

Instructions: For these weekly coding assignments, you will be asked to extend the examples from class to create custom code to answer the questions below. You will create an R code file that uses built-in datasets as the data sources. You will write the code, and an example showing that the code works. Be sure to include any packages in the code that are required for the functions to run (you may want to clear the environment in RStudio before your final check to make sure nothing is missing). The instructor will run the file to ensure that it works with no errors. Clearly label your code so it's clear which question/task is being responded to.

Submission:

A word document with any explanations (if needed), and a clearly labeled R code file.

Tasks/Questions:

1. Write a function to calculate the population standard deviation formula.
2. Write a function to calculate the spearman correlation formula.
3. Adjust the two-sample t-test from the lecture notes to:
 - a. Allow the user to input an assumed mean difference that is non-zero ($\mu_1 - \mu_2 \neq 0$).
 - b. Calculate a confidence interval on the mean difference
 - c. Calculate the unpooled version of the two-sample t-test
 - d. Calculate the test statistic and p-value for the dependent case
4. Create a function that applies several statistics to a single dataset (all columns) to create your own custom statistical summary function. Describe any limitations of your function.
5. Choose three distance metrics in addition to the Euclidean distance to apply to the iris dataset. Note: you may need to remove non-numerical or categorical encoded columns from the dataframe first. Describe the differences that you see between the methods. For some distance metrics, you may need to do additional preprocessing.
6. Apply three scaling methods to the iris dataset (see note above). Save the results as separate dataframes.
7. Describe two ways of converting a vector with only positive values into a $[-1,1]$ range. Write up the code for both methods. How do the outcomes differ? You can use mtcars or iris for data for this problem.