PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

I. Theoretical:

Task Abstraction
Validation

II. Practical

Part-to-Whole
Rates & Ratios

Week 3
Oct 27, 2015

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

I. Theoretical:

Task Abstraction
Validation

II. Practical

Part-to-Whole
Rates & Ratios

Week 3
Oct 27, 2015

The Big Picture

Who is doing **What** with a viz, and **Why**?

Why?

Actions

Analyze

- Consume
  - *Discover*
  - *Present*
  - *Enjoy*
- Produce
  - *Annotate*
  - *Record*
  - *Derive*

Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Query

- Identify
- Compare
- Summarize

Targets

All Data

- Trends
- Outliers
- Features

Attributes

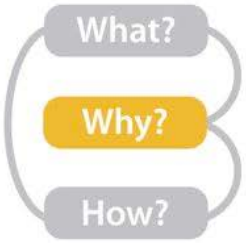
- One
  - *Distribution*
  - *Extremes*
- Many
  - *Dependency*
  - *Correlation*
  - *Similarity*

Network Data

- Topology
- *Paths*

Spatial Data

- Shape



Why?

Actions

Analyze

- Consume
  - *Discover*
  - *Present*
  - *Enjoy*
- Produce
  - *Annotate*
  - *Record*
  - *Derive*

Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Query

- Identify
- Compare
- Summarize

VERBS

Targets

All Data

- Trends
- Outliers
- Features

Attributes

- One
  - *Distribution*
  - *Extremes*
- Many
  - *Dependency*
  - *Correlation*
  - *Similarity*

Network Data

- Topology
- *Paths*

Spatial Data

- Shape

What?

Why?

How?

Why?

Actions

Analyze

- Consume
  - Discover
  - Present
  - Enjoy
- Produce
  - Annotate
  - Record
  - Derive

Search

	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Query

- Identify
- Compare
- Summarize

Targets

All Data

- Trends
- Outliers
- Features

Attributes

- One
  - Distribution
  - Extremes
- Many
  - Dependency
  - Correlation
  - Similarity

Network Data

- Topology
- Paths

Spatial Data

- Shape

What?
Why?
How?

NOUNS

Why?

Actions

Analyze

- Consume
  - *Discover*
  - *Present*
  - *Enjoy*

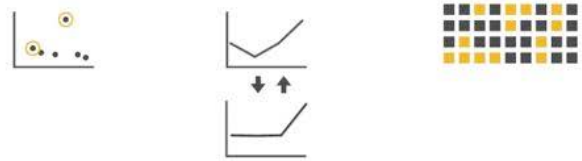
- Produce
  - *Annotate*
  - *Record*
  - *Derive*

Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Query

- Identify
- Compare
- Summarize



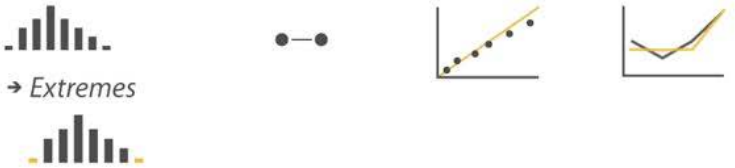
Targets

All Data

- Trends
- Outliers
- Features

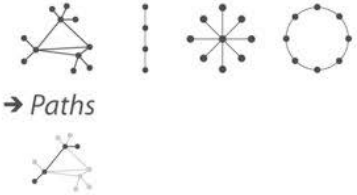
Attributes

- One
  - *Distribution*
  - *Extremes*
- Many
  - *Dependency*
  - *Correlation*
  - *Similarity*



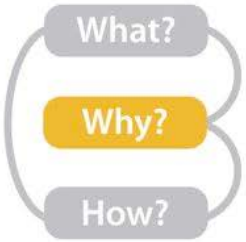
Network Data

- Topology
- *Paths*

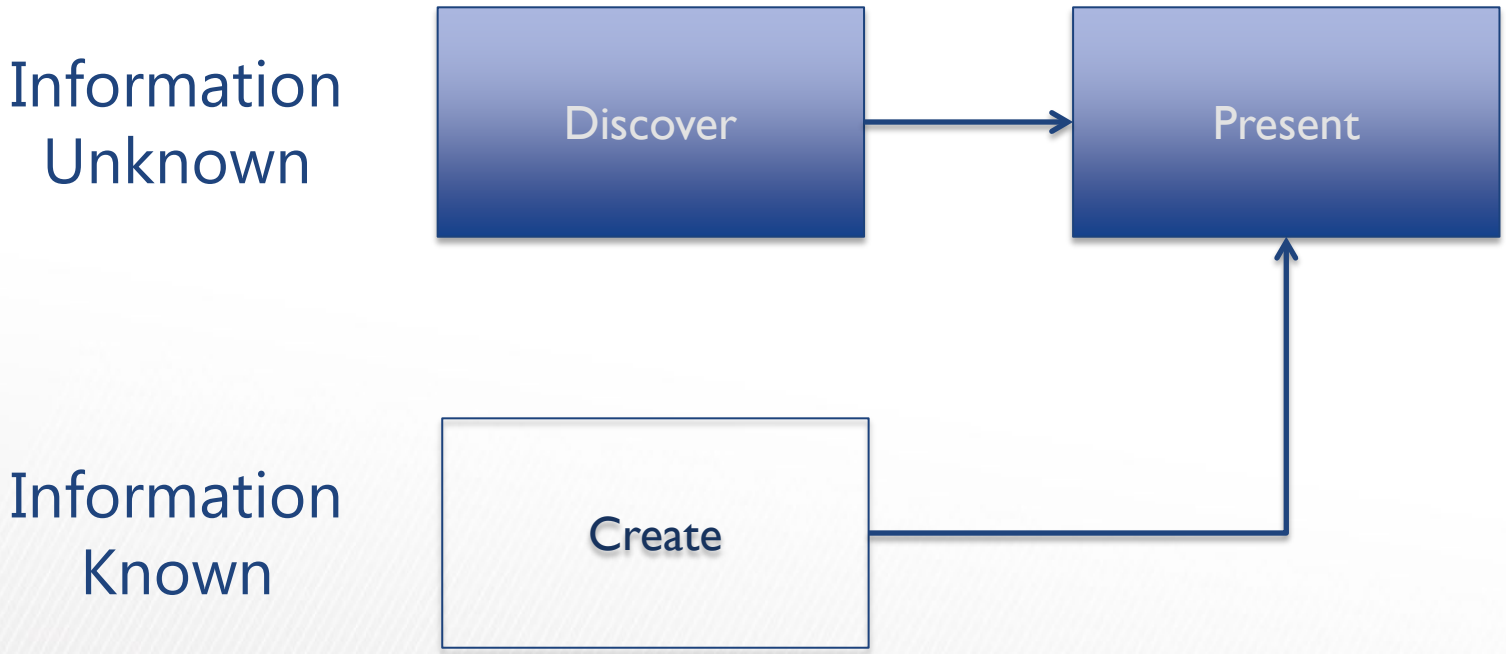


Spatial Data

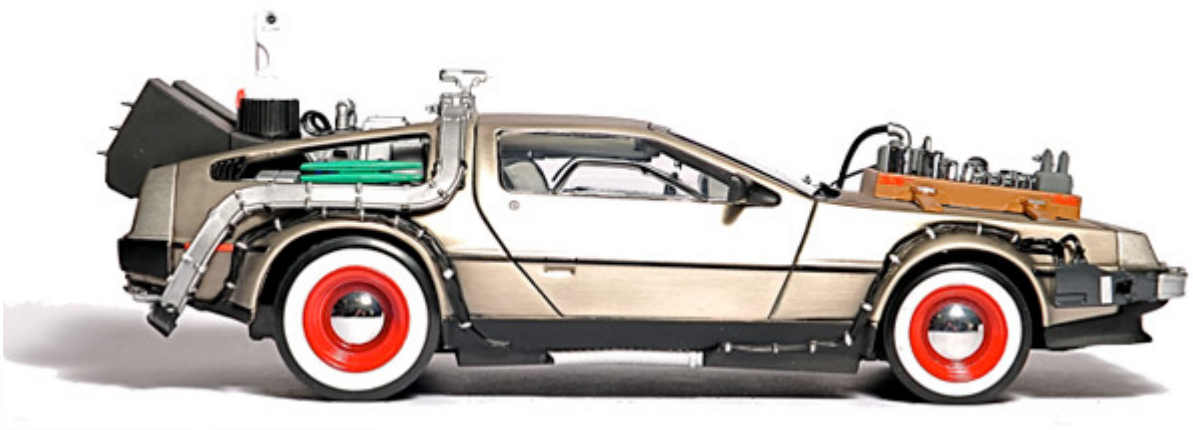
- Shape



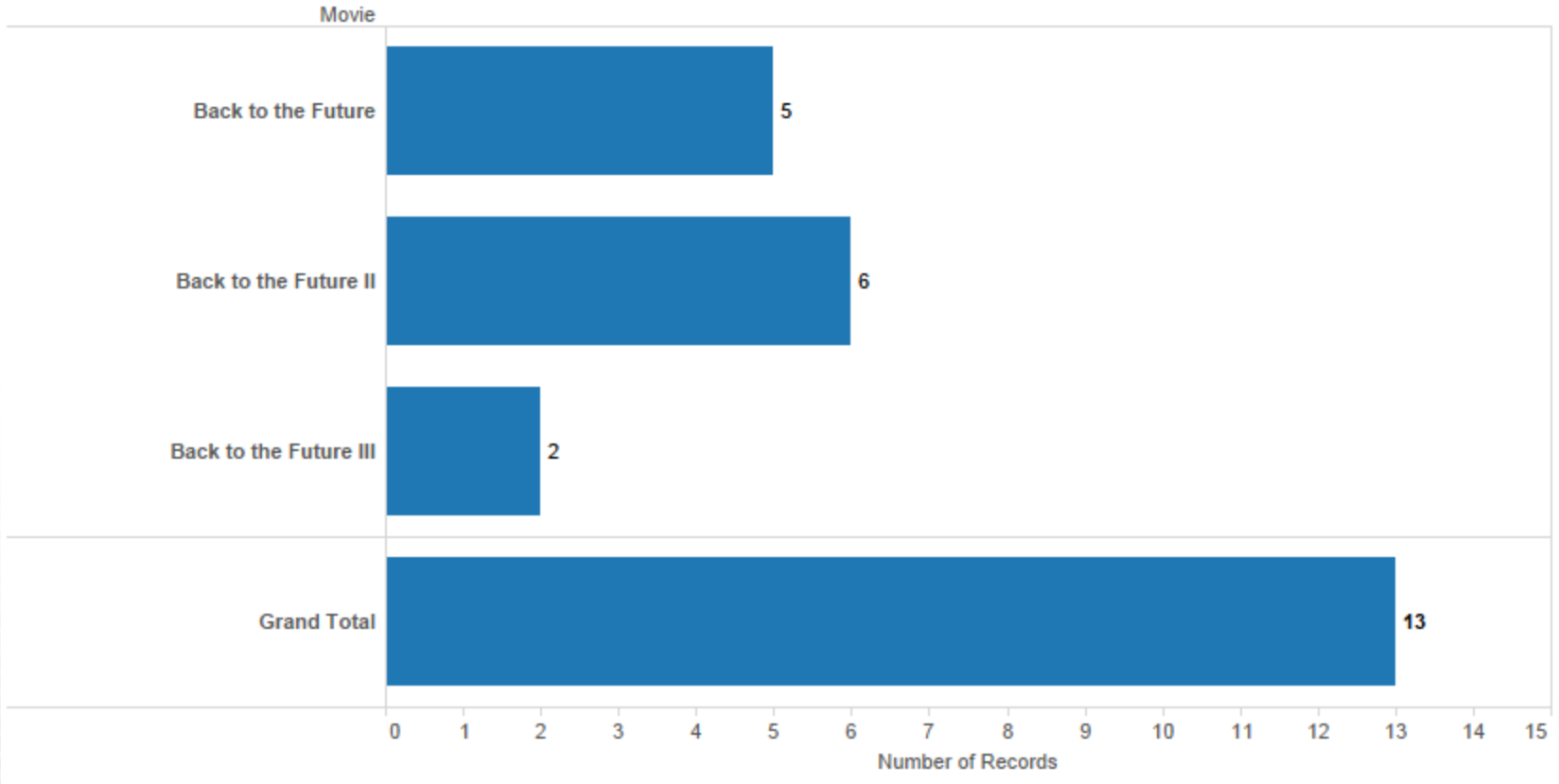
Analyze: Consume

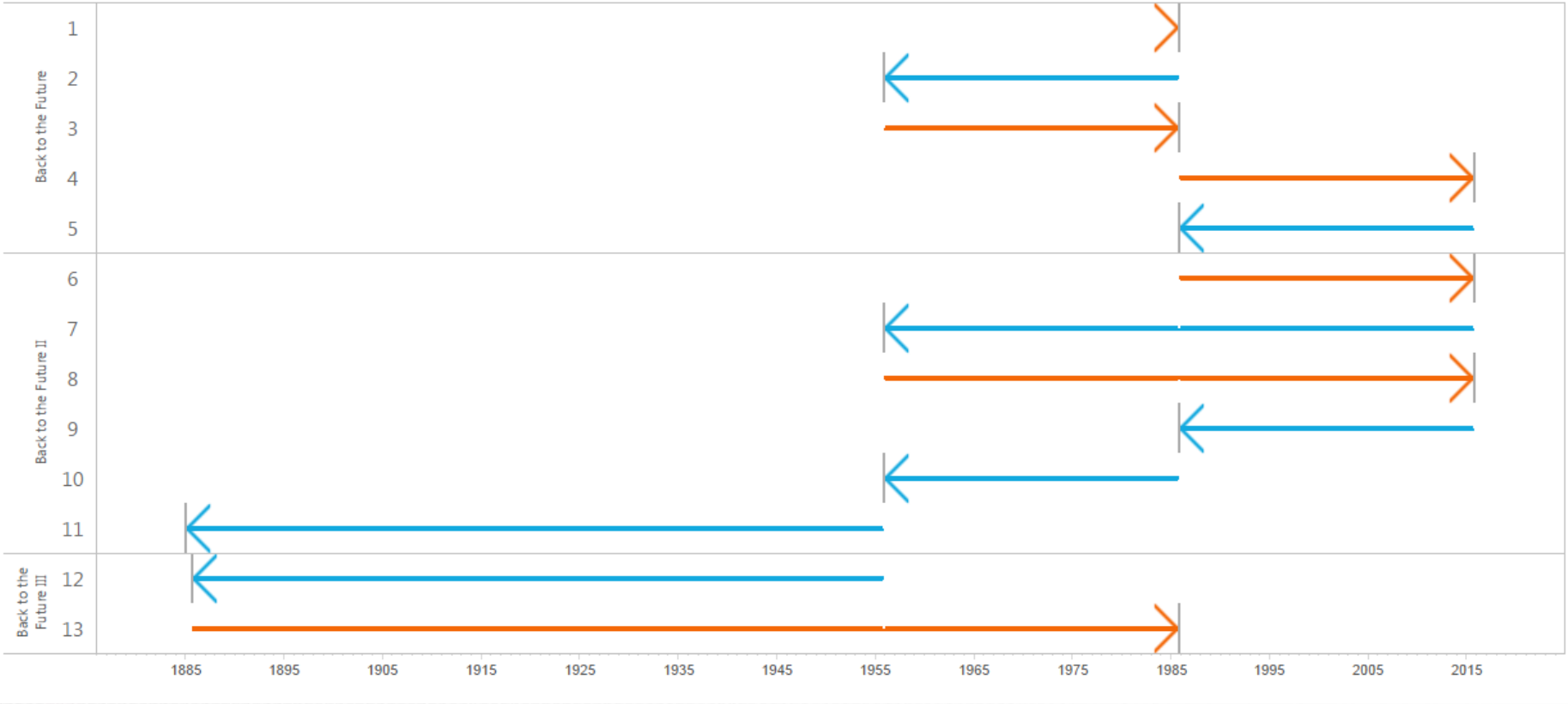


Q: How many trips did the DeLorean time machine take in the Back To The Future movies?



