STAT 440 – Homework 03

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\hw03. 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\hw03 folder elsewhere.
b. Assign the library reference **hw03** to the folder 'C:\440\hw03'. Use this library as your permanent library for this assignment. If you could not create the folder, assign the library reference **hw03** to your …\440\hw03 folder.
   Note: If you are using a folder other than 'C:\440\hw03', you must change any pathname references in your program file to 'C:\440\hw03' 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 HW03 assignment would be HW03_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 HW01 would be HW01_dunger.pdf. Only include your final set of output. Do not include output for every execution of your SAS program. Use the template file **hw01 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. The raw data file **AUaids.dat** contains data on patients diagnosed with AIDS in Australia before July 1, 1991. In order of occurrence, it contains the following variables: observation number, state of origin, sex, date of diagnosis, date of death or end of observation, status at end of observation, reported transmission category, and age at diagnosis.

    a. Create user defined formats for state of origin (either New South Wales, Victoria, Queensland, or Other), sex (female or male), and status at end of observation (alive or dead).

    b. Write a DATA step so that the **AUaids.dat** raw data file is used to create a SAS data set called **AUaids_*NetID***. Choose appropriate names, labels, formats, and other attributes as needed for the variables.

    c. Print the first 10 observations from the data portion of your new SAS data file which represent patients from Queensland, excluding observation numbers. (Include results in the HW Report.)

    d. Repeat part (a) by writing a DATA step to read in **AUaids.dat**. But this time the output data set should only contain those patients who are males age 25 and under and who contracted AIDS through the receipt of `"blood"`. The output data set should be permanent and titled **under26_*NetID***. Choose appropriate names, labels, and other attributes as needed for the variables.

    e. Print the entire data portion of your new SAS data file, excluding observation numbers. (Include results in the HW Report.)

2. You will be working with sleep data from the article "Sleep in Mammals: Ecological and Constitutional Correlates" by Allison, T. and Cicchetti, D. (1976), *Science*, November 12, vol. 194, pp. 732-734.

The fields are delimiter-separated and the records have variable length (free-format), meaning that each line does not necessarily have the same number of characters.

Variable Name	Description
Species	animal species
BodyWt	body weight in kg
BrainWt	brain weight in g
Slow	slow wave ("nondreaming") sleep in hrs/day
Para	paradoxical ("dreaming") sleep in hrs/day
Total	total sleep in hrs/day; sum of slow wave and paradoxical sleep
LifeSpan	maximum life span in years
Gestation	gestation time in days
Pred	predation index (1-5) 1 = minimum (least likely to be preyed upon) 5 = maximum (most likely to be preyed upon)
Exposure	sleep exposure index (1-5) 1 = least exposed (e.g. animal sleeps in a well-protected den) 5 = most exposed
Danger	overall danger index (1-5); based on the above two indices and other information 1 = least danger (from other animals) 5 = most danger (from other animals)

a. Using list input, write a DATA step to read in the raw data file and create a SAS data file called **sleep_*YourNetID*.**
   - Use informats where necessary to ensure the data is read properly and completely.
   - Permanently assign a format to the variables BodyWt and BrainWt so that no more than 2 decimal places are displayed for their values and so that commas separate every three digits.
   - Permanently assign a format to the three "sleep" variables so that only 1 decimal place is displayed for their values.
   - Assign labels where necessary.

b. Print the descriptor portion of your new SAS data file. (Include results in the HW Report.) It can be used to verify that the number of observations, number of variables, and variable attributes are as you'd expect.

c. Create a SAS date file called **big_*YourNetID*** that contains only animals with a body weight of at least 150 kg and brain weight greater than 80g. Include only the species name and two weight variables of interest in **big_*YourNetID***.

d. Print the data portion of your new SAS data file. Suppress observation numbers and display variable labels. (Include results in the HW Report.)

e. Create a SAS date file called **nottired_*YourNetID*** that contains only animals that get either less than 6 hours of slow wave sleep or less than 6 total hours of sleep in a day or both. Include only the species name and all three "sleep" variables in **nottired_*YourNetID***.

f. Print the data portion of your new SAS data file. Suppress observation numbers and display variable labels. (Include results in the HW Report.)