



U.S. AIRLINE TWITTER SENTIMENT

Liz Lyon
MGSC-410
Chapman University

DATA SET OVERVIEW

Background:

This data set is comprised of airline customers' sentiment tweets from February of 2015.

The six airlines mentioned are American, Delta, Southwest, United, US Airways, and Virgin America.

The three sentiment classifiers are negative, neutral, and positive.

Data source: Kaggle

Original dimensions:

Rows: 14640

Columns: 15

Columns dropped for analysis:

tweet_id

airline_sentiment_gold

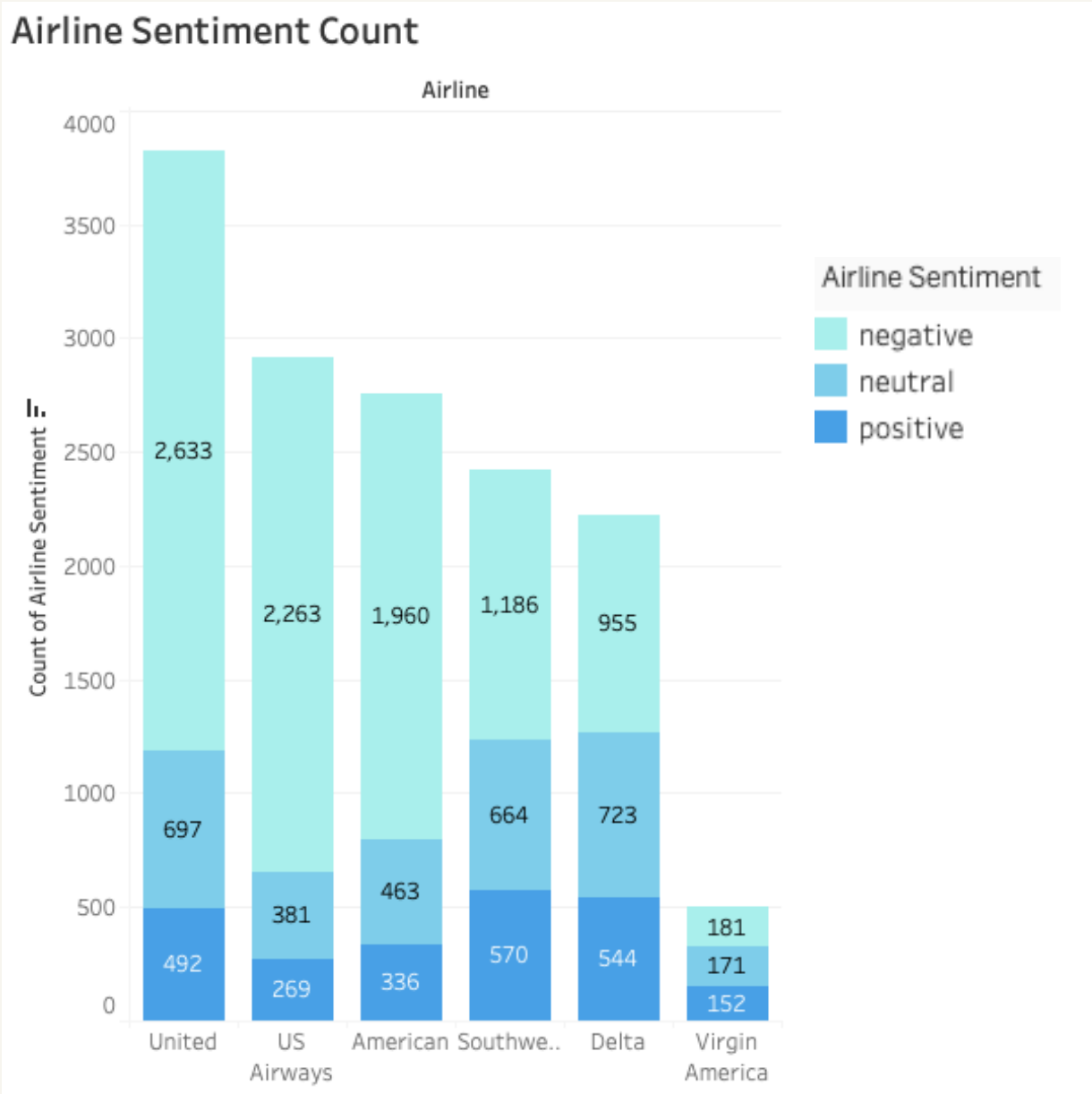
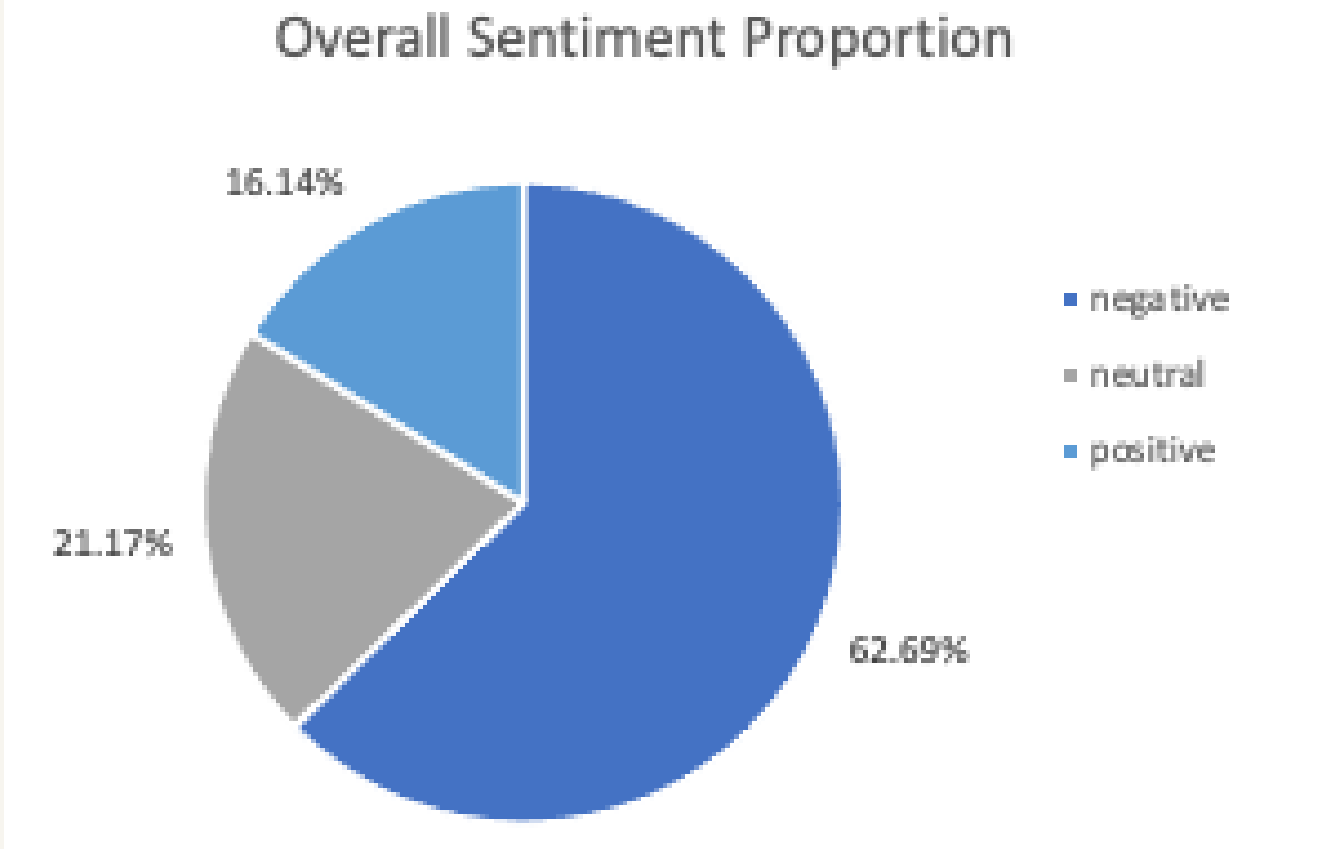
negativereason_gold

