19K41A0565

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 feautures 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 preavailable in the Azure ML Classic Studio.

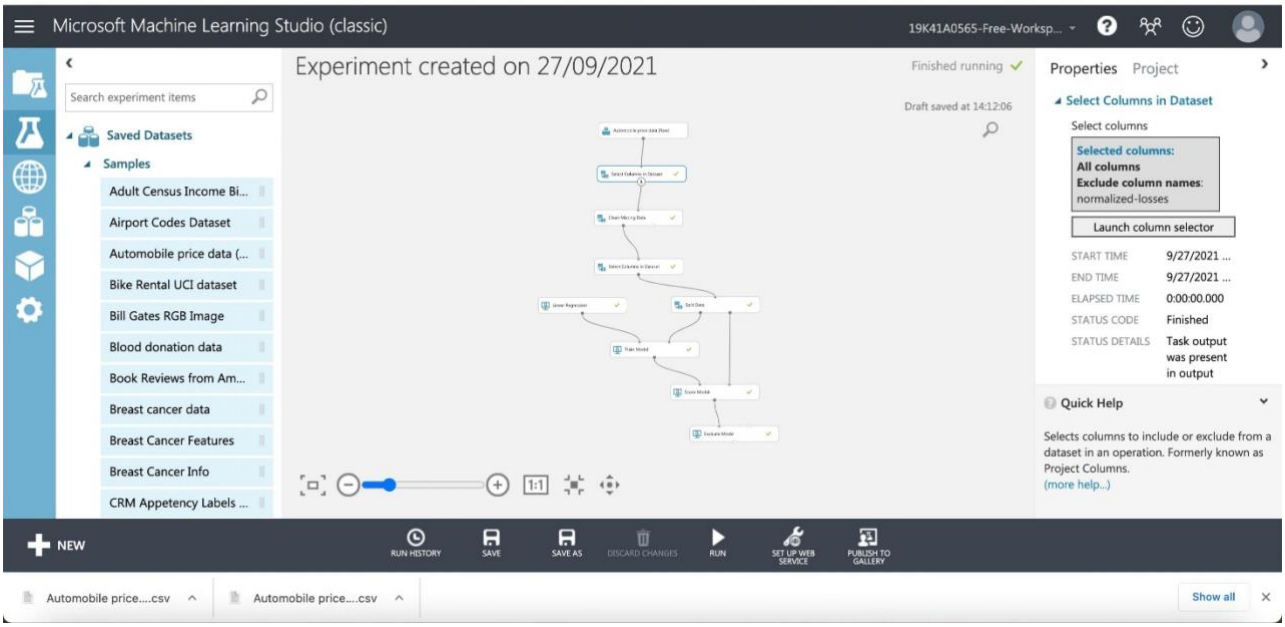
Automobile Price RAW dataset (CSV format)

- **Explore Data**

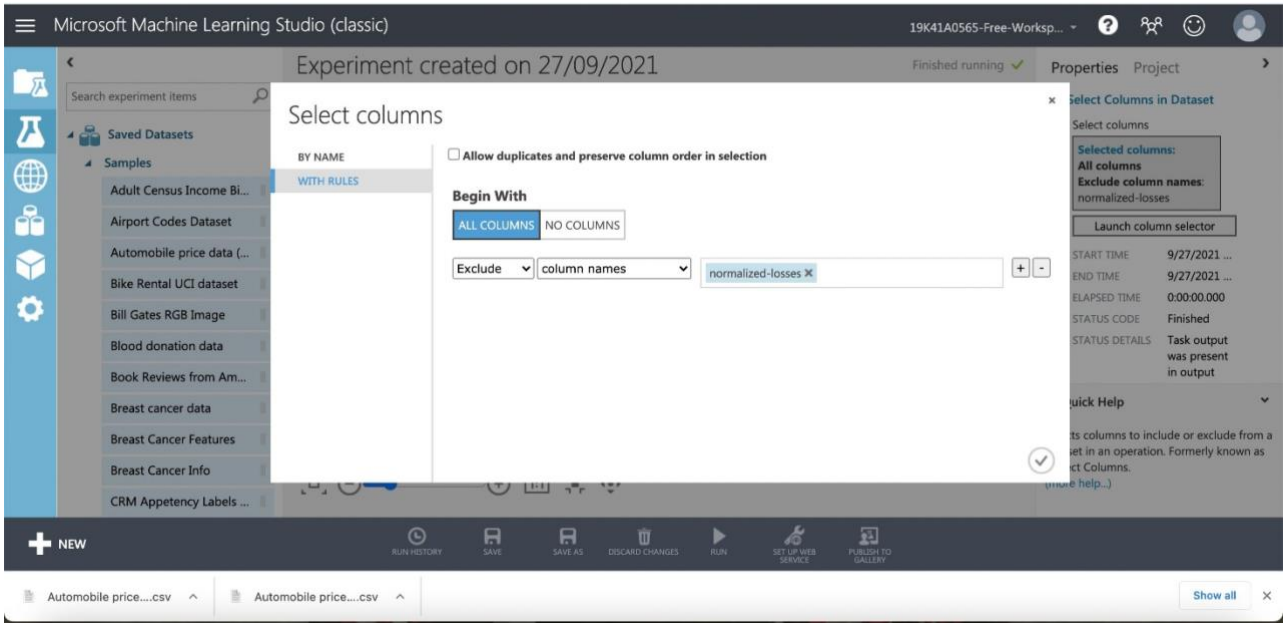
This basically includes data visualization to search for any missing values in the Dataset.

