Exploring Airline Passenger Satisfaction: A Neural Network Approach

Project Report
submitted by
Naveen Murali

ABSTRACT

This study explores the realm of airline passenger satisfaction utilizing a novel approach employing Artificial Neural Networks (ANNs). In the airline industry, understanding and enhancing passenger satisfaction are crucial for maintaining competitiveness and fostering customer loyalty. Traditional methods often fall short in capturing the complex interplay of factors influencing passenger satisfaction. Hence, this research employs ANNs to unravel intricate patterns within extensive datasets to predict and optimize passenger satisfaction levels.

The study begins with a comprehensive review of existing literature on airline passenger satisfaction and the application of ANNs in aviation. Subsequently, a detailed methodology outlines data collection, preprocessing, and the architecture of the ANN model. Key factors influencing passenger satisfaction are identified through the ANN's analysis, shedding light on nuanced variables crucial for predicting satisfaction levels accurately.

Results from the ANN model are presented, highlighting its performance metrics and comparing predicted satisfaction levels with actual passenger feedback. The discussion section delves into the interpretation of results, emphasizing the effectiveness of ANNs in discerning complex patterns and predicting passenger satisfaction more accurately than conventional methods.

Practical implications of the findings for airlines are discussed, providing insights into how airlines can leverage ANN models to enhance passenger satisfaction effectively. The study concludes with reflections on the significance of employing ANNs in understanding and optimizing airline passenger satisfaction, along with recommendations for future research directions.

INTRODUCTION

In the highly competitive airline industry, passenger satisfaction plays a pivotal role in determining the success and sustainability of airlines. As passengers increasingly prioritize their travel experiences, airlines must continuously strive to understand and meet their evolving needs and expectations. However, accurately assessing and enhancing passenger satisfaction is a complex endeavor, often confounded by numerous interconnected factors ranging from service quality to pricing and convenience.

Traditional approaches to studying passenger satisfaction often rely on surveys and statistical analyses, which may overlook subtle patterns and interactions within the data. In recent years, Artificial Neural Networks (ANNs) have emerged as a powerful tool for uncovering intricate relationships in large and complex datasets. ANNs, inspired by the human brain's neural networks, have shown remarkable capabilities in pattern recognition, making them well-suited for analyzing diverse data sources and predicting outcomes.

This project aims to explore airline passenger satisfaction using a Neural Network approach, leveraging the strengths of ANNs to delve deeper into the underlying factors influencing satisfaction levels. By employing ANNs, we seek to overcome the limitations of traditional methodologies and provide airlines with actionable insights to enhance the passenger experience.

In this introduction, we provide an overview of the project's objectives, the significance of studying airline passenger satisfaction, and the rationale behind adopting a Neural Network approach. Additionally, we outline the structure of the project, including the methodology, data sources, and expected outcomes.

GENERAL BACKGROUND

The airline industry is characterized by fierce competition and ever-changing consumer demands. In this context, passenger satisfaction holds immense significance as it directly impacts customer loyalty, brand reputation, and ultimately, the bottom line of airlines. Understanding the factors that influence passenger satisfaction is crucial for airlines seeking to differentiate themselves in a crowded market and retain a loyal customer base.

Passenger satisfaction is influenced by a myriad of factors, including but not limited to:

1. Service Quality: The quality of in-flight services, ground services, and interactions with airline staff significantly impact passenger satisfaction. This encompasses aspects such as cabin crew behavior, cleanliness of aircraft, onboard amenities, and timeliness of service delivery.
2. Flight Experience: Factors related to the overall flight experience, such as seat comfort, legroom, entertainment options, and meal quality, play a vital role in shaping passenger perceptions and satisfaction levels.
3. Booking Process: The ease and convenience of the booking process, including website usability, ticket pricing transparency, and availability of preferred seat options, contribute to overall passenger satisfaction.
4. On-time Performance: Punctuality and reliability are critical factors influencing passenger satisfaction. Delays, cancellations, and disruptions to travel plans can lead to frustration and dissatisfaction among passengers.
5. Customer Service: The responsiveness and effectiveness of customer service in addressing passenger queries, concerns, and complaints have a significant impact on overall satisfaction levels.

While these factors have traditionally been analyzed using conventional statistical methods, the increasing availability of vast and diverse datasets has led to the adoption of more advanced analytical techniques, such as Artificial Neural Networks (ANNs). ANNs offer the capability to uncover complex patterns and relationships within data, enabling more accurate predictions and insights into passenger behavior and satisfaction.

Against this backdrop, this project aims to explore airline passenger satisfaction using a Neural Network approach, leveraging the power of ANNs to unravel the multifaceted nature of passenger preferences and experiences. By doing so, we seek to provide airlines with actionable insights to enhance the passenger journey and drive greater satisfaction and loyalty among passengers.

SCOPE OF THE PROJECT

The scope of the project "Exploring Airline Passenger Satisfaction: A Neural Network Approach" encompasses the following key aspects:

1. Data Collection: The project will involve collecting relevant data pertaining to airline passenger satisfaction from diverse sources, including surveys, customer feedback, and operational data provided by airlines.
2. Data Preprocessing: Prior to analysis, the collected data will undergo preprocessing to clean and prepare it for input into the Neural Network model. This may involve tasks such as data normalization, handling missing values, and feature engineering.
3. Neural Network Architecture: The project will design and implement an appropriate Neural Network architecture tailored to the specific task of predicting passenger satisfaction. This may involve selecting the appropriate type of Neural Network (e.g., feedforward, recurrent, convolutional) and optimizing the network architecture through experimentation.
4. Training and Validation: The Neural Network model will be trained using the prepared dataset, with a portion of the data reserved for validation to assess the model's performance and prevent overfitting.
5. Feature Selection and Importance: Through the Neural Network analysis, the project aims to identify the key factors influencing airline passenger satisfaction. This will involve analyzing the weights assigned to different features by the Neural Network and assessing their importance in predicting satisfaction levels.
6. Prediction and Evaluation: The trained Neural Network model will be used to predict passenger satisfaction levels based on input features. The predictions will be evaluated using appropriate performance metrics, such as accuracy, precision, recall, and F1-score.

7. Interpretation and Insights: The project will interpret the results obtained from the Neural Network analysis, providing insights into the factors driving passenger satisfaction and their relative importance. This will enable airlines to prioritize areas for improvement and strategic interventions.
8. Recommendations: Based on the findings of the analysis, the project will provide actionable recommendations for airlines to enhance passenger satisfaction and improve the overall passenger experience.
9. Limitations: The project will acknowledge and discuss any limitations or constraints encountered during the analysis, such as data availability, model assumptions, and potential biases.
10. Future Research Directions: Finally, the project will suggest potential avenues for future research, including refining the Neural Network model, exploring additional datasets, and investigating emerging trends in airline passenger satisfaction.

By addressing these aspects, the project aims to contribute to a deeper understanding of airline passenger satisfaction and provide valuable insights for airlines to better serve their customers and enhance competitiveness in the industry.

Implementation

1. Data Collection:
    - Gather data on airline passenger satisfaction from various sources such as surveys, customer feedback, and operational databases.
    - Ensure data includes relevant features such as flight details, service quality ratings, booking experience, and demographic information.
2. Data Preprocessing:
    - Clean the collected data to handle missing values, outliers, and inconsistencies.
    - Normalize numerical features to ensure uniform scales and handle categorical variables through one-hot encoding or label encoding.
    - Split the dataset into training, validation, and testing sets.

3. Neural Network Architecture:
    - Select an appropriate Neural Network architecture based on the nature of the problem and the characteristics of the dataset.
    - Design the architecture with input nodes representing features, hidden layers for processing, and output nodes for predicting satisfaction levels.
    - Experiment with different architectures, activation functions, and regularization techniques to optimize performance.
4. Training and Validation:
    - Train the Neural Network using the training dataset and validate its performance using the validation dataset.
    - Monitor metrics such as loss function values and accuracy during training to detect overfitting or underfitting.
    - Adjust hyperparameters such as learning rate and batch size based on validation performance.
5. Feature Selection and Importance:
    - Analyze the learned weights and biases of the Neural Network to identify the importance of different features in predicting passenger satisfaction.
    - Conduct sensitivity analysis to assess the impact of individual features on prediction outcomes.
6. Prediction and Evaluation:
    - Use the trained Neural Network to predict passenger satisfaction levels for the testing dataset.
    - Evaluate the predictions using appropriate performance metrics such as accuracy, precision, recall, and F1-score.
    - Compare predicted satisfaction levels with actual feedback to assess the model's effectiveness.
7. Interpretation and Insights:
    - Interpret the results of the Neural Network analysis to understand the factors driving passenger satisfaction.
    - Provide insights into areas where airlines can focus their efforts to improve satisfaction levels.
    - Generate visualizations and summary statistics to communicate findings effectively.
8. Recommendations:
    - Based on the analysis, offer actionable recommendations for airlines to enhance passenger satisfaction.
    - Prioritize recommendations based on the importance of identified factors and their potential impact on satisfaction levels.

9. Limitations:
   - Acknowledge any limitations or constraints encountered during the implementation, such as data quality issues or model assumptions.
   - Discuss how these limitations may have influenced the results and interpretations.
10. Future Research Directions:
    - Suggest potential avenues for future research, such as incorporating additional data sources or refining the Neural Network model.
    - Identify emerging trends or technologies that could further enhance the understanding and prediction of airline passenger satisfaction.

By following these implementation steps, the project aims to effectively explore airline passenger satisfaction using a Neural Network approach and provide actionable insights for airlines to improve customer experience and satisfaction levels.

Literature Review

The airline industry is inherently customer-centric, with passenger satisfaction being a critical determinant of airlines' success and competitiveness. A comprehensive review of the existing literature reveals various studies exploring the factors influencing airline passenger satisfaction and the emerging role of Artificial Neural Networks (ANNs) in this domain.

1. Factors Influencing Airline Passenger Satisfaction:
   - Numerous studies have identified a wide range of factors influencing airline passenger satisfaction. These factors include service quality (Cronin & Taylor, 1992), flight experience (Chi & Qu, 2008), booking process (Sultan & Simpson, 2000), on-time performance (Formoso & González, 2018), and customer service (Han, 2019).
   - Research has also highlighted the importance of individual preferences, demographics, and travel purposes in shaping passenger satisfaction levels (Bieger & Laesser, 2003; Ebrahim et al., 2016).
2. Traditional Methodologies for Analyzing Passenger Satisfaction:
   - Traditional approaches to studying passenger satisfaction typically involve surveys, statistical analyses, and regression modeling. While these methods provide valuable insights, they may overlook complex nonlinear relationships and interactions among variables.

    - Moreover, traditional methodologies often struggle to handle large and diverse datasets, limiting their ability to capture the multifaceted nature of passenger satisfaction (Ouyang et al., 2020).
3. Application of Artificial Neural Networks in Aviation Research:
    - In recent years, there has been a growing interest in leveraging Artificial Neural Networks (ANNs) to analyze airline passenger satisfaction. ANNs, inspired by the human brain's neural networks, offer the capability to uncover intricate patterns and relationships within vast and complex datasets.
    - Studies have demonstrated the effectiveness of ANNs in predicting passenger satisfaction levels and identifying key factors influencing satisfaction outcomes (Park et al., 2018; Zhang et al., 2020).
    - ANNs have been applied to various aspects of airline operations, including demand forecasting, revenue management, and customer sentiment analysis, highlighting their versatility and utility in aviation research (Bulut & Isiklar, 2017; Homaifar & Schilling, 1994).
4. Challenges and Opportunities:
    - Despite the promise of ANNs in analyzing passenger satisfaction, challenges such as data quality, model interpretability, and computational complexity remain. Addressing these challenges requires interdisciplinary collaboration and methodological innovations.
    - However, the proliferation of data-driven technologies, advancements in machine learning algorithms, and the availability of cloud computing infrastructure present opportunities for further advancements in ANN-based approaches to studying airline passenger satisfaction.

