When Charts Get Dangerous

Training for the Foundations of Data Viz

Los Angeles | London | New Delhi | Singapore | Washington DC | Melbourne | Toronto

SAGE Publishing

Today’s panel

Daniela

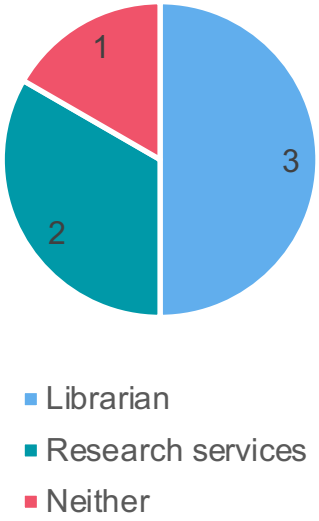
Thilani

Cameron

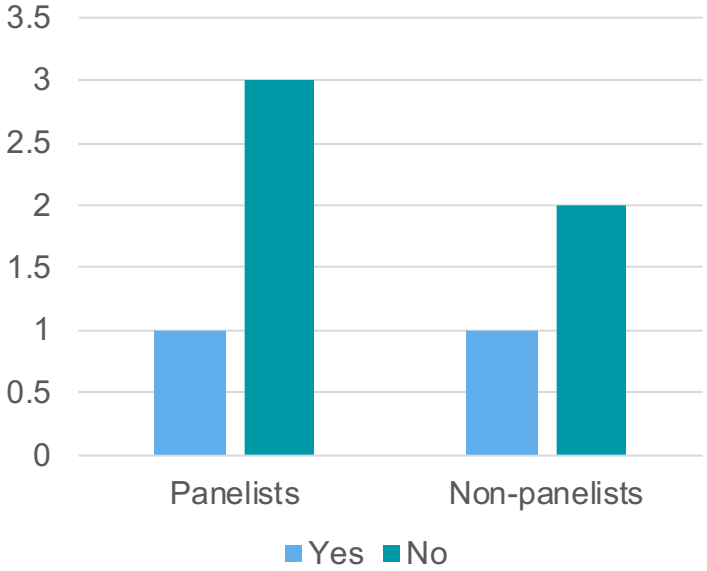
Jasmine

A bit about the panelists

1. Your role?

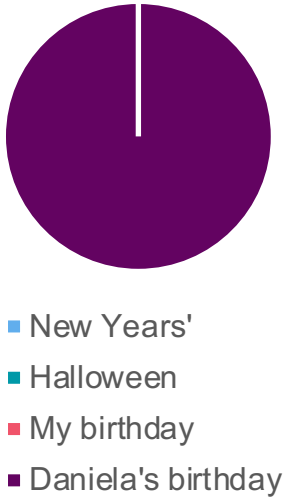


2. Did you make a COVID viz in 2020?

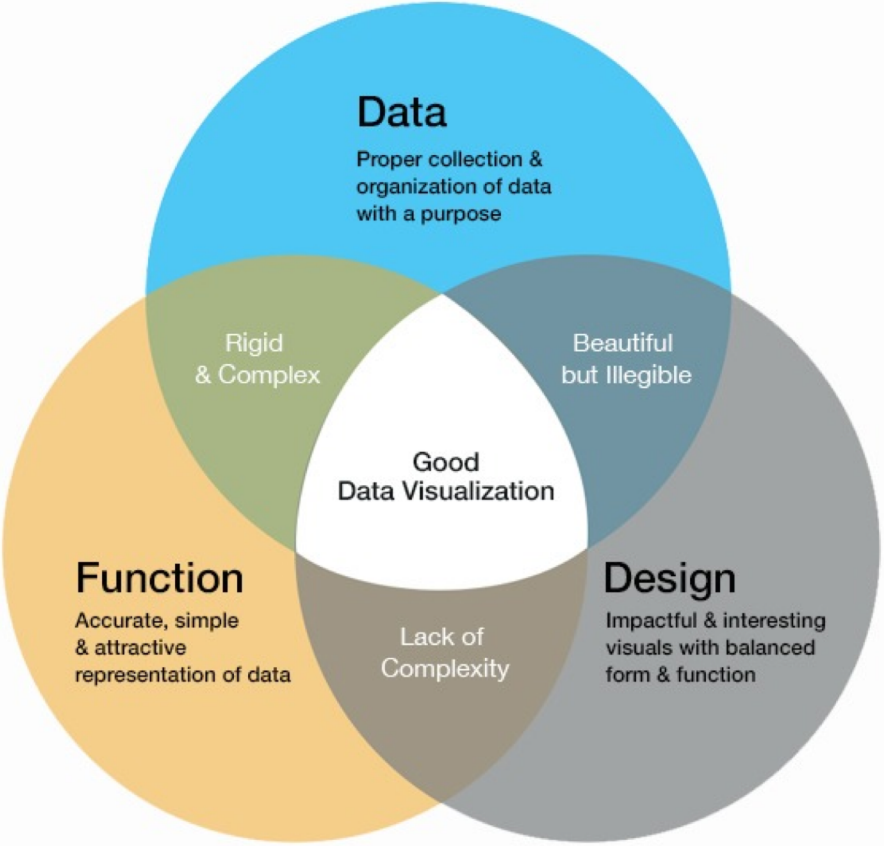


3. Your favourite holiday?

- ☐ New Years’
- ☐ Halloween
- ☐ My birthday

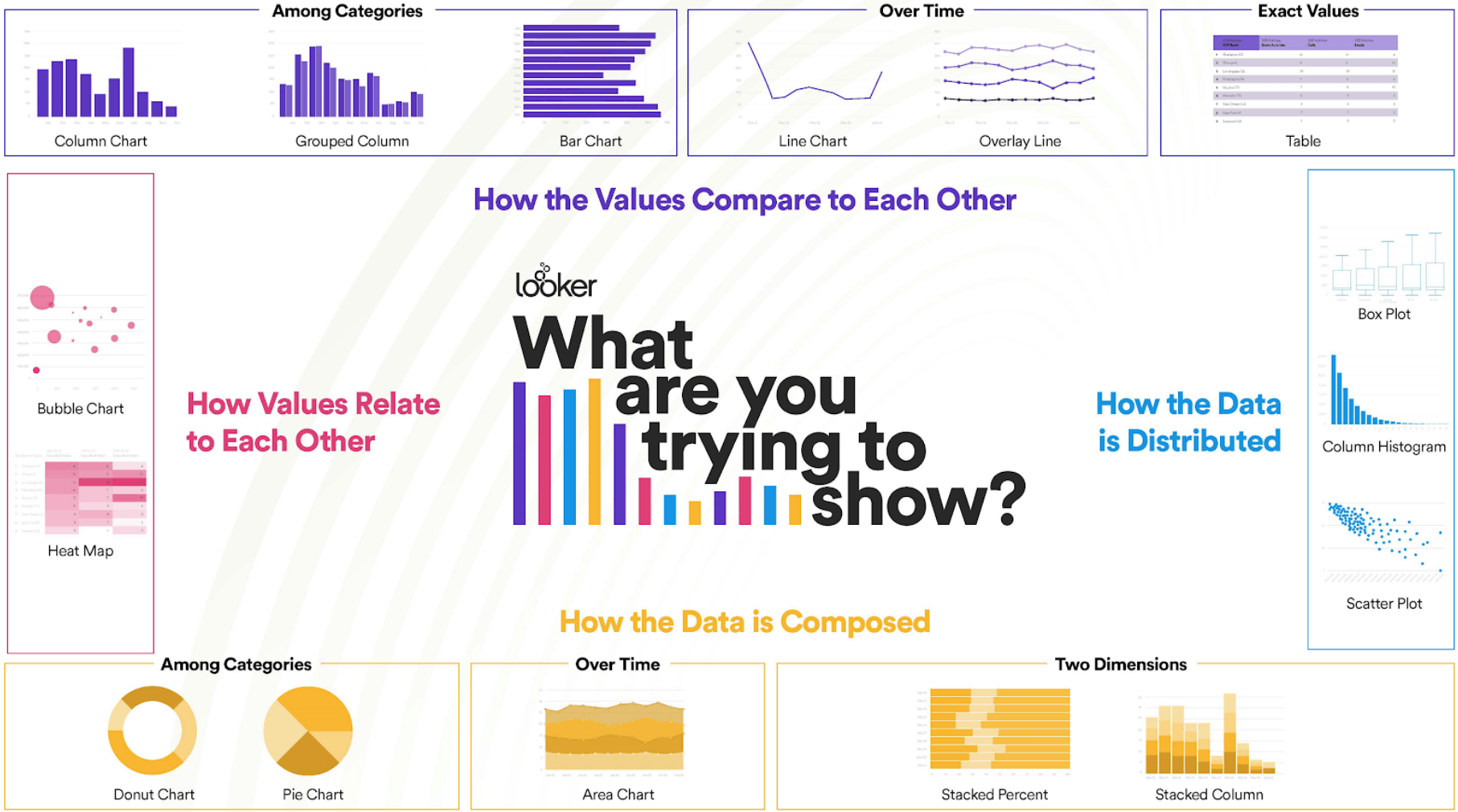


