greatlearning

Capstone Project Presentation

Flight Cost Prediction

Business Problem Understanding

Business problem :

The Business problem is to predict the prices according to the given variables in the dataset

Constraints :

Some of the variables were missing which could add impact on price prediction :

- Profits
- Date of booking
- Number of persons booking
- Gender of the passenger
- No. of clicks before actually booking

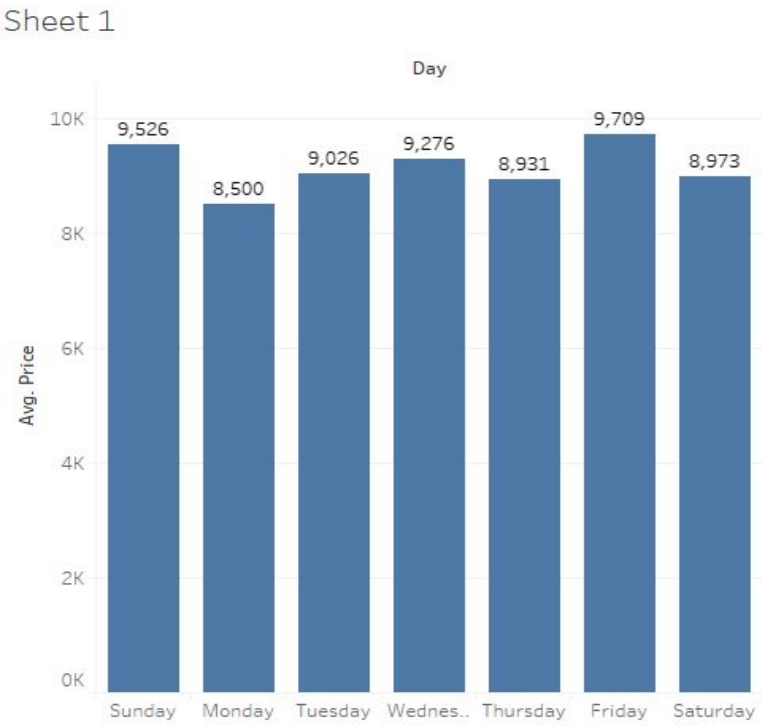
- Salary of the passenger
- Job or student
- Age
- Passenger name or ID
- flight comfort for customer (yes/no) .
- Some of the data points were given incorrect
- Prices were not given the test dataset .

Scope :

Hypothesis test

a. flight prices during weekends are more costlier than weekdays

b. flight prices during daytime or peak hours are costlier i.e 9 A.M - 9 P.M



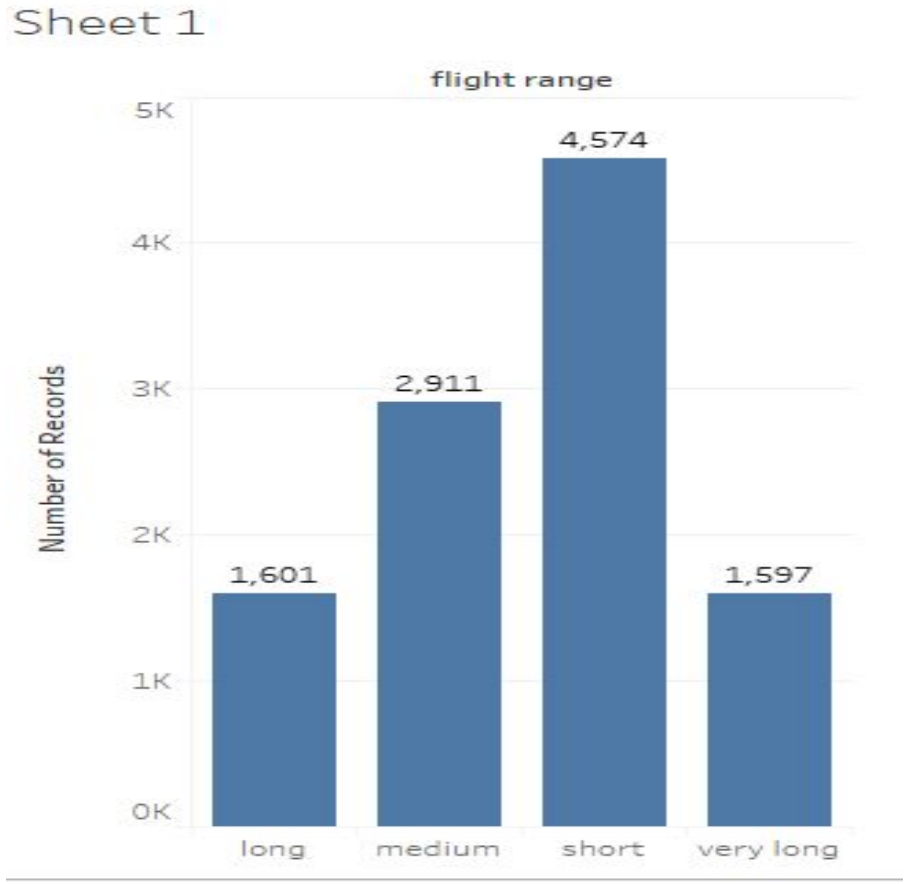
Variables test

The impact of some of the variables given in the dataset

Airlines

Flight Name	total passengers
Air Asia	319
Air India	**1753**
GoAir	194
Indigo	**2053**
Jet Airways	**3855**
Multiple carr..	1209
Spicejet	818
Vistara	482

Duration




```
   Duration
Min.    :  75
1st Qu.: 170
Median : 520
Mean   : 643
3rd Qu.: 930
Max.   :2860
```

greatlearning

Objectives :

1. To predict Prices of different airlines under given variables at different circumstances .
2. To build and run models accordingly .
3. To predict and increase efficiency of the models .

Modelling Approach Used & why

Performance table(regression models)

Model	Performance measure	Value	Rank
GLM	RMSE Multiple R-squared	1239.304 79.33	2
CART	RMSE MAPE value	2021.95 0.21	3
Random Forest	RMSE Mape value	146.8 0.004	1

These values were obtained from the final dataset with adjusted and confirmed variables .Also the models were tuned to get better and accurate results .

Insights from Analysis

- Understanding the insights from EDA (summarize)
- Flight prices during weekends are costlier than weekdays .
- Flight prices at peak hours(9 A.M-9P.M) is more .
- The details regarding flights were minimised.
- Some of the constraints were missing which could add value .
- Some of the variables were useless .
- From a business point of view we don’t have any info regarding customers .
- Duration and total stops are important factors to consider.
- Duration is directly proportional to price
- Arrival time has been important factor than departure .

Recommendations

Business side :

1. We can understand that people prefer weekends, so we can increase in prices a bit high .
2. People prefer day time flights mainly due to :
   (i) security issues
   (ii) To relax for the next day
   (iii) Transportation issues at night.
   Either way we could increase price

3. If we'd had details about passengers then we could gain better understanding of the urgency and importance of the customer .

4. To differentiate b/w primary and secondary customers .

5. For 2 stops or more duration the customers might not prefer the airline ,rather take a short flight so could reduce the price a bit .

Customer side :

1. To keep the prime customers with our clients by providing
   (i) good service
   (ii) good pricing & offers
   (iii) regular feedbacks
   (iv) Make the customer suggest the airline .

2. One time during day time (12 noon- 3 P.M) the traffic is less,could probably introduce lunch with less price along with Price of the flight.

3. Also to increase night time flights the airline company could provide a taxi service at nights with a bit less price.Also gender is a factor here .

4. Make sure the arrival time is not delayed by much .