SAS Problem Session 1

Use the Problem Session Data on Moodle to download the PS_data.zip file. Place the file in the relevant directory and unzip it. Create a library called PS_data. This problem session consists of two parts.

Part 1

Scenario 1

Directions

This scenario uses the PS_Data.Stress data set. Write a SAS program to do the following:

- Create a new temporary SAS data set that uses PS_Data.Stress and store the results in Work.Stress1.
- Remove observations with RestHR values that are greater than or equal to 70.
- Create a new variable called TotalTime. The value of TotalTime is the value of TimeMin multiplied by 60, plus the value of TimeSec.
- Remove TotalTime values that are less than 600.
- Sort the Work.Stress1 by descending TotalTime, and create a new data set called Work.Sorted

Test Your Code

1.- How many observations are in Work.Sorted?

2. What is the value of TotalTime for observation 3 in Work.Sorted?

Scenario 2

Directions

This scenario uses the PS_Data.Staff data set. Write a SAS program to do the following:

- Create a new temporary SAS data set that uses PS_Data.Staff and store the results in Work.StaffReports.
- Select observations where WageCategory is not equal to H.
- Format the variable DOB as mmddyy10.
- Create a new variable named Raise whose value is WageRate multiplied by 3%.
- Using Proc Means, determine the mean of the variable Raise

- Sort the Work.StaffReports by ascending ID, and create a new data set called Work.StaffId. Did it sort the data? Why?

Test Your Code

1. In the Work.StaffReports data set, for observation 5, what is the value of DOB?
2. In the Work.StaffReports data set, for observation 15, what is the value of Raise? Round your answer to 2 decimal places.}
3. What is the mean of the variable Raise?

Part 2

Scenario 1

Directions

Open the PS_4.sas program (posted on Moodle) and correct the errors in the program.

Example Code 1 ehs01 Program: Fix the Errors

```
data work.aprilbills drop=Total, EquipCost;                  /* #1 */
   set PS_Data.aprbills;
   if Days > 7 then Discount=(RoomCharge)*20% else 0;        /* #2 */
   TotalDue=Total-Discount;                                  /* #3 */
   format DateIn DateOut date9;                              /* #4 */
   format RoomRate RoomCharge Discount TotalDue dollar10.;   /* #5 */
   proc print data=work.aprilbills;                          /* #6 */
```

Here are instructions of what the program is trying to accomplish:

1. Drop the variables Total and EquipCost.
2. If the Days variable is greater than 7, then Discount is the value of RoomCharge multiplied by 20 %. If the Days variable is less than or equal to 7, then Discount is set to 0.
3. Create a new variable, TotalDue, with a value of Total minus Discount.
4. Format DateIn and DateOut to appear as 05APR2009.
5. Format the variables RoomRate, RoomCharge, Discount, and TotalDue to appear as $100.00.
6. Print your results.

Test Your Code

1. What is the value of the variable TotalDue in observation 4?
2. What is the value of the variable Discount in observation 5?

Scenario 2

Directions

This scenario uses the PS_Data.Temp18 data set. Write a SAS program to do the following:

- Create a temporary SAS data set that uses PS_Data.Temp18 and store the results as Work.Scenario2.
- Format the Day variable so that the date appears as 01JAN2018.
- Use a function to create a variable named Month that is equal to the numeric value of the month of the Day variable. For example, if the month is January, Month=1, if the month is February, Month=2, and so on.
- Create a one-way frequency table using the variable HighTemp.
- Use PROC MEANS to calculate the mean and standard deviation for the variables AvgHighTemp and AvgLowTemp by the new Month variable.

Test Your Code

1. What is the frequency for a HighTemp of 63?
2. What is the HighTemp on January 12, 2018?
3. What is the mean for AvgLowTemp for Month=1? (Round your answer to the nearest integer.)
4. What is the standard deviation (std) for AvgHighTemp for Month=3? (Round your answer to two decimal places.)