**Day 3 Lab Manual Part 2**

**Name:** U. Mohan Srinivas

**Reg No:** 192110320

**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**

1. **Load the dataset and Create a data frame and name it as dataframe1**

**Source Code:**

dataframe1 <- data.frame(

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

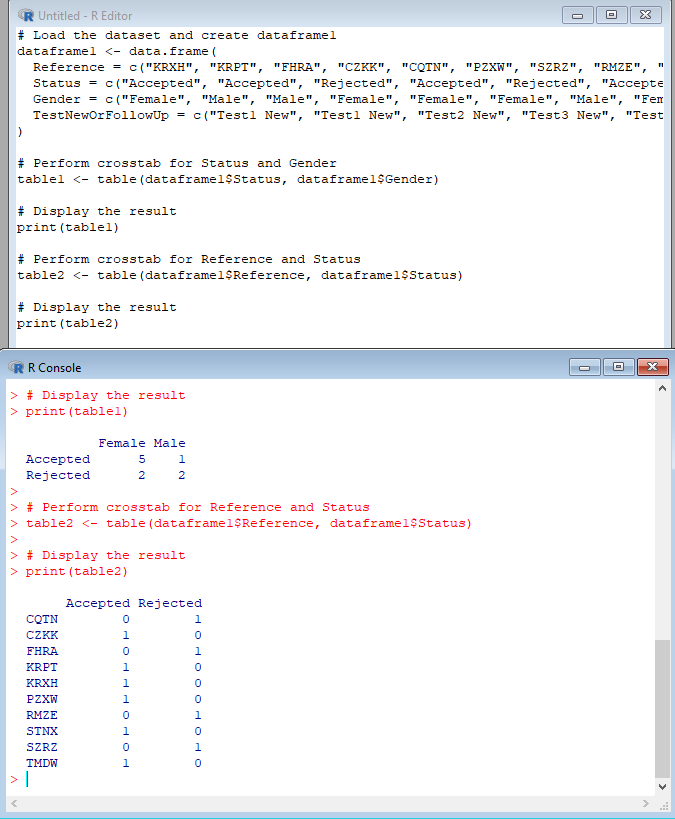
"Status" = c("Accepted", "Accepted", "Rejected", "Accepted", "Rejected", "Accepted", "Rejected", "Rejected", "Accepted", "Accepted"),

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

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

)

**OUTPUT:**

****

**Exercise: 9**

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

**dataset**

**Source Code:**

# Load the dataset

data(mtcars)

# View the first few rows of the dataset

head(mtcars)

# Explore the relationship between gear and vs variables

table1 <- table(mtcars$gear, mtcars$vs)

print(table1)

# Calculate the percentage of each combination of gear and vs variables

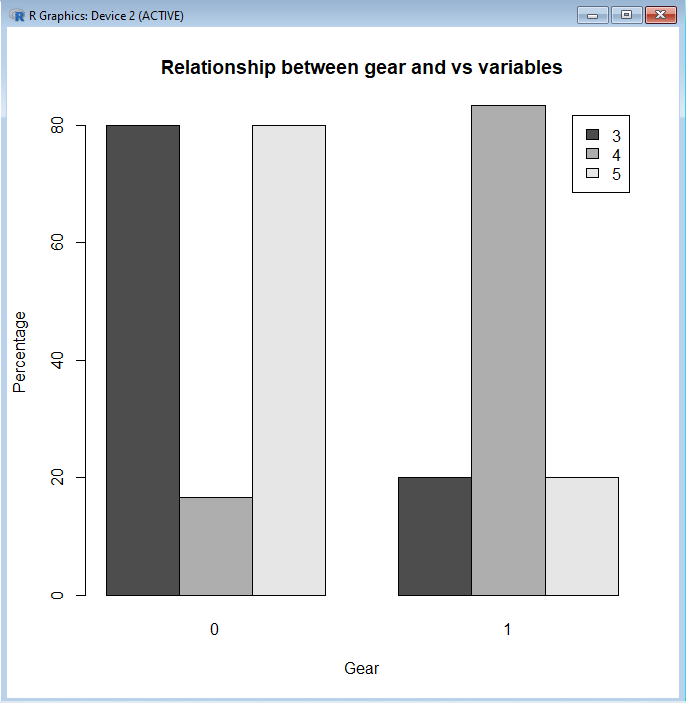
table2 <- prop.table(table1, margin = 1) \* 100

print(table2)

# Visualize the relationship between gear and vs variables using a stacked bar plot

barplot(table2, beside = TRUE, legend.text = TRUE, main = "Relationship between gear and vs variables", xlab = "Gear", ylab = "Percentage")

**OUTPUT:**



**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.**

1. **Save the file in the name of dataframe2**

**Source Code:**

# Use xtabs() to create a table showing the relationship between Reference and Status

table3 <- xtabs(~ Reference + Status, data = dataframe1)

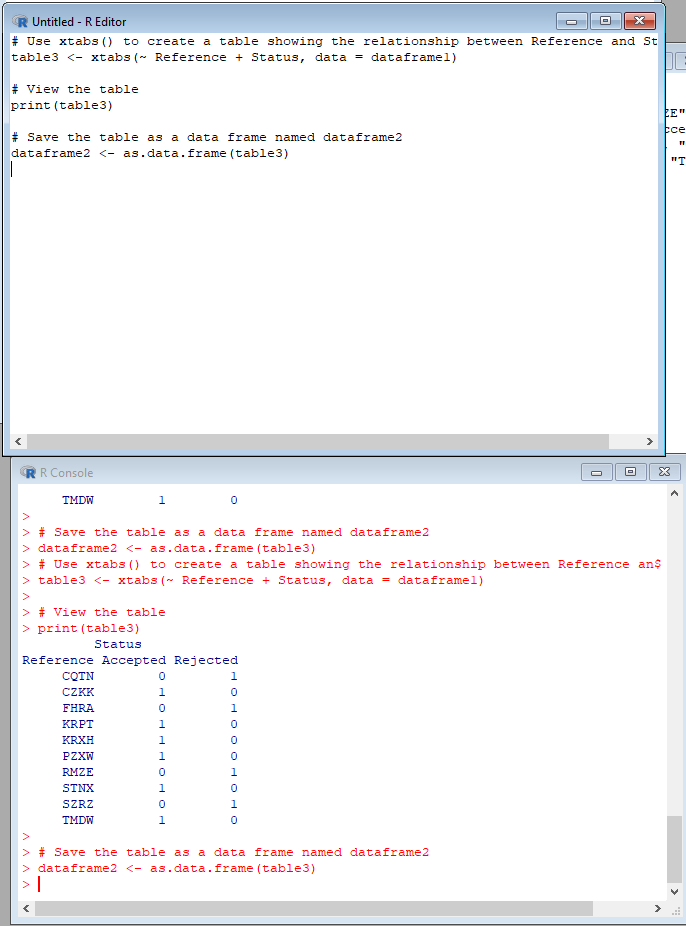
# View the table

print(table3)

# Save the table as a data frame named dataframe2

dataframe2 <- as.data.frame(table3)

**OUTPUT:**



**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“.**

**Source Code:**

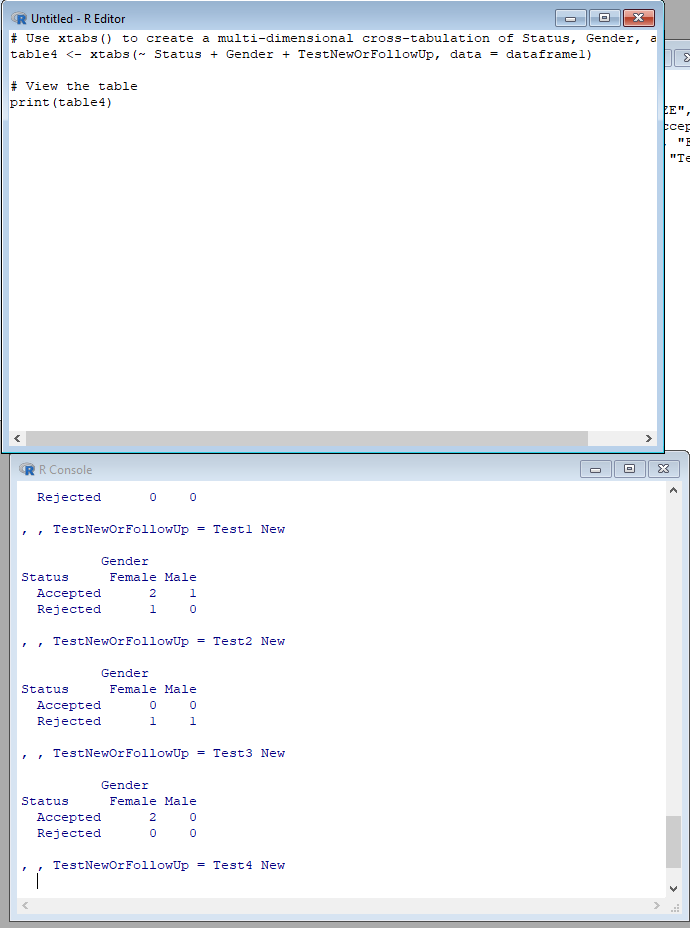
# Use xtabs() to create a multi-dimensional cross-tabulation of Status, Gender, and Test

