

AIRLINE PASSENGER SATISFACTION PROJECT

Using Machine Learning to Predict and
Improve Passenger Experience

Presented by: DANIEL MURUTHI



AGENDA

- Project Overview
- Objectives
- Dataset Description
- Data Preprocessing
- Exploratory Data Analysis
- Machine learning modelling
- Business Insights
- Future Work
- Q&A



PROJECT OVERVIEW



The goal of this project is to analyze key factors affecting airline passenger satisfaction using machine learning. The insights aim to guide strategic improvements in customer experience.



OBJECTIVES

- 01** Explore and analyze customer satisfaction data
- 02** Identify features influencing satisfaction
- 03** Build predictive models
- 04** Derive actionable business insights

DATASET DESCRIPTION

Source : Kaggle - Airline passenger satisfaction

Features :

- Demographics (Gender, Age, Loyalty)
- Travel Details (Class, Distance, Type of Travel)
- Service Ratings (e.g., Wi-Fi, Seat Comfort, Cleanliness)
- Timeliness (Arrival/Departure Delays)

Total Records : 129,880 records

Target: Satisfaction (Satisfied vs. Neutral/Dissatisfied)



DATASET PREPROCESSING

01

Handled missing values and outliers

02

Encoded categorical features

03

Normalized numerical columns

04

Train-test split (80/20)

