OntarioTech UNIVERSITY

MBAI 5300G

MBAI 5300G Programming and Data Processing Final Project

SKY-LEVEL SATISFACTION

ANALYZING KEY DRIVERS OF AIRLINE PASSENGER SATISFACTION

John de Guzman, Elaine Bhel Lagman, Renee Rae Santos, Hazal Turken

Abstract

This study analyzes the key drivers of airline passenger satisfaction using a dataset of 103,904 survey responses covering demographic attributes, service quality ratings, and operational performance indicators. After data cleaning, feature engineering, and exploratory analysis, satisfaction patterns were examined across customer types, travel classes, age groups, and delay categories. Results show that service experience is the strongest determinant of satisfaction, with online boarding, in-flight entertainment, seat comfort, and class of travel emerging as the most influential positive predictors. Operational factors such as departure and arrival delays significantly increase dissatisfaction, particularly when delays exceed 30 minutes. Business-class passengers and loyal customers report consistently higher satisfaction, while economy-class passengers and first-time travelers show lower ratings. Feature-importance analysis confirms that digital convenience and cabin comfort drive satisfaction more strongly than demographic variables. Overall, the findings highlight that improving digital processes, enhancing onboard comfort, and reducing schedule disruptions are the most effective pathways for airlines seeking to elevate customer satisfaction.

Keywords: Passenger Satisfaction; Service Quality; Feature Importance; Delay Analysis; Cabin Class; Customer Behavior

1. Introduction

The airline industry operates in a highly competitive, service-driven environment where passenger satisfaction is central to customer loyalty, operational strategy, and long-term performance. Understanding what drives satisfaction has therefore become a key priority for carriers seeking to improve service quality, reduce disruptions, and enhance the overall travel experience. By examining factors such as service interactions, cabin comfort, and operational performance, airlines can identify where improvements matter most and use those insights to guide targeted service enhancements and long-term planning.

The Canadian airline industry generated an estimated CAD 23.8 billion in revenue in 2023, underscoring both its economic significance and the intensity of domestic competition (Marketline, 2024). In this environment, low-cost and smaller carriers such as Porter and Flair have begun to capture a meaningful share of capacity from the traditional "big two" of Air Canada and WestJet, indicating that market power is no longer concentrated solely among the largest incumbents (CBC News, 2024). As these challengers succeed in attracting passengers even when competing on similar routes and price points, non-price factors—particularly airline passenger satisfaction—emerge as a critical source of competitive advantage, enabling airlines to differentiate their offerings and build loyalty in an increasingly contested market (CBC News, 2024).

This data analytics project aims to identify the most influential factors affecting airline passenger satisfaction by analyzing survey responses that capture customer demographics, service ratings, cabin experience, and operational metrics such as delays. The scope of the study includes data cleaning, feature engineering, exploratory data analysis, and feature-importance modeling to uncover meaningful patterns across different customer groups and travel classes.

The primary objectives of the report are:

- To examine how passenger characteristics and travel choices relate to satisfaction levels.
- To determine which service features most strongly influence satisfaction through exploratory data analysis
- To provide evidence-based insights and recommendations that airlines can use to improve service quality and customer experience.

By systematically analyzing both behavioral and operational dimensions, this project seeks to offer a comprehensive understanding of the key drivers shaping airline passenger satisfaction.

2. Literature Review

Airline passenger satisfaction has been studied for years because of its strong influences on airline competitiveness, market performance, and customer loyalty. Understanding how passengers judge their experience gives them an idea on their overall performance and helps them determine which areas to retain, improve, or discontinue. In this project, the researchers adopt a data-driven approach to explore passenger characteristics, operational factors, and service features to identify the strongest predictors of satisfaction. The review below summarizes recent studies that guided the analytical decisions made in this work.

Passenger segmentation is a key part of understanding how different groups evaluate airline services. Clemes et al. identified seven key service quality dimensions (timeliness, assurance, convenience, helpfulness, comfort, meals, and safety). In this study, they found out that satisfaction varies by demographic factors such as age, gender, income, occupation, and marital status [1]. This suggests that passengers evaluate their experiences differently based on who they are and why they travel.

Another major factor in how passengers assess their flight experience is operational reliability. Studies in China show that flight delays significantly reduce satisfaction and loyalty [2], and consumer complaint data from the Civil Aviation Administration of China show that delays remain a leading issue for passengers [2]. Song et al. highlight that passengers make several travel decisions during delays, thus, making them satisfied through delay-related services is a strategic move for airlines when delay happens [2]. To add, Oliveira et al. mentioned that while delays generally decrease satisfaction, positive experiences with airport Wi-Fi and food and beverage services can lessen this effect [3]. These findings support the inclusion of delay metrics and service features in this study.

Lastly, service features, especially digital and in-flight services, are also critical determinants of satisfaction. Service quality remains central in airline research, especially because airlines try to differentiate themselves in a competitive market through the service they offer [4]. Many of them have adopted digital tools such as self-service options and contactless processes to improve customer experience and streamline processes through efficiency. Studies show that these digital features, including Wi-Fi and online boarding are becoming key drivers of satisfaction. Also, with the use of data mining methods, Noviantoro and Huang found that online boarding, in-flight Wi-Fi, baggage handling, and in-flight entertainment were among the strongest contributors to positive evaluations [5]. These findings are aligned well with Du, who used a neural network model and found that Wi-Fi access and online boarding pass printing have high predictive importance. Du also notes that those flying in business class which receive a more premium service have strongly improved satisfaction [6]. Meanwhile, personal or leisure travelers tend to respond well to promotions and small service upgrades, and improving services in economy class can reduce negative experiences and encourage passengers to travel with the same airline again.

As for analyzing the importance, effect size guidelines were used. Originally, Cohen described r =.10,.30 and .50 as small, medium and large [7]. However, Hemphill revealed that above .50 correlations are uncommon, therefore, suggesting to treat > .30 as large effect [7]. Meanwhile for the mean satisfaction scores of a five-point scale, service quality studies in airlines showed that mean values of above 3 were considered satisfactory [8].

Overall, recent studies show that combined influence of passenger characteristics, operational reliability and service features work together in shaping or determining passenger satisfaction. These findings align with the methodological decisions in this project and help explain the difference observed across passenger groups in the dataset.

3. The Dataset Summary

The present analysis draws on the publicly available Airline Passenger Satisfaction dataset hosted on Kaggle (Teejmahal, 2020). This dataset comprises 103,904 individual survey responses and 25 variables capturing both customer segmentation characteristics and ratings of multiple service dimensions. Service quality is assessed using a Likert-type scale ranging from 0 to 5, where a score of 5 denotes the highest level of satisfaction.

3.1. Data Cleaning and Basic Inspection

As part of the data cleaning and basic inspection procedures, several preprocessing steps were undertaken to ensure data quality and consistency. Missing values in the arrival delay variable were imputed using the median, and an unnamed column containing only index values, which was not needed for the analysis, was removed from the dataset. All survey records were verified to be unique, ensuring that no duplicate responses were included, and the satisfaction outcome was recoded so that passengers who reported being neutral or dissatisfied were collectively categorized as dissatisfied.

During the outlier inspection stage, only flight distance, check-in service rating, departure delay, and arrival delay displayed values that were visually distant from the bulk of their respective distributions. The distributions of these variables were highly skewed, which caused a number of legitimate extreme values to appear as potential outliers. Because these observations plausibly represent genuine long-haul flights and substantial delays, they were retained in the dataset rather than being removed. Including such outliers preserves the full range of operational conditions captured by the survey and helps avoid biasing the analysis toward more typical flight experiences.

