



# **Airline Passenger Satisfaction**

# TABLE OF CONTENTS

- Objective
- Dataset Overview
- Exploratory Data Analysis
- Predictive Modeling
- Visualizations
- Clustering Analysis
- Dashboard
- Conclusion

# OBJECTIVE



**Our goal is to tackle the challenge of enhancing passenger satisfaction within the competitive airline industry by pinpointing areas where service can be improved and enriching the overall customer experience.**

# Dataset Overview

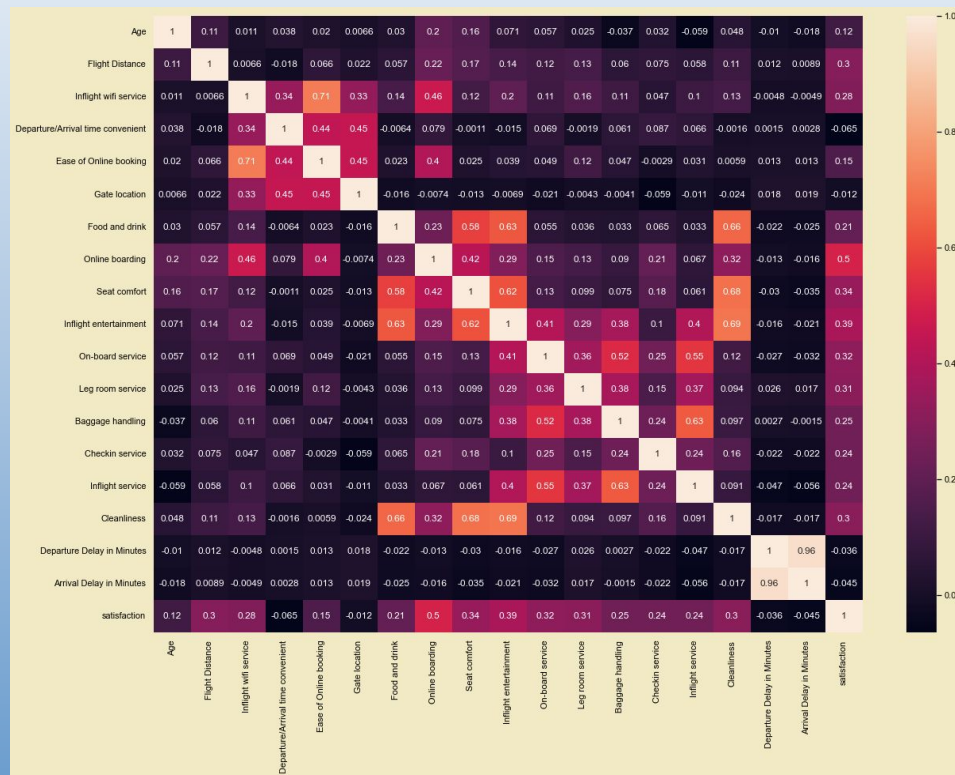
## Airline Satisfaction Data

- **Categorical Data:** Gender, Customer Type, Type of Travel, Class, 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, Satisfaction.
- **Numerical Data:** Age, Flight Distance, Departure Delay in Minutes, Arrival Delay in Minutes.

## Airlines Reviews

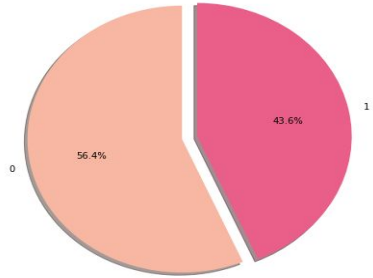
- **Text Data:** Name, Reviews, and Route.
- **Categorical Data:** Airline, Verified, Type of Traveller, Month Flown, Route, and Class.
- **Numerical Ratings:** Seat Comfort, Staff Service, Food & Beverages, Inflight Entertainment, Value For Money, and Overall Rating.
- **Binary Data:** Recommended is a binary field, likely containing values such as Yes/No.

