l\

# **Experiment No.1**

# **Title: Laboratory on Data Visualization**

**Aim:** i. To explore the dataset for different case study examples with different commands.

**ii.** To plot the Box plot and scatter plot.

1. What is the mean of “Temp” when “Month” is equal to 6?
2. How many observations are there in the given data?**Software used:** Programming language R.

**Code Statement:**

1. Write a single R code to display the answers for the following questions.

Case Study: Consider the “pollutant” data set.

1. Print last two rows of the data.
2. What is the value of Ozone in 47th row?
3. How many values are missing in Ozone column?
4. What is the mean of Ozone column excluding missing values?
5. Extract the subset of rows of the data frame where Ozone values are above 31 and Temp values are above 90. What is the mean of Solar.R in this subset?
6. What was the maximum ozone value in the month of May (i.e. Month is equal to 5)?
7. Write a single R code to display the answers the following questions

Case Study: Hair Eye color Data set

1. How many people have brown eye color?
2. How many people have Blonde hair?
3. How many Brown haired people have Black eyes?
4. What is the percentage of people with Green eyes?
5. What percentage of people have red hair and Blue eyes?
6. Write a single R code to display the answers for the following questions

Case study: Germination Data Set

1. What is the average number of seeds germinated for the uncovered boxes with level of watering equal to 4?
2. What is the median value for the data covered boxes?

Establish conclusions on the basis of available data and write them in the conclusion part.

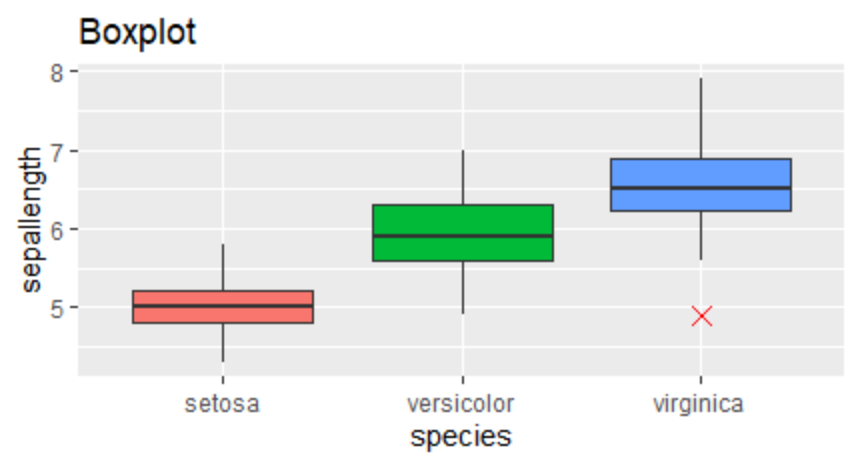
a. Association of levels of watering with the number of germinating seeds in case of covered boxes as well as uncovered boxes.

b. Association of number of germinating seeds with the fact that the boxes were covered or uncovered.

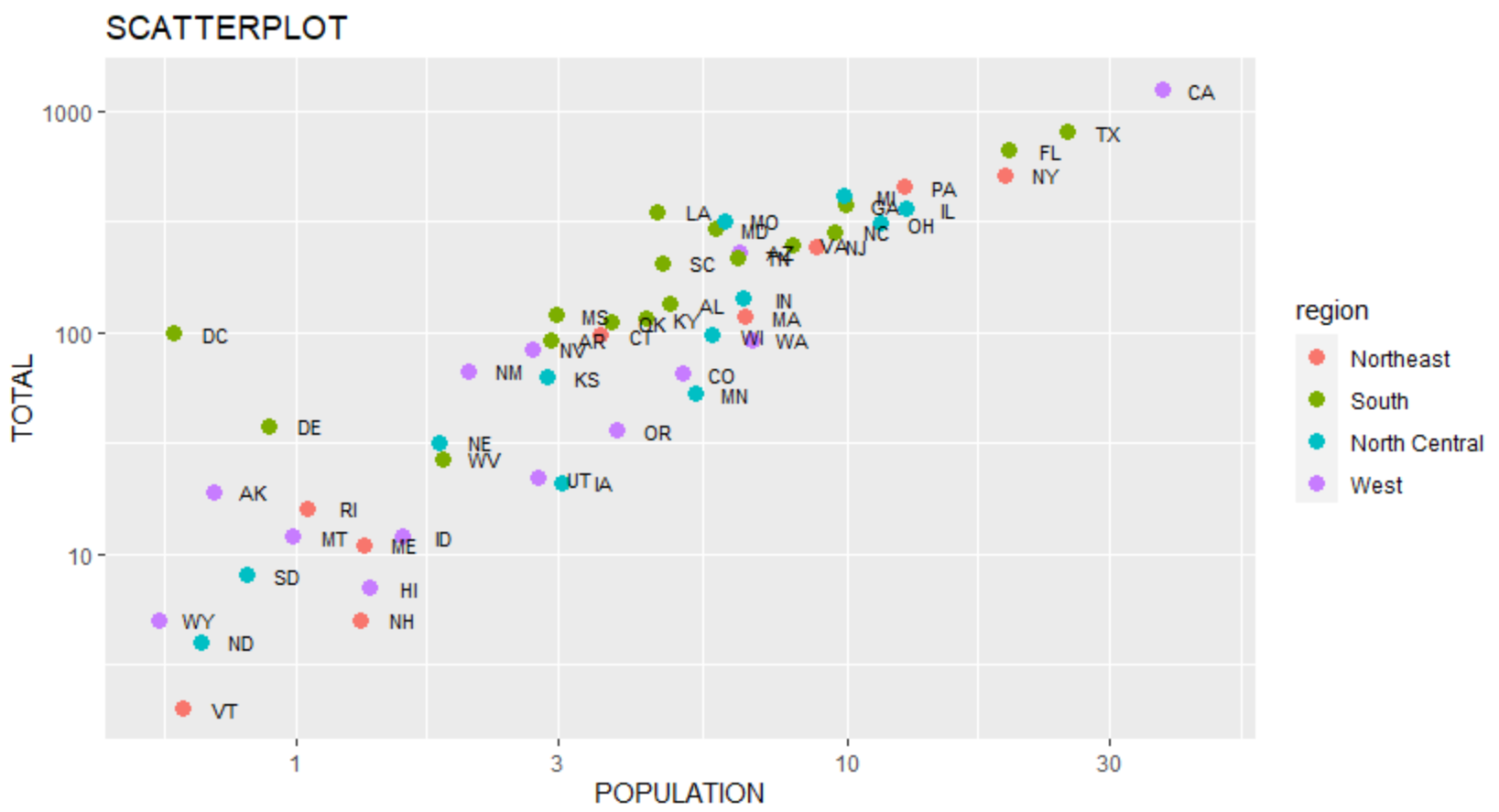
1. Write a single R code :
2. To display the Boxplot for sepal length of iris data set as shown below
3. To display the Scatter plot for murders data set present in “dslabs” package as shown below.

Give proper title, x,y axis label etc. to each plot.

**Expected Boxplot:**



**Expected Scatter Plot:**

****

**Code:** *(Write code for all above cases here)*

# **Results: (***Display the output obtained on R console for all the cases. Also add the plots which you obtained for 4th question above.* )

# **Conclusion: *(****Write the conclusion in your words).*