# A study of sentiment of COVID-19 related tweets in the USA

Truong (Jack) P. Luu, Rosangela Follmann

School of Information Technology
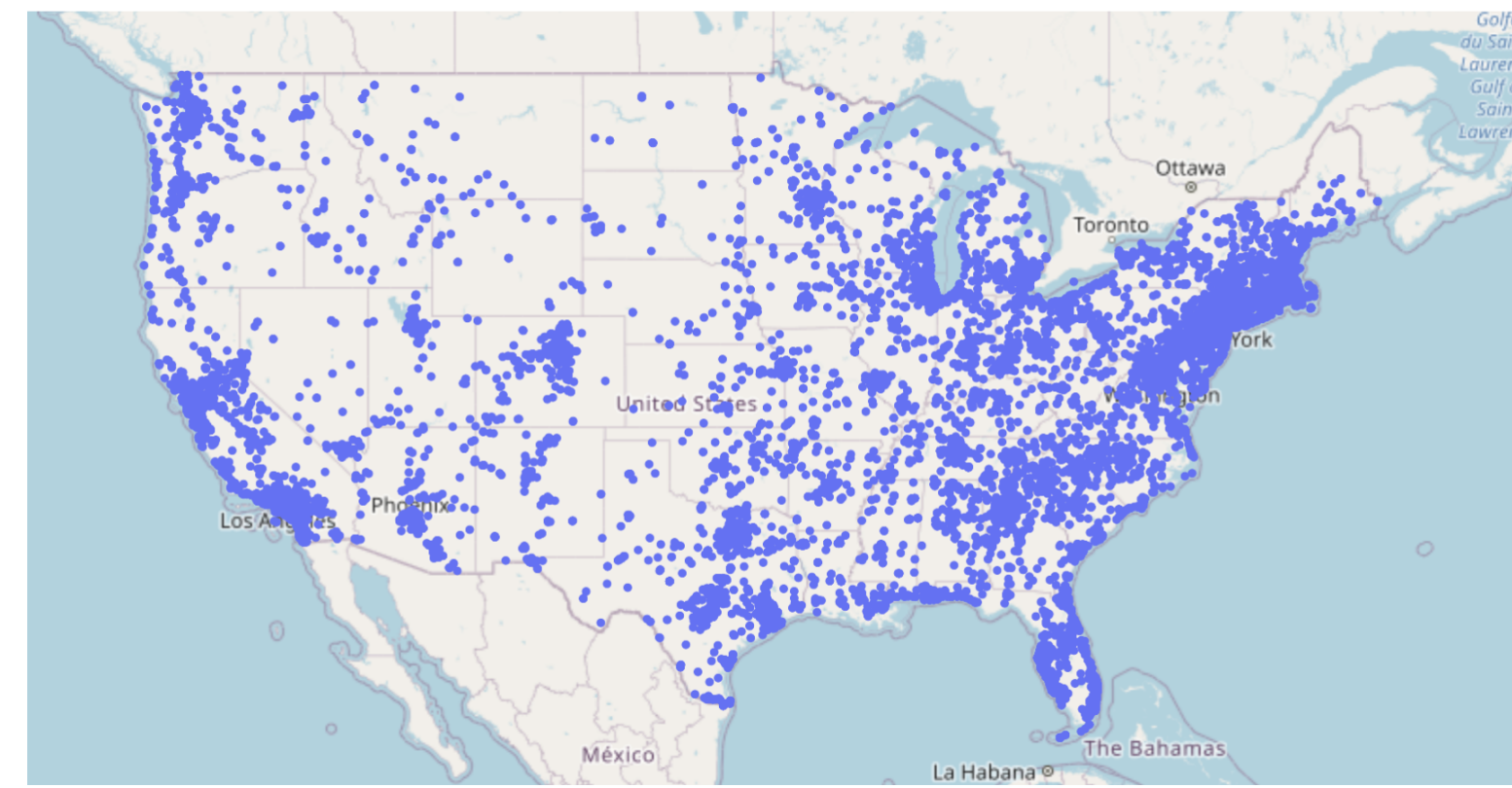Illinois State University, Normal, IL 61790

