**Finolex Academy of Management and Technology, Ratnagiri**

**Department of MCA**

**Course:- MCAL13 Advance Database Management System Lab**

**Practical No -05: Graphics and Data importing**

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

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

1. **Display the content.**

> data

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

10 10 Tara 37 Engineer 67000 NA

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

> dim(data)

[1] 10 6

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

> new\_data <- subset(data,Designation=="Clerk")

> new\_data

id Name Age Designation Salary isLocal

2 Ryan 27 Clerk 48000 NA

3 Gary 30 Clerk 54000 NA

5 Harsh 40 Clerk NA NA

7 James NA Clerk 52000 NA

1. **Get the people whose salary is greater than 55,000 and write the output in new excel file.**

> info <- subset(data, Salary > 55000)

> info

id Name Age Designation Salary isLocal

1 1 Michelle 44 Manager 72000 NA

4 4 Guru 38 Engineer 61000 NA

6 6 Brad 35 Engineer 58000 NA

8 8 Tina 48 Senior\_manager 79000 NA

9 9 Mina 50 CEO 83000 NA

10 10 Tara 37 Engineer 67000 NA

> write.csv(info,"Output.csv")

> newdata1 <- read.csv("Output.csv")

> newdata1

X id Name Age Designation Salary isLocal

1 1 Michelle 44 Manager 72000 NA

4 4 Guru 38 Engineer 61000 NA

6 6 Brad 35 Engineer 58000 NA

8 8 Tina 48 Senior\_manager 79000 NA

9 9 Mina 50 CEO 83000 NA

10 10 Tara 37 Engineer 67000 NA

1. **Summarize the above dataset**

> summary(newdata1)

X id Name

Min. : 1.000 Min. : 1.000 Length:6

1st Qu.: 4.500 1st Qu.: 4.500 Class :character

Median : 7.000 Median : 7.000 Mode :character

Mean : 6.333 Mean : 6.333

3rd Qu.: 8.750 3rd Qu.: 8.750

Max. :10.000 Max. :10.000

Age Designation Salary

Min. :35.00 Length:6 Min. :58000

1st Qu.:37.25 Class :character 1st Qu.:62500

Median :41.00 Mode :character Median :69500

Mean :42.00 Mean :70000

3rd Qu.:47.00 3rd Qu.:77250

Max. :50.00 Max. :83000

isLocal

Mode:logical

NA's:6

1. **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.**

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

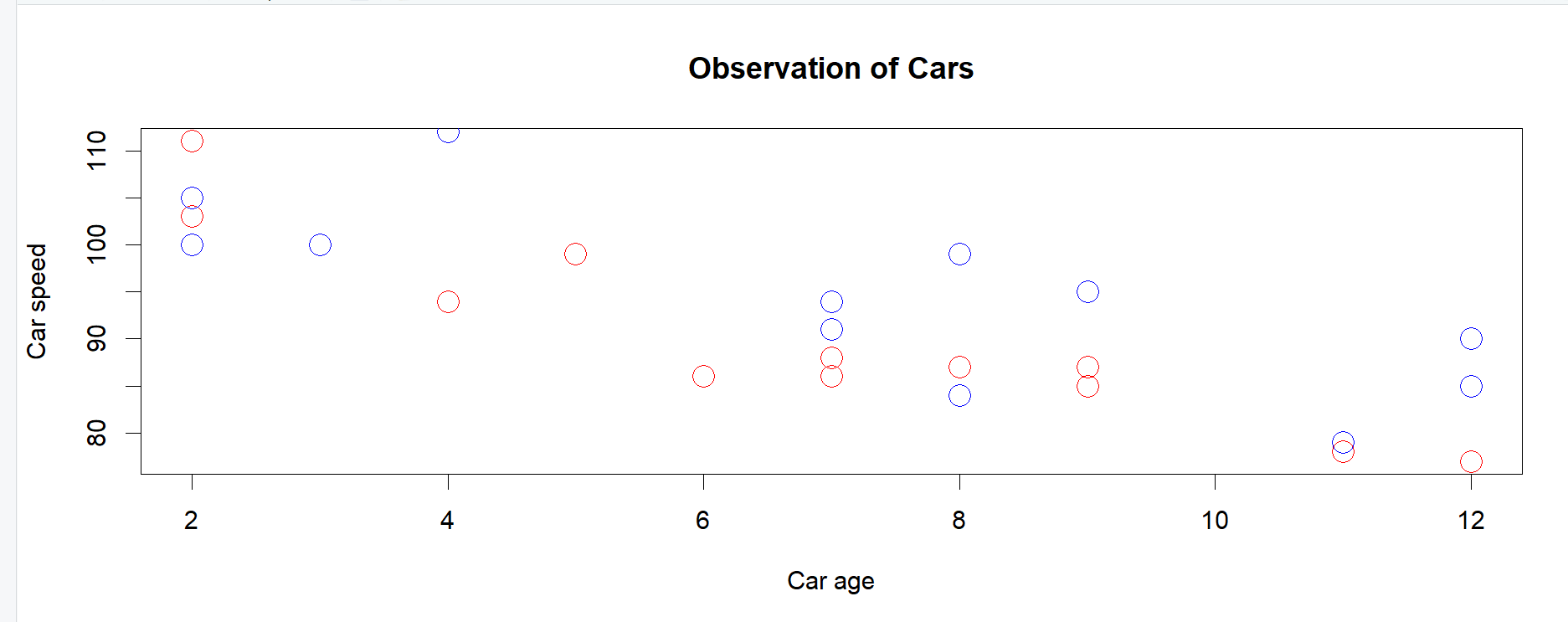
> y1 <- c(99,86,87,88,111,103,87,94,78,77,85,86)

> 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)

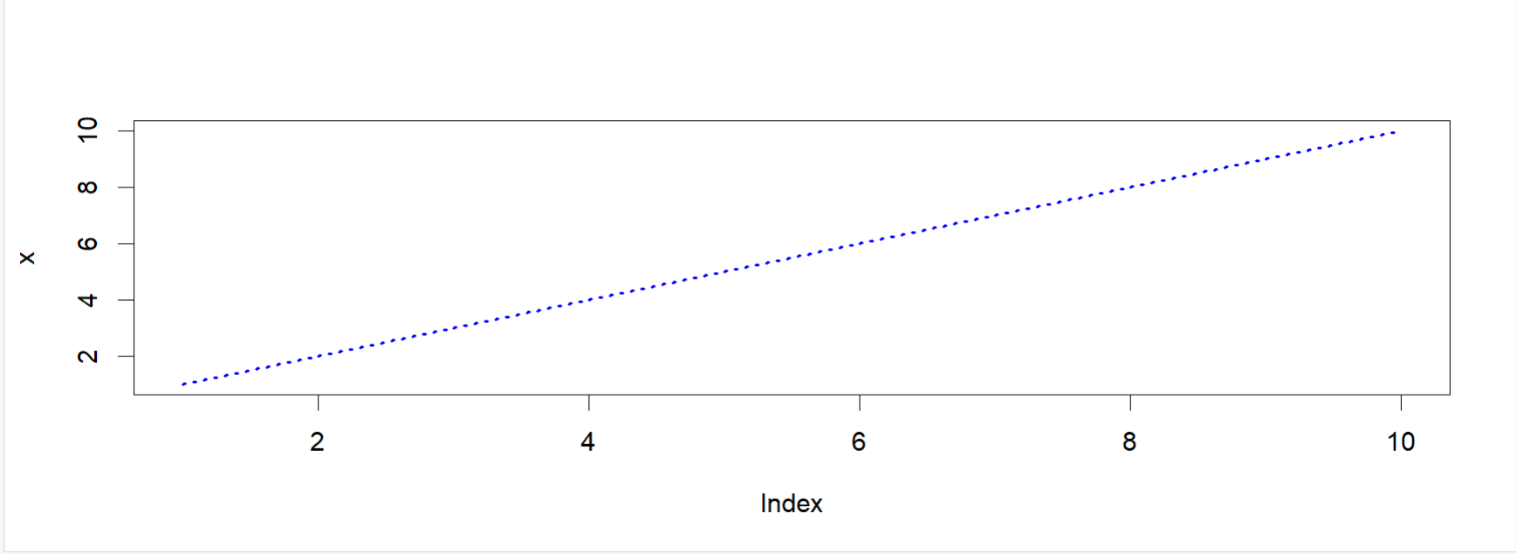
> points(x2, y2, col="blue", cex=2)