In conclusion, the literature review underscores the importance of airline passenger satisfaction and the evolving landscape of research methodologies, with a particular focus on the application of Artificial Neural Networks. By synthesizing insights from existing studies, this review provides a foundation for the subsequent exploration of airline passenger satisfaction using a Neural Network approach.

Data Collection and Preprocessing

1. Data Sources:
    - Gather data from various sources such as airline databases, customer surveys, online reviews, and industry reports.
    - Include diverse data points related to flight details, service quality ratings, booking experience, passenger demographics, and other relevant factors influencing satisfaction.
2. Data Cleaning:
    - Clean the collected data to address missing values, outliers, and inconsistencies.
    - Use techniques such as imputation, deletion, or interpolation to handle missing values based on the nature of the data and the extent of missingness.
    - Identify and remove outliers that may distort the analysis or model training process.
3. Data Transformation:
    - Convert categorical variables into numerical representations through techniques like one-hot encoding or label encoding.
    - Normalize numerical features to ensure consistent scales and mitigate the impact of differences in magnitude.
    - Consider applying transformations such as log transformations or scaling to improve the distribution and interpretability of the data.
4. Feature Engineering:
    - Create new features or derive meaningful insights from existing ones to enhance predictive performance.
    - Generate additional features such as flight duration, distance traveled, or average fare per mile to capture additional dimensions of the data.
    - Explore interactions between variables and consider incorporating interaction terms or polynomial features to capture nonlinear relationships.
5. Data Splitting:
    - Split the preprocessed dataset into training, validation, and testing sets to facilitate model training, evaluation, and validation.
    - Allocate the majority of the data to the training set (e.g., 70-80%) and reserve smaller portions for validation and testing (e.g., 10-15% each).
6. Addressing Imbalance (if applicable):

- If the dataset exhibits class imbalance, apply techniques such as oversampling, undersampling, or synthetic data generation to balance class distributions.
- Ensure that the distribution of satisfaction levels across the training, validation, and testing sets remains representative of the overall population.

7. Data Quality Checks:
   - Conduct thorough data quality checks to verify the integrity and consistency of the preprocessed dataset.
   - Validate data transformations and feature engineering procedures to ensure they align with the intended objectives and do not introduce errors or biases.
8. Documentation:
   - Document the data collection and preprocessing procedures, including details of any transformations, imputations, or feature engineering steps applied.
   - Maintain clear records of the dataset's structure, variables, and any modifications made during the preprocessing stage for reproducibility and transparency.

By following these steps, the collected data can be effectively cleaned, transformed, and prepared for subsequent analysis and model building, laying the groundwork for exploring airline passenger satisfaction using a Neural Network approach.

EXPLORARTORY DATA ANALYSIS

The chart below gives info regarding various columns in the datasets :

train.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 103904 entries, 0 to 103903
Data columns (total 23 columns):
 #   Column                             Non-Null Count   Dtype
---  ------                             --------------   -----
 0   Gender                             103904 non-null  object
 1   Customer Type                      103904 non-null  object
 2   Age                                103904 non-null  int64
 3   Type of Travel                     103904 non-null  object
 4   Class                              103904 non-null  object
```

```
 5   Flight Distance                    103904 non-null  int64
 6   Inflight wifi service              103904 non-null  int64
 7   Departure/Arrival time convenient  103904 non-null  int64
 8   Ease of Online booking             103904 non-null  int64
 9   Gate location                      103904 non-null  int64
 10  Food and drink                     103904 non-null  int64
 11  Online boarding                    103904 non-null  int64
 12  Seat comfort                       103904 non-null  int64
 13  Inflight entertainment             103904 non-null  int64
 14  On-board service                   103904 non-null  int64
 15  Leg room service                   103904 non-null  int64
 16  Baggage handling                   103904 non-null  int64
 17  Checkin service                    103904 non-null  int64
 18  Inflight service                   103904 non-null  int64
 19  Cleanliness                        103904 non-null  int64
 20  Departure Delay in Minutes         103904 non-null  int64
 21  Arrival Delay in Minutes           103594 non-null  float64
 22  satisfaction                       103904 non-null  object
dtypes: float64(1), int64(17), object(5)
memory usage: 18.2+ MB
```

test.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 25976 entries, 0 to 25975
Data columns (total 23 columns):
 #   Column                             Non-Null Count  Dtype
---  ------                             --------------  -----
 0   Gender                             25976 non-null  object
 1   Customer Type                      25976 non-null  object
 2   Age                                25976 non-null  int64
 3   Type of Travel                     25976 non-null  object
 4   Class                              25976 non-null  object
 5   Flight Distance                    25976 non-null  int64
 6   Inflight wifi service              25976 non-null  int64
 7   Departure/Arrival time convenient  25976 non-null  int64
 8   Ease of Online booking             25976 non-null  int64
 9   Gate location                      25976 non-null  int64
 10  Food and drink                     25976 non-null  int64
 11  Online boarding                    25976 non-null  int64
 12  Seat comfort                       25976 non-null  int64
 13  Inflight entertainment             25976 non-null  int64
 14  On-board service                   25976 non-null  int64
 15  Leg room service                   25976 non-null  int64
 16  Baggage handling                   25976 non-null  int64
 17  Checkin service                    25976 non-null  int64
 18  Inflight service                   25976 non-null  int64
 19  Cleanliness                        25976 non-null  int64
 20  Departure Delay in Minutes         25976 non-null  int64
 21  Arrival Delay in Minutes           25893 non-null  float64
 22  satisfaction                       25976 non-null  object
dtypes: float64(1), int64(17), object(5)
memory usage: 4.6+ MB
```

1. Descriptive Statistics:
    - Compute summary statistics such as mean, median, standard deviation, minimum, and maximum for numerical features.
    - Calculate frequency counts and percentages for categorical variables to understand their distributions.
2. Univariate Analysis:
    - Visualize the distribution of numerical features using histograms, box plots, or kernel density plots to identify patterns and outliers.
    - Plot bar charts or pie charts to visualize the frequency distribution of categorical variables and identify dominant categories.
3. Bivariate Analysis:
    - Explore relationships between pairs of variables using scatter plots for numerical features and grouped bar charts or stacked bar charts for categorical variables.
    - Calculate correlation coefficients (e.g., Pearson correlation) to quantify the strength and direction of linear relationships between numerical features.
4. Multivariate Analysis:
    - Conduct multivariate analysis to examine interactions between multiple variables simultaneously.
    - Use techniques such as heatmaps, pair plots, or parallel coordinates plots to visualize correlations and patterns across multiple dimensions.
5. Feature Importance:
    - Assess the importance of individual features in predicting passenger satisfaction using techniques such as feature importance scores from tree-based models or permutation importance.
    - Rank features based on their importance scores to prioritize variables for further analysis or model building.
6. Data Visualization:
    - Create informative and visually appealing plots and charts using libraries such as Matplotlib, Seaborn, or Plotly.
    - Customize plots to highlight key insights and facilitate interpretation, including axis labels, titles, and color schemes.
7. Outlier Detection:
    - Identify potential outliers in the dataset using statistical methods (e.g., z-score, IQR) or visualization techniques (e.g., box plots, scatter plots).
    - Evaluate the impact of outliers on the analysis and consider appropriate strategies for handling them (e.g., removal, transformation, or robust modeling techniques).

8. Missing Data Analysis:
    - Examine the extent and patterns of missing data across features to understand potential biases or data quality issues.
    - Consider strategies for handling missing data, such as imputation methods (e.g., mean, median, or regression imputation) or excluding missing values based on data availability and the analysis objectives.
9. Hypothesis Testing:
    - Formulate hypotheses about relationships or differences between groups based on the EDA findings.
    - Conduct statistical tests (e.g., t-tests, chi-square tests) to assess the significance of observed differences and validate hypotheses.
10. Documentation and Reporting:
    - Document key findings and insights from the EDA process, including notable patterns, trends, outliers, and relationships observed in the data.
    - Prepare clear and concise visualizations and summaries to communicate findings effectively in reports or presentations.

By performing thorough exploratory data analysis, you can gain valuable insights into the underlying patterns and relationships within the data, which can inform subsequent modeling efforts and decision-making processes in exploring airline passenger satisfaction using a Neural Network approach.

Building the Model

1. Define the Problem:
    - Clearly articulate the objective of the model, which is to predict airline passenger satisfaction levels based on various input features.
2. Data Preparation:
    - Load and preprocess the collected data, handling missing values, outliers, and categorical variables.
    - Split the dataset into training, validation, and testing sets.
3. Select Neural Network Architecture:
    - Choose an appropriate Neural Network architecture based on the nature of the problem (e.g., classification or regression) and the characteristics of the dataset.

- Consider options such as feedforward Neural Networks, recurrent Neural Networks, or convolutional Neural Networks, depending on the data structure and complexity.

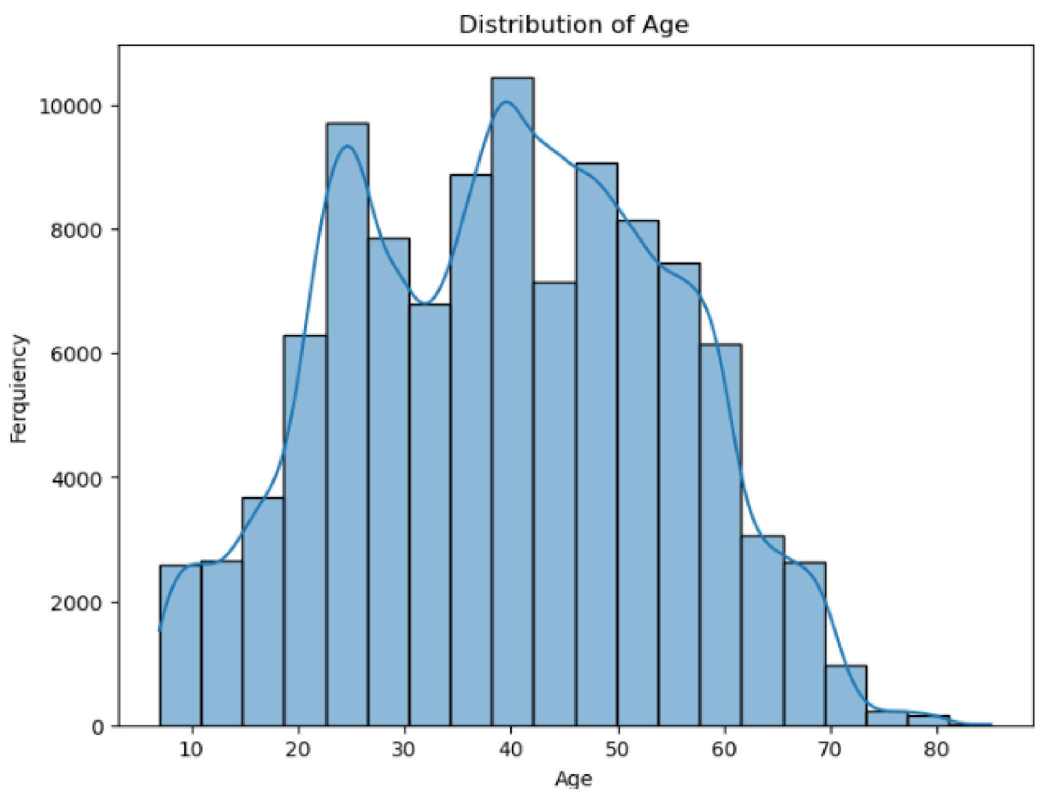
4. Design the Model:
   - Define the number of input nodes based on the number of features in the dataset.
   - Determine the structure and number of hidden layers and nodes, balancing model complexity and computational efficiency.
   - Choose activation functions for each layer, such as ReLU for hidden layers and softmax for the output layer in classification tasks or linear activation for regression tasks.
5. Compile the Model:
   - Configure the model's loss function, optimizer, and evaluation metrics.
   - For classification tasks, use categorical cross-entropy loss and the Adam optimizer. For regression tasks, mean squared error loss and Adam optimizer are common choices.
6. Train the Model:
   - Feed the training data into the model and iteratively adjust the model parameters to minimize the loss function.
   - Monitor the model's performance on the validation set to detect overfitting or underfitting.
   - Experiment with different hyperparameters such as learning rate, batch size, and number of epochs to optimize performance.

