Choosing the Right Chart

Two Questions

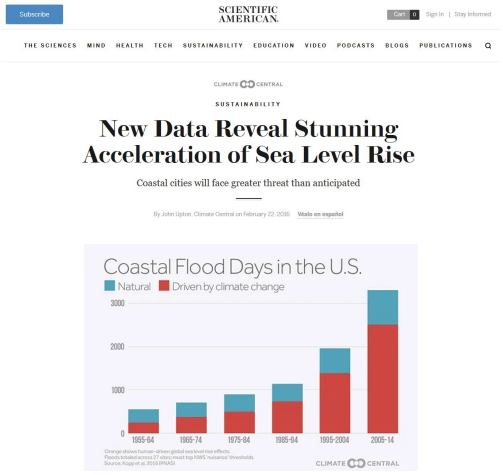
Who is your audience/what do they want?

Consider both the technical and critical thinking skills of your audience.

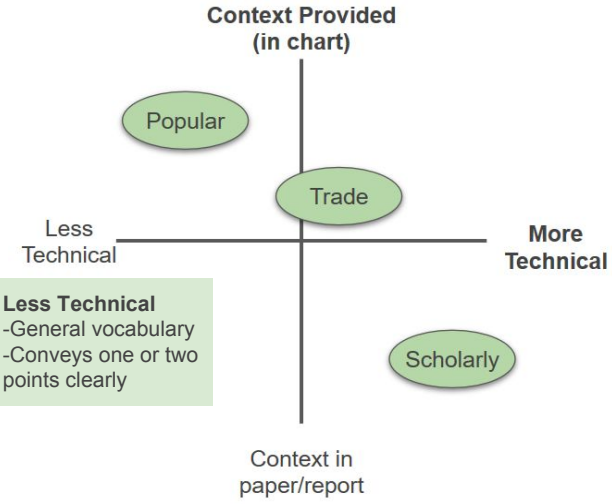
What information do you want convey?

Use data visualization to: inform, compare, show change, organize, and reveal relationships.

Is this Your Audience?



Context Provided
-Necessary information can be conveyed in chart.
-Information is simple and brief, conclusion is clear.



Less Technical
-General vocabulary
-Conveys one or two points clearly

https://www.scientificamerican.com/article/new-data-reveal-stunning-acceleration-of-sea-level-rise/

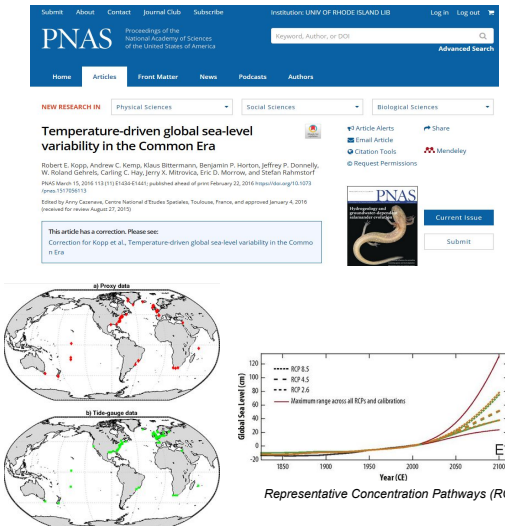
Less Technical

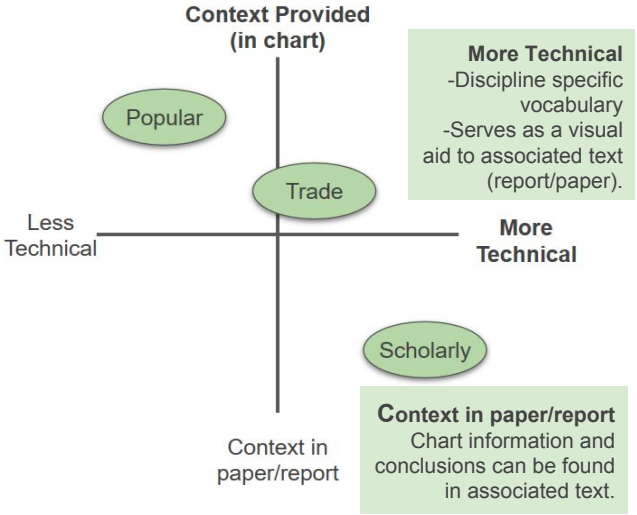
- General vocabulary
- Conveys one or two points clearly

Context Provided

- Necessary information can be conveyed in chart.
- Information is simple and brief, conclusion is clear.

Or is This?





https://www.pnas.org/content/113/11/E1434/tab-figures-data

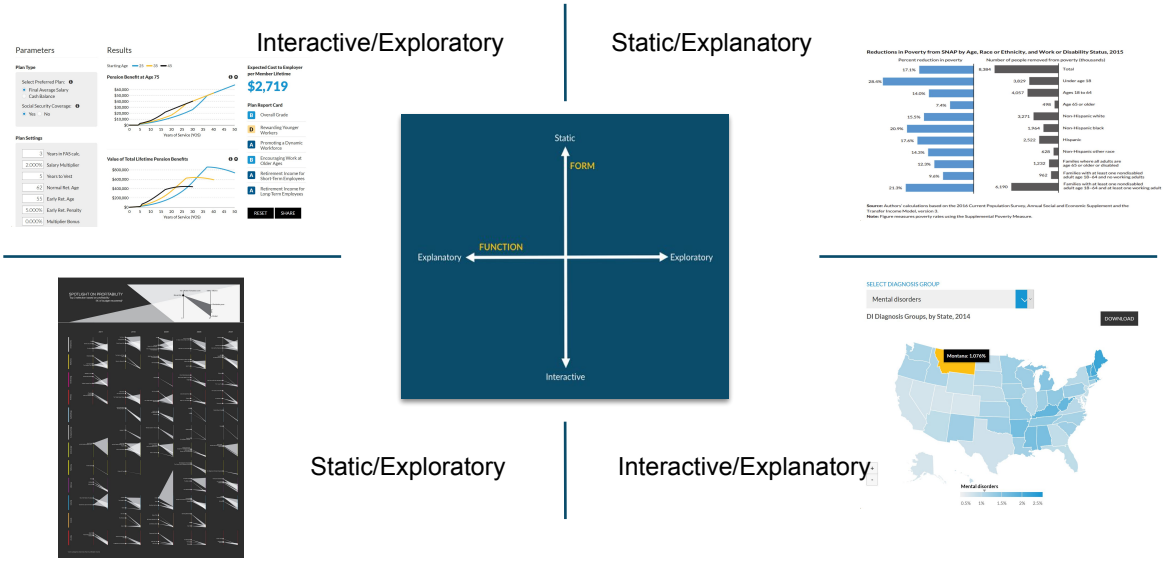
More Technical

- Discipline specific vocabulary
- Serves as a visual aid to associated text (report/paper).

Context in paper/report

- Chart information/conclusions can be found in associated text (report/paper).

What does your audience want?



Explanatory/Exploratory - Are they looking to clarify information provided in the text? Or, is the information new to them.
Static/Interactive - Do they want you to make the connections and do they want to see how the information relates on their own?

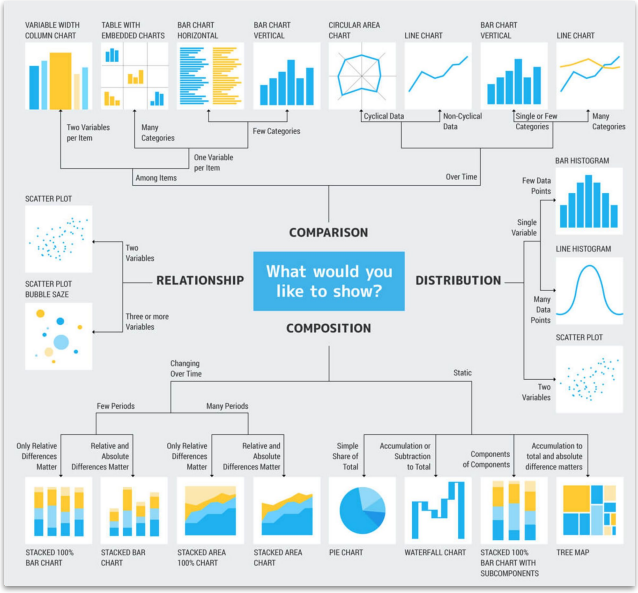
Central chart- https://medium.com/@urban_institute/form-and-function-let-your-audiences-needs-drive-your-data-visualization-choices-3c0603745822
Interactive/Explanatory - https://www.urban.org/features/11-charts-about-social-security-disability-insurance-program
Interactive/Exploratory - http://apps.urban.org/features/build-your-own-pension/
Static/Explanatory - https://www.urban.org/urban-wire/what-does-it-mean-snap-remove-people-poverty
Static/Exploratory - http://krisztinaszucs.com/

What information do you want to convey?

