Lesson 7
Storytelling Techniques

Overview

In this lesson, we will look into the various techniques and methods in telling the data journalism story. Focus will mainly be on visualization, as both a process for getting insights from data for your story and a process for delivering your data in a story. However, other storytelling techniques such as web-based visualizations and interactives will also be briefly discussed. Likewise, tips on putting the story together with narration and humanization will also be tackled here.

Duration: (3 weeks) - 6 hours

Objectives

At the end of the lesson, you should be able to:

1. Identify the two purposes of visualization in data journalism
2. Enumerate different interactive types of data visualization
3. Discuss key principles of design for data visualization

Lesson proper

Discussion flow:

1. Visualization in publishing a story
2. Design principles in data visualization
3. Interactive visualization techniques
4. Don't forget the narrative in writing the data journalism story

INTRODUCTION

Visualization is not just about the aesthetics of presenting data; it is also an essential tool in summarizing and simplifying insights derived from the data. As discussed in Lesson 1 one of the components of data journalism is the "'visualization to present data" besides the "treatment of data as a source" and the "application of statistics to interrogate data" (Howard, 2014). Hence, one of the focus of our course is on visualization in data journalism. We will try to practice visualization.

Meanwhile, as introduced in Lesson 5, visualization is one of the techniques in telling stories with data. After you have your data journalism story's hypothesis, you would have to analyze the data you have to either support or refute your hypothesis.

Data visualization is not exclusive to journalism nor is it a new technique. It has been around for quite some time but the way it is presented now has drastically changed due to the available technologies - apps and softwares. William Playfair is widely considered as the first to have used or even invented the graphics forms of presentation we know today such as line graph and bar chart as well as the pie chart and circle graph (Friendly, 2006, p.8). The development of tools (such as apps) for creating visualizations through the years have made visualizing more complex data with more creative design elements possible. This development has given journalism - data journalism - more sophisticated, impactful, and engaging stories.

One useful definition of data visualization where we can start our discussion with is Andy Kirk's, author of *Data Visualisation: A Handbook for Data Driven Design* (2016), succinct definition:

> *Data visualization is "the representation and presentation of data to facilitate understanding"* (Kirk, 2016. p.19)

Zeroing in to journalism, how do we "represent and present data to facilitate understanding"?

Visualizing the data can provide two main purposes for journalism (Gray, et al 2012, p. 186-187; Gray, et al, 2012, p. 165; and Kuek Ser Kuang Keng, 2020):

1. **Reporting phase** (verifying/checking hypothesis) or discovering insights to verify hypotheses. Usually, data collected for journalism projects are in table format. And most often than not, tables are not sufficient to give us an overview of the dataset nor can give us patterns or trends we can use for our stories. Hence, at the start of writing the data journalism project, it is useful to already start on the work of visualization to get insights and see if these are related to your hypothesis. Visualizing data at the early stages of your research can:

a. help you identify themes and questions for the rest of your reporting
b. surface trends and contrasts; hidden connections
c. identify outliers: good stories, or perhaps errors, in your data
d. help you find typical examples
e. show you holes in your reporting

For more examples of these insights, watch the short discussion of the portion on 'Finding Stories from Data' (16:50-23:40 timestamp) on Data Journalism & Visualization for Rookies: Discover Stories from Data. (2020, June 23). Note: you can watch the whole video for more insights on how visualization can provide insights to possible stories though our general approach in our class for our data journalism stories is hypothesis-based (not a "fishing expedition" on data sets).

Remember the four questions you can ask in hypothesis building, these are questions you can answer via initial visualization so that your story either proves or disproves the hypothesis.

- Problem: show how big or expensive or intense is the problem
- Impact: show who are affected, how much, disaggregation of impact
- Cause: show what are the possible causes and what factors impact/ affect the causes
- Solution: show what could be solution or how effectiveness can be measured

2. Publishing the story

a. illustrate a point made in a story in a more compelling way
b. remove unnecessarily technical information from prose
c. provide transparency about your reporting process to your readers

Visualization in publishing a story

Viz for publishing a story will be the main content of this Lesson on visualization. It is worth noting that given the purpose of visualization to a data journalism story, it is not something that should be done as an afterthought to when a story is already largely done. Hence these two purposes are part of the whole process of working with data. Visualization can and will

provide more insight into your data and thus starting the work on visualization is part of the earlier process of making the data journalism story or project.

There are many instances when a story is best told using visuals, or visuals would be used to illustrate a point in the story or provide representation of a technical information so it would no longer be discussed in prose (Gray, Bounegru, Chamber, 2012). A survey among data journalists showed that data visualization for them was used to "emphasize a point, to add empirical evidence, to enable users to explore data sets, as aesthetic attraction to stimulate interest and to offer entry into unseen stories" (Gray and Bounegru, 2021, p.101).

To start working on the visualization for your story, we'll follow the basic principles of journalism as well as of design (Chiasson, Gregory et al, 2018, She Loves Data, 2021, and Mendiola, 2020)

- We have to identify our message or the hypothesis of our story. How do you organize your data story into this story hypothesis? Is the story about learning, about a challenge, a how-to, a big idea or an impact.

- Important also is to understand who readers or audiences are - their level of data literacy, what data points are important for them, and the context to which we will tell the story of the data.

- Once the message is set and who know who we are making the story for, we need to decide which data to use and how much data to illustrate

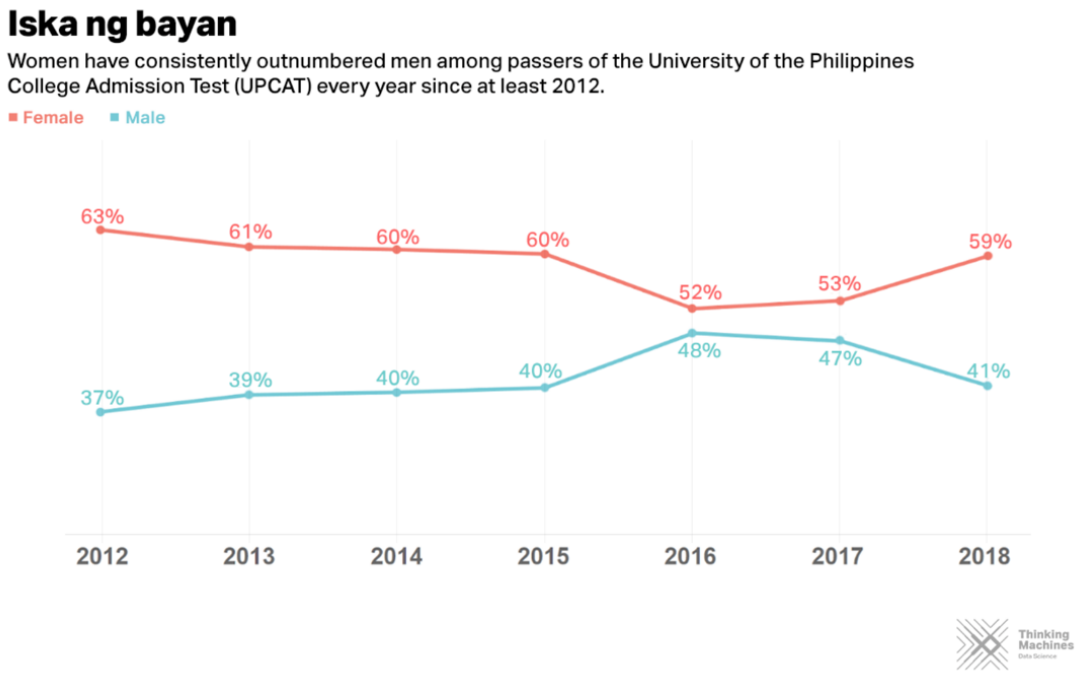
- And then we design, get feedback, edit and revise. In designing, it's not all flair and colors. It's important to know which type of visualization to use for various purposes. Let's check Andy Kirk's (2016, p. 161-209).

Picking the right graph

Not all charts are created equal. A data insight should be represented by the right visualization for it to be effective. Here's a rundown of suggested visualizations to represent common data relationships:

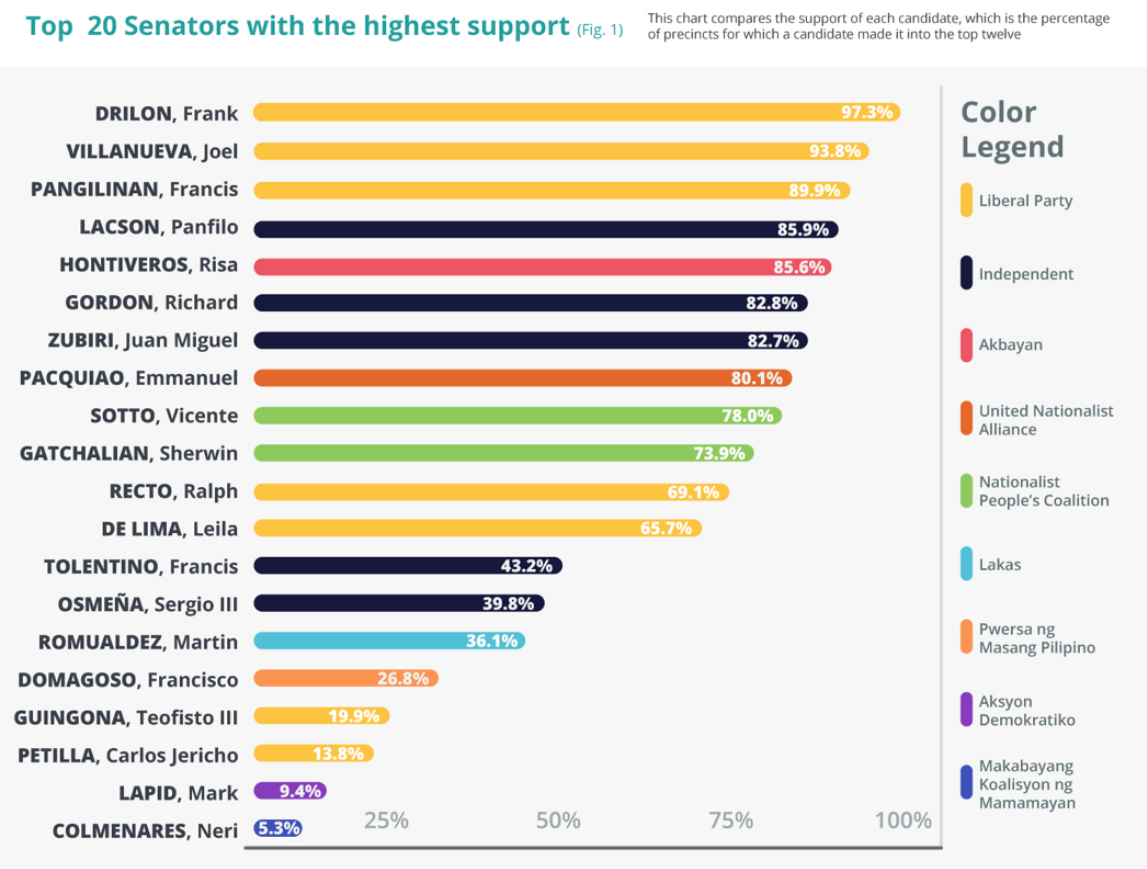
- Trend over time. Change is more observable when time is in the x-axis. Progression over time is best represented by a line graph.

  See visualization by Thinking Machines, from Mind the gap: A gender parity analysis of UPCAT passers (https://stories.thinkingmachin.es/upcat-passers-gender-gap/)



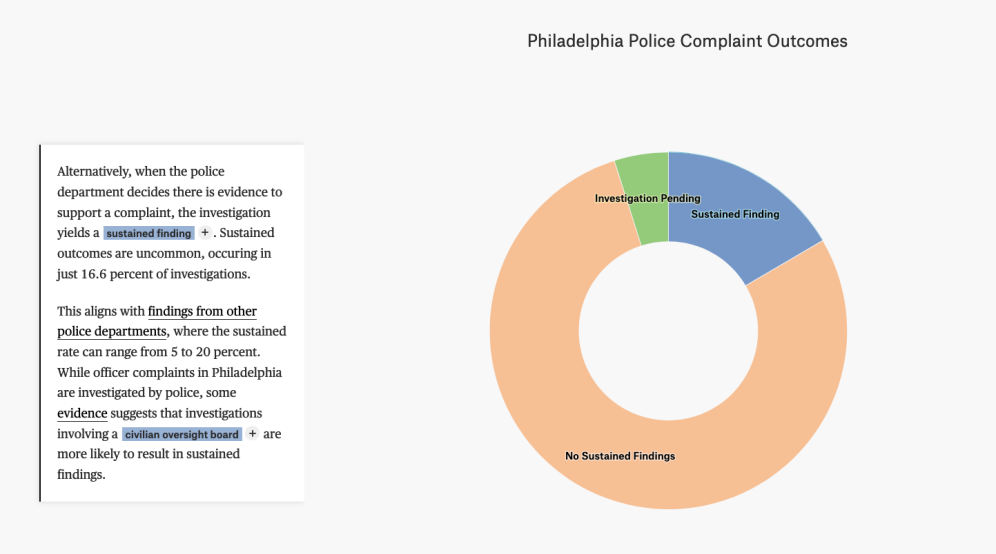
- Ranking. To compare size or magnitude of similar items at different times, bar graphs are used to easily visualize differences in quantities side by side.

  See visualization by Reina Reyes, How We Choose Our Senators (https://www.researchgate.net/publication/329835844_Shopping_for_Politicians_Insights_from_Market_Basket_Analysis_of_Senatoriables)



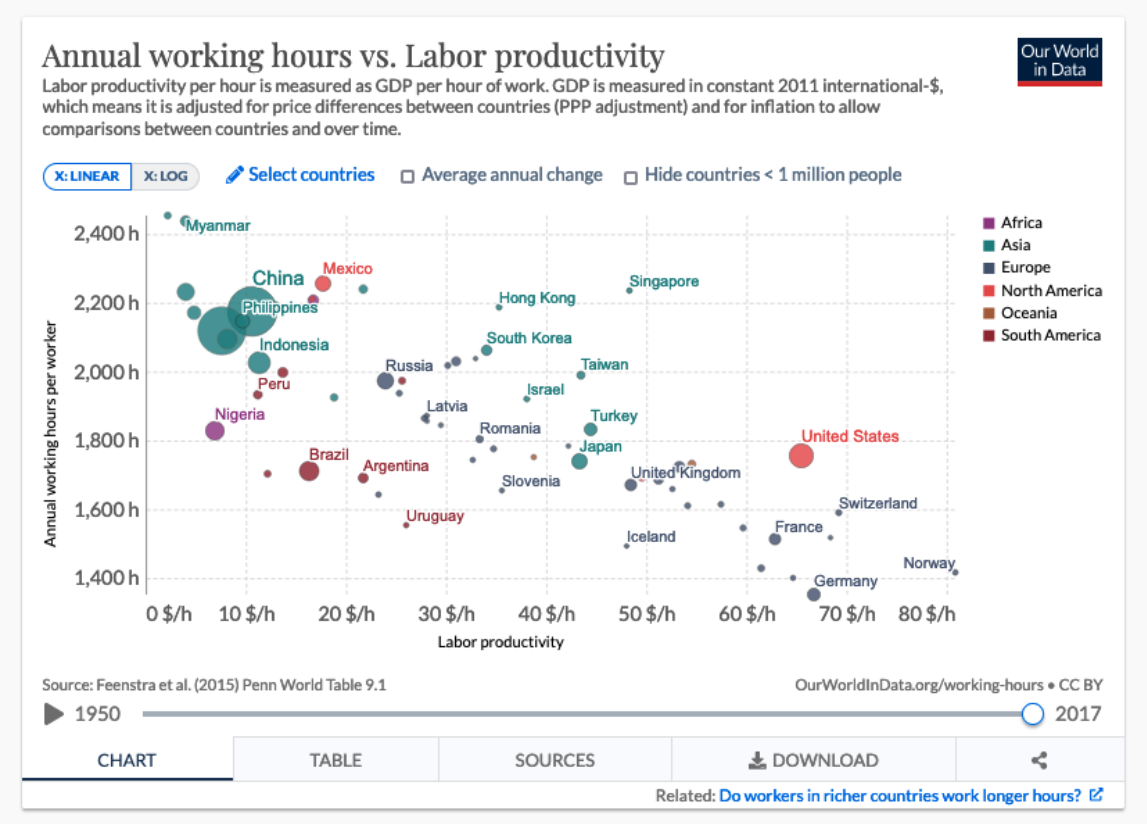
- Composition. Donut graphs, cross section diagrams, and tree maps are made to visualize parts or components of a whole.

