STAC32

Assignment 8

Due Thursday November 14 at 11:59pm

Question 2 below gives you instructions for setting up and structuring your SAS assignments. If your assignment is disorganized or otherwise difficult for the grader to deal with, you can expect to lose marks.

See `https://www.utsc.utoronto.ca/~butler/c32/quercus1.nb.html` for instructions on handing in assignments in Quercus. You can (and for SAS assignments, *should*) hand in a Word document.

Reminder: *there are no extensions due to failure to access software*. It is *your* responsibility to make sure that you allow yourself enough time to get connected to SAS Studio and to get your work done. You will be competing with a lot of other people, here and around the world, for access to SAS's servers, so it is up to you to allow enough time.

This is the last Assignment that you'll need to hand in. I will point you towards PASI-SAS for practice problems on the remaining material, and a reminder that your Project will be due at 11:59pm on the last day of classes (December 2).

1. Work through Section 2 of PASI-SAS, `http://ritsokiguess.site/STAC32/pasi-sas.pdf`. (You can also get to this from the course website, using the link at the top.) You *definitely* need to do question 2.1, or else you won't be able to do anything else.

2. To prepare yourself to hand in questions on SAS, work through Section 1 of PASI-SAS. (I probably have sections 1 and 2 the wrong way around, but so it is.) You need to *do* the first set of seven bullet points, and you need to *read and understand* the second set of three bullet points. If you don't do the first set of bullet points, you won't be able to hand in this Assignment properly.

   Hand the next question in:

3. A statistics course had two lecture sections, taught by the same instructor. One section used a typical variety of examples from different subject areas, while the other section used only sports-themed examples. The lecture sections were labelled as such, so that a student knew which type of section they were enrolling in, and could choose the one they preferred. The sections are labelled `Regular` and `Sports` in the data file. The sports-themed section had its lectures earlier in the day than the regular section.

   The data are in `https://www.utsc.utoronto.ca/~butler/c32/SportsExamples.csv`. For each student, the total number of points earned in the course was recorded, as well as which section they were in and their letter grade.

   (a) (2 marks) Read the data into SAS, and display the dataset. (It has 57 observations, which is rather longer than I like you to display normally, but the grader will check only that you have apparently the right thing.)

   (b) (2 marks) Display the mean `TotalPoints` for each lecture section.

   (c) (3 marks) Obtain a boxplot of total points for each section. Does one of the sections have a clearly higher average than the other? Explain briefly.

   (d) (2 marks) What is it about the design of this study that would make you hesitant to say that having only sports examples is a bad idea? Explain briefly.