

Project Proposal:

Determine the level of support of Republicans and Democrats to their Candidate by each State from Sentiment Analysis Results

Aim:

The aim of this project is to determine the level of support for the Republican and Democrat candidates in each state by performing a sentiment analysis of the data collected from twitter. The project will use natural language processing techniques and sentiment analysis algorithms to determine the overall sentiment expressed by each party about their candidate in each state. Then based of those values build a contingency table and use chi-square method for analysis of significance of level of support.

Methodology:

- 1. Data Collection:** Collect data on the Republican and Democrat parties from twitter. Data will be collected from various sources such as social media, news articles, and websites using web scraping techniques.
- 2. Sentiment Analysis:** The collected data will be pre-processed and then subjected to sentiment analysis using natural language processing techniques and sentiment analysis algorithms. Perform sentiment analysis on the data collected to determine the overall sentiment expressed by each party about their candidate in each state.
- 3. Contingency Table:** A contingency table will be built with sentiment scores as the independent variable and party affiliation as the dependent variable in each state. Build a contingency table with sentiment scores as the independent variable and party affiliation as the dependent variable in each state.
- 4. Chi-square Test:** A chi-square test will be performed in each state to compare the observed distribution of sentiment to a hypothetical distribution where there is the same level of support for both candidates. Perform a chi-square test in each state to compare the observed distribution of sentiment to a hypothetical distribution where there is the same level of support for both candidates. Compare the results of the chi-square tests across states to get a broader understanding of the level of support for each party and candidate at a national level.
- 5. Data Analysis and Interpretation:** The results of the chi-square tests will be compared across states to get a broader understanding of the level of support for each party and candidate at a national level.

Outcome:

The outcome of this project is expected to provide valuable insights into the level of support for the Republican and Democrat candidates in each state and as a nation whole. The results of the project can be useful for making predictions and decisions related to elections and other political events.