



Air France

Topic and Sentiment Analysis
of Customer Reviews

Justin Alger, Diego Parra, Chris Smith

Background



TEXAS A&M
UNIVERSITY®

- Our client, a leading airline company, has observed a growing trend of negative reviews across various customer feedback platforms. This shift in perception has raised concerns about the airline's reputation, customer satisfaction, and its ability to retain loyal travelers in an increasingly competitive market.
 - The client is particularly alarmed by complaints related to delays, baggage handling, customer service, and overall travel experience. While positive reviews still highlight certain strengths, the rise in dissatisfaction signals underlying issues that need to be addressed.
 - To gain a clearer understanding of the root causes, we have conducted a comprehensive topic and sentiment analysis on customer reviews. This analysis aims to identify the specific pain points and patterns driving negative perceptions and provide actionable insights to help the airline improve its services, rebuild customer trust, and enhance its overall brand reputation
-

Business Questions



TEXAS A&M
UNIVERSITY

- What are the key pain points in the customer journey?
 - Are there recurring complaints about delays, baggage handling, booking issues, or customer service?
 - What aspects of the airline experience are customers most satisfied with?
 - Are customers praising the crew, in-flight entertainment, or seating comfort?
 - What is the overall sentiment trend over time?
 - How do our customers perceive us? Has customer satisfaction improved or declined in recent years?
 - What specific words or phrases are associated with positive vs. negative reviews?
 - Are customers mentioning specific aspects of the brand that align with the airline's core values?
 - What are the main drivers of negative reviews?
 - Are there operational areas (e.g., cancellations, lost luggage) that require immediate attention?
-

Data



TEXAS A&M
UNIVERSITY

- Source: [Kaggle](#)
 - The Air France Reviews Dataset is a collection of customer reviews specifically curated to analyze feedback on Air France services. This dataset was ethically mined, ensuring compliance with web scraping guidelines and ethical standards. Additionally, all personally identifiable information (PII) has been thoroughly removed to protect user privacy, making it safe for further analysis and application in data-driven projects.
 - The dataset includes several key columns:
 - 'rating': The numerical rating given by the customer, typically ranging from 1 to 5.
 - 'title': The title of the review.
 - 'text': The main text of the review, capturing the customer's opinion.
 - 'publishedDate': The date on which the review was posted.
-

Process Overview



TEXAS A&M
UNIVERSITY®

- Data Pre-Processing
 - Topic Analysis
 - Synonyms and stop words, # topics, topic summaries
 - Sentiment Analysis
 - Overall sentiment, most positive/negative
 - Sentiment change over time
 - Clear reversal in sentiment... why?
 - Re-running analysis pre- and post-2021
 - Segmented topic analysis pre- and post-2021
 - Further Data Exploration
 - Business Impact and Recommendations
-



Topic Analysis

Topic Analysis

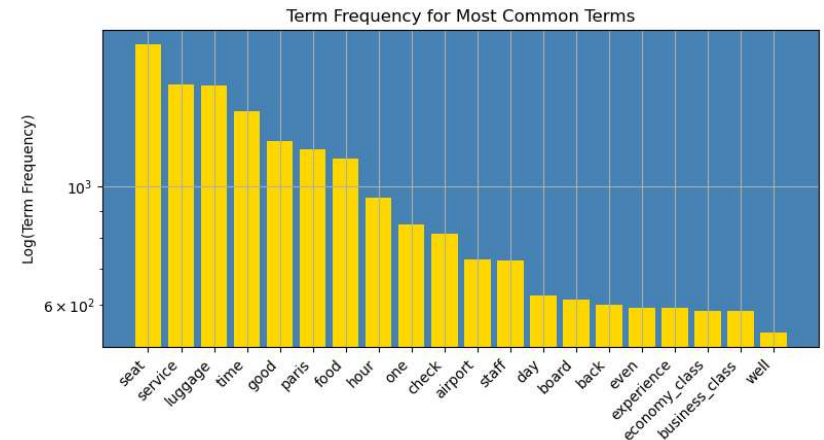


TEXAS A&M
UNIVERSITY

- Suppressed top and bottom 5%

```
***** TOPIC ANALYSIS *****  
max_df= 0.95 suppresses terms appearing in more than 2432 reviews  
min_df= 0.05 suppresses terms appearing in less than 128 reviews  
Number of Reviews..... 2560  
Number of Terms..... 147
```