What makes a good visualization?



https://www.nabler.com/articles/the-data-visualization-manual-that-you-will-want-to-bookmark/

https://golagoon.com/how-do-i-make-a-google-spreadsheet/how-to-make-a-column-or-row-visible-always-in-google-spreadsheets-do-i-spreadsheet-fit-on-o/

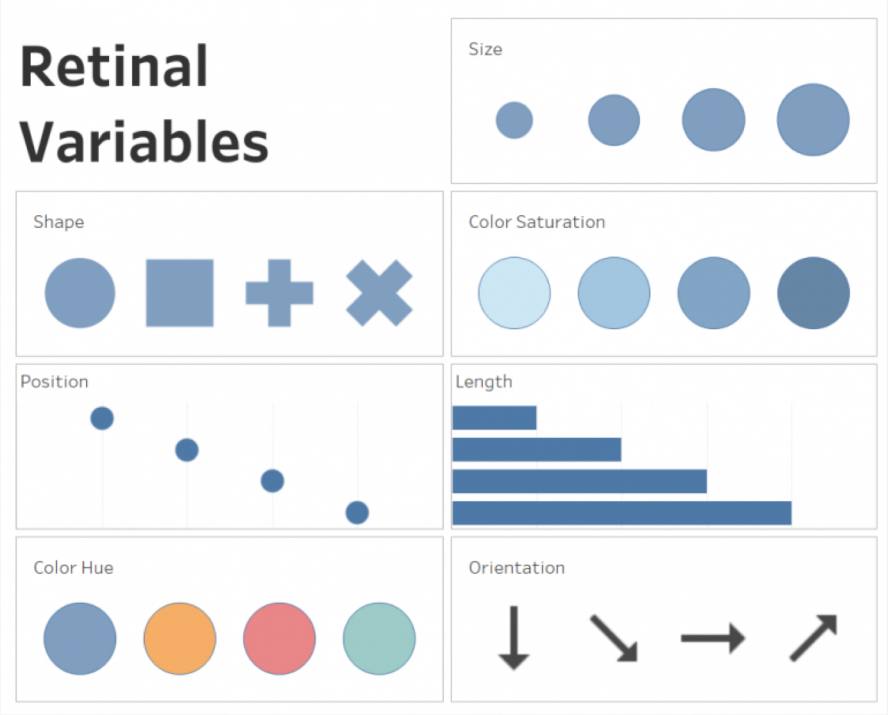


https://looker.com/blog/different-types-graphs-charts-uses

Visual encoding

- Visual encoding is converting a visual image to understand it as an object.
- Encoding of retinal variables to become visually informative.
- What are we able to decode most accurately?

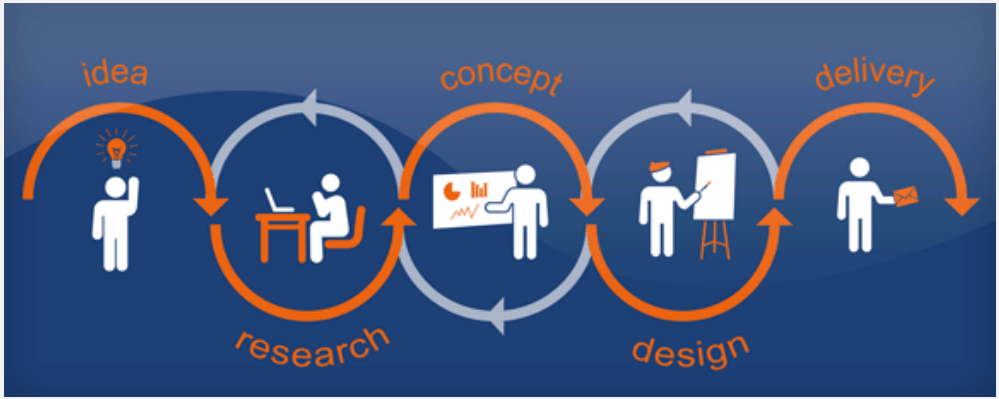
  Position, length, and size



https://vizpilgrim.com/2019/11/01/visual-encoding/

Graphic design

- Creating visual communications to convey a message effectively
- Use of design tools like Adobe Illustrator.
- Best practices, design theme, creativity, accuracy, and accessibility

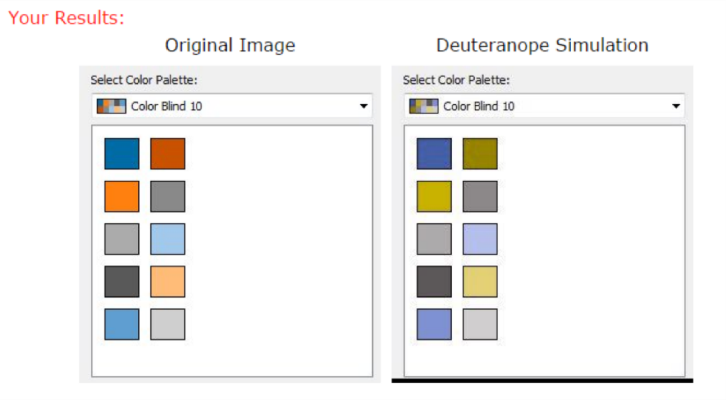


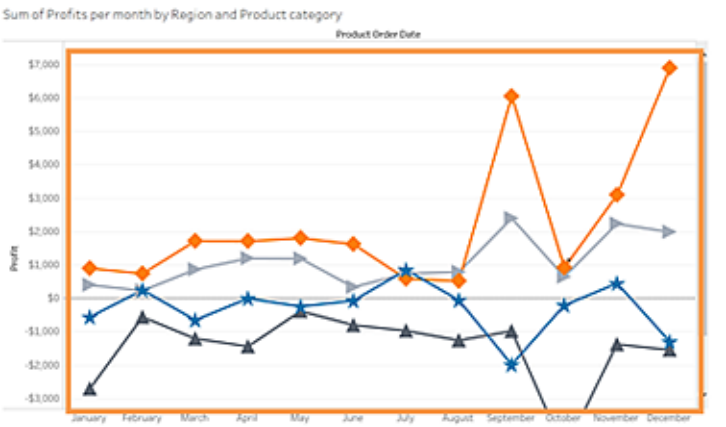
https://www.howdodesign.com/the-graphic-design-process-getting-ready/, https://help.tableau.com/current/pro/desktop/en-us/accessibility_best_practice.htm

Accessibility

- Keep it simple.
- Use color-blind palettes to create visualization.
- Use a color Blindness Simulator
- For line marks, use additional options like shapes, size, and labels.
- Image infographics require an alternative text description in the form of a transcript.

E.g.: Transcript can be displayed below an image infographic.

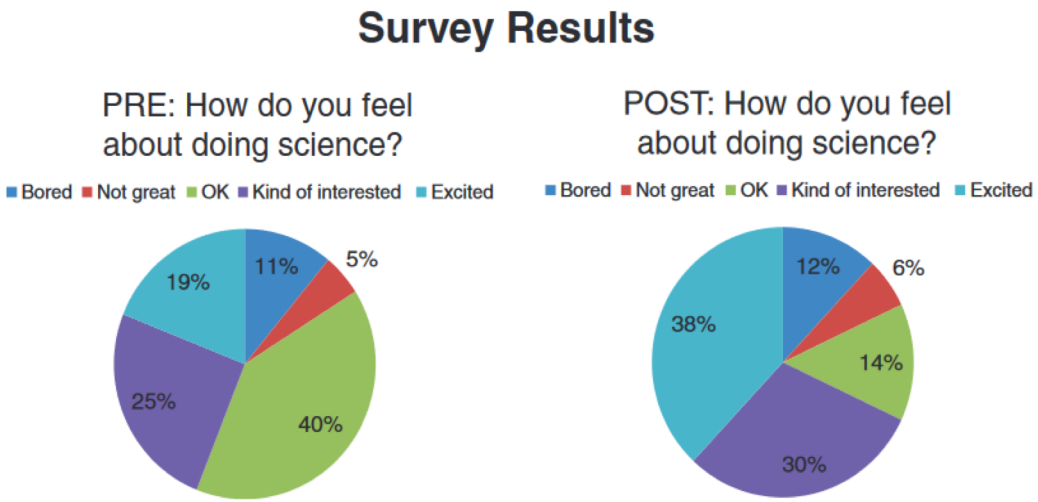


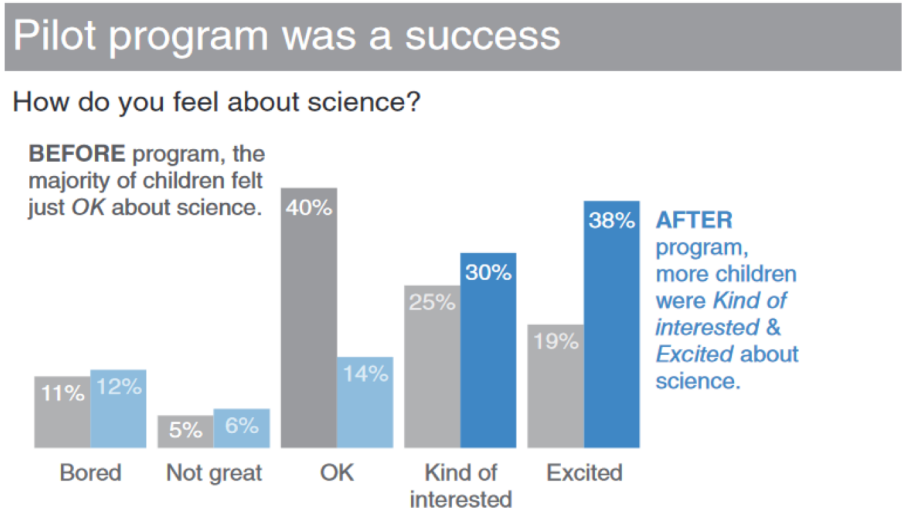


https://help.tableau.com/current/pro/desktop/en-us/accessibility_best_practice.htm, https://www.csun.edu/universal-design-center/accessible-infographics, https://www.tableau.com/about/blog/2016/4/examining-data-viz-rules-dont-use-red-green-together-53463, https://www.color-blindness.com/coblis-color-blindness-simulator/

Before
Difficult to gain insights
Reasons: pie charts, color, legends

After
A successful visualization!!
Reasons: bar chart, color, clear text





Knaflic, Cole Nussbaumer. *Storytelling with Data : A Data Visualization Guide for Business Professionals*. Hoboken, New Jersey: Wiley, 2015. Web.

Use highlighting and colors

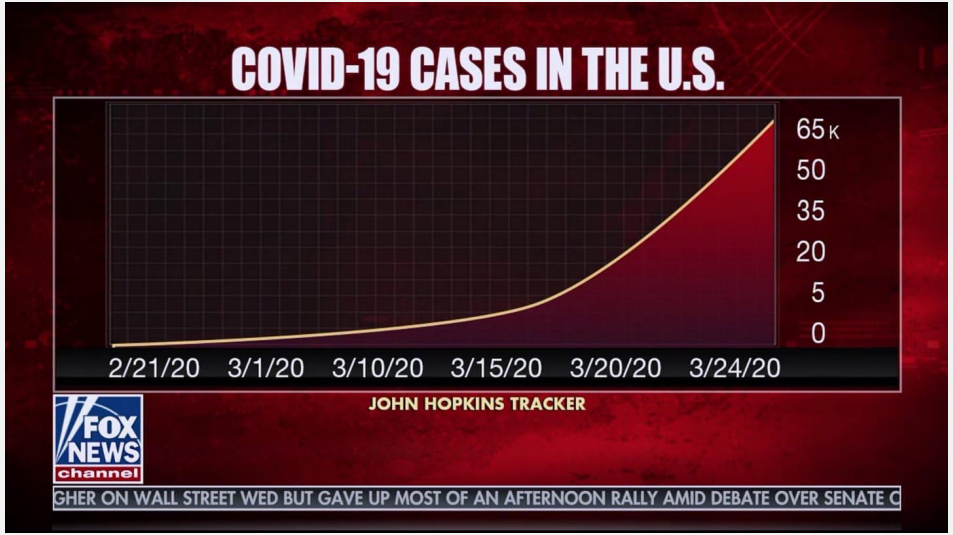
Region	January	February	March	April
South	34.00	20.00	40.00	60.00
North	20.00	35.00	30.00	30.00
West	26.00	18.00	-3.00	8.00
East	9.00	-5.00	15.00	10.00

Region	January	February	March	April
South	34.00	20.00	40.00	60.00
North	20.00	35.00	30.00	30.00
West	26.00	18.00	-3.00	8.00
East	9.00	-5.00	15.00	10.00

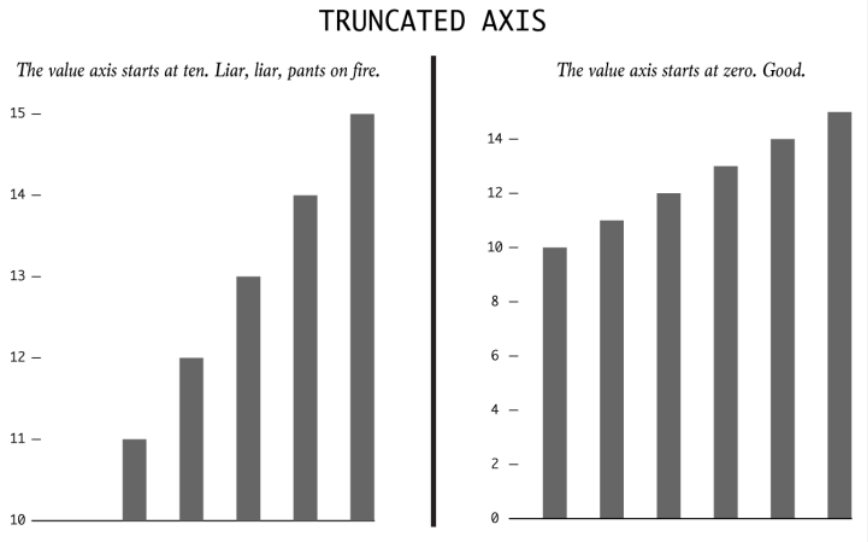
Profit % for regions from January - April

What leads to misleading visualizations?

Manipulating the Y-axis



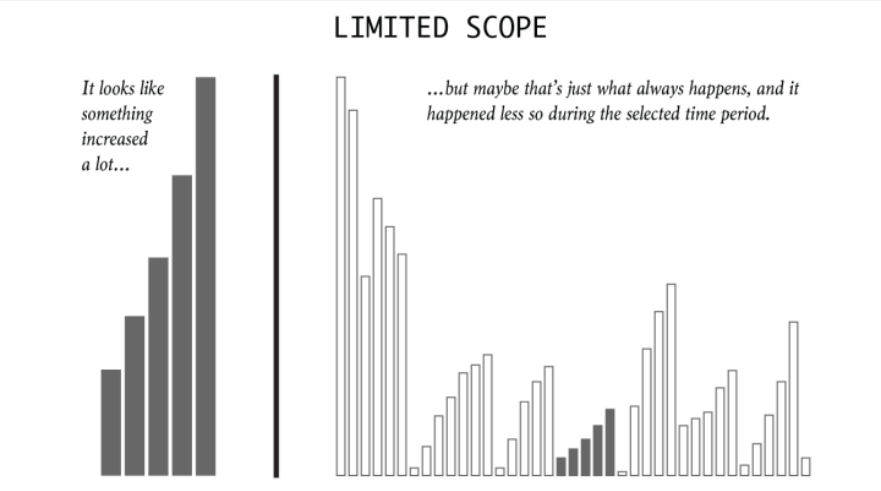
Truncated axes - omitting the baseline



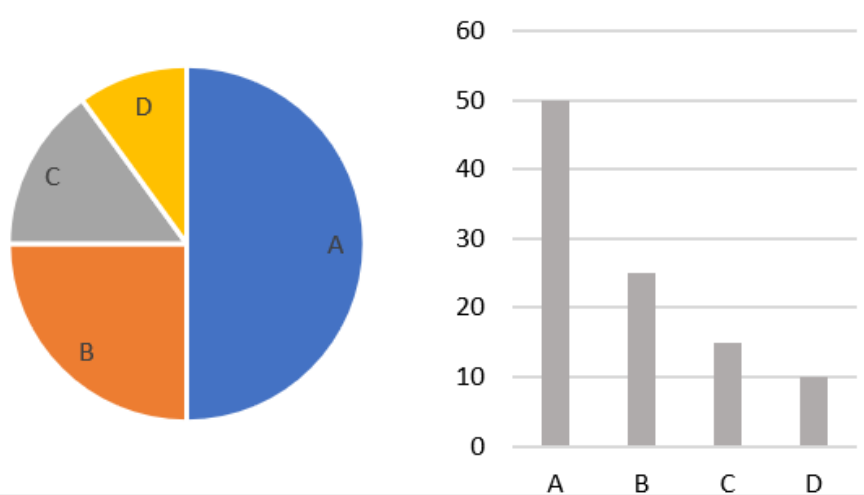
https://venngage.com/blog/misleading-graphs/, https://flowingdata.com/2017/02/09/how-to-spot-visualization-lies/

What leads to misleading visualizations?

Omitting data - trends can be misleading



Choosing the wrong graph



https://flowingdata.com/2017/02/09/how-to-spot-visualization-lies/, https://www.boldbi.com/blog/your-pie-chart-questions-answered-here

Key topics in data visualization

Cameron Riopelle, PhD
Head of Data Services
University of Miami Libraries

Ethos

Graphical excellence is:

Nearly always multivariate.

That which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.

The well-designed presentation of interesting data-a matter of substance, of statistics, and of design

(Edward Tufte, The Visual Display of Quantitative Information, 51)

To make good visualizations, you need to understand not just your audience but the **structure of your data.**

Know Your Data

What is the file format?

Is your data clean?

What is the appropriate measure of central tendency?

Measure of dispersion?

Is your data skewed?

What is the level of measurement - nominal, ordinal, interval/ratio?

Do you need to aggregate before visualizing? If you aggregate, do you lose complexity or over-simplify?

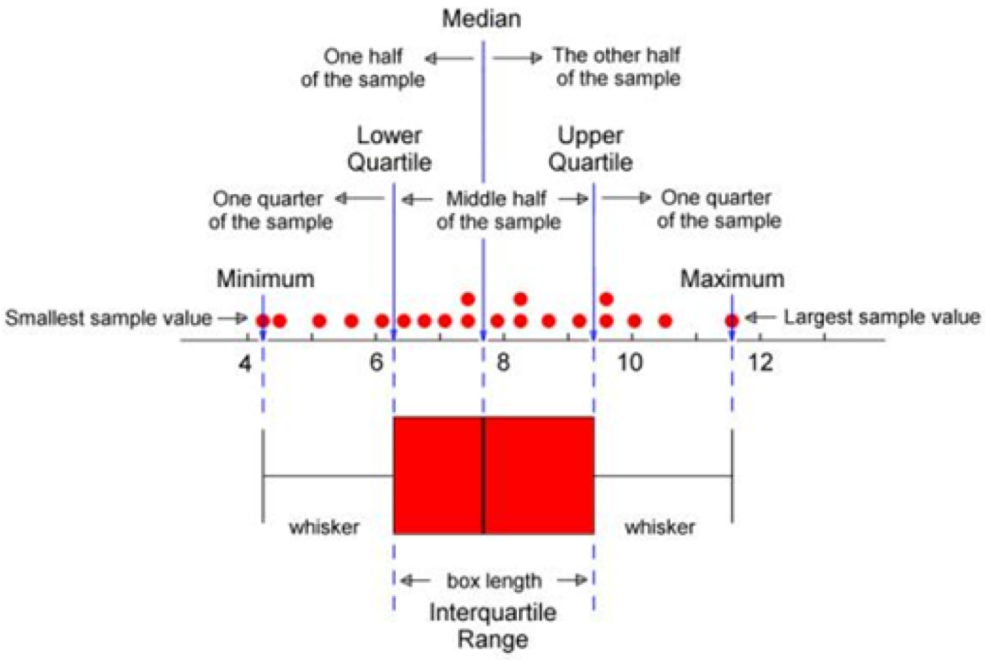


Image from http://web.pdx.edu/~stipakb/download/PA551/boxplot.html

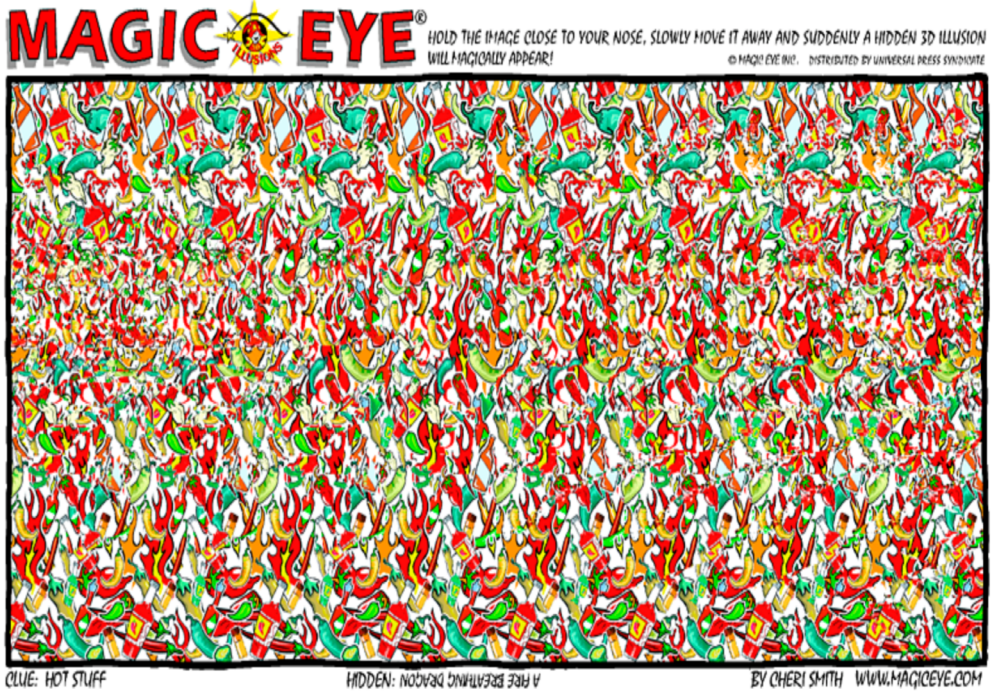
Exploratory vs. Confirmatory Data Analysis

Exploratory: Looking for patterns and trends

"The greatest value of a picture is when it forces us to notice what we never expected to see" (Tukey, Exploratory Data Analysis)

Patterns can be misleading. We identify patterns when there aren’t any.

Confirmatory: Hypothesis Testing



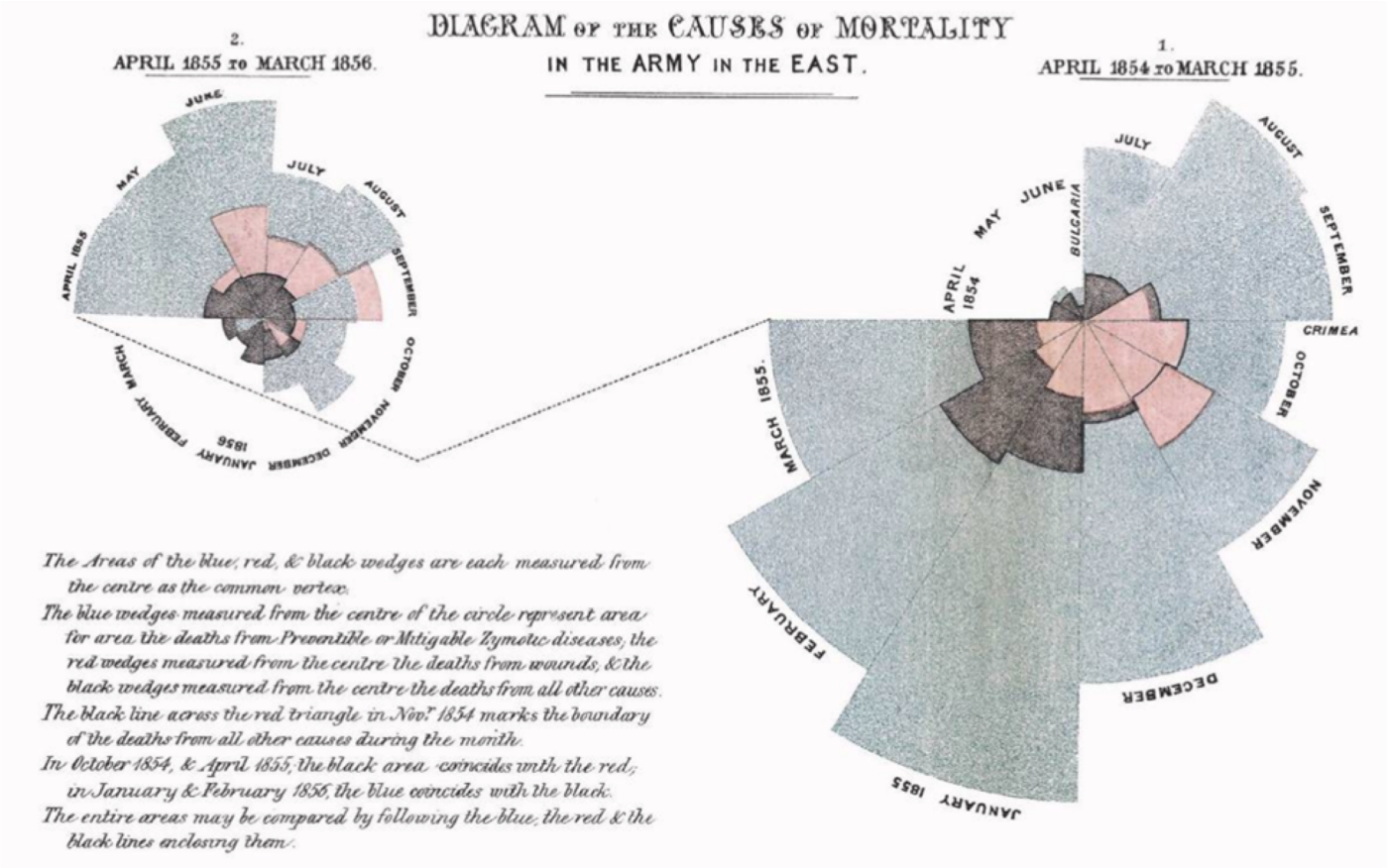
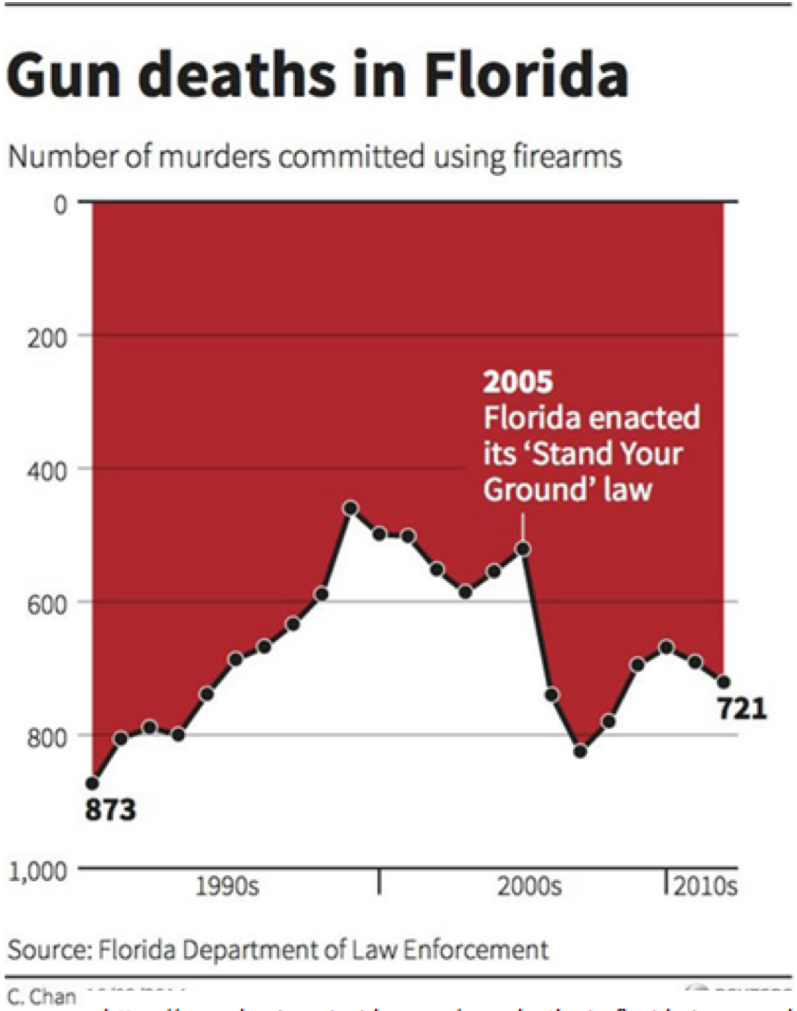


Figure I.1
Diagram of the Causes of Mortality in the Army in the East.
Florence Nightengale.

Lankow, Crooks, and Ritchie (2014): 15



https://www.businessinsider.com/gun-deaths-in-florida-increased-with-stand-your-ground-2014-2

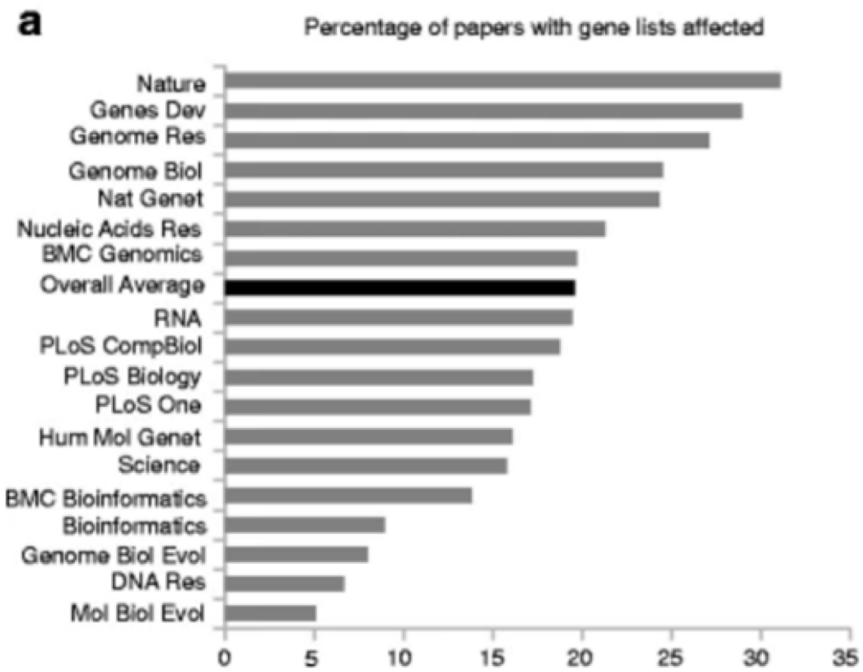
When should you use scripts?

