Social Media Analysis of Sentiment Regarding Health During the COVID-19 Pandemic

Isaac Burright, Jack Burright, Liam Townley
Creighton University School of Medicine, Omaha, NE
Creighton University College of Arts & Sciences, Omaha, NE

Creighton UNIVERSITY — School of Medicine

Creighton UNIVERSITY — College of Arts and Sciences

Background

- Throughout the COVID-19 pandemic, social media has served as a public forum for health information and discussion among users.
- With over 450 million users, Twitter hosts abundant dialogue regarding health information, disinformation, facts and opinions.
- These topics are often emotionally charged, and investigating their tone can reveal individuals' perceptions towards health; on a larger scale, analyzing these posts in summation may provide insight into societal sentiment regarding health.

Study Methods

- Python was used as a Twitter scraper to collect a total of approximately 800,000 tweets that contain the word 'health'.
- Data consisting of 50,000 tweets per date were collected from posts dated on four specific days each year from 2019 to 2022.
- Tweets were then individually preprocessed with LIWC built in processor.
- Tweets were subsequently analyzed using a sentiment analysis algorithm and given a numerical sentiment score for both positive and negative tone.
- Data were then analyzed using paired and unpaired t-tests in Excel and statistical significance was set at $\alpha = 0.05$.

Emotion of Health

Tone of Health

Results & Conclusions
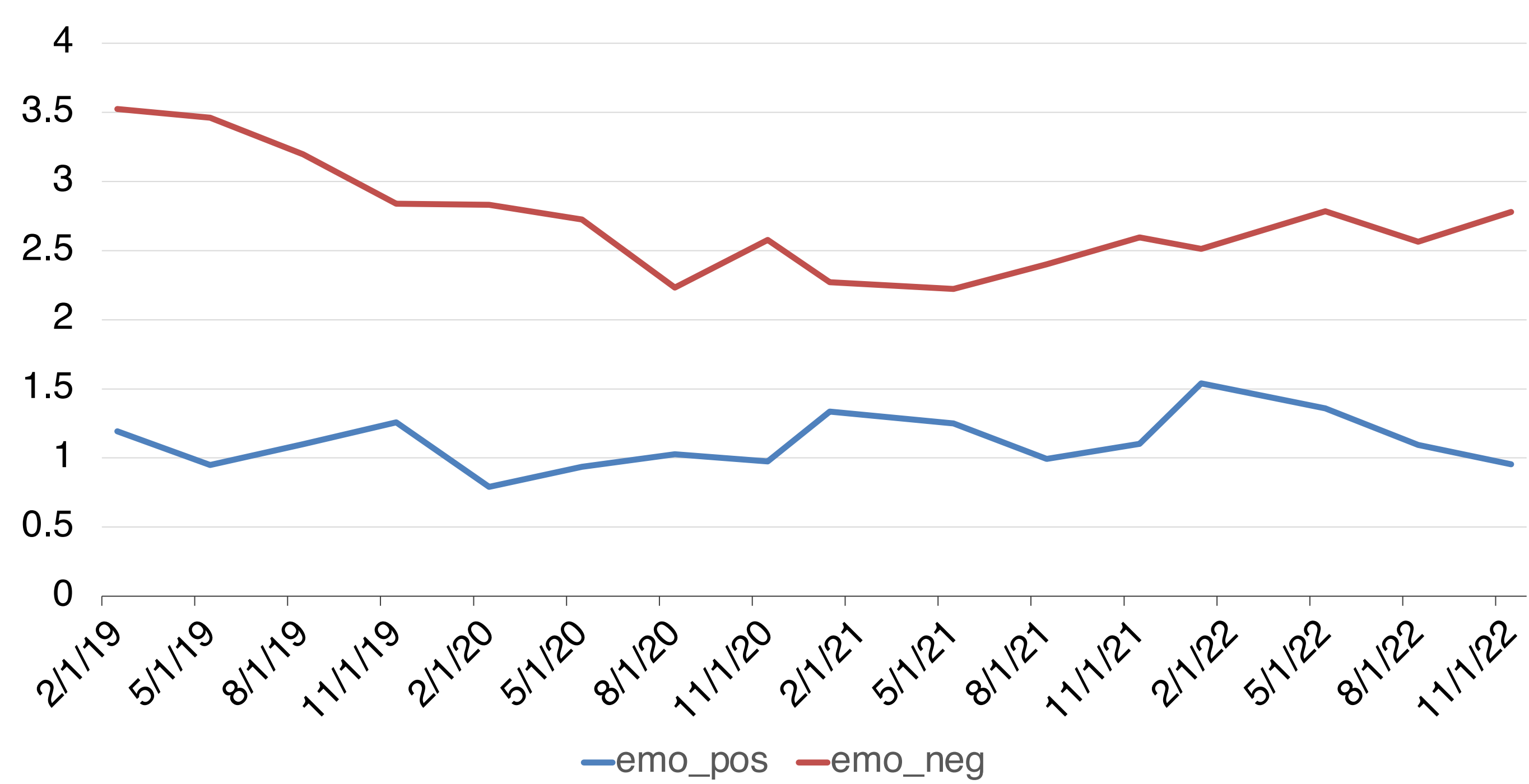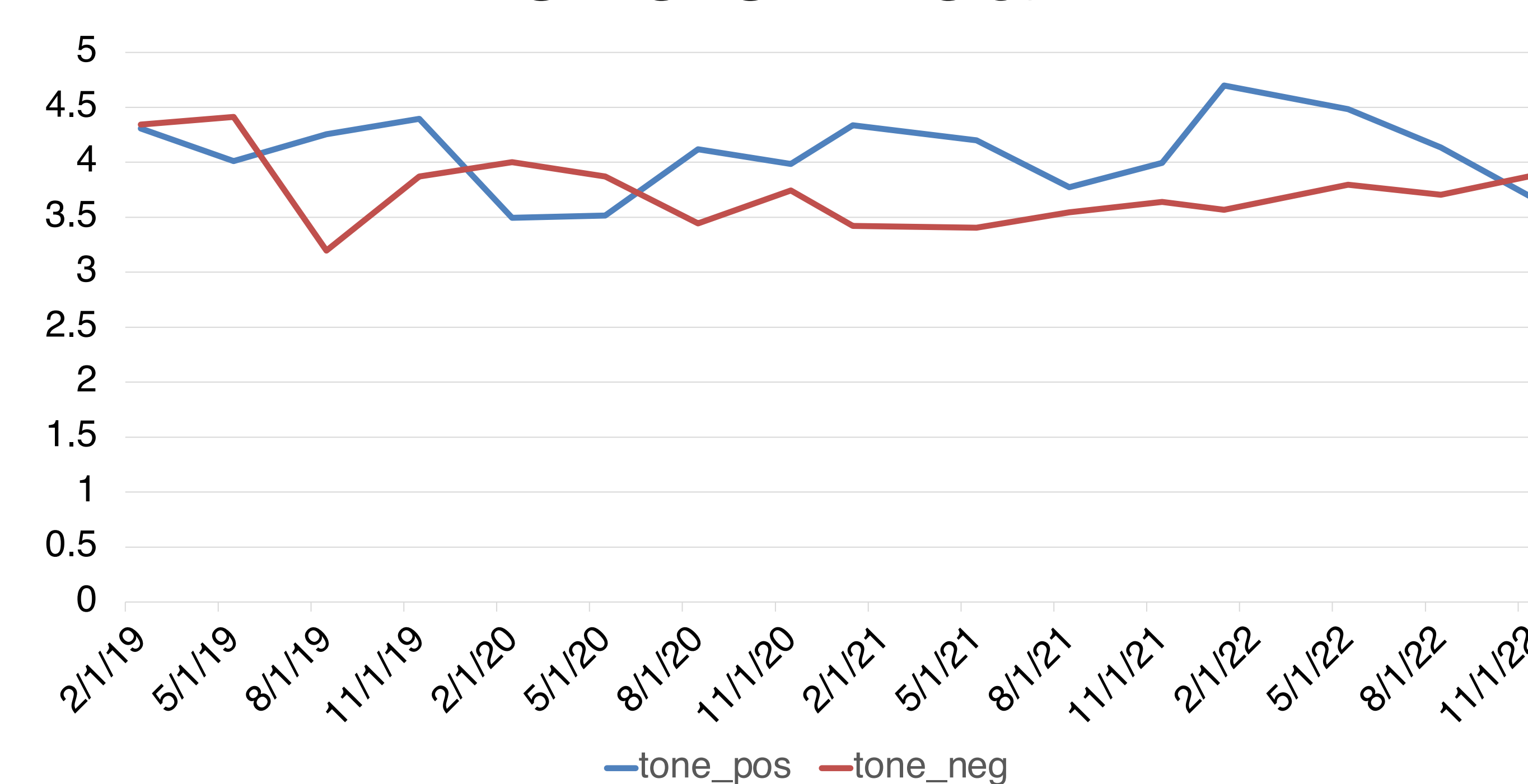
- The lowest positive tone score (3.495) occurred nearest the onset of the pandemic on February 29th, 2020.
- Positive tone remained low in the early months of the pandemic compared to pre-pandemic scores until 8/2/2020 when it recovered to 4.12, which we suspect may be due to the simultaneous announcement of the US purchasing 100 million doses of the Moderna vaccine.
- The difference in the average positive tone in the pre-pandemic period (2/4/2019 to 2/28/2020) was statistically significant (p<.001) in comparison to the post-pandemic period (2/29/2020-11/2/2020) with average positive tones of 4.24 and 3.78, respectively.
- A decrease in positive tone regarding health was expected because pandemic shutdowns resulted in isolation and hardship for many in the context of public health and safety. Future directions of study should use social media to investigate more specific health topics using new keywords including 'vaccines' and 'COVID-19'.