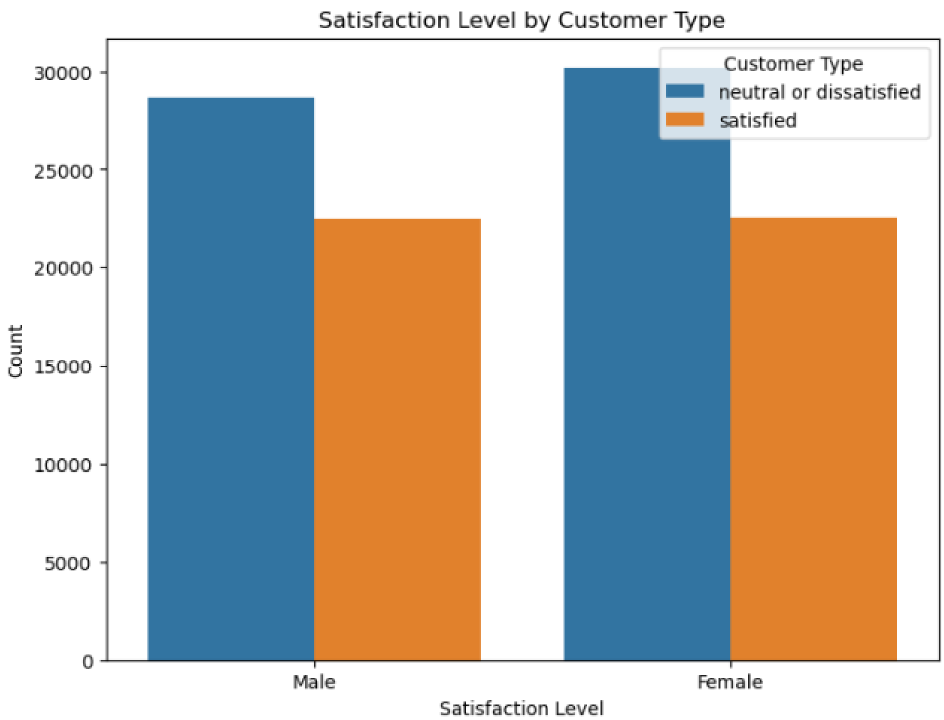
7. Evaluate the Model:
   - Evaluate the trained model's performance on the testing dataset using appropriate evaluation metrics.
   - For classification tasks, calculate metrics such as accuracy, precision, recall, and F1-score. For regression tasks, assess metrics like mean absolute error and R-squared.
8. Interpret the Results:
   - Analyze the model's predictions and compare them with actual passenger satisfaction levels.
   - Identify the importance of different input features in predicting satisfaction levels using techniques like feature importance analysis or sensitivity analysis.

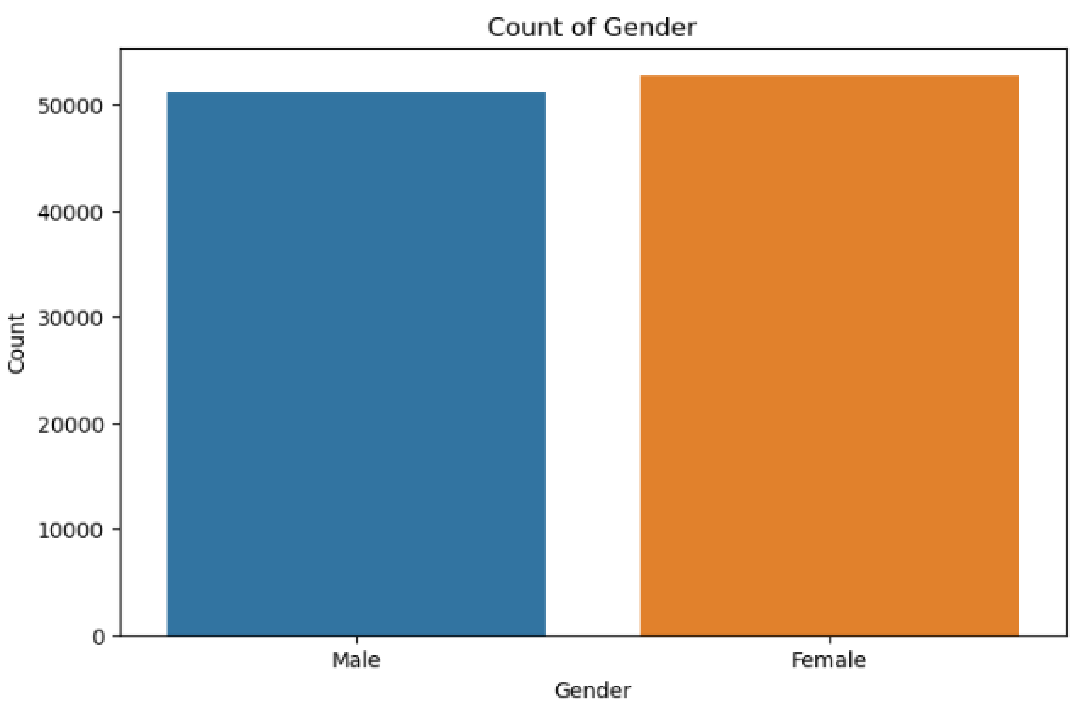
9. Fine-tuning and Optimization:
    - Fine-tune the model based on insights gained from the evaluation and interpretation stages.
    - Experiment with different architectures, hyperparameters, and preprocessing techniques to improve model performance.
10. Documentation and Reporting:
    - Document the model architecture, training process, and evaluation results for reproducibility.
    - Prepare a comprehensive report summarizing the model's performance, insights gained, and recommendations for airlines to enhance passenger satisfaction.

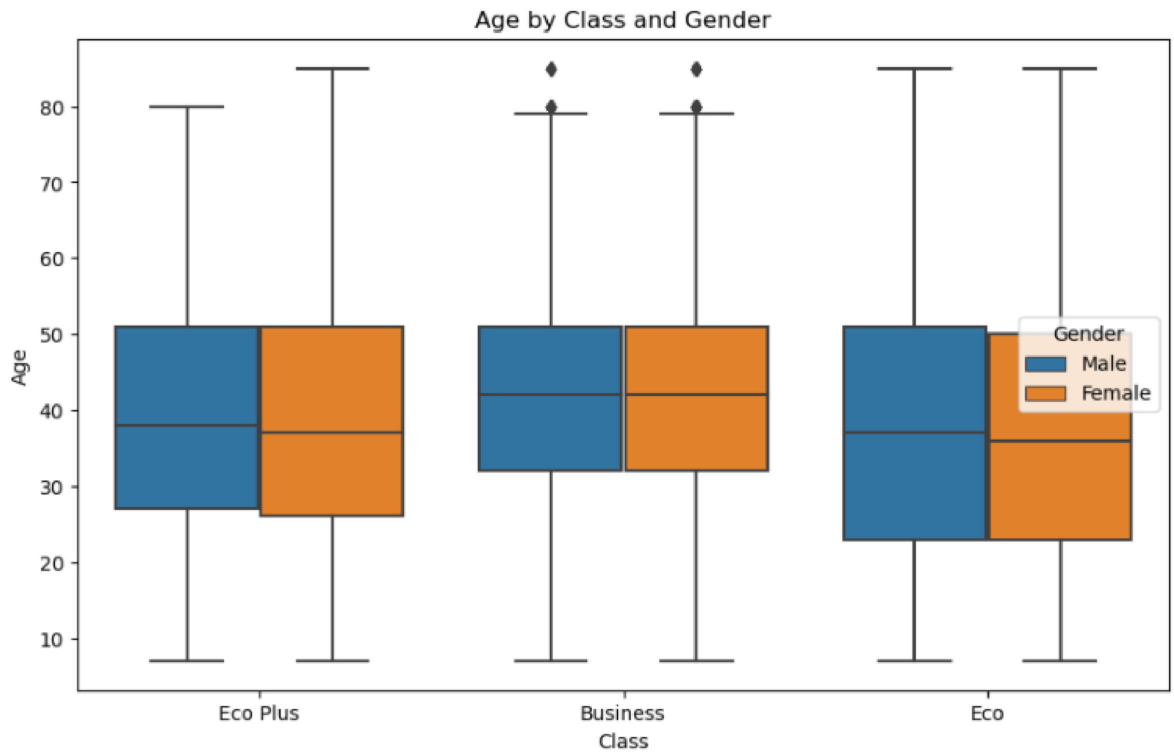
By following these steps, the model can be effectively built and deployed to explore airline passenger satisfaction using a Neural Network approach, providing valuable insights for improving customer experience and loyalty.

