**EPI 5143 Winter 2024 QUIZ 1**

**Due Tuesday February 6th, 11:59pm via Github**

**Answer all the questions by writing the necessary SAS code and producing any output from SAS procedures used. Provide the SAS code, the SAS log, and SAS output with your quiz solutions.**

***You are expected to do your own coding and submit your own independent work—no copying/collaboration is permitted for quizzes and exams, we will be monitoring this closely.***

1. Save the quiz1 data (quiz1\_data.sas7bdat) on your computer in your course data folder.

Make it read only so you don't accidentally change it.

(this is a suggestion not for marks).

2. Create a new permanent SAS library called quiz1 that points to the folder on your computer where you saved your quiz1 dataset (you could also write protect the SAS library to prevent changing the original quiz data).

(ie so to reference this dataset in SAS it would be quiz1.quiz1\_data).

LIBNAME Quiz1 '/home/u63577751/Quiz1/';

data Quiz1.quiz1;

set Quiz1.Quiz1\_data;

run;

3. Use PROC CONTENTS to find out some information about this dataset.

proc contents data=Quiz1.quiz1\_data;

run;

1. How many observations does the dataset have?

The dataset contains 1400 observations

1. How many variables does the dataset have?

The dataset has 9 variables

4. Use PROC FREQ to provide information about the variable diabetes. If this variable represents those individuals in the dataset with diabetes, what proportion of people in the dataset have diabetes?

proc freq;

table mentalhealth \* (sex community health income DHHGAGE) / nopercent norow cmh;

run;

proportion= individuals with diabetes / total sample

proportion= 190/1400

proportion= 0.13

5. Use PROC UNIVARIATE to provide information about the variable X1.

proc univariate data= Quiz1.quiz1\_data;

variable X1;

histogram X1;

run;

1. What are the mean and standard deviation of X1?

Mean= 0.81

Standard Deviation= 1.01

b) Produce a frequency histogram of X1.

A graph of a distribution of numbers

Description automatically generated

6. Create a temporary copy of the quiz1 dataset called work.quiz1.

proc univariate data= Quiz1.quiz1\_data;

variable X1;

histogram X1;

run;

The remainder of the questions involve working with the work.quiz1 dataset.

7. a) Create a new variable called sum\_V1 that is the sum of X1, X2 and X3 using mathematical operators.

data work.quiz1;

set Quiz1.quiz1\_data;

sum\_V1=X1+X2+X3;

run;

1. Create a new variable called sum\_V2 that is the sum of X1, X2 and X3 using a SAS function.

data work.quiz1;

sum\_V1=X1+X2+X3;

sum\_V2=sum(X1,X2,X3);

run;

8. Consult\_dt and Surgery\_dt are SAS dates. Create a new variable called wait\_time that calculates the time in days between consult and surgery.

data work.quiz1;

wait\_time= Surgery\_dt-Consult\_dt;

run;

9. Create a new variable called X2\_high which has a value of 1 if X2 is greater than or equal to the overall mean of X2 and 0 otherwise (you can find the mean of X2 using PROC UNIVARIATE).

data work.quiz1;

set work.

X2\_high=0;

if X2>=11.97 then X2\_high=1;

run;

10. a) Use PROC UNIVARIATE to determine the mean and standard deviation of the variable sum\_V2, and the median, minimum and maximum values for wait\_time.

Sum\_V2

Mean= 36.80

SD= 1.72

Wait\_time

Median= 48

Minimum= 0

Maximum=99

1. Use PROC FREQ to create a 2x2 frequency table for X2\_high vs. diabetes.

A screenshot of a medical procedure

Description automatically generated