

Homework 2 – Intro. to Computational Statistics

For all problems, please show all your work. As described in the Homework Guidelines, use RMarkdown to write up your work as a .Rmd file, “knit” the result to a PDF file, and submit that PDF file to Blackboard. (You can also knit to an HTML or Word document and save that as a PDF, as described in the Homework Guidelines, but knitting directly to PDF is recommended for greater compatability.) Be sure to use R code for all your calculations, and the latex equation format to write up any math. See the Homework Guidelines in Course Resources on Blackboard for more formatting details.

1.

- a. Write a function that calculates the mean of any numeric vector you give it, without using the built-in `mean()` function.
- b. Write a function that takes as its input a vector with four elements. If the sum of the first two elements is greater than the sum of the second two, the function returns the vector; otherwise it returns 0.
- c. Write a function that calculates the Fibonacci sequence up to the n th element, where n is any number input into your function (its argument). The Fibonacci sequence is: 1, 1, 2, 3, 5, 8, 13, 21..., ie, each element is the sum of the previous two elements. One way to do this is to start off with the first two elements, `c(1,1)` and set an internal variable to this sequence. Then write a loop that counts up to n , where for each new element, you first calculate it by adding the last two elements of the growing sequence, and then stick that new number onto the growing sequence using `c()`. When the loop is finished, the function should return the final vector of Fibonacci numbers.
- d. Create a 4x4 matrix of the numbers 1 through 16. Use `apply` to apply you function from (a) to each of the rows in your matrix.

2.

- a. Using the `airquality` dataset, constuct an aggregated dataset which shows the mean wind and ozone by month.
- b. Create the `authors` and `books` datasets following the example and data in the lecture, and then create a new data set by merging these two datasets by author, preserving all rows.
- c. Take the following string and replace every instance of “to” or “To” with “2” :

```
To be, or not to be -- that is the question:
Whether 'tis nobler in the mind to suffer
The slings and arrows of outrageous fortune,
Or to take arms against a sea of troubles,
And by opposing end them. To die -- to sleep --
No more...
```

3.

- a. Create a histogram using the base R graphics using some dataset or variable other than the one in the lessons. Always make sure your graph has well-labeled x and y axes and an explanatory title.
- b. Create a scatter plot using the base R graphics, again with some variable other than the one in the lessons.

- c. Create a histogram using ggplot, using some new data. In this and the later plots, feel free to tinker with the settings using the examples in <http://www.cookbook-r.com/Graphs/> to make it prettier.
- d. Create a box plot using ggplot, using some new data.
- e. Create a scatter plot using ggplot, using some new data.