July
2023

MAERSK CASE COMPETITION

SMU DSA X MAERSK

TFCC:
FOO CHUAN WEI
SHARAFINAZ BINTE SHAWAL
NGUYEN HANH TRANG
TOH JING LIN CHERYL

Table of Contents

I. Data Visualization and Analysis with Python

Overall, there is an equal representation of both genders (Figure 1), in which female makes up roughly half of the reviewers. The majority are loyal customers, and more customers travel for business reasons than personal ones.

	Gender	Customer Type	Type of Travel	Class
count	129880	129880	129880	129880
unique	2	2	2	3
top	Female	Loyal Customer	Business travel	Business
freq	65899	106100	89693	62160

Figure 1: Overview of categorical variables

Age is categorized into 5 groups and the majority of travelers are adults aged 20 to 59.

Age range	Age Category
5 to 12	Child
13 to 19	Teenager
20 to 39	Adult
40 to 59	Middle Age
60+	Senior

```
Middle Age    53602
Adult         51883
Senior        12215
Teenager       7386
Child          4794
Name: Age Group, dtype: int64
```

Figure 2: Overview of Age Groups

We investigate each Customer Type to understand what they value in their experience. Airline Passenger Satisfaction.csv is used to create stacked bar chart with Python Matplotlib. Analysis of Customer Type's ratings against the 14 factors helps identify those that may directly influence one's satisfaction or dissatisfaction. (Refer to satisfaction_reviews.ipynb)

From Figure 2, Baggage handling is a key driver that leads to customer satisfaction, with positive ratings (4-5) taking up 60% of the reviews. In contrast, Departure/Arrival time convenient, which we understand as flight delay time, causes customers the most dissatisfaction, with negative ratings (0-2) taking up ~40%. More specifically, 0 rating makes up 20% of the reviews of disloyal customers.

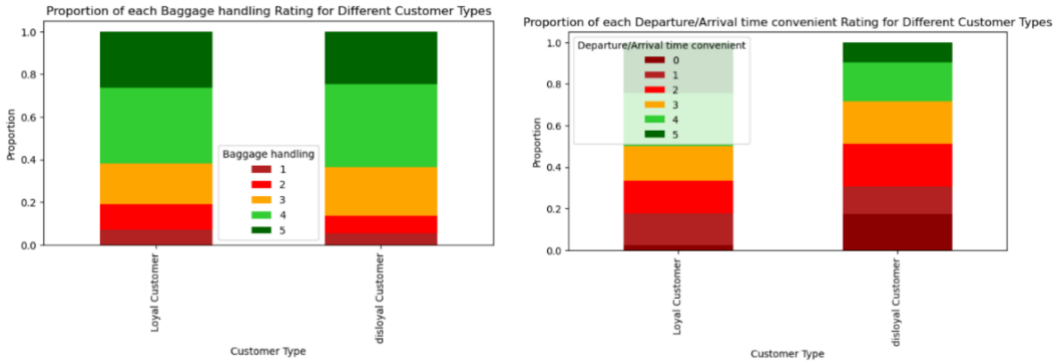

Figure 3: Customer Type rating for Baggage handling (Left) & Departure/Arrival time convenient (Right)

Comparing with the key drivers for Type of travel and Age Group (refer to Appendix A), we observe that the common factors are **Baggage handling, Departure/Arrival time convenience and Ease of Online booking.**

The visualization also indicates that most disloyal customers are business travelers (Figure 3). This could be due to how business travel can be unplanned and whichever airline fits the schedule would work.

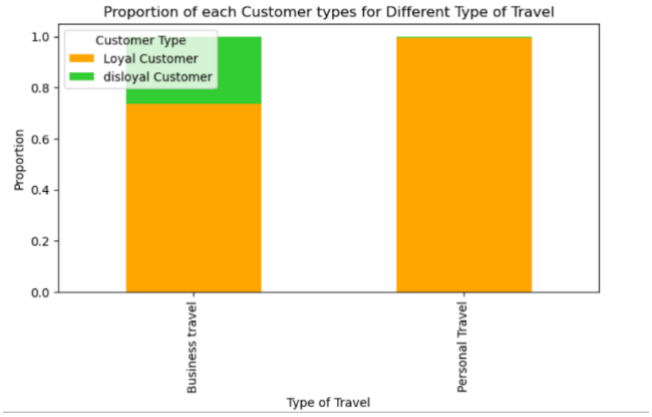

Figure 4: Proportion of Loyal vs Disloyal customer by Type of Travel (Left)

Limitation of Data Visualization

While the visuals and charts help provide a quick overview of the data, we can only show which are the possible drivers contributing to customer experience. There may be other factors contributing to the overall satisfaction and customer loyalty, but they are not fully represented by the data. Additionally, it is difficult to find correlations between factors or between different categories.

As a result, we decided to use statistical testing to further analyze our data sets.

II. Key Drivers of Satisfaction

To identify the key drivers of satisfaction, we used multiple statistical tests to bring about focus to core insights for such data results with further research.

1. Statistical Testing

The statistical testing involves 2 parts; a two-sided Welch t-test and a post-hoc test called Cohen's d. For output refer to (Test_Statistics_Outcomes.xlsx) for the specific outcomes of the test statistics.For input refer to the folder (Code for T_Tests&Cohen) and lastly do note our datasource of the code would be the wrangled Airline Passenger Satisfaction.csv in our zip folder.

The selection of this test statistic comes from the fact that the given dependent variables (the different 14 satisfaction ratings from 1-5 with 3 being treated as neutral) are that of a Likert scale. There are two schools of thought but in this case, we will be treating the dependent variables to be an interval variable and that numeric mean difference are interpretable and hence a t-test is conducted at the 95% level. Thereafter a statistically significant result will be parsed through Cohen’s d to see how significant the difference is. We followed Cohen’s d boundaries where 0-0.2 is negligible practical significance,0.2-0.5 is small, 0.5-0.8 is moderate and more than that is large. We will be sticking to them and not altering the boundaries due to lack of prior contextual knowledge. The t-test was conducted between satisfied and (Neutral + Dissatisfied) and that is further broken down to satisfied vs neutral & satisfied vs dissatisfied. There after we also looked at comparisons between loyal and disloyal customers, and the post hoc Cohen’s d was implemented for all regardless of outcome for simplicity in the code. To give an example of our implementation, from our results we can see that the comparison of mean ratings of Onboard service between satisfied and neutral is statistically significant and that difference is 0.0269. However, Cohen's d is below 0.2 and hence considered the outcome of the difference of mean to be negligible practically significant. With that we can omit the consideration of onboard service as a key driver between satisfied and neutral. Hence to sum it up, t-test tells us whether there is sufficient evidence to conclude that there is a

difference between two groups for the dependent variable in question and Cohen's d then tell us whether that difference is big enough to make practical sense or not.

2. Correlation Matrix

This statistical test helps us to understand better the relationships between the variables with a coefficient value between –1 to +1. A positive value indicates that the two variables have a positive relationship while a negative value indicates that two variables have a negative relationship.

Test Statistics Outcome – Satisfaction

Comparing the Satisfied vs Neutral, we can derive that Online Boarding with the value of 0.745153, brings about higher satisfaction that has moderate practical significance. This potentially signifies Online Boarding as one of the factors. This is also supported if we compare Satisfied vs Dissatisfied, where the value is 1.517265 for Online Boarding with a large practical significance. (Refer to T-Test Sheet in Test_Statistics_Outcomes.xlsx)

We can also view that inflight entertainment could be a potential driver of satisfaction where Cohen's result is 0.8873586 and the effect size has a large practical significance. This occurs when we compare Satisfied VS Neutral + Dissatisfied. (Refer to T-Test Sheet in Test_Statistics_Outcomes.xlsx)

To further support this, from the Correlation Matrix, the same factors Online Boarding (0.0014465) has a strong positive correlation with satisfaction score, which further strengthens its stand as the driver for Satisfaction. (Refer to Correlation Matrix Satisfaction Sheet in Test_Statistics_Outcomes.xlsx)

III. Key Drivers of Loyalty

For this section, our team has decided to look at the loyalty score which was to tabulate the number of loyal and disloyal customers giving their satisfaction score for each variable. We also considered looking at the correlation matrix of loyalty.

