**Technical Report**

**Introduction**

This project analyses airline reviews to understand customer satisfaction in the global airline industry. It begins with a broad analysis of worldwide reviews to identify key trends, followed by a regional focus on airlines in the Gulf Cooperation Council (GCC). The final part of the project presents a detailed analysis of Gulf Air reviews. By narrowing the scope step by step, the project aims to uncover specific customer preferences and highlight areas for service improvement. This analysis is essential for understanding how airlines, especially Gulf Air, can enhance customer experience.

**Problem Statement**

Airline passengers often share detailed reviews of their travel experiences, yet the specific factors that most strongly influence overall ratings and customer recommendations remain ambiguous. This project seeks to analyse airline reviews, especially written feedback, where passengers express their genuine emotions and experiences to uncover the key drivers of satisfaction. By identifying recurring themes in customer sentiment, this analysis aims to deliver actionable insights that airlines can use to enhance service quality.

**Objectives**

1. Identify the main factors influencing airline reviews using the overall ratings and written reviews.
2. To uncover recurring themes in passenger feedbacks and classifying the issues using text analysis techniques.
3. To provide data-driven insights and recommendations for improving service quality and customer experience.

**Audience**

1. Airline companies seeking to enhance customer satisfaction and service quality.
2. Gulf Air and other GCC-based airlines aiming to better understand regional customer expectations.
3. Aviation consultants and business analysts focused on performance improvement.
4. Researchers and students interested in customer sentiment analysis and aviation data analytics.

**Datasets**

In this project two datasets were used, one for the whole analysis, and the other just played a little part just to see one factor.

1. The data was taken from Kaggle: [Airline Reviews](https://www.kaggle.com/datasets/juhibhojani/airline-reviews)
2. This data was used just to take the rating for Qatar Airways using PowerBI: [Airline Reviews Dataset](https://www.kaggle.com/datasets/sujalsuthar/airlines-reviews)

The data dictionary for the first dataset can be found in the table below:

|  |  |  |
| --- | --- | --- |
| Column Name | Description | Data Type |
| Airline Name | Name of the airline. | Text |
| Overall\_Rating | Rating given by the user out of 10. | Integer |
| Review\_Title | Title of the review given by the user. | Text |
| Review Date | The date when review was entered. | Text |
| Verified | Whether the reviewer is verified or not. | Boolean |
| Review | Detailed written review given by the user. | Text |
| Aircraft | A code representing the aircraft type or model of the flight. | Text |
| Type of traveller | Description of traveller types such as solo, couple, family. | Text |
| Seat Type | Categorical class type (Economy, business,etc) | Text |
| Route | Flight source and destination | Text |
| Date Flown | Month and Year of flight | Text |
| Seat Comfort | Rating out of 5 for seat comfort | Integer |
| Cabin Staff Service | Rating out of 5 for staff service | Integer |
| Food & Beverages | Rating out of 5 for food | Integer |
| Ground Service | Rating out of 5 for ground service | Integer |
| Inflight Entertainment | Rating out of 5 for entertainment | Integer |
| Wifi & Connectivity | Rating out of 5 for internet services | Integer |
| Value for money | Rating out of 5 if the flight was worth the money users paid | Integer |
| Recommended | Whether the flight is recommended or not | Boolean |

**Data Handling**

**Note:** The data was not cleaned all at once; instead, it was cleaned progressively during the analysis as needed. Detailed cleaning steps can be found in the Jupyter Notebook

1. Removed the unnecessary columns and the columns that had a high percentage of nulls using Excel (Removed columns: Aircraf, Inflight Entertainment, Wifi & Connectivity).
2. A column was created for cleaned text of reviews.
3. A column was created to check the sentiment of the reviews.
4. A clean column was created for date of review.

**Analysis and Findings**

Can be found in the files: jupyter lab and powerpoint presentation uploaded in this repository.

**Recommendations**

1. **Be on time** : People don’t like to wait to reach their destinations, or if the flight is delayed, notify the passengers before they reach the airport.
2. **Focus on treatment** : People don’t really remember how they are fed, rather how they are treated. Be kind, respectful, and fair.
3. **Improve customer service** : Train staff to be more patient and friendly.
4. **Fix website issues** : Make sure the booking system works smoothly.
5. **Ask for feedback** : Encourage passengers to fill out satisfaction surveys.

**Limitations**

1. The data was limited, each airline had only 100 reviews.
2. There were a lot of missing values.