tweet_cord

SENTIMENT MEASURE

The pie chart illustrates the overall sentiment proportion of tweets. As indicated, negative sentiments were most apparent.

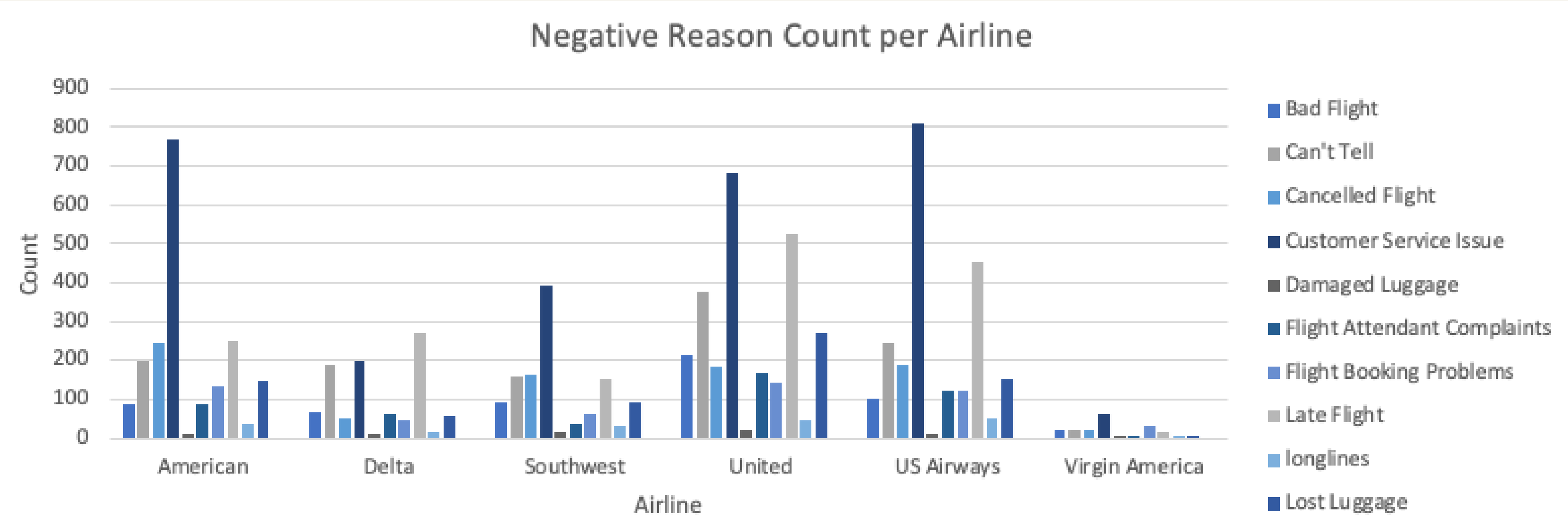
Across all airlines, United Airlines was mentioned the most in addition to having the highest overall negative sentiment count.



