Predicting Airline Passenger Satisfaction

By: Jieyi Chen, Ben Matz, Hanzhe Zhou

I. Introduction and Project Purpose

Customer satisfaction surveys are an important tool for airlines to improve their inflight experience and target customers. However, sometimes it is difficult for airlines to know which specific factors of a passenger's inflight experience lead to a passenger's satisfaction with the particular airline and which may deter the passenger from traveling with the airline again in the future. Therefore, the goal of this project is to predict whether a passenger will be satisfied or dissatisfied with their travel experience with a particular airline (name not given) based on various factors of the customer's inflight experience, the customer's purpose of travel, and characteristics of the customer.

Supervised learning methods will be used in order to predict passenger satisfaction. Models that will be used include, logistic regression, a decision tree model, a random forest model, a boosting model, and a neural networks model. Variable importance for each model will be used to determine which factors of a passenger's experience are the most important determinants of satisfaction level. Additionally, clustering will be used to group passengers into segments. Such clusters have the potential to provide useful information to the airline about which group of customers are likely to have a satisfactory travel experience and which will have a negative travel experience. The airline could use information from these clusters to examine the strength and deficiencies in its inflight service for each market segment. For example if a particular cluster is dissatisfied with inflight entertainment, then the airline could target improving inflight entertainment for this market segment to improve customer satisfaction.

II. Description of Data Set

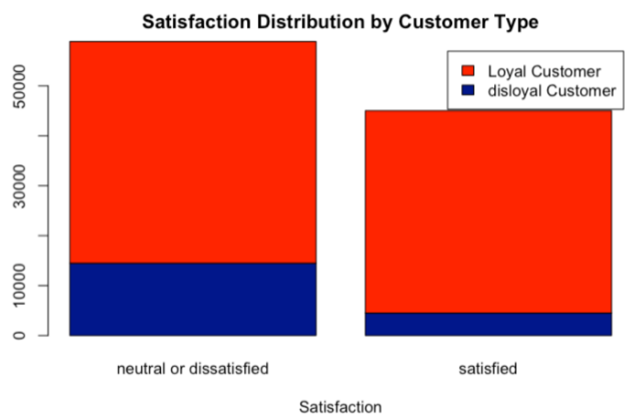
A dataset for flight reviews from an anonymous airline was obtained from Kaggle. A separate training set and test set were both obtained. The dataset includes twenty-two different variables along with the satisfaction level of the customer. Explanatory variables in the dataset include characteristics of the passenger such as age, gender, whether the customer is a loyal customer of the airline, and what class the customer's seat was in. Additionally, the dataset includes characteristics of the customer's purpose for traveling such as whether the trip was for leisure or business purposes. Other characteristics included in the dataset pertained to aspects of the customer's inflight experience including whether or not WiFi was included, the quality of entertainment, seat comfort, legroom, the quality of food and drink, and the amount of time the flight was delayed. The survey results were divided between two satisfaction levels–satisfied and neutral or dissatisfied. Ultimately, based on the various characteristics included in the dataset, models were developed to predict whether or not customers would be satisfied with their flight experience with this particular airline.

The dataset included both categorical variables and continuous variables. The categorical variables for the characteristics of the passenger's inflight experience were each assigned an integer value from 1 to 5. This value reflected the passenger's satisfaction level with that particular characteristic. Some of these categorical variables had values of 0 which were treated as missing values. Additionally, missing data from the customer satisfaction data does not provide valuable information to the airline. Other variables, such as delay time, flight distance, and the age of the passenger were treated as continuous variables. Lastly, variables such as the overall satisfaction of the customer with the airline, the purpose of travel, and the type of customer were treated as binary variables.