	A	B	C	D	E	F	G
1	**Trip**	**Movie**	**Departure Day**	**Arrival Day**	**Time Jump (days)**	**Direction**	**Passengers**
2	1	Back to the Future	10/26/1985	10/26/1985	0.000694444	Forward	Einstein
3	2	Back to the Future	10/26/1985	11/5/1955	-10948	Backward	Marty
4	3	Back to the Future	11/12/1955	10/26/1985	10941	Forward	Marty
5	4	Back to the Future	10/26/1985	10/26/2015	10957	Forward	Doc
6	5	Back to the Future	10/21/2015	10/26/1985	-10952	Backward	Doc
7	6	Back to the Future II	10/26/1985	10/21/2015	10952	Forward	Doc, Marty and Jennifer
8	7	Back to the Future II	10/21/2015	11/12/1955	-21893	Backward	Biff
9	8	Back to the Future II	11/12/1955	10/21/2015	21893	Forward	Biff
10	9	Back to the Future II	10/21/2015	10/26/1985	-10952	Backward	Doc, Marty and Jennifer
11	10	Back to the Future II	10/26/1985	11/12/1955	-10941	Backward	Doc and Marty
12	11	Back to the Future II	11/12/1955	1/1/1885	-25881	Backward	Doc
13	12	Back to the Future III	11/16/1955	9/2/1885	-25641	Backward	Marty
14	13	Back to the Future III	9/7/1885	10/27/1985	36574	Forward	Marty
15							


Movie
Back to the Future
5
Back to the Future II
6
Back to the Future III
2
Grand Total
13
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Number of Records


Back to the Future
Back to the Future II
Back to the Future III
1
2
3
4
5
6
7
8
9
10
11
12
13
1885
1895
1905
1915
1925
1935
1945
1955
1965
1975
1985
1995
2005
2015

Every DeLorean Trip
in the Back To The Future Trilogy

Click an arrow to view each movie clip:

Back to the Future
1
2
3
4
5

Back to the Future II
6
7
8
9
10
11

Back to the Future III
12
13

1885 1955 1985 2015

Credit: Vincent Ho http://i.stack.imgur.com/tVqJk.jpg

Chart type: http://dataremixed.com/2014/04/dimension-line-charts/

Why?

Actions

Analyze

- Consume
  - Discover
  - Present
  - Enjoy
- Produce
  - Annotate
  - Record
  - Derive

Search

	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Query

- Identify
- Compare
- Summarize

Targets

All Data

- Trends
- Outliers
- Features

Attributes

- One
  - Distribution
  - Extremes
- Many
  - Dependency
  - Correlation
  - Similarity

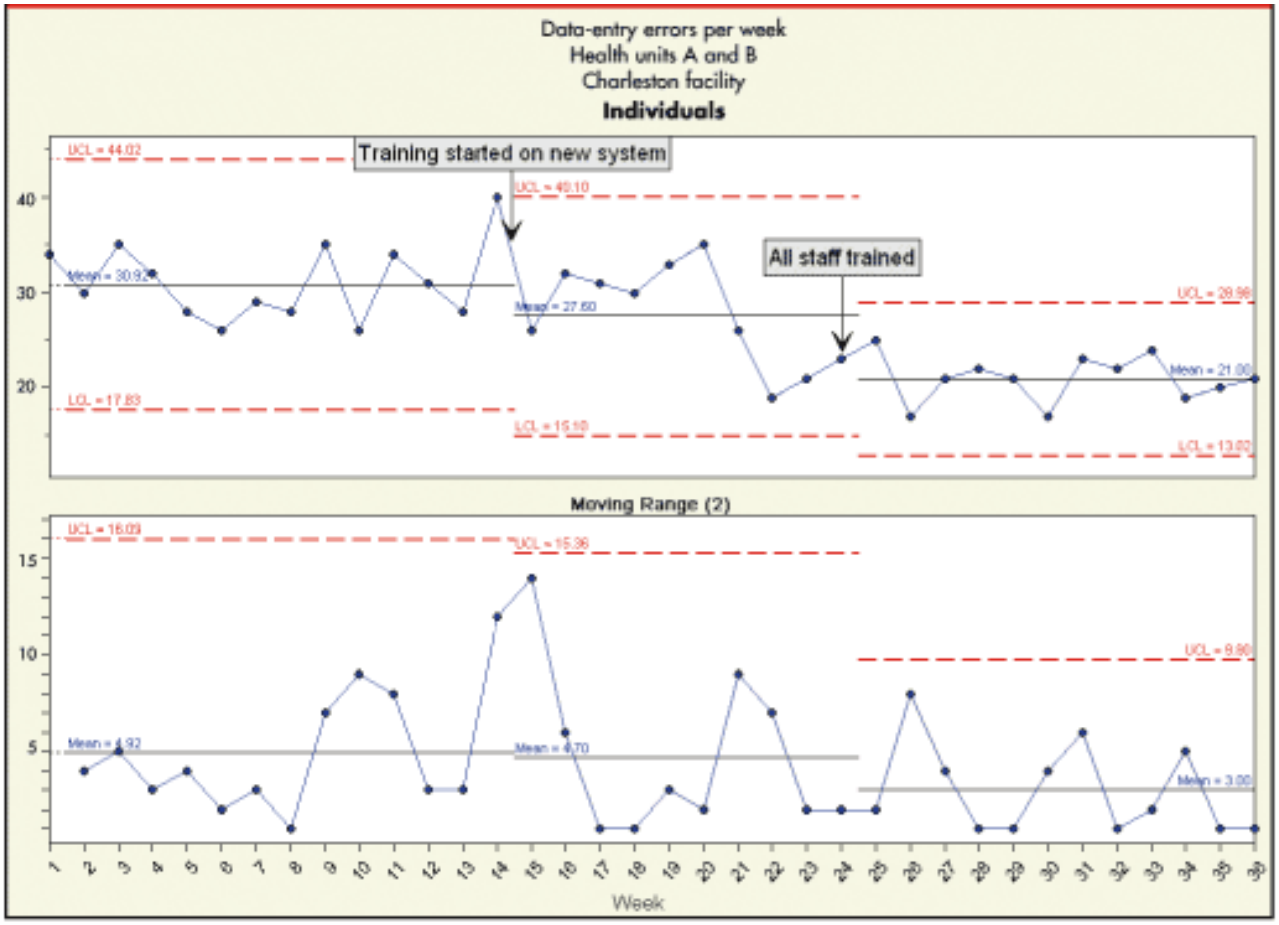
Network Data

- Topology
- Paths

Spatial Data

- Shape

What?
Why?
How?


Data-entry errors per week
Health units A and B
Charleston facility
Individuals
Training started on new system
All staff trained
UCL = 44.02
UCL = 40.10
Mean = 30.92
Mean = 27.60
UCL = 28.96
Mean = 21.00
LCL = 17.83
LCL = 15.10
LCL = 13.07
Moving Range (2)
UCL = 16.09
UCL = 15.36
UCL = 9.90
Mean = 4.92
Mean = 4.70
Mean = 3.00
Week

Why?

Actions

Analyze

- Consume
  - *Discover*
  - *Present*
  - *Enjoy*
- Produce
  - *Annotate*
  - *Record*
  - *Derive*

Search

	Target known	Target unknown
Location known	*Lookup*	*Browse*
Location unknown	*Locate*	*Explore*

Query

- Identify
- Compare
- Summarize

Targets

All Data

- Trends
- Outliers
- Features

Attributes

- One
  - *Distribution*
  - *Extremes*
- Many
  - *Dependency*
  - *Correlation*
  - *Similarity*

