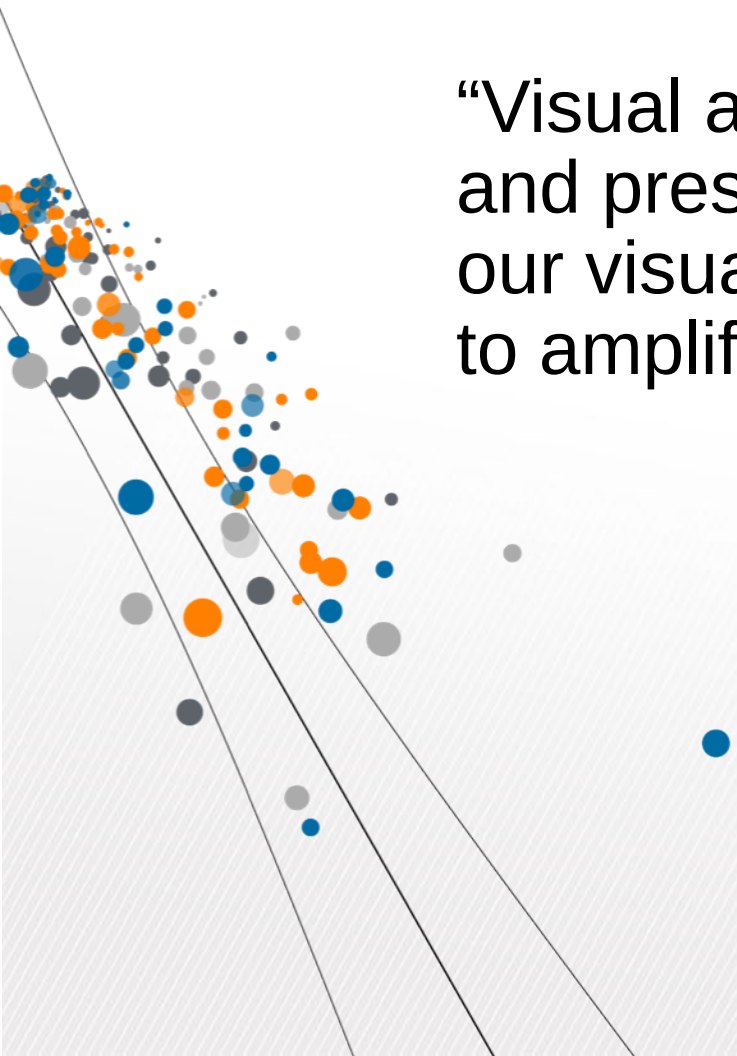
The Science of Visual Analysis

What is Analytics?

Analytics is defined as the scientific process of transforming data into insight for making better decisions.

Analytics is not just the description of information, but goes beyond that into examination and explanation of this information.

What is Visual Analytics?



"Visual analytics is the representation and presentation of data that exploits our visual perception abilities in order to amplify cognition."

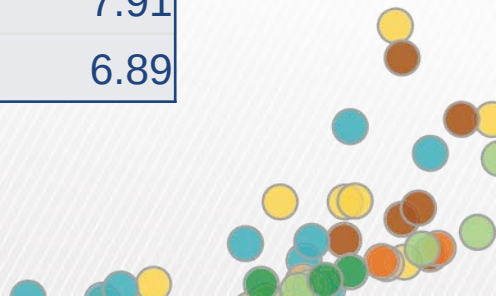
- Andy Kirk, author of "Data Visualization: a successful design process"

3 3 0 3 0 1 8 7 6 8 2 1 4 0 3 8 3 7 7 2 0 5 2 3 2 7 0 2 0
7 1 4 6 0 2 1 3 2 7 6 0 2 5 6 3 2 5 7 6 3 3 0 2 0 3 0 7 2
8 7 5 7 2 8 3 8 7 7 8 2 0 7 7 5 2 3 1 1 5 6 3 8 4 7 8 2 0
0 5 0 5 1 6 1 7 5 6 8 0 4 4 6 7 4 7 1 4 0 0 8 4 4 3 0 3 2
2 4 3 1 3 5 4 9 5 0 7 6 0 7 4 3 1 8 2 7 3 4 6 0 2 4 8 2 3
8 6 2 2 6 5 4 6 7 0 7 6 0 0 3 9 0 2 4 7 1 7 2 3 3 5 8 7 0
0 8 4 5 1 3 1 7 6 4 5 4 1 2 4 5 3 3 5 4 9 6 7 7 6 3 4 2 5
4 7 7 0 2 2 0 1 1 7 7 7 0 2 6 6 4 7 5 8 6 1 4 3 7 8 5 4 6
4 3 6 6 4 6 6 2 8 4 8 5 3 7 8 8 1 3 8 5 4 5 7 4 0 3 2 8 4
5 5 0 3 5 3 5 3 8 3 2 3 8 2 3 1 6 2 7 2 4 6 3 6 4 4 3 2 5
4 4 0 2 1 7 2 4 4 7 4 1 9 2 4 5 2 5 0 4 0 0 5 3 6 3 3 6 7
7 4 6 6 8 7 5 7 9 2 0 2 8 8 8 8 3 2 4 2 6 4 0 4 6 3 7 2 1
0 1 7 1 5 9 1 4 2 8 7 3 7 1 4 5 1 8 7 8 0 5 1 7 0 5 8 8 1
2 8 5 2 1 2 8 7 7 6 2 5 6 2 6 4 1 5 1 6 1 2 1 1 0 5 6 4 0
2 1 1 7 7 2 0 0 1 8 7 0 2 9 0 2 8 5 7 8 4 6 0 6 5 0 7 1 2
0 5 2 4 1 5 3 3 1 5 5 1 4 0 1 6 4 3 3 9 8 8 3 4 6 8 4 8 6
7 3 7 5 2 4 0 2 7 6 3 8 5 5 4 5 8 8 7 5 5 6 5 6 7 9 7 7 4
0 3 2 8 1 4 4 6 0 8 2 3 0 1 3 4 6 2 0 5 7 7 3 6 1 8 7 3 5
4 4 8 3 3 3 5 0 1 0 3 8 6 3 2 0 5 0 6 1 3 3 4 3 6 1 5 8 6
1 0 2 2 7 6 3 3 0 8 8 0 3 1 8 8 1 2 1 7 5 2 9 3 5 8 3 2 5

3 3 0 3 0 1 8 7 6 8 2 1 4 0 3 8 3 7 7 2 0 5 2 3 2 7 0 2 0
7 1 4 6 0 2 1 3 2 7 6 0 2 5 6 3 2 5 7 6 3 3 0 2 0 3 0 7 2
8 7 5 7 2 8 3 8 7 7 8 2 0 7 7 5 2 3 1 1 5 6 3 8 4 7 8 2 0
0 5 0 5 1 6 1 7 5 6 8 0 4 4 6 7 4 7 1 4 0 0 8 4 4 3 0 3 2
2 4 3 1 3 5 4 9 5 0 7 6 0 7 4 3 1 8 2 7 3 4 6 0 2 4 8 2 3
8 6 2 2 6 5 4 6 7 0 7 6 0 0 3 9 0 2 4 7 1 7 2 3 3 5 8 7 0
0 8 4 5 1 3 1 7 6 4 5 4 1 2 4 5 3 3 5 4 9 6 7 7 6 3 4 2 5
4 7 7 0 2 2 0 1 1 7 7 7 0 2 6 6 4 7 5 8 6 1 4 3 7 8 5 4 6
4 3 6 6 4 6 6 2 8 4 8 5 3 7 8 8 1 3 8 5 4 5 7 4 0 3 2 8 4
5 5 0 3 5 3 5 3 8 3 2 3 8 2 3 1 6 2 7 2 4 6 3 6 4 4 3 2 5
4 4 0 2 1 7 2 4 4 7 4 1 9 2 4 5 2 5 0 4 0 0 5 3 6 3 3 6 7
7 4 6 6 8 7 5 7 9 2 0 2 8 8 8 8 3 2 4 2 6 4 0 4 6 3 7 2 1
0 1 7 1 5 9 1 4 2 8 7 3 7 1 4 5 1 8 7 8 0 5 1 7 0 5 8 8 1
2 8 5 2 1 2 8 7 7 6 2 5 6 2 6 4 1 5 1 6 1 2 1 1 0 5 6 4 0
2 1 1 7 7 2 0 0 1 8 7 0 2 9 0 2 8 5 7 8 4 6 0 6 5 0 7 1 2
0 5 2 4 1 5 3 3 1 5 5 1 4 0 1 6 4 3 3 9 8 8 3 4 6 8 4 8 6
7 3 7 5 2 4 0 2 7 6 3 8 5 5 4 5 8 8 7 5 5 6 5 6 7 9 7 7 4
0 3 2 8 1 4 4 6 0 8 2 3 0 1 3 4 6 2 0 5 7 7 3 6 1 8 7 3 5
4 4 8 3 3 3 5 0 1 0 3 8 6 3 2 0 5 0 6 1 3 3 4 3 6 1 5 8 6
1 0 2 2 7 6 3 3 0 8 8 0 3 1 8 8 1 2 1 7 5 2 9 3 5 8 3 2 5

Let’s Look at Some Data

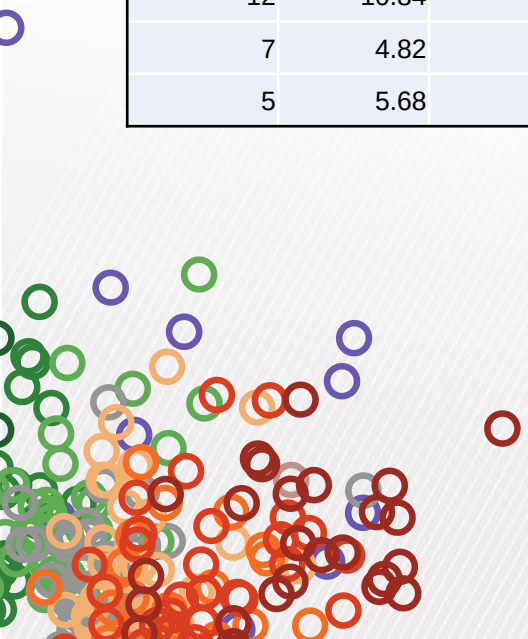
I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.1	4	5.39	19	12.5
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89



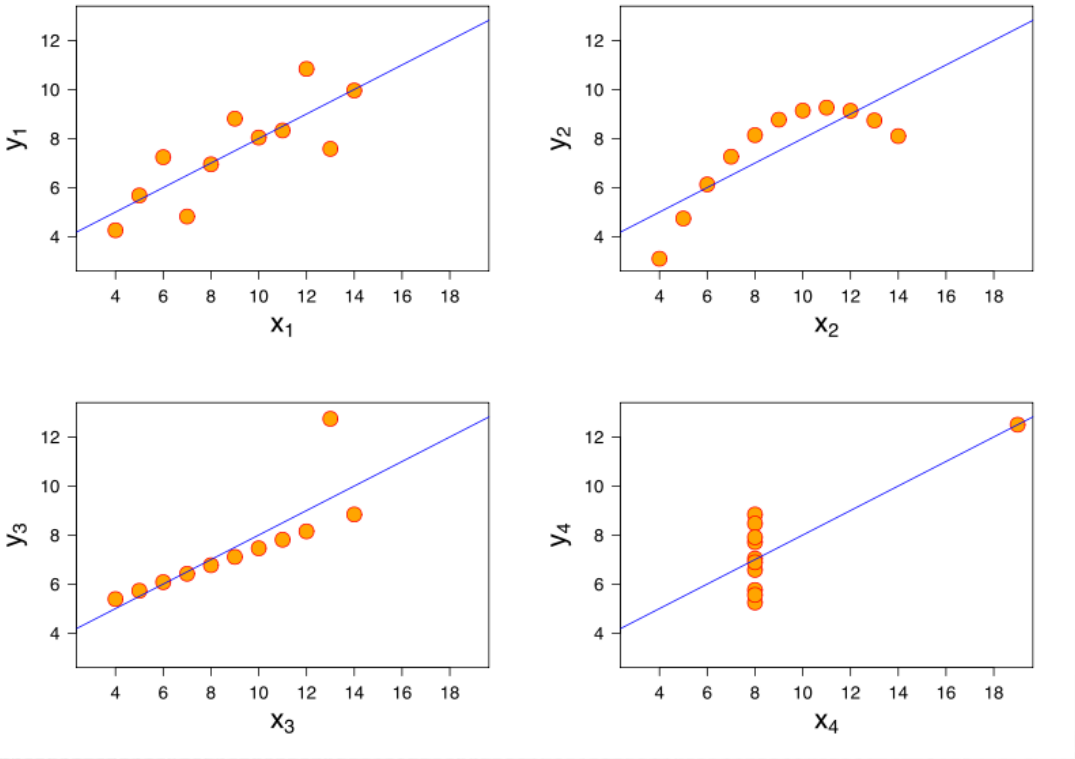
Let’s Analyze Some Data

I		II		III		IV	
x	y	x	y	x	y	x	y
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.1				
6	7.24	6	6.13				
4	4.26	4	3.1				
12	10.84	12	9.13				
7	4.82	7	7.26				
5	5.68	5	4.74				

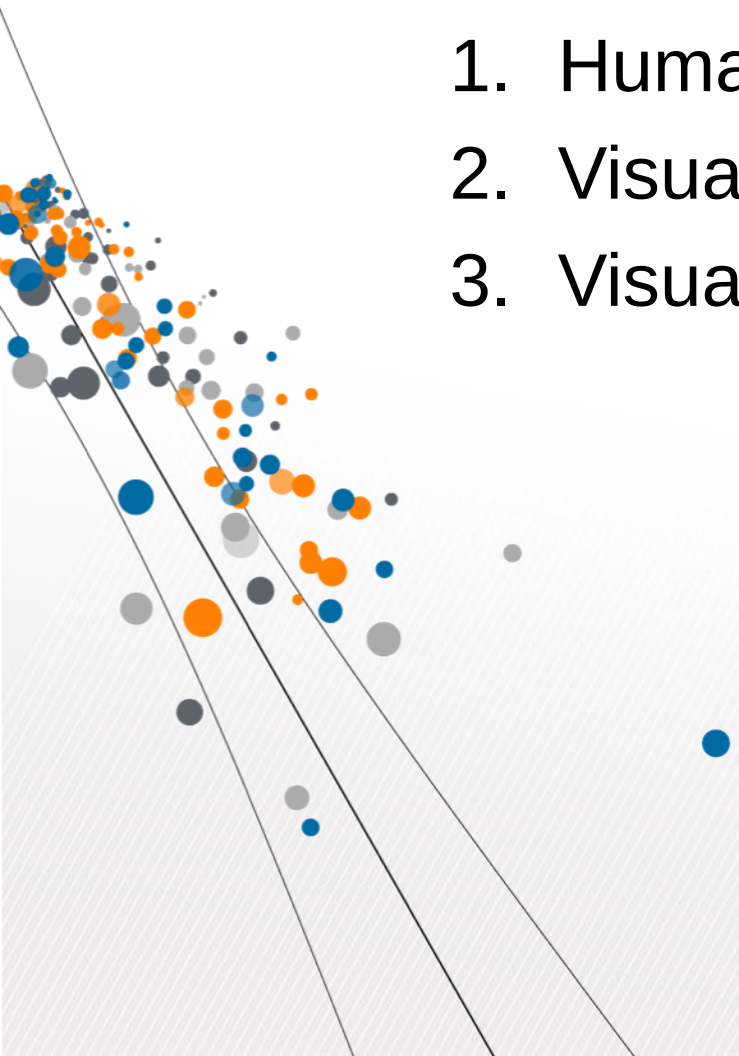
Property	Value
Mean of x in each case	9 (exact)
Variance of x in each case	11 (exact)
Mean of y in each case	7.50 (to 2 decimal places)
Variance of y in each case	4.122 or 4.127 (to 3 decimal places)
Correlation between x and y in each case	0.816 (to 3 decimal places)
Linear regression line in each case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



Let’s Look at Some Data … Visually



“Anscombe’s Quartet”
Source: Wikipedia



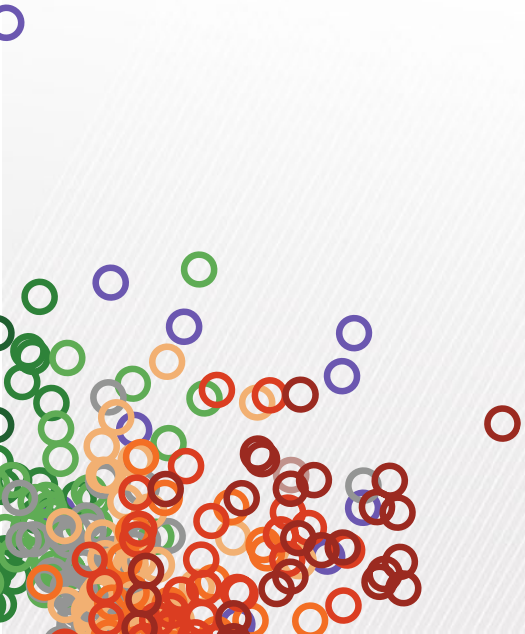
Agenda

1. Human Perception and Cognition
2. Visual Analysis Cycle
3. Visualization Best Practices

Human Perception & Cognition

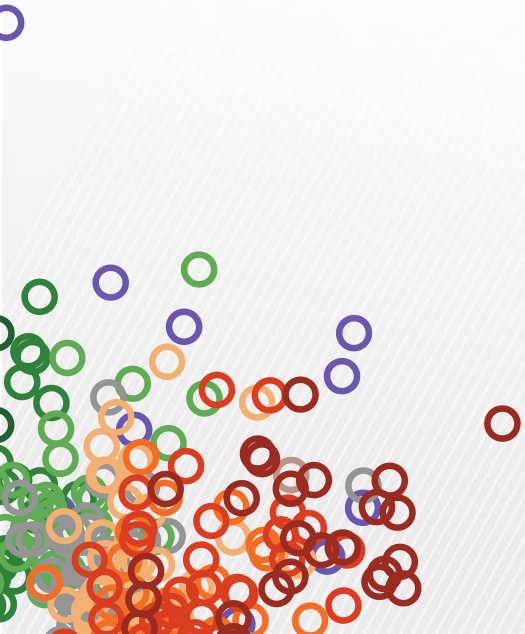
Humans Are Good at some Mental Math

$$\begin{array}{r} 3 \\ \times\ 7 \\ \hline \end{array}$$



Humans Are not so Good at other Mental Math

$$\begin{array}{r} 34 \\ \times\ 72 \\ \hline \end{array}$$

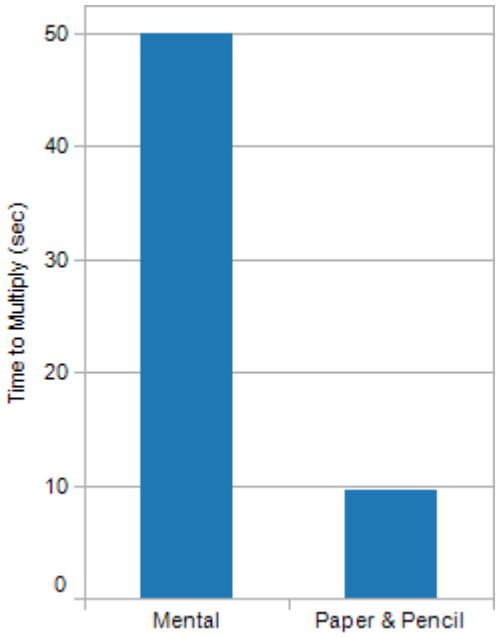


We’re Faster When We Use the Learned Tools and Techniques

$$
\begin{array}{r}
34 \\
\times\ 72 \\
\hline
68 \\
23^{1}80 \\
\hline
2448
\end{array}
$$

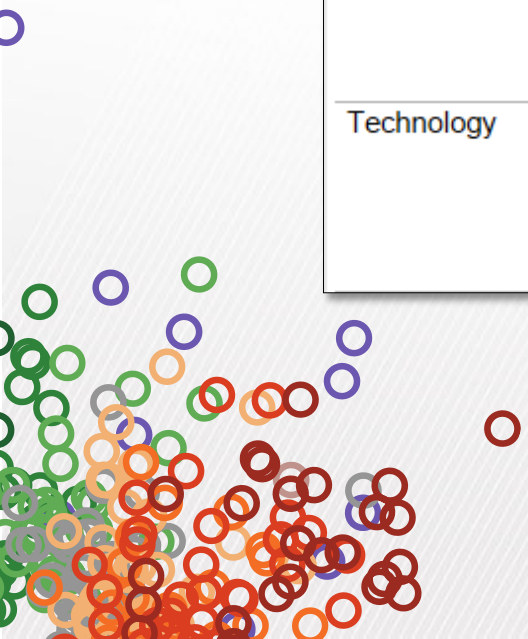
Much Faster

$$\begin{array}{r} 34 \\ \times\ 72 \\ \hline 68 \\ 23^{1}80 \\ \hline 2448 \end{array}$$



We're Faster When We Can "See" Data

		Customer Segment			
Category	Sub-Category (group)	Consumer	Corporate	Home Office	Small Business
Furniture	Bookcases	-63.02	-9,305.76	-16,610.95	-7,602.40
	Chairs & Chairmats	42,942.97	39,370.10	41,686.28	25,650.38
	Office Furnishings	12,099.80	27,374.47	42,196.25	18,757.40
	Tables	-12,251.51	-35,430.73	-43,292.40	-8,087.89
Office Supplies	Appliances	15,501.48	50,095.94	25,343.06	6,217.58
	Binders and Binder Ac..	48,035.27	125,811.27	71,674.19	61,892.69
	Envelopes, Labels, Pa..	16,907.52	31,230.67	25,508.13	33,476.65
	Pens & Art Supplies	2,621.68	1,670.40	1,580.82	1,691.88
	Rubber Bands	271.85	-353.54	-93.12	72.14
	Scissors, Rulers and ..	-558.10	-3,330.62	-2,844.06	-1,066.47
	Storage & Organization	5,752.65	-2,086.83	-23.24	3,021.57
Technology	Computer Peripherals	14,152.79	45,092.93	17,771.05	17,270.71
	Copiers and Fax	41,310.35	28,654.48	29,283.14	68,113.50
	Office Machines	51,454.78	180,356.22	39,386.23	36,515.70
	Telephones and Com..	49,781.48	120,596.92	86,788.72	59,784.52



We're Faster When We Can "See" Data

		Customer Segment			
Category	Sub-Category (group)	Consumer	Corporate	Home Office	Small Business
Furniture	Bookcases	-63.02	-9,305.76	-16,610.95	-7,602.40
	Chairs & Chairmats	42,942.97	39,370.10	41,686.28	25,650.38
	Office Furnishings	12,099.80	27,374.47	42,196.25	18,757.40
	Tables	-12,251.51	-35,430.73	-43,292.40	-8,087.89
Office Supplies	Appliances	15,501.48	50,095.94	25,343.06	6,217.58
	Binders and Binder Ac..	48,035.27	125,811.27	71,674.19	61,892.69
	Envelopes, Labels, Pa..	16,907.52	31,230.67	25,508.13	33,476.65
	Pens & Art Supplies	2,621.68	1,670.40	1,580.82	1,691.88
	Rubber Bands	271.85	-353.54	-93.12	72.14
	Scissors, Rulers and ..	-558.10	-3,330.62	-2,844.06	-1,066.47
	Storage & Organization	5,752.65	-2,086.83	-23.24	3,021.57
Technology	Computer Peripherals	14,152.79	45,092.93	17,771.05	17,270.71
	Copiers and Fax	41,310.35	28,654.48	29,283.14	68,113.50
	Office Machines	51,454.79	180,356.22	39,386.23	36,515.70
	Telephones and Com..	49,781.48	120,596.92	86,788.72	59,784.52

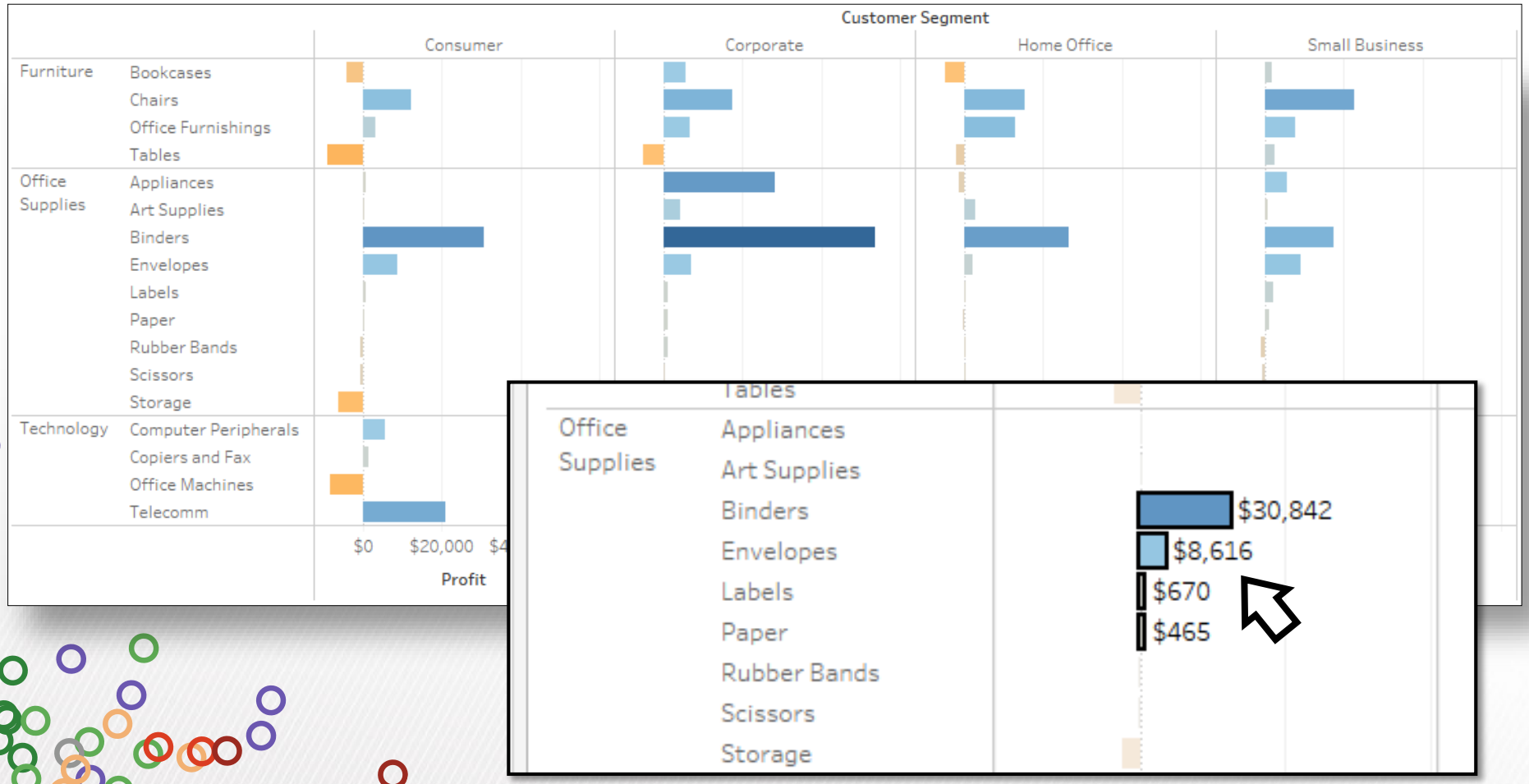
We're Faster When We Can "See" Data

		Customer Segment			
Category	Sub-Category (group)	Consumer	Corporate	Home Office	Small Business
Furniture	Bookcases	-63.02	-9,305.76	-16,610.95	-7,602.40
	Chairs & Chairmats	42,942.97	39,370.10	41,686.28	25,650.38
	Office Furnishings	12,099.80	27,374.47	42,196.25	18,757.40
	Tables	-12,251.51	-35,430.73	-43,292.40	-8,087.89
Office Supplies	Appliances	15,501.48	50,095.94	25,343.06	6,217.58
	Binders and Binder Ac..	48,035.27	125,811.27	71,674.19	61,892.69
	Envelopes, Labels, Pa..	16,907.52	31,230.67	25,508.13	33,476.65
	Pens & Art Supplies	2,621.68	1,670.40	1,580.82	1,691.88
	Rubber Bands	271.85	-353.54	-93.12	72.14
	Scissors, Rulers and ..	-558.10	-3,330.62	-2,844.06	-1,066.47
	Storage & Organization	5,752.65	-2,086.83	-23.24	3,021.57
Technology	Computer Peripherals	14,152.79	45,092.93	17,771.05	17,270.71
	Copiers and Fax	41,310.35	28,654.48	29,283.14	68,113.50
	Office Machines	51,454.78	180,356.22	39,386.23	36,515.70
	Telephones and Com..	49,781.48	120,596.92	86,788.72	59,784.52

