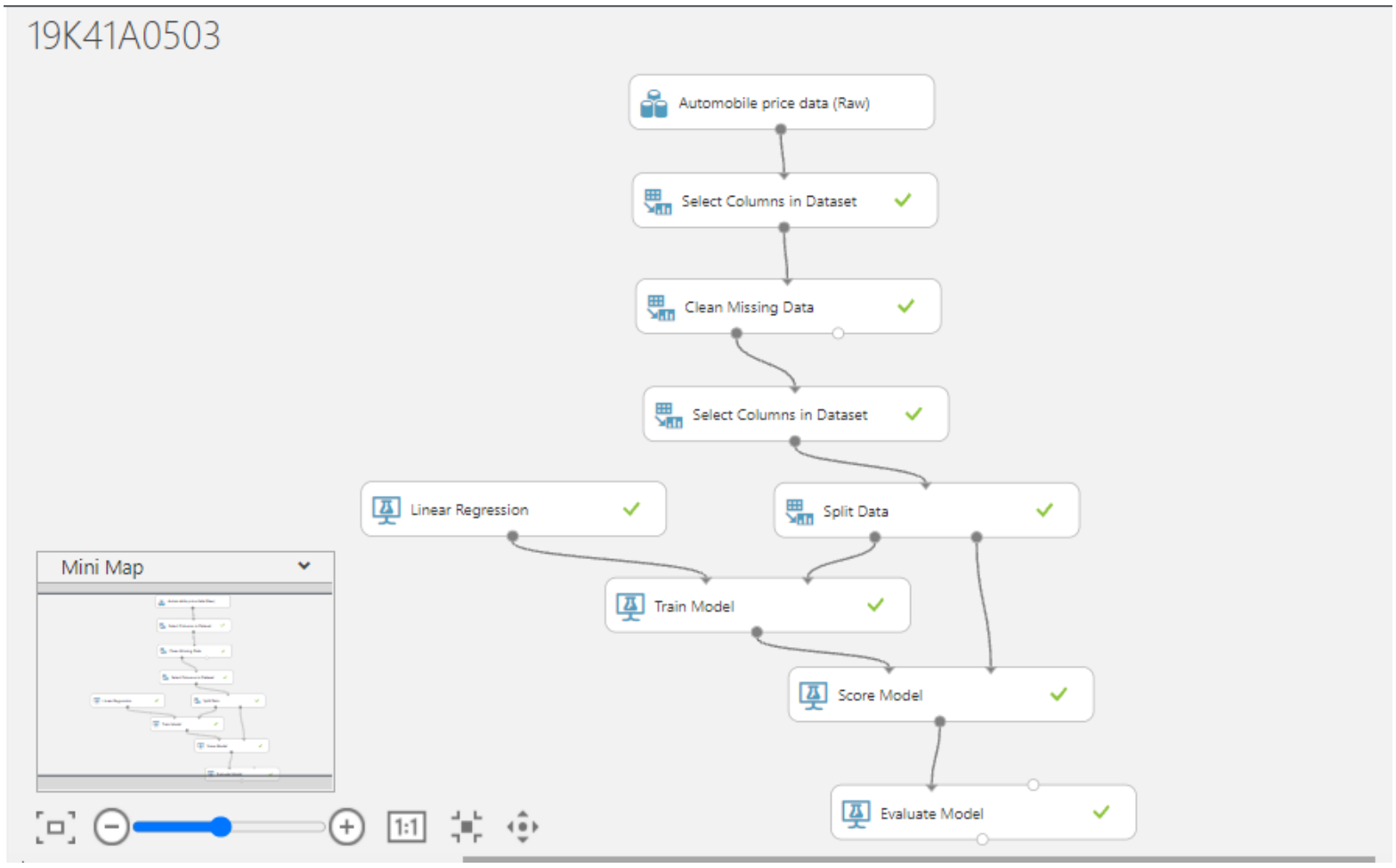
AZURE ML Classic Studio

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

This model (Pipeline) trains a linear regressor to predict a car's price based on technical features such as make, model, horsepower, and size. Because you're trying to answer the question "How much?" this is called a regression problem. However, you can apply the same fundamental steps in this example to tackle any type of machine learning problem whether it be regression, classification, clustering, and so on.

