Plane Ticket Price Analysis

Sreeja Choudhury

Introduction

An airline is a company that provides air transport services for traveling passengers and freight.

Airlines assign prices to their services in an attempt to maximize profitability. The price of an Airline Ticket is affected by a number of factors, such as flight duration, days left for departure, arrival time and departure time etc. Airline organizations may diminish the cost at the time they need to build the market and at the time when the tickets are less accessible. They may maximize the costs. The price may rely upon different factors. Each factor has its own proprietary rules and algorithms to set the price accordingly.

This analysis explores a dataset containing information about plane tickets. Understanding this dataset can provide valuable insights into the airline industry, including trends in pricing, popular flight routes, and the impact of departure times on ticket costs.

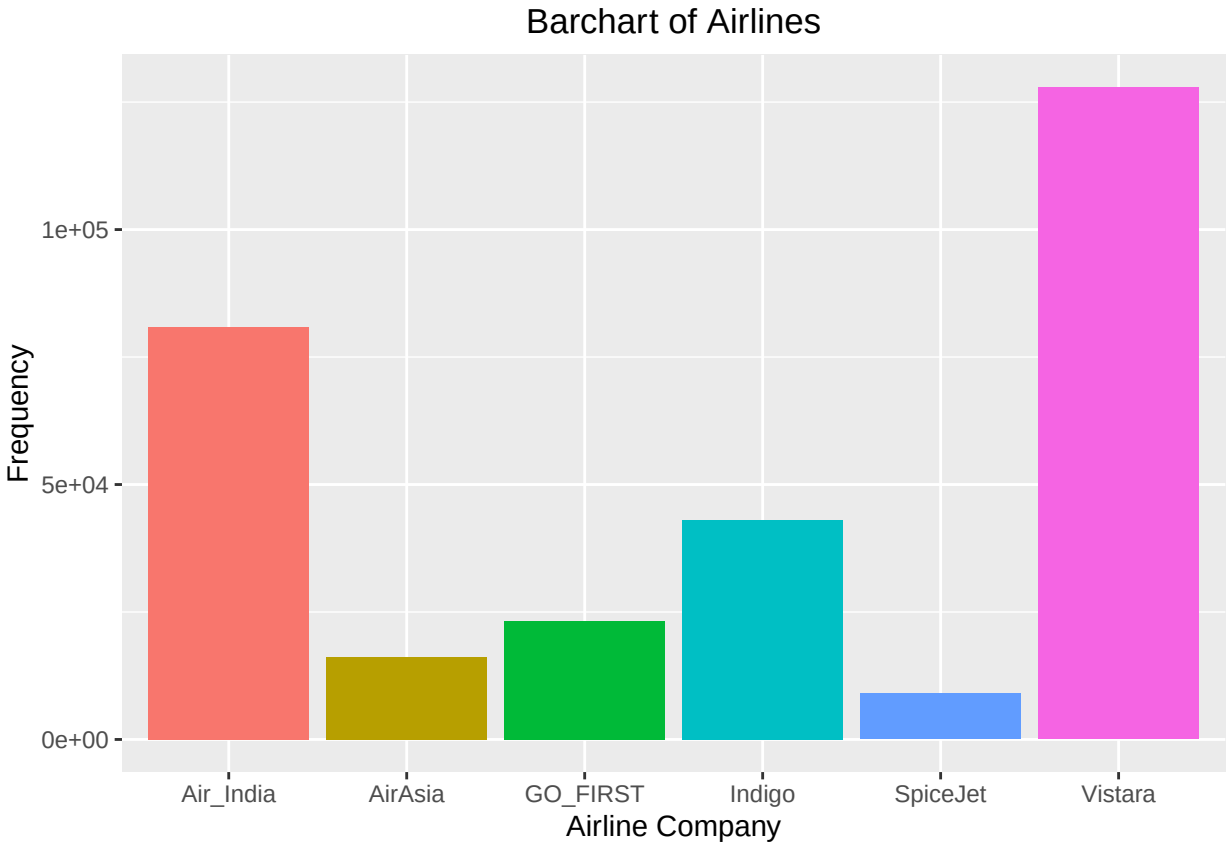
Data Description

Dataset contains information about flight booking options from the website Easemytrip for flight travel between India's top 6 metro cities. There are 300261 datapoints and 11 features in the cleaned dataset.

The various features of the cleaned dataset are explained below:
1.**Airline**: The name of the airline company is stored in the airline column. It is a categorical feature having 6 different airlines.
2.**Flight**: Flight stores information regarding the plane's flight code. It is a categorical feature.
3.**Source city**: City from which the flight takes off. It is a categorical feature having 6 unique cities.
4.**Departure time**: This is a derived categorical feature obtained created by grouping time periods into bins. It stores information about the departure time and have 6 unique time labels.
5.**Stops**: A categorical feature with 3 distinct values that stores the number of stops between the source and destination cities.
6.**Arrival time**: This is a derived categorical feature created by grouping time intervals into bins. It has six distinct time labels and keeps information about the arrival time.
7.**Destination city**: City where the flight will land. It is a categorical feature having 6 unique cities.
8.**Class**: A categorical feature that contains information on seat class; it has two distinct values: Business and Economy.
9.**Duration**: A continuous feature that displays the overall amount of time it takes to travel between cities in hours.
10.**Days left**: This is a derived characteristic that is calculated by subtracting the trip date by the booking date.
11.**Price**: Target variable stores information of the ticket price.

