SOUTHEAST AIRLINES CUSTOMER SATISFACTION SURVEY ANALYSIS

Jenny Cao

Nisha Rangnani

Rehman Sheikh

Aditya Tornekar

Yifan Wang

Syracuse University - IST 687

TABLE OF CONTENTS

Project Description

The project is based on customer survey results for Southeast Airlines and seeks to raise the Likelihood to Recommend score of the airline. The project dataset includes 10,282 observations and 32 variables. Primary goal is to increase NPS score using various attributes provided by the airlines and understand what features translate most towards gaining better recommendation scores. Additionally, we provide actionable data-driven recommendations to increase the Likelihood to Recommend, based on our findings.

Business Questions

1. Why CheapSeats, FlyFast and OursinAir?
2. Is there a gender bias in customer satisfaction ratings?
3. Is age a factor in higher or lower customer satisfaction ratings for CheapSeats customers?
4. Is Airline Status a factor in higher or lower customer satisfaction ratings?
5. What is the most frequent reason for travel for our customers?
6. Does reason for travel (type of travel) affect customer satisfaction ratings?
7. What Class is most traveled by our customers?
8. Does Class of travel affect customer satisfaction ratings?

Data Munging

As part of data cleaning process, “NA” values were found for the below variables:

1) **Departure.Delay.in.Minutes:** Departure delay in minutes was NA only when a flight was cancelled(208 observations), this was changed to 0 as the flight was cancelled.

2) **Arrival.Delay.in.Minutes:** Arrival delay had “NA” values for 235 observations, regardless of flight being cancelled or not.

   a) For cancelled flights: This was transformed and set to 0 as the flight was cancelled.

   b) For non-cancelled flights: “NA” values in this scenario were handled using departure delay time variable, as that could have been the minimum amount of possible delay.

3) **Flight.time.in.minutes:** This variable had “NA” values for 235 observations.

   a) For cancelled flights(214 rows): This was transformed and set to 0 as the flight was cancelled.

   b) For non-cancelled flights(21 rows): “NA” values in this scenario were handled by taking average of flight time of other flights which had the same flight distance.

4) **Likelihood.to.recommend:** This variable just had one “NA” value which was converted to a passive score of 7. (1-6 Detractors, 7-8 Passive, 9-10 Promoters)

The dataset also had duplicate rows which were removed, after which we were left with 9930 observations which were cleaned and were ready to use for analysis.

Modeling Techniques

A. Linear Regression

Linear regression models were created to understand the features that were most impacting the Likelihood to recommend variable in a positive or negative way.This was done to understand what were the possible conditions and scenarios were the customer would give a high/low recommendation score. Further, this analysis was used to check if these intermediate variables were affected by any other variables.

Primarily, models were created on filtered dataset of 2000 customers roughly, which consisted of 'Eco' & 'Business' class customers who were traveling for Personal or Business reasons using FlyToSun Airlines/FlyFast Airways/Oursin Airlines & travelers of status blue/silver/gold. These variables were selected based on initial linear model.

Later when checked, nearly similar trends were observed for the overall data as well.

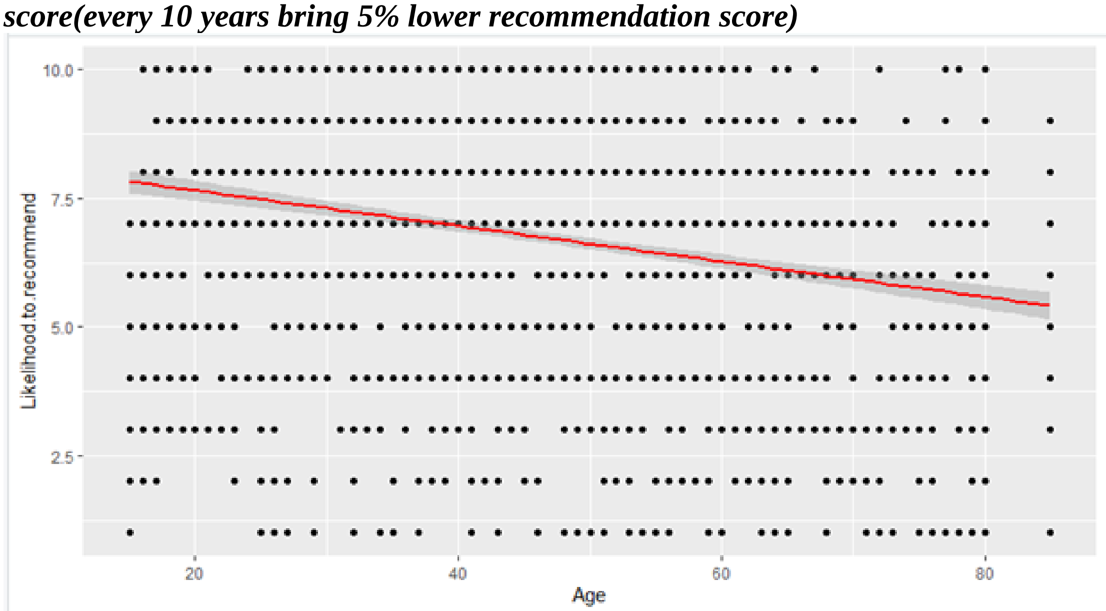
Below is a summary snippet of the created model on filtered data:

```
> summary(lm8)

Call:
lm(formula = Likelihood.to.recommend ~ Airline.Status + Age +
    Gender + Flights.Per.Year + Loyalty + Shopping.Amount.at.Airport +
    Eating.and.Drinking.at.Airport + Arrival.Delay.in.Minutes +
    olat + Class + Type.of.Travel + Partner.Code, data = ModelData)

Residuals:
   Min     1Q Median     3Q    Max
-6.893 -1.105  0.146  1.352  5.033

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                    5.6354433  0.3763810  14.973  < 2e-16 ***
Airline.StatusGold             0.7883713  0.1603692   4.916 9.59e-07 ***
Airline.StatusSilver           1.3401345  0.1099764  12.186  < 2e-16 ***
Age                           -0.0058570  0.0027895  -2.100  0.03589 *
GenderMale                     0.2168577  0.0874933   2.479  0.01328 *
Flights.Per.Year              -0.0103458  0.0041659  -2.483  0.01310 *
Loyalty                       -0.1501398  0.1135802  -1.322  0.18637
Shopping.Amount.at.Airport     0.0016704  0.0008483   1.969  0.04908 *
Eating.and.Drinking.at.Airport 0.0045878  0.0008857   5.180 2.46e-07 ***
Arrival.Delay.in.Minutes      -0.0049760  0.0010512  -4.734 2.37e-06 ***
olat                           0.0514837  0.0067763   7.598 4.70e-14 ***
ClassEco                      -0.1213940  0.1537001  -0.790  0.42974
Type.of.TravelPersonal Travel -2.5827892  0.1019770 -25.327  < 2e-16 ***
Partner.CodeEV                -0.3824199  0.1471997  -2.598  0.00945 **
Partner.CodeOU                 0.2154601  0.1468214   1.467  0.14241
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.864 on 1904 degrees of freedom
Multiple R-squared:  0.4258,    Adjusted R-squared:  0.4216
F-statistic: 100.8 on 14 and 1904 DF,  p-value: < 2.2e-16

>
```

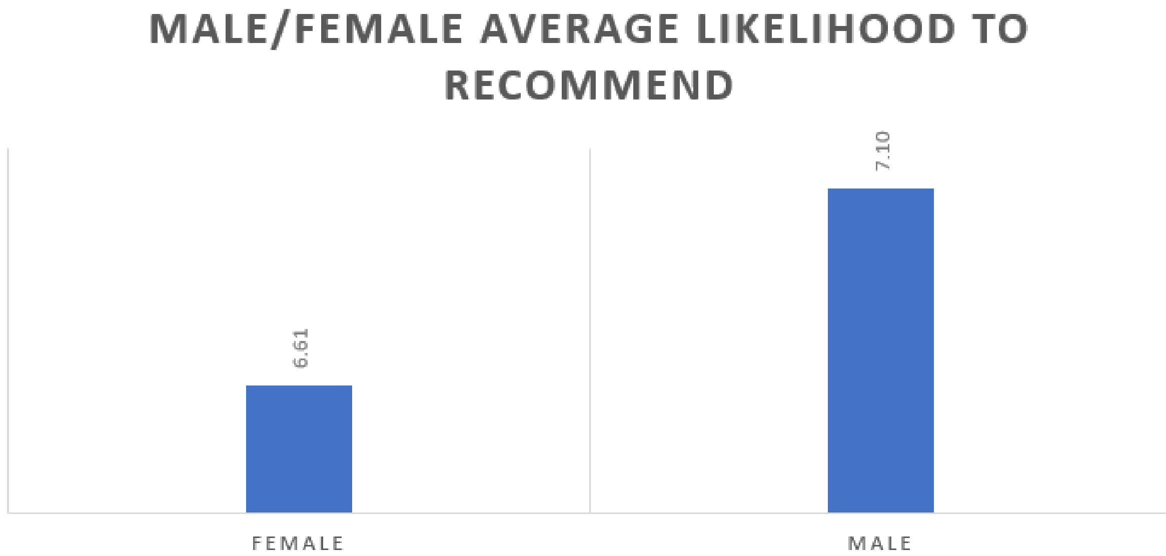
These models and variables were further explored by visualizing them using **ggplot**.
This was done to better understand the trend for the given parameters. Coefficients could have been used but plots are much better to grasp the information.

1) ***Airline Status: Silver status customers are the ones to boost the Likelihood to recommend,followed by gold,blue,platinum status customers. Each tuple of 3 silver status customers boosts the net recommendation score by 4 as compared to gold and blue status customers***

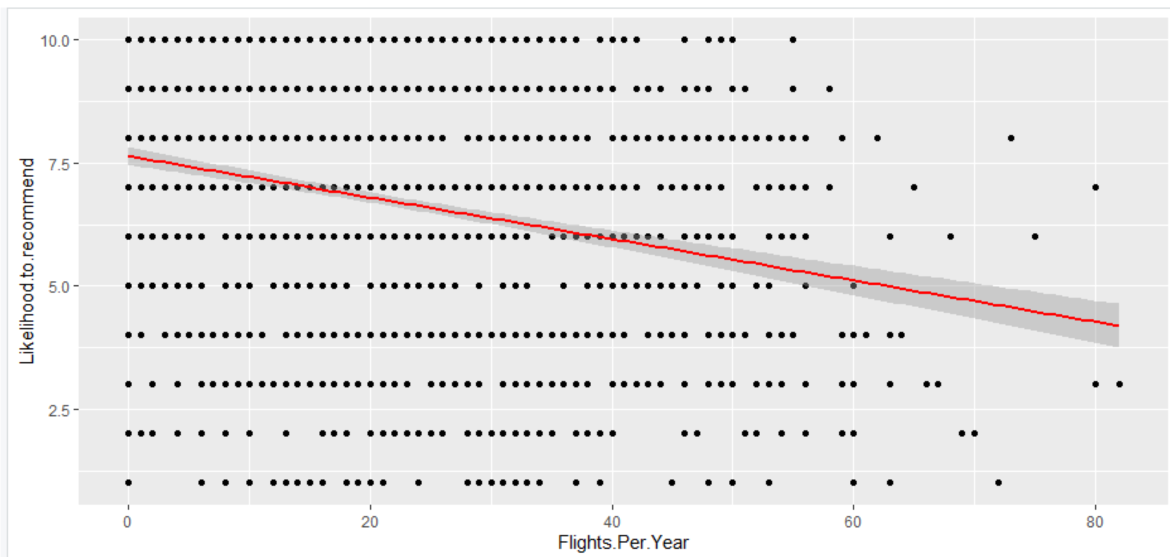
2) ***AGE: As the age increases, a customer tends to give negative recommendation score(every 10 years bring 5% lower recommendation score)***



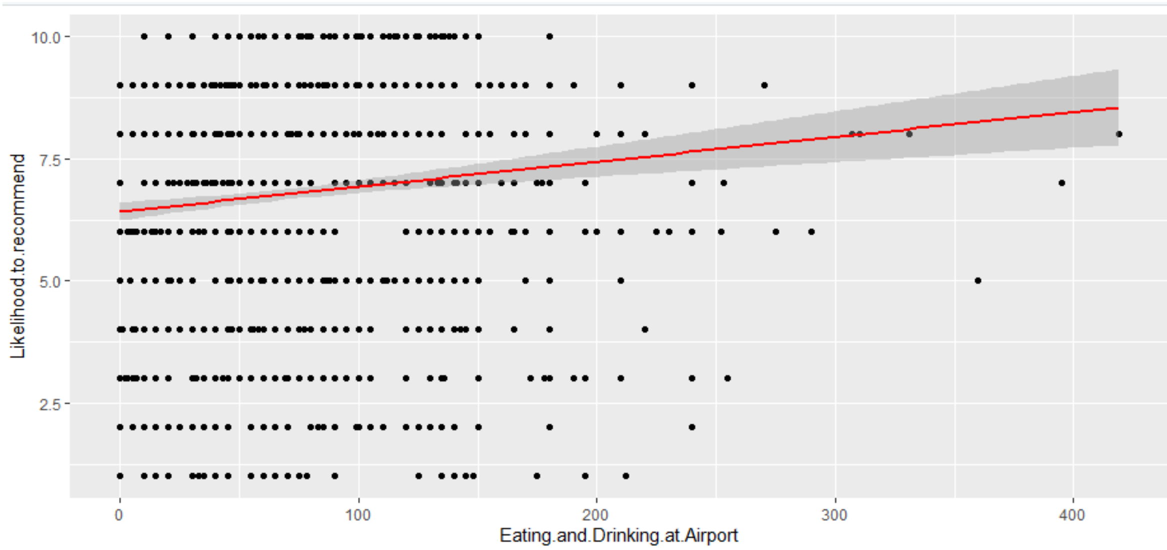
3) Gender: Males tend to give better recommendation scores



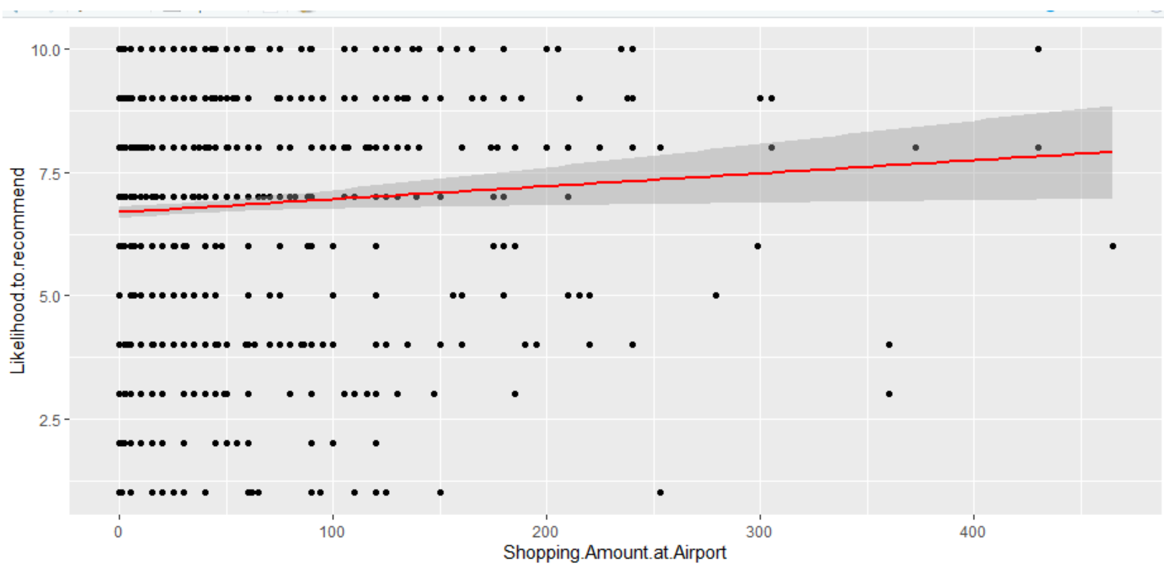
4) Flights per year: As the number of flights taken by a customer in a year increase, for each extra flight customers tend to give 0.1% lower recommendation scores.



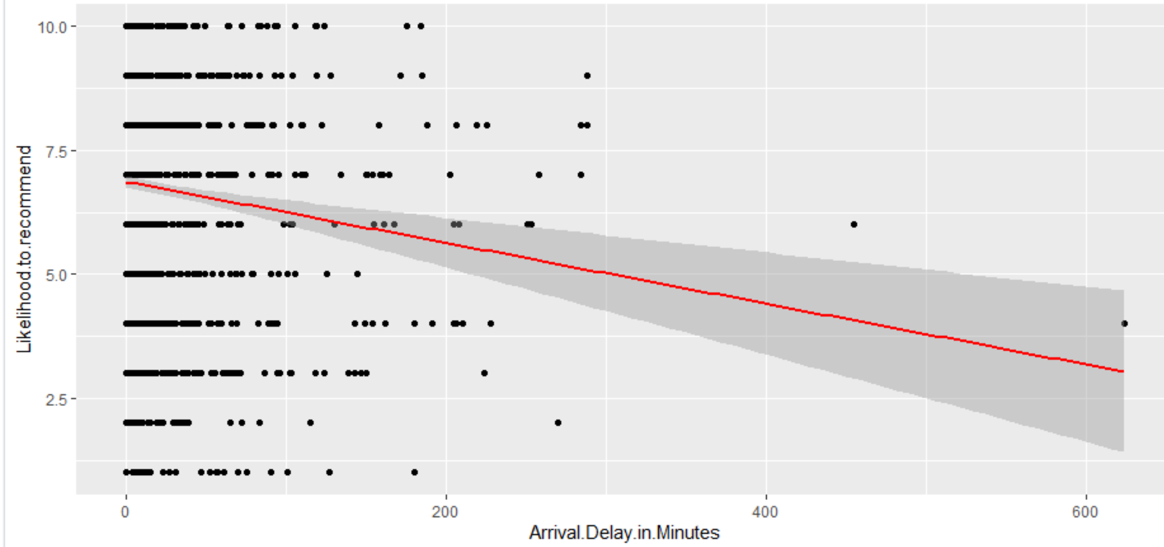
5) ***Eating & drinking at the airport: Customers who eat & drink at the airport tend to give positive recommendations. Each 100$ spent on eating & drinking at the airport translates to increased 4% recommendation score***



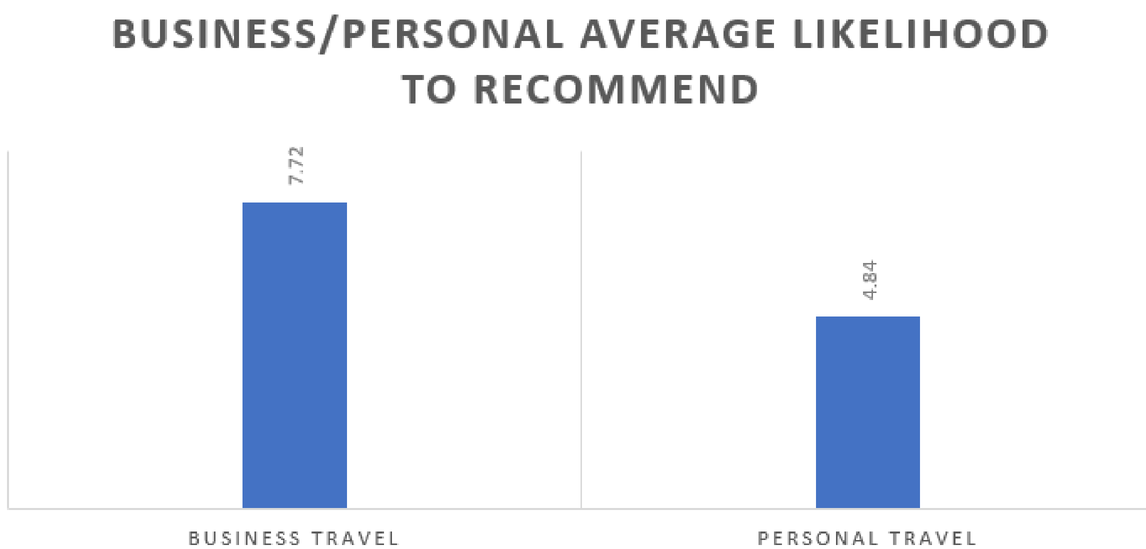
6) ***Shopping at the airport: More the amount spent at the airport on shopping better the recommendation score. Each 100$ spent on shopping translates to increased 1% recommendation score***



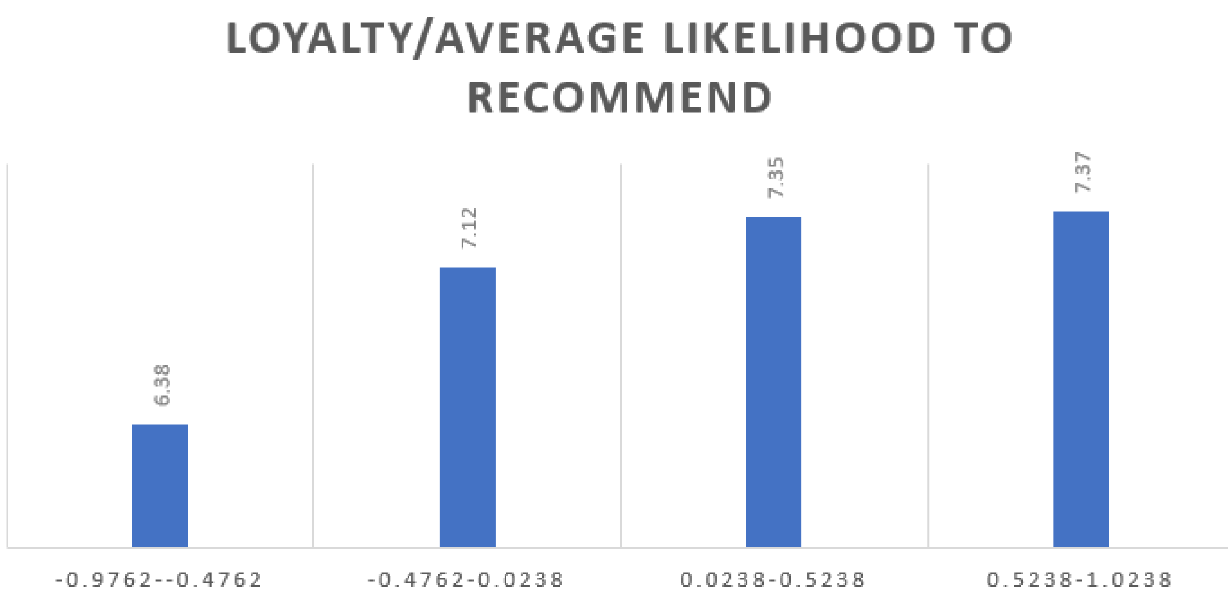
7) ***Arrival delay: Every hour of delay during arrival brings down recommendation score by 2.4%***

