Pfizer Vaccine Tweets: A Sentiment Analysis and Implications for Public Health

Team Members:

- Justin Ehly – jehly@smu.edu
- Nicole Norelli – nnorelli@smu.edu
- Mingyang Nick YU – nyu@smu.edu

Problem Statement:

Worldwide distribution of the Pfizer vaccine in response to the COVID-19 pandemic has been inconsistent, in part due to insufficient data on public perceptions of the vaccine. Deep analysis of public sentiment could assist the medical community develop a more effective distribution strategy, which would save lives.

This project will analyze real-world Twitter data related to the Pfizer vaccine to identify public perceptions of the vaccine as well as any trends that might be useful to researchers and government officials. We will use these trends to provide recommendations that could inform world-wide government actors to facilitate distribution, education, and public awareness to bolster confidence in the vaccine. The dataset is provided by Kaggle, and the resource link is listed below.

Methodology:

This project will encompass several analytic techniques to identify useful trends within public perception of the Pfizer vaccine. These techniques will include: deep data cleaning with R; EDA with R; natural language processing; Amazon cloud computing including S3, Comprehend, and other visualization tools.

The analysis may include but will not be limited to: location analysis on discussion trends, hashtag analysis on topics of high interest, and sentiment analysis on public perceptions of the Pfizer vaccine.

Milestones:

- Clean data and perform EDA on dataset to explore useful insights
- Utilize Amazon Comprehend on cleaned data and perform sentiment analysis
- Evaluate sentiment analysis, location analysis, and hashtag analysis to form recommendations and identify future research directions

Resources:

- https://www.kaggle.com/gpreda/pfizer-vaccine-tweets
- https://docs.aws.amazon.com/comprehend/latest/dg/comprehend-dg.pdf
- https://www.red-gate.com/simple-talk/sql/bi/text-mining-and-sentiment-analysis-with-r/