Experiment No. 9

Title: Case study: Big data platform / analytics as business need)

Batch:A1 Roll No.:1714012 Experiment No.:9

Title: Case Study, Machine Learning Algorithm on Azure ML Studio

---

Resources needed:

Azure ML Studio

Experiment Title:
Car Price Prediction using Linear Regression

Describe the following points with respect to the business under consideration,

1. **Problem faced by the business:**
   The system is designed such that it helps to estimate the price of a Car based upon the different features present, so it helps the business to give the correct amount of pricing to their Car Models, so that the customers are attracted and also are satisfied with price range assigned based upon it's features. So, assigning Price range accordingly is challenging without the use of ML algorithms.

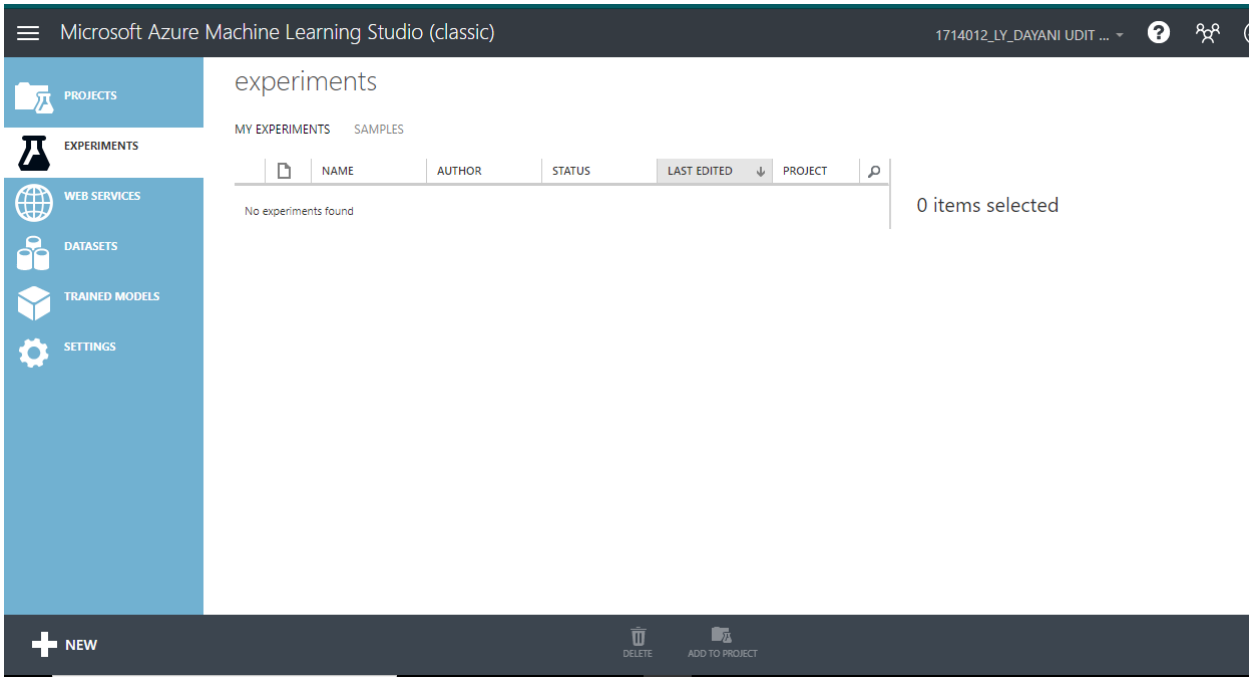
2. **Approach/ Methodology followed by the business :**
   The methodology which we followed was that we took the dataset of Automobile Price Data. Then we split the dataset into train and test. Then we used Linear Regression and trained the machine learning model with that dataset and in the end we tested that model.

3. **Skillsets, infrastructure and other impact on the business during implementation**
   Skillset: Cloud computing
   Infrastructure: Microsoft Azure Cloud services
   There were no as such impacts on the business, as any of the business service lines were not being used.

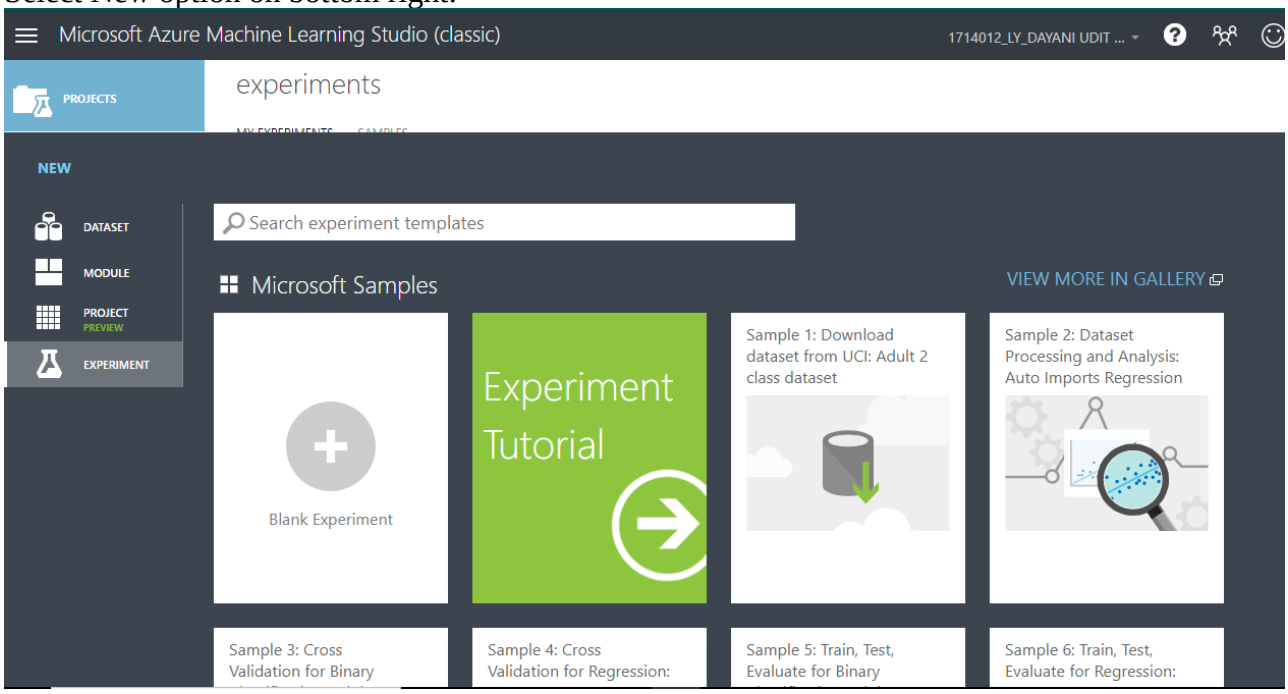
4. **Similar approaches followed by other businesses:**
   No Other business is focusing on using ML in order to estimate their car prices, and thus are not able to properly evaluate the right price range slot to ensure the customers are attracted to the Company's Cars.

Steps:

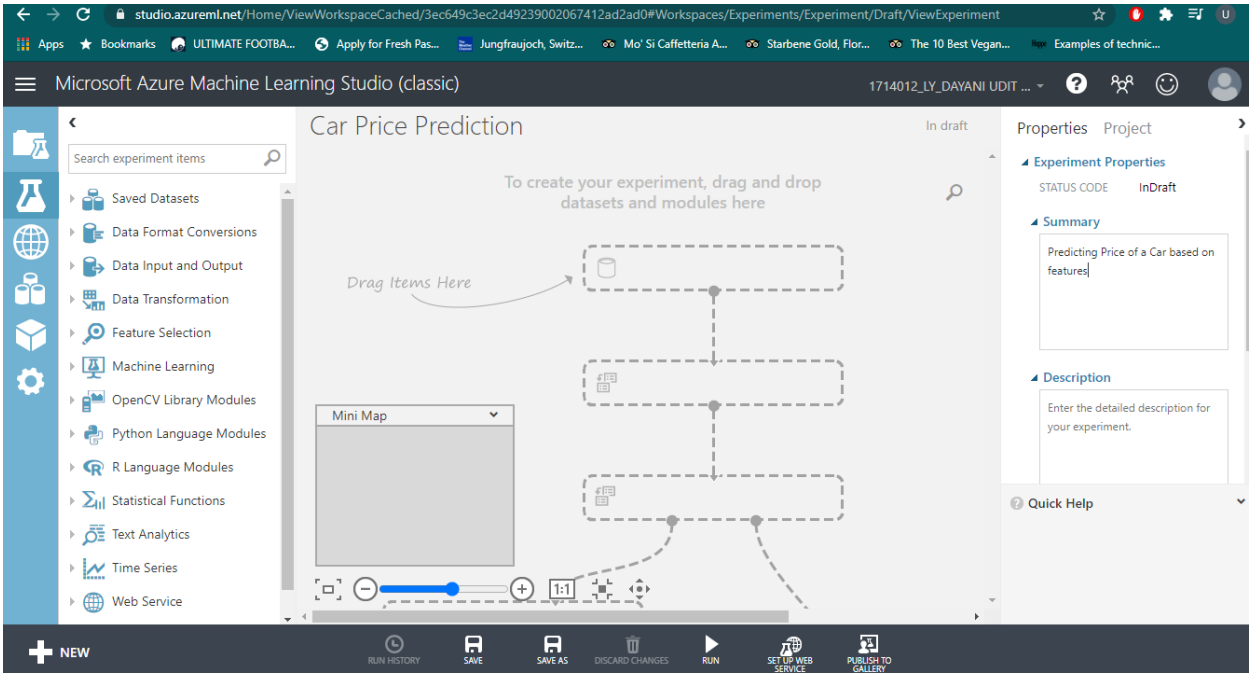
Sign-in using Microsoft account on studio.azureml.net



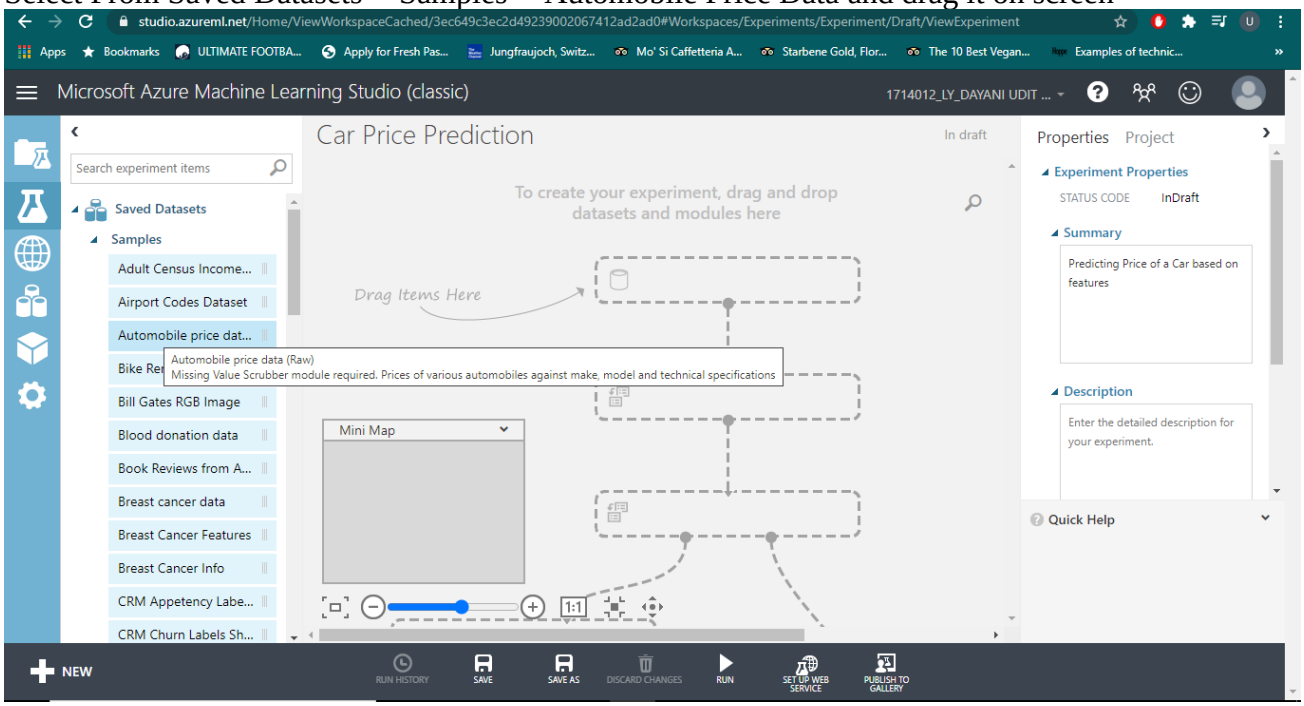
Select New option on bottom right:

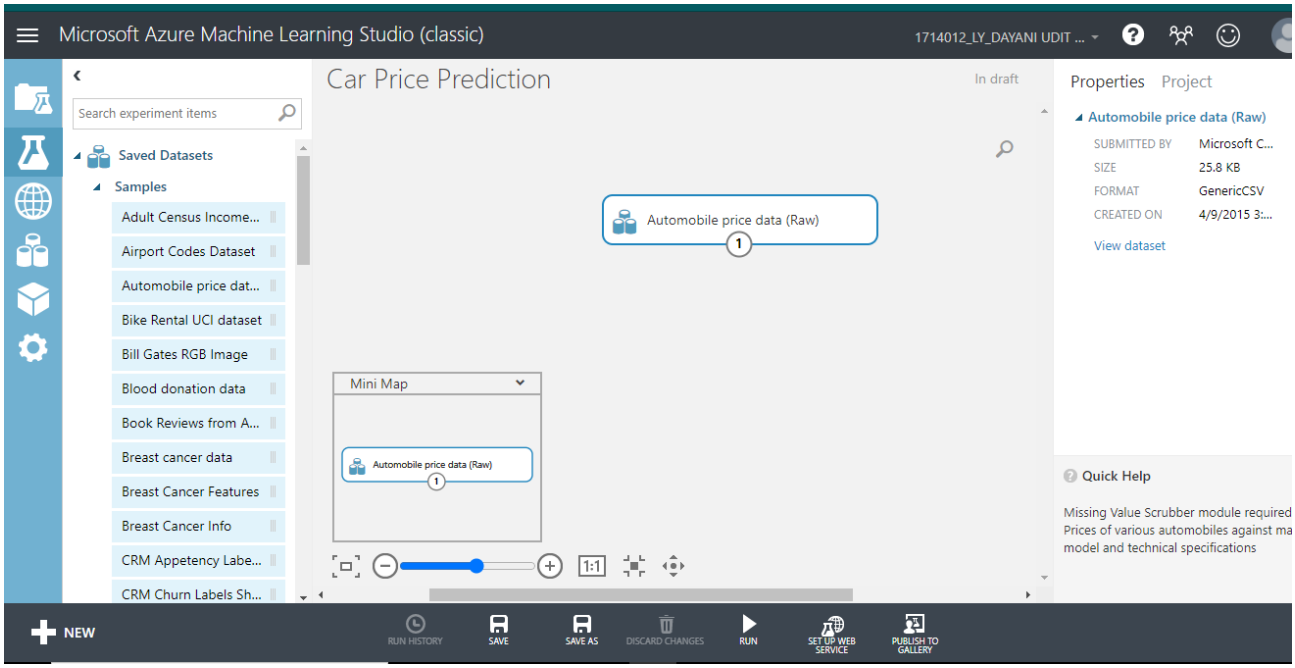


Click on Blank experiment and write name and summary of experiment

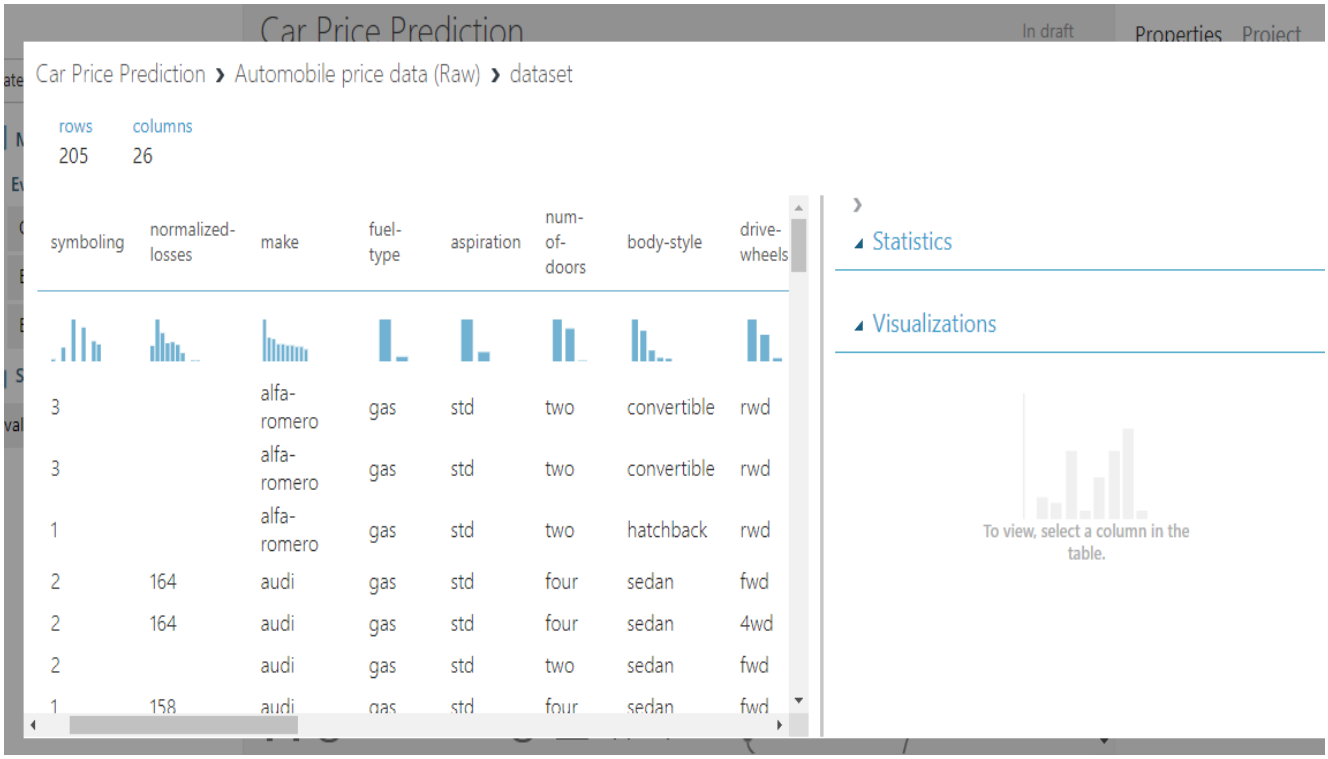


Select From Saved Datasets-> Samples-> Automobile Price Data and drag it on screen

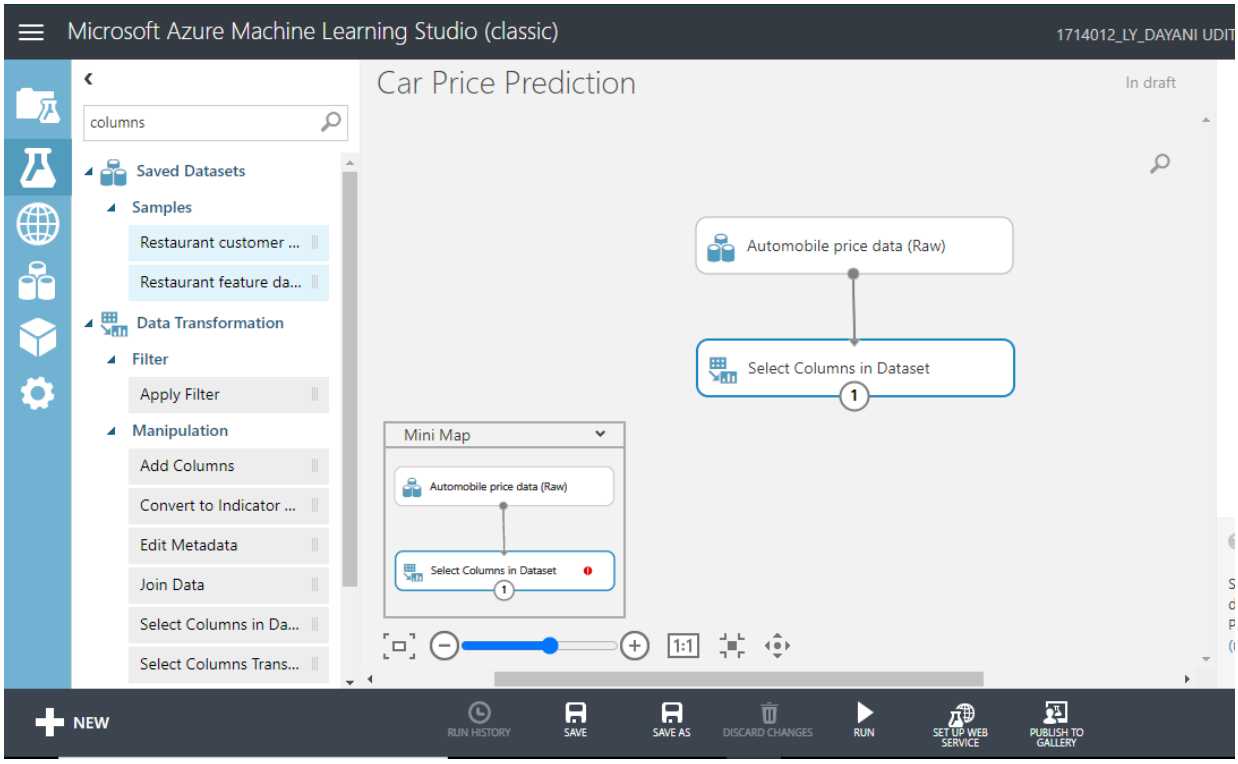




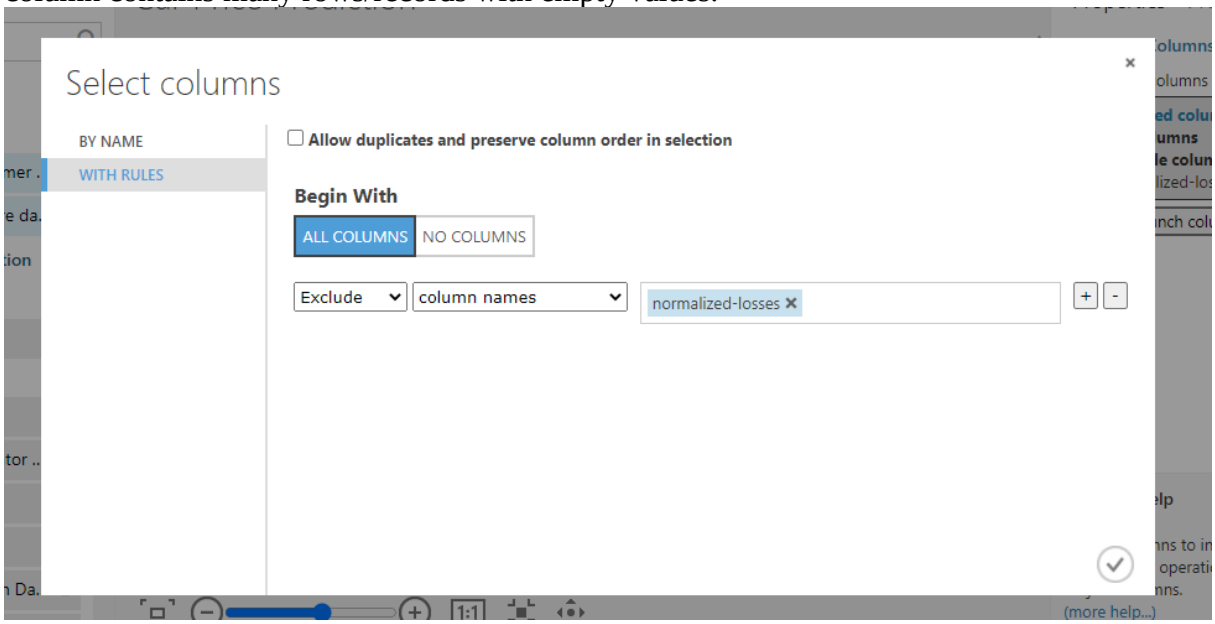
Below is the visualization of dataset selected



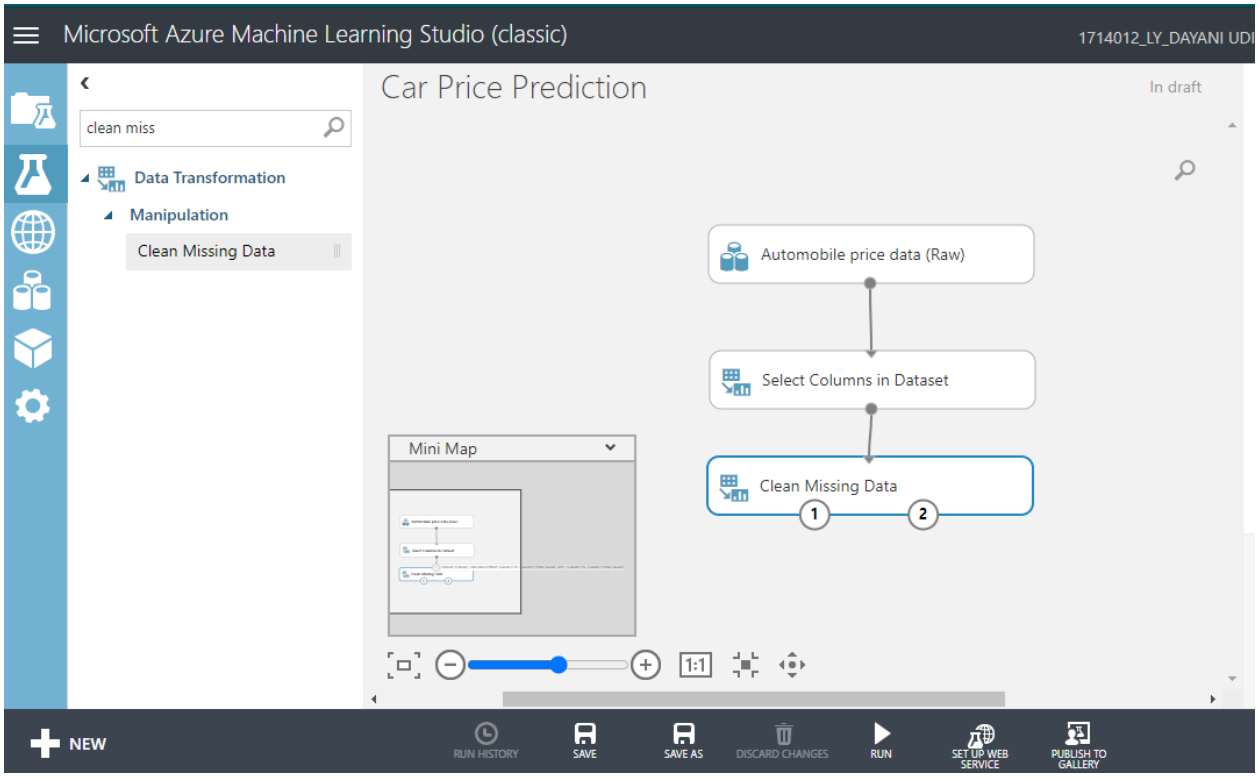
Now, search ‘Select columns in dataset’ from items and drag it



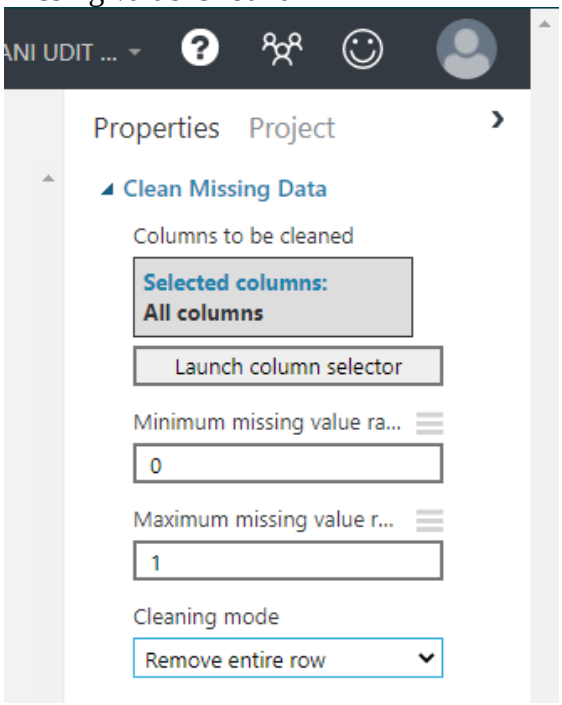
Now, click on launch column selector-> with rules->exclude column normalized-losses as that column contains many rows/records with empty values.



