DATA VISUALIZATION

Graphs


Pie I have eaten
Pie I have not yet eaten

WHAT IS A GRAPH, REALLY?

1. **graph**

A slang term for a sketching of what someone plans on tagging.
graffiti.

boy 1: Hey check out this graph i drew in science class.

boy 2: It looks good call me when you want to hit up.

by Eamz February 08, 2007

27 18

2.

Graph

A word used in the north west around the area of Manchester and Liverpool

It means "bad" or something that is "annoying" as in "mate that was well graph" or " stop being so graph"

" that was so graph"
" stop being so graph"

by oli hughes November 07, 2012

6 4

3. **graph**

Graph- Mythical 1337 Creature. The peak of gaming excellelance.

see leet uber
Antonyms non-factor

That was SO Graph.
He killed that thing with Graph like Skills.

by BoBO1 June 10, 2005

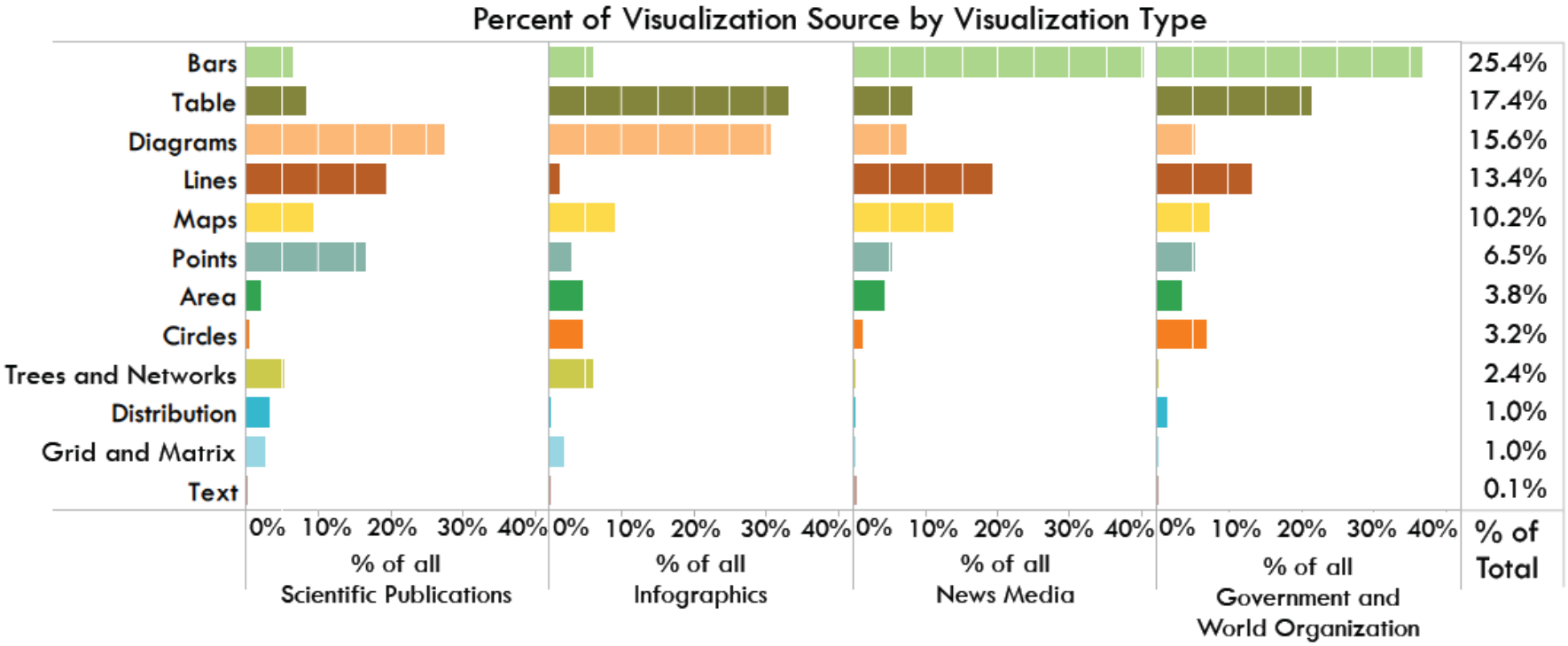
8 31

NOUN \'graf\

A DIAGRAM REPRESENTING A SYSTEM OF CONNECTIONS OR INTERRELATIONS AMONG TWO OR MORE THINGS BY A NUMBER OF DISTINCTIVE DOTS, LINES, BARS, ETC.

DICTIONARY.COM

! | TOTAL SIDEBAR HERE...

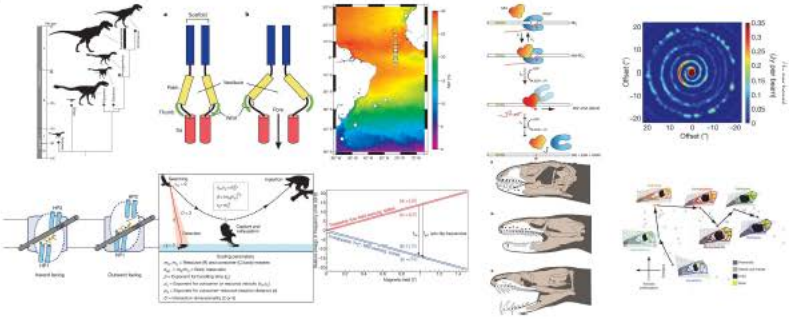


Borkin et al What makes a Visualization Memorable? 2013 IEEE

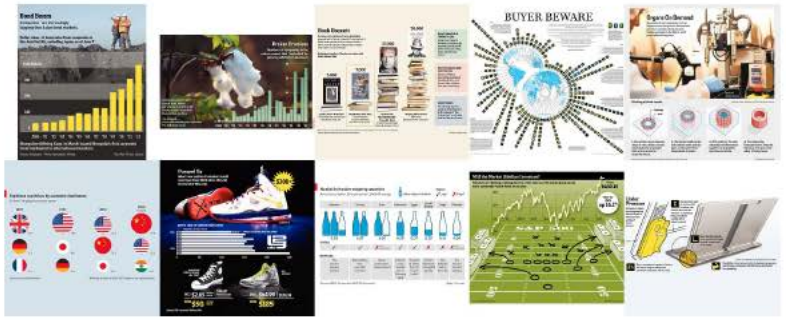
Memorability?

#1 Infographic

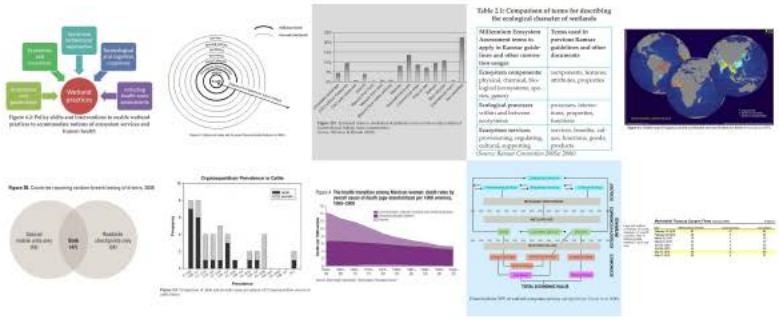
#2 Scientific Journals



#3 News media

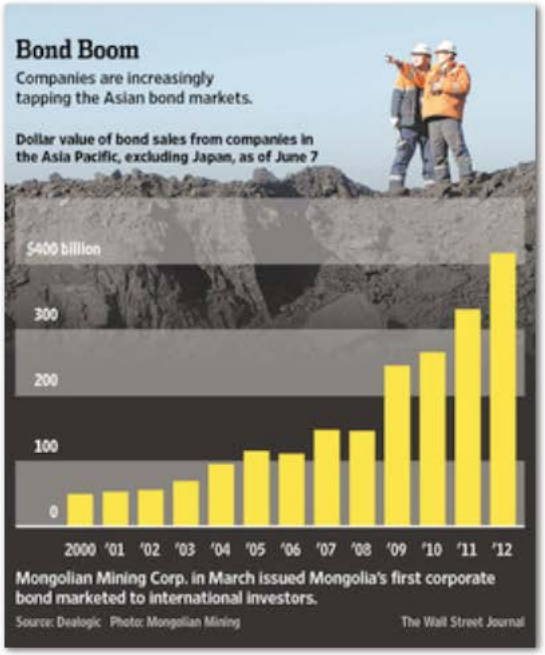


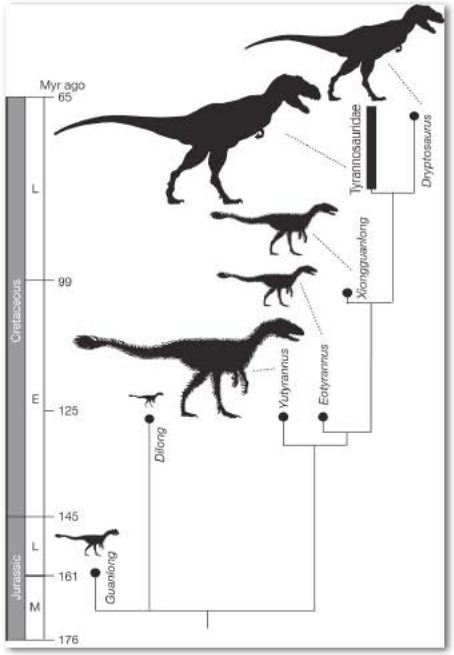
#4 Government Reports



Borkin et al What makes a Visualization Memorable? 2013 IEEE

A visualization was more memorable if...
it contains human recognizable objects.

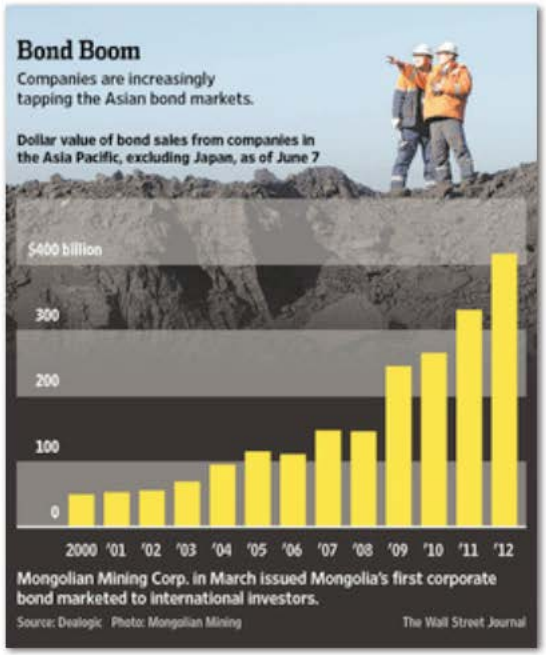


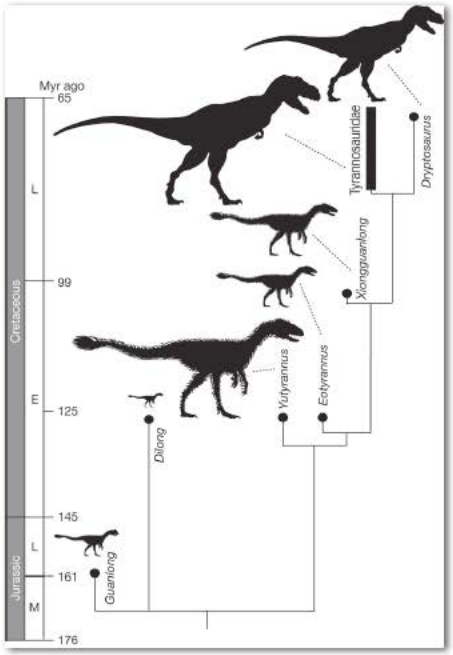


Borkin et al What makes a Visualization Memorable? 2013 IEEE

A visualization was more memorable if...

it is distinct.

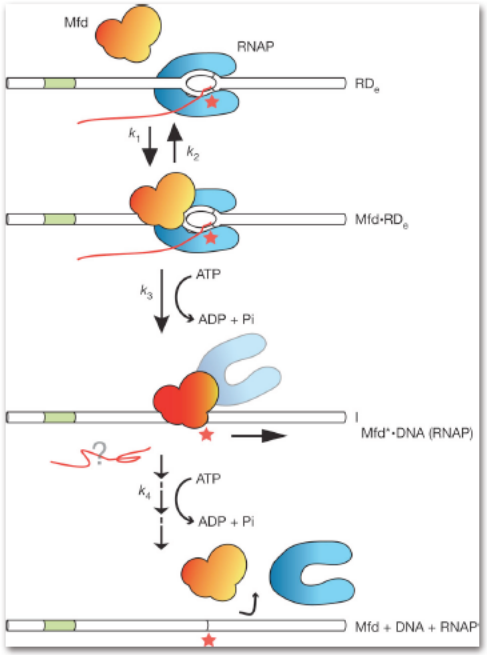


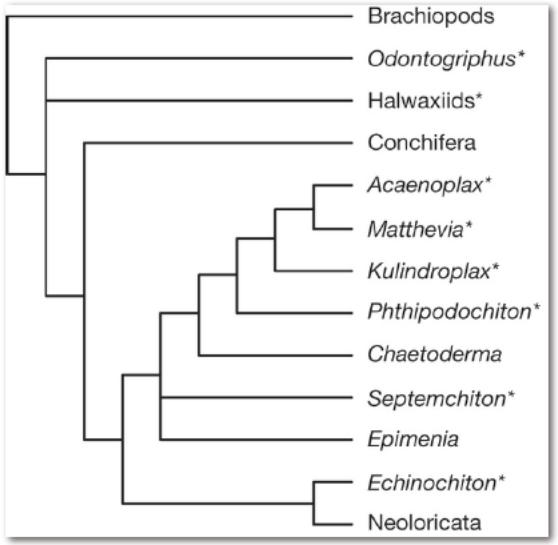


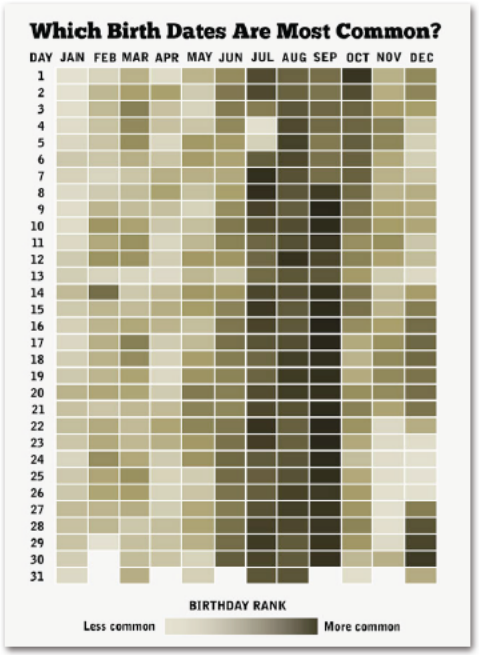
Borkin et al What makes a Visualization Memorable? 2013 IEEE

A visualization was more memorable if...

it is a distinct visualization type.







Borkin et al What makes a Visualization Memorable? 2013 IEEE

A visualization was more memorable if...

it is colorful.

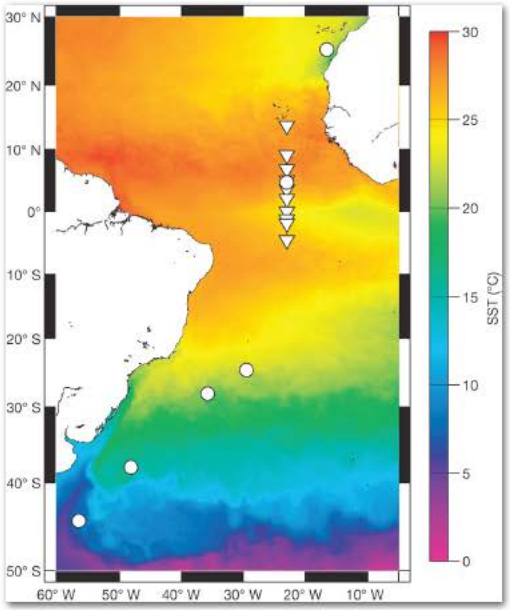
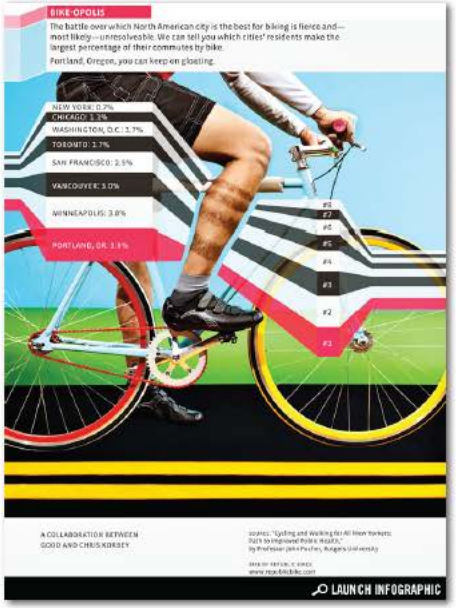


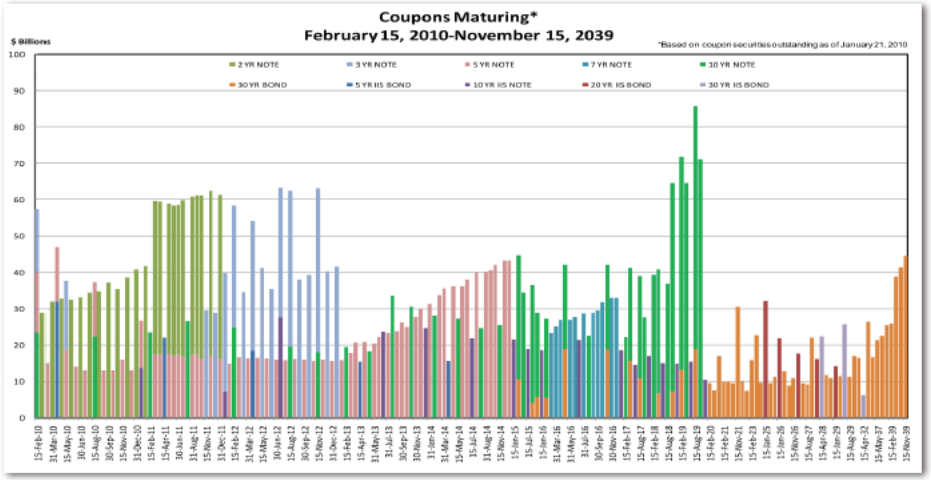
Figure 6.2: Policy shifts and interventions to enable wetland practices to accommodate notions of ecosystem services and human health

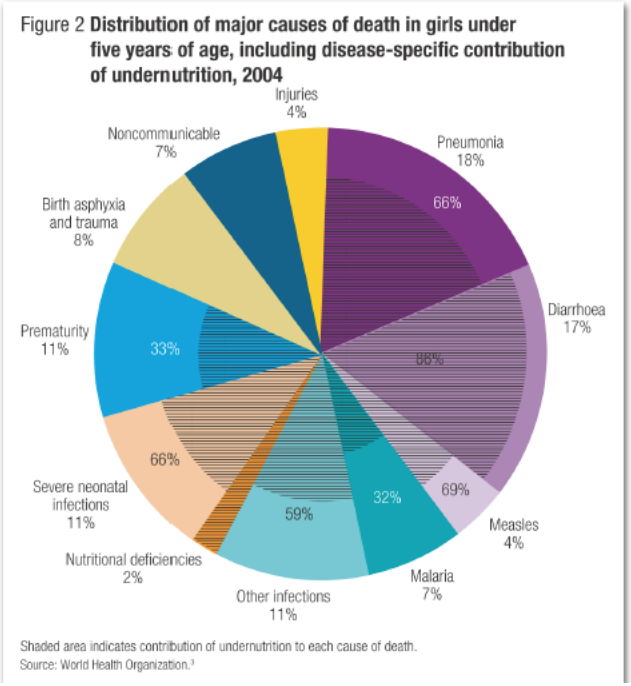


Borkin et al What makes a Visualization Memorable? 2013 IEEE

A visualization was more memorable if...