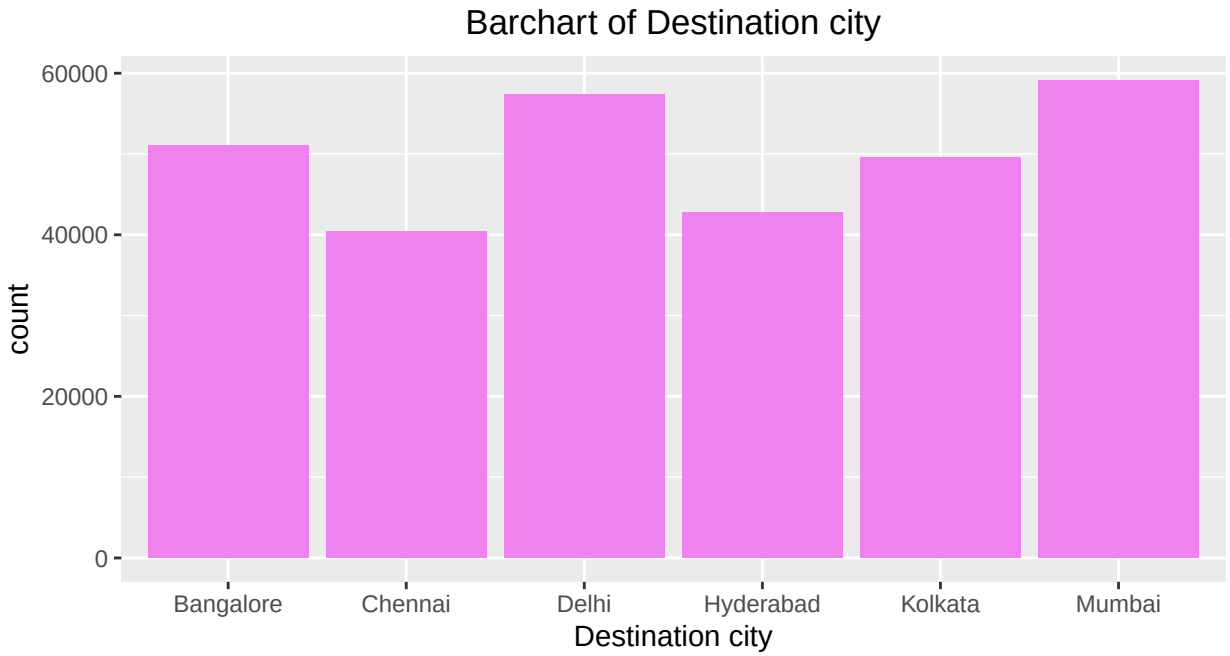
To see the number of flights operating per airline company

There are airline companies present in the dataset - SpiceJet, Air Asia, Vistara, Go First, Indigo and Air India.

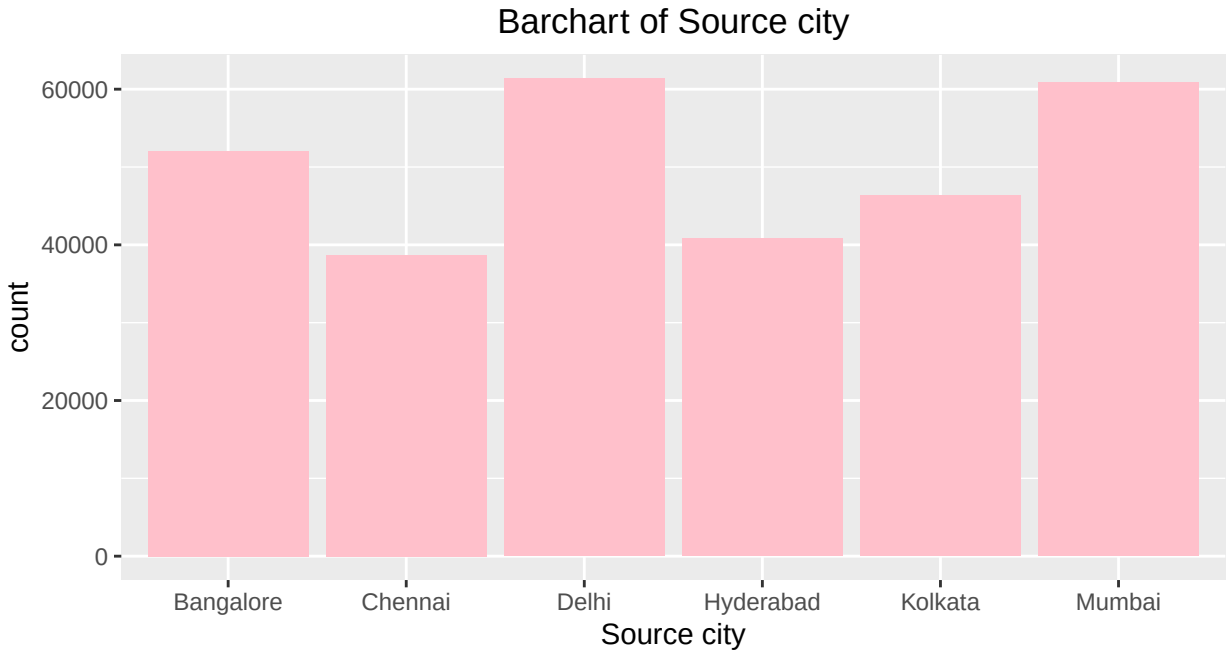


The height of each bar represents the number of flights operated by a particular airline. Here Vistara has the maximum number of flights while Spicejet has the least. This suggests that Vistara has a substantial presence in the air travel market and is preferred choices for travelers. Air India also has a substantial number of flights operating.

Distribution of planes in different Destination and Source cities

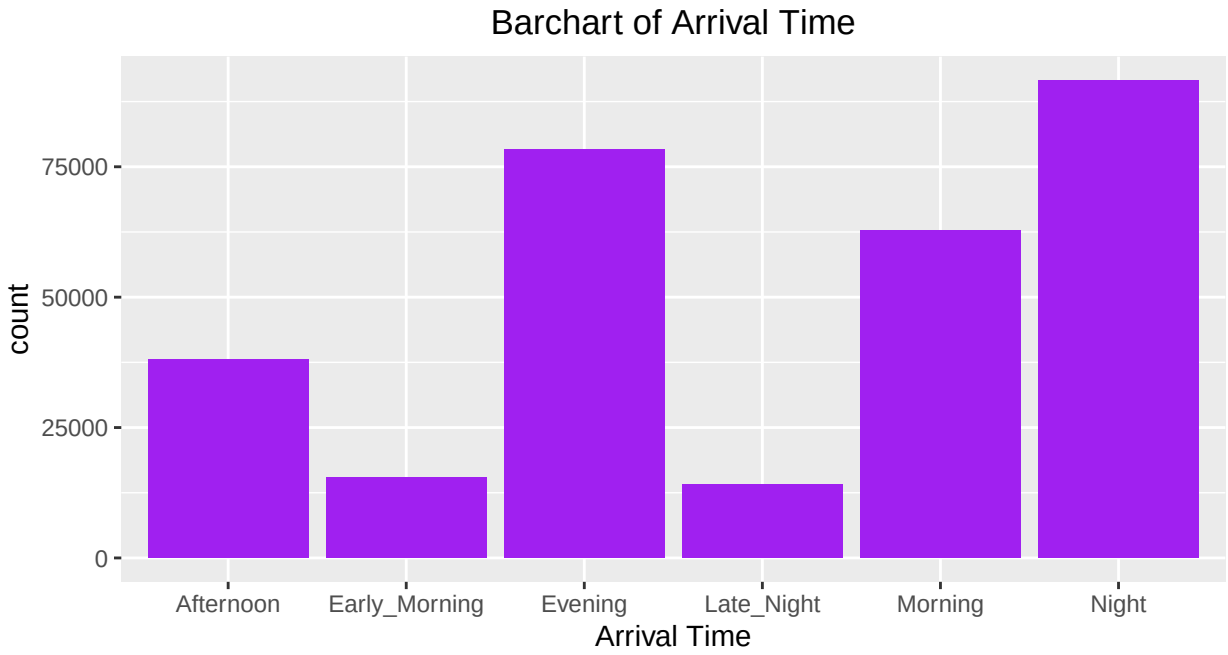


The bar chart clearly shows the distribution of travelers across different destinations. Mumbai and Delhi appear to be highly favored, as indicated by their higher frequency of ticket bookings. These popular destinations are likely to be top travel choices for a significant portion of passengers or can be a business hub. Chennai and Hyderabad are comparatively less visited by people.

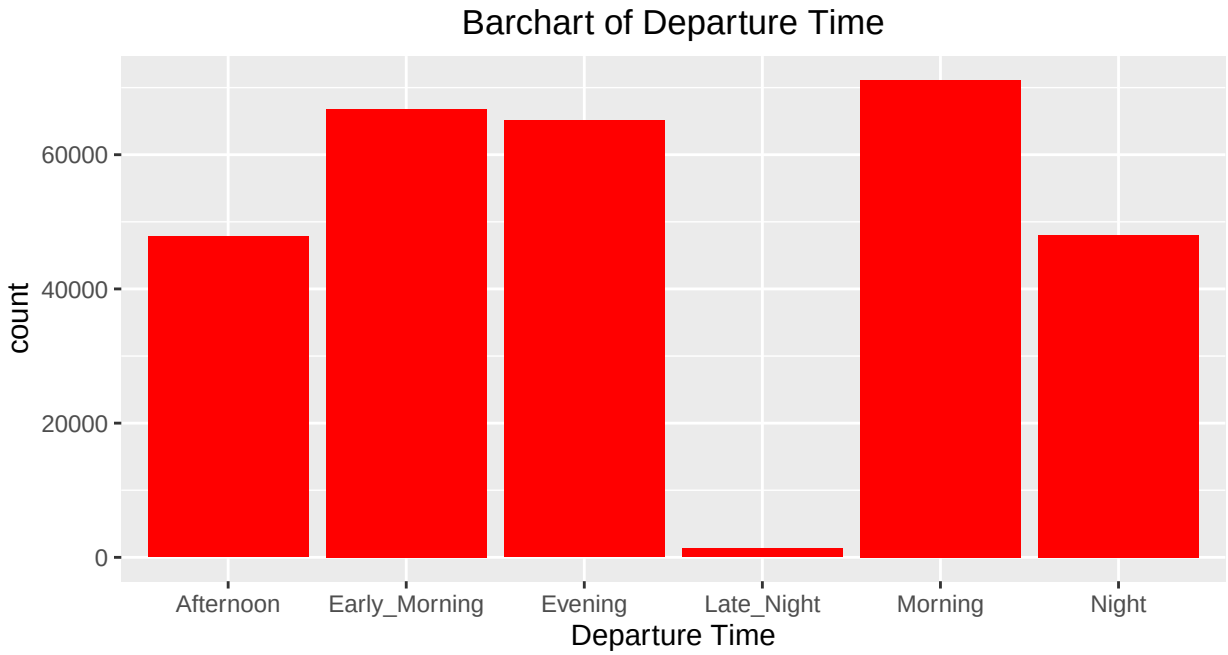


Again, Delhi and Mumbai stand out as major departure hubs. These cities are consistently chosen as the starting point for air travel by a significant number of passengers. Chennai and Hyderabad has less number of flights taking off.

