

**Flight Price Prediction Project.**

Submitted by:

*Assie Houphouet Olivier*

**ACKNOWLEDGMENT**

I want to say thank you to sir ***Keshav Bansal*** our mentor who give me this project and also all YouTube content creators and Google content creators. When I started the project it was something very confusing since there was not available dataset. This project help me a lot to understand several time while watching YouTube videos to solve a specific problem. The data is collected from India airline companies website using web scraping.

**INTRODUCTION**

* Business Problem Framing

Anyone who has booked a flight ticket knows how unexpectedly the prices vary. The cheapest available ticket on a given flight get more and less expensive over time and not way for the travellers to predict the flight.

* Conceptual Background of the Domain Problem

The success of the project will help all the travellers around the world to travel without fear , it will increase the number of travellers and also help the airline companies to manage perfectly the flight ticket price.

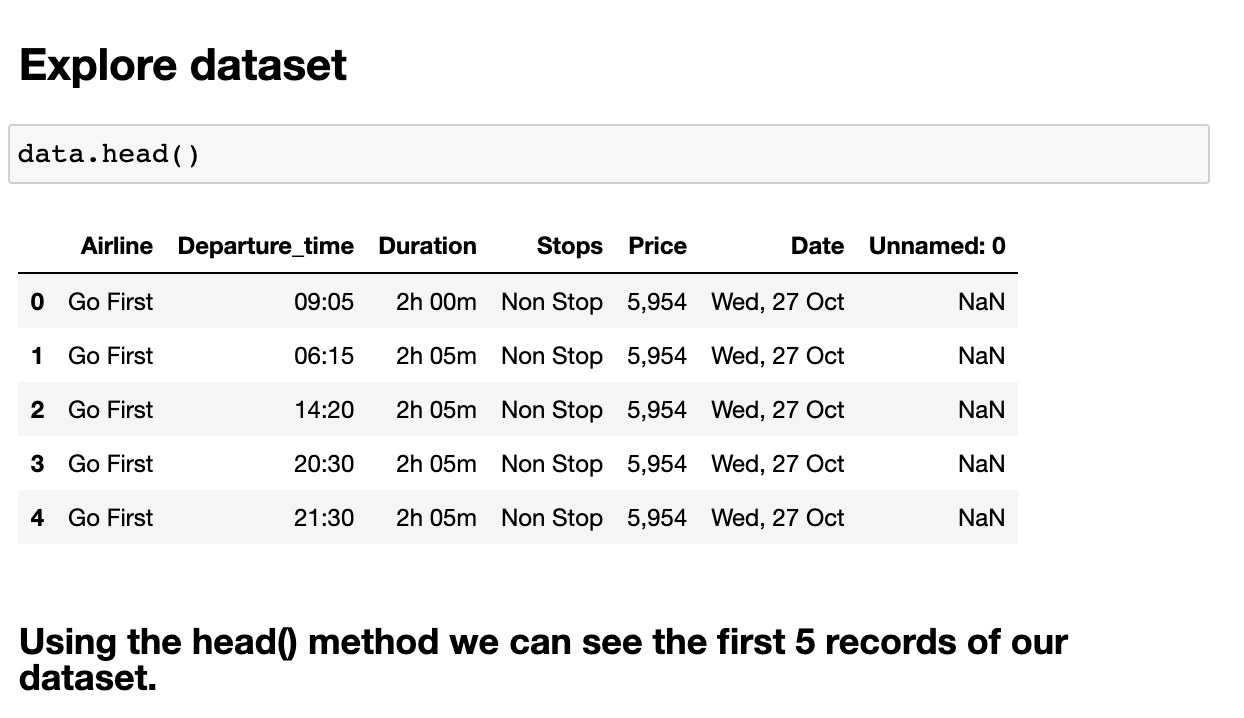
* Motivation for the Problem Undertaken

From the personal experience I can say that most people like to travel and visit the new places but their dreams are never realized because airfare is expensive and they don’t have money to afford it. So my motivation behind this project is to build a Machine Learning Model which can be used to predict the flight and guide the travellers when is good or not to book the flight ticket and help them make their dreams to be a really.

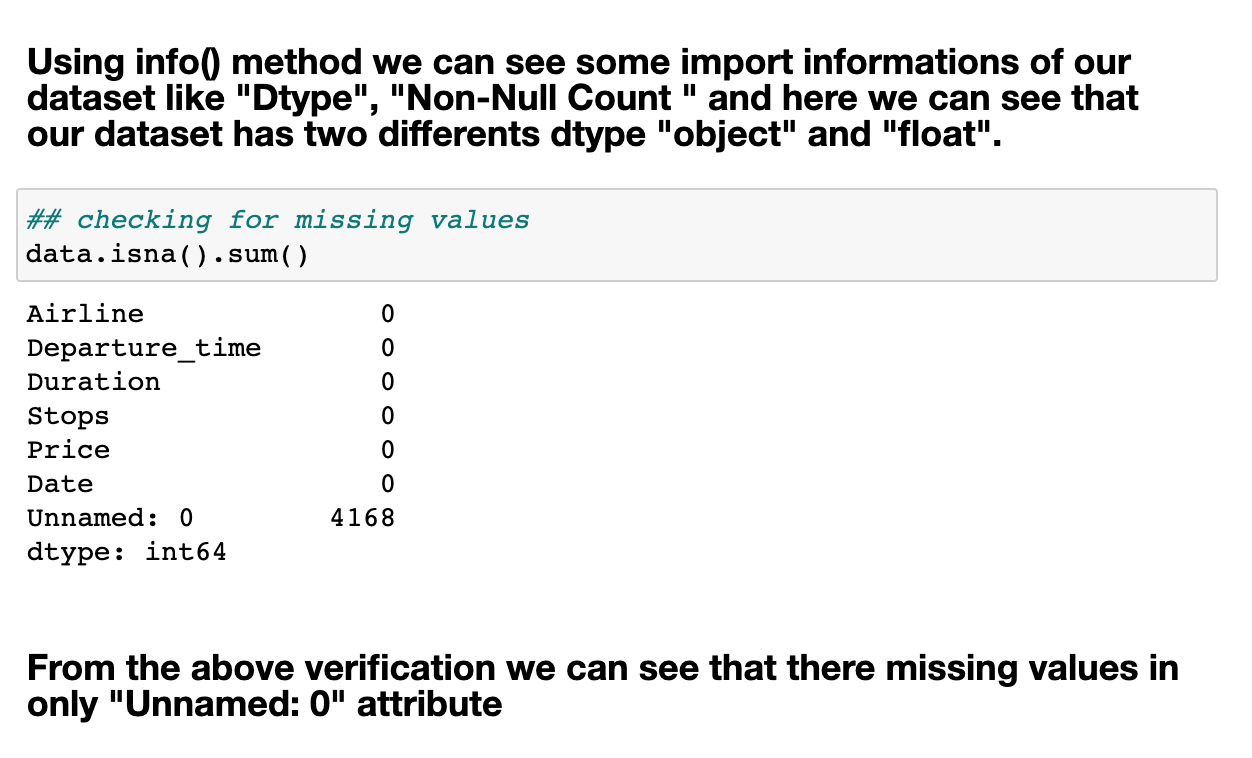
**Analytical Problem Framing**

* Mathematical/ Analytical Modelling of the Problem

First I is important to verified whether there missing values in the dataset and also the duplicated records issue after checking the shape of our train dataset.



