

Tips for Effective Data Visualization

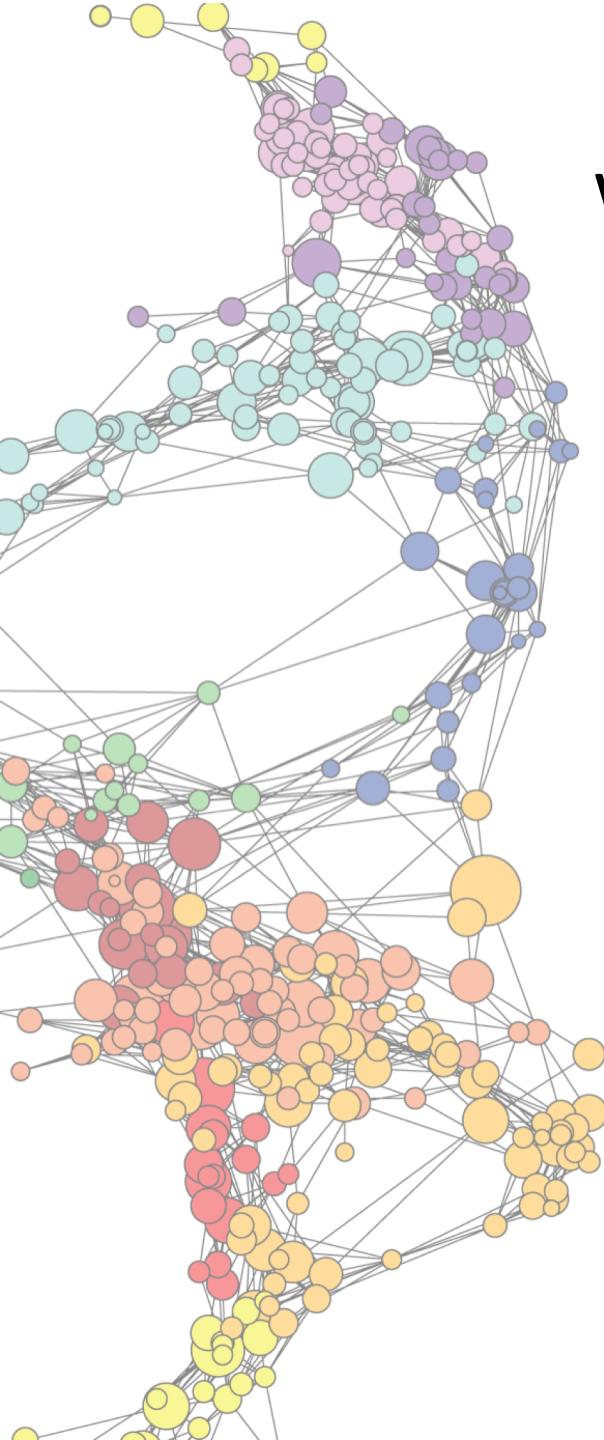


Eric E Monson, PhD

Center for Data and Visualization Sciences

STA 199 · Fall 2019

Slides: <http://bit.ly/STA199visFall2019>



What is data visualization?

**Anything that converts
data sources into a
visual representation**

*charts, graphs, maps –
even just tables!*

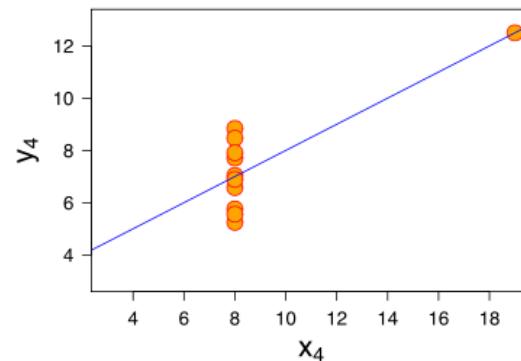
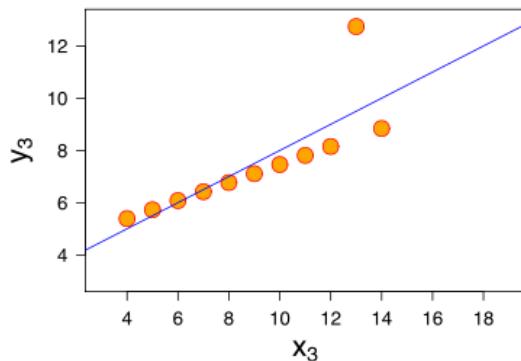
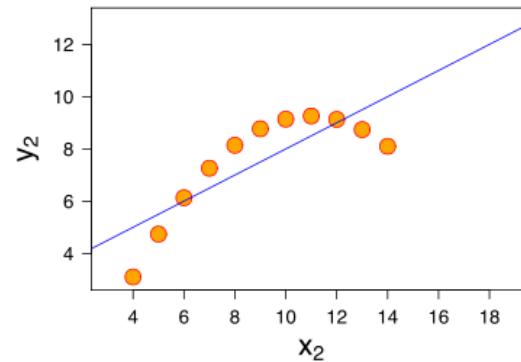
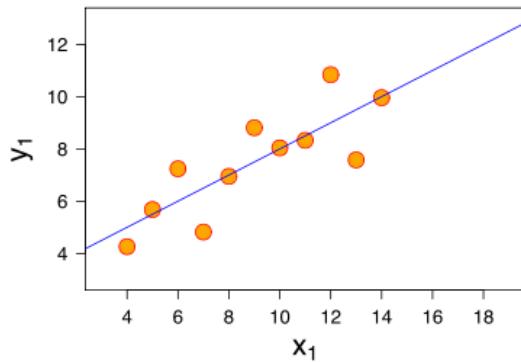
Why do we visualize?

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

Exploring data quality problems

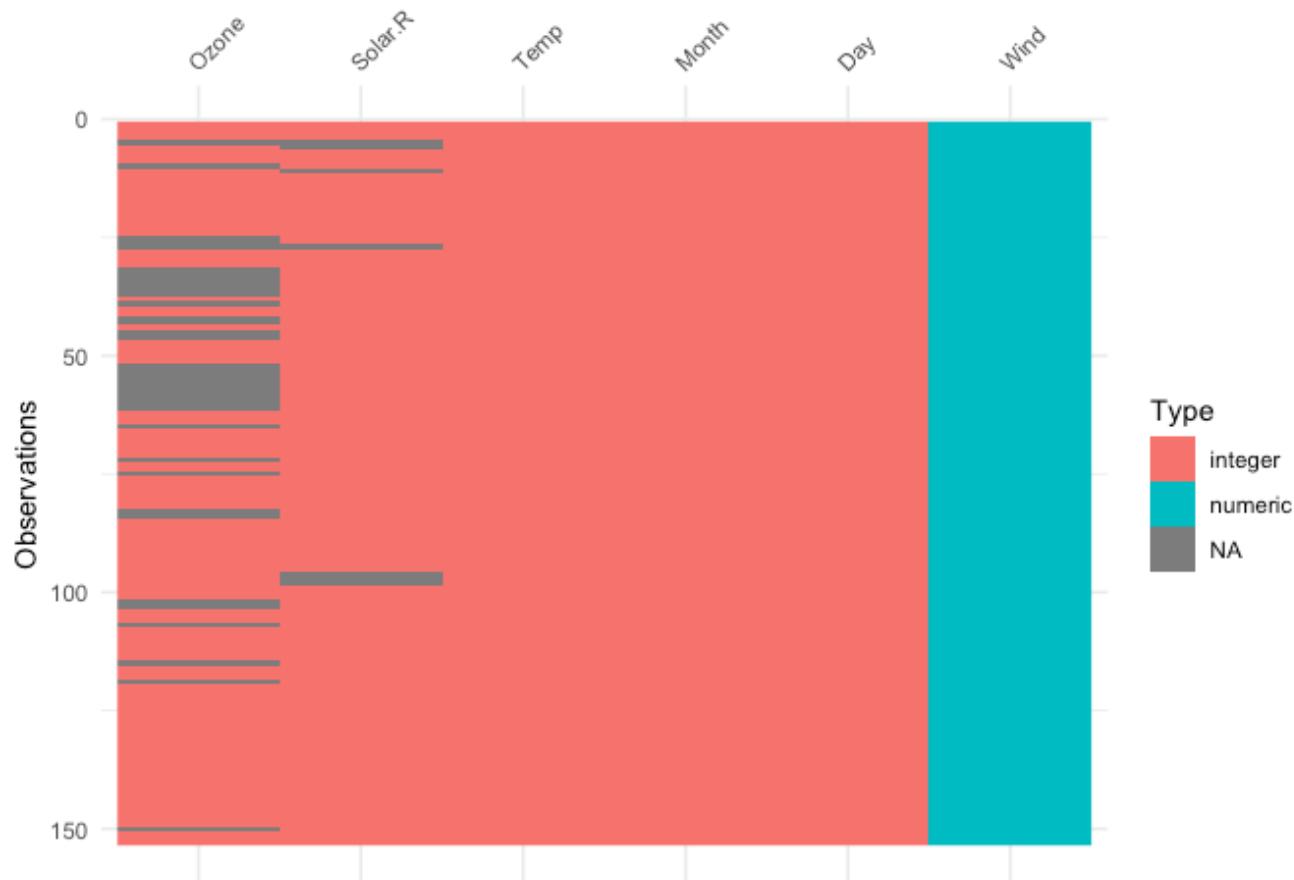


Chart choosing:
Starting points

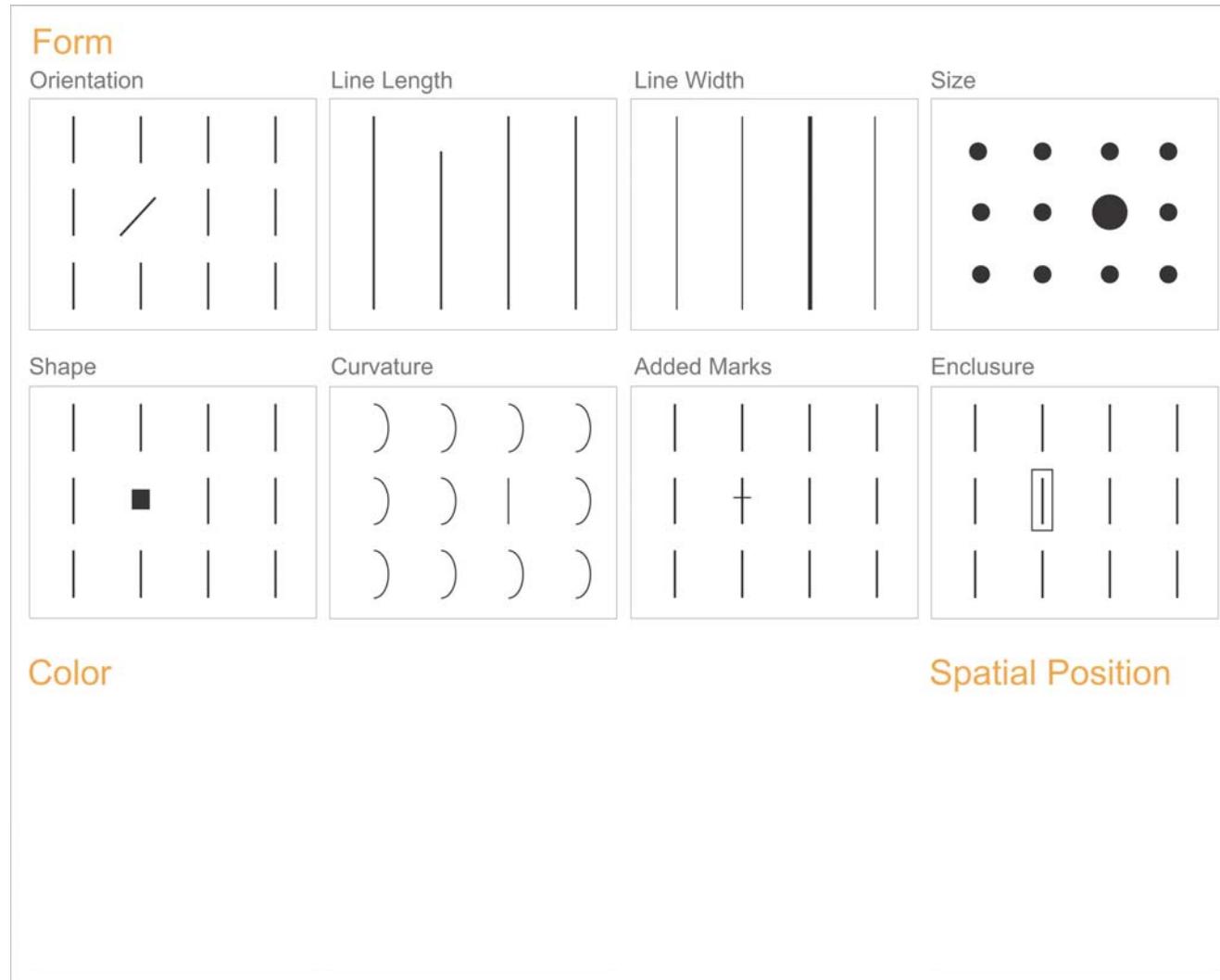
Pre-attentive visual attributes

Form

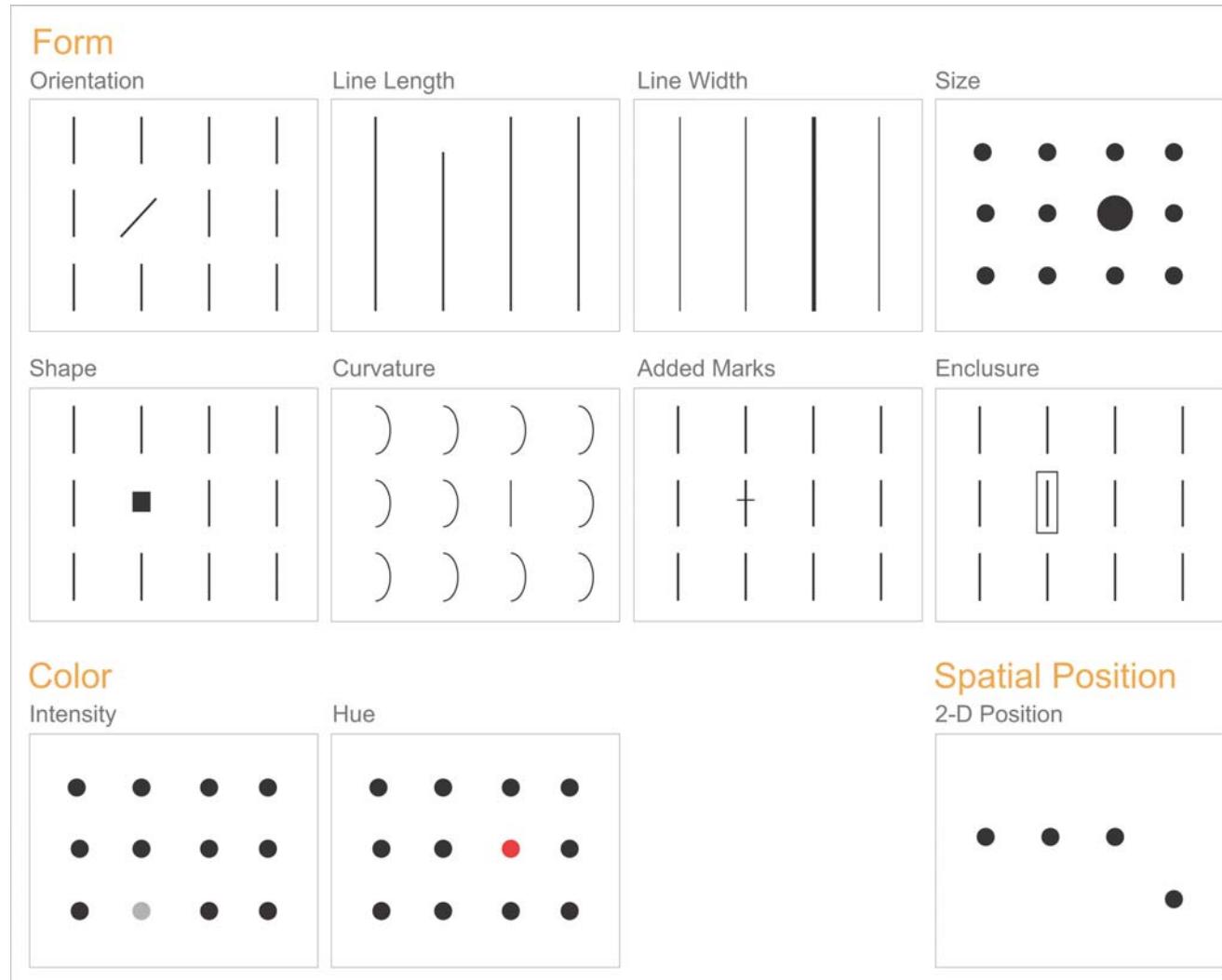
Color

Spatial Position

Pre-attentive visual attributes



Pre-attentive visual attributes



Encoding Choices

Magnitude (numerical)

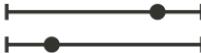
Identity (categorical)



Encoding Choices

Magnitude (numerical)

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)



Identity (categorical)

Most

Effectiveness

Same

Least

Encoding Choices

Magnitude (numerical)

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)



Identity (categorical)

Spatial region



Color hue



Motion

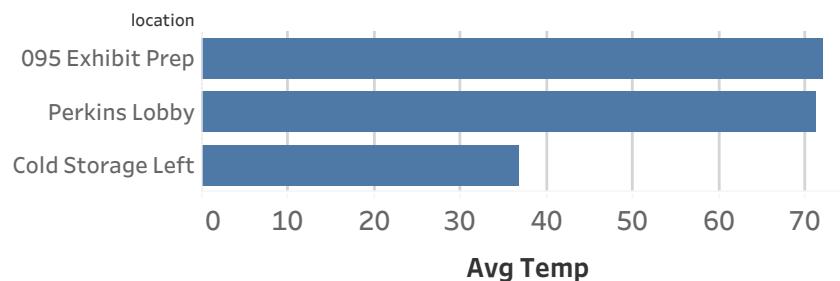


Shape

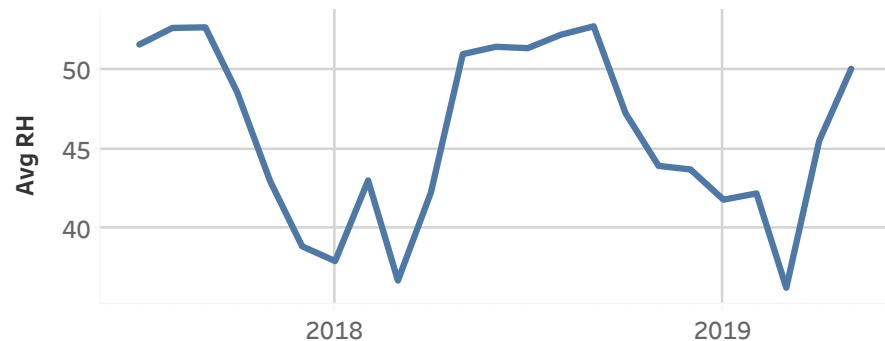


Classic charts because they work well – good starting point!

Categorical + Numerical
Bar



Date/time + Numerical
Line



Two Numerical (correlation)
Scatter

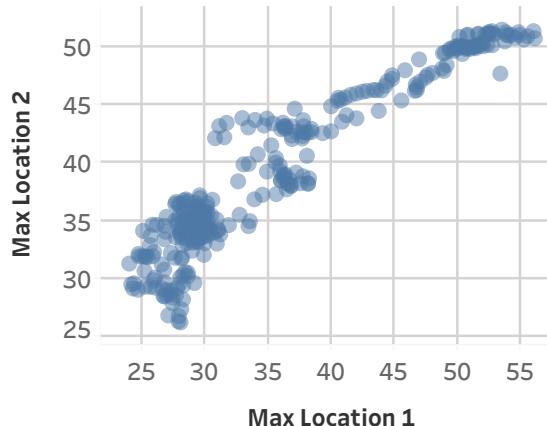


Chart choosing:

**Make the most important
comparisons easy**

You have to decide what's important!

Table

	Green	Yellow	Cheap	Tasty	Gross
Corn	6	29	18	30	7
Squash	8	27	17	13	11
Brussel sprouts	10	21	16	4	19
Green beans	20	17	16	9	7
Peas	23	5	15	19	2

Linear heatmap

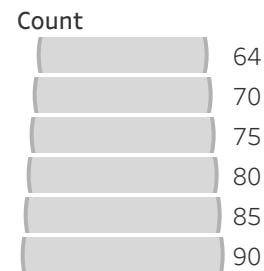
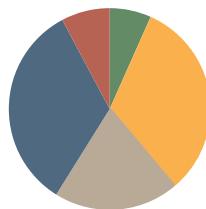


Table shapes

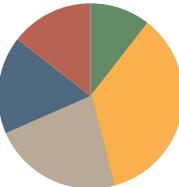


Pies

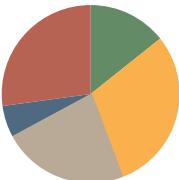
Corn



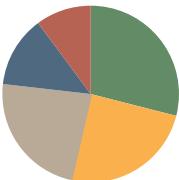
Squash



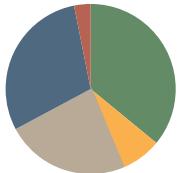
Brussel sprouts



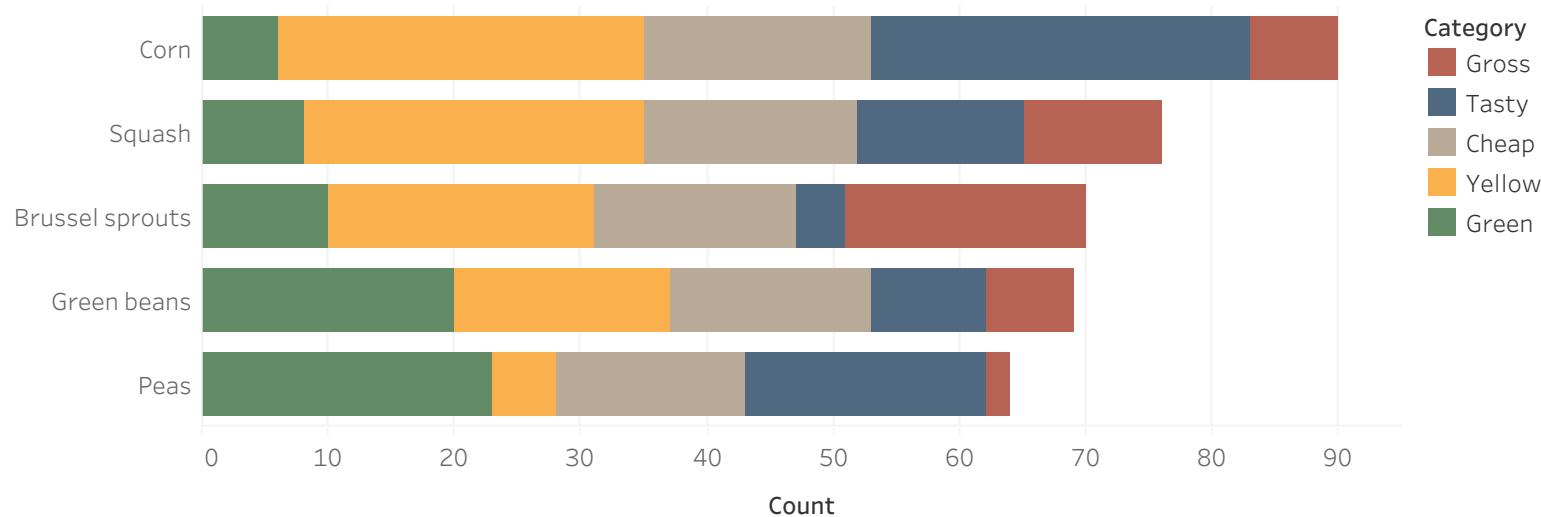
Green beans



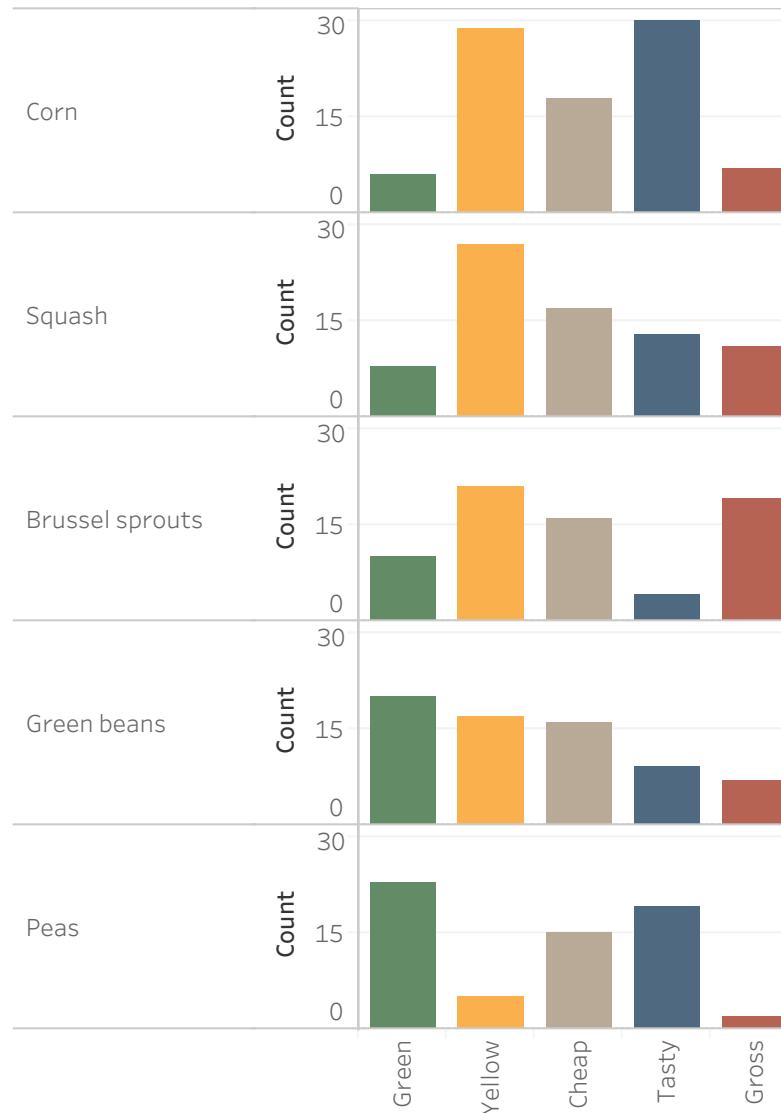
Peas



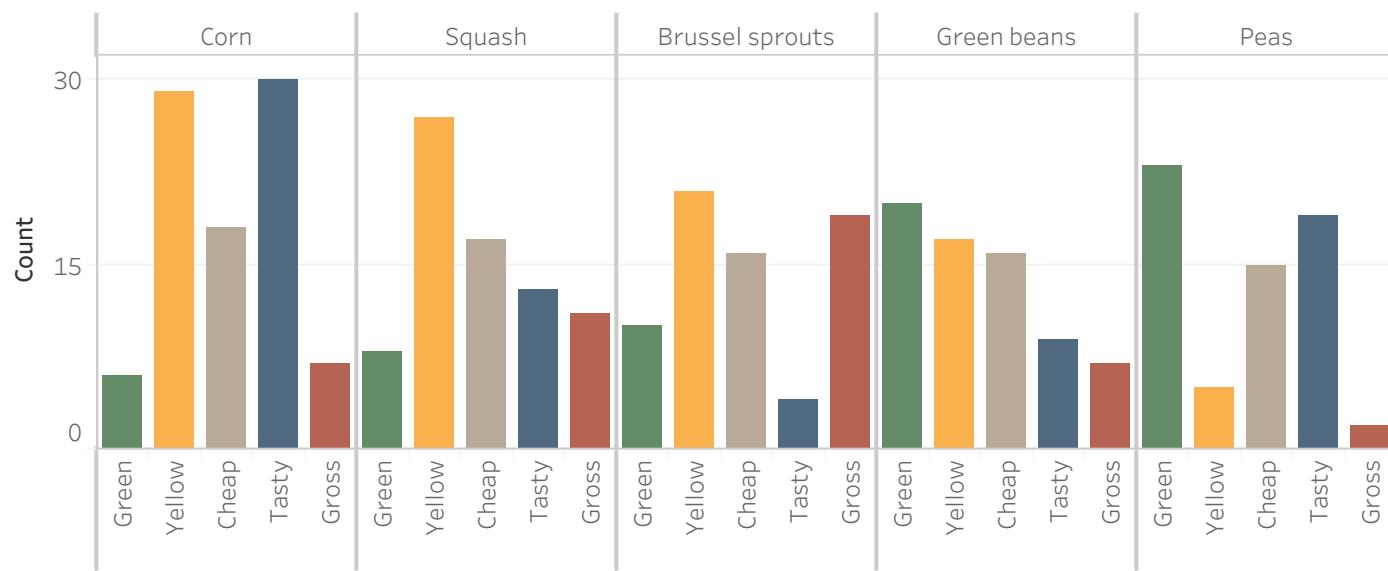
Stacked bars



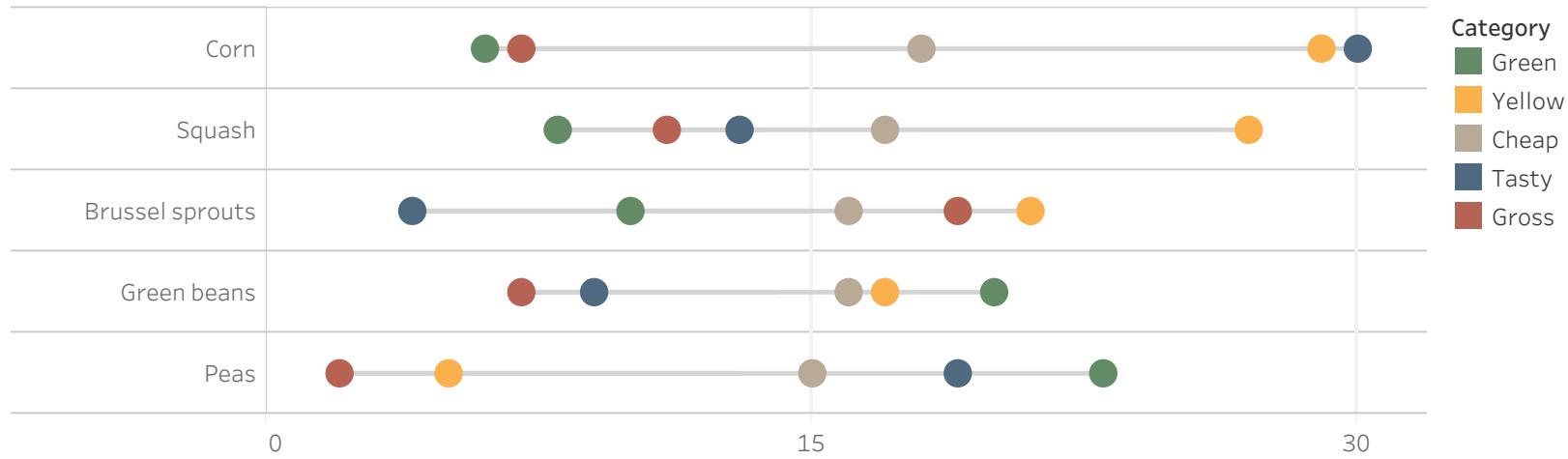
Within item bars



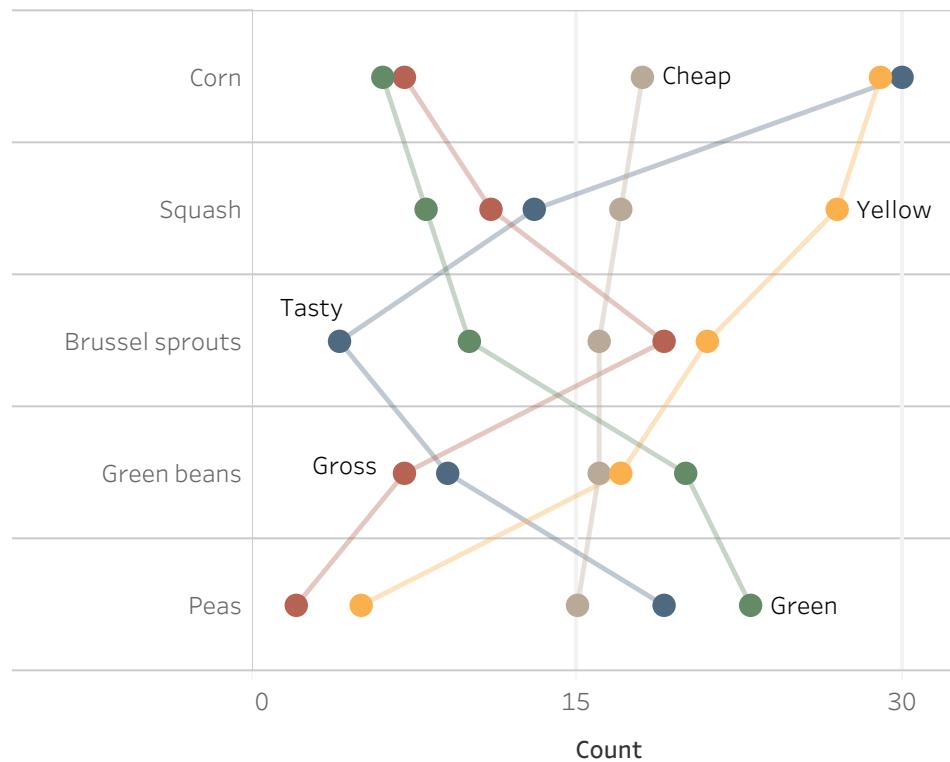
Item grouped bars



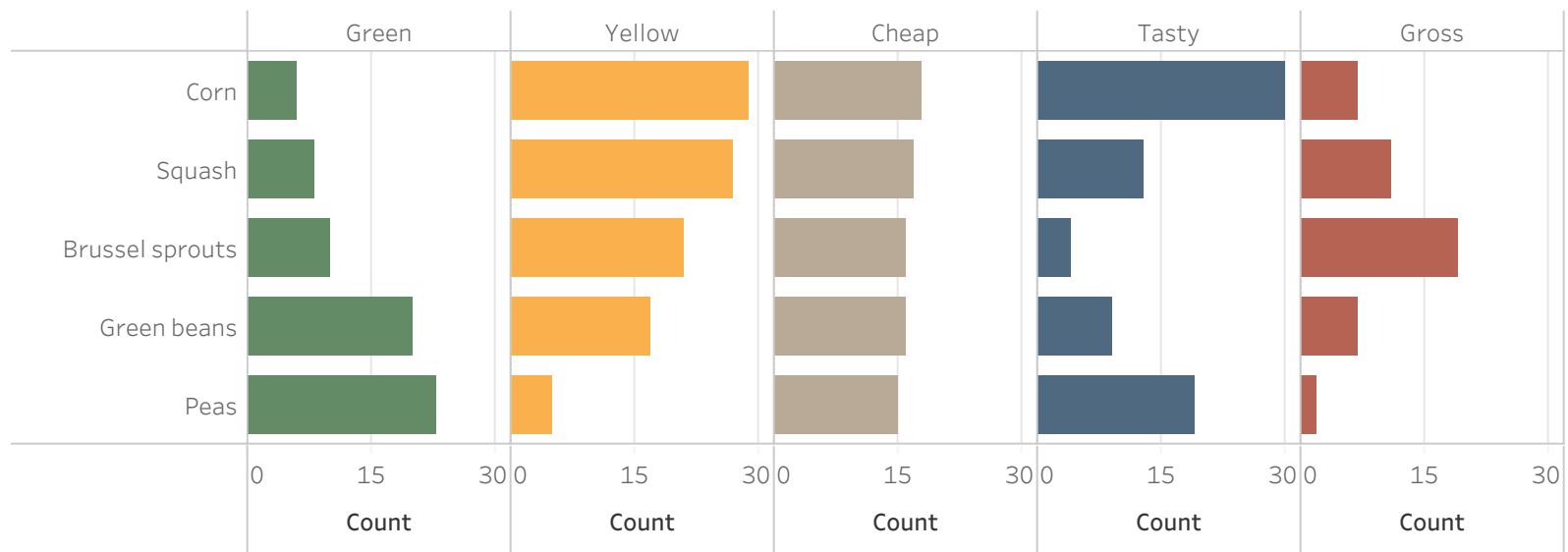
Dot plot



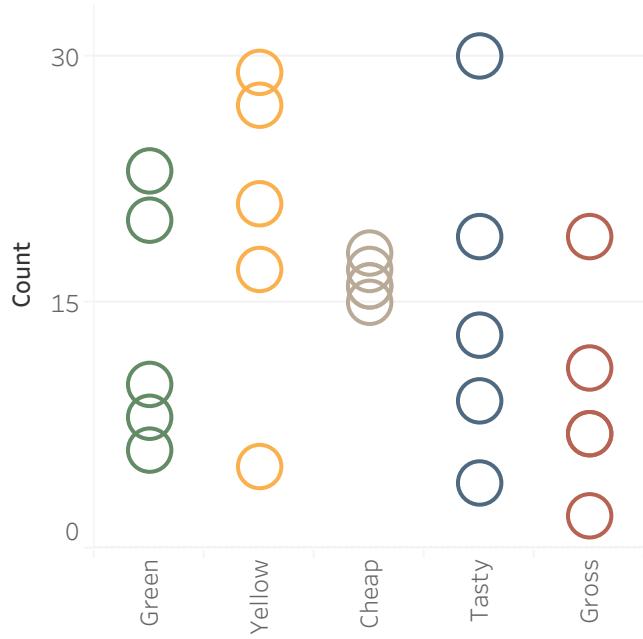
Dot plot with lines



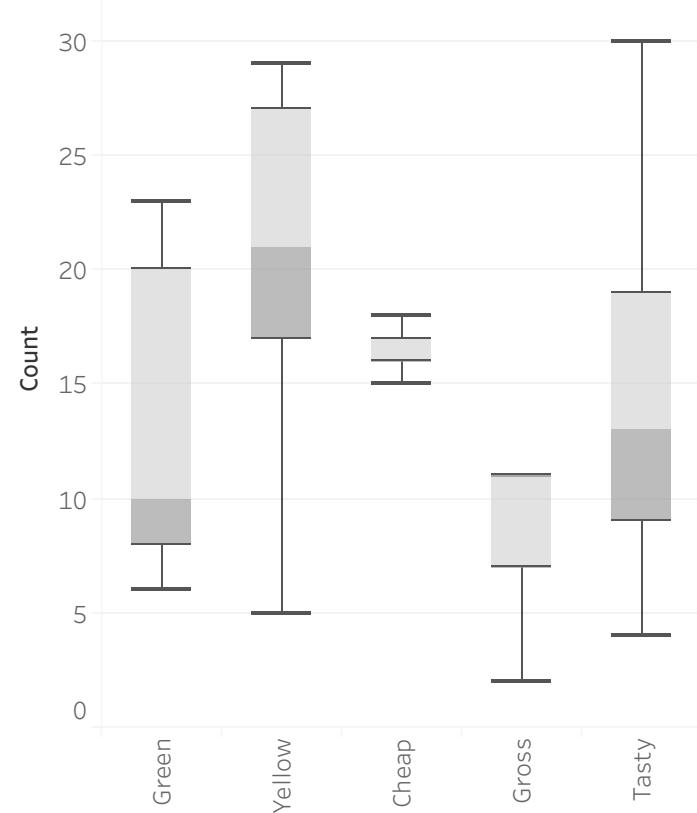
Within category bars



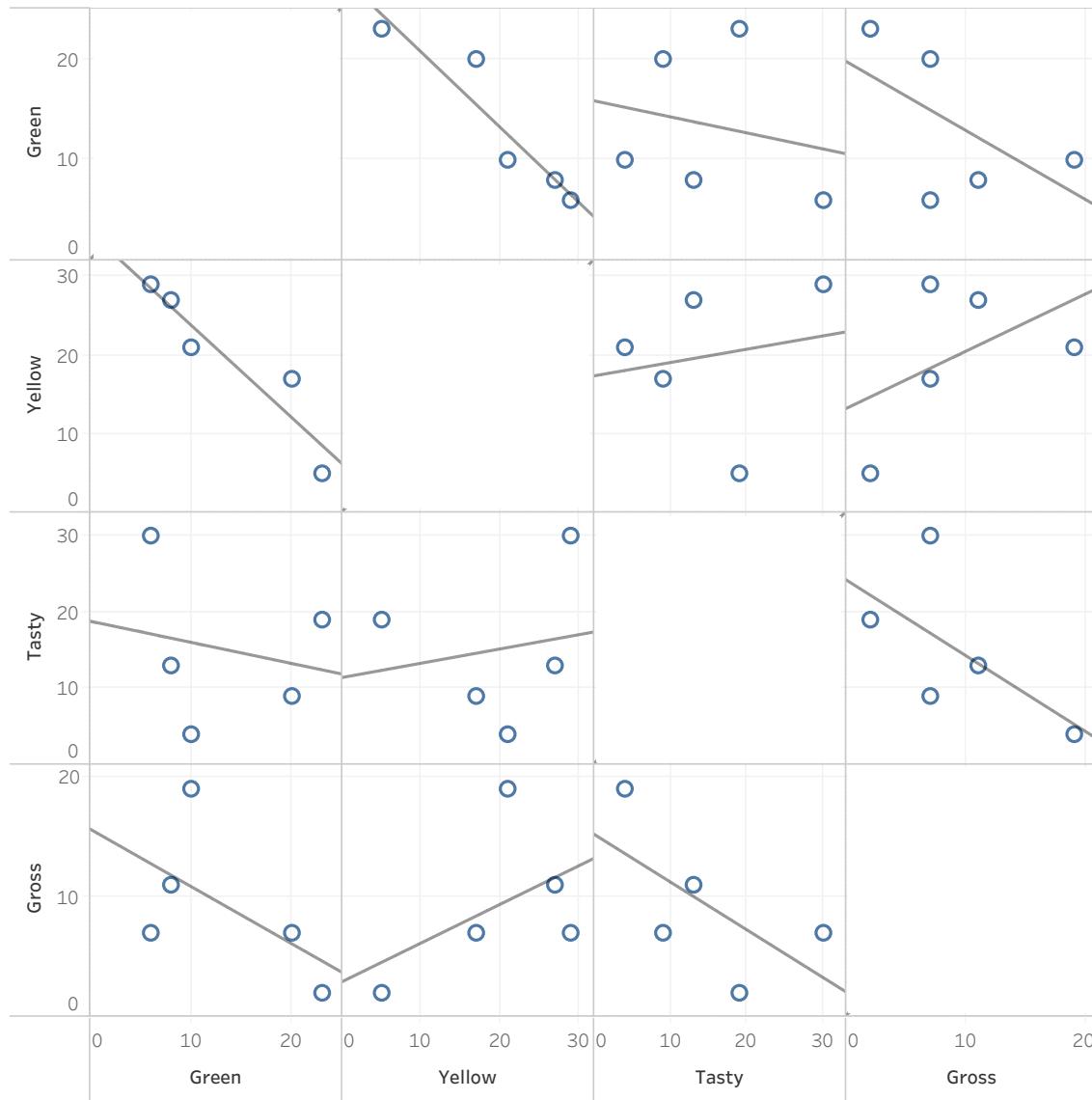
Distribution by category



Box plot by category

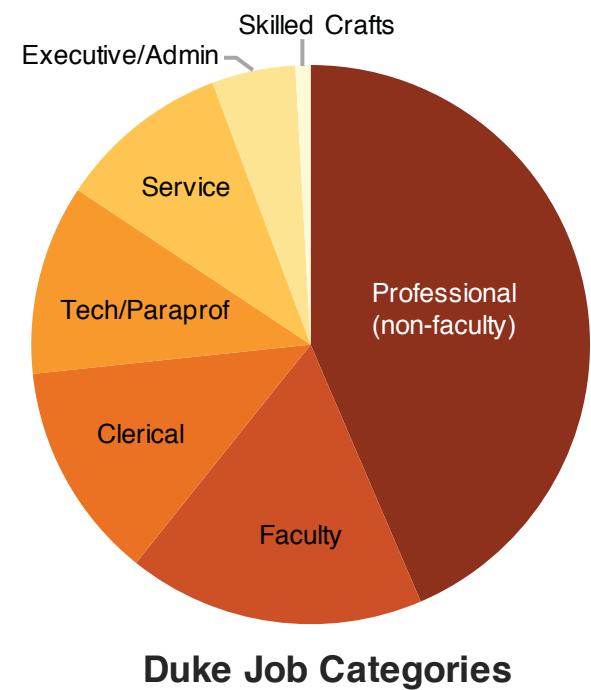
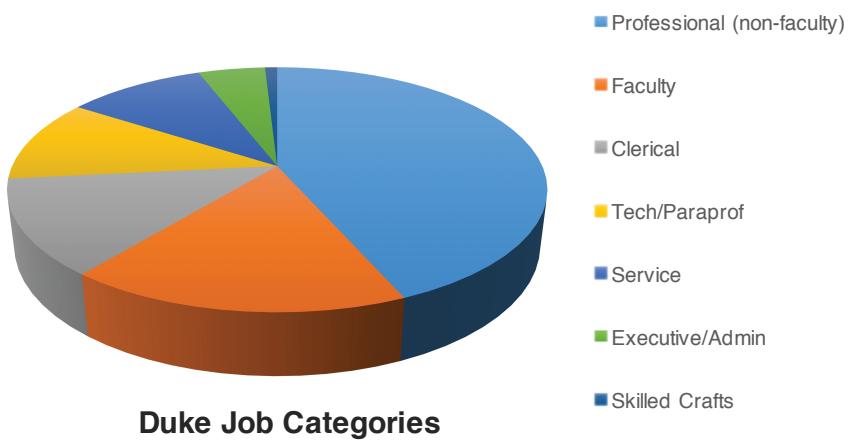


Pairs plot

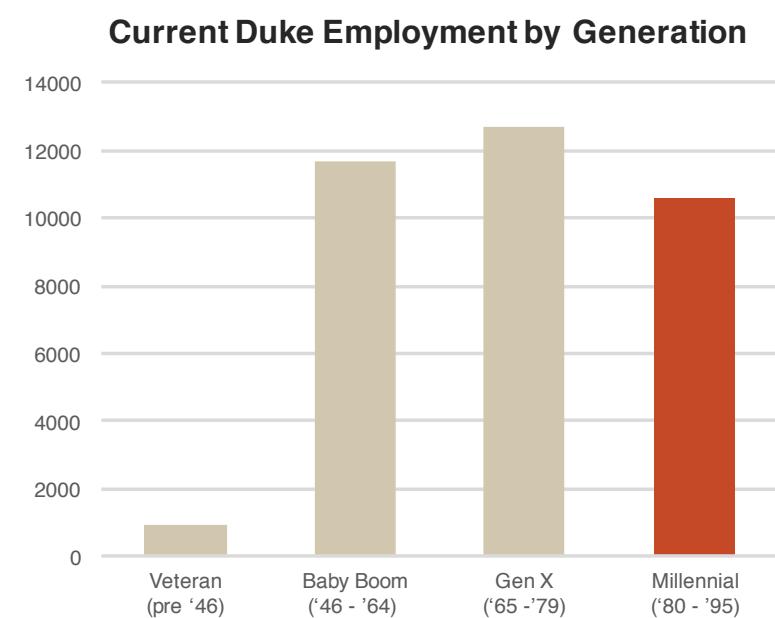
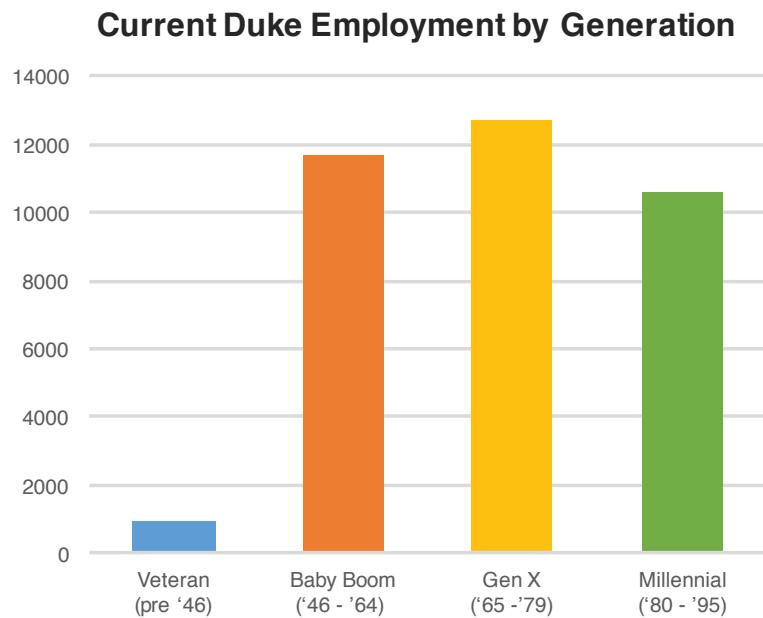


Three tips for designing effective visualizations

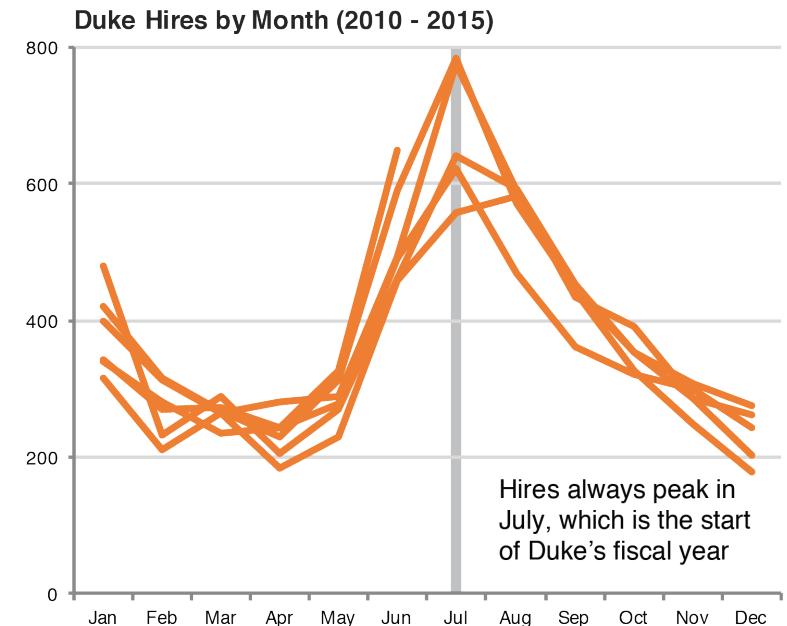
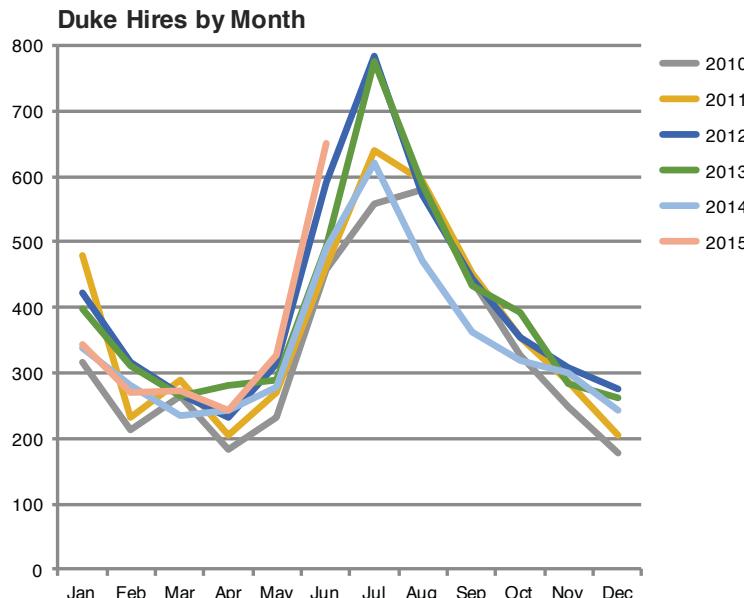
Keep it simple