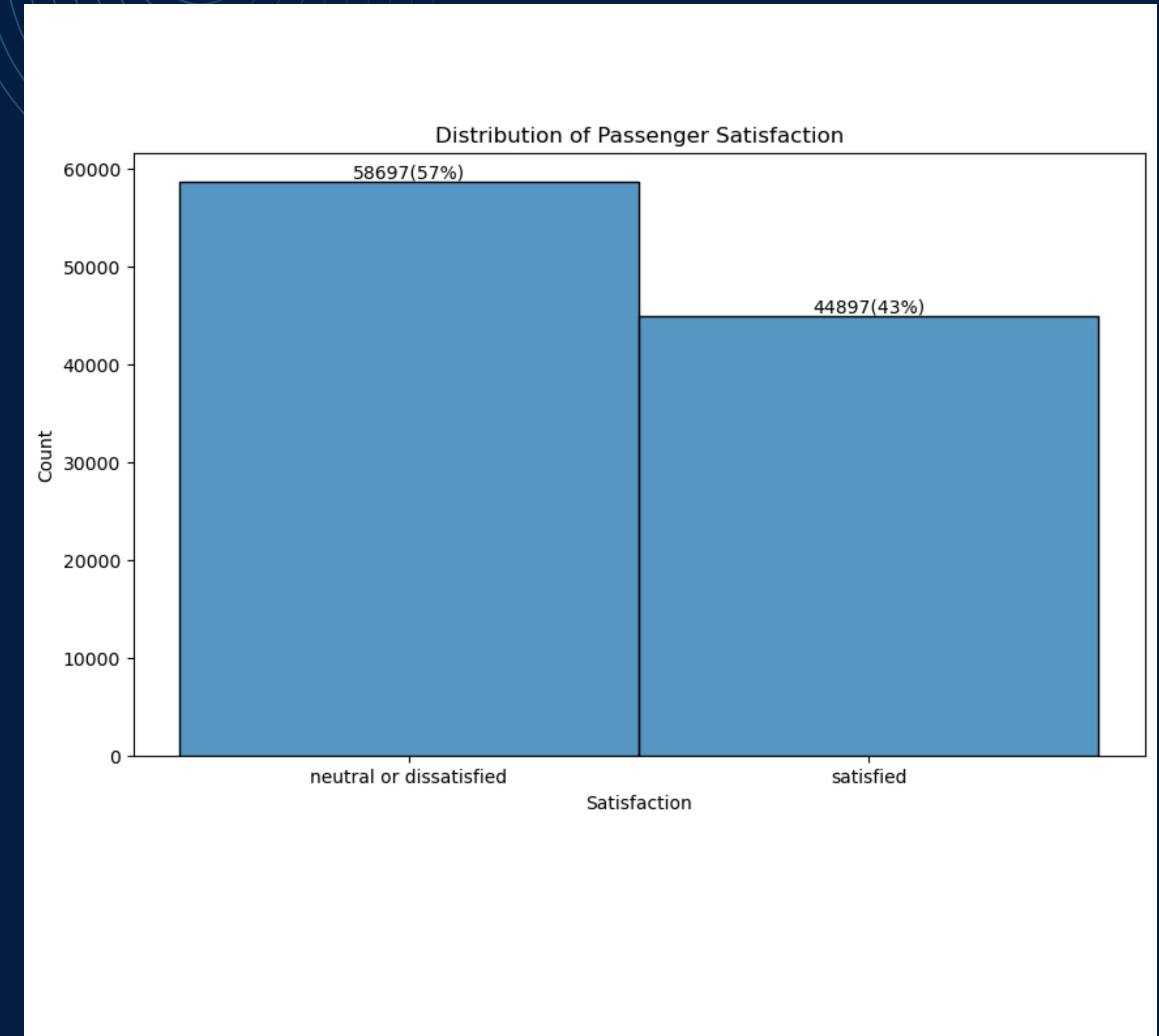


EXPLORATORY DATA ANALYSIS (EDA)



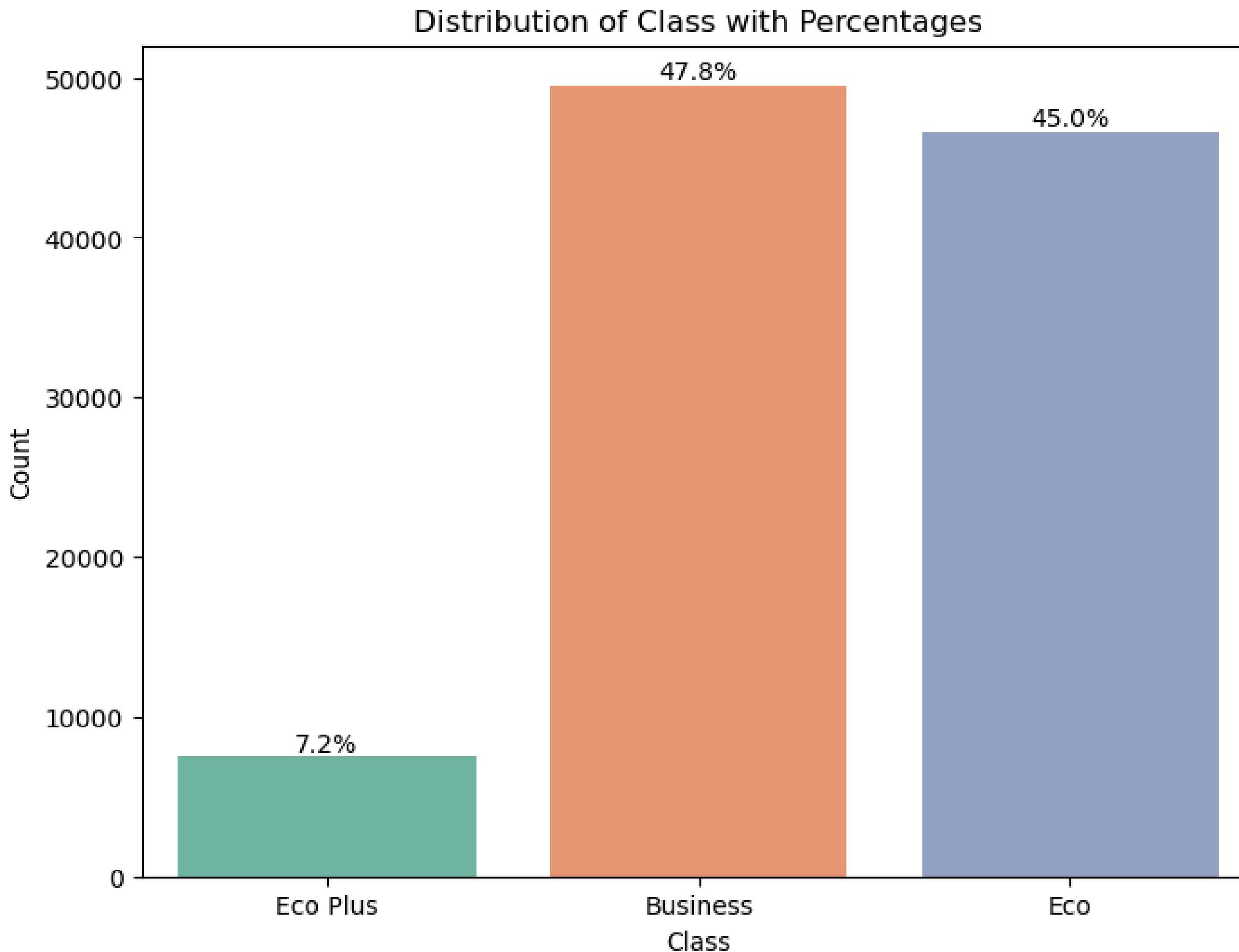
Passenger Satisfaction

- The graph shows a binary distribution of passenger satisfaction(i.e., target vector) where 57% (58,697) of passengers reported being neutral or dissatisfied with the service, while 43% (44,897) reported being satisfied. This indicates that a majority of passengers had a neutral or negative experience.
But why?



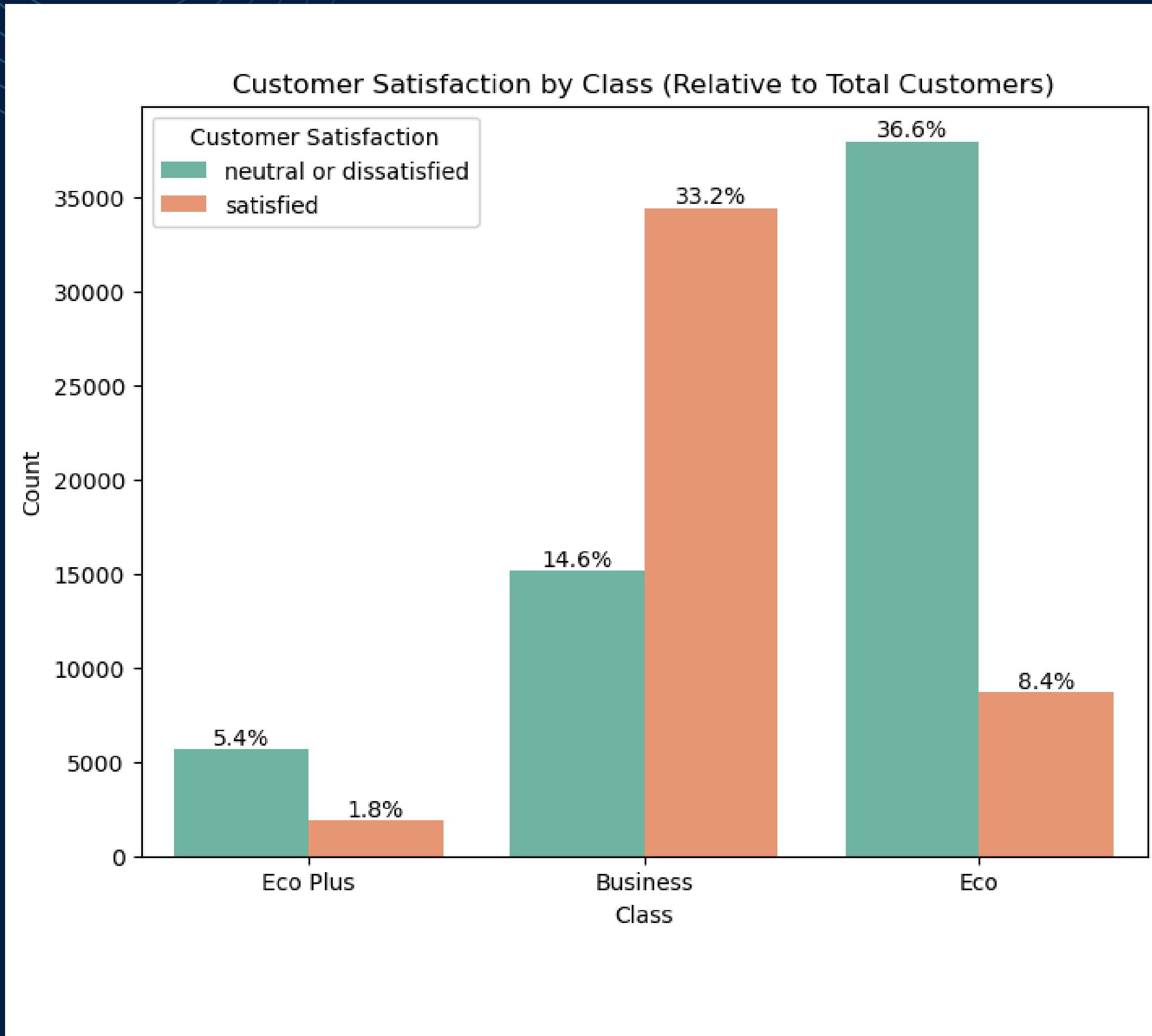
Distribution of Travel Classes

- The graph depicts the distribution of travel classes, with Business class representing the largest segment at 47.8% of passengers, followed closely by Economy class at 45.0%, while Economy Plus makes up only 7.2% of the dataset. This distribution suggests that the airline serves primarily Business and Economy travelers, with a small premium Economy segment.