  See visualization by The Pudding, How officer complaints are investigated: a case study of Philadelphia police. (https://pudding.cool/2020/10/police-misconduct/)



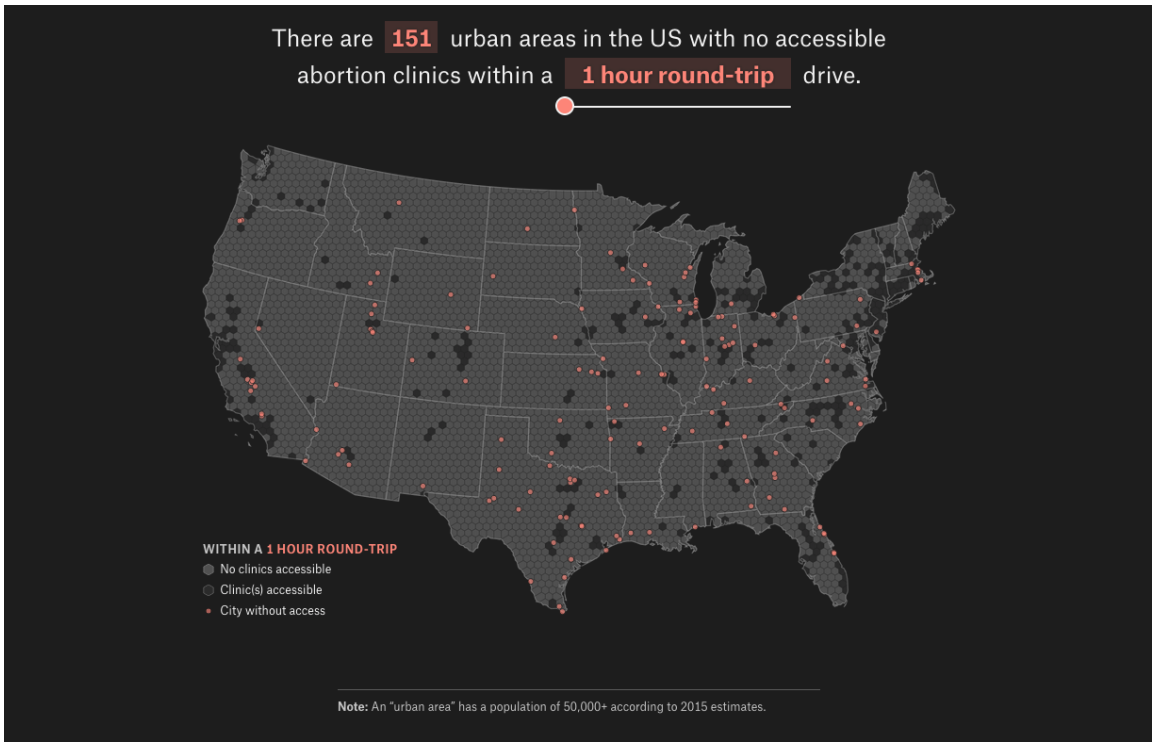
- Correlation. To illustrate how one phenomenon is related to another, scatter plots and bubble charts provide the flexibility to represent multiple data points based on two variables.

  See viz by Our World in Data, Time Use (https://ourworldindata.org/time-use)



- Geospatial. Maps are best used when comparing quantities or intensities across locations. In a geospatial visualization, location acts as a container and additional data can be added like overlaying flows, densities, or ratios.

  See visualization by The Pudding, How far is too far? An analysis of driving times to abortion clinics in the US. (https://pudding.cool/2017/09/clinics/)



Design Principles in Data Visualization

Playing with numbers and using tools to get the right visualization is part of the challenge but not the complete picture. Key design principles should be considered for a visualization in a published story to be compelling and useful.

According to Cole Nussbaumer Knaflic, author of *Storytelling with Data* (2015, p.241), "data visualization - and communicating with data in general - sits at the intersection of science and art."

Some key points to think through:

1. Three axioms of visualization and pillars of design (Steven Braun, Northeast University, 2016):

- Effective Design. leverages principles of psychology and cognitive science to inform presentation
- Ethical design. Views visualization as constructed space and presents data without bias or selectivity
- Universal design. Ensures that conclusions from data will be uniformly understood and accounts for differences in vision and learning

There are already so many data visualization guides, books, and videos out there. Some are focused on business presentations, while some focus on sharing research results, with a handful focusing on data journalism. We've filtered here some of the more relevant and easy to understand foundation lessons for visualizing data, particularly in the context of journalism.