Use color to draw attention

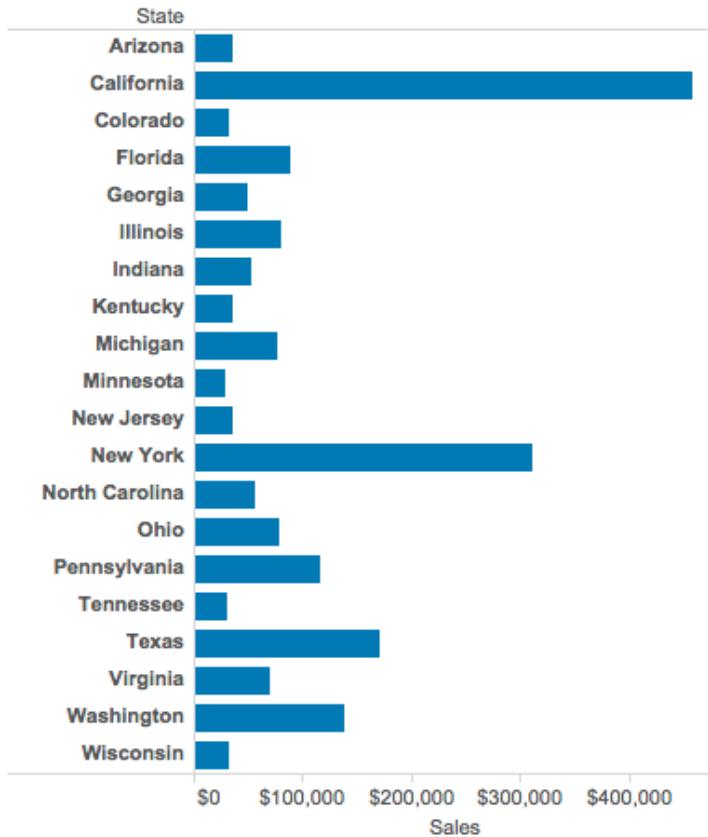


Every figure should have a purpose: tell a story or make an argument

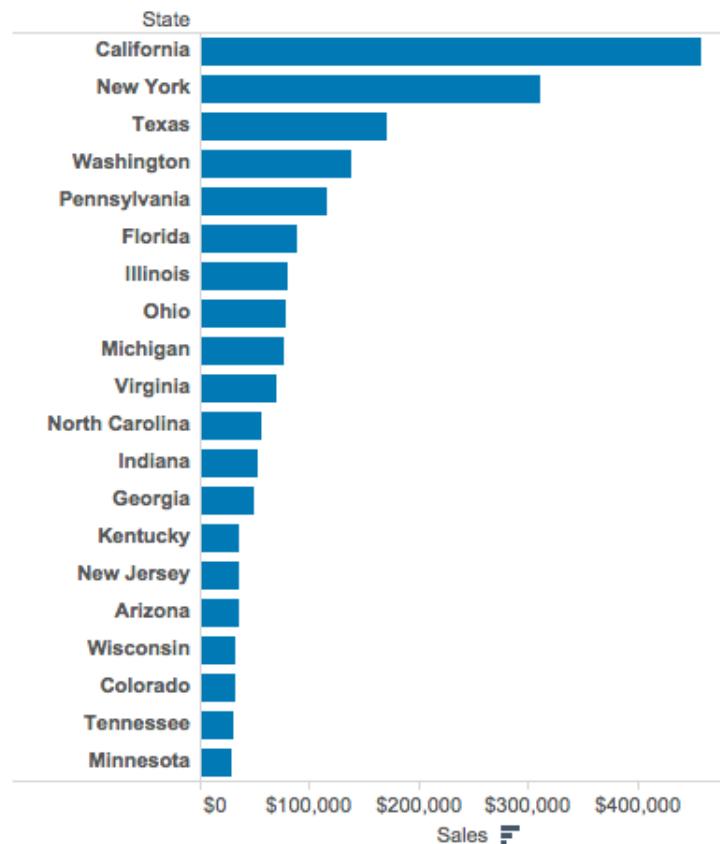
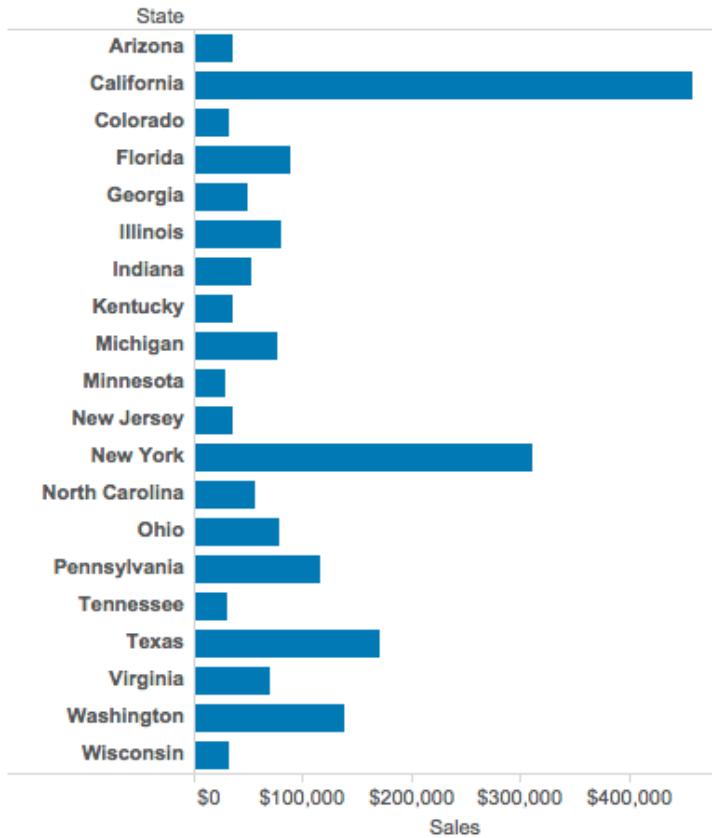


Common missteps

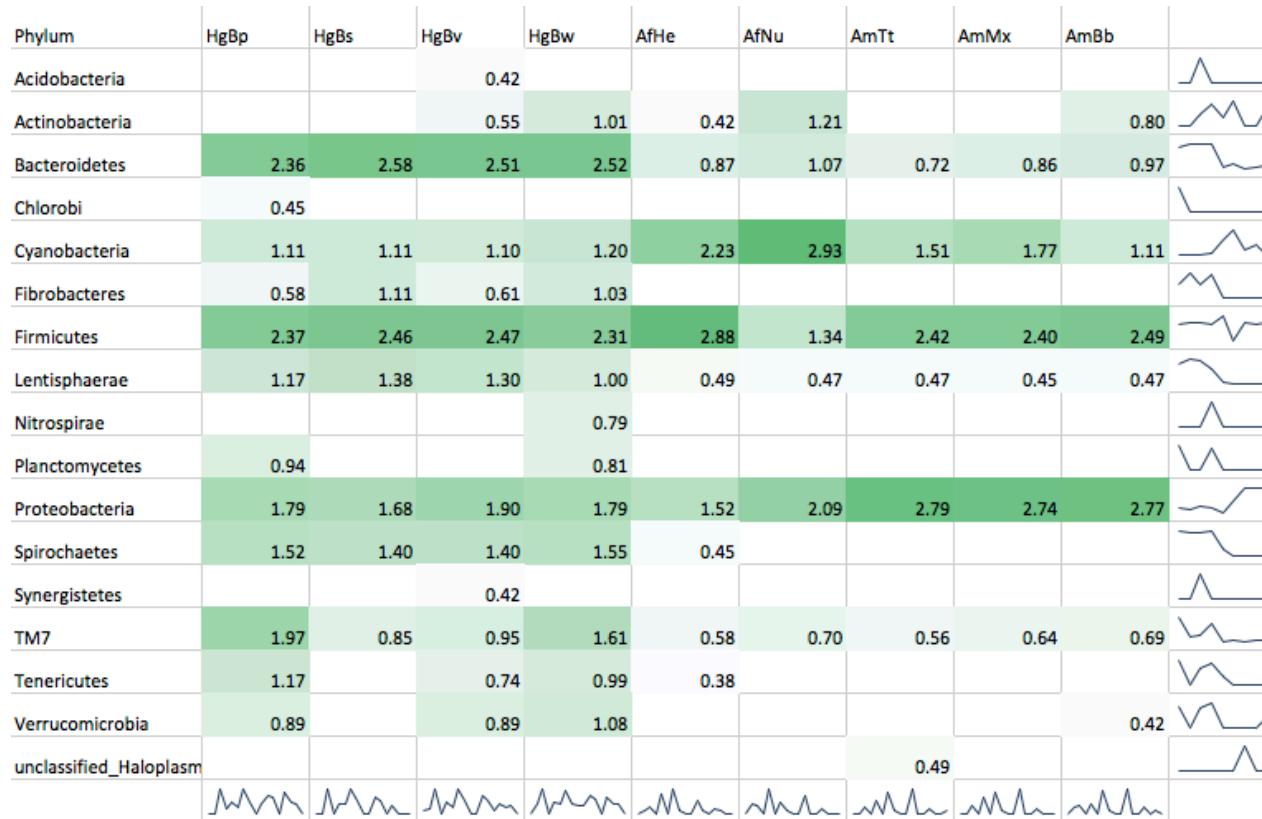
Default ordering hides patterns



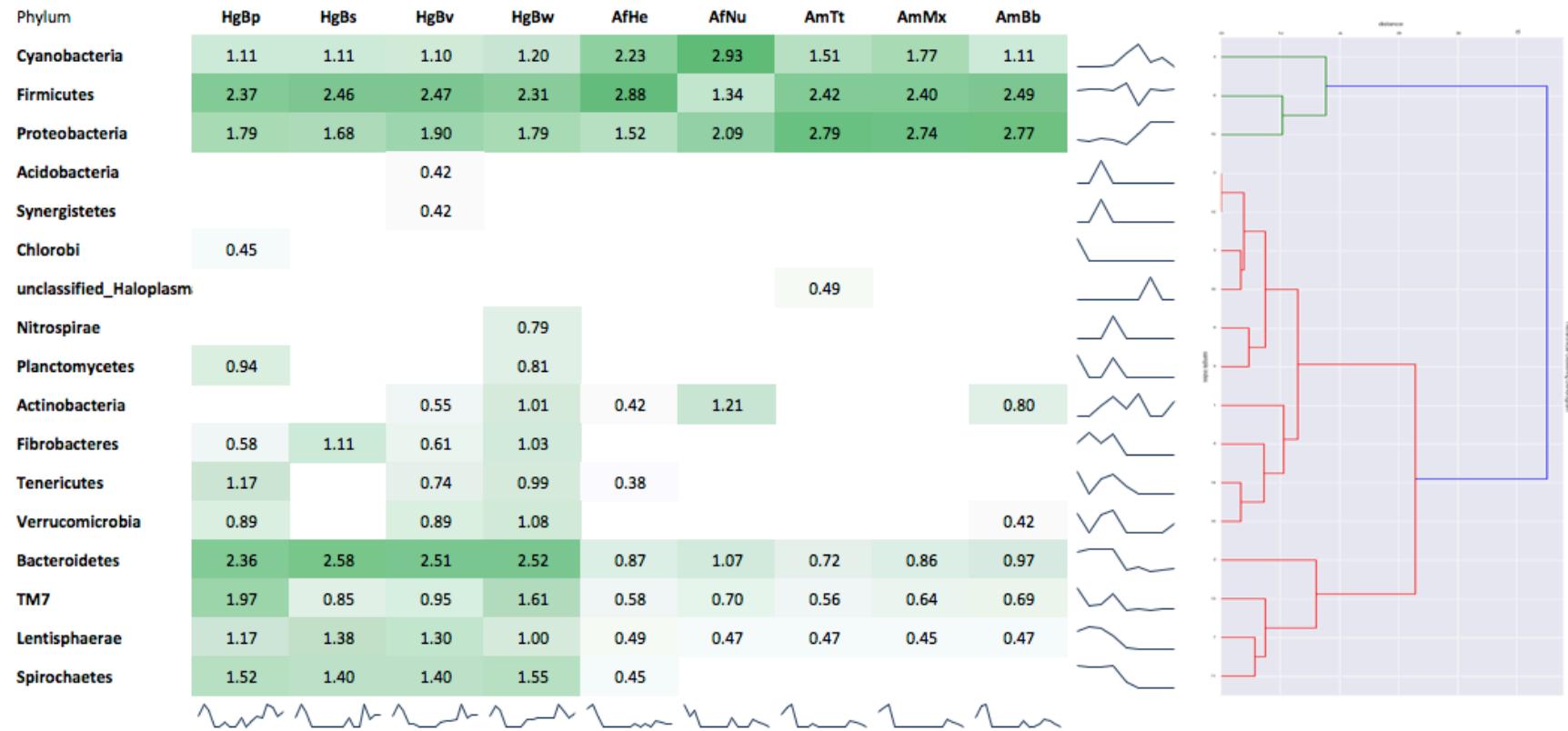
Sorting reveals patterns



Alphabetical again hides patterns



Clustering to see response groups



Tables easily hide patterns

The study group (to compare with the control group)	Cancer and age group	Type of analysis	White females	AA females
Females living in communities located in zip codes with kind of bad contions	Uterine cancer, all ages	Univariate	1.09, p=0.1027	
		Multivariate	1.24, p=0.0002	
	Uterine cancer, age 65+	Univariate	1.05, p=0.5702	1.05, p=0.6438
		Multivariate	1.18, p=0.0723	1.18, p=0.1789
	Cancer of corpus uteri, age 65+	Univariate	1.13, p=0.2861	1.07, p=0.6547
		Multivariate	1.27, p=0.0513	1.07, p=0.6763
	Uterine cancer, all ages	Univariate	1.29, p=0.0665	1.14, p=0.3630
		Multivariate	1.57, p=0.0019	1.28, p=0.1440
	Uterine cancer, age 65+	Univariate	1.55, p=0.0061	1.16, p=0.3800
		Multivariate	1.94, p=7.3x10 ⁻⁵	1.27, p=0.2300
	Cancer of corpus uteri, all ages	Univariate	1.68, p=0.0022	1.19, p=0.3820
		Multivariate	1.98, p=0.0001	1.09, p=0.7170
	Cancer of corpus uteri, age 65+	Univariate	1.91, p=0.0011	1.20, p=0.4330
		Multivariate	2.41, p=3.01x10 ⁻⁵	1.07, p=0.8120

Highlight patterns for the reader

**The study group
(to compare with
the control group)**

	Cancer	Age group	Analysis type	White females	p-value	AA females	p-value
Females living in communities located in zip codes with kind of bad conditions	Uterine	all ages	Univariate	1.09	0.1027	1.09	0.1027
			Multivariate	1.24		0.0002	0.0002
	Uterine	65+	Univariate	1.05	0.5702	1.05	0.6438
			Multivariate	1.18		1.18	
	Corpus uteri	65+	Univariate	1.13	0.2861	1.07	0.6547
			Multivariate	1.27		1.07	
Females living in communities with lots of really bad stuff	Uterine	all ages	Univariate	1.29	0.0665	1.14	0.3630
			Multivariate	1.57		0.0019	0.1440
	Uterine	65+	Univariate	1.55	7.30E-05	1.28	0.3800
			Multivariate	1.94		1.16	
	Corpus uteri	all ages	Univariate	1.68	0.0022	1.27	0.2300
			Multivariate	1.98		1.19	
	Corpus uteri	65+	Univariate	1.91	0.0011	1.09	0.3820
			Multivariate	2.41		3.01E-05	0.7170

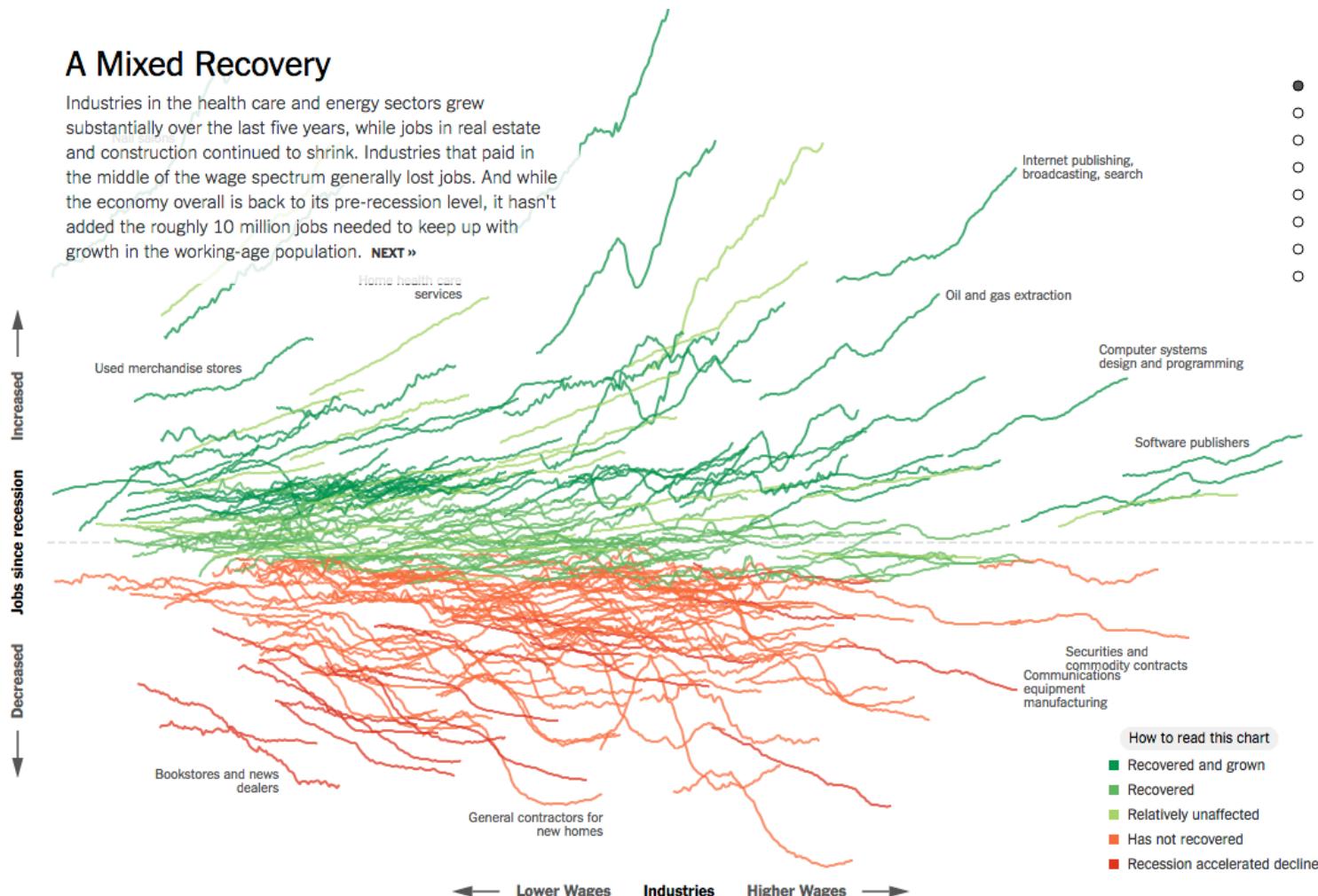
Highlight patterns for the reader

**The study group
(to compare with
the control group)**

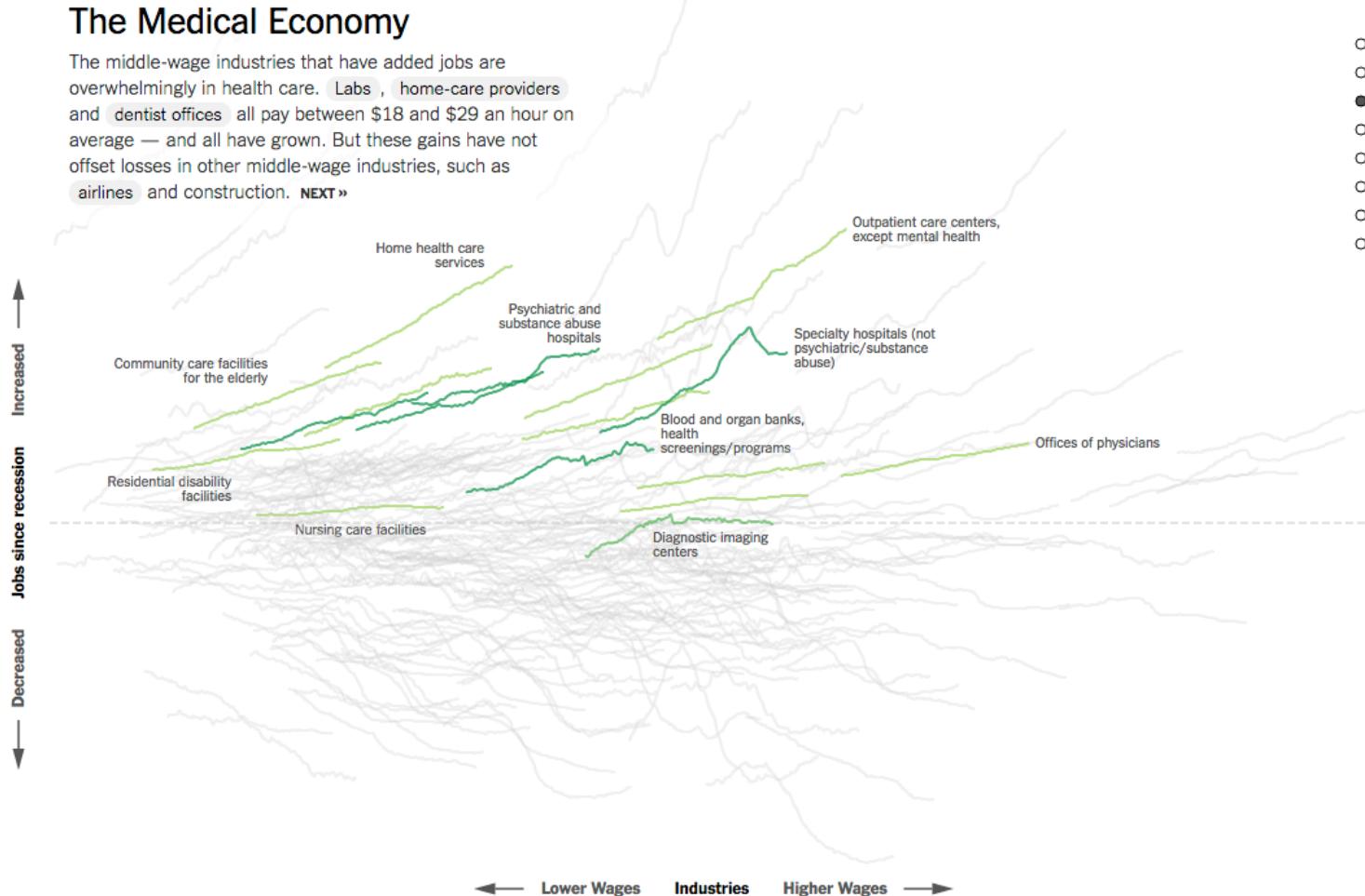
	Cancer	Age group	Analysis type	White females	p-value	AA females	p-value
Females living in communities located in zip codes with kind of bad conditions	Uterine	all ages	Univariate	1.09	0.1027	1.09	0.1027
			Multivariate	1.24	0.0002	1.24	0.0002
	Uterine	65+	Univariate	1.05	0.5702	1.05	0.6438
			Multivariate	1.18	0.0723	1.18	0.1789
	Corpus uteri	65+	Univariate	1.13	0.2861	1.07	0.6547
			Multivariate	1.27	0.0513	1.07	0.6763
	Uterine	all ages	Univariate	1.29	0.0665	1.14	0.3630
			Multivariate	1.57	0.0019	1.28	0.1440
	Uterine	65+	Univariate	1.55	0.0061	1.16	0.3800
			Multivariate	1.94	7.30E-05	1.27	0.2300
	Corpus uteri	all ages	Univariate	1.68	0.0022	1.19	0.3820
			Multivariate	1.98	0.0001	1.09	0.7170
	Corpus uteri	65+	Univariate	1.91	0.0011	1.2	0.4330
			Multivariate	2.41	3.01E-05	1.07	0.8120

Significantly impacted residents

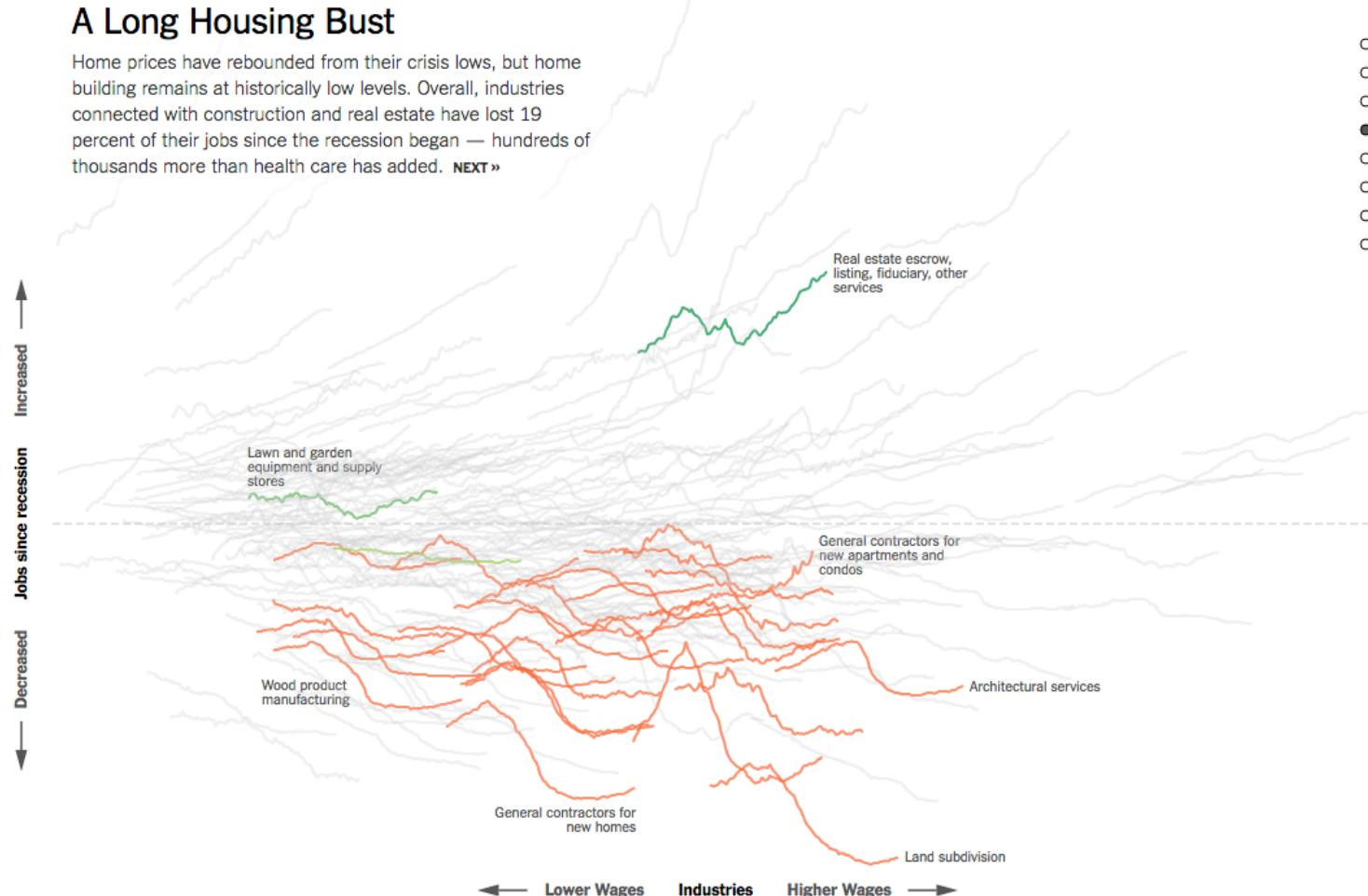
All the data doesn't tell a story



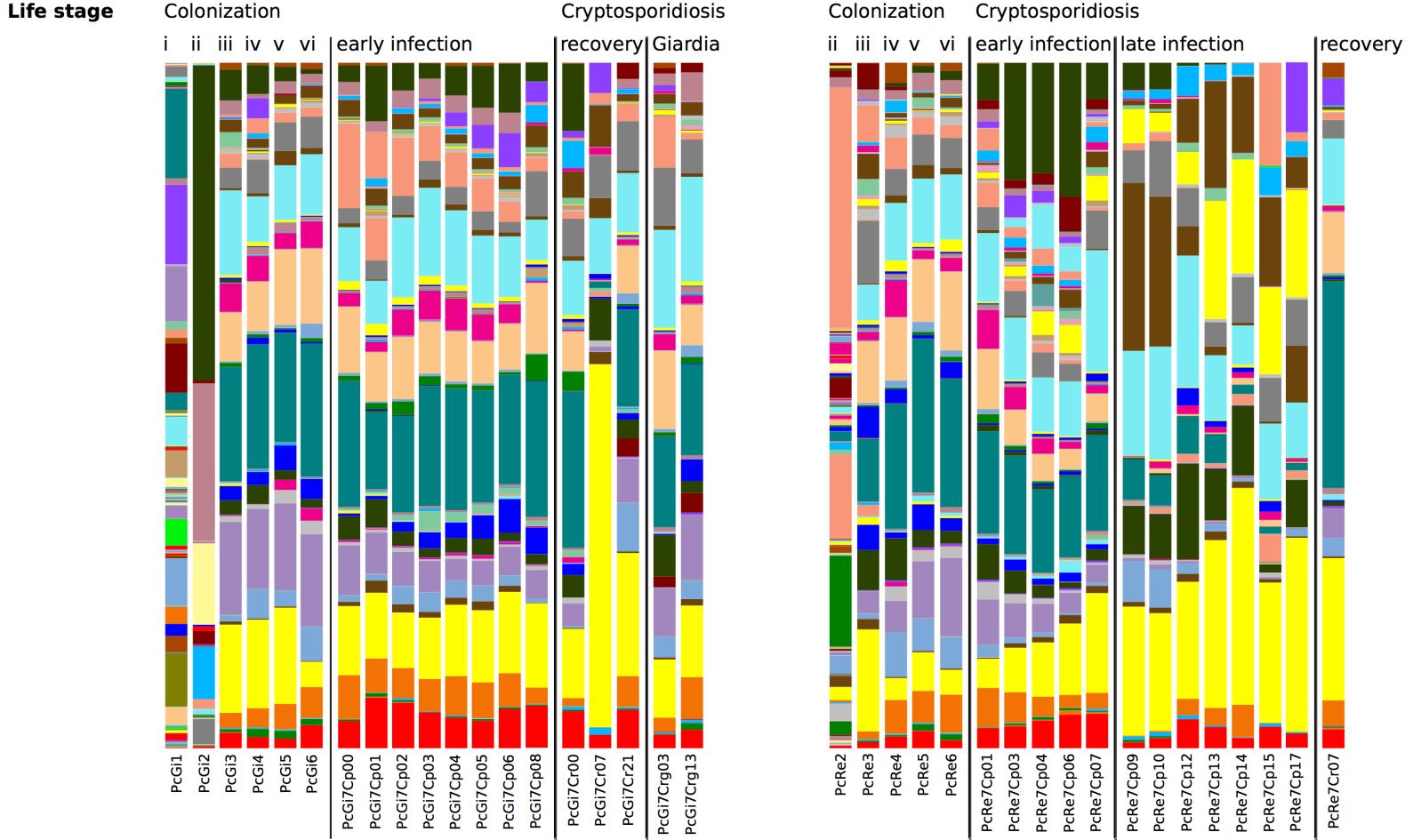
All the data doesn't tell a story



All the data doesn't tell a story

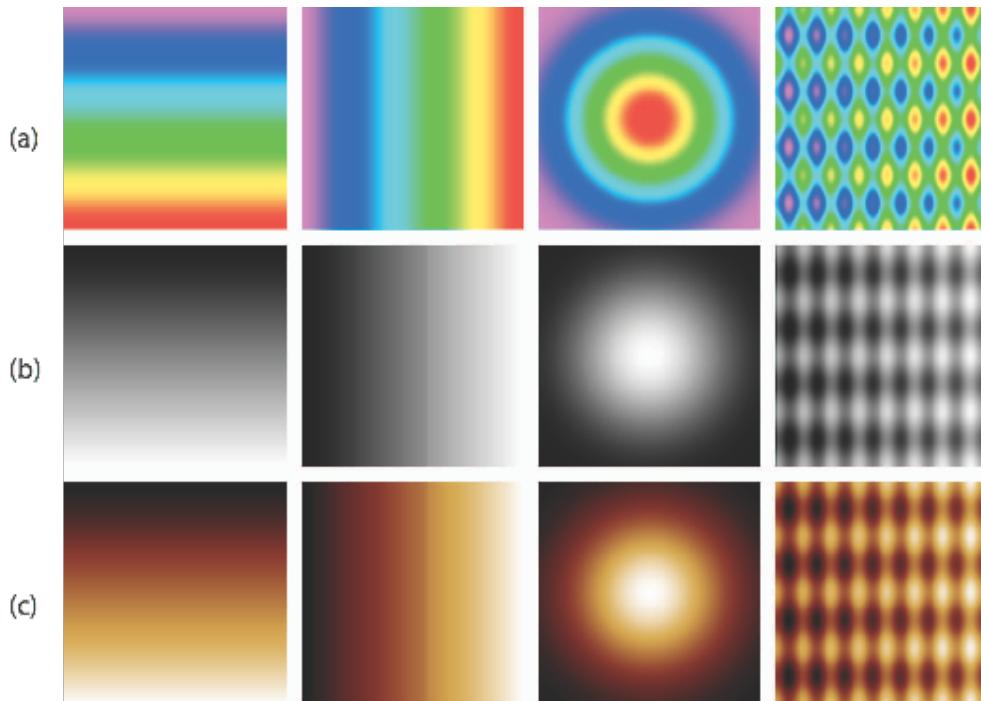


All the data doesn't tell a story



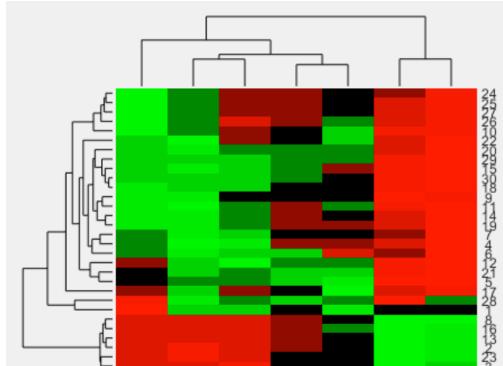
Color can be tricky

Rainbow colormaps distort

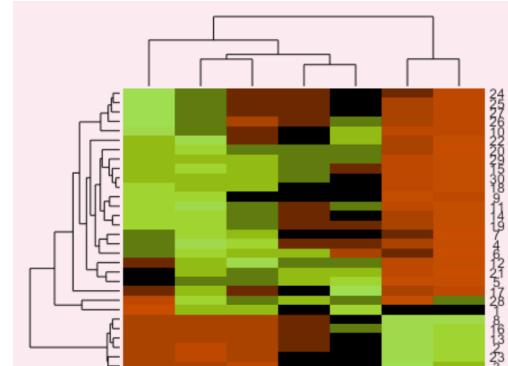


Borland, David, and Russell M. Taylor II. "Rainbow color map (still) considered harmful." *IEEE computer graphics and applications* 27.2 (2007).
<https://ieeexplore.ieee.org/document/4118486>

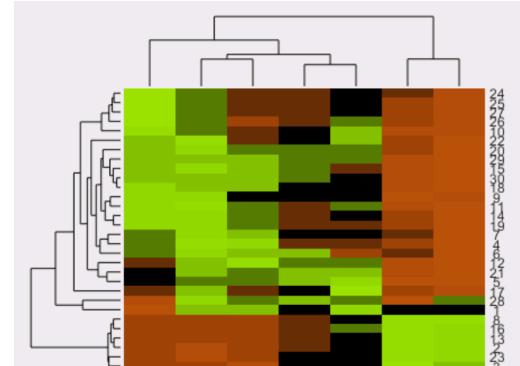
Red-green bad for color blindness



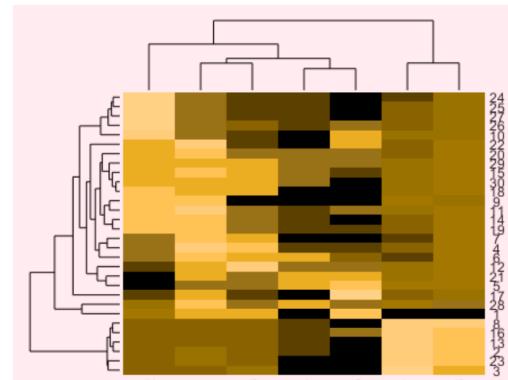
Normal



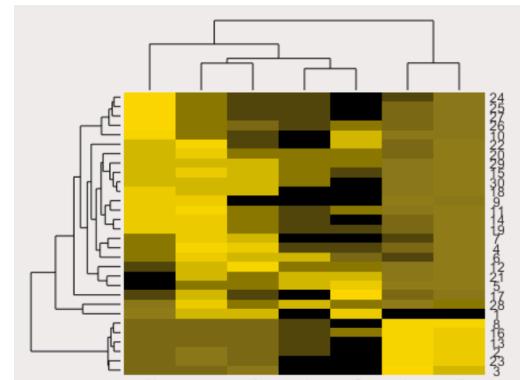
Green-weak/Deuteranomaly



Green-weak/Deuteranopia

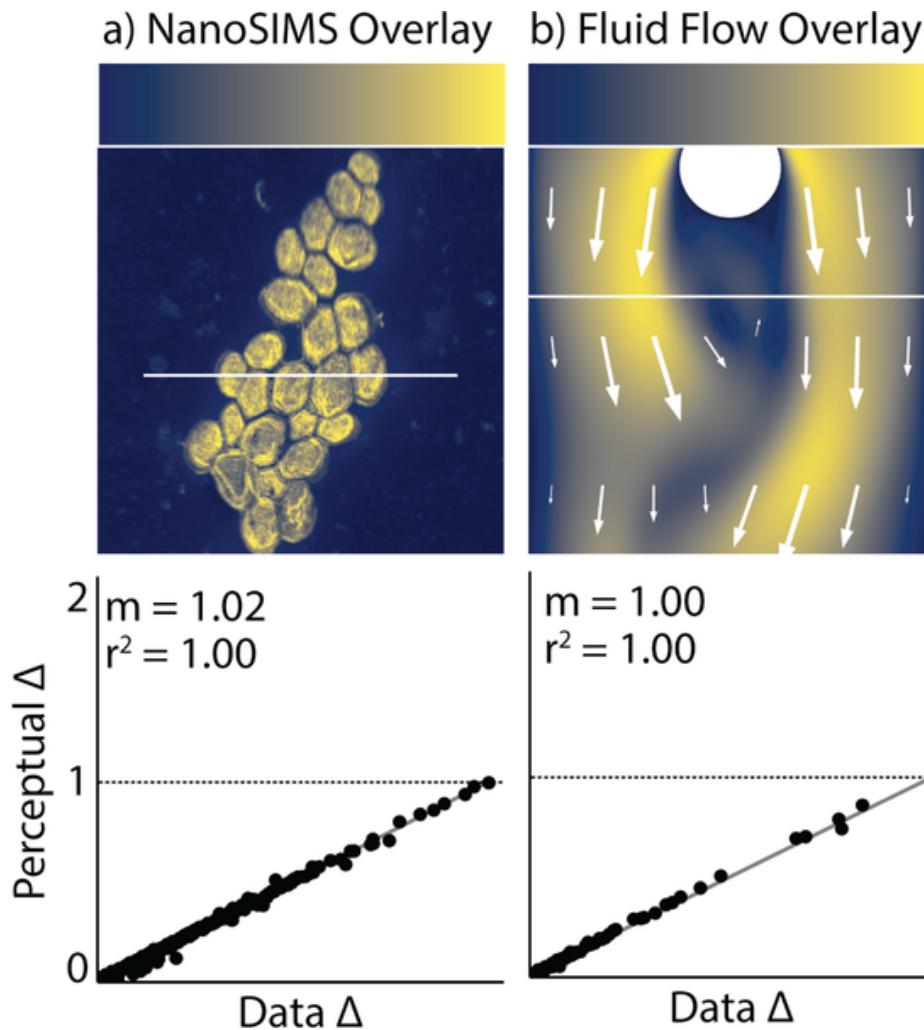


Red-weak/Protanomaly



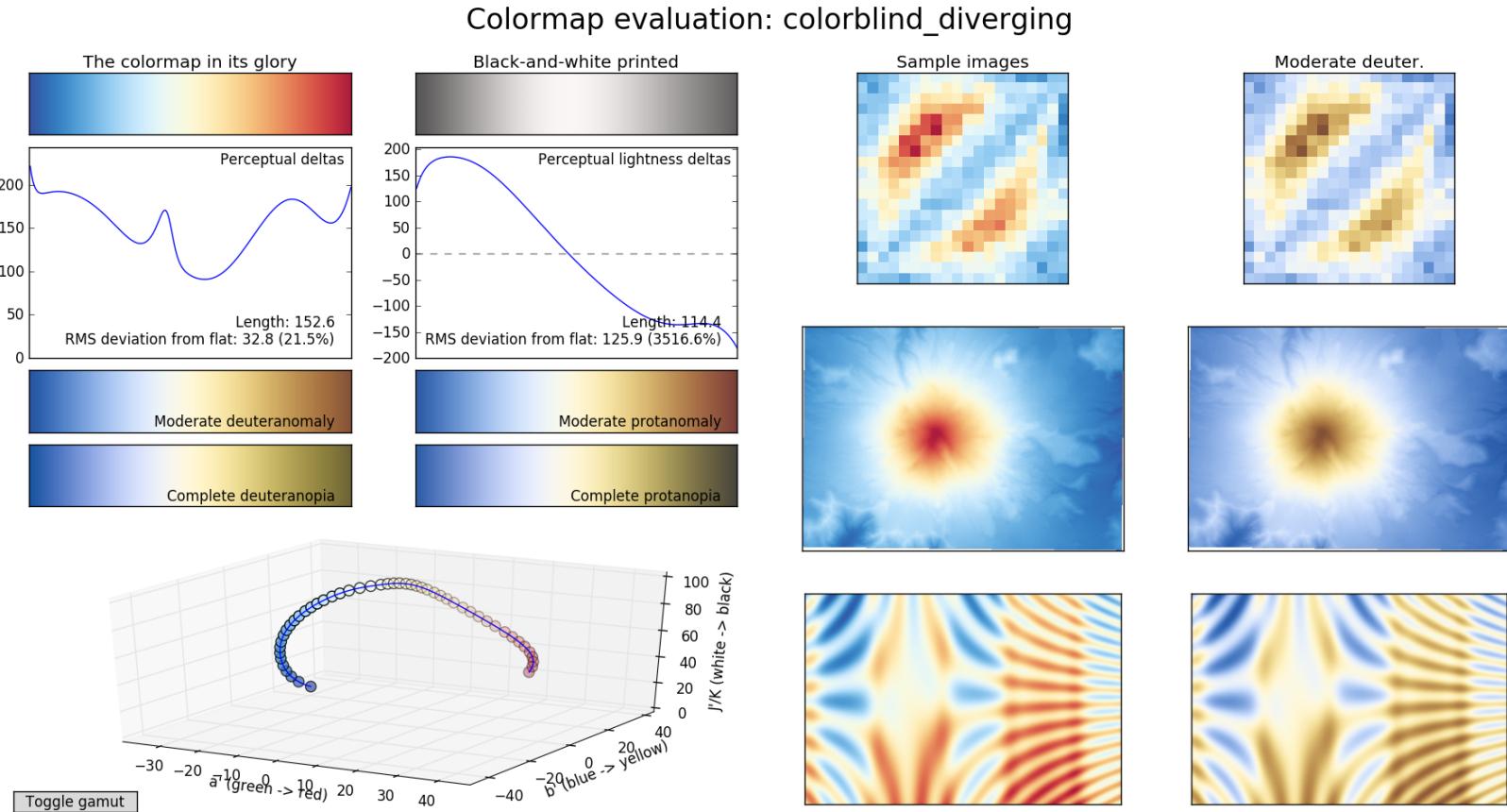
Red-weak/Protanopia

Fig 5. Our optimal colormap, cividis.



Nuñez JR, Anderton CR, Renslow RS (2018) Optimizing colormaps with consideration for color vision deficiency to enable accurate interpretation of scientific data. PLOS ONE 13(7): e0199239. <https://doi.org/10.1371/journal.pone.0199239>
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0199239>

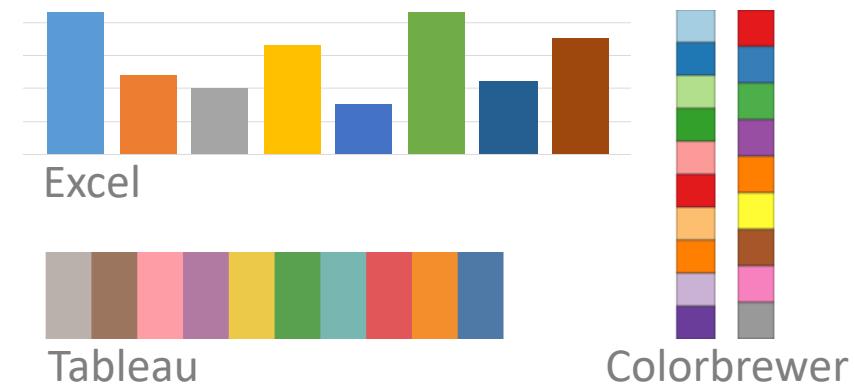
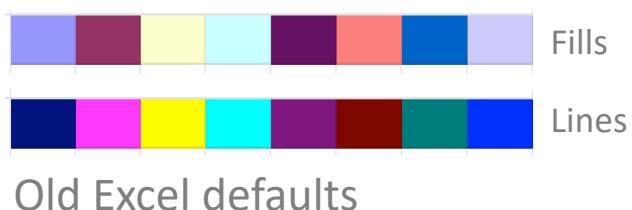
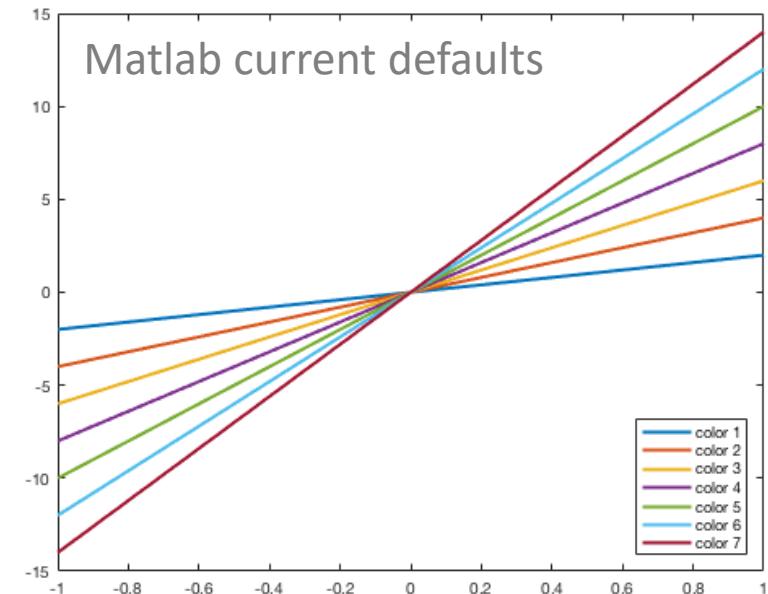
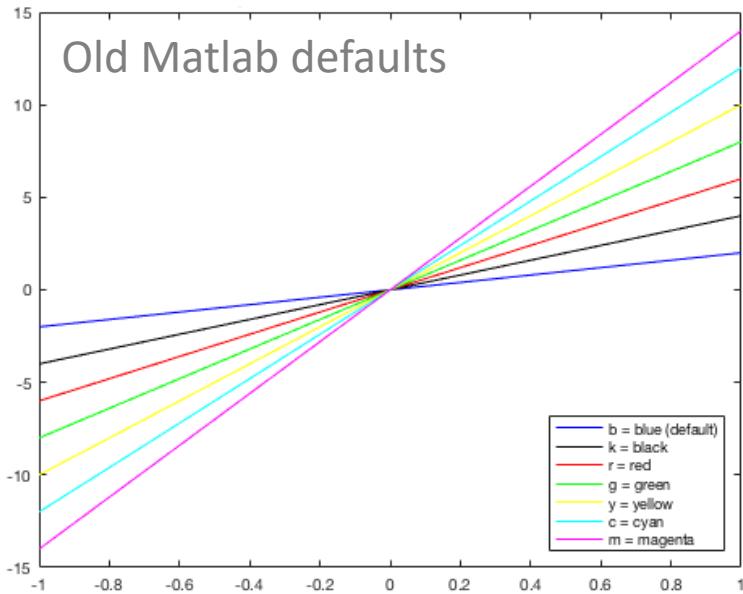
Red-x-Blue good for diverging



