

19K41A0518

AZURE ML Classic Studio

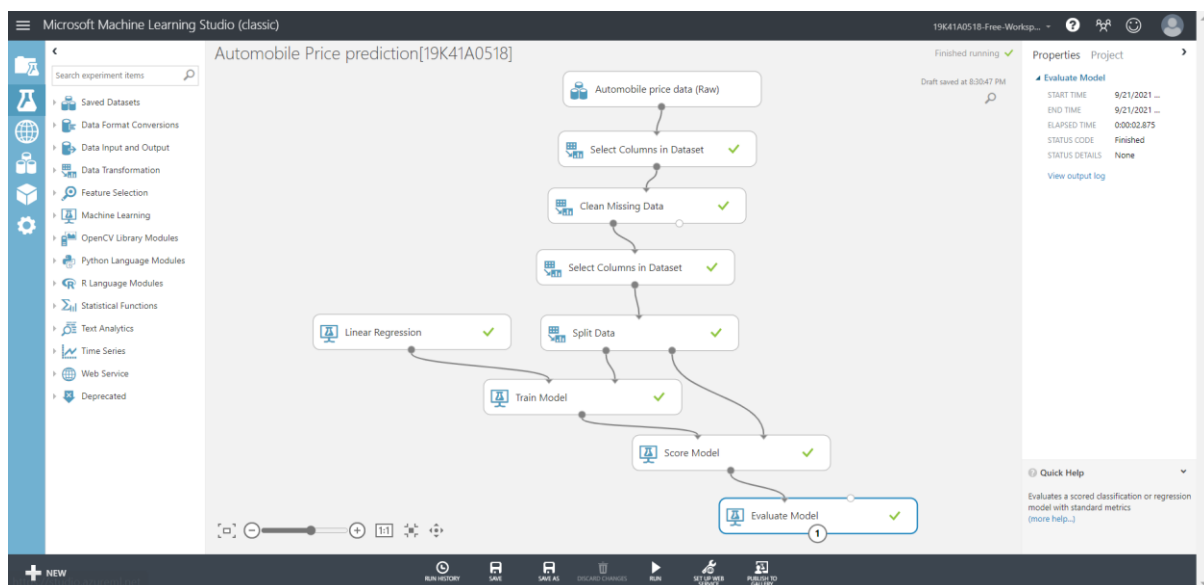
Predicting Automobile prices using Regression Model in Azure ML Classic Studio.

In this project,I created a model that automatically predicts a car's price based on some feautres such as make,model,horsepower, and more.

Project Workflow:

- 1.Load the data.
- 2.Explore data(missing values).
3. Preprocess the data.
- 4.Choose the model(Linear Regression).
5. Split the data◇Training and Testing.
- 6.Train the model.
7. Score the model.
- 8.Evaluate the model based on results.

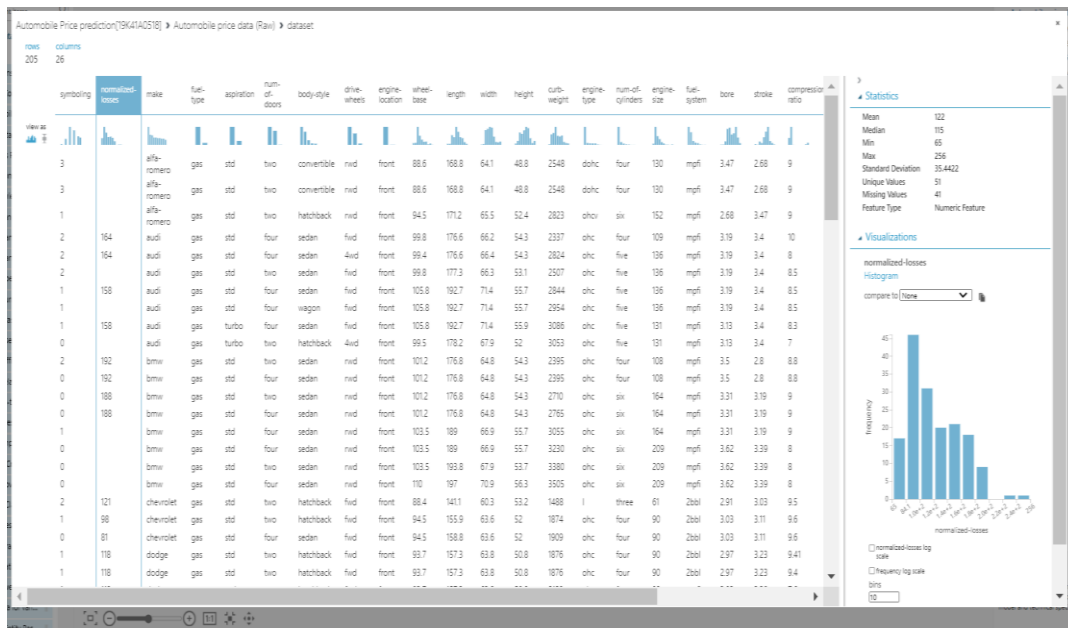
Workflow



Project Workflow

- **Import Data:**

Importing the RAW dataset which is in CSV format. The dataset is pre-available in the Azure ML Classic Studio.



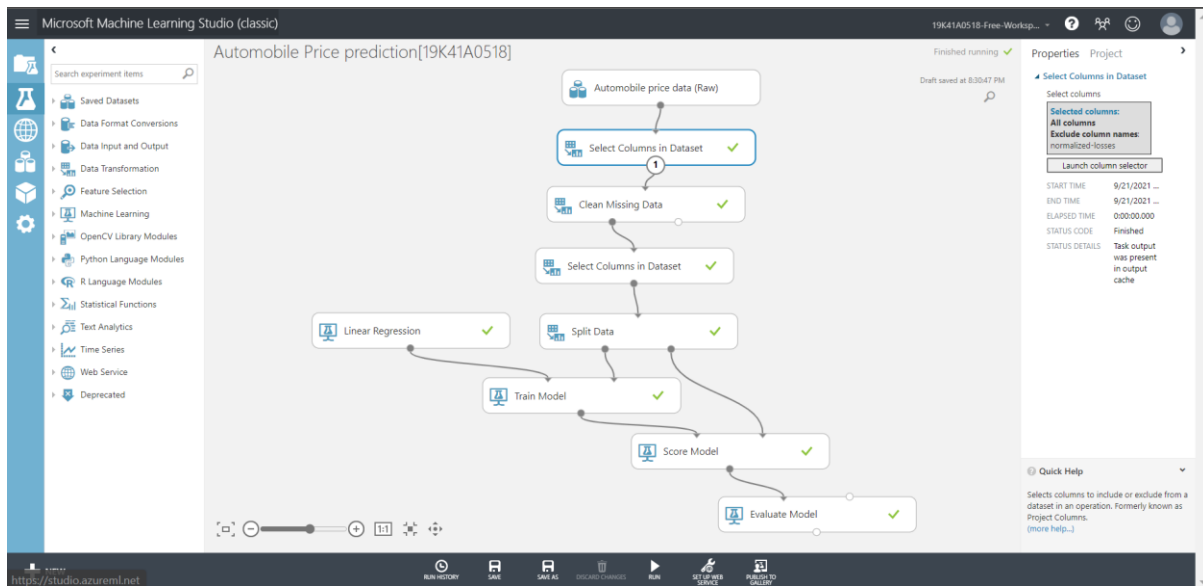
Automobile Price RAW dataset (CSV format)