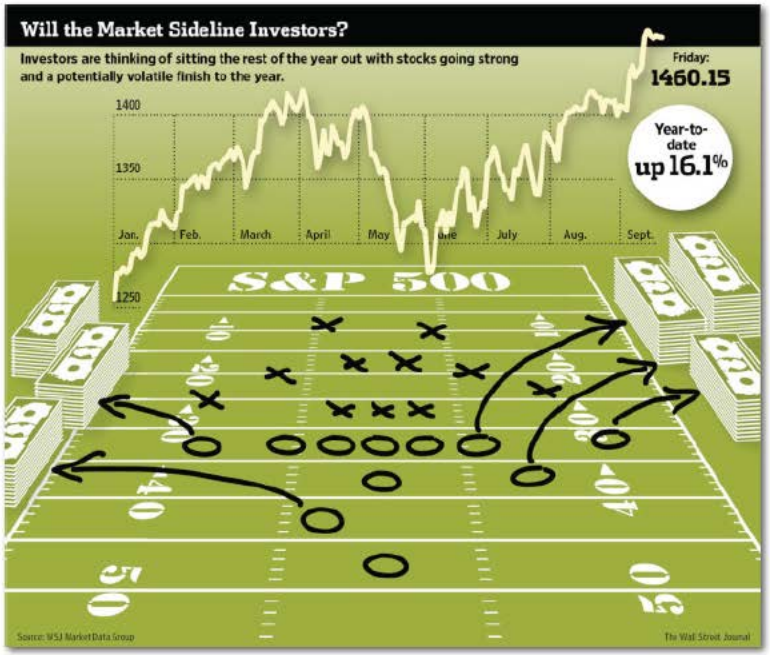
it is visually dense.

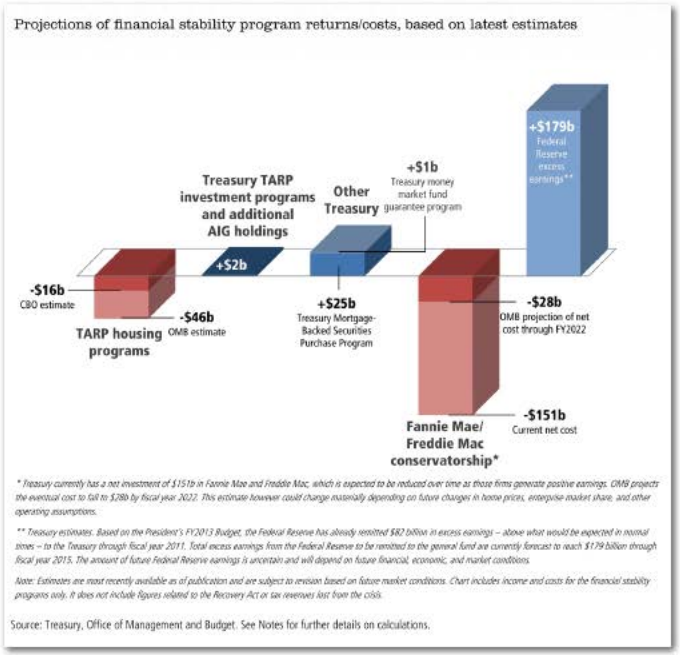




Borkin et al What makes a Visualization Memorable? 2013 IEEE

A visualization was more memorable if...
it has a low data-to-ink ratio.





Borkin et al What makes a Visualization Memorable? 2013 IEEE

1 | DATA

2 | WORKING PARTS

3 | GRAPH TYPES

4 | WHAT ARE WE TRYING TO SHOW?

5 | EXAMPLES & EXPLANATIONS

GRAPHS AREN'T ALWAYS NECESSARY

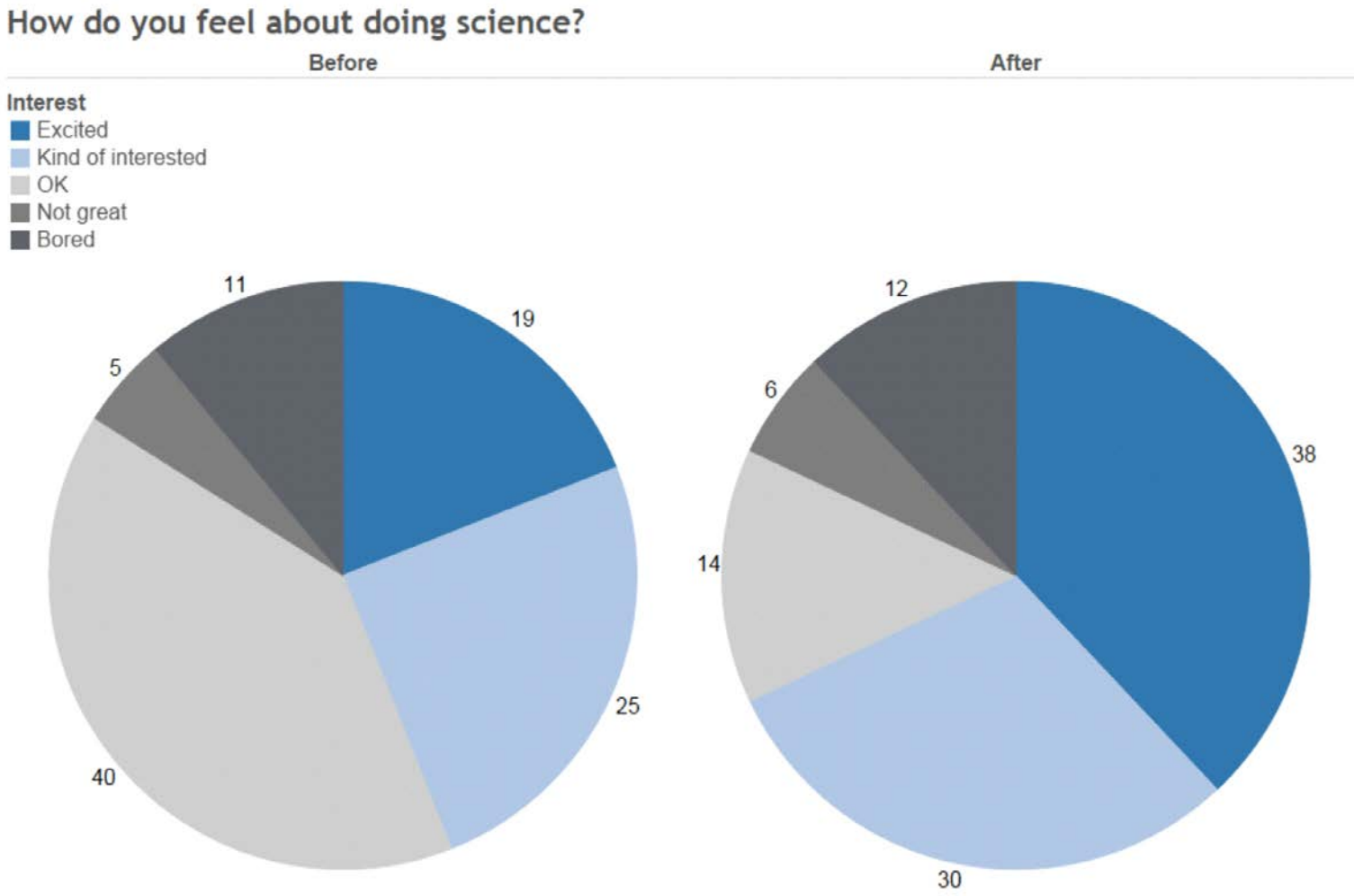
HOW DO YOU FEEL ABOUT DOING SCIENCE?

Table

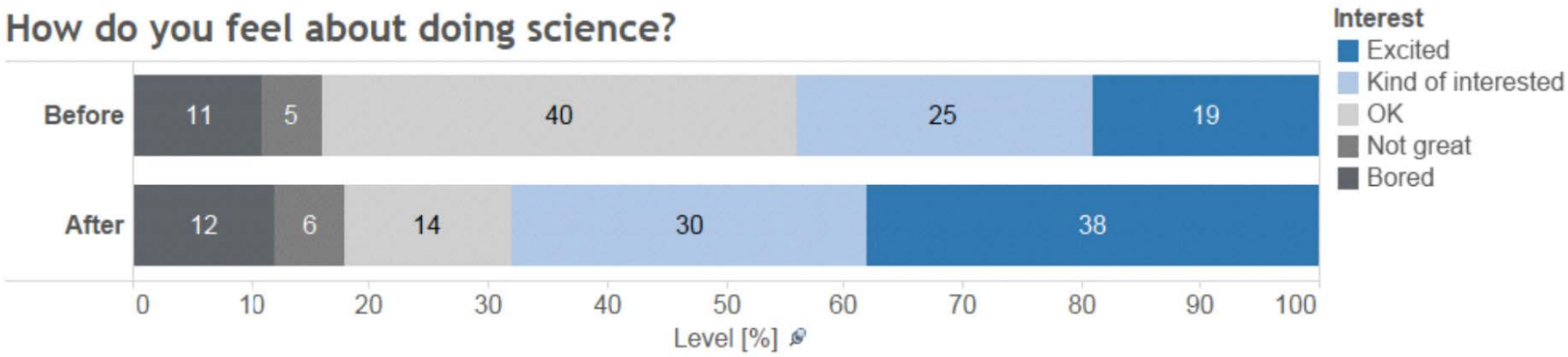
Interest	Before	After
Excited	19	38
Kind of interested	25	30
OK	40	14
Not great	5	6
Bored	11	12

Data courtesy of Cole Nussbaumer

Example courtesy of Hanspeter Pfister

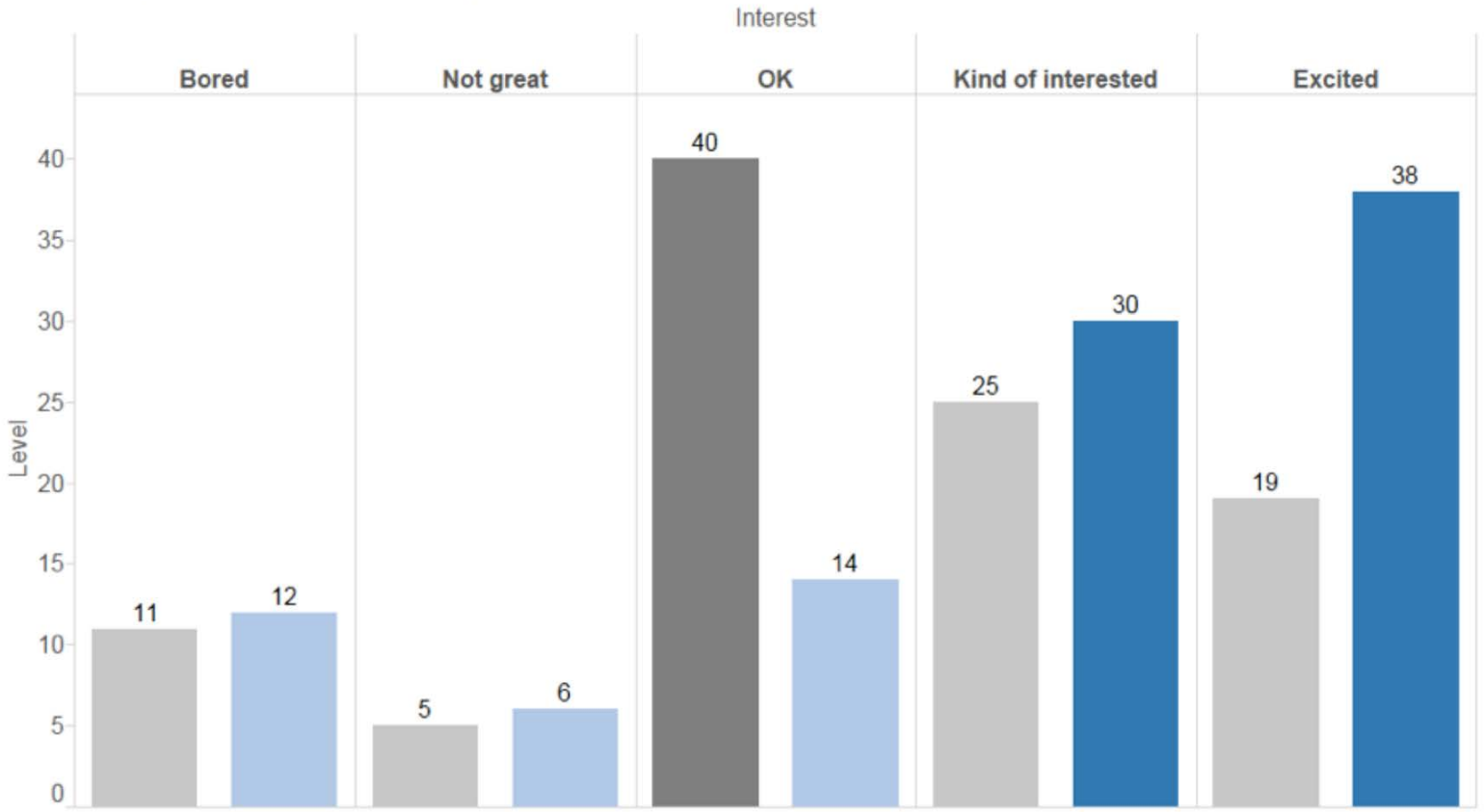


Example courtesy of Hanspeter Pfister



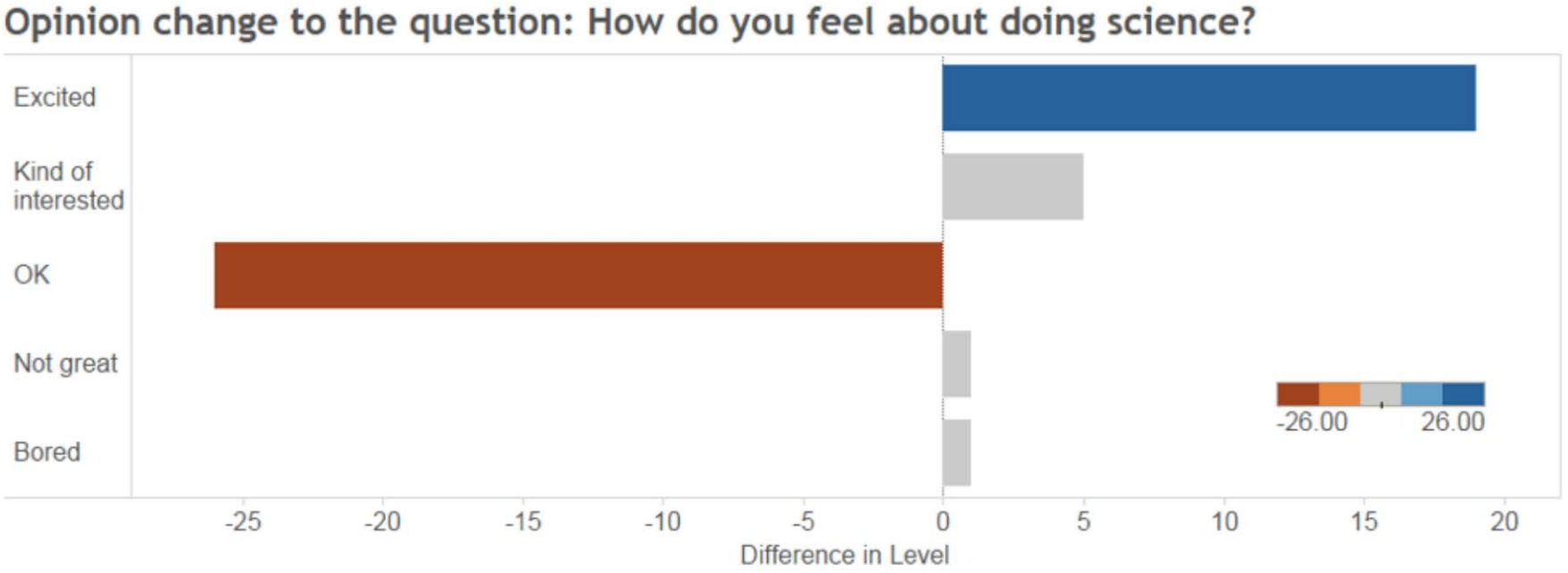
Example courtesy of Hanspeter Pfister

How do you feel about doing science?

