**CAR PRICE PREDICTION**

**Project Report**

**<Version 1.0>**

Industrial Training

(ECS 599)

**BACHELOR OF TECHNOLOGY (CSE)**

**Specialized in Artificial Intelligence**

|  |  |
| --- | --- |
| PROJECT GUIDE:  **Ms. Neeraj Kumari**  **Assistant Professor(CCSIT)** | SUBMITTED BY:  **Prasuk Jain**  **TCA1959025** |

October, 2021



**FACULTY OF ENGINEERING & COMPUTING SCIENCES**

**TEERTHANKER MAHAVEER UNIVERSITY, MORADABAD**

**ACKNOWLEDGEMENT**

I am very thankful and acknowledge to my teacher as well as our principal for their help to make my training **Machine learning** a successful one. Which also helped me in doing a lot of research and I came to know about so many new things. I am really thankful to them.

First of all, I want to express my gratitude mainly to my supervisor of intershala who helped me from the very beginning of my training in Machine Learning. And I am making this project not only for marks but to also increase my knowledge.

I feel thankful to the college staff for giving me such a big opportunity. I believe I will enroll in more such events in the coming future. I ensure that this project was done by me and is not copied.

Prasuk Jain

TCA1959025

**Place: Moradabad**

**Date: 19 December 2021**

**DECLARATION**

We hereby declare that this Project Report titled **Car Price Prediction** submitted by me and approved by our project guide,the College of Computing Sciences and Information Technology (CCSIT), Teerthanker Mahaveer University, Moradabad, is a bonafide work undertaken by us and it is not submitted to any other University or Institution for the award of any degree diploma / certificate or published any time before.

**CERTIFICATE OF TRANING**

****

**Brief About the Company**

Internshala is an internship and online training platform, based in Gurgaon, India. Founded by Sarvesh Agrawal, an IIT Madras alumnus, in 2011, the website helps studentsfind internships with organizations in India. Internshala is India's no.1 internship and training platform with 40000+ paid internships in Engineering, MBA, media, law, arts, and other streams. Currently, internships are posted for free and a fee of INR 4999 is charged from the employers for posting a full-time job. Visit Internshala Employer Section to know more. Internshala is a dot com business with the heart of dot org. We are a technology company on a mission to equip students with relevant skills & practical exposure to help them get the best possible start to their careers. Imagine a world full of freedom and possibilities. And these skills can be acquired when students work and learn through internships and jobs. In India, the concept of internships is picking pace and is progressing rapidly. Further, every Internshala training also gives you complimentary access to a soft skills training module that would be useful when you start applying for internships or jobs after your training. ... But in reality, any training certificate (be it from Internshala or anywhere else) does not mean anything in itself. We don't charge anything from anyone. But we make sure that interns are paid a stipend by companies. From day one, I was certain that we would not allow unpaid internships on our platform because I had seen companies use interns for free work. Internshala didn't want to encourage this. It means that the employer has sent you a message regarding the further procedure required in the selection process. Please check your chat box after logging to your Internshala account and respond accordingly.

Table of Contents

[1 Project Title 6](#_Toc31652427)

[2 Problem Statement 6](#_Toc31652428)

[3 Project Description 6](#_Toc31652429)

[3.1 Scope of the Work 6](#_Toc31652430)

[3.2 Project Modules 6](#_Toc31652431)

[3.3 Context Diagram (High Level) 6](#_Toc31652432)

[4 Implementation Methodology 6](#_Toc31652433)

[5 Technologies to be used 6](#_Toc31652434)

[5.1 Software Platform 6](#_Toc31652435)

[5.2 Hardware Platform 7](#_Toc31652436)

[5.3 Tools, if any 7](#_Toc31652437)

[6 Advantages of this Project 7](#_Toc31652438)

[7 Assumptions, if any 7](#_Toc31652439)

[8 Future Scope and further enhancement of the Project 7](#_Toc31652440)

[9 Project Repository Location 7](#_Toc31652441)

[10 Definitions, Acronyms, and Abbreviations 8](#_Toc31652442)

[11 Conclusion 8](#_Toc31652443)

[12 References 9](#_Toc31652444)

**Appendix**

**A:Data Flow Diagram (DFD)**

**B:Entity Relationship Diagram (ERD)**

**C:Use Case Diagram (UCD)**

**D:Data Dictionary (DD)**

**E: Screen Shots**

# Project Title

Car Price Prediction

# Problem Statement

If we want to sell our old car in present time , this project will help you to find a genuine selling price of your car. By providing the details (year, buying price, fuel type, transmission type, seller type) and model will provide you a best selling price by just giving these details.

