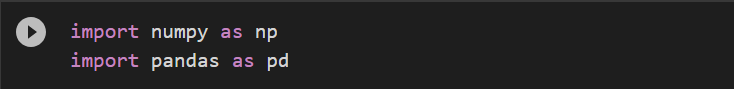
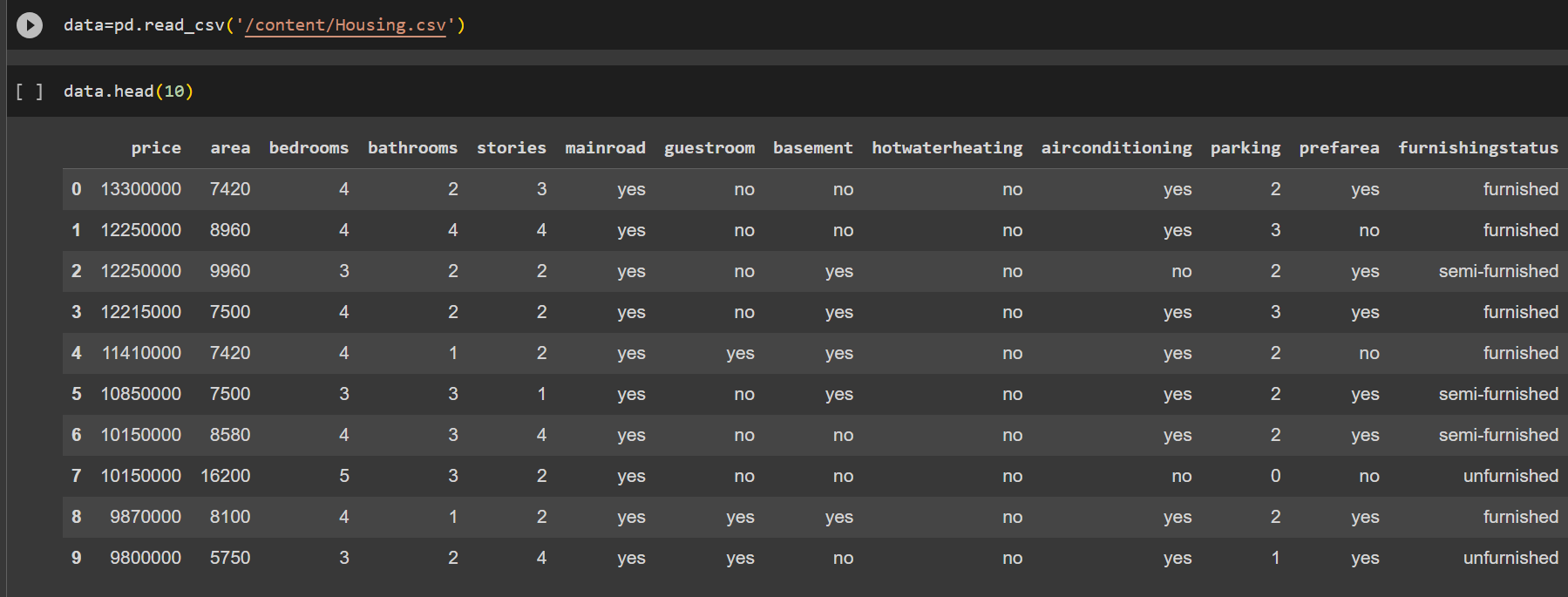
Project On Fetch to House price & Location by using Machine Learning Models

