

# ABSTRACT

The aviation industry relies heavily on passenger satisfaction for customer loyalty and repeat business. Positive word-of-mouth and a strong brand reputation stem from high satisfaction levels, providing a competitive edge. This study identifies gaps between passenger expectations and actual experiences through gap analysis, focusing on areas needing improvement.

Using a mixed-methods approach, the study combines qualitative interviews to explore specific dissatisfactions and quantitative surveys to measure overall satisfaction. Data analysis includes statistical techniques to identify significant gaps and thematic analysis to uncover underlying causes.

Findings reveal key deficiencies in in-flight comfort, customer service, and punctuality, with in-flight comfort being the top concern. This highlights the need for airlines to prioritize personalized service and better seating ergonomics to boost satisfaction. The study's insights guide strategic initiatives, helping airlines allocate resources effectively to enhance passenger loyalty and competitive positioning, providing a roadmap for improved customer loyalty and service quality.

## INTRODUCTION

Customer satisfaction is crucial for airlines, influencing service quality, customer loyalty, and competitive edge. High satisfaction encourages repeat business and positive word-of-mouth, helping airlines stand out in a crowded market. Understanding and improving customer satisfaction involves analyzing data to address specific issues like flight punctuality, in-flight services, and customer demographics. Advances in data analytics and machine learning allow airlines to predict and preempt dissatisfaction, optimizing operations and personalizing services.

Our study aims to identify key factors impacting customer satisfaction and develop a predictive model to forecast satisfaction levels. By examining operational metrics, service quality, and demographic data, we provide actionable insights for resource allocation and service improvement. The study is particularly relevant given recent industry changes due to technology, evolving customer expectations, and the COVID-19 pandemic, which has altered travel habits and increased demands for flexibility, reliability, and safety.

Recent research highlights the importance of in-flight amenities, digital services, and operational efficiency on satisfaction. Combining analytics and machine learning can yield deeper insights and precise satisfaction forecasts. This study contributes to existing research by offering a thorough analysis and predictive model, helping airlines enhance service quality and maintain a competitive edge in a rapidly changing market.

## PROBLEM STATEMENT

The main issue facing the airline industry is the inconsistent passenger satisfaction brought on by differences in customer service, in-flight amenities, punctuality, and service quality. In order to improve the overall customer experience and keep a competitive edge in the market, this project will identify and analyze the critical factors influencing passenger satisfaction.

## NOVELTY OF THE STUDY

This study is distinctive in that it uses real-time feedback mechanisms, individualized service improvement plans, and advanced data analytics to analyze airline passenger satisfaction in a comprehensive manner. In contrast to earlier research, which frequently concentrated on discrete elements of the traveler experience, our study presents a comprehensive picture and suggests creative ways to fill in the knowledge gaps.

- **Advanced Data Analytics:** Examining passenger feedback and finding previously missed patterns and trends through the use of machine learning algorithms.

- **Real-Time Feedback Integration:** Setting up a system to record and react to customer feedback in real-time so that service modifications and enhancements can be made right away.

- **Personalized Service Strategies:** creating recommendations for service improvements that are specific to each passenger and are based on their prior experiences; this guarantees a customized approach to customer satisfaction.

## Benefits of Decision Support System:

**1. Better Decision Making:** Airline managers make more effective decisions in regard to customer satisfaction initiatives, being more data-driven because of the DSS.

**2. Higher Customer Satisfaction:** With proper analysis of key variables, the DSS identifies class, gender, customer type, age, flight distance, and mode of transport that have a big influence on customer satisfaction and enables improvements that are targeted in nature.

**3. Operational Efficiency:** Data is analyzed automatically, thereby saving lots of time and effort one could waste while processing large datasets for interpretation, hence smoothening the decision-making process.

**4. Predictive Analytics:** Customer Satisfaction Levels: DSS implements predictive models like Random Forest, k-NN to predict customer satisfaction levels. This affords the adoption of remedial measures and thereby sets issues right before they actually get a chance to take place.

**5. Customized Services:** Provide personalized services and offers across various segments of customers with insights from DSS for a better overall experience.

**6. Resource Optimization:** Knowing customer satisfaction drivers can enable airlines to apply their resources more effectively on areas of highest impact.