# Exploratory Data Analysis

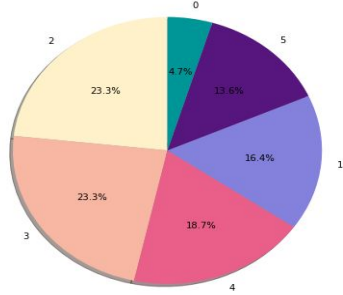


# Exploratory Data Analysis

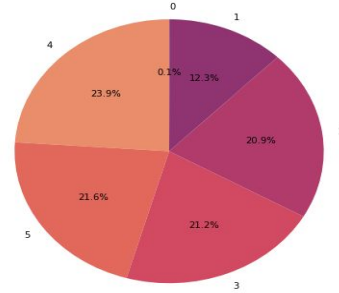
satisfaction Distribution



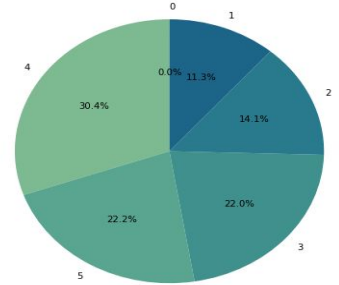
Ease of Online booking Distribution



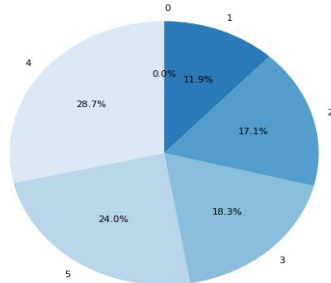
Food and drink Distribution



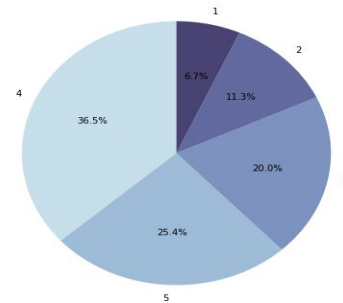
On-board service Distribution



Inflight entertainment Distribution



Baggage handling Distribution



# Predictive Modelling

```
# Selecting the specified predictor variables
predictors = ['Gender', 'Customer Type', 'Age', 'Type of Travel', 'Class', 'Flight Distance',
              'Inflight wifi service', 'Departure/Arrival time convenient', 'Ease of Online booking']
target = 'satisfaction'

# Encode categorical variables
for column in ['Gender', 'Customer Type', 'Type of Travel', 'Class']:
    df[column] = LabelEncoder().fit_transform(df[column])

# Ensure target variable is encoded if it's categorical
df[target] = LabelEncoder().fit_transform(df[target])

# Features and Target
X = df[predictors]
y = df[target]

# Splitting the dataset into training and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=42)

# Initialize and train the Logistic Regression model
model = LogisticRegression(max_iter=1000)
model.fit(X_train, y_train)

# Predicting the test set results
y_pred = model.predict(X_test)