<https://github.com/matplotlib/matplotlib/issues/7081/>

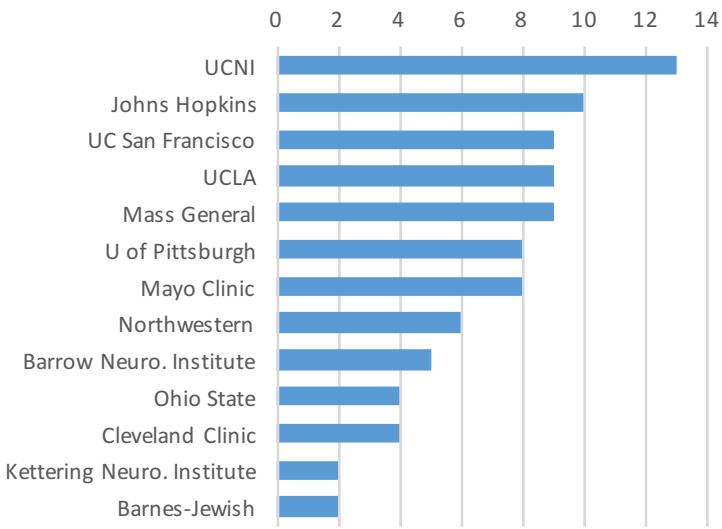
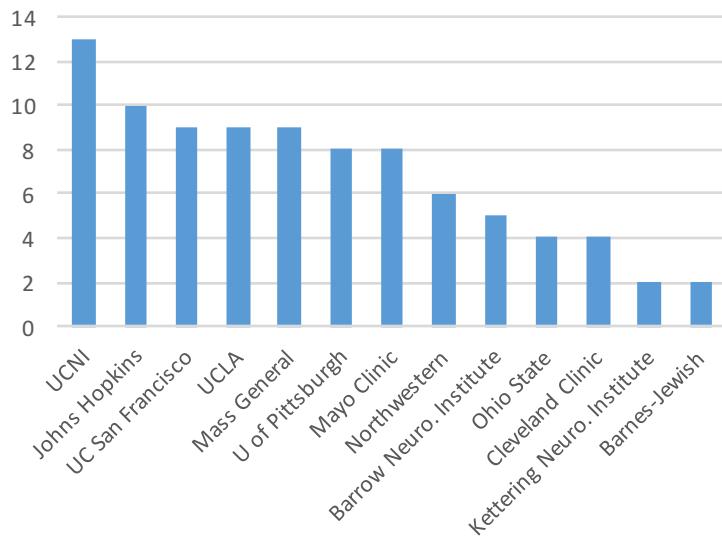
<https://github.com/matplotlib/viscm>

Avoid pure categorical colors



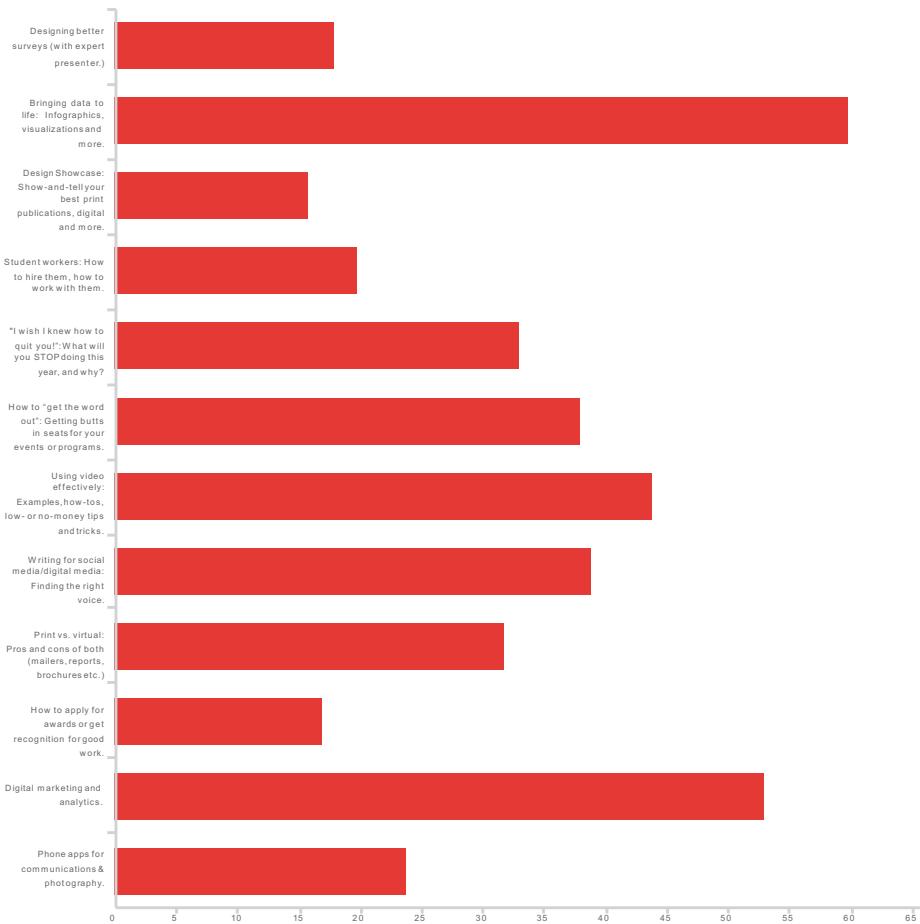
Minimal, readable text
to tell your story

Keep text horizontal

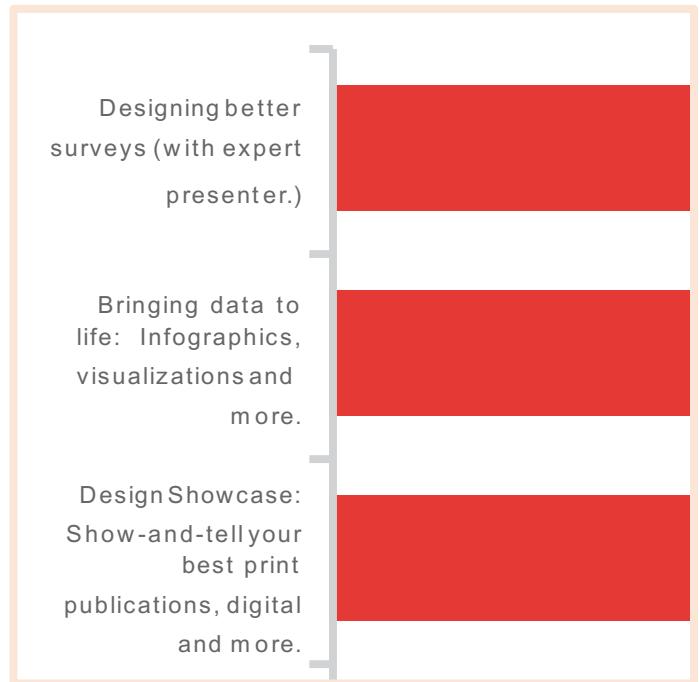


Default sizes may not be legible

Q2 - Please pick your top 5 favorite topics.

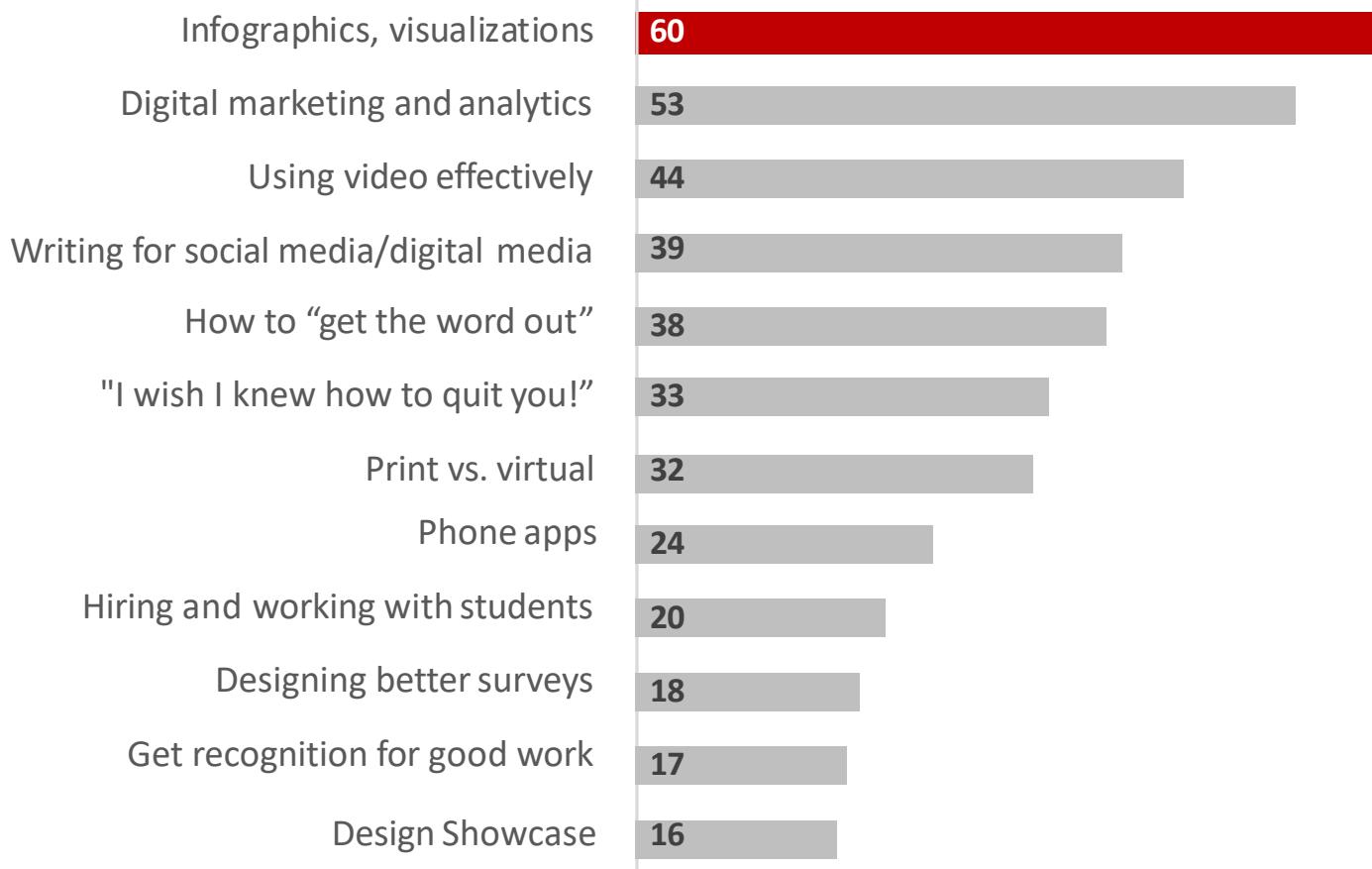


Default Qualtrics output



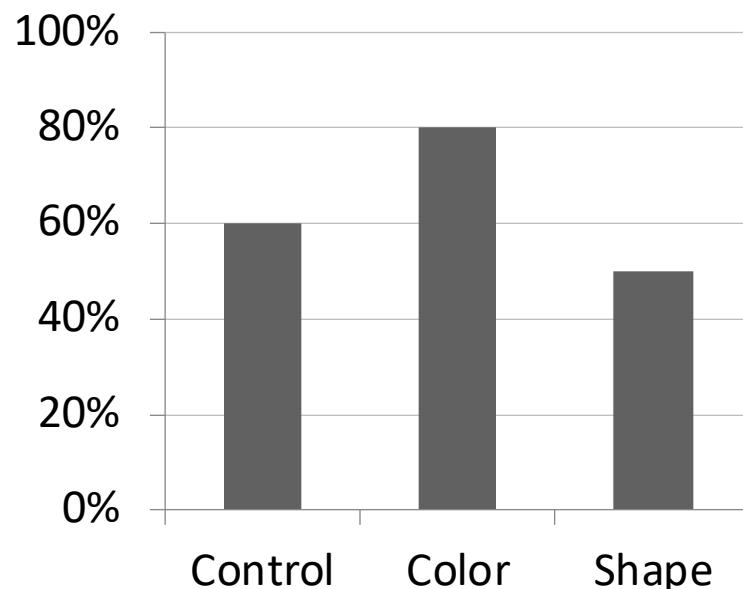
Summarize, sort & highlight

Please pick your top 5 favorite topics.

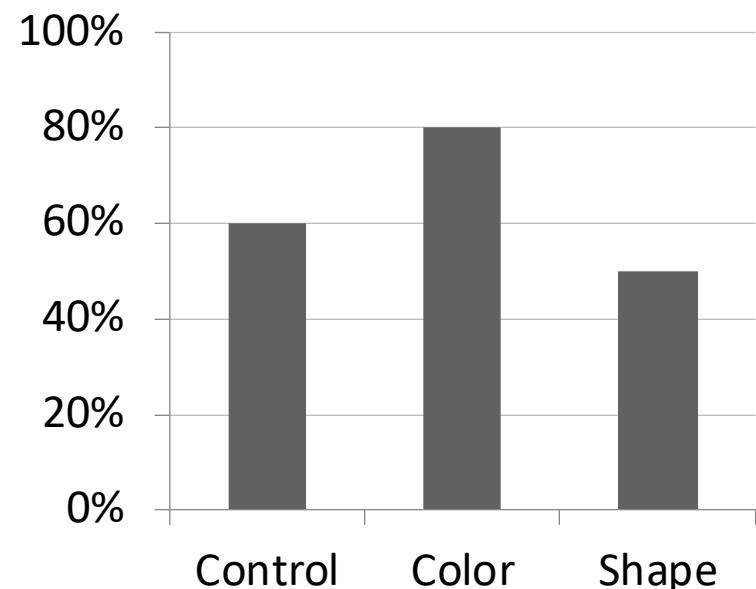


Active titles tell your story

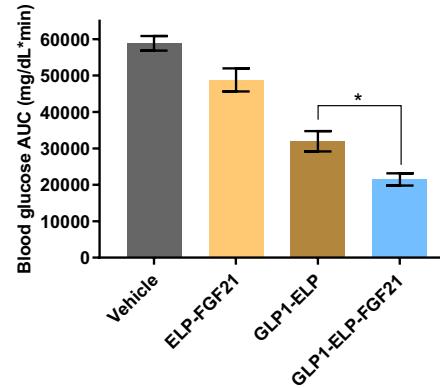
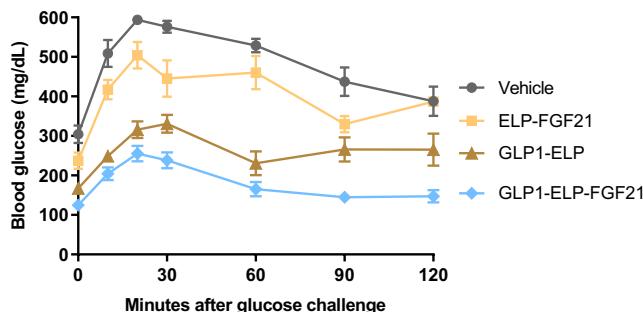
Accuracy versus Color and Shape



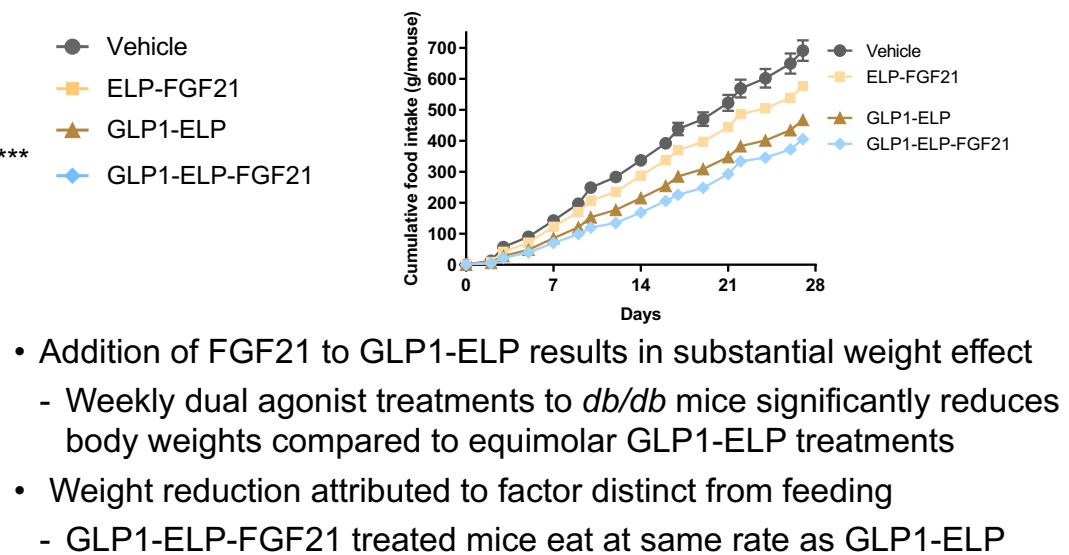
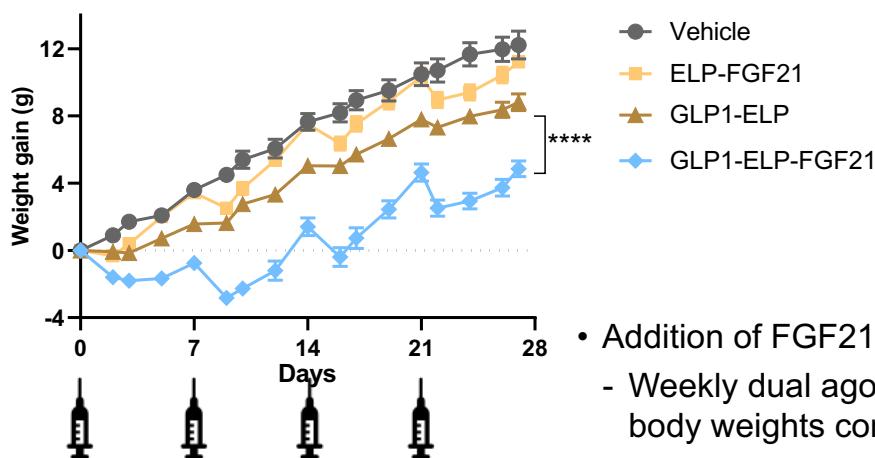
Accuracy Improved by Color, not Shape



Dual agonist outperforms GLP1 receptor agonist



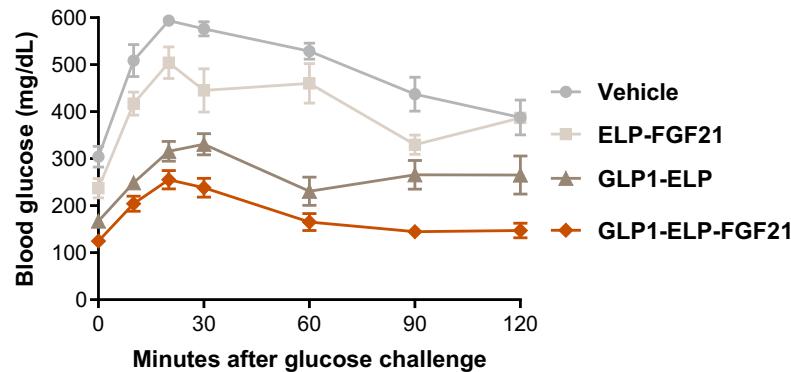
- GLP1-ELP-FGF21 treated mice display superior response to glucose challenge
 - Single treatment to *db/db* mice followed by fasted glucose bolus
 - Dual agonist group returns to baseline more quickly than equimolar dose of GLP1-ELP



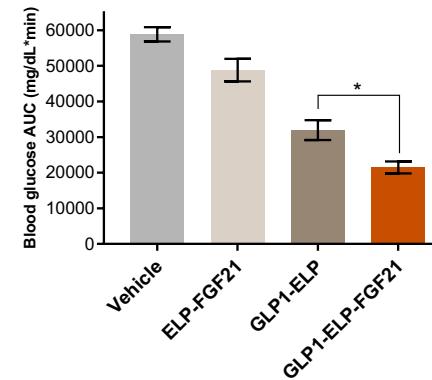
- Addition of FGF21 to GLP1-ELP results in substantial weight effect
 - Weekly dual agonist treatments to *db/db* mice significantly reduces body weights compared to equimolar GLP1-ELP treatments
- Weight reduction attributed to factor distinct from feeding
 - GLP1-ELP-FGF21 treated mice eat at same rate as GLP1-ELP

Dual agonist outperforms long-acting GLP-1 receptor agonist

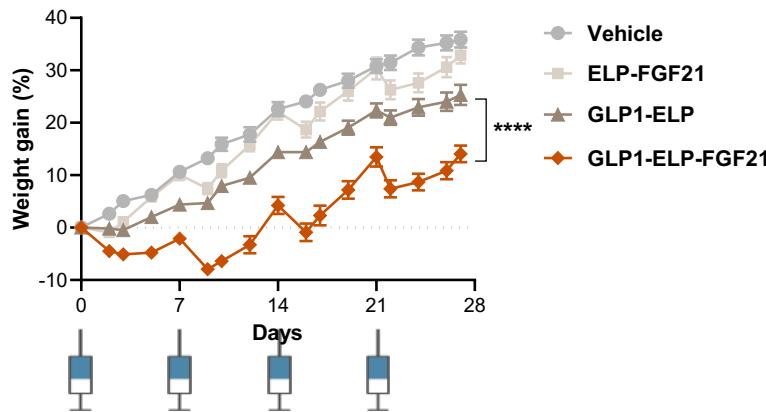
Hyperglycemic *db/db* mice challenged with a fasted glucose bolus



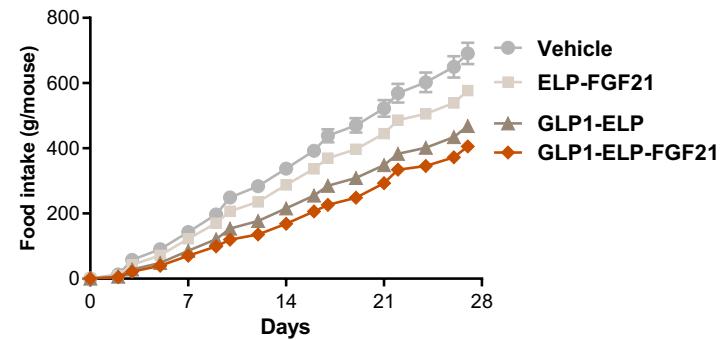
Dual agonist-treated group **responds to glucose spike more efficiently** than an equimolar dose of GLP1-ELP



Weekly dual agonist treatments to obese *db/db* mice results in **significantly lower body weights** compared to equimolar GLP1-ELP treatments



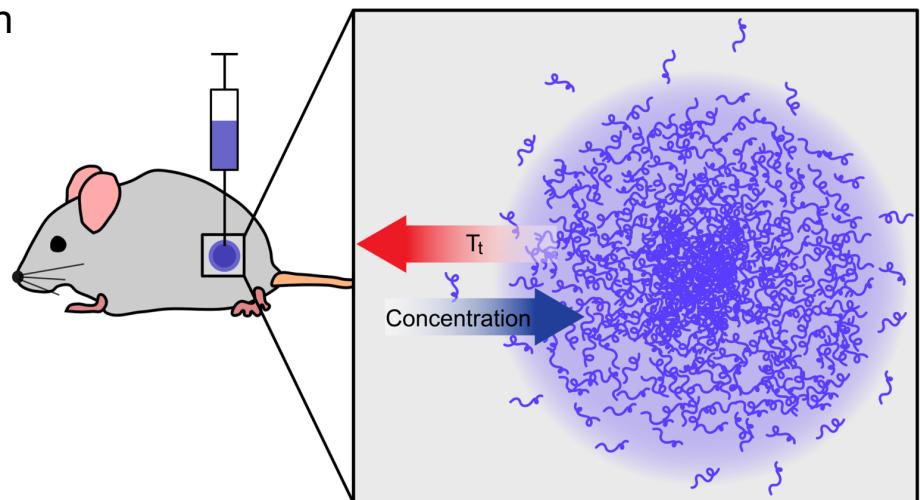
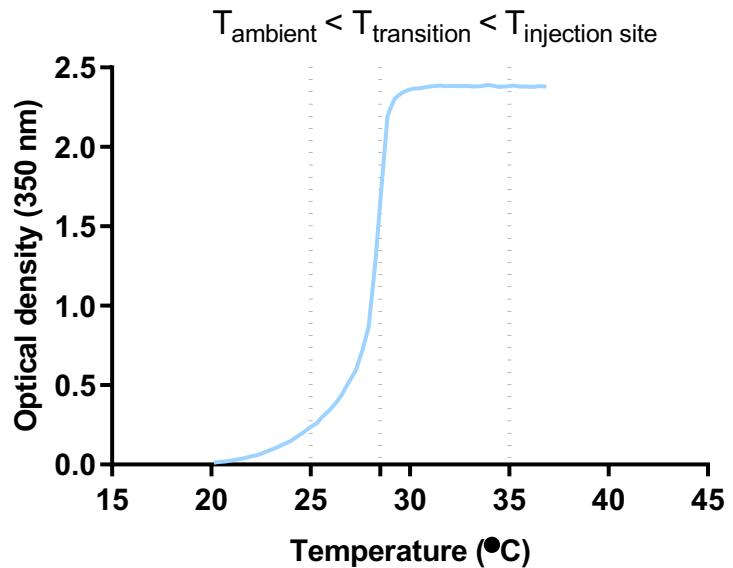
Dual agonist **inhibits weight gain without decreasing feed rate** compared to GLP1-ELP → altered energy balance



