

## SAS Training 2019

### Exercise 1

1. Consider the following dataset on 6 students:

```
001 M BIO 80 84
002 M EPI 75 73
003 F EPI 90 86
004 M BIO 83 85
005 F EPI 94 94
006 F EPI 88 84
```

- (a) Write a SAS program that reads in the data within the program. Call the variables StudentID, Gender, Major, Exam1, and Exam2. The first three variables are character and the last two are numeric. Name the dataset class.
- (b) Display the values of the variables using **proc print**. Make sure the variables have been read-in properly. Also, run **proc contents** to display information about the dataset created.
- (c) Use **proc freq** to tabulate the number of men and women, and the number of students in each department.
- (d) Use **proc means** to compute summary statistics for exam 1 and exam 2. *What is the average score for each exam?*
- (e) Go back in the data-step and add a variable called exama which is the average of exam 1 and exam 2. The line of code for this is:

```
exama = (exam1 + exam2)/2;
```

- (f) Use **proc print** to verify the new variable was computed correctly.
- (g) Use **proc means** to compute summary statistics for the variable exama.
- (h) Use **proc sgplot** to plot exam2 versus exam 1 on an x-y plot. Use the following code:

```
proc sgplot data=class;
  scatter x=exam1 y=exam2;
run;
```

- (i) Save your program to the PC. Give it the name exercise1. It will automatically have a .sas extension.
- (j) Exit SAS after you have saved the program. Then enter SAS again and open the program you saved.

2. The datafile **tomhss.csv** is a .csv file containing 38 variables from the Treatment of Mild Hypertension Study.
  - (a) Use **proc import** to read the file creating a SAS dataset called tomhs.
  - (b) Display the list of variables using **proc contents**.
  - (c) Run **proc means** on the dataset. *What is the average age of the participants?*
  - (d) Tabulate the number of men and women in the study using **proc freq**. *What percentage of participants are women?*
3. The datafile **tomhss.sas7bdat** is a SAS dataset version of the .csv file above.
  - (a) In a data-step read in the SAS dataset, creating a new SAS work dataset called tomhs. You will need to use a libname statement and a set statement.
  - (b) Run a **proc means** on the new dataset. The results should be the same as part 2.