# Evaluation metrics
print(classification_report(y_test, y_pred))
print("F1 Score:", f1_score(y_test, y_pred))

# Confusion Matrix
conf_matrix = confusion_matrix(y_test, y_pred)
print("Confusion Matrix:\n", conf_matrix)

TN = 1432
FP = 262
FN = 236
TP = 1061

Accuracy = (TP + TN) / (TP + TN + FP + FN)
Precision = TP / (TP + FP)
Recall = TP / (TP + FN)
F1_Score = 2 * (Precision * Recall) / (Precision + Recall)

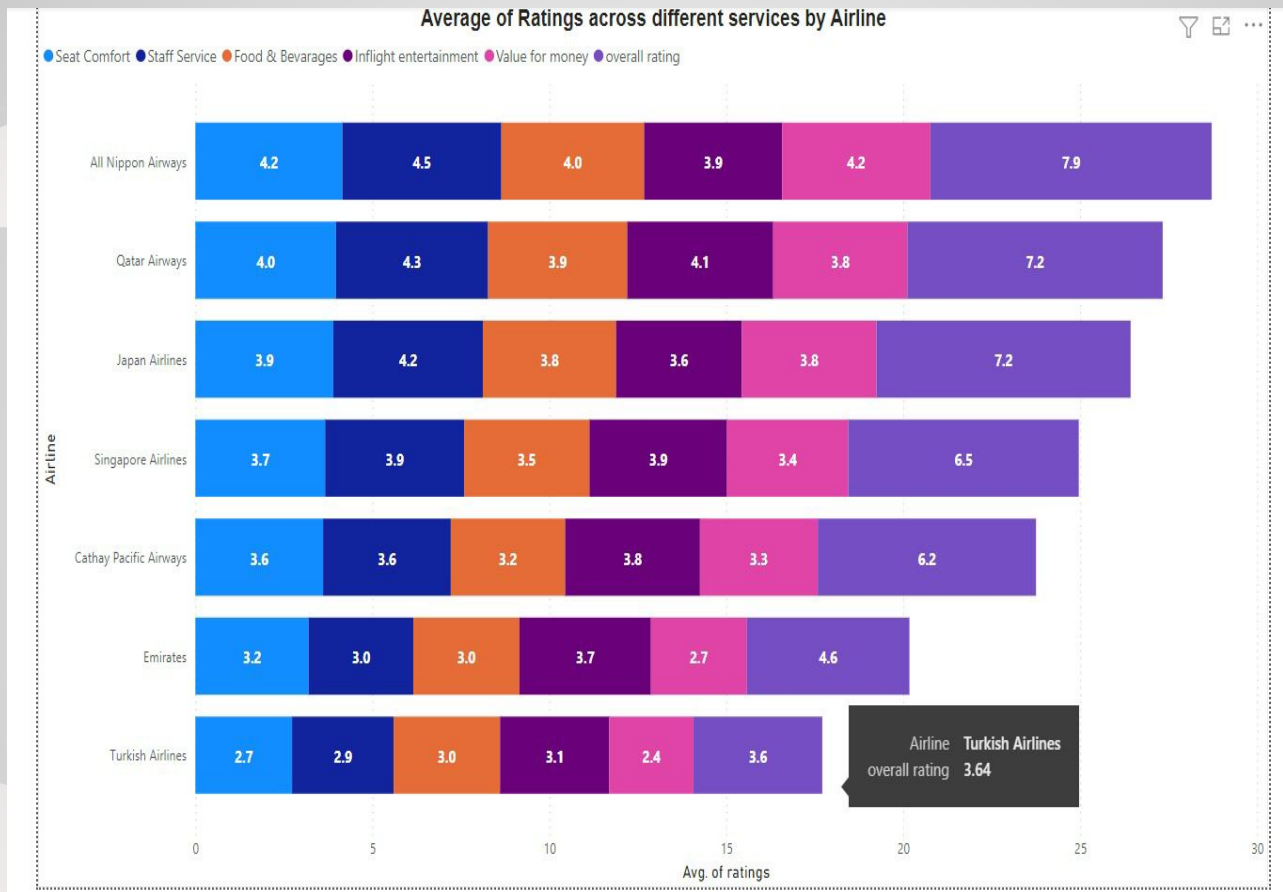
print("Accuracy:", Accuracy)
print("Precision:", Precision)
print("Recall:", Recall)
print("F1 Score:", F1_Score)

# Example Prediction
example = np.array([[1, 1, 45, 0, 1, 500, 3, 2, 3]]) # Example input for the predictors
example_prediction = model.predict(example)
print("Example Prediction:", "Satisfied" if example_prediction[0] == 1 else "Not Satisfied")
```

```
...: print("Confusion Matrix:\n", conf_matrix)
Confusion Matrix:
[[1432  262]
 [ 236 1061]]

In [102]:
...: TN = 1432
...: FP = 262
...: FN = 236
...: TP = 1061
...:
...: Accuracy = (TP + TN) / (TP + TN + FP + FN)
...: Precision = TP / (TP + FP)
...: Recall = TP / (TP + FN)
...: F1_Score = 2 * (Precision * Recall) / (Precision + Recall)
...:
...: print("Accuracy:", Accuracy)
...: print("Precision:", Precision)
...: print("Recall:", Recall)
...: print("F1 Score:", F1_Score)
Accuracy: 0.8335005015045135
Precision: 0.8019652305366591
Recall: 0.818041634541249
F1 Score: 0.8099236641221375
```

