Lesson 6
Data Stories

Overview

This lesson prepares students to put together data and other journalistic elements into a story in three formats: news, feature, and investigative articles.

Duration: 1 hour

Introduction

It can be argued that data journalism is like any other forms of journalism that involves numbers or data. On this note, although data stories usually have more visual elements and wider source of data, they also follow the structure of most journalistic pieces, i.e., head, lead, nutgraf, and quotes.

Longer pieces follow the story of a character, expose a conflict, and -- if solutions-based -- call to action.

In this lesson, you will have a review of writing straight news and longform stories but with focus on data. This lesson will also emphasize that while numbers add value to the stories, people and marginalized communities should still be at the heart of data journalism for development.

Objective

At the end of the lesson, you should be able to draft data stories.

Lesson proper

Discussion flow:
1. Data in basic news stories
2. Data in feature stories
3. Data in investigative stories

DATA IN BASIC NEWS STORIES

Format

News stories follow the inverted pyramid format of beginning with the top most important information in the lead; then, a nutgraf that connects the angle of the story to supporting details and quotes from sources or experts; then, additional details to end the story. Length usually runs from 100 to 300 words.

Context, or the bigger picture, is usually embedded in the news (e.g., a comparison of a local figure with the national or international counterpart, or a historical comparison within a 52-week or five-year or ten-year timeframe). Without context, a reporter may, for example, report an astonishing story of millions of pesos of budget allocated to healthcare and education, but in reality, this could be just a drop in the ocean if compared with other allocations in the budget pie or previous comparable figures or percentages from different counterparts in a region.

Data in news

Good data stories in news format is not just presenting numbers. Nor is any social media post with statistics in it, a data story. Data stories may have come to mean any act of journalism that touches electronically held records and statistics, or "virtually all of journalism" (Cohen, 2020), but they also add context, put happenings in a perspective, and hold leaders accountable.

DATA IN FEATURE STORIES

Format

Feature stories usually run up to 1,500 words, with lead and nutgraf "slower" in unfolding a story than a straight news does. But, the longer form gives more space for characters to develop, conflicts to arise, and actions to build. A minimum of three sources is ideal to triangulate an issue, to explore a topic, and to have a more in-depth analysis of a social problem. Digital feature stories these days are usually multimedia in format, that is accompanied with embedded video, infographics, photos or gallery of photos, and sometimes even a podcast.

In writing the lead, it is best to avoid just a plain echoing of what numbers or statistical figures institutions or organizations have reported. Instead, a better practice is to amplify what the statistics mean to a community or emphasize a public policy flaw by highlighting certain statistics. Like in news stories, data in feature stories need to be verified and assessed, even interrogated. Data journalists need to verify data: is the data true? What/who is the source? Who collected the data set? Why was the data collected? What does the source not say? (de Jong, 2021).

Once verified, the challenge for the data journalist is to unpack the cognitive load of data and tell it to the readers in a comprehensible way. Remember, most readers of journalistic pieces are "non-data-nerd" people who just want to be informed, educated, and amused. Winny de Jong in "Data Journalism Should Focus on People and Stories" calls this "limited nerdery." You want your reader or listener to think about your story, not process data. The story should prevail.

==

Learning activity

Look up the winners of the latest Sigma Awards for Data Journalism and browse the shortlisted projects' description, impact, techniques/technologies used, hardest part of the project, and lessons for others. Pick a story that you think can be replicated in the Philippine media and write a localized project description.

==

DATA IN INVESTIGATIVE STORIES

Format

Interviews, public documents, open source data and visualizations are the basic elements of a data-driven investigative story--like the feature story. However, the weight of investigative stories lay on the immediate social impact. In *The Data Journalism Handbook*, Rey (2021) said "data-driven reporting need not end when the editor clicks publish. It is proof that combining data journalism with online and offline community engagement can lead to positive policy and social change." A popular data-driven investigative story is the Panama Papers exposé by the International Consortium of Investigative Journalists.

Data in investigative stories

Storyboarding, a visual online, is a strategy in organizing story ideas and pre-writing to ensure that there is a desired flow and enough cases in the story. You can start low tech, with the use of a whiteboard, post-its, or a large notebook.

Next, data and sources need to be verified. There is a journalism adage that goes: if it's raining outside, your job as a reporter is not just to report that Person 1 said it's raining, and Person 2 said it's not; you have to go out, and check it out yourself. Once verified and triangulated, the next challenge is to present the story in a crisp, digestible manner. As Bill Kovach and Tom Rosenstiel (2007, cited in Data Journalism Handbook) said, the first task of the journalist is "to verify what information is reliable and then order it so people can grasp it efficiently."

Techniques

To produce a data-driven story, a journalist needs to filter, analyze, and discover patterns in and around structured facts and statistics. But how can you tell a data story to people who have an aversion for numbers? Duarte (2019) recommends comparing incomprehensible numbers to familiar things:

1. Attach data to something relatable. For example, when the annual earnings of Amazon boss Jeff Bezos were reported to be at $130 billion, Neil deGrasse Tyson tweeted: "… JeffBezos' 130-billion dollars, laid end-to-end, can circle Earth 200 times then reach the Moon & back 15 times then, with what's left over, circle Earth another 8 times." Or, as Duarte said: "The thickness of 130 billion stacked one-dollar bills is 8,822 miles, which is equivalent to driving back and forth across the United States 3.4 times.

   For a local example, the Great Pacific Garbage Patch, the ocean region full of trash, is 5.3 times the size of the Philippines.

2. Develop a sense of scale to help readers understand magnitude. Find something that is about the same size as the number.

   a. **Length**. Body parts (an arm or a leg), an atm card, or distances between provinces can be used to connect data to relatable size.

b. **Areas**. Relatable areas are basketball courts and football fields. In Japan, the number of tatami mats are sometimes used to measure or estimate the area of a room.

c. **Volume**. Readers can understand numbers better if they can see it or they have seen an equivalent of it. Relatable volumes are a shipping container or a gymnasium.

d. **Time**. Even saying the number of hours or days can have no impact, so a more creative and effective way is through relatable time such as the flights between cities (e.g., Manila to Ilocos), an episode of a sitcom, a TED talk, the time it takes to microwave a popcorn or boil an egg.

Human story

Finally, remember that data stories can have "a role to play in empowering marginalized communities to combat injustice, inequality, and discrimination" (Constantaras, 2021). Data journalism for development is the story of people who have no voice, people who are in the gutters of society, people who long for change.

REFERENCES

Cohen, Sarah. (2021). "Ways of Doing Data Journalism." The Data Journalism Handbook 2: Towards A Critical Data Practice. Retrieved from https://datajournalism.com/read/handbook/two/working-with-data/experiencing-data/data-methods-in-journalism

Constantaras, Eva. (2021). "Data Journalism by, about and for marginalized communities." The Data Journalism Handbook
2: Towards A Critical Data Practice. Retrieved from https://datajournalism.com/read/handbook/two/training-data-journalists/data-journalism-by-about-and-for-marginalised-communities

De Jong, (2021). "Data Journalism Shoud Focus on People and Stories." The Data Journalism Handbook 2: Towards A Critical Data Practice. Retrieved from https://datajournalism.com/read/handbook/two/working-with-data/experiencing-data/online-devices-and-their-research-affordances-for-data-investigations

Duarte, Nancy. (2019). Data Story. Canada: Ideapress.

Kayser, Bril Nicolas. (2021). "Become Data Literate in 3 Simple Steps." The Data Journalism Handbook 2: Towards A Critical Data Practice. Retrieved from https://datajournalism.com/read/handbook/one/understanding-data/become-data-literate-in-3-simple-steps

Rey, Aika. (2021). "Mapping crashing incidents to advocate for road safety in the Philippines." The Data Journalism Handbook 2: Towards A Critical Data Practice. Retrieved from

https://datajournalism.com/read/handbook/two/doing-issues-with-data/mobilising-for-road-safety-in-the-philippines

Exercise 3: Putting the story together

Overview

Your data sets from previous exercises will remain just data if you will not start writing a story. Based on your previous development journalism writing exercises, everything starts with planning on how to tell your story. At this point, one of the questions that can be looming inside your head will be how to make sense of your data set and transform it into a data story?

Making sense of that voluminous data and identifying the type of story to be produced largely depends on your appreciation and your chosen angle. It is actually not that very different when you write a news feature article. There are quite a number of story types that you can choose from depending on the data set on hand. In Devcom, we always reiterate that you should know your audience.

To help you structure your story, you will have an option to choose among the five types of data story based on what your data set can answer:

TYPE OF STORY	DATA SET SHOULD ANSWER THE CORRESPONDING QUESTION:
Factoid	Why does this one data point stand out from the others?
Interaction	Why do these aspects of the data change with each other?
Comparison	What is the meaningful difference between these parts?
Change	What made this part of data change from this to that?
Personal	How does this data relate to the lives of the audience?

Lastly, an important part of a data story is visualization.

A good visualization must facilitate understanding. Representation, in visual form, is important in communicating the key insights that will be a part of your story. How do you show an increase of budget allocation for education at a given time? Aside from representation, you must be able to present it well by adding additional visual decisions that complete your visual story.

In this exercise, you will be working with your cleaned and processed data and create intuitive data visualizations. Remember to integrate the lessons on visual design and techniques you have learned in DEVC 103: Visual and Audiovisual Media Production.

Objectives

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

1. Write an outline with a clear thesis statement.
2. Write a short narrative, with data visualizations.
3. Reflect on the process of producing a data story.

Mechanics and Deadlines

1. During the laboratory class, meet with your group mates to review your data set. Based on your previous output, identify the angles on the story each of you will be working on. Write an outline and have this approved by your professor.
2. Choose among the five types of data story and draft a development-oriented narrative.
3. Discuss with your teammates the challenges of writing the data story, what your team did to overcome this (or what your team could have done), and your overall insight/s in producing a data journalism project this semester. Share this with the class if there's still time.

TYPE OF STORY	QUESTION TO ANSWER	DESCRIPTION OF THE DATA STORY
Factoid	Why does this one data point stand out from the others?	Sometimes in large sets of data, you find the most interesting thing is the story of one particular piece of information. This could be an "outlier" (a data point not like the others), or it could be the data point that is most common. A detail about one particular piece of your data can fascinate and surprise people. It can also give them an easier way to start thinking about the whole set of data.
Interaction	Why do these aspects of the data change with each other?	When two aspects of your data seem related, you can tell a story about how they interact. The fancy name for this is "correlation". If one measure goes up, the other goes up too. If one goes down, the other goes down. In other cases, they might interact as opposites (when one goes up, the other goes down). You need to be careful not to

		guess about reasons for the interaction, but noticing the relationship itself can be a good story that connects things people otherwise don't think about together.
Comparison	What is the meaningful difference between these parts?	Comparing between sections of your data can be a good way to find a story to tell. Often one part of your data tells one story, but another part tells a totally different story. Or maybe there is a smaller portion of your data that serves as an example of an overall pattern.
Change	What made this part of data change from this to that?	People like to think about how things change over time. We experience and think about the world based on how we interact with it over time. Telling a story about change over time appeals to people's interest in understanding what causes change, and they can often remember seeing the differences.
Personal	How does data relate to the lives of the audience?	Some stories are interesting because they connect to real life. Personalizing the story creates a connection to the real-world meaning of the data and can be a powerful type of story for small audiences. Stories about someone's personal experiences can make the data seem more real.

Expected Outputs

1. Data story with a max word count of 1,200.
2. A 200-word reflection on the team's experiences in producing the data story.
3. Group presentation (optional)

Assessment

Criterion	Performance Level		
	8-10 points	**4-7 points**	**1-3 points**
Analysis of dataset and overall story (Multiplier: 4)	Data story is excellently presented in terms of logic and organization. Data is presented in a manner understandable to target audience.	Data story is technically good but inconsistencies can be observed in the quality of various components. Data presentation looks somewhat complex and needs a bit of cognitive processing to understand.	The technical quality of the data story and inconsistencies can be observed. Data is incomprehensible.
Development orientation of the output (Multiplier: 4)	The output clearly established the development orientation of the topic, localized the topic/issue, and provided sufficient details to establish the context of the topic.	The output fairly established the development orientation of the topic, localized the topic/issue, and provided sufficient details to establish the context of the topic.	The output was unable to establish the development orientation of the topic and failed to localize the topic/issue. The output did not provide enough details to establish the context of the topic.
Style and creativity (Multiplier: 2)	Excellent show of creativity and uniqueness. Style in format and language is excellent or very good.	Notably good show of creativity. Style in format and languages is good but with a few errors.	Creativity is a no-show. Style is hurting content.

Total

Perfect score=100 points