Charlee Cobb - Transcriptomics, Exercise 2

netID: cac9995

Charlee Cobb

2023-01-17

M2_1.1) Give three reasons you would want to create a function. - Writing your own function allows you to include complicated steps of an algorithm you want to employ. You can also simplify your workflow by putting regular or repetitive steps of your analysis or task into a function. And finally you can take the function you made and use the apply function to use your function on a number of elements at once.

M2_1.2) What are the three different parts of a function? - The three main parts are the name, the arguments that will pass through the function, and the expression that contains the code/commands you want to use on the arguments.

M2_1.3) What is the purpose of the return() command? - The return() function is useful so you know what is being sent back from the function you have written. Depending on your function's expression, R won't automatically return your modified values.

M2_2.1) What is the purpose of curly brackets {}? - The curly brackets {} are used to contain the expression of your function. They are the start and end of the function.

M2_2.2) What is the difference between "if else" and "ifelse" statements? - "if, else" statements are boolean arguments checking if a variable is TRUE or FALSE. When using "if else" you are checking the condition of one statement, and then the function will execute. When the condtion does not match the "if" statement, then it will go to the "else statement."ifelse" can check all the values in a vector in one line.

M2_3.1) In what scenario would you want to use a while loop instead of a for loop? - You would use a while loop instead of a for loop when you want the loop to continue until a certain condition is no longer true. For for loops, you give a set number of iterations. So if you don't know how long you want something to loop, but you know it has to continue until a condition is false, use a while loop.

M2_3.2) How do you stop a repeat loop? - To stop a repeat loop, you use the "break" command after the "if" condition.

M2_4.1) Explain in words, what will the following command do? 'apply(data, 2, min)' - The 'apply(data, 2, min)' is using the apply function to apply the function min() on every column (2 means column, 1 means row) in the dataset named "data". The min() function will return the minimun value in each column of 'data'.

M2_4.2) Why is it better to use apply() function instead of looping? - apply() is a better method than looping because it's not necessary to stop and start iterations all over again, especially in the case from above where one iteration won't impact or change the values in the next iteration. This makes going through a large data set more efficient.

M2_4.3) What is the first argument of the apply function? - The first argument of the apply function is the name of the data object you want to manipulate, likely an array or matrix.

M2_5.1) What are commands to create a pdf document with your plot? To create a pdf document with your plot, you can use the pdf() command to open a pdf file, then use dev.off() when you are finished running the command or script that creates the plot.

M2_5.2) How many plots will you get in a page if you started with: par(mfrow=c(2,3))? - There will be 6 plots drawn with the command par(mfrow=c(2,3)) because we are asking for 2 plots in the 3 rows of the page.

M2_5.3) How do I draw a line on a pre-existing plot? - You can add addtional lines with the lines() function and add points in the plot as your arguments. The abline() function will draw straight lines over your plot, in diagnol, vertical, and horizontal directions.

M2_6.1) What is the difference between a barplot and a histogram? - A barplot is great for comparing data values by category, while a histogram shows numerical data that's binned in a given set of ranges.

M2_6.2) What statistical values are presented about the data in a boxplot? - In a box plot, you can see the distribution your data in the form of quartile ranges. The boxes represent median and inter-quartile ranges, and the whiskers show the outer percentiles.

M2_7.1) What command is useful to look at the first few rows of a matrix or dataframe? - You can use the head() command to look at the first couple rows of a matrix or data frame. By default, it's 6 rows, but you can add arguments to see more rows. tail() will look at the last 6 rows.

M2_7.2) Which command will be useful to identify all occurrence of a pattern? - To find all occurrences of a pattern, use the grep() function to find the pattern of a string, which will give us the occurrences of that pattern. By default, it returns the position of the pattern, but you can pass the argument value = T to get the actual words that contain the pattern you're searching for.