Big Data and Analytical Lab

Lab Assignment – 06

(BCSE0183)

Additional Problems in Packages for R Programming Language

Submitted by:- Vishal Dixit

Sec- B (62)

University Roll No :- 201500792

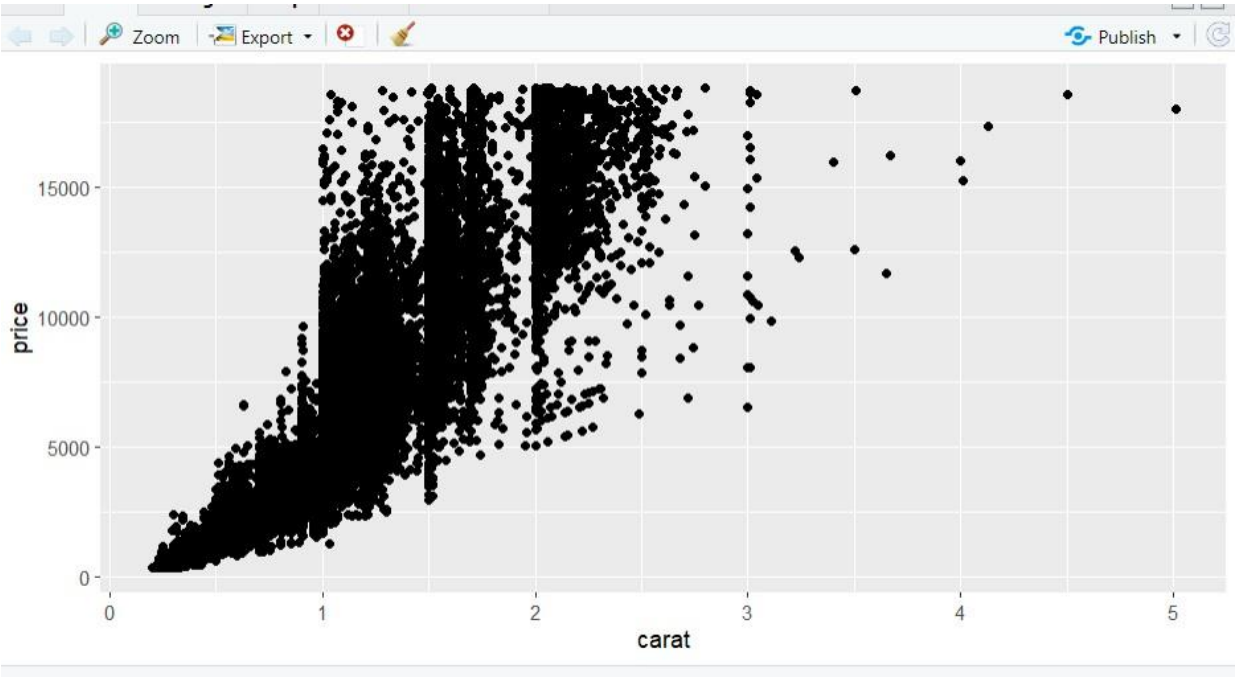
Submitted to: Dr. Robin Singh Bhadoria

1) Use ‘diamonds’ dataset and qplot ( ) of ggplot2 Package for following operations:

```
> # vishal
> library(ggplot2)
> View(diamonds)
```

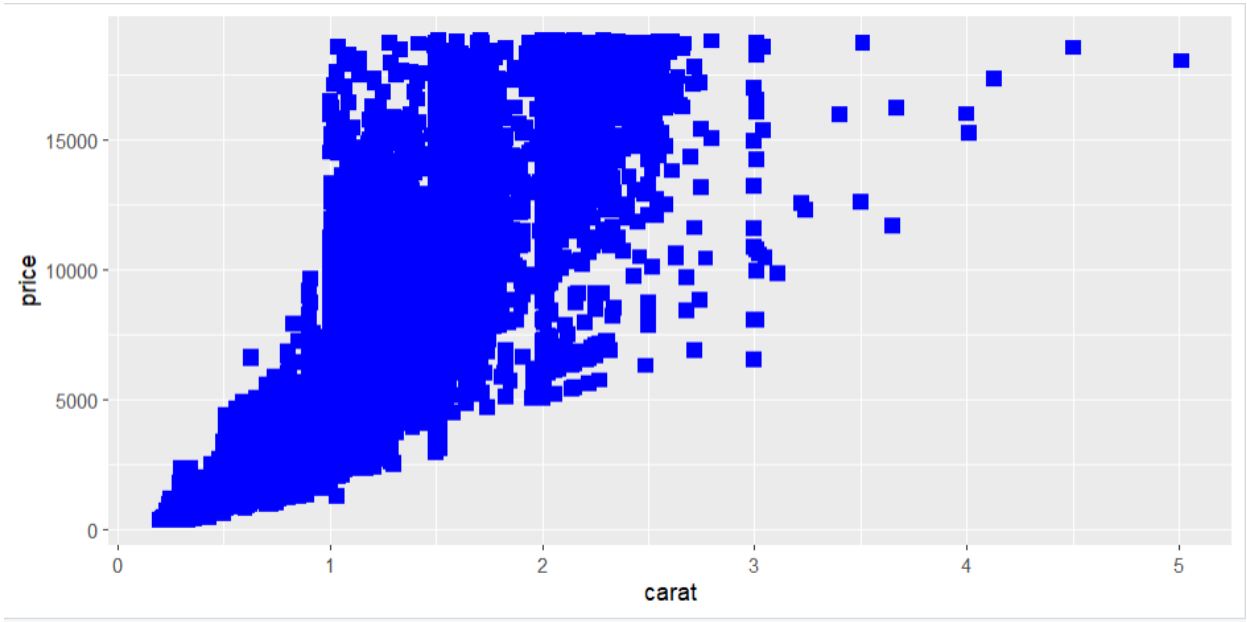
(a) Plot a relationship between parameters like price and carats (weight)

```
> qplot(x = carat, y = price, data = diamonds)
```