NEGATIVE SENTIMENT REASONS

Issues with customer service was the leading negative reason mentioned across tweets.

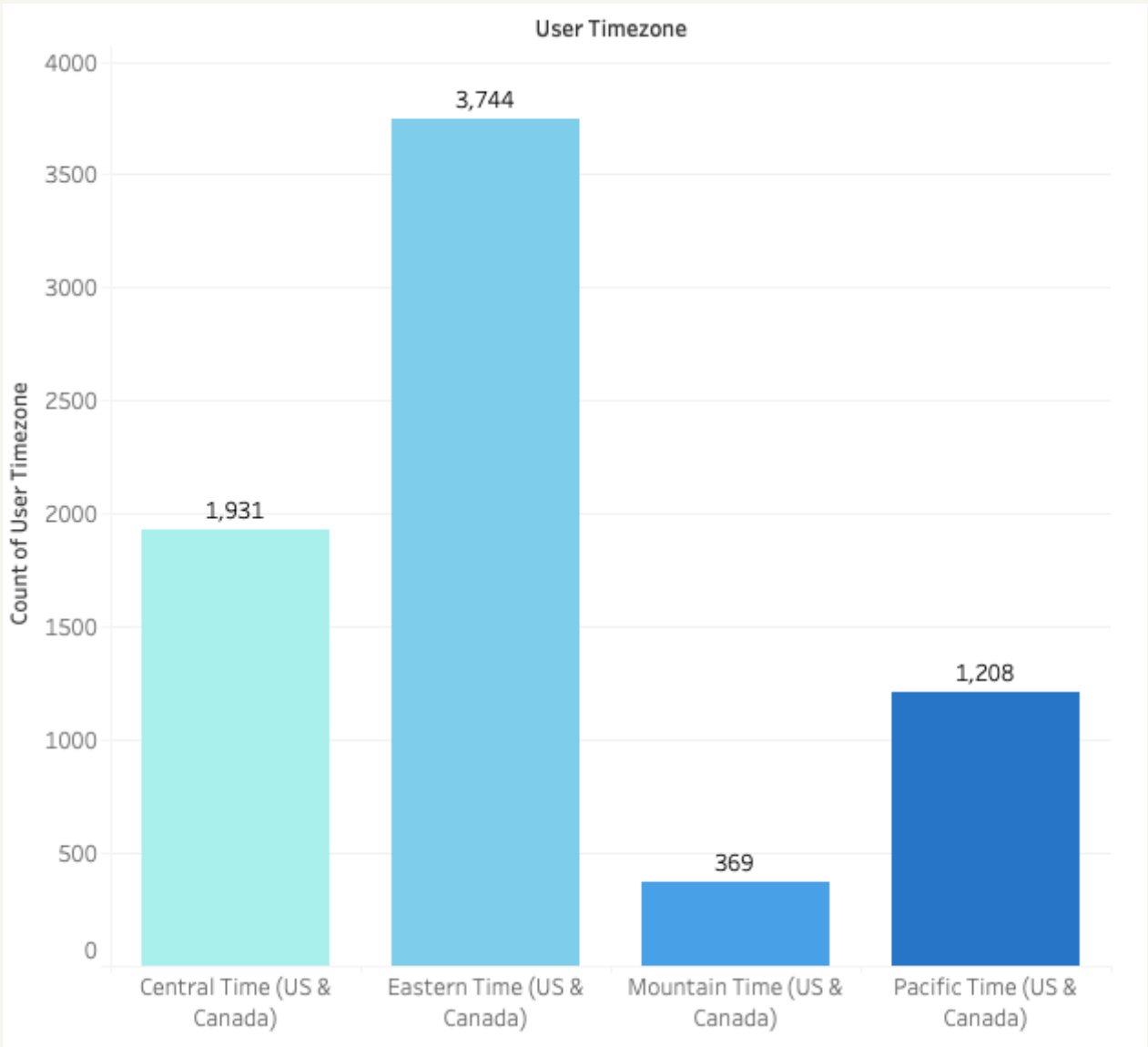
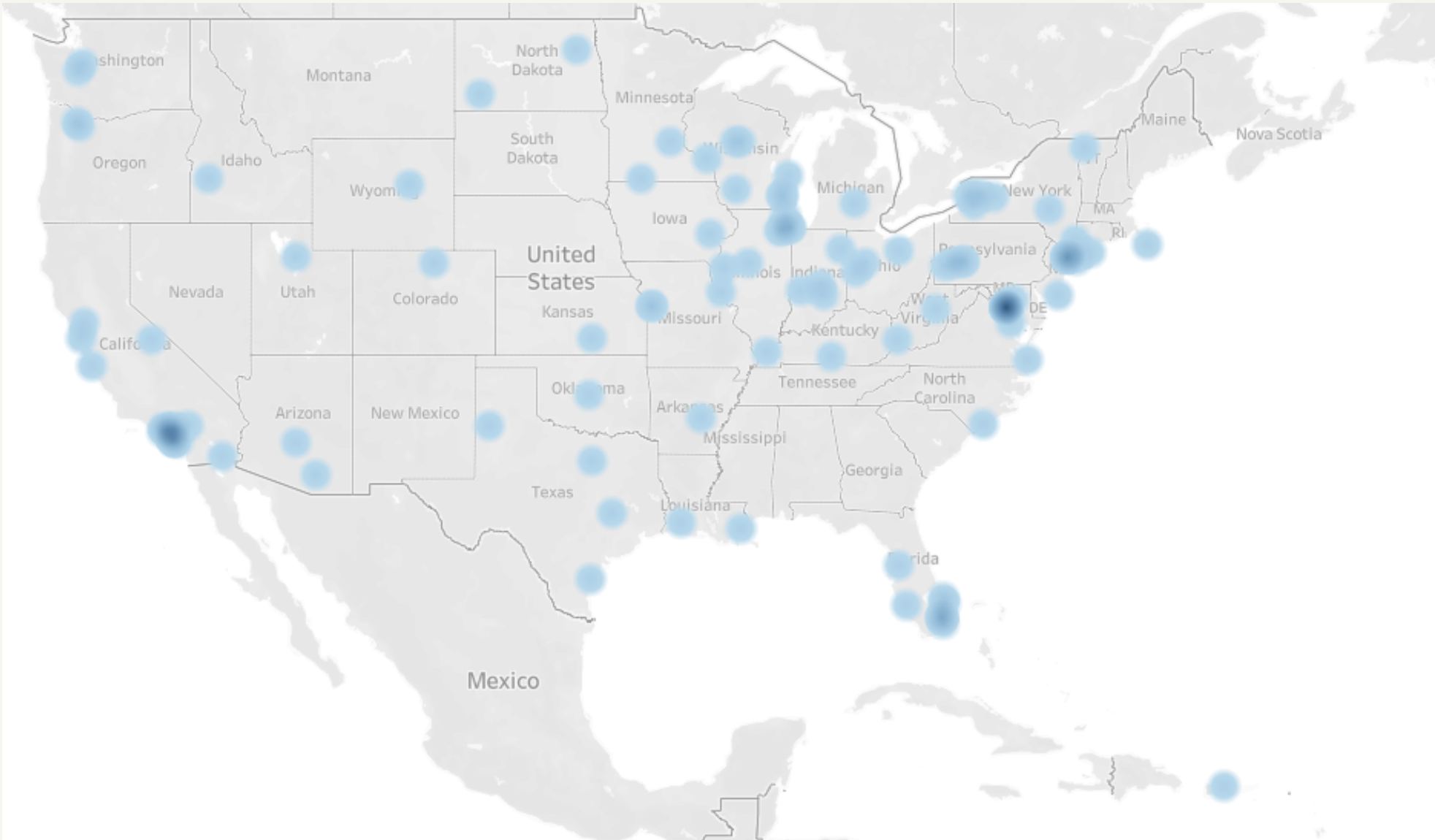
US Airways had the most negative sentiments for customer service while American was a close second.



