DAB 303 Marketing Analytics

Natural Language Processing

Prepared By:
Group 10

Sumbitted to:
Dr. Andreas S. Maniatis

Table Of Contents

List of materials that we will discuss about the project.

Introduction and Purpose

Welcome to a journey through the intricate world of social media sentiments! In this marketing project, we delve into the realms of exploratory data analysis (EDA), data visualization, and the fascinating principles of Natural Language Processing (NLP).

Our first pitstop involves unraveling insights from a rich social media dataset. Through meticulous EDA, we aim to decode patterns, trends, and anomalies within the data. By employing various statistical and graphical methods, we uncover the stories hidden within the numbers, shedding light on the pulse of the online landscape.

Data and Mission

The dataset comprises Twitter data reflecting customer sentiments towards an airline

Data

The dataset comprises Twitter data with unique tweet IDs, revealing sentiments (positive, negative, neutral) towards an airline. It includes sentiment confidence, airline details, reasons for negativity, user information, tweet content, coordinates, timestamp, location, and timezone.

Mission

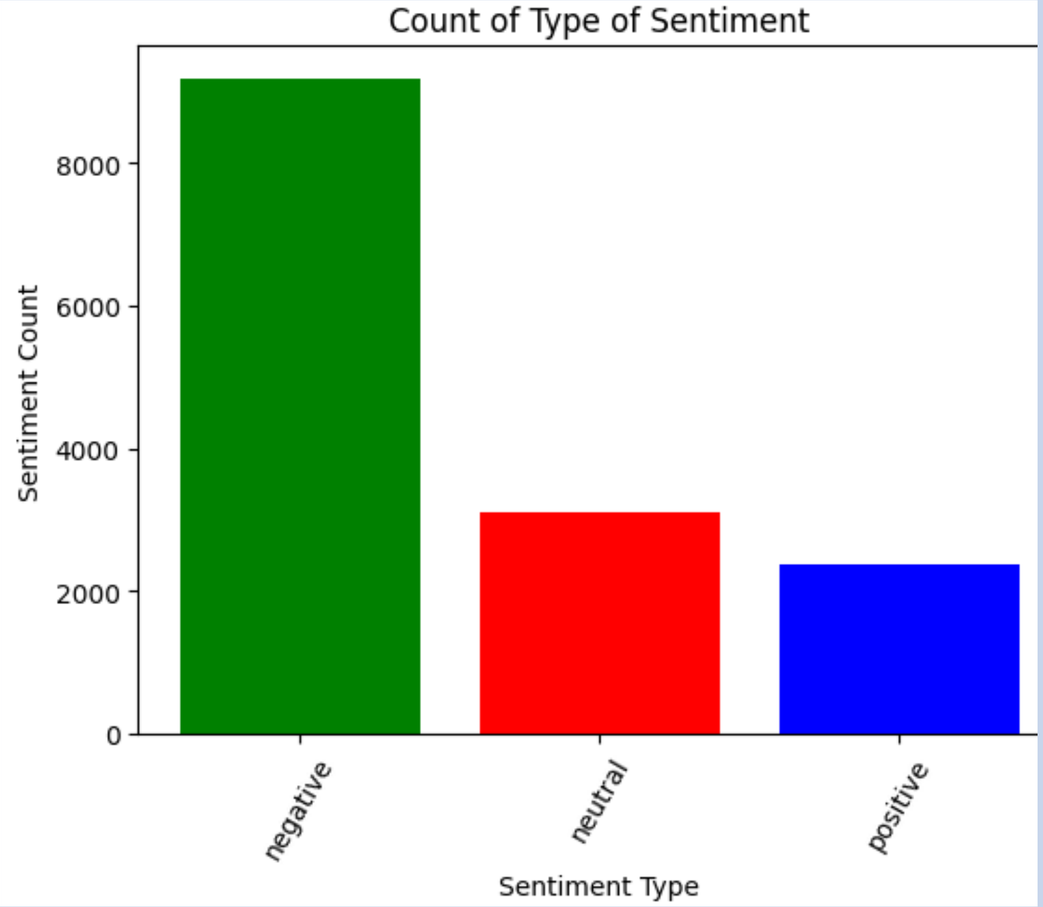
Our mission is to extract valuable insights from this data using EDA and NLP. We aim to understand sentiment distribution, identify key negative reasons, evaluate sentiment confidence, uncover geographic trends, and leverage NLP for deeper textual insights. The goal is to inform strategic decisions, enhance customer satisfaction, and strengthen the airline-passenger relationship.

