Falcon Airlines: Identifying Dissatisfied Customers to Uncover the Issues

Joenell Leina'ala Dietmeyer
UT Austin/Great Learning
October 26, 2021

Executive Summary

Problem Statement

In a highly competitive climate, airlines fight for the loyalty of repeat customers in order to increase profits and maintain stability. To motivate these customers from booking flights with their competition, Falcon airlines must maintain a high level of customer satisfaction, especially for those that travel frequently for business and pleasure. Falcon needs to find how passengers value each amenity from all parameters listed in the satisfaction survey and try to quicky improve upon those offerings that are most important. These improvements will help keep Falcon customers booking with them and not their competitors.

Proposed Solution

Through data cleaning, pre-processing and model building, we will find the most important features derived from the two customer surveys provided by Falcon Airlines. We will create hyper-tuned decision tree models and train ensemble models in order to accomplish our goals of identifying the main features offered by Falcon that are falling short of customer expectations and causing their dissatisfaction. These discoveries will be sent to the executive teams within Falcon to prioritize and develop plans to quickly improve upon.

Methods

Data was derived from two customer data sets. One was a general customer and flight information questionnaire, while the other was a rated satisfaction survey. Target variable is Overall satisfaction; Satisfied or Dissatisfied/Neutral. The two datasets were joined together to perform proper EDA, cleaning, pre-processing and model building.

After careful evaluation and research, I had decided that Decision tree classification was the model most suited for this data set because:

- Target is a binary categorical feature
- Most dataset features are categorical
- There is a non-linear relationship among features

Ensemble methods such as Bagging, Random Forest, AdaBoost etc, using Decision tree criteria will help build even more robust models.

Insights and Recommendations

- The analysis shows that ***seat comfort and leg room*** are big factors that contribute to customer dissatisfaction. Though still present in business class, the EDA shows that dissatisfaction is more prominent in economy class. The company needs to work with partners to restructure their seat designs and possibly find a way to make a little more legroom at the same time.
    - Falcon needs to make sure the new designs are part of every new plane being built.
    - Then work on their current fleet of airplanes at a rate that makes sense financially and rebuild the seating plane by plane.
- ***Inflight entertainment*** is another important feature that needs to be immediately addressed. Similar to what Easyjet and Ryanair are currently doing with Panasonic, Falcon airlines can find an electronics and/or entertainment company to brainstorm different ways to revamp their entire inflight entertainment and follow through.
- ***Food and drink*** showed up in the analysis as needing attention. The airline can conduct another passenger survey to ask what types of food and beverage items they would like to see offered. Do more research and investigate what other airlines and even hotels are offering, then work with their current vendors or find new ones to accomplish their discovered goals.
- Though not apparent in the EDA, ***gate location*** appeared as important in the model. The airline can work with airports, especially their hub airports, to try and relocate to a better, more convenient area.
- Start small, complete tasks and then conduct another sat survey to work on the next set of obstacles.

Data and Description of Features

The dataset was derived from two separate surveys/questionnaires: Flight/Customer data and Survey data. The flight data consisted of basic customer information and flight performance. The second set was a post-flight satisfaction survey that allowed passengers to rate the services they received throughout their entire experience, from booking to baggage claim, as well as their satisfaction overall. The data came from 90,917 random passengers that participated in this study.

Flight/Customer Data

This data provides basic information about the passenger and their flight. The data types initially displayed were object and numerical. The following features were included in this set.

1. Unique ID
2. Gender: Male/Female
3. Customer Type: Loyal/Disloyal
4. Travel Type: Personal/Business
5. Class: Eco/EcoPlus/Business
6. Age
7. Flight distance
8. Departure Delay in minutes
9. Arrival delay, in minutes

Customer Satisfaction Data

This dataset consists mostly of features that were rated by the passenger, ranging from extremely poor to excellent. The variable, overall satisfaction, gave an either-or option: satisfied or dissatisfied/neutral. Gate location rating options were slightly different ranging from extremely inconvenient to very convenient. There were only object datatypes in this dataset and the following features were part of this dataset.

1. Unique ID
2. Satisfaction (Target variable)
3. Seat comfort
4. Departure and arrival time convenience
5. Food and drink
6. Gate location
7. Inflight WIFI
8. Inflight entertainment
9. Online support
10. Ease of online booking

11. Onboard service
12. Legroom
13. Baggage handling
14. Check-in service
15. Cleanliness
16. Boarding

It's important to point out that a few of these variables were not clear and assumptions had to be made. Initially, the feature for legroom showed in the raw data as legroom service. Here we must assume that this simply means legroom and rated was based on how happy the customer was with the amount of legroom they had on the flight.