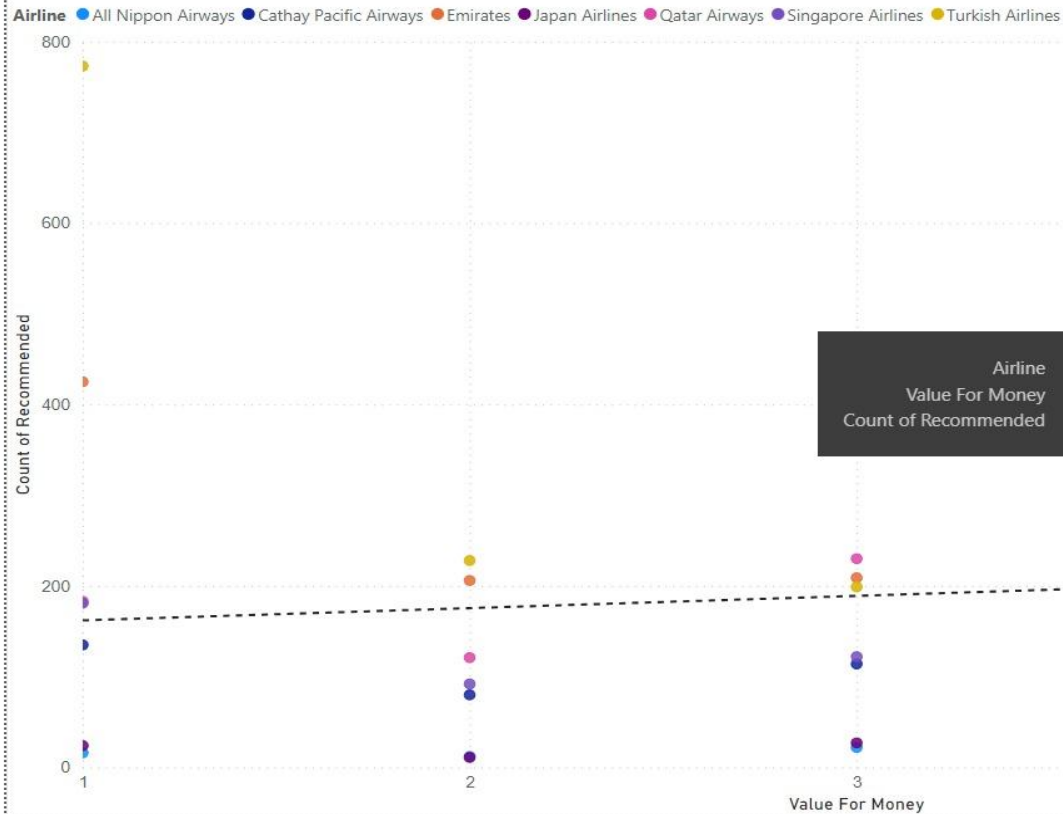
Prediction of airline customer satisfaction using pre-flight and demographic variables



This stacked bar chart effectively compares airlines across multiple service metrics, highlighting strengths and weaknesses in areas like seat comfort and staff service, and providing a quick visual summary of overall customer satisfaction and value for money.



Count of Recommended by Airline and Value For Money



This scatter plot charts the relationship between 'Value for Money' ratings and the count of recommendations for different airlines. The chart is beneficial for quickly assessing which airlines are perceived as providing good value and are likely to be recommended by passengers.

# Clustering Analysis

## Identified 3 main customer personas

### Short Distance Flyer

- Younger
- Flies shorter distances
- Mostly Econ Class
- Biggest share of disloyal
- Mixed business and personal travel
- **Biggest share of dissatisfied**



### Moderate Flyer

- Moderate in all aspects
- **Average satisfaction**

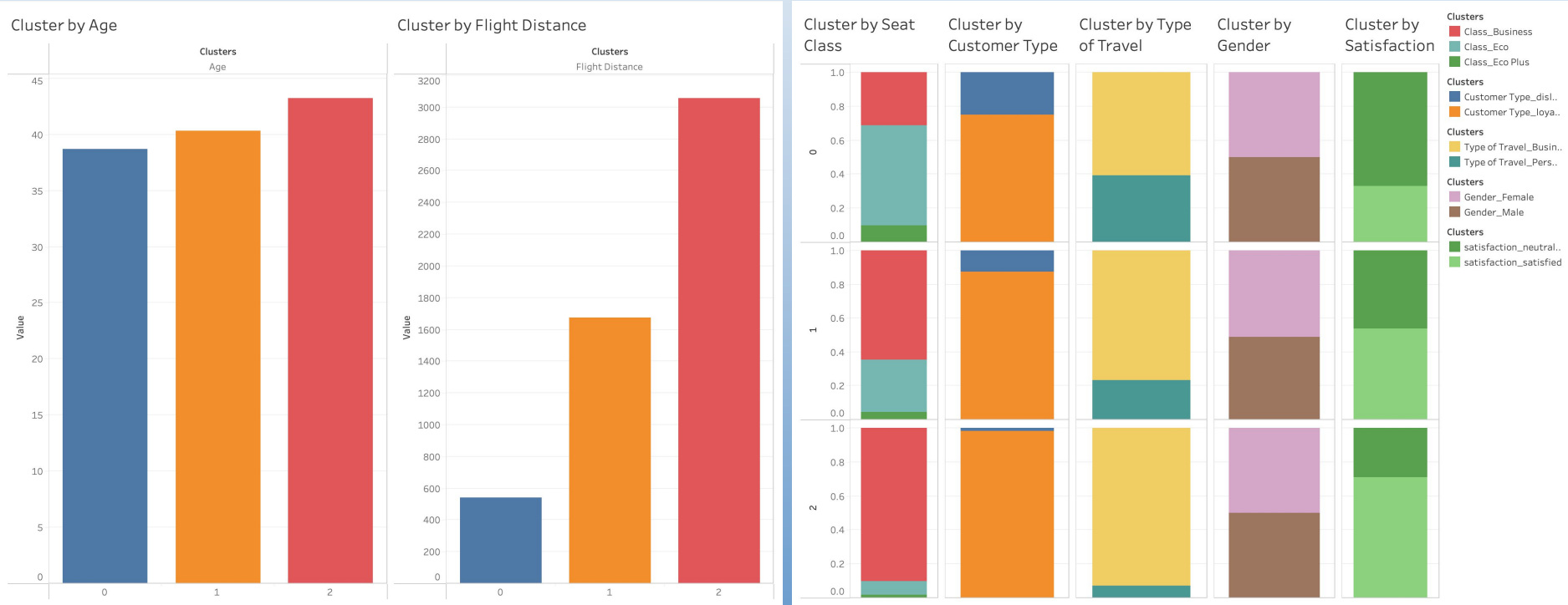


### Seasoned Traveler

- Older
- Long haul flights
- Mostly Business Class
- Very loyal
- Predominant business travel
- **Biggest share of satisfied**

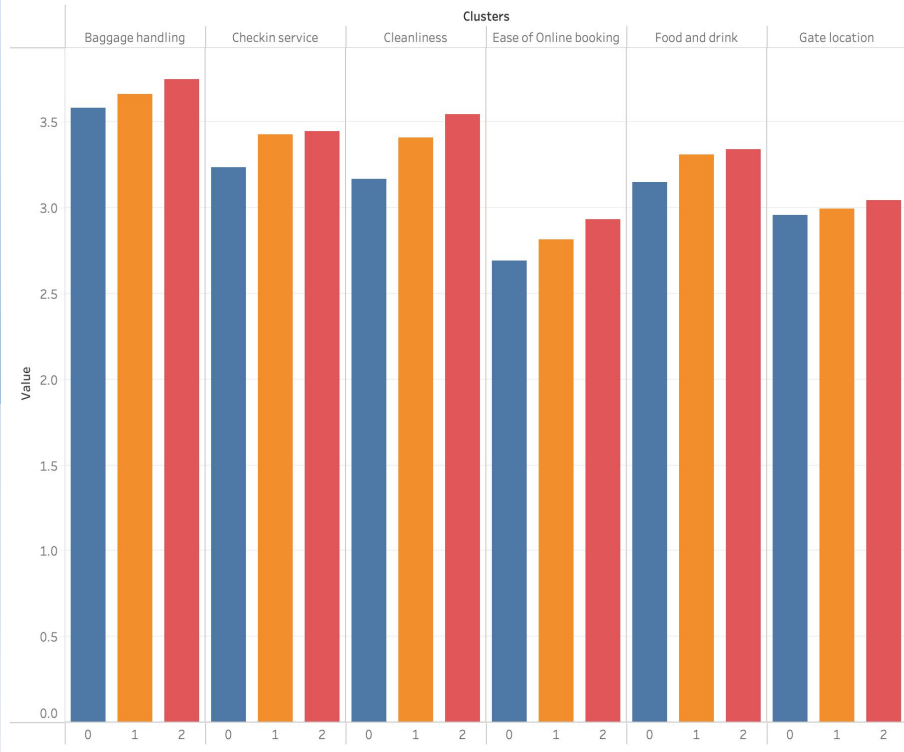


# Persona Features

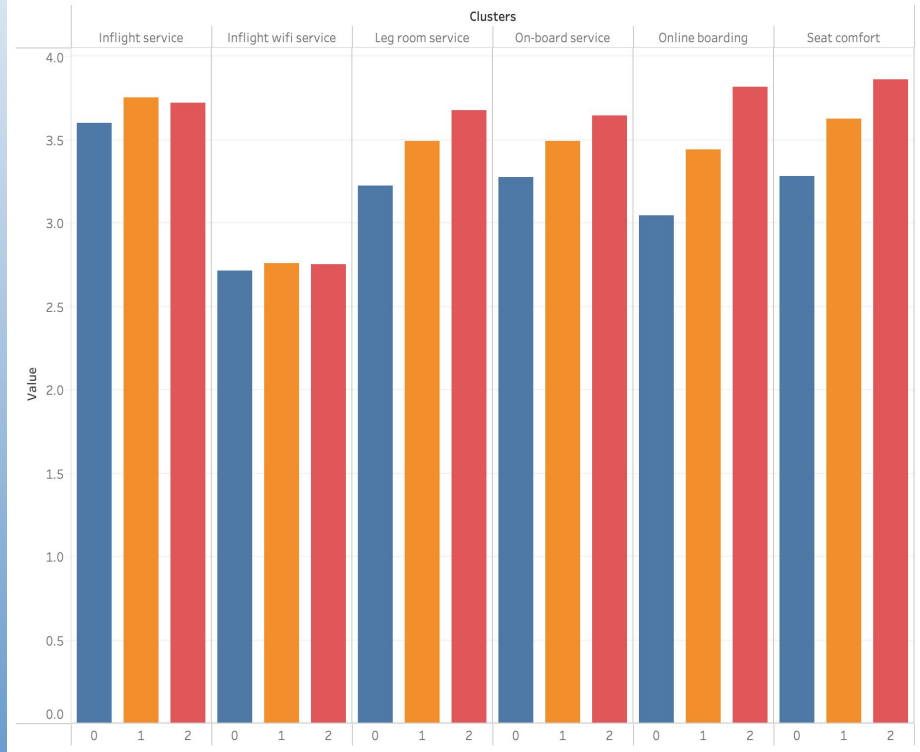


# Persona & Attributes

Satisfaction per Cluster and Attribute



Satisfaction per Cluster and Attribute



# Clustering - Main Takeaways

## Biggest Satisfaction gaps

- Leg room service
- On-board service
- Online boarding
- Seat Comfort

## Areas to improve

- In flight wifi service
- Ease of online booking
- Gate location

## Suggestions

- Focus on lower satisfaction areas
- Explore consumer behavior by creating targeted approaches

## Further Research

- Relationship between satisfaction and willingness to pay

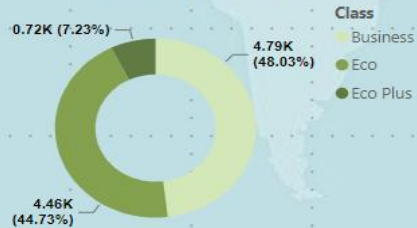
# POWER BI DASHBOARD

Total Passengers  
9,970 passengers

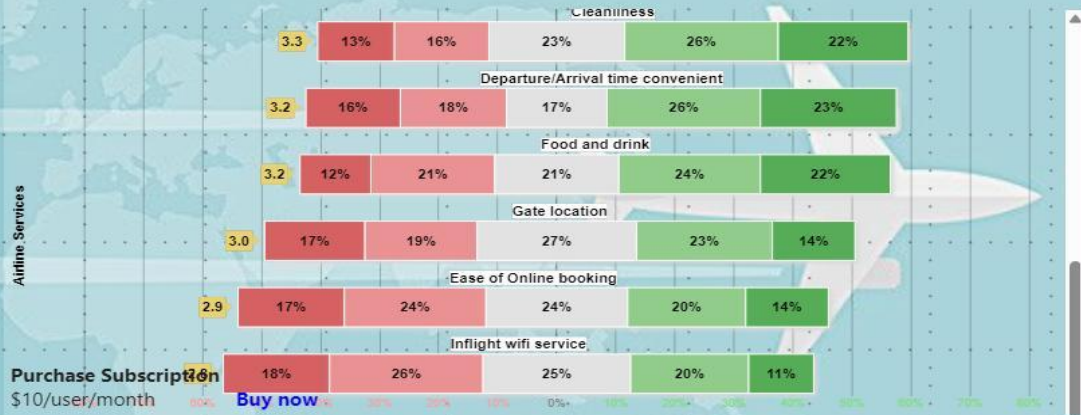


**Inflight WIFI service and Online Ticket Booking services have the highest level of dissatisfaction among passengers, with the majority of responses falling into the lowest satisfaction categories**

Satisfaction by Class



Services with Less Satisfaction Ratings among Passengers



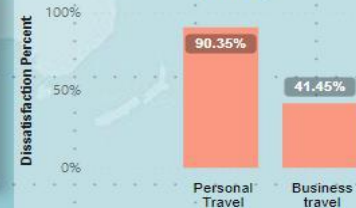
14.83

Average Arrival Delay in Mi...

