

```

/*****
*****
*****
*****/

/*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 'Moneyball Deploy Score Code';

/* Create a shortcut for the Moneyball Data */
data m_test;
    set mydata.MONEYBALL_TEST;
run;

proc print data = m_test (obs=10);

proc means data=m_test n nmiss mean max min stddev p25 p50 p75 p99;
    var Team_Batting_H Team_Batting_2B Team_Batting_3B Team_Batting_HR
    Team_Batting_BB
    Team_Batting_HBP Team_Batting_SO Team_Baserun_SB Team_Baserun_CS
    Team_Fielding_E
    Team_Fielding_DP Team_Pitching_BB Team_Pitching_H Team_Pitching_HR
    Team_Pitching_SO;
run;
/* Replace missing values */
data impute_mb;
    set m_test;

    drop team_batting_HBP; /* missing 90% of values so not using */

    Imp_Team_Batting_SO = Team_Batting_SO;
    F_Team_Batting_SO = 0; /* Missing value flag */
    if missing (Team_Batting_SO) then do;
        Imp_Team_Batting_SO = 735.61; /*Normal Distro; impute with mean */
        F_Team_Batting_SO = 1;
    end;
    drop Team_Batting_SO;

```

```

Imp_Team_Baserun_SB = Team_Baserun_SB;
F_Team_Baserun_SB = 0;
if missing (Team_Baserun_SB) then do;
    Imp_Team_Baserun_SB = 65; /* Left-skewed; impute with mode */
    F_Team_Baserun_SB = 1;
end;
drop Team_Baserun_SB;

Imp_Team_Baserun_CS = Team_Baserun_CS;
F_Team_Baserun_CS = 0;
if missing (Team_Baserun_CS) then do;
    Imp_Team_Baserun_CS = 52.80; /*Normal Distro; impute with mean */
    F_Team_Baserun_CS = 1;
end;
drop Team_Baserun_CS;
Imp_Team_Fielding_DP = Team_Fielding_DP;
F_Team_Fielding_DP = 0;
if missing (Team_Fielding_DP) then do;
    Imp_Team_Fielding_DP = 146.39; /* Normal Distro; impute with mean */
    F_Team_Fielding_DP = 1;
end;
drop Team_Fielding_DP;
Imp_Team_Pitching_SO = Team_Pitching_SO;
F_Team_Pitching_SO = 0;
if missing (Team_Pitching_SO) then do;
    Imp_Team_Pitching_SO = 817.73; /* Normal Distro; impute with mean */
    F_Team_Pitching_SO = 1;
end;
drop Team_Pitching_SO;

run;
/* Transform Data */
data trans_mb;
    set impute_mb;
    Trans_Team_Batting_3B = log(Team_Batting_3B + 1);
    Trans_Team_Baserun_SB = log(Imp_Team_Baserun_SB + 1);
    Trans_Team_Fielding_E = abs(log(Team_Fielding_E/246.48)); /*Standardize and log*/
    Trans_Team_Pitching_BB = log(Team_Pitching_BB/553 + 1)/0.678; /*Double
Standardize and log*/
    Trans_Team_Pitching_H = Team_Pitching_H;
    if Trans_Team_Pitching_H <= 627 then
Trans_Team_Pitching_H = 627;
    else if Trans_Team_Pitching_H >= 2475 then Trans_Team_Pitching_H =
2475;
    Trans_Team_Pitching_H = log(Trans_Team_Pitching_H);

```

```

Trans_Team_Pitching_SO = log(Imp_Team_Pitching_SO/817.73 + 1);
run;

/* Score Data*/
data m_score;
    set trans_mb;
    P_TARGET_WINS = 165.02689 + 0.05473*TEAM_BATTING_H + -
0.03031*TEAM_BATTING_2B
    + 0.07642*TEAM_BATTING_3B + 0.02652*TEAM_BATTING_BB +
0.00154*TEAM_PITCHING_H
    + 0.05785*TEAM_PITCHING_HR + -0.05941*TEAM_FIELDING_E + -
0.01650*Imp_Team_Batting_SO
    + 0.06180*Imp_Team_Baserun_SB + 39.42668*F_Team_Baserun_SB + -
0.10891*Imp_Team_Fielding_DP
    + 2.92487*F_Team_Fielding_DP + 9.08696*F_Team_Pitching_SO + -
1.60493*Trans_Team_Baserun_SB
    + 6.84819*Trans_Team_Fielding_E + -20.09114*Trans_Team_Pitching_H;
run;

data m_scored;
    set m_score;
    drop TEAM_BATTING_H    TEAM_BATTING_2B  TEAM_BATTING_3B
    TEAM_BATTING_HR  TEAM_BATTING_BB  TEAM_PITCHING_H
    TEAM_PITCHING_HR TEAM_PITCHING_BB TEAM_FIELDING_E Imp_Team_Batting_SO
F_Team_Batting_SO  Imp_Team_Baserun_SB
    F_Team_Baserun_SB Imp_Team_Baserun_CS F_Team_Baserun_CS
    Imp_Team_Fielding_DP F_Team_Fielding_DP
    Imp_Team_Pitching_SO F_Team_Pitching_SO    Trans_Team_Batting_3B
Trans_Team_Baserun_SB    Trans_Team_Fielding_E
    Trans_Team_Pitching_BB    Trans_Team_Pitching_H Trans_Team_Pitching_SO;
run;

/* Print Results */
proc print data = m_scored; run;

proc means data = m_scored N NMISS MIN MAX;
var P_TARGET_WINS;
run;

```