Nicholas McFadden

Mcnichol

6/12/2022

Final Project: School Mass Shooting Twitter Comparison

I choose to focus my project on tweets from the two weeks after both the Uvalde and Parkland school shootings. Tragically mass shootings have become more frequent, and I was curious to see how the public has reacted differently between the two. Not to say any mass shooting is exactly like another I chose to compare Parkland to Uvalde because Parkland was the most recent school shooting with a similar death toll. I attempted to pull 1000 tweets from each event by the following queries:

Uvalde: '(uvalde OR "robb elementary" OR #uvalde OR #robbelementary) AND (#rip OR robb OR elementary OR mass OR school OR shooting OR massacre OR shooter ) min\_faves:10 filter:has\_engagement until:2022-06-07 since:2022-05-24')

Parkland: '(parkland OR #parkland OR #parklandhighschool) AND (#rip OR "high school" OR mass OR school OR shooting OR massacre OR shooter) min\_faves:10 filter:has\_engagement since:2018-02-14 until:2018-02-28'

They both have identical words except for the respective names/hashtags and the word elementary vs high, since Parkland is a high school and Uvalde an elementary school. My pull from twitter resulted in the following counts of the twitter data:

* Table

  Description automatically generatedTable

  Description automatically generated

The questions I am looking to answer from my analysis are:

* What hashtags are the most used for Uvalde vs Parkland?
  + Field being used: ’hashtags’
  + Once isolated a frequency distribution was applied to the hashtags and only the top 10 are shown
  + Graphical user interface, table

    Description automatically generated
  + Analysis: Not surprising the top hashtags were their respective city names (as we made to search for these specifically). Interestingly we can see that the top Uvalde hashtags are less about gun control/reform compared to the Parkland hashtags. Potentially signaling the fight for gun control losing its steam over the many years without reform.
* From what platforms were people using the most and which has the most engagement?
  + Fields being used: ‘sourceLabel’, retweetCount’, ‘quoteCount’, ‘likeCount’
  + Once isolated a frequency distribution was applied to the hashtags and only the top 10 are shown
  + Then grouped by source the tweets’ sum and avg are calculated for engagement coming from each platform
  + Graphical user interface

    Description automatically generated with medium confidence Graphical user interface

    Description automatically generated with medium confidence
  + Table

    Description automatically generated Table

    Description automatically generated
  + Table

    Description automatically generated Table

    Description automatically generated
  + Analysis: Iphones seems to be a popluar choice under the Web App in across both events . This makes sense since people often choose to tweet from their phones and maybe signals that of the tweets there are less coming from businesses which is good for our analysis.
* What hashtags to the two events have in common and how many times are they each reference respectively?
  + Fields: Dataframes from the already calculated hashtag count
  + The similar hashtags are pulled and joined and then their frequency distribution is applied to see the difference is amount of use
  + Table

    Description automatically generated with low confidence
* Were the respective Governors in the top 10 mentions?
  + Fields: ‘mentionedUser’
  + The mentioned users were within an object so once extracted the instances are put into a frequency distribution and compared.
  + Graphical user interface, application

    Description automatically generatedGraphical user interface, application

    Description automatically generated
  + Analysis: Abbott has the most mentions from the Uvalde while Trump was the most from Parkland and Ron DeSantis is not in the top 10.
* What are the most used words surrounding the topics?
  + Fields being used: ‘content’
  + The text is cleaned and then put through tokenization to determine the frequency of words
  + Graphical user interface, application

    Description automatically generatedGraphical user interface, application

    Description automatically generated
  + Analysis: Interestingly the shooting as a key word drops to 3rd in Uvalde bu the other top words are more associate with the tragedy such as ‘gun’ and ‘mass’ .
* Which is the difference in sentiment between the two events?
  + Fields being used: ‘content’
  + Tokenization is applied to the text but leaving in the hashtags to provide extra context then the VADER sentiment analyzer is applied and descriptive statistics are applied to each
  + Uvalde:
  + Table

    Description automatically generated
  + Parkland:
  + Table

    Description automatically generated
  + Analysis: Uvalde’s average sentiment is more negative than that of Parkland but Parkland has more volatility across the sentiment profiles.

This program scrapes twitter for tweets regarding the Uvalde and Parkland school shootings. It enumerates over these tweets and appends portions of the data into specific lists. The lists are then put into pandas dataframes for later analysis. When the data is viewed, we can see that the ‘hashtags’ and ‘mentionedUsers’ field are a nested list and an object so in order to run calculations we will need to extract the original values. Next, we apply the groupings to the data based on the questions being asked and apply frequency distributions. Text content is tokenized and the VADER sentiment analyzer is applied. The sentiments are gathered, and descriptive statistics are applied. Then the hashtags that are found are matched and placed in the same dataframe to see the occurrences comparatively. Finally, exporting my dataframes to an excel spreadsheet but in order output the respective ‘all\_tweets\_df’ the time zones from the ‘date’ field must be removed.

My conclusions from this analysis are that when Parkland happened there was a larger debate specifically around gun control and around the NRA and more conservative political terms (aka ‘2a’). With the conversation try to fight potential gun control and how Trump at the time would be handling gun rights issues. While the public seems more fed up and angry this time around while the conversation is moving away from the national gun control conversation and being directed towards their local public officials to try to make change that hasn’t happened since Parkland. As terrible as it sounds, I am glad that the overall sentiment is getting more negative because these events are terrible and maybe with more people speaking out about how these things need to change maybe we will actually see some change. The resistance to gun control could either be wanning or as I said earlier maybe the conversations will look for other methods of reform.