Number of flights at different arrival times and departure times

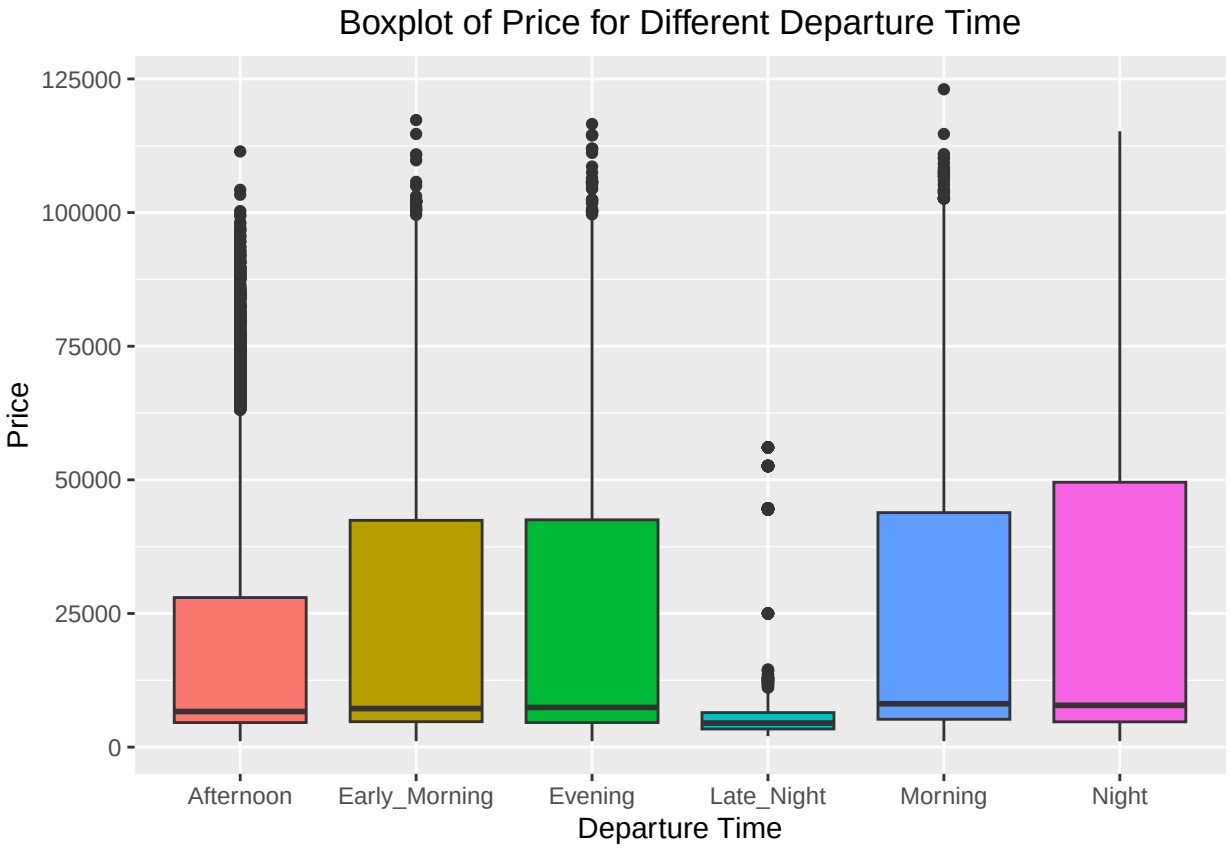


The chart clearly illustrates peak arrival times.There are two prominent peaks: one in the night and another in the evening. The morning peak often consists of arrivals from overnight flights, while the evening peak and night peak accommodates travelers returning from business trips or vacations.



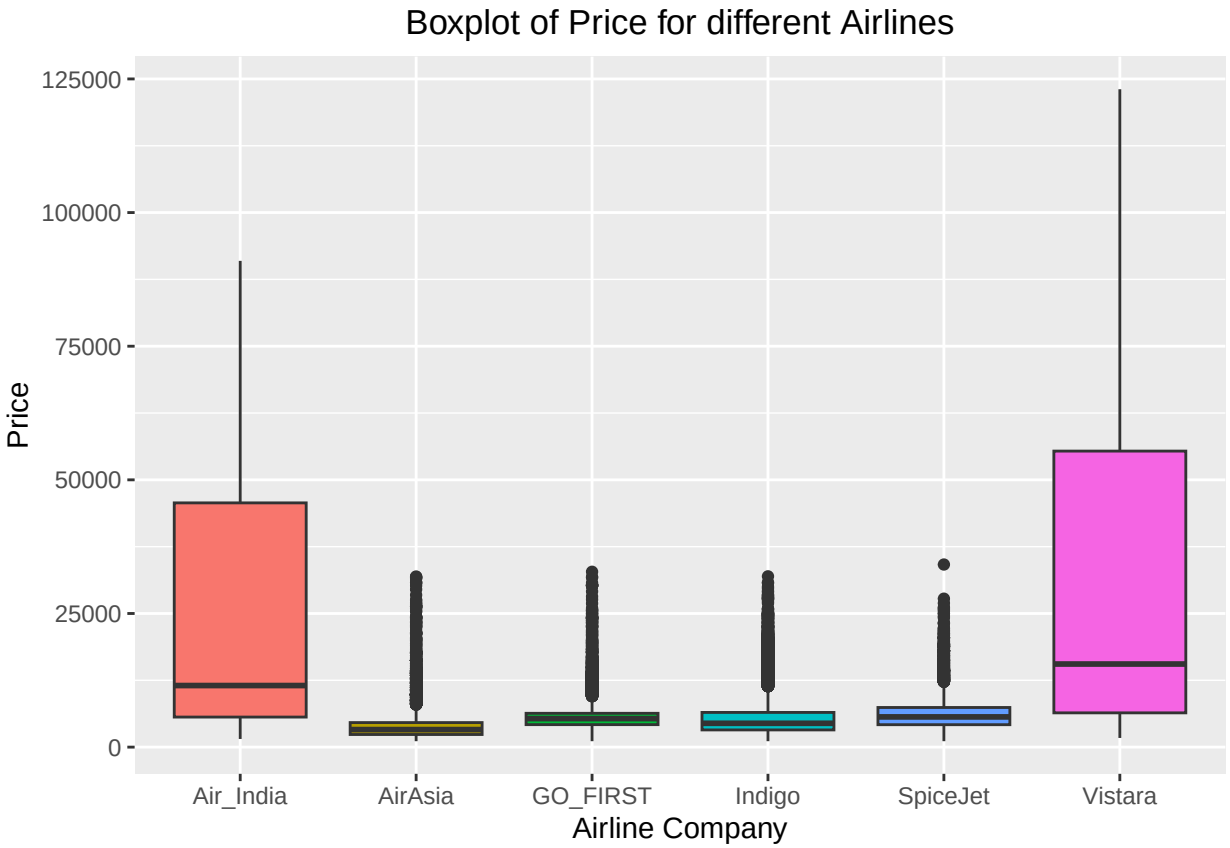
The chart clearly identifies peak departure times when the highest number of flights take off. The high peaks in this graph corresponds to early morning, morning and evening hours, aligning with the travel patterns of both business and leisure travelers. We also observe that very less number of flights take off at late night.

To see how the price changes with change in Departure time



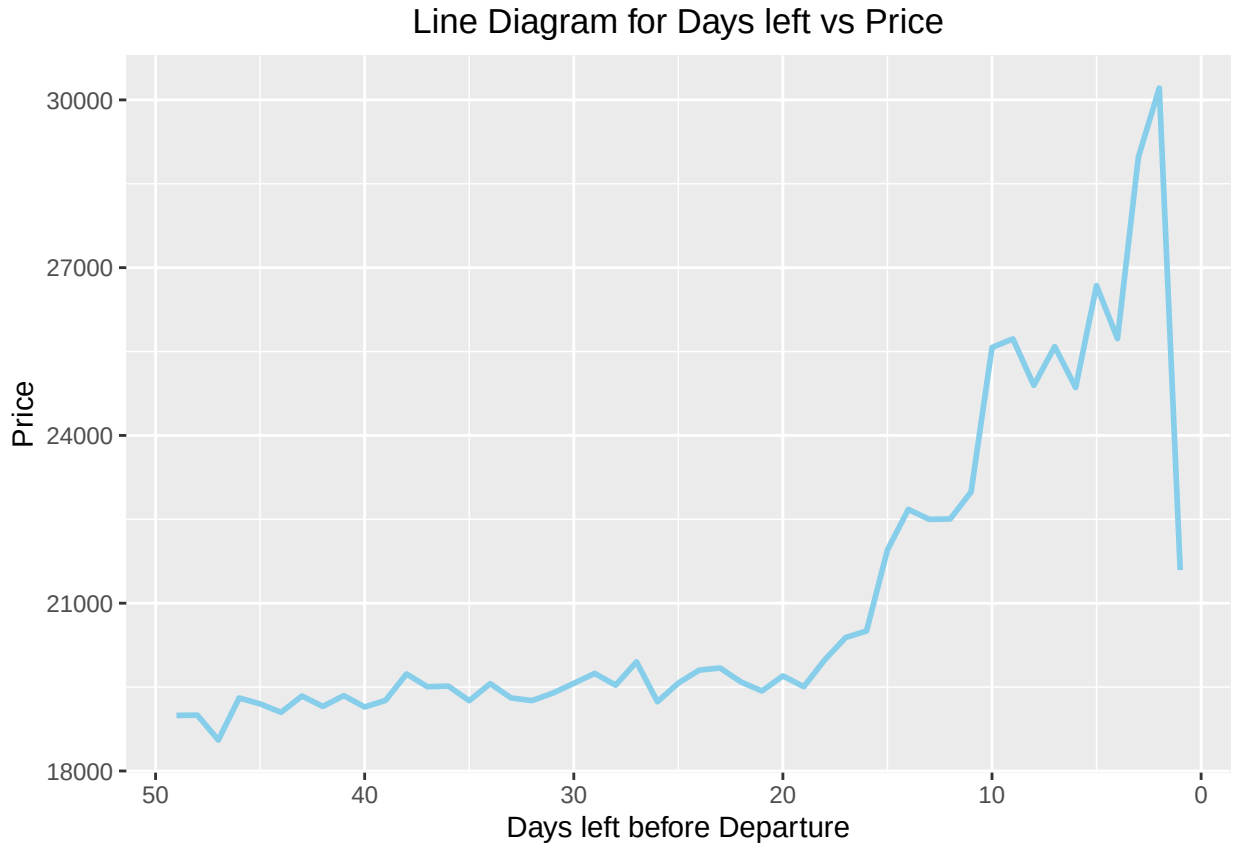
The median price for flights is almost same for every departure time, except for the late night flights, which is slightly less. There is a wider variability in price for the night flights but very less variability for the late night plane ticket prices, although there are some outlines. Travelers departing during peak times, like morning or evening may encounter higher ticket prices, while those departing during off-peak hours, i.e. at late night, tend to enjoy more competitive fares.