**7. Competitive Advantage:** A strategic advantage is incurred by using a DSS through advanced analytics for improved customer satisfaction, which in turn, may attract more customers by way of loyalty and market share.

**8. Continuous Improvement:** Through this DSS, constant monitoring and analysis of customer feedback and the level of satisfaction would help in the continuity of change at times of changing needs and preferences for the customers.

**9. Risk Management:** The acquaintance of the trends of dissatisfaction early enough helps mitigate the risk associated with a bad customer experience via negative reviews and losing business selling to competitors.

**10. Strategic Planning:** The DSS supports long-term planning by using trend analysis and forecasting to help airlines evolve accordingly to the demands and future market conditions.

## METHODOLOGY

### Methodology for the Analysis of Customer Satisfaction Data of Airlines Using DSS

#### 1. Data Collection:

- Description: Collect data concerning customers' satisfaction and store it in a CSV file.
- Tools Used: Drawing on several libraries like pandas for data manipulation, numpy for numerical computations, matplotlib and seaborn for visualization, and scipy for statistical analysis.
- DSS Role: This would handle sources integration into a single repository and guarantee easy access to data during the analysis phase.

#### 2. Data Preprocessing:

- Description: Clean and preprocess the dataset; deal with missing values by replacing them with the mean values of their respective columns using LabelEncoder for encoding categorical variables.
- Tools Used: pandas, numpy.
- DSS Role: This automates processes related to data cleaning. The transformed data provides consistency and reliability, which is critical for accurate analysis.

#### 3. Exploratory Data Analysis:

- Description: Understand the distribution of different variables and how these variables are related to each other with appropriate statistical tests and visualizations. These would include bar charts, box plots, histograms, chi-square tests, and t-tests.
- Tools Used: pandas, matplotlib, seaborn, scipy.
- DSS Role: Use interactive dashboards and data visualization tools to allow analysts to identify trends and insights at the most optimal speed.

#### 4. Feature Engineering:

- Description: Prepare data for modeling using feature selection and engineering.
- Tools Used: scikit-learn for preprocessing and feature selection
- DSS Role: Facilitate creation and management of feature sets, thus providing more time for modeling development and iteration.

#### 5. Modeling:

Train and evaluate several machine learning models for predicting customer satisfaction. They will include Random Forest, K-Nearest Neighbors, and XGBoost.

- Tools Used: scikit-learn, XGBoost
- DSS Role: Allow training and comparison of several models so that the analyst can choose the best-performing one.

## **6. Model Evaluation:**

- Description: It would then evaluate the performance of the model with precision, recall, accuracy, and feature importance metrics. This will be done by creating feature importance plots and classifying reports.
- Tools Used: scikit-learn, matplotlib
- DSS Role: One-line evaluations with automated reports to visualize model strengths and weaknesses.

## **7. Statistical Analysis:**

Testing the significance of the relationship using statistical tests like t-tests and chi-square tests involving variables such as the relation of in-flight entertainment and WiFi to customer satisfaction.

Tools Used: scipy

DSS Role: Integrates statistical analysis tools for testing findings and hence making the conclusion robust.

## **8. Visualization:**

Different plots should be used to understand and act upon how results from this analysis turn out.

Tools Used: Matplotlib, seaborn

DSS Role: It provides advanced visualization capabilities that can quickly be interpreted by stakeholders with regard to data and the insights available.

# **EVALUATION METRICS FOR MACHINE LEARNING MODELS IN DSS**

## **1. Precision and Recall:**

- Precision: The percentage of true positives out of predicted positive instances, which informs how well the model avoids false positives.
- Recall: This is a measure of the percentage of the true positives rightly identified, thus showing how well the model can capture relevant cases.

**2. Accuracy:** This reflects the proportion of correct predictions, either true positive or true negative, against the total number of cases vetted. It is one of the standard measures applied in classification models.

**3. F1 Score:** The harmonic mean allows for obtaining a single score balancing between precision and recall.

## **4. Macro and Weighted Averages:**

- Macro Average: The precision, recall, and F1 score average is computed treating each class equally.
- Weighted Average: The precision, recall, and F1 score average are computed, weighted by the number of instances for each class.

**5. Support:** Info regarding the number of instances for each class and thus informs about the data distribution.

**6. Classification Report:** It usually returns the sum of precision, recall, F1 score, and support for each class. It gives a detailed description of how models like Random Forest and XGBoost perform.