- Most used words:
 - Seat
 - Service
 - Luggage



Topic Analysis



TEXAS A&M
UNIVERSITY

Topic #1: Lost Luggage and Poor Customer Service

This topic revolves around customer frustrations with lost luggage and inadequate support. Words like *luggage*, *lose*, *ask*, *help*, and *check* highlight issues with mishandled baggage, while *customer*, *service*, *company*, and *call* suggest dissatisfaction with customer service responses. Negative terms like *bad*, *never*, and *care* further emphasize complaints about lack of assistance and care from the airline.



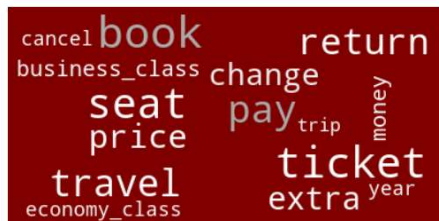
Topic #2: Airport Delays and Connection Issues

This topic captures frustrations with airport processes and flight connections. Words like *hour*, *minute*, *wait*, *late*, and *line* suggest delays and long waiting times. *Gate*, *connection*, *check*, and *board* highlight issues with navigating airports and boarding procedures, while *arrive* and *day* indicate timing concerns. The inclusion of *Paris* and *Delta* could point to specific airports or airline hubs frequently discussed in reviews.



Topic #3: Booking, Pricing, and Ticketing Issues

This topic highlights passenger concerns related to booking and ticketing. Words like *book*, *ticket*, *cancel*, and *change* point to issues with reservations and modifications. Terms such as *price*, *pay*, *money*, and *extra* suggest complaints about costs, surcharges, or perceived overpricing. Mentions of *business_class* and *economy_class* indicate comparisons between seating classes, with *trip* and *travel* pointing to broader concerns about the overall journey experience.



Topic #4: Positive In-Flight Experience

This topic focuses on positive reviews about the in-flight experience. Words like *good*, *great*, *well*, and *comfortable* highlight passenger satisfaction, while *service*, *crew*, *attendant*, and *time* emphasize praise for the airline staff. Mentions of *seat*, *food*, *meal*, *drink*, and *entertainment* reflect approval of amenities across *economy_class* and *business_class*, suggesting a well-rounded experience for travelers in different seating categories.





Sentiment Analysis

Sentiment Analysis



TEXAS A&M
UNIVERSITY

- Average sentiment 0.58*, slightly positive
- 7 sentiment words per review
- Most negative: -2.67
- Most positive: 3.38

```
=====
AIR FRANCE SENTIMENT
Length of Longest Review--- 8727
Number of Reviews----- 2560
Total Sentiment Words----- 18233
Sentiment Words/Review----- 7.12
Average Sentiment----- 0.58
=====
```

```
***** SENTIMENT CALCULATIONS *****
Corpus Average Sentiment: 0.578724709332742
Most Negative with 5 or more Sentiment Words:
Review 276 Sentiment is -2.67
Most Positive with 5 or more Sentiment Words:
Review 588 Sentiment is 3.38
Number of Reviews with zero sentiment: 26
```

bad good best
nice
help
delayed comfortable
excellent no friendly
lost
better great like
pay