# Project Description

# This model will helps to find better selling price of your old car. And this is done by mining the Big Data of the previous records of the car to train the machine and predict the price of a car. the machine was trained using the machine learning model which give the most accurate result. The main objective of this project is to predict the selling price of cars. This project is divided into four sections

# (i)Data Collection (ii) Comparison of machine learning models on collected data (iii) Training of system on most promising model (iv) Testing

## Scope of the Work

The main scope of this Machine learning model is

* To build a perfect ML model for Car Price Prediction
* It is used to make comfort ability for people for predict best selling price

## Project Modules

Machine Learning: Machine learning (ML) is a type of artificial intelligence (A.I.) that allows software applications to become more accurate at predicting outcomes without being explicitly programmed to do so. Machine learning algorithms use historical data as input to predict new output values.

* **Datasets**: A collection of instances is a dataset and when working with machine learning methods we typically need a few datasets for different purposes.
* **Training Dataset**: A dataset that we feed into our machine learning algorithm to train our model.
* **Testing Dataset**: A dataset that we use to validate the accuracy of our model but is not used to train the model. It may be called the validation dataset.

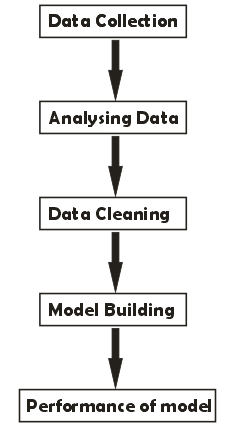
This project uses the following software and Python libraries:

1. **Python 3.7**
2. **Numpy**
3. **Pandas**
4. **Scikit- learn**
5. **Jupyter notebook**
6. **Matplotlib.pyplot**
7. **seaborn**

This project contains three files:

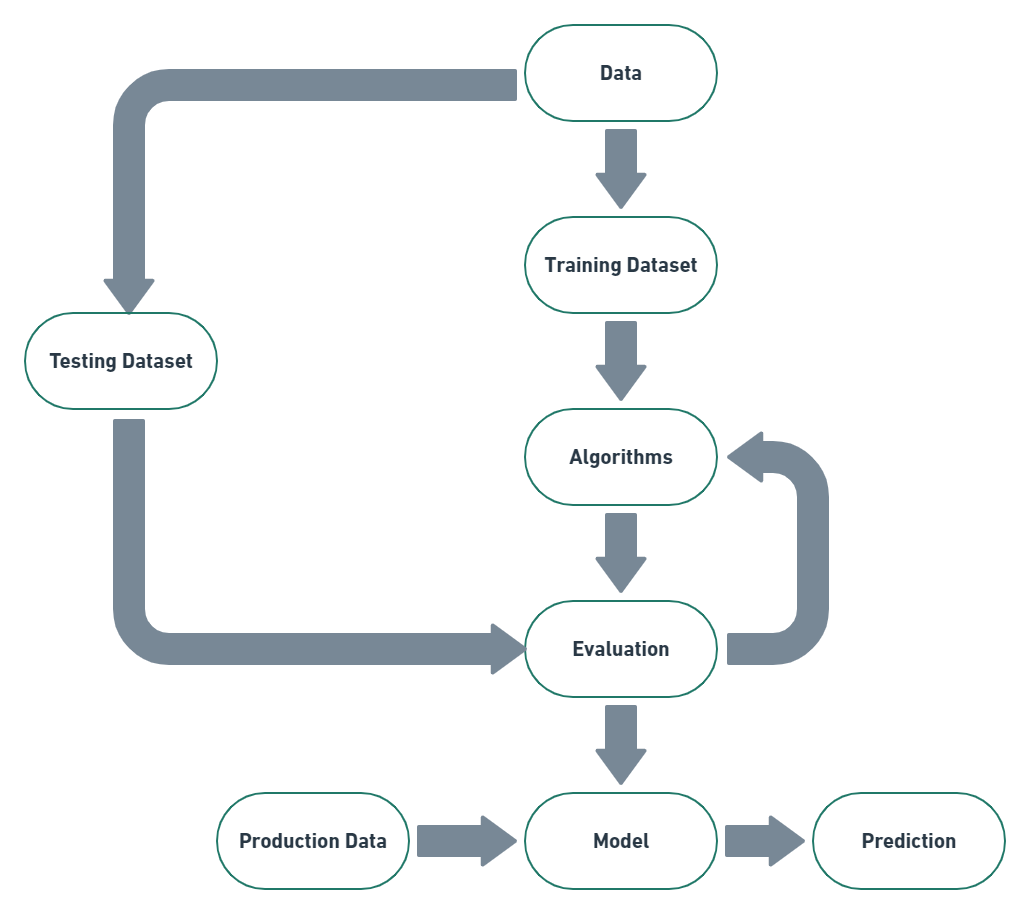
* **car\_ price\_prediction.ipynb:** This is the main file where you will be performing our work(code) on the project.
* **car data.csv:** The project dataset. we'll load this data in the notebook.
* **joblib.py:** for giving inputs from the user.

## Context Diagram (High Level)



# Implementation Methodology

For predict the selling price of a car by help of the computer and the Computer Science is capable of making such a system using Artificial Intelligence, which can make this decision accurately and quickly. Using data science and machine learning, which is responsible to deal with the large amount of data efficiently, and some algorithms of Machine Learning, a prediction system is made which, on the basis of some training data sets capable of identifying the selling price of a old car. Machine Learning algorithms like Decision Tree ,linear Regression , lasso regression etc. are used for the analysis. These are efficient algorithms that are followed for data analysis and prediction making. The system will look into some basic information of the car such as buying year, transmission type of a car, fuel type of a car , buying price etc., and after analyzing all this information, using visualization and machine learning algorithms, it will come to a decision.



# Technologies to be used

## Software Platform

* Python 3.7
* Jupyter Notebook

## Hardware Platform

CPU : I3 PROCESSOR

TOTAL RAM : 2 GB

HARD DISK : 100 GB

OS : 32 Bit Windows

## Tools

* Python idle 3.7.1.
* Visual Studio Code (Version 1.6.0 and above)
* Jupyter Notebook

# Advantages of this Project

* User convinces
* Minimal Documentation
* No physical submission
* Faster process
* User friendly

# Assumptions

NONE

# Future Scope and further enhancement of the Project

Future enhancement is to add some more additional functionality such as-

* + Chat box for customer queries.

# Project Repository Location

| **S#** | **Project Artifacts (softcopy)** | **Location** (Mention Lab-ID, Server ID, Folder Name etc.) | **Verified by Project Guide** | **Verified by Lab In-Charge** |
| --- | --- | --- | --- | --- |
|  | Project Synopsis Report  (Final Version) | https://docs.google.com/document/d/1pRCYvwcAIsCEwN9HuKlqhEX57cOdlDYEbLX\_sdqw4UQ/edit?usp=sharing |  |  |
|  | Project Progress updates | NONE |  |  |
|  | Project Requirement specifications | NONE |  |  |
|  | Project Report  (Final Version) | NONE |  |  |
|  | Test Repository | NONE |  |  |
|  | Any other document, give details | https://docs.google.com/spreadsheets/d/1m0-a9Huz5K30o1yJ7xIy4dqXDLcJWIRx1JvDqUDDctE/edit?usp=sharing |  |  |

# Definitions, Acronyms, and Abbreviations

|  |  |
| --- | --- |
| **Abbreviation** | **Description** |
| GUI | Graphical User Interface |
| RAM | Random Access Memory |
| OS | Operating System |
| CPU | Central Processing Unit |
| Dataset | A collection of instances is a dataset and when working with machine learning methods we typically need a few datasets for different  purposes. |
| Training Dataset | A dataset that we feed into our machine learning algorithm to train our model. |
| Testing Dataset | A dataset that we use to validate the accuracy of our model but is not used to train the model. It may be called the validation dataset. |
| Accuracy | Accuracy is defined as the percentage of correct predictions for the test data. It can be calculated easily by dividing the number of correct predictions by the number of total predictions. |
| ML | Machine learning |
| AI | Artificial Intelligence |
| PY | It is a extension of saving python file. |

# Conclusion

* Without any manual calculation you can find easily better selling price of your old car by providing some information to the machine. There is need of automation of this system so that people can find easily there car price with out any type of problems. In this project some ML algorithms like logistic Regression, Decision tree,random forest etc. this project focuses on making the tedious task mechanized. All the required algorithms were implemented successfully and accurate results were generated.

# References

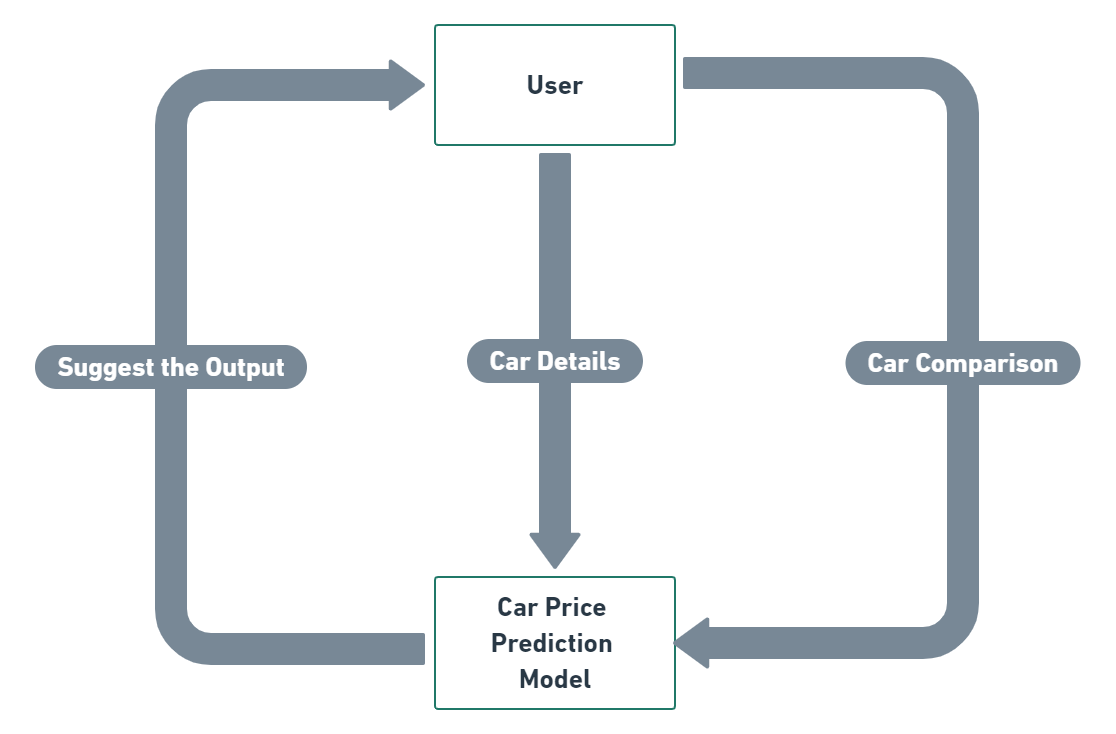
* ***https://www.udemy.com***
* ***https://internshala.com***
* ***https://www.analyticalvidhya.com***
* ***https://www.codingninjas.com***
* ***https://www.youtube.com***

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **S#** | **Reference Details** | **Owner** | **Version** | **Date** |
|  | Project Synopsis | Prasuk jain | 1.0 | 19-12-21 |
|  | Project Requirements | NONE |  |  |
|  |  |  |  |  |

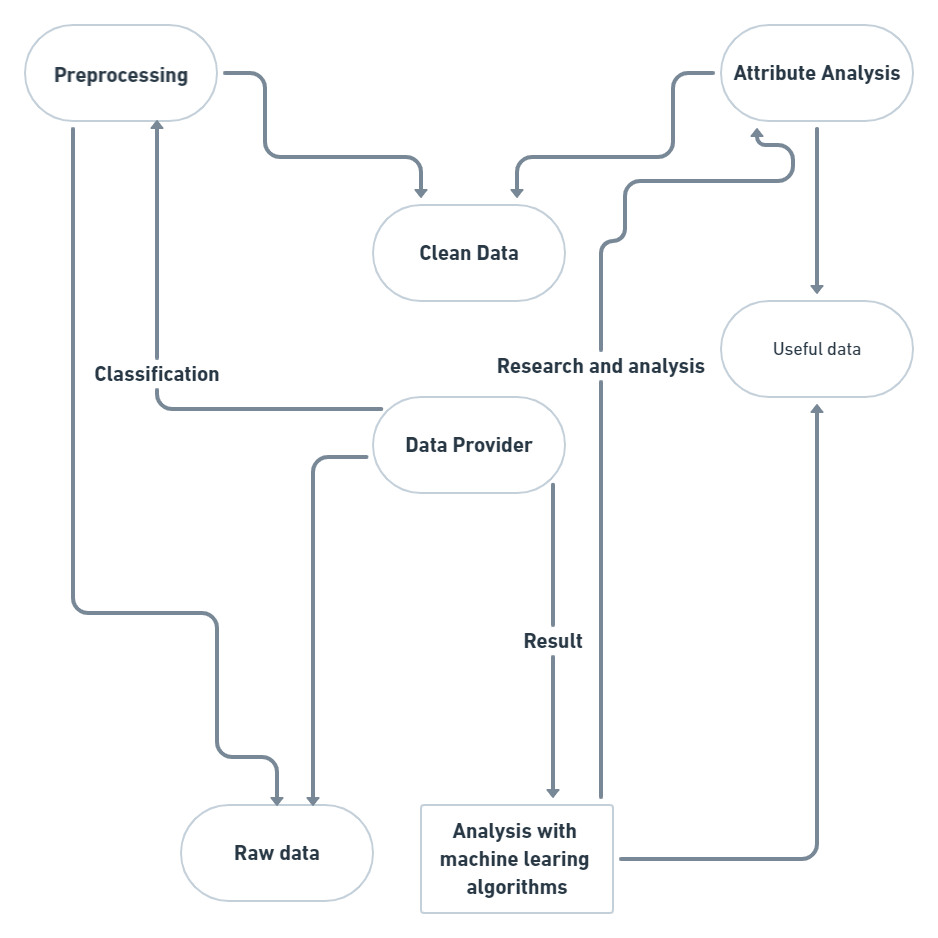
**Annexure A**

**Data Flow Diagram (DFD)-1**

**(Mandatory)**

****

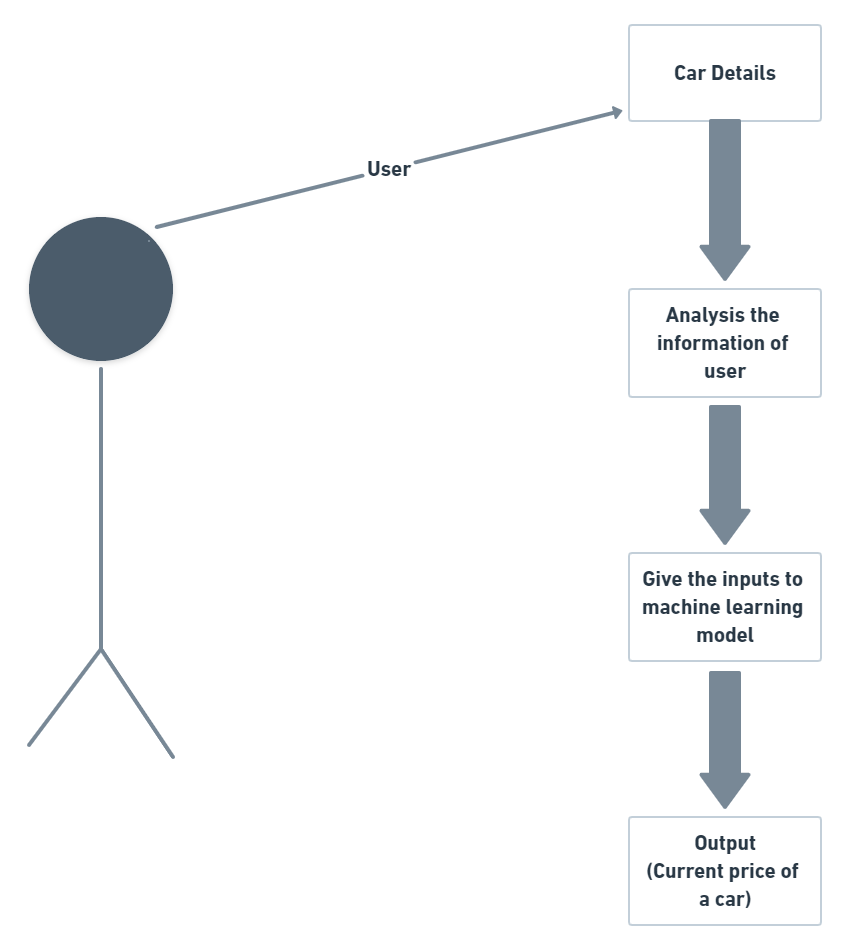
**Data Flow Diagram (DFD)-2**

****

**Annexure C**

**Use-Case Diagram (UCD)**

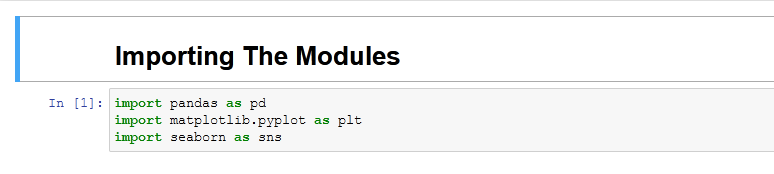
**(Optional)**

****

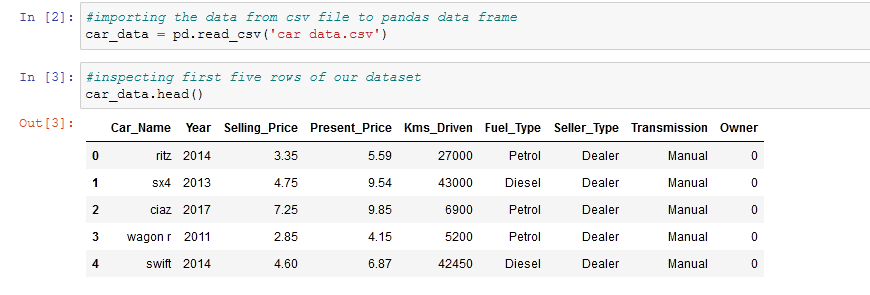
**Annexure E**

**Screen Shots**

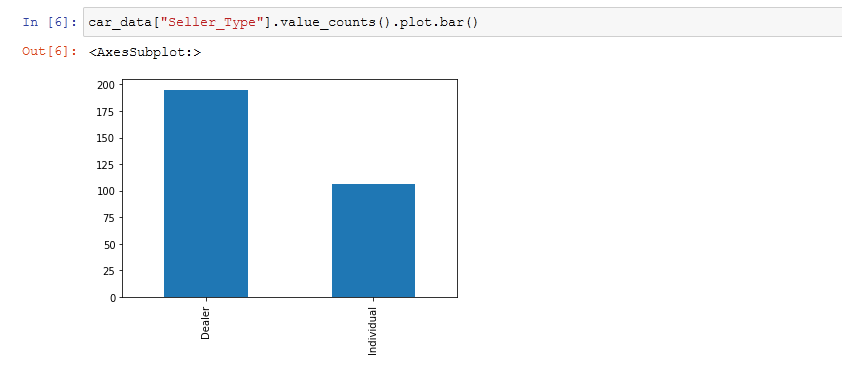
**Importing of modules:**

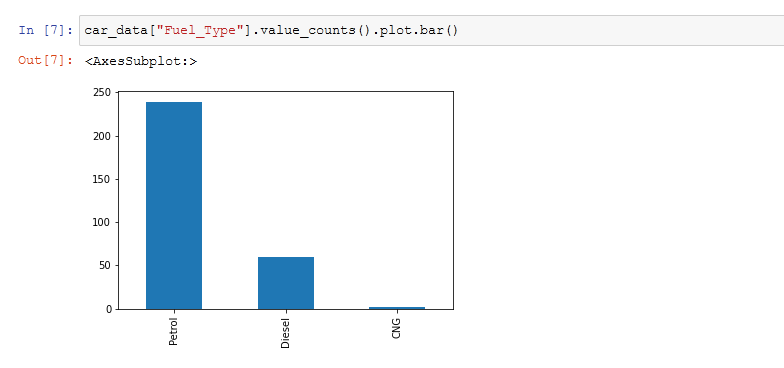
****

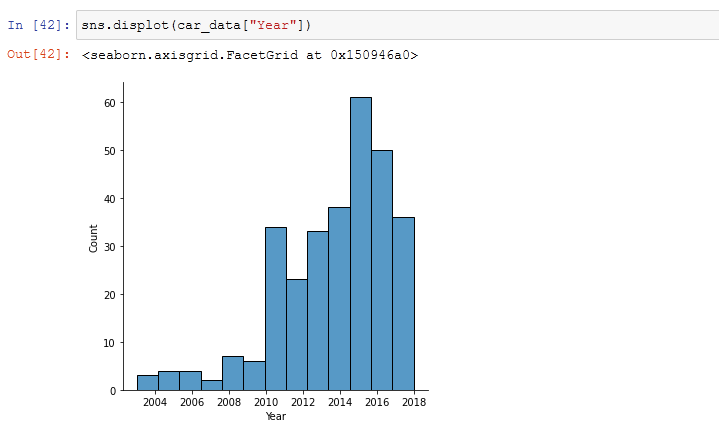
**Data Collecting and Processing:**

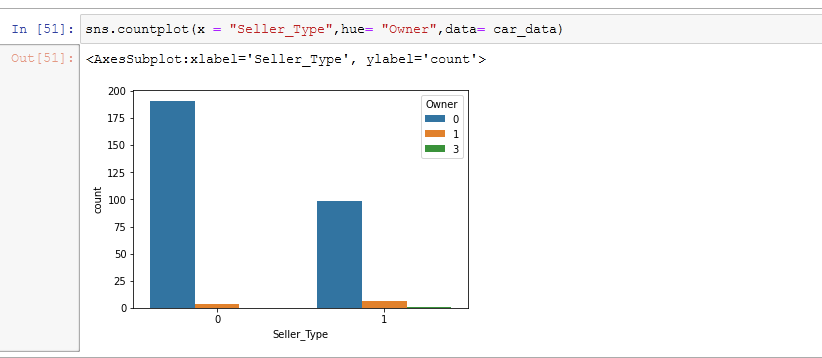
****

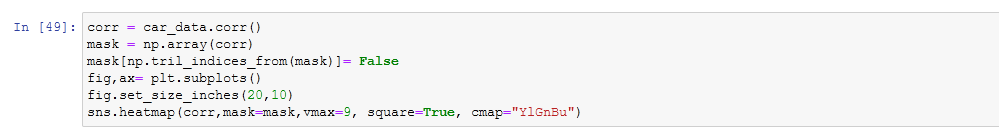
**Visualizing and Analysis of Data:**

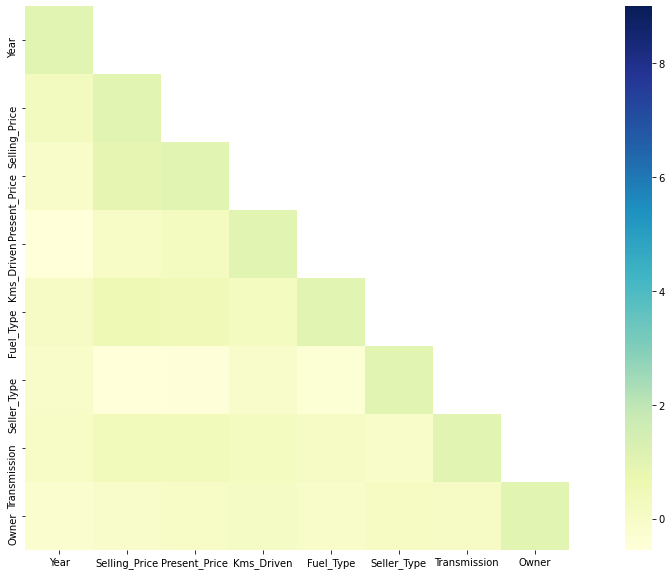
****

****

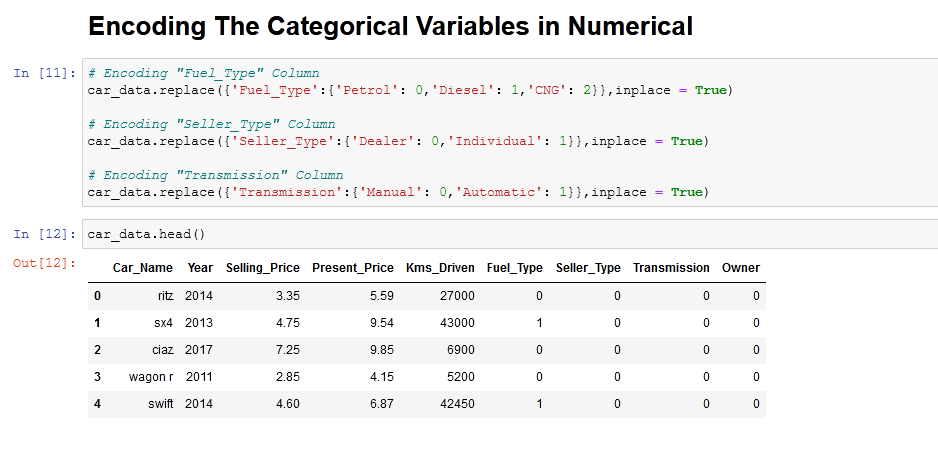
