PROFESSIONAL &
CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

I.
Theoretical:

What is Vis?
Why do it?

II.
Practical

Data Types
Datasets

Week 2
Oct 20, 2015

PROFESSIONAL &
CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

I.

Theoretical:

What is Vis?
Why do it?

II.

Practical

Data Types
Datasets

Week 2
Oct 20, 2015

What is Vis?

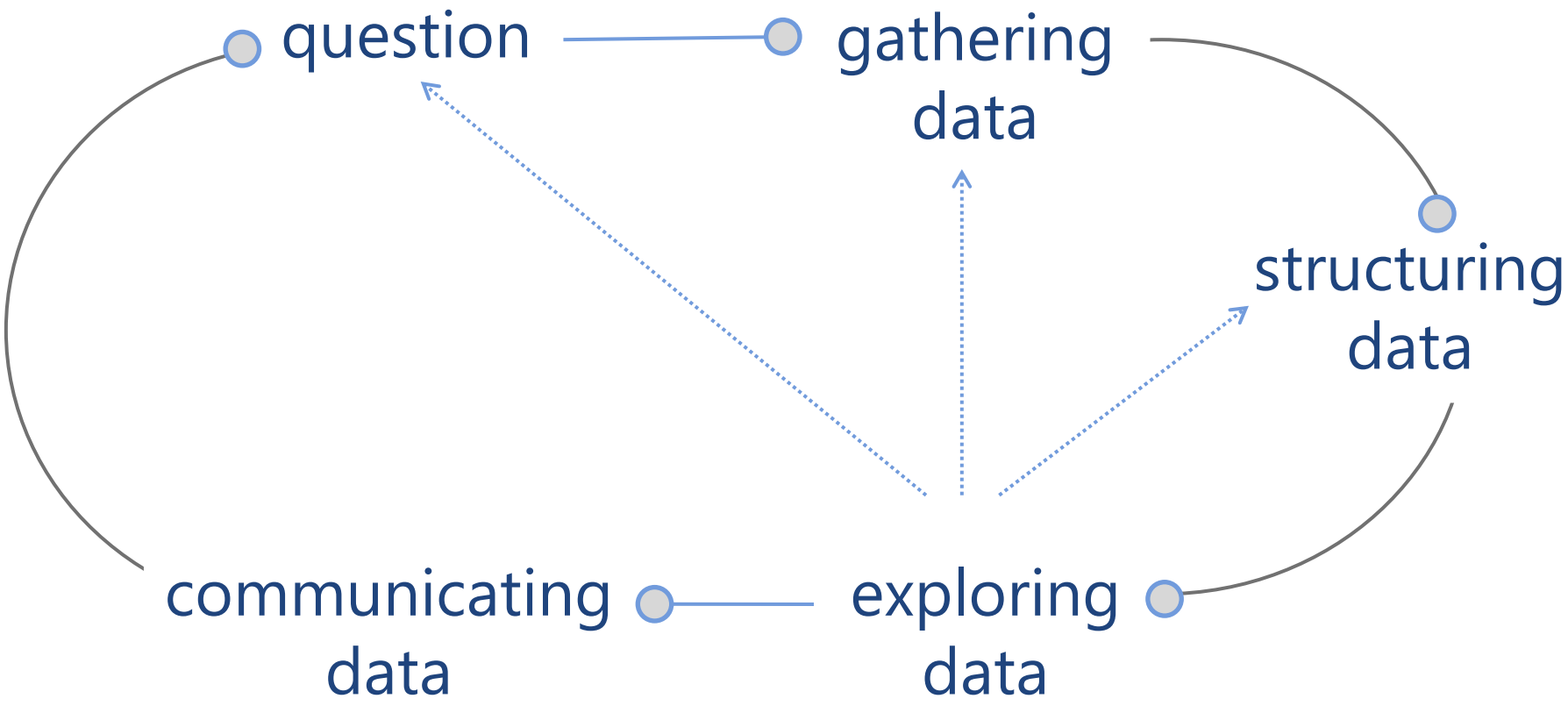
"Computer-based visualization systems provide visual representations of datasets designed to help people carry out **tasks** more effectively."

Visualization is suitable when there is a need to augment human capabilities rather than replace people with computational **decision-making** methods."

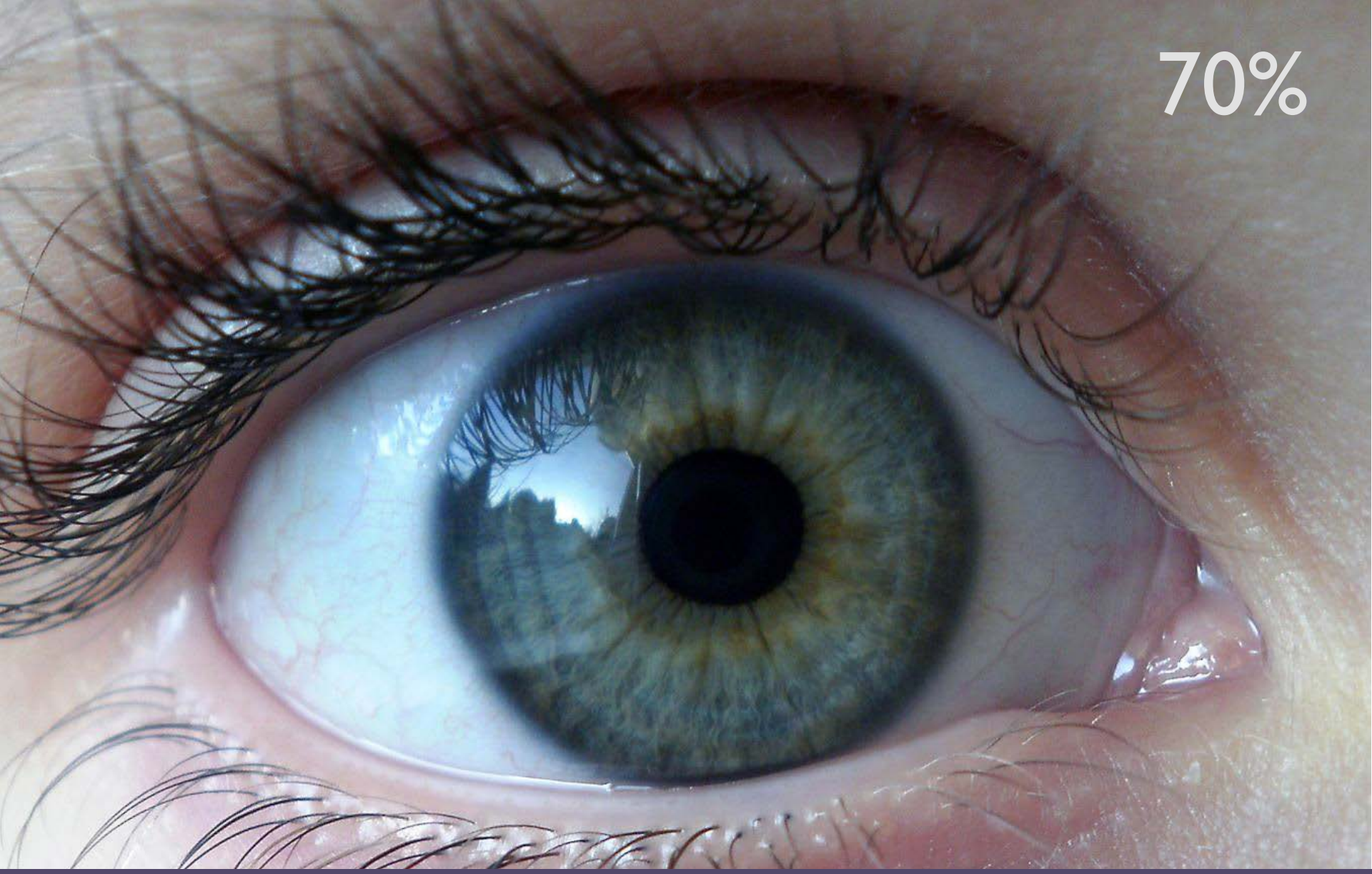
Munzner, 2014

images about numbers

The Data Discovery Horse Track



augmenting human capabilities



"The visual system provides a very high bandwidth channel to our brains." Munzner, 2014

Colin Ware on Visual Perception

"The visual system has its own rules, We can easily see patterns presented in certain ways, but if they are presented in other ways, they become invisible...When data is presented in certain ways, the patterns can be readily perceived. If we can understand how perception works, our knowledge can be translated into rules for displaying information. Following perception-based rules, we can present our data in such a way that the important and informative patterns stand out. If we disobey the rules, our data will be incomprehensible or misleading."

Source: Colin Ware (2004) *Information Visualization: Perception for Design,* Second Edition. Morgan Kauffman Publishers

How many nines are there?

7	4	7	7	5	5	2	7	4	7	1	7
2	4	9	2	5	7	7	2	6	1	7	2
7	1	7	6	9	3	4	7	5	1	2	7
4	5	1	6	3	3	8	4	8	6	6	4
8	6	5	6	4	9	3	8	9	1	9	8
3	3	8	1	5	2	2	3	6	3	9	3
7	4	6	4	5	6	3	7	7	9	1	7
3	9	1	3	3	6	1	3	3	1	8	3
8	8	1	1	8	7	5	8	1	7	4	8
3	3	6	9	2	8	9	3	7	5	7	3
2	4	4	4	2	8	2	2	9	2	8	2

How many nines are there?

Anscombe’s Quartet

Can you describe these four datasets (x,y)?