table4 <- xtabs(~ Status + Gender + TestNewOrFollowUp, data = dataframe1)

# View the table

print(table4)

**OUTPUT:**

****

**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“.**

**Source Code:**

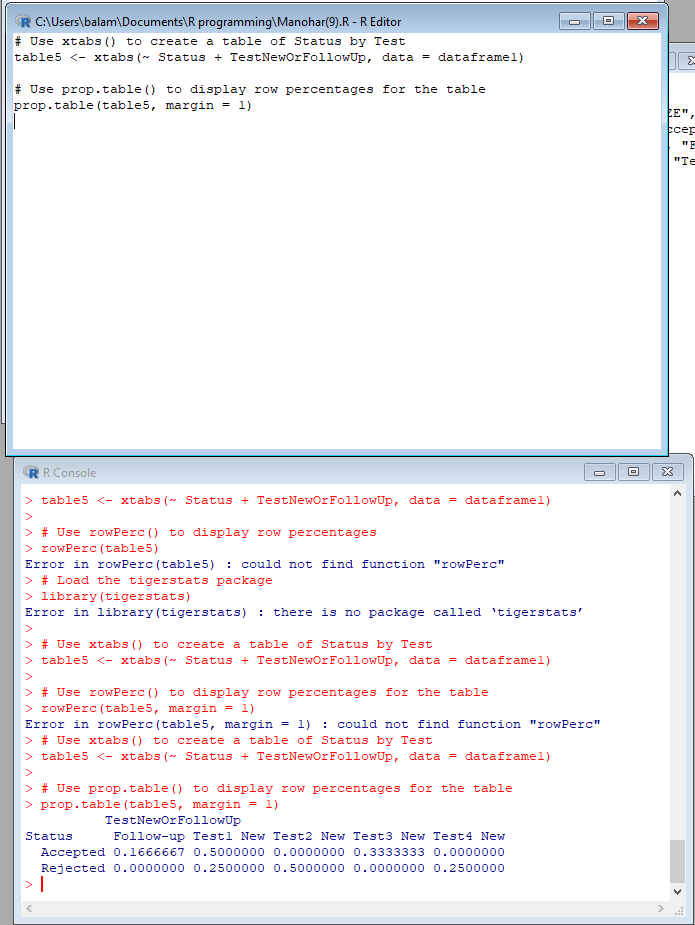
# Use xtabs() to create a table of Status by Test

table5 <- xtabs(~ Status + TestNewOrFollowUp, data = dataframe1)

# Use prop.table() to display row percentages for the table

prop.table(table5, margin = 1)

