



# Twitter Airline Sentiment Analysis

By: Ryan King

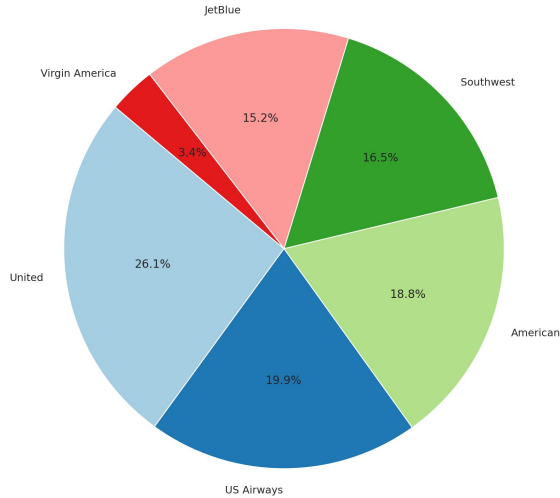


# Data Preparation/Quality

- Imported necessary Python libraries for data manipulation and visualization
  - (e.g., pandas, numpy, plotline, matplotlib, wordcloud)
- Removed unnecessary columns (airline\_sentiment\_gold, negativereason\_gold and tweet\_coord) due to large amounts of null value to reduce data dimensionality
- Updated entries where airline was labeled 'Delta' to be 'JetBlue'
  - Originally, almost all sentiments relating to Delta airline contained 'JetBlue' in their tweet text
- Moderate amount of missing values for the variables (negativereason, negativereason\_confidence, tweet\_location, and user\_timezone) so removed them

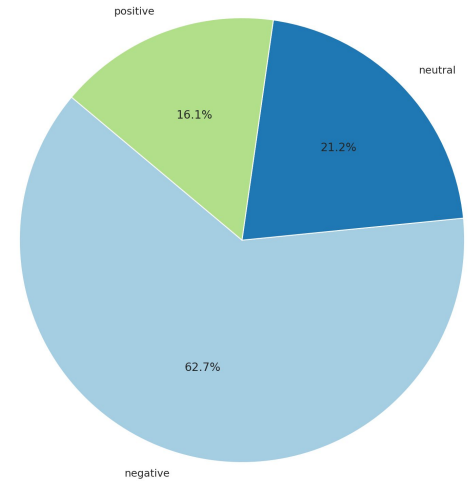
# Overall Sentiment Distribution

Percentage of Tweets per Airline



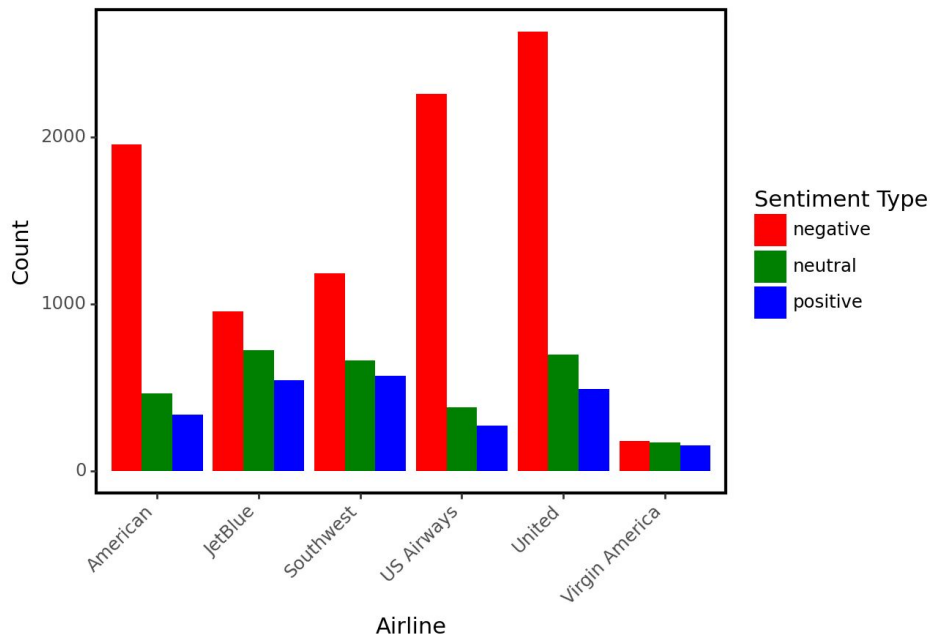
- 62.7% of tweets express negative sentiments, highlighting major customer dissatisfaction and an area needing urgent attention
- Humans often weight negative experiences more heavily, explaining the skewed negative distribution
- United, US Airways and American airlines were the most popular amongst users