EDA

Next, we will discuss the analytical insights we identified from the data

EDA 1

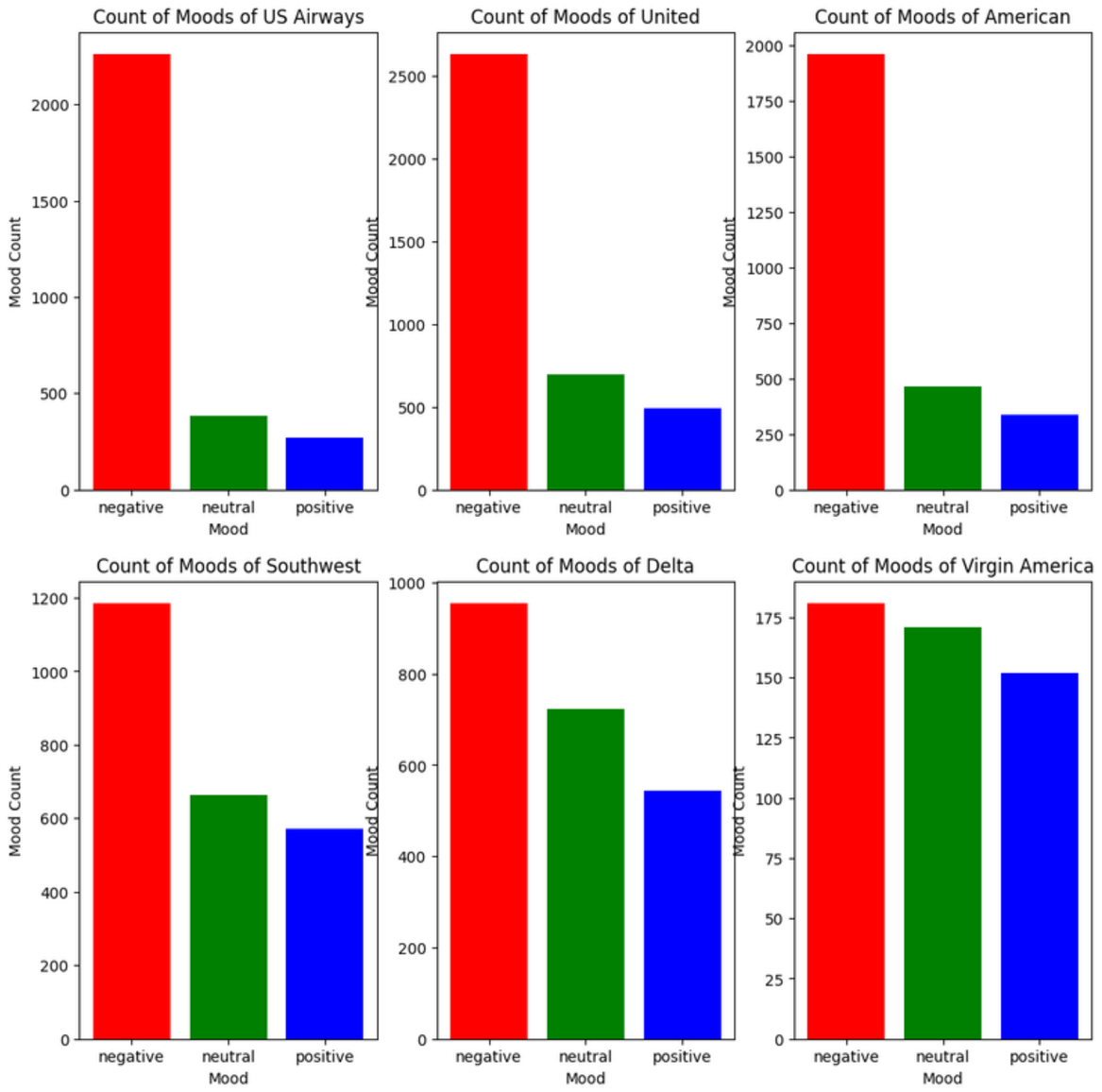
Sentiment Tally



- Sentiment Distribution Overview:
  - the prevalence of negative sentiments, followed by neutral and positive sentiments.
- Color Coding for Clarity:
  - Green for negative, red for neutral, and blue for positive sentiments,
- Actionable Takeaway from the Visual:
  - Emphasizing the need to address negative sentiments for improved customer satisfaction.
- Future Strategic Considerations:
  - Utilize visual distribution insights to inform future strategies, focusing on resolving issues expressed in negative sentiments and maintaining positive sentiment trends.

