STA302H1F / 1001HF Autumn 2020 Assignment # 1
Posted by: Dr. Shivon Sue-Chee on Sunday, September 13, 2020

Due: In **peerScholar** by 8pm on Saturday, September 26, 2020.
Late assignments will be subjected to a penalty of 20% per day late. Submissions will not be accepted beyond 48 hours of the due date. Email submissions are not allowed.

Instructions:

- Use R Studio to create two files:
  1. An Rmarkdown file with your codes according to the standard format in the A1_a20_RMForm.Rmd file.
  2. A corresponding report in html or pdf format.
- Two separate submissions into peerScholar are required. Please find the links in our website.
- This assignment will be subjected to a peer review. For the sake of privacy, please avoid revealing your full identity in your written work; instead use your initials and up to the last four digits of your student number.
- Presentation of your report is very important. Do not show R codes unless it is required for your solutions. Only required numbers and plots should be shown. Extraneous output should be hidden. Use options, include=FALSE, echo=FALSE, message=FALSE, where necessary.
- Write and submit **your own work**. For instance, personalized your code as much as possible, using your initials. **All plots produced must be given a title with the last 4 digits of your student number**.
- If you worked with a group to obtain data, **specify your group code or name** in your introduction.
- Use a benchmark significance level of 5%. Report p-values to 4 decimal places.

Grading Scheme: A grading rubric will be provided for this assignment.

The Data
The data should be the forearm length and height measurements, in centimetres, of 9 or 10 or 11 adults of our class, including yours. *You may get data during live class or a data file provided by September 17.*

Based on your sample data, complete an RMarkdown file and the corresponding report with the following sections.

I. Introduction section.

- Using words, give a title to summarise what your assignment #1 is about. This title should also be the title of your RMarkdown file.
- Describe how your sample data was collected.
- Specify your explanatory variable and response variable and, give a brief explanation of your choice.

II. Exploratory data analysis section.

- Draw at most 3 plots to visually describe your data. Is your response variable approximately Normal?

- Numerically describe the centre, spread and any unusual points of your variables/data.

III. Methods and Model section.

- Fit and describe a linear regression model between forearm length and height.
- Are the regression parameters zero?
- Interpret the estimates of the regression parameters.

IV. Discussion and Limitations section.

- Identify a potential lurking variable. Describe at least one other issue or limitation of your fit. You may include at most one plot to support your answer here.
- Identify another pair of variables to explore a simple linear regression model. Specify the response variable.