

AZURE ML Classic Studio

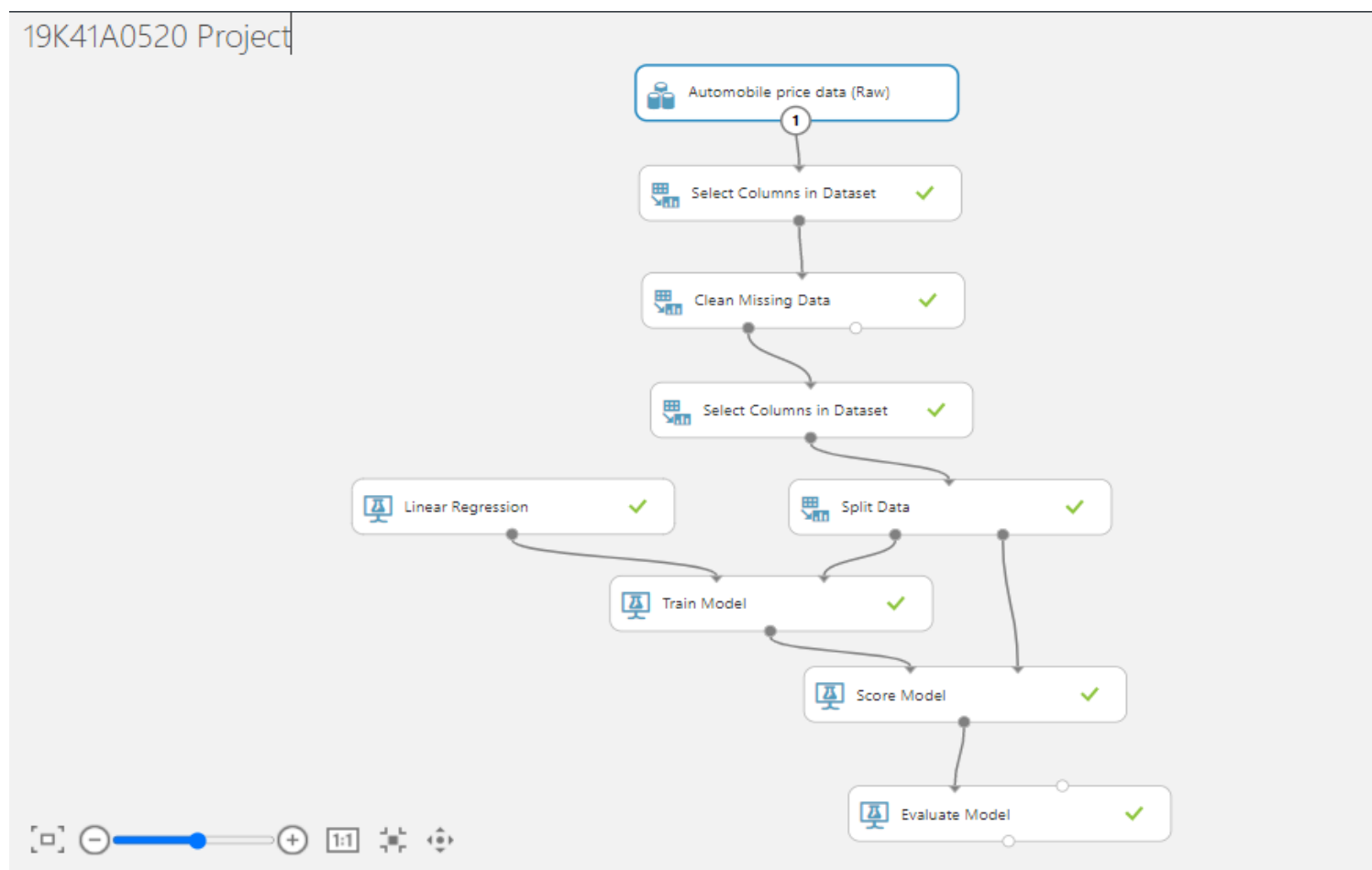
Predicting Automobile prices using Regression Model in Azure ML Classic Studio.