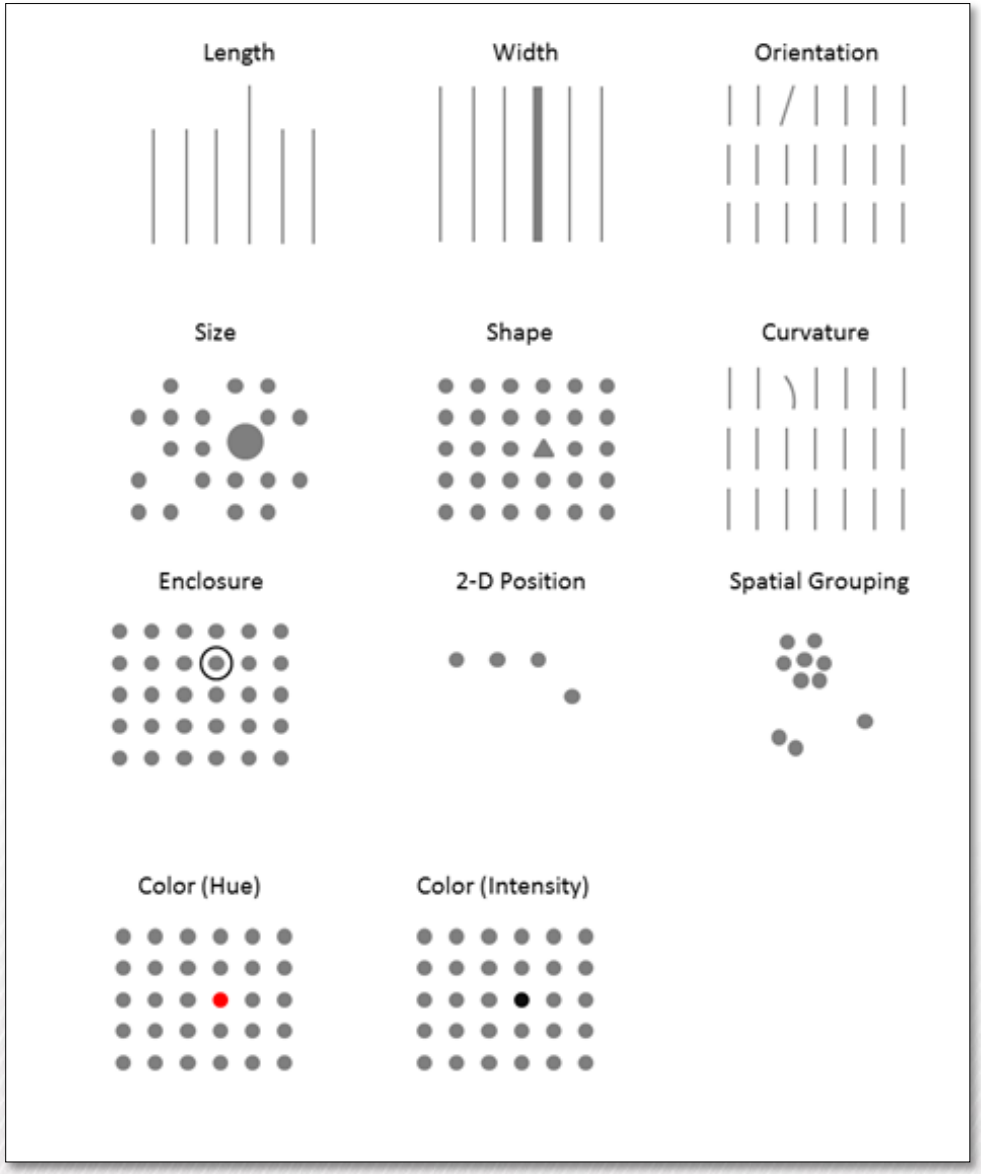
We're Faster When We Can "See" Data

		Customer Segment			
Category	Sub-Category	Consumer	Corporate	Home Office	Small Business
Furniture	Bookcases	-63.02	-9,305.76	-16,610.95	-7,602.40
	Chairs & Chairmats	42,942.97	39,370.10	41,686.28	25,650.38
	Office Furnishings	12,099.80	27,374.47	42,196.25	18,757.40
	Tables	-12,251.51	-35,430.73	-43,292.40	-8,087.89
Office Supplies	Appliances	15,501.48	50,095.94	25,343.06	6,217.58
	Binders and Binder Ac..	48,035.27	125,811.27	71,674.19	61,892.69
	Envelopes	6,731.55	15,082.58	10,848.34	15,520.13
	Labels	1,349.23	5,608.87	3,073.87	3,645.20
	Paper	8,826.74	10,539.22	11,585.92	14,311.32
	Pens & Art Supplies	2,621.68	1,670.40	1,580.82	1,691.88
	Rubber Bands	271.85	-353.54	-93.12	72.14
	Scissors, Rulers and ..	-558.10	-3,330.62	-2,844.06	-1,066.47
	Storage & Organization	5,752.65	-2,086.83	-23.24	3,021.57
Technology	Computer Peripherals	14,152.79	45,092.93	17,771.05	17,270.71
	Copiers and Fax	41,310.35	28,654.48	29,283.14	68,113.50
	Office Machines	51,454.78	180,356.22	39,386.23	36,515.70
	Telephones and Com..	49,781.48	120,596.92	86,788.72	59,784.52

We're Faster When We Can "See" Data

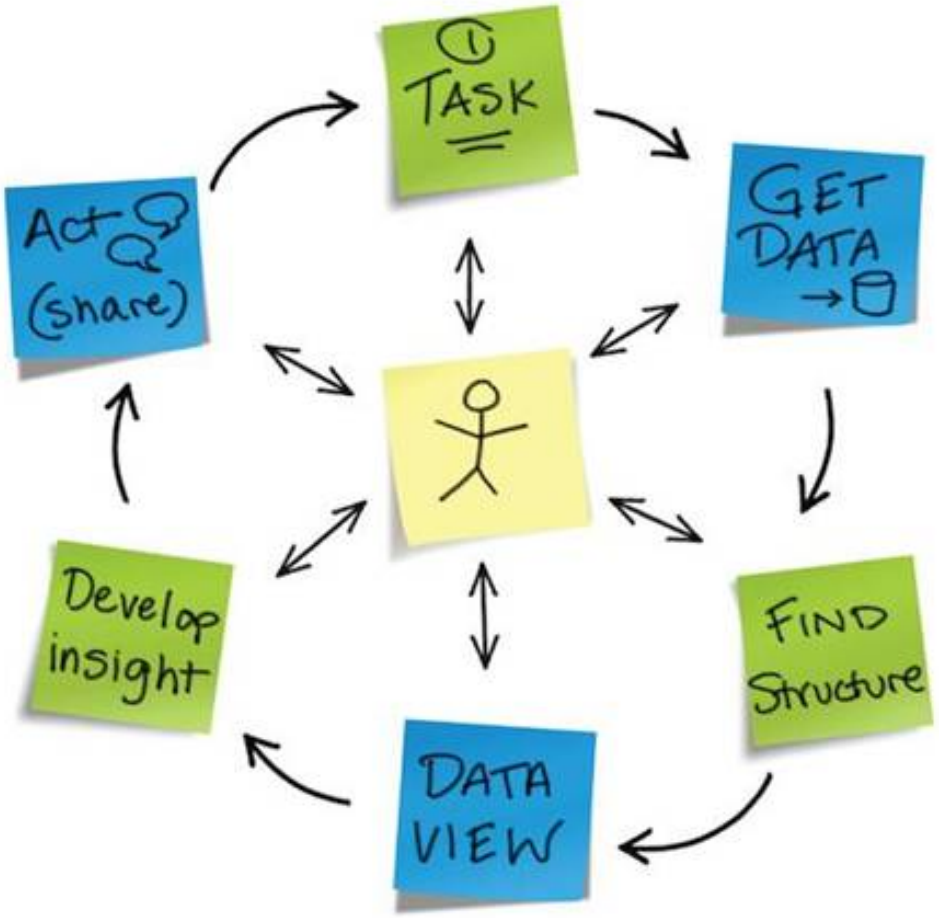


Pre-attentive Visual Attributes



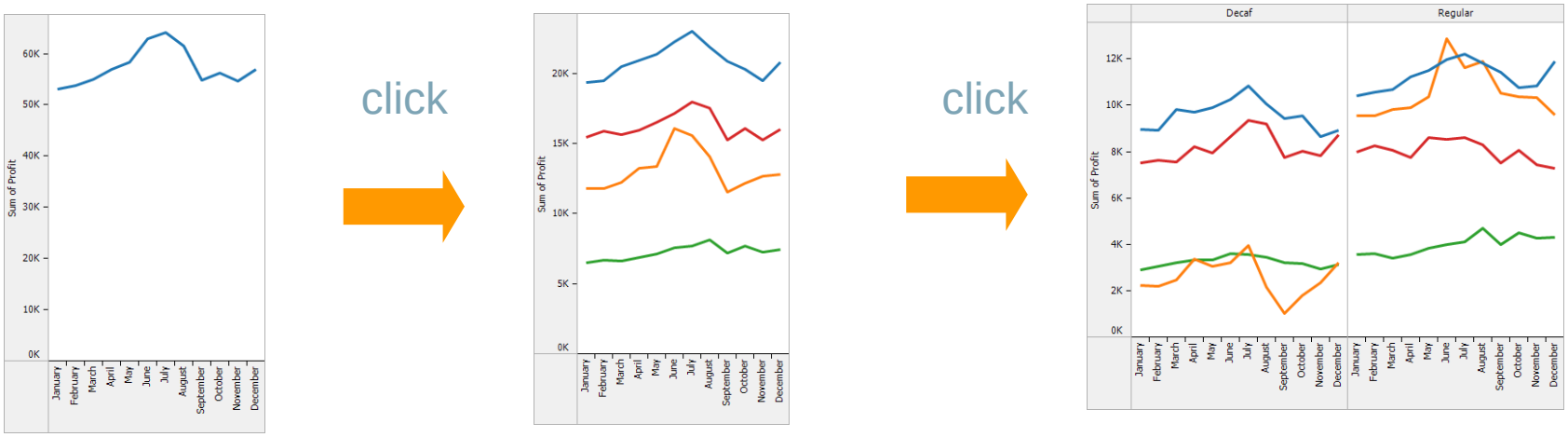
The Cycle of Visual Analysis

The Cycle of Visual Analysis



Supporting the Cycle

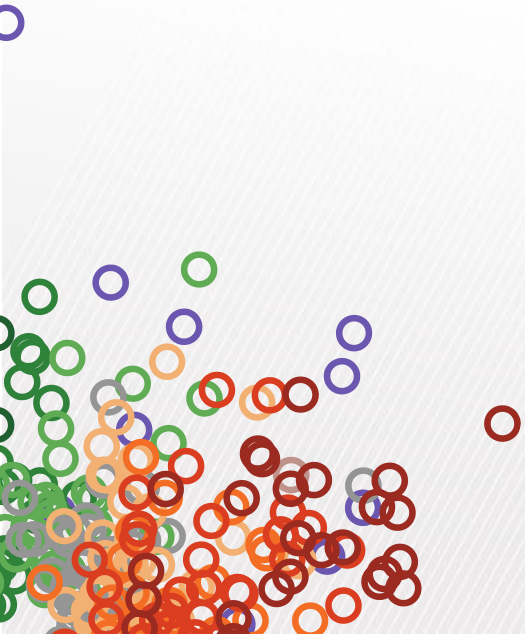
- *Incremental*: allow people to easily and incrementally **change the data** and **how they are looking at it**
- *Expressive:* there is **no single view** for all tasks and all data
- *Unified*: leverage the **revolutionary changes** in database technology
- *Direct*: **make the tool disappear** so the user can **directly interact** with the data



Visualization Best Practices

Best Practices Overview

1. Representing data for humans
2. Color
3. Maps
4. Creating dashboards



Types of Data

Qualitative (nominal)

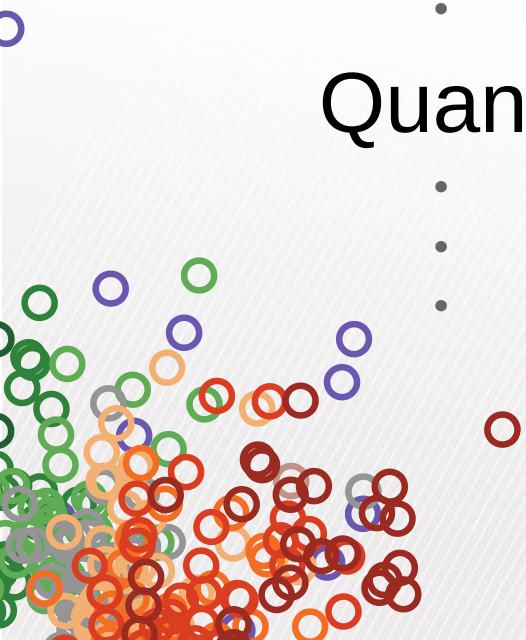
- Arizona, New York, Texas
- Sarah, John, Maria
- Coors, Bud Light, Stella Artois

