



ACADGILD

SESSION 7: Basic Statistics

Assignment 2

Table of Contents

1.Introduction
3

2.Objective
3

3.Prerequisites
3

4.Associated Data Files
3

5.Problem Statement
3

6.Expected Output
3

7.Approximate Time to Complete Task
3

1. Introduction

This assignment will help you understand the concepts learnt in the session.

2. Objective

This assignment will test your skills on basic statistics.

3. Prerequisites

Not applicable.

4. Associated Data Files

Not applicable.

5. Problem Statement

1. Write a program to create **barplots** for all the categorical columns in **mtcars**.

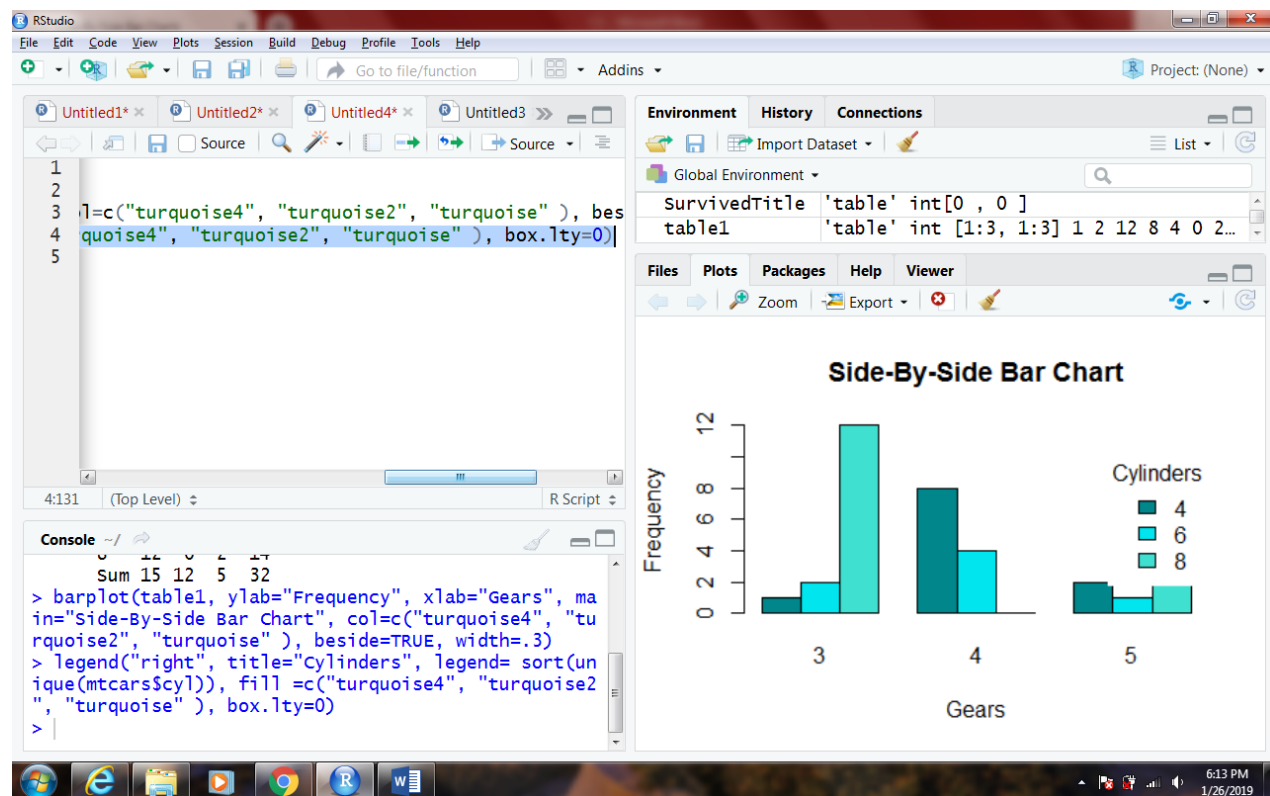
Ans –

```
table1 <- table(mtcars$cyl, mtcars$gear, dnn=c("Cylinders", "Gears"))  
addmargins(table1)
```

```
barplot(table1, ylab="Frequency", xlab="Gears", main="Side-By-Side Bar  
Chart", col=c("turquoise4", "turquoise2", "turquoise" ), beside=TRUE,  
width=.3)
```

```
legend("right", title="Cylinders", legend= sort(unique(mtcars$cyl)), fill
=c("turquoise4", "turquoise2", "turquoise" ), box.lty=0)
```

The two categorical variables, cylinders and gears are used to show how to create a bar chart.