I	
x	y
10.0	8.04
8.0	6.95
13.0	7.58
9.0	8.81
11.0	8.33
14.0	9.96
6.0	7.24
4.0	4.26
12.0	10.84
7.0	4.82
5.0	5.68

II	
x	y
10.0	9.14
8.0	8.14
13.0	8.74
9.0	8.77
11.0	9.26
14.0	8.10
6.0	6.13
4.0	3.10
12.0	9.13
7.0	7.26
5.0	4.74

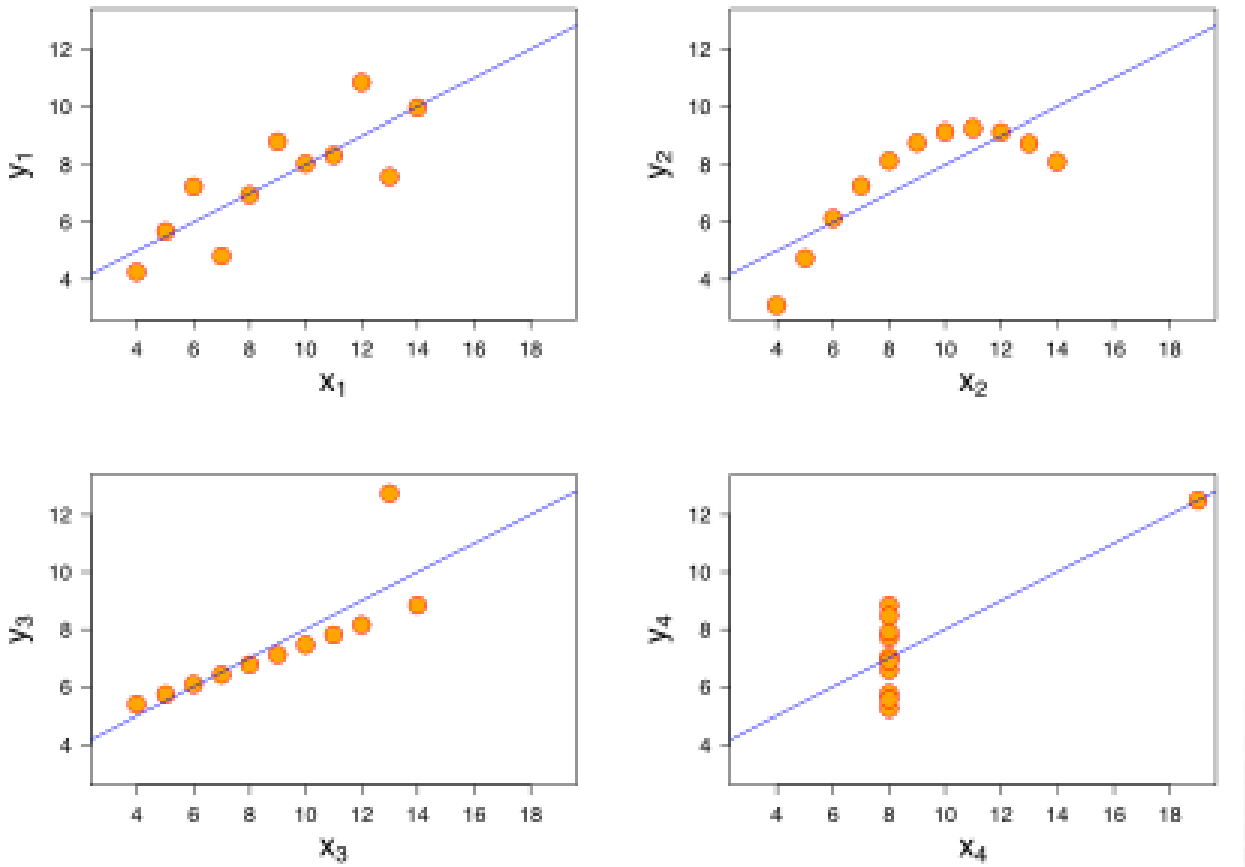
III	
x	y
10.0	7.46
8.0	6.77
13.0	12.74
9.0	7.11
11.0	7.81
14.0	8.84
6.0	6.08
4.0	5.39
12.0	8.15
7.0	6.42
5.0	5.73

IV	
x	y
8.0	6.58
8.0	5.76
8.0	7.71
8.0	8.84
8.0	8.47
8.0	7.04
8.0	5.25
19.0	12.50
8.0	5.56
8.0	7.91
8.0	6.89

Statistics to the rescue?

Property	Value
Mean of x in each case	9 (exact)
Sample variance of x in each case	11 (exact)
Mean of y in each case	7.50 (to 2 decimal places)
Sample variance of y in each case	4.122 or 4.127 (to 3 decimal places)
Correlation between x and y in each case	0.816 (to 3 decimal places)
Linear regression line in each case	$y = 3.00 + 0.500x$ (to 2 and 3 decimal places, respectively)

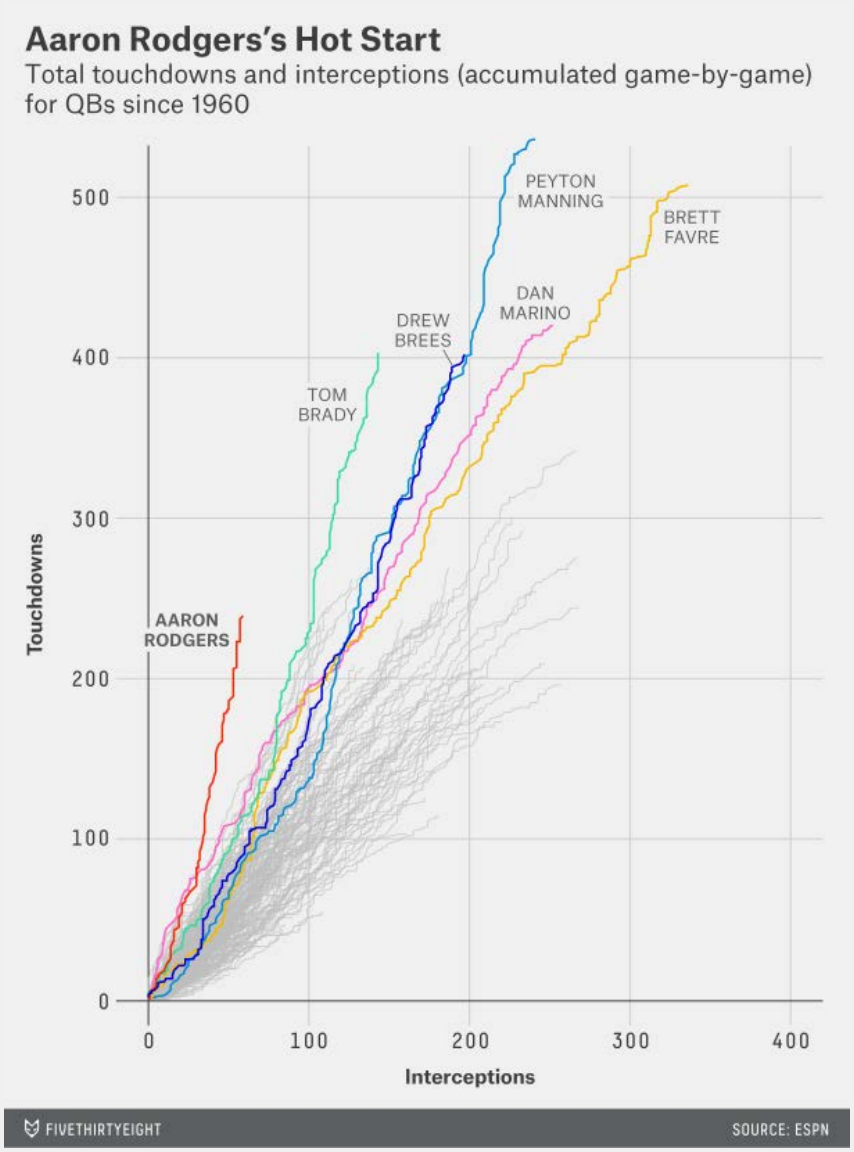
How about visualizing them?



Francis Anscombe

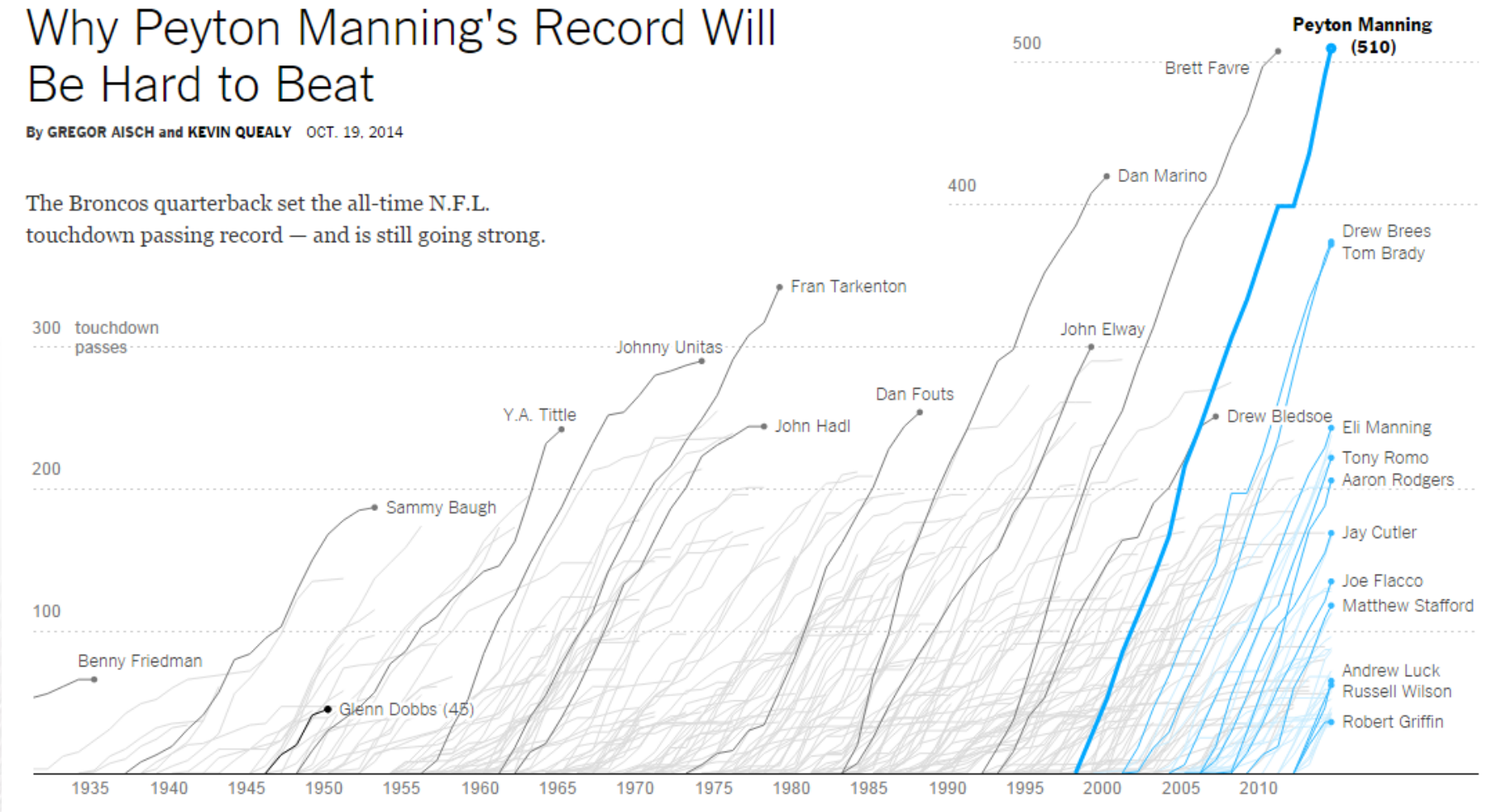
Static vs. Interactive

Static



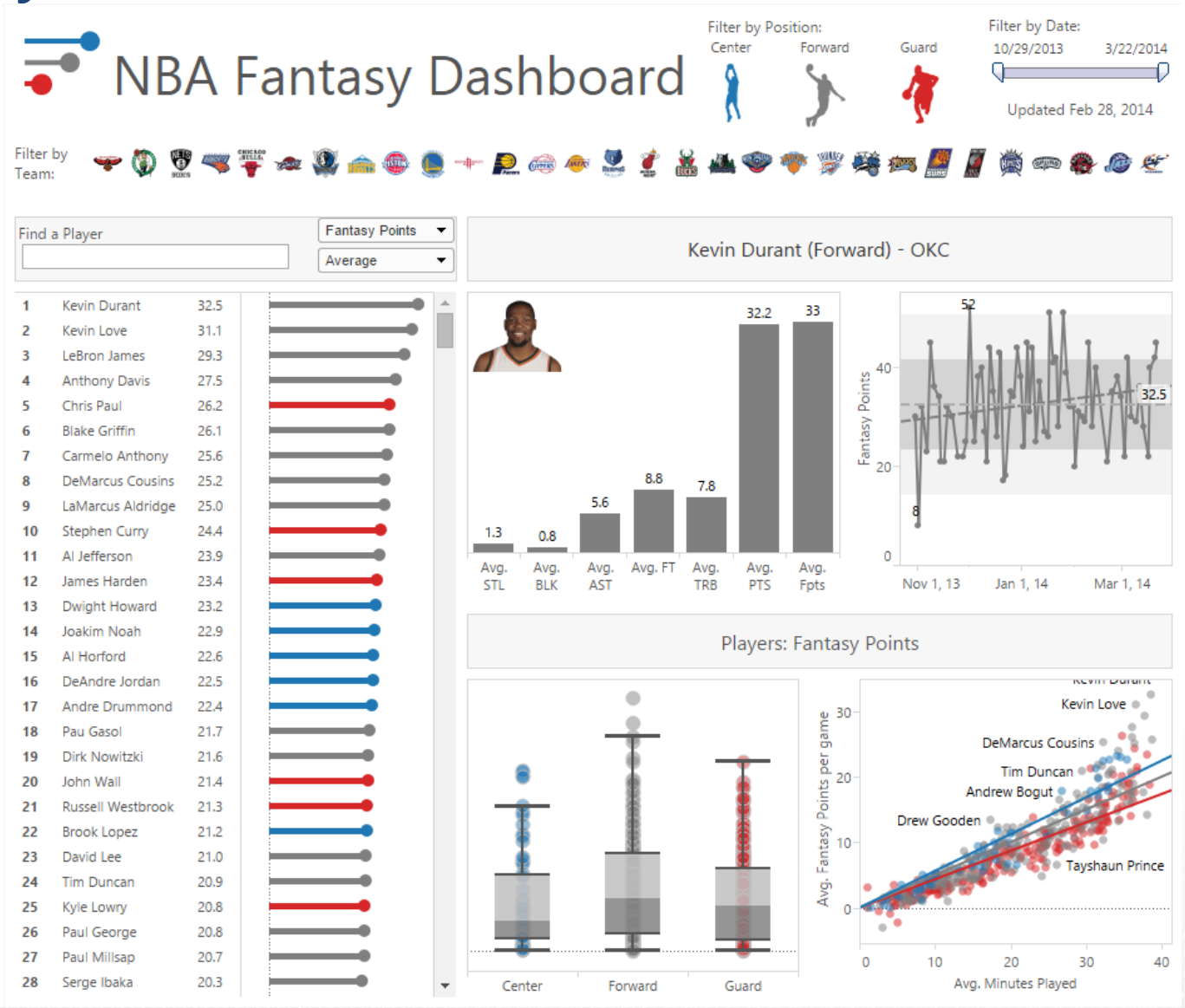
http://fivethirtyeight.com/datalab/skeptical-football-i-love-aaron-rodgers/

Interactive



http://www.nytimes.com/interactive/2014/10/19/upshot/peyton-manning-breaks-touchdown-passing-record.html?_r=0&abt=0002&abg=0

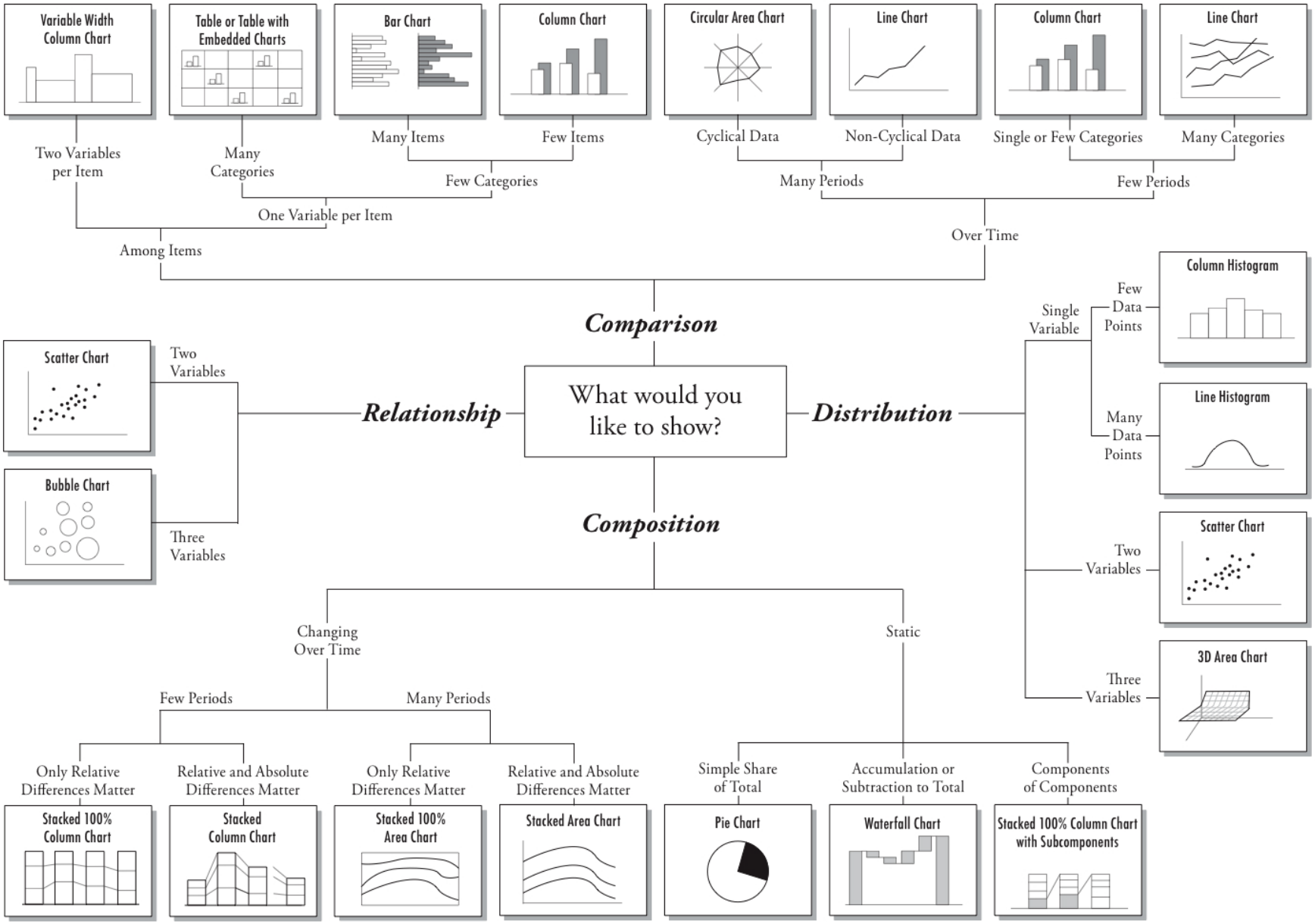
Richly Interactive



http://public.tableau.com/profile/ben.jones/#!/vizhome/FantasyNBA/FantasyNBA

The Design Space

Chart Suggestions—A Thought-Starter

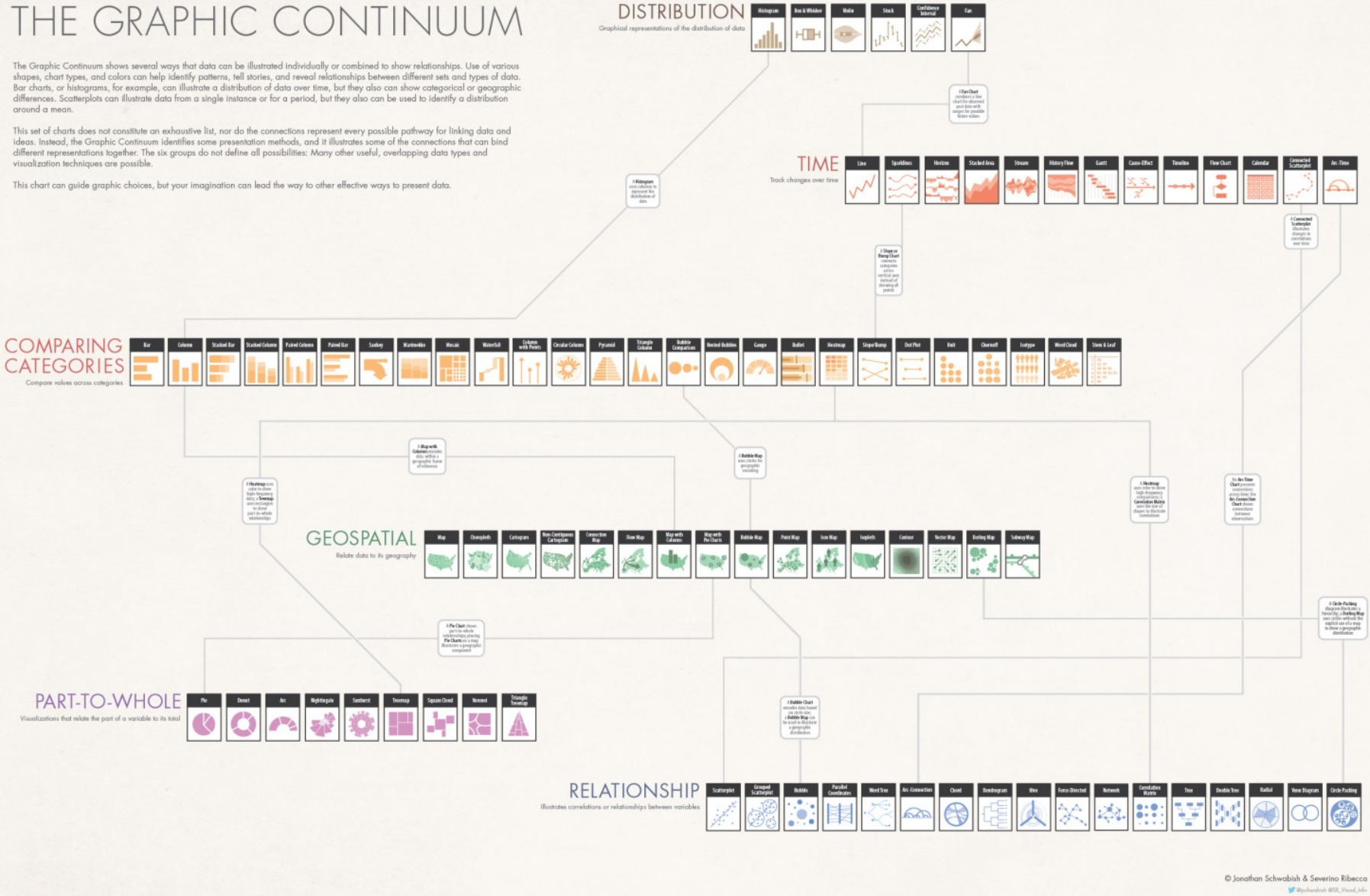


THE GRAPHIC CONTINUUM

The Graphic Continuum shows several ways that data can be illustrated individually or combined to show relationships. Use of various shapes, chart types, and colors can help identify patterns, tell stories, and reveal relationships between different sets and types of data. Bar charts, or histograms, for example, can illustrate a distribution of data over time, but they also can show categorical or geographic differences. Scatterplots can illustrate data from a single instance or for a period, but they also can be used to identify a distribution around a mean.

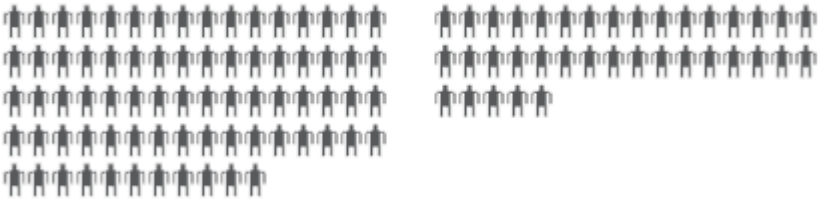
This set of charts does not constitute an exhaustive list, nor do the connections represent every possible pathway for linking data and ideas. Instead, the Graphic Continuum identifies some presentation methods, and it illustrates some of the connections that can bind different representations together. The six groups do not define all possibilities: Many other useful, overlapping data types and visualization techniques are possible.

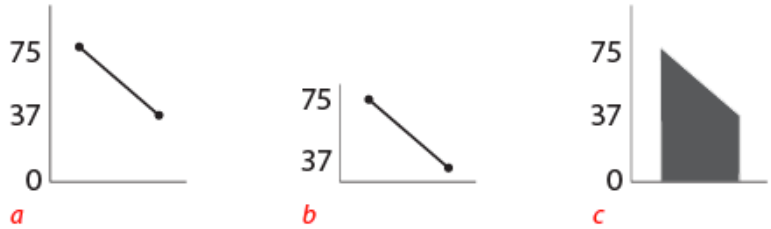
This chart can guide graphic choices, but your imagination can lead the way to other effective ways to present data.

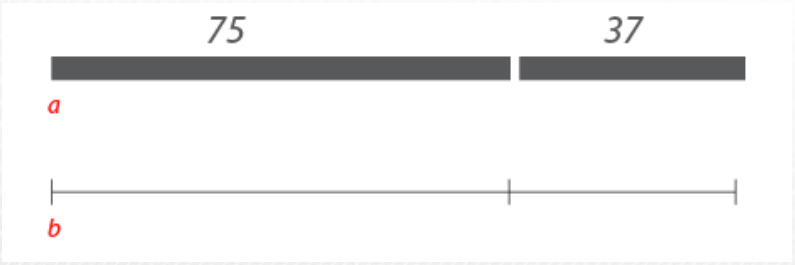


@jschwabish @SR_Visual_Info

“45 Ways to Communicate Two Quantities”





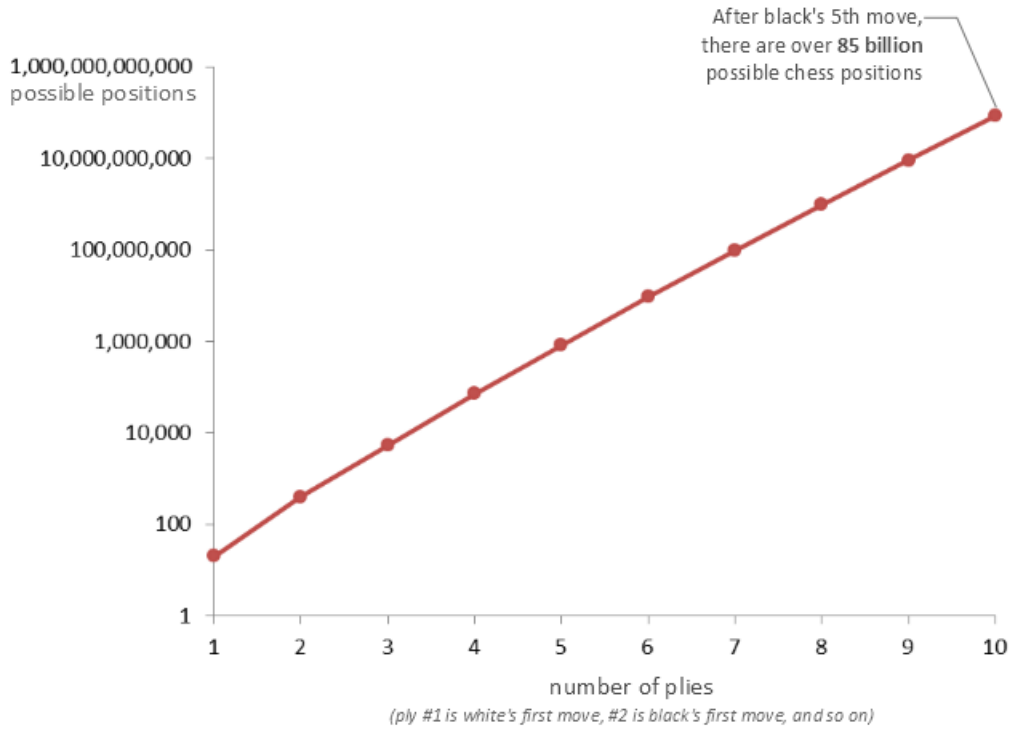


http://blog.visual.ly/45-ways-to-communicate-two-quantities/

The Design Space

Chess Positions Galore

François Labelle has calculated the number of distinct chess positions by ply, or move. The number increases dramatically with each successive move as the below line chart (with logarithmic y-axis) depicts. After just 5 moves each (or, by ply #10), the possibilities have reached staggering proportions



