STAT 440 – Homework 9

Students are encouraged to work together on homework. However, sharing or copying any part of the homework is an infraction of the University's rules on Academic Integrity.

Final submissions must be uploaded to our Compass 2g site on the Homework page. No email, hardcopy, or late submissions will be accepted.

Getting the program file ready

a. Create a folder on the hard drive with the following pathname – C:\440\hw9. Save all data files accompanying this assignment in that folder. If you cannot create the folder because you are working on a university computer and don't have permission, create the …\440\hw9 folder elsewhere.
b. Assign the library reference **hw9** to the folder 'C:\440\hw9'. Use this library as your permanent library for this assignment. If you could not create the folder, assign the library reference **hw9** to your …\440\hw9 folder.
   Note: If you are using a folder other than 'C:\440\hw9', you must change any pathname references in your program file to 'C:\440\hw9' before submitting your homework.

Submitting your work to Compass 2g

You are to submit two (and only two) files for your homework submission.

1. Your SAS program file which should be saved as **HW*n_YourNetID*.sas**. For example, my file for the HW1 assignment would be HW1_dunger.sas. All program statements and code should be included in one program file.

2. Your Report including all relevant output to address the exercises. For this homework, use ODS to send your results to a Portable Document Format (PDF) file called **HW*n_YourNetID*.pdf**. For example, my file for HW1 would be HW1_dunger.pdf. Only include your final set of output. Do not include output for every execution of your SAS program. Use the template file **hw1 template.sas** as your guide.

You have an unlimited number of submissions, but only the last one will be viewed and graded. Homework submissions must always come as a pair of files, as described above.

1. You will be working with the SAS data files **demographic**, **survey1**, and **survey2**.

    a. Merge the **demographic** and **survey1** data sets based on a common identifier to create a new, temporary SAS data set called **demo1_*NetID***. In **demographic**, the identifier is called ID. In **survey1**, the identifier is called Subj. Both are character variables.
    b. Print the data portion of **demo1_*NetID***. (Include results in the HW Report.)

    c. Merge the **demographic** and **survey2** data sets based on a common identifier to create a new, temporary SAS data set called **demo2_*NetID***. In **demographic**, the identifier is called ID and it is a character variable. In **survey2**, the identifier is also called ID, but it is a numeric variable.
    d. Print the data portion of **demo2_*NetID***. (Include results in the HW Report.)

2. You will be working with the SAS data set **fivepeople** which contains four variables pertaining to five observations.

    a. Read the data set **fivepeople** into SAS to create a temporary SAS data set called **updated_*NetID***.
        - Update ID so that it is a numeric variable instead of character.
        - Update the values in Name so that they appear as "lastname, firstname". For example, "John Doe" should instead appear as "Doe, John".
        - The variable Phone should be converted from character to numeric and should contain only digits, e.g. (123)456-7890 would become 1234567890.
        - Create a character variable called HtSymbol that rewrites the values in the Height variable using symbols: single quotes (') for feet and double quotes (") for inches. For example, the height 6ft. 5in. should appear as 6' 5".
        - Create a numeric variable called HtInches that is the person's height in inches. For example, the height 6ft. 5in. should appear as 77.
        - Notes: The units in the variable Height can be either upper or lowercase, and there may or may not be a period following the units. Also, one of the Height values is missing an inches value. Depending on your solution, this may contribute to an invalid data error in the Log. That's okay as long as your solution successfully creates HtInches.
        - Create a numeric variable called WtPounds that is the person's weight converted from the mixed rational notation of Weight. WtPounds should be in decimal notation and rounded to the nearest .001.
        - Drop the Height variable.

    b. Print the data portion of **updated _*NetID*** with all variables displayed as such: ID, Name, Phone, HtSymbol, HtInches, Weight, WtPounds. (Include results in the HW Report.)

    c. Print the descriptor portion of **updated _*NetID***. (Include results in the HW Report.)