Introduction

During the COVID-19, many have expressed their feelings and opinions related to aspects of their lives via Twitter. In this study, we analyze COVID-19 related tweets that were generated in the USA from April to August 2020. We test the correlation between users' sentiment and COVID-19 cases across the USA, and investigate the effect of specific COVID-19 milestones on the sentiment scores.
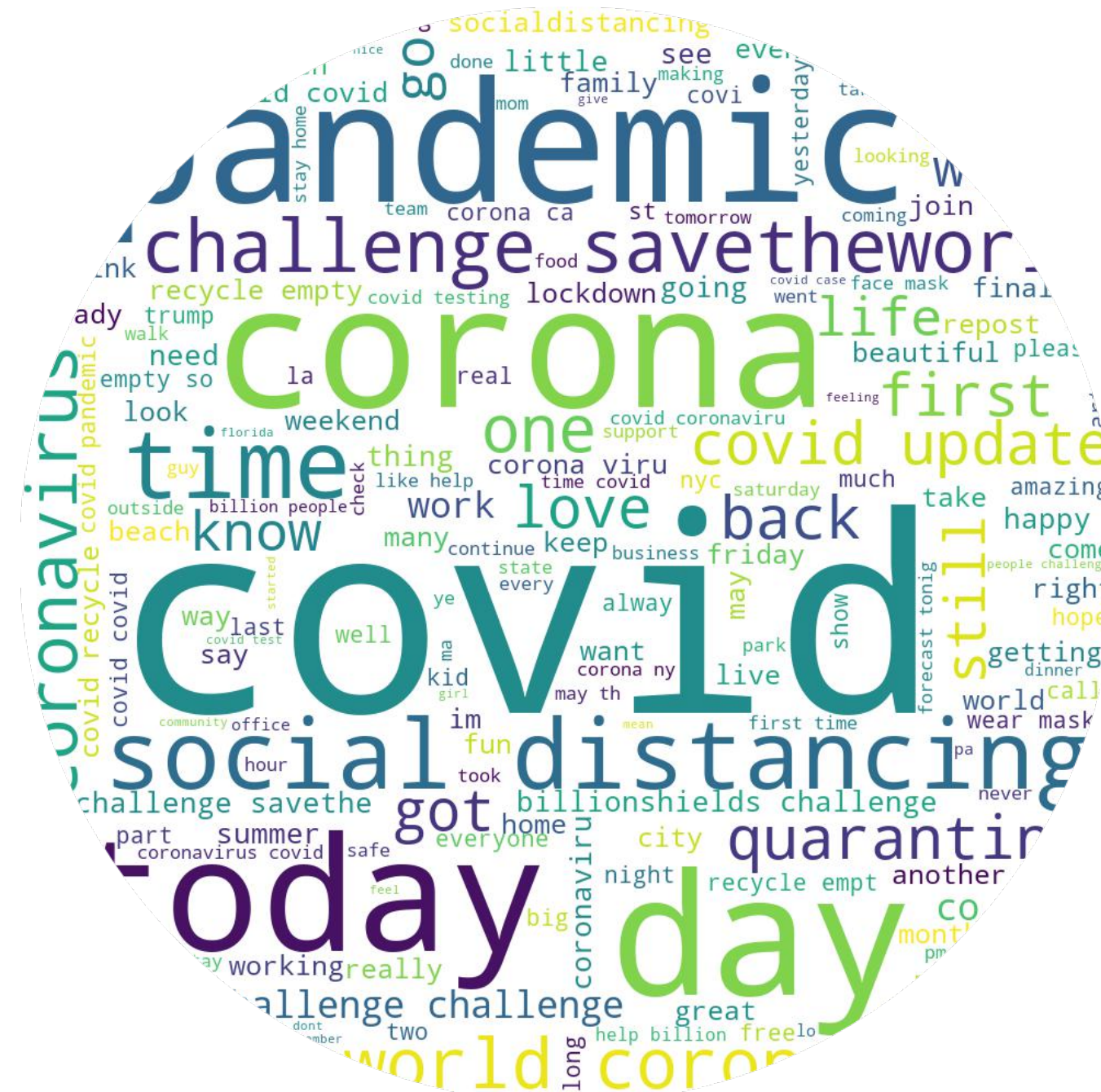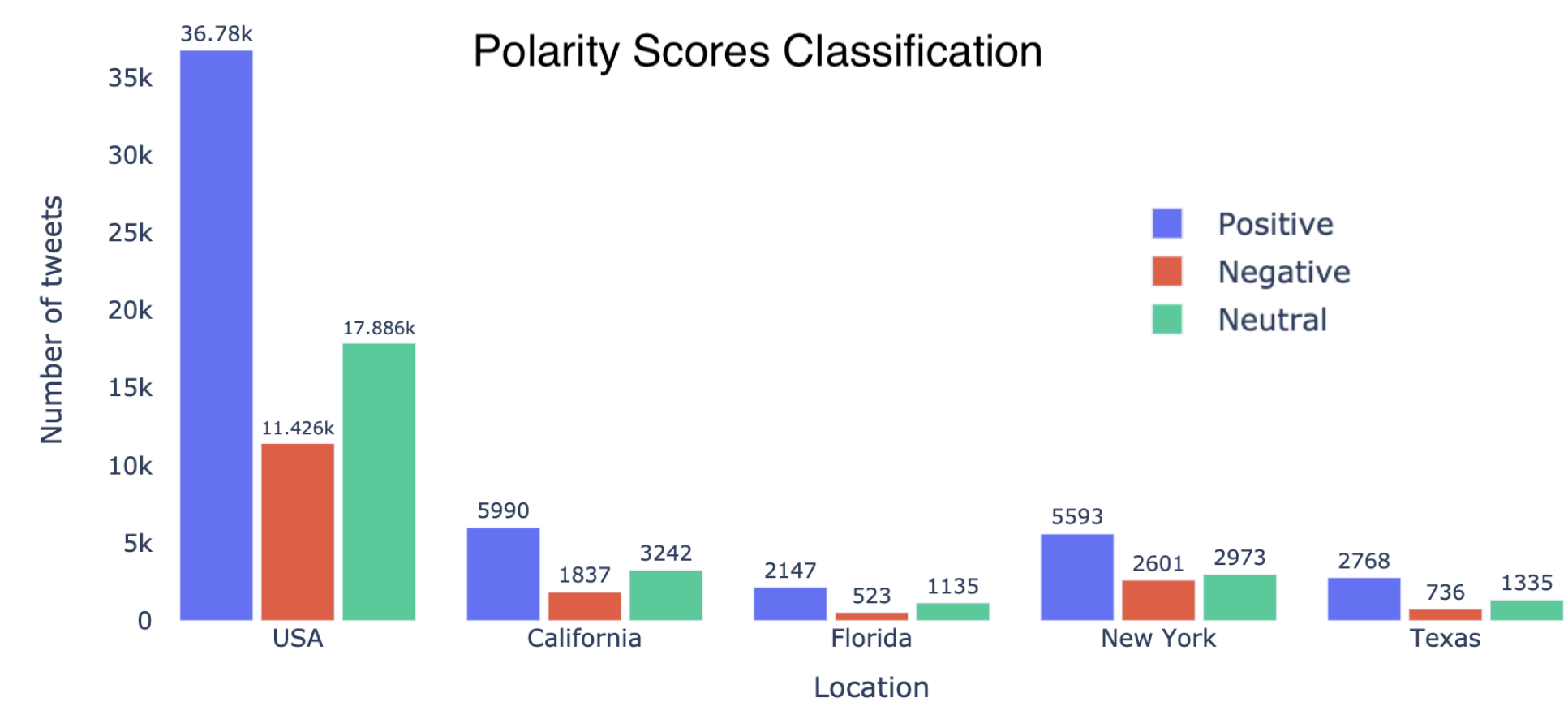

Word cloud of tweet messages.

Methodology

TextBlob - Open source python library for sentiment analysis. It can be used to determine the polarity (positive or negative) of a text along with its subjectivity [1].
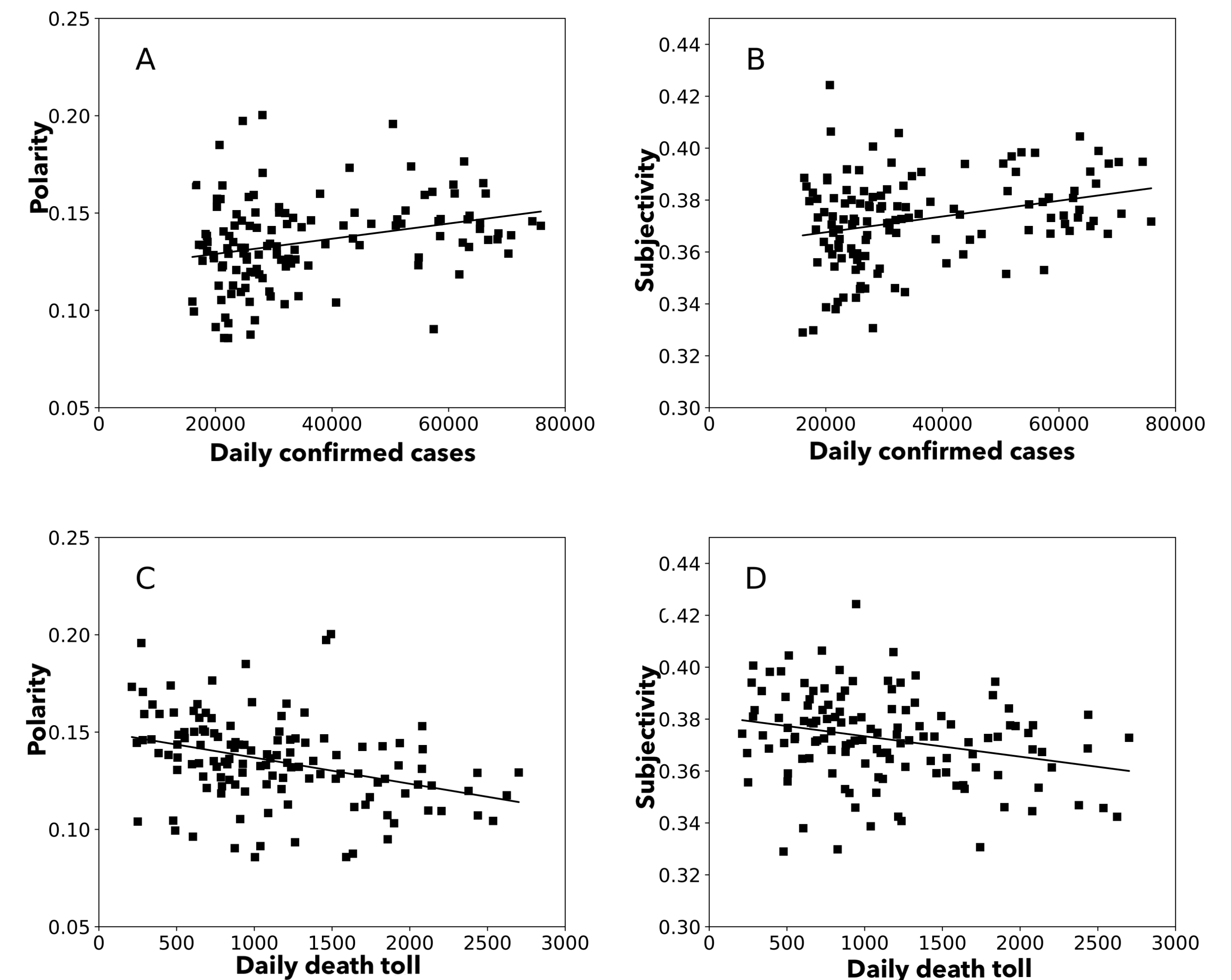
Tweets Distribution and Sentiment Score

