Sentiment Analysis of COVID-19 Omicron Tweets

Project Overview

This project is an analysis of tweets related to the COVID-19 Omicron variant to understand public sentiment over time. The objective is to classify tweets into Positive, Negative, or Neutral sentiment categories. The analysis uses the VADER (Valence Aware Dictionary and sEntiment Reasoner) sentiment analysis tool, a rule-based method designed for social media text. The results are visualized through monthly sentiment distribution pie charts to gain insight into public opinion trends.

Technique Used

This project uses VADER for sentiment analysis:

- VADER (Valence Aware Dictionary and sEntiment Reasoner)
  - Description: VADER is a lexicon and rule-based sentiment analysis tool that is particularly effective on social media text. It is fast, interpretable, and handles slang, emoticons, and punctuation cues common in tweets.
  - Advantages: Quick to run, effective for social media language, and requires minimal computational resources.
  - Limitations: Lacks deep contextual understanding, so it may miss subtleties in complex or nuanced language.

Dataset

The dataset consists of hydrated tweets across seven CSV files, each corresponding to a specific month. The files contain tweet text labeled with the following months:

- November
- December
- January
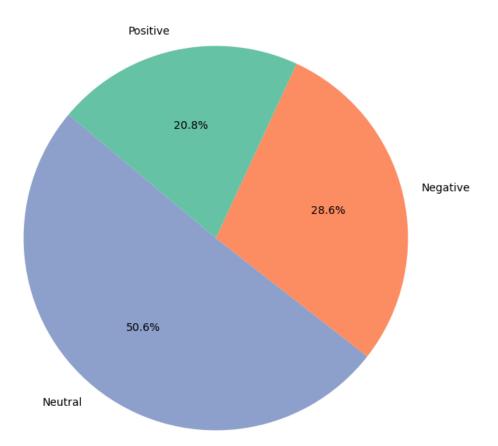- February
- March
- April
- May

Dependencies

The project requires the following Python libraries:

- pandas – for data manipulation.
- glob – for file handling.
- os – for file path operations.
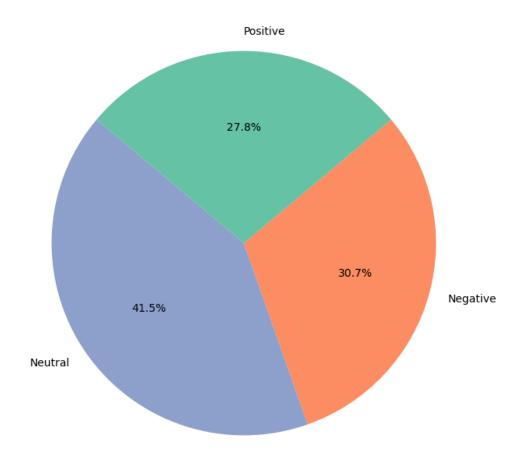- matplotlib – for data visualization (pie charts).

- vaderSentiment - for VADER-based sentiment analysis.
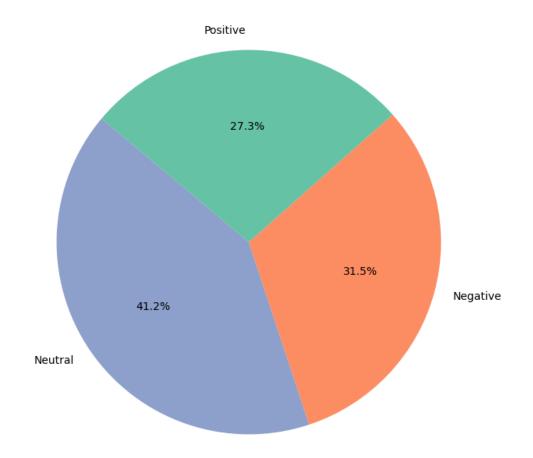
OPUPUT:
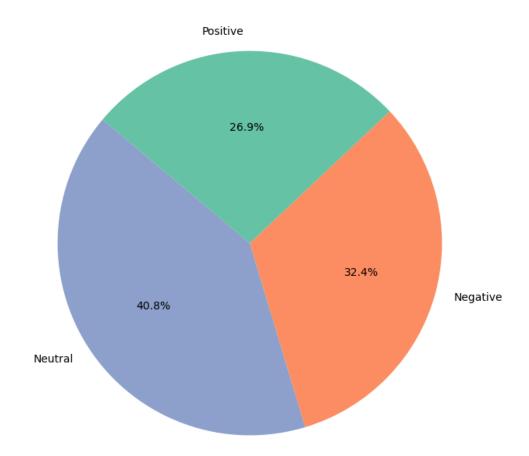
Sentiment Distribution for November
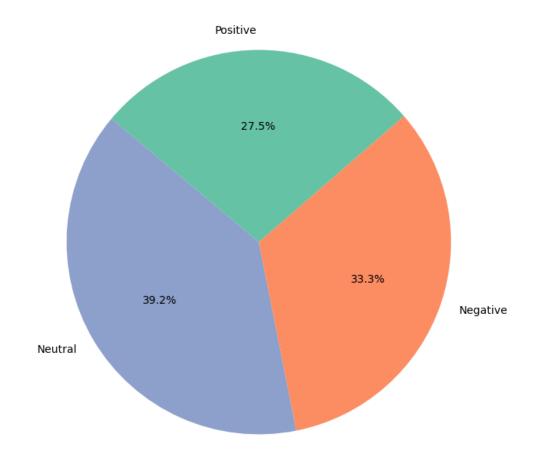
Sentiment Distribution for April
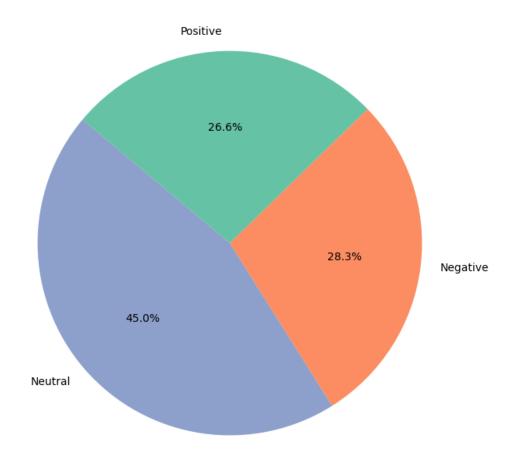
Sentiment Distribution for December

Sentiment Distribution for January

Sentiment Distribution for February

Sentiment Distribution for May

Sentiment Distribution for March