Source: http://wismuth.com/chess/statistics-positions.html

Tasks, Effectiveness, and Validation

MY FANTASY HOOPS DRAFT ASSISTANT

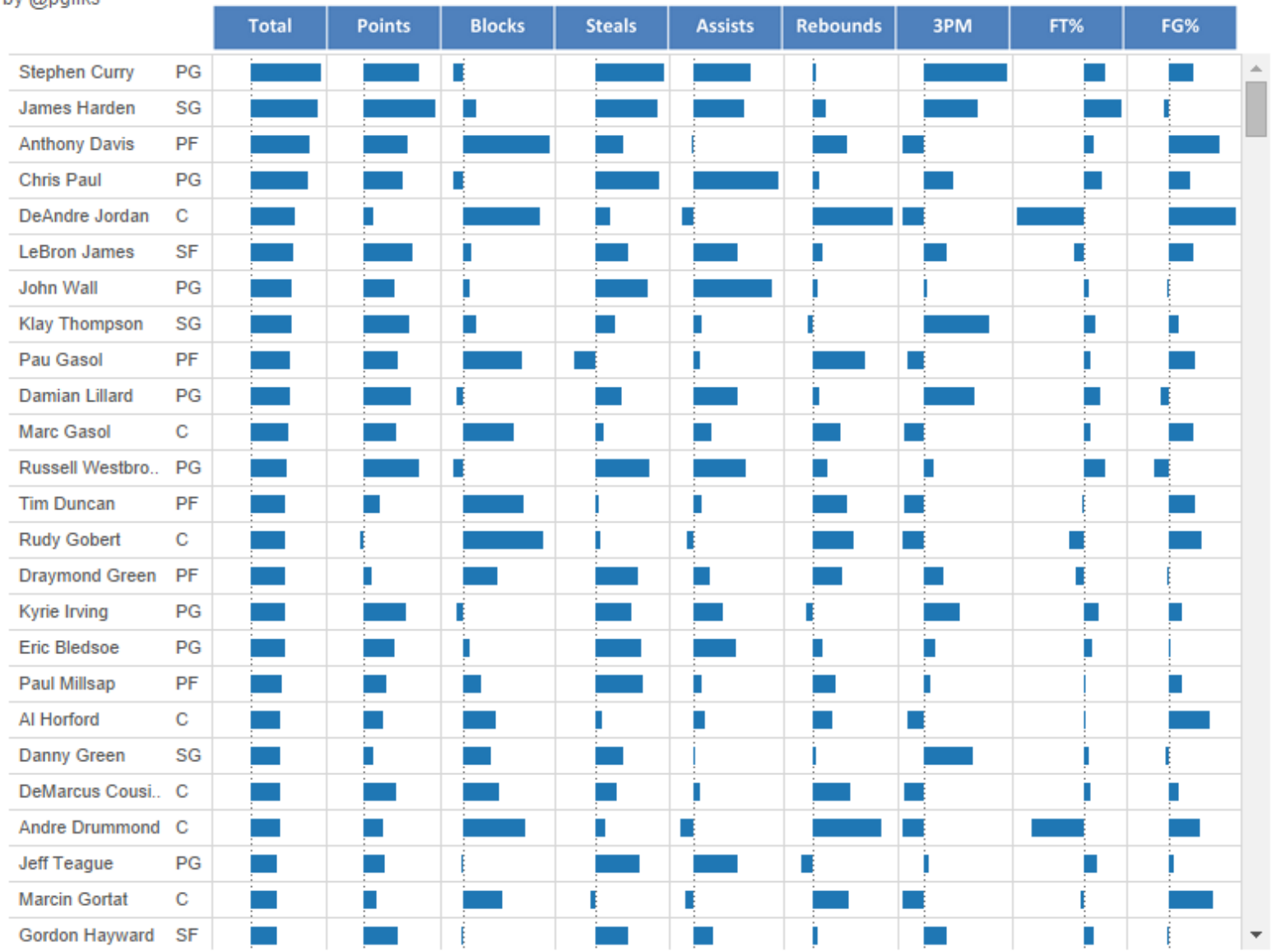
by @pgilks

Unselected Player

How does this work?

Sort players by: Total

View players by a position: All



Player 1: Select Player

Player 2: Select Player

Player 3: Select Player

Player 4: Select Player

Player 5: Select Player

Player 6: Select Player

Player 7: Select Player

Player 8: Select Player

Player 9: Select Player

Player 10: Select Player

Player 11: Select Player

Player 12: Select Player

Hiking the Washington Trails

Bellingham

Mount Constitution

click a tile to see a detailed trail map

Everett

Shoreline

Redmond

Seattle

Renton

Kent

Auburn

Tacoma

Lakewood

Olympia

Denny Creek

© OpenStreetMap contributors

MOUNT CONSTITUTION

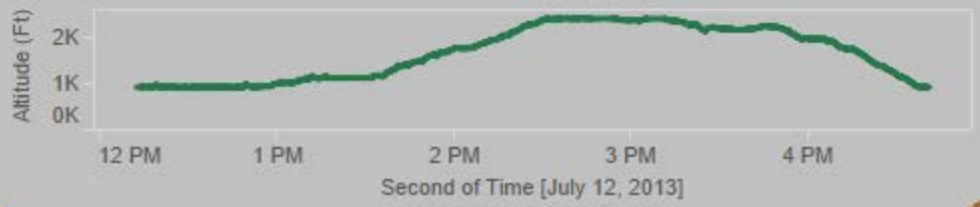
Jul 12, 2013

04:29:04 | 9.39 mi | 1,489 ft

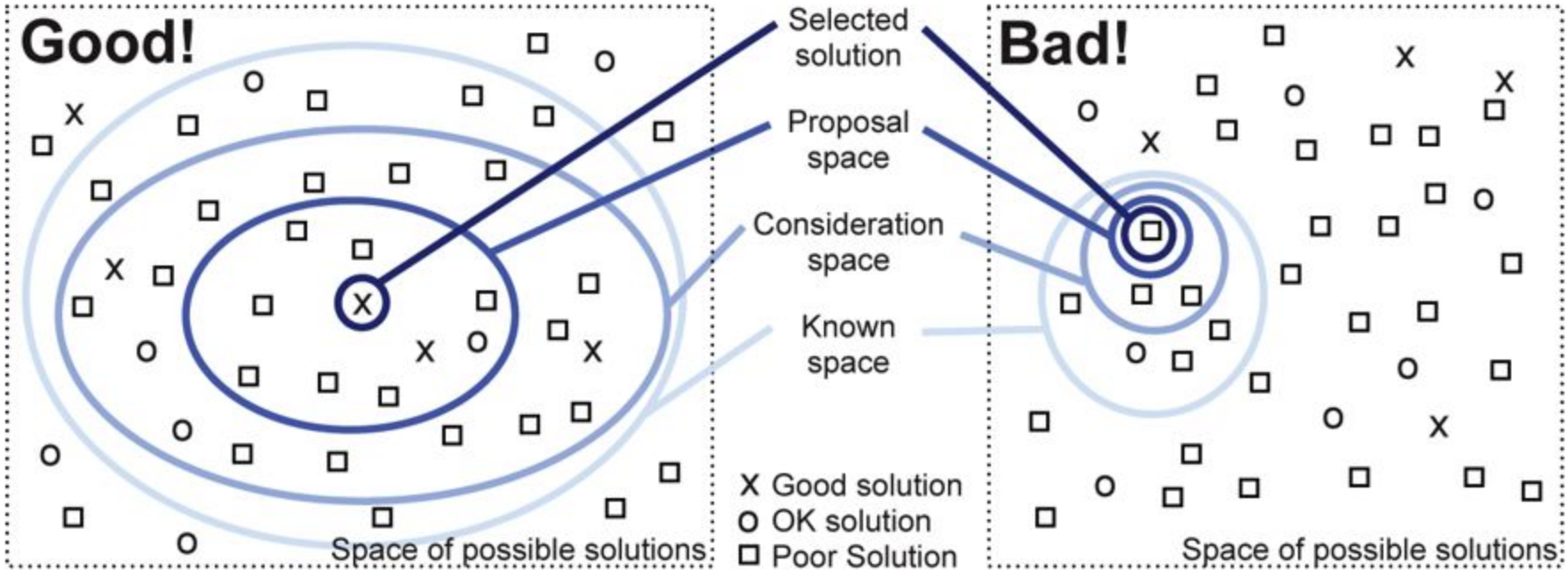
"Simon leading the way"

DISTANCE

ALTITUDE



How to find an effective design



6 Principles of Communicating Data

1 know your **goal**

- [ ] Who? identify your target audience
- [ ] What? articulate your intended message
- [ ] Why? define the desired effect

2 use the right **data**

- [ ] Necessary: get data that supports your point
- [ ] Sufficient: enough data to draw conclusions
- [ ] Relevant: eliminate any extraneous data

3 select suitable **visualizations** [1]

- [ ] Quantitative: position > length > angle >area
- [ ] Ordinal: position > gray ramp > color ramp
- [ ] Nominal: position > shape > color hue

4 design for **aesthetics**

- [ ] Clean up fonts, borders, gridlines, alignment
- [ ] Add graphic elements that aid cognition
- [ ] Ensure data is not occluded or distorted

5 choose **medium** & **channel**

- [ ] Medium: static, interactive, or animated?
- [ ] Channel: standalone, recorded, remote, live
- [ ] Mode: broadcast or directed?

6 check the **results**

- [ ] Reach: did you reach your target audience?
- [ ] Understanding: did they get it?
- [ ] Impact: did they react in the desired way?

From Communicating Data With Tableau, chapter 1

[1] Ref: Tableau, http://bit.ly/great-vis

PROFESSIONAL &
CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

I.

Theoretical:

What is Vis?
Why do it?

II.

Practical

Data Types
Datasets

Week 2
Oct 20, 2015

“‘Data! Data! Data!’ he cried impatiently.
‘I can’t make bricks without clay!’”

