

STSCI 5010 – Lab 1
(9/21/2018)

General instruction:

- Download the SAS dataset “**nutrition.sas7bdat**” from the course website. Set your SAS preference to produce both listing and HTML outputs. Use SAS comments for 1) marking the title of each exercise, 2) briefly documenting what you do in your SAS program at the beginning of each exercise and, 3) including your answer to each question at the end of your relevant SAS code segment unless specified differently in the question. Specify a title for each output. In Exercises 2 and 3, you need to specify a footnote too.
- Specify a proper page size and a proper line size so that your listing output is well organized: 1) your listing output from the same procedure is on the same page, 2) each observation is displayed on the same line and, 3) the footnote is properly aligned with the main text (e.g., not too much white space in between).
- Submit four files (combined with 7-zip or a similar tool, named lab1_last-name_first-name.7z, e.g., lab1_Yang_Xiaolong.7z) *to the course website*:
 - Your SAS code (in one single file), named lab1_code_last-name_first-name.sas, e.g., lab1_code_Yang_Xiaolong.sas. (For the points, see below Exercises 1 to 3)
 - Your SAS listing output, named lab1_listing_output_last-name_first-name.lst, e.g., lab1_listing_output_Yang_Xiaolong.lst. **(5 points)**
 - Your SAS HTML output, named lab1_HTML_output_last-name_first-name.mht, e.g., lab1_HTML_output_Yang_Xiaolong.mht. **(5 points)**
 - Your SAS log file, named lab1_log_last-name_first-name.log. **(5 points)**
- Lab report deadline: 11:59PM, 9/23/2108.

Exercise 1 (35 points):

- Assign a new SAS library named Lab1 and keep the nutrition dataset in this library.
- Use a SAS procedure to find out the number of variables, number of observations. What are the variables of a char data type?
- Use a SAS procedure to display the whole nutrition dataset and inspect the variables and the values of the variables. Comment on what you found.
- Use the data set option to only display the observations from row 10 to row 20.

Exercise 2 (25 points):

Use a DATA step to create a new permanent SAS dataset of only **male** participants whose KCAL (daily calorie intake) was 3000 or more. Call this new dataset “males3000kcal” and store it in your Lab1 library. Create a report that is sorted in descending order by KCAL, and display in your report only the first 15 observations. Include only the following variables in the report: GENDER, KCAL, KCAL_FAT, KCAL_CHO, and KCAL_PRO, and include at least one title and one footnote that describe your report.

Exercise 3 (25 points):

Sort the **nutrition** dataset in descending order by IRON and then by FIBER and output a temporary dataset to your WORK library called “sorted_IRON_FIBER,” which is basically the same data as the original dataset just with a different order of the rows. Create a report from this new dataset for only the **female** participants whose IRON was less than 4 and FIBER was less than 4, including the following variables: GENDER, KCAL, VIT_A, VIT_D, FIBER, IRON and PROTEIN. Again include relevant titles and footnotes for your report. How many women meet these criteria, and are hence included in your report?

Optional (but important): Practice all the SAS code covered in the class.