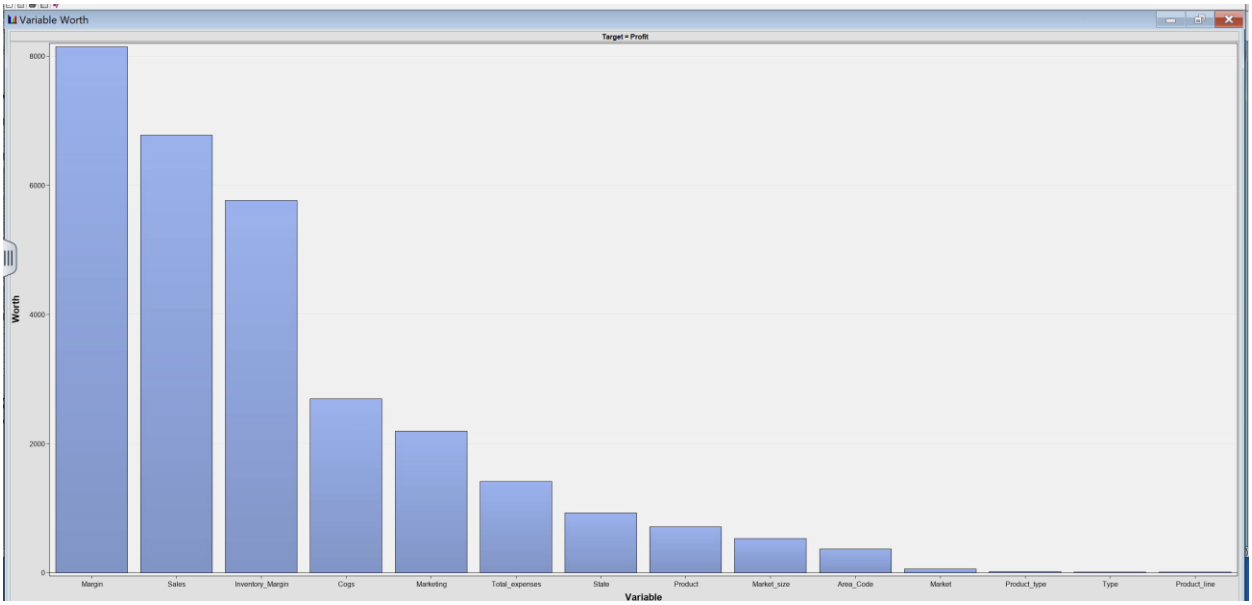
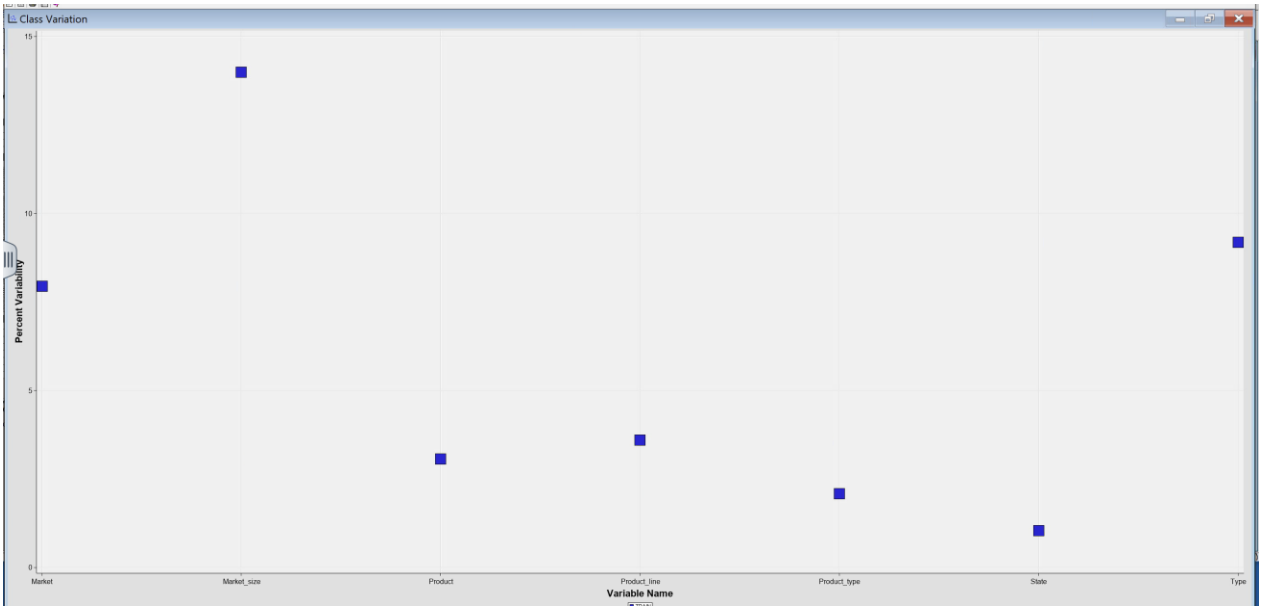
APPENDIX:

Stat Explorer: Variable Worth of Variables:



Class Variation:

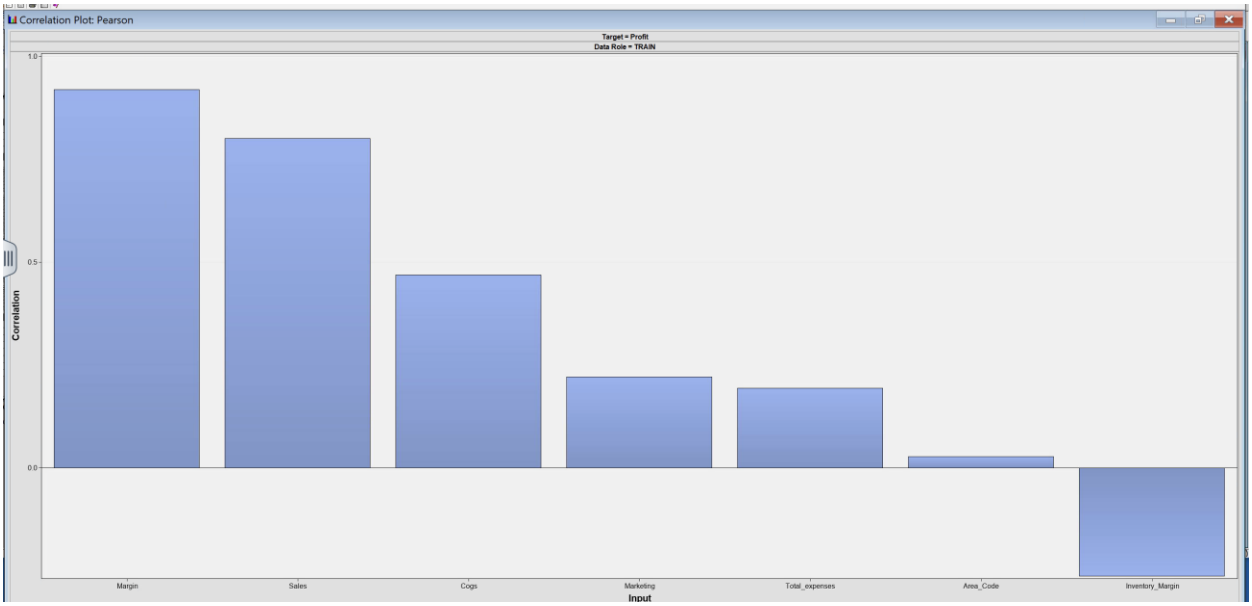


Correlation Plot:

Correlation Statistics
(maximum 500 observations printed)

Data Role=TRAIN Type=PEARSON Target=Profit

Input	Correlation
Margin	0.91855
Sales	0.79993
Cogs	0.46935
Marketing	0.22133
Total_expenses	0.19366
Area_Code	0.02721
Inventory_Margin	-0.26216



There are varying degrees of correlation between the variables and the target variable (Profit). Margin and Sales show strong positive correlations, indicating a positive relationship with Profit. Cogs, Marketing, and Total Expenses show weaker positive correlations. The Area Code has a very weak positive correlation, and Inventory Margin has a moderate negative correlation with Profit.

Variable Clustering:

Variable Selection Table

Cluster	Variable	Label	R-Square With Own Cluster Component	Next Closest Cluster	R-Square with Next Cluster Component	Type	1-R2 Ratio	Variable Selected
CLUS1	CLUS1	Cluster 1	1	CLUS3	0.133794	ClusterComp	0	YES
CLUS1	SALES	Sales	0.885412	CLUS3	0.042741	Variable	0.119704	NO
CLUS1	COGS	Cogs	0.873647	CLUS3	0.330619	Variable	0.188519	NO
CLUS1	MARKETING	Marketing	0.810298	CLUS3	0.233701	Variable	0.247556	NO
CLUS1	TOTAL_EXPENSES	Total expenses	0.780156	CLUS3	0.183775	Variable	0.269343	NO
CLUS1	MARGIN	Margin	0.662108	CLUS3	0.009324	Variable	0.341077	NO
CLUS2	CLUS2	Cluster 2	1	CLUS3	0.006117	ClusterComp	0	YES
CLUS2	AREA_CODE		1	CLUS3	0.006117	Variable	0	NO
CLUS3	CLUS3	Cluster 3	1	CLUS1	0.133794	ClusterComp	0	YES
CLUS3	INVENTORY_MARGIN		1	CLUS1	0.133794	Variable	0	NO

Variable Summary

Role	Measurement Level	Frequency Count
INPUT	INTERVAL	7
INPUT	NOMINAL	7
REJECTED	INTERVAL	5
TARGET	INTERVAL	1
TIMEID	INTERVAL	1

Oblique Principal Component Cluster Analysis

Observations	1062	Proportion	0
Variables	7	Maxeigen	1

Clustering algorithm converged.

Cluster Summary for 1 Cluster

Cluster	Members	Cluster Variation	Variation Explained	Proportion Explained	Second Eigenvalue
1	7	7	4.193011	0.5990	1.2571

Total variation explained = 4.193011 Proportion = 0.5990

Cluster 1 will be split because it has the largest second eigenvalue, 1.25706, which is greater than the MAXEIGEN=1 value.

Clustering algorithm converged.

Cluster Summary for 2 Clusters

Cluster	Members	Cluster Variation	Variation Explained	Proportion Explained	Second Eigenvalue
1	6	6	4.184371	0.6974	1.2528
2	1	1	1	1.0000	

Total variation explained = 5.184371 Proportion = 0.7406

Cluster 1 will be split because it has the largest second eigenvalue, 1.25706, which is greater than the MAXEIGEN=1 value.

Clustering algorithm converged.

Cluster Summary for 2 Clusters

Cluster	Members	Cluster Variation	Variation Explained	Proportion Explained	Second Eigenvalue
1	6	6	4.184371	0.6974	1.2528
2	1	1	1	1.0000	

Total variation explained = 5.184371 Proportion = 0.7406

2 Clusters		R-squared with Own Cluster	Next Closest	1-R**2 Ratio	Variable Label
Cluster 1	Cogs	0.9146	0.0119	0.0865	Cogs
	Inventory_Margin	0.2197	0.0061	0.7851	
	Margin	0.5729	0.0020	0.4279	Margin
	Marketing	0.8396	0.0036	0.1610	Marketing
	Sales	0.8382	0.0059	0.1627	Sales
	Total_expenses	0.7993	0.0020	0.2011	Total_expenses
Cluster 2	Area_Code	1.0000	0.0066	0.0000	

Cluster 1 will be split because it has the largest second eigenvalue, 1.252813, which is greater than the MAXEIGEN=1 value.

Clustering algorithm converged.

Cluster Summary for 3 Clusters

Cluster	Members	Cluster Variation	Variation Explained	Proportion Explained	Second Eigenvalue
1	5	5	4.011821	0.8024	0.7591
2	1	1	1	1.0000	
3	1	1	1	1.0000	

Total variation explained = 6.011821 Proportion = 0.8588

3 Clusters		R-squared with Own Cluster	Next Closest	1-R**2 Ratio	Variable Label
Cluster 1	Cogs	0.8738	0.3308	0.1885	Cogs
	Margin	0.6621	0.0093	0.3411	Margin
	Marketing	0.8103	0.2337	0.2476	Marketing
	Sales	0.8854	0.0427	0.1197	Sales
	Total_expenses	0.7802	0.1838	0.2693	Total_expenses
Cluster 2	Area_Code	1.0000	0.0061	0.0000	
Cluster 3	Inventory_Margin	1.0000	0.1338	0.0000	

No cluster meets the criterion for splitting.

Number of Clusters	Total Variation Explained by Clusters	Proportion of Variation Explained by Clusters	Minimum Proportion Explained by a Cluster	Maximum Second Eigenvalue in a Cluster	Minimum R-squared for a Variable	Maximum 1-R**2 Ratio for a Variable
1	4.193011	0.5990	0.5990	1.257060	0.0113	
2	5.184371	0.7406	0.6974	1.252813	0.2197	0.7851
3	6.011821	0.8588	0.8024	0.759074	0.6621	0.3411

Backward Regression Output:

utput

Intercept LG10_Inventory_Margin LG10_Margin LG10_Marketing LG10_Sales LG10_Total_expenses OPT_Area_Code Product State

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	38	7613851	200365	132.57	<.0001
Error	811	1225764	1511.422525		
Corrected Total	849	8839615			

Model Fit Statistics

R-Square	0.8613	Adj R-Sq	0.8548
AIC	6260.7625	BIC	6266.6001
SBC	6445.8267	C(p)	38.0928

Type 3 Analysis of Effects

Effect	DF	Sum of Squares	F Value	Pr > F
LG10_Inventory_Margin	1	122264.762	80.89	<.0001
LG10_Margin	1	165352.601	109.40	<.0001
LG10_Marketing	1	68148.7216	45.09	<.0001
LG10_Sales	1	1513104.99	1001.11	<.0001
LG10_Total_expenses	1	61100.0740	40.43	<.0001
OPT_Area_Code	2	26593.7779	8.80	0.0002
Product	12	113213.315	6.24	<.0001
State	19	124093.647	4.32	<.0001

Analysis of Maximum Likelihood Estimates

Parameter		DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept		1	-531.6	42.7302	-12.44	<.0001
LG10_Inventory_Margin		1	-68.4992	7.6160	-8.99	<.0001
LG10_Margin		1	133.2	12.7336	10.46	<.0001
LG10_Marketing		1	-90.0077	13.4043	-6.71	<.0001
LG10_Sales		1	355.8	11.2450	31.64	<.0001
LG10_Total_expenses		1	-107.0	16.8266	-6.36	<.0001
OPT_Area_Code	01:low-906.5, MISSING	1	25.3850	6.4026	3.96	<.0001
OPT_Area_Code	02:906.5-914.5	1	-52.4470	12.5075	-4.19	<.0001
Product	Amaretto	1	-8.0500	6.3168	-1.27	0.2029
Product	Caffe Latte	1	-2.7522	6.2071	-0.44	0.6576
Product	Caffe Mocha	1	-3.1564	4.0231	-0.78	0.4329
Product	Chamomile	1	-0.0212	4.6496	-0.00	0.9964
Product	Colombian	1	5.7748	4.2186	1.37	0.1714
Product	Darjeeling	1	-1.4262	4.4015	-0.32	0.7460
Product	Decaf Espresso	1	-8.6772	4.4336	-1.96	0.0507
Product	Decaf Irish Cream	1	-8.5762	4.5620	-1.88	0.0605
Product	Earl Grey	1	-4.6376	4.9675	-0.93	0.3508
Product	Green Tea	1	-0.1958	5.2493	-0.04	0.9703
Product	Lemon	1	-2.8791	3.9296	-0.73	0.4640
Product	Mint	1	-34.6301	5.9749	-5.80	<.0001
State	California	1	-4.6431	5.2010	-0.89	0.3723
State	Colorado	1	-9.0540	5.3315	-1.70	0.0898
State	Connecticut	1	-4.3552	6.6575	-0.65	0.5132
State	Florida	1	-5.5240	5.8788	-0.94	0.3477
State	Illinois	1	5.5390	6.0801	0.91	0.3626
State	Iowa	1	32.3425	6.0288	5.36	<.0001
State	Louisiana	1	-12.3670	7.0230	-1.76	0.0786
State	Massachusetts	1	-12.6037	7.7804	-1.62	0.1056
State	Missouri	1	2.4970	6.0393	0.41	0.6794
State	Nevada	1	22.2184	5.4264	4.09	<.0001
State	New Hampshire	1	1.8385	7.2852	0.25	0.8008
State	New Mexico	1	15.3477	7.4850	2.05	0.0406
State	New York	1	26.4833	6.6867	3.96	<.0001
State	Ohio	1	-6.2477	5.5335	-1.13	0.2592
State	Oklahoma	1	-12.9426	6.8728	-1.88	0.0600
State	Oregon	1	-7.7523	5.4675	-1.42	0.1566
State	Texas	1	-9.2347	7.4147	-1.25	0.2133
State	Utah	1	-3.9295	4.9970	-0.79	0.4319

Fit Statistics:

Fit Statistics

Target	Target Label	Fit Statistics	Statistics Label	Train	Validation
Profit	Profit	AIC	Akaike's Information Criterion	6260.763	
Profit	Profit	ASE	Average Squared Error	1442.075	1098.042
Profit	Profit	AVERR	Average Error Function	1442.075	1098.042
Profit	Profit	DFE	Degrees of Freedom for Error	811	
Profit	Profit	DFM	Model Degrees of Freedom	39	.
Profit	Profit	DFT	Total Degrees of Freedom	850	
Profit	Profit	DIV	Divisor for ASE	850	212
Profit	Profit	ERR	Error Function	1225764	232784.9
Profit	Profit	FPE	Final Prediction Error	1580.77	
Profit	Profit	MAX	Maximum Absolute Error	258.6106	137.4439
Profit	Profit	MSE	Mean Square Error	1511.423	1098.042
Profit	Profit	NOBS	Sum of Frequencies	850	212
Profit	Profit	NW	Number of Estimate Weights	39	.
Profit	Profit	RASE	Root Average Sum of Squares	37.97466	33.13672
Profit	Profit	RFPE	Root Final Prediction Error	39.7589	
Profit	Profit	RMSE	Root Mean Squared Error	38.87702	33.13672
Profit	Profit	SBC	Schwarz's Bayesian Criterion	6445.827	
Profit	Profit	SSE	Sum of Squared Errors	1225764	232784.9
Profit	Profit	SUMW	Sum of Case Weights Times Freq	850	212

Model Comparison:

Looking at ASE to draw comparisons between different models:

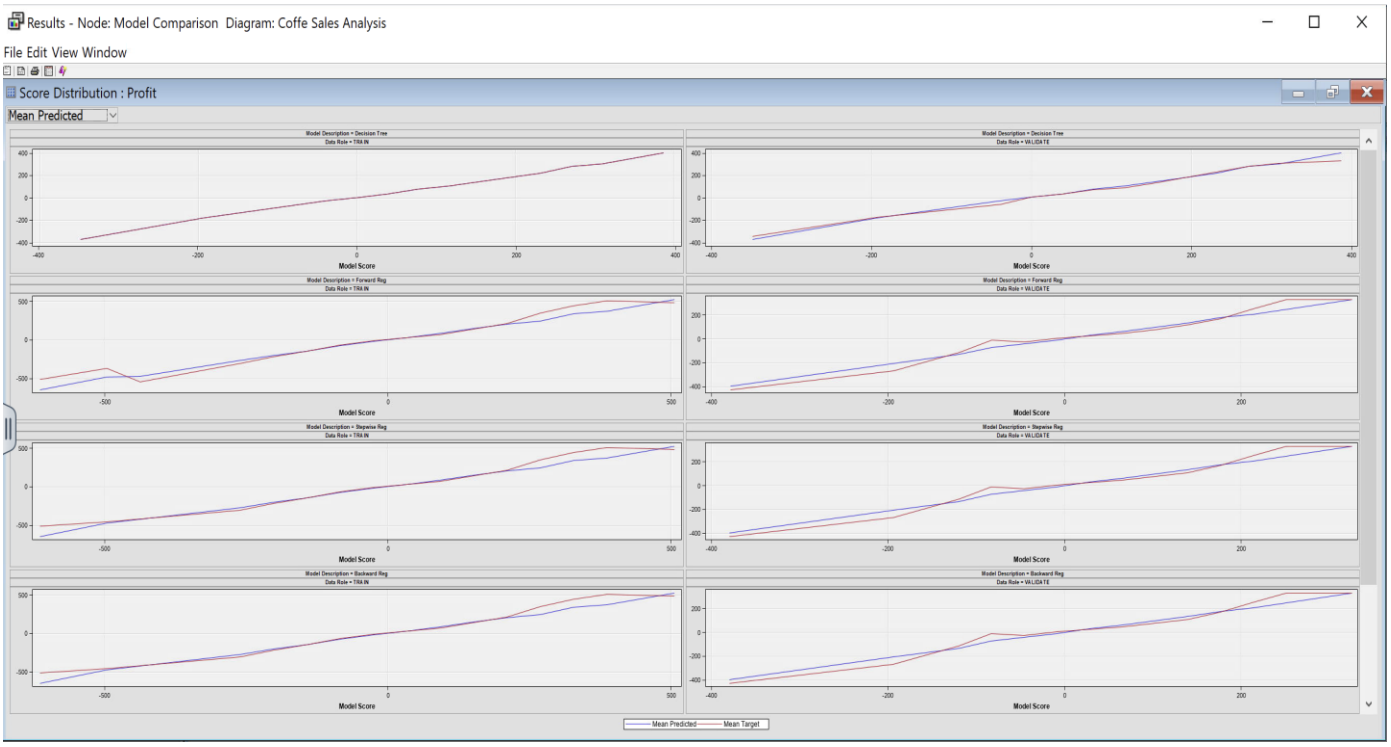
Results - Node: Model Comparison Diagram: Coffe Sales Analysis

File Edit View Window

Fit Statistics

Selected Model	Predecessor Node	Model Node	Model Description	Target Variable	Target Label	Selection Criterion: Valid: Average Squared Error	Train: Sum of Frequencies	Train: Maximum Absolute Error	Train: Sum of Squared Errors	Train: Average Squared Error	Valid: Average Squared Error	Train: Root Average Squared Error	Train: Divisor for ASE	Train: Total Degrees of Freedom	Valid: Sum of Frequencies	Valid: Maximum Absolute Error	Valid: Sum of Squared Errors	Valid: Root Average Squared Error	Valid: Divisor for VASE	Train: Degrees of Freedom for Error	Train: Model Degrees of Freedom	Train: Number of Estimated Weights	Train: Akaike's Information Criterion	Train: Schwarz's Bayesian Criterion	Train: Sum of Case Weights Times Freq	Train: Final Prediction Error	Train: Mean Squared Error	Train: Root Final Prediction Error	Train: Root Mean Squared Error	Train: Average Error Function
Y	Tree	Tree	Decision Tree	Profit	Profit	695...	850	23...	6998...	823.3666	695.4908	28.6...	850	850	212	165...	147...	26.3...	212											
	Rea	Rea	Backward Rea	Profit	Profit	1098...	850	258...	122...	1442.075	1098.042	37.9...	850	850	212	137...	2327...	33.1...	212	811	39	396260	6445...	850	158...	1511...	39.7	38.8...	1442...
	Rea2	Rea2	Stepwise Rea	Profit	Profit	1098...	850	258...	122...	1442.075	1098.042	37.9...	850	850	212	137...	2327...	33.1...	212	811	39	396260	6445...	850	158...	1511...	39.7	38.8...	1442...
	Rea3	Rea3	Forward Rea	Profit	Profit	1099...	850	258...	122...	1441.91	1099.003	37.9...	850	850	212	137...	2329...	33.1...	212	810	40	406262	6452...	850	1584...	1513...	39.8	38.8...	144...
	Neural	Neural	Neural Network	Profit	Profit	1753...	850	484...	186...	2195.537	1753.588	46.8...	850	850	212	384...	3717...	41.8...	212	705	145	1456830	7518...	850	3098...	2647...	55.6	51.4...	2195...

Score Distribution:



Score Rankings: