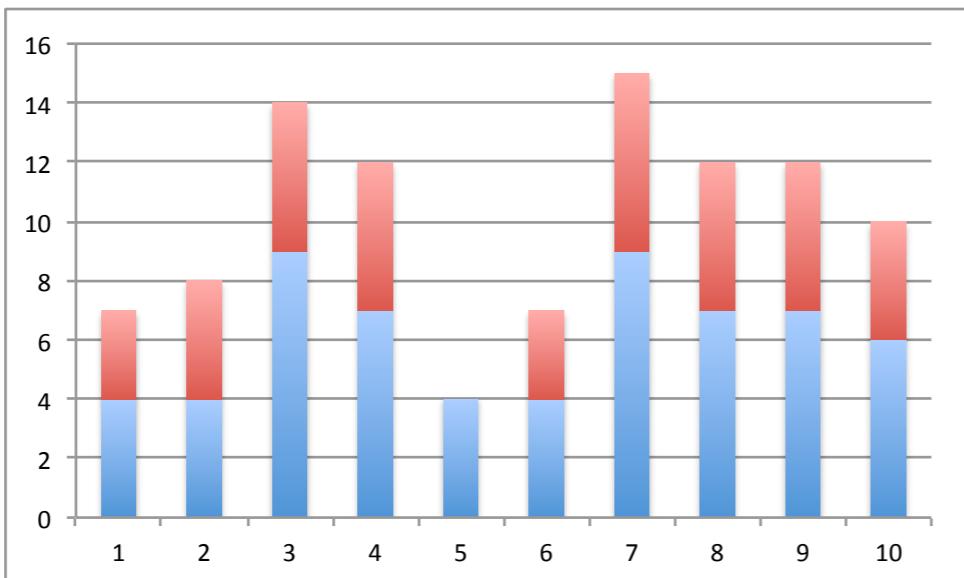
stacked bar



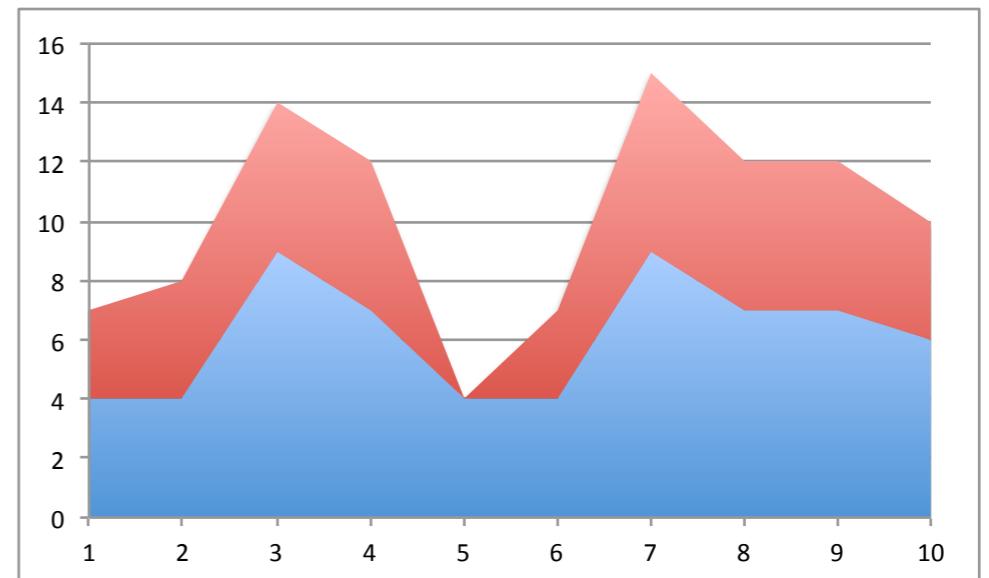
lines



scatter



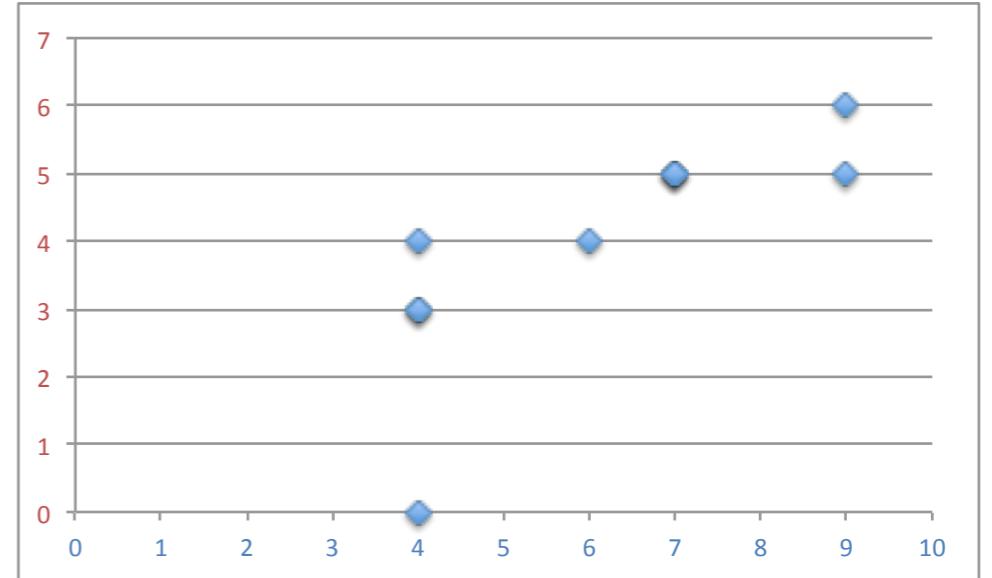
stacked bar



stacked area

## comparing dimensions

"multivariate"

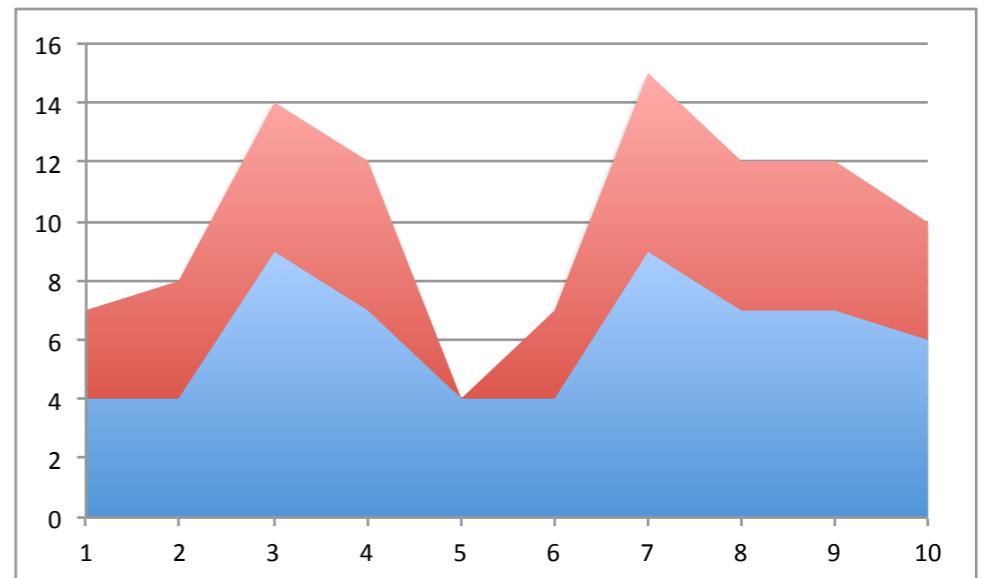


lines

scatter

## totals

"multidimensional"

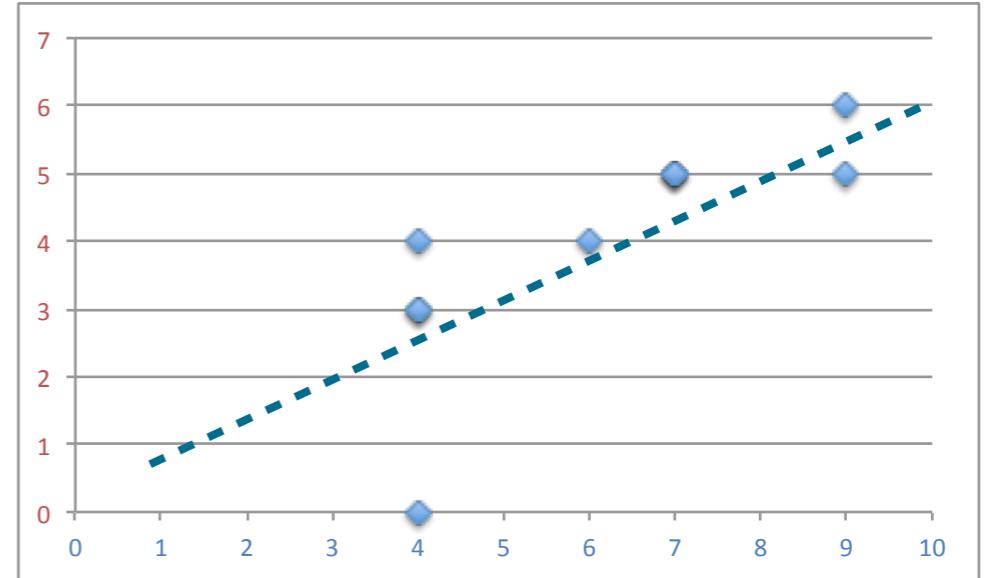


stacked bar

stacked area

## comparing dimensions

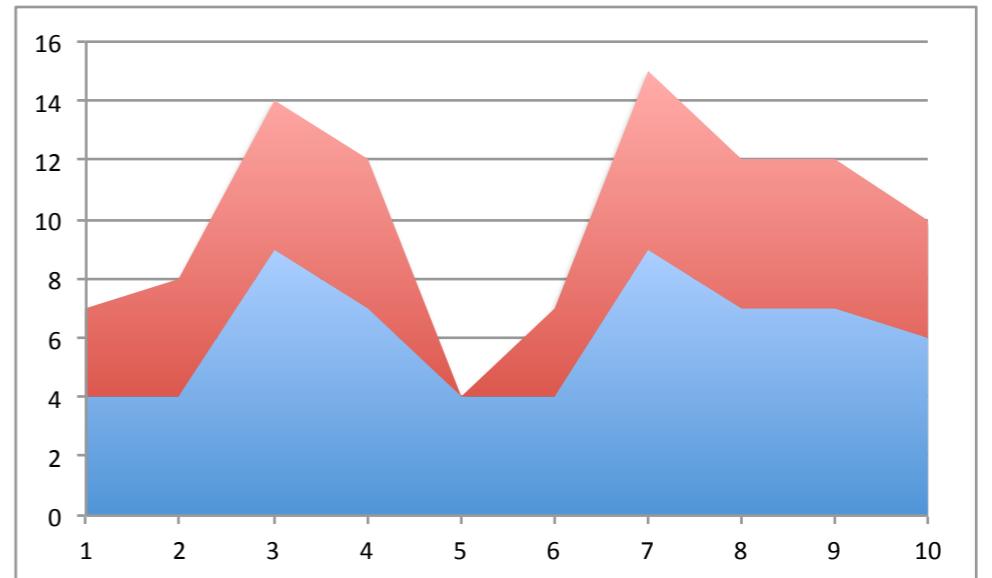
"multivariate"



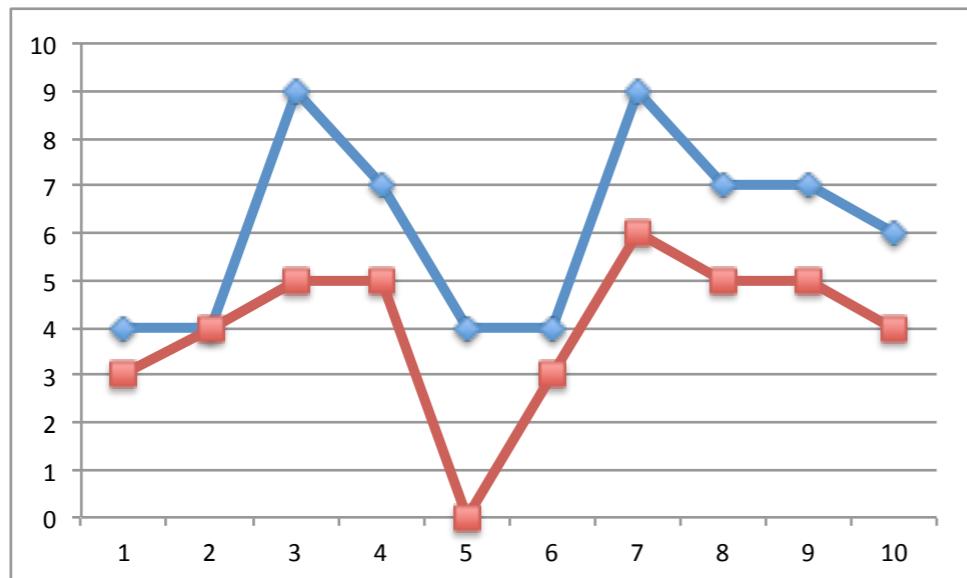
scatter

"multidimensional"

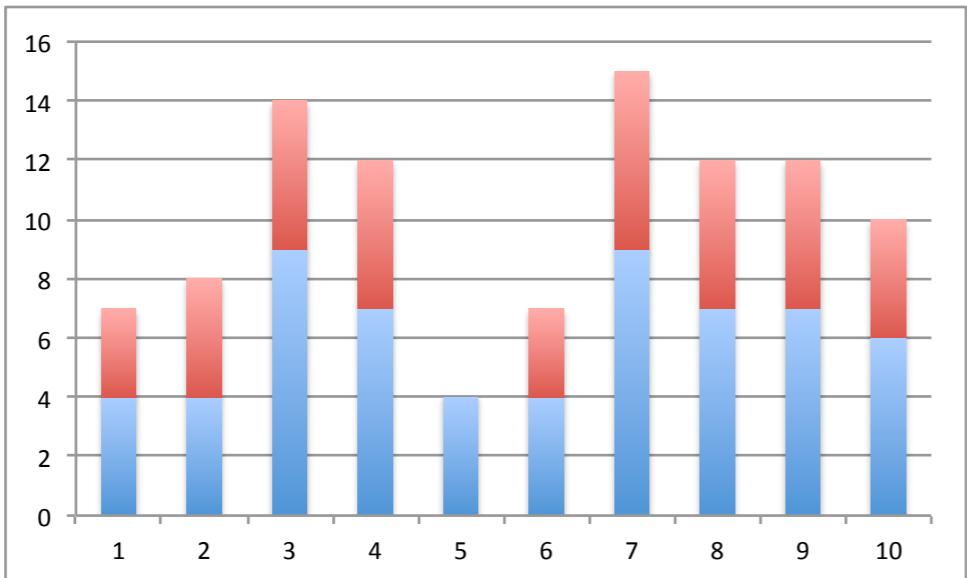
## totals



stacked area



lines



stacked bar

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

...

# *visual dimension type*

# *data dimension types*

# *visual dimension type*

position

relative location  
centrality

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

...

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

...

# *visual dimension type*

position

relative location  
centrality

shape

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

...

# *visual dimension type*

position

relative location  
centrality

shape

colour

saturation  
opacity

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

...

# *visual dimension type*

position

relative location  
centrality

shape

colour

saturation  
opacity

size

width  
height

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

...

# *visual dimension type*

position

relative location  
centrality

shape

colour

saturation  
opacity

size

width  
height

orientation

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

# *visual dimension type*

position

relative location  
centrality

shape

colour

saturation  
opacity

size

width  
height

orientation

stroke

colour  
pattern,  
thickness

...

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

# *visual dimension type*

position

relative location  
centrality

shape

colour

saturation  
opacity

size

width  
height

orientation

stroke

colour  
pattern,  
thickness

opacity

...

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

# *visual dimension type*

position

relative location  
centrality

shape

colour

saturation  
opacity

size

width  
height

orientation

stroke

colour  
pattern,  
thickness

opacity

texture

...

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

relational

...

# *visual dimension type*

position

relative location  
centrality

shape

colour

saturation  
opacity

size

width  
height

orientation

stroke

colour  
pattern,  
thickness

opacity

texture

movement

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

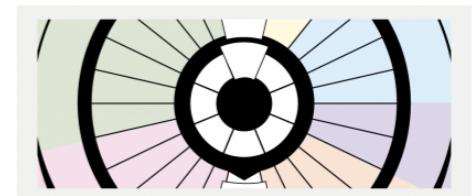
relational

...

# *visual dimension type*

position

relative location  
centrality



shape



colour

saturation  
opacity

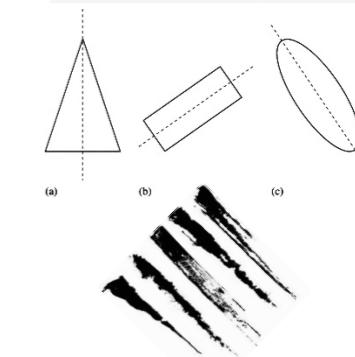


size

width  
height



orientation

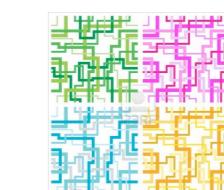


stroke

colour  
pattern,  
thickness

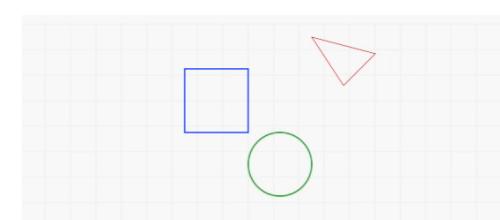


opacity



texture

movement



juxtaposition

# *data dimension types*

integral

fixed point

alpha(-numeric)

fractions of a population

categorical

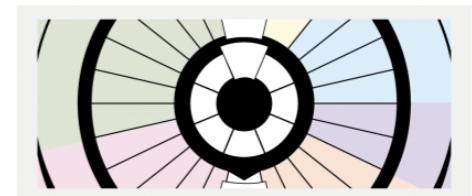
relational

...

# *visual dimension type*

position

relative location  
centrality



shape



colour

saturation  
opacity

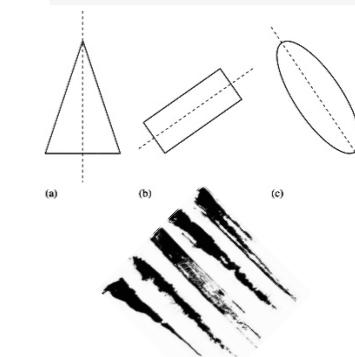


size

width  
height



orientation

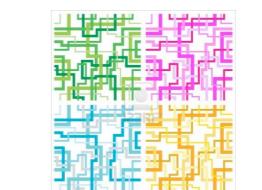


stroke

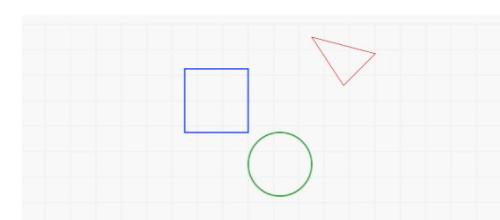
colour  
pattern,  
thickness



opacity



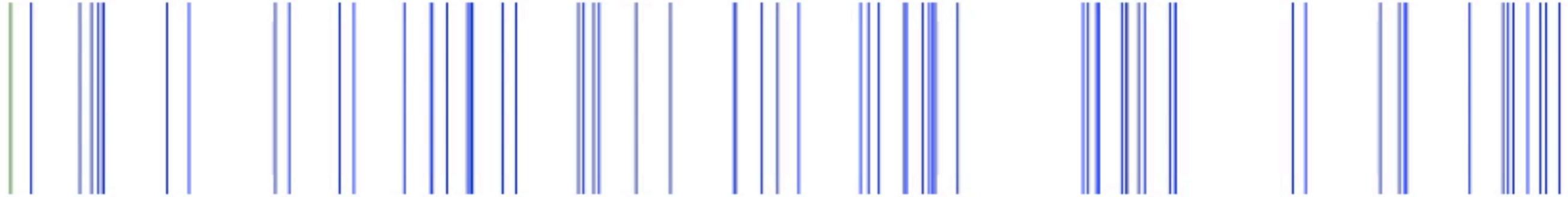
texture



movement

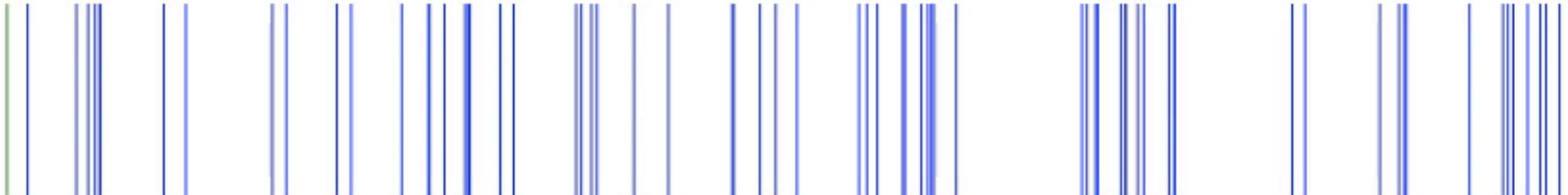
juxtaposition

# position

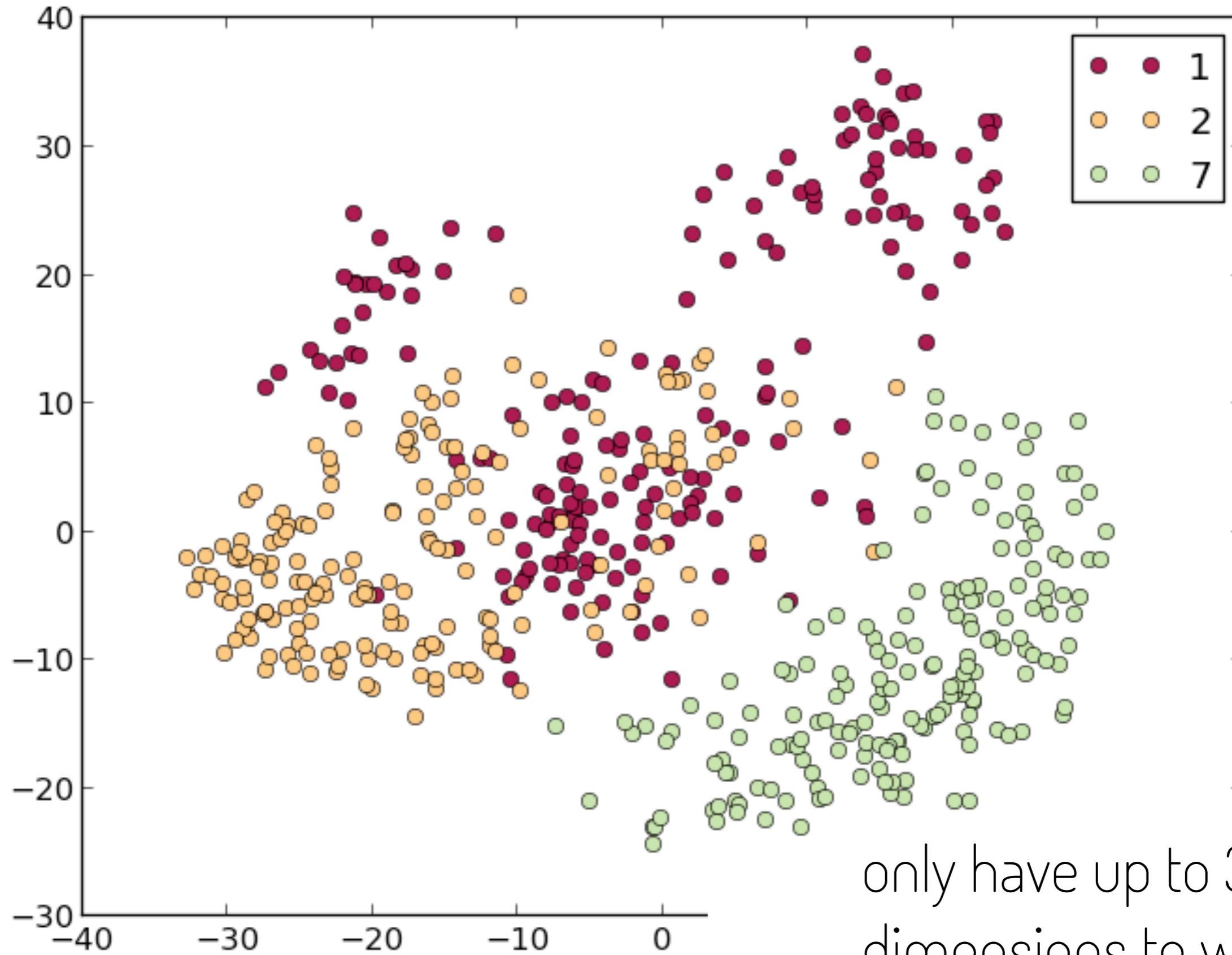


# position

linear mapping of values  
logarithmic..  
bin and count..



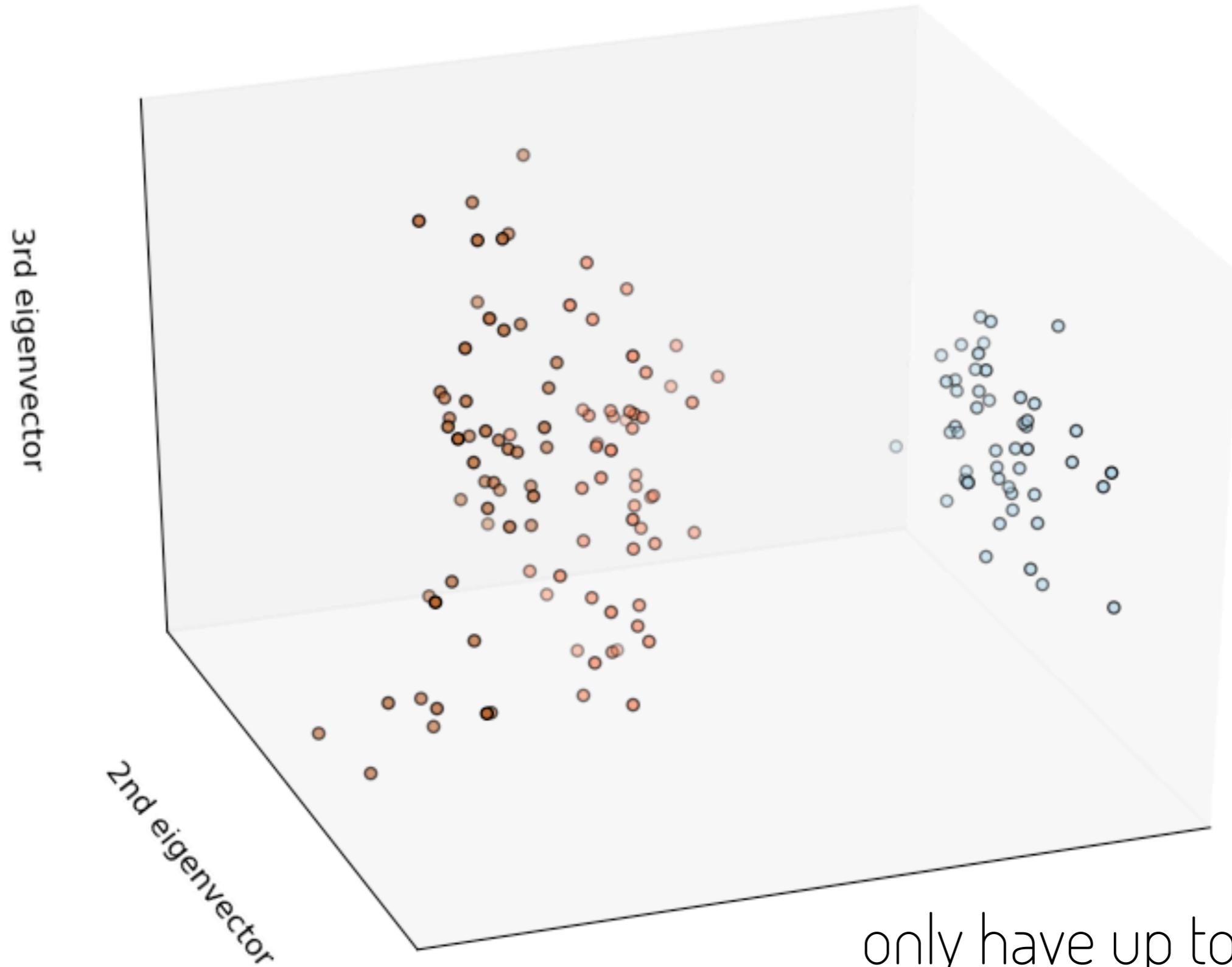
# position



only have up to 3 spatial  
dimensions to work with

# position

First three PCA directions



only have up to 3 spatial dimensions to work with

# orientation

orientation

range-limited

# orientation

range-limited



# orientation

range-limited

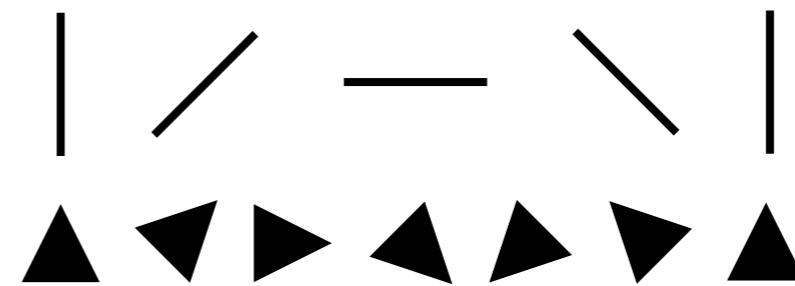
symmetry properties of the  
geometry



# orientation

range-limited

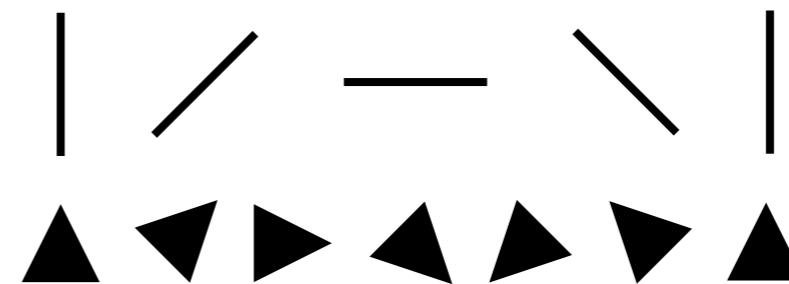
symmetry properties of the  
geometry



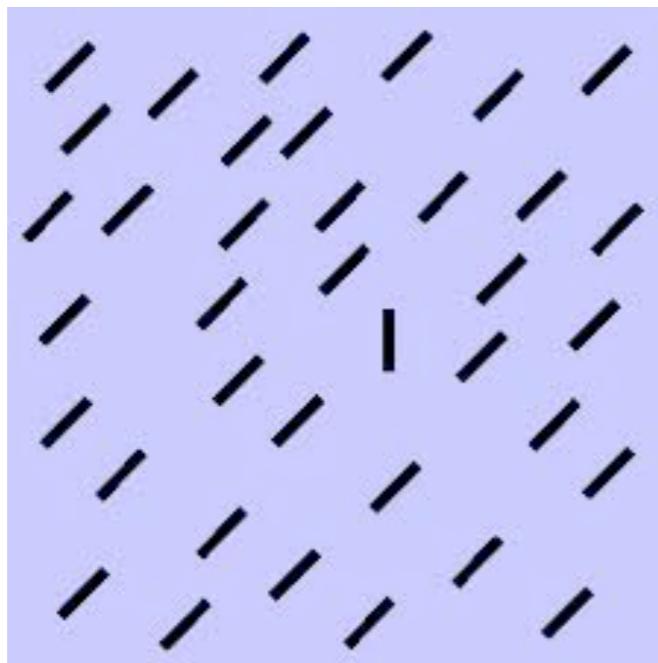
# orientation

range-limited

symmetry properties of the  
geometry



pop-out



TTTT  
TTTT  
TTTT  
TTTT

TTTT  
T~~T~~TT  
TTTT  
TTTT

FFFF  
F~~E~~FF  
FFFF  
FFFF

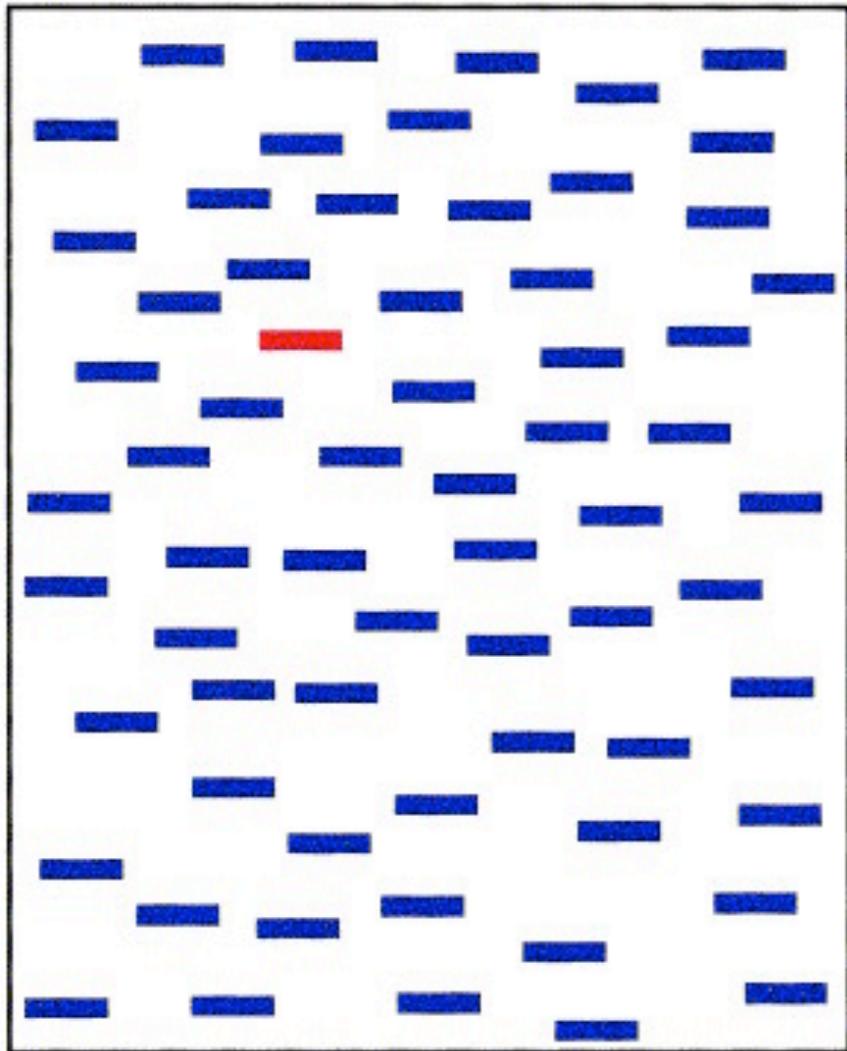
FFFF  
F~~A~~FF  
FFFF  
FFFF

UUUU  
U~~O~~UU  
UUUU  
UUUU

UUUU  
U~~G~~UU  
UUUU  
UUUU

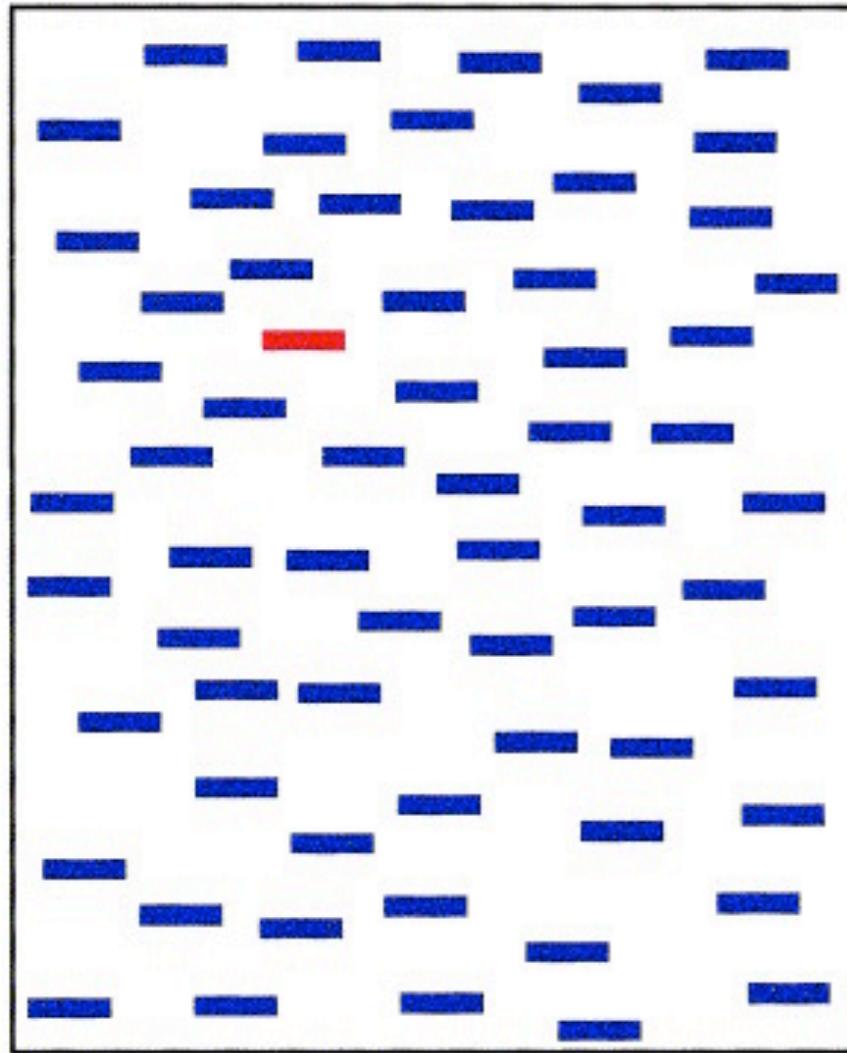
orientation  
popouts using multiple dimensions

