JBG030- DBL Data Challenge
Jikun Shen (1833847)
Thomas Heller (1796887))

Annebel Hattingh-Haasbroek (1779265)
Hetvi Chaniyara (1840932)
Iris Shi (1778676)
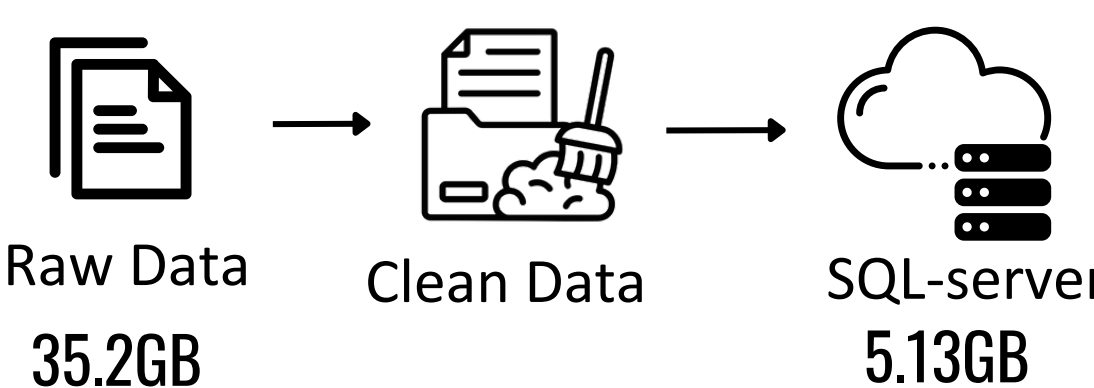
LUFTHANSA
CUSTOMER SERVICE ON TWITTER

1) CONTEXT AND OBJECTIVES

- Present findings
- Reflect on work done
- Gain client buy-in/approval
- Skills development
- Obtain feedback & input

- Twitter account reach
- Customer service quality
- Common problems faced by users
- User attitude in a conversation
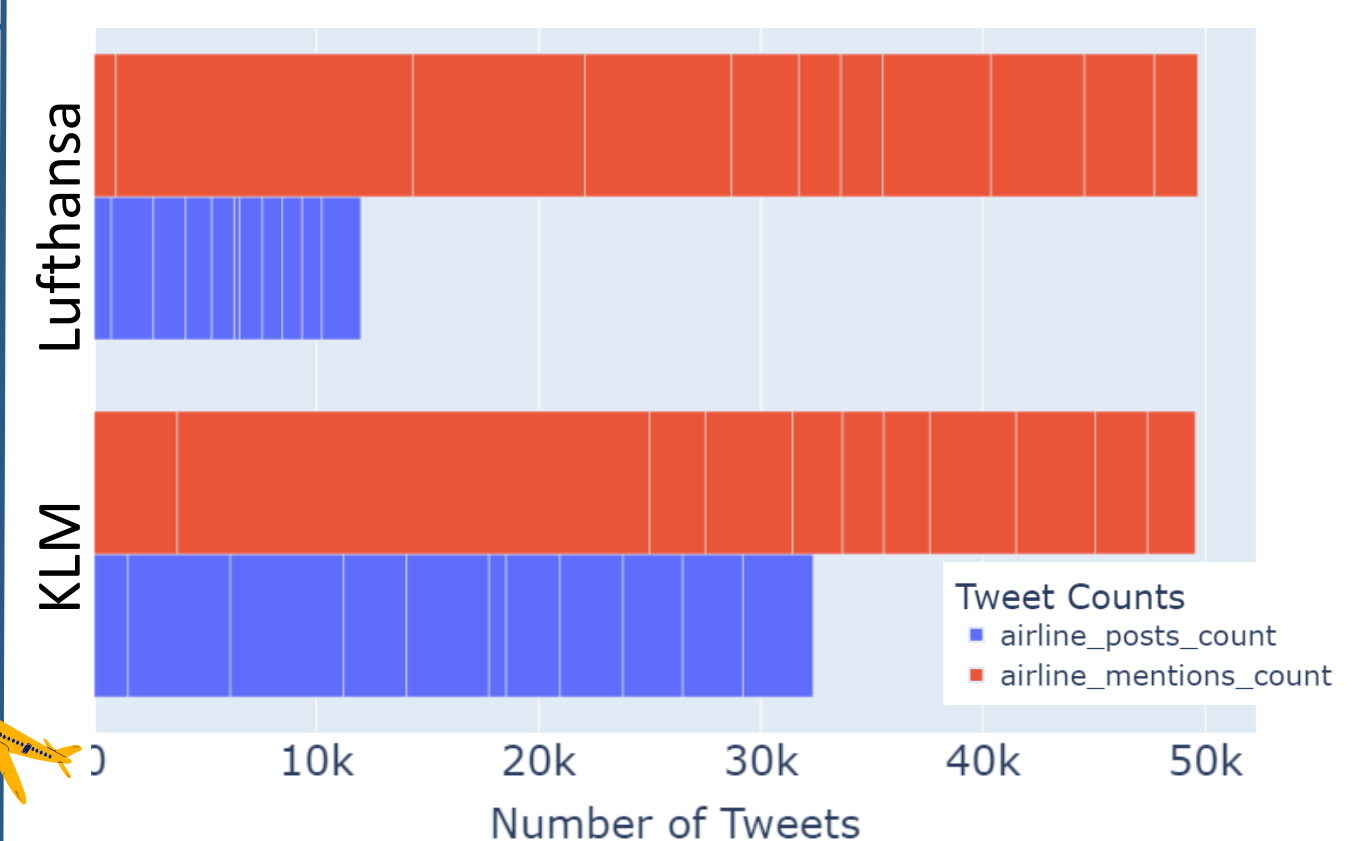- Comparison with competitors
- Efficiency during crises

