

BDA Experiment No.: 07

Aim: Data Visualizations using R.

Theory:

Data visualization is the technique used to deliver insights in data using visual cues such as graphs, charts, maps, and many others. This is useful as it helps in intuitive and easy understanding of the large quantities of data and thereby make better decisions regarding it.

Data Visualization in R Programming Language

The popular data visualization tools that are available are Tableau, Plotly, R, Google Charts, Infogram, and Kibana. The various data visualization platforms have different capabilities, functionality, and use cases. They also require a different skill set. This article discusses the use of R for data visualization.

R is a language that is designed for statistical computing, graphical data analysis, and scientific research. It is usually preferred for data visualization as it offers flexibility and minimum required coding through its packages.

Types of Data Visualizations

Some of the various types of visualizations offered by R are:

1. Heat Map

Heatmap is defined as a graphical representation of data using colors to visualize the value of the matrix. `heatmap()` function is used to plot heatmap.

Syntax: `heatmap(data)`

Parameters: data: It represent matrix data, such as values of rows and columns **Return:** This function draws a heatmap.

2. Scatter Plot

A scatter plot is composed of many points on a Cartesian plane. Each point denotes the value taken by two parameters and helps us easily identify the relationship between them.

Scatter Plots are used in the following scenarios:

- To show whether an association exists between bivariate data.
- To measure the strength and direction of such a relationship.

3. Bubble chart

A bubble chart is basically just a scatterplot where the point size is proportional to the values of a third quantitative variable.

Using the `mtcars` dataset, let's plot car weight vs. mileage and use point size to represent horsepower.

4. Map Visualization

The latest thing in R is data visualization through Javascript libraries. Leaflet is one of the most popular open-source JavaScript libraries for interactive maps. It is based at <https://rstudio.github.io/leaflet/>

5. Box Plot

The statistical summary of the given data is presented graphically using a boxplot. A boxplot depicts information like the minimum and maximum data point, the median value, first and third quartile, and interquartile range.

Box Plots are used for:

- To give a comprehensive statistical description of the data through a visual cue.
- To identify the outlier points that do not lie in the inter-quartile range of data.

6. Bar Plot

There are two types of bar plots- horizontal and vertical which represent data points as horizontal or vertical bars of certain lengths proportional to the value of the data item. They are generally used for continuous and categorical variable plotting. By setting the **horiz** parameter to true and false, we can get horizontal and vertical bar plots respectively. Bar plots are used for the following scenarios:

- To perform a comparative study between the various data categories in the data set.
- To analyze the change of a variable over time in months or years.

Program:

create a heatmap

```
data(mtcars) library(superheat)
superheat(mtcars, scale = TRUE)
# sorted heat map
superheat(mtcars,          scale
= TRUE,
left.label.text.size=3,
bottom.label.text.size=3,
bottom.label.size = .05,
          row.dendrogram = TRUE )
```

create a scatterplot3d

```
library(scatterplot3d)
with(mtcars, {
scatterplot3d(x = disp,
              y = wt,
              z = mpg,
              # filled blue circles
              color="blue",          pch=19,
```

```

      # lines to the horizontal plane
      type = "h",
      main = "3-D Scatterplot Example ",
xlab = "Displacement (cu. in.)",
ylab = "Weight (lb/1000)",      zlab =
"Miles/(US) Gallon")
})

```

create a bubble plot

```

data(mtcars) library(ggplot2)
ggplot(mtcars,
      aes(x = wt, y = mpg, size = hp)) +
  geom_point()

```

create a map visualization

```

library(magrittr)
library(leaflet) m
<- leaflet() %>%
  addTiles() %>% # Add default OpenStreetMap map tiles
  addMarkers(lng=77.2310, lat=28.6560, popup="The delicious food of chandni chowk") m
# Print the map

```

#create 3d graph

```

# Adding Titles and Labeling Axes to Plot
cone <- function(x, y){ sqrt(x ^ 2 + y ^
2)
}

```

```

# prepare variables. x <- y <-
seq(-1, 1, length = 30) z <-
outer(x, y, cone)

```

```

# plot the 3D surface
# Adding Titles and Labeling Axes to Plot persp(x,
y, z,
  main="Perspective Plot of a Cone",
  zlab = "Height",
theta = 30, phi = 15,
  col = "orange", shade = 0.4)

```

create a box plot

```

boxplot(iris$Petal.Length~iris$Species)
#Creating Box Plot between two variable
data(iris) par(mfrow=c(2,2))

```

```

boxplot(iris$Sepal.Length,col="yellow") boxplot(iris$Sepal.Length~iris$Species,col="yellow")
boxplot(iris$Sepal.Length~iris$Species,col=heat.colors(3))
boxplot(iris$Sepal.Length~iris$Species,col=topo.colors(3))

```

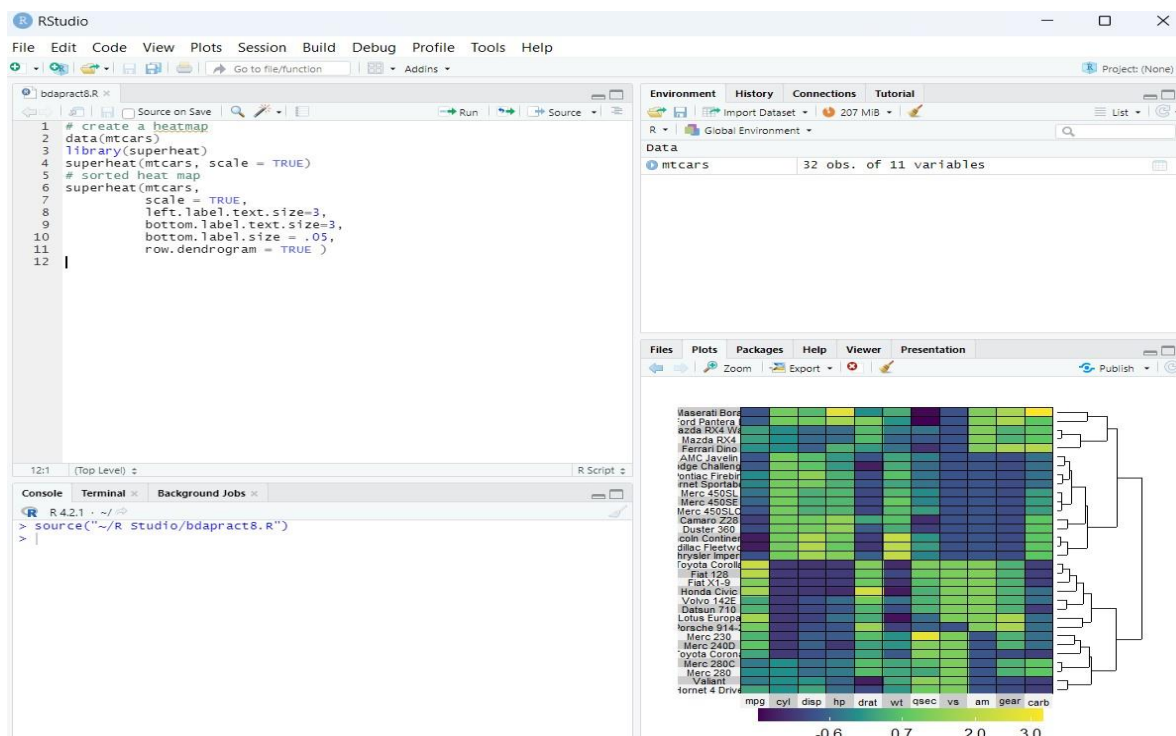
create a bar plot

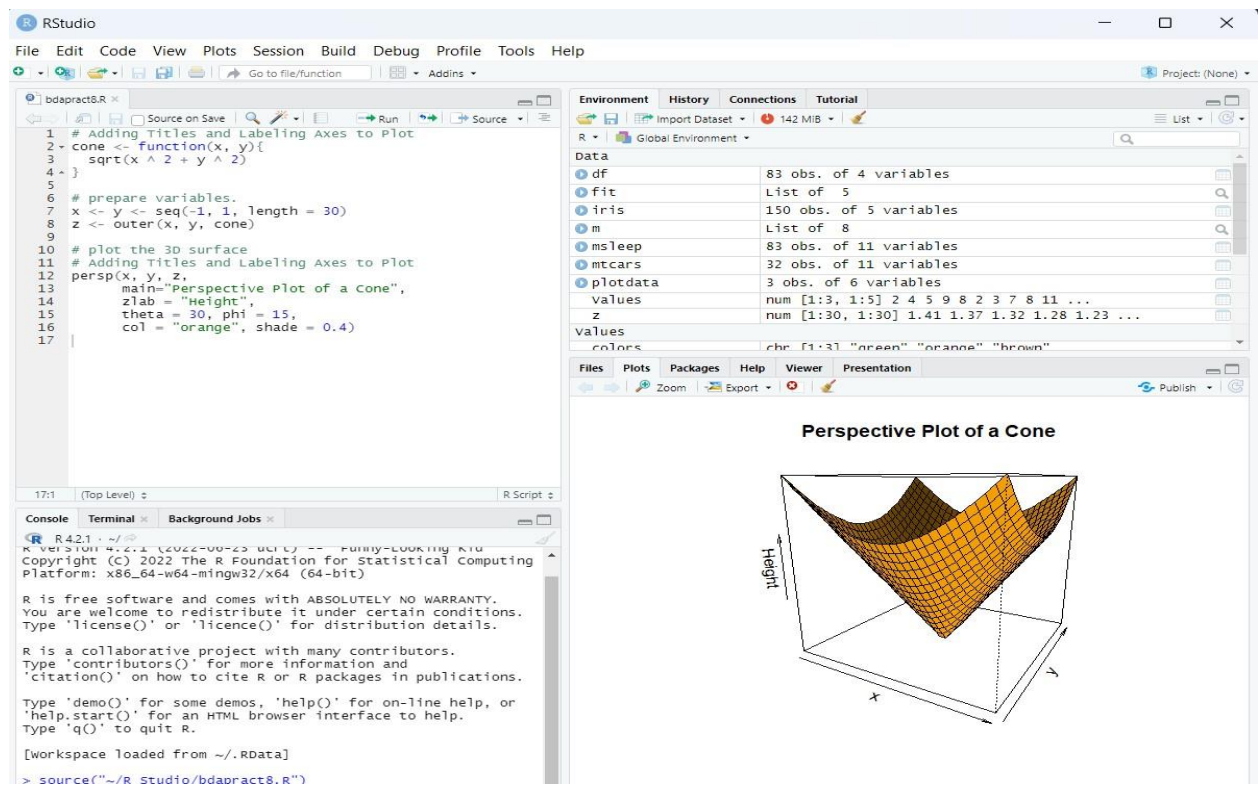
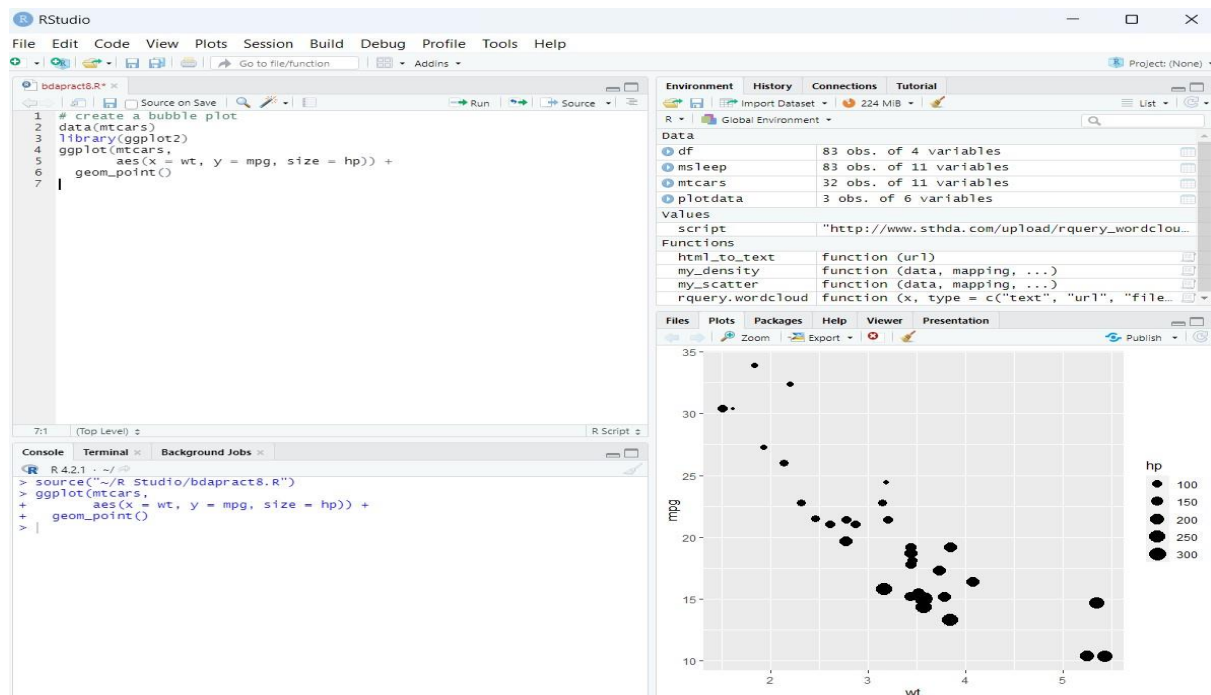
```

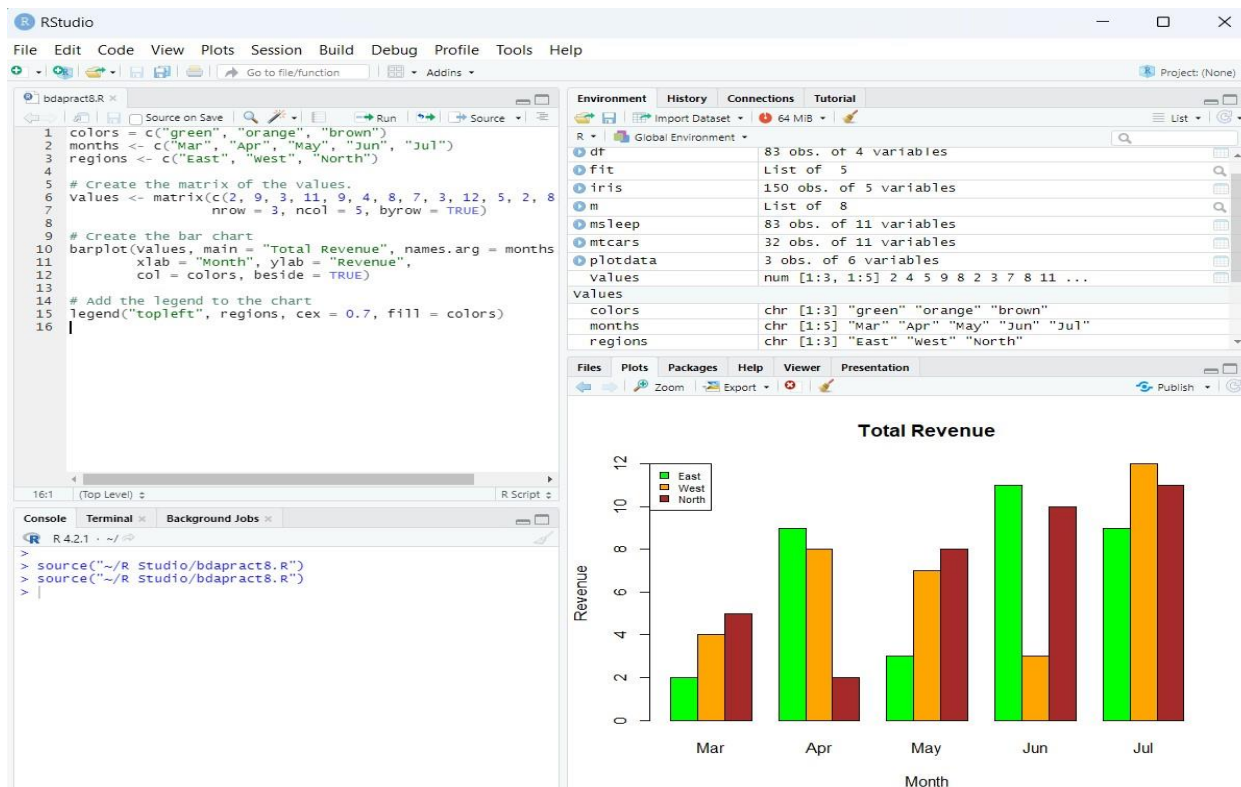
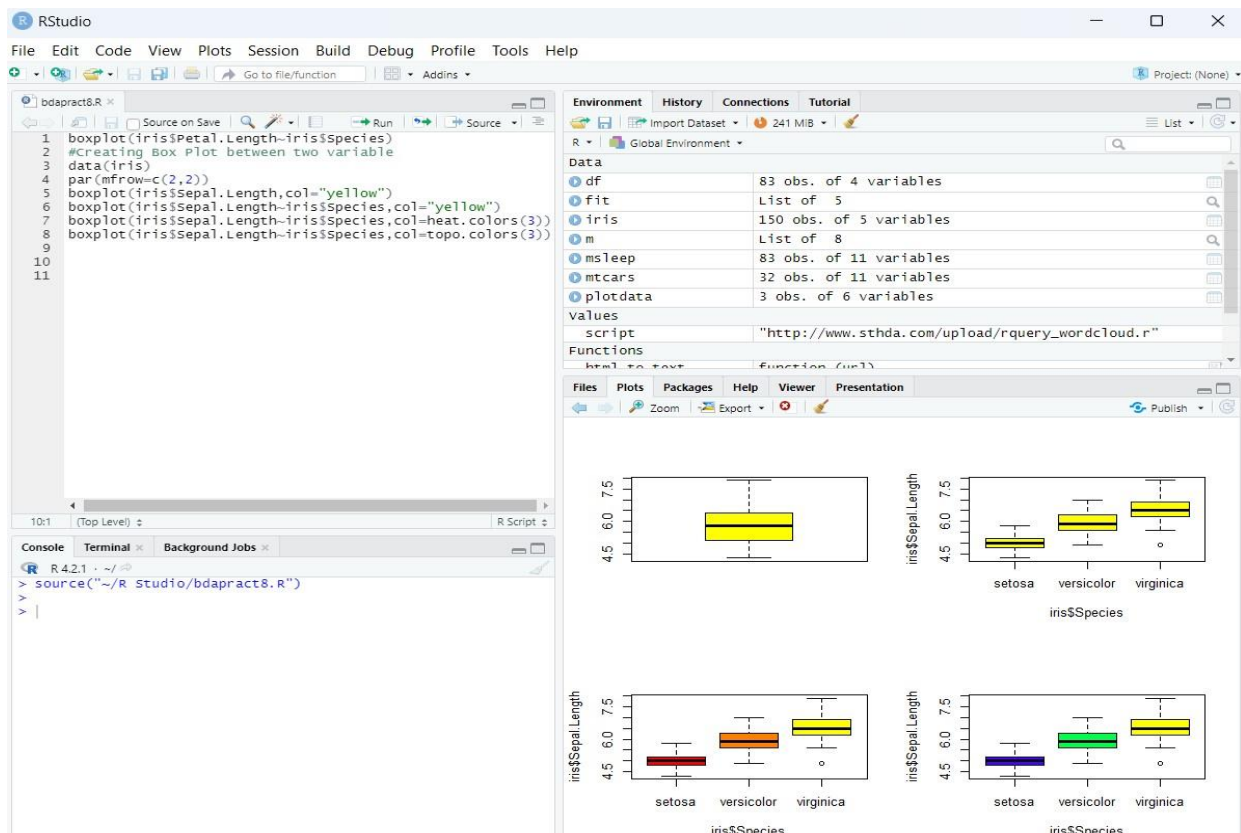
colors = c("green", "orange", "brown") months
<- c("Mar", "Apr", "May", "Jun", "Jul") regions
<- c("East", "West", "North") # Create the
matrix of the values.
Values <- matrix(c(2, 9, 3, 11, 9, 4, 8, 7, 3, 12, 5, 2, 8, 10, 11),
nrow = 3, ncol = 5, byrow = TRUE)
# Create the bar chart
barplot(Values, main = "Total Revenue", names.arg = months,
       xlab = "Month", ylab = "Revenue",
       col = colors, beside = TRUE) # Add the
legend to the chart
legend("topleft", regions, cex = 0.7, fill = colors)

```

Output:







Conclusion : Hence , we have successfully implemented experiment on Data Visualizations using R.