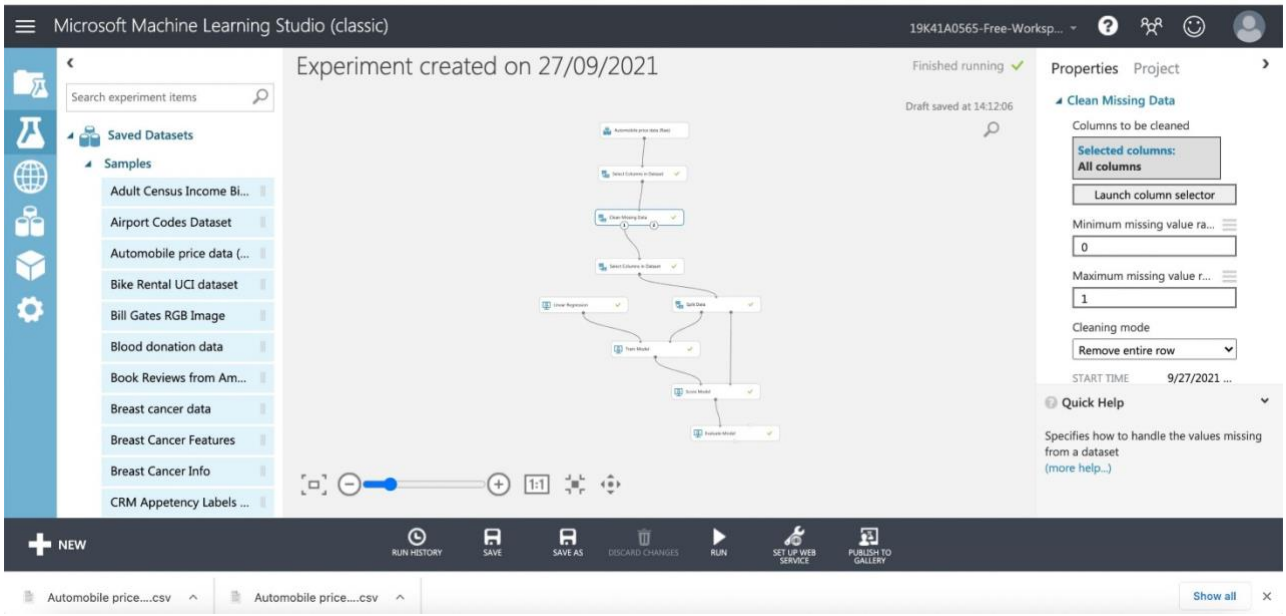
if any missing values are found, then they needs to be cleaned.

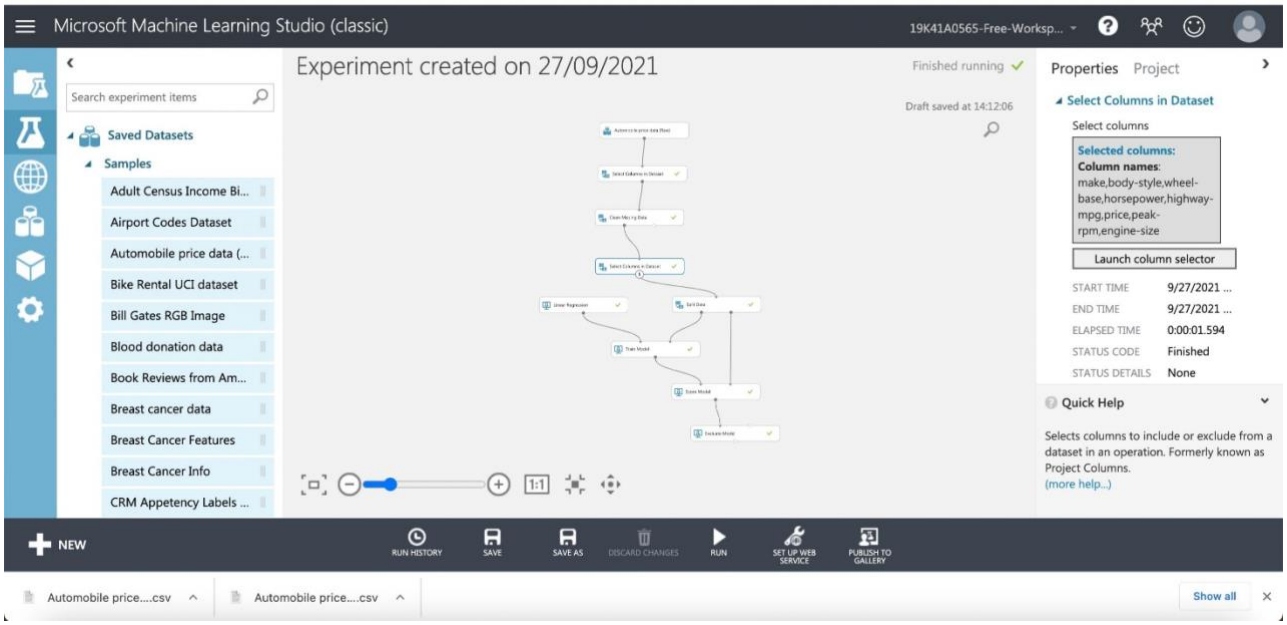
selecting the required columns and clean the data using the Clean Missing Value module (Just Drag n' Drop)

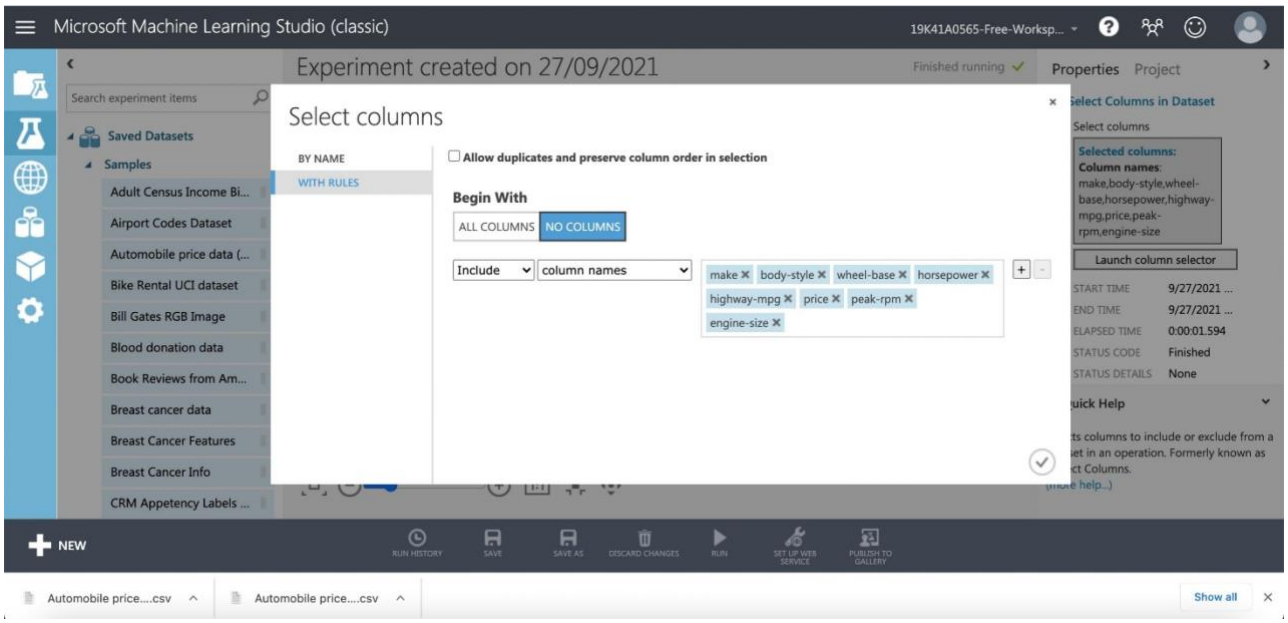


as the normalized loss has 41 missing values in the Dataset, those missing values are to be cleaned.