III. Preliminary Data Analysis

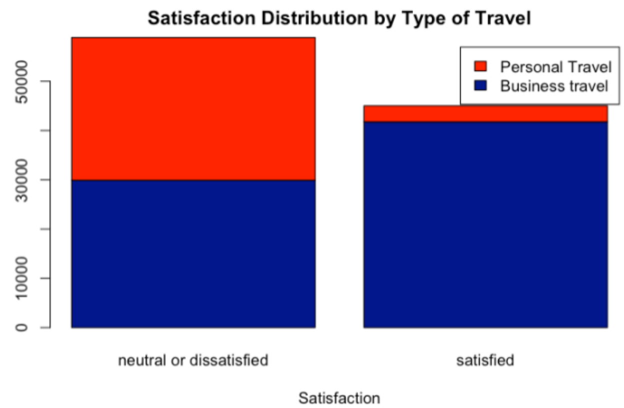
The training dataset was first analyzed to see which types of customers were represented in the dataset. The data had more passengers that were neutral or dissatisfied with their flight experience with the airline compared to passengers who were satisfied with the airline. Additionally, most of the customers in the training data set were loyal customers of the airline. This is logical as it is plausible that loyal customers respond to customer satisfaction surveys from the airline at higher rates than disloyal customers. As expected, more disloyal customers were dissatisfied with their flight experience with the airline.

Figure 1: Customer satisfaction by customer type.



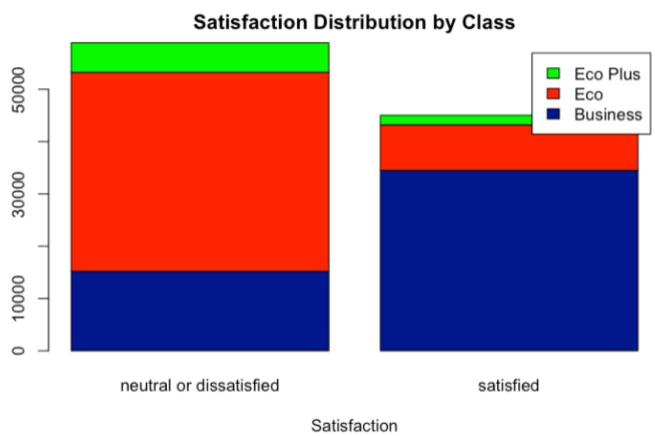
Moreover, within the training data set, most of the customers who were satisfied with the airline traveled for business related reasons. By contrast, the customers who were dissatisfied or neutral with their flight experience with the airlines were split between passengers who traveled for business purposes and those who traveled for personal reasons. It is likely that those who traveled for business purposes traveled in business class and had a better experience, which might explain why this group makes up the majority of satisfied customers.

Figure 2: Passenger satisfaction by purpose of travel.



Unsurprisingly, the majority of satisfied customers traveled in first class, while the majority of dissatisfied or neutral passengers traveled in economy class. Therefore, the enhanced characteristics of first class–premium food, extra seat space, and extra care– likely led to higher satisfaction rates among first class passengers. Economy plus passengers make up the minority of both satisfaction groups. It is possible that economy plus seating is not offered on all of the airlines' flights or there are the fewest number of available seats in this class. From Figure 3, it is evident that the characteristics of seat class are important determinants in a passenger's satisfaction level.

Figure 3: Passenger satisfaction by class.



Since several of the values for the categorical variables for the customer satisfaction levels were missing, these observations were removed from the training and test data sets. Additionally, observations that had missing values for the arrival time delay and departure time delay were also removed. After filtering the data on these conditions, most of the data in the training and test sets was still retained. As a result, imputation was not used for data preprocessing. Sampling weighting was not applied since the number of satisfied customers was roughly equal to the number of dissatisfied passengers.