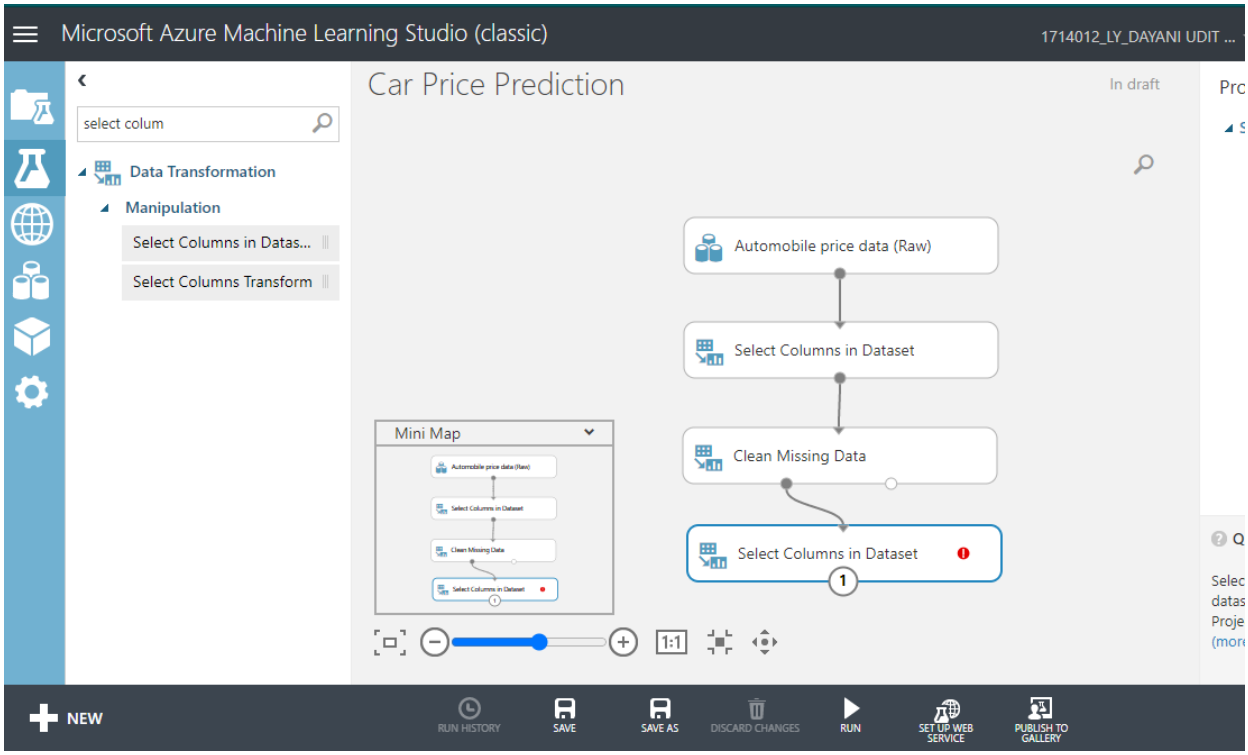
Search and select ‘Clean Missing Data’ from items list



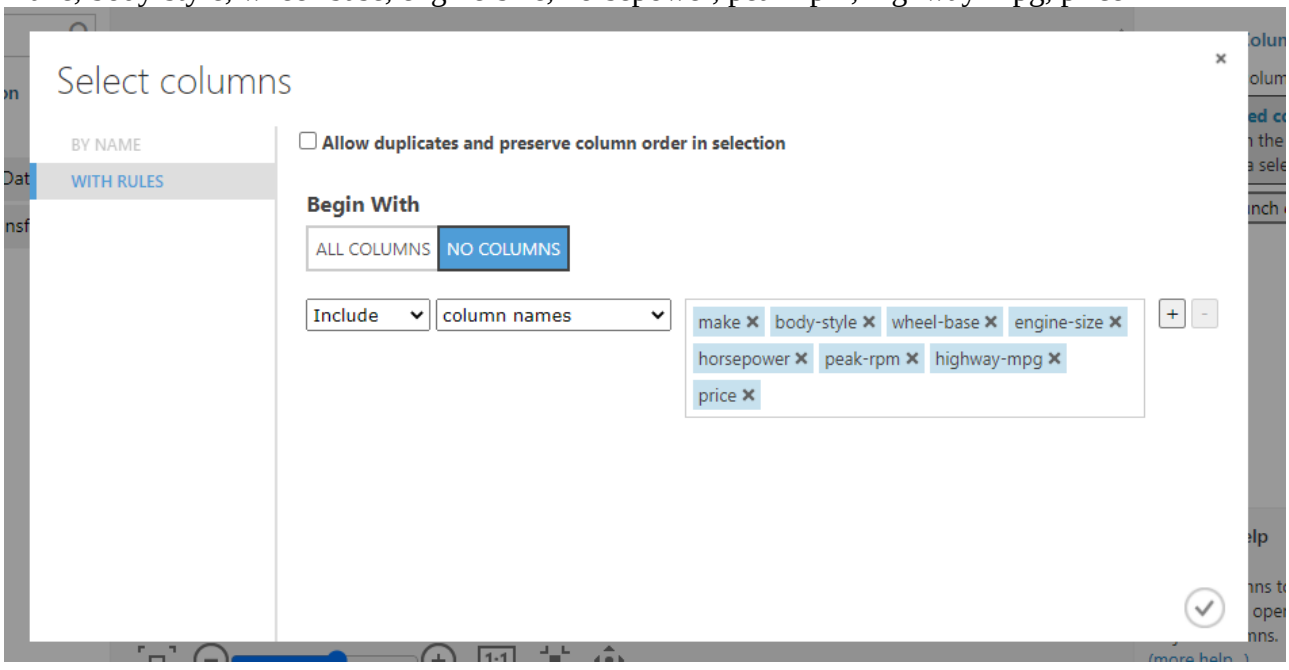
Now, select cleaning mode -> Remove entire row as it will remove the entire row wherever missing value is found



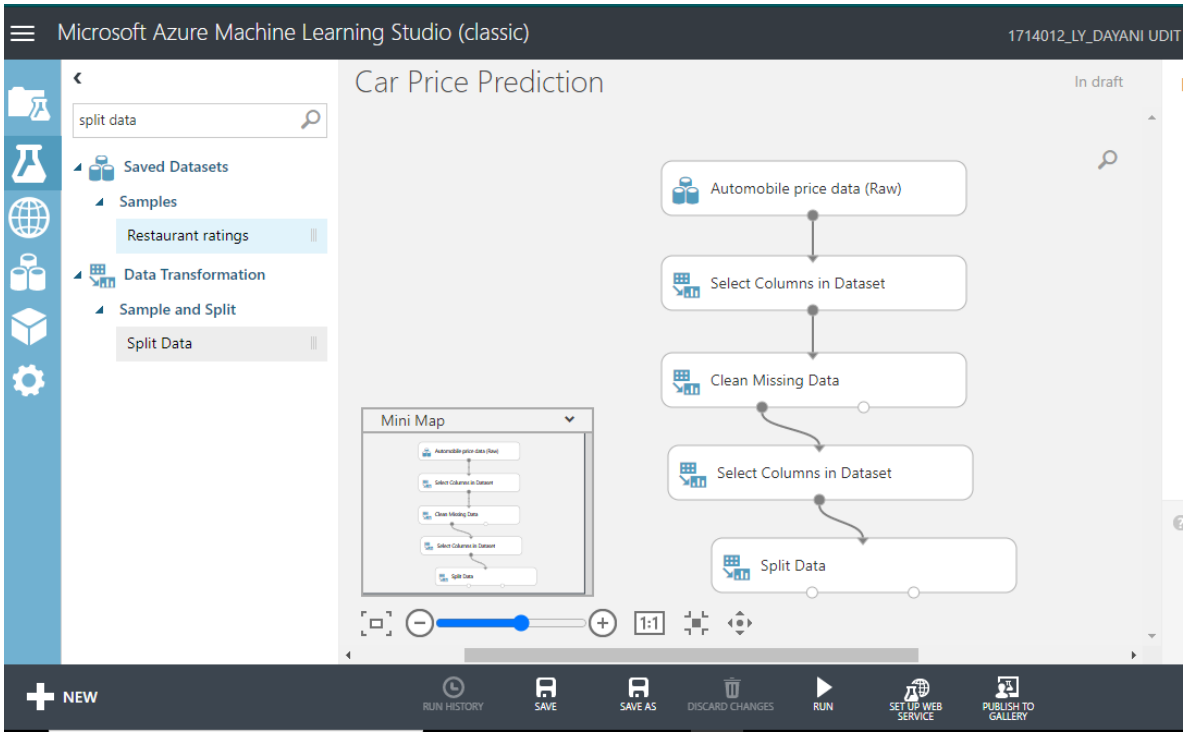
Again choose ‘select columns in dataset’



