Reducing Customer Churn in the airline Industry

Team : Jimit Mistry, Vasundra Patil, Sanjana Bhot, Sanyukta Sandhilya, Sourabh Ghosh

BUSINESS QUESTIONS

- Which characteristics contribute towards making the Customer a Promoter or a Detractor for the airline?

- In which areas does the airline needs to improve, and how it can improve to reduce customer churn?

- How can the Customer be changed from a Detractor or a Passive customer to a Promoter?

Course of Action :

1. Identifying Target label (Detractor and Promoter) and visualising it as a classification problem
2. Cleaning the data
3. Exploratory Data Analysis to understand and generate distribution graphs
4. Feature Engineering and data transformations
5. Implementation of classification and regression models to get actionable insights
6. Conclusion and Inferences based on characteristic features
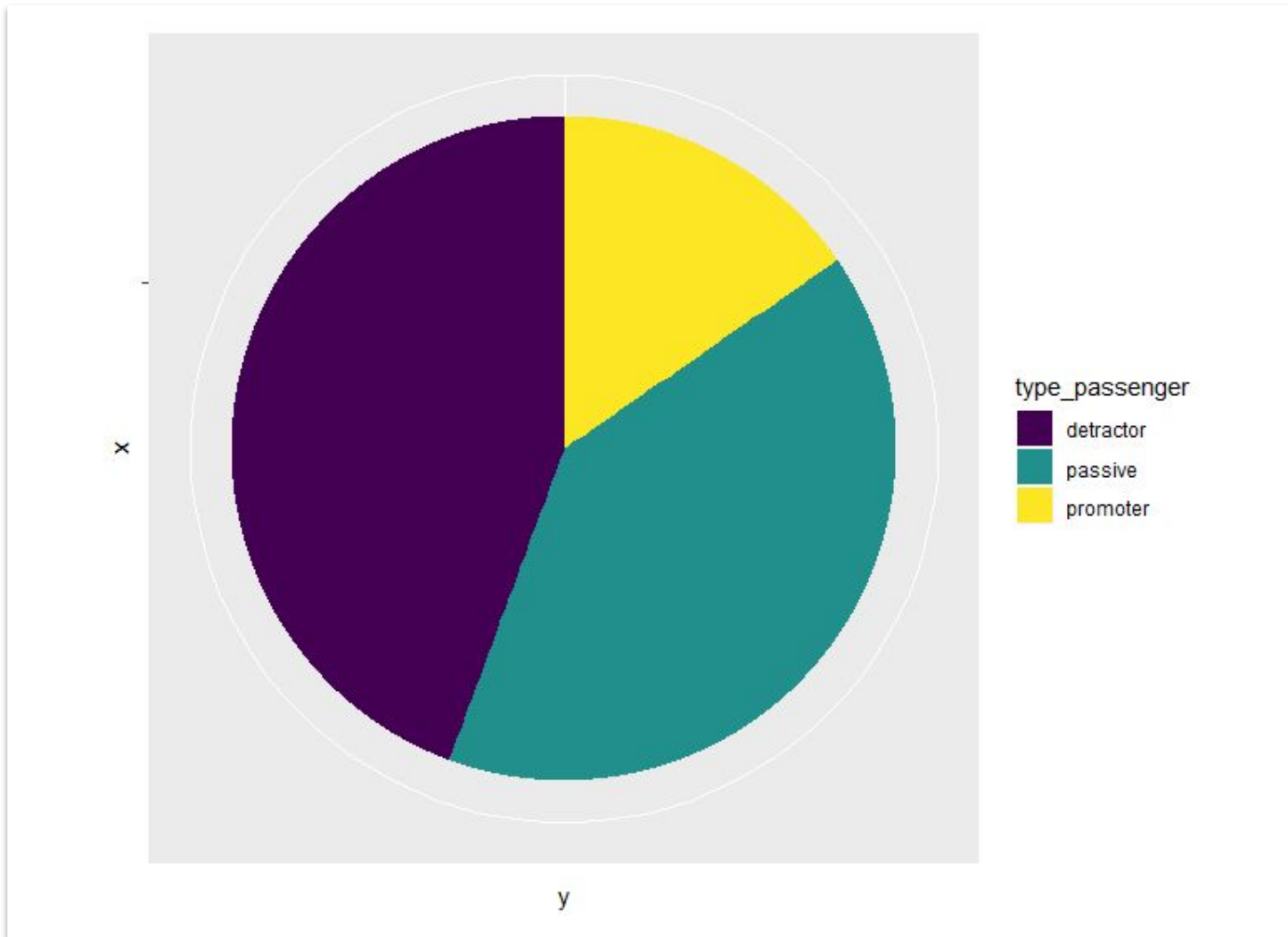
Recommender Type

Identifying the labels:
- Detractor: Someone who is least likely to return and recommend the airline
- Passive: Someone who has an unbiased opinion of the airline
- Promoter: Someone who will most likely return, and also recommend the airline for increasing profits
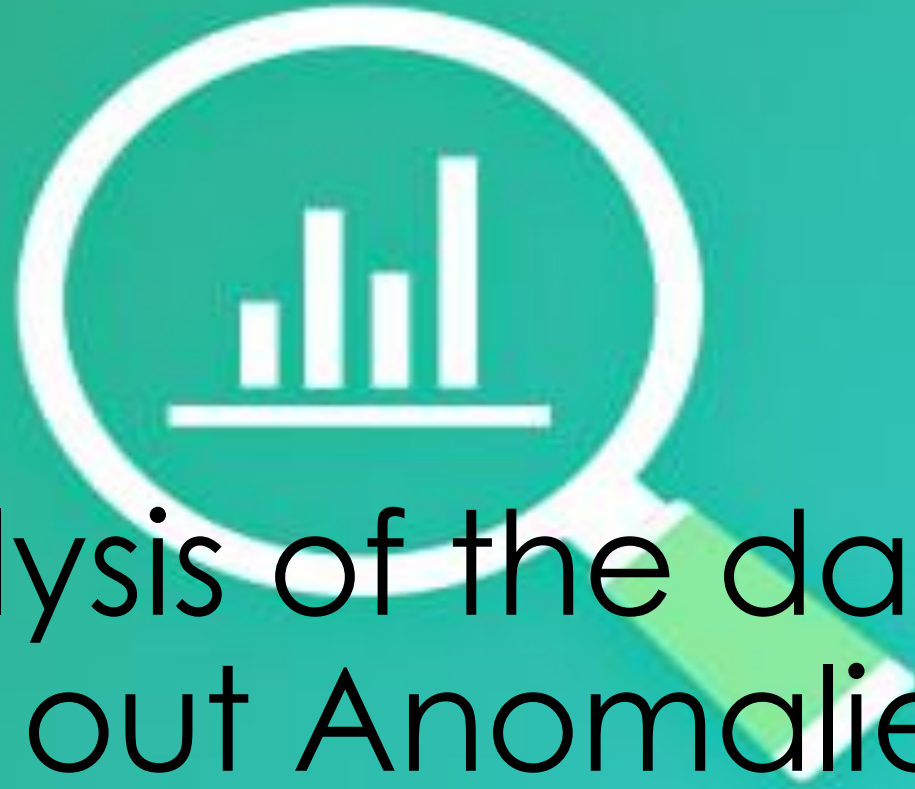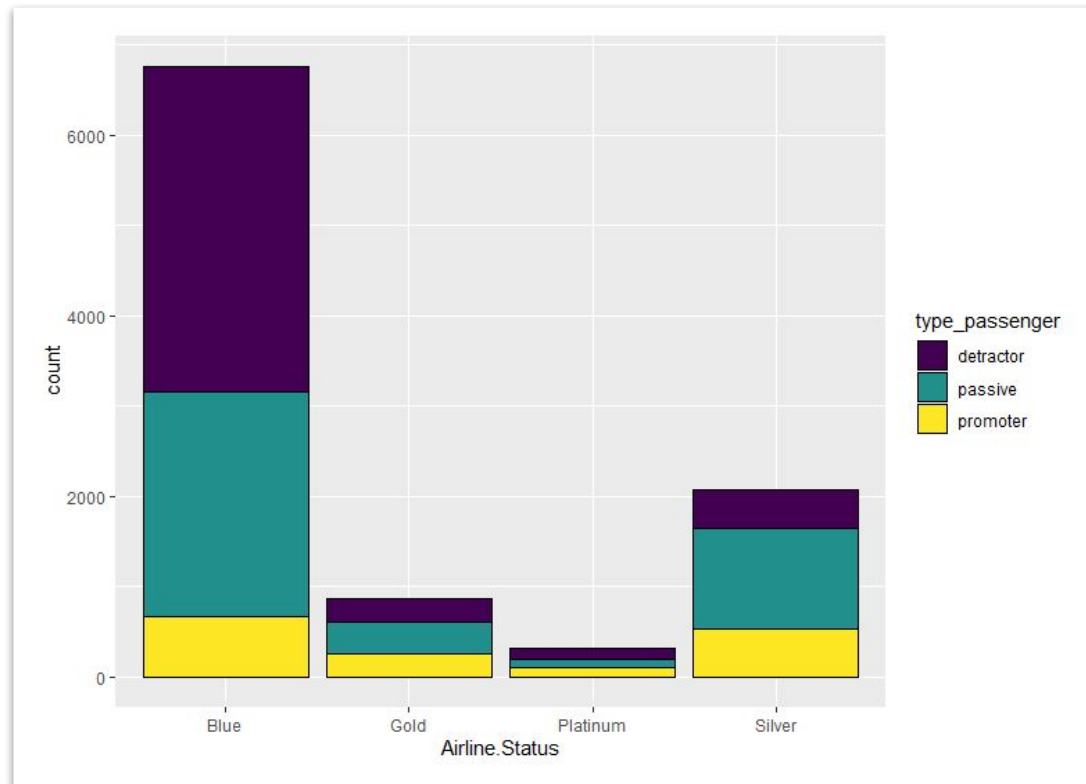
Analysis of the data to figure out Anomalies and the general characteristics
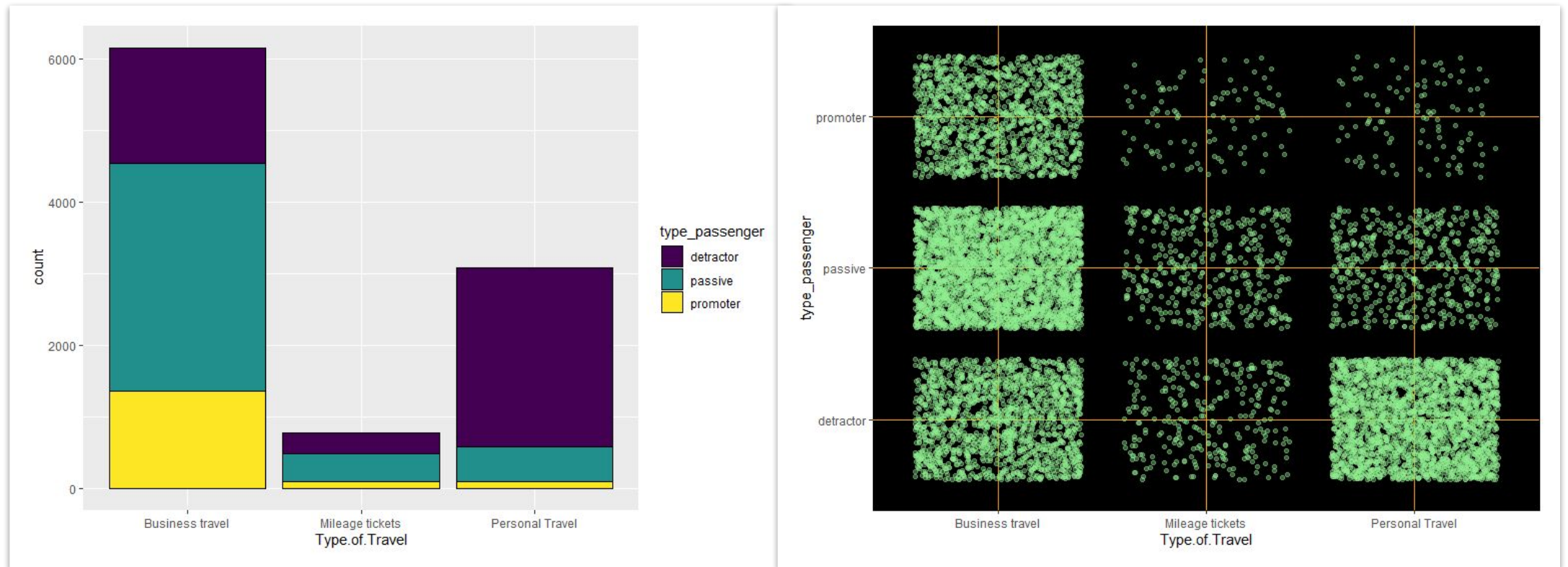
Loyalty Scores
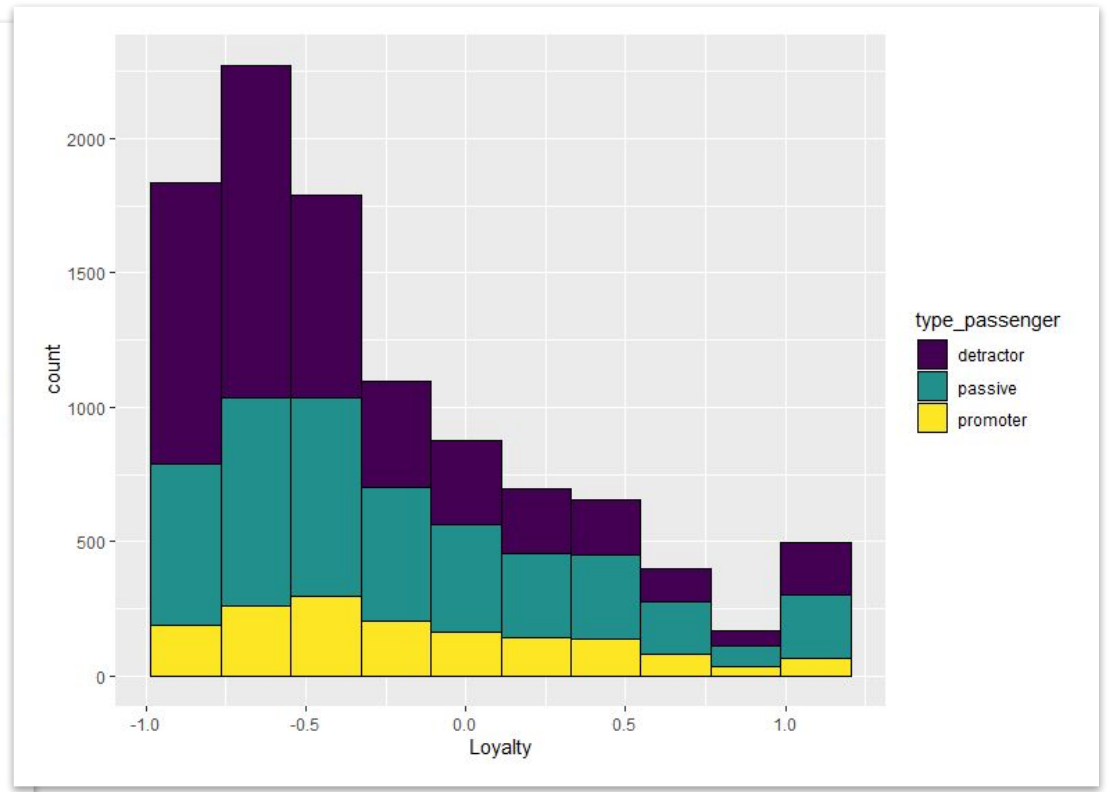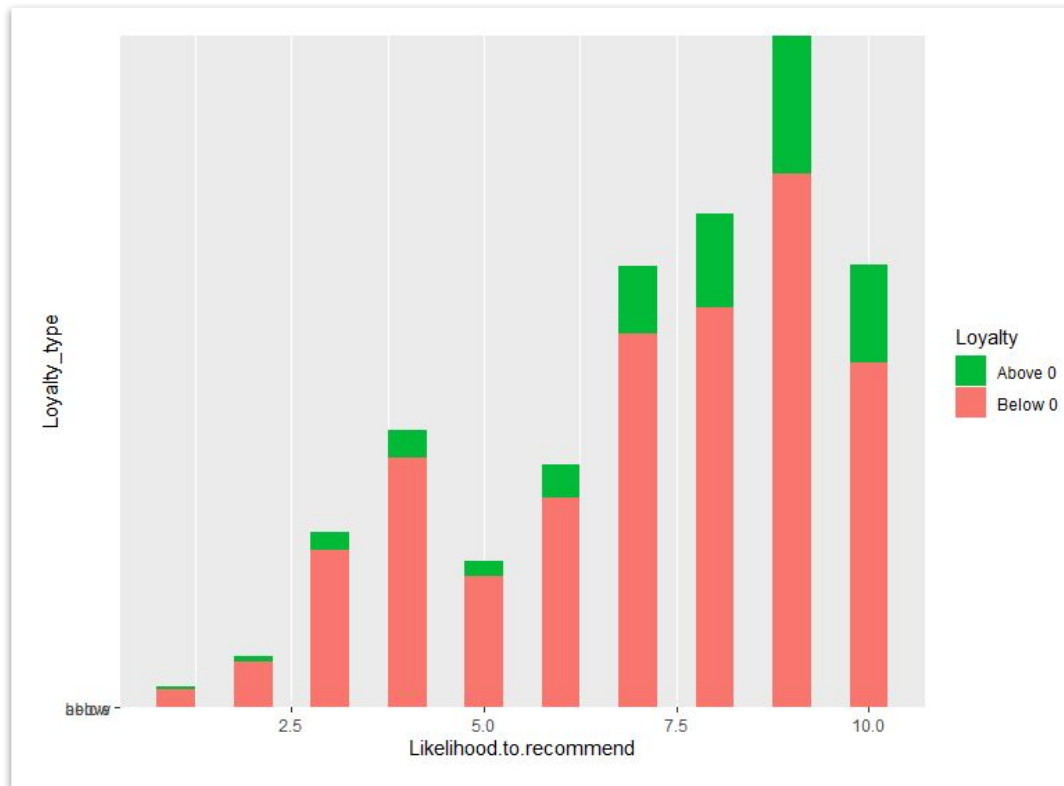
An index of loyalty ranging from -1 to 1 that reflects the proportion of flights taken on other airlines versus flight taken on this airline. A higher means more loyalty.
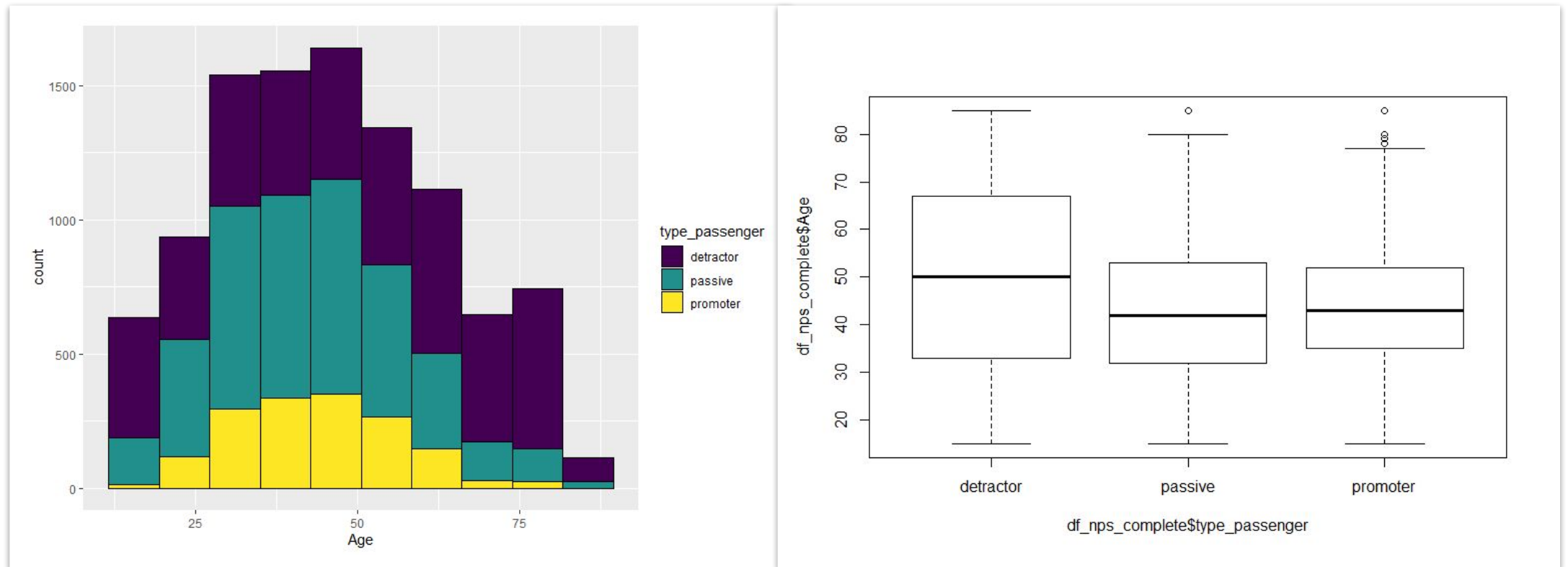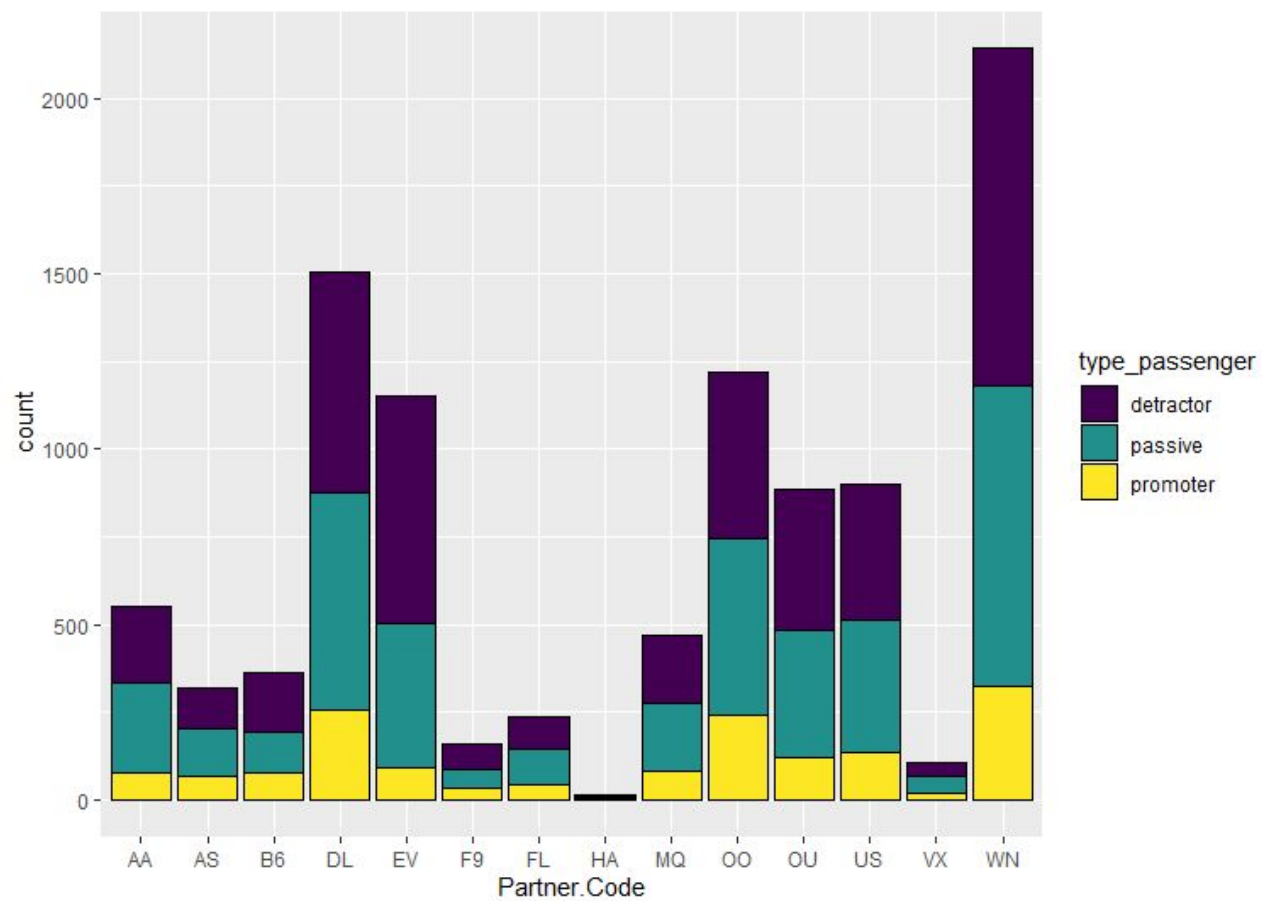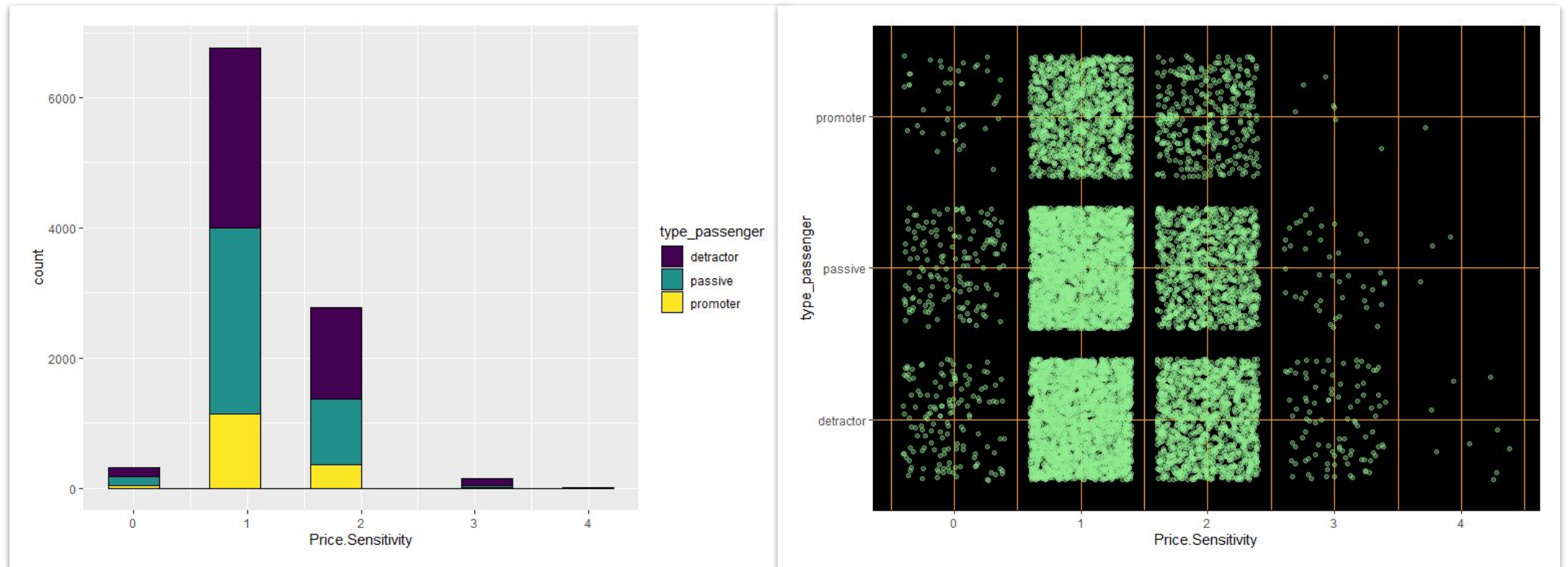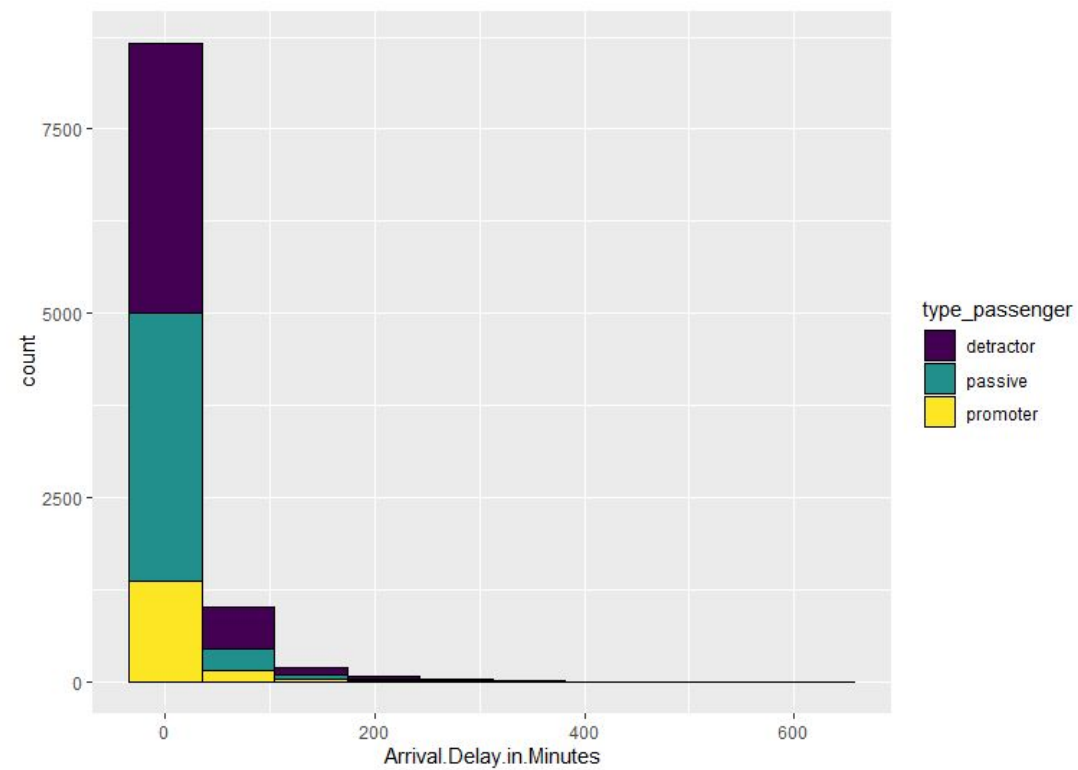
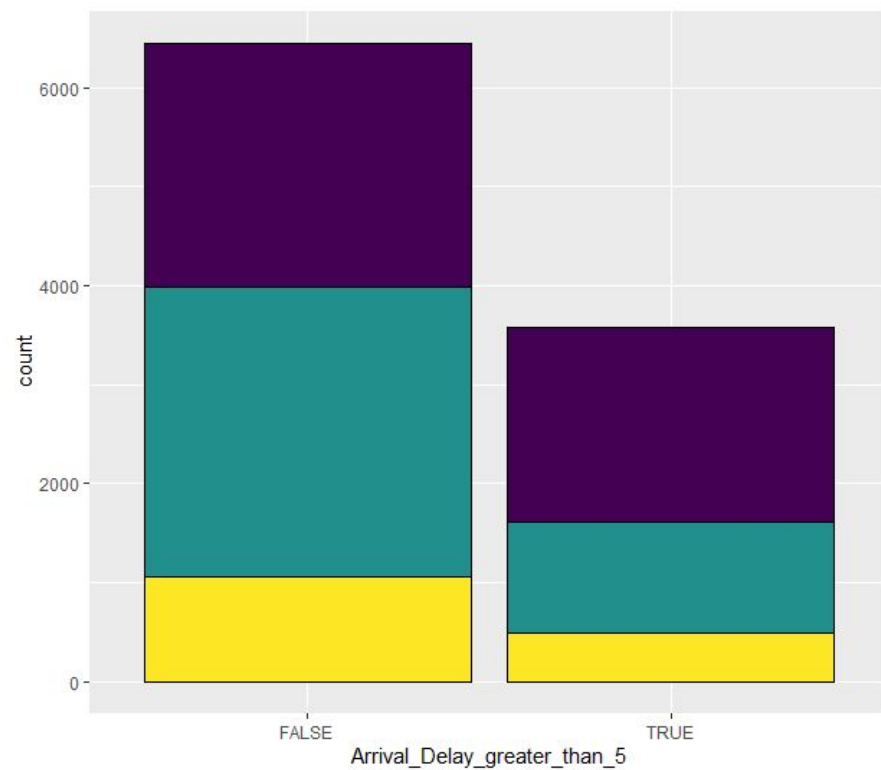# Airline Status
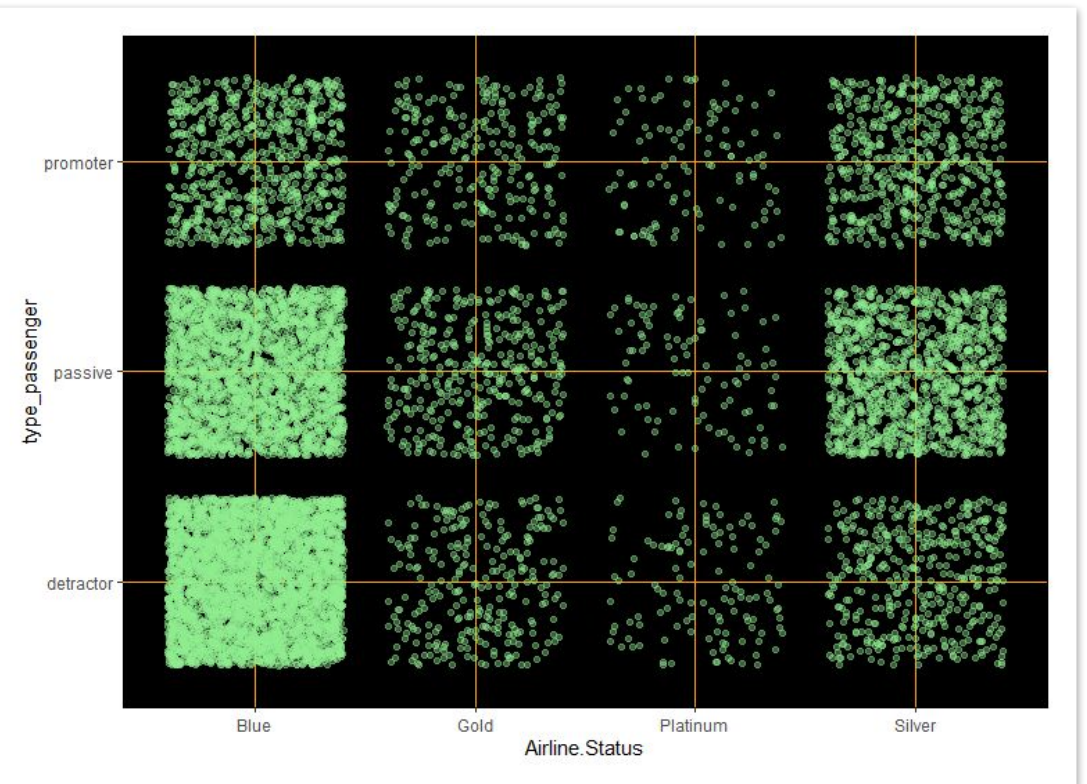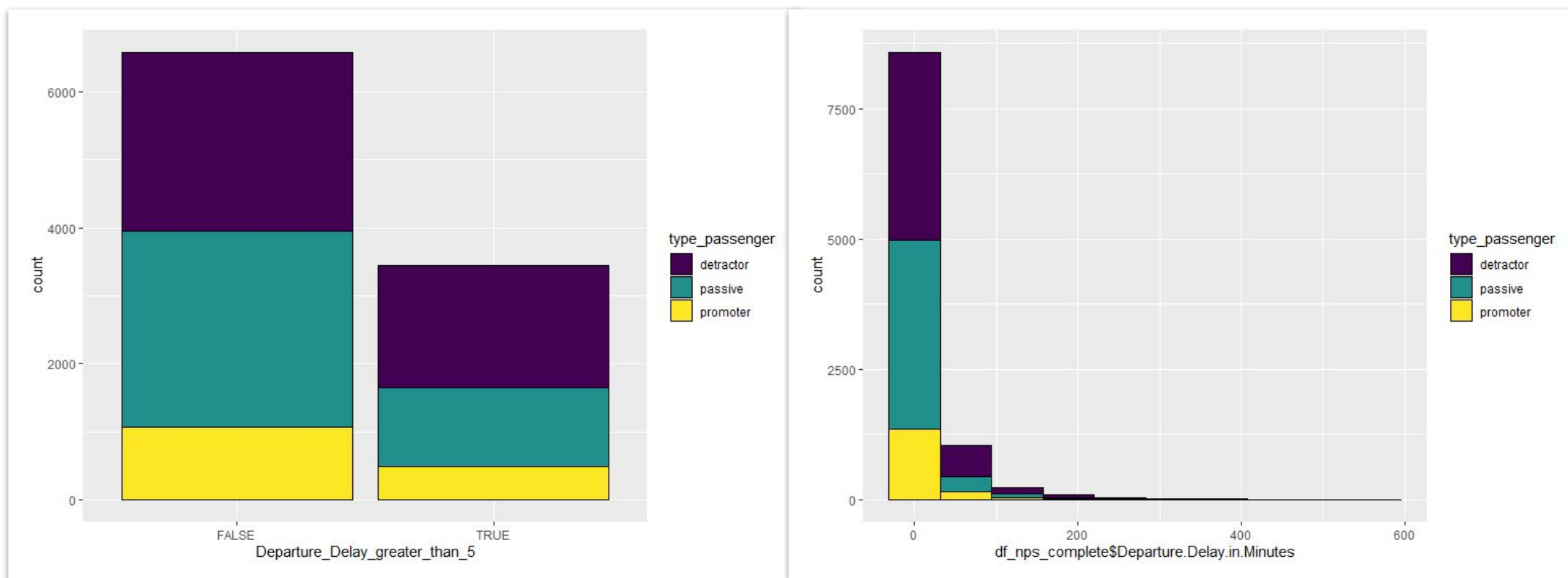
Type of Travel

Age

Partner Name

Price Sensitivity

Arrival Delay

Departure Delay

State wise analysis
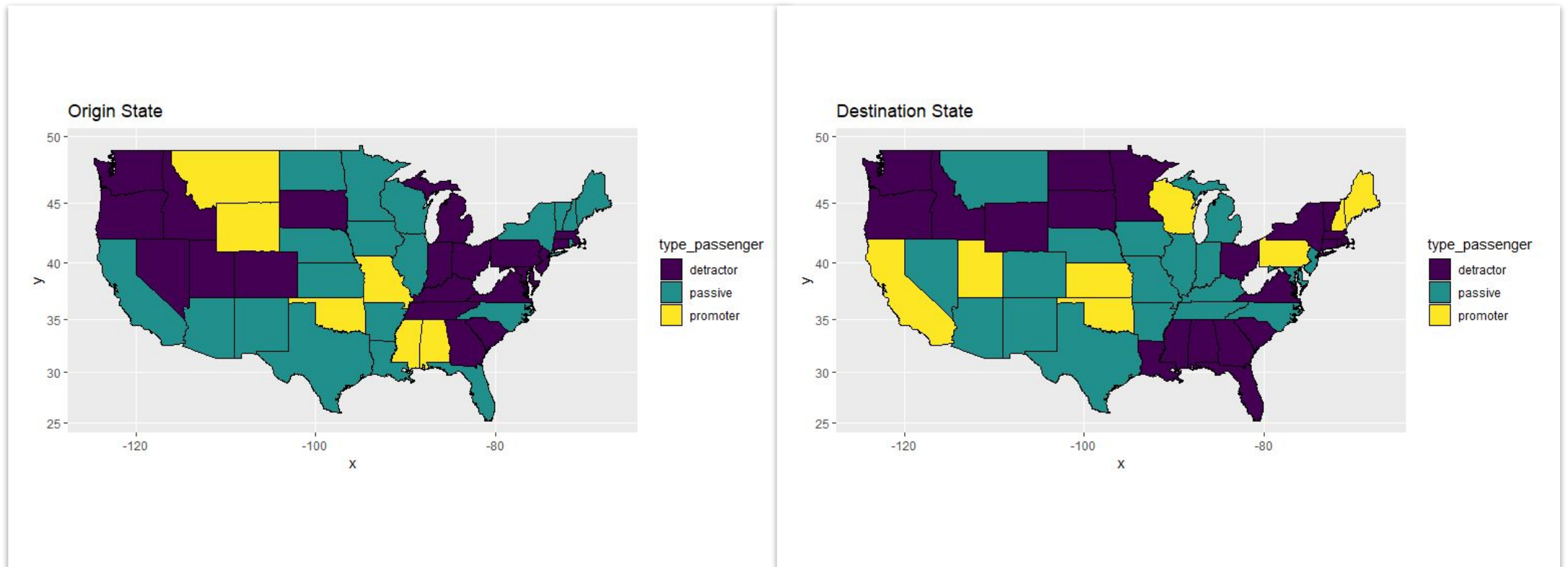
Feature Engineering:

1. Data Transformation :

●Removed variables which have more than multiple levels Date, Latitude/Longitudes, to avoid model ambiguity

●Bucketed the features as per their distribution to obtain clear insights Total variables 16 at the End of this Step

2. Feature Importance Generation to find the most significant Features

In order to get optimum Feature Importance, we perform incremental modelling technique

●Logistic Regression

●Random Forest Classifier

●XGBoost Classifier

●We compare the Relative Importance in each of the models to figure out the best combination of variables

Logistic Regression

The model performed with 75.87% Accuracy when validated using a test data set

Generalization Model Performance : 0.8451

Dominant features after tune grid (Lambda: 0.00072, Alpha: 1):

- Partner.Name
- Age
- Gender
- Airline.Status
- Price.Sensitivity
- Loyalty
- Type.of.Travel
- Departure.Delay.in.Minutes
- Arrival.Delay.in.Minutes

Random Forest Classifier

- **Performed hyper parameter tuning to avoid overfitting (Trees : 200)**

- **The model performed with 76.10% Accuracy when validated using a test data set**

- **Genaralization Model Performance AUC Score : 0.8471**

Feature Importance
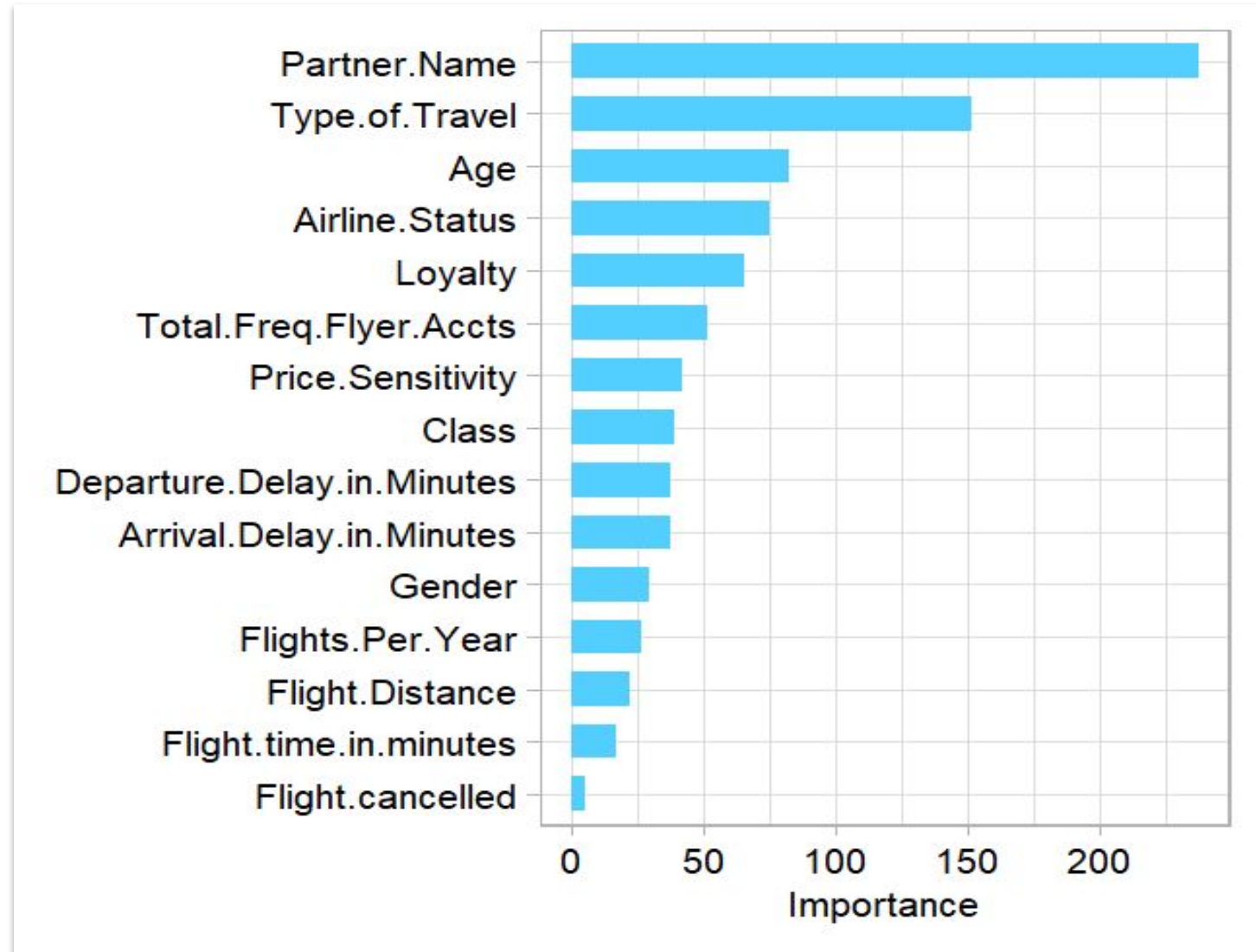
XGBoost Classifier

- **The model performed with 76% Accuracy when validated using a test data set**
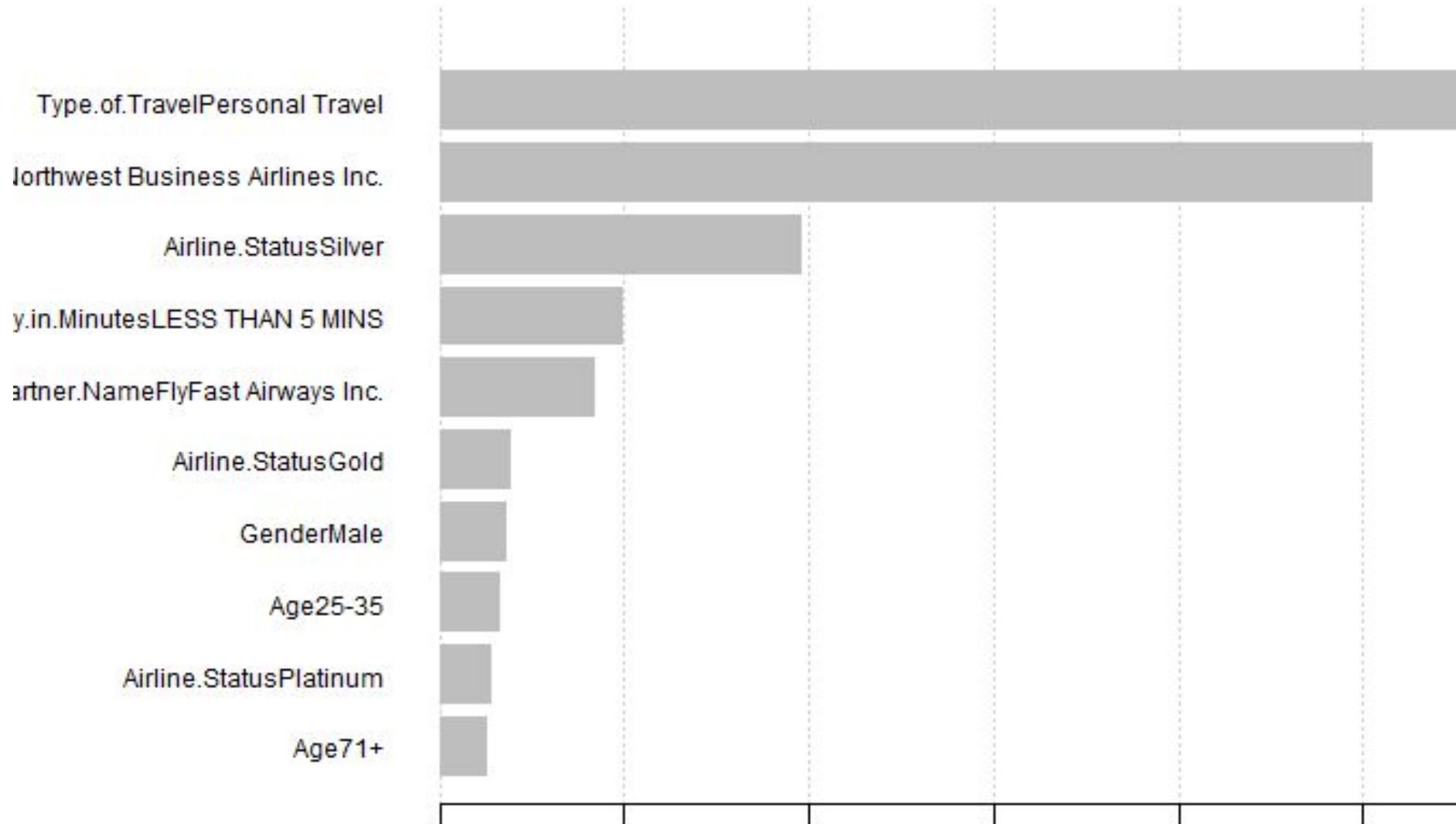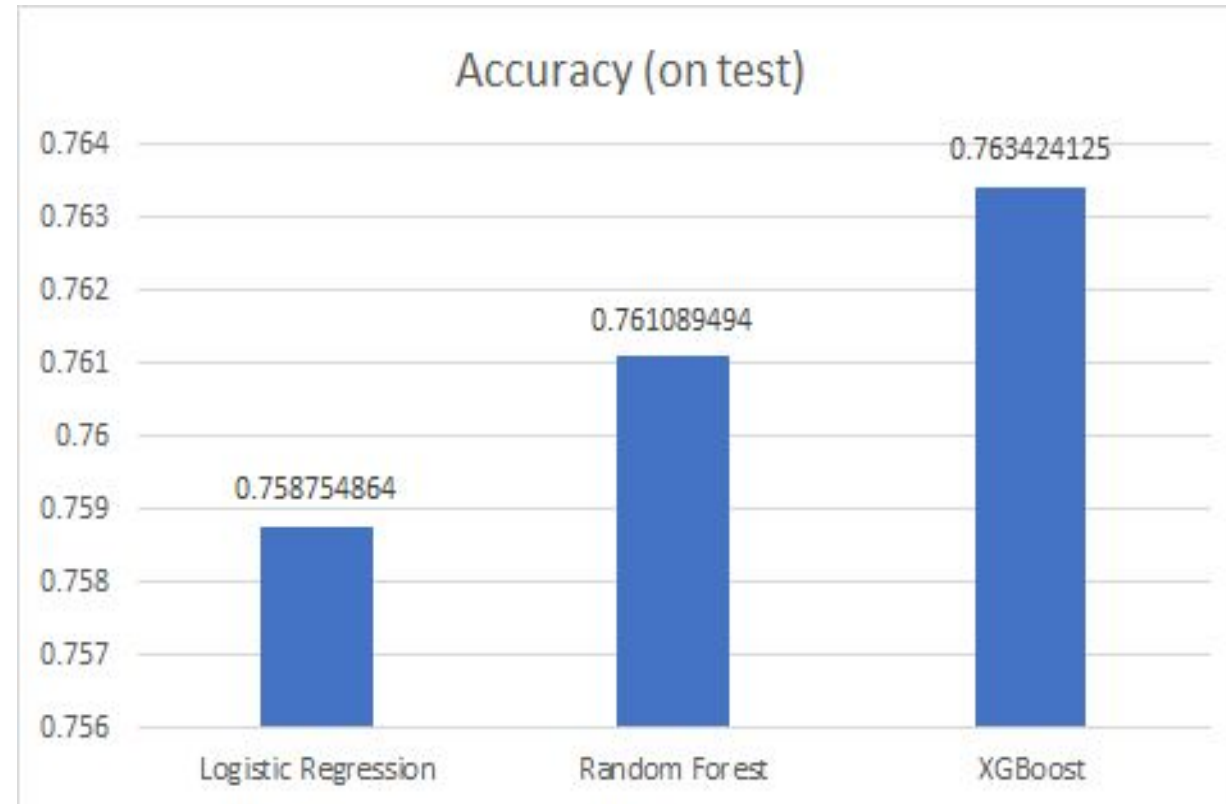- **Generalization Model Performance AUC Score: 0.8398**

Performance Comparison

We also combine the feature importance of Random Forest and XGboost with best AUC and Accuracy respectively to gain insight improving NPS Scores

Feature interpretation from logistic regression with best AUC score

AUC Scores (on test)

- Logistic Regression: 0.8451386
- Random Forest: 0.8471364
- XGBoost: 0.8398

Accuracy (on test)

- Logistic Regression: 0.758754864
- Random Forest: 0.761089494
- XGBoost: 0.763424125
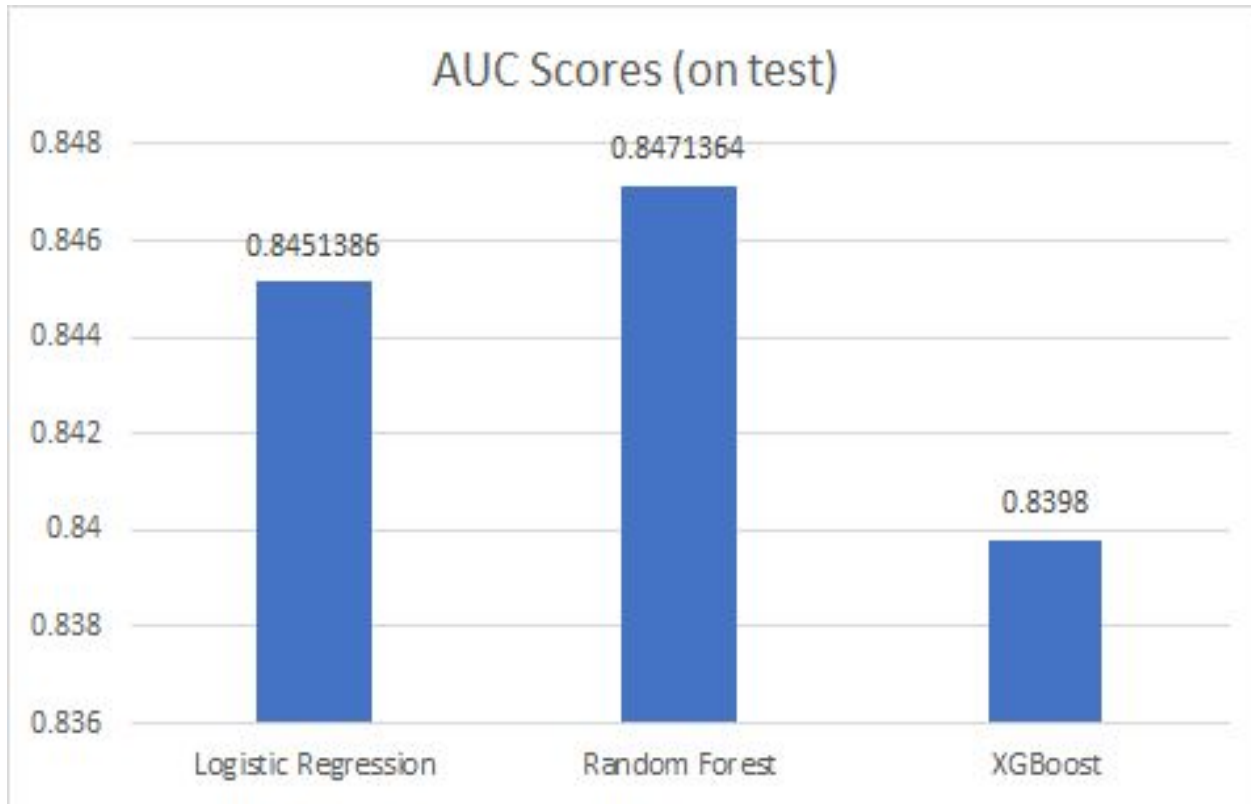
We select top 10 variables with high feature importance across all the 3 classifiers:

- Partner Name
- Type of travel
- Age
- Airline Status
- Loyalty
- Total Freq. Flyer. Accts
- Price Sensitivity
- Class
- Departure Delay in Minutes
- Arrival Delay in Minutes

Apriori Ruleset

Detracter	
Characteristics	**Values**
Type of Travel	Personal Travel
Age	71 years +
	57 to 71 years
Airline Status	Blue
Loyalty	-0.75 to -0.45 (The users generally don't travel by this airline)
Flights per Year	26+

| Promoter	
Characteristics	**Values**
Type of Travel	Business Travel
Departure Delay	Less than 5 min
Arrival Delay	Less than 5 min
Airline Status	Silver
Price Sensitivity	1 (from a range of 0 to 5,i.e. not much sensitive)

Results and Inferences

(Based on Logistic, XGBoost Apriori and K-prototype Clustering)
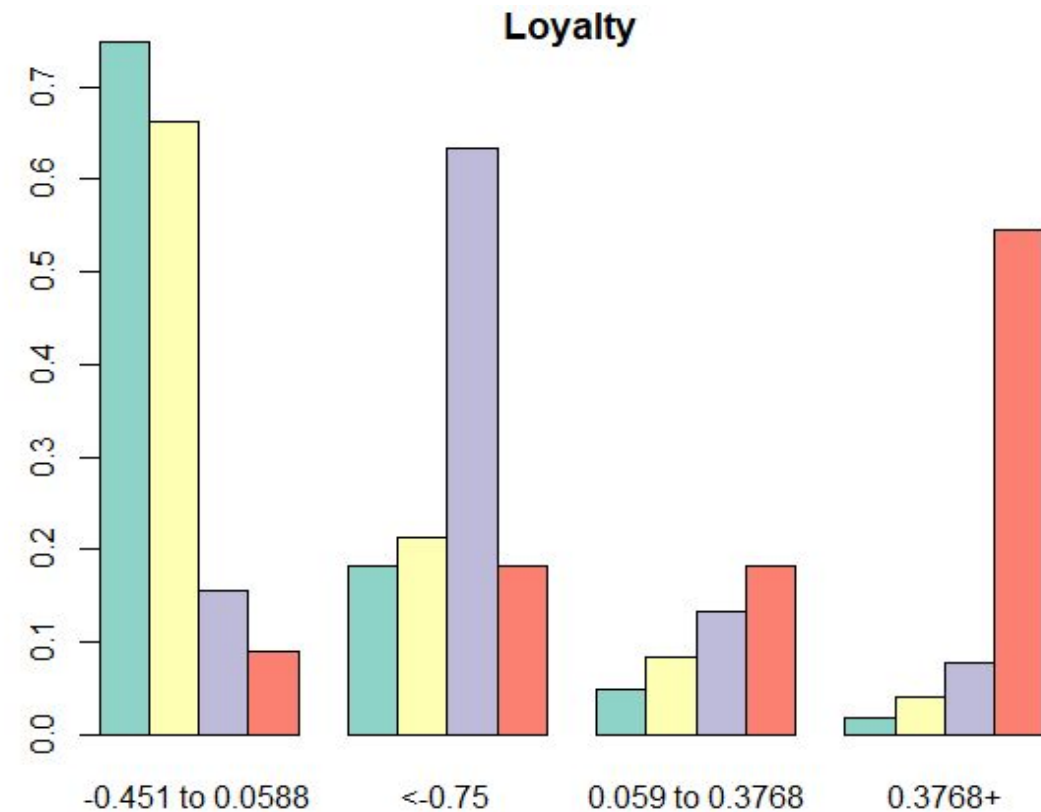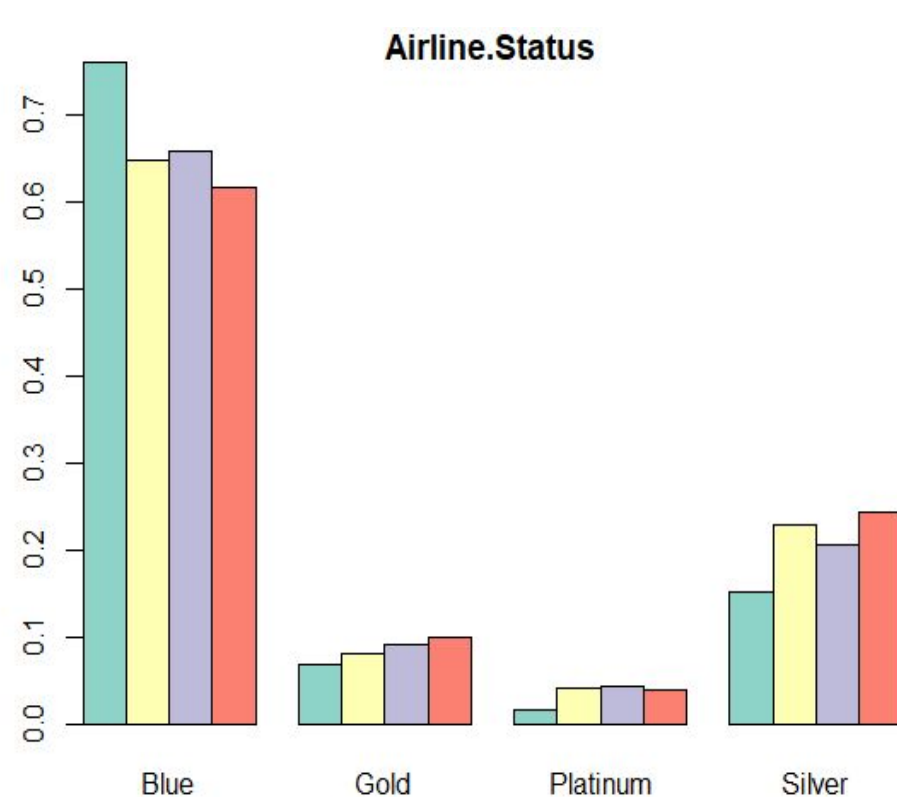
Customer whose :

- **Loyalty score is Less than –0.75**, is 1.4% less likely to be a detractor.
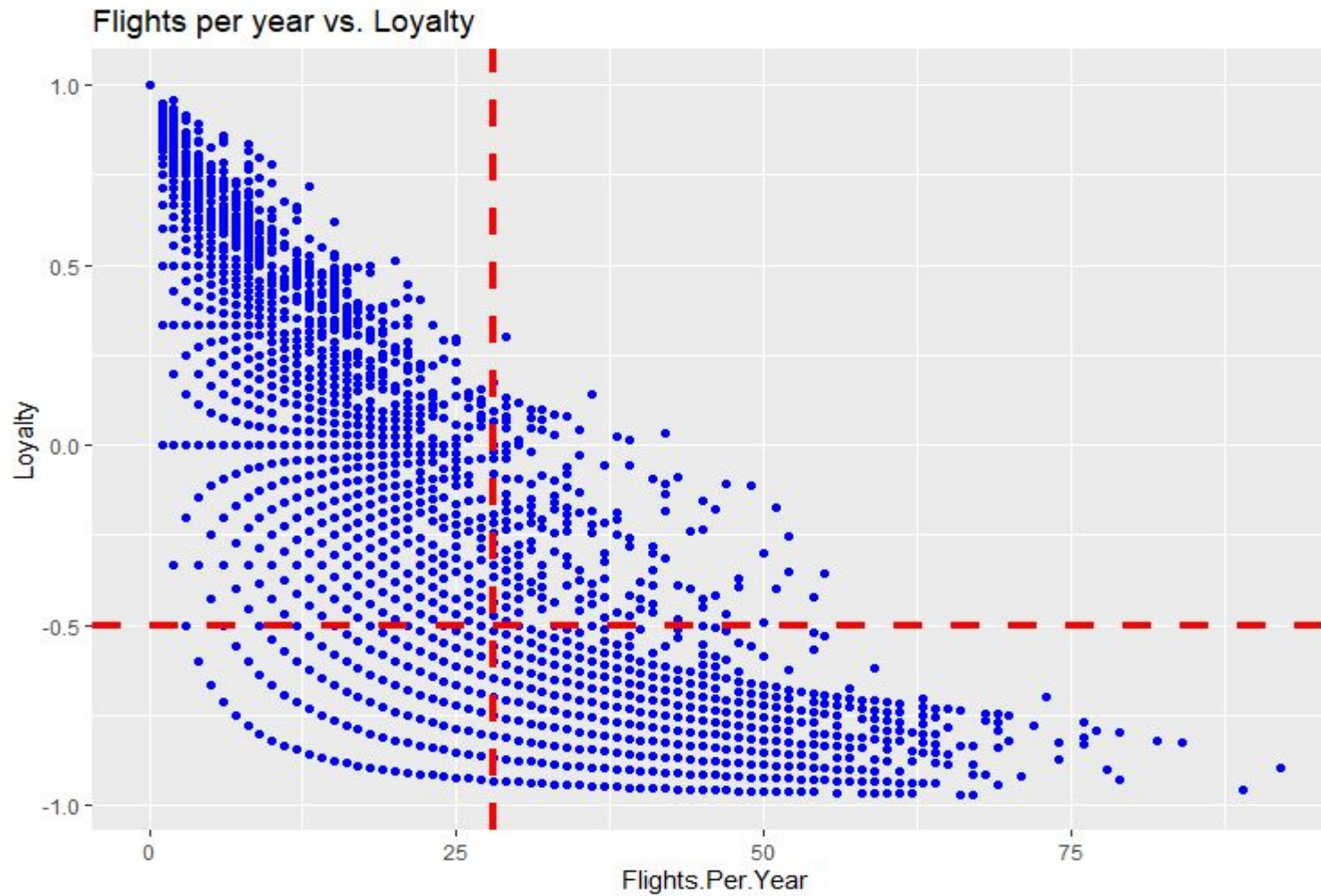- **Airline status is Silver**, is **12.4%** less likely to be a detractor
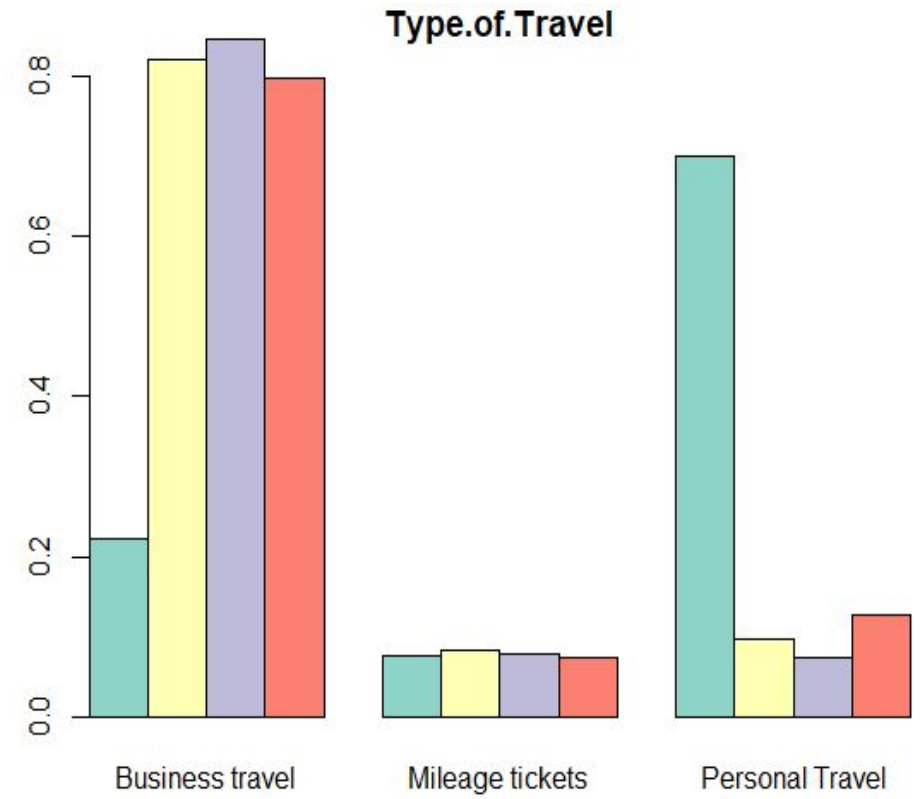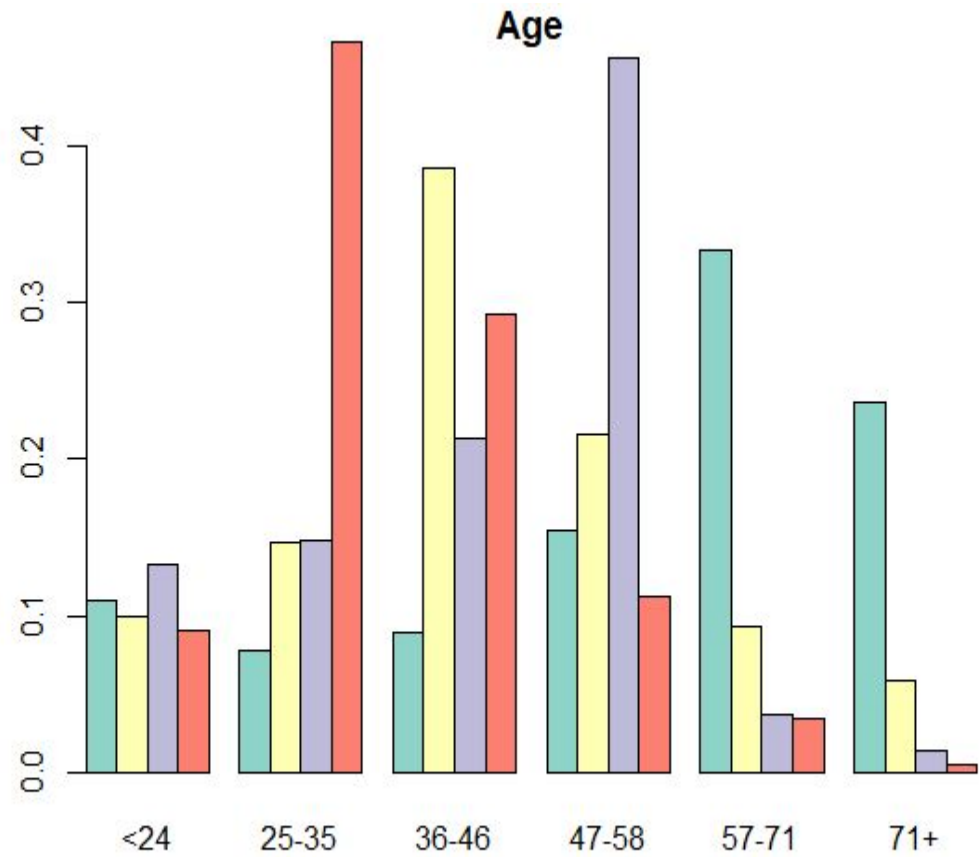
Customers :

- with Personal Travel is **21 % more likely** to be a detractor
- whose **age is greater than 71 years**, we predict that the probability of being a detractor **increases** by **2.8%**

Recommendations

Areas of Improvement: Frequent flyers


Flights per year vs. Loyalty

Areas of Improvement: Frequent flyers

- Frequent flyers with 28+ flights per year don't prefer this flight favourably.
  - We consider 25% to 30% of those flights are with Southeast, i.e. 6-9+ flights per year with Southeast
  - To make the customers return, they must feel that they will receive the most optimum offer, if they continue to travel for more than 6-7 flights with Southeast per year.
  - We advice to increase the Mileage points to offer some sort of discount and offer complimentary food to regular flyers

# Areas of Improvement (Personal travellers)

- The most dissatisfied customers of the flight are with the intention of Personal Travel.
- Our advice:
    - Increase efforts to target families,
      individual flyers
   - Advertise family plans, especially
      holidays
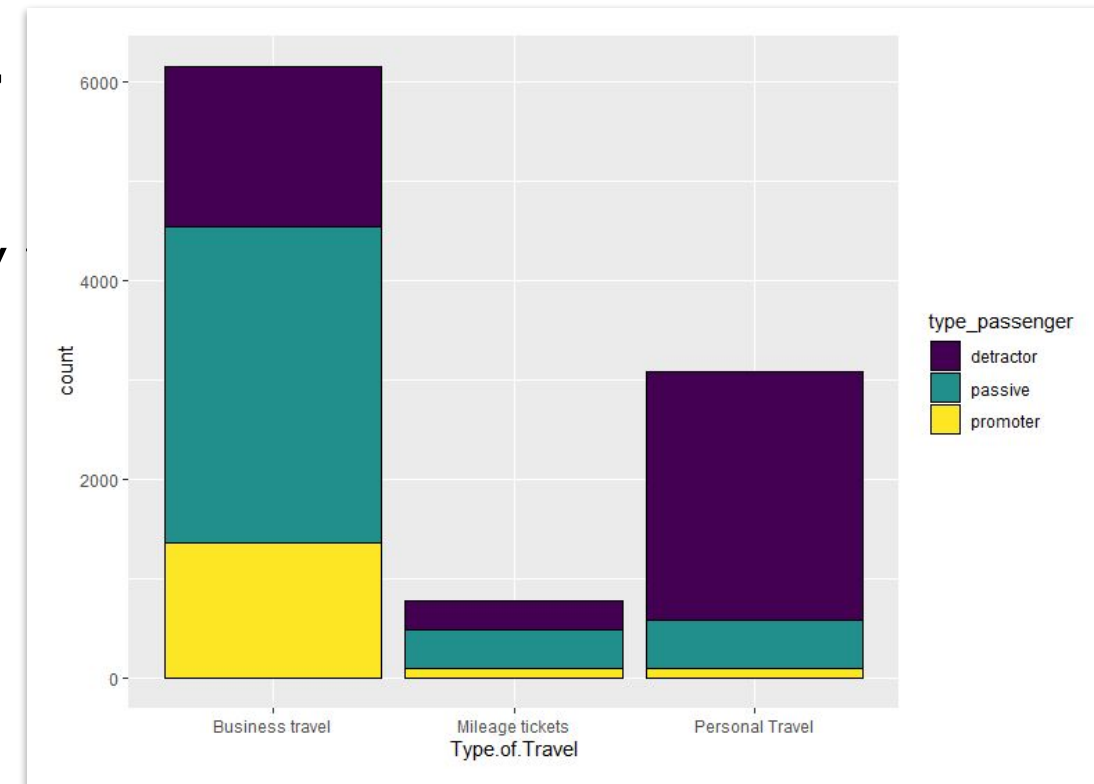
Areas of Improvement (Silver/Blue airline Status)

- Most of the promoters have Silver airline status. And even Gold and Platinum have less proportion of promoters than Silver.
- Our Advice:
  - Increase the number of Silver Status flights
  - At the same time, improve the services of the Blue status flights, as most of the passengers travel by those
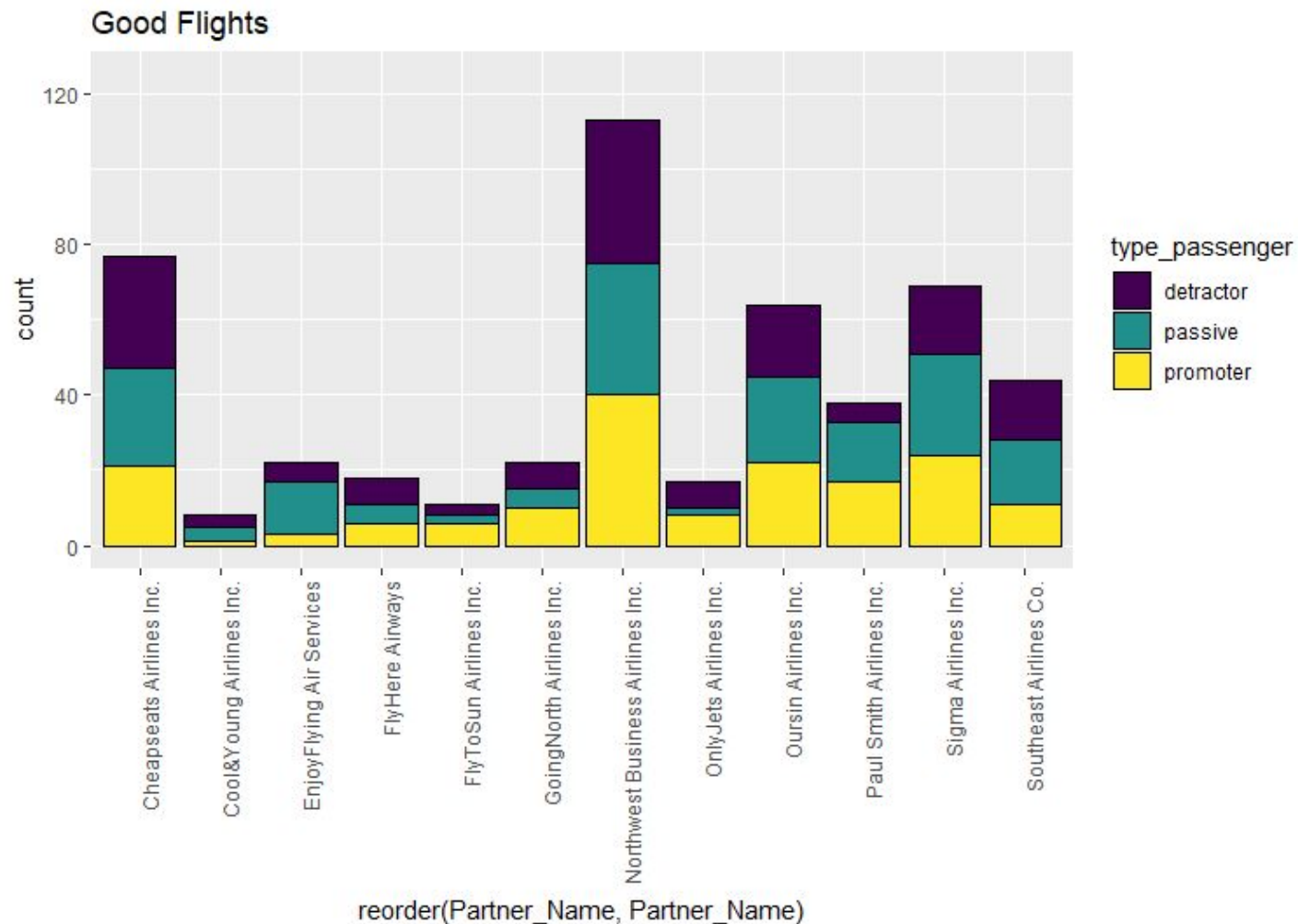
Area of Improvement (Partner Airlines)
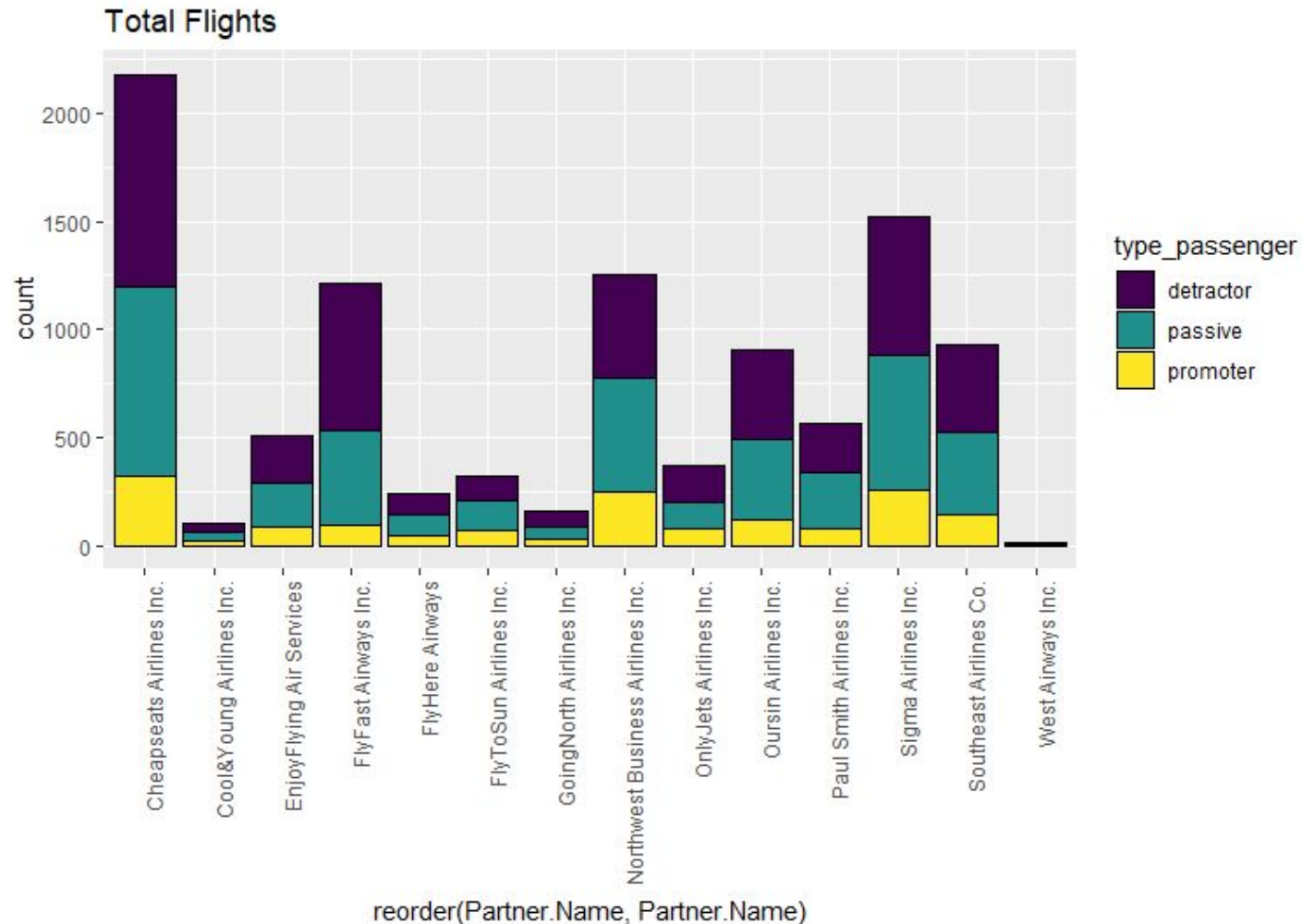
Areas of Improvement(Partner Airlines)
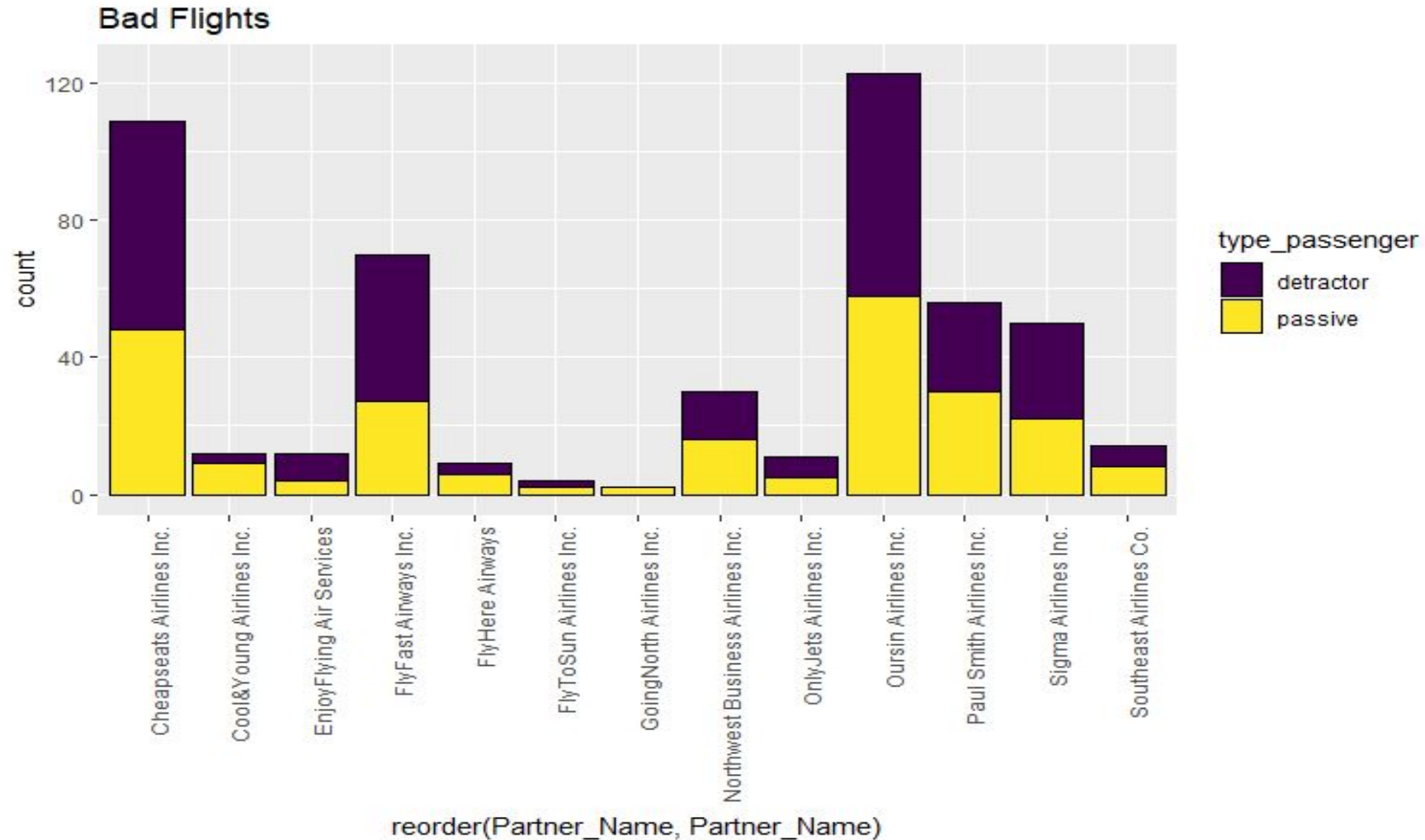
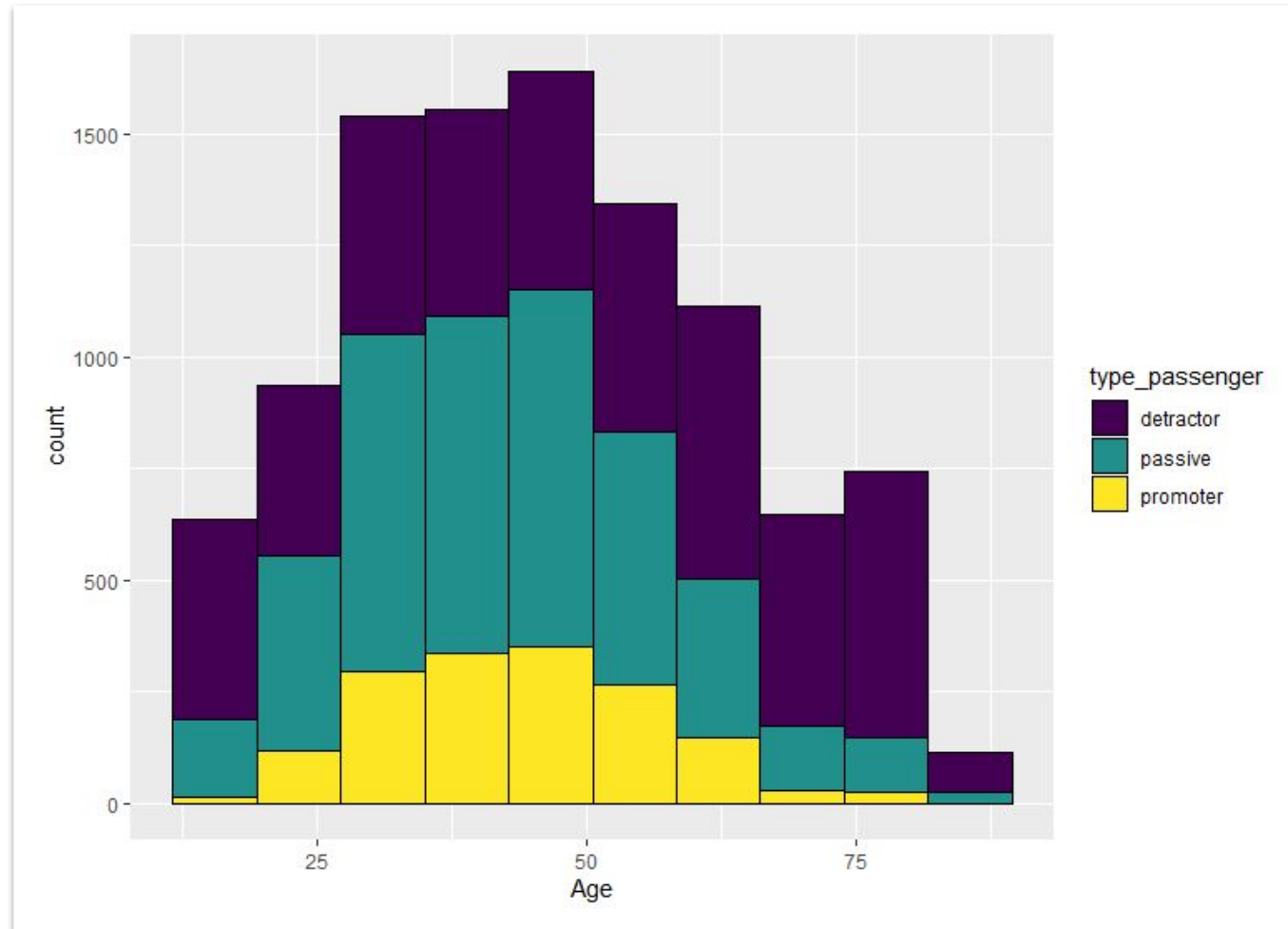# Area of Improvement (Partner Airlines)

Put these Partner airlines under Probation or Remove them

- Cheapseats Airline Inc.
- FlyFast Airways Inc.
- Oursin Airlines Inc.

Audit and ask Northwest Business Airlines to make their services more consistent in services for all their flights, as they provide both the most number of promoters as well as detractors.

Provide better services to senior citizens

Questions ?