****

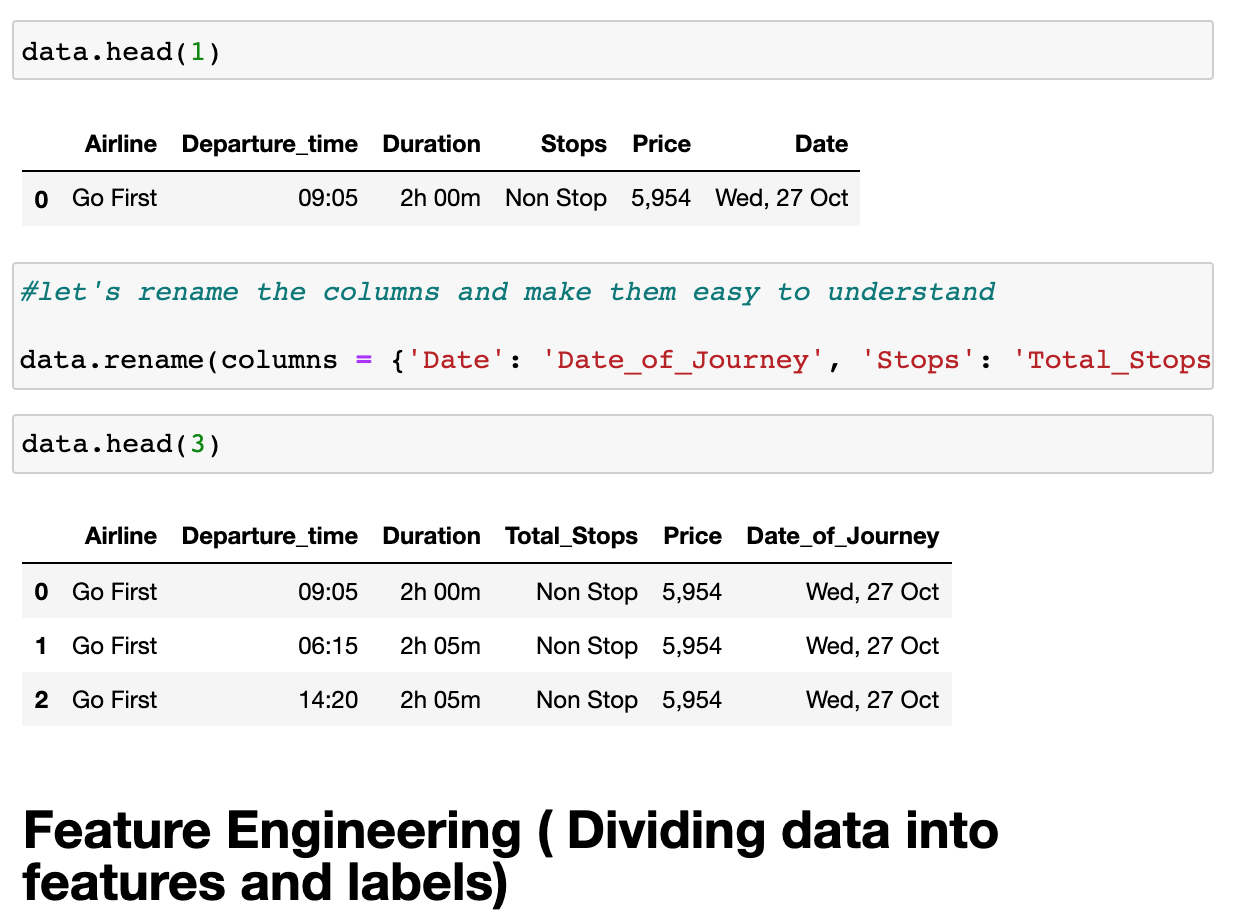
****

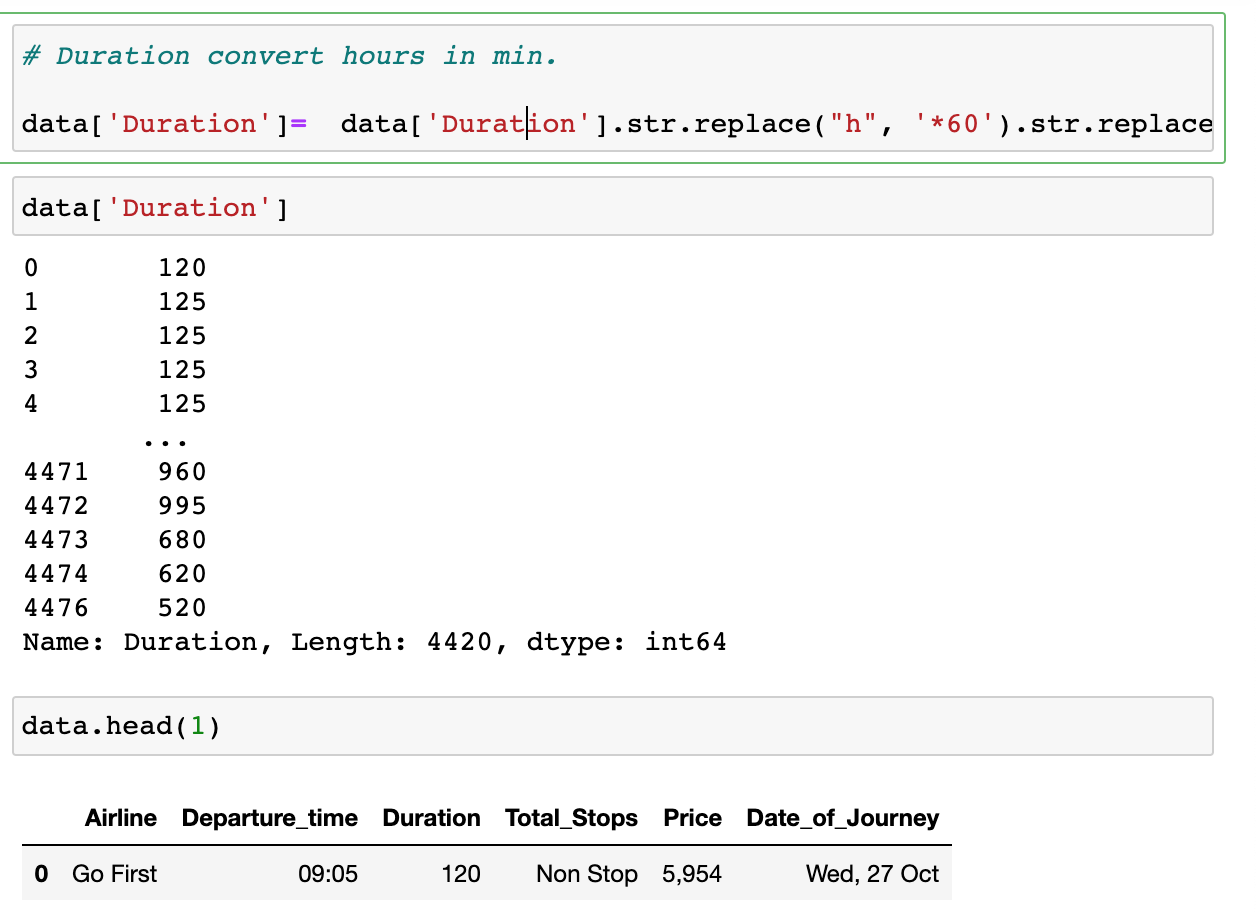
* Data Sources and their formats

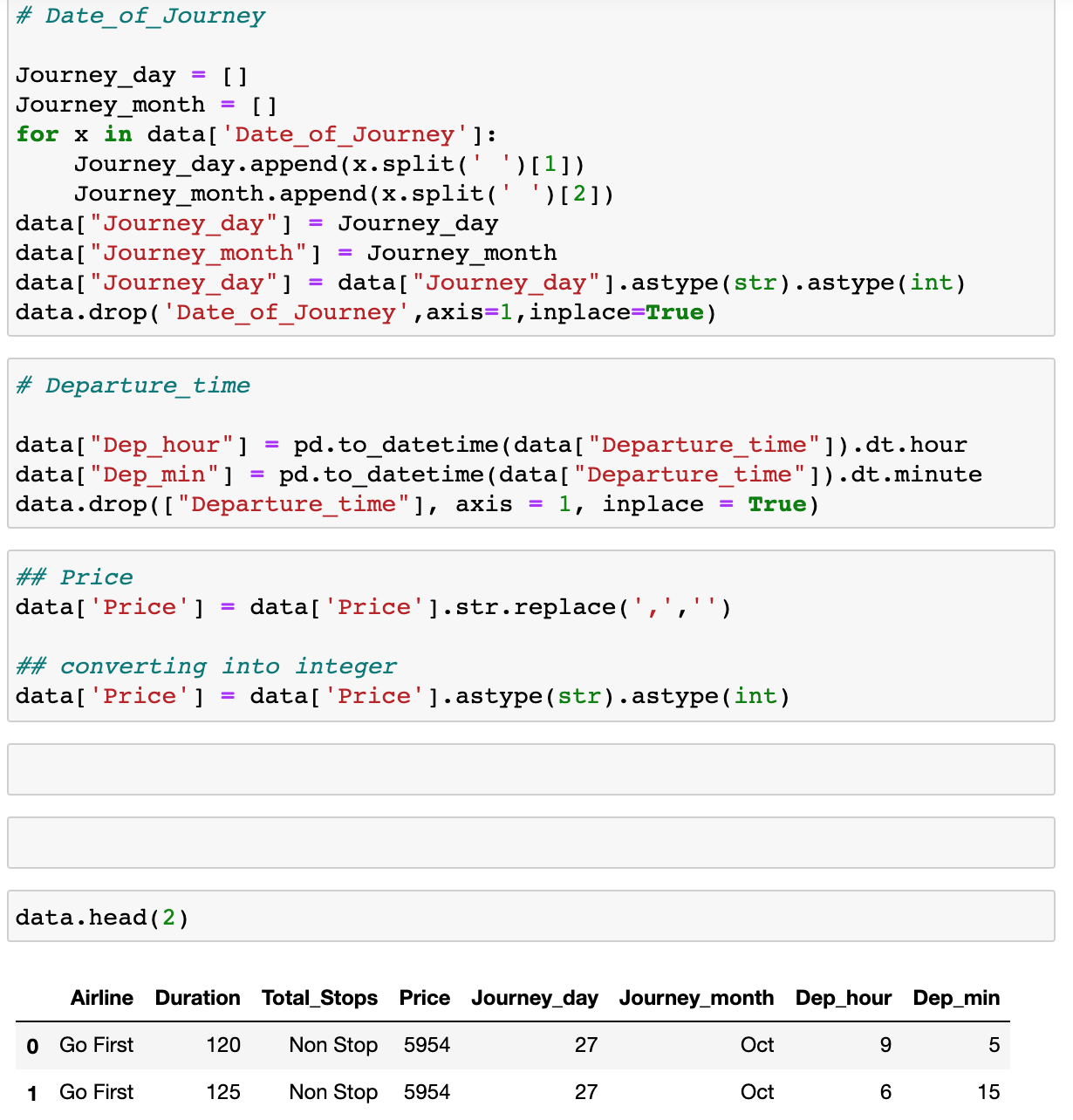
The data is collected from Indian Airline companies websites and it is converted in csv (comma separated values) format. Most of our attribute of our data have dtype object as we can see above using the info method.