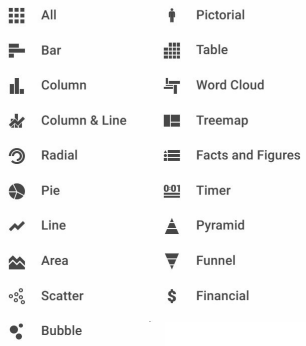
- **Inform**: convey a single important message or data point that doesn't require much context to understand
- **Compare**: show similarities or differences among values or parts of a whole
- Show **Change:** visualize trends over time or space
- **Organize:** show groups, patterns, rank or order
- **Reveal Relationships:** show correlations among variables or values

https://venngage.com/blog/how-to-choose-the-best-charts-for-your-infographic/

https://venngage.com/blog/how-to-choose-the-best-charts-for-your-infographic/



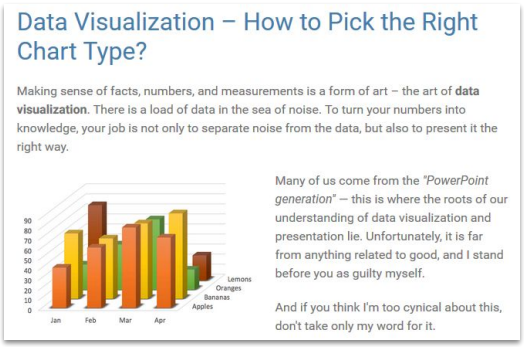
https://www.crazyegg.com/blog/wp-content/uploads/2016/12/what-to-show-chart1.jpg



Examples of each of these types of charts can be found at **Infogram.com**

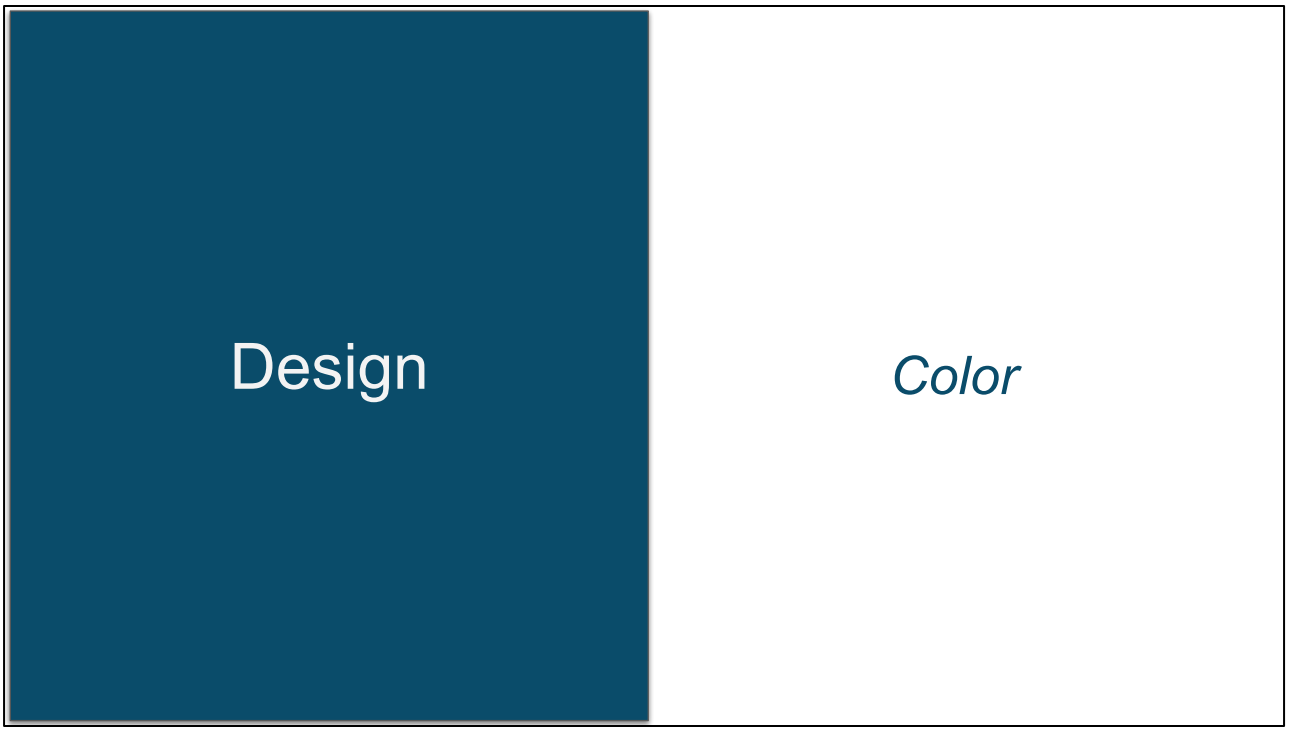
https://infogram.com/examples

Not picutured - https://serialmentor.com/dataviz/directory-of-visualizations.html



https://eazybi.com/blog/data_visualization_and_chart_types/

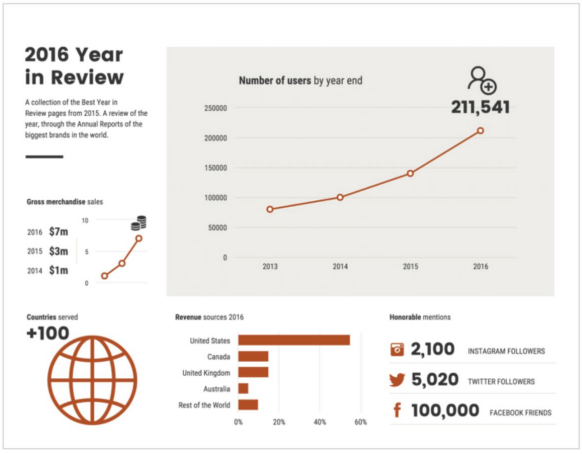
https://visme.co/blog/types-of-graphs/


Design
Color

Use color for function, not decoration

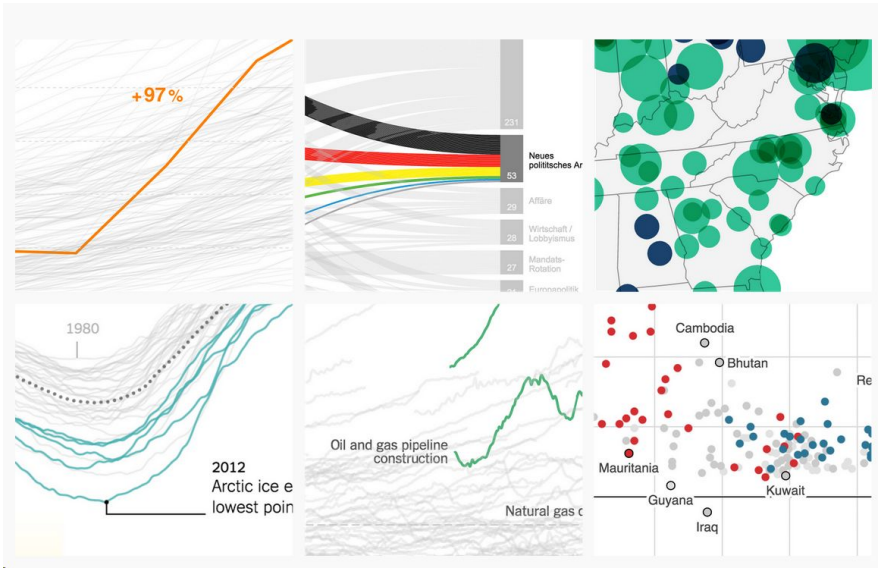
Color use plays a significant role in how “busy” your infographic will appear.

To make sure things don’t get too messy, don’t use more than six colors in any gi chart. When you’ve got more than one chart in an infographic, use those colors consistently across all of your charts.



https://venngage.com/blog/how-to-choose-the-best-charts-for-your-infographic/

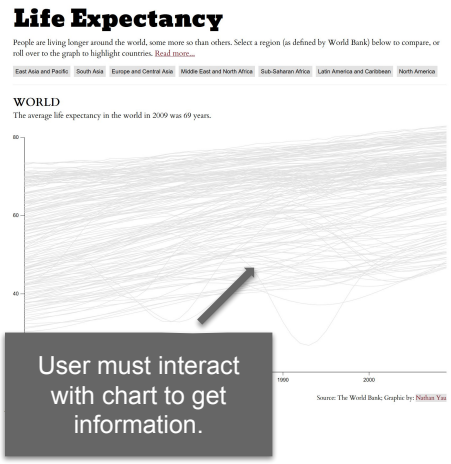
Use color to highlight and differentiate information



https://blog.datawrapper.de/better-charts/

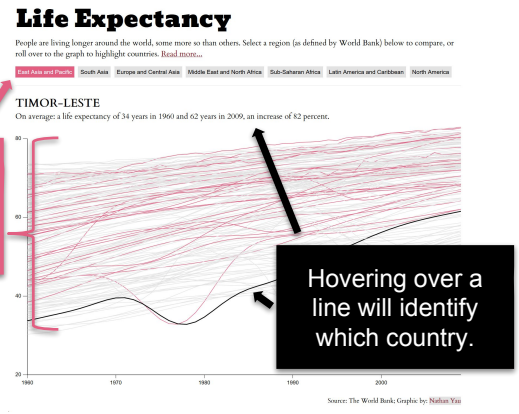
Color and Interactivity

*Color is used to **convey** information.*



User must interact with chart to get information.

Selecting a region will highlight those countries in that location.

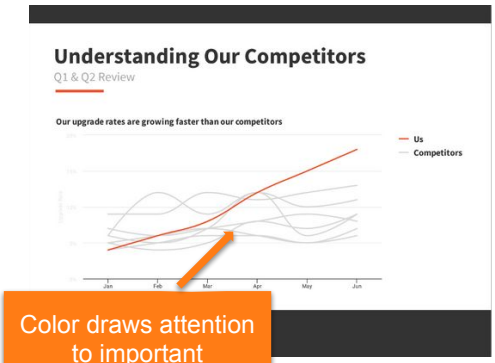


Hovering over a line will identify which country.

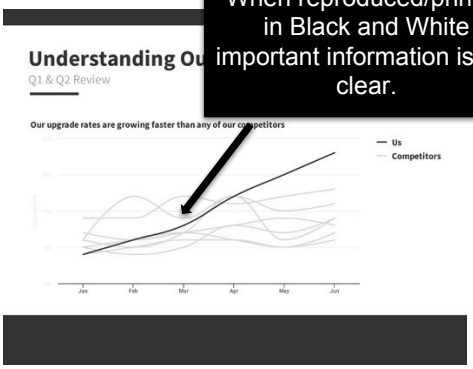
http://projects.flowingdata.com/life-expectancy/

Color to Draw Attention

*Color is used to **highlight** information.*



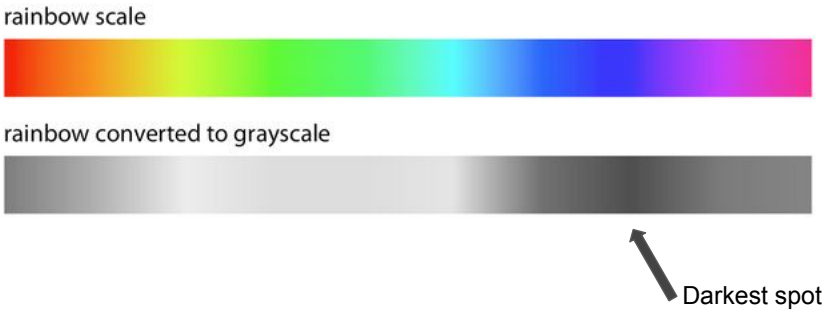
Color draws attention to important information.



When reproduced/printed in Black and White important information is still clear.

https://venngage.com/blog/how-to-choose-the-best-charts-for-your-infographic/

This is how color is reproduced in black and white.



https://serialmentor.com/dataviz/color-pitfalls.html

Design

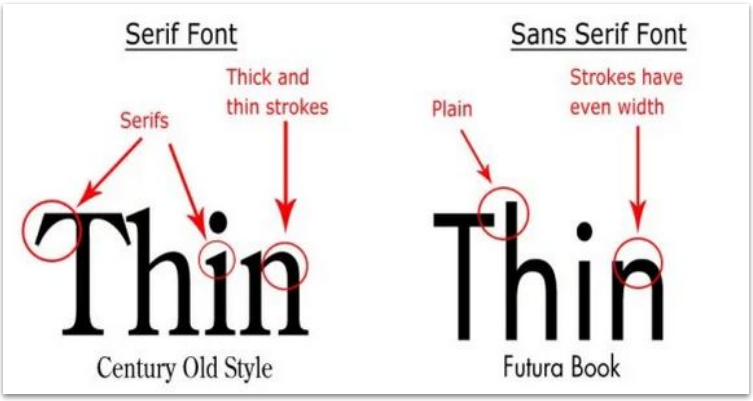
Text

Titles, Axes, Legends and Annotations

Font

Use a **Serif** font for:

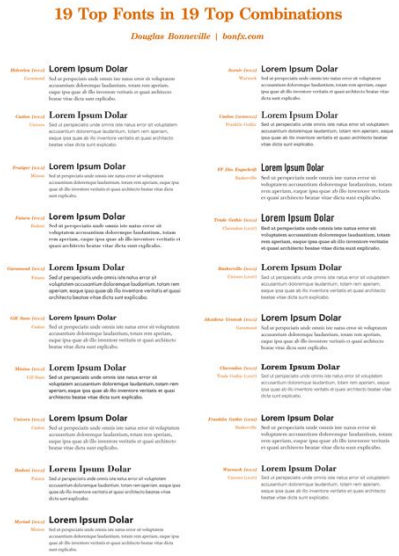
- main text



Use a **Sans Serif** font for:

- titles
- headings
- tables, and
- figures

http://hypsypops.com/fonts-wisdom-text-can-make-break-graph/



Don’t use more than two fonts in a document, make sure they match.

Use different **weights** and *italics* for emphasis, don’t underline.

https://bonfx.com/19-top-font-in-19-top-combinations/

https://macademic.org/2013/11/25/typography_essentials/

Title

Tell the viewer what the chart is about, don't over explain.

Example 1
Here's the Composition of the User Acquisition Pattern Between the Months of January 2016 and December 2016 with a Focus on the Share of Display Ads

vs

Example 2
User Acquisition by Channel (2016)

Include units of representation, measurement and time periods.

Profits in 2016 (Jan. – Dec.)
In millions (USD)

Avoid using a, an, (or) the.

Example 1
The Average Revenue per Each of the Days in September 2015

vs

Example 2
Daily Revenue (September 2015)

Avoid adjectives.

Example 1
The Massive Daily Revenue We Pulled in September 2015 – Our Most Amazing Month

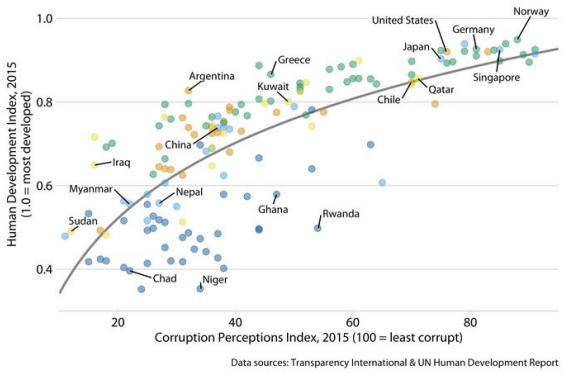
vs

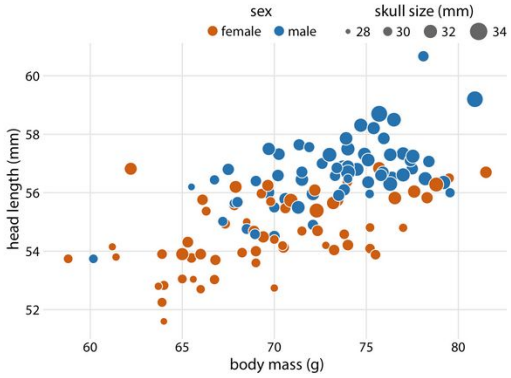
Example 2
Daily Revenue (September 2015)

https://www.fusioncharts.com/resources/charting-best-practices/5-tips-for-writing-chart-captions

https://www.anychart.com/blog/2017/04/05/chart-captions-title-graph-tips/

Axes



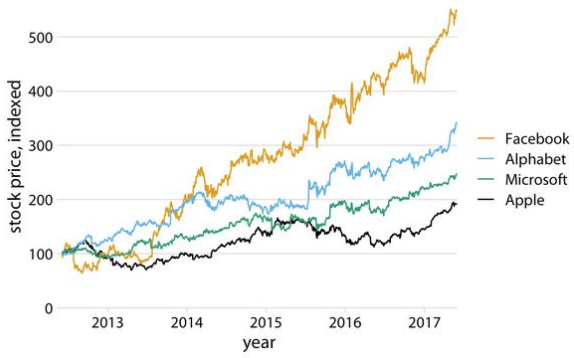


Indicates what the axes represents and the unit of measurement.

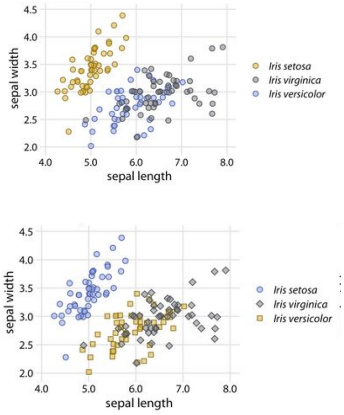
Corruption and Skull size tables - https://serialmentor.com/dataviz/figure-titles-captions.html#figure-titles-and-captions

Legends

Keep visual order of the chart (if the chart has a visual order).



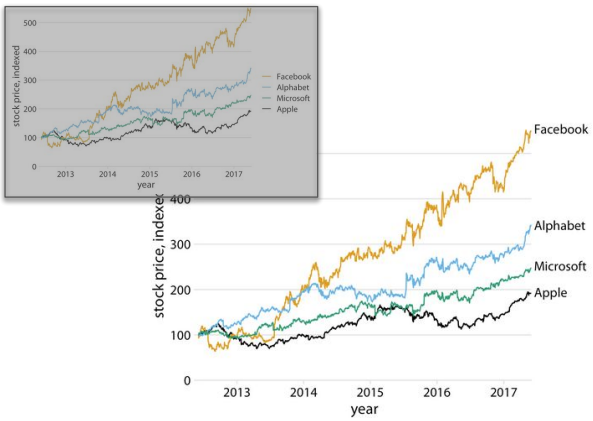
Assign symbols as well as color.

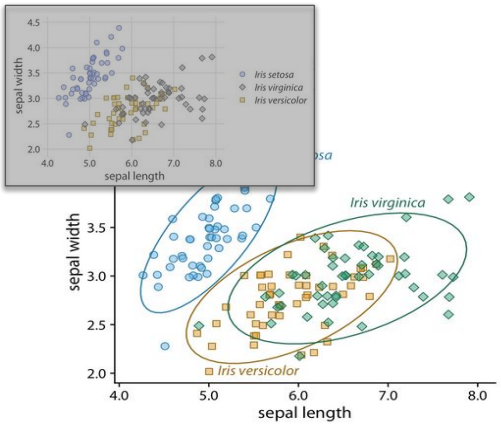


https://serialmentor.com/dataviz/redundant-coding.html#designing-legends-with-redundant-coding

No Legends

Fundamentals of Data Visualization and multiple other sources recommend labeling data directly and avoiding legends when possible.

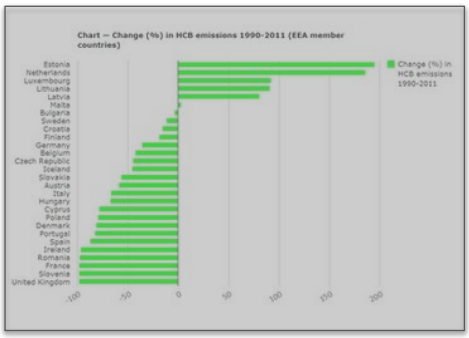


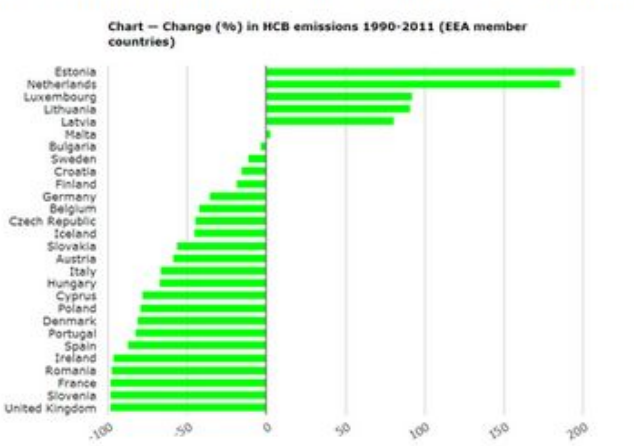


https://serialmentor.com/dataviz/redundant-coding.html#designing-legends-with-redundant-coding

No Legends

If there is only one category a legend isn't needed.

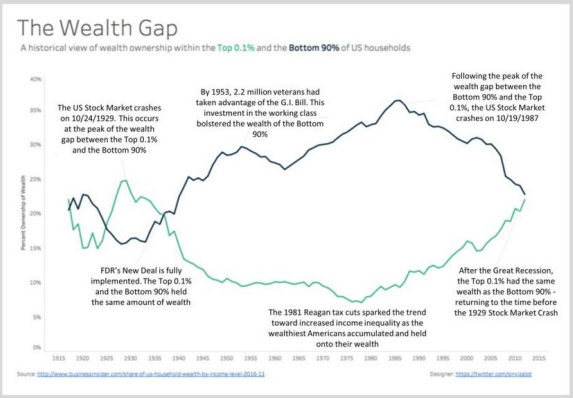




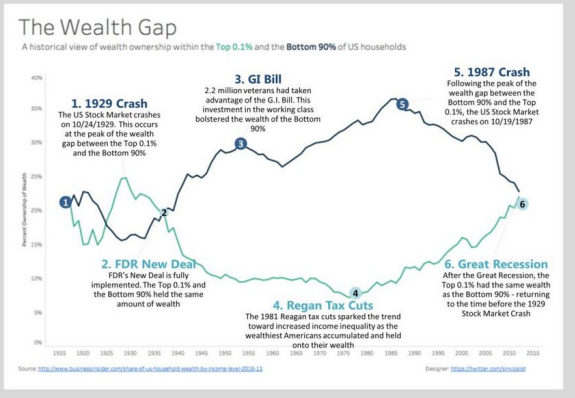
https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts

Annotations

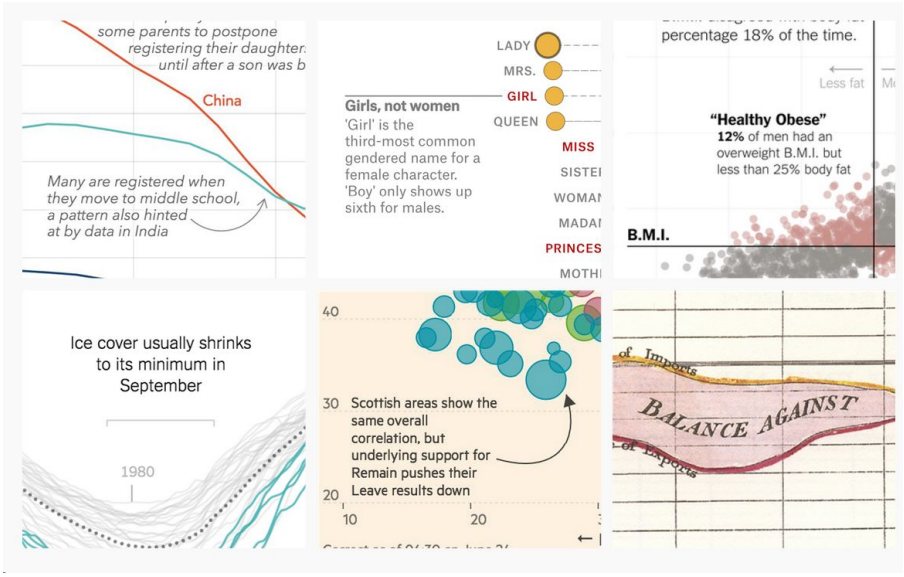
Original Before Makeover



Added 1) Signposts 2) Color Matching 3) Numbers (optional)



http://speakingppt.com/10-rules-for-graph-annotations/



Some other annotation examples

Design

Layout

Order, labels, and presentation

Table format

a ugly

Rank	Title	Amount
1	Star Wars	$71,565,498
2	Jumanji	$36,169,328
3	Pitch Perfect 3	$19,928,525
4	Greatest Showman	$8,805,843
5	Ferdinand	$7,316,746

b ugly

Rank	Title	Amount
1	Star Wars	$71,565,498
2	Jumanji	$36,169,328
3	Pitch Perfect 3	$19,928,525
4	Greatest Showman	$8,805,843
5	Ferdinand	$7,316,746

c

Rank	Title	Amount
1	Star Wars	$71,565,498
2	Jumanji	$36,169,328
3	Pitch Perfect 3	$19,928,525
4	Greatest Showman	$8,805,843
5	Ferdinand	$7,316,746

d

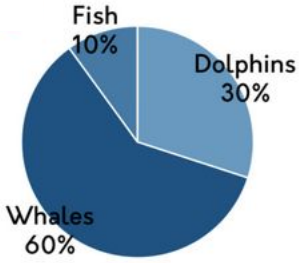
Rank	Title	Amount
1	Star Wars	$71,565,498
2	Jumanji	$36,169,328
3	Pitch Perfect 3	$19,928,525
4	Greatest Showman	$8,805,843
5	Ferdinand	$7,316,746

Figure 22.7: Examples of poorly and appropriately formatted tables, using the data from Table 6.1 in Chapter 6. (a) This table violates numerous conventions of proper table formatting, including using vertical lines, using horizontal lines between data rows, and using centered data columns. (b) This table suffers from all problems of Table (a), and in addition it creates additional visual noise by alternating between very dark and very light rows. Also, the table header is not strongly visually separated from the table body. (c) This is an appropriately formatted table with a minimal design. (d) Colors can be used effectively to group data into rows, but the color differences should be subtle. The table header can be set off by using a stronger color. Data source: Box Office Mojo (http://www.boxofficemojo.com/). Used with permission

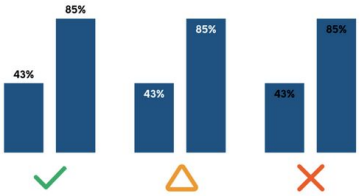
https://serialmentor.com/dataviz/figure-titles-captions.html#axis-and-legend-titles

Labels

Don’t overlap



Use contrasting colors



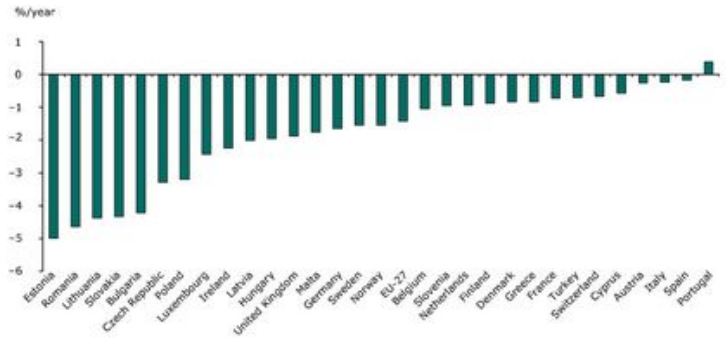
Keep long labels horizontal if possible

http://hypsypops.com/fonts-wisdom-text-can-make-break-graph/
Rotate bar chart -
https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts

Order

Use logical order

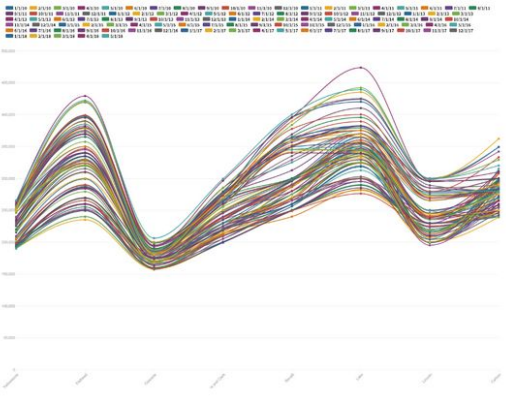
- Big to small
- Time left to right
- Oldest newest



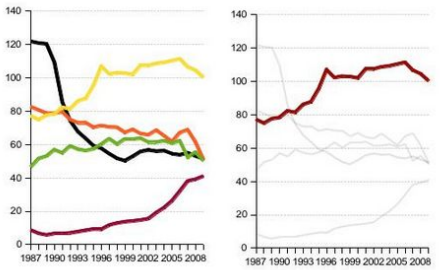
https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts

Presentation

Don’t overload



Draw the focus to what is important

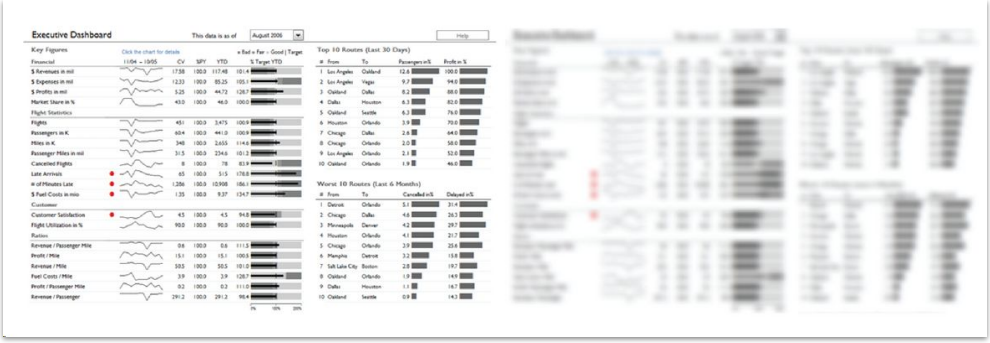


https://www.learningsolutionsmag.com/articles/misleading-data-visualizations-can-confuse-deceive-learners

https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts

Squint Test

Test your layout and organization with a squint test.



You squint your eyes and make an assessment on the overall layout, of elements that stand out, the visual balance and other characteristics of an effective user interface.

