**What rhetoric are United States Senators on Twitter when talking about others?**

United States Senators are using Twitter in different ways. Some senators use Twitter as a platform to communicate upcoming events and policy updates. Other senators use Twitter to criticize or praise others. The dashboard below demonstrates how each senator is using Twitter individually as well as by region and party affiliation.

**Methodological Approach:**

The goal of the dashboard is to address the question, “What rhetoric are United States using when talking about others on Twitter?” First, the Twitter API was used to download senators’ Tweets. Next, a portion of the Tweets were categorized manually, and finally, Naïve Bayes was conducted to predict the category of the remaining Tweets.

**Data Collection:**

The first step in downloading senators’ Tweets was creating a list of all 100 senators’ Twitter handles. Afterward, the maximum allowable requests through the Twitter API were utilized to download as many Tweets per senator as possible. The result is approximately 115 Tweets per senator, which was scraped on April 28th, 2022. The final dataset includes 11,287 Tweets from all 100 senators. While 11,500 Tweets were expected, some senators did not have 115 Tweets, and Bernie Sander’s Twitter handle recently changed so only one Tweet was included.

In addition, not all senators’ names were in the same format. For example, some senators chose Twitter names like “Sen. Susan Collins,” or “Senator Bob Menendez,” but others chose solely their name. All names were updated to reflect solely the senator’s first and last name.

**Methods of Analysis:**

Categorizing senatorial Tweets into three distinct categories was a difficult endeavor. For simplicity, the categories chosen are only reflective of their rhetoric about another individual or group.

Below is an example of how the Tweets were categorized manually:

“Climate change is bad” – Neutral

“Climate change is good” – Neutral

“The Biden Administration’s handling of climate change is bad.” – Critical

“The Biden Administration’s handling of climate change is good.” – Praise

In addition to adding a “category” column, “region” and “party” columns were also added manually for dashboard filtering.

In order to use Naïve Bayes to predict the majority of Tweets, approximately 20 Tweets per senator were manually labeled as “critical,” “neutral,” or “praise. Once these Tweets were appropriately labeled, Naïve Bayes was conducted to predict the rest. In the end, the predictions were added to a dataframe with the “name,” “region,” and “party” variables. This dataset is used to display the results in a bar plot in the above dashboard.

**Methodology Evaluation:**

Once over 15% of the Tweets were manually categorized, the Naïve Bayes model performed well in predicting the category. Comparing the actual (manually-entered) category with the predicted category shows that majority of the Tweets were predicted correctly.

By proxy, manually-entered categories are subject to subconscious bias. The three categories were specifically chosen to maintain a bipartisan point of view. The Tweet labels “critical,” “neutral,” and “praise” fit objectively in those categories.

In addition, the results are not from a single timeframe and were scraped statically. Some senators Tweet much more often than others – and as a result, individual senators’ Tweets should be assessed carefully.

**Suggestions for Improvement:**

In the future, this dashboard could be improved by:

* Expanding the Twitter developer API access to allow more Tweets to be downloaded per senator
* Scraping the Tweets dynamically based on the date the dashboard is accessed
* Fixing Bernie Sander’s Twitter handle to include his Tweets
* Adding a date range filter to look at Tweets during a specific day, week, or month
* Adding an additional label for current events so rhetoric around a specific issue could be assessed