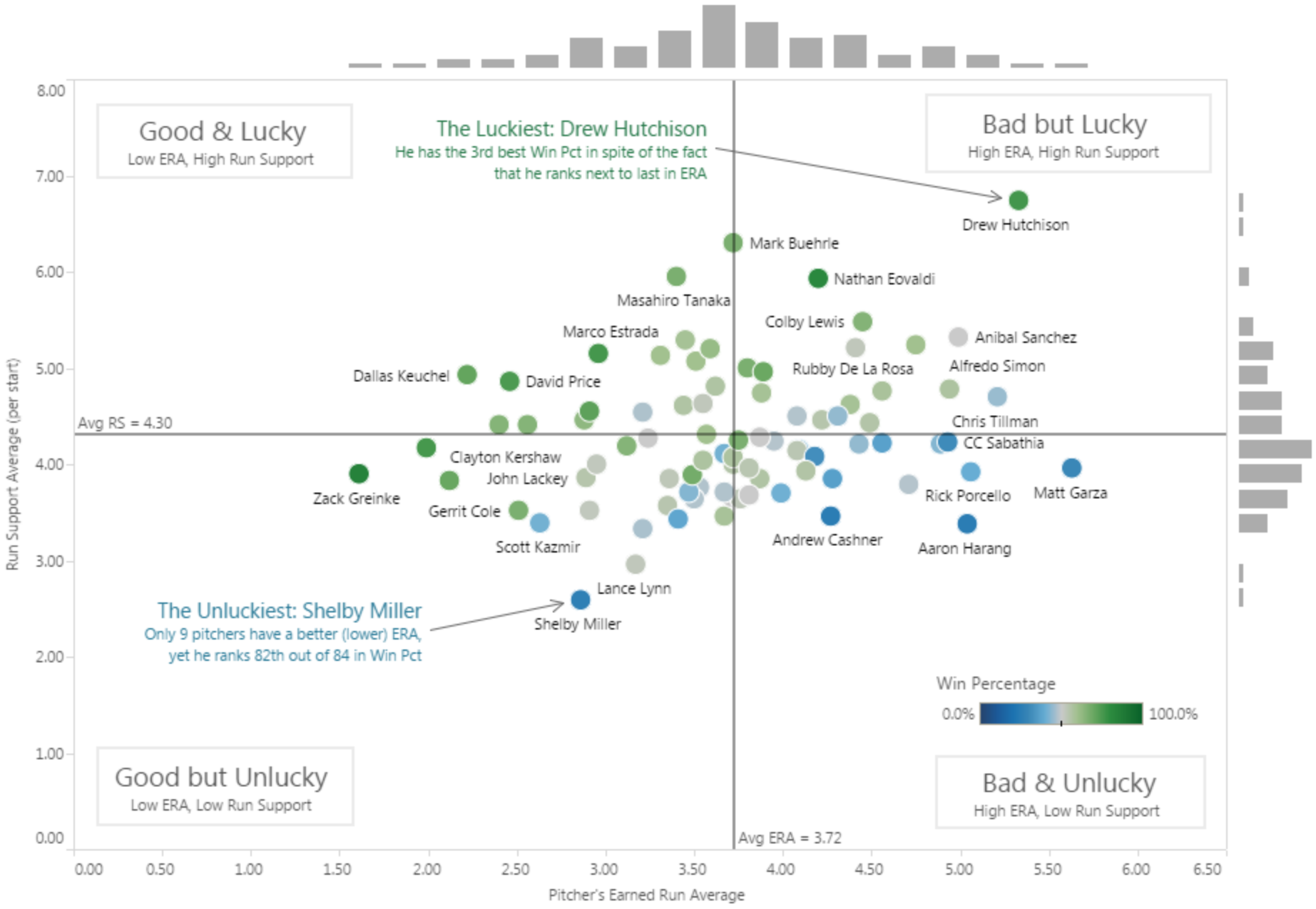
Network Data

- Topology
- *Paths*

Spatial Data

- Shape

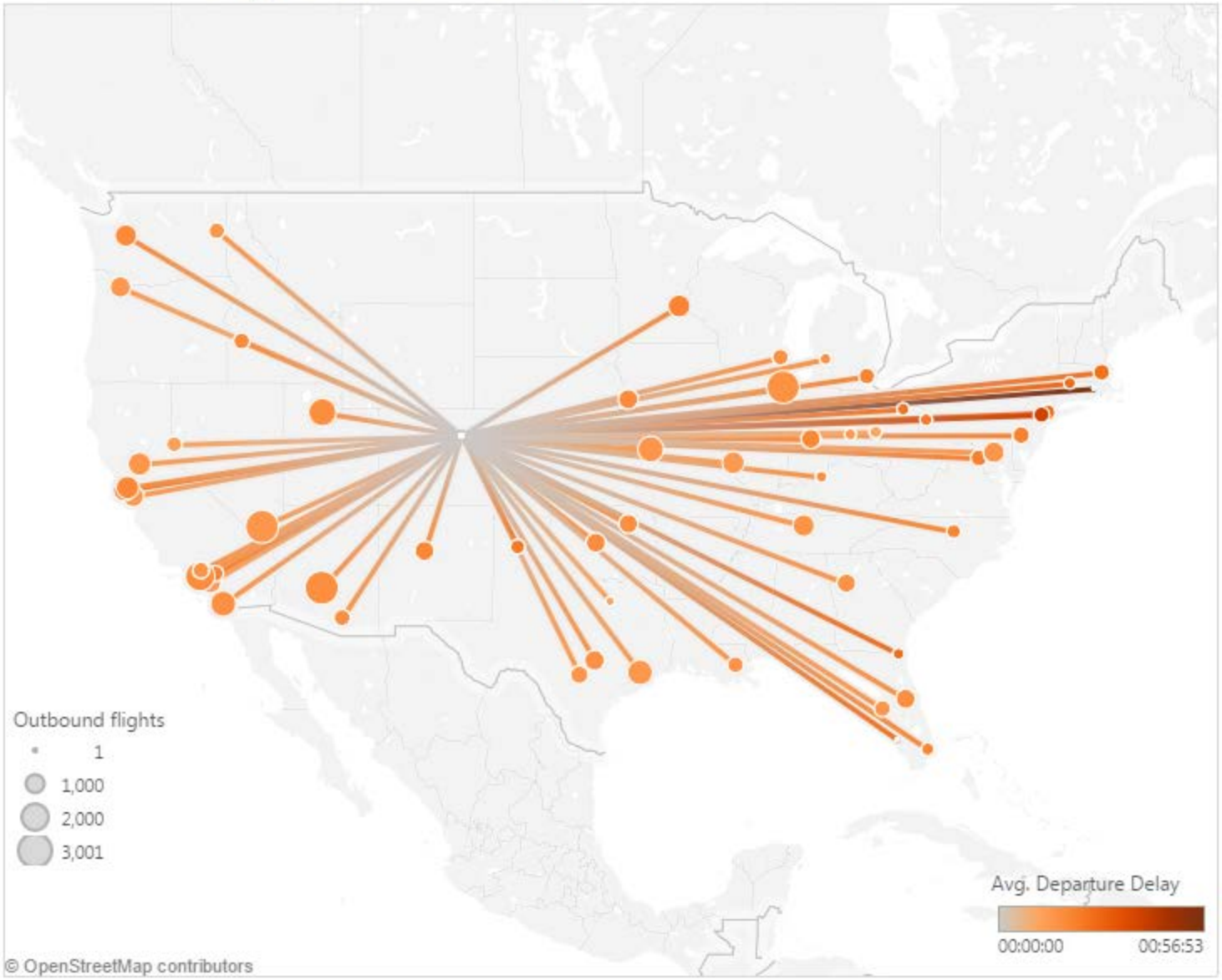
What?
Why?
How?


Good & Lucky
Low ERA, High Run Support
The Luckiest: Drew Hutchison
He has the 3rd best Win Pct in spite of the fact that he ranks next to last in ERA
Bad but Lucky
High ERA, High Run Support
Drew Hutchison
Mark Buehrle
Nathan Eovaldi
Masahiro Tanaka
Colby Lewis
Marco Estrada
Anibal Sanchez
Rubby De La Rosa
Alfredo Simon
Dallas Keuchel
David Price
Avg RS = 4.30
Chris Tillman
CC Sabathia
Clayton Kershaw
John Lackey
Zack Greinke
Gerrit Cole
Rick Porcello
Matt Garza
Scott Kazmir
Andrew Cashner
Aaron Harang
Lance Lynn
The Unluckiest: Shelby Miller
Only 9 pitchers have a better (lower) ERA, yet he ranks 82th out of 84 in Win Pct
Shelby Miller
Win Percentage
0.0%
100.0%
Good but Unlucky
Low ERA, Low Run Support
Bad & Unlucky
High ERA, Low Run Support
Avg ERA = 3.72
Run Support Average (per start)
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
0.00
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00
4.50
5.00
5.50
6.00
6.50
Pitcher's Earned Run Average

2014 U.S. Flight Departure Delays

Flights leaving from: DEN on airline: Southwest

Denver, CO



SOUTHWEST

Flights: 57,145

1	DEN→PHX	3,002
2	DEN→LAS	2,915
3	DEN→MDW	2,881
4	DEN→LAX	2,323
5	DEN→SLC	2,017
6	DEN→MCI	1,767
7	DEN→SAN	1,713
8	DEN→HOU	1,677
9	DEN→SMF	1,401
10	DEN→OAK	1,377
11	DEN→MSP	1,316

Avg delay: 00:17:43

1	DEN→PVD	00:56:53
2	DEN→EWR	00:38:43
3	DEN→LGA	00:28:23
4	DEN→BDL	00:27:25
5	DEN→JAX	00:26:23
6	DEN→SFO	00:25:42
7	DEN→BOS	00:25:29
8	DEN→RSW	00:24:48
9	DEN→CAK	00:24:35
10	DEN→AMA	00:23:06
11	DEN→PIT	00:22:29

Why?

Actions

Analyze

- Consume
  - Discover
  - Present
  - Enjoy
- Produce
  - Annotate
  - Record
  - Derive

Search

	Target known	Target unknown
Location known	Lookup	Browse
Location unknown	Locate	Explore

Query

- Identify
- Compare
- Summarize

Targets

All Data

- Trends
- Outliers
- Features

Attributes

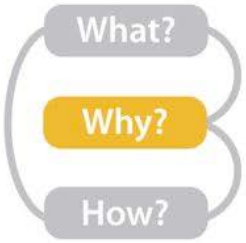
- One
  - Distribution
  - Extremes
- Many
  - Dependency
  - Correlation
  - Similarity

Network Data

- Topology
- Paths

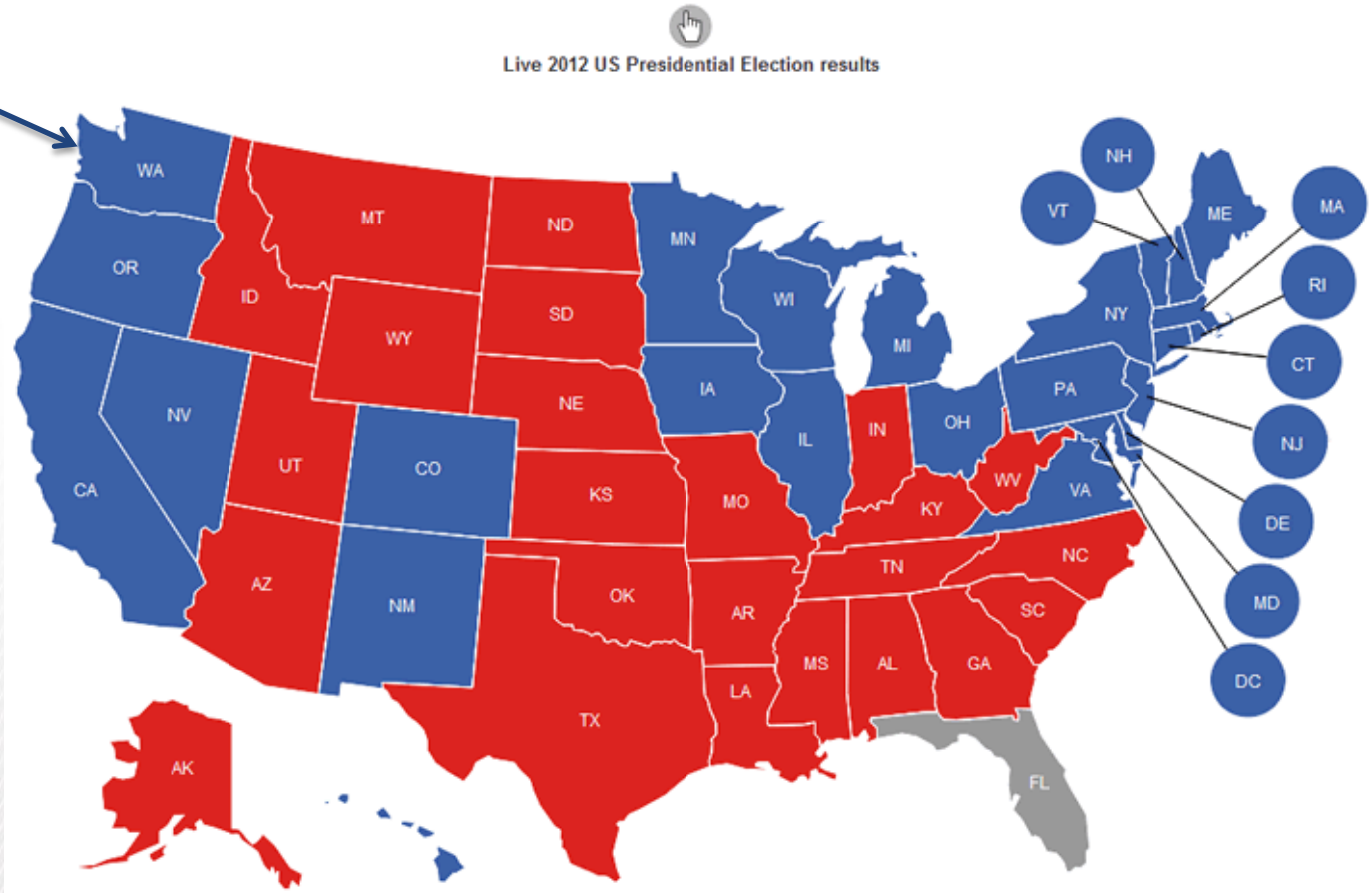
Spatial Data

- Shape



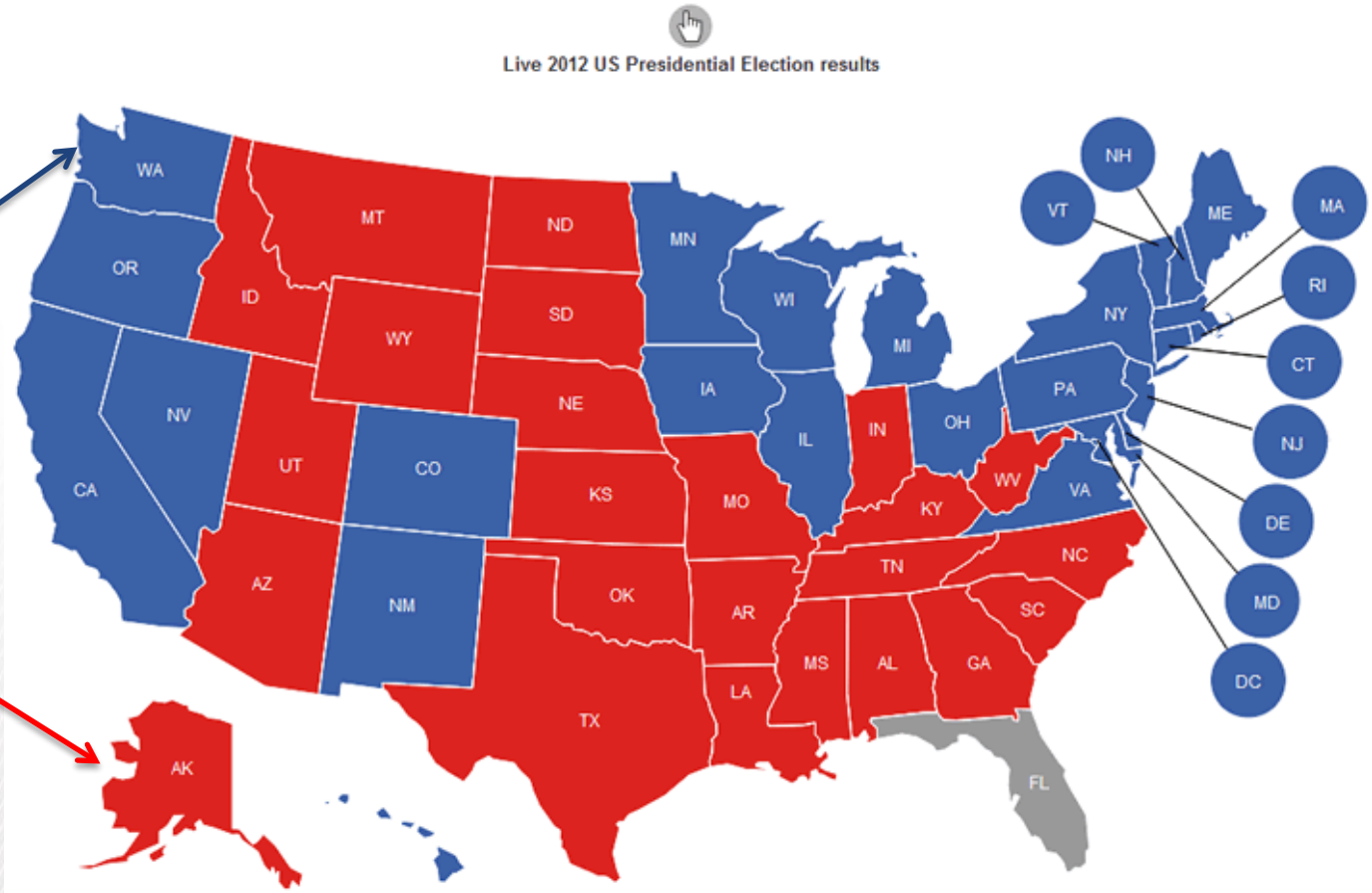
Query: Identify

How did WA vote?