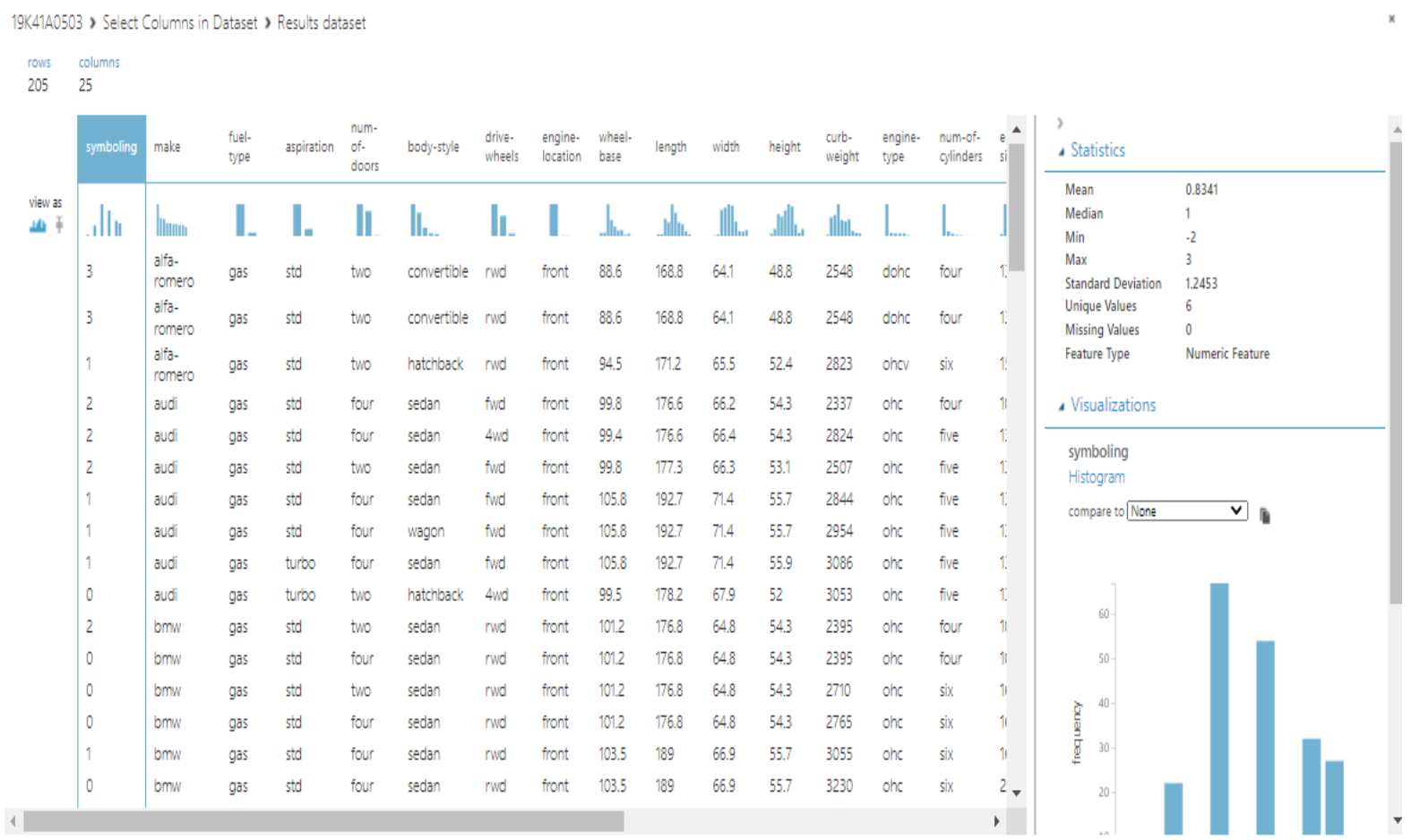
1. Load the data.
2. Explore data(missing values, outlier treatment, normalization).
3. Preprocess the data.
4. Choose the model(Linear Regression).
5. Split the data(Training and Testing).
6. Train the model.
7. Score the model.
8. Evaluate the model based on results.