Test Statistics Outcome – Customer Loyalty

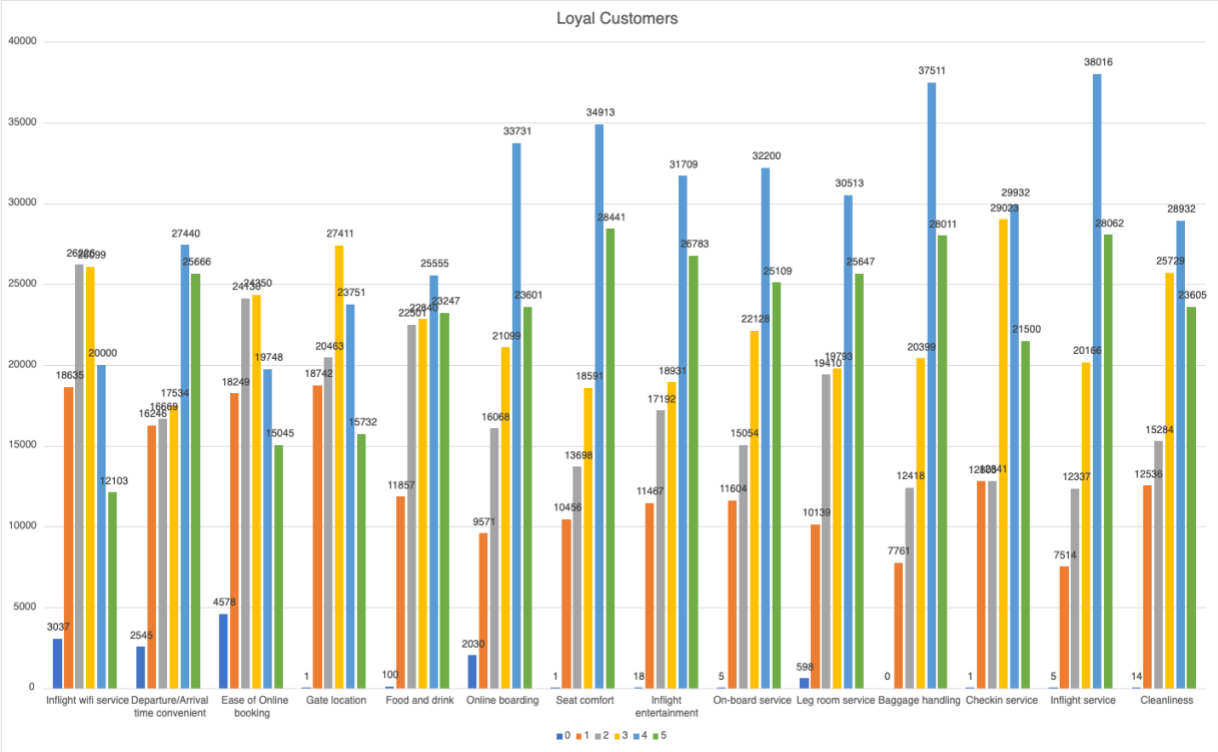

Figure 5. Loyal Customers Score

From Figure 5 we see that 61.75% and 62.3% of loyal customers gave a rating of 4 or 5 to baggage handling and inflight service respectively. We initially believe these could be potential drivers for loyalty. However, when comparing Figure 6, 63.5% and 64.3% of disloyal customers gave a rating of 4 or 5 to baggage handling and inflight services respectively.

This shows that despite the high ratings, they are not key drivers for customers to be loyal.

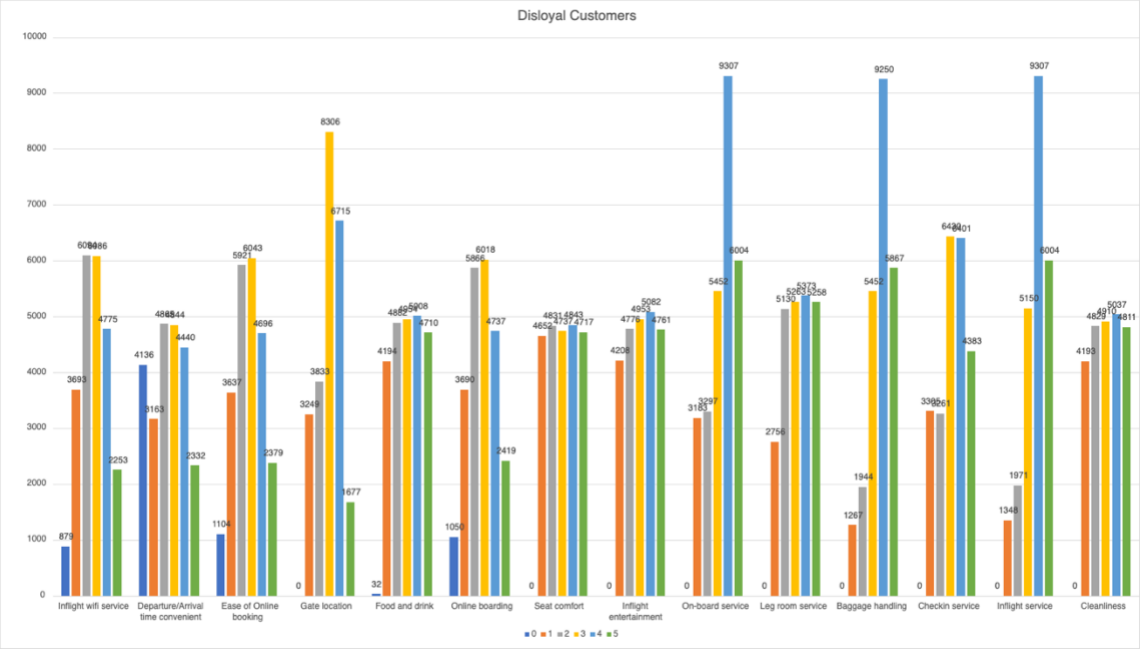

Figure 6. Disloyal Customers Score

We then found factors such as Online Boarding, Seat Comfort, In-flight entertainment, and cleanliness brought in high positive satisfaction from loyal customers and low positive satisfaction from disloyal customers. (Refer to Loyalty Score Sheet in Test_Statistics_Outcomes.xlsx)

From the correlation matrix, we found that factors such as Online Boarding, Seat Comfort, In-flight entertainment, and cleanliness brought in high positive correlation values with satisfaction. This shows that such factors are the main drivers of loyalty. (Refer to Correlation Matrix Loyalty Sheet in Test_Statistics_Outcomes.xlsx)

IV. Natural Language Processing (NLP) Keyword Extraction Model

The Voice of Customer dataset is placed into a keyword extraction model using NLP in python. The objective is to obtain several important keywords from the reviews given by each customer by tokenization and stop words. Afterwards, a new function is created to obtain the distinct words from the voice of customer column by removing the stop words and join via a string of commas. By using the keywords, we can group the reviews into different categories, i.e., baggage handling and identify what are the top service

categories the airlines are lacking in. This process is done through Excel and Python software, refer to keywords_extraction.xlsx and reviews.ipynb.

Category	Count
Arrival / departure delay	377
Gate location	18
Food and Drinks	18
Seat & legroom, storage comfort	81
Inflight entertainment	34
On-board/ inflight service	211
Baggage handling	120
Cleanliness	69
Nil	12770
Others	1275
Total	14973

Table 1. Total number of customer reviews in each category.

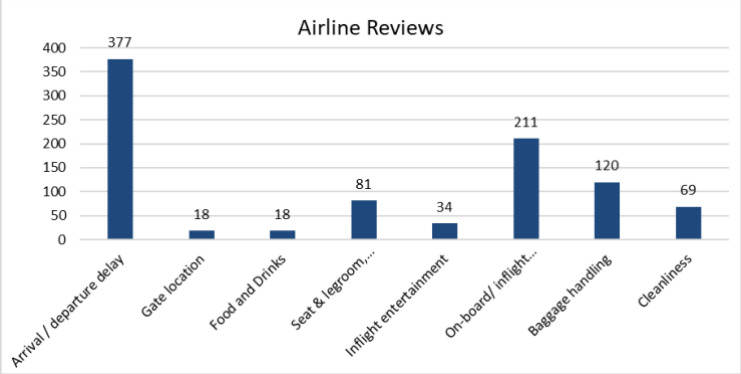

Figure 7. Total counts for airline categories based on customer reviews.

The top 3 categories are delay in arrival or departure, in-flight services, and baggage handling issues. The airline can focus on these 3 aspects based on feedback to improve their passenger service.

Limitation of the NLP model

As observed the dataset consists of 14,933 entries, the model obtains 14,973 entries after classification. As the reviews are open-ended, some words are non-unique i.e., delay, customers used delay for arrival time and baggage. Hence, there are double counts for 2 categories or more. A suggestion would be to provide fixed category names for selection and customers can input their comments. With the fixed category, we can better identify the aspects of customer service that are lacking.

V. Recommendation

While comparing commercial airline and airfreight cargo, we observed a similarity in their main goal, which is to transport passengers/items of interest to the desired destination while maintaining safety. In addition, they are similar in terms of offering quality service, like ensuring the passengers' needs are met, or making sure that cargo items are properly packed, loaded and delivered undamaged. However, commercial flights are prearranged trips in which passengers often make bookings in advance. In contrast, cargo flight bookings can be made at short notice, normally 2-3 days beforehand.

With the goal of improving customer service, our team would focus on the **Inflight service**, and **Arrival/Departure delay variables** and suggest 2 improvements to Maersk Air Freight.