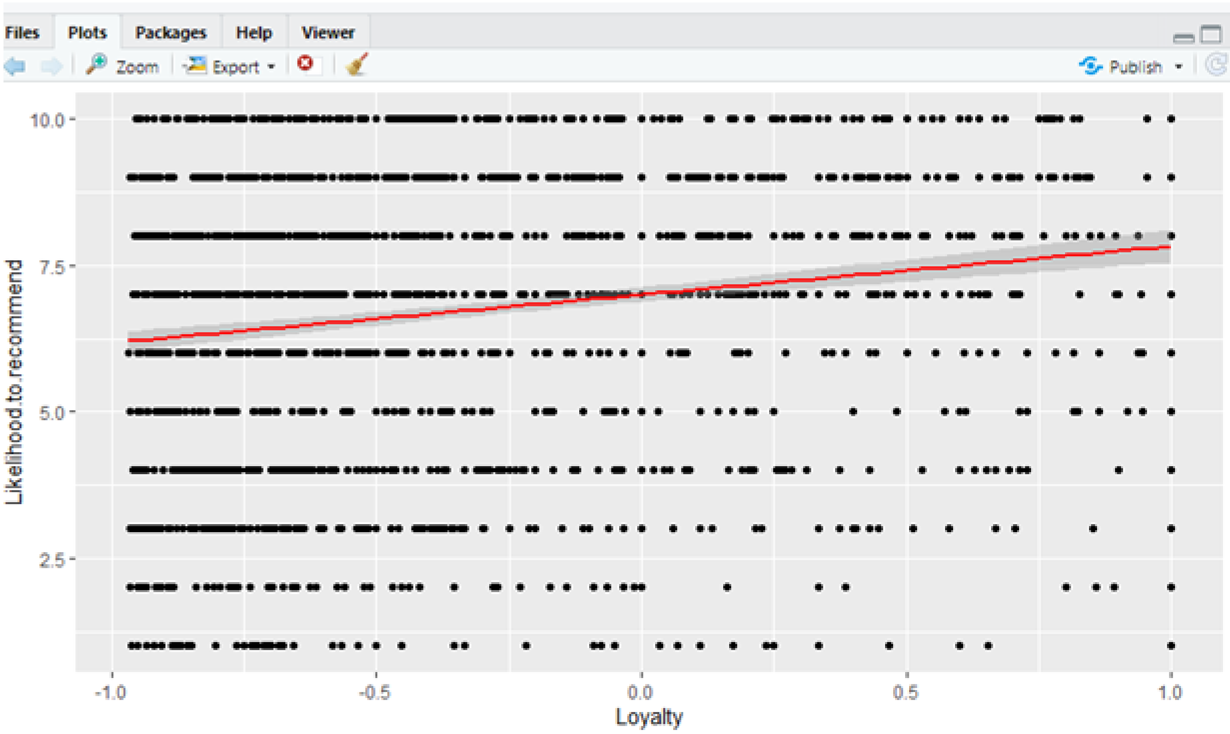


8) ***Type of travel: Customers traveling for personal work tend to give a negative 2.58 as compared to customers traveling for business reasons. The below graph shows average recommendation scores for each type of travel considered.***

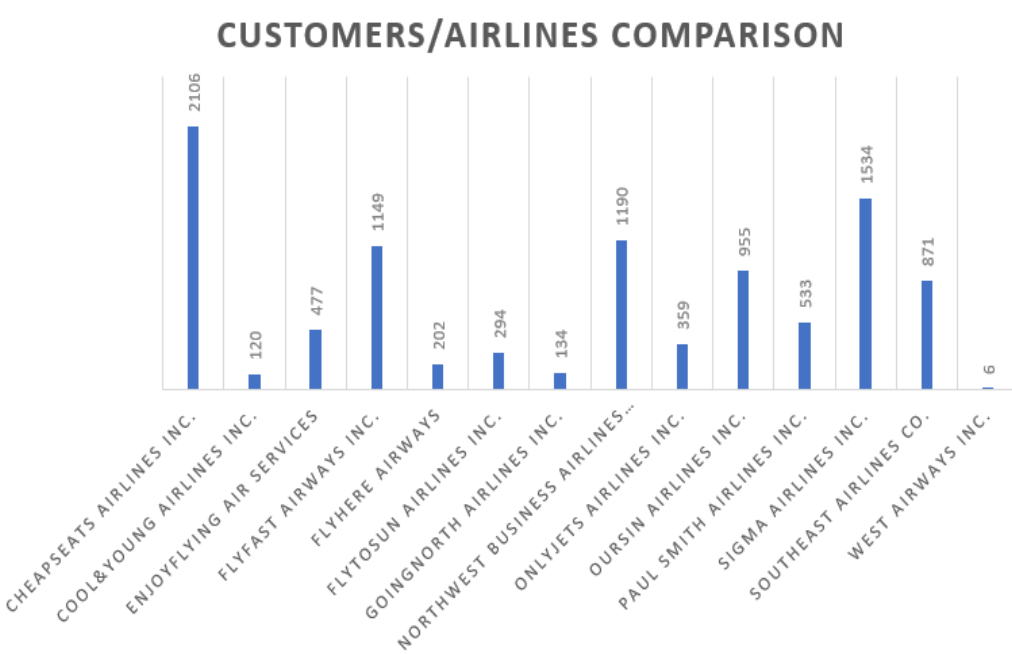


9) Loyalty: Overall Loyal customers tend to give better recommendation scores.





Top 3 airlines which had maximum customers were analyzed to understand the mass behaviour of customers:



Cheapseats Airlines Inc (2106 customers)
Sigma Airlines Inc (1534 customers)
FlyFast Airways Inc (1149 customers)

Common major parameters for the above three airlines impacting the Likelihood to recommend.

1) ***Airline.Status***
2) ***Gender***
3) ***Type.of.Travel***
4) ***Eating.and.Drinking.at.Airport***

Recommendations based on Linear Regression

1) *Higher the age lower the score, this might be resolved if travelers of old age are provided better resting facilities at the airport and better entertainment during the flight.*

2) *Increased number of flights lowers the score, which suggests that these customers are not happy but are still traveling anyways. Here the airline needs to give some kind of loyalty points to the more frequently flying customers, for better recommendation score.*

3) *If customers are given food options through coupons/points before & after flight then it might save food costs for the airlines for short flights, if the airlines can manage to cut off the food and just partner with other food services at the airport. This way the airline might still get a good recommendation score at lower costs.*

4) *If customers are given more options to shop at the airport via booked tickets through expiring coupons for next time it might increase the NPS and customer loyalty and still might be good for airlines if the customer is not a frequent flyer.*

B. Sentiment Analysis

The following codes create sentiment analysis for all data, specific companies, and SVM models that help us to offer insights.

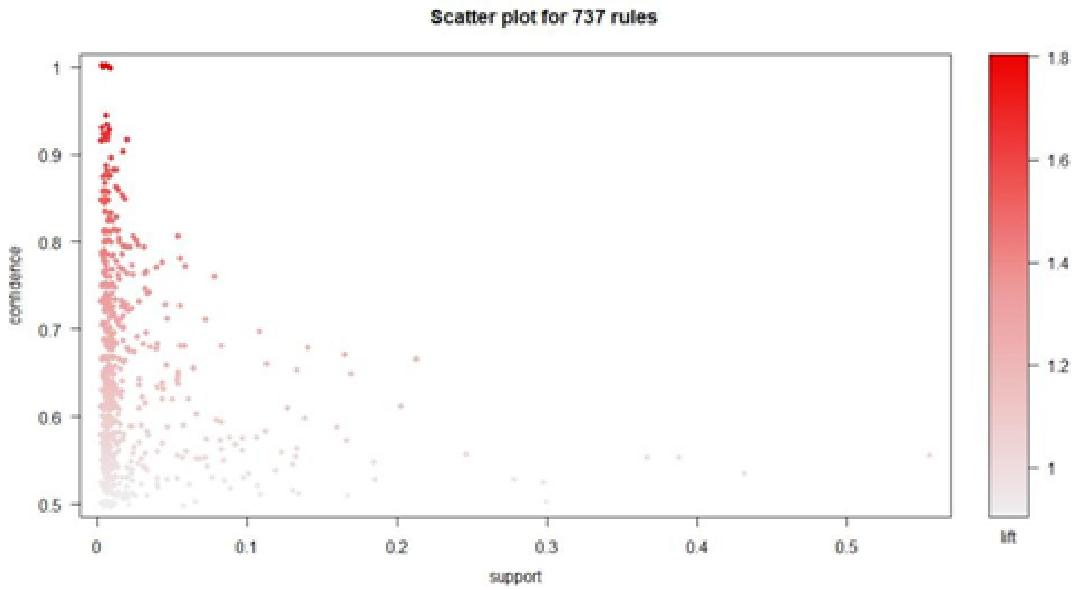
#This model is to analyze the entire customers who provided comments (free text) with likelihood less than 7 to see what caused this issue, and compare with the previous sentiment analysis to determine our insights.

#similarly, we created subsets for cheapSeats, FlyFast & OursinAir to perform sentiment analysis and Association rules

C.1 Association Rules

A support vector machine model designed to combine with and support the results of our sentiment analysis**.**

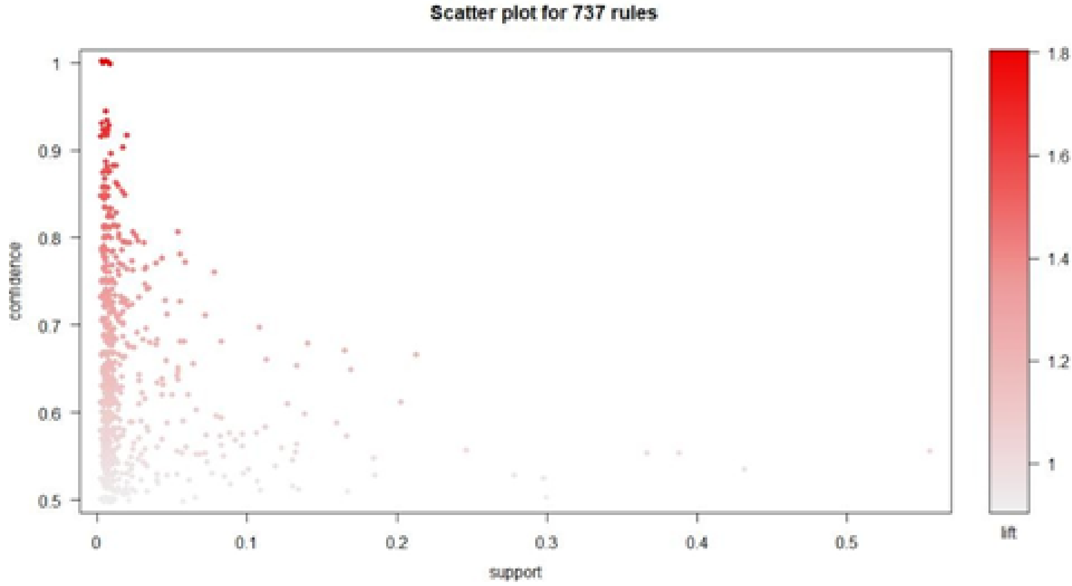
 # Support vector machine for CheapSeats



 # Contingency table to show the relationship between gender and price sensitivity

```
> #Contingency table to show the relationship between gender and price sensitivity
> CT1 <-table(cheaps$Gender,cheaps$Price.Sensitivity)
> CT1 <-prop.table(CT1)
> CT1

                0            1            2            3            4
  Female 0.0155109489 0.3681569343 0.1646897810 0.0068430657 0.0004562044
  Male   0.0168795620 0.2969890511 0.1222627737 0.0082116788 0.0000000000
```

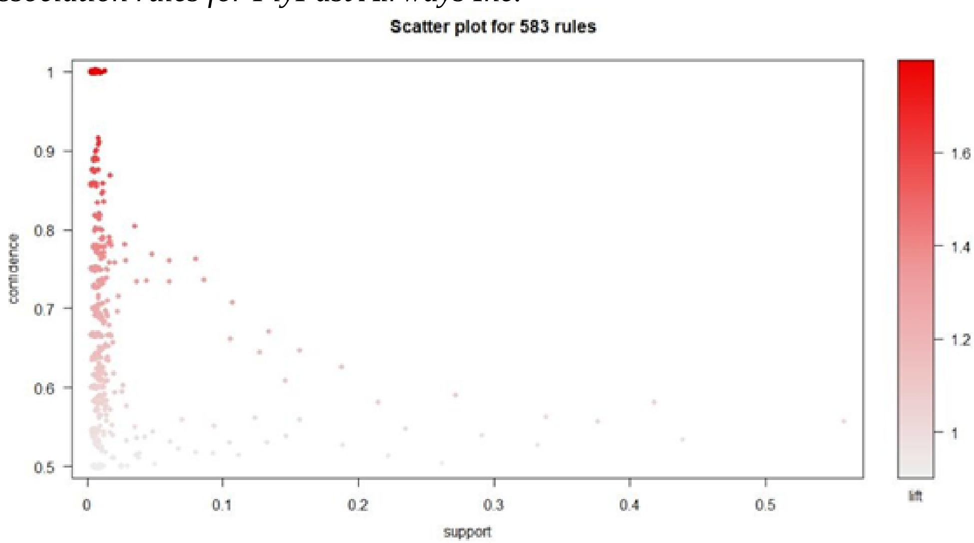
#Create the association rules for Oursin Airlines Inc.



#Create the contingency table between gender and price sensitivity

```
> OCT1 <- table(oursin$Gender,oursin$Price.Sensitivity)
> OCT2 <- prop.table(OCT1)
> OCT2

                   0           1           2           3           4
  Female 0.019230769 0.362348178 0.162955466 0.007085020 0.001012146
  Male   0.013157895 0.308704453 0.115384615 0.009109312 0.001012146
>
```

#Create association rules for FlyFast Airways Inc.



#Create the contingency table between gender and price sensitivity

```
> #Create the contingency table between gender and price sensitivity
> FFCT1 <- table(flyfast$Gender,flyfast$Price.Sensitivity)
> FFCT2 <- prop.table(FFCT1)
> FFCT2

                   0            1            2            3            4
  Female 0.0187553282 0.3776641091 0.1568627451 0.0042625746 0.0000000000
  Male   0.0127877238 0.2992327366 0.1236146633 0.0059676044 0.0008525149
> |
```

C.2 Association Rules

Description

After determining what the most significant factors are with respect to the Likelihood to Recommend, we decided to validate the findings through Association Rules. Additionally, in this model, we removed factors like Age and Gender to focus on which variables SouthEast can control or facilitate. The specific variables used were: Price Sensitivity, Loyalty, Shopping Amount, Eating and Drinking Amount, Departure Delay, Arrival Delay, and Likelihood to Recommend.

The outcome of the association rules were used in addition to validate, to determine which variables are strongly associated with giving a higher/lower NPS.

Before running the Association Rules on the CheapSeats, FlyFast, and OursInAir partners, the data was further prepared by creating breaks by grouping values. For example, the Likelihood to Recommend values were converted from anything below 7 to 0, and anything 7 and above as 1. This means a 0 value indicates a poor Likelihood to Recommend score, while a 1 is acceptable. The Loyalty values were converted to anything 0 and below to 0, and anything above 0 to 1. The Shopping Amount at Airport and Eating and Drinking at Airport values were converted from $0 spent to 0, $1-25 spent to 1, $26-50 spent to 2, $51 - 100 spent to 3, $101 - 200 spent to 4, $201 - 300 spent to 5, and $301+ spent to 6. The Arrival and Departure delay values were converted from 0 minute delay to 0, 1-25 minute delay to 1, 26-50 minute delay to 2, 51 - 100 minute delay to 3, 101 - 200 minute delay to 4, and 201+ minute delay to 5. The Price Sensitivity values were converted from 0-2 to 1, and 3-5 to 0. Afterwards, each column was converted to a factor, using the "as.factor," function to create the Association Rules.

CheapSeats Apriori Rules Output

Rules output with RHS = 1 (Likelihood to Recommend value 7 and over)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
	All	All	All	All	All	All
[481]	{Airline.Status=Silver,Shopping.Amount.at.Airport=4,Eating.and.Drinking.at.Airport=3}	{Likelihood.to.recommend=1}	0.005	1.000	1.524	11.000
[482]	{Airline.Status=Silver,Shopping.Amount.at.Airport=4,Arrival.Delay.in.Minutes=0}	{Likelihood.to.recommend=1}	0.008	1.000	1.524	17.000
[559]	{Airline.Status=Silver,Loyalty=1,Arrival.Delay.in.Minutes=3}	{Likelihood.to.recommend=1}	0.005	1.000	1.524	11.000
[598]	{Airline.Status=Gold,Shopping.Amount.at.Airport=2,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.006	1.000	1.524	14.000
[680]	{Shopping.Amount.at.Airport=1,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=3}	{Likelihood.to.recommend=1}	0.005	1.000	1.524	11.000
[689]	{Airline.Status=Silver,Eating.and.Drinking.at.Airport=2,Departure.Delay.in.Minutes=3}	{Likelihood.to.recommend=1}	0.005	1.000	1.524	11.000
[740]	{Shopping.Amount.at.Airport=1,Eating.and.Drinking.at.Airport=3,Class=Business}	{Likelihood.to.recommend=1}	0.008	1.000	1.524	18.000
[747]	{Airline.Status=Silver,Eating.and.Drinking.at.Airport=3,Class=Business}	{Likelihood.to.recommend=1}	0.010	1.000	1.524	21.000
[834]	{Airline.Status=Silver,Loyalty=1,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=1}	0.005	1.000	1.524	11.000
[835]	{Airline.Status=Silver,Eating.and.Drinking.at.Airport=3,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=1}	0.007	1.000	1.524	15.000

Showing 1 to 10 of 10,744 entries Previous 1 2 3 4 5 … 1075 Next

Rules output with RHS = 0 (Likelihood to Recommend value under 7)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
	All	All	All	All	All	All
[409]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[410]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[837]	{Price.Sensitivity=1,Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[838]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=2,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[839]	{Price.Sensitivity=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[840]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Arrival.Delay.in.Minutes=2,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[1316]	{Price.Sensitivity=1,Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=2,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[1317]	{Price.Sensitivity=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Arrival.Delay.in.Minutes=2,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.007	0.938	2.725	15.000
[832]	{Airline.Status=Blue,Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=0}	0.006	0.933	2.713	14.000
[833]	{Airline.Status=Blue,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=0}	0.006	0.933	2.713	14.000

FlyFast Apriori Rules Output

Rules output with RHS = 1 (Likelihood to Recommend value 7 and over)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
	All	All	All	All	All	All
[255]	{Loyalty=1,Shopping.Amount.at.Airport=4,Departure.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.008	1.000	1.746	9.000
[258]	{Loyalty=1,Shopping.Amount.at.Airport=4,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.009	1.000	1.746	11.000
[366]	{Loyalty=1,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=2}	{Likelihood.to.recommend=1}	0.005	1.000	1.746	6.000
[448]	{Shopping.Amount.at.Airport=1,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.006	1.000	1.746	7.000
[686]	{Loyalty=1,Shopping.Amount.at.Airport=4,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.006	1.000	1.746	7.000
[689]	{Loyalty=1,Shopping.Amount.at.Airport=4,Departure.Delay.in.Minutes=1,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.008	1.000	1.746	9.000
[690]	{Price.Sensitivity=1,Loyalty=1,Shopping.Amount.at.Airport=4,Departure.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.008	1.000	1.746	9.000
[692]	{Loyalty=1,Shopping.Amount.at.Airport=4,Arrival.Delay.in.Minutes=1,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.009	1.000	1.746	11.000
[693]	{Price.Sensitivity=1,Loyalty=1,Shopping.Amount.at.Airport=4,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.009	1.000	1.746	11.000
[824]	{Loyalty=1,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=2,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.005	1.000	1.746	6.000

Showing 1 to 10 of 1,830 entries Previous 1 2 3 4 5 … 183 Next

Rules output with RHS = 0 (Likelihood to Recommend value under 7)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
	All	All	All	All		All
[748]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=4,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=0}	0.009	0.917	2.146	11.000
[856]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=4,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.009	0.917	2.146	11.000
[857]	{Price.Sensitivity=1,Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=4,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=0}	0.009	0.917	2.146	11.000
[887]	{Price.Sensitivity=1,Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=4,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.009	0.917	2.146	11.000
[425]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=2,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=0}	0.006	0.875	2.049	7.000
[526]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=4,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=0}	0.012	0.875	2.049	14.000
[667]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=2,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=0}	0.006	0.875	2.049	7.000
[668]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=2,Arrival.Delay.in.Minutes=1,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.006	0.875	2.049	7.000
[669]	{Price.Sensitivity=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=2,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=0}	0.006	0.875	2.049	7.000
[749]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=4,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.012	0.875	2.049	14.000

Showing 1 to 10 of 890 entries Previous 1 2 3 4 5 … 89 Next

OursInAir Apriori Rules Output

Rules output with RHS = 1 (Likelihood to Recommend value 7 and over)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
	All	All	All	All	All	All
[405]	{Class=Business,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.010	1.000	1.499	10.000
[406]	{Eating.and.Drinking.at.Airport=3,Class=Business,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.008	1.000	1.499	8.000
[452]	{Loyalty=1,Departure.Delay.in.Minutes=2,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=1}	0.005	1.000	1.499	5.000
[502]	{Loyalty=1,Eating.and.Drinking.at.Airport=4,Departure.Delay.in.Minutes=2}	{Likelihood.to.recommend=1}	0.005	1.000	1.499	5.000
[538]	{Shopping.Amount.at.Airport=2,Eating.and.Drinking.at.Airport=3,Class=Eco Plus}	{Likelihood.to.recommend=1}	0.007	1.000	1.499	7.000
[543]	{Loyalty=1,Shopping.Amount.at.Airport=2,Eating.and.Drinking.at.Airport=4}	{Likelihood.to.recommend=1}	0.006	1.000	1.499	6.000
[550]	{Shopping.Amount.at.Airport=2,Eating.and.Drinking.at.Airport=2,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.005	1.000	1.499	5.000
[557]	{Loyalty=1,Shopping.Amount.at.Airport=2,Eating.and.Drinking.at.Airport=3}	{Likelihood.to.recommend=1}	0.008	1.000	1.499	8.000
[564]	{Shopping.Amount.at.Airport=2,Eating.and.Drinking.at.Airport=2,Departure.Delay.in.Minutes=0}	{Likelihood.to.recommend=1}	0.006	1.000	1.499	6.000
[616]	{Loyalty=1,Class=Eco Plus,Arrival.Delay.in.Minutes=1}	{Likelihood.to.recommend=1}	0.006	1.000	1.499	6.000

Showing 1 to 10 of 5,289 entries Previous 1 2 3 4 5 … 529 Next

Rules output with RHS = 0 (Likelihood to Recommend value under 7)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
	All	All	All	All	.	All
[112]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Departure.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.006	1.000	3.003	6.000
[274]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Departure.Delay.in.Minutes=3,Arrival.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.005	1.000	3.003	5.000
[295]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Departure.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.005	1.000	3.003	5.000
[296]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Departure.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.006	1.000	3.003	6.000
[297]	{Price.Sensitivity=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Departure.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.006	1.000	3.003	6.000
[298]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Departure.Delay.in.Minutes=3,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.006	1.000	3.003	6.000
[327]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Arrival.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.006	1.000	3.003	6.000
[516]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Departure.Delay.in.Minutes=3,Arrival.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.005	1.000	3.003	5.000
[517]	{Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Departure.Delay.in.Minutes=3,Arrival.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.005	1.000	3.003	5.000
[518]	{Price.Sensitivity=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Departure.Delay.in.Minutes=3,Arrival.Delay.in.Minutes=3}	{Likelihood.to.recommend=0}	0.005	1.000	3.003	5.000

Showing 1 to 10 of 825 entries Previous 1 2 3 4 5 … 83 Next

All 3 Partners Combined Apriori Rules Output

Rules output with RHS = 1 (Likelihood to Recommend value 7 and over)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
[3445]	{Loyalty=1,Eating.and.Drinking.at.Airport=3,Class=Eco,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0}	{Likelihood.to.recommend=1}	0.008	0.972	1.528	35.000
[4211]	{Loyalty=1,Eating.and.Drinking.at.Airport=3,Class=Eco,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.008	0.972	1.528	35.000
[4212]	{Price.Sensitivity=1,Loyalty=1,Eating.and.Drinking.at.Airport=3,Class=Eco,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0}	{Likelihood.to.recommend=1}	0.008	0.972	1.528	35.000
[4490]	{Price.Sensitivity=1,Loyalty=1,Eating.and.Drinking.at.Airport=3,Class=Eco,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.008	0.972	1.528	35.000
[3444]	{Loyalty=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0}	{Likelihood.to.recommend=1}	0.006	0.963	1.514	26.000
[4209]	{Loyalty=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.006	0.963	1.514	26.000
[4210]	{Price.Sensitivity=1,Loyalty=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0}	{Likelihood.to.recommend=1}	0.006	0.963	1.514	26.000
[4489]	{Price.Sensitivity=1,Loyalty=1,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=3,Departure.Delay.in.Minutes=1,Arrival.Delay.in.Minutes=0,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.006	0.963	1.514	26.000
[537]	{Loyalty=0,Shopping.Amount.at.Airport=2,Class=Eco Plus}	{Likelihood.to.recommend=1}	0.006	0.929	1.460	26.000
[1454]	{Loyalty=0,Shopping.Amount.at.Airport=2,Class=Eco Plus,Flight.cancelled=No}	{Likelihood.to.recommend=1}	0.006	0.929	1.460	26.000

Showing 1 to 10 of 4,559 entries Previous 1 2 3 4 5 … 456 Next

Rules output with RHS = 0 (Likelihood to Recommend value under 7)

Show 10 entries Search:

	LHS	RHS	support	confidence	lift	count
[210]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=2}	{Likelihood.to.recommend=0}	0.005	0.786	2.159	22.000
[397]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Arrival.Delay.in.Minutes=2,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.005	0.786	2.159	22.000
[43]	{Loyalty=0,Eating.and.Drinking.at.Airport=5,Class=Eco}	{Likelihood.to.recommend=0}	0.007	0.738	2.028	31.000
[155]	{Loyalty=0,Eating.and.Drinking.at.Airport=5,Class=Eco,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.007	0.732	2.011	30.000
[156]	{Price.Sensitivity=1,Loyalty=0,Eating.and.Drinking.at.Airport=5,Class=Eco}	{Likelihood.to.recommend=0}	0.007	0.732	2.011	30.000
[146]	{Price.Sensitivity=0,Loyalty=0,Shopping.Amount.at.Airport=0,Class=Eco}	{Likelihood.to.recommend=0}	0.006	0.727	1.999	24.000
[354]	{Price.Sensitivity=0,Loyalty=0,Shopping.Amount.at.Airport=0,Class=Eco,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.006	0.727	1.999	24.000
[357]	{Price.Sensitivity=1,Loyalty=0,Eating.and.Drinking.at.Airport=5,Class=Eco,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.007	0.725	1.992	29.000
[493]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Departure.Delay.in.Minutes=1}	{Likelihood.to.recommend=0}	0.010	0.717	1.969	43.000
[616]	{Loyalty=0,Shopping.Amount.at.Airport=0,Eating.and.Drinking.at.Airport=1,Class=Eco,Departure.Delay.in.Minutes=1,Flight.cancelled=No}	{Likelihood.to.recommend=0}	0.010	0.717	1.969	43.000

Showing 1 to 10 of 694 entries Previous 1 2 3 4 5 … 70 Next

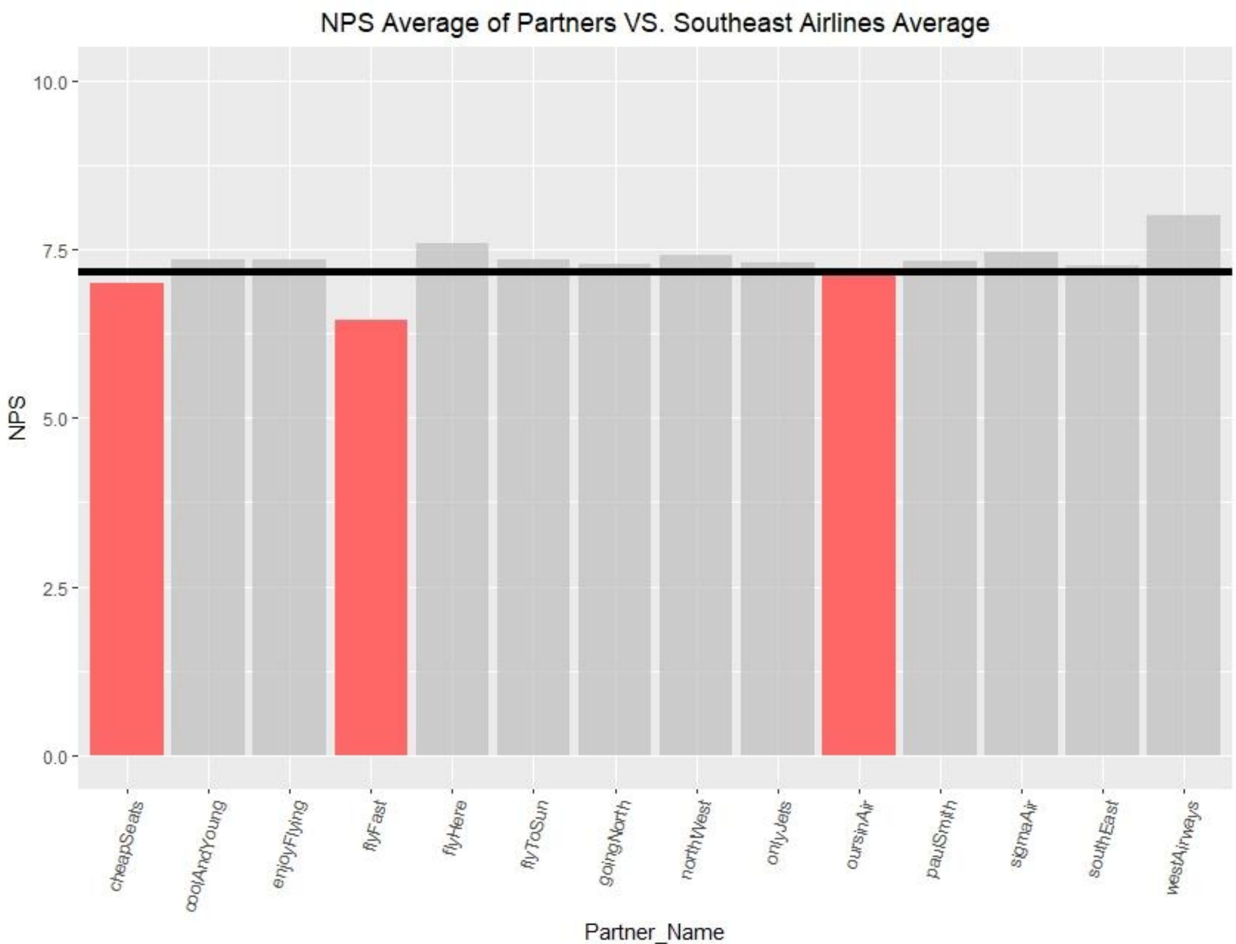
Association Rules Outcome

Based on the outcomes, we felt our findings from previous models were sufficiently validated. Additionally, we observed a consumer behavior trend between the 3 airline partners. According to the apriori outputs, customers of the airline partners tend to spend money on eating and drinking, as well as shopping in the airport. Furthermore, the outputs indicate loyal customers are more likely to give a higher Likelihood to Recommend score, while the inverse is true.

Business Questions

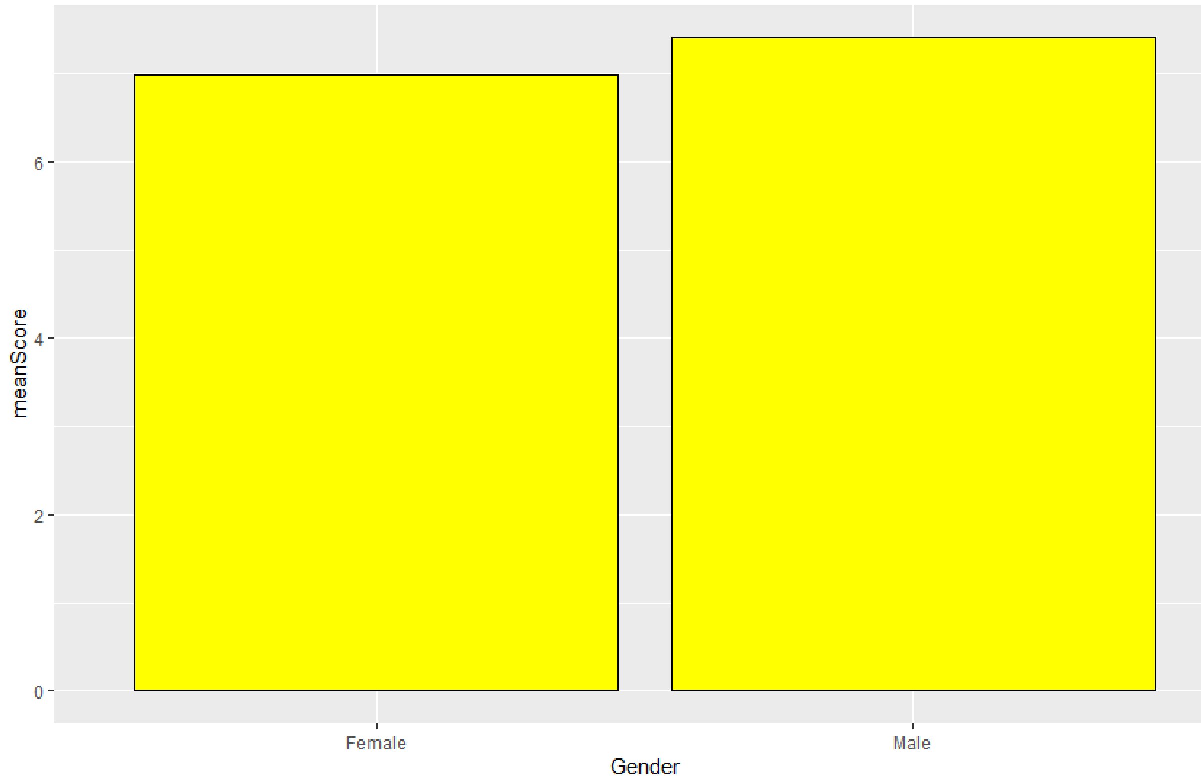
1. Why CheapSeats, FlyFast and OursinAir?

The graph below depicts the average customer satisfaction ratings of all partners (Black solid horizontal line) VS. each partner's average customer satisfaction ratings. The three airlines - CheapSeats, FlyFast and OursinAir have average of Likelihood to Recommend score below the combined average of all partners. Therefore, we chose to focus on these three airlines.



2. Is there a gender bias in customer satisfaction ratings?

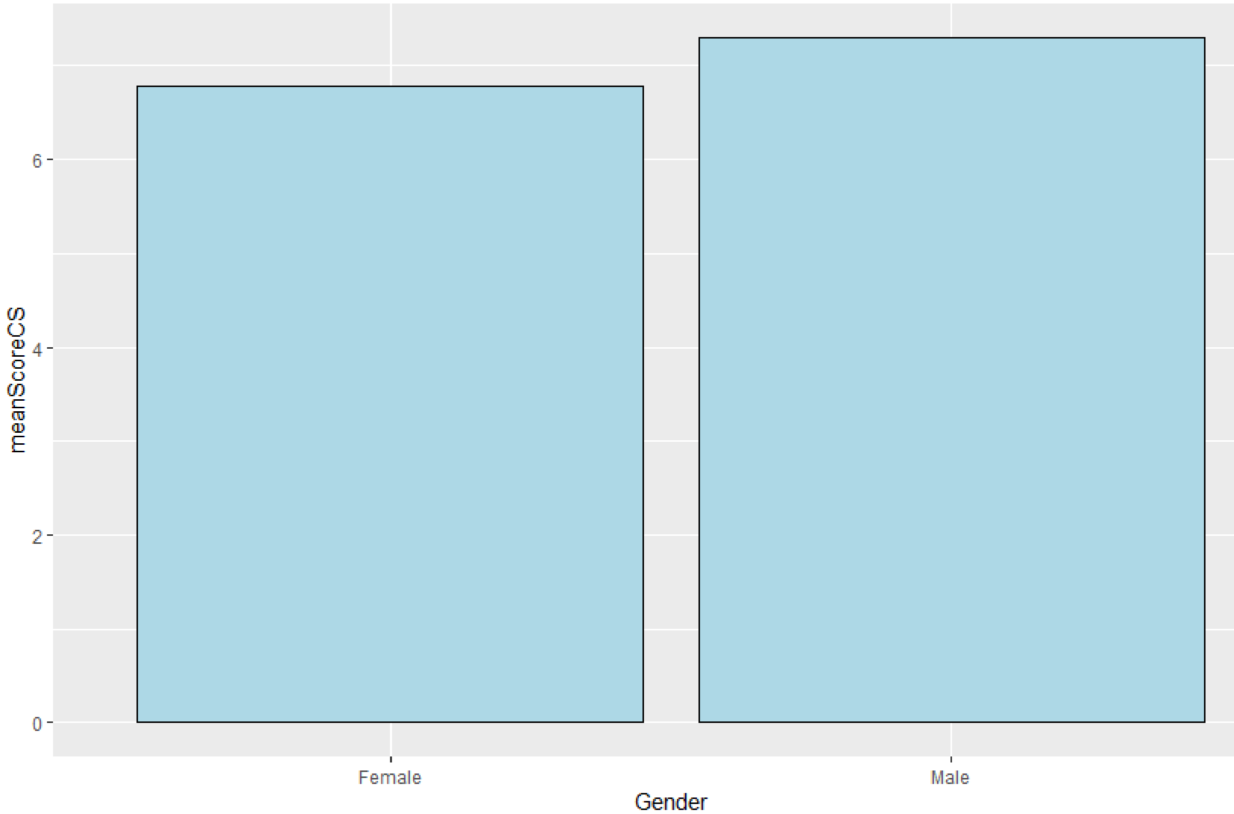
Yes, according to the overall statistics for gender and their respective average of Likelihood to Recommend score; we can see that Females tend to give a lower score.



