



PROFESSIONAL &
CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

I.

Theoretical:

Marks and
Channels

II.

Practical

Central Tendency
& Dispersion

Week 4

Nov 3, 2015

Feedback

Pluses:

- Theoretical topics
- Practical examples
- Time for questions
- Demos with Tableau
- Good sense of humor(?)

Deltas:

- More hands-on practicing
- Give files ahead of time
- Better email response
- Earlier video postings
- Go slower & don't jump around (demos)
- In-Class comments are hard to hear
- More Tableau resources



PROFESSIONAL &
CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

I.

Theoretical:

Marks and
Channels

II.

Practical

Central Tendency
& Dispersion

Week 4
Nov 3, 2015

Channels: Expressiveness Types and Effectiveness Ranks

➔ Magnitude Channels: Ordered Attributes

Position on common scale 

Position on unaligned scale 

Length (1D size) 

Tilt/angle 

Area (2D size) 

Depth (3D position) 

Color luminance 

Color saturation 

Curvature 

Volume (3D size) 

Most

Effectiveness

Least

➔ Identity Channels: Categorical Attributes

Spatial region 

Color hue 

Motion 

Shape 

Group →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
↓ Period																		
1	1 H																	2 He
2	3 Li	4 Be											5 B	6 C	7 N	8 O	9 F	10 Ne
3	11 Na	12 Mg											13 Al	14 Si	15 P	16 S	17 Cl	18 Ar
4	19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
5	37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
6	55 Cs	56 Ba	*	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
7	87 Fr	88 Ra	**	104 Rf	105 Db	106 Sg	107 Bh	108 Hs	109 Mt	110 Ds	111 Rg	112 Cn	113 Uut	114 Fl	115 Uup	116 Lv	117 Uus	118 Uuo
		*	57 La	58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu	
		**	89 Ac	90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr	



Designed by
Datafemised

Interactive Periodic Table of Elements with Variable Sizing

Size By

Atomic weight

State at STP

☒ (All)

☐ Gas

☐ Liquid

☐ Solid

☐ Unknown

Occurrence

☒ (All)

☐ Primordial

☐ Synthetic

☐ Transient

Description

☒ Actinide

☐ Alkali metal

☐ Alkaline earth metal

☐ Halogen

☐ Lanthanide

☐ Metal

☐ Metalloid

☐ Noble gas

☐ Non-metal

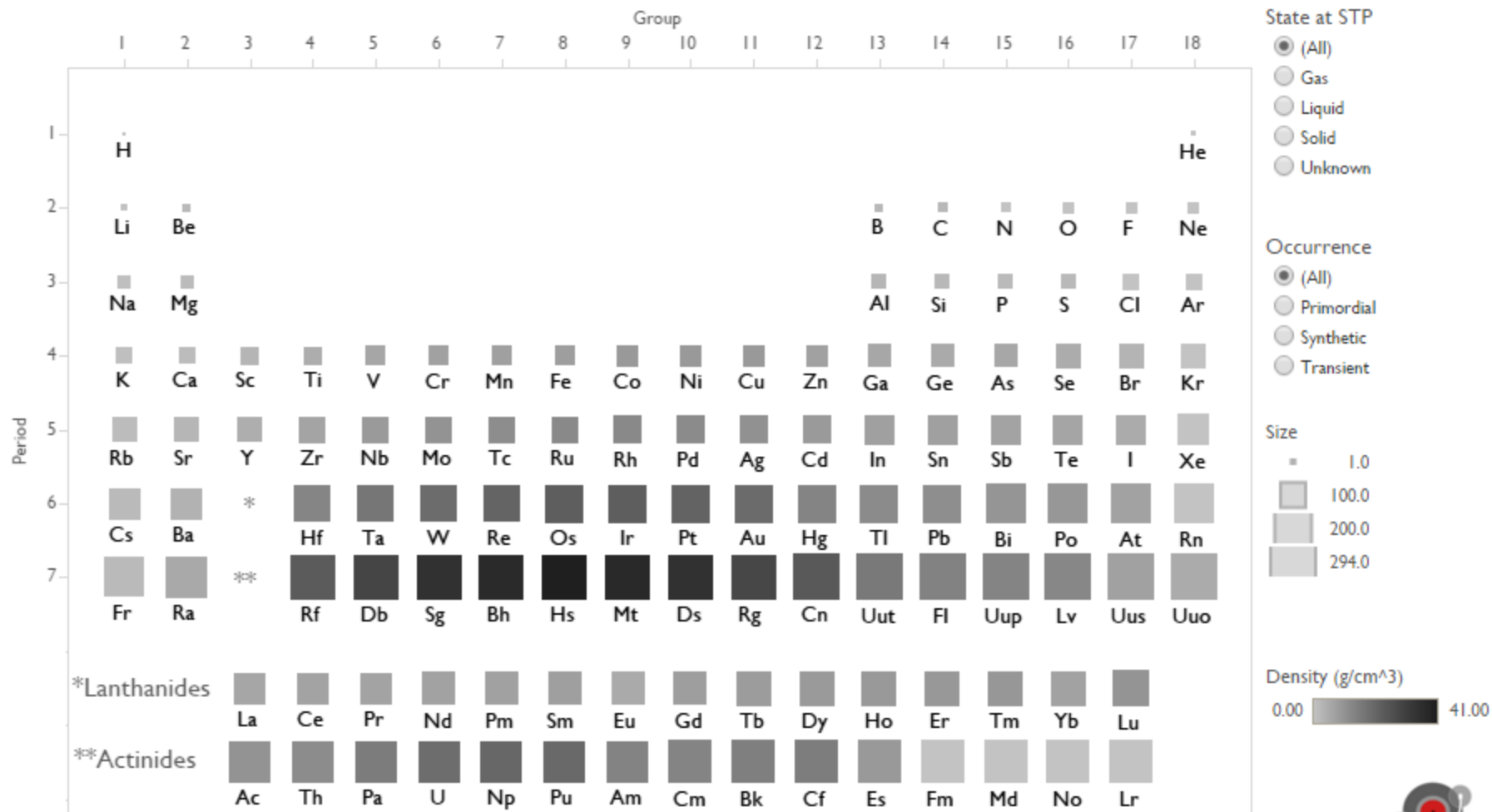
☐ None

☐ Transition metal

	Group																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	H																	He
2	Li	Be											B	C	N	O	F	Ne
3	Na	Mg											Al	Si	P	S	Cl	Ar
4	K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br	Kr
5	Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I	Xe
6	Cs	Ba	*	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At	Rn
7	Fr	Ra	**	Rf	Db	Sg	Bh	Hs	Mt	Ds	Rg	Cn	Uut	Fl	Uup	Lv	Uus	Uuo
*Lanthanides			La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu	
**Actinides			Ac	Th	Pa	U	Np	Pu	Am	Cm	Bk	Cf	Es	Fm	Md	No	Lr	

Source: http://en.wikipedia.org/wiki/List_of_elements

Interactive Periodic Table of Elements with Variable Sizing



Source: http://en.wikipedia.org/wiki/List_of_elements



Interactive Periodic Table of Elements with Variable Sizing

Size By

Atomic weight

Shapes

State at STP

Occurrence

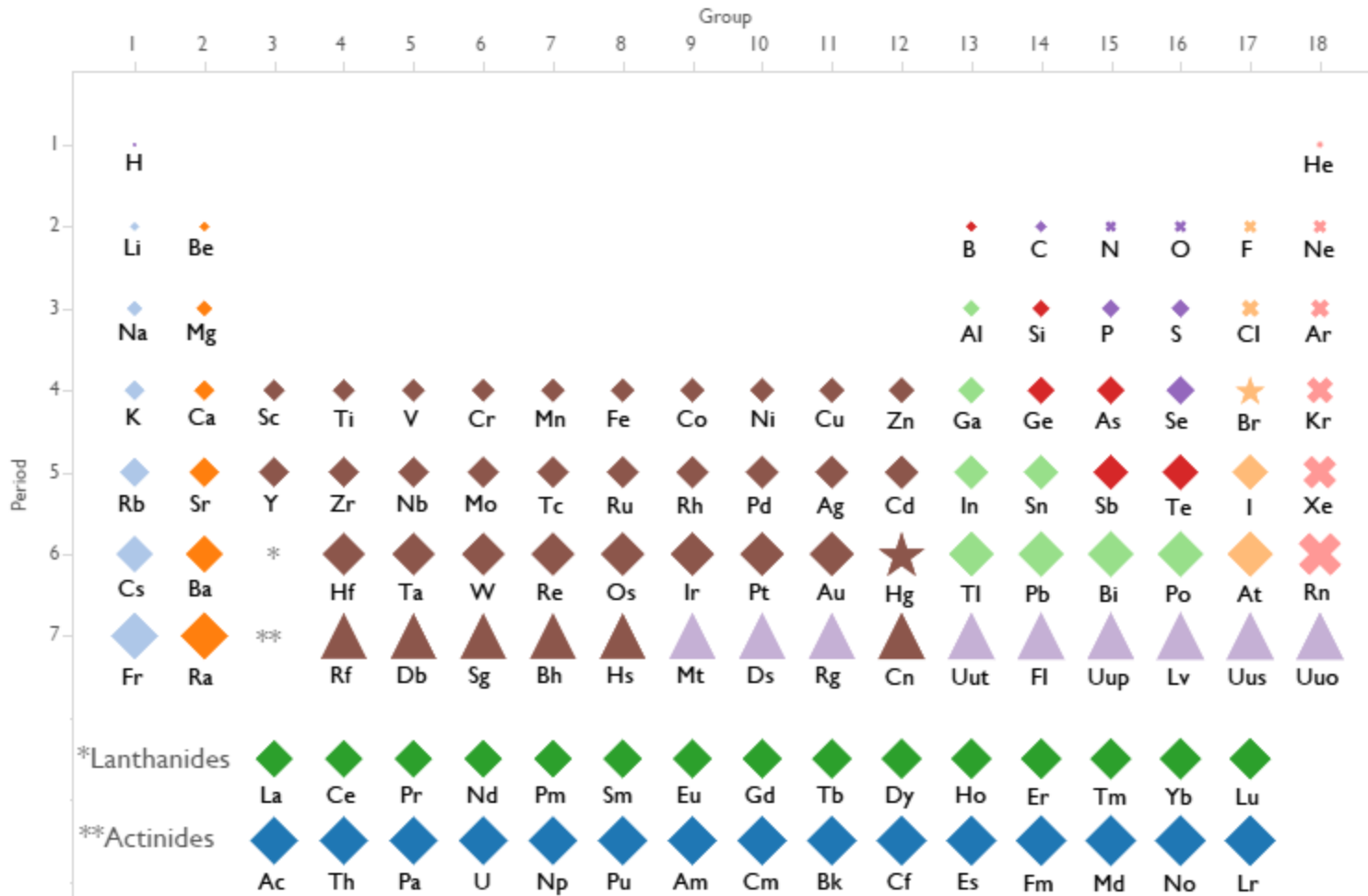
- (All)
- Primordial
- Synthetic
- Transient

Shapes

- Gas
- Liquid
- Solid
- Unknown

Description

- Actinide
- Alkali metal
- Alkaline earth metal
- Halogen
- Lanthanide
- Metal
- Metalloid
- Noble gas
- Non-metal
- None
- Transition metal



Source: http://en.wikipedia.org/wiki/List_of_elements

Expressiveness, Effectiveness and Accuracy

The **expressiveness principle** dictates that the visual encoding should express all of, and only, the information in the dataset attributes. The most fundamental expression of this principle is that ordered data should be shown in a way that our perceptual system intrinsically senses as ordered. Conversely, unordered data should not be shown in a way that perceptually implies an ordering that does not exist.

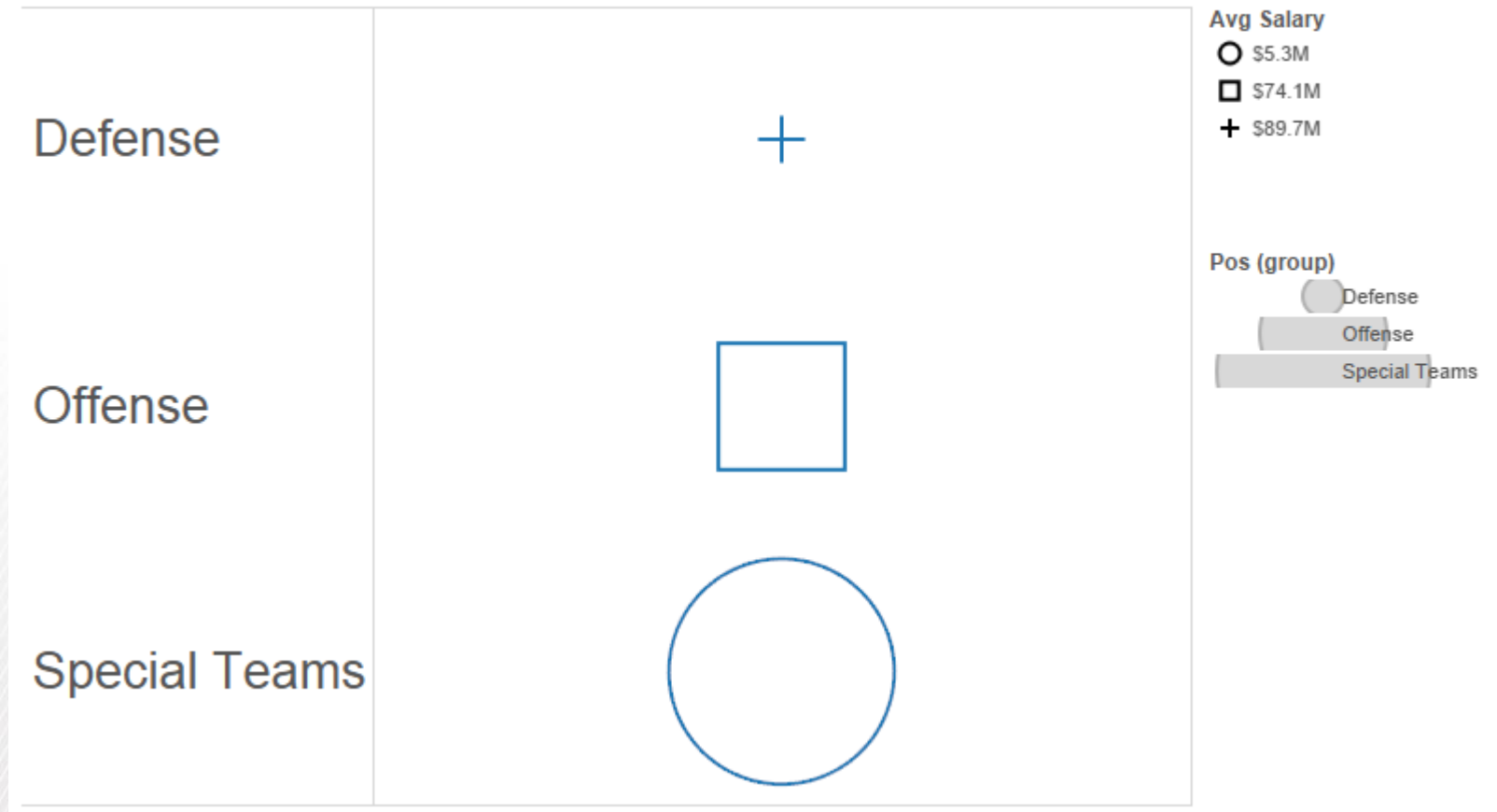
The **effectiveness principle** dictates that the importance of the attribute should match the salience of the channel; that is, its noticeability. In other words, the most important attributes should be encoded with the most effective channels in order to be most noticeable, and then decreasingly important attributes can be matched with less effective channels.