orientation  
popouts using multiple dimensions

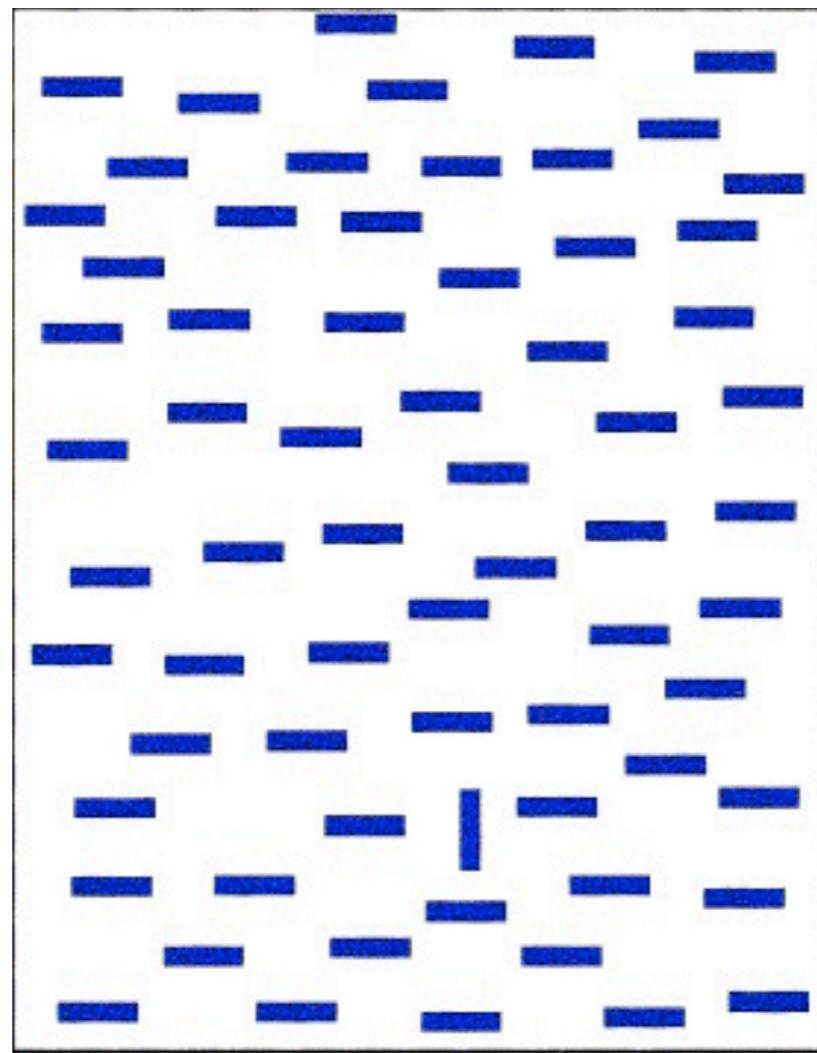


1D colour

orientation  
popouts using multiple dimensions

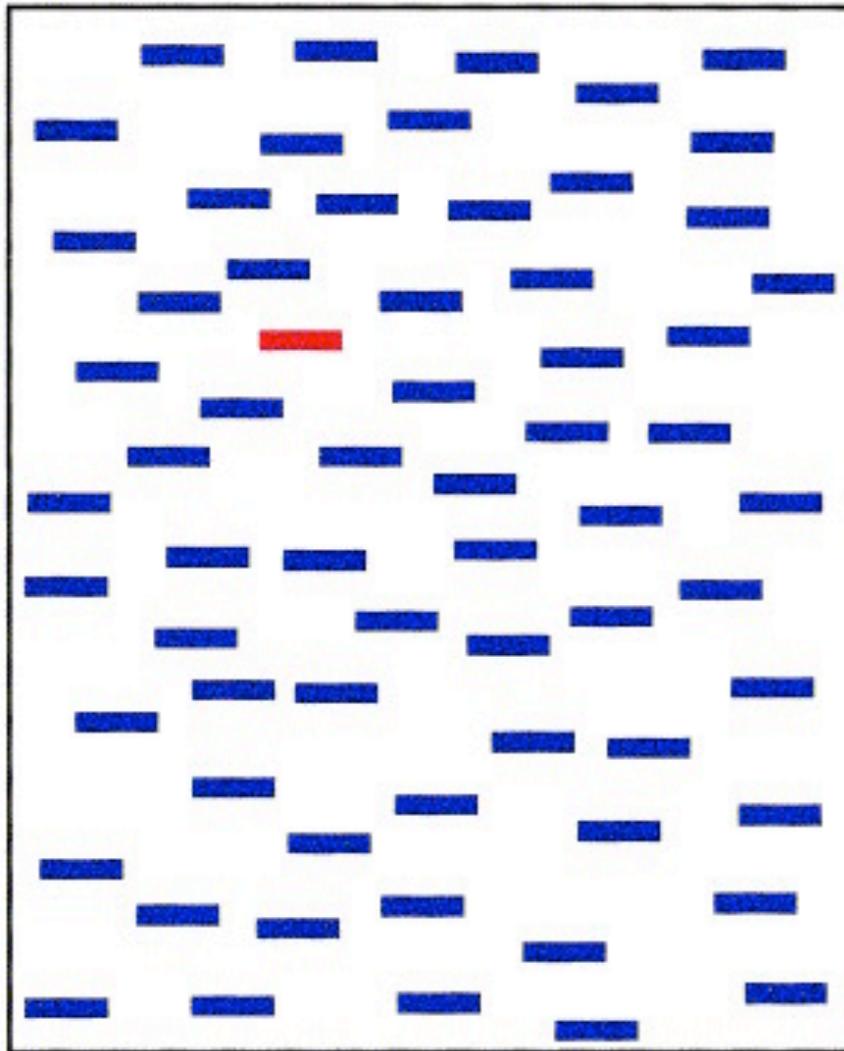


1D colour

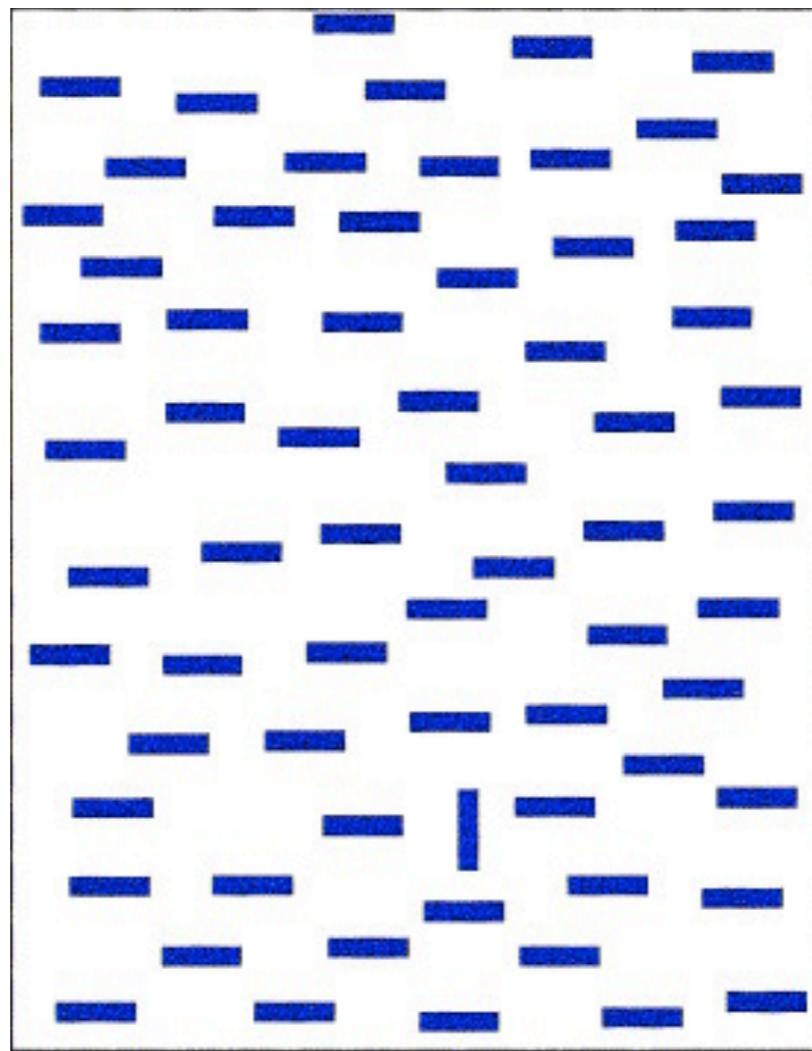


1D orientation

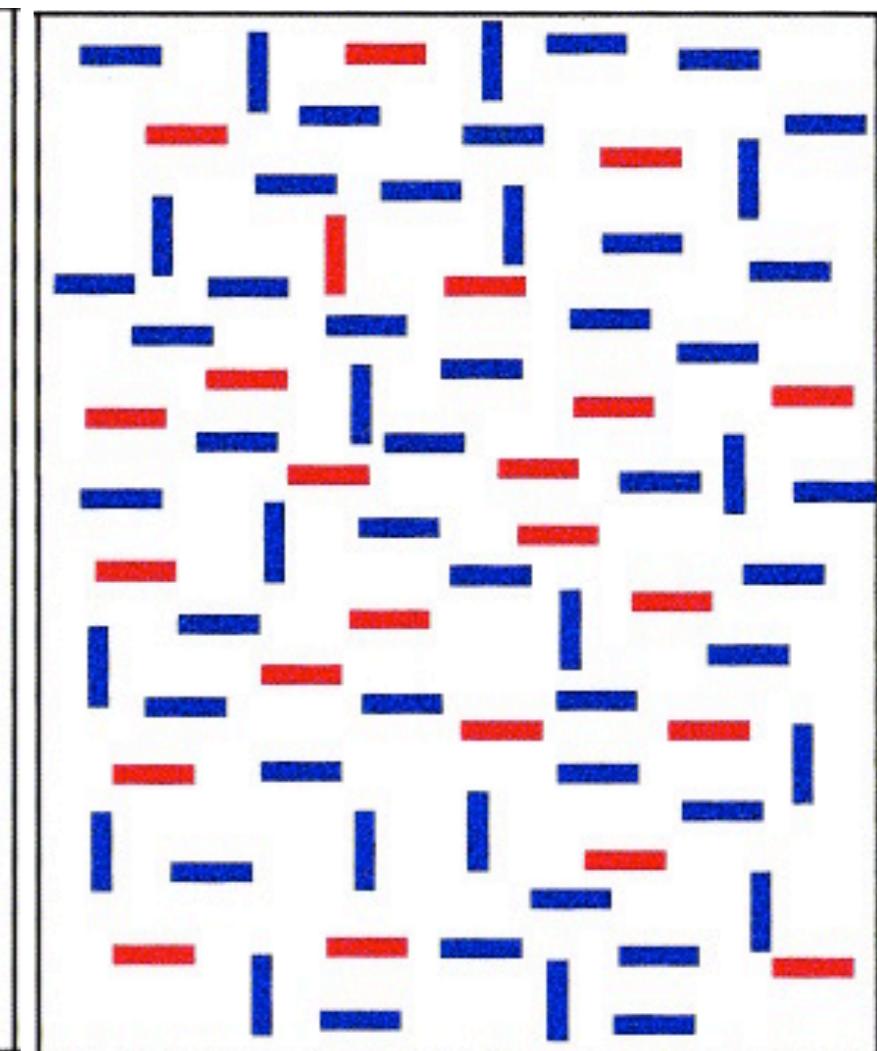
orientation  
popouts using multiple dimensions



1D colour

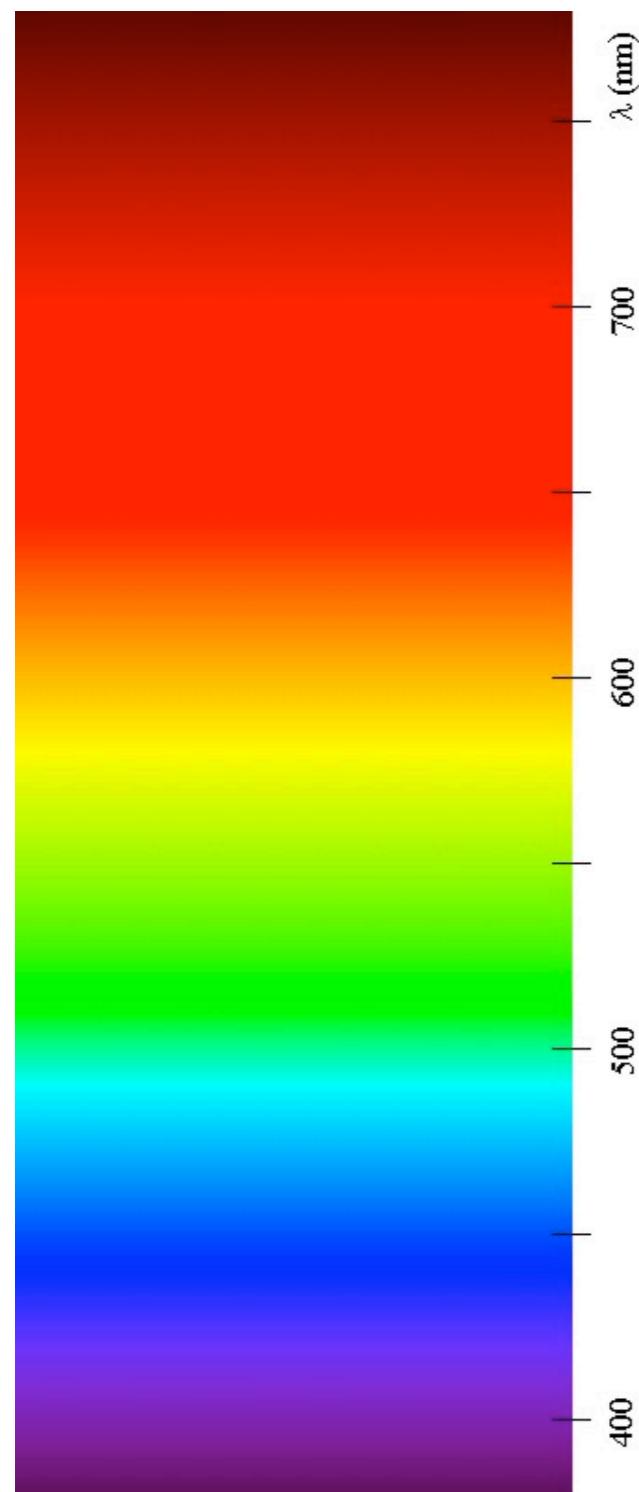


1D orientation

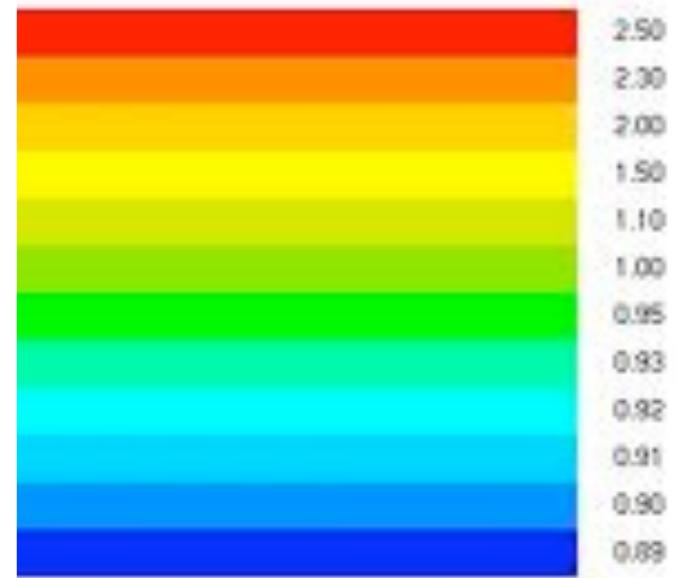
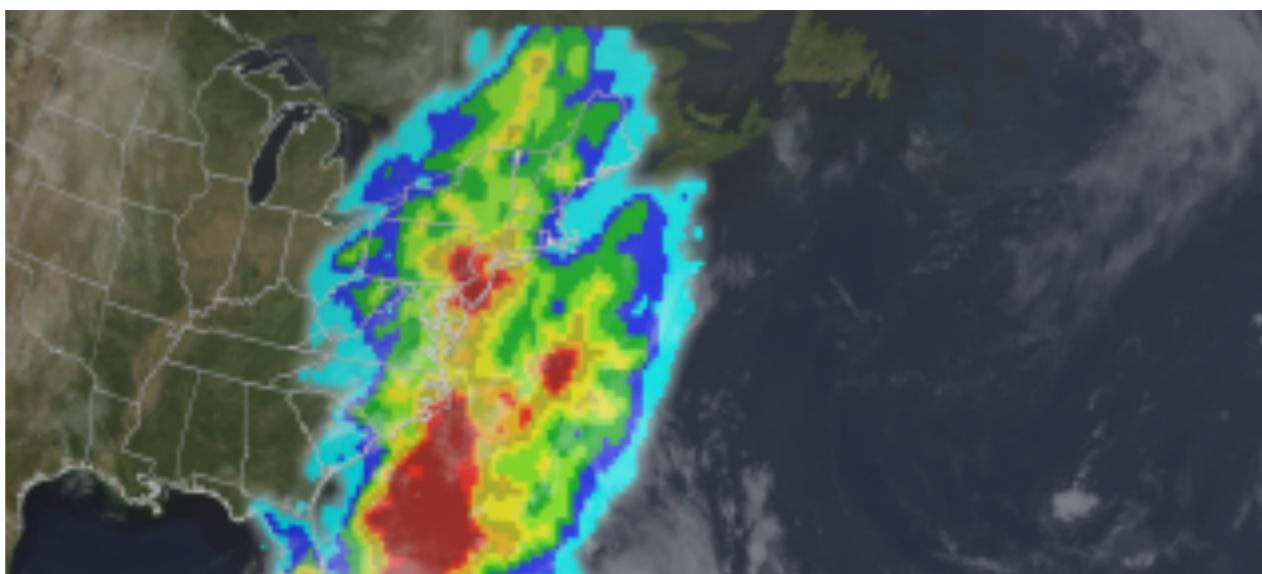
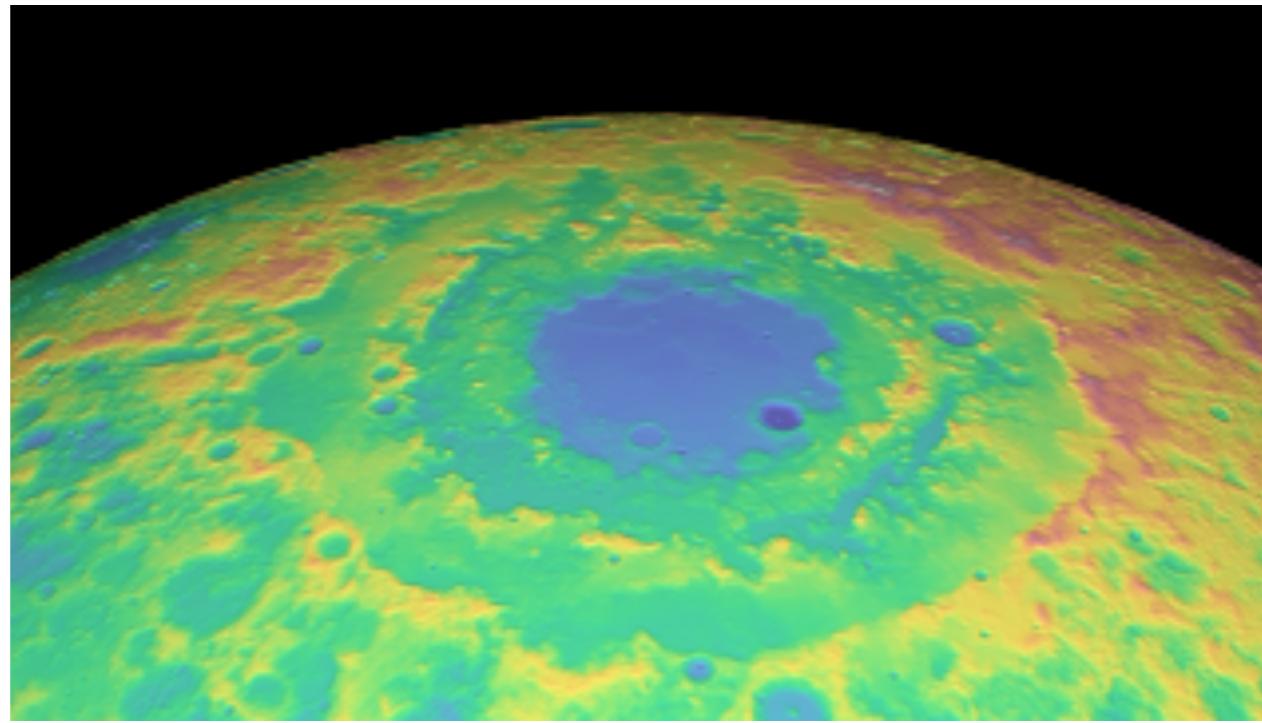
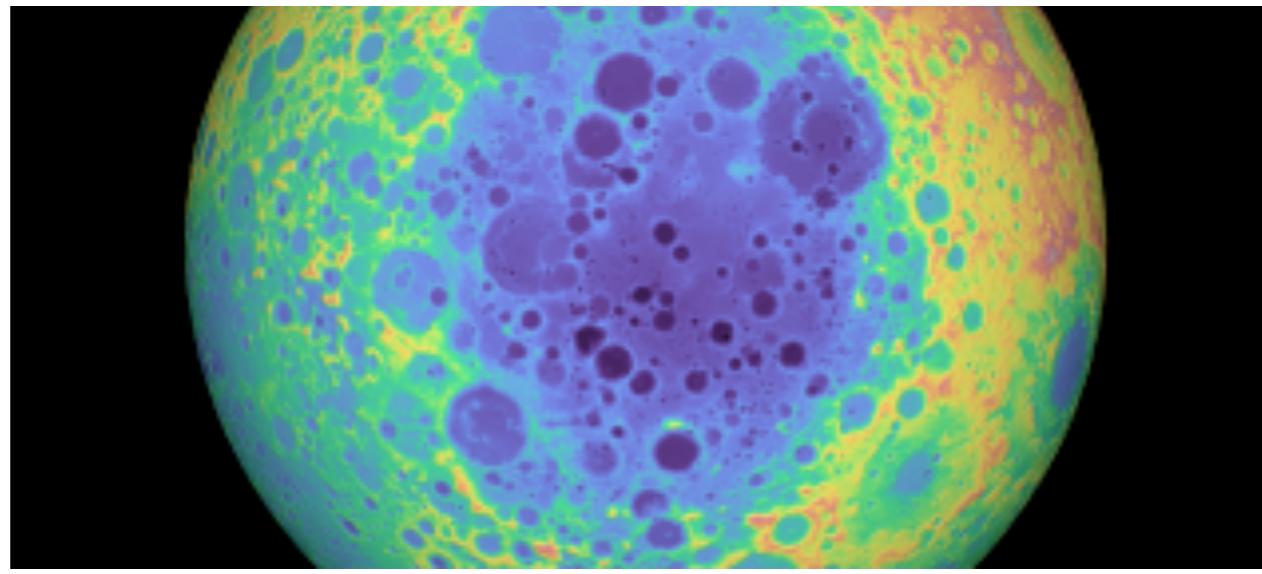
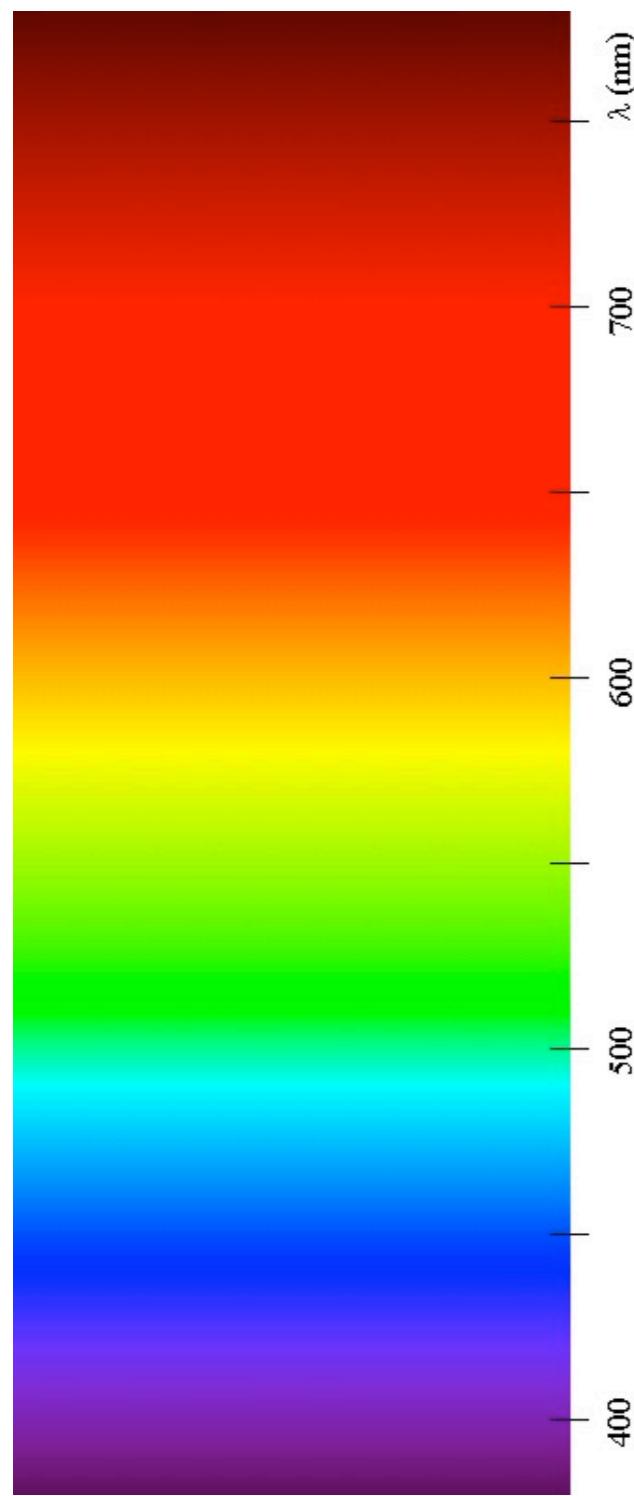


2D color/  
orientation

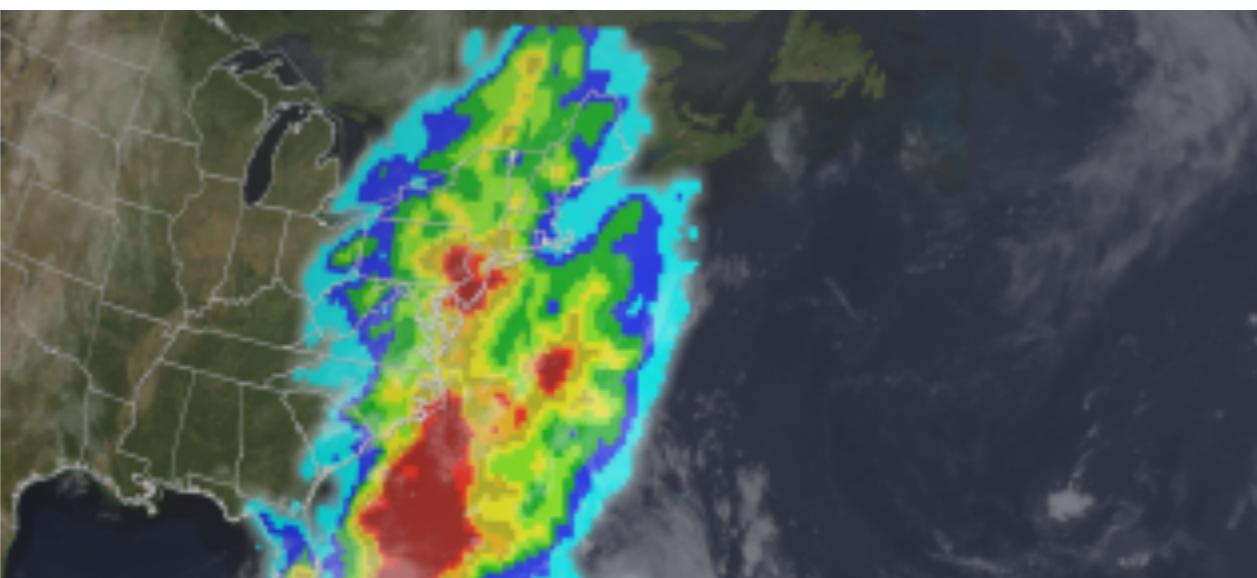
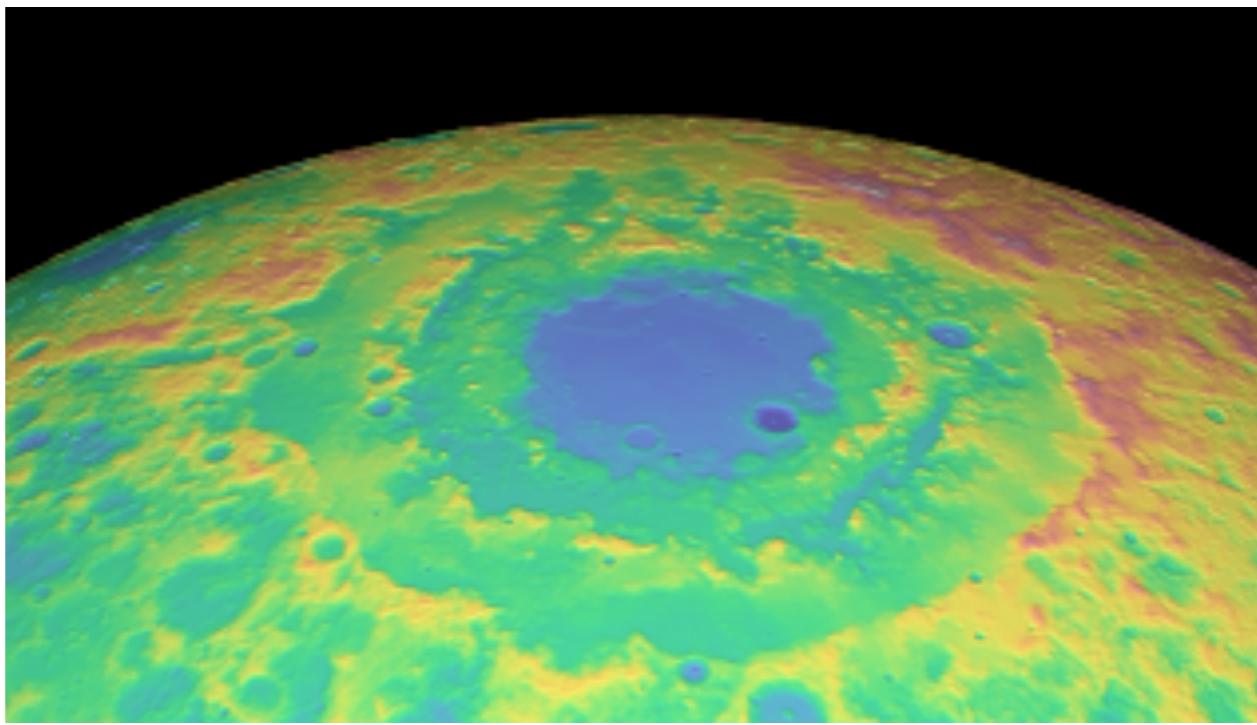
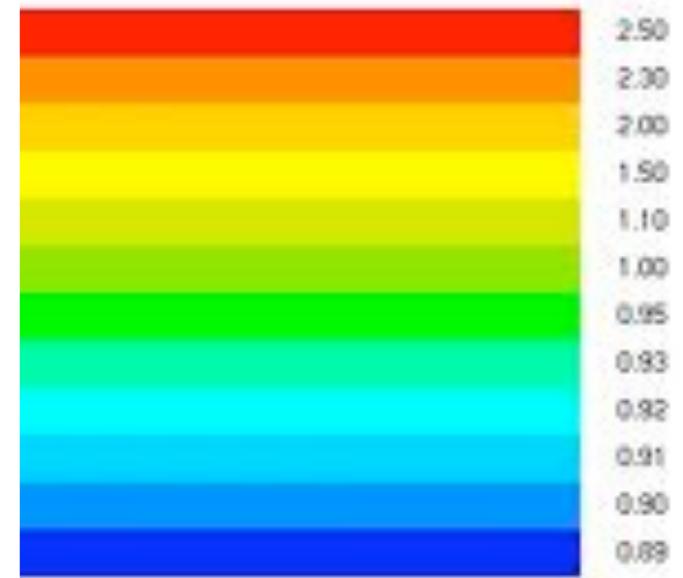
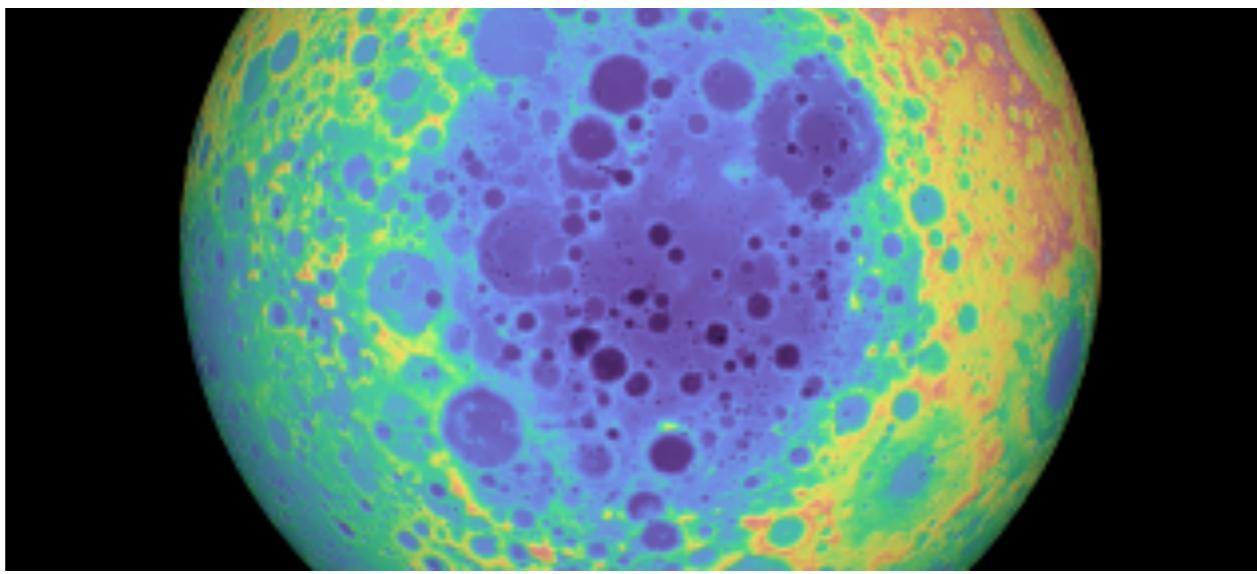
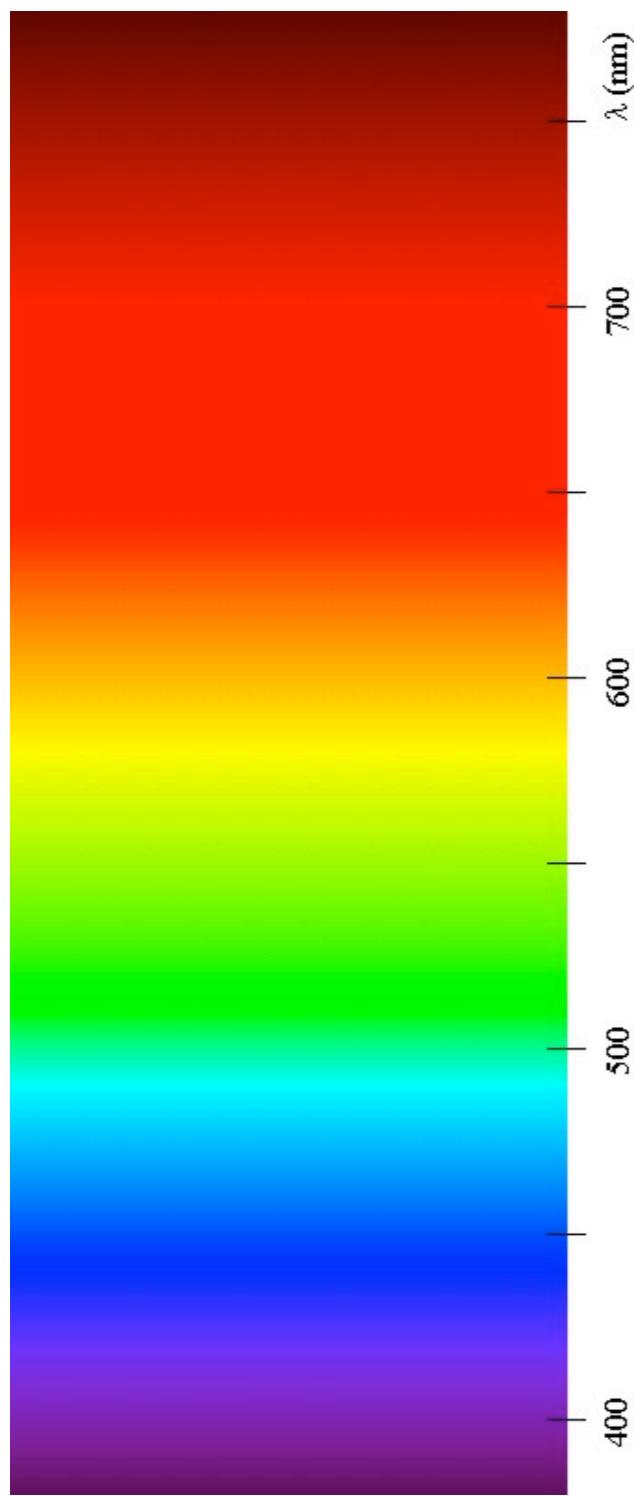
# Using colour for continuous values



# Using colour for continuous values



# Using colour for continuous values



# Using colour for continuous values

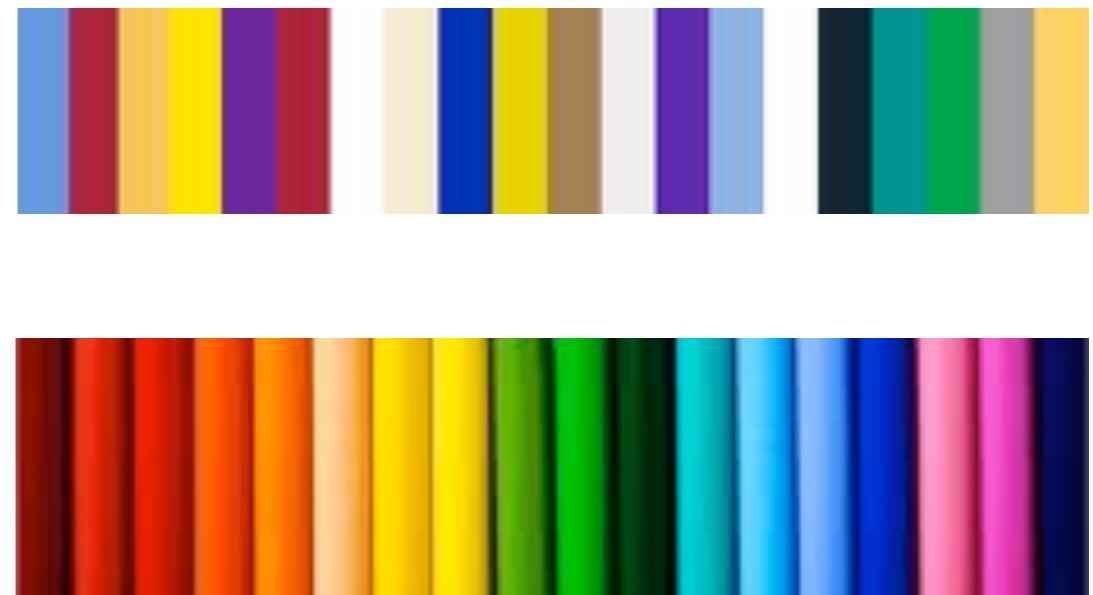
problem 1: No natural ordering

# Using colour for continuous values



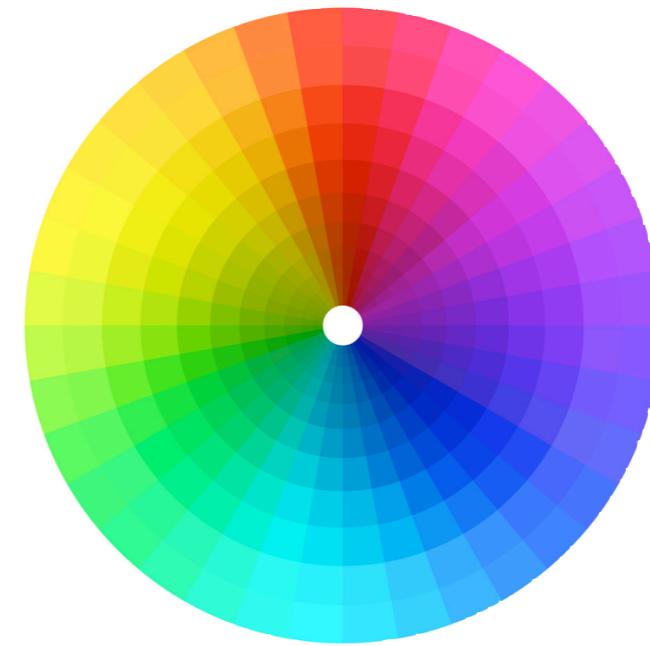
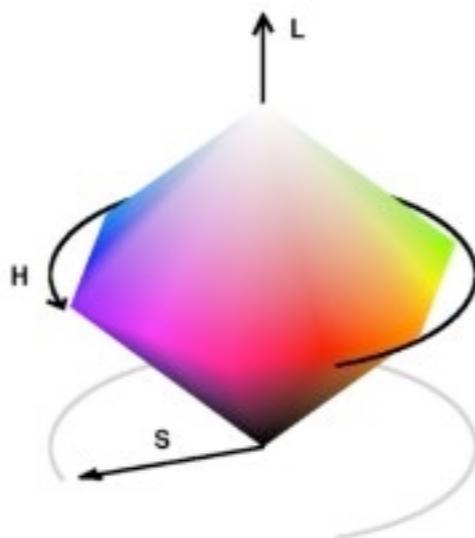
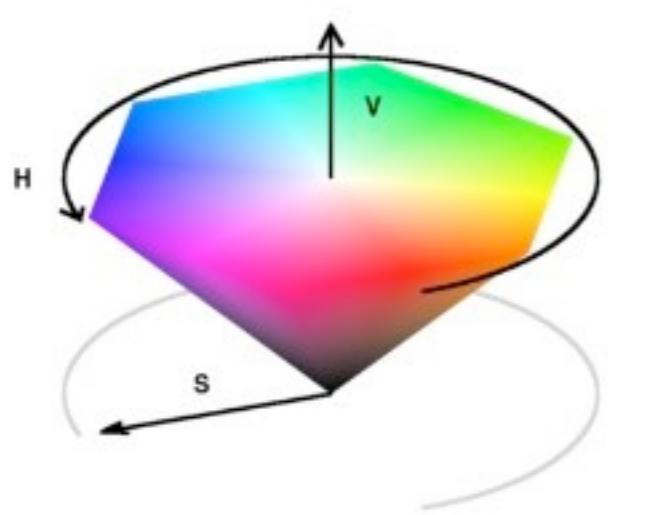
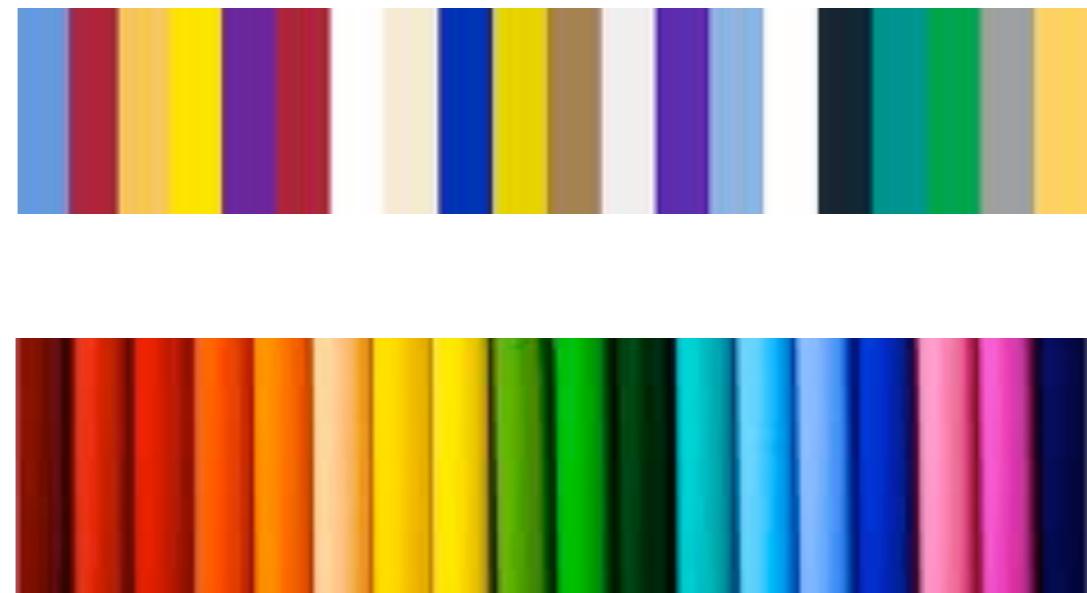
problem 1: No natural ordering

# Using colour for continuous values



problem 1: No natural ordering

# Using colour for continuous values



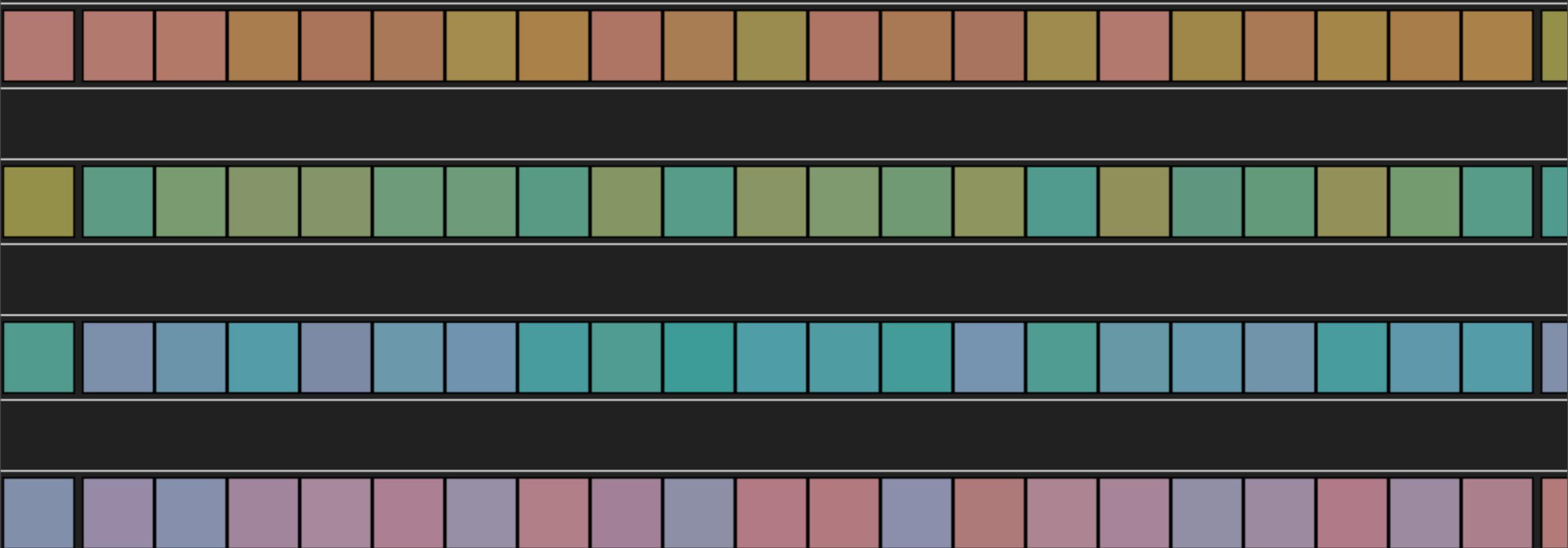
problem 1: No natural ordering

Kostenlos 1000x400 px  
Foto: Heric. JPG file download  
www.4sharedgraphics.com

# Using colour for continuous values

Drag and drop the colors in each row to arrange them by hue order.