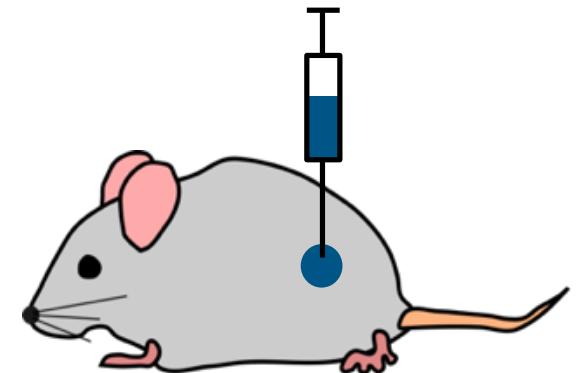
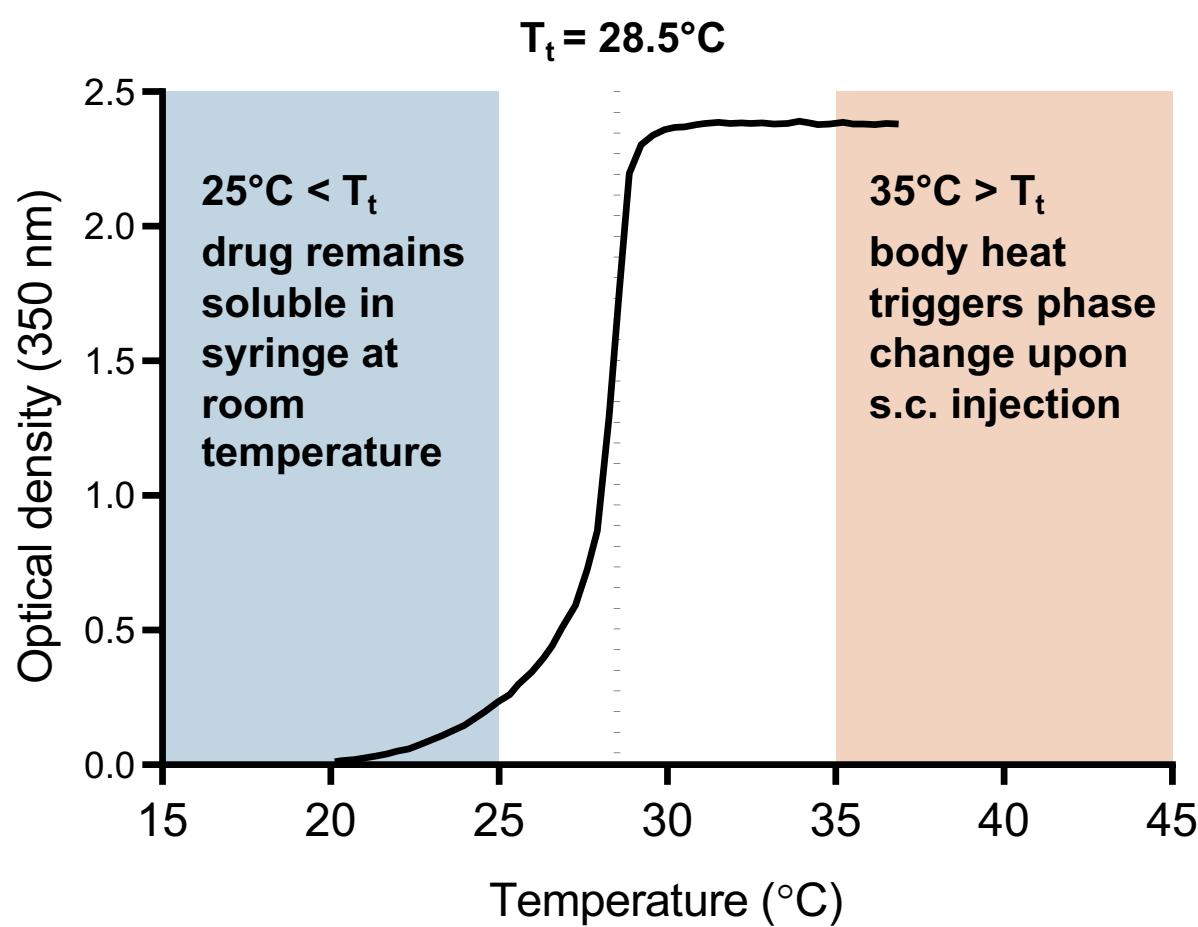
Replacing Text with Visuals

Depot formation

- GLP1-ELP-FGF21 designed to form an *in vivo* drug depot
 - GLP1-ELP and ELP-FGF21 previously optimized as depot-forming single agonist treatments [2,3]
 - $25^{\circ}\text{C} < T_t \rightarrow$ drug remains soluble in syringe at room temperature
 - $35^{\circ}\text{C} > T_t \rightarrow$ body heat triggers phase change upon s.c. injection
 - T_t identified by monitoring ELP solution turbidity during temperature ramping
- ELP T_t inversely dependent on concentration
 - Core of depot represents injection concentration
 - Depot boundary slowly hydrated
 - Concentration decreases $\rightarrow T_t$ increases
 - When T_t increases above 35°C , fusion unimers resolubilize and leave depot

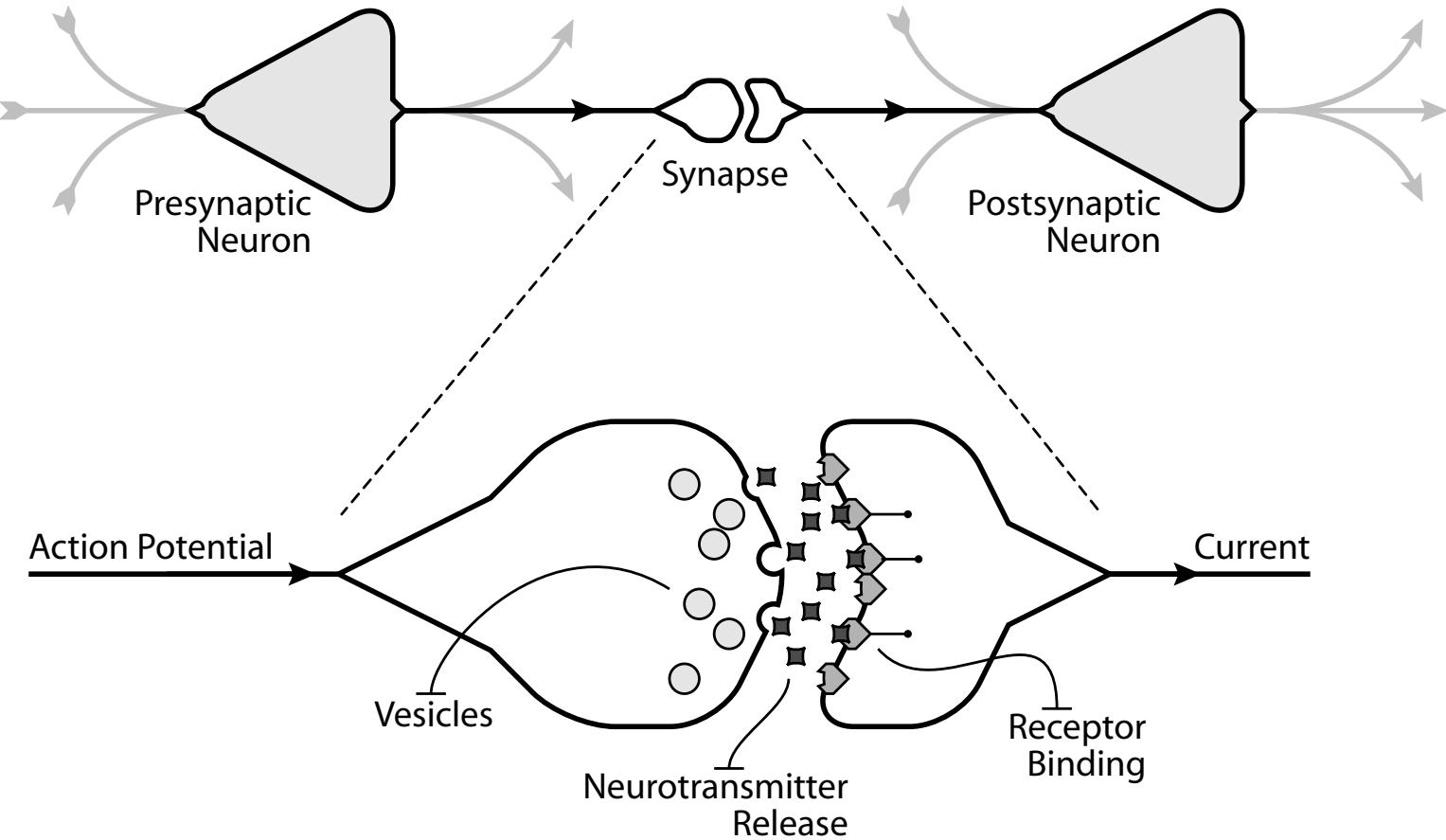


Dual agonist designed to form an *in vivo* drug depot

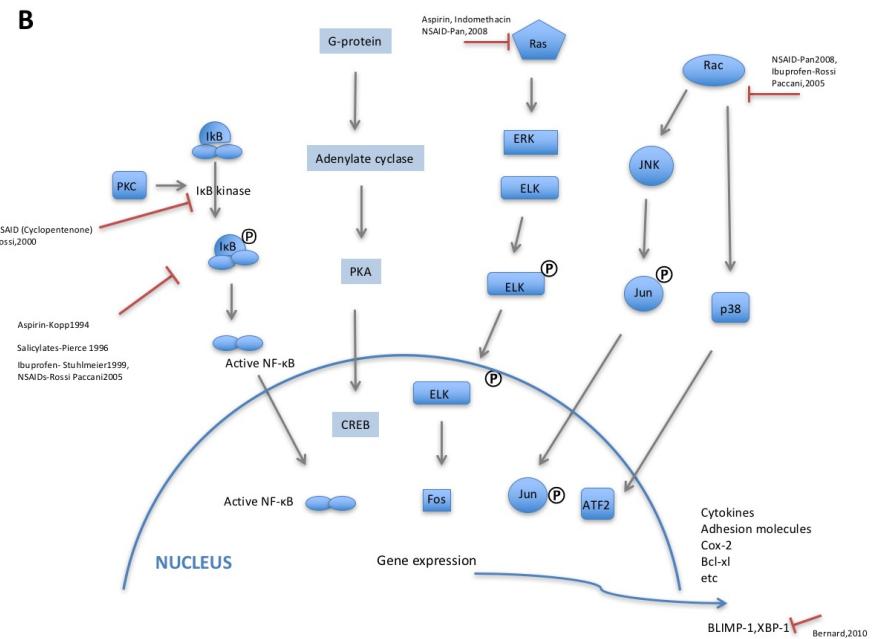
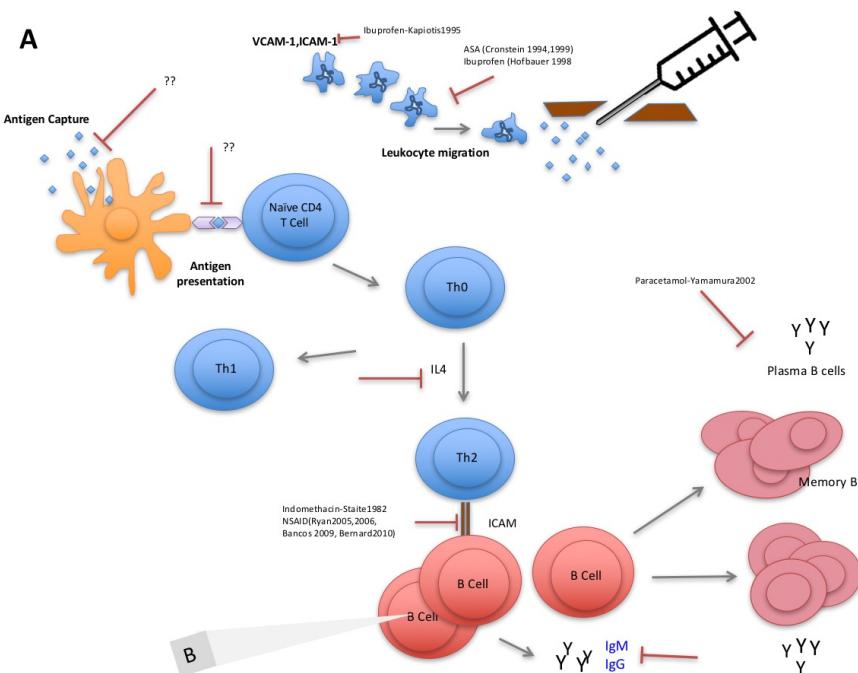


Depot slowly dissolves as it gets hydrated

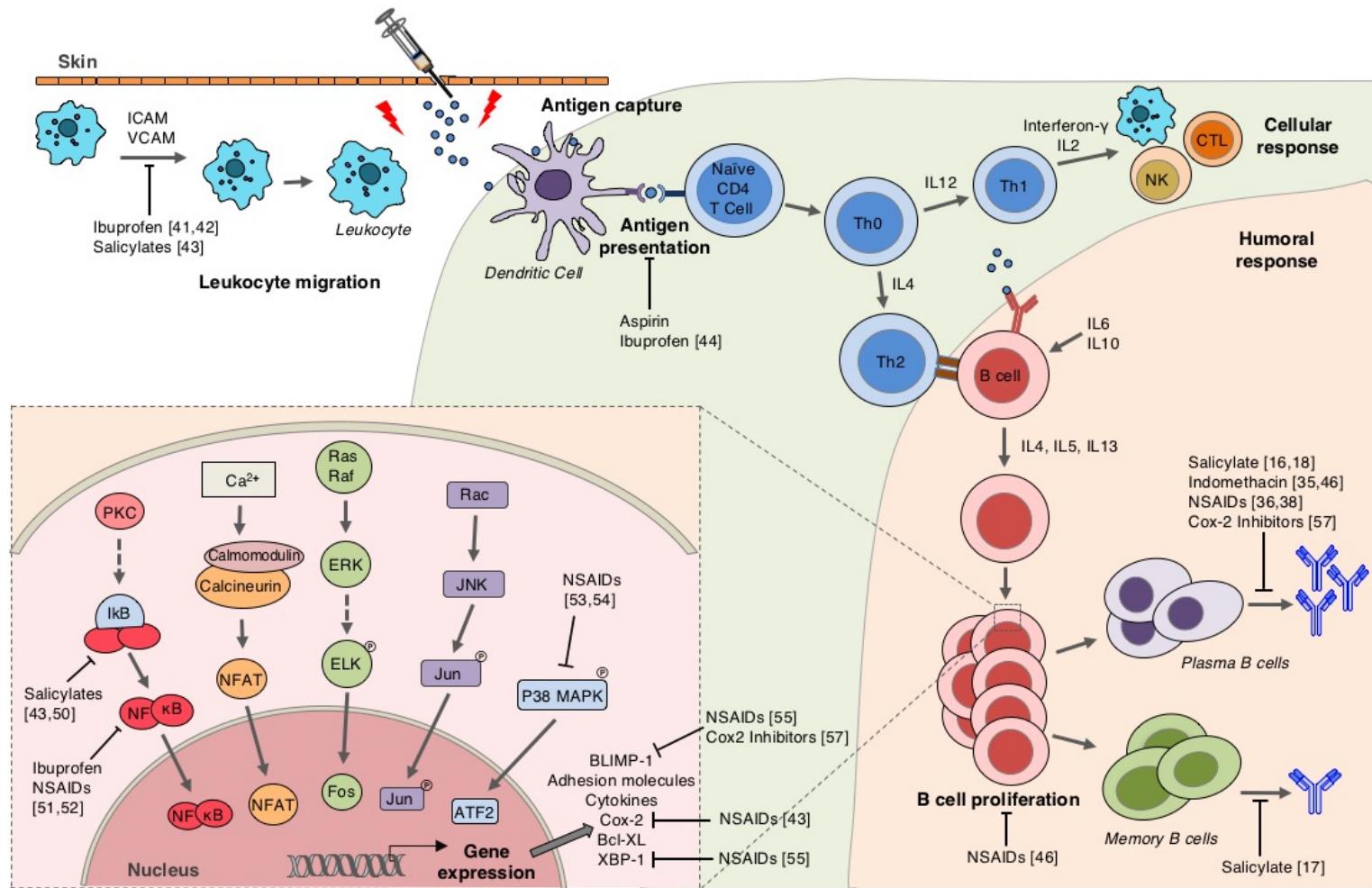
Adobe Illustrator for diagrams



PowerPoint for diagrams



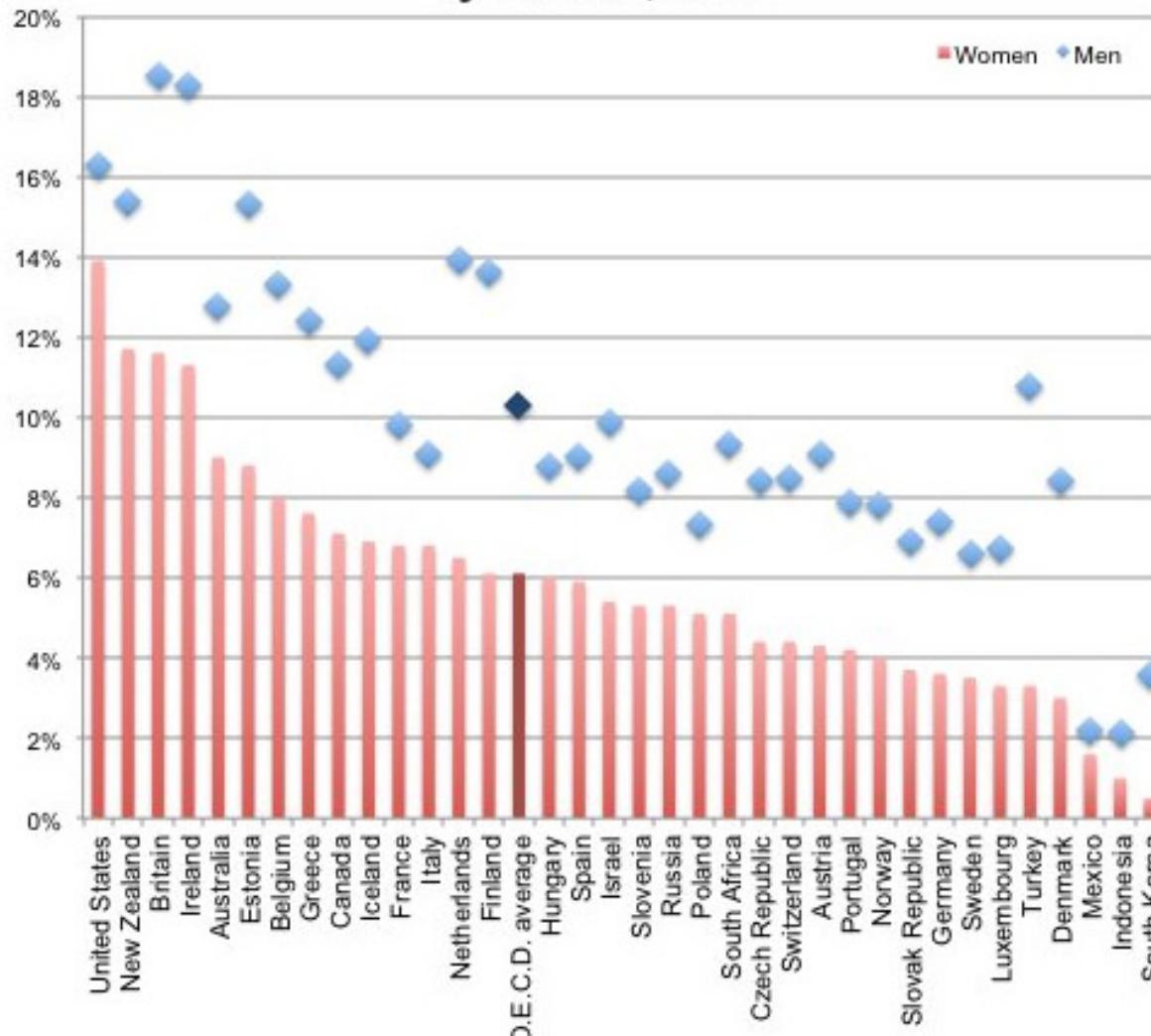
PowerPoint for diagrams



Ezzeldin Saleh

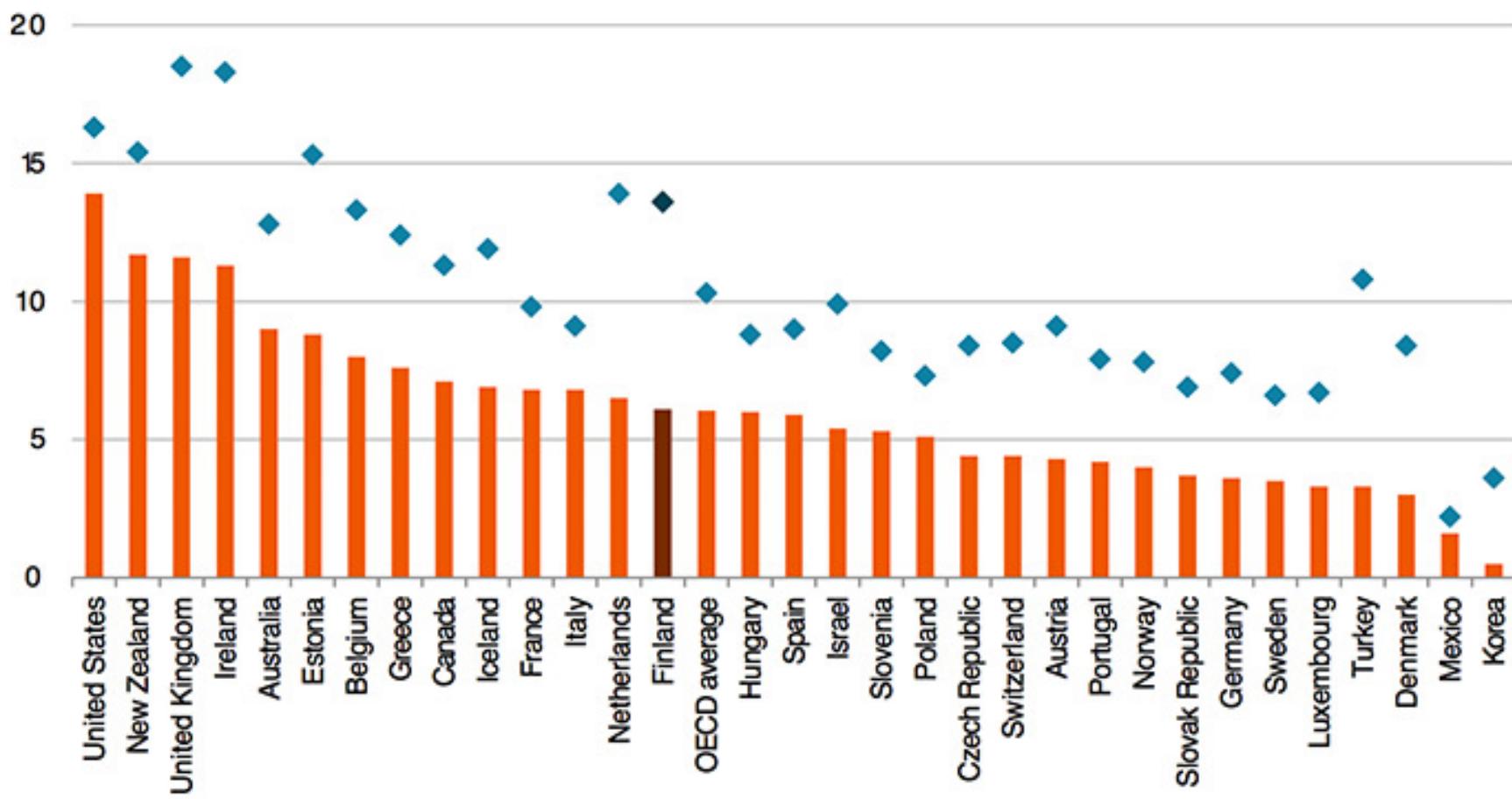
Figure critique & reworks

Percentage of Employed Who Are Senior Managers, by Gender, 2008

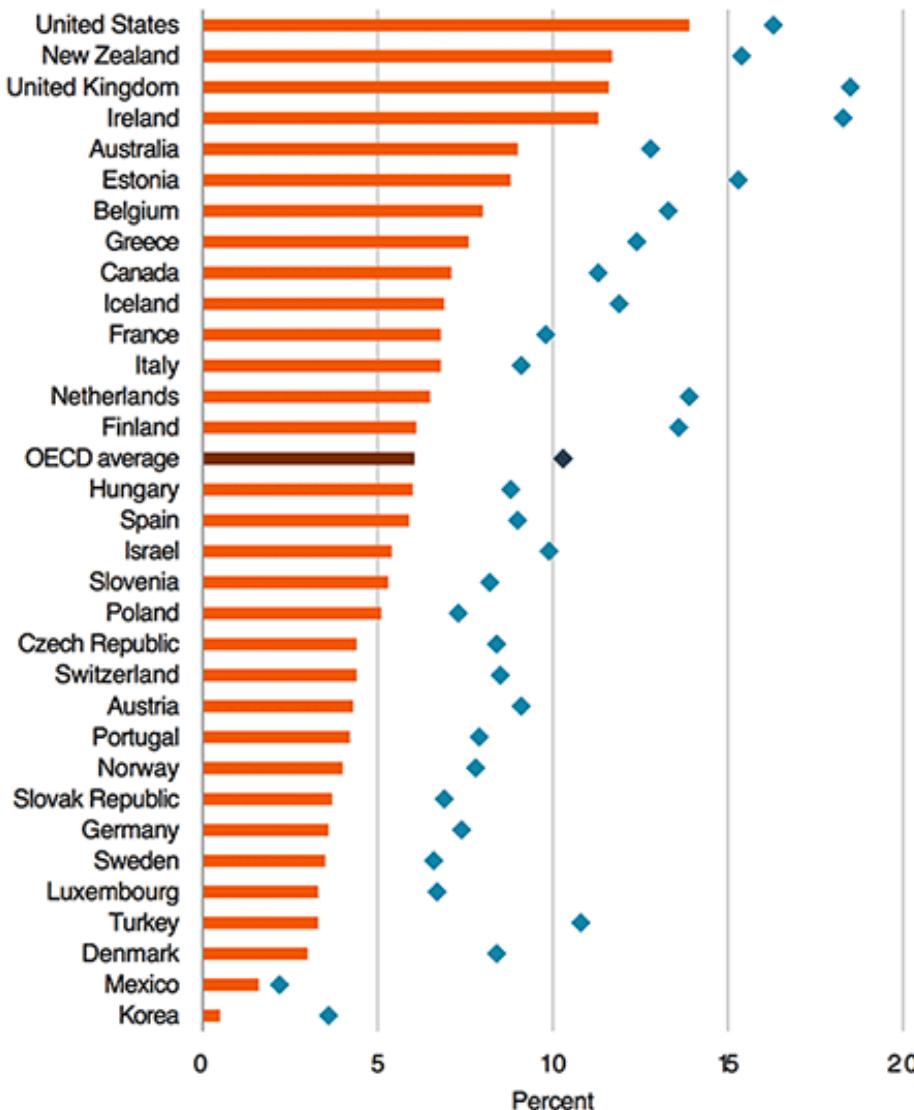


Percentage of Employed Who are Senior Managers, by Gender, 2008

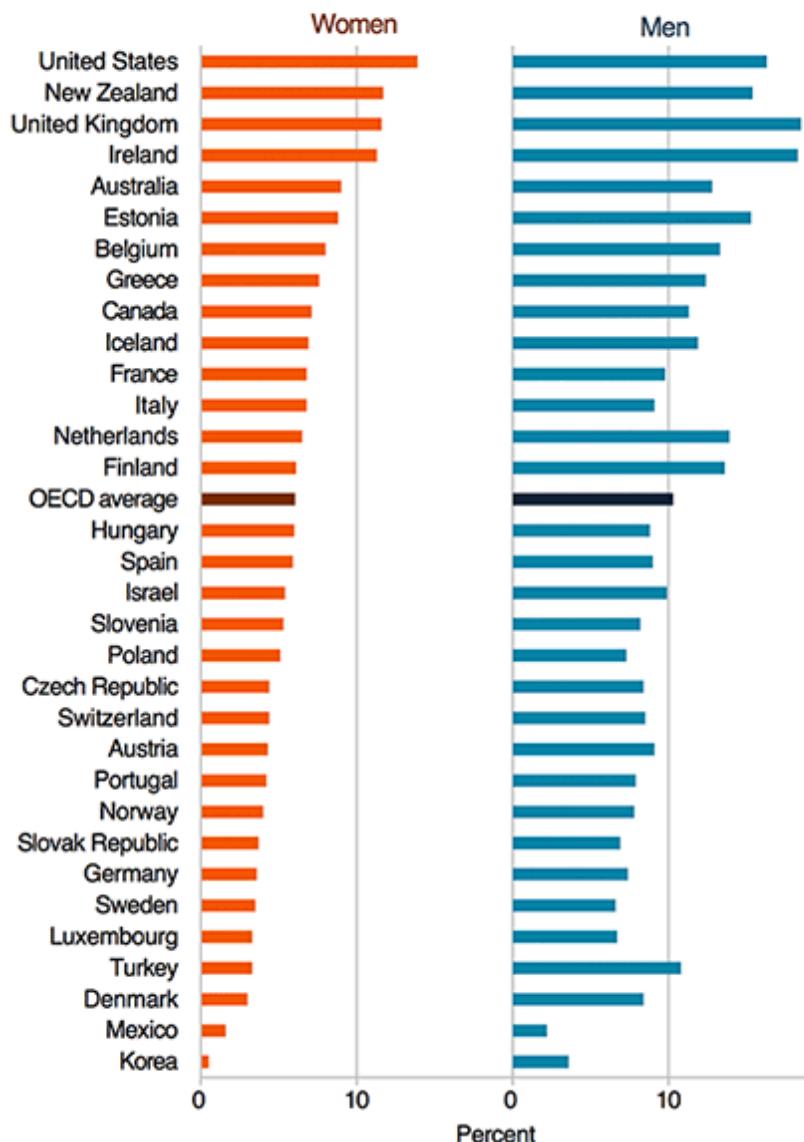
(Percent) □ Women □ Men



Percentage of Employed Who are Senior Managers,
by Gender, 2008
(Percent) □ Women □ Men

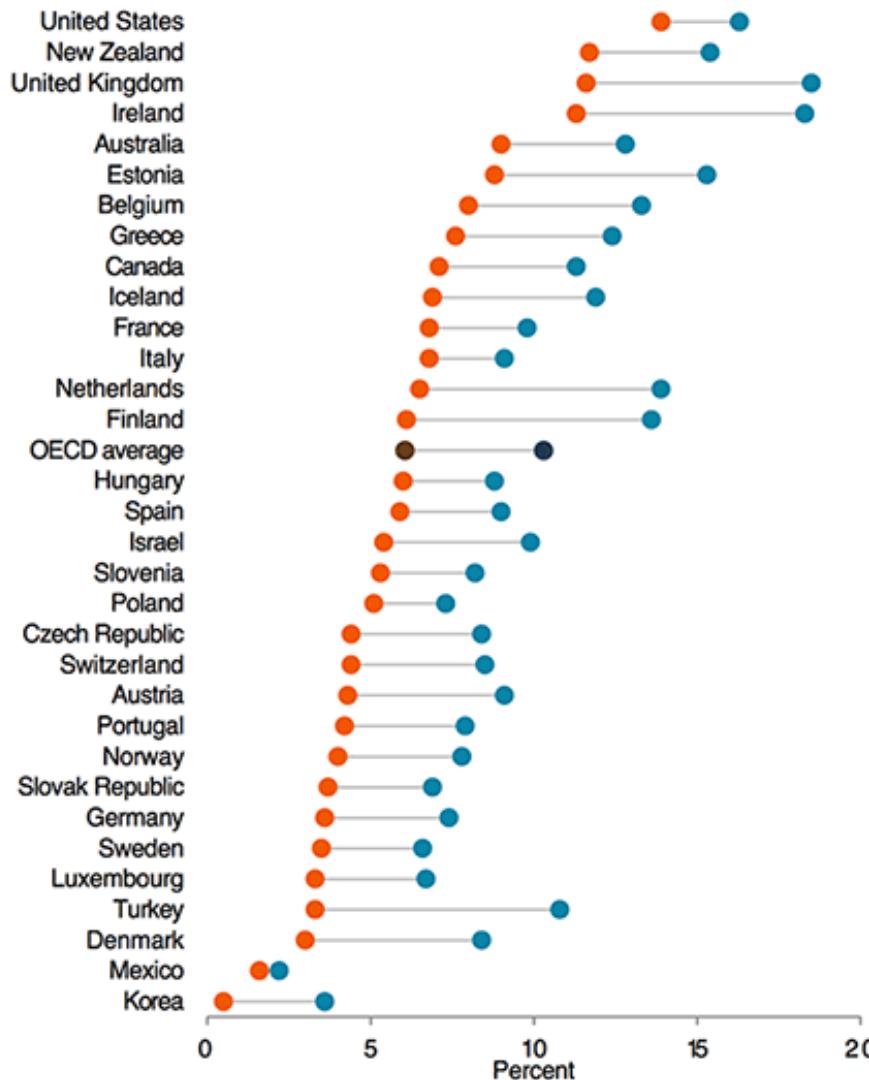


Percentage of Employed Who are Senior Managers,
by Gender, 2008
(Percent)

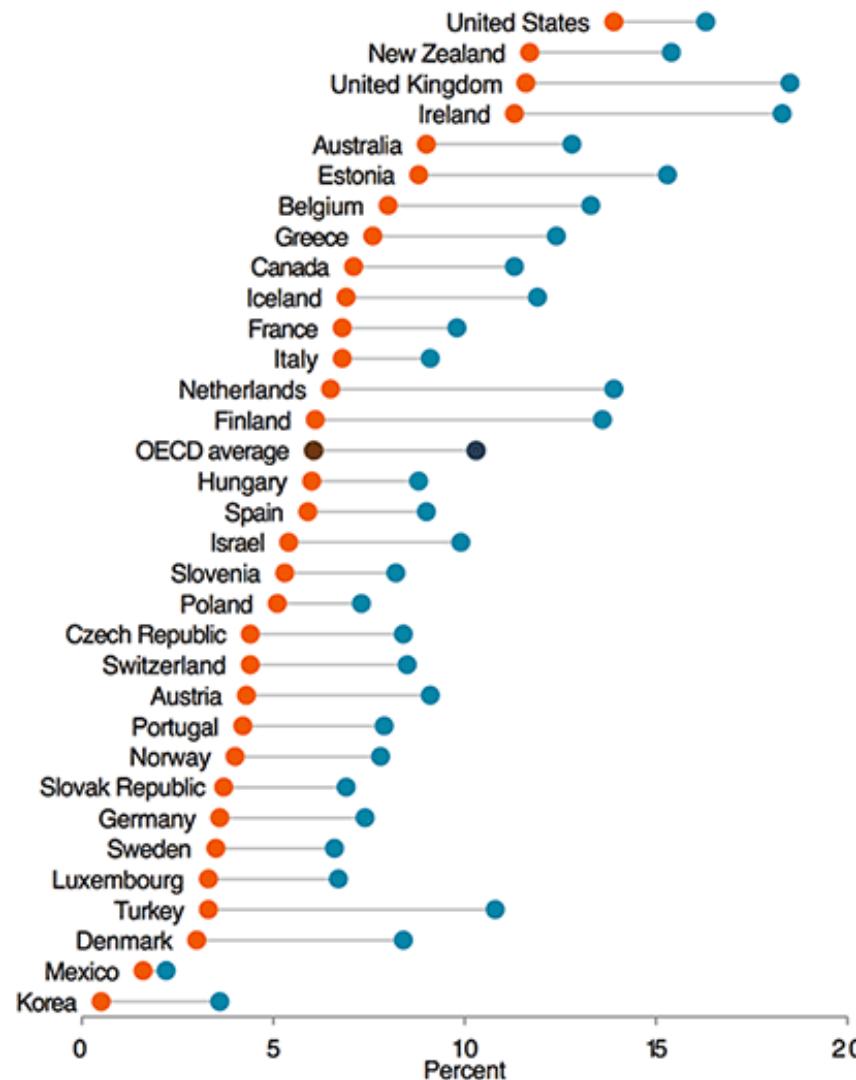


Jon Schwabish: <http://thewhyaxis.info/gap-remake/>

Percentage of Employed Who are Senior Managers,
by Gender, 2008
(Percent) ● Women ● Men

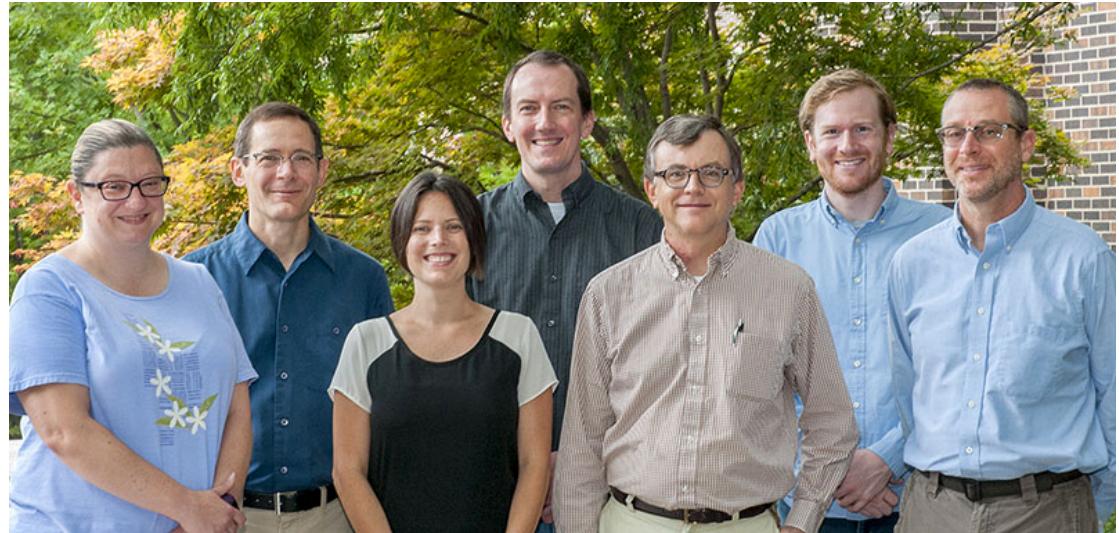


Percentage of Employed Who are Senior Managers,
by Gender, 2008
(Percent) ● Women ● Men



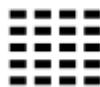


askdata@duke.edu



<http://library.duke.edu/data>

Support Areas



Data Sources



Data Science



Mapping and GIS



Data Management



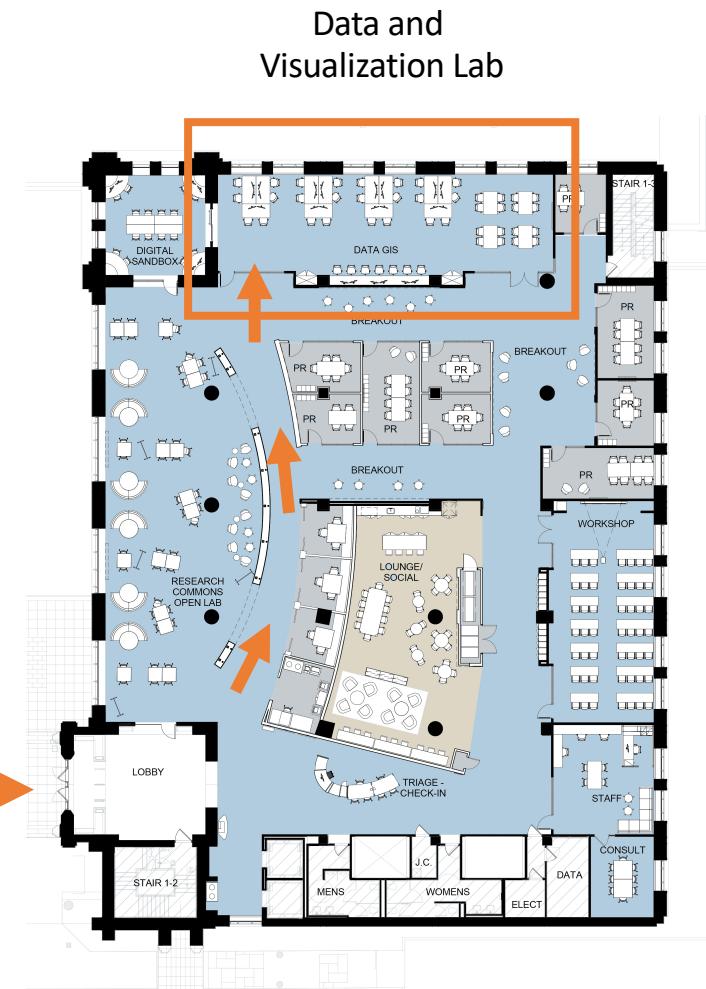
Data Visualization

Brandaleone Family Lab for Data and Visualization Services

<http://library.duke.edu/data/about/lab>

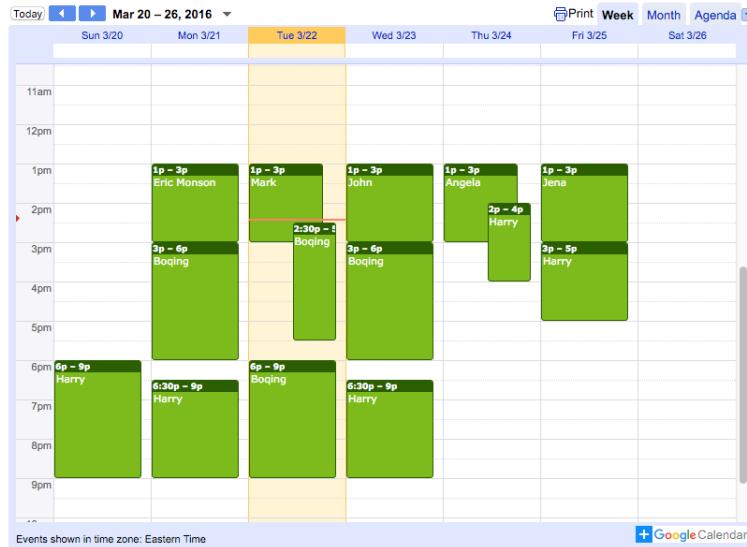
- **The Edge** (*1st floor of Bostock*)
- Open whenever the library is open
- 12 high-powered Dell workstations
- 3 Bloomberg financial workstations
- Software for data analysis, GIS, and visualization

Entrance to
Bostock



Consulting

Office Hours - Data and Visualization Services Lab

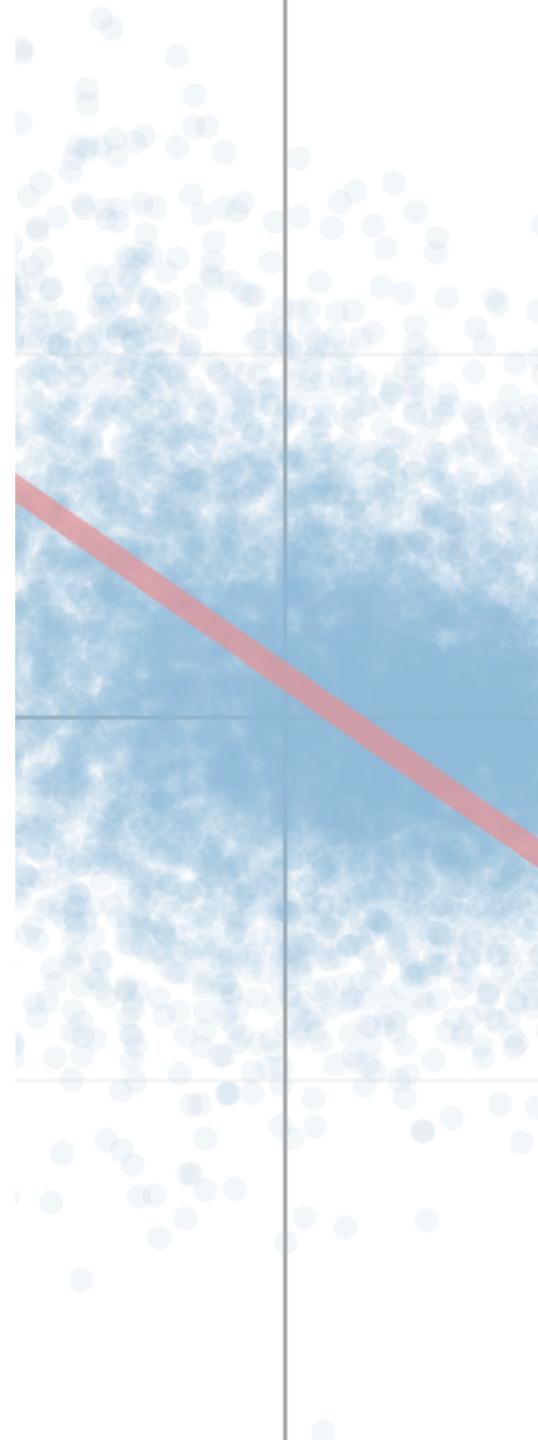


<http://library.duke.edu/data/about/schedule>

...or email
askdata@duke.edu
to make an appointment

Types of visualization consulting

- Look at data and brainstorm about the best visualization
- Recommend appropriate tools
- Troubleshoot software problems
- Help with cleaning and structuring data
- Offer graphic design advice for figures, diagrams, slides and posters



Inhibiting Abelson (ABL) family kinase activity as a potential treatment for leptomeningeally disseminated medulloblastoma

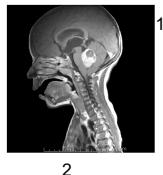
Jill K. Jones^{1,2}

¹Trinity College of Arts & Sciences, Duke University, Durham, NC, USA, ²Duke University Medical Center, Department of Neurosurgery, Durham, NC, USA

BACKGROUND

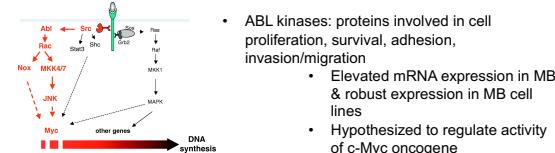
Medulloblastoma (MB)

- Most common pediatric brain tumor
- Arises from various progenitor cells in cerebellum
- 80% of patients with recurrent MB diagnosed with **leptomeningeal dissemination**: the dissemination of cancerous cells into the delicate membranes (meninges) surrounding the brain



What mediates leptomeningeal dissemination in MB?

Abelson (ABL) kinases and c-Myc?

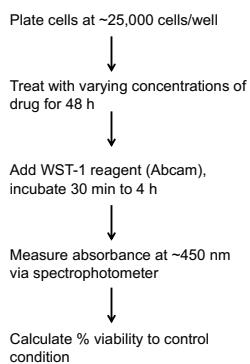


- ABL kinases: proteins involved in cell proliferation, survival, adhesion, invasion/migration
 - Elevated mRNA expression in MB & robust expression in MB cell lines
 - Hypothesized to regulate activity of c-Myc oncogene

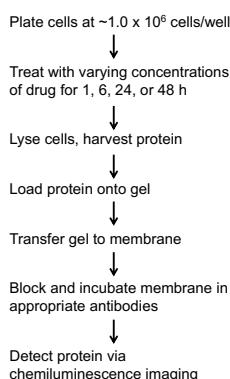
Does inhibiting ABL kinase activity affect MB cell proliferation and protein expression?

METHODS

WST-1 assay



Western blotting



RESULTS

Cell viability after treatment with GNF-5, according to WST-1 assay

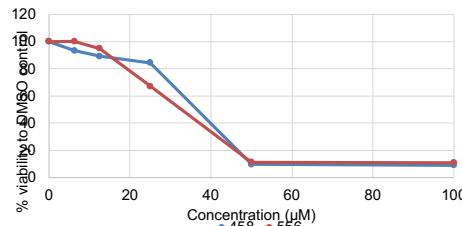


Figure 1. GNF-5 exerts cytotoxic effects on patient-derived medulloblastoma cell lines. Cell viability to DMSO control after treatment with 0, 25, 50, 75, or 100 μM GNF-5 for 48 h.

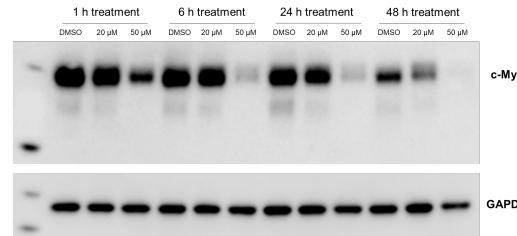


Figure 2. GNF-5 treatment decreases c-Myc expression in a time- and dose-dependent manner. Western blot of 458 cell lysate using c-Myc antibody (Cell Signaling Technology); GAPDH = loading control.

DISCUSSION

FUTURE DIRECTIONS

- Replicate c-Myc blot in multiple cell lines
 - Look into Rac, NOX, JNK signaling pathways
- Next year: move to *in vivo* mouse model of medulloblastoma
- Other ABL kinase inhibitors:
 - GNF-2
 - GNF-7
 - ABL-001
- LITERATURE REVIEW



ACKNOWLEDGEMENTS

Thank you to Duke's Undergraduate Research Support Office for providing me with two semesters' worth of funding for this project through Neuroscience independent Study Grants, and to the Charles Lafitte Foundation for providing me with two semesters' worth of funding through URS Matching Grants. Thank you to Dr. Eric Thompson for his mentorship and guidance as my principal investigator, and thank you to Dr. Wafa Hassen, my bench mentor, for answering my seemingly endless questions and being such a role model for the scientist and person I want to become one day.

REFERENCES

1. Polkinghorn, W. R., & Tarbell, N. J. (2007). Medulloblastoma: Tumorigenesis, current clinical paradigm and efforts to improve risk stratification. *Nature Clinical Practice Oncology*, 4(5), 295-304. doi:10.1038/ncponc0794
2. Boureux, A., Furstoss, O., Simon, V., & Roche, S. (2005). Abl tyrosine kinase regulates a Rac/JNK and a Rac/Nox pathway for DNA synthesis and Myc expression induced by growth factors. *Journal of Cell Science*, 118(16), 3717-3726. doi:10.1242/jcs.02491

Inhibiting ABL kinases: A potential treatment for pediatric brain tumors?

Jill K. Jones^{1,2}

¹Trinity College of Arts & Sciences, Duke University, Durham, NC, USA, ²Duke University Medical Center, Department of Neurosurgery, Durham, NC, USA

Background

What is medulloblastoma (MB)?

- Most common pediatric brain tumor
 - Leading cause of cancer-related pediatric deaths
- Arises from various progenitor cells in cerebellum



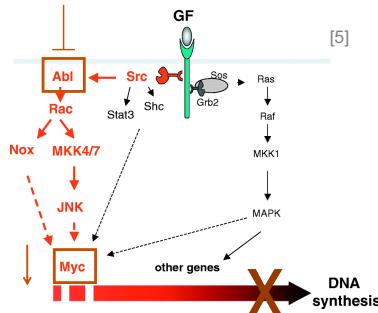
What mediates its dissemination?