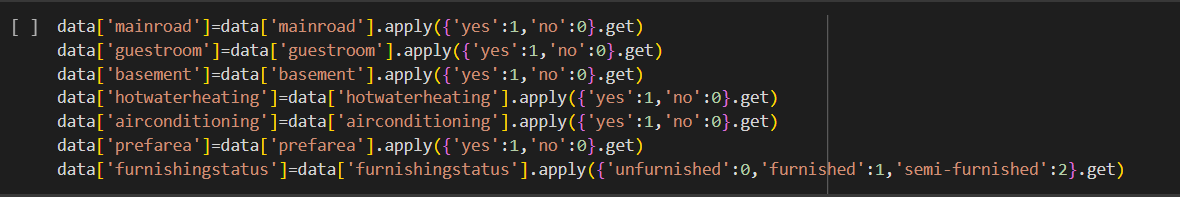
* AIM: To create a Data Science project, where we'll be predicting the price of a new house. & that house is available on main road or not. So, here I have used machine learning models with the help of csv-dataset(s) provided which contains 6 numerical & 7 categorical values of different-different facilities. It will give us idea of how much cost of new house & what kind of facilities they can get.
* Steps to be taken in the project is sub-divided into the following sections. These are:
* Loading necessary libraries such as 'numpy','pandas','sklearn.model' etc.
* Loading Dataset(s) as a csv file. Here we are using two different files for training & testing the models.
* Data cleaning was performed by changing string values to integer values.
* Visualisation of According to price & area getting facilities using Tableau.
* Splitting the dataset into independent & dependent sets (only train dataset was taken in use).
* Importing the train\_test\_split model from sklearn.model for splitting data into train & test sets.
* Importing one Regression model & different kinds of Classification models & then training those models with the help of fit().
* Predicting the trained models & then checking their accuracy of the model using confusion matrix & accuracy score.
* Then recalled test\_dataset & splitted the dataset into testing & training sets using X1\_train & X1\_test.
* Then, trained the test\_dataset with train\_dataset with the help of better accuracy’s model.
* Finally, predicted in Regression\_model the cost price, & Classification\_model the facilities by help of test\_dataset.
* Steps of creating ML model:
* Step-1: importing numpy as np & pandas as pd for loading & reading the data-set.