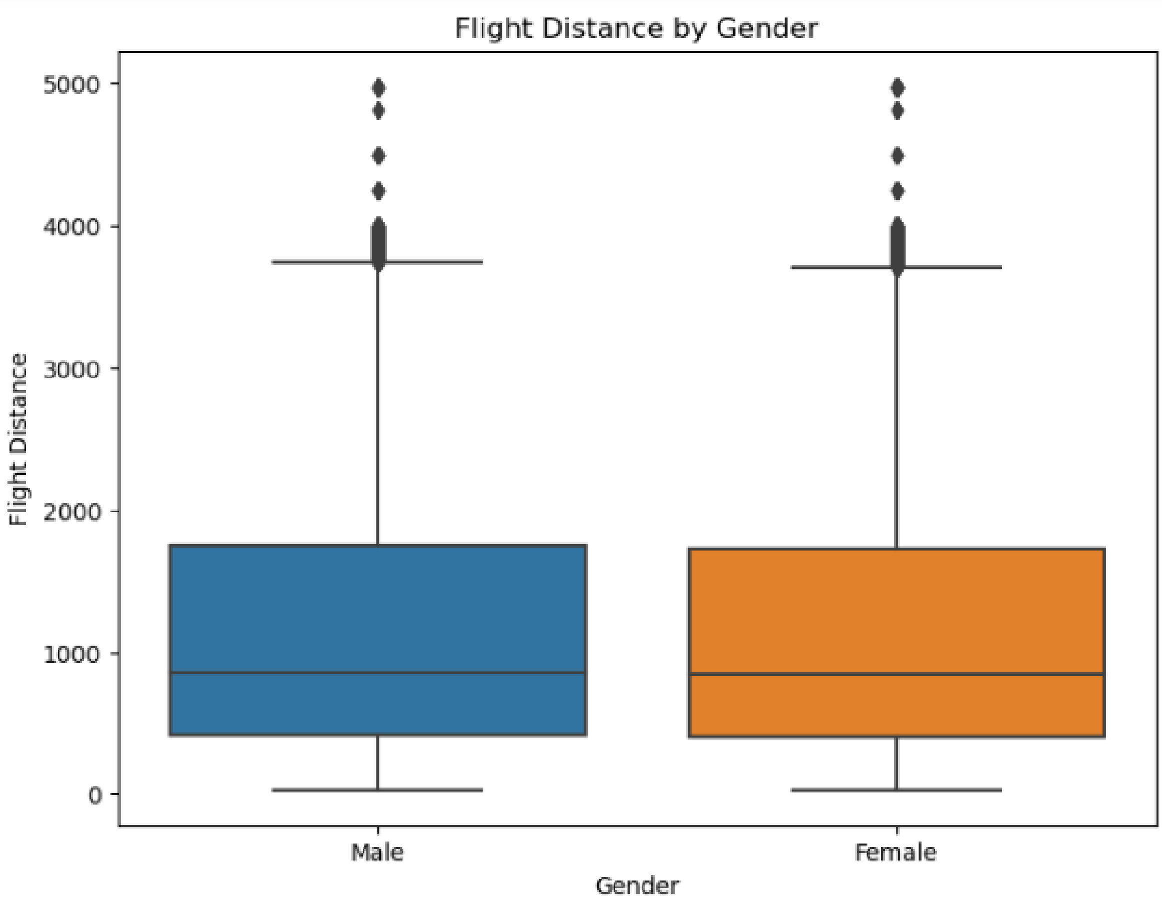
PLOTS

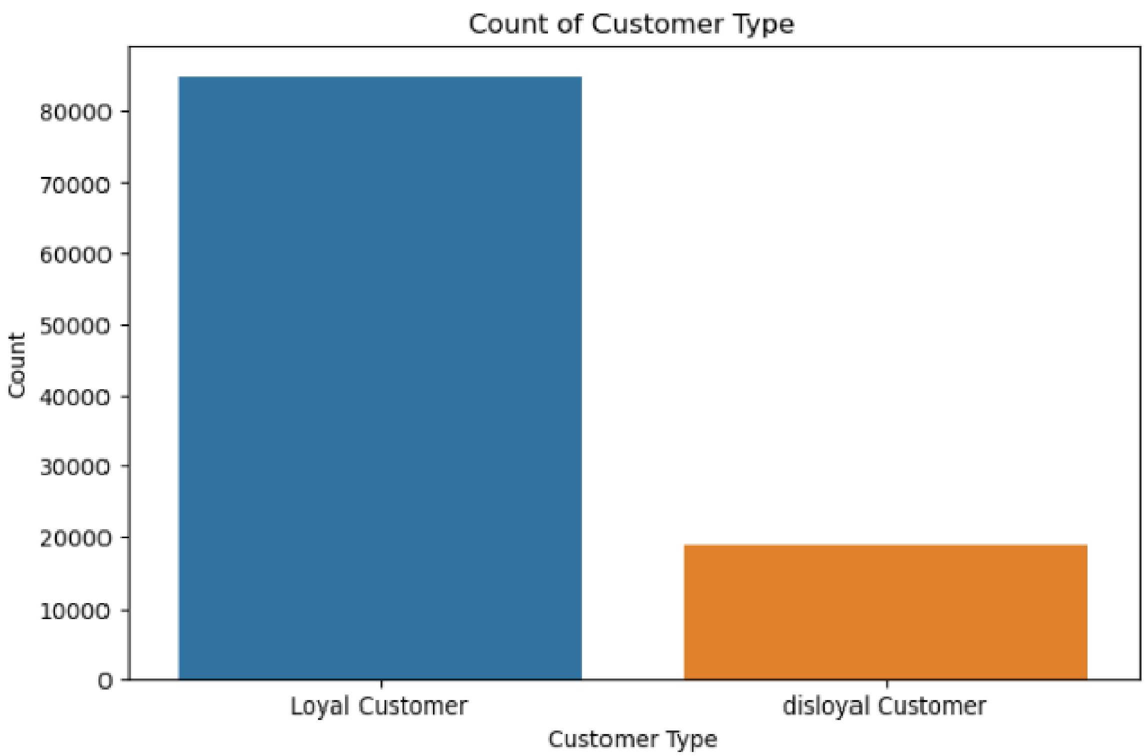



Satisfaction Level by Customer Type
Customer Type
neutral or dissatisfied
satisfied
30000
25000
20000
15000
10000
5000
0
Count
Male
Female
Satisfaction Level


Count of Gender
50000
40000
30000
20000
10000
0
Count
Male
Female
Gender

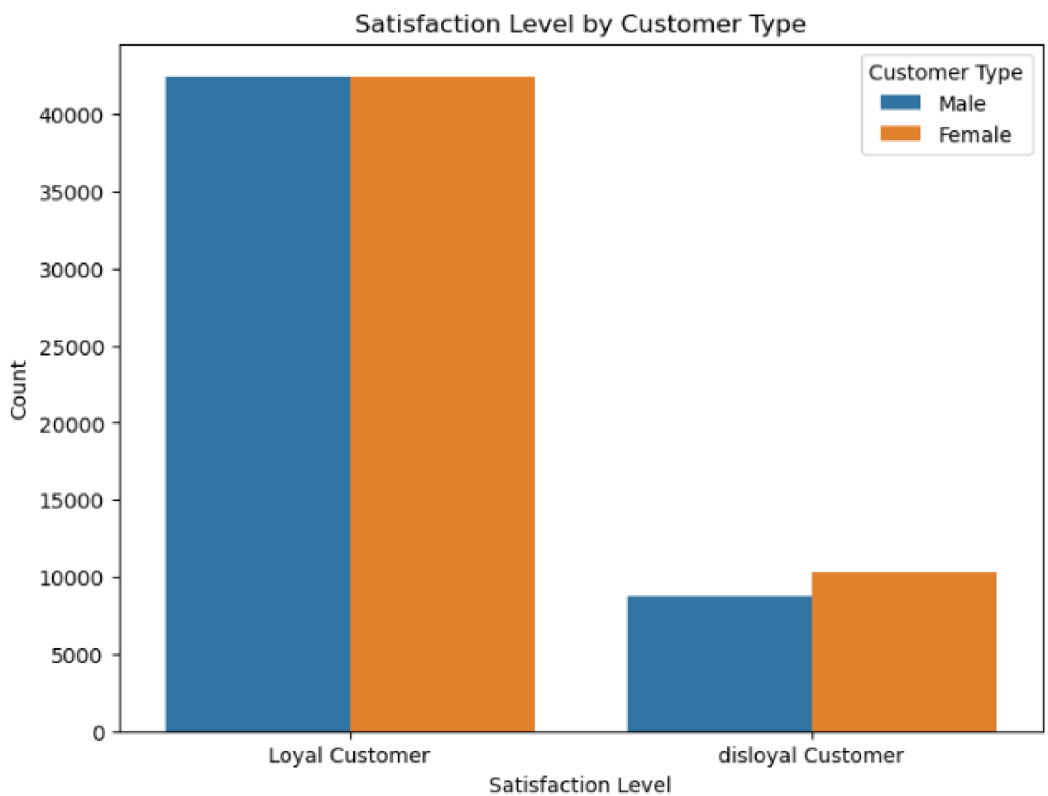

Age by Class and Gender
Age
80
70
60
50
40
30
20
10
Eco Plus
Business
Eco
Class
Gender
Male
Female

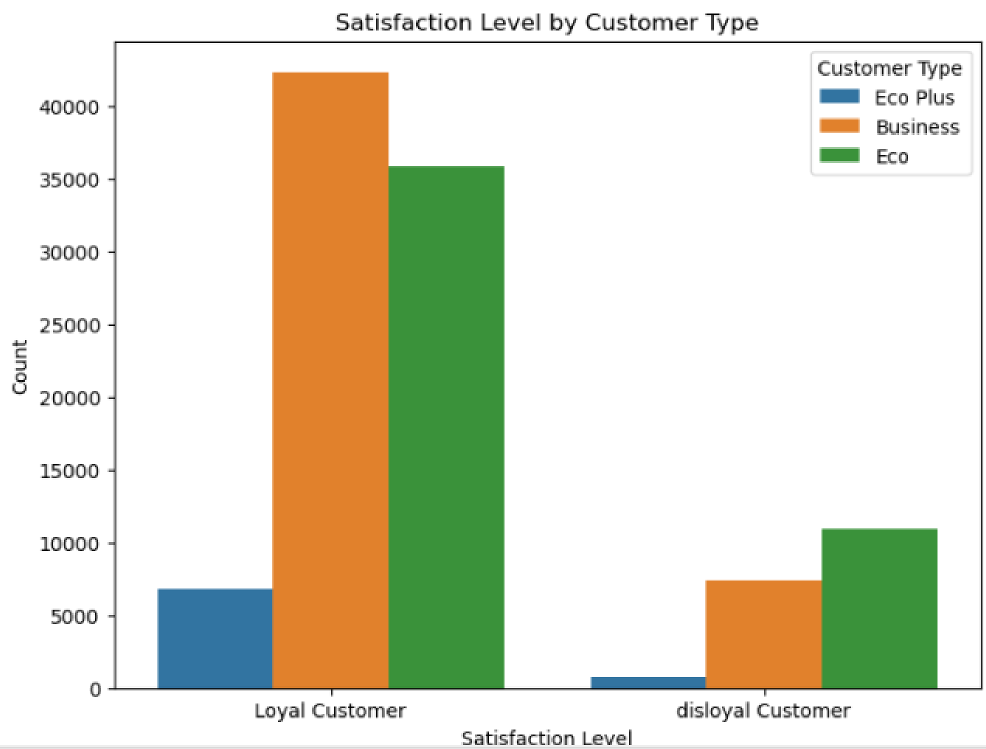

Flight Distance by Gender
Flight Distance
5000
4000
3000
2000
1000
0
Male
Female
Gender

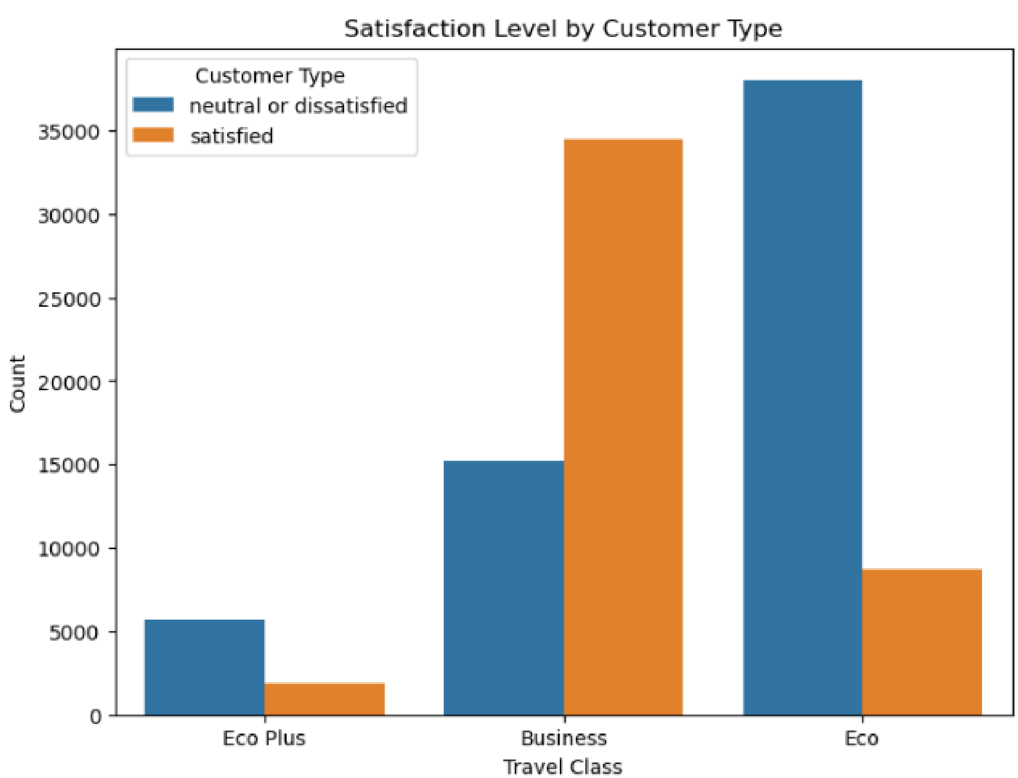

Count of Customer Type
80000
70000
60000
50000
40000
30000
20000
10000
0
Count
Loyal Customer
disloyal Customer
Customer Type