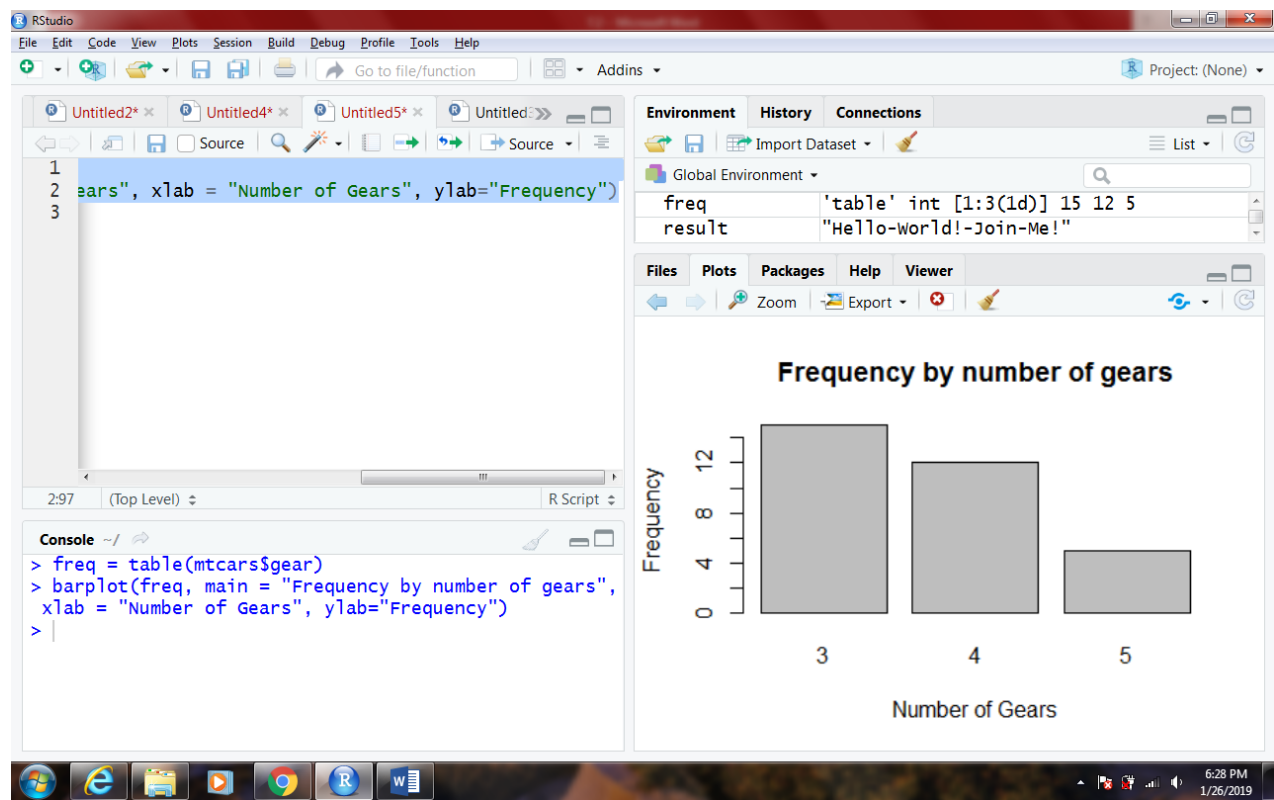


Bar graph showing no of each gear type in mtcars

```
freq = table(mtcars$gear)
```

