**FINOLEX ACADEMY OF MANAGEMENT AND TECHNOLOGY, RATNAGIRI**

**DEPARTMENT OF MCA**

**PRACTICAL NO .05**

**INTRODUCTION TO R GRAPHICS AND DATA IMPORTING**

QUE.1. Import employee.csv file and perform following -

1. Display the content.
2. Find the dimensions of the data in the above imported dataset.
3. Get all the people with designation “clerk”.
4. Get the people whose salary is greater than 55,000 and write the output in new excel file.
5. Summarize the above dataset

ANS.

1. Display the content.

> dfdata<-read.csv("employee.csv")

> dfdata

id Name Age Designation Salary isLocal

1 1 Michelle 44 Manager 72000 NA

2 2 Ryan 27 Clerk 48000 NA

3 3 Gary 30 Clerk 54000 NA

4 4 Guru 38 Engineer 61000 NA

5 5 Harsh 40 Clerk NA NA

6 6 Brad 35 Engineer 58000 NA

7 7 James NA Clerk 52000 NA

8 8 Tina 48 Senior\_manager 79000 NA

9 9 Mina 50 CEO 83000 NA

1. 10 Tara 37 Engineer 67000 NA
2. Find the dimensions of the data in the above imported dataset.

> dim(dfdata)

[1] 10 6

1. Get all the people with designation “clerk”.

> details<-subset(dfdata,Designation=="Clerk")

> print(details)

id Name Age Designation Salary isLocal

2 2 Ryan 27 Clerk 48000 NA

3 3 Gary 30 Clerk 54000 NA

5 5 Harsh 40 Clerk NA NA

1. 7 James NA Clerk 52000 NA
2. Get the people whose salary is greater than 55,000 and write the output in new excel file.

> info\_sal<-subset(dfdata,Salary>55000)

> write.xlsx(info\_sal,"output3.xlsx",col.names = TRUE,row.names = TRUE,sheetName = "Sheet1",append = TRUE)

> new\_sal\_xlsx<-read.xlsx("output3.xlsx",sheetIndex = 1)

> new\_sal\_xlsx

NA. id Name Age Designation Salary isLocal

1 1 1 Michelle 44 Manager 72000 NA

2 4 4 Guru 38 Engineer 61000 NA

3 6 6 Brad 35 Engineer 58000 NA

4 8 8 Tina 48 Senior\_manager 79000 NA

5 9 9 Mina 50 CEO 83000 NA

6 10 10 Tara 37 Engineer 67000 NA

1. Summarize the above dataset

> summary(dfdata)

id Name Age

Min. : 1.00 Length:10 Min. :27.00

1st Qu.: 3.25 Class :character 1st Qu.:35.00

Median : 5.50 Mode :character Median :38.00

Mean : 5.50 Mean :38.78

3rd Qu.: 7.75 3rd Qu.:44.00

Max. :10.00 Max. :50.00

NA's :1

Designation Salary isLocal

Length:10 Min. :48000 Mode:logical

Class :character 1st Qu.:54000 NA's:10

Mode :character Median :61000

Mean :63778

3rd Qu.:72000

Max. :83000

NA's :1

QUE.2. The age and speed of 12 cars observed on day 1 are age1(5,7,8,7,2,2,9,4,11,12,9,6), speed1(99,86,87,88,111,103,87,94,78,77,85,86) and on day 2 following values are observed age2(2,2,8,1,15,8,12,9,7,3,11,4,7,14,12), speed2(100,105,84,105,90,99,90,95,94,100,79,112,91,80,85). Write a R program to draw a scatterplot that compares observations of the two days.

ANS.

> x1 <- c(5,7,8,7,2,2,9,4,11,12,9,6)

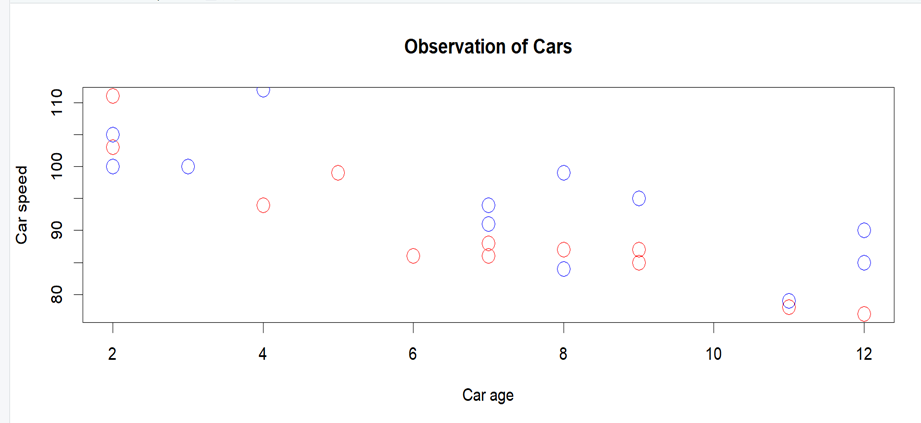
> y1 <- c(99,86,87,88,111,103,87,94,78,77,85,86)

> x2 <- c(2,2,8,1,15,8,12,9,7,3,11,4,7,14,12)

> y2 <- c(100,105,84,105,90,99,90,95,94,100,79,112,91,80,85)

> plot(x1, y1, main="Observation of Cars", xlab="Car age", ylab="Car speed", col="red", cex=2)

> points(x2, y2, col="blue", cex=2)



QUE.3. Write a R program to create a vector with numerical values in a sequence from 1 to 10 and draw a blue colored dotted line of width 2 for the above vector.

ANS.

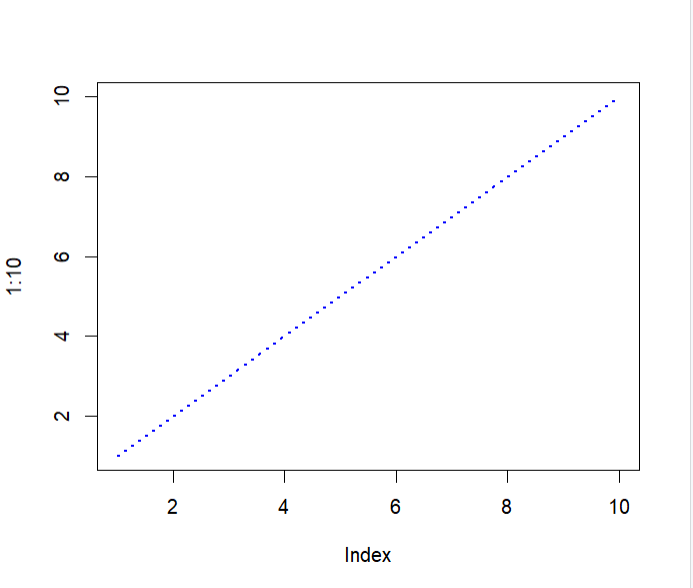
> #creating a vector

> a<-seq(1:10)

> a

[1] 1 2 3 4 5 6 7 8 9 10

> plot(1:10, type="l",col="blue", lwd=2,lty=3)



QUE.4. Write a R program to read the excel file "input.xlsx" and perform following

1. Display the content.
2. Find the dimensions of the data in the above imported dataset.
3. Get all the people working in IT department
4. Get the people who joined on or after 2014 and write the output in new excel file.
5. Summarize the above dataset

ANS.

1. Display the content.

> details<-read.xlsx("input.xlsx",sheetIndex = 1)

> details

id name salary start\_date dept

1 1 Rick 623.30 2012-01-01 IT

2 2 Dan 515.20 2013-09-23 Operations

3 3 Michelle 611.00 2014-11-15 IT

4 4 Ryan 729.00 2014-05-11 HR

5 5 Gary 843.25 2015-03-27 Finance

6 6 Nina 578.00 2013-05-21 IT

7 7 Simon 632.80 2013-07-30 Operations

8 8 Guru 722.50 2014-06-17 Finance

1. Find the dimensions of the data in the above imported dataset.

> dim(details)

[1] 8 5

1. Get all the people working in IT department.

> IT\_dept<-subset(details,dept=="IT")

> IT\_dept

id name salary start\_date dept

1 1 Rick 623.3 2012-01-01 IT

3 3 Michelle 611.0 2014-11-15 IT

1. 6 Nina 578.0 2013-05-21 IT
2. Get the people who joined on or after 2014 and write the output in new excel file.

> info\_join\_date<-subset(details,as.Date(start\_date)>=as.Date("2014-01-01"))

> write.xlsx(info\_join\_date,"output2.xlsx",col.names = TRUE,row.names = TRUE,sheetName = "Sheet1",append = TRUE)

> new\_xlsx<-read.xlsx("output2.xlsx",sheetIndex = 1)

> new\_xlsx

NA. id name salary start\_date dept

1 3 3 Michelle 611.00 2014-11-15 IT

2 4 4 Ryan 729.00 2014-05-11 HR

3 5 5 Gary 843.25 2015-03-27 Finance

4 8 8 Guru 722.50 2014-06-17 Finance

1. Summarize the above dataset

> summary(details)

id name salary

Min. :1.00 Length:8 Min. :515.2

1st Qu.:2.75 Class :character 1st Qu.:602.8

Median :4.50 Mode :character Median :628.0

Mean :4.50 Mean :656.9

3rd Qu.:6.25 3rd Qu.:724.1

Max. :8.00 Max. :843.2

start\_date dept

Min. :2012-01-01 Length:8

1st Qu.:2013-07-12 Class :character

Median :2014-01-16 Mode :character

Mean :2013-12-13

3rd Qu.:2014-07-24

Max. :2015-03-27

> summary(new\_xlsx)

NA. id

Length:4 Min. :3.00

Class :character 1st Qu.:3.75

Mode :character Median :4.50

Mean :5.00

3rd Qu.:5.75

Max. :8.00

name salary

Length:4 Min. :611.0

Class :character 1st Qu.:694.6

Mode :character Median :725.8

Mean :726.4

3rd Qu.:757.6

Max. :843.2

start\_date dept

Min. :2014-05-11 Length:4

1st Qu.:2014-06-07 Class :character

Median :2014-08-31 Mode :character

Mean :2014-09-24

3rd Qu.:2014-12-18

Max. :2015-03-27

QUE.5.Create a pie chart for favourite movie categories (comedy,action,drama,romance,sci-fi). Consider appropriate percentages for creating pies. Add a list of explanation for each pie

ANS.

> x <- c(10,20,20,10,40)

> mylabel <- c("Comedy","Action","Drama","Sci-fi","Romance")

> pie(x, label=mylabel, main="Favourite Movie Categories",col=rainbow(length(x)))

>legend("topright",c("Comedy","Action","Drama","Romance",”Sci-fi”),cex=0.6,fill=rainbow(length(x)))