2) DATA CLEANING & STORING

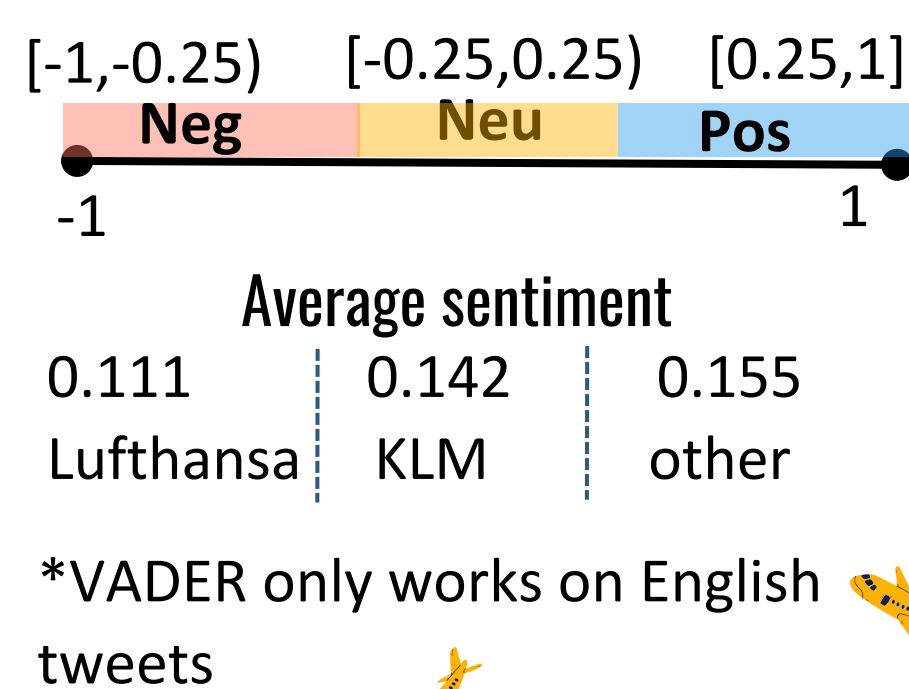Raw Data 35.2GB → Clean Data → SQL-server 5.13GB

3) INITIAL FINDINGS

- Lower number of tweets by Lufthansa compared to KLM.
- Low number of tweets mentioning Lufthansa
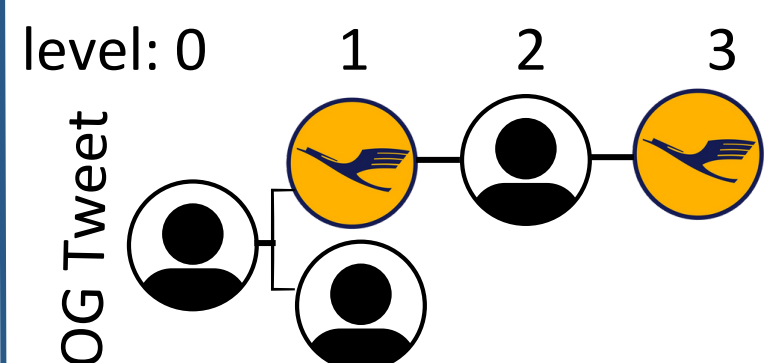- Lower number of followers of Lufthansa
- 76.3% of all tweets are in English

Number of posts by the airline VS number of mentions

5.1) VADER MODEL

[-1,-0.25) [-0.25,0.25] [0.25,1]

| Neg | Neu | Pos |
| -1 | | 1 |

Average sentiment

| 0.111 | 0.142 | 0.155 |
| Lufthansa | KLM | other |

*VADER only works on English tweets

4) CONVERSATIONS

level: 0 1 2 3

OG Tweet

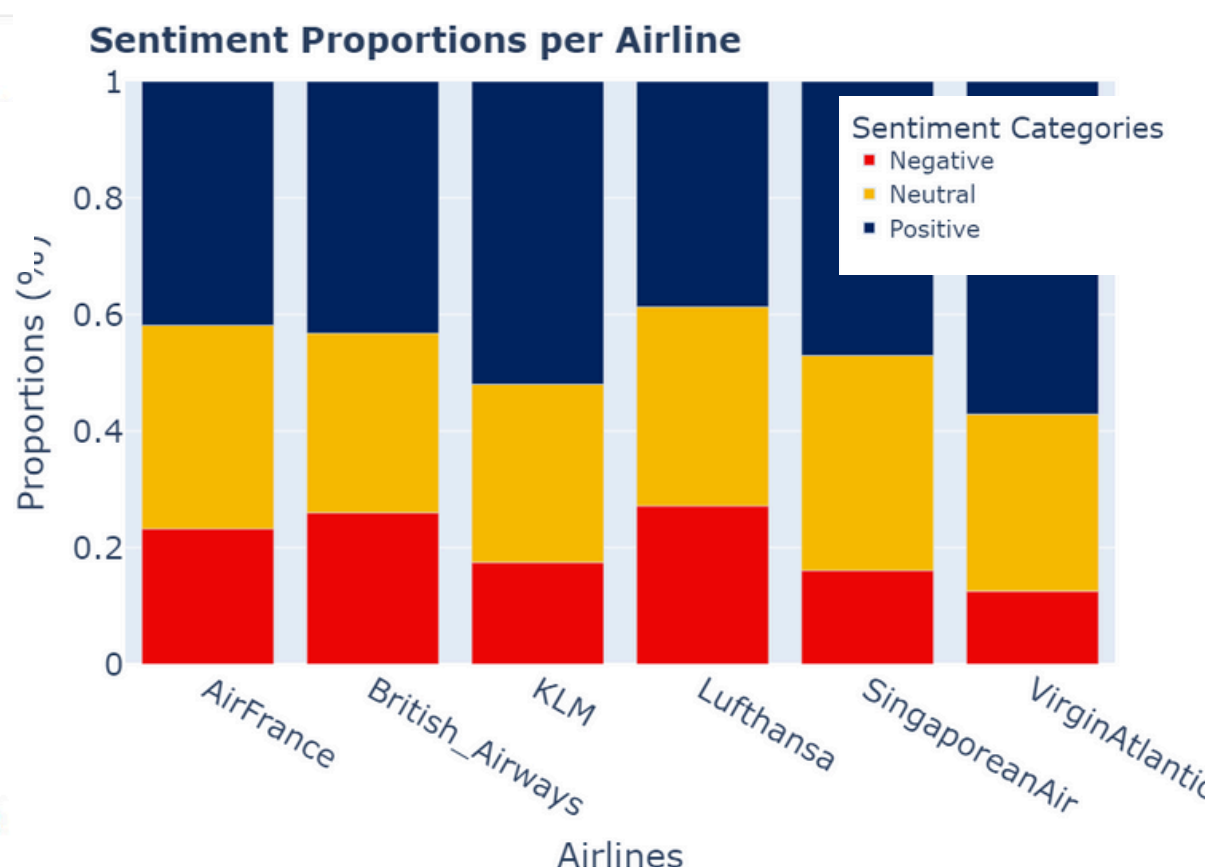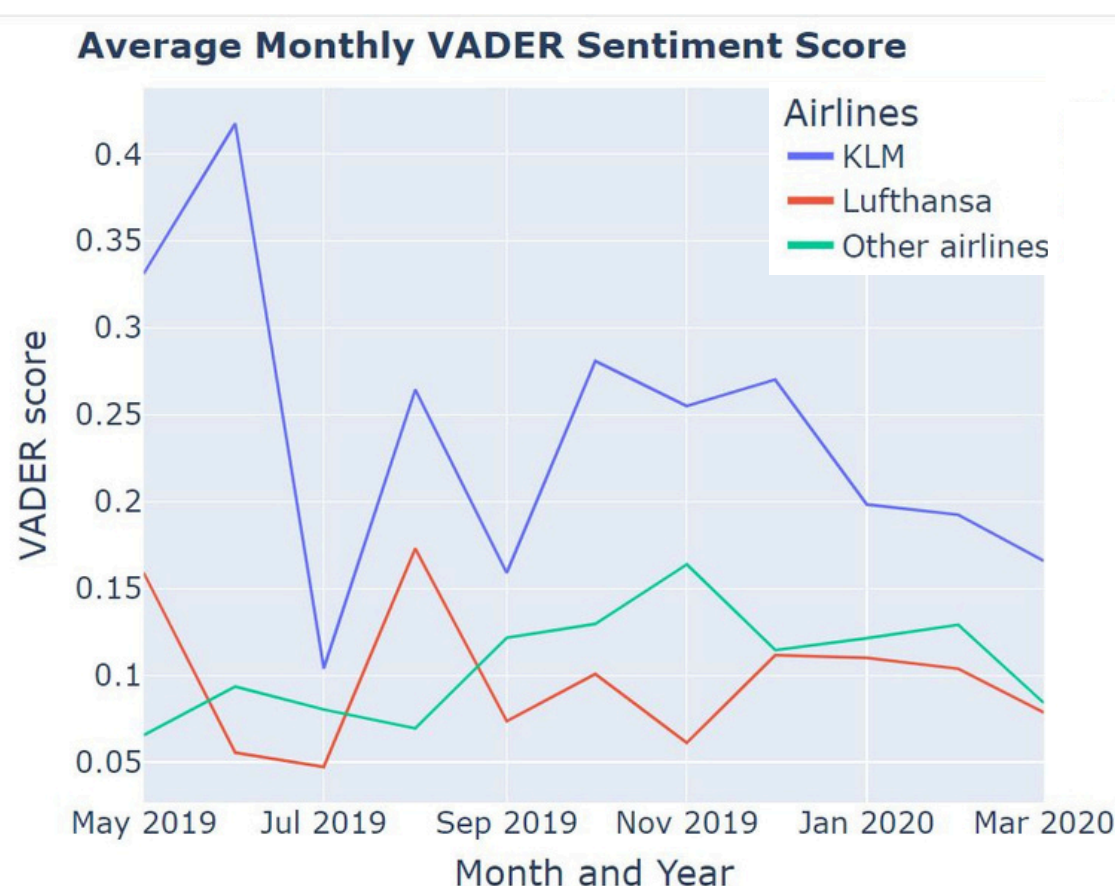| Airline	Mean Length	Conversation Total number	Unanswered Tweets	Level1 Reply	Level2 Reply
American Air	2,005	31500	14923	31640	0
British Airways	2,025	14662	8082	15023	0
KLM	2,144	4243	4025	4848	5
Virgin Atlantic	2,023	2912	2686	2976	3
Lufthansa	2,024	1864	2560	1909	0
AirFrance	2,210	1202	1775	1425	29

- Start: original tweet by user
- User has to @airline
- Reply: user & airline mentioned
- Airline reply at least once

- Lufthansa: many unanswered tweets
- KLM: many conversations & unanswered tweets
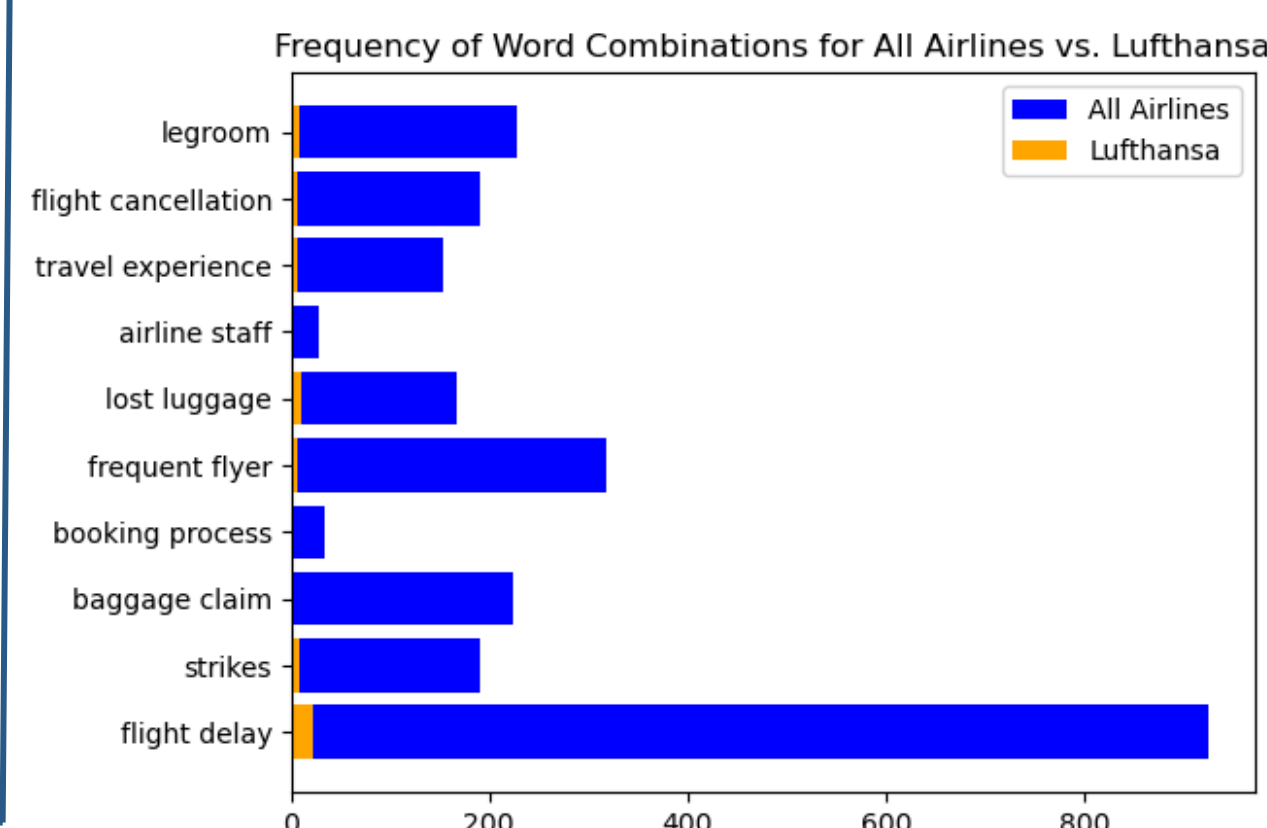- User does not reply to the airline again often