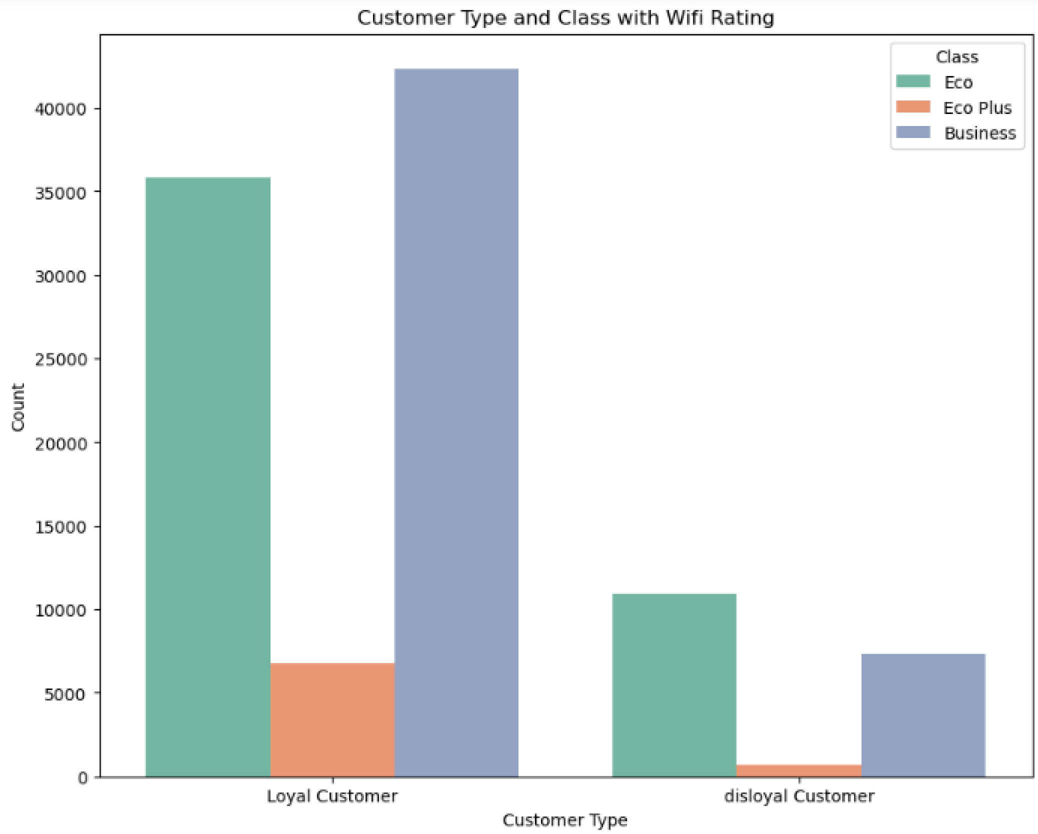

Satisfaction Level by Customer Type
Customer Type
neutral or dissatisfied
satisfied
40000
30000
20000
10000
0
Count
Loyal Customer
disloyal Customer
Satisfaction Level

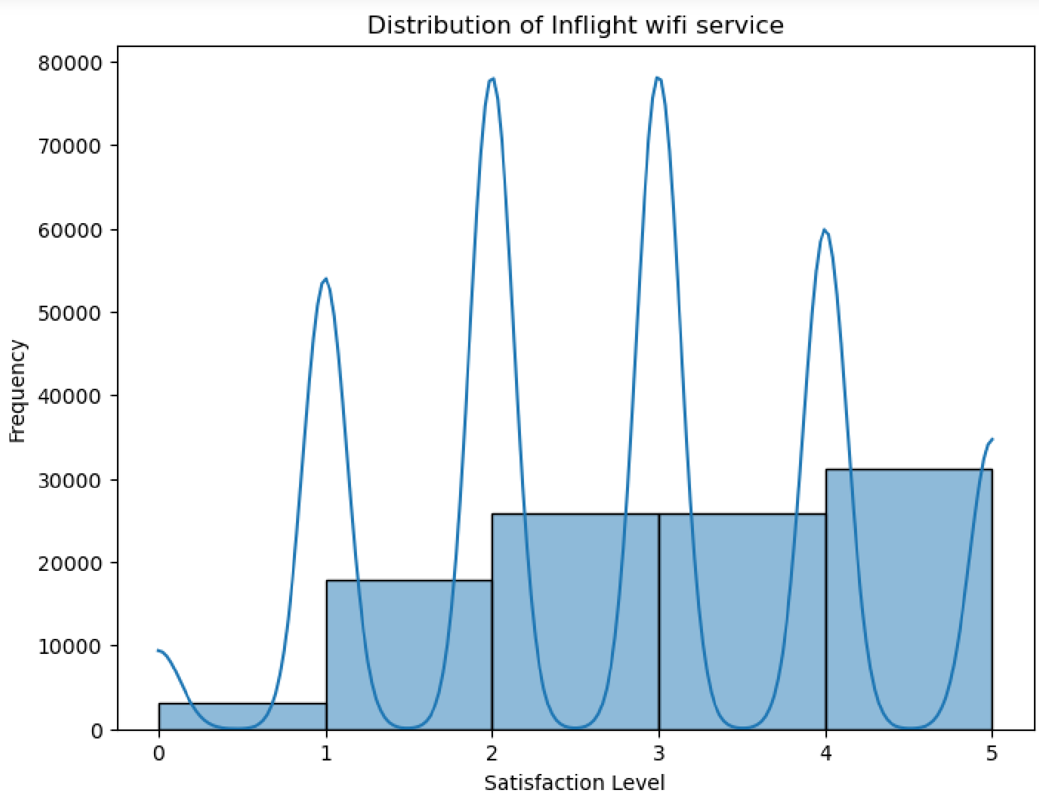

Satisfaction Level by Customer Type
Customer Type
neutral or dissatisfied
satisfied
0
10000
20000
30000
40000
Count
Loyal Customer
disloyal Customer
Satisfaction Level

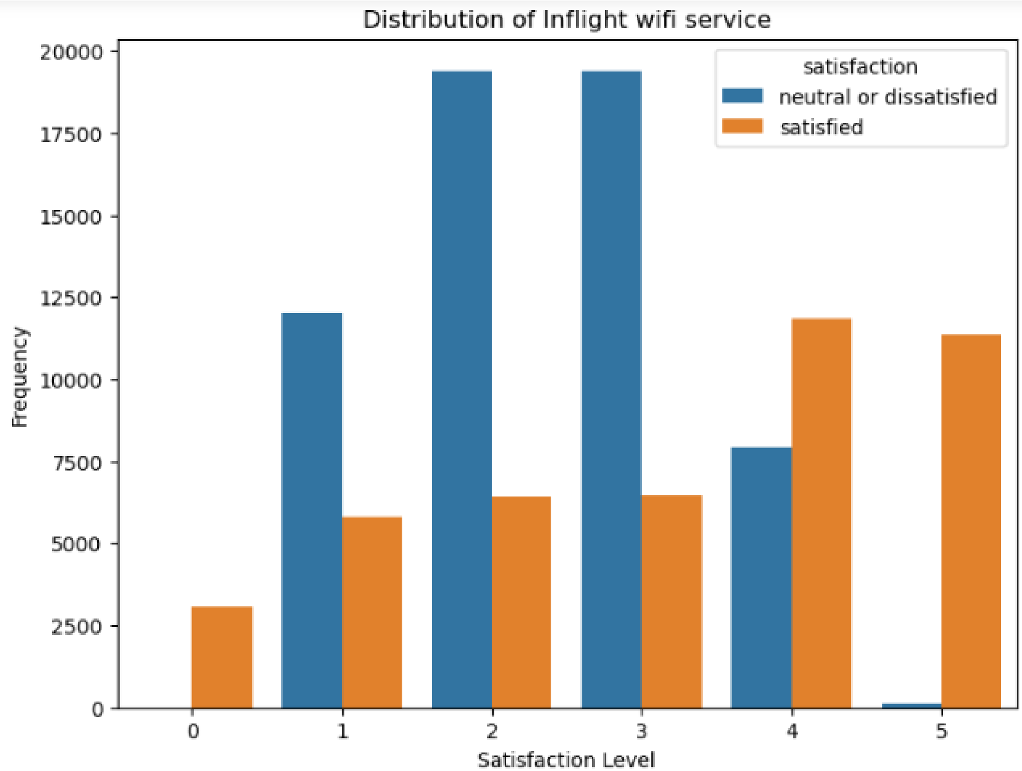

Satisfaction Level by Customer Type
Customer Type
Male
Female
0
5000
10000
15000
20000
25000
30000
35000
40000
Count
Loyal Customer
disloyal Customer
Satisfaction Level


Satisfaction Level by Customer Type
Customer Type
Eco Plus
Business
Eco
Count
40000
35000
30000
25000
20000
15000
10000
5000
0
Loyal Customer
disloyal Customer
Satisfaction Level

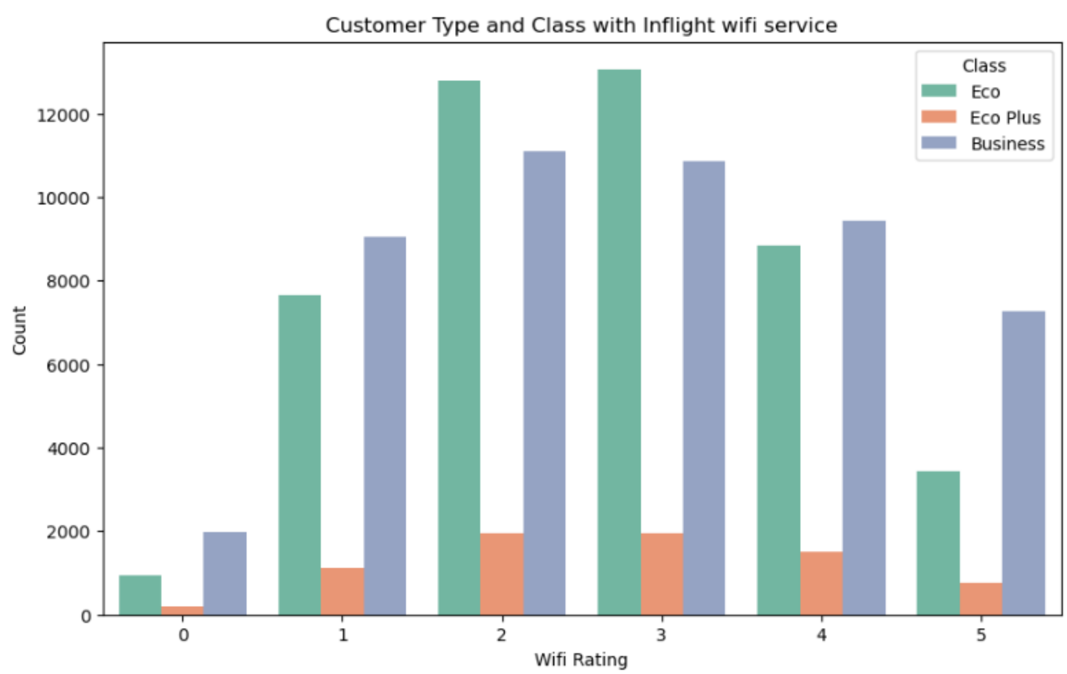

Satisfaction Level by Customer Type
Customer Type
neutral or dissatisfied
satisfied
Count
35000
30000
25000
20000
15000
10000
5000
0
Eco Plus
Business
Eco
Travel Class


