**EPI 5143 Winter 2022 QUIZ 1**

**Due Tuesday February 8th, 5pm via Github**

**Answer all the questions by writing the necessary SAS code and providing the relevant output from SAS procedures used. Provide both the SAS code and SAS output with your quiz solutions.**

**Retrieve the Quiz1 data from this link:**

[**https://www.dropbox.com/s/oigi44pji5km2ru/quiz1\_data.sas7bdat?dl=0**](https://www.dropbox.com/s/oigi44pji5km2ru/quiz1_data.sas7bdat?dl=0)

1. Save the quiz1 data 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 quiz and specify the

physical directory on your computer where you saved your quiz1 dataset.

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

3. Create a new permanent dataset in your quiz library which is an (approximate) 50% subsample of the quiz dataset named quiz.quiz1\_data\_subset (use the ranuni function to selectively read in observations to the new dataset. Specify a positive integer seed so the random subset can be reproduced).

4. 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?

5. 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?

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

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

b) Produce a frequency histogram of X1.

7. Create a copy of the quiz1\_data\_subset dataset called work.quiz1.

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

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

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

9. 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.

10. Create a new variable called X1\_level which has a value of 1 if X1 is greater than or equal to 1.25 and 0 otherwise.

11. a) Use PROC UNIVARIATE to find out the mean of sum\_V1, sum\_V2, and wait\_time for the work.quiz1 dataset.

b) Use PROC FREQ to find out the frequency count for the X1\_level variable.