20K45A0240
GUNDAPU RAMA KRISHNA

Azure Microsoft Machine Learning Studio

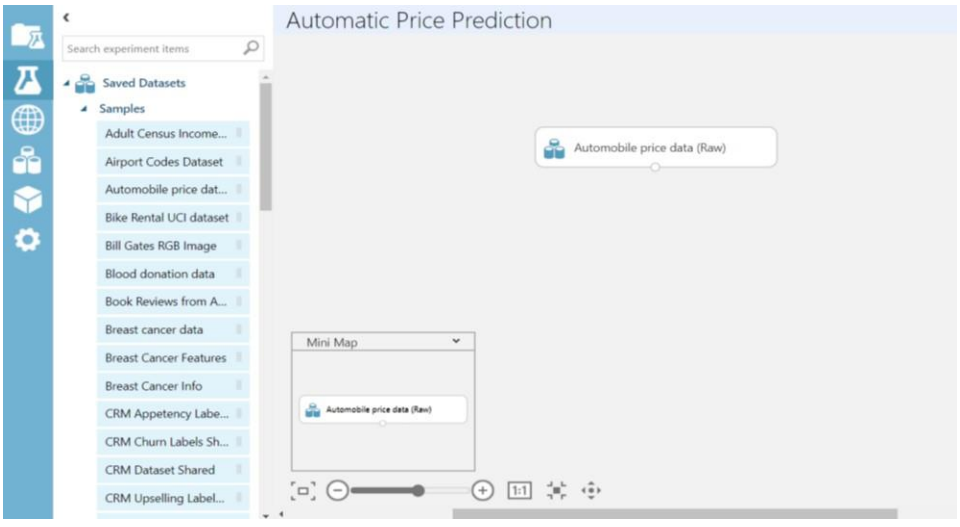
Automatic price prediction Model

- In this project, we tried to create a model which automatically predicts the price of automobile based on the data sets that are already available in Azure Microsoft Machine Learning Studio.

Workflow:

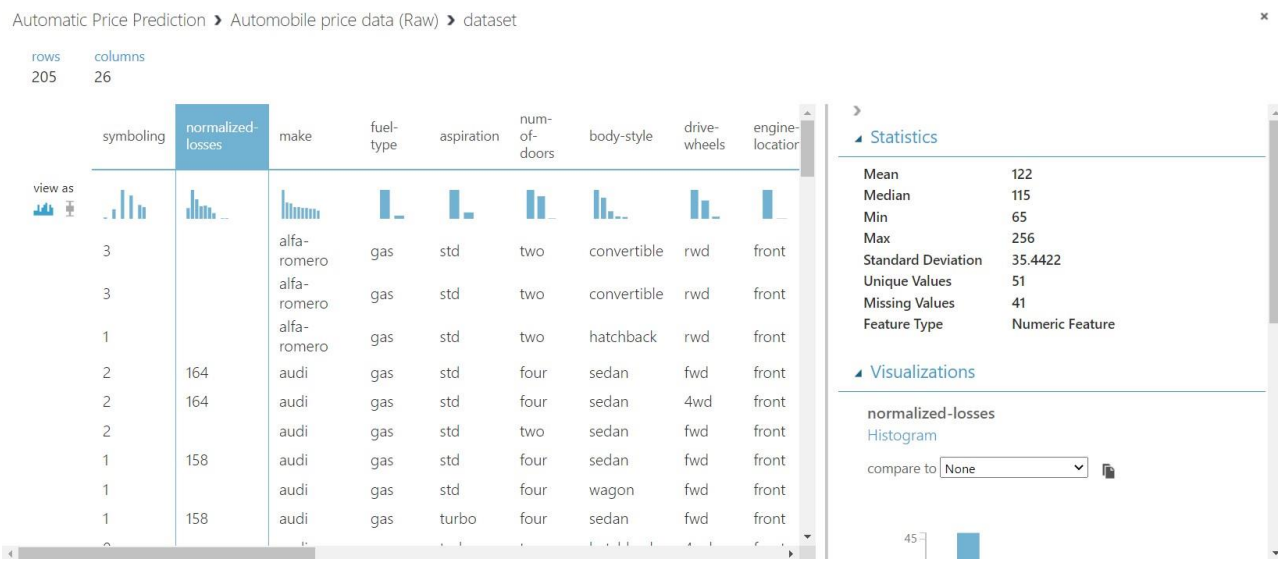
1. Import the Data
2. Identify the columns that has more missing values
3. Missing Values Treatment
4. Split the data
5. Train the model
6. Test the model
7. Evaluate of the model