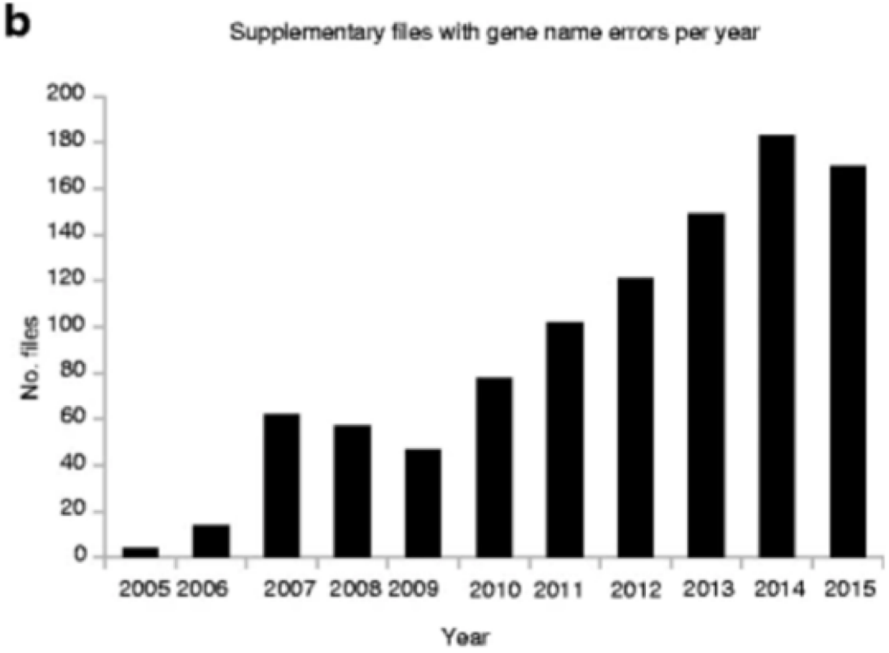
	B	C	I	J	K	L	M
2				Real GDP growth			
3				Debt/GDP			
4	Country	Coverage	30 or less	30 to 60	60 to 90	90 or above	30 or less
26			3.7	3.0	3.5	1.7	5.5
27	Minimum		1.6	0.3	1.3	-1.8	0.8
28	Maximum		5.4	4.9	10.2	3.6	13.3
29							
30	US	1946-2009	n.a.	3.4	3.3	-2.0	n.a.
31	UK	1946-2009	n.a.	2.4	2.5	2.4	n.a.
32	Sweden	1946-2009	3.6	2.9	2.7	n.a.	6.3
33	Spain	1946-2009	1.5	3.4	4.2	n.a.	9.9
34	Portugal	1952-2009	4.8	2.5	0.3	n.a.	7.9
35	New Zealand	1948-2009	2.5	2.9	3.9	-7.9	2.6
36	Netherlands	1956-2009	4.1	2.7	1.1	n.a.	6.4
37	Norway	1947-2009	3.4	5.1	n.a.	n.a.	5.4
38	Japan	1946-2009	7.0	4.0	1.0	0.7	7.0
39	Italy	1951-2009	5.4	2.1	1.8	1.0	5.6
40	Ireland	1948-2009	4.4	4.5	4.0	2.4	2.9
41	Greece	1970-2009	4.0	0.3	2.7	2.9	13.3
42	Germany	1946-2009	3.9	0.9	n.a.	n.a.	3.2
43	France	1949-2009	4.9	2.7	3.0	n.a.	5.2
44	Finland	1946-2009	3.8	2.4	5.5	n.a.	7.0
45	Denmark	1950-2009	3.5	1.7	2.4	n.a.	5.6
46	Canada	1951-2009	1.9	3.6	4.1	n.a.	2.2
47	Belgium	1947-2009	n.a.	4.2	3.1	2.6	n.a.
48	Austria	1948-2009	5.2	3.3	-3.8	n.a.	5.7
49	Australia	1951-2009	3.2	4.9	4.0	n.a.	5.9
50							
51			4.1	2.8	2.8	=AVERAGE(L30:L44)	

Economic Policy

Is the evidence for austerity based on an Excel spreadsheet error?

https://www.washingtonpost.com/news/wonk/wp/2013/04/16/is-the-best-evidence-for-austerity-based-on-an-excel-spreadsheet-error/





Nearly 20% of scientific papers that used Excel for genomic data analysis contain errors related to Excel's automatic date format conversion (Ziemann, Eren, and El-Osta, 2016)

Article and images from https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1044-7

Takeaways

Small ideas:

Visualizations should appeal to different types of learners

Labels are helpful

Colorblindedness

Color can be used to accentuate key points

Standardize units and axes when possible

Big ideas:

Don’t lose track of your key message

Pilot your visualization

“Graphical excellence requires telling the truth about the data.” (Tufte, 51)

Bringing Data Visualization Instruction to Your Library

Jasmine Kirby

Outline

- Data literacy as part of information literacy
- Interdisciplinarity
- Providing training
- Managing expectations

Data literacy as part of information literacy

Spotlight on the framework

- Can teach all of the ACRL framework but especially:
  - Information Creation as a Process

"Framework for Information Literacy for Higher Education", American Library Association, February 9, 2015. http://www.ala.org/acrl/standards/ilframework (Accessed February 3, 2021)
Document ID: b910a6c4-6c8a-0d44-7dbc-a5dcbd509e3f

Building good habits and expectations for data visualization early on

- Letting people know they have data to visualize
- Teaching people how to find and cite datasets and how to critically evaluate these sources

Story time: **What to do when finding the perfect dataset seems inconceivable…**

- One shot data literacy program for upperclassmen in a statistics course, originally an in person session now a LibWizard tutorial
- Based on finding data on the countries and rodents of unusual size from the movie *The Princess Bride*
- Encourages students to take a more critical view of data and data visualizations they will encounter while researching

Interdisciplinarity

Library as a truly interdisciplinary space

- Many of us find ourselves liaising for a bunch of unrelated subjects
- We know about and study different research methodologies

Tools

- We familiarize ourselves with the latest software tools to support our departments and in our own research
- People don’t always use the right tool for what they are trying to do
- Getting people to ask if what they are doing requires that tool is important

Yes you can provide training

Even when you are also learning

Higher education: the learning never stops

- A lot of what we offer is lifelong learning and training and professional development, self-directed learning
- So many free open resources available online

Getting people to try new things, together

- Hosting workshops that bring together people from a variety of disciplines and backgrounds

- Advertising software tools and learning resources available through the library that people may not have known about

Story time: Slytherin House Initiative at Iowa State University

- A book club about learning Python open to the entire library that took over a computer lab on Friday afternoons
- Also brought Software Carpentry to campus for the library
- Resulted in a couple of projects based on using Python to solve various library issues

Managing Expectations

Improving awareness of how practical stuff is by having people try to do it

- Main goal of Slytherin house was to show people what was possible with using Python and bring people from different parts of the library together to cross pollinate knowledge

- Skunkworks and making space for people to figure out what is a good or a bad idea before getting too invested

Morgan, P. (2020, January 14). Reconsidering Skunkworks. *Paige Morgan Blog*. http://blog.paigemorgan.net/articles/20/reconsidering-skunkworks.html

Nowviskie, B. (2013). Skunks in the Library: A Path to Production for Scholarly R&D. *Journal of Library Administration*, *53*(1), 53–66. https://doi.org/10.1080/01930826.2013.756698

Story time: ArcGIS Story Maps workshop

- Part of a GIS day celebration in collaboration with campus and local GIS user groups
- Chosen because it is a very easy to use tool with excellent documentation
- Help people learn how much GIS they can realistically learn in a one hour library workshop
- Had to switch to Esri ArcGIS Story Maps because Knight Lab's StoryMap JS did not offer any accessibility features or workarounds at the time. You have the power.

Don’t let data visualization out of the purview of library instruction!