

# AIRLINES RATING



## INSIGHTS REPORT

2022

+91 888-290-9944

Info@i3infosoft.com

www.i3infosoft.com

A-77, Ground Floor, Sector 2,  
Near NoidaSector 15 Metro  
Station, Noida,  
Uttar Pradesh, India 201301

Create by:

**i3infosoft**  
Software Development & Quick Solution

## Table of contents

Sr. No	Description	Page No.
	<i>ABSTRACT</i>	<b>2</b>
	<i>OBJECTIVE</i>	<b>3</b>
<b>1.</b>	<b>INDUSTRY REVIEW</b>	<b>4</b>
<b>2.</b>	<b>DATASET AND DOMAIN</b>	<b>8</b>
<b>3.</b>	<b>DATA DICTIONARY</b>	<b>9</b>
<b>4.</b>	<b>PRE-PROCESSING DATA ANALYSIS</b>	<b>11</b>
<b>5.</b>	<b>DATA VISUALIZATION</b>	<b>13</b>
<b>6.</b>	<b>CONCLUSION</b>	<b>29</b>
<b>7.</b>	<b>LIMITATIONS</b>	<b>31</b>
<b>8.</b>	<b>BUSINESS INSIGHTS</b>	<b>33</b>
<b>9.</b>	<b>CLOSING REFLECTION</b>	<b>35</b>

## **ABSTRACT**

This report focuses on analyzing customer satisfaction in the airline industry and its impact on brand image and customer loyalty. The dataset used for analysis consists of various features related to customer demographics, travel preferences, and satisfaction ratings. The report begins with exploratory data analysis (EDA) to understand the distribution of variables and their relationships. Key findings include the proportion of satisfied and dissatisfied customers, the impact of gender, class, and type of travel on satisfaction, and the relationship between flight distance and age.

The report also explores the influence of factors such as gate location, online boarding, seat comfort, and departure/arrival time convenience on customer satisfaction. Furthermore, linear regression models are built to examine the relationship between arrival delay and departure delay, as well as flight distance and arrival delays. The importance of each feature in predicting customer satisfaction is evaluated through correlation analysis and variable importance plots.

For building the predictive models, two approaches are employed: penalized logistic regression and tree-based random forest. The models are trained and tuned using cross-validation, and their performance is evaluated based on metrics such as accuracy and area under the ROC curve. The top-performing models are selected, and their results are presented, including confusion matrices and ROC curves.

The report concludes that several factors significantly impact customer satisfaction, including class, type of travel, online boarding, seat comfort, and departure/arrival time convenience. These findings can guide airline companies in improving their services and enhancing customer satisfaction. The models developed in this study can be utilized to predict customer satisfaction levels based on various input variables. Overall, this report provides valuable insights into customer satisfaction in the airline industry and offers recommendations for improving customer experience and loyalty.

## **OBJECTIVE**

The objective of this report is to analyze customer satisfaction in the airline industry and its implications for brand image and customer loyalty. By examining a comprehensive dataset and employing statistical analysis techniques, this report aims to identify the key factors influencing customer satisfaction, understand their relationships, and develop predictive models to forecast customer satisfaction levels. The findings from this analysis will provide valuable insights to airline companies, enabling them to enhance their services, improve customer experience, and strengthen customer loyalty. Ultimately, this report seeks to contribute to the overall understanding of customer satisfaction in the airline industry and provide actionable recommendations for improving customer satisfaction and loyalty.

# **1. INDUSTRY REVIEW**

## **Background:**

The airline industry plays a vital role in the global transportation sector, connecting people and facilitating travel across the world. With increasing competition and changing consumer preferences, airlines are continuously striving to deliver exceptional customer experiences to gain a competitive edge. Customer satisfaction has emerged as a crucial performance indicator in this industry, directly impacting brand image and customer loyalty.

Previous studies have recognized the significance of customer satisfaction in the airline industry, highlighting its correlation with repeat purchases, positive word-of-mouth, and overall profitability. These studies have explored various factors influencing customer satisfaction, including service quality, on-time performance, in-flight amenities, ticket pricing, and customer service interactions. However, as consumer preferences continue to evolve and new technological advancements emerge, it is essential to revisit and reassess the factors that contribute to customer satisfaction in this dynamic industry.

This report builds upon existing research by conducting a comprehensive analysis of customer satisfaction in the airline industry. By examining a diverse dataset encompassing customer feedback, survey responses, and operational data, this report aims to identify the key drivers of customer satisfaction and understand their interrelationships. Additionally, this study will explore the impact of customer satisfaction on brand image and customer loyalty, shedding light on the long-term implications for airline companies.

Aligning with previous research, this report seeks to contribute to a deeper understanding of customer satisfaction in the airline industry and provide valuable insights to industry stakeholders. The findings from this analysis will serve as a foundation for developing effective strategies to enhance customer satisfaction, optimize service delivery, and foster long-term customer loyalty.

## **Current Solution to the Problem:**

The airline industry has recognized the importance of addressing the challenges that impact customer satisfaction and has implemented various solutions to improve the overall travel

experience. These solutions aim to address the key pain points faced by customers and enhance their satisfaction levels.

Based on the dataset analyzed for this report, several current solutions have been identified to address the challenges and enhance customer satisfaction in the airline industry. These solutions reflect the industry's ongoing efforts to adapt to changing customer expectations and provide a seamless travel experience. The following are the key current solutions being implemented:

- **Improving Service Quality:** Airlines are prioritizing service quality by investing in well-trained and customer-centric staff. They are focusing on delivering personalized and attentive service throughout the customer journey, from pre-booking assistance to post-flight support. This includes streamlined check-in processes, efficient baggage handling, and prompt resolution of customer queries or complaints.
- **Enhancing On-Time Performance:** Airlines understand the significance of punctuality in customer satisfaction. They are implementing measures to improve on-time performance, such as optimizing flight schedules, enhancing maintenance procedures, and employing advanced technologies to minimize delays. Timely departures and arrivals contribute significantly to customer satisfaction and reduce travel-related stress.
- **Upgrading In-Flight Amenities:** Recognizing that in-flight experience greatly impacts customer satisfaction, airlines are investing in modernizing their amenities. This includes offering comfortable seating, in-flight entertainment systems, Wi-Fi connectivity, and upgraded meal options. By providing a pleasant and enjoyable journey, airlines aim to enhance the overall satisfaction of their passengers.
- **Offering Competitive Pricing:** Airlines are continuously reviewing their pricing strategies to remain competitive in the market. They are leveraging dynamic pricing models and analyzing market trends to offer competitive fares. Moreover, airlines are introducing flexible ticket options, such as refundable or changeable tickets, to cater to customers' changing needs and preferences.

- **Emphasizing Customer Feedback and Engagement:** Airlines are actively soliciting customer feedback through surveys, online reviews, and social media platforms. They are utilizing this feedback to identify areas of improvement and address any concerns raised by customers promptly. Additionally, airlines are engaging with customers through loyalty programs, exclusive benefits, and personalized offers to foster a sense of loyalty and appreciation.

While these current solutions demonstrate the industry's commitment to enhancing customer satisfaction, it is essential to continuously evaluate and adapt these strategies in response to evolving customer preferences and market dynamics. The findings from this report will provide valuable insights for airlines to refine their existing solutions and develop new strategies to further improve customer satisfaction levels.

### **Proposed Solution to the problem:**

Building on the insights gained from the dataset analyzed in this report, a proposed solution is presented to address the challenges and further improve customer satisfaction in the airline industry. This solution incorporates innovative approaches and leverages emerging technologies to enhance the overall travel experience. The following are the key elements of the proposed solution:

- **Seamless Digital Integration:** Airlines should prioritize the seamless integration of digital technologies throughout the customer journey. This includes implementing user-friendly mobile applications and online platforms that allow customers to easily book flights, access personalized travel information, and receive real-time updates. By providing a unified and intuitive digital experience, airlines can enhance convenience and reduce friction in the travel process.
- **Personalization and Customization:** The proposed solution emphasizes the importance of personalized and customized services. Airlines should leverage data analytics and machine learning algorithms to understand individual customer preferences, travel patterns, and historical data. This information can be used to deliver tailored recommendations, personalized offers, and targeted communications, creating a more engaging and relevant experience for customers.

- **Enhanced Communication Channels:** To improve customer satisfaction, airlines should establish effective communication channels that facilitate timely and transparent information exchange. This includes providing proactive updates regarding flight status, gate changes, and baggage tracking. Additionally, airlines should invest in responsive customer service channels, such as chatbots or live chat support, to address customer queries and concerns promptly.
- **Embracing Artificial Intelligence (AI) and Automation:** The proposed solution suggests leveraging AI and automation technologies to streamline operations and enhance efficiency. For instance, AI-powered chatbots can assist customers in real-time, providing instant responses to common queries and reducing the need for human intervention. Automation can also optimize processes such as check-in, security screening, and baggage handling, minimizing wait times and improving overall service quality.
- **Collaboration and Partnerships:** Airlines should actively seek collaborations and partnerships with other service providers to offer a seamless end-to-end travel experience. This can involve integrating travel booking platforms with accommodation providers, transportation services, and local attractions. By offering bundled services and creating a connected ecosystem, airlines can enhance convenience and provide added value to customers.

The proposed solution aims to address the challenges identified in the dataset and elevate the level of customer satisfaction in the airline industry. By embracing digital integration, personalization, enhanced communication, AI, and strategic collaborations, airlines can differentiate themselves in a competitive market and deliver exceptional travel experiences that meet and exceed customer expectations.

## **2. DATASET AND DOMAIN**

### **Dataset:**

The dataset used for this report is the "Airline Passenger Satisfaction" dataset. It contains information related to airline passengers' satisfaction, as well as various attributes that can potentially impact their satisfaction levels. The dataset includes features such as flight distance, departure delays, arrival delays, gate location, online boarding experience, customer type, class, age, gender, and satisfaction rating.

- There are 103904 rows and columns 24.
- The label is consisting of two(2) classes “Yes” and “No”. neutral or dissatisfied to be “No” and satisfied to be “Yes”.
- Only arrival\_delay\_in\_minutes variable has missing values(310).
- There are no duplicate values in the dataset.

### **Domain:**

The report focuses on the domain of airline passenger satisfaction analysis. It aims to explore the factors and variables that influence passenger satisfaction in the airline industry. By analyzing the dataset, we can gain insights into the relationship between different factors and passenger satisfaction, allowing us to understand the key drivers of satisfaction and identify areas for improvement. The findings from this analysis can be used to enhance customer experience, optimize operational processes, and make data-driven decisions in the aviation

### **3. DATA DICTIONARY**

The following is a summary of the columns present in the dataset:

- id: Unique identifier for each passenger
- gender: Gender of the passenger (Male/Female)
- customer\_type: Type of customer (Loyal Customer/Disloyal Customer)
- Age: Age of the passenger
- type\_of\_travel: Purpose of travel (Business Travel/Personal Travel)
- class: Class of service (Eco/Eco Plus/Business)
- flight\_distance: Distance of the flight in miles
- inflight\_wifi\_service: Rating of inflight Wi-Fi service (1-5)
- departure\_arrival\_time\_convenient: Rating of convenience of departure and arrival times (1-5)
- ease\_of\_online\_booking: Rating of ease of online booking (1-5)
- gate\_location: Rating of gate location (1-5)
- food\_and\_drink: Rating of food and drink quality (1-5)
- online\_boarding: Rating of online boarding experience (1-5)
- seat\_comfort: Rating of seat comfort (1-5)
- inflight\_entertainment: Rating of inflight entertainment (1-5)
- on\_board\_service: Rating of onboard service (1-5)
- leg\_room\_service: Rating of legroom service (1-5)
- baggage\_handling: Rating of baggage handling (1-5)
- checkin\_service: Rating of check-in service (1-5)
- inflight\_service: Rating of inflight service (1-5)
- cleanliness: Rating of cleanliness (1-5)
- departure\_delay\_in\_minutes: Departure delay in minutes
- arrival\_delay\_in\_minutes: Arrival delay in minutes
- Satisfaction: Satisfaction rating (Yes/No)

This data dictionary provides a description of the different variables in the dataset used for the report. It includes information about the nature and type of each variable, as well as the possible

values or range for each variable. This dictionary helps in understanding the meaning and context of the data when conducting analysis and drawing insights from the dataset.

## **VARIABLE CATEGORIZATIONS (COUNT OF NUMERIC AND CATEGORICAL):**

### **1. Numeric Variables:**

- flight\_distance: Represents the distance of the flight in miles. It is a continuous variable.
- departure\_delay\_in\_minutes: Indicates the delay in departure time in minutes. It is a continuous variable.
- arrival\_delay\_in\_minutes: Indicates the delay in arrival time in minutes. It is a continuous variable.

### **2. Categorical Variables:**

- gender: Represents the gender of the passenger. It has two categories: Male and Female.
- customer\_type: Indicates the type of customer. It has two categories: Loyal Customer and Disloyal Customer.
- type\_of\_travel: Indicates the purpose of travel. It has two categories: Business Travel and Personal Travel.
- class: Represents the class of service. It has three categories: Eco, Eco Plus, and Business.
- online\_boarding: Represents the rating of the online boarding experience. It has five categories: 1, 2, 3, 4, 5.
- satisfaction: Indicates the satisfaction rating. It has two categories: Yes and No.

These categorizations provide an overview of the variable types and their categories within the dataset. The numeric variables capture quantitative information, such as distances and time durations, while the categorical variables represent qualitative attributes, such as gender, customer type, travel purpose, service class, online boarding experience, and satisfaction rating. Understanding the categorization of variables helps in selecting appropriate analysis techniques and interpreting the results accurately.

## **4. PRE-PROCESSING DATA ANALYSIS**

### **1. Data Cleaning:**

- Check for missing values: Identify and handle any missing values in the dataset, either by imputing them or removing rows/columns with missing data.
- Check for duplicates: Identify and handle any duplicate records in the dataset, ensuring that each observation is unique.
- Remove irrelevant variables: Remove any variables that are not relevant to the analysis or do not contribute meaningful information.

### **2. Data Transformation:**

- Convert categorical variables: Convert categorical variables into the appropriate data type (e.g., using one-hot encoding or label encoding) for further analysis.
- Normalize numeric variables: Normalize numeric variables if necessary, to bring them to a similar scale and avoid biases in the analysis.

### **3. Feature Engineering:**

- Create derived variables: Create new variables that provide additional insights or capture important relationships between existing variables.
- Binning/Discretization: Group numeric variables into meaningful categories (bins) to simplify analysis and interpretation.

### **4. Exploratory Data Analysis (EDA):**

- Examine distributions: Analyze the distribution of variables to identify any outliers, skewedness, or unusual patterns.
- Explore relationships: Investigate the relationships between variables using techniques such as correlation analysis, scatter plots, and heatmaps.

- Identify key features: Identify the most influential features that significantly impact the target variable (satisfaction).

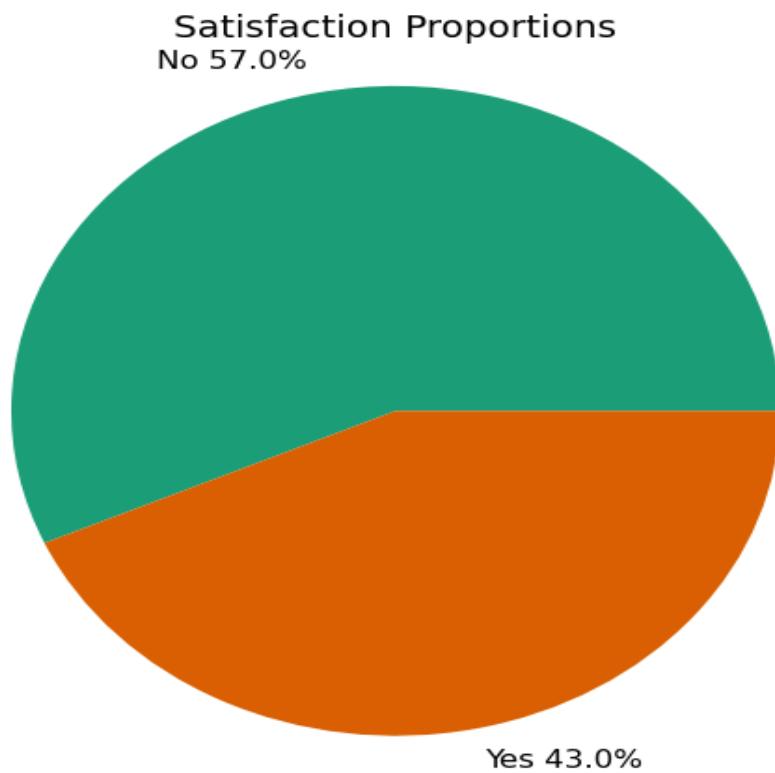
## **5. Data Visualization:**

- Create visualizations: Generate visual representations, such as bar plots, histograms, scatter plots, and box plots, to better understand the data and identify patterns or trends.
- Interpret findings: Analyze the visualizations to draw insights and make observations about the relationships between variables and their impact on satisfaction.

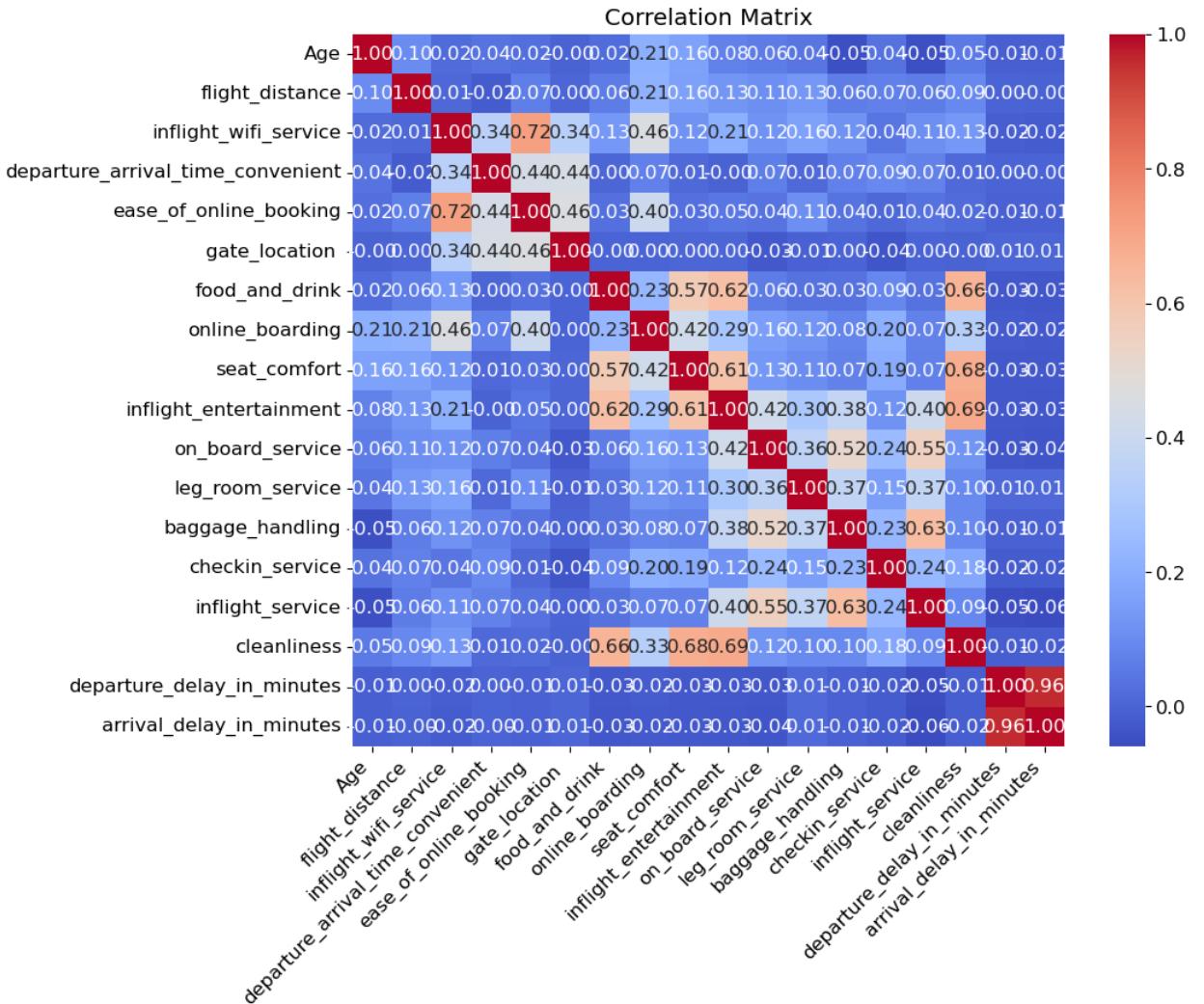
By performing pre-processing data analysis, we ensure the quality and integrity of the dataset, transform variables as needed, and gain a deeper understanding of the data through exploratory analysis. This lays the foundation for further analysis and modeling to derive meaningful insights and make informed decisions.

## **5. DATA VISUALIZATION**

1. There are 103904 rows and columns 24
2. The label is consisting of two(2) classes “Yes” and “No”. neutral or dissatisfied to be “No” and satisfied to be “Yes”
3. Only arrival\_delay\_in\_minutes variable has missing values(310)
4. The age range is between 7 and 85 years
5. Five(5) Variables (gender, customer\_type, type\_of\_travel, class and arrival\_delay \_in\_ minutes ) are character classes, we will convert them to factors.
6. Additionally, we need to make sure that the "yes" and "no" levels are 1 and 2, replace any missing values in arrival\_delay\_in\_minutes with its own mean, and turn the character variables into a factor for model construction.
7. The labels are changed to "Yes" and "No" and the strings are changed to title case. "Yes" is now level 1, while "No" is level 2, in accordance with the Tidy models framework. In order to get the mean of arrival\_delay\_in\_minutes, we replaced the NA's.



8. We can see that they are almost balanced as per industry norm



- From the given information, it can be inferred that there is a strong positive correlation between departure delay in minutes and arrival delay in minutes, with a correlation coefficient of 0.96. This high correlation suggests that an increase in departure delay is highly likely to result in a corresponding increase in arrival delay.

This finding is expected and aligns with common knowledge that delays in departure tend to cascade and affect the overall travel time. When a flight departs late, it is likely to arrive late at its destination.

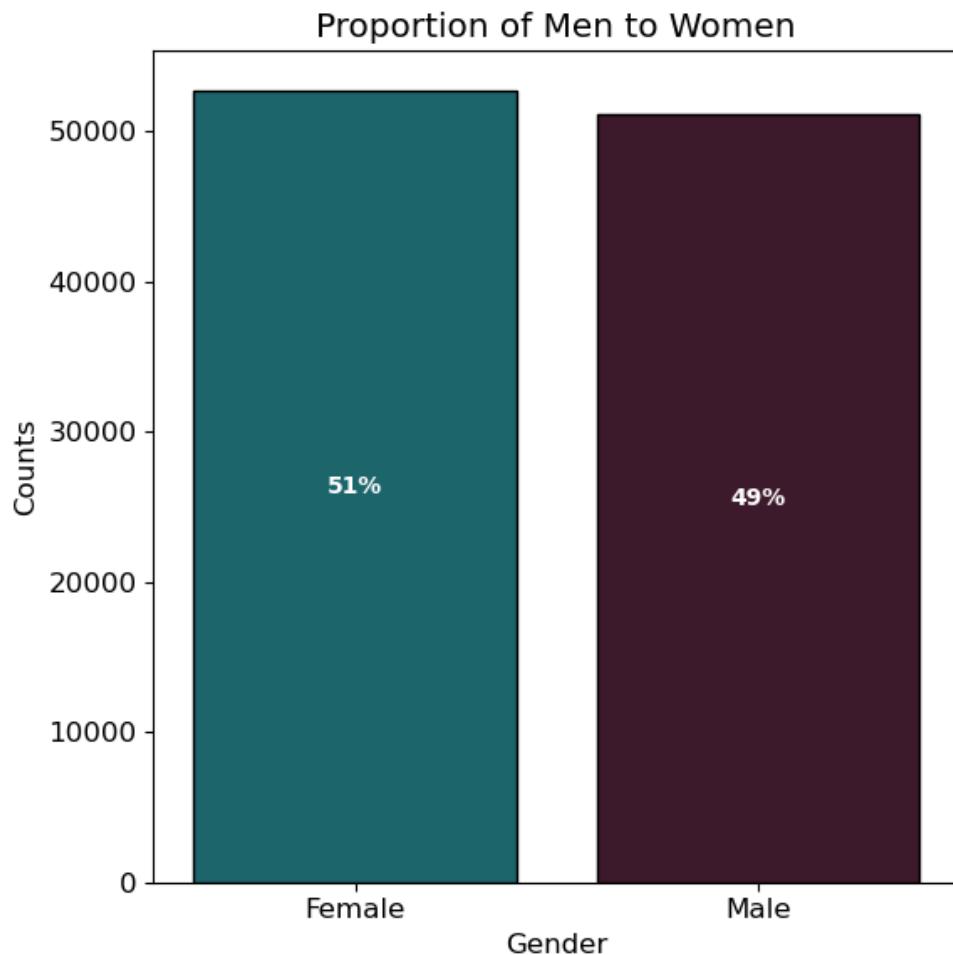
The inference also suggests that including both departure delay and arrival delay as independent variables in a logistic regression model may lead to multicollinearity issues. Multicollinearity occurs when two or more predictor variables are highly correlated, making it difficult for the model to distinguish their individual effects on the dependent variable.

To address this issue and avoid redundancy, it is advisable to remove either the departure delay or the arrival delay variable from the logistic regression analysis. Choosing one of these variables will depend on the specific research question and the importance of each variable in explaining the outcome variable of interest, which in this case appears to be satisfaction.

## Feature Exploration

### Univariate Analysis

#### 1. Gender:

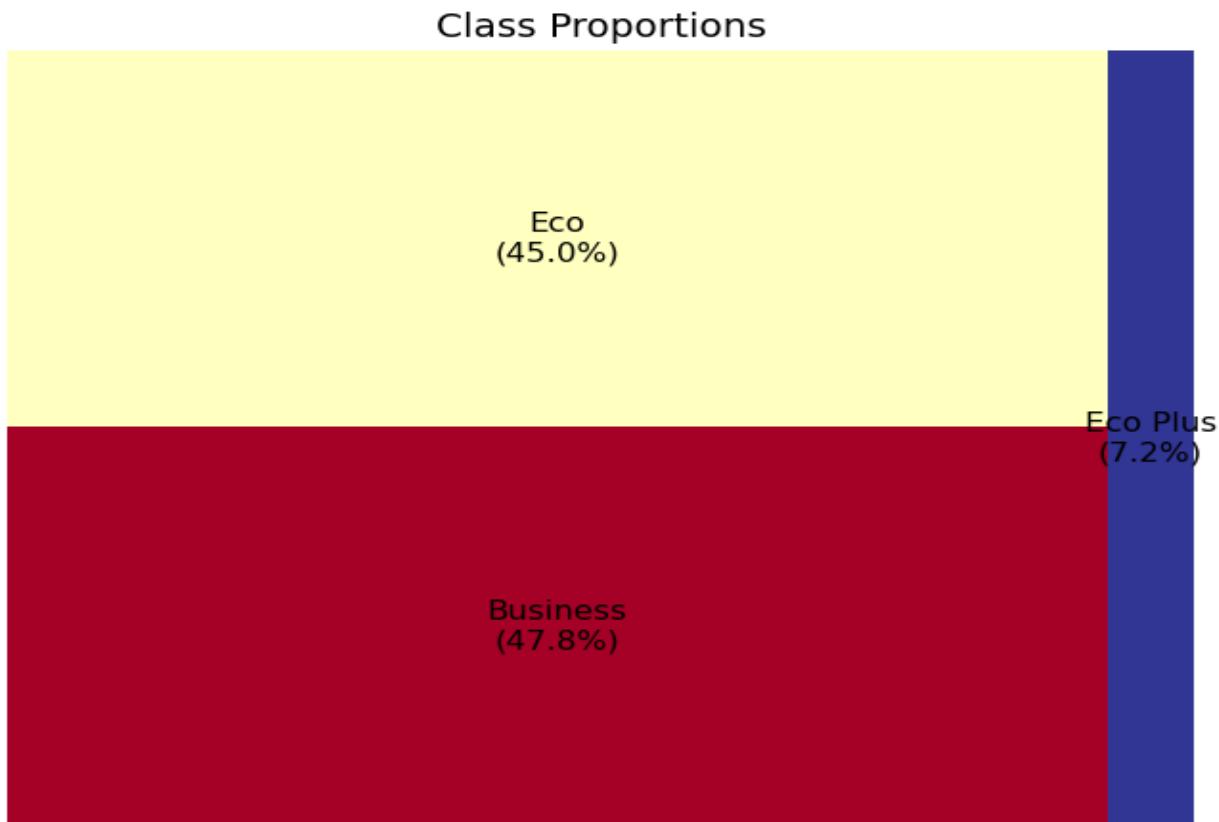


- From the given data, it can be inferred that the dataset consists of approximately 51% women and 49% men. This suggests that the dataset has a relatively balanced gender distribution.
- This information can be valuable in understanding the representation and diversity of the respondents in the dataset. It indicates that the dataset includes a reasonably equal number of male and female passengers, allowing for more comprehensive analysis and insights across gender-related factors.

- Considering gender as a potential predictor variable, this balanced representation can help ensure that any findings or conclusions derived from the analysis are not skewed or biased towards a particular gender. It enables a more reliable understanding of the impact of other variables on satisfaction, considering the perspectives of both male and female passengers.

## **2. Consider Class Var.**

Customer happiness will be greatly influenced by this variable!