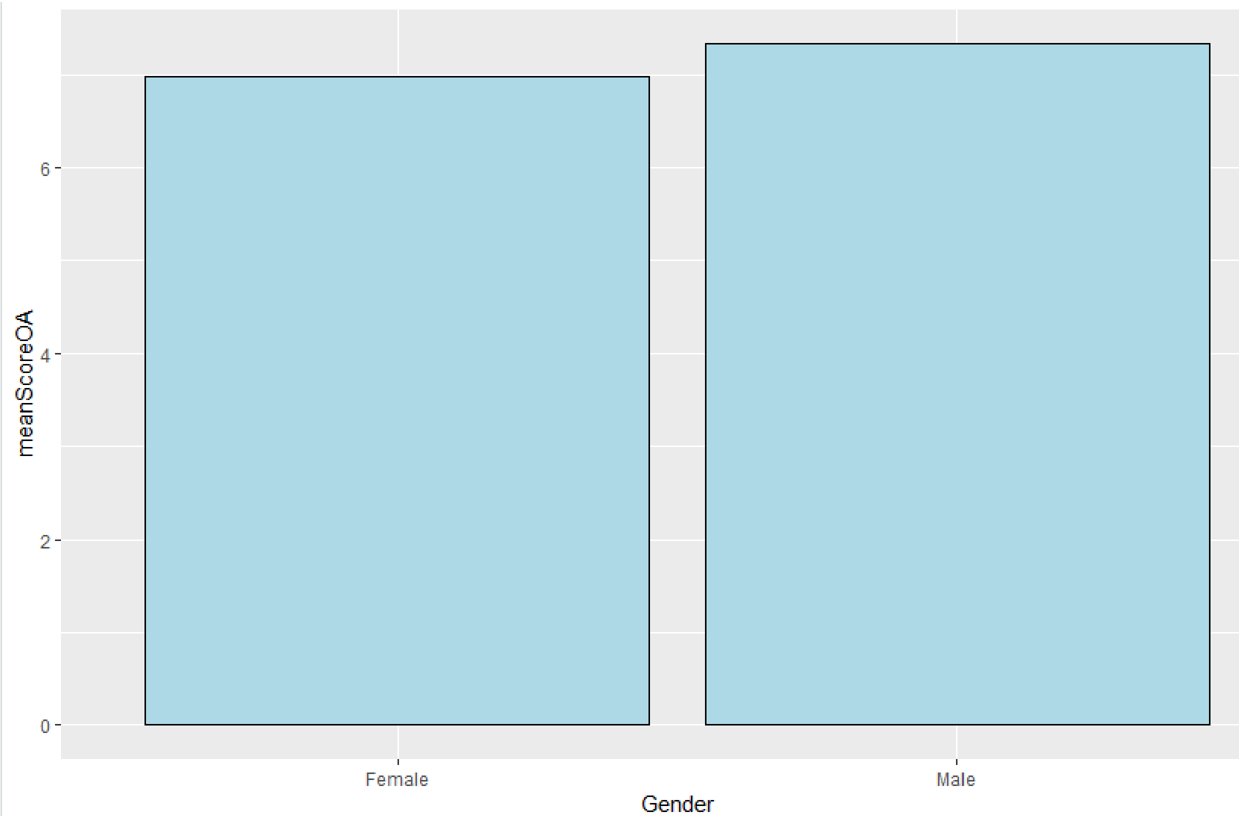
Even when we individually check average of Likelihood to recommend for CheapSeats, OursinAir and FlyFast, we can see the gender bias.

```
dfGenCheapS <- dfnew %>%
                    filter(str_trim(Partner.Name)=="Cheapseats Airlines Inc.") %>%
                    group_by(Gender) %>%
                    summarise(meanScoreCS = mean(Likelihood.to.recommend))
gg2<- ggplot(dfGenCheapS, aes(x=Gender, y=meanScoreCS)) + geom_col(fill='Light Blue', color='black')
gg2
.
```

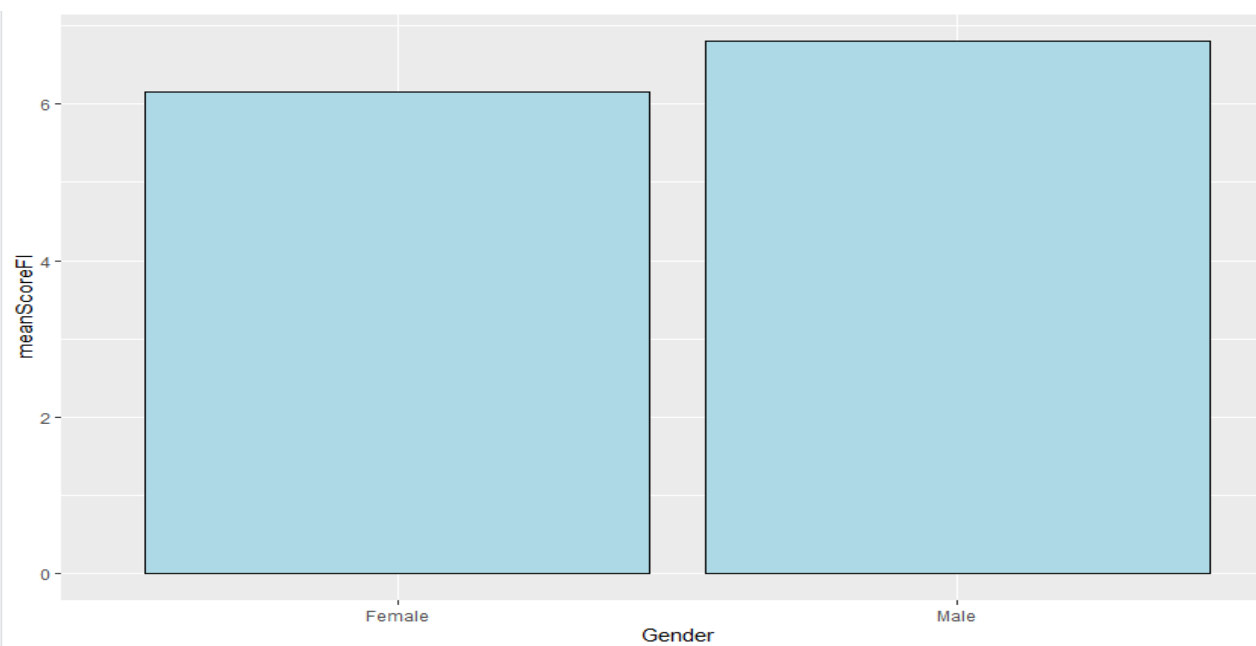
Gender VS. average Likelihood to Recommend for just **CheapSeats** passengers:



Gender VS. average Likelihood to Recommend for just **OursinAir** passengers:

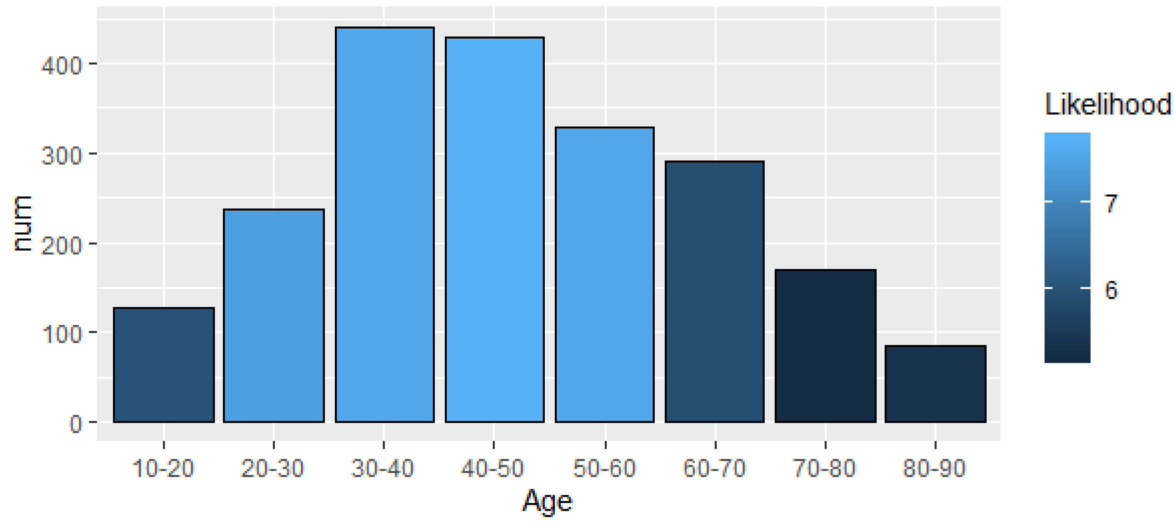


Gender VS. average Likelihood to Recommend for just **FlyFast** passengers:



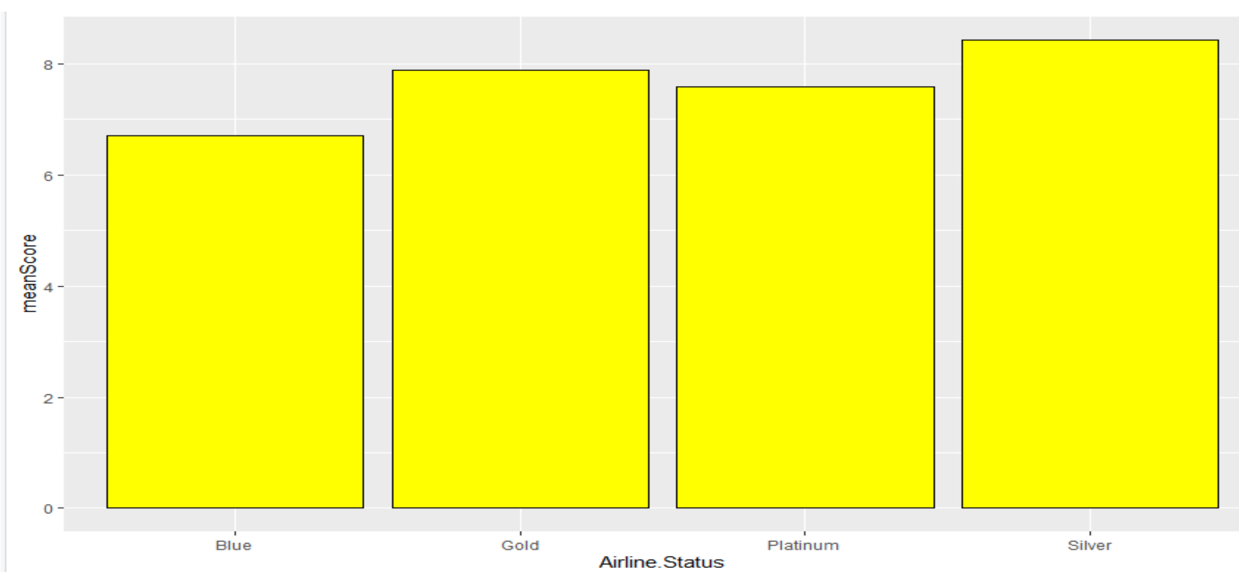
3. Is age a factor in higher or lower customer satisfaction ratings for CheapSeats customers?

Yes, age affects customer satisfaction ratings for CheapSeats passengers. As we can observe, as Age goes over 60, average customer satisfaction rating starts decreasing (changes to a darker shade of blue).

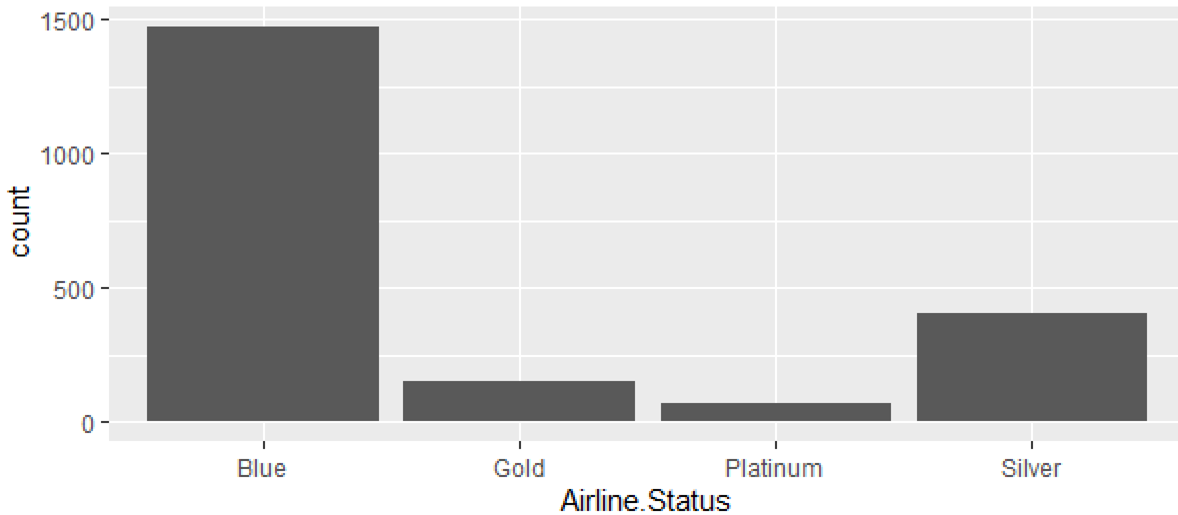


4. Is Airline Status a factor in higher or lower customer satisfaction ratings?

Yes, Blue customers give the least Likelihood to Recommend score; whereas, Silver customers tend to give highest Likelihood to Recommend score.

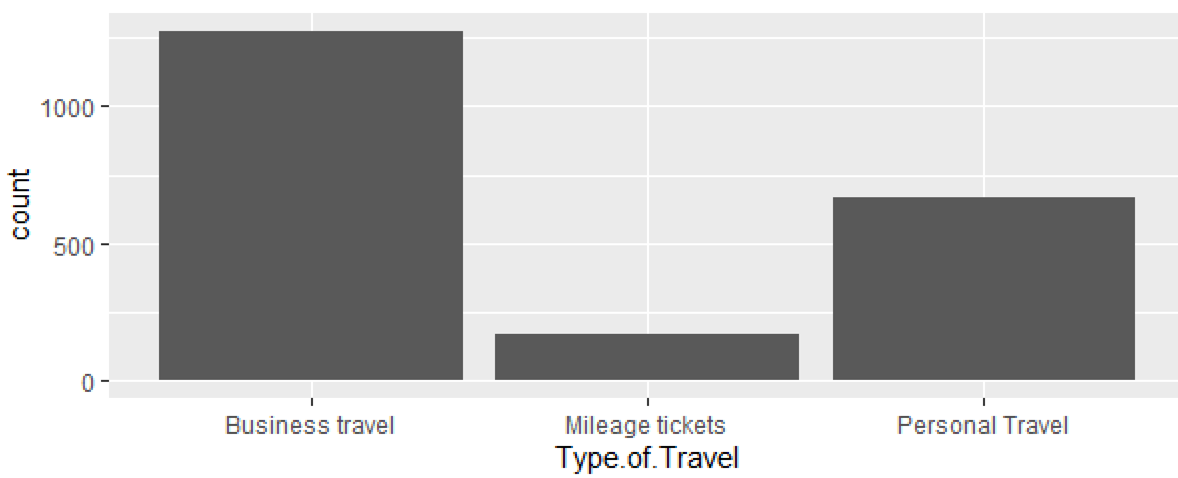


Blue and Silver constitute our most number of travelers:



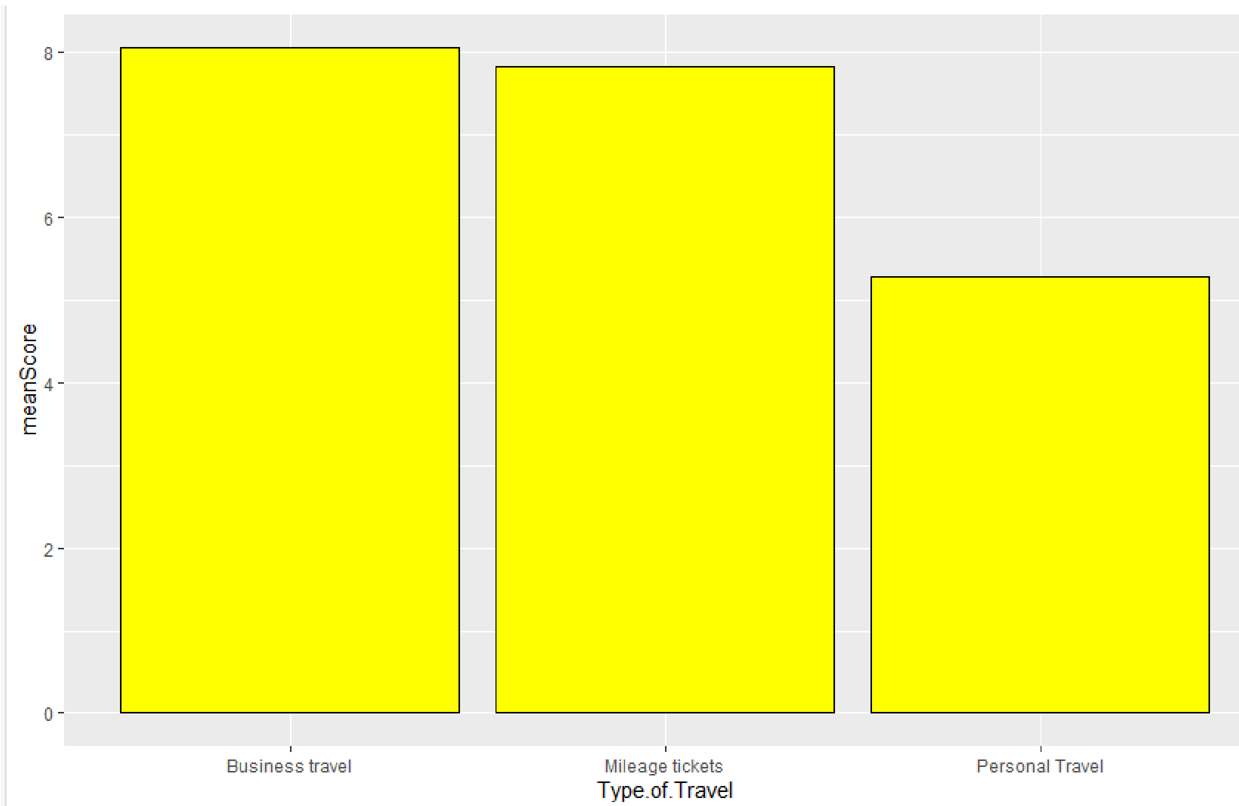
5. What is the most frequent reason for travel for our customers?

Most of our travelers travel for business purposes.



6. Does reason for travel (type of travel) affect customer satisfaction ratings?

Yes, Business travelers tend to give highest ratings. Whereas, travelers traveling for personal reasons give comparatively very low satisfaction ratings.



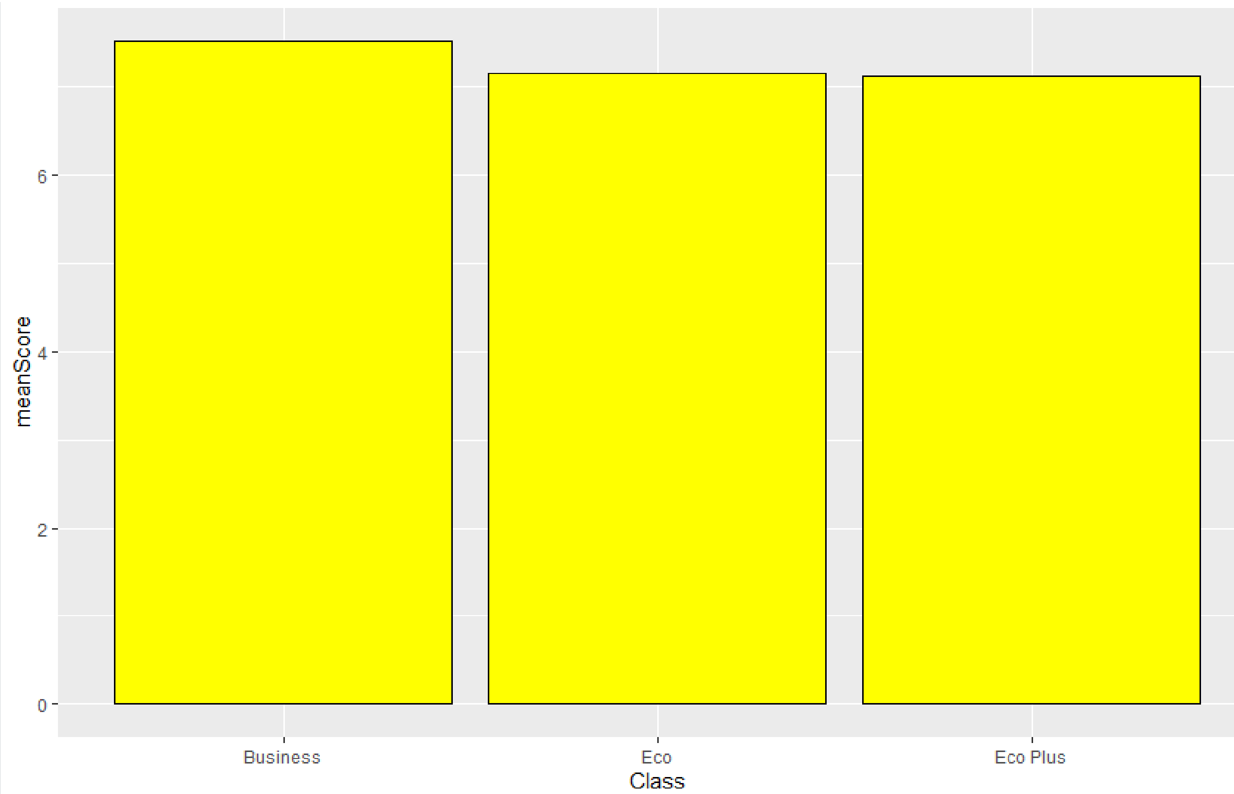
7. What Class is most traveled by our customers?

Economy class is most traveled by our customers, followed by Economy Plus, followed by Business class.

	Class	count
3	Eco	8099
2	Eco Plus	1037
1	Business	794

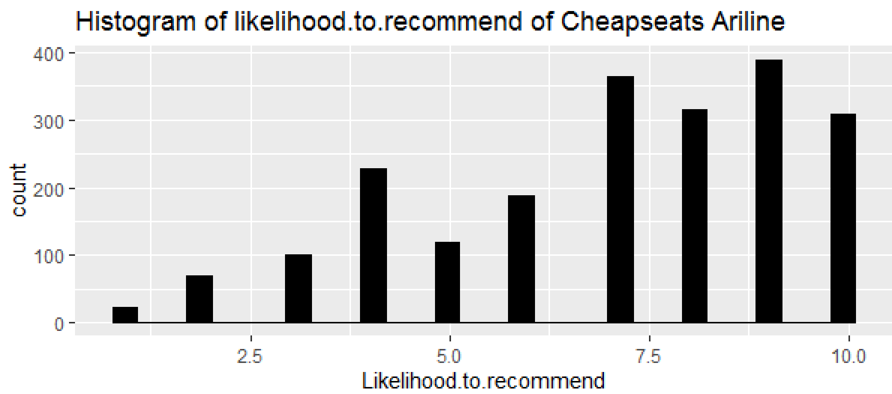
8. Does Class of travel affect customer satisfaction ratings?

Business class travelers give highest customer satisfaction rating.



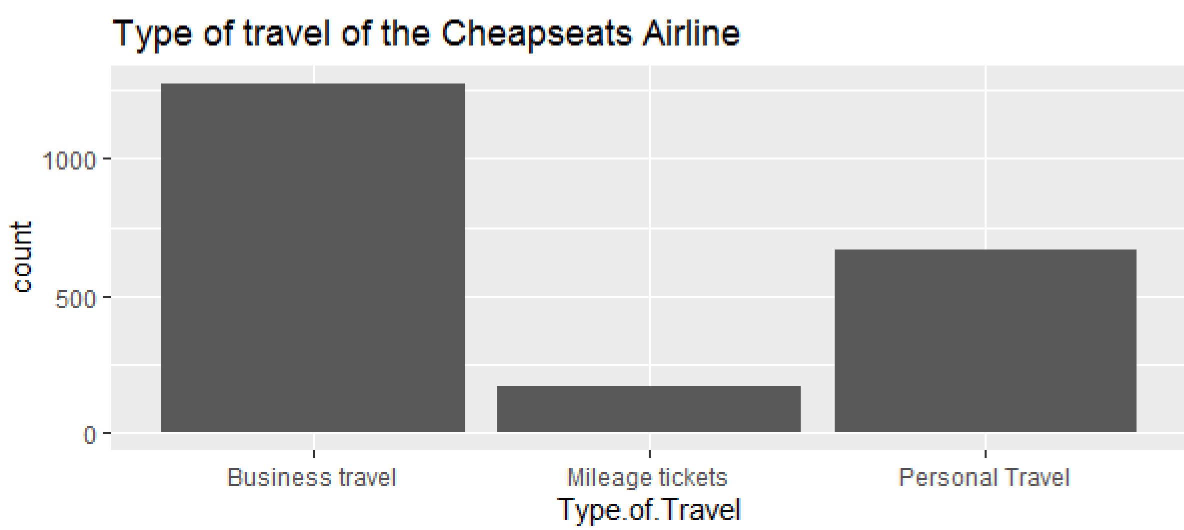
ADDITIONAL VISUALIZATIONS

Histogram of likelihood.to.recommend of Cheapseats Airline



This graph shows the likelihood distribution of Cheapseats Airline Inc. The number of likelihood from 1 to 10 is 23, 140, 306, 916, 590, 1122, 2548, 2520, 3501, and 3090. Therefore, the number of promoters is 6591.

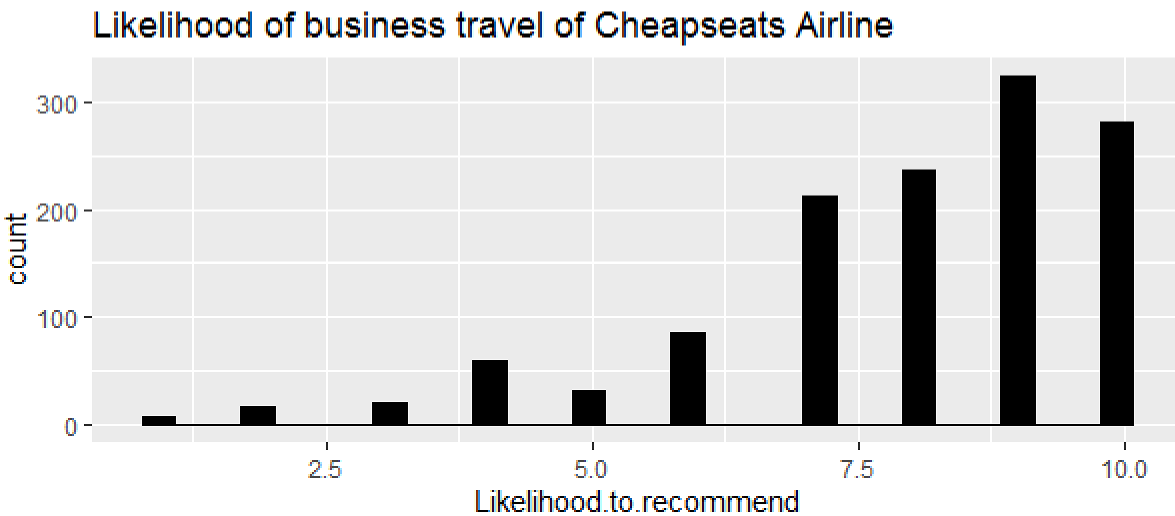
Type of travel of the Cheapseats Airline



Description

This graph shows the distribution of the travel type of Cheapseats Airline Inc. There are 1274 business travels, 167 mileage tickets and 665 personal travels. The business travel occupies a large proportion. And the second is personal travel. Therefore, we will focus on these two types of travel.

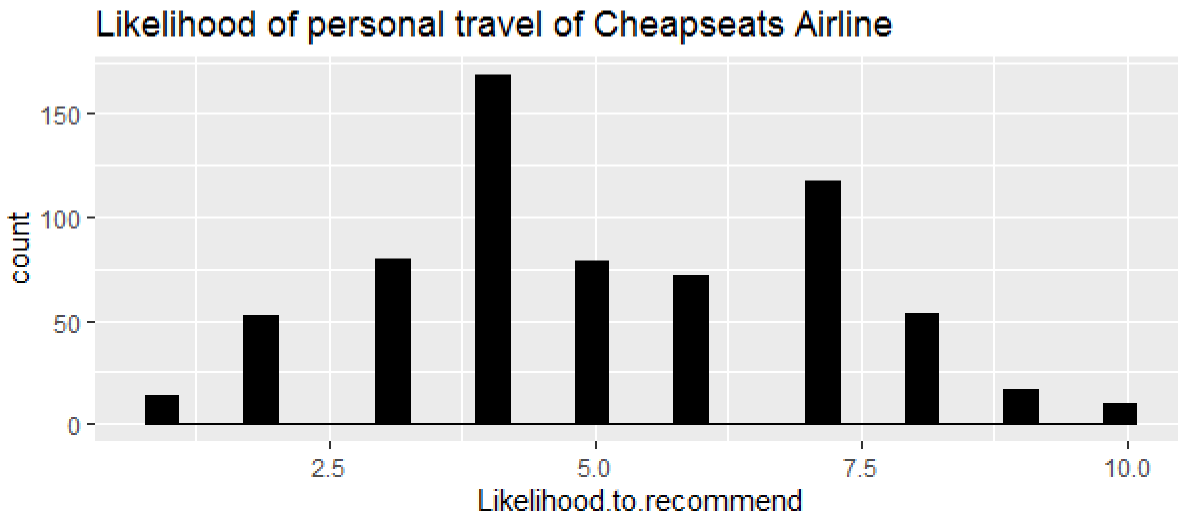
Likelihood of business travel of Cheapseats Airline



Description

This graph shows the distribution of likelihood of business type of cheapseats airline inc. The counts of each likelihood is 7, 34, 60, 236, 160, 510, 1491, 1888, 2916 and 2810. So the promoters counts is 5726 and the proportion is 56.63%. It seems like a person who is on a business type of travel is more likely to recommend the airline to others.

Likelihood of personal travel of Cheapseats Airline



Description

This graph shows the distribution of personal type of travel and the counts of each category is 14, 106, 240, 676, 395, 432, 819, 432, 153 and 100. The number of promoters is 253 which occupies 7.51%. Therefore, we prefer to only focus on the people going on a business type of travel.

INTERPRETATIONS & RECOMMENDATIONS

Sentiment Analysis

We ran the sentiment analysis for the entire data set where the likelihood to recommend was less than 7, to find out the top 10 words of their reviews.

```
> head(wordCounts4,10)
   flight southeast    service    airline    delayed    luggage      worst        one      seats    flights
      116        77         47         47         34         28         28         24         24         24
```

The word cloud of the entire data set where likelihood to recommend less than seven

According to our result of head command and word cloud, we found the word service is the most appeared word in these three partners' comments. Since we already knew female passengers are more sensitive to price, so we created the SVM models based on each of the three companies with the results of sentiment analysis to see what type of female passengers these three partners should focus on and provide better services.

Our SVM plots and association rules tell us the group of elder females who are blue airline status, personal travel, and price sensitivity equals to one are the group the companies need to concentrate on.

The attributes that we used in the models are the Airline. Status, Gender, Age, Price.Sensitivity, Loyalty, Type of Travel, and Class.

CheapSeats Word Cloud

OursinAir Word Cloud

FlyFast Word cloud

Our sentiment results show that the difference between positive and negative words is less than 0.01, which is not useful to decide if the customer's reviews are a factor that impacted the companies' customers' loyalty and the NPS score.

However, we could use the results of our SVM models combine with the results of text mining to provide an insight in order to enhance the company's NPS score. We recommend that these companies ought to provide better services to elder females because they are the majority results of our association rules.

Insights

Overall, we suggest that those three partners should provide more professional training of serving elder groups, especially for the cabin crews and boarding services crews. Moreover, we strongly recommend these three partners develop marketing campaigns that focus on female passengers. For instance, they may provide baby seats, baby foods(adequate water for babies' bottles on board), and baby bassinet services for international travelers. For domestic travelers, they can place free diapers in the washrooms on board.

Association Rules

The outcome of the association rules that were based on the significant variables that SouthEast can control or facilitate indicated loyalty to the airline was the determining factor in whether a high or low NPS was received. The Association Rules algorithm was used on the three airline partners (cheapSeats, flyFast, oursinAir) that were identified as underperformers. We observed a similar pattern throughout the three airlines. Based on the code provided, we observed that the following attributes were most associated with receiving a higher NPS:

- Loyal to the airline
- Not sensitive to pricing
- Spends $51-100 on eating and drinking
- Spends $26-50 on shopping
- At most a 25-minute departure delay

Additionally, we observed the following attributes were most associated with receiving a lower NPS:

- Not loyal to the airline
- Spends $26-50 on eating and drinking
- Spends $26-50 on shopping
- Arrived at destination at least 26 minutes late

Based on these findings, we understand that individuals are more likely to spend money on food and shopping, while loyalty and arrival delay are the attributes that may determine a high or low NPS.

Recommendations

From the findings of the association rules, we can produce two recommendations to increase loyalty and to ease price sensitivity on consumers.

Coupon-Based Approach

Customers tend to spend money on food and drink, as well as shopping at the airport before their flight. It would be beneficial to partner with restaurants and stores in the airport to offer customers coupons during the purchasing phase to these selected restaurants and stores. There is a low switching cost between Southeast and competitors. By providing customers with in-airport discounts, we may lock-in customers and in return boost the NPS.

Revamping Loyalty Program

Based on the Project Overview document, we understand the current SouthEast Airlines Loyalty program is not as effective as it can be. Additionally, our findings indicate that customers that are identified as a “Blue”, or entry level airline status, are more likely to give a lower NPS. To combat a lower NPS from this customer segment, it would be beneficial to revamp the current loyalty program. Southeast needs to rethink how it values its customers, and may do that by expanding the loyalty program to outside of the airport. Southeast may partner with credit card companies, restaurants, and stores to provide customers with loyalty points with select purchases outside of the airport. This will allow customers to build loyalty points on their daily transactions, gain a higher airline status above “Blue,” and use the loyalty points towards the cost of their airline ticket. This will lock-in customers as the loyalty program in more encompassing and ultimately boost the NPS.