

# STAT 480 Assignment #3

(Due July 23rd, Thursday, 11:59 PM)

**Instructions:** Read lectures 4, 5 and sample codes I used in class carefully before starting your homework. You need to submit a .sas file on ANGEL to answer the following questions (Lessons > Homework > Homework#3 > dropbox for HW#3).

**1. (IF-THEN-ELSE Statement and Create a List Report).** We have the following small dataset called *HealthReport*

Subj	Fname	Lname	Gender	Height	Weight	Date	Cal
1024	Alice	Smith	Female	65	125	12/01/95	2,036
1167	Maryann	White	Female	68	140	12/02/95	1,800
1168	Thomas	Jones	Male		190	01/05/96	2,302
1201	Benedictine	Arnold	Male	68	190	11/30/95	2,432
1302	Felicia	Ho	Female	63	115	01/01/96	1,972
1344	James	Hunter	Male	70	188	12/15/95	1,869

Table 1: Health Examination Report

The columns, from left to right, represent the following variables:

Variable Name	Descriptions	Variable Type
Subj	Subject number	Numeric
Fname	First name	Character
Lname	Last name	Character
Gender	Gender	Character
Height	Height (lbs.)	Numeric
Weight	Weight (inches)	Numeric
Date	Date of measurement	Numeric
Cal	Calorie	Numeric

Table 2: Descriptions of variables

- (a) Read the dataset above into SAS (pay attention to length and format of variables), and change the SAS system settings as follows:

linesize = 64, no current date and time, no page number

- (b) Use assignment statement to create a new numeric variable *BMI*. BMI (Body mass index) is a person's weight divided by the square of his/her height. In other words, it is defined to be

$$BMI = \frac{Weight}{Height^2}$$

where the weight unit is kilogram and the height unit is meter. So first you need to convert pound into kilogram and convert inch into meter (1 kilogram = 2.205 pounds, 1 inch = 0.0254 meter), and then calculate BMI.

- (c) Use **IF-TEHN-ELSE** statement to create a new character variable *Result*. The value of variable *Result* is determined by the value of BMI according to the following table.

Value of BMI	Value of Result
Missing	'Missing'
$\leq 25$	'Underweight'
$> 25$ and $\leq 30$	'Normal'
$> 30$	'Overweight'

**Note:** the length of the variable *Result* can be longer than 8. You should specify an appropriate length for variable *Result* with **length** statement. The **length** statement should be placed before the **input** statement.

- (d) Use **SORT** procedure to sort the dataset *HealthReport* first by the variable *Gender* in an ascending order, and then by the variable *Lname* in a descending order. The sorted dataset should be saved in another SAS dataset named *SrtdHealthReport*.
- (e) Print the the sorted dataset *SrtdHealthReport*. When printing the dataset,
1. Use variable *Lname* and *Fname* as the identification variables (Use **ID** statement).
  2. Split the dataset into two groups by variable *Gender*. Calculate the sum of *Cal* for each group.
  3. Use **LABEL** statement to assign labels to all variables to make your output more readable. Particularly, use the descriptions in Table 2 as the labels. You can place the **LABEL** statement either in the DATA step or in the PRINT procedure. Use labels rather than the variable names when printing the dataset.
  4. Add a **TITLE** ('Health Examination Report Sorted by Last Name') and **FOOTNOTE** for this dataset. The content of the **FOOTNOTE** is: 'BMI = Weight(kg)/(Height(m))^2'.
  5. Use the correct **FORMAT** for variables *Date* and *Cal*.
- (f) Use **Where** statement to print the first and last names of female students who are underweight.