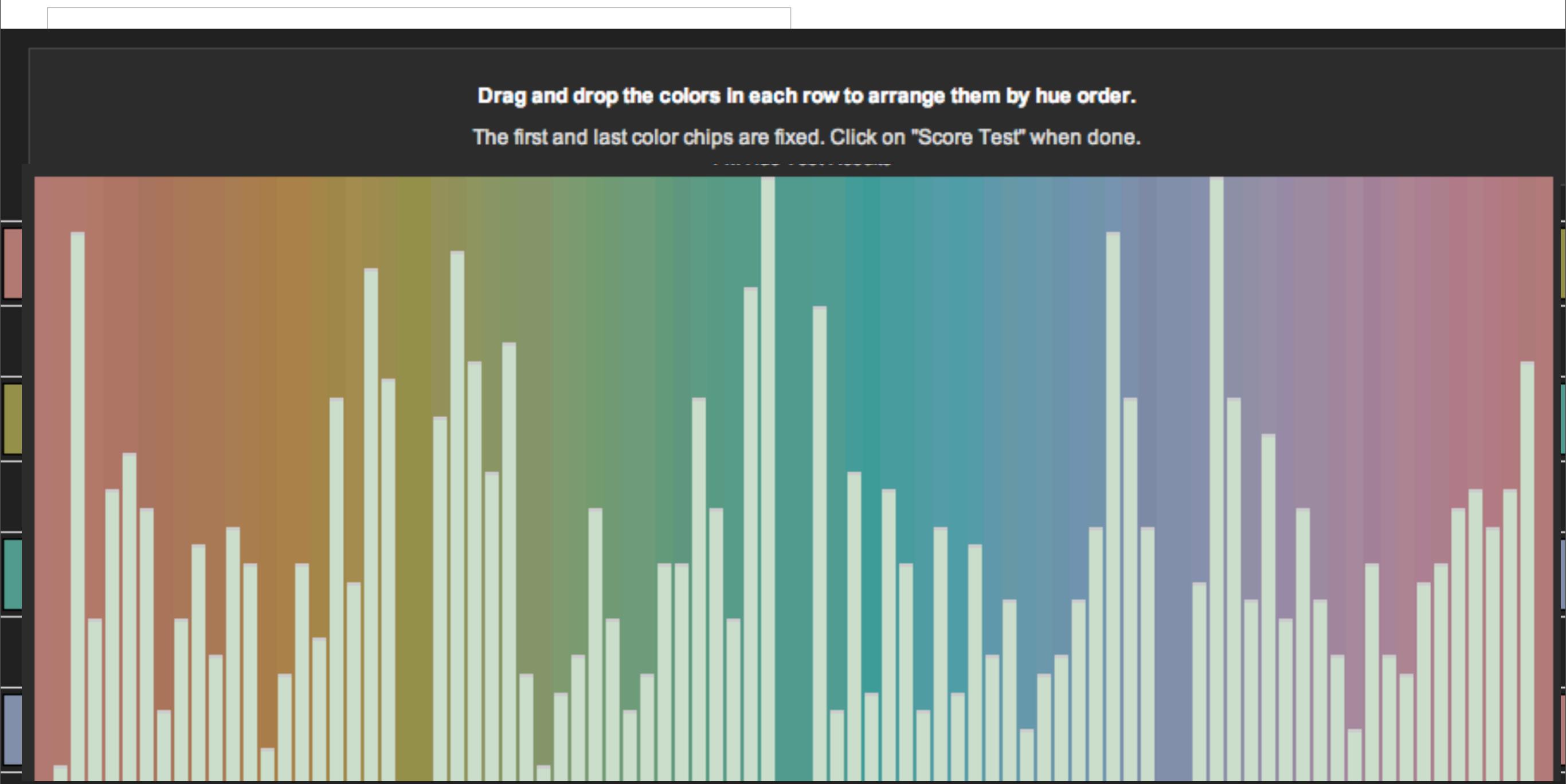
The first and last color chips are fixed. Click on "Score Test" when done.



[http://www.colormunki.com/game/huetest\\_kiosk](http://www.colormunki.com/game/huetest_kiosk)

problem 1: No natural ordering

# Using colour for continuous values

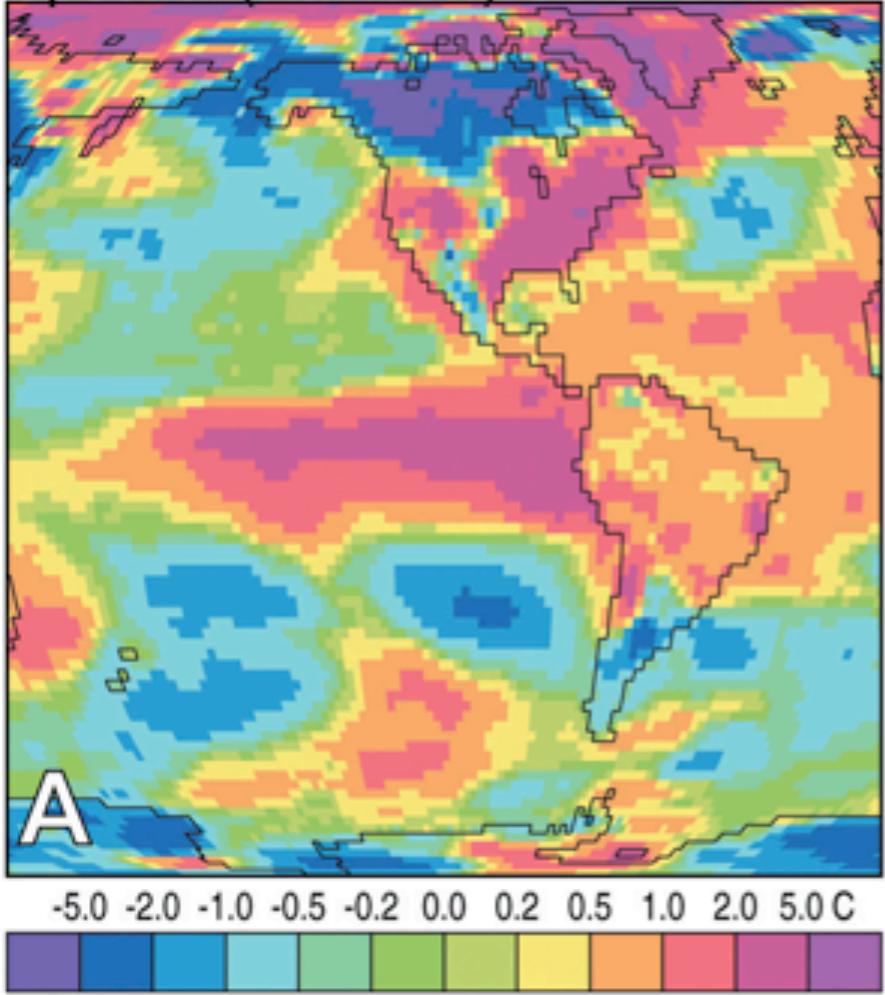


[http://www.colormunki.com/game/huetest\\_kiosk](http://www.colormunki.com/game/huetest_kiosk)

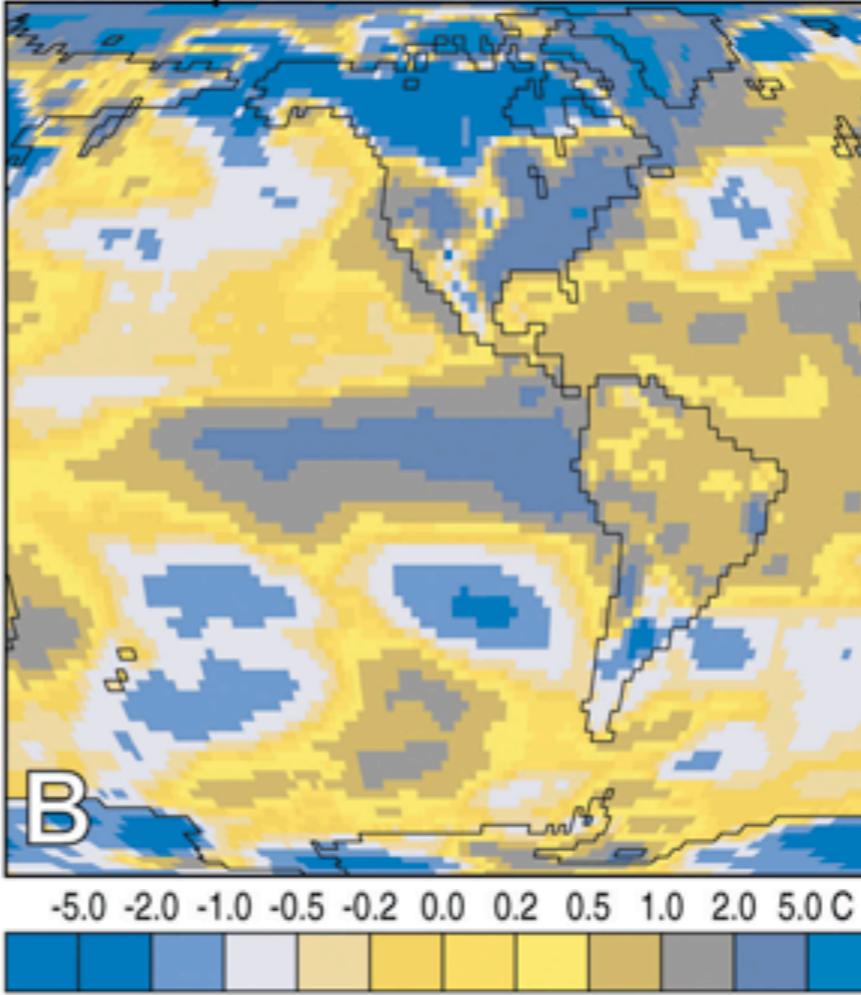
problem 1: No natural ordering

# Using colour for continuous values

Spectral (Rainbow) Color Scale



Protanopic Simulation

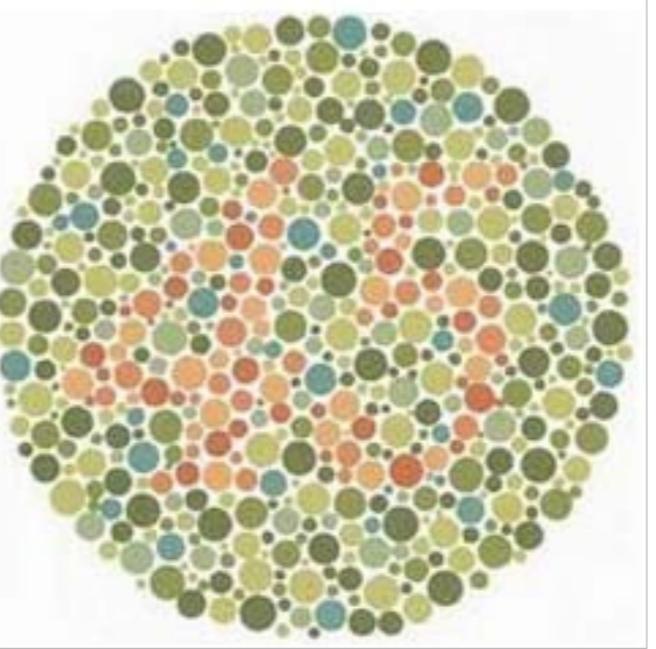
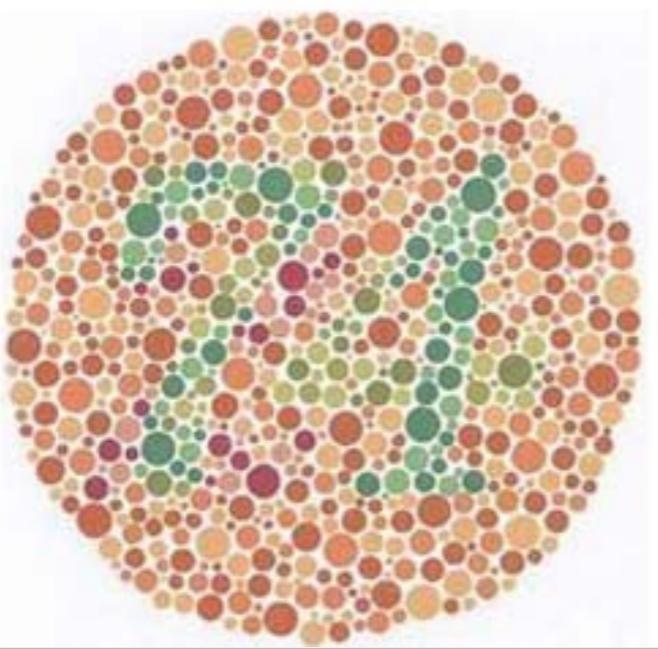
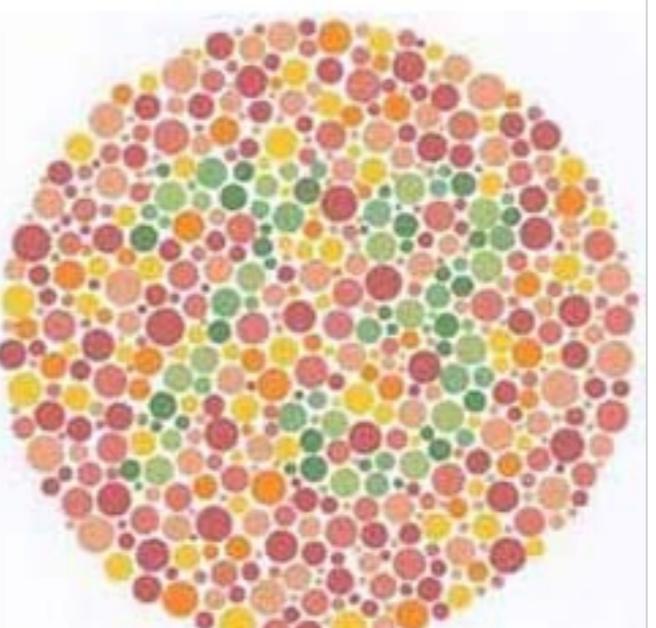
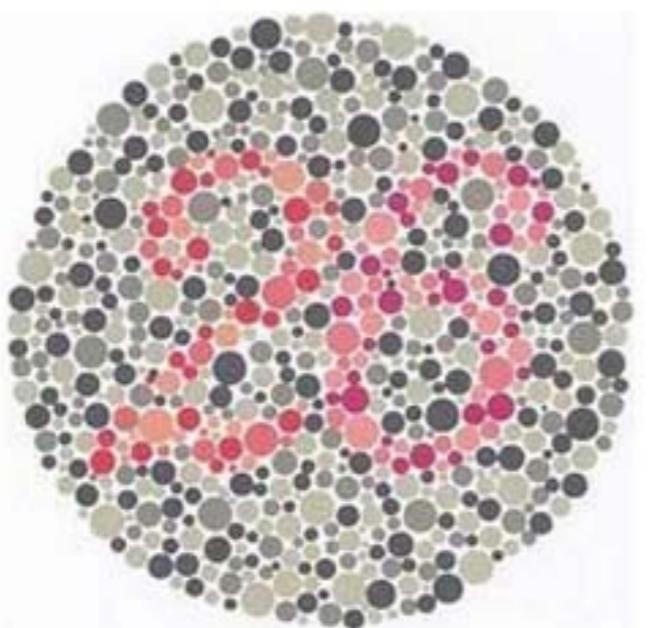
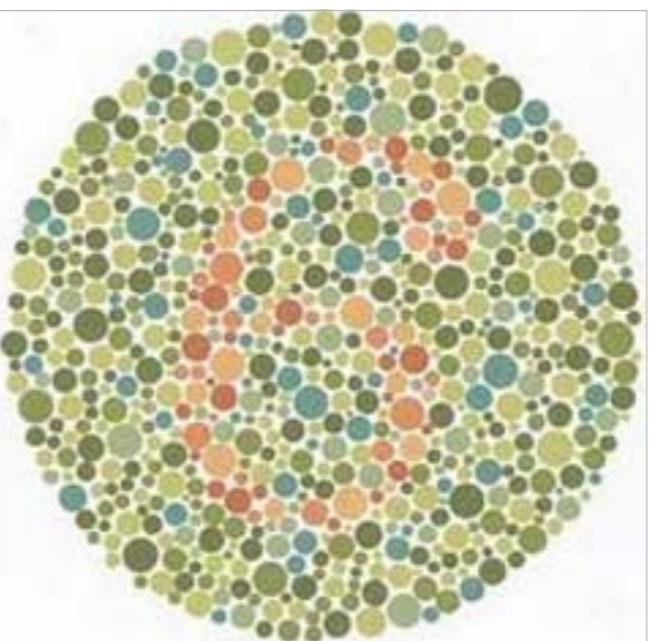
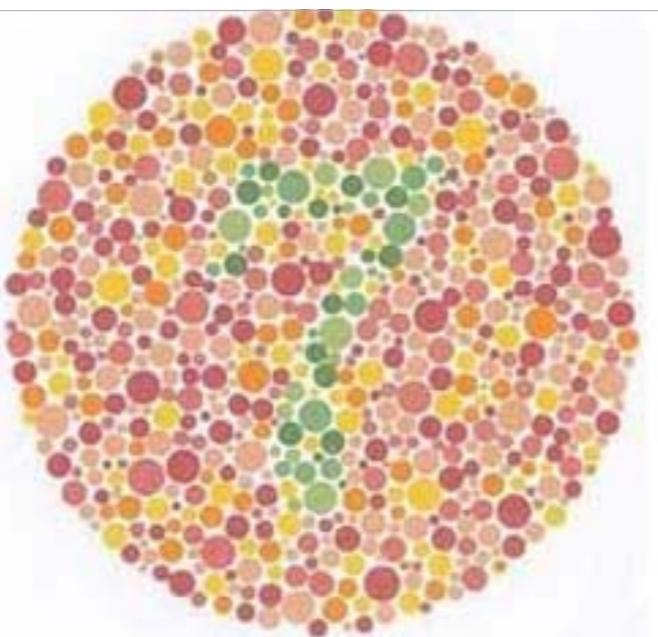


protanopia



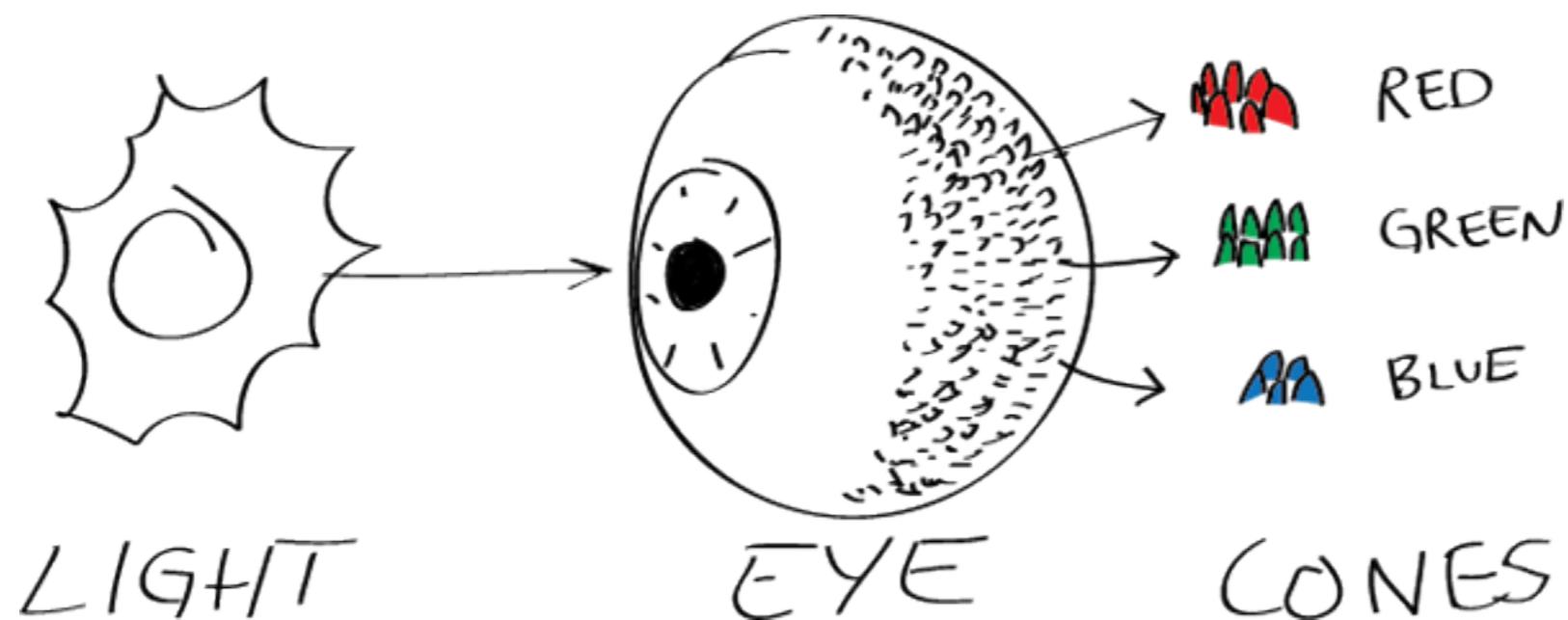
Protanopia affects 8% of males, 0.5% females  
of Northern European ancestry

problem 2: colour sensitivity



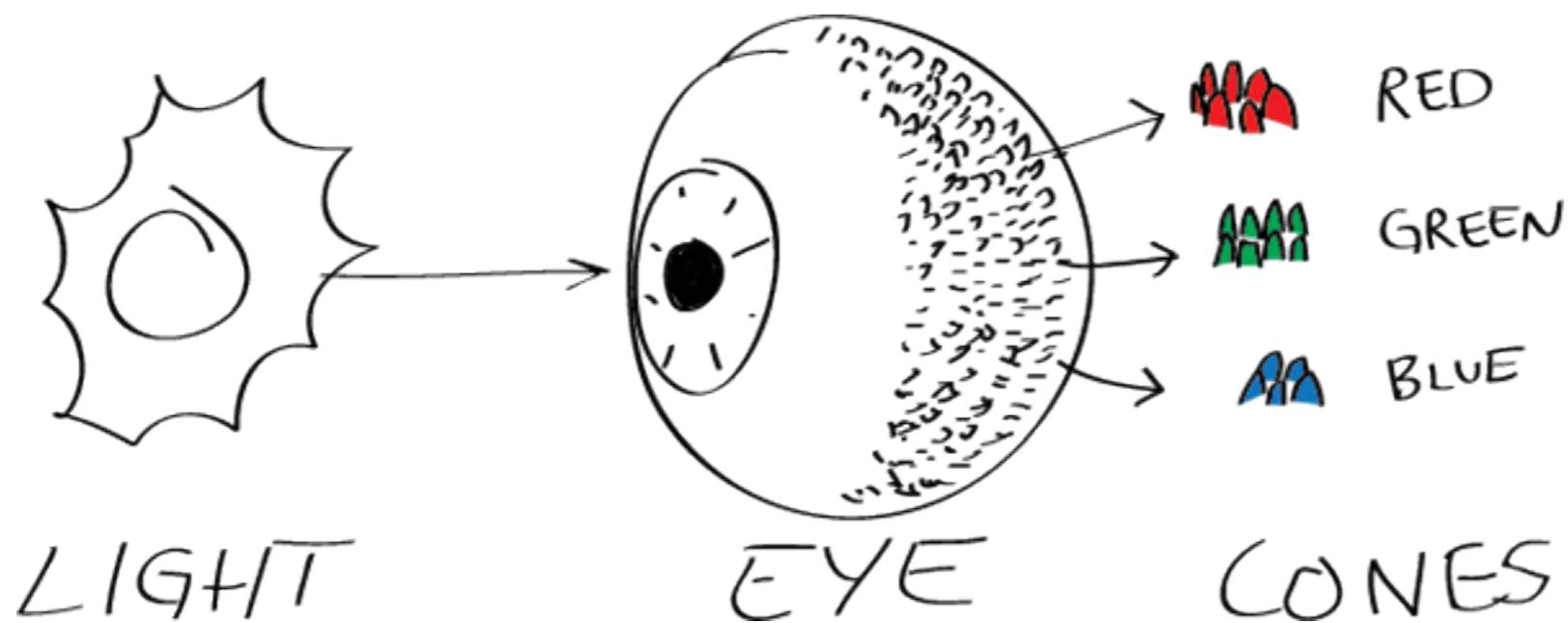
Using colour for continuous values

problem 3: yellow is special

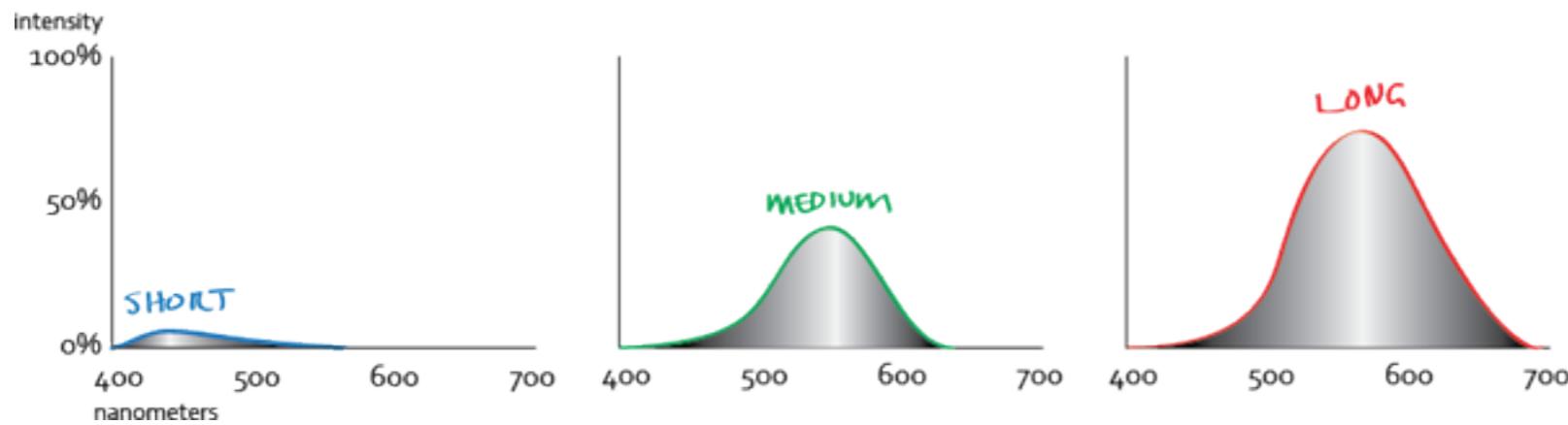


# Using colour for continuous values

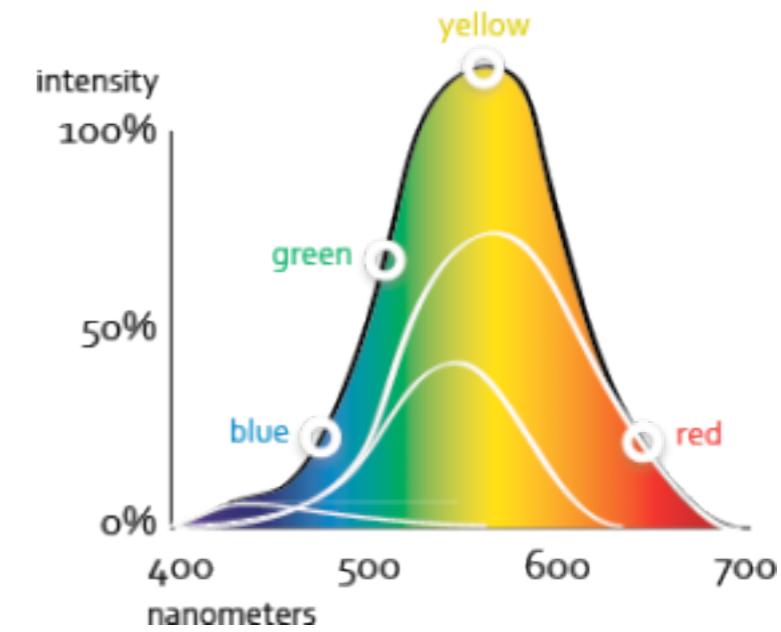
## problem 3: yellow is special



RELATIVE SENSITIVITY TO LIGHT WAVELENGTHS

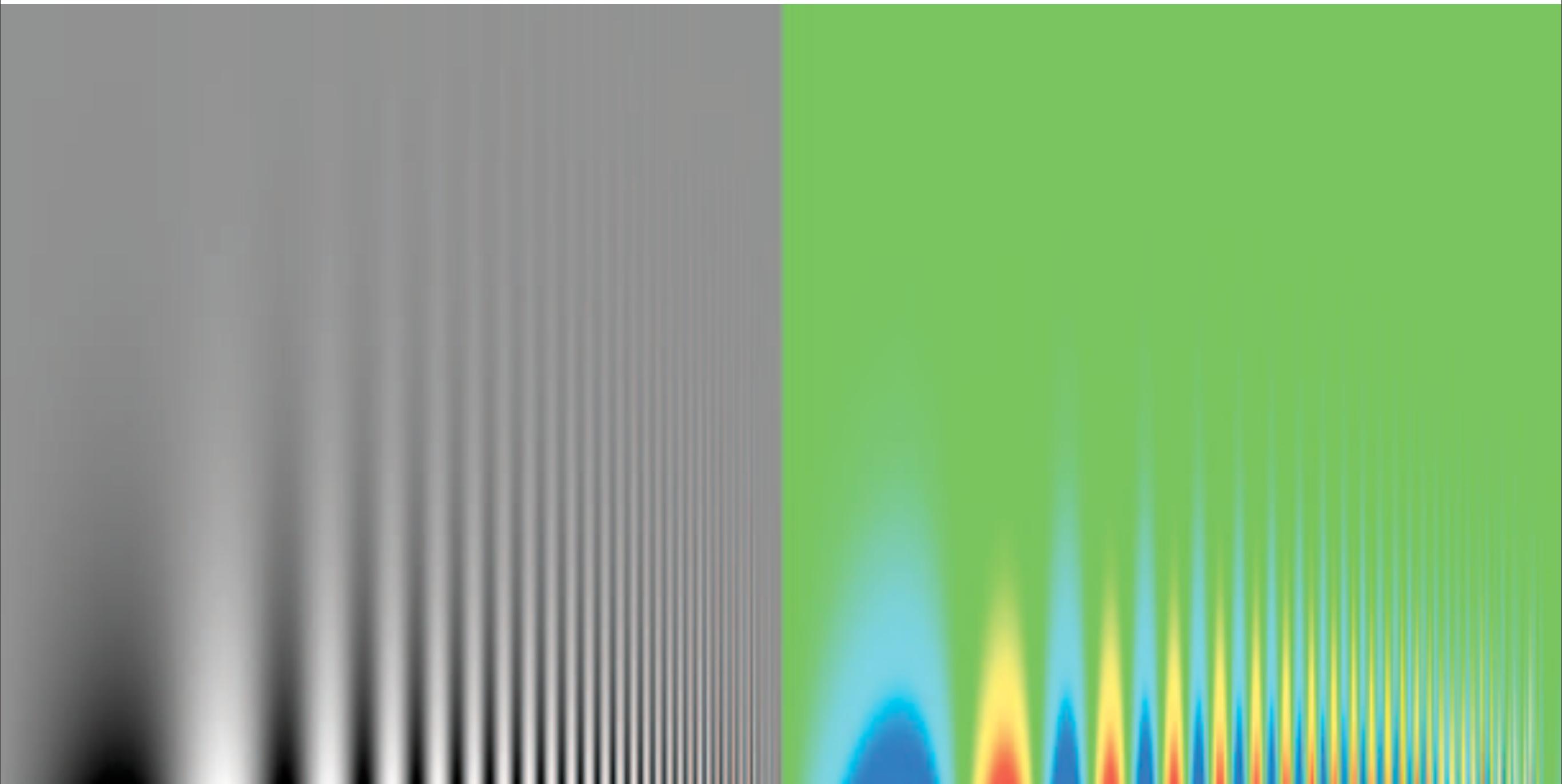


PUTTING IT ALL TOGETHER



# Using colour for continuous values

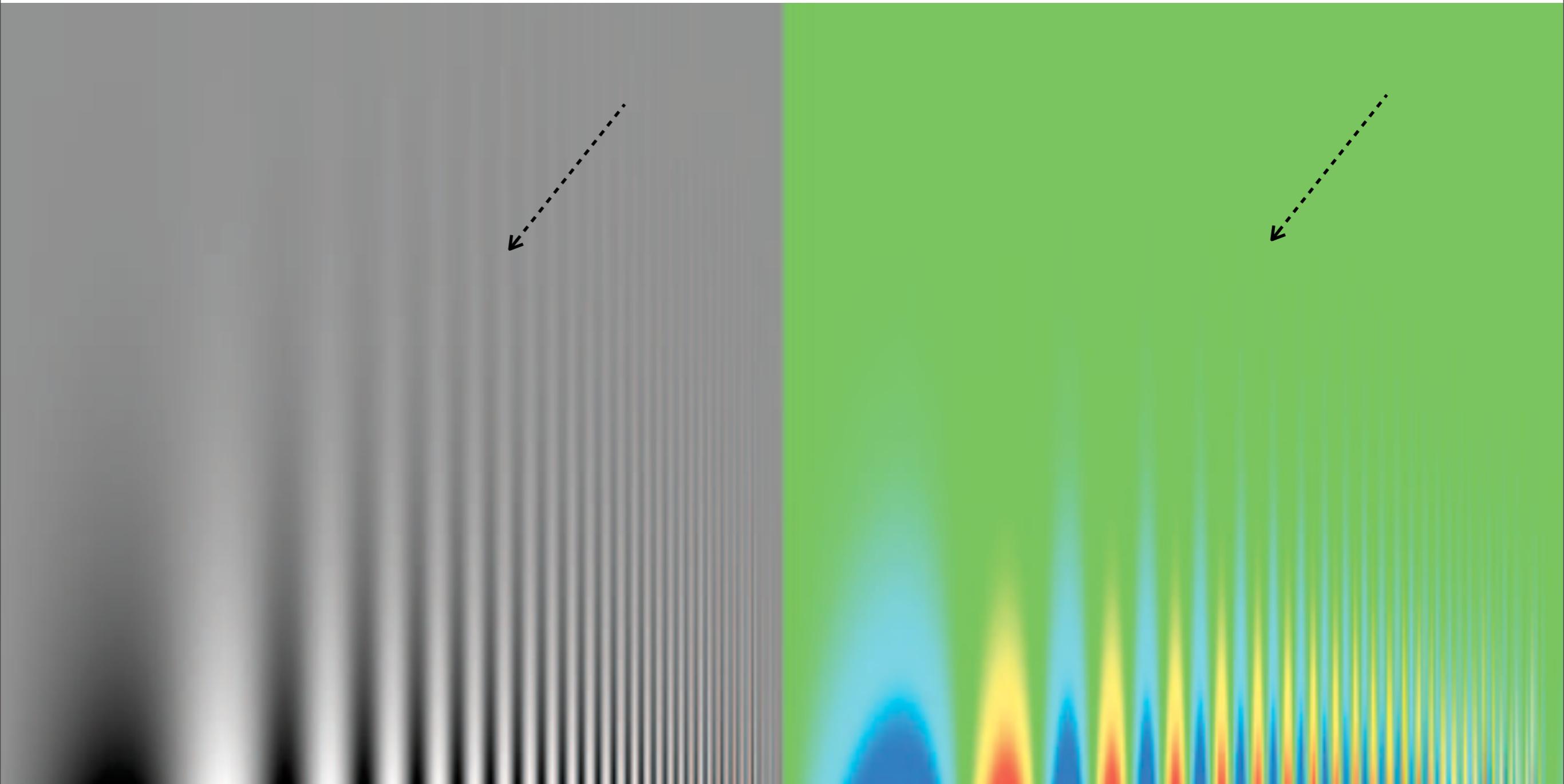
## problem 4: Details: overemphasised or obscured



hue ‘borders’ overemphasise small changes, hue ‘middles’ blend potentially important details

