MODEL CODE:

```
proc import out=nba datafile="/home/u63401111/sasuser.v94/Final Project/NBA.xlsx"
dbms=xlsx replace; sheet="Data";
run;

/* Scatterplot of Players by Their Total Points */

proc sgplot data=nba;
        scatter x=age y= w;
        title "Scatterplot of Age by Total Games Won";
        xaxis label="Age";
        yaxis label="Total Games Won";

/* Bar Chart of Positions Count */

proc sgplot data=nba;
        vbar pos / categoryorder=respdesc;
        title "Bar Chart of Player Positions";
run;

/* Correlation Matrix */
proc corr data=nba;
        var gp age min fgm fga _3pm _3pa ftm fta oreb dreb ast tov stl blk pf fp dd2 td3
efficiency FGP _3PP FTP REB L;
run;

/* Boxplot */
proc sgplot data=nba;
  vbox w / category=pos;
  title "Box Plot of Points Scored by Basketball Position";
  xaxis label="Position";
  yaxis label="Games Won";
run;

/* Numerical Summaries*/

proc means data=nba n mean median min max maxdec=3;
        var W;
        class pos;
run;

proc means data=nba n mean median min max maxdec=3;
```

```
    var w;
    class team;
run;

proc means data=nba n mean median min max maxdec=3;
    var gp;
    class team;
run;

proc means data=nba n mean median min max maxdec=3;
    var gp;
    class pos;
run;


/* BUILDING MODELS */

/* Creating Dummy Variables */
data nba;
    set nba;
    if POS = "PG" then POS_PG = 1; else POS_PG = 0;
    if POS = "SG" then POS_SG = 1; else POS_SG = 0;
    if POS = "SF" then POS_SF = 1; else POS_SF = 0;
    if POS = "PF" then POS_PF = 1; else POS_PF = 0;
    if POS = "F" then POS_F = 1; else POS_F = 0;
    if POS = "G" then POS_G = 1; else POS_G = 0;
run;

/* Partition Data - 80/20 */
proc surveyselect data=nba samprate=.8 method=srs outall out=nba_part seed=12345;
run;

/* Linear Regression 1 - all predictors ------------------------------------------------------ */

proc hpreg data=nba_part;
    partition rolevar=selected(train='1' validate='0');
    model W = Age GP MIN PTS FGM FGA FGP _3PM _3PA _3PP FTM FTA FTP OREB DREB REB
AST TOV STL BLK PF FP DD2 TD3 EFFICIENCY POS_PG POS_SG POS_SF POS_PF POS_F
POS_G/vif;
    selection method=stepwise;
    output out=linear_pred1 p=W_predict r=W_resid copyvar=(W selected);
run;
```

```
/* Evaluate Linear Regression Model Performance */
data linear_pred1;
   set linear_pred1;
   if selected = 1 then do;
      mape_fit = (abs(W_resid) / W) * 100;
      mae_fit = abs(W_resid);
      mse_fit = W_resid**2;
   end;
   else if selected = 0 then do;
      mape_acc = (abs(W_resid) / W) * 100;
      mae_acc = abs(W_resid);
      mse_acc = W_resid**2;
   end;
run;

/* Performance Metrics */
proc means data=linear_pred1 n mean maxdec=3;
   var mape_fit mape_acc mae_fit mae_acc mse_fit mse_acc;
   title "Performance Metrics: MAPE, MAE, and MSE for Player Wins (W)";
run;

/* Linear Regression 2 */

proc hpreg data=nba_part;
   partition rolevar=selected(train='1' validate='0');
   model W = AGE GP FGM _3PM FTM OREB DREB AST STL BLK EFFICIENCY/vif;
   selection method=stepwise;
   output out=linear_pred2 p=W_predict r=W_resid copyvar=(W selected);
run;

/* Evaluate Linear Regression Model Performance */
data linear_pred2;
   set linear_pred2;
   if selected = 1 then do;
      mape_fit = (abs(W_resid) / W) * 100;
      mae_fit = abs(W_resid);
      mse_fit = W_resid**2;
   end;
   else if selected = 0 then do;
      mape_acc = (abs(W_resid) / W) * 100;
      mae_acc = abs(W_resid);
```

```
    mse_acc = W_resid**2;
  end;
run;

/* Performance Metrics */
proc means data=linear_pred2 n mean maxdec=3;
  var mape_fit mape_acc mae_fit mae_acc mse_fit mse_acc;
  title "Performance Metrics: MAPE, MAE, and MSE for Player Wins (W)";
run;

/* CART Model 1 --------------------------------------------------------- -*/

proc hpsplit data=nba_part nodes=detail;
        partition rolevar=selected(train='1' validate='0');
        class POS;
        model W = POS Age GP MIN PTS FGM FGA FGP _3PM _3PA _3PP FTM FTA FTP OREB
DREB REB AST TOV STL BLK PF FP DD2 TD3 EFFICIENCY;
        grow rss;
        prune cc;
        output out= nbaout_cart1;
        id selected;
run;

/* CART Model - Error Measures */
data nbaout_cart1;
        set nbaout_cart1;
        if selected=1 then
                do;
                        mape_fit=abs((W-P_W)/W)*100;
                        mae_fit=abs(W-P_W);
                        mse_fit=(W-P_W)**2;
                end;
        else if selected=0 then
                do;
                        mape_acc=abs((W-P_W)/W)*100;
                        mae_acc=abs(W-P_W);
                        mse_acc=(W-P_W)**2;
                end;
run;

/* CART Model 2 */
```

```
proc hpsplit data=nba_part nodes=detail;
        partition rolevar=selected(train='1' validate='0');
        model W = AGE GP FGM _3PM FTM OREB DREB AST STL BLK EFFICIENCY;
        grow rss;
        prune cc;
        output out= nbaout_cart2;
        id selected;
run;

/* CART Model 2 - Error Measures */
data nbaout_cart2;
        set nbaout_cart2;
        if selected=1 then
                do;
                        mape_fit=abs((W-P_W)/W)*100;
                        mae_fit=abs(W-P_W);
                        mse_fit=(W-P_W)**2;
                end;
        else if selected=0 then
                do;
                        mape_acc=abs((W-P_W)/W)*100;
                        mae_acc=abs(W-P_W);
                        mse_acc=(W-P_W)**2;
                end;
run;

proc means data=nbaout_cart2 n mean;
        var mape_fit mape_acc mae_fit mae_acc mse_fit mse_acc;
run;

/* Neural Network Model 1 – all predictors ---------------------------------------------- */

proc hpneural data=nba_part;
        partition rolevar=selected(train=1);
        target W/level=int;
        input AGE GP FGM _3PM FTM OREB DREB AST STL BLK EFFICIENCY/level=int;
        hidden 11;
        train numtries=10 maxiter=1000;
        id w selected;
        score out=nbaoutneural1;
run;
```

```
data nbaoutneural1;
        set nbaoutneural1;
        if selected=1 then
                do;
                        mape_fit=(abs(w-p_w)/w)*100;
                        mae_fit=abs(w-p_w);
                        mse_fit=(w-p_w)**2;
                end;
        else if selected=0 then
                do;
                        mape_acc=(abs(w-p_w)/w)*100;
                        mae_acc=abs(w-p_w);
                        mse_acc=(w-p_w)**2;
                end;
run;

proc means data=nbaoutneural1 n mean maxdec=2;
        var mape_fit mape_acc mae_fit mae_acc mse_fit mse_acc;
run;

/* Neural Network Model 2 */

proc hpneural data=nba_part;
        partition rolevar=selected(train=1);
        target w/level=int;
        input AGE GP FGM _3PM FTM OREB DREB AST STL BLK EFFICIENCY/level=int;
        hidden 11;
        train numtries=10 maxiter=1000;
        id w selected;
        score out=nbaoutneural2;
run;


data nbaoutneural2;
        set nbaoutneural2;
        if selected=1 then
                do;
                        mape_fit=(abs(w-p_w)/w)*100;
                        mae_fit=abs(w-p_w);
                        mse_fit=(w-p_w)**2;
                end;
        else if selected=0 then
        do;
            mape_acc=(abs(w-p_w)/w)*100;
            mae_acc=abs(w-p_w);
            mse_acc=(w-p_w)**2;
        end;
run;

proc means data=nbaoutneural2 n mean maxdec=2;
    var mape_fit mape_acc mae_fit mae_acc mse_fit mse_acc;
run;
```