

# Flight Ticket Price Predictor

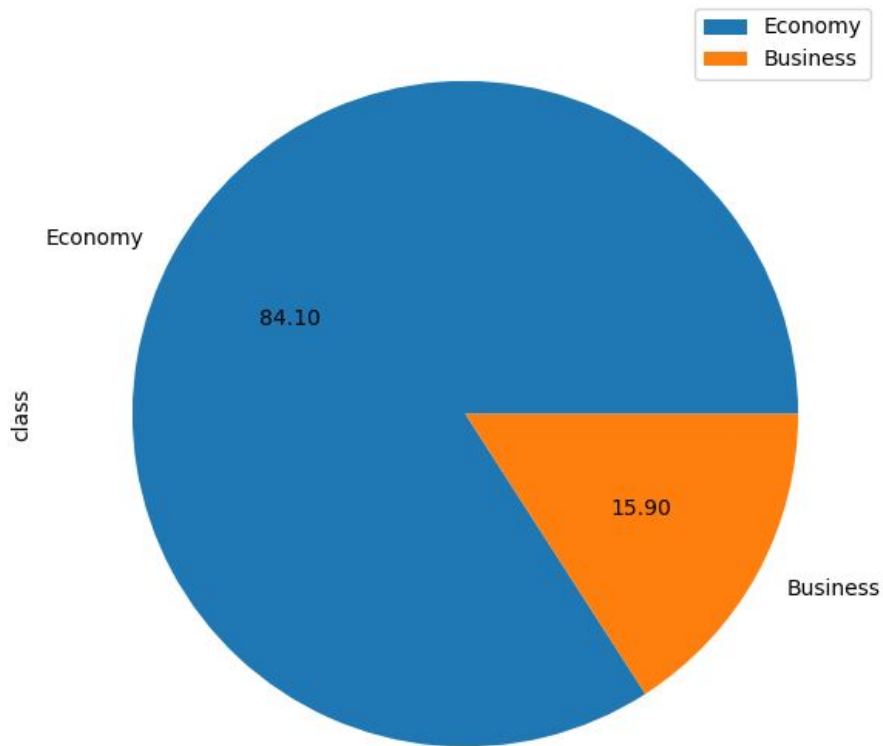
Final Project  
DA - Ironhack  
Clara Balcells

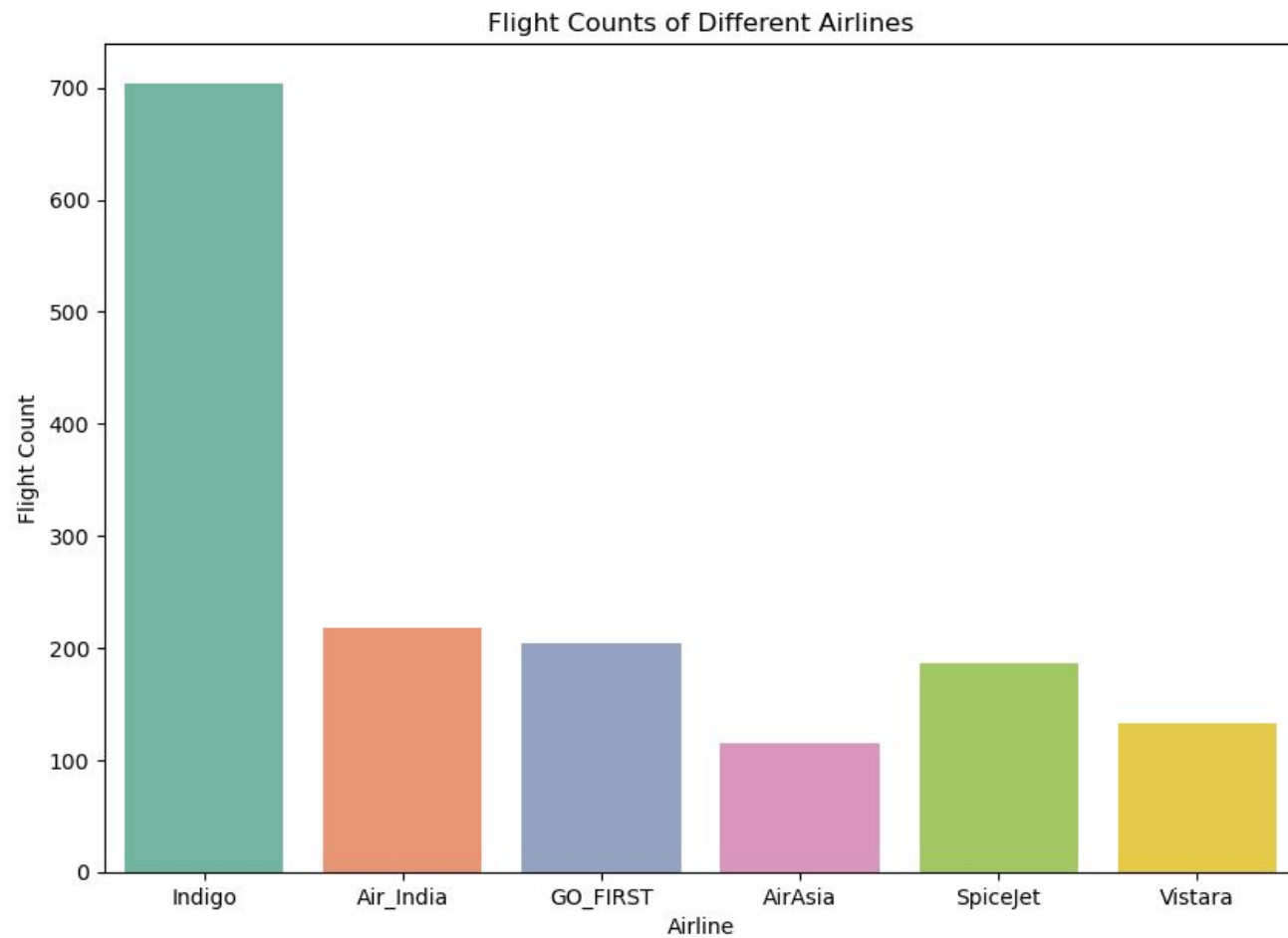


# Business Questions

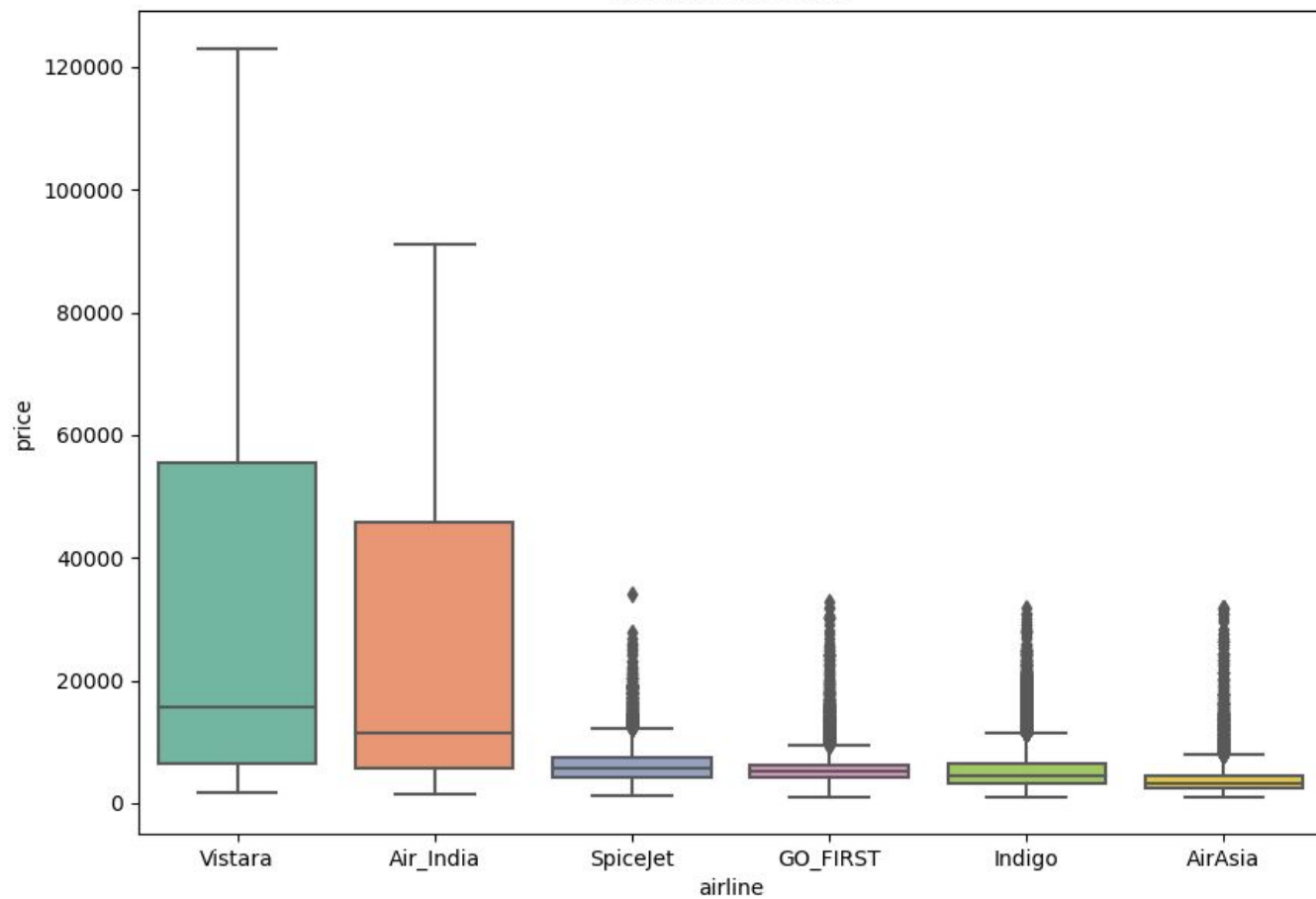
- How does the ticket price vary between Economy and Business class?
- Does price vary within Airlines?
- How the price changes when tickets are bought before departure?
- Does ticket price change based on the departure time and arrival time?
- How the price changes based on Source and Destination?
- Does the number of stops influences the price?