Online boarding was another vague feature. Was this meant to say inline boarding, the boarding positions they received online when checking in or just boarding in general?

Lastly, Loyal customers was another attribute that is slightly unclear. Does it mean they have booked repeatedly with the airlines, do they have a frequent flyer account with a certain number of accumulated points that automatically put them in that category, or did self-assess?

These uncertainties would need to be clarified with domain experts or data managers at Falcon to properly process and analyze the data.

Combining these two datasets using unique ID and then dropping the ID column will set us in the right direction to begin exploring the data and building models. The merged customer flight data and flight survey data, contains 90917 rows and 23 columns.

Missing data included:

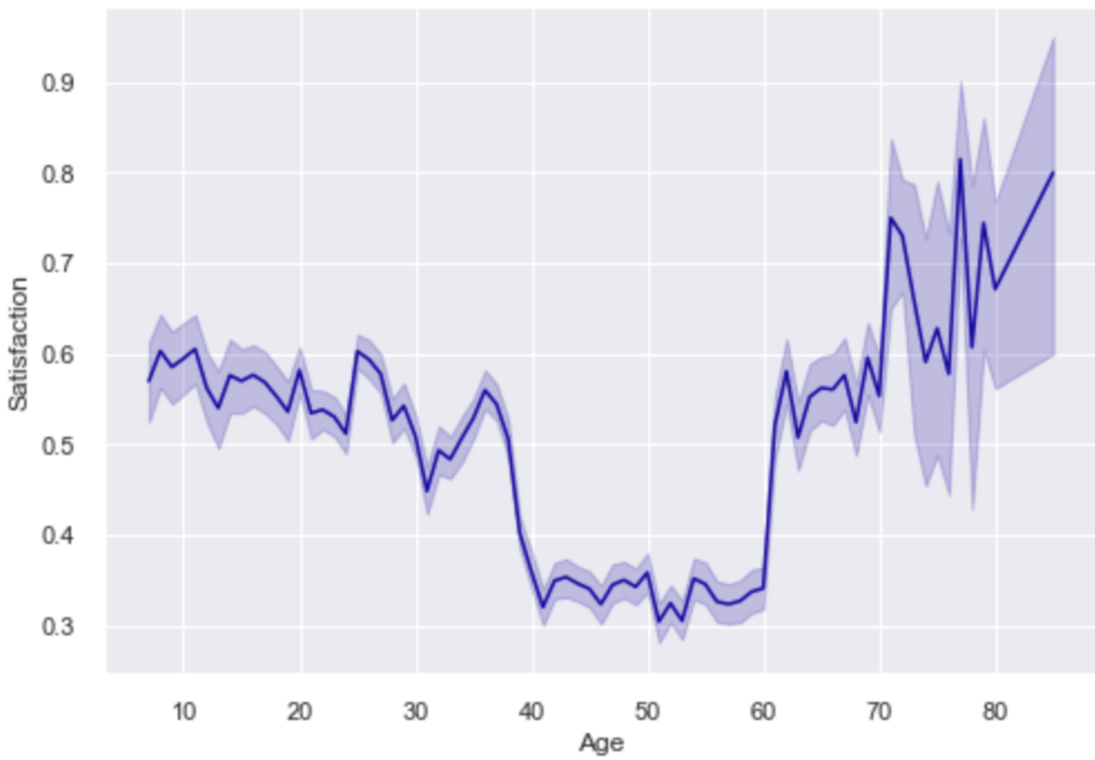
- Customer type-9099
- Travel type-9088
- Depart/arrival time convenience – 8244
- Food/drink-8181
- Onboard service – 7179
- Arrival delay in minutes - 284.

Exploratory Data Analysis

It is extremely important to visually compare features to one another before implementing feature engineering and model building. This way we can see what we need, what we can drop, what needs further investigation and what needs adjusting. This dataset contained many attributes during the EDA that appeared to be significant in contributing toward customer dissatisfaction. Here I will only highlight and summarize a few EDA points that were essential to my end model analytical insights.

The study showed a normal distribution of age with two tall peaks around age 26 and 46. With 1 representing dissatisfied/neutral customers, Figure 1 shows that more dissatisfied customers tend to be under 35 and over 60.

Figure 1: Age related to target



Figures 2 and 3 show that two of the main factors contributing to the mentioned age groups unhappiness are ***Seat comfort*** and ***Inflight entertainment***. It's important to recognize that these two features are important to all age groups, but particularly notice that age groups 36 to 60 show to be overwhelmingly happier when the rating is good and above. The other age groups are split or more dissatisfied when they rate good, and ratings less than that.

Figure 2: Age/Seat comfort satisfaction

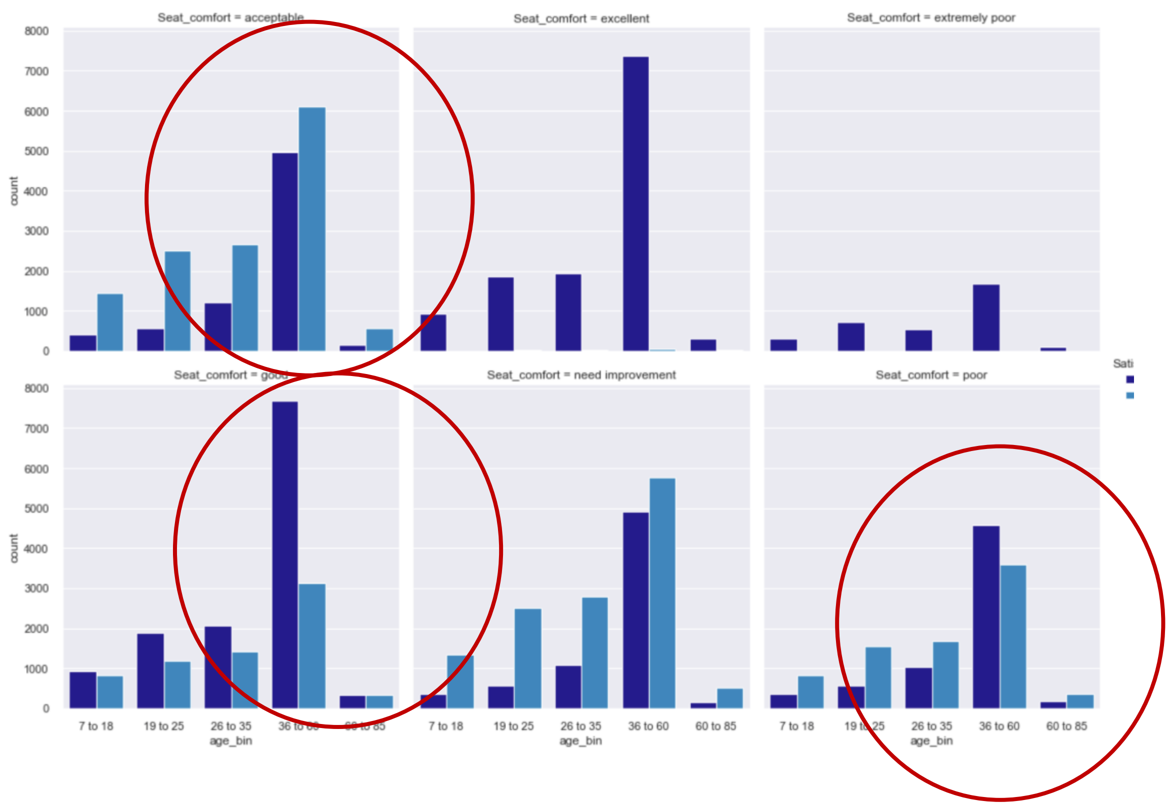
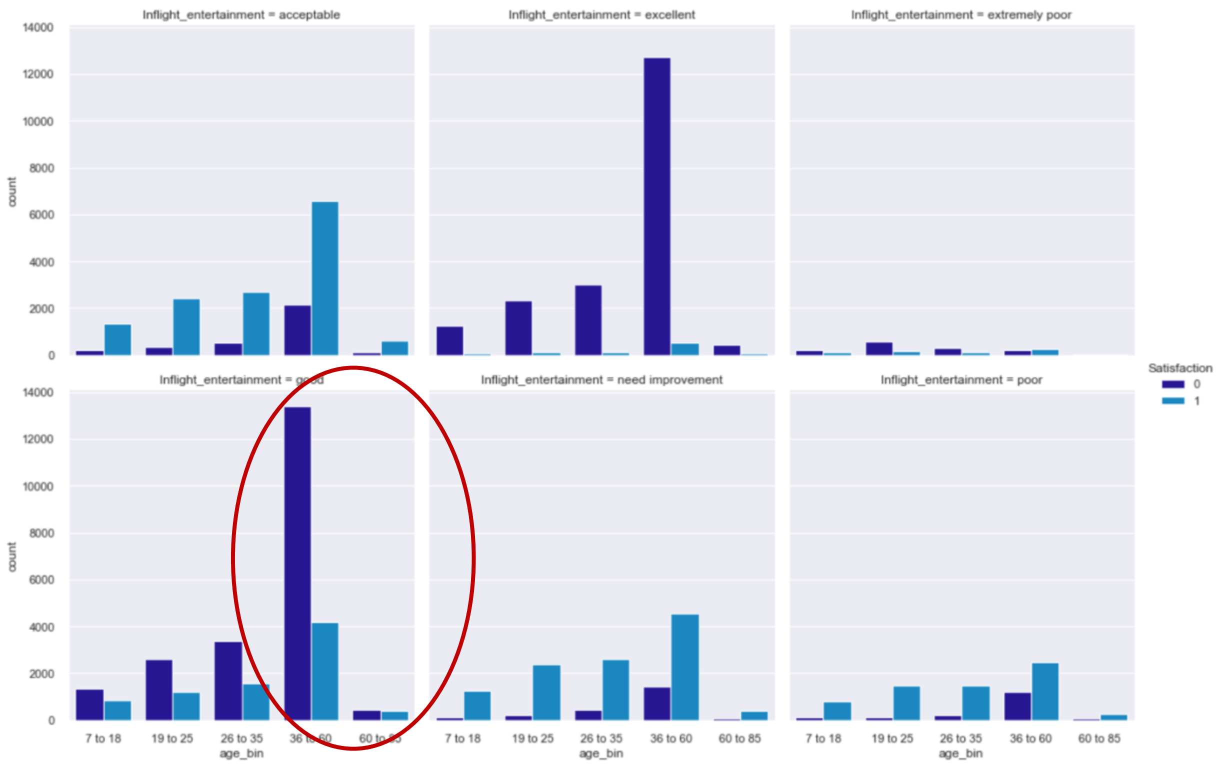