To see how price varies among different airlines



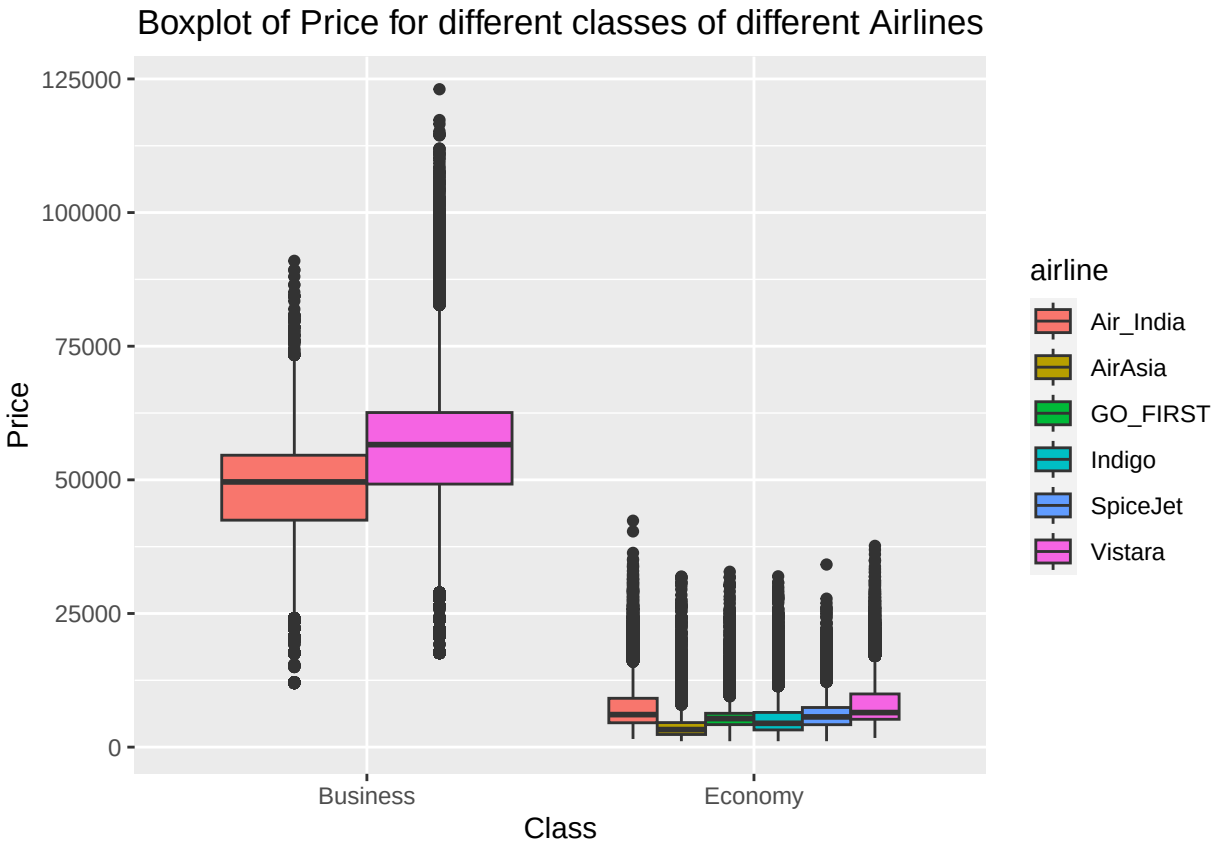
Vistara and Air India have a wider variability in the price range. This may be due to the fact that both these airlines have economy and business class. For this reason, the median price of these two airlines is slightly higher than the others. The other airlines have very less variability, having a large number of outliers. These outliers may be associated with specific factors such as last-minute bookings.

To see how is the price affected when tickets are bought in just 2-5 days before departure



We can observe that as the days left before departure decreases, the price of flight increases.This is so because as days before departure decreases, number of seats available in a certain plane decreases, hence its demand increases. The ticket price bought in just 2-5 days before departure is very high.

To see how ticket price varies between Economy and Business class



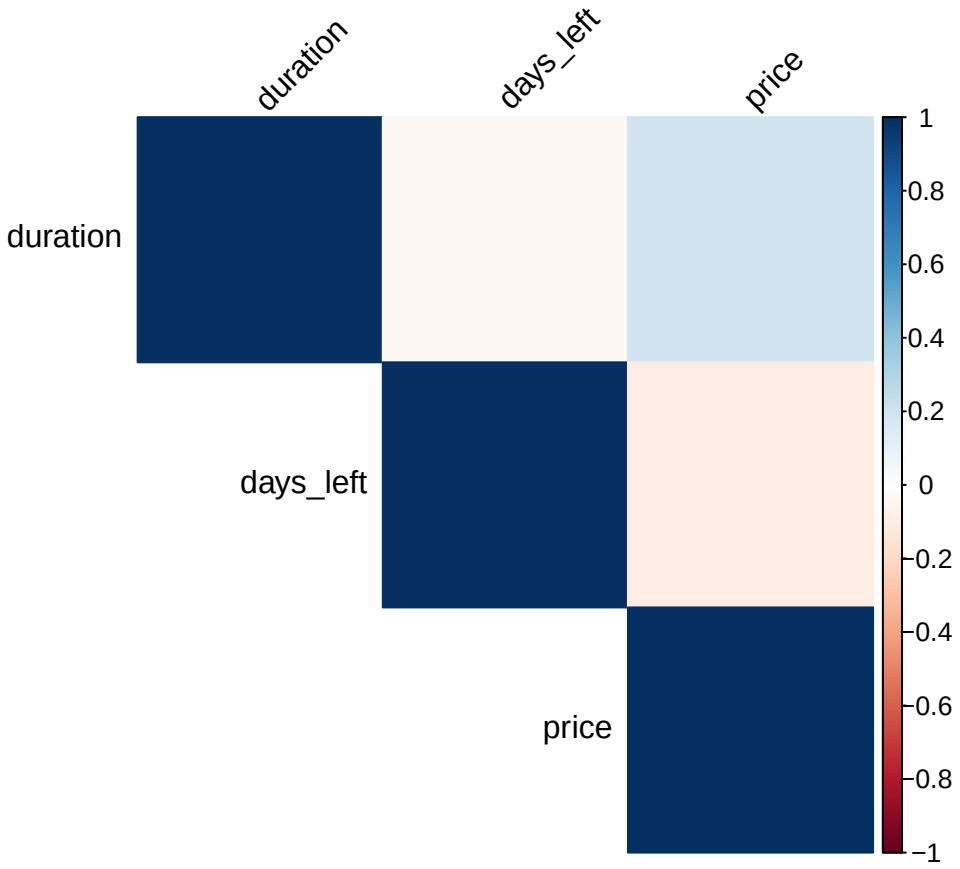
Only Vistara and Air India has Business class. The median ticket price for this class is quite high. All the six airlines provide Economy class, with a low median ticket price. There are outliers in all the situations.