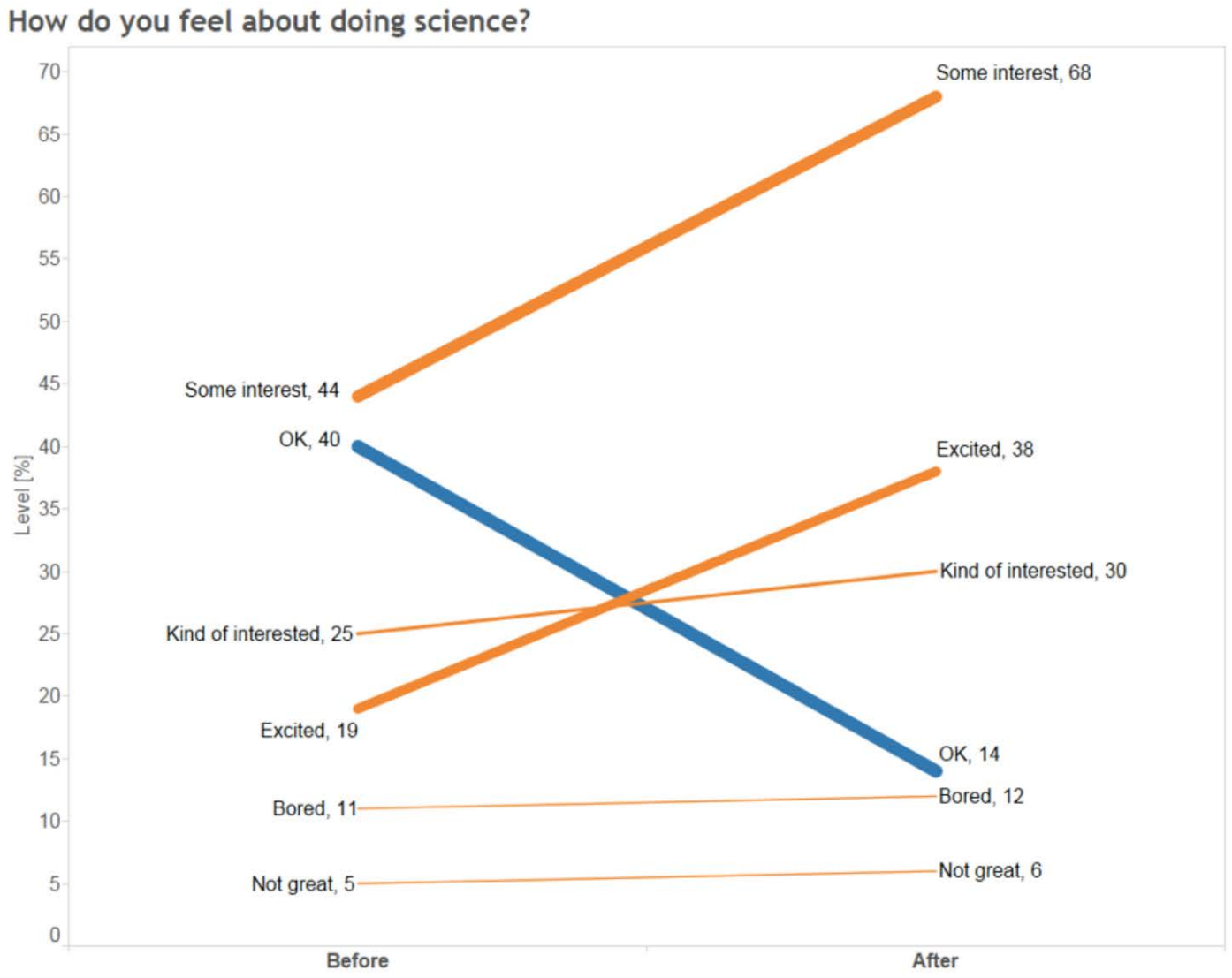


Before the program, the majority of children felt just ***OK*** about science. After the program, more children were ***Kind of interested*** and ***Excited*** about science.

Example courtesy of Hanspeter Pfister



Example courtesy of Hanspeter Pfister



Example courtesy of Hanspeter Pfister

After the pilot program,

68%

of kids expressed interest towards science,
compared to 44% going into the program.

Example courtesy of Hanspeter Pfister

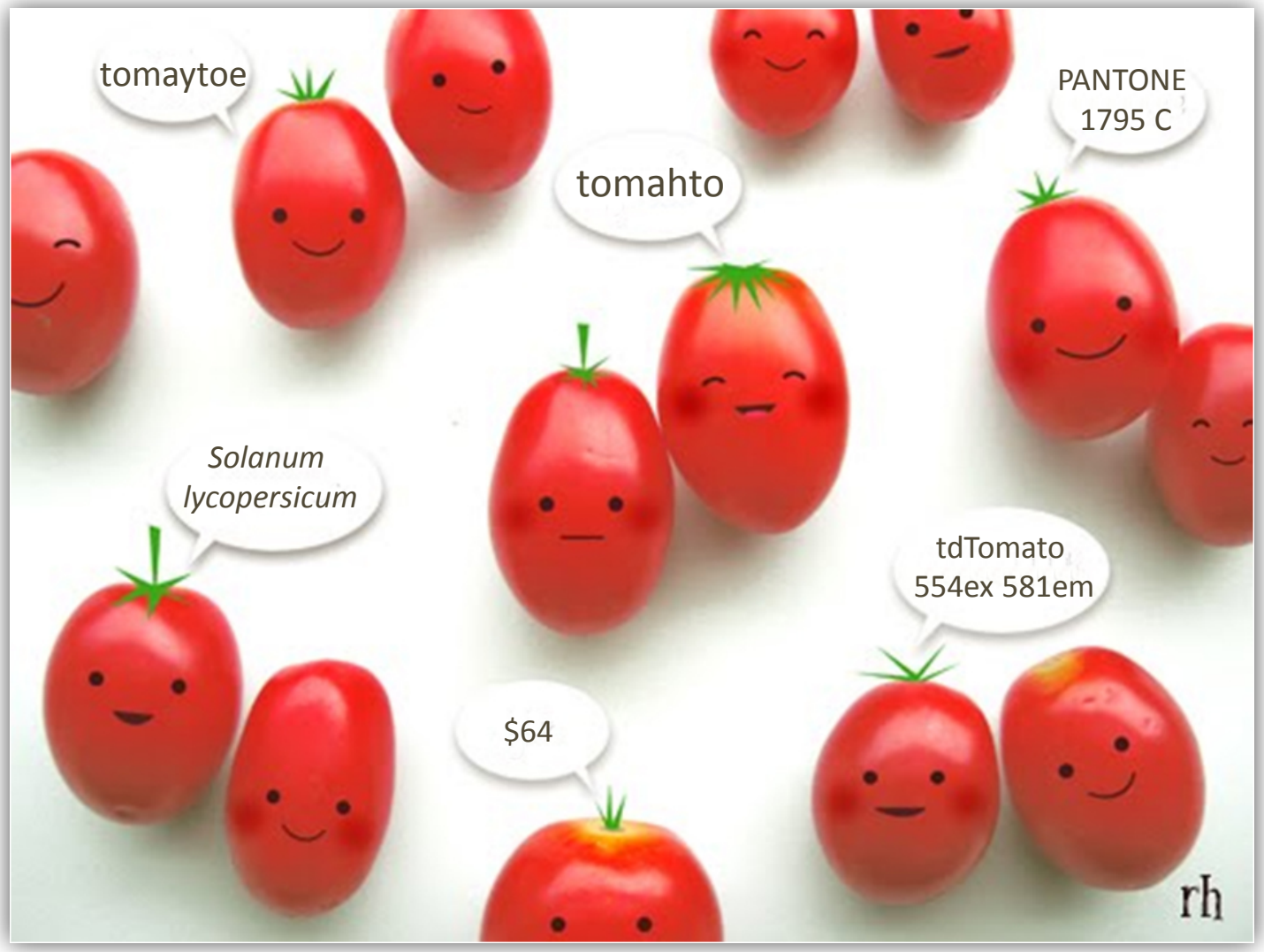
1 | DATA

2 | WORKING PARTS

3 | GRAPH TYPES

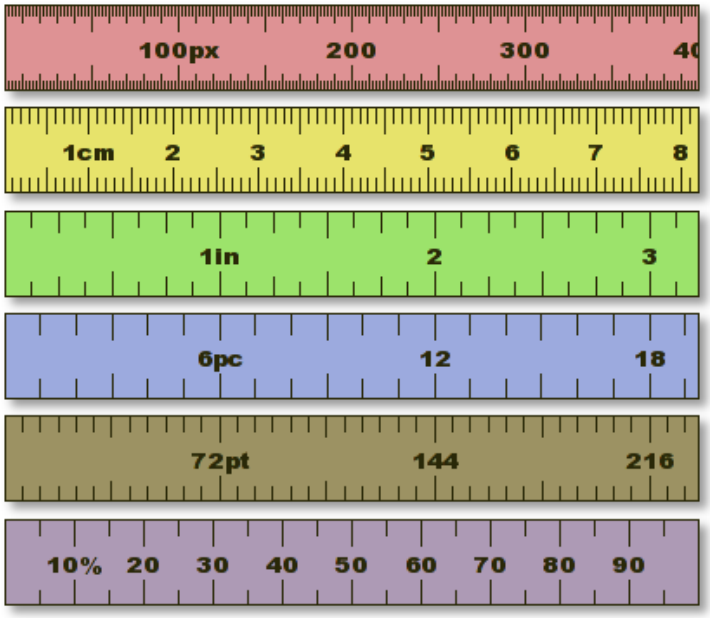
4 | WHAT ARE WE TRYING TO SHOW?

5 | EXAMPLES & EXPLANATIONS


tomaytoe
tomahto
PANTONE
1795 C
Solanum
lycopersicum
tdTomato
554ex 581em
$64
rh

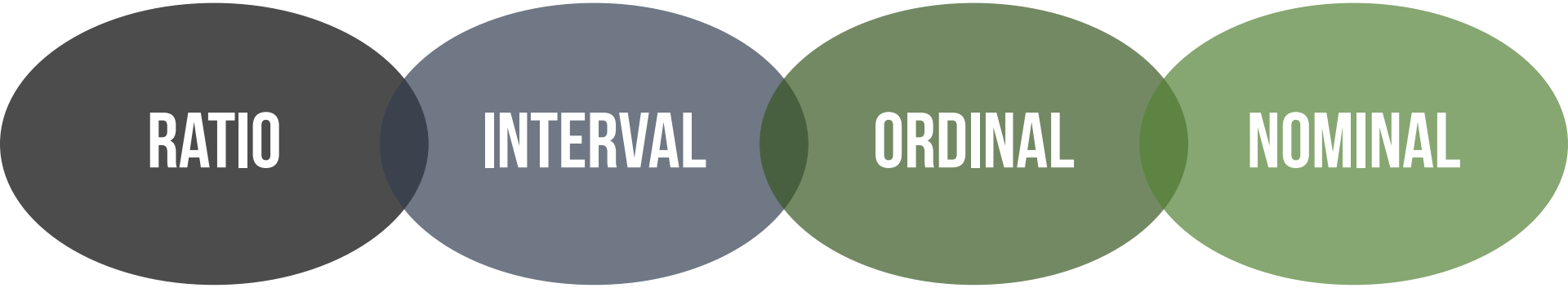
DATA DEFINITIONS

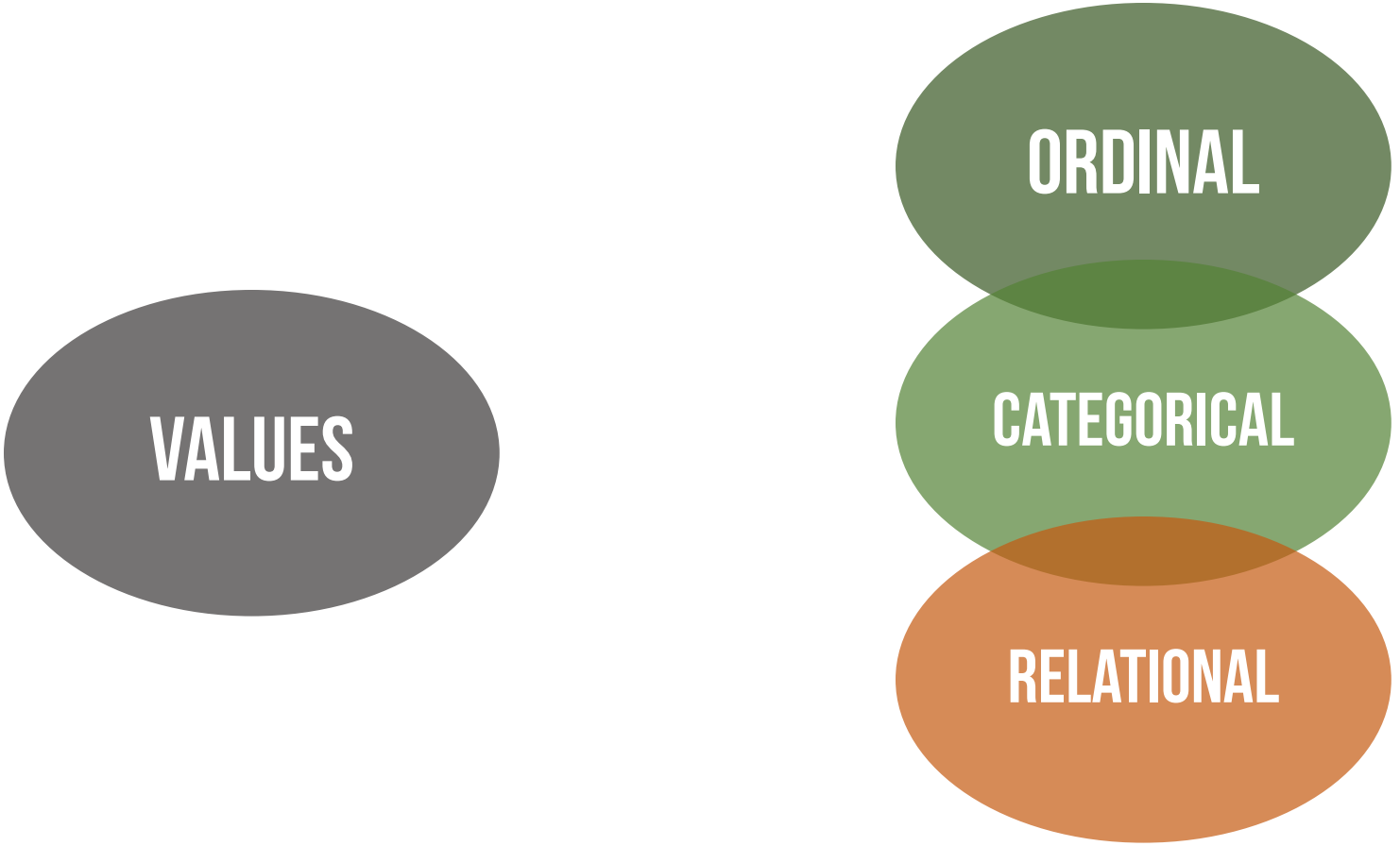
QUANTITATIVE | QUALITATIVE



QUANTITATIVE | QUALITATIVE

QUANTITATIVE QUALITATIVE

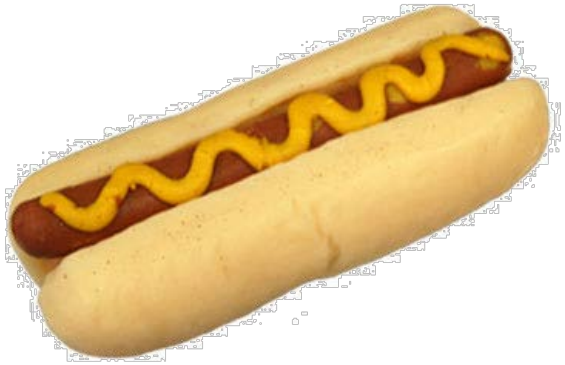



VALUES
ORDINAL
CATEGORICAL
RELATIONAL

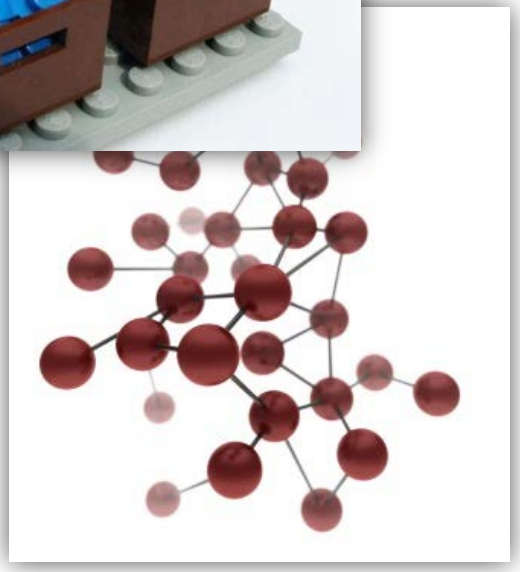
SO WHAT?

IMPLIED GRAPHICAL CONTENT

VALUES



RELATIONSHIPS



1 | DATA

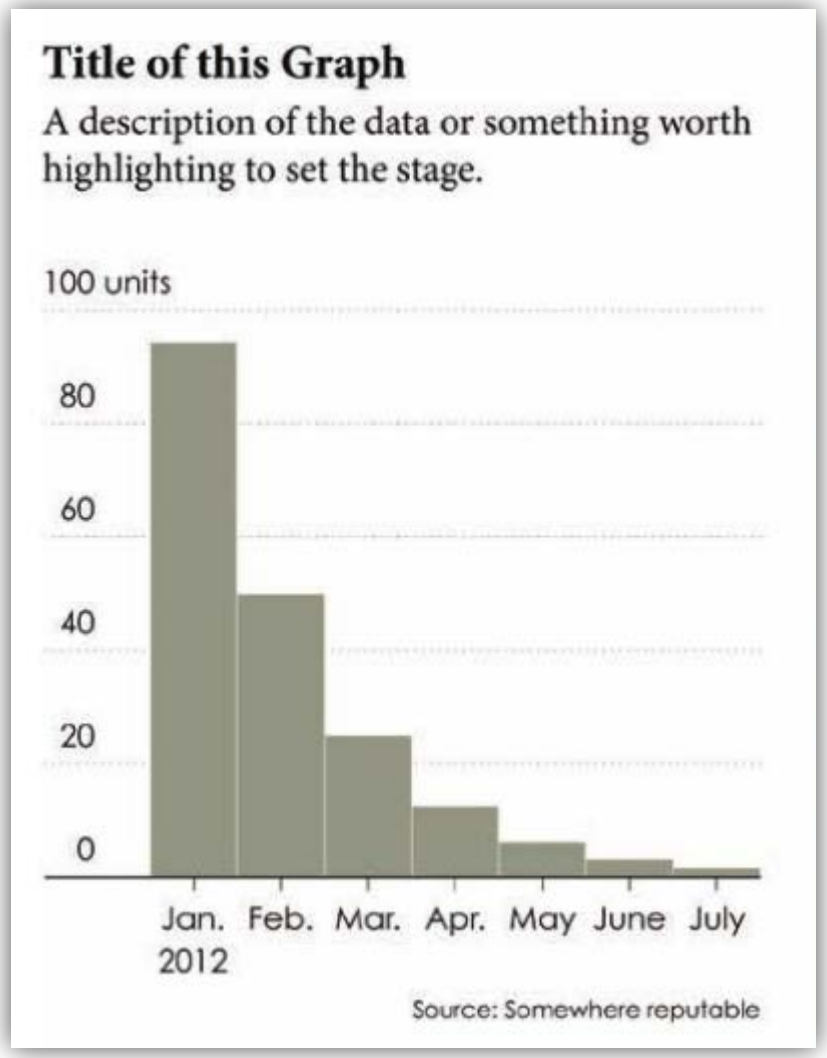
2 | WORKING PARTS

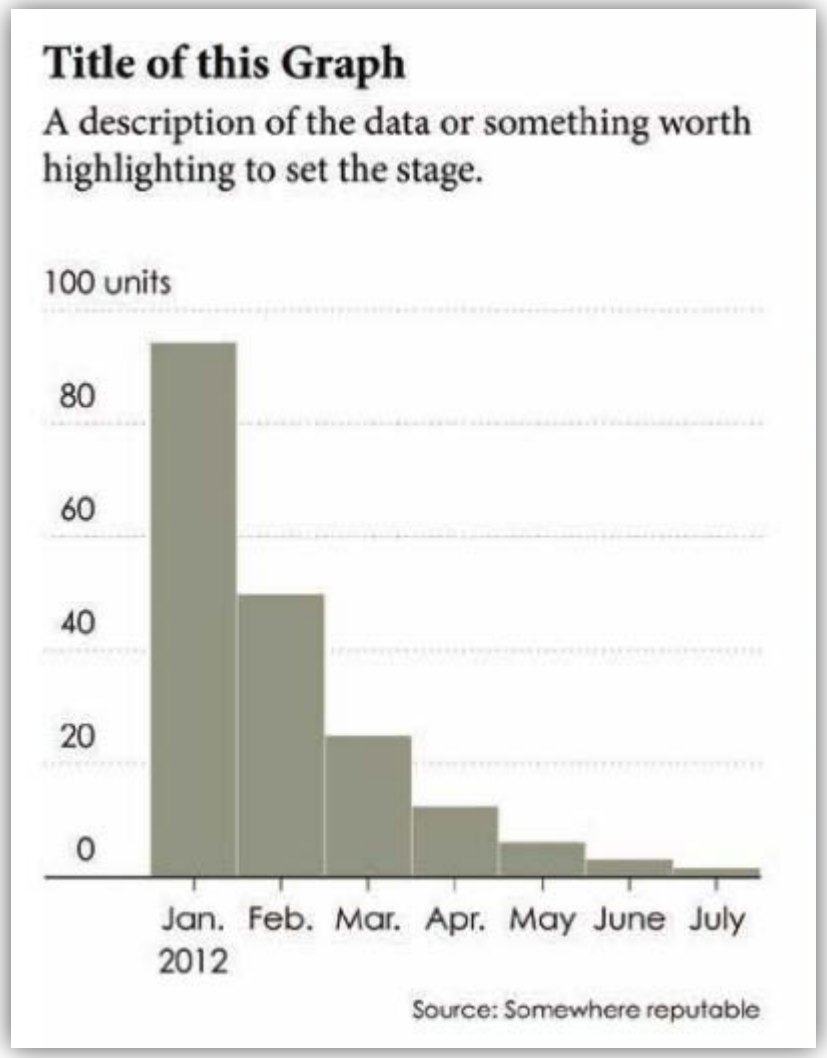
3 | GRAPH TYPES

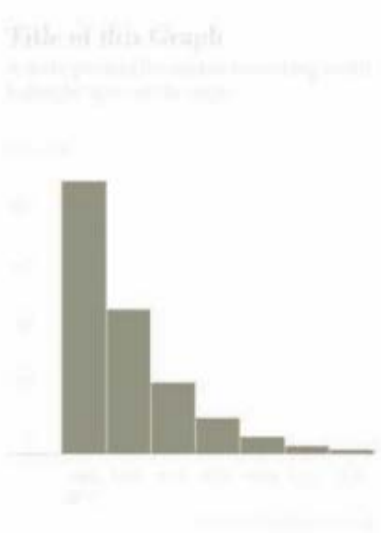
4 | WHAT ARE WE TRYING TO SHOW?

5 | EXAMPLES & EXPLANATIONS

GRAPHS:
WORKING PARTS


Title of this Graph
A description of the data or something worth highlighting to set the stage.
100 units
80
60
40
20
0
Jan.
2012
Feb.
Mar.
Apr.
May
June
July
Source: Somewhere reputable



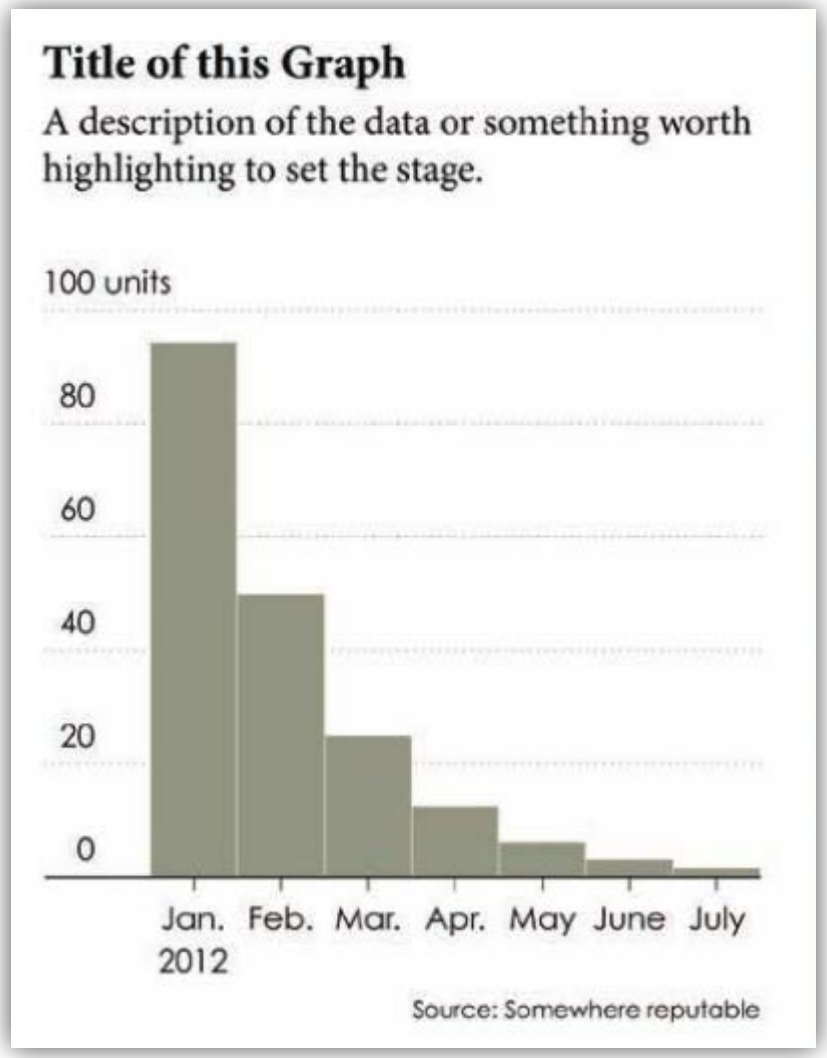


Visual Cues

Visualization involves encoding data with shapes, colors, and sizes. Which cues you choose depends on your data and your goals.

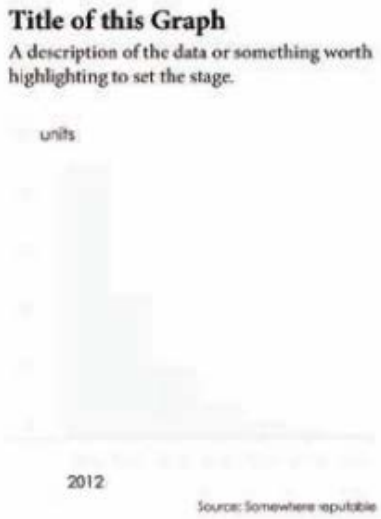
Coordinate System

You map data differently with a scatterplot than you do with a pie chart. It's x- and y-coordinates in one and angles with the other; it's cartesian versus polar.



Scale

Increments that make sense can increase readability, as well as shift focus.



Context

If your audience is unfamiliar with the data, it's your job to clarify what values represent and explain how people should read your visualization.

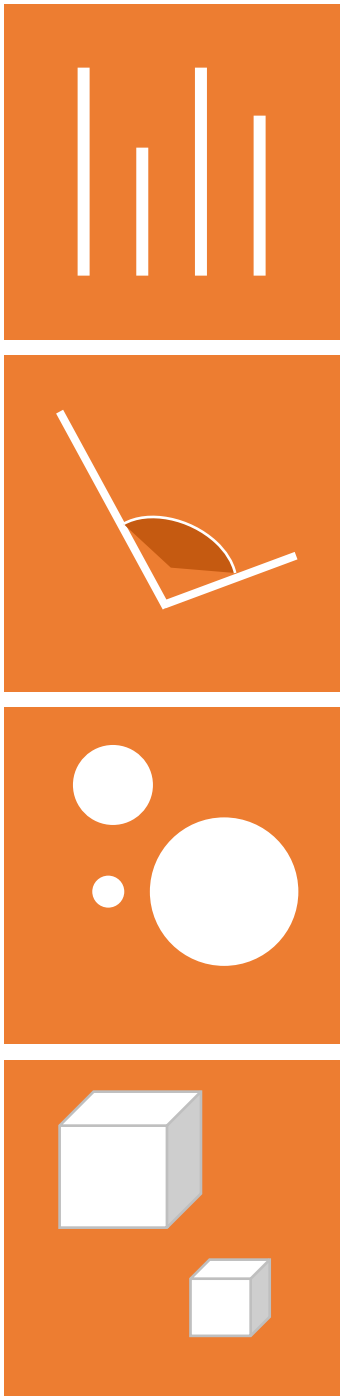
GRAPHING: VISUAL CUES

PROXIMITY

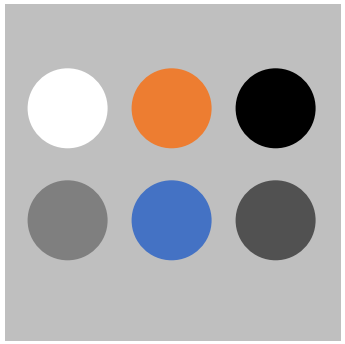
VALUES

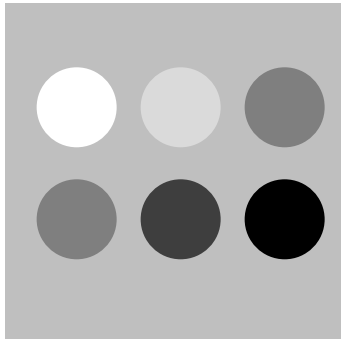
COLOR

PROXIMITY



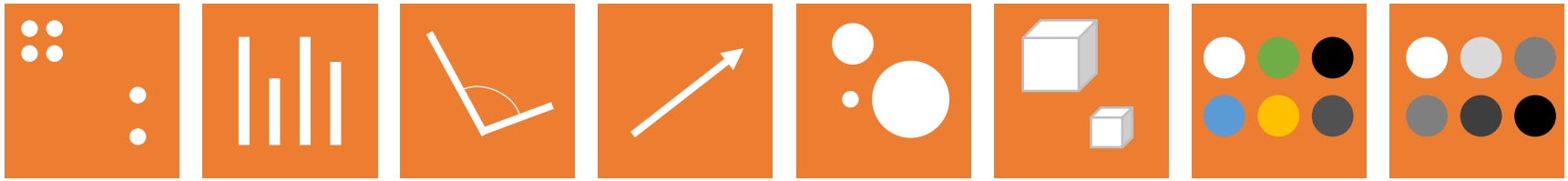
VALUES





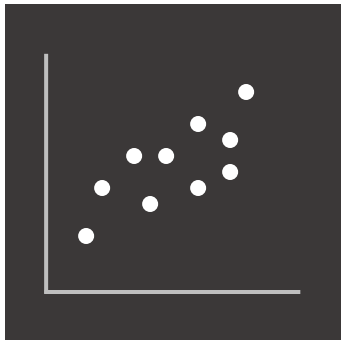
COLOR

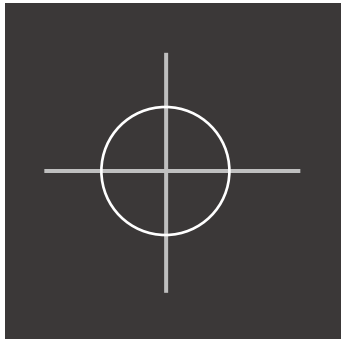
QUANTITATIVE QUALITATIVE



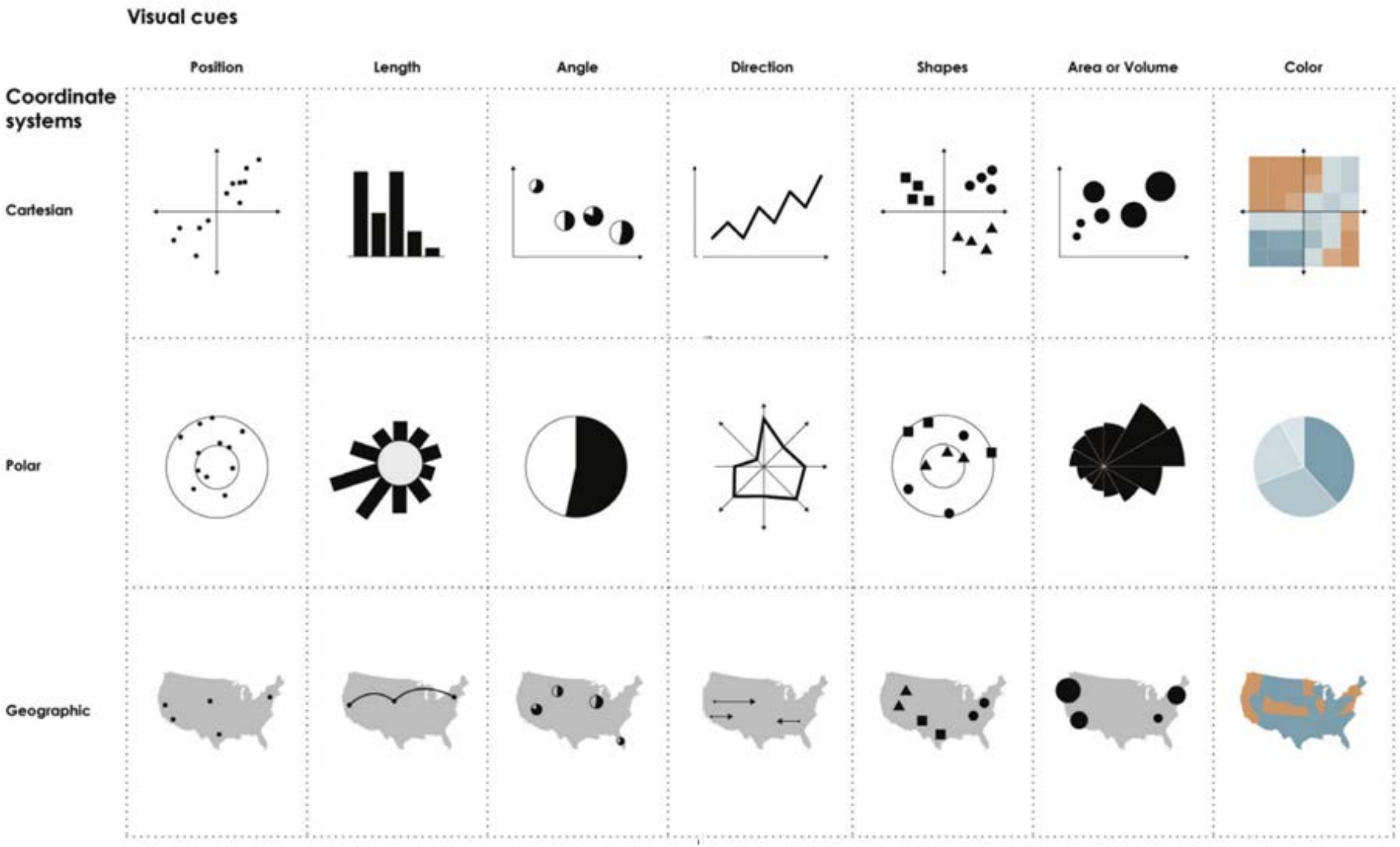
ASSOCIATION DIFFERENTIATION

GRAPHING: COORDINATE SYSTEMS





VISUAL CUES + COORDINATE SYSTEMS

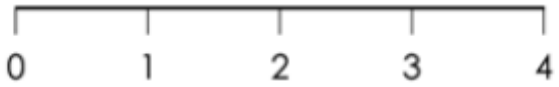


Nathan Yau, Data Points

GRAPHING: SCALE

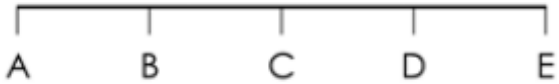
Linear

Values are evenly spaced



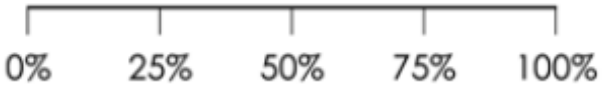
Categorical

Discrete placement in bins



Percent

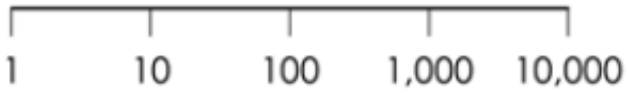
Representing parts of a whole



GRAPHING: SCALE

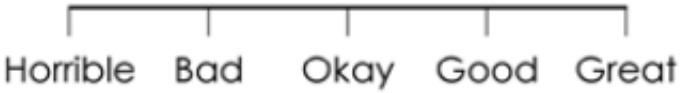
Logarithmic

Focus on percent change



Ordinal

Categories where order matters



Time

Units of months, days, or hours

GRAPHING: CONTEXT

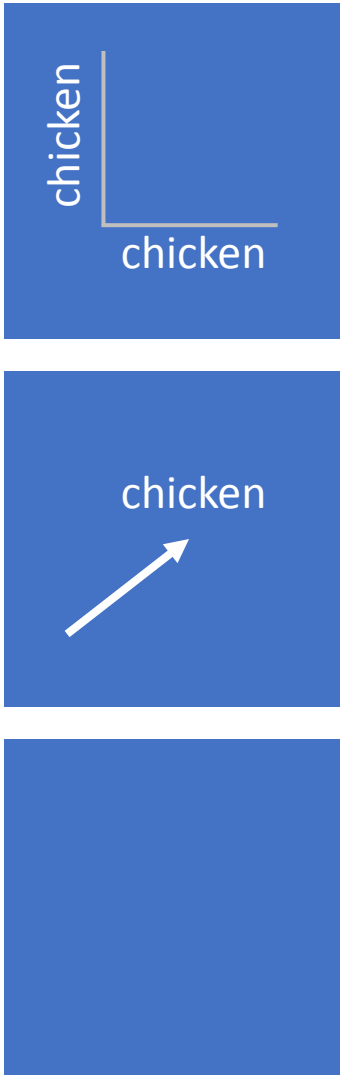


Figure 1. chicken

1 | DATA

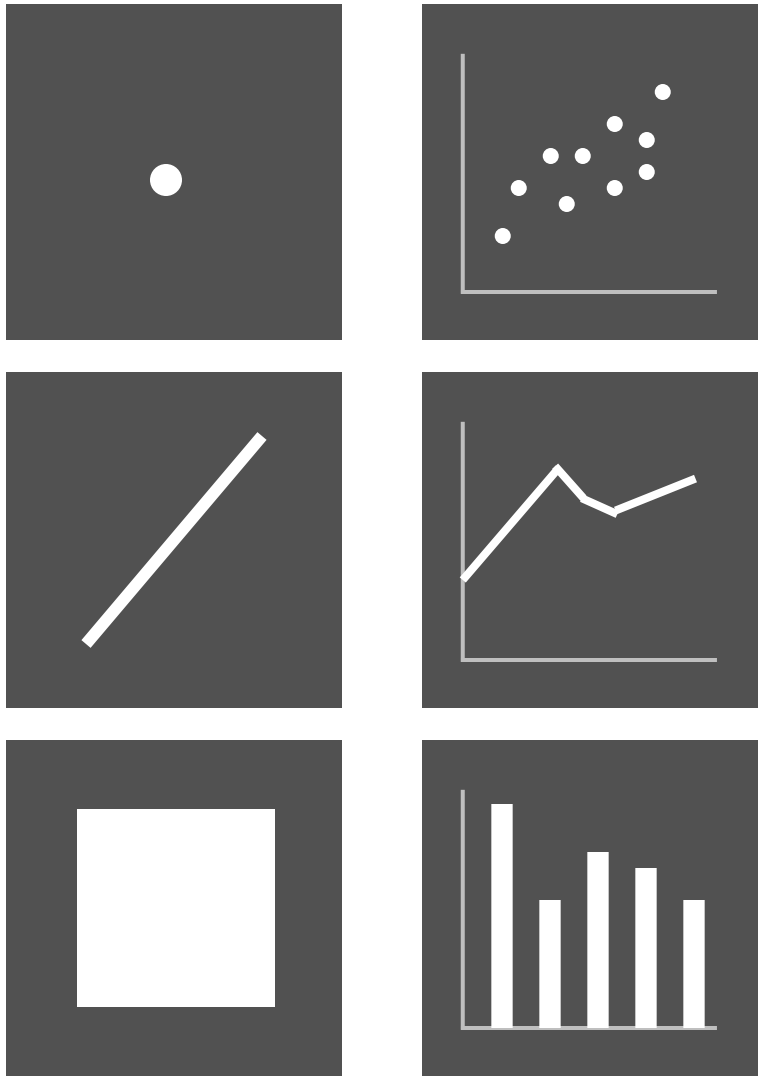
2 | WORKING PARTS

3 | GRAPH TYPES

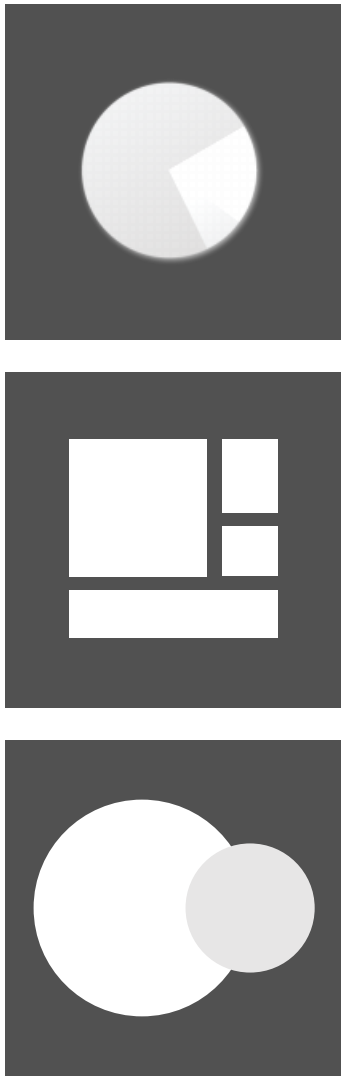
4 | WHAT ARE WE TRYING TO SHOW?

