Assignment 4

Assignment Objective

In this assignment will be demonstrating your skills in manipulating, combining, and reporting your data in SAS.

Deliverables

All requested submissions for this assignment must be uploaded as attached files in the assignment folder. Any additional written responses required to the assignment questions must be provided by filling out the 'Text Submission' section when you are uploading your assignment files. To ensure accuracy of feedback, please ensure that written responses to questions are clearly identified by question number.

In assignment 4 you will be producing a written report containing the following deliverables:

- Deliverable 1: A copy of your program submission components.
  When requested, for each question in the assignment that requires the development and running of SAS code, the student must include the following in the report:
  1. A complete listing of the code being run
  2. An unedited copy of the log report that includes all run time information

- Deliverable 2: An electronic copy of all SAS programs
  Executable programs must be included as a separate file attachment with all assignments. Use the following naming convention for all SAS programs when you submit them for evaluation. XX_A4_QY.sas Where XX is your first and last initial, replace Y with the number of the question that this program is referring to. Please note that all individual programs that are submitted must be fully executable on a standalone basis in a SAS on Demand Enterprise Guide Session. For example, if you developed the program assuming some temporary files or library references were run previously, you must copy those steps so that each program you submit runs with a single execution of the program.
- Deliverable 3: Written responses to Assignment Questions
  When requested, you may also be required to provide written responses to Assignment Questions that will assess your ability to interpret program output.

Instructions:

Question 1 (Sub-setting Observations) **(2 marks)**

a) Create a temporary data set, **work.bigdonations**, using **employee_donations** as input.

b) Create a new variable named **Total** that is equal to the sum of **Qtr1, Qtr2, Qtr3**, and **Qtr4**.

c) Create a new variable named **NoDonation** that is equal to the count of missing values in **Qtr1, Qtr2, Qtr3,** and **Qtr4**. Use the NMISS function.

d) The final dataset should contain only observations that meet the following two conditions:
- **Total** values greater than or equal to 50
- **NoDonation** values equal to 0

Use an IF-THEN DELETE statement to eliminate the observations where the conditions are not met

e) Write a PROC PRINT step to create the report below. Partial Proc Print

Obs	Employee_ID	Qtr1	Qtr2	Qtr3	Qtr4	Total	NoDonation
1	1202670	15	15	15	15	60	
2	1202690	20	20	20	20	80	
3	1202710	20	20	20	20	80	
4	1202750	15	15	15	15	60	
5	120660	25	25	25	25	100	0

Submit your program file as XX_A4_Q1.sas (replace XX with your first and last initials)
Submit a copy of the log as XX_A4_Q1_log.txt (replace XX with your first and last initials)
Submit a copy of the results as XX_A4_Q1_results.pdf (or .txt is acceptable) (replace XX with your first and last initials)

Question 2 (Using the APPEND or SET Statements) **(2 marks)**

a) Write and submit three PROC CONTENTS steps to compare the variables in **shoes_eclipse**, **shoes_tracker**, and **shoes**.

b) Write a PROC DATASETS step with the appropriate APPEND or SET statements to append shoes_eclipse and shoes_tracker to shoes.

c) Submit the PROC DATASETS step followed by a PROC PRINT to confirm that **shoes** contains 34 observations (10 original observations plus 14 observations from **shoes_eclipse** and 10 observations from **shoes_tracker**.

(Reminder: When working with procedures to modify data in the library, always develop your code using a temporary dataset in the work library and ensure that it is working as required before running it on shoes. If you develop and test your code using the shoes dataset as the target dataset, you may find that you will have created additional duplicate rows in shoes If this happens, you will have to re-create shoes back to the original 10 records by rerunning the creation of this dataset from the original source code)

Submit your program file as XX_A4_Q2.sas (replace XX with your first and last initials)
Submit a copy of the log as XX_A4_Q2_log.txt (replace XX with your first and last initials)
Submit a copy of the results as XX_A4_Q2_results.pdf (or .txt is acceptable) (replace XX with your first and last initials)

Question 3 (Interleaving Data Sets) **(2 marks)**

Copy the following starter program into a new program.

```
proc sort data=shoes_eclipse
out=work.eclipsesort;    by
Product_Name; run;
```

a) Add a PROC SORT step after the PROC SORT step to sort **shoes_tracker** by **Product_Name** to create a new data set named **work.trackersort**.

b) Add a DATA step after the two PROC SORT steps to interleave the two sorted data sets by **Product_Name** to create a new data set named **work.e_t_shoes**. Include only the following variables: **Product_Group, Product_Name**, and **Supplier_ID**.

c) Create the report shown below.

Partial PROC PRINT Output

Obs	Product_Group	Product_Name	Supplier_ID
1	Eclipse Shoes	Atmosphere Imara Women's Running Shoes	1303
2	Eclipse Shoes	Atmosphere Shatter Mid Shoes	1303
3	Eclipse Shoes	Big Guy Men's Air Deschutz Viii Shoes	1303
4	Eclipse Shoes	Big Guy Men's Air Terra Reach Shoes	1303
5	Eclipse Shoes	Big Guy Men's Air Terra Sebec Shoes	1303

Submit your program file as XX_A4_Q3.sas (replace XX with your first and last initials)
Submit a copy of the log as XX_A4_Q3_log.txt (replace XX with your first and last initials)
Submit a copy of the results as XX_A4_Q3_results.pdf (or .txt is acceptable) (replace XX with your first and last initials)

Question 4 (Subsetting Combining Data sets) **(2 marks)**

Run the program here to create a SAS data set called Markup:

```
data Markup;
input Manuf : $10. Markup;
datalines;
Cannondale 1.06
Trek 1.08
;
```

Edit the above program to combine this data set with the Bicycles data set so that each observation in the Bicycles data set now has a markup value of 1.06 or 1.08, depending on whether the bicycle is made by Cannondale or Trek. In this new data set (call it Markup_Prices), create a new variable (NewTotal) computed as TotalSales times Markup.

Submit your program file as XX_A4_Q4.sas (replace XX with your first and last initials)
Submit a copy of the log as XX_A4_Q4_log.txt (replace XX with your first and last initials)
Submit a copy of the results as XX_A4_Q4_results.pdf (or .txt is acceptable) (replace XX with your first and last initials)

Question 5 (Arrays, Numeric and Character Functions) **(2 marks)**

Note that data sets needed for this question are available in your BAN130 Directory

a. The SAS data set Psych contains an ID variable, 10 question responses (Ques1-Ques10), and 5 scores (Score1-Score5). You want to create a new, temporary SAS data set (Evaluate) containing the following:

i). A variable called QuesAve computed as the mean of Ques1-Ques10. Perform this computation only if there are seven or more nonmissing question values.

ii). If there are no missing Score values, compute the minimum score (MinScore), the maximum score (MaxScore), and the second highest score (SecondHighest).

b. List all patients in the Medical data set where the word antibiotics is in the comment field (Comment.

c. Data set Survey2 has five numeric variables (Q1-Q5), each with values of 1, 2, 3, 4, or 5. You want to determine for each subject (observation) if they responded with a 5 on any of the five questions. This is easily done using the OR or the IN operators. However, for this question, use an array to check each of the five questions. Set variable (ANY5) equal to Yes if any of the five questions is a 5 and No otherwise.

Submit your program file as XX_A4_Q5.sas (replace XX with your first and last initials)
Submit a copy of the log as XX_A4_Q5_log.txt (replace XX with your first and last initials)
Submit a copy of the results as XX_A4_Q5_results.pdf (or .txt is acceptable) (replace XX with your first and last initials)