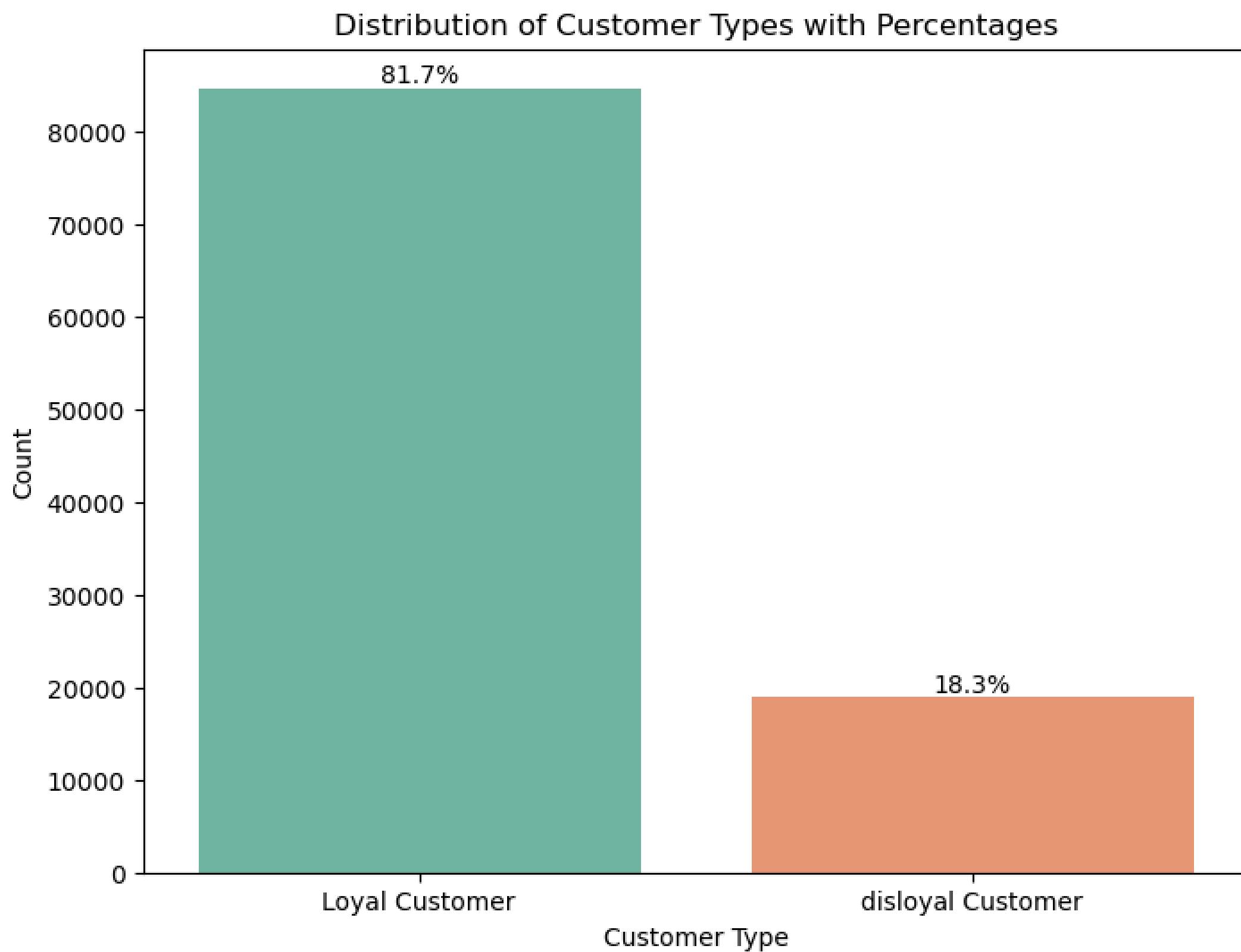


Distribution of Satisfaction across Travel Classes

- Business class passengers report the highest satisfaction rate with 33.2% of all customers being satisfied. Business travelers, while Economy class shows the highest dissatisfaction with 36.6% of all customers being neutral or dissatisfied. Economy travelers. Economy Plus has the lowest overall impact on satisfaction metrics, comprising just 7.2% of customers total (5.4% dissatisfied, 1.8% satisfied).



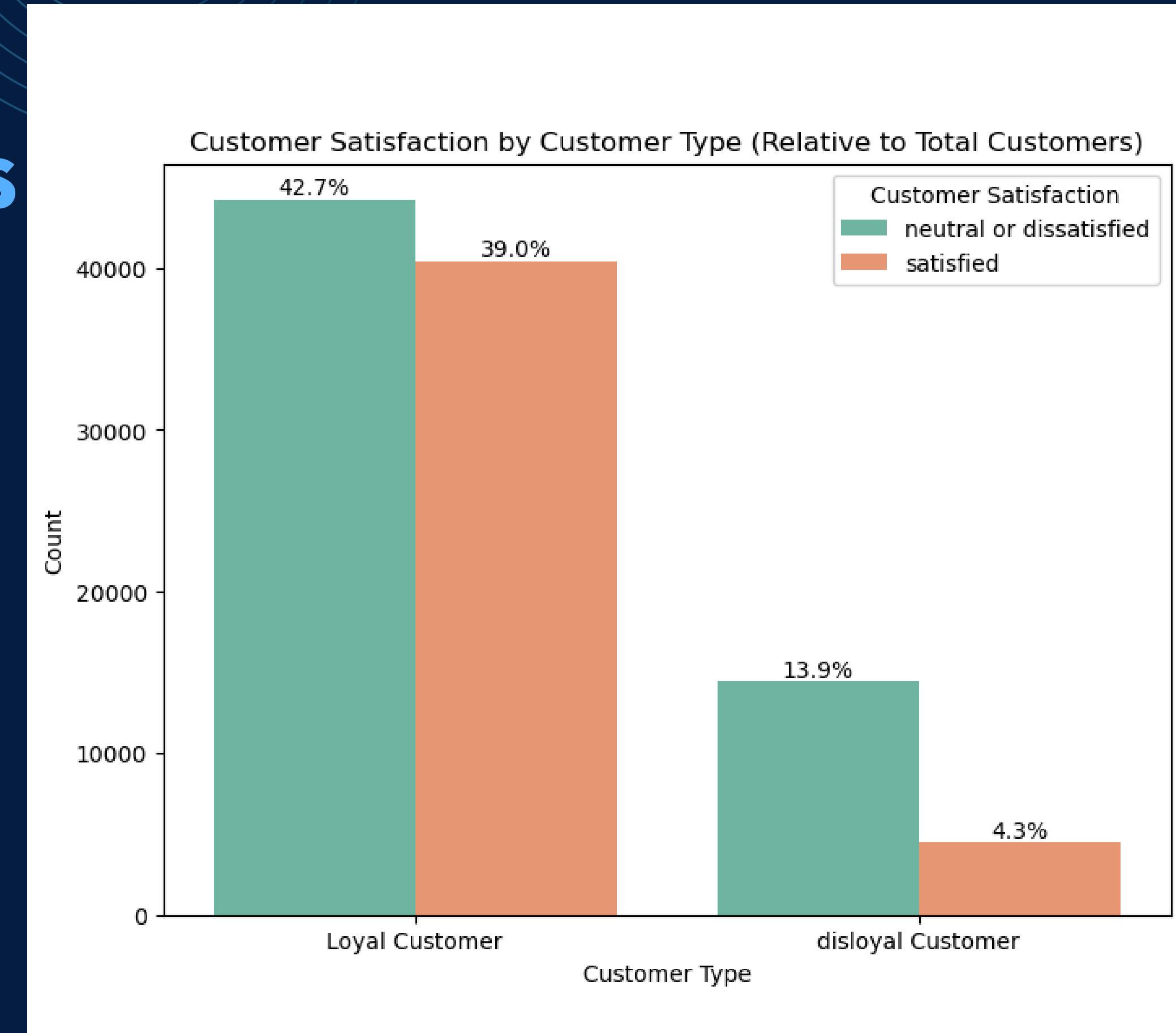
Distribution of Customer Types (loyal vs disloyal)