Using colour for continuous values

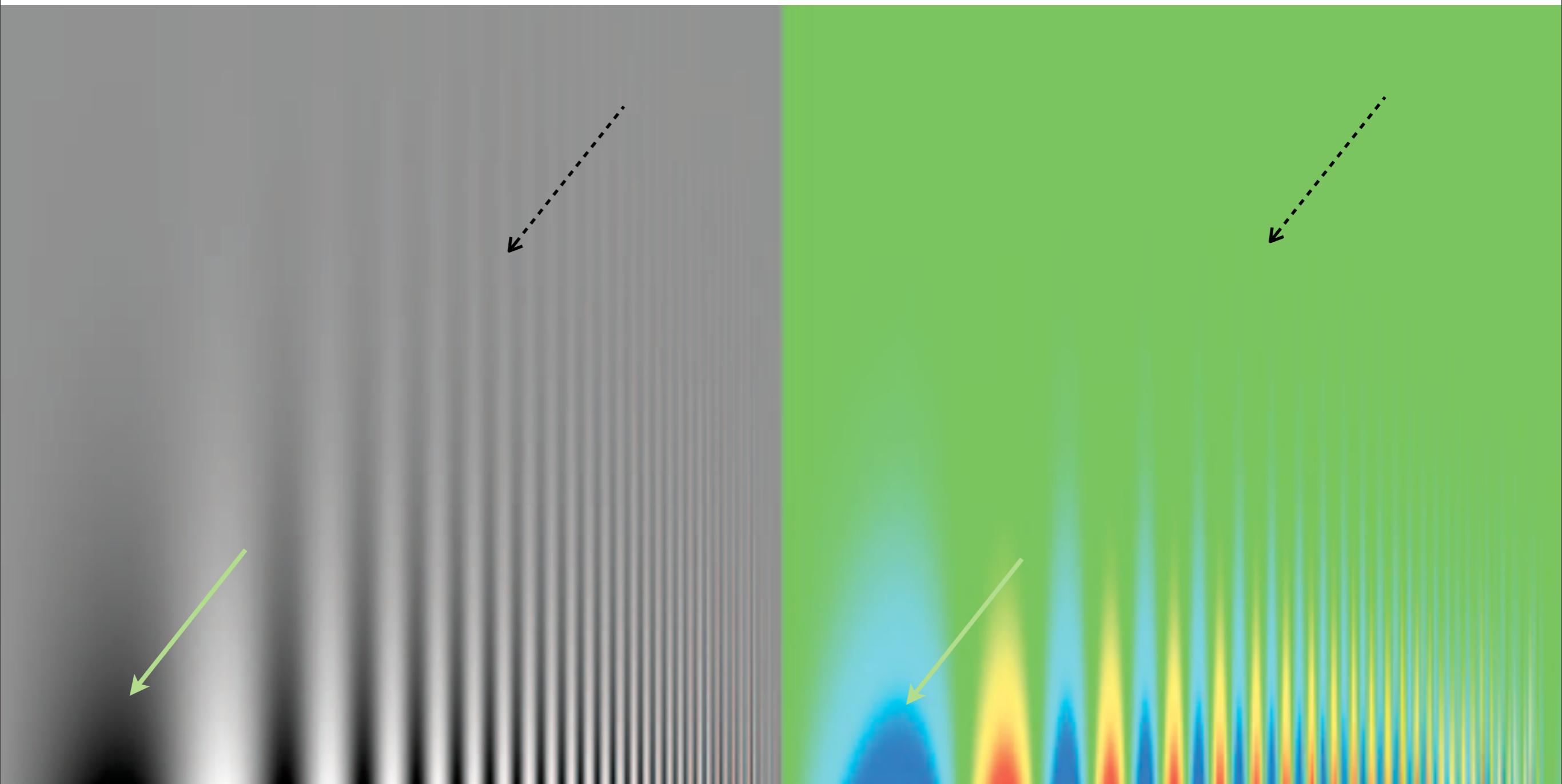
problem 4: Details: overemphasised or obscured



hue ‘borders’ overemphasise small changes, hue ‘middles’ blend potentially important details

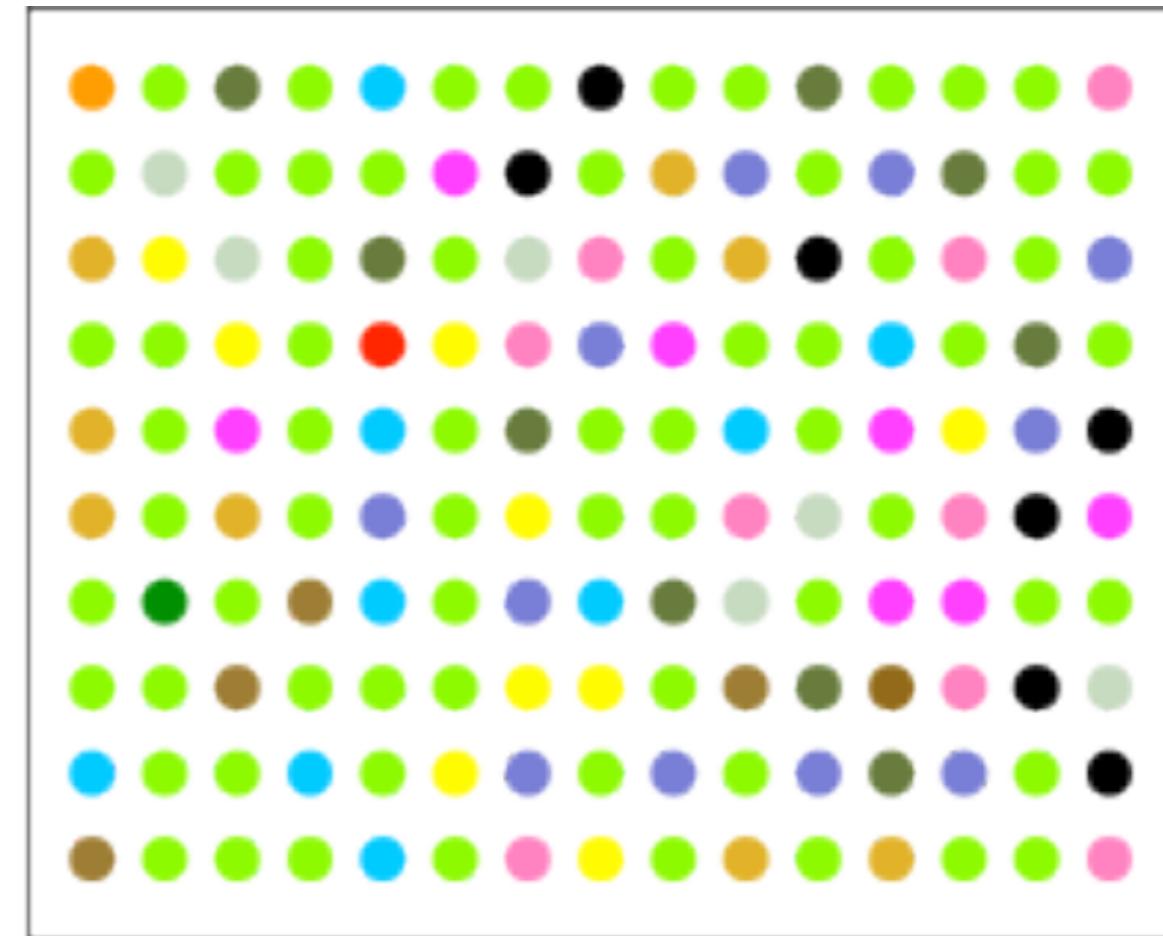
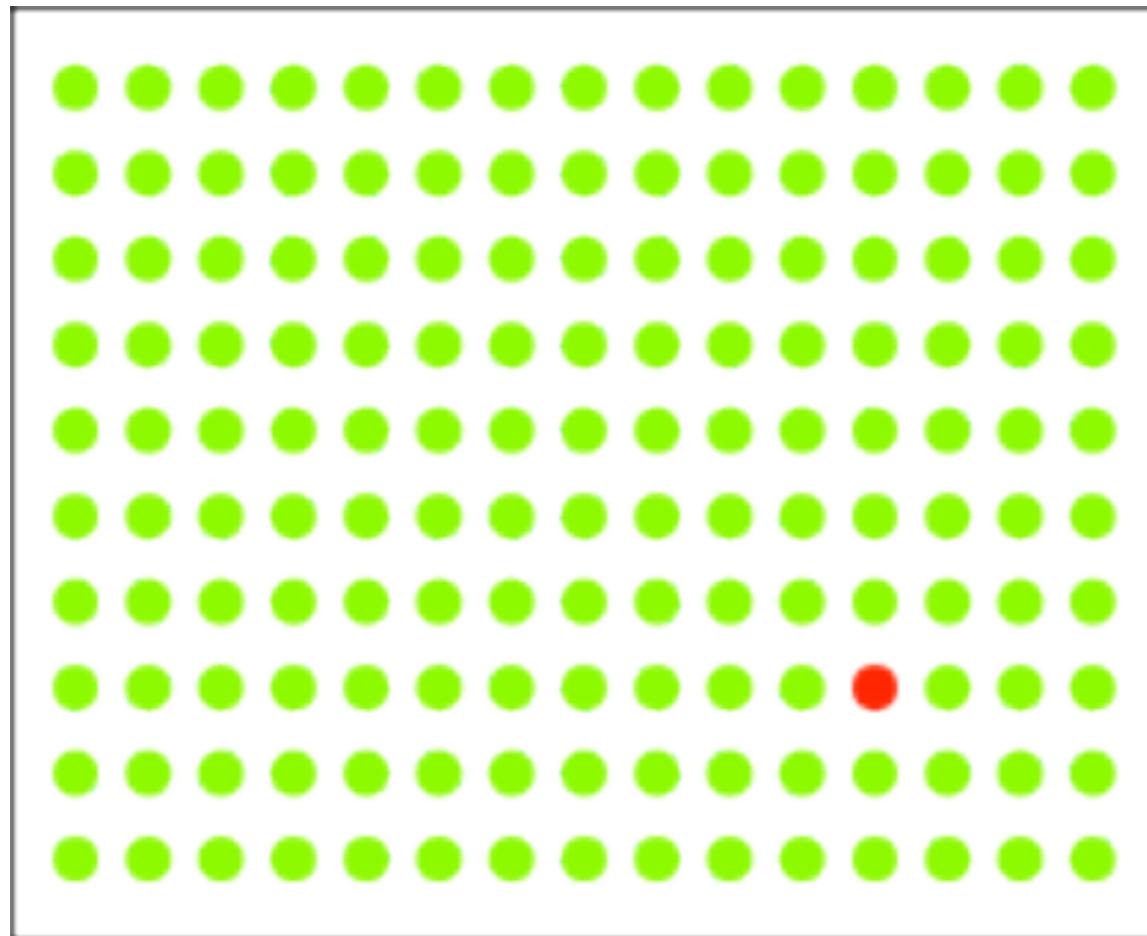
# Using colour for continuous values

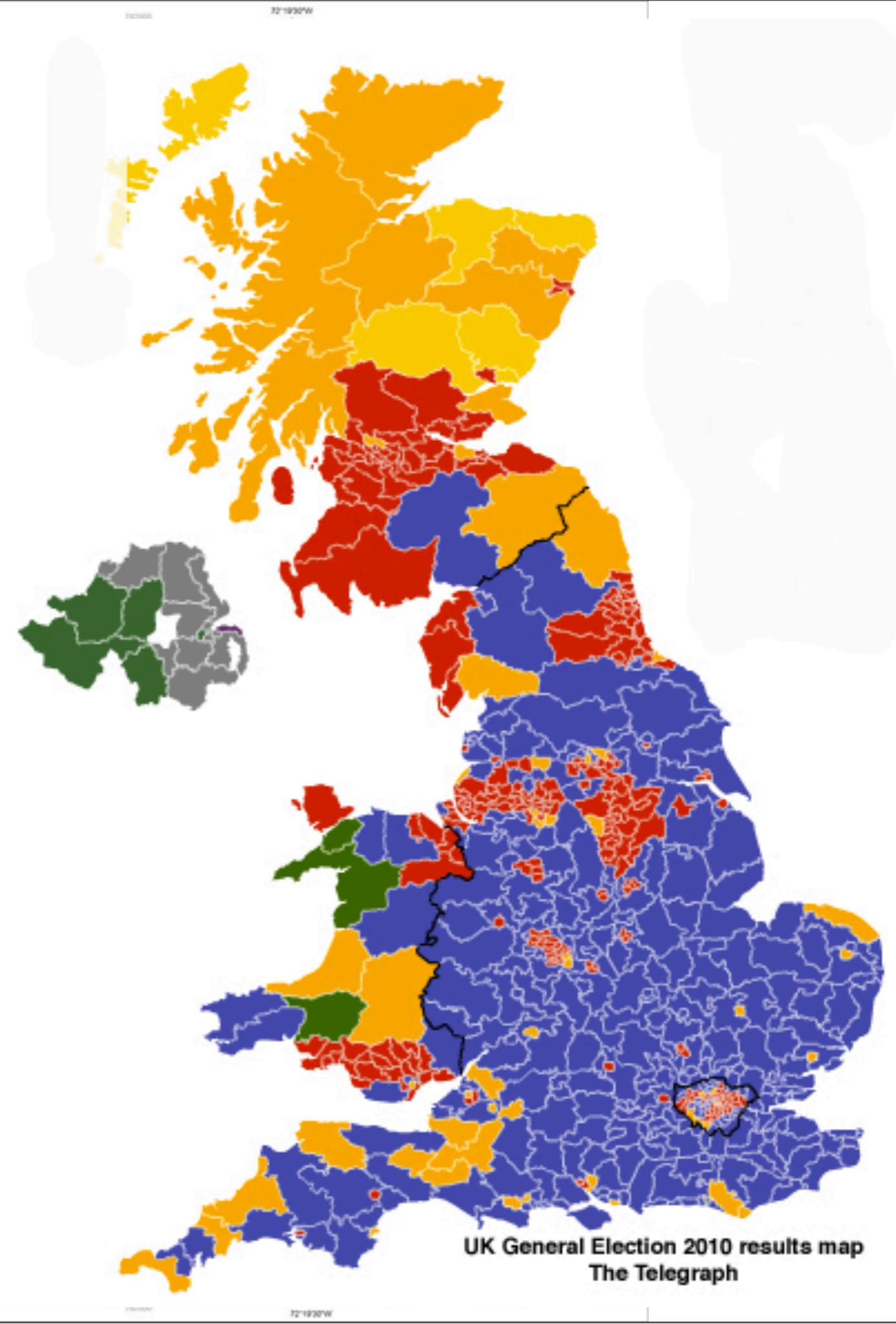
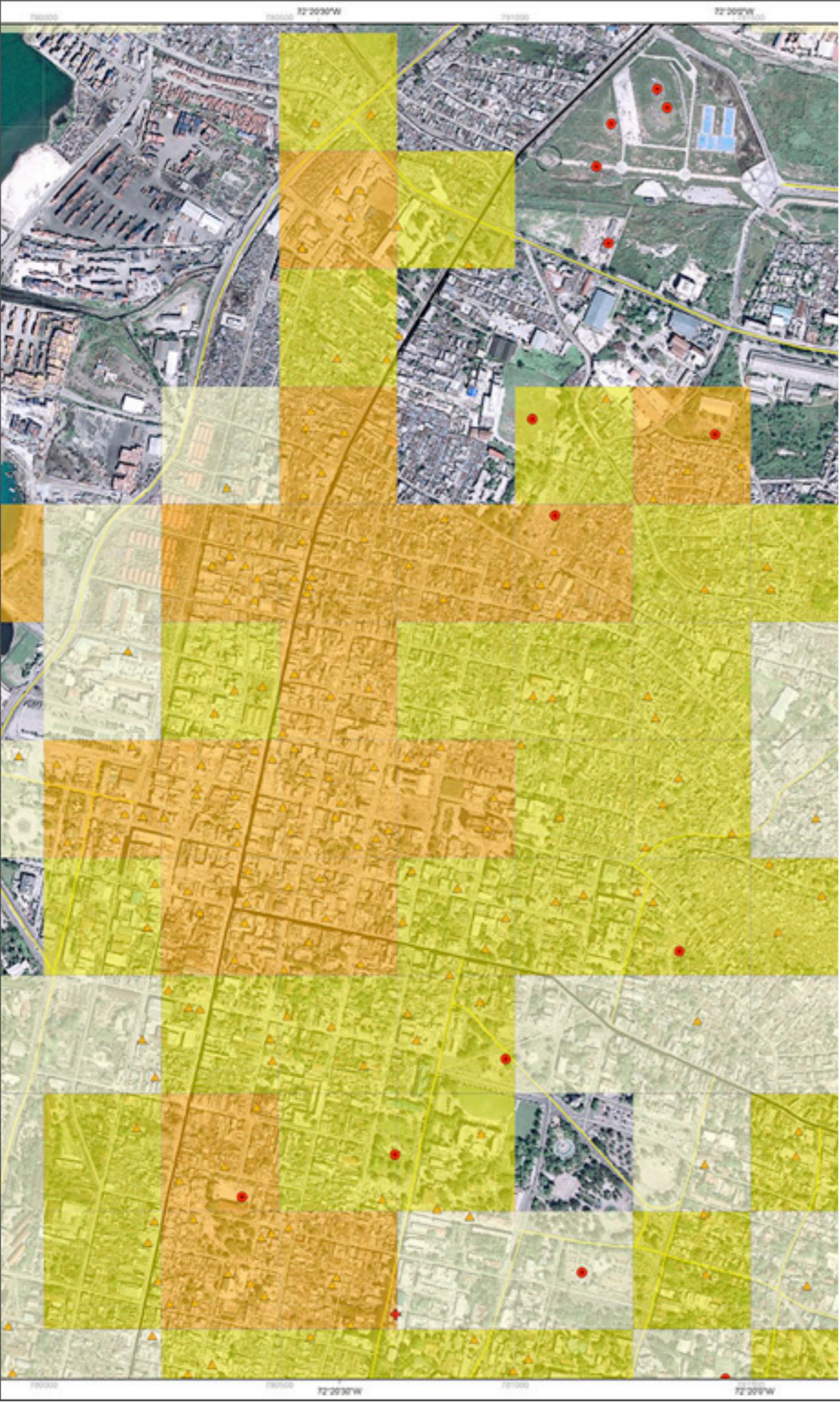
## problem 4: Details: overemphasised or obscured



hue ‘borders’ overemphasise small changes, hue ‘middles’ blend potentially important details

Using colour for continuous values  
problem 5: **pop out** can drown out





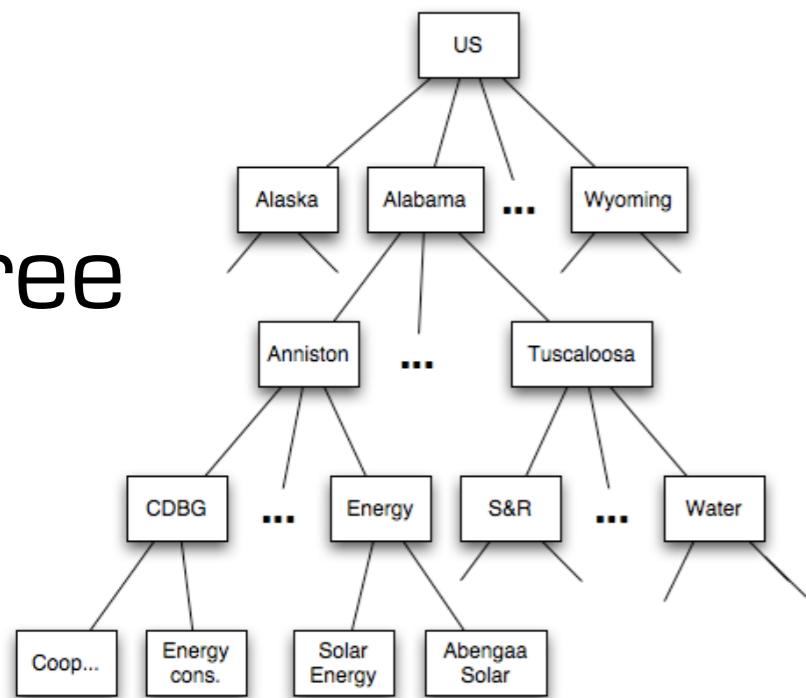
categorical, hierarchical values?



# multivariate relational data: hierarchical

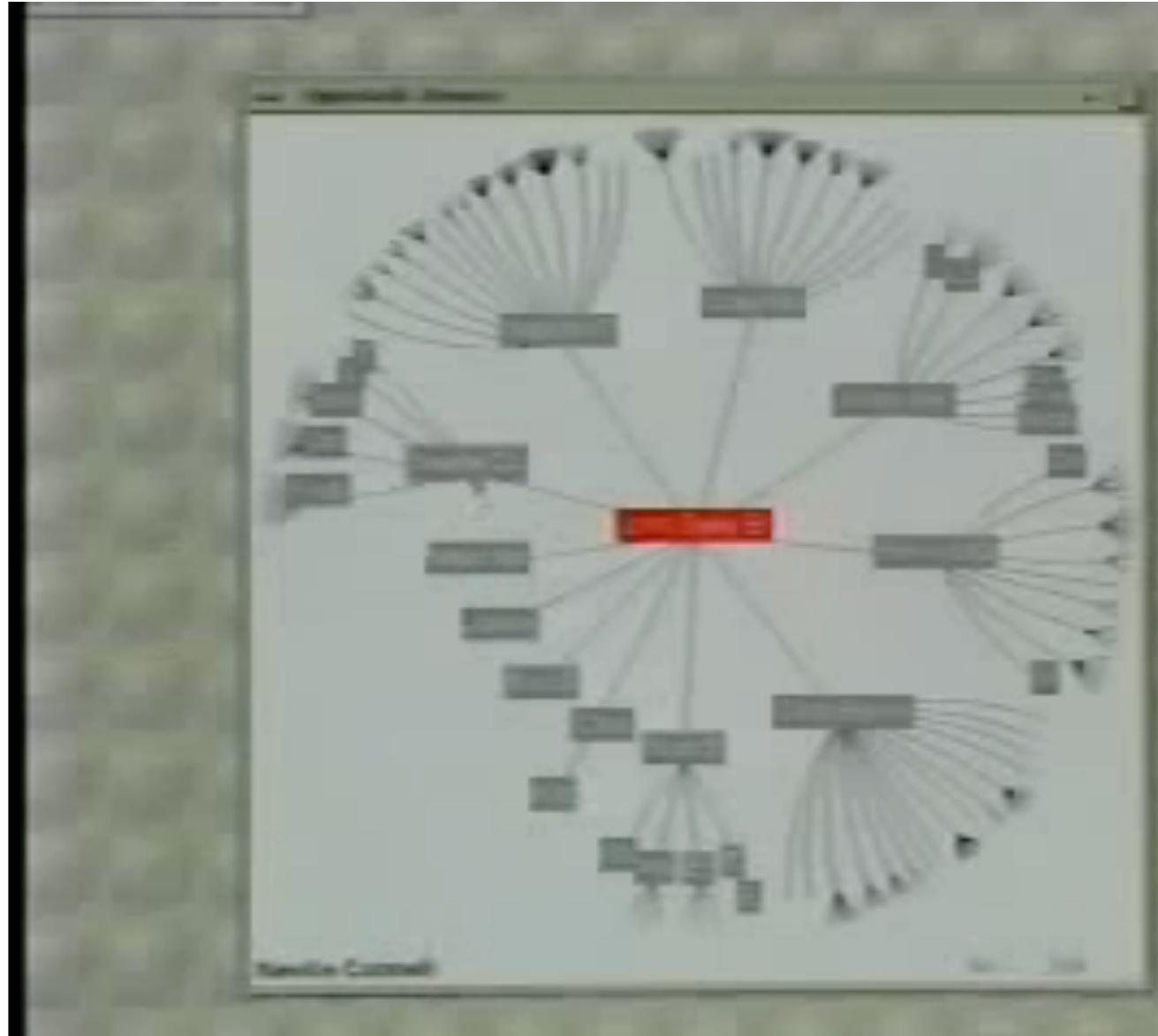
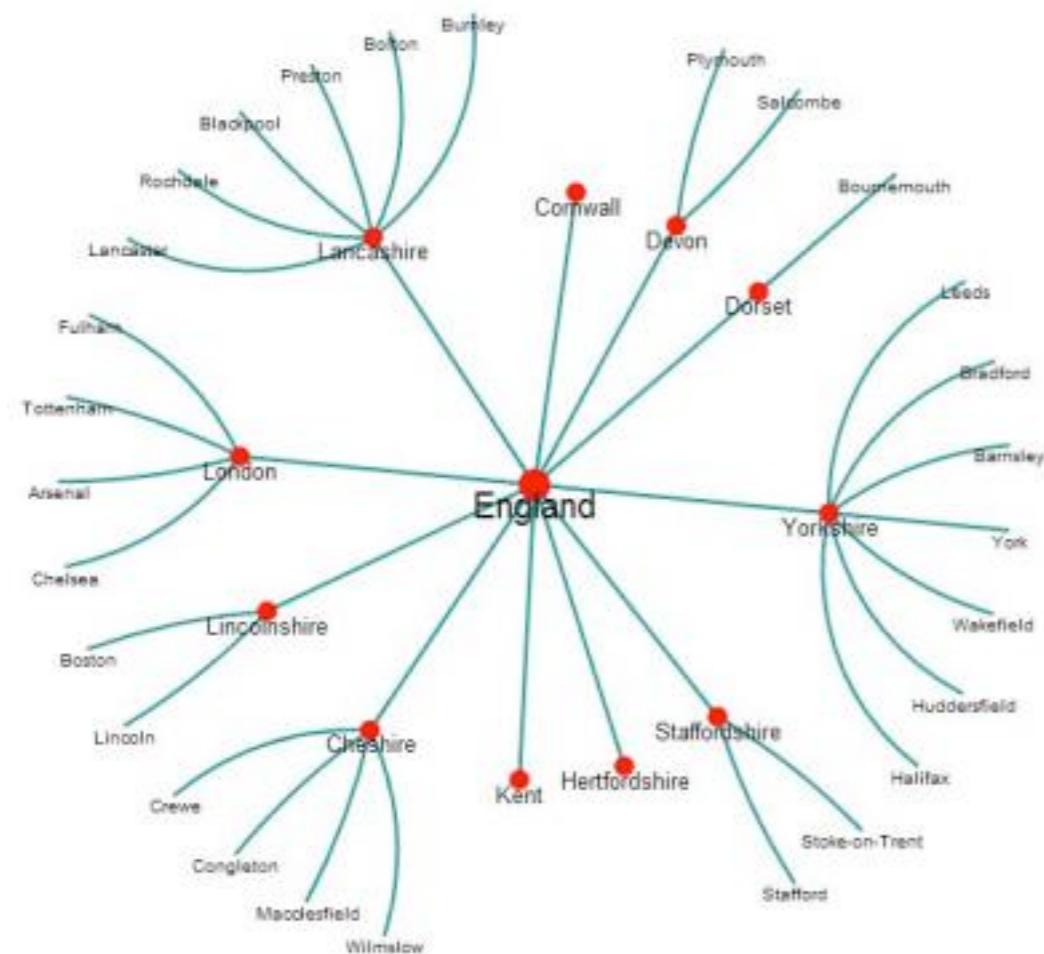
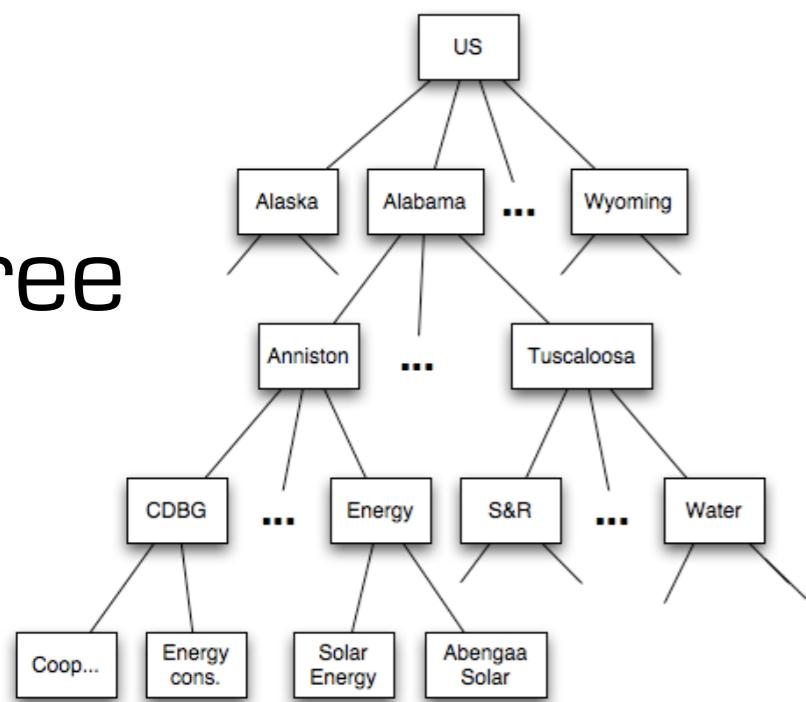
# multivariate relational data: hierarchical

tree



# multivariate relational data: hierarchical

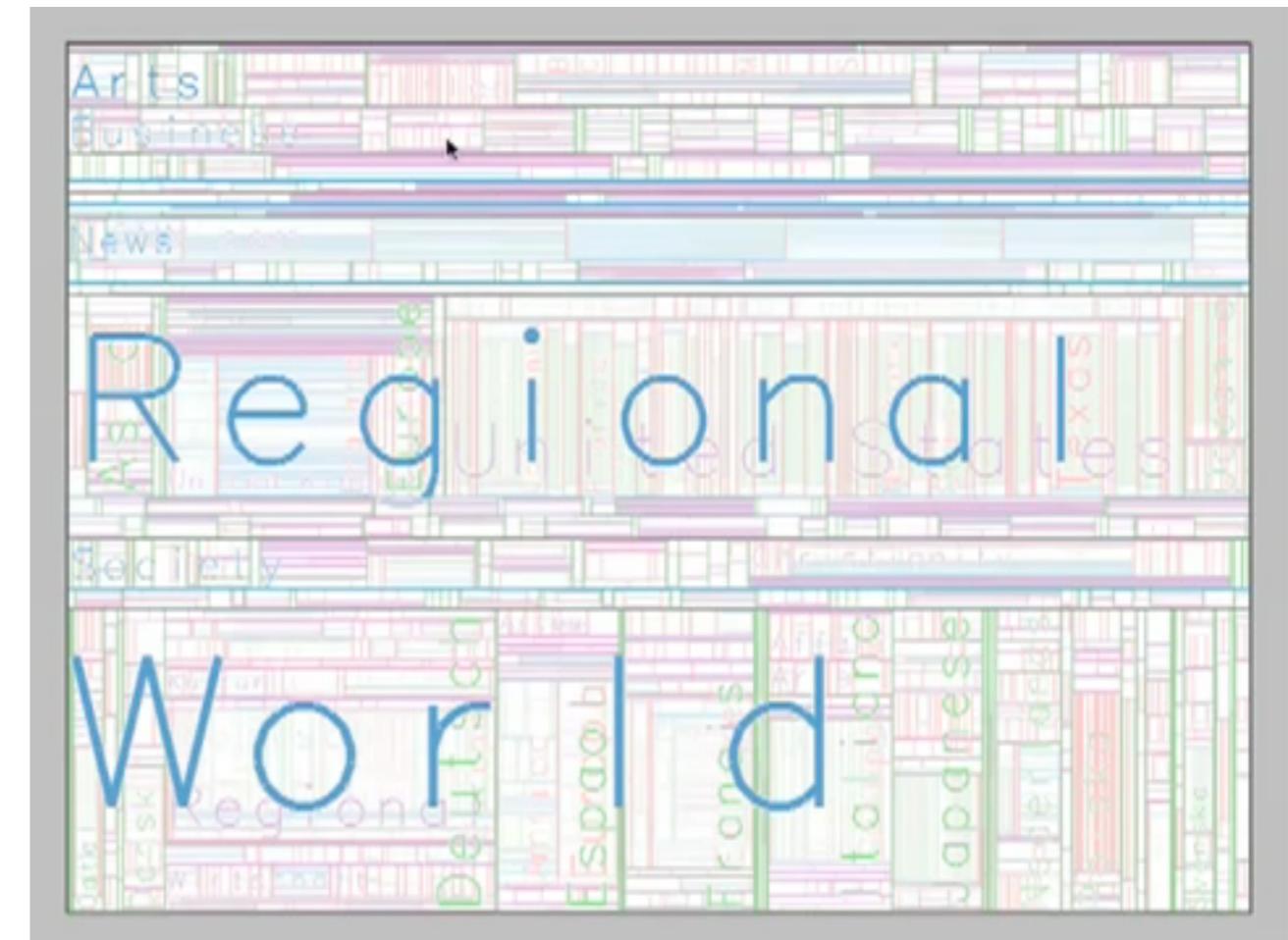
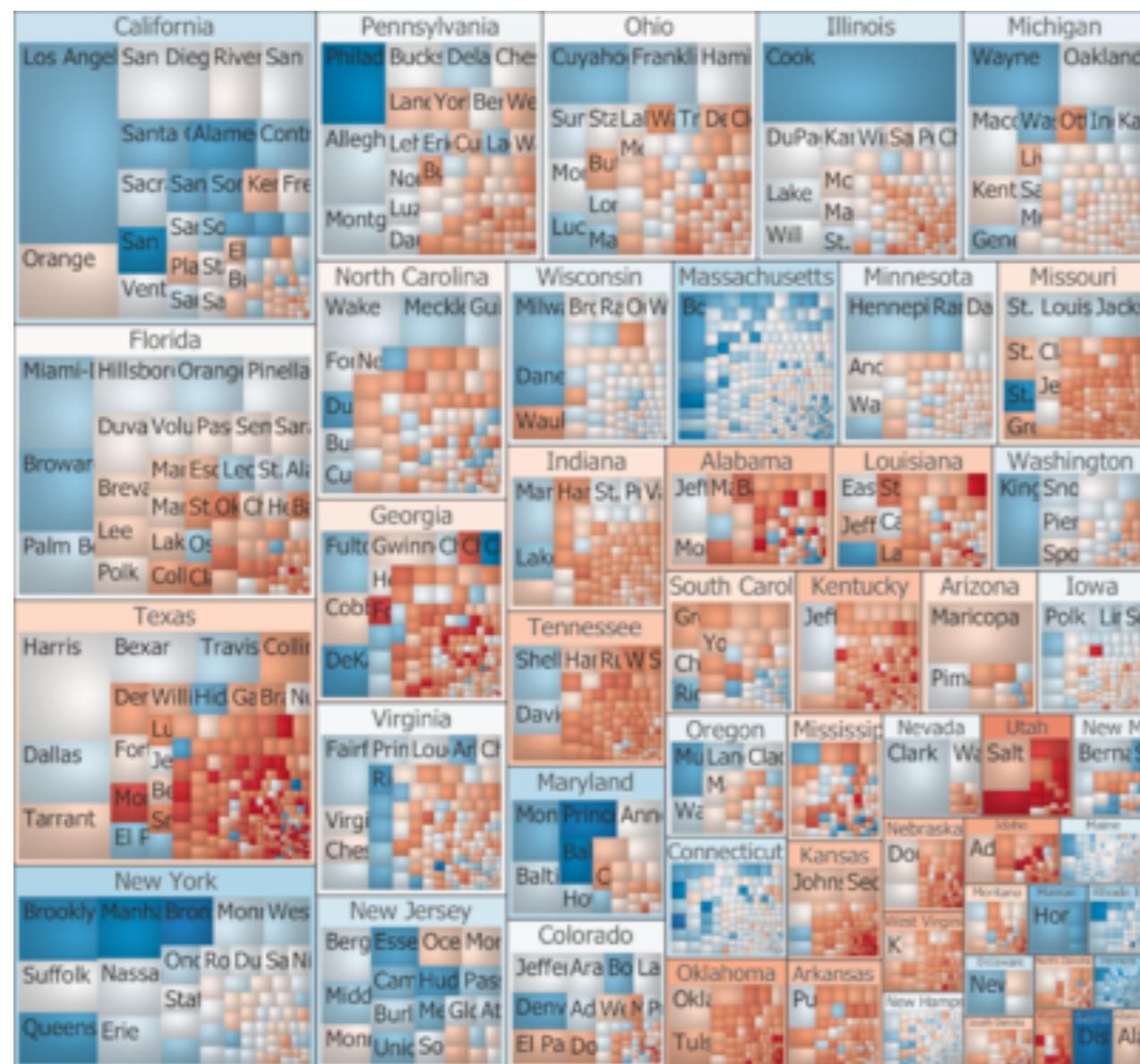
# tree



## hyperbolic tree

# multivariate relational data: hierarchical

# treemap

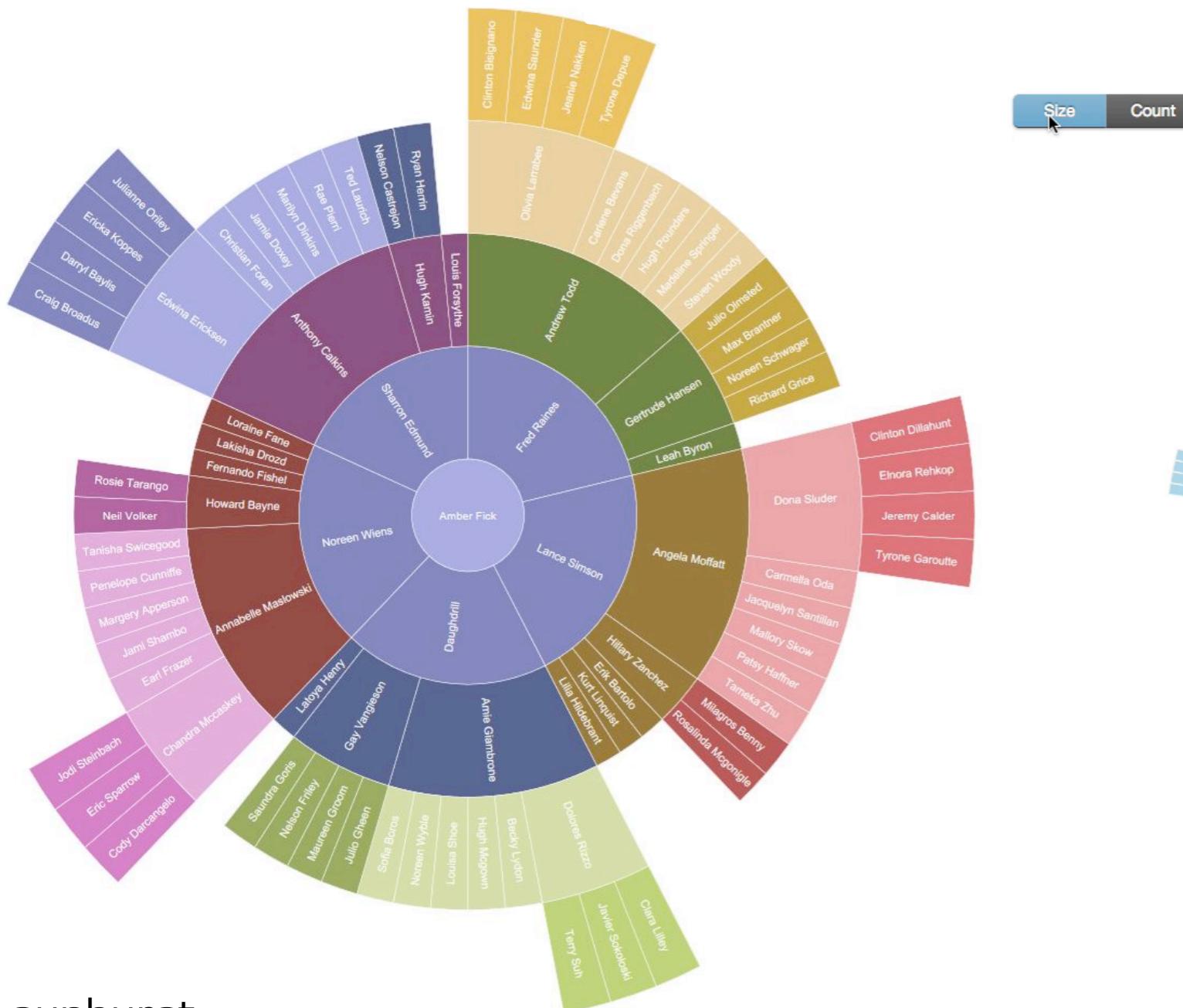


# multivariate relational data: hierarchical



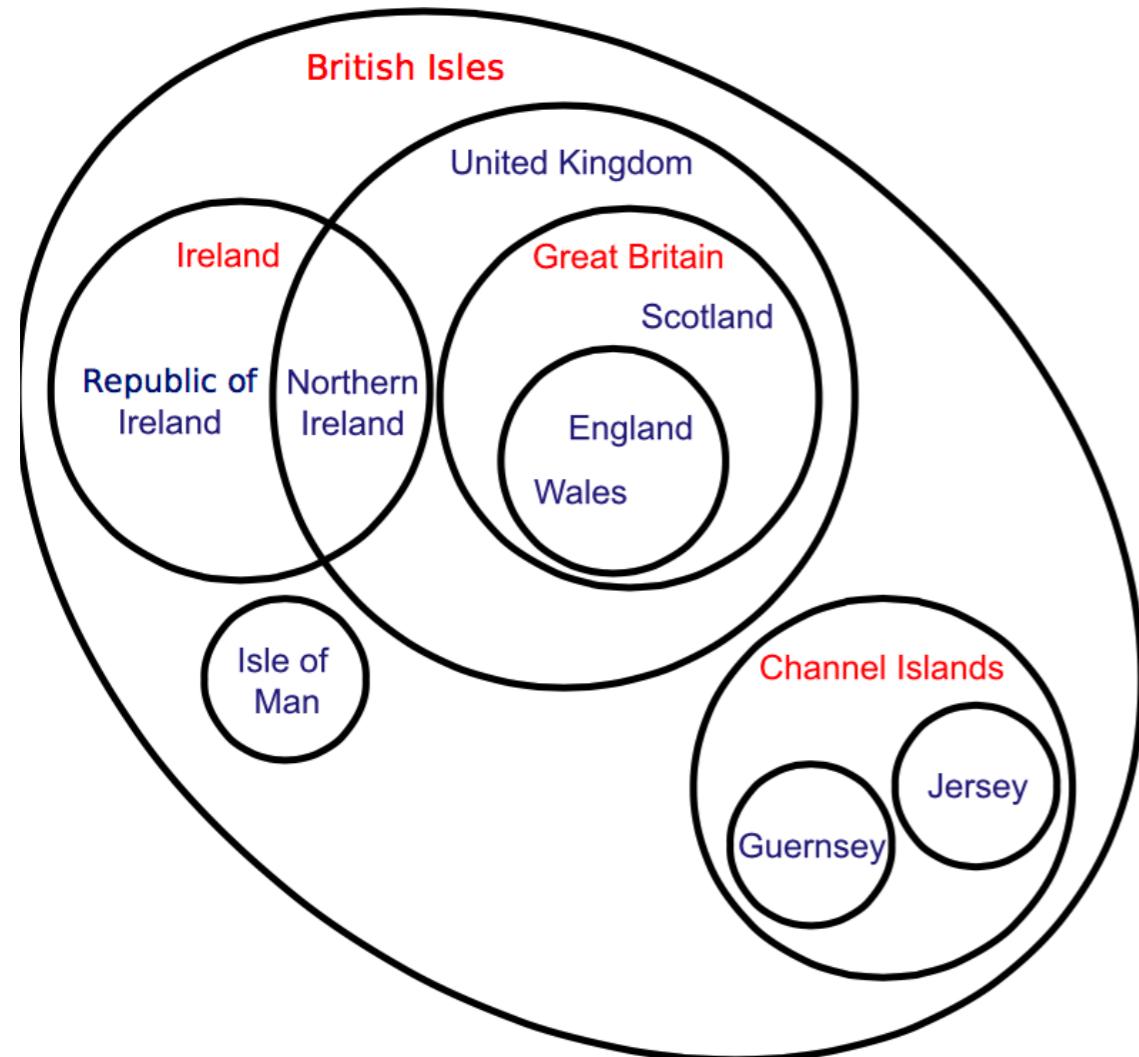
sunburst

# multivariate relational data: hierarchical



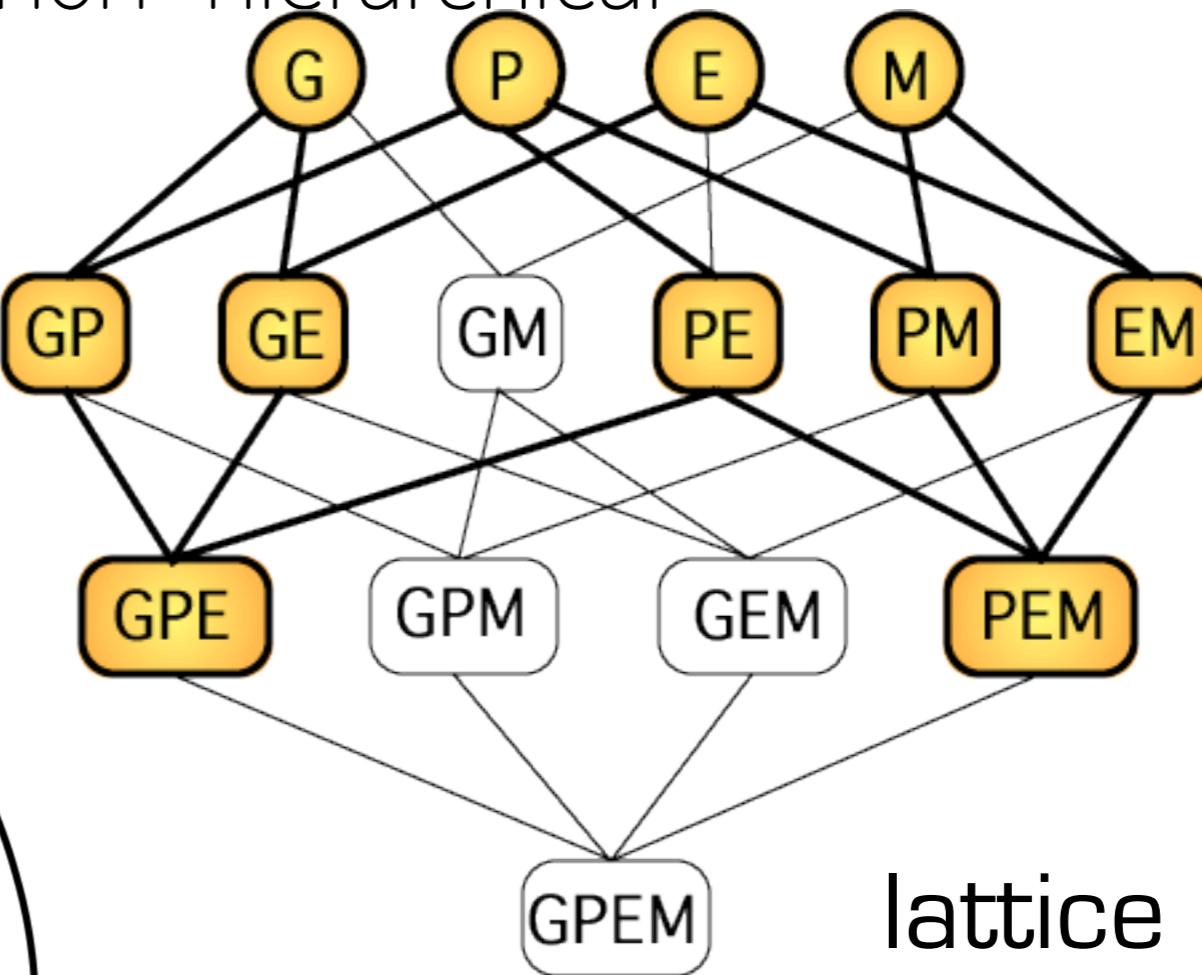
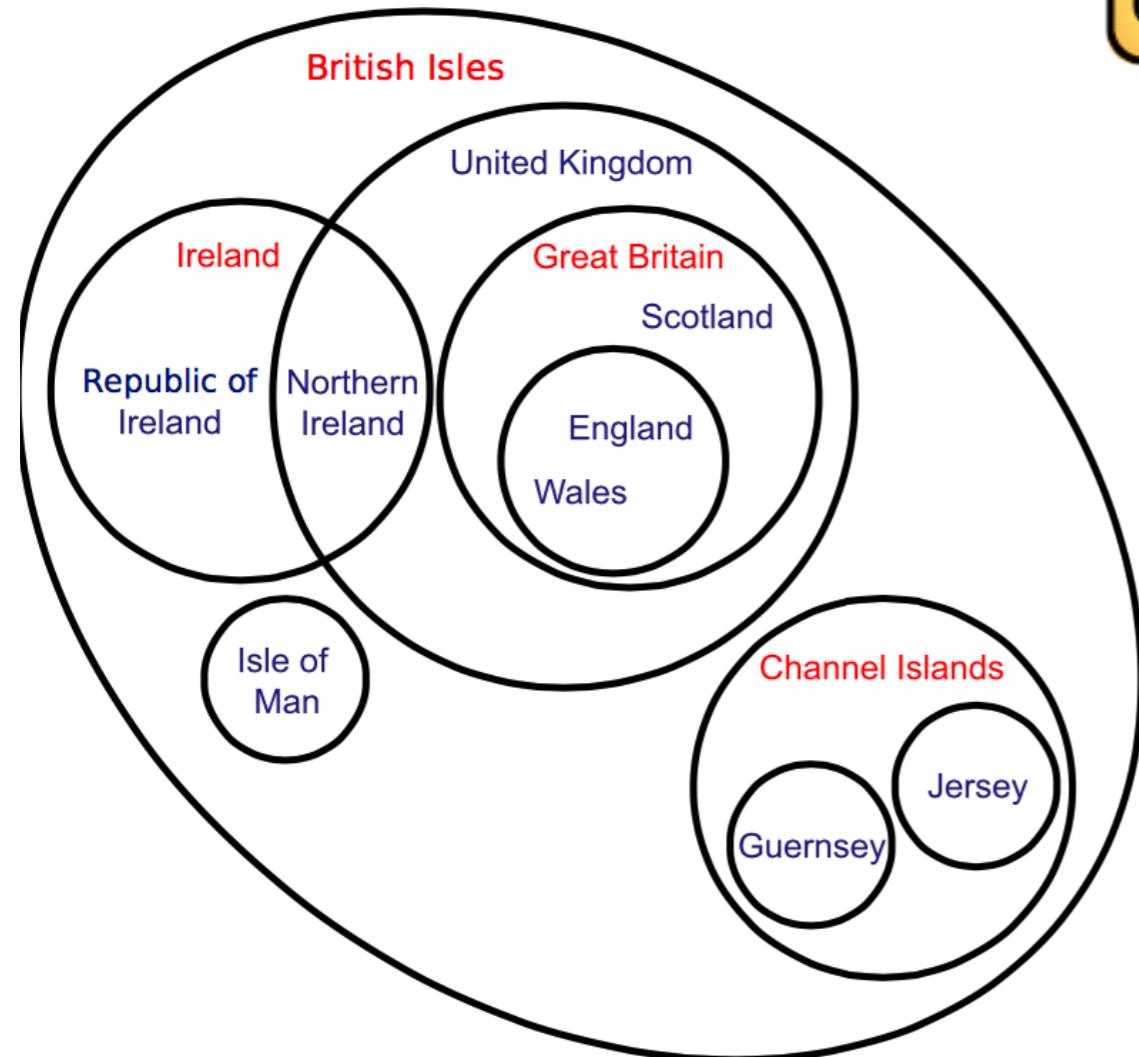
# sunburst

# multivariate relational data: non-hierarchical



venn diagram

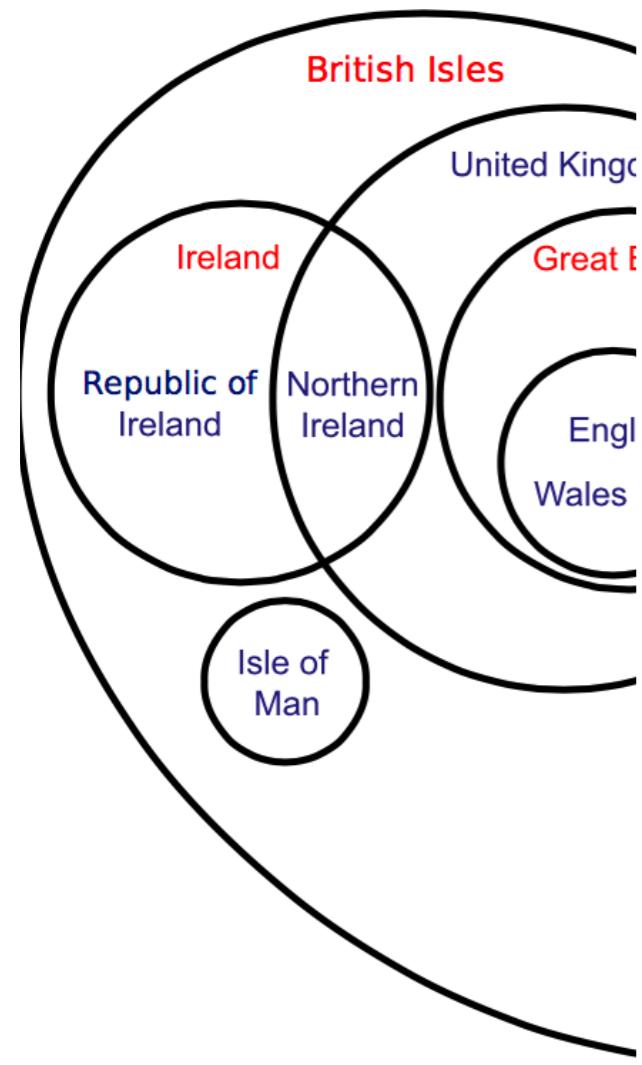
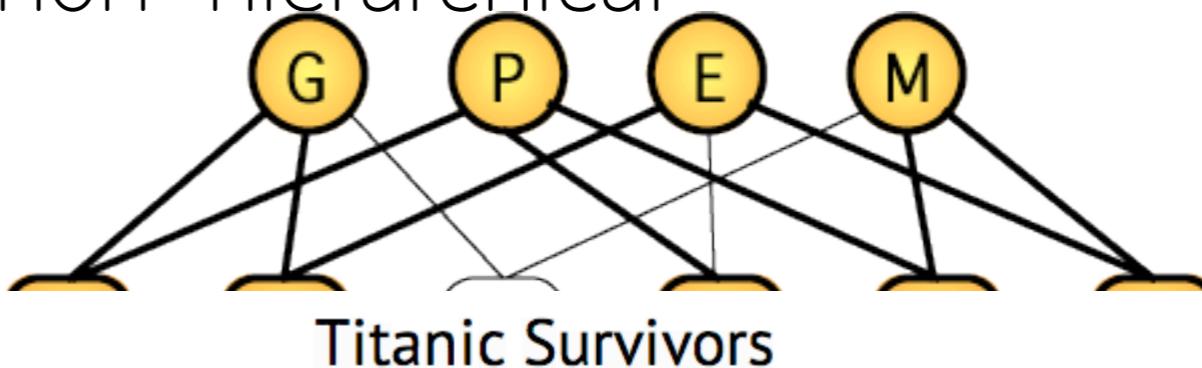
multivariate relational data: non-hierarchical



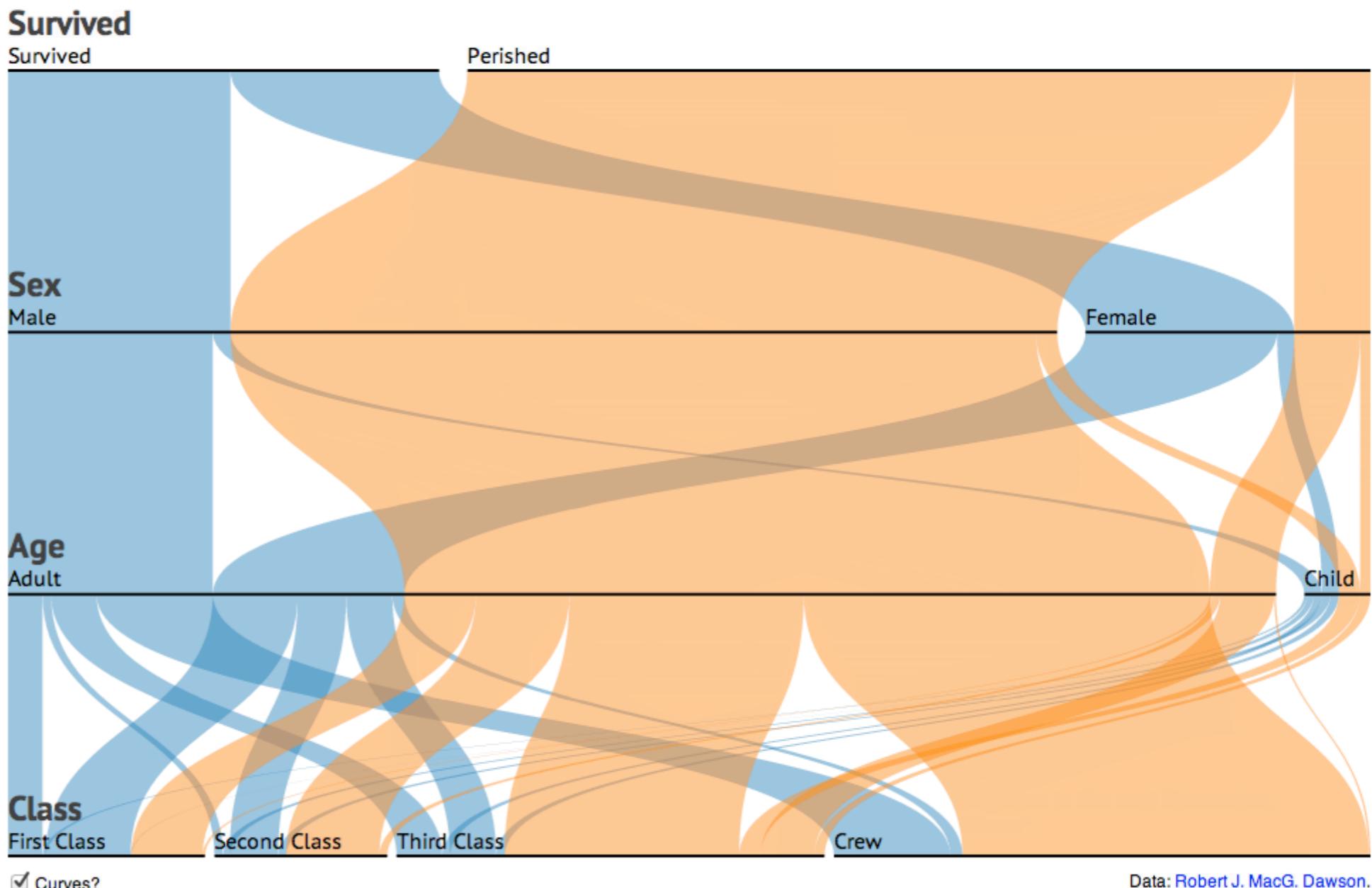
lattice

venn diagram

# multivariate relational data: non-hierarchical

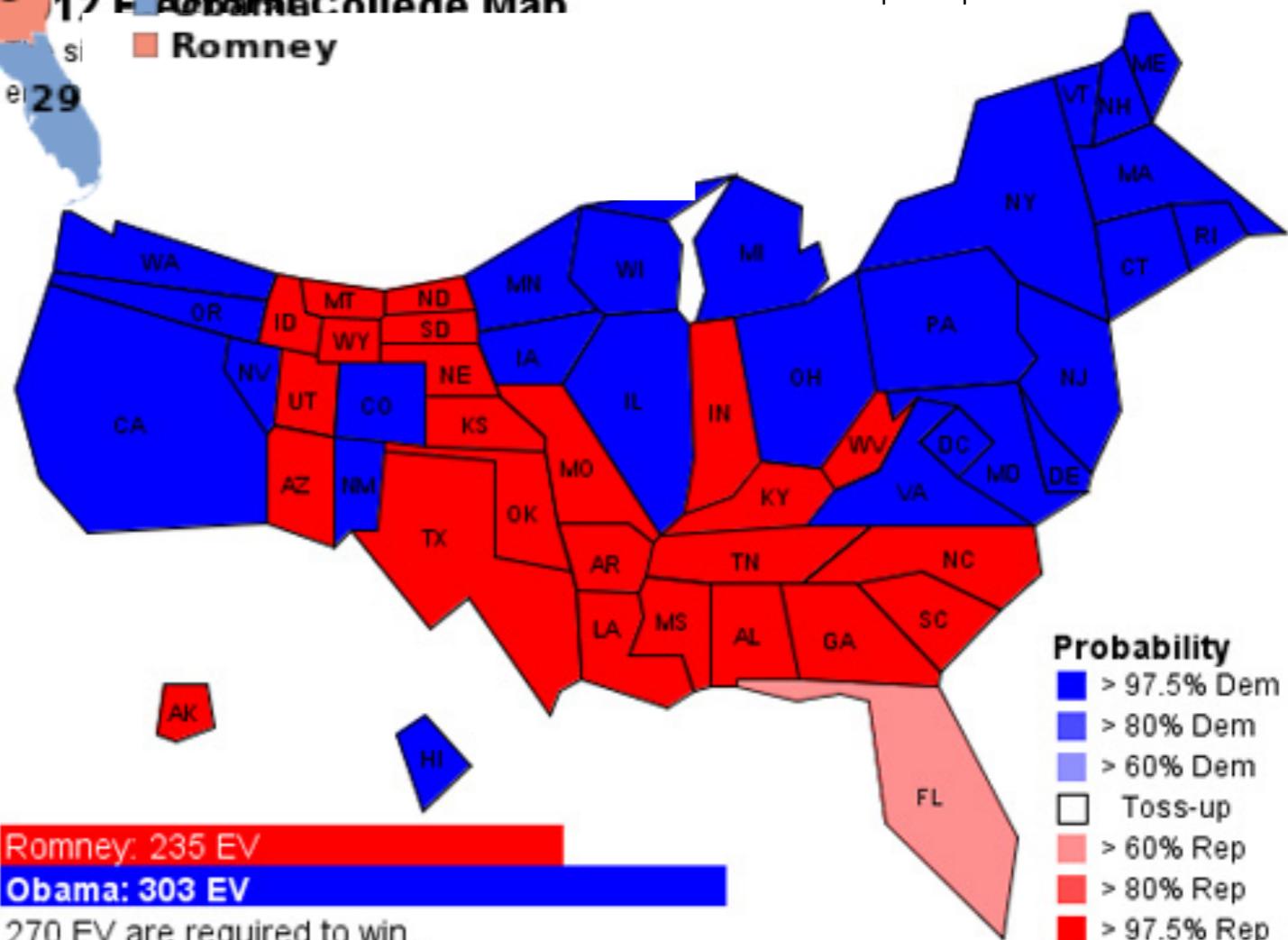
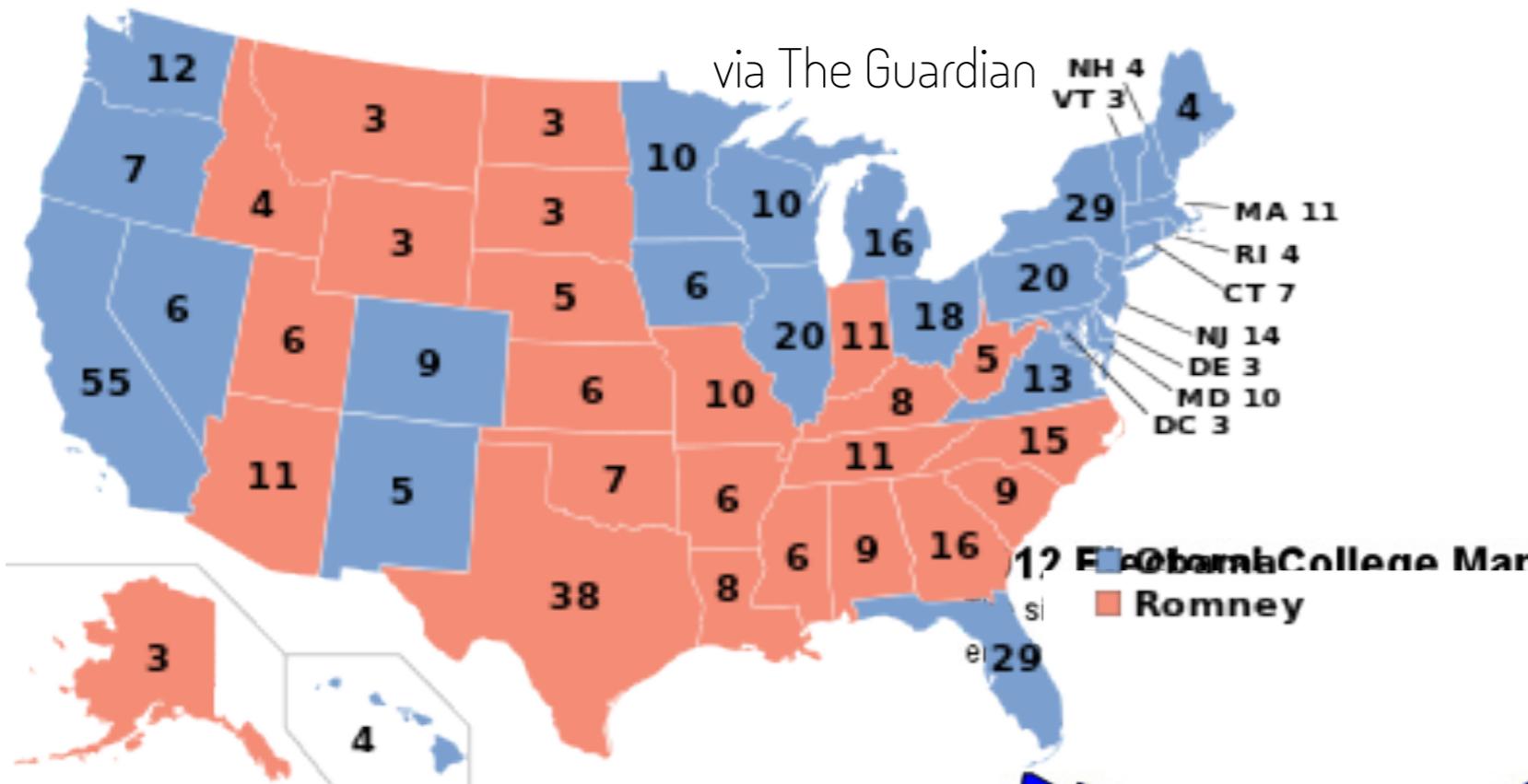


## venn diagram



## parallel sets

# multivariate geospatial data

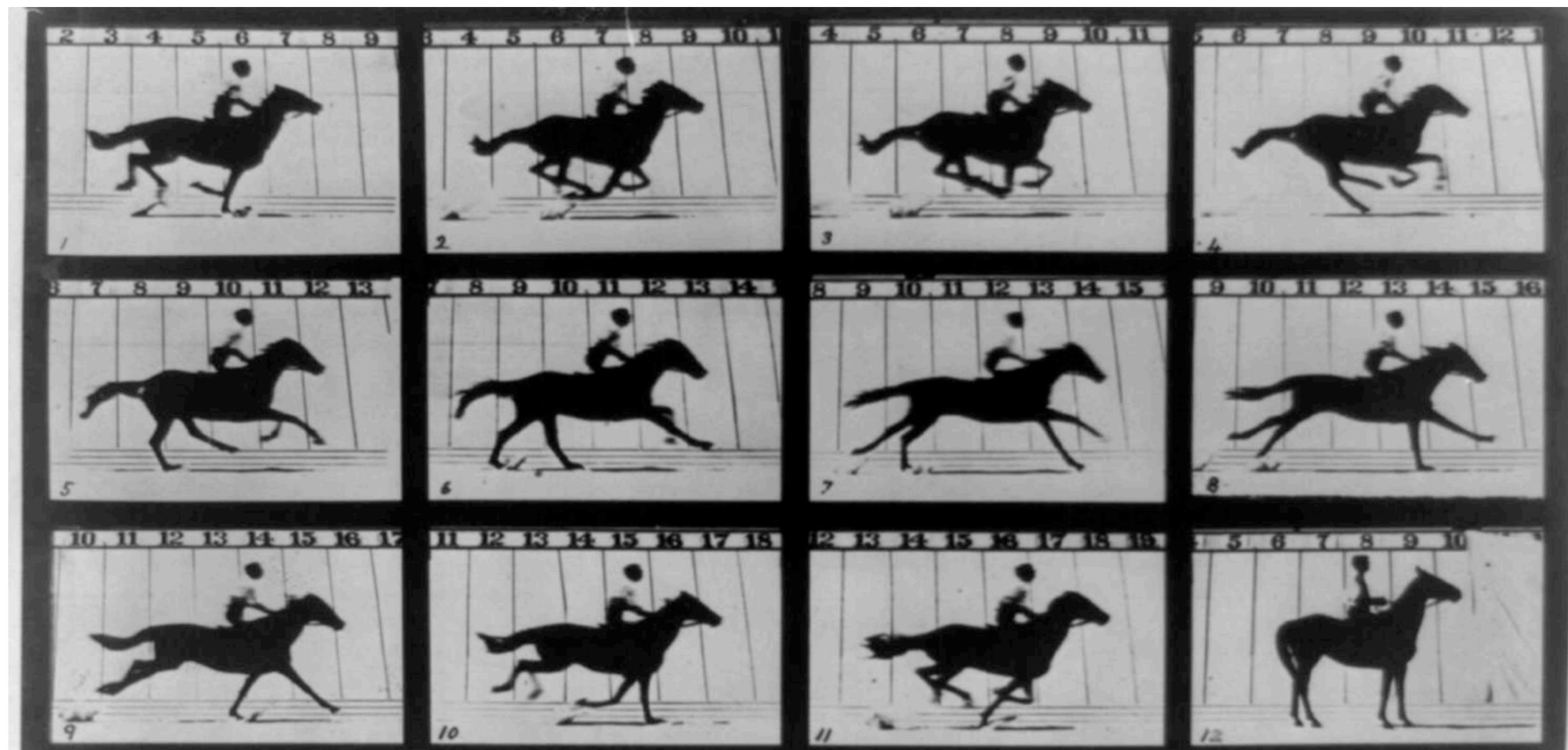
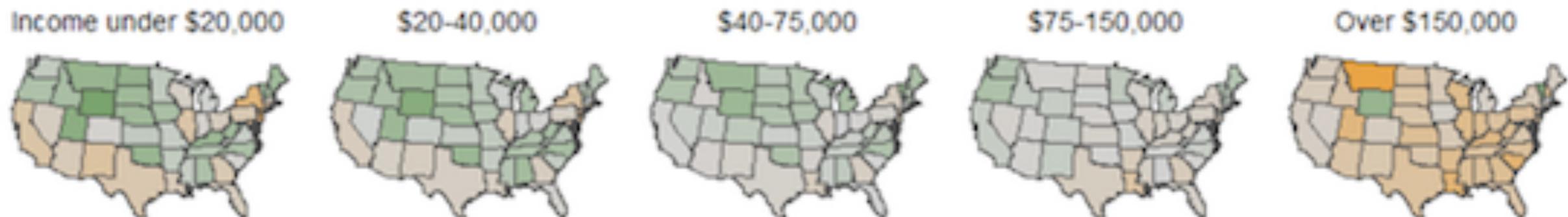


(via <http://zompist.wordpress.com/>)

Obama-Romney 2012  
victories by state

distortion added to make  
area proportional to votes

# time series (static) – small multiples



Copyright, 1878, by MUYBRIDGE.

MORSE'S Gallery, 417 Montgomery St., San Francisco

THE HORSE IN MOTION.

Illustrated by

MUYBRIDGE

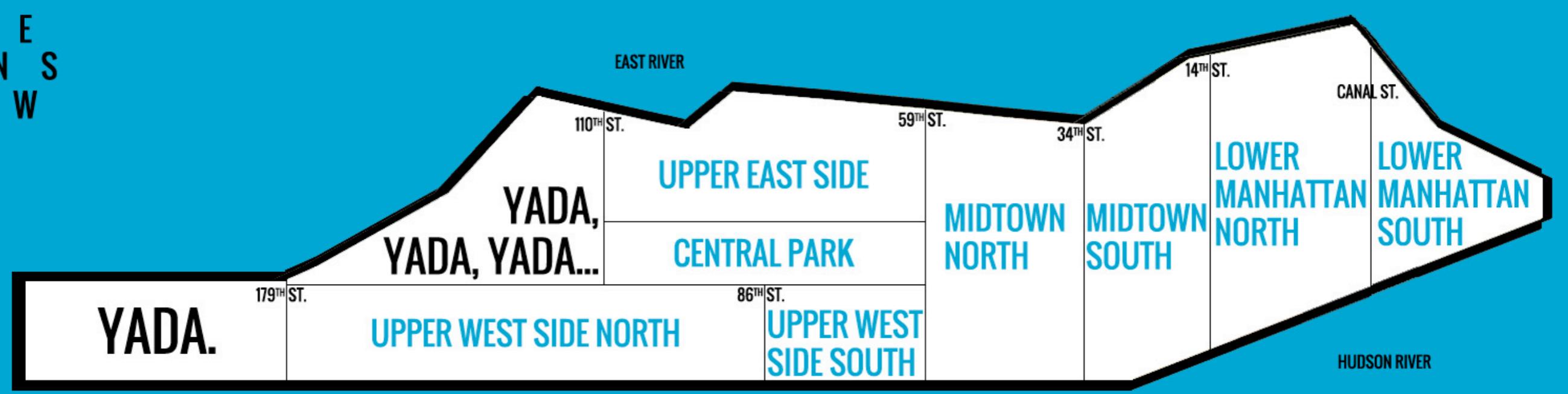
Patent for apparatus applied for

INTERNATIONAL EXHIBITION, PHILADELPHIA, 1876.

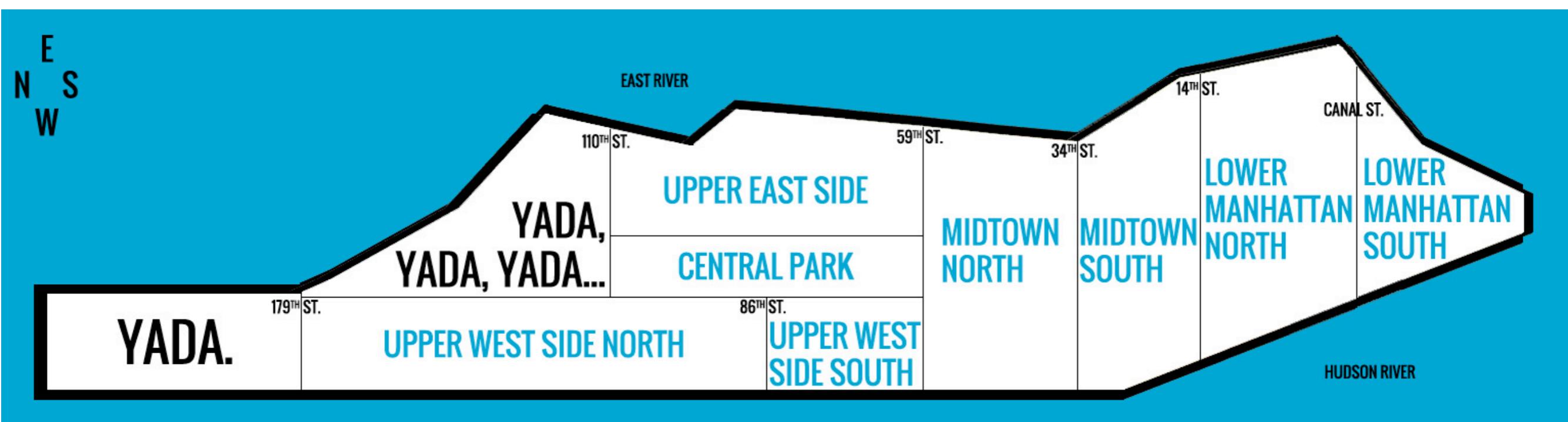
# time series (animation)

aaron koblin - flight patterns

E  
N  
S  
W



visualisation is communication



visualisation is designed narrative

*Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.*

Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.

Paris, le 20 Novembre 1869.

Les nombres d'hommes perdus sont représentés par les larges des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui ont péri en Russie; le noir ceux qui en sont sortis. — Les renseignements qui ont servi à dresser la carte ont été pris dans les ouvrages de M. Chier, de Segur, de Fezensac, de Chambray et le journal intime de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Smolensk et qui rejoignirent Ossatch et Witebsk, avaient toujours marché avec l'armée.

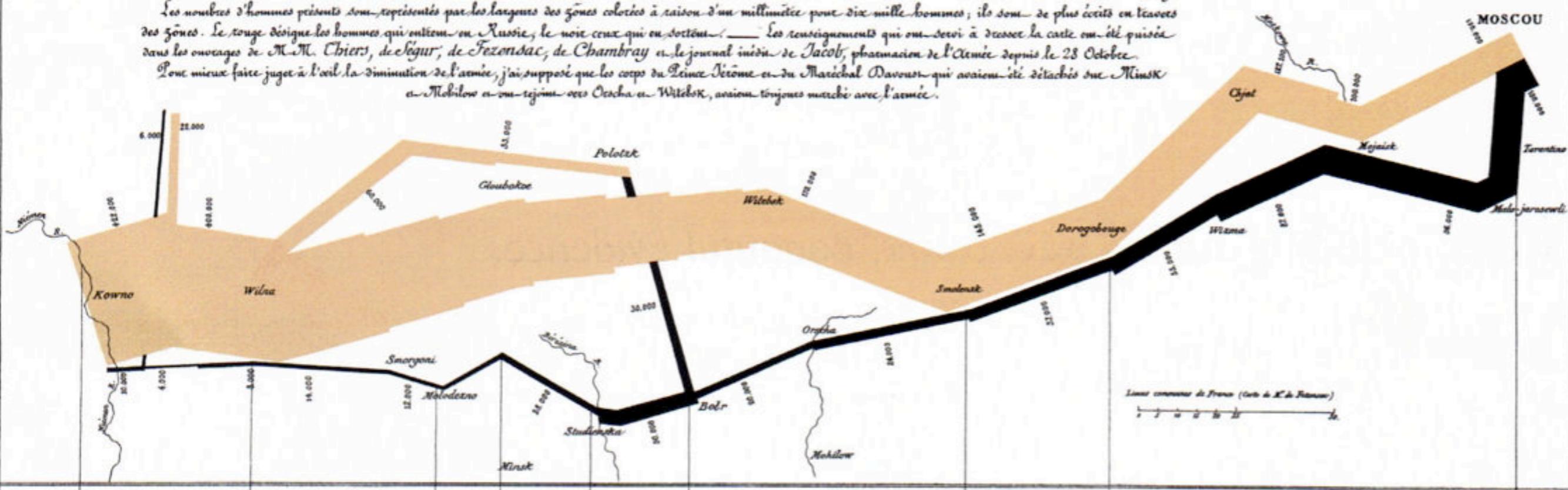


TABLEAU GRAPHIQUE de la température en degrés du thermomètre de Réaumur au dessous de zéro.