5 | EXAMPLES & EXPLANATIONS

GRAPH TYPES: HEAVER HITTERS



GRAPH TYPES: OTHER



GRAPH TYPES: EXOTIC

1 | DATA

2 | WORKING PARTS

3 | GRAPH TYPES

4 | WHAT ARE WE TRYING TO SHOW?

5 | EXAMPLES & EXPLANATIONS

PATTERNS AND SHAPE

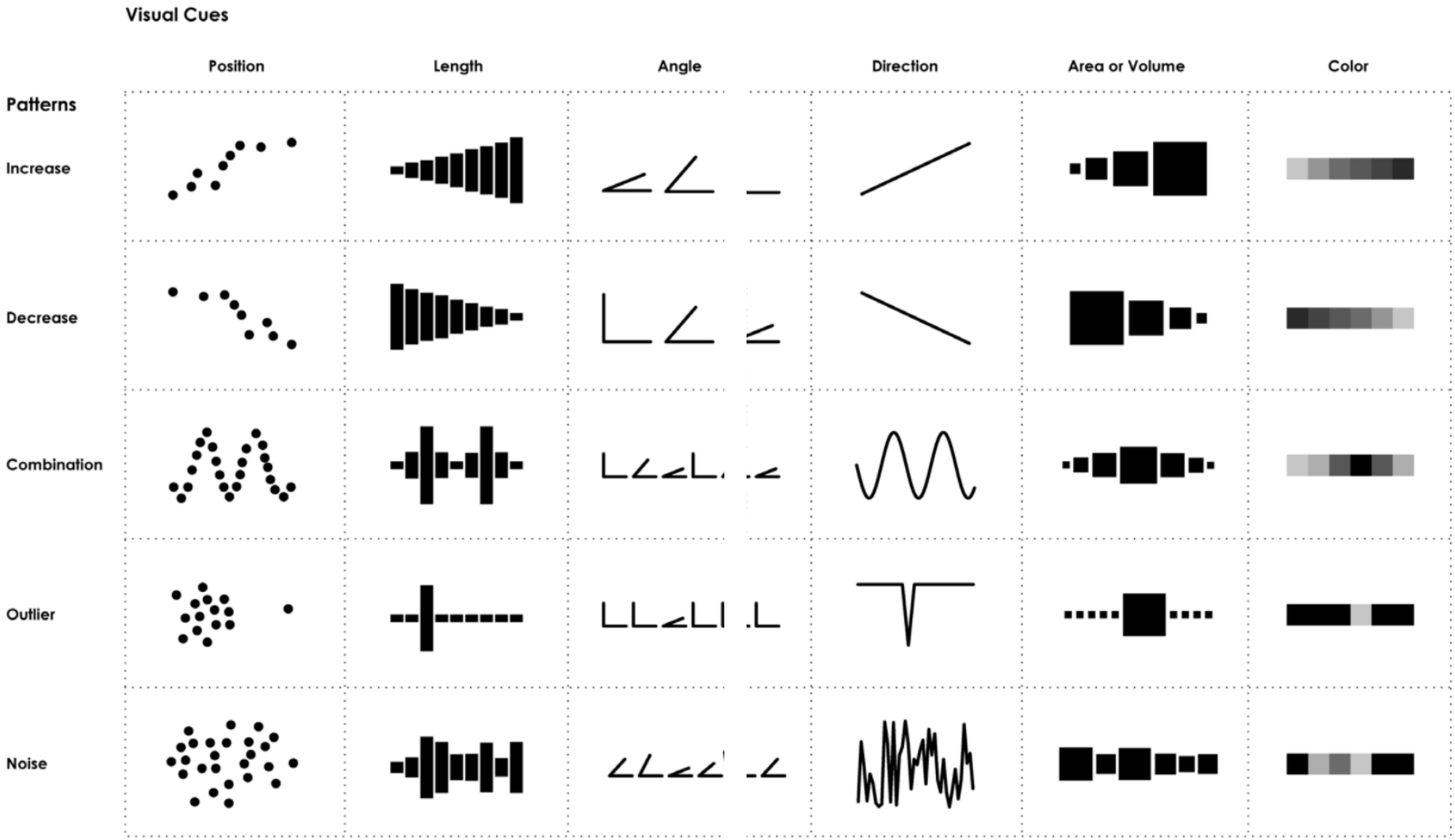
INCREASE

DECREASE

COMBINATION

OUTLIER

NOISE



Nathan Yau, Data Points

CONCEPTS

COMPARISON

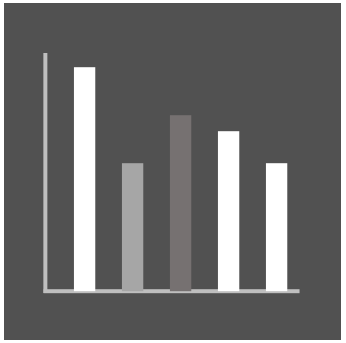
TRENDS

DISTRIBUTIONS

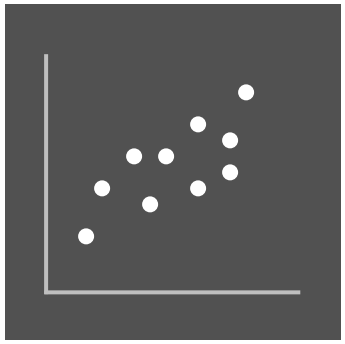
PROPORTIONS

CORRELATIONS

COMPARE ITEMS



TRENDS OVER TIME

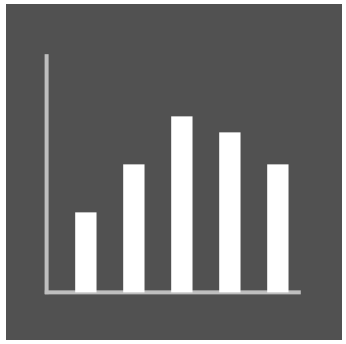


CORRELATION

Comparison

Relationship — What would you like to show? — *Distribution*

Composition



DISTRIBUTION

PROPORTIONS

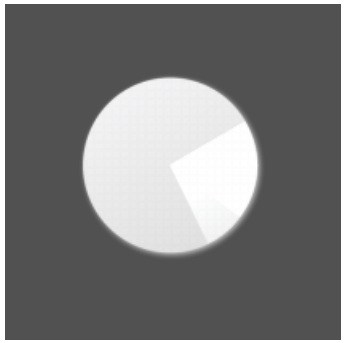
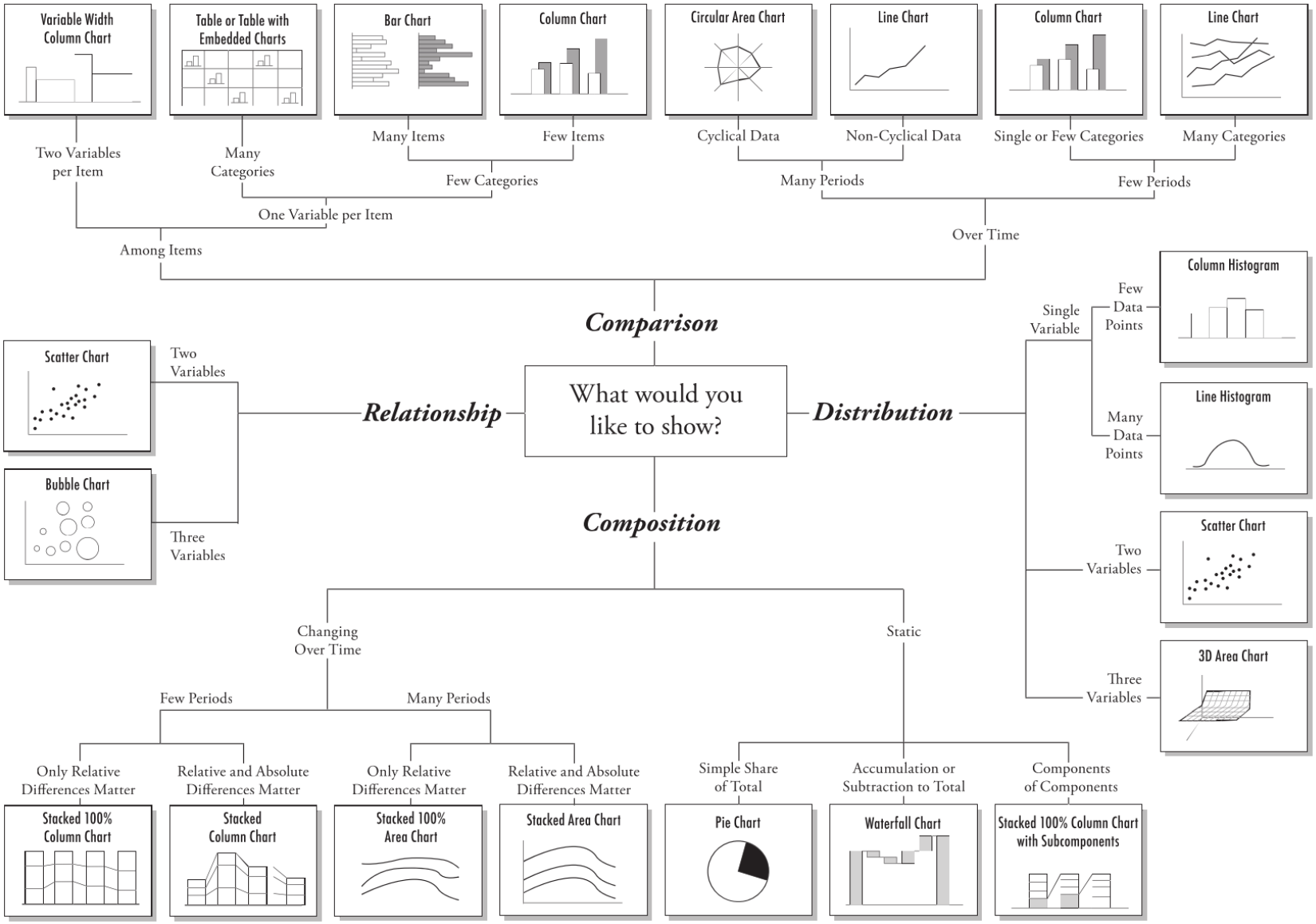


Chart Suggestions—A Thought-Starter

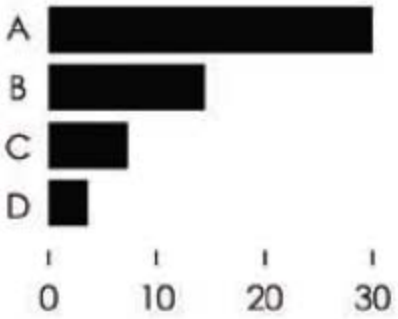


CATEGORICAL DATA:
COMPARISON/COMPOSITION

Categories

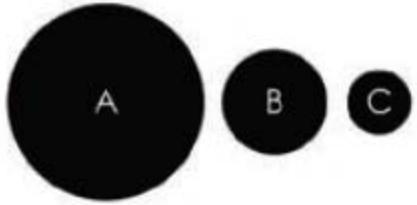
When your data is straightforward, with a value for each category, these are easy to read and create.

Bar graph



With length as visual cue, useful for straightforward comparisons

Symbol plot

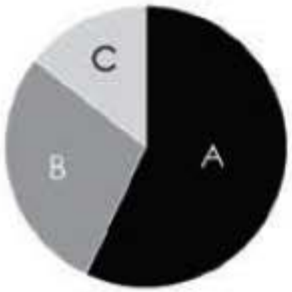


Can be used in place of bars, but can be hard to see small differences

Parts of a whole

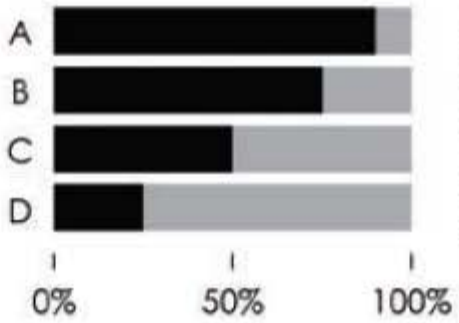
The categorical breakdown within a population can be interesting, and you might want to keep the groups together, although often not essential.

Pie chart



Parts add to 100 percent, typically sorted clockwise for readability

Stacked bar chart

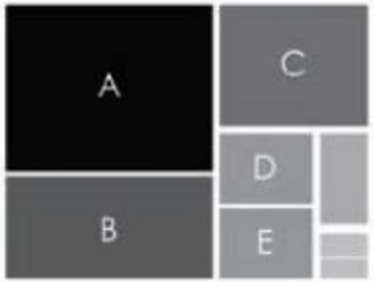


Often used to show poll results and can also be used for raw counts

Subcategories

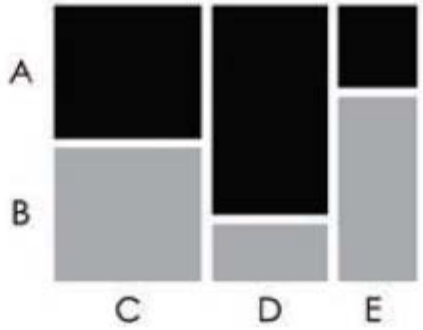
Data can have a hierarchical structure, which can be important in data interpretation and it often allows for different points of view.

Treemap



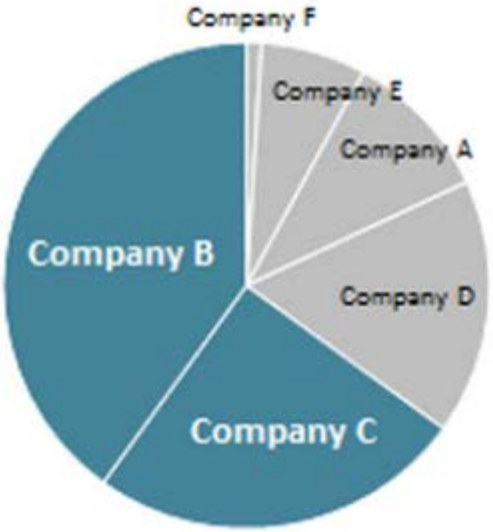
Shows hierarchical structure in a compact space, area often combined with color

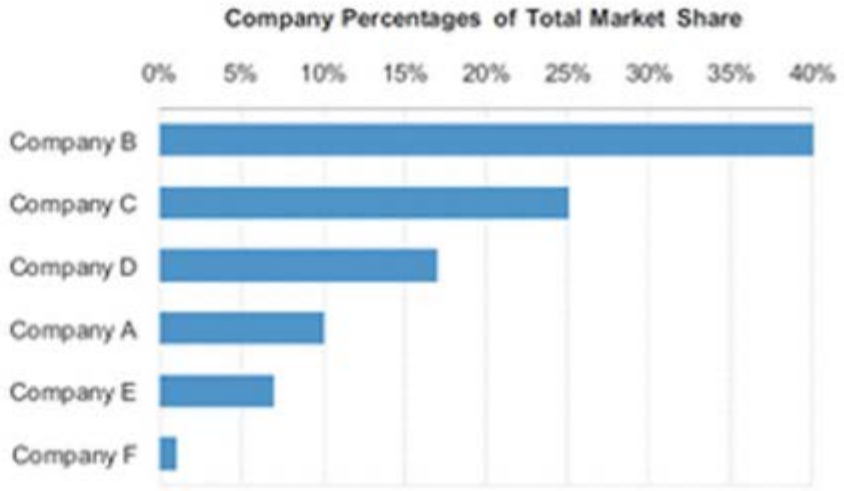
Mosaic plot

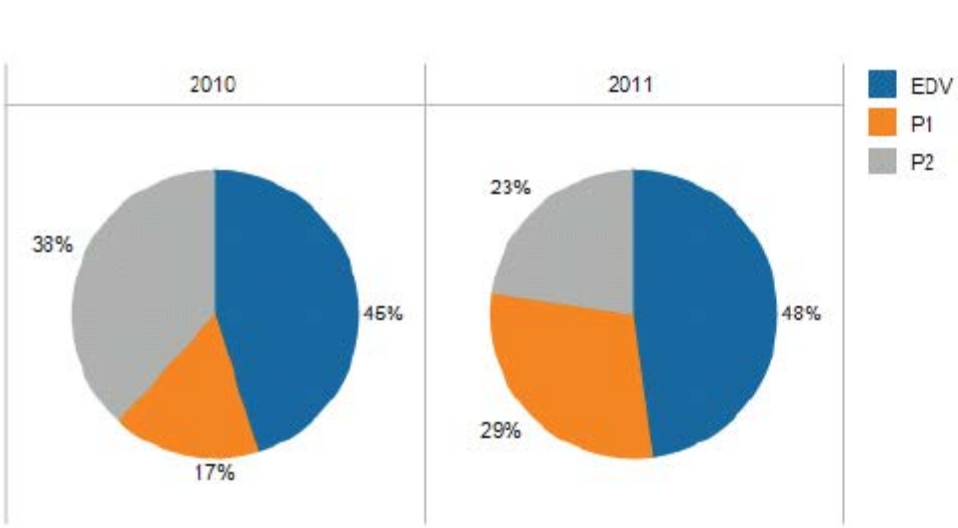


Allows comparisor across multiple categories in one view

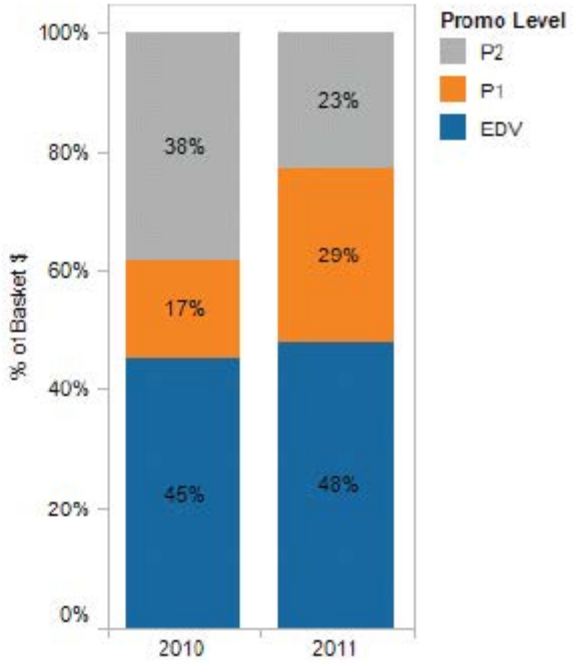
65% of the market is controlled by companies B and C