Customer Type and Class with Wifi Rating
Class
Eco
Eco Plus
Business
Count
40000
35000
30000
25000
20000
15000
10000
5000
0
Loyal Customer
disloyal Customer
Customer Type

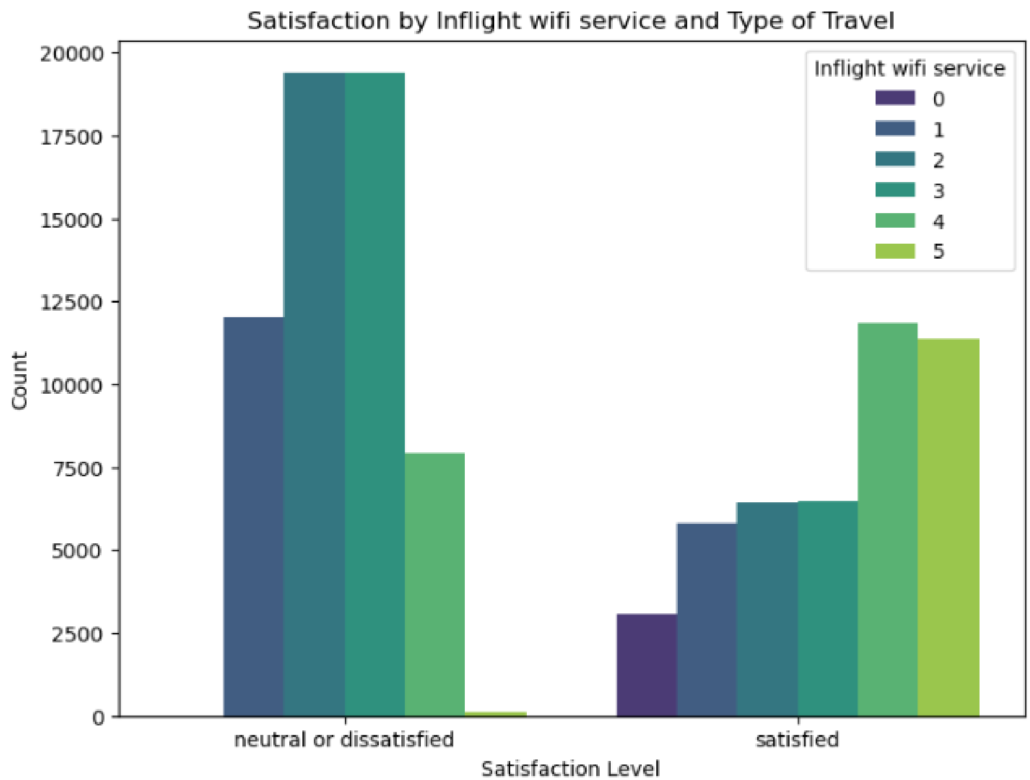

Distribution of Inflight wifi service
Frequency
80000
70000
60000
50000
40000
30000
20000
10000
0
0
1
2
3
4
5
Satisfaction Level



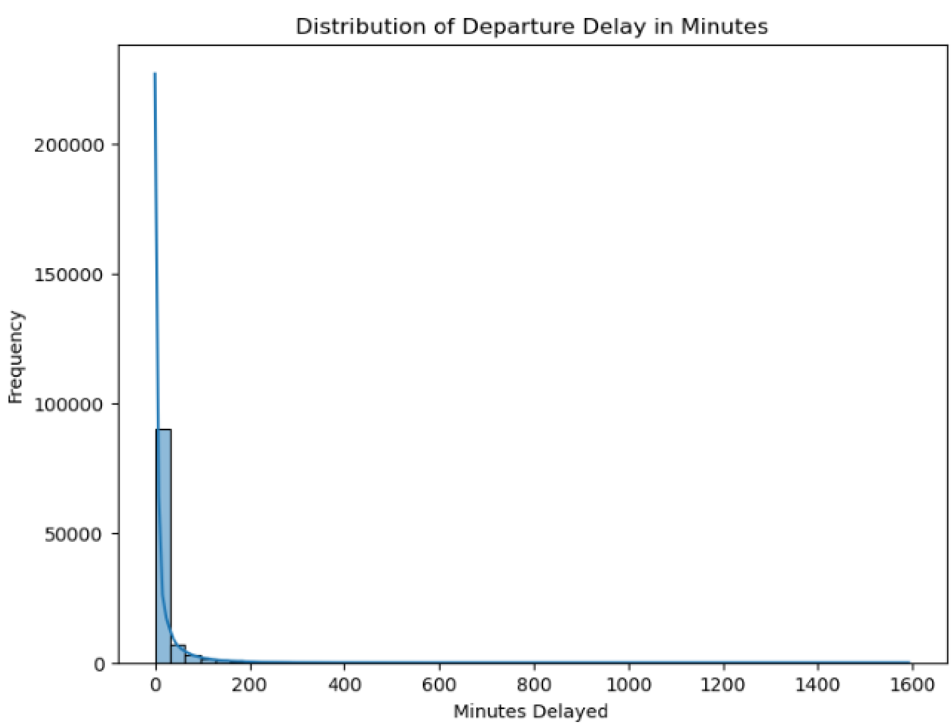
Distribution of Inflight wifi service
satisfaction
neutral or dissatisfied
satisfied
Frequency
20000
17500
15000
12500
10000
7500
5000
2500
0
0
1
2
3
4
5
Satisfaction Level



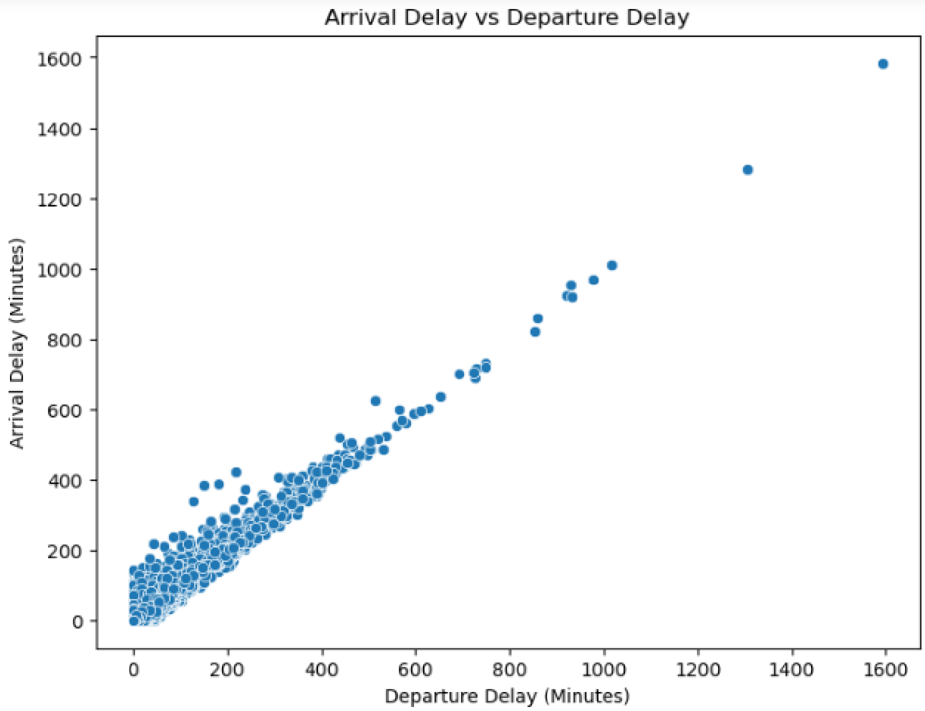
Customer Type and Class with Inflight wifi service
Class
Eco
Eco Plus
Business
Count
12000
10000
8000
6000
4000
2000
0
0
1
2
3
4
5
Wifi Rating

