Assignment 02

Out: 03/22
Due: 04/05

Instructions

Collaboration:

Collaboration on solving the assignment is allowed, after you have thought about the problem sets on your own. It is also OK to get clarification (but not solutions) from online resources, again after you have thought about the problem sets on your own. There are two requirements:

- Cite your collaborators fully and completely (e.g., "XXX explained to me what is asked in problem set 3"). Or cite online resources (e.g., "I got inspired by reading XXX") that helped you.
- Write your scripts and report independently - the scripts and report must come from you only.

Late Submission:

Late submissions will not receive full credit. Half credit will be awarded to correct solutions submitted within 24 hours of the original deadline. Otherwise, no credit will be given.

Submitting your assignment:

Submit your scripts and report via mail (pudc2020@mail.sustech.edu.cn) .

---

1. [15 points] [RS] *The Statistical Sleuth: A Course in Methods of Data Analysis*
   P51 2.7 EXERCISES 15.

2. [20 points] Load rivers data sets (length of 141 rivers in North America)

(1) Make a density plot, a Q-Q plot, and a box plot;

(2) Is the sample from a normal distribution? If not, can you make a normal transformation for the sample?

3. [25 points] Load beaver1 and beaver2 data sets (body temperature data of 2 beavers every 10 minutes)

(1) Visualize these two data sets in multiple ways;

(2) Check the normality of the two samples;

(3) We need to check whether the body temperatures of the two beavers are the same. What is the H0 and H1? Can H0 and H1 be exchanged?

(4) Do you use a Z-test or a t-test? Report your findings.

4. [15 points] Suppose the scores of students obey a normal distribution, from which the scores of 36 students are randomly selected, the average score is 66.5, the standard deviation is 15, and the question is whether it can be considered that the average score of all candidates in this test is 70 points at the significance level $\alpha = 0.05$? And give the inspection process.

5. [25 points] Analyze the data you selected in the seventh question of the first assignment again.

(1) Select one or more features and visually analyze the data based on tidyr and ggplot packages.
(2) Propose a hypothesis test problem and solve it.