**Building Regression Model on Flight Ticket Prediction Model**

Hello all! On this Article, I will cover the basic of creating your own regression model using Python. I will try to explain and demonstrate to you step-by-step from preprocessing your data, training your model, optimising the model, and how to solve it for better use.

**Introduction**

In Machine learning, Linear regression is a supervised learning model. In this, Models generated are represented in form of equation, where equation can be:

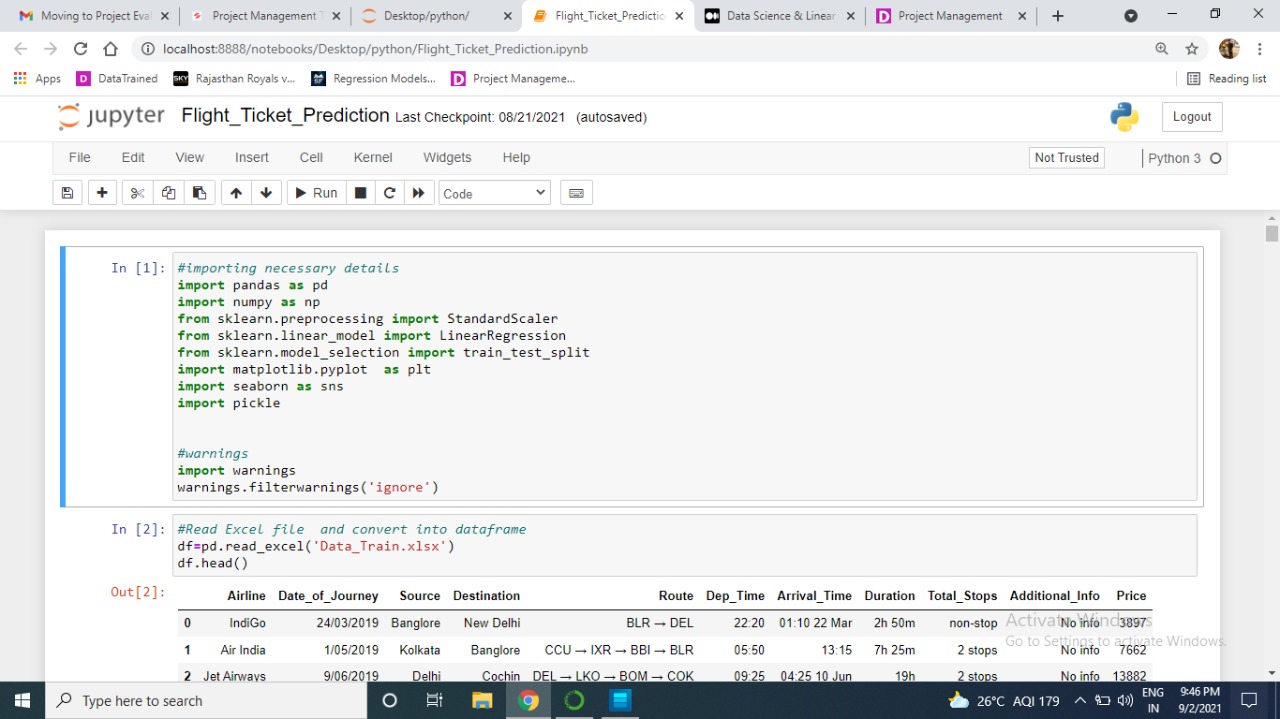
* 2-Dimension->Equation of line
* 3-Dimension->Equation of Plane
* n-Dimension->Equation of Hyper Plane

Using this model, Regression tells that we will predict some real number values. The term ‘Linear’ refers to method used to generate model which uses linear combination of Independent Variable i.e. y is combination of xi independent variables having different weights of same degree.

**Equation of Y:**

**Y=Intercept + Coefficient \* Independent variable**

In this article, we will use the Customer\_Retention\_Project to build a model to predict the price of flight ticket based on some attributes. We will try to build a model with good accuracy. We will also use some technique and try to enhance the accuracy of the model.



**Data Pre-processing:**

Before we can begin to create our first model we first need to load and preprocess. This will ensure that our model will receive a good data to learn from, as they said “a model is only as good as its data”.

The data preprocessing will be divided into few steps as explained below:

Loading Data:

In this first step we will load our dataset.

**Input variables:**

FEATURES:

Airline: The name of the airline.

Date\_of\_Journey: The date of the journey

Source: The source from which the service begins.

Destination: The destination where the service ends.

Route: The route taken by the flight to reach the destination.

Dep\_Time: The time when the journey starts from the source.

Arrival\_Time: Time of arrival at the destination.

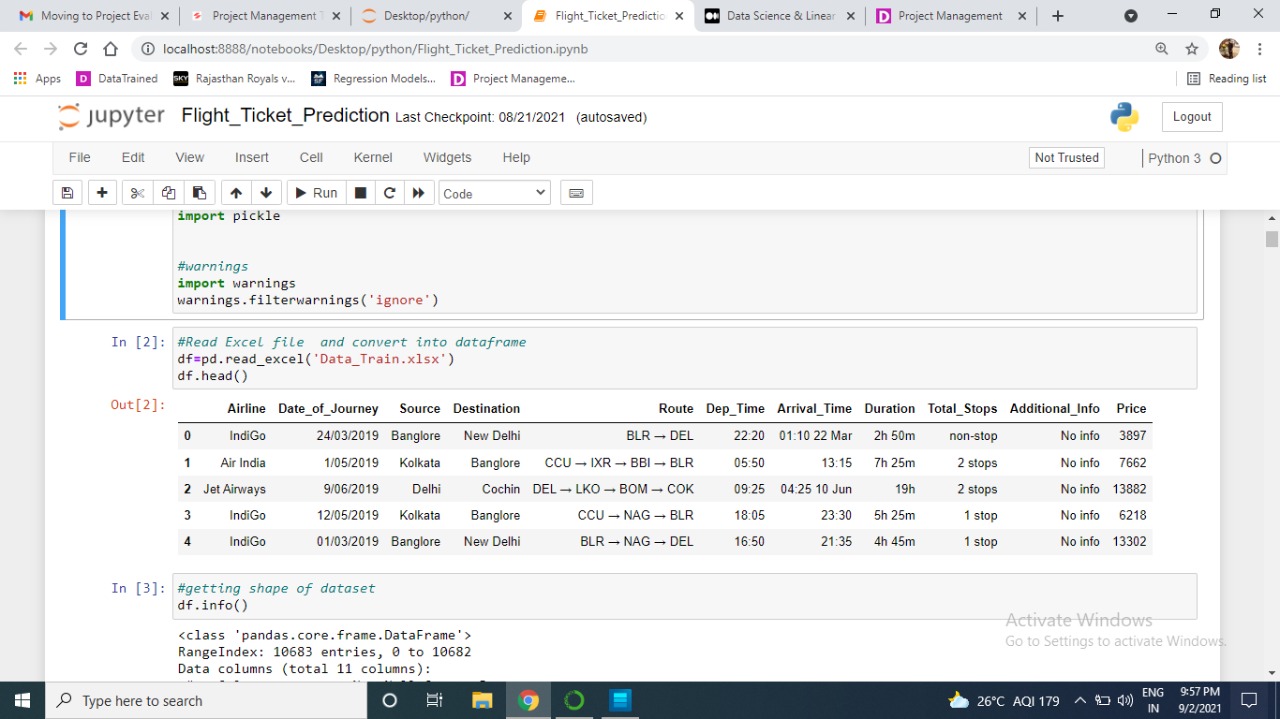
Duration: Total duration of the flight.

Total\_Stops: Total stops between the source and destination.

Additional\_Info: Additional information about the flight

**Output variables:**

Price: The price of the ticket.



After loading dataset we have to get info of whole dataset.