– Munzner, Tamara. *Visualization Analysis and Design*. CRC Press, 09/2014. VitalBook file.

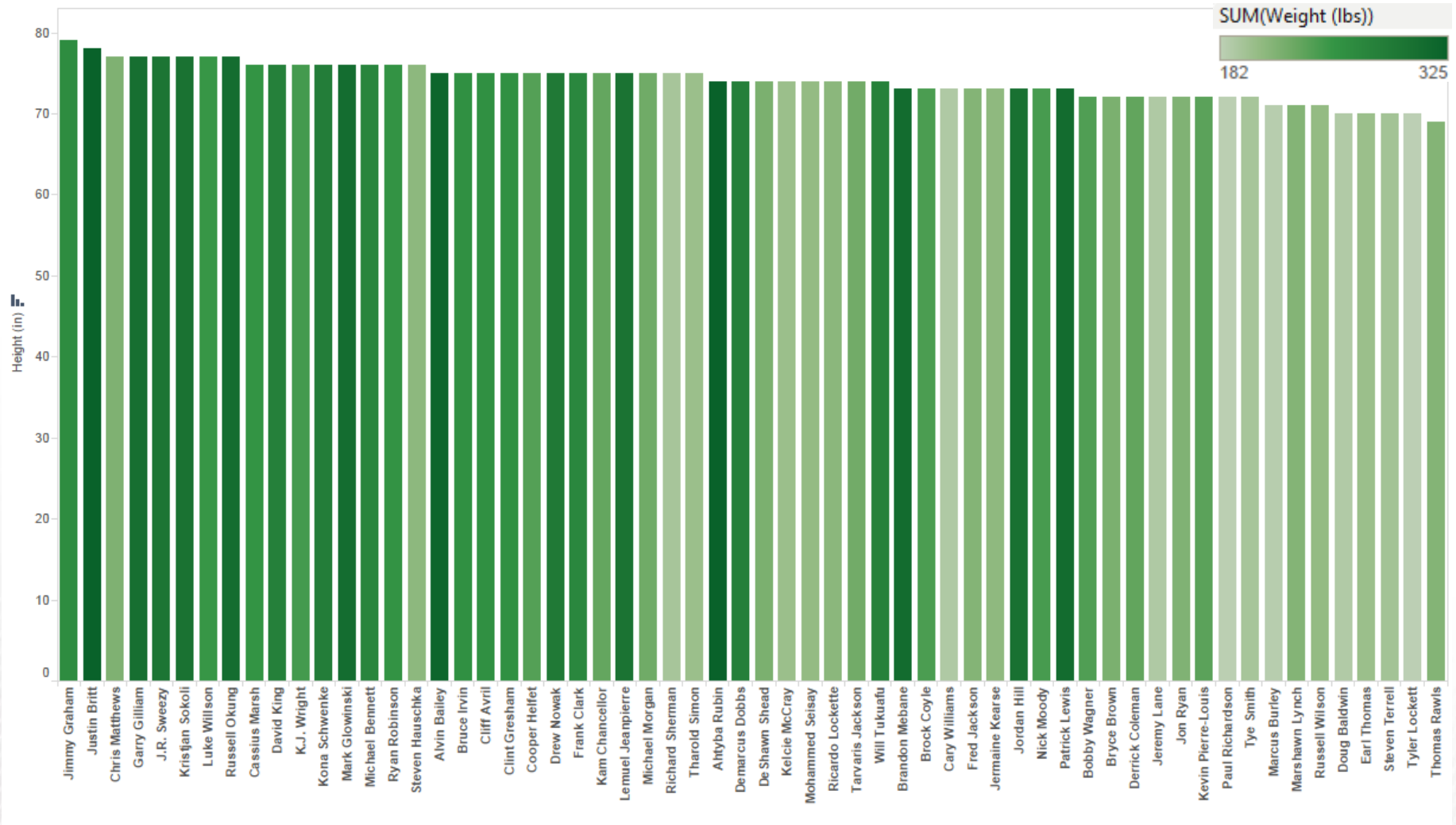
The citation provided is a guideline. Please check each citation for accuracy before use.

Expressiveness: Does this work?

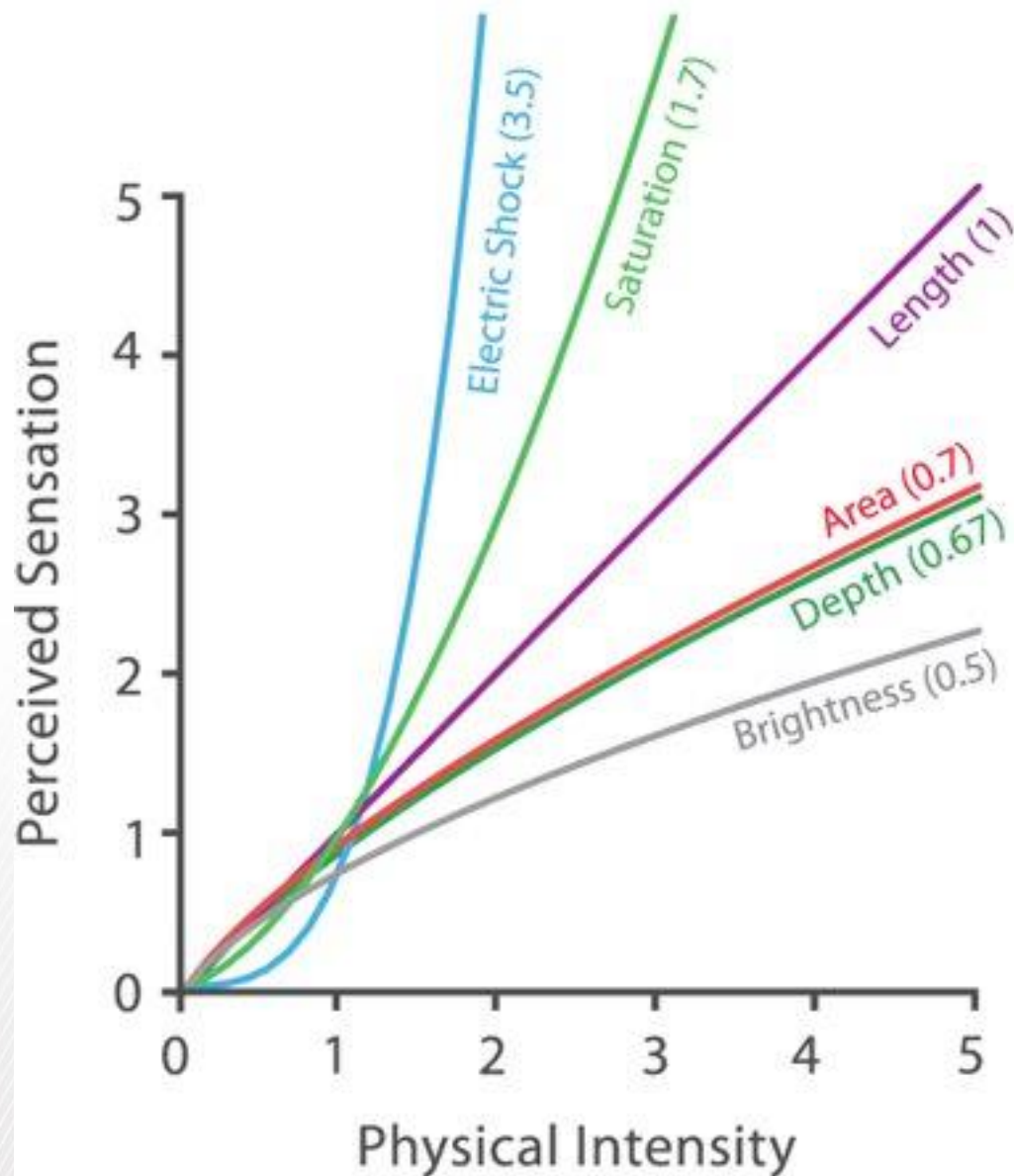
How Much the Seattle Seahawks Spend On....



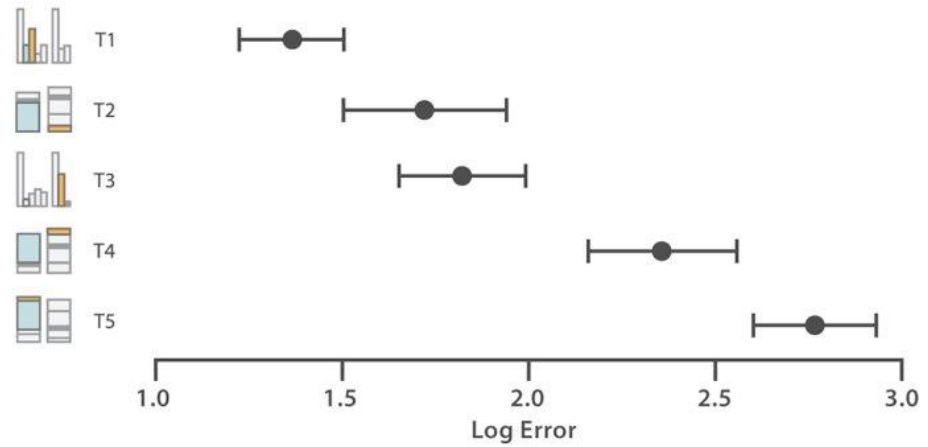
Effectiveness: Is it easier to tell relative **height** or **weight**?



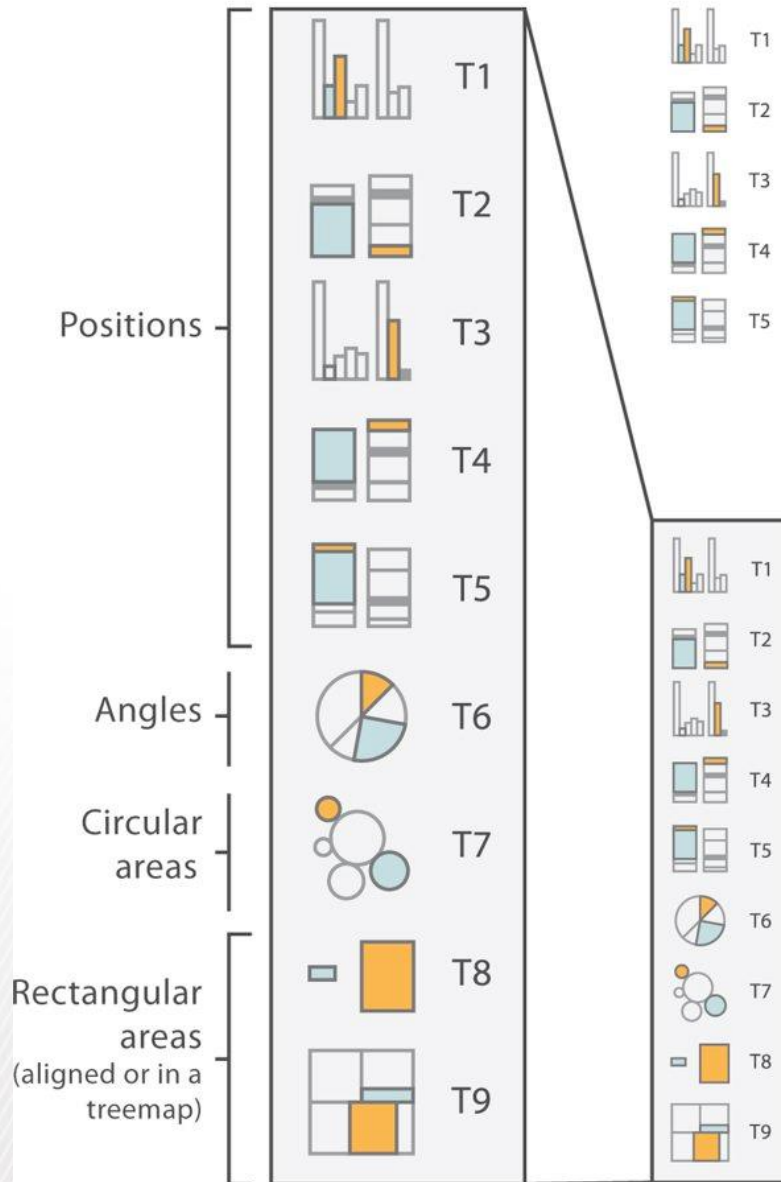
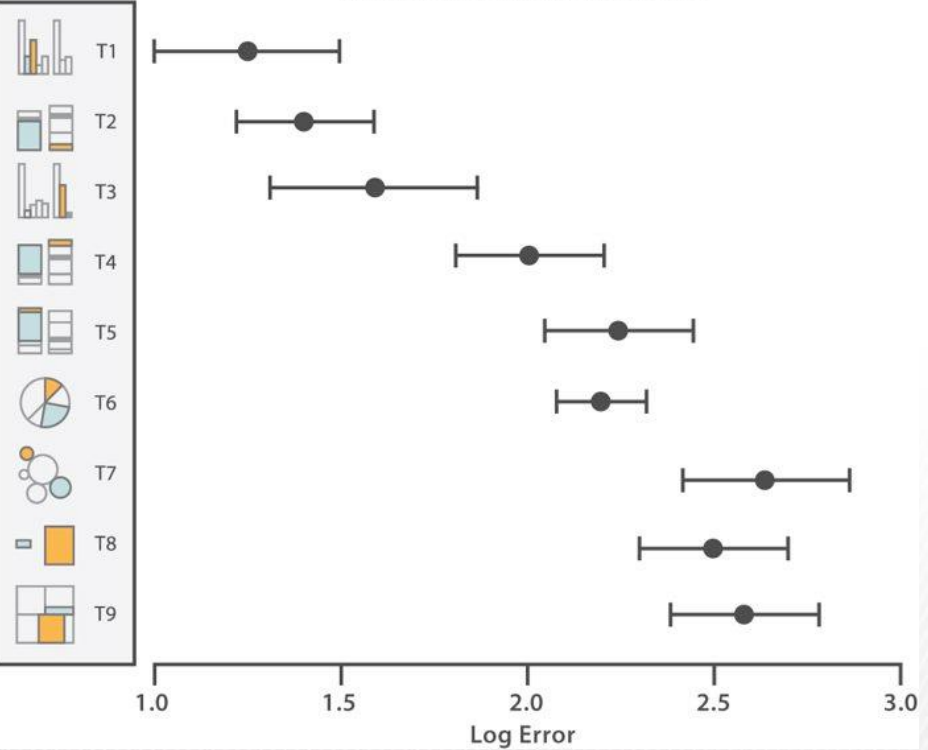
Steven's Psychophysical Power Law: $S = I^N$



Cleveland & McGill's Results



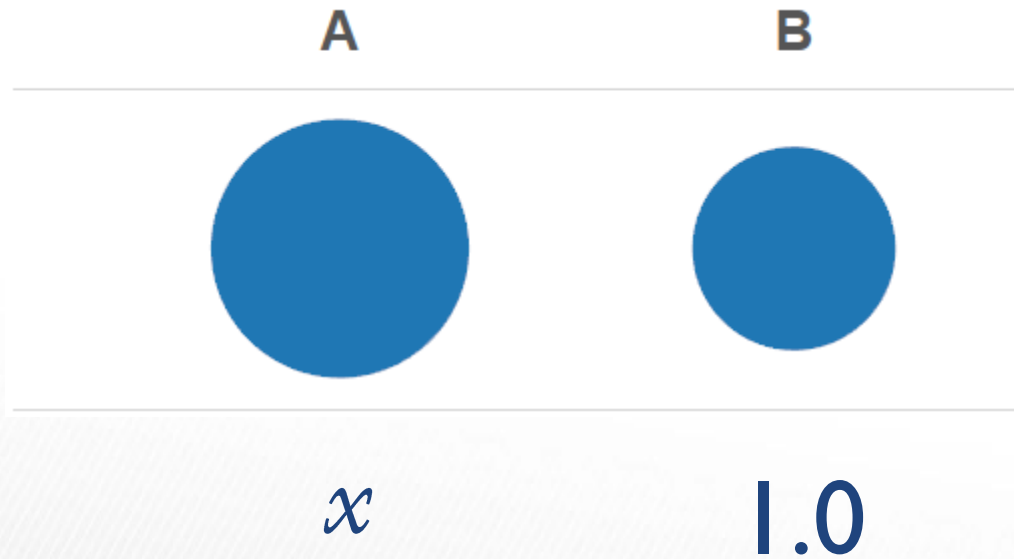
Crowdsourced Results



Which is bigger?



How much bigger?



Which is darker?

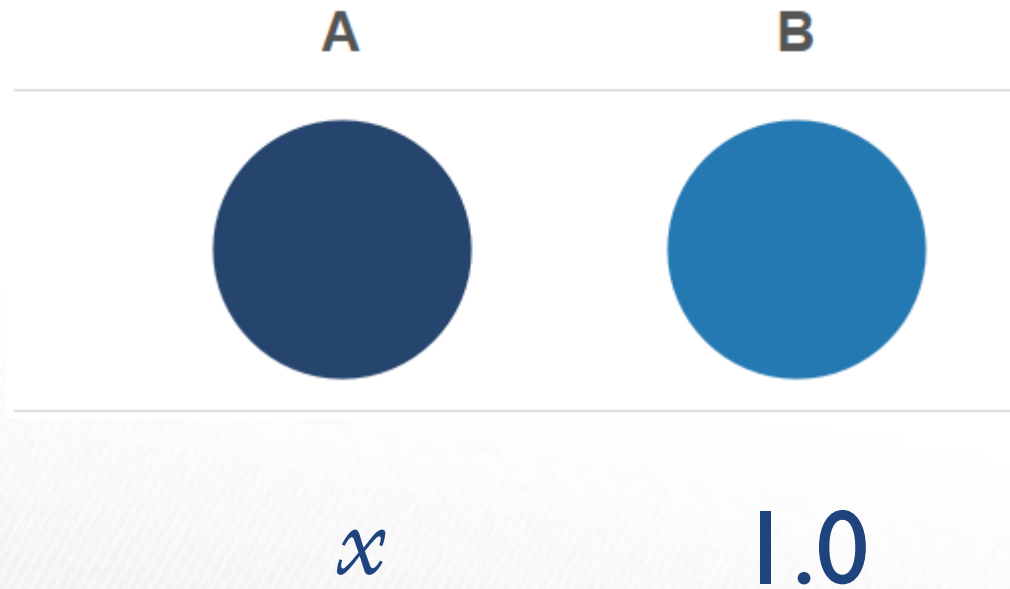
A



B



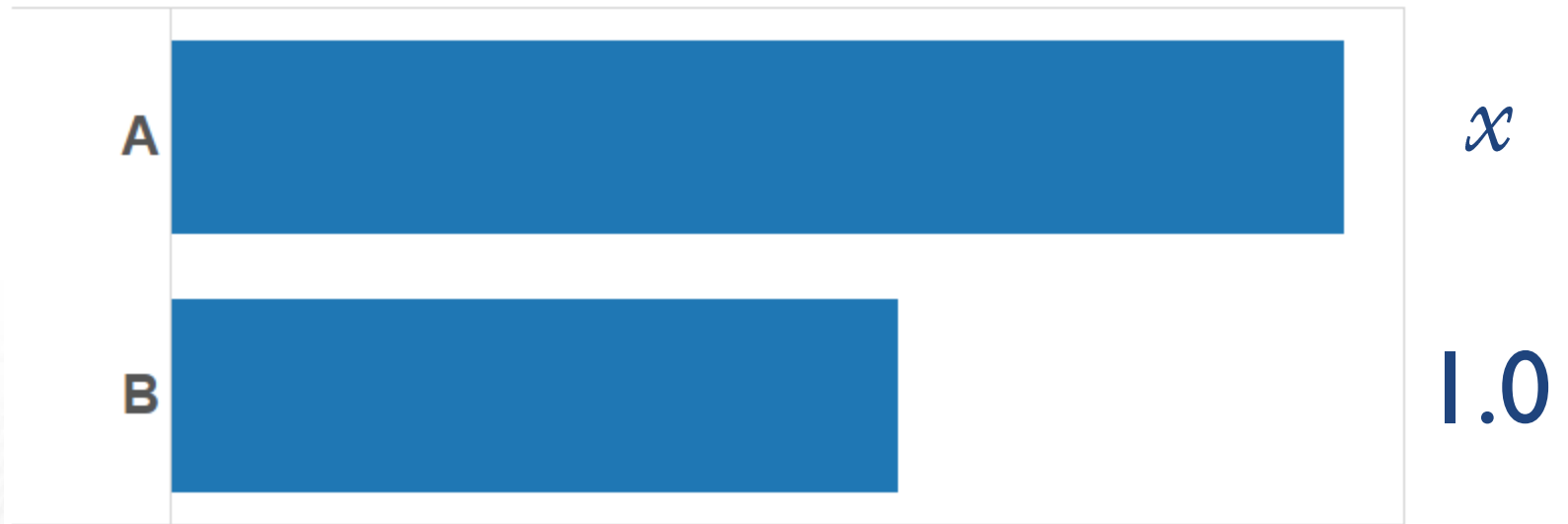
How much darker?



Which is longer?

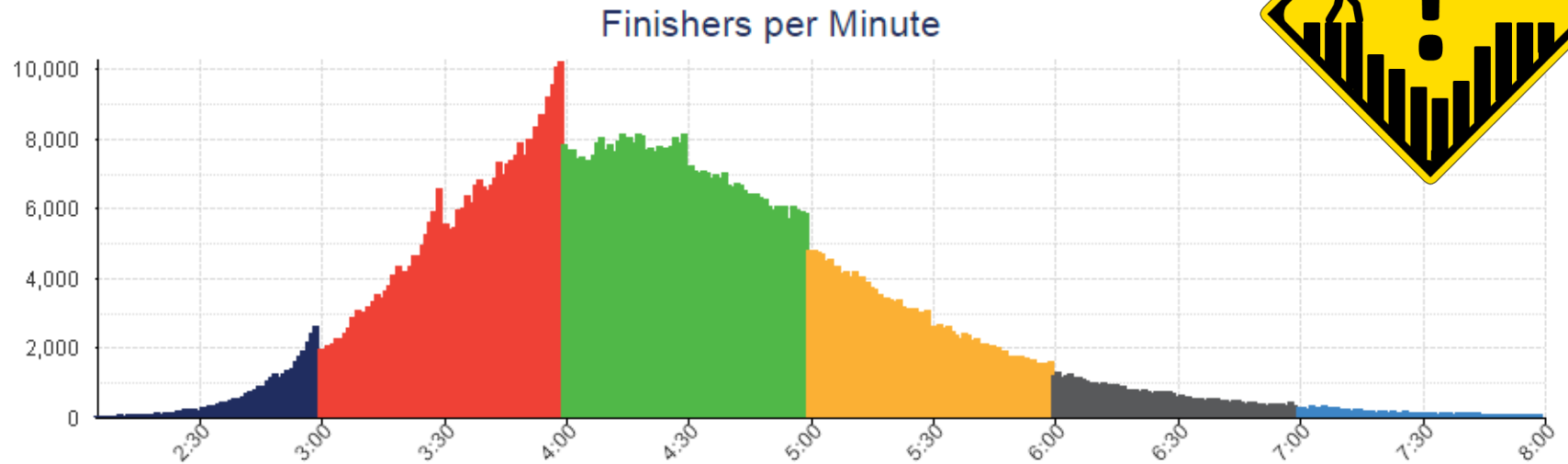


How much longer?

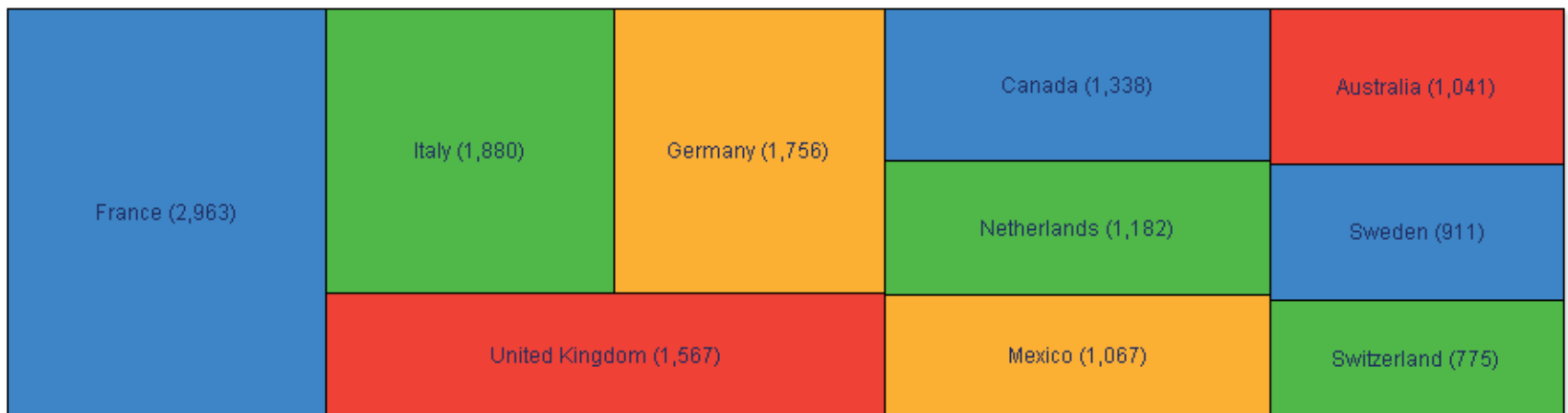


A word on Color Hue and Color Saturation

What's wrong with this?



Top 10: Growth in Number of Finishers in 2014 vs. 1970 by Country (non-US)



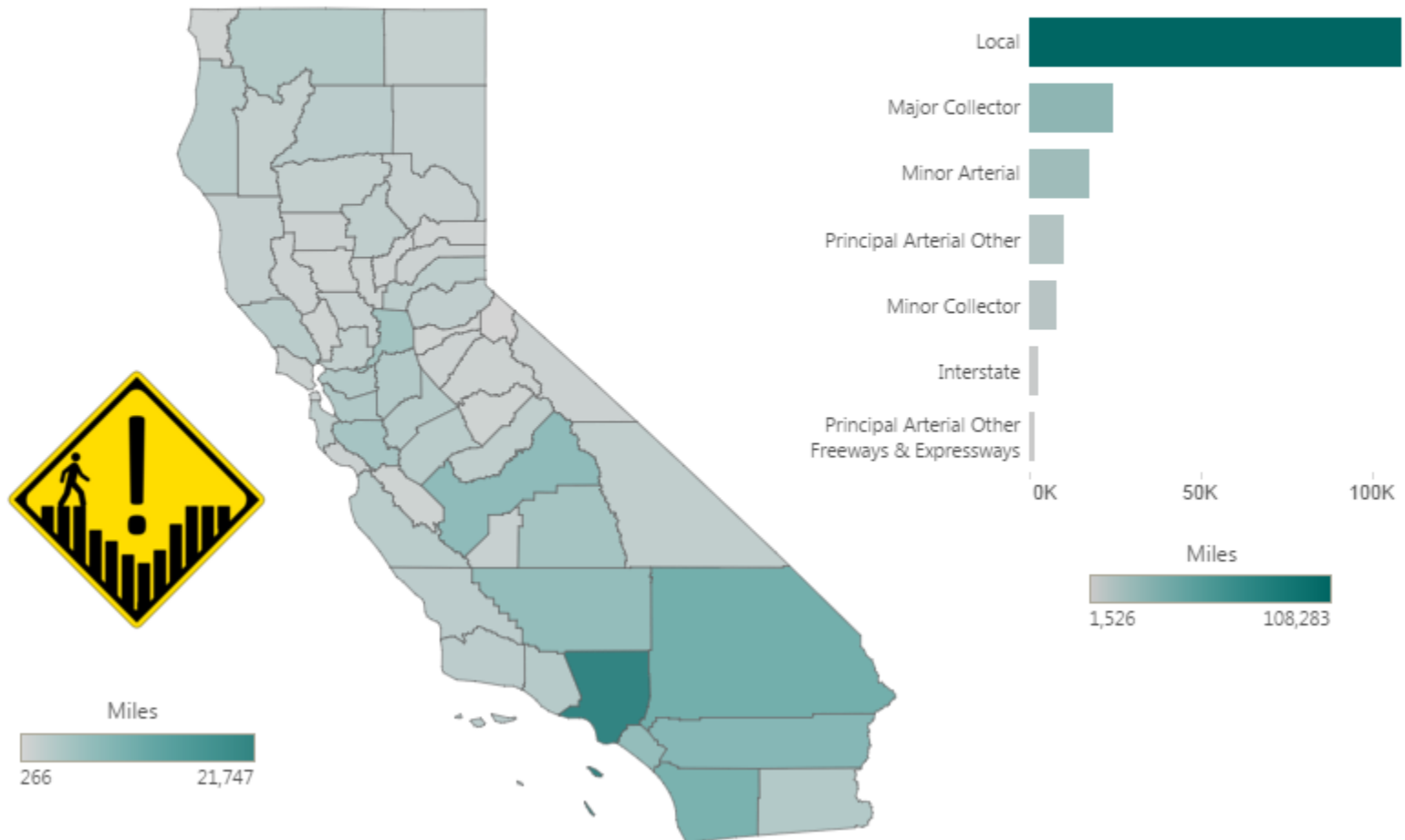
What's wrong with this?

A California Road-Trip

miles of roads, by county and type

Miles

(All)



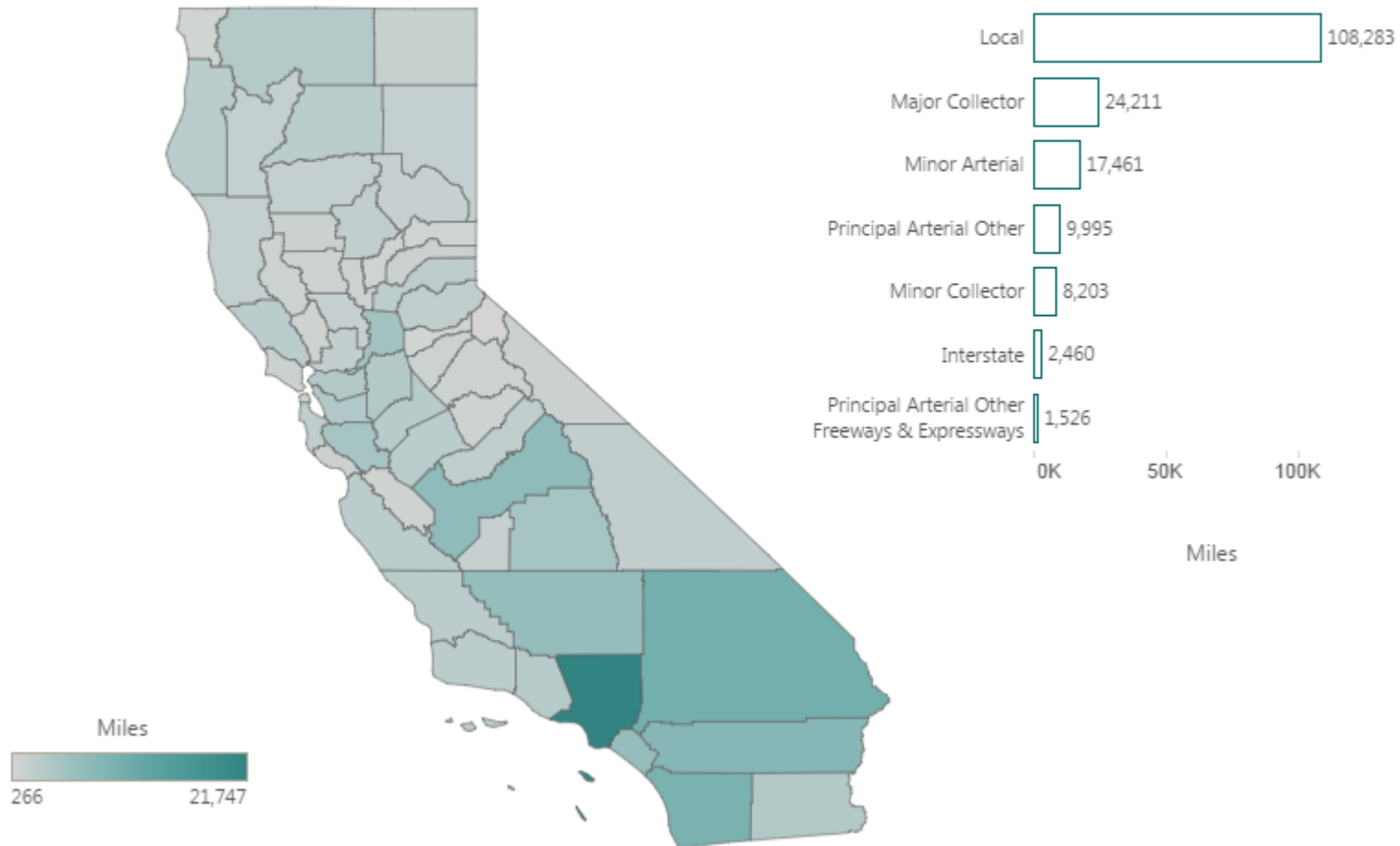
Improved

A California Road-Trip

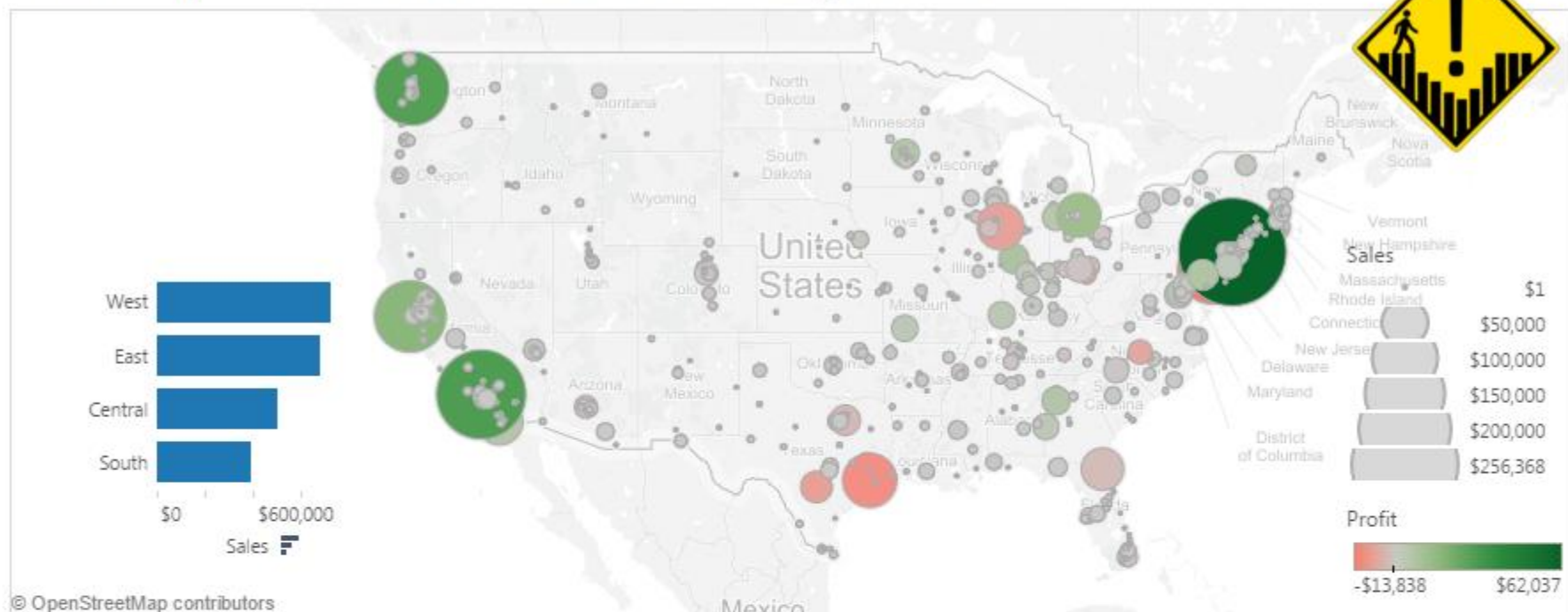
miles of roads, by county and type

Miles ▼

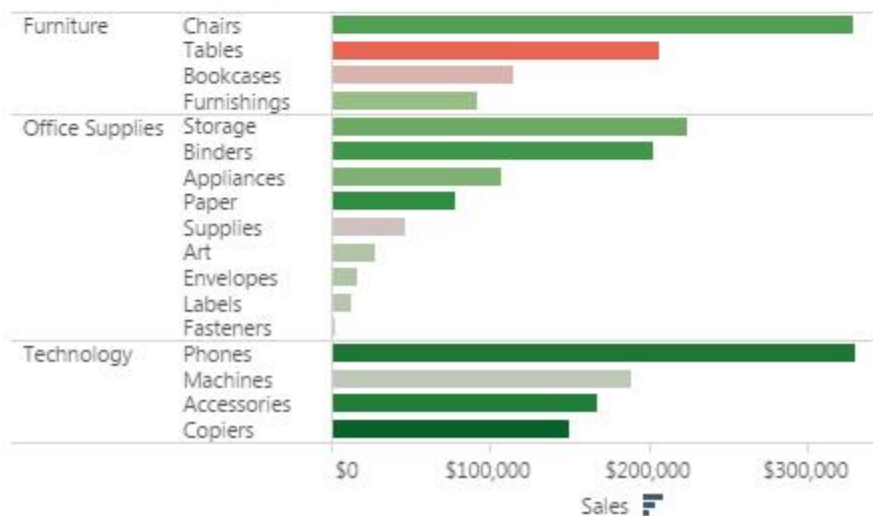
(All) ▼



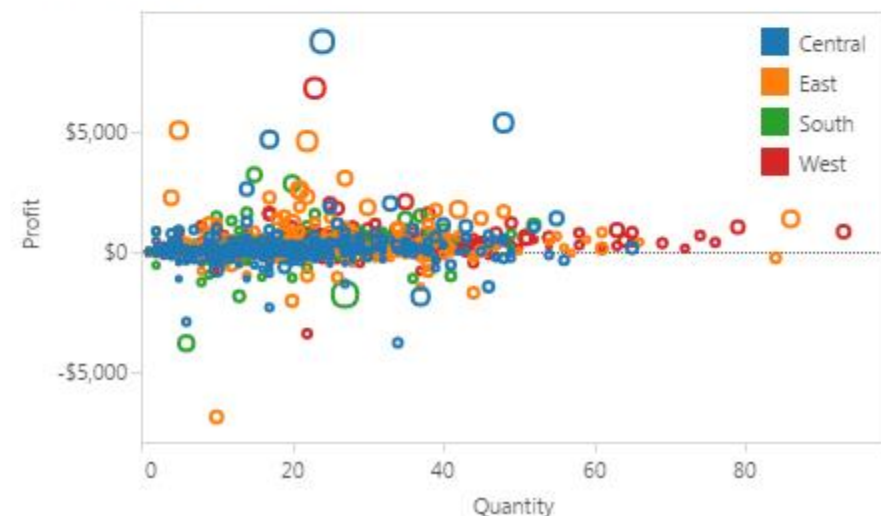
Sales Superstore Dashboard: Confusing Colors



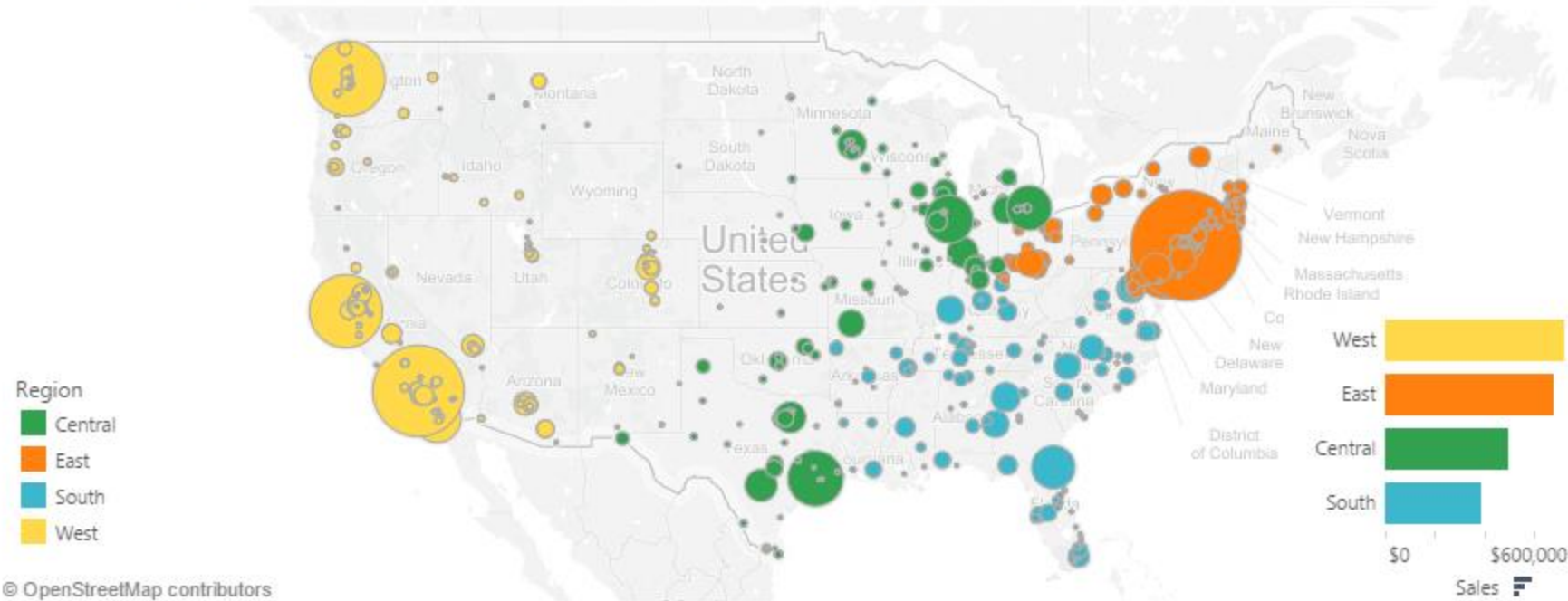
Sales by Product Category



Quantity and Profit by Customer



Sales Superstore Regional Dashboard: Single Color Palette

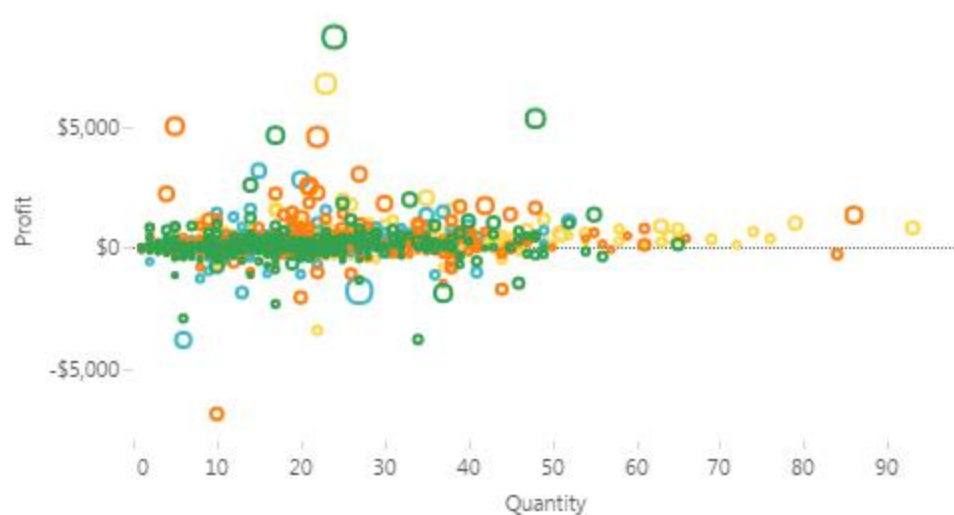


© OpenStreetMap contributors

Sales by Product Category



Quantity and Profit by Customer

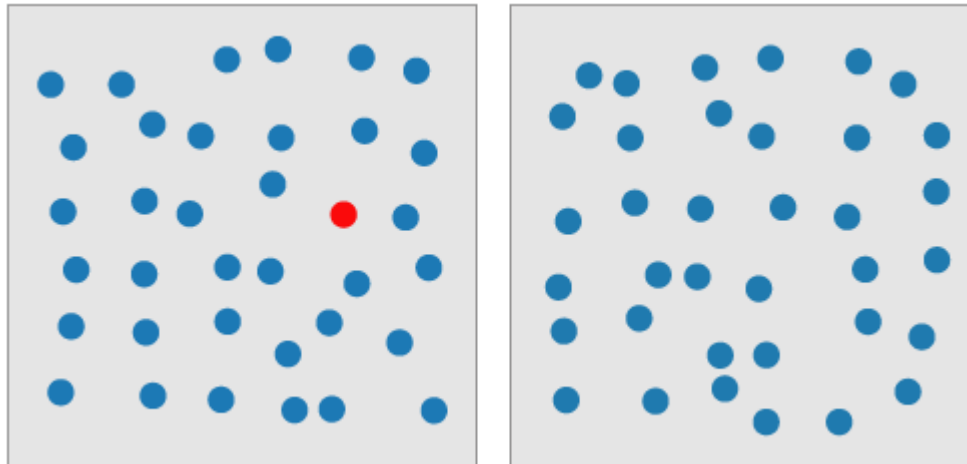


The Color Channels: Rules of Thumb

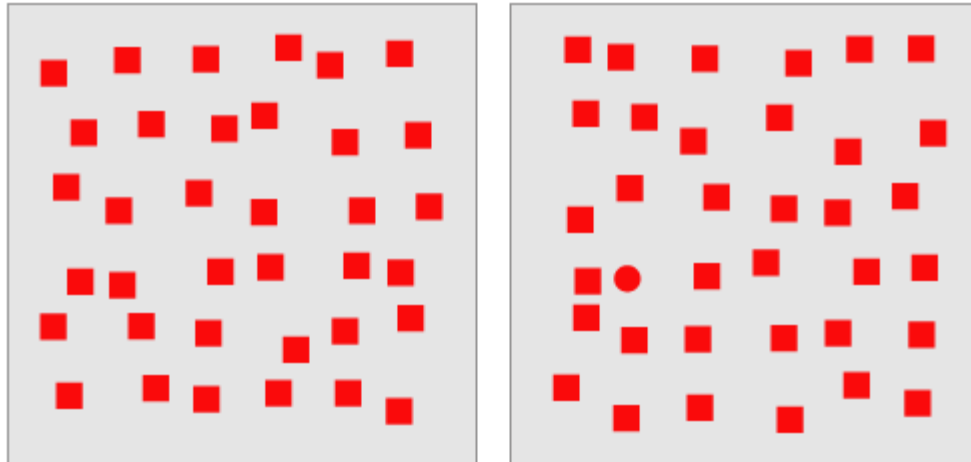
1. DON'T use the same color hue for two different variables
2. DON'T use the same color saturation for different magnitudes of the same variable
3. DON'T use too many color encodings on one dashboard
4. When possible, DO use one and only one color encoding per view

“Preattentive” Processing

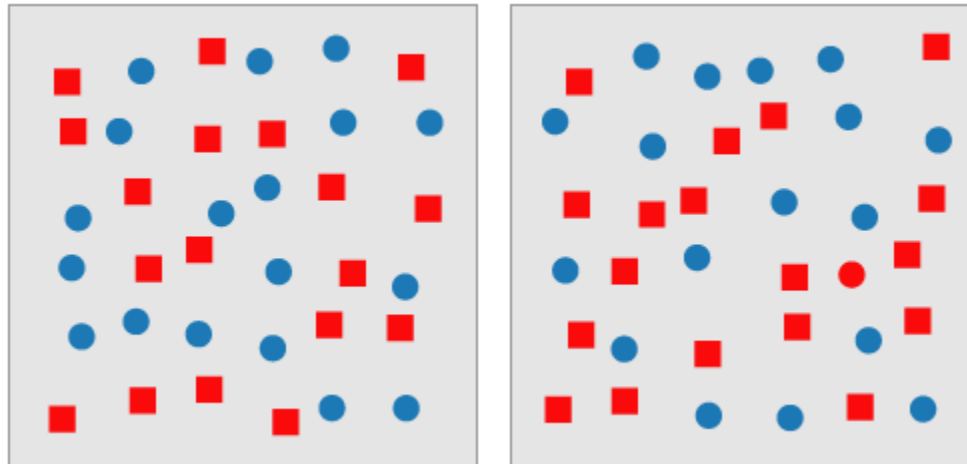
Find the box with the red circle



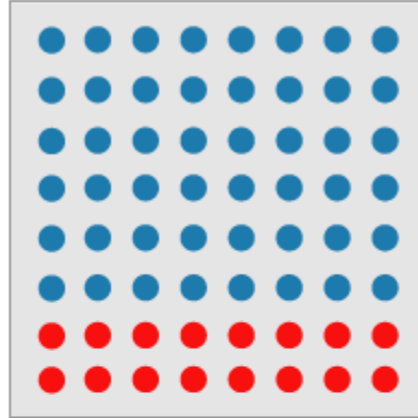
Find the box with the red circle



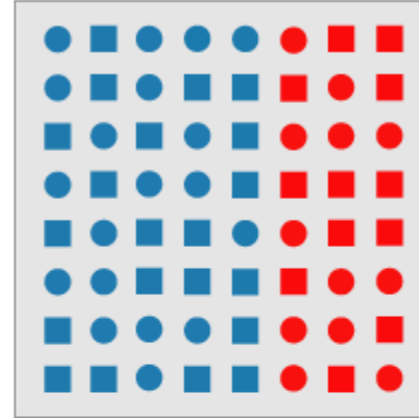
Find the box with the red circle



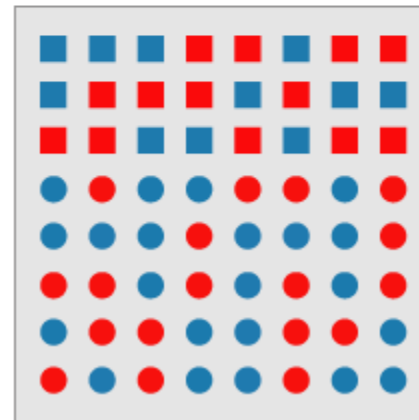
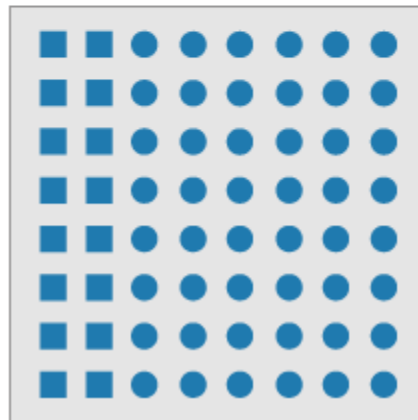
Find the boundary line



(a)



(b)



Use of Preattentive Features

Target detection:

- users rapidly and accurately detect the presence or absence of a "target" element with a unique visual feature within a field of distractor elements

Boundary detection:

- users rapidly and accurately detect a texture boundary between two groups of elements, where all of the elements in each group have a common visual property

Region tracking:

- users track one or more elements with a unique visual feature as they move in time and space, and

Counting and estimation:

- users count or estimate the number of elements with a unique visual feature.



PROFESSIONAL &
CONTINUING EDUCATION
UNIVERSITY *of* WASHINGTON

I.

Theoretical:

Marks and
Channels

II.

Practical

Central Tendency
& Dispersion

Week 4

Nov 3, 2015

Measures of Central Tendency

What does “typical” look like?



