**EPI 5143 Winter 2023 QUIZ 1**

**Due Friday February 10th, 5pm via Github**

**Answer all the questions by writing the necessary SAS code and producing any output from SAS procedures used. Provide both the SAS code file, the SAS log file and requested 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.

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

See SAS code, lines 6-7

3. Use PROC CONTENTS to find out some information about this dataset. How many observations does the dataset have? How many variables does the dataset have?

See SAS code, lines 9-11

This dataset has 2334 observations and 7 variables.

Table

Description automatically generated

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? (provide the frequency table with your answers).

There are 31 people with diabetes in the total population (n=1150). Hence, those with diabetes represents only 2.70% of the total sample.

Table

Description automatically generated

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

a) What are the mean and standard deviation of X1?

For variable X1, mean=1.17 and std=0.96

Table

Description automatically generated

b) Produce a frequency histogram of X1 (provide with your answers).

Chart, histogram

Description automatically generated

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

See SAS code, lines 35-38

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

7. a) Create a new variable called mean\_V1 that is the mean of X1, X2 and X3 using mathematical operators. See SAS code, lines 42-46

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

See SAS code, lines 50-54

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.

See SAS code, lines 58-62

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

See SAS code, lines 68-78

10. a) Use PROC UNIVARIATE to find out the mean values of the variables of mean\_V1, and mean\_V2, and the median, minimum and maximum values for wait\_time.

See SAS codes, lines 83-86

The mean of variables 'meanv1' and 'meanV2' is 12.74.

The median and min/max values for var ‘wait\_time’ are:

* Median: 52.00
* Min value: 0
* Max value: 99

b) Use PROC FREQ to create a 2x2 frequency table for X1\_high vs. diabetes. (provide frequency table with your answers).

Table

Description automatically generated

See SAS code, lines 93-96