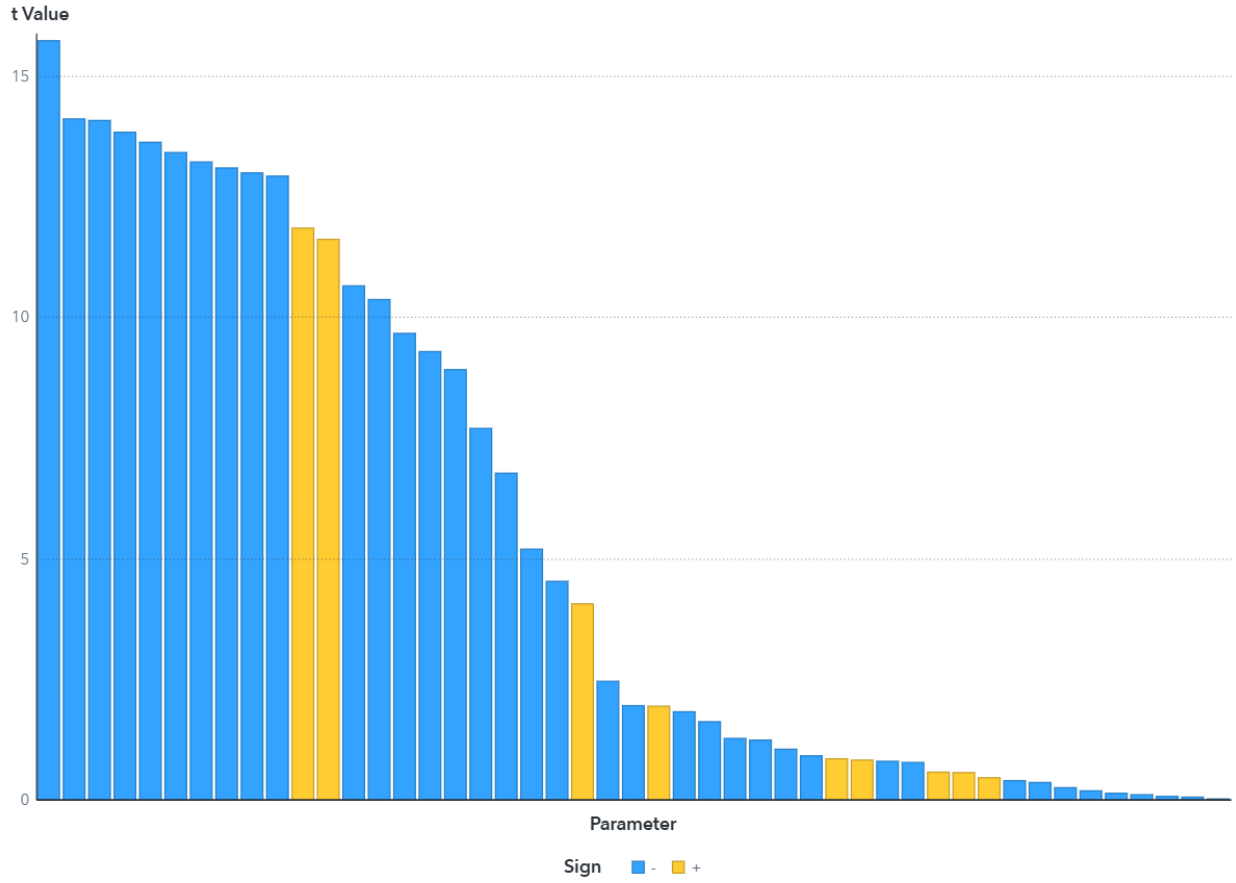
UAS

"Linear Regression" Results

by: vincent.richie@student.umn.ac.id

Contents

t Values by Parameter



This plot displays the absolute value of the t value for each parameter estimate in the model. Larger values indicate more significant parameters. The bar that represents the parameter is colored by the sign of the estimate. Bars that are colored as positive (+) correspond to a positive parameter estimate, which indicates an increase in the predicted target value as the parameter value increases. Bars that are colored as negative (-) correspond to a negative parameter estimate, which indicates a decrease in the predicted target value as the parameter value increases. The most significant parameter is Fuel Type Gasoline with a t value of -15.733.

Parameter Estimates

Effect	Parameter	t Value	Sign
Fuel Type	Fuel Type Gasoline	15.7328	-
Color	Color White	14.1128	-
Fuel Type	Fuel Type Diesel	14.0815	-
Color	Color Silver	13.8382	-
Color	Color Brown	13.6302	-
Color	Color Gold	13.4182	-
Color	Color Red	13.2224	-
Color	Color Black	13.0984	-
Color	Color Blue	12.9961	-
Color	Color Grey	12.9295	-
Intercept	Intercept	11.8519	+
Cylinders	Cylinders 12	11.6197	+
Color	Color Burgundy	10.6573	-
Color	Color Green	10.3736	-
Color	Color Other Color	9.6735	-
Fuel Type	Fuel Type Electric	9.2962	-
Color	Color Beige	8.9229	-
Color	Color Orange	7.7067	-
Color	Color Purple	6.7787	-
Color	Color Teal	5.2061	-
Color	Color Tan	4.5395	-
Body Type	Body Type Sports Car	4.0733	+
Cylinders	Cylinders 4	2.4696	-
Cylinders	Cylinders 6	1.9661	-
Cylinders	Cylinders 10	1.9535	+
Cylinders	Cylinders 3	1.8390	-
Body Type	Body Type Pick Up	1.6332	-

Effect	Parameter	t Value	Sign
	Truck		
Body Type	Body Type Sedan	1.2871	-
Body Type	Body Type Crossover	1.2535	-
Body Type	Body Type Hatchback	1.0636	-
Cylinders	Cylinders 5	0.9268	-
Body Type	Body Type Soft Top Convertible	0.8651	+
Body Type	Body Type Hard Top Convertible	0.8386	+
Body Type	Body Type Utility Truck	0.8136	-
Body Type	Body Type Van	0.7899	-
Cylinders	Cylinders None	0.5882	+
Cylinders	Cylinders 8	0.5794	+
Location	Location Dubai	0.4731	+
Body Type	Body Type SUV	0.4134	-
Location	Location Ras Al Khaimah	0.3740	-
Body Type	Body Type Coupe	0.2667	-
Location	Location Sharjah	0.2018	-
Location	Location Ajman	0.1514	-
Location	Location Abu Dhabi	0.1221	-
Location	Location Fujeirah	0.0873	-
Body Type	Body Type Other	0.0714	-
Location	Location Al Ain	0.0348	-

Estimate	Absolute Estimate	Standard Error	Pr > |t|
-1,604,437.9215	1,604,437.9215	101,980.6259	0.0000
-956,199.0548	956,199.0548	67,753.8877	0.0000

Estimate	Absolute Estimate	Standard Error	Pr > |t|
-1,562,618.7594	1,562,618.7594	110,969.9337	0.0000
-965,170.4957	965,170.4957	69,746.7945	0.0000
-1,038,856.9824	1,038,856.9824	76,217.5097	0.0000
-1,019,242.5553	1,019,242.5553	75,959.5826	0.0000
-927,491.1743	927,491.1743	70,145.3797	0.0000
-891,107.2708	891,107.2708	68,031.8057	0.0000
-910,319.8901	910,319.8901	70,045.3708	0.0000
-886,979.5995	886,979.5995	68,601.0092	0.0000
2,825,785.0559	2,825,785.0559	238,423.7054	0.0000
1,005,597.4083	1,005,597.4083	86,542.5549	0.0000
-993,393.8855	993,393.8855	93,212.3639	0.0000
-821,373.4233	821,373.4233	79,179.1345	0.0000
-913,381.3156	913,381.3156	94,420.9555	0.0000
-1,631,766.5291	1,631,766.5291	175,531.0702	0.0000
-858,544.3524	858,544.3524	96,218.1824	0.0000
-742,010.0253	742,010.0253	96,280.9786	0.0000
-841,997.7099	841,997.7099	124,212.1242	0.0000
-954,273.2171	954,273.2171	183,297.3498	0.0000
-903,427.7644	903,427.7644	199,012.6596	0.0000
366,849.2890	366,849.2890	90,062.9619	0.0000
-200,305.0600	200,305.0600	81,106.9605	0.0136
-159,445.6802	159,445.6802	81,096.0397	0.0493
221,711.1569	221,711.1569	113,494.7756	0.0508
-196,318.8480	196,318.8480	106,752.7877	0.0660
-142,835.9791	142,835.9791	87,455.6118	0.1025
-108,529.2724	108,529.2724	84,319.6672	0.1981
-122,955.0873	122,955.0873	98,093.3293	0.2101
-95,172.8494	95,172.8494	89,484.1844	0.2876
-100,066.6994	100,066.6994	107,965.0699	0.3540

