Working with Data in R

Open RStudio and create a new project under your Module 1 folder and call it **Mod1Assign2.** After you have created the project, you will need to create a script to put all the code in. Remember: A script is created by clicking on File -> New File… -> R Script. Name this new script: **Mod1Assign2Script** and put the following information at the top of the script using comments (#):

```
# Name: First Last
# Module 1: Assignment 2
```

Note: All code should be written in the script which appears in the upper left box in RStudio. Do not type your code in the Console (the lower left box) as this will not save the code to be uploaded to Canvas.

Complete the following questions in your script using the example labels below (note: where it says "R Code for Question1 goes here" that this is where you are to put the code to answer Question 1):

```
#==Question 1
R Code for Question 1 goes here

#==Question 2
R Code for Question 2 goes here
....
```

Question 1: **Simple Vector Assignment**

1) Create an object called **student_num** and assign it the values 1 through 75.

2) Write the code to display the contents of **student_num**

Question 2: **Loading Packages and Plotting Data**

You should have installed the tidyverse in the discussion/homework for this module.

1) Write the code to load the tidyverse package.

2) Create an object called **sales.**

3) Add the following values to sales (in this order): **50, 150, 200, 100, 300, 250**

4) Create another object called **month**.

5) Add the following values to month (in this order): 1,2,3,4,5,6

6) Use the ***qplot*** function to create a scatterplot with **sales** on the y-axis and **month** on the x-axis.

7) Add the following comment to your script and answer the question on the same line:

```
#What month had the largest sales? What was the amount?
```

Question 3: **Working with Vectors and Data Types**

1) Create a vector called **months** and add the months of the year (e.g., Jan, Feb, Mar, etc.).

2) Add the following comment to your script and answer the question on the same line:

```
#What type of vector is this (hint: use typeof() or look in environment pane)?
```

3) Create a vector called **days** and add the days of the week (Mon.-Sat.) except for Sunday where you should enter the number 7.

4) Add the following comments to your script and answer the question on the same line:

```
#How is the number 7 treated in the vector (as a double or character)?
#Why?
```

5) Write the following code:

```
Days <- c(1:7)
```

6) Add the following comments to your script and answer the question on the same line:

```
#Do Days and days contain the same values?
#If they are different, what data type is Days and days?
```

Question 4: **Data Frames**

7) Create a new data frame called **Yearly_Sales.** This data frame will contain two vectors with name for the first vector being *month* and the second vector being *sales.* For *month*, type out the abbreviation for the 12 months of the year (e..g, Jan). For *sales*, add the following values: 150.25, 258.54, 268.55, 122.52, 987, 458.82, 667.23, 845.54, 586.78, 888.58, 756.12, 456.84.

8) Add the following comments to your script and answer the question on the same line

```
#Open data frame from the environment pane.
#Which month had the most sales?
#Which month had the least sales?
```

9) Finally, add the code to the script that would create a **qplot** with month on the x-axis and sales on the y-axis (hint: to select a column of data you use the $ sign as in *yearly_sales$sales*).

10) The last step is to upload the file named **Mod1Assign2Script.R** into Canvas. If you do not have a file that ends in .R, you have probably typed into the Console area and not the Script area. Email me and I can help you in correcting this problem.