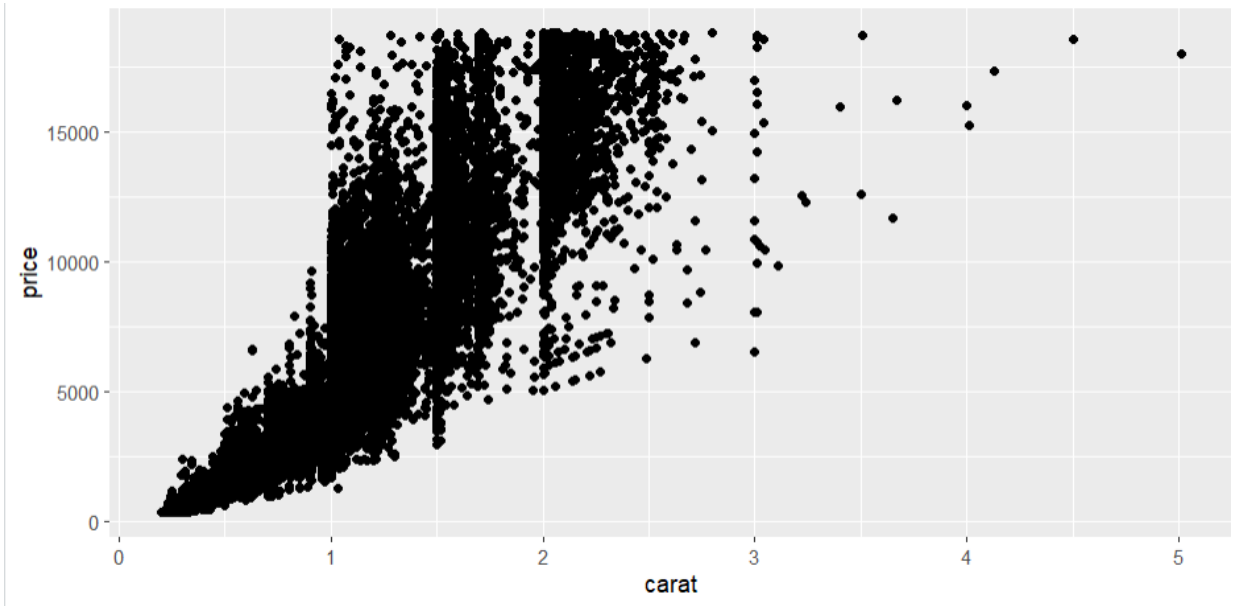
(b) Customize colour, size, & shape

```
> qplot(x = carat, y = price, data = diamonds, color = color, size = I(3),
shape = I(15))
```



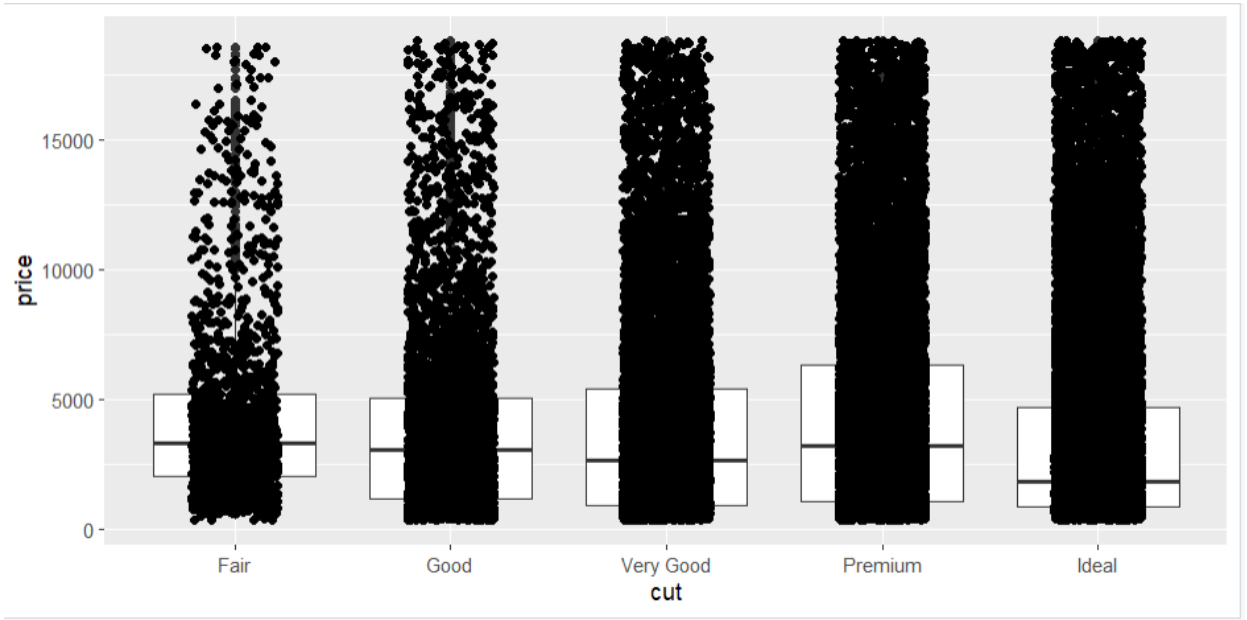
(c) Adding a smoother to a plot using geoms

```
> qplot(x = carat, y = price, data = diamonds) + geom_smooth(method = "loess")
```



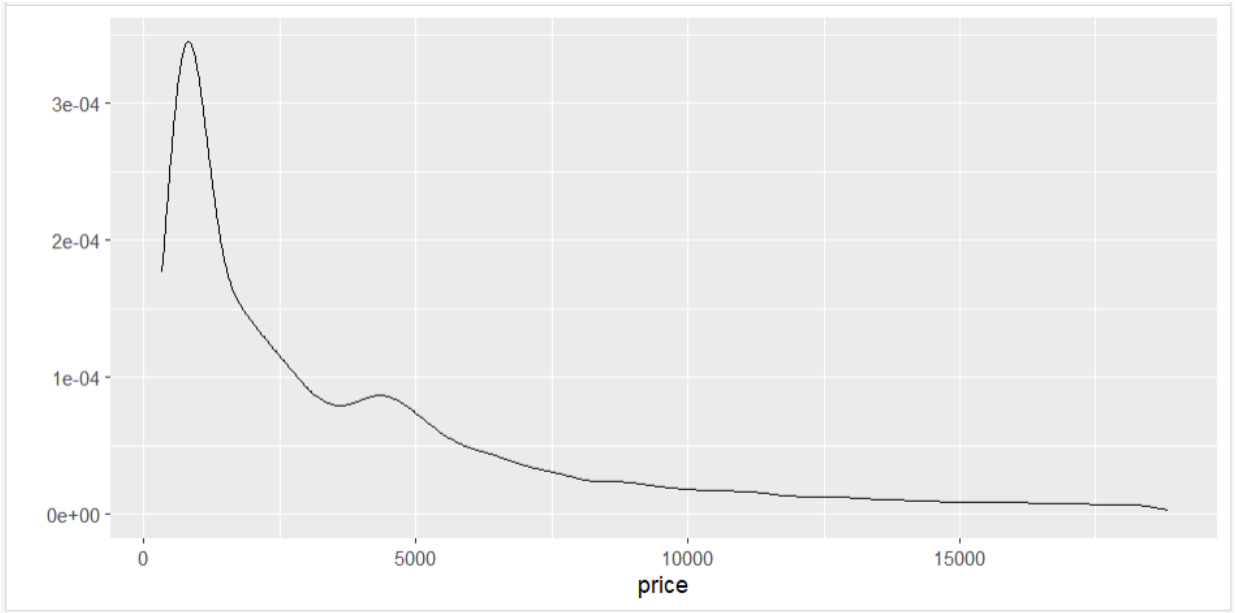
(d) Boxplots and jittered points using geoms

```
> qplot(x = cut, y = price, data = diamonds, geom = "boxplot") +
geom_jitter(width = 0.2)
```