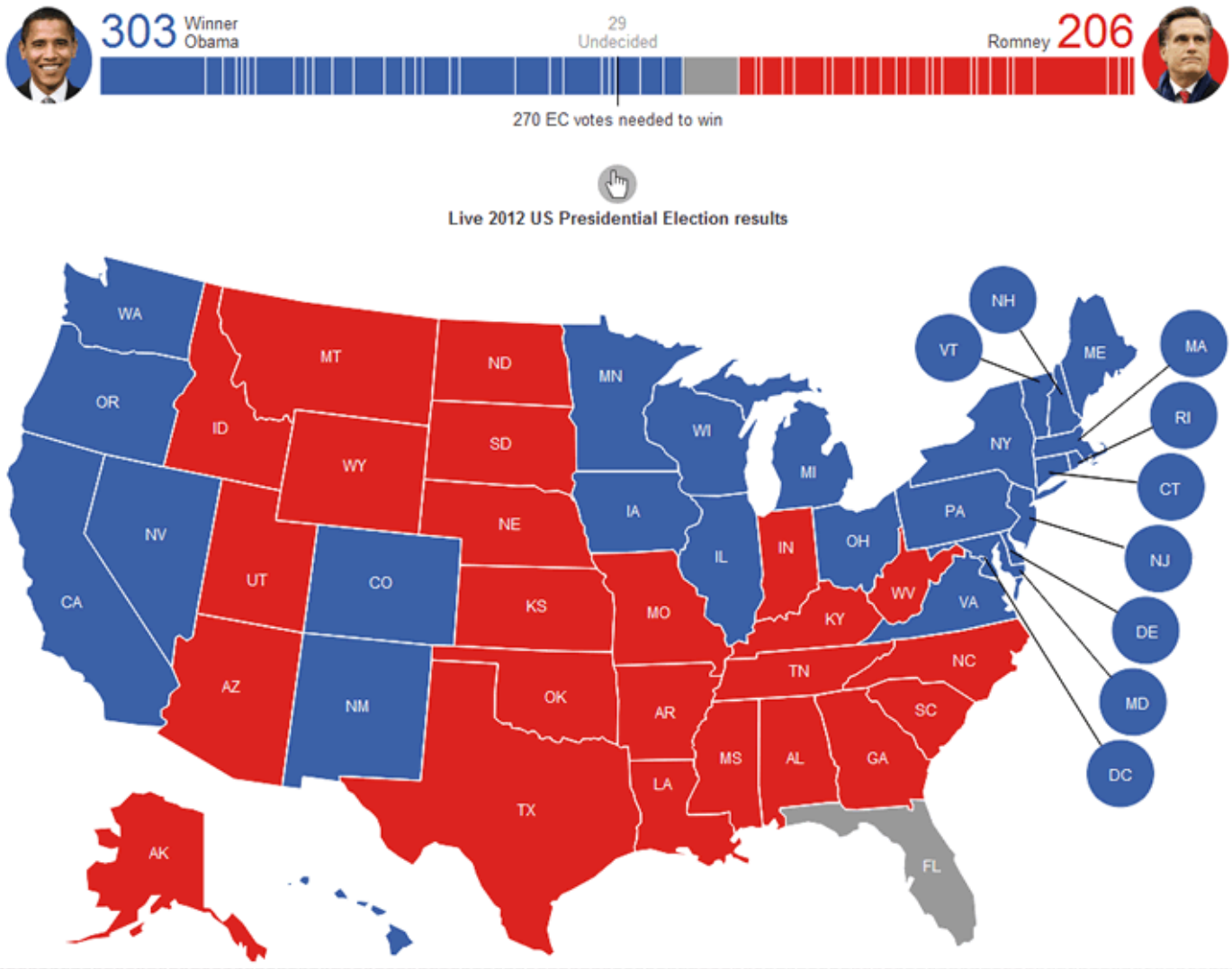
Query: Compare

How did WA vote compared to AK?

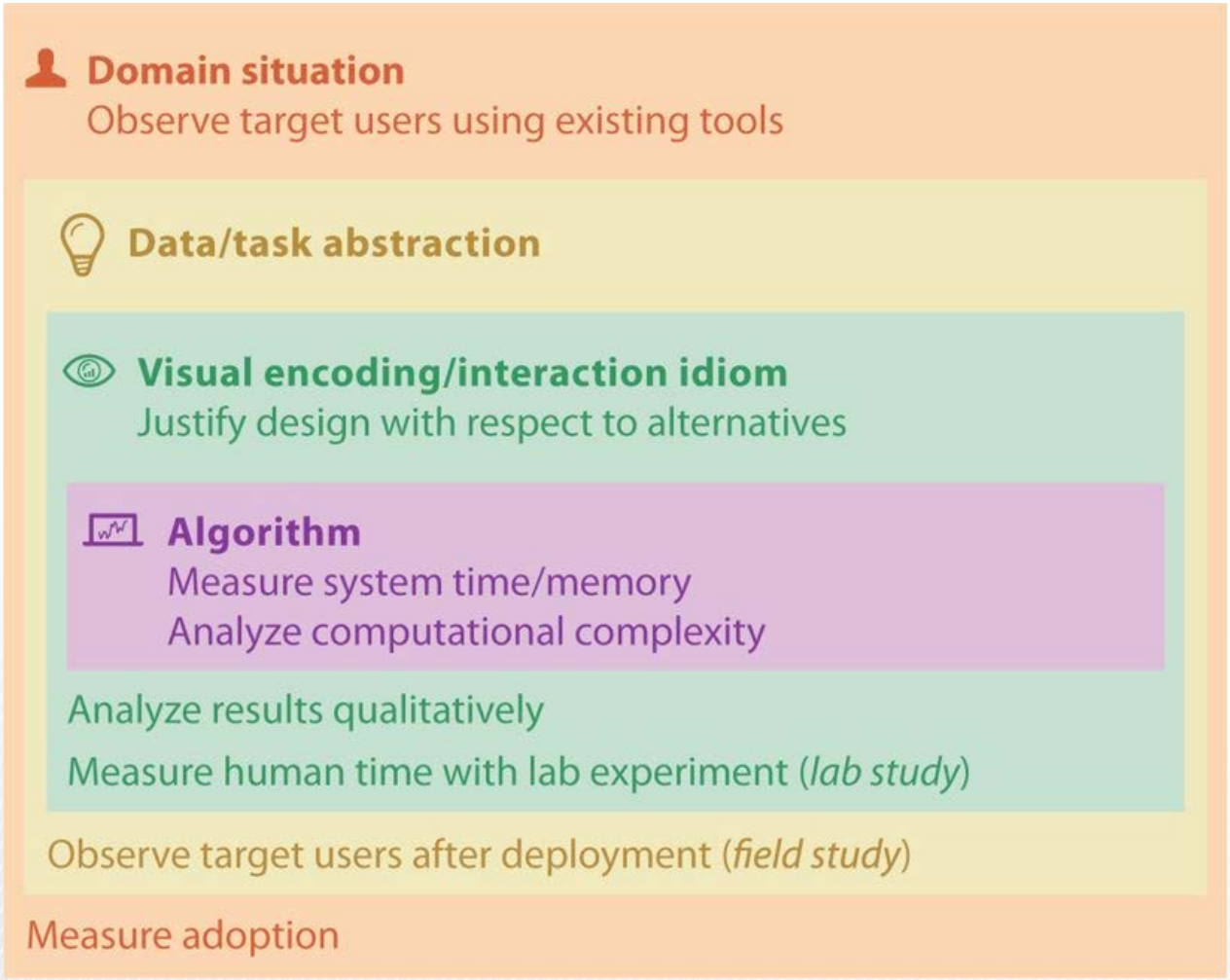


Query: Summarize

Who won the most votes **overall**?



Validation



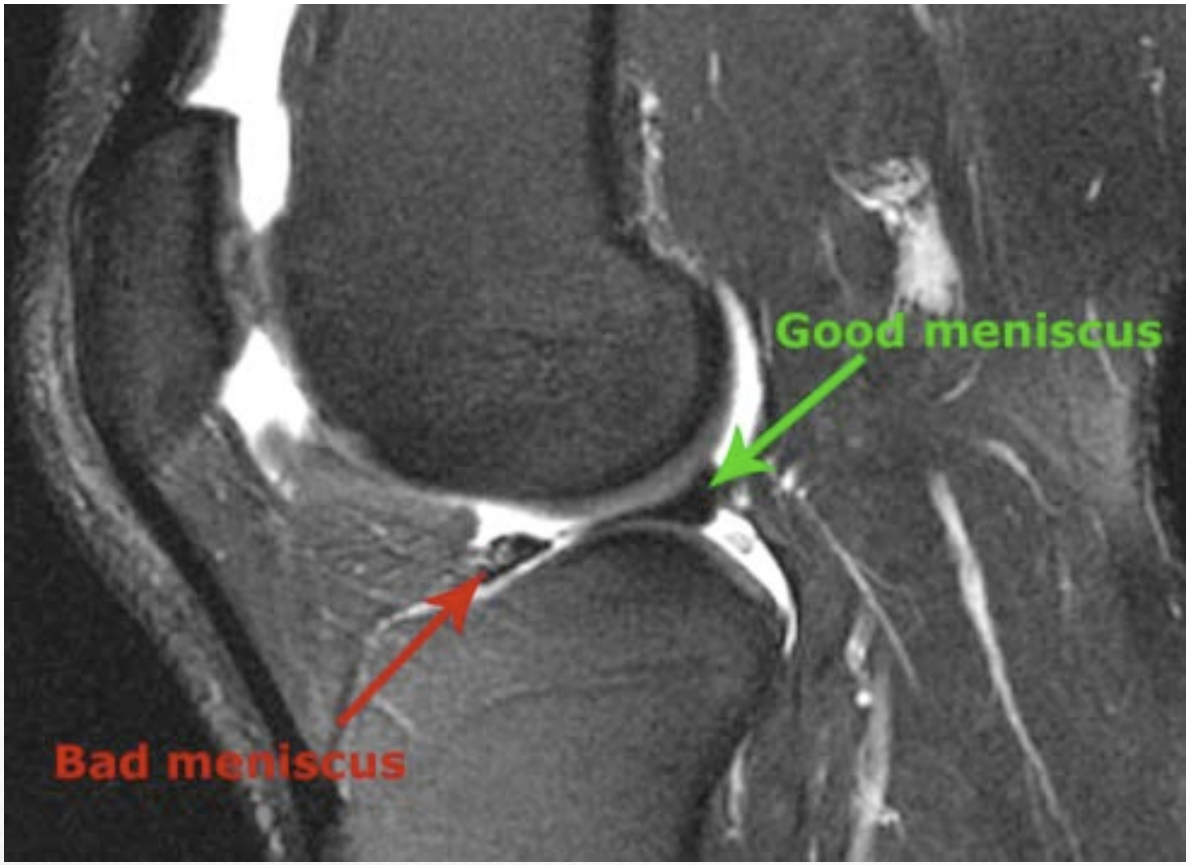
Who are they and **Why** do they care?

What data do they have and **what** are they doing?

How is the data shown to them?

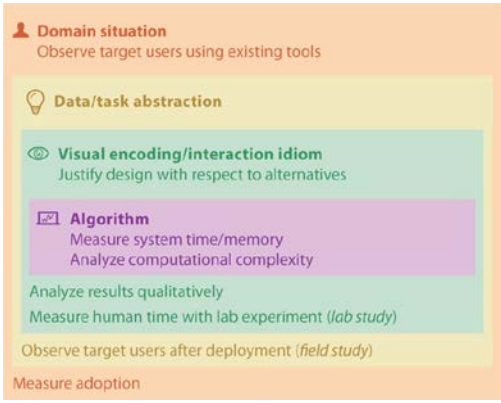
Does it work?

The Doctor Example



Group Exercise

Come up with your own example:



1. Who is the user and why do they care?

2. What data do they have and what are they trying to do with it?

3. How can the data be shown to them to help them do their job effectively?

4. What needs to happen so that the visual system doesn't fail them?

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

I.

Theoretical:

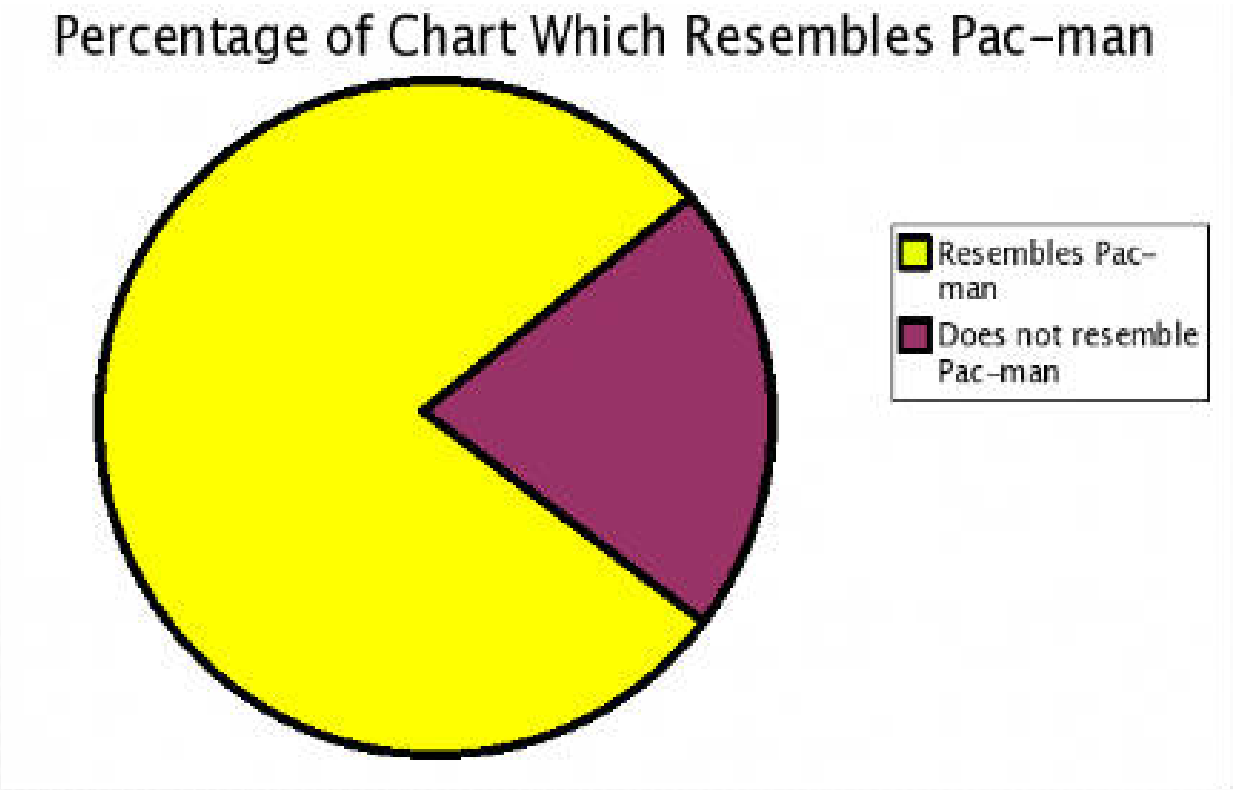
Task Abstraction
Validation

II.

Practical

Part-to-Whole
Rates & Ratios

Week 3
Oct 27, 2015


Percentage of Chart Which Resembles Pac-man
Resembles Pac-man
Does not resemble Pac-man

Comparative Quotients: 4 Types

ratios

proportions

rates

percentages

Comparative Quotients: Formulas

ratios

$$\frac{\text{Value \#1}}{\text{Value \#2}}$$

proportions

$$\frac{\text{Partial amount}}{\text{Whole amount}}$$

rates

$$\frac{\text{Value \#1}}{\text{Value \#2}} \text{ (different units)}$$

percentages

$$\frac{\text{Partial amount}}{\text{Whole amount}} \times 100$$

Comparative Quotients: Examples

ratio

Screen resolution

proportion

Baseball batting average

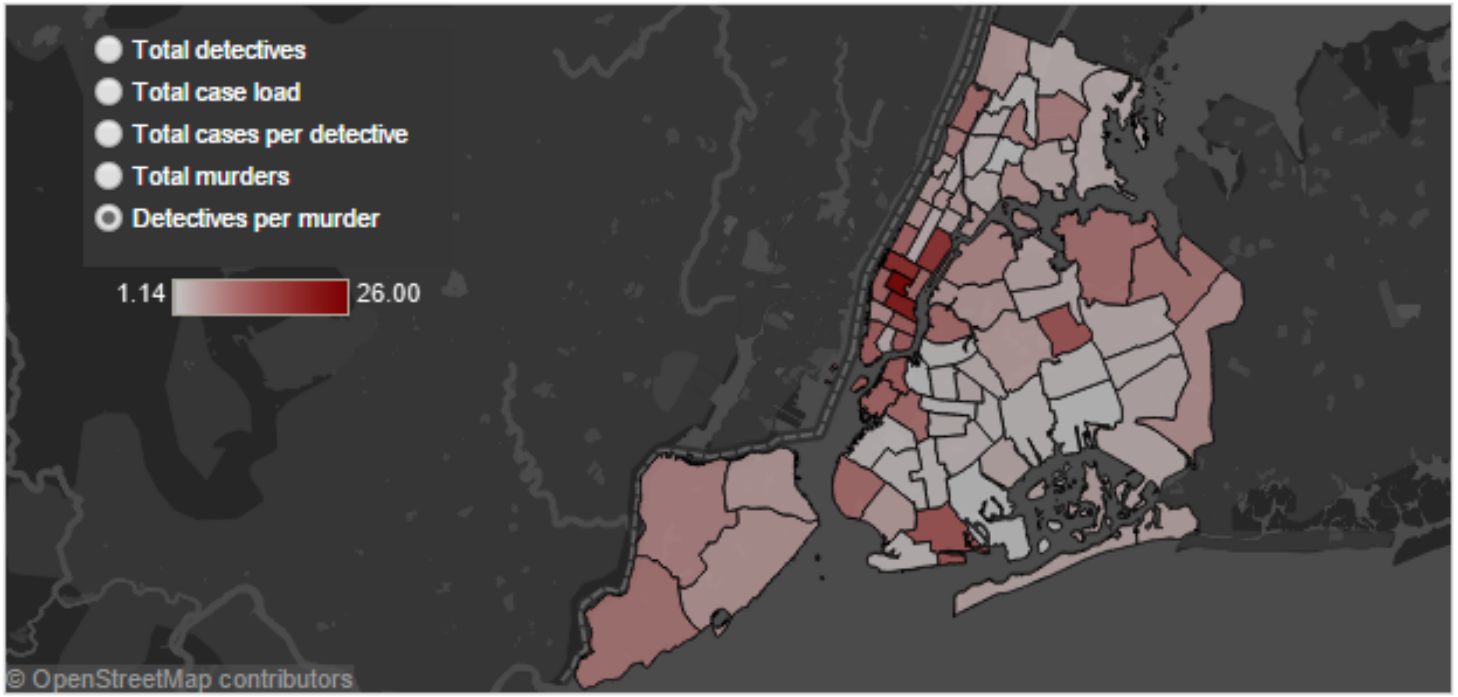
rate

Miles per hour

percentage

Market share

PRECINCT DETECTIVES AND MURDERS: The NYPD's detective force has dropped by about 2,000 since Sept. 11, and specialists were done away with in the precincts, so detectives could "catch" anything from a mugging to a murder. This chart looks at the number of detectives in each precinct and their caseload in 2013.



Precinct Name	Detectives	Total cases	Cases per detective	Detectives per murder	Total Murders	OPEN CLOSED
East Flatbush	28	3800	136	2.3	12	
Crown Heights (north)	19	2900	153	1.6	12	
East New York	32	6200	194	1.9	17	
Brownsville	27	4200	156	2.1	13	
Soundview	28	4300	154	3.5	8	

Detective staffing numbers are from a police source. Murders are compiled using information provided by the NYPD. "Total case load" is an estimate. "Open murders" indicates no arrest has been made, "closed murders" indicates an arrest; however the suspect could later be deemed innocent. The News does not have records for one murder.

DAILY NEWS

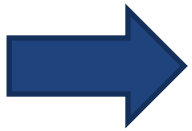
http://www.nydailynews.com/new-york/nyc-crime/forgotten-record-murder-rate-cases-unsolved-article-1.1566572

Visualizing Part-to-Whole

Market Share Data

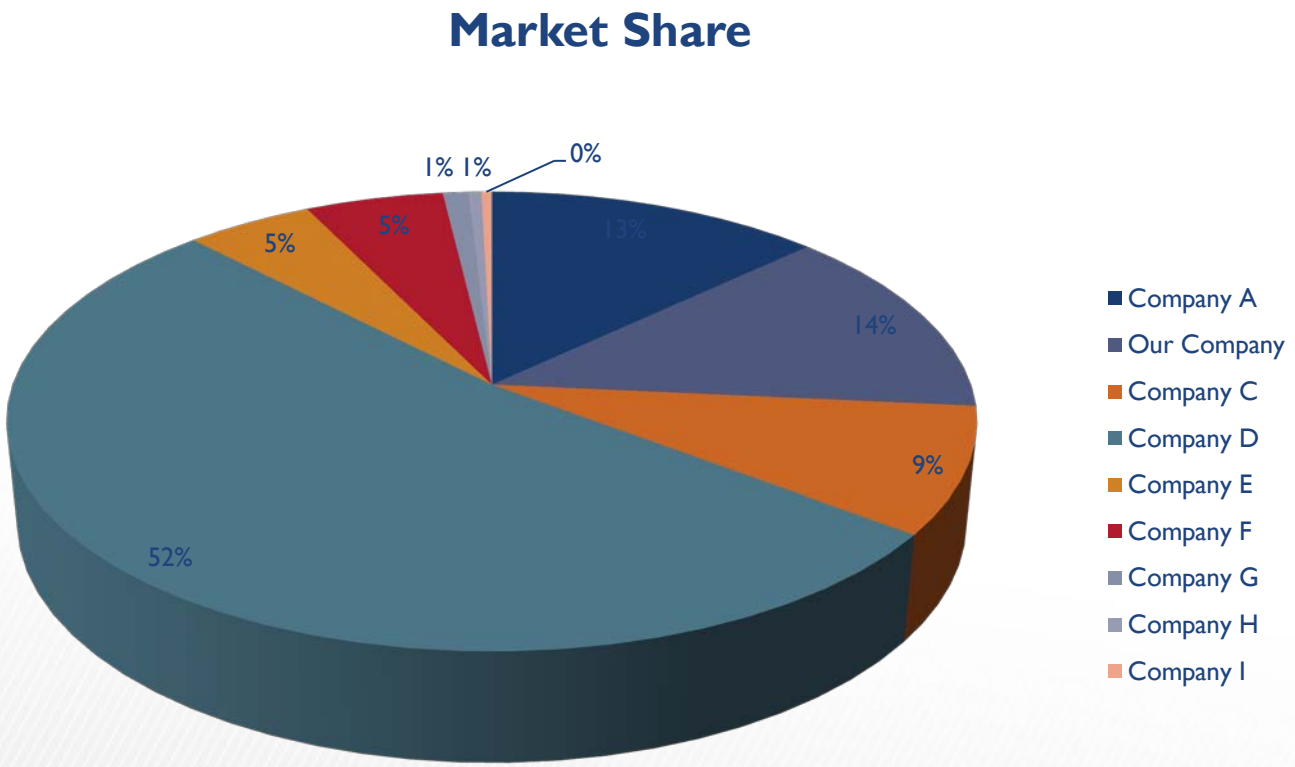
(Before)

Company A	0.1303
Our Company	0.1346
Company C	0.0885
Company D	0.5244
Company E	0.0495
Company F	0.0537
Company G	0.01
Company H	0.005
Company I	0.004

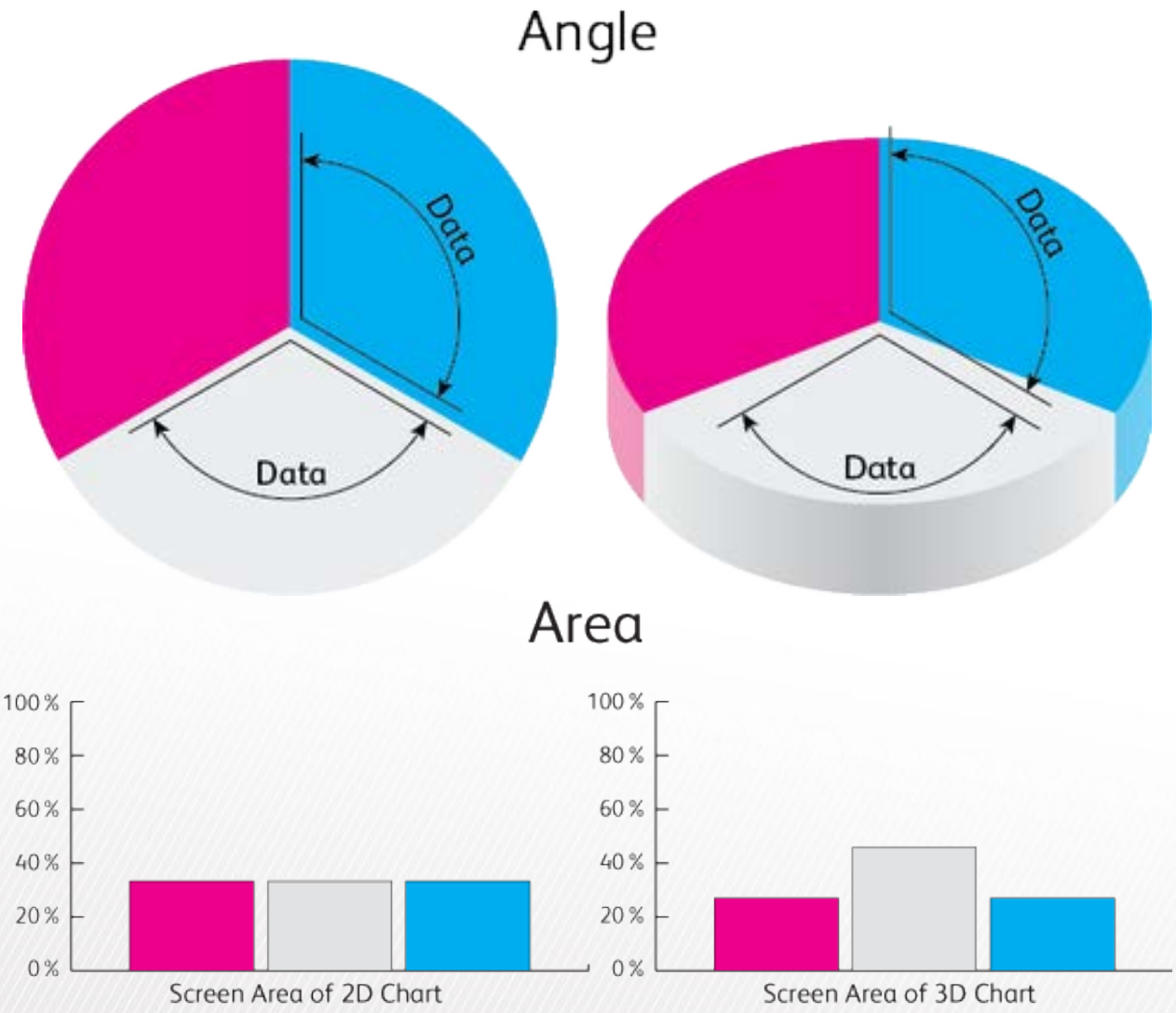


(After)

3D pie charts

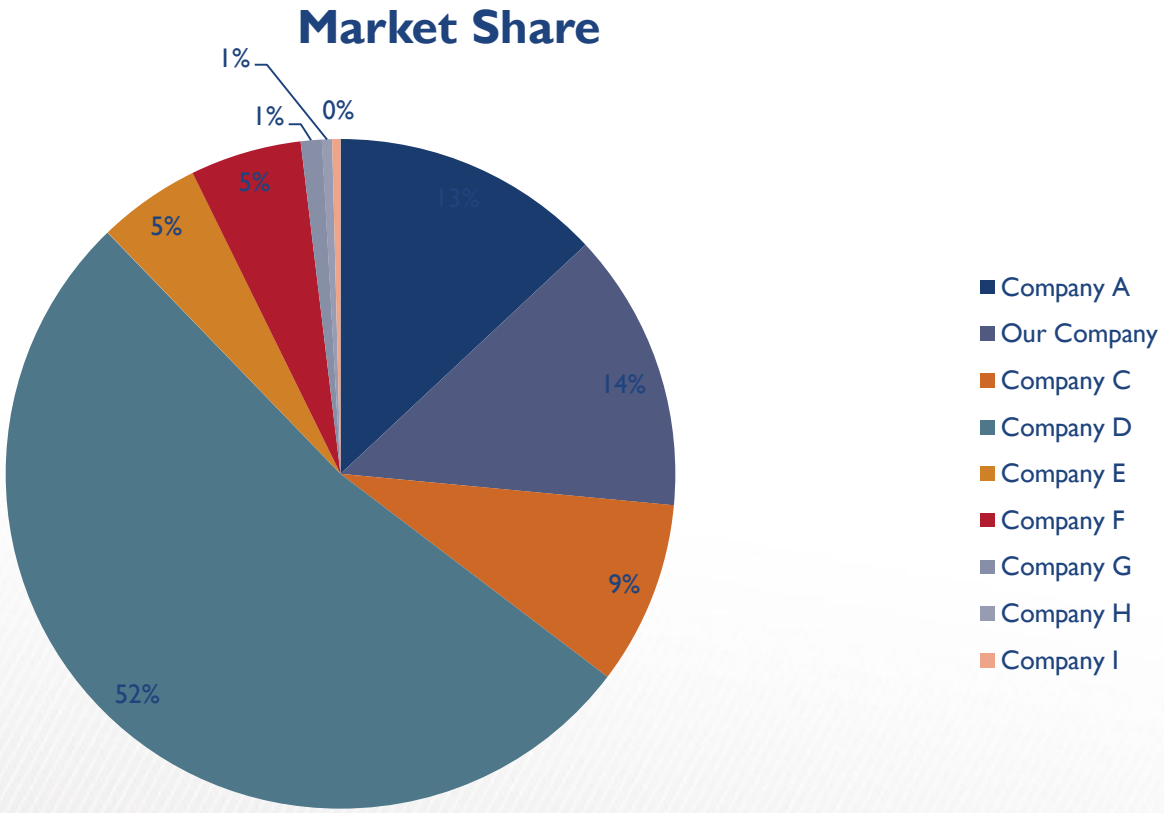


Why Not?