U.S. CUSTOMER GEOGRAPHIC BREAKDOWN

Population density map of the U.S. cities where tweets took place.

Eastern Time Zone had the highest tweet volume.

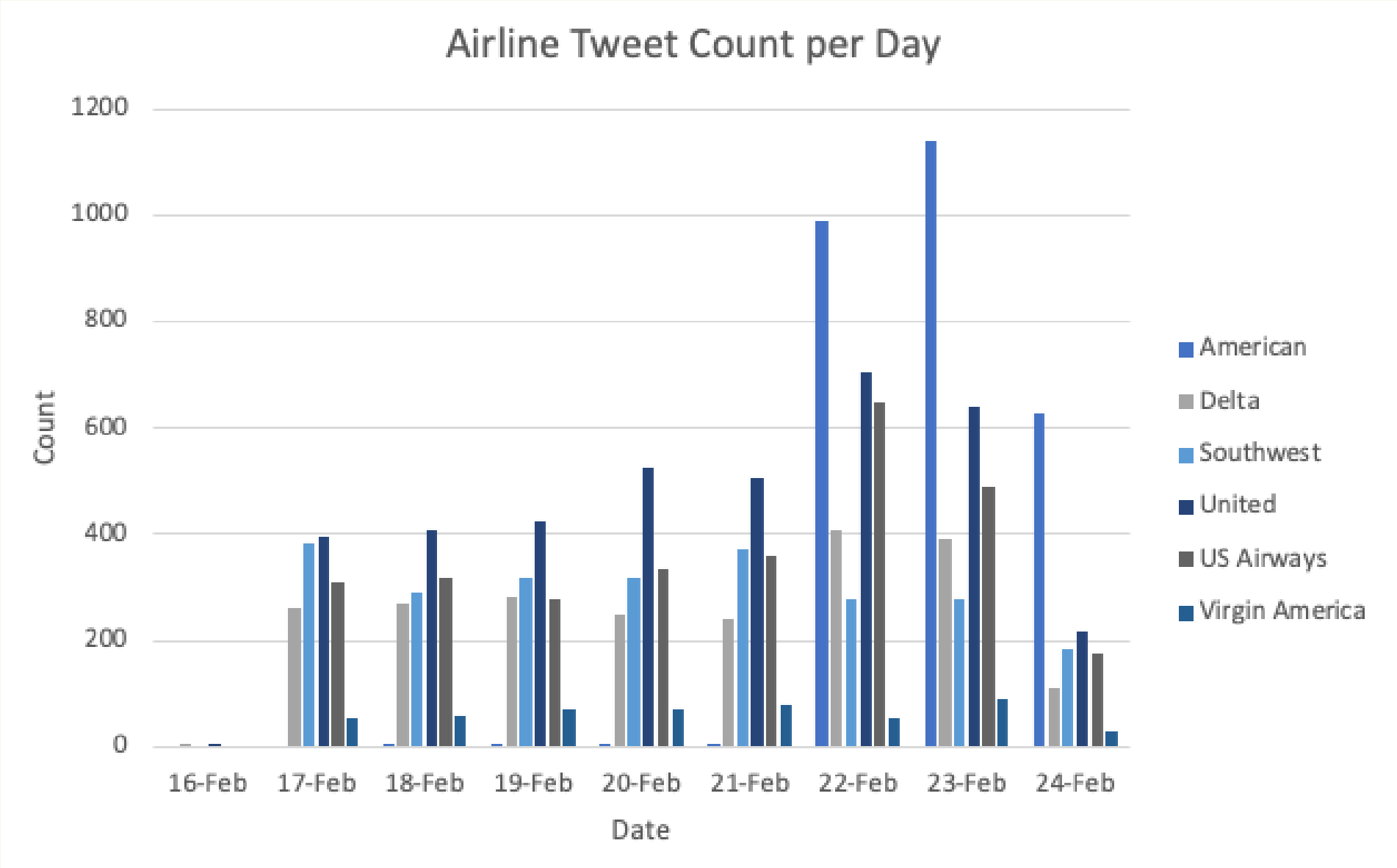


DAILY TWEET VOLUME

Tweets were collected from Monday the 16th - Tuesday the 24th in the year 2015.

Highest tweet volume was on Sunday the 22nd and Monday the 23rd.

United had the highest cumulative tweets amount while Virgin America had the lowest.



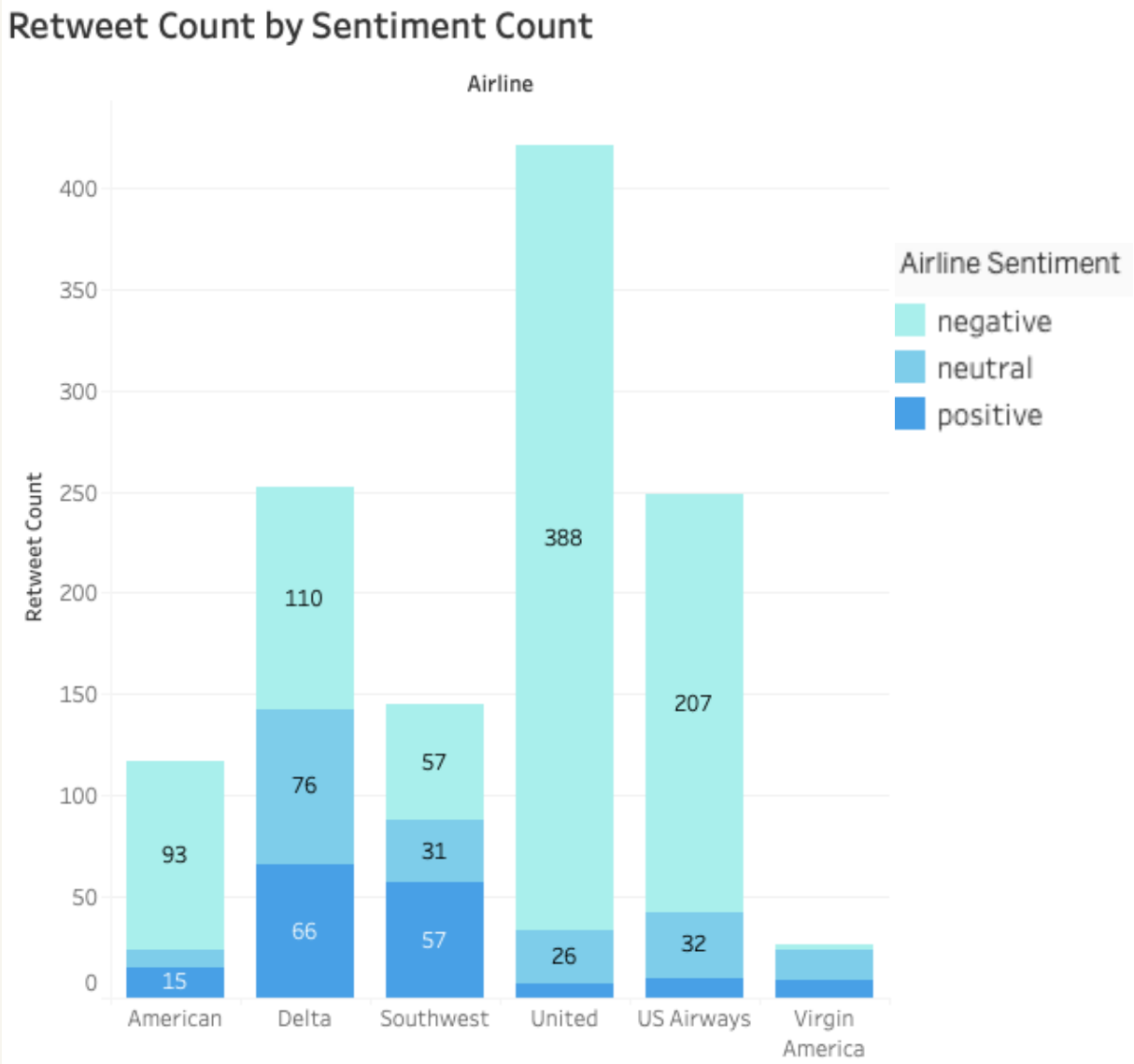
Airline	Day of Tweet Created									Grand Total
	February 16, 2015	February 17, 2015	February 18, 2015	February 19, 2015	February 20, 2015	February 21, 2015	February 22, 2015	February 23, 2015	February 24, 2015	
American			1	1	1	1	988	1,141	626	2,759
Delta	2	263	268	283	251	243	408	391	113	2,222
Southwest		384	292	317	319	370	279	276	183	2,420
United	2	396	406	426	526	506	703	641	216	3,822
US Airways		311	317	279	333	360	648	488	177	2,913
Virgin America		54	60	70	70	77	53	91	29	504
Grand Total	4	1,408	1,344	1,376	1,500	1,557	3,079	3,028	1,344	14,640

RETWEET INFORMATION

In terms of retweets, United Airlines had both the most retweets and the most negative retweets which indicates a shared experience amongst 388 customers.

The average number of retweets was 201.67. Delta, United, and US Airways rise above the average.

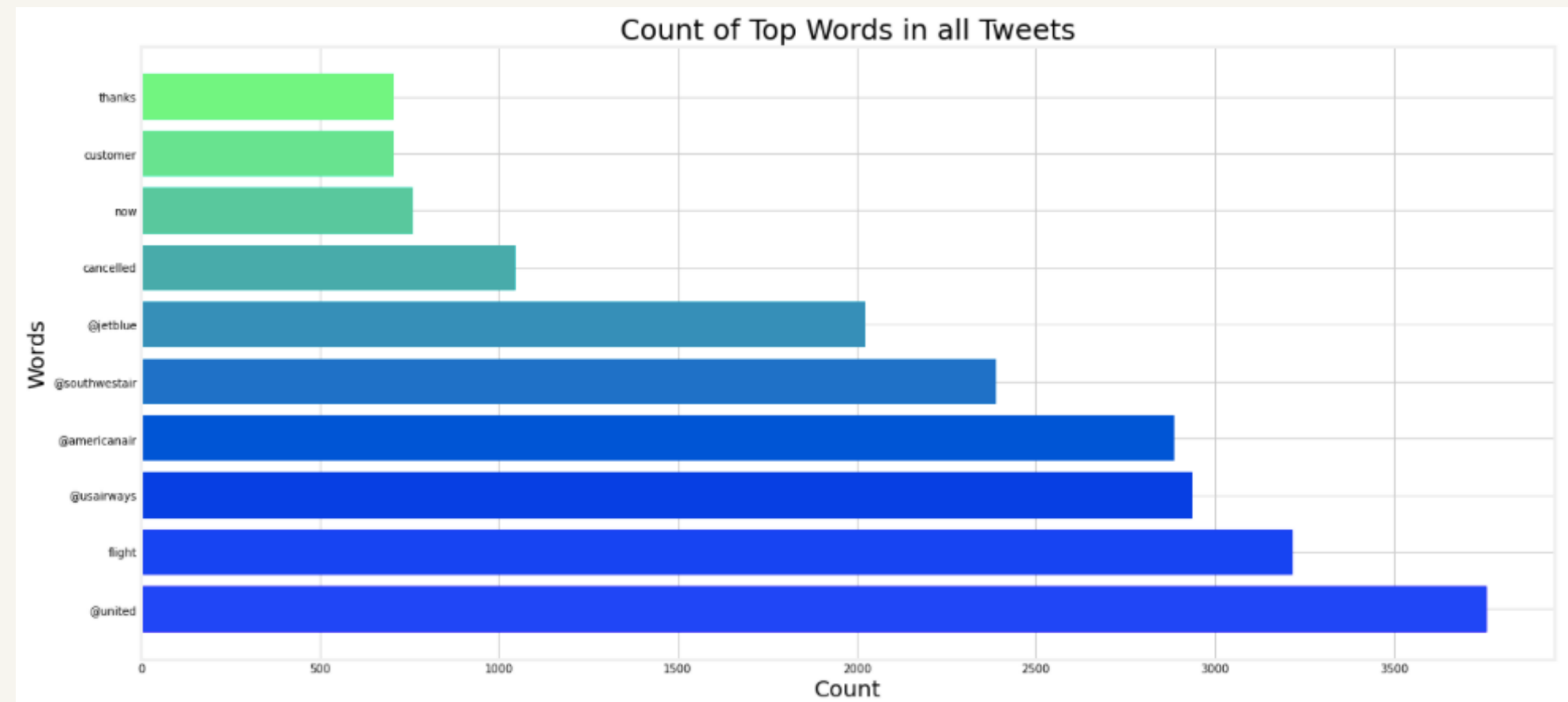
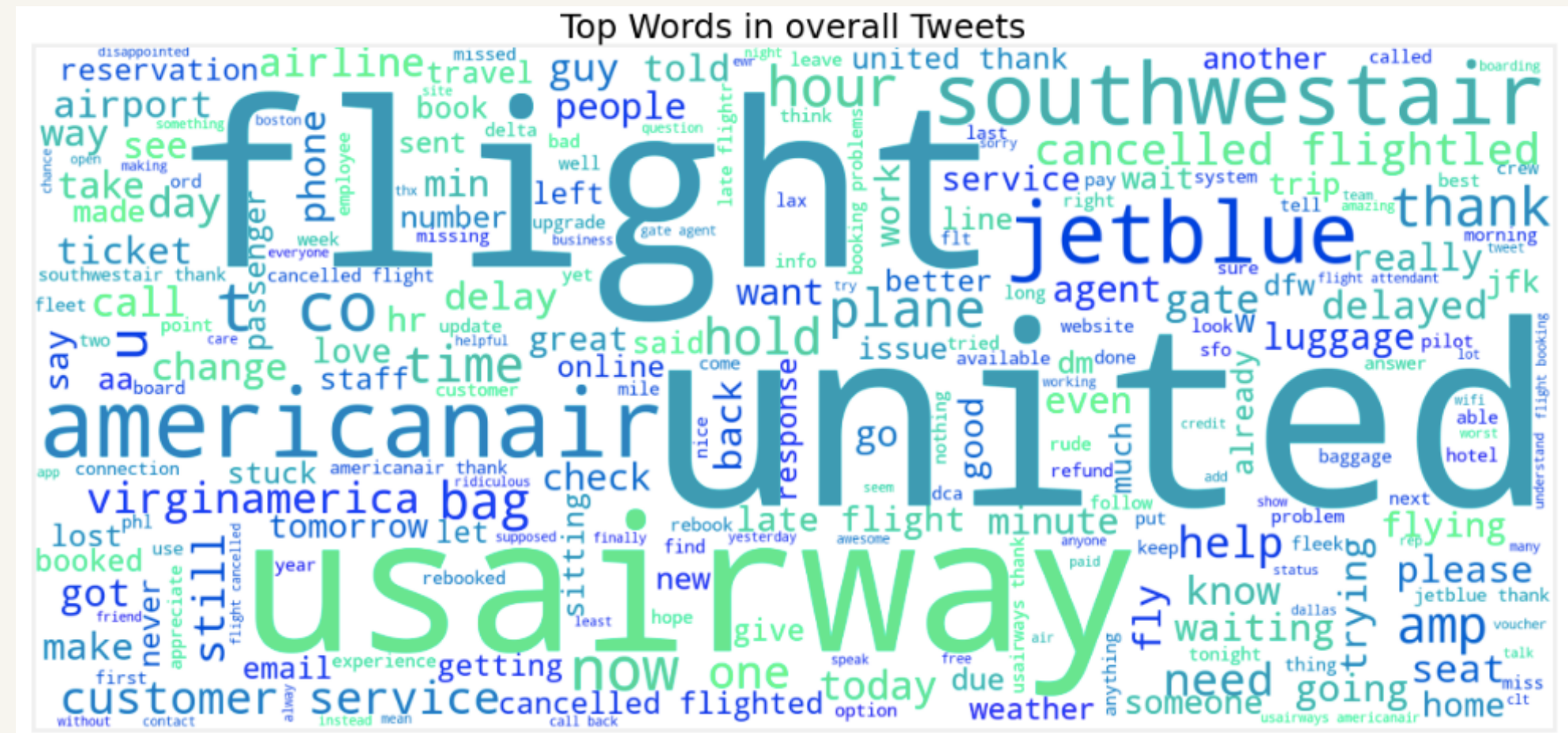
Airline	
American	9.67%
Delta	20.83%
Southwest	11.98%
United	34.79%
US Airways	20.58%
Virgin America	2.15%
Grand Total	100.00%



OVERALL WORD FREQUENCY

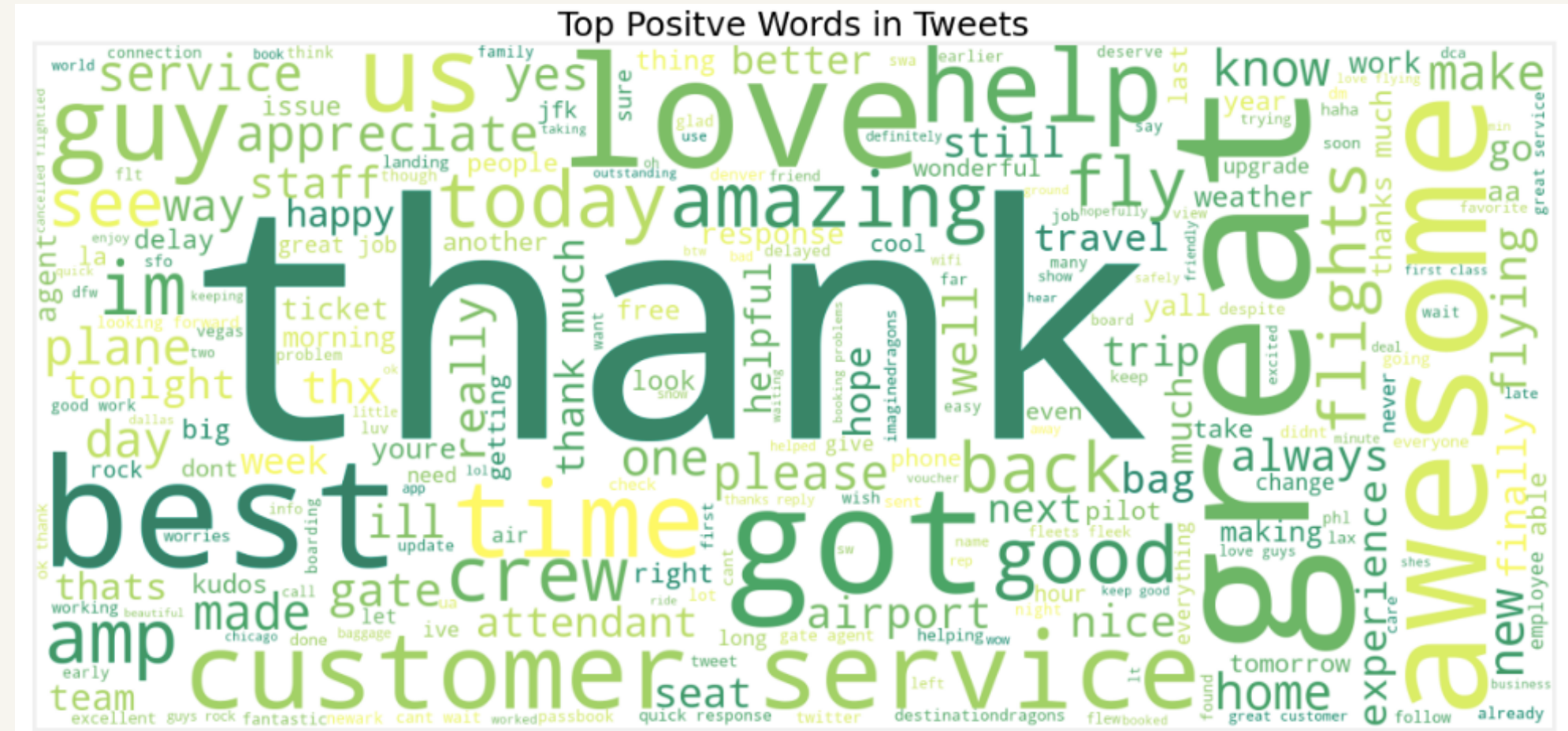
Used the Python WordCloud package, on uncleaned data, to output the most frequent words in all tweets.