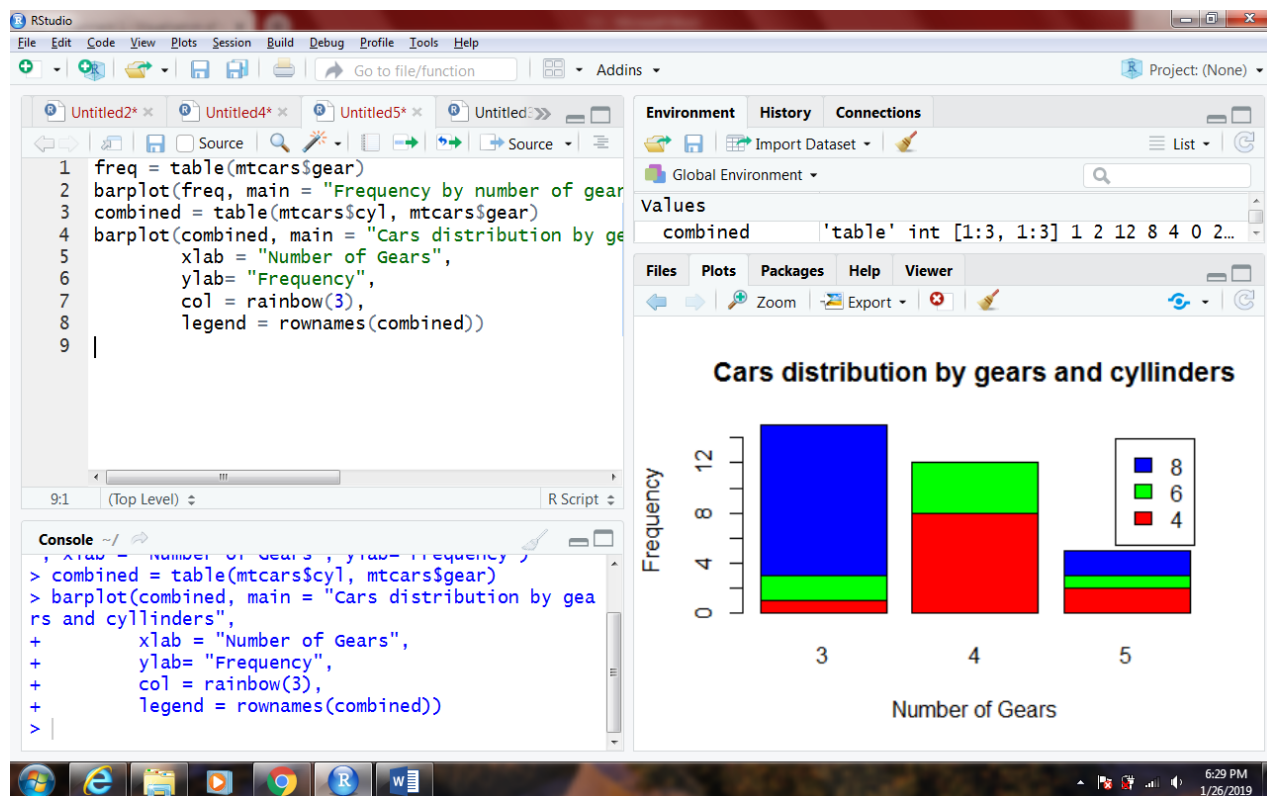
```
barplot(freq, main = "Frequency by number of gears", xlab = "Number of Gears",
ylab="Frequency")
```

Data Analytics



Data Analytics

Stacked bar graph of the number of each gear type and how they are further divided out by cyl