Distribution of Sentiments Across All Tweets



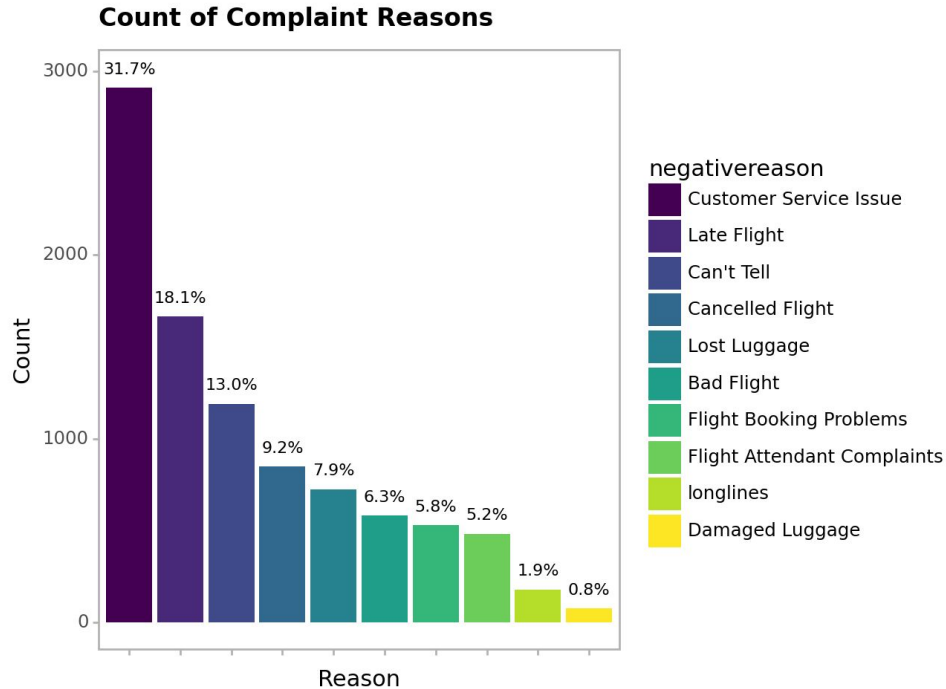
# Airline Sentiment Analysis

Sentiment Distribution Across Airlines



- American, US Airways, and United airlines all received an extremely high amount of negative feedback
  - This might imply potential systemic issues in various factors such as customer service or flight experiences
- Virgin America and, to a lesser extent, JetBlue and Southwest, likely have a relatively positive public perception given their more balanced sentiment distributions
  - This could translate to better customer loyalty and potentially higher repeat business

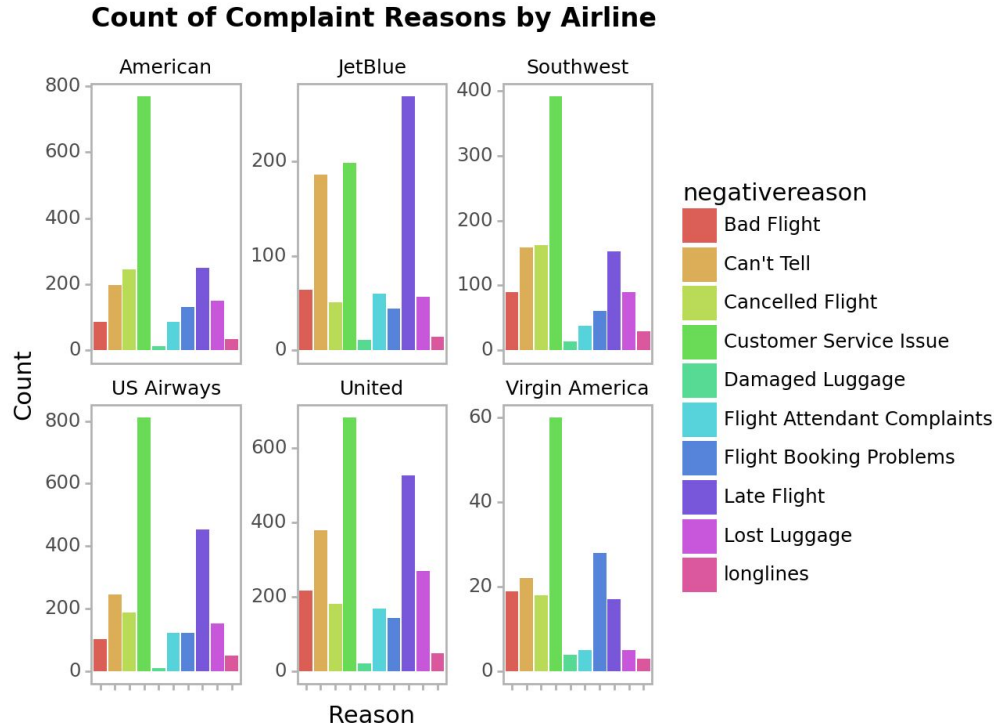
# Reasons for Negative Complaints



- A staggering 31.7% of negative feedback was attributed to customer service issues, highlighting a critical area for airlines to address in order to improve customer relations
- Late and cancelled flights collectively accounted for 27.3% of negative sentiments, indicating potential operational inefficiencies or external challenges affecting timeliness and reliability of airlines



