

# STSCI 5010 HW1

Due: 11:59 PM, Sep. 27, 2018

## General instructions:

- *What and how to submit:*
    - *Submit your homework with a single zipped file (hw1\_lastName\_firstName.7z) to the course website on Blackboard in a single zipped file:*
      - *A single file of your SAS code, named hw1\_SAS\_code\_your last name\_your first name.sas, for example, hw1\_SAS\_code\_Yang\_Xiaolong.sas. Give your answer(s) to the questions (if any) by attaching these to the end of each SAS code segment as comments.*
      - *Your listing outputs, which are obtained by running all your working SAS code together at once and are saved to a single file named listing\_outputs\_lastName\_firstName.lst.*
      - *and submitted through the course Blackboard account. We will run your code.*
  - *When you do SAS coding, you are required to include comments to explain what you are doing for all the sections of your code. If the comments are missing, up to 5 points will be taken away.*
  - *You are required to use titles and footnotes for all the output produced in this homework (i.e., in all the PROC PRINT procedures). The titles should be some texts about what you are listing (or reporting) and the footnotes should be always a phrase like “Produced by XX” (where XX represents your full name). Adjust the page sizes of your output so that you do not leave too much white space between your result and footnotes and you can display your output on the same page. In this homework, you should only submit the **listing output** (not HTML format).*
1. Download the SAS dataset on the course website called **admit.sas7bdat**. Assign a new SAS library named **HW1** and keep the dataset in this library. Open the dataset in SAS VIEWTABLE (Hint: you can use the SAS Explorer to find and open it) to view the dataset so that you know what is in the dataset. Use a SAS procedure to find out the numbers of observations and variables in the dataset. When was the dataset created? How many character variables are there? How many numeric variables are there? (10 points)
  2. Create three new datasets from the original **admit.sas7bdat** dataset. First, create a new dataset that only includes those whose activity levels (ActLevel) were “HIGH.” Call this new dataset “al\_high”, and save it in the WORK library. Second, create two more new datasets for those whose activity levels were “LOW” or “MOD” with a dataset name of “al\_low” or “al\_mod” respectively. Save these files in the WORK

library as well. Display these three datasets you just created (using listing output). (21 points)

Next time you start a new SAS session after closing the current session, where can you find these three new SAS datasets you just created? (4 points)

3. From the Admit dataset, create a listing output for each sub-questions.
  - a) List all the patients who have HIGH or MOD ActLevel and pay a Fee within the range of [100, 130]. (15 points)
  - b) List the ID, name, gender and age of the patients whose names contains the letters “J” and “R”. (10 points)
  - c) List all the observations where the female patients who pay a fee that is greater than \$100. Use the *DOLLARw.d* format with 2 decimal points. In the output, use their IDs to replace Obs and label the variable ActLevel as “Activity Level” so that it is more meaningful to the people who read the report. (20 points)
  
4. From the Admit dataset, create a new data set named OldMale in the HW1 library containing all the males who are at least 65 years old. Give the variable **Fee** a permanent called “Fee charged at the time of admission (\$)” and a permanent format w.d (where d=2) to variable **Height** (make sure your format can represent all the values from the original data set). Display the OldMale data set with a print procedure that does not contain label and format statements (20 points).