- The graph shows that loyal customers form the vast majority of the airline's customer base at 81.7%, while disloyal customers represent only 18.3%. This substantial difference indicates the airline has built a strong base of returning customers, though there remains a notable segment of non-repeat travelers to address.

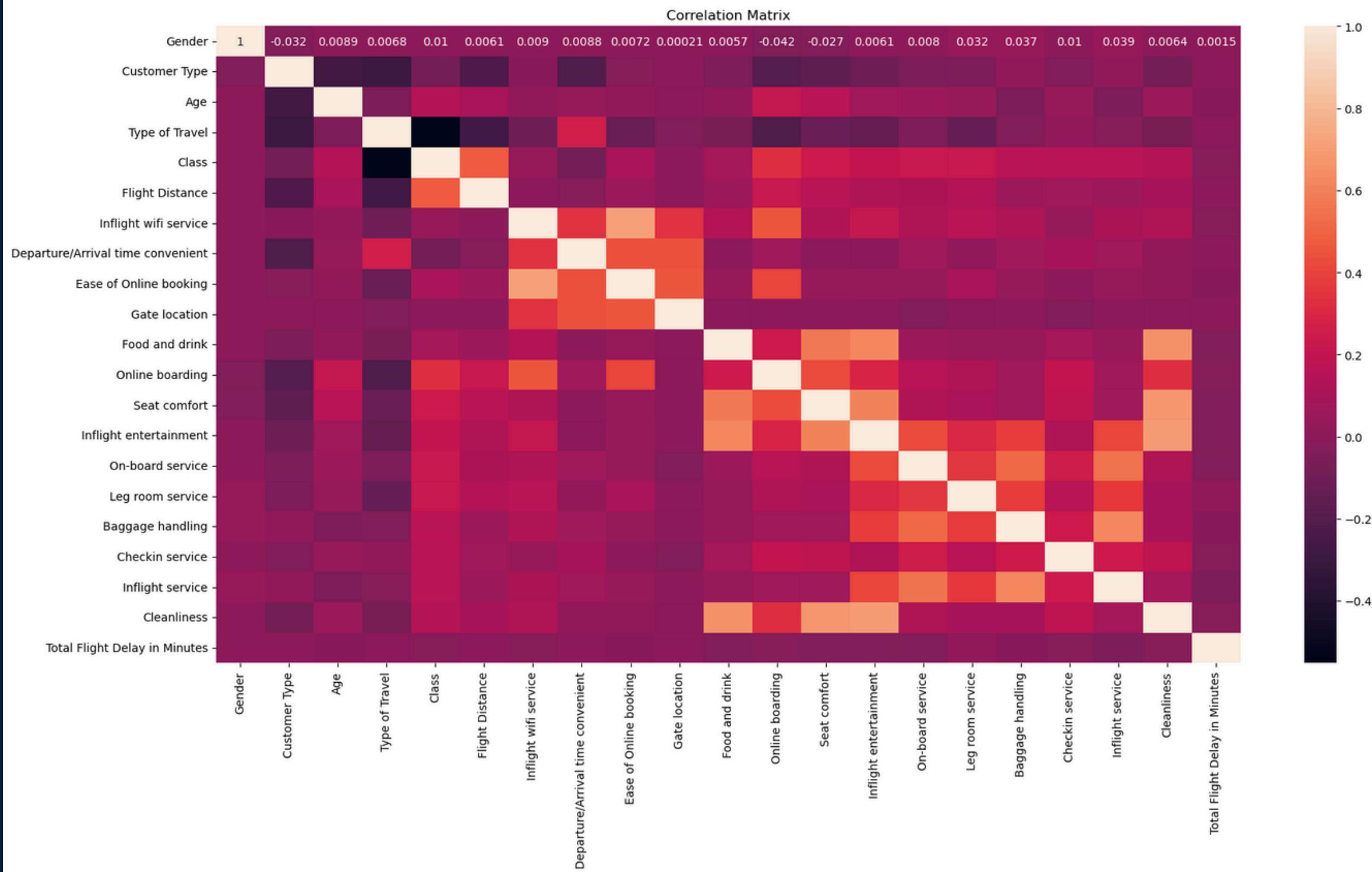
Distribution of Satisfaction across Customer Types

