BINF 5210:

Project 1: Due Dec 16th (12pm)

A study was conducted to determine whether a "DRUG A" helps to reduce fasting sugar level. There are 3400 patients were enrolled for this study and before and after fasting sugar level were taken at specific interval.

About Raw files:

There are two files "First Study_Project1" and "Second Study_Project1".

Initial Study_Project1:

Variables	Datatype
Patient ID	???
Age	???
State	????
Length of Stay	????
Total Charge	???
Initial_Sugar	???

Second Study_Project1:

Variables	Datatype
Patient ID	???
After_Sugar	???

Follow steps:

1). Merge both the files

2). sampling: Method SRS (Random sampling), Sample size 1000 (Don't use seed)

3). Run descriptive statistics

4). Design your research question.

5). Create PowerPoint presentation

- Explain your findings (descriptive statistics)
- Write your hypothesis (Null and alternative) and explain your results.

Submission: (3 files)

1). Data set (After sampling)

2). SAS Code

3). Report(PowerPoint slide)

Please submit in the drop box

Hint**

Design your research question:

To formulate your research question, you need to evaluate these files and ask your self-following question?

1) What kind of information will you be able to get from this dataset? 2) Can you use any comparison method (ANOVA, t-test or Chi sq)?

You have to write your research question in the form of a hypothesis. If you are using T-test, you need to explain why you are using T-test? What kind of assumption you made? Is your data normally distributed? (run univariate test)? [Note: please assume that your data is normally distributed and start your analysis…]

Note: I am not going to evaluate your result. But I want to see your understanding for different methods