Data Visualization - Basics

The Visualization Process
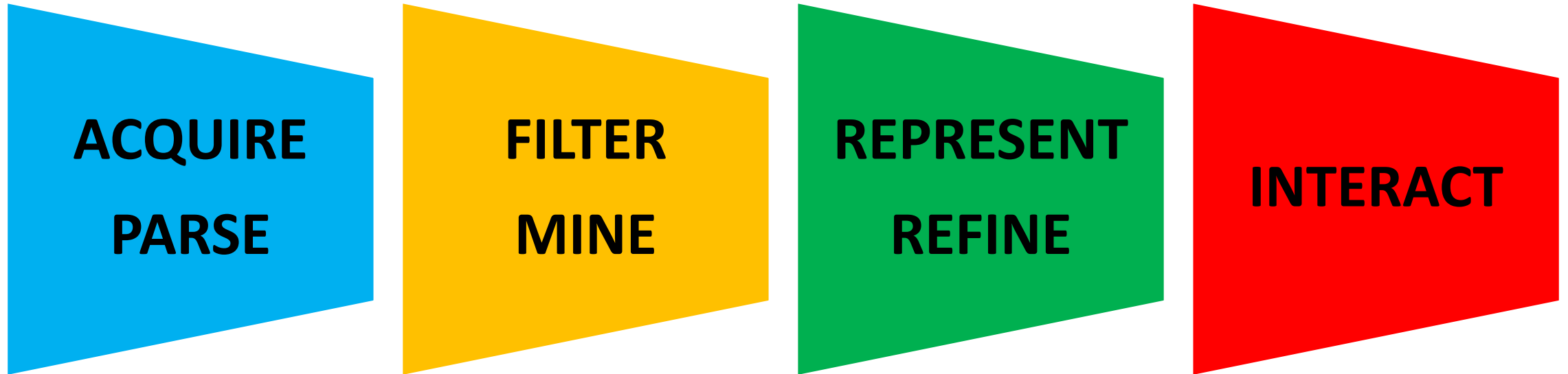
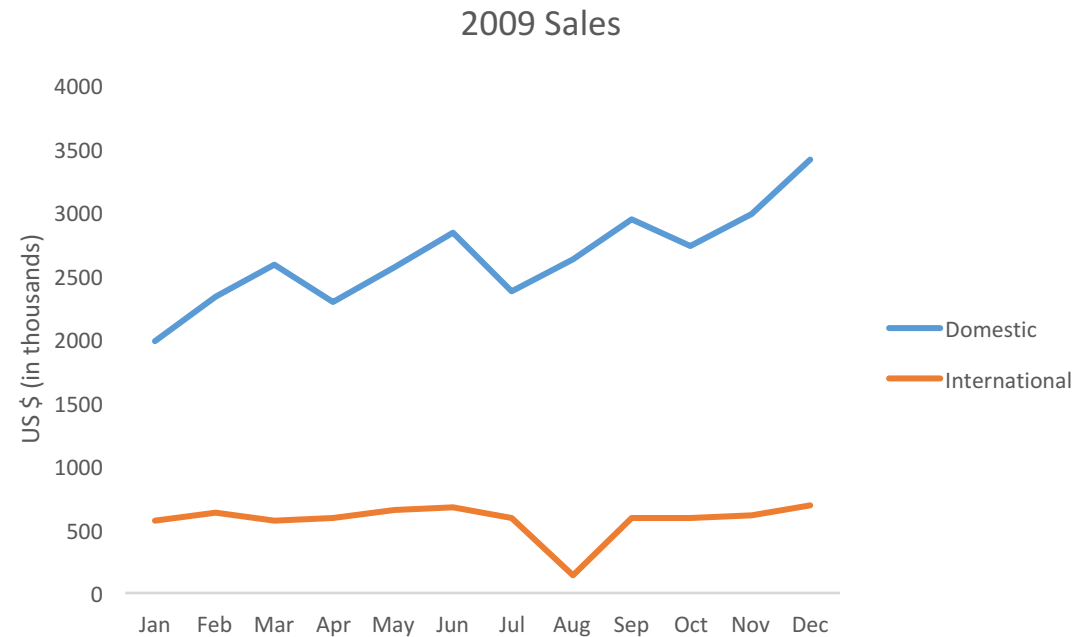# Sales Data
(US $ in thousands)

Region	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Domestic	1983	2343	2593	2283	2574	2838	2382	2634	2938	2739	2983	3413
International	574	636	573	593	644	679	593	139	599	583	602	690

Sales Data
(US $ in thousands)

Region	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Domestic	1983	2343	2593	2283	2574	2838	2382	2634	2938	2739	2983	3413
International	574	636	573	593	644	679	593	139	599	583	602	690

2009 Sales

Causes of Deaths in US in 2007

Cause of Death	Deaths per year	%
Heart Disease	616067	25.42
Cancer	562875	23.22
Stroke	135952	5.61
Chronic Lower Respiratory Diseases	127924	5.28
Accidents	123706	5.10
Alzheimer's Disease	74632	3.08
Diabetes	71382	2.95
Influenza and Pneumonia	52717	2.18
Nephritis and Nephrosis	46448	1.92
Septicemia	34828	1.44
Other Causes	577181	23.81
Totals	2423712	

Causes of Deaths in US in 2007

Cause of Death	Deaths per year	%
Heart Disease	616067	25.42
Cancer	562875	23.22
Stroke	135952	5.61
Chronic Lower Respiratory Diseases	127924	5.28
Accidents	123706	5.10
Alzheimer's Disease	74632	3.08
Diabetes	71382	2.95
Influenza and Pneumonia	52717	2.18
Nephritis and Nephrosis	46448	1.92
Septicemia	34828	1.44
Other Causes	577181	23.81
Totals	2423712	

Total Deaths in America by Cause

- Heart Disease
- Cancer
- Stroke
- Chronic Lower Respiratory Diseases
- Accidents
- Alzheimer's Disease
- Diabetes
- Influenza and Pneumonia
- Nephritis and Nephrosis
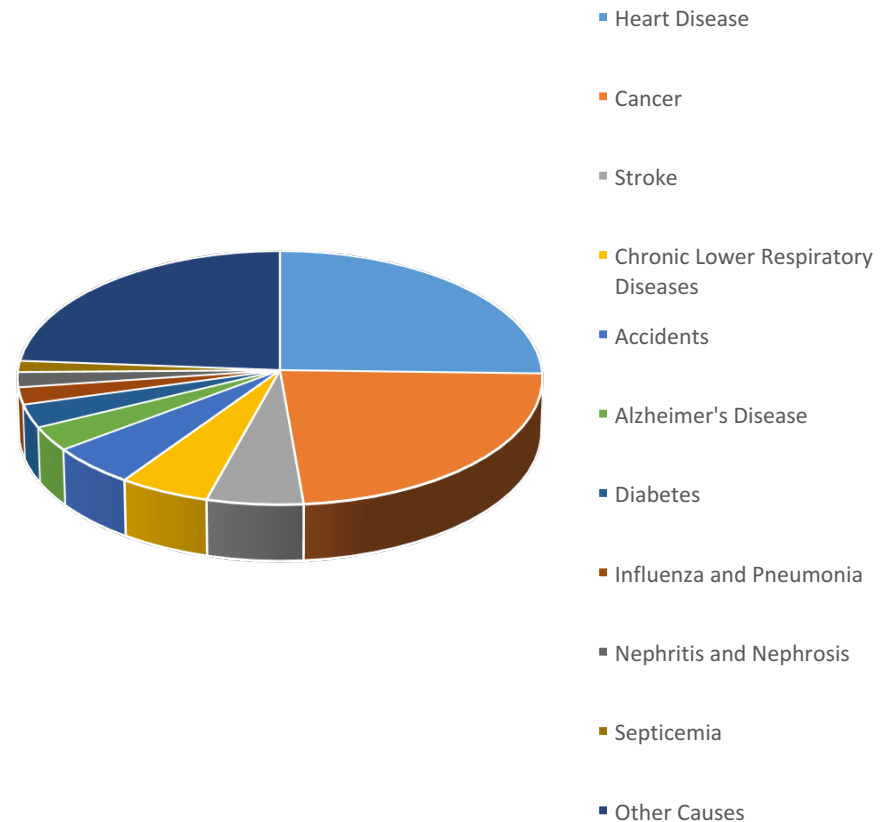- Septicemia
- Other Causes

Causes of Deaths in US in 2007

Total Deaths in America by Cause

Criteria for effectiveness

- Clearly indicates the nature of the relationship?

- Represents the quantities accurately?

- Makes it easy to compare quantities?

- Makes it easy to see ranked order of values?

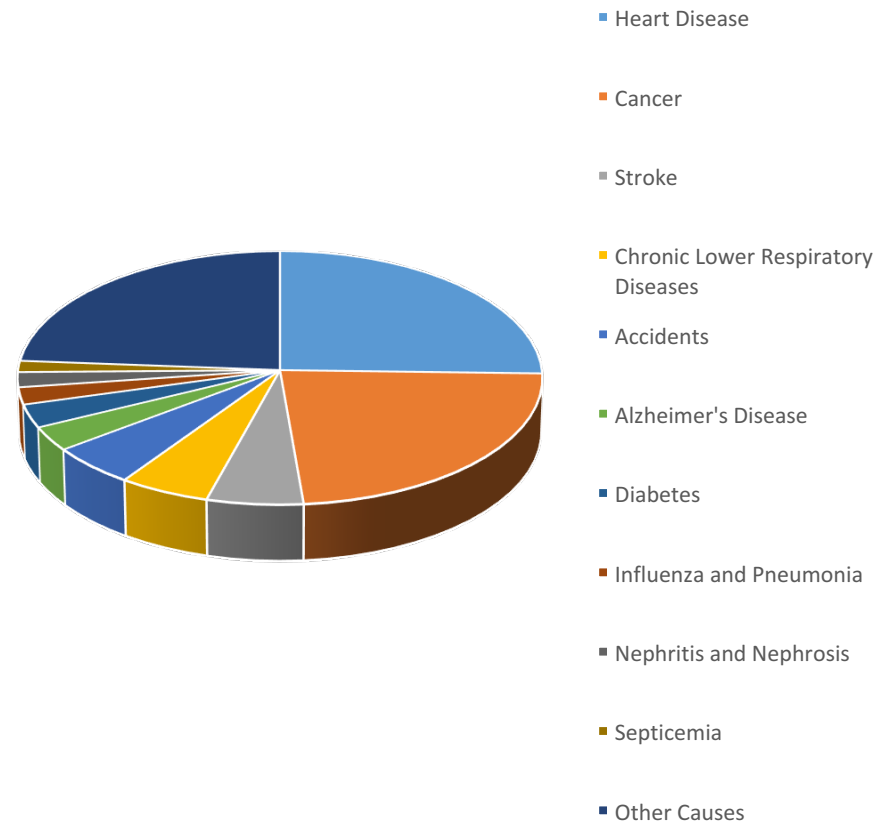- Makes obvious how people should use the information?

Legend:
- Heart Disease
- Cancer
- Stroke
- Chronic Lower Respiratory Diseases
- Accidents
- Alzheimer's Disease
- Diabetes
- Influenza and Pneumonia
- Nephritis and Nephrosis
- Septicemia
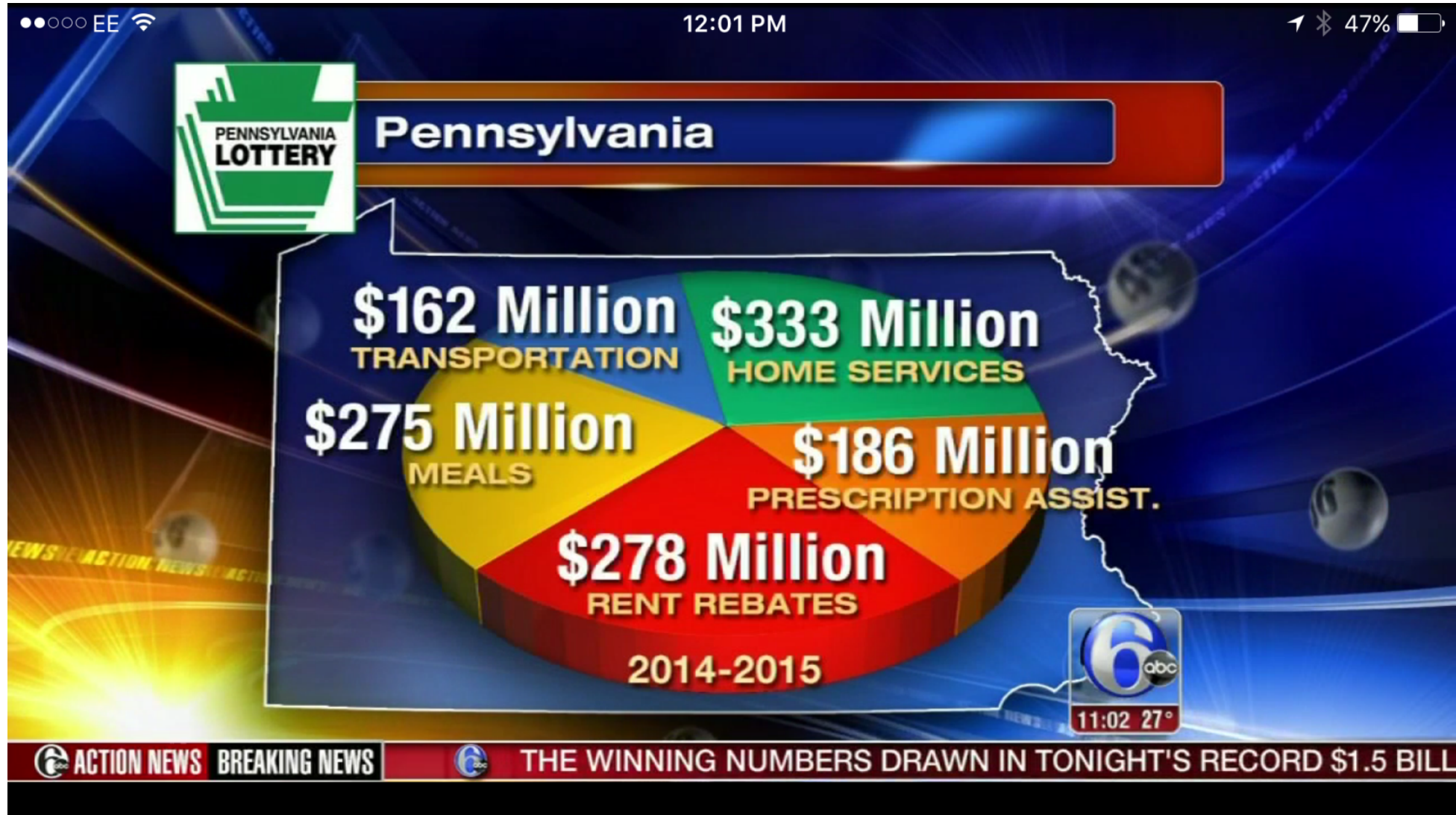- Other Causes

Causes of Deaths in US in 2007

Total Deaths in America by Cause

Criteria for effectiveness

- Clearly indicates the nature of the relationship? Yes

- Represents the quantities accurately? No

- Makes it easy to compare quantities? No

- Makes it easy to see ranked order of values? No

- Makes obvious how people should use the information? Partially

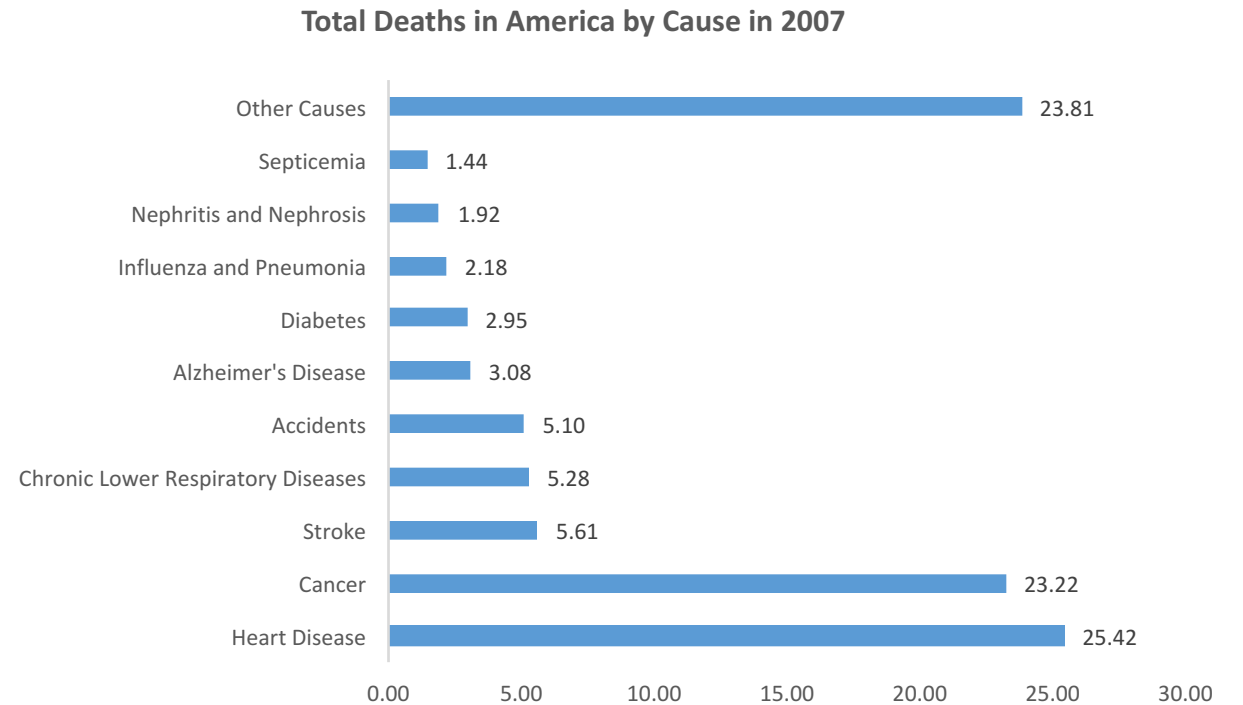- Heart Disease
- Cancer
- Stroke
- Chronic Lower Respiratory Diseases
- Accidents
- Alzheimer's Disease
- Diabetes
- Influenza and Pneumonia
- Nephritis and Nephrosis
- Septicemia
- Other Causes

3-D Pie Charts – The Worst!

Causes of Deaths in US in 2007

Cause of Death	Deaths per year	%
Heart Disease	616067	25.42
Cancer	562875	23.22
Stroke	135952	5.61
Chronic Lower Respiratory Diseases	127924	5.28
Accidents	123706	5.10
Alzheimer's Disease	74632	3.08
Diabetes	71382	2.95
Influenza and Pneumonia	52717	2.18
Nephritis and Nephrosis	46448	1.92
Septicemia	34828	1.44
Other Causes	577181	23.81
Totals	2423712	

Total Deaths in America by Cause in 2007

Cause	%
Other Causes	23.81
Septicemia	1.44
Nephritis and Nephrosis	1.92
Influenza and Pneumonia	2.18
Diabetes	2.95
Alzheimer's Disease	3.08
Accidents	5.10
Chronic Lower Respiratory Diseases	5.28
Stroke	5.61
Cancer	23.22
Heart Disease	25.42

Causes of Deaths in US in 2007

Criteria for effectiveness

- Clearly indicates the nature of the relationship?

- Represents the quantities accurately?

- Makes it easy to compare quantities?

- Makes it easy to see ranked order of values?

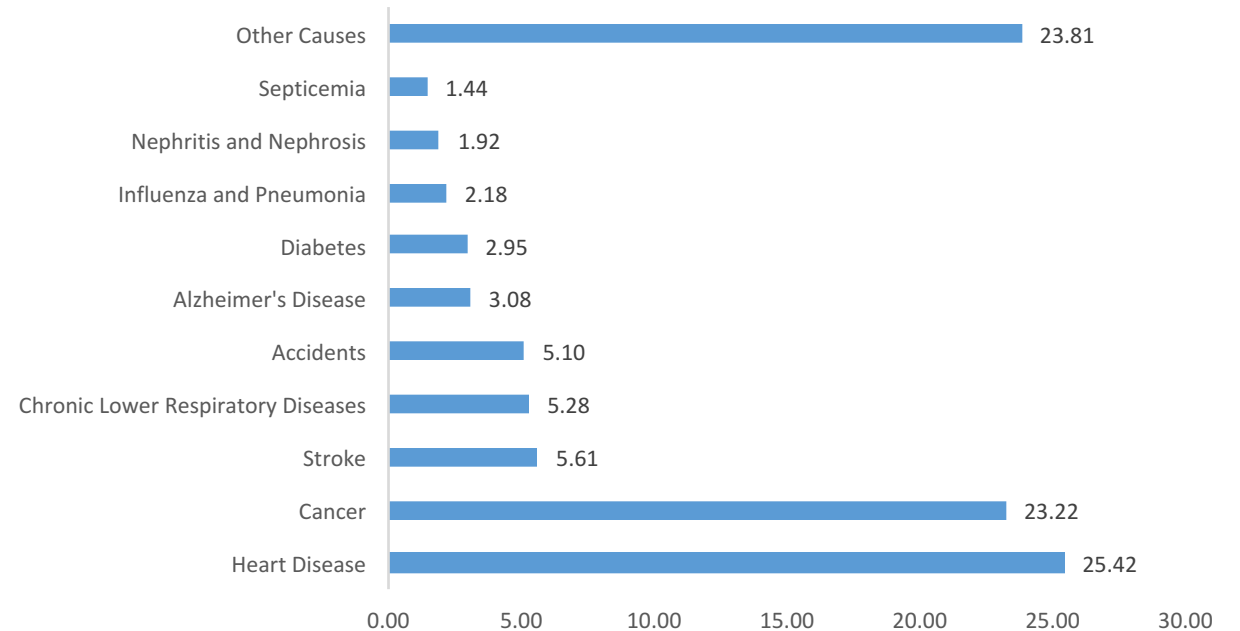- Makes obvious how people should use the information?

Total Deaths in America by Cause in 2007

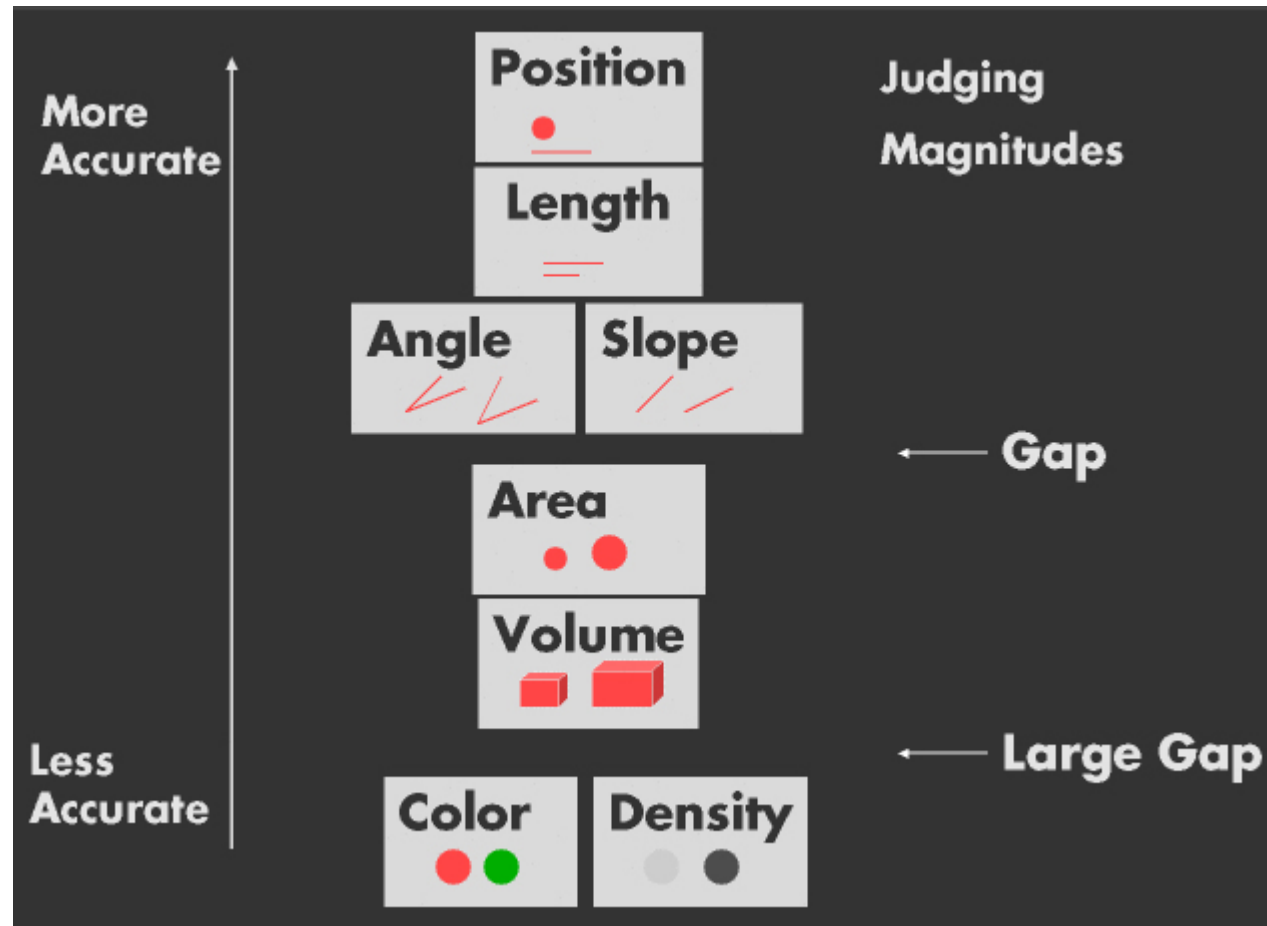| Cause	Value
Other Causes	23.81
Septicemia	1.44
Nephritis and Nephrosis	1.92
Influenza and Pneumonia	2.18
Diabetes	2.95
Alzheimer's Disease	3.08
Accidents	5.10
Chronic Lower Respiratory Diseases	5.28
Stroke	5.61
Cancer	23.22
Heart Disease	25.42

Causes of Deaths in US in 2007

Criteria for effectiveness

- Clearly indicates the nature of the relationship? Yes

- Represents the quantities accurately? Yes

- Makes it easy to compare quantities? Yes

- Makes it easy to see ranked order of values? Yes

- Makes obvious how people should use the information? Yes

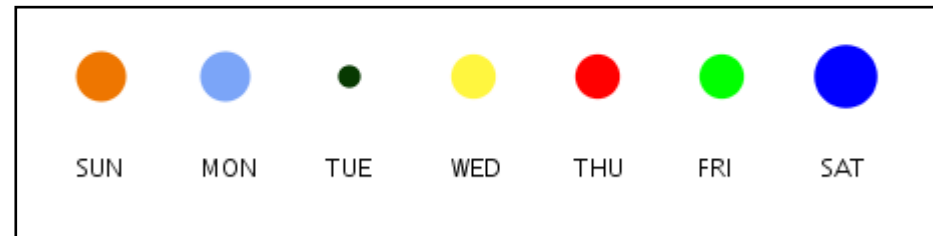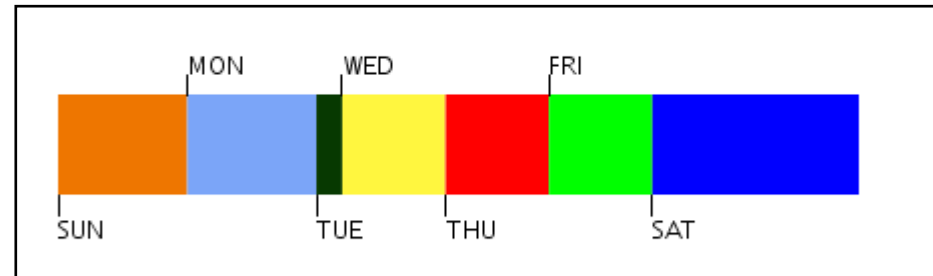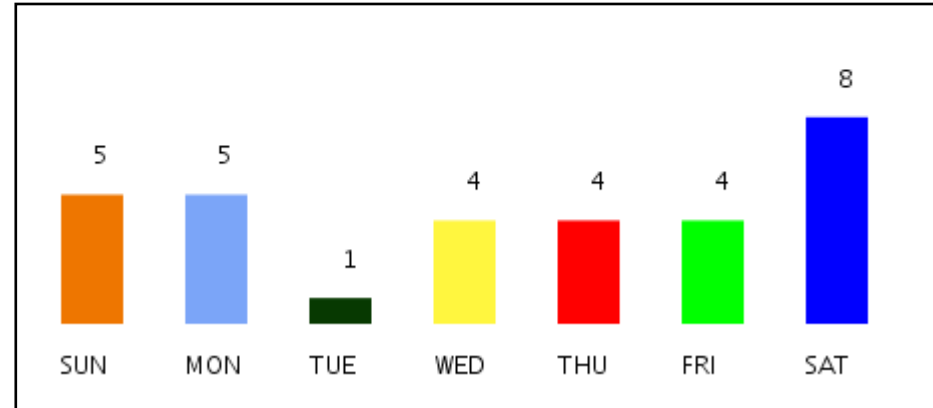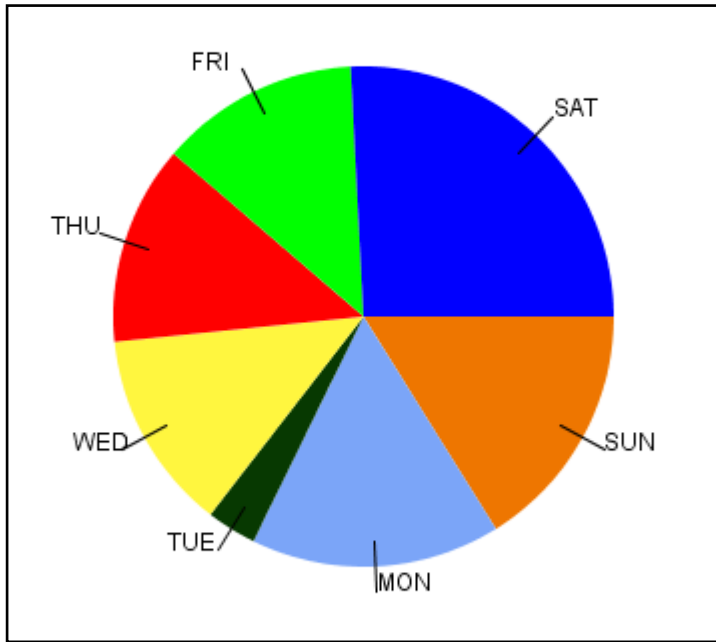**Total Deaths in America by Cause in 2007**

Cause	Value
Other Causes	23.81
Septicemia	1.44
Nephritis and Nephrosis	1.92
Influenza and Pneumonia	2.18
Diabetes	2.95
Alzheimer's Disease	3.08
Accidents	5.10
Chronic Lower Respiratory Diseases	5.28
Stroke	5.61
Cancer	23.22
Heart Disease	25.42

Accuracy Ranking of Basic Charts

From: Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods
William S. Cleveland; Robert McGill, Journal of the American Statistical Association, Vol. 79, No. 387. (Sep., 1984), pp. 531-554.

% Births by Day of Week

Top Medals in Olympics by Country (1992-2016)

Country	1992	1996	2000	2004	2008	2012	2016
United States of America	108	01	92	103	110	104	121
People's Republic of China	54	50	59	63	100	88	70
Russian Federation	112	63	88	92	72	82	56
Great Britain	20	15	28	30	47	65	67
Australia	27	41	58	49	46	35	29
Germany	82	65	56	49	41	44	42
France	29	37	38	33	40	34	42
Republic of Korea	29	27	28	30	31	28	21
Japan	22	14	18	37	25	38	41
Italy	19	35	34	32	27	28	28

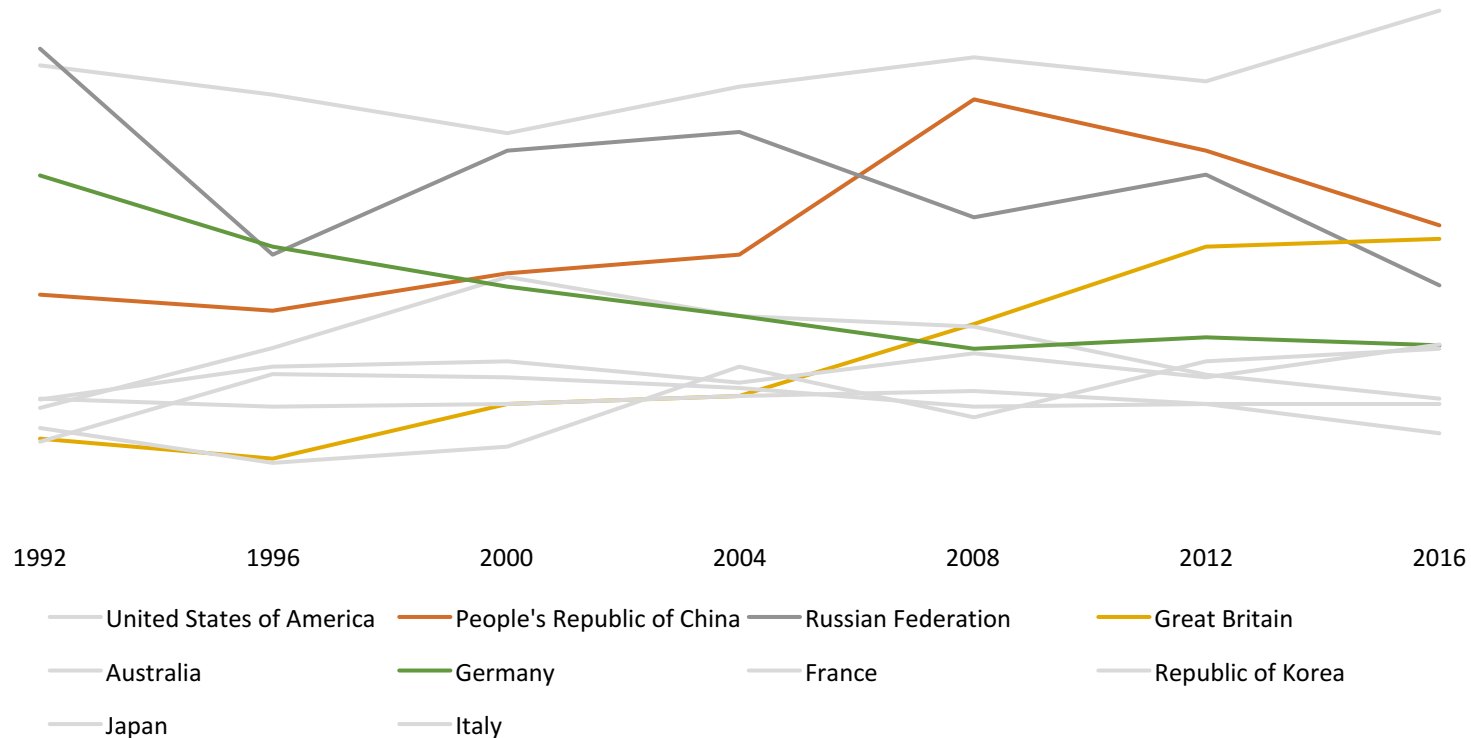# Top Medals in Olympics by Country (1992-2016)

Olympics Medals Tally

Focus on
- Russia
- China
- Great Britain
- Germany

| 1992 | 1996 | 2000 | 2004 | 2008 | 2012 | 2016 |

— United States of America   — People's Republic of China   — Russian Federation   — Great Britain

— Australia   — Germany   — France   — Republic of Korea

— Japan   — Italy

Top Medals in Olympics by Country (1992-2016)

Country	1992	2016
United States of America	108	121
People's Republic of China	54	70
Russian Federation	112	56
Great Britain	20	67
Australia	27	29
Germany	82	42
France	29	42
Republic of Korea	29	21
Japan	22	41
Italy	19	28
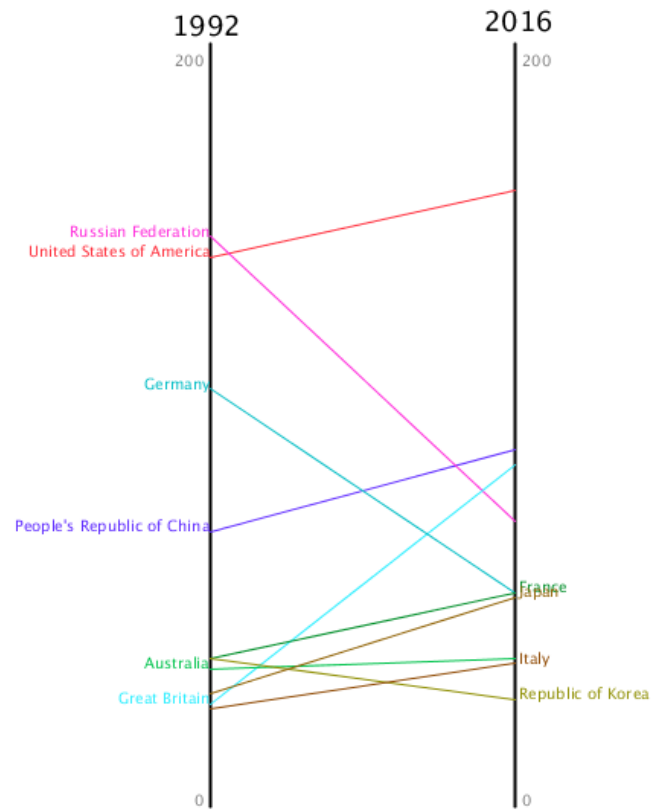
Mapping Numbers
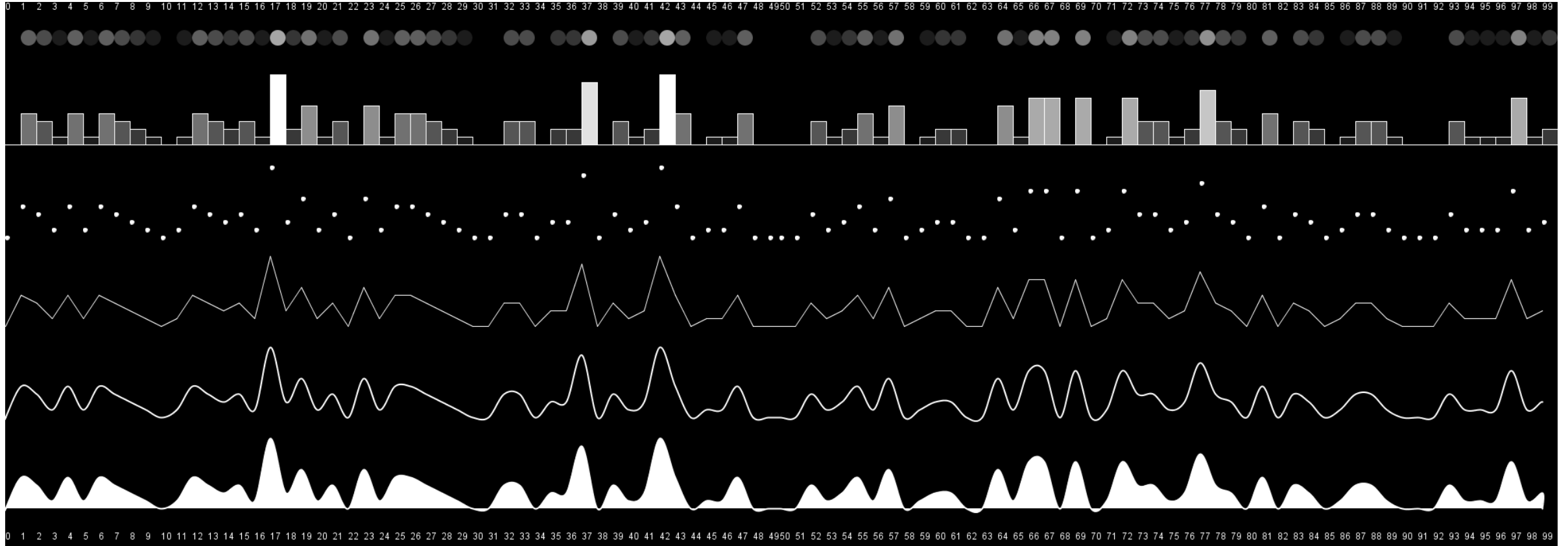
Mapping Numbers: Heat Map

Permeating Data Visualization in CS Courses
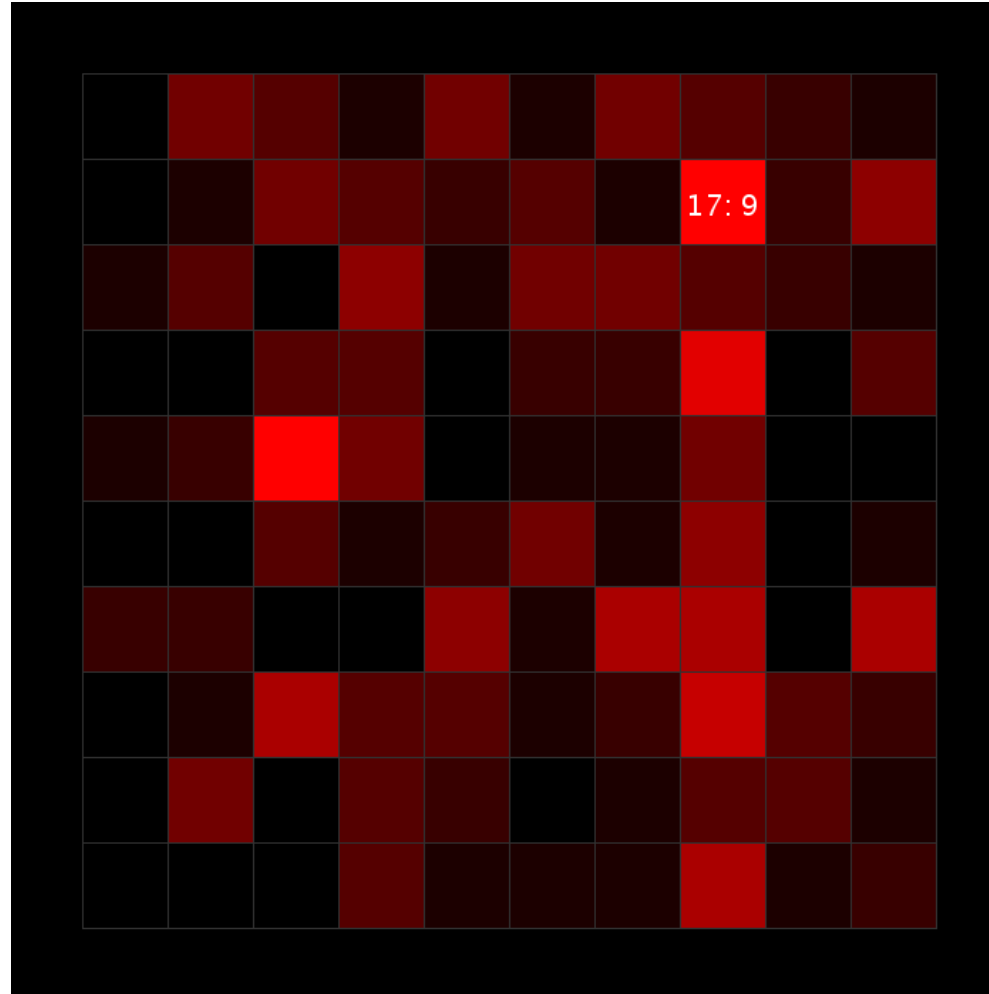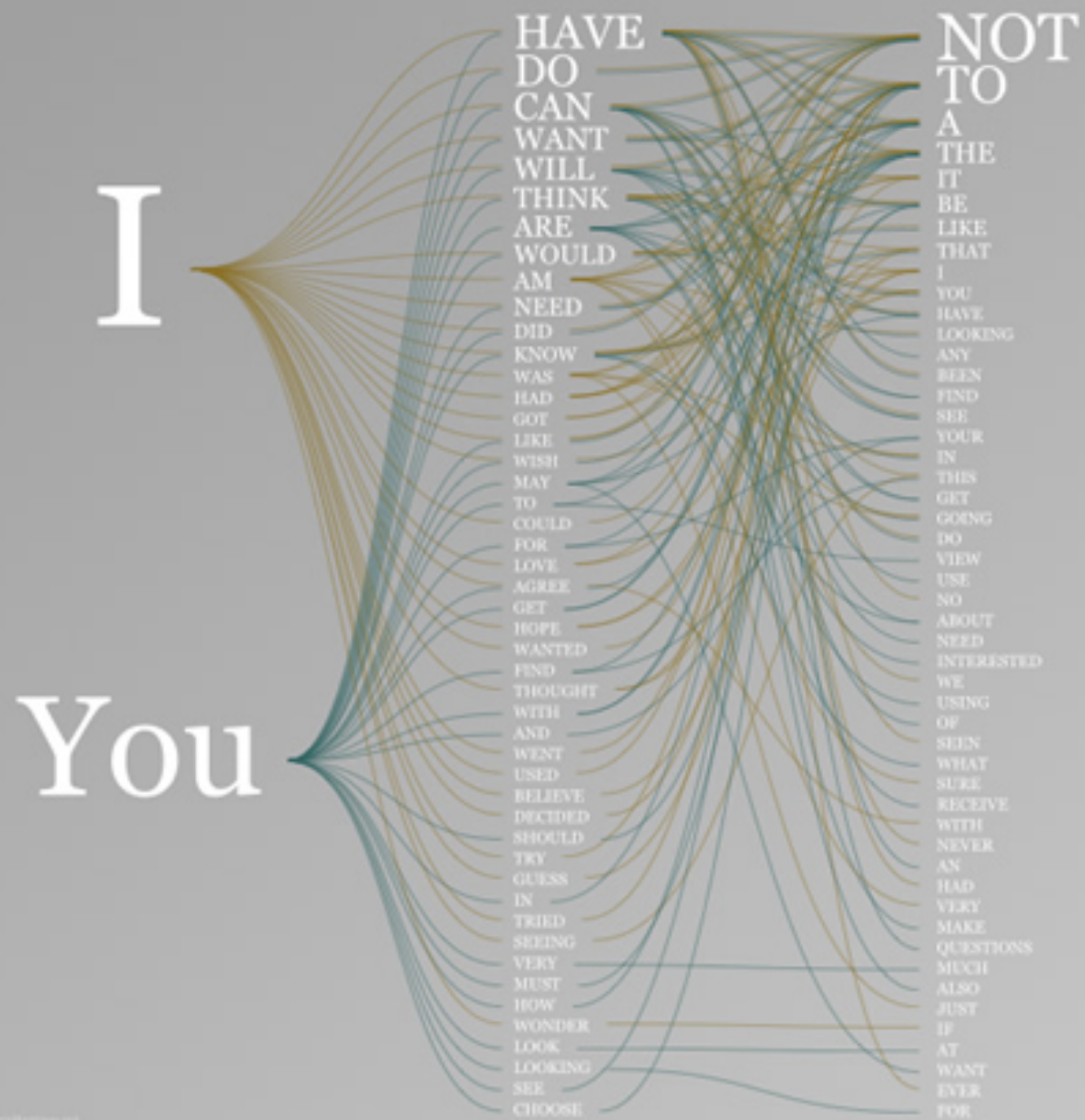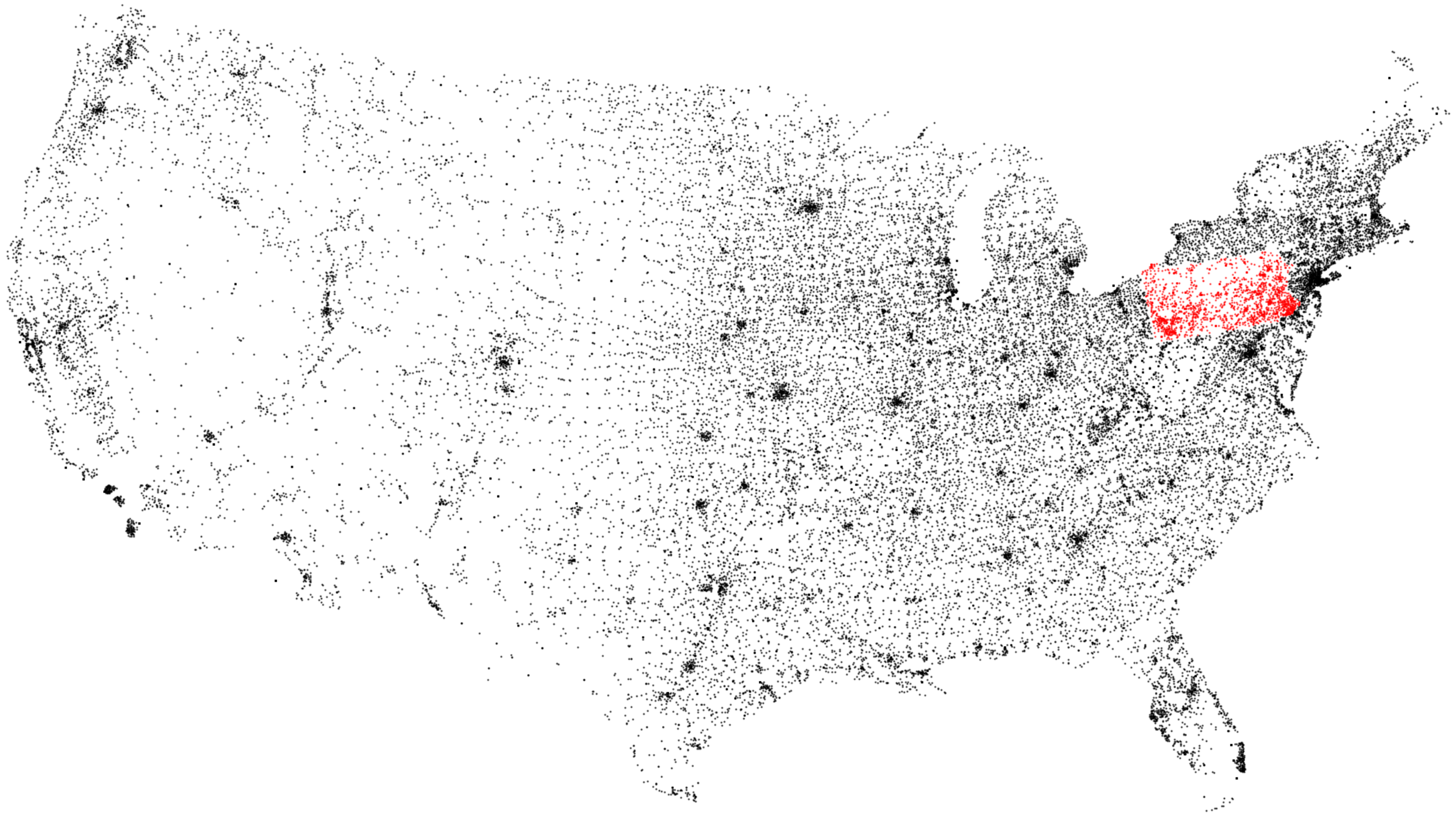
Proportional Symbols
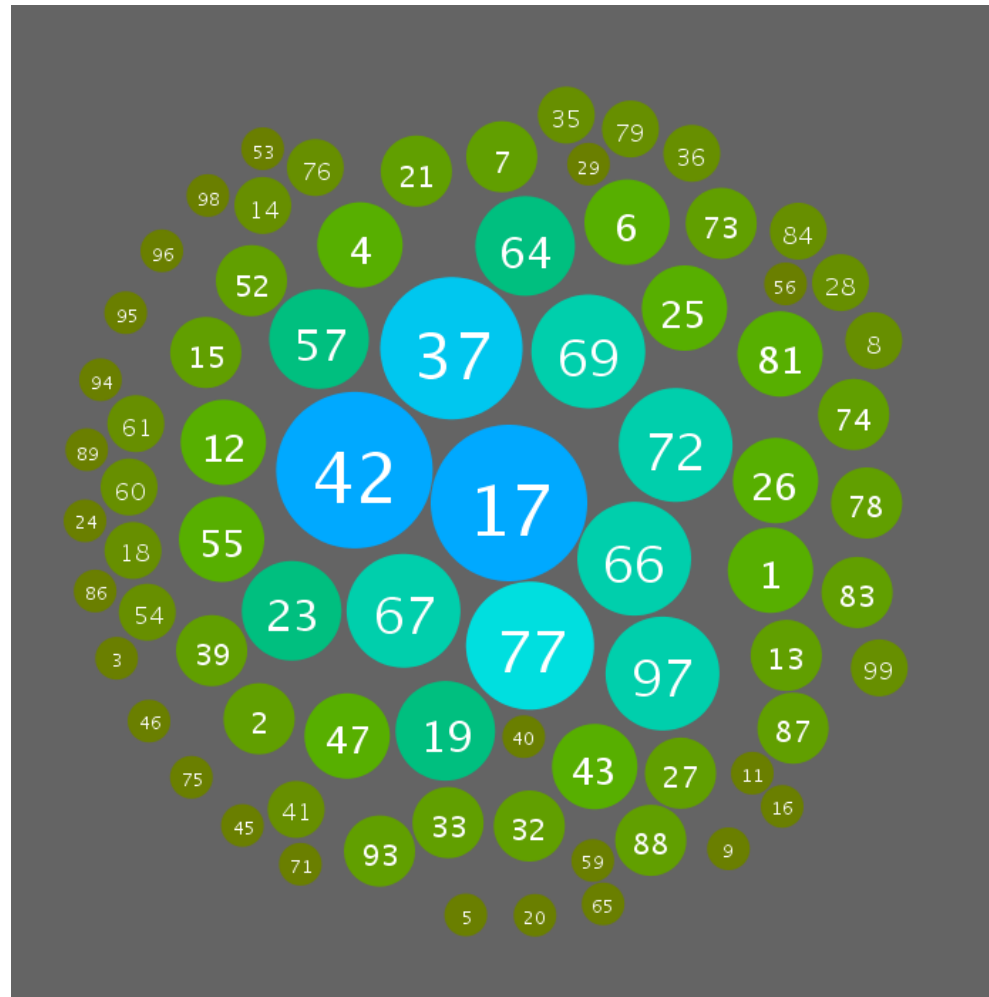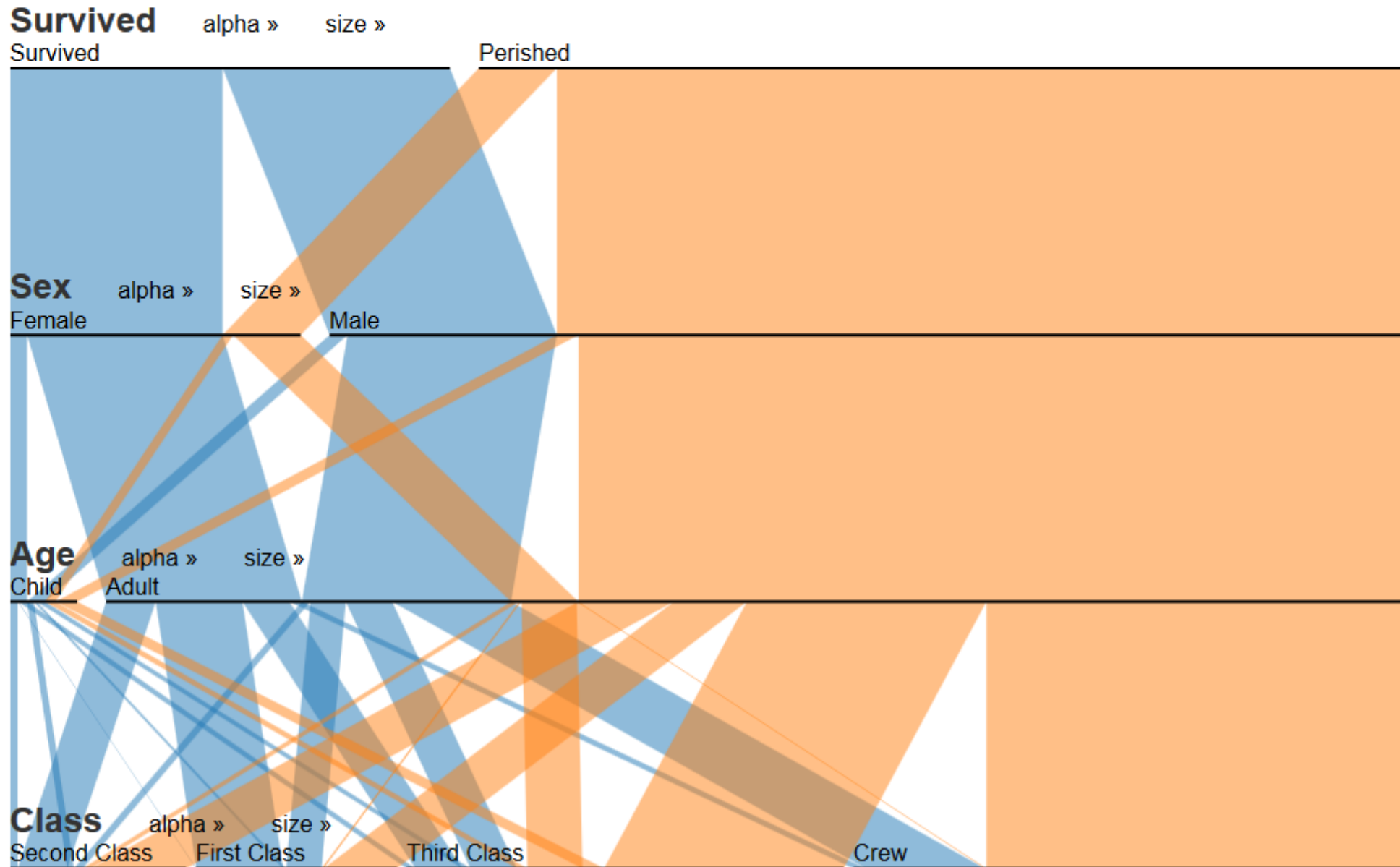
UZ Zip Codes + Lat & Long Data

```
…
19010  0.27167493    0.69789624  Bryn Mawr, PA
19012  0.27438772    0.70036817  Cheltenham, PA
19013  0.27077052    0.6963129   Chester, PA
19014  0.2713161     0.6962156   Aston, PA
19015  0.27156553    0.6970415   Brookhaven, PA
…
```

Titanic Survivors

Survived alpha » size »

Survived   Perished

Sex alpha » size »

Female   Male

Age alpha » size »

Child   Adult

Class alpha » size »

Second Class   First Class   Third Class   Crew

☐ Curves?

Titanic Survivors

Survived   alpha »   size »

Survived   Survived   Perished
            711 (32%)

Sex   alpha »   size »

Female   Male

Age   alpha »   size »

Child   Adult

Class   alpha »   size »

Second Class   First Class   Third Class   Crew

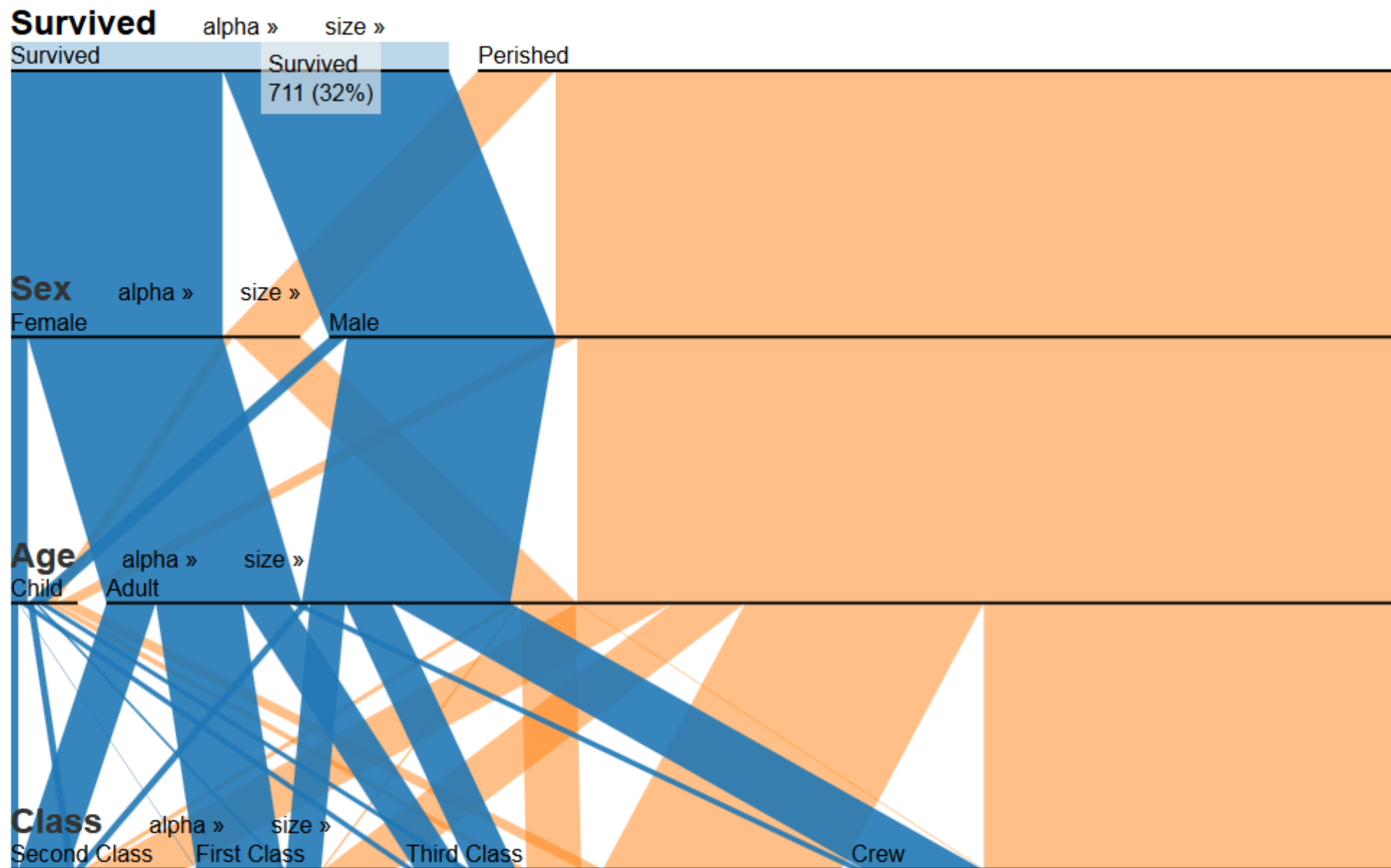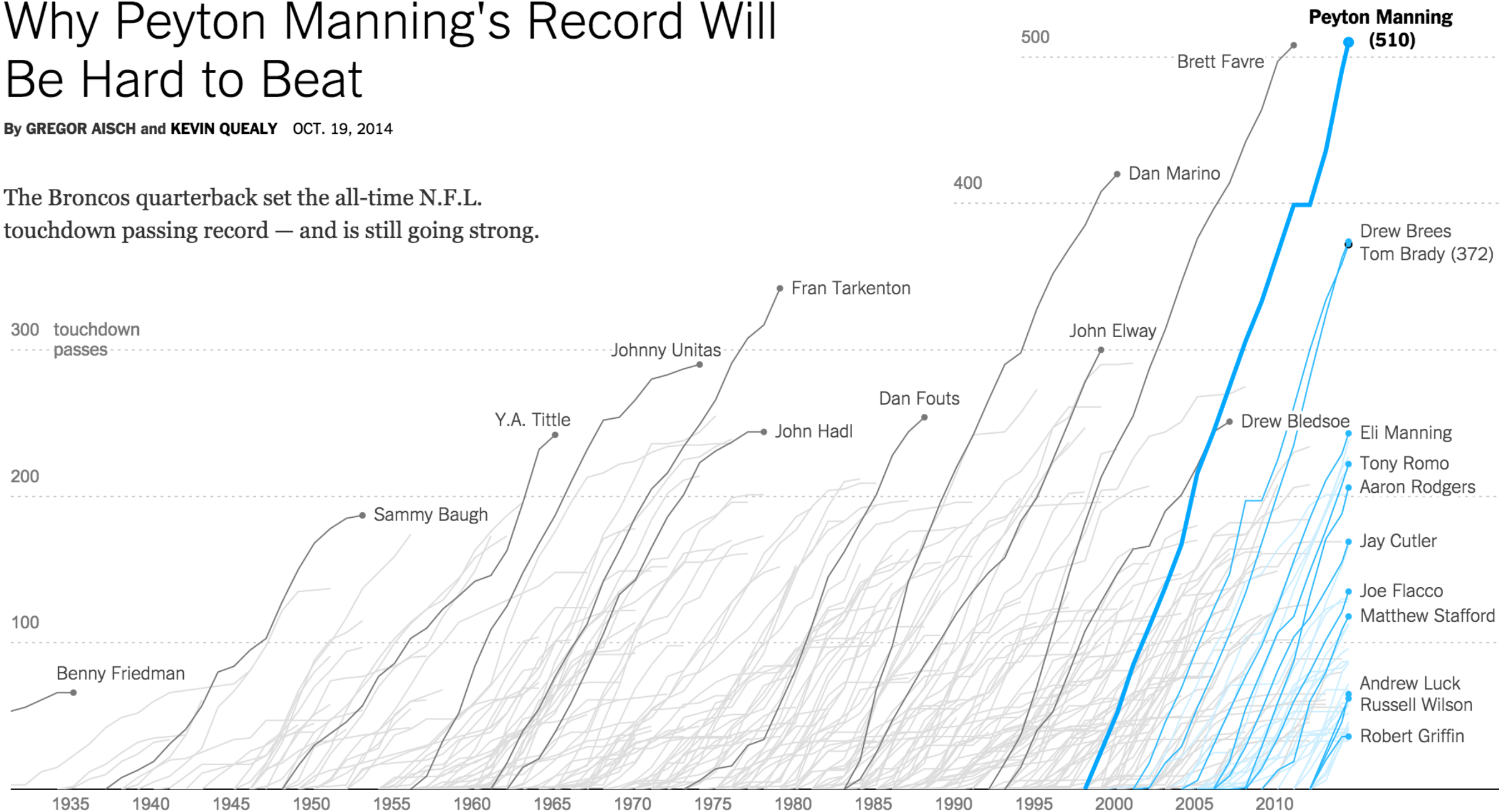☐ Curves?

Why Peyton Manning's Record Will Be Hard to Beat

By GREGOR AISCH and KEVIN QUEALY OCT. 19, 2014

The Broncos quarterback set the all-time N.F.L. touchdown passing record — and is still going strong.

Peyton Manning (510)

Brett Favre

Dan Marino

Drew Brees
Tom Brady (372)

Fran Tarkenton

Johnny Unitas

John Elway

300 touchdown passes

Dan Fouts

Y.A. Tittle

John Hadl

Drew Bledsoe

Eli Manning

Tony Romo
Aaron Rodgers

200

Jay Cutler

Sammy Baugh

Joe Flacco
Matthew Stafford

100

Andrew Luck
Russell Wilson

Benny Friedman

Robert Griffin

500

400

1935 1940 1945 1950 1955 1960 1965 1970 1975 1980 1985 1990 1995 2000 2005 2010

Data Visualization - Horizons