

STAT 517: Takehome 2

Provide code, output, and answers in either Word, PDF, RTF, or text format; try to combine files as much as possible before sending them to me. Work independently.

1. The Excel file contains two separate measurements on the number of tree rings in core samples taken from different trees. Each tree was cored twice, with the second core taken at a right angle to the first core. Different measurements occur because the rings may be hard to count, or the core may miss the center of the tree.
 - (a) Import the data from the Excel file into SAS, and save it in a SAS library other than WORK. Read it into a datafile in your WORK library.
 - (b) Create a new variable called *Avage*, which is the average of the last two variables' values for each tree. Don't worry about missing data.
 - (c) Sort the resulting data set by the values of *Avage*, from greatest to least.
 - (d) Print the resulting sorted data in the output window with a descriptive title. Do the variables have attractive and readable formats? Modify the formats if necessary.
 - (e) In addition to the above steps, graduate students should complete the following: *compute, store, and print* averages for complete data records only (no missing cores). Their final output should not contain the `Observation` variable.

2. Students were recently polled on 6 separate items measuring the availability and cost of texts at the University bookstore. Their responses are 1="Strongly Disagree", 2="Disagree", 3="Neutral", 4="Agree", 5="Strongly Agree", 6="Missing". These coded answers for variables Y1-Y6 are stored in the Excel worksheet, along with student IDs.
 - (a) Read the data into SAS as the data set `Survey`. Graduate students: Under `Options...`, be sure to click on the box that asks whether you want to save numeric data in a mixed column as character data; this helps with the Student ID variable.
 - (b) Use arrays and `DO` loops to create the character variables Q1-Q6; these variables correspond to the descriptive labels provided for the numeric variables Y1-Y6.
 - (c) Calculate the sum of Y1-Y6 after this recoding; don't worry about missing data.
 - (d) Create a format `Likertfmt`, which will print the character strings shown above in place of each of the coded values 1-5, and "Missing" for missing values. Print a frequency table for variable Y1 using the `Likertfmt` in your table.
 - (e) Grad students: Forms included reverse-coded responses, indicated by the leading R in the ID variable. Reverse the scores for these students (only for responses 1-5) in SAS, then repeat the steps in the exercise above. I can think of awkward ways to handle this using methods we already know, but it would help to look up the `SUBSTR(variable, startposition, length)` command.

3. Midlands Tech faculty members are paid once a month. Suppose the scheduled paydates are 1/1/07-5/1/07,9/1/07-12/1/07. Read these dates into SAS (`DATALINES` may be easier than an Excel Import here).
 - (a) Change any payday that occurs on a weekend to the first weekday preceding it (this will always be a Friday), using the `WEEKDAY` function (1=Sunday,...,7=Saturday). Print a pay schedule for 2007 including weekday and date—you can print using either a combined or separate format variable(s).
 - (b) Graduate students should print an additional variable that notes when the payday would have initially fallen on a weekend.