IV. Predictive Modeling

A. Variable Selection

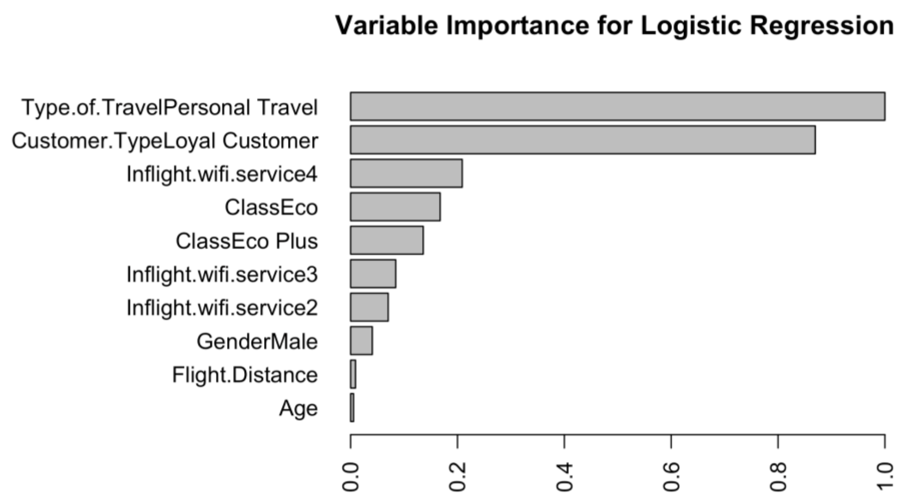
In order to determine the most important variables for predicting passenger satisfaction of the airline, lasso regression was used for variable selection. Ultimately, twenty-one explanatory variables were deemed important for the model. The only variable that was dropped was departure delay time, because arrival delay time captures this effect. Arrival delay time and departure delay time are highly correlated and have a correlation of 0.965. Selected variables include class, type of travel, and the passengers' satisfaction levels with various aspects of the inflight experience. Some of these characteristics included inflight entertainment, seat comfort, legroom, and arrival delay time. Several predictive models were developed with these selected variables.

Explanatory variables: Gender, customer type (loyal or disloyal), age, type of travel (personal or leisure), class (economy, economy plus, business), flight distance, inflight WiFi service, departure/arrival time convenience, ease of online booking, food and drink, online boarding, seat comfort, inflight entertainment, on board service, leg room service, baggage handling, checkin service, inflight service, cleanliness, arrival delay time

B. Variable Importance for Models

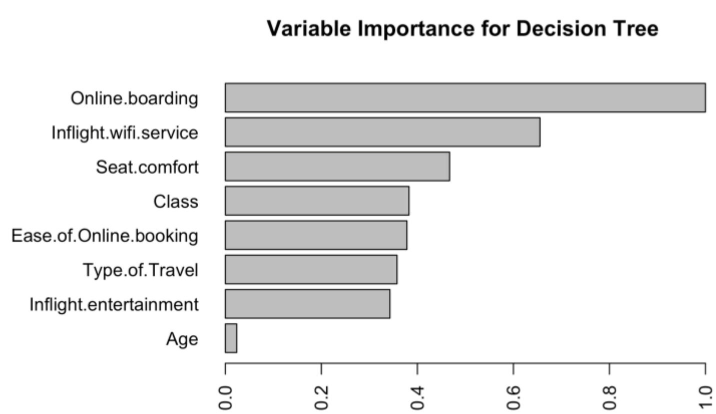
Several models were developed to predict passenger satisfaction with the airline including a logistic regression, a decision tree model, a random forest model, a boosting model, and a neural network model. A k-nearest neighbors model was not developed since this model does not provide feature importance plots and because of the different scale for our explanatory variables. The optimal parameters for the decision tree and boosting models were determined by finding the combination of parameters that yield the lowest cross validation error. For the random forest model, the combination of variables which yielded the lowest out of bag error was selected. By contrast, for the neural networks model, the optimal combination of parameters was chosen based on which model yielded the lowest log loss. Additionally, feature importance plots were generated for each model to determine the most important features of passenger satisfaction and provide useful information to the airline and will also be covered later in the clustering analysis section.

Figure 4: Variable importance for logistic regression.



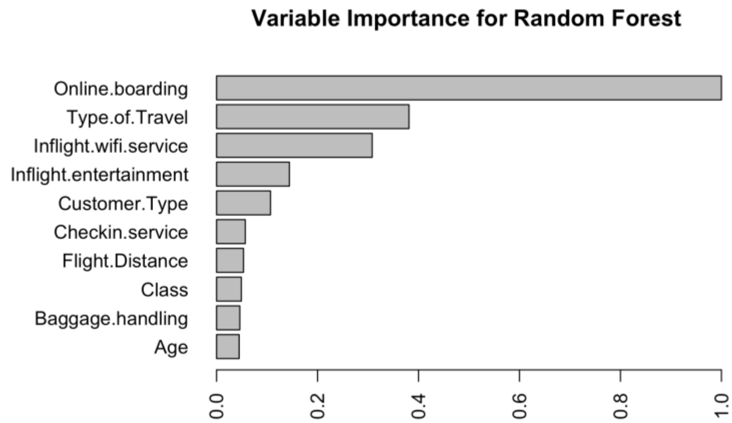
The first model that was developed was a logistic regression. From this model, the most important variables were type of travel, type of customer (loyal or unloyal customer), inflight WiFi satisfaction level, and the class the passenger flew in. Surprisingly, seat comfort, onboard service satisfaction level, and satisfaction with inflight entertainment were deemed the least important variables in predicting a particular customer's overall satisfaction with the airline.