1. Import the data



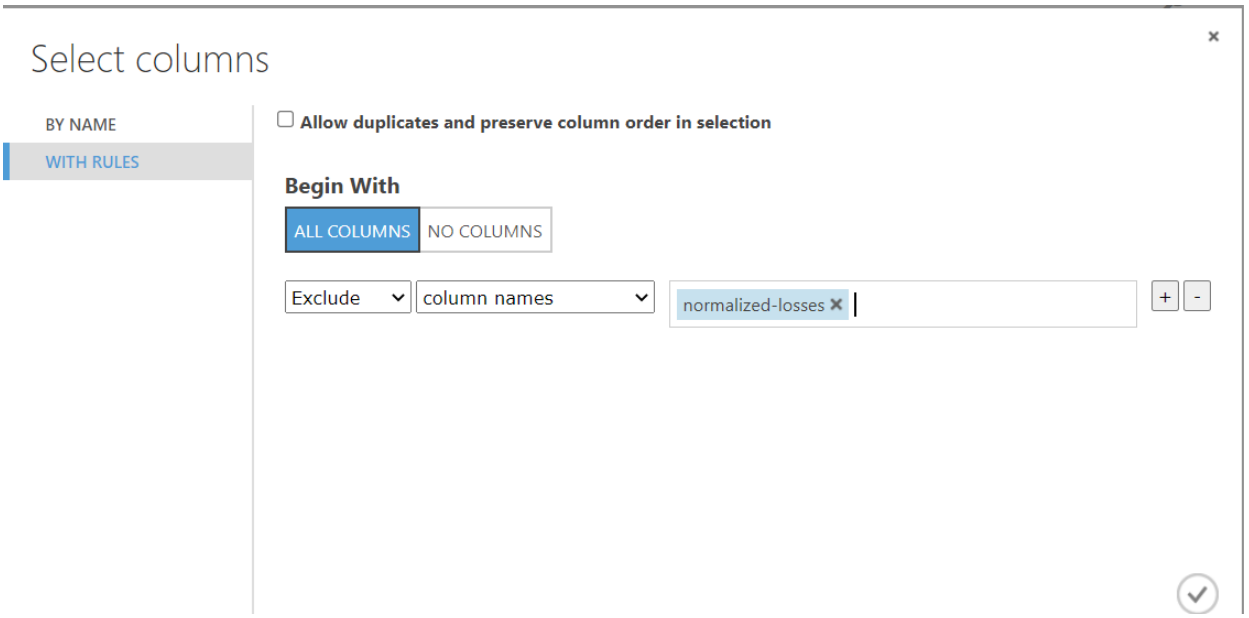
2. Identify the columns that has more missing values

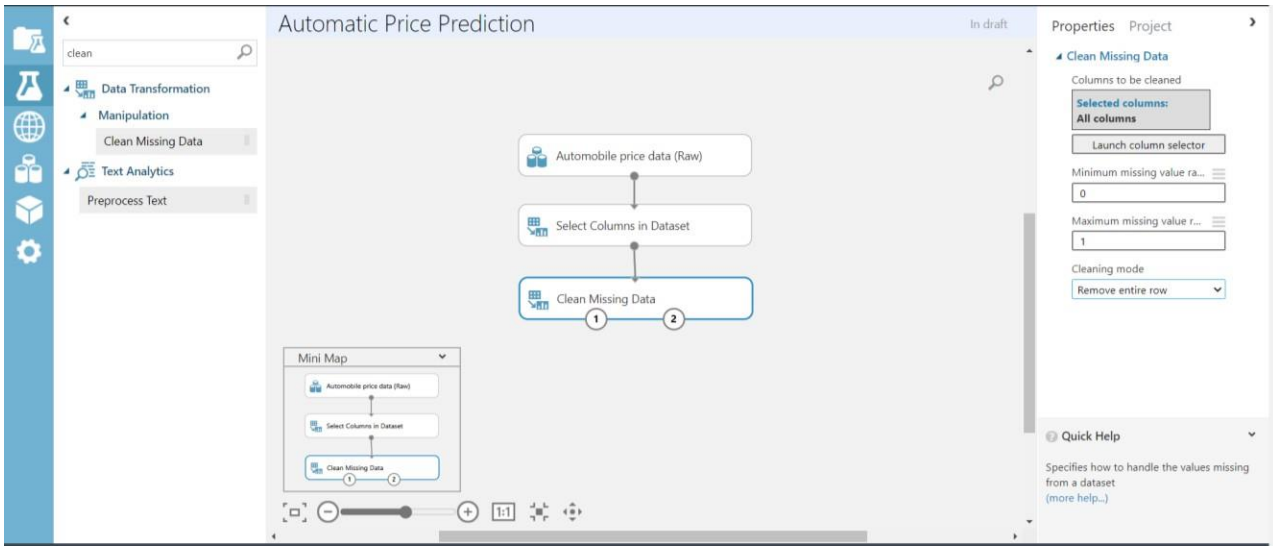
Here we observe that the column normalised-losses has more missing values.



3. Missing Values Treatment

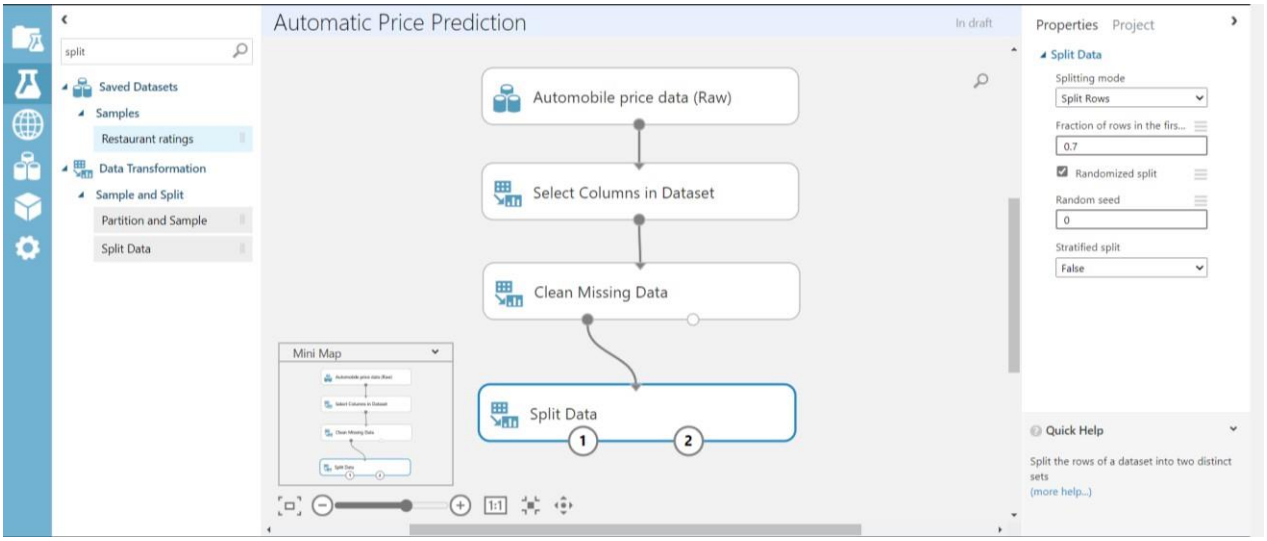
Since the column normalized-losses has more missing values we can exclude that column.





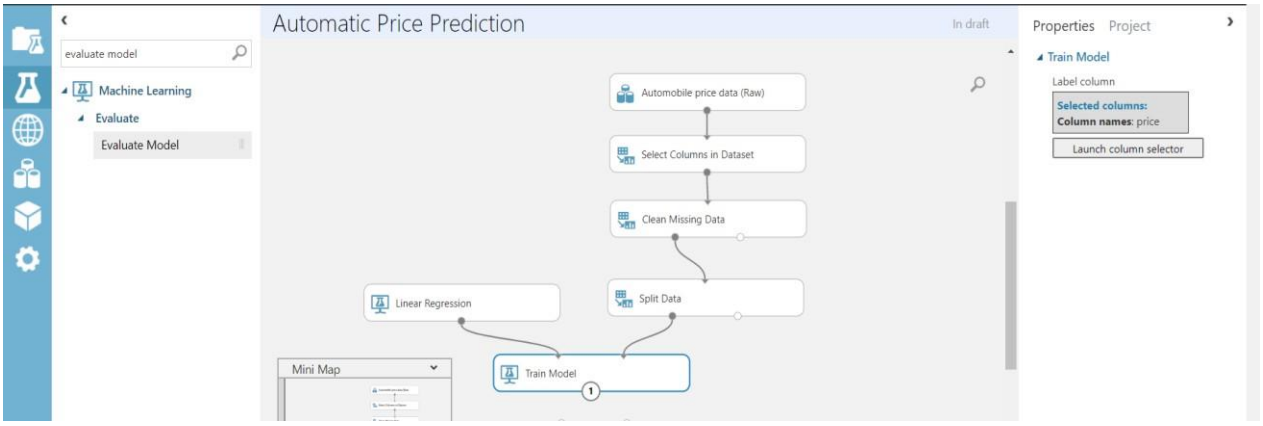
4. Split the data

We usually split the data for training and testing randomly. Here we used 70% of data for training and the remaining for testing.

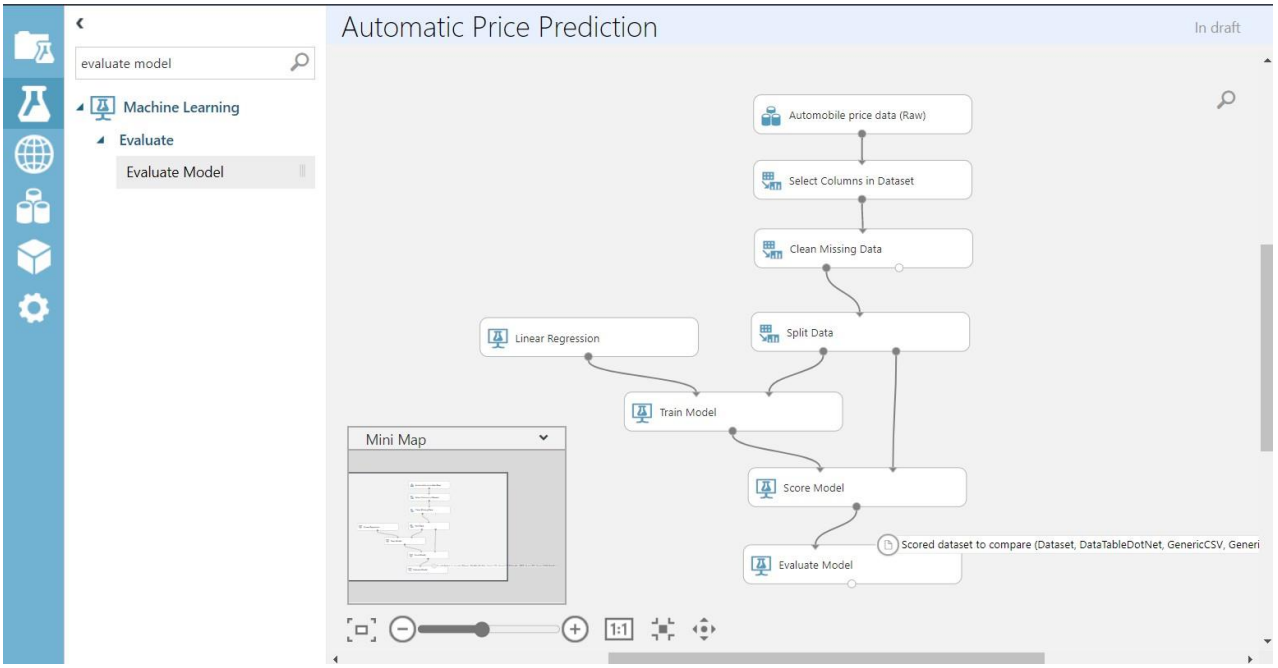


5. Train the Model

We used linear regression model for the analysis.

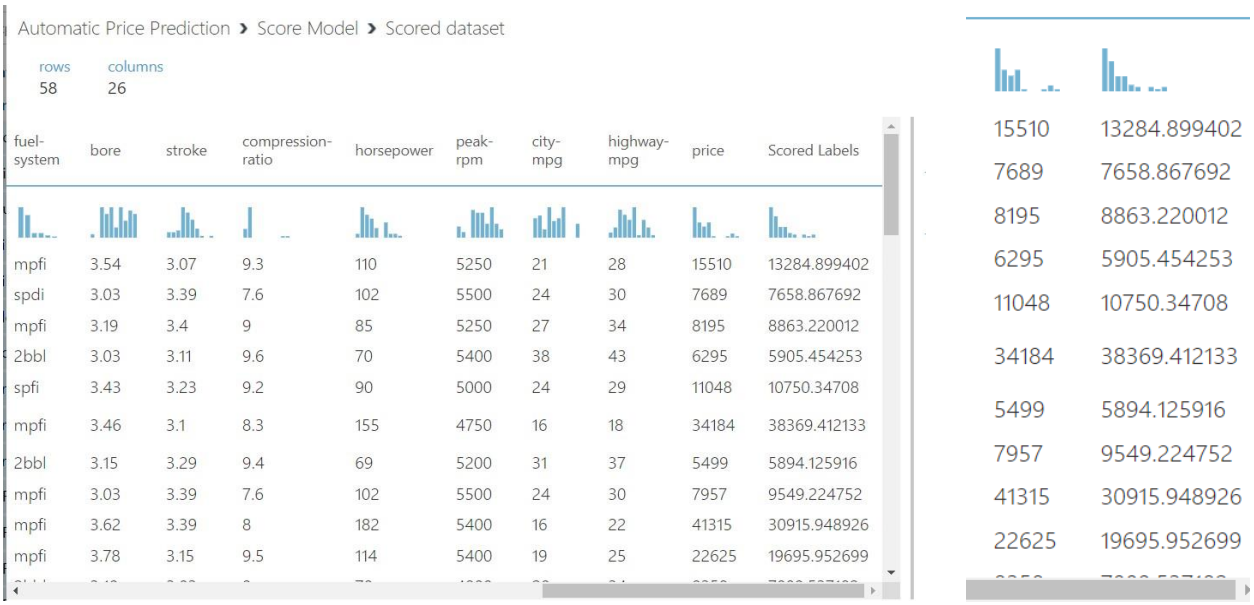


6. Test the model



7. Evaluate of the model

These are the results after evaluating the model.



Automatic Price Prediction › Score Model › Scored dataset

rows 58 columns 26

fuel-system	bore	stroke	compression-ratio	horsepower	peak-rpm	city-mpg	highway-mpg	price	Scored Labels
mpfi	3.54	3.07	9.3	110	5250	21	28	15510	13284.899402
spdi	3.03	3.39	7.6	102	5500	24	30	7689	7658.867692
mpfi	3.19	3.4	9	85	5250	27	34	8195	8863.220012
2bbl	3.03	3.11	9.6	70	5400	38	43	6295	5905.454253
spfi	3.43	3.23	9.2	90	5000	24	29	11048	10750.34708
mpfi	3.46	3.1	8.3	155	4750	16	18	34184	38369.412133
2bbl	3.15	3.29	9.4	69	5200	31	37	5499	5894.125916
mpfi	3.03	3.39	7.6	102	5500	24	30	7957	9549.224752
mpfi	3.62	3.39	8	182	5400	16	22	41315	30915.948926
mpfi	3.78	3.15	9.5	114	5400	19	25	22625	19695.952699

price	Scored Labels
15510	13284.899402
7689	7658.867692
8195	8863.220012
6295	5905.454253
11048	10750.34708
34184	38369.412133
5499	5894.125916
7957	9549.224752
41315	30915.948926
22625	19695.952699

Automatic Price Prediction › Evaluate Model › Evaluation results

Metrics

Mean Absolute Error	1605.514464
Root Mean Squared Error	2385.271889
Relative Absolute Error	0.266248
Relative Squared Error	0.083112
Coefficient of Determination	0.916888

Error Histogram