Day 3 Lab Manual Part 2

BIVARIATEANALYSIS IN R -COVARIANCE, CORRELATION, CROSSTAB

Exercise: 8

```
   Reference    Status Gender  TestNewOrFollowUp
1        KRXH Accepted Female Test1             New
2        KRPT Accepted   Male Test1             New
3        FHRA Rejected   Male Test2             New
4        CZKK Accepted Female Test3             New
5        CQTN Rejected Female Test1             New
6        PZXW Accepted Female Test4       Follow-up
7        SZRZ Rejected   Male Test4             New
8        RMZE Rejected Female Test2             New
9        STNX Accepted Female Test3             New
10       TMDW Accepted Female Test1             New
```
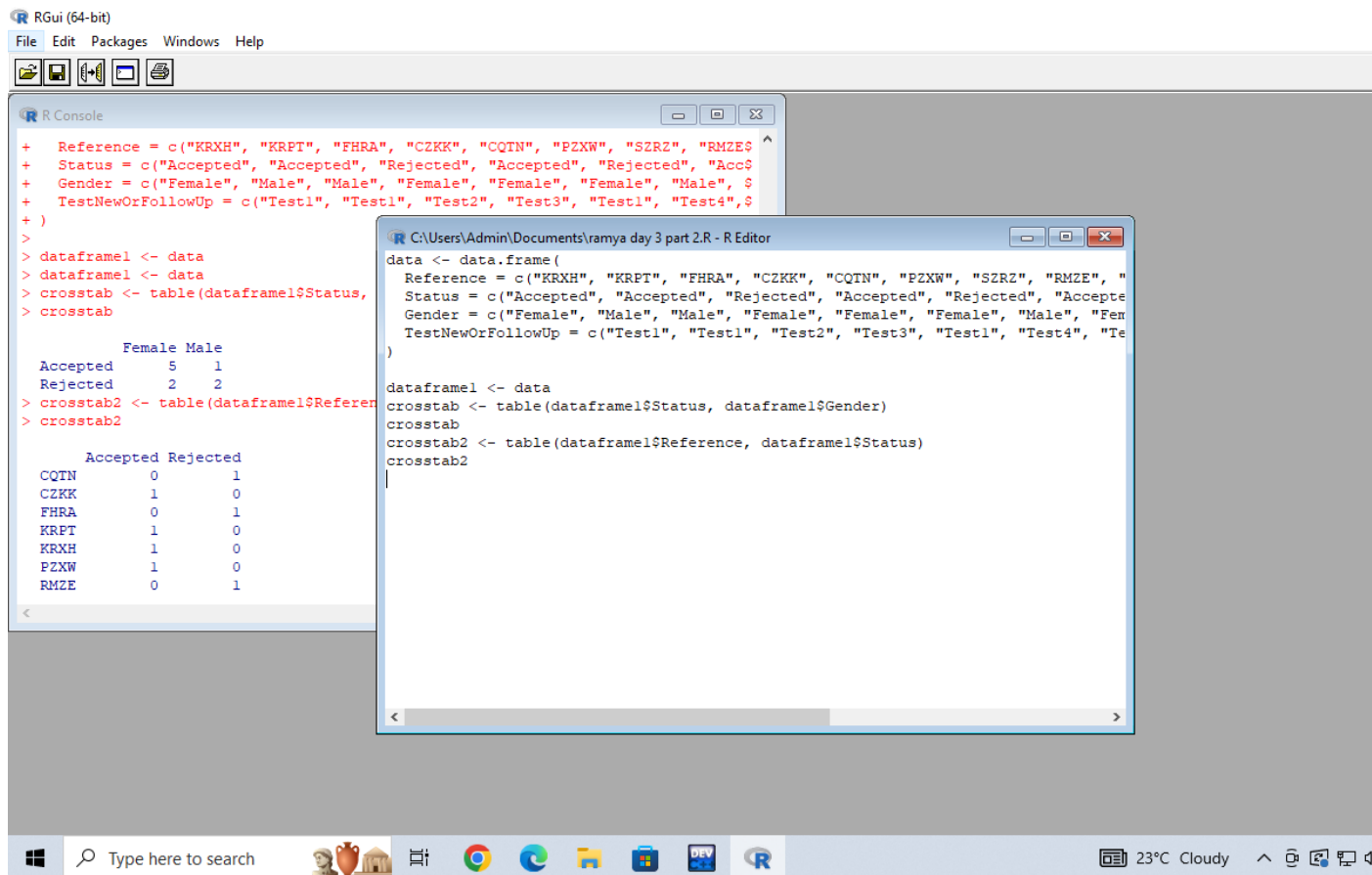
    i)   Load the dataset and Create a data frame and name it as dataframe1

    ii)   Load the function for crosstab

Note: Perform status+gender

```
          Gender
Status     Female Male
  Accepted      5    1
  Rejected      2    2
```

Note: Reference+Status

```
          Status
Reference Accepted Rejected
     CQTN        0        1
     CZKK        1        0
     FHRA        0        1
     KRPT        1        0
     KRXH        1        0
     PZXW        1        0
     RMZE        0        1
     STNX        1        0
     SZRZ        0        1
     TMDW        1        0
```

program:

```
data <- data.frame(
  Reference = c("KRXH", "KRPT", "FHRA", "CZKK", "CQTN", "PZXW", "SZRZ", "RMZE", "STNX", "TMDW"),
  Status = c("Accepted", "Accepted", "Rejected", "Accepted", "Rejected", "Accepted", "Rejected", "Rejected", "Accepted", "Accep
  Gender = c("Female", "Male", "Male", "Female", "Female", "Female", "Male", "Female", "Female", "Female"),
  TestNewOrFollowUp = c("Test1", "Test1", "Test2", "Test3", "Test1", "Test4", "Test4", "Test2", "Test3", "Test1")
)

dataframe1 <- data
crosstab <- table(dataframe1$Status, dataframe1$Gender)
crosstab
crosstab2 <- table(dataframe1$Reference, dataframe1$Status)
crosstab2
```

Exercise: 9

    i)   Use Two Categorical Variables and Discover the relationships within a dataset

program:

```r
cross_table <- table(dataframe1$Reference, dataframe1$Status)

chi_square <- chisq.test(dataframe1$Reference, dataframe1$Status)

library(ggplot2)

ggplot(dataframe1, aes(x = Reference, fill = Status)) +

  geom_bar(position = "fill") +

  labs(title = "Relationship between Reference and Status")
```
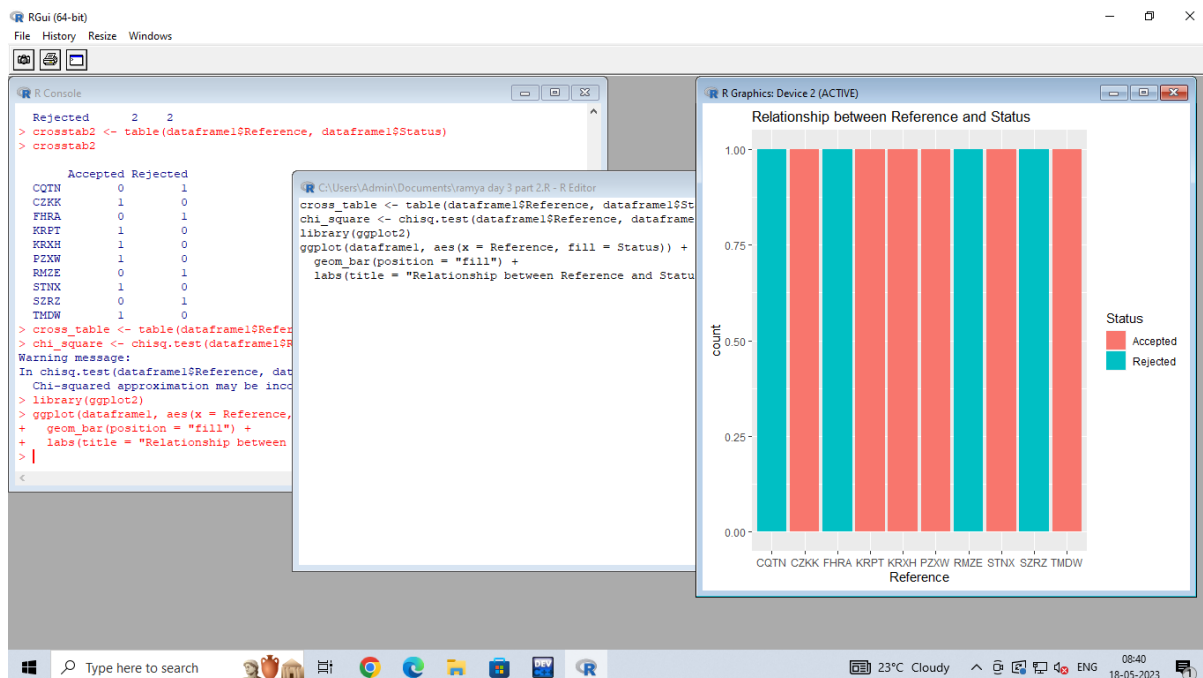

ii) Next, using the xtabs() function, apply two variables from "dataframe1 ", to create a table delineating the relationship between the "Reference" category, and the "Status" category.

program:

```r
dataframe1 <- data.frame(

  Reference = c("KRXH", "KRPT", "FHRA", "CZKK", "CQTN", "PZXW", "SZRZ", "RMZE", "STNX", "TMDW"),

  Status = c("Accepted", "Accepted", "Rejected", "Accepted", "Rejected", "Accepted", "Rejected", "Rejected", "Accepted", "Accepted"),

  Gender = c("Female", "Male", "Male", "Female", "Female", "Female", "Male", "Female", "Female", "Female"),

  TestNewOrFollowUp = c("Test1", "Test1", "Test2", "Test3", "Test1", "Test4", "Test4", "Test2", "Test3", "Test1")

)
```

dataframe2 <- xtabs(~ Reference + Status, data = dataframe1)

dataframe2

iii)   Save the file in the name of dataframe2

Exercise: 10

Use the same data frame using three Categorical Variables create a Multi-Dimensional Table

Apply three variables from "dataframe1" to create a Multi-Dimensional Cross-Tabulation of "Status", "Gender", and "Test".
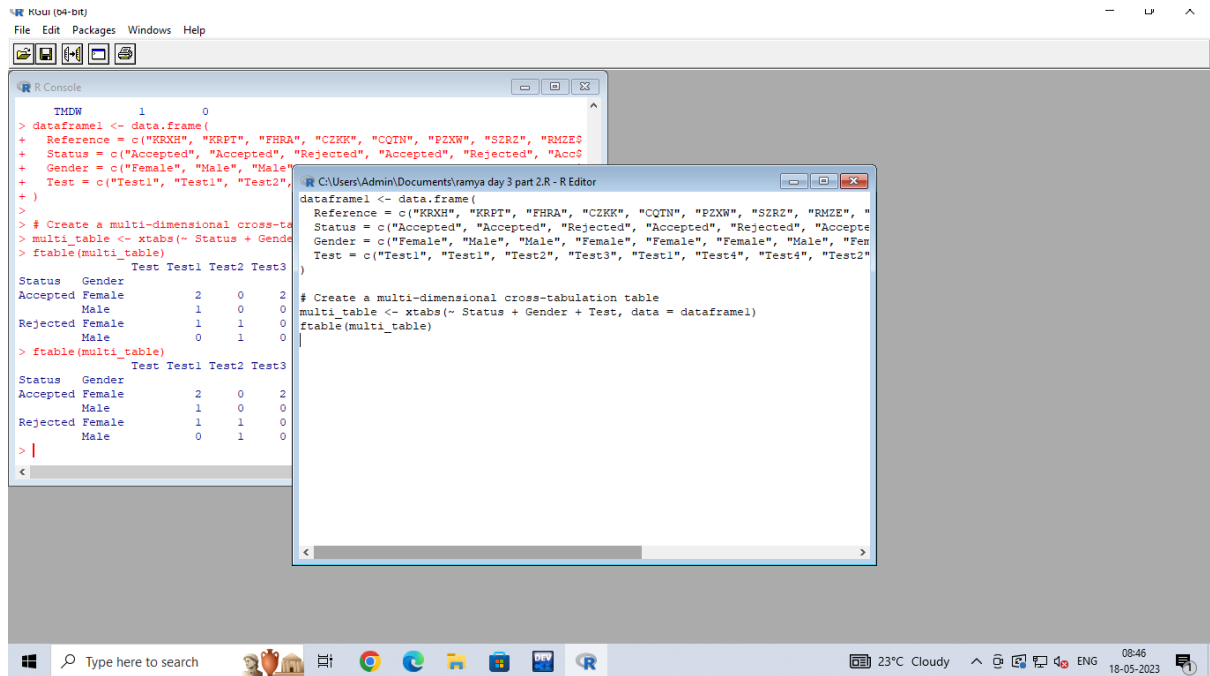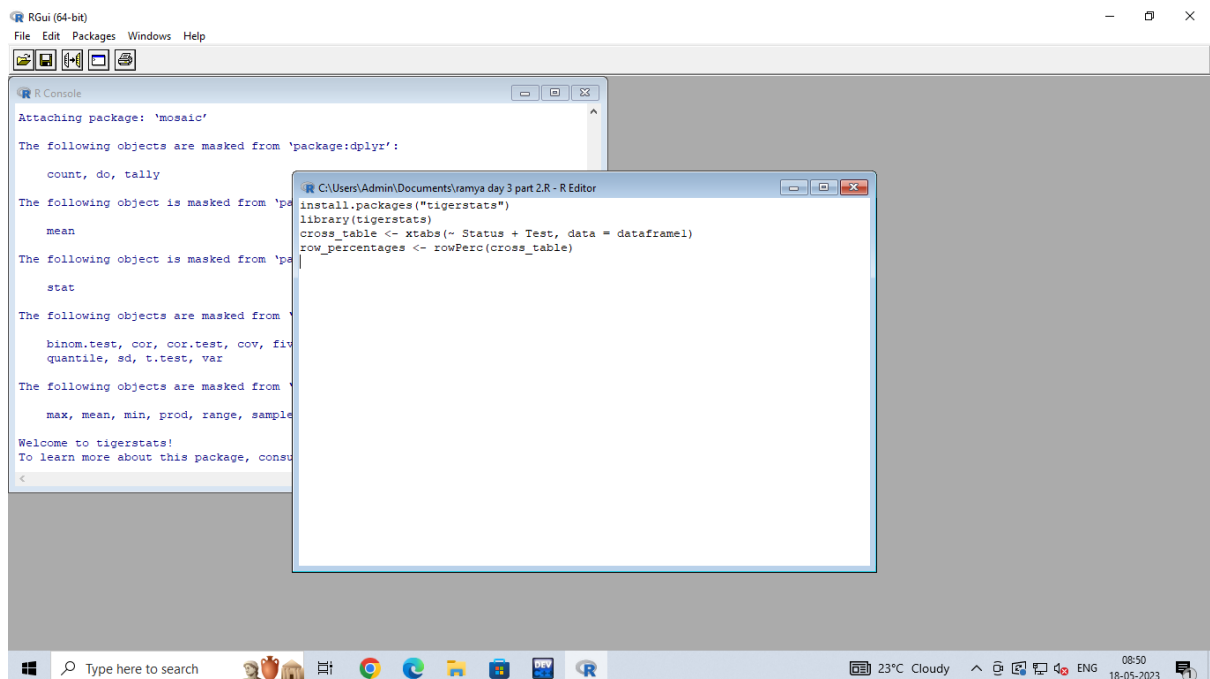
Exercise: 11

Row Percentages

The R package "tigerstats" is required for the next two exercises.

1) Create an xtabs() formula that cross-tabulates "Status", and "Test".

2) Enclose the xtabs() formula in the tigerstats function, "rowPerc()" to display row percentages for "Status" by "Test".

Exercise 12

Column Percentages

1) Create an xtabs() formula that cross-tabulates "Status", and "Test".

2) Enclose the xtabs() formula in the tigerstats function, "colPerc()" to display row percentages for "Status" by "Test".