Les Cosaques passent au galop  
le Nilovka gelé.

- 26° le 7 X.  
- 30° le 6 X.  
- 24° le 1<sup>er</sup> X.  
- 20° le 28 9<sup>me</sup>.  
- 11°.

- 21° le 14 9<sup>me</sup>.  
- 8° le 9 9<sup>me</sup>.

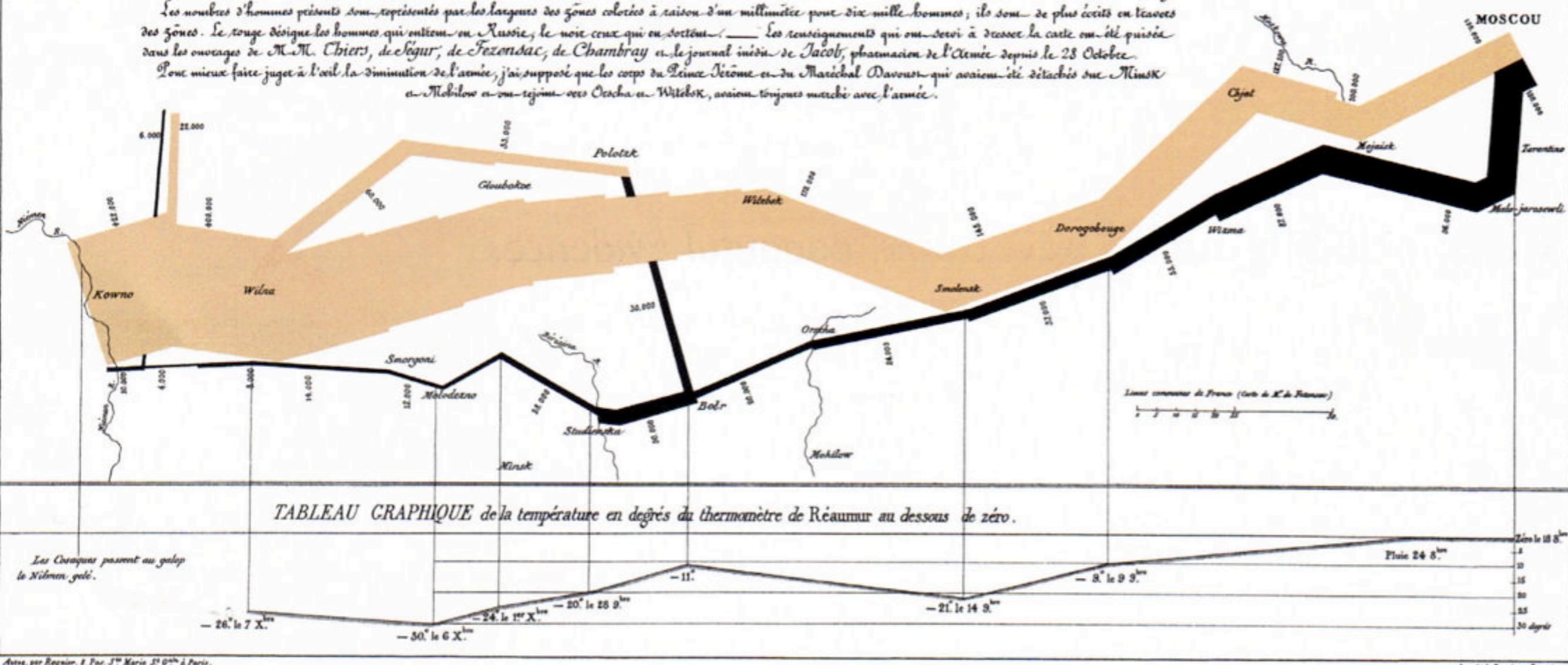
Zéro le 15 8<sup>me</sup>.  
- 1  
- 2  
- 3  
- 4  
- 5  
- 6  
- 7  
- 8  
- 9  
- 10  
- 11  
- 12  
- 13  
- 14  
- 15  
- 16  
- 17  
- 18  
- 19  
- 20  
- 21  
- 22  
- 23  
- 24  
- 25  
- 26  
- 27  
- 28  
- 29  
- 30 degrés.

*Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.*

dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite — Paris, le 20 Novembre 1869.

Les nombres d'hommes perdus sont représentés par les larges des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui ont péri en Russie; le noir ceux qui en sont sortis. — Les renseignements qui ont servi à dresser la carte ont été pris dans les ouvrages de M. Chier, de Segur, de Fezensac, de Chambray et le journal intime de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Malibow aient rejoint vers Ossaka en Witlob, avant de toujours marcher avec l'armée.



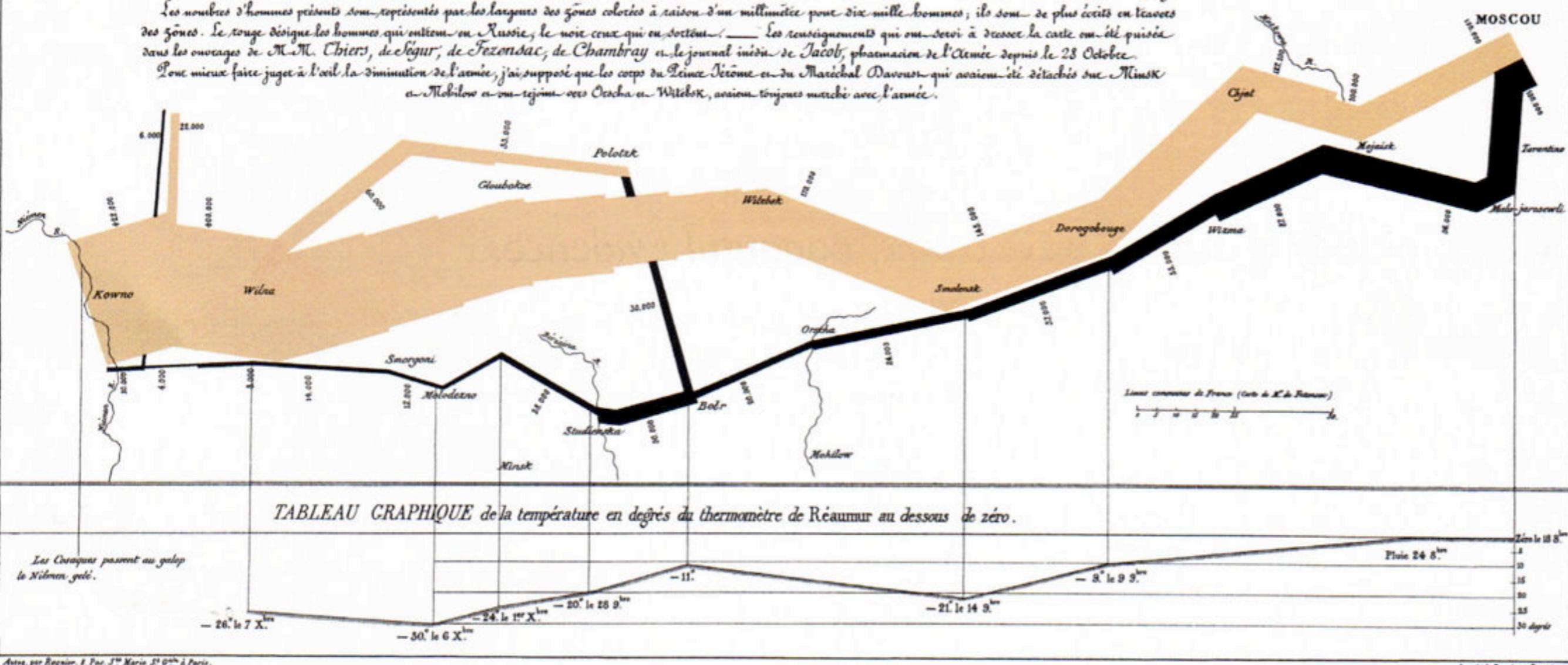
napoleon's march to moscow  
charles joseph minard

*Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.*

dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite — Paris, le 20 Novembre 1869.

Les nombres d'hommes perdus sont représentés par les larges des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui ont péri en Russie; le noir ceux qui en sont sortis. — Les renseignements qui ont servi à dresser la carte ont été pris dans les ouvrages de M. Chiers, de Segur, de Fezensac, de Chambray et le journal intime de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Malibow aient rejoint vers Ossaka en Witlobk, avaient toujours marché avec l'armée.



how many dimensions can you find?

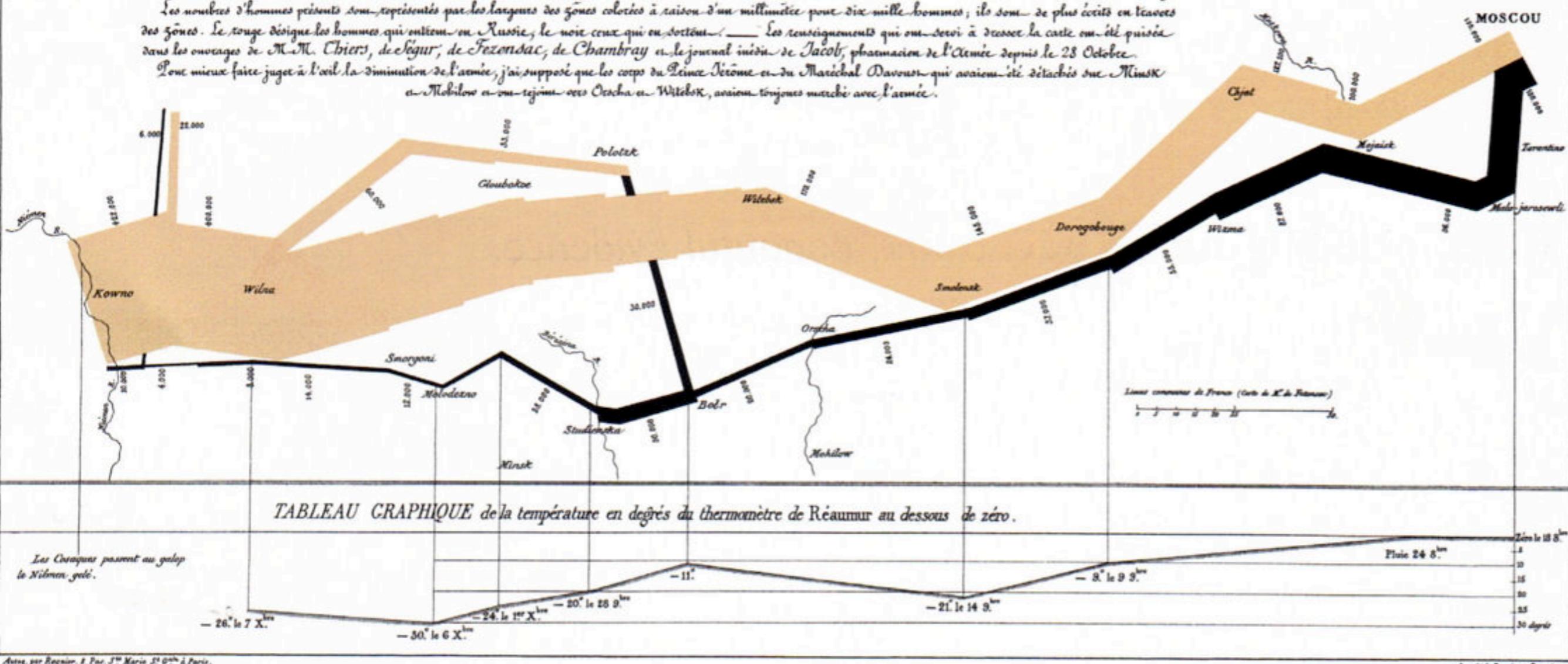
napoleon's march to moscow  
charles joseph minard

*Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.*  
 Dessiné par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite

Paris, le 20 Novembre 1869.

Les nombres d'hommes perdus sont représentés par les larges des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en lettres des zones. Le rouge désigne les hommes qui ont péri en Russie; le noir ceux qui en sont sortis. — Les renseignements qui ont servi à dresser la carte ont été pris dans les ouvrages de M. Chiers, de Segur, de Fezensac, de Chambray et le journal intime de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout qui avaient été détachés sur Minsk et Malibow aient rejoint vers Ossaka en Witlobk, avaient toujours marché avec l'armée.

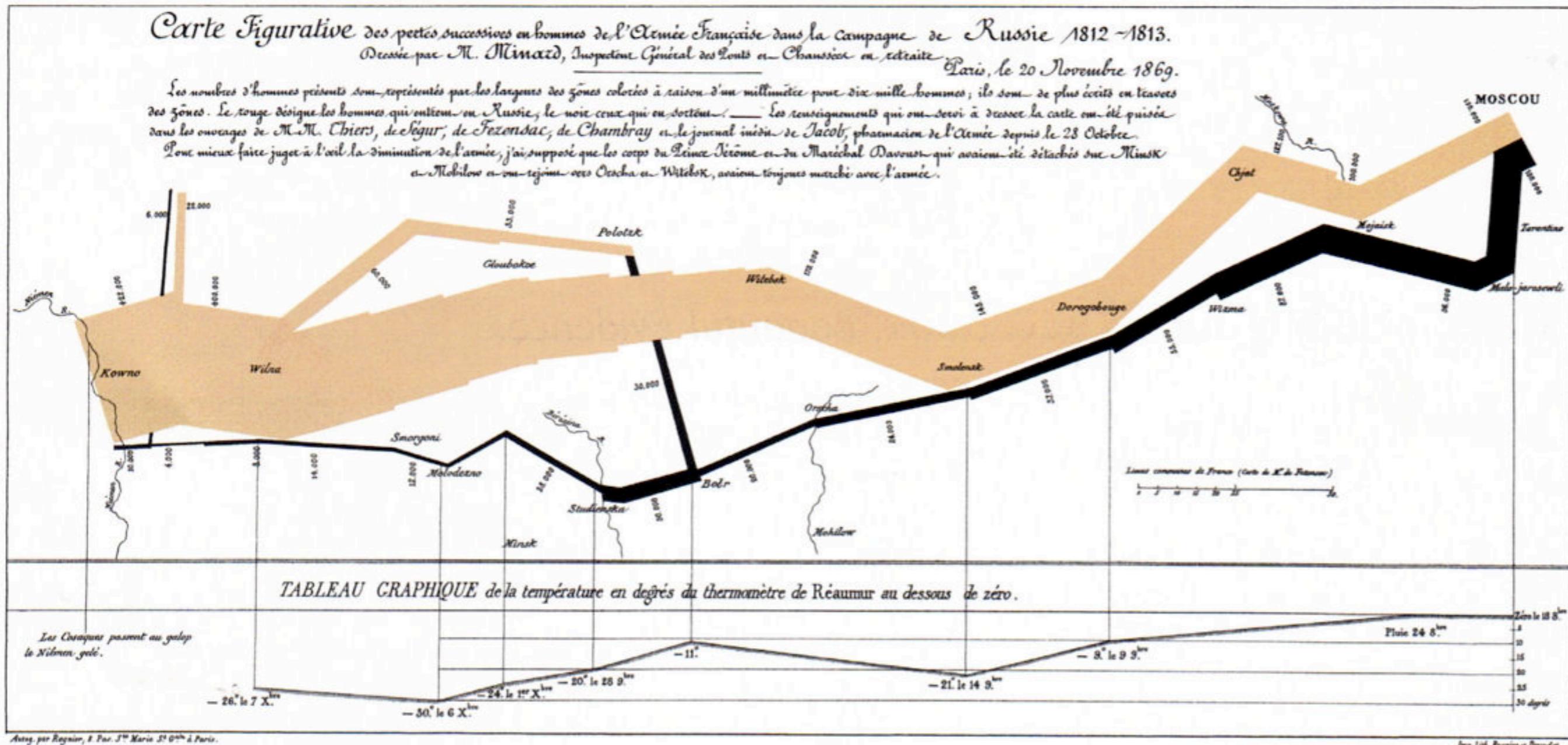


how many dimensions can you find?

- ans: 1) size of the army 2-3) path (lat/lng) taken on a map  
 4) direction army was traveling 5) temperature 6) dates army reached particular locations

napoleon's march to moscow  
 charles joseph minard

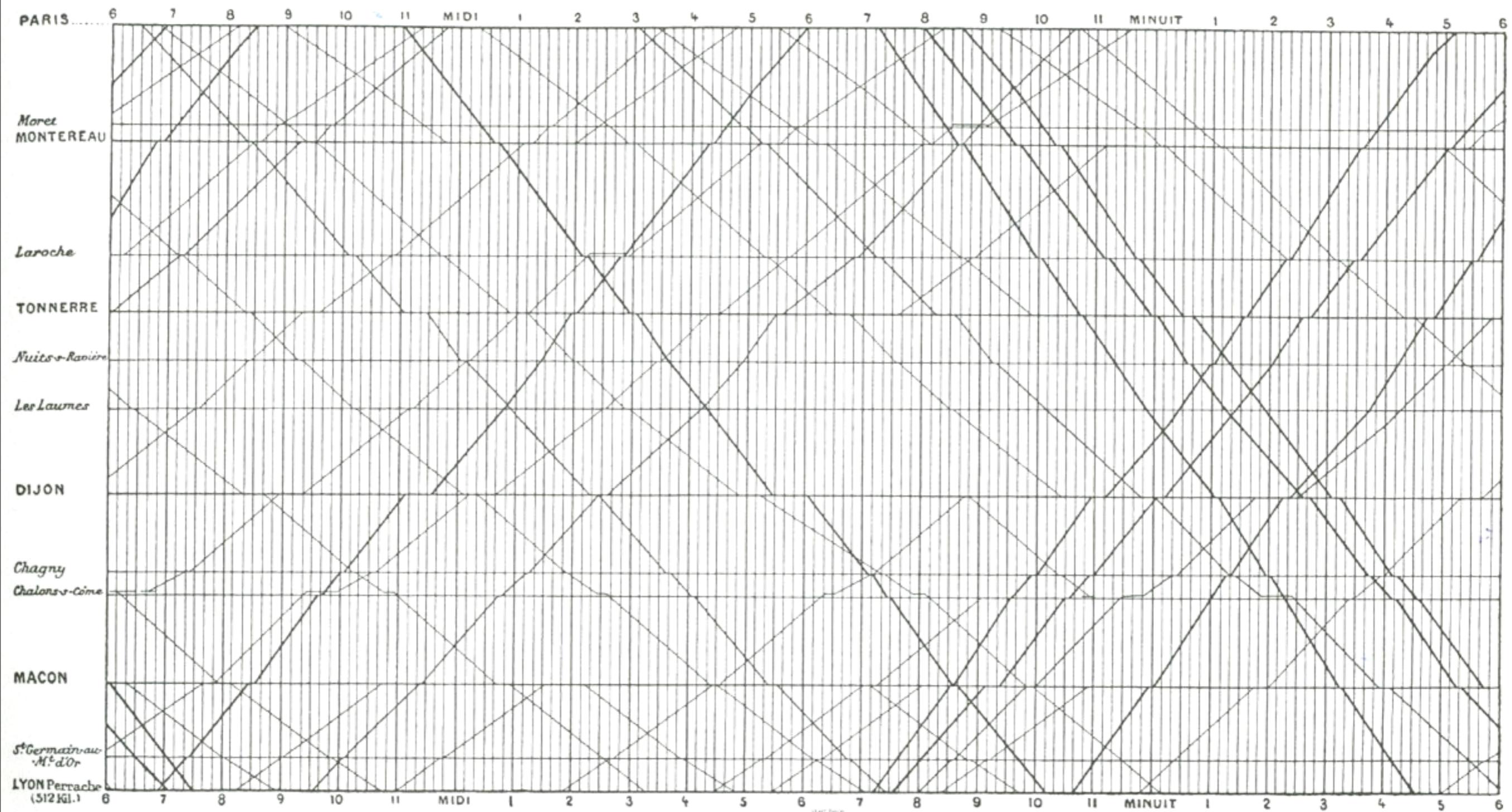
# geo-temporal-relational-continuous multivariate?!



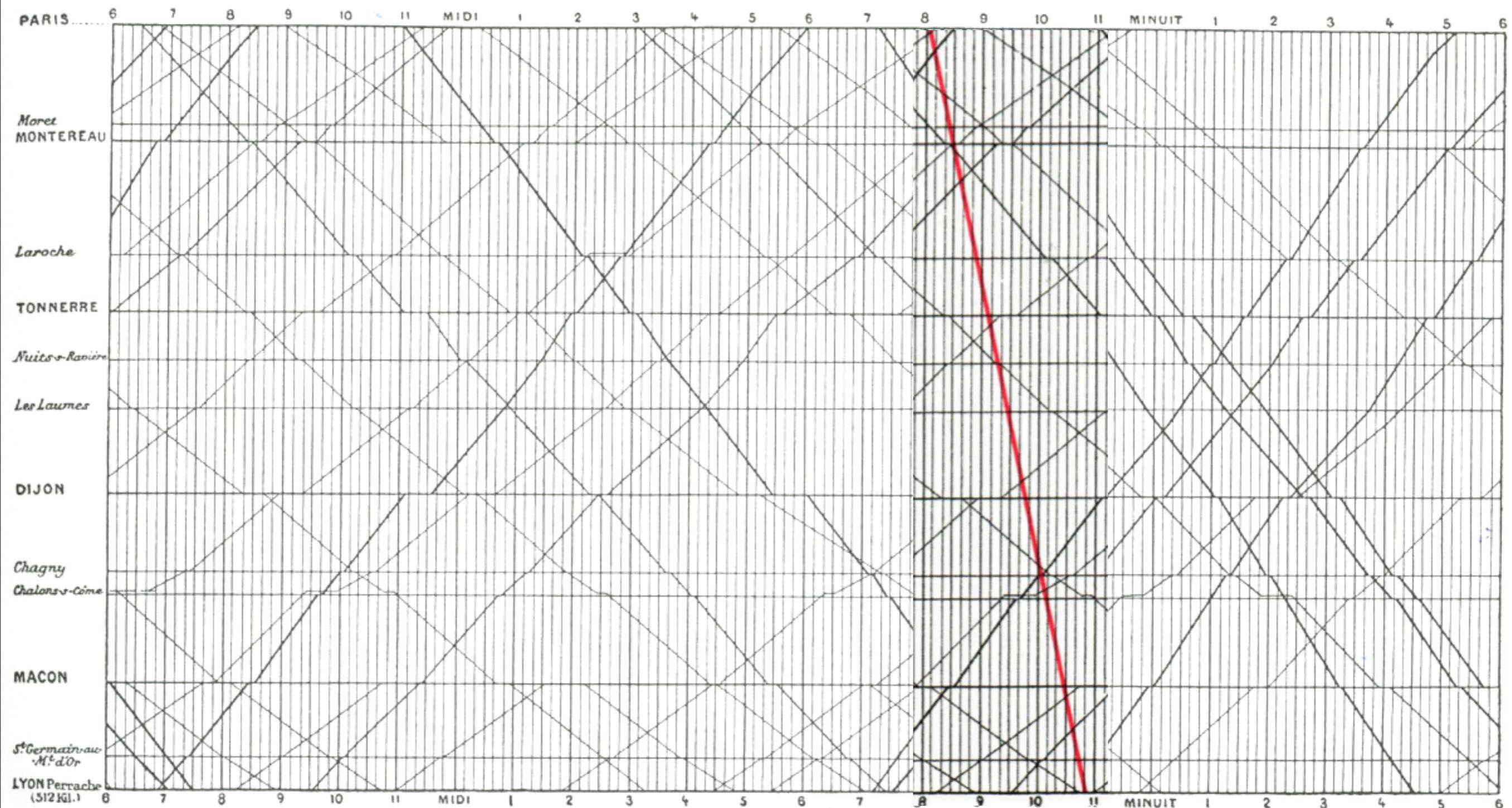
how many dimensions can you find?

- ans: 1) size of the army
- 2) path (lat/lng) taken on a map
- 3) direction army was traveling
- 4) temperature
- 5) dates army reached particular locations

napoleon's march to moscow  
 charles joseph minard

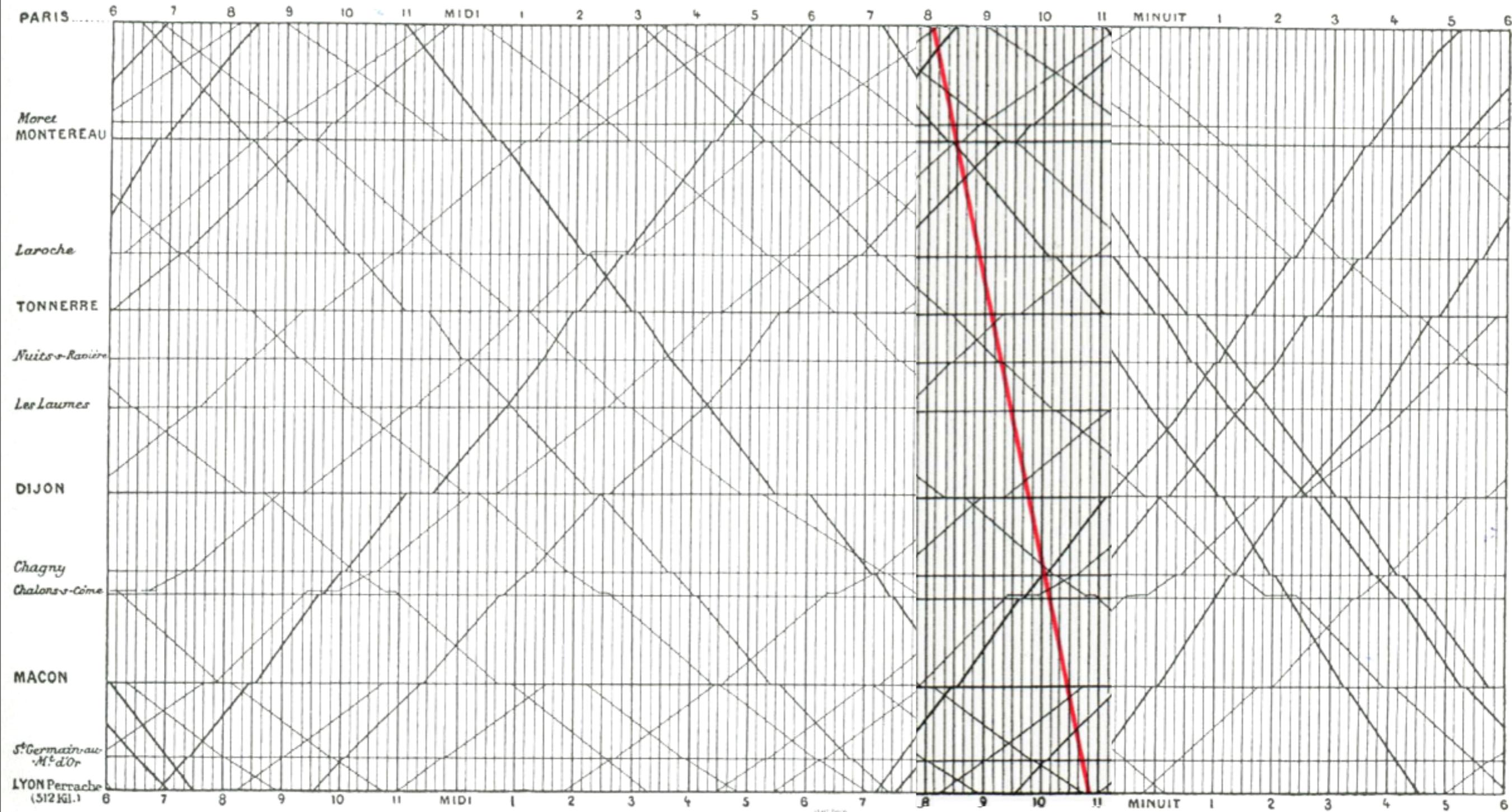


E.J. Marey  
La méthode graphique  
(1885)



E.J. Marey  
La méthode graphique  
(1885)

# TGV Paris-Lyon



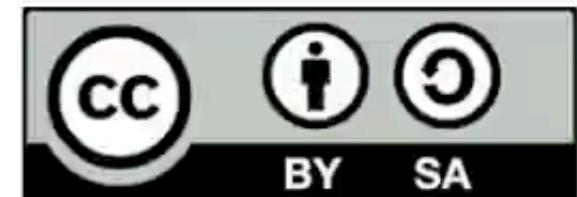
E.J. Marey  
La méthode graphique  
(1885)

time series (animation)

# **200 years that changed the world**

with Hans Rosling

Free to redistribute



[www.gapminder.org](http://www.gapminder.org)

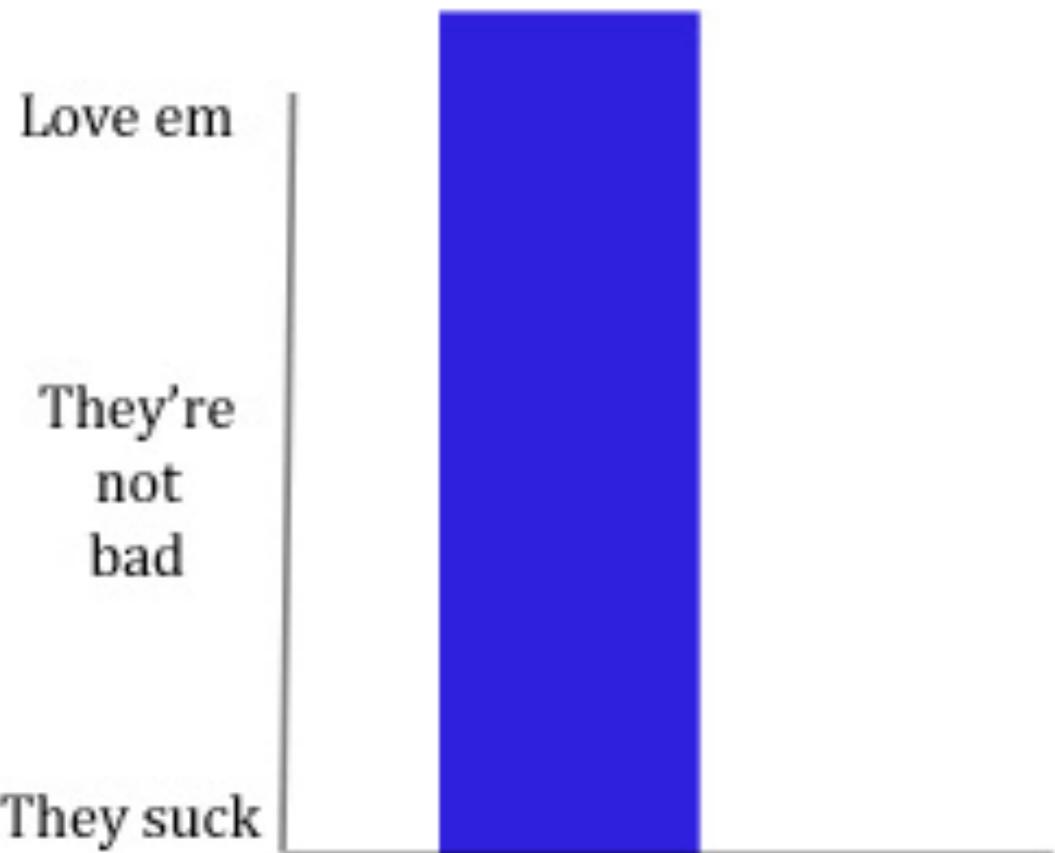
Plenty of other interesting visualisations....

Some favourites I didn't mention?

send them to: [max@hip.cat](mailto:max@hip.cat)

and I'll compile a list on a datavis wiki

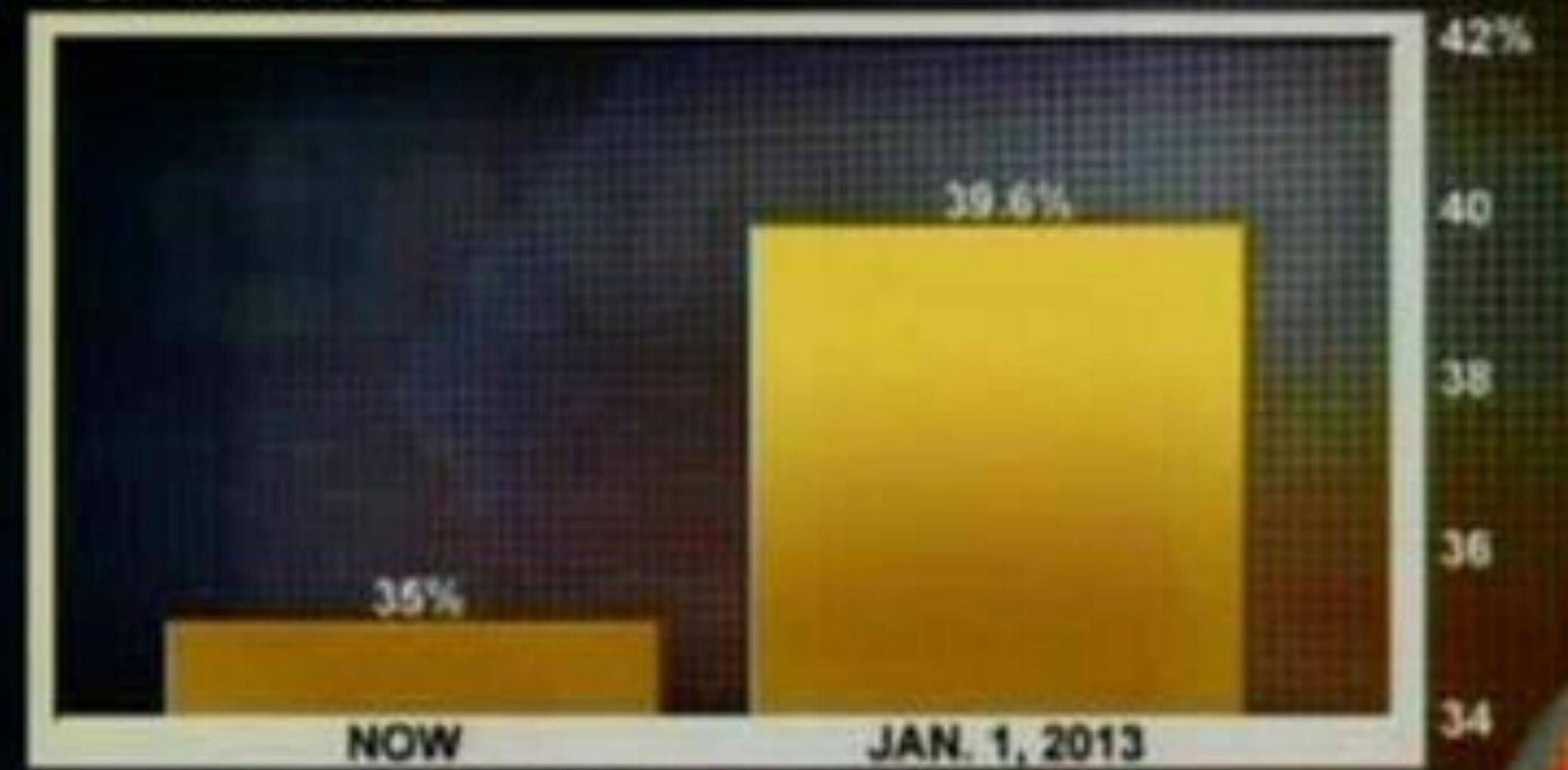
## How You Feel About Bar Charts



infographic fails:  
visual + statistical sleight  
of hand to mislead the  
audience

