**Optimizing Flight Booking Decisions through Machine Learning Price Predictions**



SUBMITTED BY

MAHALAKSHMI R

MCA I YEAR

**ABSTRACT**

People who work frequently travel through flight will have better knowledge on best discount and right time to buy the ticket. For the business purpose many airline companies change prices according to the seasons or time duration. They will increase the price when people travel more. In this project, we estimating the highest prices of the airlines data for the route is collected with features such as Duration, Source, Destination, Arrival and Departure. Features are taken from chosen dataset and in the price wherein the airline price ticket costs vary overtime. We have implemented flight price prediction for users by using KNN, decision tree and random forest algorithms. Random Forest shows the best accuracy of 80% for predicting the flight price. Also, we have done correlation tests and metrics for the statistical analysis.

**Languages used:**

* Front – End : HTML, CSS
* Back – End : Python
* Technology : Machine Learning
* Dataset : Data\_Train.xslx

1. **INTRODUCTION**

**1.1 Overview:**

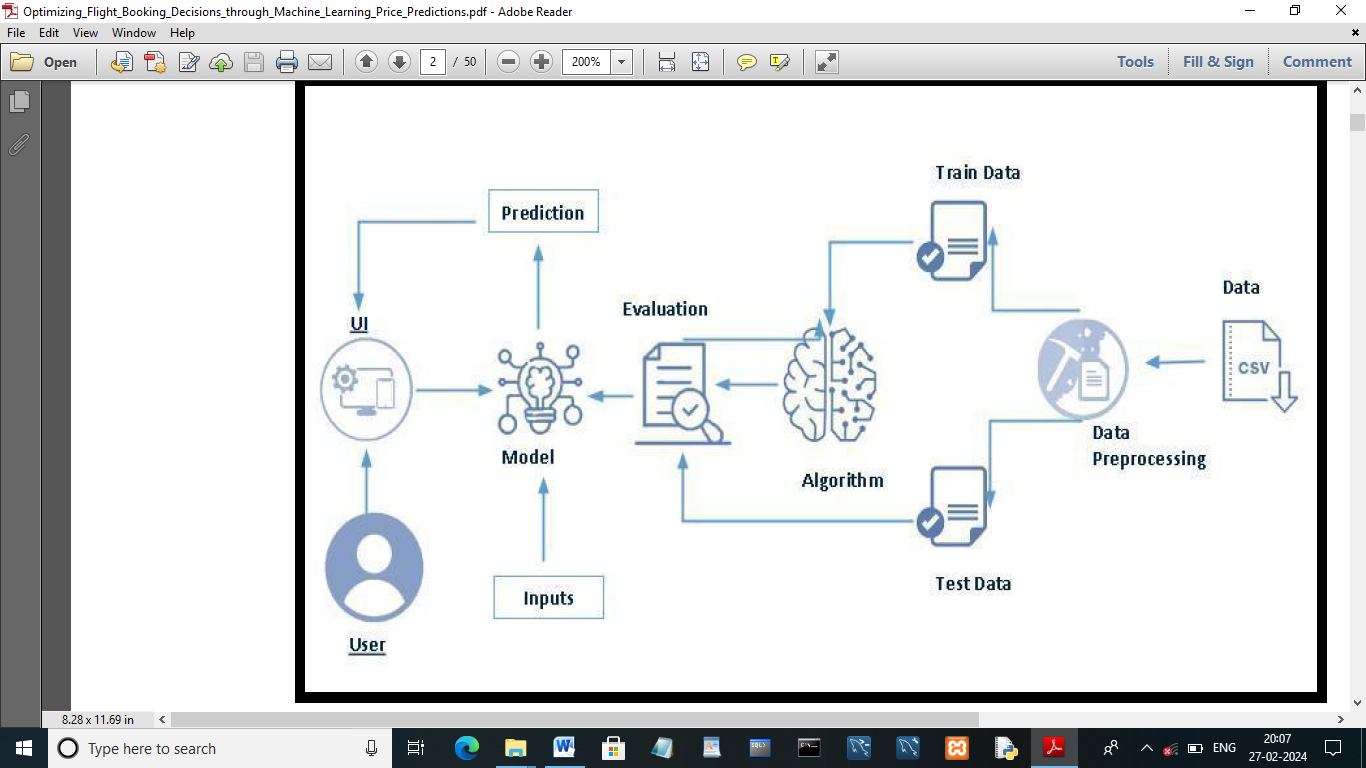
Perfect time for purchasing plane ticket by the passenger’s view is difficult since passengers get very less information of future business price rates. Different models figure out future business price on plane and categorise the best time to obtain flight ticket. Airlines use different strategies of pricing for their tickets, later taking the decision on price because order shows higher value for the approximation models. The causes behind the difficult system is each Planes has limited number of seats to be filed, so airlines must regulate demand. Suppose when demand is expected to increase capacity, the airline may increase prices, to decrease the rate at which seats fill.

Also, seating arrangements in flight which is not occupied shows the loss of the amount invested for the business airline companies and making them purchase the ticket to fill the seats for any price this would be the best idea to get profit in loss too. Passengers should be compatible with the airline companies to get adjusted for the increase and decrease of the price. Passengers or customers should make their own planning to get the best offers available on different airlines and travel through less price. Planes ticket prices changes as time passes, pulling out the elements which creates the difference. Reporting the correlated and models which is used to price the flight tickets. Then, using that information, building the model which helps passengers to make pull out the ticket to buy and predicting air ticket prices which progresses in the future. Duration, Arrival time, Price, Source, Destination and much more these are the attribute used for flight price prediction.

**1.2 Purpose**

* The primary purpose of flight price prediction is to help travelers save money by predicting when prices are likely to go up or down. By using this information, travelers can time their purchases to get the best possible deals on flights.
* A Flight price prediction application which predicts fares of flight for a particular date based on various parameters like source, Destination, Stops and Airlines. The prediction will help a traveller to decide a specific airline as per his/her budget.
* The main purpose of out system is to predict the flight prices with comparison of today to another any day because of this customer can be book their tickets of Flight according to their comfortably, according to their affordability, Means whichever cheaper cost they want they can be easily choose. The purpose is to provide customers with the relevant information they need to decide the best time to purchase a ticket.
* By helping travelers save money and plan their trips more effectively, flight price prediction tools can improve overall travel experiences. This reduces stress and enhances enjoyment, making travel more accessible and enjoyable for everyone.

**2. TECHNICAL ARCHITECTURE**

****

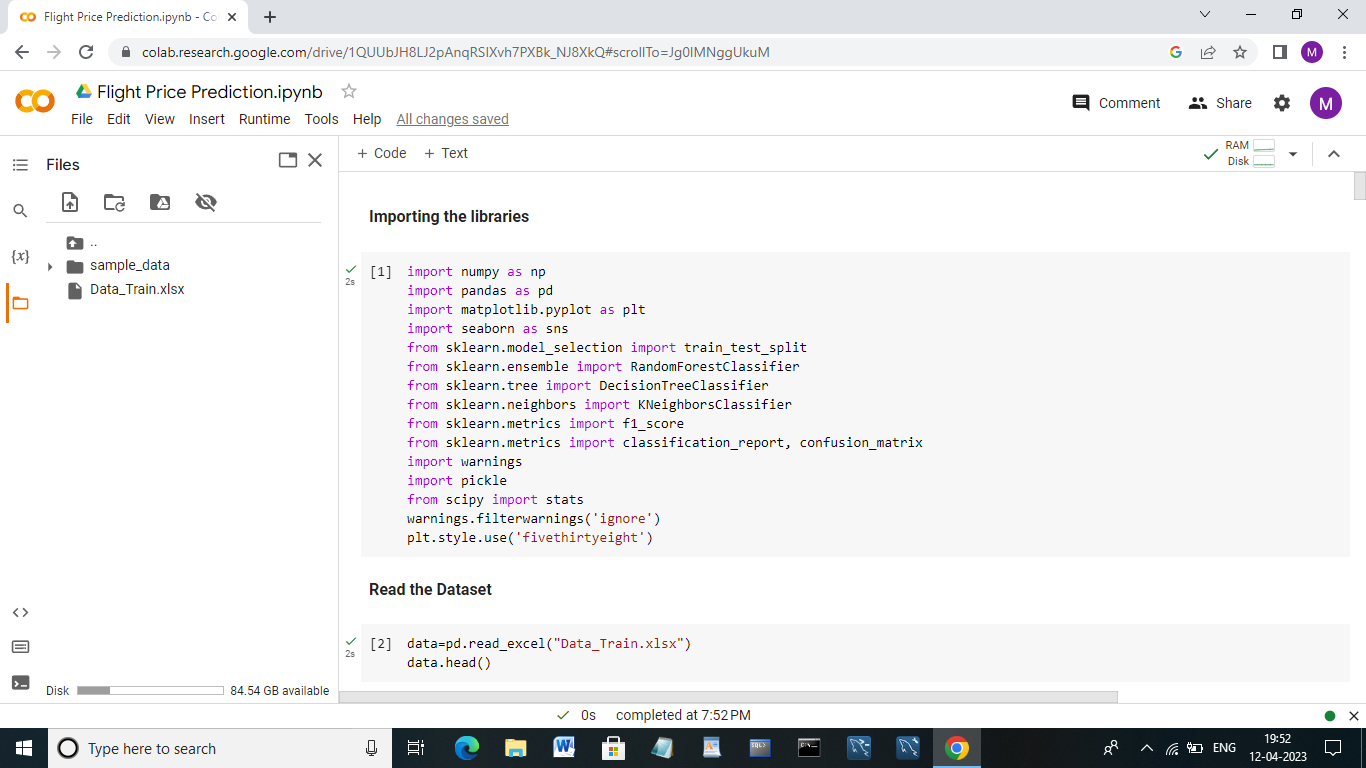
**3. DATA COLLECTION AND PREPARATION**

**3.1 COLLECT THE DATASET:**

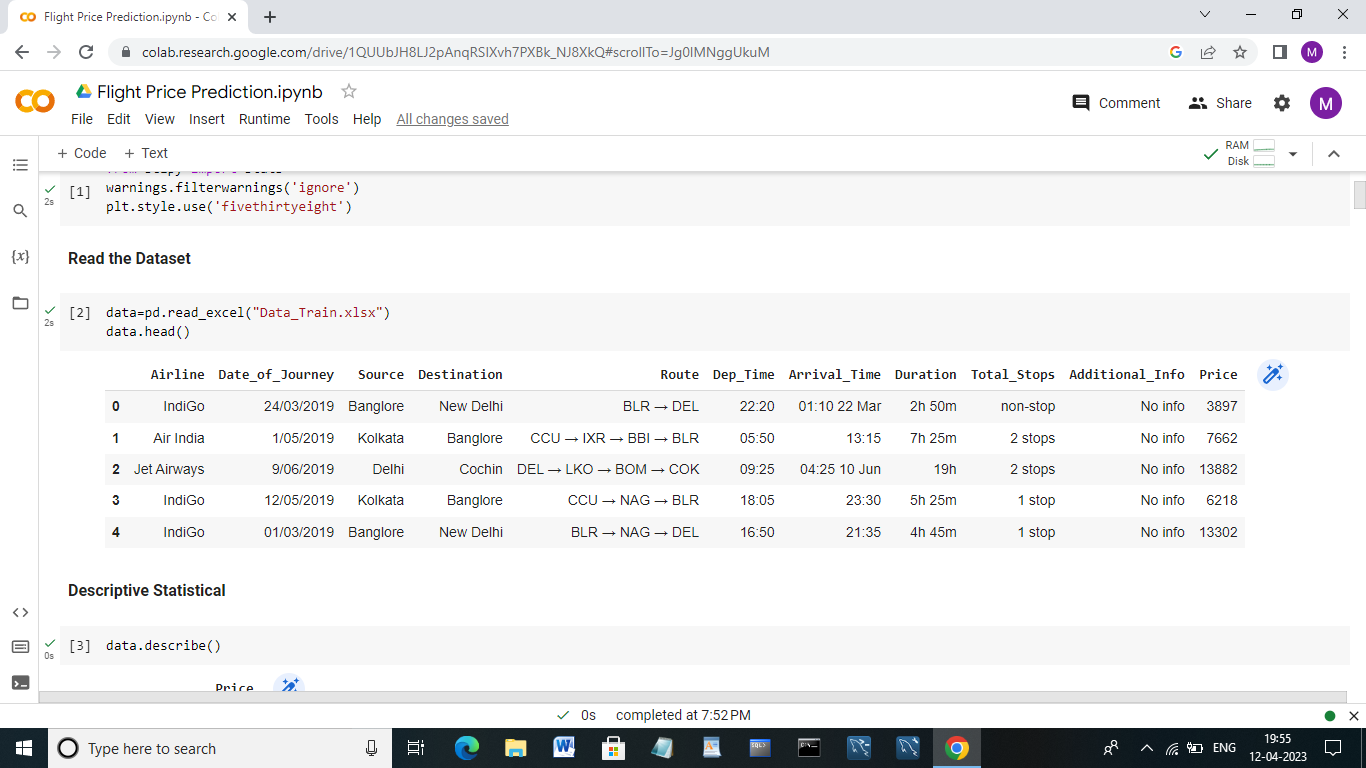
In this project we have used .csv data. This data is downloaded from kaggle.com.

Link: <https://www.kaggle.com/code/anshigupta01/flight-price-prediction/data>

**Import the libraries:**

****

**Read the data set:**

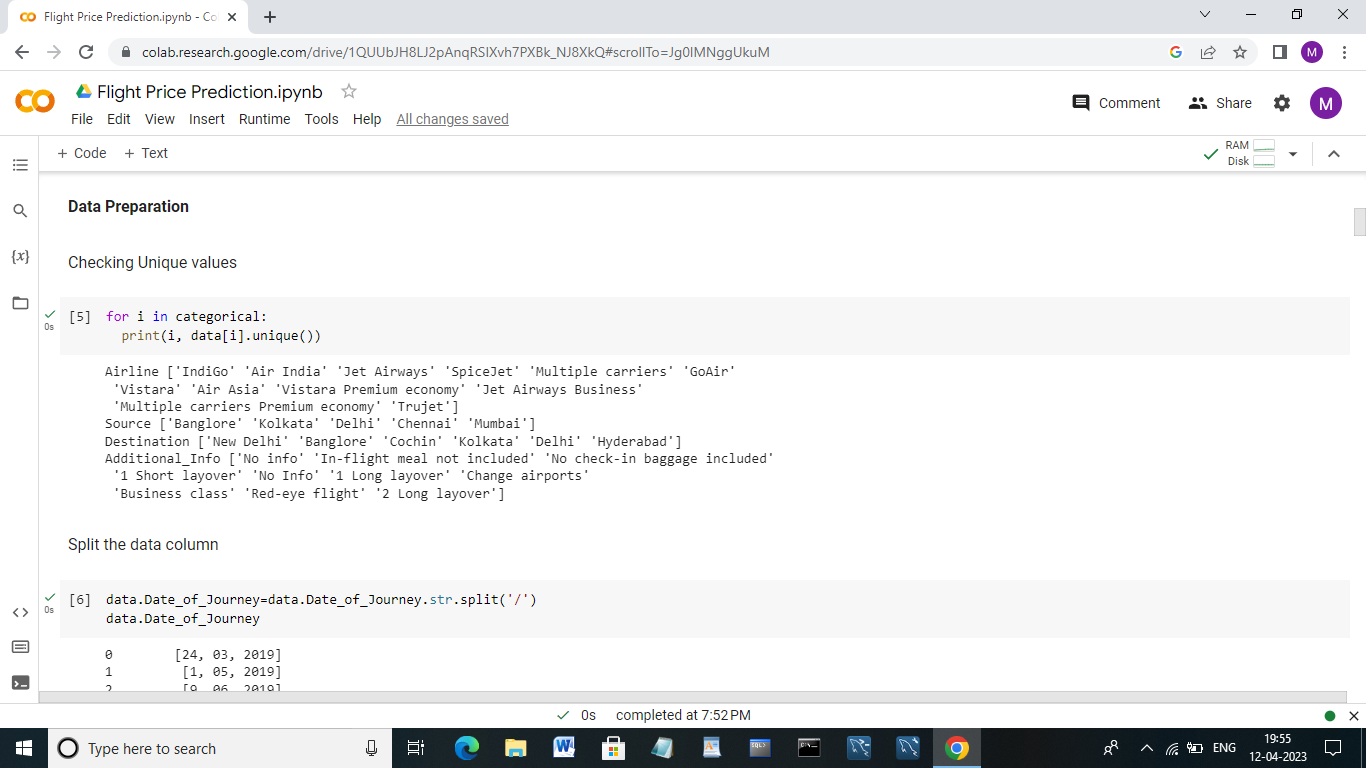


**Descriptive Statistical:**

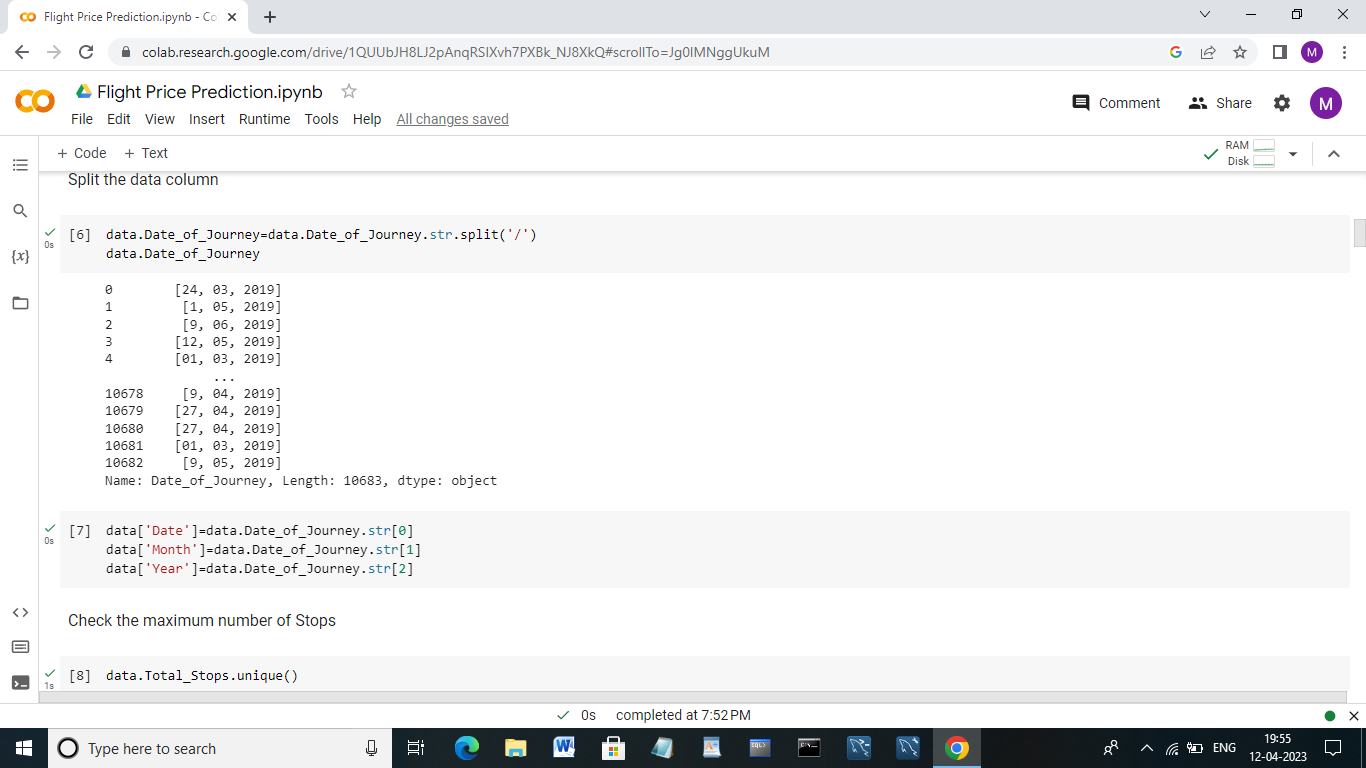


**Data Preparation:**

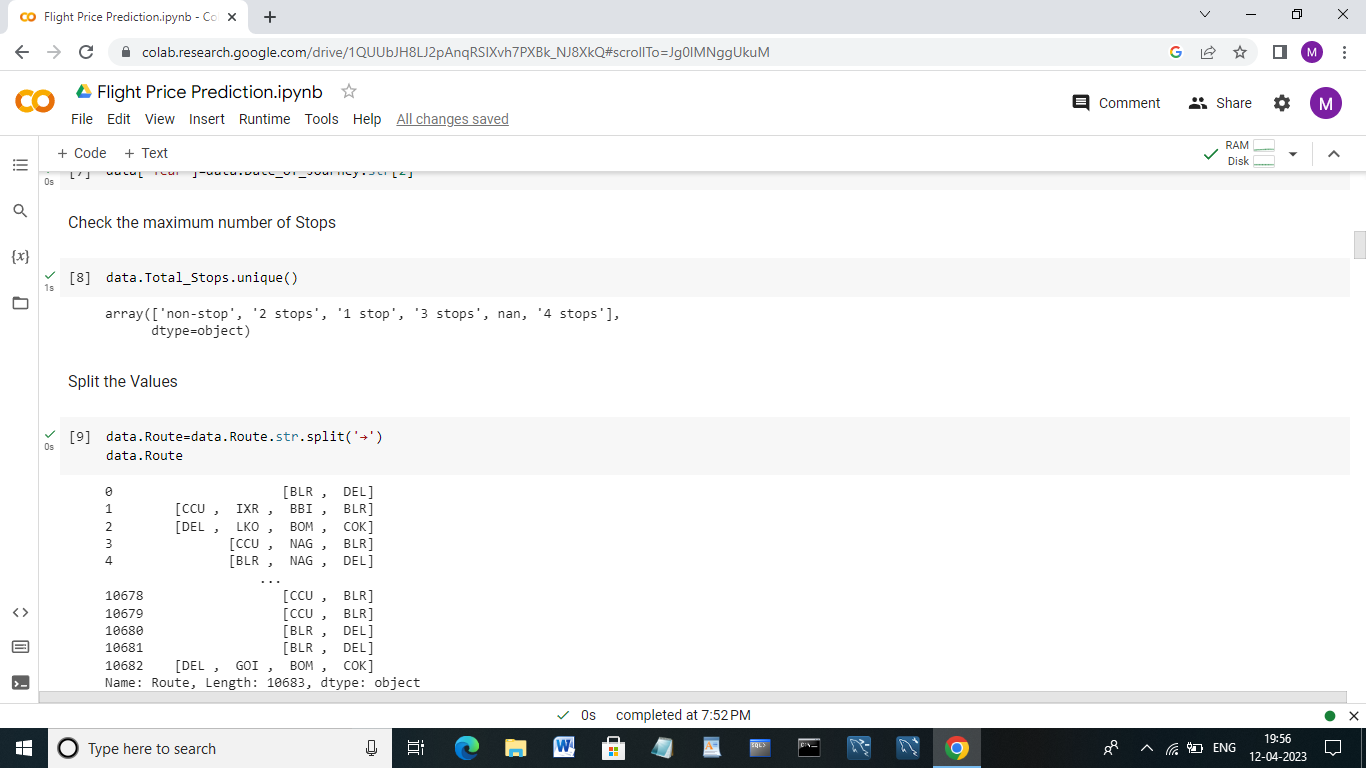
**Check the unique values:**



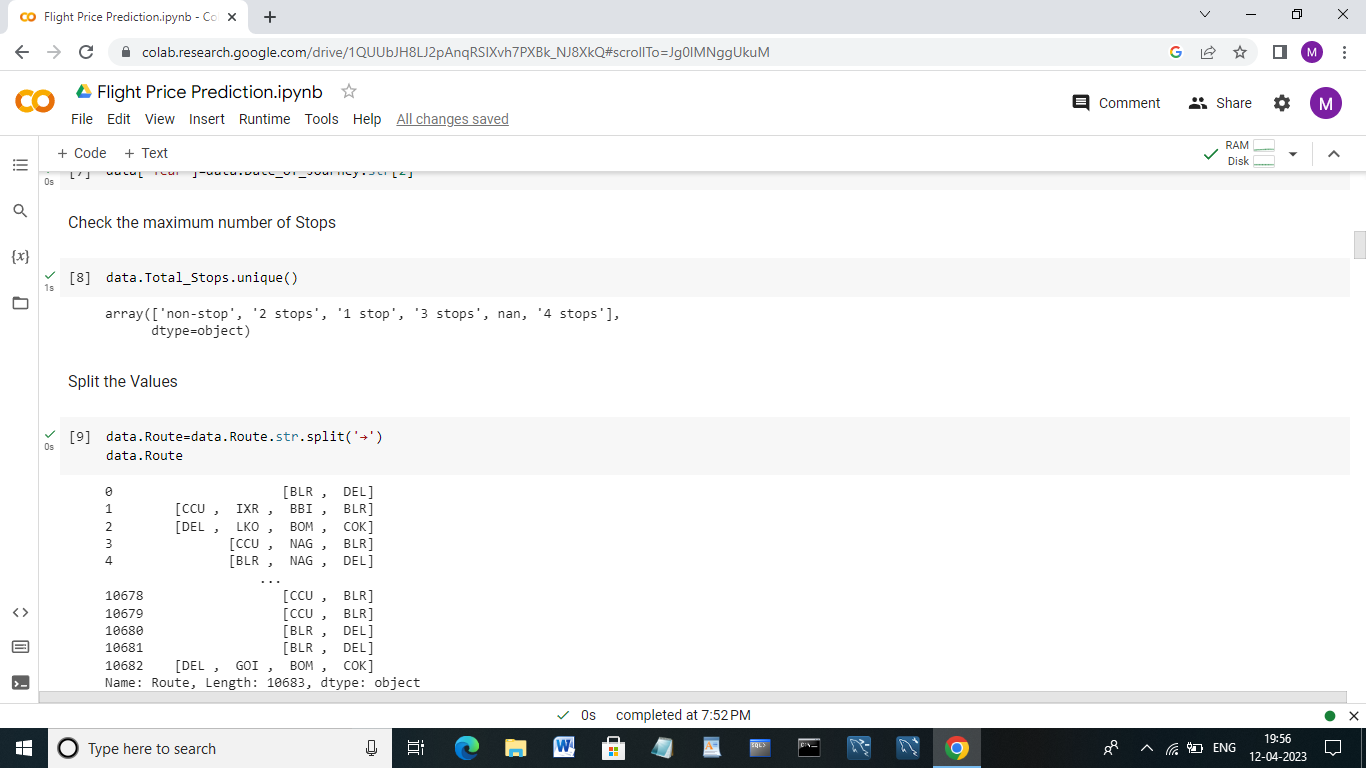
**Split the data column:**

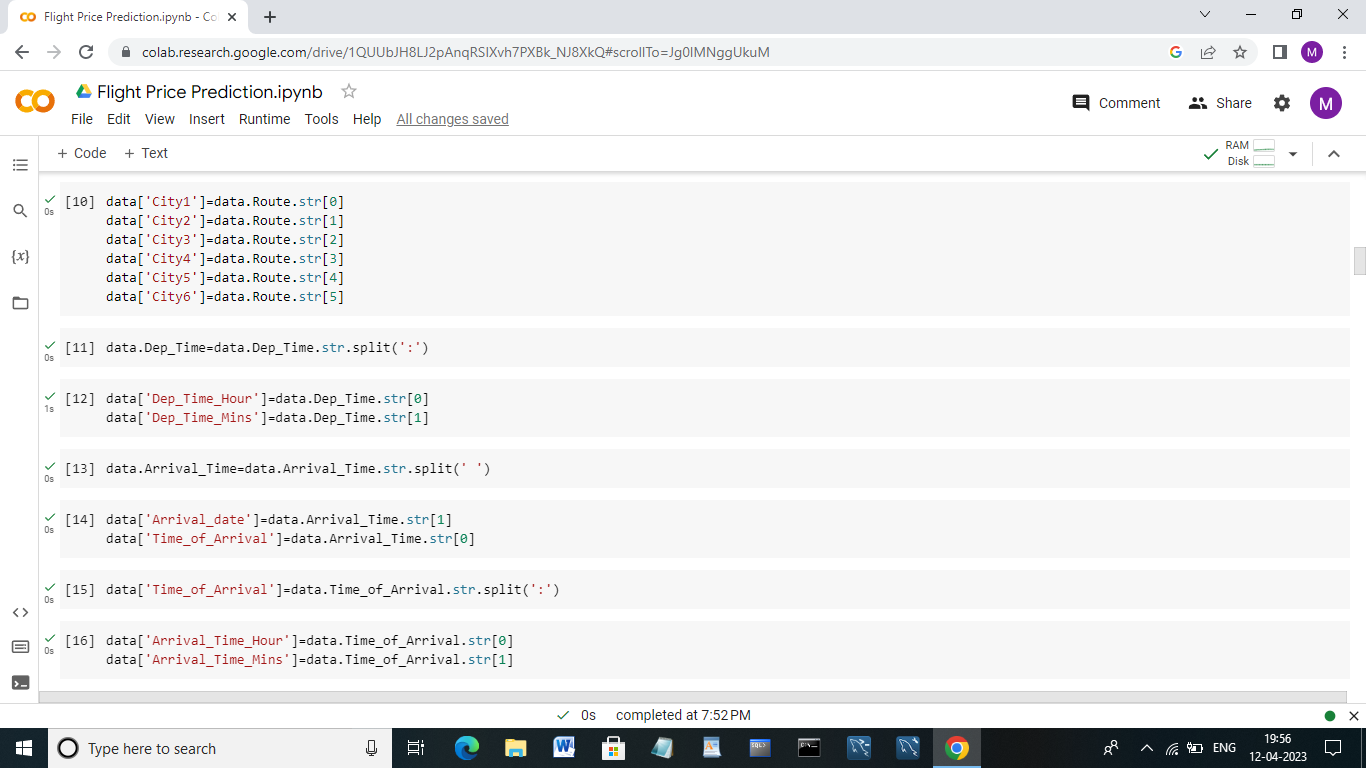
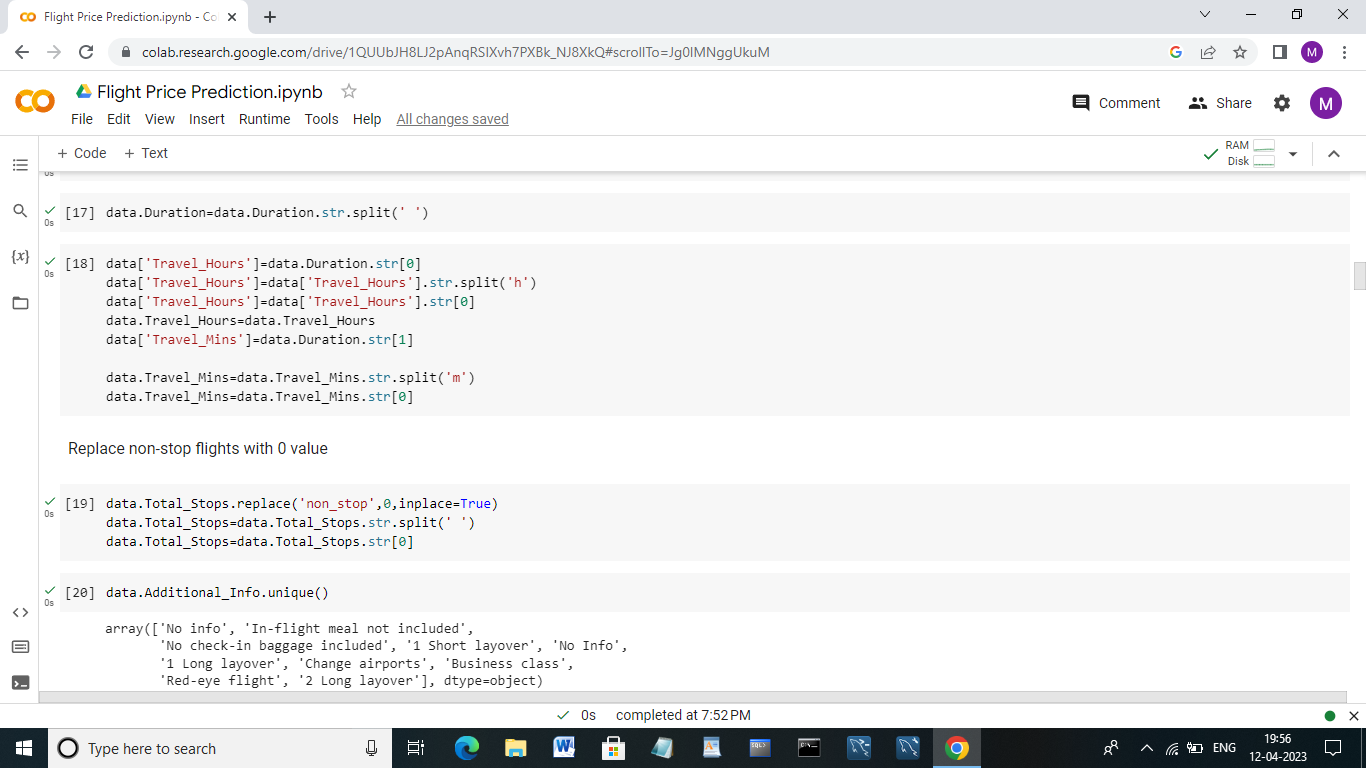


**Check the maximum Number of stops:**

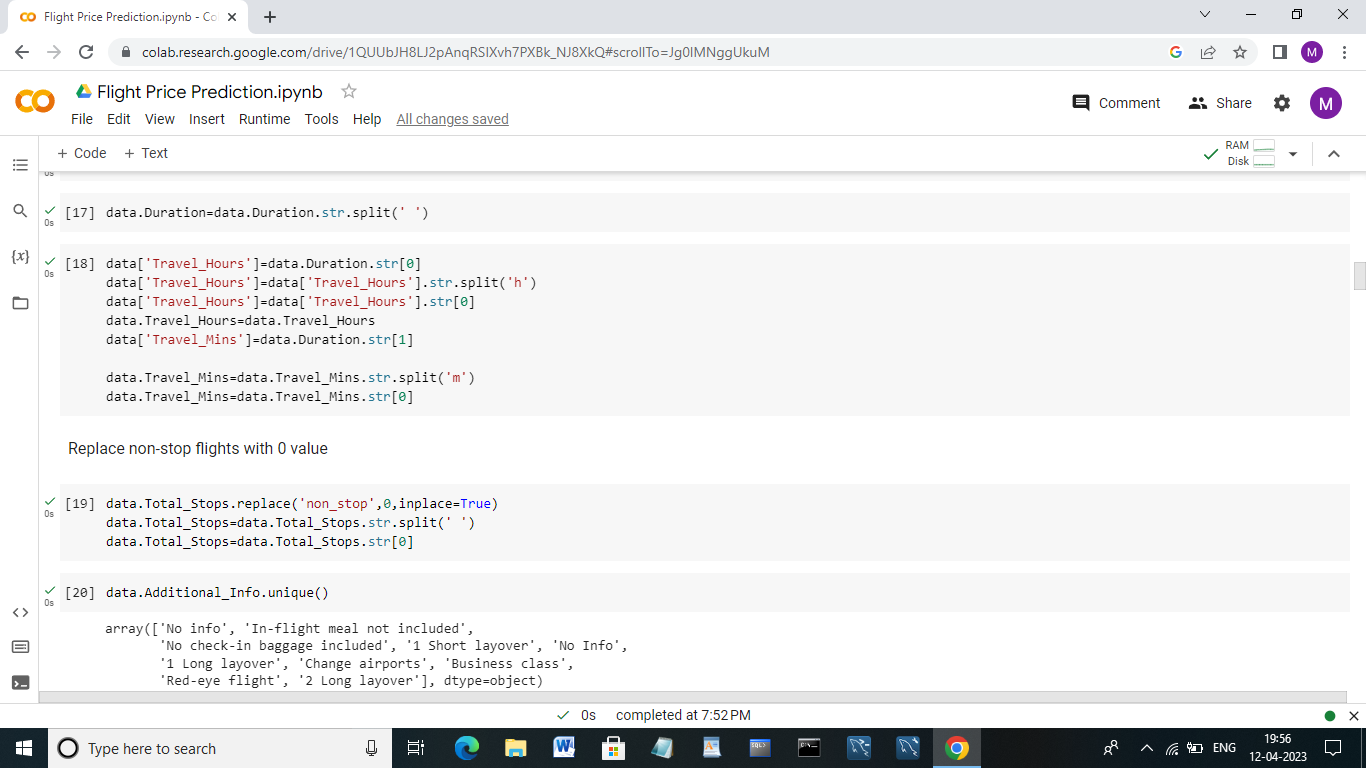


**Split the Values:**

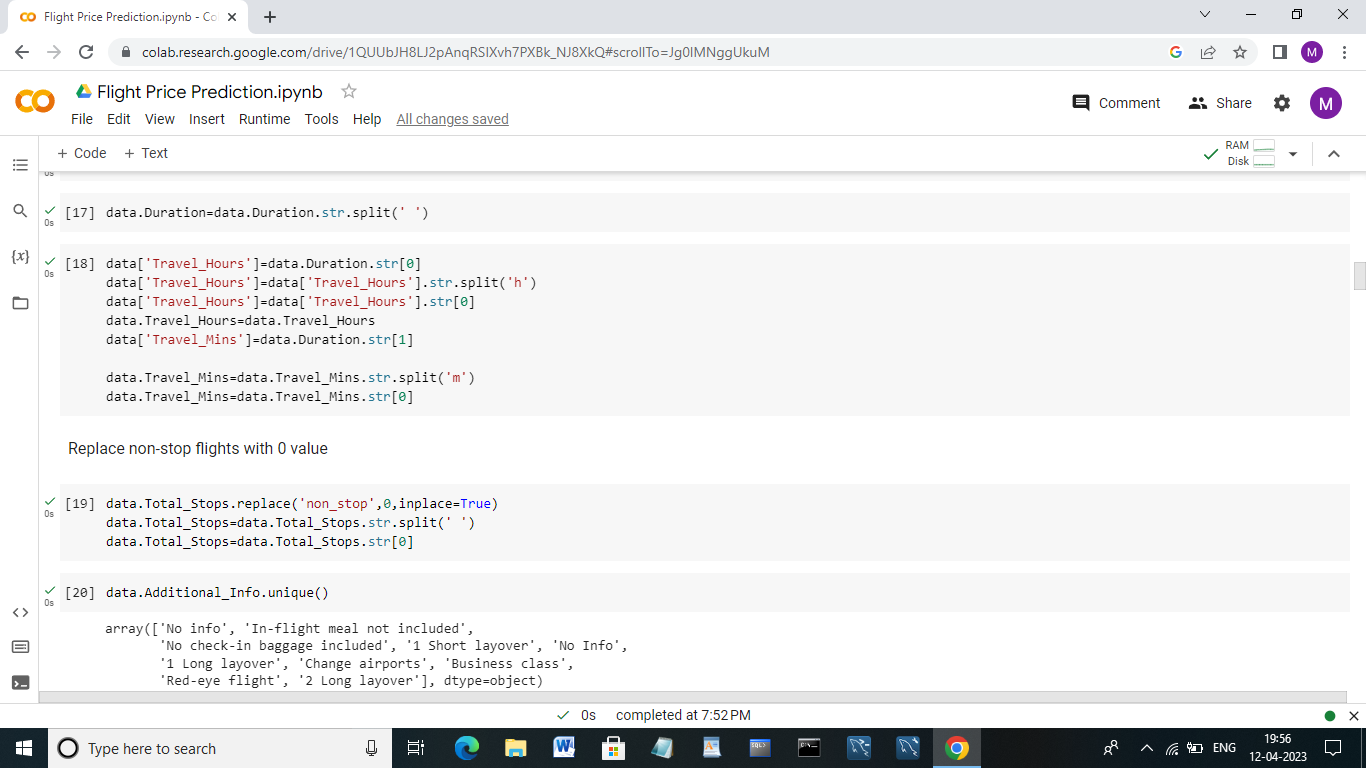
****

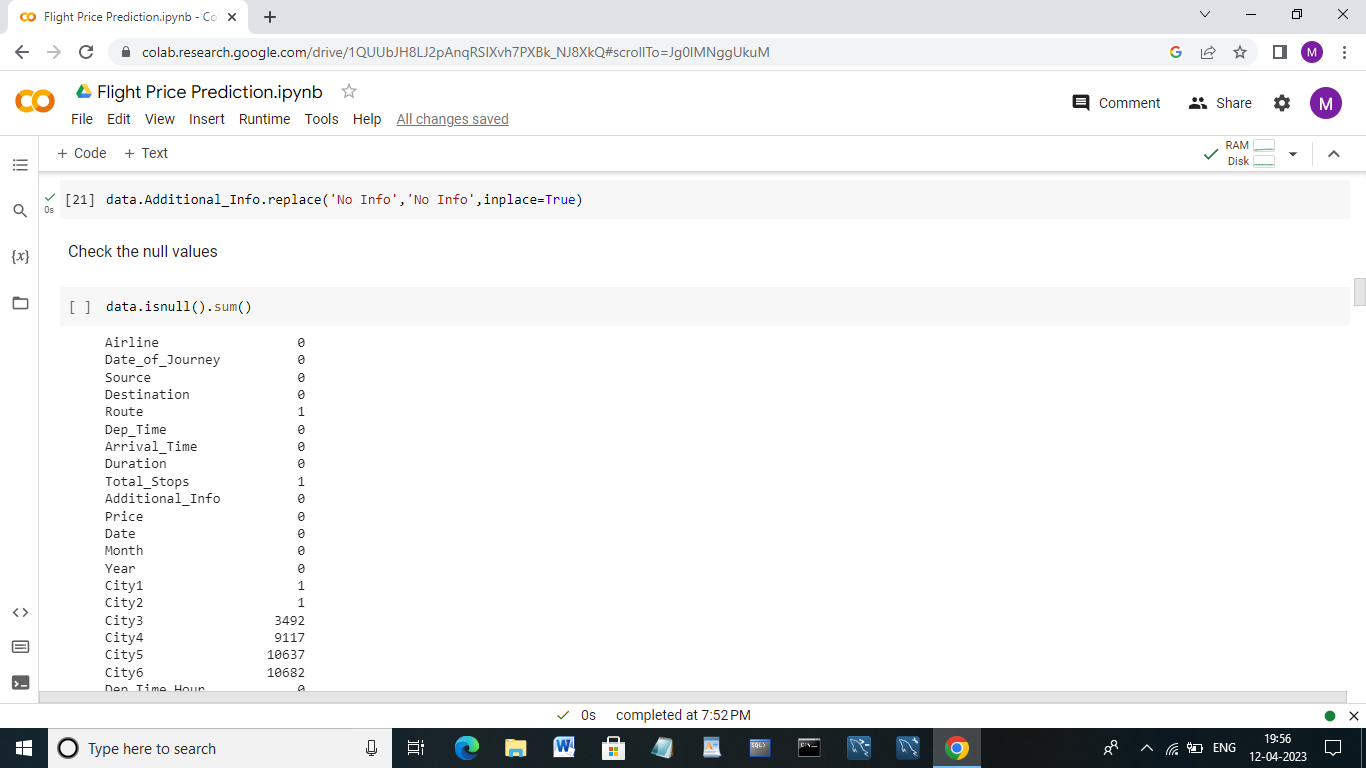
****

**Replace non-stop flights with o value:**

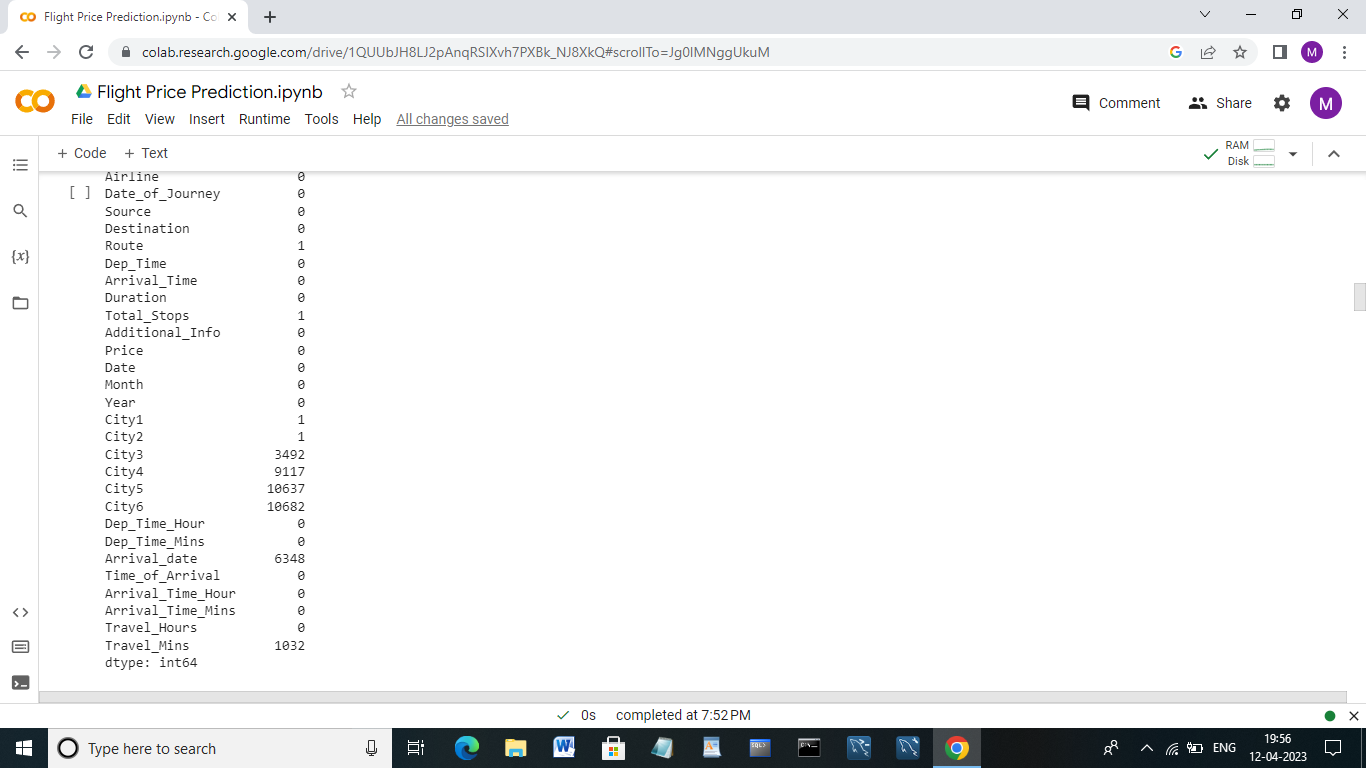
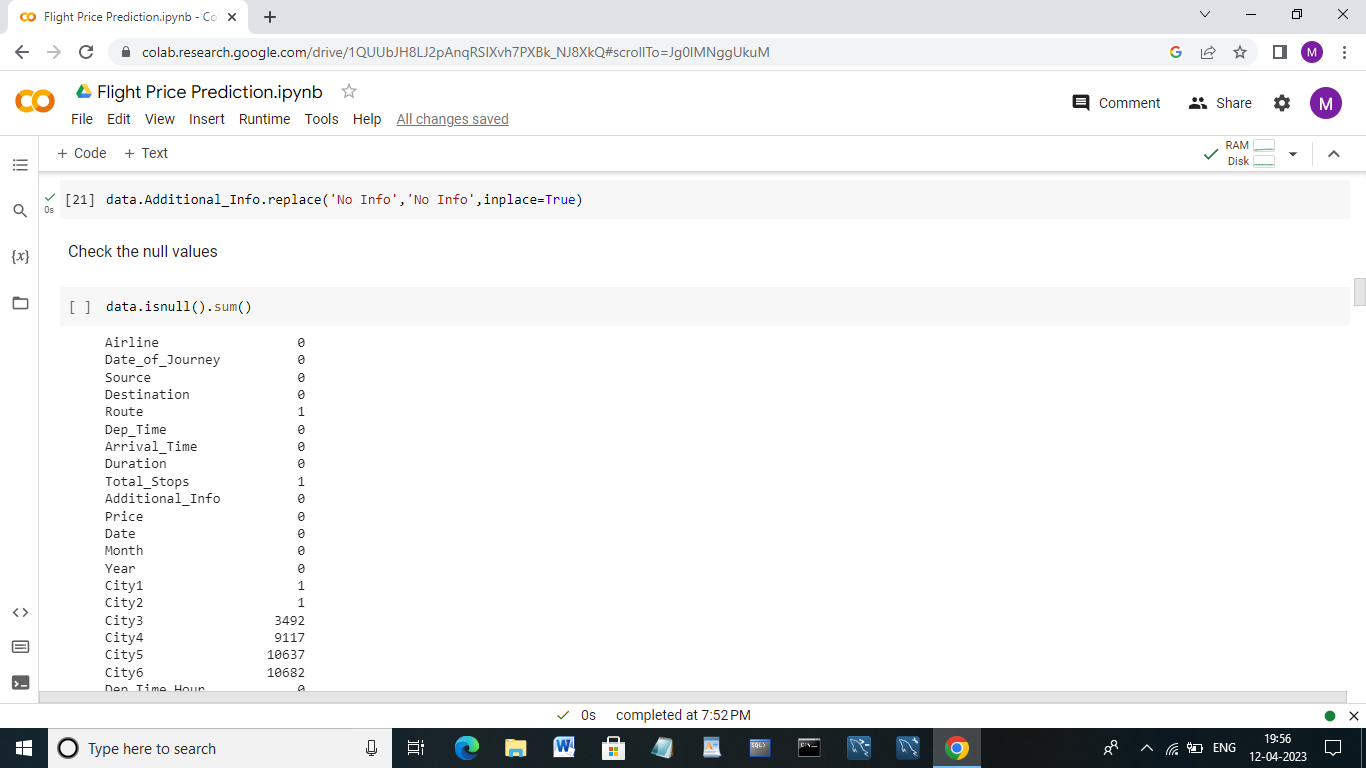
****

**Check the additional information:**

****

****

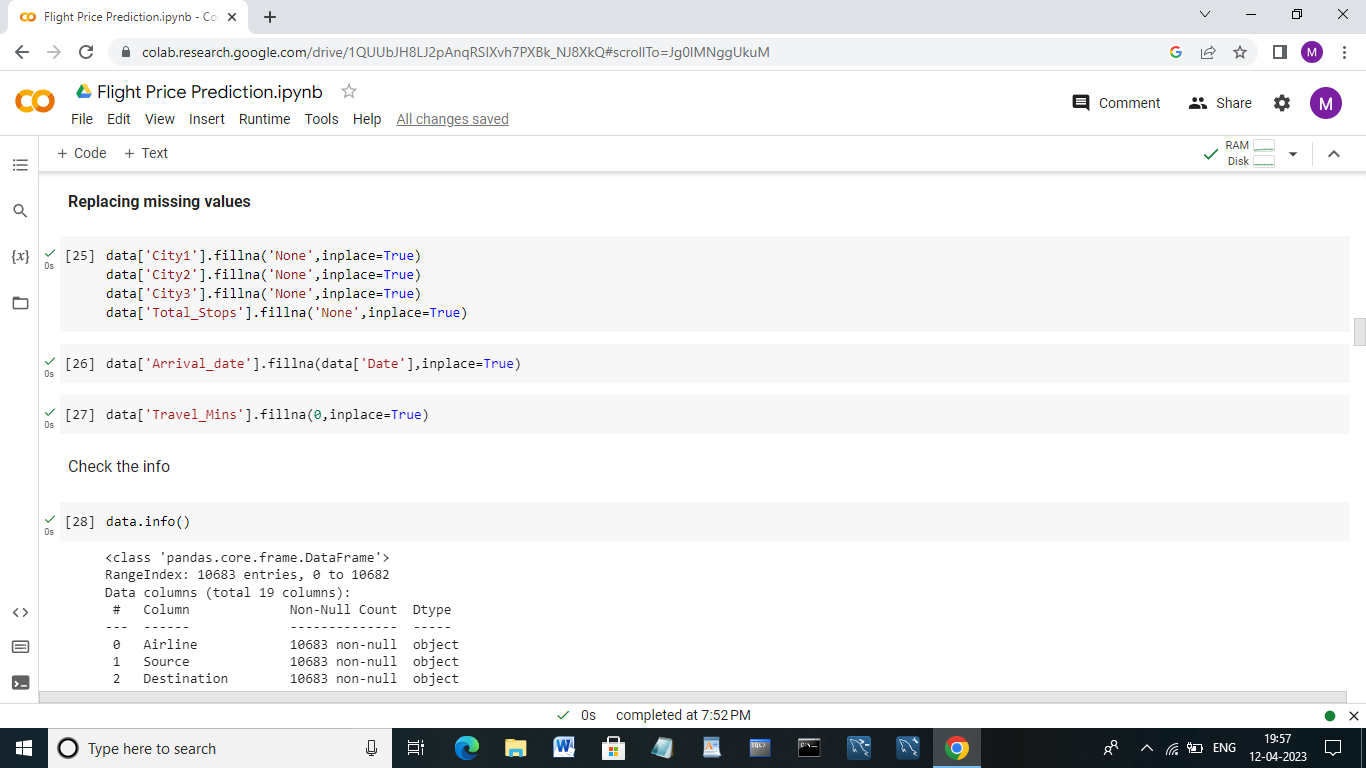
**Check the null values:**



**Drop the columns:**



**Replacing missing values:**

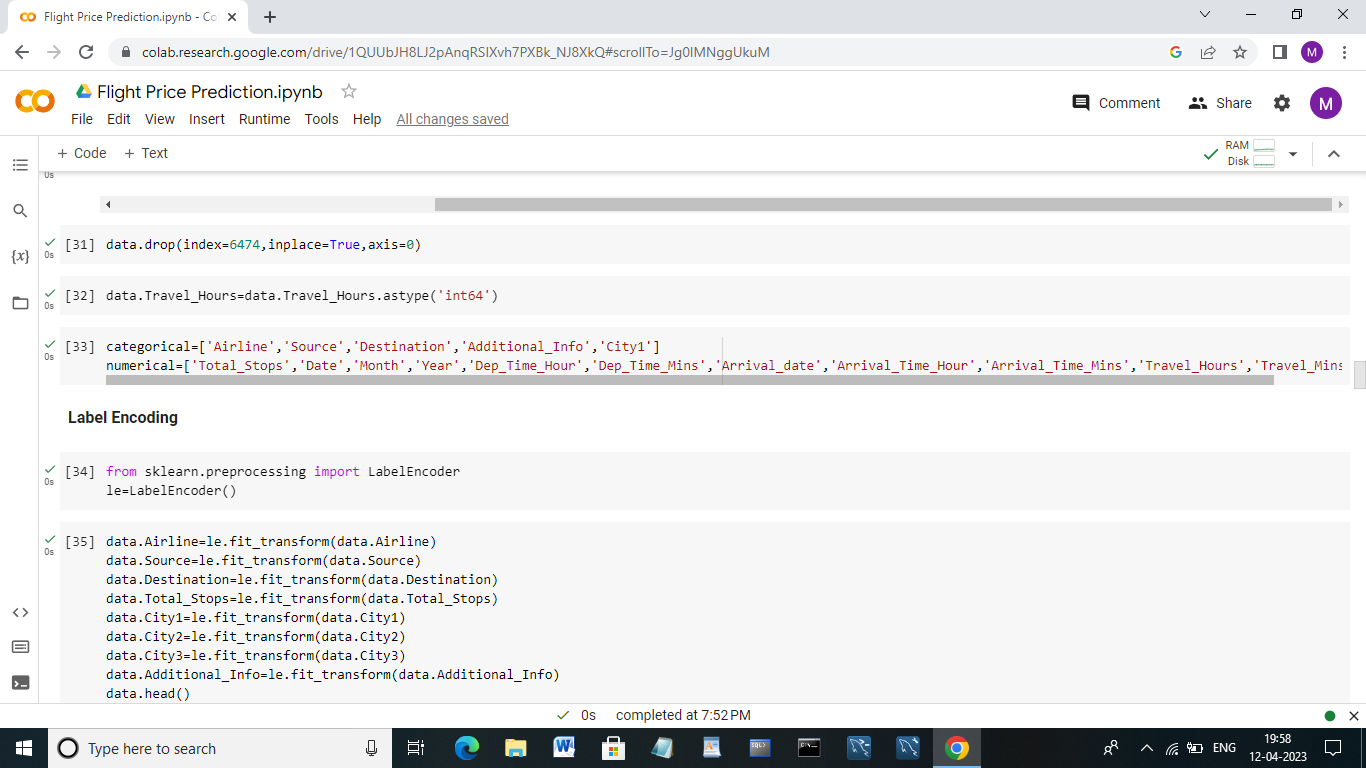


**Check Info:**

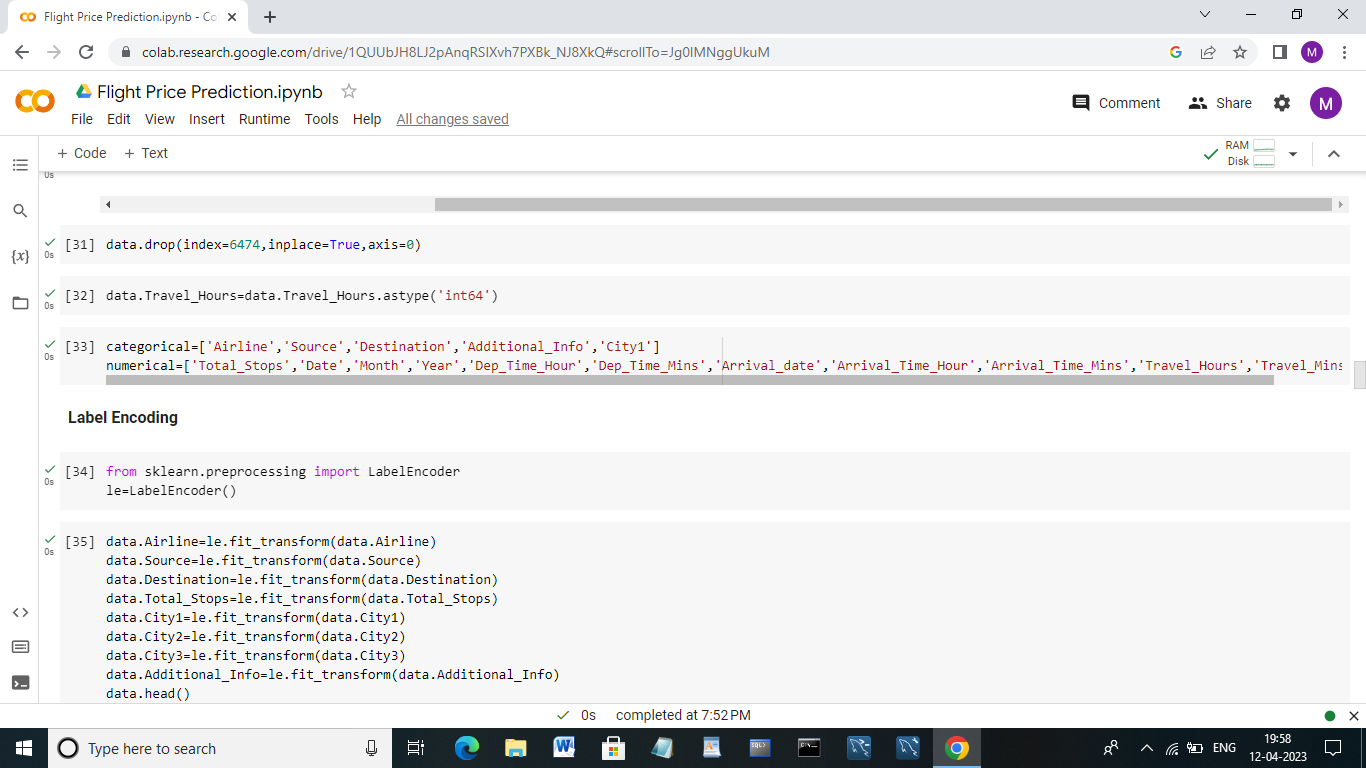
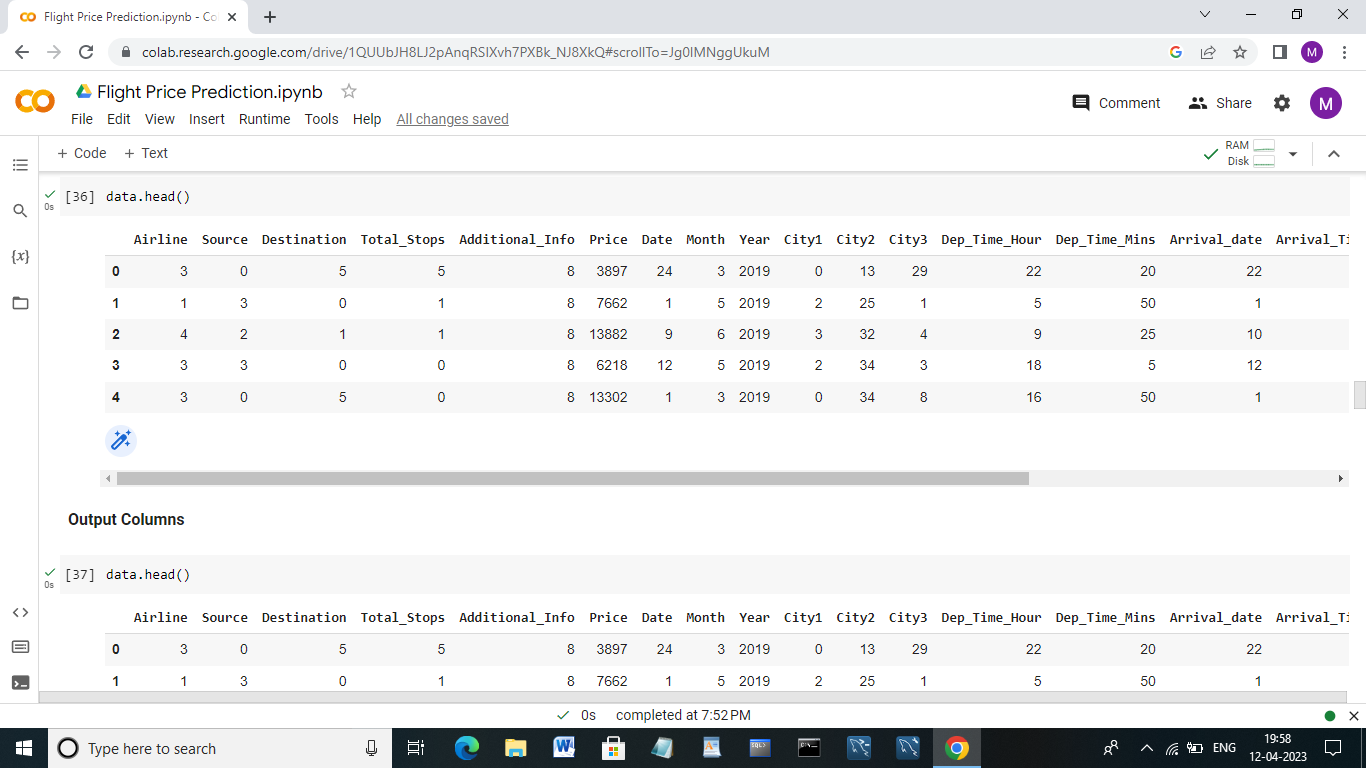
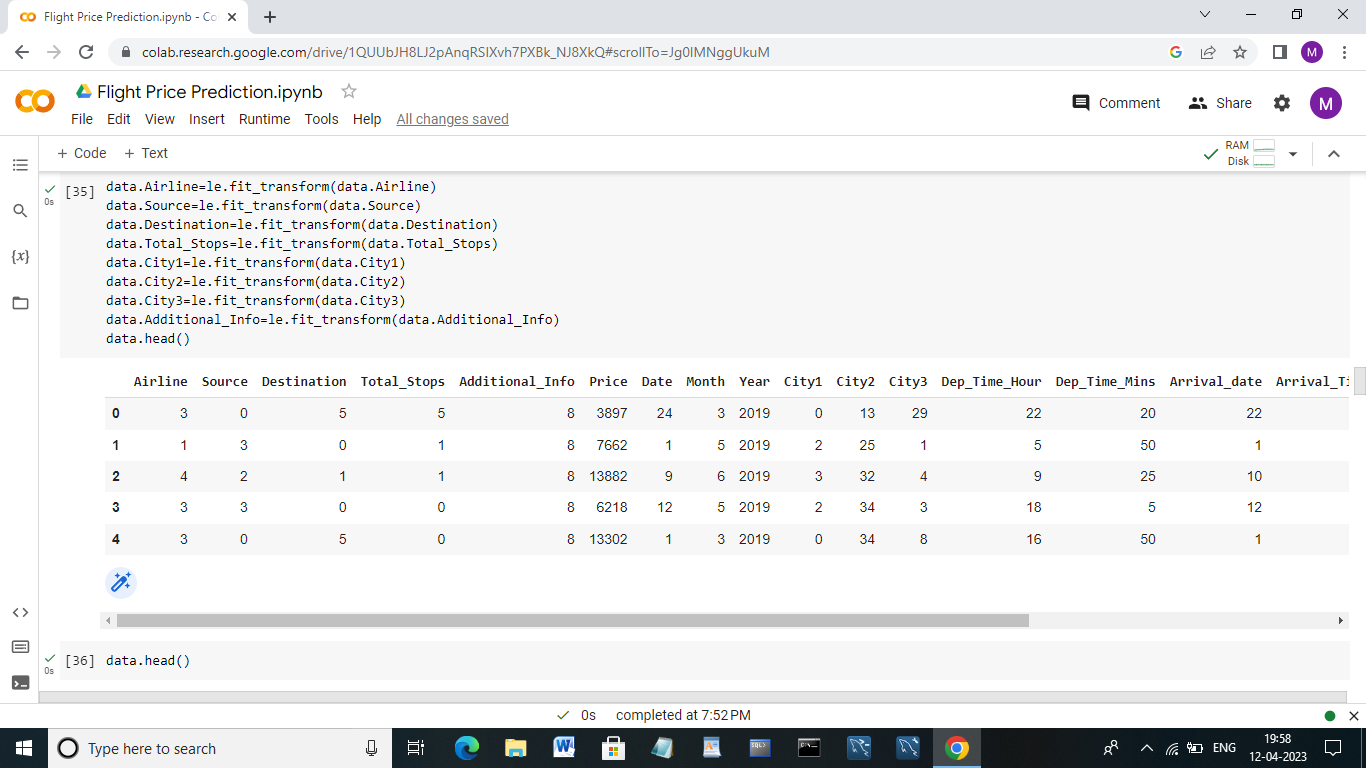


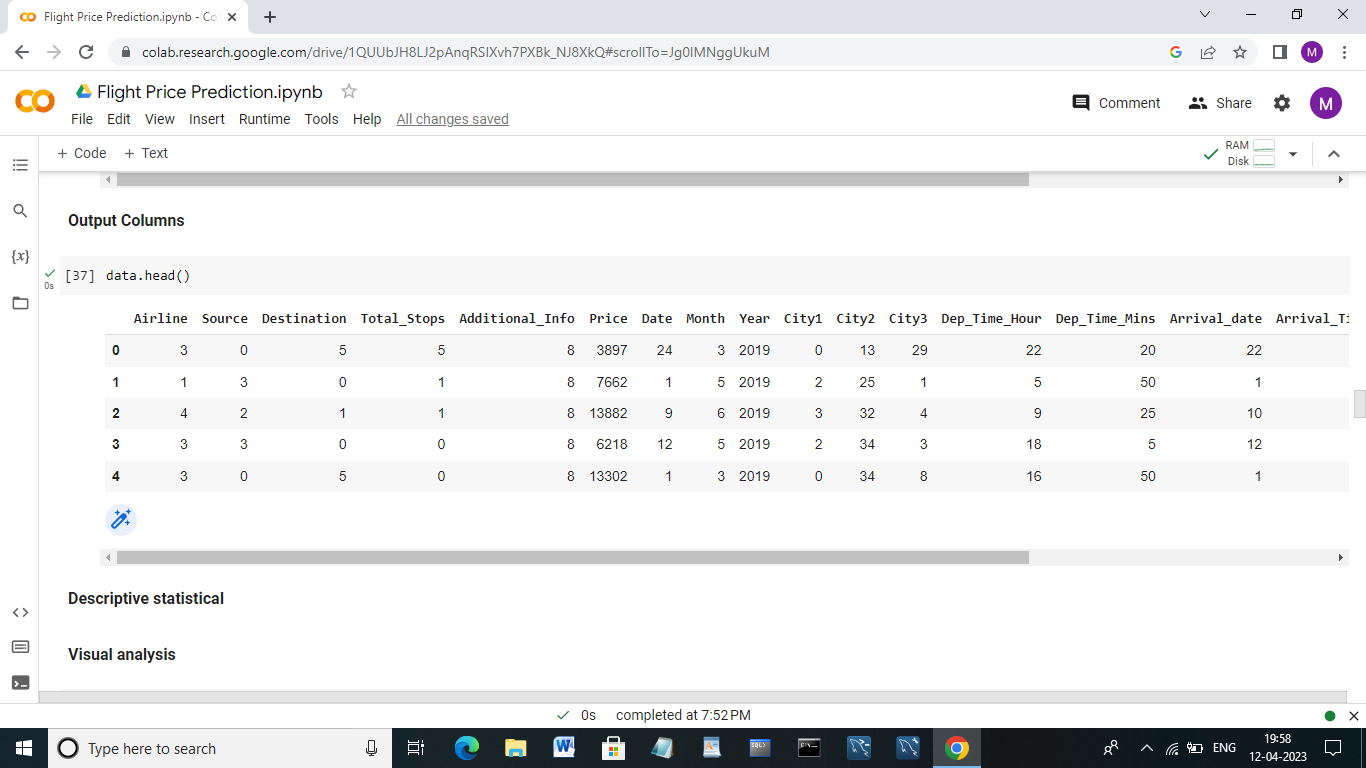
**Change the Datatype:**





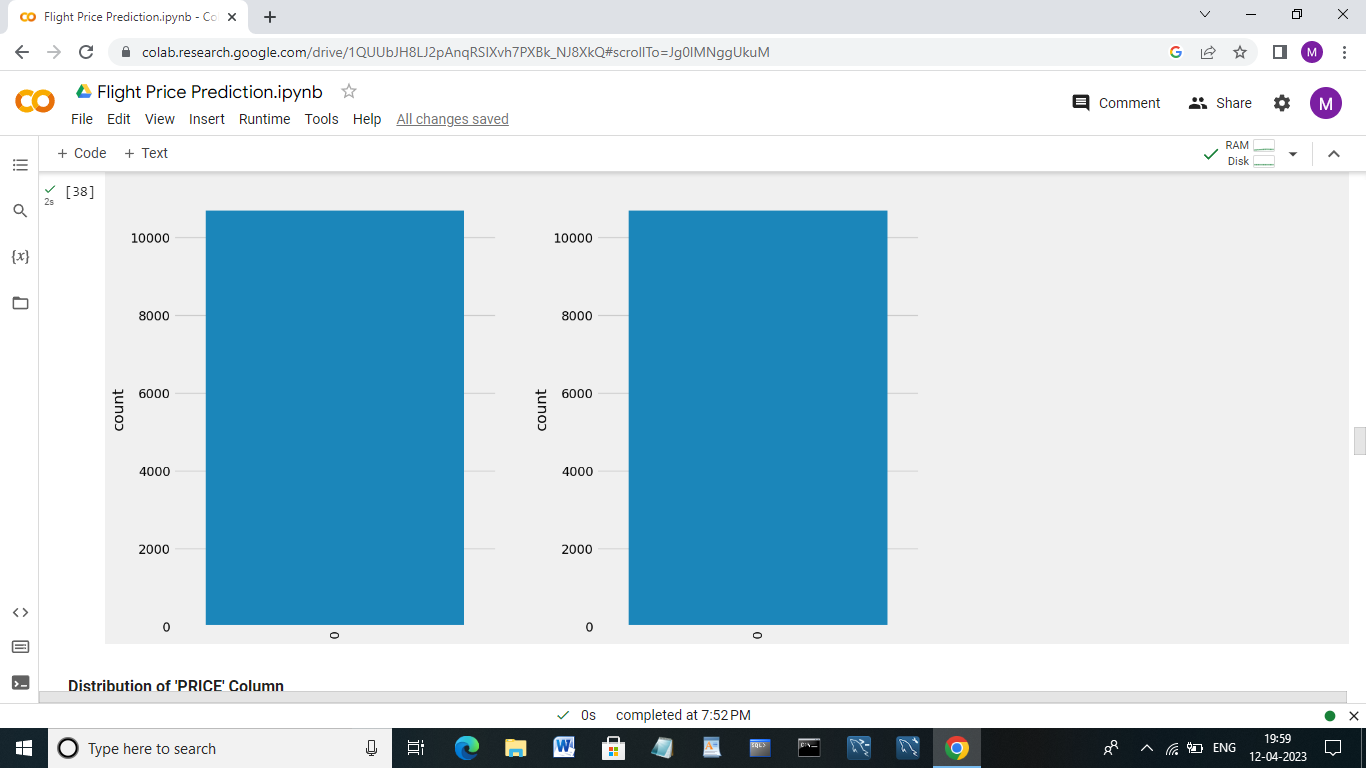
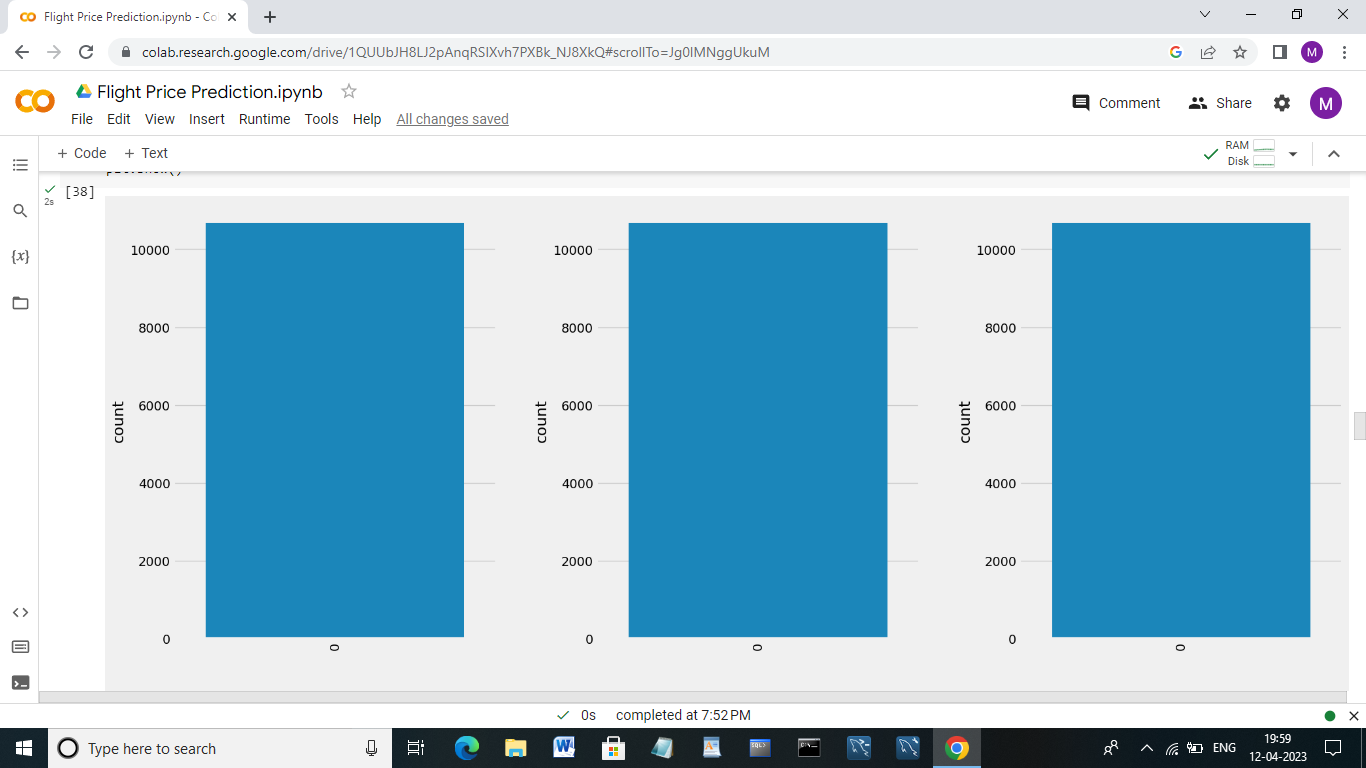
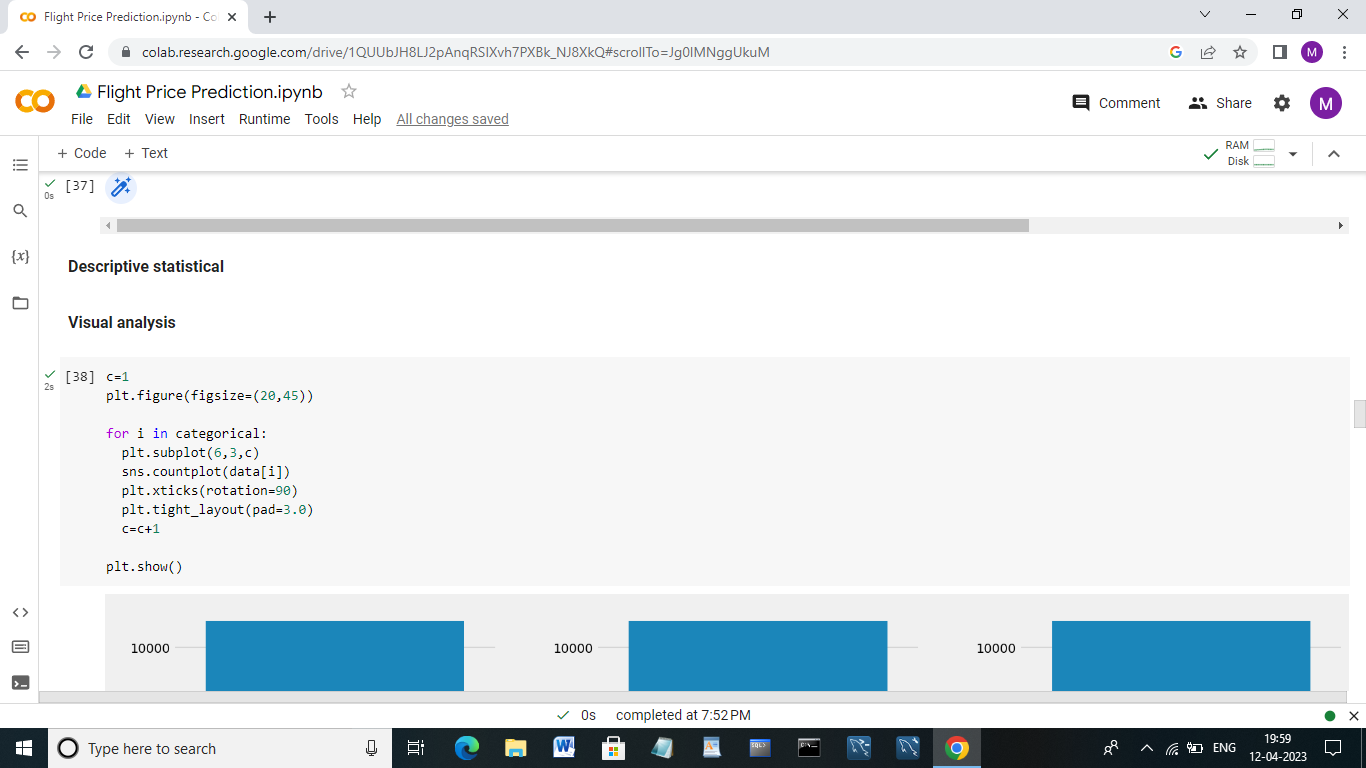
**Label Encoding:**

****

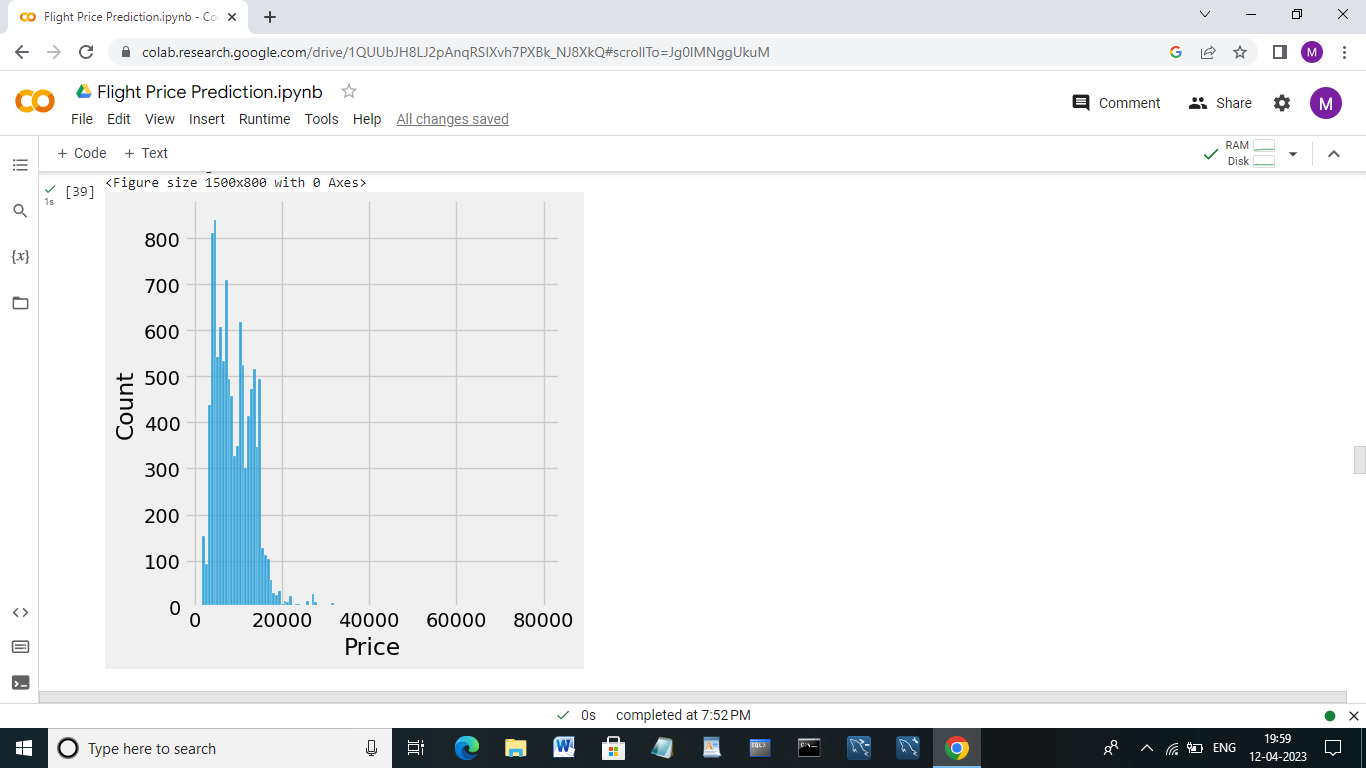
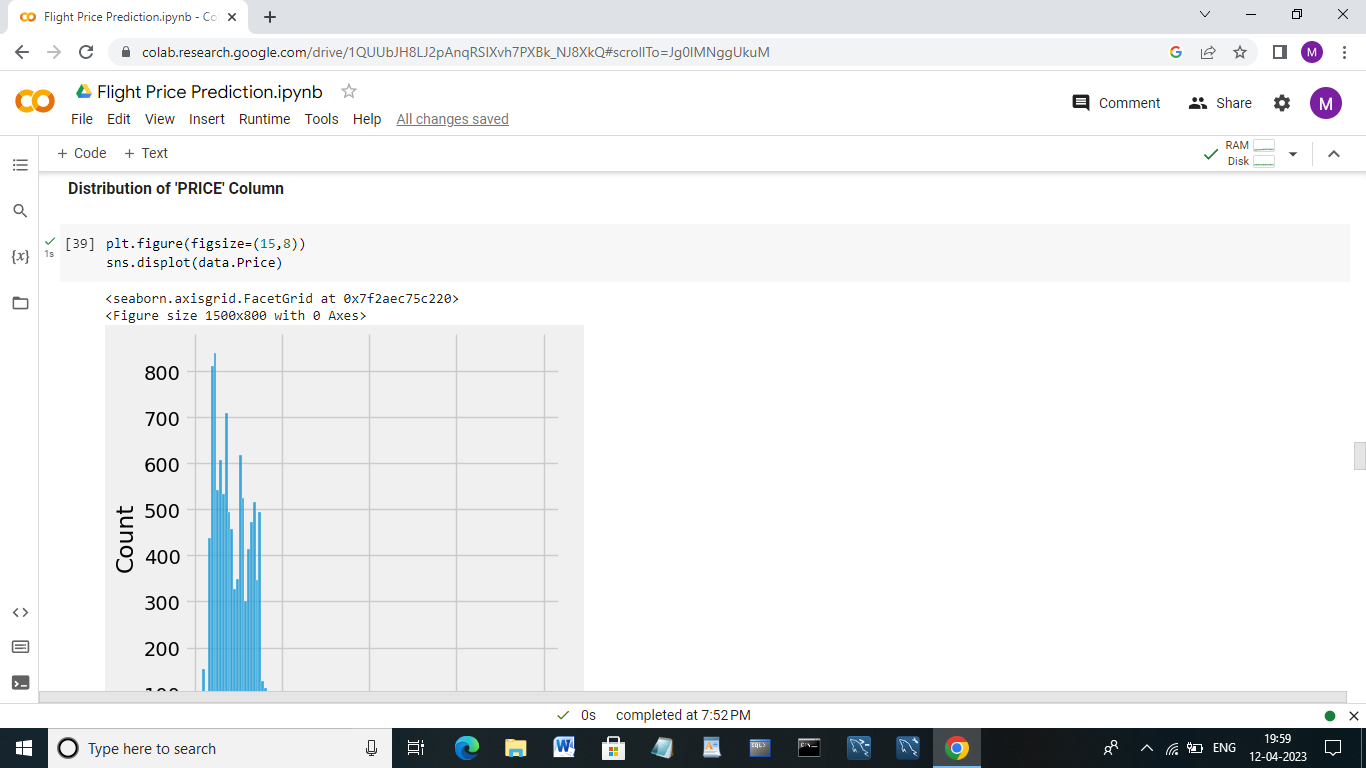
**Output columns:**

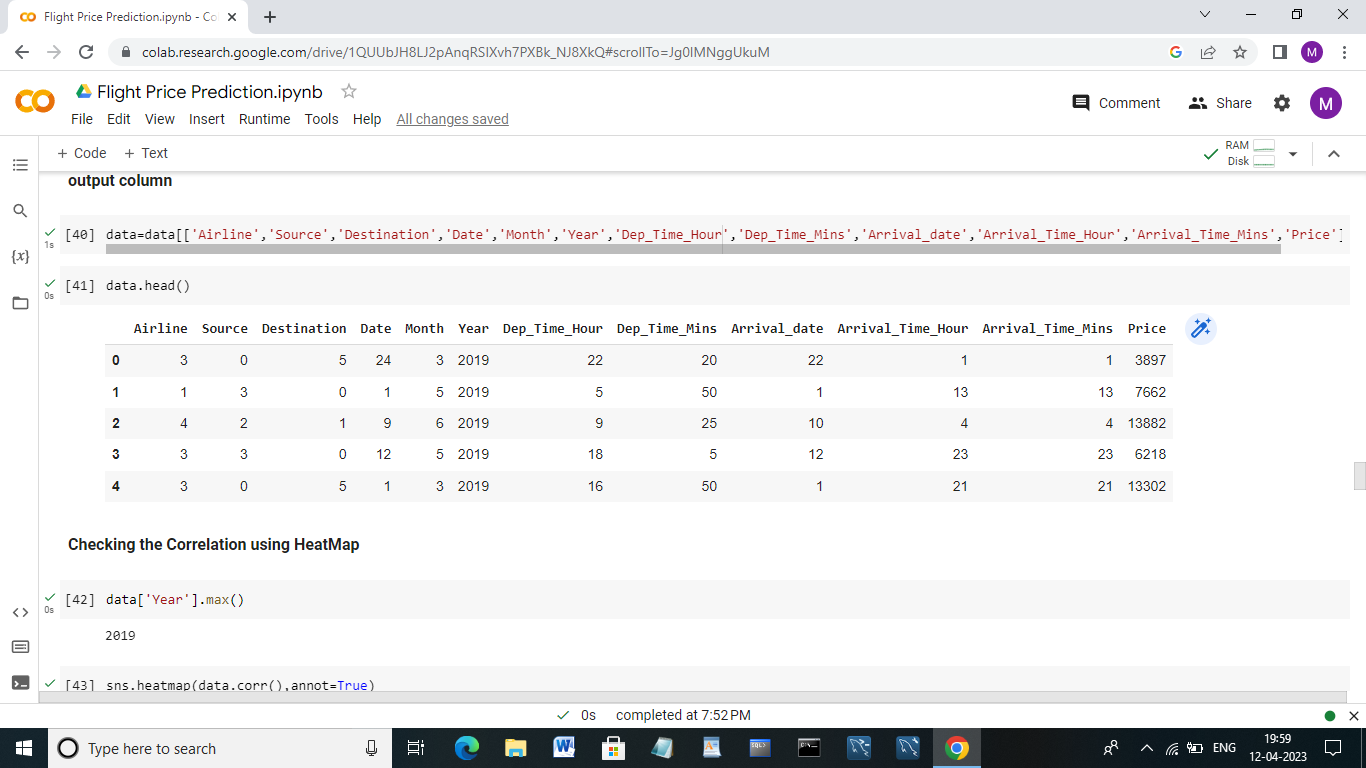
**Descriptive Statistical:**

**Visual Analysis:**

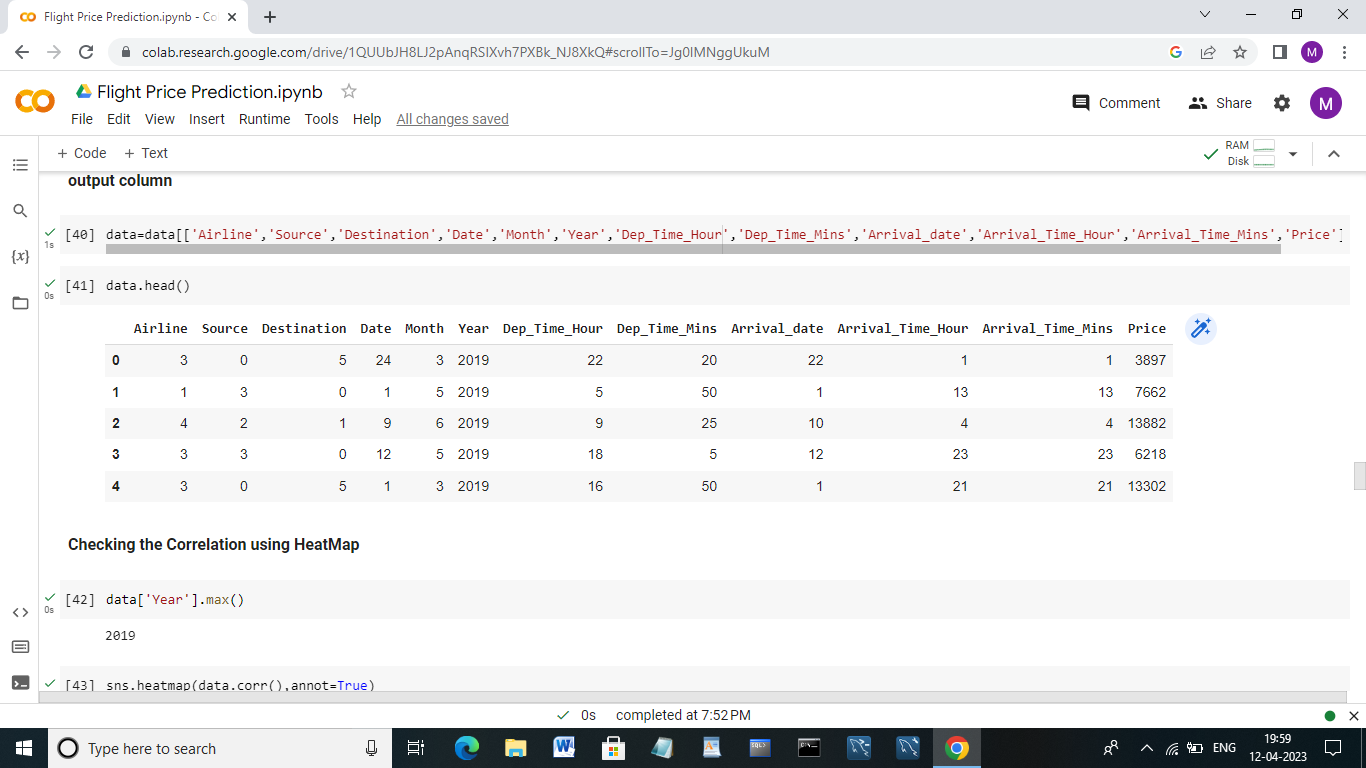


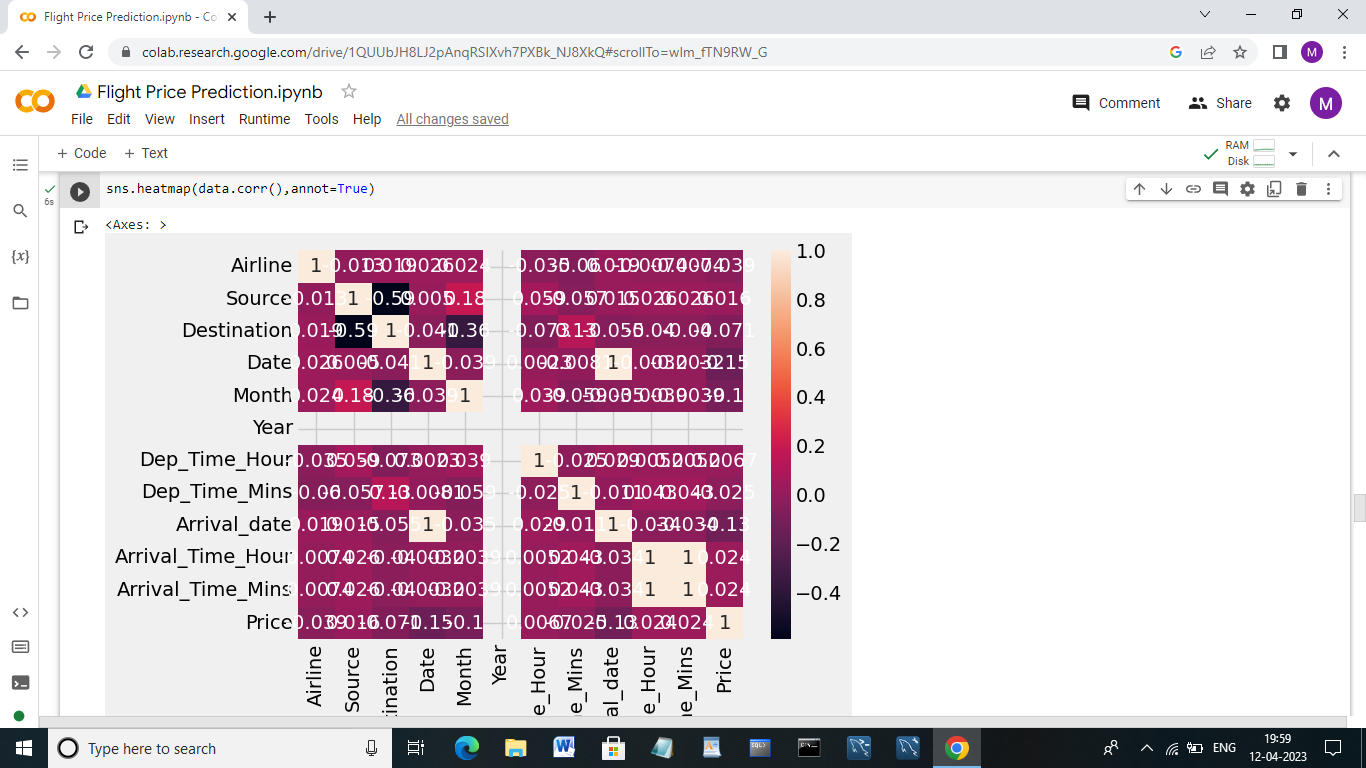
**Distribution of ‘PRICE’ Column:**

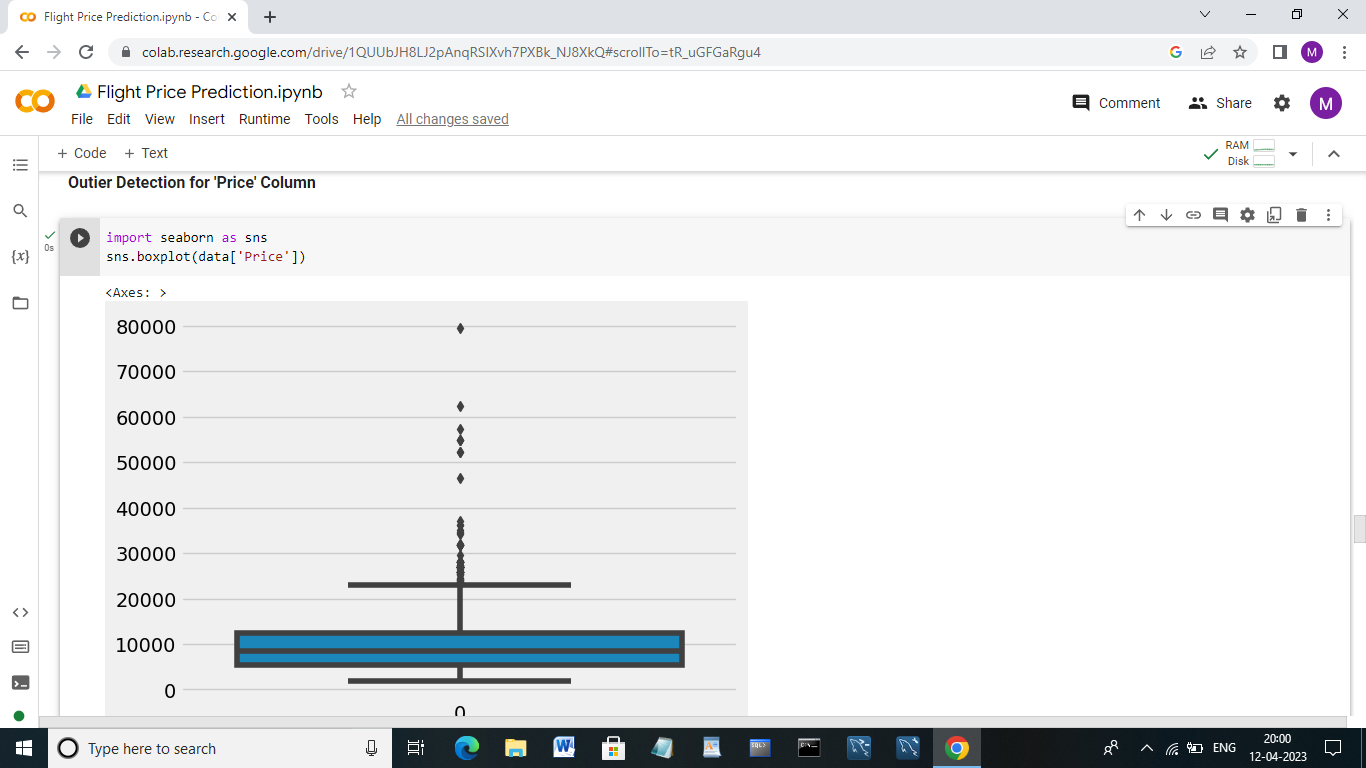


**Output Column:**

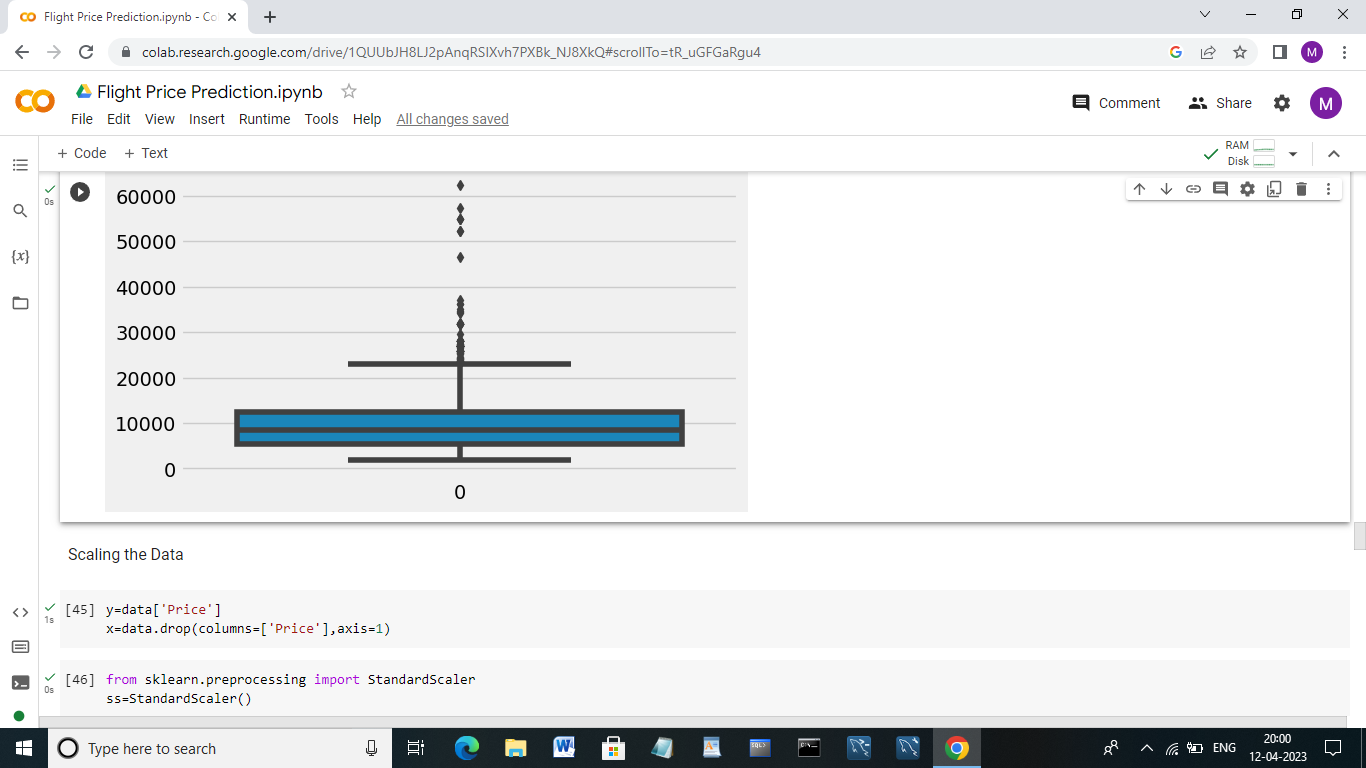
**Checking the correlation using HeatMap:**

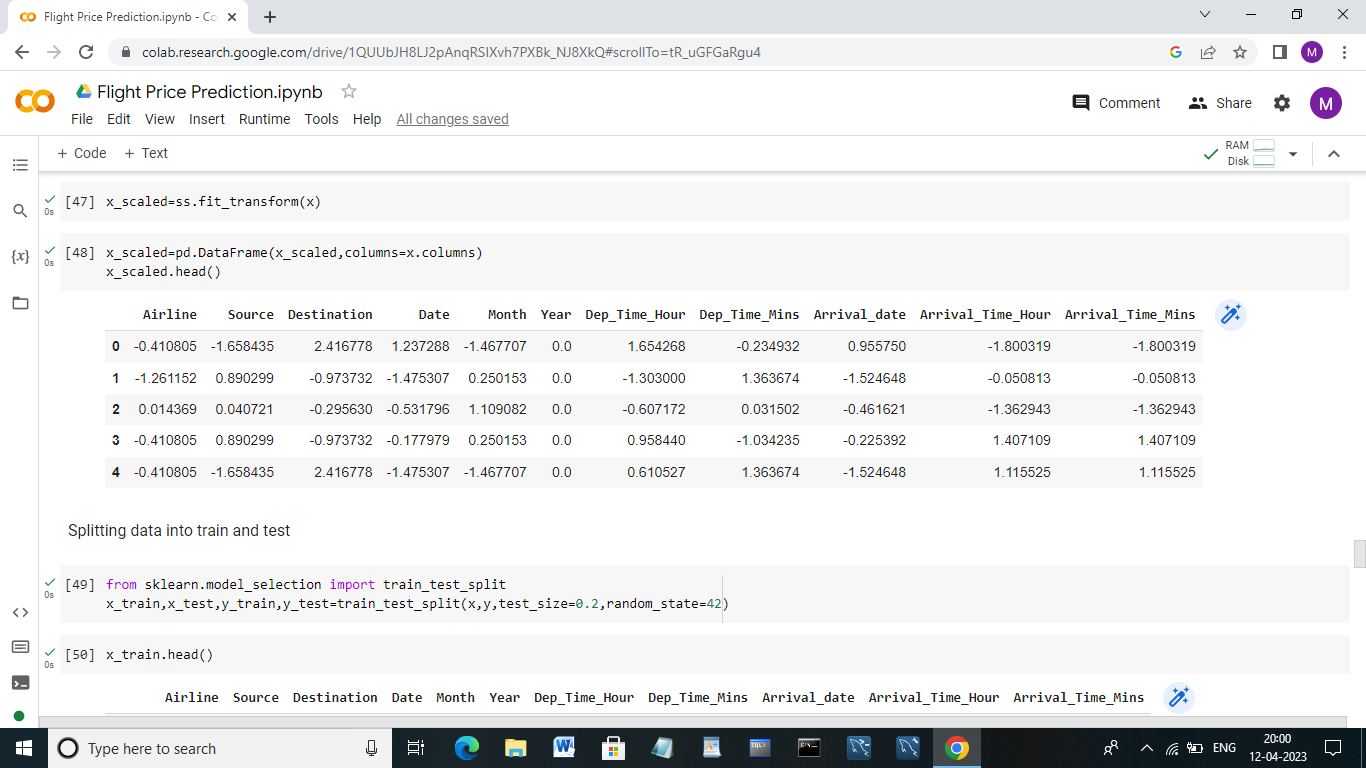
****



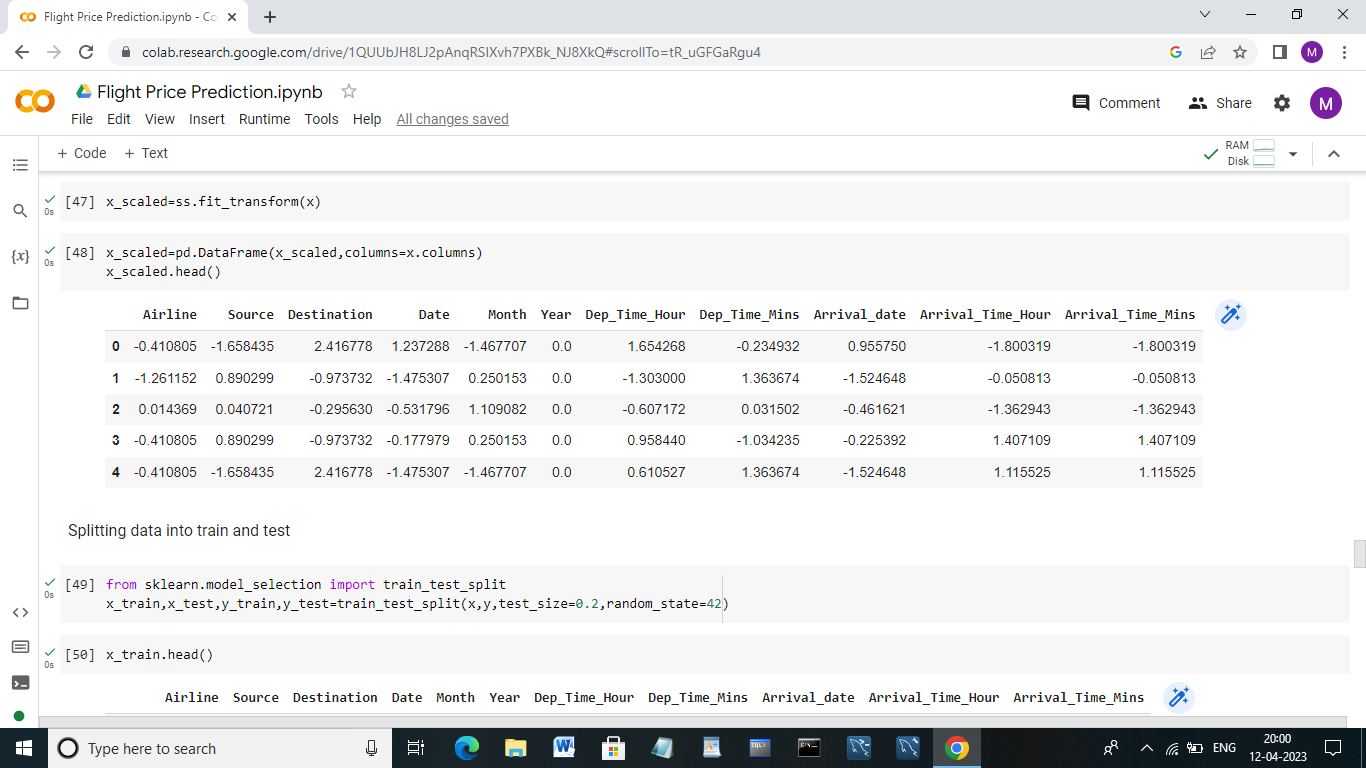
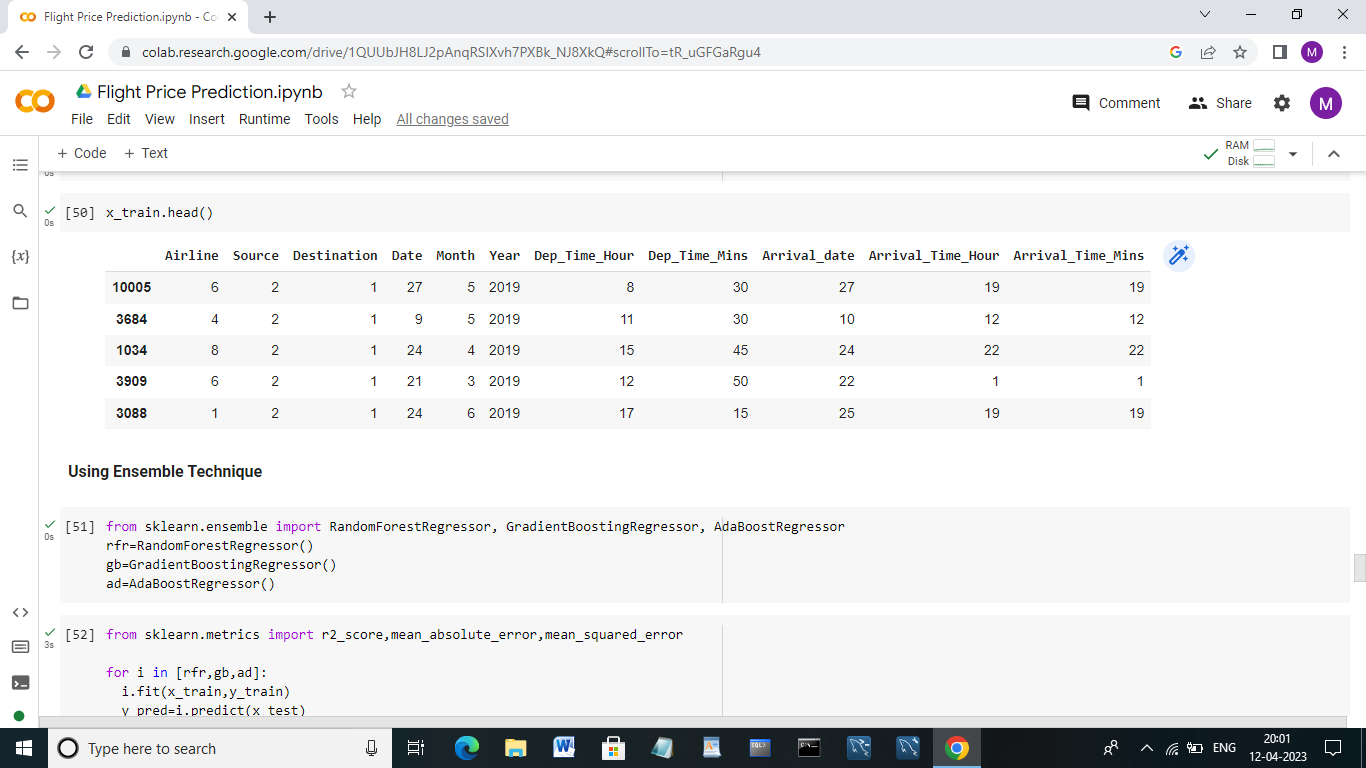
**Outlier Detection for ‘PRICE’ column:**

**Scaling the data:**

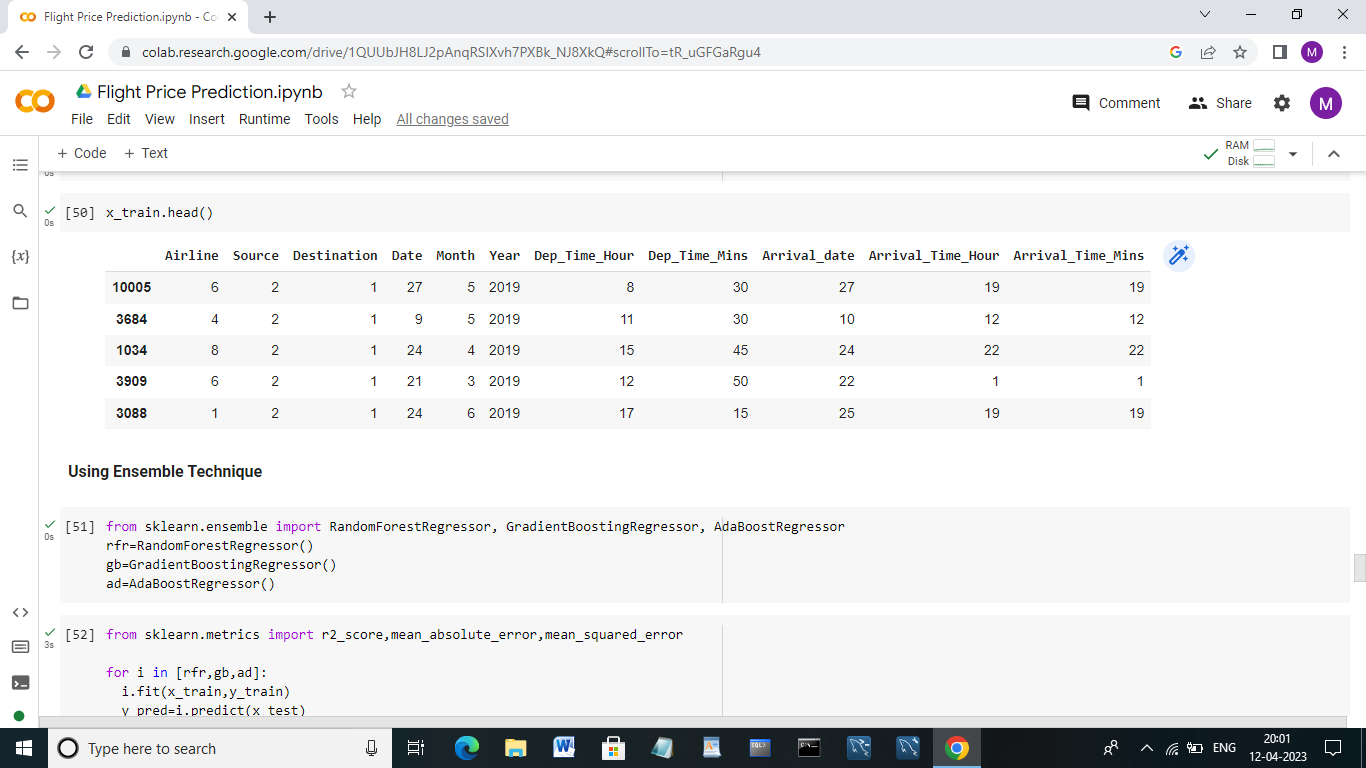
****

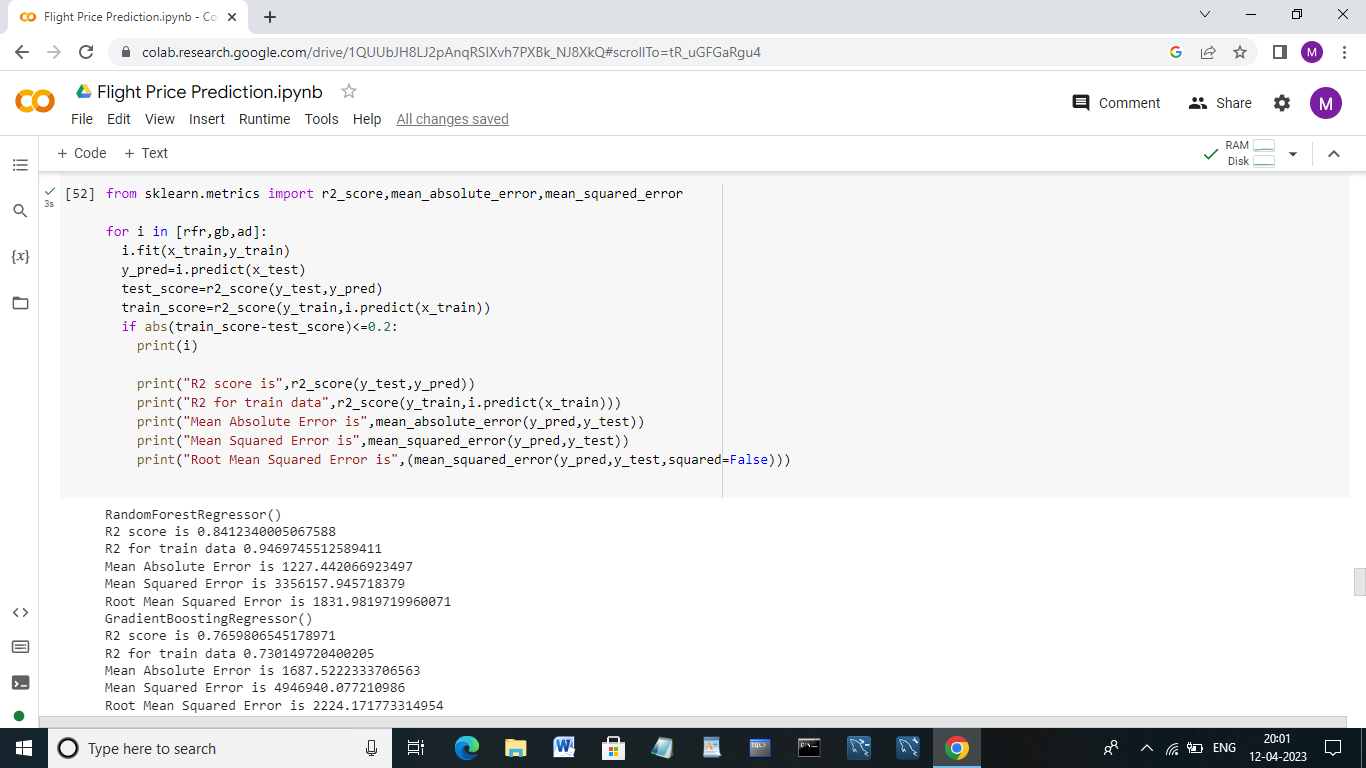


**Splitting data into train and test:**

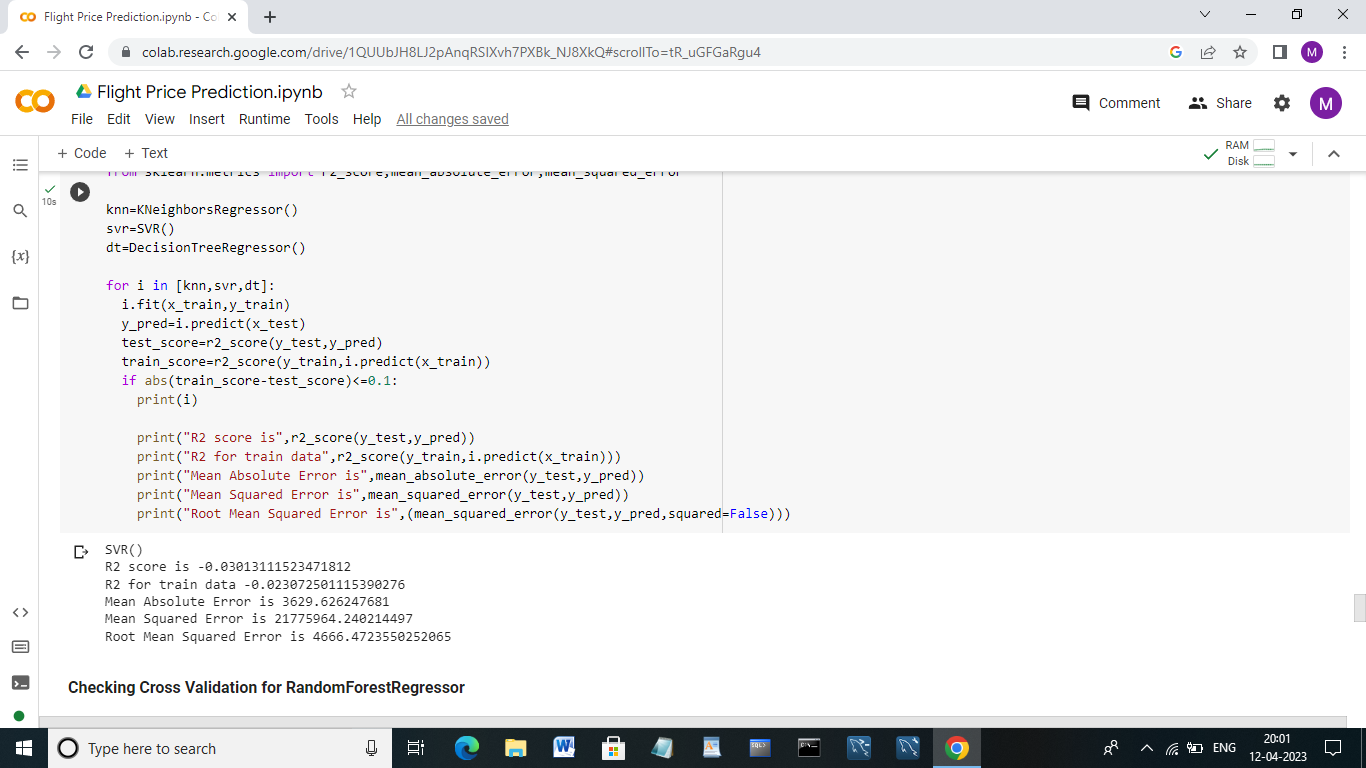
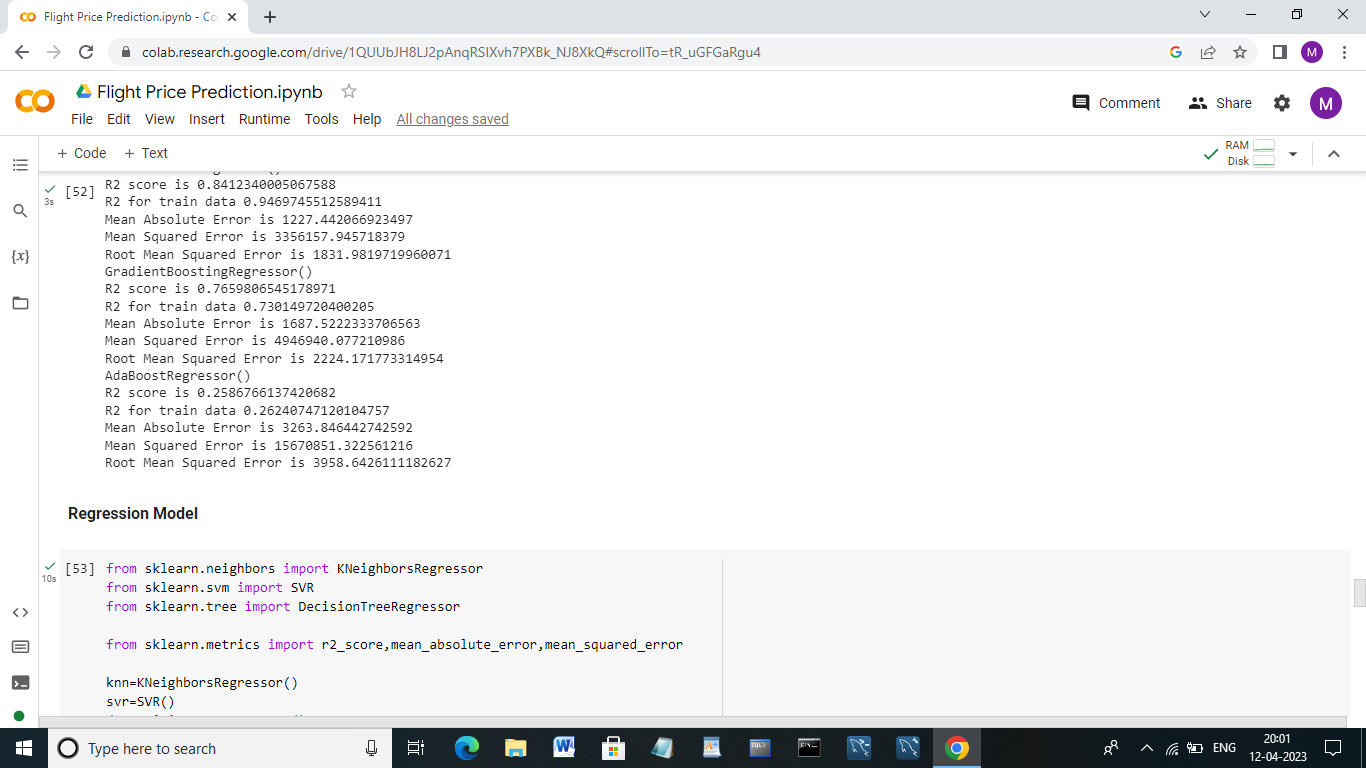
****

**Using Ensemble techniques:**

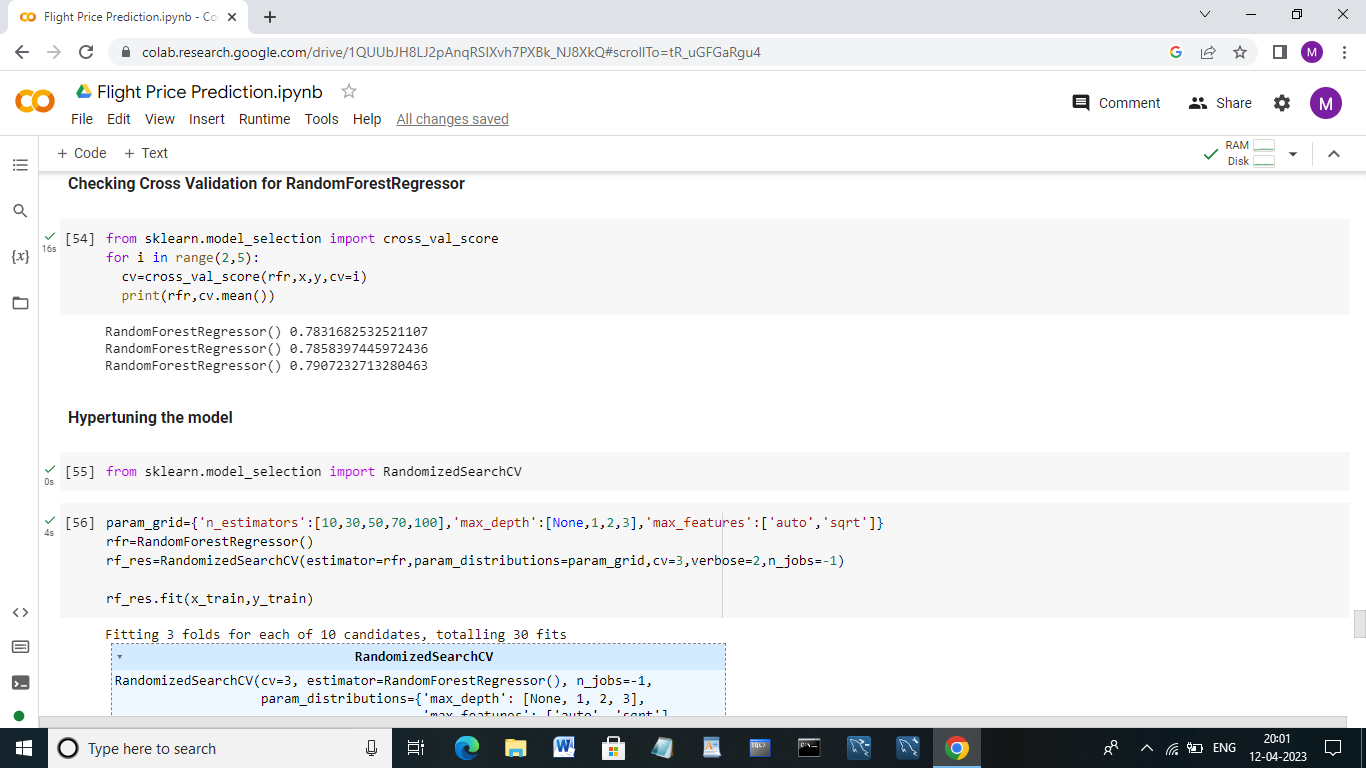
****



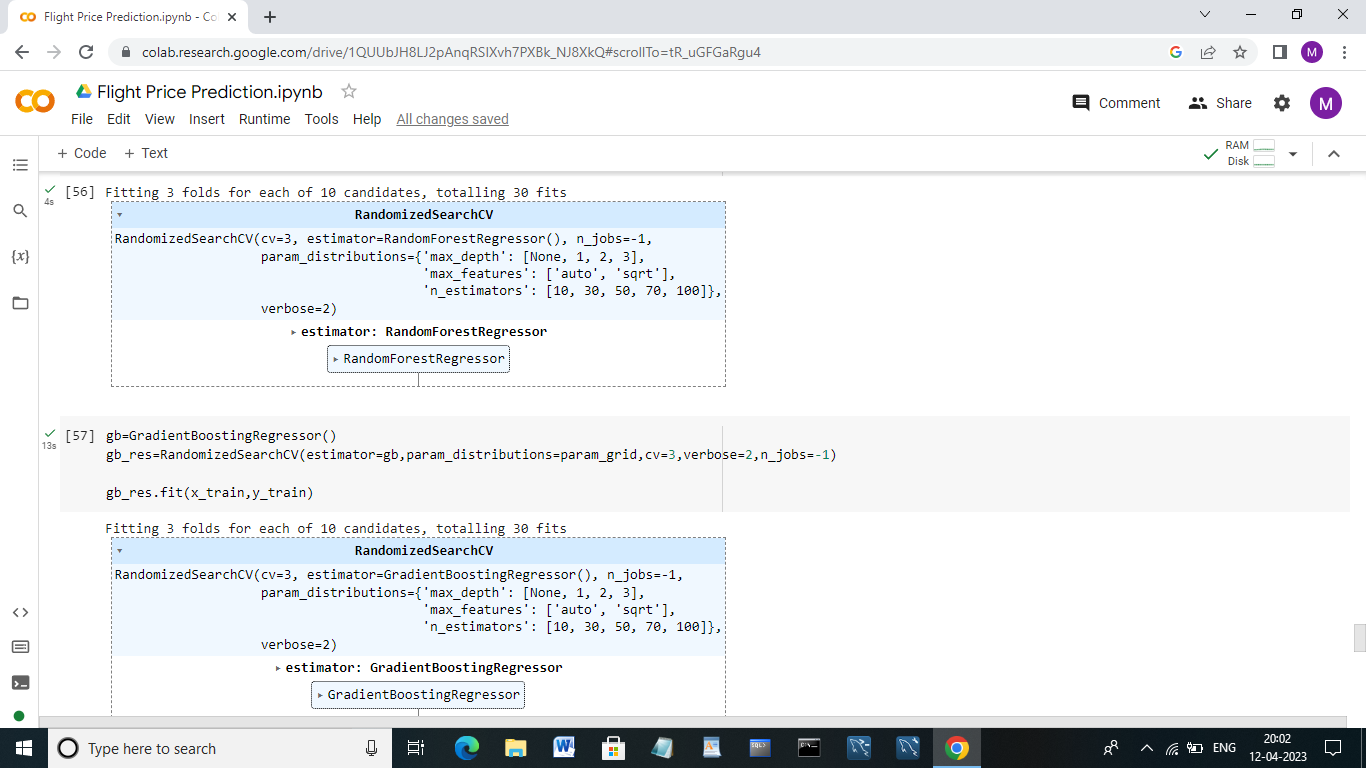
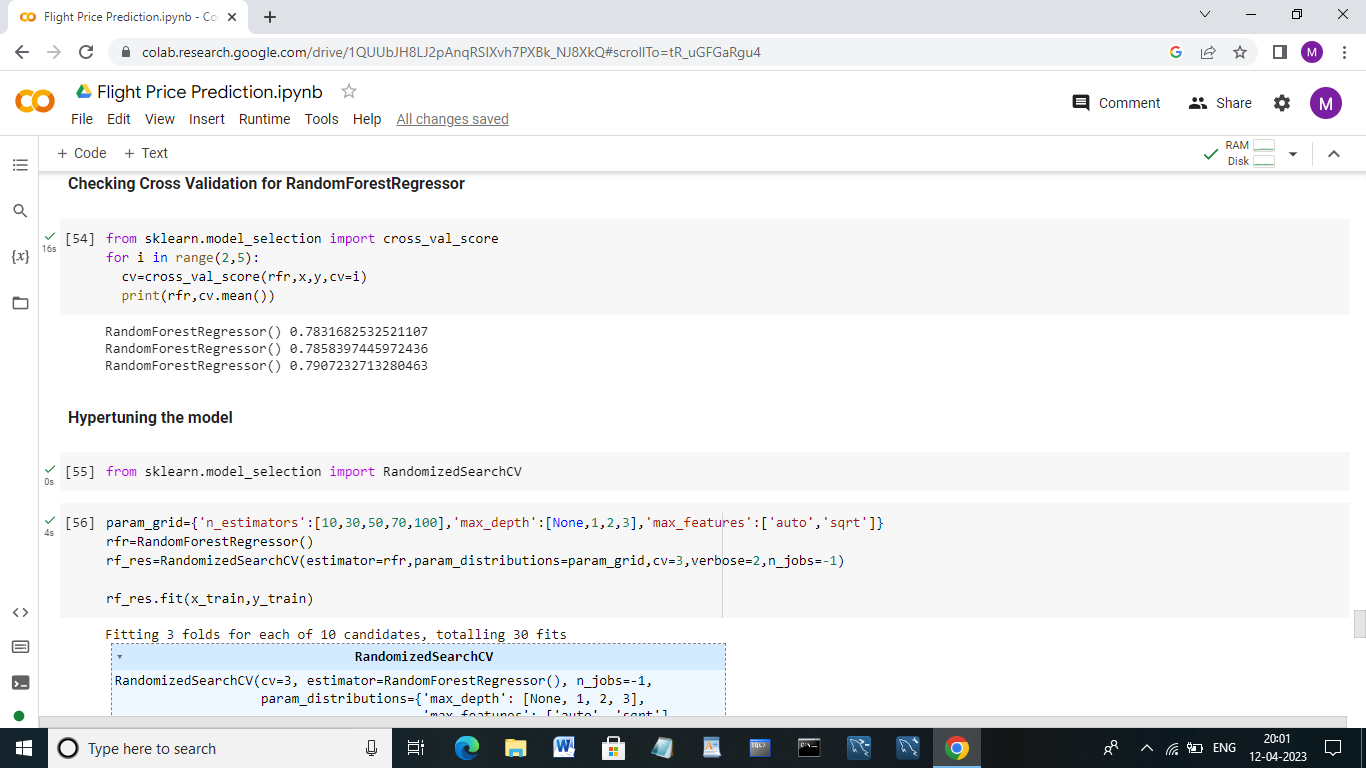
**Regression Model:**

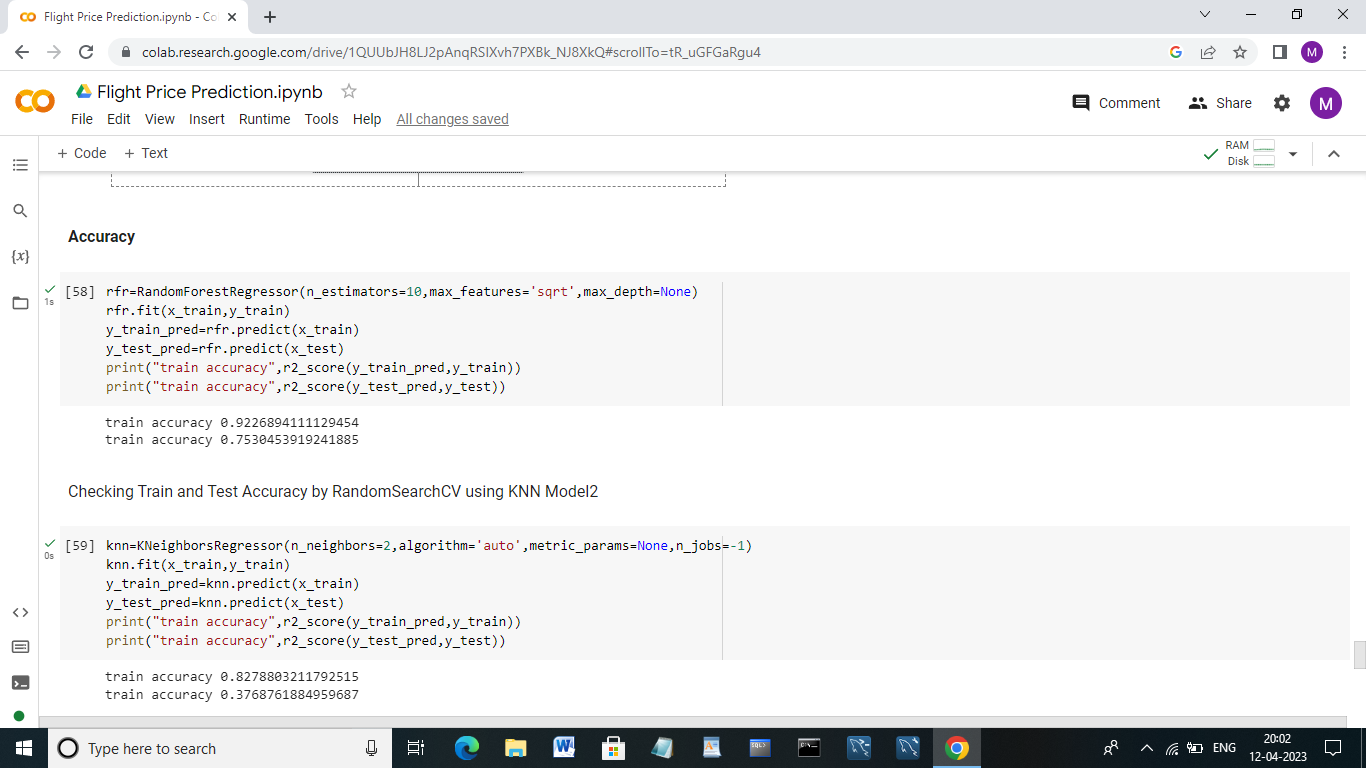


**Checking cross validation for RandomForestRegressor:**

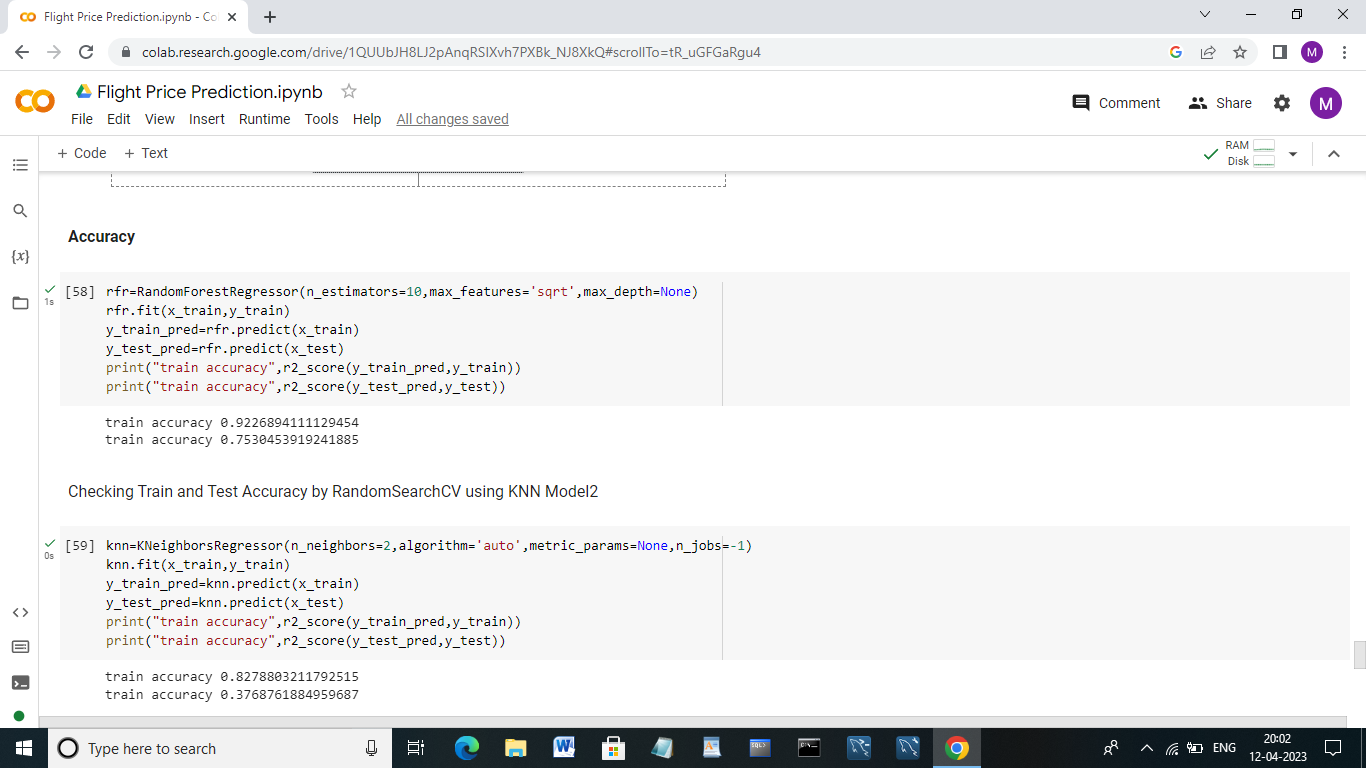
****

**Hypertuning the Model:**

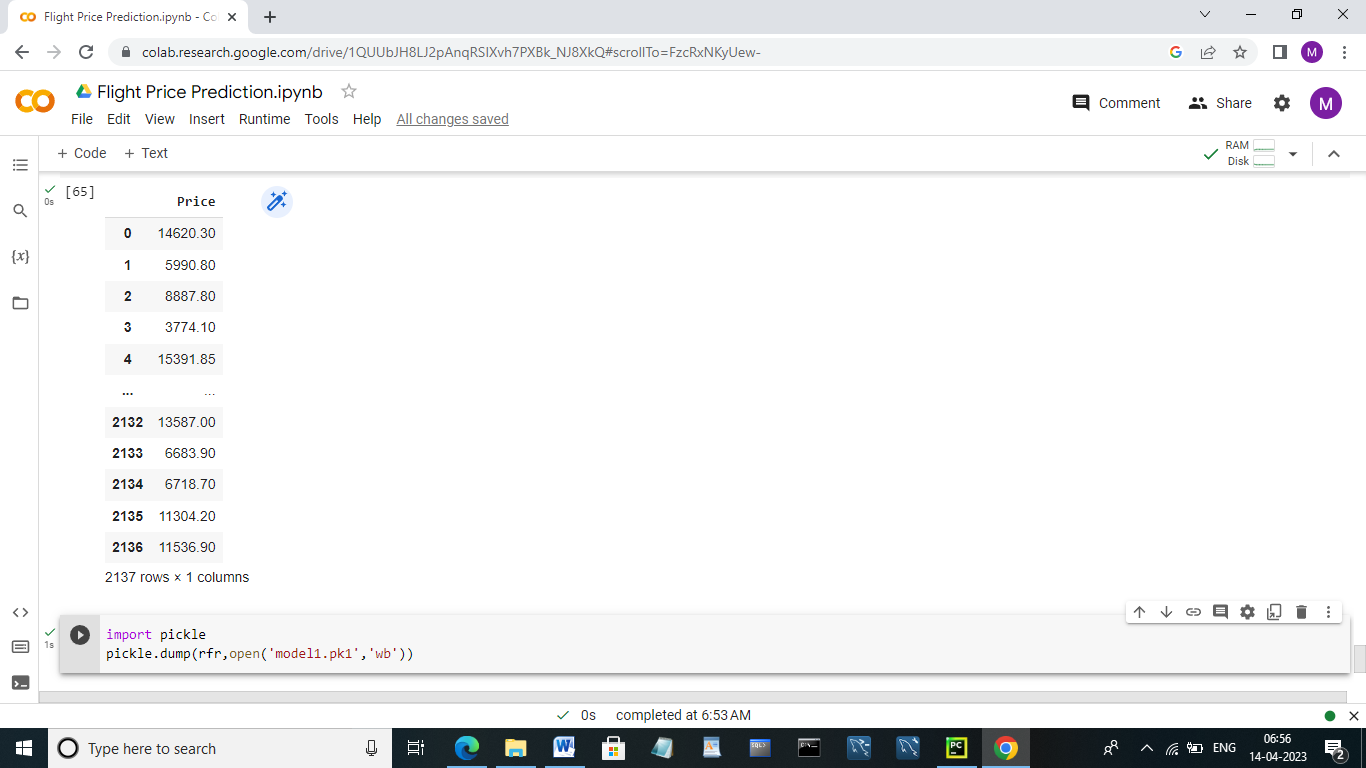
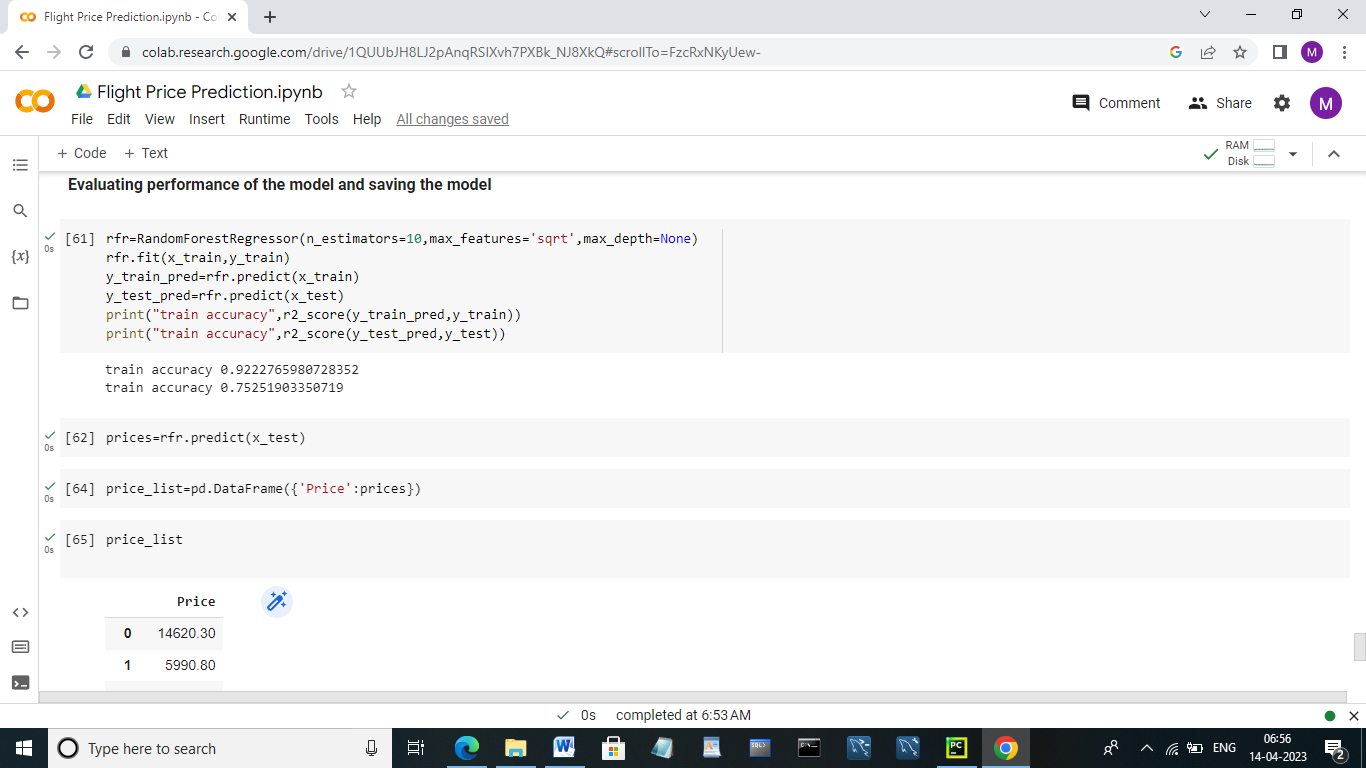


**Accuracy:**

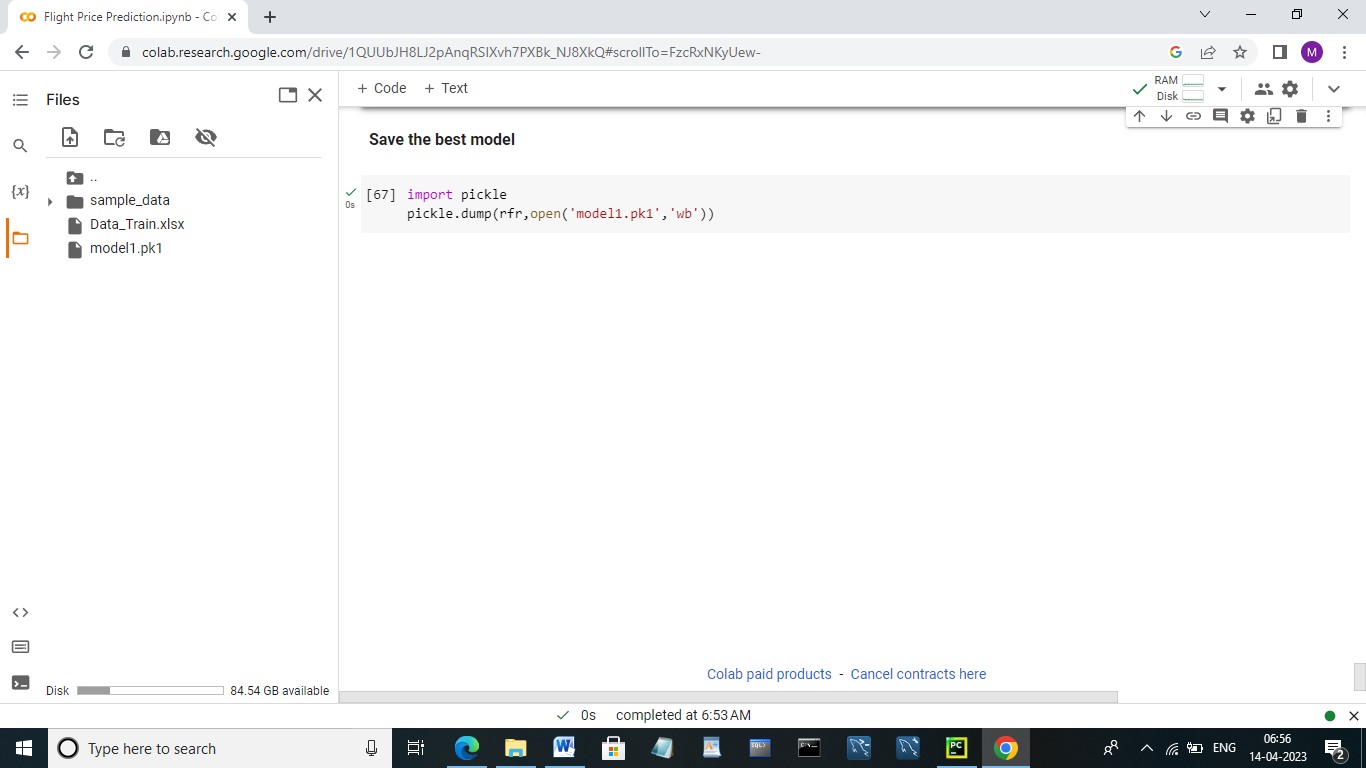
**Checking Train and Test Accuracy by RandomSearch using KNN Model2:**

****

**Evaluating performance of the model and saving the model:**



**Save the best model:**

****

**4. ADVANTAGES AND DISADVANTAGES**

**Advantages:**

* Traveler get the fare prediction handy using which it’s easy to decide the airlines.
* Saves time in searching/deciding for airlines.
* Flight price prediction tools can help you save money on your flights by alerting you to when the prices are expected to drop, allowing you to book your flights as the optimal time.
* Flight price prediction tools allow you to track multiple flights and airlines at once, so you can easily compare price and choose the best option for you.
* Rather than spending hours searching for the best deals on flights, you can use a flight price prediction tool to do the work for you, giving you more time to focus on planning the rest of your trip.
* If you have flexible travel dates, flight price prediction tools can help you find the cheapest times to travel, allowing you to adjust your plans accordingly.
* Knowing that you got the best deal possible on your flights can provide a sense of satisfaction and peace of mind, making your travel experience more enjoyable overall.

**Disadvantages:**

* Improper data will result in incorrect fare predictions.
* Flight price prediction tools use historical data and algorithms to predict future prices, but there are many factors that can influence the accuracy of these predictions, including unexpected events such as natural disasters or political unrest.
* Not all airlines or flights are covered by flight price prediction tools, so travelers may not be able to get accurate predictions for every flight they are interested in.
* Flight price prediction tools can provide travelers with information about when to book their flights, but they cannot control the actual prices or availability of flights. This means that even if a prediction suggests that prices will drop in the future, there is no guarantee that this will happen.
* Some travelers may become too reliant on flight price prediction tools and may miss out on good deals if they rely solely on these tools rather than actively searching for the best deals.
* Some flight price prediction tools may collect personal data from travelers, raising concerns about privacy and data protection.

**APPLICATIONS**

Flight price prediction has several applications for both individual travelers and travel companies:

* **Personal travel planning:** Individuals can use flight price prediction tools to find the best deals on flights for personal travel. By monitoring prices and predicting future changes, travelers can save money and plan their trips more effectively.
* **Business travel planning:** Companies can use flight price prediction tools to find the best deals on flights for business travel. This can help reduce travel costs and increase overall efficiency.
* **Travel agency services:** Travel agencies can use flight price predictions tools to offer their clients the best possible deals on flights. This can help increase customer satisfaction and loyalty.
* **Airline revenue management:** Airlines can use flight price prediction tools to adjust prices in real-time based on predicted demand. This can help airlines optimize revenue and increase profitability.
* **Tourism industry:** Flight price prediction tools can help the tourism industry plan and promote travel packages and events more effectively. By understanding future price changes, travel companies can tailor their offerings to meet the needs of their customers.

**6. CONCLUSION**

Evaluating the algorithmic rule, a dataset is collected, pre-processed, performed data modelling and studied a value difference for the number of restricted days by the passengers for travelling. Machine Learning algorithms with square measure for forecasting the accurate fare of airlines and it gives accurate value of plane price ticket at limited and highest value. Information is collected from Kaggle websites that sell the flight tickets therefore restricting data which are often accessed. The results obtained by the random forest and decision tree algorithm has better accuracy, but best accuracy is predicted by random Forest algorithm to fitting/split the train & test data, (x\_train, y\_train) accuracy= 93%, (x\_test, y\_test)accuracy = 77%.

**7. FUTURE SCOPE**

In Upcoming days when huge amount of information is accessed as in detailed information in the dataset, the expected results in future are highly correct. For further research anyone desire to expand upon it ought to request different sources of historical data or be a lot of organized in collection knowledge manually over amount of your time to boot, a lot of different combination of plane are going to be traversed. There is whole possibility that planes differ their execution ideas consisting characteristics of the plane. At last, it is curious to match our model accuracy with that of the business models accuracy offered nowadays.