Qualitative (ordinal)

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

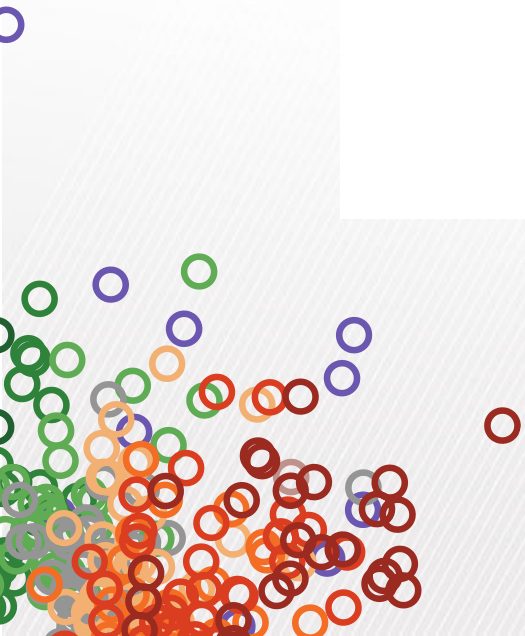
Quantitative

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

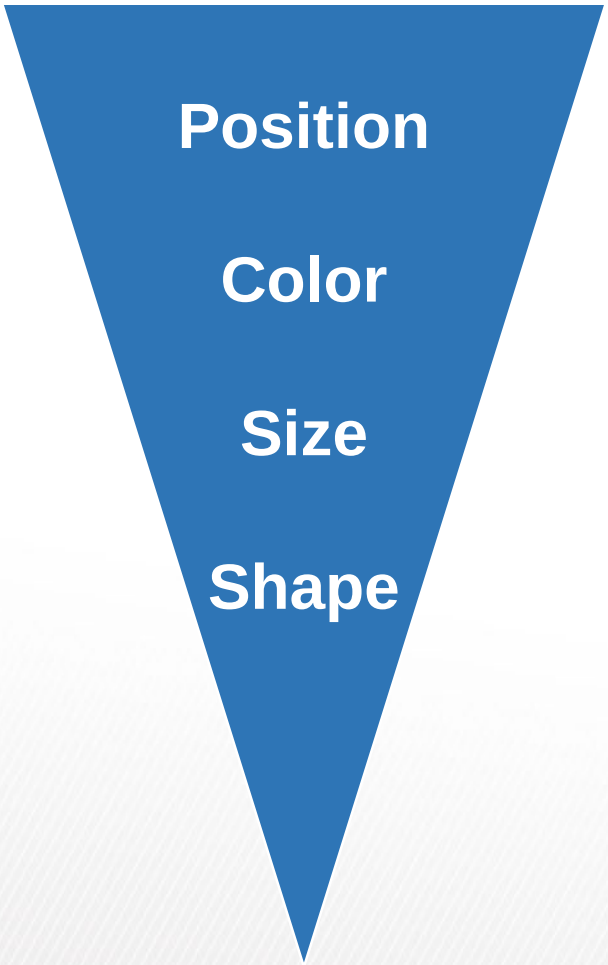


How Do Humans Like Their Data?

Quantitative	Ordinal	Categorical
Position	Position	Position
Length	Size	Shape
Size	Color Intensity	Different Colors
Color Intensity	*Different Colors	
	*Shape	

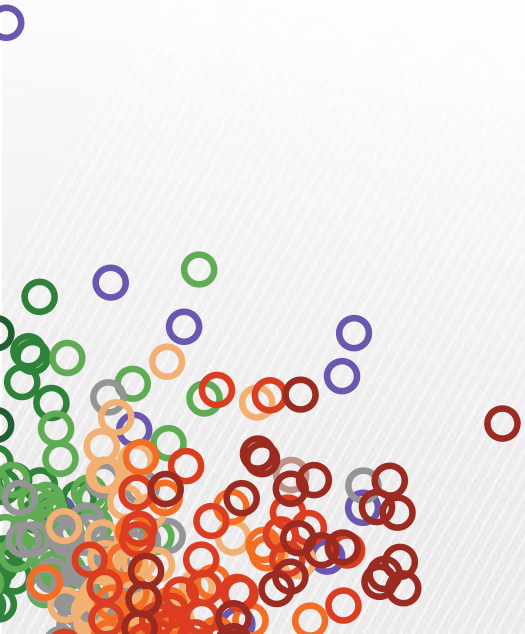


How Do Humans Like Their Data?



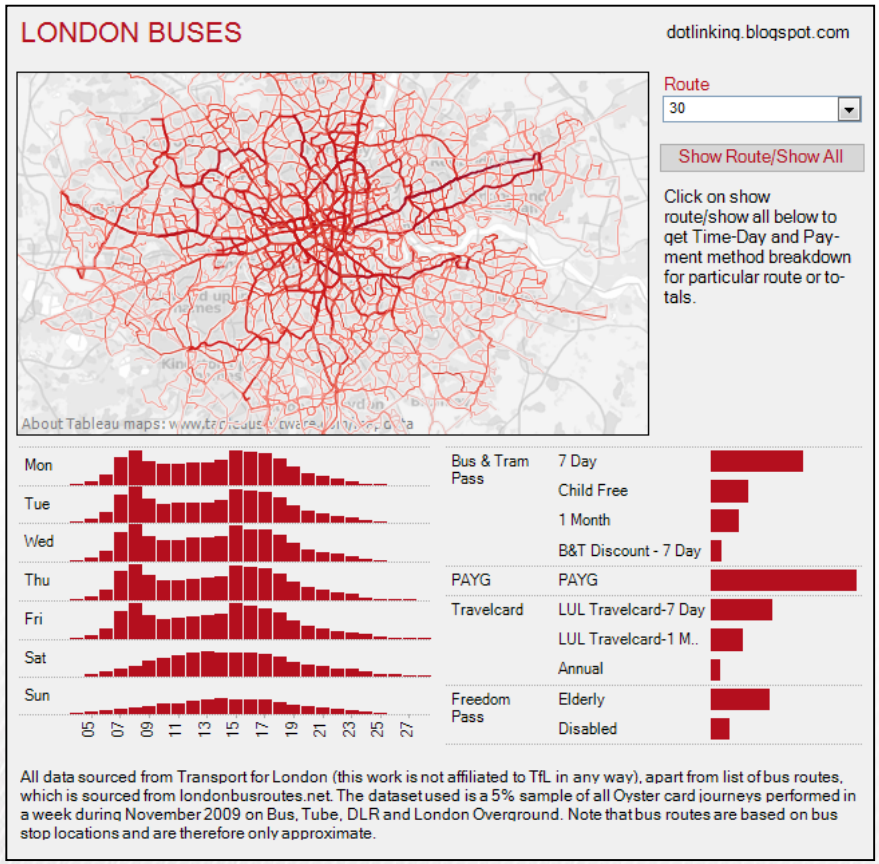
More important

Less important



How Do Humans Like Their Data?

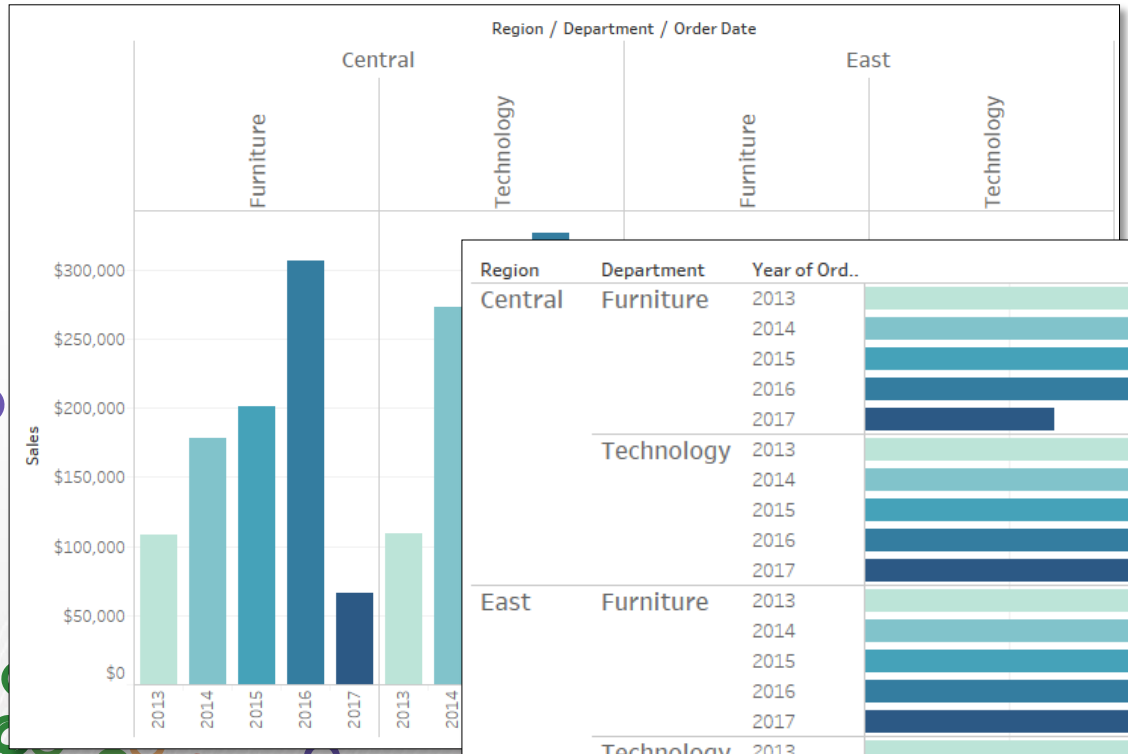
- *Exploring relationships*: scatter plot
- *Relative proportions*: treemap
- *Comparing values*: bar chart
- *Time*: on an x-axis
- *Location*: on a map



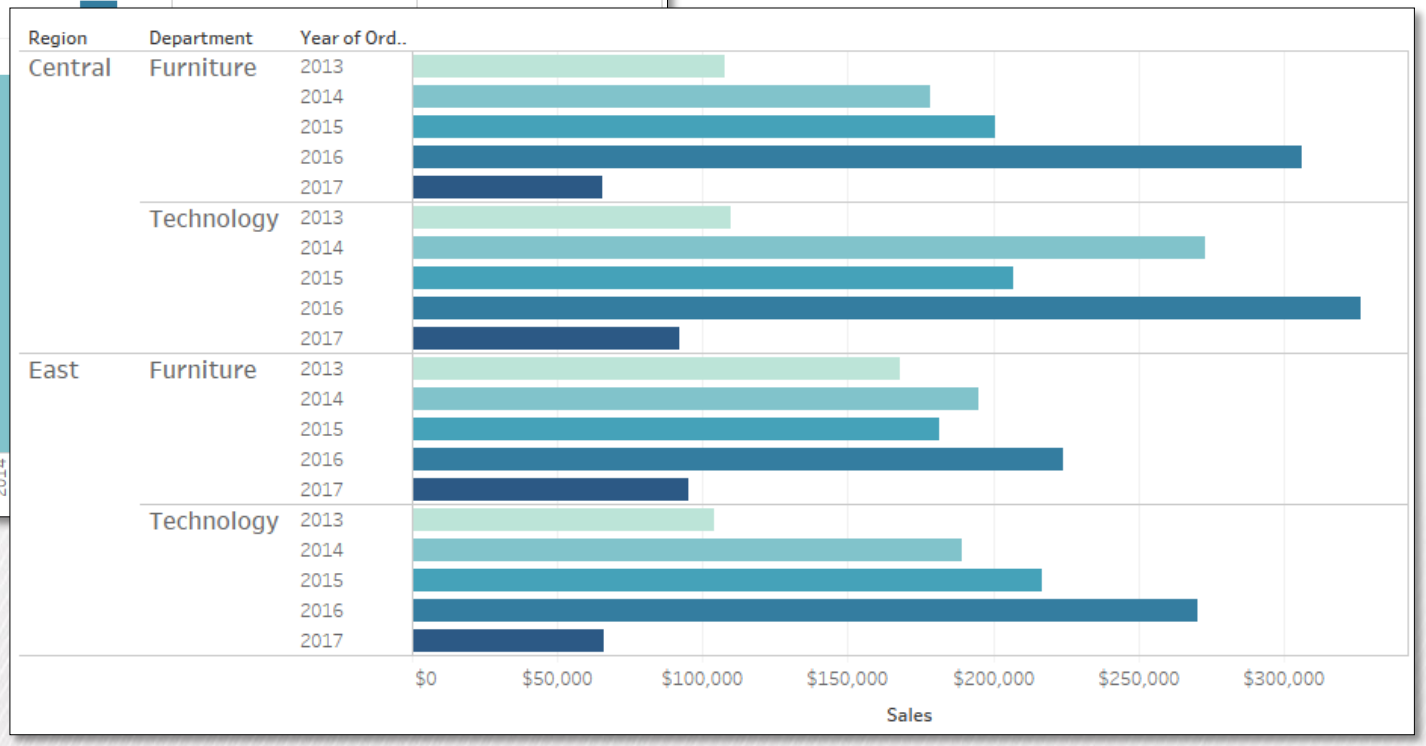
How Do Humans Like Their Data?

