**Gabriel Andres Soto, Sam Cohen, Daniel Cardenas Sanchez**

**PPOL 5203**

**Final Project Proposal**

Introduction and Research Question

With the advent of the internet and the increase in Smartphone usage across the world, the speed at which individuals can now obtain information and news has never been faster. This interconnectedness comes with problems, however. The spread of misinformation, political extremism, and polarization are just a few consequential examples. This phenomenon provides researchers with a looking glass into public sentiment regarding certain issues, which can often be found in some of the darkest places on the web– Youtube’s comments sections.

Our research will attempt to answer the following research question: **do individuals react differently to different types of shocking news on Youtube?** Will those comments be mostly “positive” (i.e. concerned, empathetic, well-intentioned) if the news is more serious? Will they be mostly negative (ill-intentioned, bullying, insulting)? We will also include the following when writing our report and carrying out our analysis:

1. What is shocking news?
2. Difference in shocking news on Youtube? Are the kinds we are talking about or others?
3. Why will these types of comments be different from others on and off Youtube?
4. What patterns are there in Youtube comments: How is this platform unique, and does this foster different reactions?
5. How do human reactions differ in regards to horrific, scandalous, and/or pop culture news?

Understanding this type of reaction can help policymakers create better strategies to mitigate polarization and extremism in social media.

Methodology

We will use Natural Language Processing (NLP) and sentiment analysis to analyze Youtube comments pertaining to 3 different types of shocking news: “horrific” news (Uvalde school shooting, though in previous iterations of our proposal we were considering the outbreak of the Russia-Ukraine war), “scandalous” news (the indictment of Robert Menendez), and a “pop” news (Taylor Swift dating Travis Kelce). Reaction to popular videos of non-shocking news (MrBeast channel) to compare with the others.

We will use Youtube’s API to retrieve data on comments for news videos regarding these three stories. We will analyze the first 100 videos that come from a Google search on the topic, use Youtube's API to retrieve all the comments of each one, and use NLP and sentiment analysis to answer our research question. For the bulk of our analysis, we will analyze comments that were made within the first week of when the video was posted, due to what we predict will be more emotionally-driven comments. We will then compare older comments to newer comments

Data

To answer our research question, we will use data on Youtube comments from Youtube’s API.

100 videos per each bin (Google’s API restriction) \* 4 bins (categories of videos) = 400 videos in total. For all of those we will get and process all the comments.

Anticipated Findings (Hypothesis)

* We anticipate that reactions to shocking news will tend to be more extreme than those from a normal video (i.e. our control).
* We also predict that reactions to horrific situations will be more positive (empathetic) than those of the serious news and both more so than the pop culture shocking news.
* Reactions on the first week of a video will be more extreme than those afterwards.
* Newer comments (after the first week) on shocking videos will be more similar to the non-shocking video.

Anticipated Problems

One of the biggest problems we will likely face is the fact that we are choosing concrete situations and videos that could be biased. Not only is it likely that we will observe comments on videos from biased news sources– it is also possible that the reactions on one news video will be substantially different from those of others, even if both videos cover “horrific” news. In addition, it is possible that situations are very specific, and certain aspects of them cause the reactions seen and not others (i.e. people love Taylor Swift and it could be shocking and serious news for them whoever she is dating). We are also only analyzing comments in English, which may give us a biased perspective: comments on a news story regarding Russia-Ukraine may be more supportive of Ukraine than comments in Russian, for example.

What a successful project will look like

Success, in this case, would mean that our data paints a clear picture of how people decide to comment on Youtube. This can either mean that our predictions will be true (i.e. more horrific news pushes people to comment more positively), or that our data will give us a different picture on how and why people comment the way they do. For example, if our prediction is false, can we say with confidence that horrific news actually influences people to comment negatively?