Mean, Median and Mode

3, 7, 10, 8, 31, 10, 2

$$\text{Mean (avg)} = \frac{3 + 7 + 10 + 8 + 31 + 10 + 2}{7} = \frac{71}{7}$$

↓
10.14

7 numbers

$$\text{Median} = 2, 3, 7, 8, 10, 10, 31$$

↓
8

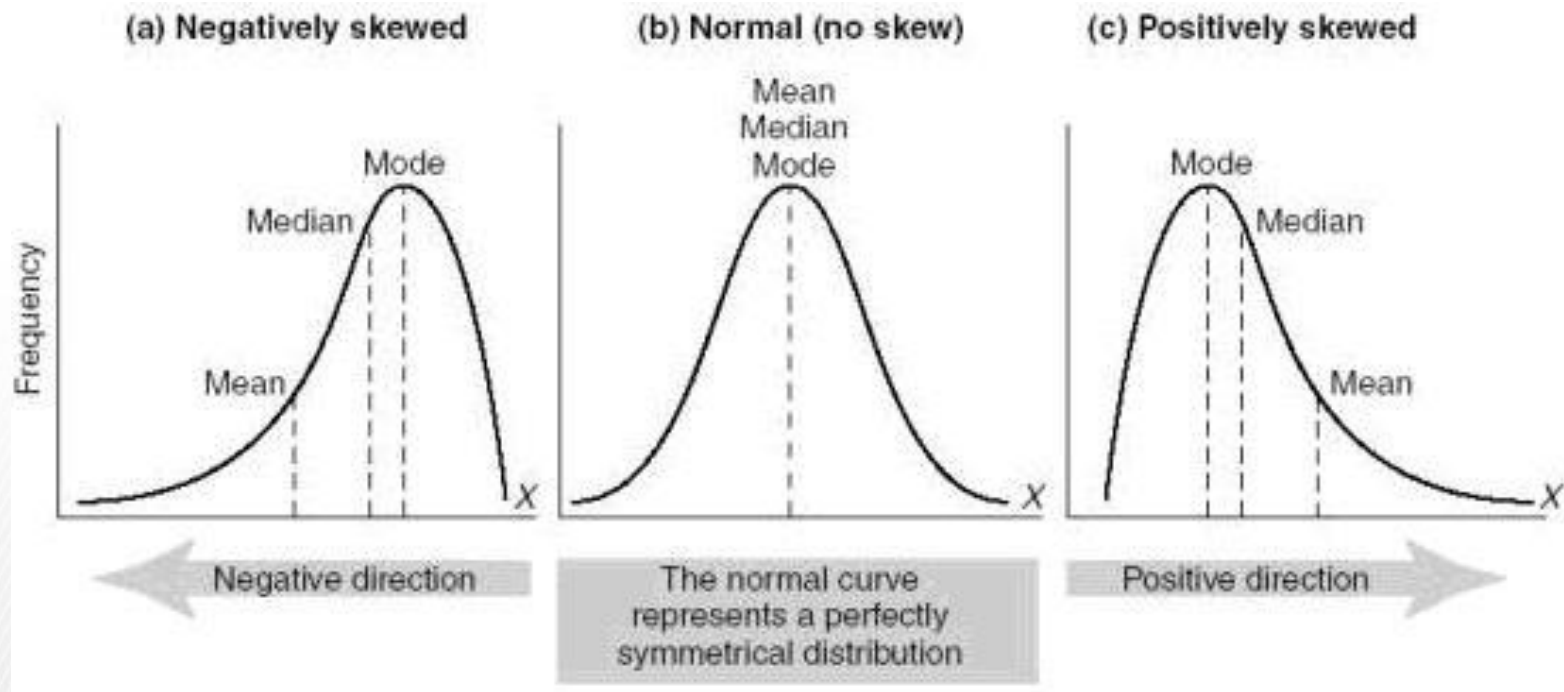
↑
middle

Mode

3, 7, (10), 8, 31, (10), 2

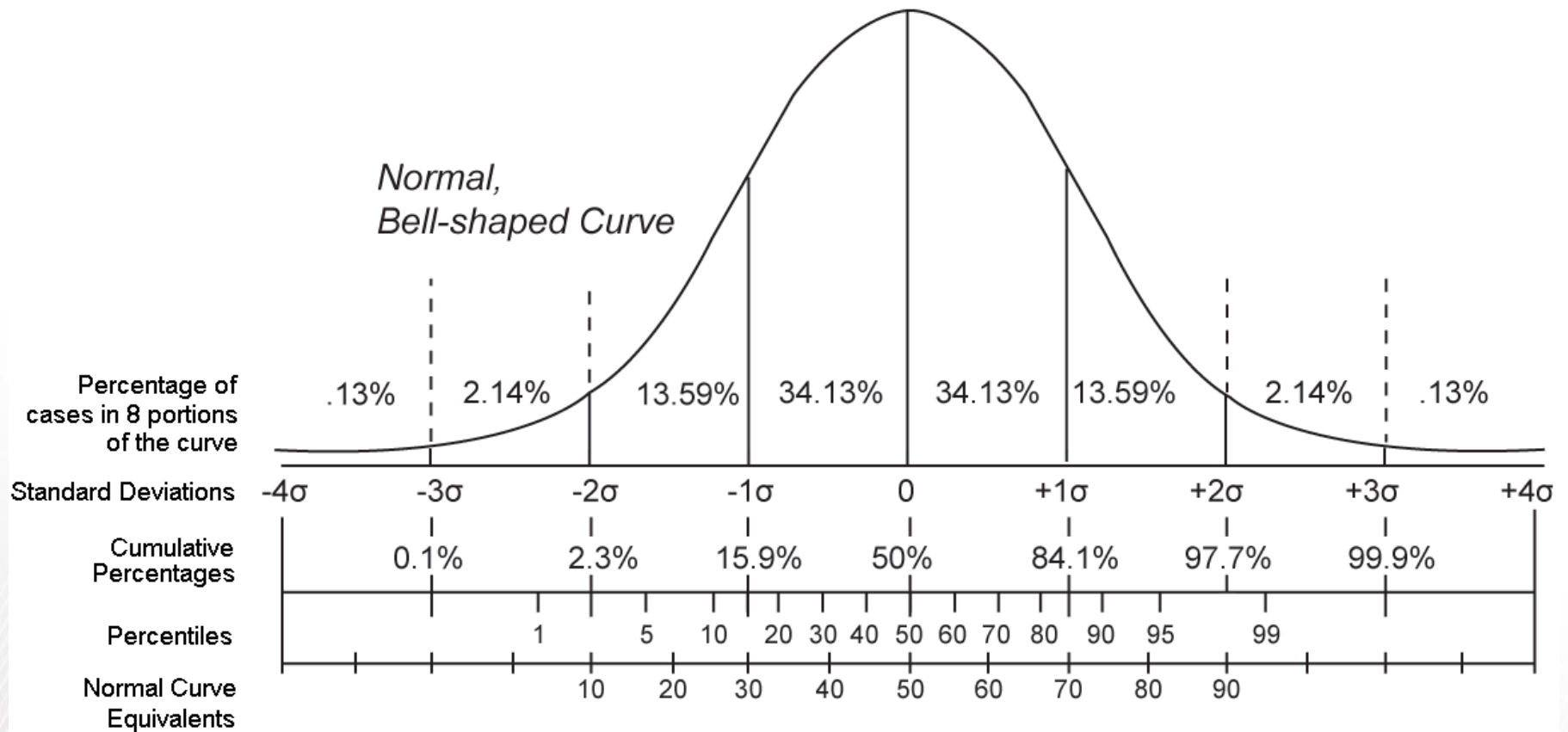
↓
10

Comparing measures of central tendency

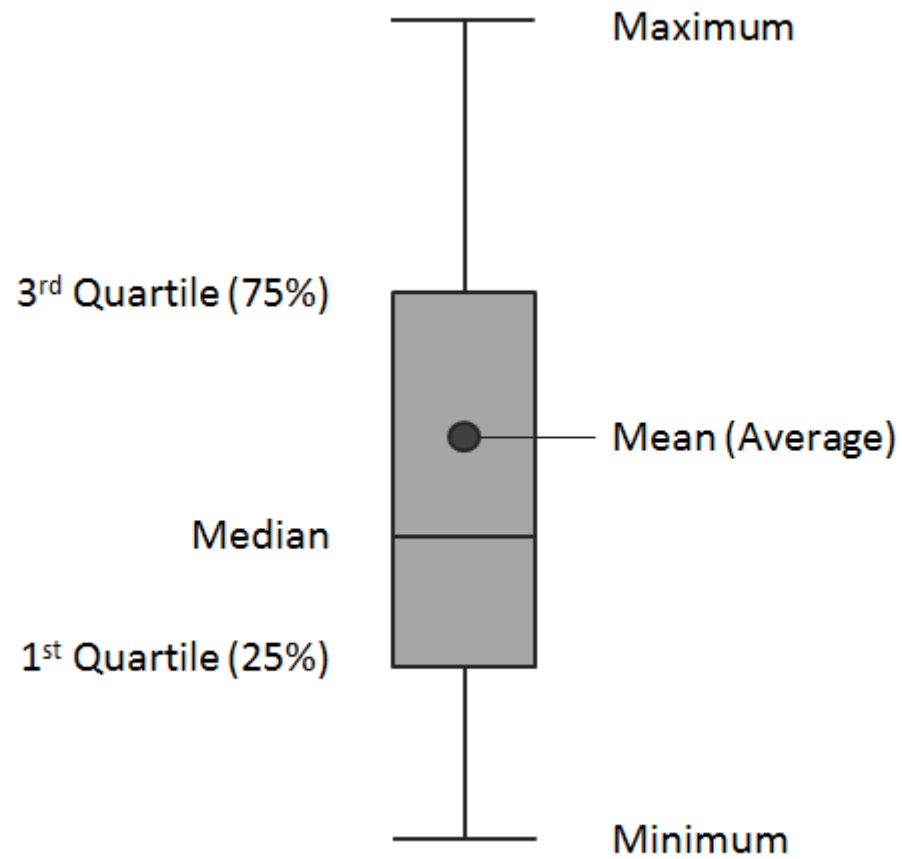


Measures of Dispersion

The Standard Deviation

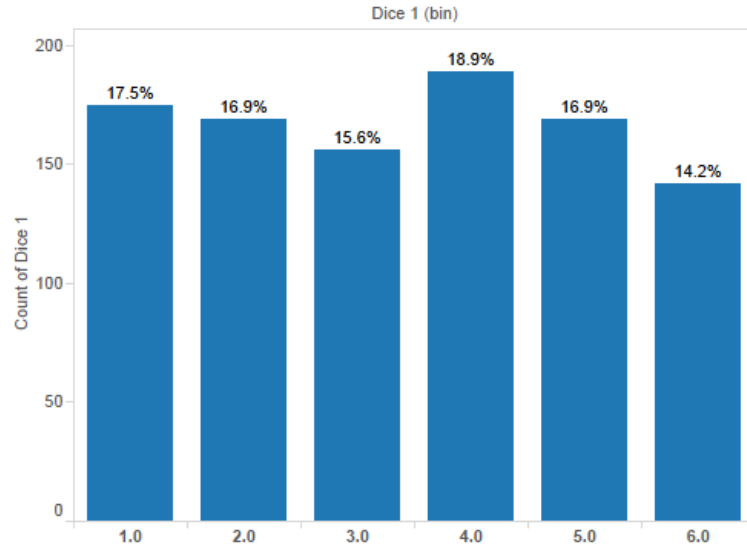


Quartiles

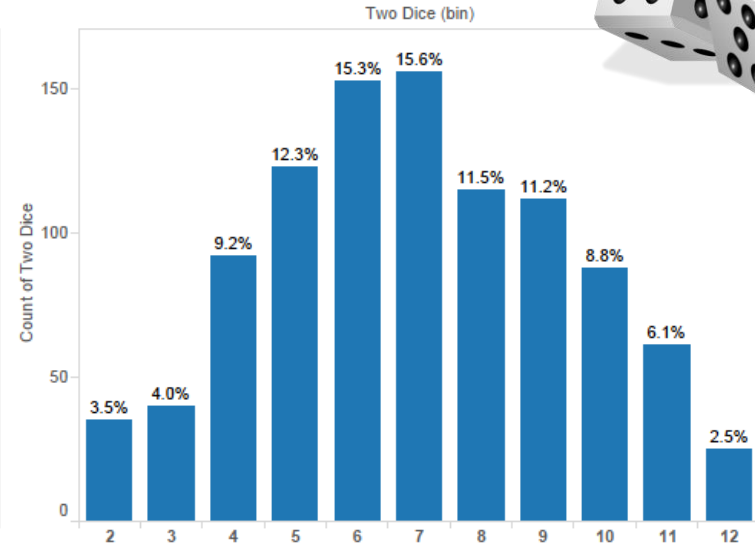


Visualizing Variation

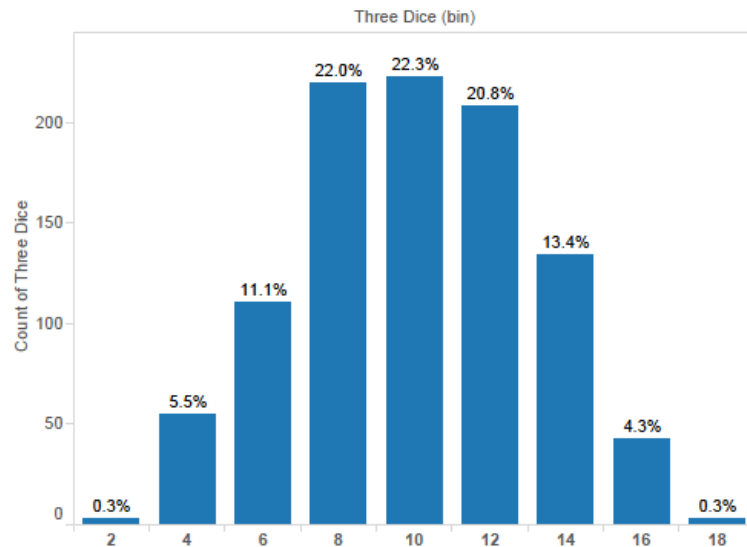
One Die



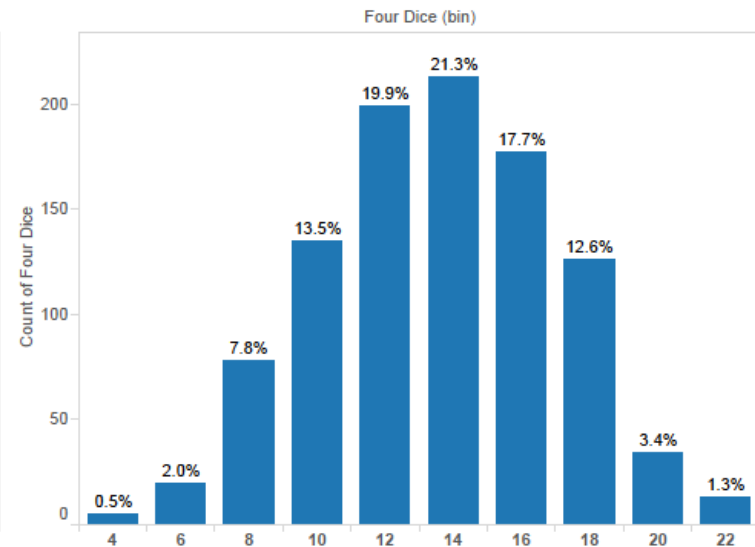
Two Dice



Three Dice



Four Dice





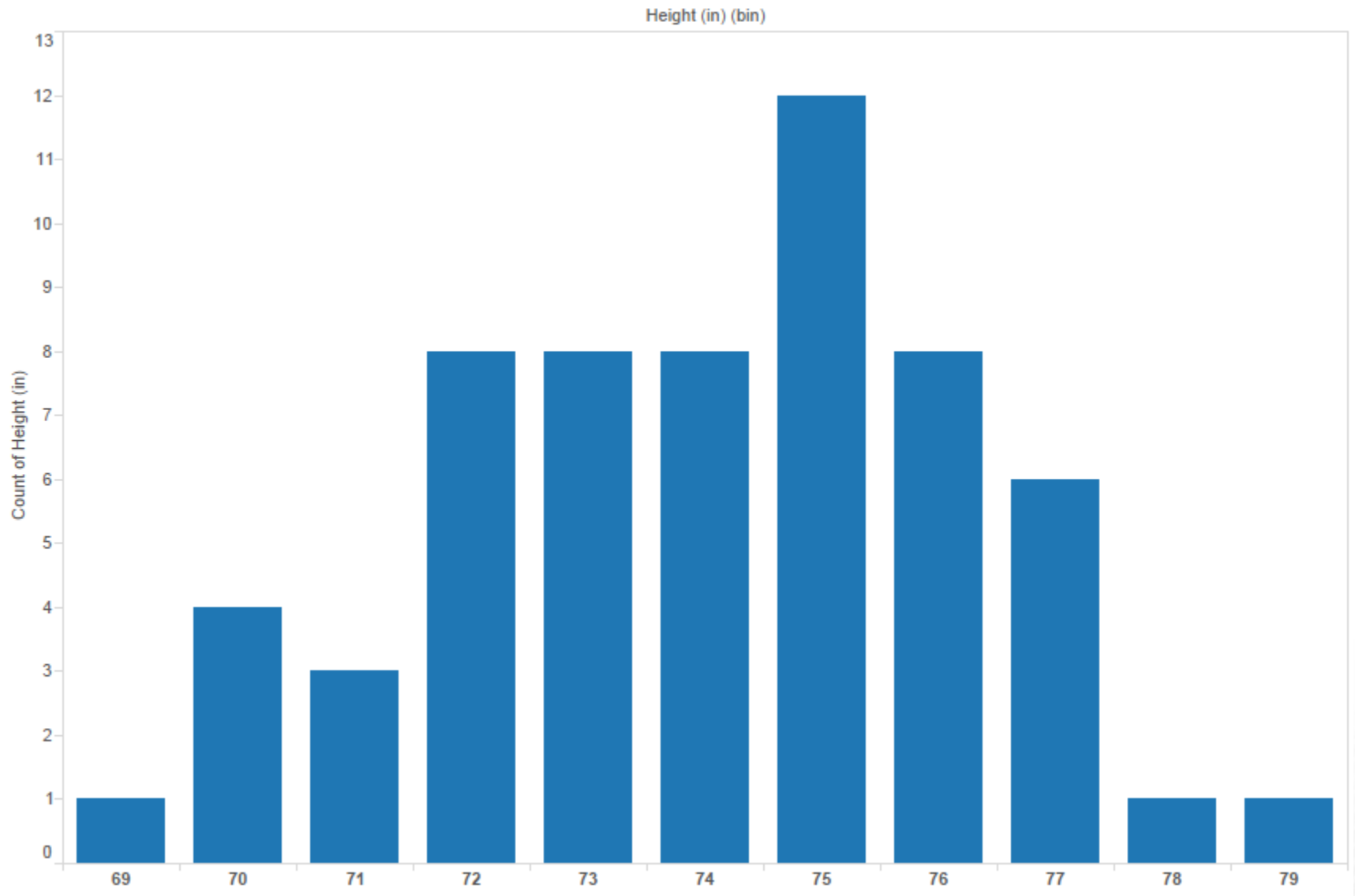
SEAHAWKS
SEATTLE

2015 Seahawks Roster

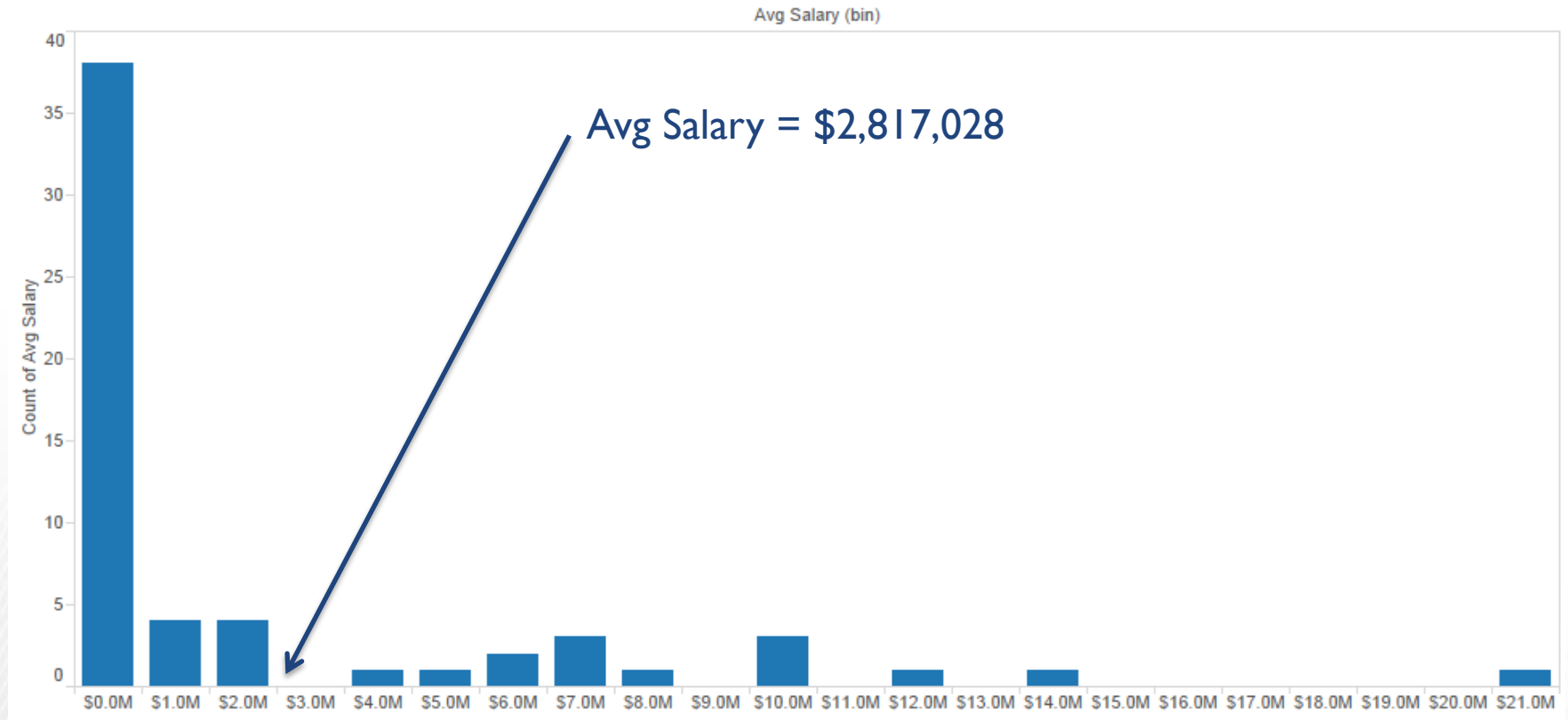
	A	B	C	D	E	F	G	H	I	J	
1	No	Player	Avg Salary	Pos	Status	Height (in)	Weight (lbs)	Birthdate	Exp	College	
2	56	Cliff Avril	\$ 7,125,000	DE	ACT	75	260	4/8/1986	8	Purdue	
3	78	Alvin Bailey	\$ 497,000	G	ACT	75	320	8/26/1991	3	Arkansas	
4	89	Doug Baldwin	\$ 4,333,333	WR	ACT	70	189	9/21/1988	5	Stanford	
5	72	Michael Bennett	\$ 7,125,000	DE	ACT	76	274	11/13/1985	7	Texas A&M	
6	68	Justin Britt	\$ 886,733	OG	ACT	78	325	5/29/1991	2	Missouri	
7	30	Bryce Brown	\$ 660,000	RB	ACT	72	220	5/14/1991	4	Kansas State	
8	28	Marcus Burley	\$ 465,000	DB	ACT	71	185	7/16/1990	3	Delaware	