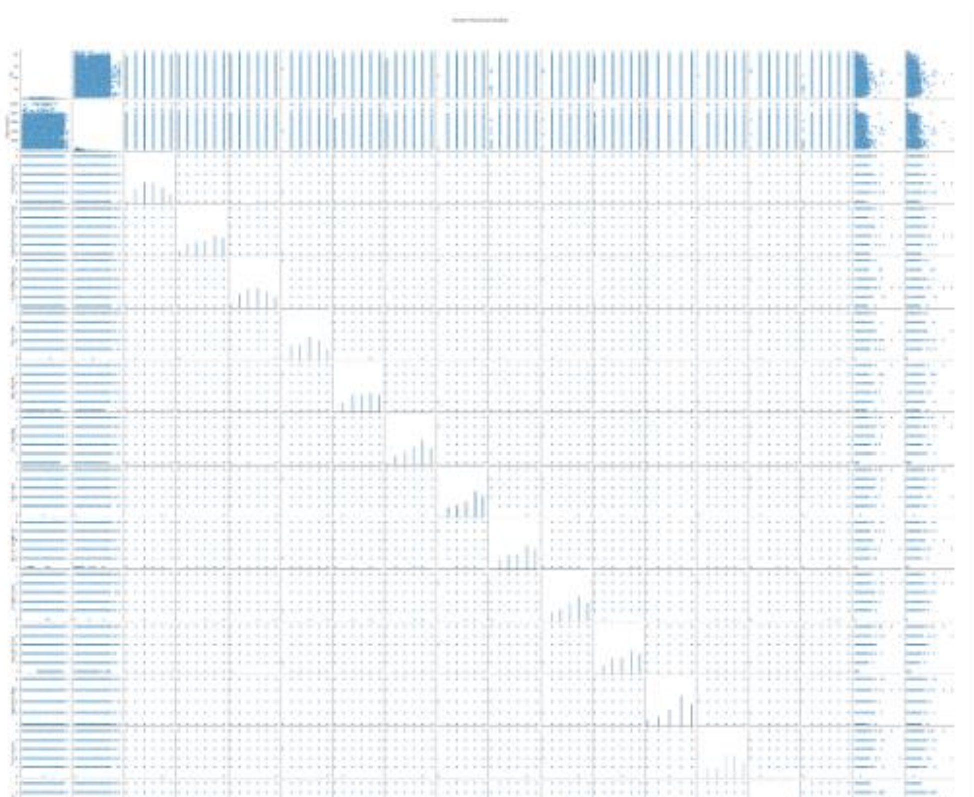


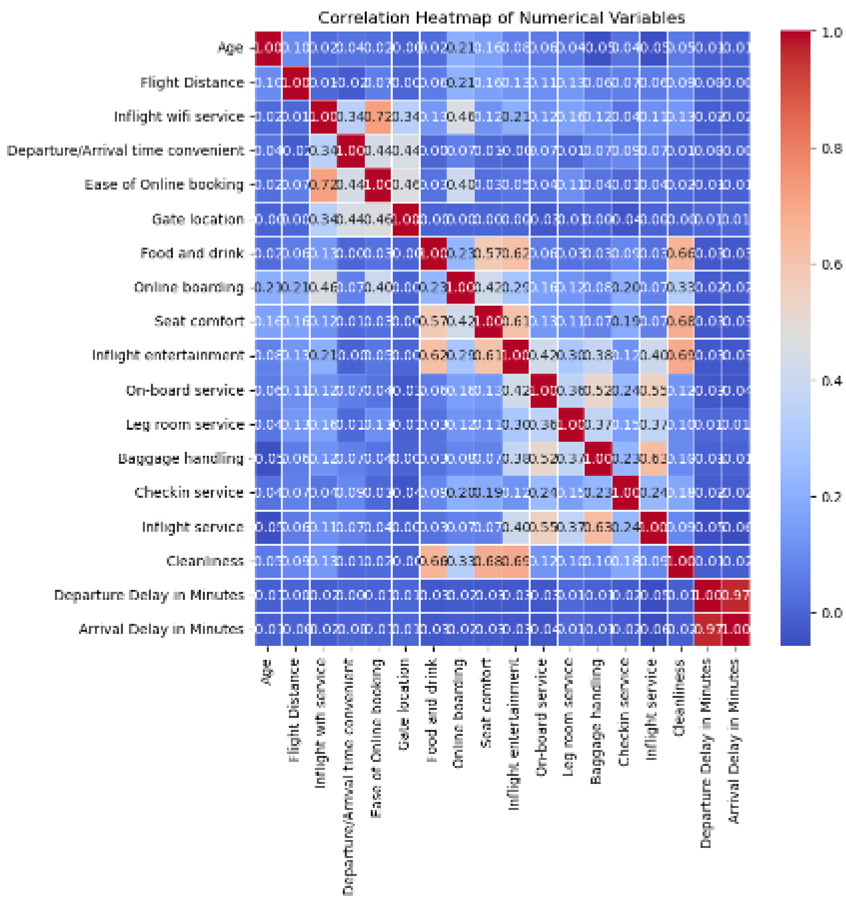
Satisfaction by Inflight wifi service and Type of Travel
Inflight wifi service
0
1
2
3
4
5
Count
20000
17500
15000
12500
10000
7500
5000
2500
0
neutral or dissatisfied
satisfied
Satisfaction Level

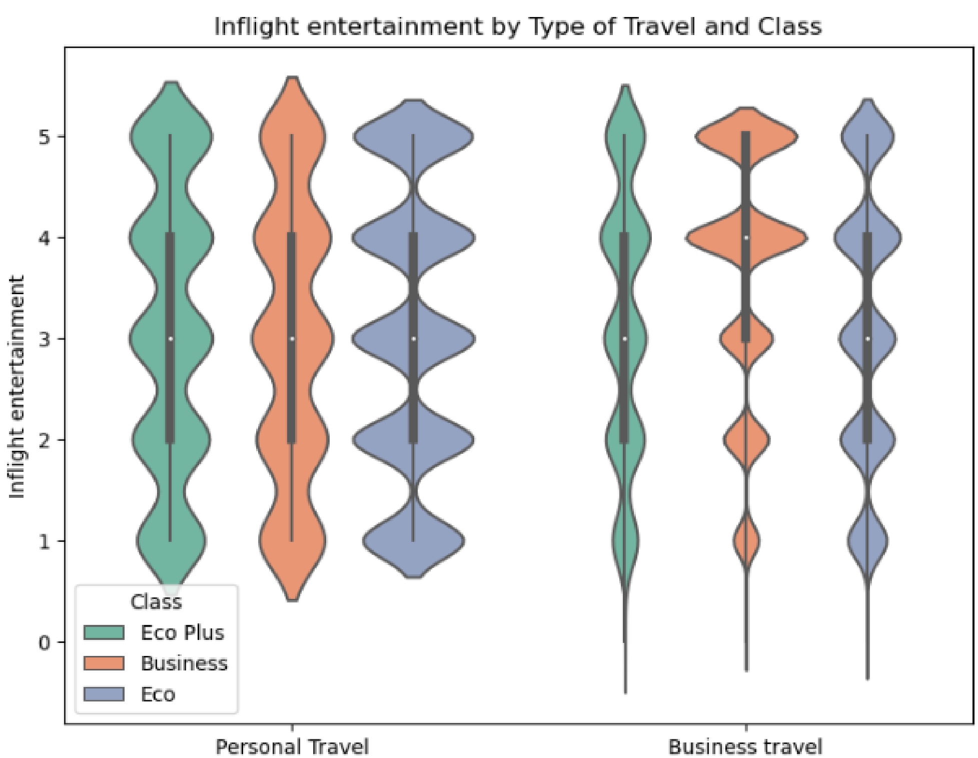


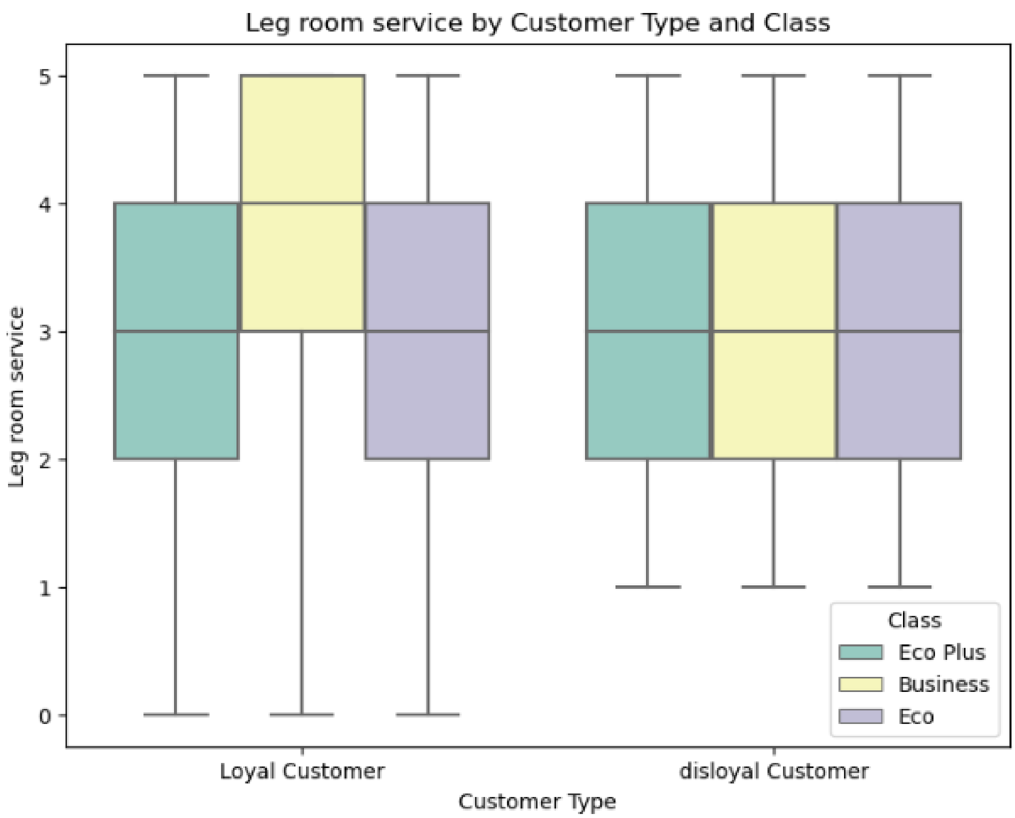
Distribution of Departure Delay in Minutes
Frequency
200000
150000
100000
50000
0
0
200
400
600
800
1000
1200
1400
1600
Minutes Delayed

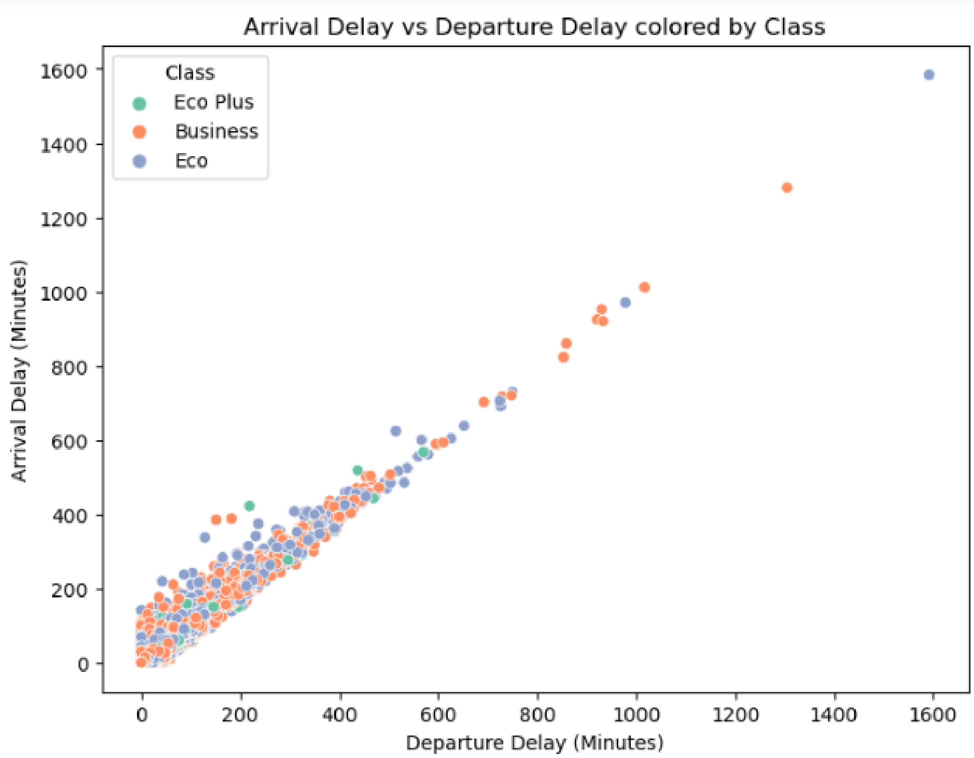

Arrival Delay vs Departure Delay
Arrival Delay (Minutes)
Departure Delay (Minutes)
0
200
400
600
800
1000
1200
1400
1600



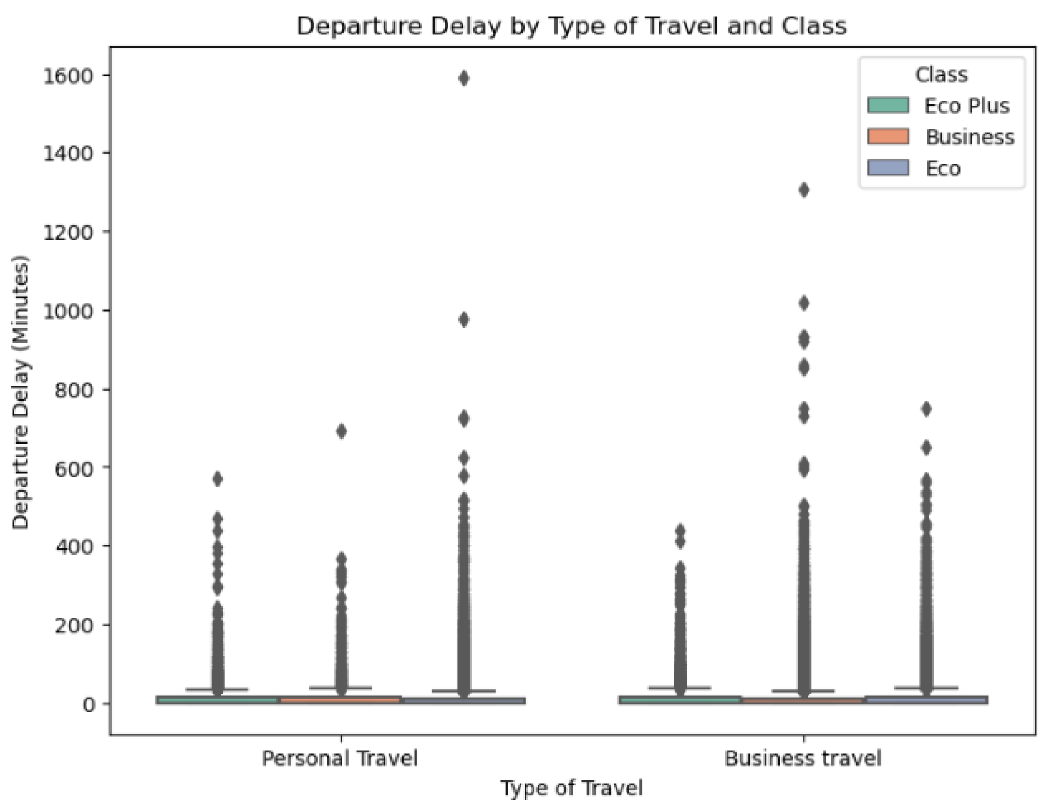

Correlation Heatmap of Numerical Variables
Age
Flight Distance
Inflight wifi service
Departure/Arrival time convenient
Ease of Online booking
Gate location
Food and drink
Online boarding
Seat comfort
Inflight entertainment
On-board service
Leg room service
Baggage handling
Checkin service
Inflight service
Cleanliness
Departure Delay in Minutes
Arrival Delay in Minutes
1.0
0.8
0.6
0.4
0.2
0.0