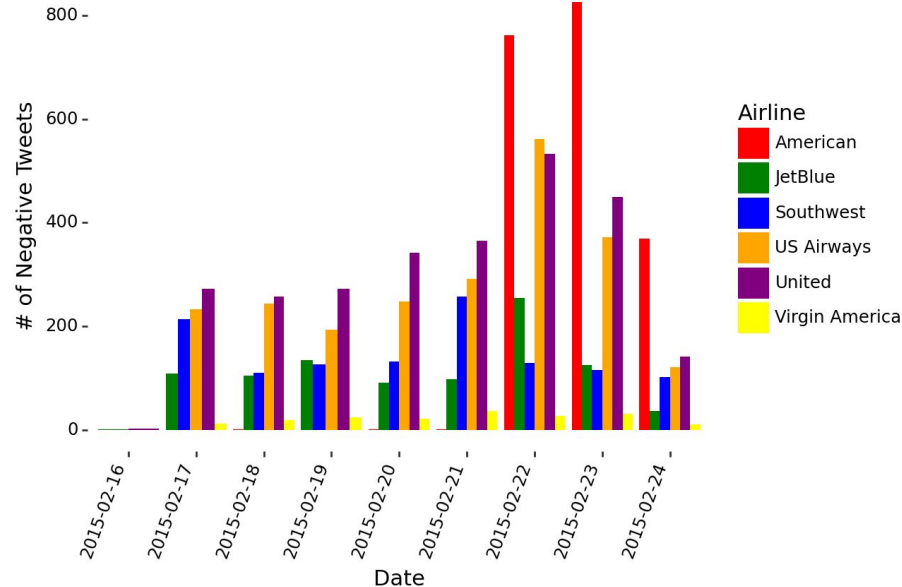
# Negative Sentiment Causes per Airline



- The universal struggle with customer service across all airlines pinpoints the need for a collective industry revamp in customer support strategies
- US Airways and American Airlines led in customer struggles, each nearing 800 negative responses
- The reason all these airlines follow the same trend might be because travel is inherently stressful, making passengers more susceptible to frustration when issues arise

# Negative Sentiment Trend Over Week

Relationship between Negative Sentiments & Date



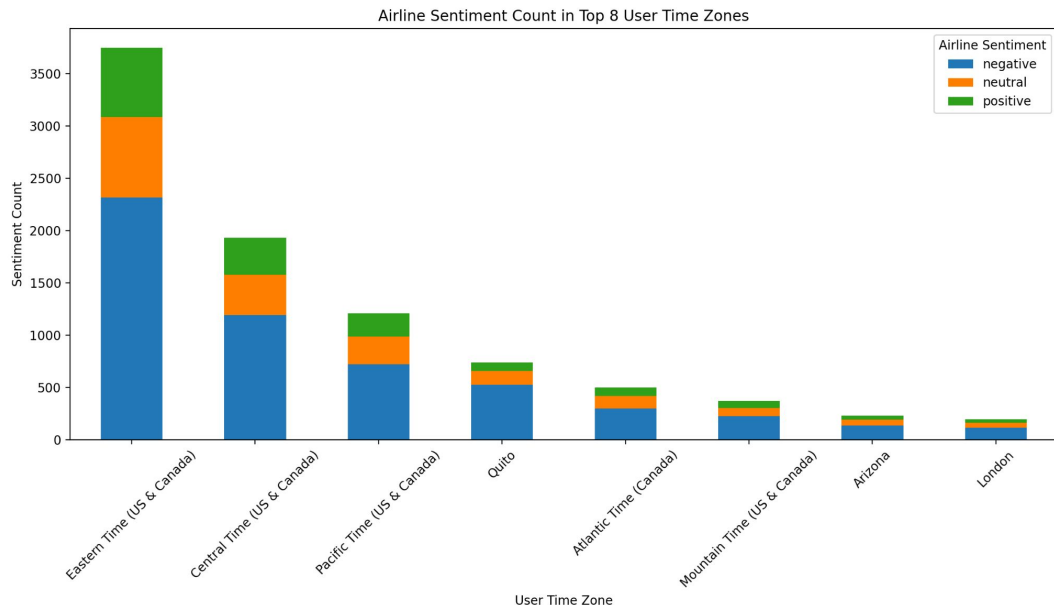
- The surge negative tweets for the airlines on 2/22 and 2/23 suggests an industry-wide challenge such as technical or weather-related disruptions
- The airlines kept their negative complaints relatively low until Sunday the 22nd
- American, US Airways, and United faced the strongest wave of customer dissatisfaction during this period



- **Negative Insights**
  - **Timing Issues:** “Delayed” and “cancelled” suggests prevalent scheduling struggles
  - **Service Challenges:** The appearance of “service” and “help” indicates shortcomings in customer support
- **Positive Insights**
  - **Staff Interactions:** Recognition and approval of staff efforts are visible, with a common use of appreciative words like “thank” and “appreciate,” coupled with compliments directed toward “crew”



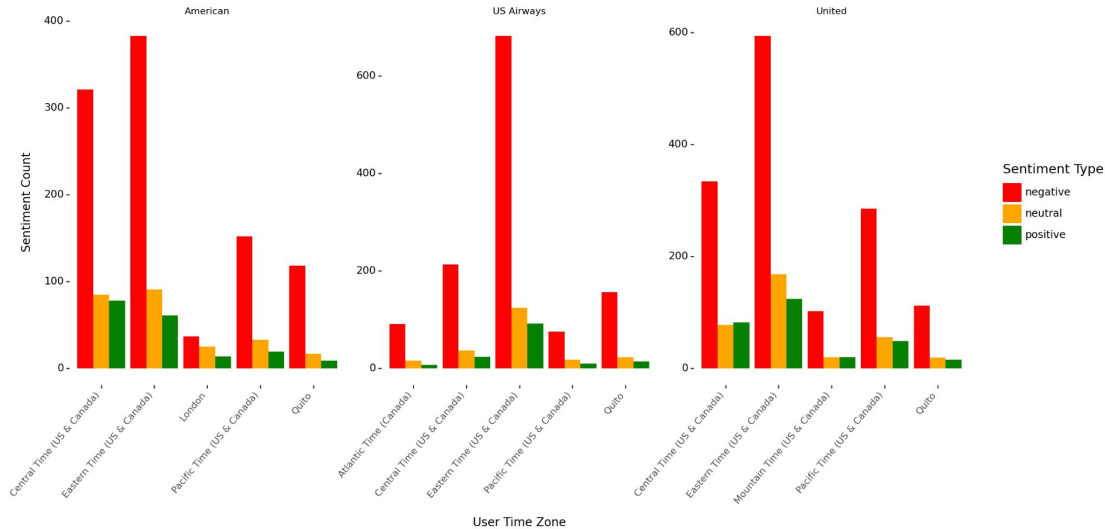
# Sentiment Analysis by Timezone



- The majority of sentiments arise from U.S. time zones, highlighting busy markets
  - There was a storm in the U.S. during the time of these flights which might be responsible for the notable influx of sentiments observed
- The predominance of U.S. users in this category may indicate that the majority of these airlines' clientele is located within the United States

# Timezone Analysis for Top Airlines

Top 5 Time Zones for Top Airline Sentiment Counts



- American, US Airways, and United airlines exhibited notably high negative sentiments
  - Particularly in the predominant U.S time zones (Central and Eastern)
- High sentiment counts align with the presence of some of the world's busiest airports in these regions:
  - JFK, Dallas, and Atlanta



# Executive Summary

- **In markets where customer service issues are prevalent, airlines that innovate and drastically improve their customer support can create a unique opportunity to become a preferred choice among travelers**
- **Analysis revealed that specific geographical regions, such as Central and Eastern U.S. time zones, generated higher sentiment counts. This can guide marketing and customer engagement strategies, recognizing the importance of these regions in operations**
- **The notable anomaly that occurred on February 22nd and 23rd underscores the significance of prioritizing proactive planning to effectively handle unforeseen events**
- **A majority of customer sentiments were negative, highlighting the need for refined airline customer experiences among major players like American, US Airways and United**





# Appendix - Action Plan

1. Loaded and previewed the data using pandas to understand its contents
2. Explored the data looking at different variables and checking for null values
3. Cleaned the data by deleting irrelevant columns and altering variable info to get ready for data manipulation
4. Conducted exploratory data analysis, extracting key statistical summaries, generating various plots to visually comprehend the data, and focused on top categories for specific subgroups
5. Documented discoveries from the exploratory quality assessment phases, summarizing key findings and insights, and data characteristics
6. Developed a presentation to properly showcase the findings

