# DAT 510 Using Your Lab Environment: Introduction to R

|  |  |
| --- | --- |
| **Purpose:** | This lab introduces you to the use of the R statistical package within the Data Science and Big Data Analytics environment. After completing the tasks in this lab, you should be able to write and modify a simple script, execute a script, save files and open files, and exit R. |
|  | |
| **Tasks:** | You will complete the following tasks in this lab:   * Invoke the R environment and examine the R workspace * Write a simple script and observe output * Modify script and observe output * Exit the R environment   You are to answer the questions highlighted in **yellow** in rows 3, 5, 7, 8, and 9 in this document. This document will be the final submission for this lab assignment. When you save this document, please rename it as follows (substitute your last name for the word *Name*):  Name\_DAT-510\_Lab\_1.docx |
|  | |
| **References:** | References used in this lab are located in your [***Student Resource Guide Appendix***](http://snhu-media.snhu.edu/files/course_repository/graduate/dat/dat510/dat_510_student_resource_guide_appendix.pdf)*.* See the appendix for:   * R Commands – Quick Reference |

|  |  |
| --- | --- |
| **Step** | **Action** |
| 1 | **Invoke the R Environment:**  Log in to SNHU’s Virtual Lab environment.  Once logged into the SNHU Virtual Lab, you can access the RStudio environment by following these steps:   1. RStudio is accessed through the Start menu. 2. Double-clicking this icon will invoke the R environment. Make sure you access RStudio through the Start menu. Clicking on the **R** icon on the desktop will bring you to a different interface that will not be useful for this exercise. Screenshot depicting how RStudio is accessed through the Start menu |
|  | 1. Verify that you see the following once you have invoked the R environment: |
| 2 | **Write a script to create a Hello World program**   1. **Create a New File:**    * 1. In the upper left-hand corner of the RStudio window, click on **File**.      2. Choose **New File** and then **R Script**. This will open a new window, where you will be entering your code.   Screenshot of the RStudio window with the File, New File, and R Script options highlighted   1. **Write the Script**    * 1. At the blinking cursor, enter the following text exactly as below in the text box, paying close attention to spaces, capital letters, and symbols.   Hello <- function( name ) {  sprintf( "Hello, %s", name );  }   1. **Save the Script**    * 1. Click on the disc icon to save your script.   Screenshot of the RStudio window. An arrow points to the highlighted Disc icon.   * + 1. Choose the R folder, and name your file **hello**. Click **Save**. You can close out the window where you composed your script once it is saved.   Screenshot of the Save File window showing the R folder location. The file name hello is also highlighted. The Save button at the bottom is highlighted, indicating it needs to be clicked. |
| 3 | **Execute the Script:** For this part of the lab, you will be working in the console window of RStudio. In order to execute the script, you will be selecting the source file and then executing a command.   1. At the top of the RStudio screen, click on **Code**. 2. Choose **Source File...**   Screenshot of the RStudio window with the Code tab and Source File option highlighted   1. From the R folder, select **Hello** and click **Open**. 2. In the **console window**, at the prompt, enter Hello(“Your name”), entering your own name between the quotation marks. For example, enter hello(“Bob”) 3. You should see a statement appear. Enter that statement on the line below. \_\_\_\_\_\_\_”Hello, Michael"\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_   Congratulations! You have successfully written and executed a script in R! |
| 4 | **Modify a Script, Part 1 –** You will now go back into your original script, modify it, and observe the different results that you obtain when running the program.   1. Open your original file by clicking on **File** and then **Open File**. Select hello.R from the R folder by selecting it and clicking **Open**.   Screenshot of the RStudio window with the File tab and Open File option highlighted(   1. Using your mouse to navigate, modify your original script in the following way:      1. Save your changes using the file icon, using the same process as before. |
| 5 | **Execute Your Modified Script –** Following the same steps as above:   1. Click on **Code** and then **Source File**. Open the R folder and select the file you would like to run, and then click **Open**. 2. At the prompt in the console window, execute your script by typing Hello(“YourName”). For example, Hello(“Bob”) 3. What is the output? Enter it in the space below. \_\_\_\_\_\_”Hello, Michael, how is your day going?”\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ |
| 6 | **Modify a Script, Part 2 –** You will now go back into your original script, modify it again to introduce a second variable, and observe the different results that you obtain when running the program.   1. Open your original file by clicking on **File** and then **Open File**. Select hello.R from the R folder by selecting it and clicking **Open**. 2. Using your mouse to navigate, modify your original script in the following way, paying careful attention to punctuation, capitalization, and spaces:   Hello <- function( name, age ) {  sprintf( “Hello, %s, how is your day going – rumor has it you just turned %i years old!”, name, age);  }  (Note – all of the script that starts on the second line [highlighted in gray] will be located on ONE line within R. You should not start a new line. Enter it in as one long line of text.) |
| 7 | **Execute Your Modified Script –** Following the same steps that you used in activities 3 and 5, run your modified script.   1. Click on **Code** and then **Source File**. Open the R folder and select the file you would like to run, and then click **Open**. 2. At the prompt in the console window, execute your script by typing Hello(“YourName”, Age). For example, Hello(“Bob”, 22) 3. What is the output? Enter it in the space below.   \_”Hello, Michael, how is your day going – rumor has it you just turned 25 years old!”\_\_\_ |
| 8 | **Questions to Answer: For each of the following questions, enter your answer directly into the document.**  **Question 1:**  The original code only had one parameter between the parentheses – Function (name). Describe what that parameter was used for and your impression of how it worked when you ran the code. For example, after you ran the script, you were asked to type something like “Hello(“Gwen”)” to execute it. How does Hello(“Gwen”) relate to the original code you typed and the output you saw when you ran the script?  The parameter in the original code was name. The name parameter was the one object of the Hello function in the original code. I noticed when executing the function that the name parameter was placed in quotation marks. In the output, the entire statement was in quotation marks, due to the function we created, but the quotation marks around the name parameter disappear. |
| 9 | **Question 2:**  When you modified the code, you added an additional parameter – Function (name, **age**). How did this change affect what happened with the program? How were the execution and contents of the script different when you added a second parameter?  When I added the age parameter, the function required multiple inputs. There were now 2 instances in the resulting statement that could be modified by the input. The script was essentially only changed by adding age inside the first parentheses and by including %i in the function itself. |
| 10 | **Question 3:**  Why were “%i” and “%s” in the script? What purpose do you think those two symbols served? Why?  It seemed %s and %i were in the script to serve as placeholders for the parameters in the statement. Since the function was created to print the statement and fill in certain spaces with the parameters, the %s and the %i stood in to show the locations within the statement where the parameters will be filled. I am not sure the letters themselves matter or not, but I was thinking that perhaps the %s is a placeholder for text objects and the %i is a placeholder for numeric objects. |
| 11 | **Exit R:**   1. Execute the following command:   **q()**   1. R will ask you if you want to save your workspace. Answer no. |