The top 5 words include "united",
"flight", "usairways", "americanair",
and "southwestair".

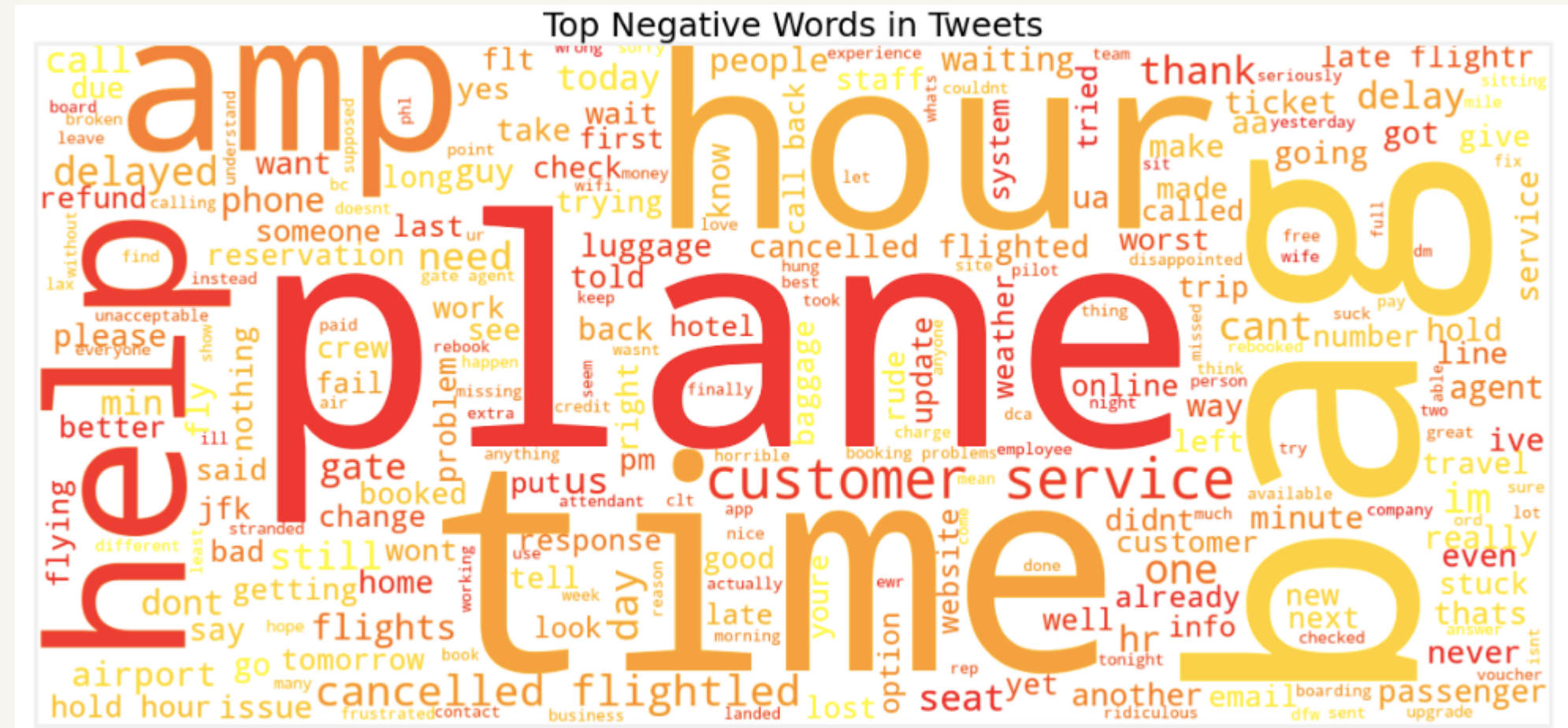


POSITIVE & NEGATIVE WORD FREQUENCY

The top right figure displays the most frequent words in positive sentiment tweets on cleaned data.



The bottom right figure displays the most frequent words in negative sentiment tweets on cleaned data.



ACTION STEPS

1. Asked questions about the data
2. Cleaned the data
3. Analyzed the data
4. Created visualizations based on analysis
5. Interpreted the visualizations

EXEC SUMMARY

1. 66% of the overall customers reported a negative airline sentiment.
2. On average, the leading customer complaint was customer service.
3. The days of the week with the highest tweet volume was Sunday and Monday.
4. Tweets that mentioned 'United Airlines' had the most retweets and of those tweets, they were largely associated with a negative sentiment.
5. 'Thank' was the most frequent positive tweet while 'Plane' was the most frequent negative tweet.

