**Covid-19: (con)textualized**

By Ami’s Angels

(Alena Abella, Emma Broback, Linus Jen, Josh Kong, and Tyler Wu)

<https://docs.google.com/presentation/d/1JkBR8qh8Ys-yPHNkGEgc128_lA1CMlrhyDzQG_uJVB0/edit?usp=sharing>

**Introduction**

Our group decided to analyze how different politicians changed their social media presence with the development of COVID-19. We chose to analyze the Instagram profiles of Barack Obama, Donald Trump, Mike Pence, Bernie Sanders, Elizabeth Warren, and Ted Cruz, though much of our analysis does not include Ted Cruz’s online presence.

To run our analysis, politicians’ Instagram posts were scraped to include likes, views, captions, post type, and top comments. Models comparing likes to number of followers were made, as were comparisons between follower count, likes, and changes to these values throughout time. From the captions, trends regarding changes in key term usage was compared, as were bigrams showing common links between terms. Using the afinn, bing, and nrc lexicons, we compared the different sentiments expressed in the captions of each politician. This allowed us to visualize what kinds of feelings were being conveyed to their followers, and how these feelings changed over time as COVID-19 gained more attention. We also compared the change in terms after key events that occurred. A full list of these events were pulled from the New York Times, though we shortened the list to the most important dates affecting the United States.

**Assumptions**

We made several key assumptions that defined how we handled the data. Because Twitter limited the tweets that could be pulled, we used Python to manually scrape the captions of each politician’s captions. Because captions were scraped, the images or videos themselves share emotions not captured in the captions. When running our sentiment analysis, we used the afinn, bing, and nrc lexicons. While widely agreed upon as correctly labeling words with either a value or an emotional term, there may be terms that may hold different meanings than how they were defined in the lexicon. For example, “trump” was defined as positive in the bing lexicon, but the President’s surname could be used both in a positive or negative manner, depending on context. We removed as many terms that may be interpreted incorrectly out of context. Lastly, all posts analyzed started from December 1st, 2019, and ended on May 10th, 2020. Thus, findings cannot capture the sentiments of politicians before December 1st, 2019.

**Citations:**

“Getting Started.” *Gganimate*, gganimate.com/articles/gganimate.html.

“Instagram.” *Instagram*, [www.instagram.com/](http://www.instagram.com/).

Shearer, Elisa. “Social Media Outpaces Print Newspapers in the U.S. as a News Source.” *Pew Research Center*, Pew Research Center, 10 Dec. 2018, [www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/](http://www.pewresearch.org/fact-tank/2018/12/10/social-media-outpaces-print-newspapers-in-the-u-s-as-a-news-source/).

Silge, Julia, and David Robinson. “Text Mining with R.” *Text Mining with R*, 7 Mar. 2020, [www.tidytextmining.com/](http://www.tidytextmining.com/).

Taylor, Derrick Bryson. “How the Coronavirus Pandemic Unfolded: a Timeline.” *The New York Times*, The New York Times, 13 Feb. 2020, [www.nytimes.com/article/coronavirus-timeline.html](http://www.nytimes.com/article/coronavirus-timeline.html).

*Trackalytics*, www.trackalytics.com.