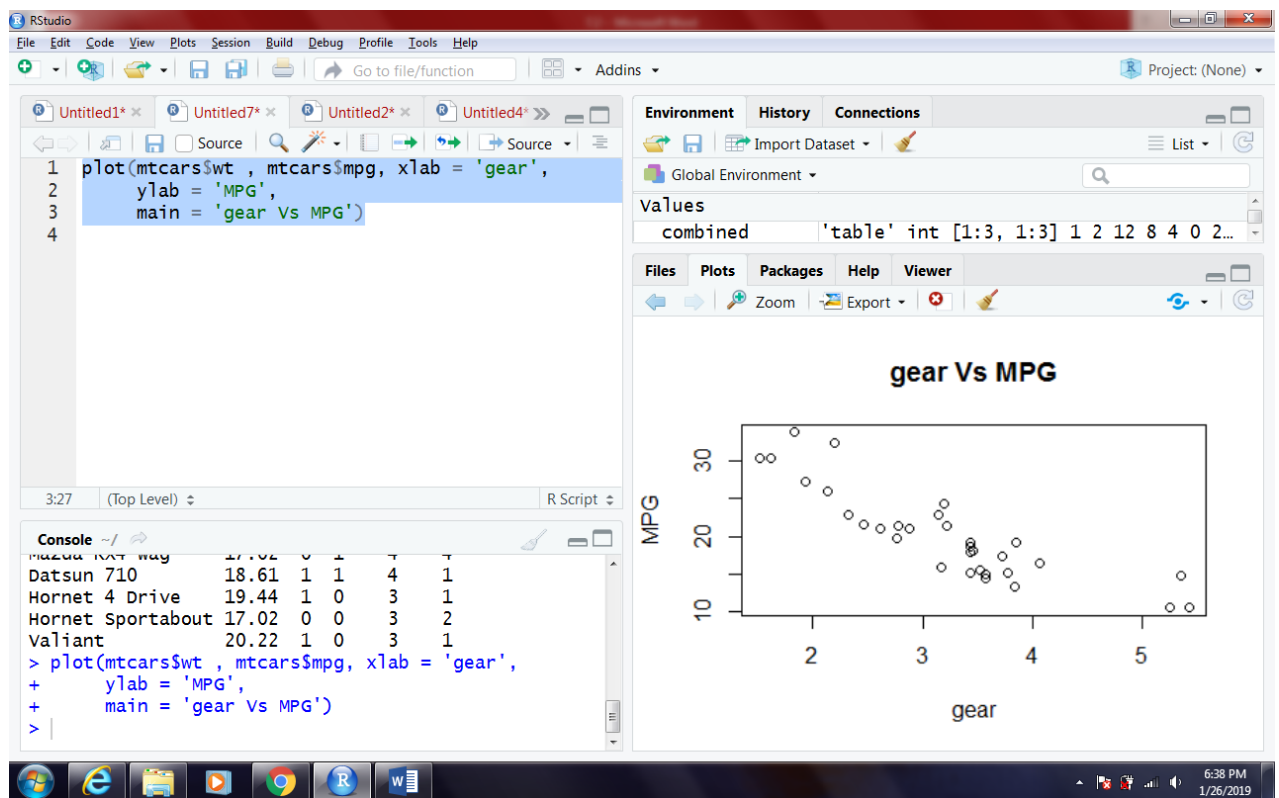


```
combined = table(mtcars$cyl, mtcars$gear)
barplot(combined, main = "Cars distribution by gears and cyllinders",
        xlab = "Number of Gears",
        ylab = "Frequency",
        col = rainbow(3),
        legend = rownames(combined))
```

2. Create a **scatterplot** matrix by gear types in **mtcars** dataset.

```
plot(mtcars$wt, mtcars$mpg, xlab = 'gear',
     ylab = 'MPG',