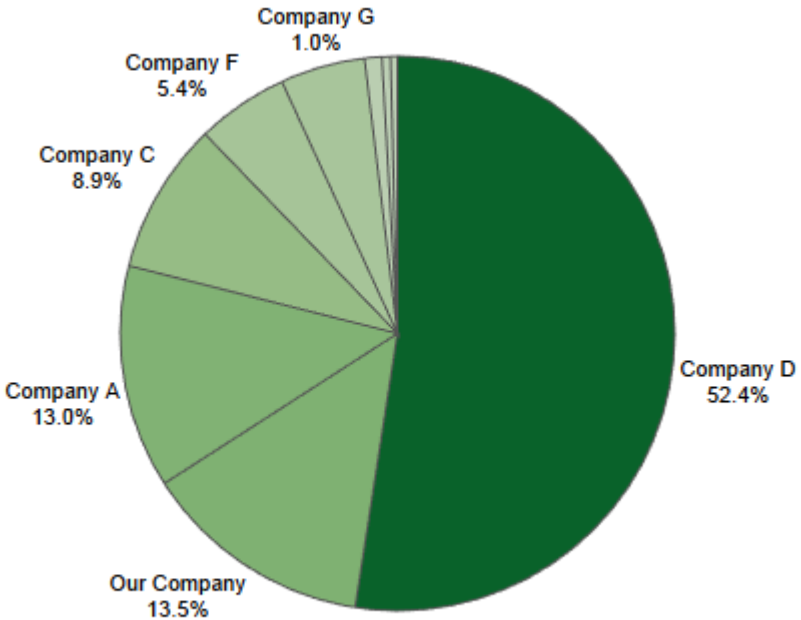


http://blog.visual.ly/2ds-company-3ds-a-crowd/

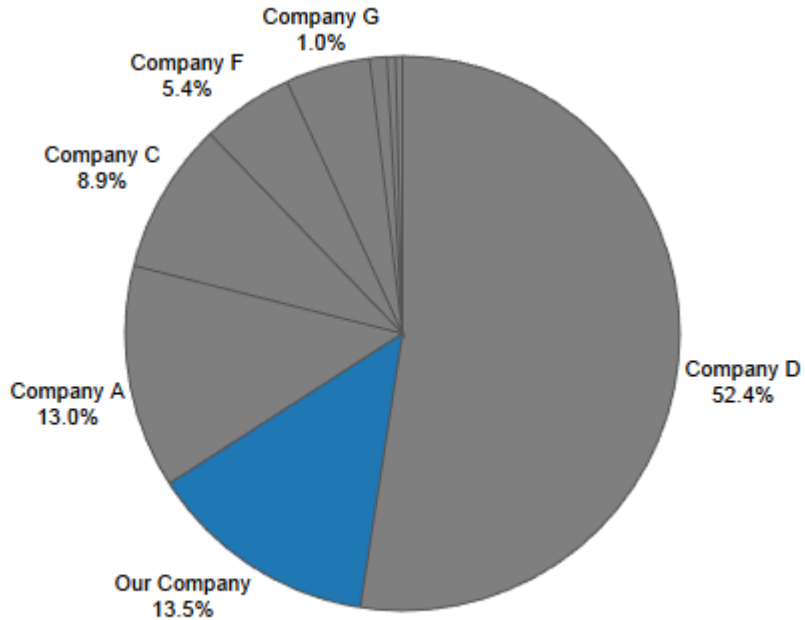
Pie charts


Market Share
1%
1%
0%
5%
5%
13%
14%
9%
52%
Company A
Our Company
Company C
Company D
Company E
Company F
Company G
Company H
Company I

Suggestions for Pie Charts

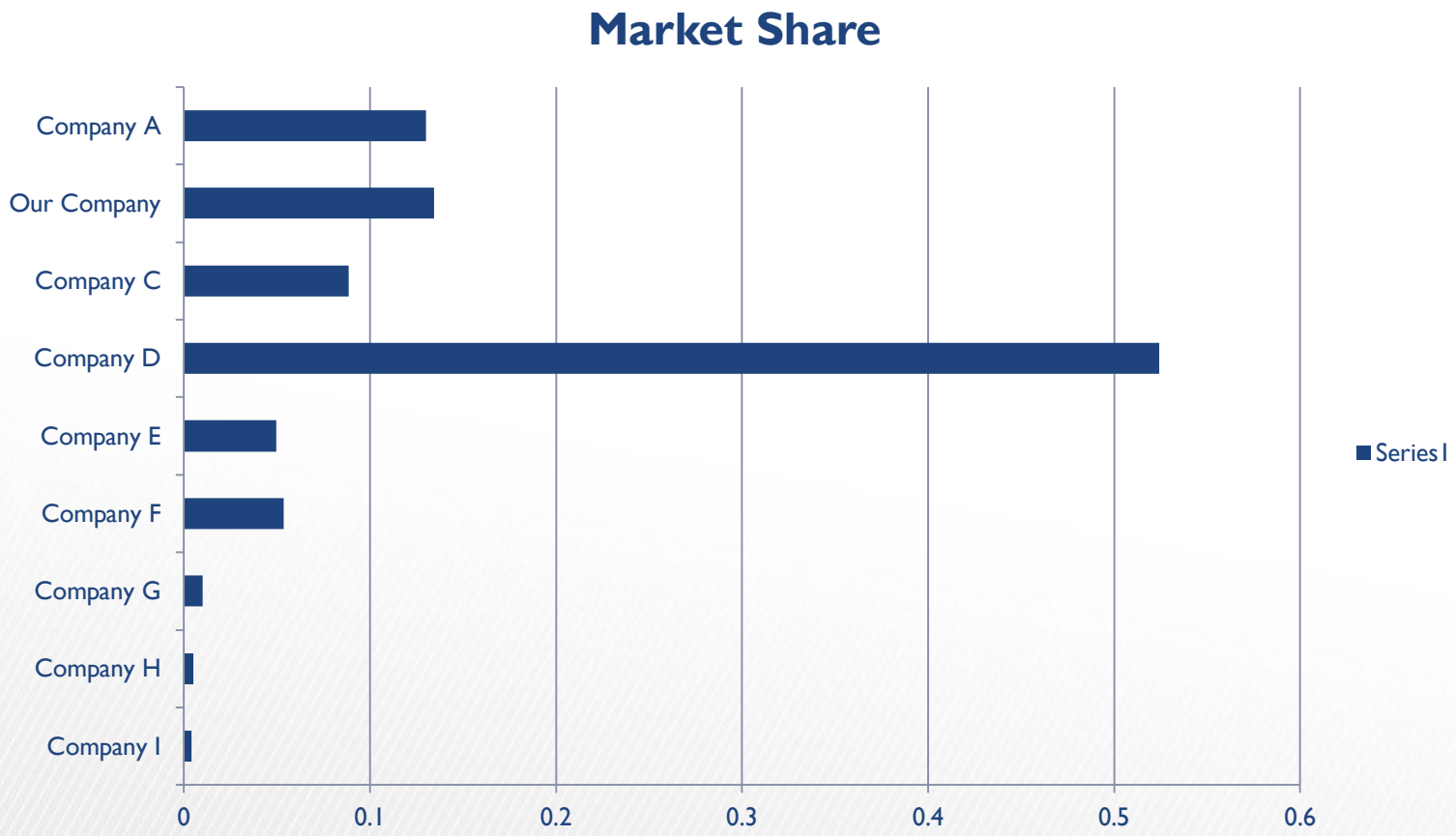


Order slices from largest to smallest
Color slices by magnitude (dual-encoded)

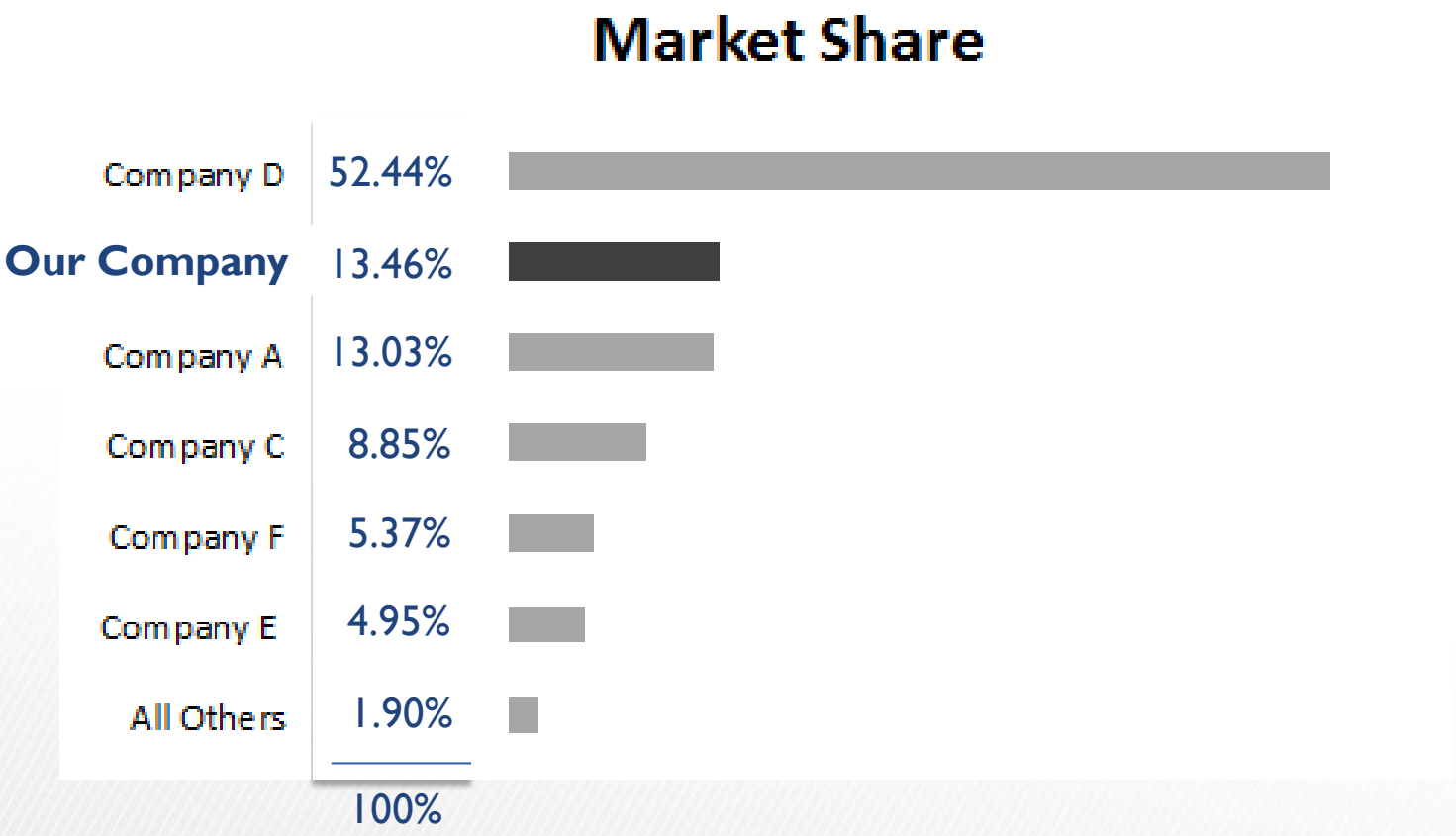


If one slice matters more, highlight it

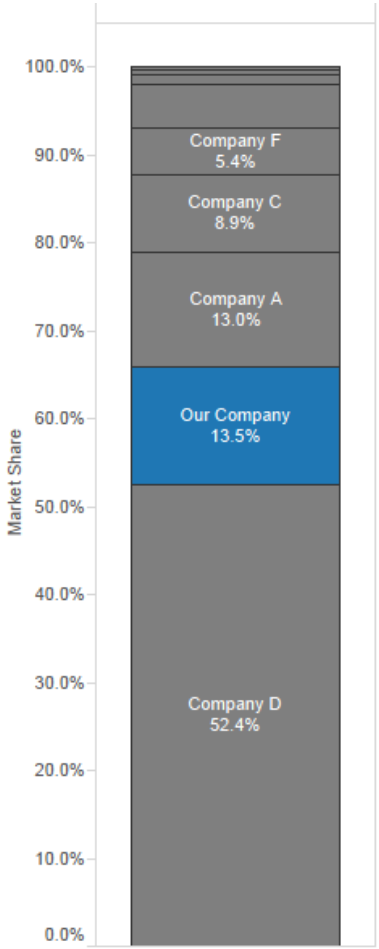
Bar charts are easier to compare

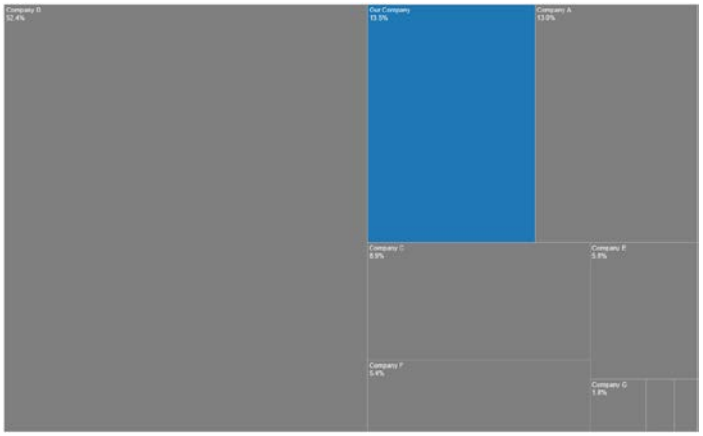


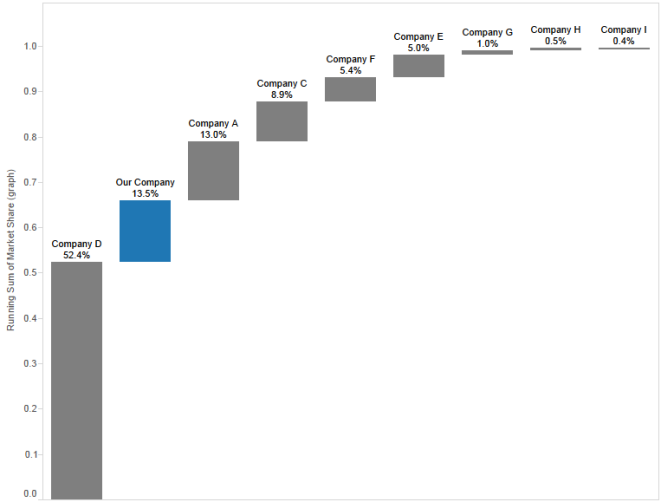
Sort in descending order

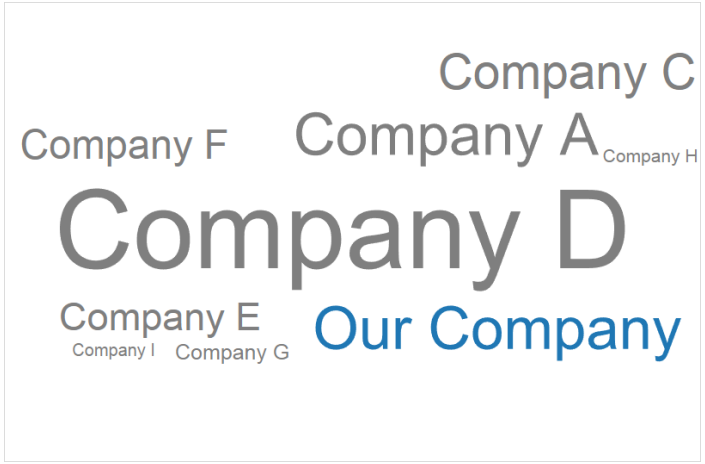


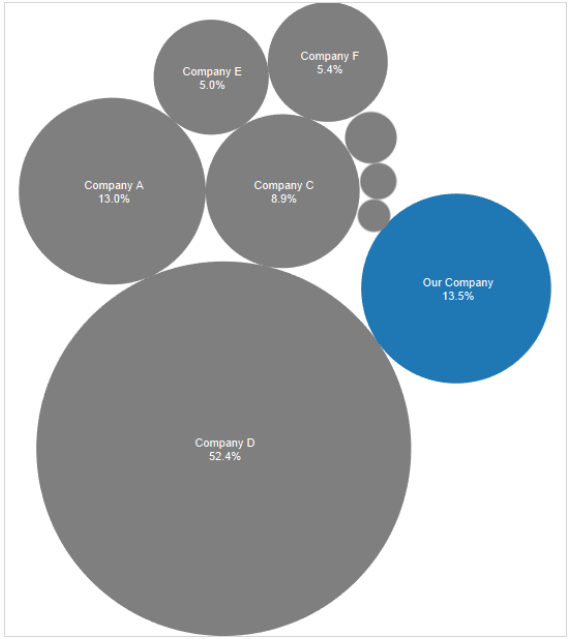
Other Part-to-Whole Visualizations

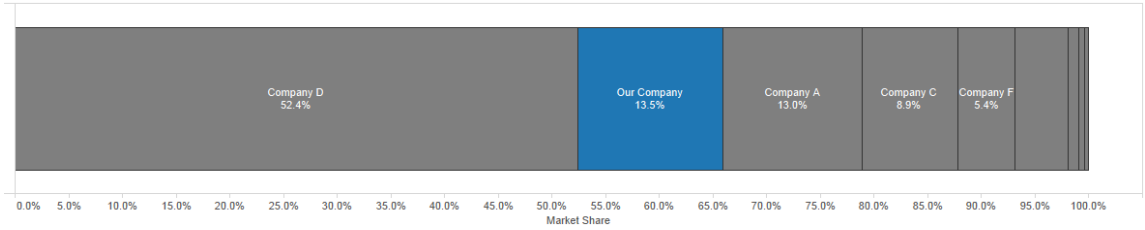








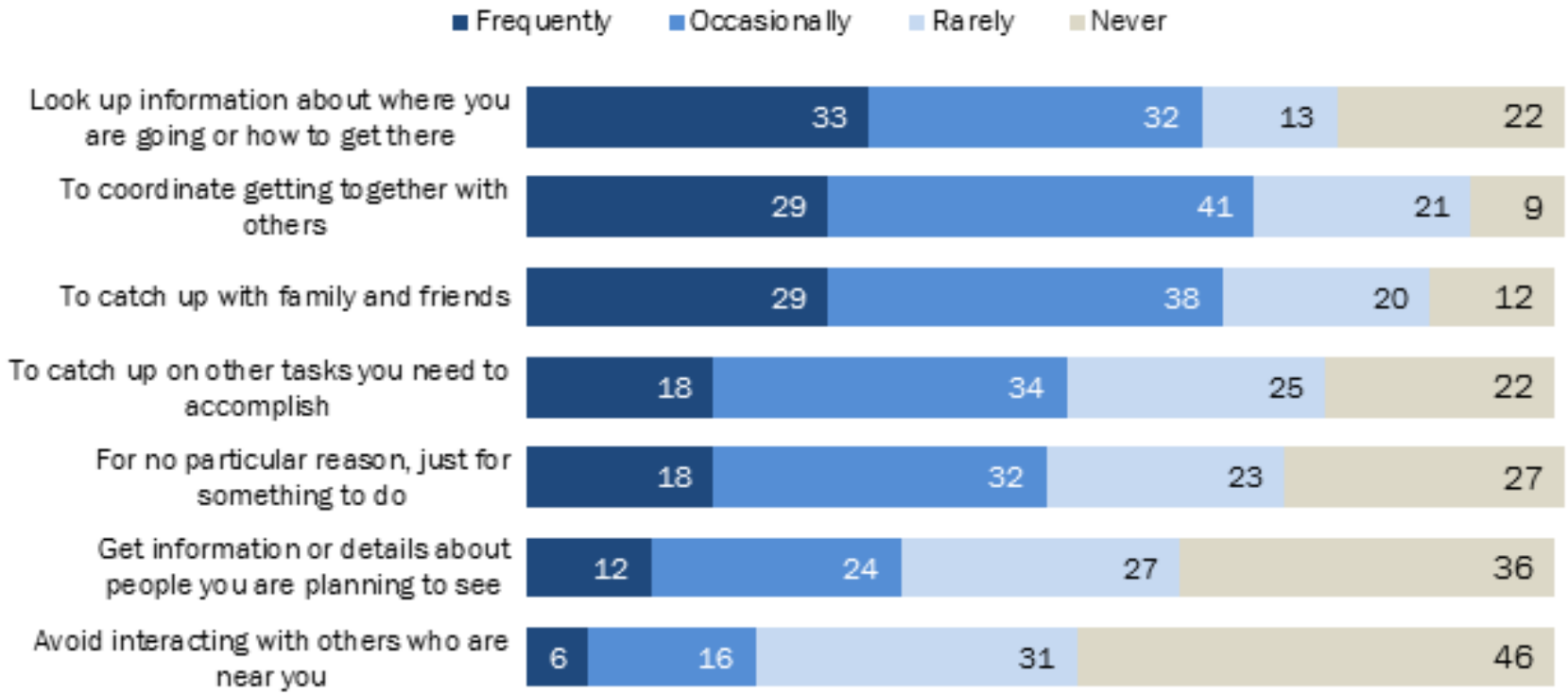




Visualizing Part-to-Whole

People Use Their Cellphones in Public for a Variety of Purposes

% of cellphone owners who do these things in public with their phones



Source: Pew Research Center American Trends Panel survey, May 30-June 30, 2014. N=3,042 cell users

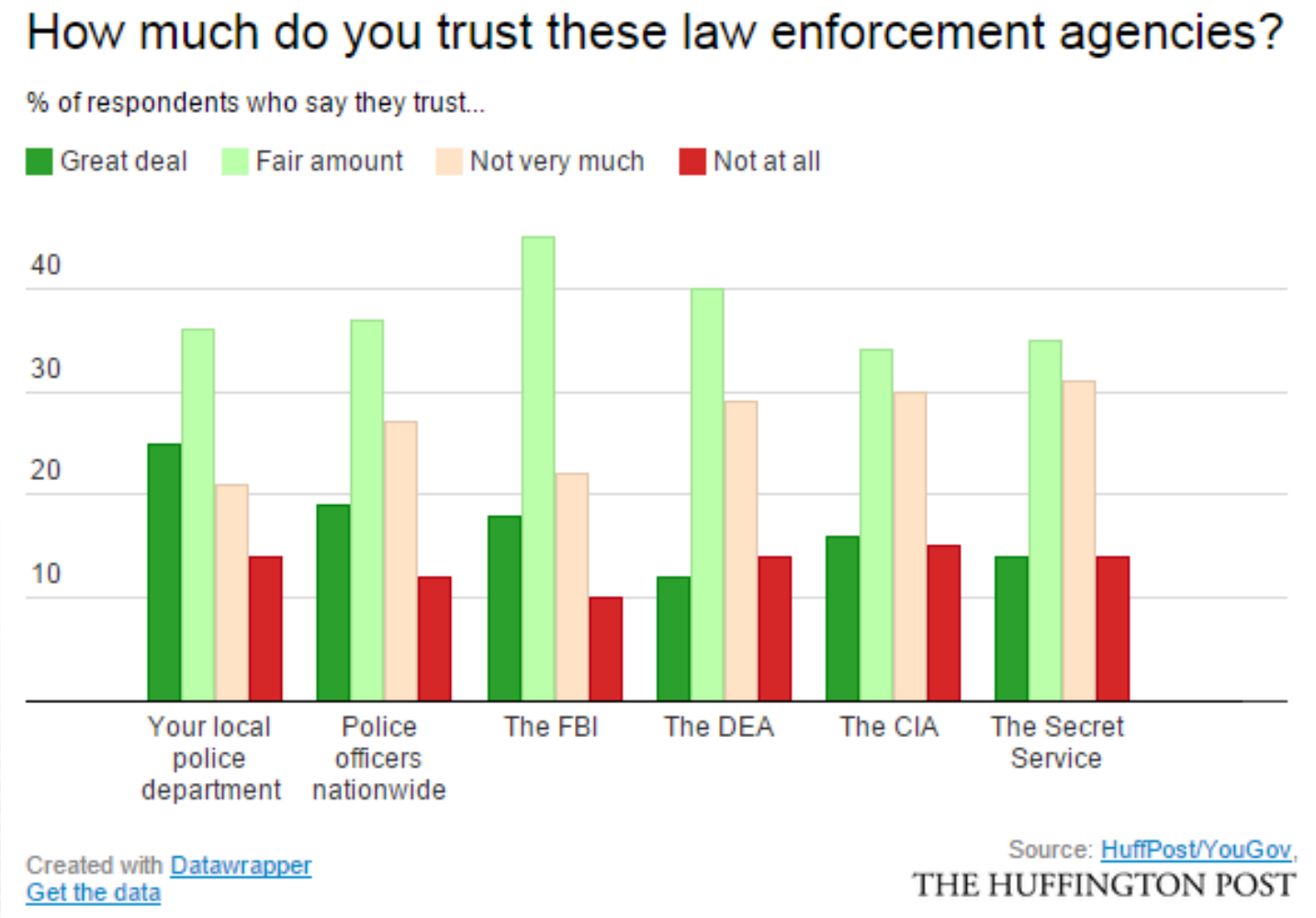
PEW RESEARCH CENTER

http://www.pewinternet.org/2015/08/26/chapter-2-phone-use-in-public-areas/

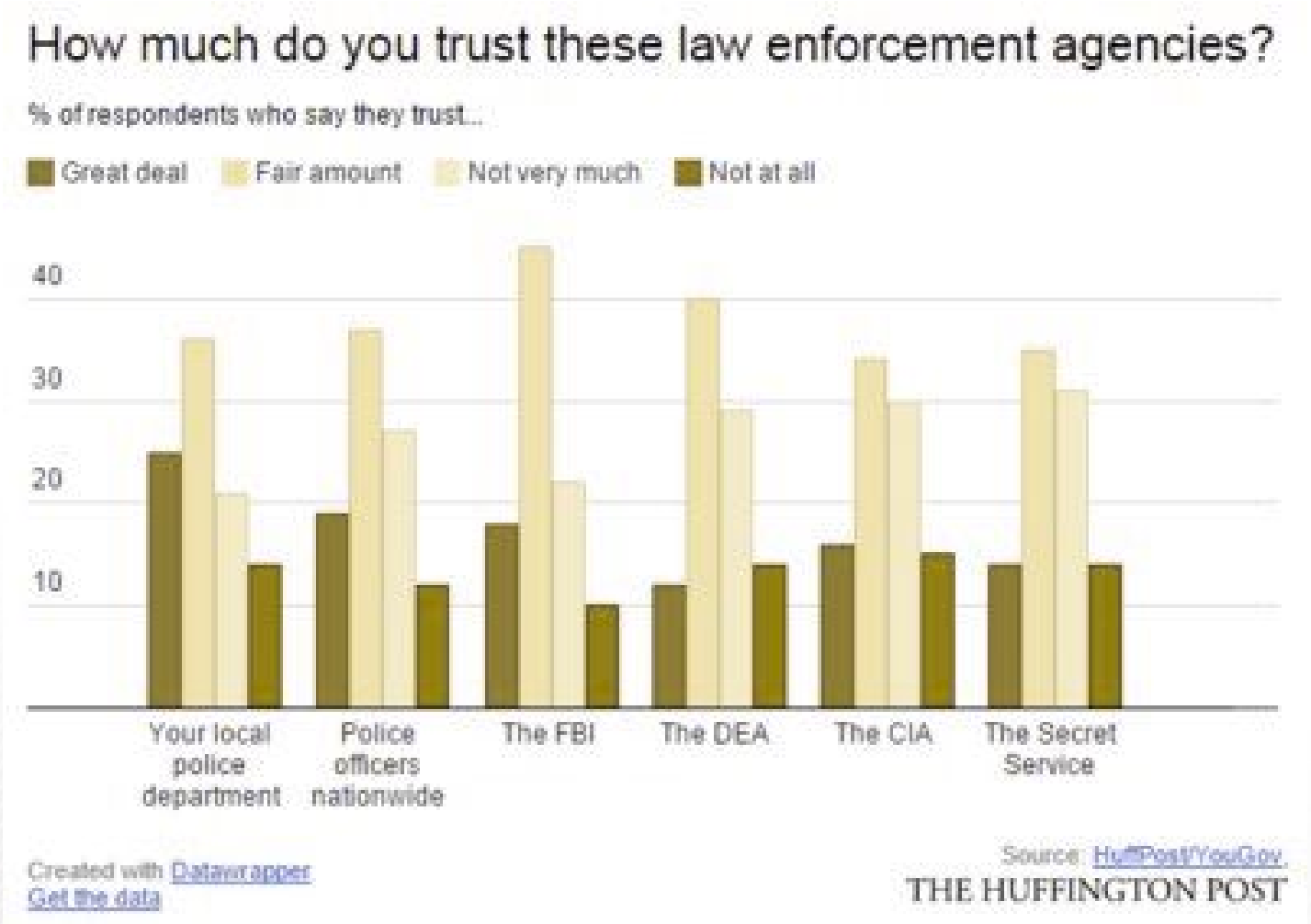

POLICE

YouGov

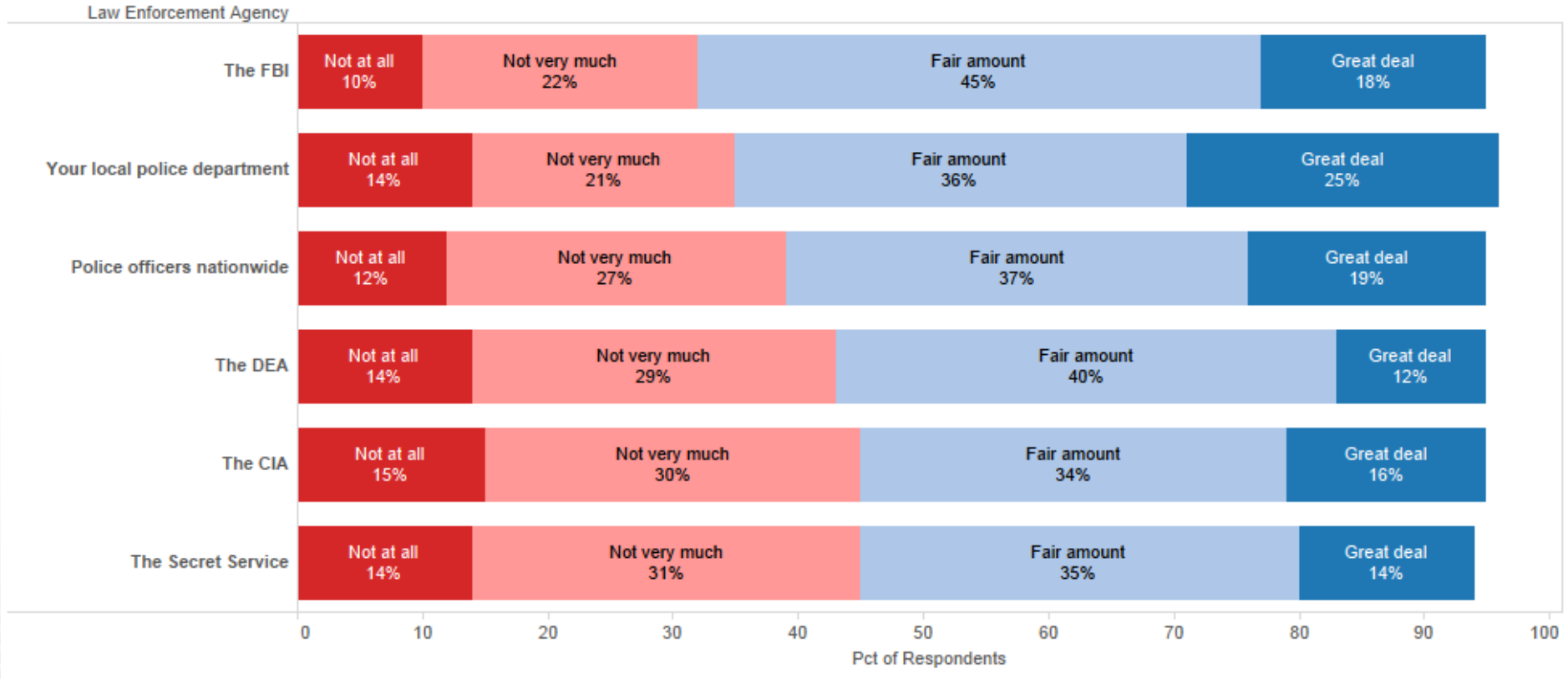
Sample	1000 Adult Interviews
Conducted	April 16 - 20, 2015
Margin of Error	±4.2%

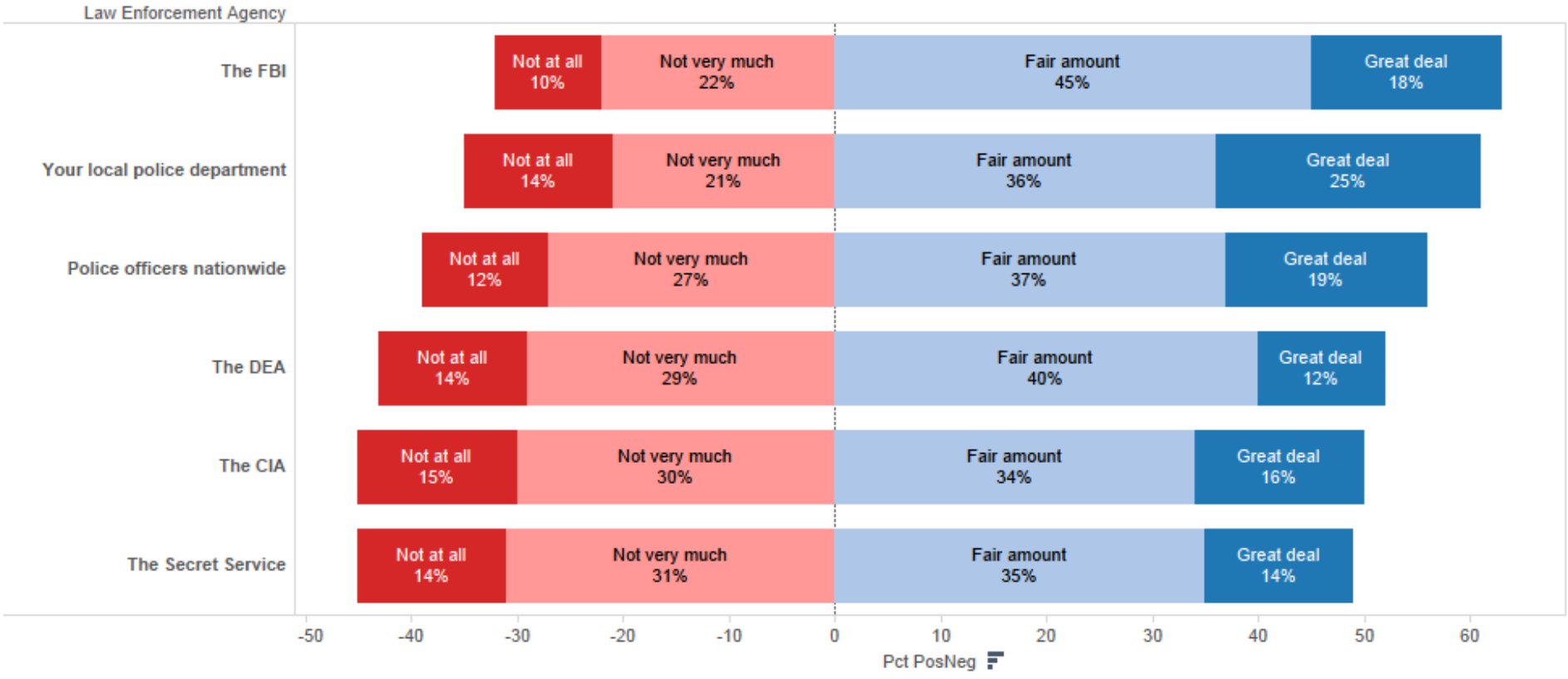


http://www.huffingtonpost.com/2015/04/22/law-enforcement-trust-poll_n_7118634.html

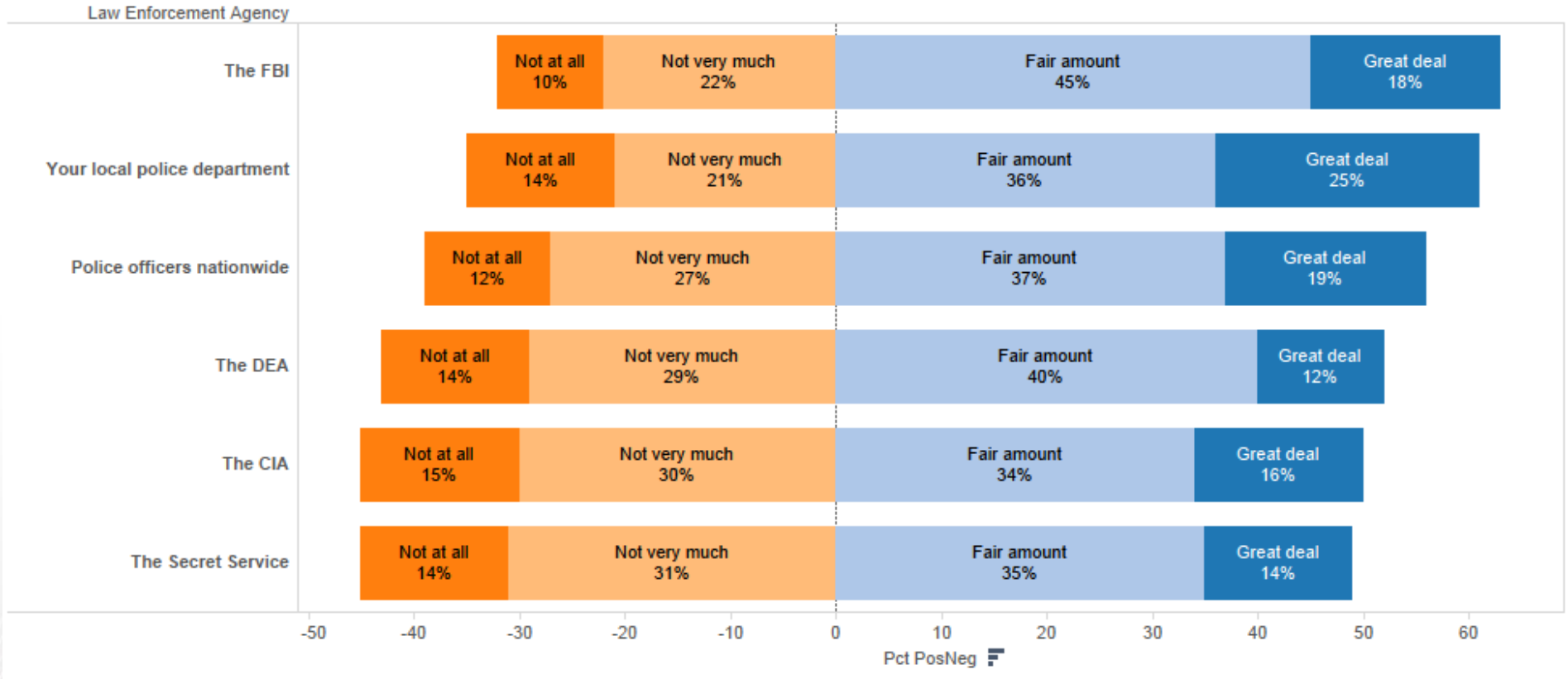


Using http://www.vischeck.com/


Law Enforcement Agency
The FBI
Not at all 10%
Not very much 22%
Fair amount 45%
Great deal 18%
Your local police department
Not at all 14%
Not very much 21%
Fair amount 36%
Great deal 25%
Police officers nationwide
Not at all 12%
Not very much 27%
Fair amount 37%
Great deal 19%
The DEA
Not at all 14%
Not very much 29%
Fair amount 40%
Great deal 12%
The CIA
Not at all 15%
Not very much 30%
Fair amount 34%
Great deal 16%
The Secret Service
Not at all 14%
Not very much 31%
Fair amount 35%
Great deal 14%
0
10
20
30
40
50
60
70
80
90
100
Pct of Respondents



http://www.huffingtonpost.com/2015/04/22/law-enforcement-trust-poll_n_7118634.html


Law Enforcement Agency
The FBI
Not at all 10%
Not very much 22%
Fair amount 45%
Great deal 18%
Your local police department
Not at all 14%
Not very much 21%
Fair amount 36%
Great deal 25%
Police officers nationwide
Not at all 12%
Not very much 27%
Fair amount 37%
Great deal 19%
The DEA
Not at all 14%
Not very much 29%
Fair amount 40%
Great deal 12%
The CIA
Not at all 15%
Not very much 30%
Fair amount 34%
Great deal 16%
The Secret Service
Not at all 14%
Not very much 31%
Fair amount 35%
Great deal 14%
-50 -40 -30 -20 -10 0 10 20 30 40 50 60
Pct PosNeg

Week 3 Homework

- Read the chapters for the Week 4 Lecture:
  - Munzner, *Visualization Analysis & Design*, Ch. 5
  - Jones, *Communicating Data with Tableau*, Ch. 6-7

- Visualization
  - Find a dataset that includes either a rate, ratio, percent or proportion.
  - Create a visualization using this dataset (static or interactive)
  - Write a single paragraph describing an imaginary user, their task, why the visual encoding that you created supports that task, and how the system needs to work to be effective