|  |  |
| --- | --- |
| /\*Setting data set to a new data\*/ | |
|  | |  | |
|  | data HBAT\_200\_1; | |
|  | set WORK.HBAT\_200; | |
|  | x19logsatis = log(x19satis); | |
|  | run; | |
|  |  | |
|  | proc contents data = HBAT\_200\_1 order=varnum; | |
|  | run; | |
|  |  | |
|  | /\*exploratory analysis of numeric variables\*/ | |
|  |  | |
|  | proc means data= hbat\_200\_1 N NMISS STDDEV MEAN MAX MIN ; | |
|  | run; | |
|  | proc freq data=HBAT\_200\_1; | |
|  | run; | |
|  | proc freq data=HBAT\_200\_1; | |
|  | table x19satis ; | |
|  | run; | |
|  |  | |
|  | /\*scatter plot\*/ | |
|  | proc sgplot data=HBAT\_200\_1; | |
|  | scatter x=x19satis y=x6prodqual; | |
|  | run; | |
|  | /\*scatter plot matrix for product variables\*/ | |
|  | proc sgscatter data=HBAT\_200\_1; | |
|  | title 'Scatter plot'; | |
|  | matrix x19satis x6prodqual x11prodline x13compric x15newprod x17pricflex/ group=x19satis diagonal=(histogram kernel); | |
|  | where (x6prodqual< 9.9); | |
|  | run; | |
|  | proc freq data=HBAT\_200\_1; | |
|  | where (x6prodqual< 9.9); | |
|  | table x6prodqual ; | |
|  | run; | |
|  |  | |
|  | title; | |
|  | /\*scatter plot matrix for all ECOM MARKETING variables\*/ | |
|  | proc sgscatter data=HBAT\_200\_1; | |
|  | title 'Scatter plot'; | |
|  | matrix x19satis x7ecomm x10adv x12sales/ group=x19satis diagonal=(histogram kernel); | |
|  | run; | |
|  | title; | |
|  | /\*scatter plot matrix for all customer satisfaction variables\*/ | |
|  | proc sgscatter data=HBAT\_200\_1; | |
|  | title 'Scatter plot'; | |
|  | matrix x19satis x8techsup x14warran x15order x18delsped/ group=x19satis diagonal=(histogram kernel); | |
|  | run; | |
|  | title; | |
|  |  | |
|  | /\*scatter plot matrix for all variables\*/ | |
|  | proc sgscatter data=HBAT\_200\_1; | |
|  | title 'Scatter plot'; | |
|  | matrix x19satis x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped/ group=x19satis diagonal=(histogram kernel); | |
|  | run; | |
|  | title; | |
|  |  | |
|  | /\*scatter plot matrix for x6 to x11 variables-cHECK FOR MODES\*/ | |
|  | proc sgscatter data=HBAT\_200\_1; | |
|  | title 'Scatter plOT'; | |
|  | matrix x19satis x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline/ group=x19satis diagonal=(histogram kernel); | |
|  | run; | |
|  | /\*scatter plot matrix for x12 to x18 variables-cHECK FOR MODES\*/ | |
|  | proc sgscatter data=HBAT\_200\_1; | |
|  | title 'Scatter plot'; | |
|  | matrix x19satis x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped/ group=x19satis diagonal=(histogram kernel); | |
|  | run; | |
|  | title; | |
|  | ods graphics on; | |
|  | ods trace on; | |
|  | proc | |
|  | /\*bOX plot X1\*/ | |
|  | proc sort data = hbat\_200\_1; | |
|  | by x1custype; | |
|  | run; | |
|  | goptions reset = all; | |
|  | proc boxplot data = hbat\_200\_1; | |
|  | title1 'Boxplot of Satis Score vs. X1'; | |
|  | plot x19satis\*x1custype /boxconnect=mean boxstyle=schematic ; | |
|  | run; | |
|  | /\*bOX plot x2\*/ | |
|  | proc sort data = hbat\_200\_1; | |
|  | by x2industype; | |
|  | run; | |
|  | goptions reset = all; | |
|  | proc boxplot data = hbat\_200\_1; | |
|  | title1 'Boxplot of Satis Score vs. 2'; | |
|  | plot x19satis\*x2industype /boxconnect=mean boxstyle=schematic ; | |
|  | run; | |
|  | /\*bOX plot x3\*/ | |
|  | proc sort data = hbat\_200\_1; | |
|  | by x3firmsize; | |
|  | run; | |
|  | goptions reset = all; | |
|  | proc boxplot data = hbat\_200\_1; | |
|  | title1 'Boxplot of Satis Score vs. X3'; | |
|  | plot x19satis\*x3firmsize /boxconnect=mean boxstyle=schematic ; | |
|  | run; | |
|  | /\*bOX plot x4\*/ | |
|  | proc sort data = hbat\_200\_1; | |
|  | by x4region; | |
|  | run; | |
|  | goptions reset = all; | |
|  | proc boxplot data = hbat\_200\_1; | |
|  | title1 'Boxplot of Satis Score vs. X4'; | |
|  | plot x19satis\*x4region /boxconnect=mean boxstyle=schematic ; | |
|  | run; | |
|  | /\*Box plot x5\*/ | |
|  | proc sort data = hbat\_200\_1; | |
|  | by x5distsys; | |
|  | run; | |
|  | goptions reset = all; | |
|  | proc boxplot data = hbat\_200\_1; | |
|  | title1 'Boxplot of Satis Score vs. X5'; | |
|  | plot x19satis\*x5distsys/boxconnect=mean boxstyle=schematic ; | |
|  | run; | |
|  | proc insight data =HBAT\_200\_1; | |
|  | scatter x19satis x6prodqual x7ecomm x8techsup \* | |
|  | x19satis x6prodqual x7ecomm x8techsup ; | |
|  | run; | |
|  | quit; | |
|  | /\*correlation matrix for product data\*/ | |
|  | proc corr data=HBAT\_200\_1 noprob plots=matrix(histogram); | |
|  | variable x19satis x6prodqual x11prodline x13compric x15newprod x17pricflex; | |
|  | run; | |
|  | ods graphics on; | |
|  | /\*correlation matrix for ecomm data\*/ | |
|  | proc corr data=HBAT\_200\_1 noprob plots=matrix(histogram); | |
|  | variable x19satis x7ecomm x10adv x12sales ; | |
|  | run; | |
|  | ods graphics on; | |
|  | /\*correlation matrix for customer satisfaction data\*/ | |
|  | proc corr data=HBAT\_200\_1 noprob plots=matrix(histogram); | |
|  | variable x19satis x8techsup x14warran x15order x18delsped ; | |
|  | run; | |
|  | ods graphics on; | |
|  | /\*correlation matrix for customer satisfaction data\*/ | |
|  | proc corr data=HBAT\_200\_1 noprob plots=matrix(histogram); | |
|  | run; | |
|  | ods graphics on; | |
|  |  | |
|  | proc univariate data=hbat\_200\_1 all; | |
|  | run; | |
|  | proc reg data=hbat\_200\_1; | |
|  | model x19satis = x6prodqual x7ecomm; | |
|  | PLOT x19satis\*x6prodqual x19satis\*x7ecomm; | |
|  | run; | |
|  |  | |
|  | ods graphics on; | |
|  | proc corr data=work.wine\_data cov plots(maxpoints= NONE)=matrix(Histogram); | |
|  | var quality alcohol chlorides 'citric acid'n density; | |
|  | ods select Matrixplot; | |
|  | run; | |
|  | ods graphics off; | |
|  | ods graphics on; | |
|  | proc univariate data=work.wine\_data plots normal noprint; | |
|  | histogram ; | |
|  | qqplot; | |
|  | run; | |
|  |  | |
|  | ods graphics off; | |
|  | /\*Variation inflation method \*/ | |
|  | /\* it was found that density, residual sugar, fixed acidity and alcohol have multicollinarity effect\*/ | |
|  | ods graphics on; | |
|  | /\* vif and Residual\*/ | |
|  | proc reg data=; | |
|  | model quality= alcohol chlorides 'citric acid'n density 'fixed acidity'n 'free sulfur dioxide'n pH 'residual sugar'n sulphates 'total sulfur dioxide'n 'volatile acidity'n/ vif ; | |
|  | plot residual.\*predicted.; | |
|  | run; | |
|  | ods graphics off; | |
|  | ods graphics on; | |
|  | /\*Stepwise\*/ | |
|  | proc reg data=HBAT\_200\_1 plots(only)= All; | |
|  | model x19satis = x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline/ selection= stepwise sle=0.15 sls=0.15 ; | |
|  | run; | |
|  | ods graphics off; | |
|  | ods graphics on; | |
|  |  | |
|  |  | |
|  |  | |
|  | /\* Coment on CL\*/ | |
|  | proc reg data=work.wine\_data plots=ResidualByPredicted ; | |
|  | model quality = alcohol / r clm cli ; | |
|  | run; | |
|  | ods graphics off; | |
|  |  | |
|  |  | |
|  |  | |
|  | ods graphics on; | |
|  | /\* cHECK glm vs Reg\*/ | |
|  | proc glmselect data=work.wine\_data plots=all; | |
|  | model quality= alcohol|chlorides|'citric acid'n|density|'fixed acidity'n|'free sulfur dioxide'n|pH 'residual sugar'n|sulphates|'total sulfur dioxide'n|'volatile acidity'n @2 /Selection= stepwise( select=sl sle=0.15 sls=0.15) stats=all; | |
|  | run; | |
|  |  | |
|  | ods graphics off; | |
|  |  | |
|  | ODS GRAPHICS ON ; | |
|  | /\* after univariate-lEVERAGE COVARIANCE COOKS\*/ | |
|  |  | |
|  | PROC GLM DATA =work.wine\_data PLOTS=all ; | |
|  | MODEL quality = alcohol / P ; | |
|  | OUTPUT OUT = stat P=pred R=Residual DFFITS=dffits COOKD=cookd H=hatvalue PRESS=res\_del ; | |
|  | RUN ; | |
|  | ODS GRAPHICS OFF ; | |
|  | ODS GRAPHICS ON; | |
|  |  | |
|  | PROC TRANSREG DATA = work.wine\_data TEST; | |
|  | MODEL BOXCOX(quality) = IDENTITY(alcohol density pH chlorides); | |
|  | RUN; | |
|  | ODS GRAPHICS OFF; | |
|  |  | |
|  | /\*Test for muticollinearity \*/ | |
|  |  | |
|  | proc reg data=HBAT\_200\_1 plots(only)= All; | |
|  | model x19satis = x2industype x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped/tol vif collin; | |
|  | run; | |
|  |  | |
|  | /\*Test for indepenpendance\*/ | |
|  |  | |
|  | proc reg data=HBAT\_200\_1; | |
|  | model x19satis = x2industype x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex/ dw; | |
|  | output out=res3 (keep = snum r id) residual=r; | |
|  | run; | |
|  | quit; | |
|  | goptions reset=all; | |
|  | proc gplot data=res3; | |
|  | plot r\*id; | |
|  | run; | |
|  | quit; | |
|  | /\* residual are normally \*/ | |
|  | proc reg data=HBAT\_200\_1; | |
|  | model x19satis = x2industype x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex; | |
|  | output out=elem1res (keep= r fv) residual=r predicted=fv; | |
|  | run; | |
|  | quit; | |
|  | proc kde data=elem1res out=den; | |
|  | var r; | |
|  | run; | |
|  |  | |
|  | proc sort data=den; | |
|  | by r; | |
|  | run; | |
|  |  | |
|  | goptions reset=all; | |
|  | symbol1 c=blue i=join v=none height=1; | |
|  | proc gplot data=den; | |
|  | plot density\*r=1; | |
|  | run; | |
|  | quit; | |
|  | goptions reset=all; | |
|  | proc univariate data=elem1res normal; | |
|  | var r; | |
|  | qqplot r / normal(mu=est sigma=est); | |
|  | run; | |
|  |  | |
|  | /\* Test for Hetro\*/ | |
|  | proc reg data=hbat\_200\_1; | |
|  | model x19satis = x2industype x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex; | |
|  | plot r.\*p.; | |
|  | run; | |
|  | quit; | |
|  | /\*WHITE TEST\*/ | |
|  | proc reg data=hbat\_200\_1; | |
|  | model x19satis = x2industype x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex / spec; | |
|  | run; | |
|  | quit; | |
|  | /\* Models \*/ | |
|  | proc reg data=HBAT\_200\_1 plots(only)= All; | |
|  | var x19satis x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped; | |
|  | /\* Estimating the effect of only ecomm/marketing variables\*/ | |
|  | M1: Model x19satis = x12sales X7ecomm X10adv; | |
|  | /\* Explains 22% -X12 is only significant and alone explains 22% variance \*/ | |
|  | run; | |
|  | /\* Only Product\*/ | |
|  | M2: Model x19satis= x11prodline x6prodqual x15newprod ; | |
|  | /\* Explains 47% -X6 and x11 are only significant. x11 alone explains 41% variance, x6 plus x11 explains 46% \*/ | |
|  | /\* Only Customer support2\*/ | |
|  | M3: Model x19satis= x18delsped x9compl x15order; | |
|  | /\* Explains 45%. x9 alone explains 35%, x18 alone explains 39% \*/ | |
|  | run; | |
|  | ods graphics off; | |
|  | /\*Lean Model\*/ | |
|  | M4: Model x19satis=x9compl x11prodline x12sales; | |
|  | run; | |
|  | ods graphics on; | |
|  | /\*PROC GLM\*/ | |
|  | proc GLM data=HBAT\_200\_1 plots(only)= All; | |
|  | CLASS x1custype; | |
|  | Model x19SATIS = x1custype x12sales X7ecomm X10adv; | |
|  | /\* Explains 22% -X12 is only significant and alone explains 22% variance \*/ | |
|  | run; | |
|  | /\*Backward\*/ | |
|  | DATA WORK.SORTTempTableSorted &\_EG\_DSTYPE\_ / VIEW=WORK.SORTTempTableSorted; | |
|  | SET WORK.HBAT\_200(KEEP=x19satis x1custype x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped &\_DSTYPE\_VARS\_); | |
|  | RUN; | |
|  | TITLE; | |
|  | TITLE1 "Linear Regression Results"; | |
|  | FOOTNOTE; | |
|  | FOOTNOTE1 "Generated by the SAS System (&\_SASSERVERNAME, &SYSSCPL) on %TRIM(%QSYSFUNC(DATE(), NLDATE20.)) at %TRIM(%SYSFUNC(TIME(), TIMEAMPM12.))"; | |
|  | PROC REG DATA=WORK.SORTTempTableSorted | |
|  | PLOTS(ONLY)=DIAGNOSTICSPANEL | |
|  | ; | |
|  | Linear\_Regression\_Model: MODEL x19satis = x1custype x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped | |
|  | / SELECTION=backward | |
|  | SLE=0.15 | |
|  | SLS=0.15 | |
|  | INCLUDE=0 | |
|  | ; | |
|  | RUN; | |
|  | QUIT; | |
|  |  | |
|  | /\* ------------------------------------------------------------------- | |
|  | End of task code. | |
|  | ------------------------------------------------------------------- \*/ | |
|  | RUN; QUIT; | |
|  | %\_eg\_conditional\_dropds(WORK.SORTTempTableSorted, | |
|  | WORK.TMP1TempTableForPlots); | |
|  | TITLE; FOOTNOTE; | |
|  | ODS GRAPHICS OFF; | |
|  | DATA WORK.SORTTempTableSorted &\_EG\_DSTYPE\_ / VIEW=WORK.SORTTempTableSorted; | |
|  | SET WORK.HBAT\_200(KEEP=x19satis x1custype x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped &\_DSTYPE\_VARS\_); | |
|  | RUN; | |
|  | TITLE; | |
|  |  | |
|  | TITLE1 "Linear Regression Results"; | |
|  | FOOTNOTE; | |
|  | FOOTNOTE1 "Generated by the SAS System (&\_SASSERVERNAME, &SYSSCPL) on %TRIM(%QSYSFUNC(DATE(), NLDATE20.)) at %TRIM(%SYSFUNC(TIME(), TIMEAMPM12.))"; | |
|  | /\*Forward\*/ | |
|  | PROC REG DATA=WORK.SORTTempTableSorted | |
|  | PLOTS(O/NLY)=DIAGNOSTICSPANEL | |
|  | ; | |
|  | Linear\_Regression\_Model: MODEL x19satis = x1custype x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped | |
|  | / SELECTION=forward | |
|  | SLE=0.15 | |
|  | SLS=0.15 | |
|  | INCLUDE=0 | |
|  | ; | |
|  | RUN; | |
|  | QUIT; | |
|  |  | |
|  | /\* ------------------------------------------------------------------- | |
|  | End of task code. | |
|  | ------------------------------------------------------------------- \*/ | |
|  | RUN; QUIT; | |
|  | %\_eg\_conditional\_dropds(WORK.SORTTempTableSorted, | |
|  | WORK.TMP1TempTableForPlots); | |
|  | TITLE; FOOTNOTE; | |
|  | ODS GRAPHICS OFF; | |
|  | /\*Stepwise\*/ | |
|  | DATA WORK.SORTTempTableSorted &\_EG\_DSTYPE\_ / VIEW=WORK.SORTTempTableSorted; | |
|  | SET WORK.HBAT\_200(KEEP=x19satis x1custype x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped &\_DSTYPE\_VARS\_); | |
|  | RUN; | |
|  | TITLE; | |
|  | TITLE1 "Linear Regression Results"; | |
|  | FOOTNOTE; | |
|  |  | |
|  |  | |
|  | PROC REG DATA=WORK.SORTTempTableSorted | |
|  | PLOTS(ONLY)=DIAGNOSTICSPANEL | |
|  | ; | |
|  | Linear\_Regression\_Model: MODEL x19satis = x1custype x2industype x3firmsize x4region x5distsys x6prodqual x7ecomm x8techsup x9compl x10adv x11prodline x12sales x13compric x14warran x15newprod x15order x17pricflex x18delsped | |
|  | / SELECTION=NONE | |
|  | SLE=0.15 | |
|  | SLS=0.15 | |
|  | INCLUDE=0 | |
|  | ; | |
|  | RUN; | |
|  | QUIT; | |