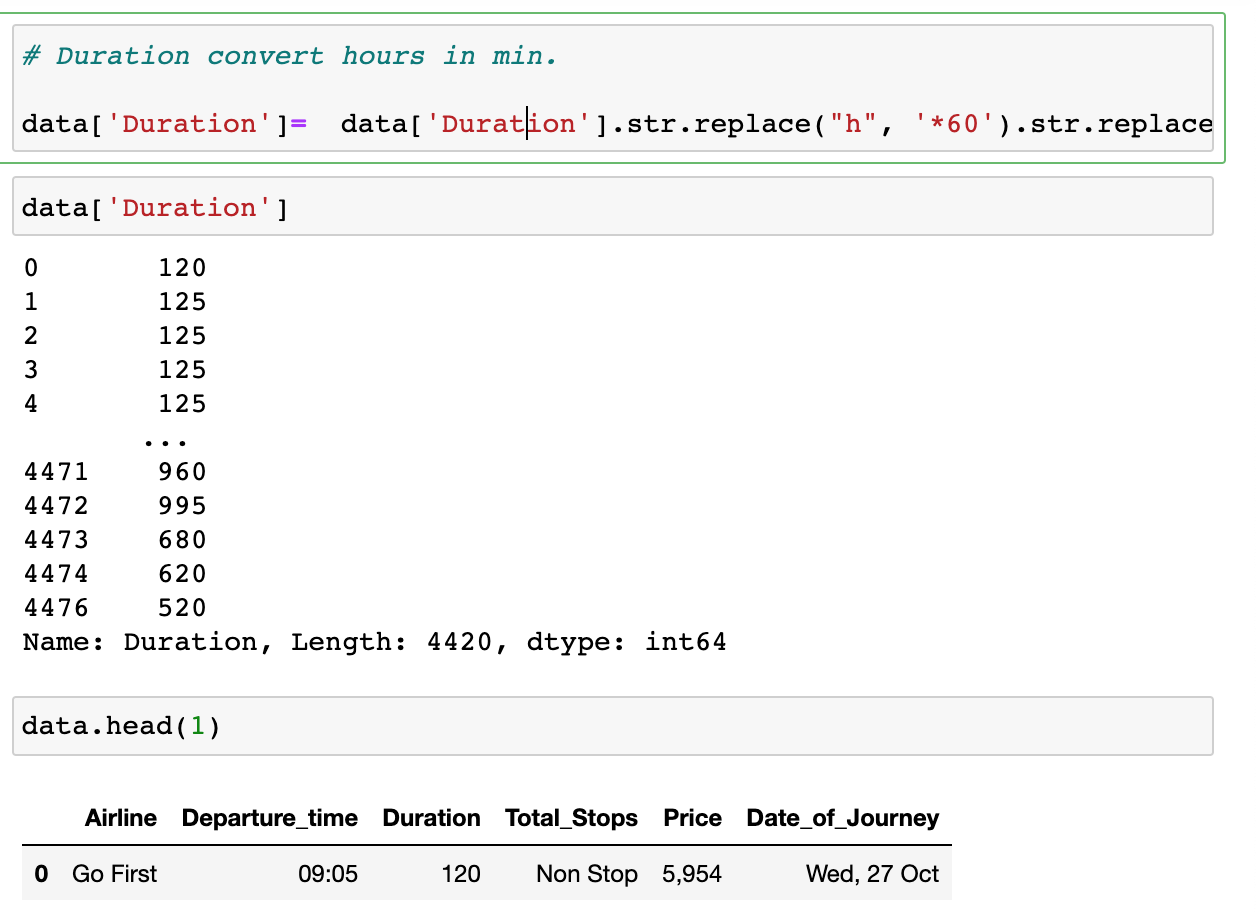
* Data Pre-processing Done

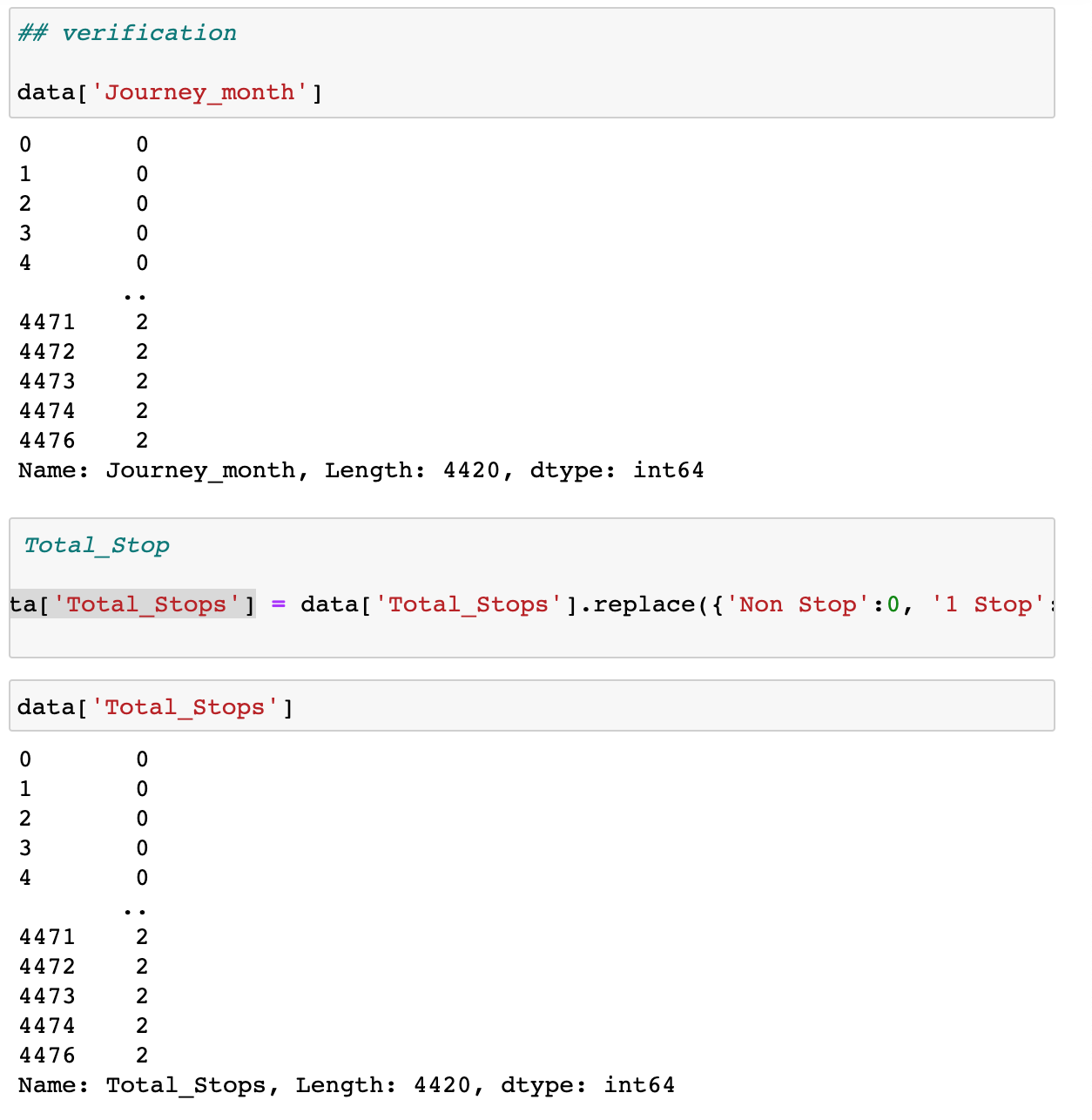
****

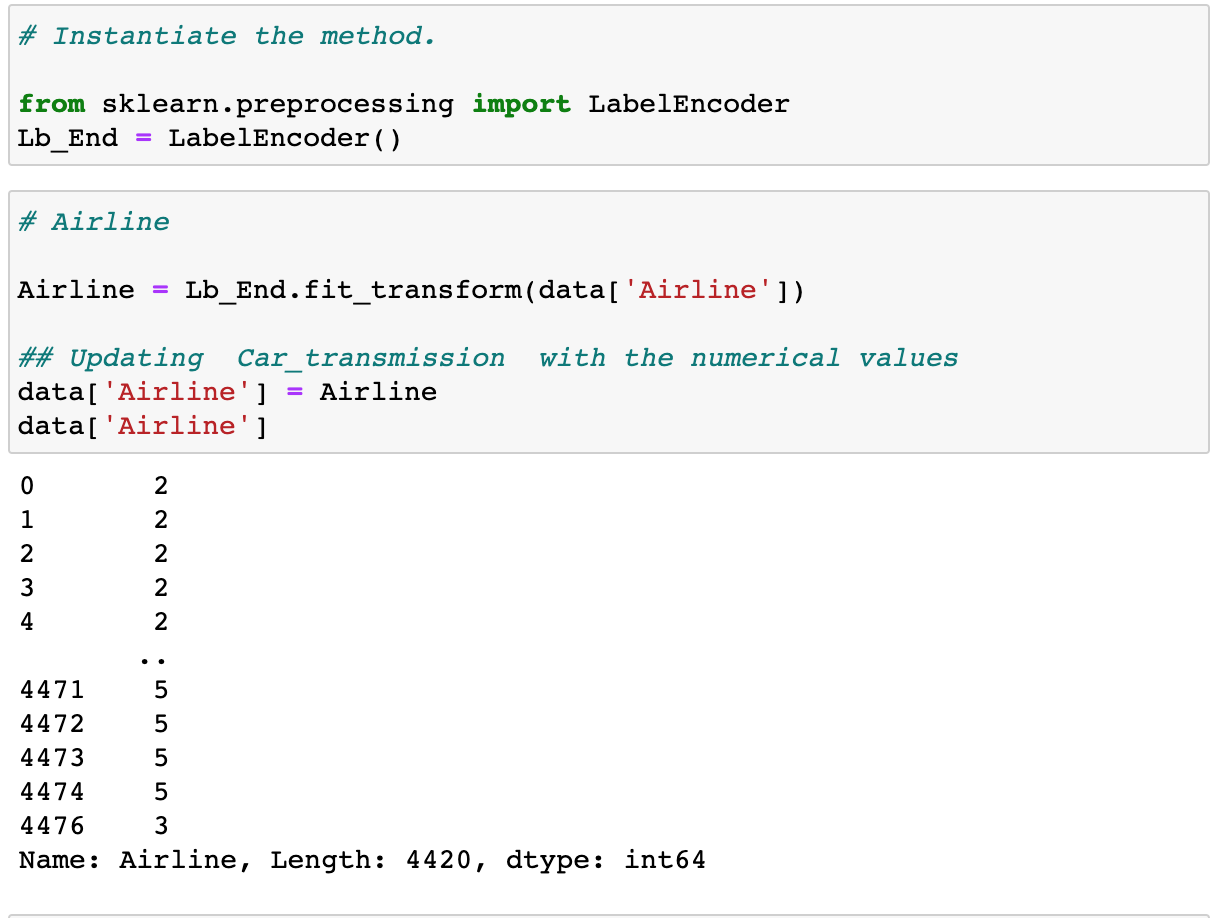
****

****

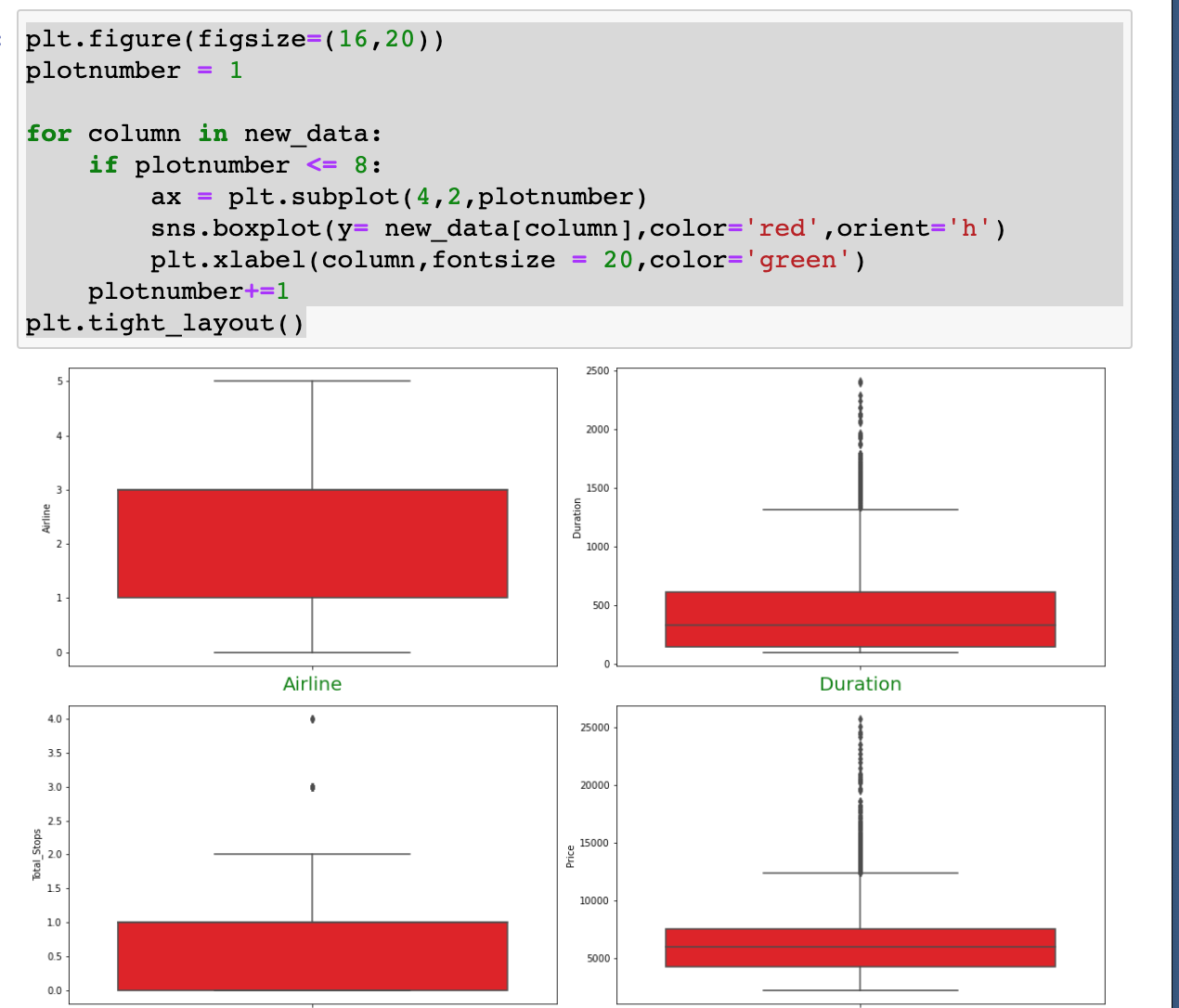
****

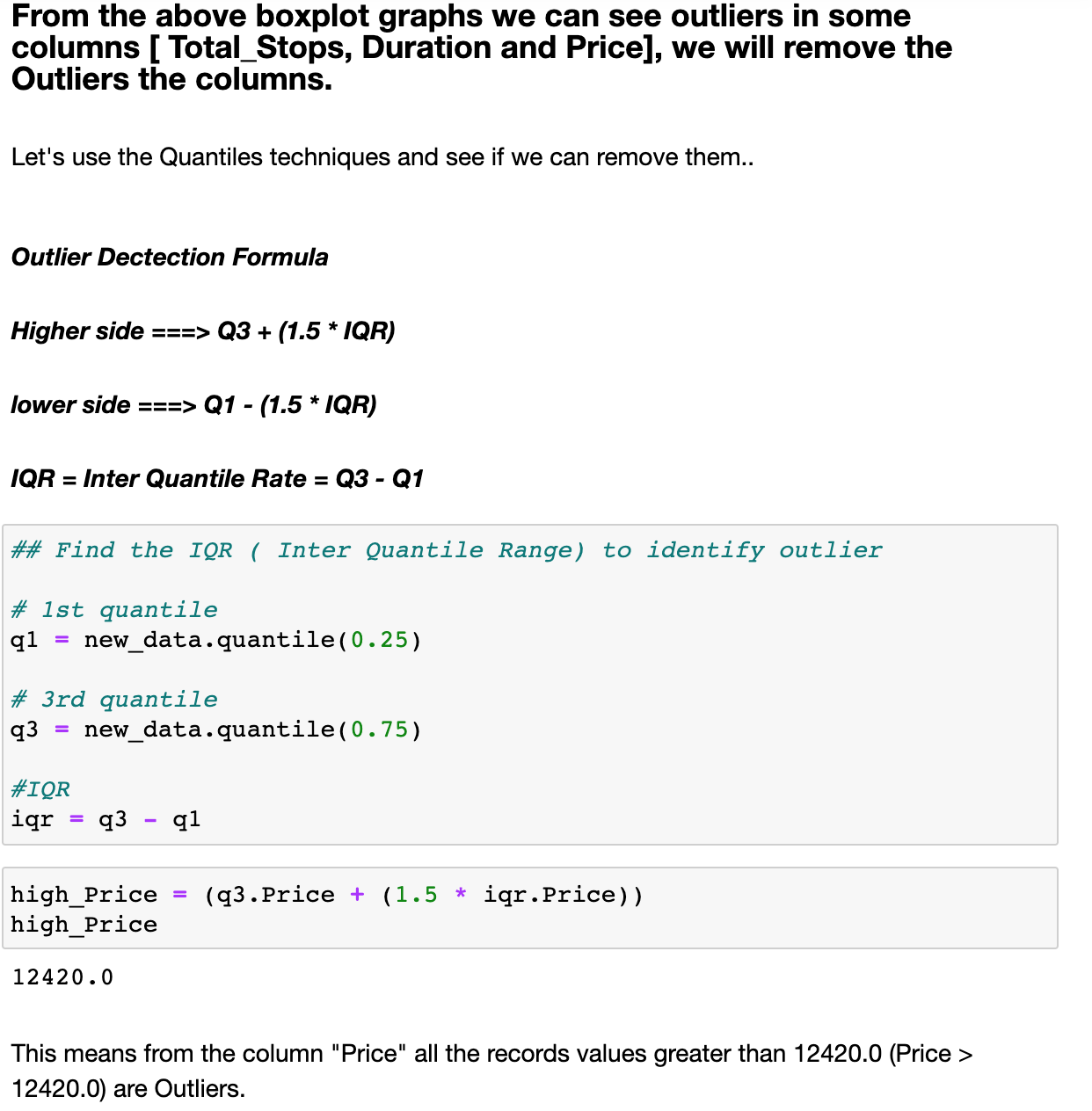
****

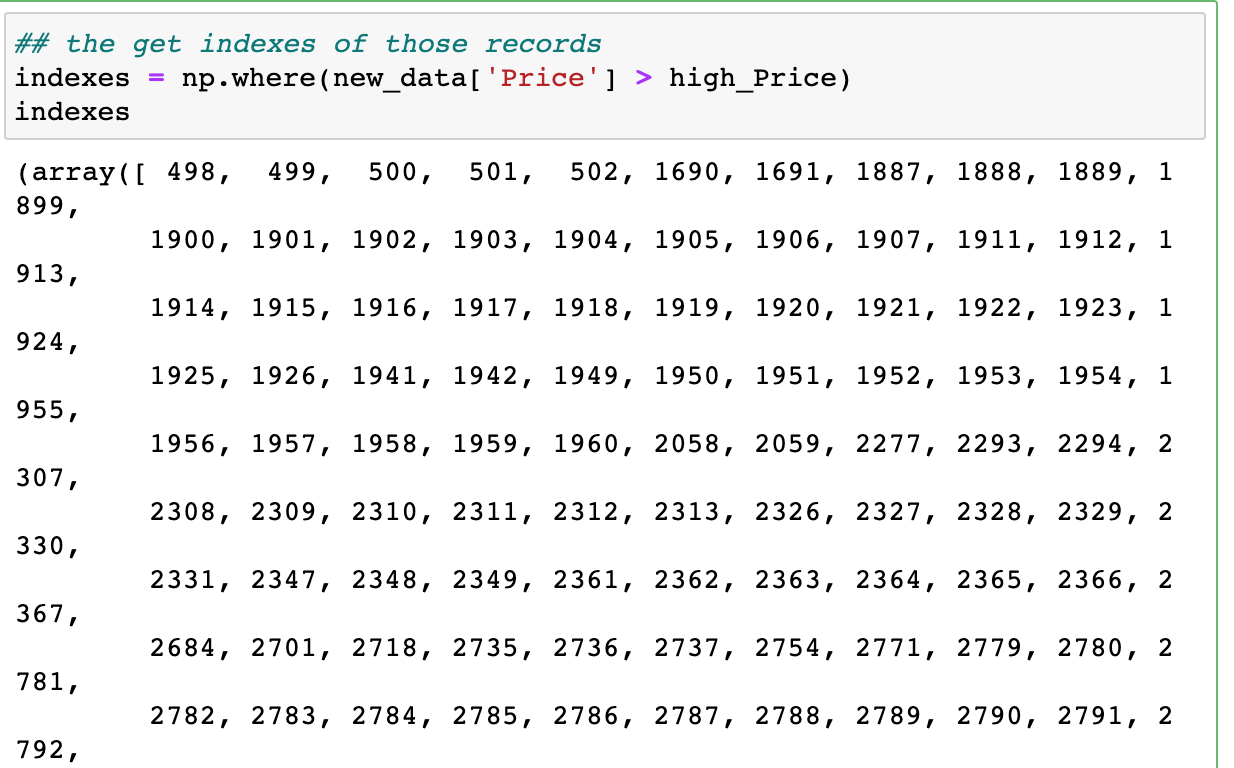
****

****

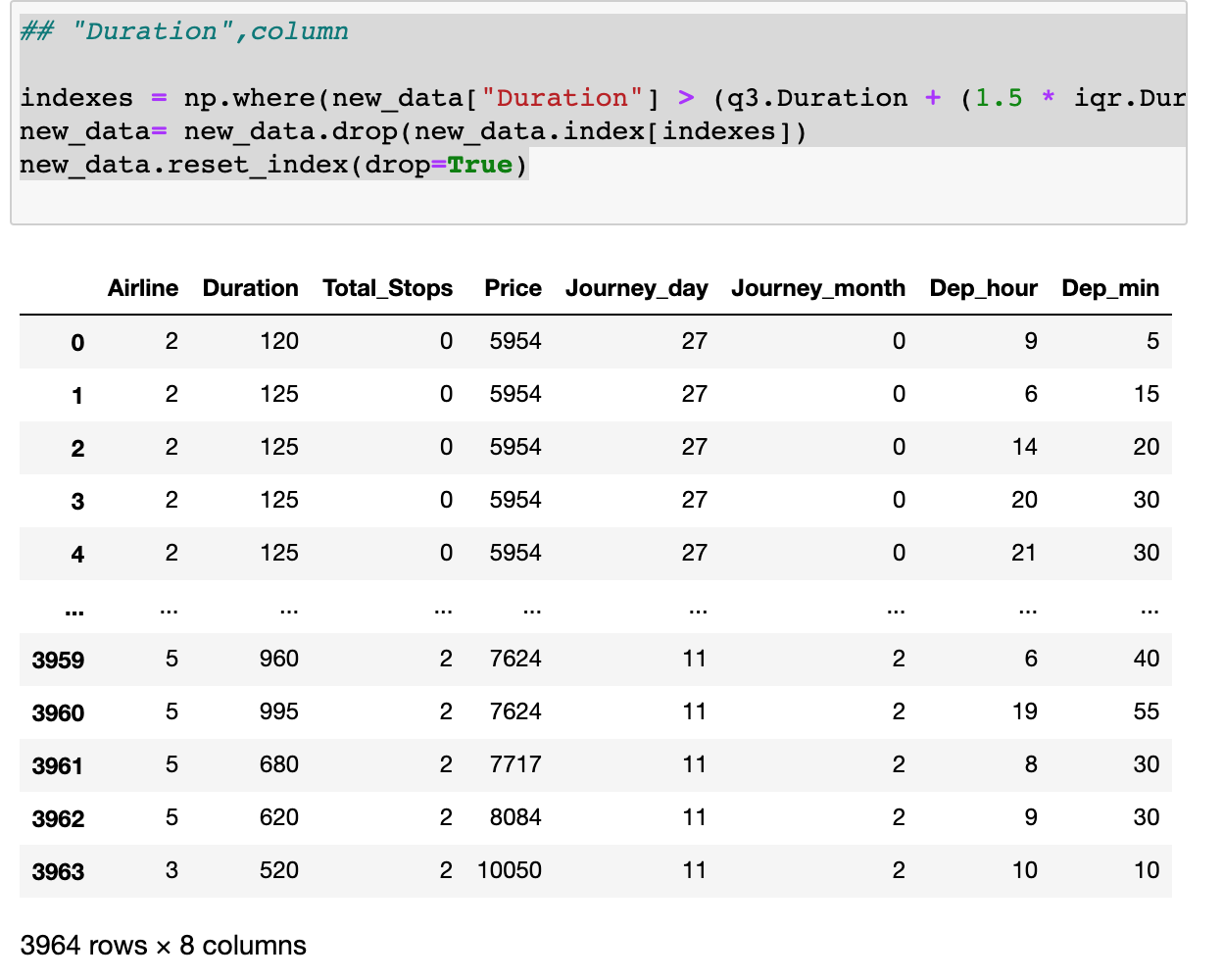
**Verification for Outliers and data skewness**

