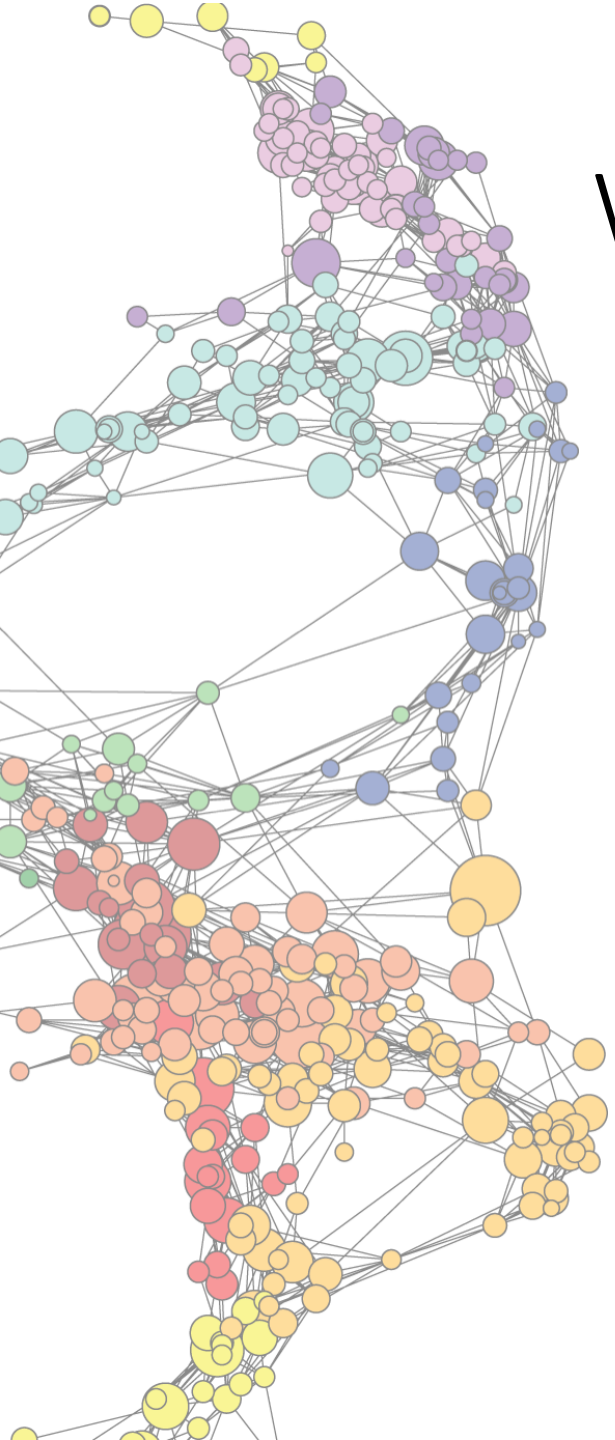
Tips for Effective Data Visualization

Angela Zoss · Eric Monson
Data and Visualization Services

STA 112FS · Fall 2017

Slides: http://bit.ly/STA112FSVisFall2017



What is data visualization?

Anything that converts data sources into a visual representation

charts, graphs, maps, even just tables

http://guides.library.duke.edu/datavis

Why do we visualize?

1		2		3		4	
x	y	x	y	x	y	x	y
10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89

Almost identical summary statistics:

x & y mean

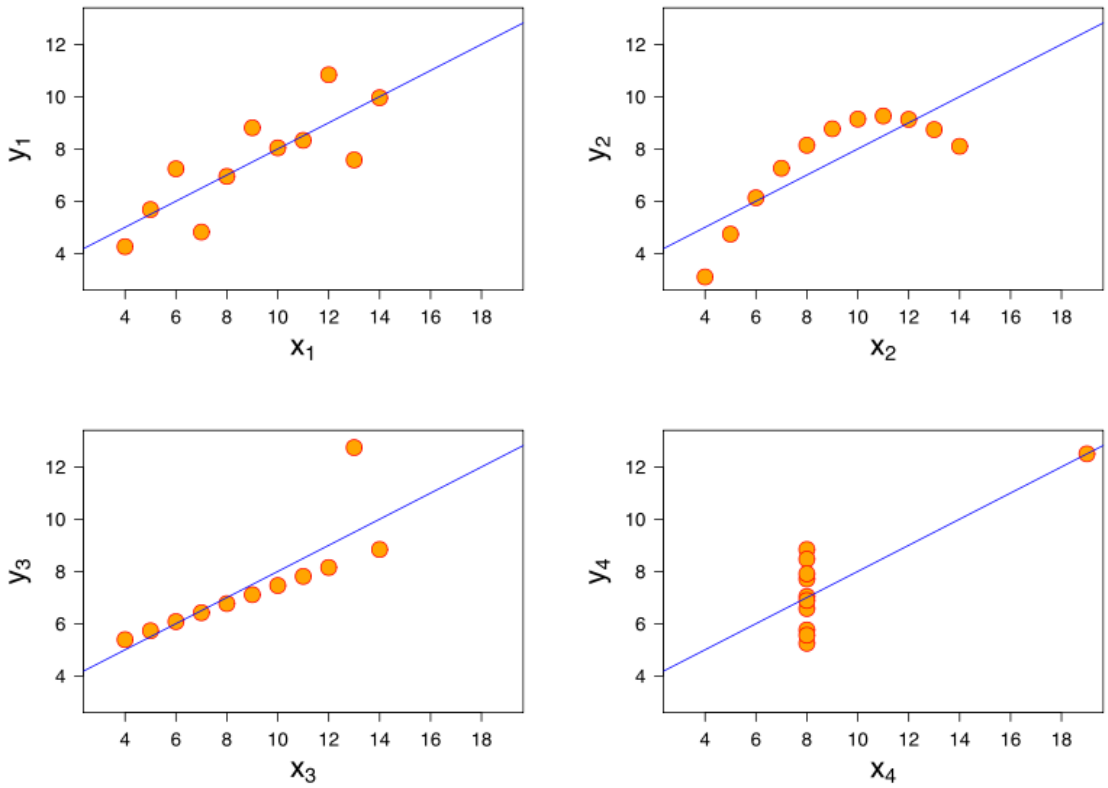
x & y variance

x-y correlation

x-y linear regression

https://en.wikipedia.org/wiki/Anscombe%27s_quartet

We visualize to see patterns

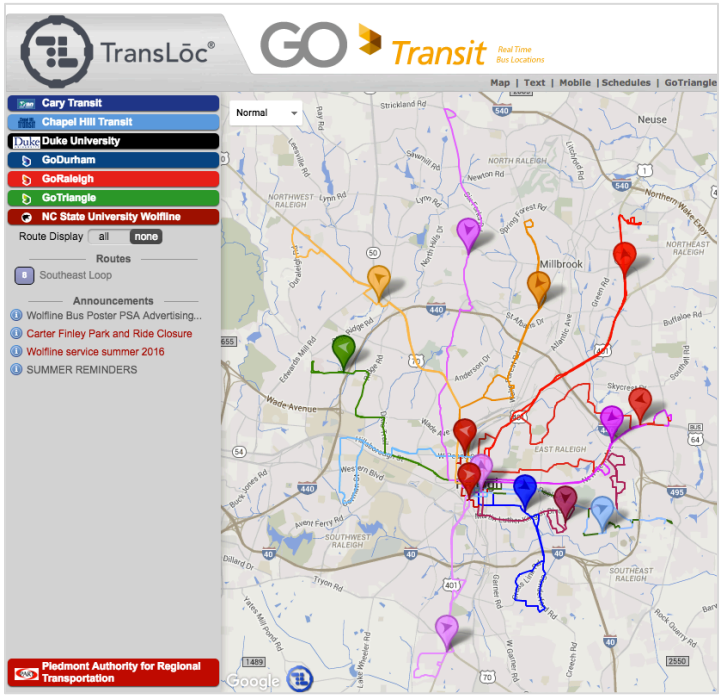


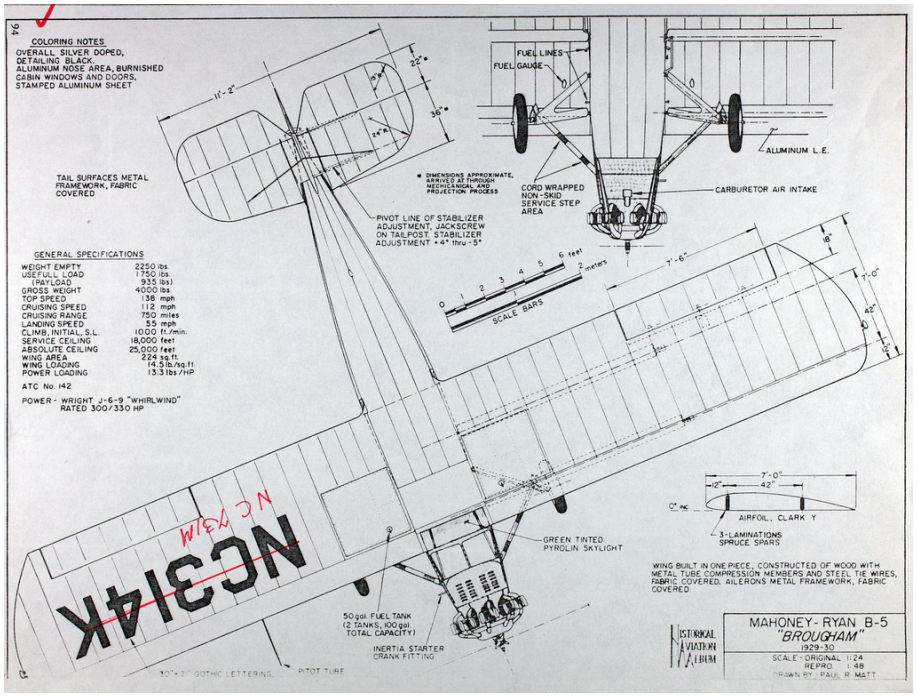
Anscombe's Quartet

http://en.wikipedia.org/wiki/Anscombe%27s_quartet

Visualization Intention

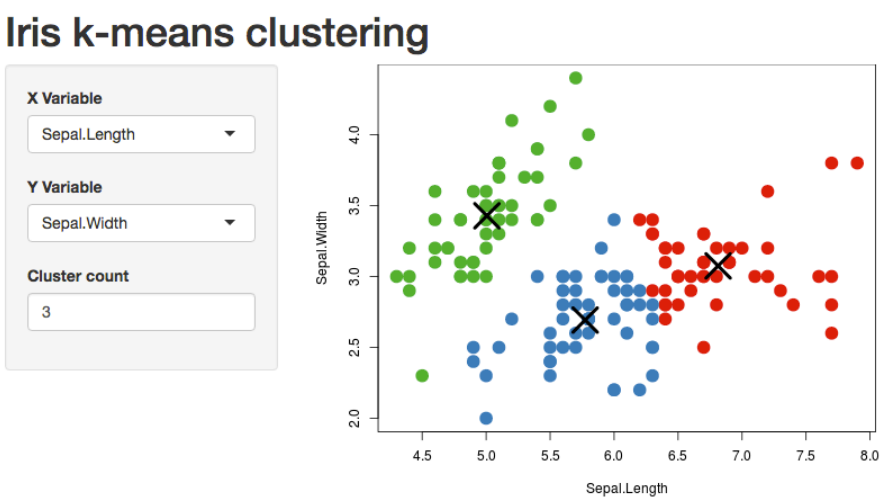
Reference/Look-up

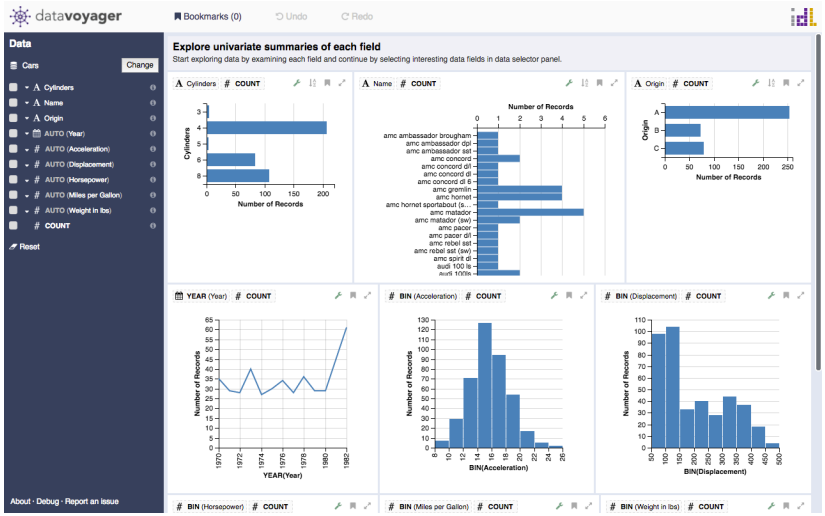




https://triangle.transloc.com/

Exploration/Analysis

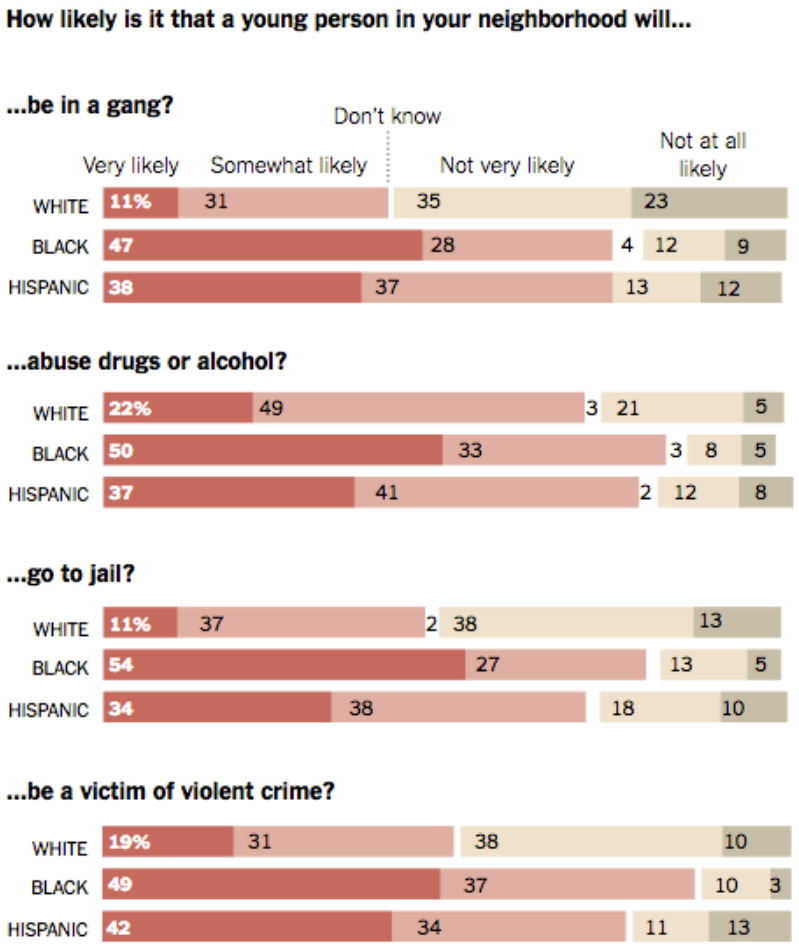


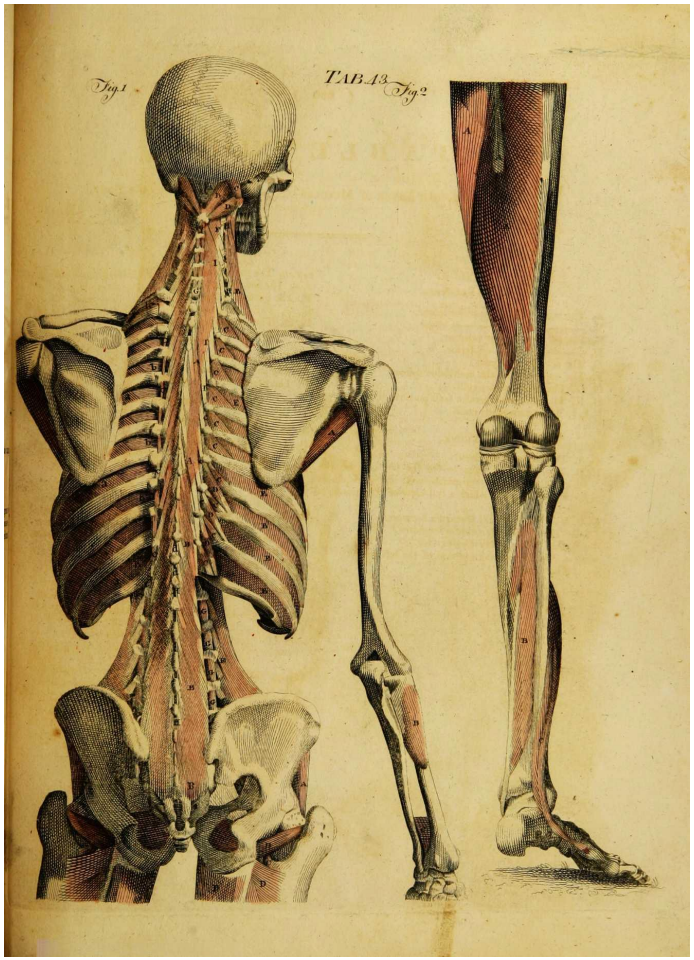


http://shiny.rstudio.com/gallery

https://vega.github.io/voyager

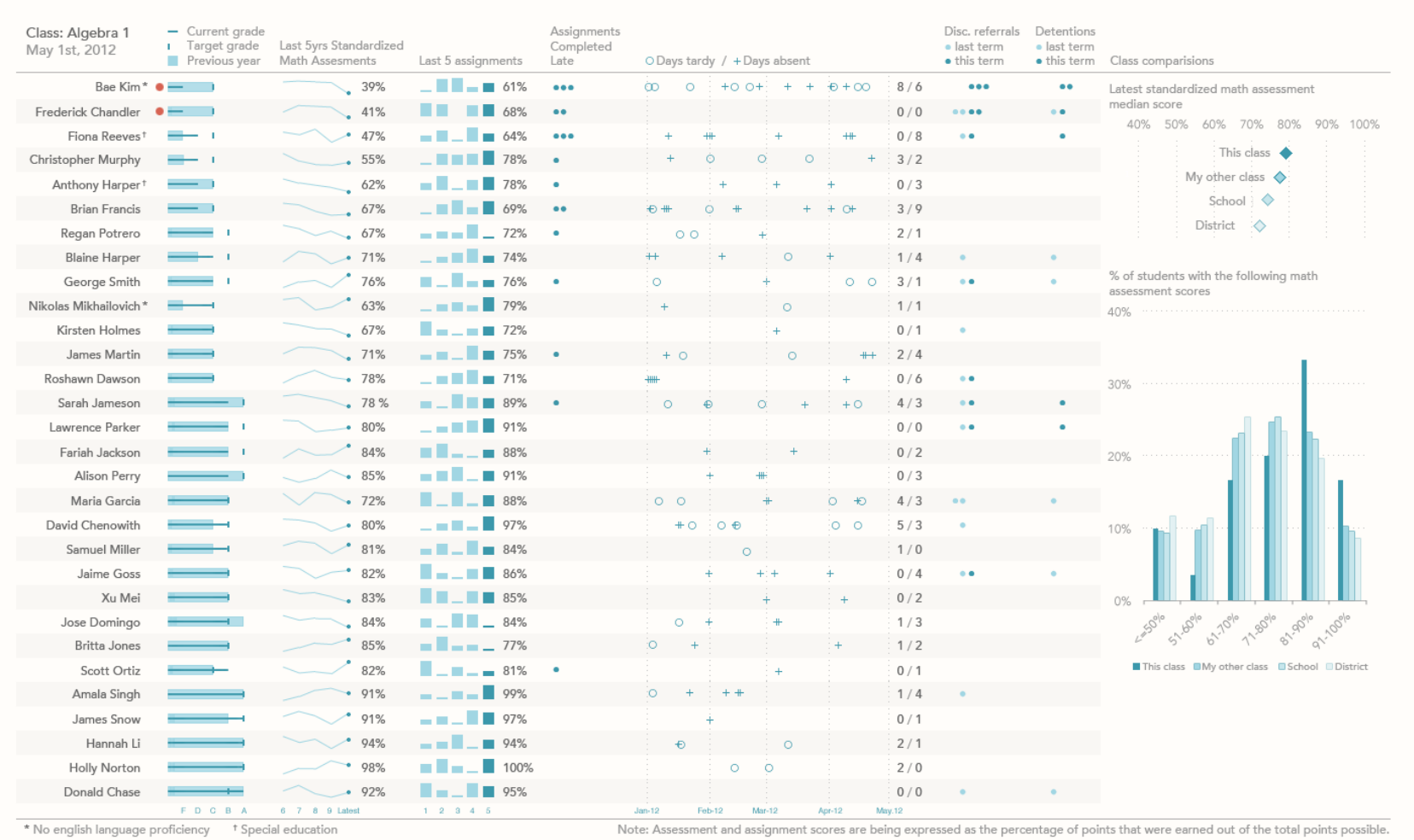
Presentation/Communication





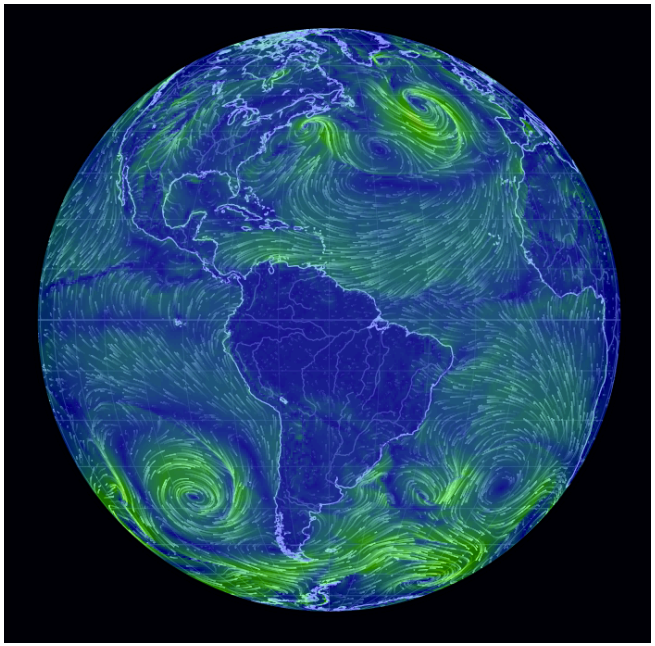
https://t.co/03h2Ipppni

Monitoring

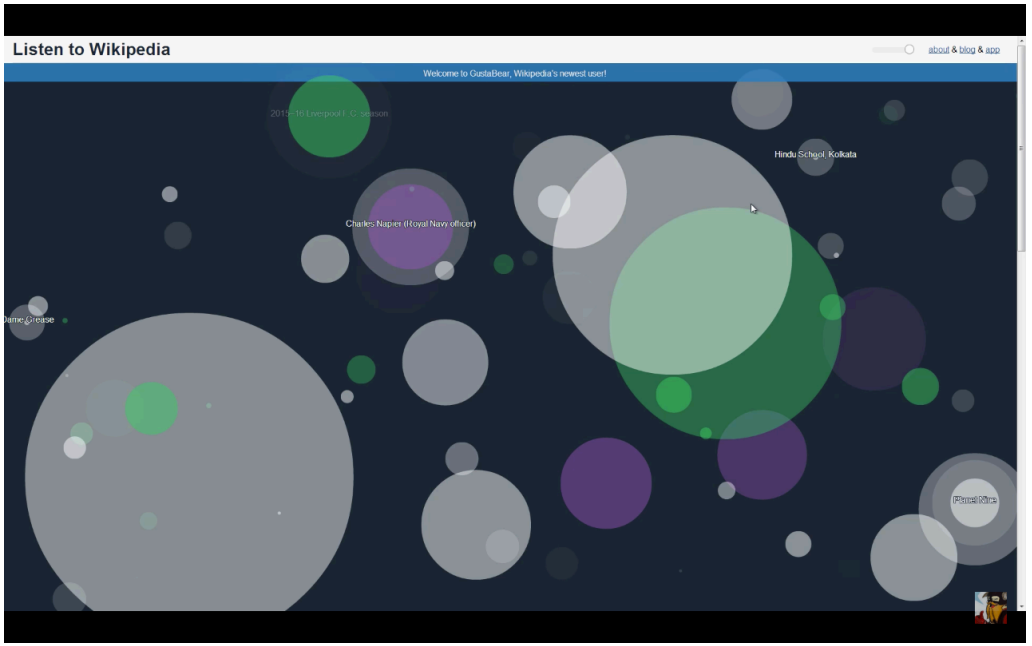


https://www.perceptualedge.com/blog/?p=1374

Enjoyment/Aesthetics

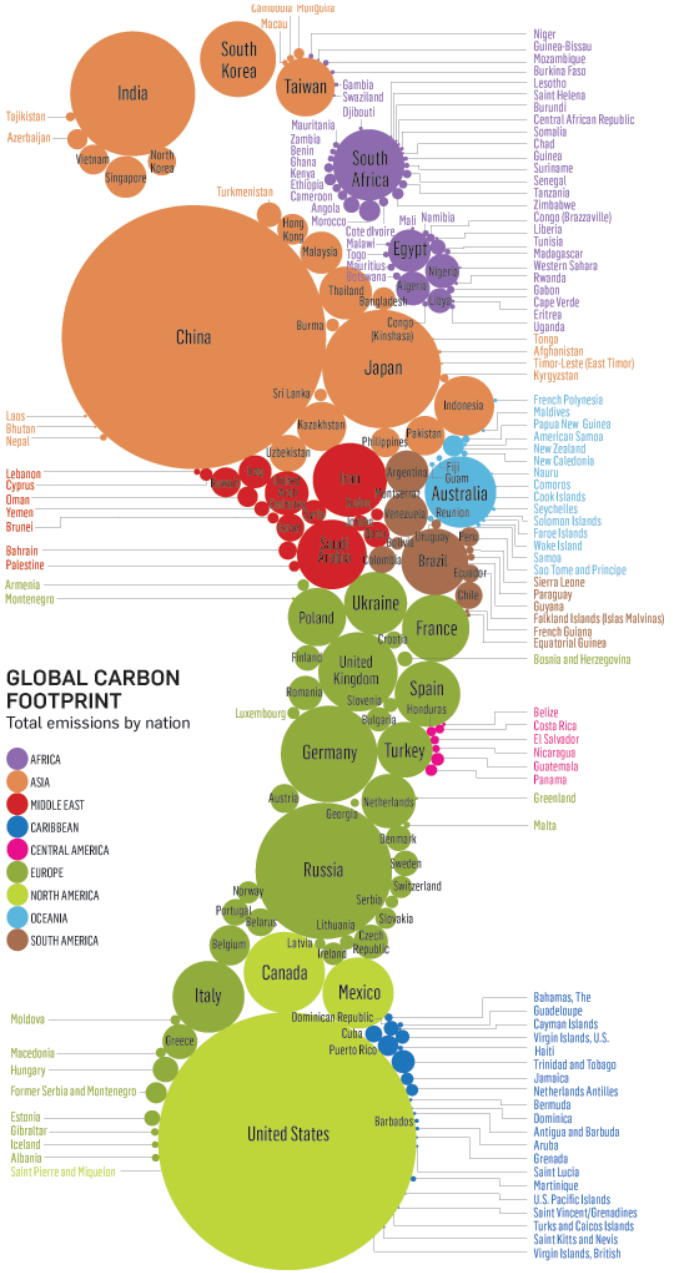


https://earth.nullschool.net

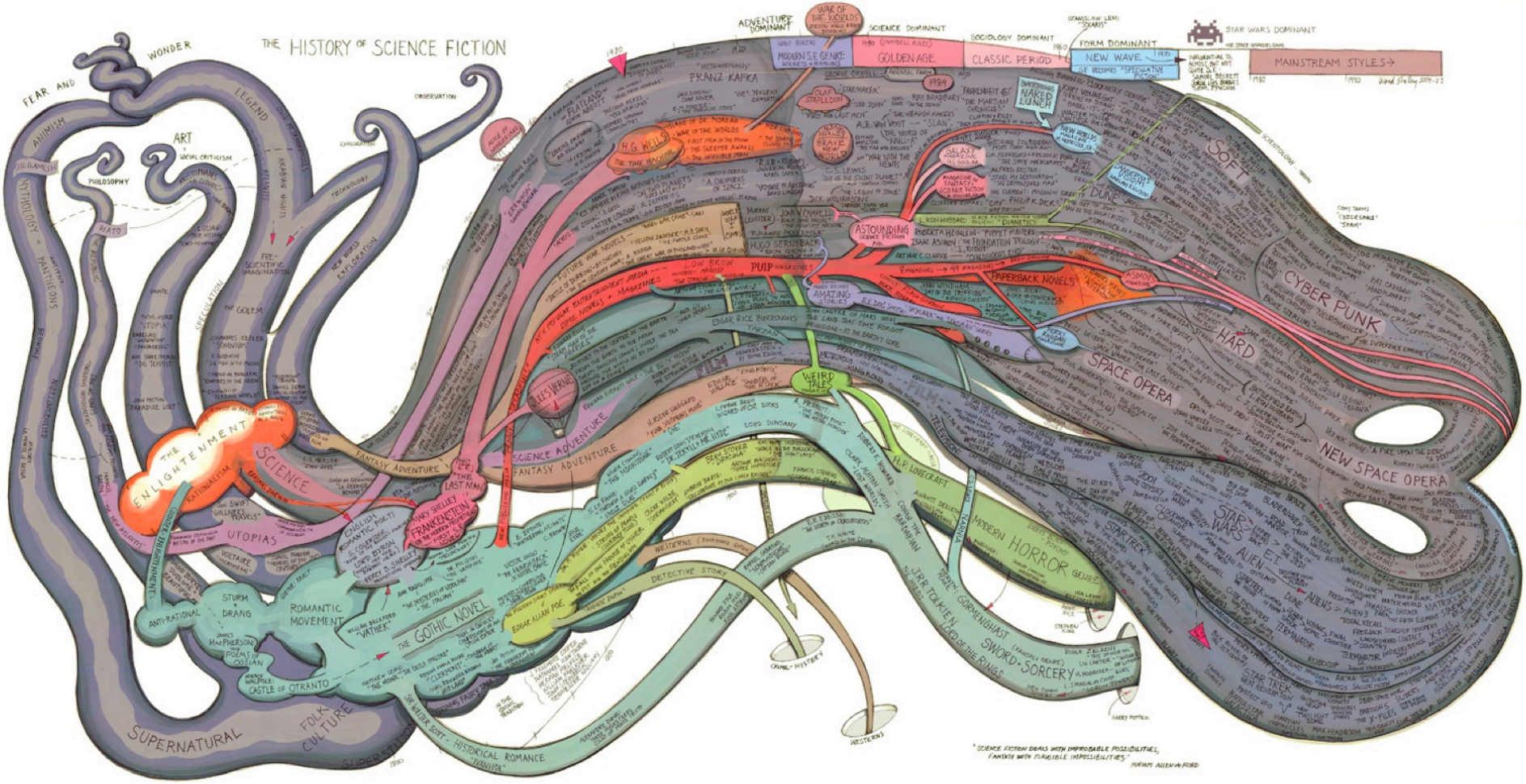


https://youtu.be/F-1CwdqGGPM

Engagement

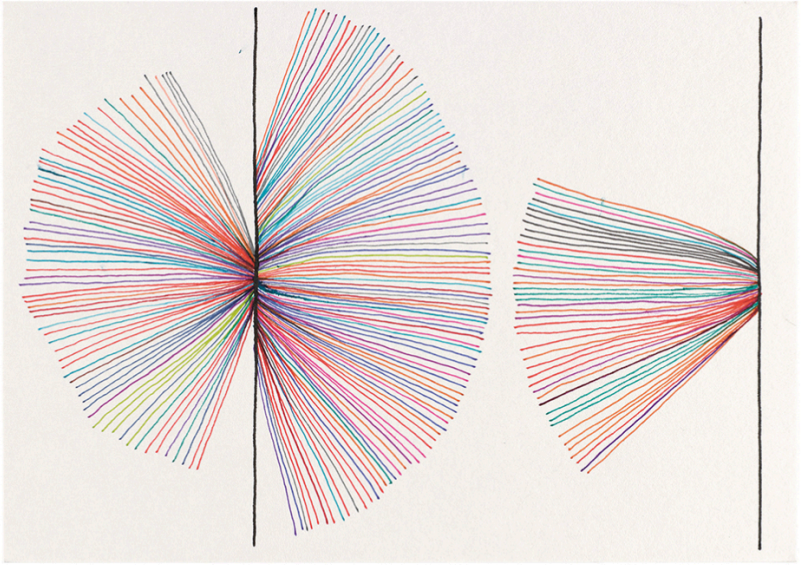


http://www.stanfordkaystudio.com/information.html

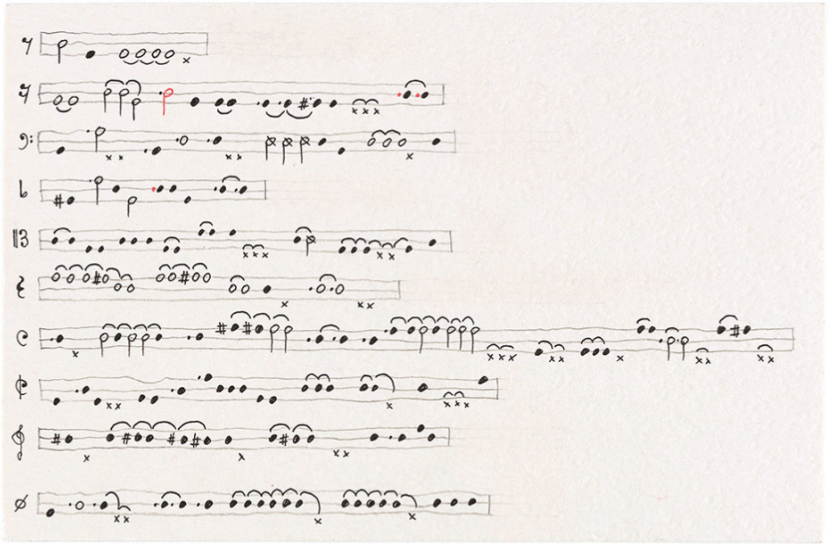


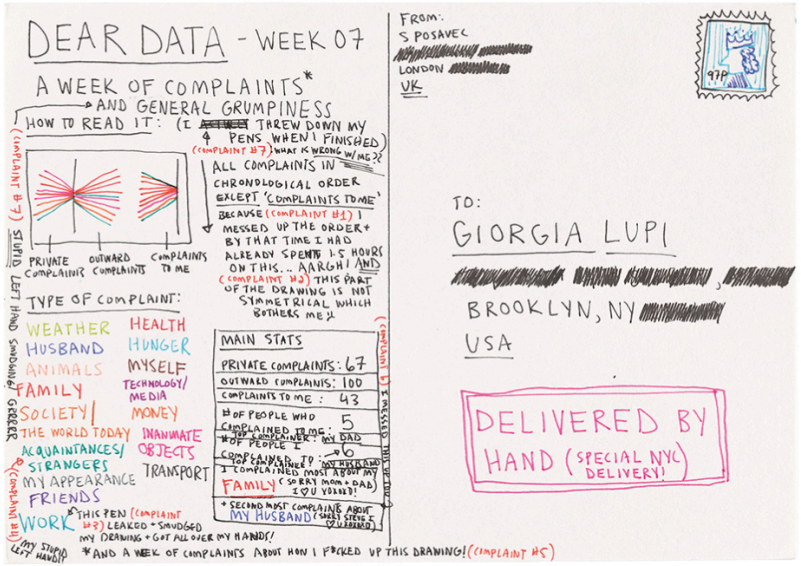
History of Science Fiction (2011), by Ward Shelley
http://www.scimaps.org/maps/map/history_of_science_f_132/

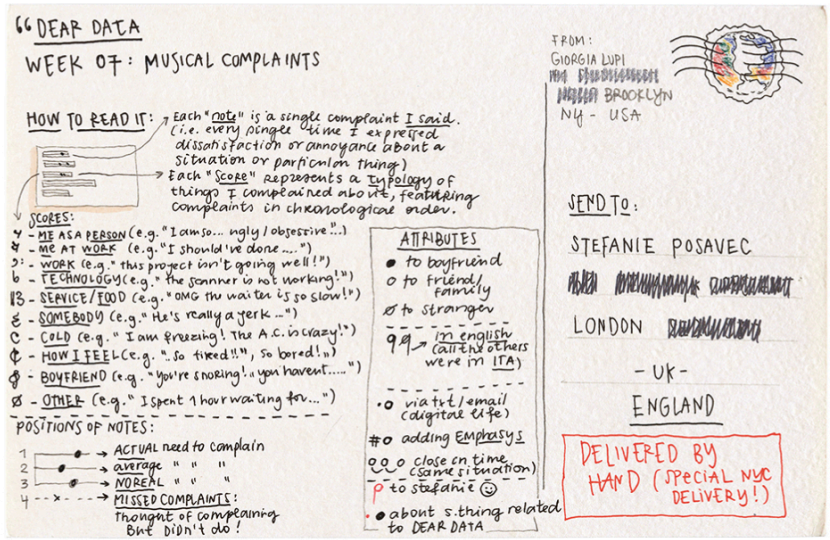
a week of complaints — Stefanie



GIORGIA — week seven



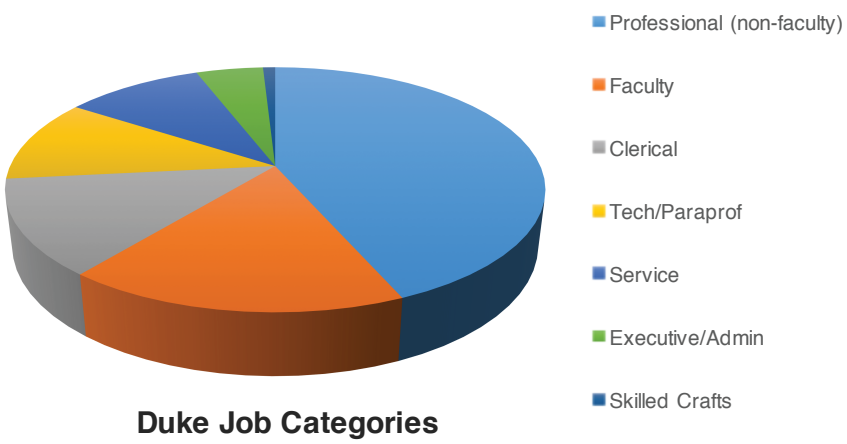


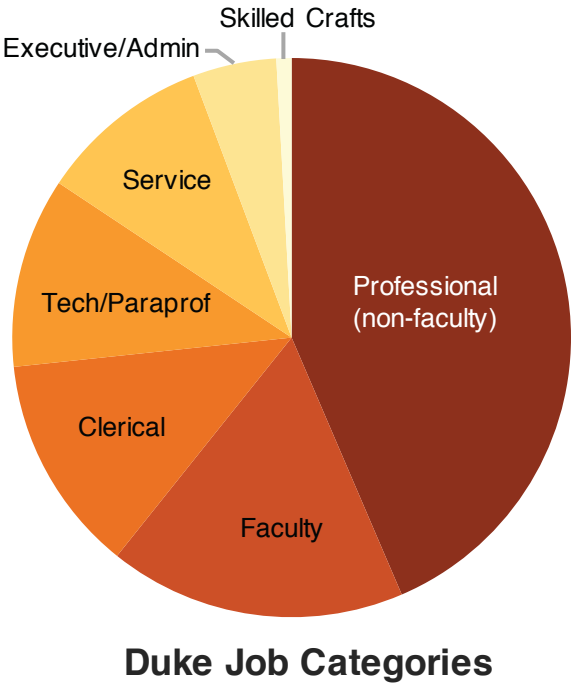


http://www.dear-data.com/theproject

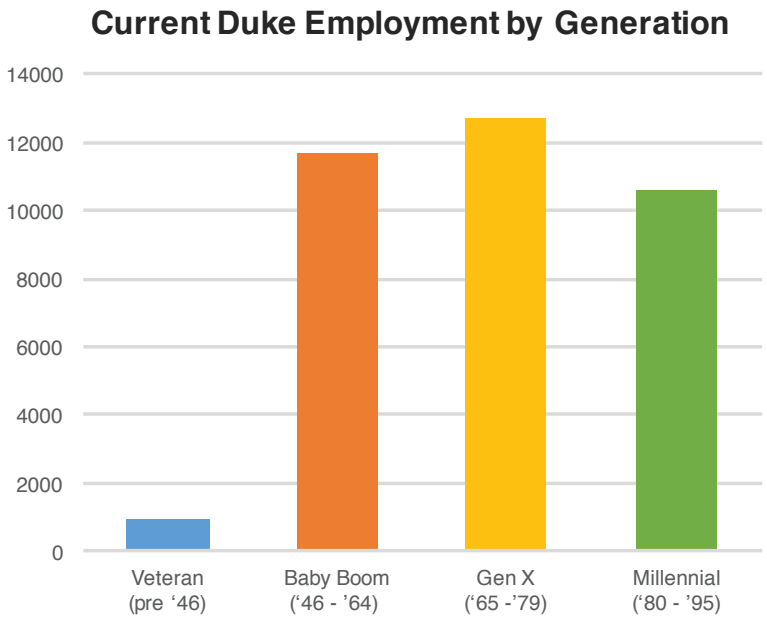
Designing effective visualizations

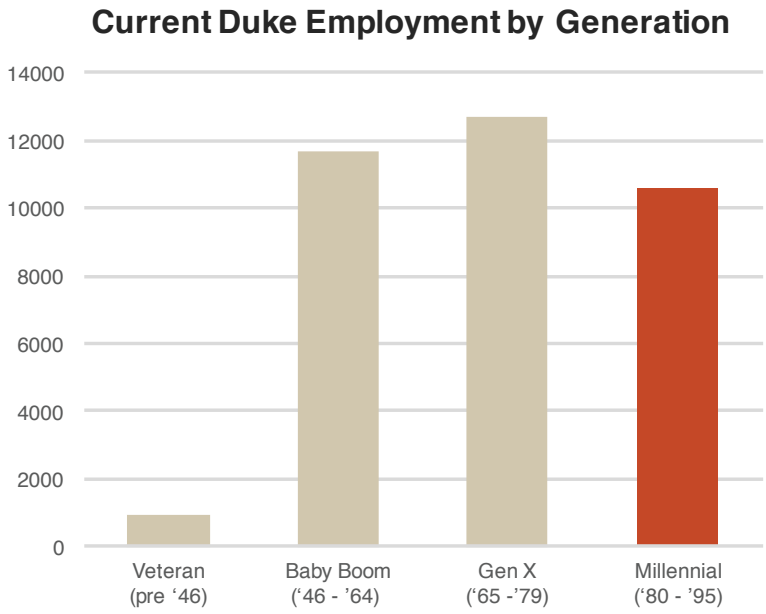
Keep it simple



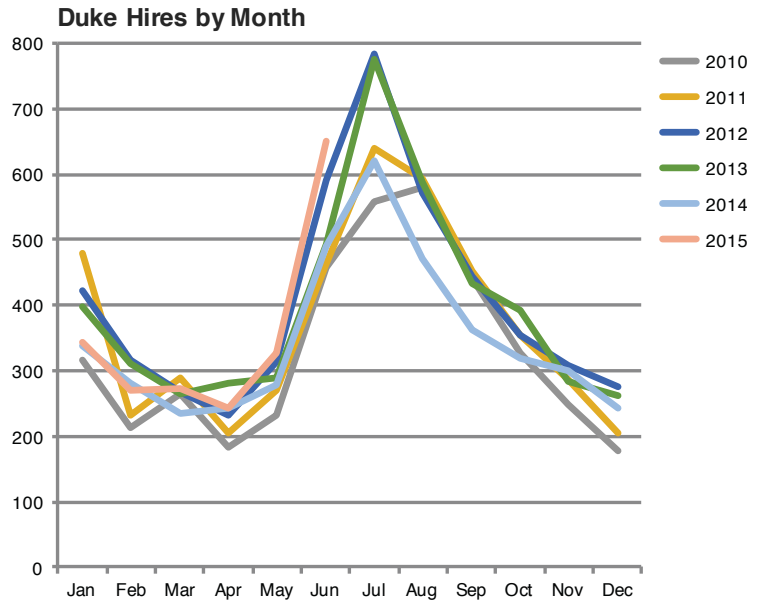


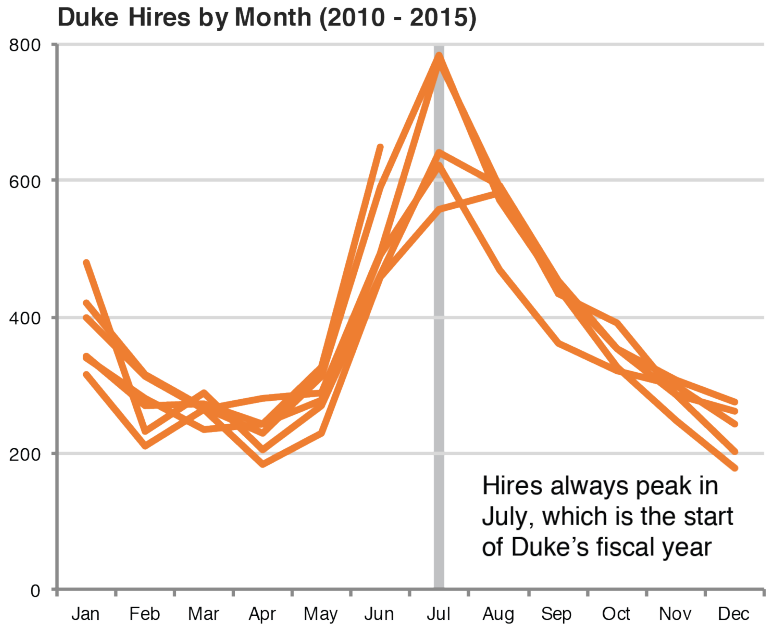
Use color to draw attention

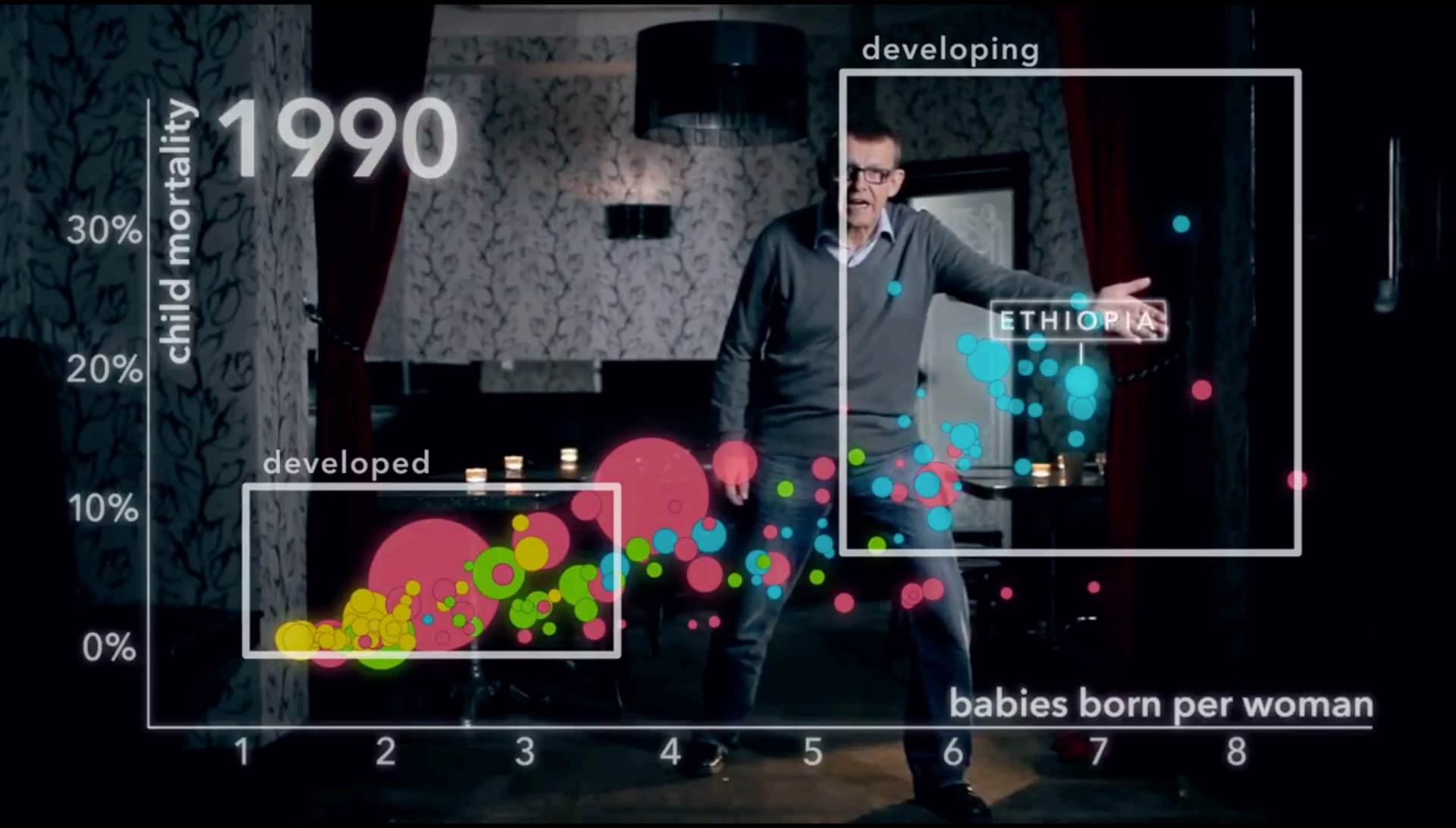




Tell a story


Duke Hires by Month
800
700
600
500
400
300
200
100
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
2010
2011
2012
2013
2014
2015

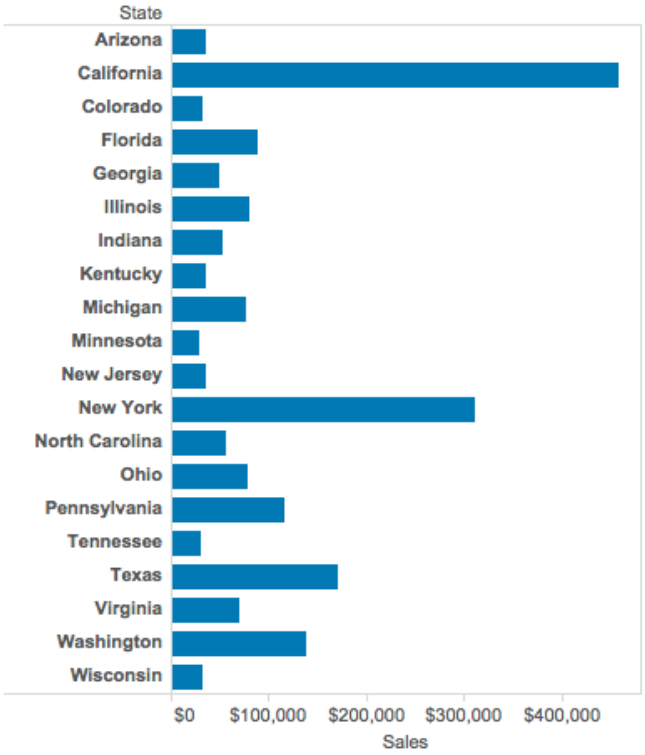

Duke Hires by Month (2010 - 2015)
800
600
400
200
0
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Hires always peak in July, which is the start of Duke's fiscal year



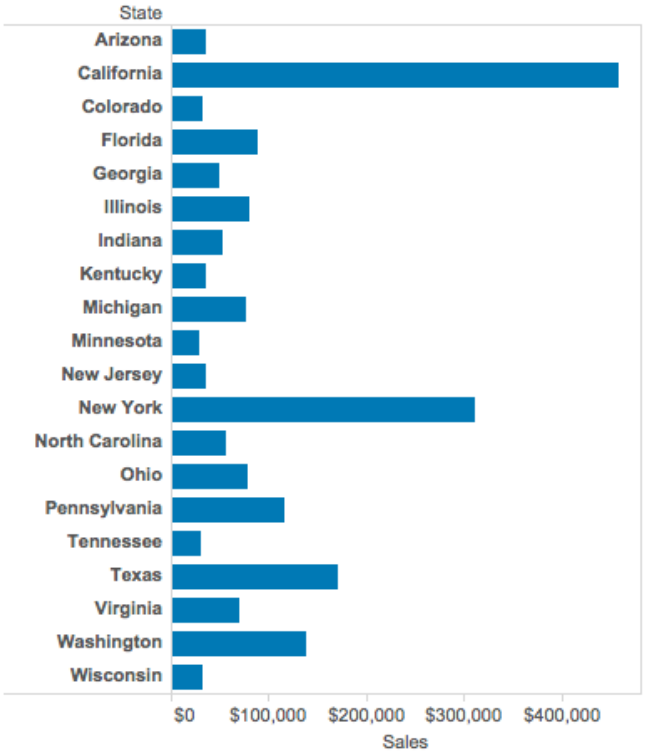
http://www.youtube.com/watch?v=OwII-dwh-bk