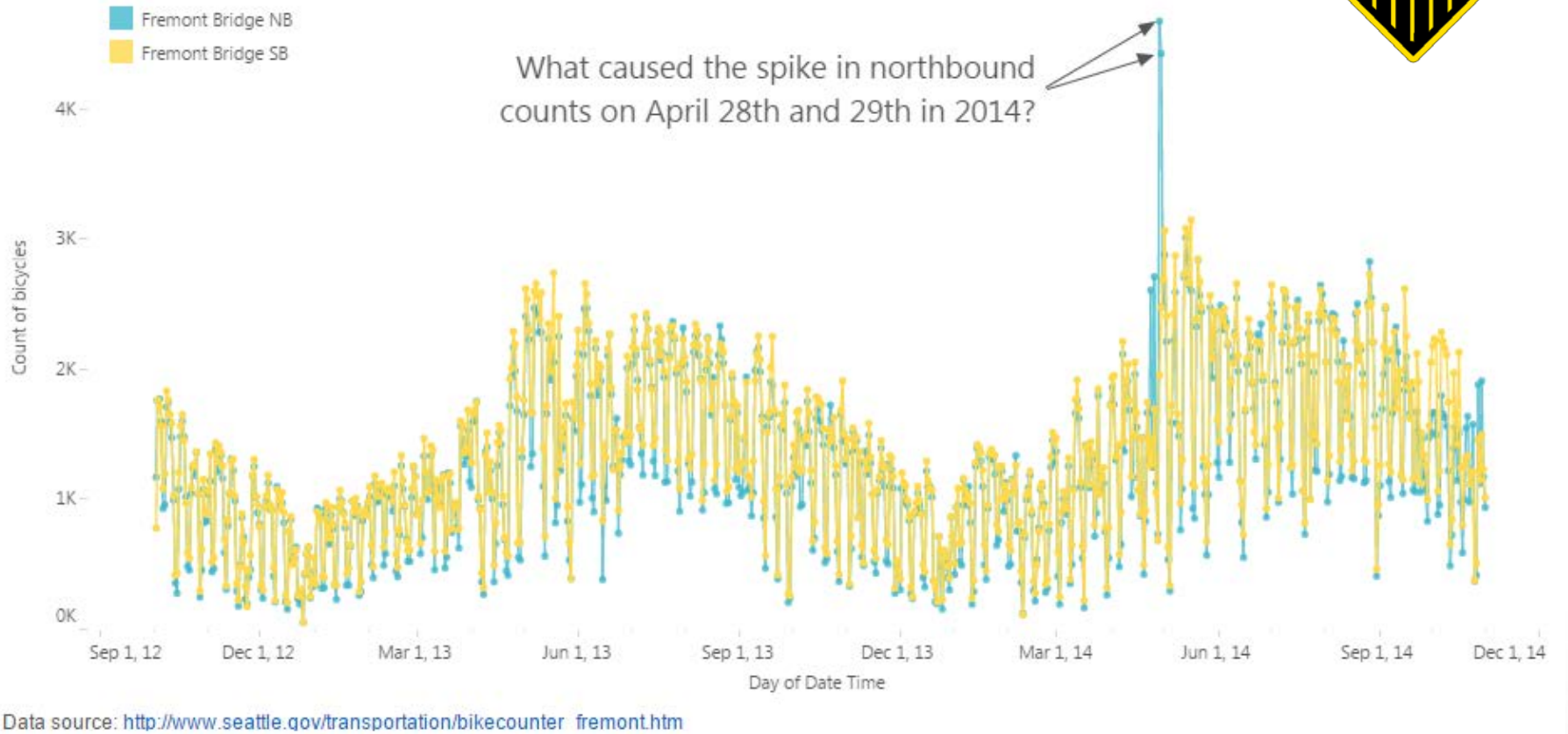
Semantics…

```
Basil      7      S      Pear
```

Semantics matter

ID	Name	Age	Shirt Size	Favorite Fruit
1	Amy	8	S	Apple
2	Basil	7	S	Pear
3	Clara	9	M	Durian
4	Desmond	13	L	Elderberry
5	Ernest	12	L	Peach
6	Fanny	10	S	Lychee
7	George	9	M	Orange
8	Hector	8	L	Loquat
9	Ida	10	M	Pear
10	Amy	12	M	Orange

Fremont Bridge Bike Counter Time Series, Oct 2012 - Oct 2014

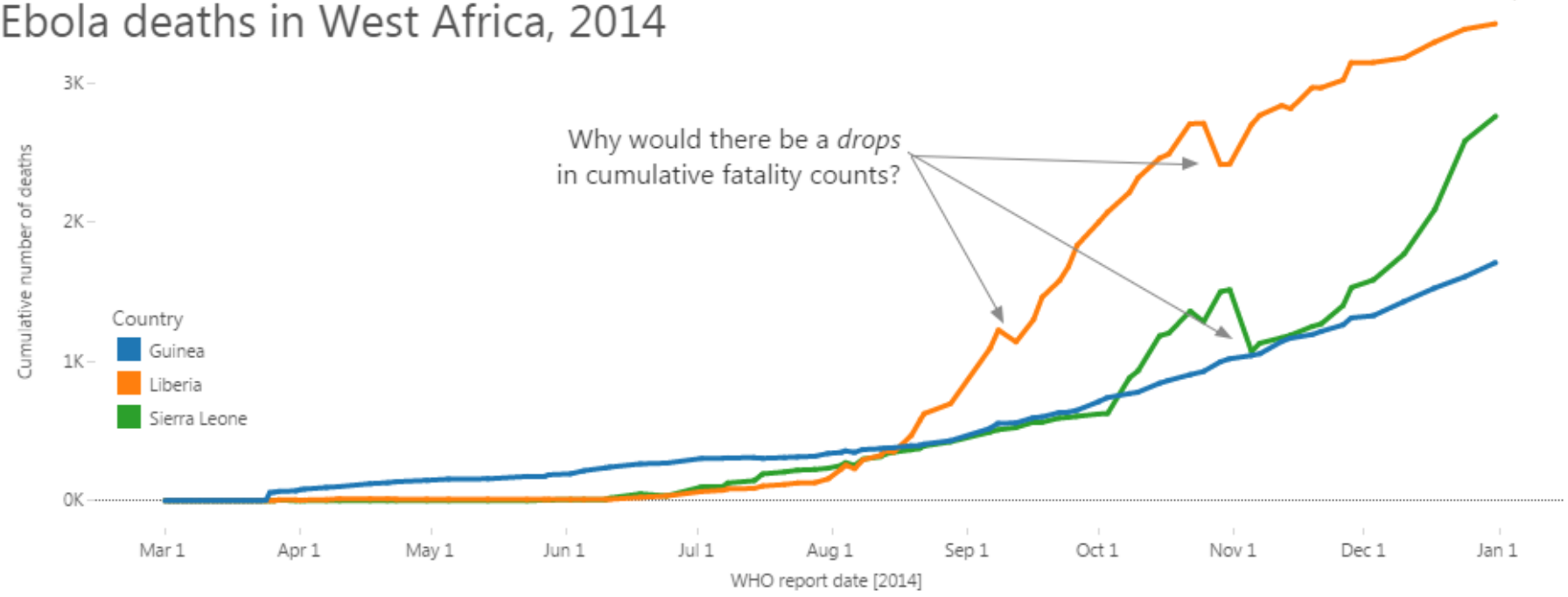


Data source: http://www.seattle.gov/transportation/bikecounter_fremont.htm

Not bicycle passes, *registered* bicycle passes


Ebola

Ebola deaths in West Africa, 2014



Data Source: http://www.cdc.gov/vhf/ebola/outbreaks/2014-west-africa/cumulative-cases-graphs.html

Not deaths, *reported* deaths

What?

Datasets

Data Types

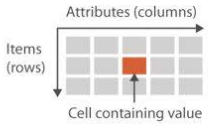
→ Items → Attributes → Links → Positions → Grids

Data and Dataset Types

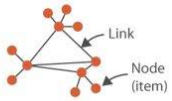
Tables	Networks & Trees	Fields	Geometry	Clusters, Sets, Lists
Items	Items (nodes)	Grids	Items	Items
Attributes	Links	Positions	Positions	
	Attributes	Attributes		

Dataset Types

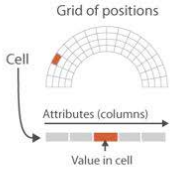
→ Tables



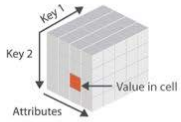
→ Networks



→ Fields (Continuous)



→ *Multidimensional Table*



→ *Trees*

→ Geometry (Spatial)

Dataset Availability

→ Static

→ Dynamic

Attributes

Attribute Types

→ Categorical

→ Ordered

→ *Ordinal*

→ *Quantitative*

Ordering Direction

→ Sequential

→ Diverging

→ Cyclic

Types of data

- Qualitative (nominal)

  Arizona, New York, Texas
  Sarah, John, Maria
  Coors, Bud Light, Stella Artois

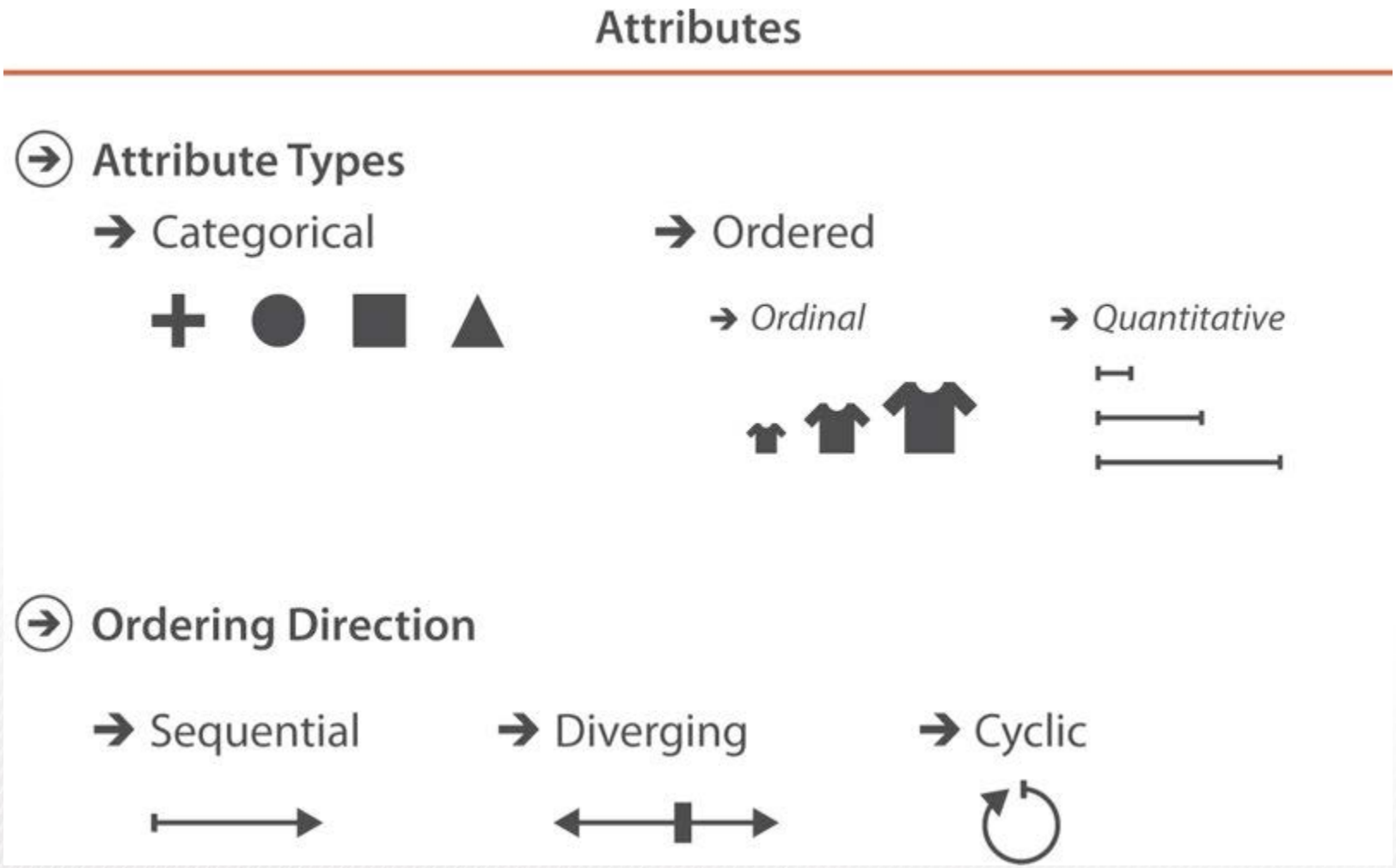
- Qualitative (ordinal)

  Gold, silver, bronze
  Excellent health, good health, poor health
  Love it, like it, hate it

- Quantitative

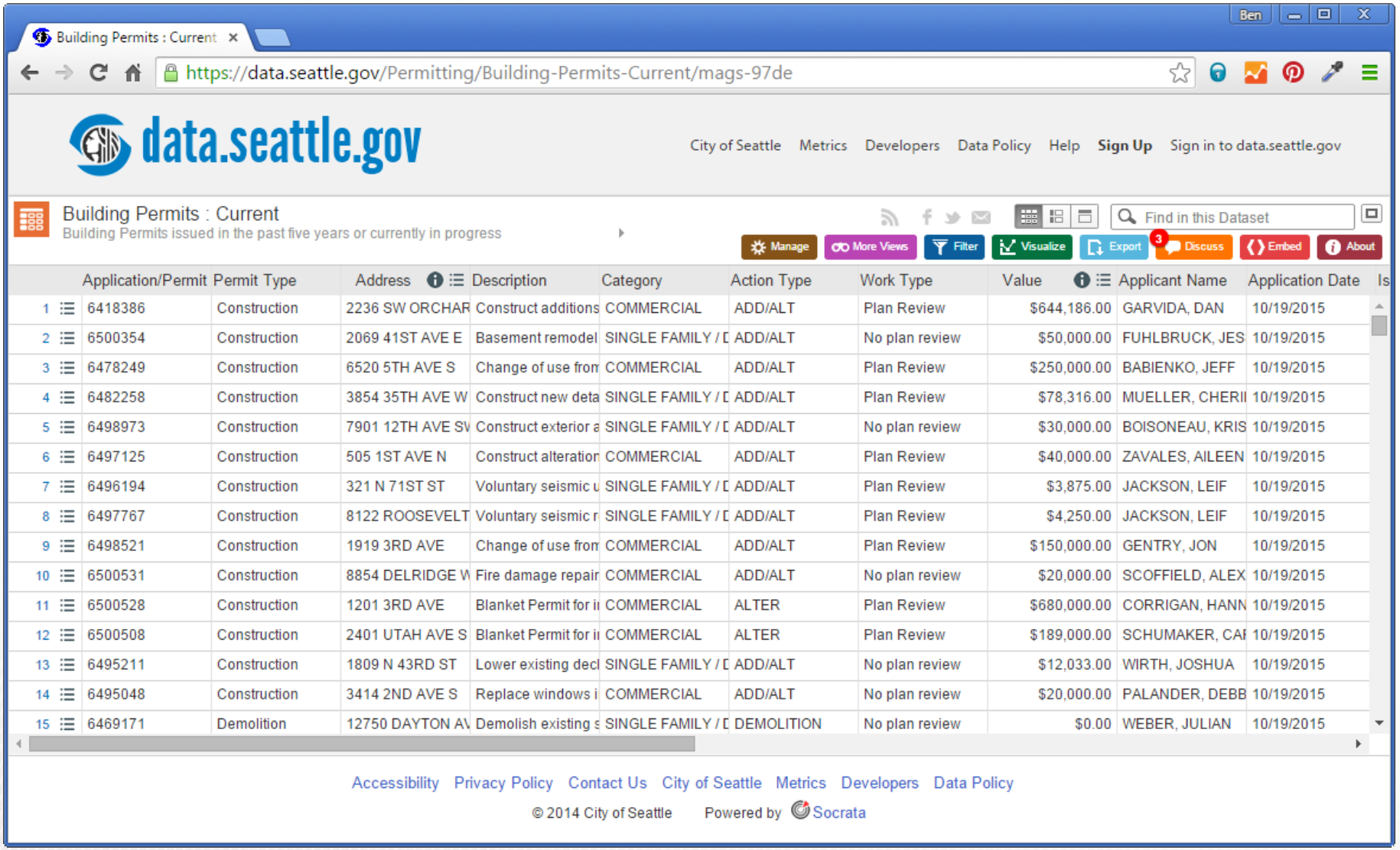
  Weight (10 lbs, 20 lbs, 5000 lbs)
  Cost ($50, $100, $0.05)
  Discount (5%, 10%, 12.8%)

Attributes

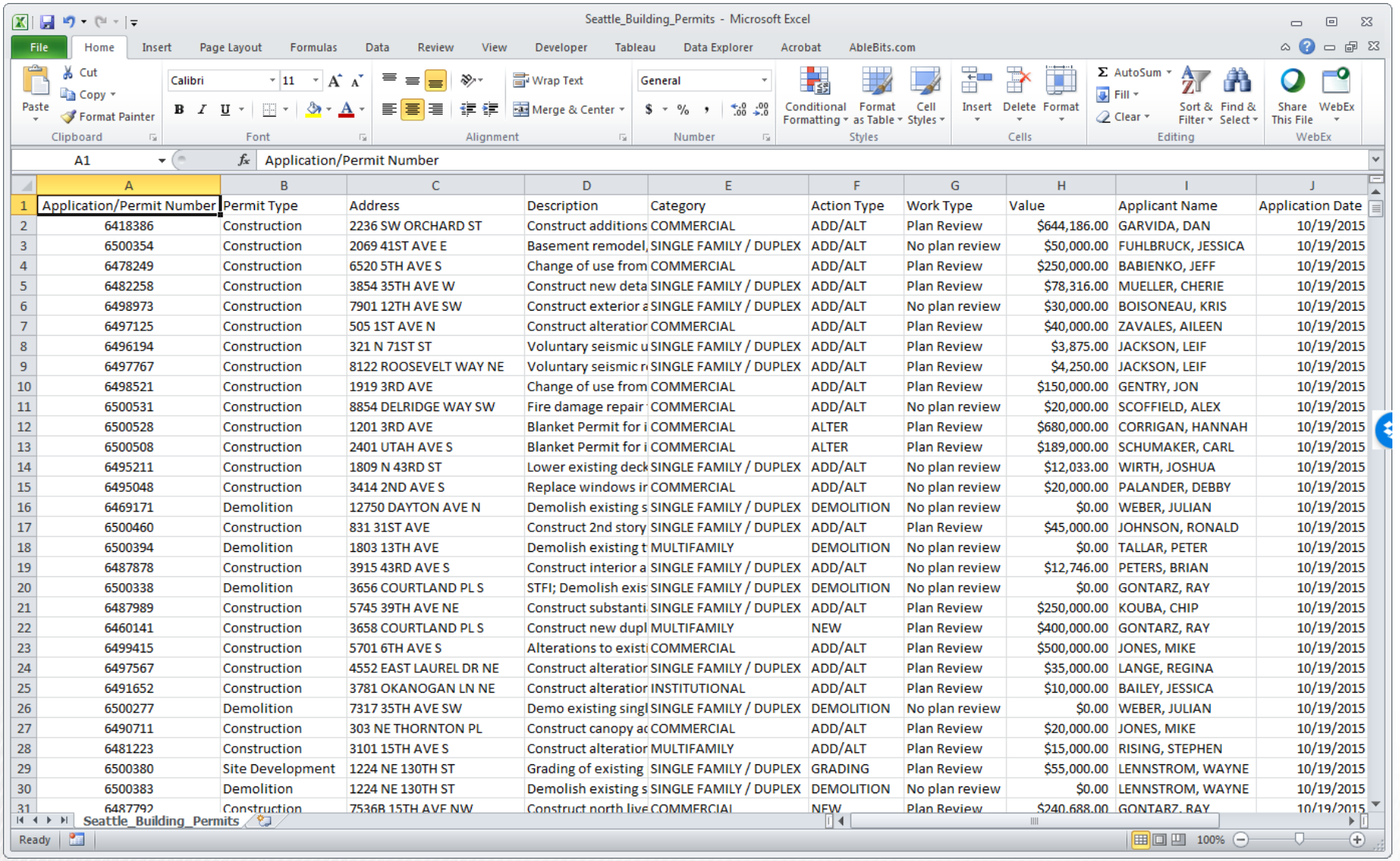

Attributes
Attribute Types
Categorical
Ordered
Ordinal
Quantitative
Ordering Direction
Sequential
Diverging
Cyclic

How Much vs. How Many


Link-Belt
BUCYRUS



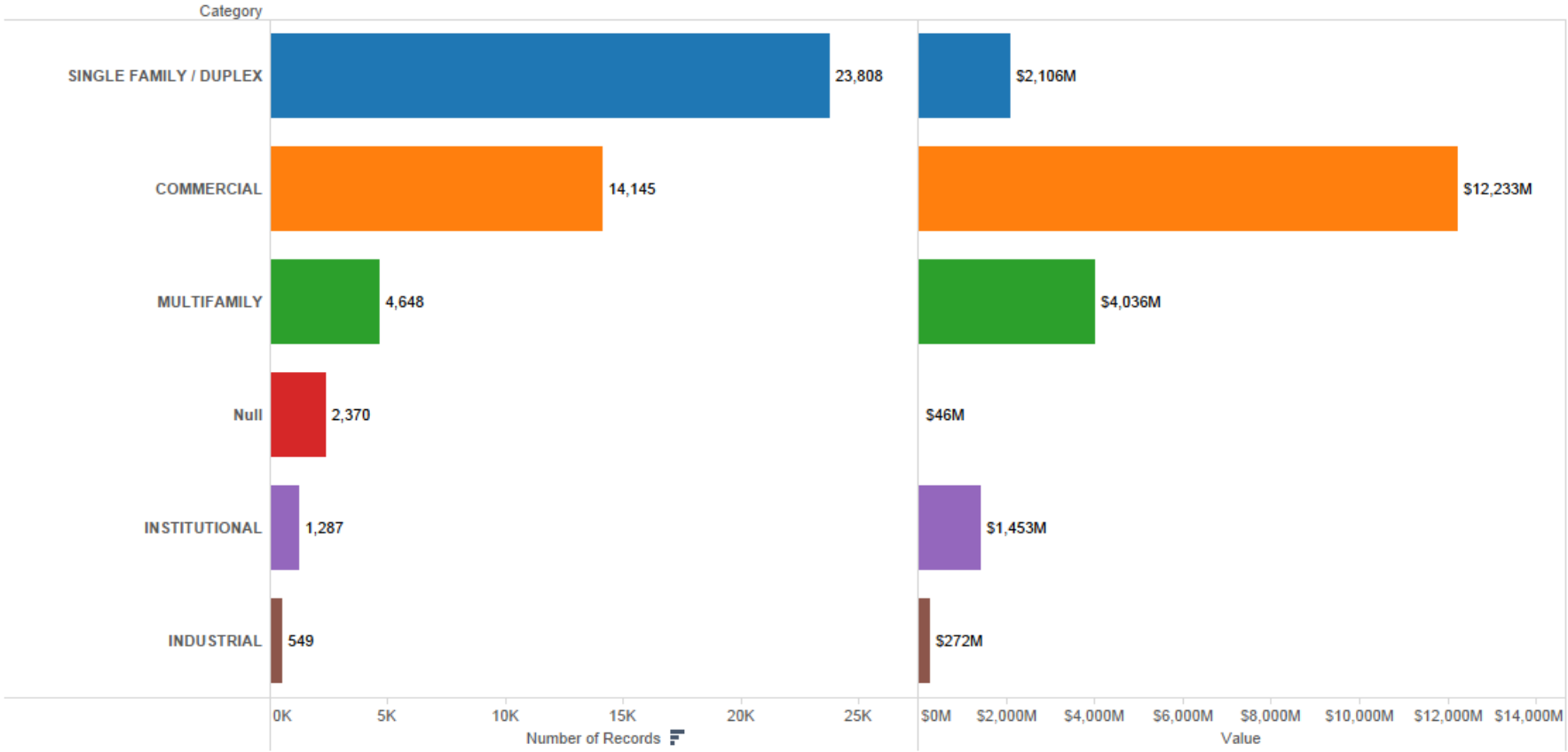
https://data.seattle.gov/Permitting/Building-Permits-Current/mags-97de



	A	B	C	D	E	F	G	H	I	J
1	Application/Permit Number	Permit Type	Address	Description	Category	Action Type	Work Type	Value	Applicant Name	Application Date
2	6418386	Construction	2236 SW ORCHARD ST	Construct additions	COMMERCIAL	ADD/ALT	Plan Review	$644,186.00	GARVIDA, DAN	10/19/2015
3	6500354	Construction	2069 41ST AVE E	Basement remodel,	SINGLE FAMILY / DUPLEX	ADD/ALT	No plan review	$50,000.00	FUHLBRUCK, JESSICA	10/19/2015
4	6478249	Construction	6520 5TH AVE S	Change of use from	COMMERCIAL	ADD/ALT	Plan Review	$250,000.00	BABIENKO, JEFF	10/19/2015
5	6482258	Construction	3854 35TH AVE W	Construct new deta	SINGLE FAMILY / DUPLEX	ADD/ALT	Plan Review	$78,316.00	MUELLER, CHERIE	10/19/2015
6	6498973	Construction	7901 12TH AVE SW	Construct exterior a	SINGLE FAMILY / DUPLEX	ADD/ALT	No plan review	$30,000.00	BOISONEAU, KRIS	10/19/2015
7	6497125	Construction	505 1ST AVE N	Construct alteration	COMMERCIAL	ADD/ALT	Plan Review	$40,000.00	ZAVALES, AILEEN	10/19/2015
8	6496194	Construction	321 N 71ST ST	Voluntary seismic u	SINGLE FAMILY / DUPLEX	ADD/ALT	Plan Review	$3,875.00	JACKSON, LEIF	10/19/2015
9	6497767	Construction	8122 ROOSEVELT WAY NE	Voluntary seismic r	SINGLE FAMILY / DUPLEX	ADD/ALT	Plan Review	$4,250.00	JACKSON, LEIF	10/19/2015
10	6498521	Construction	1919 3RD AVE	Change of use from	COMMERCIAL	ADD/ALT	Plan Review	$150,000.00	GENTRY, JON	10/19/2015
11	6500531	Construction	8854 DELRIDGE WAY SW	Fire damage repair	COMMERCIAL	ADD/ALT	No plan review	$20,000.00	SCOFFIELD, ALEX	10/19/2015
12	6500528	Construction	1201 3RD AVE	Blanket Permit for i	COMMERCIAL	ALTER	Plan Review	$680,000.00	CORRIGAN, HANNAH	10/19/2015
13	6500508	Construction	2401 UTAH AVE S	Blanket Permit for i	COMMERCIAL	ALTER	Plan Review	$189,000.00	SCHUMAKER, CARL	10/19/2015
14	6495211	Construction	1809 N 43RD ST	Lower existing deck	SINGLE FAMILY / DUPLEX	ADD/ALT	No plan review	$12,033.00	WIRTH, JOSHUA	10/19/2015
15	6495048	Construction	3414 2ND AVE S	Replace windows ir	COMMERCIAL	ADD/ALT	No plan review	$20,000.00	PALANDER, DEBBY	10/19/2015
16	6469171	Demolition	12750 DAYTON AVE N	Demolish existing s	SINGLE FAMILY / DUPLEX	DEMOLITION	No plan review	$0.00	WEBER, JULIAN	10/19/2015
17	6500460	Construction	831 31ST AVE	Construct 2nd story	SINGLE FAMILY / DUPLEX	ADD/ALT	Plan Review	$45,000.00	JOHNSON, RONALD	10/19/2015
18	6500394	Demolition	1803 13TH AVE	Demolish existing t	MULTIFAMILY	DEMOLITION	No plan review	$0.00	TALLAR, PETER	10/19/2015
19	6487878	Construction	3915 43RD AVE S	Construct interior a	SINGLE FAMILY / DUPLEX	ADD/ALT	No plan review	$12,746.00	PETERS, BRIAN	10/19/2015
20	6500338	Demolition	3656 COURTLAND PL S	STFI; Demolish exis	SINGLE FAMILY / DUPLEX	DEMOLITION	No plan review	$0.00	GONTARZ, RAY	10/19/2015
21	6487989	Construction	5745 39TH AVE NE	Construct substanti	SINGLE FAMILY / DUPLEX	ADD/ALT	Plan Review	$250,000.00	KOUBA, CHIP	10/19/2015
22	6460141	Construction	3658 COURTLAND PL S	Construct new dupl	MULTIFAMILY	NEW	Plan Review	$400,000.00	GONTARZ, RAY	10/19/2015
23	6499415	Construction	5701 6TH AVE S	Alterations to existi	COMMERCIAL	ADD/ALT	Plan Review	$500,000.00	JONES, MIKE	10/19/2015
24	6497567	Construction	4552 EAST LAUREL DR NE	Construct alteration	SINGLE FAMILY / DUPLEX	ADD/ALT	Plan Review	$35,000.00	LANGE, REGINA	10/19/2015
25	6491652	Construction	3781 OKANOGAN LN NE	Construct alteration	INSTITUTIONAL	ADD/ALT	Plan Review	$10,000.00	BAILEY, JESSICA	10/19/2015
26	6500277	Demolition	7317 35TH AVE SW	Demo existing singl	SINGLE FAMILY / DUPLEX	DEMOLITION	No plan review	$0.00	WEBER, JULIAN	10/19/2015
27	6490711	Construction	303 NE THORNTON PL	Construct canopy ac	COMMERCIAL	ADD/ALT	Plan Review	$20,000.00	JONES, MIKE	10/19/2015
28	6481223	Construction	3101 15TH AVE S	Construct alteration	MULTIFAMILY	ADD/ALT	Plan Review	$15,000.00	RISING, STEPHEN	10/19/2015
29	6500380	Site Development	1224 NE 130TH ST	Grading of existing	SINGLE FAMILY / DUPLEX	GRADING	Plan Review	$55,000.00	LENNSTROM, WAYNE	10/19/2015
30	6500383	Demolition	1224 NE 130TH ST	Demolish existing s	SINGLE FAMILY / DUPLEX	DEMOLITION	No plan review	$0.00	LENNSTROM, WAYNE	10/19/2015
31	6487792	Construction	7536B 15TH AVE NW	Construct north live	COMMERCIAL	NEW	Plan Review	$240,688.00	GONTARZ, RAY	10/19/2015

How Many

How Much



For Next Week

Readings

Read Munzner's *Visualization Analysis & Design"* ch 3-4

Read Jones's *"Communicating Data with Tableau"* ch 4-5

Homework Assignment #2 (Due before class on 10/27)

- Find one dataset that includes at least one quantitative and one categorical attribute (data type).
- Create a visualization that shows "how much" and "how many"
- Publish your visualization to Tableau Public, and upload the Viz URL