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

Project Workflow:

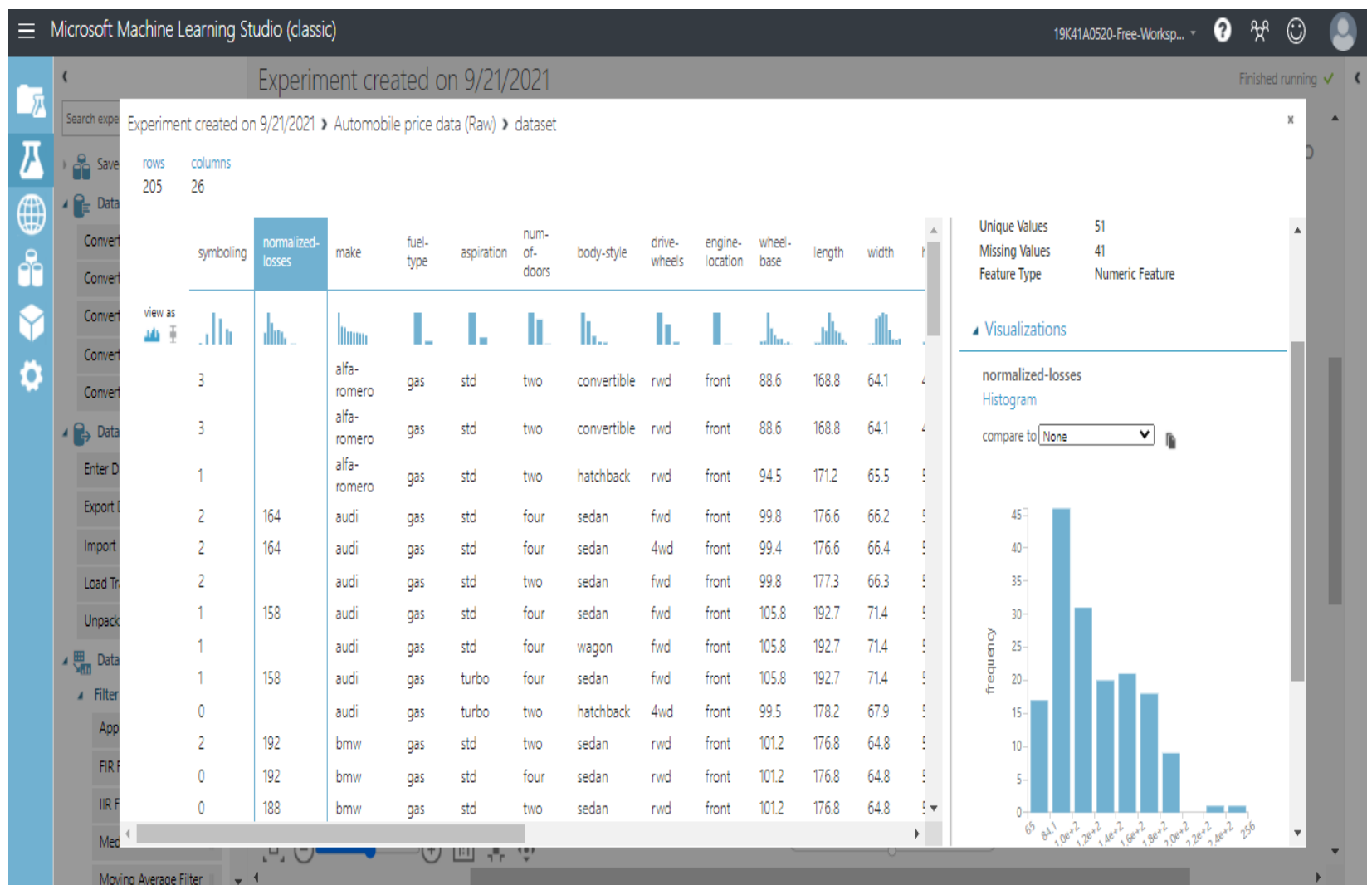
1. Load the data.
2. Explore data (missing values).
3. Preprocess the data.
4. Choose the model (Linear Regression).
5. Split the data → Training and Testing.
6. Train the model.
7. Score the model.
8. Evaluate the model based on results.

Workflow:-

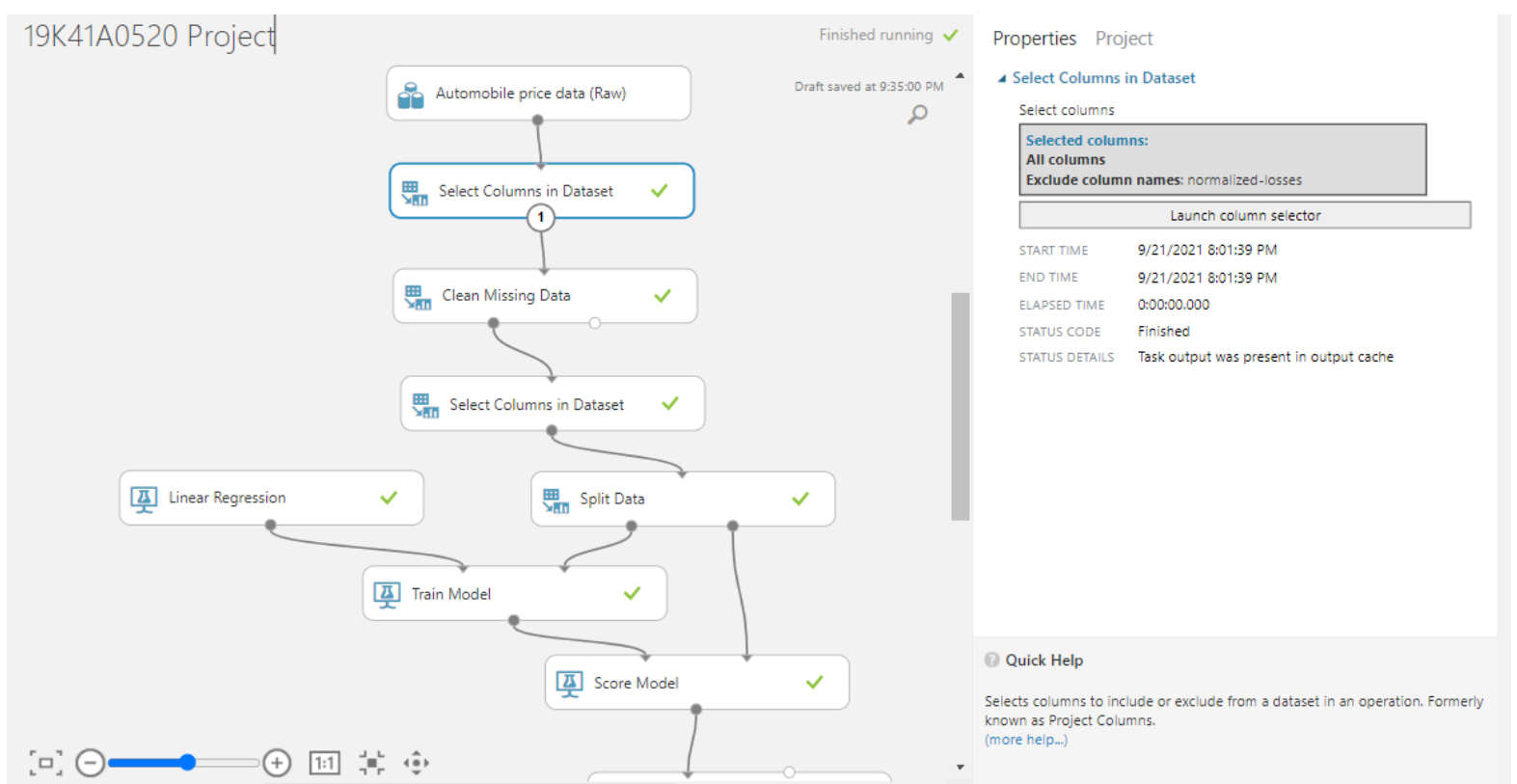


- **Import data:**

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



- **Remove the column that has more number of null values.**



- Clean missing data

19K41A0520 Project

Finished running ✓

Draft saved at 9:35:12 PM

Automobile price data (Raw)

Select Columns in Dataset ✓

Clean Missing Data ✓

Select Columns in Dataset ✓

Linear Regression ✓

Split Data ✓

Train Model ✓

Score Model ✓

Properties Project

Clean Missing Data

Columns to be cleaned

Selected columns: All columns

Launch column selector

Minimum missing value ratio 0

Maximum missing value ratio 1

Cleaning mode Remove entire row

START TIME 9/21/2021 8:01:39 PM

END TIME 9/21/2021 8:01:39 PM

ELAPSED TIME 0:00:00.000

STATUS CODE Finished

STATUS DETAILS Task output was present in output cache

Quick Help

Specifies how to handle the values missing from a dataset (more help...)

19K41A0520 Project

Finished running ✓

Automobile price data (Raw)

Select Columns in Dataset ✓

Clean Missing Data ✓

Select Columns in Dataset ✓

Linear Regression ✓

Split Data ✓

Train Model ✓

Score Model ✓

Evaluate Model ✓

Properties Project

Select Columns in Dataset

Select columns

Selected columns: All columns

Column names: price

Launch column selector

START TIME 9/25/2021 7:18:28 PM

END TIME 9/25/2021 7:18:29 PM

ELAPSED TIME 0:00:01.503

STATUS CODE Finished

STATUS DETAILS None

View output log

Quick Help

Selects columns to include or exclude from a dataset in an operation. Formerly known as Project Columns. (more help...)

RUN HISTORY SAVE SAVE AS DISCARD CHANGES RUN SET UP WEB SERVICE PUBLISH TO GALLERY

- **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 data set contains 30% of the original data.

19K41A0520 Project Finished running ✓ Draft saved at 9:35:12 PM

Properties Project

Split Data

Splitting mode: Split Rows

Fraction of rows in the first output dataset: 0.7

☒ Randomized split

Random seed: 0

Stratified split: False

START TIME: 9/21/2021 8:01:40 PM

END TIME: 9/21/2021 8:01:40 PM

ELAPSED TIME: 0:00:00.000

STATUS CODE: Finished

STATUS DETAILS: Task output was present in output cache

Quick Help

Split the rows of a dataset into two distinct sets (more help...)

- **Linear Regression to train model**

19K41A0520 Project Finished running ✓ Draft saved at 9:35:12 PM

Properties Project

Linear Regression

Solution method: Ordinary Least Squares

L2 regularization weight: 0.001

☒ Include intercept term

Random number seed:

☒ Allow unknown categorical levels

START TIME: 9/21/2021 8:01:39 PM

END TIME: 9/21/2021 8:01:39 PM

ELAPSED TIME: 0:00:00.000

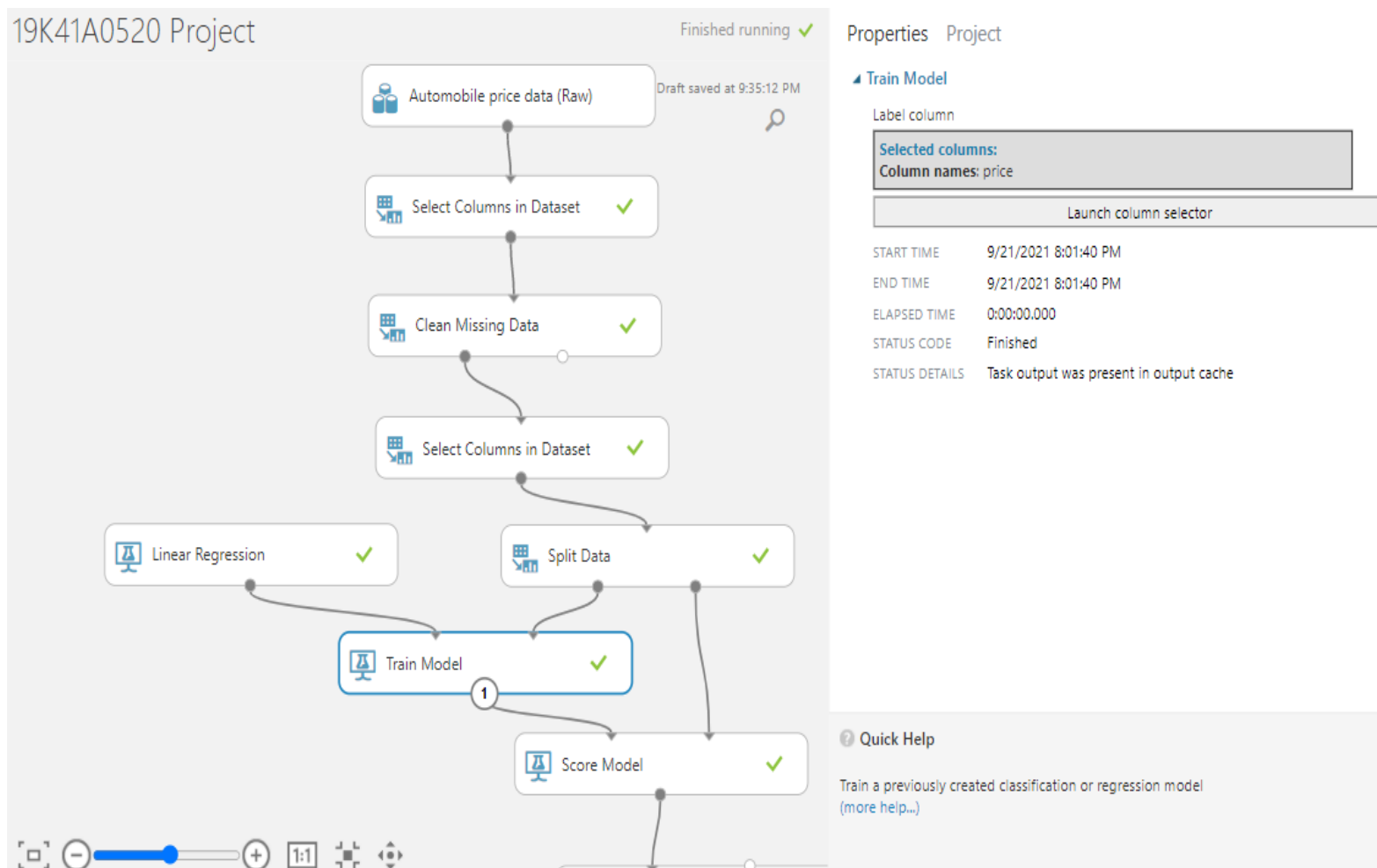
STATUS CODE: Finished

STATUS DETAILS: Task output was present in output cache

Quick Help

Creates a linear regression model (more help...)

• Model Training



• Score Model and Evaluate Model

After the model is trained, we can use the Score Model and Evaluate Model modules to generate predicted results and evaluate the models.

19K41A0520 Project > Score Model > Scored dataset

rows58columns26

	symboling	make	fuel-type	aspiration	num-of-doors	body-style	drive-wheels	engine-location	wheel-base	length	width	height	curb-weight	engine-type	num-of-cylinders	engine-size	fuel-system	bore	stroke	compression-ratio	horsepower
view as																					
2	saab	gas	std	four	sedan	fwd	front	99.1	186.6	66.5	56.1	2758	ohc	four	121	mpfi	3.54	3.07	9.3	110	
1	mitsubishi	gas	turbo	two	hatchback	fwd	front	93	157.3	63.8	50.8	2145	ohc	four	98	spdi	3.03	3.39	7.6	102	
2	volkswagen	gas	std	four	sedan	fwd	front	97.3	171.7	65.5	55.7	2212	ohc	four	109	mpfi	3.19	3.4	9	85	
1	chevrolet	gas	std	two	hatchback	fwd	front	94.5	155.9	63.6	52	1874	ohc	four	90	2bbl	3.03	3.11	9.6	70	
2	isuzu	gas	std	two	hatchback	rwd	front	96	172.6	65.2	51.4	2734	ohc	four	119	spfi	3.43	3.23	9.2	90	
-1	mercedes-benz	gas	std	four	sedan	rwd	front	115.6	202.6	71.7	56.5	3740	ohcv	eight	234	mpfi	3.46	3.1	8.3	155	
1	nissan	gas	std	two	sedan	fwd	front	94.5	165.3	63.8	54.5	1889	ohc	four	97	2bbl	3.15	3.29	9.4	69	
1	dodge	gas	turbo	two	hatchback	fwd	front	93.7	157.3	63.8	50.8	2128	ohc	four	98	mpfi	3.03	3.39	7.6	102	
0	bmw	gas	std	two	sedan	rwd	front	103.5	193.8	67.9	53.7	3380	ohc	six	209	mpfi	3.62	3.39	8	182	
-1	volvo	gas	turbo	four	sedan	rwd	front	109.1	188.8	68.9	55.5	3062	ohc	four	141	mpfi	3.78	3.15	9.5	114	
0	toyota	gas	std	four	hatchback	fwd	front	95.7	166.3	64.4	52.8	2122	ohc	four	98	2bbl	3.19	3.03	9	70	
-1	volvo	gas	std	four	wagon	rwd	front	104.3	188.8	67.2	57.5	3042	ohc	four	141	mpfi	3.78	3.15	9.5	114	
1	alfa-romero	gas	std	two	hatchback	rwd	front	94.5	171.2	65.5	52.4	2823	ohcv	six	152	mpfi	2.68	3.47	9	154	
1	dodge	gas	std	four	hatchback	fwd	front	93.7	157.3	63.8	50.6	1967	ohc	four	90	2bbl	2.97	3.23	9.4	68	
2	toyota	gas	std	two	hatchback	rwd	front	98.4	176.2	65.6	52	2551	ohc	four	146	mpfi	3.62	3.5	9.3	116	
2	volkswagen	diesel	std	four	sedan	fwd	front	97.3	171.7	65.5	55.7	2264	ohc	four	97	idi	3.01	3.4	23	52	
1	plymouth	gas	std	two	hatchback	fwd	front	93.7	157.3	63.8	50.8	1918	ohc	four	90	2bbl	2.97	3.23	9.4	68	

Statistics and Visualizations

Evaluation Results:

19K41A0520 Project > Evaluate Model > Evaluation results

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