Estimate	Absolute Estimate	Standard Error	Pr > |t|
85,082.7992	85,082.7992	98,354.7681	0.3870
79,603.7835	79,603.7835	94,926.0743	0.4017
-97,956.0836	97,956.0836	120,394.3568	0.4159
-75,237.4233	75,237.4233	95,252.1765	0.4296
91,556.8318	91,556.8318	155,664.1492	0.5564
47,059.9603	47,059.9603	81,220.1928	0.5623
80,851.5975	80,851.5975	170,887.6195	0.6361
-34,691.4526	34,691.4526	83,915.8626	0.6793
-80,738.3435	80,738.3435	215,880.7342	0.7084
-22,838.7335	22,838.7335	85,634.1285	0.7897
-34,626.3188	34,626.3188	171,625.6839	0.8401
-26,600.9414	26,600.9414	175,707.1601	0.8797
-20,984.1839	20,984.1839	171,890.7394	0.9028
-20,304.7059	20,304.7059	232,568.2966	0.9304
-6,948.8060	6,948.8060	97,310.1079	0.9431
-6,431.4631	6,431.4631	184,905.7657	0.9723

Variance Inflation	Degrees of Freedom
7.4313	1
34.9480	1
6.5496	1
12.5048	1
4.4381	1
4.5031	1
10.0378	1
26.8202	1
10.9134	1
18.6343	1
0	1

Variance Inflation	Degrees of Freedom
7.5564	1
2.0831	1
3.4391	1
2.0364	1
2.8265	1
1.9571	1
1.9070	1
1.4154	1
1.1578	1
1.1375	1
7.3487	1
47.0354	1
50.5814	1
2.1371	1
2.3443	1
11.1734	1
48.5698	1
3.6646	1
8.5104	1
2.2654	1
3.6842	1
4.5416	1
1.9919	1
4.3703	1
1.8070	1
47.9179	1
161.8181	1
59.8176	1

Variance Inflation	Degrees of Freedom
2.6748	1
20.1104	1
90.0005	1
17.6221	1
72.2381	1
2.1741	1
3.8191	1
6.8392	1

Selection Summary

Step	Effect Entered	Number of Effects	SBC
0	Intercept	1	156,247.7651
1	Cylinders	2	154,843.1653
2	Fuel Type	3	154,601.6708
3	Color	4	154,424.2428
4	Body Type	5	154,310.4239
5	Location	6	154,307.3319

Optimal SBC
0
0
0
0
0
1

Regression Fit Statistics

Statistic	Description	Value
RMSE	Root MSE	417,063.1004
RSQUARE	R-Square	0.3253
ADJRSQ	Adj R-Sq	0.3200
AIC	AIC	159,942.8638
AICC	AICC	159,943.6612
SBC	SBC	154,307.3319
TRAIN_ASE	ASE (Train)	1.7256717E11
VAL_ASE	ASE (Validate)	1.1966907E11
TEST_ASE	ASE (Test)	1.0275235E11

Score Inputs

Name	Role	Variable Level	Type
Body Type	INPUT	NOMINAL	C
Color	INPUT	NOMINAL	C
Cylinders	INPUT	NOMINAL	C
Fuel Type	INPUT	NOMINAL	C
Location	INPUT	NOMINAL	C

Variable Type	Variable Label	Variable Format	Variable Length
varchar			20
varchar			11
varchar			7
varchar			8
varchar			15

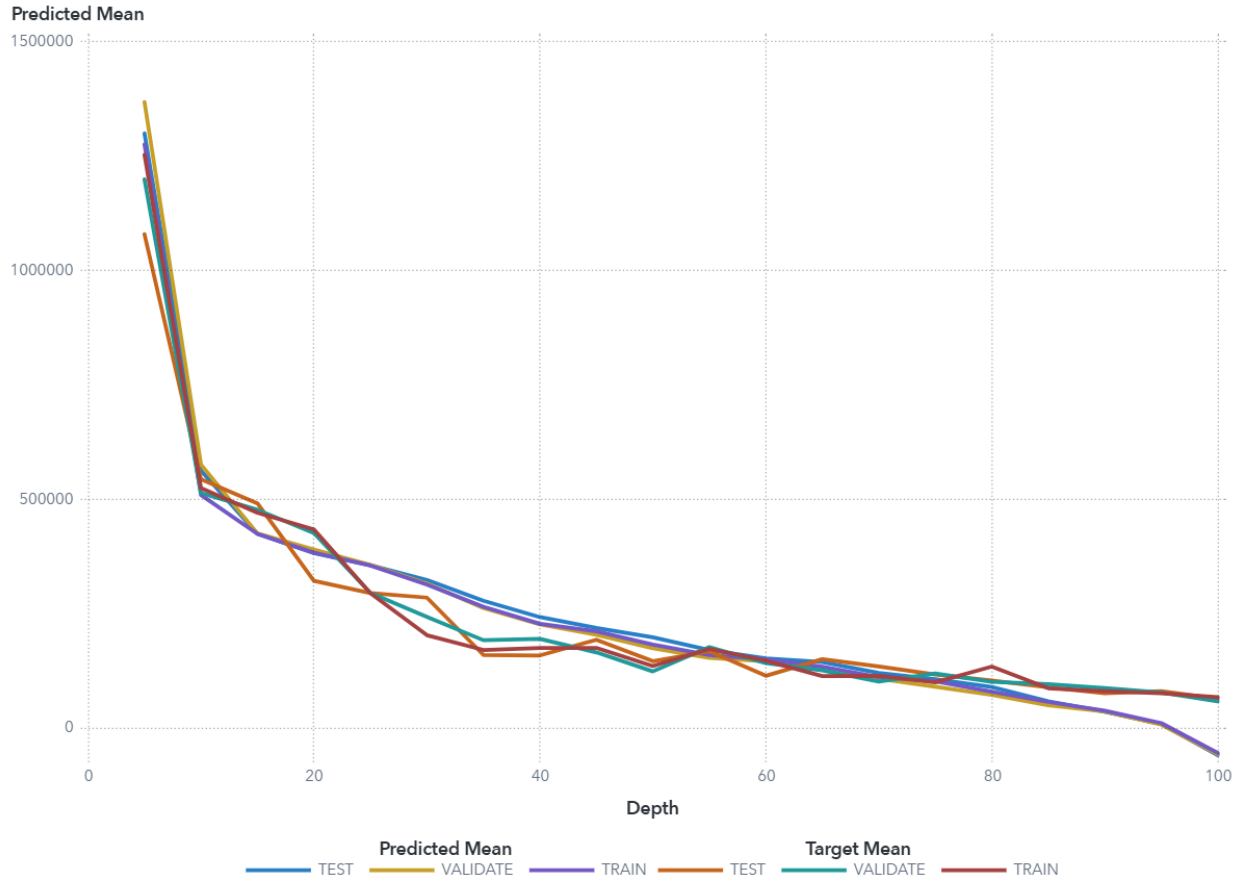
Score Outputs

Name	Role	Type	Variable Type
EM_PREDICTION	PREDICT	N	double
P_Price	PREDICT	N	double

Variable Label	Variable Format	Variable Length	Creator
Predicted: Price		8	linearreg
Predicted: Price		8	linearreg

Function	Creator GUID
PREDICT	1ddcd56a-d971-4fa3-bfec-4e6aa7f971cc
PREDICT	1ddcd56a-d971-4fa3-bfec-4e6aa7f971cc

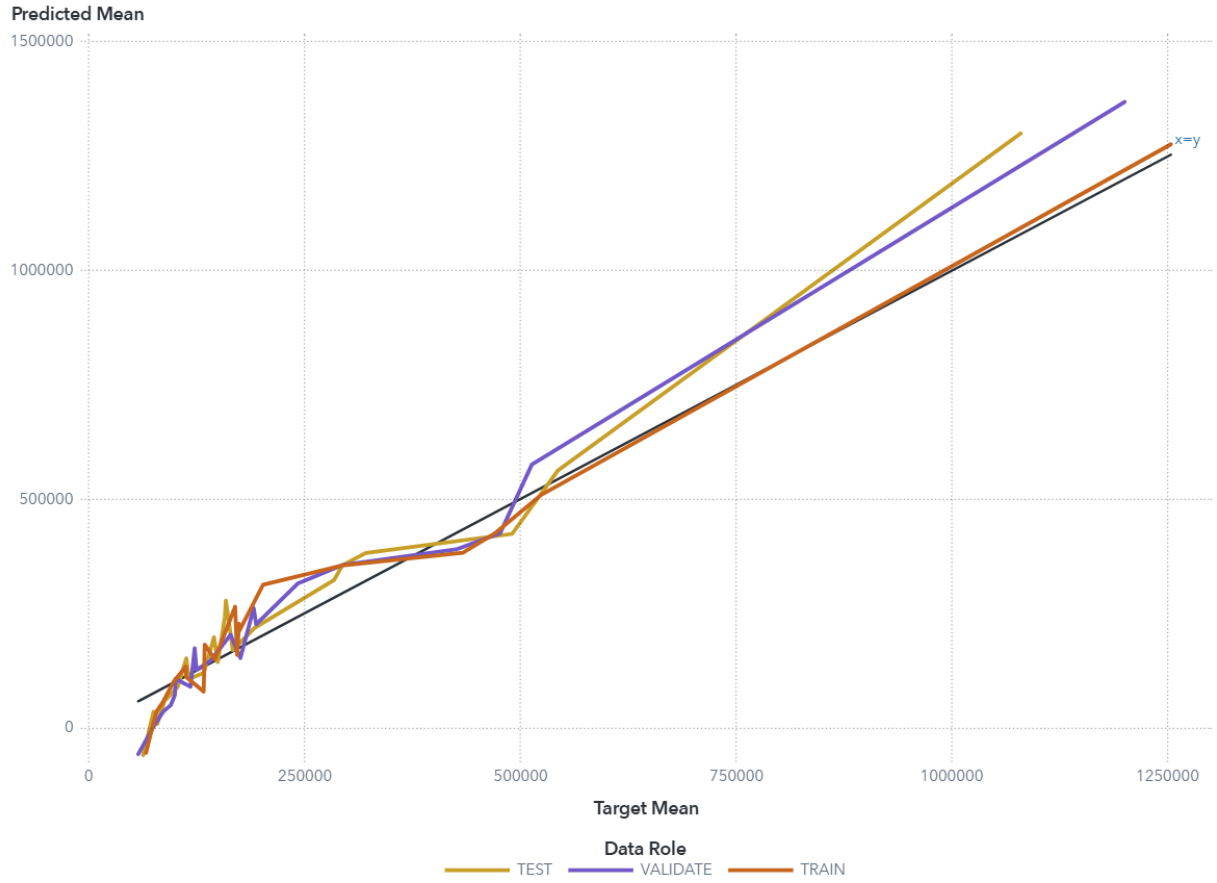
Actual and predicted by depth



For this plot, each partition of the data is sorted by the predicted target P_Price for the actual target Price, in descending order. The data is then divided into 20 quantiles (demi-deciles, with 5% of the data in each), and the mean of the predicted target and actual target are calculated and plotted for each quantile (depth in increments of 5).

The greatest difference between the actual and predicted target means is 220,387.461 and occurs for the TEST partition at depth 5.

Predicted by actual



The "Predicted by actual" plot is a different representation of the data in the "Actual and predicted by depth" plot. In this plot, the predicted mean is plotted against the target mean. Each partition of the data is sorted in descending order by the predicted target P_Price for the target Price. The data is divided into 20 quantiles (demi-deciles, with 5% of the data in each), and the means of the predicted target and actual target are calculated and plotted. The black, diagonal line on the plot indicates where the predicted mean and the actual mean are equal. Thus, the points for a perfect model would correspond to that line. Points that are plotted above the line over-predict the target and points that are plotted below the line under-predict the target.

The greatest difference between the actual and predicted target means is 220,387.461 and occurs for the TEST partition at depth 5 where the target is being under-predicted.

Fit Statistics

Target Name	Data Role	Partition Indicator	Formatted Partition
Price	VALIDATE	0	0
Price	TRAIN	1	1
Price	TEST	2	2

Number of Observations	Average Squared Error	Divisor for ASE	Root Average Squared Error
3,000	1.1891137E11	3,000	344,835.2851
6,000	1.7115686E11	6,000	413,711.0859
1,000	1.0207978E11	1,000	319,499.2587

Mean Absolute Error	Root Mean Absolute Error	Mean Squared Logarithmic Error	Root Mean Squared Logarithmic Error
194,525.2200	441.0501	1.3467	1.1605
194,748.2019	441.3028	1.3885	1.1783
194,057.2986	440.5194	1.4100	1.1875

R-Square
0.3420
0.3252
0.3372

Properties

Property Name	Property Value
chooseCriterion	SBC
chooseFcp	SBC
classCoding	GLM
codeLocation	mlearning
dataMiningVersion	V2024.09
exactPctlLift	true
explainFidelity	false
explainInfo	false
factorInteractions	false
factorSplit	false
fullDatasetReconstitution	false
hierarchy	NONE
icePlots	false
informativeMiss	false
lsCoeffs	true
maxEffects	0
maxL2	1
maxNumShapVars	20
maxSteps	0
maxTime	600
minEffects	0
minL2	0
missAsLvl	false
nBins	50
numAlpha	4
numL2	50
numLambda	10

Property Name	Property Value
optTech	MILP
parmValueAlpha	SEARCH
parmValueL2	SEARCH
parmValueLambda	SEARCH
pdNumImportantInputs	5
pdObsSamples	1,000
pdPlots	false
performKernelShap	false
performLime	false
performVI	false
polynomialDegree	2
seedId	12,345
selectCriterion	SBC
selectMethod	STEPWISE
slEntry	0.0500
slStay	0.0500
specifyRows	RANDOM
stopCriterion	SBC
suppressIntercept	false
templateRevision	3
train	true
truncateLl	5
truncateUl	95
usePolynomial	false
useSpline	false
useSplineSplit	false

Output

The SAS System

The REGSELECT Procedure

Model Information	
Data Source	DM_AM5XLJRL8K9G8SC5M37HLYPZ4
Response Variable	Price
Predicted Response	P_Price

Selection Information	
Selection Method	Stepwise
Select Criterion	SBC
Choose Criterion	SBC
Stop Criterion	SBC
Effect Hierarchy Enforced	None
Stop Horizon	3

Number of Observations Read	10000
Number of Observations Used	9920
Number of Observations Used for Training	5948
Number of Observations Used for Validation	2979
Number of Observations Used for Testing	993

Class Level Information

Class	Levels	Values
Body Type	13	Coupe Crossover Hard Top Convertible Hatchback Other Pick Up Truck SUV Sedan Soft Top Convertible Sports Car Utility Truck Van Wagon
Color	17	Beige Black Blue Brown Burgundy Gold Green Grey Orange Other Color Purple Red Silver Tan Teal White Yellow
Customer Rating	5	1 2 3 4 5
Cylinders	9	10 12 3 4 5 6 8 None Unknown
Fuel Type	4	Diesel Electric Gasoline Hybrid
Latitude	16	-84.4252913839482 -84.0960226301587 -83.0535188511359 -62.3251238214748 -52.0187954926453 -50.4499049801286 -33.3120190573922 -8.41288094534527 -5.67697732432443 24.6685596306529 37.6525266829928 46.3444291301585 50.7603286069094 60.9519233156502 71.1180052319239 73.052890199875
Location	8	Abu Dhabi Ajman Al Ain Dubai Fujeirah Ras Al Khaimah Sharjah Umm Al Qawain
Longitude	16	-133.567593146221 -121.810430237547 -107.877664405281 -81.0375630274484 -28.6792598896204 -22.0781891509024 5.84921635202175 35.3901872505667 55.7192704623337 67.7854794338118 83.4138981803345 102.222906515222 105.188377779088 136.397212528626 151.821601002598 173.754868137684
Transmission	2	Automatic Transmission Manual Transmission

Dimensions	
Number of Effects	12
Number of Parameters	93

The SAS System

The REGSELECT Procedure

Selection Details

Selection Summary

Step	Effect Entered	Number Effects In	SBC
0	Intercept	1	156247.765
1	Cylinders	2	154843.165
2	Fuel Type	3	154601.671
3	Color	4	154424.243
4	Body Type	5	154310.424
5	Location	6	154307.332*

* Optimal Value Of Criterion

Stepwise selection stopped because adding or removing an effect does not improve the SBC criterion.

The model at step 5 is selected where SBC is 154307.3.

Selected Effects: Intercept Body Type Color Cylinders Fuel Type Location

The SAS System

The REGSELECT Procedure

Selected Model

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	46	4.948188E14	1.075693E13	61.84	<.0001
Error	5901	1.02643E15	1.739416E11		
Corrected Total	5947	1.521248E15			

Root MSE	417063
R-Square	0.32527
Adj R-Sq	0.32001
AIC	159943
AICC	159944
SBC	154307
ASE (Train)	1.725672E11
ASE (Validate)	1.196691E11
ASE (Test)	1.027524E11

Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	2825785	238424	11.85	<.0001	0
Body Type Coupe	1	-22839	85634	-0.27	0.7897	20.11040
Body Type Crossover	1	-122955	98093	-1.25	0.2101	3.66463
Body Type Hard Top Convertible	1	79604	94926	0.84	0.4017	4.54161
Body Type Hatchback	1	-95173	89484	-1.06	0.2876	8.51035
Body Type Other	1	-6948.806003	97310	-0.07	0.9431	3.81905
Body Type Pick Up Truck	1	-142836	87456	-1.63	0.1025	11.17336
Body Type SUV	1	-34691	83916	-0.41	0.6793	59.81757
Body Type Sedan	1	-108529	84320	-1.29	0.1981	48.56985
Body Type Soft Top Convertible	1	85083	98355	0.87	0.3870	3.68419
Body Type Sports Car	1	366849	90063	4.07	<.0001	7.34871
Body Type Utility Truck	1	-97956	120394	-0.81	0.4159	1.99188
Body Type Van	1	-75237	95252	-0.79	0.4296	4.37033
Body Type Wagon	0	0	.	.	.	.
Color Beige	1	-858544	96218	-8.92	<.0001	1.95705
Color Black	1	-891107	68032	-13.10	<.0001	26.82024
Color Blue	1	-910320	70045	-13.00	<.0001	10.91340
Color Brown	1	-1038857	76218	-13.63	<.0001	4.43812
Color Burgundy	1	-993394	93212	-10.66	<.0001	2.08312
Color Gold	1	-1019243	75960	-13.42	<.0001	4.50305
Color Green	1	-821373	79179	-10.37	<.0001	3.43913
Color Grey	1	-886980	68601	-12.93	<.0001	18.63433
Color Orange	1	-742010	96281	-7.71	<.0001	1.90697
Color Other Color	1	-913381	94421	-9.67	<.0001	2.03640
Color Purple	1	-841998	124212	-6.78	<.0001	1.41539
Color Red	1	-927491	70145	-13.22	<.0001	10.03784
Color Silver	1	-965170	69747	-13.84	<.0001	12.50475
Color Tan	1	-903428	199013	-4.54	<.0001	1.13753
Color Teal	1	-954273	183297	-5.21	<.0001	1.15777
Color White	1	-956199	67754	-14.11	<.0001	34.94804
Color Yellow	0	0	.	.	.	.
Cylinders 10	1	221711	113495	1.95	0.0508	2.13709
Cylinders 12	1	1005597	86543	11.62	<.0001	7.55637
Cylinders 3	1	-196319	106753	-1.84	0.0660	2.34434
Cylinders 4	1	-200305	81107	-2.47	0.0136	47.03539
Cylinders 5	1	-100067	107965	-0.93	0.3540	2.26544
Cylinders 6	1	-159446	81096	-1.97	0.0493	50.58144
Cylinders 8	1	47060	81220	0.58	0.5623	47.91792
Cylinders None	1	91557	155664	0.59	0.5564	1.80704
Cylinders Unknown	0	0	.	.	.	.
Fuel Type Diesel	1	-1562619	110970	-14.08	<.0001	6.54963
Fuel Type Electric	1	-1631767	175531	-9.30	<.0001	2.82654
Fuel Type Gasoline	1	-1604438	101981	-15.73	<.0001	7.43126
Fuel Type Hybrid	0	0	.	.	.	.
Location Abu Dhabi	1	-20984	171891	-0.12	0.9028	72.23811
Location Ajman	1	-26601	175707	-0.15	0.8797	17.62215
Location Al Ain	1	-6431.463117	184906	-0.03	0.9723	6.83915
Location Dubai	1	80852	170888	0.47	0.6361	161.81811
Location Fujeirah	1	-20305	232568	-0.09	0.9304	2.17412
Location Ras Al Khaimah	1	-80738	215881	-0.37	0.7084	2.67481
Location Sharjah	1	-34626	171626	-0.20	0.8401	90.00047
Location Umm Al Qawain	0	0	.	.	.	.

Task Timing

Task	Seconds	Percent
Setup and Parsing	0.01	9.08%
Levelization	0.01	7.65%
Model Initialization	0.00	0.30%
SSCP Computation	0.10	74.04%
Model Selection	0.00	2.48%
Producing Score Code	0.00	0.73%
Cleanup	0.01	5.65%
Total	0.14	100.00%