```
/************************************************************
************************************************************
************************************************************
************************************************************/
/*Daren Purnell, 2017SP_PREDICT_411-DL_SEC60*/
/* Connect Predict411 Data */
libname mydata "/scs/wtm926/" access=readonly;
proc datasets library=mydata;
run;
ods graphics on;
title 'Wine Sales Score Code';
/* Create a copy of wine sales test data */
data test_data;
	set mydata.wine_test;
	TARGET_FLAG = ( TARGET > 0 ); /* 1 if cases sold; 0 if no cases sold */
	TARGET_AMT = TARGET - 1;
	if TARGET_FLAG = 0 then TARGET_AMT = .;
run;
/*IMPUTATIONS*/
data imp_data;
	set test_data;
	IMP_STARS 			= STARS;
	IMP_Sulphates 			= Sulphates;
	IMP_Alcohol 			= Alcohol;
	IMP_TotalSulfurDioxide 	= TotalSulfurDioxide;
	IMP_Chlorides 			= Chlorides;
	IMP_FreeSulfurDioxide 	= FreeSulfurDioxide;
	IMP_ResidualSugar 		= ResidualSugar;
	IMP_pH 				= pH;

	/*Missing values for STARS seems significant, based off PROC FREQ, and warrants a
flag*/
	F_STARS 			= 0;

	if missing(STARS) 		then do;IMP_STARS 		= 1;
F_STARS = 1; end;
	if missing(Sulphates) 		then IMP_Sulphates 		= 0.5271118;
	if missing(Alcohol) 		then IMP_Alcohol 		= 9.4;
	if missing(TotalSulfurDioxide)then IMP_TotalSulfurDioxide = 125.0000000;
	if missing(Chlorides) 		then IMP_Chlorides 		= 0.0548225;
	if missing(FreeSulfurDioxide) then IMP_FreeSulfurDioxide  = 30.8455713;
	if missing(ResidualSugar) 	then IMP_ResidualSugar 		= 5.4187331;
	if missing(pH) 			then IMP_pH
	= 3.2076282;
```

```
        keep
        TARGET Index
                TARGET_FLAG
                TARGET_AMT
                AcidIndex
                IMP_Alcohol
                IMP_Chlorides
                CitricAcid
                Density
                FixedAcidity
                IMP_FreeSulfurDioxide
                LabelAppeal
                IMP_ResidualSugar
                IMP_STARS
                F_STARS
                IMP_Sulphates
                IMP_TotalSulfurDioxide
                VolatileAcidity
                IMP_pH
         ;
run;
/*TRANSFORMATIONS*/
data xfer_data;
        set imp_data;
        FixedAcidity = sqrt(abs(FixedAcidity) + 1);
        VolatileAcidity = log(abs(VolatileAcidity));
        CitricAcid = sqrt(abs(CitricAcid));
        IMP_ResidualSugar = log(abs(IMP_ResidualSugar)+ 1);
        IMP_Chlorides = sqrt(abs(IMP_Chlorides));
        IMP_FreeSulfurDioxide = log(abs(IMP_FreeSulfurDioxide) + 1) ;
        IMP_TotalSulfurDioxide = log(abs(IMP_TotalSulfurDioxide) + 1);
run;
data score_log;
        set xfer_data;
        P_TARGET_FLAG = 0.7802
   + AcidIndex * -0.3819
   + IMP_Alcohol * -0.0184
   + CitricAcid * 0.2617
   + IMP_FreeSulfurDioxide * 0.0783
   + LabelAppeal * -0.4649
   + IMP_STARS * 2.5383
   + F_STARS * -1.8202
   + IMP_ResidualSugar * 0.0547
```

```
   + IMP_Sulphates * -0.1056
   + IMP_TotalSulfurDioxide * 0.2201
   + VolatileAcidity * -0.1497
   + IMP_pH * -0.1799
         ;
         P_TARGET_FLAG = exp(P_TARGET_FLAG);
         P_TARGET_FLAG = P_TARGET_FLAG / (1.0 + P_TARGET_FLAG);
run;
data score_poi;
         set score_log;
         P_TARGET_AMT = 0.7874
   + AcidIndex * -0.0205
   + IMP_Alcohol * 0.009
   + CitricAcid * 0.0083
   + IMP_FreeSulfurDioxide * 0.0052
   + LabelAppeal * 0.2952
   + IMP_ResidualSugar * -0.0021
   + IMP_STARS * 0.1211
   + F_STARS * -0.0866
   + IMP_Sulphates * 0.0003
   + IMP_TotalSulfurDioxide * -0.0049
   + VolatileAcidity * -0.0132
   ;
         P_TARGET_AMT = exp(P_TARGET_AMT) + 1;
         P_SCORE_HURDLE = P_TARGET_FLAG * P_TARGET_AMT;
run;
data score_data;
         set score_poi;
         keep INDEX
                  P_SCORE_HURDLE
                  ;
run;
proc print data = score_data; run; quit;
```