Patricia Miranda
Data Science Springboard Career Track

Final Report:
Airline Passenger Satisfaction Analysis

Problem Statement:

Travel is an important part of our human nature to explore and go on adventures. To experience and enjoy traveling, we have to factor in how we arrive to our destination, which is often through flying in an airplane.

This project aims at showcasing basic data science skills in an attempt to answer the most common business problem: What determines customer satisfaction? With the use of data cleaning, EDA, and Machine learning, several models will be used in order to predict if a passenger was Satisfied or Unsatisfied.

Data Wrangling:

The raw dataset from Kaggle website contained 25893 rows and 25 columns. The majority of the dataset was usable as it did not contain many null values. The only null values were contained in the column “Arrival Delay in Minutes” and attributed to 310 values. These values were dropped and no other missing values were left in the dataset.

Target Client(s):

My client would be an Airline company. Although this dataset is based on an anonymous Airline company, any company can use the models to predict Passenger Satisfaction, since many airlines have similar metrics.

Data Collection:

My dataset is included as a CSV file, but can be accessed on Kaggle through the following link:

Airline Passenger Data

Questions to Answer:

-What are the different types of classes when flying and do they affect passenger satisfaction?

-What is the age range of passengers? Is there an age group that responds differently to their airline experience and has a different satisfaction?

-What is the overall satisfaction result (Satisfaction, neutral or dissatisfaction)?

-Based on data of airline, can satisfaction be predicted?