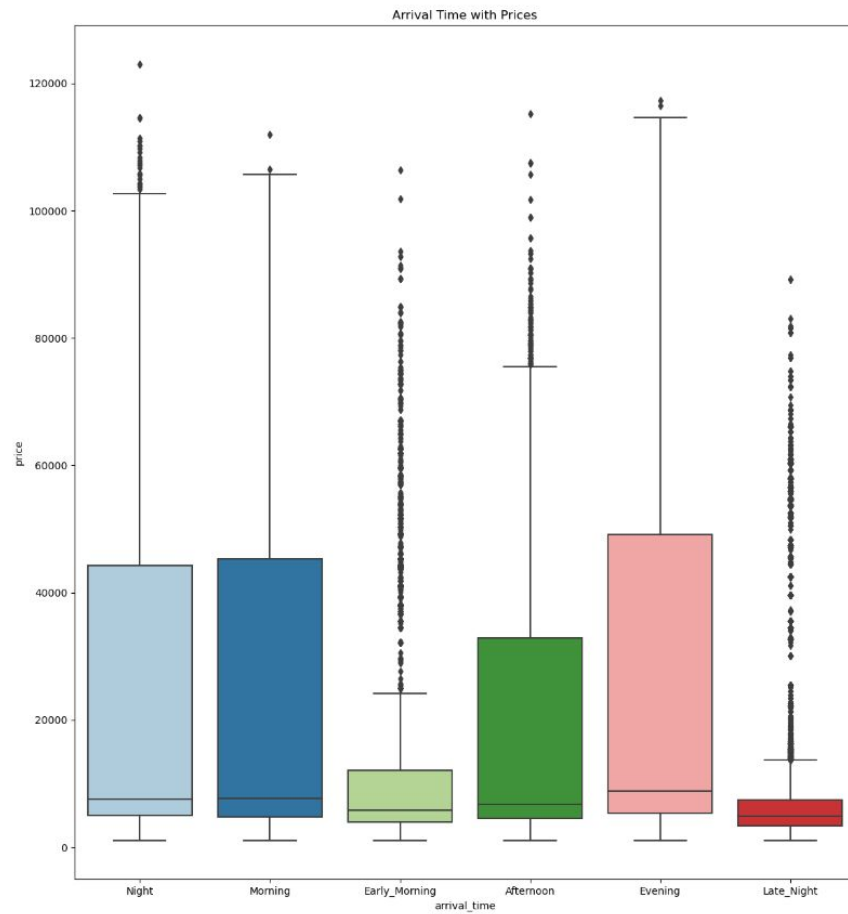
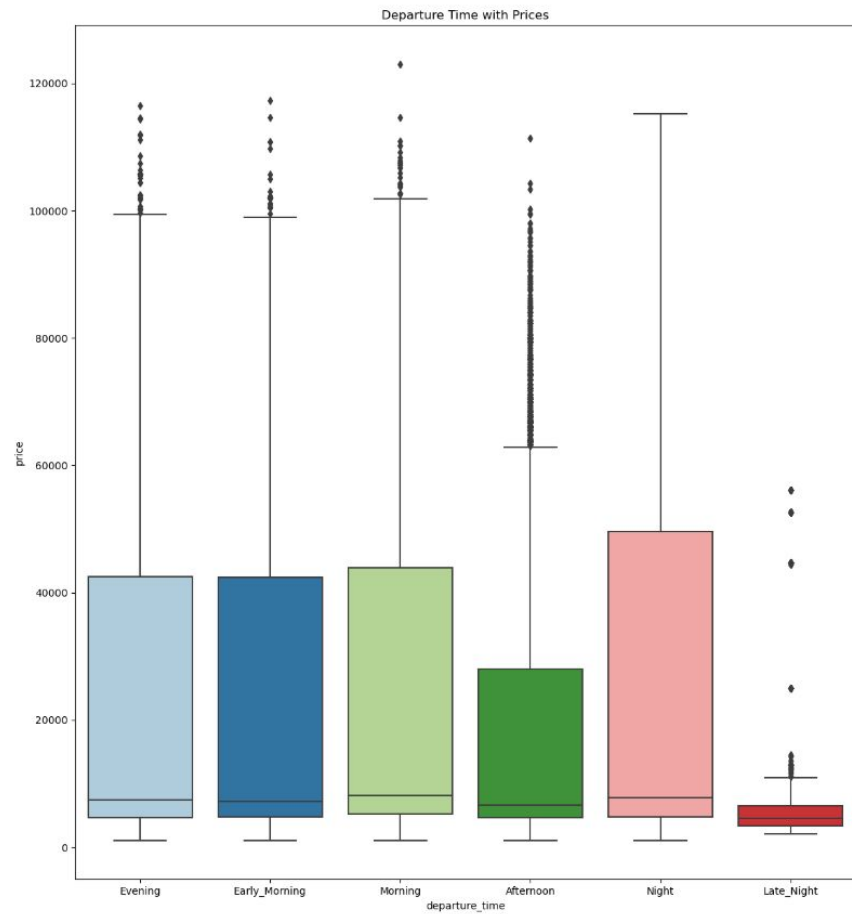
# Initial EDA

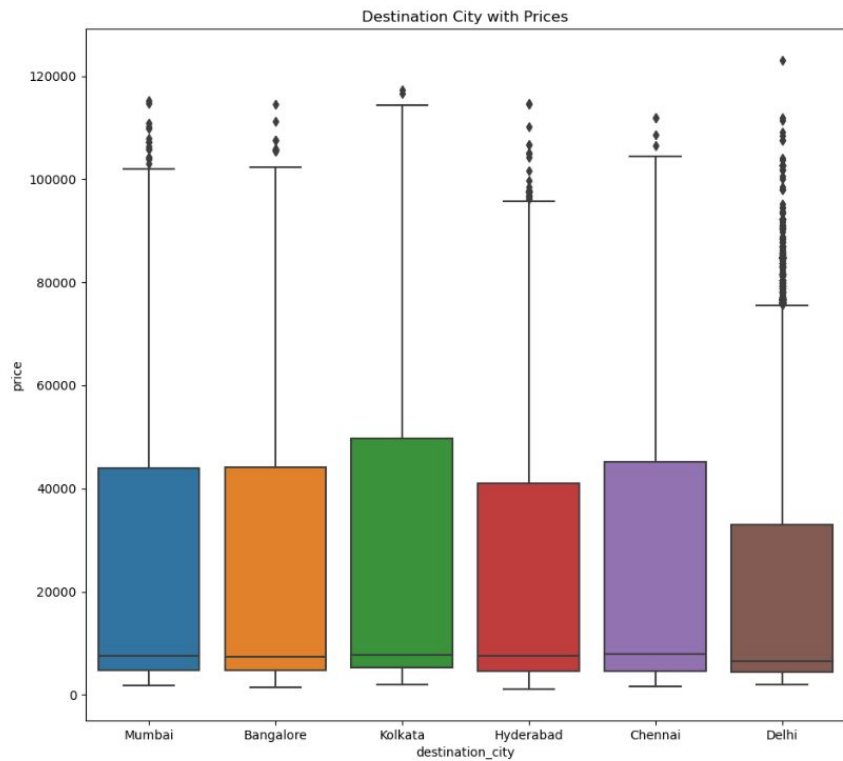
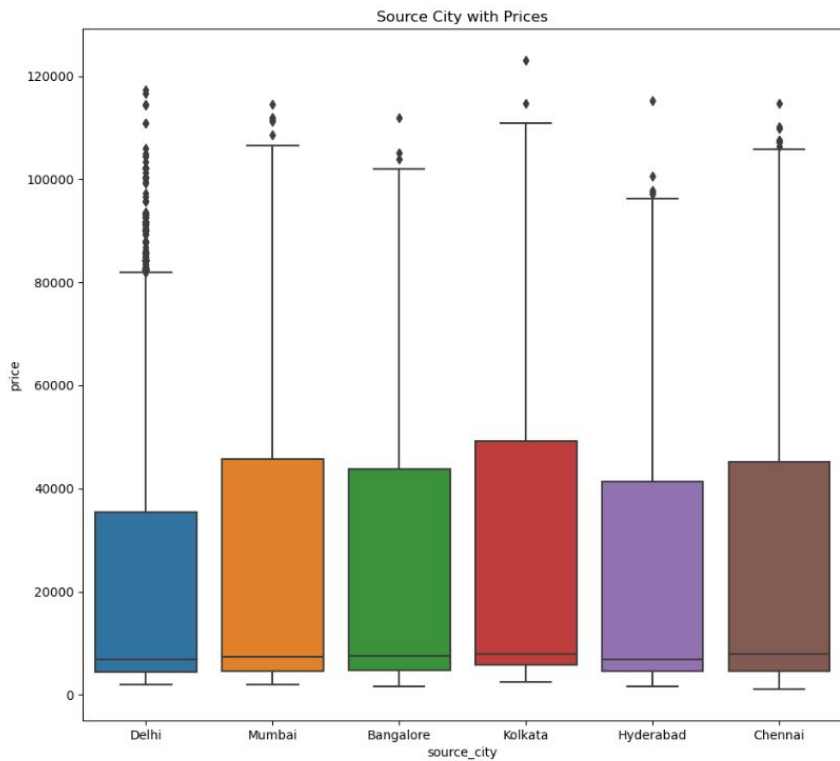




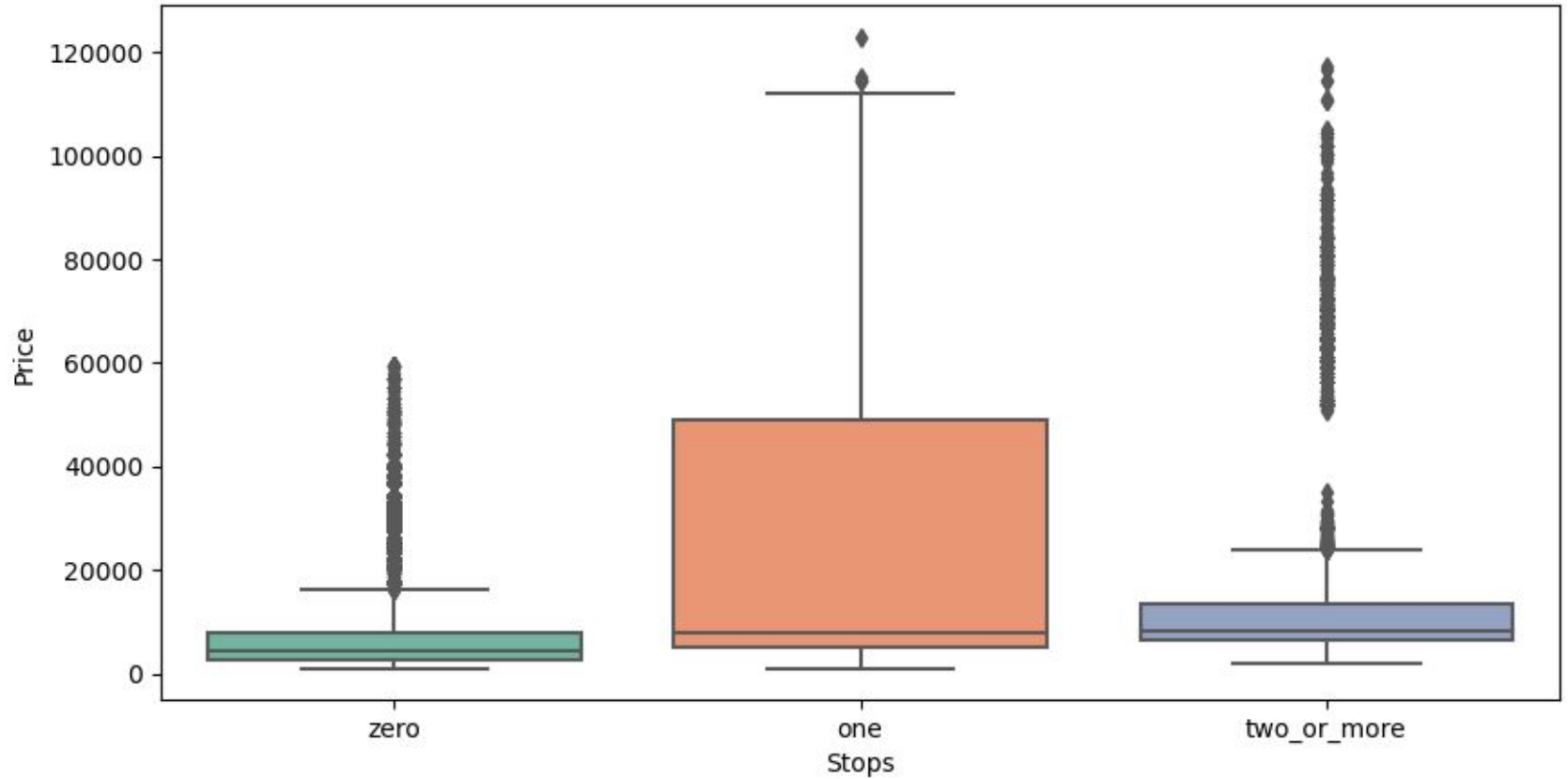
Airlines with Prices