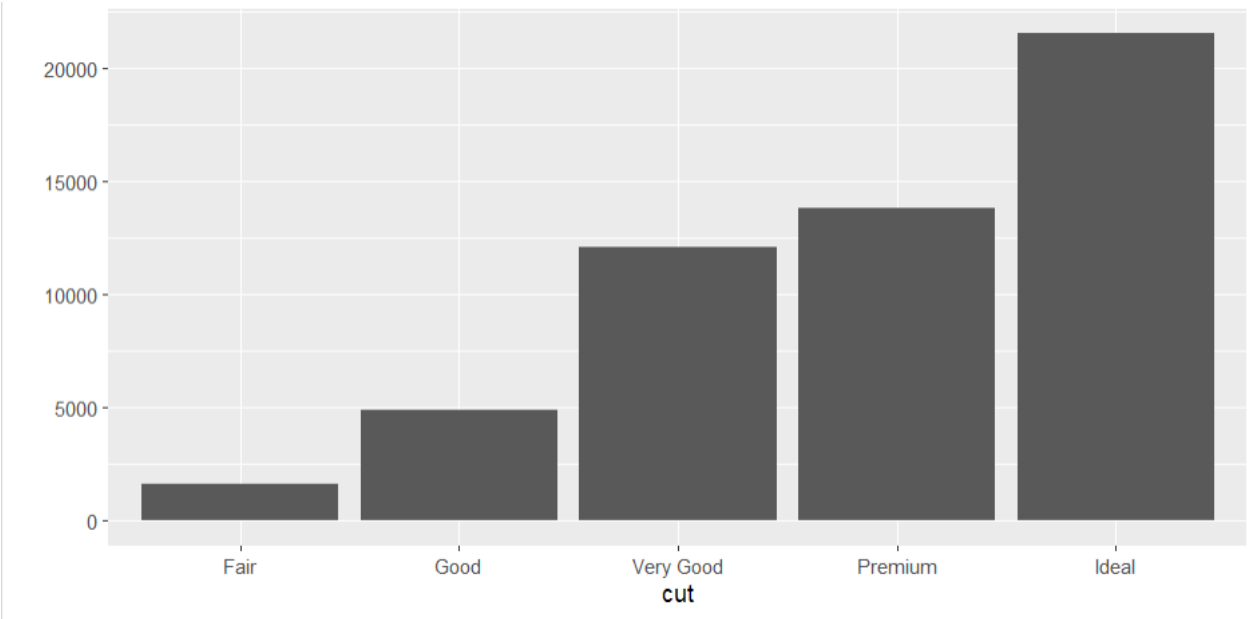
(e) Histogram and density plots using geoms

```
> qplot(x = price, data = diamonds, geom = "histogram", bins = 30)
> qplot(x = price, data = diamonds, geom = "density")
```



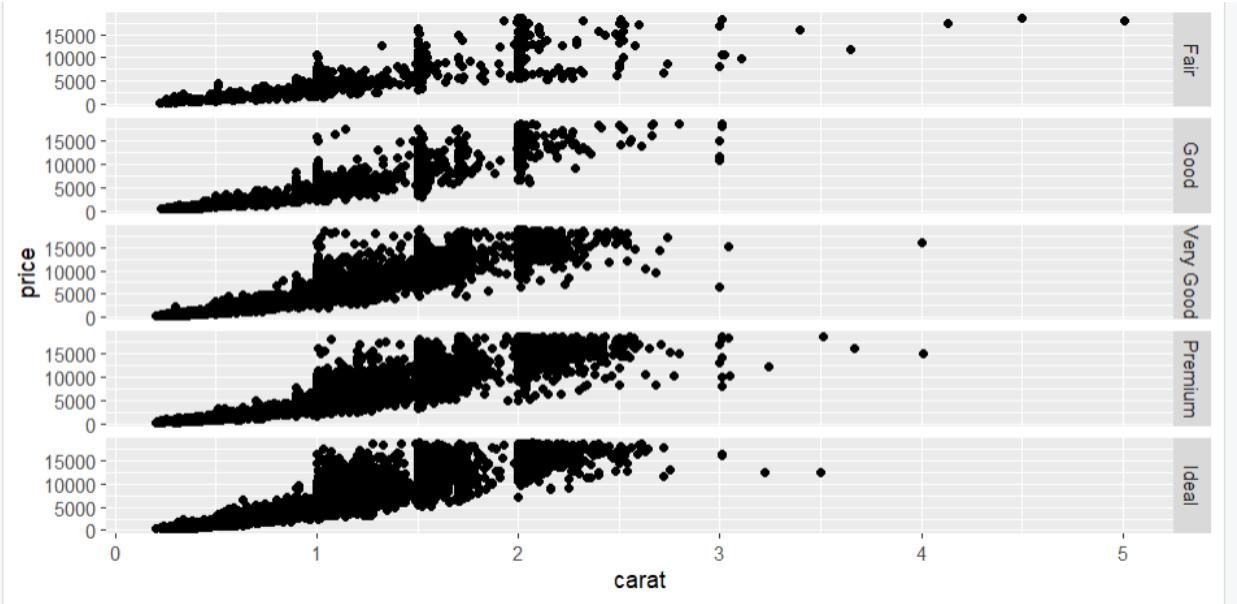
(f) Bar charts using geoms

```
> qplot(x = cut, data = diamonds, geom = "bar")
```



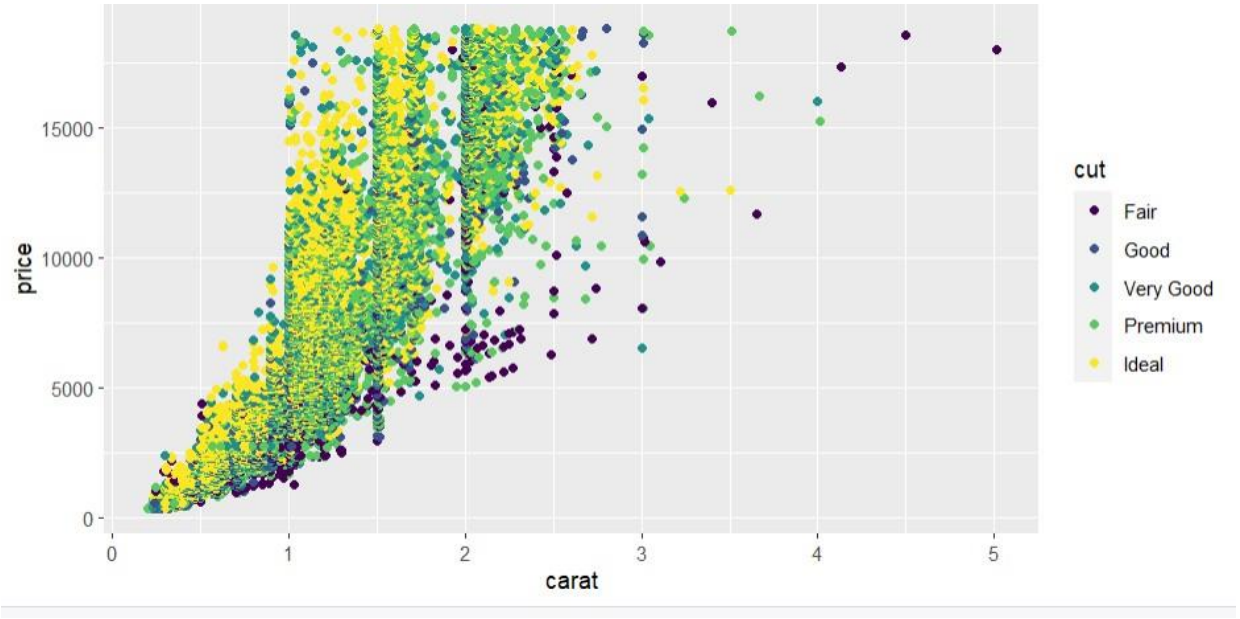
(g) faceting method in qplot()

```
> qplot(x = carat, y = price, data = diamonds, facets = cut ~ .)
```



(h) Building a scatterplot using factor ( ) function

```
> qplot(x = carat, y = price, data = diamonds, color = cut)
```



2) Use ‘mtcars’ dataset and ggplot ( ) of ggplot2 Package for following operations:

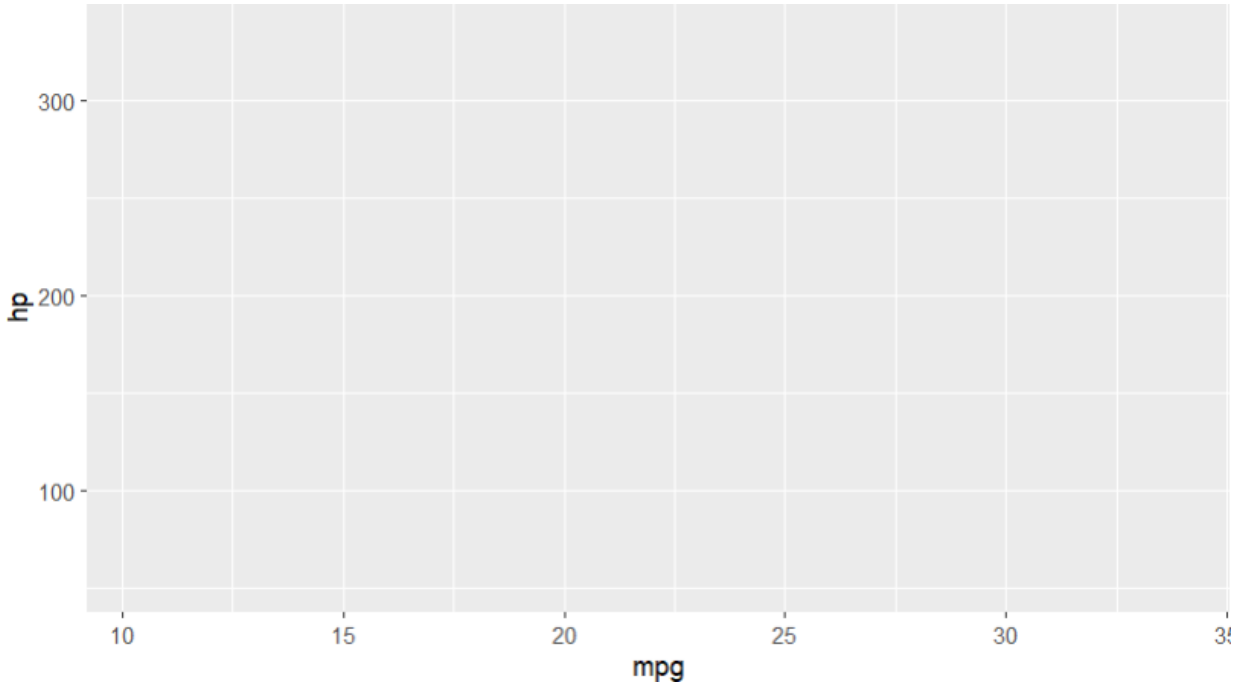
```
> # vishal
> Install.packages("dplyr")
> library(ggplot2)
> library(dplyr)
> View(mtcars)
```

(a) Data Layer

```
> ggplot(data = mtcars)
```

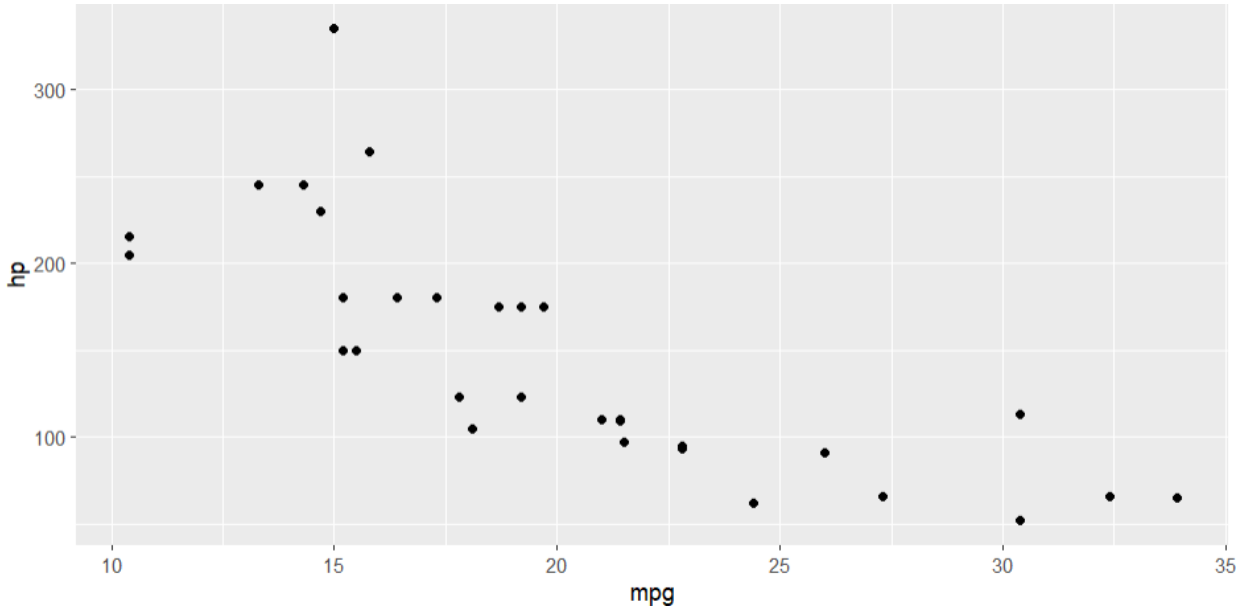
(b) Aesthetic Layer

```
> ggplot(data = mtcars, aes(x = mpg, y = hp))
```