==========

Learning activities

WATCH For this part, let's revisit the discussion from the Data Journalism Workshop in 2020. The main topic is from 8:04-1:8:46 (it's a one hour talk). https://www.youtube.com/watch?v=DGD664zXRC0&list=PLUVjjEFpwcL9lAF-2dVjG2licEHt_iBsX&index=6)

READ and explore the samples in the following readings
Chapter 14: Anatomy of a Graphic by Amanda Makulec.
https://trinachi.github.io/data-design-builds/ch14.html. Chiasson, T., Gregory., et al. Data + Design: A simple introduction to preparing and visualizing information

For more inspiration and best practices, visit these sites:
David MacCandless' https://informationisbeautiful.net/
Edward Tufte's
https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=0002w4&topic_id=1&topic=
Data Viz Done Right. http://www.datavizdoneright.com/
Simple way of changing data viz for clarity. http://www.perceptualedge.com/example19.php

TASK Look 1-2 samples of recent and local data viz in journalism gone wrong (misleading, confusing, etc) and briefly explain based on the ethics and design principles of visualizing data.

==========

Interactive Visualization Techniques

Recent developments in the past few years have seen 3D graphics and even virtual reality (VR) being considered and explored in journalism, in general. However, these considerations should still be hinged on how to best tell the story and what the readers really need.

Web based visualization / Interactive visualization

Visualization for data journalism that makes use of the web is often referred to as interactive visualization. Elliot Bentley (in Gray and Bounegru, 2021) talked about how the web can provide a rich platform for data visualization in journalism.

Allows for huge datasets which readers can explore

The web can allow opening data sets or providing news apps or big interactive maps for readers to check the data themselves or explore the data relevant to them only. Another possibility is to provide visualization to "show readers where they fall in a data set… by asking a couple of personal details." One example is the BBC's UK Fat Scale Calculator https://www.bbc.com/news/health-43697948

News application (or News App) is a "big interactive database" which is part of a story or is the story itself (p. 185). either open and searchable databases or interactive visualizations which allows readers / web users to interact or engage with the data (Gray, Bounegru, Chamber, 2012).

Some examples of these are:
ABS-CBN's 2018 'The Pinoy Tax Reform Calculator"
https://news.abs-cbn.com/pinoytaxreformcalculator

ProPublica's 2019 Dollars for Docs: How industry dollars reached your doctors.
https://projects.propublica.org/docdollars
The goal of news apps is to create products of "lasting value" beside the daily news cycle or even the noise of an issue has died down. The main considerations if news apps would be the best option for a story is to again consider who are the audience and their needs as well as logistical concerns in building something as complex and durable as news apps. The journalist could need more than just him/herself crunching numbers or visualizing on static charts. Building news apps would need developers and designers on the team. It can be 'time consuming' and 'expensive' (p. 184). For news organizations, the challenge is if they

are building something a news app, it has a strong 'return on investment' for the organization.

Guiding the reader through complex charts

The web also allows users to explore the data set - zoom in and out or travel through time or check other data; samples include "scrollytelling"

"Scrollytelling"

This concept is fairly simple: readers will see new content or content changes through various transitions when they scroll up or down the web page (Amabili, 2019). The main advantage of this feature is that scrolling while on web pages whether on our laptops or our mobile phones has been sort of a natural action for us nowadays. Second, scrolling and getting new content as they scroll "engages readers because of the sense of control, exploration, and discoverability." Third, according to Amabili, there are existing libraries or sample codes which journalists can use to implement these features.

Examples:
Swissinfo.ch's 2019 story on 'What people in Switzerland worry about",
https://www.swissinfo.ch/eng/2019-elections_what-people-in-switzerland-worry-about/44997722

The Pudding's 2021 virtual essay / article: Follow the Science: A look at the global research effort to combat the coronavirus pandemic by Jeff Macinnes,
https://pudding.cool/2021/03/covid-science/

Allows for data driven games and simulations

Using data sets and programming, some news agencies tried to provide a 'video game' feel to presenting data. It lets readers explore the data more. There are only a few success stories in this type of visualizing data journalism as these require a high level of engagement from web readers - most of which are only up for scrolling.

Some of the best examples are the following:

- Five Thirty Eight's 'How to Win a Trade War'
  https://fivethirtyeight.com/features/how-to-win-a-trade-war/

- One example is Upshot in NYTimes 'You Draw it" https://www.nytimes.com/interactive/2015/05/28/upshot/you-draw-it-how-family-income-affects-childrens-college-chances.html

The Seeingdata.org research surfaced some understanding on what makes visualization effective. These factors, however, "carry different weight in relation to different visualizations, contexts, and purposes" (Gray and Bounegru, 2021).