Figure 5: Variable importance for decision tree.



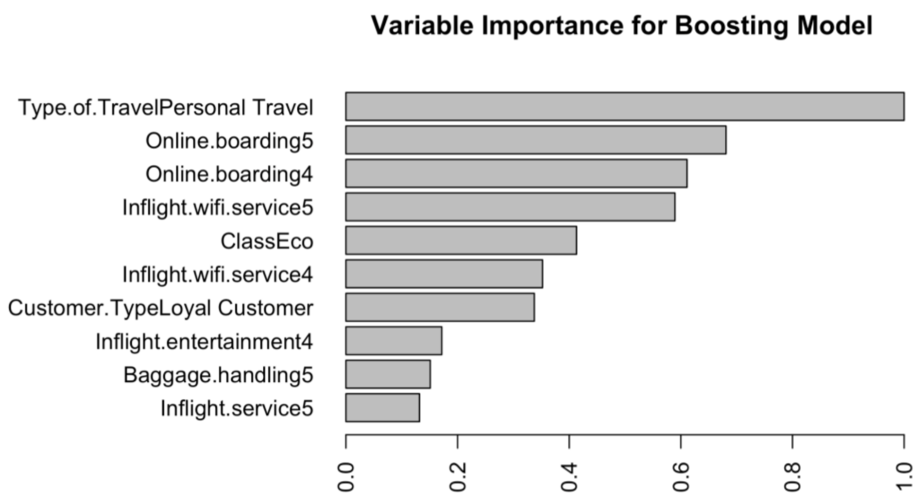
Next, a decision tree model was fitted. A hypergrid was used to find the optimal parameters including the tree size and the number of observations to split on. As shown in Figure 5, based on the optimal decision tree model, the most important variables in determining passenger satisfaction are passenger satisfaction with online boarding pass, satisfaction level with inflight WiFi service, seat comfort, the ease of online booking, and the type of travel. By contrast to the logistic regression model, passenger characteristics including type of travel and customer type were deemed less important.

Figure 6: Variable importance for random forest model.



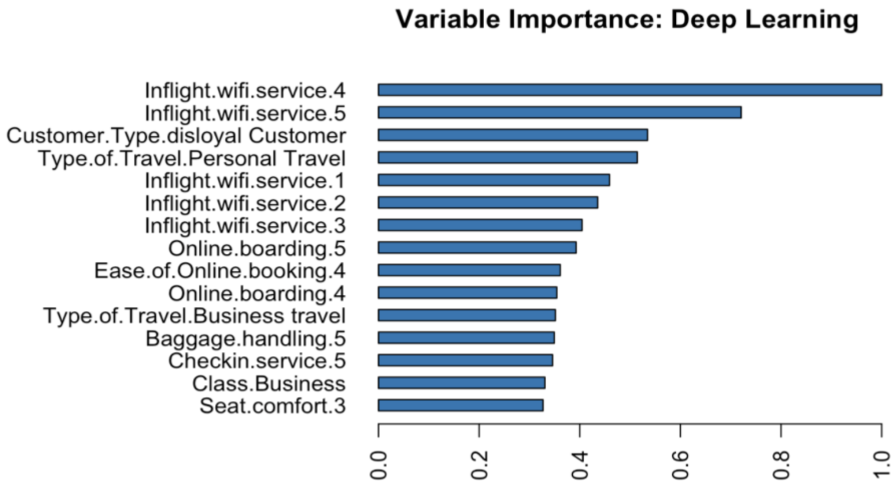
Subsequently, a random forest model was fitted on the data. As shown in Figure 6, the variable importance for the random forest model was similar to the decision tree model. However, type of traveler and passenger satisfaction with inflight entertainment were deemed to be more important predictors while seat comfort and ease of online booking were deemed to be less important predictors. The difference can be attributed to the fact that random forest models reduce variance by fitting multiple trees and using majority vote for classification. Overall, the variable importance between the optimal decision tree model and optimal random forest model were similar as passenger satisfaction with the online boarding pass was the most important predictor. Since random forest models are derived from decision tree models the similarity in variable importance is not surprising.

Figure 7: Variable importance for boosting model.



As shown in Figure 7, the variable importance for the boosting model is quite similar to the decision tree and random forest models. Passenger satisfaction with the availability of an online boarding pass, type of travel, and inflight WiFi were deemed the most important variables. However, customer type–whether the customer is a loyal customer of the airline or not–had greater importance in the boosting model than in the random forest and decision tree models. Surprisingly, the amount of the time the flight was delayed did not have any significant predictive power and was not deemed one of the top 10 most important variables in predicting customer satisfaction in any of the models.

Figure 8: Variable importance for neural networks.



The last model that was fitted on the data was a neural networks model. As displayed in Figure 8, by contrast to the previous models, passenger satisfaction with the online boarding pass was not the most important predictor. However, similar to most of the other models, passenger satisfaction with inflight WiFi, customer type, and type of travel were deemed important variables in predicting passenger satisfaction.

Broadly, across all models, the most important variables in predicting customer satisfaction were passenger satisfaction with the online boarding pass, satisfaction with inflight WiFi service, type of travel, customer type, and class. Therefore, satisfaction levels are based on the characteristics of the customers such as whether they flew for business purposes or leisure and whether they flew in first class or economy. Passenger satisfaction levels with various characteristics of the inflight experience had less predictive power overall.

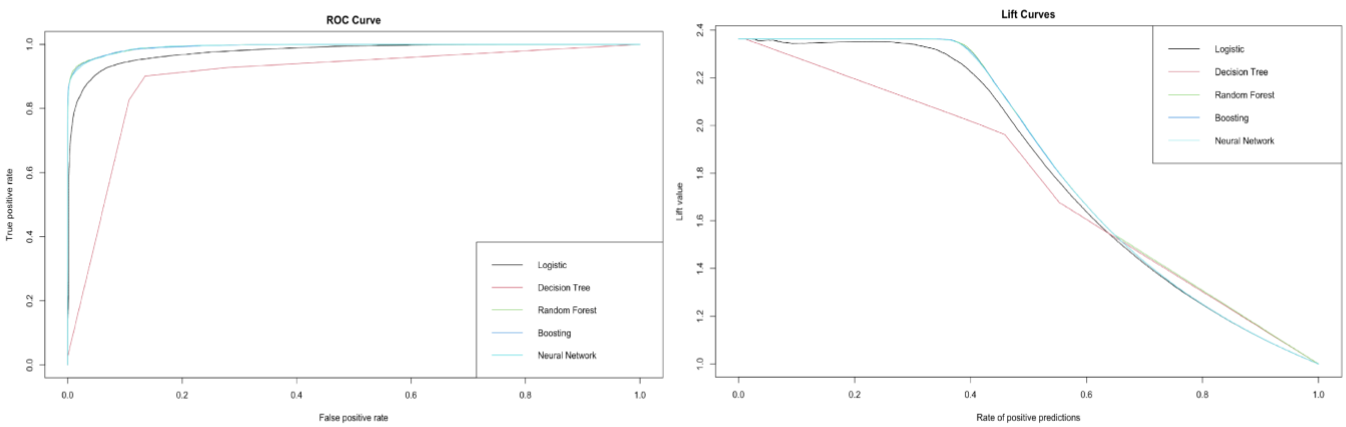
C. Choosing the Optimal Model

Figure 9: Metrics for each of the different models.

Model	**Deviance**	**AUC**	**Accuracy**
Logistic	17609	0.977	0.931
Decision Tree	33331	0.896	0.88
Random Forest	9163	0.993	0.961
Boosting	9577	0.993	0.959
Neural Network	9739	0.993	0.958

For each of the five different models fitted on the data, deviance, AUC, and accuracy –were obtained. The original training data set was first split into a training data set and a validation set. Each of these metrics is based on how the fitted model–with the optimal parameters–performed on the validation set. As reflected in Figure 9, the random forest, boosting, and neural network models performed the best in terms of AUC. However, the random forest model performed best in terms of deviance and accuracy.

Figure 10: ROC curves (left) and lift curves (right) for each of the five models fitted on the data.



Additionally, the ROC and lift curves were obtained for each of the five models. The random forest, boosting and neural network models performed the best in terms of the ROC and lift curves. By contrast, the decision tree model performed the worst.

D. Fitting Model on Test Data

Figure 11: Confusion Matrix for Random Forest Model

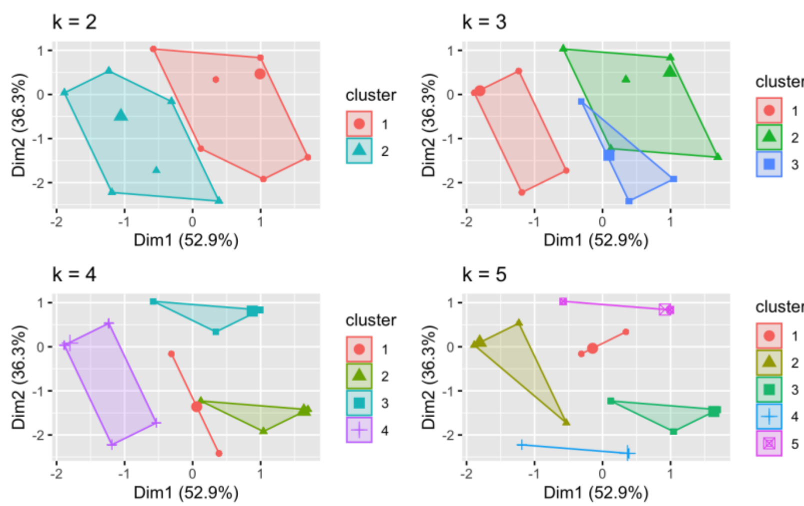
	Reference	
Prediction	0	1
0	13243	627
1	310	9609

Based on the metrics discussed above, the random forest model was chosen for predicting customer satisfaction with the test data set. We used the random forest model with the optimal tuning parameters to predict satisfaction in the test data. The model yielded accuracy 0.9606 and AUC 0.99.

V. Clustering Analysis

Clustering on passenger characteristics was used in order to determine which aspects of the inflight experience each cluster was most satisfied with and least satisfied with. These characteristics were deemed the most important by the predictive models and are intuitive since these describe the passengers. Additionally, clustering on these variables with age and sex did not create clear clusters. As a result, we decided to only use three variables to create clusters which were customer type, type of travel, and class. The varying satisfaction levels within each cluster could provide the airline with useful information in order to improve the overall satisfaction for each customer type.

Figure 12: Clusters formed by grouping on customer type, type of travel, and the class the passenger's seat was in.



As displayed in Figure 12, dividing the data into three clusters was optimal. The characteristics of each cluster was examined further to better understand what type of passengers were captured in each cluster. A description of each cluster is shown below in Figures 13 and 14.

Figure 13: Histograms showing the passenger characteristics (seat class, type of travel, and customer type) for each cluster.

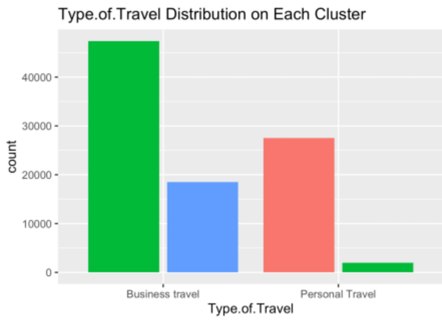
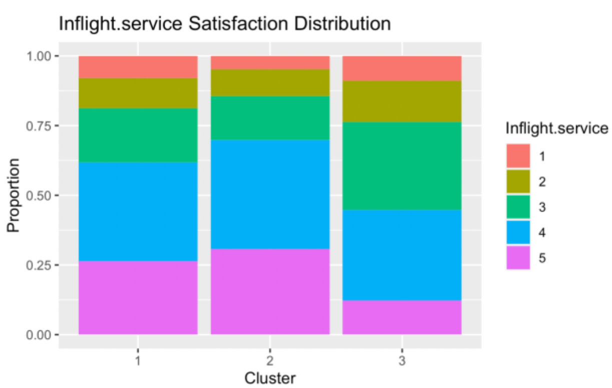
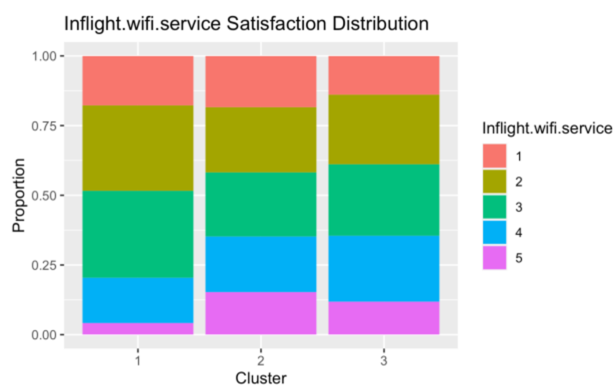
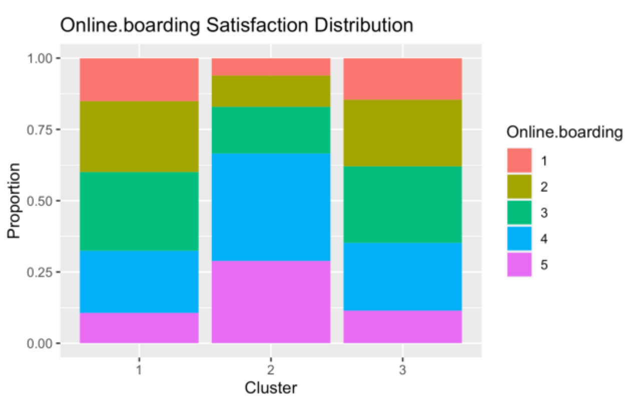


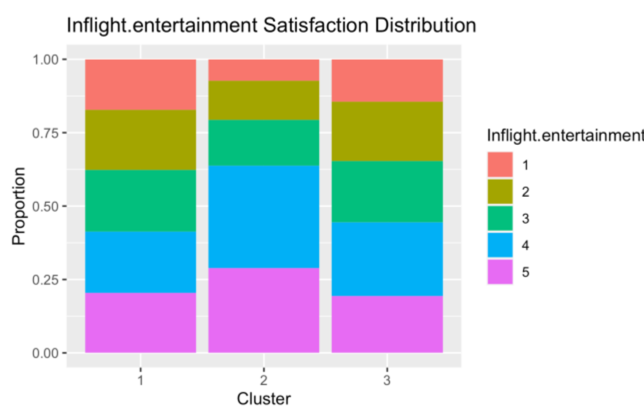
Figure 14: Summary of passenger satisfaction levels by cluster.

Cluster	Profile	High Satisfaction Aspects	Low Satisfaction Aspects
1	- economy class - personal reasons - loyal customers	- inflight service	- inflight WiFi - online boarding
2	- business class - business reasons - loyal customers	- inflight service - inflight entertainment - online boarding	- inflight WiFi
3	- economy class - business reasons	- inflight entertainment	- inflight WiFi - online boarding





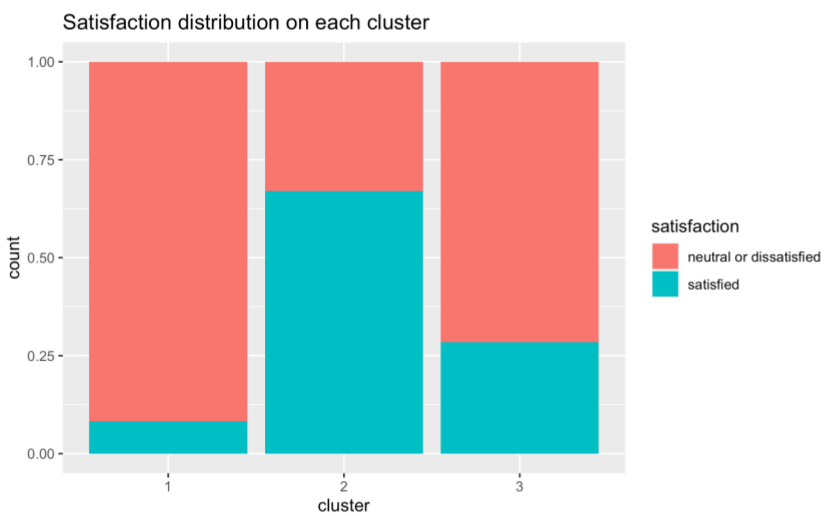




The satisfaction levels of each cluster for various aspects of the passenger's inflight experience were explored to see which aspects of the inflight experience each cluster was most satisfied with and which aspects each cluster was most dissatisfied with. The aspects of a passenger's flight experience that were deemed important by the models discussed previously were examined more closely.