- Leptomeningeal dissemination:** spread of tumor cells into delicate membranes (*meninges*) surrounding the brain
 - Reported to occur in ~80% of patients with recurrent MB³
- ABL kinases:** proteins involved in cell proliferation, survival & adhesion
 - Elevated mRNA expression in MB
 - Robust expression in several MB cell lines
 - May upregulate **c-Myc** oncogene
 - Inhibition suppresses non-small cell lung cancer metastasis⁴



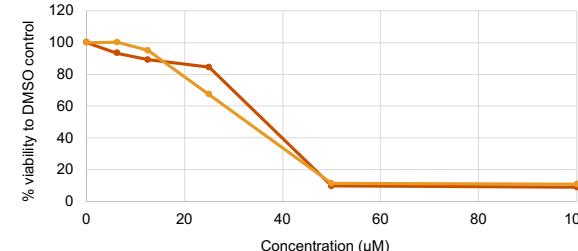
How might we stop it?

INHIBIT WITH SELECTIVE DRUG (GNF-5)



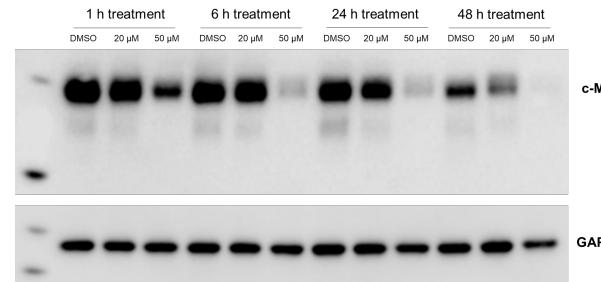
Methods, results, and discussion

GNF-5 kills patient-derived MB cells



Note: Each line represents results obtained in a different cell line. Blue = MB458; orange = MB556.

GNF-5 decreases c-Myc expression before cells die



What this means:

ABL kinases do seem to play a role in MB cell survival through regulation of the c-Myc oncogene

So **inhibiting ABL kinases holds potential as an alternative treatment** for pediatric brain tumor patients

Future directions

- Attempt to replicate findings in multiple patient-derived MB cell lines
- Explore effects of ABL inhibition on Rac, Nox, and JNK protein expression
- Treat cells with other ABL-selective inhibitors (GNF-2, GNF-7, ABL-001) to assess drug efficacy

References

- Polkington, W. R., & Tarbell, N. J. (2007). Medulloblastoma: Tumorigenesis, current clinical paradigm and efforts to improve risk stratification. *Nat Clin Pract Oncol*, 4(5), 295-304. doi:10.1038/ncponc0794
- Hottiger, A. F., et al. (2018). Phase I study of sorafenib combined with radiation therapy and temozolamide as first-line treatment of high-grade glioma. *Br J Cancer*, 118(6). doi:10.1038/bjc.2017.469
- Koschmann, C., et al. (2016). Survival After Relapse of MB. *J Pediatr Hematol Oncol*, 38(4), 269-273. doi:10.1097/MPH.00000000000000547
- Gu, J. J., et al. (2016). Inactivation of ABL kinases suppresses non-small cell lung cancer metastasis. *JCI Insight*, 1(21). doi:10.1172/jci.insight.89647
- Boureux, A., et al. (2005). Abl tyrosine kinase regulates a Rac/JNK and a Rac/Nox pathway for DNA synthesis and Myc expression induced by growth factors. *J Cell Sci*, 118(16), 3717-3726. doi:10.1242/jcs.02491

Acknowledgements

Thank you to Duke's Undergraduate Research Support Office and the Charles Lafite Foundation for providing me with two Neuroscience Independent Study and two URS Matching Grants. Thank you to Dr. Eric Thompson for his mentorship and guidance as my principal investigator, and thank you to Dr. Wafa Hassen, my bench mentor, for answering my endless questions and being such a role model for the scientist and person I want to become one day. Thank you also to all the pediatric patients who have battled or are currently battling MB - we are fighting for your cure.

Workshops

<http://library.duke.edu/data/workshops>

- Typically toward the beginning of the semester
- 1-2 hours, often hands-on



For announcements, sign up for our listserv:
<https://lists.duke.edu/sympa/subscribe/dvs-announce>

Fall 2019 workshops sample

Workshop	Date	Time
Python for Data Science: Pandas and Jupyter Lab	Sep 12	10:00 AM - 12:00 PM
Intro to R: Data Transformations, Data Structures, and the Tidyverse	Sep 12	1:00 PM - 3:00 PM
R Open Lab: semi-structured opportunities for learning R	Sep 16	6:00 PM - 7:00 PM
Research Data Management 101+ Disciplinary Discussions	Sep 17	1:00 PM - 3:00 PM
Visualization in R using ggplot2	Sep 19	1:00 PM - 3:00 PM
OSF + TIER Protocol: Designing a reproducible workflow	Sep 24	10:00 AM - 11:00 AM
Python for Data Science: Visualization with Altair	Sep 26	10:00 AM - 12:00 PM
Intro to R: Data Transformations, Data Structures, and the Tidyverse	Oct 3	1:00 PM - 3:00 PM
Intro to Tableau	Oct 9	10:00 AM - 12:00 PM
Visualization in R using ggplot2	Oct 9	1:00 PM - 3:00 PM
R Open Lab: semi-structured opportunities for learning R	Oct 14	6:00 PM - 7:00 PM
An Introduction to Duke's Research Data Repository	Oct 16	10:00 AM - 11:00 AM
Mapping in Tableau	Oct 23	1:00 PM - 3:00 PM
Advanced ggplot2 techniques	Oct 25	9:30 AM - 11:30 AM
Research Data Management 101 + Tool Demonstrations	Oct 31	1:30 PM - 3:30 PM
Research Data Management 201: Preparing Your Data for Publishing	Nov 14	1:30 PM - 3:30 PM

Videos of past workshops

<http://bit.ly/DVSvideos>

The screenshot shows a Panopto video player interface. At the top left is the Panopto logo and the title "Panopto® Figures and Posters". To the right is the date "March 4, 2016" and the category "DVS Training". On the far right are "Help" and "Sign in" buttons. The main video frame shows two people in a classroom setting: a woman in a dark shirt and a man in a plaid shirt standing near a whiteboard. Below the video frame is a search bar with the placeholder "Search this recording" and a magnifying glass icon. Underneath the search bar are "Discussion" and "Sign in" buttons. The video title "Designing Academic Figures and Posters" is displayed prominently in large black text. Below the title is the date "March 4, 2016". A link to "Slides: <http://duke.box.com/PostersSpring2016>" is provided. Two speakers are identified: "Angela Zoss" and "Eric Monson". Their roles are listed as "Data Visualization Coordinator" and "Data Visualization Analyst" respectively, both from "Data and Visualization Services". The video progress bar shows a timestamp of "0:03". Below the progress bar are three video thumbnails: "Good Posters" (1:32), "Causal Inference" (4:32), and "Purpose of a poster" (7:32). The bottom right corner of the video player has a vertical toolbar with icons for "C", "P", "V", and other controls.

Visualization Friday Forum

<http://vis.duke.edu/FridayForum>

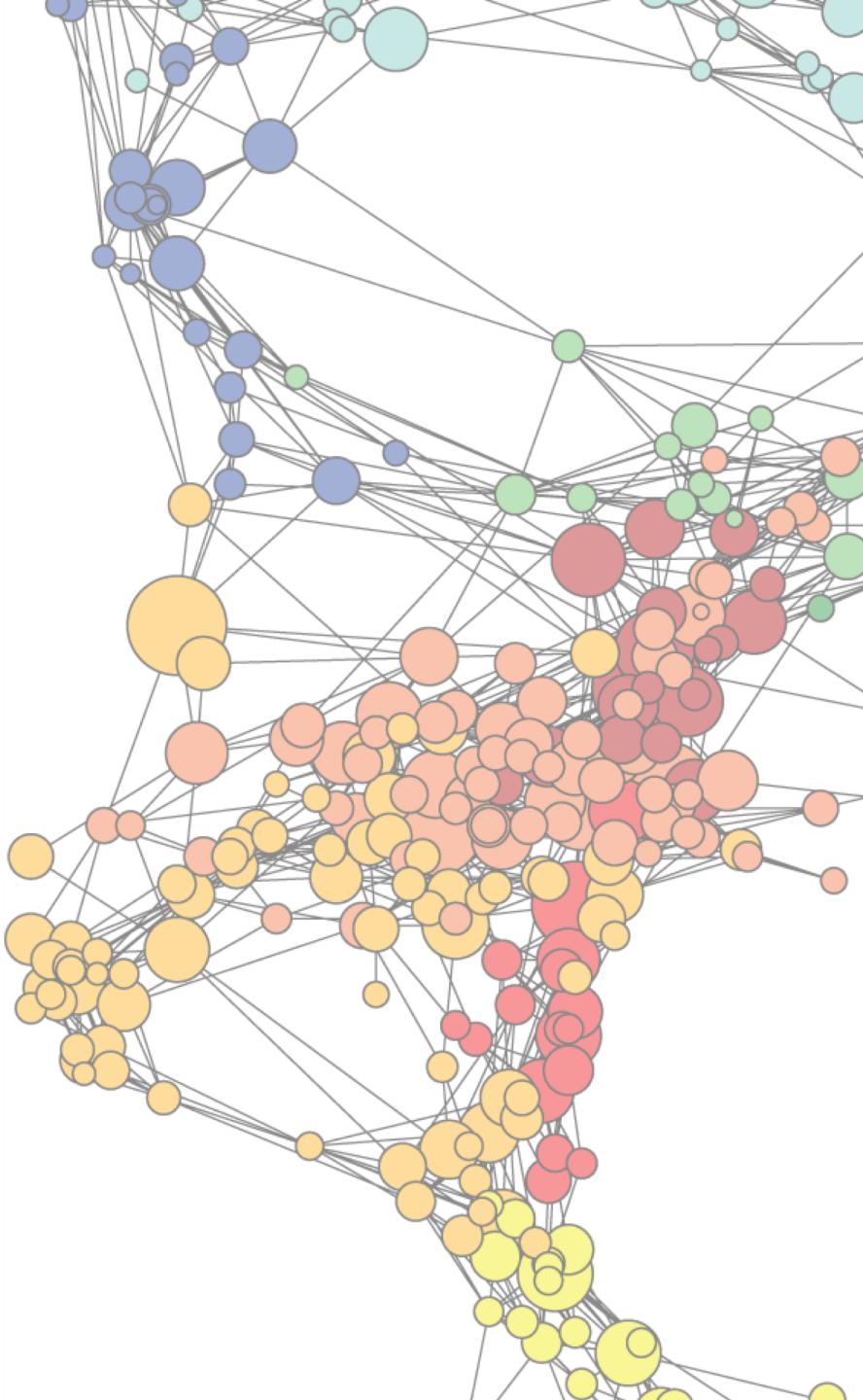
- First Friday of every month
- Noon in LSRC D106
- Free lunch (sssh!)
- Live streamed and recorded
- Email me to get on the mailing list



Questions

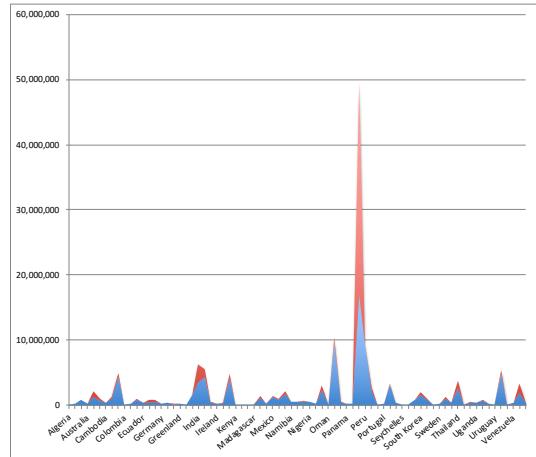
askdata@duke.edu

Slides: <http://bit.ly/STA199visFall2019>

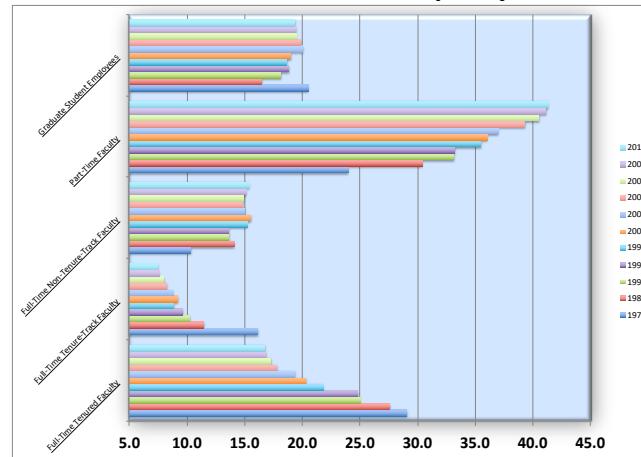


Visualization Rework Activity

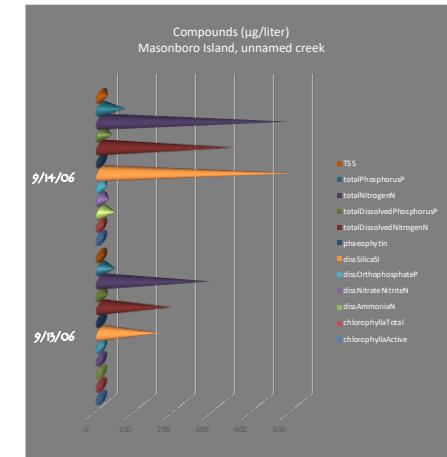
Fishing Industry



Instructional Staff Employment



NERRS Nutrients



<https://www2.stat.duke.edu/courses/Fall19/sta199.001/hw/hw-01.html>

Extras

Chart choosing helper sites

<https://datavizcatalogue.com/>

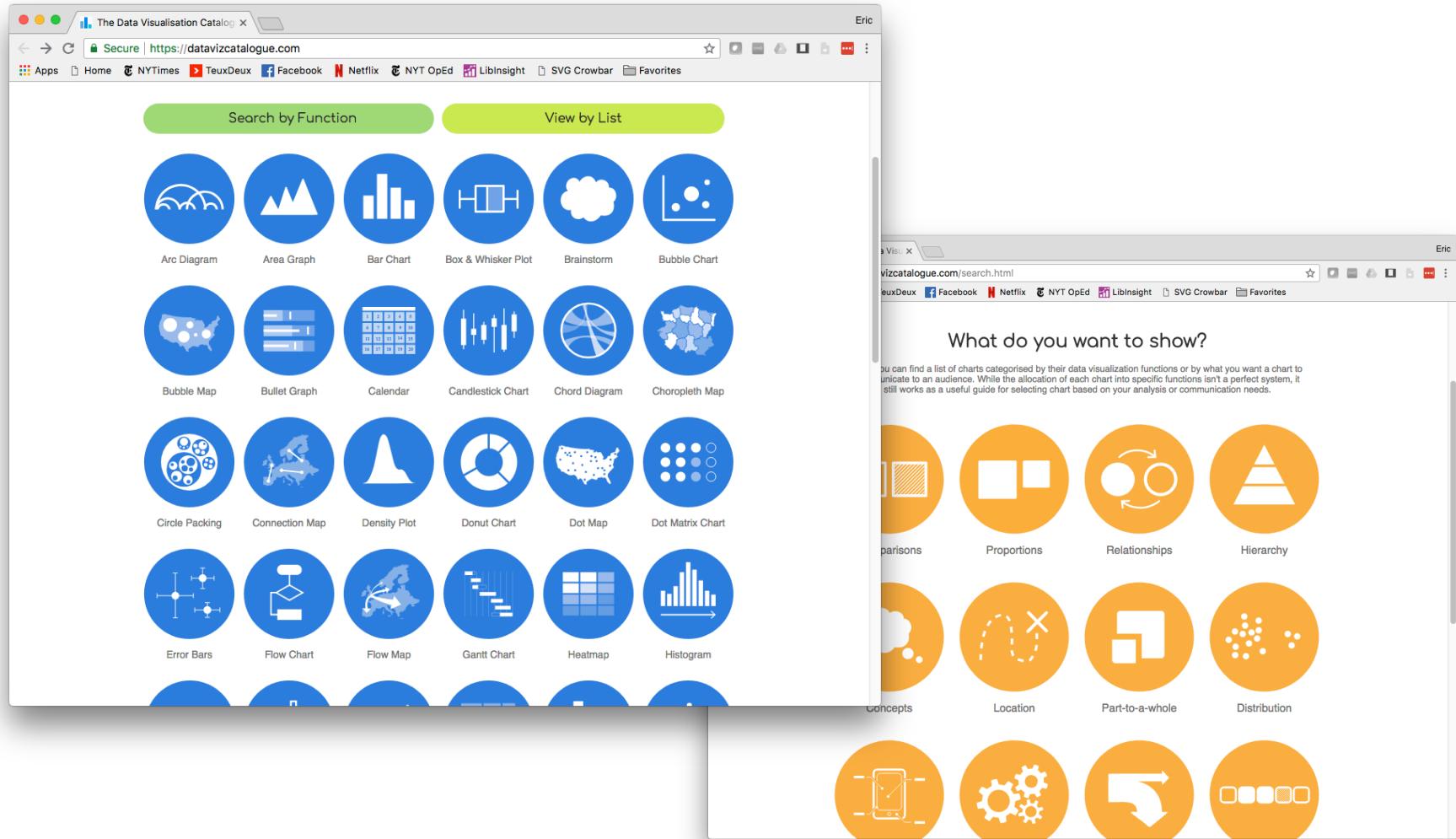


Chart choosing helper sites

<http://datavizproject.com/>

Screenshot of the Data Viz Project website (<http://datavizproject.com/>) showing various chart types and their data.

The site features a navigation bar with categories: ALL, FAMILY, INPUT, FUNCTION, SHAPE, and a search bar.

Key charts displayed include:

- Sankey Diagram**: A flow diagram with nodes A, B, C, D, E, F.
- Alluvial Diagram**: A flow diagram with nodes A, B, C.
- Donut Chart**: A donut chart with segments labeled A, B, C, D.
- Radial Bar Chart**: A radial chart with segments labeled A, B, C, 200, 400, 600.
- Radial Histogram**: A sunburst-style histogram with values 0, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20.
- Sorted Stream Graph**: A stream graph with nodes A, B, C, D over months JAN, FEB, MAR, APR.
- Fishbone Diagram**: A cause-and-effect diagram with nodes PEOPLE, MACHINES, METHODS, MATERIALS, CAUSE, OUTCOME.
- Matrix Diagram**: A simple matrix with columns 1, 2, 3.
- Matrix Diagram (Roof Shaped)**: A roof-shaped matrix diagram with node A.
- Arc Diagram**: An arc diagram showing connections between nodes.
- Hexagonal Binning**: A scatter plot using hexagonal bins.
- Radial Line Graph**: A radial line graph with concentric rings.
- Line Graph**: A line graph with data points at x-coordinates 2, 4, 6, 8, 10.
- Cluster Analysis**: A scatter plot with two clusters of points.
- Area Chart**: A line chart with filled areas under the lines.
- Bagplot**: A bagplot showing outliers.
- Radial Area Chart**: A radial area chart with segments labeled 1, 2.
- Spline Graph**: A line graph with a smooth curve.

The site also includes a sidebar with tables and dropdown menus for input, function, and shape selection.

Chart choosing helper sites

<https://github.com/ft-interactive/chart-doctor/tree/master/visual-vocabulary>

The screenshot shows a GitHub repository page for 'chart-doctor/visual-vocabulary'. The main view displays a large grid of data visualization examples categorized into ten groups: Deviation, Correlation, Ranking, Distribution, Change over Time, Magnitude, Part-to-whole, Spatial, Flow, and a final column. Below this grid, the page features several sections with detailed descriptions and examples:

- Visual vocabulary**: A section with a brief introduction and a link to ft.com/vocabulary.
- Ranking**: Describes how to use position in an ordered list. Examples include FT's wealth deprivation league tables and constituency election results.
- Distribution**: Focuses on showing values in a dataset over time. Examples include income distribution, population, and age-sex distribution.
- Change over Time**: Emphasizes trends. Examples include share price movements and economic time series.
- Financial Times Visual Vocabulary**: A summary statement about the poster's purpose: "A poster (available in English, Japanese, traditional Chinese and simplified Chinese) and we journalists to select the optimal symbology for data visualisations, by the Financial Times Visual Vocabulary team."
- Deviation**, **Correlation**, **Ranking**, **Distribution**, **Change over Time**, **Magnitude**, **Part-to-whole**, **Spatial**, and **Flow**: These sections provide detailed explanations and examples for each category, often including small diagrams or charts.

Chart choosing helper sites

<https://policyviz.com/shop/>

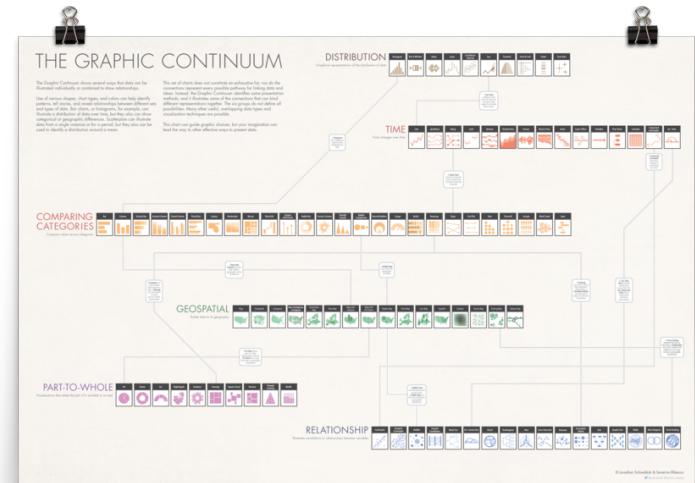
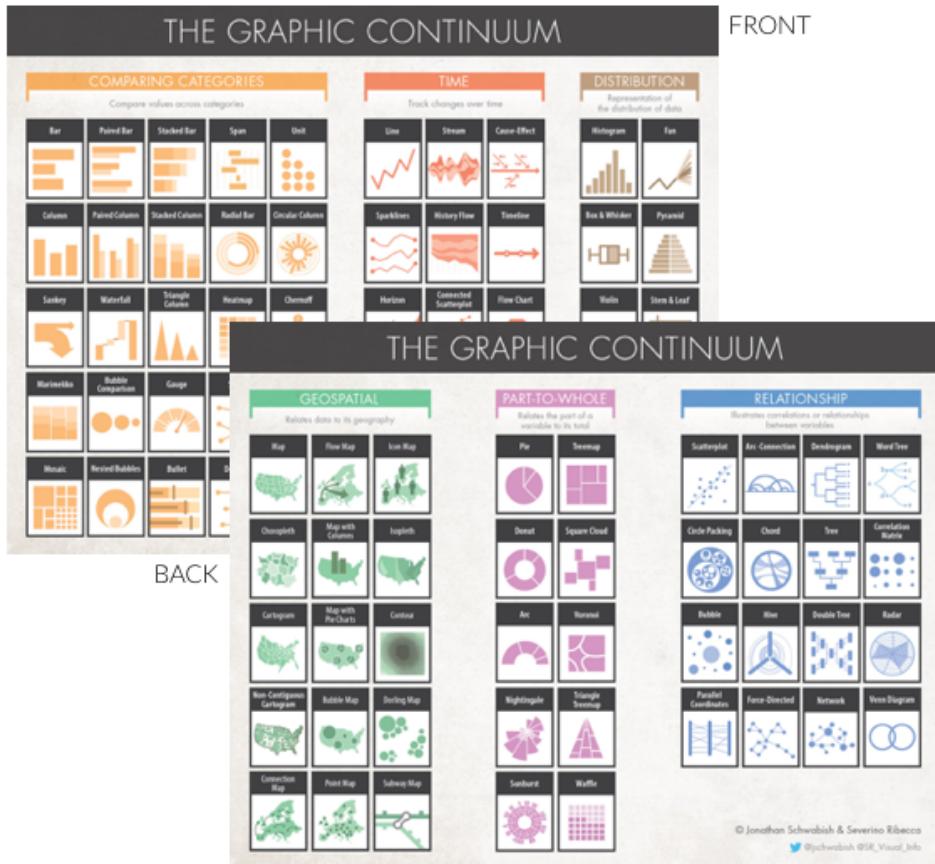


Chart choosing helper sites

<http://chartmaker.visualisingdata.com/>

The screenshot shows a web browser window titled "The Chartmaker Directory". The URL in the address bar is <http://chartmaker.visualisingdata.com/>. The page features a navigation bar with links like "Home", "NYTimes", "TeuxDeux", "Facebook", "Netflix", "NYT OpEd", "LibInsight", "SVG Crowbar", and "Favorites". A sidebar on the right includes social sharing icons for Twitter, Facebook, Google+, and LinkedIn.

The main content area is titled "THE CHARTMAKER DIRECTORY" and contains a grid of chart types. The columns represent different tools: Amazon QuickSight, ArcGIS, ChartJS, D3.js, Data Illustrator, Datawrapper, Flourish, FusionCharts, Gephi, Google Charts, and Highcharts. The rows list various chart types: Bar chart, Clustered bar chart, Bullet chart, Connected dot plot, Pictogram, Bubble chart, Word cloud, and Radar chart. Each cell in the grid contains a visual representation of the chart's availability, consisting of black dots of varying sizes and shapes (circles, triangles, diamonds) on a light gray grid background.

	Amazon QuickSight	ArcGIS	ChartJS	D3.js	Data Illustrator	Datawrapper	Flourish	FusionCharts	Gephi	Google Charts	Highcharts
Bar chart	●			●●●	○	●●●	○	○		●●	●●○
Clustered bar chart	●				●	●●●○	○	○		●●	
Bullet chart					●	●●			○		
Connected dot plot				●●		●●●○					
Pictogram					○						
Bubble chart				●●●○	○		○○	○		●	
Word cloud					●						
Radar chart			○	●				○		○	

Chart choosing helper sites

<https://xeno.graphics/>