1. provoke questions/desire to engage in discussions with others
2. create empathy for other humans in the data
3. generate enough curiosity to draw the user in
4. reinforce or back up existing knowledge
5. provoke surprise
6. persuade or change minds
7. present something new
8. lead to new confidence in making sense of dataviz
9. present data useful for one's own purposes
10. enable an informed or critical engagement with a topic
11. be a pleasurable experience
12. provoke a strong emotional response

Don't forget the narrative in the Data Journalism Story

As much as visualization - whether the static ones or the interactive ones - are exciting to work with in data journalism. Some aspects of coming up with a data journalism story should not be forgotten.

Paul Bradshaw (2011) worked on a framework of how to communicate a data journalism story. He outlined six ways but for our purposes, we have highlighted the visualization and some examples of personalization and utility in the previous discussions for Lesson 7. Here we highlight the importance of the other significant part of the data journalism story - humans and the context for such data.

Bradshaw noted that "making stories relevant to people" is a key point in storytelling in data journalism. Ways to go about bringing your data to the level of relatable is to interview sources whose lives were affected by the data. Ilia Blinderman of The Pudding suggests

starting the story with a character or a single data point (2020). Providing the readers "a human experience" they can relate the statistics to draws and moves them through a story. It also makes it easier for readers to understand the bigger picture and the other data presented in the story.

Having a smaller scale to look at is usually key when the story's hypothesis is complex and needs contextualization.

Local stories:
https://news.abs-cbn.com/spotlight/02/22/21/dissecting-data-covid-19-is-killing-filipinos-at-its-fastest-rate-in-7-months

https://news.abs-cbn.com/business/03/15/21/dissecting-data-the-economy-after-a-year-into-the-covid-19-pandemic

Ethics of Data Viz

- How to spot a misleading graph - Lea Gaslowitz. July 6, 2017. TED Talks. https://www.youtube.com/watch?v=E91bGT9BjYk
- Chapter 17 (Perception Deception) and 18 (Common Visualization Mistakes) Chiasson, T., Gregory., et al. Data + Design: A simple introduction to preparing and visualizing information

SUMMARY OF LESSON 7

The process of visualizing data in a story is not exactly linear. Many times you will have to revisit the considerations of who the audience are, what's the main message, what is the right data and how much do we need, etc. Consequently, data visualization is both an art and a science where design principles should be reviewed and applied.

REFERENCES

Amabili, Lorenzo, Aug 2019. From Storytelling to Scrollytelling: A Short Introduction and Beyond. Medium. Data Accessed August 2021.
https://medium.com/nightingale/from-storytelling-to-scrollytelling-a-short-introduction-and-beyond-fbda32066964

Blinderman, Ilia. (Sept 3, 2020) How to make dope shit Part 3: Storytelling. The Pudding. https://pudding.cool/process/how-to-make-dope-shit-part-3/

Bradshaw, Paul. 2020. 6 ways of communicating data journalism. Online Journalism Blog. Date Accessed August 2021 https://onlinejournalismblog.com/2011/07/13/the-inverted-pyramid-of-data-journalism-part-2-6-ways-of-communicating-data-journalism/

Elements of Information Design. 2016. https://repository.library.northeastern.edu/downloads/neu:m039ws550?datastream_id=content

Friendly, M. 2006. A brief history of data visualization in Handbook of Computational Statistics: Data Visualization, Vol. 3. Springer-Verlag. (p 1-43)

Gray, J., & Bounegru, L. (Eds.). (2021). Data Visualisations: Newsroom Trends and Everyday Engagements. In The Data Journalism Handbook 2: Towards a Critical Data Practice (pp. 101–107). European Journalism Centre and Google News Initiative. https://doi.org/10.5117/9789462989511

Kirk, Andy. (2016). Data Visualization: A Handbook for Data Driven Design. SAGE

Kuek Ser Kuang Keng and Google News Initiative. (2020, June 23). Data Journalism & Visualization for Rookies: Discover Stories from Data. YouTube. ["Visualization as exploration tool" (1:04:13) https://www.youtube.com/watch?v=pAD87LvYagw]