**Day 3 Lab Manual Part 2**

**NAME : B. Pradeep**

**REG NO : 192121092**

**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**

* **Load the dataset and Create a data frame and name it as dataframe1**
* **Load the function for crosstab**

**xtabs(~colname , data=Data frame name )**

**Program:**

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"),

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

NewOrFollowUp = c("New", "New", "New", "New", "New", "Follow-up", "New", "New", "New", "New"))

library(tidyr)

xtabs(~Status + Gender, data = dataframe1)

**Exercise: 9**

* **Use Two Categorical Variables and Discover the relationships within a dataset**
* **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.**
* **Save the file in the name of dataframe2**

**Program:**

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

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“.**

**Program:**

xtabs(~Status+Gender+Test,data=dataframe1)

**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“.**

**Program:**

library(tigerstats)

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"),

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

NewOrFollowUp = c("New", "New", "New", "New", "New", "Follow-up", "New", "New", "New", "New"))

dataframe3<-xtabs(~Status+Test,data=dataframe1)

rowperc(dataframe3)

**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“.**

**Program:**

library(tigerstats)

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"),

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

NewOrFollowUp = c("New", "New", "New", "New", "New", "Follow-up", "New", "New", "New", "New"))

dataframe3<-xtabs(~Status+Test,data=dataframe1)

colperc(dataframe3)

**VISUALIZATION IN R**

* 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.

**Program:**

library(ggplot2)

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

df<-data.frame(values,labels=c("London","New York","Singapore","Mumbai"))

ggplot(df, aes(x = "", y = values, fill = labels)) +

geom\_bar(width = 1, stat = "identity") +

coord\_polar("y", start = 0) +

ggtitle("City pie-chart") +

xlab("") +

ylab("") +

guides(fill = guide\_legend(title = "", nrow = 2, direction = "horizontal")) +

theme(legend.position = c(0.85, 0.9),

plot.title = element\_text(hjust = 0.5),

axis.line = element\_blank(),

axis.text = element\_blank(),

axis.ticks = element\_blank(),

panel.grid = element\_blank(),

plot.background = element\_blank(),

legend.key = element\_blank(),

legend.title = element\_text(size = 10))



* 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.
* 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”.

**Program:**

library(ggplot2)

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

M<-c("Mar","Apr","May","Jun","Jul")

df<-data.frame(Month=M,Revenue=H)

ggplot(df, aes(x = Month, y = Revenue)) +

geom\_col(fill = "Blue") +

ggtitle("Revenue Chart") +

xlab("Month") +

ylab("Revenue") +

theme(plot.title = element\_text(hjust = 0.5))

**Output:**



* 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:**

library(ggplot2)

data("AirPassengers")

ggplot(data = AirPassengers, aes(x = AirPassengers)) +

geom\_histogram(binwidth = 200, color = "orange", fill = "blue") +

scale\_x\_continuous(limits = c(100, 700), breaks = seq(100, 700, 200)) +

ggtitle("AirPassengers Histogram") +

xlab("Passengers") +

ylab("Frequency") +

theme(plot.title = element\_text(hjust = 0.5))

* 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.

**Program:**

boxplot(mtcars$mpg,ylab="Miles Per Gallon(MPG)",xlab=”No.of.Cylinders”,col=”Purple”)

**Output:**

