The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: Recommend

15 Variables:	Recommend Prod_Qual Ecommerce Tech_support Complaint Adv Prod_Line Sales_Image Pricing Warranty New_Prod Ordering Price_Flex Del_Speed Industry_type

Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
Recommend	100	7.02000	1.04330	702.00000	4.60000	9.90000
Prod_Qual	100	7.81000	1.39628	781.00000	5.00000	10.00000
Ecommerce	100	3.67200	0.70052	367.20000	2.20000	5.70000
Tech_support	100	5.36500	1.53046	536.50000	1.30000	8.50000
Complaint	100	5.44200	1.20840	544.20000	2.60000	7.80000
Adv	100	4.01000	1.12694	401.00000	1.90000	6.50000
Prod_Line	100	5.80500	1.31529	580.50000	2.30000	8.40000
Sales_Image	100	5.12300	1.07232	512.30000	2.90000	8.20000
Pricing	100	6.97400	1.54506	697.40000	3.70000	9.90000
Warranty	100	6.04300	0.81974	604.30000	4.10000	8.10000
New_Prod	100	5.15000	1.49305	515.00000	1.70000	9.50000
Ordering	100	4.27800	0.92884	427.80000	2.00000	6.70000
Price_Flex	100	4.61000	1.20600	461.00000	2.60000	7.30000
Del_Speed	100	3.88600	0.73444	388.60000	1.60000	5.50000
Industry_type	100	0.48000	0.50212	48.00000	0	1.00000

Descriptive statistics for the metric independent variable Recommend, 13 metric dependent variables, and the non-metric variable Industry_type will be used for the purpose of building a multiple regression model.

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: Recommend

Pearson Correlation Coefficients, N = 100 Prob > |r| under H0: Rho=0							
	Recommend	**Prod_Qual**	**Ecommerce**	**Tech_support**	**Complaint**	**Adv**	**Prod_Line**
Recommend	1.00000	0.44717 <.0001	0.22191 0.0265	0.18807 0.0610	0.45064 <.0001	0.17191 0.0872	0.44526 <.0001
Prod_Qual	0.44717 <.0001	1.00000	-0.13716 0.1736	0.09560 0.3441	0.10637 0.2922	-0.05347 0.5972	0.47749 <.0001
Ecommerce	0.22191 0.0265	-0.13716 0.1736	1.00000	0.00087 0.9932	0.14018 0.1642	0.42989 <.0001	-0.05269 0.6026
Tech_support	0.18807 0.0610	0.09560 0.3441	0.00087 0.9932	1.00000	0.09666 0.3387	-0.06287 0.5343	0.19263 0.0549
Complaint	0.45064 <.0001	0.10637 0.2922	0.14018 0.1642	0.09666 0.3387	1.00000	0.19692 0.0496	0.56142 <.0001
Adv	0.17191 0.0872	-0.05347 0.5972	0.42989 <.0001	-0.06287 0.5343	0.19692 0.0496	1.00000	-0.01155 0.9092
Prod_Line	0.44526 <.0001	0.47749 <.0001	-0.05269 0.6026	0.19263 0.0549	0.56142 <.0001	-0.01155 0.9092	1.00000
Sales_Image	0.38250 <.0001	-0.15181 0.1316	0.79154 <.0001	0.01699 0.8668	0.22975 0.0215	0.54220 <.0001	-0.06132 0.5445
Pricing	-0.14543 0.1488	-0.40128 <.0001	0.22946 0.0216	-0.27079 0.0064	-0.12795 0.2046	0.13422 0.1831	-0.49495 <.0001
Warranty	0.15075 0.1344	0.08831 0.3823	0.05190 0.6081	0.79717 <.0001	0.14041 0.1635	0.01079 0.9151	0.27308 0.0060
New_Prod	-0.00169 0.9867	0.02699 0.7898	-0.02743 0.7865	-0.07358 0.4669	0.05940 0.5572	0.08417 0.4051	0.04616 0.6483
Ordering	0.43866 <.0001	0.10430 0.3017	0.15615 0.1208	0.08010 0.4282	0.75687 <.0001	0.18424 0.0665	0.42441 <.0001
Price_Flex	0.01469 0.8847	-0.49314 <.0001	0.27067 0.0065	-0.18610 0.0638	0.39451 <.0001	0.33355 0.0007	-0.37797 0.0001
Del_Speed	0.44620 <.0001	0.02772 0.7843	0.19164 0.0561	0.02544 0.8016	0.86509 <.0001	0.27586 0.0055	0.60185 <.0001
Industry_type	0.00656 0.9484	-0.11065 0.2731	0.10177 0.3137	0.00105 0.9917	-0.03689 0.7156	0.00571 0.9550	-0.09544 0.3449

Pearson Correlation Coefficients, N = 100 Prob > |r| under H0: Rho=0								
	Sales_Image	**Pricing**	**Warranty**	**New_Prod**	**Ordering**	**Price_Flex**	**Del_Speed**	**Industry_type**
Recommend	0.38250 <.0001	-0.14543 0.1488	0.15075 0.1344	-0.00169 0.9867	0.43866 <.0001	0.01469 0.8847	0.44620 <.0001	0.00656 0.9484
Prod_Qual	-0.15181 0.1316	-0.40128 <.0001	0.08831 0.3823	0.02699 0.7898	0.10430 0.3017	-0.49314 <.0001	0.02772 0.7843	-0.11065 0.2731

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: Recommend

Pearson Correlation Coefficients, N = 100 Prob > |r| under H0: Rho=0								
	Sales_Image	**Pricing**	**Warranty**	**New_Prod**	**Ordering**	**Price_Flex**	**Del_Speed**	**Industry_type**
Ecommerce	0.79154 <.0001	0.22946 0.0216	0.05190 0.6081	-0.02743 0.7865	0.15615 0.1208	0.27067 0.0065	0.19164 0.0561	0.10177 0.3137
Tech_support	0.01699 0.8668	-0.27079 0.0064	0.79717 <.0001	-0.07358 0.4669	0.08010 0.4282	-0.18610 0.0638	0.02544 0.8016	0.00105 0.9917
Complaint	0.22975 0.0215	-0.12795 0.2046	0.14041 0.1635	0.05940 0.5572	0.75687 <.0001	0.39451 <.0001	0.86509 <.0001	-0.03689 0.7156
Adv	0.54220 <.0001	0.13422 0.1831	0.01079 0.9151	0.08417 0.4051	0.18424 0.0665	0.33355 0.0007	0.27586 0.0055	0.00571 0.9550
Prod_Line	-0.06132 0.5445	-0.49495 <.0001	0.27308 0.0060	0.04616 0.6483	0.42441 <.0001	-0.37797 0.0001	0.60185 <.0001	-0.09544 0.3449
Sales_Image	1.00000	0.26460 0.0078	0.10746 0.2873	0.03164 0.7547	0.19513 0.0517	0.35224 0.0003	0.27155 0.0063	-0.01696 0.8670
Pricing	0.26460 0.0078	1.00000	-0.24499 0.0140	0.02316 0.8191	-0.11457 0.2564	0.47111 <.0001	-0.07287 0.4712	0.14124 0.1610
Warranty	0.10746 0.2873	-0.24499 0.0140	1.00000	0.03520 0.7281	0.19707 0.0494	-0.17025 0.0904	0.10939 0.2786	-0.02120 0.8341
New_Prod	0.03164 0.7547	0.02316 0.8191	0.03520 0.7281	1.00000	0.06854 0.4980	0.09413 0.3516	0.10575 0.2950	-0.24926 0.0124
Ordering	0.19513 0.0517	-0.11457 0.2564	0.19707 0.0494	0.06854 0.4980	1.00000	0.40697 <.0001	0.75100 <.0001	-0.08109 0.4226
Price_Flex	0.35224 0.0003	0.47111 <.0001	-0.17025 0.0904	0.09413 0.3516	0.40697 <.0001	1.00000	0.49669 <.0001	0.05872 0.5617
Del_Speed	0.27155 0.0063	-0.07287 0.4712	0.10939 0.2786	0.10575 0.2950	0.75100 <.0001	0.49669 <.0001	1.00000	-0.03638 0.7194
Industry_type	-0.01696 0.8670	0.14124 0.1610	-0.02120 0.8341	-0.24926 0.0124	-0.08109 0.4226	0.05872 0.5617	-0.03638 0.7194	1.00000

Correlation matrix

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: Recommend

Number of Observations Read	100
Number of Observations Used	100

Research Objective: To predict or understand the relationship between several metric independent variables and the dependent variable 'Recommend'.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	14	64.16104	4.58293	8.93	<.0001
Error	85	43.59896	0.51293		
Corrected Total	99	107.76000			

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$;
H_1: not all $\beta_j = 0$

$F_{14,85}$=8.93

P-Value <0.0001. Reject is Null.

At least one independent variable helps to explain recommend.

Root MSE	0.71619	**R-Square**	0.5954
Dependent Mean	7.02000	**Adj R-Sq**	0.5288
Coeff Var	10.20214		

H_0: $\beta_k = 0$;
H_1: $\beta_k \neq 0$

Where k = each independent variable.

At α=0.10, Ecommerce, Tech_support, Sales_Image, Warranty, and Ordering are significant.

VIF

Prod_Line, Price_Flex, and Del_Speed have VIF values greater than 10 indicating high multicollinearity and are highly correlated with each other. They may be problematic in the model.

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	1	1.11028	1.42016	0.78	0.4365	0
Prod_Qual	1	0.32424	0.06667	4.86	<.0001	1.67276
Ecommerce	1	-0.37695	0.17573	-2.14	0.0348	2.92502
Tech_support	1	0.21655	0.08212	2.64	0.0099	3.04869
Complaint	1	-0.07202	0.13119	-0.55	0.5845	4.85066
Adv	1	-0.10256	0.07952	-1.29	0.2007	1.55020
Prod_Line	1	0.21527	0.33768	0.64	0.5255	38.07313
Sales_Image	1	0.65173	0.13073	4.99	<.0001	3.79302
Pricing	1	0.01782	0.06136	0.29	0.7721	1.73447
Warranty	1	-0.36242	0.15901	-2.28	0.0252	3.27946
New_Prod	1	-0.00570	0.05150	-0.11	0.9121	1.14102
Ordering	1	0.27729	0.13290	2.09	0.0399	2.94089
Price_Flex	1	0.09543	0.34484	0.28	0.7827	33.38133
Del_Speed	1	0.04277	0.65101	0.07	0.9478	44.12323
Industry_type	1	0.24445	0.15371	1.59	0.1155	1.14972

Measures of fit

The **R^2** value of 0.5954 suggests that the regression model explains a good amount of the variability in the dependent variable.

The **adjusted R^2** value of 0.5288 is lower than the **R^2** value of 0.5954 suggesting that some of the predictors may not be contributing much to explain the variance in the dependent variable.

The **Root MSE** value suggests that the regression model's predictions are off by 0.71619 which is relatively low, indicating that predictions are close to the actual values.

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: Recommend

Number of Observations Read	100
Number of Observations Used	100

Removing a variable to improve VIF

Removing Del_Speed from the model because it has the highest VIF.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	13	64.15883	4.93529	9.73	<.0001
Error	86	43.60117	0.50699		
Corrected Total	99	107.76000			

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$;
H_1: not all $\beta_j = 0$

$F_{13,86}$=9.73

P-Value <0.0001. Reject is Null.

At least one independent variable helps to explain recommend.

Root MSE	0.71203	**R-Square**	0.5954
Dependent Mean	7.02000	**Adj R-Sq**	0.5342
Coeff Var	10.14291		

H_0: $\beta_k = 0$; Where k = each independent variable.
H_1: $\beta_k \neq 0$

At α=0.10, Ecommerce, Tech_support, Prod_Line, Sales_Image, Warranty, and Ordering are significant.

VIF

All VIF values are below 10, next step is to run variable selection methods. Stepwise and c(P) selection methods will be used.

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	1	1.05605	1.14892	0.92	0.3606	0
Prod_Qual	1	0.32418	0.06628	4.89	<.0001	1.67253
Ecommerce	1	-0.37860	0.17291	-2.19	0.0313	2.86489
Tech_support	1	0.21670	0.08161	2.66	0.0094	3.04616
Complaint	1	-0.07135	0.13003	-0.55	0.5846	4.82129
Adv	1	-0.10291	0.07889	-1.30	0.1956	1.54343
Prod_Line	1	0.23615	0.11345	2.08	0.0403	4.34768
Sales_Image	1	0.65395	0.12554	5.21	<.0001	3.53848
Pricing	1	0.01729	0.06047	0.29	0.7756	1.70451
Warranty	1	-0.36318	0.15767	-2.30	0.0237	3.26204
New_Prod	1	-0.00596	0.05105	-0.12	0.9074	1.13464
Ordering	1	0.27828	0.13127	2.12	0.0369	2.90318
Price_Flex	1	0.11633	0.13247	0.88	0.3823	4.98363
Industry_type	1	0.24498	0.15261	1.61	0.1121	1.14660

Measures of fit

The **R^2** value of 0.5954 suggests that the regression model explains a good amount of the variability in the dependent variable.

The **adjusted R^2** value of 0.5342 improved with removal of Del_Speed.

The **Root MSE** value suggests that the regression model's predictions are off by 0.71619 which is relatively low, indicating that predictions are close to the actual values.

The SAS System

The REG Procedure
Model: MODEL1
Dependent Variable: Recommend

Summary of Stepwise Selection								
Step	**Variable Entered**	**Variable Removed**	**Number Vars In**	**Partial R-Square**	**Model R-Square**	**C(p)**	**F Value**	**Pr > F**
1	Complaint		1	0.2031	0.2031	71.4230	24.97	<.0001
2	Prod_Qual		2	0.1612	0.3643	39.5541	24.60	<.0001
3	Sales_Image		3	0.1339	0.4982	13.4310	25.61	<.0001
4	Ordering		4	0.0144	0.5125	12.4116	2.80	0.0975
5	Ecommerce		5	0.0118	0.5243	11.9363	2.33	0.1304
6		Complaint	4	0.0102	0.5141	12.0854	2.02	0.1584
7	Prod_Line		5	0.0196	0.5337	9.9596	3.96	0.0495
8	Industry_type		6	0.0144	0.5481	8.9444	2.95	0.0890

The stepwise selection summary yields a model with C(p) values greater than the number of variables. The next step will be to use an adjusted R-squared method for variable selection to improve the model.

Number in Model	Adjusted R-Square	R-Square	Variables in Model
9	0.5496	0.5905	Prod_Qual Ecommerce Tech_support Adv Prod_Line Sales_Image Warranty Ordering Industry_type
10	0.5478	0.5935	Prod_Qual Ecommerce Tech_support Adv Prod_Line Sales_Image Warranty Ordering Price_Flex Industry_type
8	0.5478	0.5843	Prod_Qual Ecommerce Tech_support Prod_Line Sales_Image Warranty Ordering Industry_type

The first model with 9 variables is selected because it has the highest adjusted R-squared value of 0.5496 and an R-square of 0.5905 because approximately 54.96% of the variance in the dependent variable is explained by the independent variables included in the model, after adjusting for the number of predictors and the sample size. This is a moderate level of explanatory power, indicating that the model captures a substantial portion of the variability in the dependent variable, but there is still room for improvement.

The next step is to run the model with the selected variables **Prod_Qual, Ecommerce, Tech_support, Adv, Prod_Line, Sales,_Image, Warranty, Ordering, Industry_type.**

The SAS System

The REG Procedure
Model: MODEL2
Dependent Variable: Recommend

Number of Observations Read	100
Number of Observations Used	100

Reduced model with the highest adj R-squared method, including 9 suggested variables. Option 1.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	9	63.63228	7.07025	14.42	<.0001
Error	90	44.12772	0.49031		
Corrected Total	99	107.76000			

H₀: $\beta_1 = \beta_2 = \beta_3 = 0$;
H₁: not all $\beta_j = 0$

$F_{9,90}$=14.42

P-Value <0.0001. Reject is Null.

At least one independent variable helps to explain recommend.

Root MSE	0.70022	**R-Square**	0.5905
Dependent Mean	7.02000	**Adj R-Sq**	0.5496
Coeff Var	9.97464		

H₀: $\beta_k = 0$; Where k = each independent variable.
H₁: $\beta_k \neq 0$

At α=0.10, Prod_Qual, Ecommerce, Tech_support, Prod_Line, Sales_Image, Warranty, Ordering, and Industry_type are significant.

VIF
All VIF values are less than 10 indicating low multicollinearity.

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	1	1.80264	0.80729	2.23	0.0280	0
Prod_Qual	1	0.29739	0.05852	5.08	<.0001	1.34820
Ecommerce	1	-0.37396	0.16785	-2.23	0.0284	2.79152
Tech_support	1	0.20929	0.07745	2.70	0.0082	2.83675
Adv	1	-0.08760	0.07507	-1.17	0.2463	1.44518
Prod_Line	1	0.15547	0.06938	2.24	0.0275	1.68118
Sales_Image	1	0.66019	0.11905	5.55	<.0001	3.29083
Warranty	1	-0.37470	0.14896	-2.52	0.0137	3.01074
Ordering	1	0.31632	0.08775	3.60	0.0005	1.34127
Industry_type	1	0.25593	0.14410	1.78	0.0791	1.05710

Measures of fit

The **R^2** value of 0.5905 suggests that the regression model explains a good amount of the variability in the dependent variable.

The **adjusted R^2** value of 0.5496 is lower than the **R^2** value of 0.5905 suggesting that some of the predictors may not be contributing much to explain the variance in the dependent variable.

The **Root MSE** value suggests that the regression model's predictions are off by 0.70022 which is relatively low, indicating that predictions are close to the actual values.

The SAS System

The REG Procedure
Model: MODEL2
Dependent Variable: Recommend

Number of Observations Read	100
Number of Observations Used	100

Reduced model with the second highest adj R-squared method, including 10 suggested variables. Option 2.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	10	63.95023	6.39502	12.99	<.0001
Error	89	43.80977	0.49224		
Corrected Total	99	107.76000			

H₀: $\beta_1 = \beta_2 = \beta_3 = 0$;
H₁: not all $\beta_j = 0$

$F_{10,89}$=12.99

P-Value <0.0001. Reject is Null.

At least one independent variable helps to explain recommend.

Root MSE	0.70160	**R-Square**	0.5935
Dependent Mean	7.02000	**Adj R-Sq**	0.5478
Coeff Var	9.99432		

H₀: $\beta_k = 0$; Where k = each independent variable.
H₁: $\beta_k \neq 0$

At α=0.10, Prod_Qual, Ecommerce, Tech_support, Prod_Line, Sales_Image, Warranty, Ordering, and Industry_type are significant.

VIF
All VIF values are less than 10 indicating low multicollinearity.

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	1	1.30518	1.01853	1.28	0.2034	0
Prod_Qual	1	0.31966	0.06486	4.93	<.0001	1.64932
Ecommerce	1	-0.36385	0.16865	-2.16	0.0337	2.80716
Tech_support	1	0.20722	0.07764	2.67	0.0090	2.83988
Adv	1	-0.10092	0.07703	-1.31	0.1935	1.51540
Prod_Line	1	0.18864	0.08084	2.33	0.0219	2.27394
Sales_Image	1	0.64588	0.12061	5.36	<.0001	3.36411
Warranty	1	-0.35223	0.15185	-2.32	0.0227	3.11636
Ordering	1	0.25004	0.12055	2.07	0.0410	2.52150
Price_Flex	1	0.08308	0.10338	0.80	0.4237	3.12601
Industry_type	1	0.24842	0.14469	1.72	0.0895	1.06153

Measures of fit

The **R^2** value of 0.5935 suggests that the regression model explains a good amount of the variability in the dependent variable.

The **adjusted R^2** value of 0.5478 is lower than the **R^2** value of 0.5935 suggesting that some of the predictors may not be contributing much to explain the variance in the dependent variable.

The **Root MSE** value suggests that the regression model's predictions are off by 0.70160 which is relatively low, indicating that predictions are close to the actual values.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

Number of Observations Read	100
Number of Observations Used	100

Final Model
Reduced model with the second highest adj R-squared method, including 8 suggested variables. Option 3.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	8	62.96470	7.87059	15.99	<.0001
Error	91	44.79530	0.49226		
Corrected Total	99	107.76000			

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$;
H_1: not all $\beta_j = 0$

$F_{8,91}$=15.99

P-Value <0.0001. Reject is Null.

At least one independent variable helps to explain recommend.

Root MSE	0.70161	**R-Square**	0.5843
Dependent Mean	7.02000	**Adj R-Sq**	0.5478
Coeff Var	9.99444		

H_0: $\beta_k = 0$;
H_1: $\beta_k \neq 0$
Where k = each independent variable.

At α=0.10, all included variables are significant. Industry_type would not be significant at an α=0.05.

VIF
All VIF values are less than 10 indicating low multicollinearity.

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	1	1.73063	0.80652	2.15	0.0346	0
Prod_Qual	1	0.29482	0.05860	5.03	<.0001	1.34629
Ecommerce	1	-0.37291	0.16818	-2.22	0.0291	2.79144
Tech_support	1	0.21454	0.07747	2.77	0.0068	2.82719
Prod_Line	1	0.15685	0.06950	2.26	0.0264	1.68069
Sales_Image	1	0.61085	0.11152	5.48	<.0001	2.87581
Warranty	1	-0.37500	0.14926	-2.51	0.0138	3.01073
Ordering	1	0.30644	0.08751	3.50	0.0007	1.32877
Industry_type	1	0.25092	0.14432	1.74	0.0855	1.05616

Measures of fit

The **R^2** value of 0.5843 suggests that the regression model explains a good amount of the variability in the dependent variable.

The **adjusted R^2** value of 0.5478 is lower than the **R^2** value of 0.5843 suggesting that some of the predictors may not be contributing much to explain the variance in the dependent variable.

The **Root MSE** value suggests that the regression model's predictions are off by 0.70161 which is relatively low, indicating that predictions are close to the actual values.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

Output Statistics					
Obs	**Residual**	**RStudent**	**Hat Diag H**	**Cov Ratio**	**DFFITS**
1	0.4781	0.7115	0.0879	1.1514	0.2209
2	-0.2517	-0.3771	0.1034	1.2146	-0.1281
3	-0.009883	-0.0145	0.0627	1.1784	-0.0037
4	-0.3183	-0.4663	0.0617	1.1519	-0.1196
5	-0.4131	-0.5964	0.0322	1.1015	-0.1088
6	0.6559	0.9719	0.0752	1.0873	0.2772
7	1.5884	2.5097	0.1389	0.6978	**1.0078**
8	-0.3373	-0.4961	0.0687	1.1573	-0.1347
9	0.4526	0.6656	0.0663	1.1319	0.1774
10	-0.7408	-1.0968	0.0712	1.0553	-0.3037
11	0.0449	0.0649	0.0390	1.1489	0.0131
12	-0.0188	-0.0284	0.1176	1.2517	-0.0104
13	0.4799	0.7444	0.1598	1.2440	0.3247
14	-0.6785	-1.0090	0.0811	1.0864	-0.2998
15	0.1903	0.2803	0.0733	1.1826	0.0788
16	-0.1353	-0.2011	0.0907	1.2099	-0.0635
17	1.1397	1.6908	0.0582	0.8851	0.4202
18	-0.006318	-0.009318	0.0764	1.1959	-0.0027
19	0.2192	0.3275	0.0987	1.2125	0.1084
20	0.6200	0.9293	0.0972	1.1228	0.3050
21	0.0527	0.0789	0.1041	1.2322	0.0269
22	0.4707	0.7447	0.1922	1.2939	0.3633
23	-0.3856	-0.5623	0.0518	1.1287	-0.1314
24	0.4334	0.7165	0.2605	1.4192	0.4253
25	-1.2792	-1.9095	0.0617	0.8234	-0.4897
26	-0.1687	-0.2548	0.1186	1.2451	-0.0935
27	-0.0514	-0.0756	0.0725	1.1902	-0.0211
28	-0.6718	-0.9838	0.0530	1.0593	-0.2327
29	-0.4646	-0.6898	0.0837	1.1496	-0.2085
30	-1.0947	-1.6381	0.0759	0.9175	-0.4693

Observation #7 has the highest DFFITS value of **1.0078** but is less than 2. Influence acceptable levels and will remain in the model.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

Output Statistics					
Obs	**Residual**	**RStudent**	**Hat Diag H**	**Cov Ratio**	**DFFITS**
31	0.1074	0.1595	0.0894	1.2099	0.0500
32	-0.4215	-0.6244	0.0803	1.1552	-0.1845
33	-0.1997	-0.2923	0.0613	1.1666	-0.0747
34	0.2616	0.3831	0.0618	1.1602	0.0983
35	-0.1717	-0.2606	0.1265	1.2560	-0.0992
36	-1.2016	-1.8410	0.1118	0.8916	-0.6531
37	1.1331	1.6591	0.0342	0.8721	0.3121
38	1.6645	2.5192	0.0610	0.6371	0.6422
39	-0.3064	-0.4460	0.0501	1.1399	-0.1024
40	-0.8330	-1.2567	0.1017	1.0514	-0.4229
41	-0.1387	-0.2097	0.1205	1.2504	-0.0776
42	0.2243	0.3307	0.0745	1.1805	0.0938
43	-0.1382	-0.2083	0.1157	1.2437	-0.0753
44	0.4876	0.7417	0.1264	1.1969	0.2821
45	-1.2616	-1.9399	0.1146	0.8629	-0.6979
46	1.8790	2.8790	0.0654	0.5348	0.7614
47	1.3351	2.0947	0.1441	0.8408	0.8594
48	-0.1798	-0.2743	0.1358	1.2686	-0.1088
49	-0.3205	-0.4857	0.1229	1.2299	-0.1818
50	0.0896	0.1313	0.0646	1.1789	0.0345
51	1.1133	1.6411	0.0477	0.8896	0.3674
52	0.3898	0.5856	0.1065	1.1946	0.2022
53	1.3679	2.0739	0.0842	0.7923	0.6287
54	0.5738	0.8428	0.0615	1.0965	0.2158
55	1.0687	1.5891	0.0658	0.9217	0.4216
56	-0.4173	-0.6106	0.0575	1.1291	-0.1508
57	0.1689	0.2517	0.0951	1.2129	0.0816
58	0.2738	0.3988	0.0512	1.1459	0.0927
59	-0.1804	-0.2768	0.1459	1.2834	-0.1144
60	0.6605	1.0202	0.1482	1.1693	0.4255

