Eras Tour – Taylor Swift Twitter Analytics

Author: Anahita Darban
Course: COMP 651 – Social Media Analytics
Instructor: Professor Manan Patel
Date: 2025-08-11

Introduction

This project is an examination of Twitter traffic on Taylor Swift based on the Eras Tour data. It aimed to investigate sentiment, engagement rates, and frequency per topicality of time by a mix of preprocessing of the records through Python, and visual analytics in Tableau.

The data retrieved (TaylorSwift13.csv and processed derivatives) was obtained in Kaggle, where the most significant addition of value was the usage of sentiment analysis and topic modeling to comprehend online conversation dynamics better. The described analysis can facilitate the establishment of engagement patterns of fans, the times when fans exhibit the highest engagement rates, and the most popular discussion topics.

Data Description

The analysis was conducted upon three processed CSV files namely daily_trends.csv, which included the aggregated tweet volumes and sentiment per day; posts_enriched.csv that contained the individual tweets with the engagement data containing the likes, retweet, and replies, as well as the sentiment scores and detected topics; and topics_daily.csv that provided the daily mentions of the topics to be plotted on specific keywords. The auto_report.md indicates that 756 posts were examined, and sentiment was divided into 419 positive, 260 neutral and 77 negative. The analysis has determined that there are 6 modelled topics, where the busiest day was on June 3, 2021, and the 7-day forecast of tweet volume predicts 0.0 on the next day.

Data Preparation

The data cleaning procedure divided into special characters removal, text normalization, and rewriting to lower case all the names of the variables. Rule-based VADER sentiment-based model was then utilized to give a sentiment score and label. In the case of topic modeling, the tweets were clustered into six discussion topics, and the aggregation processes were implemented to get daily summaries of the tweet data reflecting the sentiment, engagement, and frequencies of topics.

Visualizations in Tableau

The preprocessed and analyzed data was exported and then imported into Tableau Public as CSV files (daily_trends.csv, posts_enriched.csv and topics_daily.csv). This data was then utilized to create an interactive dashboard which consists of 6 major visualizations: Sentiment & Tweets

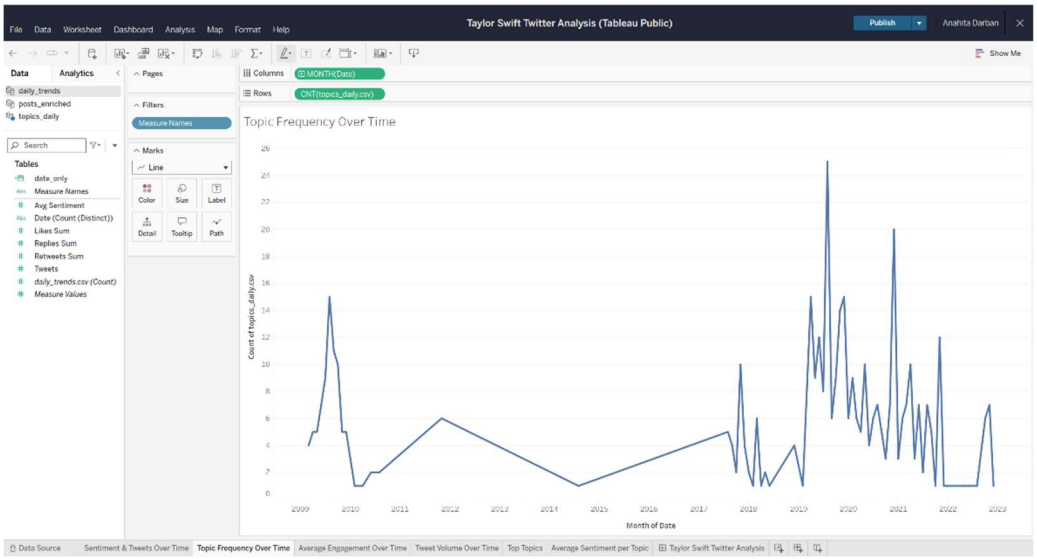
Over Time, Topic Frequency Over Time, Tweet Volume Over Time, Top Topics, Average Sentiment per Topic, and Average Engagement Over Time. All the visualizations were developed to reveal certain analytical findings of the data that has passed through the processing process so that the sentiment trends, topic changes, engagement patterns, and the event-driven sources of the fan activity surges could be clearly seen. In appendix, the mandatory Tableau dashboard connection is available to interact.

The following are the principal visualizations generated and how they were interpreted.

- Monthly Sentiment & Tweet Volume

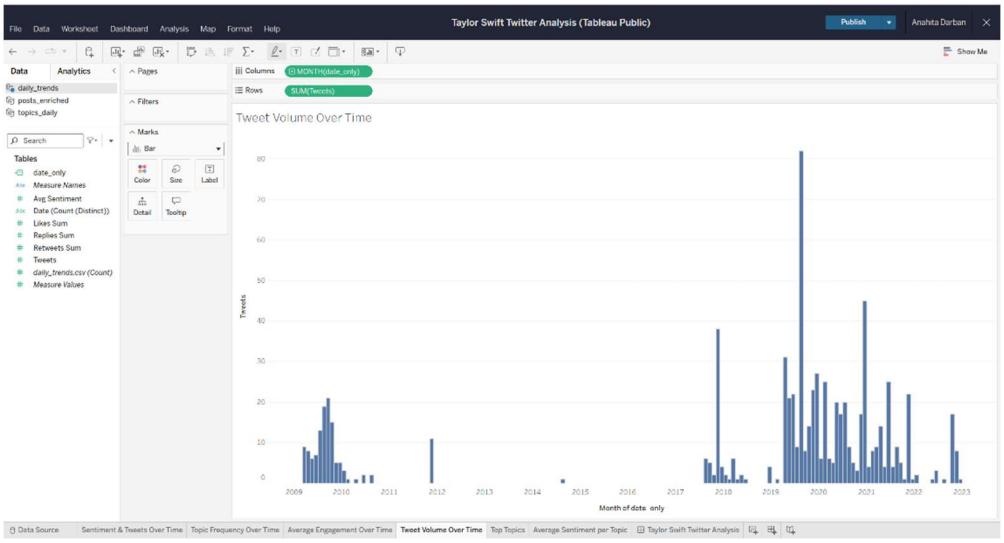
This diagram indicates the change in sentiment according to the number of tweets over the course of time. Most of the months are dominated by positive sentiments, although over the top they have occurred during big events.

- Monthly Topic Mentions



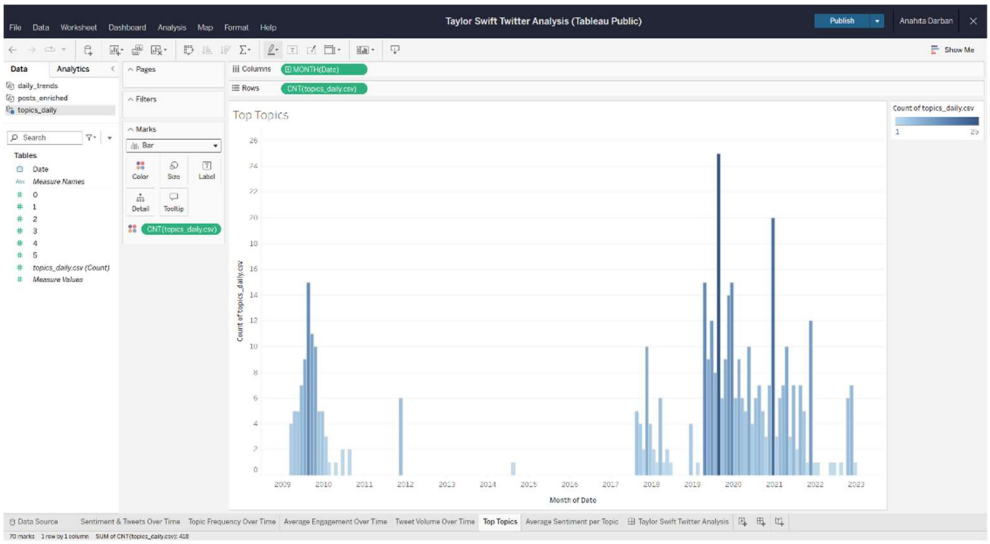
Shows how frequent each topic category is in each month. Certain issues were seasonal in nature and scheduled with Eras Tour announcements often in mind.