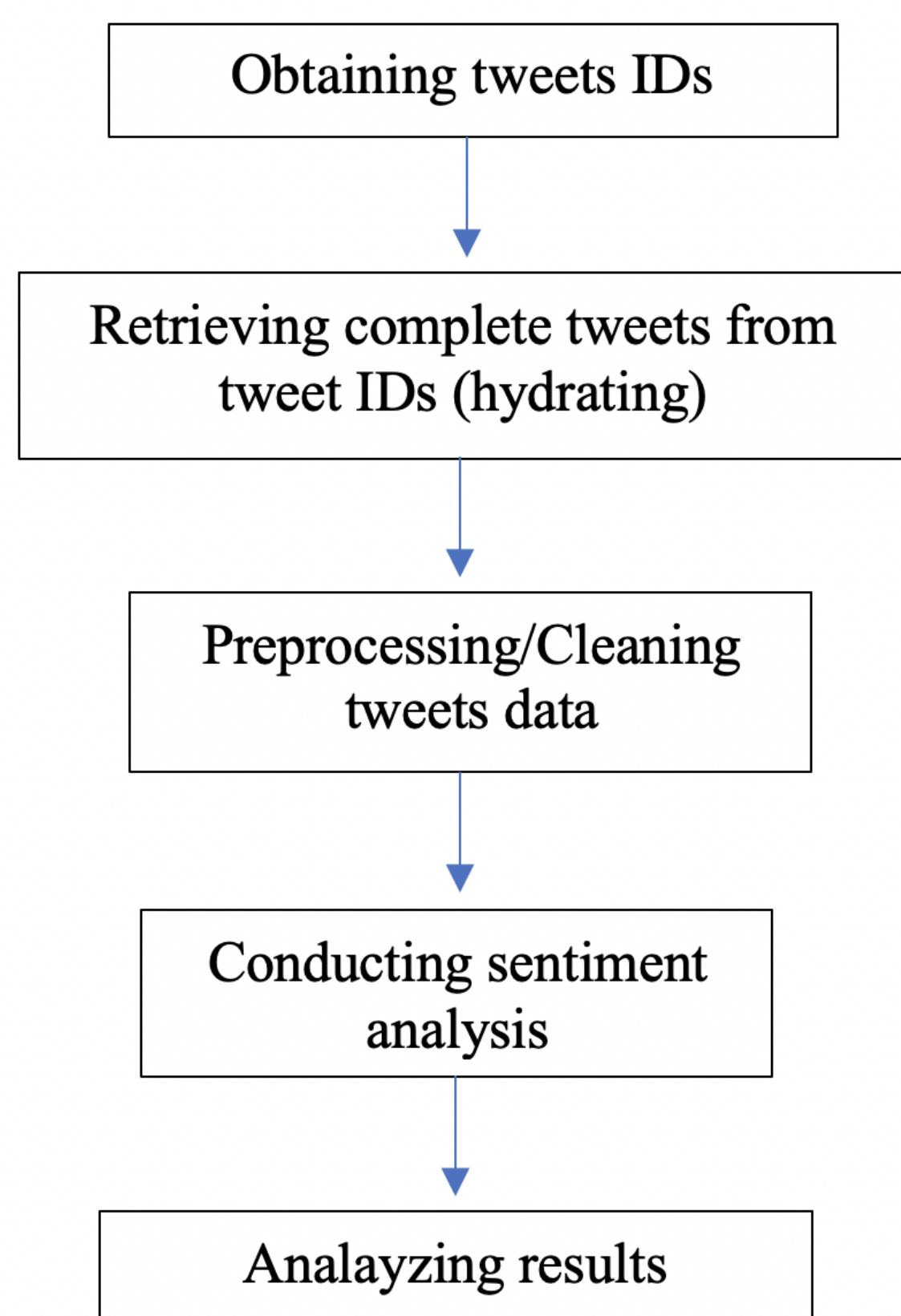Highly populated areas yield dense tweets distribution

The top four populous States (California, Texas, Florida, and New York) account for almost 50% of the COVID-19 related tweets. California: 11,069 tweets (16.7%), Florida: 3,805 tweets (5.8%), New York: 11,167 (16.9%), and Texas: 4839 (7.3%).


Polarity versus Subjectivity

Skewed distribution showing more points towards the positive polarity and higher subjectivity (>0.5), suggesting that the more positive-oriented a tweet is, the more opinion-oriented its meaning will be.
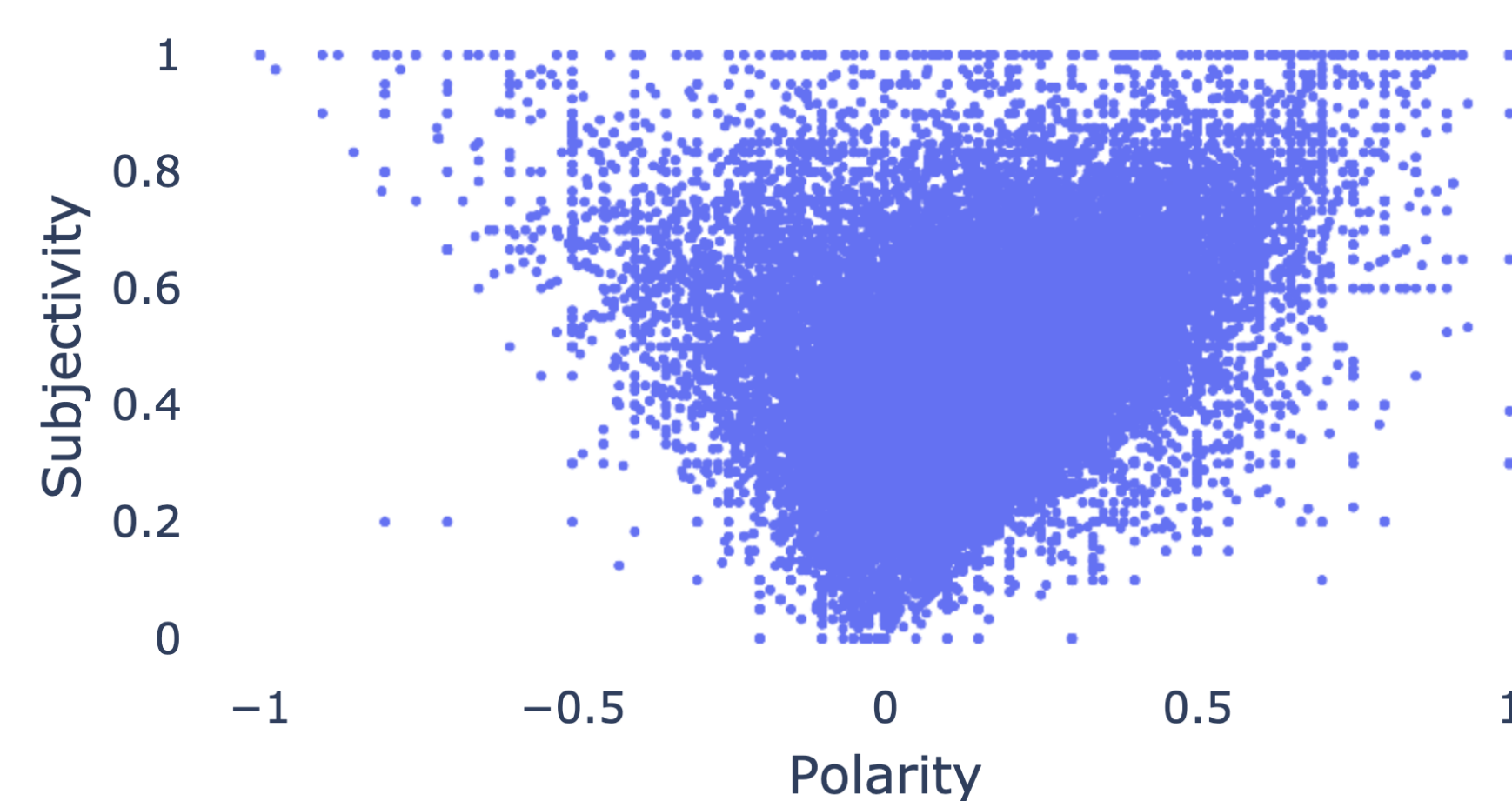
Subjectivity >0.5	Tweet counts	Percent
Polarity >0	15,640	78.3%
Polarity <0	3,612	18.1%
Polarity = 0	732	3.6%