In the future, the airline should improve inflight WiFi for all customer segments and online boarding for customers in clusters 1 and 3. It is plausible that cluster 2 has higher satisfaction with online boarding because the mobile experience is better for passengers who travel in business class. The airline should consider making its mobile app and online boarding experience more friendly to economy class passengers. However, the airline should consider the costs of improving its mobile app or online platform.

Figure 15: Overall satisfaction with the airline by cluster.



In terms of the overall satisfaction level with the airline, cluster 2 had a greater proportion of satisfied passengers compared to clusters 1 and 3. However, this may be due to the fact that most of the passengers in cluster 2 travel in first class and therefore have an overall better flight experience. By contrast, passengers in cluster 1 and 3 have a higher proportion of dissatisfied customers. This might be related to the fact that economy passengers are more dissatisfied with features such as online boarding and WiFi, which the airline should consider improving. Perhaps, the airline should consider providing economy passengers with limited free connection time. Similarly, since passenger satisfaction with inflight service is lower in clusters 1 and 3, the airline should consider ameliorating the inflight service that is provided to passengers who travel in economy class. This might include providing more food options or more options for inflight entertainment. It is plausible that improving these features for economy class passengers would boost the satisfaction levels of clusters 1 and 3.

VI. Analysis and Implication of Results

Across the five different models that were fitted on the data, the most important variables in determining passenger satisfaction were customer type, type of travel, class, satisfaction with inflight WiFi, satisfaction with inflight entertainment, and satisfaction with online boarding (i.e. online boarding pass through a mobile app).

Factors which were less influential in determining customer satisfaction included the quality of food and drink, arrival time delay, and baggage handling. In order to broadly boost customer satisfaction, the airline should focus on improving inlight WiFi service, online boarding, and inflight entertainment since these were important in predicting overall customer satisfaction. Therefore, the models that were developed indicate to the airline which aspects of its business model it should focus on improving.

Clustering was used in order to determine how the most important features that drive customer satisfaction differed within each segment. Ultimately clustering revealed a divide in passenger satisfaction based on various characteristics of the passengers including which class they flew in, purpose of travel, and whether or not they were loyal customers of the airline. Those who flew in business class had higher levels of overall satisfaction with the airline while passengers who flew in economy had lower levels of overall satisfaction. However, business class passengers had low satisfaction with inflight WiFi. Similarly, passengers who flew in economy were also dissatisfied with WiFi service, but also online boarding. Perhaps the mobile app or online platform used for online boarding is not tailored towards economy class passengers, which the airline should consider improving. Additionally, the airline should consider improving WiFi for all passenger classes in order to improve overall satisfaction. One possible solution would be for the airline to remove any fees that it might charge passengers for WiFi or to provide more options in terms of inflight entertainment.

VII. Future Work

The dataset analyzed for this project only contained results from one airline's customer surveys. However, the only information about the customers' characteristics and satisfaction with various aspects of the airline's inflight experience were included in the dataset. In the future, it would be useful to analyze the financial gains of retaining satisfied customers against the financial losses of having dissatisfied customers switch to a different airline. This would provide strategists at the airline with useful information about how to maximize their profits.

Moreover, the airlines should consider the benefits, costs, and impact on profits in terms of improving inflight WiFi for all customers and the online boarding experience for economy passengers. If improving inflight WiFi does help the airline retain customers or improve overall satisfaction, then it might not be worth it for the airline to improve inflight WiFi. Similarly, there are many difficulties improving online boarding. If the airline improves its mobile app or platform so satisfaction is higher among economy passengers, then satisfaction among business class passengers may drop. As a result, it might be difficult for the airline to improve its online boarding experience without deterring some customers.