

SAS Programming

Practice #5

1. Use the combined NBA data which contains the conference variable. You created this data set in Practice #3. You will use this data set often in the rest of the course. Beginning with these NBA data steps, add lines of SAS code to produce a single program that does all of the following.
 - a. In this same DATA step that combines the two years of data, create a variable called Points. The total points scored are the sum of points scored on 2-point field goals plus points scored on 3-point field goals, plus the number of free throws. Recall "field goals" is the number of combined 2-point and 3-point field goals, not the number of points. For example –

Team	Field	Field	3-Point	
	Goals	Goal Attempts	Field Goals	Free Throws
Denver Nuggets	3339	6983	521	2818

In this case, the Denver Nuggets would have 10017 points.

- b. Using the combined years' basketball data, BY year rank the teams with respect to POINTS created in question 1a. The team with the best (highest) POINTS value should have rank 1. For teams that are tied, assign the lowest rank number. For example, if two teams are tied for second, assign rank 2 to the two teams. No third place (or rank 3) would then be assigned. Print the team name, POINTS value and the rank in the ranked order by year. Do NOT use IF . . . THEN DELETE statements in the DATA step to control the print out.
2. Soybeans are typically planted in the spring in Oklahoma. A researcher is examining the effects of late planting on a variety of soybeans. The yield data are:

Early Spring:	4101, 3809, 2764, 5407, 3328
Late Spring:	4482, 3610, 3824, 3685, 4076
Summer 1:	1944, 2348, 2938, 2008, 2145
Summer 2:	1384, 1492, 1776, 1288, 1559

- a. Use SAS/NPAR1WAY and Wilcoxon (rank sum) methods to analyze the data for differences among the four planting dates. Experiment with the ODS Graphics available for this method.
 - b. Use SAS/GLM to perform an ANOVA on the above data. Include a meanplot and a residual plot panel in your analysis. ANOVA is appropriate when the data are normally distributed. One of the places you can visually check this is on the residual plot. Do the data appear normal?
Option: Rather than a panel of plots for the residual analysis produce larger individual plots. How can you do this?