# IF BUSH TAX CUTS EXPIRE

TOP TAX RATE



8:01 p ET

**FOX**  
BUSINESS

TOP STORIES

TECHNOLOGY

CONSUMER

WITH THE JUSTICE DEPARTMENT AND ACQUIRES FULL T

DOW 13008.68

64.33

S&P

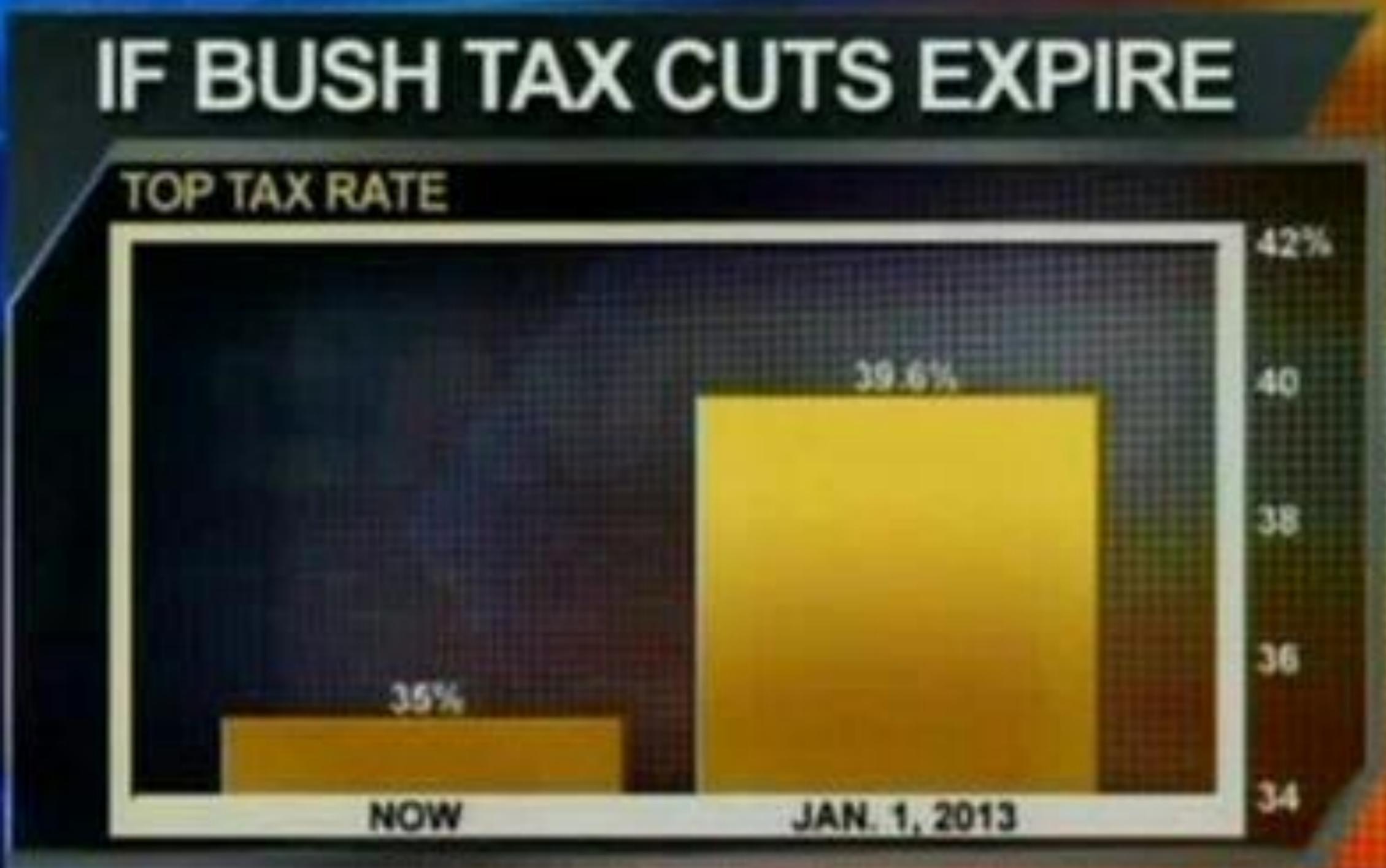
1379.32

5.98

NASDAQ 2939.52

6.32

# 1. Barchart baseline fail



8:01 p ET

**FOX**  
BUSINESS

TOP STORIES

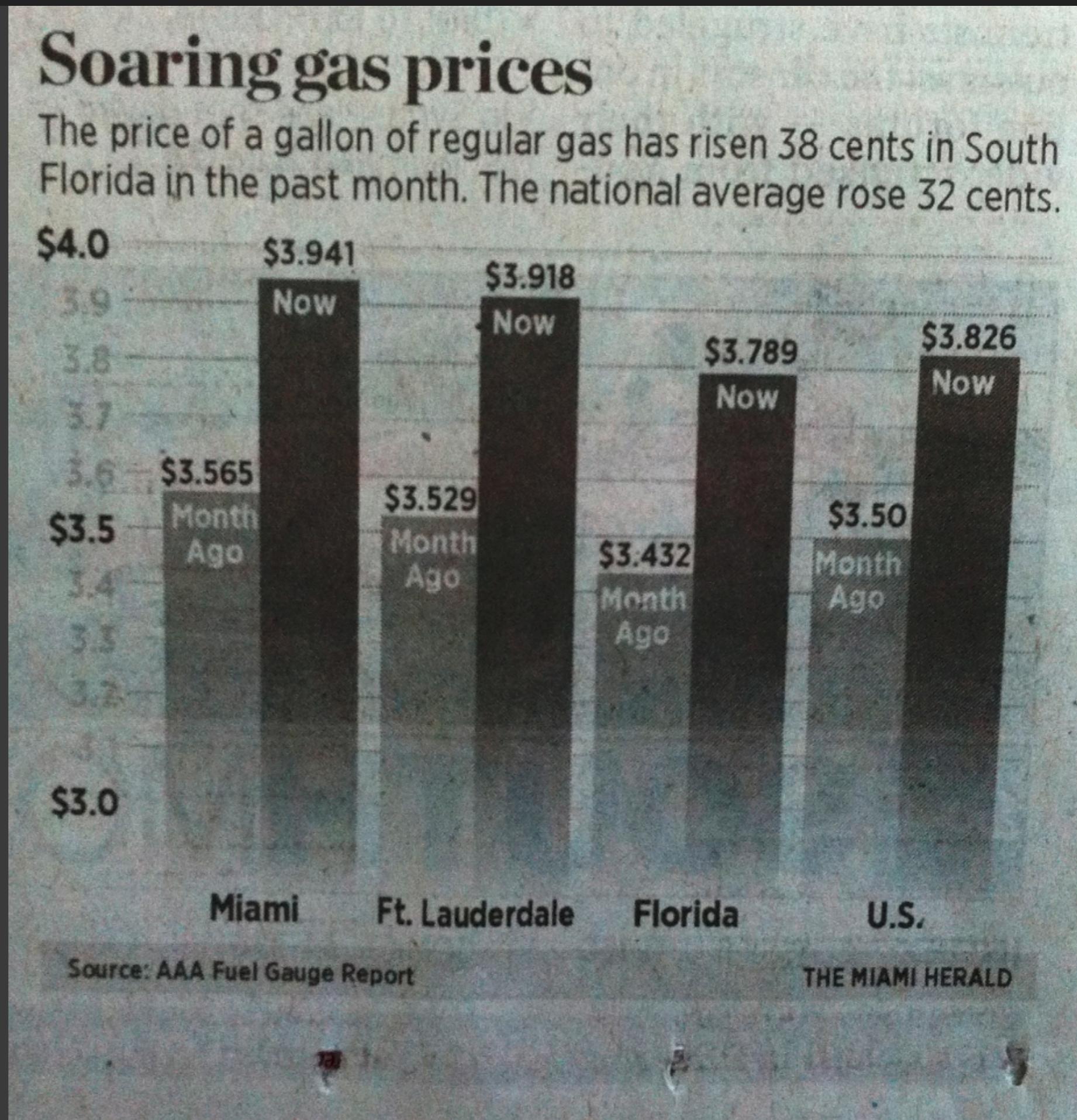
TECHNOLOGY

CONSUMER

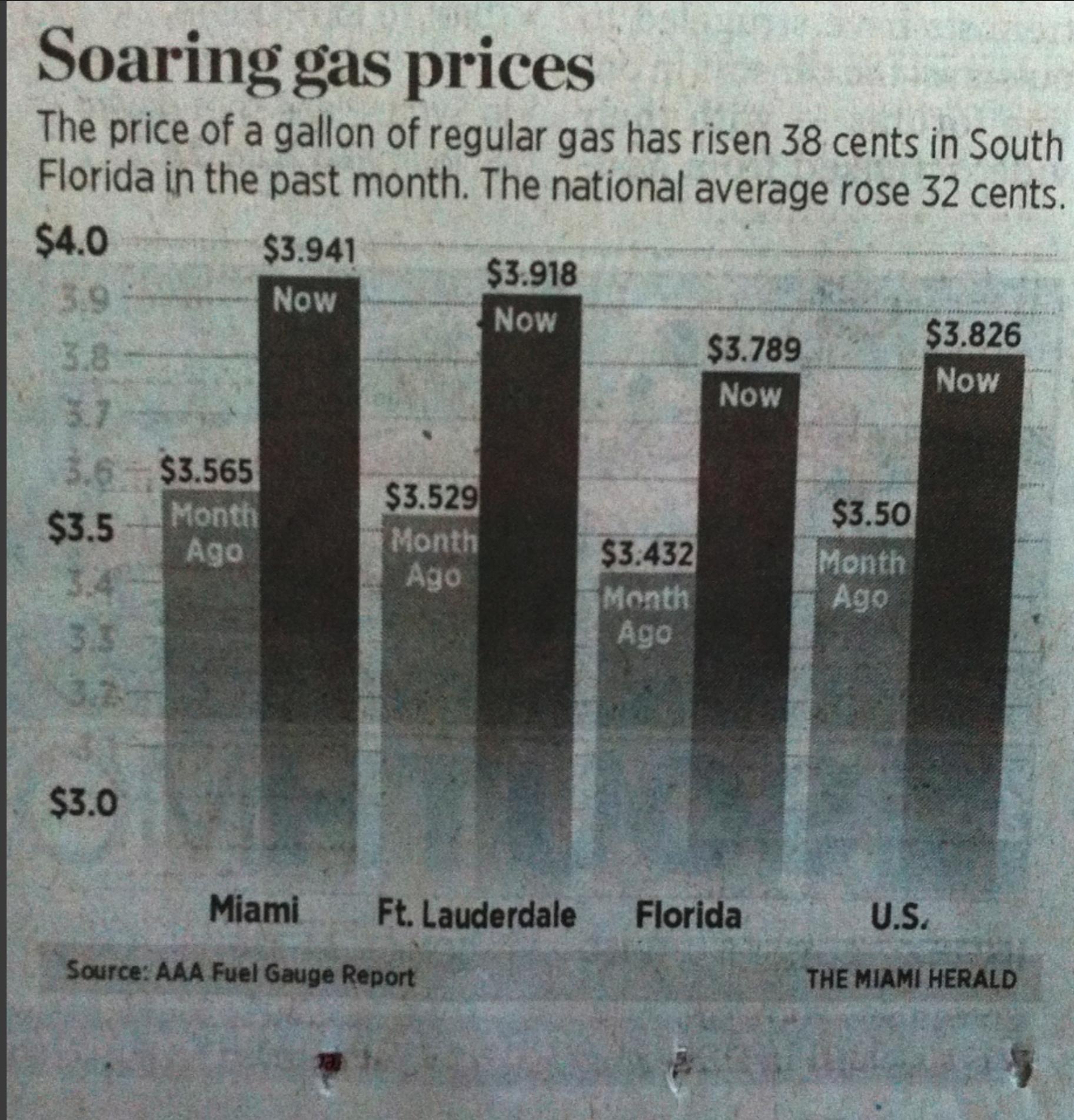
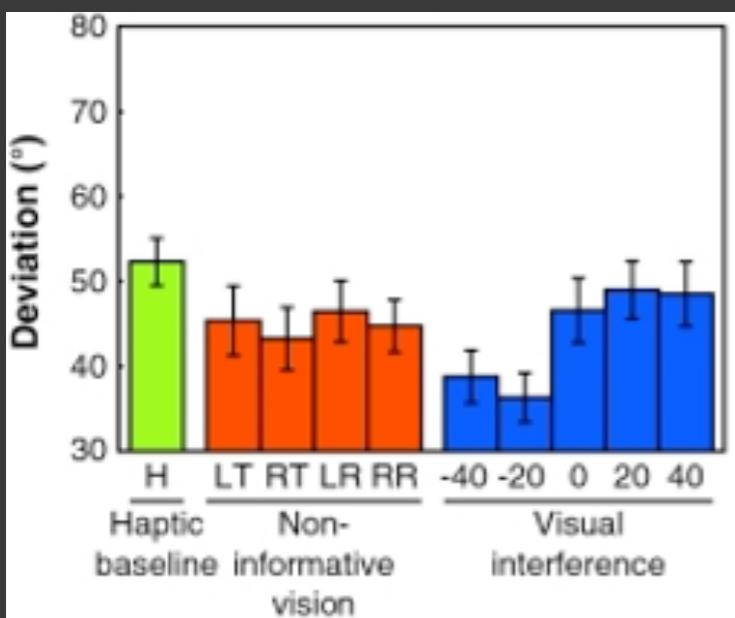
WITH THE JUSTICE DEPARTMENT AND ACQUIRES FULL T

DOW 13008.68 ▲ 64.33 S&P 1379.32 ▲ 5.98 NASDAQ 2939.52 ▲ 6.32

# 1. Barchart baseline fail



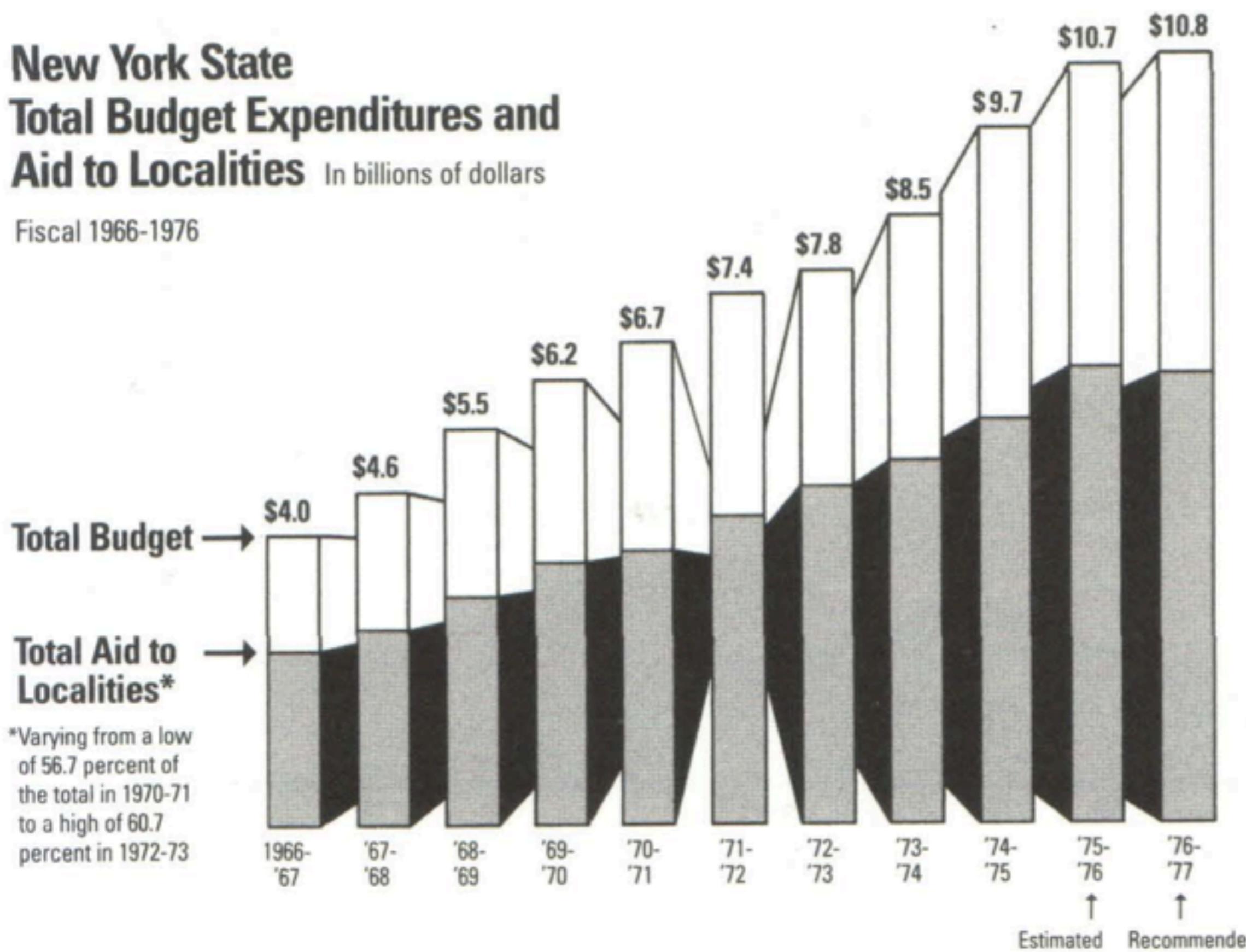
# 1. Barchart baseline fail



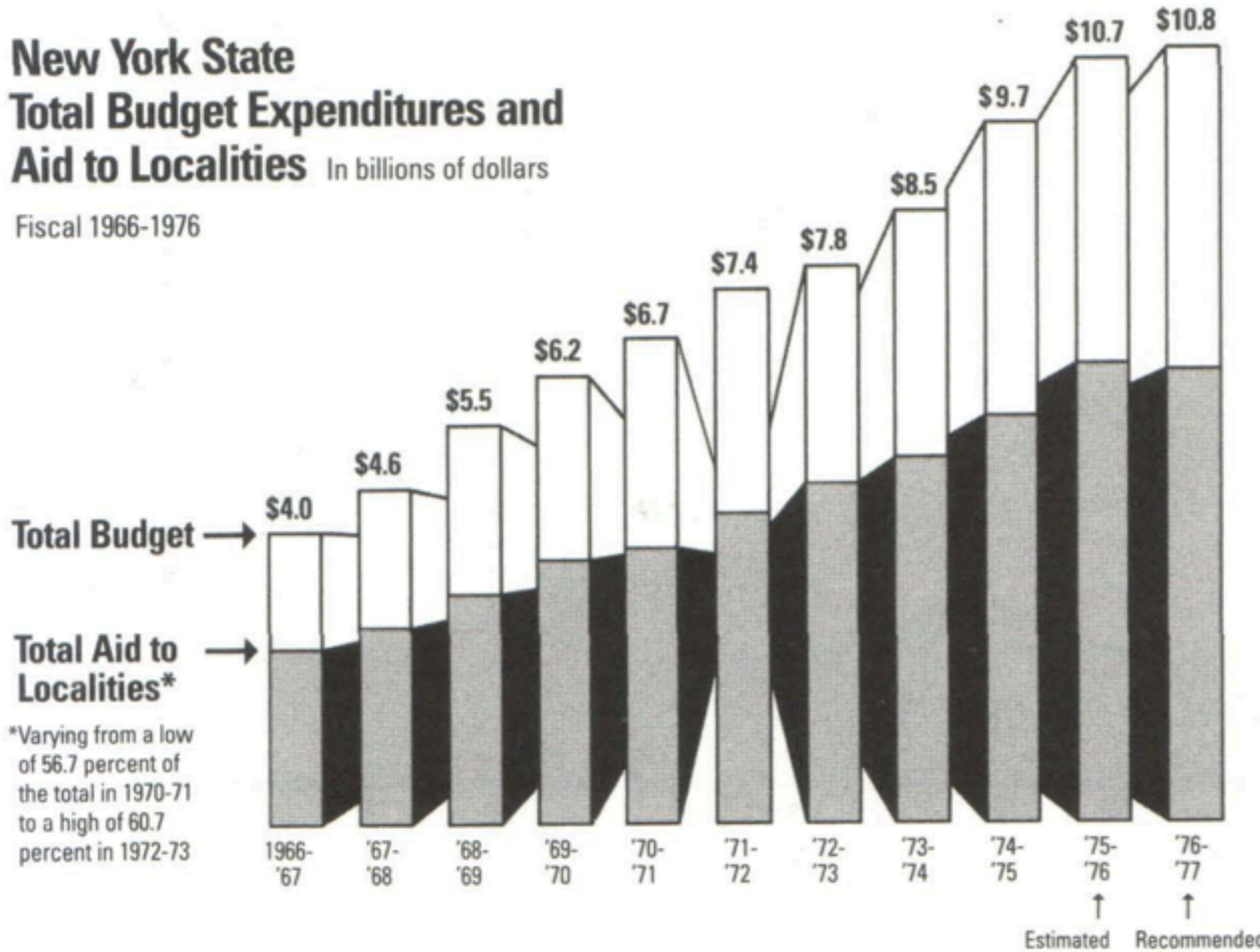
# New York State Total Budget Expenditures and Aid to Localities

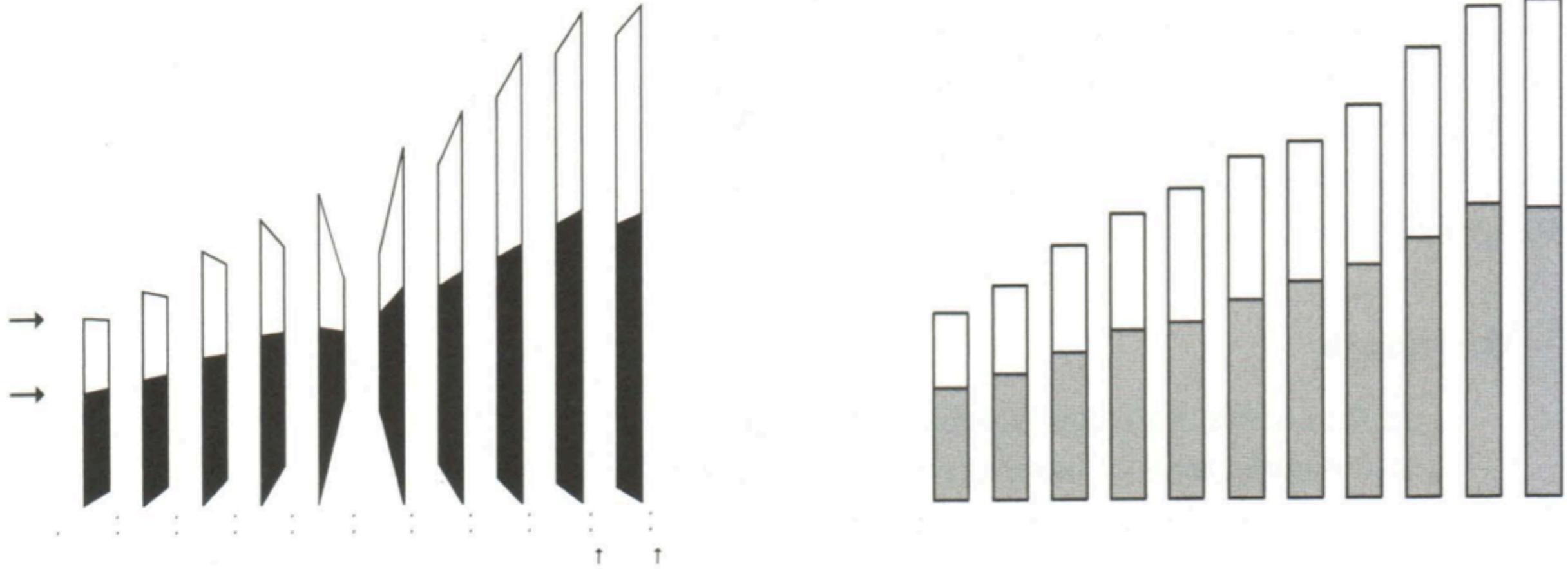
In billions of dollars

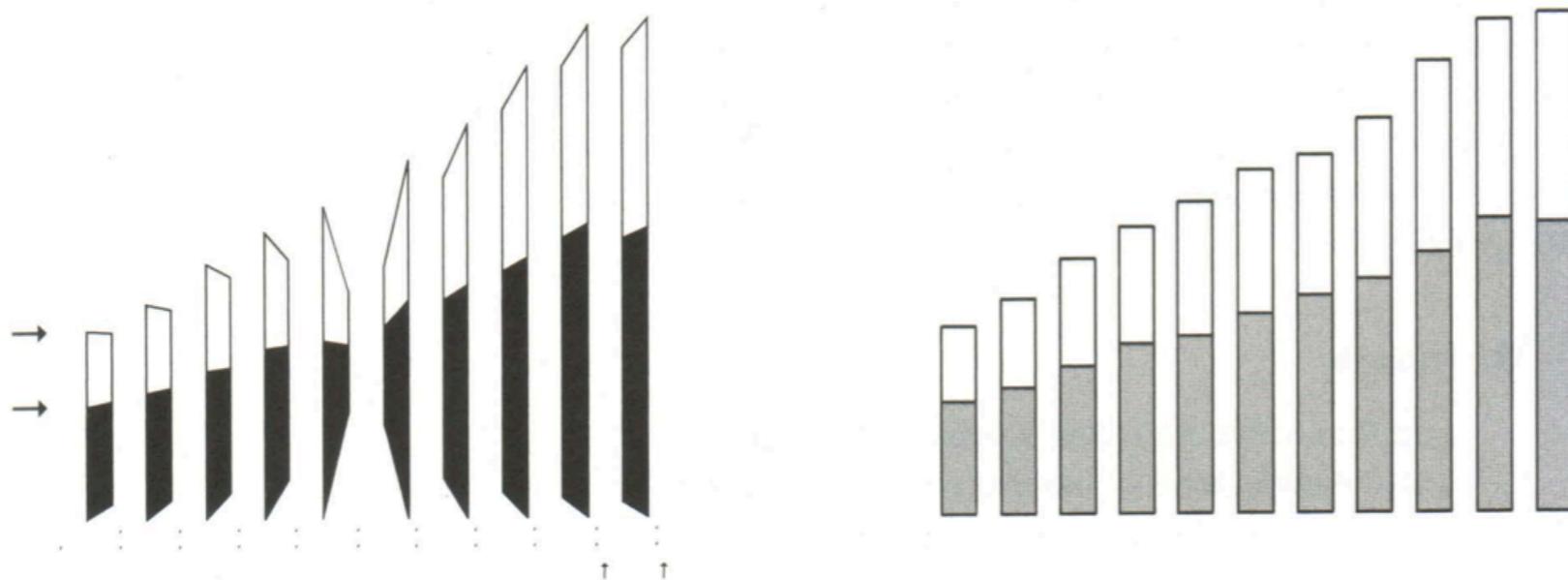
Fiscal 1966-1976



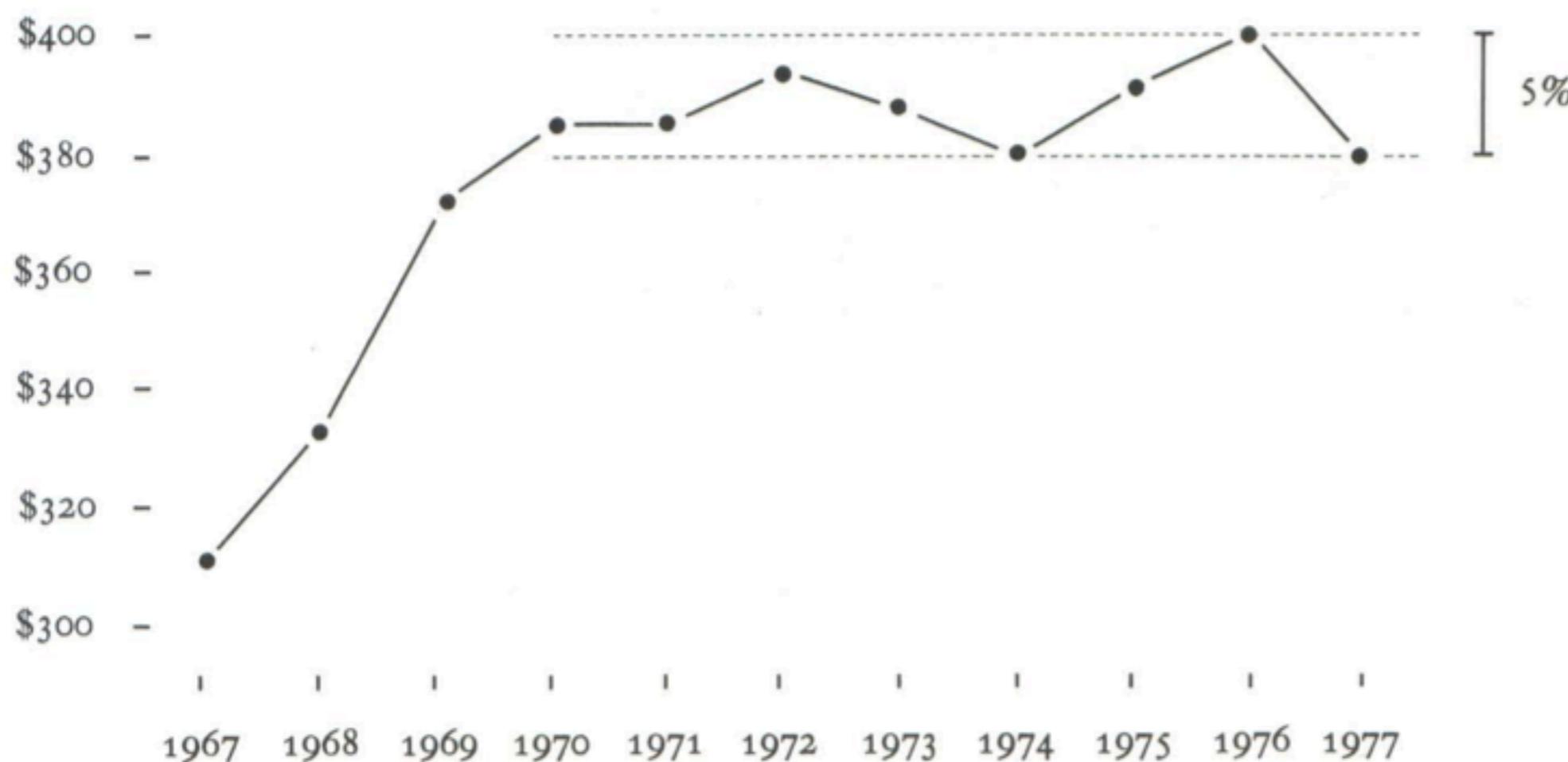
## 2. Perspective and measurement fail







Per capita  
budget expenditures,  
in constant dollars

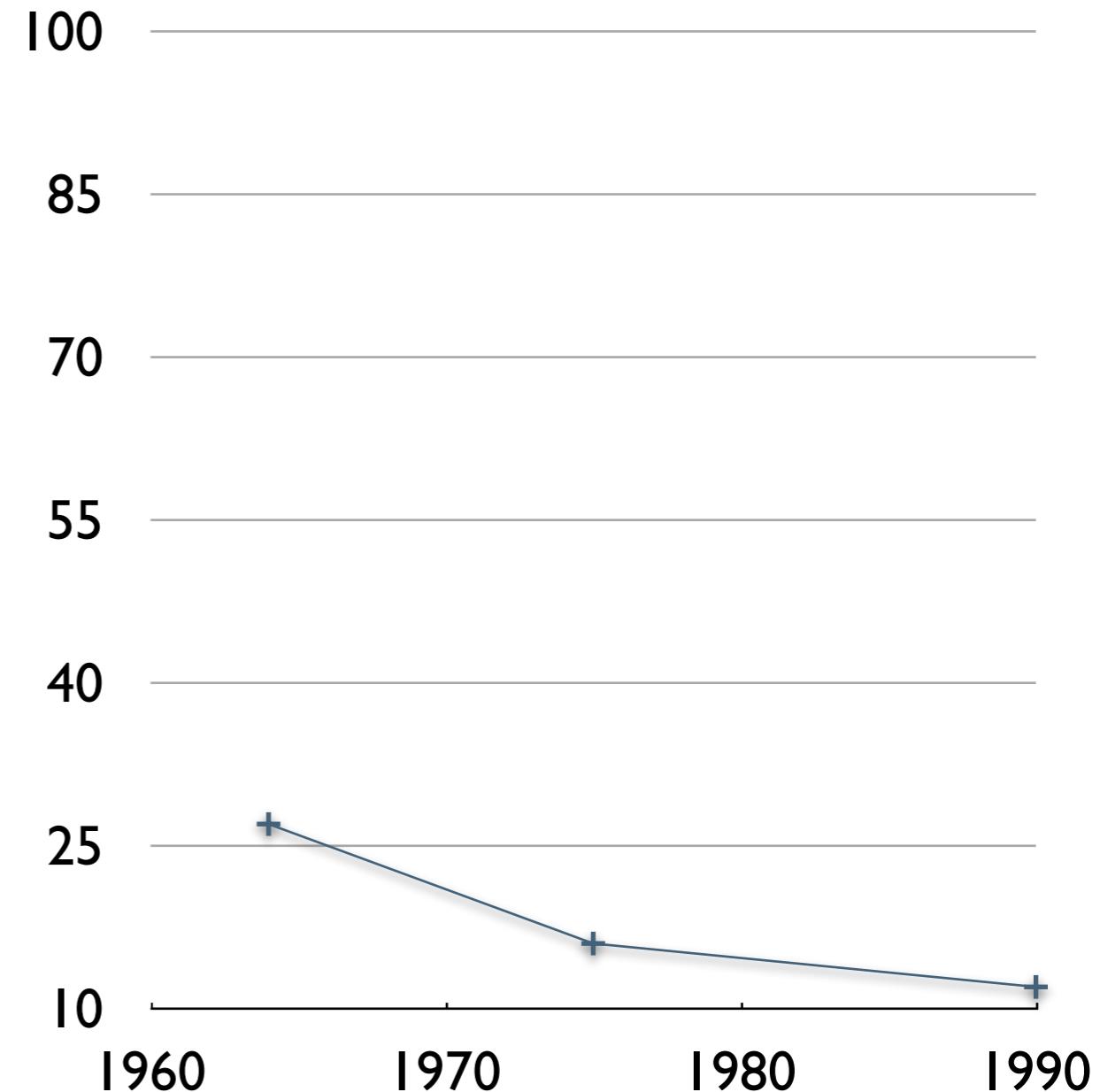
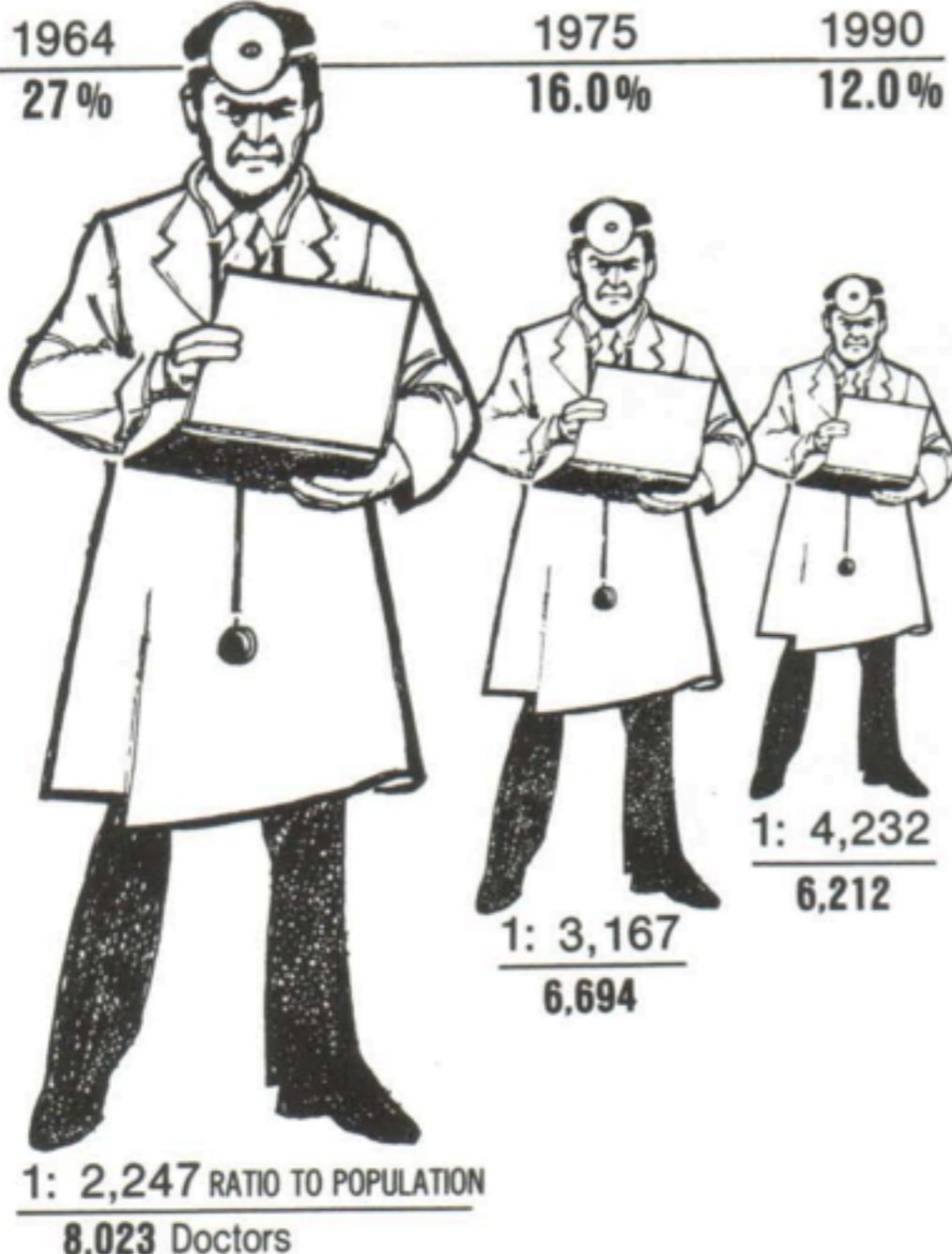


## 2. “Huge differences” fail

### THE SHRINKING FAMILY DOCTOR In California

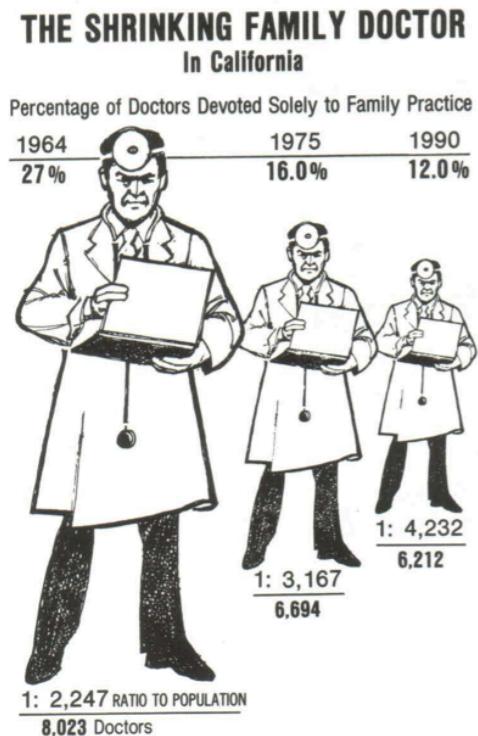
Percentage of Doctors Devoted Solely to Family Practice

1964	1975	1990
27%	16.0%	12.0%



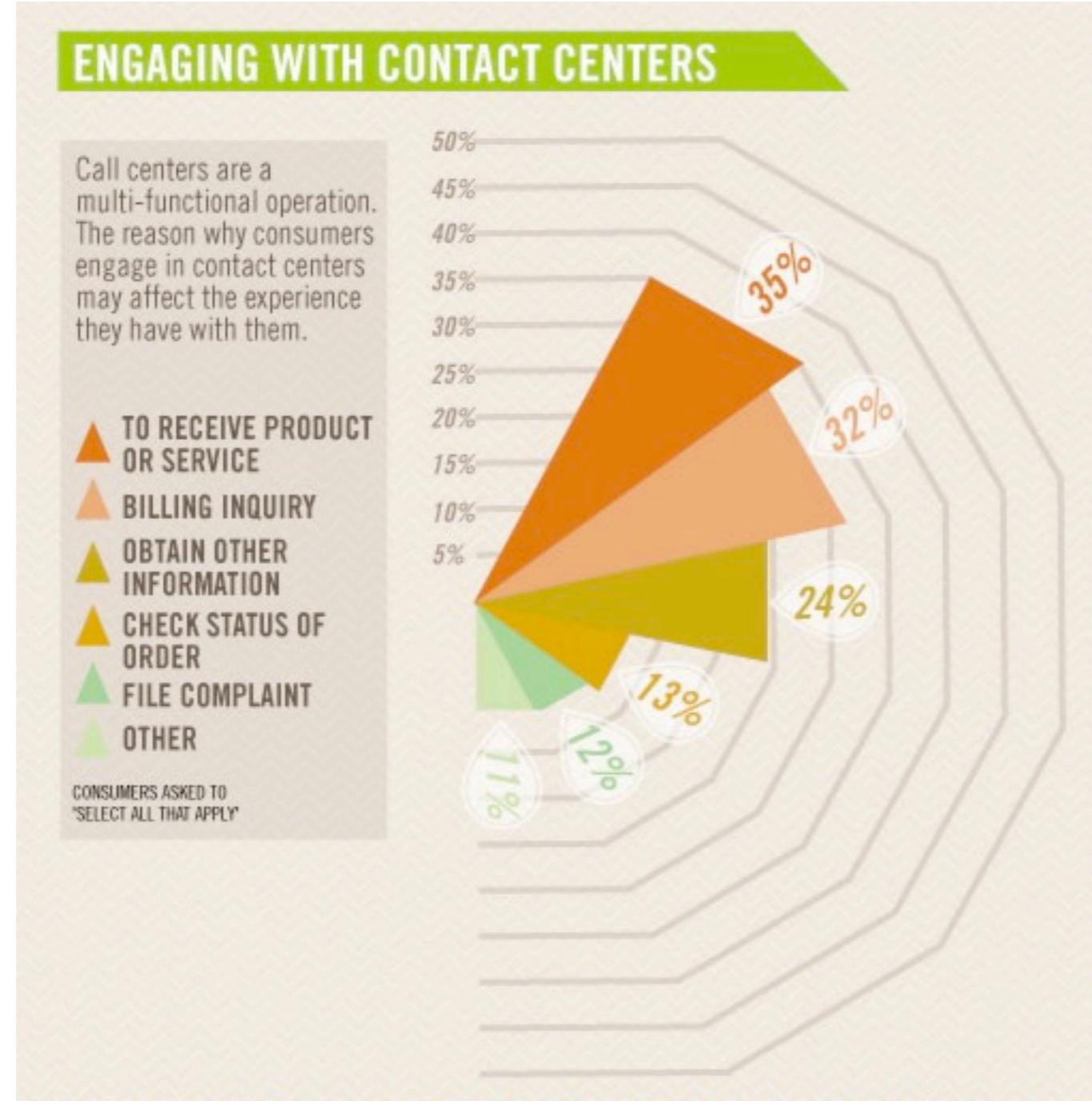
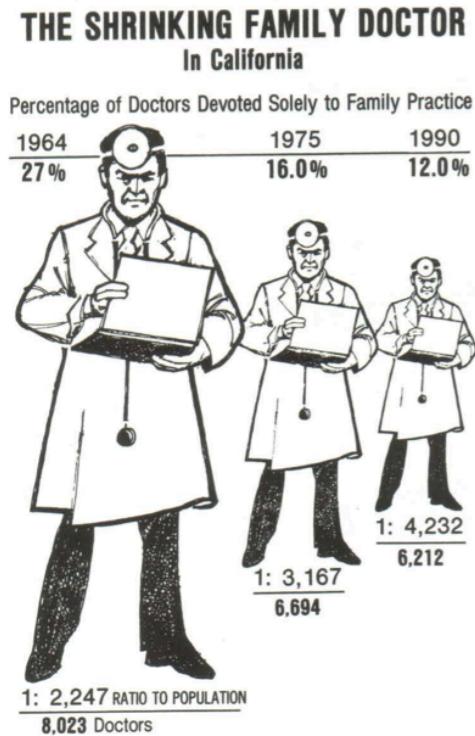
using area (2 dimensions) to represent one dimension

## 2. “Huge differences” fail



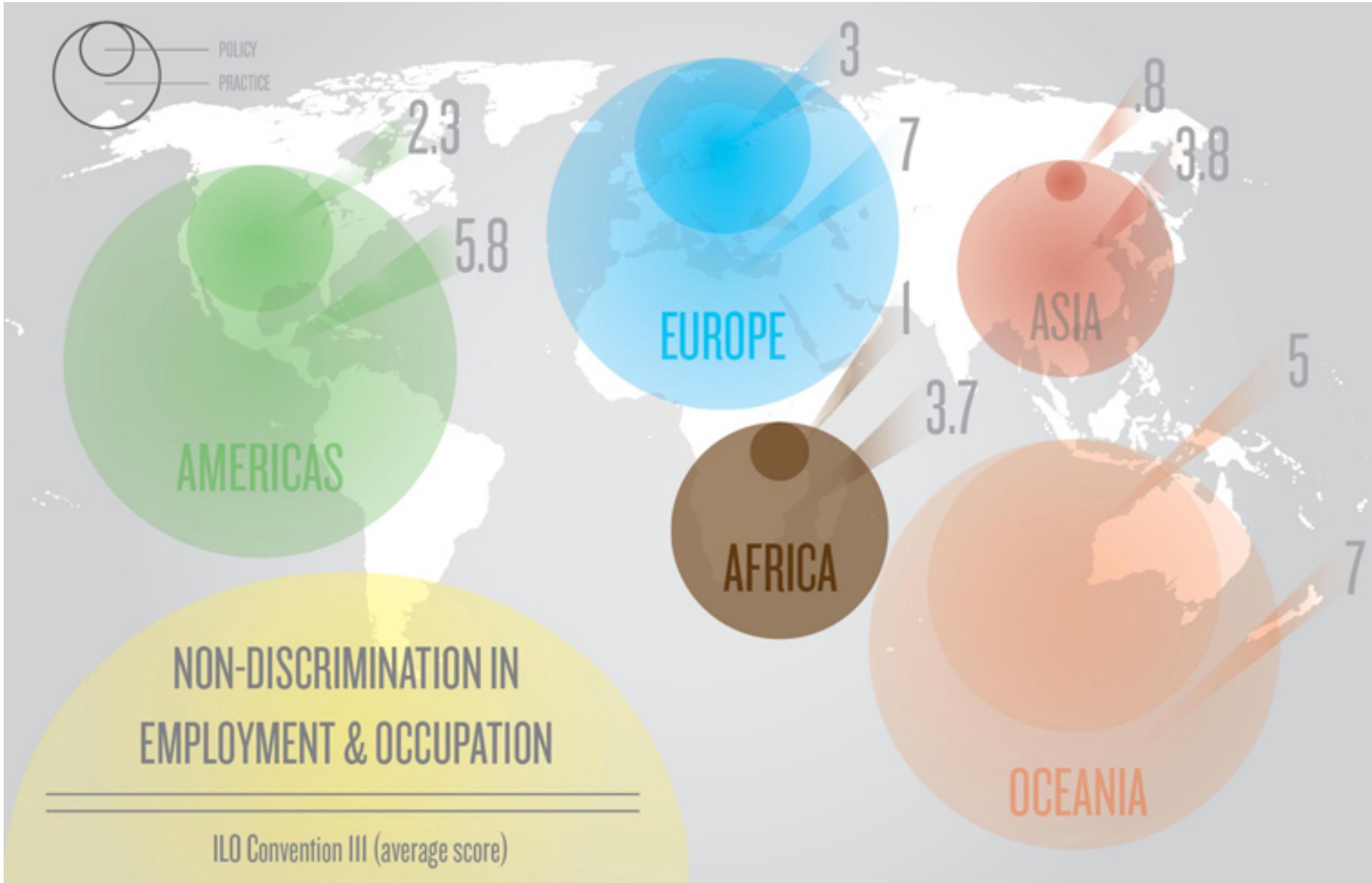
using area to represent one dimension

## 2. “Huge differences” fail



using area to represent one dimension

## 2. “Huge differences” fail



using area to represent one dimension

# Quiz: How does this fail?

## THE ISSUE OF TRUST

### ACCENTS AND DISTRUST

Another reason why accents affects customer service is the question of credibility. If I can not understand you, then I can not trust you.