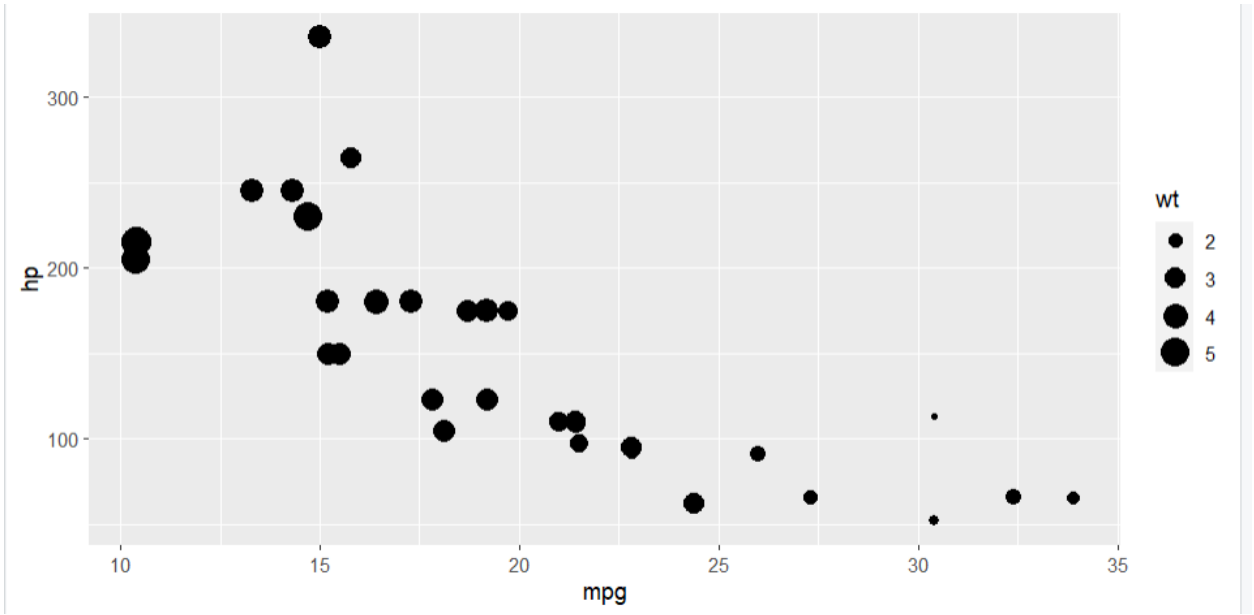
(c) Geometric layer

```
> ggplot(data = mtcars, aes(x = mpg, y = hp)) +
+     geom_point()
```



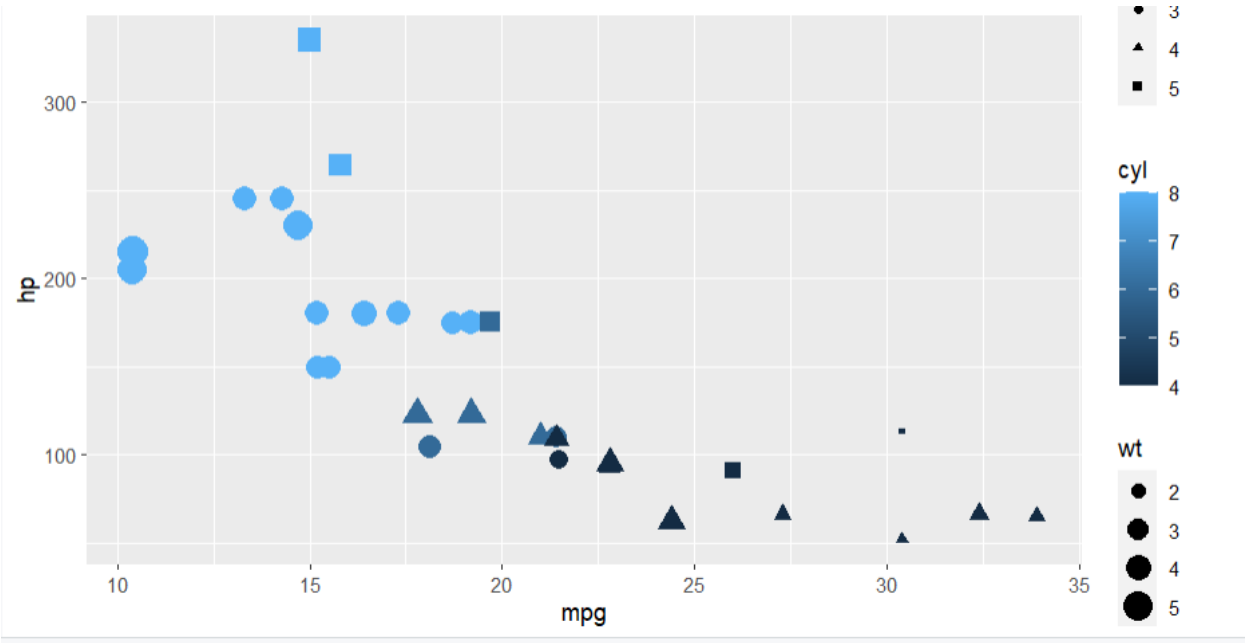
(d) Adding size

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



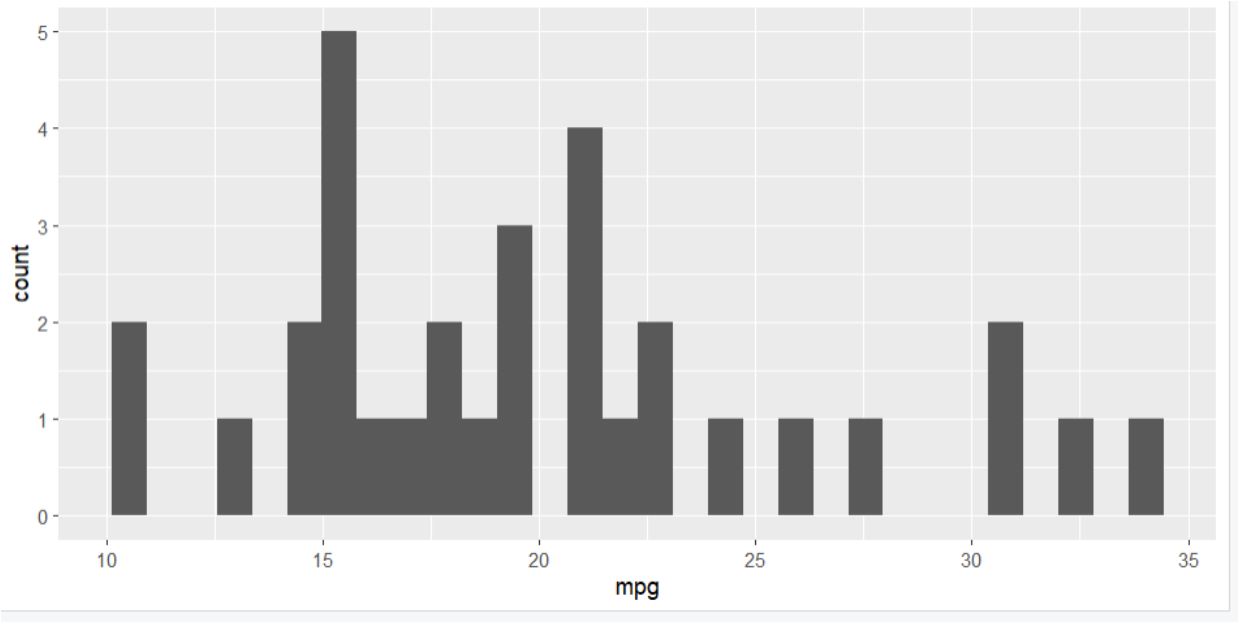
(e) Adding color and shape

```
> ggplot(data = mtcars, aes(x = mpg, y = hp, size = wt, color = cyl, shape = factor(gear))) +
+     geom_point()
```



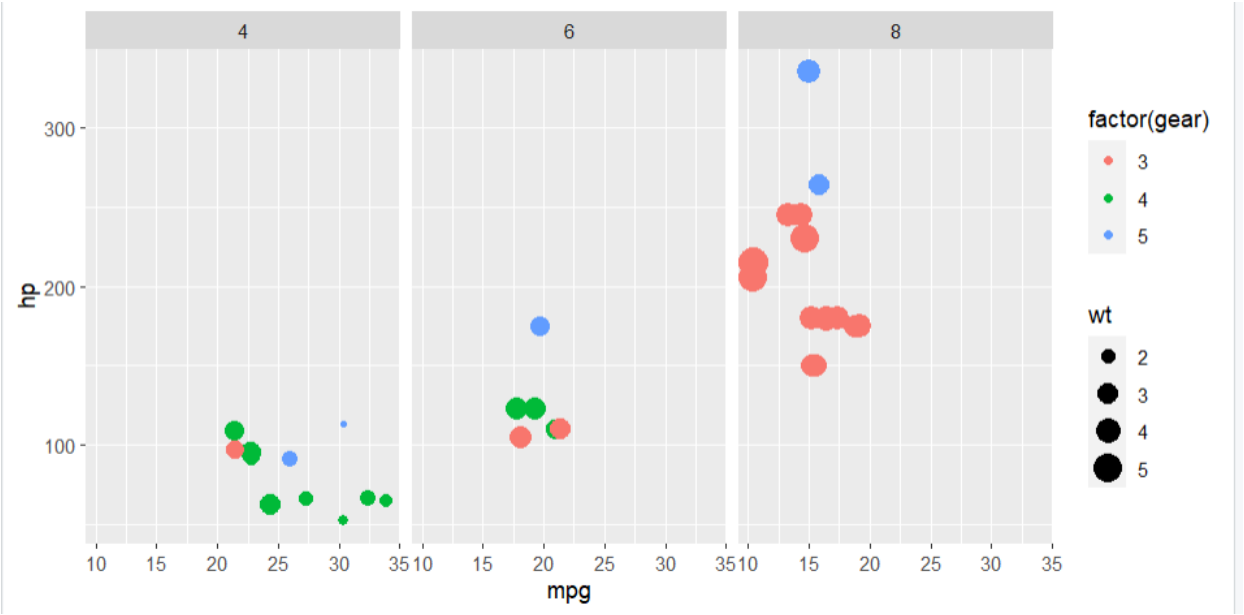
(f) Histogram plot

```
> ggplot(data = mtcars, aes(x = mpg)) +
+     geom_histogram()
```



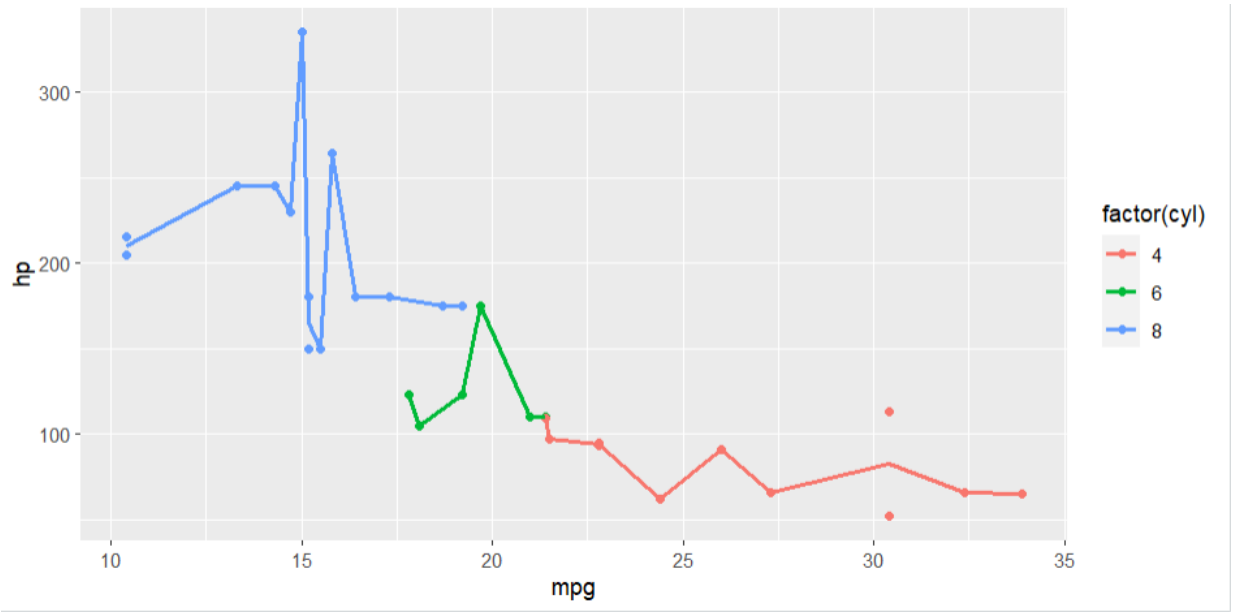
(g) Facet Layer

```
> ggplot(data = mtcars, aes(x = mpg, y = hp, size = wt, color = factor(gear)))
+     geom_point() +
+     facet_wrap(~cyl)
```



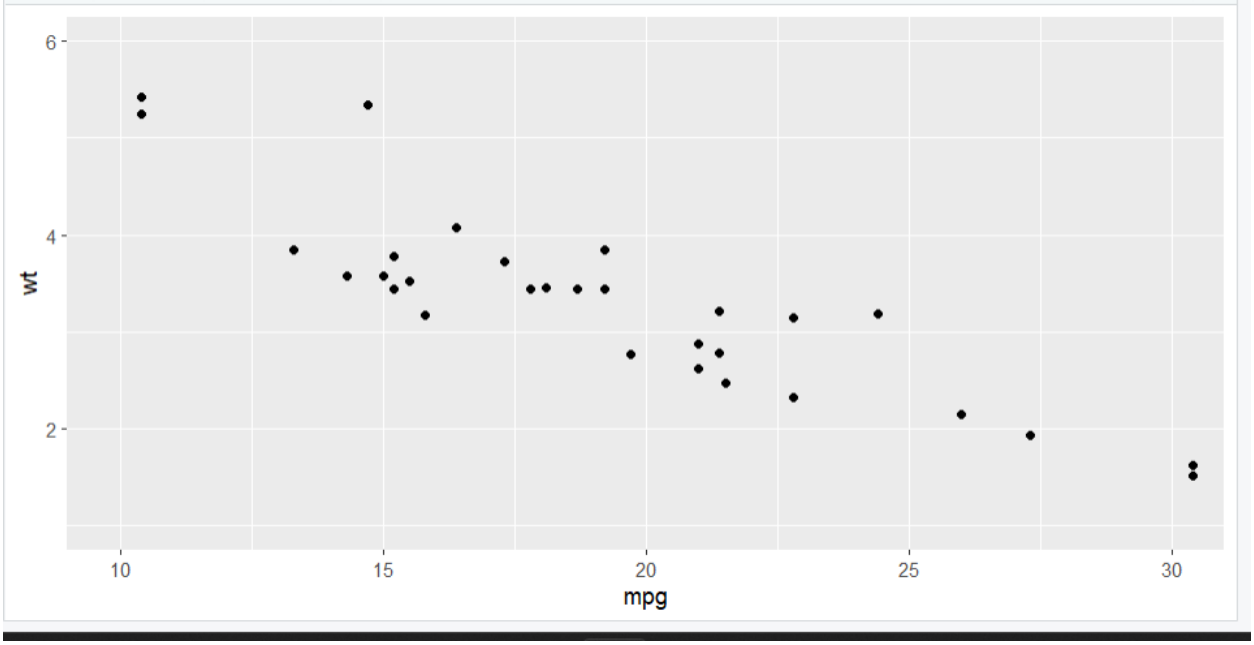
(h) Statistics layer

```
> ggplot(data = mtcars, aes(x = mpg, y = hp, color = factor(cyl))) +
+     geom_point() +
+     stat_summary(fun = mean, geom = "smooth")
```



(i) Coordinates layer: Control plot dimensions

```
> ggplot(mtcars, aes(x = mpg, y = wt)) +
+     geom_point() +
+     coord_cartesian(xlim = c(10, 30), ylim = c(1, 6))
```



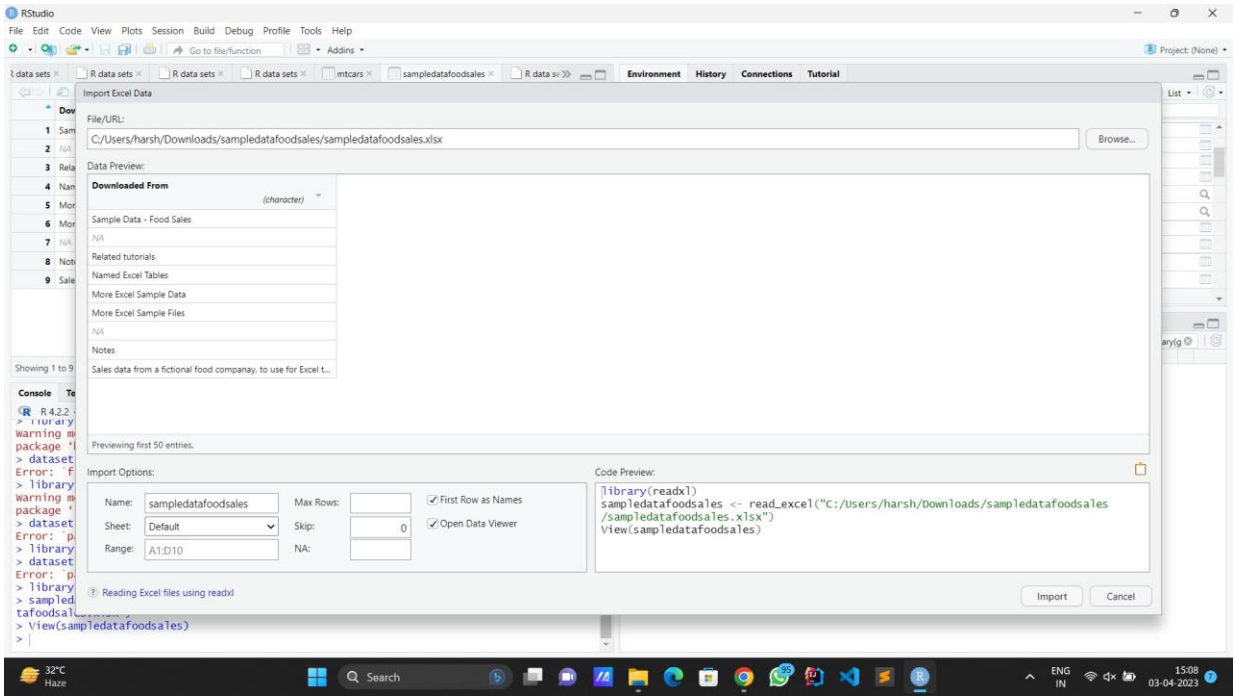
3) Display Import and load dataset from external sources like:

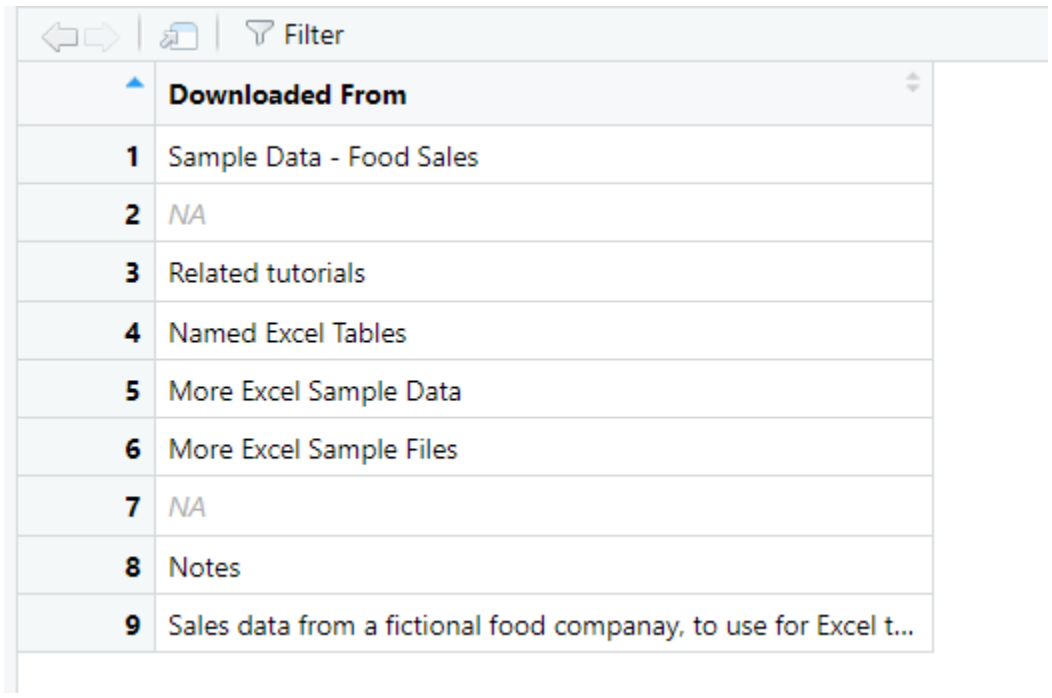
(a) Excel File

(b) Statistical Analysis System (SAS) File

```
mydata <- read_sas("path/to/mydata.sas7bdat")
```

(c) Text File (base)





	Downloaded From
1	Sample Data - Food Sales
2	NA
3	Related tutorials
4	Named Excel Tables
5	More Excel Sample Data
6	More Excel Sample Files
7	NA
8	Notes
9	Sales data from a fictional food companay, to use for Excel t...

4) Read file using below function commands:
(a) read.delim( ) and read.delim2( )
(b) read.table( )
(c) read.csv( )
(d) read.xlsx( ) using packages xlsx

```
my_file1<- read. delim (file = "C:/Users/Glau/Desktop/cancer_csv.csv")
my_file2<- read. delim2 (file = "C:/Users/Glau/Desktop/cancer_csv.csv")

df <- read.table("adi.txt", header = TRUE)

dfcsv<- read.csv("C:/Users/Glas/Desktop/cancer_csv.csv", header=TRUE) dfcsv
```

5) Read XML file in R using the function xmlParse().

```
> xml_data <- xmlParse("C:\Users\vishaldixit\Downloads\myXMLFile0 (1).xml")
```

6) The tidyverse package is designed to make it easy to install and load core packages.
Make following operations:

(a) Install tidyverse

```
> #vishal
> install.packages("tidyverse")
```

(b) Pivoting

```
> # Create a sample data frame
> df <- data.frame(id = 1:3,
+                  year_2000 = c(10, 20, 30),
+                  year_2001 = c(15, 25, 35),
+                  year_2002 = c(18, 28, 38))
>
> # Pivot the data frame longer
> df_long <- df %>%
+     pivot_longer(cols = starts_with("year_"),
+                  names_to = "year",
+                  values_to = "value")
>
> # View the result
> df_long
# A tibble: 9 × 3
     id year      value
  <int> <chr>     <dbl>
1     1 year_2000    10
2     1 year_2001    15
3     1 year_2002    18
4     2 year_2000    20
5     2 year_2001    25
6     2 year_2002    28
7     3 year_2000    30
8     3 year_2001    35
9     3 year_2002    38
>
>
```

(c) Rectangling

```
> # Create a sample data frame
> df <- data.frame(id = 1:3,
+                  year_2000 = c(10, 20, 30),
+                  year_2001 = c(15, 25, 35),
+                  year_2002 = c(18, 28, 38))
>
> # Rectangle the data frame
> df_rect <- df %>%
+     gather(key = "year", value = "value", starts_with("year_"))
>
> # View the result
> df_rect
  id      year value
1  1 year_2000    10
2  2 year_2000    20
3  3 year_2000    30
4  1 year_2001    15
5  2 year_2001    25
6  3 year_2001    35
7  1 year_2002    18
8  2 year_2002    28
9  3 year_2002    38
>
>
```

(d) Nesting

```
> # Create a sample data frame
> df <- data.frame(id = 1:3,
+                  x = c(10, 20, 30),
+                  y = c(15, 25, 35),
+                  z = c(18, 28, 38))
>
> # Nest the data frame
> df_nested <- df %>%
+     nest(data = c(x, y, z))
>
> # View the result
> df_nested
# A tibble: 3 x 2
     id data
  <int> <list>
1     1 <tibble [1 x 3]>
2     2 <tibble [1 x 3]>
3     3 <tibble [1 x 3]>
>
>
```

(e) Splitting

```
#vishal
library(tidyverse)

# create a sample data frame with nested data
df_nested <- tibble(
    group = rep(c("A", "B"), each = 3),
    data = list(tibble(x = 1:3, y = 4:6), tibble(x = 7:9, y = 10:12))
```

7) Write a program in Java for counting the number of words.

```
7  ********************************************************************************/
8  import java.util.*;
9  public class Main
10 {
11     public static void main(String[] args) {
12         Scanner sc = new Scanner(System.in);
13         String str="Hello my name is Vishal Dixit";
14         int count=1;
15         for(int i=0;i<str.length();i++){
16             if((str.charAt(i)==' ') && (str.charAt(i)==' ')){
17                 count++;
18             }
19         }
20 
21         System.out.println("Number of words in a String is : " +count);
22     }
23 }
24 
```

input

```
Number of words in a String is : 6


...Program finished with exit code 0
```