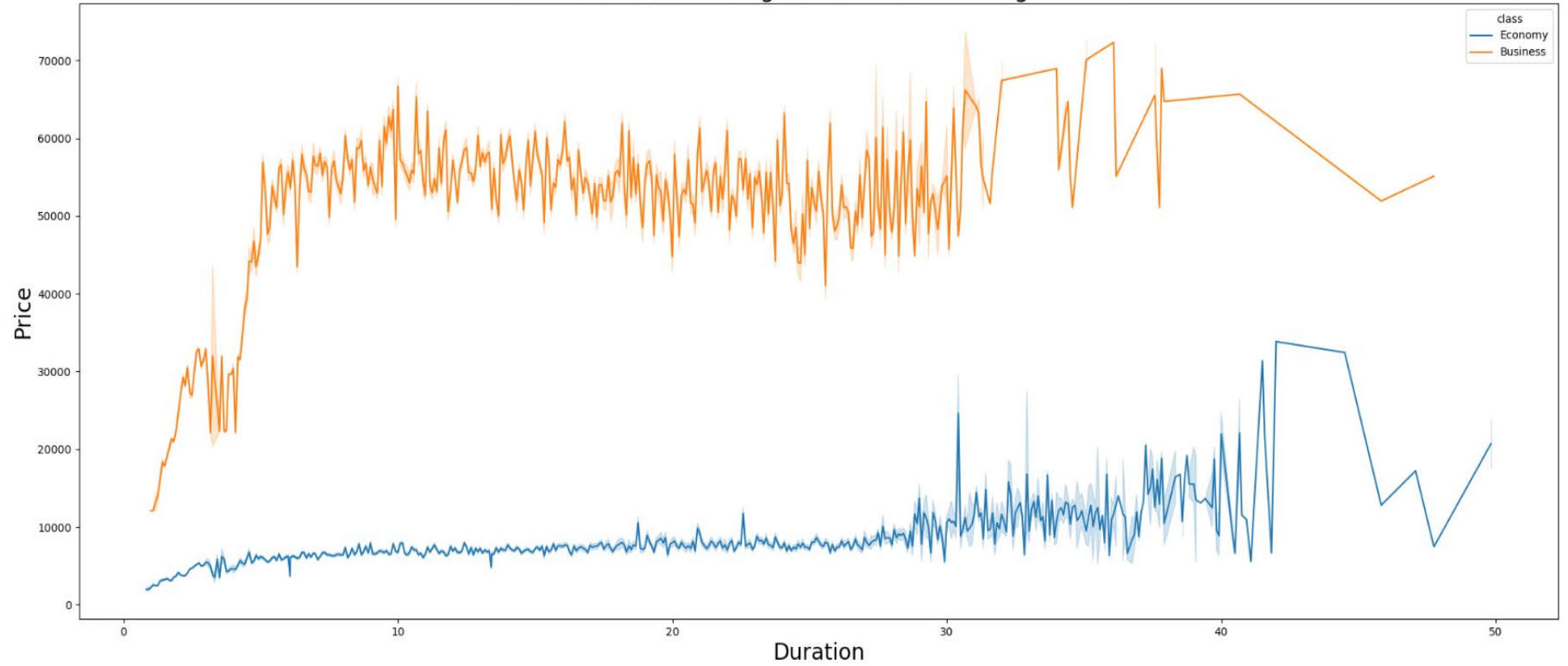


Stops Vs Ticket Price

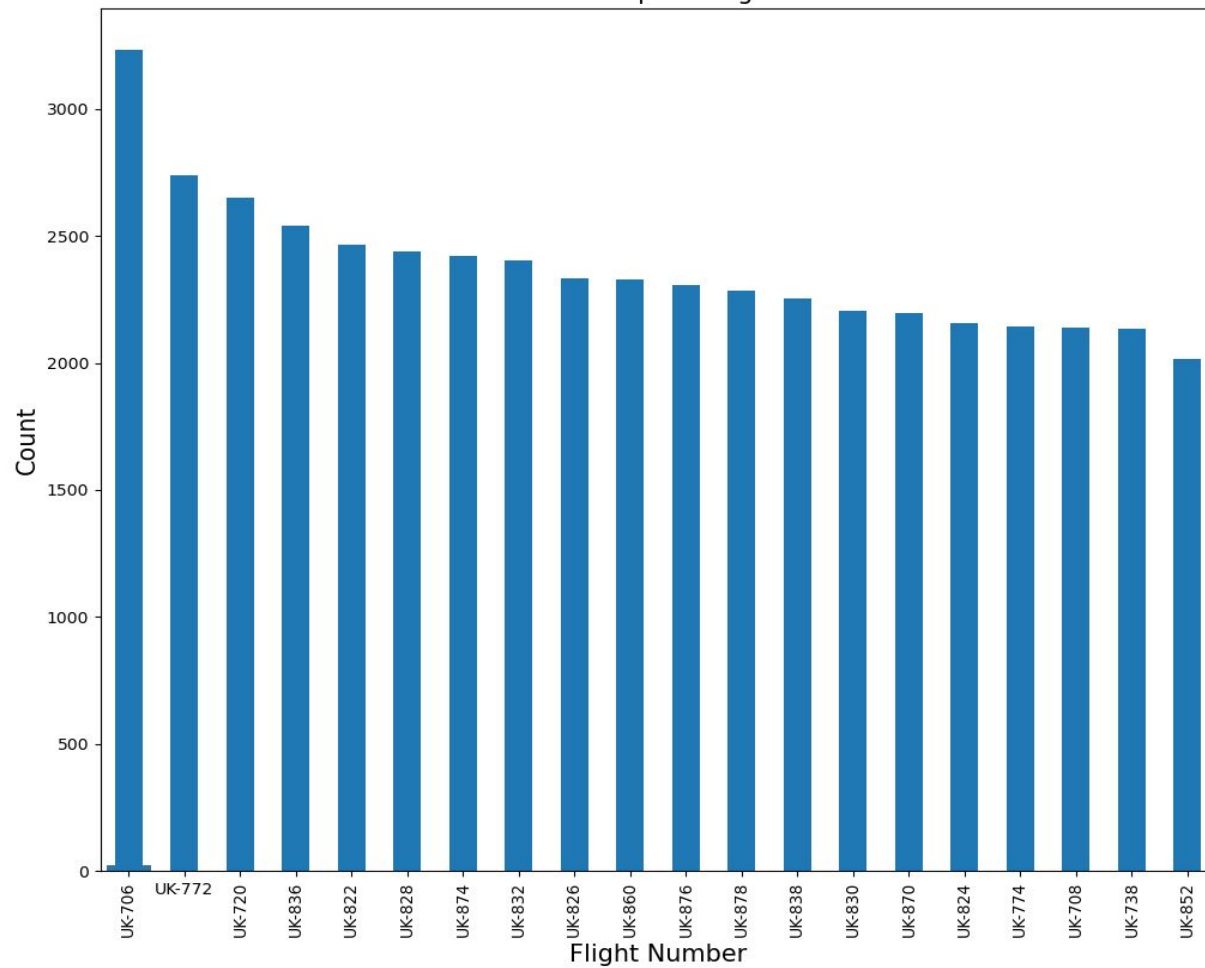


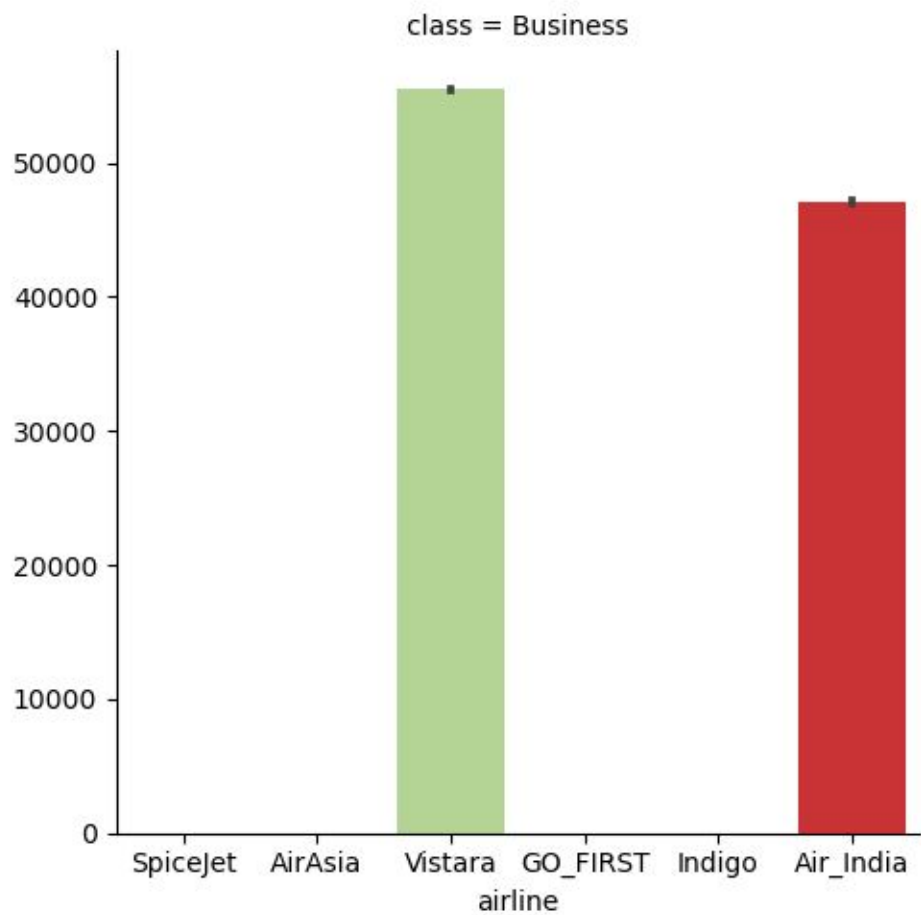
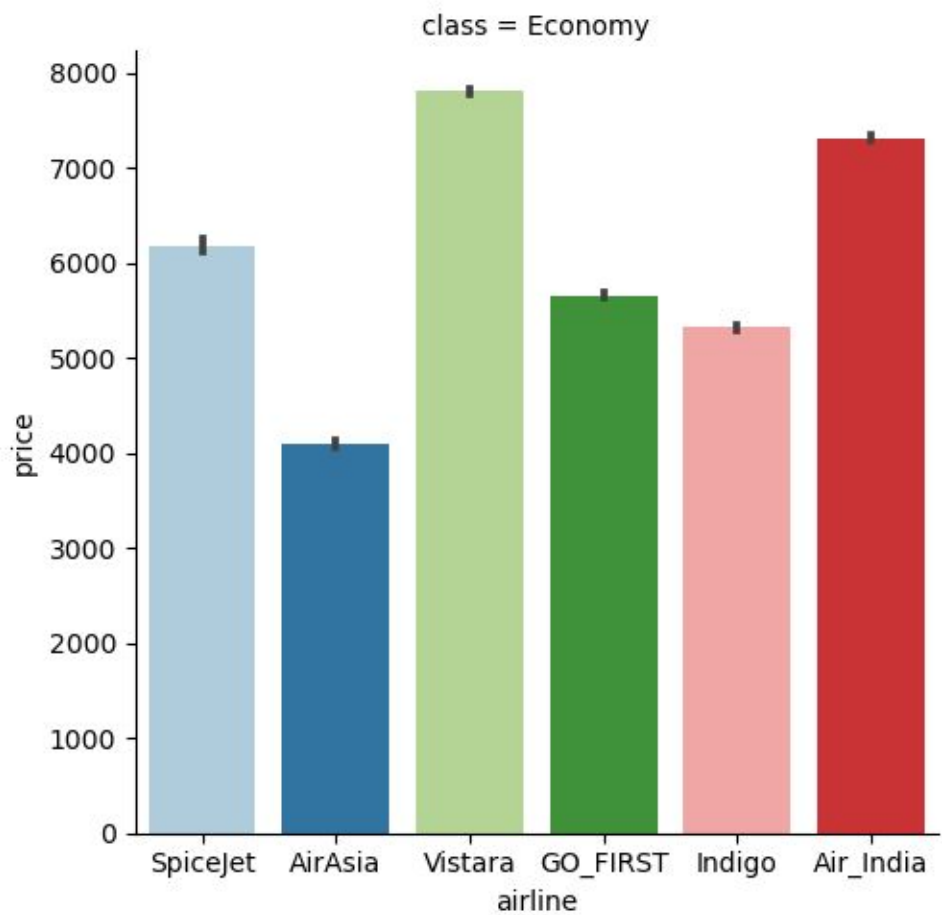


Ticket Price Versus Flight Duration according to Class

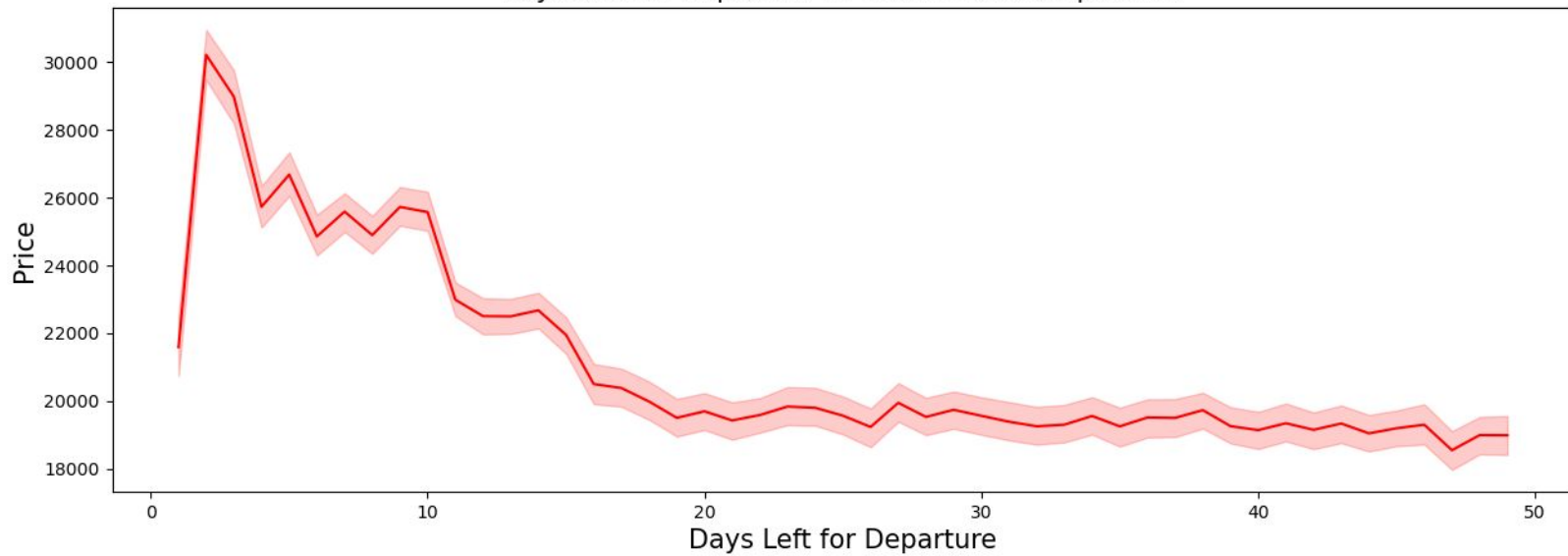


Most Frequent Flights





Days Left For Departure vs Ticket Price Comparative



# EDA Conclusions

- There is a big gap between flight tickets in business and economy. Business tickets are 6 times more expensive than economy tickets.
- The most expensive airlines appear to be Vistara and AirIndia, while AirAsia is the least expensive. There are just two airlines that provide business class tickets, Vistara and AirIndia, and Vistara is significantly more expensive.
- Prices generally increase gradually until 20 days prior to the flight, at which point they dramatically increase. However, there are frequently open seats that haven't been filled a day before takeoff. As a result, tickets can be found for three times less than the previous day.
- It appears that early morning and late night departures are less expensive, whereas evening departures are more expensive. It appears that early departure, late departure, and late departure are less expensive than night time departure.
- Flights to Chennai are slightly more expensive than those to other cities, but flights from Delhi are the least expensive overall.
- Generally, the cost of an airline ticket increases with the number of stops.

# Models used:

- Linear Regression Model
- Decision Tree Model
- Gradient Boosting Model
- Random Forest Regression Model
- KNeighbors Regression
- Bagging
- Extra Trees Regression

# Best model performance:

## Random Forest Regression Model

- R2 Score: 0.9894574805166351
- Mean Squared Error is 5398550.098449273
- Mean Absolute Error is 888.2363460939549
- Root Mean Squared Error is 2323.478017638487

# Random Forest Model Prediction