5.2) MEAN SENTIMENT

- Others: lowest in summer holiday season
- Lufthansa: highest in Dec 2019 & Feb 2020
- Lufthansa: high proportion of negative tweets compared to others

Average Monthly VADER Sentiment Score

Sentiment Proportions per Airline

7) TOPIC ANALYSIS

Frequency of Word Combinations for All Airlines vs. Lufthansa

6) SENTIMENT EVOLUTION & ACCURACY

Airline	Neg-> Neu	Neu->Pos	Neg->Pos	Pos->Neu	Neu->Neg	Pos->Neg	Pos_same	Neu_same	Neg_same
KLM	0	0	67%	0	0	33%	0	0	0
Air France	7%	17%	41%	0	0	0	24%	10%	0
Singapore Air	11%	22%	30%	2%	0	0	33%	2%	0
Virgin Atlantic	0	0	0	33%	0	0	33%	0	33%

- Many positive & neutral tweets stayed the same
- Generally increase in sentiment

Airline	Mean Diff
KLM	0.1517
Air France	-0.0459
Singapore Air	0.1559
Virgin Atlantic	-0.1398

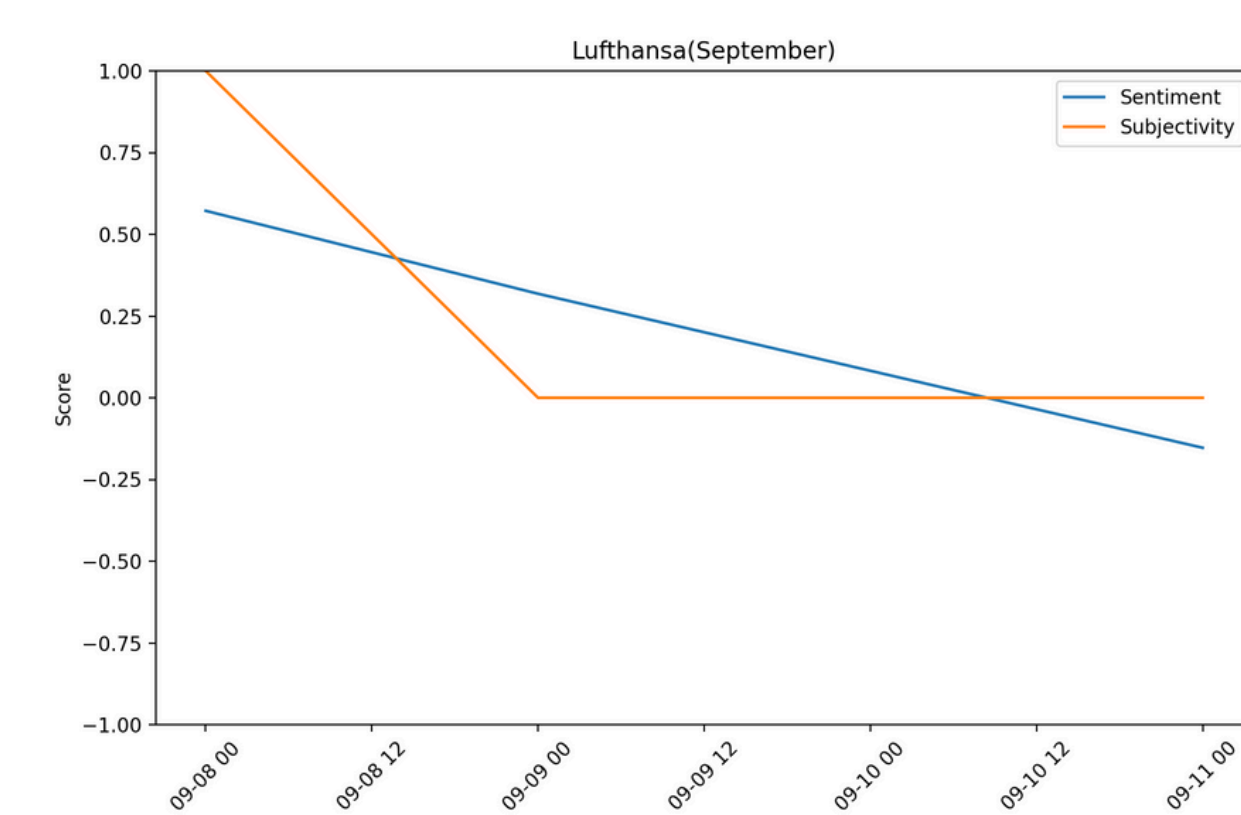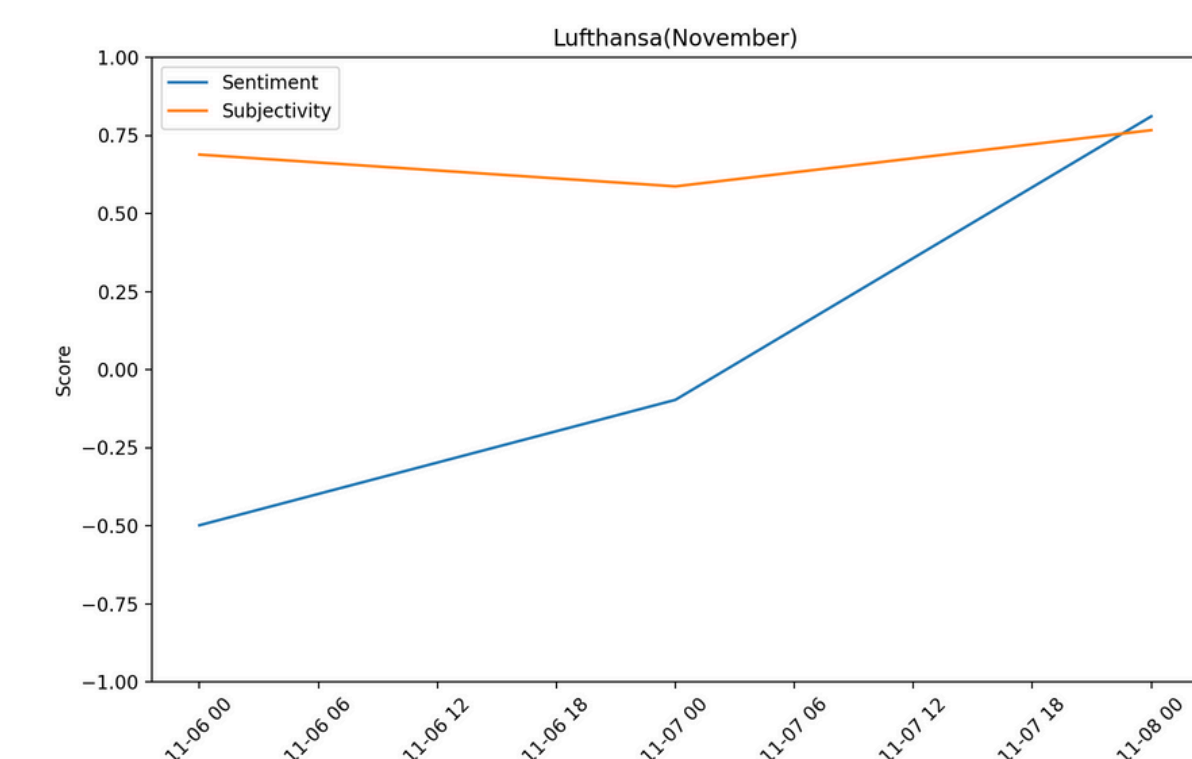- Mean change is neutral
- Biased result due to low level 2 replies

