

Statistics 3080
Homework 5
Due: Wednesday, February 28

Complete the following problems in a R Markdown file and submit your compiled PDF.

Problem 1: For each part of this problem, you should first recreate the graphic identified and save it as `Graphic1X`, where `X` is the part letter. If your original graphic was not fully correct, you should modify the code that you use to create `Graphic1X`, so that you start with the correct graphic. Next, make the changes detailed using layers that are added to `Graphic1X`. You may need to reuse select layers from `Graphic1X`.

- (a) Recreate the graphic from Homework 3 Problem 2f. Modify the graphic such that its overall description is “Evening activity at home”, the x-axis is “Time sitting (minutes)”, the y-axis is “Time watching TV (minutes)”, the legend description is “Regular exercise”, the legend values are labeled “Yes” and “No”, and the legend is within the graphic at the top left, but does not cover any points.
- (b) Recreate the graphic from Homework 4 Problem 1c. Modify the graphic such that its overall description is “Reaction times after days of sleep deprivation” and the x-axis markers are the integers from 0 to 9.
- (c) Recreate the graphic from Homework 4 Problem 2b. Modify the graphic such that its overall description is “Yield of mature plants”, the x-axis is “Treatment”, the y-axis is “Median yield with interquartile range”, and the categories are labeled “Control”, “Leaf reduction 25%”, “Leaf reduction 50%”, “Root trimming 10%”, “Root trimming 5%”, respectively. Additionally, rotate these labels 45 degrees and shift them down such that no part of the label is within the gray area of the graphic.
- (d) Recreate the graphic from Homework 4 Problem 3. Modify the graphic such that its overall description is “Distribution of durations of Old Faithful eruptions”, the x-axis is “Eruption duration”, and the y-axis is “Density”.

Problem 2: Note: Generalize as much of your code as possible.

- (a) Draw a random sample of 15 individuals from a Normal distribution with a mean of 77 and a standard deviation of 31. Save the sampled values in a vector named `samp.15`.
- (b) Draw a random sample of 30 individuals from the same Normal distribution. Save the sampled values in a vector named `samp.30`.
- (c) Draw a random sample of 45 individuals from the same Normal distribution. Save the sampled values in a vector named `samp.45`.
- (d) Use each set of sample values to conduct a z-test for whether the true mean is equal to 77 or not. Determine whether the null hypothesis would be rejected or not in each of the three tests and explain how you came to each conclusion.
Note: You should **not** use a built-in testing function.

Problem 3:

- (a) Using a single loop, repeat Problem 2 10,000 times. Record only whether the null hypothesis would be rejected or not for each sample size in each iteration. In other words, you should have 30,000 recorded values when the loop is finished - 10,000 for each sample size. These values do not all need to be stored in the same data structure.
- (b) Determine the proportion of times the null hypothesis would be rejected for each sample size.
- (c) Theoretically, what should the proportions be? Explain.