     main = 'gear Vs MPG')
```

Data Analytics



```
library(car)
```

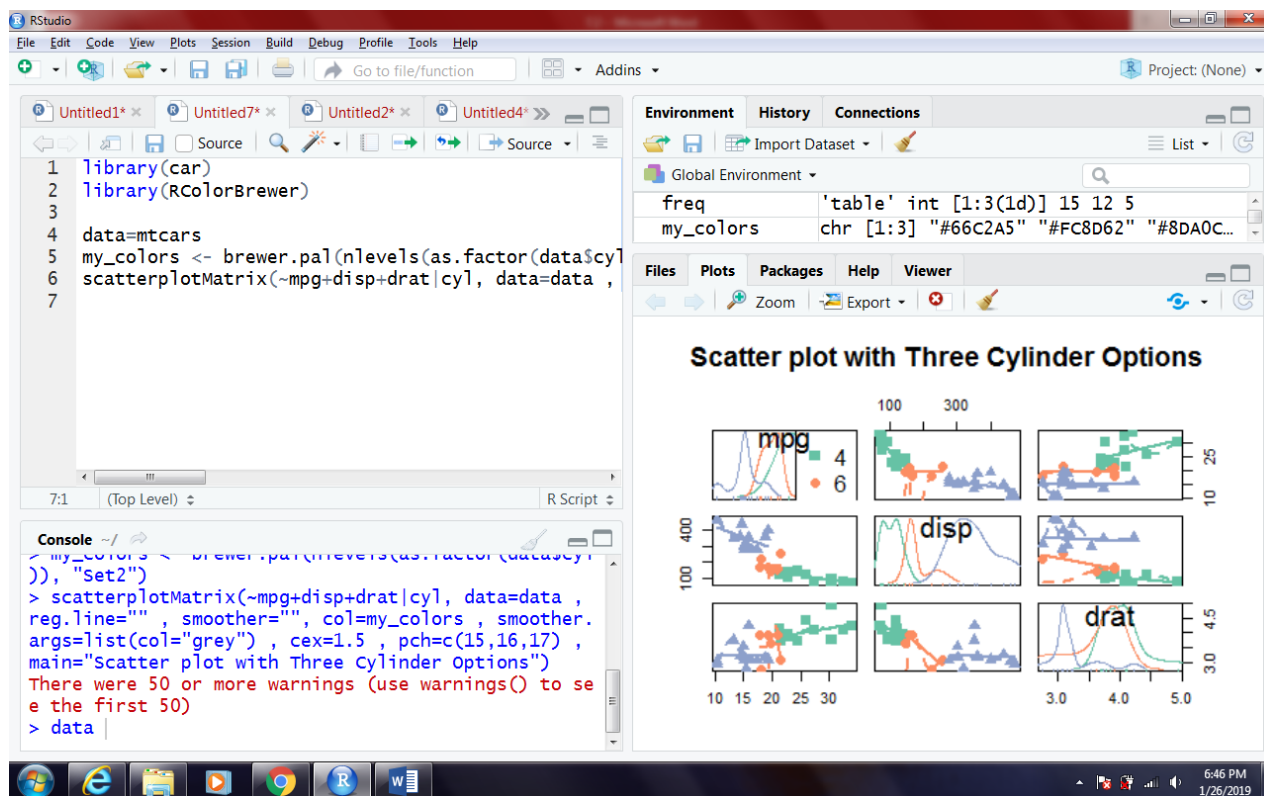
```
library(RColorBrewer)
```

```
data=mtcars
```

```
my_colors <- brewer.pal(nlevels(as.factor(data$cyl)), "Set2")
```

```
scatterplotMatrix(~mpg+disp+drat|cyl, data=data, reg.line="",  
smoother="", col=my_colors, smoother.args=list(col="grey"), cex=1.5,  
pch=c(15,16,17), main="Scatter plot with Three Cylinder Options")
```

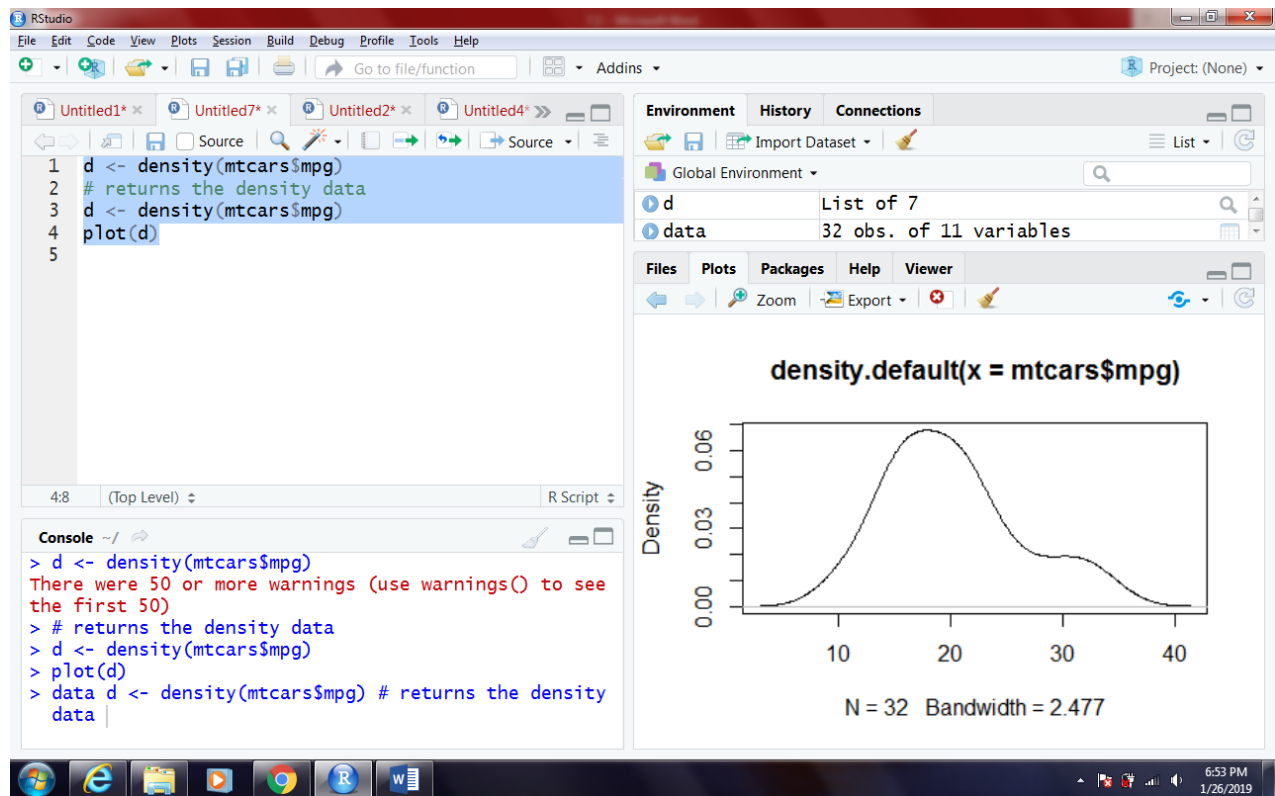
Data Analytics



3. Write a program to create a **plot density** by class variable.

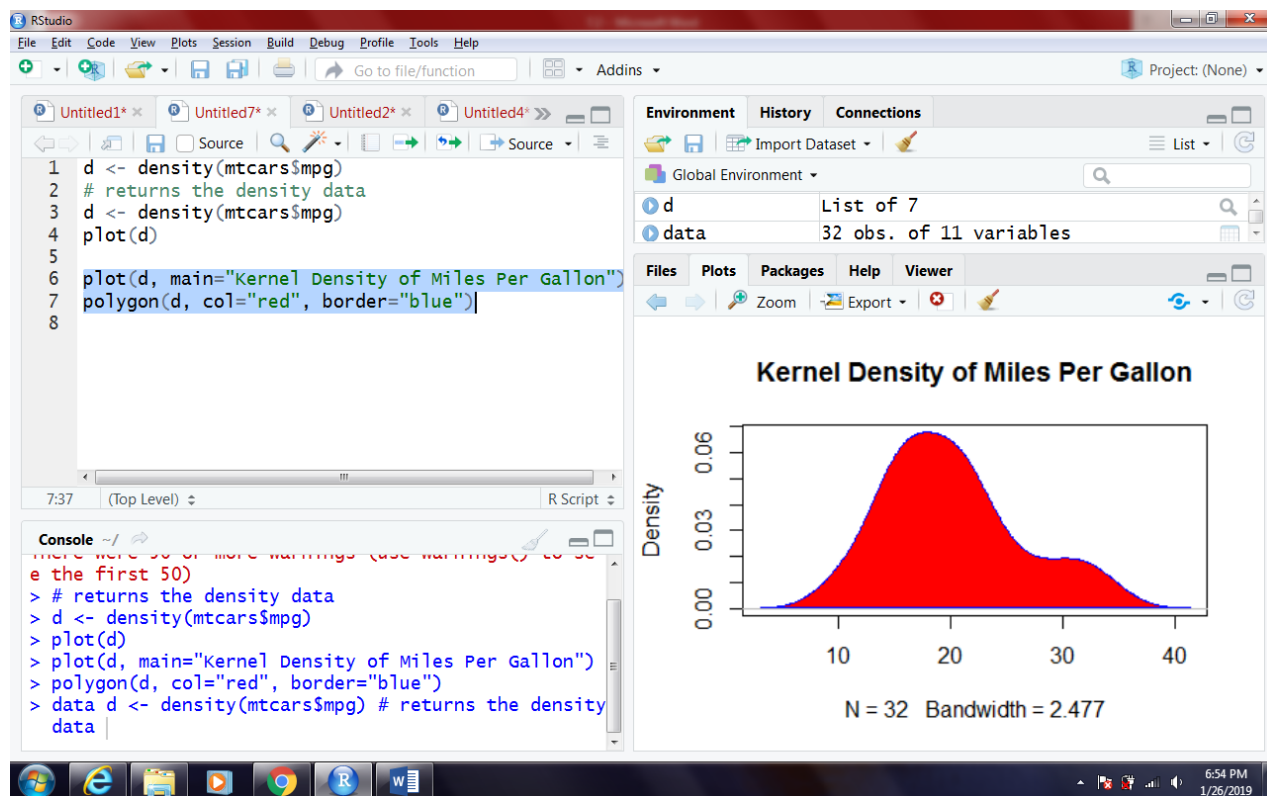
```
d <- density(mtcars$mpg)
# returns the density data
d <- density(mtcars$mpg)
plot(d)
```


Data Analytics



```
plot(d, main="Kernel Density of Miles Per Gallon")
polygon(d, col="red", border="blue")
```

Data Analytics



6. Expected Format

1. R file should be submitted where applicable.
2. R file should be in PDF or in .r format
3. Proper screenshots of the outputs should be submitted as well
4. The r codes, if submitted in any other format, will be subjected to deduction in marks

Note: Your solution will not be entertained if it is any other format, e.g., .zip, .doc, .rtf etc.

7. Approximate Time to Complete Task

30 mins.