- **Explore Data**

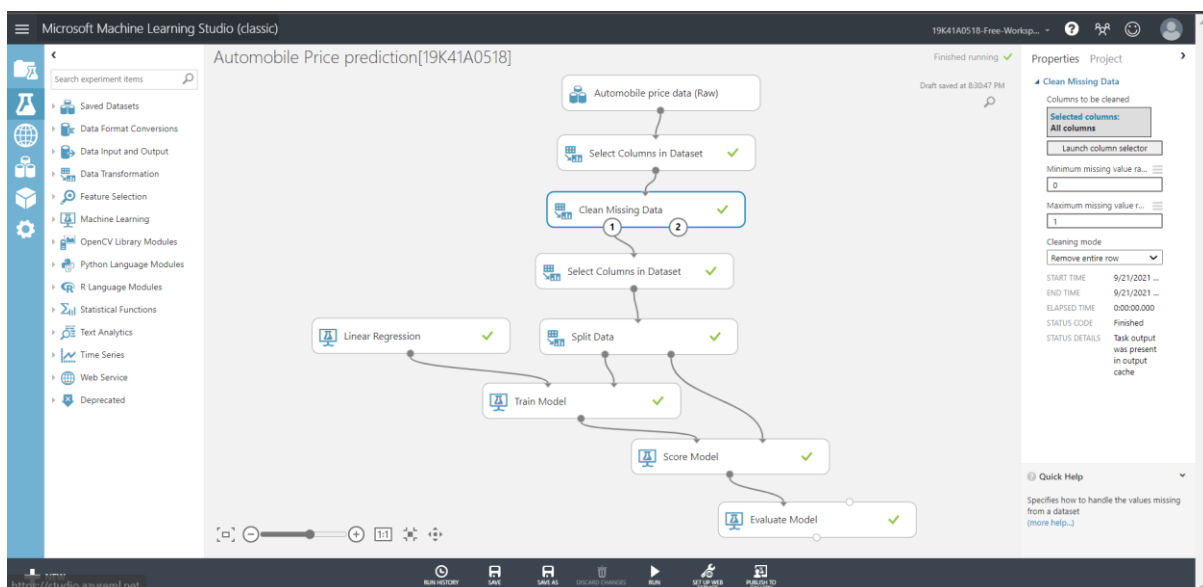
This basically includes data visualization to search for any missing values in the Dataset.

if any missing values are found, then they needs to be cleaned.

selecting the required columns and clean the data using the Clean Missing Value module (Just Drag n' Drop)



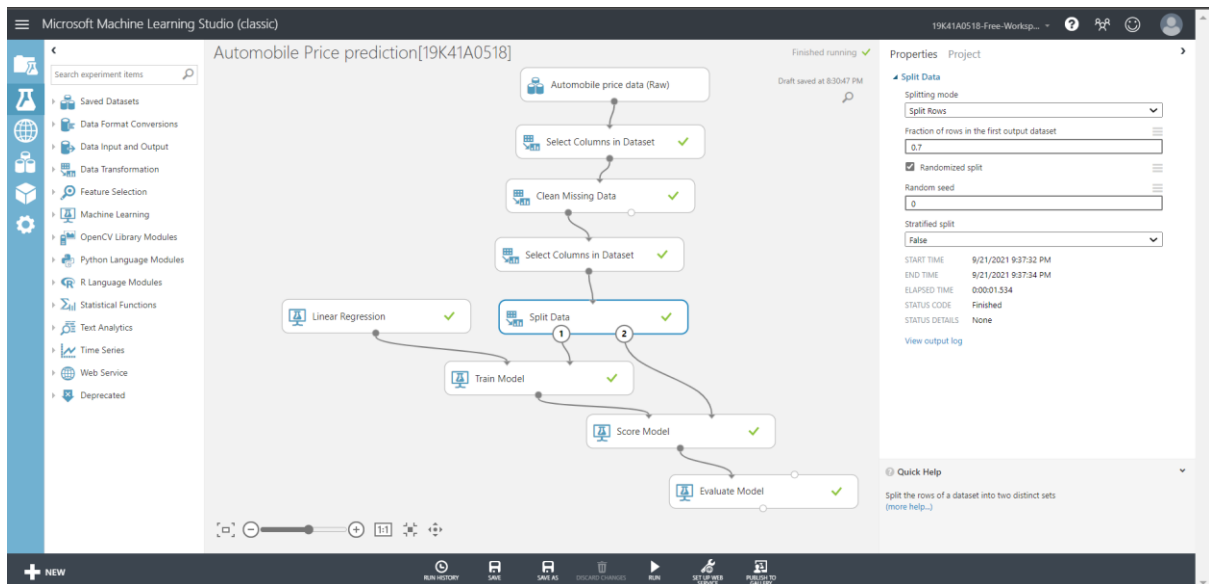
as the normalized loss has 41 missing values in the Dataset, those missing values are to be cleaned.