VISUALIZATION IN R

13. Write a program for creating a pie-chart in R using the input vector(21,62,10,53). Provide labels for the chart as 'London', 'New York', 'Singapore', 'Mumbai'. Add a title to the chart as 'city pie-chart' and add a legend at the top right corner of the chart.
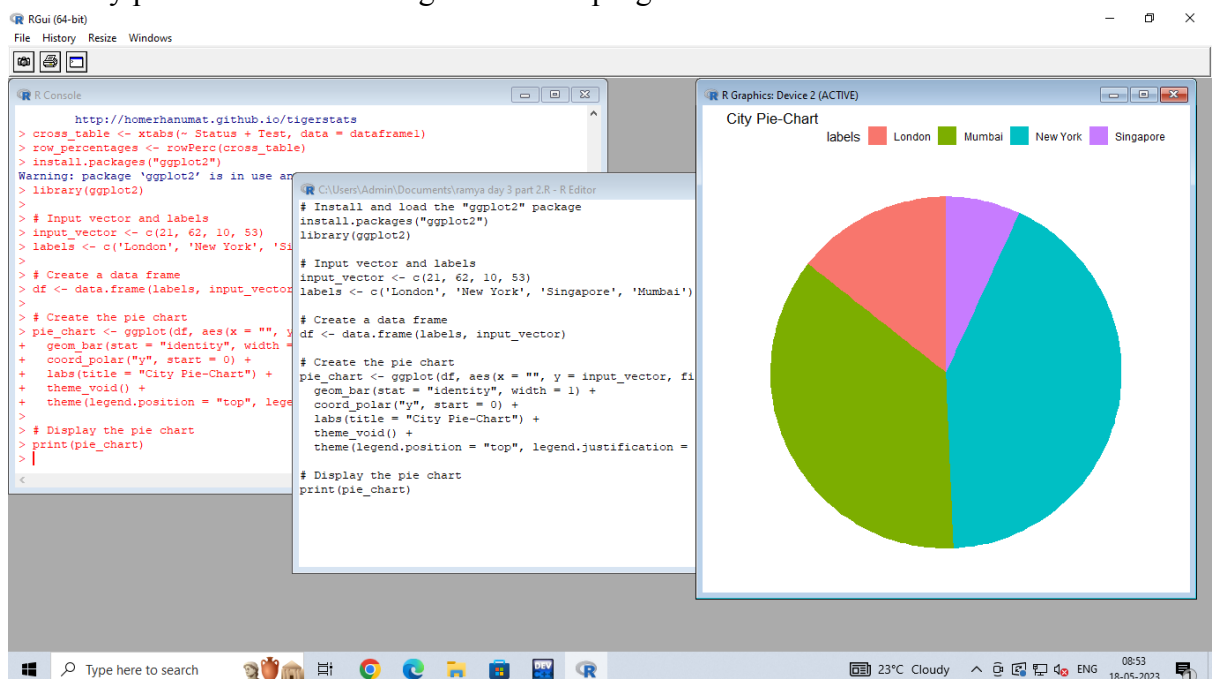
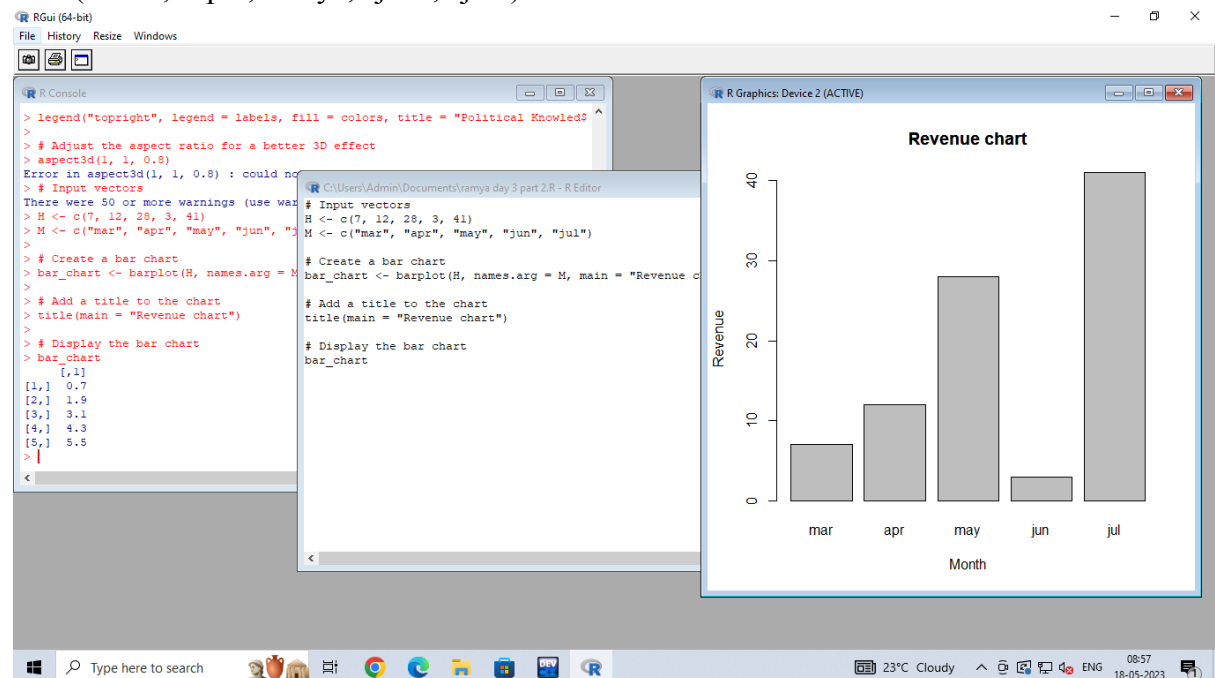14. Create a 3D Pie Chart for the dataset "political Knowledge" with suitable labels,colours and a legend at the top right corner of the chart.
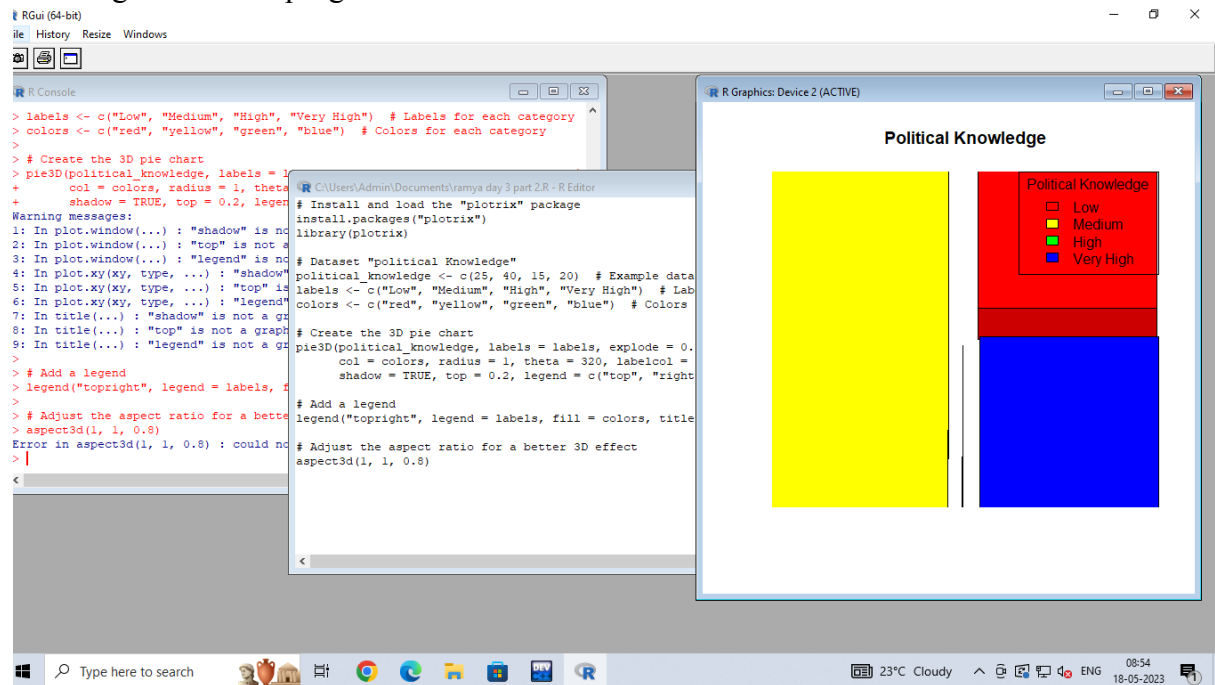
15. Write a program for creating a bar chart using the vectors H=c(7,12,28,3,41) and M=c("mar", "apr", "may", "jun", "jul"). Add a title to the chart as "Revenue chart".

16. Make a histogram for the "AirPassengers"dataset, start at 100 on the x-axis, and from values 200 to 700, make the bins 200 wide

program;

```
# Load the "datasets" package (if not already loaded)
library(datasets)
```

```
# Access the "AirPassengers" dataset
data("AirPassengers")

# Set the breaks for the histogram bins
breaks <- seq(200, 800, by = 200)

# Create the histogram
hist(AirPassengers, xlim = c(100, 800), ylim = c(0, 100), breaks = breaks, main = "Histogram
of AirPassengers",
    xlab = "Passenger Count", ylab = "Frequency")

# Add a title to the histogram
title(main = "Histogram of AirPassengers")
```

17. Create a Boxplot graph for the relation between "mpg"(miles per galloon) and "cyl"(number of Cylinders) for the dataset "mtcars" available in R Environment.