iii) **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.**

> x <- 1:10

> plot(x, type="l",col="blue",lty="dotted", lwd=2)



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

install.packages(xlsx)

library(xlsx)

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

1. **Display the content.**

> data

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(data)

[1] 8 5

1. **Get all the people working in IT department**

> new\_data <- subset(data, dept=="IT")

> new\_data

id name salary start\_date dept

1 1 Rick 623.3 2012-01-01 IT

3 3 Michelle 611.0 2014-11-15 IT

6 6 Nina 578.0 2013-05-21 IT

1. **Get the people who joined on or after 2014 and write the output in new excel file.**

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

> retval <- subset(data, as.Date(start\_date) > as.Date("2014-01-01"))

> write.xlsx(retval,"output.xlsx")

> new <- read.xlsx("output.xlsx",sheetIndex = 1)

> new

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(new)

NA. id name

Length:4 Min. :3.00 Length:4

Class :character 1st Qu.:3.75 Class :character

Mode :character Median :4.50 Mode :character

Mean :5.00

3rd Qu.:5.75

Max. :8.00

salary start\_date

Min. :611.0 Min. :2014-05-11

1st Qu.:694.6 1st Qu.:2014-06-07

Median :725.8 Median :2014-08-31

Mean :726.4 Mean :2014-09-24

3rd Qu.:757.6 3rd Qu.:2014-12-18

Max. :843.2 Max. :2015-03-27

dept

Length:4

Class :character

Mode :character

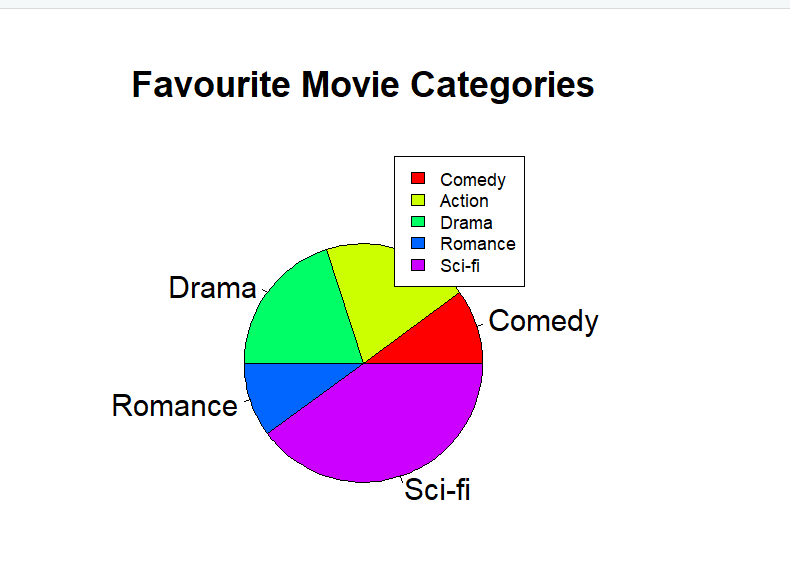
1. **Create a pie chart for favourite movie categories (comedy,action,drama,romance,sci-fi). Consider appropriate percentages for creating pies. Add a list of explanation for each pie**

> 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)))

****