**OUTPUT:**



**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“.**

**Source Code:**

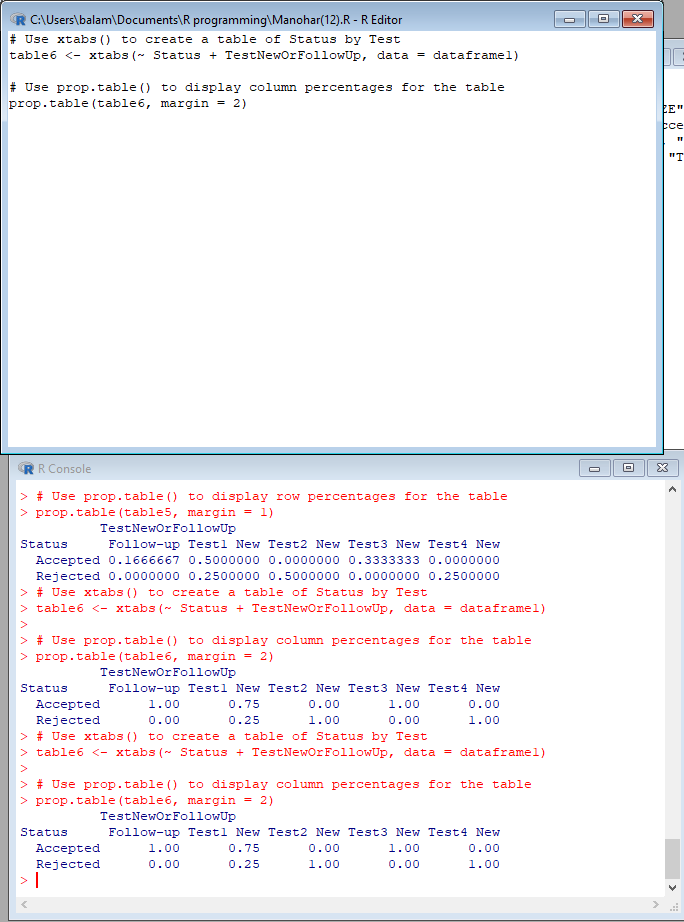
# Use xtabs() to create a table of Status by Test

table6 <- xtabs(~ Status + TestNewOrFollowUp, data = dataframe1)

# Use prop.table() to display column percentages for the table

prop.table(table6, margin = 2)

**OUTPUT:**



**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.**

**Source Code:**

input <- c(21, 62, 10, 53)

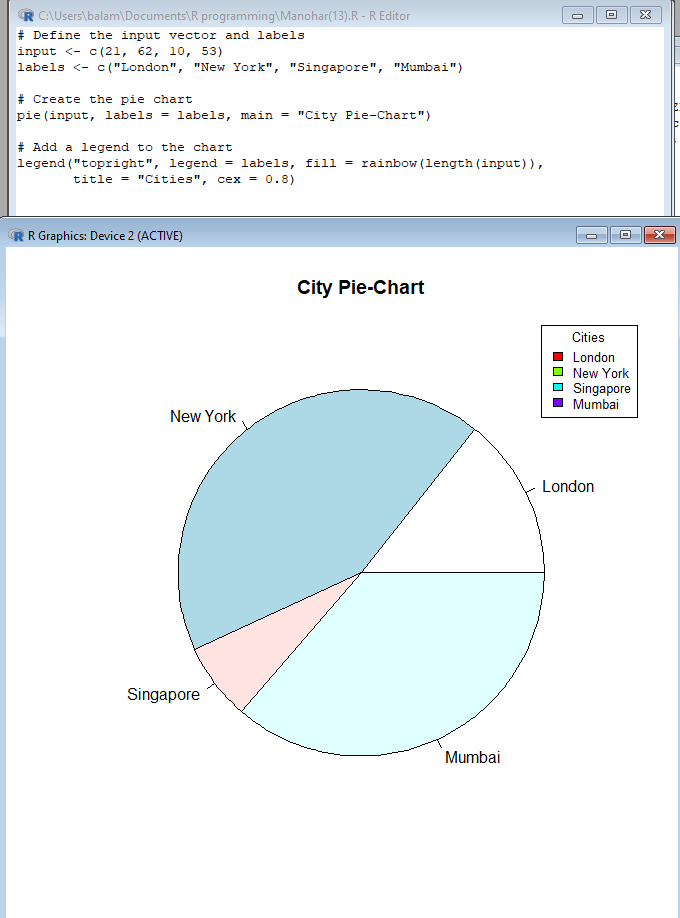
labels <- c("London", "New York", "Singapore", "Mumbai")

pie(input, labels = labels, main = "City Pie-Chart")

legend("topright", legend = labels, fill = rainbow(length(input)),

title = "Cities", cex = 0.8)