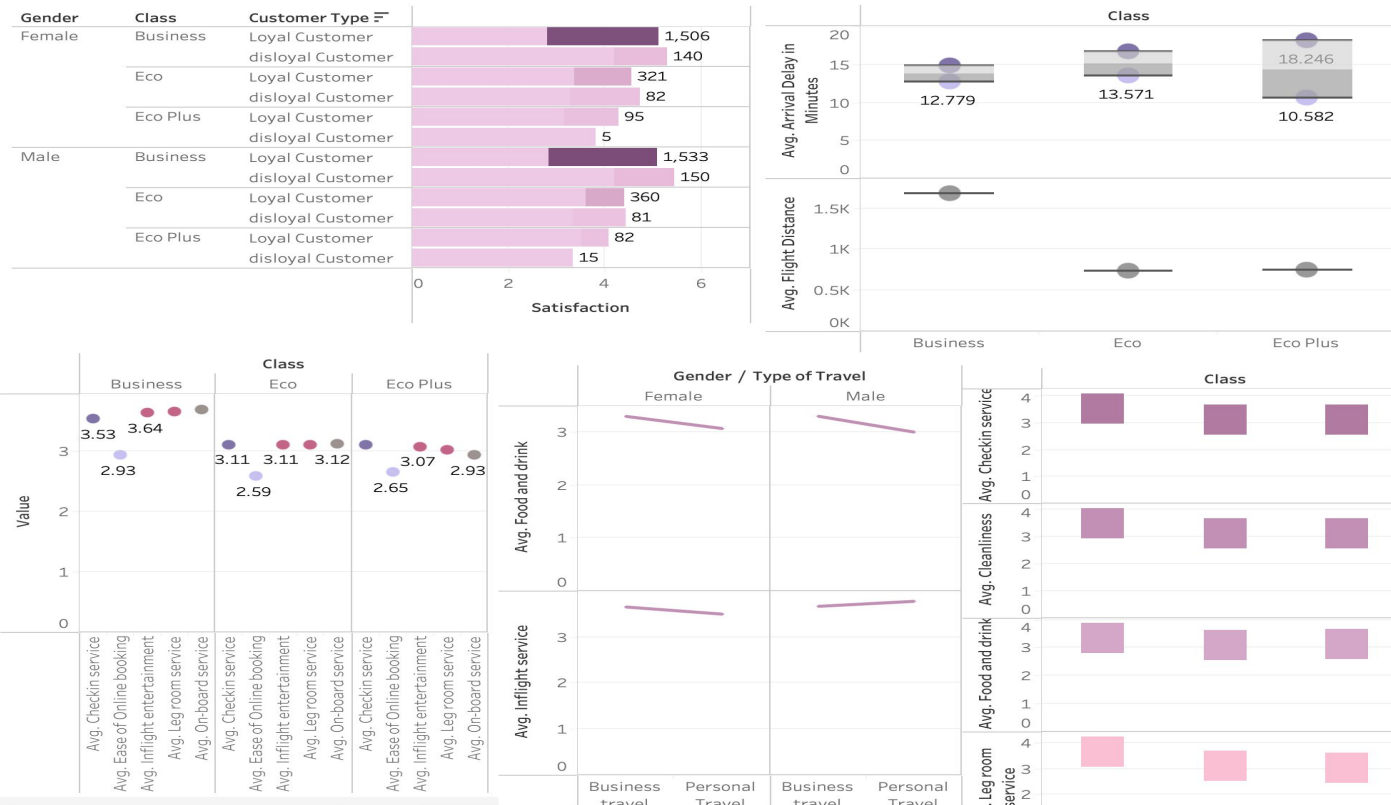
14.43

Average of Departure Delay in Minutes

Dissatisfaction Percent by Type of Travel



# TABLEAU DASHBOARD



# Findings

## Impact of Service Quality on Satisfaction:

- Key factors: Inflight WiFi, online booking, seat comfort.
- High satisfaction correlates with exceeding expectations in these areas.

## Demographic and Travel Type Influences:

- Business class and loyal customers show higher satisfaction.
- Business travel rated better than personal travel, highlighting the need for tailored services.

## Operational Efficiency and Delays:

- On-time performance and baggage handling are crucial for satisfaction.
- Delays, particularly in Eco Plus, negatively impact satisfaction.





# Recommendations

**Enhance Inflight Services:** Prioritize WiFi and seat comfort improvements for higher satisfaction.

**Operational Improvements:** Boost satisfaction by ensuring on-time flights and efficient baggage handling.

**Marketing Strategies:** Emphasize reliability and value in campaigns to appeal to price-sensitive segments.

**Feedback Mechanisms:** Establish a system for gathering and analyzing feedback, focusing on key improvement areas.

**Data-Driven Decision Making:** Leverage analytics for ongoing service adaptation and satisfaction monitoring.



**Thank You**