Vivian Ellis

DSCI 449

Multiple Regression Part 1

Multiple regression will take a set of independent variables as possible predictors of a single dependent variable. Using the HBAT data set we will choose the likely hood of recommending HBAT to other firms (recommend) as the dependent variable and 13 metric variables that represent the perceptions of HBAT as the independent variables. Bock MR-1 describes the simple statistics for all the metric variables.

**Block MR-1 Statistics**

|  |  |
| --- | --- |
| **14 Variables:** | Recommend Prod\_Qual Ecommerce Tech\_support Complaint Adv Prod\_Line Sales\_Image Pricing Warranty New\_Prod Ordering Price\_Flex Del\_Speed |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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 |

The correlation matrix in Block MR-1 Correlation Matrix.

Block MR-1 Correlation Matrix

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **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 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Pearson Correlation Coefficients, N = 100 Prob > |r| under H0: Rho=0** | | | | | | | |
|  | **Sales\_Image** | **Pricing** | **Warranty** | **New\_Prod** | **Ordering** | **Price\_Flex** | **Del\_Speed** |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |
| **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 |

We will focus on the Pearson’s Correlation coefficient between the dependent variable, recommend, and all the independent variables. Where r = 0.44717 is the correlation coefficient for recommend and product quality. This positive correlation means that when a customer rates the level of quality of HBAT’s paper products, on a 10-point scale, they are about 44.71% likely to recommend HBAT products to other firms. With a p value of < .0001 and,

H\_0 : there exists no correlation between recommend and product quality

H\_a : there exists a correlation between recommend and product quality

we reject H\_0 and accept the alternative. Recommend and product quality are related. The closer r, the correlation coefficient, is to 1 the stronger the correlation is. The following variables have a p-value <.0001 meaning that there exists a positive correlation with recommend, but they are not a *strong* correlation. Recommend and complaint are 45.06% correlated, recommend and product line are 44.52% correlated, recommend and ordering are 43.86% correlated and del speed and recommend are 44.62% correlated.

There only exists two negatively correlated variables to recommend, pricing and new product. However, r= -0.00169, the negative correlation between recommend and new product, is closer to 0 meaning that there exists no relationship between the two variables. The p-value is 0.9867 and,

H\_0 : there exists no correlation between recommend and new product

H\_a : there exists a correlation between recommend and new product

we accept the null hypothesis that recommend, and new product are not related. The p-value for pricing is .1488, thus we can conclude that there is not relationship between recommend and pricing.

Block MR-2 gives the full model regression, with adjusted , , and a root MSE of .72253. There are 13 independent variables with p=# of parameters=# of variables +1, thus p=14. Of the independent variables, product line, price flex and deliver speed, have a variance inflation greater than 10. With,

H\_0 :

H\_a : not all

, p-value <.0001

We reject the null hypothesis meaning, at least one variable explains recommend.

Block MR-2

|  |  |
| --- | --- |
| **Number of Observations Read** | 100 |
| **Number of Observations Used** | 100 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Analysis of Variance** | | | | | |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 13 | 62.86375 | 4.83567 | 9.26 | <.0001 |
| **Error** | 86 | 44.89625 | 0.52205 |  |  |
| **Corrected Total** | 99 | 107.76000 |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Root MSE** | 0.72253 | **R-Square** | 0.5834 |
| **Dependent Mean** | 7.02000 | **Adj R-Sq** | 0.5204 |
| **Coeff Var** | 10.29245 |  |  |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Parameter Estimates** | | | | | | |
| **Variable** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** | **Variance Inflation** |
| **Intercept** | 1 | 1.26703 | 1.42928 | 0.89 | 0.3778 | 0 |
| **Prod\_Qual** | 1 | 0.32156 | 0.06724 | 4.78 | <.0001 | 1.67169 |
| **Ecommerce** | 1 | -0.32464 | 0.17416 | -1.86 | 0.0657 | 2.82256 |
| **Tech\_support** | 1 | 0.21387 | 0.08283 | 2.58 | 0.0115 | 3.04740 |
| **Complaint** | 1 | -0.06125 | 0.13217 | -0.46 | 0.6442 | 4.83774 |
| **Adv** | 1 | -0.09721 | 0.08016 | -1.21 | 0.2285 | 1.54742 |
| **Prod\_Line** | 1 | 0.18849 | 0.34024 | 0.55 | 0.5810 | 37.97842 |
| **Sales\_Image** | 1 | 0.61187 | 0.12944 | 4.73 | <.0001 | 3.65361 |
| **Pricing** | 1 | 0.02892 | 0.06150 | 0.47 | 0.6393 | 1.71202 |
| **Warranty** | 1 | -0.34774 | 0.16015 | -2.17 | 0.0327 | 3.26841 |
| **New\_Prod** | 1 | -0.02534 | 0.05044 | -0.50 | 0.6167 | 1.07545 |
| **Ordering** | 1 | 0.25530 | 0.13334 | 1.91 | 0.0589 | 2.90906 |
| **Price\_Flex** | 1 | 0.07442 | 0.34763 | 0.21 | 0.8310 | 33.33234 |
| **Del\_Speed** | 1 | 0.09665 | 0.65589 | 0.15 | 0.8832 | 44.00376 |

Block MR-3 is available in the Appendix. To determine the “best” model the criteria is as follows, the highest adjusted , lowest standard error, and parsimony. Figure 1 represents the fit criterion for recommend,



Figure 1

Where step 7 model produces the highest adjusted . In figure 2 shows that the 7th step model has the “best” cp statistic.



Figure 2

The “best” regression model is,

At step 7, the bounds on condition number: 2.7469, 48.346. 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. The model produces a 5.4904, and an adjusted 0.5389.

**Block MR-4**

|  |  |
| --- | --- |
| **Number of Observations Read** | 100 |
| **Number of Observations Used** | 100 |

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 6 | 57.69526 | 9.61588 | 17.86 | <.0001 |
| **Error** | 93 | 50.06474 | 0.53833 |  |  |
| **Corrected Total** | 99 | 107.76000 |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Root MSE** | 0.73371 | **R-Square** | 0.5354 |
| **Dependent Mean** | 7.02000 | **Adj R-Sq** | 0.5054 |
| **Coeff Var** | 10.45171 |  |  |

| **Parameter Estimates** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** | **Variance Inflation** |
| **Intercept** | 1 | 0.97516 | 0.66173 | 1.47 | 0.1440 | 0 |
| **Complaint** | 1 | 0.06214 | 0.10700 | 0.58 | 0.5628 | 3.07424 |
| **Prod\_Qual** | 1 | 0.30180 | 0.06196 | 4.87 | <.0001 | 1.37644 |
| **Sales\_Image** | 1 | 0.53230 | 0.11680 | 4.56 | <.0001 | 2.88493 |
| **Ecommerce** | 1 | -0.28197 | 0.17364 | -1.62 | 0.1078 | 2.72096 |
| **Prod\_Line** | 1 | 0.11915 | 0.07996 | 1.49 | 0.1396 | 2.03392 |
| **Ordering** | 1 | 0.22590 | 0.12211 | 1.85 | 0.0675 | 2.36581 |

In Block MR-4 all variance of inflations is less than 10. The regressions diagnostics may be seen in Block MR-4 of the Appendix. In plot 7 below, the histogram appears to be normal.