Correlation between duration, days left and departure

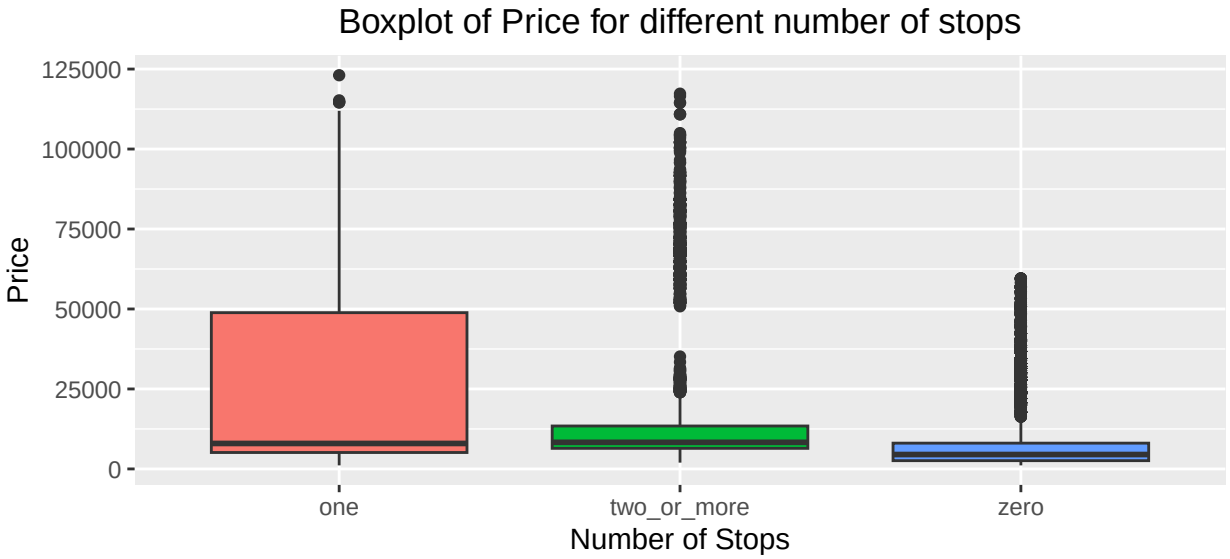
The correlation matrix is as follows -

```
##              duration   days_left       price
## duration   1.00000000 -0.03915688  0.20422237
## days_left -0.03915688  1.00000000 -0.09194853
## price      0.20422237 -0.09194853  1.00000000
```

Correlation matrix heatmap -



The correlation matrix shows that there is a positive correlation between number of days left before departure and price of plane ticket, which we also concluded from the line graph before. We also observe that duration of flight and price of plane ticket are positively related, which is obvious because as duration of a certain plane ride increases, the cost of plane fare should increase.



The median price of non-stop flights is sightly lower than one or two stop flights. The distribution of ticket price for one stop flights is positively skewed, with a wide range of variation.

Results -

The main findings from our analysis are that we observed significant price variations among different airlines. Some airlines consistently offered lower fares, while others tended to be more expensive. Passengers may benefit from exploring multiple airline options to find the best deals.Prices varied considerably depending on the specific routes. Some routes were consistently more affordable, while others commanded premium fares. Understanding the pricing dynamics of popular routes can help travelers make informed decisions. Advanced Booking: Our analysis confirmed that booking flights well in advance generally resulted in lower ticket prices. Last-minute bookings tended to be more expensive. Travelers seeking cost-effective options should plan their trips and book tickets with sufficient lead time.We noticed that certain airlines employed dynamic pricing strategies, adjusting ticket prices based on factors such as demand, time to departure, and seat availability. These pricing tactics underscore the importance of flexible travel plans for budget-conscious travelers. Ticket prices also exhibited regional variations, with certain destinations consistently offering more affordable airfare options than others. Travelers can leverage this information when selecting their travel destinations.

Conclusion -

In conclusion, understanding the factors influencing plane ticket prices is essential for travelers seeking to make informed choices and optimize their travel budgets. By considering the insights gained from this analysis and remaining vigilant for deals and discounts, passengers can enhance their overall travel experience while minimizing expenses.