Accuracy of VADER: 85%

Based on:
- Research
- Sampling: compare VADER scores with our own

Average Sentiment Scores for Word Combinations
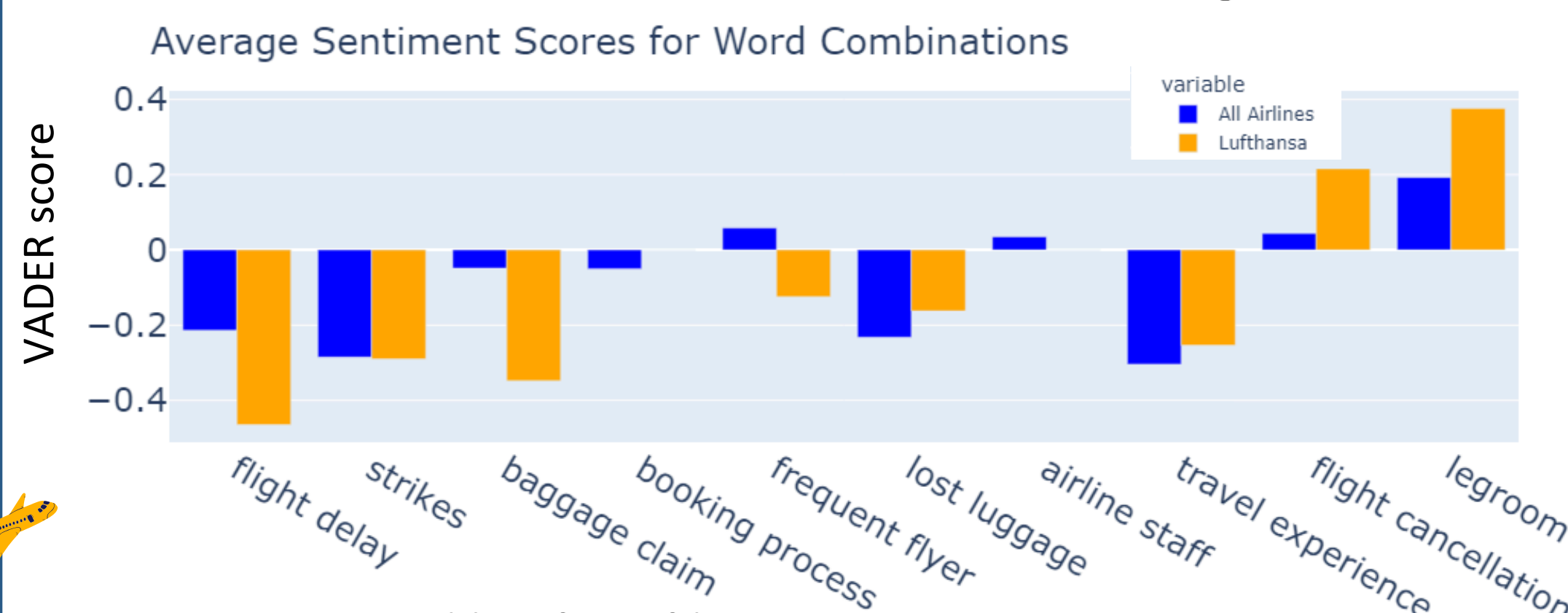
- Most common problem for Lufthansa Strikes, flight delay, flight cancellation, lost luggage.
- Surprisingly high score for legroom and flight cancellation

8) REAL LIFE EVENTS & THEIR EFFECT

Lufthansa(November)

Lufthansa(September)

Response time

Lufthansa(September)

Lufthansa(November)