1. 24/7 MAERSK Chatbot

While the Maersk website provides live chat with agents, the service is only available during office hours. There may be urgent times when agents are unavailable, questions come in after office hours or clients have problems finding information. Often, the problem lies in delays of transportation and customers do not know how to check for status.

We recommend creating a 24/7 chatbot, in which customers can input their questions and the bot will show possible matching FAQs with high accuracy. Then, they will be directed to the corresponding website with instructions on how to check their cargo's information. If their questions are not found in FAQs, chatbot will automatically draft an inquiry that customers can edit accordingly and be sent immediately.

2. POV Video

Another recommendation we would like to propose is to include a POV Video that displays the behind-the-scenes of the cargo airlines, i.e., the handling of the cargo in terms of the packaging to protect it from any damage. From our findings, we believe that customers would have uncertainty and would seize for knowledge in understanding how their items are being packed and transported.

Such a video would help in providing assurance and keeping it transparent between the company and customers. It can also help provide clear guidelines and a description of how cargo airlines work. This can aid in providing greater satisfaction. The mood board below helps to visualize better on what possible contents can be covered.

Figure 8. Possible information to include in the video

Appendix

Appendix A: Python Data Visualization with Stacked Bar Chart

For Type of Travel, Baggage handling is a key driver for customer satisfaction with 60% positive reviews while Departure/Arrival time convenient contributes to dissatisfaction of service, especially for business travelers where negative reviews taking up to 40%

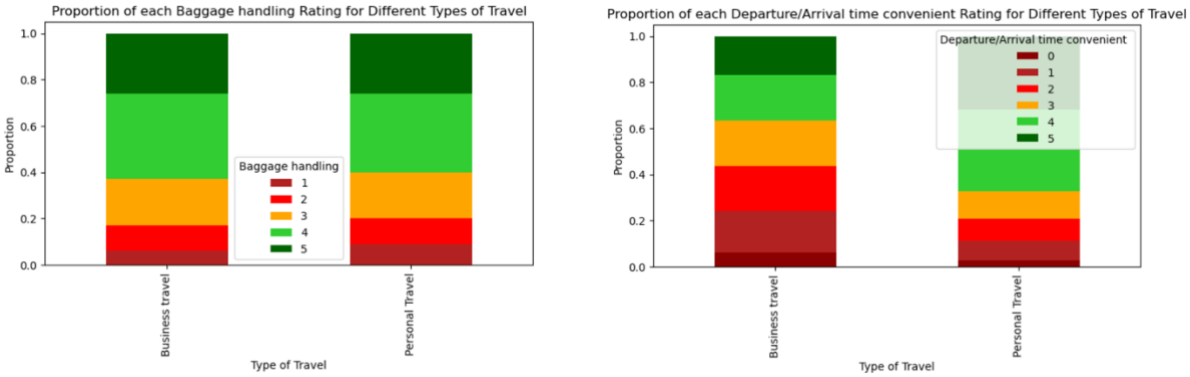

Figure 1: Type of Travel rating for Baggage handling (Left) & Departure/Arrival time convenient (Right)

Baggage handling remains a positively rated driver and the older the passenger, the more they value baggage services. On the other hand, Ease of Online booking's negative ratings are up to 50% of the reviews across the different groups, indicating an important factor.

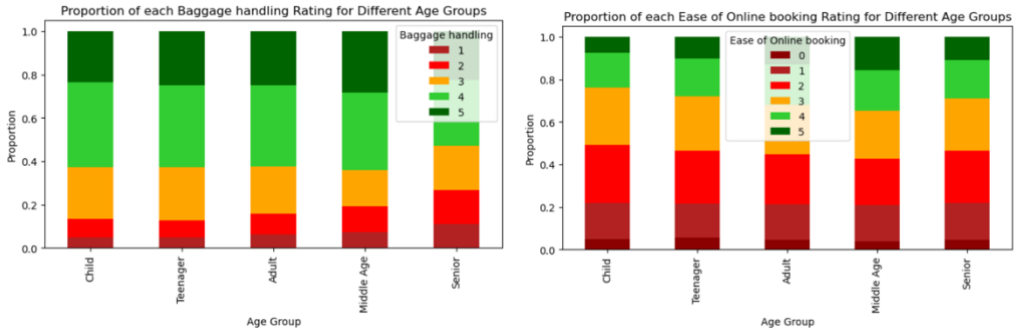

Figure 2: Age Group rating for Baggage handling (Left) & Ease of Online booking (Right)