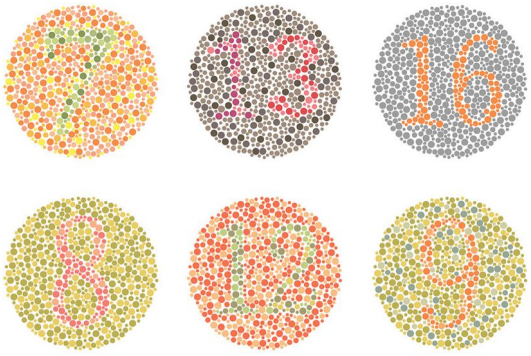
https://usabilla.com/blog/the-squint-test-how-quick-exposure-to-design-can-reveal-its-flaws/

https://blog.xlcubed.com/2008/08/the-dashbord-squint-test/

Accessibility

Color Blindness and Visual Impairment

Color Blindness



https://www.tableau.com/about/blog/2016/4/examining-data-viz-rules-dont-use-red-green-together-53463

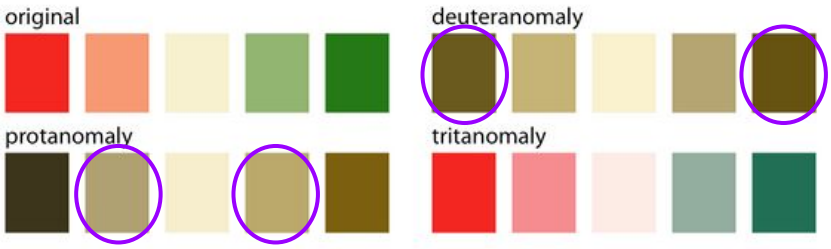


Figure 19.7: A red–green contrast becomes indistinguishable under red–green cvd (deuteranomaly or protanomaly).

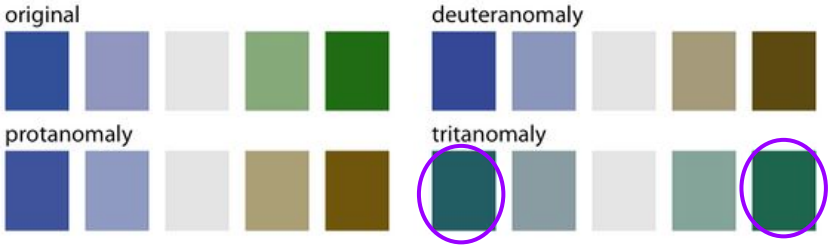
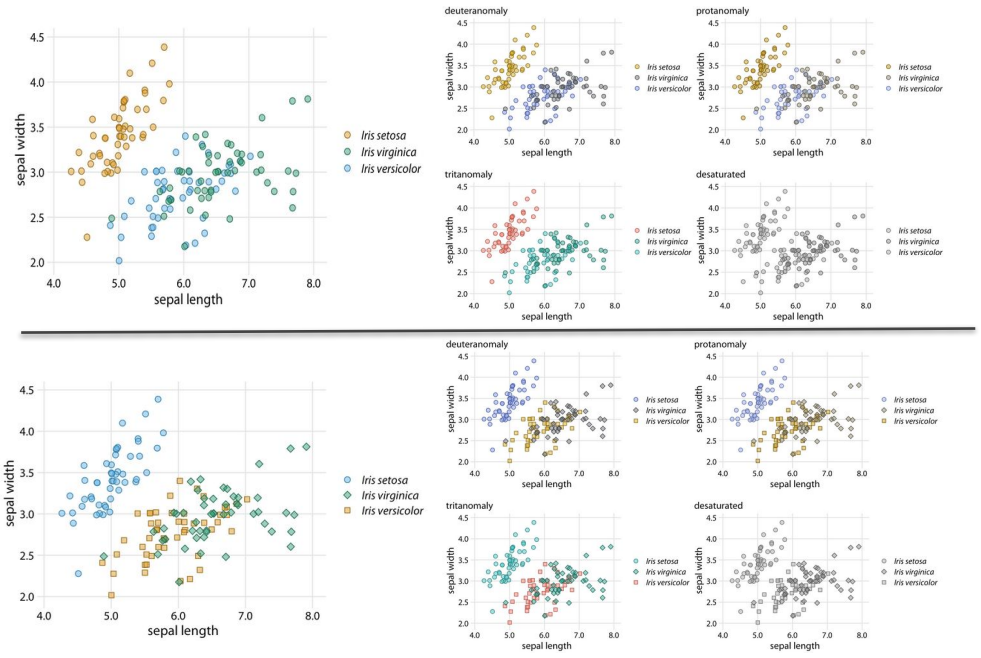


Figure 19.8: A blue–green contrast becomes indistinguishable under blue–yellow cvd (tritanomaly).

https://serialmentor.com/dataviz/color-pitfalls.html#not-designing-for-color-vision-deficiency

Use shapes AND colors to distinguish categories



https://serialmentor.com/dataviz/redundant-coding.html#designing-legends-with-redundant-coding

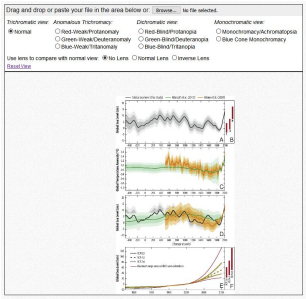
Colblindor

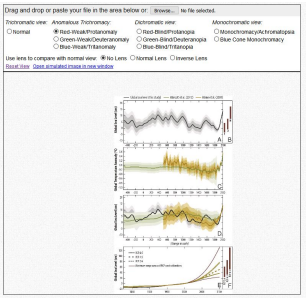
Color Blindness Simulator

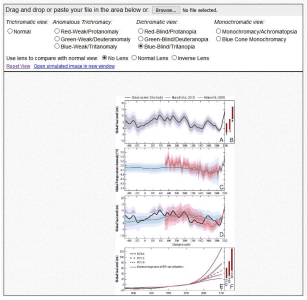
Original Colors

Red-Weak (Protanomaly)

Blue-Blind (Tritanopia)







Charts from: *Temperature-driven global sea-level variability in the Common Era.*

https://www.color-blindness.com/coblis-color-blindness-simulator/

https://www.pnas.org/content/113/11/E1434/tab-figures-data

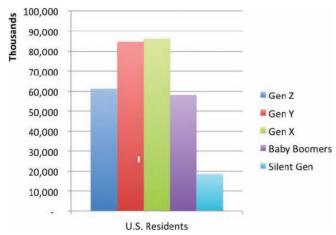
Visual Impairment

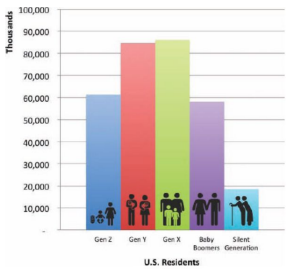
Use a simple, large font when possible.

Avoid similar colors

Don’t overcrowd with gridlines or textboxes.

Be attentive to backgrounds colors/constrasts.





How to write alt text -
http://www.perkinselearning.org/technology/blog/how-write-alt-text-and-image-descriptions-visually-impaired

Dealing with Data

Proportionality and Messy/Missing Data

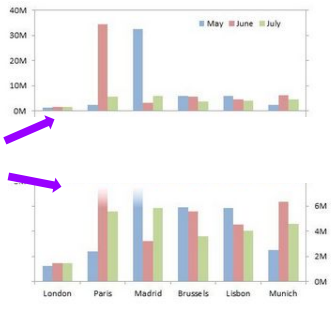
Proportionality

Visualization should represent data

https://www.shinobicontrols.com/blog/6-common-mistakes-with-data-visualization/

Representing Amount

If this happens

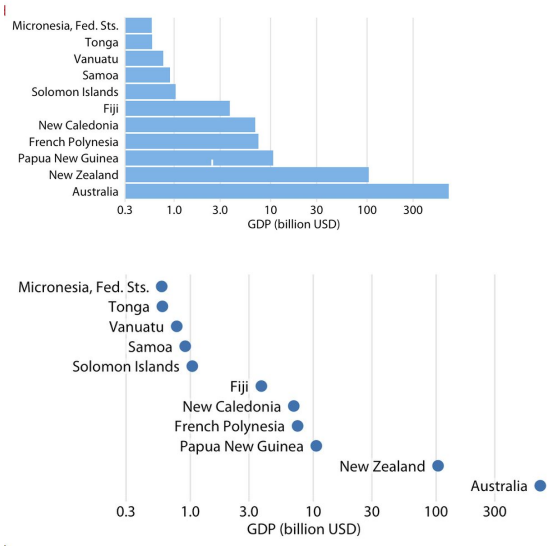


'Break' the axis

https://peltiertech.com/broken-y-axis-in-excel-chart/

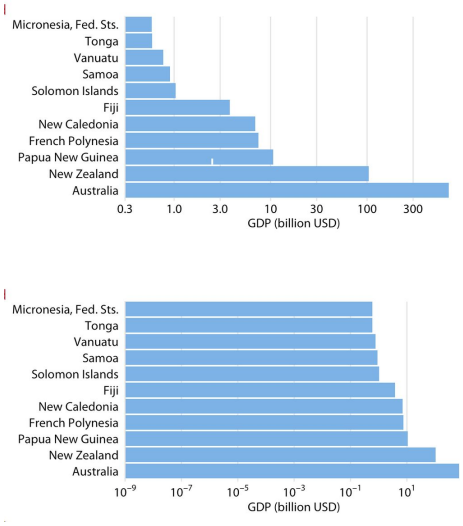
Representing Amount

If the chart clear, consider a different format.



https://serialmentor.com/dataviz/proportional-ink.html#visualizations-along-logarithmic-axes

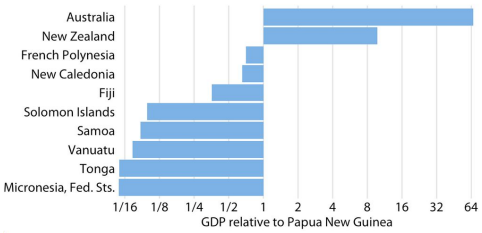
Logarithmic scale might not represent amounts well

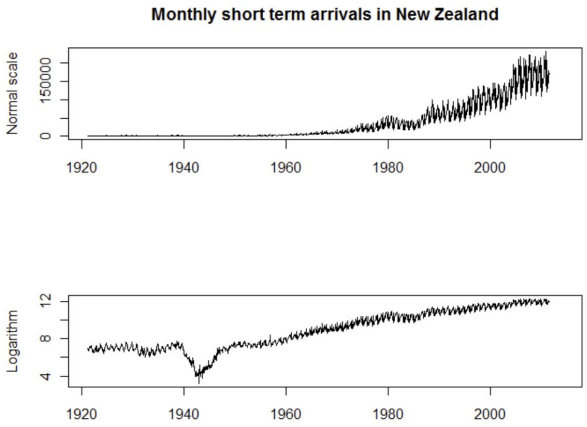


https://serialmentor.com/dataviz/proportional-ink.html#visualizations-along-logarithmic-axes

Representing Ratio

Use a logarithmic scale to visualize ratios





https://serialmentor.com/dataviz/proportional-ink.html#visualizations-along-logarithmic-axes

http://archive.stats.govt.nz/infoshare/

Messy/Missing Data

11. Dealing with Messy Data

- Analysis with missing data
  - Visualizing missing data
  - Types of missing data
    - So which one is it?
  - Unsophisticated methods for dealing with missing data
    - Complete case analysis
    - Pairwise deletion
    - Mean substitution
    - Hot deck imputation
    - Regression imputation
    - Stochastic regression imputation
    - Multiple imputation
      - So how does mice come up with the imputed values?
        - Methods of imputation
    - Multiple imputation in practice
- Analysis with unsanitized data
  - Checking for out-of-bounds data
  - Checking the data type of a column
  - Checking for unexpected categories
  - Checking for outliers, entry errors, or unlikely data points
  - Chaining assertions
- Other messiness
  - OpenRefine
  - Regular expressions
  - tidyr

Book available through URI 'Safari' database (instructions for access in resources).

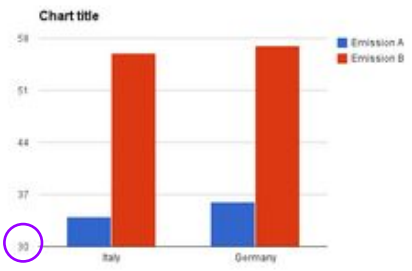
http://uri-primo.hosted.exlibrisgroup.com/01URI:Books_More:01URI_ALMA51175282870002396

Transparency

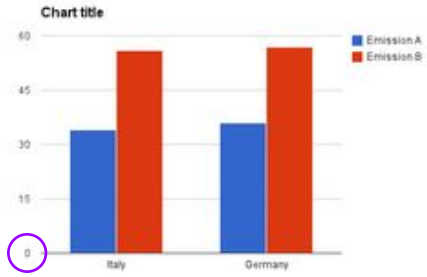
Clear Communication, Sources, and Credits

Clear Communication

Wrong

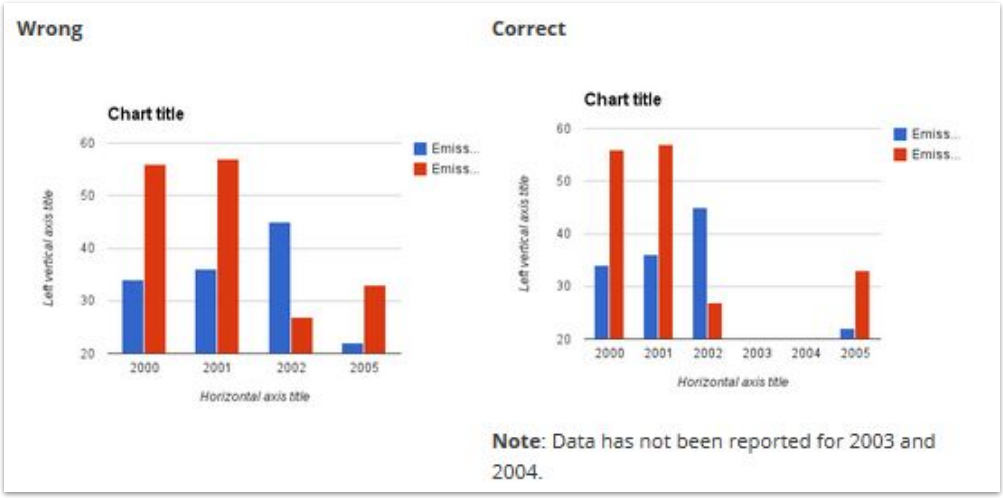


Correct



Cover the full set of data.

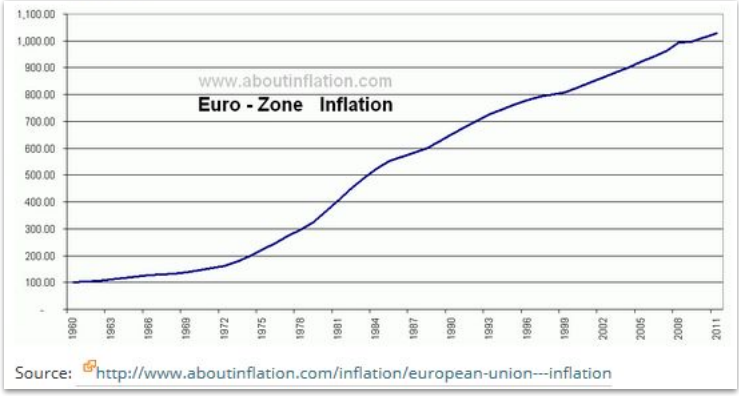
https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts



Represent and identify gaps in data.

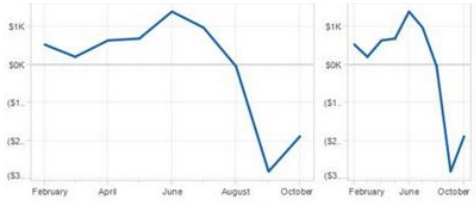
https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts

Adjust for inflation

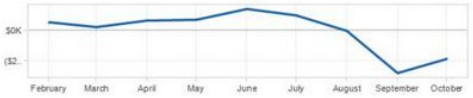


https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts

Use correct aspect ratios

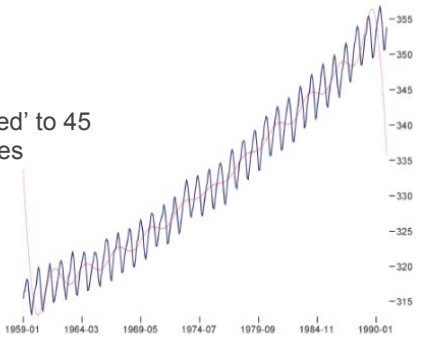


‘Banking’ to 45 degrees is accepted practice

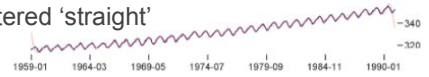


Unless ‘Banking’ obscures overall trends

‘Banked’ to 45 degrees



Chartered ‘straight’



Two plots of monthly atmospheric carbon dioxide measurements, taken from 1959 to 1990. The first plot, with an aspect ratio of 1.17, reveals an accelerating increase in CO_2 levels. The second plot, with an aspect ratio of 7.87, facilitates closer inspection of seasonal fluctuations, revealing a gradual attack followed by a steeper decay. Source: Computer Science Division, University of California, Berkeley (http://vis.berkeley.edu/papers/banking/)

https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts

Sources and Credits

BEST PRACTICES

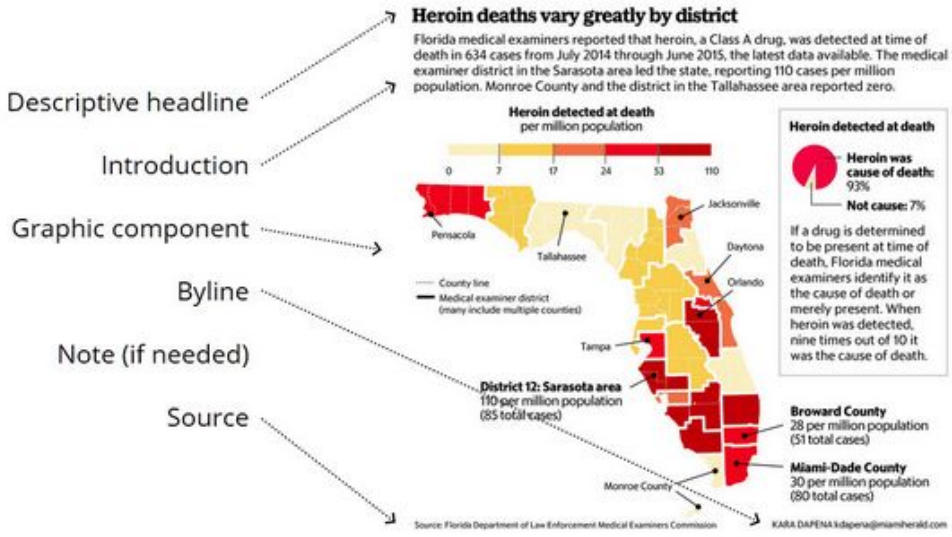
An infographic design that is upfront about its data sources is instantly perceived to be more credible. There are a few best practices that are common to good infographics.

- Track down and cite the original source of the data, not the news article that quoted the data.
- List the data source references either inline with the data or in the footer at the bottom of the infographic.
- List the URL to the specific report or data set, not just the host site.
- Include the relevant date or year the data was published to establish relevancy.

In addition, one of the best practices is for an infographic author to make the source data available to readers online for download as a spreadsheet. This is often accomplished with Google Docs spreadsheets[4] that are accessible to the public. The infographic would include the URL directly to the data spreadsheet, which lists the data and shows any relevant sources or calculations made as part of the design. Readers can access the data directly and even use the data on their own.

https://learning.oreilly.com/library/view/cool-infographics-effective/9781118582305/14_chapter-06.html

Parts that make the whole



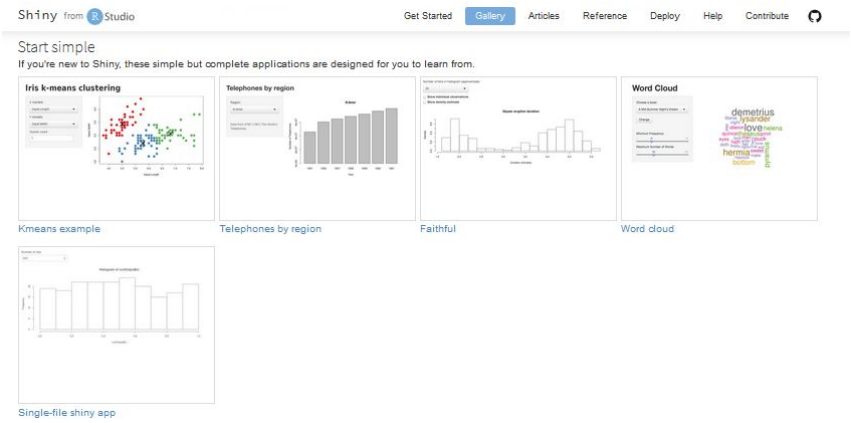
http://www.snpa.org/stories/using-data-visualization-to-tell-a-story,2704509

Learning Resources

Shiny.rstudio, Safari (O'Reilly) database, and web

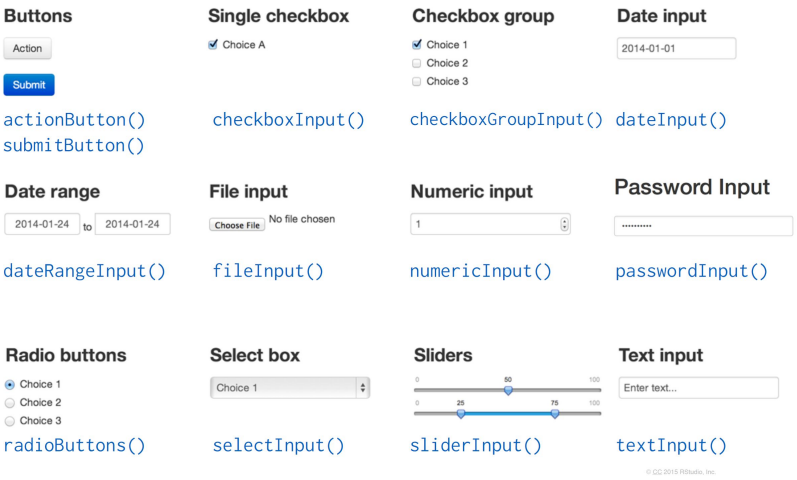
Shiny.rstudio

Getting Started



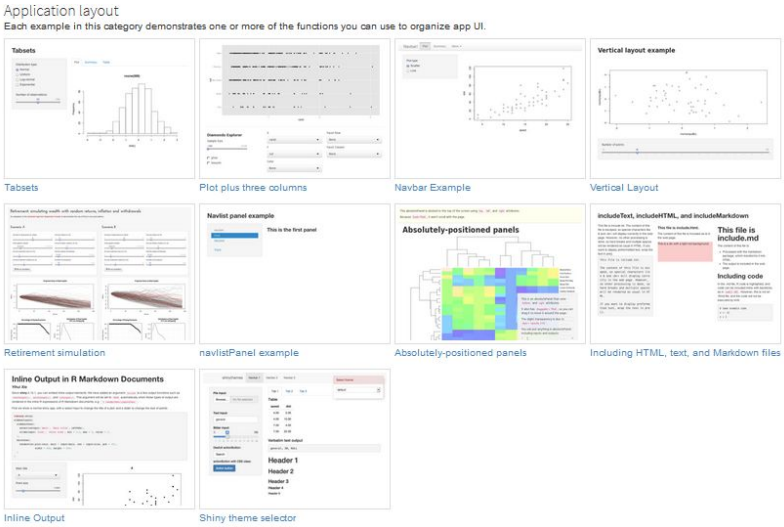
https://shiny.rstudio.com/gallery/

Widgets



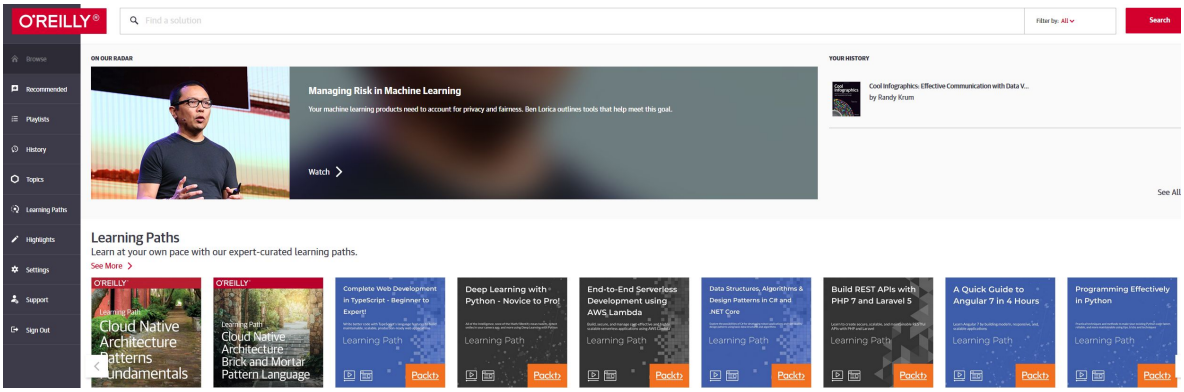
https://shiny.rstudio.com/gallery/

Layout



https://shiny.rstudio.com/gallery/

Safari (O'Reilly) database



To log into Safari (O'Reilly) e-books, videos and learning paths, go to:

https://uri.libguides.com/az.php

Find 'Safari' in list of databases. Use your URI.edu email to login/setup account.

Getting Started

Shiny R

by Jared P. Lander

Continue

4+ Hours of Video Instruction

Create visual, interactive, data-based web applications without having to be a web developer

Estimated time to complete:
4h 52m

Topics:
Shiny

Published by:
Addison-Wesley Professional 2017

Lesson 1: Getting Started with Markdown

Markdown is an excellent tool for writing documents. It enables you to focus on content and not be distracted by design. Highly flexible, it can render elegant HTML files, professional PDFs, editable Word documents, web slideshows, and more. You learn the basic structure of Markdown documents and how easy it is to write page after page.

Lesson 2: Integrating Code into Markdown

Weaving together R code with prose has been a hallmark of scientific writing in R for many years thanks to S-Weave and knitr. Now with RMarkdown it is easier than ever before. We go over the necessary steps to integrate R code into Markdown documents for easy blending of results and explanations.

Lesson 3: Shiny in RMarkdown

Getting started with Shiny can seem daunting at first, having to deal with server-side and UI-side code. This endeavour does not have to be intimidating, however, because Shiny can be integrated into RMarkdown with ease. We learn the basic concepts of Shiny, such as user inputs and rendering outputs, all within a single-file RMarkdown document.

Lesson 4: Reactive Expressions

The underpinning of Shiny is reactivity. This is how objects are updated based on changes in other objects. We cover the basics of what is needed to know in order to avoid trouble while building Shiny apps.

Lesson 5: flexdashboard

The flexdashboard package enables the creation of stunning web layouts, powered by R, written entirely in RMarkdown. We create a document that looks and feels like a full website and add Shiny components for interactivity.

Lesson 6: shinydashboard

Traditional Shiny apps depend upon UI and server files which separate form from function. With shinydashboard we can quickly build professional web apps that are highly customizable. We go through layout considerations and how to incorporate input and output objects. We then integrate HTML widgets, use shinyjs to incorporate JavaScript functionality and enable the creation of user-generated reports.

https://learning.oreilly.com/videos/shiny-r/9780134863320

Widgets

Hands-On Dashboard Development with Shiny

0 REVIEWS

by Chris Beeley

Publisher: Packt Publishing

Release Date: August 2018

ISBN: 9781789611557

Topic: Shiny

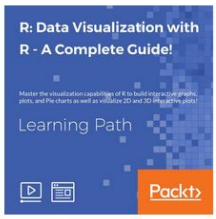
- Dashboards
  - Building your first dashboard
  - Laying out a dashboard
  - Adding icons to your dashboards
  - Adding notifications, messages, and tasks
  - Info boxes and value boxes
  - Adding Google Charts to your dashboard
  - Summary

https://learning.oreilly.com/library/view/web-application-development/9781782174349/ch02s03.html

LEARNING PATH

R: Data Visualization with R - A Complete Guide!

Instructor Carol Quadros



Continue

TIME TO COMPLETE:
7h 28m

TOPICS:
R

PUBLISHED BY:
Packt Publishing

CREATED:
October 2018

What will you learn

- *Deepen your knowledge by adding bar-charts, scatterplots, and time series plots using ggplot2.*
- *Create a fully-featured website using Shiny with real-time features such as adding and controlling functionalities.*
- *Create simple and quick visualizations using the basic graphics tools in R.*
- *Introduce users to basic R functions and data manipulation techniques while creating meaningful visualizations.*
- *Add elements, text, animation, and colors to your plot to make sense of data. Perform predictive modeling and create animated applications.*

Resources: Code downloads and errata:

- Learning R for Data Visualization
- R Data Visualization - Basic Plots, Maps, and Pie Charts
- R Data Visualization - Word Clouds and 3D Plots

PATH PRODUCTS

This path navigates across the following products (in sequential order):

- Learning R for Data Visualization (1h 59m)
- R Data Visualization - Basic Plots, Maps, and Pie Charts (3h)
- R Data Visualization - Word Clouds and 3D Plots (2h 29m)

https://learning.oreilly.com/learning-paths/learning-path-r/9781789950717/

Learning Shiny

0 REVIEWS

by Hernán G. Resnizky
Publisher: Packt Publishing
Release Date: October 2015
ISBN: 9781785280900
Topic: Shiny

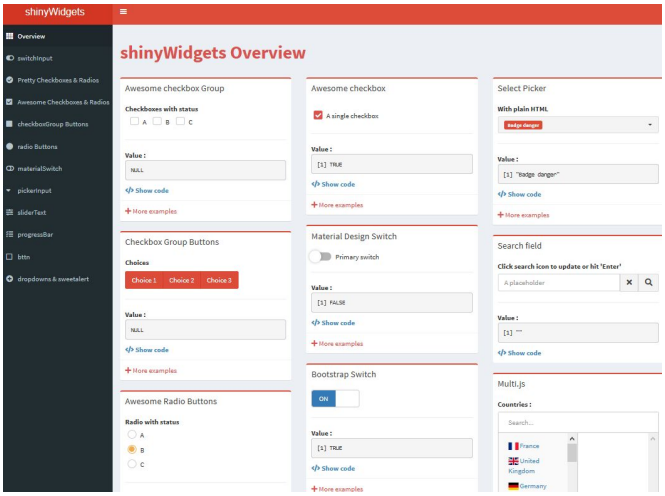
What You Will Learn

- Comprehend many useful functions, such as lapply and apply, to process data in R
- Write and structure different files to create a basic dashboard
- Develop graphics in R using popular graphical libraries such as ggplot2 and GoogleVis
- Mount a dashboard on a Linux Server
- Integrate Shiny with non-R-native visualization, such as D3.js
- Design and build a web application

https://learning.oreilly.com/library/view/learning-shiny/9781785280900/

Web

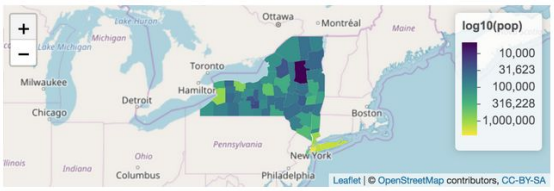
Packages for widgets



https://github.com/dreamRs/shinyWidgets

Charting packages

Leaflet



GeoJSON map (RStudio)

Like highcharter, Leaflet for R is another charting packaged based on a hugely-popular JavaScript library of the same name.

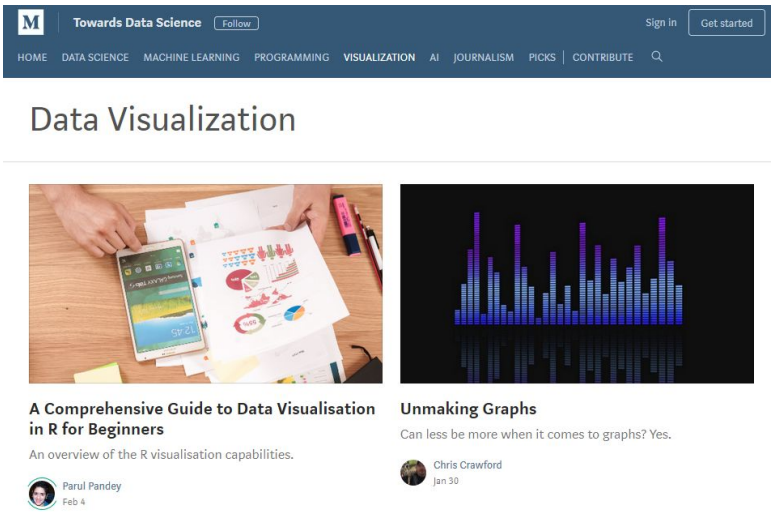
Leaflet offers a lightweight but powerful way to build interactive maps, which you've probably seen in action (in their JS form) on sites ranging from The New York Times and The Washington Post to GitHub and GIS specialists like Mapbox and CartoDB.

The R interface for Leaflet was developed using the htmlwidgets framework, which makes it easy to control and integrate Leaflet maps right in R Markdown documents (v2), RStudio, or Shiny apps.

Created by: Joe Cheng, Bhaskar Karambelkar, Yihui Xie
Where to learn more: Leaflet for R

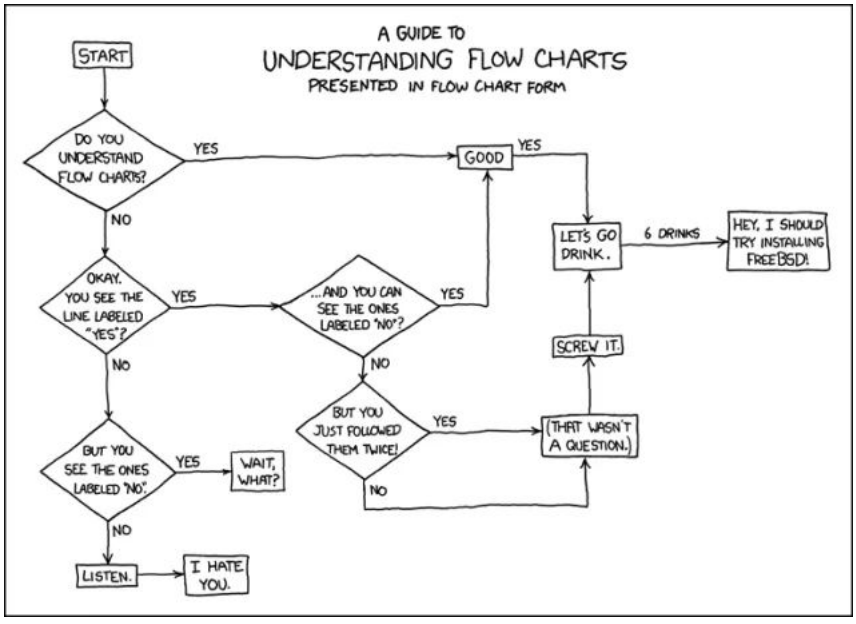
https://mode.com/blog/r-data-visualization-packages

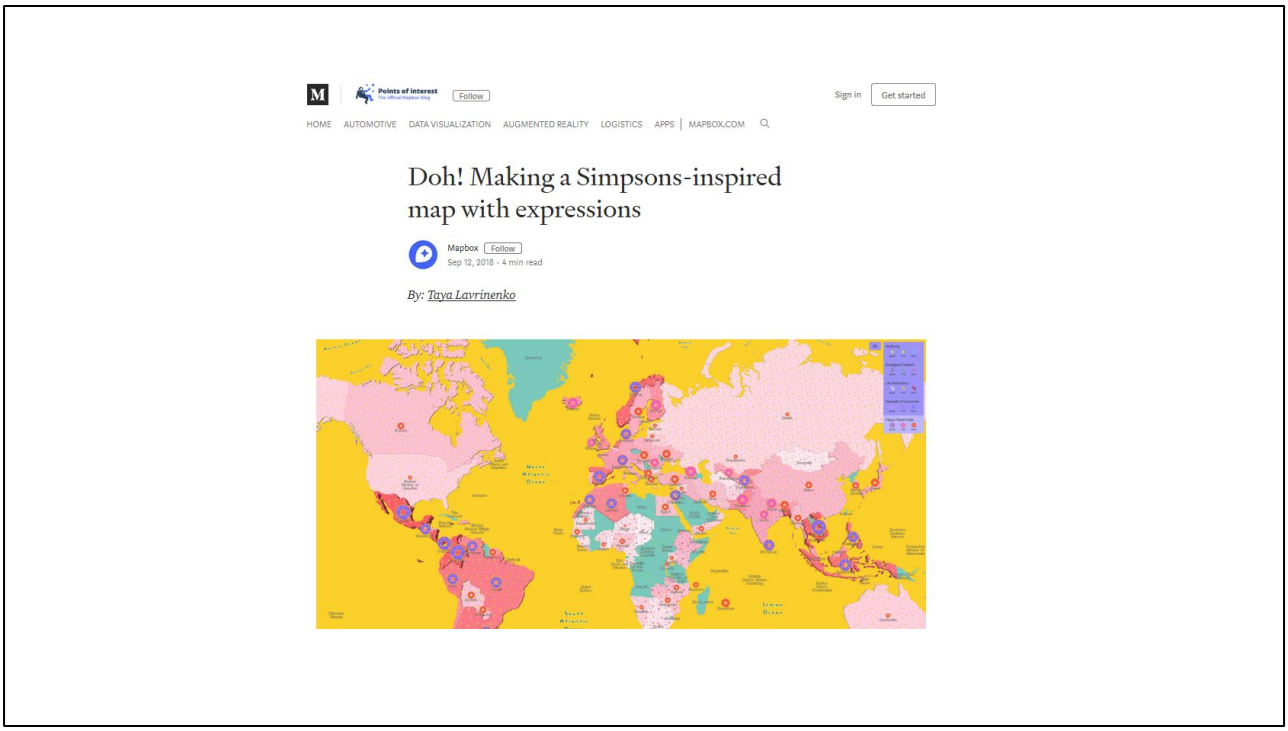
General resources



https://towardsdatascience.com/data-visualization/home

Fun


A GUIDE TO
UNDERSTANDING FLOW CHARTS
PRESENTED IN FLOW CHART FORM
START
DO YOU UNDERSTAND FLOW CHARTS?
YES
GOOD
YES
NO
LET'S GO DRINK.
6 DRINKS
HEY, I SHOULD TRY INSTALLING FREEBSD!
OKAY. YOU SEE THE LINE LABELED "YES"?
YES
...AND YOU CAN SEE THE ONES LABELED "NO"?
YES
NO
SCREW IT.
NO
BUT YOU JUST FOLLOWED THEM TWICE!
YES
(THAT WASN'T A QUESTION.)
BUT YOU SEE THE ONES LABELED "NO".
YES
WAIT, WHAT?
NO
LISTEN.
I HATE YOU.



https://blog.mapbox.com/doh-making-a-simspons-inspired-map-with-expressions-86e633b61ede

Optimized Brewery Road Trip, With Genetic Algorithm

Visit the best American breweries of 2018, based on RateBeer rankings, while minimizing travel time and distance.

Rājekumāravenkataperumālrāzumbahadūrvāripeta and other place I'd like to visit

https://medium.com/@tjukanov/places-and-their-names-observations-from-11-million-place-names-8ea34cf61da4