**OUTPUT:**



**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.**

**Source Code:**

# Load the plotrix package

library(plotrix)

# Define the dataset

political\_knowledge <- c(20, 30, 40, 10)

labels <- c("Low", "Medium", "High", "Very High")

colors <- c("red", "orange", "yellow", "green")

# Create the 3D pie chart

pie3D(political\_knowledge, labels = labels, explode = 0.1, col = colors,

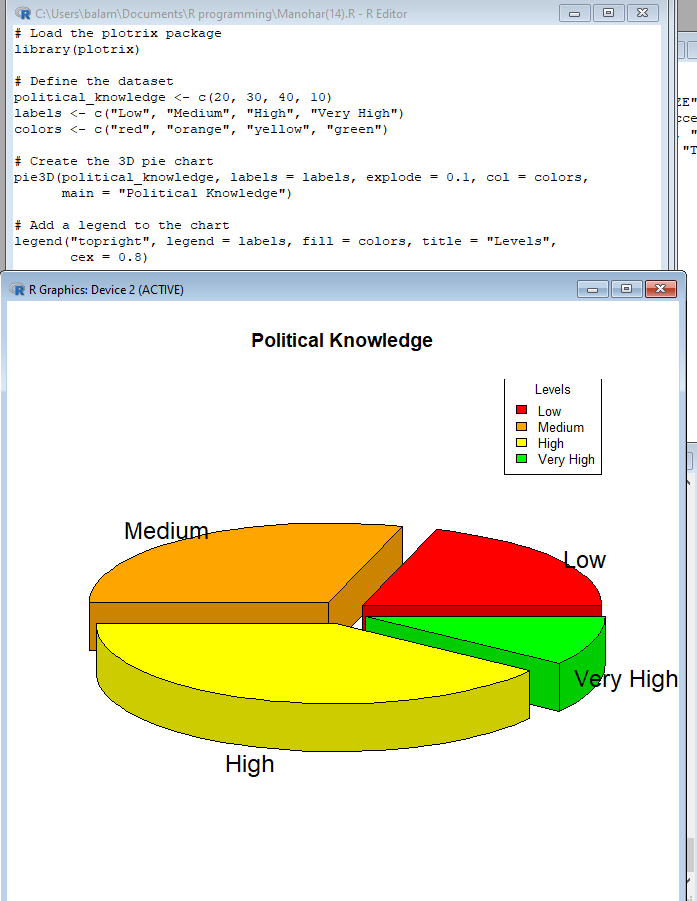
main = "Political Knowledge")

# Add a legend to the chart

legend("topright", legend = labels, fill = colors, title = "Levels",

cex = 0.8)

**OUTPUT:**



**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”.**

**Source Code:**

# Define the vectors

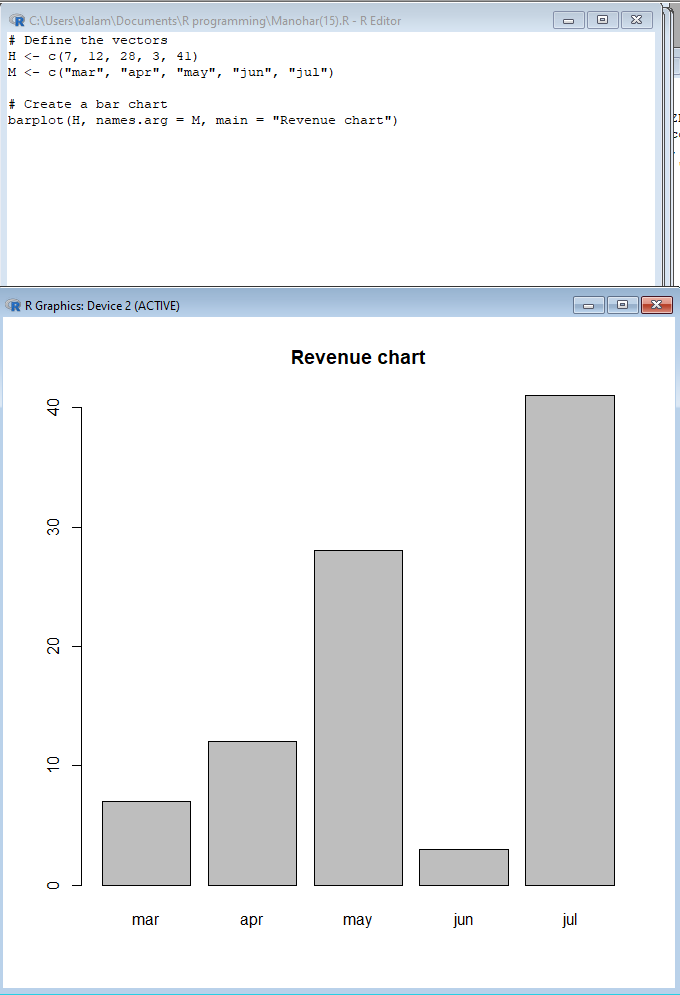
H <- c(7, 12, 28, 3, 41)