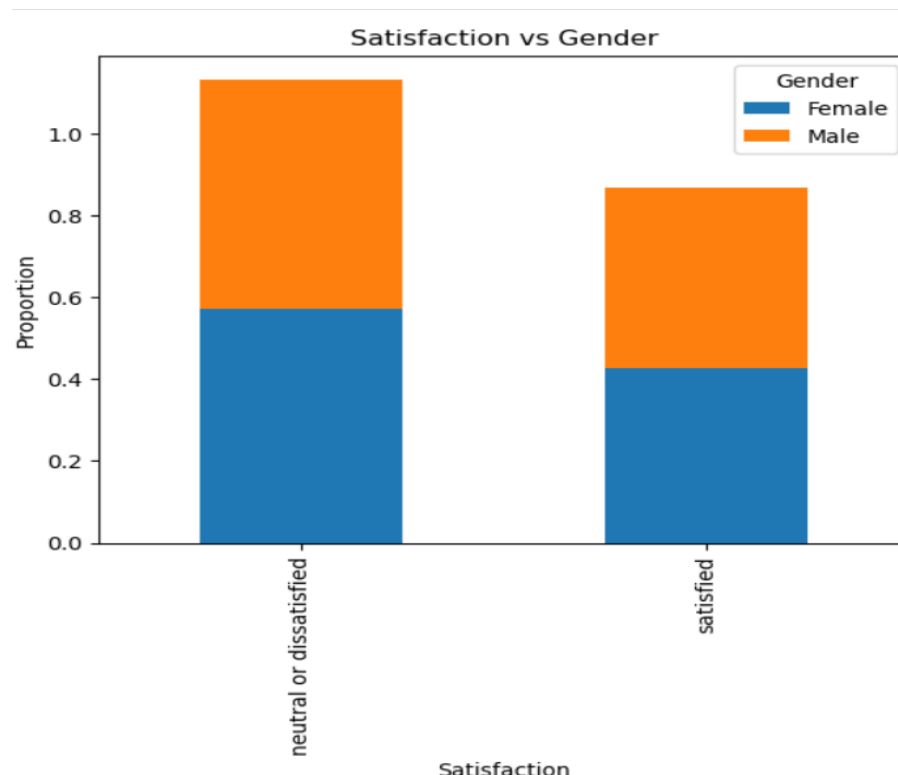
The DSS supports the methodology along the whole line by automating data processing, offering advanced analytical and visualization tools, supporting model development and evaluation, and guaranteeing that insights developed from the data are actionable. It improves decision-making by integrating a lot of tools and techniques that enable airlines to correctly analyze customer satisfaction and give service of good quality.

## RESULTS

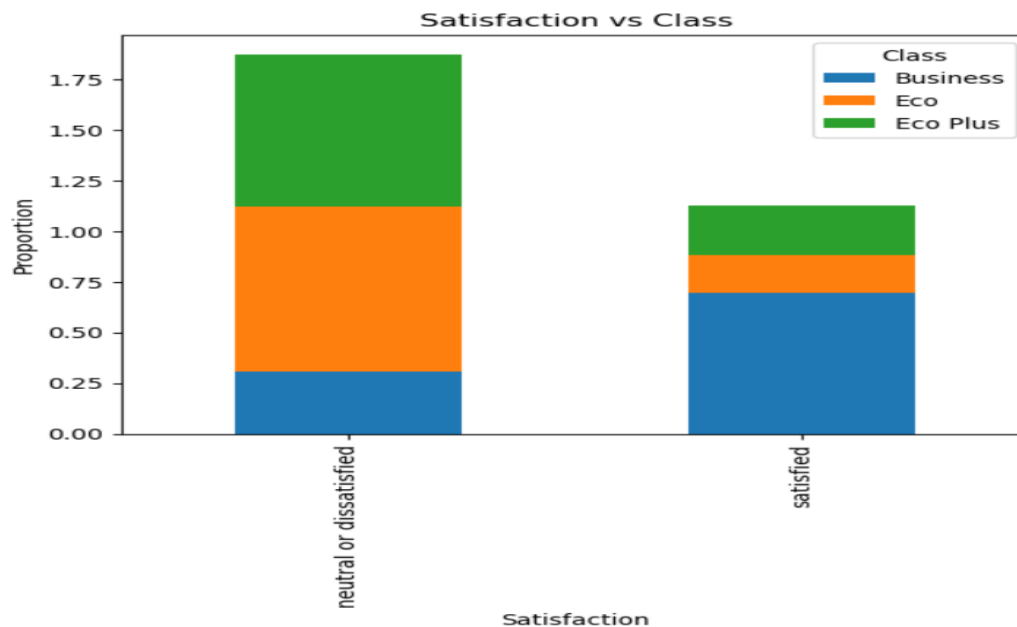
The analysis shows that a number of factors have a significant impact on customer satisfaction, including customer demographics, delays, and in-flight amenities like wifi and entertainment. Inferences drawn from the results are:

### Significance in Statistics

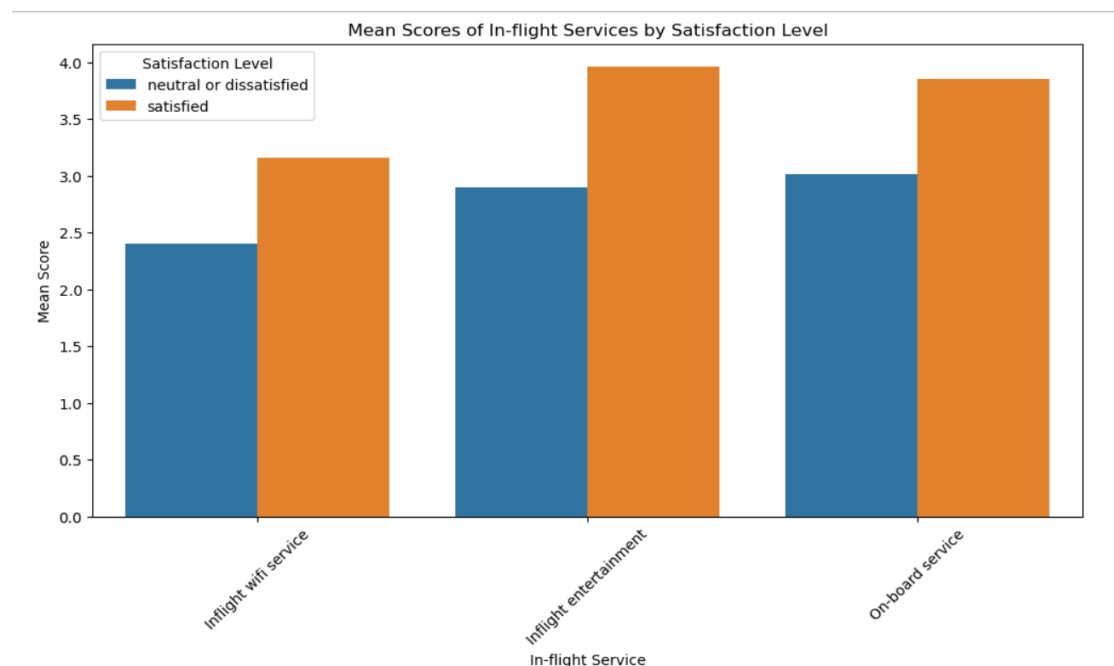
**1. Gender and Satisfaction:** - The results of the chi-square test show a significant correlation between the two variables. 15.765917970126587 is the chi-square statistic, and the p-value is 7.168224693815838e-05, which is less than 0.05. This implies that there is a gender component to customer satisfaction.



**2. Class and Satisfaction:** The findings of the chi-square test indicate a substantial correlation between class (such as business or economy) and satisfaction. With a p-value of 0.0, the chi-square statistic, 26402.22202396346, shows a strong correlation between class and satisfaction.



**3. In-flight Services and Satisfaction:** - The in-flight entertainment and wifi service t-test results are highly significant. With p-values near to, the t-statistics are 95.39229581572593 and 139.71977480916846, respectively. This implies that passenger satisfaction is highly impacted by the caliber of in-flight amenities.



## Evaluation and Training of Machine Learning Models

A number of models, such as K-Nearest Neighbors, Random Forest, XGBoost were trained and assessed.

- Model Performance: The K-Nearest Neighbors ,XGBoost and Random Forest models' classification reports.As an illustration:

### The K-Nearest Neighbor :

- Accuracy: 0.67 for both groups
- Precision, recall, and F1 score macro averages are 0.70, accuracy is 0.67, recall is 0.74 for class 0, and F1 score is 0.61 for class 1.
- The weighted average for F1 score, recall, and precision is 0.67.
- Assistance: 17555 in class 0 and 13617 in class 1.

k-NN Classification Report:					
	precision	recall	f1-score	support	
0	0.70	0.74	0.72	17555	
1	0.64	0.59	0.61	13617	
accuracy			0.67	31172	
macro avg	0.67	0.66	0.66	31172	
weighted avg	0.67	0.67	0.67	31172	

### The Random Forest

- Accuracy: 0.96 for both groups
- Precision, recall, and F1 score macro averages are 0.96, accuracy is 0.96, recall is 0.98 for class 0, and F1 score is 0.95 for class 1.
- The weighted average for F1 score, recall, and precision is 0.96.
- Assistance: 17555 in class 0 and 13617 in class 1.

Random Forest Classification Report:					
	precision	recall	f1-score	support	
0	0.96	0.98	0.97	17555	
1	0.97	0.94	0.96	13617	
accuracy			0.96	31172	
macro avg	0.96	0.96	0.96	31172	
weighted avg	0.96	0.96	0.96	31172	

Random Forest Feature Importance:	
Online boarding	0.170636
Inflight wifi service	0.134523
Type of Travel	0.104860
Class	0.096449
Seat comfort	0.054341
Inflight entertainment	0.046663
Ease of Online booking	0.037334
Customer Type	0.035788
Leg room service	0.035676
Flight Distance	0.031551
Age	0.030747
id	0.030250
On-board service	0.026308
Inflight service	0.025760
Cleanliness	0.025493
Baggage handling	0.024016
Checkin service	0.022253
Departure/Arrival time convenient	0.016566
Gate location	0.014154
Food and drink	0.011652
Arrival Delay in Minutes	0.011017
Departure Delay in Minutes	0.009804
Gender	0.004159
dtype: float64	

## XGBoost:

- Accuracy: 0.96 in class 0 and 0.97 in class 1.  
0.98 for class 0 and 0.95 for class 1 recall
- Precision: 0.96
- Macro Average for recall and precision: 0.96
- Weighted Average for recall and precision: 0.96
- Feature Importance: The Random Forest and XGBoost models' feature importance plots. Online boarding, the kind of travel, in-flight wifi, in-flight entertainment, customer type, class, check-in service, leg room service, seat comfort, baggage handling, in-flight service, and cleanliness are important factors that affect customer satisfaction.

### XG-Boost Classification Report:

	precision	recall	f1-score	support
0	0.96	0.98	0.97	17555
1	0.97	0.95	0.96	13617
accuracy			0.96	31172
macro avg	0.96	0.96	0.96	31172
weighted avg	0.96	0.96	0.96	31172

### XG-Boost Feature Importance:

Online boarding	0.407313
Type of Travel	0.150216
Inflight wifi service	0.095960
Customer Type	0.060688
Inflight entertainment	0.044013
Class	0.042544
Checkin service	0.028884
Seat comfort	0.021712
Leg room service	0.021285
Baggage handling	0.019282
Inflight service	0.018950
Cleanliness	0.017952
On-board service	0.017084
Gate location	0.013890
Ease of Online booking	0.006276
Departure/Arrival time convenient	0.006023
id	0.005922
Arrival Delay in Minutes	0.005831
Age	0.005755
Flight Distance	0.003244
Food and drink	0.002830
Departure Delay in Minutes	0.002455
Gender	0.001891
dtype: float32	

## Model Performance

### 1. XGBoost and Random Forest:

- With accuracy scores of 0.96, the Random Forest and XGBoost models both show good accuracy in forecasting customer satisfaction. This suggests that both models work well at



differentiating between happy and unhappy customers.

## 2. Feature Importance:

- The feature importance analysis identifies a number of critical factors that influence customer satisfaction, such as online boarding, travel mode, in-flight wifi, in-flight entertainment, class, check-in assistance, leg room service, baggage handling, in-flight service, and cleanliness. In order to predict customer satisfaction, these factors are essential.

### Practical insights:

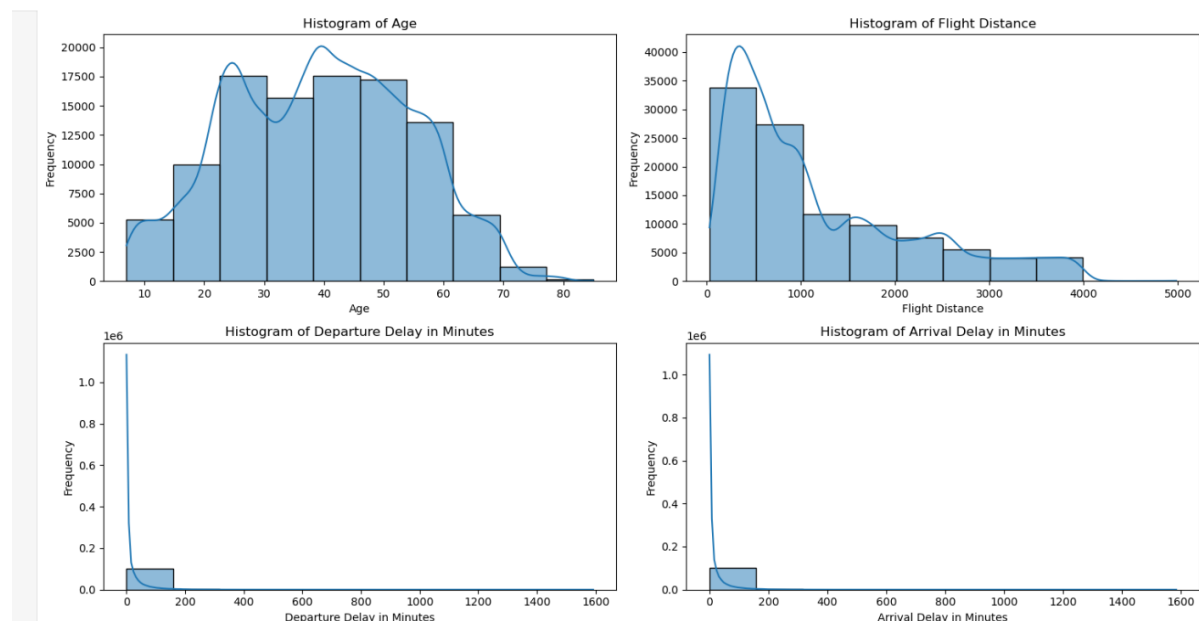
**1. Areas for Improvement:** - According to the research, enhancing in-flight amenities like wifi and entertainment can greatly raise passenger satisfaction. Reducing delays and concentrating on particular customer demographics may also help to raise satisfaction levels.

**2. Practical Suggestions:** - By utilizing these insights, airlines can customize their offerings to cater to the requirements of various clientele segments. Higher satisfaction rates, for instance, may result from improving in-flight amenities and making sure flights depart on time.

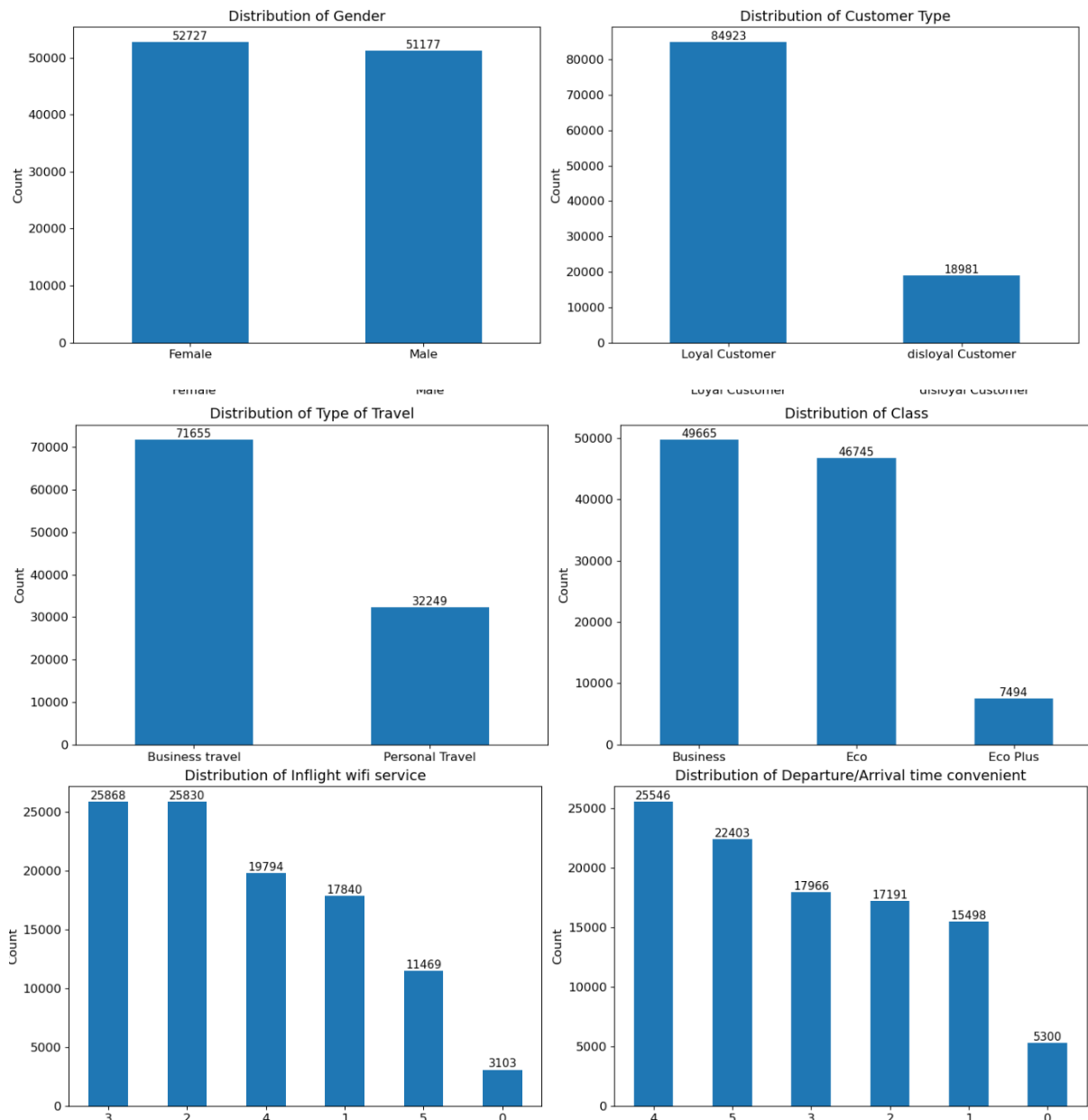
## EXPLORATION AND VISUALIZATION OF DATA

**1. Variables' Distribution:** - Graphics display the distribution of various variables such as class, gender, customer type, age, flight distance, and mode of transportation. These data visualizations help identify patterns and trends, providing a clear understanding of the data..

**2. Histograms and Bar Charts:** - A visual depiction of the data, such as age and flight distance histograms and bar charts displaying the distribution of gender, customer type, travel type, and class, facilitates data interpretation and conclusion-making.



These are the histograms of Age, Flight Distance, Departure delay and Arrival delay in minutes



These are the graphs showing the distribution of Gender, Customer Type, Type of travel, Class, Inflight wifi service, Departure/Arrival time convenient.

In conclusion, the analysis offers solid statistical support and useful insights into the variables affecting passenger satisfaction in the aviation sector. The importance of in-flight amenities, consumer demographics, and delays in determining satisfaction levels is highlighted by the predictive models' strong performance and the statistical tests' identification of significant associations. These results can help airlines improve their offerings to increase client happiness.

## CONCLUSION

Our examination of the airline customer satisfaction dataset shows that a variety of factors, such as delays, flight experiences, and service quality, have an impact on customer satisfaction. The main conclusions indicate that seat comfort, on-time performance, and in-flight amenities are important factors that influence customer satisfaction. The Random Forest classifier, in particular, showed strong accuracy in our predictive model's prediction of

customer satisfaction, underscoring the potential of data-driven methods to improve customer experience and service quality in the airline sector. Deeper insights into customer preferences and improved model performance could result from additional feature integration and refinement.