Microsoft Machine Learning Studio (classic)
19K41A0565-Free-Worksp...
Experiment created on 27/09/2021
Finished running
Properties Project
Search experiment items
Saved Datasets
Samples
Adult Census Income Bi...
Airport Codes Dataset
Automobile price data (...
Bike Rental UCI dataset
Bill Gates RGB Image
Blood donation data
Book Reviews from Am...
Breast cancer data
Breast Cancer Features
Breast Cancer Info
CRM Appetency Labels ...
Select columns
BY NAME
WITH RULES
Allow duplicates and preserve column order in selection
Begin With
ALL COLUMNS
NO COLUMNS
Exclude
column names
normalized-losses
Select Columns in Dataset
Select columns
Selected columns:
All columns
Exclude column names:
normalized-losses
Launch column selector
START TIME 9/27/2021 ...
END TIME 9/27/2021 ...
ELAPSED TIME 0:00:00.000
STATUS CODE Finished
STATUS DETAILS Task output was present in output
Quick Help
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY
Automobile price....csv
Automobile price....csv
Show all

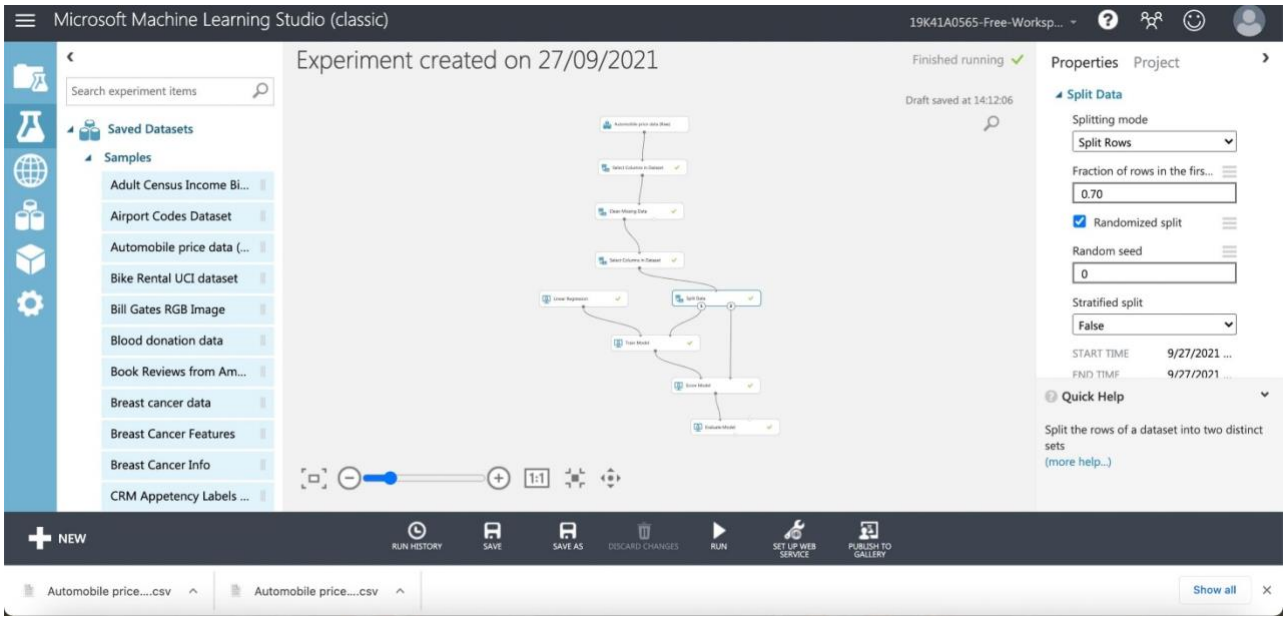

Microsoft Machine Learning Studio (classic)
19K41A0565-Free-Worksp...
Experiment created on 27/09/2021
Finished running
Draft saved at 14:12:06
Properties Project
Search experiment items
Saved Datasets
Samples
Adult Census Income Bi...
Airport Codes Dataset
Automobile price data (...
Bike Rental UCI dataset
Bill Gates RGB Image
Blood donation data
Book Reviews from Am...
Breast cancer data
Breast Cancer Features
Breast Cancer Info
CRM Appetency Labels ...
Clean Missing Data
Columns to be cleaned
Selected columns:
All columns
Launch column selector
Minimum missing value ra...
0
Maximum missing value r...
1
Cleaning mode
Remove entire row
START TIME 9/27/2021 ...
Quick Help
Specifies how to handle the values missing from a dataset
(more help...)
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY
Automobile price....csv
Automobile price....csv
Show all





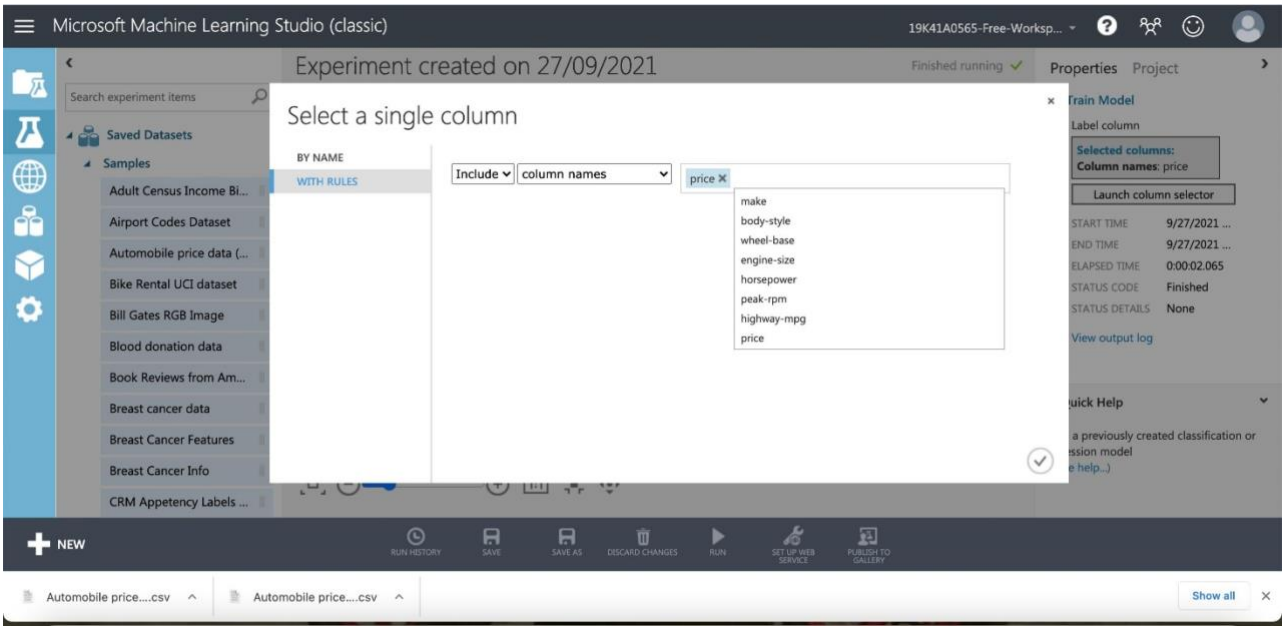
- Split Data

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.



Splitting data

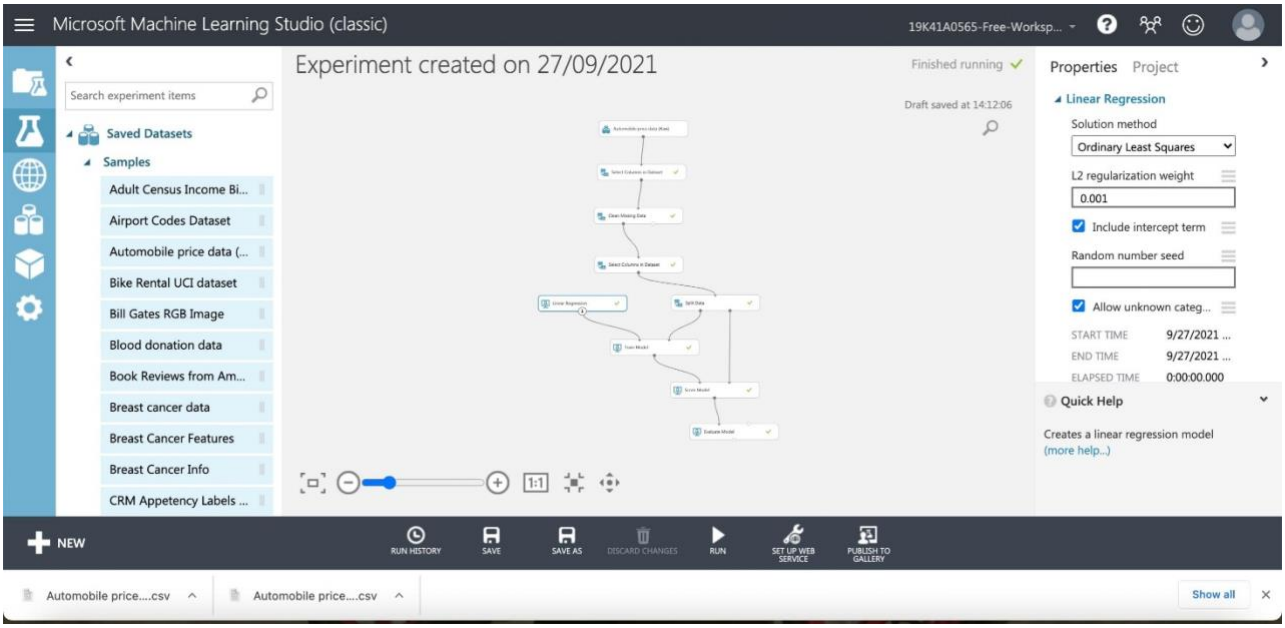
- Model Training and Algorithm



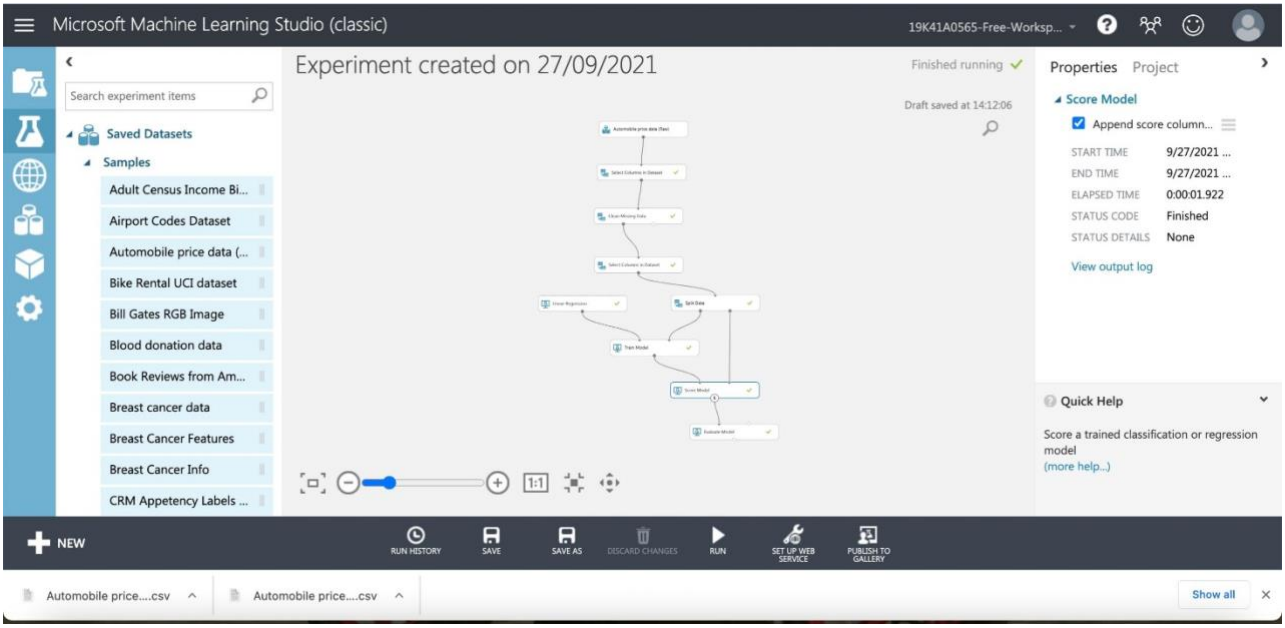
Model Training

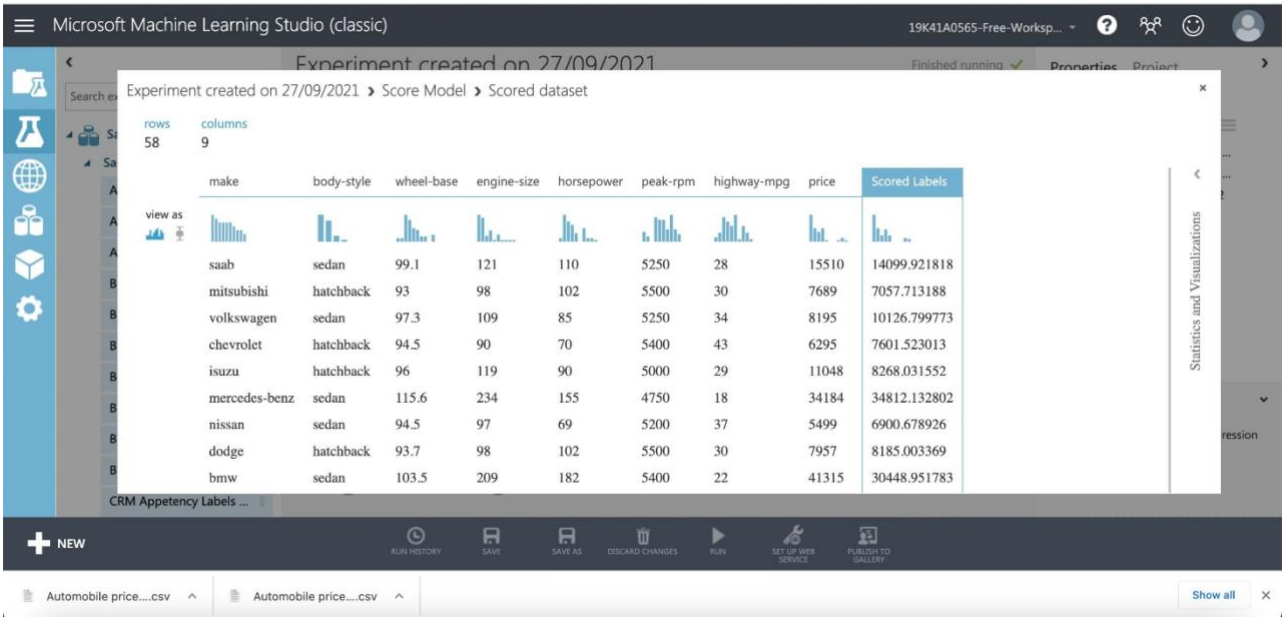
- using Linear regression to train the model

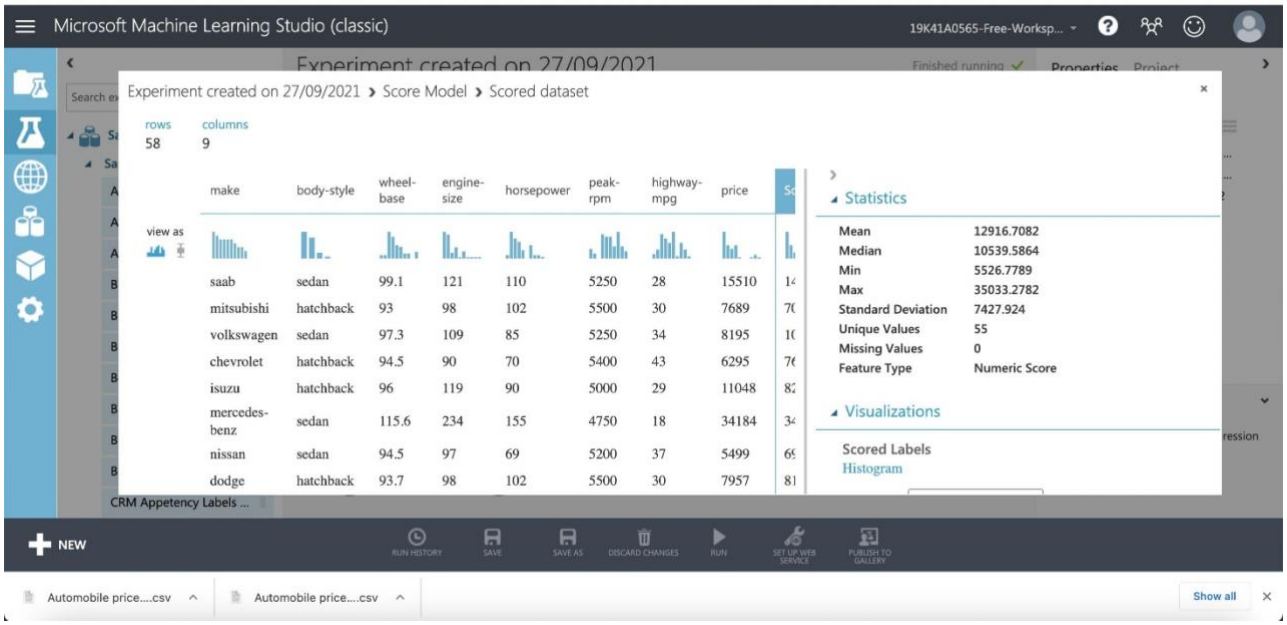
Since the goal of this sample is to predict automobile prices, and because the label column (price) is continuous data, a regression model can be a good choice. We use Linear Regression for this pipeline.



• Score Model and Evaluate Model

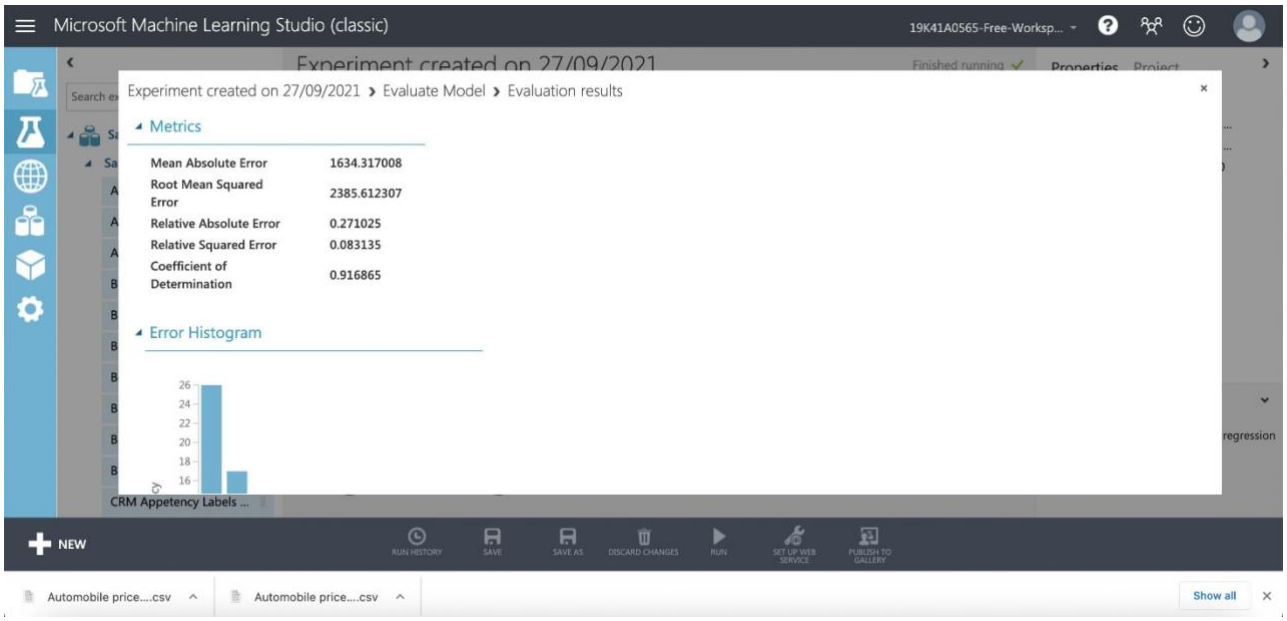





Microsoft Machine Learning Studio (classic)
19K41A0565-Free-Worksp...
Experiment created on 27/09/2021
Experiment created on 27/09/2021 › Score Model › Scored dataset
rows 58
columns 9
make body-style wheel-base engine-size horsepower peak-rpm highway-mpg price
view as
saab sedan 99.1 121 110 5250 28 15510
mitsubishi hatchback 93 98 102 5500 30 7689
volkswagen sedan 97.3 109 85 5250 34 8195
chevrolet hatchback 94.5 90 70 5400 43 6295
isuzu hatchback 96 119 90 5000 29 11048
mercedes-benz sedan 115.6 234 155 4750 18 34184
nissan sedan 94.5 97 69 5200 37 5499
dodge hatchback 93.7 98 102 5500 30 7957
Statistics
Mean 12916.7082
Median 10539.5864
Min 5526.7789
Max 35033.2782
Standard Deviation 7427.924
Unique Values 55
Missing Values 0
Feature Type Numeric Score
Visualizations
Scored Labels
Histogram
CRM Appetency Labels ...
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY
Automobile price....csv
Automobile price....csv
Show all


Microsoft Machine Learning Studio (classic)
19K41A0565-Free-Worksp...
Experiment created on 27/09/2021
Experiment created on 27/09/2021 › Score Model › Scored dataset
rows 58
columns 9
make body-style wheel-base engine-size horsepower peak-rpm highway-mpg price
view as
saab sedan 99.1 121 110 5250 28 15510
mitsubishi hatchback 93 98 102 5500 30 7689
volkswagen sedan 97.3 109 85 5250 34 8195
chevrolet hatchback 94.5 90 70 5400 43 6295
isuzu hatchback 96 119 90 5000 29 11048
mercedes-benz sedan 115.6 234 155 4750 18 34184
nissan sedan 94.5 97 69 5200 37 5499
dodge hatchback 93.7 98 102 5500 30 7957
frequency
Scored Labels
CRM Appetency Labels ...
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY
Automobile price....csv
Automobile price....csv
Show all

- Evaluation Results




Microsoft Machine Learning Studio (classic)
19K41A0565-Free-Worksp...
Experiment created on 27/09/2021
Finished running
Properties
Project
Experiment created on 27/09/2021 › Evaluate Model › Evaluation results
Error Histogram
frequency
Error
CRM Appetency Labels ...
NEW
RUN HISTORY
SAVE
SAVE AS
DISCARD CHANGES
RUN
SET UP WEB SERVICE
PUBLISH TO GALLERY
Automobile price....csv
Automobile price....csv