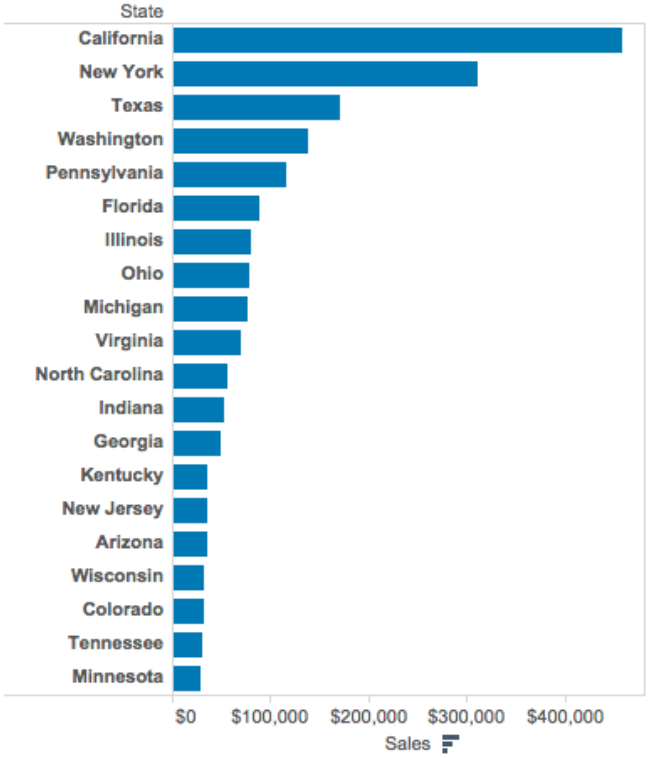
Common missteps

Default ordering hides patterns

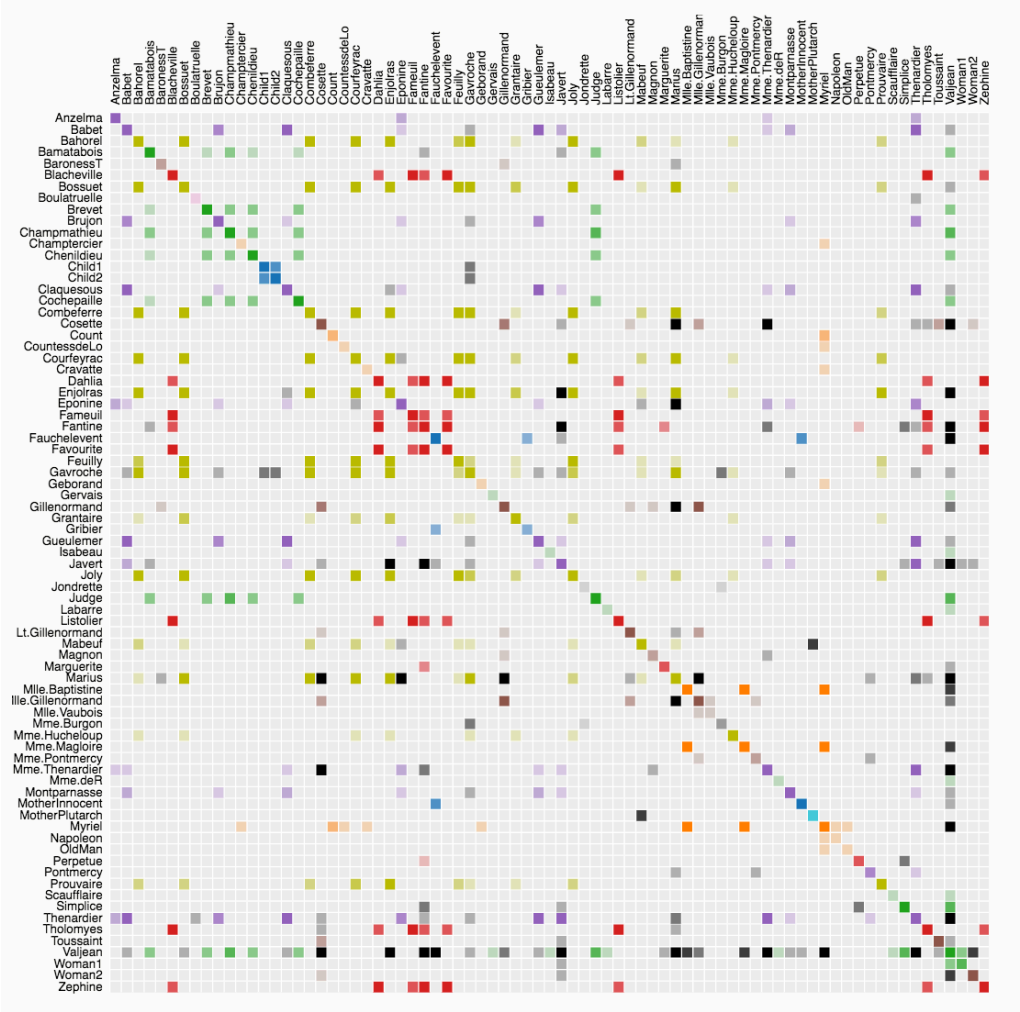


Sorting reveals patterns



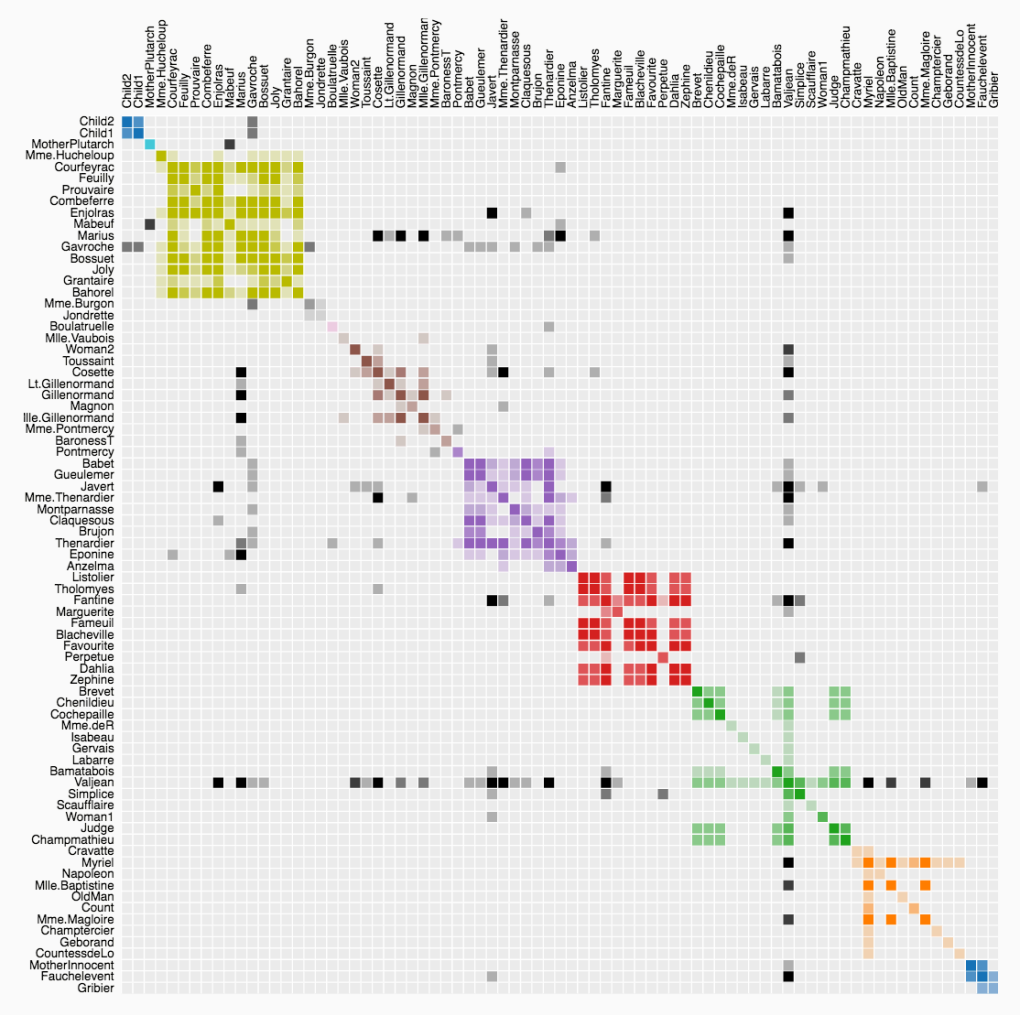


Default ordering hides patterns



https://bost.ocks.org/mike/miserables/

Cluster ordering reveals patterns



https://bost.ocks.org/mike/miserables/

Trust rank	Index rank	Borough	Amount approved (£)	Number of grants
1	3	Tower Hamlets	£9,692,642	269
2	2	Hackney	£7,809,608	225
3	12	Southwark	£7,266,118	232
4	14	Camden	£6,140,419	136
5	4	Islington	£5,424,137	156
6	8	Lambeth	£5,257,941	156
7	2	Newham	£5,217,075	154
8	13	Hammersmith and Fulham	£4,085,708	109
9	29	Merton	£3,656,112	113
10	20	Croydon	£3,629,066	127
11	9	Lewisham	£3,537,049	144
12	17	Westminster	£3,357,911	100
13	15	Ealing	£3,057,709	84
14	30	Bromley	£3,038,621	131
15	19	Kensington and Chelsea	£2,979,468	74
16	11	Brent	£2,898,224	85
17	10	Greenwich	£2,837,658	87
18	24	Barnet	£2,796,587	99
19	21	Wandsworth	£2,592,453	89
20	5	Waltham Forest	£2,505,730	131
21	28	Sutton	£2,468,511	87
22	18	Hounslow	£2,383,393	75
23	7	Haringey	£2,360,290	101
24	22	Redbridge	£2,285,173	75
25	33	Rechmond upon Thames	£2,249,983	133
26	23	Hullingdon	£2,181,566	103
27	16	Enfield	£2,145,800	86
28	6	Barking and Dagenham	£1,943,597	68
29	25	Havering	£1,934,424	95
30	26	Bexley	£1,631,415	103
31	27	Harrow	£1,516,193	62
32	31	Kingston upon Thames	£1,353,125	55
33	32	City of London	£402,060	11
		Several Additional Inner Bouroughs	£18,704,677	481
		Several Additional Outer Boroughs	£6,392,100	164
		Other	£28,566,830	566
		London-wide	£86,583,750	1214
		Total	£252,883,123	6180

Tables easily hide patterns

Total grants spend by London Borough

September 1995 to March 2011

http://www.storytellingwithdata.com/blog/2012/02/grables-and-taphs

Total grant spend by London Borough

September 1995 - March 2011

Borough	Trust rank	Index rank	Number of grants	Amount approved (£)
Tower Hamlets	1	3	269	9,692,642
Hackney	2	2	225	7,809,608
Southwark	3	12	232	7,266,118
Camden	4	14	136	6,140,419
Islington	5	4	156	5,424,137
Lambeth	6	8	156	5,257,941
Newham	7	2	154	5,217,075
Hammersmith and Fulham	8	13	109	4,085,708
Merton	9	29	113	3,656,112
Croydon	10	20	127	3,629,066
Lewisham	11	9	144	3,537,049
Westminster	12	17	100	3,357,911
Ealing	13	15	84	3,057,709
Bromley	14	30	131	3,038,621
Kensington and Chelsea	15	19	74	2,979,468
Brent	16	11	85	2,898,224
Greenwich	17	10	87	2,837,658
Barnet	18	24	99	2,796,587
Wandsworth	19	21	89	2,592,453
Waltham Forest	20	5	131	2,505,730
Sutton	21	28	87	2,468,511
Hounslow	22	18	75	2,383,393
Haringey	23	7	101	2,360,290
Redbridge	24	22	75	2,285,173
Rechmond upon Thames	25	33	133	2,249,983
Hullingdon	26	23	103	2,181,566
Enfield	27	16	86	2,145,800
Barking and Dagenham	28	6	68	1,943,597
Havering	29	25	95	1,934,42[illegible]
Bexley	30	26	103	1,631,4[illegible]15
Harrow	31	27	62	1,516,[illegible]93
Kingston upon Thames	32	31	55	1,353[illegible]125
City of London	33	32	11	[illegible]02,060
Several Additional Inner Bouroughs			481	18,704,677
Several Additional Outer Boroughs			164	6,392,100
Other			566	28,566,830
London-wide			1,214	86,583,750
Total			**6,180**	**252,883,123**

Help people see patterns in tables

http://www.storytellingwithdata.com/blog/2012/02/grables-and-taphs

Help viewers interpret tables

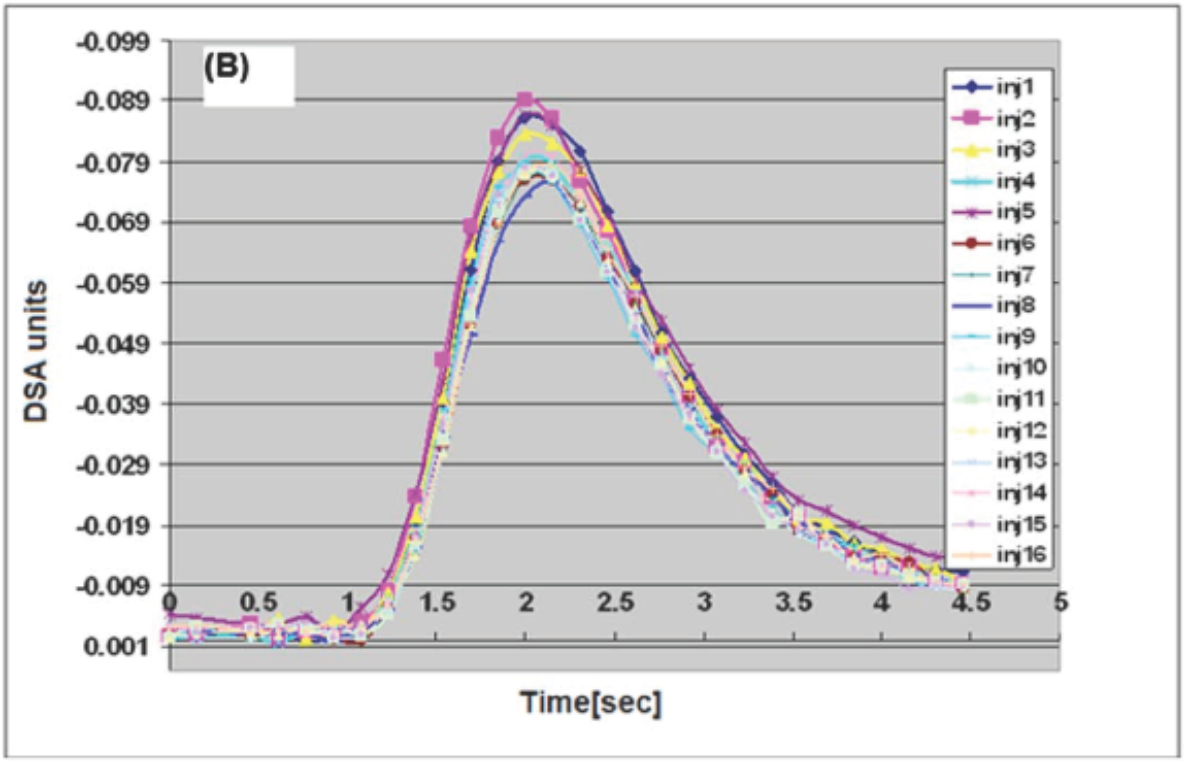
- Limit and standardize decimal places
- Emphasize important values with color, bold text and annotations
- Sort rows by values
- Turn table into another chart or a handout

MDG 4 Progress 1990-2013: Low Income Countries

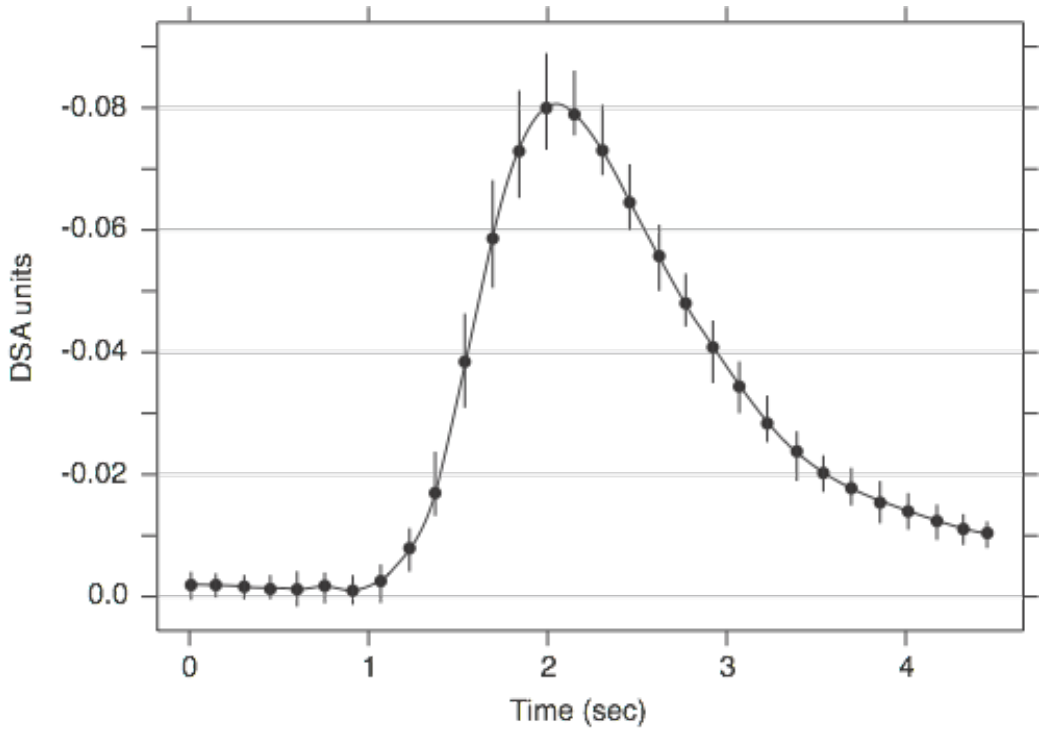
Country	1990 (Per 1,000)	2013 (per 1,000)	Percent Change	2015 Target	Percent of Target Met
Afghanistan	179	97	46%	60.9	69%
Bangladesh	144	41	72%	49	108%
Benin	179	85	53%	60.9	80%
Burkina Faso	202	98	51%	68.7	78%
Burundi	171	83	51%	58.1	78%
Cambodia	118	38	68%	40.1	103%
Central African Republic	177	139	21%	60.2	33%
Chad	215	148	31%	73.1	47%
Comoros	125	78	38%	42.5	57%
Congo, Democratic Republic of	176	119	32%	59.8	49%
Eritrea	151	50	67%	51.3	101%
Ethiopia	205	64	69%	69.7	104%
Gambia, The	170	74	56%	57.8	86%
Guinea	238	101	58%	80.9	87%
Guinea-Bissau	225	124	45%	76.5	68%
Haiti	145	73	50%	49.3	75%
Kenya	99	71	28%	33.7	43%
Korea, DPR	43	27	37%	14.6	56%
Liberia	248	71	71%	84.3	108%
Madagascar	161	56	65%	54.7	99%
Malawi	245	68	72%	83.3	109%
Mali	254	123	52%	86.4	78%
Mozambique	237	87	63%	80.6	96%
Myanmar	109	51	53%	37.1	81%
Nepal	142	40	72%	48.3	109%
Niger	327	104	68%	111.2	103%
Rwanda	152	52	66%	51.7	99.70%
Sierra Leone	268	161	40%	91.1	60%
Somalia	180	146	19%	61.2	29%
Tajikistan	108	48	56%	39.8	84%
Tanzania	167	52	69%	56.8	104%
Togo	146	85	42%	49.6	63%
Uganda	179	66	63%	60.9	96%
Zimbabwe	75	89	-19%	25.5	-28%