Follower count fluctuation

- Lufthansa handled flight cancellations really well in November strikes

completely objective — highly subjective
0 —————————— 1

- Lufthansa did not handle flight cancellations well in September strike
- Strange: Subjectivity drops with Sentiment
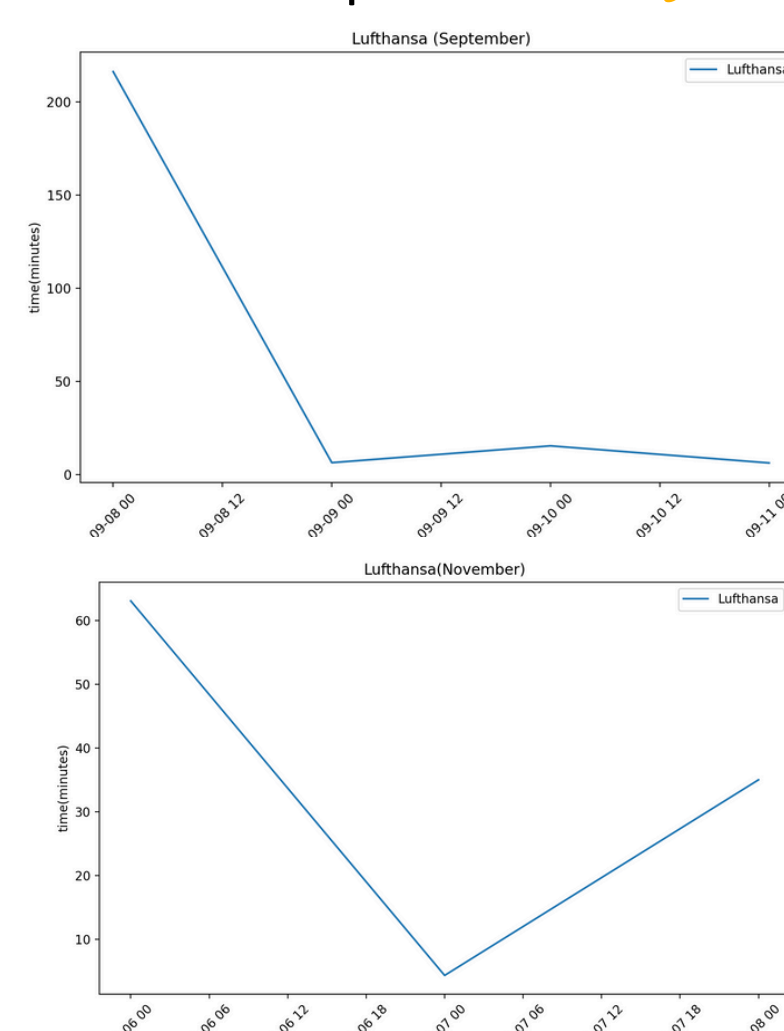
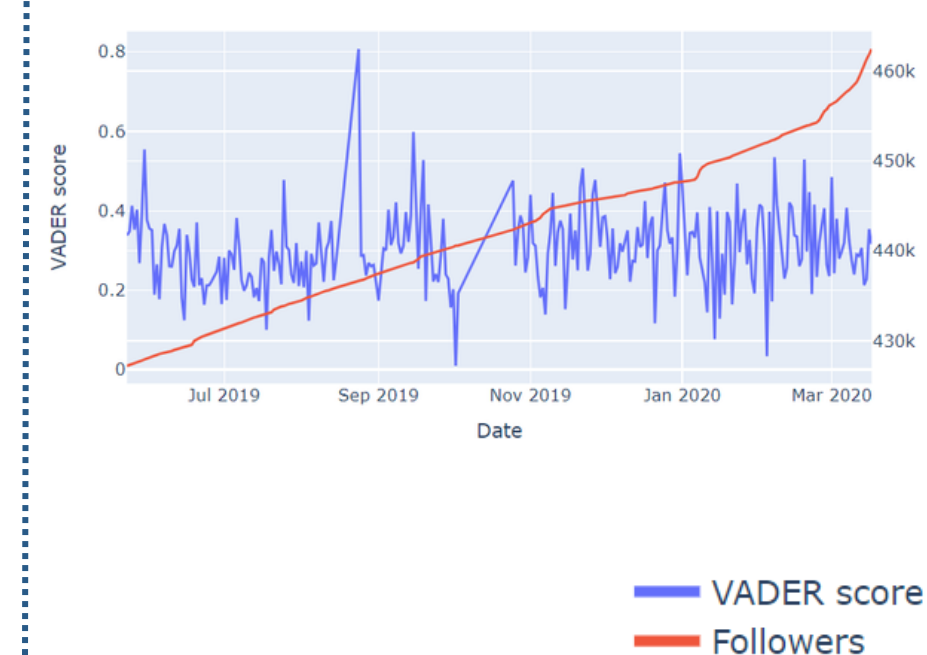- Lufthansa reply time increases after strike in September but decreases in November.

- The Sentiment fluctuates a lot with a dip in the summer
- There is a steady increase in follower count despite strikes and other events

9) CONCLUSION

Our findings:

- Lufthansa has a relatively low follower count
- They do not reply that often
- Mean sentiment of Lufthansa lower than KLM
- KLM has more followers and tweets
- Steady increase in followers for Lufthansa
- Biggest issues Baggage Claims and Flight Delays
- Positives: Legroom and Flight Cancellations
- Strikes would lead to sentiment drop
- Generally after strikes, airlines take longer to respond

Advice:
- Keep the account active
  - Steady increase in follower
  - Will increase popularity and reach
  - Allows extensive insight into public's opinion
- Advertise twitter customer service
- Prepare well for strikes so that user sentiment does not decrease too much

Why hire us again?
- We could use different models for sentiment analysis
- Look at other languages
- Look into geographical data
- Look into a bias concerning location and language

10) DISCUSSION

What we learned:
- How to work with big data files
- Decide direction based on client's wishes
- Analyzing data to see correlations
- Concluding based on results
- Advising client based on results
- Plan and work as a team
- Presenting the project and delivering our message to the audience