Bias Over Time
Captain-Callen Shutts, NetID: cshutts2

My project is a free project trying to understand how much we influence politicians versus politicians influencing us. For instance, vaccine hesitancy was roughly equal before the pandemic with a slight Republican lean. However, currently, vaccine hesitancy is strongly correlated with political ideology. What changed the first public opinion on vaccines or the official party line? I am more interested in more niche topics like Trump's tariff policy or views on the Chinese economy. While polling can answer these questions polling typically only begins once a topic is of a certain level of importance.

Practically this means I will attempt to identify a large number of users active on social media who are clearly a republican or a democrat. Then attempt to track the sentiment of the large group on a topic over time. I will attempt to identify inflection points in which their party takes a view on a previously obscure topic or switches its view on a topic. Then we can understand did the party officials drive this change or did their party members' views change so they had to change their policy. This will be applied to between five and ten topics.

I plan to do this work in Python utilizing a series of web scraping tools. Currently, I plan to scrape Meta’s Facebook, but this may change. To identify clearly republican or democratic users I plan to simply identify users that follow a lot of political accounts and make pollical posts on the topic of choice using simple TF-IDF searches. Once I have identified users for a topic I will try a series of sentiment analysis tools, such as BERT, VADER, and AFINN lexicon analysis. I will track what percent of each group makes posts on the topic and their overall sentiment. The outcome will be sentiment score for each party over time. I will then provide overlaying annotations with current events to help understand which group is influencing this behavior, the party or the people.

This work will be evaluated by studying the relationship between the party and its constituent’s sentiment. For some topics, it is assumed the party includes the people and vice versa. However, if the sentiment moves widely and randomly the data processing will likely be flawed. One would also assume that data aligns loosely with point-in-time polling.

Plan
· Identifying members of each party on a social media platform ~5hr
· Building sentiment analysis pipeline ~5hr
· Applying to a certain topic ~1-2hr X 5-10
Total ~20 hours