Cleaning data

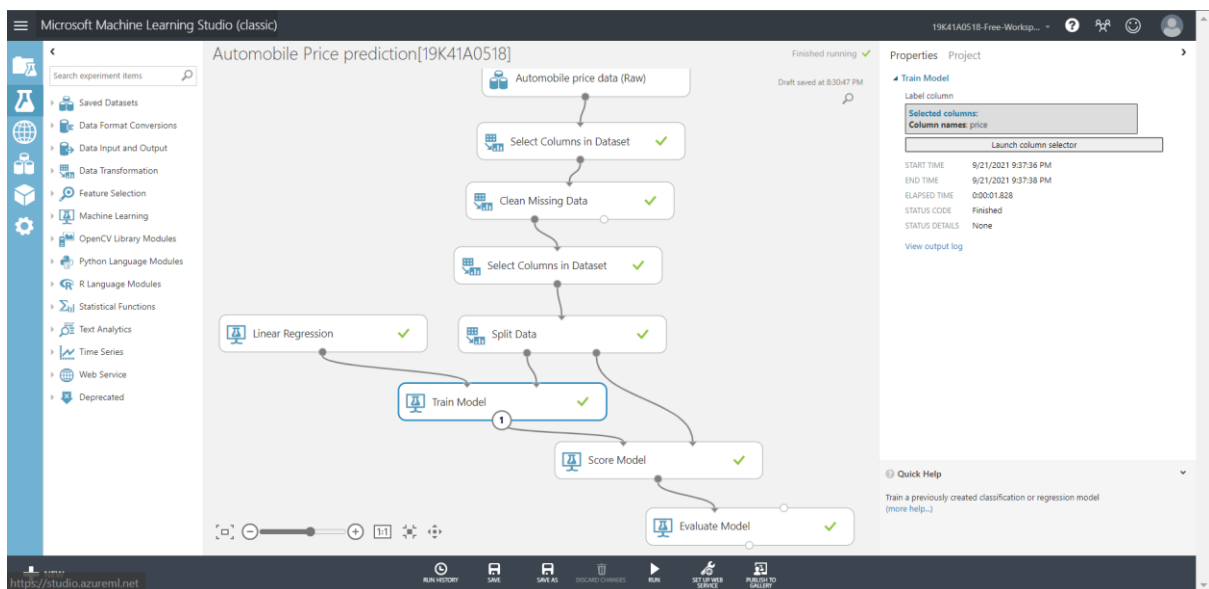
• Split Data

Use the Split Data module to randomly divide the input data so that the training dataset contains 70% of the original data and the testing dataset contains 30% of the original data.



Data Splitting

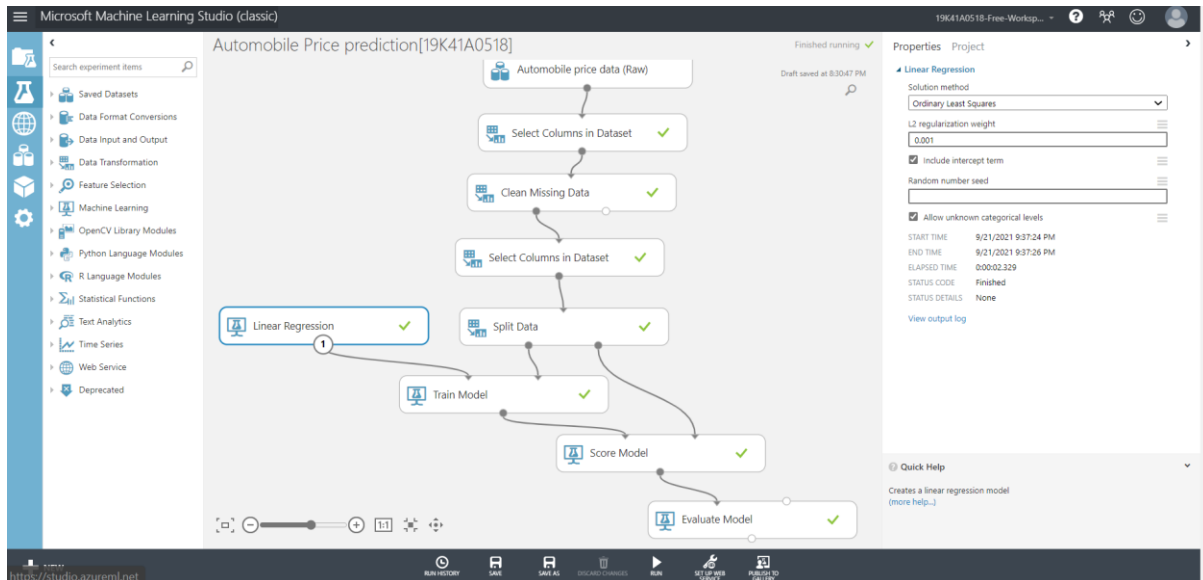
• Model Training and Algorithm



Model Training

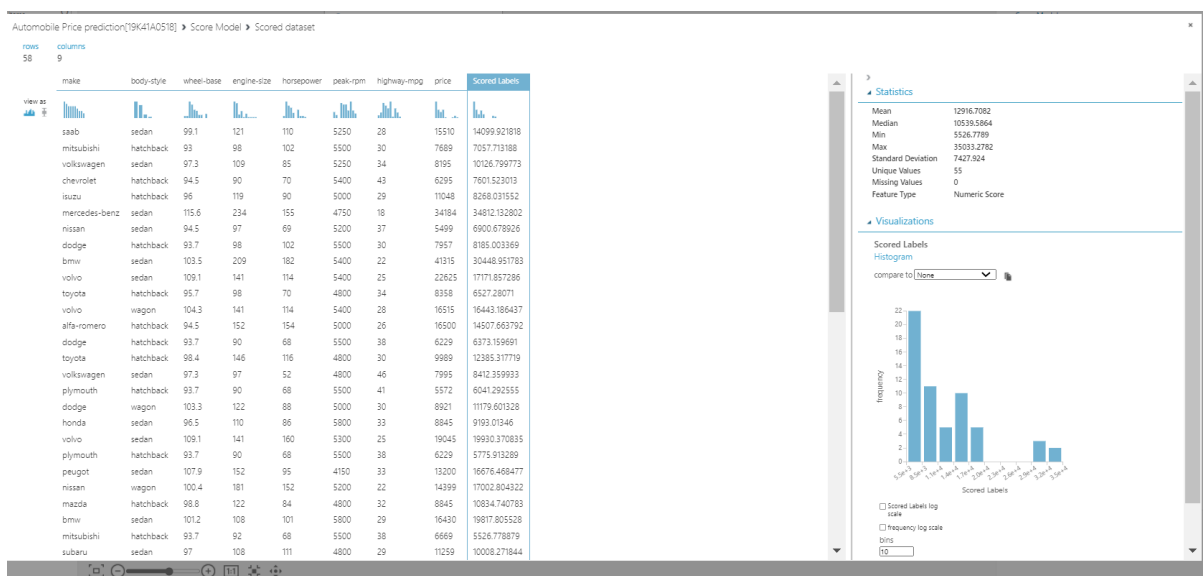
- using Linear regression to train the model

Since the goal of this sample is to predict automobile prices, and because the label column (price) is continuous data, a regression model can be a good choice. We use Linear Regression for this pipeline.



Linear Regression

- Score Model and Evaluate Model



Score labels

● Evaluation Results

Automobile Price prediction[19K41A0518] ➤ 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