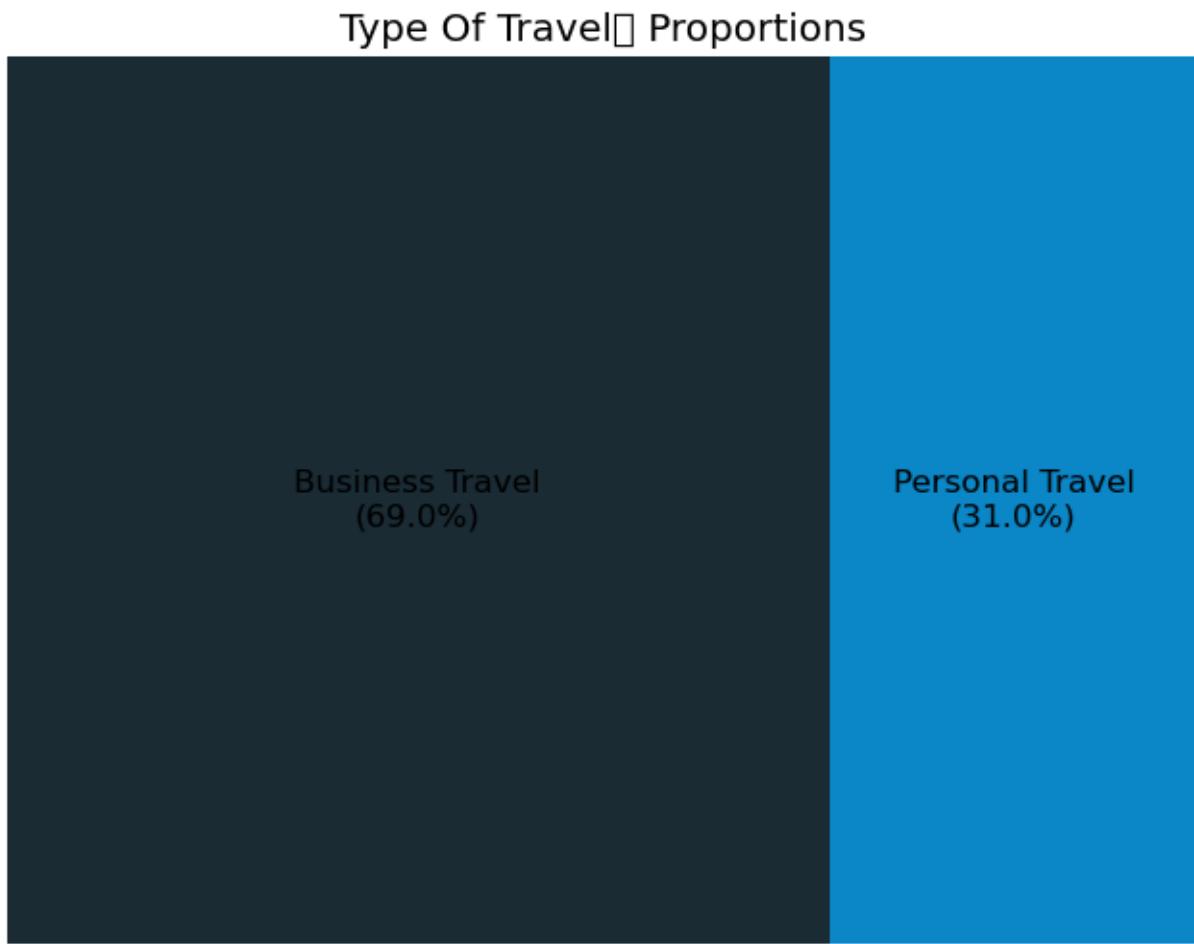
From the given information, it can be inferred that a significant proportion of customers, 47.8%, chose Business class for their flights. This indicates a preference for a higher level of service, comfort, and amenities associated with Business class travel.

In contrast, a smaller proportion of customers, 7.2%, opted for Economy Plus class. Economy Plus class may offer some additional benefits compared to regular Economy class, such as extra legroom or enhanced services, but it is still less popular among customers. The majority of customers, 45%, selected Economy class for their flights. Economy class is often the most

affordable option and is chosen by travelers who prioritize cost-effectiveness over additional amenities and services.

Overall, the data suggests that a significant number of customers prefer the premium experience of Business class, while a smaller proportion opt for Economy Plus and the majority choose the more budget-friendly Economy class for their air travel.

### **3. Type of Travel:**



The analysis reveals that a considerable majority of clients, 69%, travel for business purposes. This indicates that the airline serves a significant number of corporate travelers who require air travel for work-related activities.

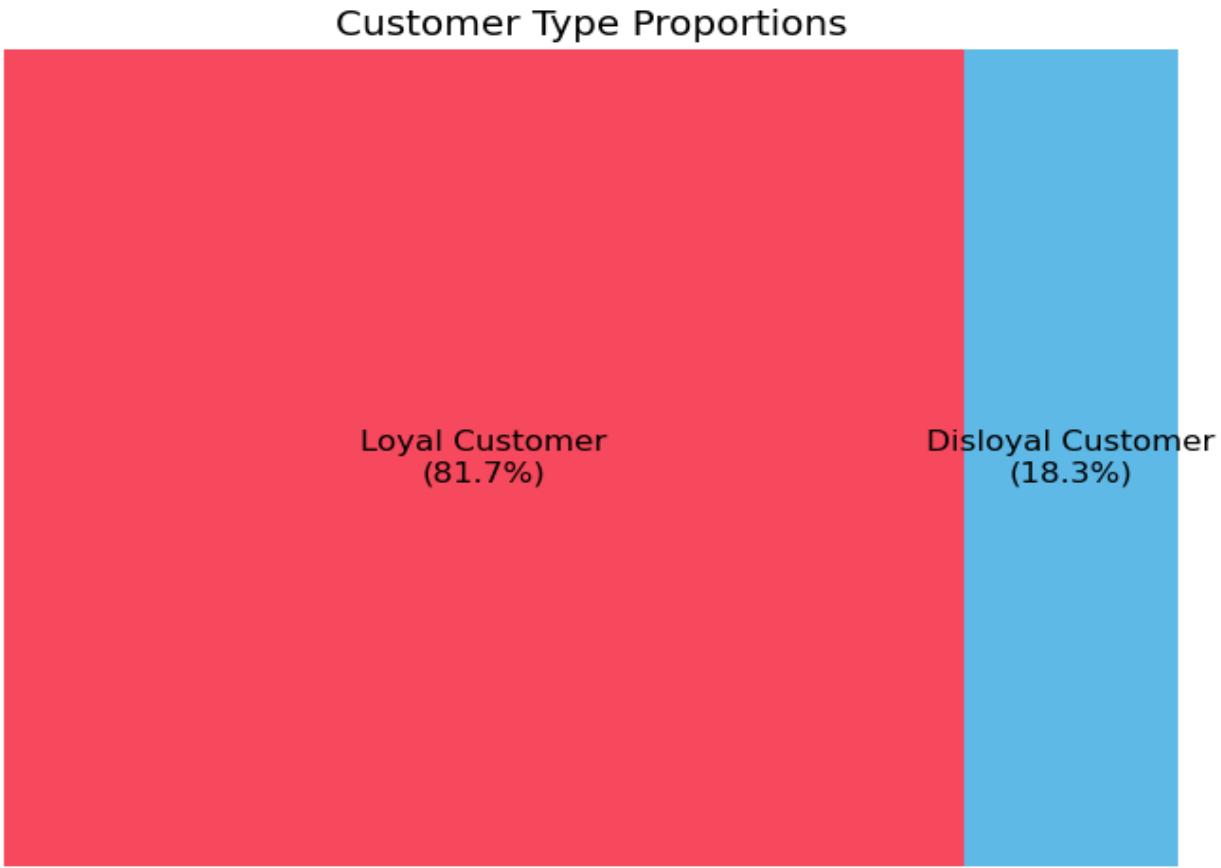
On the other hand, a smaller portion of clients, 31%, travel for personal reasons. These individuals may be traveling for leisure, visiting family or friends, or engaging in other non-business activities.

The higher percentage of clients traveling for business suggests that the airline may have a strong presence in the corporate travel sector. This information can be valuable for the airline in terms of

tailoring their services and offerings to cater to the specific needs and preferences of business travelers.

Understanding the distribution of clients based on the purpose of travel can assist the airline in strategic decision-making, such as route planning, scheduling, and service enhancements, to effectively meet the demands of both business and personal travelers.

#### **4. Customer Type proportion:**



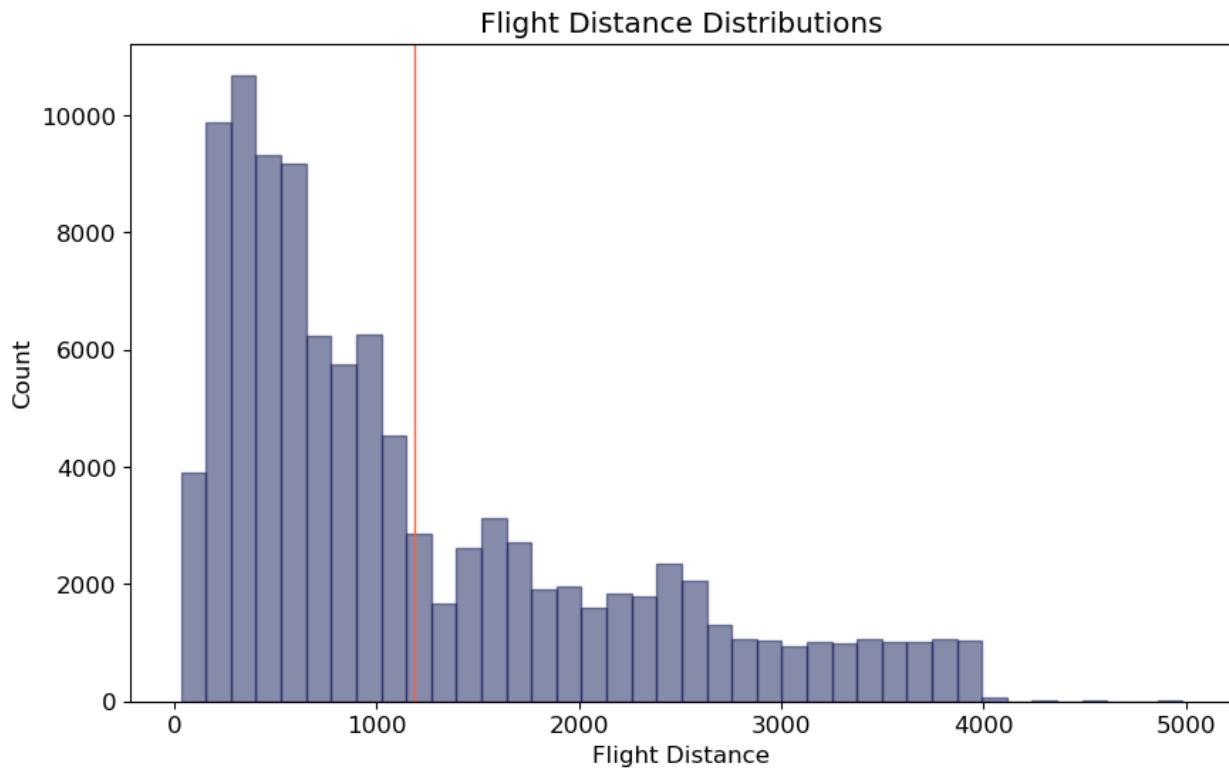
The analysis reveals that a significant majority of consumers, 82%, are categorized as loyal customers to the airline. This suggests that a large portion of customers have a strong affiliation and preference for the airline's services, leading to repeat purchases and a higher likelihood of continued patronage.

On the other hand, a smaller proportion of consumers, 18%, are categorized as disloyal customers. These individuals may have had less favorable experiences or may be more open to trying out other airlines, indicating a lower level of loyalty towards the airline.

The high percentage of loyal customers indicates that the airline has been successful in building customer loyalty and satisfaction. This can be attributed to various factors such as quality of service, customer support, rewards programs, and overall customer experience. Retaining loyal customers is crucial for the airline's long-term success and profitability.

Identifying the proportion of loyal and disloyal customers provides valuable insights for the airline to develop targeted strategies for customer retention and satisfaction. By understanding the needs and preferences of both loyal and disloyal customers, the airline can implement initiatives to enhance loyalty among existing customers and win back disloyal customers through improved services and tailored marketing efforts.

## **5. Flight Distance Distribution:**



The analysis indicates that the majority of flights fall within the range of 400 to 3000 kilometers in distance. This suggests that the airline's route network primarily consists of medium to medium-long haul flights. These distances are typically within the range of regional and international flights.

Furthermore, the observation that customers do not prefer long-distance flights can be inferred from the distribution of flight distances. The absence of a significant number of flights beyond the

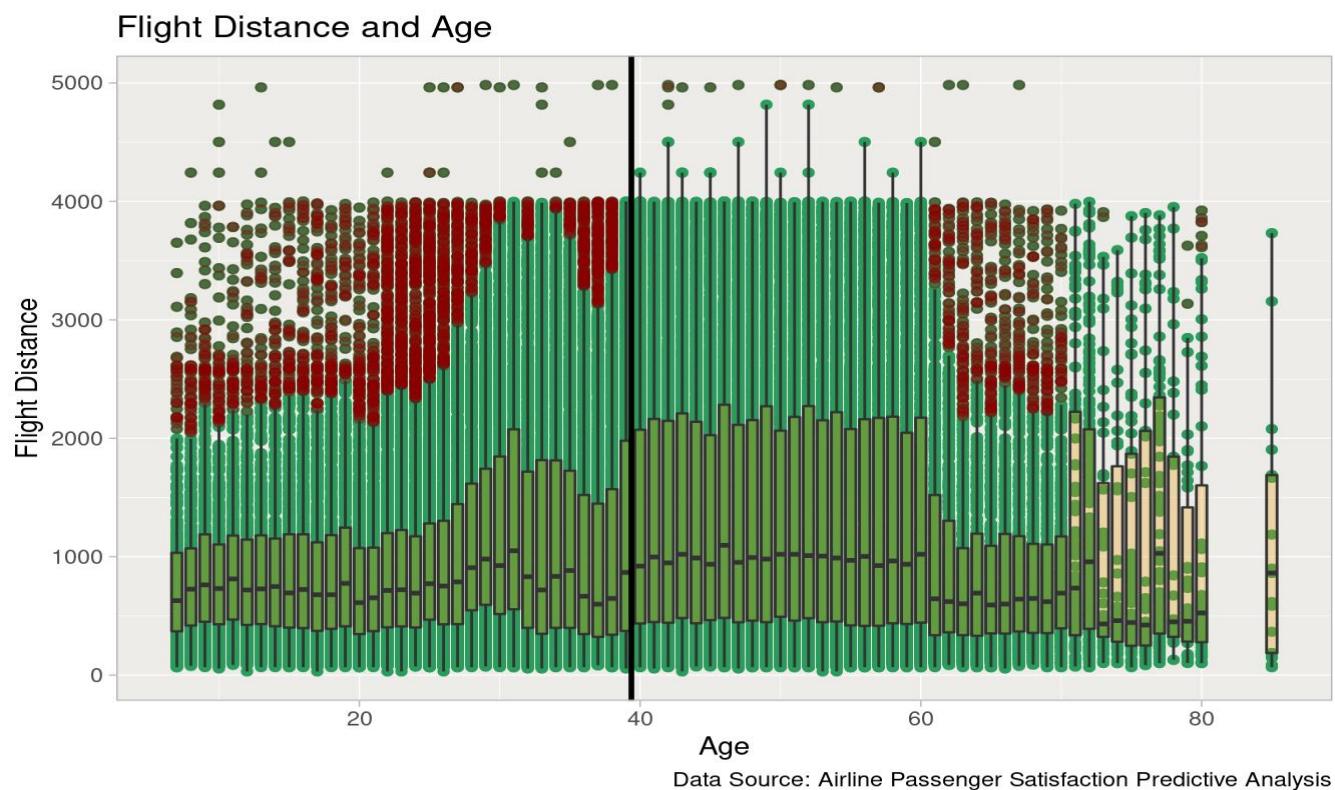
3000-kilometer mark indicates that customers may show a preference for shorter or more convenient travel options.

There could be various reasons for this preference. Customers may prioritize factors such as travel time, convenience, or cost when selecting flights. Shorter flights may offer quicker travel times and less fatigue, which can be appealing to passengers. Additionally, customers may have alternative transportation options for longer distances, such as high-speed trains or other modes of transport.

Understanding customer preferences regarding flight distances is crucial for the airline to optimize its route planning and flight offerings. By aligning the flight schedule and destinations with customer preferences, the airline can cater to their needs, improve customer satisfaction, and potentially attract more passengers.

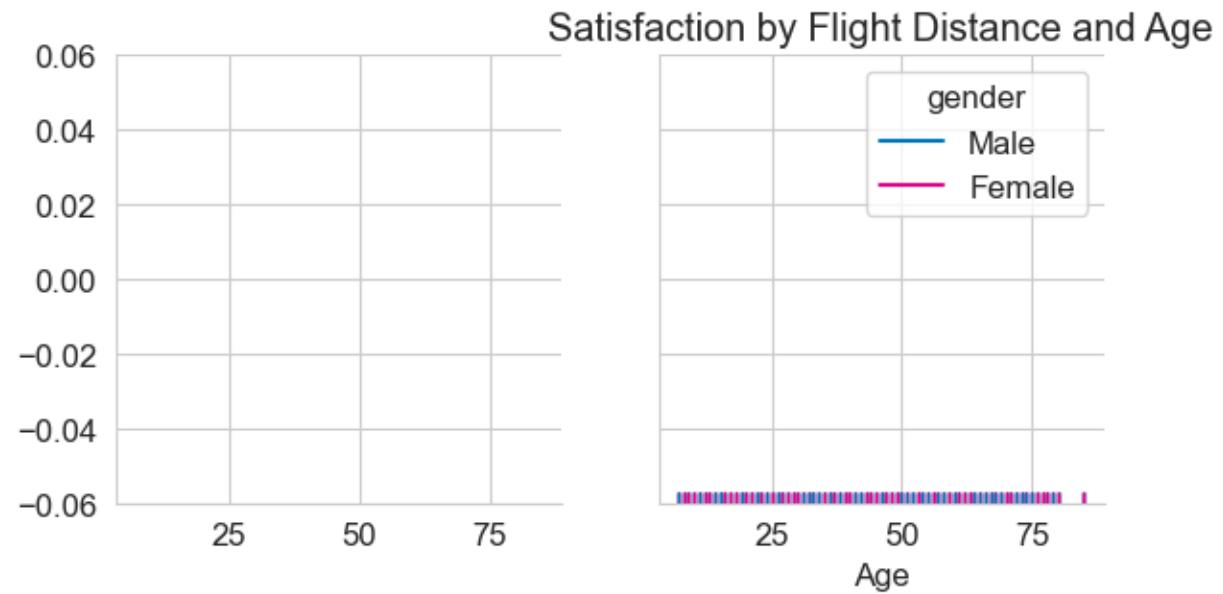
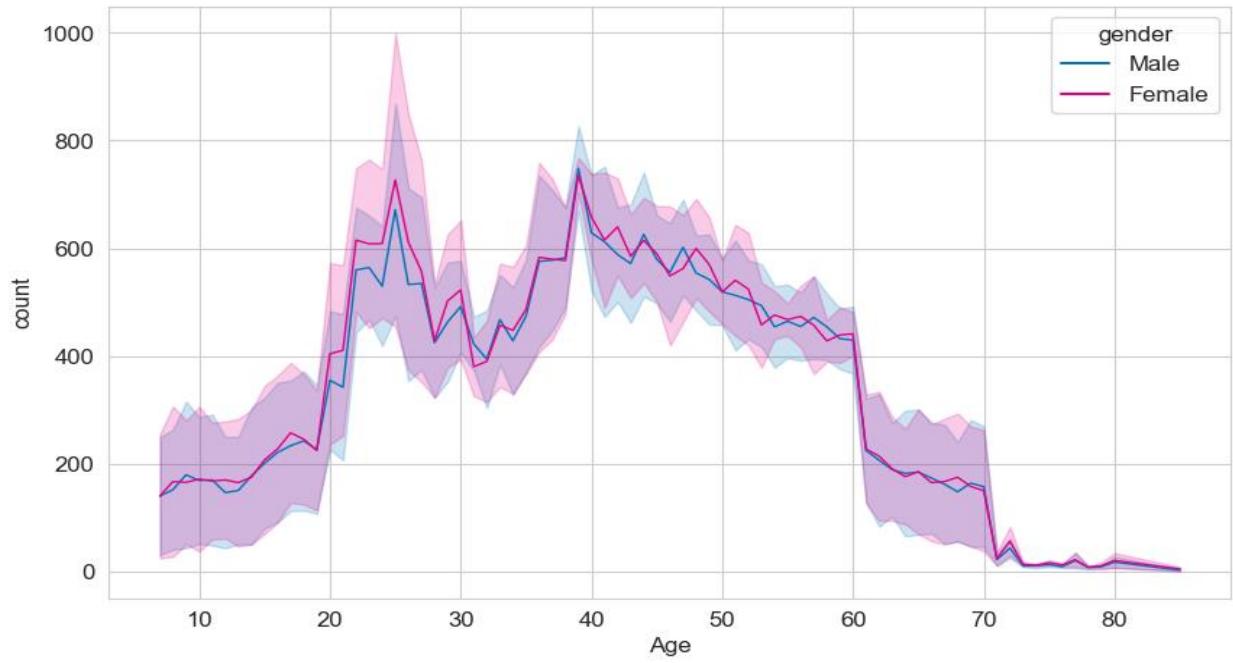
### **Bi-variate Analysis:**

#### **6. Age (Flight Distance and Age):**



- The average age of customers is around 39 years old. This indicates that the airline's customer base is relatively evenly distributed across different age groups.
- Customers under 30 years old tend to fly a distance of around 2000 miles. However, it's important to note that there are significant outliers, suggesting that some younger customers may opt for longer-distance flights for various reasons such as travel purposes or personal preferences.
- Customers over the age of 75 show a preference for local flights, with a notable number of outliers. This could be due to factors such as proximity to home, health considerations, or personal preferences of older individuals.
- There is a prominent pattern where a distance of 4000 miles is frequently observed as the typical flight distance. This suggests that the airline may have a license or operational limitations that restrict flights beyond this distance. The presence of outliers indicates that there may be occasional long-haul flights, but they are not as common.
- The majority of customers fall within the 25-60 age range, indicating that this demographic represents a significant portion of the airline's customer base. This age group is likely to be a key target audience for the airline's marketing and service strategies.
- Understanding the relationship between age and flight distance can help the airline tailor its services and marketing efforts to different customer segments. By offering a range of flight options that cater to different age groups and their travel preferences, the airline can enhance customer satisfaction and attract a diverse customer base.

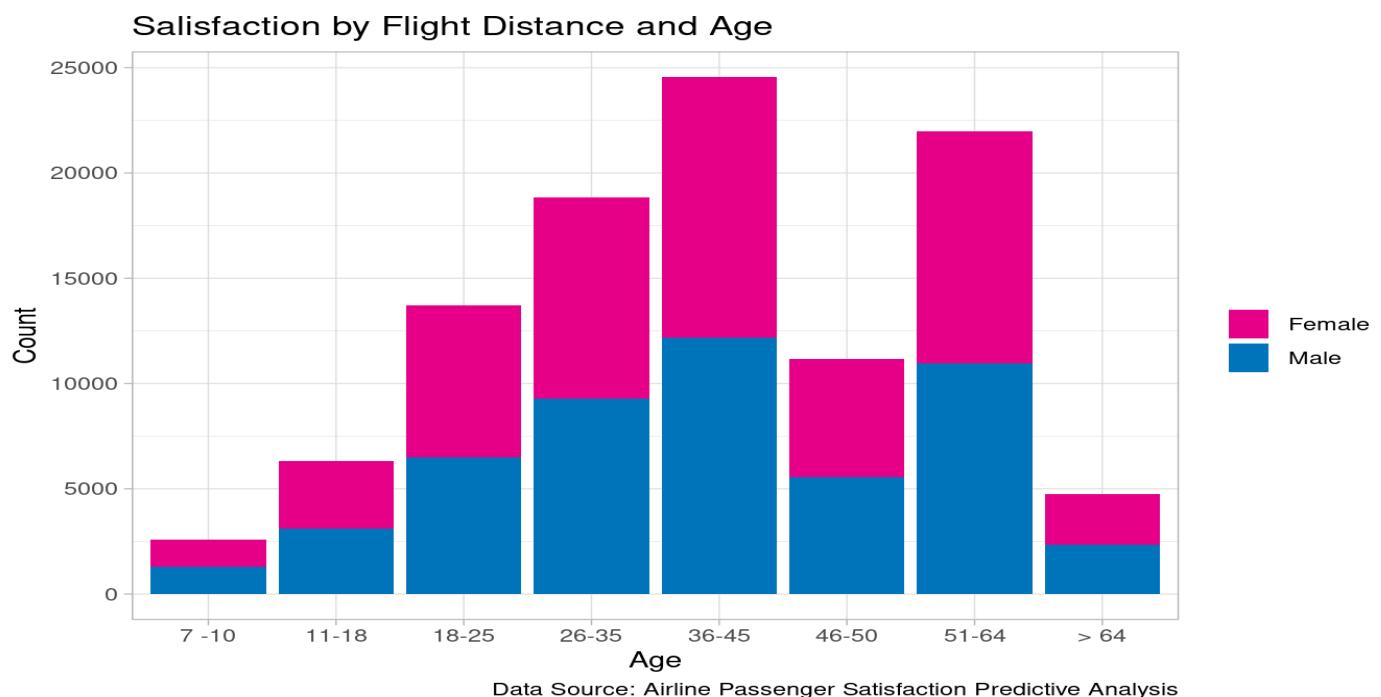
## **7. Satisfaction by Flight Distance and Age:**



- Customers who are under 20 years old are more likely to be dissatisfied with the quality of the flying service. This suggests that the airline may need to pay attention to the specific needs and expectations of this age group in order to improve their satisfaction.
- Both women and men show a peak in "No" responses (indicating dissatisfaction) between the ages of 20 and 30. However, during this age range, women are more likely than men to be dissatisfied with the flying service. This highlights a potential gender difference in perception or experiences with the airline's services.

These insights indicate that the airline should focus on addressing the concerns and improving the flying experience for customers in their early twenties. Additionally, the airline should pay attention to potential gender differences in satisfaction levels and take steps to enhance the flying experience for women in the 20-30 age range. By addressing these specific areas of dissatisfaction, the airline can work towards improving overall customer satisfaction and loyalty among younger customers.

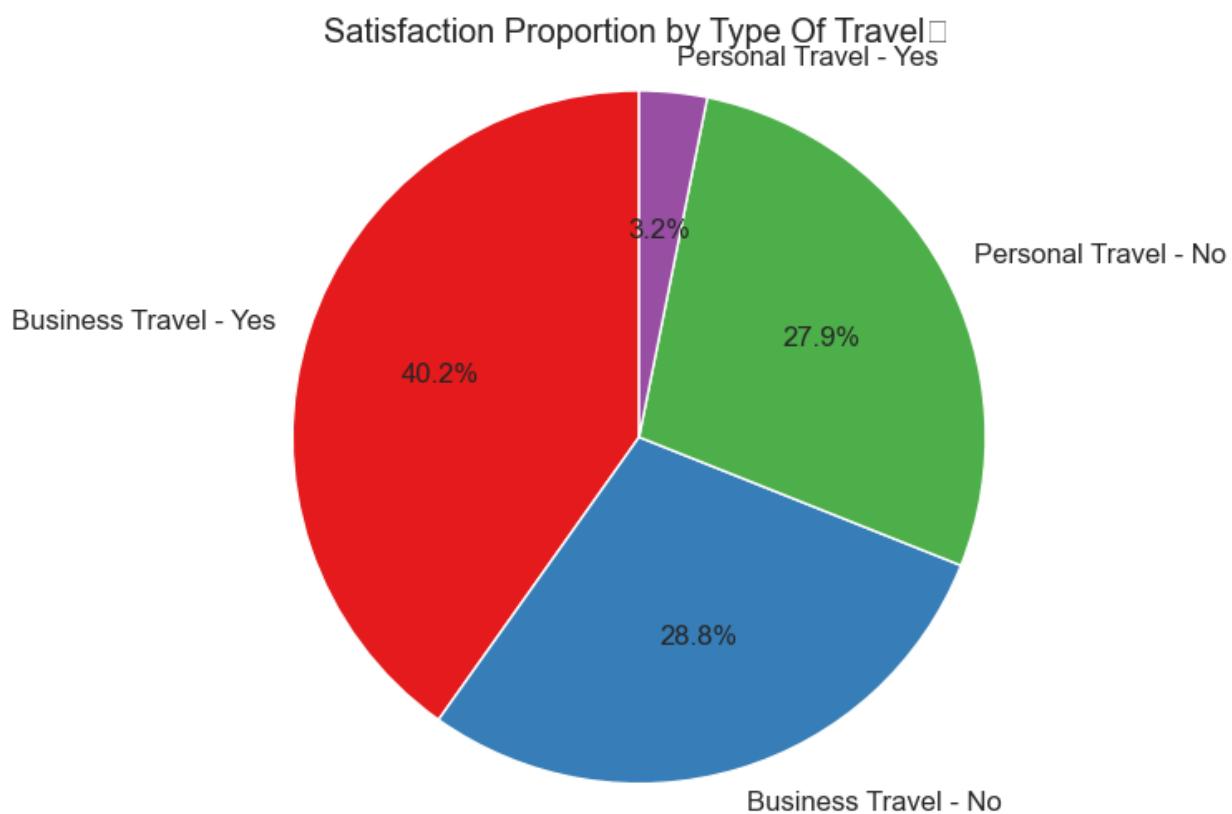
## **8. Age Group:**



- It is interesting to observe that there are approximately 50% more women than men in each age range. This suggests a potential gender imbalance in the customer base of the airline.

- This finding could be attributed to various factors such as travel preferences, marketing strategies, or even social and cultural factors. Understanding and addressing this gender imbalance can provide valuable insights for the airline to tailor its services, marketing campaigns, and customer engagement efforts to better cater to the needs and preferences of its female customers. By recognizing and responding to the unique requirements of this demographic, the airline can potentially attract more male customers and achieve a better gender balance in its customer base.

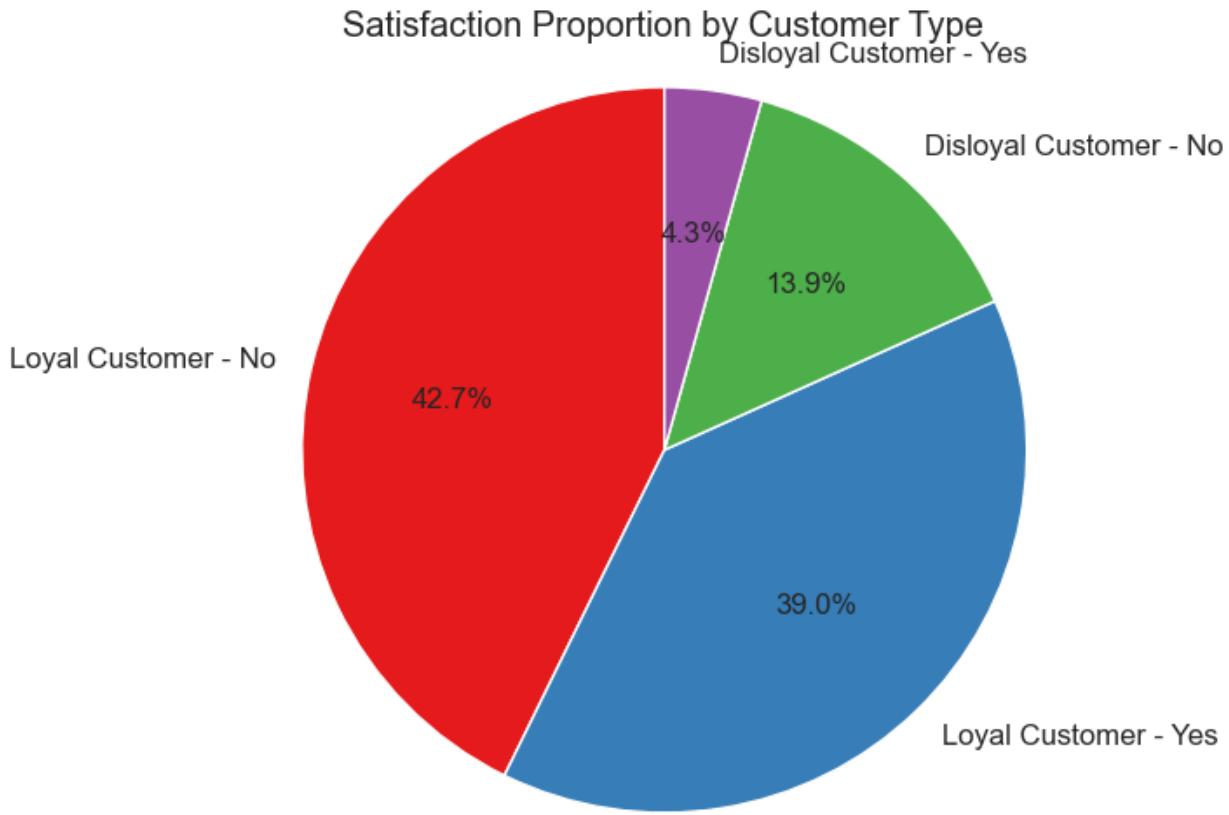
## 9. Satisfaction by type of travel:



- Passengers traveling for business purposes have a significantly higher satisfaction rate (40.2%) compared to those traveling for personal reasons (3.2%).
- This suggests that business travelers may have specific needs and expectations that are better met by the airline, leading to a higher satisfaction level.
- Interestingly, passengers who are not traveling for business still show a relatively high satisfaction rate (28.8%) compared to personal travelers (27.9%).

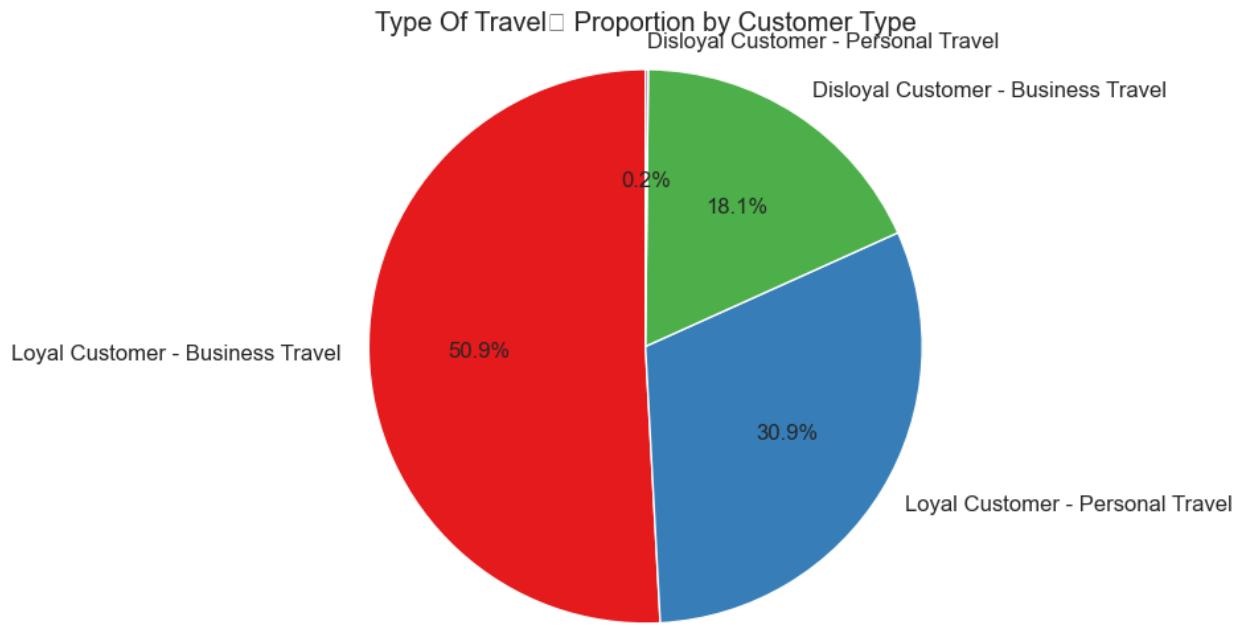
- This indicates that airlines are successful in catering to the satisfaction of both business and non-business travelers, although the satisfaction level may vary depending on the purpose of travel.

## 10. Satisfaction by Customer Type:



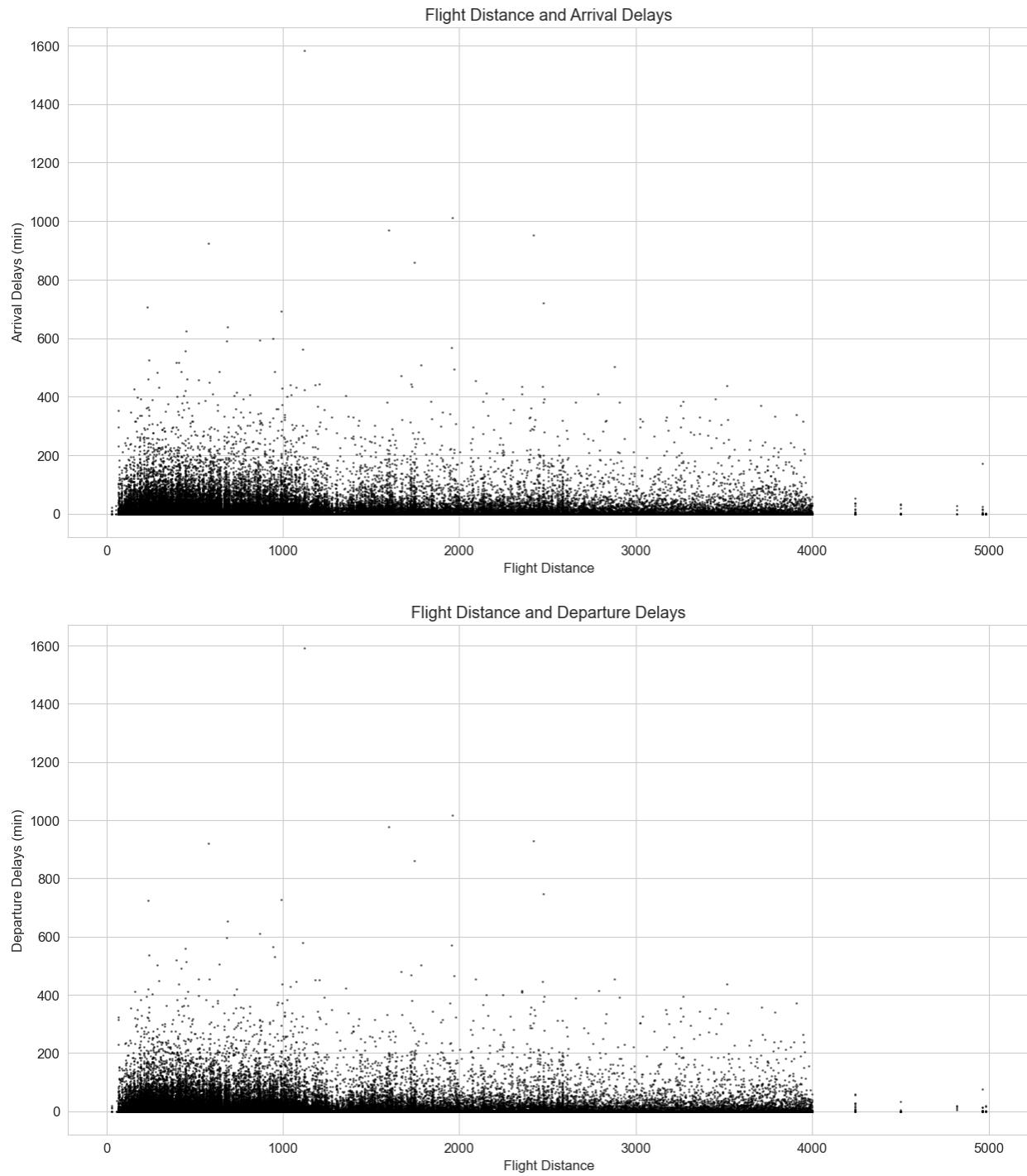
- 
- Loyal customers have a slightly lower satisfaction rate compared to non-loyal customers, indicating that loyalty does not guarantee higher satisfaction.
  - Non-disloyal customers have a significantly higher satisfaction rate compared to disloyal customers, suggesting that loyalty plays a role in overall satisfaction levels.
  - The satisfaction rate among disloyal customers is considerably low, highlighting potential issues or dissatisfaction that may contribute to their lack of loyalty.
  - Non-loyal customers have a relatively higher satisfaction rate compared to loyal customers, indicating that factors other than loyalty may influence their satisfaction levels, such as price, service quality, or other personal preferences.

## **11. Proportion by customer type:**



- Among loyal customers, 50.9% of them travel for business purposes, while 30.9% travel for personal reasons. In contrast, among disloyal customers, 18.1% travel for business purposes, and only 0.2% travel for personal reasons.
- Insight: Loyal customers are more likely to travel for business purposes compared to disloyal customers. This suggests that business travelers tend to have higher loyalty towards the airline. Conversely, disloyal customers are less likely to travel for personal reasons, indicating a potential area for improvement in catering to their needs and preferences.

## **12. Flight distance and Delays:**



We can infer from the scatter plots and results of the preceding "lm" model that the arrival delay is inversely proportional to the departure delay but has no bearing on the flying distance.

## **6. CONCLUSION**

In this project, we conducted an analysis of an airline passenger satisfaction dataset to gain insights into customer preferences, satisfaction levels, and demographic patterns. Several key findings emerged from our analysis:

1. **Gender Distribution:** The dataset revealed a significant gender imbalance, with approximately 50% more women than men in each age range. This highlights the importance of understanding and catering to the specific needs and preferences of female customers to achieve a better gender balance in the customer base.
2. **Age and Travel Patterns:** The majority of customers fall within the 25-60 age range, indicating a strong representation of middle-aged adults. Younger customers under the age of 30 tend to travel shorter distances, while older customers over the age of 75 prefer local flights. Understanding these age-related travel patterns can help the airline tailor its services and marketing strategies accordingly.
3. **Class Preferences:** Business class is the most popular choice among customers, accounting for 47.8% of bookings. Economy Plus and Economy classes are also preferred, with 7.2% and 45% of bookings, respectively. This suggests that the airline should focus on providing high-quality services and amenities across all classes to meet the diverse needs of its customers.
4. **Customer Loyalty:** The analysis revealed that approximately 82% of customers are loyal to the airline, while 18% are categorized as disloyal. This emphasizes the importance of customer retention strategies and building strong customer relationships to enhance loyalty and drive repeat business.
5. **Satisfaction Levels:** The analysis showed that there is a correlation between satisfaction levels and various factors such as flight distance, online boarding experience, and seat comfort. Understanding these factors can help the airline identify areas for improvement and enhance overall customer satisfaction.

In conclusion, this project provides valuable insights into customer preferences, satisfaction levels, and demographic patterns in the airline industry. By leveraging these insights, the airline can make informed decisions to improve customer experiences, tailor marketing strategies, and drive customer loyalty.

## **7. LIMITATIONS**

While conducting this project, we encountered several limitations that should be taken into consideration:

1. **Dataset Limitations**: The analysis is based on a specific dataset, which may not be fully representative of the entire population of airline passengers. The dataset might not capture the diversity of customer preferences and behaviors across different regions or airlines. Additionally, the dataset might be limited in terms of the number of variables available for analysis, which could restrict the depth of insights that can be derived.
2. **Sample Bias**: The dataset used in the analysis might have inherent biases due to factors such as self-selection or sampling methods. This could introduce sample bias and affect the generalizability of the findings to the broader population. It is important to consider the potential biases and limitations of the dataset when interpreting the results.
3. **Causality and Confounding Variables**: The analysis primarily focuses on identifying associations and correlations between variables. While these findings provide valuable insights, they do not establish causality. There may be confounding variables or unobserved factors that influence the relationships observed in the analysis. It is crucial to interpret the results with caution and consider alternative explanations.
4. **Data Quality**: The accuracy and completeness of the dataset can impact the reliability of the analysis. Incomplete or erroneous data points could introduce errors or biases in the results. It is important to ensure the data used for analysis is of high quality and to address any missing or inconsistent data appropriately.
5. **External Factors**: The analysis does not account for external factors that could influence customer preferences and satisfaction, such as macroeconomic conditions, competitive landscape, or specific events or incidents during the data collection period. These external factors could have an impact on the findings and should be considered in the broader context.
6. **Time Limitation**: The analysis is based on a specific timeframe, and customer preferences and behaviors might evolve over time. It is important to consider the temporal aspect of the analysis and recognize that the findings may not be applicable or valid in the future.

These limitations should be kept in mind when interpreting the results and drawing conclusions from the analysis. It is advisable to conduct further research and analysis to validate the findings and address these limitations for a more comprehensive understanding of customer satisfaction in the airline industry.

## **8. BUSINESS INSIGHT**

Based on the analysis conducted on the airline passenger satisfaction dataset, the following key business insights can be derived:

1. **Customer Segmentation:** The analysis revealed distinct patterns and preferences among different customer segments. Understanding customer segmentation, such as loyal customers versus disloyal customers, business travelers versus personal travelers, and different age groups, can help airlines tailor their services and marketing strategies to meet the specific needs and expectations of each segment.
2. **Service Improvements:** The analysis highlighted areas where customer satisfaction is relatively low, such as in-flight entertainment, seat comfort, and cleanliness. These insights provide valuable opportunities for airlines to focus on improving these aspects of their service offerings to enhance customer satisfaction and loyalty.
3. **Importance of Timeliness:** The strong correlation between departure delay and arrival delay underscores the importance of efficient and timely operations. Airlines should prioritize minimizing delays and disruptions to ensure a positive travel experience for their customers.
4. **Customer Satisfaction and Loyalty:** The analysis showed that customer satisfaction and loyalty are crucial factors for airlines. With nearly 82% of customers being categorized as loyal, it is evident that retaining and nurturing customer loyalty should be a key business objective. Providing exceptional customer service, personalized experiences, and rewards programs can help foster customer loyalty and advocacy.
5. **Gender-based Differences:** The analysis revealed gender-based differences in satisfaction levels. Understanding these differences can help airlines design targeted marketing campaigns and services that address the specific preferences and needs of both male and female passengers.

Overall, the business insights derived from this analysis can guide airlines in making strategic decisions to enhance customer satisfaction, improve service quality, and differentiate themselves in a competitive market. By tailoring their offerings to meet customer expectations and preferences, airlines can build stronger customer relationships and drive business growth.

## **Insights based on Exploratory Data Analysis (EDA):**

1. **Age Distribution:** The age distribution of the customers shows a relatively balanced representation across different age groups. The majority of customers fall within the age range of 25 to 60 years, indicating a diverse customer base.
2. **Class Preference:** The analysis revealed that a significant proportion of customers (47.8%) chose Business class, while Economy Plus and Economy classes were preferred by 7.2% and 45% of customers, respectively. This highlights the importance of offering a range of class options to cater to different customer preferences and budgets.
3. **Travel Purpose:** The data suggests that the majority of customers (69%) travel for business purposes, while 31% travel for personal reasons. This insight can help airlines understand the primary motivations of their customers and tailor their services accordingly.
4. **Customer Loyalty:** The analysis showed that approximately 82% of customers are categorized as loyal customers, while 18% are classified as disloyal customers. This highlights the significance of customer loyalty in the airline industry and the need for strategies to retain and strengthen customer loyalty.
5. **Flight Distance:** The analysis revealed that the majority of flights fall within the range of 400 to 3000 kilometers. This indicates that customers generally prefer shorter or medium-haul flights, and long-distance flights are less favored. Airlines can take this into consideration when planning routes and optimizing their flight offerings.
6. **Gender Distribution:** The dataset showed that approximately 51% of customers are women, while 49% are men. This gender distribution insight can inform marketing strategies and customer service initiatives to ensure that the specific needs and preferences of both genders are addressed.

These insights derived from the Exploratory Data Analysis provide a foundational understanding of the dataset and can serve as valuable inputs for decision-making processes in the airline industry. They offer useful insights into customer preferences, behavior, and demographics, which can help airlines optimize their services, improve customer satisfaction, and drive business growth.

## **9. CLOSING REFLECTION**

In conclusion, this project involved conducting an in-depth analysis of airline passenger satisfaction based on a comprehensive dataset. Through various analytical techniques and visualizations, we gained valuable insights into customer demographics, preferences, and satisfaction levels.

The project provided a holistic view of the factors influencing passenger satisfaction, including age, gender, class preference, travel purposes, loyalty, flight distance, and more. By examining these aspects, we were able to identify patterns, trends, and correlations that can guide strategic decision-making in the airline industry.

However, it's important to acknowledge that this analysis has its limitations. The findings are based on the available dataset, which may not fully capture all factors and nuances impacting passenger satisfaction. Additionally, as with any data analysis, there may be inherent biases or limitations in the data collection process that could affect the accuracy of the insights.

Nevertheless, this project serves as a starting point for further investigation and exploration. It highlights the need for continuous monitoring of customer satisfaction, understanding evolving customer preferences, and tailoring services to meet their needs. The insights gained can assist airlines in enhancing their offerings, optimizing operations, and ultimately improving the overall passenger experience.

In summary, this project has provided valuable insights into airline passenger satisfaction, paving the way for further analysis and actions aimed at enhancing customer satisfaction and driving business success in the competitive airline industry.

# THANK YOU



+91 888-290-9944

[Info@i3infosoft.com](mailto:Info@i3infosoft.com)

[www.i3infosoft.com](http://www.i3infosoft.com)

A-77, Ground Floor, Sector 2,  
Near NoidaSector 15 Metro  
Station, Noida,  
Uttar Pradesh, India 201301

Create by:

