# Twitter Airlines Data: Insight Report & Key Executive Summary

Mark Hyun

Overview

- Introduction
- Analysis of Twitter data: Insignificant variables
- Analysis of Twitter data: Airline Sentiment
- Analysis of Twitter data: Type of Complaint
- Analysis of Twitter data: Time zones
- Analysis of Twitter data: Problems & Shortcomings Encountered
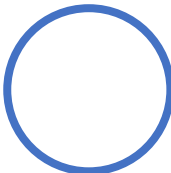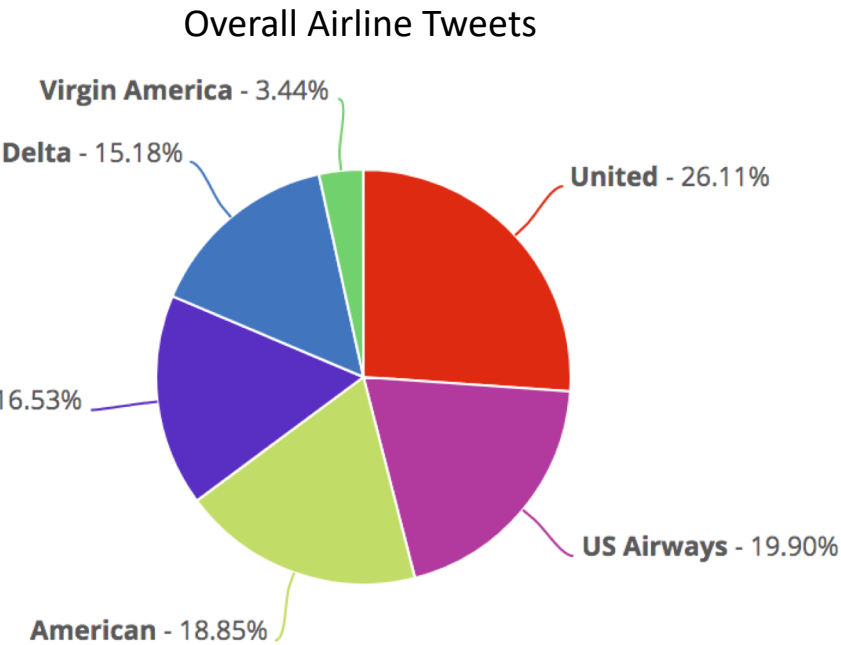- Executive Summary
- Appendix
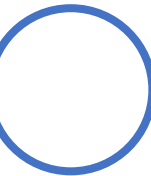
Introduction: Assessing Quality of Data

The Twitter Airlines data set provides insight on customer feedback by tweets on twitter.

- This data set contains 15 attributes and 14,640 rows.

- The Overall Airline Tweets pie chart shows the majority of the tweets were directed towards United Airlines (26.11%) and US Airways (19.9%).

Overall Airline Tweets

Virgin America - 3.44%

Delta - 15.18%

United - 26.11%

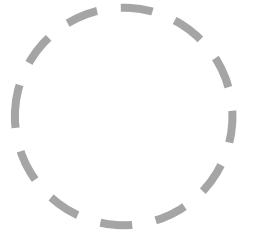Southwest - 16.53%

US Airways - 19.90%

American - 18.85%

Insignificant Variables

- Name: Twitter names do not provide significant information needed for visualizing the twitter data. Goal of the data analysis is to view the data from a more wholistic approach and find correlations.

- Tweet Coord: This variable gave no insight when trying to visual the locations of where types of complaint may be more frequently prevalent. Variable had no significance as its only use is to pinpoint possible correlations to areas.

- Tweet ID: This variable gave no insight when trying to visual the locations of where types of complaint may be more frequently prevalent because there was inconsistencies with the values (see . Variable had no significance as its only use is to pinpoint possible correlations to areas.
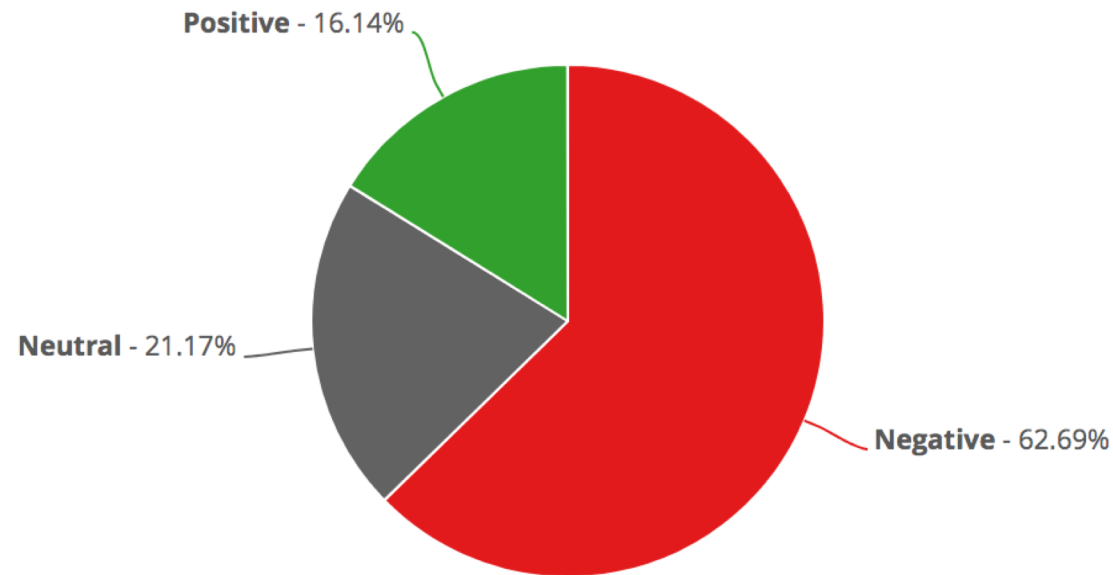
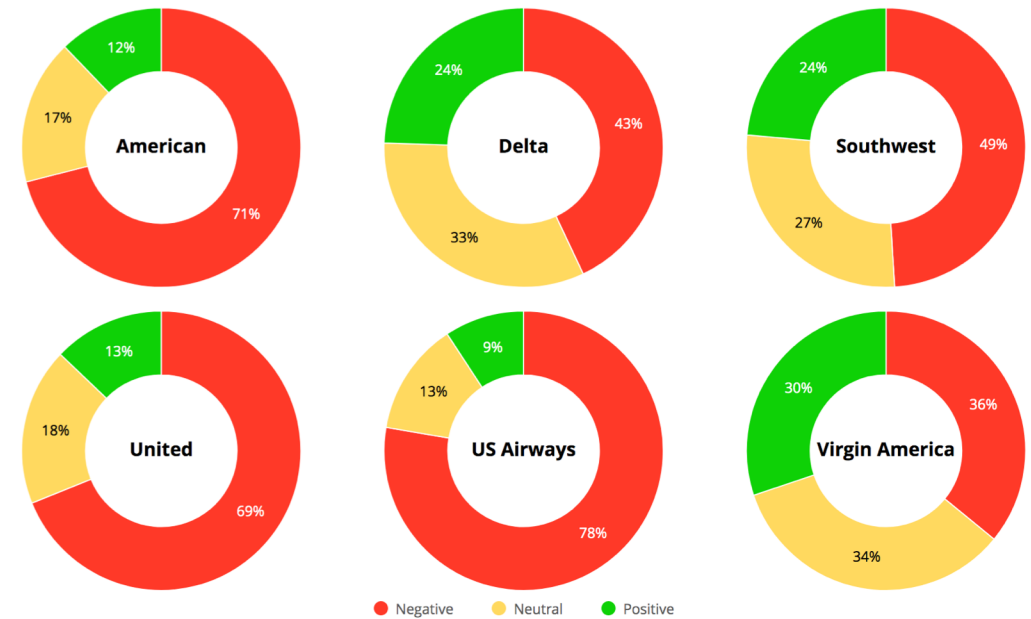|	
1	Boston, MA
2	New York, NY
3	Washington, DC
4	New York
5	USA
6	Chicago
7	Los Angeles, CA
8	New York City
9	NYC
10	San Francisco, CA
11	San Francisco

Airline Sentiment

- The Overall Airline Sentiment pie chart shows that majority of the people are most likely to post about their negative experience compared to that of a positive or neutral experience.