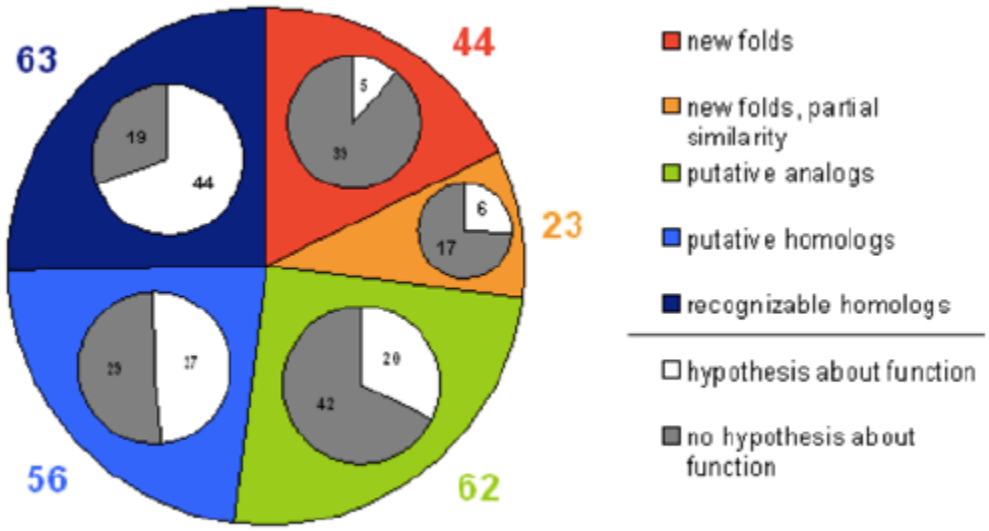





2010
2011
EDV
P1
P2
38%
46%
17%
23%
48%
29%

V.S.


Promo Level
P2
P1
EDV
% of Basket $
100%
80%
60%
40%
20%
0%
38%
17%
45%
23%
29%
48%
2010
2011


63
44
23
56
62
19
44
5
39
6
17
29
37
20
42
new folds
new folds, partial similarity
putative analogs
putative homologs
recognizable homologs
hypothesis about function
no hypothesis about function

RELATIONAL DATA

CORRELATION

Should You Trust That Doctor?

CollegeHumor

TRUST WITH YOUR LIFE

NO PROFESSIONAL TRAINING

POSSESSES A PROFESSIONAL DOCTORATE IN MEDICAL STUDIES

RUN. RUN VERY FAR.

The Doctor

Dr. Seuss

Dr. Strange

Dr. Quinn

Dr. Huxtable

Doogie Howser, MD

Doc Brown

J.D.