****

****

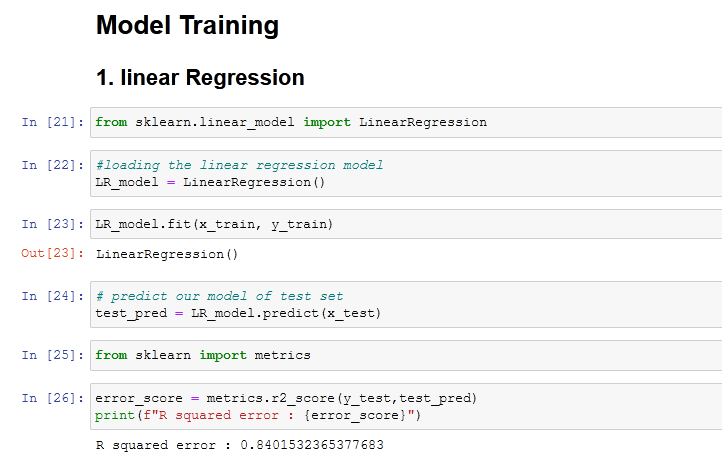
****

****

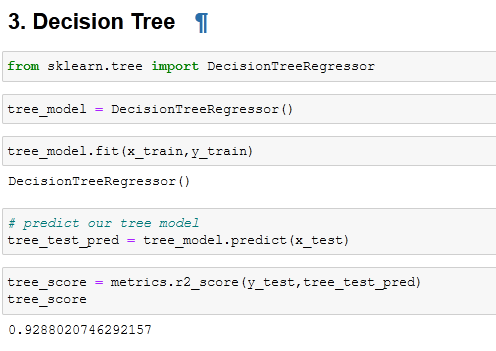
**Encoding And Decoding of Data:**

****

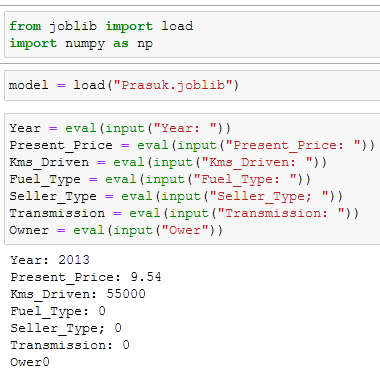
**Building Machine learning Model:**

****

****

****

**Input For Analysis The Car Price:**

****