9	31	Kam Chancellor	\$ 7,000,502	SS	ACT	75	232	4/3/1988	6	Virginia Tech	
10	55	Frank Clark	\$ 933,056	DE	ACT	75	272	6/14/1993	0	Michigan	
11	40	Derrick Coleman	\$ 495,000	FB	ACT	72	233	10/18/1990	4	UCLA	
12	52	Brock Coyle	\$ 511,000	MLB	RES	73	243	10/12/1990	2	Montana	
13	95	Demarcus Dobbs	\$ 825,000	DE	ACT	74	282	11/30/1987	5	Georgia	
14	79	Garry Gilliam	\$ 514,000	T	ACT	77	306	11/26/1990	2	Penn State	
15	63	Mark Glowinski	\$ 656,423	OG	ACT	76	310	5/3/1992	0	West Virginia	
16	88	Jimmy Graham	\$ 10,000,000	TE	ACT	79	265	11/24/1986	6	Miami (Fla.)	
17	49	Clint Gresham	\$ 901,667	LS	ACT	75	260	8/24/1986	6	Texas Christian	
18	4	Steven Hauschka	\$ 2,850,000	K	ACT	76	210	6/29/1985	8	North Carolina State	
19	84	Cooper Helfet	\$ 465,000	TE	ACT	75	239	6/2/1989	4	Duke	
20	97	Jordan Hill	\$ 691,035	DT	ACT	73	303	2/8/1991	3	Penn State	
21	51	Bruce Irvin	\$ 2,335,550	OLB	ACT	75	260	11/1/1987	4	West Virginia	
22	7	Tarvaris Jackson	\$ 1,500,000	QB	ACT	74	225	4/21/1983	10	Alabama State	

Data: <https://www.dropbox.com/s/po2qfa9oo6a4k6j/SeahawksRoster2014.xlsx?dl=0>

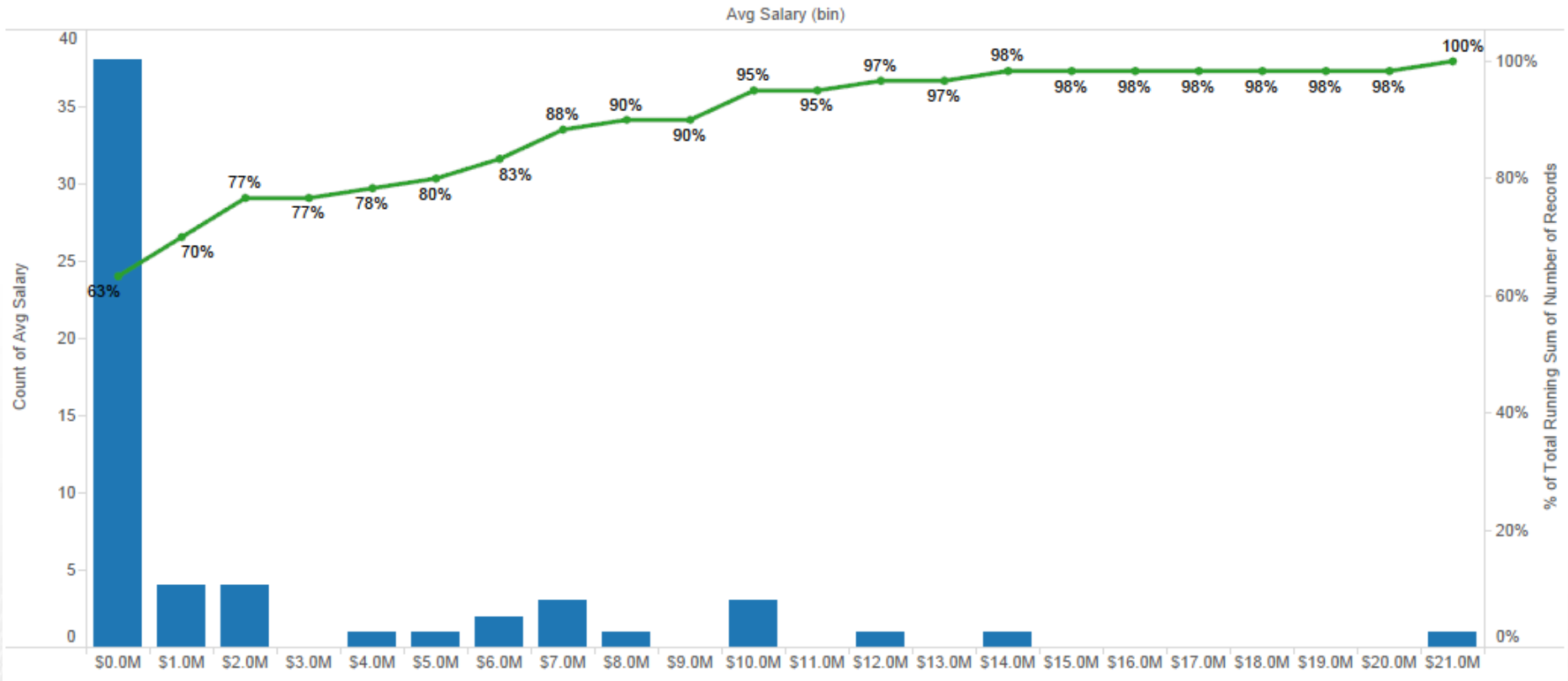
Normal: Heights of Seahawk Players



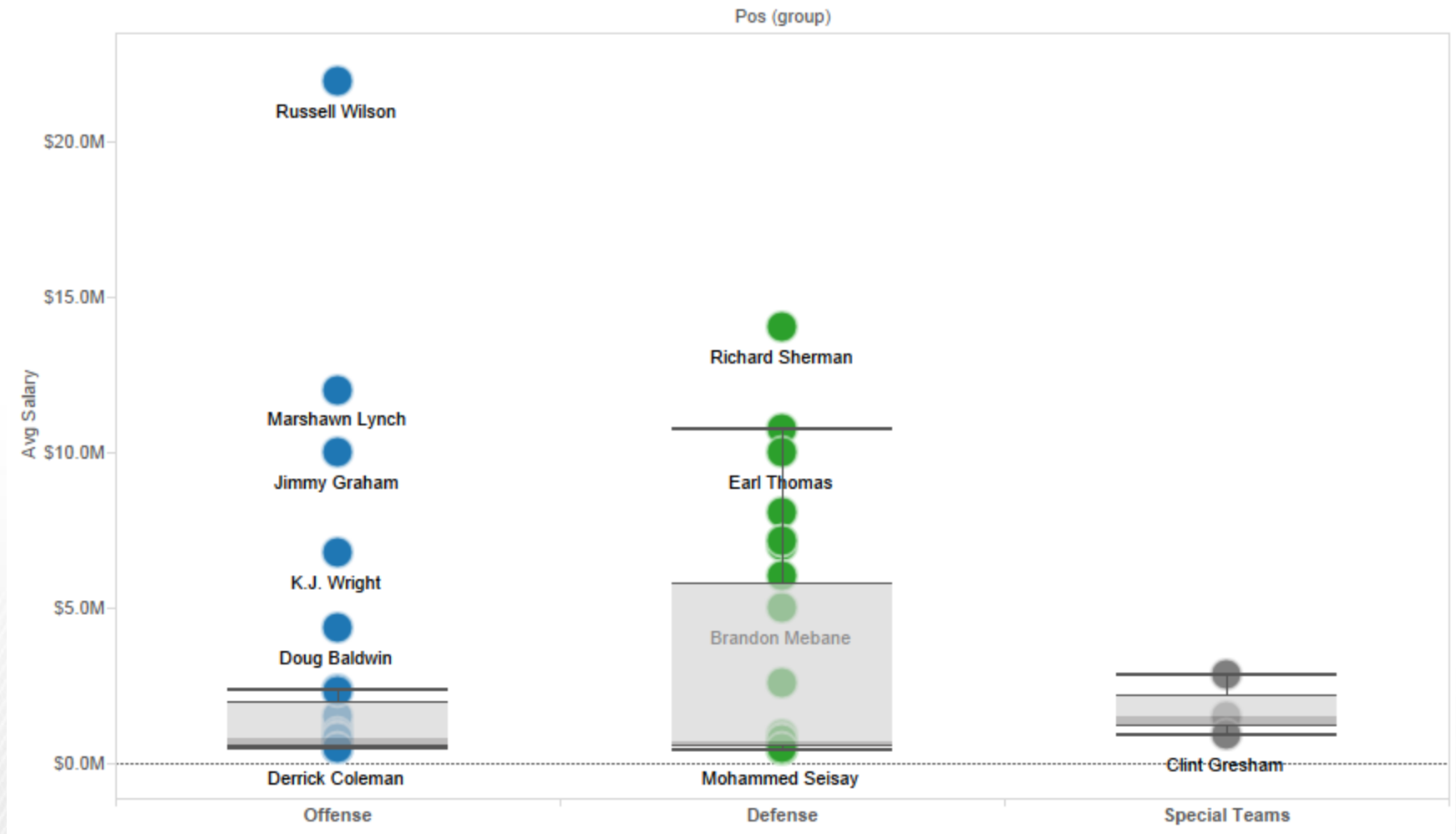
Non-normal: Seahawk Salaries



Adding a Cumulative Distribution Plot



Box Plots: Salary by Position Group



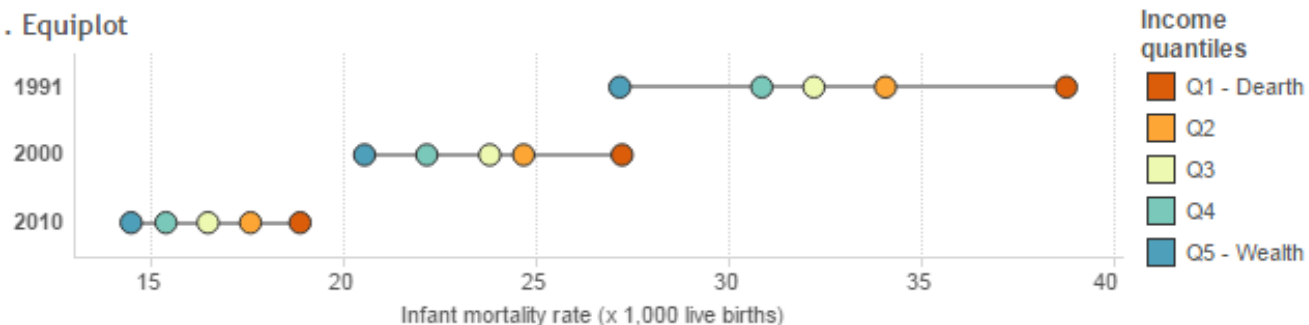
The “Equiplot”