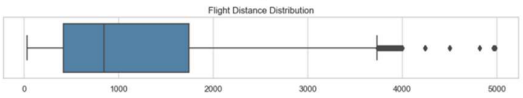

Figure 3.1 1 Flight Distance Box plot

Figure 3.1 1 Check-In Service Box plot

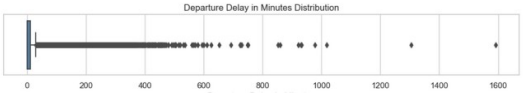

Figure 3.1 3 Departure Delay Box plot

Figure 3.1 2 Arrival Delay Box plot

3.2. Feature Engineering

For feature engineering, key variables were recoded into discrete segments to enhance interpretability and reduce the dimensionality of the predictors. Age was grouped into ordered life-stage categories (0–18 years as youth, 19–34 as young adults, 30–45 as adults, 46–60 as mature adults, and above 60 as seniors). Flight distance and delay in minutes were similarly binned into short, medium, long, and very long ranges, and travel class was treated as an ordered categorical variable (Economy, Eco Plus, Business). This segmentation approach allowed the analysis to focus on substantively meaningful groups rather than numerous individual values, many of which would have offered limited additional insight.

Table 1 Variable Categories for Segmentation

Variable	**Category label**	**Definition (range or level)**
Age	Youth	0–18 years
	Young adults	19–30 years
	Adults	31–45 years
	Mature adults	46–60 years
	Seniors	Above 60 years
Distance	Short	Short haul (< 500 km)
	Medium	Medium haul (500–1500 km)
	Long	Long haul (1501–3000 km)
	Very long	Very long haul (> 3000 km)
Delay (minutes)	Short delay	No or short delay (< 30 min)
	Medium delay	31–60 minutes
	Long delay	1–2 hours
	Very long delay	More than 2 hours
Class	Economy	Standard economy cabin
	Eco Plus	Enhanced economy cabin
	Business	Business-class cabin

4. Data Analysis

The data analysis part of this project is subdivided into 4 major categories, namely, data distribution, passenger characteristics, operational factors, and feature importance. These steps play an important role in understanding the underlying patterns that may influence passenger satisfaction. A comprehensive understanding of how travelers differ in demographics, travel behavior, and flight experiences can be done through examining both numerical and categorical features. This analysis serves as a foundation for identifying which factors meaningfully affect or shape customer satisfaction.

4.1. Data Distribution

The first part of the analysis is the data distribution, which shows the distribution of basic numerical variables in the dataset alongside the target variable. This step helps in providing an initial overview of the data's structure and offers a glimpse into the underlying patterns present.

A. Numerical Distribution Analysis

This section examines the dataset's numerical variables such as age, flight distance, and the different measures of delay.

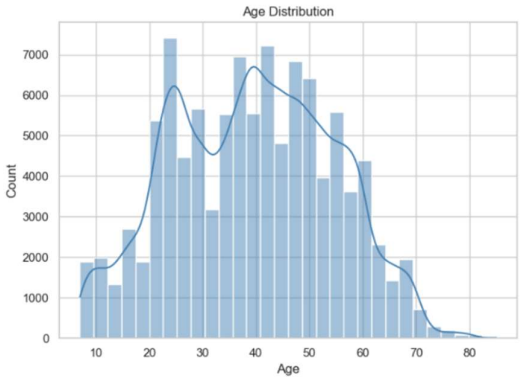

Figure 4.1 2 Age Distribution

As shown in Figure 4.1 1, the age distribution approximates a normal distribution centered on working-age adults, with the majority of passengers aged 25-50. This pattern indicates that working adults dominate the airline passengers, with few numbers in the younger (0-18) and order (60+) age groups. It is slightly skewed to the right, which reflects that as age increases, there is a gradual decline in the number of passengers.

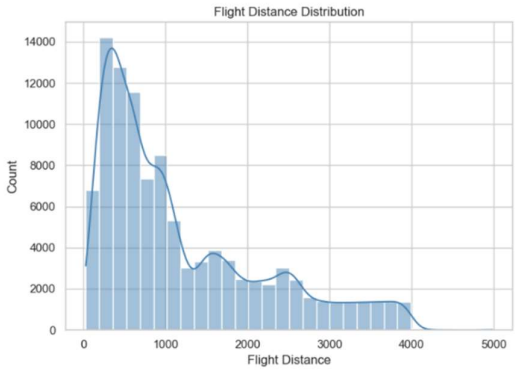

Figure 4.1 1 Distribution of Flight Distances

Figure 4.1 2 shows that most flights fall within short and medium distances, with relatively few long-haul trips. The majority of distances are concentrated below 1500 kilometers, resulting to the right-skewed pattern. This is primarily due to a small number of long-distance outliers that extend the upper tail, highlighting the imbalance between the large volume of shorter routes and the relatively limited number of long distance flights.

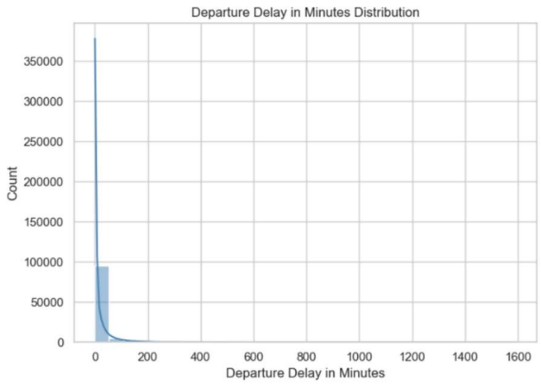

Figure 4.1 3 Delay Distribution for airline passengers: Departure Delay in Minutes

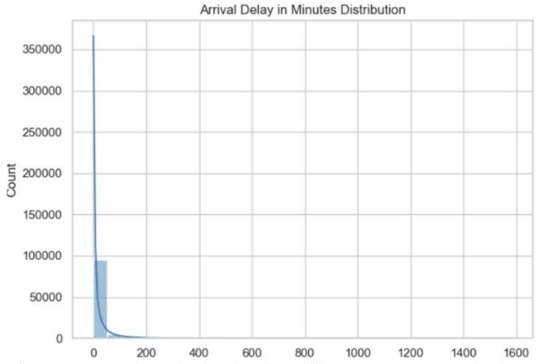

Figure 4.1 4 Delay Distribution for airline passengers: Arrival Delay in Minutes

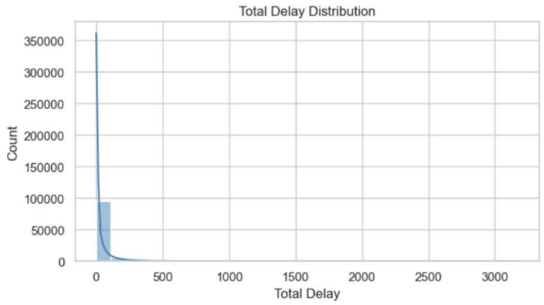

Figure 4.1 5 Delay Distribution for airline passengers: Total Delay in Minutes

It can be seen in Figure 4.1 3-5, that delay metrics, including departure, arrival, and total delays, are heavily concentrated near zero, indicating that the majority of the flights only experiences minor disruptions, while long delays are infrequent, they can reach substantial magnitudes. This small number of extreme delay cases stretched the tail to the extreme right.

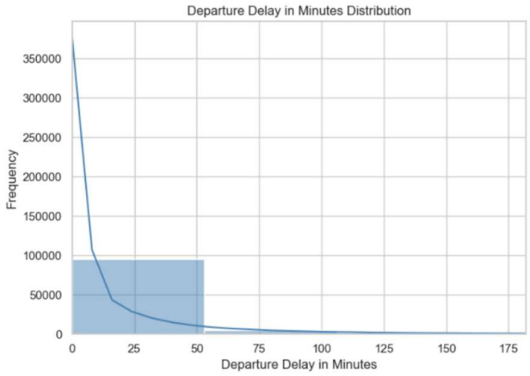

Figure 4.1 6 Zoomed-In Delay Distribution: Departure Delay

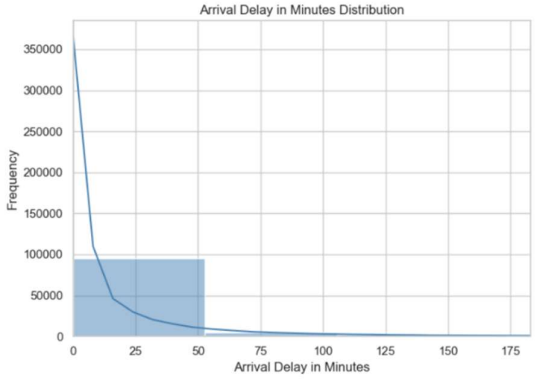

Figure 4.1 7 Zoomed-In Delay Distribution: Arrival Delay

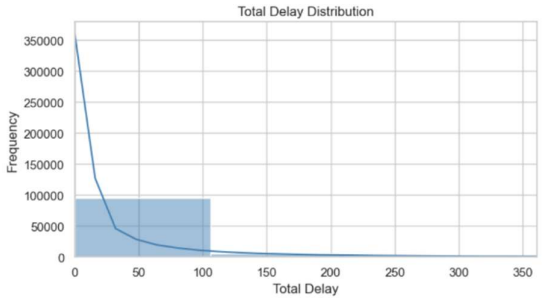

Figure 4.1 8 Zoomed-In Delay Distribution: Total Delay

Figure 4.1 6-8 provides a closer look at the central range of delay values, highlighting that most delays fall within the first 50 minutes, indicating that operational disruptions are generally minor, with only occasional severe delays impacting the passenger experience. Overall, the frequency dropped sharply as delay duration increases.

B. Target Variable (Satisfaction) Analysis

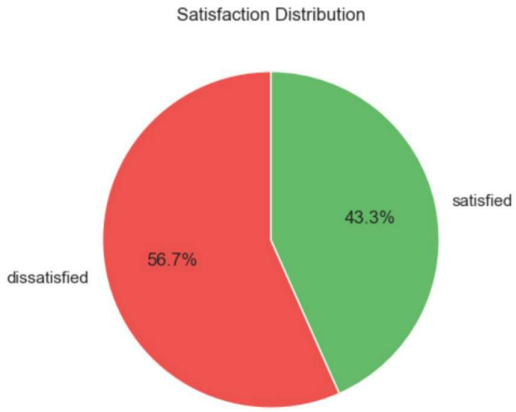

Figure 4.1 9 Distribution of satisfaction levels among Passengers

It can be clearly seen from Figure 4.1 9 that the passenger satisfaction demonstrates an uneven distribution, with 56.7% reporting dissatisfaction, and 43.3% reporting satisfaction. While the difference is noticeable, it does not represent an extreme imbalance. This suggests that a portion of passengers experience unmet expectations or service inconsistencies, as reflected in the relatively close distribution between satisfied and dissatisfied passengers.

4.2. Passenger Characteristics Analysis

The second part of the analysis examines how passenger attributes relate to satisfaction outcomes. By comparing and analyzing demographic and behavioral characteristics, patterns can be identified that can help explain differences in how passengers differ in judging their flight experiences. This section

is divided into 3 parts: core characteristics, combined characteristics, and age-related characteristics.

A. Satisfaction by Core Characteristics

This subsection evaluates satisfaction across passenger core characteristics, including gender, customer type, travel purpose, and cabin class. These variables represent the primary ways passengers differ within the dataset. Thus, comparing satisfaction across these groups provides insights on how various travelers judge their flight experience and helps determine which specific characteristics are associated with higher or lower satisfaction.

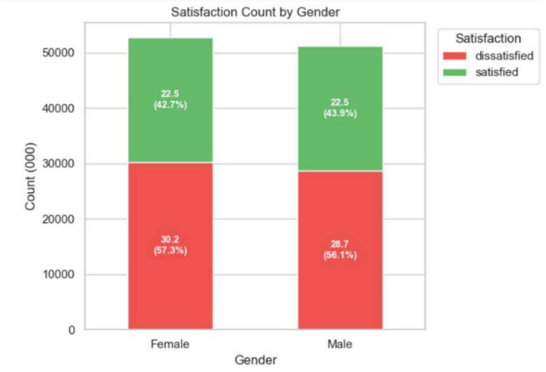

Figure 4.2 1 Distribution of Passenger Satisfaction by Gender

Figure 4.2 1 shows a minimal difference in the satisfaction between female and male passengers. Both groups show a higher proportion of dissatisfaction, suggesting that service issues affect all genders similarly and that gender does not substantially influence passenger satisfaction in this dataset.

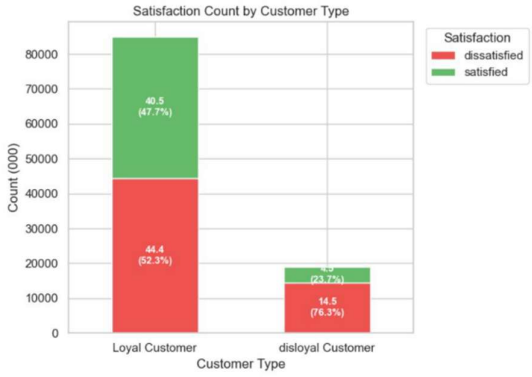

Figure 4.2 2 Distribution of passenger satisfaction by Customer Type

As depicted in Figure 4.2 2, there is a clear difference in satisfaction between loyal and disloyal travelers. Loyal customers have a more balanced distribution (52.3% dissatisfied, 47.7% satisfied), while disloyal customers report much lower satisfaction (76.3% dissatisfied, 23.7% satisfied). Although loyalty status is linked to better outcomes, even loyal customers show high dissatisfaction, suggesting that their expectations are not being met and experiencing service inconsistency despite continued patronage. On the other hand, disloyal ones experience even greater dissatisfaction, indicating weaker experiences for first-time or occasional passengers.

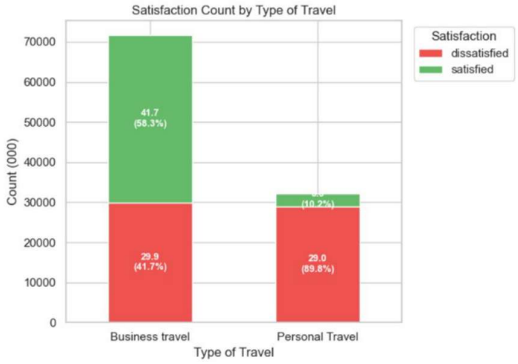

Figure 4.2 3 Distribution of passenger satisfaction by Travel Type

The chart in Figure 4.2 3 compares satisfaction levels between business and personal travelers and illustrates that business travelers report higher satisfaction than personal travelers. This indicates that travel purpose is closely associated with satisfaction, likely because airline services are more aligned with the needs and expectations of business passengers.

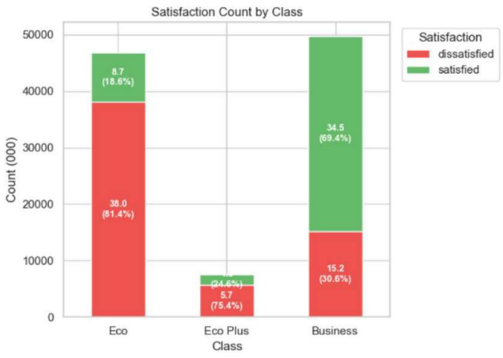

Figure 4.2 4 Distribution of passenger satisfaction by Cabin Class

It clearly shows in Figure 4.2 4 that there is a substantial difference in satisfaction across cabin classes. Business class passengers report the highest satisfaction, while Eco and Eco Plus' satisfaction decreased notably, with Eco being the lowest. With this, it can be said that cabin class plays an essential role in shaping satisfaction, with higher classes

associated with more positive passenger experience. In short, the result points to a clear gap in service quality across cabin types with more premium cabins receiving a more premium service.

B. Combined Characteristics Analysis

After analyzing satisfaction in relation to individual variables, the study dived deeper by combining passenger characteristics to identify potential interaction effects. In this case, customer type and cabin class, and travel purpose and cabin class. This approach helps determine whether certain groups consistently benefit from higher cabin classes.

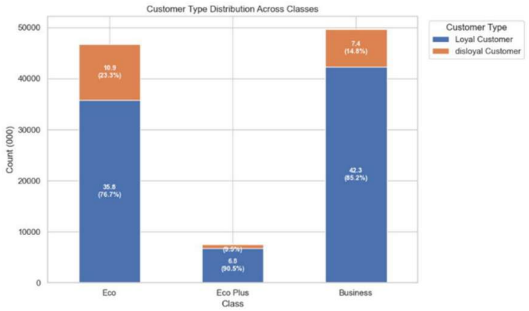

Figure 4.2 5 Customer Characteristics by Cabin Classes: Distribution of Customer Types

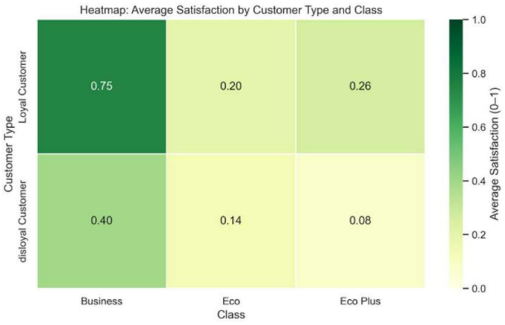

Figure 4.2 6 Customer Characteristics by Cabin Classes: Average Satisfaction by Customer Type and Class

The 2 charts in Figure 4.2 5-6 examine the relationship between customer type and cabin class and how these 2 factors affect the passenger's satisfaction. Figure 4.2 5 shows that the majority of travelers in each class are loyal customers, particularly in Business Class. Meanwhile, disloyal passengers are most concentrated in economy class, followed by business, then eco plus.

Figure 4.2 6 presents the average satisfaction scores by customer type and class. Here, a clear pattern emerges with loyal customers achieving their highest satisfaction levels in Business Class, demonstrating that premium cabins deliver stronger service experiences. In contrast, disloyal customers flying Eco show the lowest satisfaction overall. Satisfaction declines for both loyal and disloyal passengers in Eco and Eco Plus, showing that cabin class has a stronger influence on satisfaction than loyalty status. While loyal customers generally evaluate their experience more positively, their loyalty alone cannot offset the lower satisfaction associated with flying lower cabin classes.

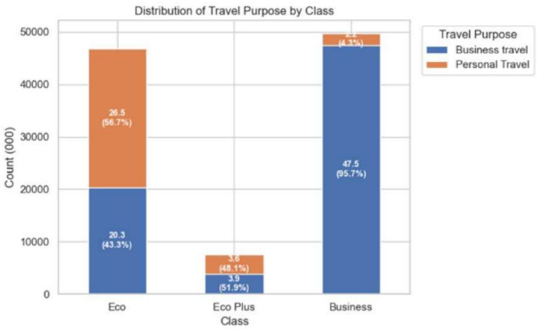

Figure 4.2 7 Travel Purpose characteristics across Cabin Classes: Distribution of Travel Purposes

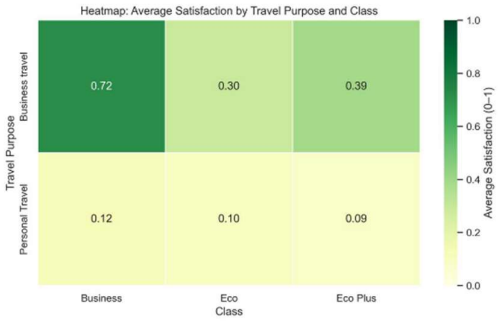

Figure 4.2 8 Travel Purpose Characteristics across Cabin Classes: Average Satisfaction by Travel Purpose and Class

As shown in Figure 4.2 7-8, the relationship between travel purpose and cabin class indicates that travelers predominantly select Business Class for work-related reasons, whereas Economy is chosen for both personal and some business travel. Eco Plus is infrequently used regardless of travel purpose, indicating that the middle-tier class may not offer sufficient value or appeal to most passengers. Figure 4.2 8 shows that business travellers report higher satisfaction in all classes, with the highest levels in Business Class. This pattern suggests that cabin class has a stronger influence on satisfaction than travel purpose.

C. Age Group Analysis

The third subsection is about how satisfaction varies across age groups and how these patterns interact with cabin class.

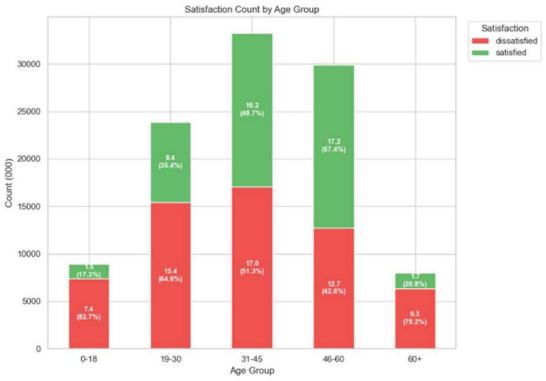

Figure 4.2 9 Distribution of Passenger Satisfaction across Age Groups

It is depicted in Figure 4.2 9 that satisfaction varies significantly across age groups. Passengers aged 31-45 and 46-60 report the highest satisfaction proportions, reaching 48.7% and 57.4%, respectively while younger travelers show notably higher dissatisfaction, and the 60+ group records low satisfaction. This shows that middle-aged passengers tend to have more positive flight experiences, as compared to younger and older passengers.

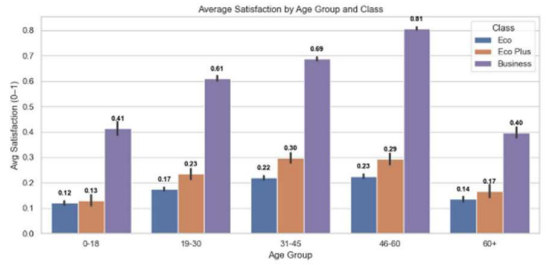

Figure 4.2 10 Age-related satisfaction patterns: Average satisfaction by age group and cabin class

Figure 4.2 11 Age-related satisfaction patterns: heatmap of average satisfaction across age groups and cabin classe

Figure 4.2 10-11 presents that Business Class produces the highest satisfaction across all age groups, with the 46-60 group having the highest ratings. This group also flies business class more often and is one of the main travelers in Eco Plus, alongside 31-45 group.

Across all ages, Eco and Eco Plus yield the lowest satisfaction. Supporting the insight that cabin class has a stronger influence on satisfaction than age, with age-related differences being smaller as compared to the gaps between different cabin classes.

4.3. Operational Performance Analysis

The third part of the data analysis is the operational performance analysis which explores how key-flight related factors, such as flight delay and flight distance, influence customer experience. Assessing these variables help in identifying whether operational disruptions or travel length contribute meaningfully to passenger satisfaction and experience.

A. Flight Delay Impact

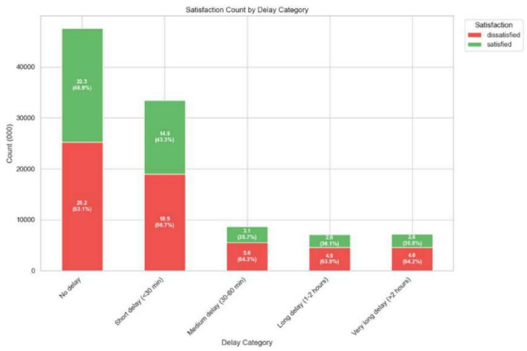

Figure 4.2 12 Delay-related satisfaction patterns: Distribution of satisfaction across delay categories

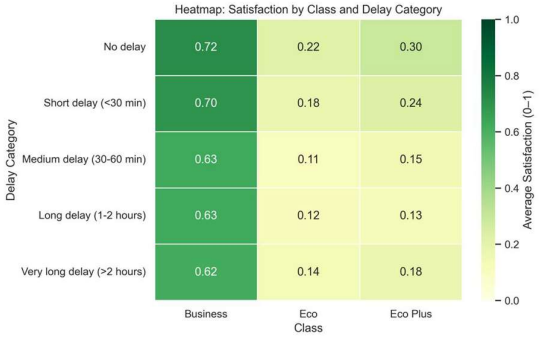

Figure 4.2 13 Delay-related satisfaction patterns: average satisfaction by cabin class and delay category

Figure 4.2 12-13 depicts that the duration of flight delays exhibits a threshold effect on passenger satisfaction. It shows that passenger satisfaction declines sharply once delays reach 30 minutes. Across all delay categories of 30 minutes or more, dissatisfaction rates are estimated to be 64%, establishing this interval as a key operational benchmark. In contrast, flights with no delays have relatively higher satisfaction levels, underscoring

punctuality as a primary expectation for airline passengers and a key determinant of perceived service quality. 4.2 13, the heatmap further illustrates that business class passengers remain the most resilient to delays, although satisfaction still decreases as delay duration increases.

B. Flight Distance Relationship

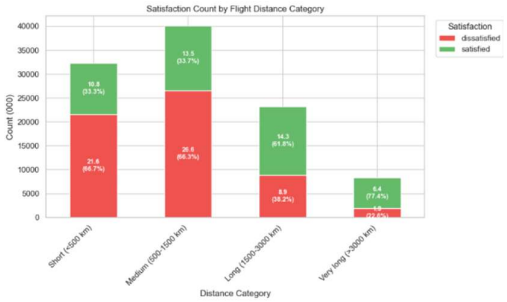

Figure 4.2 14 Flight Distance-related satisfaction patterns: Distribution of satisfaction across flight distance categories

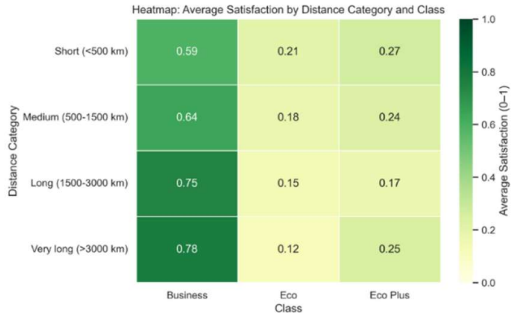

Figure 4.2 15 Flight Distance-related satisfaction patterns: average satisfaction by cabin class and flight distance category

The results in Figure 4.2 14-15 indicate a positive relationship between flight distance and passenger satisfaction. Longer flights consistently receive higher satisfaction ratings, whereas shorter journeys are associated with the highest dissatisfaction rates. This trend may be attributed to the distribution of cabin classes by route length, as longer flights generally include a greater proportion of Business Class passengers, who report higher satisfaction. For Business Class travelers specifically, satisfaction increases with flight distance, indicating that these passengers place particular value on comfort and service quality during extended journeys.

The combined analysis of operational factors indicates that punctuality and journey length interact with cabin class to influence the overall passenger experience. Although airlines cannot manage all operational variables, these findings identify specific performance thresholds and passenger segments where targeted improvements could yield substantial increases in satisfaction.

5. Feature Importance Analysis

In this section, correlation of both categorical and numerical variables is analyzed through feature importance. This gives insights on preferences and and correlation to satisfaction.

Figure 5.0 1 Correlation of service features and passenger characteristics with satisfaction

Feature	Mean
Flight Distance	1189.448375
Age	39.379706
Inflight service	3.640428
Baggage handling	3.631833
Seat comfort	3.439396
On-board service	3.382363
Inflight entertainment	3.358158
Leg room service	3.351055
Checkin service	3.304290
Cleanliness	3.286351
Online boarding	3.250375
Food and drink	3.202129
Departure/Arrival time convenient	3.060296
Gate location	2.976883
Ease of Online booking	2.756901
Inflight wifi service	2.729683
Class	1.028103
Customer Type	0.817322
Type of Travel	0.689627

Figure 5.0 2 Mean ratings of service features and passenger characteristics

As shown in Figure 5.0 1, Online Boarding is the strongest positive driver of satisfaction, indicating that

passengers place a high value on smooth and efficient digital processes. Class and Type of Travel also show strong positive correlations, with higher cabin classes and business travel consistently associated with higher satisfaction.

Figure 5.0 2, instead of showing importance, presents the mean ratings for each service feature. These averages help contextualize the feature-importance results by showing how passengers actually rated each attribute. For instance, many mid-tier service features, such as Inflight Service, Baggage Handling, Seat Comfort, and On-board Service, have moderate mean scores, which aligns with their moderate correlations in Figure 5.0 1.

Gender and Gate Location show both low correlation (Figure 5.0 1) and lower mean importance relative to other features (Figure 5.0 2), indicating they are low-priority improvement areas. Meanwhile, Departure/Arrival Time Convenience exhibits a negative correlation in Figure 5.0 1 but maintains a relatively average mean score in Figure 5.0 2, suggesting that schedule inconvenience affects satisfaction even when passengers rate the service itself neutrally.

5.1. Feature Importance by Variable

A. Customer Features Correlation

As shown in the figure 5.1 1, Class is the strongest customer-related predictor of satisfaction, indicating that higher cabin classes consistently yield better experiences. Type of Travel also has a substantial positive effect, with business travelers reporting higher satisfaction than personal travelers. Flight Distance and Customer Type contribute moderately, while Age has a weak influence. Gender, as seen in the chart, shows almost no meaningful relationship with satisfaction.

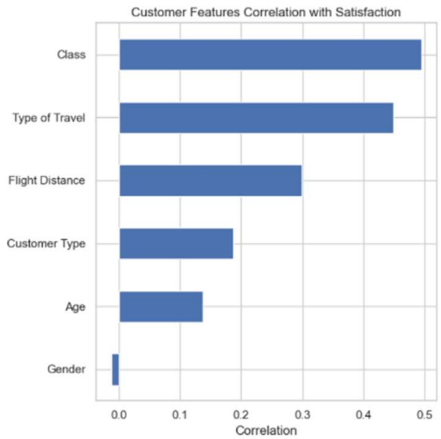

Figure 5.1 1 Correlation of customer-related features with passenger satisfaction

B. Service Features Correlation

As shown in Figure 4.3, service-feature correlations indicate that Online Boarding is the strongest driver of satisfaction, followed by Inflight Entertainment and Seat Comfort, highlighting the importance of digital convenience and onboard comfort. Mid-level factors such as On-board Service, Leg Room Service, and Cleanliness show moderate contributions, while ease of online booking and gate location have minimal influence. Departure and Arrival Time Convenience shows a negative correlation, meaning poorly timed schedules directly decrease satisfaction.

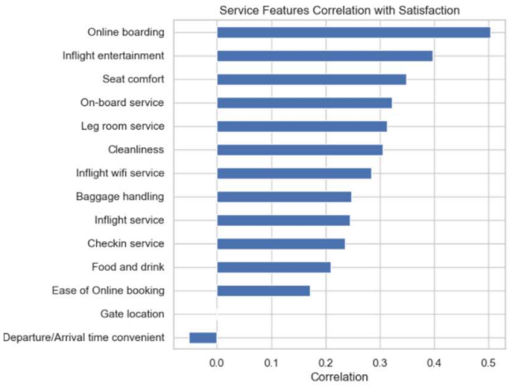

Figure 5.1 2 Correlation of service-related features with passenger satisfaction

5.2. Feature Importance by Class

A. Eco Class Correlation

As shown in Figure 5.2 1, Inflight Wi-Fi is the strongest satisfaction driver for Eco passengers, followed by Online Boarding and Ease of Online Booking, highlighting the importance of digital touchpoints in this class. Inflight Entertainment, Food and Drink, and Cleanliness show moderate influence on satisfaction, while Seat Comfort, On-board Service, and Leg Room Service contribute only modestly. Features such as Baggage Handling, Gate Location, and Inflight Service have minimal impact, and Departure/Arrival Time Convenience displays a slight negative correlation, suggesting that scheduling issues reduce satisfaction even within the Eco class. As shown in Figure 5.2 2, the mean scores for many of these features remain moderate, indicating room for service improvement.

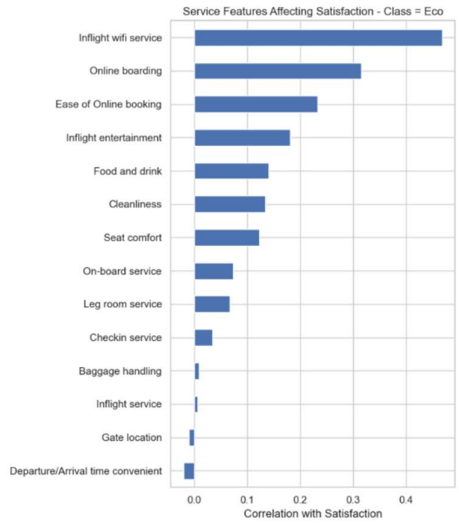

Figure 5.2 1 Correlation of service features with passenger satisfaction in Eco Class

	Eco
Inflight service	3.463921
Baggage handling	3.450551
Departure/Arrival time convenient	3.199123
Seat comfort	3.138838
Checkin service	3.122002
On-board service	3.120355
Cleanliness	3.108097
Inflight entertainment	3.098256
Food and drink	3.086277
Leg room service	3.085720
Gate location	2.971954
Online boarding	2.812985
Inflight wifi service	2.675067
Ease of Online booking	2.605241

Figure 5.2 2 Mean ratings of service features for Eco class

B. Eco Plus Class Correlation

As shown in Figure 5.2 3, in-flight Wi-Fi is the strongest satisfaction driver for Eco Plus passengers, while Figure 5.2 4 shows that its mean score remains relatively low, revealing a clear improvement gap. Online Boarding and Inflight Entertainment also have high importance, and their mean ratings in Figure 5.2 4 are only moderate, suggesting that passengers value these services more than they currently receive. Cleanliness, Food and Drink, and Seat Comfort show mid-level correlations supported by mid-range averages, reflecting an adequate but not exceptional service experience. Gate Location and Baggage Handling have minimal importance and similarly average means in Figure 5.2 4, indicating they are low-impact areas. Finally, Departure/Arrival Time Convenience shows a slight negative correlation, and its mean score in Figure 5.2 4 suggests that schedule-related issues continue to reduce satisfaction for Eco Plus travelers.

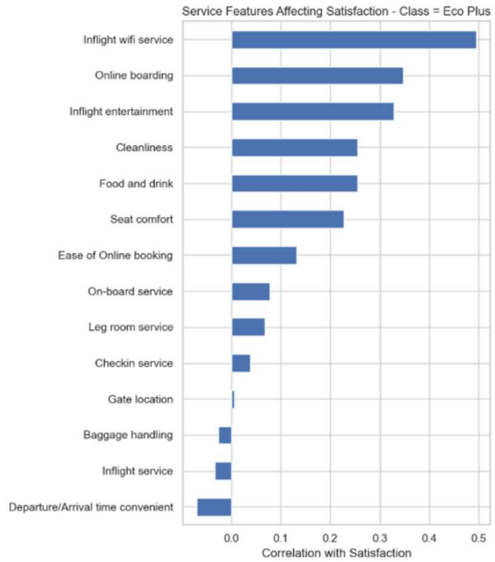

Figure 5.2 3 Correlation of service features with passenger satisfaction in Eco Plus Class

	Eco Plus
Inflight service	3.388444
Baggage handling	3.363758
Departure/Arrival time convenient	3.217507
Seat comfort	3.183747
Inflight entertainment	3.141713
Cleanliness	3.130771
Food and drink	3.122631
Leg room service	3.061382
On-board service	3.047638
Checkin service	3.017214
Gate location	2.967574
Online boarding	2.889245
Inflight wifi service	2.767948
Ease of Online booking	2.661996

Figure 5.2 2 Mean ratings of service features for Eco Plus Class

C. Business Class Correlation

As shown in Figure 5.2 5, Online Boarding and Inflight Entertainment are the strongest satisfaction drivers for Business Class passengers, followed by On-board Service, Leg Room Service, and Seat Comfort, indicating that premium travelers place high value on both digital efficiency and onboard comfort. Features such as Cleanliness, Inflight Service, and Baggage Handling contribute moderately, while Ease of Online Booking and Inflight Wi-Fi have relatively low impact. Figure 5.2 6 shows that most service-related mean scores are already high in Business Class, aligning with the high correlations and reflecting consistently strong service delivery. Departure/Arrival Time Convenience and Gate Location show minimal influence. Unlike other classes, no features are negatively correlated, suggesting that improvements in any area tend to enhance, rather than reduce, satisfaction for Business Class passengers.

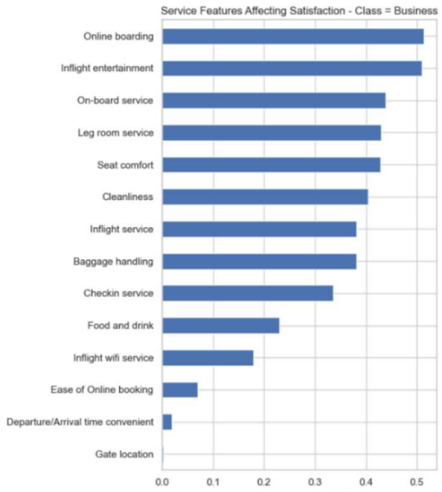

Figure 5.2 3 Correlation of service features with passenger satisfaction in Business Class

	Business
Inflight service	3.844579
Baggage handling	3.842907
Seat comfort	3.760858
Online boarding	3.716541
On-board service	3.679472
Leg room service	3.644498
Inflight entertainment	3.635437
Checkin service	3.519178
Cleanliness	3.477600
Food and drink	3.323165
Gate location	2.982926
Ease of Online booking	2.913964
Departure/Arrival time convenient	2.905910
Inflight wifi service	2.775315

Figure 5.2 4 Mean ratings of service features for Business Class

5.3. Feature Importance by Customer Type

A. Loyal Customer Correlation

As shown in Figure 5.3 1, Online Boarding and Inflight Entertainment are the strongest satisfaction drivers for loyal customers, indicating that digital convenience and high-quality entertainment remain central to their experience. Seat Comfort, Leg Room, and Cleanliness also show strong positive correlations, reinforcing the idea that comfort-related factors continue to shape satisfaction even among repeat travelers. In contrast, Inflight Wi-Fi and Ease of Online Booking display relatively low influence, and Figure 5.3 2 shows that their mean scores remain modest—highlighting ongoing improvement opportunities. Baggage Handling, On-board Service, and Check-in Service contribute moderately, while Gate Location shows minimal effect. Departure/Arrival Time Convenience is negatively correlated, suggesting that scheduling issues reduce satisfaction even for highly engaged, loyal passengers.

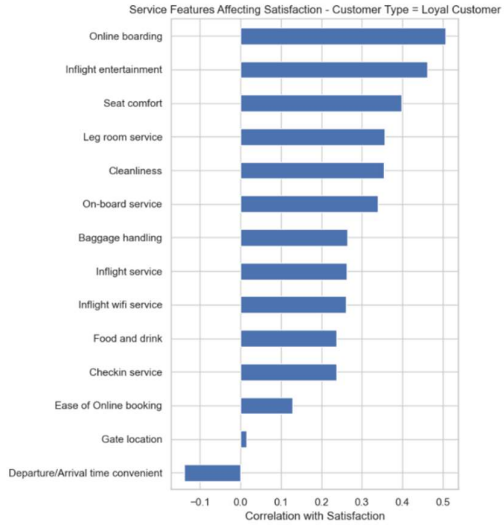

Figure 5.3 1 Correlation of service-related features with satisfaction among loyal passengers

Service Area	Loyal Mean
Inflight service	3.627828
Baggage handling	3.618030
Seat comfort	3.538853
Inflight entertainment	3.427434
On-board service	3.416813
Leg room service	3.380827
Online boarding	3.371124
Cleanliness	3.338224
Checkin service	3.323540
Food and drink	3.239499
Departure/Arrival time convenient	3.209449
Gate location	2.973211
Ease of Online booking	2.769803
Inflight wifi service	2.734418

Figure 5.3 2 Mean service ratings reported by loyal passengers

B. Disloyal Customer Correlation

As shown in Figure 5.3 1 Inflight Wi-Fi is the strongest satisfaction driver for disloyal customers, indicating that digital connectivity plays a central role in shaping their experience. Ease of Online Booking and Online Boarding also show high positive correlations, suggesting that smooth digital processes are especially important for passengers who are not repeat customers. Check-in Service, Baggage Handling, and On-board Service contribute moderately, while Seat Comfort, Cleanliness, and In-flight Entertainment have minimal influence. Gate Location and Departure/Arrival Time Convenience both show negative correlations, indicating that these operational issues frustrate disloyal passengers the most. As reflected in Figure 5.3 2, mean scores for most features remain low to moderate, confirming that disloyal customers tend to rate nearly all aspects of the service lower compared to loyal travelers.

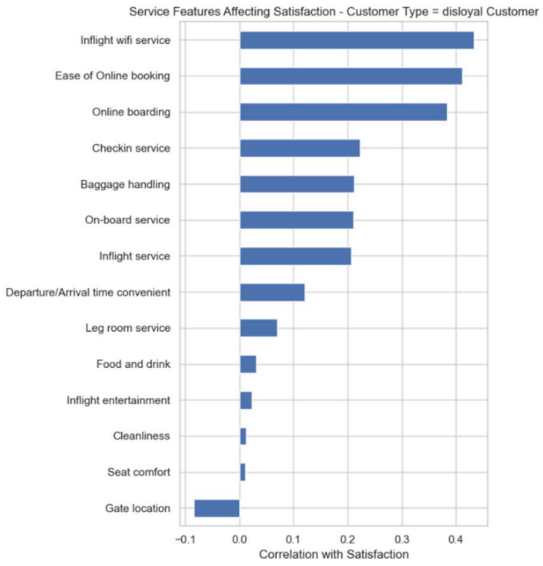

Figure 5.3 3 Correlation of service-related features with satisfaction among disloyal passengers

Service Area	Disloyal Mean
Inflight service	3.696802
Baggage handling	3.693588
On-board service	3.228228
Checkin service	3.218166
Leg room service	3.217849
Cleanliness	3.054265
Inflight entertainment	3.048206
Food and drink	3.034930
Seat comfort	2.994415
Gate location	2.993309
Online boarding	2.710131
Inflight wifi service	2.708498
Ease of Online booking	2.699173
Departure/Arrival time convenient	2.392972

Figure 5.3 4 Mean service ratings reported by loyal passengers

6. Discussion

A. Objective 1: To examine how passenger characteristics relate to satisfaction levels

Our group findings indicate that passenger characteristics influence satisfaction levels, supporting the view by Clemes et al. [1] that different groups of people evaluate airline service differently. However, our analysis showed that service-related factors have stronger and more consistent impact as compared to demographics alone. Though it's interesting to see that cabin class stands as the primary factor that determines passenger satisfaction, people taking the Business cabin class demonstrate higher satisfaction levels than economy and eco plus passengers, telling us that service levels are more important than demographic differences. Also, EDA revealed that customer type plays a pivotal role too, by showing that loyal customers are less affected by service inconsistencies when compared to disloyal ones. However, the loyalty status cannot outweigh the higher service quality offered by premium cabins. Premium cabins can draw new customers, but economy class fails to meet the expectations of even dedicated customers. Furthermore, travel reasons also affect satisfaction as it is clearly seen that business travelers show higher satisfaction rates than personal travelers in all three cabin classes. Age shows an intricate pattern with satisfaction, with passengers, aged 31-45 and 46-60 having the highest satisfaction, while the opposite happens for younger and older passengers. Gender, however, did not have a meaningful impact on the satisfaction level in this dataset.

B. Objective 2: To determine which service features or travel choices most strongly influence satisfaction through exploratory data analysis

Just like what existing literature highlighted that digital convenience and operational reliability are important drivers of airline satisfaction [2-5]. The EDA of our project supported this pattern; all passengers show the highest positive correlation between online boarding and satisfaction (r=0.45). For Eco, in-flight Wi-Fi emerged as their top satisfaction despite receiving the lowest ratings. Meanwhile for business class passengers, they value other service features more such as online boarding, in-flight entertainment, and seat comfort. Here, it can be said that airlines need to develop separate service improvements for different passenger segments because their needs differ significantly.

As for operations, delay analysis confirms that delays trigger significant dissatisfaction among all passengers when delays exceed 30 minutes because

dissatisfaction reaches 64%, though business class remains tolerant due to better service conditions as compared to lower classes. Delay-related scheduling creates a negative relationship with passenger satisfaction at a rate of ~-0.21 which demonstrates that flight timing directly influences passenger emotions regardless of service quality. The combination of poor scheduling and inadequate digital services leads to extreme dissatisfaction among disloyal customers.

In contrast to what happens when delay is longer, the longer routes flights lead to higher passenger satisfaction, which can be attributed to passengers flying business class when the flight is longer, which has more premium service.

C. To provide evidence-based insights and recommendations that airlines can use to improve service quality and customer experience

As a recommendation, the Airline can benchmark the services to focus on high priority services for possible improvement or at least maintain top-tier services. Based on industry standards, services with satisfaction correlation of 0.3 higher have high importance and the corresponding expected average mean for each service is 3.0 or higher. Given this standard [7,8], we adjusted the benchmark to 3.5 in this project to get more meaningful level of passenger satisfaction. This benchmark analysis was used to focus on what is important to give way for possible service overhaul for passengers. For any high-priority service feature that is less 3.5 average mean satisfaction, these should be immediately reviewed and improved with a goal of lifting the satisfaction mean to the benchmark of 3.5. On the other hand, for high priority service feature 3.5 or higher, it is imperative to maintain the service level or even top the current satisfaction level. The goal is to have consistency within the benchmark. As depicted in the benchmark analysis, it is evident that the business class are very much satisfied with services while the rest needs catching up with industry benchmark.

A. Benchmark Analysis for High Priority Service Features per Class Type

Table 2 Benchmark Analysis of High Priority Services Features by Class Type

'Service Features with Correlation >= 0.3 for Class = Eco Plus:'

	Service Feature	Mean (Class Subset)	Correlation	Performance Status
0	Inflight wifi service	2.768	0.495	Needs Improvement
1	Online boarding	2.889	0.348	Needs Improvement
2	Inflight entertainment	3.142	0.328	Needs Improvement

'Service Features with Correlation >= 0.3 for Class = Business:'

	Service Feature	Mean (Class Subset)	Correlation	Performance Status
0	Online boarding	3.717	0.513	Maintain Service Performance
1	Inflight entertainment	3.635	0.510	Maintain Service Performance
2	On-board service	3.679	0.438	Maintain Service Performance
3	Leg room service	3.644	0.430	Maintain Service Performance
4	Seat comfort	3.761	0.429	Maintain Service Performance
5	Cleanliness	3.478	0.404	Needs Improvement
6	Inflight service	3.845	0.381	Maintain Service Performance
7	Baggage handling	3.843	0.381	Maintain Service Performance
8	Checkin service	3.519	0.336	Maintain Service Performance

'Service Features with Correlation >= 0.3 for Class = Eco:'

	Service Feature	Mean (Class Subset)	Correlation	Performance Status
0	Inflight wifi service	2.675	0.467	Needs Improvement
1	Online boarding	2.813	0.316	Needs Improvement

B. Benchmark Analysis for High Priority Service Features per Customer Type

Table 3 Benchmark Analysis for High Priority Service Features per Customer Type

'Service Features with Correlation >= 0.3 for Customer Type = Loyal Customer:'

	Service Feature	Mean (Customer Type Subset)	Correlation	Performance Status
0	Online boarding	3.371	0.506	Needs Improvement
1	Inflight entertainment	3.427	0.461	Needs Improvement
2	Seat comfort	3.539	0.398	Maintain Service Performance
3	Leg room service	3.381	0.357	Needs Improvement
4	Cleanliness	3.338	0.355	Needs Improvement
5	On-board service	3.417	0.339	Needs Improvement

'Service Features with Correlation >= 0.3 for Customer Type = disloyal Customer:'

	Service Feature	Mean (Customer Type Subset)	Correlation	Performance Status
0	Inflight wifi service	2.708	0.433	Needs Improvement
1	Ease of Online booking	2.699	0.411	Needs Improvement
2	Online boarding	2.710	0.383	Needs Improvement

In essence, the recommendation moves around the Start, Continue and Stop approach. First, The airline should conduct a full review of the Delay Data to reduce occurrence. As previously discussed, Delay has the highest negative correlation to satisfaction which means passengers' satisfaction is very sensitive to delay. As it seems, the airline can explore if there is any measures done for the Business Class that can be emulated for other cabin classes to compensate for delays with appropriate cost-benefit analysis. Second, Improve or add Loyalty perks to Loyal Customers. As shown in the benchmark analysis, the high priority services for Loyal Customers are mostly below benchmark (except Seat Comfort) and needs improvement. In this case, the airline needs to improve

these services to satisfy loyal customers and retain loyalty. Offering extra benefits for online boarding and inflight entertainment can enhance loyalty and retention among these valued passengers. Third, explore upgrading services for Eco Class passengers. To improve satisfaction among Eco Class passengers, the airline should prioritize enhancing inflight Wi-Fi reliability, simplifying online boarding and booking processes. As per the benchmark analysis for Eco classes, the current satisfaction level for in-flight service, which has the highest priority for Eco classes, is very low from the benchmark satisfaction average. This clearly shows the dissatisfaction among Eco Class passengers, which should be carefully reviewed and remedied. Lastly, promote Business Class and top-tier services to both loyal or Eco Classes Passengers to convert more passengers to Business Class. Although the Business Class has the highest satisfaction average from all the cabin classes, it is important to note that the current satisfaction levels still has below benchmark for cleanliness. The airline should definitely ensure to maintain and enhance online boarding and inflight services to retain satisfaction among existing loyal business class passengers. Introduce targeted loyalty incentives (point system) and service upgrades aimed at encouraging both economy passengers and non-loyal business travellers to become loyal business class customers. This way, the top-tier services can entice non-business passengers' loyalty and conversion to business class.

In general, the aim is to Start converting passengers to higher cabin class to improve loyalty and satisfaction levels, Continuously improve Business Class services to the benchmark levels to retain business class passengers, and Stop occurrence of delays in flights by reviewing data and develop strategies to avoid, if not possible, compensate for uncontrollable flight delays. The strategy provides a strong foundation of capturing and retaining market share in a competitive aviation industry.

7. Conclusion

In summary, this study demonstrates that passenger satisfaction in the airline industry is shaped far more by service experience and operational reliability than by demographic characteristics. Service-related attributes, particularly Online Boarding, emerged as the strongest positive predictors of satisfaction across almost all customer segments, reinforcing the growing centrality of seamless digital touchpoints in modern air travel. Cabin class was the most influential customer factor: Business Class passengers consistently reported higher satisfaction, likely due to enhanced comfort, priority handling, and more personalized services, whereas Economy and Eco Plus passengers exhibited pronounced dissatisfaction, especially with In-flight Wi-Fi—despite this feature being identified as their most critical driver of satisfaction, revealing a substantial expectation–performance gap. Operational performance further intensified these dynamics, with delays exceeding 30 minutes associated with at least 64% dissatisfaction and only limited evidence of negative effects among Business Class passengers, suggesting that premium service offerings partially buffer the impact of disruptions. Loyalty patterns added another layer of complexity, as loyal customers, while generally more satisfied, still registered suboptimal evaluations, and disloyal customers proved highly sensitive to shortcomings in Wi-Fi quality, booking convenience, and schedule reliability, underscoring the fragility of brand perception among infrequent or switching passengers. As the benchmark analysis tells more about the current status of each service feature, this is the airline's tool to improve the overall performance and target goals. Taken together, the moderate mean service rating of 3.24 out of 5, coupled with a high dissatisfaction rate of 56.7%, portrays an industry that reliably meets basic expectations yet falls short of delivering consistently positive experiences, highlighting a pivotal moment for airlines to re-examine how effectively their service systems convert operational performance and digital interactions into enduring customer satisfaction. With a benchmark of 3.5 for the industry, the airline must push for better services to secure a place in the aviation industry and maintain its position.

References

[1] Marketline. (2024). Airlines in Canada. Research and Markets. https://www.researchandmarkets.com/report/canada-airline-market

[2] CBC News. (2024, June 29). The new airline rivalries: Air Canada vs. Porter, WestJet vs. Flair. *CBC News*. https://www.cbc.ca/news/canada/montreal/airline-rivalries-air-canada-porter-westjet-flair-1.7251143

[3] M.D. Clemes, C. Gan, T.-H. Kao, M. Choong, An empirical analysis of customer satisfaction in international air travel, Innovative Marketing 4(2) (2008) 50–62.

[4] C. Song, J. Xie, J. Zhang, X. Wu, The adverse impact of flight delays on passenger satisfaction: An innovative prediction model utilizing wide and deep learning, Journal of Air Transport Management 110 (2023) 102154.

[5] A.V.M. Oliveira, B.F. Oliveira, M. Vassallo, Airport service quality perception and flight delays, arXiv Preprint (2024). https://arxiv.org/abs/2401.02139

[6] M. Bakır, E. Özdemir, Ş. Akan, Ö. Atalık, A bibliometric analysis of airport service quality, Journal of Air Transport Management 103 (2022) 102273. https://doi.org/10.1016/j.jairtraman.2022.102273

[7] T. Noviantoro, J.-P. Huang, Investigating airline passenger satisfaction: Data mining method, Research in Transportation Business & Management 43 (2022) 100726. https://doi.org/10.1016/j.rtbm.2021.100726

[8] W. Du, Neural network in aircraft customer satisfaction prediction, International Conference on Artificial Intelligence and Data Engineering (2023). https://www.researchgate.net/publication/375551968_Neural_Network_in_Aircraft_Customer_Satisfaction_Prediction

[9] G.E. Gignac, E.T. Szodorai, Effect size guidelines for individual differences researchers, Personality and Individual Differences 102 (2016) 74–78.

[10] C.M. Kei, A study of customer satisfaction in airlines, International Journal of Trade, Economics and Finance 10(6) (2019) 144–148. https://doi.org/10.18178/ijtef.2019.10.6.652

[11] Teejmahal. (2020). Airline passenger satisfaction [Data set]. Kaggle. https://www.kaggle.com/datasets/teejmahal20/airline-passenger-satisfaction