Overview

The attached data set contains NBA team stats from the last eleven regular seasons. Your task is to fit a linear regression model predicting the number of wins. A typical regular season consists of 82 games (41 at home and 41 on the road) for each team. The teams are divided into Western and Eastern Conferences (and into six divisions, 3 in each conference). A team faces opponents in its division four times yearly (16 games). Each team plays six of the teams from the other two divisions in its conference four times (24 games) and the remaining four teams three times (12 games). Finally, each team plays all the teams in the other conference twice apiece (30 games).

Data Description

Variable		Description
Season		Regular season
Conference		Western or Eastern
Age		Players' average age on Feb 1 of the season
W		Wins
SOS		Strength of Schedule; a rating of strength of schedule. The rating is denominated in points above/below average, where zero is average
Pace		Pace Factor - An estimate of possessions per 48 minutes
FTr		Free Throw Attempt Rate - Number of FT Attempts Per FG Attempt
3PAr		3-Point Attempt Rate - Percentage of FG Attempts from 3-Point Range
Offense Four Factors	***eFG%***	Effective Field Goal Percentage - This statistic adjusts for the fact that a 3-point field goal is worth one more point than a 2-point field goal
	TOV%	Turnover Percentage - An estimate of turnovers committed per 100 plays
	ORB%	Offensive Rebound Percentage - An estimate of the percentage of available offensive rebounds a player grabbed while they were on the floor
	FT/FGA	Free Throws Per Field Goal Attempt
Defense Four Factors	***OppeFG%***	Opponent Effective Field Goal Percentage
	OppTOV%	Opponent Turnover Percentage (forcing the turnover)
	DRB%	Defensive Rebound Percentage - An estimate of the percentage of available defensive rebounds a player grabbed while they were on the floor
	OppFT/FGA	Opponent Free Throws Per Field Goal Attempt (should be minimal)
	ORtg	**offensive rating, an estimate of points scored per 100 possessions. (<90 be bad)**

Your Task

1. Import the data to SAS.
2. Using SAS, fit a multiple linear regression model using Wins (W) as the dependent variable and other variables (except team names) as predictors.
    a. Is the model statistically significant at the 0.05 significance level? State the hypotheses and include supporting figures that provide the decision for the test.
    b. Discuss the model's goodness of fit.

   c. Which predictors, if any, are significant and which are not? Include the supporting figures in your answer.
   d. Examine the residuals. Do regression assumptions hold?
   e. Is there any evidence of multicollinearity?
   f. Are there any influential observations present?
3. Fit a regression model using the stepwise selection procedure using Mallow's C(p) as the selection criterion.
   a. Is the model significant?
   b. Discuss the model's goodness of fit.
   c. Discuss the significance of predictors.
   d. Examine regression assumptions.
   e. Is multicollinearity a problem?
4. Are models from parts 1 & 3 different? If so, discuss the differences.
5. Interpret the results from part 3. Note that reporting the figures from the regression result set in SAS is not equivalent to interpreting the results. If in doubt, think of a scenario where you act as a consultant to the general manager of a basketball team who is trying to assess what factors the team should focus on for it to have a successful season.

Submission Guidelines

Your submission should include the following:

1. One Word document with your answers and supporting figures included.
2. One SAS program file with all the code used in the problem. Any tables that you create should be created in the WORK library.