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

Output Statistics					
Obs	**Residual**	**RStudent**	**Hat Diag H**	**Cov Ratio**	**DFFITS**
61	-0.5038	-0.7556	0.1012	1.1610	-0.2536
62	0.1165	0.1808	0.1664	1.3208	0.0808
63	-1.3285	-1.9965	0.0710	0.8050	-0.5519
64	-0.1073	-0.1587	0.0817	1.1998	-0.0473
65	1.0174	1.4972	0.0492	0.9309	0.3405
66	-0.0703	-0.1037	0.0780	1.1967	-0.0302
67	-0.9629	-1.4504	0.0937	0.9899	-0.4662
68	-0.0351	-0.0511	0.0540	1.1673	-0.0122
69	-0.6079	-0.8886	0.0515	1.0765	-0.2070
70	0.4063	0.6031	0.0845	1.1635	0.1832
71	-0.7047	-1.0659	0.1108	1.1095	-0.3762
72	-0.0828	-0.1240	0.1045	1.2316	-0.0424
73	-0.3231	-0.4797	0.0861	1.1812	-0.1473
74	0.1743	0.2635	0.1195	1.2458	0.0970
75	-0.5638	-0.8220	0.0477	1.0844	-0.1841
76	-0.4594	-0.6908	0.1068	1.1791	-0.2388
77	0.5235	0.7992	0.1318	1.1938	0.3114
78	-0.1304	-0.1911	0.0641	1.1759	-0.0500
79	-1.1177	-1.6591	0.0602	0.8962	-0.4200
80	-0.9556	-1.4463	0.1024	1.0008	-0.4886
81	0.1562	0.2364	0.1224	1.2515	0.0883
82	-0.006968	-0.0100	0.0329	1.1421	-0.0019
83	-0.1906	-0.2791	0.0619	1.1683	-0.0717
84	-0.3970	-0.6186	0.1691	1.2796	-0.2791
85	0.2727	0.4006	0.0675	1.1657	0.1078
86	-0.3634	-0.5301	0.0529	1.1340	-0.1253
87	0.5840	0.9105	0.1660	1.2195	0.4062
88	0.1212	0.1806	0.0952	1.2168	0.0586
89	-0.7564	-1.1042	0.0445	1.0242	-0.2384
90	-0.4174	-0.6447	0.1538	1.2523	-0.2748

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

Output Statistics					
Obs	**Residual**	**RStudent**	**Hat Diag H**	**Cov Ratio**	**DFFITS**
91	0.2910	0.4283	0.0703	1.1666	0.1178
92	-0.2049	-0.3120	0.1319	1.2601	-0.1216
93	0.000629	0.000935	0.0894	1.2130	0.0003
94	-0.5771	-0.8468	0.0594	1.0934	-0.2128
95	-0.3437	-0.5037	0.0622	1.1484	-0.1297
96	0.4221	0.6224	0.0722	1.1453	0.1736
97	-0.9240	-1.3697	0.0667	0.9830	-0.3662
98	1.3866	2.1560	0.1261	0.8033	0.8190
99	-0.8342	-1.2332	0.0650	1.0160	-0.3252
100	0.6689	0.9857	0.0648	1.0723	0.2594

Output Statistics									
	DFBETAS								
Obs	**Intercept**	**Prod_Qual**	**Ecommerce**	**Tech_support**	**Prod_Line**	**Sales_Image**	**Warranty**	**Ordering**	**Industry_type**
1	0.0154	0.0798	-0.0312	-0.0560	-0.0815	0.0577	-0.0317	0.0834	-0.0562
2	-0.0480	0.0335	0.0041	-0.0171	-0.0831	0.0331	0.0355	0.0283	-0.0420
3	0.0006	-0.0008	0.0020	-0.0004	-0.0008	-0.0022	0.0006	-0.0010	0.0007
4	-0.0048	0.0305	0.0117	-0.0072	0.0417	0.0195	-0.0374	-0.0236	-0.0448
5	0.0135	-0.0488	-0.0201	0.0205	0.0301	0.0381	-0.0212	-0.0203	0.0560
6	0.1435	-0.0551	-0.0524	-0.1114	-0.0539	-0.0487	0.0418	0.0007	0.0980
7	0.4955	0.1538	-0.1394	0.3780	-0.4223	0.2294	-0.4271	-0.3027	0.2283
8	-0.0615	0.0222	0.0349	0.0246	0.0733	-0.0148	-0.0064	-0.0394	0.0496
9	0.0559	-0.1082	-0.0867	0.0089	0.0666	0.0895	-0.0242	-0.0067	0.0796
10	-0.1116	0.1375	-0.1547	-0.0963	-0.0837	0.0781	0.1312	0.0403	0.1498
11	0.0022	0.0007	-0.0019	-0.0034	0.0058	0.0006	0.0010	-0.0058	-0.0056
12	-0.0026	0.0021	-0.0046	-0.0042	0.0012	0.0010	0.0031	0.0035	0.0045
13	-0.2264	0.1126	0.2152	-0.1263	-0.0018	-0.1022	0.1160	0.0173	0.0636
14	0.1344	-0.0141	-0.0822	0.0818	-0.1783	0.0708	-0.0890	0.0717	-0.1111
15	0.0025	-0.0467	0.0340	-0.0136	0.0387	-0.0171	0.0047	-0.0044	-0.0347
16	0.0049	0.0060	-0.0155	0.0358	-0.0290	0.0275	-0.0266	0.0136	0.0205
17	0.0467	-0.2671	0.1100	0.0666	0.0660	-0.1145	0.0424	-0.0193	0.1293

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

Output Statistics									
	DFBETAS								
Obs	Intercept	Prod_Qual	Ecommerce	Tech_support	Prod_Line	Sales_Image	Warranty	Ordering	Industry_type
18	-0.0003	0.0017	-0.0007	-0.0002	-0.0005	0.0006	0.0002	-0.0012	0.0011
19	0.0517	-0.0574	-0.0542	0.0358	0.0238	0.0479	-0.0506	0.0304	0.0418
20	-0.1233	0.1488	-0.0229	-0.1346	-0.0860	0.1286	0.0827	0.0001	-0.0726
21	0.0050	-0.0138	0.0062	0.0005	-0.0028	-0.0070	0.0085	-0.0076	-0.0115
22	-0.1527	0.0597	0.1057	0.0943	0.1654	0.0774	-0.0873	-0.1088	0.0828
23	0.0072	0.0064	-0.0295	-0.0678	-0.0395	0.0349	0.0332	-0.0113	0.0561
24	-0.1135	0.0746	-0.1661	-0.2789	-0.0551	0.0606	0.2195	0.1941	-0.0200
25	-0.0004	0.2414	-0.2038	0.1450	0.0393	0.1980	-0.1958	-0.0741	0.2445
26	-0.0222	0.0079	0.0071	-0.0399	0.0541	0.0017	0.0264	-0.0636	0.0270
27	0.0006	0.0027	-0.0026	-0.0048	-0.0046	0.0093	-0.0002	-0.0066	0.0072
28	0.0071	0.0177	-0.0135	0.0843	0.0936	0.0860	-0.1207	-0.0510	-0.0934
29	-0.0475	0.0253	-0.0169	-0.0618	-0.0504	0.0772	0.0760	-0.1039	0.0628
30	-0.0832	-0.1075	-0.0411	0.1274	0.2495	0.1067	0.0132	-0.2275	-0.1666
31	-0.0047	-0.0173	-0.0039	0.0100	0.0129	-0.0013	0.0129	-0.0160	0.0159
32	-0.0486	-0.0112	0.0581	-0.0332	0.0325	-0.0022	-0.0258	0.0671	0.0589
33	-0.0389	0.0220	-0.0021	-0.0519	-0.0080	0.0010	0.0455	0.0145	-0.0254
34	0.0154	0.0288	-0.0236	0.0403	-0.0561	0.0060	-0.0310	0.0507	0.0444
35	-0.0118	0.0233	0.0293	0.0115	-0.0294	-0.0558	-0.0120	0.0715	0.0251
36	-0.1297	-0.0310	-0.3594	-0.0609	-0.0472	0.4915	-0.0072	0.2105	0.2653
37	-0.0395	0.1240	0.0606	-0.0242	-0.0455	-0.0952	0.0772	-0.0696	-0.1640
38	-0.1531	0.1743	0.1744	0.2447	0.1919	-0.0495	-0.2500	0.1190	-0.2270
39	-0.0101	-0.0382	0.0120	0.0097	0.0464	-0.0043	0.0180	-0.0479	-0.0514
40	-0.0387	0.0589	-0.2012	0.1196	0.1364	0.0878	-0.0227	-0.0731	0.1758
41	-0.0313	-0.0214	0.0248	-0.0225	0.0205	-0.0233	0.0496	-0.0274	-0.0283
42	0.0164	0.0332	0.0090	0.0369	0.0177	-0.0001	-0.0614	-0.0031	0.0371
43	0.0086	-0.0205	-0.0450	-0.0175	-0.0088	0.0152	0.0323	-0.0001	0.0261
44	-0.0818	-0.1249	-0.0040	-0.0167	0.0241	0.1012	0.0699	0.0064	0.0743
45	-0.3426	0.1885	0.1044	-0.1479	-0.3874	0.1277	0.2963	0.0131	-0.2335
46	0.0787	-0.5264	0.2128	-0.0010	0.2233	-0.1792	0.0428	0.1725	-0.3403
47	-0.3572	0.4825	0.4803	0.1425	-0.4199	-0.4685	0.0385	0.3137	0.1761

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

Output Statistics									
	DFBETAS								
Obs	Intercept	Prod_Qual	Ecommerce	Tech_support	Prod_Line	Sales_Image	Warranty	Ordering	Industry_type
48	-0.0027	0.0388	0.0311	-0.0317	0.0122	-0.0258	0.0187	-0.0755	-0.0356
49	0.0566	-0.0706	0.0866	0.0725	-0.0237	-0.1218	-0.0366	0.0271	0.0201
50	0.0066	-0.0045	-0.0080	0.0042	0.0203	0.0147	-0.0150	-0.0088	0.0160
51	0.2180	-0.1068	0.0610	0.1902	-0.0786	-0.0830	-0.1657	0.0276	-0.1962
52	0.0528	-0.0447	-0.0055	-0.0439	0.1118	-0.0613	-0.0107	-0.0036	0.0692
53	-0.0472	0.1281	-0.4194	-0.2121	0.1131	0.2669	0.1635	0.0065	-0.0893
54	0.0346	-0.0947	-0.0250	-0.0307	0.1183	0.0091	0.0462	-0.1356	0.0843
55	0.0457	0.1326	-0.0968	-0.0791	-0.2296	0.0133	-0.0112	0.2363	0.1930
56	0.0419	-0.0652	0.0224	-0.0724	-0.0040	-0.0436	0.0359	0.0192	-0.0741
57	-0.0351	-0.0163	0.0368	-0.0233	0.0219	-0.0073	0.0177	0.0158	-0.0283
58	-0.0329	-0.0126	-0.0266	-0.0490	0.0227	0.0228	0.0531	-0.0041	0.0493
59	0.0591	-0.0624	-0.0668	0.0452	0.0376	0.0614	-0.0579	0.0126	-0.0218
60	-0.0091	0.1908	-0.0361	0.1851	-0.1782	-0.0817	-0.1276	0.2599	0.1368
61	0.0176	0.0219	-0.0540	-0.1139	-0.0481	0.1092	0.0331	-0.0733	0.0782
62	0.0333	0.0124	-0.0117	0.0581	-0.0019	0.0275	-0.0576	-0.0261	-0.0190
63	-0.0702	-0.2501	-0.2257	0.0148	0.2749	0.2104	0.1124	-0.1566	0.2285
64	-0.0277	0.0006	0.0197	-0.0146	0.0116	-0.0002	0.0079	0.0056	0.0145
65	0.0655	-0.1036	0.0753	-0.0594	-0.0680	-0.1021	0.1335	-0.1025	-0.1844
66	-0.0076	0.0160	0.0112	0.0081	-0.0085	-0.0126	0.0004	-0.0028	-0.0124
67	0.0640	0.2156	-0.1025	-0.0682	0.0379	0.1356	-0.1189	-0.1250	-0.1175
68	-0.0057	0.0085	-0.0005	-0.0004	-0.0021	0.0018	0.0019	-0.0023	-0.0045
69	-0.0680	-0.0440	-0.0127	0.0262	0.1113	0.0450	0.0219	-0.0504	-0.0850
70	0.0446	-0.0180	-0.0126	-0.0895	-0.0863	-0.0023	0.0418	0.0651	-0.0620
71	0.2393	-0.1852	0.0203	0.0802	0.1899	-0.1181	-0.1459	-0.1211	0.0824
72	-0.0151	0.0165	0.0074	0.0176	-0.0233	-0.0037	-0.0053	0.0243	-0.0134
73	-0.0498	0.0639	-0.0242	0.0246	-0.0624	-0.0065	0.0395	0.0045	0.0530
74	-0.0355	0.0354	-0.0460	0.0155	0.0100	0.0670	0.0054	-0.0264	-0.0134
75	-0.0069	-0.0816	0.0321	-0.0488	0.0866	-0.0703	0.0566	-0.0355	0.0706
76	0.0742	-0.1103	0.0622	0.0532	0.0374	-0.0466	-0.1187	0.1067	0.0528
77	-0.1256	0.0364	0.0326	-0.0863	-0.1603	-0.0640	0.2206	0.0396	-0.0905

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
RED = Outliers
GREEN = Leverage

Output Statistics									
	DFBETAS								
Obs	Intercept	Prod_Qual	Ecommerce	Tech_support	Prod_Line	Sales_Image	Warranty	Ordering	Industry_type
78	0.0326	-0.0226	-0.0098	0.0083	0.0059	0.0100	-0.0251	0.0035	-0.0208
79	0.1253	-0.0985	0.2054	-0.0549	-0.0997	-0.2413	0.0255	-0.0621	0.0903
80	-0.0552	0.0371	0.2661	0.1050	-0.0671	-0.3050	-0.0986	0.2994	-0.1871
81	0.0217	0.0334	-0.0271	0.0584	-0.0021	0.0273	-0.0478	-0.0237	-0.0184
82	-0.0003	-0.0001	0.0003	-0.0007	-0.0005	-0.0000	0.0004	0.0002	0.0009
83	-0.0216	-0.0103	0.0063	-0.0028	-0.0042	0.0064	-0.0042	0.0461	0.0283
84	0.0022	-0.0011	0.0647	0.0634	0.1032	-0.0434	-0.1461	0.1252	-0.0623
85	0.0343	-0.0367	-0.0578	0.0332	0.0424	0.0249	-0.0195	0.0059	-0.0282
86	-0.0652	-0.0204	-0.0062	-0.0170	0.0458	0.0175	0.0460	0.0064	-0.0484
87	0.2299	-0.1906	0.0263	-0.1693	0.0817	-0.0249	0.0273	-0.1911	-0.1284
88	0.0138	-0.0164	-0.0266	0.0305	0.0139	0.0104	-0.0155	0.0129	-0.0129
89	0.1163	-0.1150	-0.0381	0.0800	0.0509	0.0675	-0.1027	-0.0448	-0.1223
90	0.0971	0.0908	-0.0262	-0.0522	-0.0105	-0.0813	-0.0145	-0.0499	-0.0638
91	-0.0787	0.0470	-0.0151	-0.0296	-0.0188	0.0167	0.0700	-0.0057	0.0536
92	-0.0147	-0.0115	-0.0644	0.0313	0.0338	0.0568	-0.0282	0.0586	-0.0148
93	-0.0001	0.0000	0.0001	-0.0001	0.0002	-0.0001	0.0001	-0.0001	0.0001
94	0.0640	-0.0486	0.1030	0.0333	-0.0717	-0.1258	-0.0237	0.0158	0.0469
95	0.0082	-0.0524	0.0202	0.0183	-0.0331	-0.0432	0.0192	0.0499	-0.0626
96	-0.0489	0.0501	0.1118	0.0435	-0.0186	-0.0634	-0.0483	0.0472	0.0517
97	-0.1954	-0.0197	0.0192	0.0680	0.1184	0.0146	0.0883	-0.0768	0.1376
98	0.2343	0.0420	0.4159	-0.2121	0.1181	-0.5675	0.0357	-0.2547	-0.3024
99	0.0867	-0.0630	-0.0335	-0.0862	0.0309	-0.1285	0.0801	0.0108	-0.1274
100	0.0460	0.0763	-0.1382	-0.0649	-0.1100	0.0939	0.0462	-0.0324	0.1238

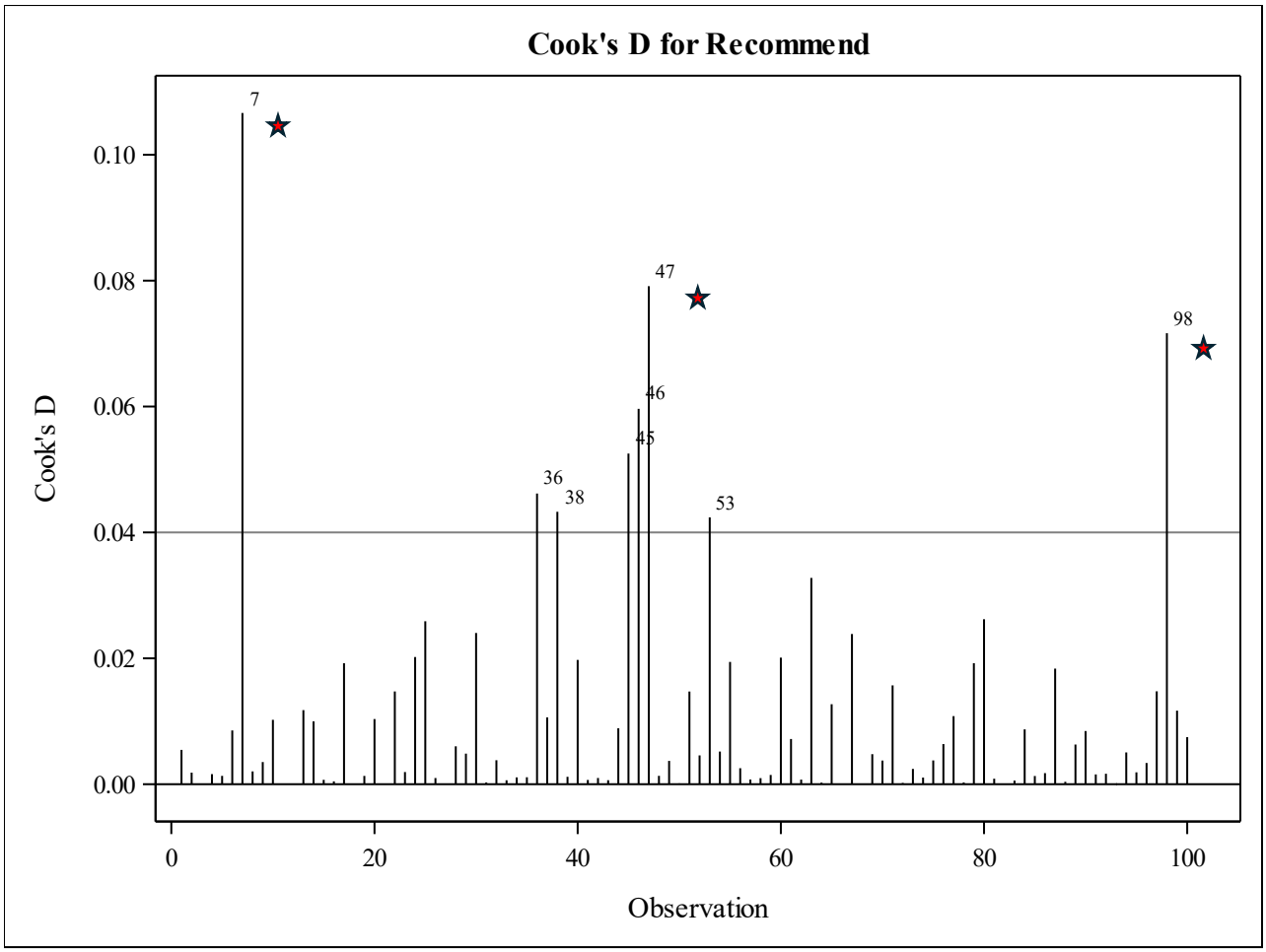
The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

The SSE of 44.7953 and the corresponding PRESS value of 53.86246 suggest that while my model explains a certain amount of variance within the sample data, it may not generalize well to new observations. This indicates potential limitations in the model's predictive performance on unseen data, which could be further investigated and addressed to improve the overall robustness and applicability of the regression model.

Sum of Residuals	0
Sum of Squared Residuals	44.79530
Predicted Residual SS (PRESS)	53.86246

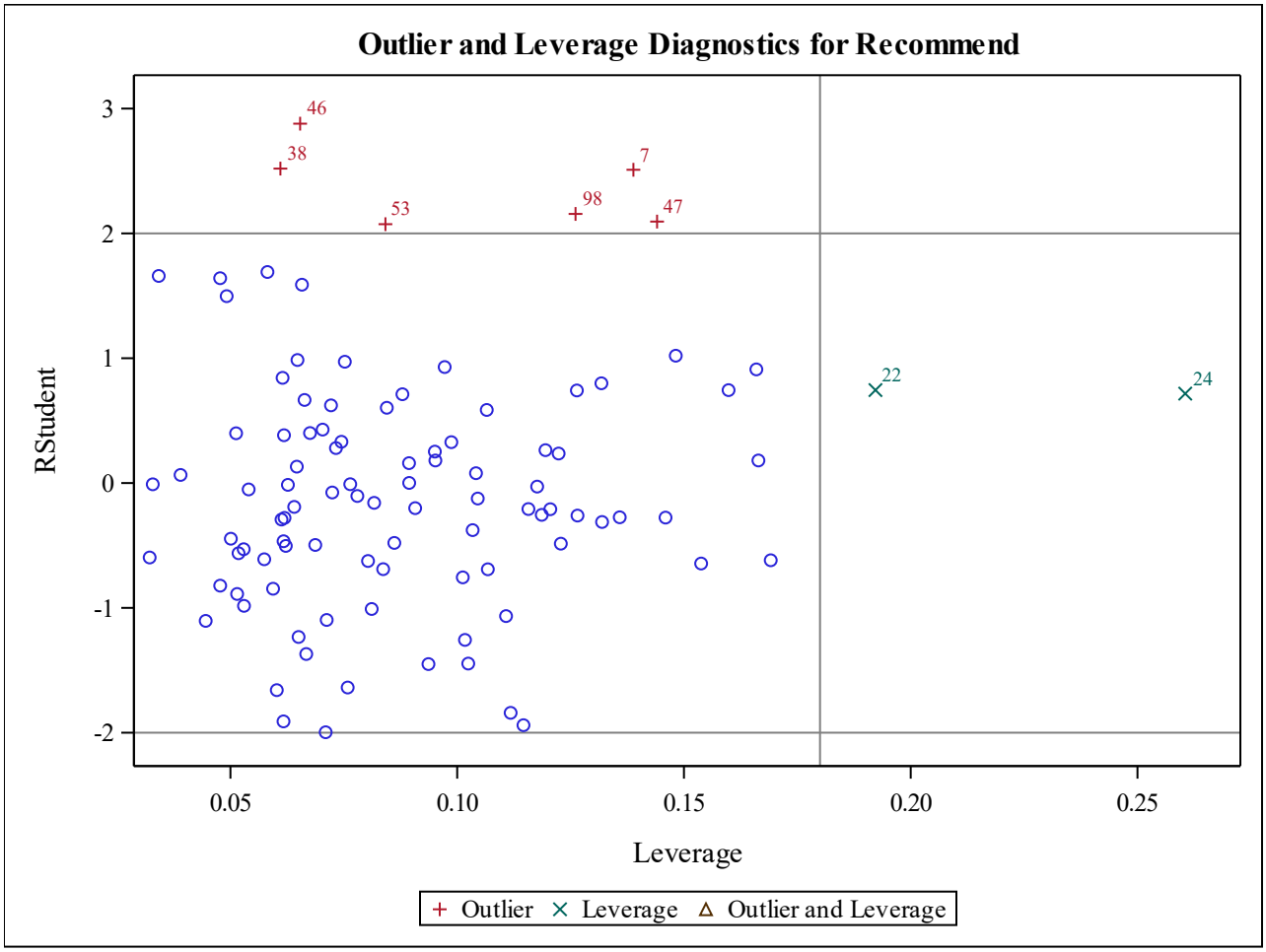


The observations annotated with a red star are highlighted in DFBETAS above. The observations 7, 47, and 98 do not have values greater than two allowing them to remain in the model.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

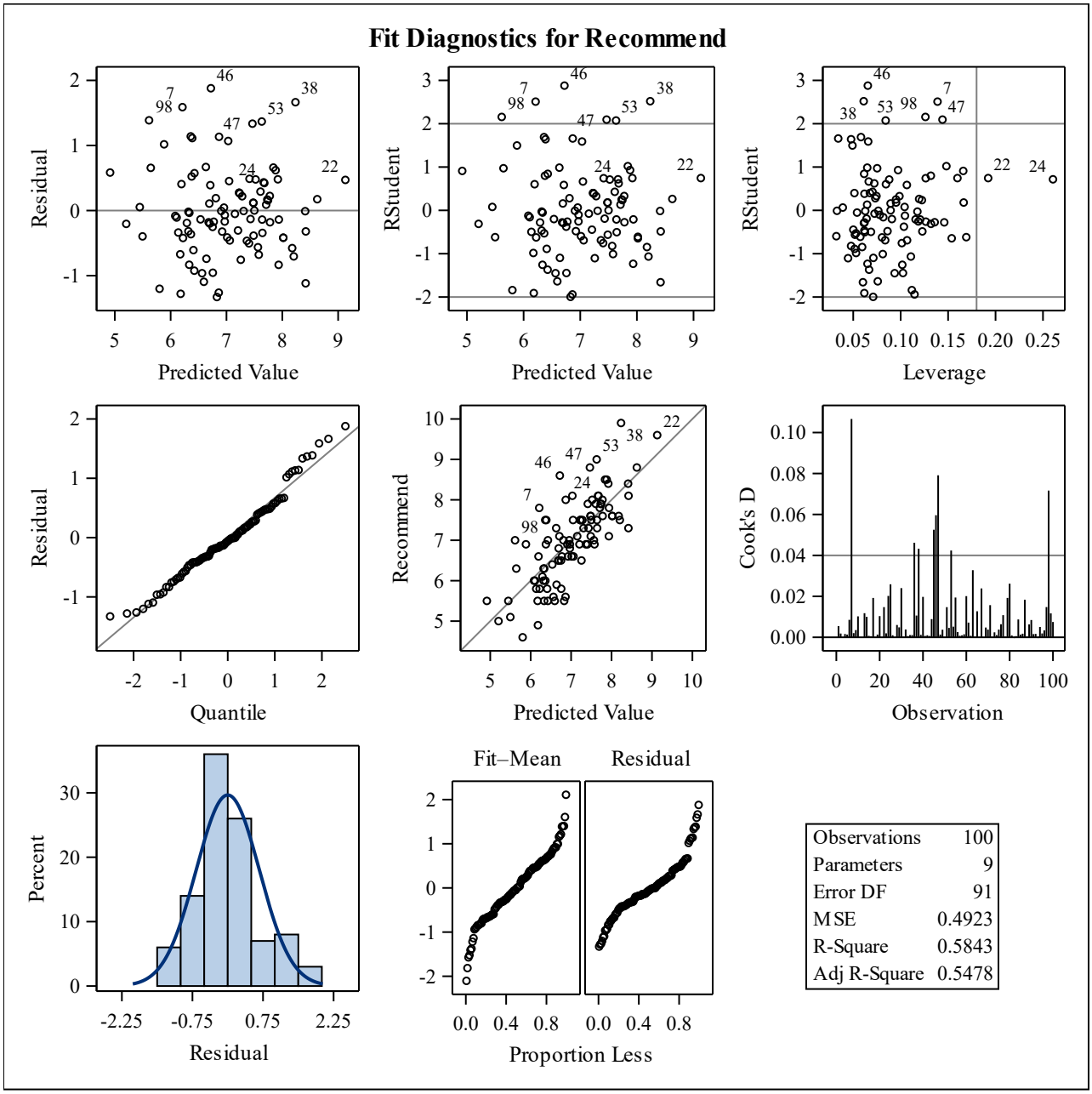


Observations 7, 38, 46, 47, 53, and 98 are outliers and will be highlighted in red in the DFBETAS table. Observations 22 and 24 are Leverage and will be highlighted in green in the DFBETAS table. There are no observations that are Outliers and Leverage.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
RED = Outliers
GREEN = Leverage

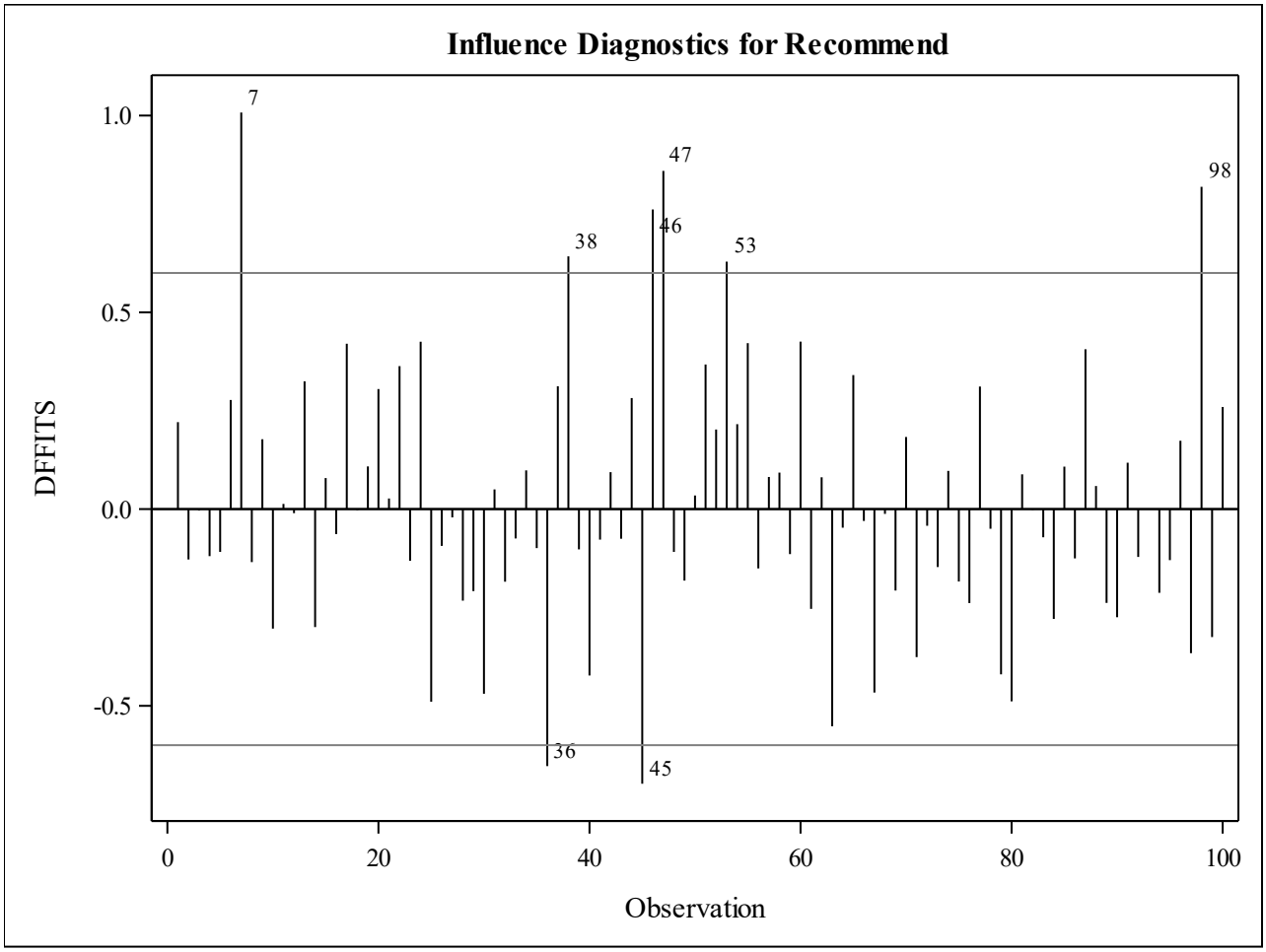


There is a positive linear relationship in the normal probability plot and the slope is close to 1, residuals versus quantiles. The skew is nearly symmetrical, and the data appears normally distributed. The histogram appears to have a normal distribution. The RF plot consisting of side-by-side quantile plots of the centered fit and the residuals shows that the spread in the residuals is no greater than the spread in the centered fit. The residual versus predicted value appears to be randomized with no noticeable pattern, suggesting homoscedasticity, upholding the assumption of constant variance.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

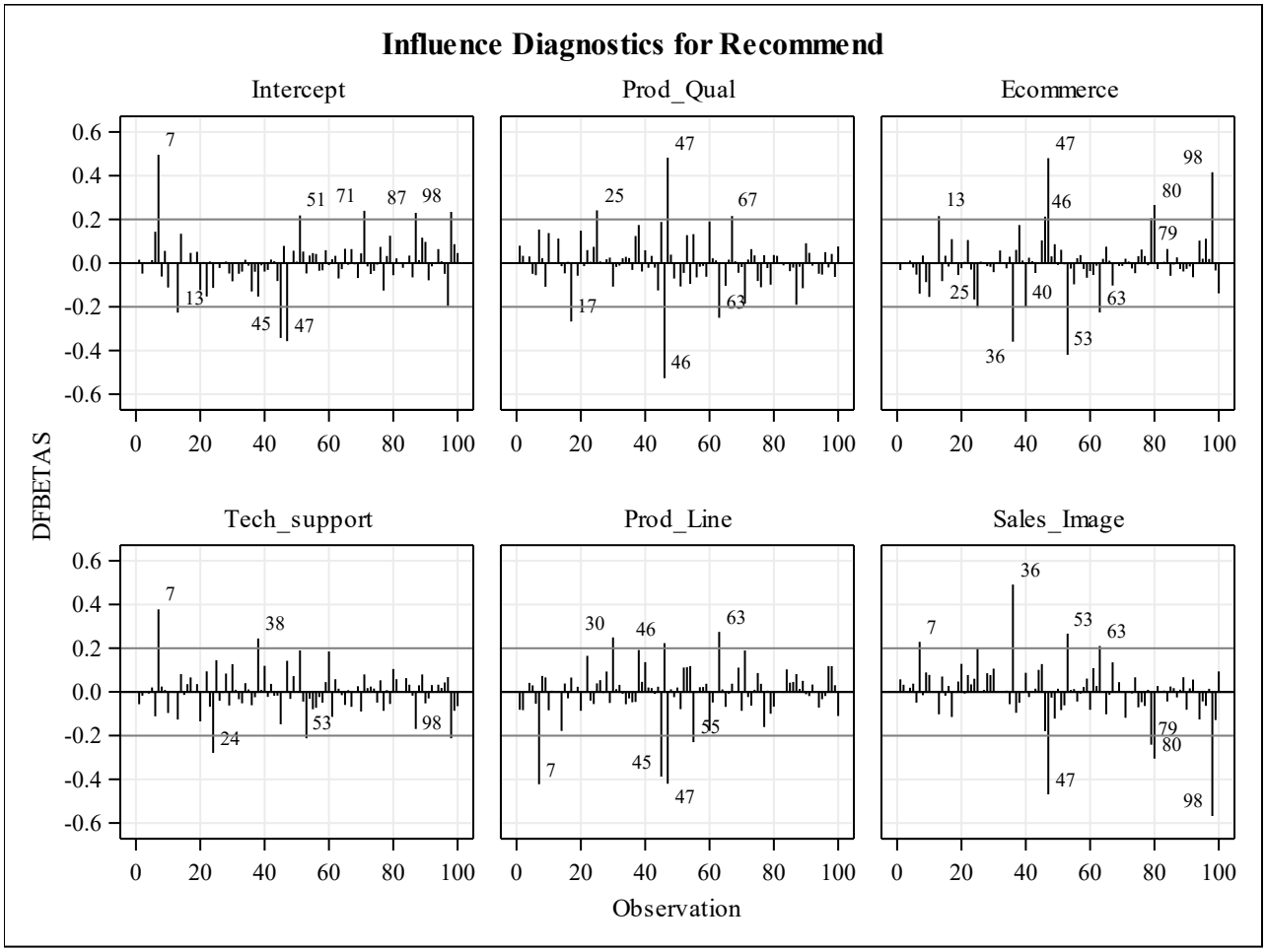


Observation 7 has a value of 1.0078 so a validation model will be run after the final model to check for variable stability across subsets.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
***RED** = Outliers*
***GREEN** = Leverage*

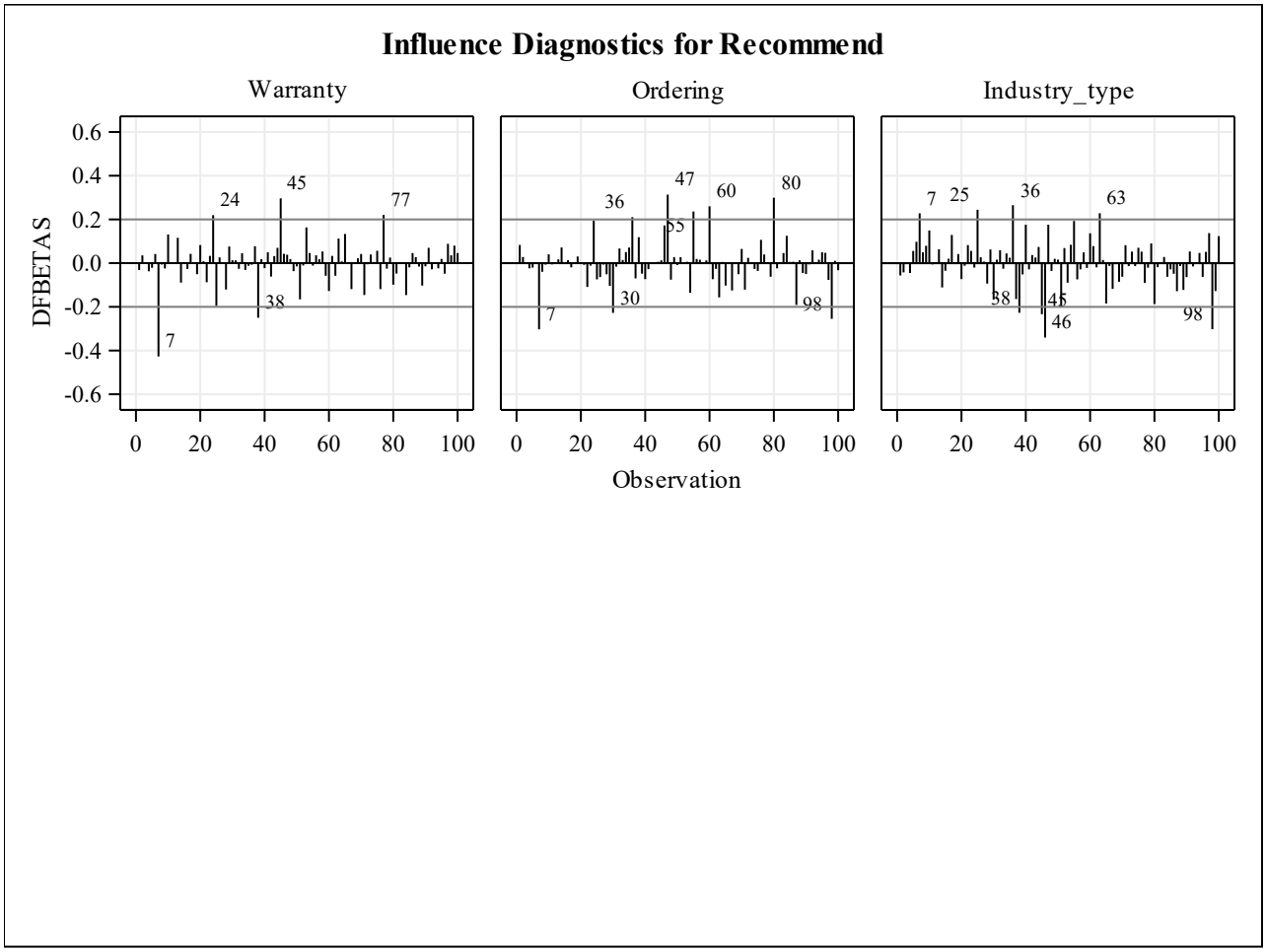


No observations greater than two are observed.

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
RED = Outliers
GREEN = Leverage

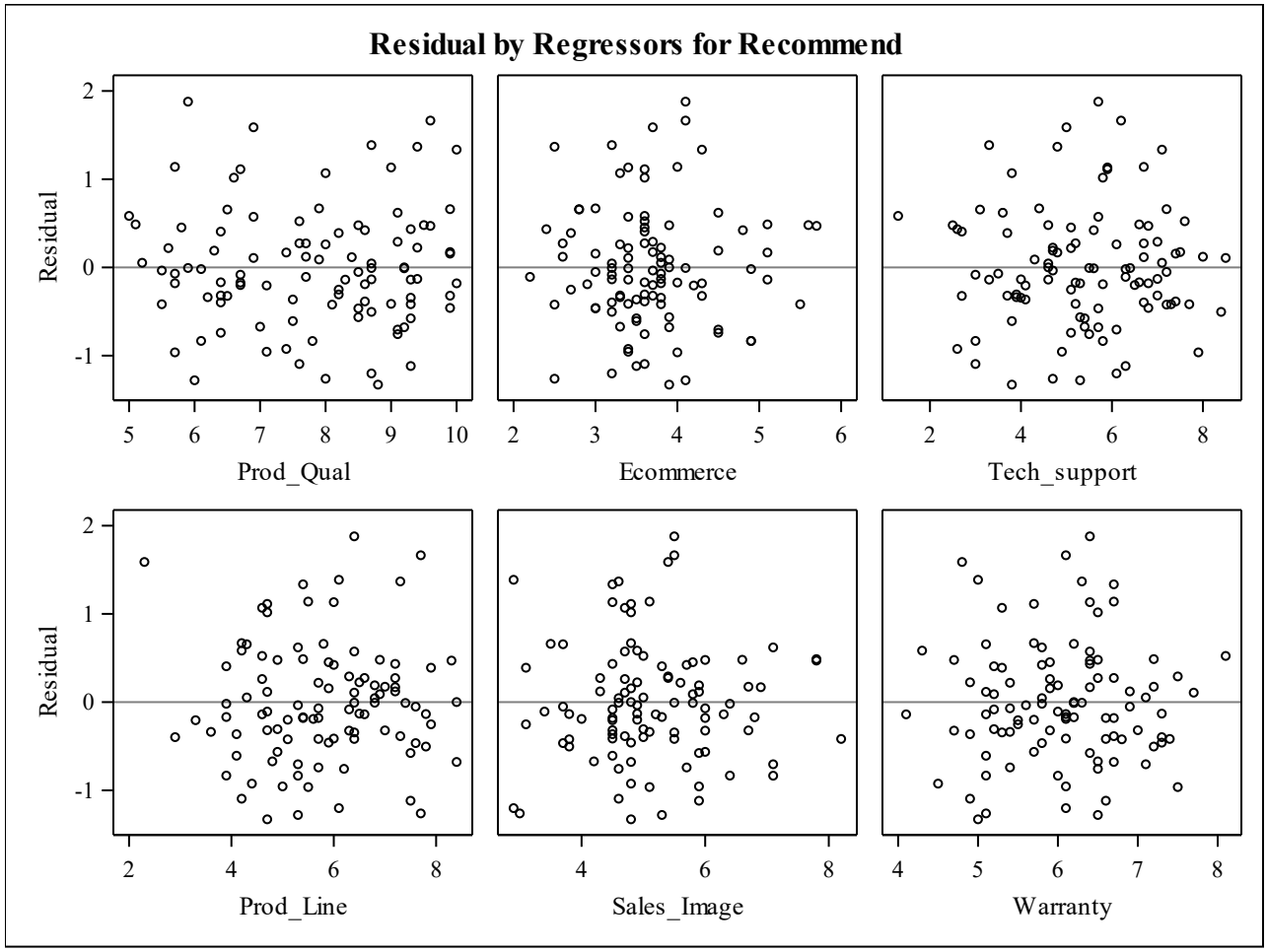


No observations greater than two are observed.

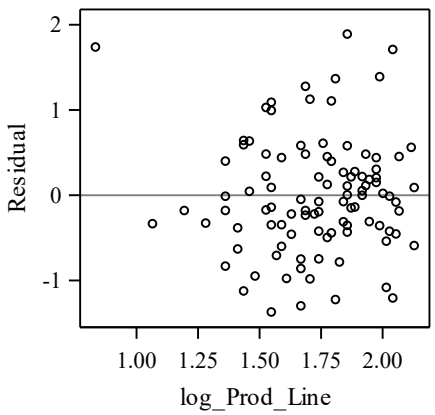
The SAS System

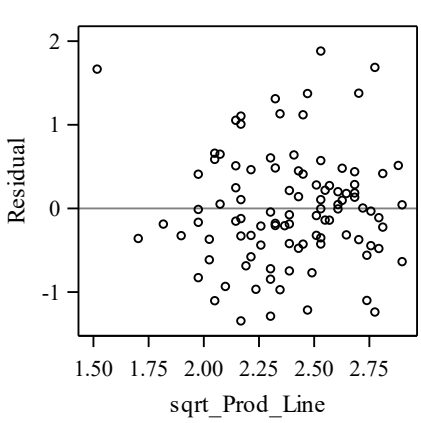
The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
RED = Outliers
GREEN = Leverage



Residuals for the metric variables Prod_Qual, Ecommerce, Tech_support, Prod_Line, Sales_Image, and Warranty are normally distributed with about 68% of the residuals should be within 1 standard deviation of the mean of 0, 95% within 2 standard deviations, and about 99% within 3 standard deviations. Prod_Line shows some coning and may not uphold constant variance, but square root and log transformations did not improve the output.

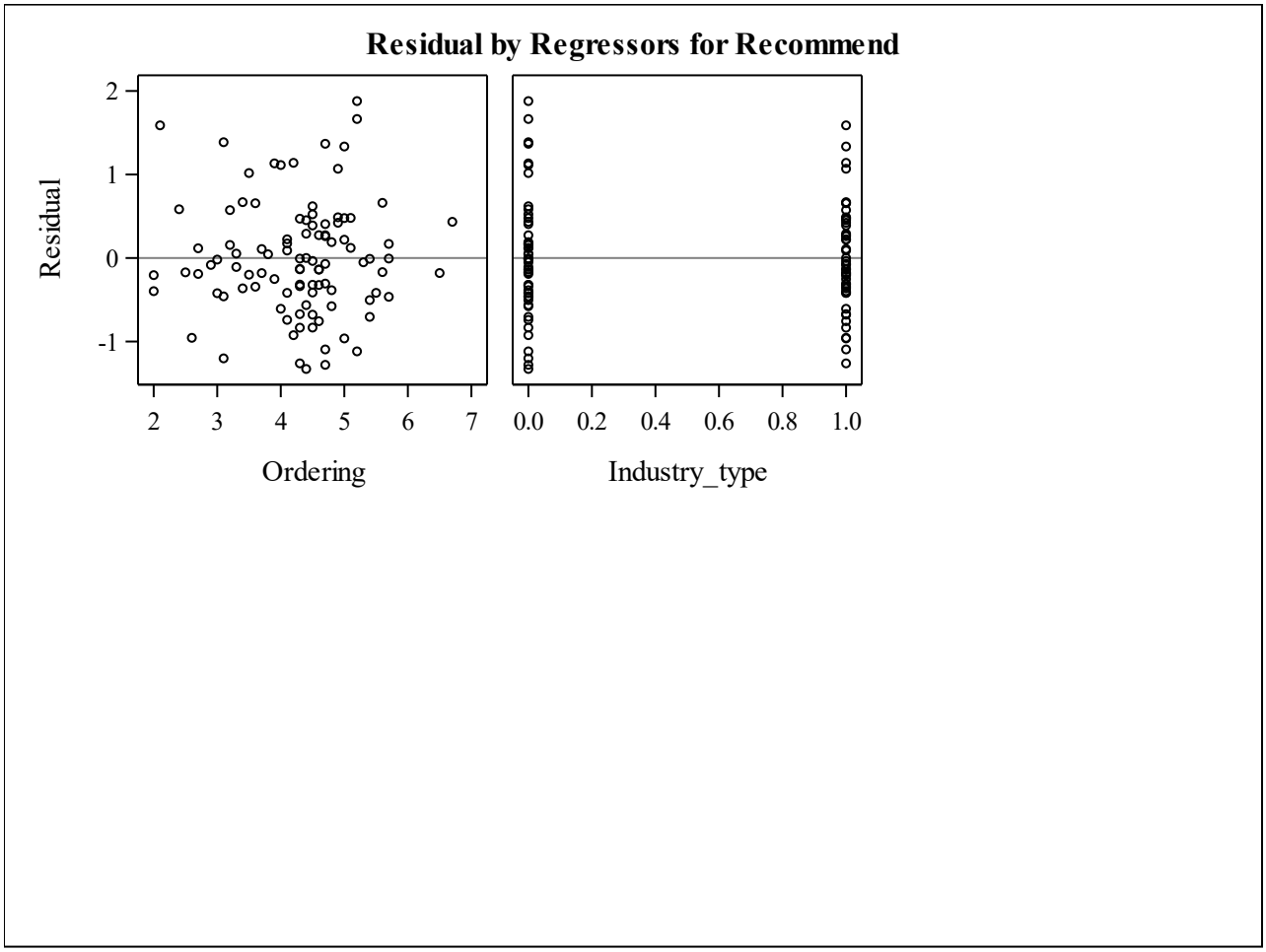




The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

DFBETAS Key
RED = Outliers
GREEN = Leverage



Residuals for the metric variables Warranty and Ordering are normally distributed with about 68% of the residuals should be within 1 standard deviation of the mean of 0, 95% within 2 standard deviations, and about 99% within 3 standard deviations. Residuals for the non-metric qualitative variable Industry_type indicate a linear relationship.

The SAS System

The REG Procedure
Model: TESTING DUMMY VARIABLES ON FINAL MODEL
Dependent Variable: Recommend

Number of Observations Read	100
Number of Observations Used	100

MR2 adding two non-metric dummy variables, newcust, and medcust to determine if they are significantly related to the dependent variable.

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	10	64.05224	6.40522	13.04	<.0001
Error	89	43.70776	0.49110		
Corrected Total	99	107.76000			

H₀: $\beta_1 = \beta_2 = \beta_3 = 0$;
H₁: not all $\beta_j = 0$

$F_{10,89}$=13.04

P-Value <0.0001. Reject is Null.

At least one independent variable helps to explain recommend.

Root MSE	0.70078	**R-Square**	0.5944
Dependent Mean	7.02000	**Adj R-Sq**	0.5488
Coeff Var	9.98268		

H₀: $\beta_k = 0$; Where k = each independent variable.
H₁: $\beta_k \neq 0$

At α=0.10, Prod_Qual, Tech_support, Sales_Image, Warranty, and Ordering are significant. Newcust and medcust have to be tested as a group.

VIF
All VIF values are less than 10 indicating low multicollinearity.

Parameter Estimates						
Variable	**DF**	**Parameter Estimate**	**Standard Error**	**t Value**	**Pr > |t|**	**Variance Inflation**
Intercept	1	2.28626	1.08635	2.10	0.0382	0
Prod_Qual	1	0.27688	0.06953	3.98	0.0001	1.90004
Ecommerce	1	-0.29099	0.17835	-1.63	0.1063	3.14669
Tech_support	1	0.19464	0.07896	2.47	0.0156	2.94401
Prod_Line	1	0.12736	0.07818	1.63	0.1068	2.13137
Sales_Image	1	0.54619	0.12043	4.54	<.0001	3.36211
Warranty	1	-0.33071	0.15297	-2.16	0.0333	3.16958
Ordering	1	0.24213	0.09763	2.48	0.0150	1.65775
Industry_type	1	0.22240	0.14554	1.53	0.1300	1.07652
newcust	1	-0.29590	0.28737	-1.03	0.3060	3.65913
medcust	1	0.02483	0.23171	0.11	0.9149	2.48713

Measures of fit

The R^2 value of 0.5944 suggests that the regression model explains a good amount of the variability in the dependent variable.

The **adjusted R^2** value of 0.5488 is lower than the R^2 value of 0.5944 suggesting that some of the predictors may not be contributing much to explain the variance in the dependent variable.

The **Root MSE** value suggests that the regression model's predictions are off by 0.70078 which is relatively low, indicating that predictions are close to the actual values.

Test to include Customer_Type in the Final Model.

$F_{2,89}$=0.855142
P-Value 0.35760 > α=0.10. fail to reject the null. Type 2 error is possible.
Newcust and medcust do not help explain Recommend and will not be included.

The SAS System

The REG Procedure
Model: VALIDATION MODEL / SPLIT PT 1
Dependent Variable: Recommend

Stepwise Selection: Step 1

Number of Observations Read	50
Number of Observations Used	50

Stepwise Selection: Step 1

Variable Prod_Qual Entered: R-Square = 0.1873 and C(p) = 11.0135

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	12.19204	12.19204	11.06	0.0017
Error	48	52.90796	1.10225		
Corrected Total	49	65.10000			

Variable	**Parameter Estimate**	**Standard Error**	**Type II SS**	**F Value**	**Pr > F**
Intercept	4.45412	0.80338	33.88098	30.74	<.0001
Prod_Qual	0.34361	0.10332	12.19204	11.06	0.0017

Bounds on condition number: 1, 1

Stepwise Selection: Step 2

Variable Ecommerce Entered: R-Square = 0.3380 and C(p) = 2.4436

The SAS System

The REG Procedure
Model: VALIDATION MODEL / SPLIT PT 1
Dependent Variable: Recommend

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	2	22.00078	11.00039	12.00	<.0001
Error	47	43.09922	0.91700		
Corrected Total	49	65.10000			

Variable	**Parameter Estimate**	**Standard Error**	**Type II SS**	**F Value**	**Pr > F**
Intercept	2.01749	1.04499	3.41797	3.73	0.0596
Prod_Qual	0.36466	0.09446	13.66787	14.90	0.0003
Ecommerce	0.60590	0.18526	9.80874	10.70	0.0020

Bounds on condition number: 1.0047, 4.0187

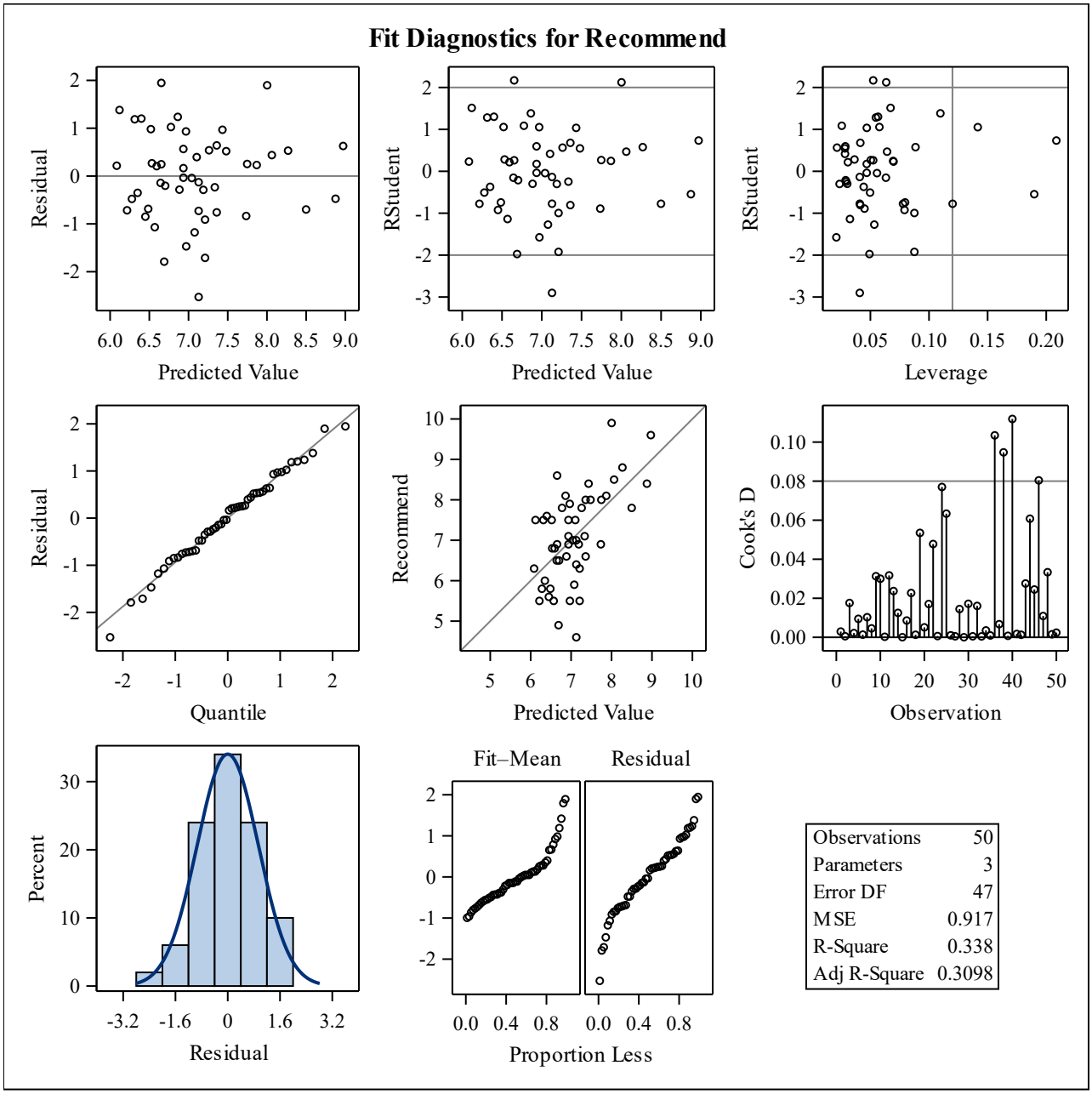
All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Selection								
Step	**Variable Entered**	**Variable Removed**	**Number Vars In**	**Partial R-Square**	**Model R-Square**	**C(p)**	**F Value**	**Pr > F**
1	Prod_Qual		1	0.1873	0.1873	11.0135	11.06	0.0017
2	Ecommerce		2	0.1507	0.3380	2.4436	10.70	0.0020

The SAS System

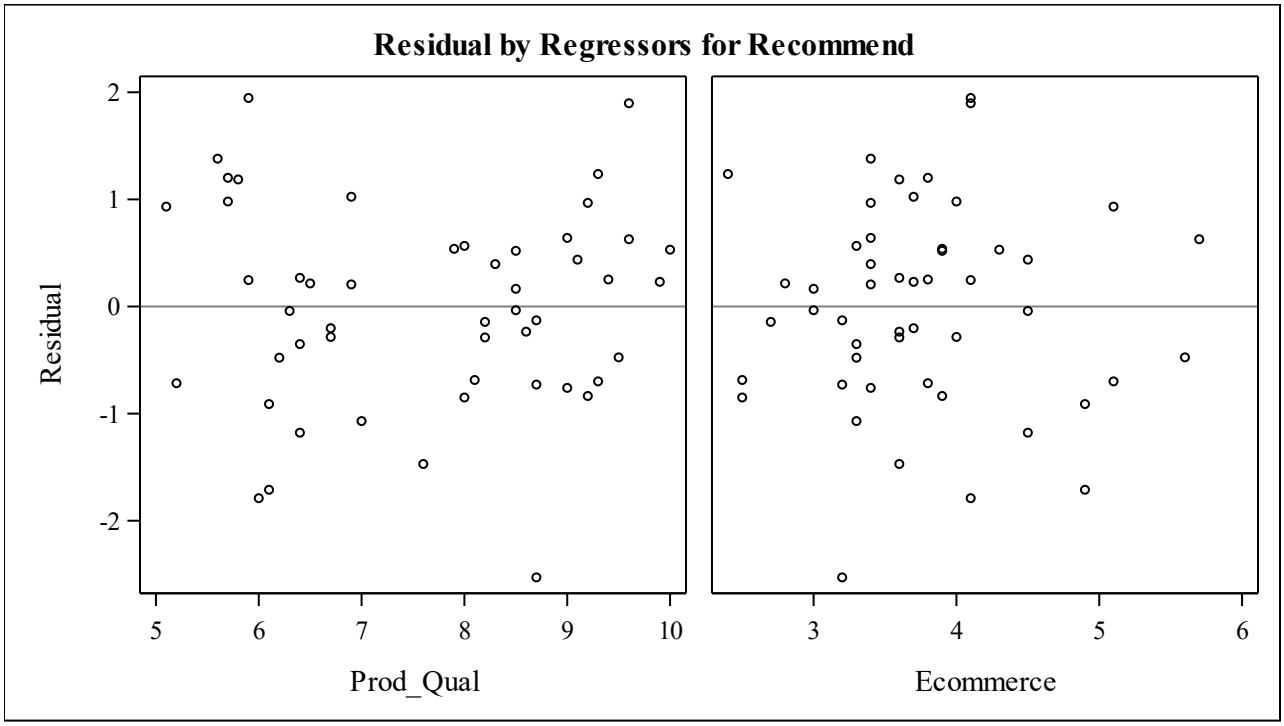
The REG Procedure
Model: VALIDATION MODEL / SPLIT PT 1
Dependent Variable: Recommend



There is a positive linear relationship in the normal probability plot and the slope is close to 1, residuals versus quantiles. The skew is nearly symmetrical, and the data appears normally distributed. The histogram appears to have a normal distribution with a positive kurtosis. The RF plot consisting of side-by-side quantile plots of the centered fit and the residuals shows that the spread in the residuals is no greater than the spread in the centered fit. The residual versus predicted value appears to be randomized with no noticeable pattern, suggesting homoscedasticity, upholding the assumption of constant variance.

The SAS System

The REG Procedure
Model: VALIDATION MODEL / SPLIT PT 1
Dependent Variable: Recommend



Residuals for Prod_Qual and Ecommerce are normally distributed with about 68% of the residuals should be within 1 standard deviation of the mean of 0; 95% within 2 standard deviation, and about 99% within 3 standard deviations. Validating the assumption of central tendency.

The SAS System

The REG Procedure
Model: VALIDATION MODEL / SPLIT PT 2
Dependent Variable: Recommend

Number of Observations Read	50
Number of Observations Used	50

Stepwise Selection: Step 1

Variable Ordering Entered: R-Square = 0.2831 and C(p) = 1.1965

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	1	11.97719	11.97719	18.96	<.0001
Error	48	30.32281	0.63173		
Corrected Total	49	42.30000			

Variable	**Parameter Estimate**	**Standard Error**	**Type II SS**	**F Value**	**Pr > F**
Intercept	4.80419	0.50770	56.56496	89.54	<.0001
Ordering	0.52504	0.12058	11.97719	18.96	<.0001

Bounds on condition number: 1, 1

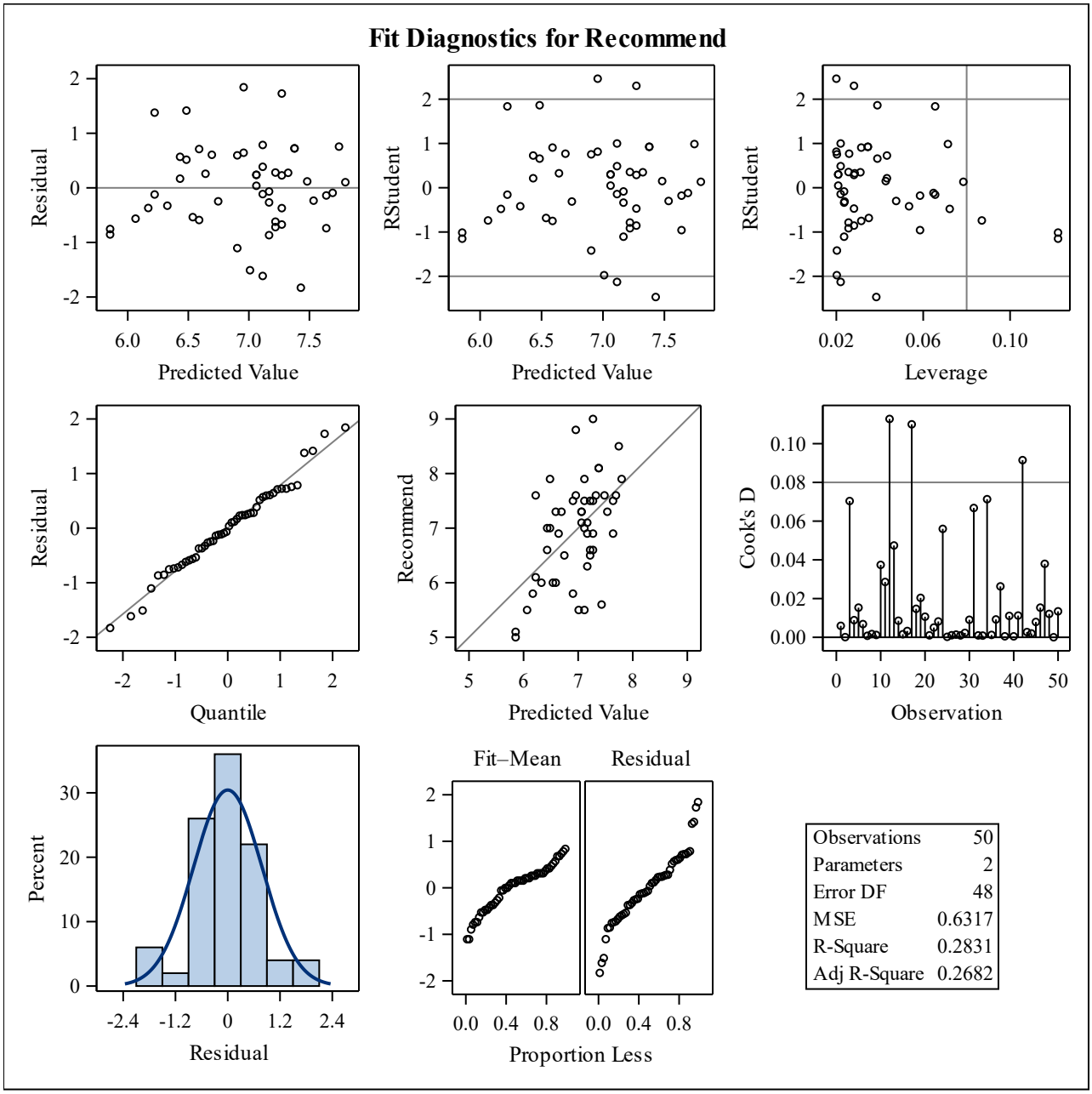
All variables left in the model are significant at the 0.1500 level.

No other variable met the 0.1500 significance level for entry into the model.

Summary of Stepwise Selection								
Step	**Variable Entered**	**Variable Removed**	**Number Vars In**	**Partial R-Square**	**Model R-Square**	**C(p)**	**F Value**	**Pr > F**
1	Ordering		1	0.2831	0.2831	1.1965	18.96	<.0001

The SAS System

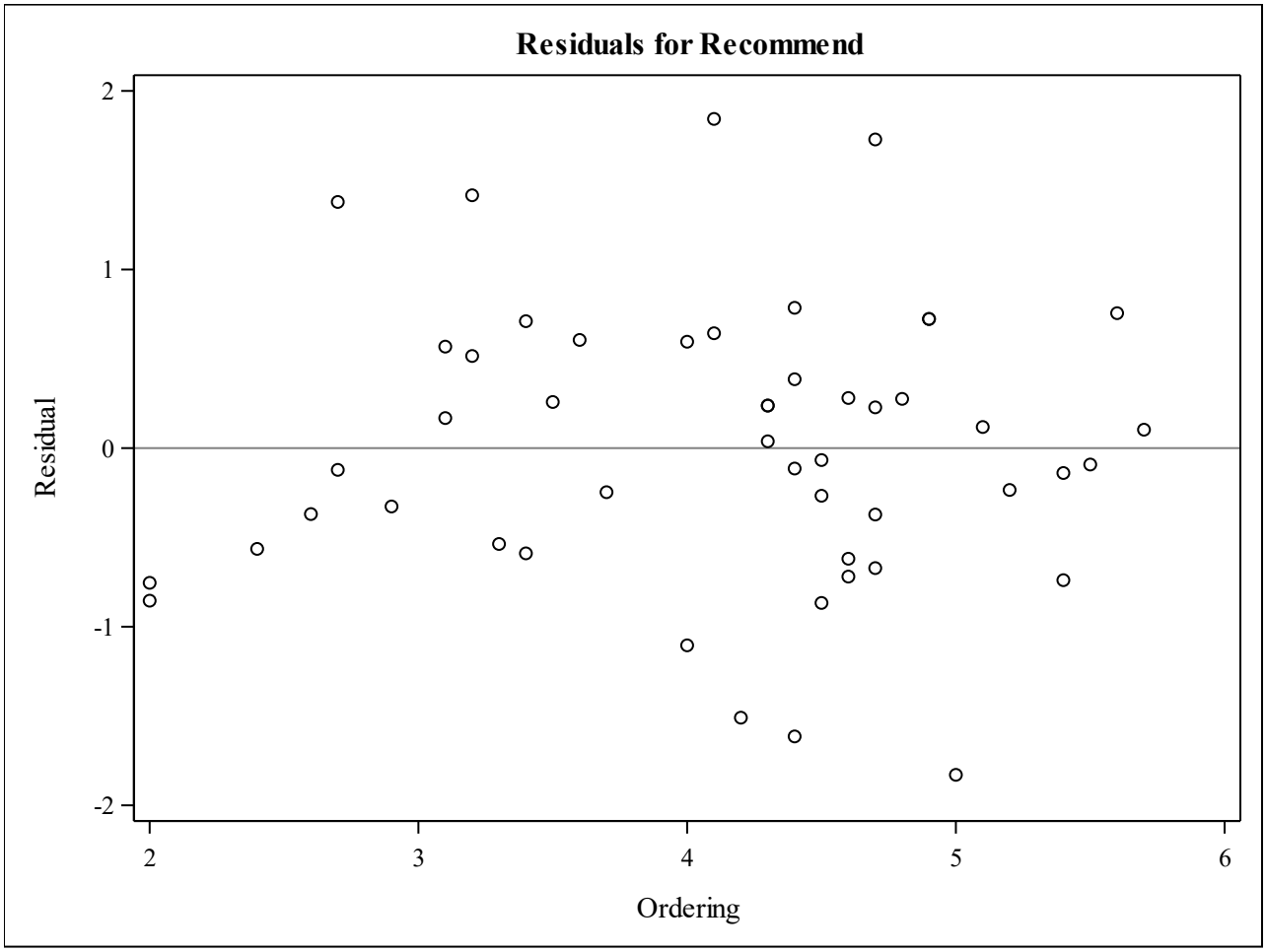
The REG Procedure
Model: VALIDATION MODEL / SPLIT PT2
Dependent Variable: Recommend



There is a positive linear relationship in the normal probability plot and the slope is close to 1, residuals versus quantiles. The skew is nearly symmetrical, and the data appears normally distributed. The histogram appears to have a normal distribution. The RF plot consisting of side-by-side quantile plots of the centered fit and the residuals shows that the spread in the residuals is no greater than the spread in the centered fit. The residual versus predicted value appears to be randomized with no noticeable pattern, suggesting homoscedasticity, upholding the assumption of constant variance.

The SAS System

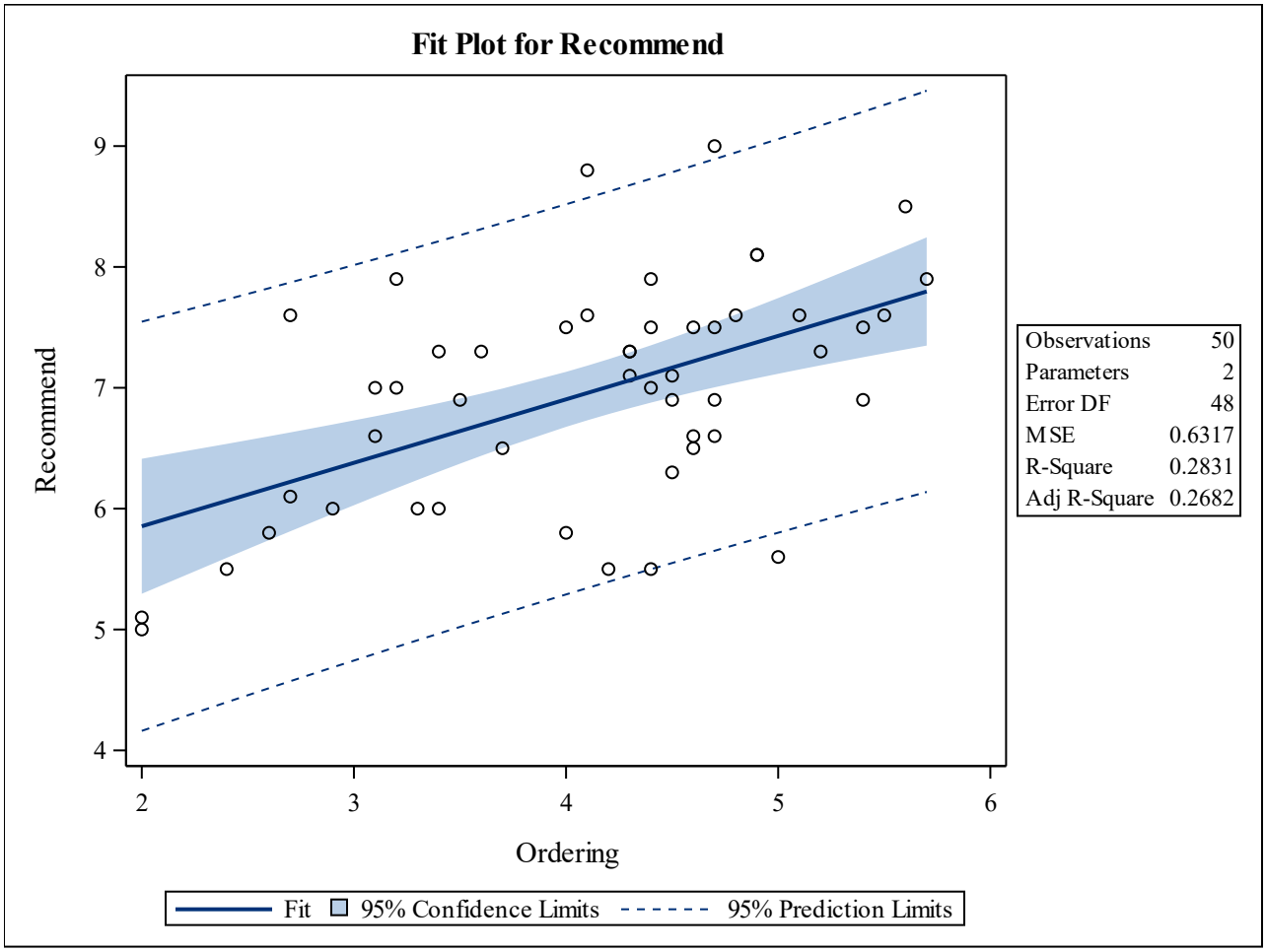
The REG Procedure
Model: VALIDATION MODEL / SPLIT PT2
Dependent Variable: Recommend



Residuals are normally distributed with about 68% of the residuals should be within 1 standard deviation of the mean of 0; 95% within 2 standard deviation, and about 99% within 3 standard deviations. Validating the assumption of central tendency.

The SAS System

The REG Procedure
Model: VALIDATION MODEL / SPLIT PT2
Dependent Variable: Recommend



The fitted plot shows normality with three outliers.

End of Split / Validation

The SAS System

The REG Procedure
Model: FINAL MODEL
Dependent Variable: Recommend

Final Model Summary

Research Objective: To determine which metric independent variables can be used as a set of predictors to best explain the value of the dependent variable 'Recommend'. Using multiple regression to determine which model will best explain the likeliness a customer will recommend HBAT industries.

Regression Equation:
$\hat{Y}_i = 1.73063 + 0.29482Prod_{Qual} - 0.37291Ecommerce +$ 0.21454Tech_support + 0.15685Prod_Line + 0.61085Sales_Image - 0.375Warranty + 0.30644Ordering + 0.25092Industry_type

The variables in the Final Model; Prod_Qual, Ecommerce, Tech_support, Prod_Line, Sales,_Image, Warranty, Ordering, Industry_type are approximately normally distributed, and it measures the average amount of time it takes to deliver paper products once an order has been confirmed. The regression equation suggests a willingness to recommend at 1.73063 when all of the dependant variables are zero. The variable with the most impact on recommend is Sales_Image.

The final model selected includes the variables **Prod_Qual, Ecommerce, Tech_support, Prod_Line, Sales,_Image, Warranty, Ordering, Industry_type.** The variable **Adv** was removed because it may not be contributing much to explain the variance in the dependent variable. There were some noteworthy observations found in the final model Cook's D plot that warranted further study, so the model was split into two sets of 50 observations and run for validation purposes. The findings of the validation model were…?

Regression Equation:
$\hat{Y}_i = 1.73063 + 0.29482Prod_{Qual} - 0.37291Ecommerce +$ 0.21454Tech_support + 0.15685Prod_Line + 0.61085Sales_Image - 0.375Warranty + 0.30644Ordering + 0.25092Industry_type

The SSE of 44.7953 and the corresponding PRESS value of 53.86246 suggest that while my model explains a certain amount of variance within the sample data, it may not generalize well to new observations. This indicates potential limitations in the model's predictive performance on unseen data, which could be further investigated and addressed to improve the overall robustness and applicability of the regression model.