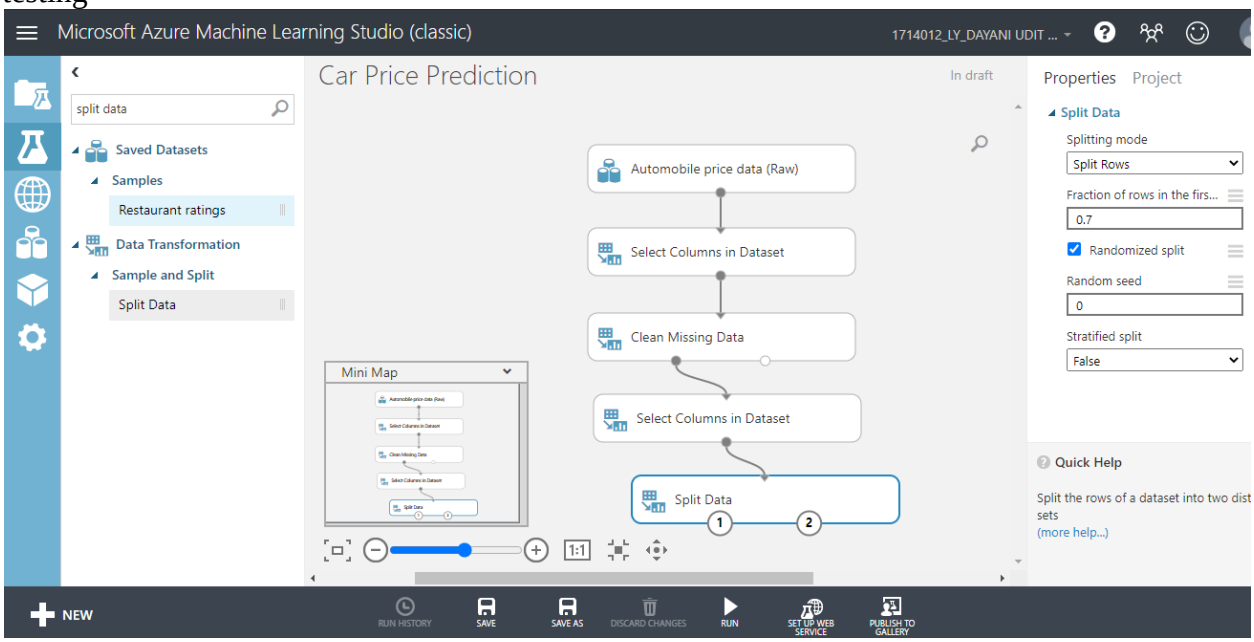
Now, launch column selector and include all the columns based on which prediction is to be done: make, body-style, wheel-base, engine-size, horsepower, peak-rpm, highway-mpg, price



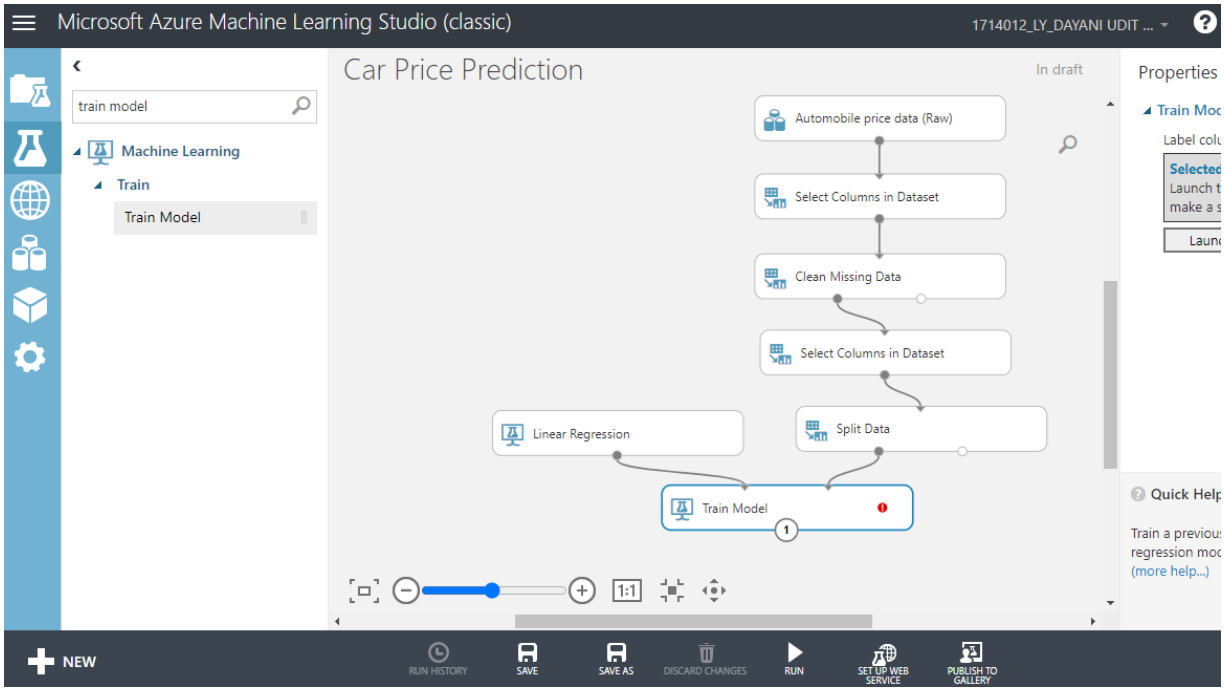
Now, select ‘split data’ from list and drag it



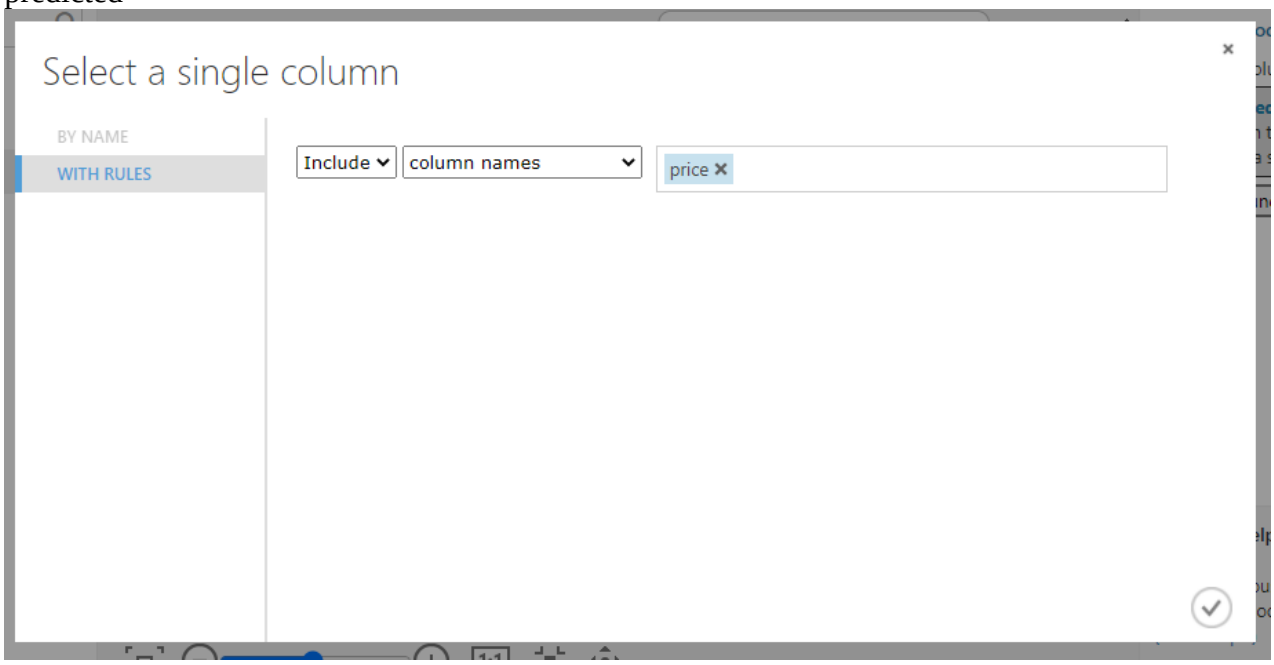
For Split data, enter the fraction of data which is needed for training while rest will be used for testing



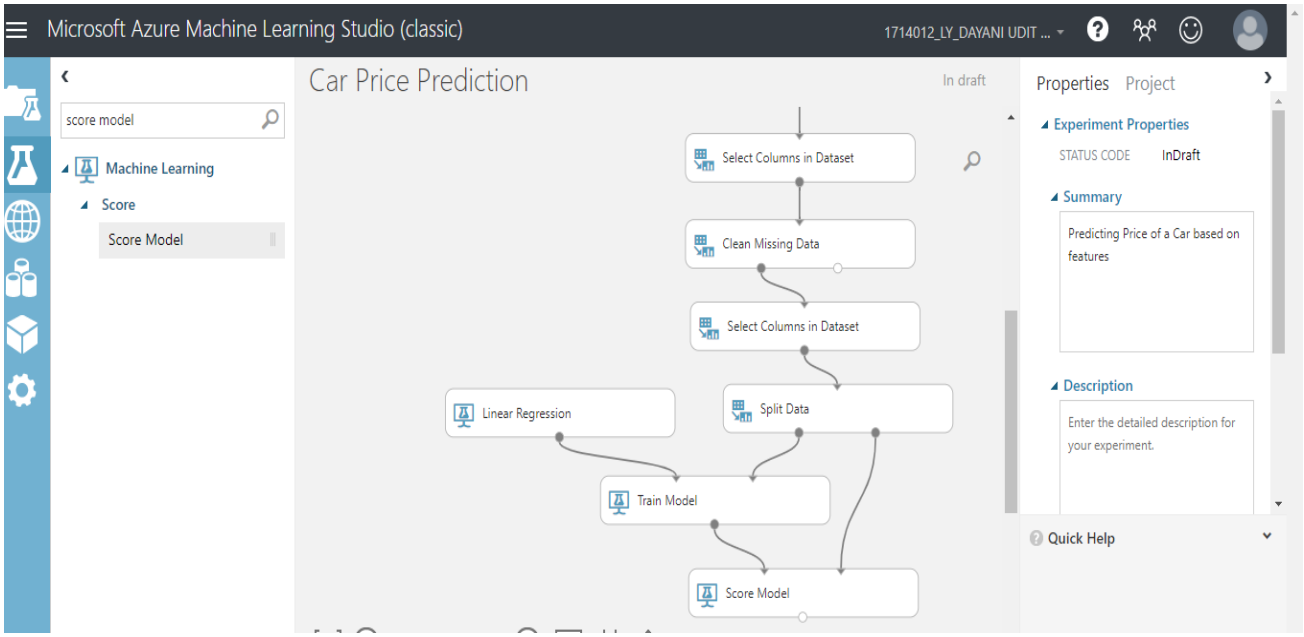
Now, Select ‘Linear Regression’ as the algorithm to be used and ‘Train Model’ from list



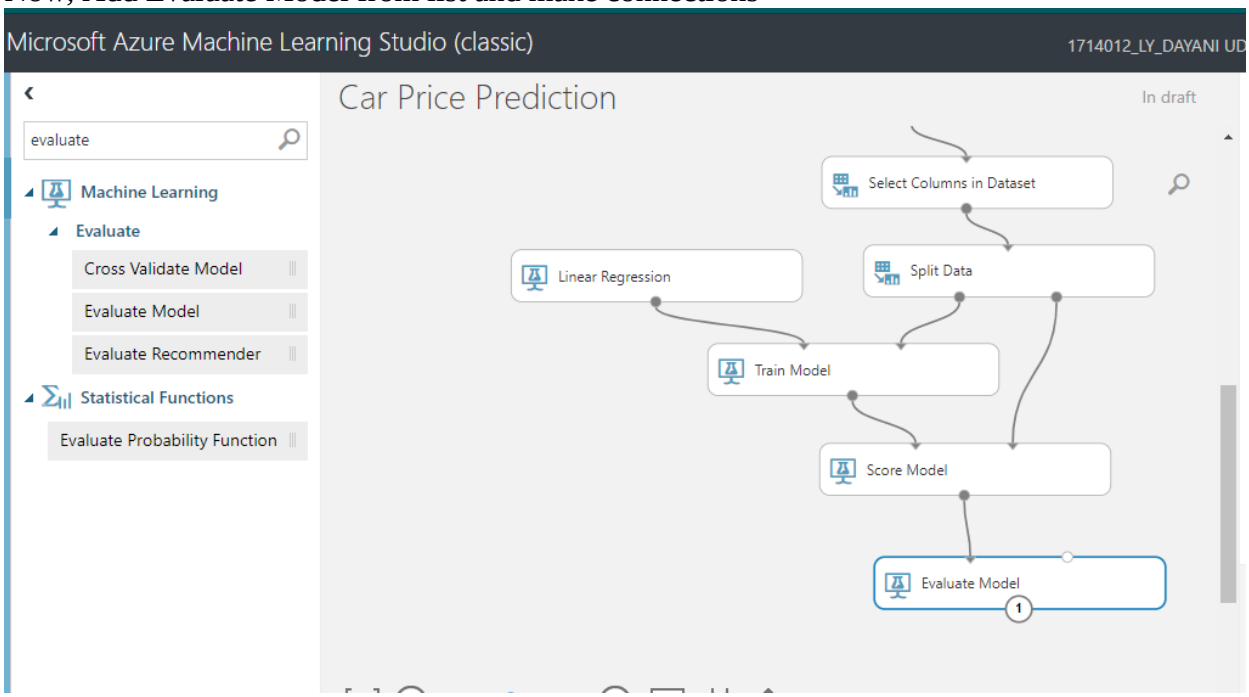
For training model, click on launch column selector, include price column as Price is what is to be predicted



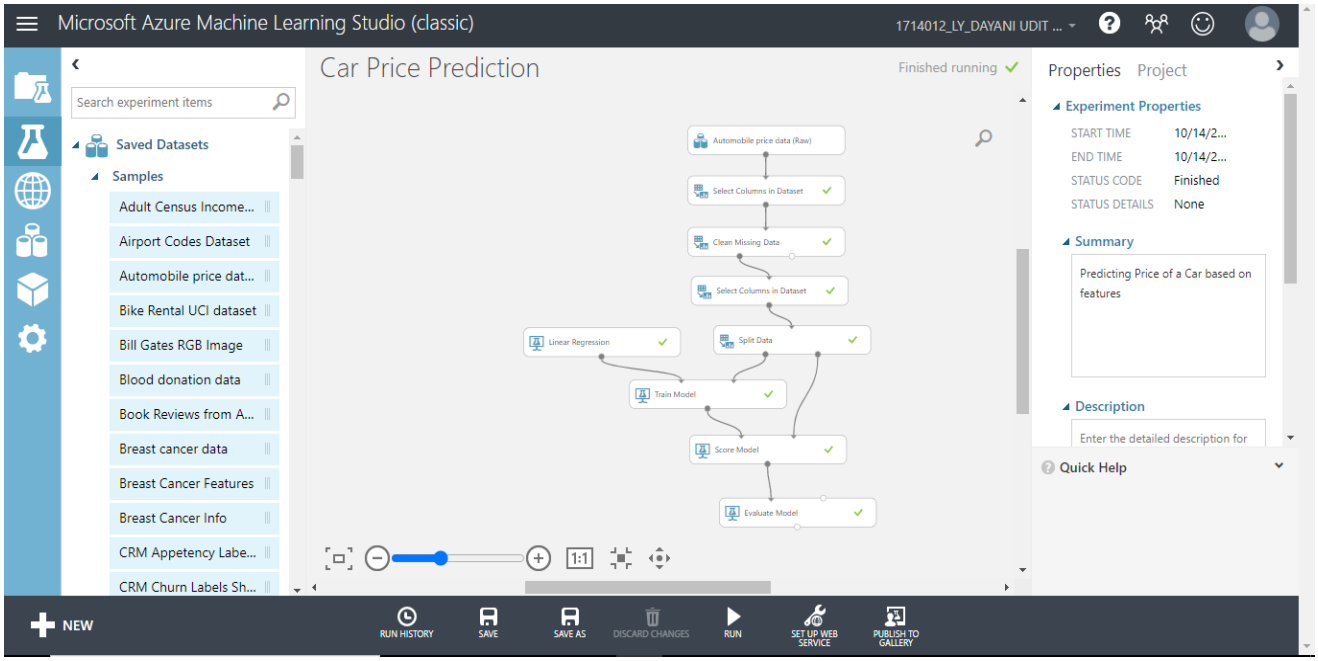
Add Score Model from list drag it and make connections



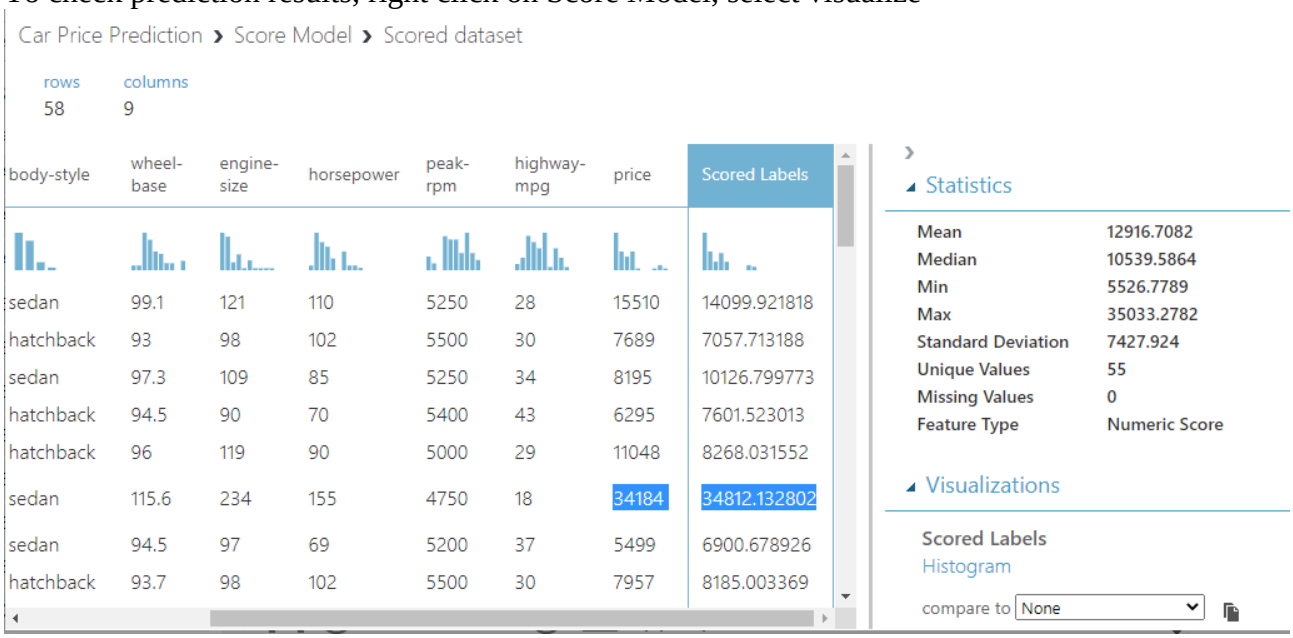
Now, Add Evaluate Model from list and make connections



Now, Click on Run



To check prediction results, right click on Score Model, select visualize



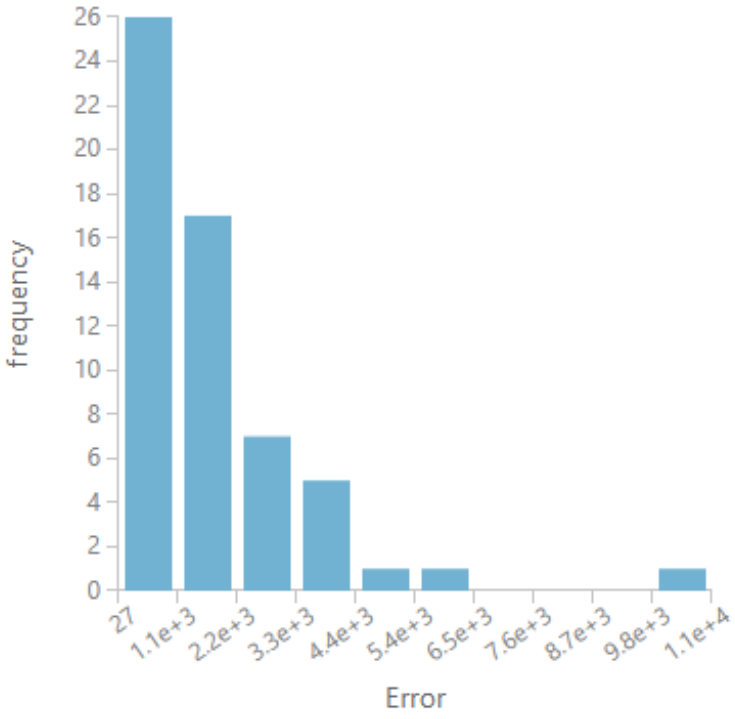
To check Evaluation results, right click on Evaluation Model, select visualize

Car Price Prediction › Evaluate Model › Evaluation results

Metrics

Mean Absolute Error	1634.317008
Root Mean Squared Error	2385.612307
Relative Absolute Error	0.271025
Relative Squared Error	0.083135
Coefficient of Determination	0.916865

Error Histogram



Questions:

Discuss the tangible and intangible benefits the business has observed after the implementation.

- **The right Price range for Cars.**
- **The trained model can used for any type of Car.**
- **This model can be used now anywhere in the enterprise.**

Outcomes:

Realize adequate perspectives of big data analytics in various applications.

Conclusion: (Conclusion to be based on the objectives and outcomes achieved)

I was able to perform an experiment on Azure ML Studio using Machine Learning algorithm(linear regression).
Experiment Link: https://gallery.azure.ai/Experiment/Car-Price-Prediction-BDA-exp-9-1714012
Github Link: https://github.com/uditdayani/BDA

Grade: AA / AB / BB / BC / CC / CD /DD

Signature of faculty in-charge with date

Reference

Books/ Journals/ Websites: