Applied Data Science Internship

Project Report

Project Title:

**DYNAMIC PRICING PREDICTION FOR TRAVEL USING IBM WATSON**

****

Team Members:

|  |  |  |
| --- | --- | --- |
| **S.No** | **Name** | **Register Number** |
| 1 | Siddharth Garg | 18BCE0525 |
| 2 | K Satya Deva Hanumanth Charan | 18BEE0140 |
| 3 | S Sudhir Sai | 18BEE0184 |
| 4 | Ch Sai Vineeth | 18BEE0185 |

Table of Contents

[Introduction: 3](#_Toc4662)

[a. Overview 3](#_Toc25016)

[b. Purpose 3](#_Toc15644)

[Literature Survey : 3](#_Toc17957)

[a. Existing problem 3](#_Toc11933)

[b. Proposed solution 3](#_Toc10981)

[Theoretical Analysis : 4](#_Toc5038)

[a. Block Diagram : 4](#_Toc1125)

[b. Software Designing 4](#_Toc9537)

[Experimental Investigation : 5](#_Toc19734)

[Flowchart: 5](#_Toc19467)

[Result: 5](#_Toc27331)

[Advantages & Disadvantages: 5](#_Toc32518)

[Applications: 5](#_Toc26477)

[Conclusion: 6](#_Toc7001)

[Future Scope: 6](#_Toc25118)

[Bibliography : 6](#_Toc10335)

[Appendix : 6](#_Toc32605)

[a. Source Code: 7](#_Toc20689)

[b. UI/Output Screenshot: 12](#_Toc7710)

## Introduction:

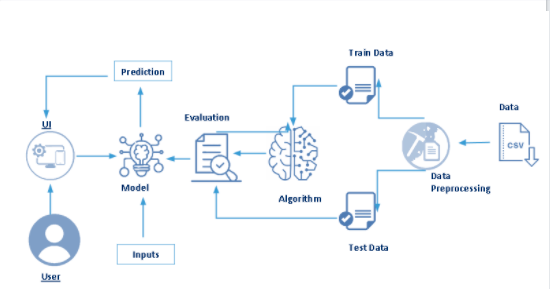
1. Overview: Dynamic pricing is one of the most commonly used methods for revenue management. It allows companies to increase their revenues by combining its supplies and demands, responding to changing patterns of demand. In this project we predict the prices of cab rides whose prices are dynamic. We have two datasets named: ‘weather.csv’ and ‘cab\_rides.csv’ using these datasets we build our prediction model. To do so we perform data wrangling(a pre-processing step), analyse the data, select an appropriate algorithm for prediction ,train and test the model and then Deploy it on a User Interface.
2. Purpose : Demand response plays an important role in modern life. Flat rate tariffs are no longer a viable solution. It is essential to make changes in prices of any service to meet the demand. Travel industry is one place where we see such pricing models. If you were a frequent traveller you might have observed that the flight rates, cab rates and even bus rates change depending on the surge and season. In such a case we need to be able to know the price of the ticket/ride at a particular point of time in future, so that we can make necessary decisions. Our objective is to build such a model and UI such that a user can give the necessary details and predict the price with good accuracy.

## Literature Survey :

1. Existing problem:The fluctuating prices in cab-rides, plane tickets can be a burden for us. We sometimes cannot make a decision whether the tickets at a certain point of time in future will be at an affordable price or not. Because of this we cannot make a decision whether to travel or wait till the prices come down. We should be able to know the price at the expected travel date. During high demand surge pricing is implemented by travel companies, the drivers need to know well in advance in such cases, so that they can concentrate on the area of high demand and the travel agency can deploy more vehicles there. To meet such supply/demand situations we need a learning model that will be able to accurately predict the price of a ride when the proper constraints are given.
2. Proposed solution:With the help of weather records in a particular area, the available cab Rides between different routes i.e., source and destination, we find the factors affecting the surge pricing of available cab services in both of the datasets and then used the random forest regression algorithm after data pre-processing for building the model in order to predict the surge prices of different cab services available at that time. We then build a Machine Learning model and deploy it on the IBM cloud. We will also be using the Flask web framework to create a web application.

## Theoretical Analysis :

### **Block Diagram** :

****

1. Software Designing:The cab-rides and weather datasets are collected and the datasets undergo pre-processing. Under pre-processing we have many tasks, but first we study the data by performing exploratory data analysis. This helps us to decide on what data can be used as input features and in pointing out the anomalies. Under pre-processing we merge data-frames, fill missing values, remove outliers, remove categorical data and split the data into train and test models. Then we use the best fitting ML algorithm after studying the data and we train the data on it. The model is then evaluated for its accuracy and if desired results are not met then we make appropriate changes to our model. The model training and testing is done on IBM Watson and the web application is built using Flask.

## Experimental Investigation :

## Flowchart:

## Result:

## We used random forest regression to build our machine learning model. We got an accuracy of \_\_95 %\_\_\_\_. The model was trained in ibm watson and deployed in flask framework.

## Advantages & Disadvantages:

|  |  |
| --- | --- |
| · Higher Profit & Sales  · Adjusting to the Competition  · Flexibility  · Better Inventory Management | · Customer Dissatisfaction  · Loss of Sales  · Gaming the system  · Not Applicable Everywhere  · Price Fluctuation |

## Applications:

Every day the price of travel was changed due to the demand for public uses. The framework developed for the price prediction is analyzed for the travel plans. For the same travel plan offered at a fixed price for a particular group of customers, our proposed model saw a final fare with a lesser number of errors in predicting customer planning. As time progresses and more data are collected, the supervised learning will produce more accurate results and will be helpful in determining fare optimizer and dynamic availability of adjustments and continuously improve future recommendations.

## Conclusion:

Demand side response programs with dynamic pricing have significant potential to optimize the demand-supply of the customer. In our project we built an application that takes various input features such as distance, cab-type, destination, etc and predicts the results. The traveller can enter such details and get the predicted price easily in an easy to use UI. The model accuracy which we got is around 95%. The performance of the model is validated from the results. This approach can be used for any application/business that uses dynamic pricing.

## Future Scope:

Further this work can be carried out using different machine learning algorithms and techniques in order to get higher efficiency and lower errors. And as time passes we will be having more and more data, which will help in getting a better training of models.

Overall the project has a great scope in any industry which implements dynamic pricing as it helps the user in predicting the price as well as the organization as unpredictability of the prices could lead to loss of customer loyalty.

## Bibliography :

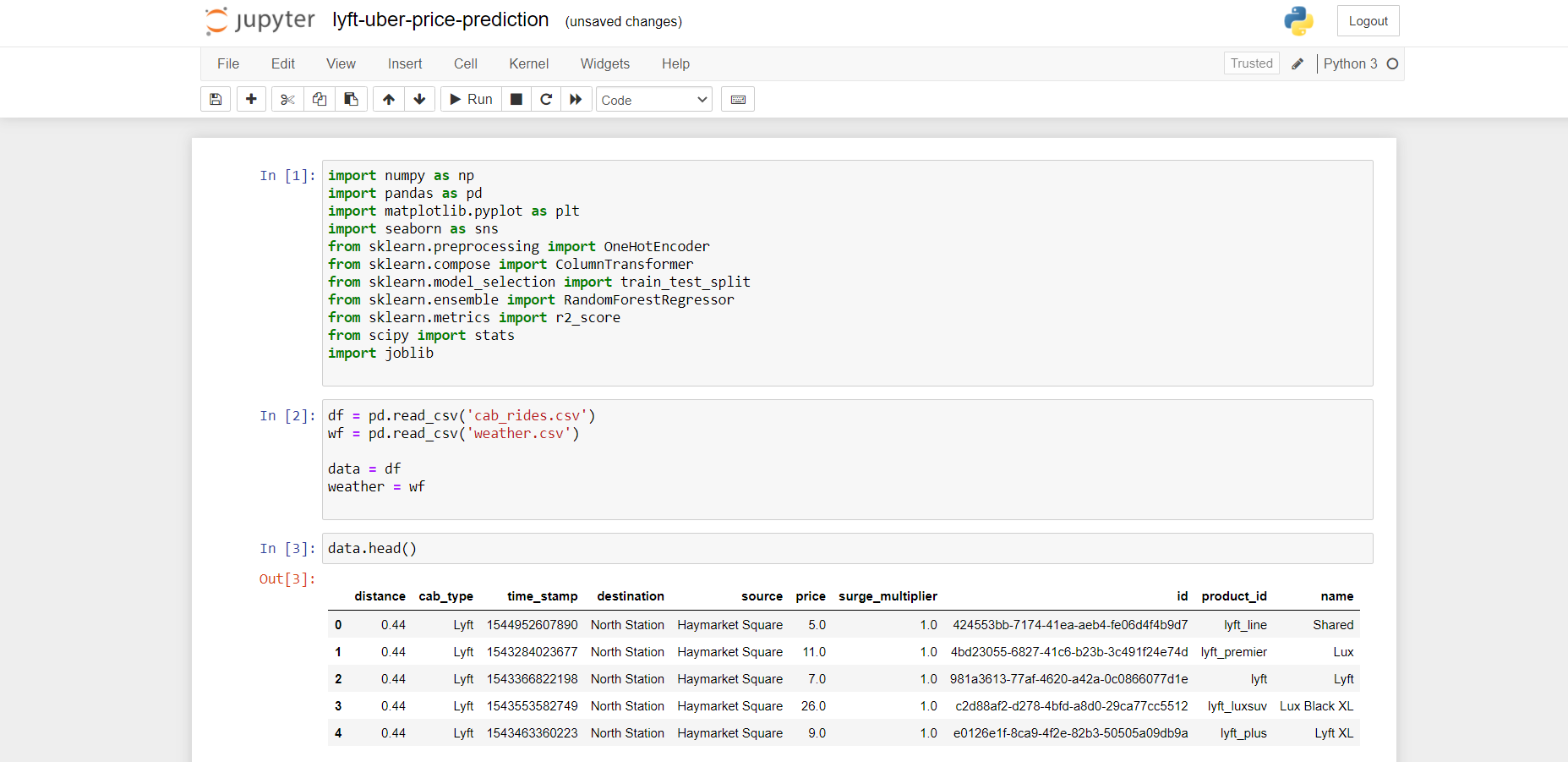
## Major references: <https://smartinternz.com/Student/guided_projects>

## Other References: Kunal, Arora & Kaur, Sharanjit & Sharma, Vinod. (2021). Prediction of Dynamic Price of Ride-On-Demand Services Using Linear Regression. International Journal of Computer Applications & Information Technology. 13. 376-389.

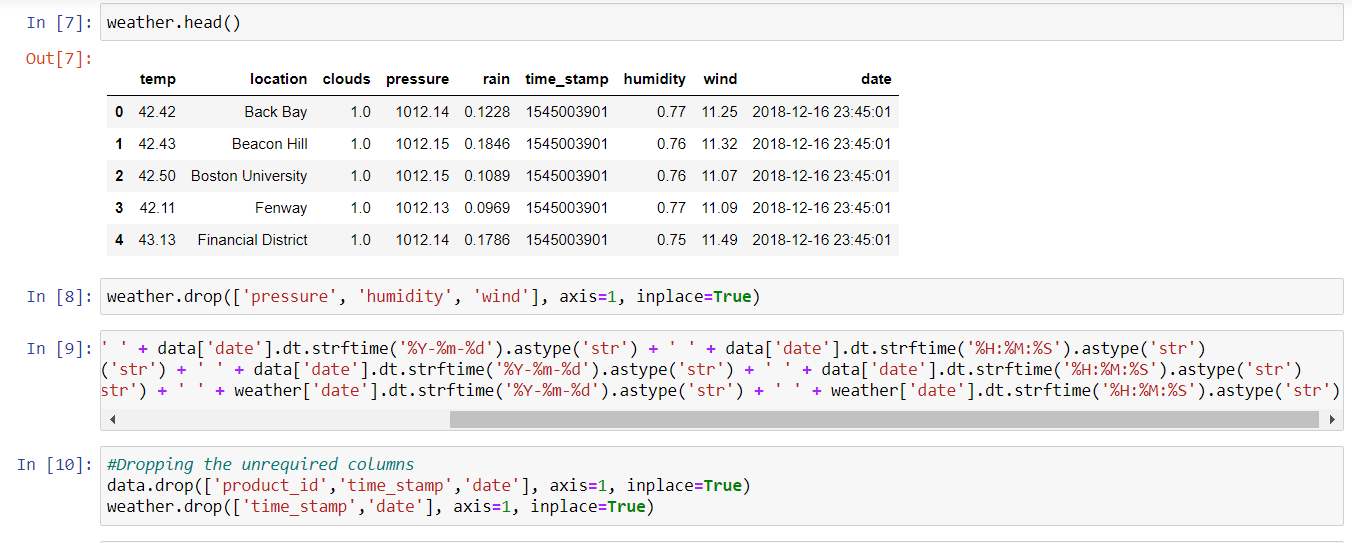
## Appendix :

## Source Code:

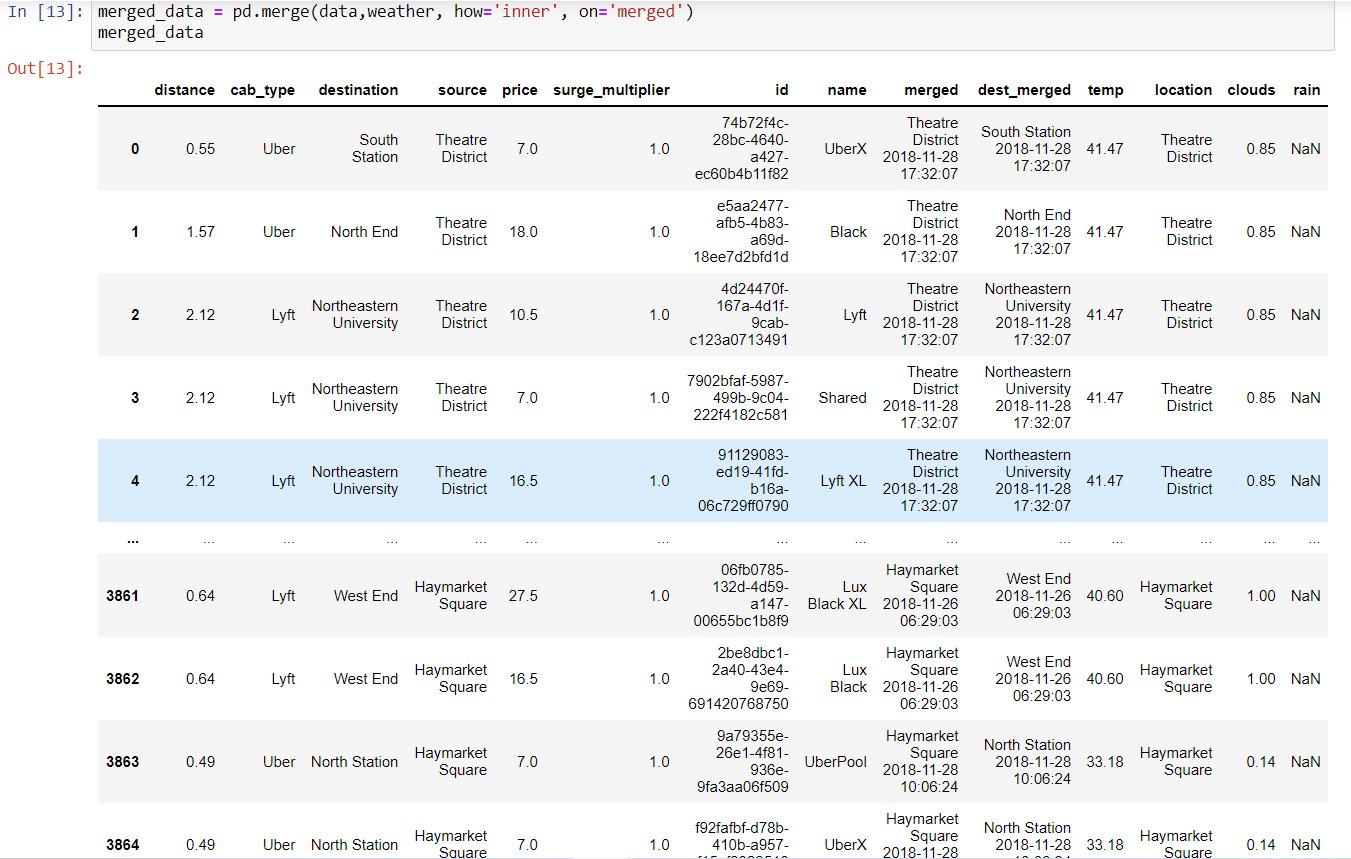
**Importing Libraries:**

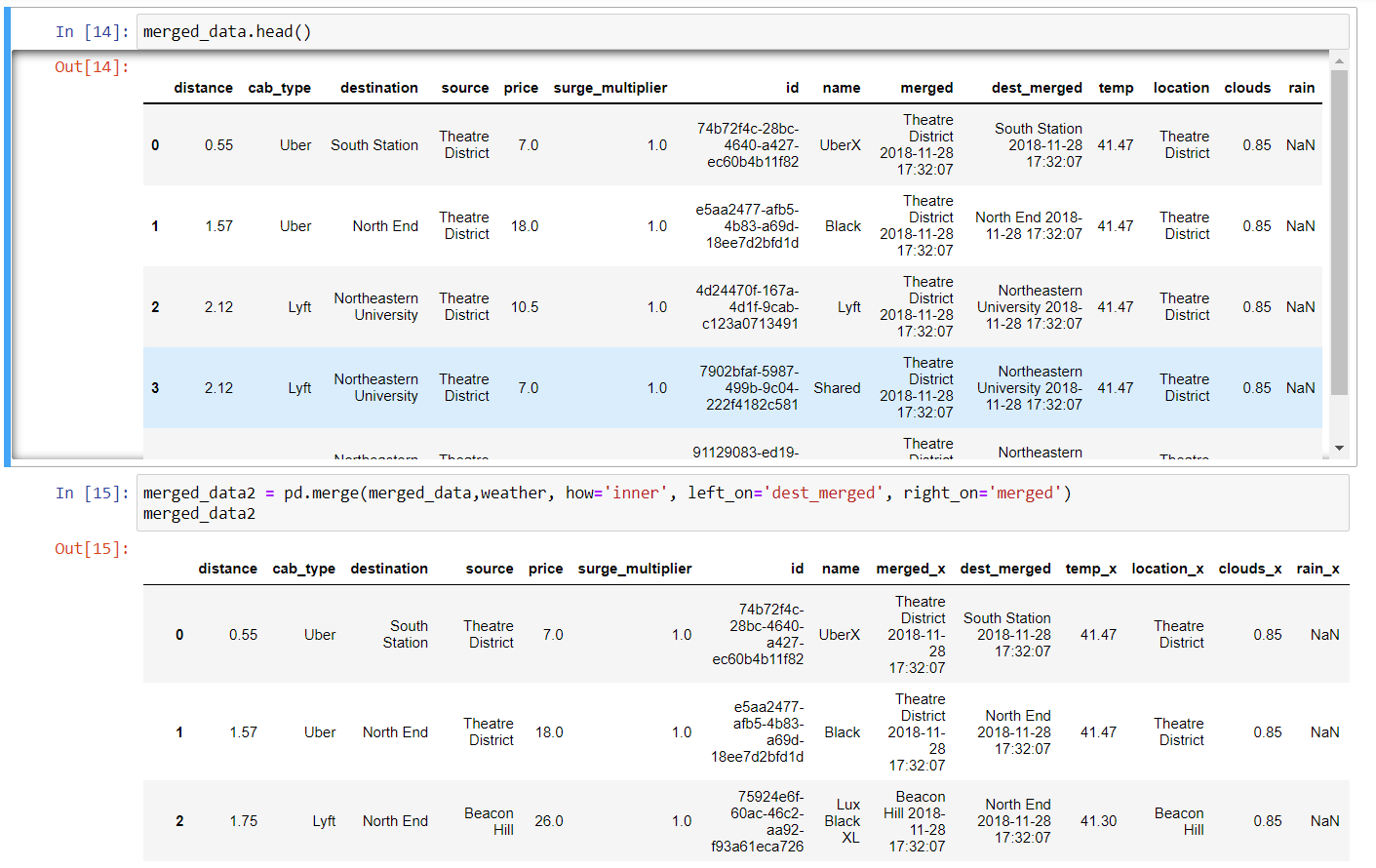




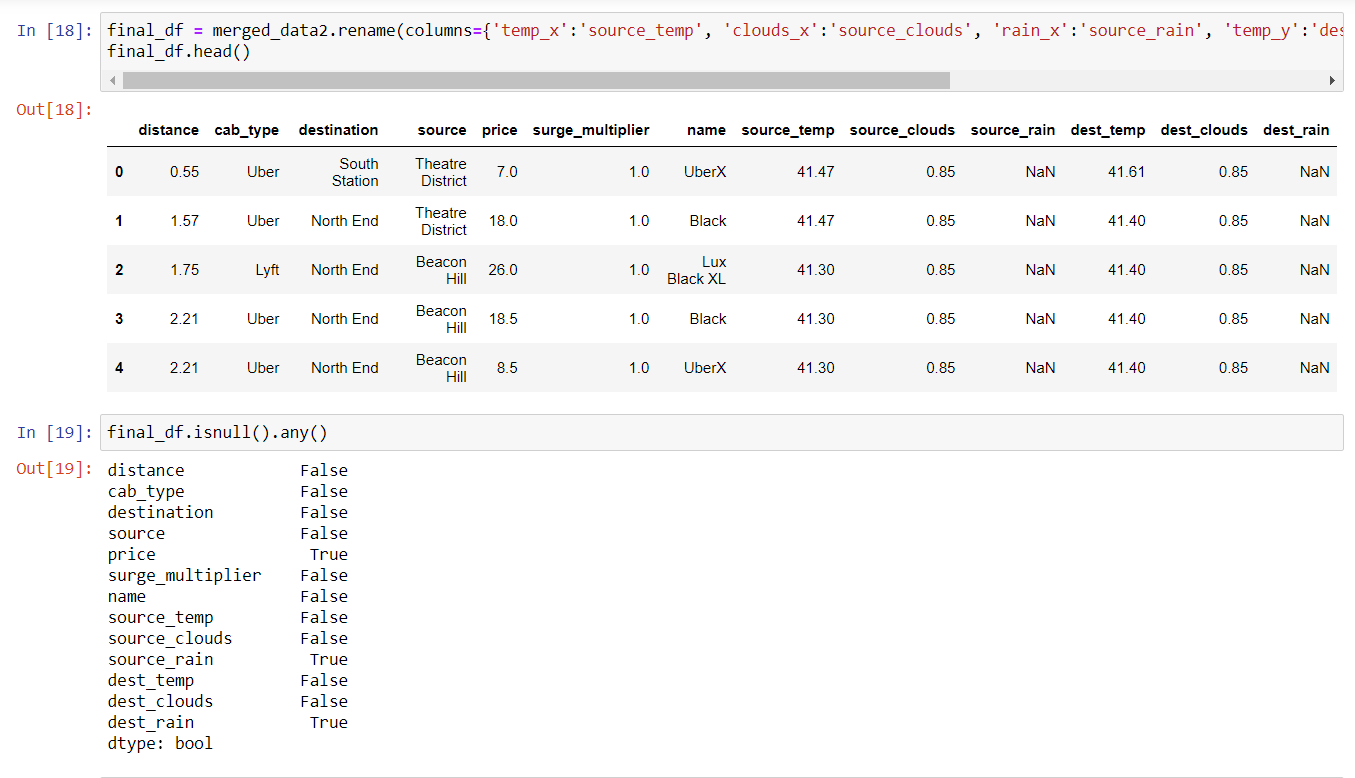


**Merged Data:**

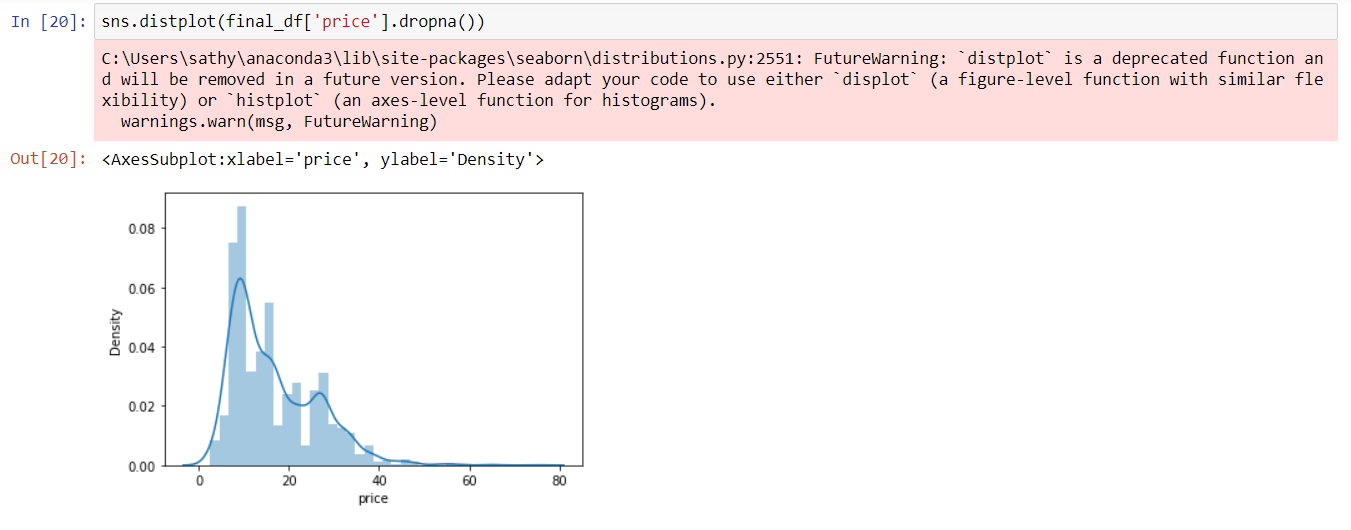




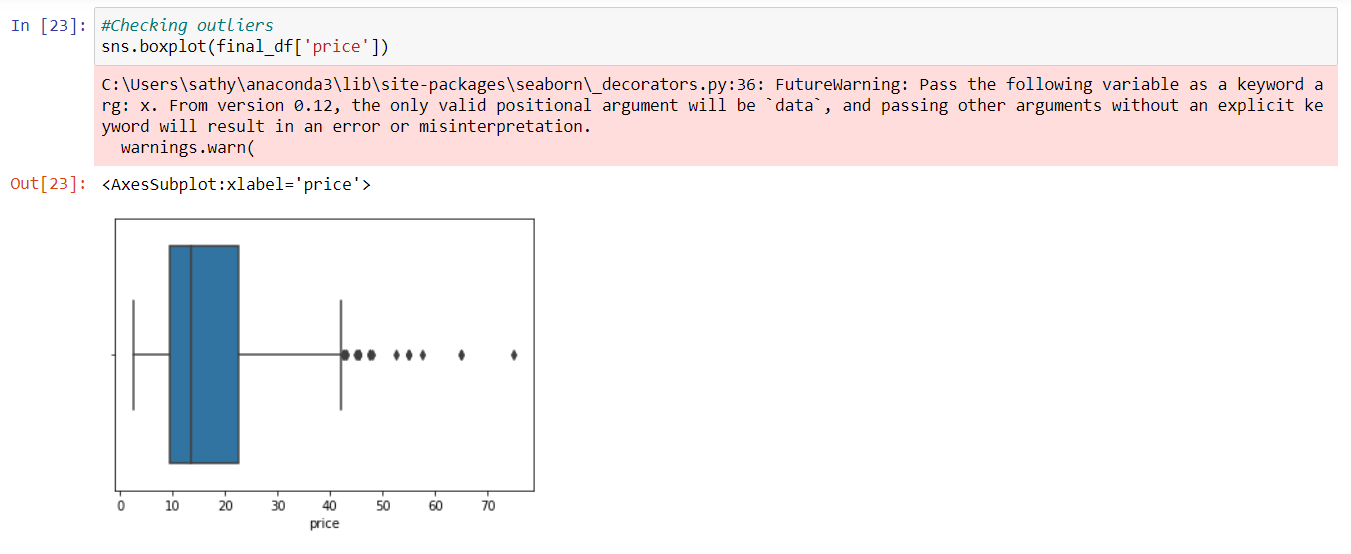
**Final Dataframe before checking outliers:**



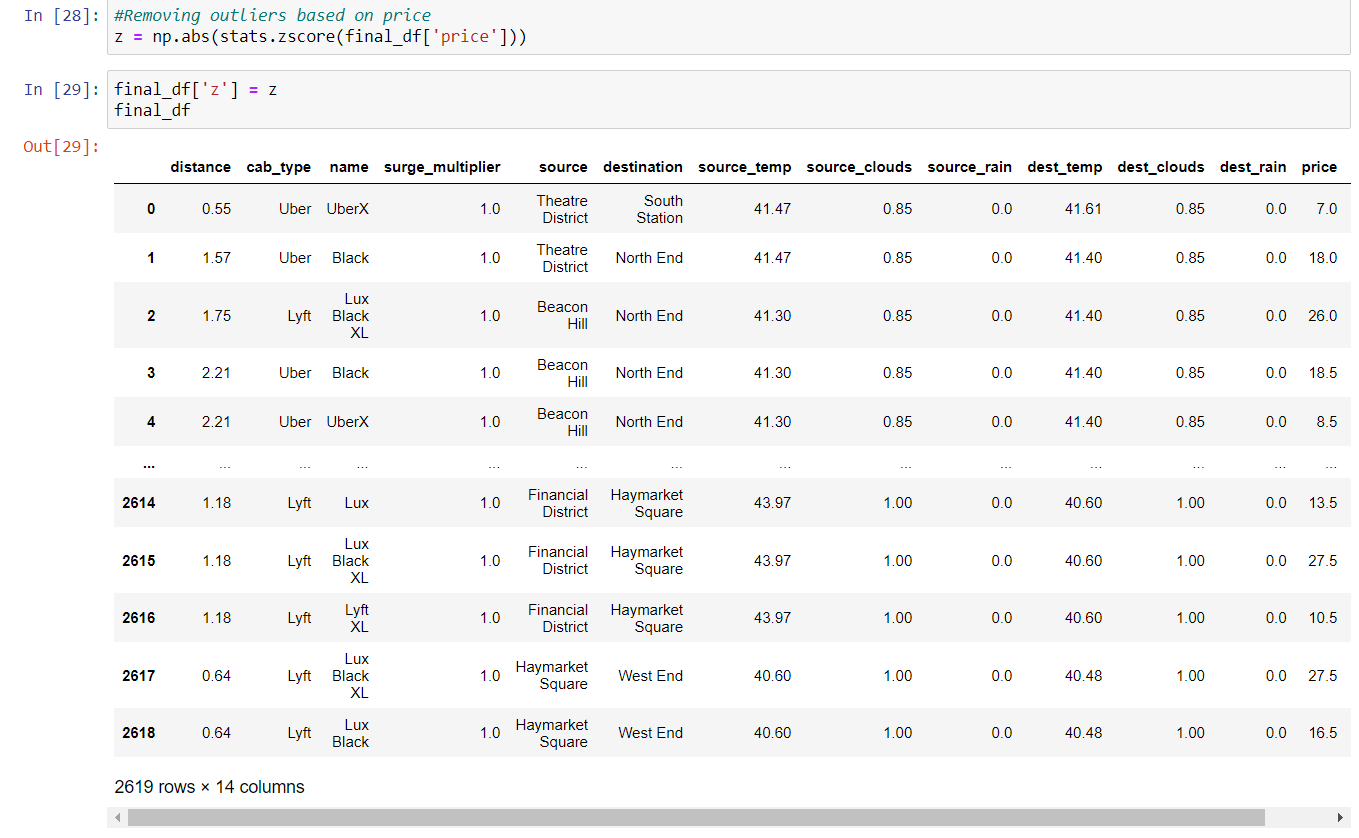
**Checking the type of distribution of “price” column to fill the NULL values:**



