

Research Essay

What factors lead to customer satisfaction and dissatisfaction in choosing an airline?

1) Introduction:

Picture this: you are sitting in your seat, waiting for the plane to take off. The engines are whirring, and you are counting down the minutes until your flight departs. Is it going to be a smooth and stress-free journey? Or is it going to be bumpy and full of delays that will leave you grumbling “never again”?

Today we are going to dive deep into all the factors that shape a passenger’s perception of an airline. We’ll look at everything from how easy it is to book a flight to the smile on a flight attendant’s face, the comfort of your seat, the quality of in-flight food and beverages, and all the frustrations that come with flight delays and high fees. So, buckle up and join me as we explore the ins and outs of pre-flight anxiety, in-flight comfort, post-flight discomforts, and more.

But this research isn’t just an academic exercise. Knowing these things helps airlines and travelers alike. Airlines learn to put passenger’s comfort and convenience first, creating experiences that build loyalty and to stand out from the crowd. Travelers, on the other hand, gain the knowledge they need to make better decisions and navigate the often-bumpy skies with more ease and less anxiety. So, buckle up and get ready for takeoff, because we are going to dive deep into the world of what brings airline passengers joy (or grumpiness) from start to finish.

2) Related Work/Theory:

Low-Cost Airline (Budget Airline; Low-Cost Carrier) is a unique business which is distinguished by a dynamic host of the competitive scheme in aviation industry. In general perspective, LCA were considered as a convenient, affordable, and alternative way to travel. **Tanomsin, P., & Chen, C. (2018)**. It is necessary to draw distinctions between business and leisure travellers when studying the determinants of customer satisfaction and customer loyalty. **Jiang, H., & Zhang, Y. (2016)**. Then certain customer's seat features, such as row location, position of aisle/window, relative location, and position with the same group member can be different before/after the change, and it can lead to huge customer dissatisfaction. **Ko, Y. D., Kwag, S. I., & Oh, Y. (2020)**. *Security and safety* is considered to be the greatest asset of the airline industry, and airlines are making great efforts to guarantee their passengers the security and safety they require. **Hussain, R., Al Nasser, A., & Hussain, Y. K. (2015)**.

3) Data:

A dataset is a collection of data. While selecting dataset for this topic, keeping ethical issues into consideration, I thought that Kaggle is the best site to get my dataset – “Airline Review”. Kaggle is an open source where lots of datasets are available for free and there are no restrictions on how we can use the data. It is also used by many people to download datasets, explore it, do some analysis, and publish it. The data in this dataset is large and tells us about people’s opinion about different Airlines in the form of Reviews. There are ratings for different facilities provided by the Airlines such as Seat Comfort, Cabin Staff Service, Food & Beverages, Ground Service, Inflight Entertainment, Wi-Fi & Connectivity. They have also given rating for value for money out of 5 and have given recommendations in the form of yes or no.

4) Methodology:

The methodology I am using for my dataset are Sentiment analysis and Topic modelling.

Sentiment analysis or opinion mining is a computational study of people’s opinions, attitudes, and emotions towards an entity. The entity can represent individuals, events, or topics. These topics are most likely to be covered by reviews. **Medhat, W., Hassan, A., & Korashy, H. (2014)**

Word cloud is basically a visual representation of text. Word clouds are mainly made with a body of text. Word clouds helps to know that how much similar an information is for specific research. **Kabir, A. I., Ahmed, K., & Karim, R. (2020).**

In Sentiment analysis I am performing Text Blob method to get word cloud of the sentiment scores. The Text Blob is a python library. Here the positive, negative and neutral score is calculated based on the polarity score. If the polarity score is greater than 0 then it is a positive sentiment. If the polarity score is less than 0 then it is a negative sentiment. If not, then it is a neutral sentiment. These sentiments are applied to the review column which contains text.

Here the stop words are removed (stop words are words like and, the) these don’t have any positive or negative feeling in the sentences.

Topic modelling using BERTopic approach. BERTopic, a topic model that extends this process by extracting coherent topic representation through the development of a class-based variation of TF-IDF. More specifically, BERTopic generates document embedding with pre-trained transformer-based language models, clusters these embeddings, and finally, generates topic representations with the class-based TF-IDF procedure. BERTopic generates coherent topics and remains competitive across a variety of benchmarks involving classical models and those that follow the more recent clustering approach of topic modelling. **Grootendorst, M. (2022).**

The `topic_model.fit_transform(docs)` takes all the information about our dataset and embeds and predicts where they will be placed and in which common group. The `topic_model.get_topic_info()` helps us in getting all topic information. If we just want top 10 topics then we can just give `head(10)`.