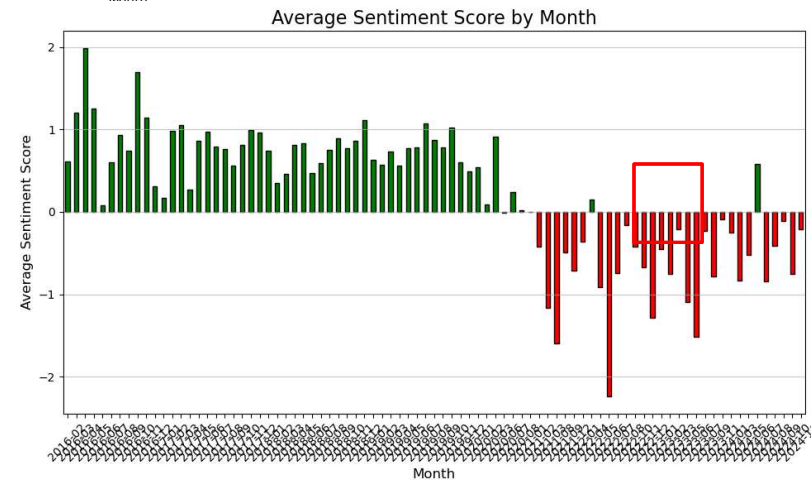
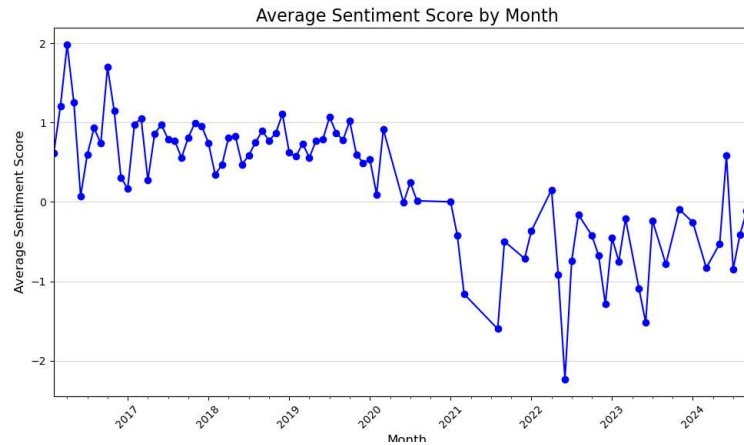
*on scale from (-5, 5)

Sentiment Analysis



TEXAS A&M
UNIVERSITY®

- Decreasing trend in average sentiment scores
- Clear positive to negative flip occurs around 2021



Sentiment Analysis – Pre-2021



TEXAS A&M
UNIVERSITY

- Prior to 2021:
 - Avg sentiment 0.77
 - 6.8 sentiment words/review
 - Sentiment words mostly positive

```
***** PRE-CUTOFF: *****

Corpus Average Sentiment: 0.7705731542336476

Most Negative with 5 or more Sentiment Words:
Review 87 Sentiment is -2.60
Review 1167 Sentiment is -2.60

Most Positive with 5 or more Sentiment Words:
Review 208 Sentiment is 3.38

Number of Reviews with zero sentiment: 25

=====
AIR FRANCE SENTIMENT - PRE-CUTOFF
Length of Longest Review--- 8727
Number of Reviews----- 2180
Total Sentiment Words----- 14850
Sentiment Words/Review----- 6.81
Average Sentiment----- 0.77
=====
```

A word cloud on a dark red background featuring various sentiment words. The words are arranged in a circular pattern, with 'good' and 'best' at the top, 'friendly' and 'help' in the upper right, 'delayed' on the far right, 'no' and 'comfortable' in the middle right, 'like' and 'lost' at the bottom right, 'great' and 'excellent' at the bottom, 'nice' and 'bad' in the middle left, 'helpful' and 'better' on the far left.

Sentiment Analysis – Post-2021



TEXAS A&M
UNIVERSITY

- 2021 and after:
 - Avg sentiment: -0.52
 - 8.9 sentiment words/review
 - Sentiment words mostly negative
 - Fewer reviews

***** POST-CUTOFF: *****

Corpus Average Sentiment: -0.5218795272040325

Most Negative with 5 or more Sentiment Words:
Review 276 Sentiment is -2.67

Most Positive with 5 or more Sentiment Words:
Review 24 Sentiment is 2.83

Number of Reviews with zero sentiment: 1

```
=====
AIR FRANCE SENTIMENT - POST-CUTOFF
Length of Longest Review--- 5207
Number of Reviews----- 380
Total Sentiment Words----- 3383
Sentiment Words/Review---- 8.90
Average Sentiment----- -0.52
=====
```

A word cloud on a dark red background featuring various sentiment words. The words are arranged in a circular pattern, with 'no' being the largest and most prominent. Other visible words include 'bad', 'worst', 'missed', 'avoid', 'help', 'great', 'cancelled', 'lost', 'rude', 'like', 'good', 'delayed', 'poor', and 'pay'.