Workflow:-

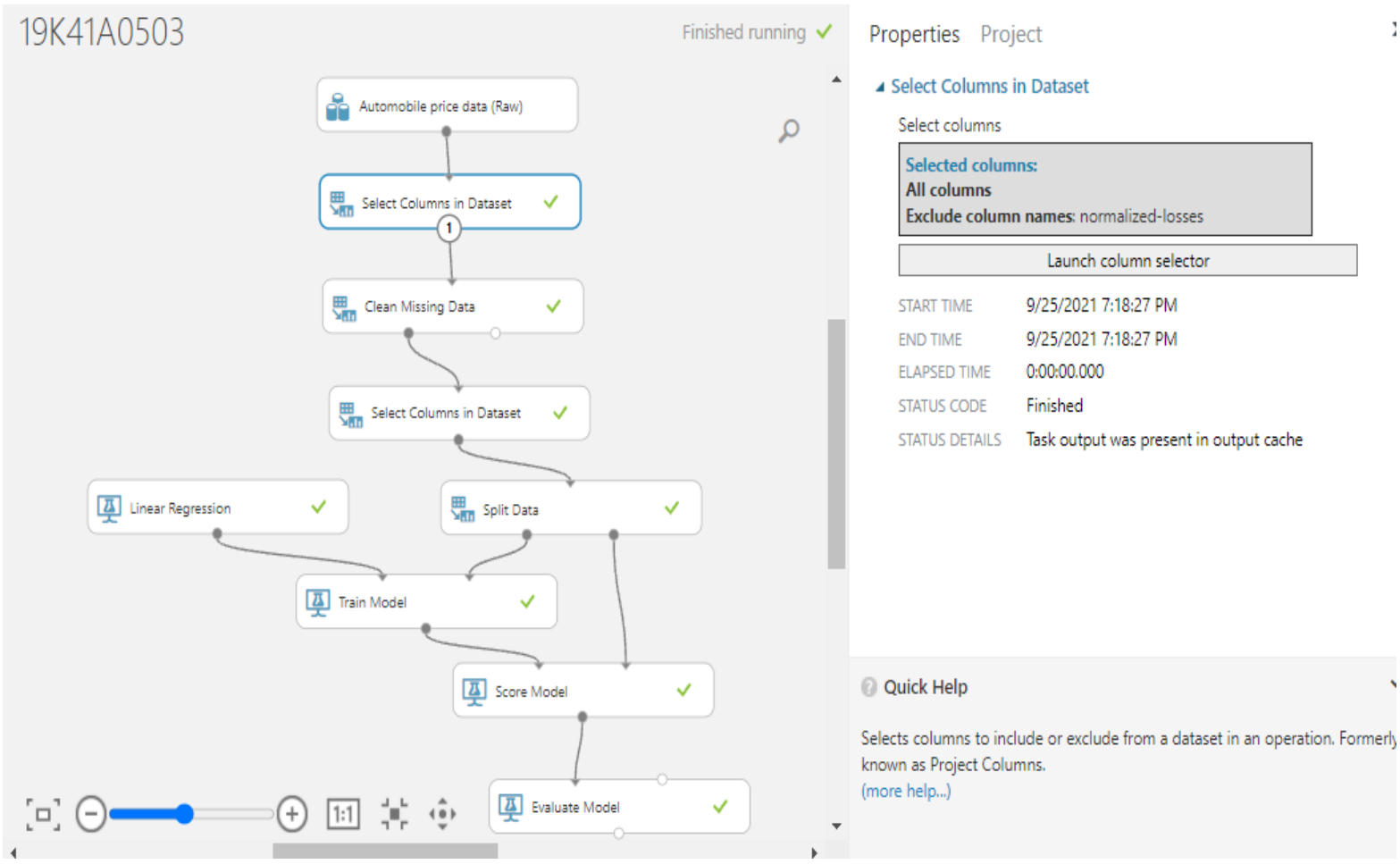


- **Import data:**

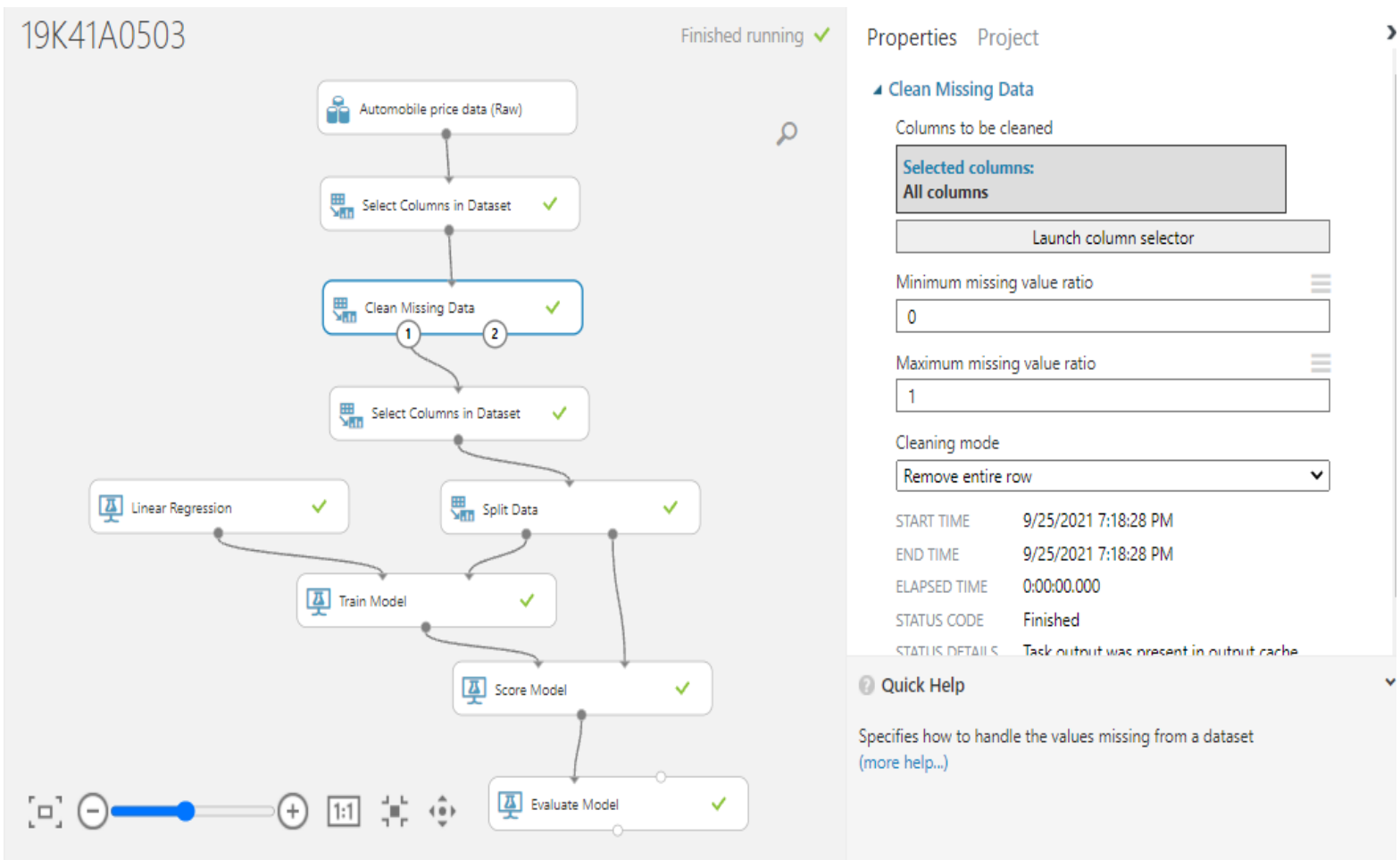
Importing the RAW dataset which is in CSV format.The dataset is pre-available in the Azure ML Classic Studio.

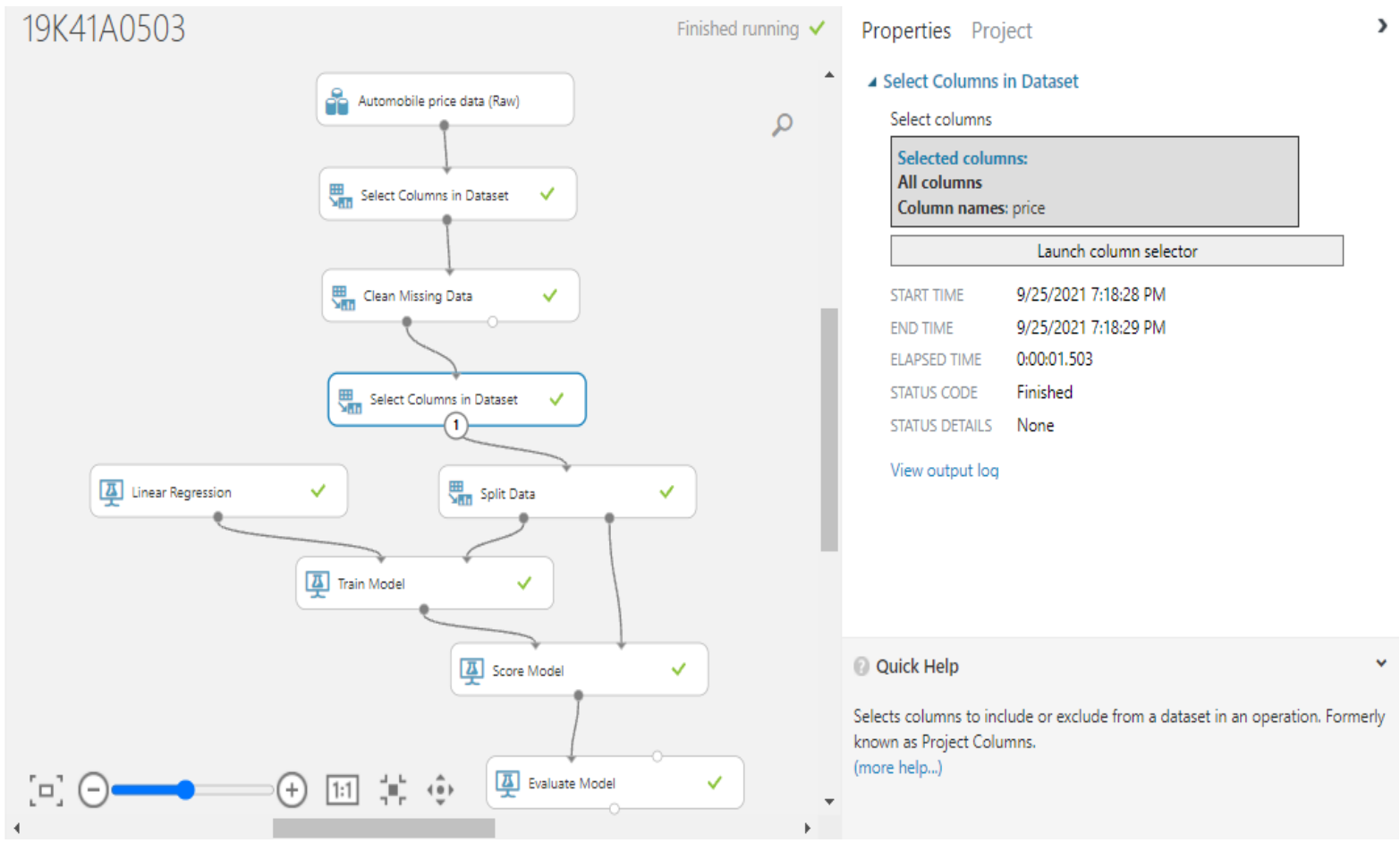


- **Remove the column that has more number of null values.**

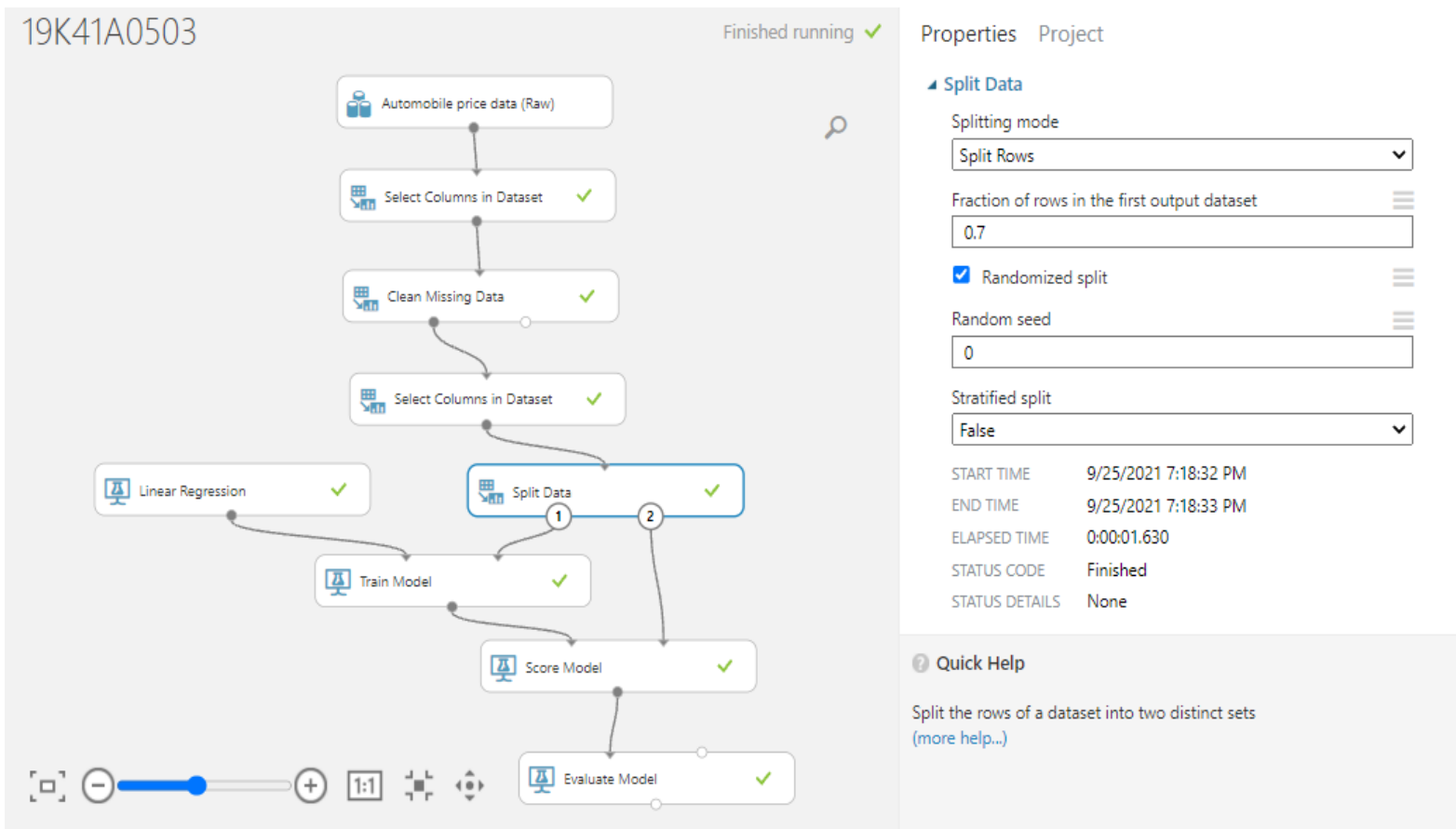


- **Clean missing data**

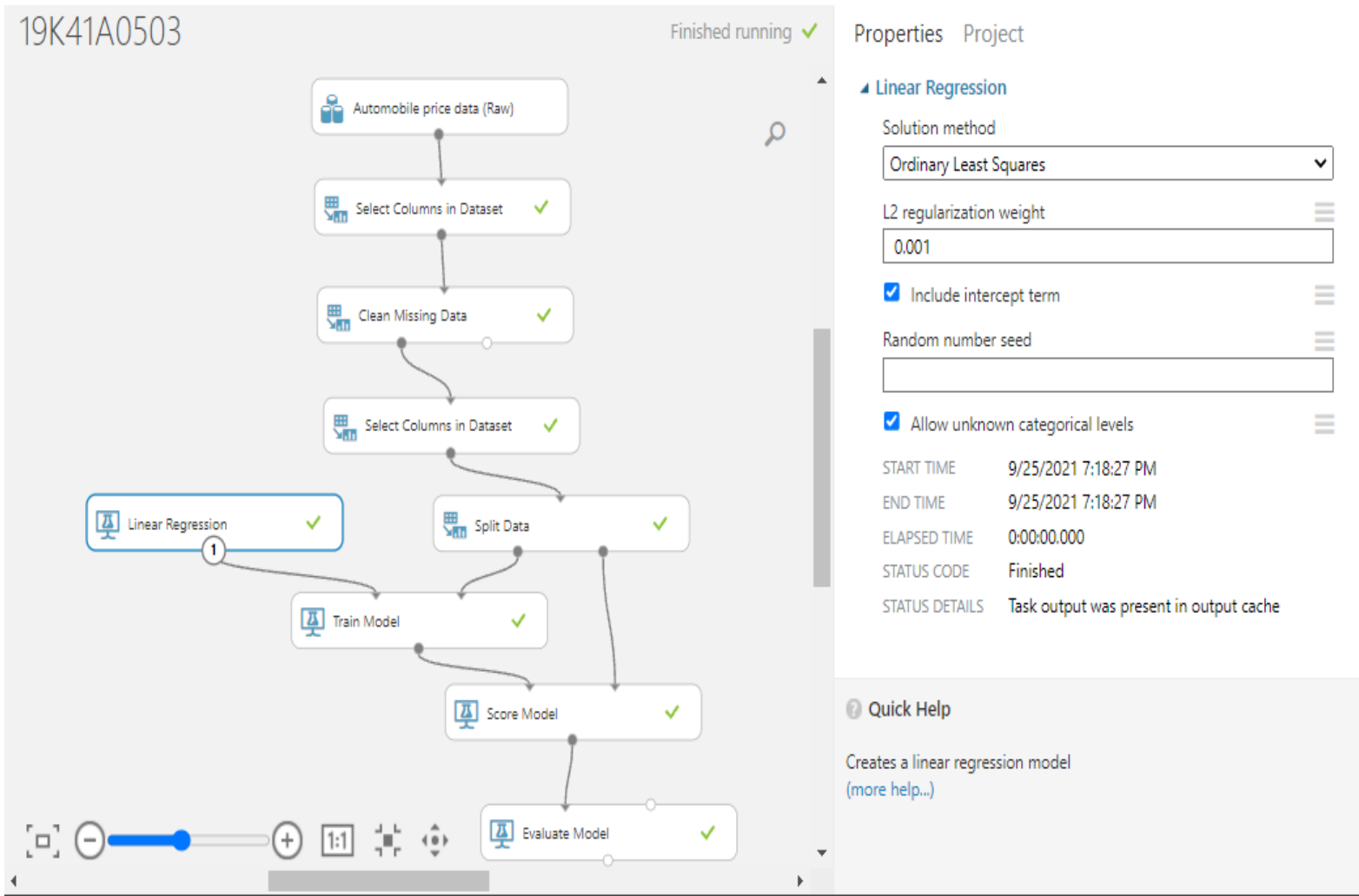




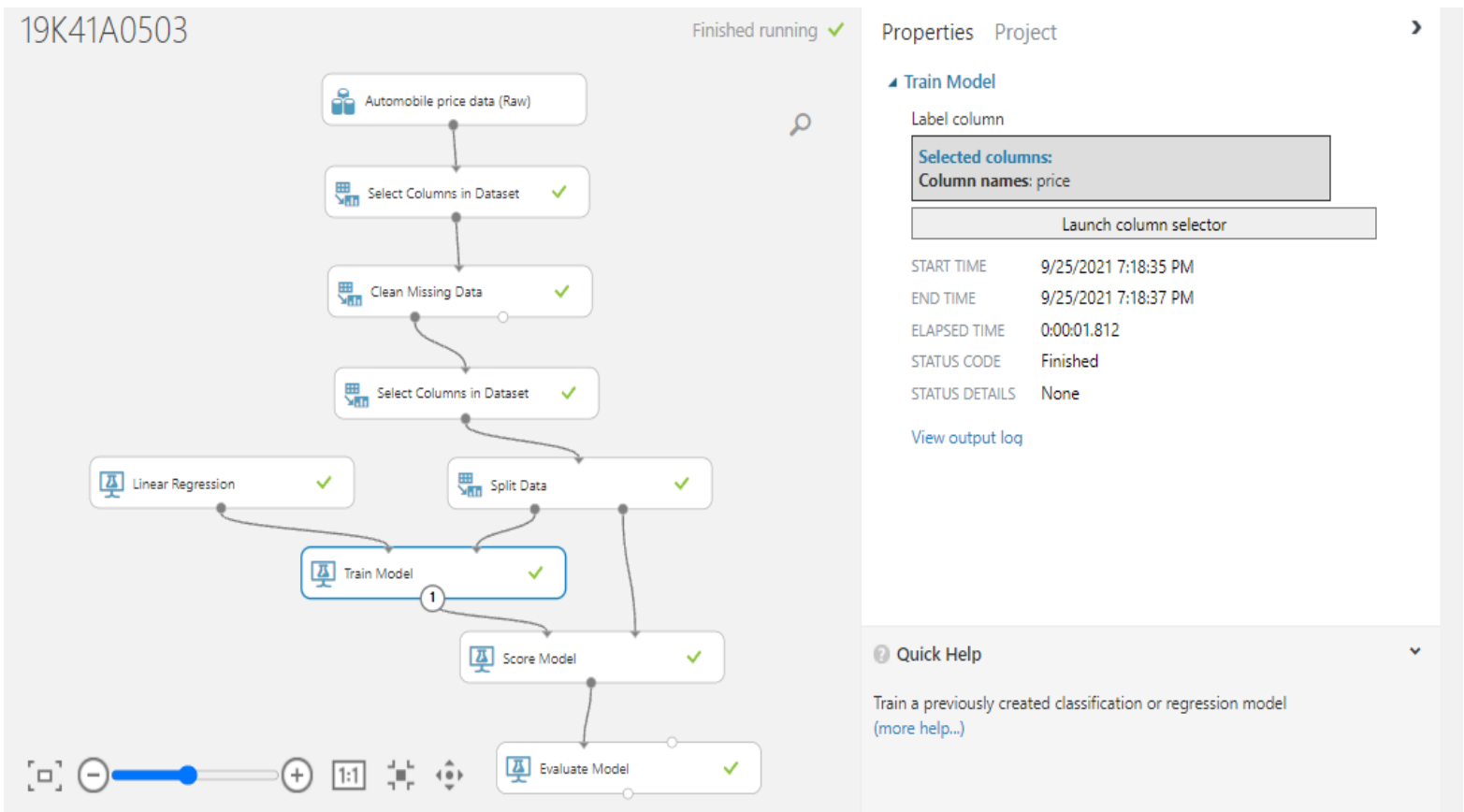
- **Split data:**Use the split data module to randomly divide the input data so that the training dataset contains 70% of the original data and the testing data set contains 30% of the original data.



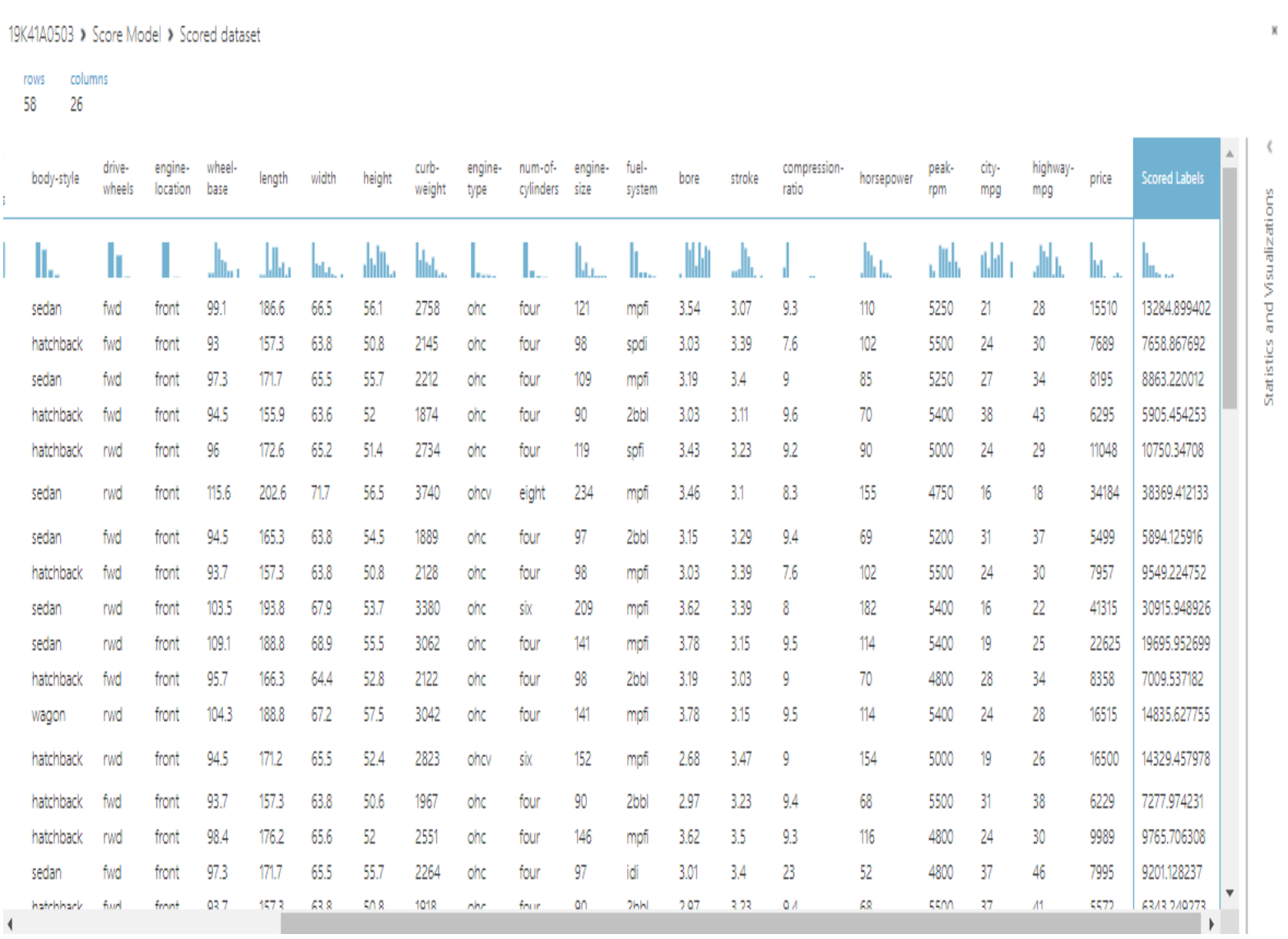
- **Linear Regression**



- **Model Training**



- **Score Model and Evaluate Model**



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	peak-rpm	city-mpg	highway-mpg	price	Scored Labels
sedan	fwd	front	99.1	186.6	66.5	56.1	2758	ohc	four	121	mpfi	3.54	3.07	9.3	110	5250	21	28	15510	13284.899402
hatchback	fwd	front	93	157.3	63.8	50.8	2145	ohc	four	98	spdi	3.03	3.39	7.6	102	5500	24	30	7689	7658.867692
sedan	fwd	front	97.3	171.7	65.5	55.7	2212	ohc	four	109	mpfi	3.19	3.4	9	85	5250	27	34	8195	8863.220012
hatchback	fwd	front	94.5	155.9	63.6	52	1874	ohc	four	90	2bbl	3.03	3.11	9.6	70	5400	38	43	6295	5905.454253
hatchback	rwd	front	96	172.6	65.2	51.4	2734	ohc	four	119	spfi	3.43	3.23	9.2	90	5000	24	29	11048	10750.34708
sedan	rwd	front	115.6	202.6	71.7	56.5	3740	ohcv	eight	234	mpfi	3.46	3.1	8.3	155	4750	16	18	34184	38369.412133
sedan	fwd	front	94.5	165.3	63.8	54.5	1889	ohc	four	97	2bbl	3.15	3.29	9.4	69	5200	31	37	5499	5894.125916
hatchback	fwd	front	93.7	157.3	63.8	50.8	2128	ohc	four	98	mpfi	3.03	3.39	7.6	102	5500	24	30	7957	9549.224752
sedan	rwd	front	103.5	193.8	67.9	53.7	3380	ohc	six	209	mpfi	3.62	3.39	8	182	5400	16	22	41315	30915.948926
sedan	rwd	front	109.1	188.8	68.9	55.5	3062	ohc	four	141	mpfi	3.78	3.15	9.5	114	5400	19	25	22625	19695.952699
hatchback	fwd	front	95.7	166.3	64.4	52.8	2122	ohc	four	98	2bbl	3.19	3.03	9	70	4800	28	34	8358	7009.537182
wagon	rwd	front	104.3	188.8	67.2	57.5	3042	ohc	four	141	mpfi	3.78	3.15	9.5	114	5400	24	28	16515	14835.627755
hatchback	rwd	front	94.5	171.2	65.5	52.4	2823	ohcv	six	152	mpfi	2.68	3.47	9	154	5000	19	26	16500	14329.457978
hatchback	fwd	front	93.7	157.3	63.8	50.6	1967	ohc	four	90	2bbl	2.97	3.23	9.4	68	5500	31	38	6229	7277.974231
hatchback	rwd	front	98.4	176.2	65.6	52	2551	ohc	four	146	mpfi	3.62	3.5	9.3	116	4800	24	30	9989	9765.706308
sedan	fwd	front	97.3	171.7	65.5	55.7	2264	ohc	four	97	idi	3.01	3.4	23	52	4800	37	46	7995	9201.128237
hatchback	fwd	front	93.7	157.3	63.8	50.8	1918	ohc	four	90	2bbl	2.97	3.23	9.4	68	5500	37	41	5572	6343.249273

Evaluation Results:

19K41A0503 › 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