Sentiment score versus COVID-19 cases

Reasoning between polarity score and the COVID-19 milestones
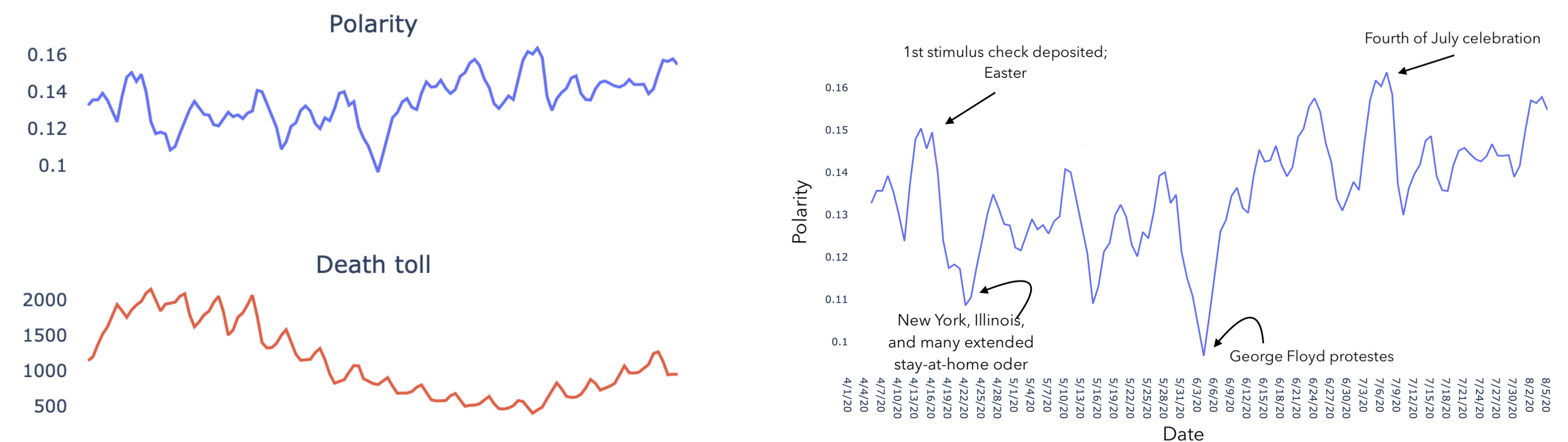
Significant events such as new regulations from the government, in this case are the stimulus check or stay at home orders, celebration of certain holidays such as Easter or the Fourth of July, and social conflicts as in George Floyd protests, can directly affect the sentiment of the public. The changes in sentiment can be captured through the change of the sentiment score.

Concluding Remarks

There is a connection between sentiment scores, and COVID-19 confirmed cases and death toll in the USA. Significant events, such as new regulations from the government, celebration of important holidays, and social conflicts, can directly affect the public's sentiment. As the future work, we plan to diversify the analysis by collecting a Twitter data that contain Emoji to build a sentiment classifier.

References and Acknowledgments

[1] Lamsal R. Coronavirus (covid-19) geo-tagged tweets dataset, 2020.
[2] Textblob: Simplified text processing — textblob 0.16.0 documentation.https://textblob.readthedocs.io/en/dev/.

Contact: pluu@ilstu.edu