BINF 5210:

Project 3: 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. They were randomly assigned to one of three treatment groups: Low-Drug A, High_Drug A and Placebo. After 10 weeks their reading were taken. Patients were placed in to three diff sites.

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	???

Second Study_Project1:

Variables	Datatype
Patient ID	???
Site	???
Group	???
Test_Score	???

Follow steps:

1). Merge both the files

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

3). Run descriptive statistics (if needs)

4). Design your research question.

5). Power point Presentation
- Write your hypothesis (Null and alternative) and explain your results.

Submission: (3 files)

1). Data set (After sampling)

2). SAS Code

3). Powerpoint

Hint**

Design your research question:

1) What kind of information will you be able to get from this dataset?
2) Is there any difference in the mean test scores among the three groups across all three sites?

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 normally distributed? (run univariate test)? [Note: Please assume 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 of different methods.