- To no surprise, United, American, & US Airways, also the 3 most tweeted about airlines, have the highest rate of negative tweets.
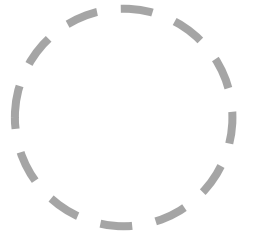

Overall Airline Sentiment

Positive - 16.14%
Neutral - 21.17%
Negative - 62.69%

Airline Sentiment by Airline

American: 12%, 17%, 71%
Delta: 24%, 33%, 43%
Southwest: 24%, 27%, 49%
United: 13%, 18%, 69%
US Airways: 9%, 13%, 78%
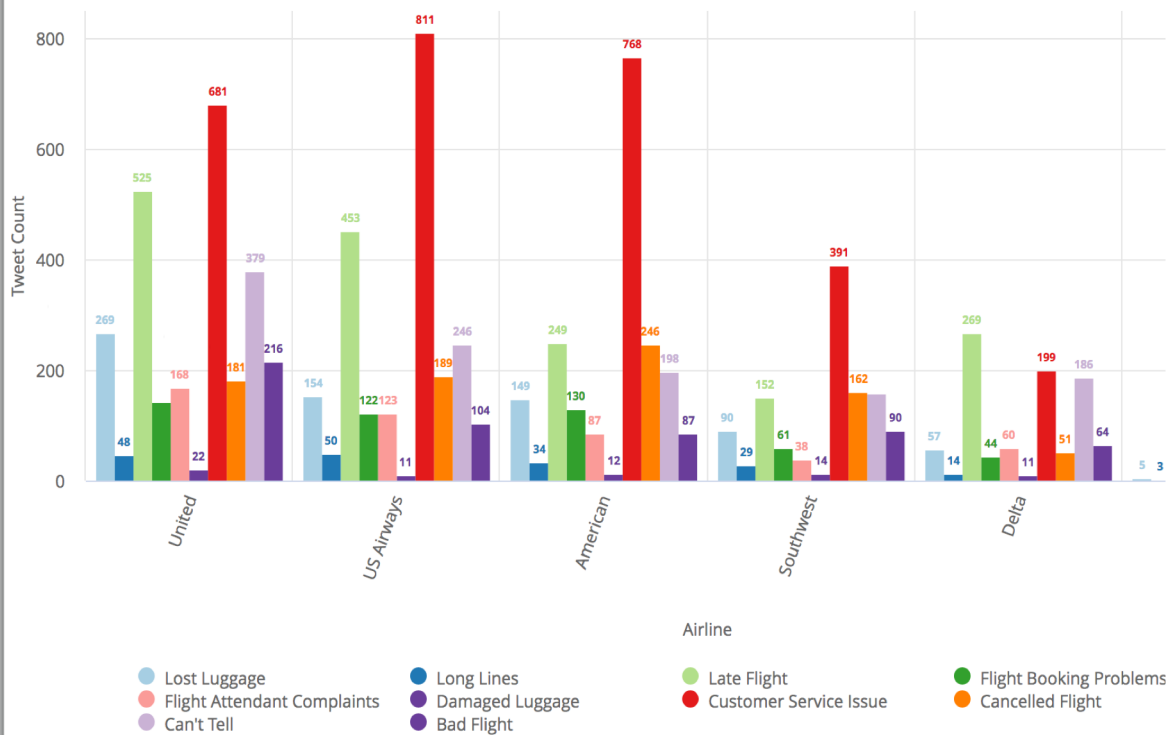Virgin America: 30%, 34%, 36%

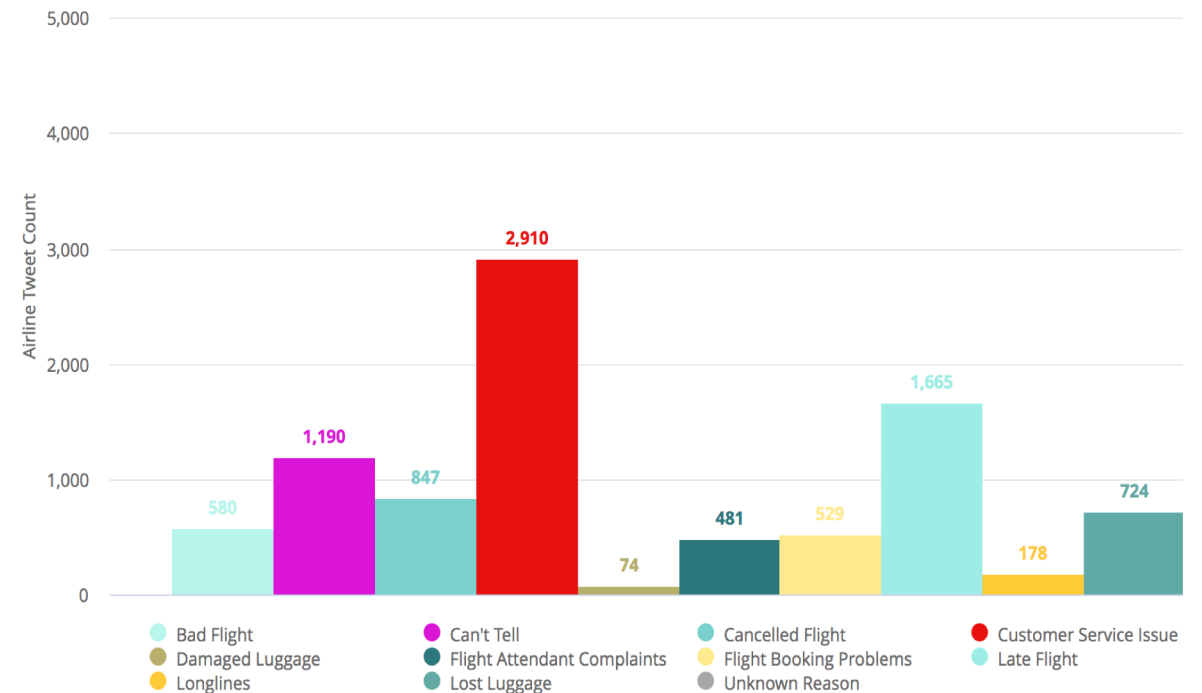Negative Neutral Positive

Types of Complaints

- Taking a closer look at the types of complaints customers are tweeting about. The airline customer service seems to be the biggest problem to the customers in both overall and individual airlines.

- United, US Airways, and American seem to have the most prominent customer service issues.
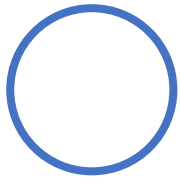
Time zones

- Based off the graph, most of the tweets seem to be based in Eastern Time (US & Canada).

- Arizona and Alaska identified as its own time zone but part of the U.S.

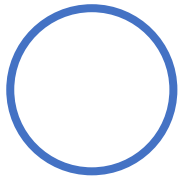- Possibility that tweets time records are based in US time zone.
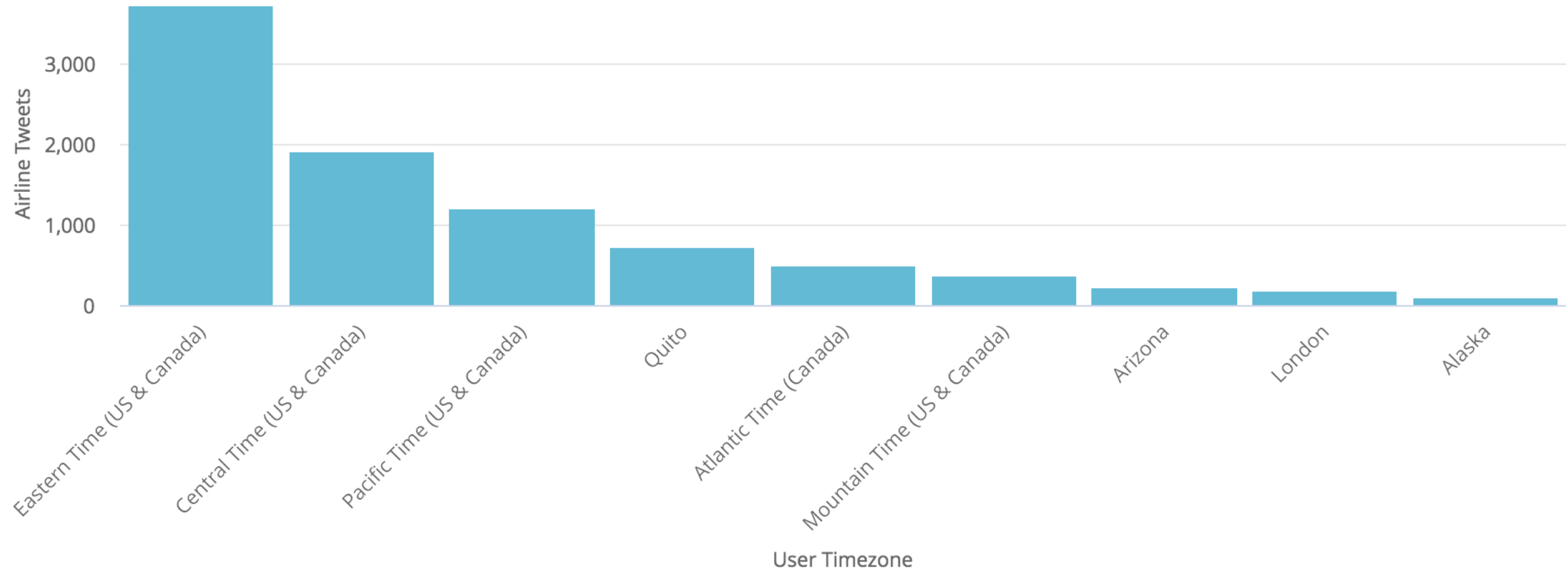
Problems & Shortcomings Encountered

- Types of Complaints – "Can't Tell": When trying to clean our data to make our analysis as accurate as possible, the count of "Can't Tell had a pretty high frequency as shown in the graph. If the "Can't Tell" values were identifiable, the may be different despite the high frequency of "Customer service issues."

- Accustomization: As a data analyst, getting accustomed and experience to resources and tools such as Looker, etc. may inhibit to maximize performance/potential for depth visualizations.

- Location/Time zone: Throughout this data analysis, correlated a relationship with location of negative feedback proves difficult due to the inconsistent and possible inaccurate data.

Executive Summary

After analyzing the Twitter Data...

- The 3 attributes that proved insignificant in our data set were name, Tweet Coord, and Tweet ID due to inconsistencies and lack of relevancy.

- Of all the airlines, United Airlines seems to received the most negative responses.

- Most of the tweets were negative due to customer service issues.

- Majority of tweets seems to be in the based in Eastern Time (US & Canada). However, there seemed to be some inconsistencies in the data.

- Shortcomings and inconsistencies in data may have an affect on affect on performance and accuracy of the analysis.

Appendix
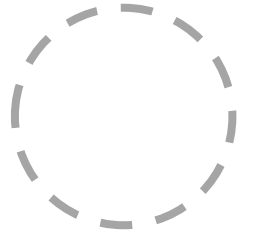
- Action Steps
    1. Clean Data
    2. Perform Exploratory Analysis
    3. Identify problems
    4. Results

    Resources used: https://purseia.looker.com/dashboards/20