M <- c("mar", "apr", "may", "jun", "jul")

# Create a bar chart

barplot(H, names.arg = M, main = "Revenue chart")

**OUTPUT:**



**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**

**Source Code:**

# Load the AirPassengers dataset

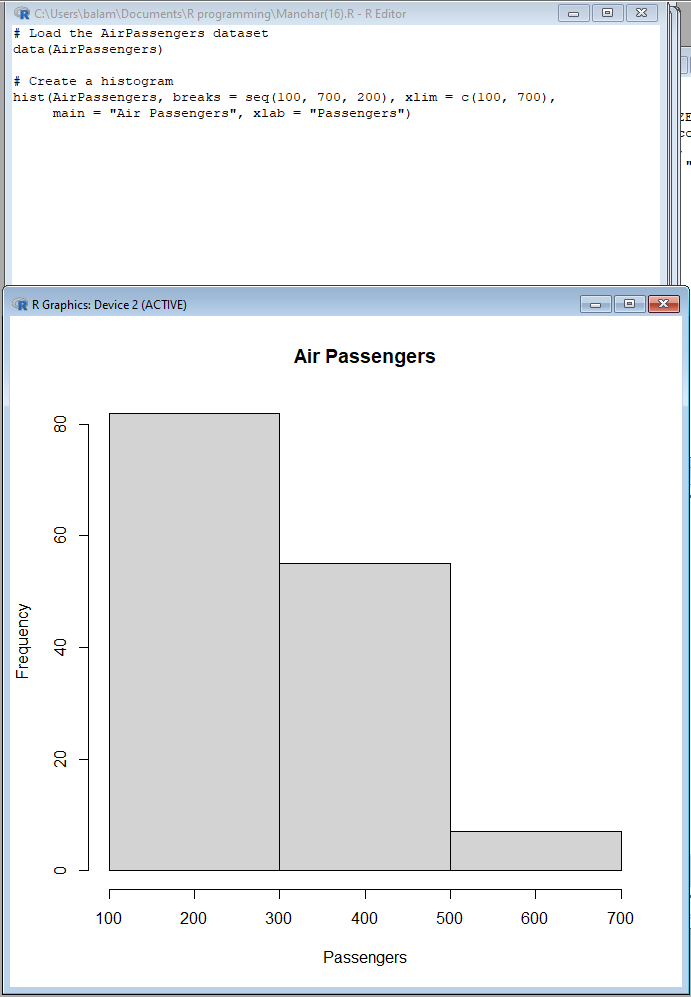
data(AirPassengers)

# Create a histogram

hist(AirPassengers, breaks = seq(100, 700, 200), xlim = c(100, 700),

main = "Air Passengers", xlab = "Passengers")

**OUTPUT;**

****

**17. Create a Boxplot graph for the relation between &quot;mpg&quot;(miles per galloon) and**

**&quot;cyl&quot;(number of Cylinders) for the dataset &quot;mtcars&quot; available in R Environment.**

**Source Code:**

# Load the "mtcars" dataset

data(mtcars)

# Create a boxplot graph for the relationship between "mpg" and "cyl"

boxplot(mpg ~ cyl, data = mtcars, main = "Miles per Gallon by Number of Cylinders",

xlab = "Number of Cylinders", ylab = "Miles per Gallon")

**OUTPUT;**