Before sentiment analysis the overall word cloud for Review title is calculated as

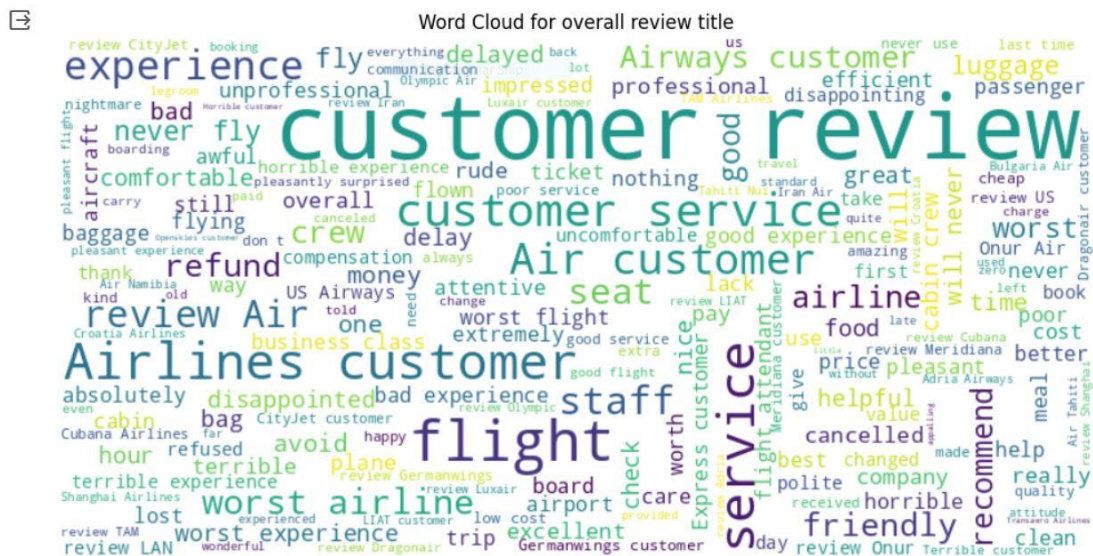


Fig. 1 Word Cloud for overall review title

The overall word cloud for review is calculated as



Fig.2 Word cloud for overall review

After performing sentiment analysis by using Text blob. The review column will have sentiment score given for each review. This sentiment analysis for overall rating is calculated and the bar

chart is shown below. Here we can see that the positive sentiment score is more than negative and neutral sentiment score.

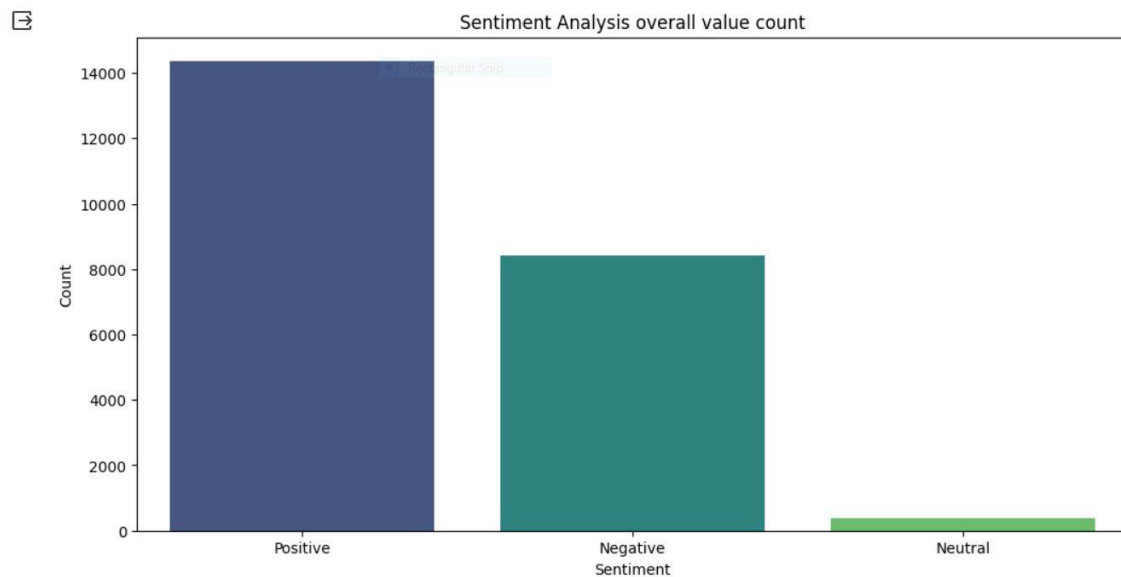


Fig.3 Sentiment analysis for overall value count

We can split the positive, negative, and neutral sentiment score into different word clouds.

For the positive sentiment score the graph is as follows:

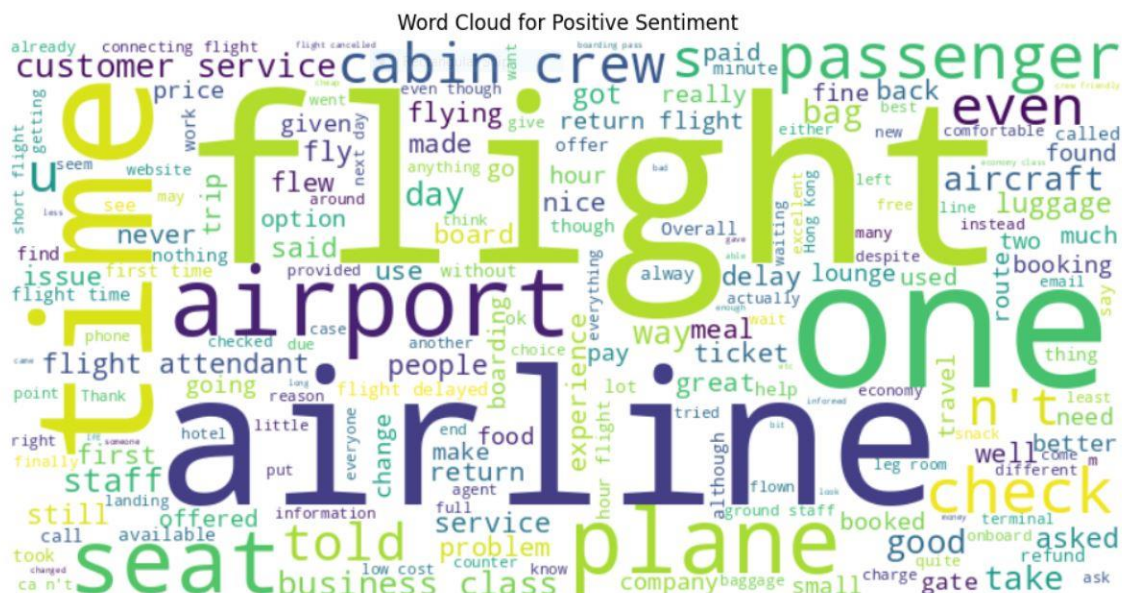
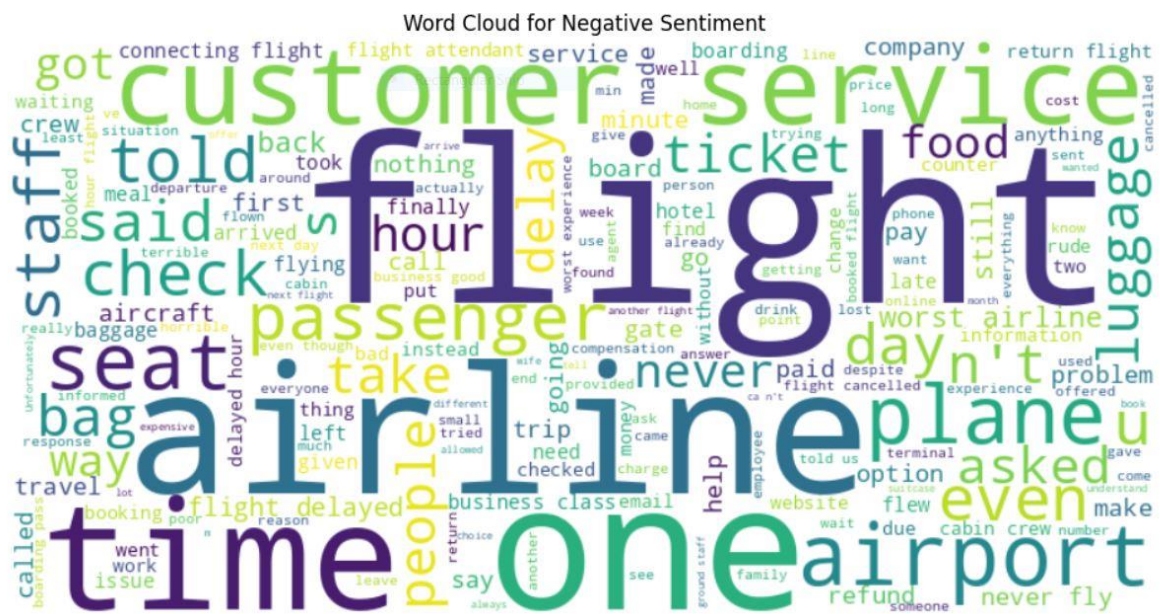


Fig.4 Word cloud for Positive Sentiment

After doing the sentiment analysis to identify which words appear frequently in the review column, we get the above image which tells us about the positive sentiment score. Here the font size in the word cloud image shows us how many times these keywords has appeared in

the reviews. These positive sentiment words like – Airline, seat, check, plane, time, flight, passenger, customer service, cabin crew and all the other words shows customer's opinion towards the airline. The less frequent words which are smaller in font size should not be ignored as they also give information about the positive sentiment of people.



Here we get the above image which tells us about the negative sentiment score.

For neutral sentiment score:

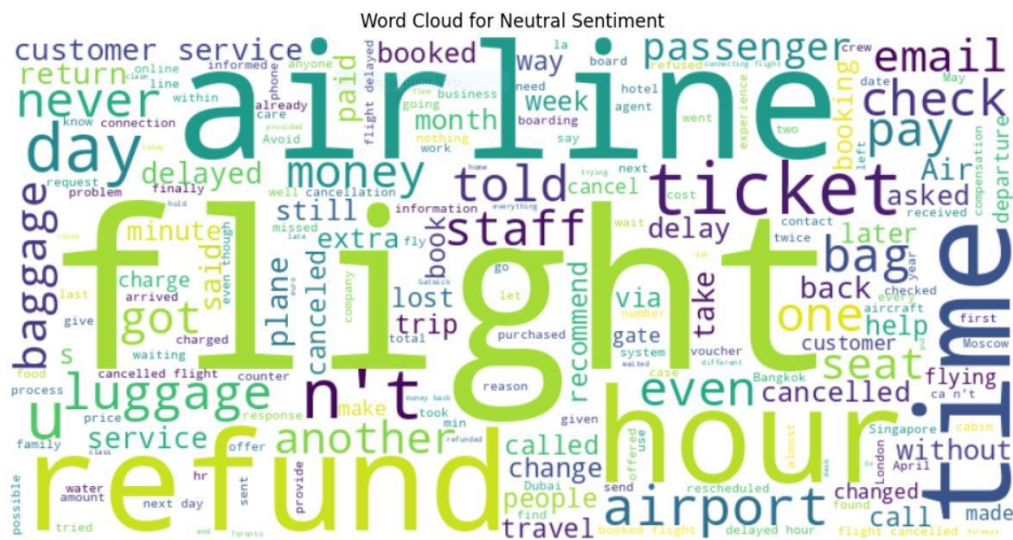


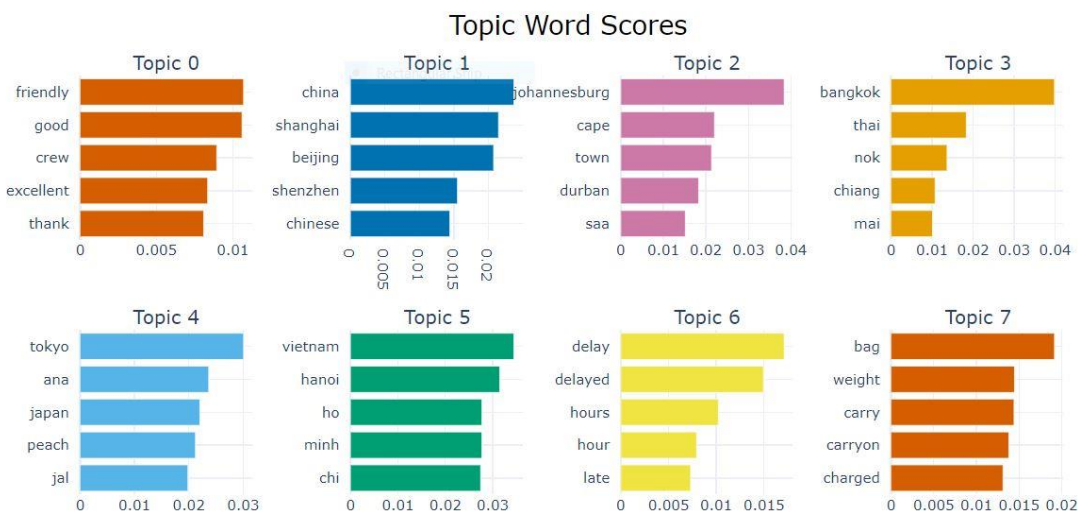
Fig.6 Word cloud for neutral sentiment

Here we get the above image which tells us about the neutral sentiment score.

Most frequently appeared keywords for neutral sentiment are- refund, baggage, email check, time, ticket, cancelled and other words which are in the image.

For Topic modelling:

After Performing BERTopic analysis, it can be visualized by showing it in Bar charts. There are 10 topics shown in that bar chart.



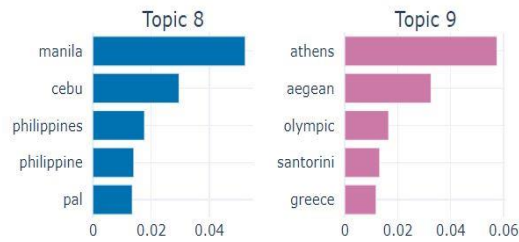


Fig.7 Topic word scores

If we consider Topic 0 – it has friendly, good, crew, excellent, thank these corresponds to positive sentiment of the customer towards the airline.

Topic 1 – China, shanghai, Beijing, Shenzhen, Chinese these all are having some similarity.

Topic 2, Topic 3, Topic 4, Topic 5, Topic 8 all have some similarity in their group of words.

Topic 6 – delay, delayed, hours, hour, late. This tells about negative sentiment of the customer towards the airline.

Topic 7 – bag, weight, carry, carryon, charged also tells us about the negative sentiment of customer towards the airline.

6) Conclusion:

We have done basic sentiment analysis here using word cloud using tokenization and removal of stop words but still if we see in the word cloud images there are some keywords which are repeating for both positive and negative sentiment word cloud images. These can be improved by removing special characters and methods such as vectorization as these are advanced process this can be taken for future related work. Even the topic model can be improved by classifying to which category it will go to using predictive models by telling which category the text belongs to such as refund related or charge related. Overall customer's satisfaction or dissatisfaction depends on many parameters. For customer satisfaction – flight arrival on time, cheap flight tickets, customer service, good cabin crew, check in happening in time, seat comfort all these adds up to customer satisfaction. For customer's dissatisfaction – delay in flight, gate changes at last moment, issues with the baggage and getting charged extra for it and problems with check in process all these lead to customer dissatisfaction towards the airline. The airline should make sure that the customers are satisfied because if they are satisfied only, they will travel in that airline again. So, these methods can be used to understand the sentiment of people towards the airline by using their reviews or by the scores given by the customer.

Word count: 1570

Bibliography:

- <https://www.kaggle.com/datasets/juhibhojani/airline-reviews>
- Grootendorst, M. (2022). BERTopic: Neural topic modelling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.
- Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. Ain Shams engineering journal, 5(4), 1093-1113.
- Kabir, A. I., Ahmed, K., & Karim, R. (2020). Word cloud and sentiment analysis of Amazon earphones reviews with R programming language. Informatica Economica, 24(4), 55-71.
- Tanomsin, P., & Chen, C. (2018). Factors affecting customer satisfaction and loyalty towards low-cost airline in Thailand. International Journal of Accounting, 3(7), 104-114.
- Jiang, H., & Zhang, Y. (2016). An investigation of service quality, customer satisfaction and loyalty in China's airline market. Journal of air transport management, 57, 80-88.
- Ko, Y. D., Kwag, S. I., & Oh, Y. (2020). An efficient airline seat reallocation algorithm considering customer dissatisfaction. Journal of Air Transport Management, 85, 101792.
- Hussain, R., Al Nasser, A., & Hussain, Y. K. (2015). Service quality and customer satisfaction of a UAE-based airline: An empirical investigation. Journal of Air Transport Management, 42, 167-175.

Code reference:

- <https://youtu.be/v3SePt3fr9g?si=laCCGkElr-Xxh6Ln>
- <https://youtu.be/eHxWshJpoRk?si=pZa1QAH5GISrgnId>