- Tweet Volume Over Time



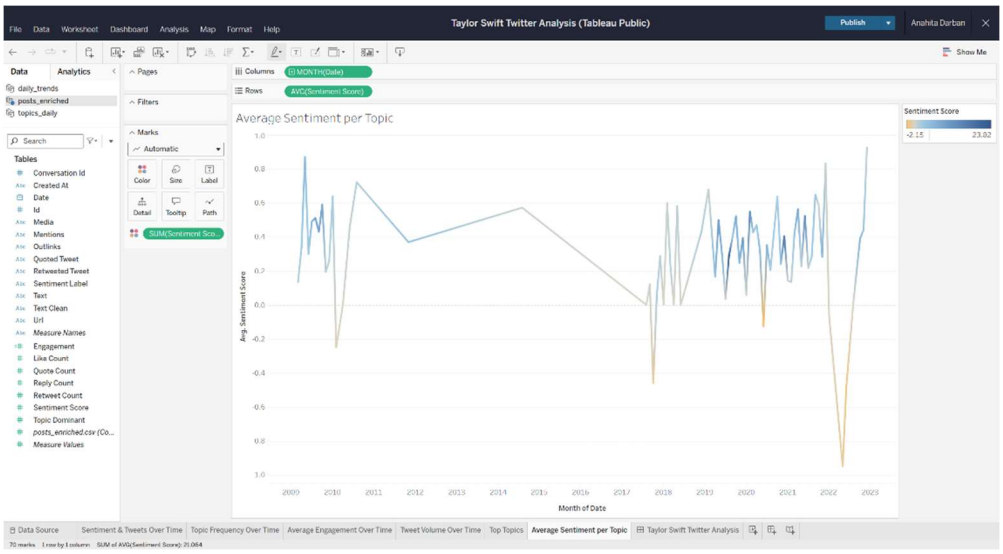
It underlines the number of tweets made each day and it is not hard to note the high activity days like June 3, 2021.

- Top Topics



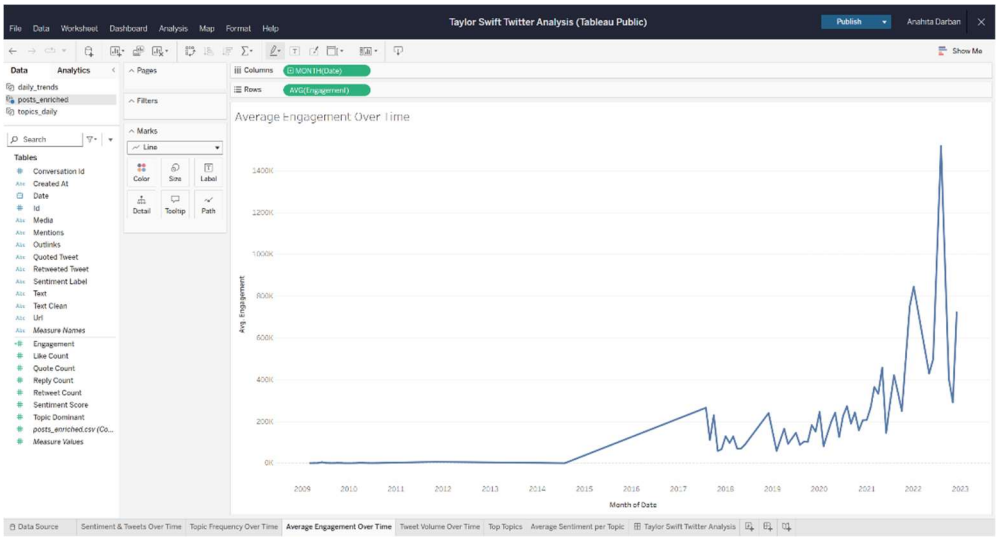
Rank of the most common topics on bar chart. The first 3 held a majority of the conversation.

- Average Sentiment per Topic



Shows what subjects were most positively or negatively toned by the fans.

- Average Engagement Over Time



The like and retweet trends, and reply trends depict that likes and tweets and their positive sentiment usually matched with increased engagement.

Key Insights

The conclusions showed that the good feelings prevailed since the percentage of positive tweets was 55 percent and it is linked to high enthusiasm of the fans. The activity is peaked on June 3, 2021, which can be related to some weighting action or news. There were six key discussion topics identified through analysis of topic trends with some producing a higher level of positivity than others with engagement patterns showing that positive tweets enjoyed a higher level of engagement most of the time. This implies that inspirational or encouraging materials are more likely to strike a chord with the reader and thus increase their popularity and contribution. In general, the results show the very close relationship between the topic relevance, event-driven spikes, and the audience engagement patterns.

Conclusion

Using Python-based preprocessing and Tableau visualization allowed extracting actionable knowledge about the way of the fan and the perception of it. Potential future enhancements would have involved the addition of real-time streaming to enable the live dashboard operation, the coverage of other platforms where analysis may be performed, including Instagram and TikTok, and the use of a more advanced NLP model, BERT, to provide better and more detailed sentiment classification. Such improvements would make it possible to analyze information timelier, in its entirety, and more accurately, which would ultimately better inform the audience dynamics and trends.

Appendix

Tableau Public Link:
https://public.tableau.com/views/TaylorSwiftTwitterAnalysis/TaylorSwiftTwitterAnalysis?:language=en-US&:sid=&:redirect=auth&:display_count=n&:origin=viz_share_link

Dataset Source: https://www.kaggle.com/datasets/thedevastator/taylor-swift-s-social-media-performance-tweet-en

Files Used: daily_trends.csv, posts_enriched.csv, topics_daily.csv