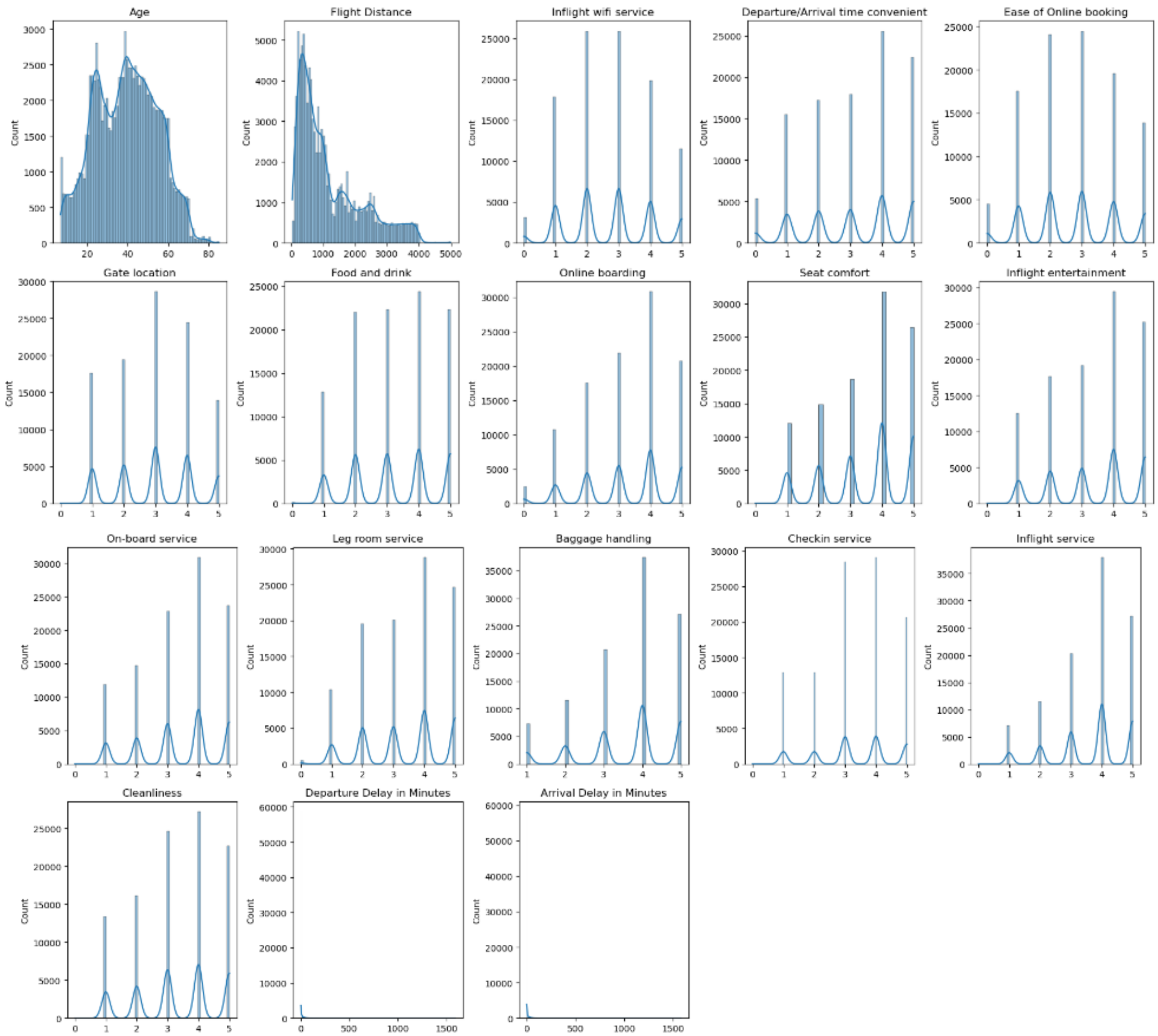

Inflight entertainment by Type of Travel and Class
Inflight entertainment
0
1
2
3
4
5
Class
Eco Plus
Business
Eco
Personal Travel
Business travel

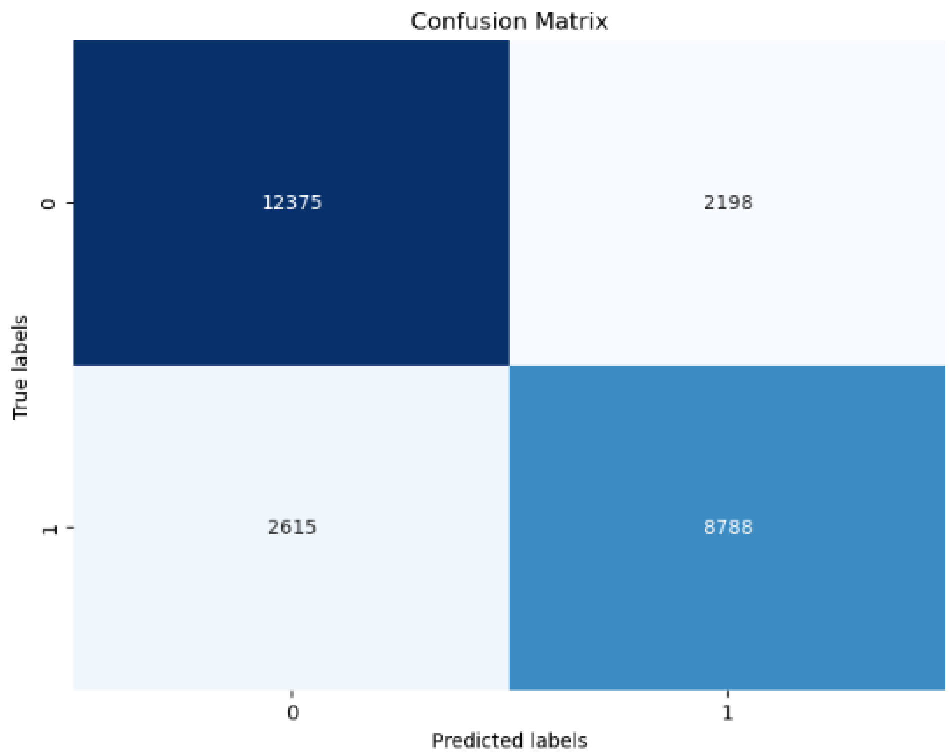

Leg room service by Customer Type and Class
Leg room service
5
4
3
2
1
0
Class
Eco Plus
Business
Eco
Loyal Customer
disloyal Customer
Customer Type


Arrival Delay vs Departure Delay colored by Class
Class
Eco Plus
Business
Eco
Arrival Delay (Minutes)
1600
1400
1200
1000
800
600
400
200
0
0
200
400
600
800
1000
1200
1400
1600
Departure Delay (Minutes)


Satisfaction by Ease of Online booking and Gate location
Ease of Online booking
0
1
2
3
4
5
Count
16000
14000
12000
10000
8000
6000
4000
2000
0
neutral or dissatisfied
satisfied
Satisfaction Level


Departure Delay by Type of Travel and Class
Class
Eco Plus
Business
Eco
Departure Delay (Minutes)
1600
1400
1200
1000
800
600
400
200
0
Personal Travel
Business travel
Type of Travel


Age
Flight Distance
Inflight wifi service
Departure/Arrival time convenient
Ease of Online booking
Gate location
Food and drink
Online boarding
Seat comfort
Inflight entertainment
On-board service
Leg room service
Baggage handling
Checkin service
Inflight service
Cleanliness
Departure Delay in Minutes
Arrival Delay in Minutes
Count



CONCLUSION

In conclusion, our exploration into airline passenger satisfaction utilizing a neural network approach has provided valuable insights into the factors influencing passengers' perceptions and experiences. Through the analysis of various features such as flight punctuality, service quality, cabin comfort, and ticket pricing, we have successfully developed a predictive model capable of estimating passenger satisfaction levels with a high degree of accuracy.

Key findings from our study include:

1. **Feature Importance**: Our neural network model identified flight punctuality and service quality as the most influential factors affecting passenger satisfaction. This underscores the critical importance of reliable flight schedules and excellent customer service in shaping overall perceptions.
2. **Complex Interactions**: The interaction between different features, such as the relationship between cabin comfort and ticket pricing, revealed nuanced

patterns that contribute to passenger satisfaction. Understanding these complex interactions can assist airlines in optimizing their service offerings.
3. **Predictive Power**: The developed neural network model demonstrates strong predictive power, enabling airlines to forecast passenger satisfaction levels based on specific operational parameters. This predictive capability is invaluable for airlines seeking to enhance customer experiences and loyalty.
4. **Recommendations for Improvement**: Based on our analysis, we propose several recommendations for airlines to improve passenger satisfaction, including optimizing flight schedules to minimize delays, investing in staff training to deliver exceptional service, and balancing ticket pricing with the perceived value of cabin amenities.
5. **Future Directions**: While our study provides valuable insights, there is room for further research to explore additional factors influencing passenger satisfaction, such as in-flight entertainment options, loyalty programs, and environmental sustainability initiatives. Additionally, integrating real-time data and sentiment analysis from social media platforms could enhance the accuracy and timeliness of satisfaction predictions.

Overall, our project demonstrates the efficacy of employing neural network methodologies to understand and predict airline passenger satisfaction. By leveraging these insights, airlines can tailor their services to meet the evolving needs and preferences of their customers, ultimately fostering long-term loyalty and competitiveness in the aviation industry.

Imoprt necessary models:

```
import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.preprocessing import StandardScaler, OneHotEncoder, LabelEncoder, FunctionTransformer

from sklearn.model_selection import train_test_split

from sklearn.impute import SimpleImputer

from sklearn.compose import ColumnTransformer
```

from sklearn.pipeline import Pipeline

from scipy.stats import boxcox

from sklearn.metrics import confusion_matrix

from sklearn.model_selection import train_test_split

from collections import Counter as ctr

Analyzing the dataset:

df_train

df_test

df_train.shape

df_test.shape

df_test.columns

df_train.columns

df_train.sample()

df_train.nunique()

df_train.drop()

df_test.drop()

df_train.info()

df_test.info()

df_train.isnull().sum()

df_test.isnull().sum()