- The graph reveals that loyal customers have significantly higher satisfaction rates with 39.0% of all customers being satisfied loyal customers, compared to just 4.3% satisfied disloyal customers. However, loyal customers also represent the largest segment of neutral or dissatisfied travelers at 42.7%, while disloyal customers account for 13.9% of dissatisfaction, suggesting an opportunity to improve experiences for the airline's core customer base.



Service Quality Features Show Strong Interdependence

- The correlation matrix reveals strong positive correlations between service-related features (in-flight service, on-board service, seat comfort), suggesting these aspects of the customer experience are interconnected.
 - There's also moderate correlation between booking convenience features (online booking, departure/arrival time), while demographic factors like gender show minimal correlation with other variables.
 - This indicates that service quality measures tend to be evaluated together by passengers.



MACHINE LEARNING MODELLING

ML Models:

- Logistic Regression (base model)
- Support Vector Machine
- Random Forest

Evaluation Metrics

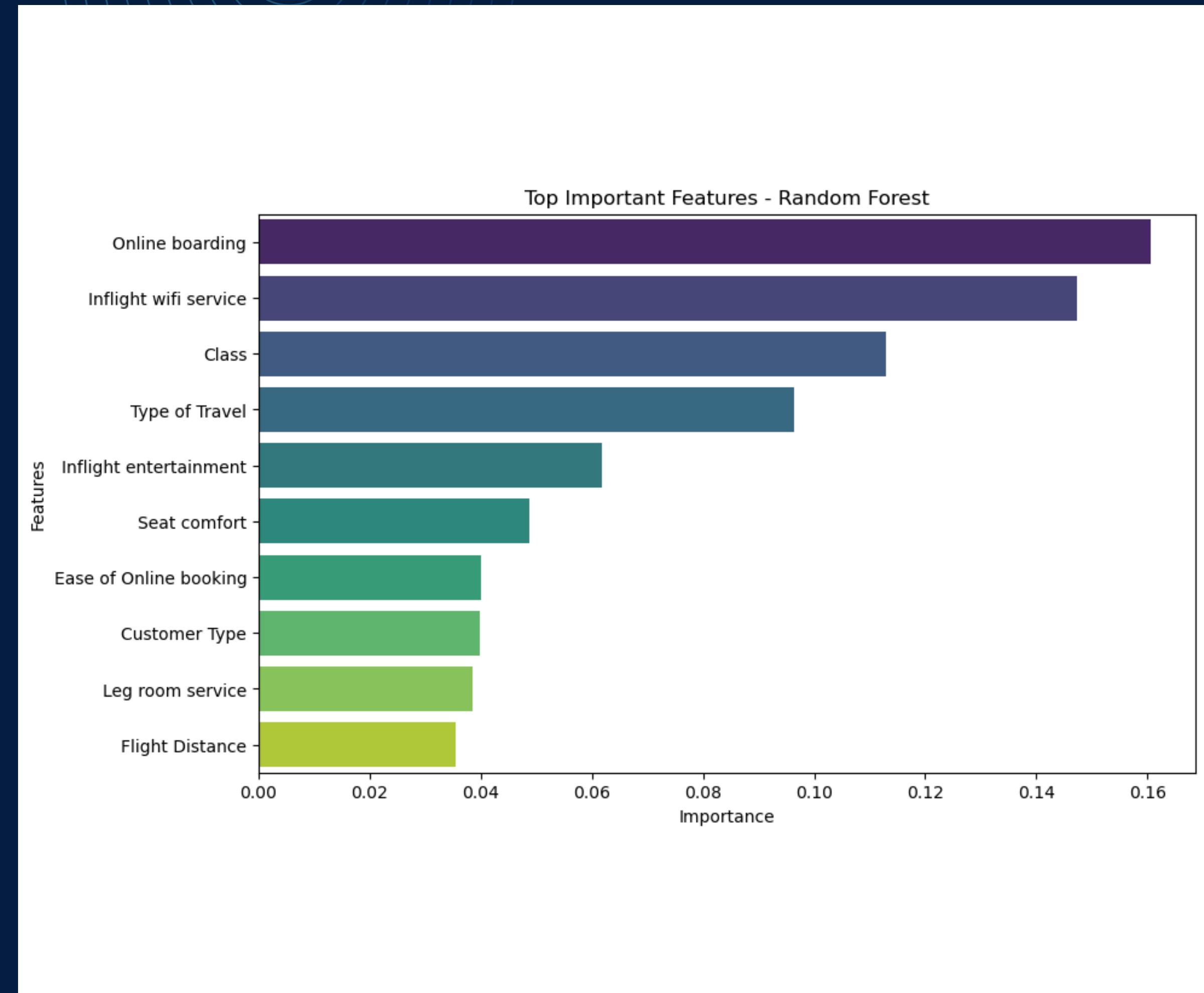
- Accuracy
- Precision
- Recall
- F1 Score

MODEL PERFORMANCE

MODEL	ACCURACY	PRECISION	RECALL	F1-SCORE
LOGISTIC REGRESSION	87.28%	87.21%	83.11%	85.11%
SUPPORT VECTOR MACHINE	95.53%	96.14%	93.54%	94.83%
RANDOM FOREST	96.24%	97.37%	93.93%	95.62%
RANDOM FOREST (TUNED)	96.25%	97.38%	93.97%	95.64%

Feature Importance

- The Random Forest model identifies online boarding and in-flight WiFi service as the most important predictors of passenger satisfaction, followed by travel class and type of travel. Digital touchpoints (boarding, WiFi) have substantially higher importance scores than physical comfort features (seat comfort, leg room), suggesting technological aspects of the journey significantly influence overall satisfaction ratings.





BUSINESS INSIGHTS & STRATEGIC RECOMMENDATIONS

Optimize Digital Experience

- Revamp online check-in and booking systems
- Integrate real-time support (e.g., chatbots)

Enhance Inflight Connectivity

- Partner with premium Wi-Fi providers
- Offer tiered access: basic (free) and premium (paid)

Upgrade Economy Class Amenities

- Improve seat comfort, legroom, and onboard service
- Rebrand Economy Plus with value-added benefits

Segment Customer Services

- Tailor services for business (efficiency) vs. leisure (comfort)
- Use targeted promotions based on travel type and loyalty

Implement Feedback Loops

- Collect feedback on key services post-flight
- Drive continuous improvement through customer insights

FUTURE WORK

Strategic Next Steps for Model Enhancement and Deployment

Advanced Model Optimization

- Apply XGBoost and LightGBM for higher accuracy and generalization

Interactive Dashboard Deployment

- Build a web-based tool using Streamlit or Dash
- Allow teams to visualize and predict satisfaction in real time

Explainable AI Integration (SHAP / LIME)

- Increase model transparency with explainability tools
- Empower stakeholders to understand why certain predictions are made

Thank You

Connect with me.



[project repository](#)



adinomuruthil@gmail.com



RESOURCE PAGE