Equiplot: Inequalities in Infant Mortality by Income

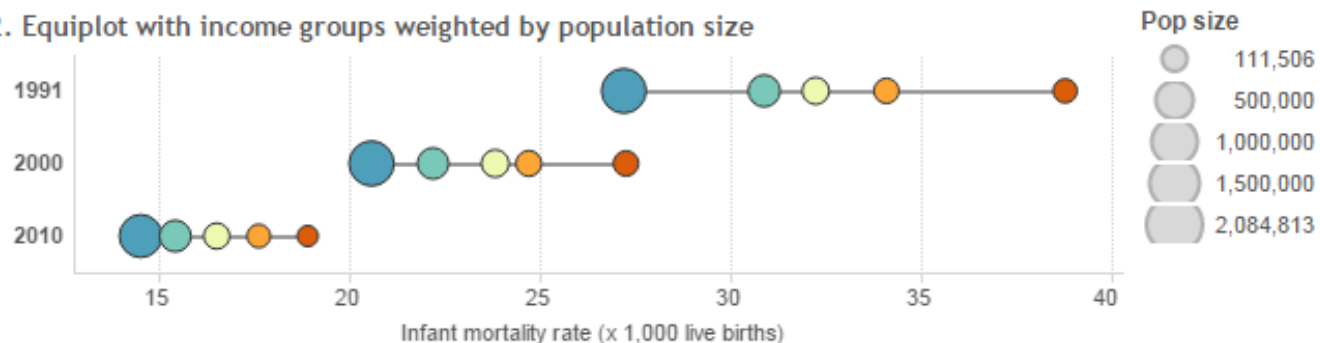
Infant mortality level and disparities across income groups and trends from 1991 to 2010

Hover the mouse over any dot and/or line for addition information, including summary disparity measures such as absolute and relative differences

Panel 1. Equiplot

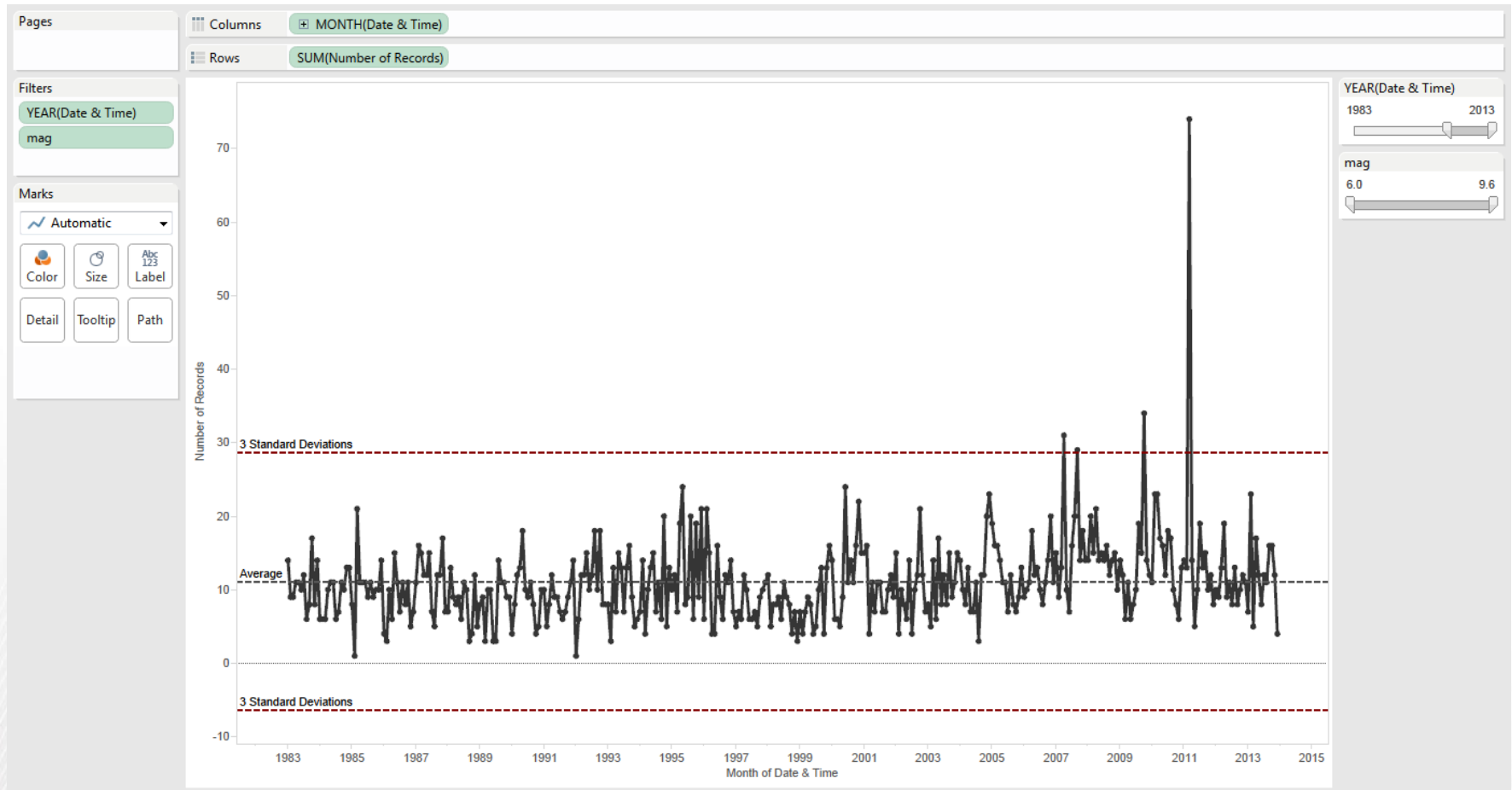


Panel 2. Equiplot with income groups weighted by population size



Source: Infant Mortality rate per 1,000 live birth. Hypothetical data to illustrate the Equiplot chart.

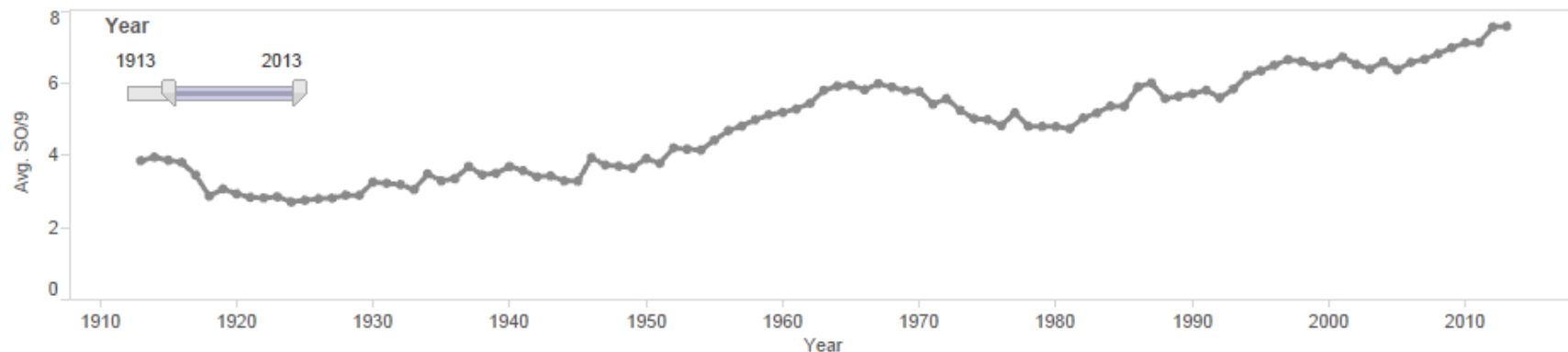
Variation over time: Control Charts



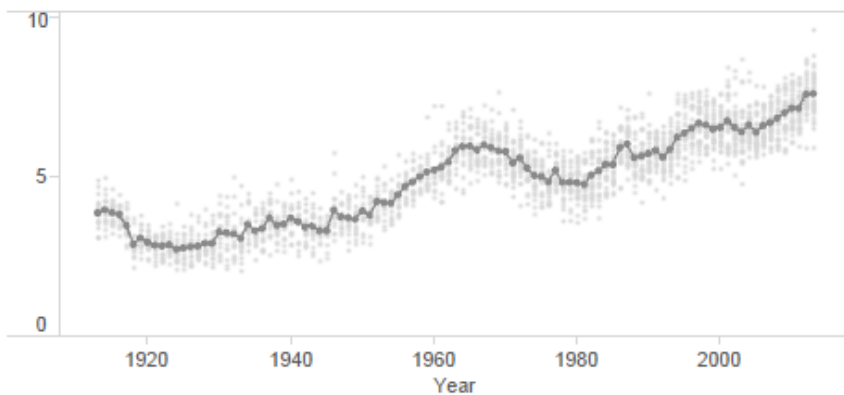
Don't "hit them over the head" with variation



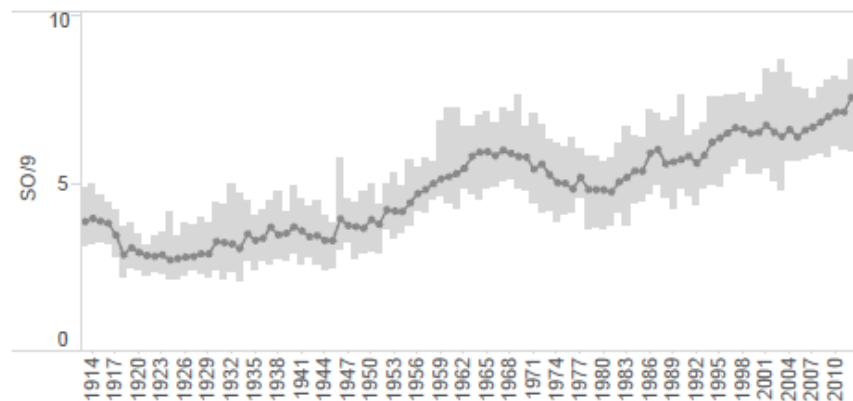
Average Only



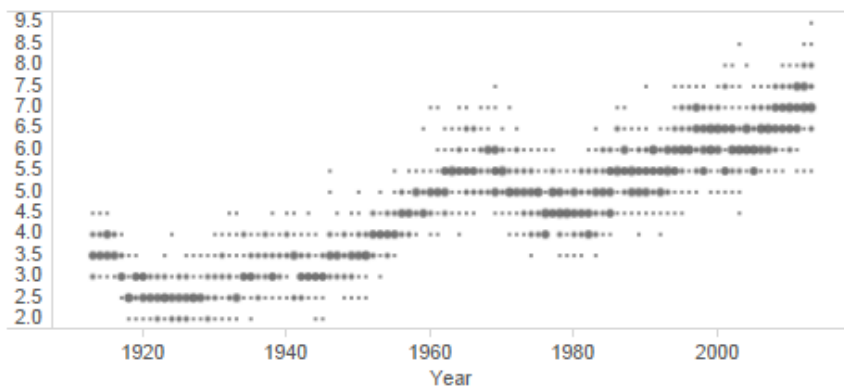
Data Points



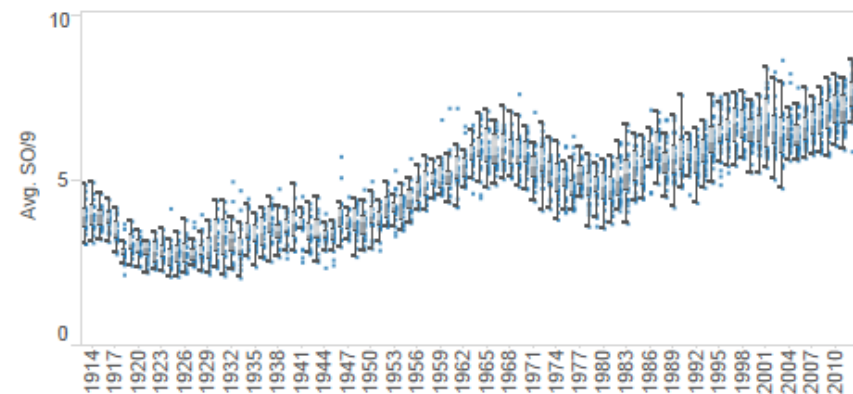
Bands



Circle Histograms



Box Plot



Exercise:
How Tall are We?

Find the mean, median, mode,
standard deviation, and quartiles

Visualize Class Heights

Form: <http://bit.ly/1iByFRM>

Data: <http://bit.ly/1OmOavb>

Named by *Fortune*
ONE OF THE SMARTEST BOOKS OF ALL TIME

FOLLOLED *BY* *RANDOMNESS*

*The Hidden Role of Chance
in Life and in the Markets*

NASSIM NICHOLAS TALEB
SECOND EDITION, UPDATED BY THE AUTHOR



Two types of variation



**DON'T BE A
TURKEY!**



Week 4 Homework

- Read the chapters for the Week 5 Lecture:
 - Munzner, *Visualization Analysis & Design*, Ch. 6
 - Jones, *Communicating Data with Tableau*, Ch. 8
- Visualization
 - Find (or create) a data set and create a visualization that clearly shows measures of central tendency and variation