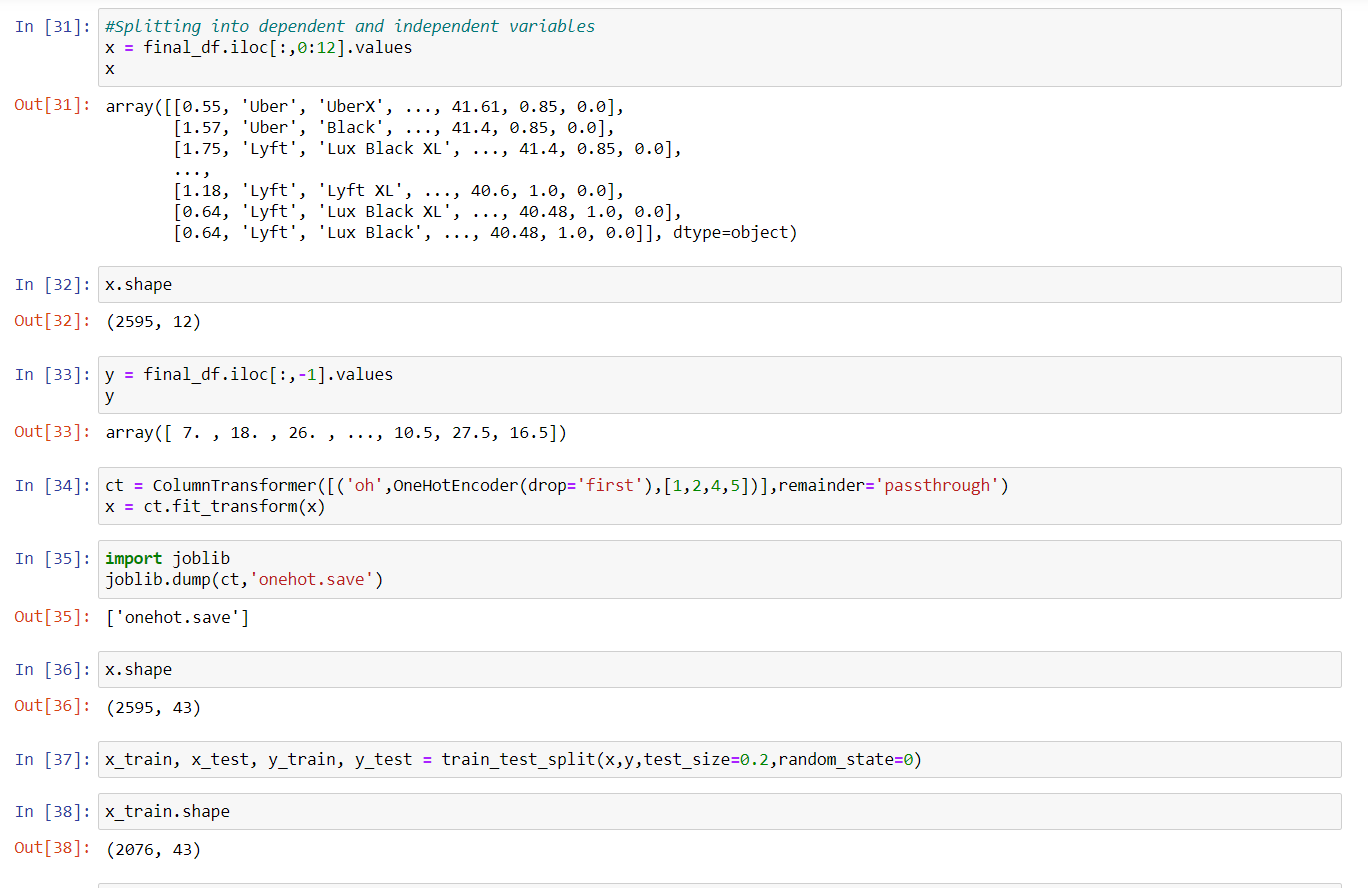
**Checking the outliers:**



**Removing outliers:**



**Splitting Data into independent and dependent variables and also splitting the data as train data and test data:**



**Model building and saving the model:**



## UI/Output Screenshot:

