

Sentiment Analysis of Twitter/X Posts on a Specific Topic

Name: Ho Phuoc Lanh

ID: 2102114

1. Objective

- Collect posts (tweets) from Twitter/X related to a trending topic (e.g., "Taylor Swift").
- Analyze the sentiment of these tweets, classifying them into Positive, Negative, or Neutral categories.
- Visualize the data using charts to display the sentiment distribution clearly.
- Tell a story through the data visualization to reveal public opinion trends toward the chosen topic.

2. Reason for Choosing This Topic

I chose this topic because sentiment analysis is a highly relevant and practical application of data visualization and data science. Social media platforms like Twitter/X reflect real-time public opinion, making them a rich source of insights for various fields such as marketing and social studies.

3. Scope

- Topic: The final topic will be confirmed depending on the availability of recent data.
- Data volume: at least **3,000 tweets** collected.

4. Methodology

Step 1: Data Collection

- Use snsrape (Python library).
- Fields to collect:
 - Timestamp
 - Tweet content (text)
 - Engagement metrics (likes, retweets) if available.

Step 2: Data Cleaning

- Remove:
 - Duplicate tweets.
 - Tweets without meaningful content (spam, promotional links, etc.).
- Standardize the text:
 - Lowercasing
 - Removing excessive emojis or irrelevant characters (if necessary).

Step 3: Sentiment Analysis

- Apply a pre-trained sentiment analysis model:
 - VADER (Valence Aware Dictionary and sEntiment Reasoner) — suitable for social media text.
- Each tweet will be labeled as:
 - Positive
 - Negative
 - Neutral

Step 4: Data Visualization

- Create multiple types of charts

Step 5: Conclusion (Storytelling)

- Summarize key findings