Sentiment Exploration – *Live Demo*



TEXAS A&M
UNIVERSITY®

35.7%
1-2 Star Ratings (914)

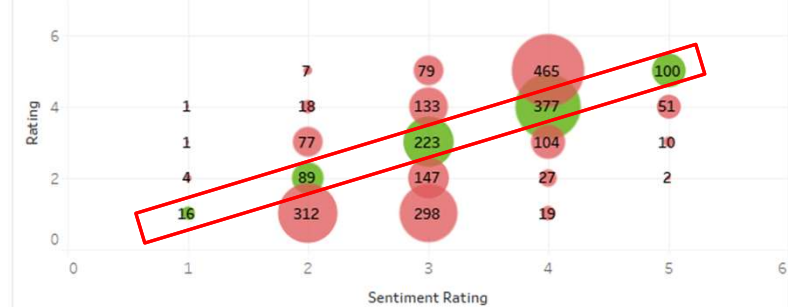
16.2%
3 Star Ratings (415)

48.1%
4-5 Star Ratings (1,231)

Average Sentiment Over Time



Rating vs. Sentiment-Derived Rating



Reviews *Select time period above to view reviews*

Month, Day, Year of Published Date	review_id	Rating	Sentiment Rating	Title	Text	sentiment_words	Sentiment	N Words
February 18, 2016	2559	5	5	AIR FRANCE HAS SPOILED ME!	planet. Special remark: When flying First, you must take some time at La Prem...	best		1
	2560	4	4	Good value for money	This was the best priced option that we cold find. Service was good and the Premium Economy class is fine for a daytime flight.	best good fine	2.67	3
February 19, 2016	2552	5	4	Best airline for me.	I have travelled fairly regularly for the last two years with Air France and I really don't understand the criticisms from some passengers. This one of the best if not the best airline I have flown with ever. I have been on Economy and ...	best best good best excellent smiling helpful great improve rewards rec..	2.55	11
	2553	3	3	Tight leg room.	This was roughly a 3 hour flight in Economy. Just as well it wasn't longer, as the leg room was really tight for average people. The snack of Coffee and Macarons was tasty though...	Null	0.00	0
	2554	4	3	Good service	The flight with Air France was fine and comfortable, especially the flight from Paris to St.Maarten. They use for the flight the Airbus A340 which a modern and comfortable plane. For your entertainment are prepared new movies and ...	fine comfortable comfortable prepared nice complain problem big chao..	0.56	9

Topic Analysis (Pre/Post-2021)



TEXAS A&M
UNIVERSITY

Pre-2021 (2180 reviews)

Post-2021 (380 reviews)

seat year small full
paris always sleep best
first french screen one old
business_class experience

claim call luggage
phone customer need book
day ticket tell

customer luggage
tell lose paris
hour arrive day time
next charles airport
delay gate
leave

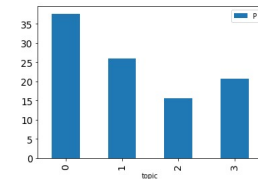
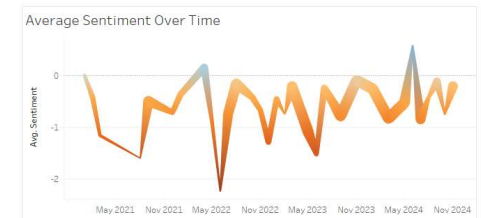
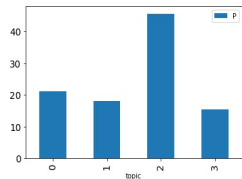
food attendant
passenger ask bad
experience
seat check
service time

nice good comfortable
crew premium_class meal
economy_class
food excellent drink
service seat leg
entertainment great

good passenger
late business_class
seat
next cdg delta crew
available

seat service
comfortable great leg
premium business
first good class food
economy small
excellent choice

company luggage
carry paris
staff long
seat two
rude check





Business Impact

Business Recommendations



TEXAS A&M
UNIVERSITY®

- Revamp Customer Service Operations
 - Enhance the In-Flight Experience
 - Boost morale and team spirit among staff
 - Rebuild trust with a loyalty campaign
-



Thank you!