Writing Workshop Day 2

Summarize your manuscript

The Storm Events database collects information from the National Weather Service regarding weather events that occur in the United States. Many researchers use this database to conduct studies on how weather events impact human life, property and crop damage, and commerce. Though the database is quite extensive, several concerns and biases may be present in its data.

It is important to understand these biases so that studies using this dataset can accurately portray trends and changes caused by weather events.

Identify key Introduction components in an example paper

Twitter:

1. What is the larger problem that the paper will be tackling?
    - Arguing for the use of twitter in the science community
2. What is the relevant context for that problem?
    - The increasing popularity of social media (twitter in particular-320 mil active users) and the potential for the use of twitter to increase collaboration and spread information
3. What are the key characters?
    - Twitter and its users, researchers, the science community in general
4. What background information is given to help the reader understand the specific work in the paper?
    - It describes what twitter is and how it is used in general and in the science community
    - It gives some background on the authors of the paper
5. What is the "challenge" ("the specific hypotheses/questions/goals of the current work")?
    - educating researchers on how to use twitter to their benefit within science

Shwartz Ecological Fallacy:

1. What is the larger problem that the paper will be tackling?
    - defining and examining ecological fallacy as a general validity probblem
2. What is the relevant context for that problem?
    - epidemiology's wide use of ecological studies
    - epidemiology's use of ecological fallacy to describe the flaws in their studies
3. What are the key characters?
    - epidemiology, epidemiologists, Robinson, the use of ecological fallacy as a concept
4. What background information is given to help the reader understand the specific work in the paper?
    - epidemiology's use of ecological studies
    - previous research and papers that have tried to define ecological fallacy
    - examples of previous consequneces of ecological fallacy
5. What is the "challenge" ("the specific hypotheses/questions/goals of the current work")?
    - redefining ecological fallacy as a general validity issue
    - showing how this new definition will dissect ecological fallacy into three distinct fallacy notions

Diagnose the Opening for an example paper

Shwartz Ecological Fallacy: Who do you think is the intended audience for the paper? What is the larger issue the paper will address? Do you think that they are properly "advertising" what they will later cover in the paper?

I think that this article is trying to target epidemiologists and other researchers that use population data or are interested in this field of science. The first paragraph brings up epidemiology, ecological studies, and ecological fallacy. This leads the reader to believe that the paper will be about these topics, but it is unclear that the rest of the paper will be about redefining ecological fallacy.

Define Opening components for your paper.

What is the target audience? Is it broad / interdisciplinary or targeted to researchers in a certain field? What is the larger issue that the manuscript will address? In writing this, be sure to be clear about the scope with which the manuscript will cover this issue. What are a few elements of the issue that are interesting but that your manuscript will not address?

My manuscript is attempting to target scientists using the NOAA storm events database for research. It is broad in terms of how it discusses general bias in data, but specific in that it focuses on the NOAA database and its weather data. My manuscript covers what the NOAA Storm Events database is, examples of how it has been used in research, and 4-5 biases that occur in the dataset. Though my writing will cover 4-5 broad biases and include some examples of each type, there are several other examples that could also showcase the biases. There may also be other forms of bias in the data that I do not cover.

Revise the Opening for your manuscript.

Does your Opening need the whole paragraph, or just the first sentence or two? Does this first paragraph include any signals to clarify what audience you expect for the manuscript? Does the paragraph give readers clues on the larger issue that the manuscript will address? Does the paragraph "overpromise", indicating that the manuscript will cover a larger scope than it does?

What is the Storm Events Database

The National Oceanic & Atmospheric Administration (NOAA) Storm Events database collects information on storms, significant weather phenomena, rare or unusual weather, and other significant meteorological events across the United States. NOAA stores this information in csv files for each year. It has been recording weather information since 1950. However, the database has undergone several changes since then. In 1950, NOAA only recorded tornado events and now it records 48 different event types. Following these changes, the NOAA Storm Events website states that they have reformatted and standardized events types without changing any specific values or details [@stormeventsdetails].

The database assigns each weather phenomenon with a location, date, event type, event ID, episode ID, property damage estimate, crop damage estimate, county name, state, event narative, and episode narrative. The database categorizes large storms as episodes which contain several individual events. For example, a hurricane will be given one episode ID, and the rain, wind, floods, etc associated with that hurricane will be given event IDs that fall under that episode.

The US National Weather Service (NWS) collects weather data from storm trackers, federal agencies, the media, the public, and several other sources. The NWS then uses this data to create the NOAA Storm Events database that is released as a monthly publication.

Concerns with the Database

Though the NOAA Storm Events database is quite extensive, several biases affect its data. These biases fall under five major categories which are common in weather data. These include hazard bias, temporal

bias, threshold bias, accounting bias, and geographic bias. These biases result from various factors including structural changes over time, reporting errors, inherent bias, and others.

The Storm Events website gives a disclaimer that "an effort is made to use the best available information but because of time and resource constraints, information from [the] sources may be unverified by the NWS" [@stormdatafaq].