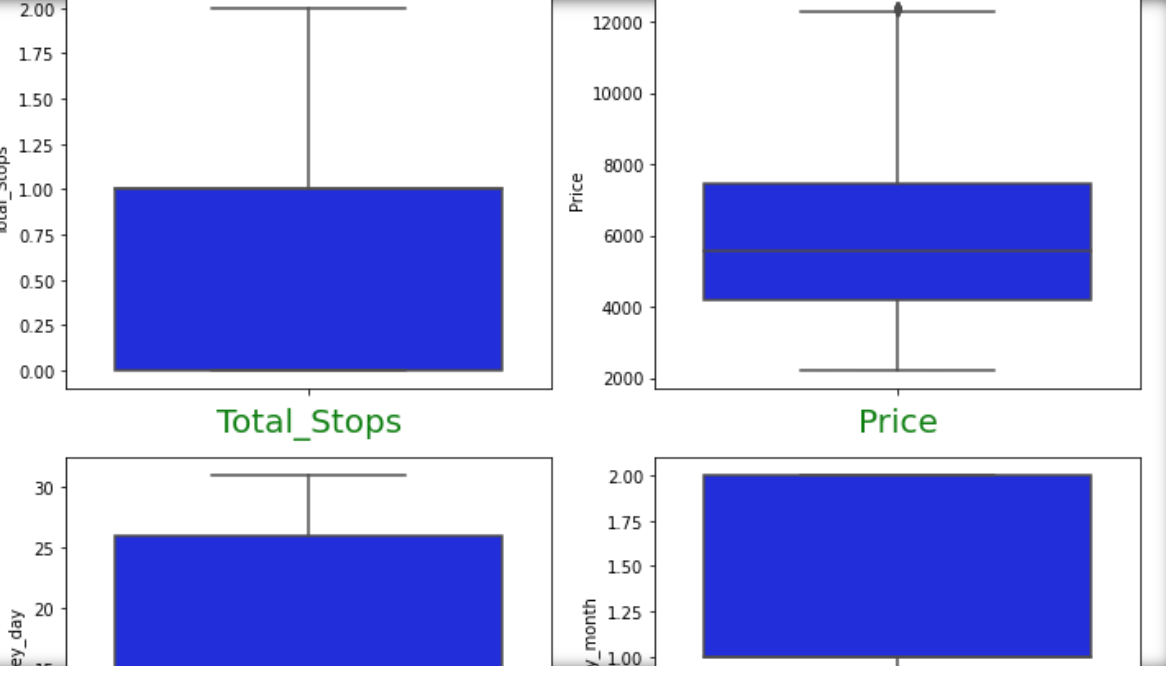
****

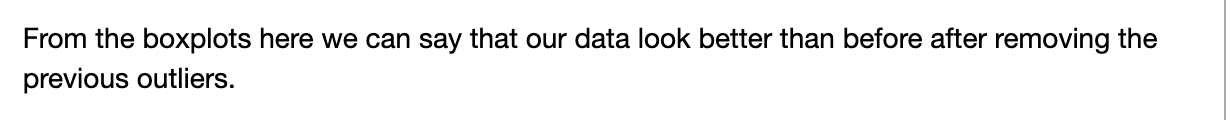
****

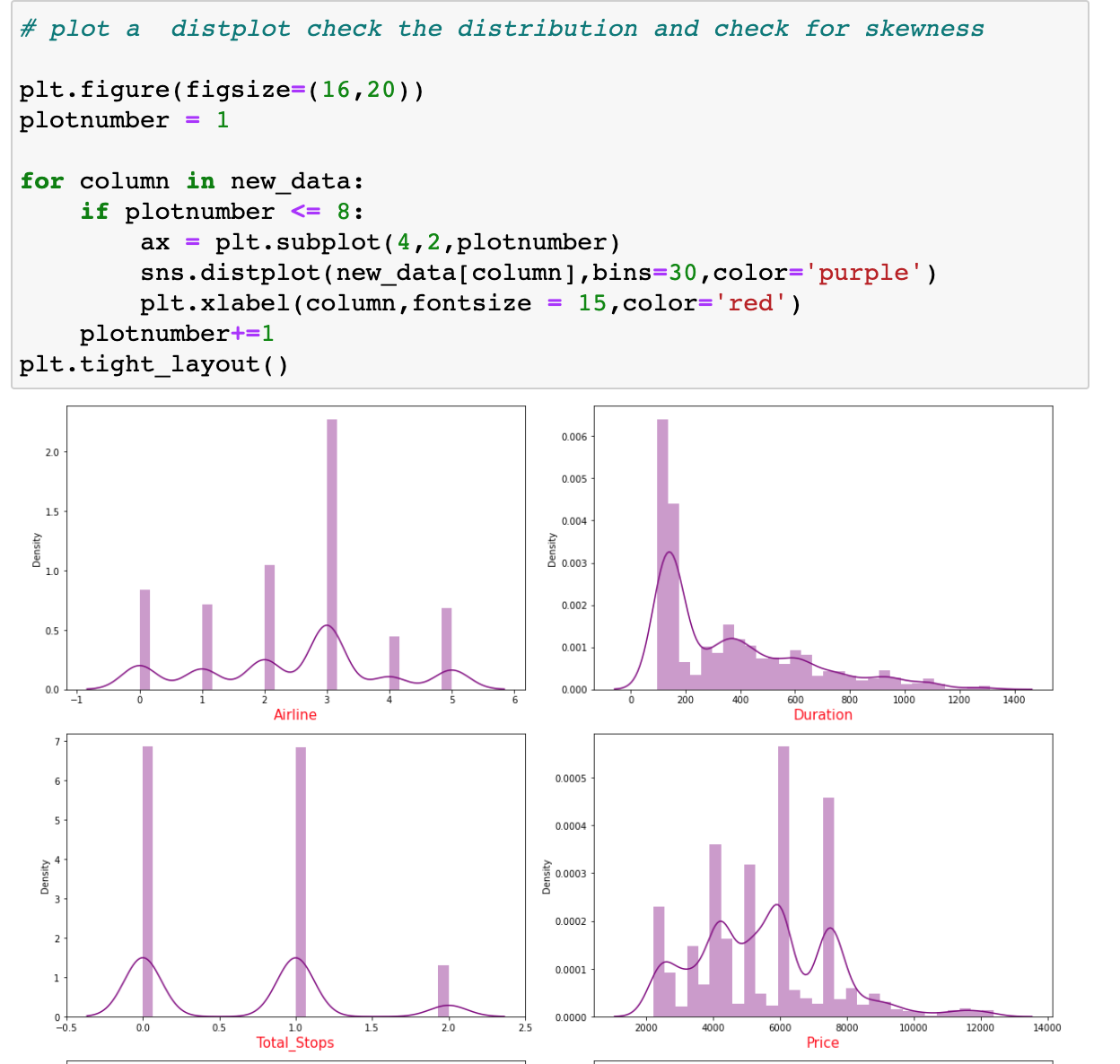
****

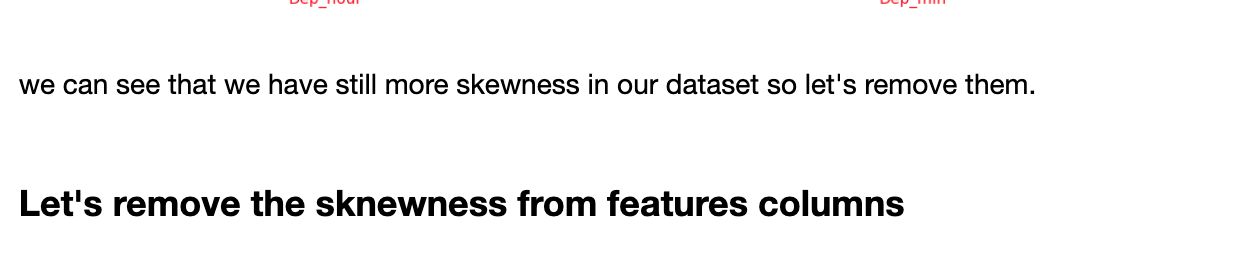
****

****

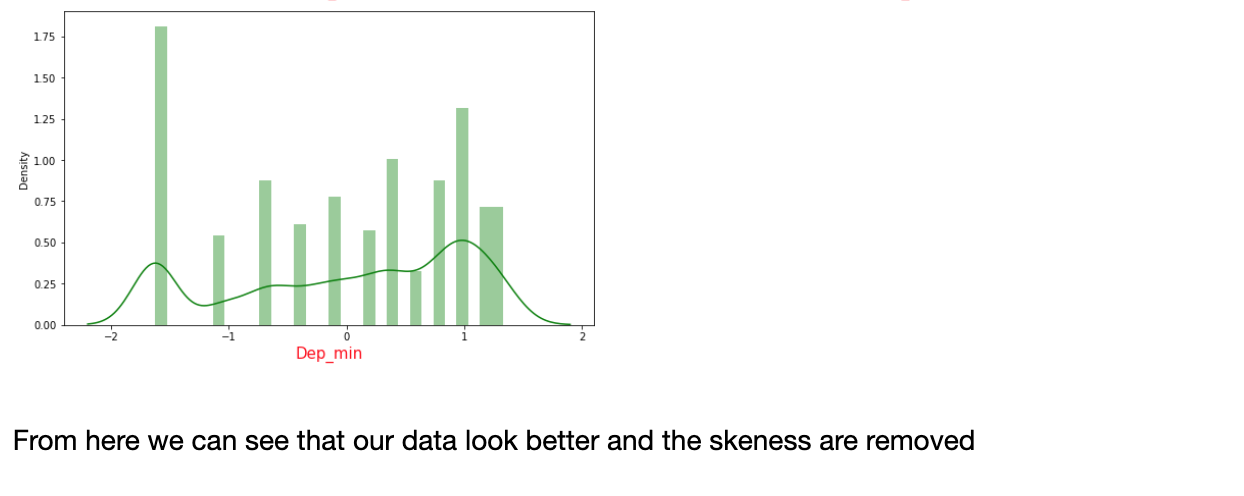
****

****

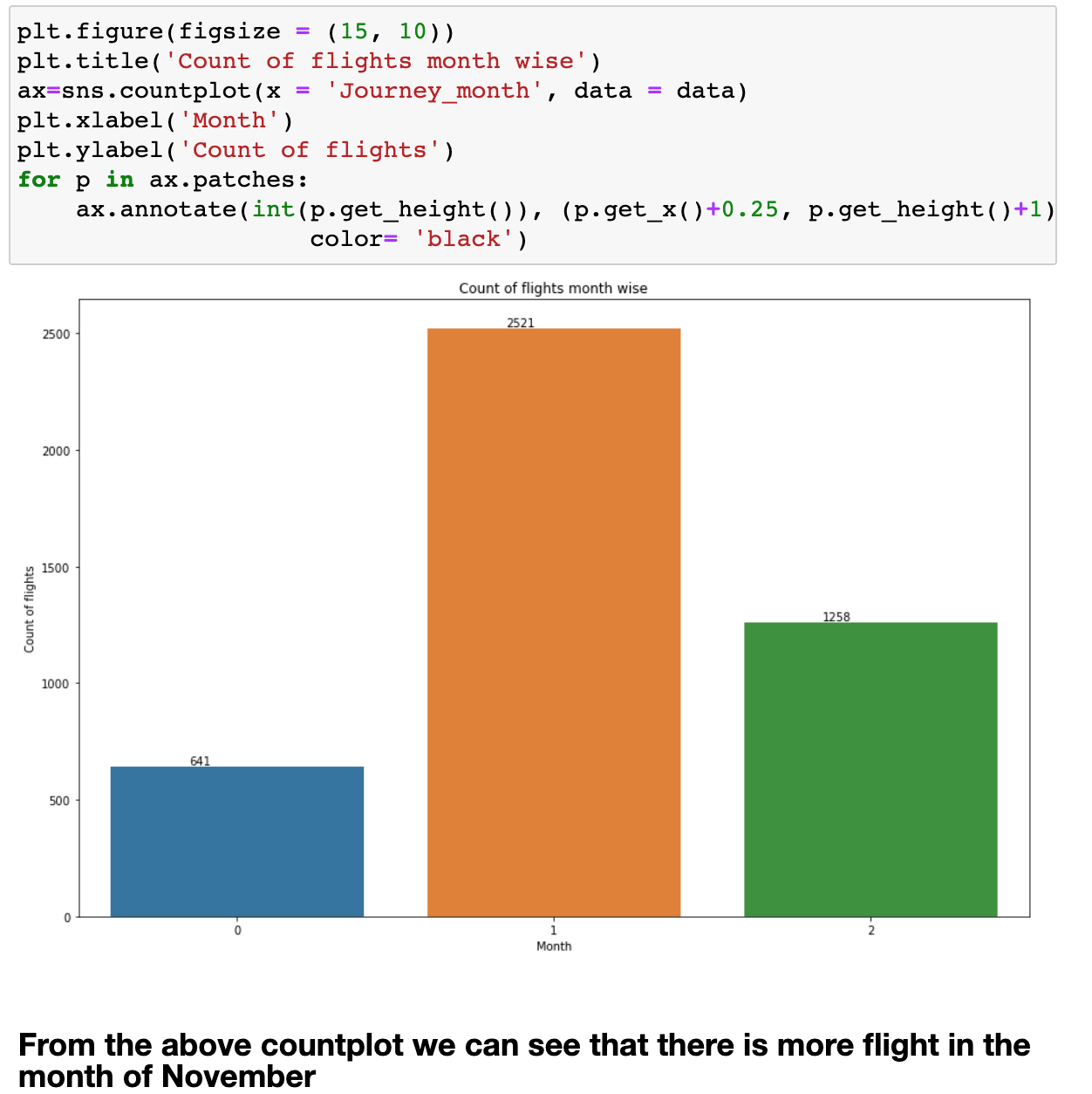
****

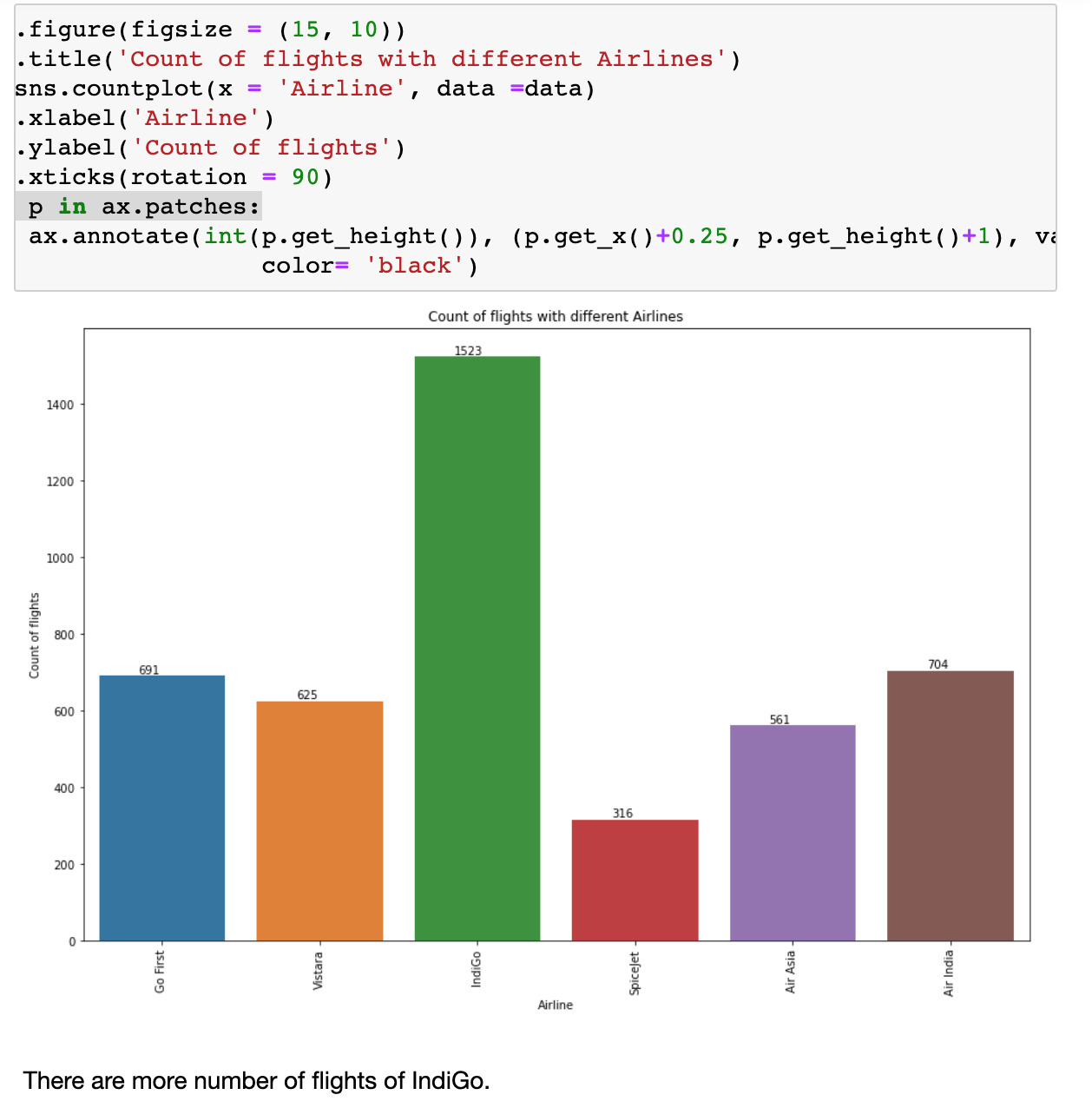
****

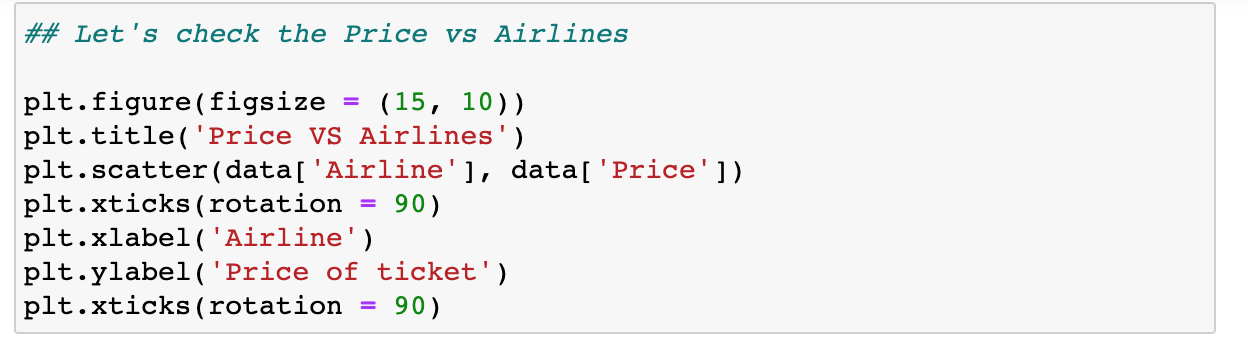
****

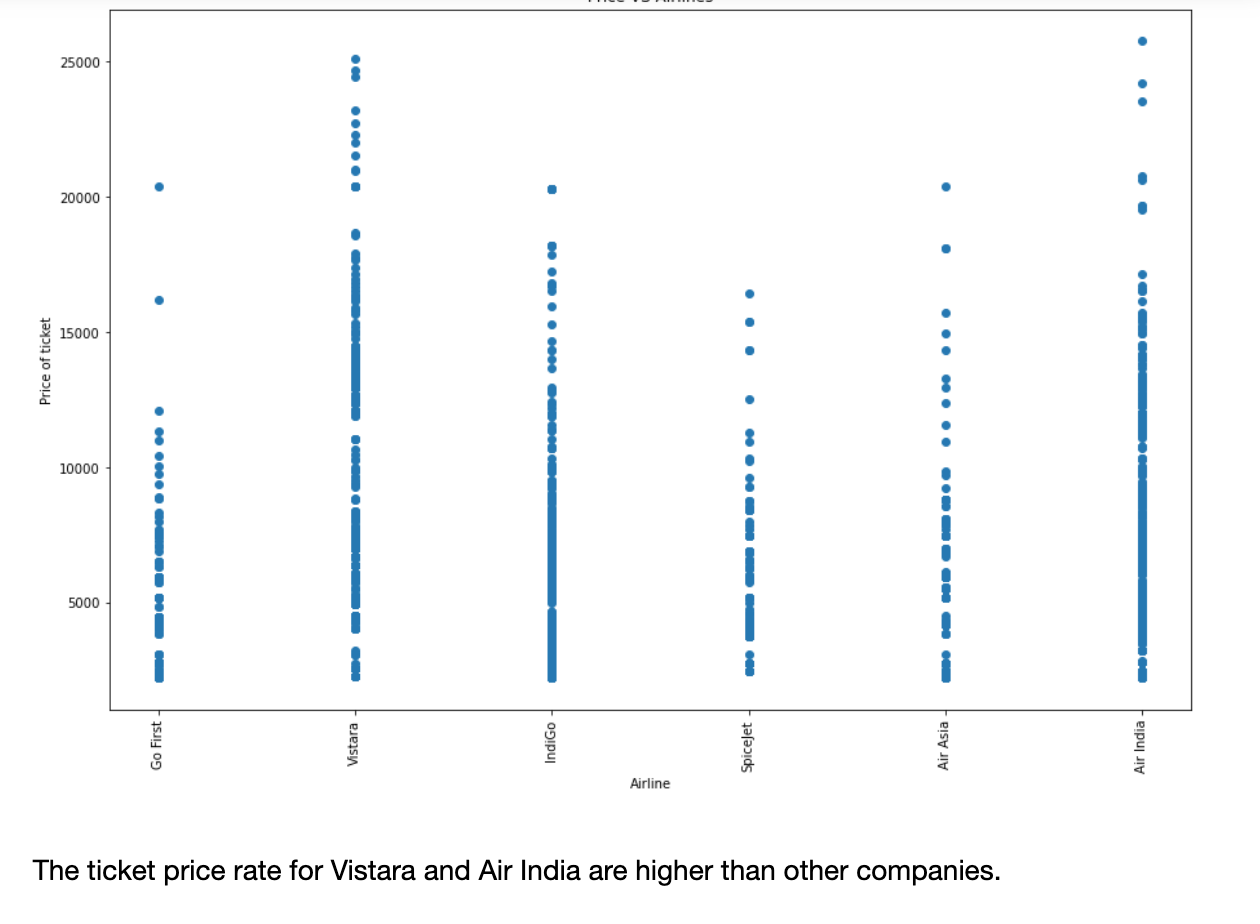
****

* Data Inputs- Logic- Output Relationships

****

****

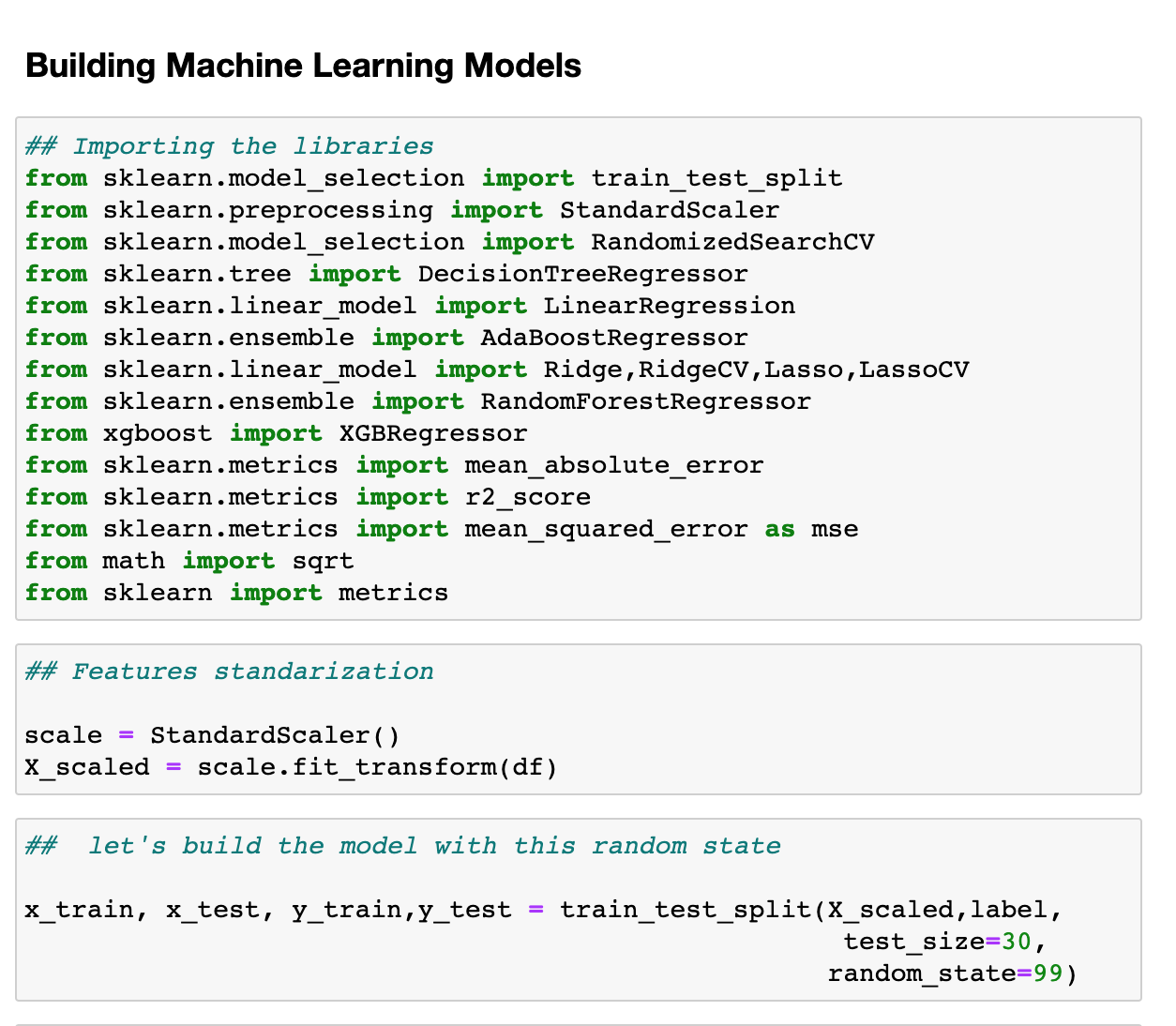
****

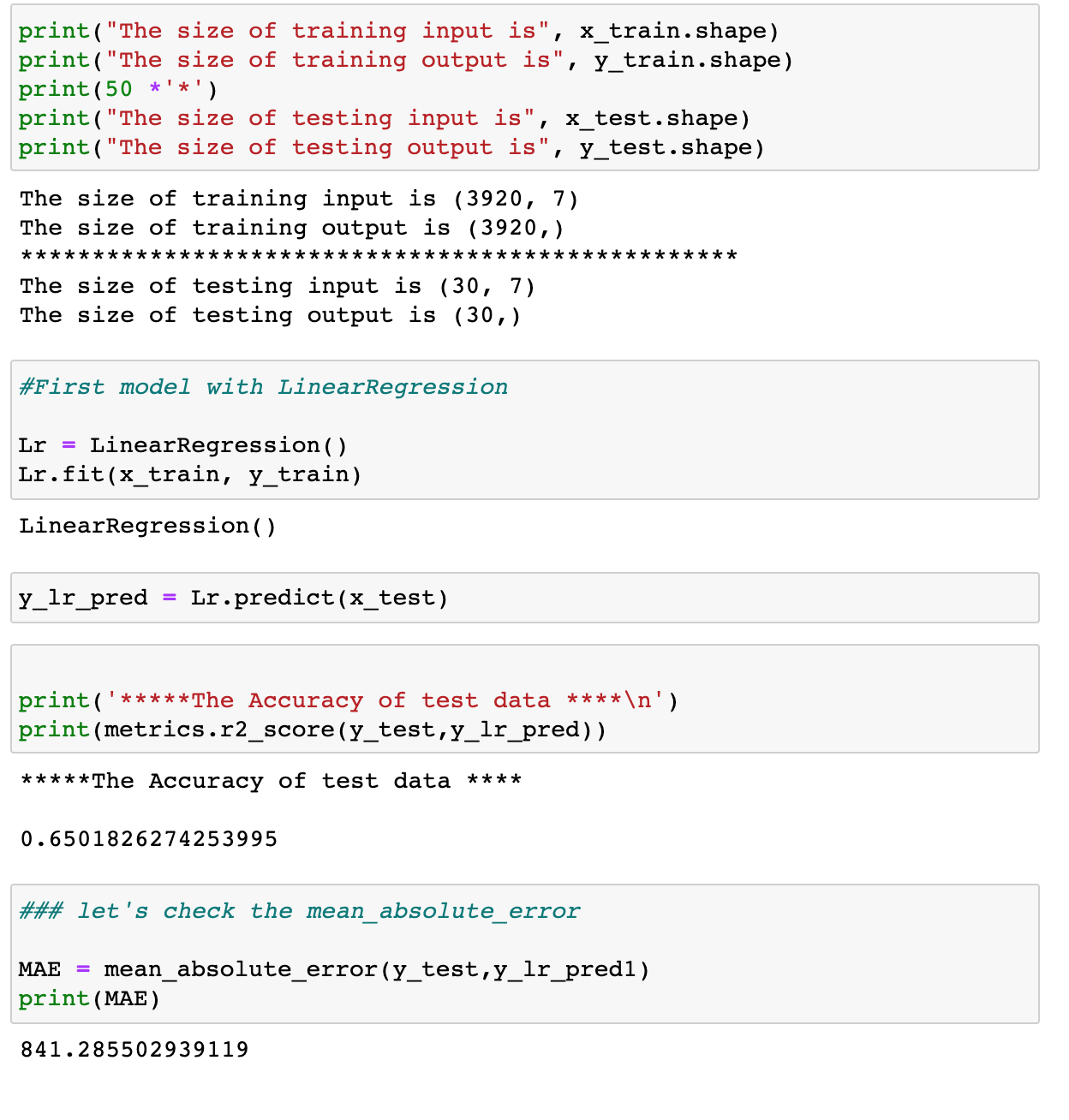
****

* Hardware and Software Requirements and Tools Used
* **1- Pandas** and **NumPy**
* **2 -Seaborn** and **Matplotlib for plotting graphs**

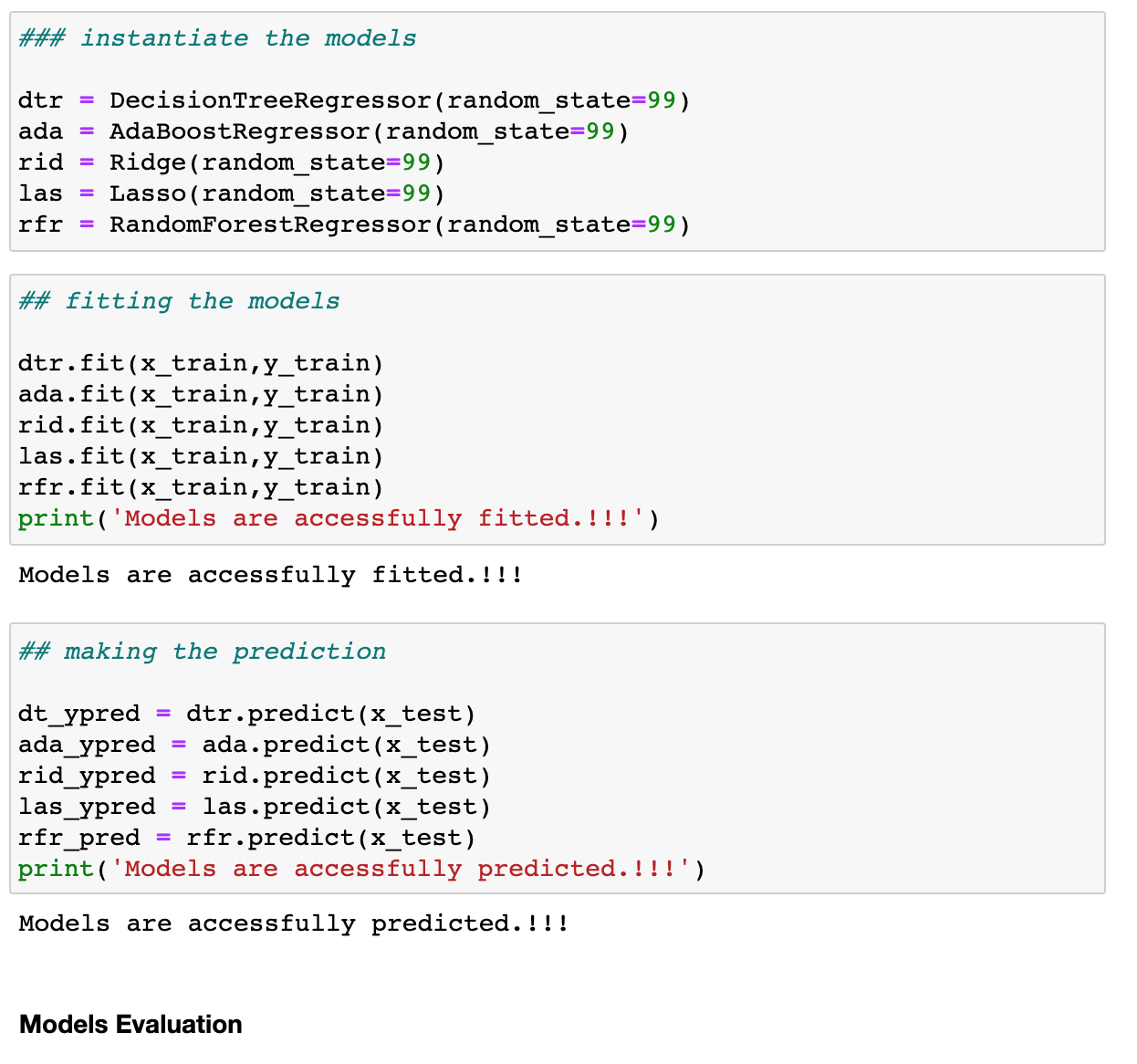
**Model/s Development and Evaluation**

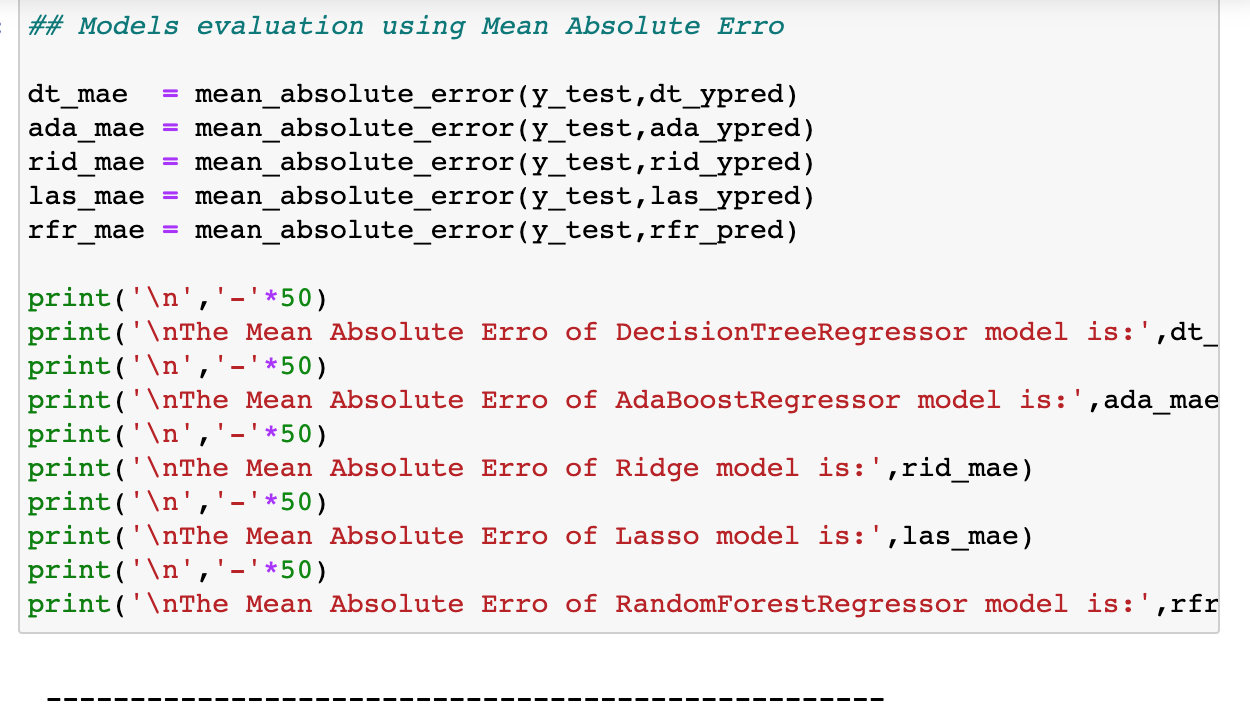
***Importing the necessary libraries***

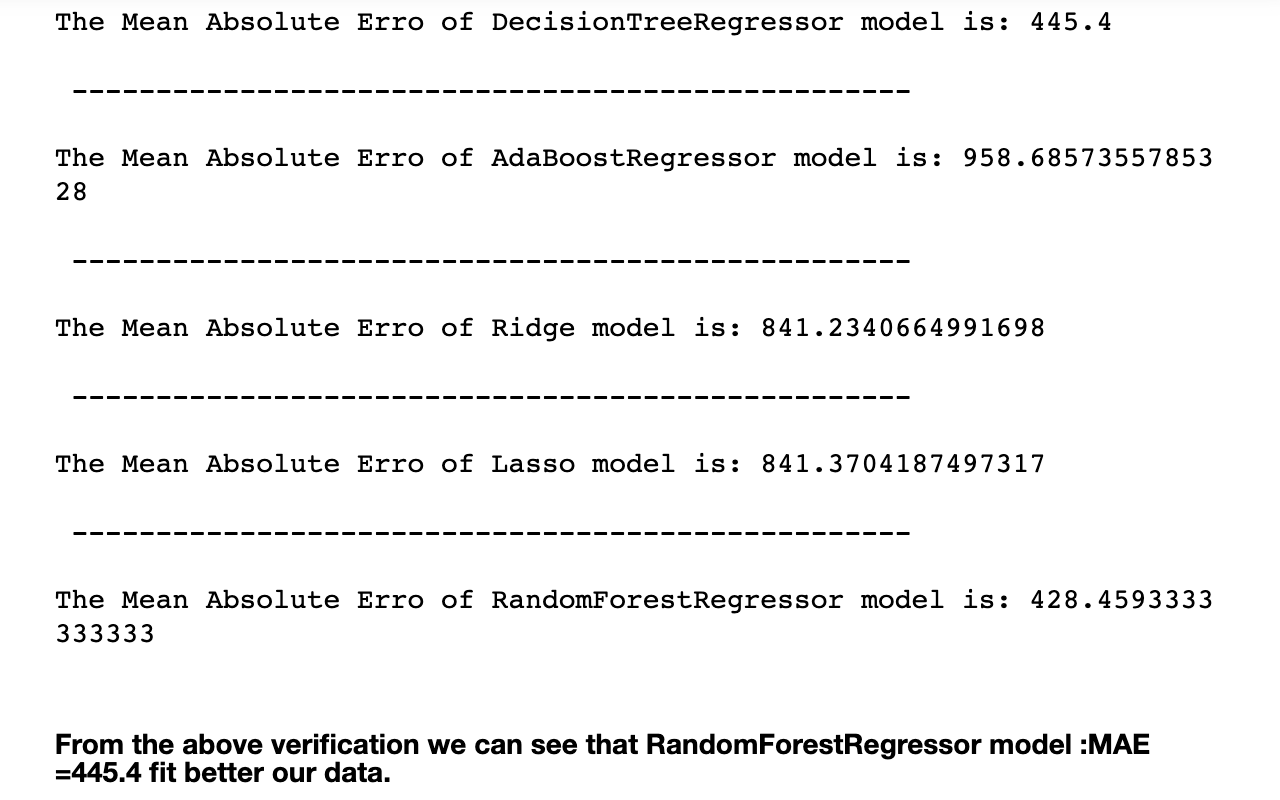
****

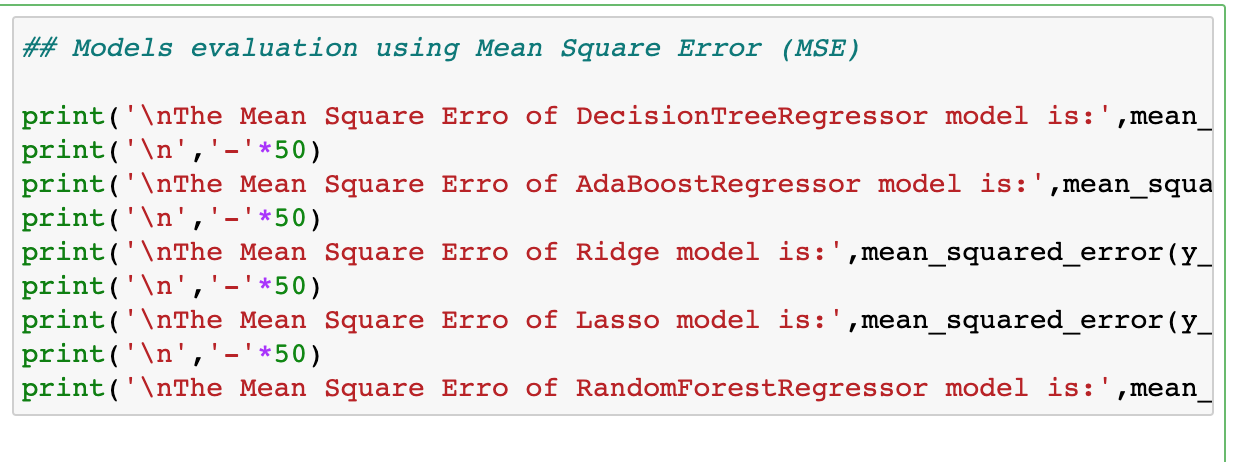
****



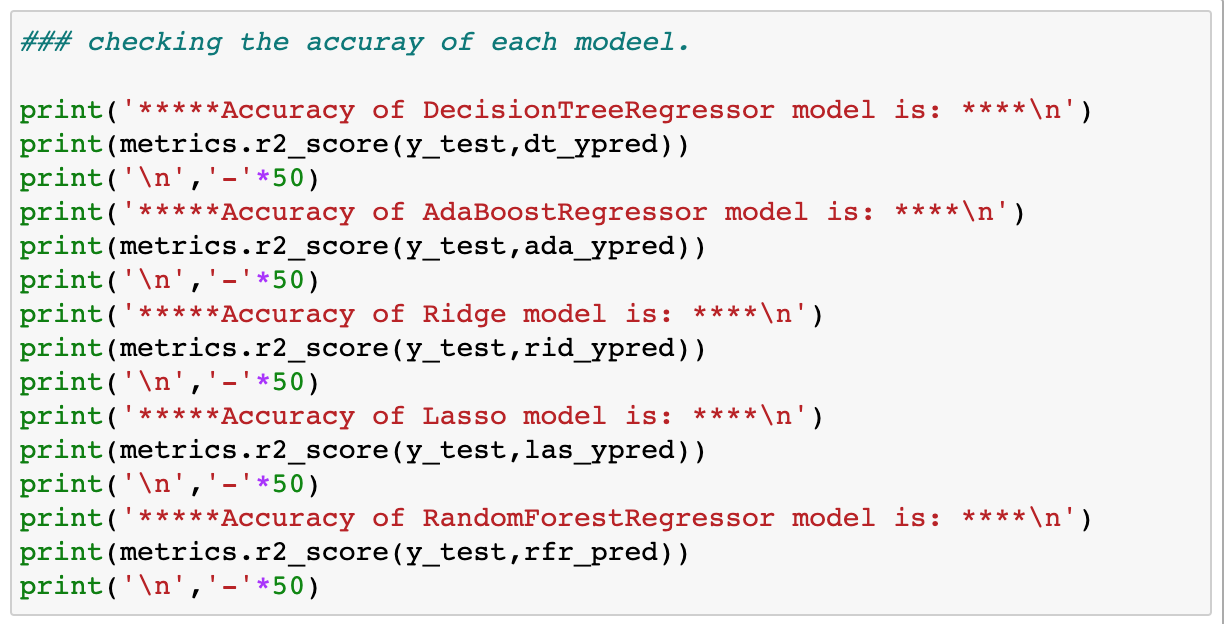


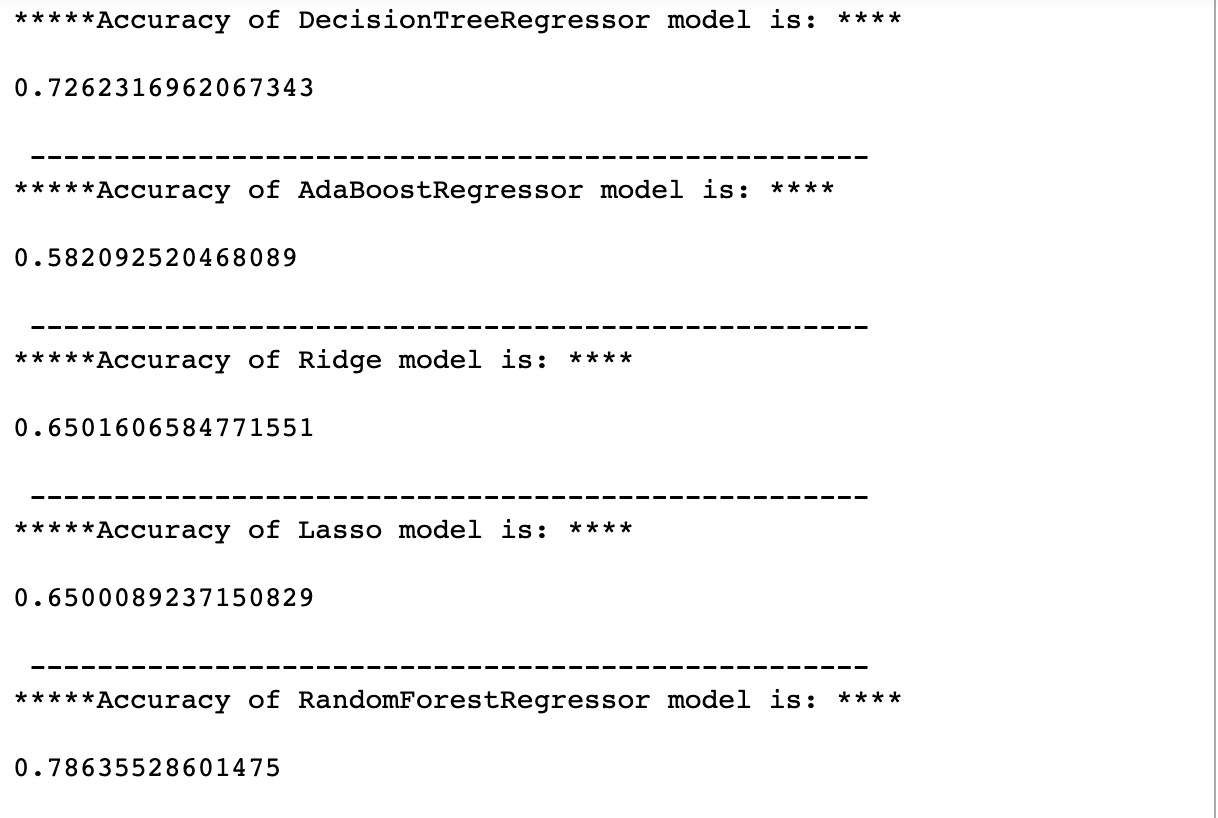


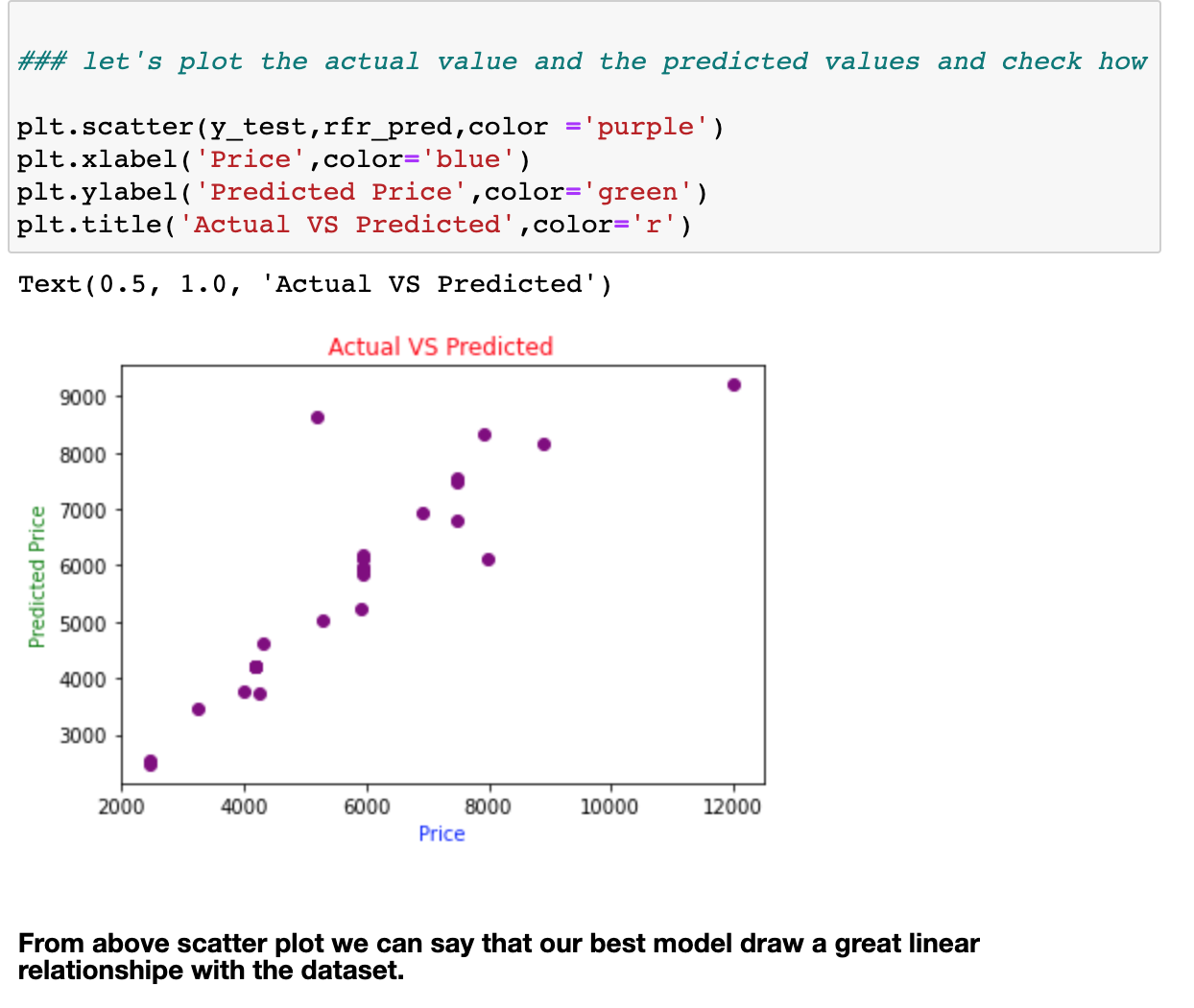


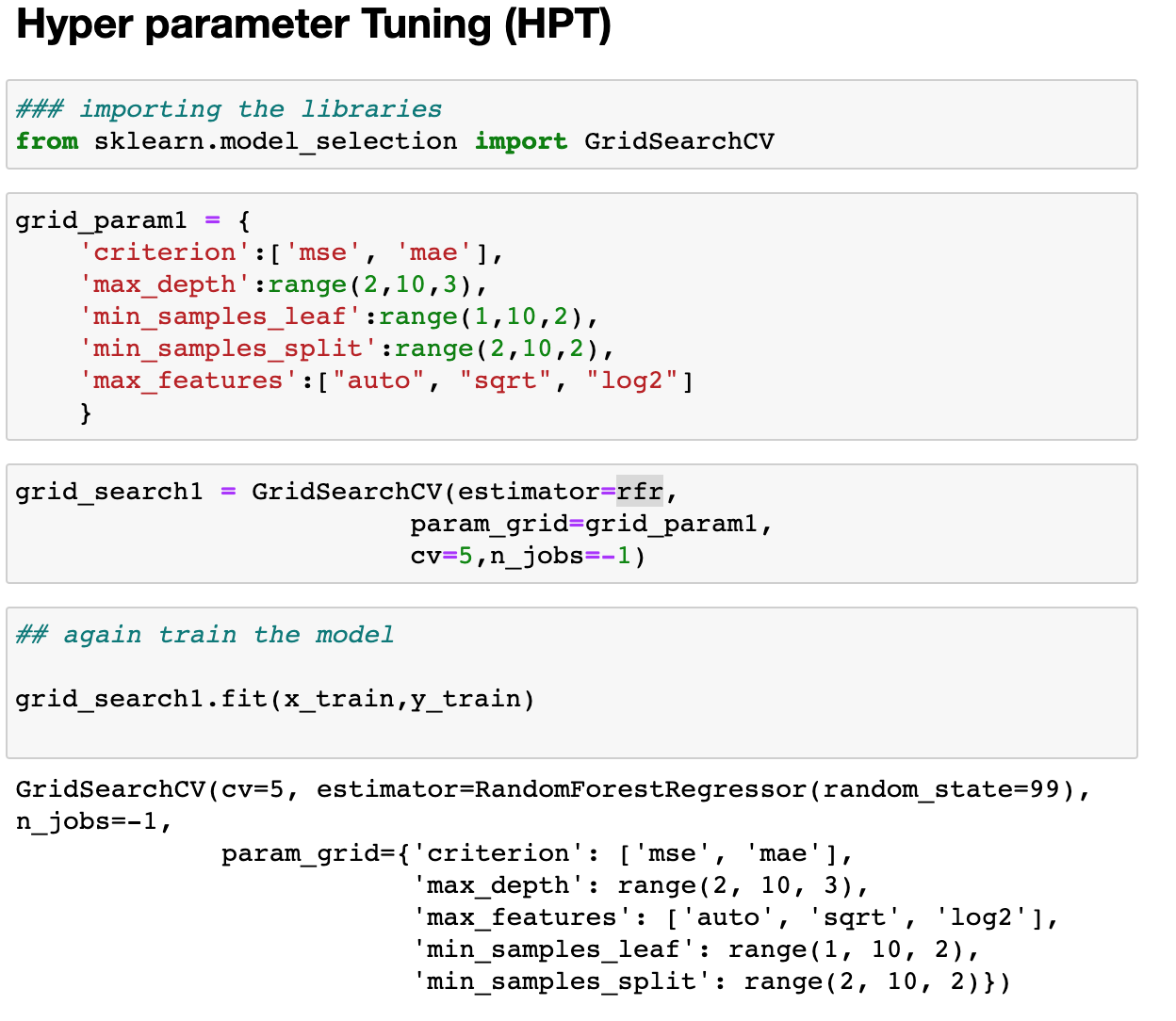


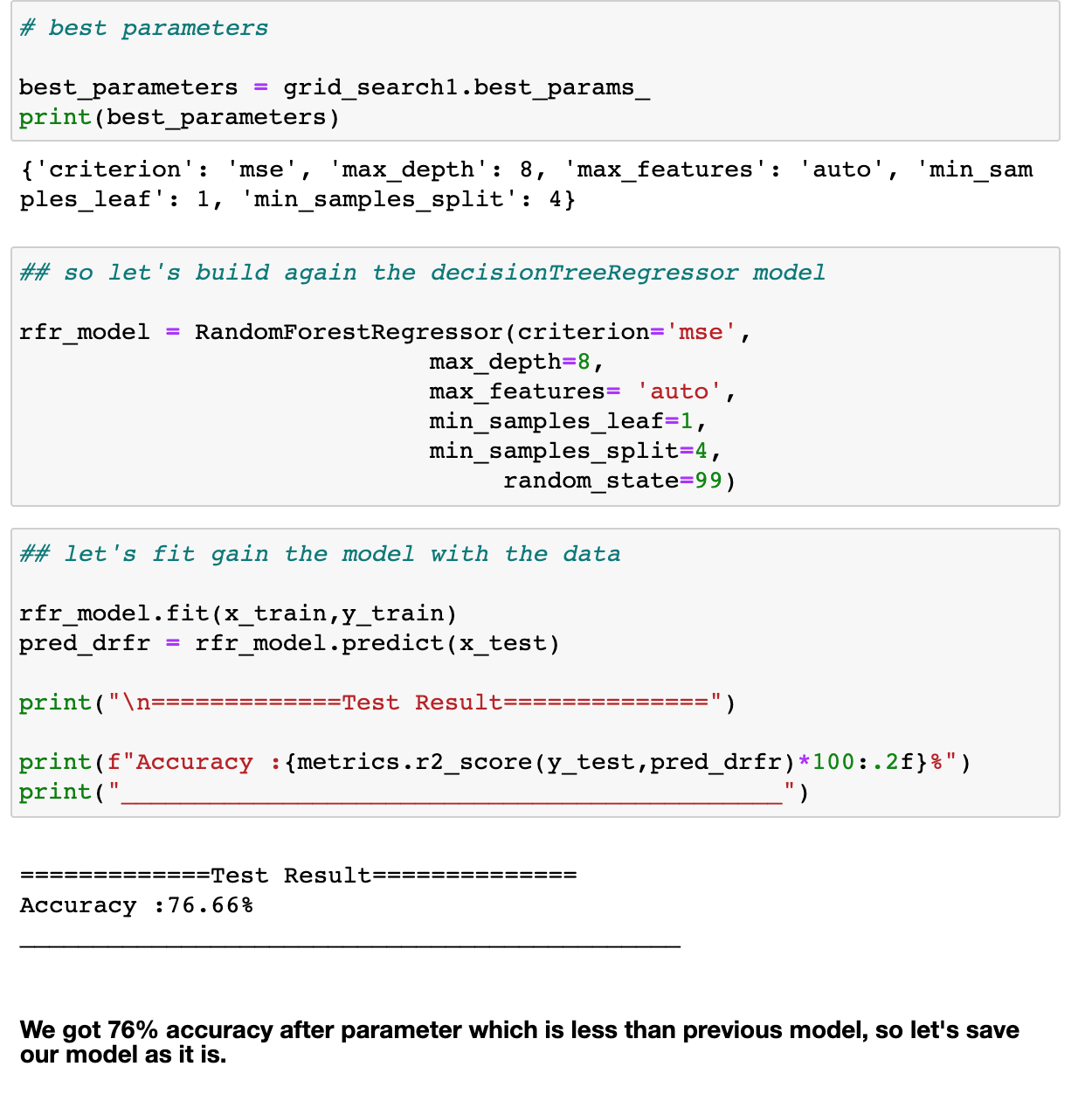


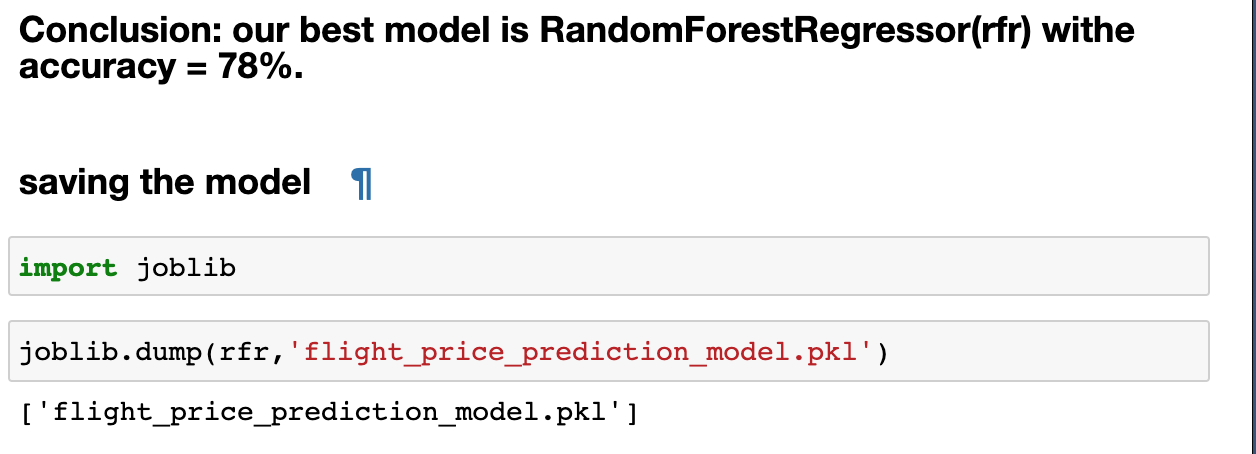












**CONCLUSION**

The key Finding and Conclusion of the Study

The principal key finding was that the model is built and it is capable to predict the target variable Price

And this can help Airline companies to well manage the ticket price and help traveller to get the flights for cheapest prices.