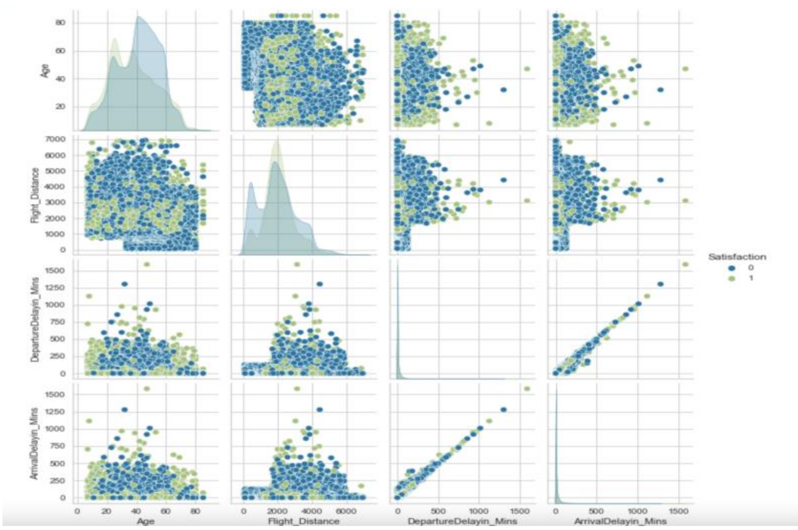
Figure 3: Age/Inflight entertainment satisfaction



Similar outputs were given on other important features such as Legroom, Food and Drink, Gate Location and WIFI.

The heatmap, figure 3, reveals one strong correlation, departure delay and arrival delay. The values shown were given in minutes. Not only was there a strong correlation, but they both had many outliers and arrival delay had quite a few missing values. It made sense to remove the arrival delay attribute from the dataset altogether, especially since we are capturing what we need from departure delay. Otherwise, it would have just contributed to capturing noise.

Figure 4: Correlation heatmap



Modeling Approach

We want to find the passengers that are dissatisfied and find common attributes that they are unhappy with.

- Incorrectly labeling a customer is satisfied when they are not will not allow us to detect the accurate features that need improvement. This will inevitably lead to more unhappy customers who will then look elsewhere for flights, thus loss of revenue.
- There could be a loss if we mislabel a satisfied customer as dissatisfied, because those numbers could contribute to the truly dissatisfied customers which could sway feature importance. Which could mean that money would be spent on improving features that aren't most important.
- We used Recall as the metric to evaluate our model and try to minimize the number of false negatives. The greater the Recall, the lesser the chances of false negatives.
- Though recall is our most important metric, it's important that precision and F1 scores are at an acceptable level as well.
- We built different models to train and test the algorithms - Logistic Regression, Decision Tree Classifier, RandomForest Classifier, Bagging Classifier, AdaBoost Classifier and XGBClassifier.
- We also performed hyperparameter tuning for these models and evaluated their performance using different metrics and confusion matrix.

Decision tree was the preferred model choice but for comparison's sake I wanted to look at how logistic regression would perform. Although the LG model performed well, it still wasn't as robust as other models.

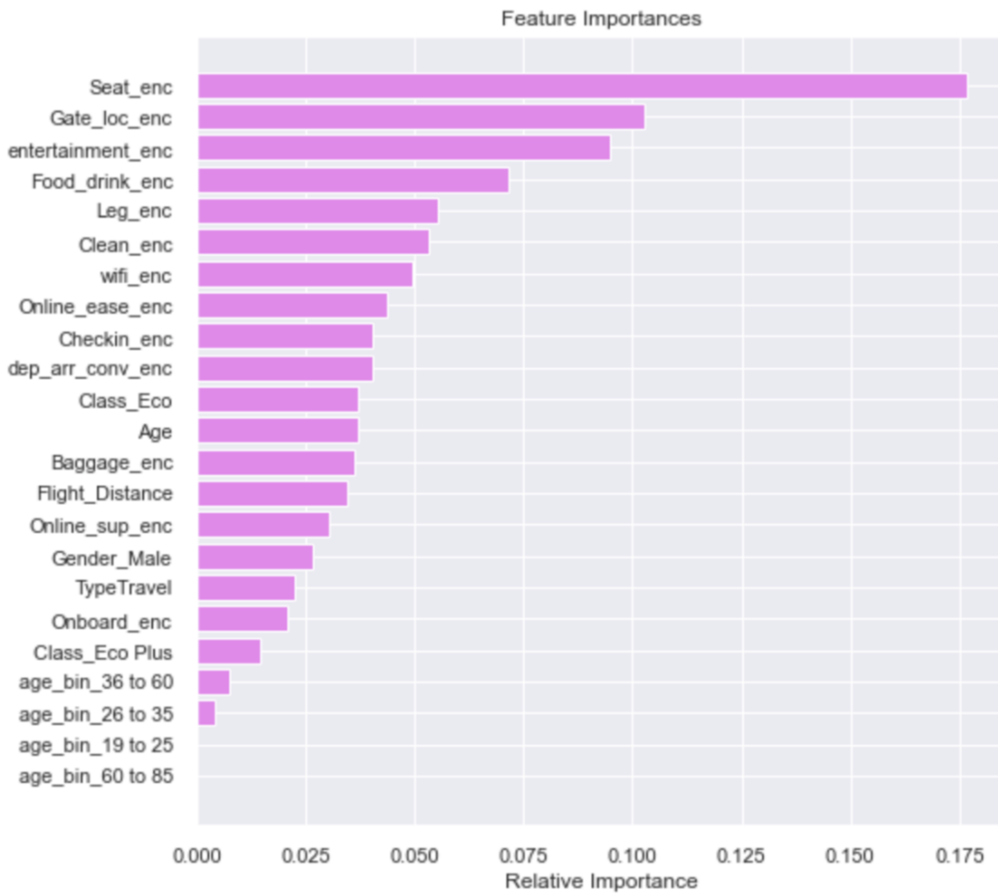
Figure 5: Model comparison

Model	Train_Accuracy	Test_Accuracy	Train_Recall	Test_Recall	Train_Precision	Test_Precision
Decision Tree Tuned	0.900	0.900	0.900	0.890	0.890	0.890
Random Forest Tuned	0.930	0.920	0.920	0.910	0.920	0.920
Bagging Classifier	1.000	0.950	1.000	0.940	1.000	0.940
Ada Boost Tuned	0.930	0.930	0.930	0.920	0.920	0.920
XGBoost Tuned	0.900	0.890	0.990	0.990	0.820	0.820
lg2	0.820	0.822	0.796	0.796	0.805	0.805

- All models performed fairly well, including the logistic regression models
- The Ada Boost model performed great and generalized best, like the random forest model
- Xgboost gave the highest recall with other scores still in a good range
- Even though xgboost has the highest recall, the Adaboost model proved to be the most robust. It generalized best, there was no overfitting, and all scores were very high. Also, the first feature importance didn't completely take over. The feature ratings were distributed somewhat evenly.

Adaboost models feature importance are measured against the target variable, 'satisfaction'. These measurements exhibit an adequate number of top features with the second and third attributes not trailing too far behind the first. Meaning, it's putting a lot of weight into the number one feature but not to the point where its overwhelmingly unbalanced, like all other models.

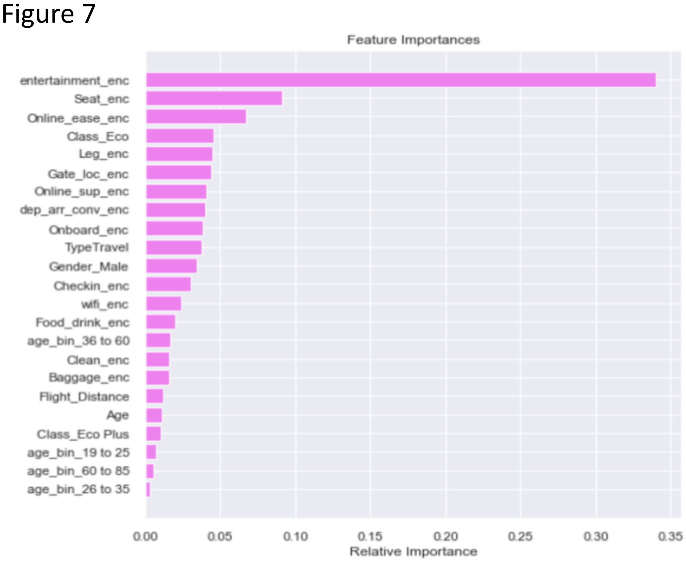
Figure 6: AdaBoost Feature importance



It’s important to recognize all the discoveries during the EDA process even if they don’t prove important on the final model decision because they do add to overall analysis, especially when they show up as significant on other models. For instance, the analysis from our Adaboost model shows that seat comfort and entertainment are among the most important features, while the grouped ages are least important. The EDA shows that customers ages 60 and up as well as 35 and younger are more dissatisfied in general and especially when it comes to these two features. Which is why it is highlighted in the EDA. The 36 to 60 year old age group makes up the fifty-five percent of the survey data and sixty-three percent of them are satisfied, that number can easily go up making the appropriate changes. All other age groups reveal the exact opposite where they are closer to sixty percent dissatisfied/neutral.

Just to expound further, Figure 7 displays the XGBoost model feature importance. As stated previously, this models recall was the highest. Here *economy class* and *ease of online booking* show up in the top five, and that was recognized in the EDA. Here, *entertainment* is number one by a long shot. But either way, *seat comfort, entertainment and legroom* are high ranking features for both models and should be validated.

Figure 7



Since the remaining age groups make up around fifty percent of the data it would be in Falcon’s best interest to do better at making these customers happier. Marketing to retirees who wish to travel could be a large revenue source for Falcon, but they need to make sure they can make that age group content. The same goes for family vacation travel. If the children are happy during the flights, the parents will most likely want to rebook with that airline.

Conclusion

Model feature importance ranked.

1. Seat comfort
2. Gate location
3. Inflight entertainment
4. Food and drink
5. Legroom

Seat comfort and legroom can be addressed at the same time, while redesigning and revamping airplanes. The company needs to work with partners to restructure their seat designs and possibly find a way to make a little more legroom at the same time. They also need to make sure new planes are made with new designs. Old planes need to be on rotation to undergo rehab.

Inflight entertainment is another important feature that needs to be immediately addressed and if it does mean adding personal televisions to every seat, that can be done at the same time as seat/legroom renovation. Another suggestion is to hand out personal devices like kindles or ipads for anyone that doesn't have a personal device to log onto.

As stated in the executive summary part of this report, the airline can conduct another passenger survey, ask what types of food and beverage items they would like to see offered. Implement more research and investigate what other airlines offer, then work with their current vendors or find new ones to accomplish their discovered goals.

Though not apparent in the EDA, gate location showed up as important in the model. The airline needs to work with airports, especially their hub airports, to try and relocate to a better, more convenient area.

All in all, if the airline works on improving a few important areas at a time, they will get happier customers. If the comfort and amenities are present, along with the right pricing, they would become less annoyed by other disadvantages. Start in segments, conduct another satisfaction survey, then work their way through the next set of hurdles.

Appendix

Link to html jupyter notebook/python codes google drive

https://drive.google.com/file/d/1RY5_h5pUT4E89eCUQtwo6G4Tu_YEw4ED/view?usp=sharing