**AIRFARE FLUCTUATION PREDICTION MODEL**

**Project Proposal:**

Using Predictive analytics to build a model to forecast and to understand the reasons of airfare fluctuation throughout the given period for domestic flights & domestic travel within India.

**Business: Airline Industry Domestic Travel within India**

It is in airlines’ best interest to make sure people can afford to fly with them, so they want to offer good prices. But they also need to find the right balance between (a) making sure they fill up a flight and (b) earning enough money to justify all the costs that go into operating the journey.

For this reason, ticket prices will drop and rise very regularly, depending on many different factors that are specific to that particular flight. Those things can include:

Of course, this depends on a ton of different and ever-changing factors, like how well that particular flight or route is selling.

**Business Problem Statement: Airline Industry Domestic Travel within India**

Flight ticket prices can be something hard to guess, today we might see a price, check out the price of the same flight tomorrow, it will be a different story. We might have often heard travelers saying that flight ticket prices are so unpredictable. Here you will be provided with prices of flight tickets for various airlines between the months of March and June of 2019 and between various cities

**Applied Business case for the Problem Statement:**

Our analysis shows that flight prices go through a weekly & Monthly cycle . Typically, the lowest prices are made available earlier in the week, and the highest prices are offered later in the week.

Midweek. Flights that take off and land on weekends, or Mondays and Fridays generally cost more. So aim for Tuesdays, Wednesdays and Thursdays.

Prices are based on demand and availability. The aim of each airline is to fly with fully loaded airplanes and have sold each seat as expensive as possible. So, they cannot overdo it (in price) as this would mean they would fly empty. Airlines have so-called Yield Management where the loading and the pricing of each flight and each seat are reviewed on a regular basis. They can lock back at past years to see how the situation was on this particular day or period one, two, or three years ago.

If the plane is half empty, they will release a few cheaper seats, if the plane is rather full, they will take the cheaper seats out of the sale and show only more expensive fares hoping to find passengers willing to pay more.

This Yield Management is done mainly by complex computer programs with some manual interventions by Yield-Management specialists reviewing at a regular basis the load factors and the revenue or profit generated by each flight.

Airline ticket prices fluctuate because of yield management, computer programs that aim to maximize the revenue an airline gets from a seat. If you look very closely at airline schedules you may notice subtle capacity changes from week to week or month to month.

**Source:**

[**https://machinehack.com/hackathon/predict\_the\_flight\_ticket\_price\_hackathon/overview**](https://machinehack.com/hackathon/predict_the_flight_ticket_price_hackathon/overview)

**Data Dictionary:**

Data has 10683 rows & 11 columns and collected

<class 'pandas. core. frame.DataFrame'>

Range Index: 10683 entries, 0 to 10682

Data columns (total 11 columns):

# Column Non-Null Count Dtype

--- ------ -------------- -----

0 Airline 10683 non-null object

1 Date\_of\_Journey 10683 non-null datetime64[ns]

2 Source 10683 non-null object

3 Destination 10683 non-null object

4 Route 10682 non-null object

5 Dep\_Time 10683 non-null object

6 Arrival\_Time 10683 non-null object

7 Duration 10683 non-null object

8 Total\_Stops 10682 non-null object

9 Additional\_Info 10683 non-null object

10 Price 10683 non-null int64

dtypes: datetime64[ns](1), int64(1), object(9)

memory usage: 918.2+ KB

**Details in the Columns:**

**0 Airline - String /Categorical**

* Name of all the Airlines is given

1. **Date\_of\_Journey -Datetime /Numeric**

* Date of Journey is or date of arrival is mentioned.

**2 Source -String / Categorical**

* Departure City is mentioned.

**3 Destination -String / Categorical**

* Arrival City is mentioned.

**4 Route -String/Categorical**

* Route of the flight is mentioned.

**5 Dep\_Time -String/-Datetime /Numeric**

* Departure time of the flight is mentioned.

**6 Arrival\_Time - String/-Datetime /Numeric**

* Arrival time of the flight is mentioned.

**7 Duration - String /Numerical**

* Travel duration of the flight

**8 Total\_Stops -Str /Numerical**

* Stopover of the flights

**9 Additional\_-Str/ Categorical**

* IF Meal/ Baggage is included or not

**10 Price -Int/Numeric**

* Airfare price
* Factors affecting

EDA question?

1. Which Airline has the maximum and minimum travel?
2. Which City has the maximum and minimum departures?
3. Which City has the maximum and minimum Arrivals?
4. Which month has the maximum/minimum departures?
5. What is the highest & lowest Ticket price ranges fall in?
6. On which day in a week has the maximum/minimum departure booking?
7. What is the minimum & maximum duration of the travel?
8. Month-wise is Revenue Maximum & minimum?
9. On which day of the Week is Revenue Maximum & minimum?
10. Maximum & minimum duration of Flights?
11. Which Airline gives maximum and Minimum Revenue
12. At which hour of day maximum/ minimum departures are there?
13. At which hour of day maximum/ minimum arrival are there?
14. From which city, which airline departs most?

**Insights:**

1. Jet Airways has the Maximum travel with a count of 3849 throughout the season in the given data.

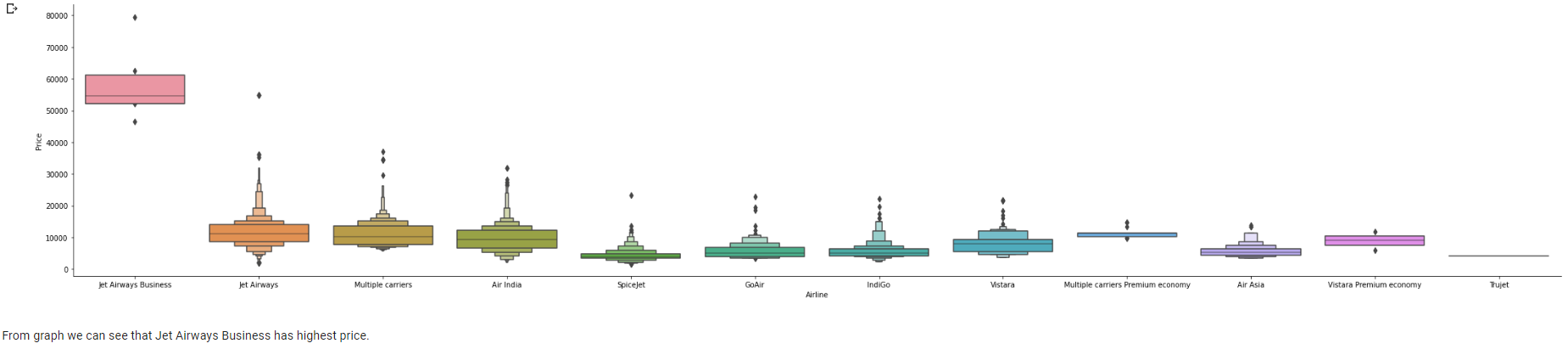
2. Top 5 Airlines with the highest count of Travel as below

|  |  |  |
| --- | --- | --- |
| **Sr.no** | **Airlines** | **No of Trips** |
| 1 | Jet Airways | 3849 |
| 2 | IndiGo | 2053 |
| 3 | Air India | 1752 |
| 4 | Multiple carriers | 1196 |
| 5 | SpiceJet | 818 |

3. Airline with the lowest no of Trips throughout the season

|  |  |  |
| --- | --- | --- |
| **Sr.no** | **Airlines** | **No of Trips** |
| 1 | GoAir | 194 |
| 2 | Multiple carriers Premium economy | 13 |
| 3 | Jet Airways Business | 6 |
| 4 | Vistara Premium economy | 3 |
| 5 | Trujet | 1 |

4. Delhi is the departure city with the highest count of departures which is 4537 trips then follows by Kolkata =2871, Bangalore=2197, Mumbai=697, Chennai=381.

5. Airline VS Price:

From the graph, we can see that Jet Airways Business has the highest price

6. Cochin is the city with the highest count of Arrivals with 4537 then follows Bangalore =2871, Delhi= 2197, Hyderabad =697, Kolkata=381

7. In Weekday wise stats it shows passengers had traveled most on Wednesday with a count of 2079 follows Monday 1850, Thursday 1832

8. Least Travel on weekdays was on Tuesday & Friday

9. In Month wise bifurcation In the month of May (count 3466) highest num of Passengers were booked then follows June(count 3414)

10. In the Month of April lowest passengers were booked (count 1079) followed by March(2724)

11. Month of May generated the highest revenue of INR 3,16,35,040 & April gave the b lowest INR 6226744

12. Flights with the highest Revenue is Jet Airways INR 44817461 then follows Air India INR 16838841, Multiple carriers INR 13039603, IndiGo INR 11648071.

13. Flights with lowest Revenue Trujet INR 4140 then follows Vistara Premium economy INR 26887, Multiple carriers’ Premium economy INR 148445

14. As Wednesday we have maximum departures it gives the highest revenue INR 19287952 follows Thursday INR 16362287 & Tuesday with Lowest Revenue INR 7843724 follows Friday INR 8833659

15. Flights are with a stopover and 1 stopover gives the highest Revenue of INR 59591945 follows 2 stop over INR 19328028, 0 stop over generates INR 17541927 and lowest revenue was generated by 4 stop over INR 17686 & 3 stop over INR 590040