Orient data so people can read it easily

Good



Better

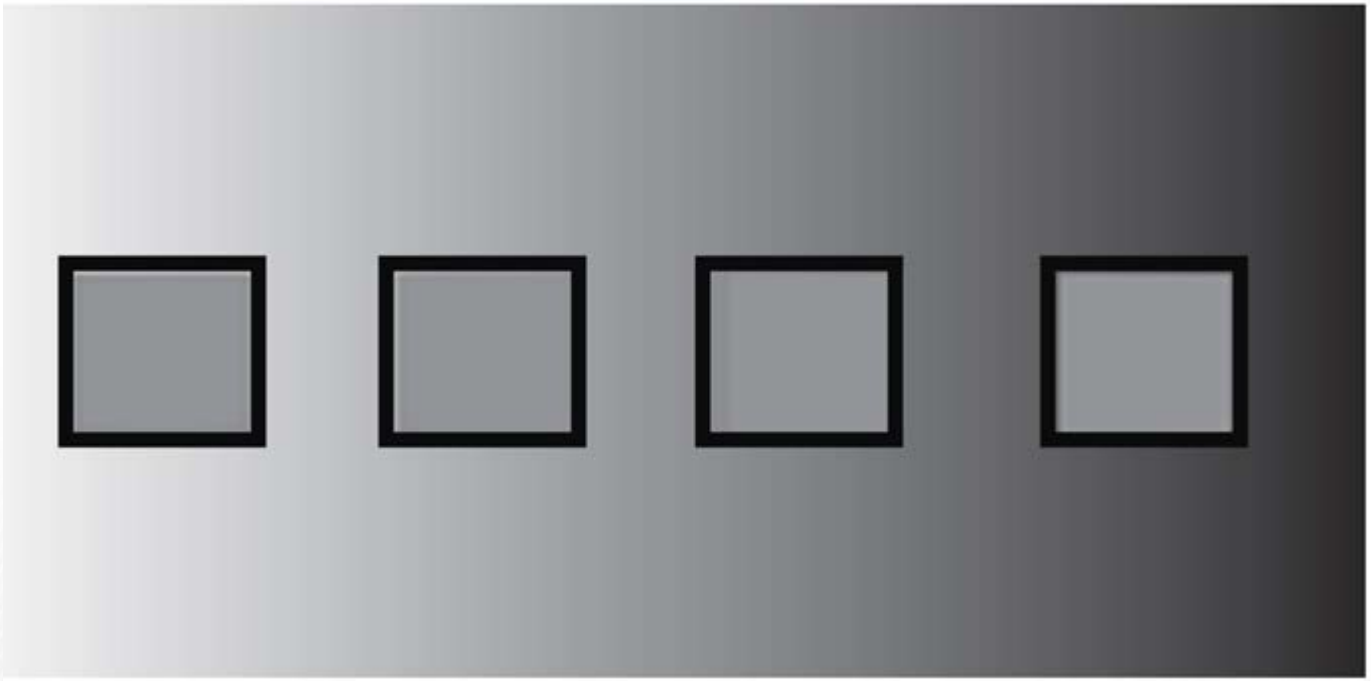


Color Me Impressed

Color perception is relative, not absolute

Color Me Impressed

Provide a consistent background



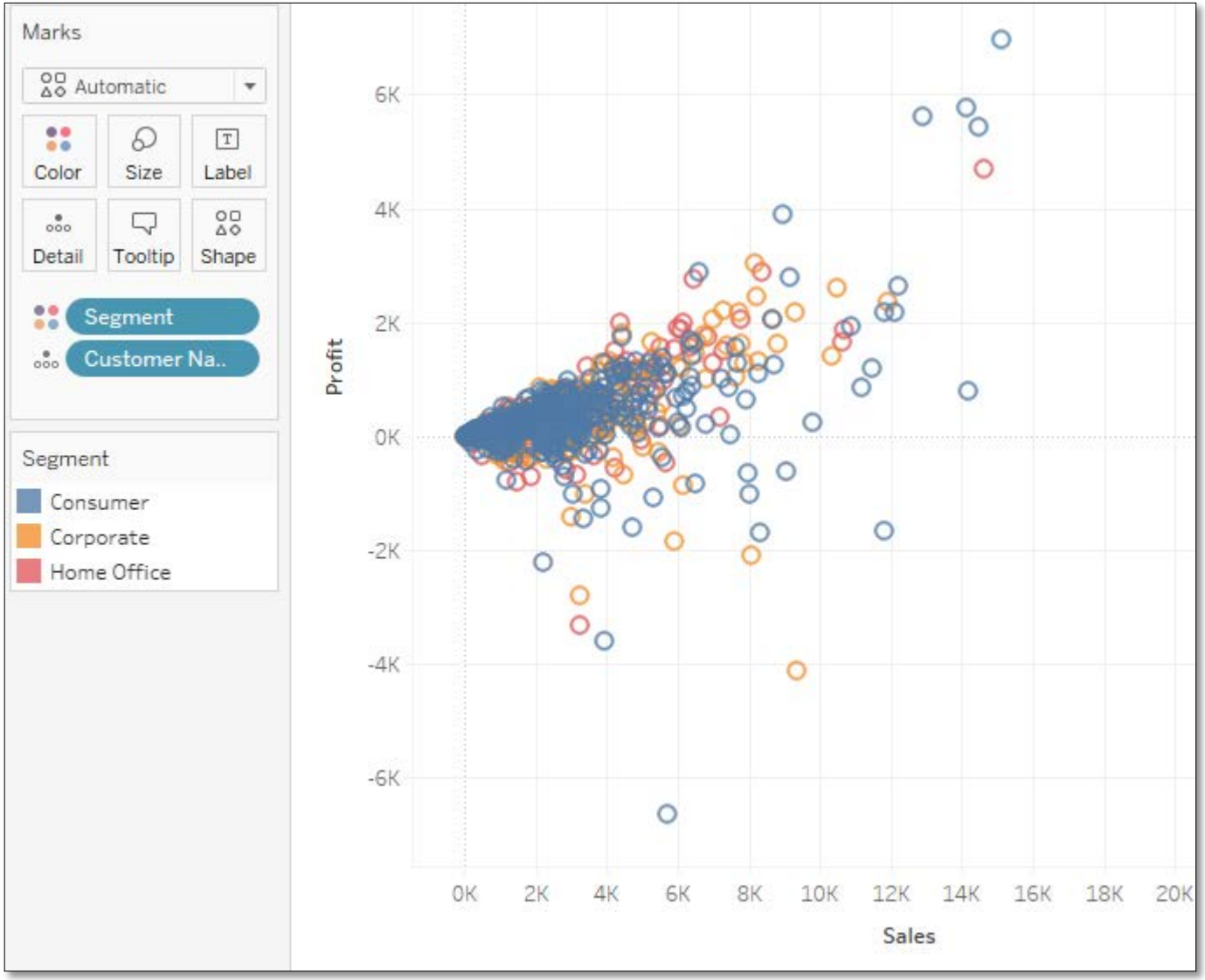
Color Me Impressed

Humans can only distinguish ~8 colors

This is not helpful.

Color Me Impressed

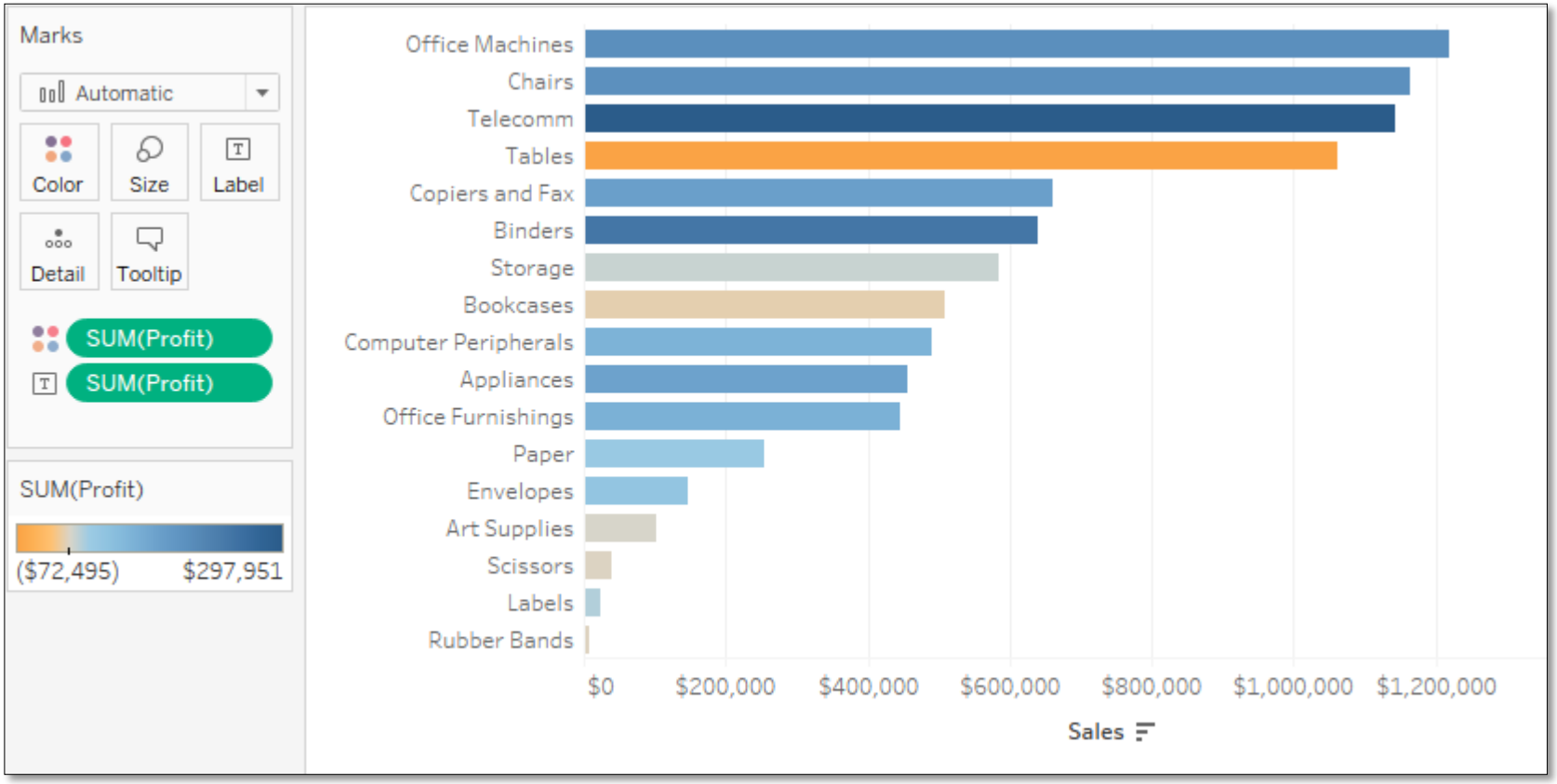
Humans can only distinguish ~8 colors



This is helpful.

Color Me Impressed

For quantitative data, color intensity and diverging color palettes work well



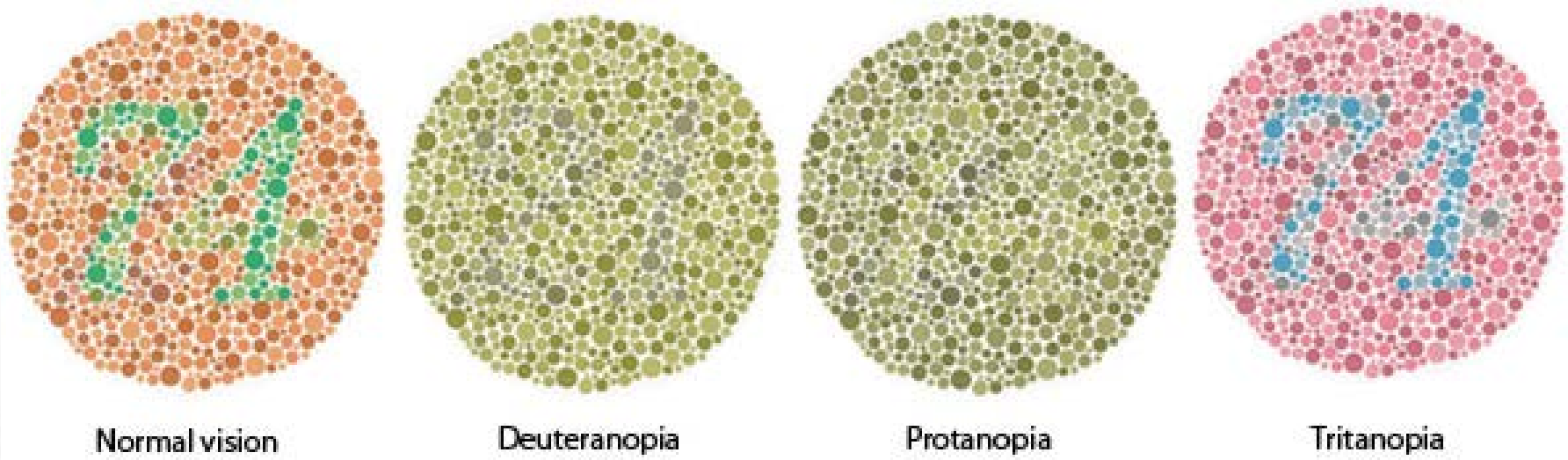
A word on Color and Color Blindness

On Tuesday, September 16th 2014, fans of Liverpool FC were baffled during the clash in the Champions League with Ludogorets.

> “So this Liverpool game isn't great for me being colour blind. Green pitch green kit red kit all look the same. It's just floating heads.”
>
> - Elliot Heard (on Twitter)

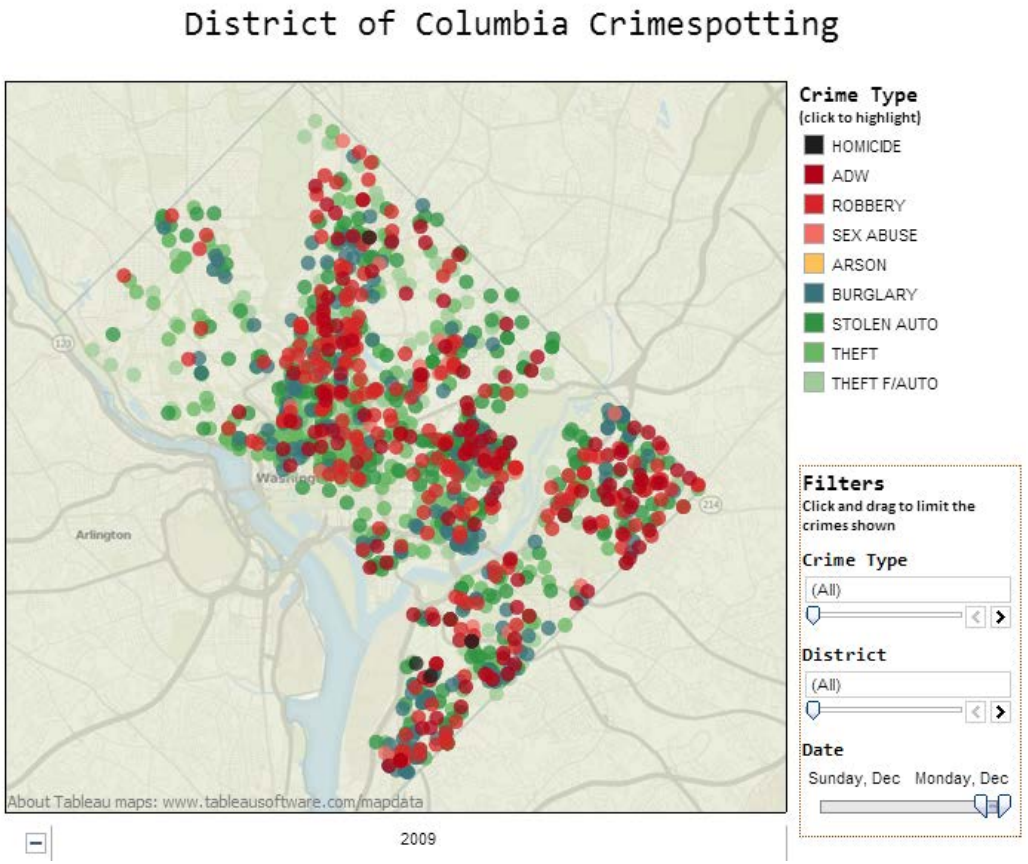
A word on Color and Color Blindness

7-10% of Males has some form of Color Blindness while 0.5% of Females have.



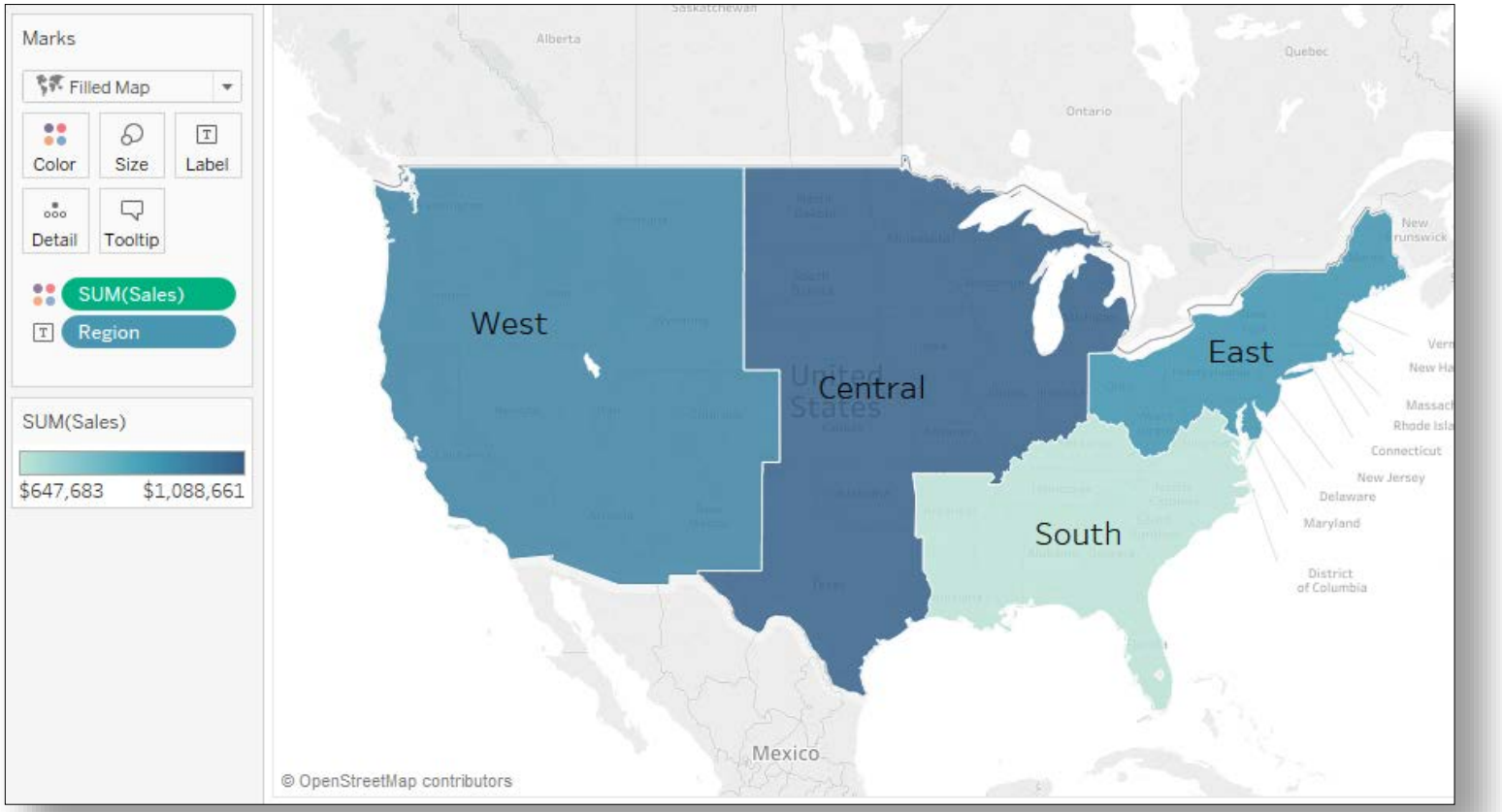
Mapping to Insight

Use maps when location is relevant



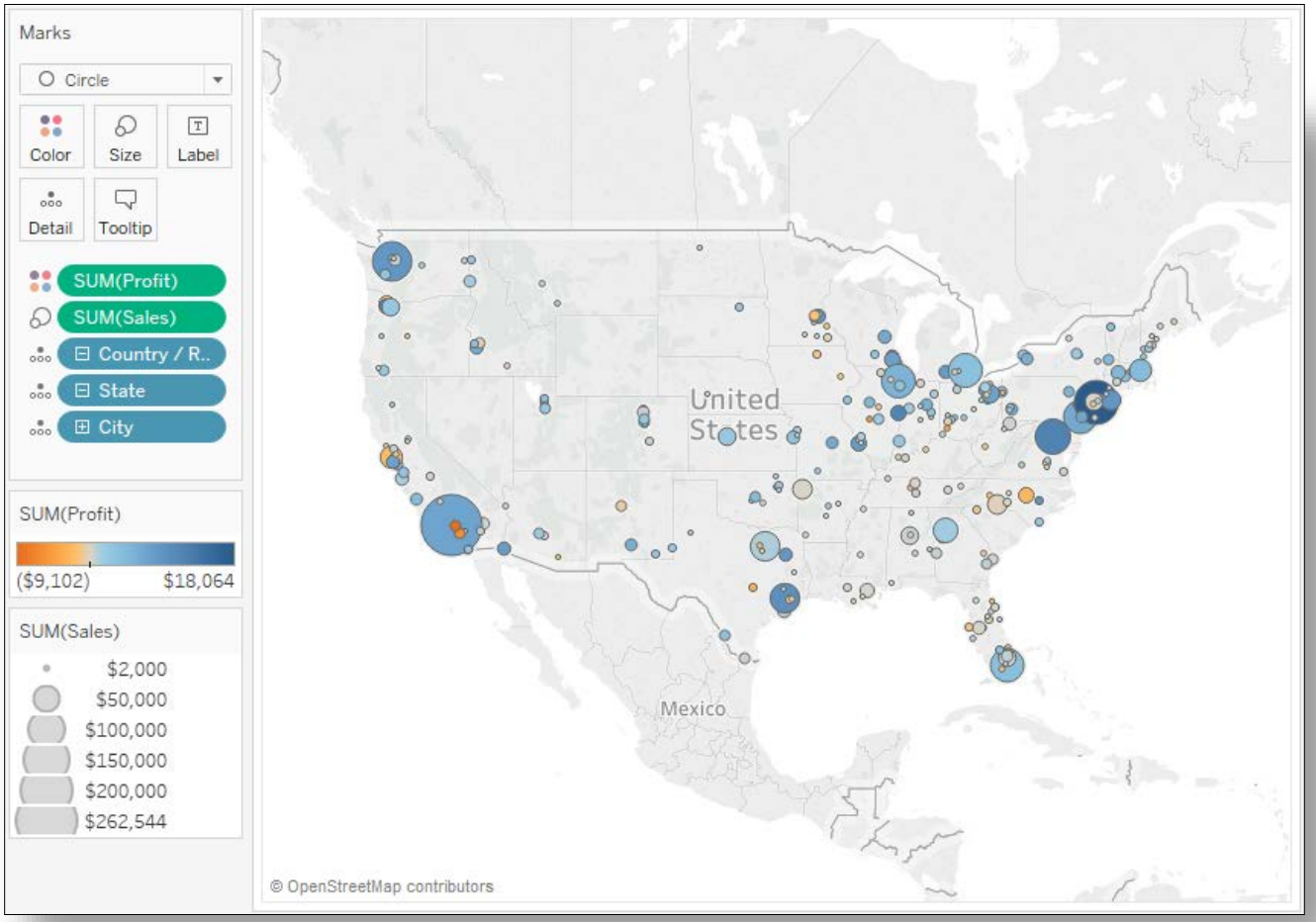
Mapping to Insight

Use filled maps (“cloropleths”) for defined areas and only ONE measure



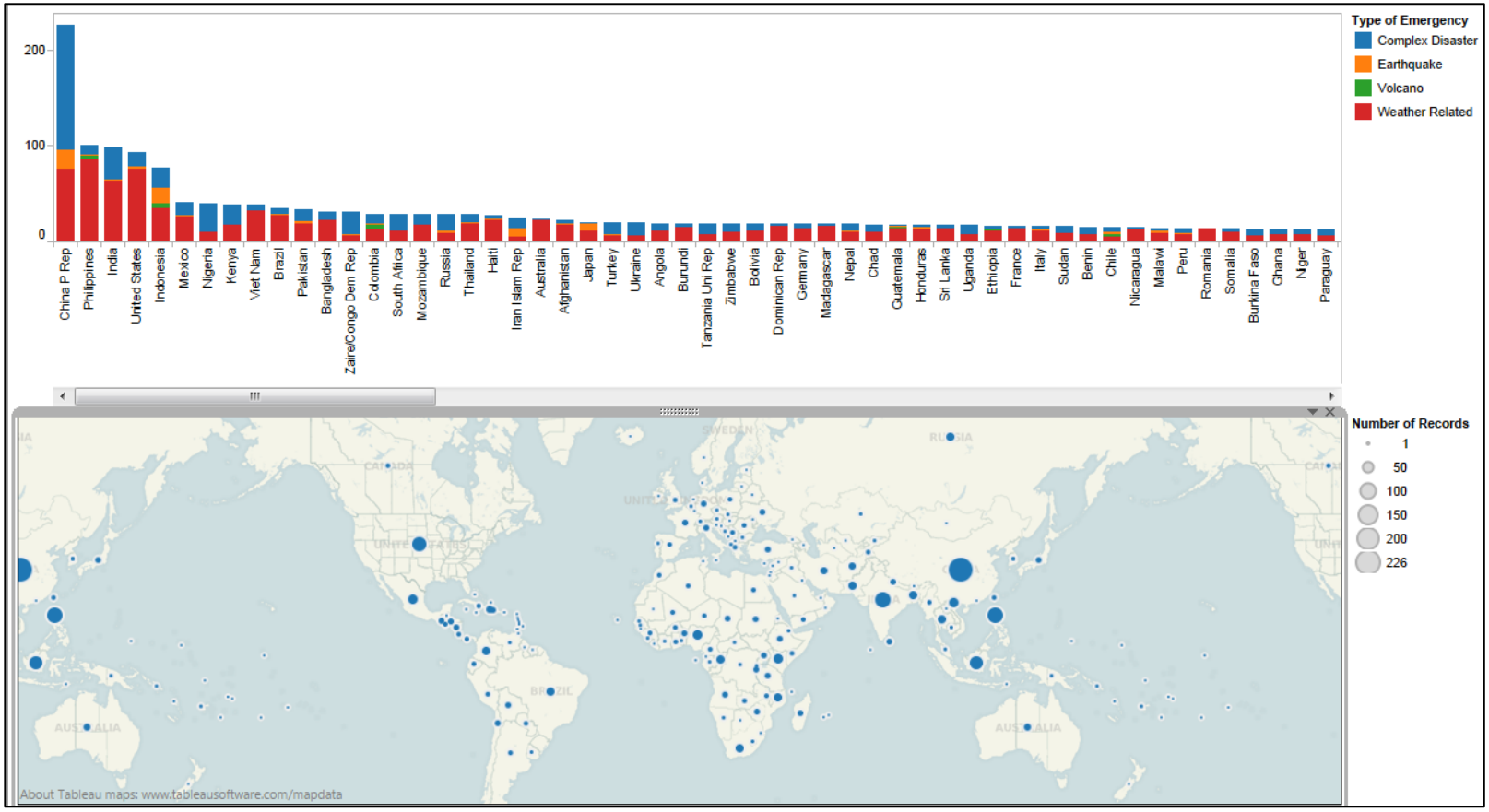
Mapping to Insight

Filled maps won’t work for multiple measures



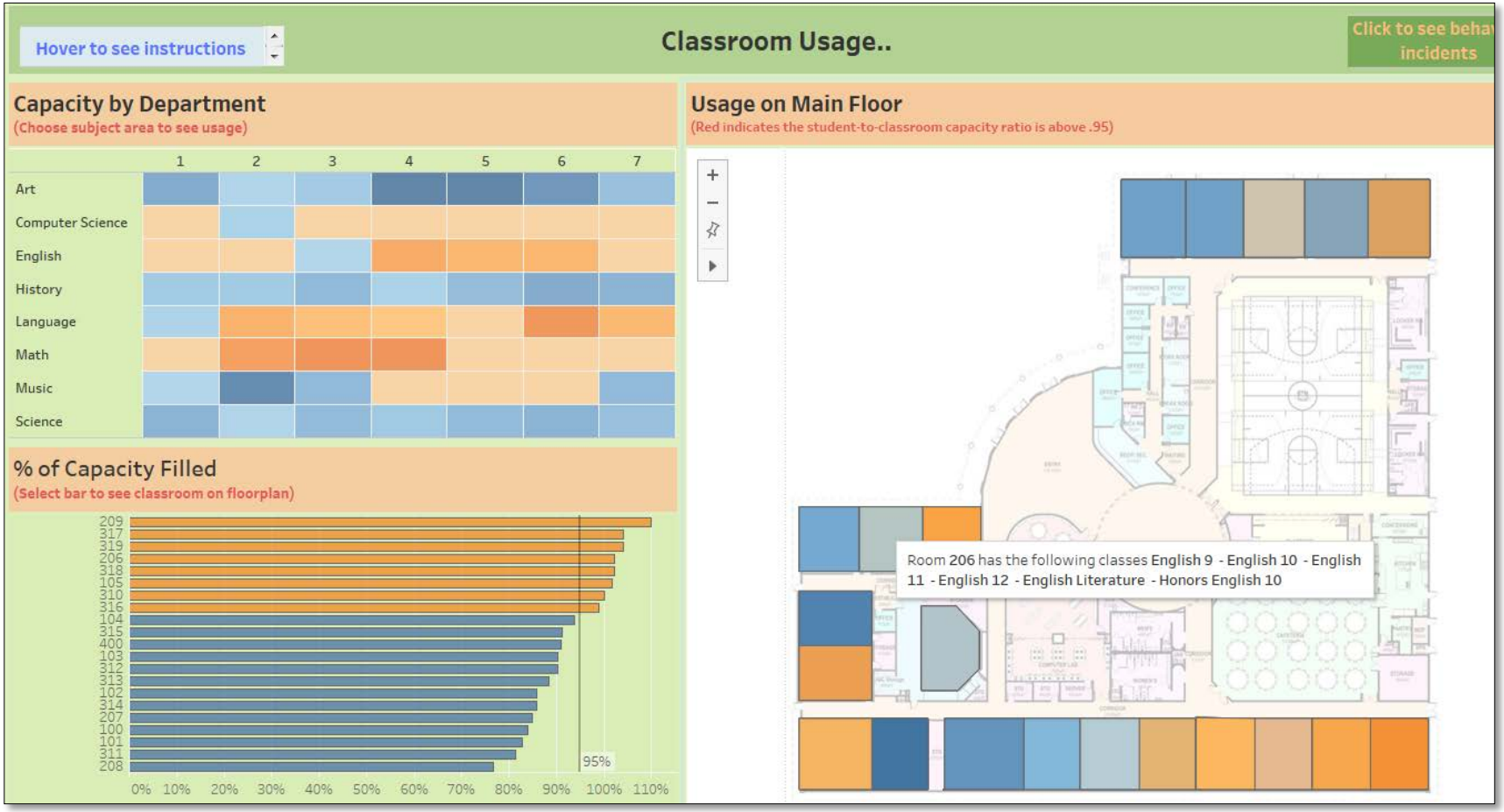
Mapping to Insight

Don’t use maps just because you can



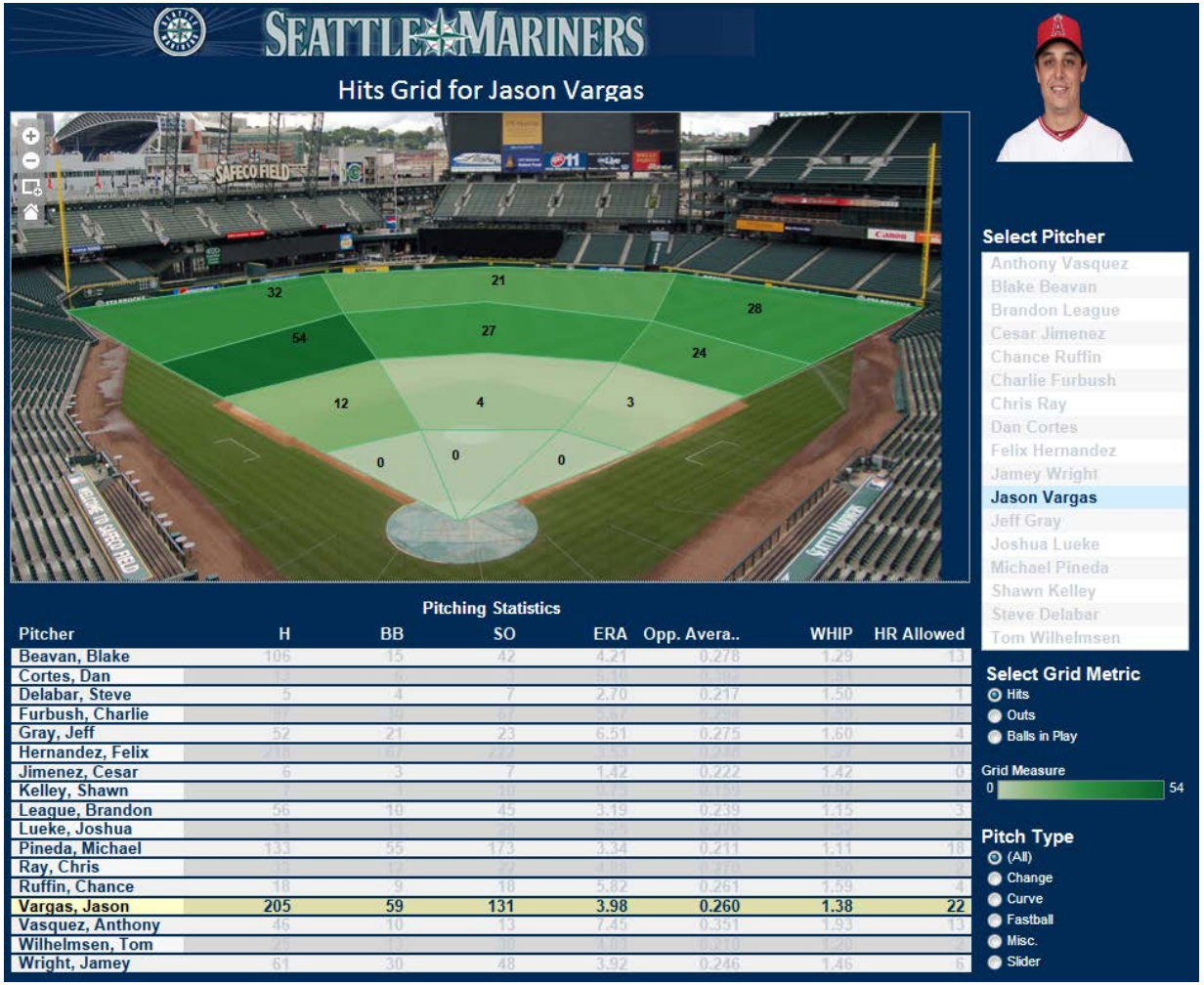
Mapping to Insight

Maps don’t have to be geographic



Mapping to Insight

Maps don’t have to be geographic

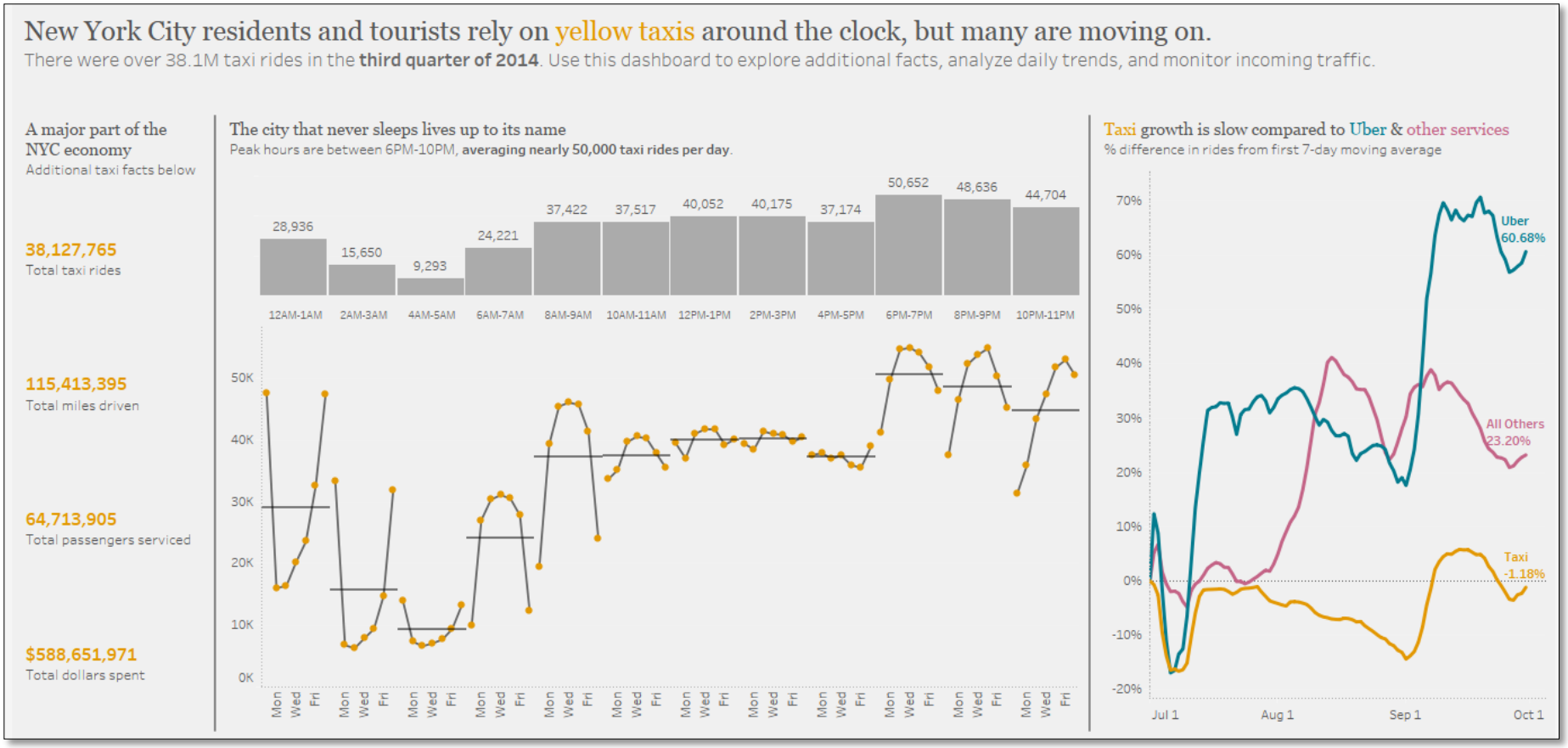


Pitching Statistics

Pitcher	H	BB	SO	ERA	Opp. Avera..	WHIP	HR Allowed
Beavan, Blake	106	15	42	4.21	0.278	1.29	13
Cortes, Dan	13	6	3	6.10	0.302	1.84	3
Delabar, Steve	5	4	7	2.70	0.217	1.50	1
Furbush, Charlie	87	38	61	5.67	0.293	1.93	15
Gray, Jeff	52	21	23	6.51	0.275	1.60	4
Hernandez, Felix	205	61	222	3.45	0.240	1.21	18
Jimenez, Cesar	6	3	7	1.42	0.222	1.42	0
Kelley, Shawn	7	1	10	9.45	0.189	0.95	1
League, Brandon	56	10	45	3.19	0.239	1.15	3
Lueke, Joshua	33	11	29	6.26	0.270	1.52	7
Pineda, Michael	133	55	173	3.34	0.211	1.11	18
Ray, Chris	33	12	22	4.86	0.210	1.50	2
Ruffin, Chance	18	9	18	5.82	0.261	1.59	4
Vargas, Jason	**205**	**59**	**131**	**3.98**	**0.260**	**1.38**	**22**
Vasquez, Anthony	46	10	13	7.45	0.351	1.93	13
Wilhelmsen, Tom	29	15	30	4.00	0.218	1.20	2
Wright, Jamey	61	30	48	3.92	0.246	1.46	6

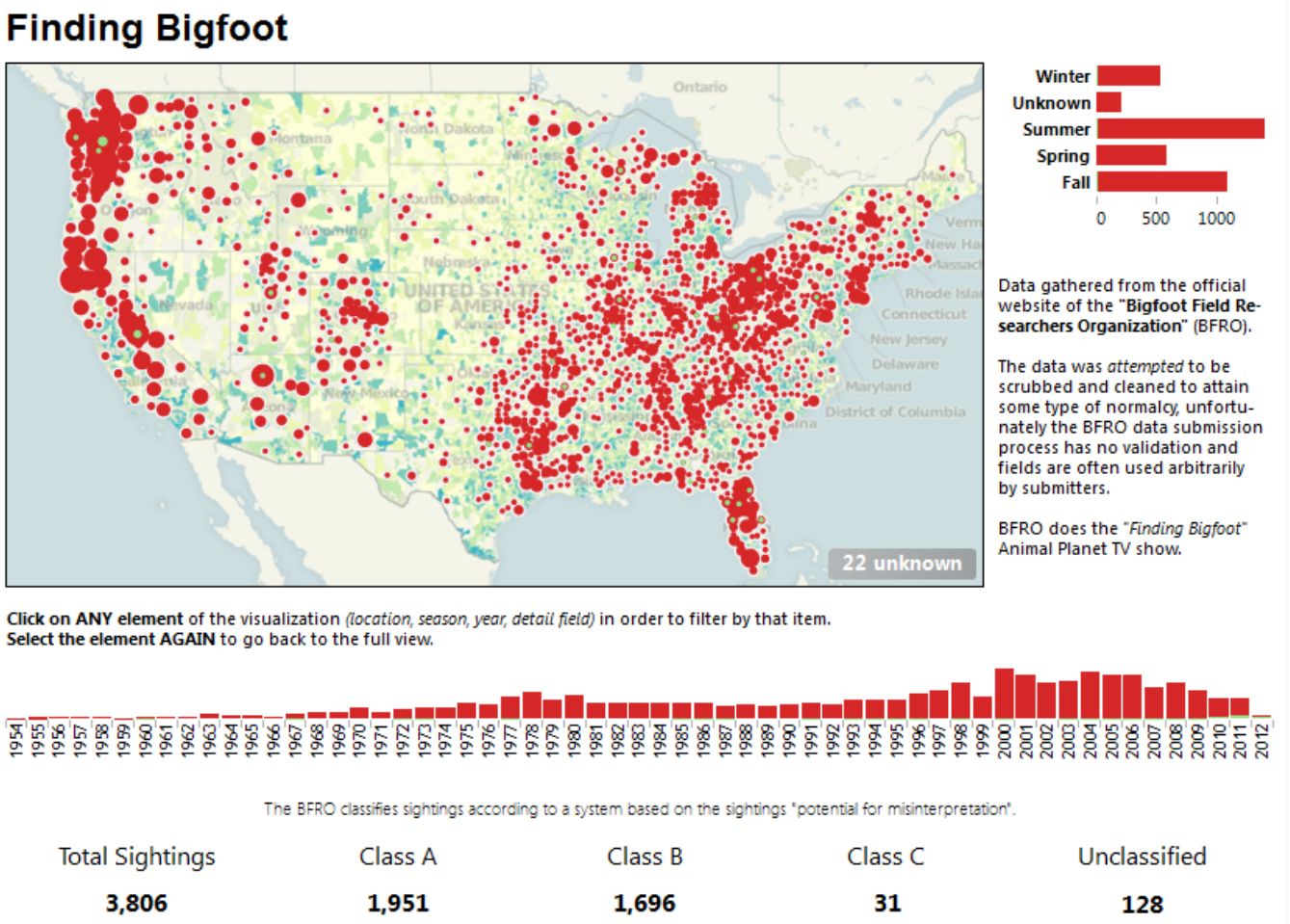
Dashboards

Dashboards bring together multiple views



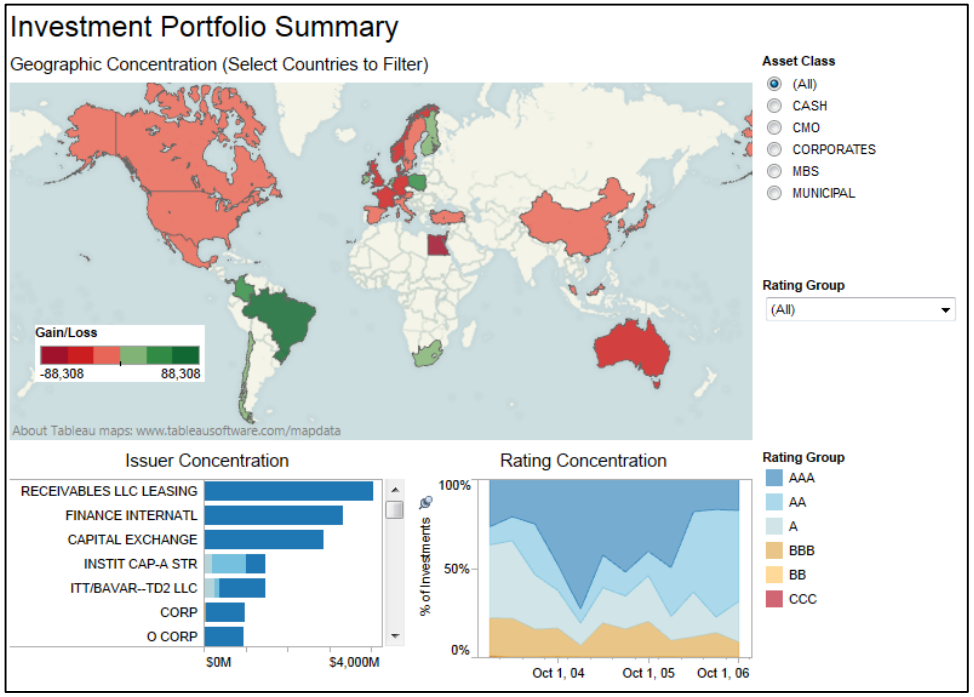
Dashboards

Dashboards should pass the 5-second test



The Five-second Dashboard Test

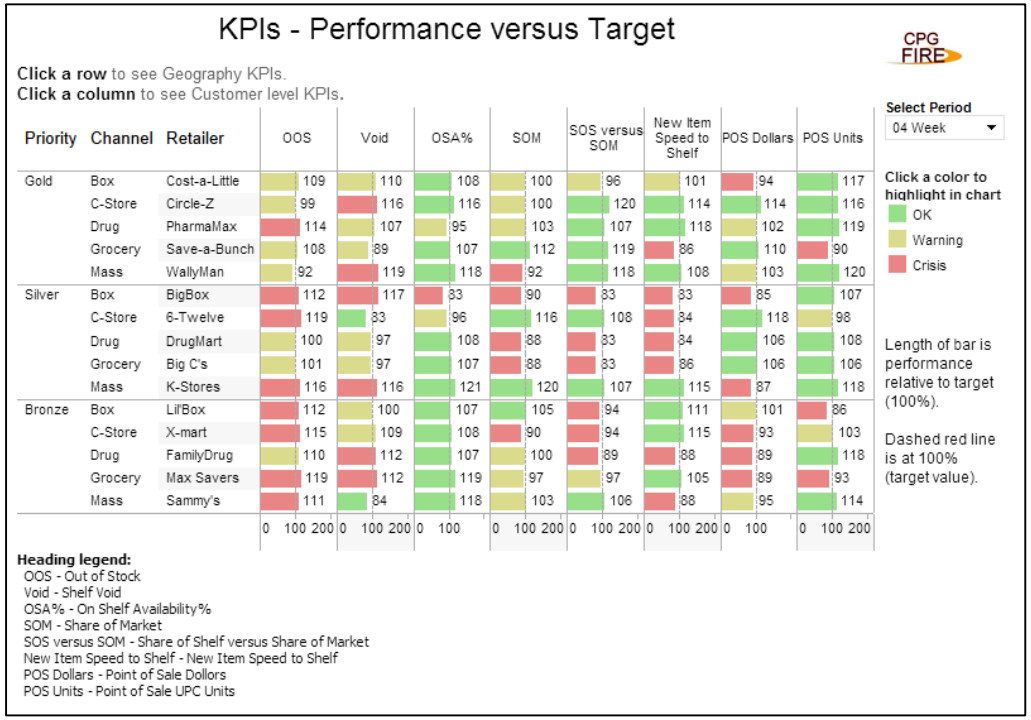
- Most important view goes on top or top-left
- Legends go near their views
- Avoid using multiple color schemes on a single dashboard
- Use 5 views or fewer in dashboards
- Provide interactivity



The Five-second Dashboard Test

Use your words!

- Titles
- Axes
- Key facts and figures
- Units
- Remove extra digits in numbers
- Great tooltips



KPIs - Performance versus Target

CPG FIRE

Click a row to see Geography KPIs.
Click a column to see Customer level KPIs.

Priority	Channel	Retailer	OOS	Void	OSA%	SOM	SOS versus SOM	New Item Speed to Shelf	POS Dollars	POS Units
Gold	Box	Cost-a-Little	109	110	108	100	96	101	94	117
	C-Store	Circle-Z	99	116	116	100	120	114	114	116
	Drug	PharmaMax	114	107	95	103	107	118	102	119
	Grocery	Save-a-Bunch	108	89	107	112	119	86	110	90
	Mass	WallyMan	92	119	118	92	118	108	103	120
Silver	Box	BigBox	112	117	83	90	83	83	85	107
	C-Store	6-Twelve	119	83	96	116	108	84	118	98
	Drug	DrugMart	100	97	108	88	83	84	106	108
	Grocery	Big C's	101	97	107	88	83	86	106	106
	Mass	K-Stores	116	116	121	120	107	115	87	118
Bronze	Box	Lil'Box	112	100	107	105	94	111	101	86
	C-Store	X-mart	115	109	108	90	94	115	93	103
	Drug	FamilyDrug	110	112	107	100	89	88	89	118
	Grocery	Max Savers	119	112	119	97	97	105	89	93
	Mass	Sammy's	111	84	118	103	106	88	95	114

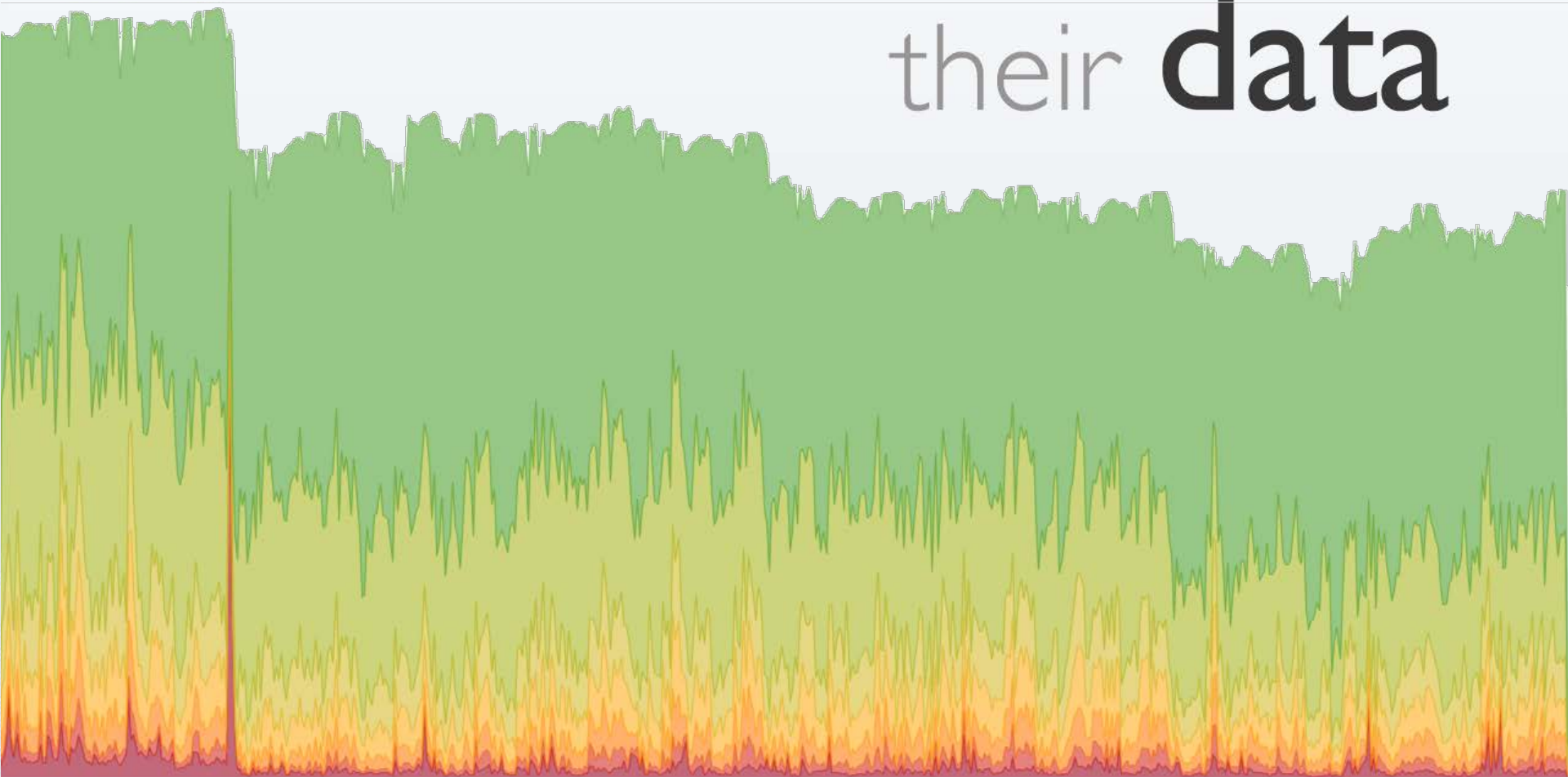
0 100 200 | 0 100 200 | 0 100 | 0 100 200 | 0 100 200 | 0 100 200 | 0 100 | 0 100 200

Select Period
04 Week

Click a color to highlight in chart
OK
Warning
Crisis

Length of bar is performance relative to target (100%).

Dashed red line is at 100% (target value).

Heading legend:
OOS - Out of Stock
Void - Shelf Void
OSA% - On Shelf Availability%
SOM - Share of Market
SOS versus SOM - Share of Shelf versus Share of Market
New Item Speed to Shelf - New Item Speed to Shelf
POS Dollars - Point of Sale Dollors
POS Units - Point of Sale UPC Units

Help people see and understand their data