An experiment conducted by the University of Chicago demonstrated this aspect. The question posed, do trivia statements sound less true when spoken by a non-native speaker? Furthermore, listeners were told in advance that all of the trivia questions were provided by the experimenter. This way, even listeners who were knowingly prejudice against non-native accents should not have been affected.

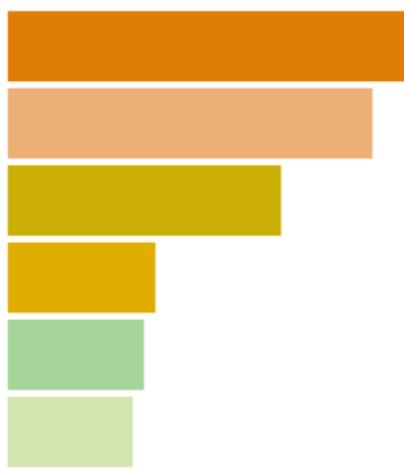
The results showed that the heavier the accent the less trust worthy the person became.

- ▲ NATIVE ACCENT
- ▲ MILD ACCENT
- ▲ HEAVY ACCENT





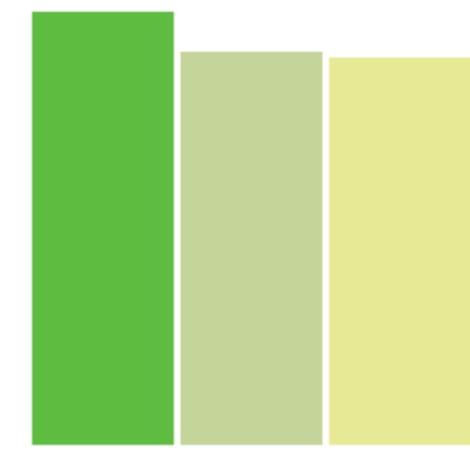
Distortion factor: 2.5



True data



Distortion factor: 5.0



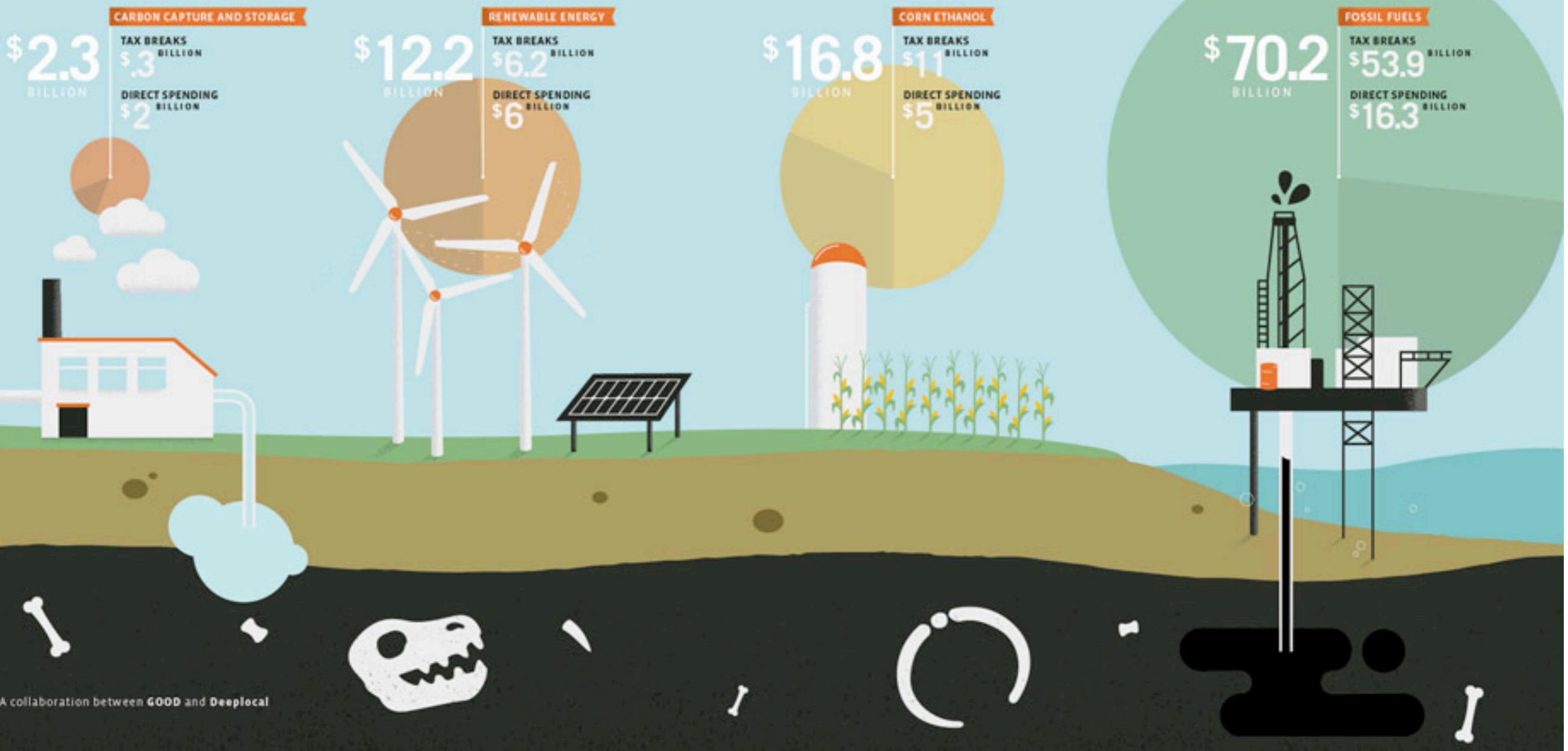
True data

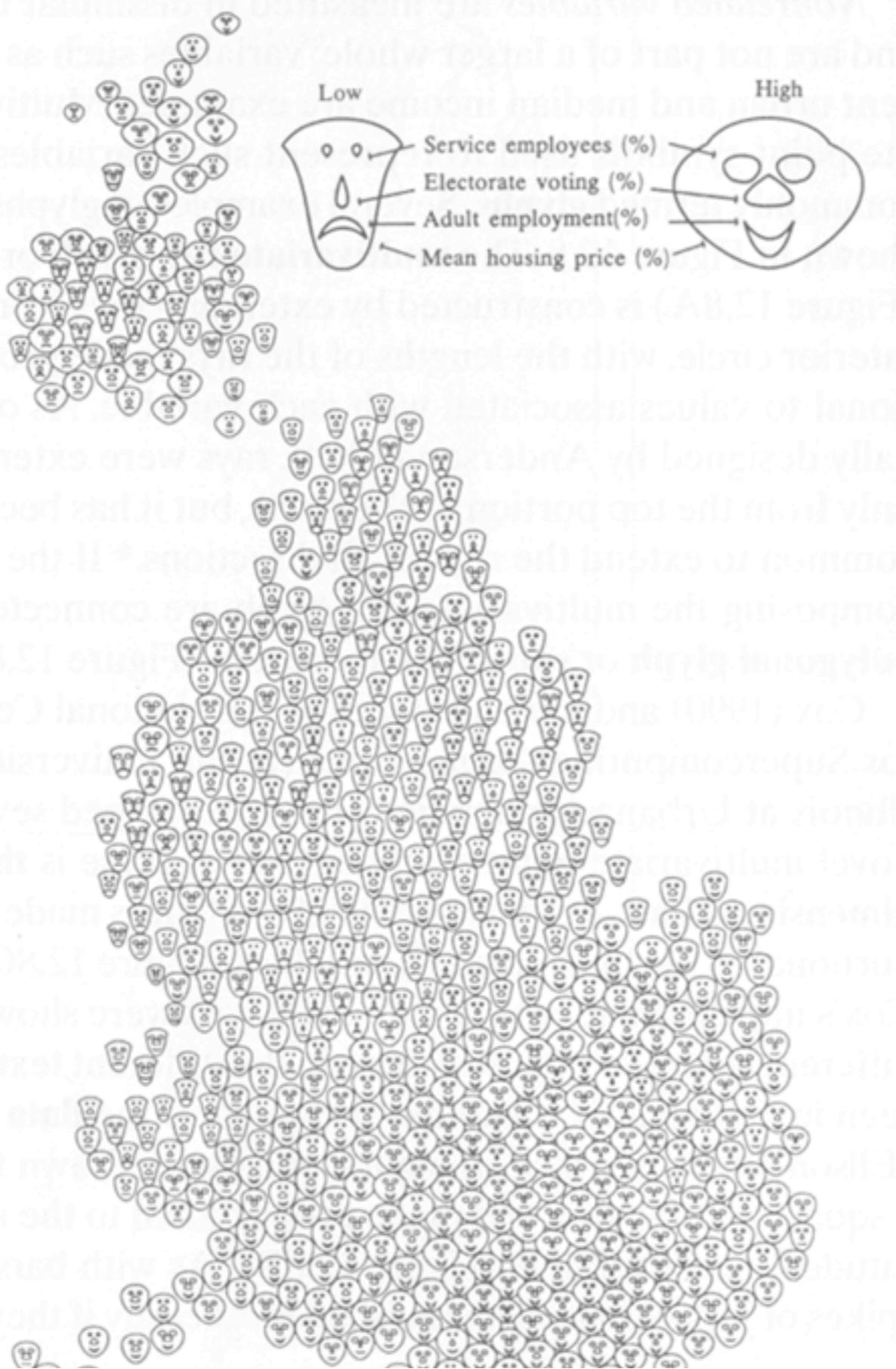
# SUBSIDIZE THIS

THE PRICE THAT YOU PAY FOR ENERGY—WHETHER ELECTRICITY AT YOUR HOUSE OR GAS AT THE PUMP—isn't actually the price that the market would set for that energy.

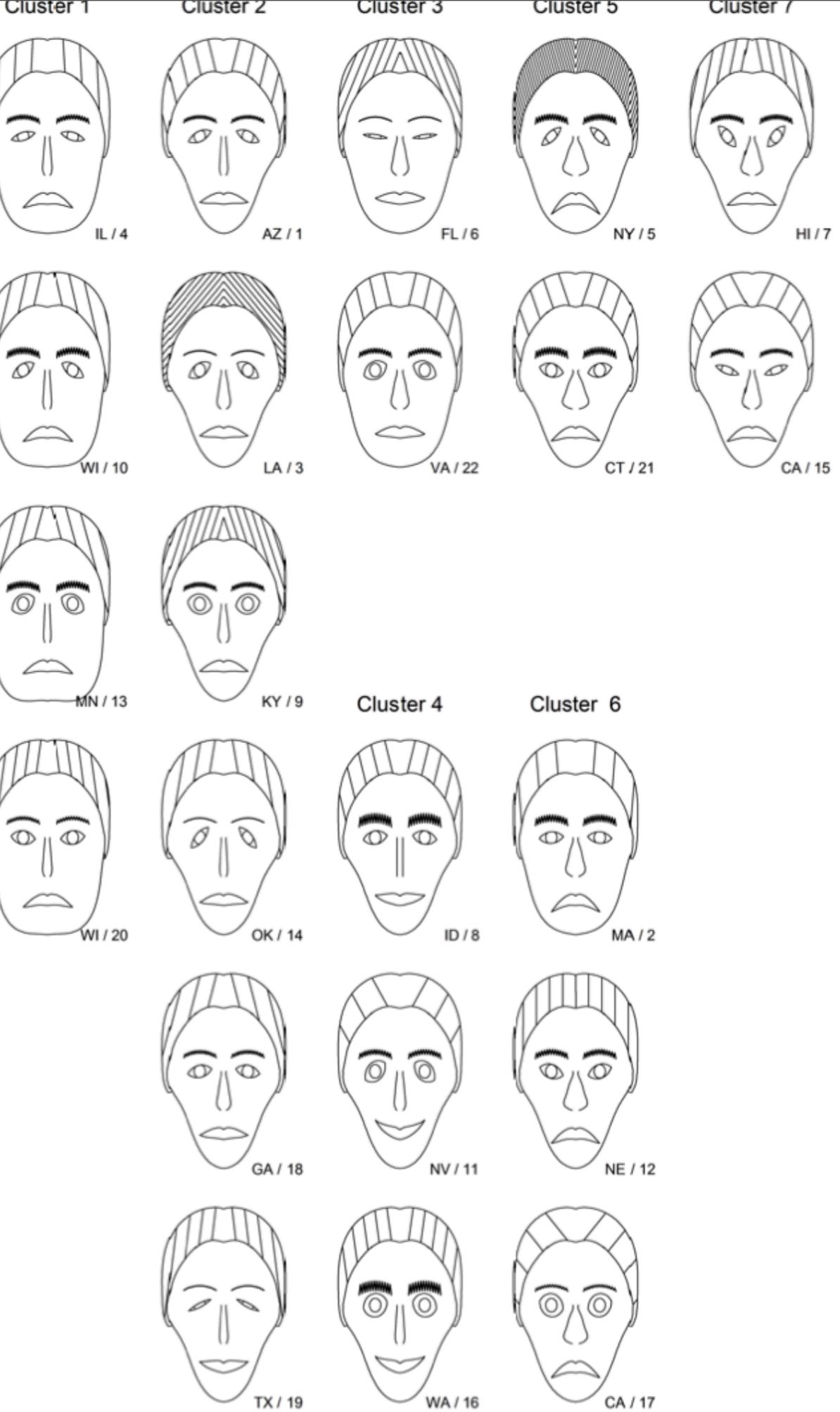
The government spends billions of dollars to support the energy industry, which allows it to make energy cheaper than it should cost on the open market. These subsidies—either in the form of tax breaks or direct funding—favor some types of energy over others, giving our country a skewed sense of what each gallon of gas or wind-powered electron costs. This is a look at where the government directed its subsidy dollars from 2002 to 2008.

SOURCE "Estimating U.S. Government Subsidies to Energy Sources" by the Environmental Law Institute



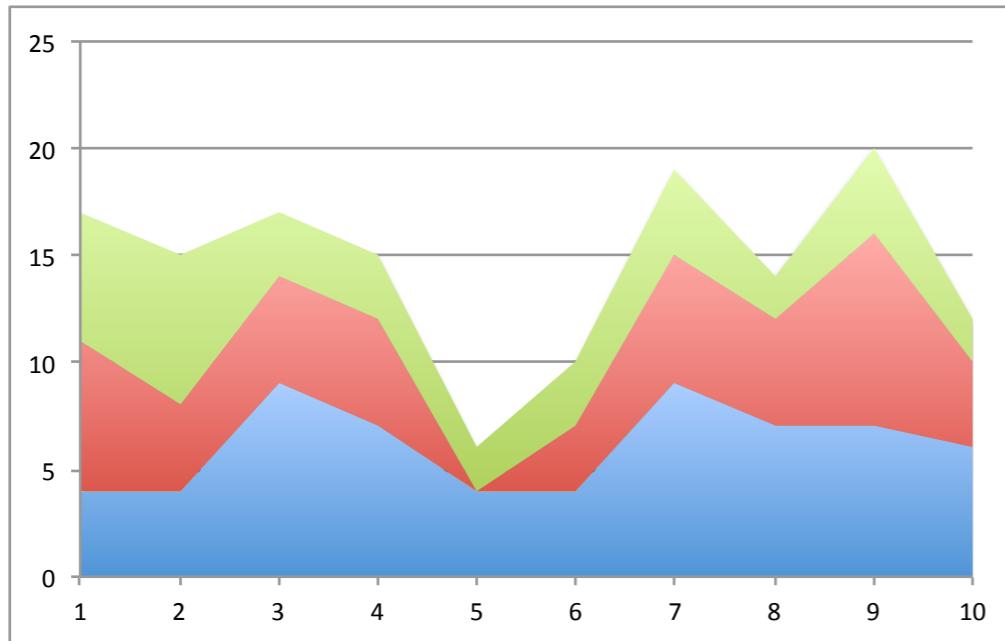


# Chernoff Faces

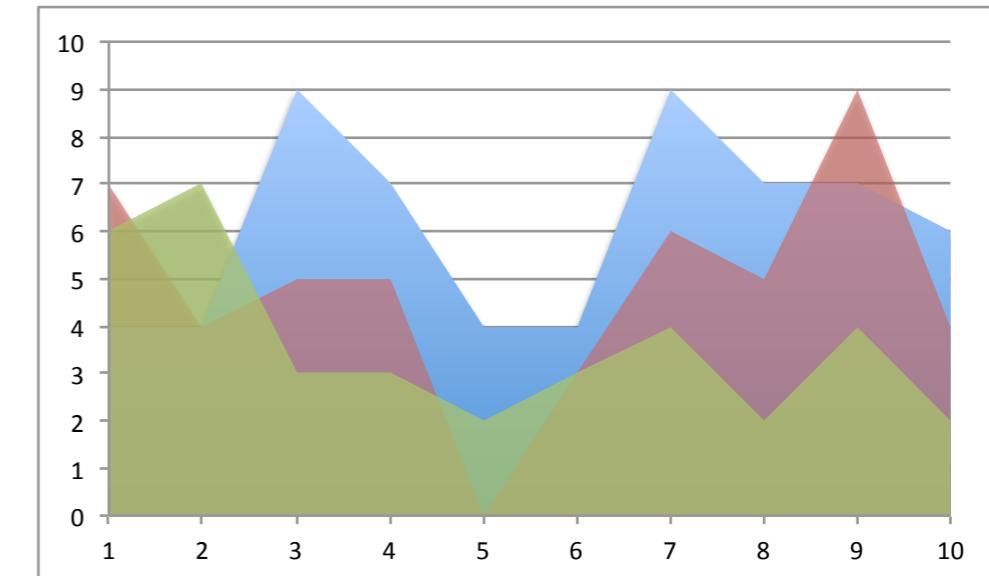
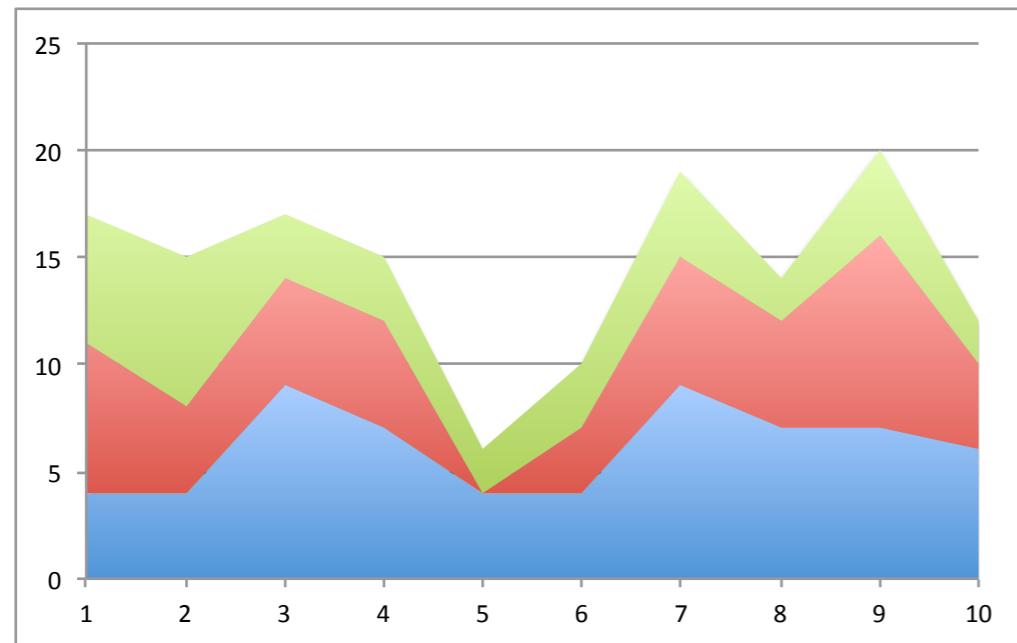


“streamgraphs”: double-stacked areas of horror

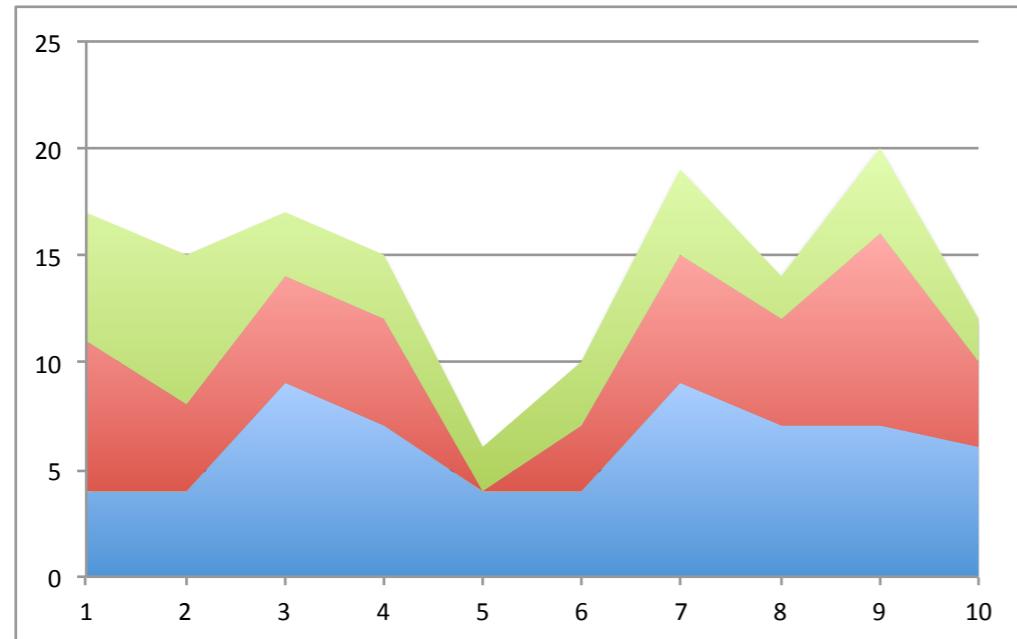
# “streamgraphs”: double-stacked areas of horror



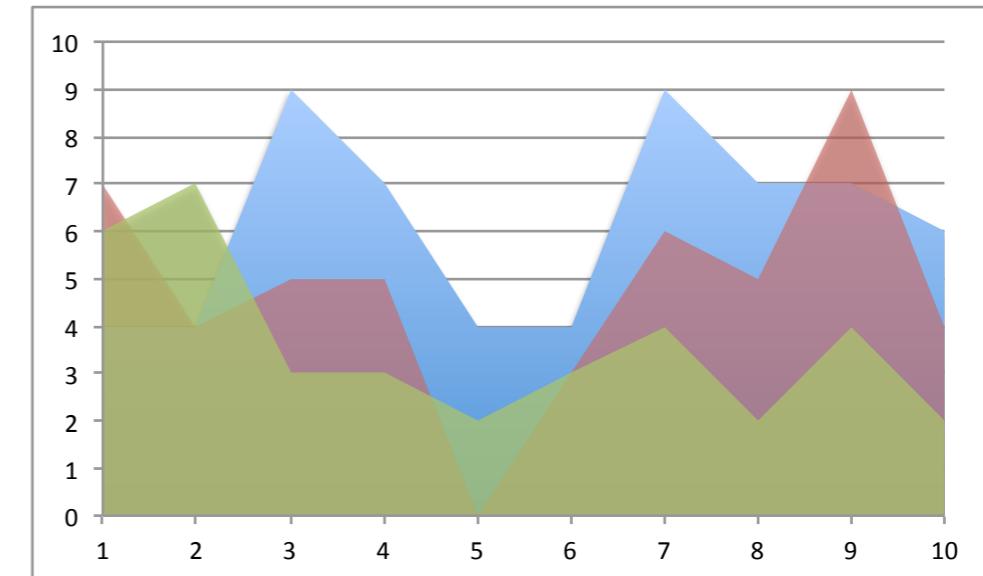
# “streamgraphs”: double-stacked areas of horror



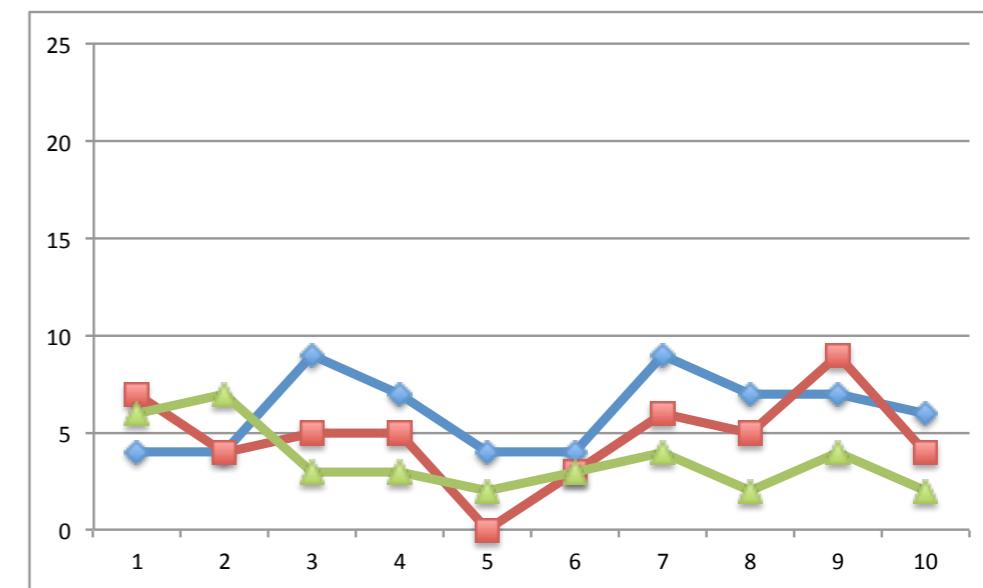
# “streamgraphs”: double-stacked areas of horror



?

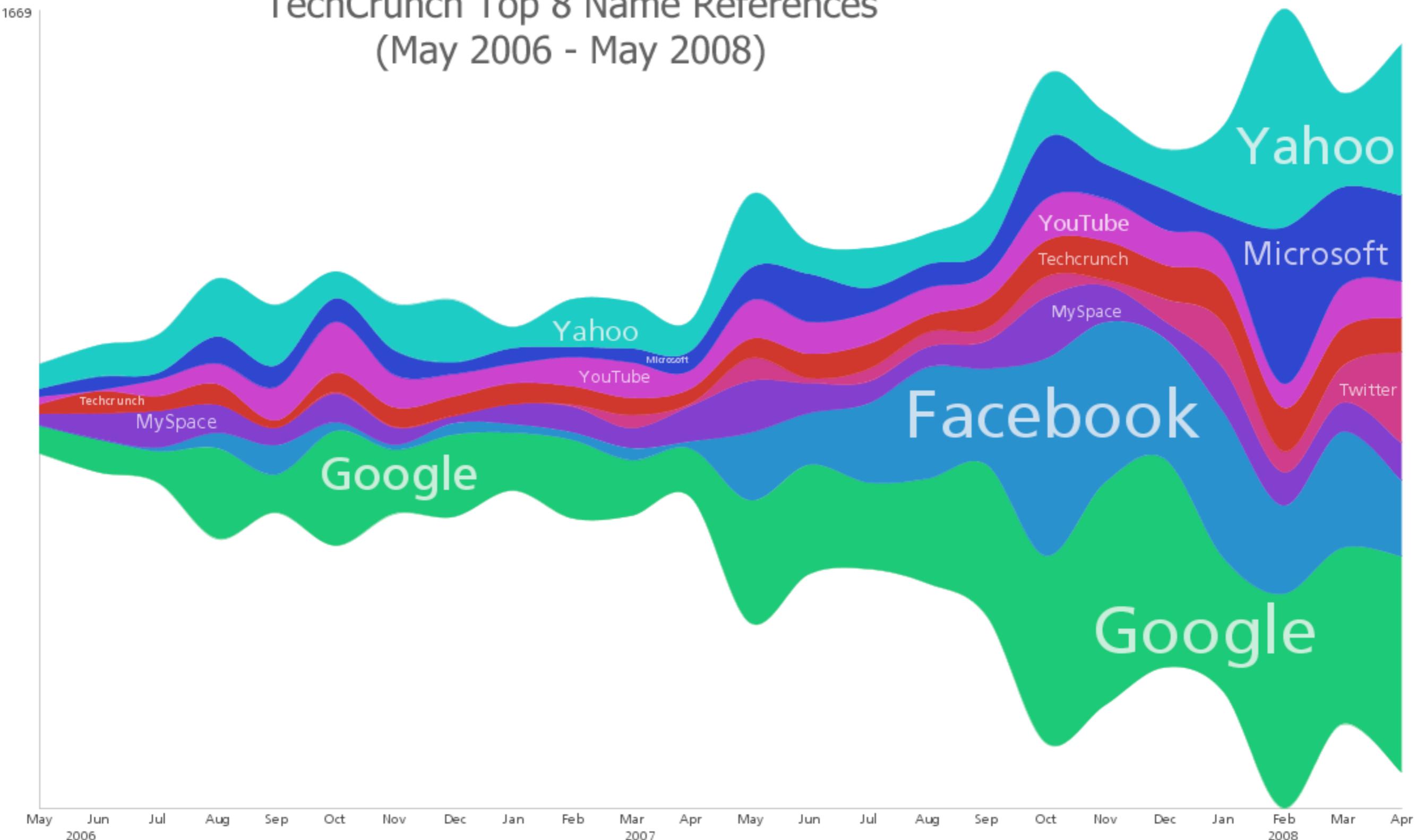


↓



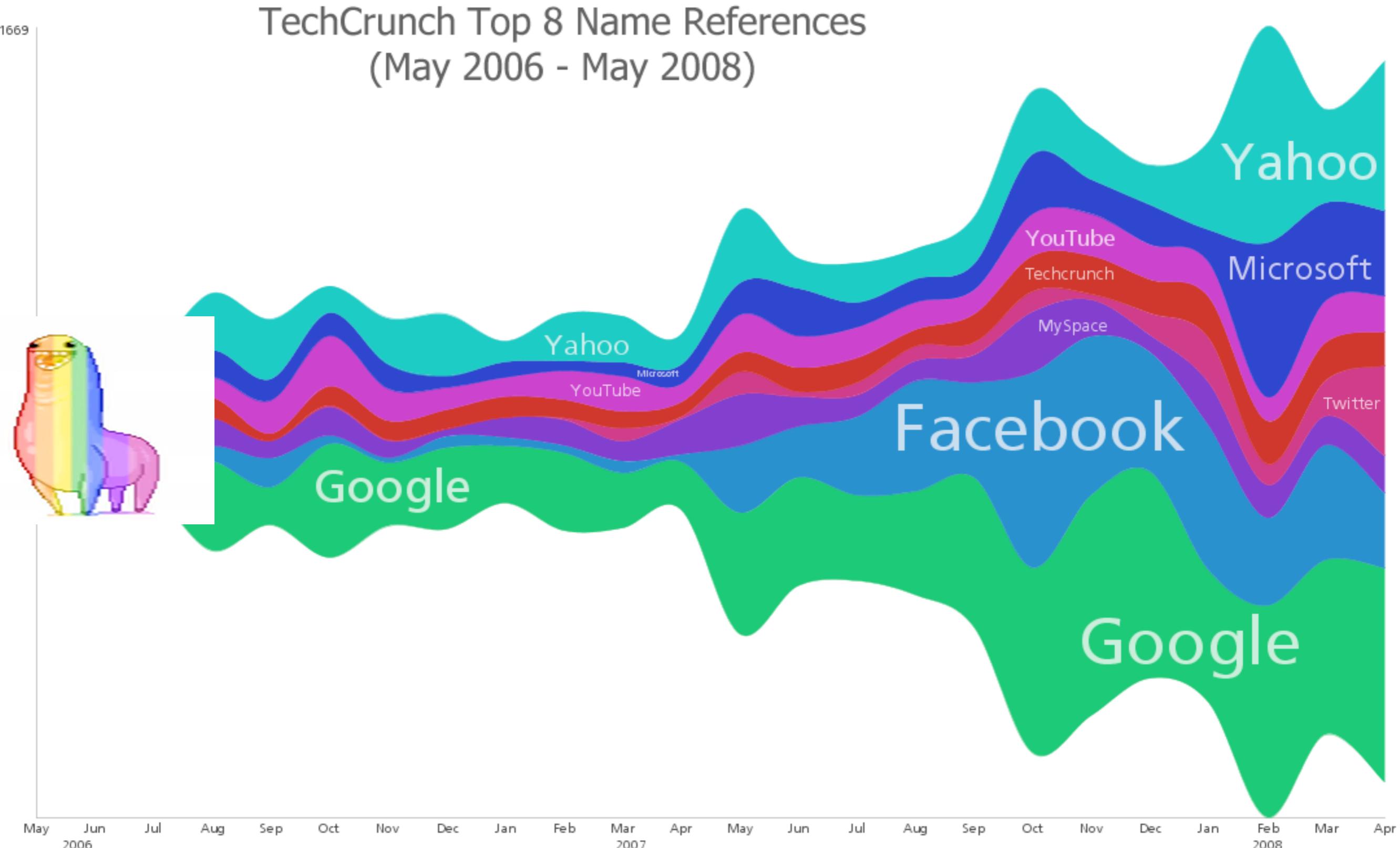
“streamgraphs”: double-stacked areas of horror

TechCrunch Top 8 Name References  
(May 2006 - May 2008)

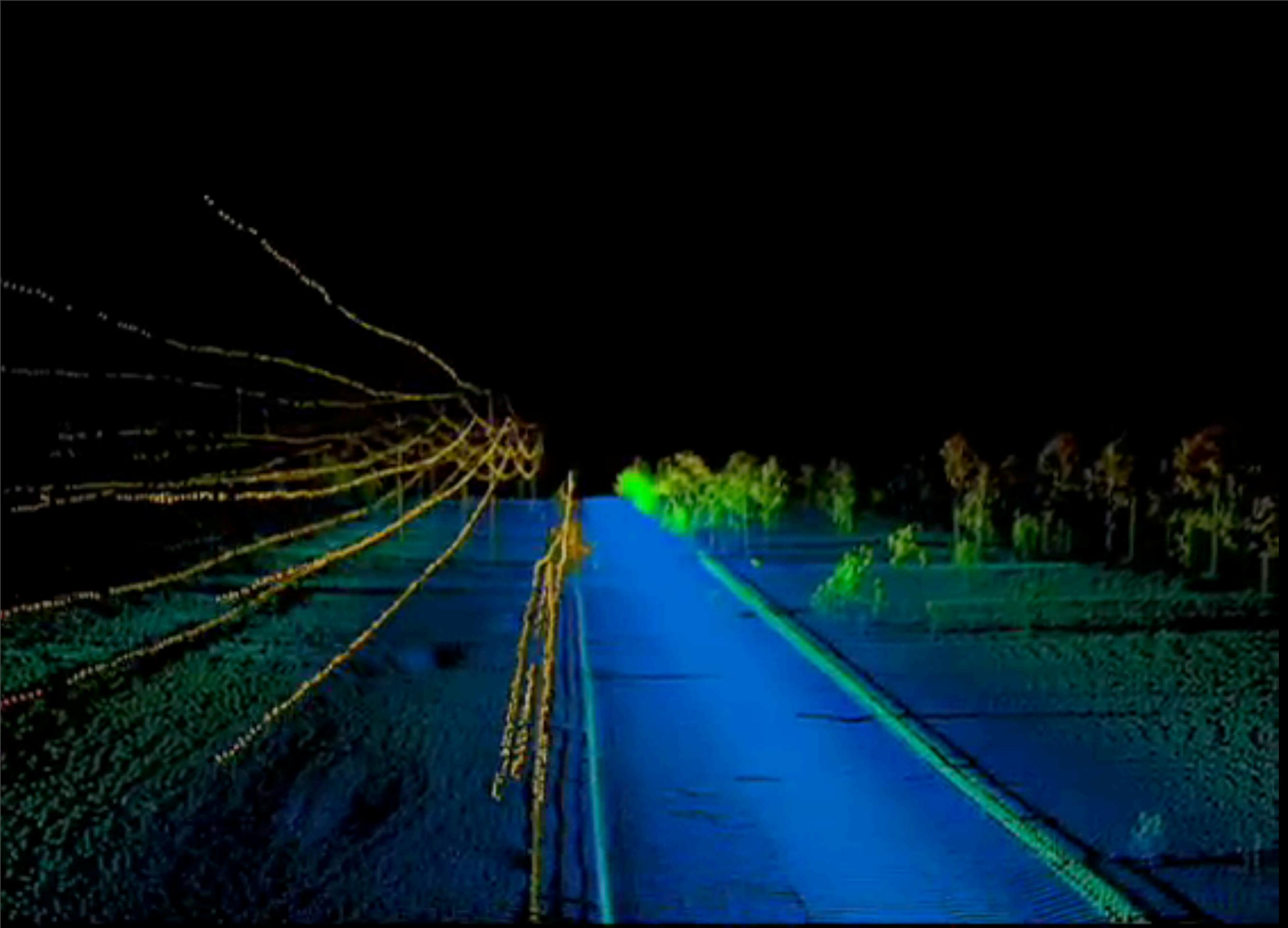


“abandon all hope ye who vieweth”

“streamgraphs”: double-stacked areas of horror



“abandon all hope ye who vieweth”



# In conclusion

Designing effective infographics

is about effectively conveying or facilitating an understanding of relationships in data

offloading “heavy lifting” to our trained neural circuitry.

While still an art, many design principles grounded in usability can provide guidance: natural mappings, simplicity, & avoiding distortion

