Post Graduate Program in Data Science

Module 4 – R for Data Science – Graded Assignment 1

1. Create a data Frame including:

• Random data of 8 Participants including Names, Department, Experience, Age, and Mobile Number

• Store the data frame as “MHG”

• Select first 3 rows as first and last 3 rows as last

1. Using the given values, create a vector in R = c(11,10,22,-31,-15,8,13)

• Store the vector with your name.

• Find the length of the vector.

• Sort the vector in descending order

• Print the value “22” as an output.

1. Import the csv files named as April Revenue, June Revenue, May Revenue.
2. Create a new file called as ‘Total\_Revenue’ using rbind.
3. From Total\_Revenue, find the number of orders that is made on a Monday.
4. Find the Order Id which has the maximum amount.
5. Create a new dataframe named as ‘filter\_val’ and filter out the rows with order total greater than 7000.
6. Import the excel file named as ‘global\_sales’. Import all the 3 sheets separately. For the following questions use the sheet named ‘Orders’.
7. Using dplyr function find the number of orders ‘Southern Asia’ region.
8. Using dplyr function group the data based on total sales.
9. Using dplyr function find the top 3 ‘Market’ based on average profit.
10. Using dplyr select only the Customer Name, City and State. Store the output as ‘Location’.

**Instructions:**

1. This is a graded assignment.
2. Once the assignment is done submit the R file in the LMS.
3. For the questions 3b, 3c, 4a, 4b, 4c the answers should be written in a separate word file and submitted in LMS. All the best!