EDA 2

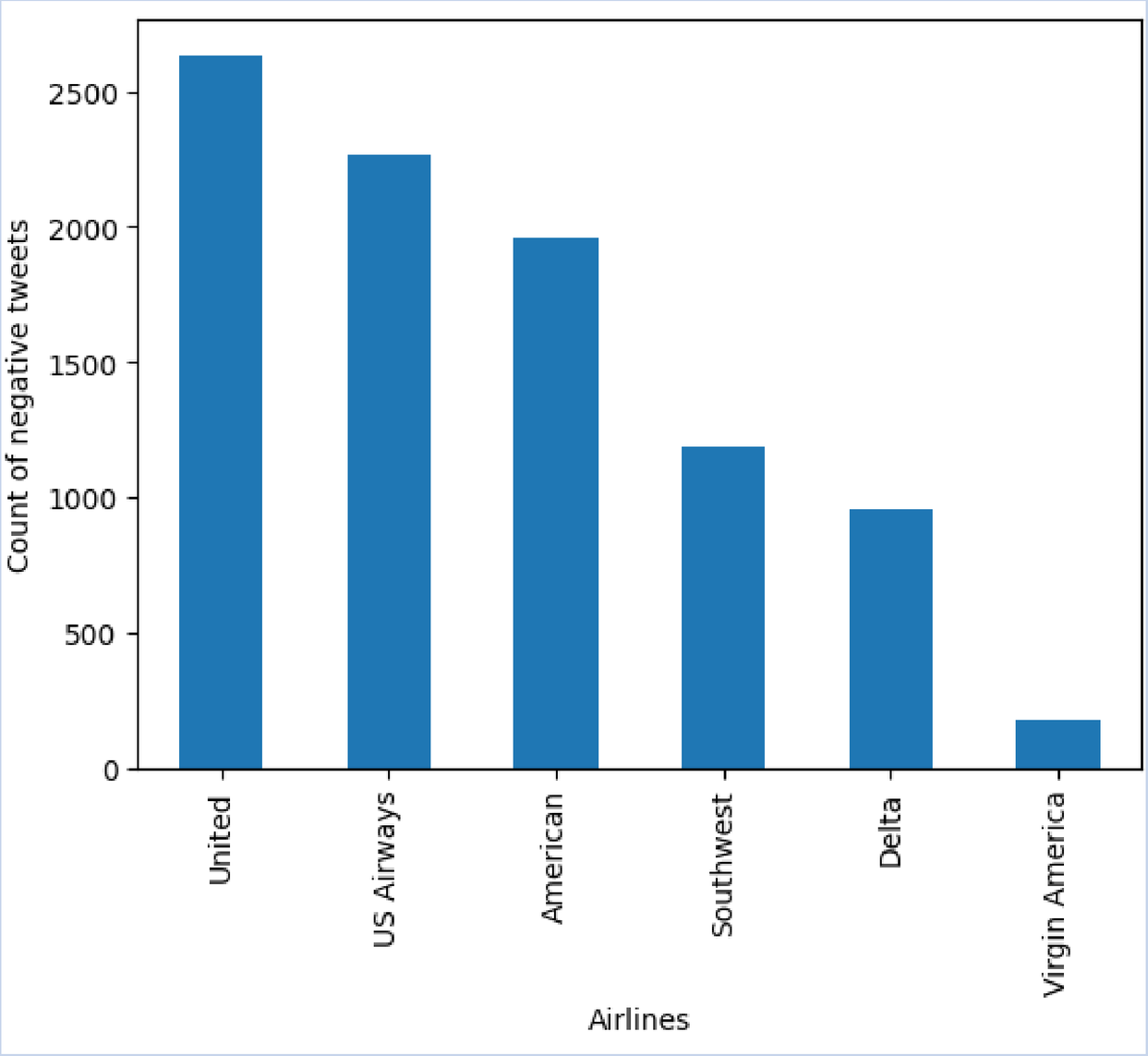
Sentiment Tally



Airline-Specific Sentiment Breakdown:

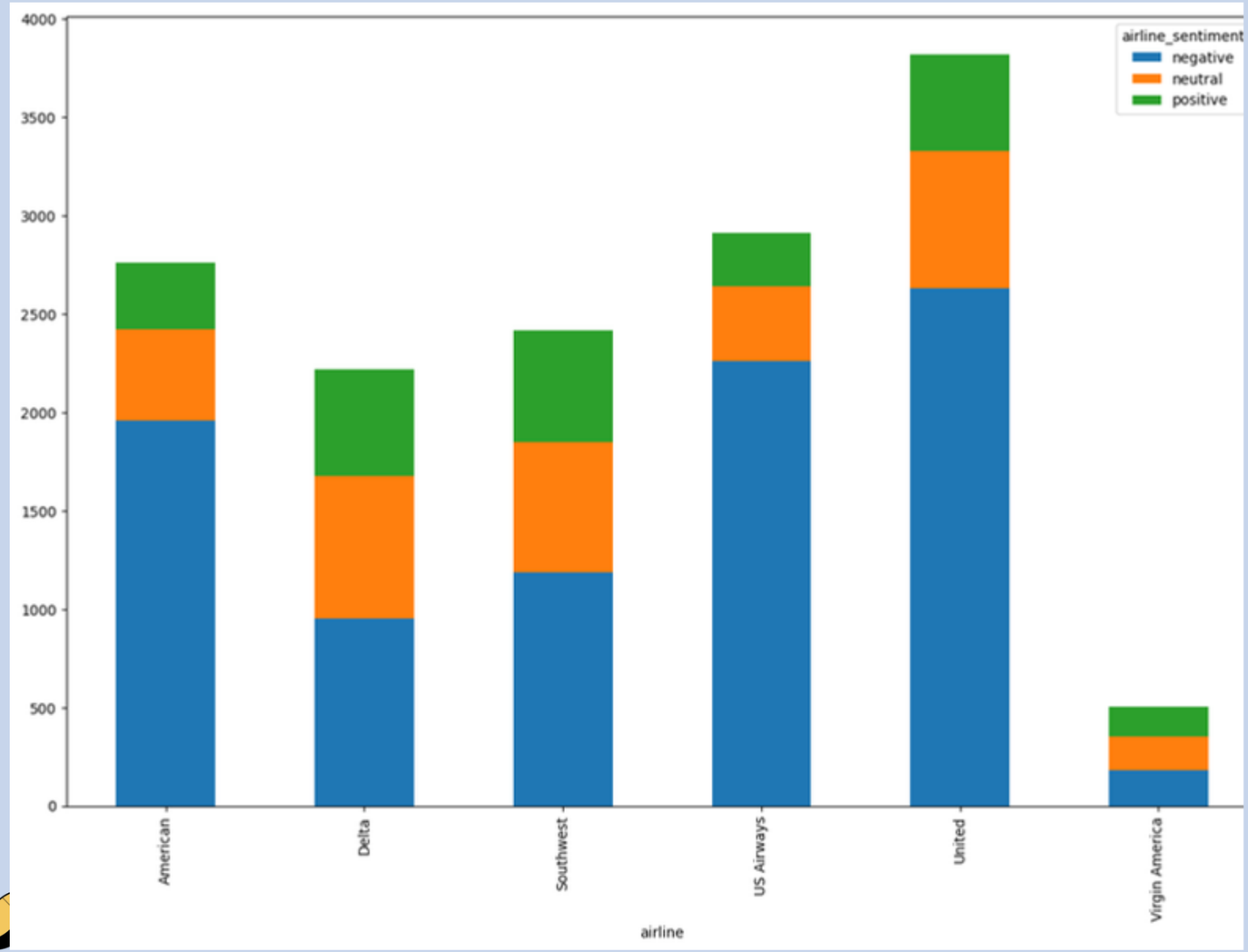
- Individual charts provide a detailed breakdown of sentiments for each airline, offering a nuanced understanding of customer perceptions.
- Comparative Analysis Across Airlines:
  - Side-by-side comparisons enable stakeholders to assess how each airline fares in terms of sentiment distribution, identifying areas of strength and improvement.
- Insights for Tailored Strategies:
  - By examining mood counts for each airline, tailored strategies can be formulated to address specific sentiment trends and enhance overall customer satisfaction.