Spin Doctors

Doc

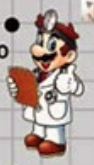
Dr. Bunsen Honeydew

Dr. Fishman

Dr. Dre

Zoidberg

Dr. Mario



Dr. Phil

Dr. Oz

Dr. Horrible

Doctor Venture

Dr. Spaceman

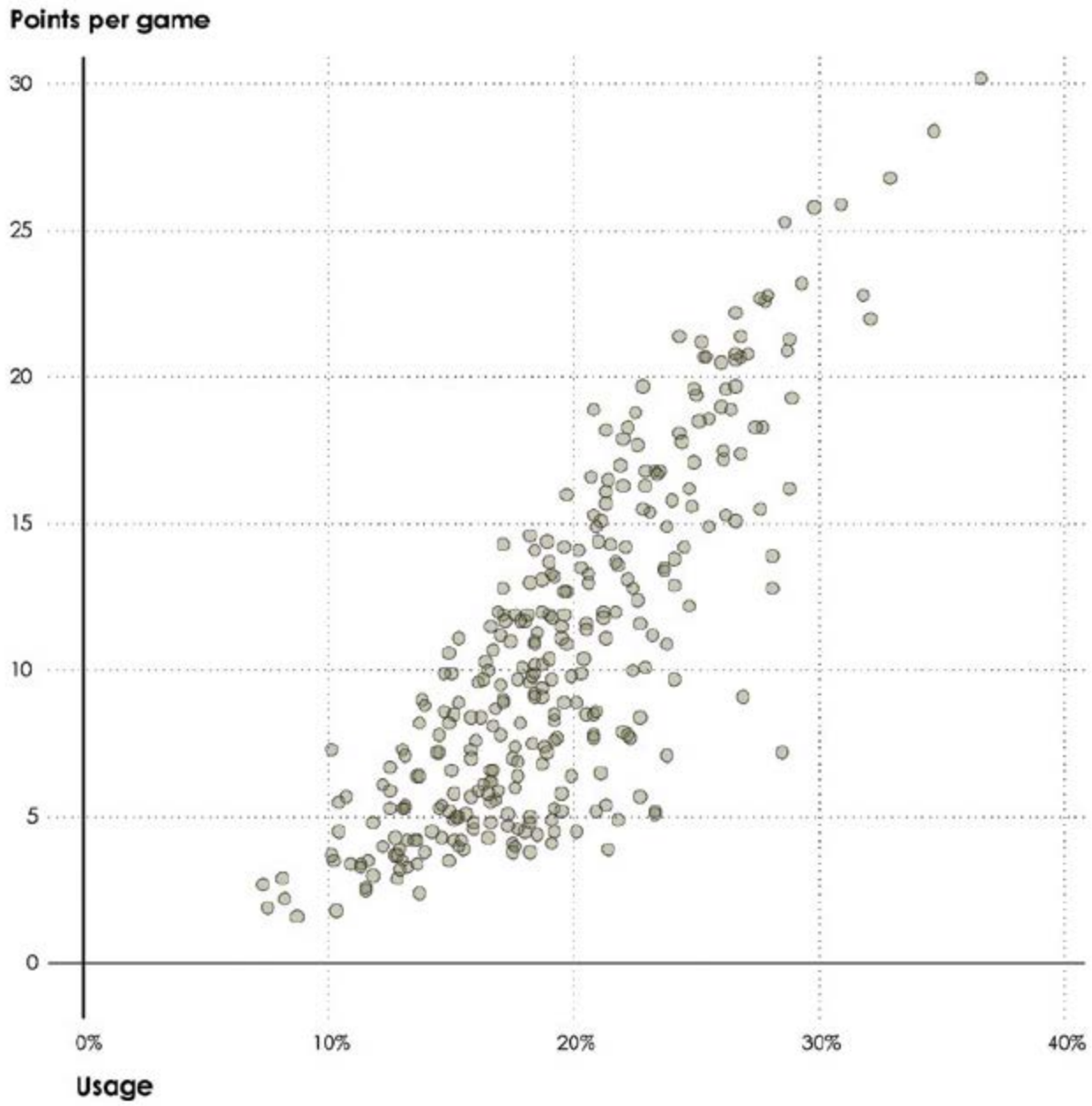
Doctor Octopus

Dr Pepper

Doctor Evil

Dr. Robotnik

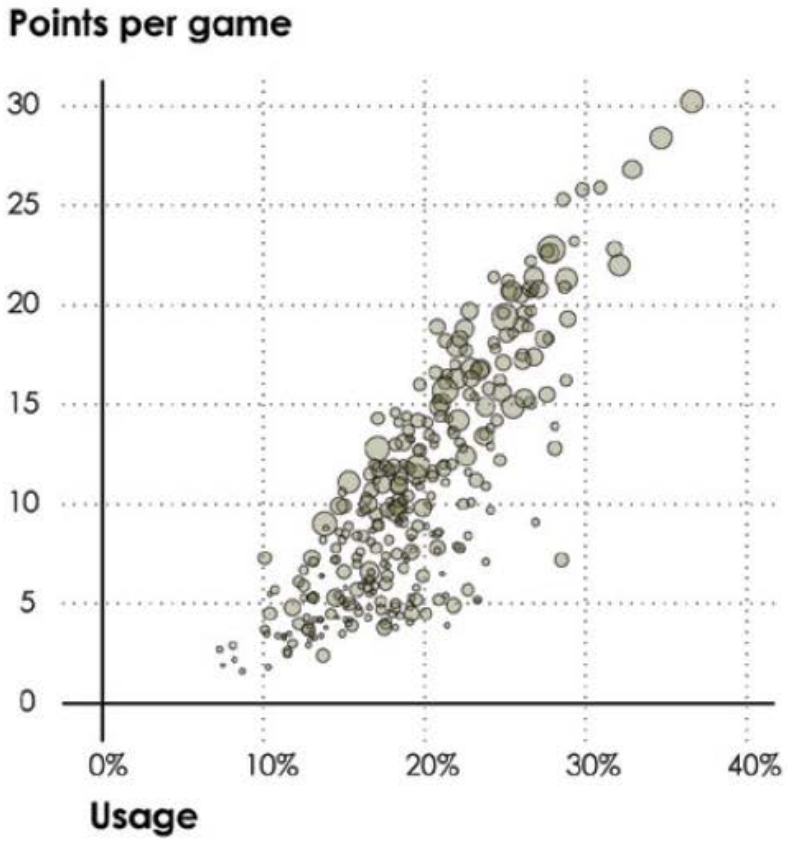
Dr. Hannibal Lector

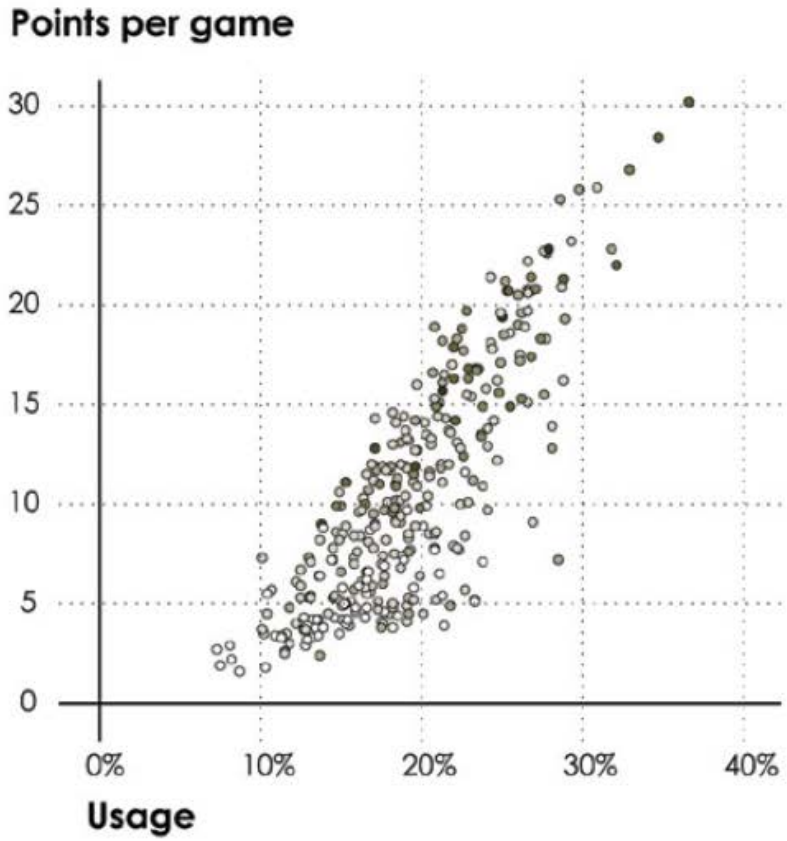


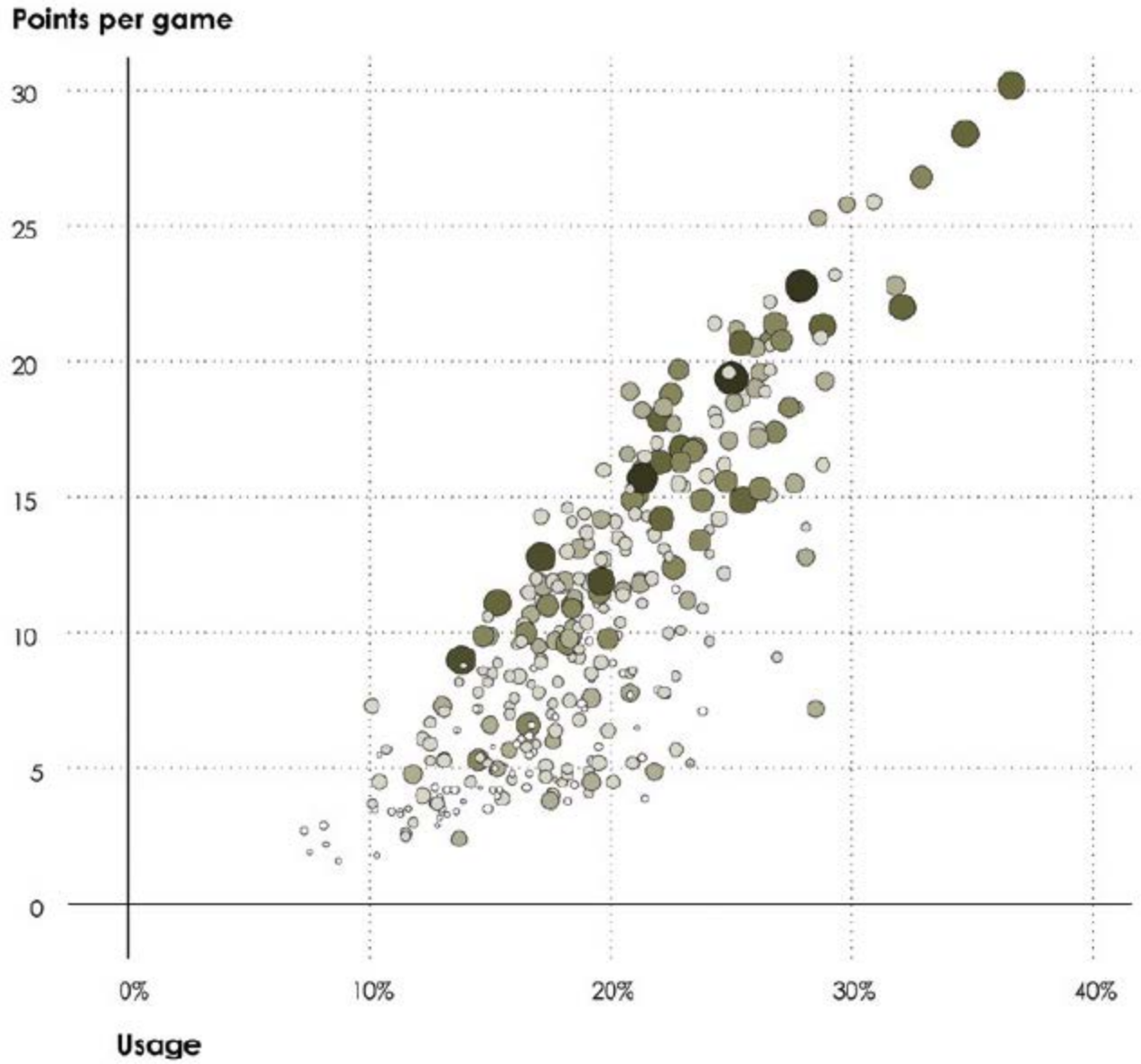
Nathan Yau, Data Points

RELATIONAL DATA

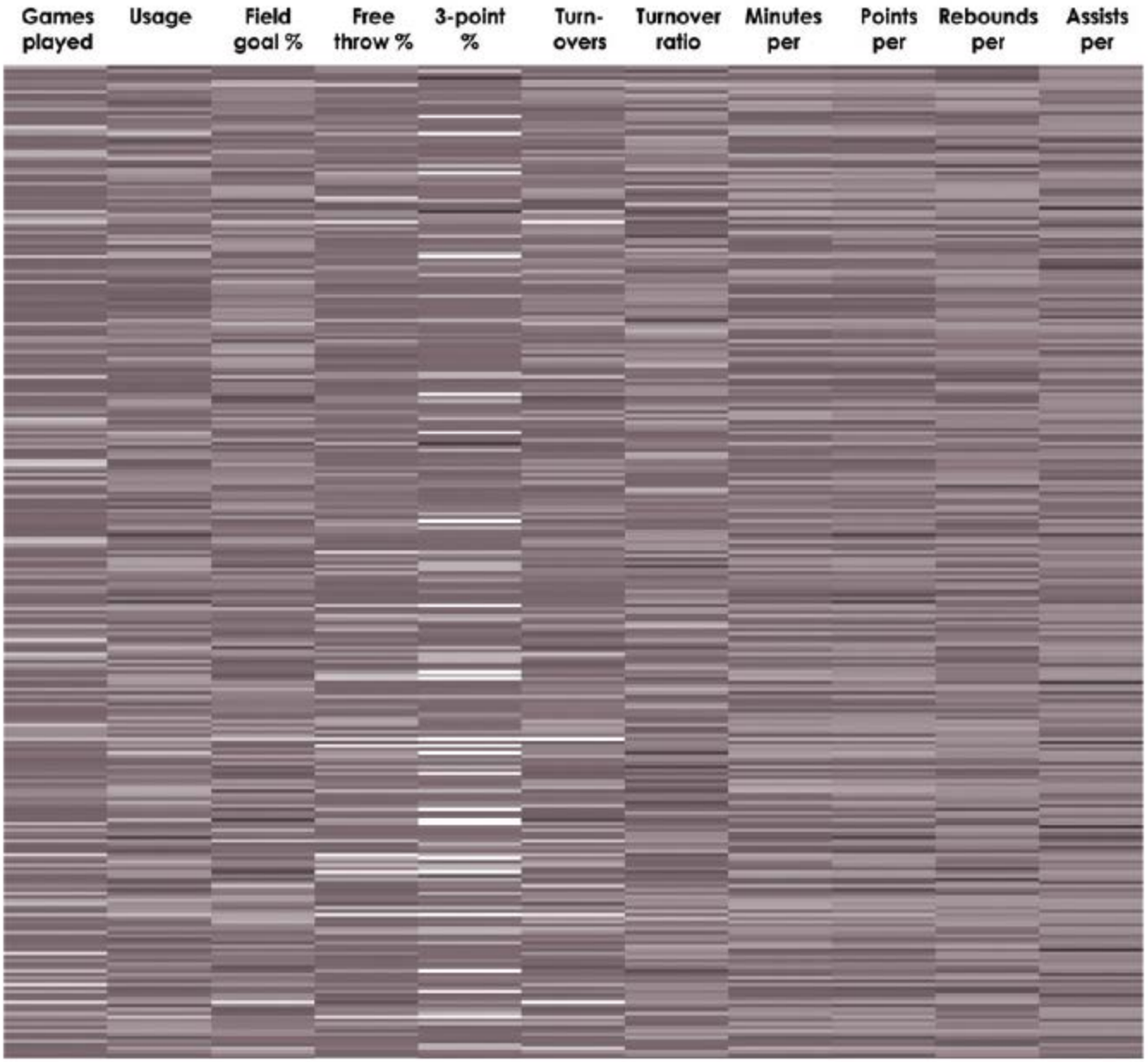
MULTIPLE VARIABLES


Points per game
30
25
20
15
10
5
0
0%
10%
20%
30%
40%
Usage

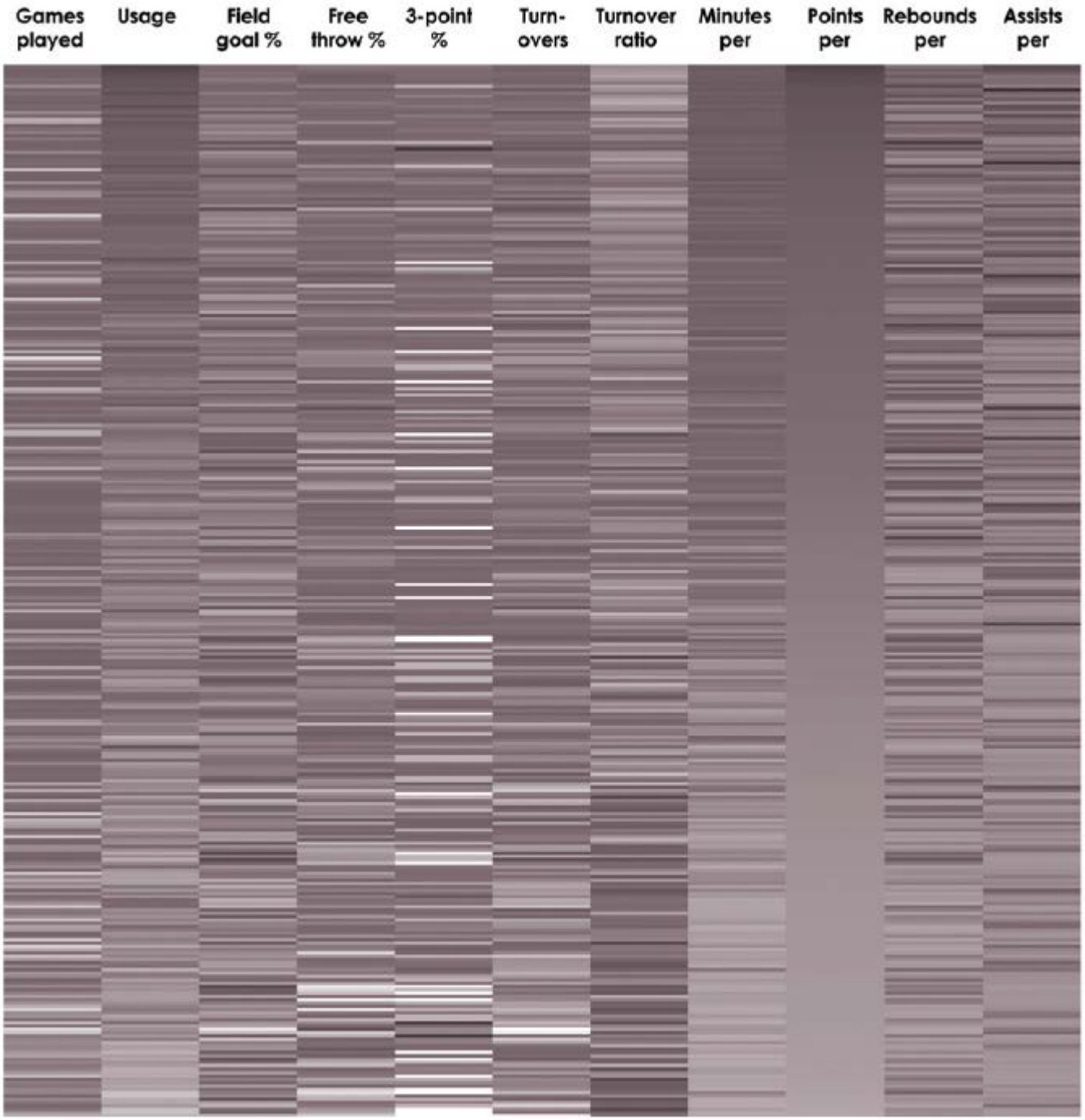

Points per game
30
25
20
15
10
5
0
0%
10%
20%
30%
40%
Usage

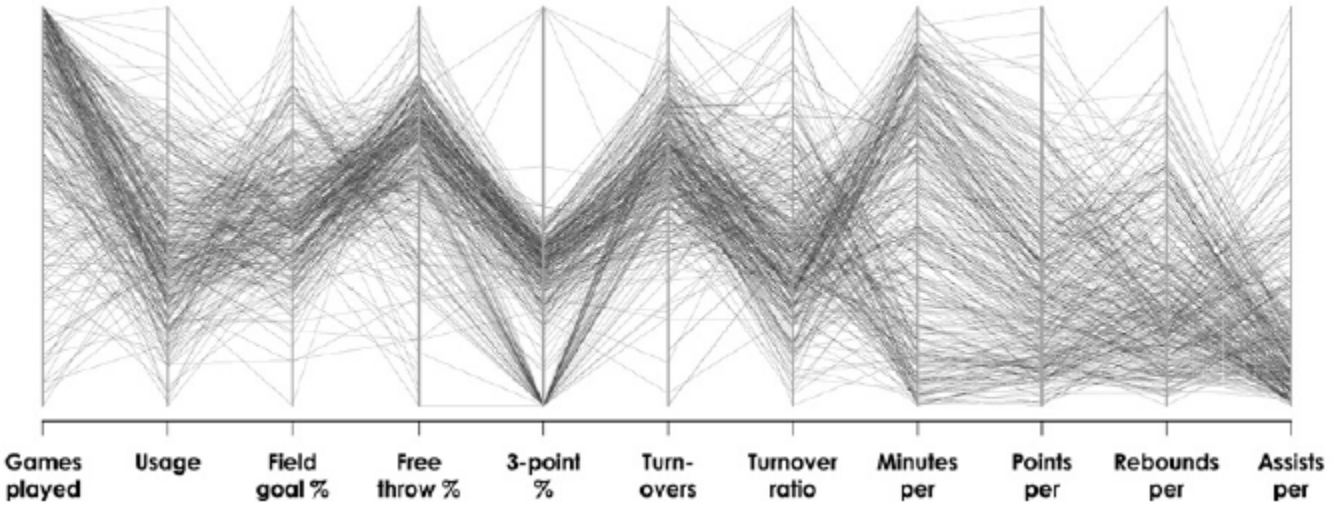


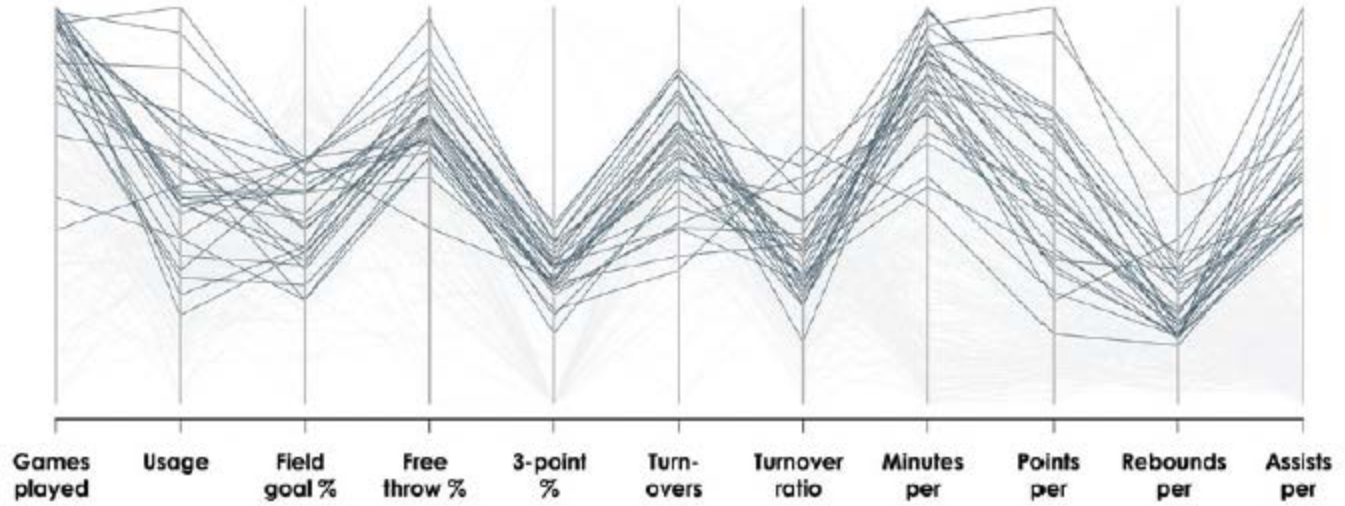
Nathan Yau, Data Points



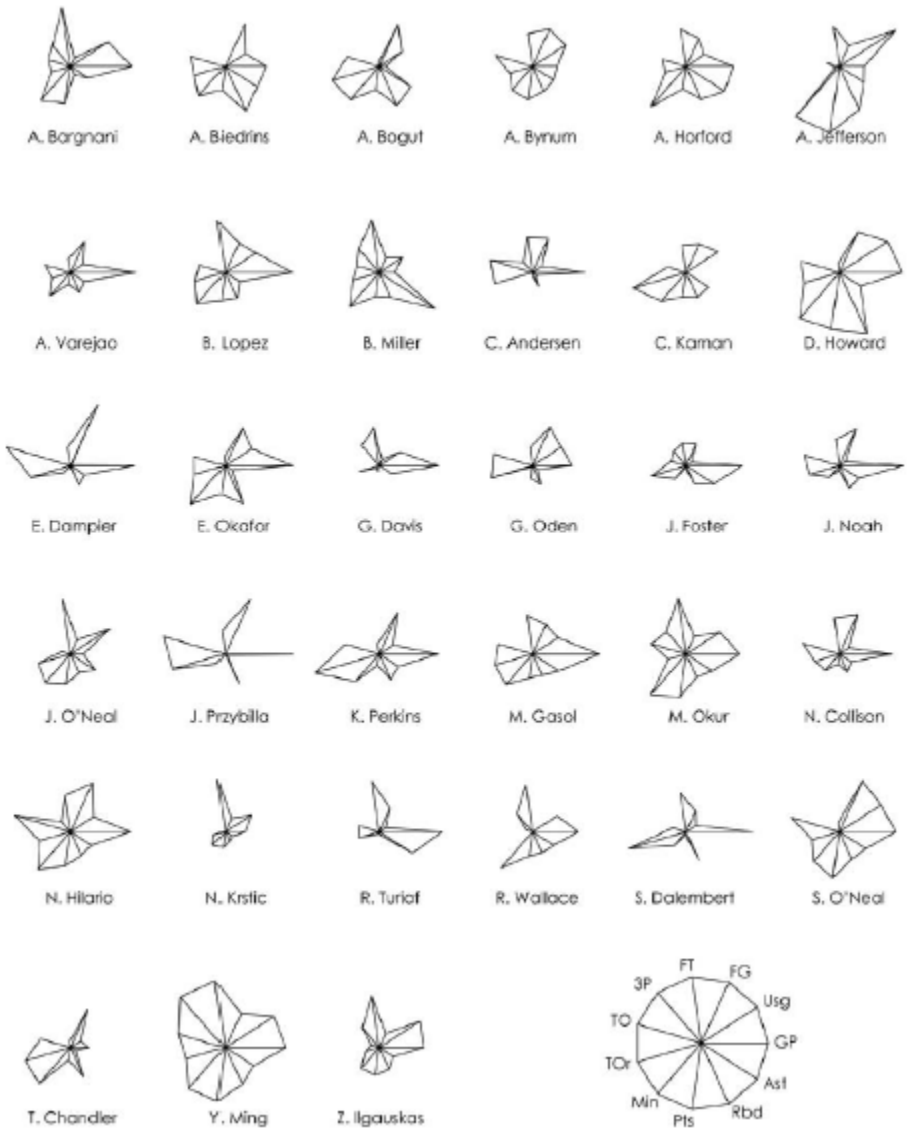
Nathan Yau, Data Points


Games played
Usage
Field goal %
Free throw %
3-point %
Turn-overs
Turnover ratio
Minutes per
Points per
Rebounds per
Assists per

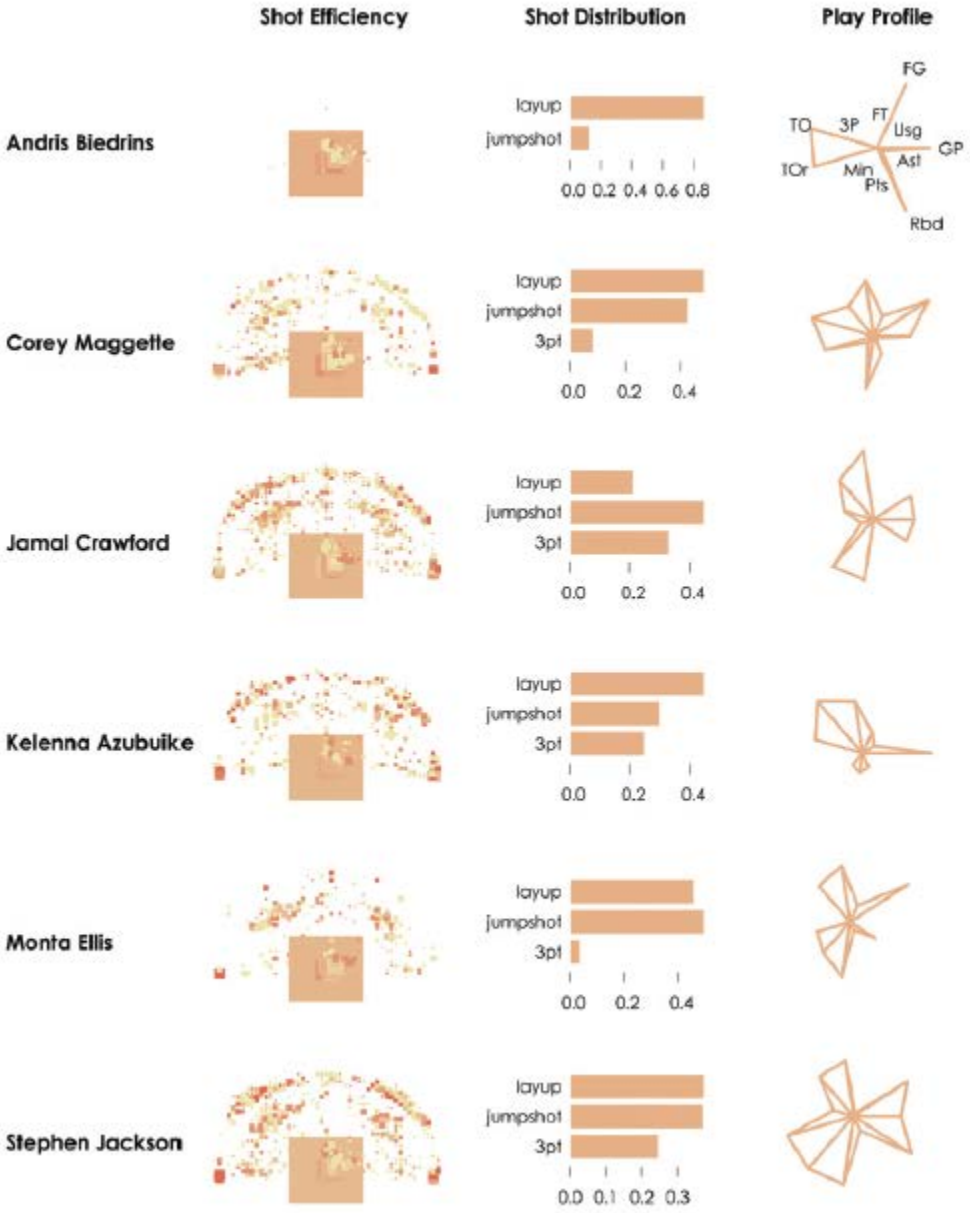




Nathan Yau, Data Points



Nathan Yau, Data Points



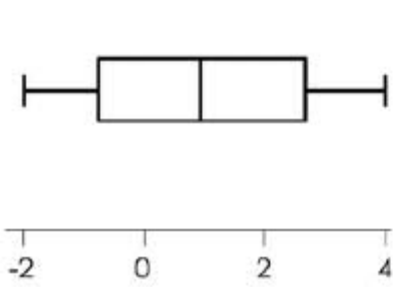
Nathan Yau, Data Points

DISTRIBUTIONS

Distribution Summary

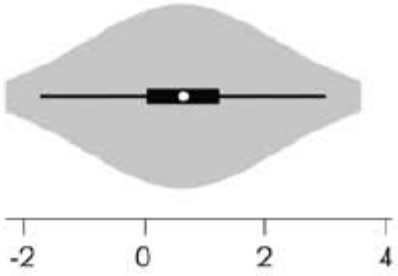
You can visualize data at different granularities with the charts above. These show key values for a less specific view of distributions.

Box plot



Shows range, median and quartiles

Violin plot

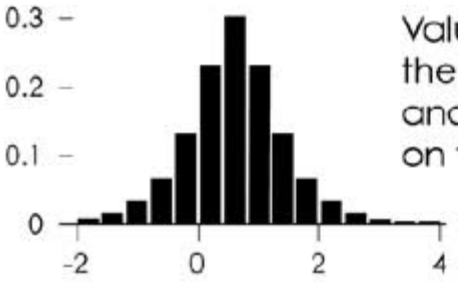


A combination of a box plot and density plot

Distribution of one variable

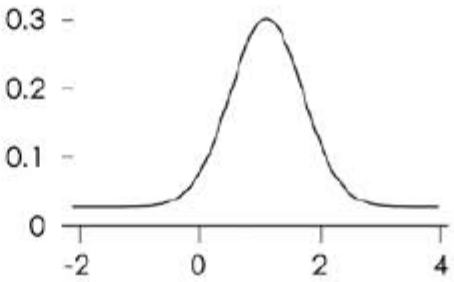
You can see where data is clustered and see any outliers by keeping track of where they sit on a value axis.

Histogram



Value axis on the horizontal and density on the vertical

Density plot

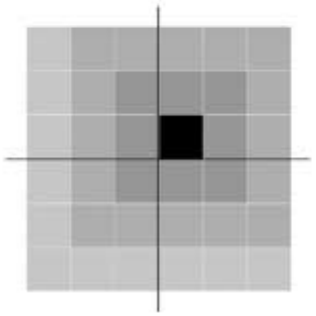


Like histogram but continuous instead of bins

Distribution of multiple variables

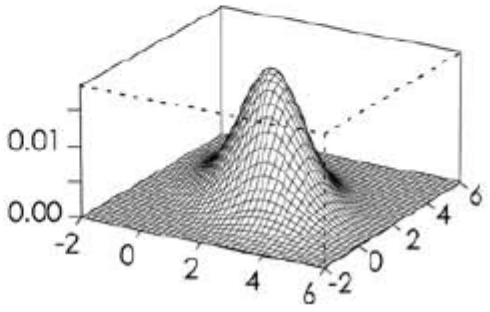
Sometimes values come as pairs, and it makes sense to show both values at the same time.

Heat map



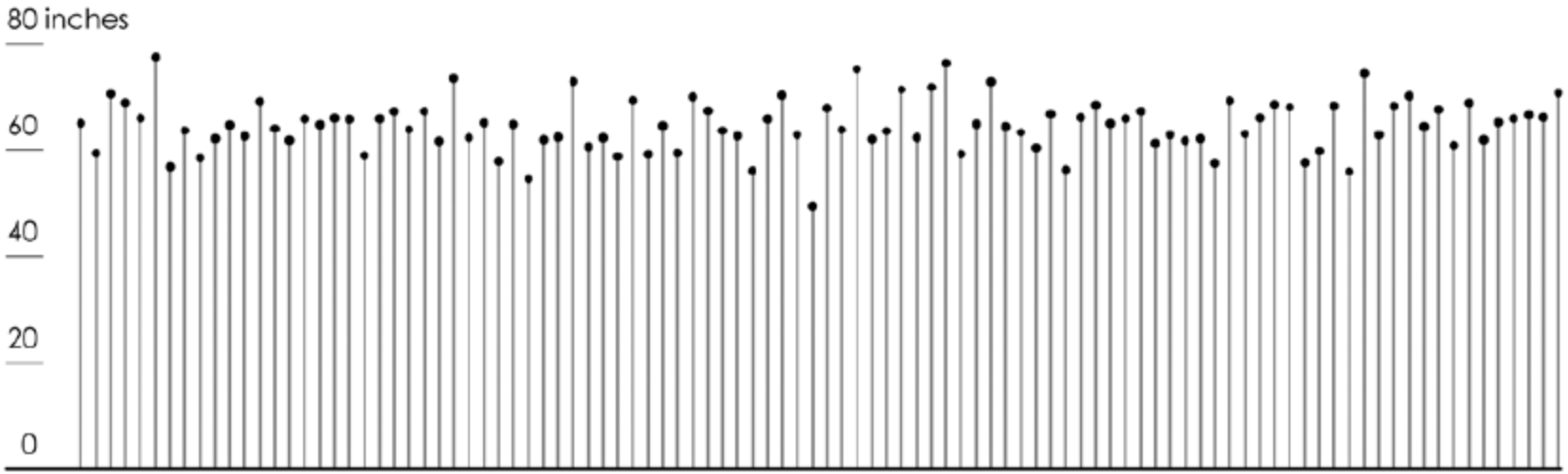
Density on a a 2-D plane, using color as visual cue

Surface plot

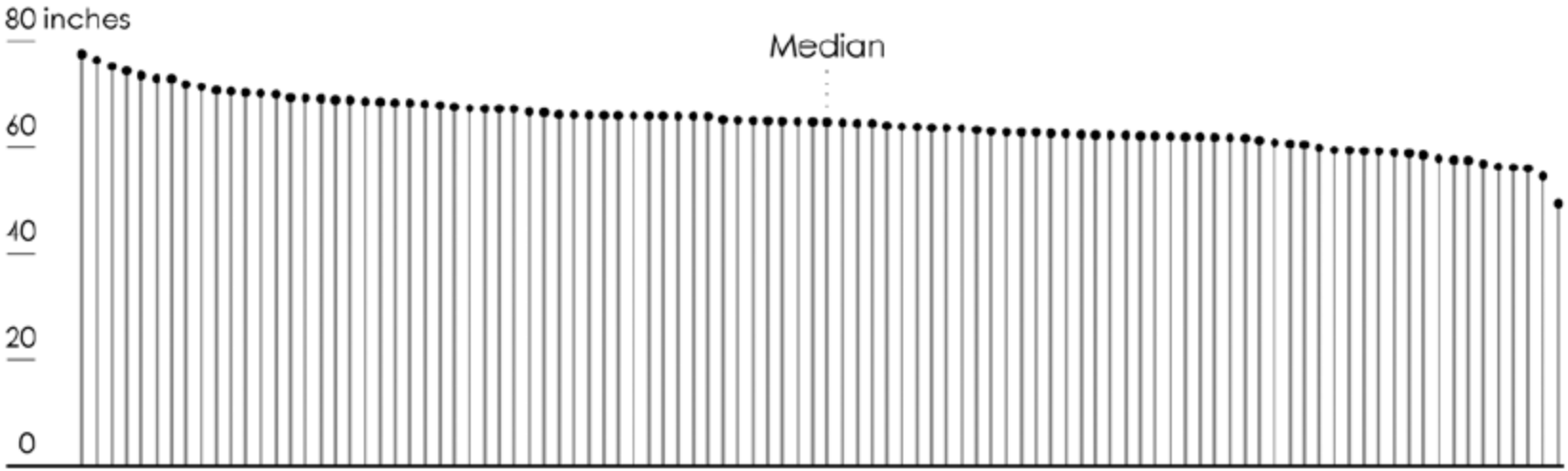


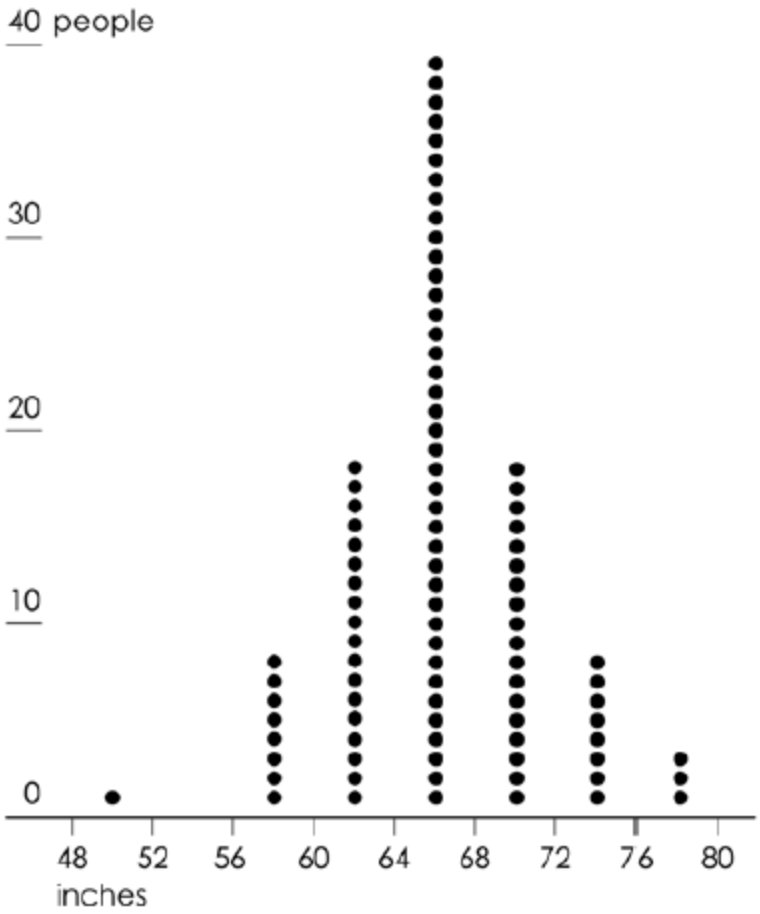
Shows same patterns as heat map, but uses height instead of color

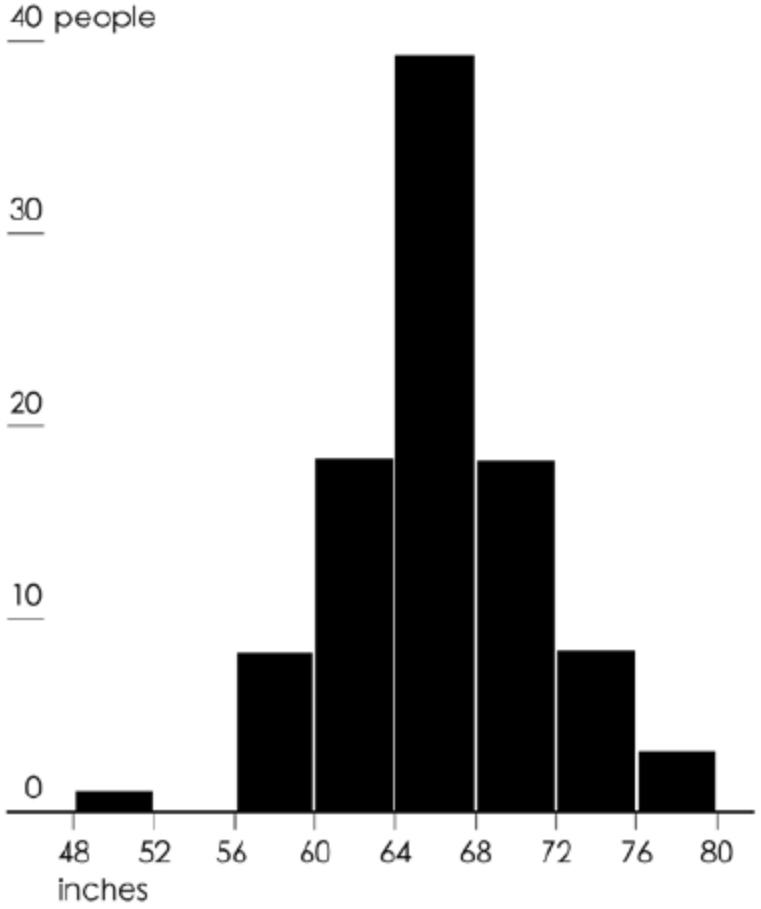
100 IMAGINARY PEOPLE

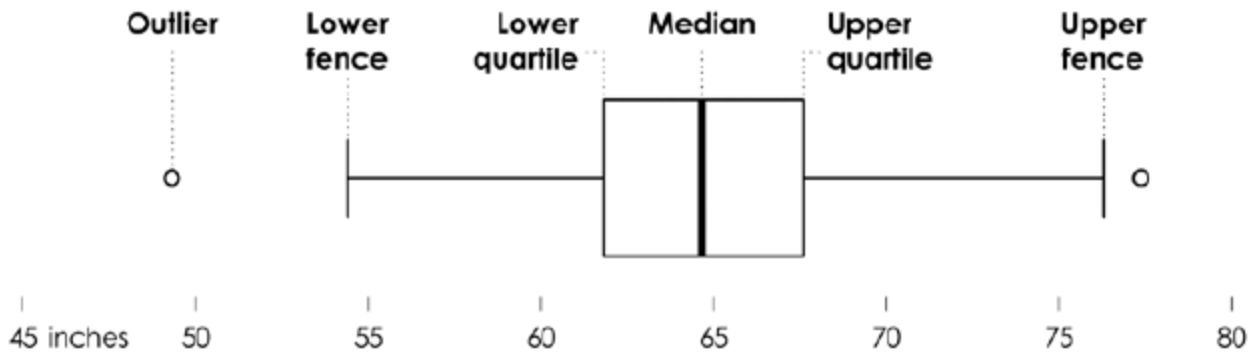


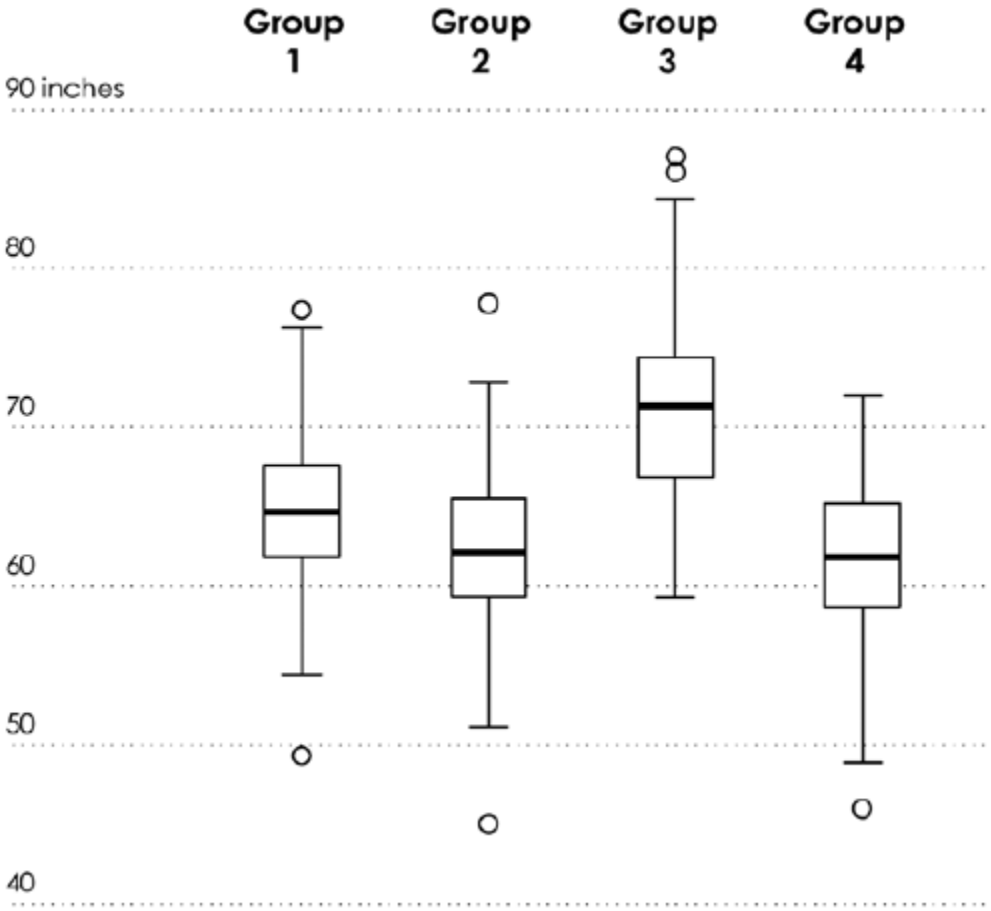
Nathan Yau, Data Points


80 inches
60
40
20
0
Median


40 people
30
20
10
0
48
52
56
60
64
68
72
76
80
inches

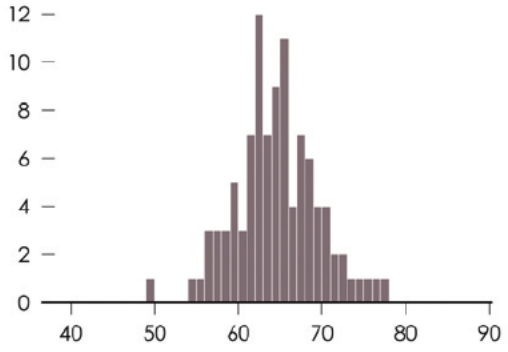

40 people
30
20
10
0
48
52
56
60
64
68
72
76
80
inches


Outlier
Lower fence
Lower quartile
Median
Upper quartile
Upper fence
45 inches
50
55
60
65
70
75
80


Group 1
Group 2
Group 3
Group 4
90 inches
80
70
60
50
40

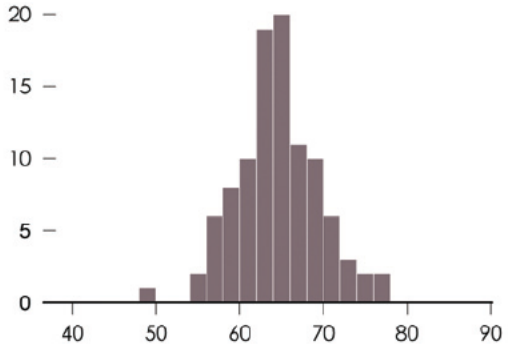
One-inch bins

Small bins shows variations at higher granularity.



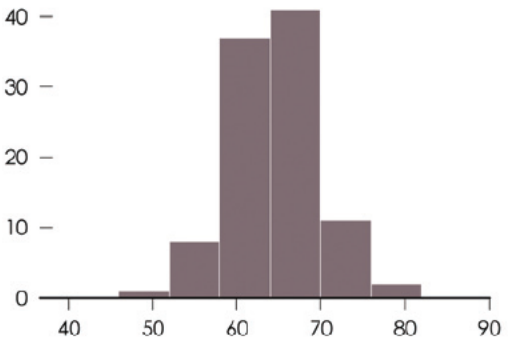
Two-inch bins

You see less variation, but the distribution around the median is more obvious.



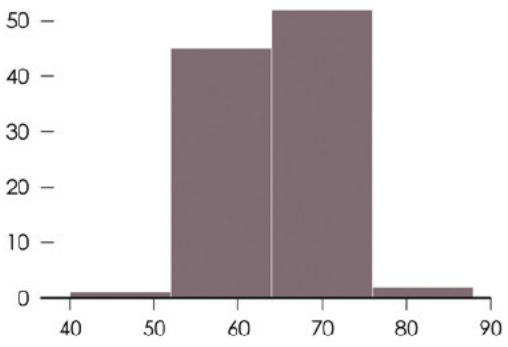
Half-foot bins

You can see distribution around the median, but you can only see some variation.



One-foot bins

The spread of the data isn't as obvious. because the larger bins show less detail.



1 | DATA

2 | WORKING PARTS

3 | GRAPH TYPES

4 | WHAT ARE WE TRYING TO SHOW?

5 | SPECIFIC EXAMPLES & EXPLANATIONS

EXAMPLES

WHAT IS A GRAPH, REALLY?

A graph is a means to an end. You need to understand the end to find the best means...

Jackie, hot air