https://dpt.duhs.duke.edu/files/Group%2011.pdf

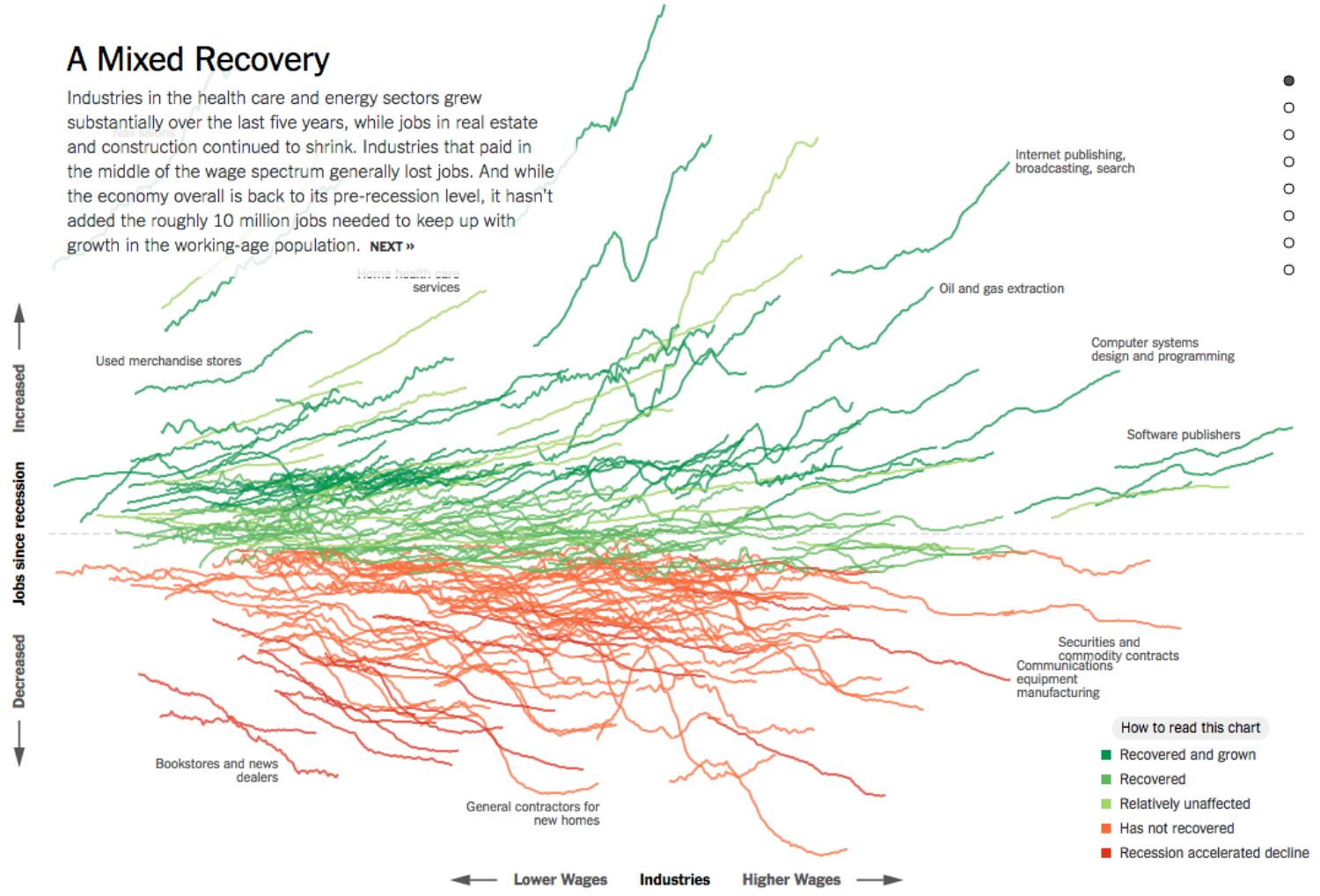
Leave out non-story details



Leave out non-story details

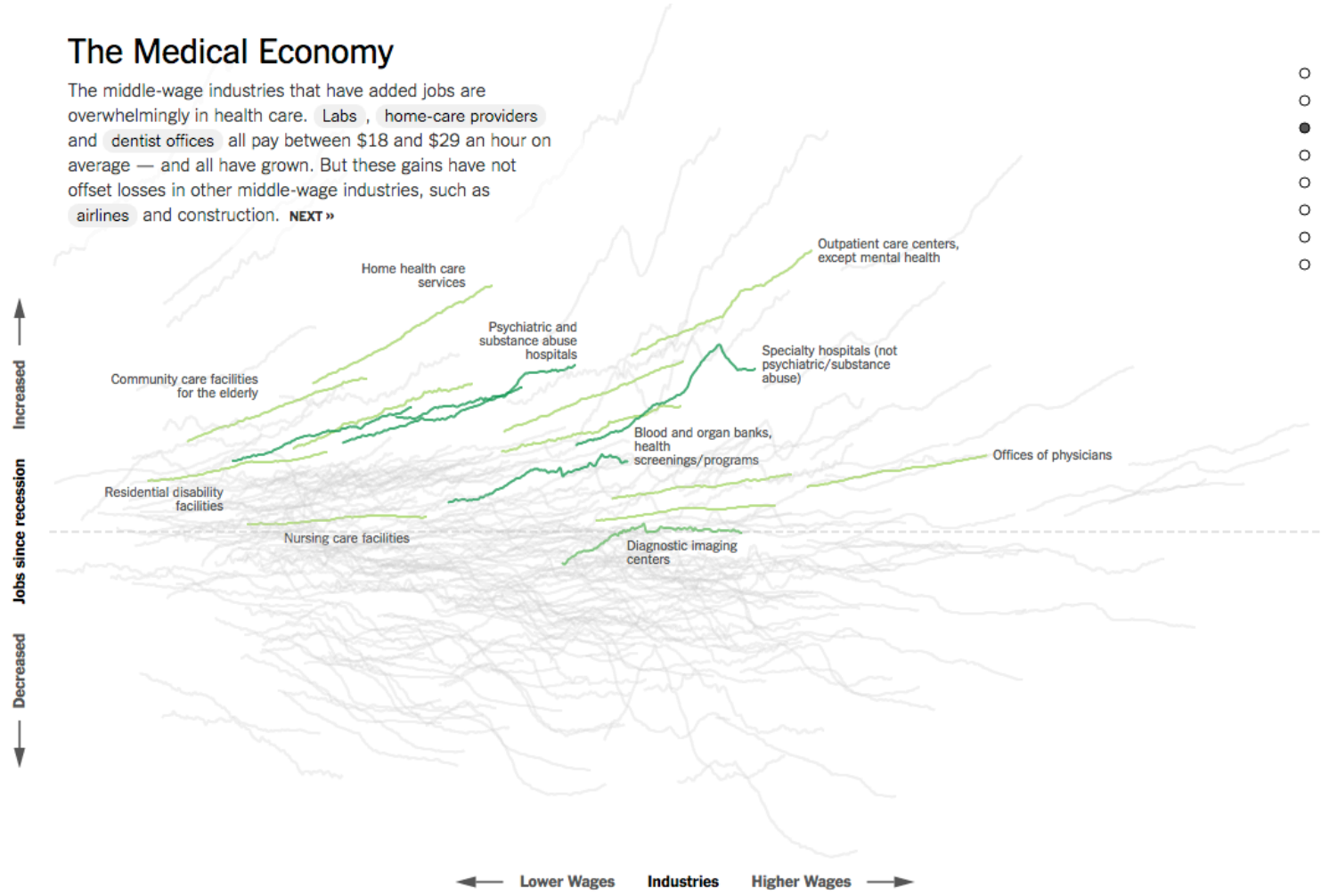


All the data doesn't tell a story



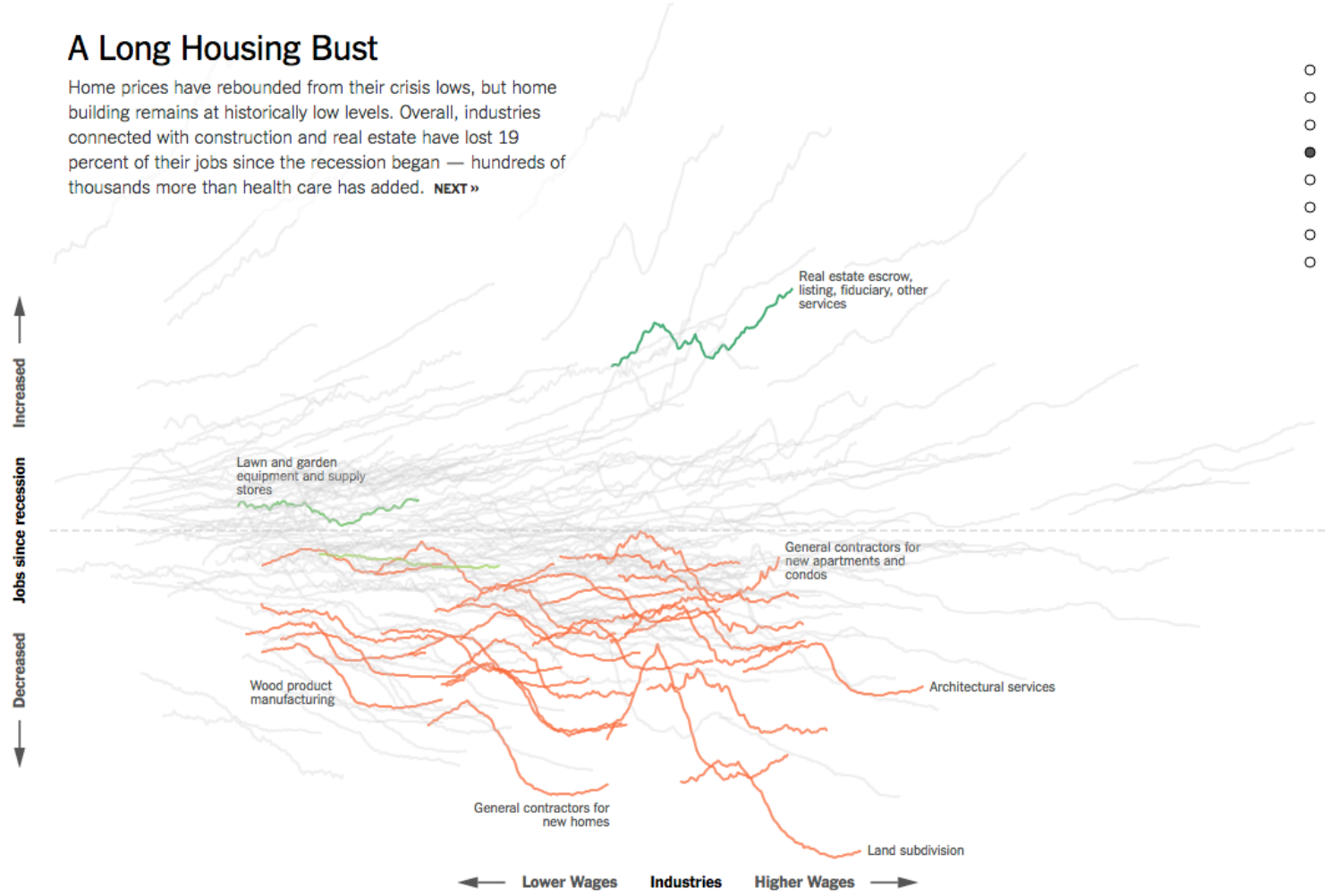
http://www.nytimes.com/interactive/2014/06/05/upshot/how-the-recession-reshaped-the-economy-in-255-charts.html

All the data doesn't tell a story



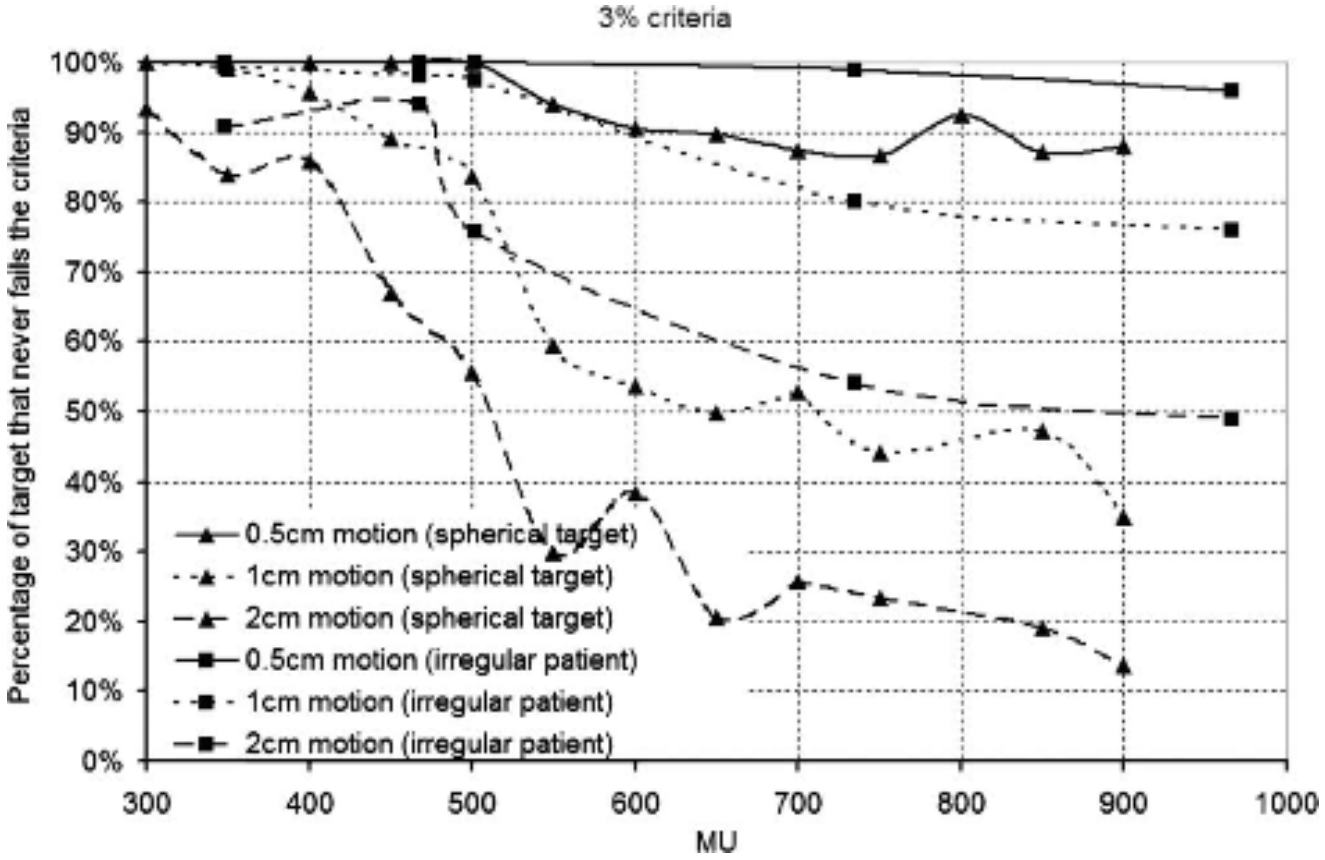
http://www.nytimes.com/interactive/2014/06/05/upshot/how-the-recession-reshaped-the-economy-in-255-charts.html

All the data doesn't tell a story

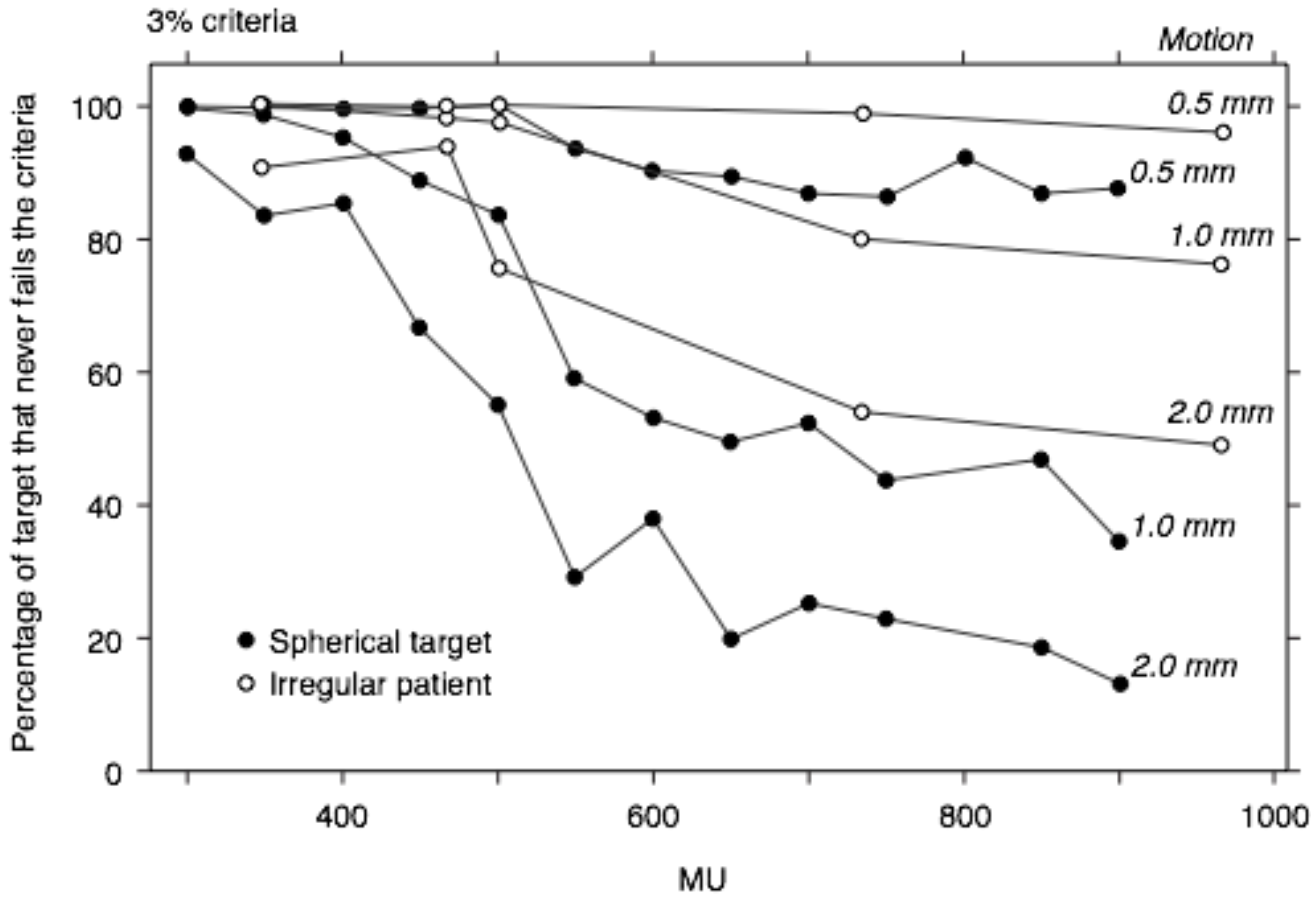


http://www.nytimes.com/interactive/2014/06/05/upshot/how-the-recession-reshaped-the-economy-in-255-charts.html

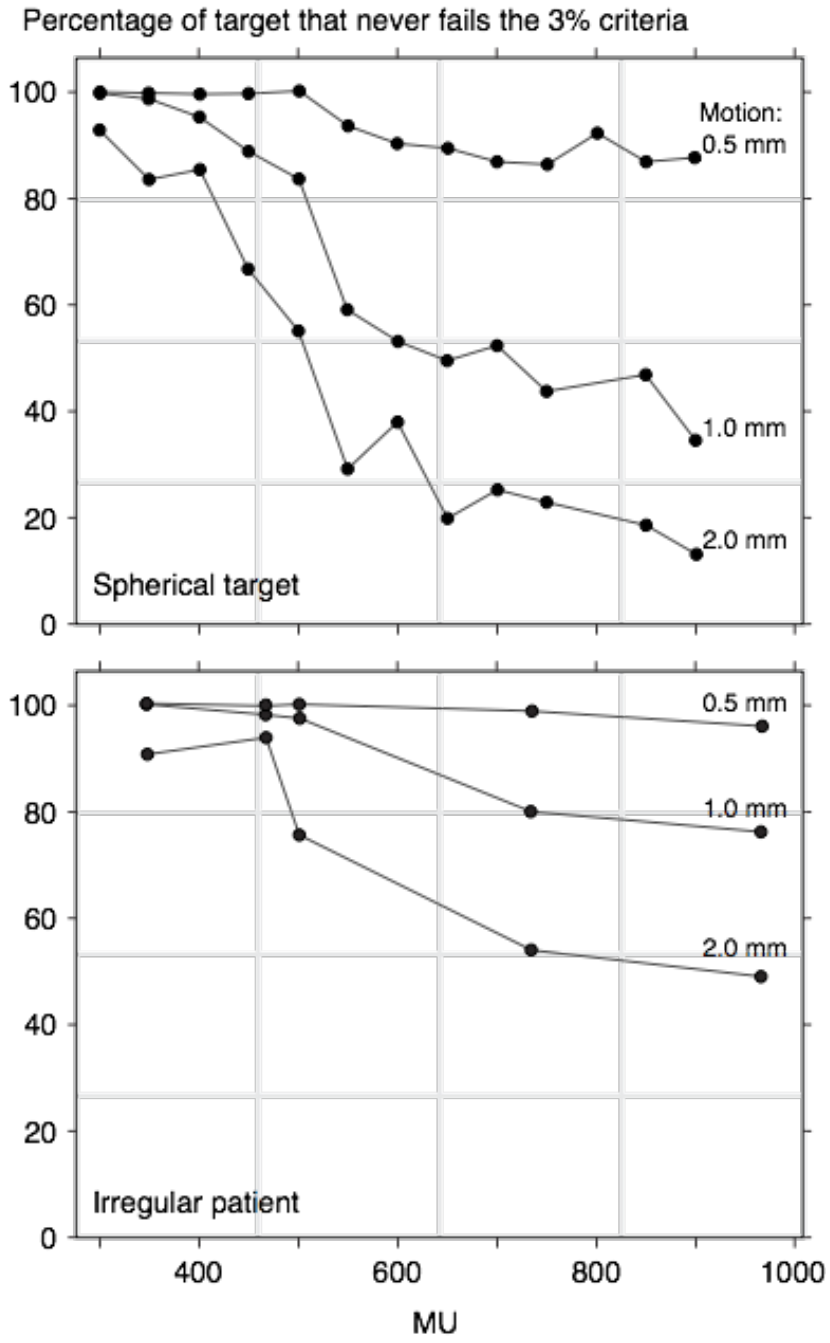
Original


3% criteria
Percentage of target that never fails the criteria
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
300
400
500
600
700
800
900
1000
MU
0.5cm motion (spherical target)
1cm motion (spherical target)
2cm motion (spherical target)
0.5cm motion (irregular patient)
1cm motion (irregular patient)
2cm motion (irregular patient)

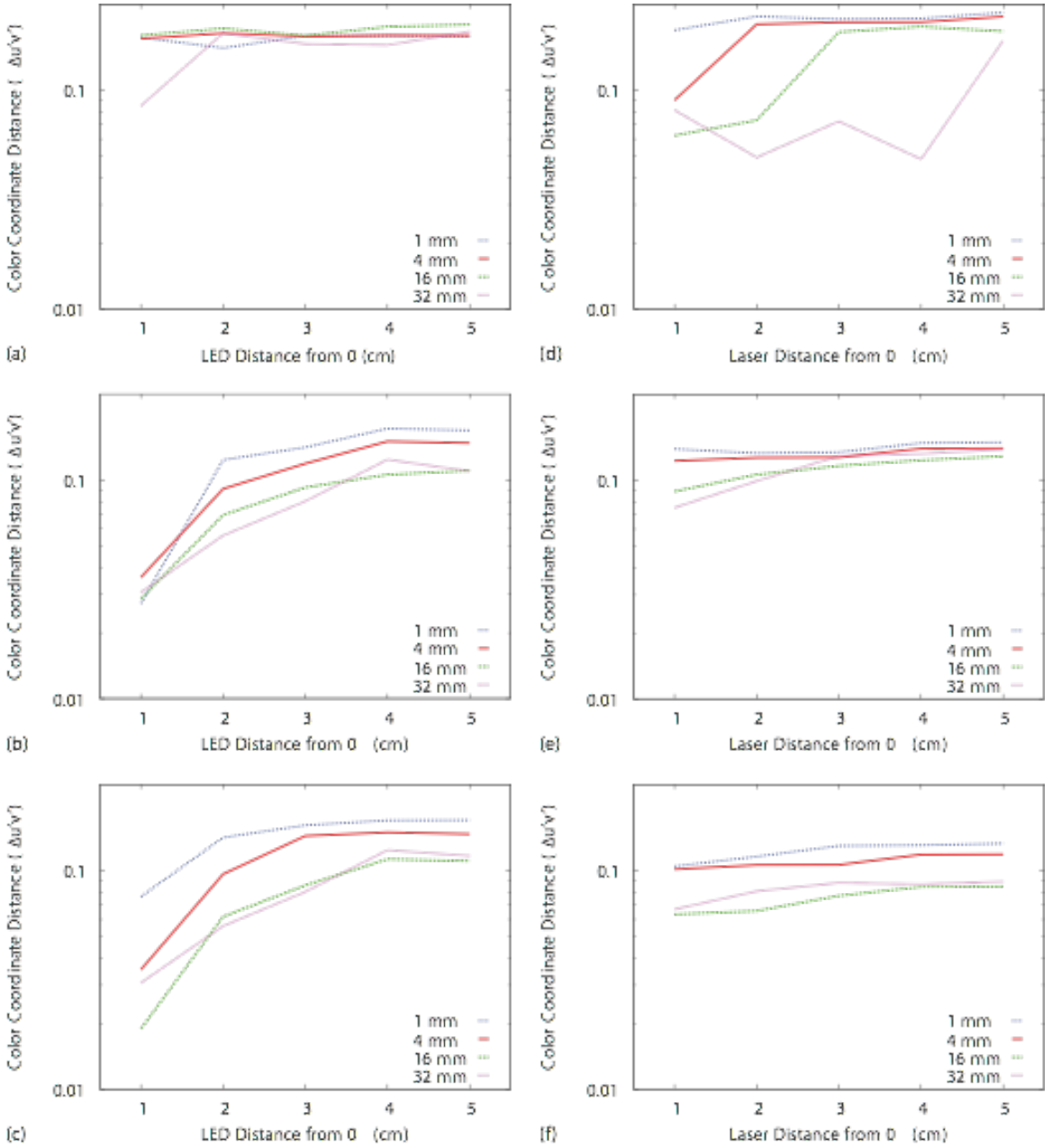
Reworked as single plot



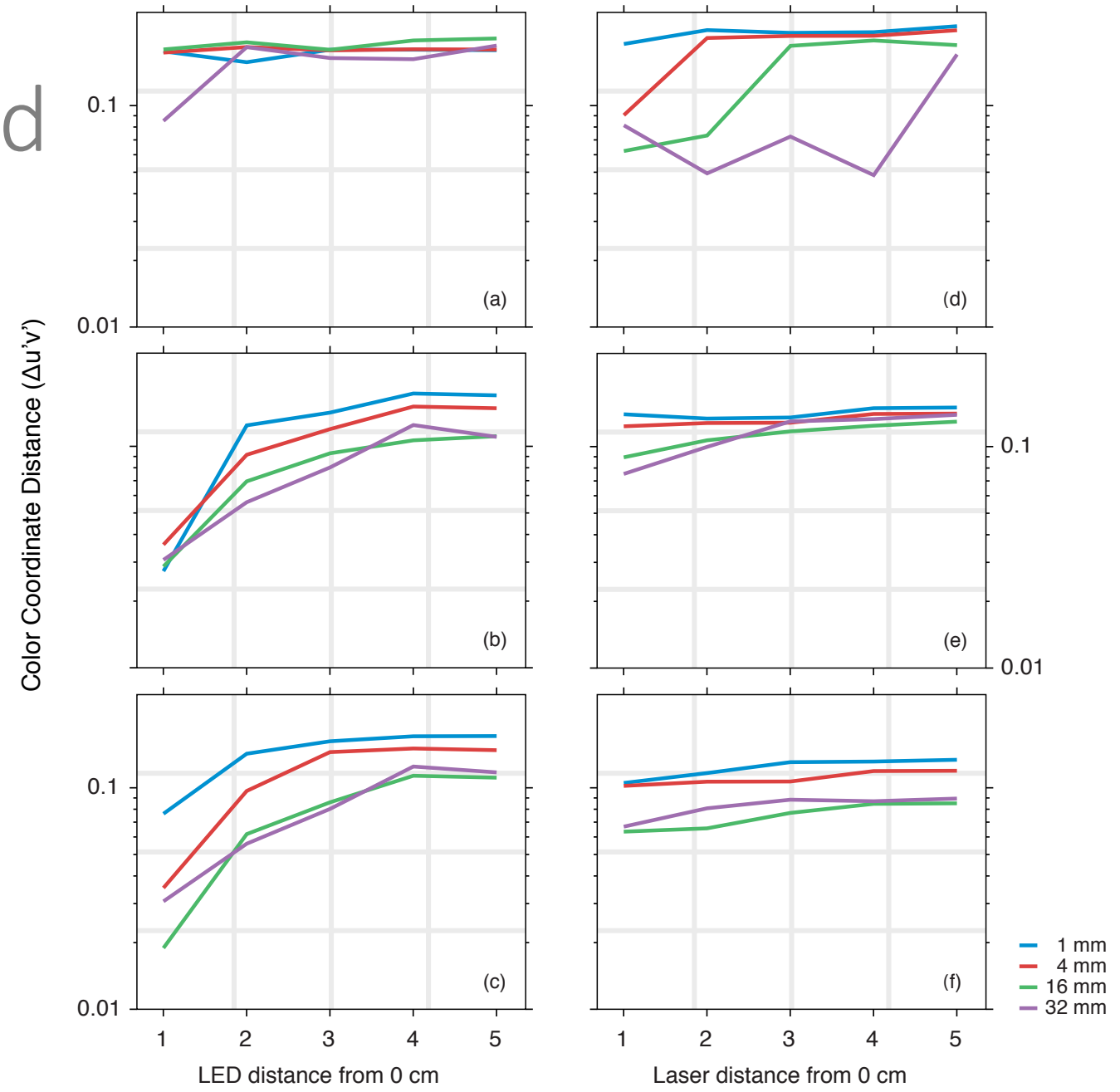
Reworked as small multiples



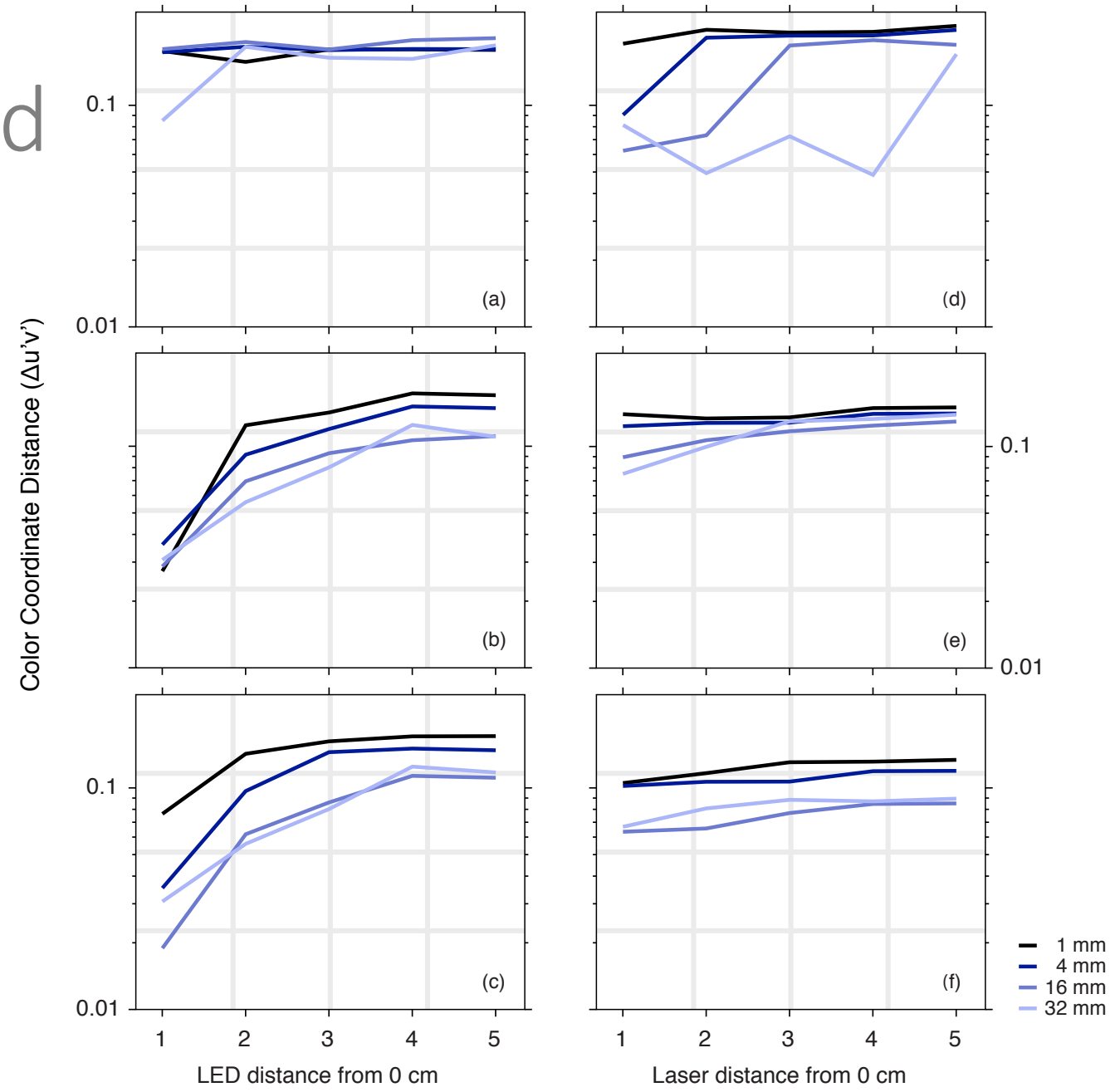
Original

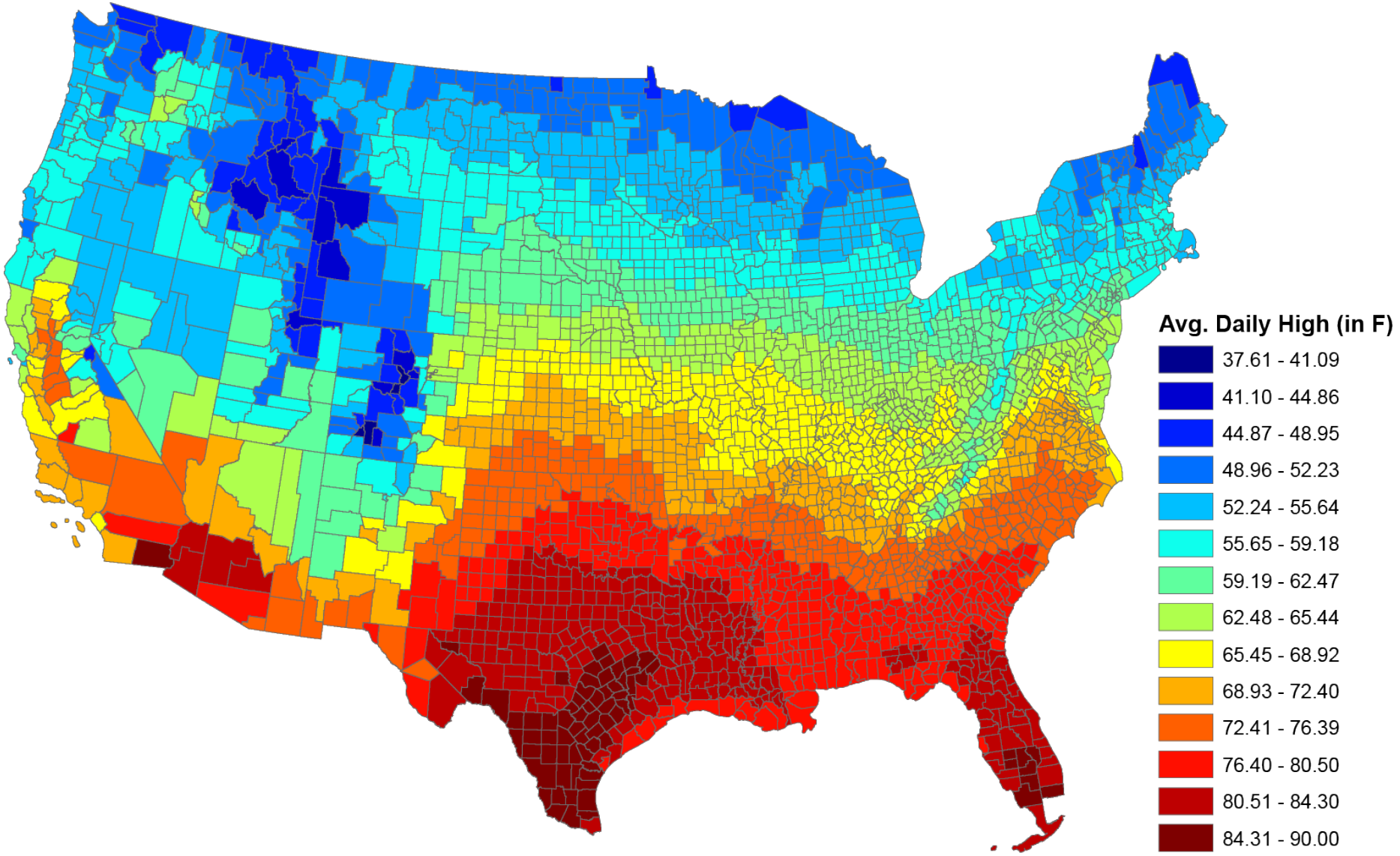

Color Coordinate Distance (Δu'v')
0.1
0.01
1 mm
4 mm
16 mm
32 mm
1
2
3
4
5
(a)
LED Distance from 0 (cm)
(d)
Laser Distance from 0 (cm)
(b)
LED Distance from 0 (cm)
(e)
Laser Distance from 0 (cm)
(c)
LED Distance from 0 (cm)
(f)
Laser Distance from 0 (cm)

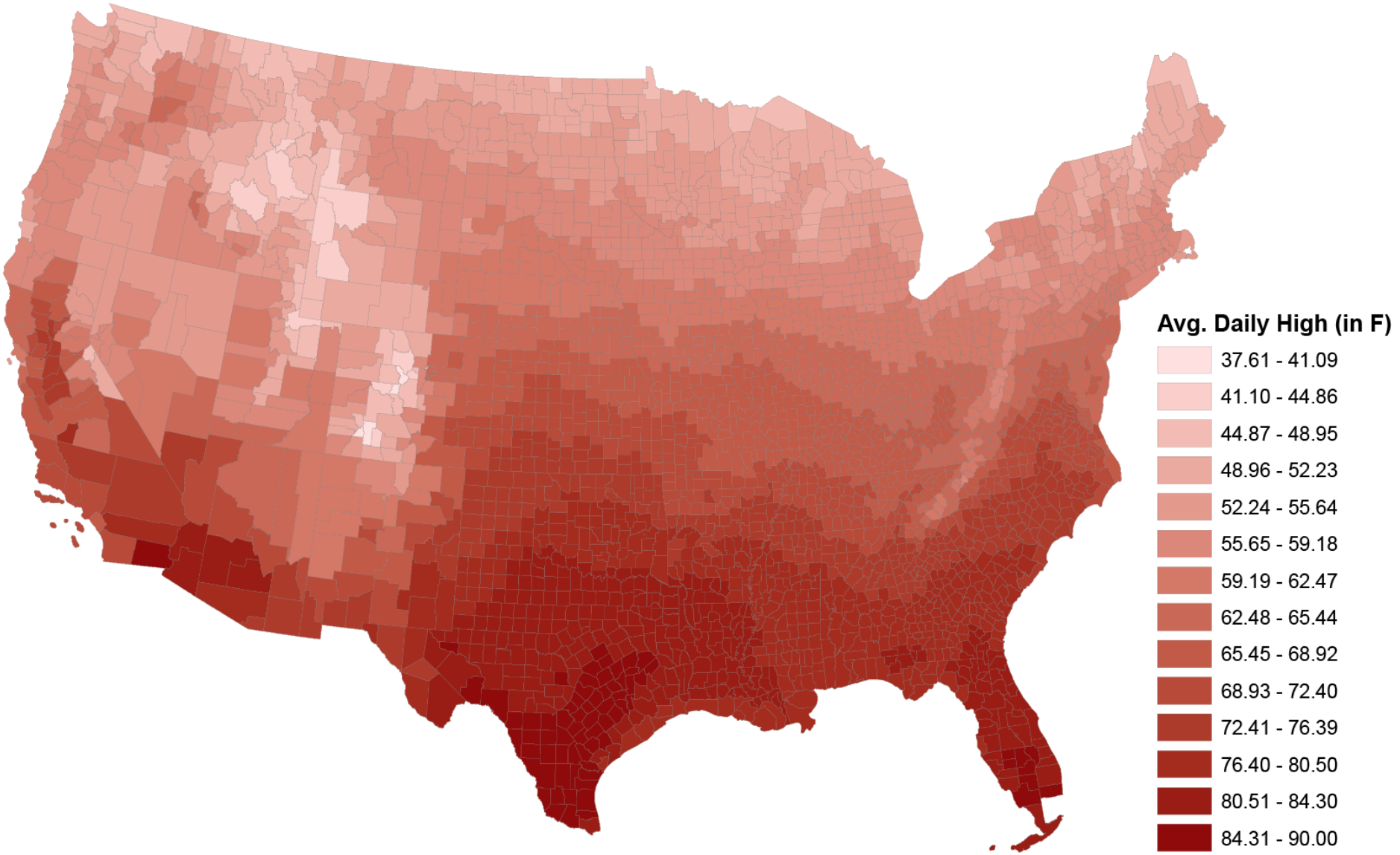
Reworked


Color Coordinate Distance (Δu'v')
0.1
0.01
(a)
(b)
(c)
(d)
(e)
(f)
1
2
3
4
5
LED distance from 0 cm
Laser distance from 0 cm
1 mm
4 mm
16 mm
32 mm

Reworked

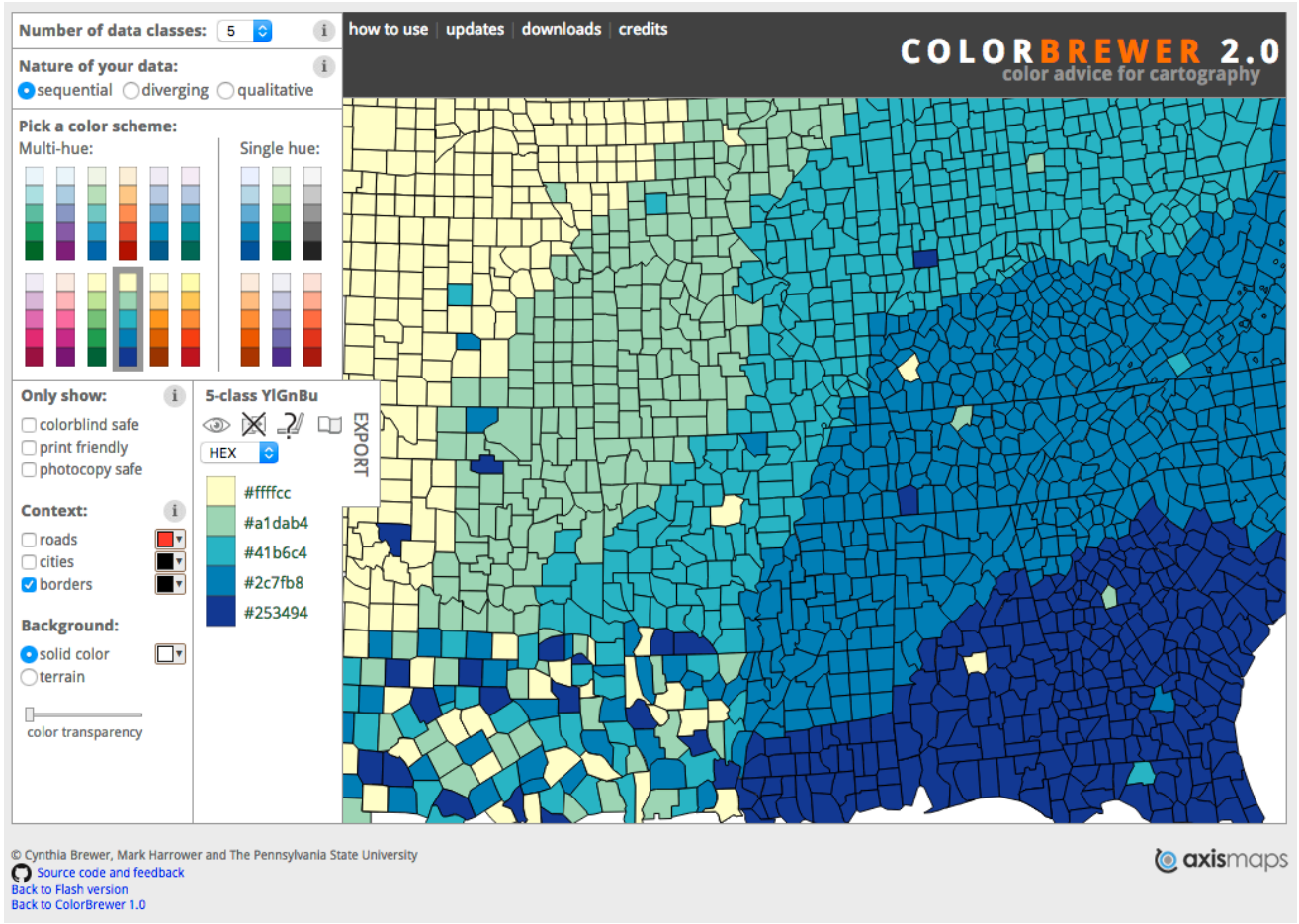

Color Coordinate Distance (Δu'v')
0.1
0.01
(a)
(b)
(c)
(d)
(e)
(f)
1
2
3
4
5
LED distance from 0 cm
Laser distance from 0 cm
1 mm
4 mm
16 mm
32 mm

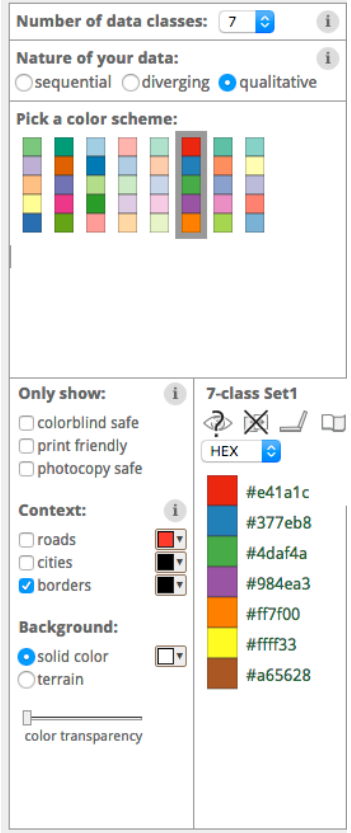

Avg. Daily High (in F)
37.61 - 41.09
41.10 - 44.86
44.87 - 48.95
48.96 - 52.23
52.24 - 55.64
55.65 - 59.18
59.19 - 62.47
62.48 - 65.44
65.45 - 68.92
68.93 - 72.40
72.41 - 76.39
76.40 - 80.50
80.51 - 84.30
84.31 - 90.00


Avg. Daily High (in F)
37.61 - 41.09
41.10 - 44.86
44.87 - 48.95
48.96 - 52.23
52.24 - 55.64
55.65 - 59.18
59.19 - 62.47
62.48 - 65.44
65.45 - 68.92
68.93 - 72.40
72.41 - 76.39
76.40 - 80.50
80.51 - 84.30
84.31 - 90.00

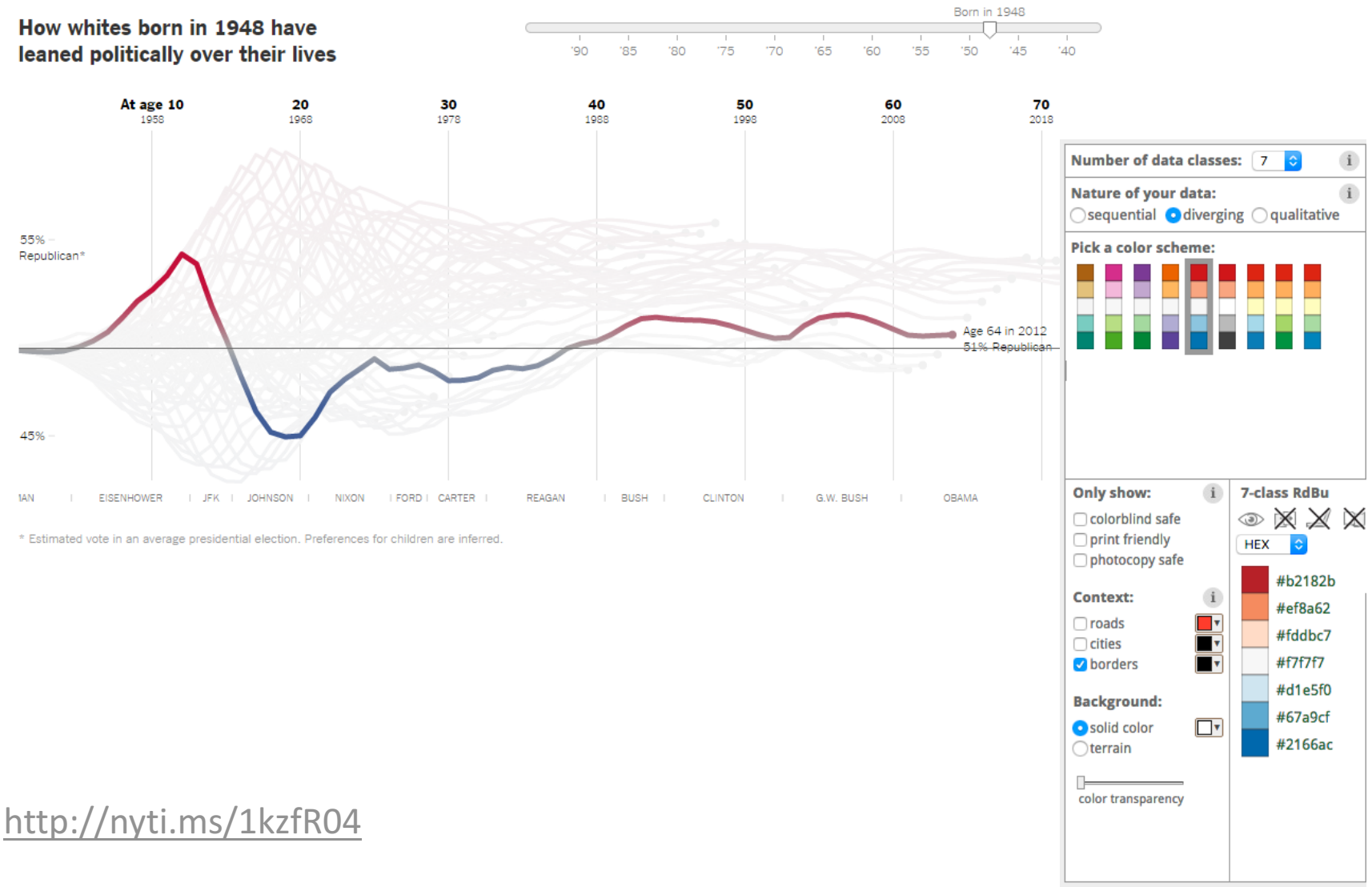
ColorBrewer for good colormaps

http://colorbrewer2.org/





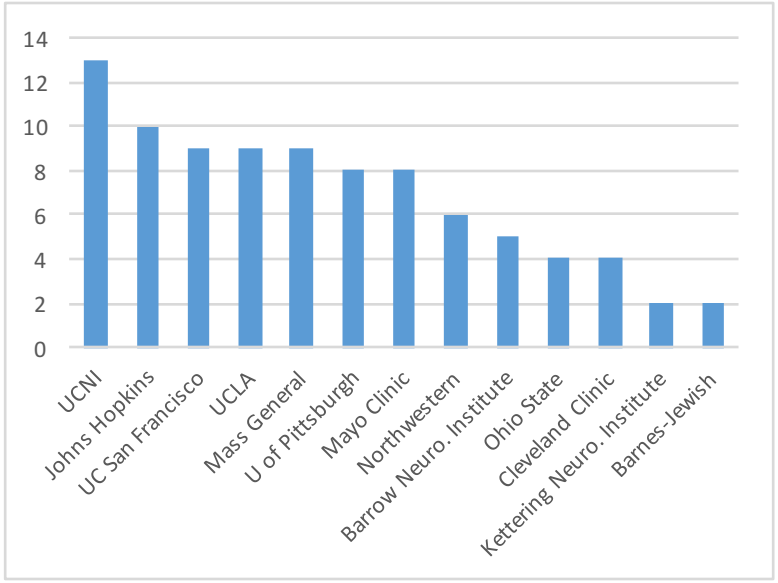
Do data have a natural center?

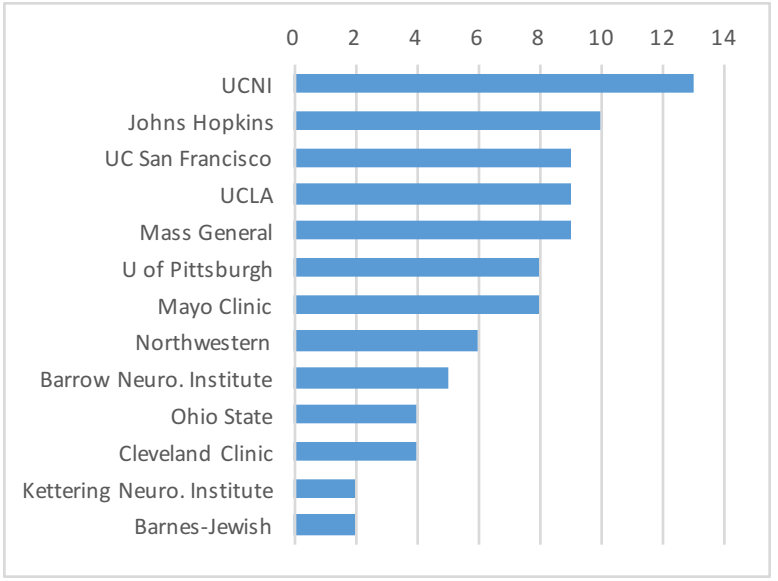


http://nyti.ms/1kzfR04

Text to clarify

Keep text horizontal

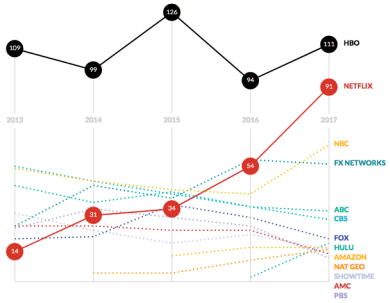




http://www.storytellingwithdata.com/2012/09/some-finer-points-of-data-visualization.html

Annotate figures directly

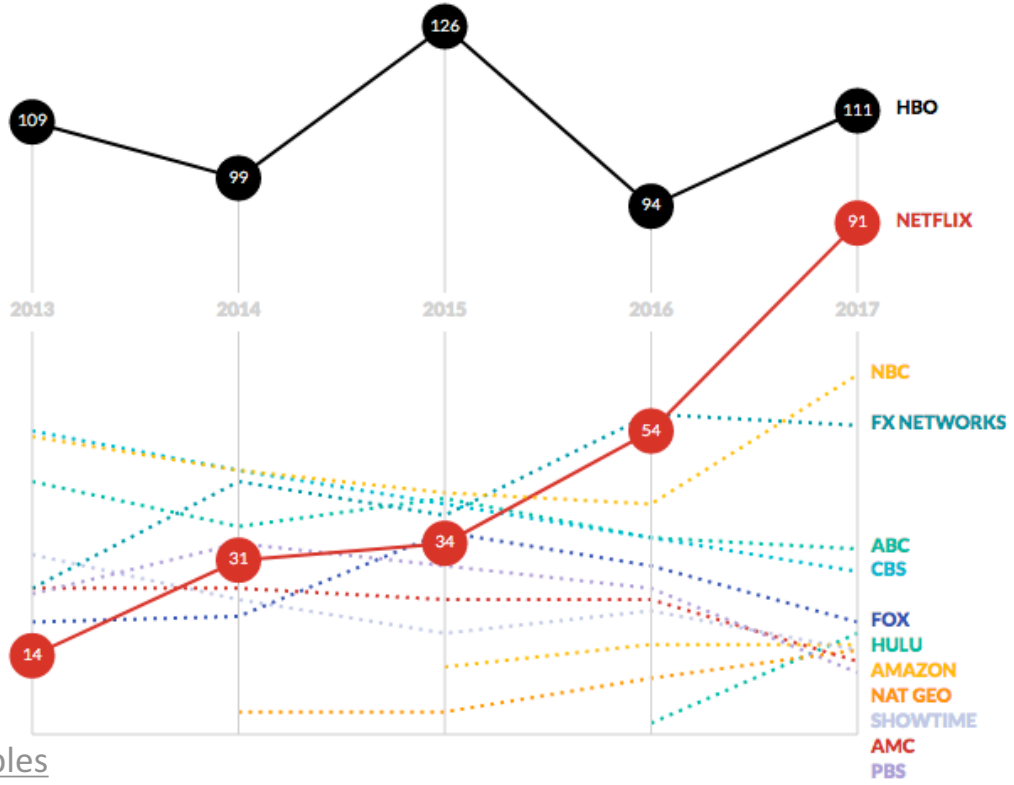
Netflix Challenges **HBO** at the 2017 Emmys



http://d3-annotation.susielu.com/#examples

Annotate figures directly

Netflix Challenges **HBO** at the 2017 Emmys

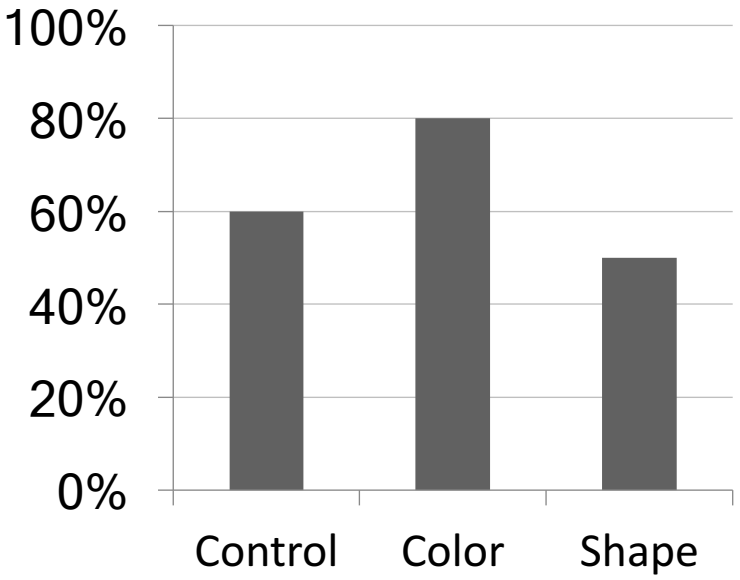


http://d3-annotation.susielu.com/#examples

Use descriptive titles

Active titles summarize trends in the figure and reinforce your message.

Accuracy versus Color and Shape



Accuracy Improved by Color, not Shape

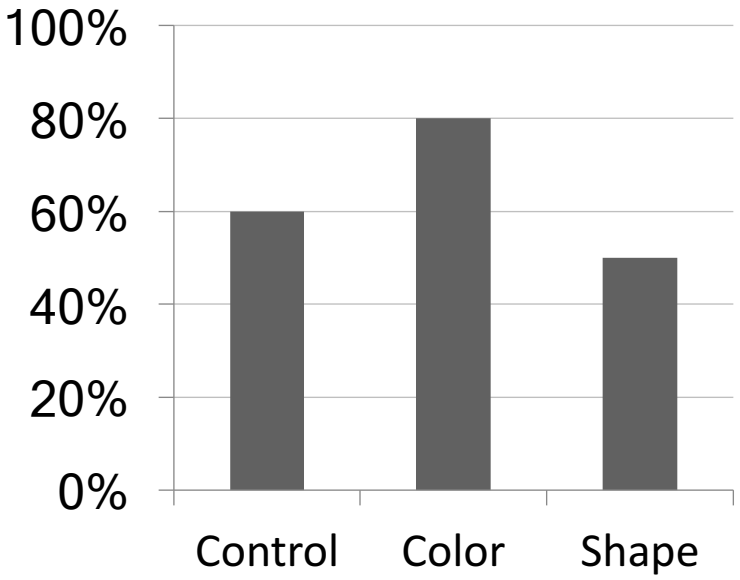
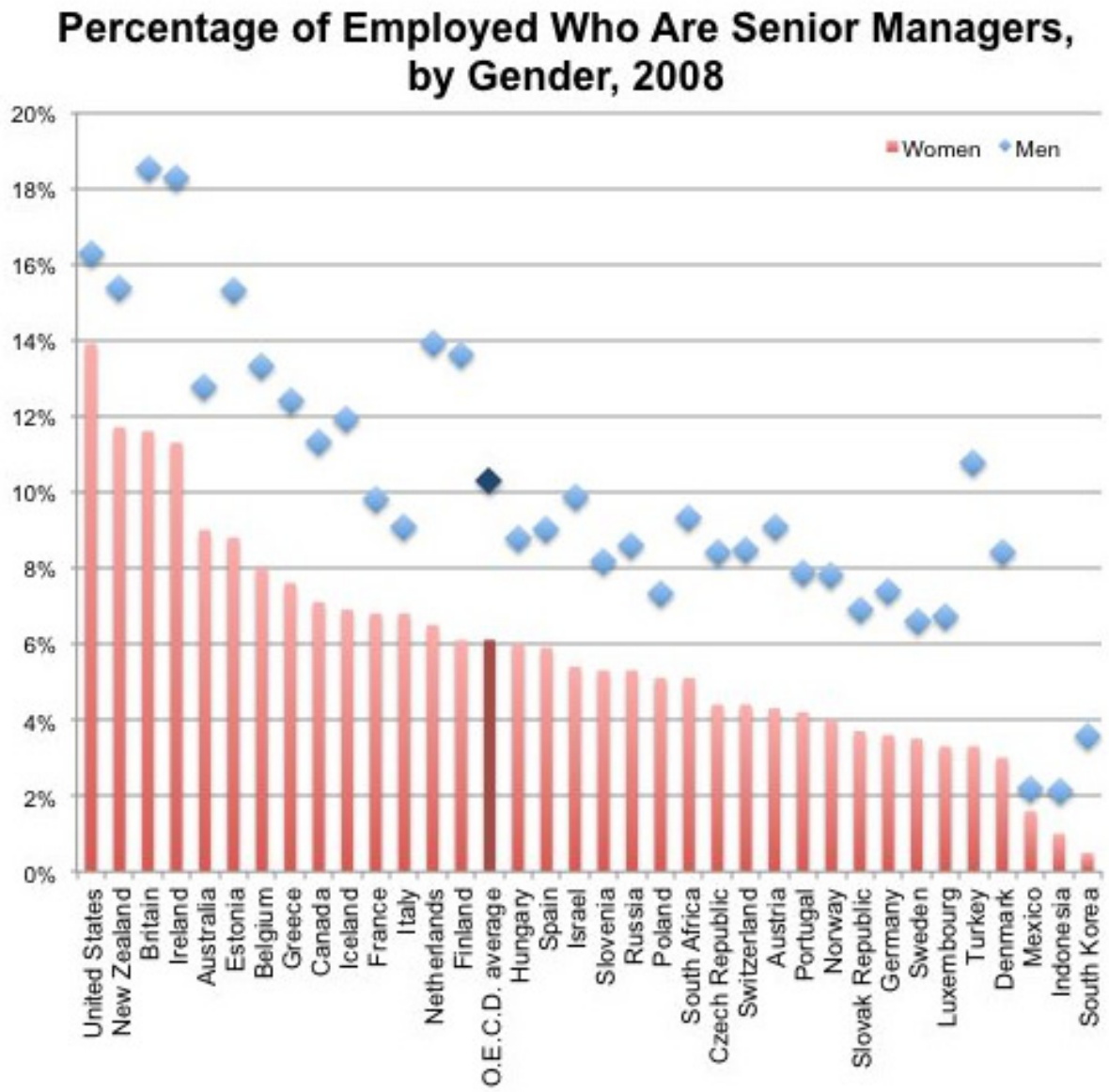
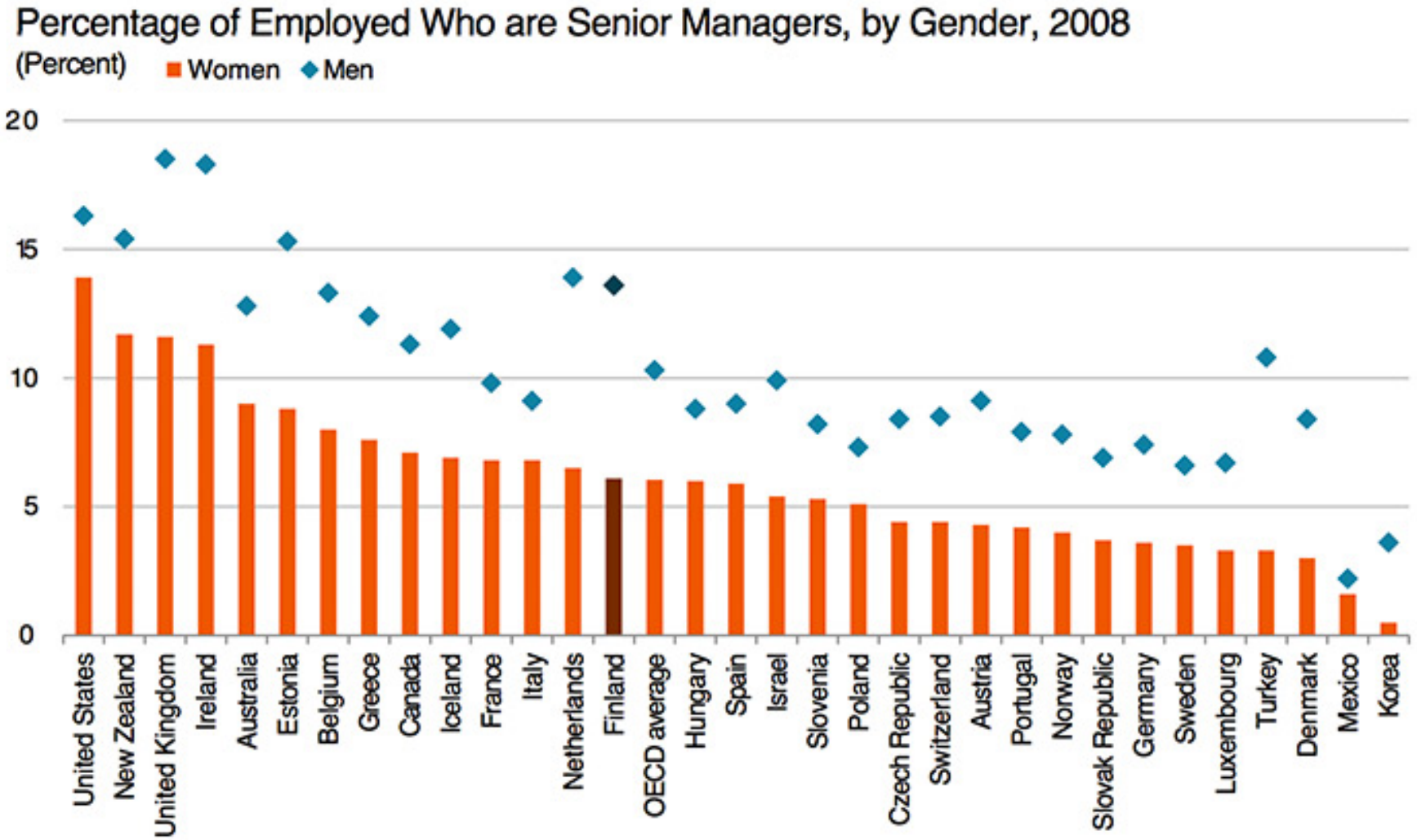


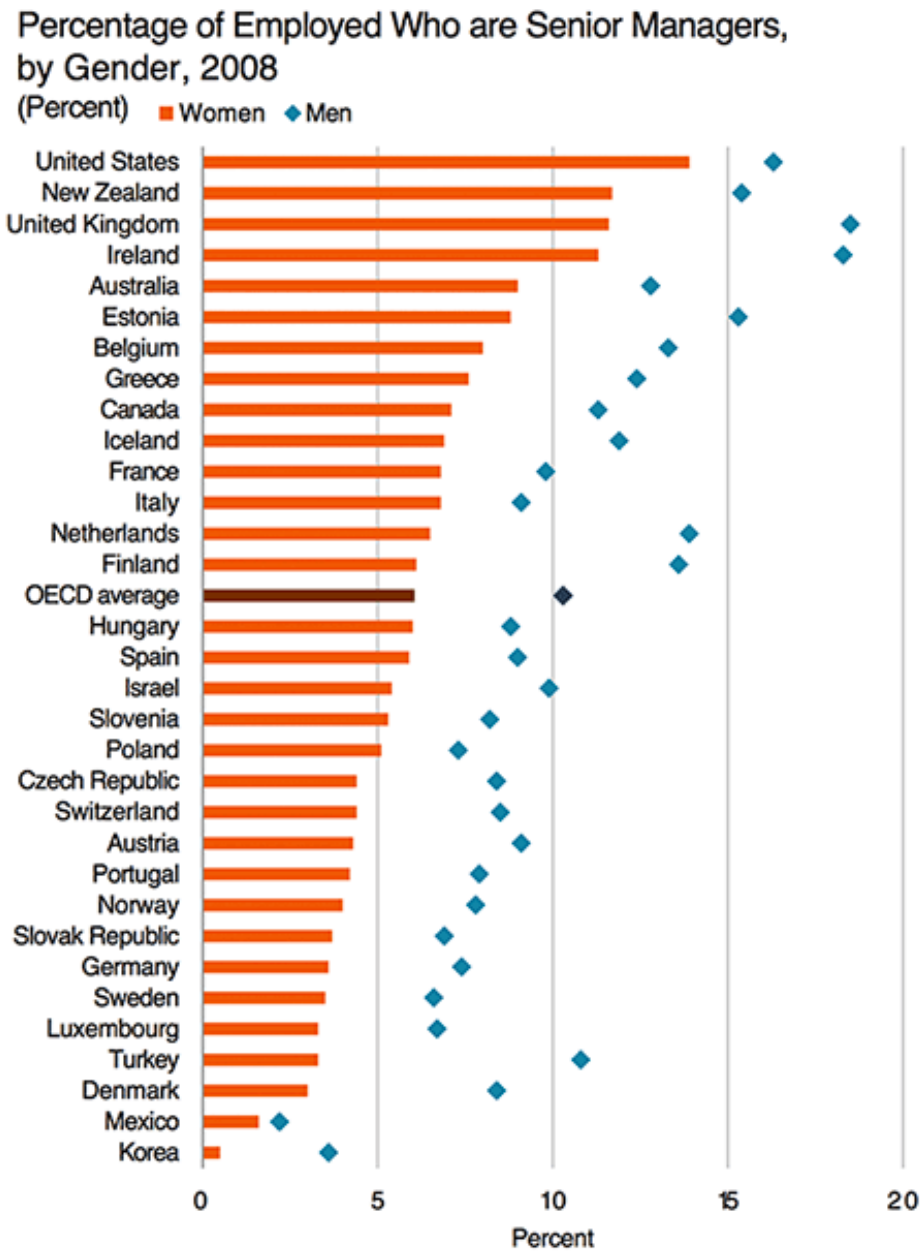
Figure rework



Jon Schwabish: http://thewhyaxis.info/gap-remake/



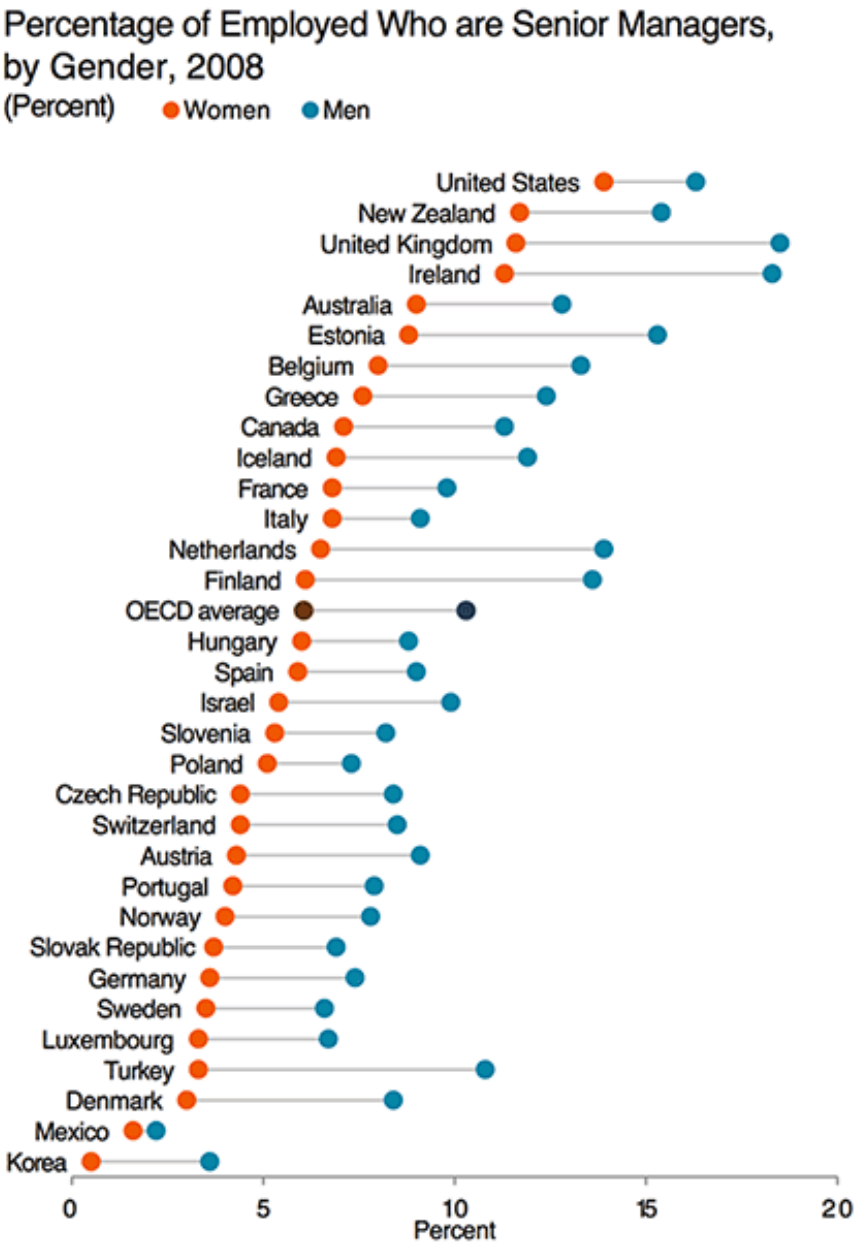
Jon Schwabish: http://thewhyaxis.info/gap-remake/



Jon Schwabish: http://thewhyaxis.info/gap-remake/

Jon Schwabish: http://thewhyaxis.info/gap-remake/

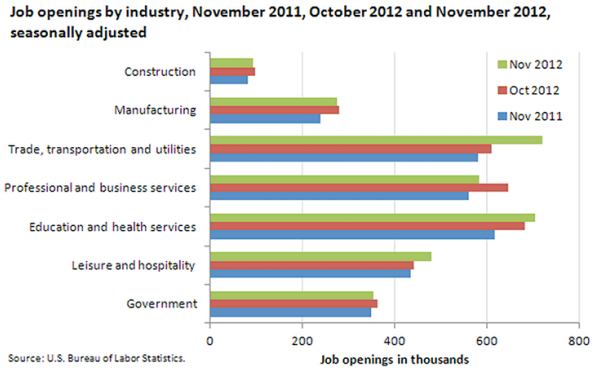
Jon Schwabish: http://thewhyaxis.info/gap-remake/

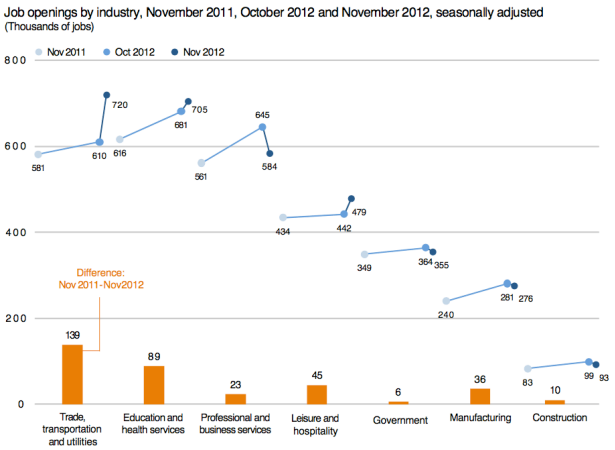


Jon Schwabish: http://thewhyaxis.info/gap-remake/

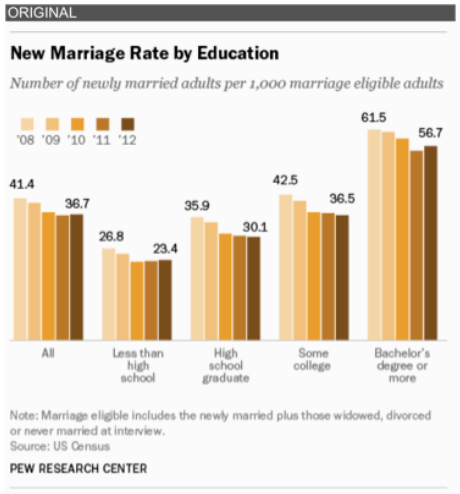
Other chart makeover examples

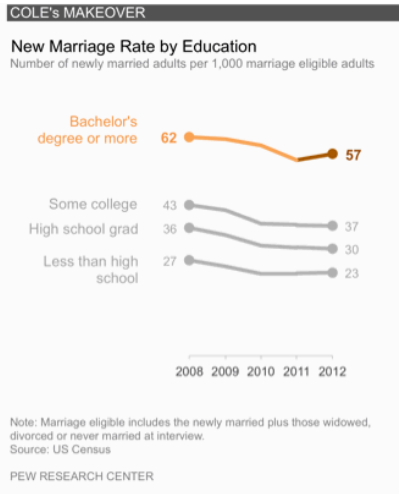
The Why Axis chart remakes
http://thewhyaxis.info/remakes/





Storytelling With Data visual makeovers:
http://www.storytellingwithdata.com/?tag=Visual+Makeover





Excel rework activity