EDA 3



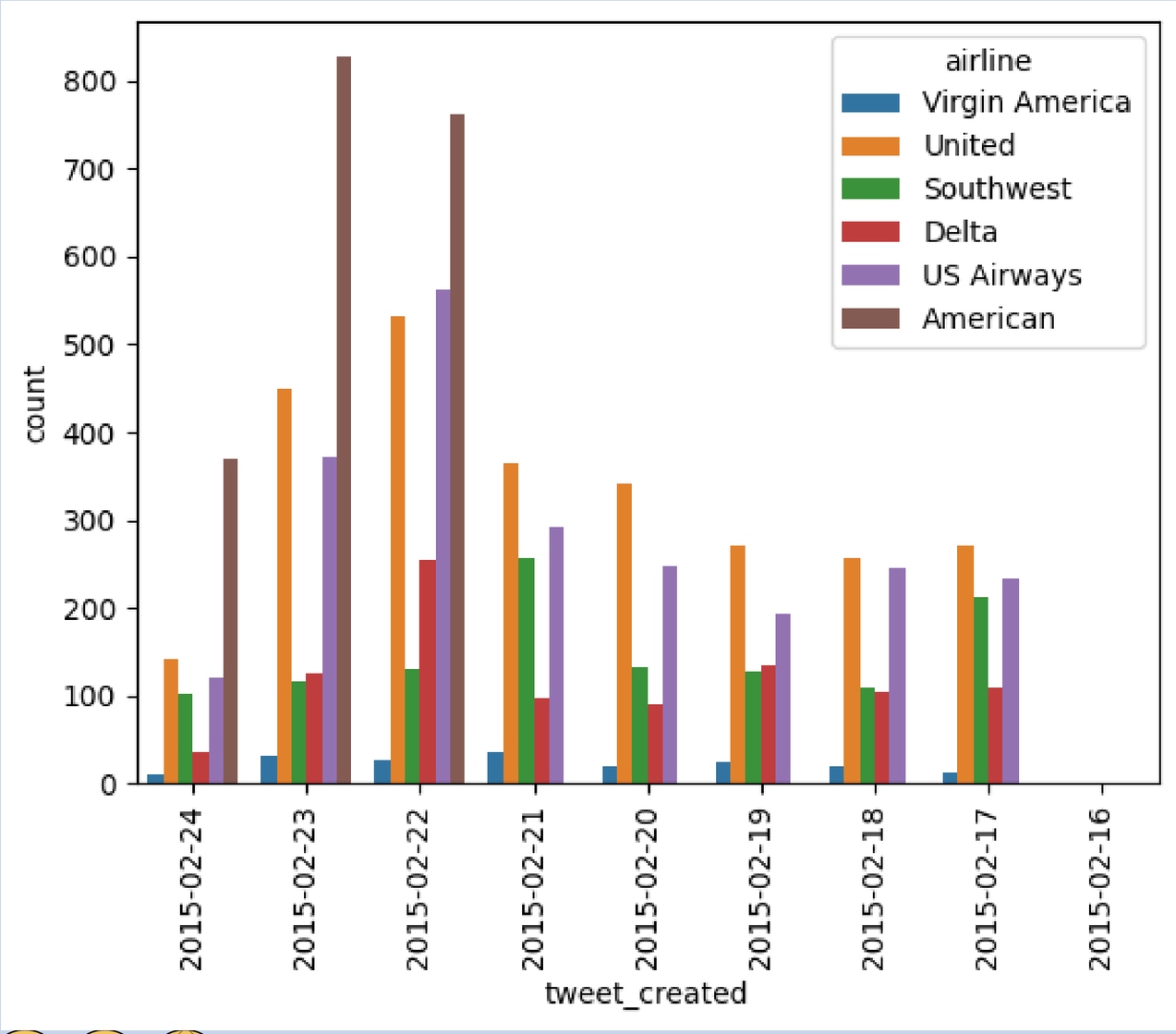
- Airline-Specific Negative Sentiment Analysis:
  - Visual representation showcasing the distribution of negative sentiments across different airlines.
- Identifying Airlines with Higher Negative Sentiments:
  - Quick insights into which airlines receive a higher count of negative tweets, allowing for targeted attention and improvements.
- Informing Customer Service Strategies:
  - Valuable data to inform customer service strategies by addressing issues highlighted in negative tweets and improving overall customer experiences.

EDA 4



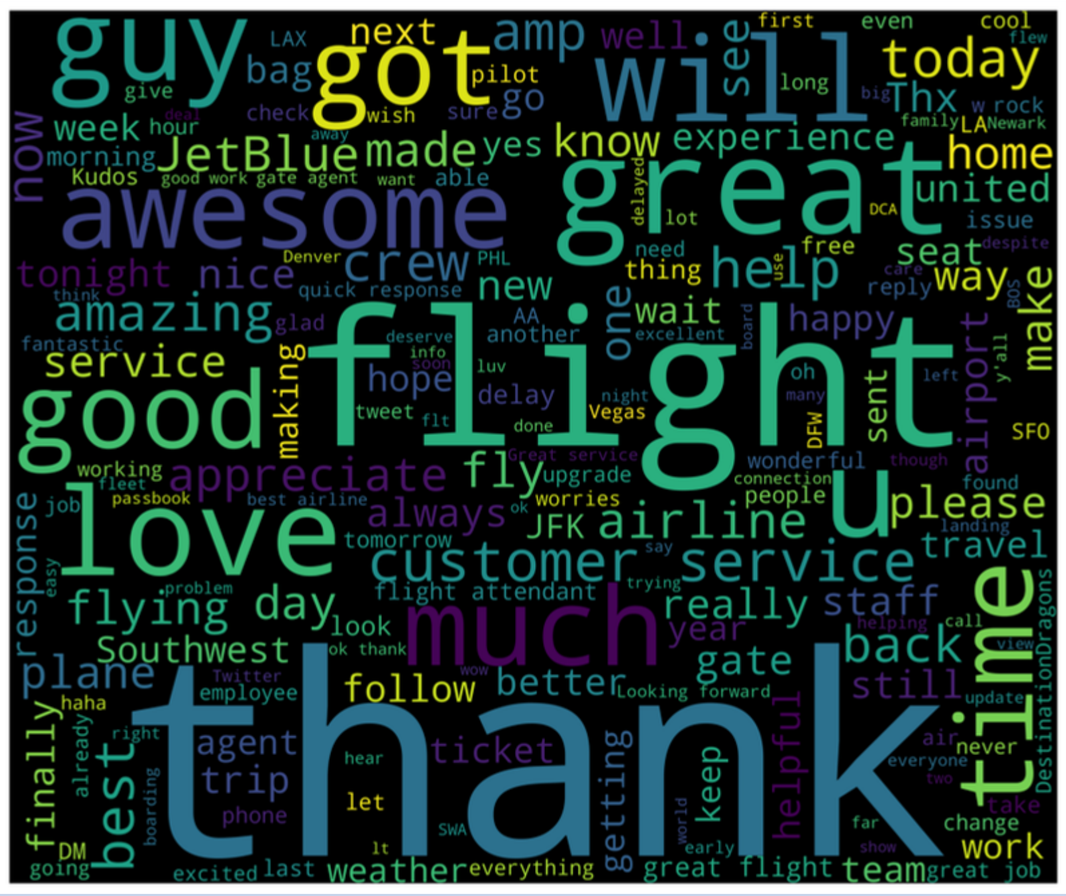
- Visualizing Sentiment Proportions:
  - Clearly depicts the proportion of each sentiment category for every airline, aiding in quick comparative analysis.
- Highlighting Sentiment Balance:
  - Effectively illustrates how negative, neutral, and positive sentiments contribute to the overall sentiment balance for each airline.
- Facilitating Cross-Airline Sentiment Comparison:
  - Enables easy comparison of sentiment distributions across multiple airlines, providing insights into relative customer satisfaction levels.

EDA 5



- Temporal Analysis of Negative Tweets:
  - Utilizes a countplot to visualize the distribution of negative tweets over time, offering insights into temporal patterns of dissatisfaction.
- Identifying Spikes or Trends:
  - Facilitates the identification of spikes or consistent trends in negative tweets for each airline, aiding in timely response and strategic interventions.

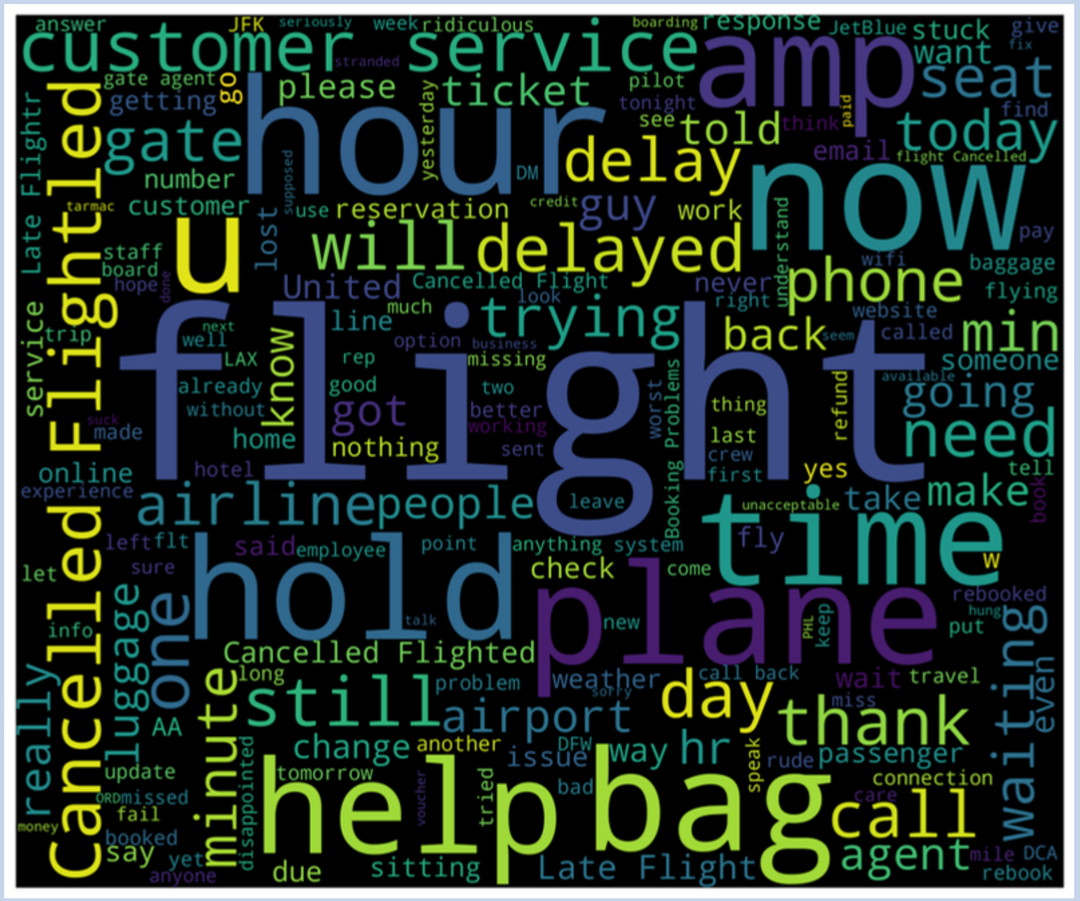
EDA 6



Positive Sentiment Word Cloud:

- Generated using the text from positive sentiment tweets, providing a visually appealing representation of frequently occurring words.

EDA 7



Negative Sentiment Word Cloud:

- Utilizes the text from negative sentiment tweets to create a visually striking word cloud, offering insights into frequently used words in negative contexts.

Conclusion

1. **Optimal Model Selection:**
   - The Logistic Regression model, coupled with CountVectorizer, emerged as the most effective for predicting tweet sentiments.
   - Achieving an accuracy of 78.921%, it stands out as our top performer in this sentiment analysis task.
2. **United Airlines Focus:**
   - A substantial 26% of tweets were directed at United Airlines, shedding light on a significant portion of the conversation.
   - Common complaints centered around issues such as bad service and late flights, pinpointing areas for improvement.
3. **Dominance of Negative Sentiments**:
   - A notable finding is that over 60% of tweets expressed negative emotions.
   - Users predominantly conveyed dissatisfaction or frustration, emphasizing the importance of addressing customer concerns.
4. **.Tweet Length Insight:**
   - An intriguing correlation was observed between tweet length and sentiment.
   - Longer tweets tended to carry more negative sentiments, suggesting a potential trend worth exploring further.

Thank You!

Thank you for your attention during the presentation time.