



Predict Ariline Customer Satisfaction

Brief

Most of airline organizations seeking for his customers satisfaction but reaching to the customer satisfaction depends on many factors. When the airline organizations recognize these factors, they will produce their services as the customers preference. So, they will reach customer 's satisfaction and gain their loyalty. In this project data associated with factors keep customer satisfaction will be analyzed. Particularly the factors that affect the customer satisfaction will be identified.

Dataset Description

The dataset collects information from an airline passenger satisfaction survey. The survey done by some airline organization on their customers who have already flown with them. The feedback of the customers on various context and their flight data has been consolidated.

This dataset contains 129880 surveys feedback from airline customers with their detail information based on 23 factors the organization want to focus on it .

The main purpose of this dataset is to predict whether a future customer would be satisfied with their service given the details of the other parameters values.Also, What factors are highly correlated to a satisfied (or dissatisfied) passenger? Can you predict passenger satisfaction?. So, the airline organizations have to be emphasized more to generate more satisfied customers.

Project Target

What is the important factor of services offered by airline organization have to be emphasize to generate more satisfied customers? Create a model to predict whether a future customer would be satisfied with their service given the details of the other parameters values?

Data Definition

- Satisfaction: Airline satisfaction level (Satisfaction, dissatisfaction)
- Gender : Gender of the passengers (Female, Male)
- Customer Type : The customer type (Loyal customer, disloyal customer)
- Age : The actual age of the passengers
- Type of Travel : Purpose of the flight of the passengers (Personal Travel, Business Travel)
- Class : Travel class in the plane of the passengers (Business, Eco, Eco Plus)
- Flight Distance : The flight distance of this journey
- Seat comfort : Satisfaction level of Seat comfort (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Departure/Arrival time convenient : Satisfaction level of Departure/Arrival time convenient (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Food and drink : Satisfaction level of Food and drink (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Gate location : Satisfaction level of Gate location (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Inflight wifi service : Satisfaction level of the inflight wifi service (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Inflight entertainment : Satisfaction level of inflight entertainment (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Online support : Satisfaction level of online support ((0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Ease of Online booking : Satisfaction level of online booking (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- On-board service : Satisfaction level of On-board service (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Leg room service : Satisfaction level of Leg room service (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Baggage handling : Satisfaction level of baggage handling (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Checkin service : Satisfaction level of Check-in service (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Cleanliness : Satisfaction level of Cleanliness (0:Not Applicable;1:Least Satisfied ;2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)

- Online boarding : Satisfaction level of online boarding (0:Not Applicable;1:Least Satisfied ; 2:Partly Satisfied;;3:neutral ;4:Satisfied; 5: is Most Satisfied)
- Departure Delay in Minutes : Minutes delayed when departure
- Arrival Delay in Minutes : Minutes delayed when Arrival

Data Cleaning

Data cleaning is the process of identifying incomplete, incorrect, inaccurate or irrelevant parts of the data and then replacing, modifying, or deleting the dirty data.

In the data set we found some missing data:

- Edit the 'Arrival Delay in Minutes'column datatype to be integer.
- Delete rows which has null value in 'Arrival Delay in Minutes' column.

Algorithms

- Classification

Questions for Analysis

This analysis aims to outline some most factors which makes airline passenger more satisfaction,as well as makes the airline organizations more awareness about his customer preferences .

To accomplish this, we will first try to understand the dataset and the context on which they were collected. Then we will exploratory data analysis.

There is some questions needs to investigate:

- What is the percentage of customer who is satisfied vs. who is dissatisfied ?
- Is one gender more satisfied than other gender?
- what is the customer age range are more satisfied?
- Is the customer type (loyal, disloyal) effect the satisfaction?
- What is the type of traveler more satisfied?
- What is the most factor make customer satisfaction?

Tools

- Pandas: a library offers data structures and operations for manipulating numerical tables and time series.
- Numpy: a library used for working with arrays. It also has functions for working in domain of linear algebra, fourier transform, and matrices.
- Matplotlib: a plotting library for the Python programming language and its numerical mathematics extension NumPy.
- Seaborn: a data visualization library built on top of matplotlib and closely integrated with pandas data structures in Python .