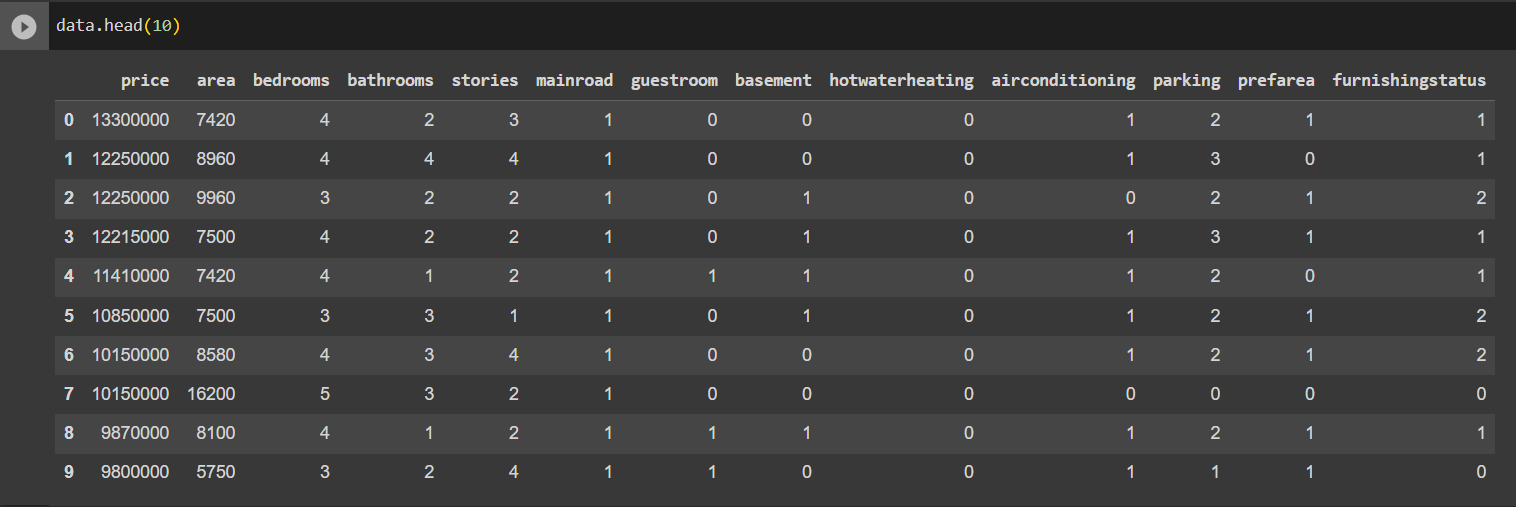


* Step-2: Loading the CSV-dataset in the variable name(s) ‘train\_data’ & ‘test\_data’. Then viewing the the data(s) with data.head(10). Here’ 10 means we are fetching the first 10 colums only.



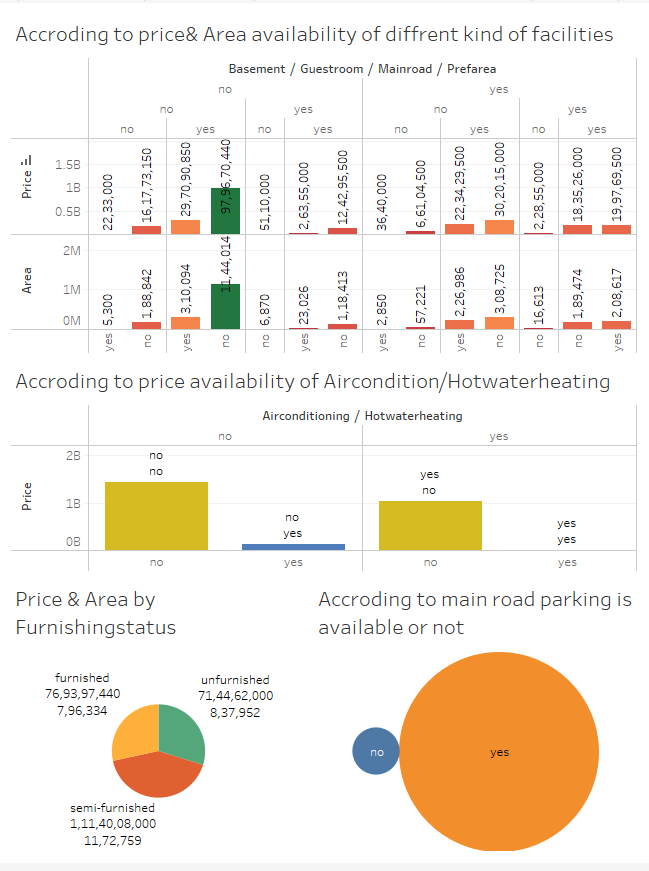
* Step-3: Cleaning the dataset by changing any categorical values to numerical values.





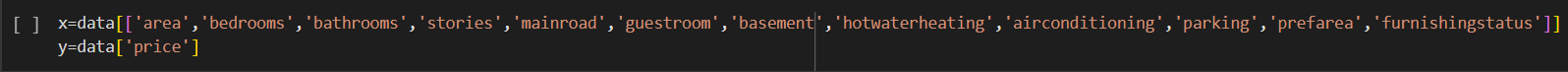
-Viewing dataset after using (One-Hot Encoding)

* Step-4: visualizing the different kind of facilities are available according price & area.

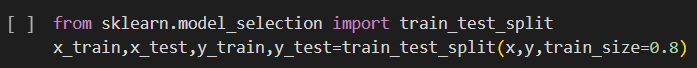


-Insights of Price & area consideration we fetch the houses furnishing status, different kind of facilities & main is we can see the houses are near by main road or not. Also we can fetch more insights from this dataset.

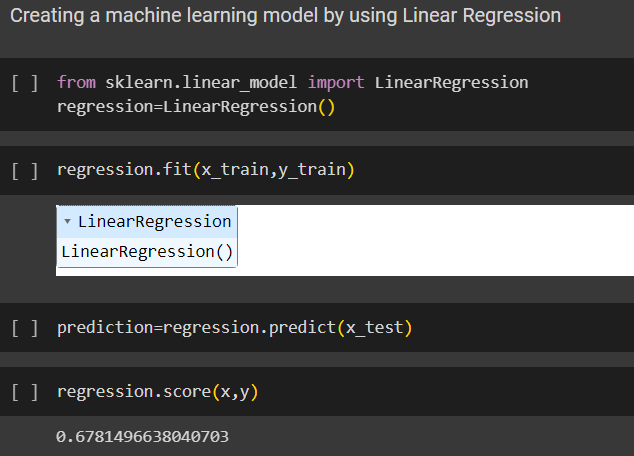
* Step-5: Splitting the dataset into dependent & independent sets.



* Step-6: Importing train\_test\_split from sklearn.model library for splitting the data into train & test sets.

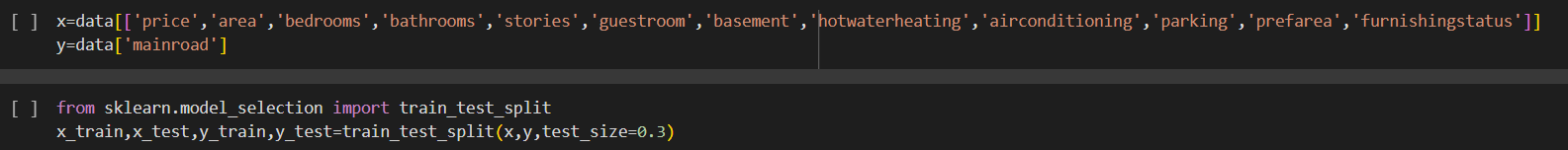


* Step-7: Importing Linear regression from sklearn.linear\_model. Then train the model by providing train & test sets as x & y. and finally predict the train model & checked score of the Linear model.

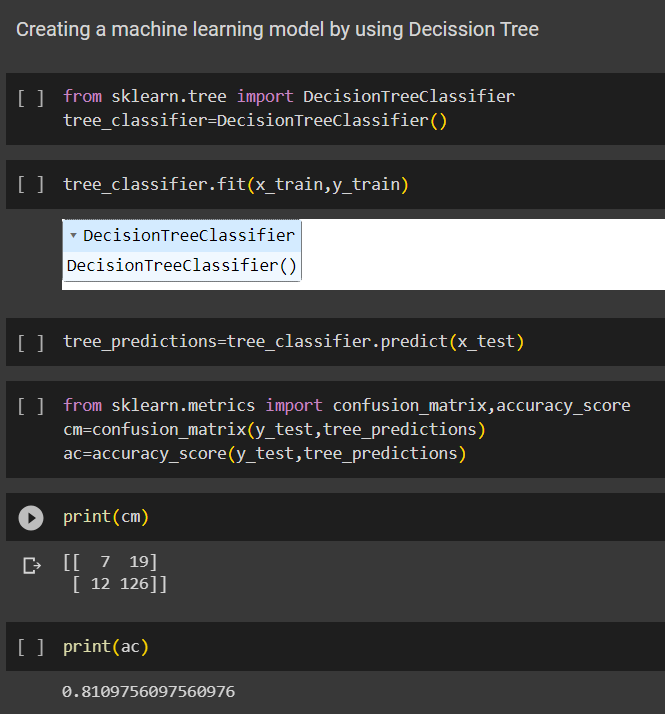


-In the above model we can see that the score obtained is 67% which is quite good. By using this Linear score we can predict the approx price of the new house.

* Now we can fetch the house is available on main road or not by using different-different Classification models:

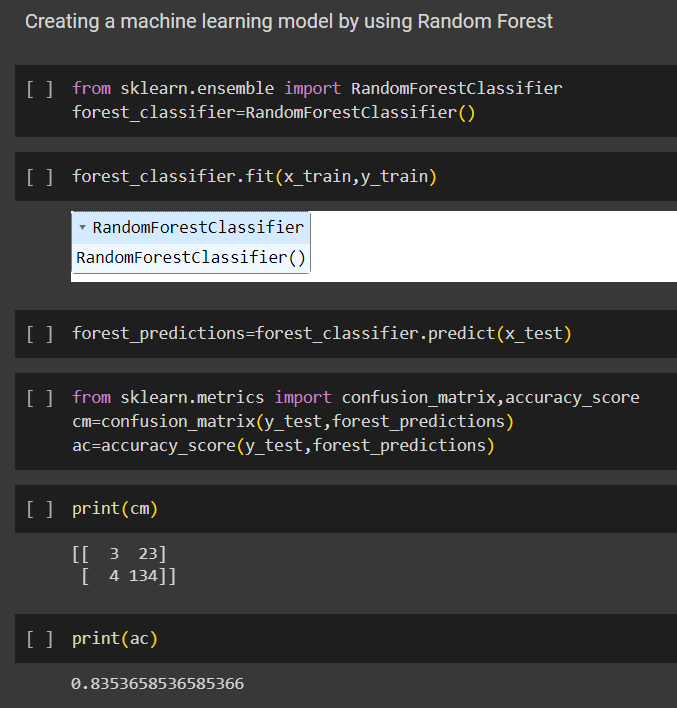


* Step-9: Importing DecisionTreeClassifier from sklearn.tree & then activating it by storing into the variable name tree. Then used tree\_classifier.fit() to train the model by providing train & test sets as x & y. And then predicted the trained model & the checked accuracy of the model using confusion\_matrix & accuracy\_score.



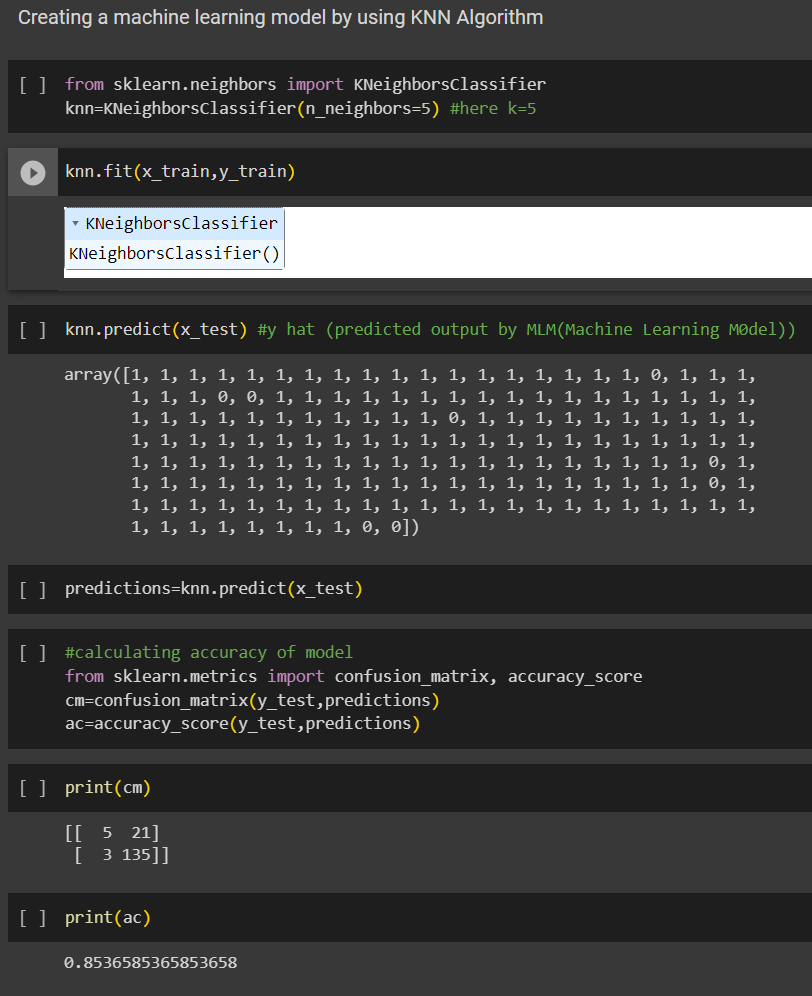
-In the above model we can see that the accuracy obtained is 81% which is quite good but we can also try using different models to see if we can get better accuracy than this or not.

* So I have also used RandomForestClassifier & KNN for obtaining better accuracy of the model.
* Model using RandomForestClassifier.



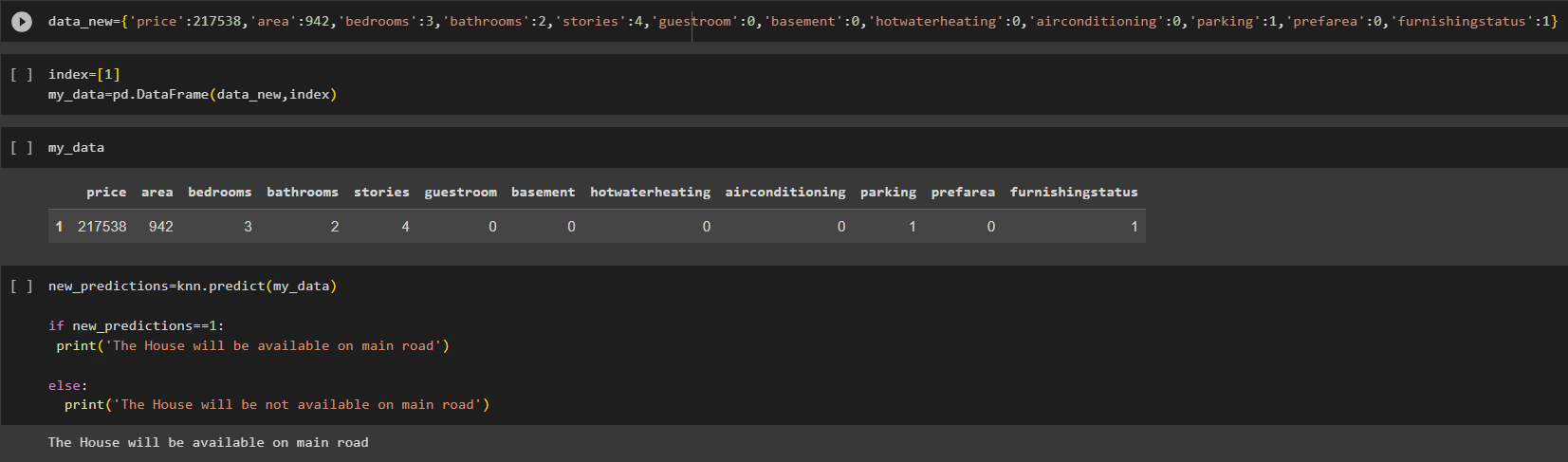
-In the above model we have obtained accuracy of 83% using RandomForestClassifier which is more accurate than DecisionTreeClassifier.

* Now model using KNN.



-Here In this model we have obtained accuracy of 85%. Which is more than the both of model’s i.e. RandomForestClassifier and DecisionTreeClassifier.

* Step-10: Predicting the new house data with KNN(high accuracy model). To identify that is it available in main road or not.



-Insight of this Prediction we are able to know that the house is available on main road(with 85% of accuracy).

* Conclusion: From this project we have analysed and visualized the cost of House price by using linear regression. And by help of different models of classification we can fetch the house is available on main road or not by considering the different types of facilities & furnishing status.

THANK YOU