

MATHS LAB EXPERIMENT 1 TO 5

NAME: ADARSH GAURAV

REG NO: 21BCE1628

LAB-1

Introduction: Understanding Data types; importing/exporting data

CODE

```
#Generate data
```

```
1:10
```

```
#Assign variable name to the value
```

```
X = 10
```

```
# Combine numeric values into vector
```

```
c(1,2,5)
```

```
#Arithmetic operations on vectors
```

```
a = c(1, 3, 5, 7)
```

```
b = c(2, 4, 6, 8)
```

```
#Addition
```

```
a+b
```

```
#Subtraction
```

```
a-b
```

```
#Constant multiplication
```

```
5*a
```

```
#Product
```

```
a*b
```

```
#Division
```

```
a/b
```

```
#String values
```

```
X = as.character(5.2)
```

```
X
```

OUTPUT

```
[1] 1 2 3 4 5 6 7 8 9 10
[1] 1 2 5
[1] 3 7 11 15
[1] -1 -1 -1 -1
[1] 5 15 25 35
[1] 2 12 30 56
[1] 0.5000000 0.7500000 0.8333333 0.8750000
[1] "5.2"
```

CODE

#Generate data

1:10

#Assign variable name to the value

X <- 5

Combine numeric values into vector

c(12,43,14,54,100)

c

#Arithmetic operations on vectors

a = c(12, 43, 11, 14, 9, 10)

b = c(59, 110, 2, 19, 6, 39)

#Addition

a+b

#Subtraction

a-b

#Constant multiplication

10*a

#Product

a*b

#Division

a/b

#String values

X = as.character(12.9)

X

#Concatenation of strings

paste("Hello", "World")

OUTPUT

```
[1] 1 2 3 4 5 6 7 8 9 10
[1] 12 43 14 54 100
function (...) .Primitive("c")
[1] 71 153 13 33 15 49
[1] -47 -67 9 -5 3 -29
[1] 120 430 110 140 90 100
[1] 708 4730 22 266 54 390
[1] 0.2033898 0.3909091 5.5000000 0.7368421 1.5000000 0.2564103
[1] "12.9"
[1] "Hello World"
```

Computing Summary Statistics /plotting and visualizing data using Tabulation and Graphical Representations

CODE

```
empid
age = c(30, 37, 45, 32, 50, 60, 35, 32, 34, 43, 32, 30, 43, 50, 60)
age
gender = c(0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, 0, 0)
gender
x
status = c(1, 1, 2, 2, 1, 1, 1, 2, 2, 1, 2, 1, 2, 1, 2)
status
empinfo = data.frame(empid, age, gender, status)
empinfo

empinfo$gender = factor(empinfo$gender, labels=c("male", "female"))
empinfo
empinfo$status = factor(empinfo$status, labels=c("staff", "faculty"))
empinfo$status
empinfo
male = subset(empinfo, empinfo$gender=="male")
male
female = subset(empinfo, empinfo$gender=="female")
female
staff = subset(empinfo, empinfo$status=="staff")
staff
faculty = subset(empinfo, empinfo$status=="faculty")
faculty
faculty_and_female = subset(female, female$status=="faculty")
faculty_and_female
summary(empinfo)
summary(male)
summary(faculty_and_female)
summary(age)
table1 = table(empinfo$gender)
table1
table2 = table(empinfo$status)
table2
table3 = table(empinfo$gender, empinfo$status)
table3
```

```
plot(empinfo$age, type="l", main="Age of employees", xlab="empid", ylab="age in years",  
col="blue")
```

```
pie(table1)
```

```
pie(table2)
```

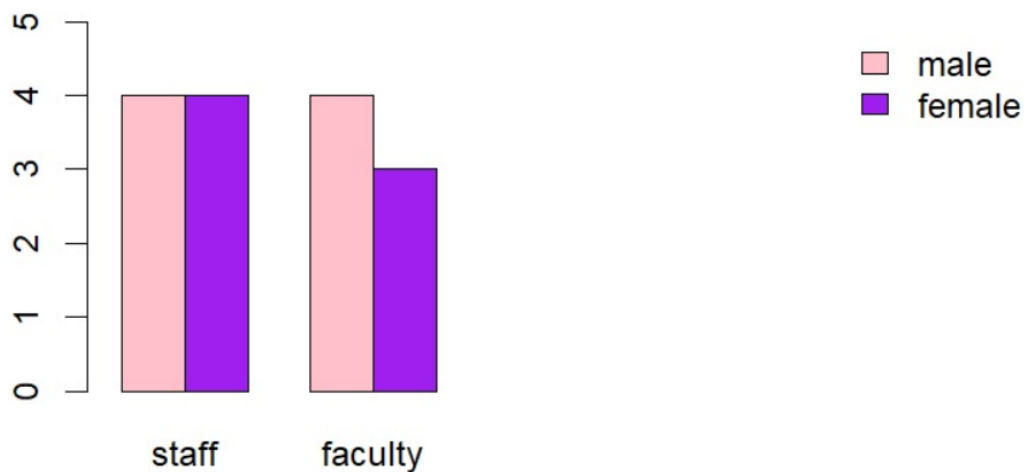
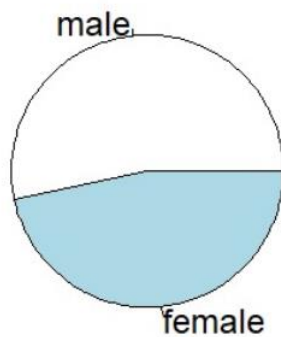
```
pie(table3)
```

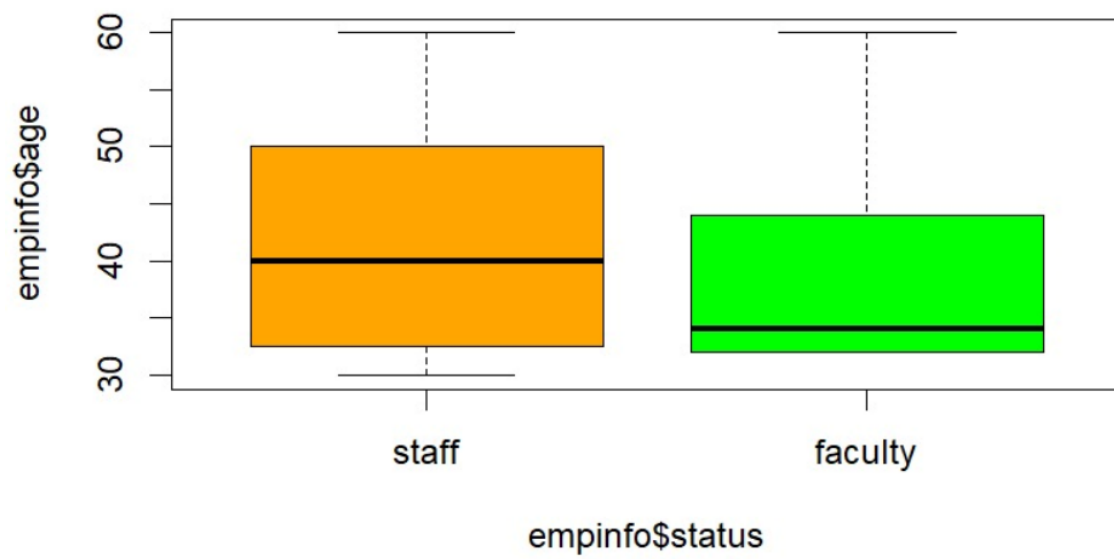
```
barplot(table3, beside=T, xlim=c(1, 15), ylim=c(0, 5), col=c("pink", "purple"))
```

```
legend("topright", legend=rownames(table3), fill=c('pink', 'purple'), bty="n")
```

```
boxplot(empinfo$age~empinfo$status, col=c('orange', 'green'))
```

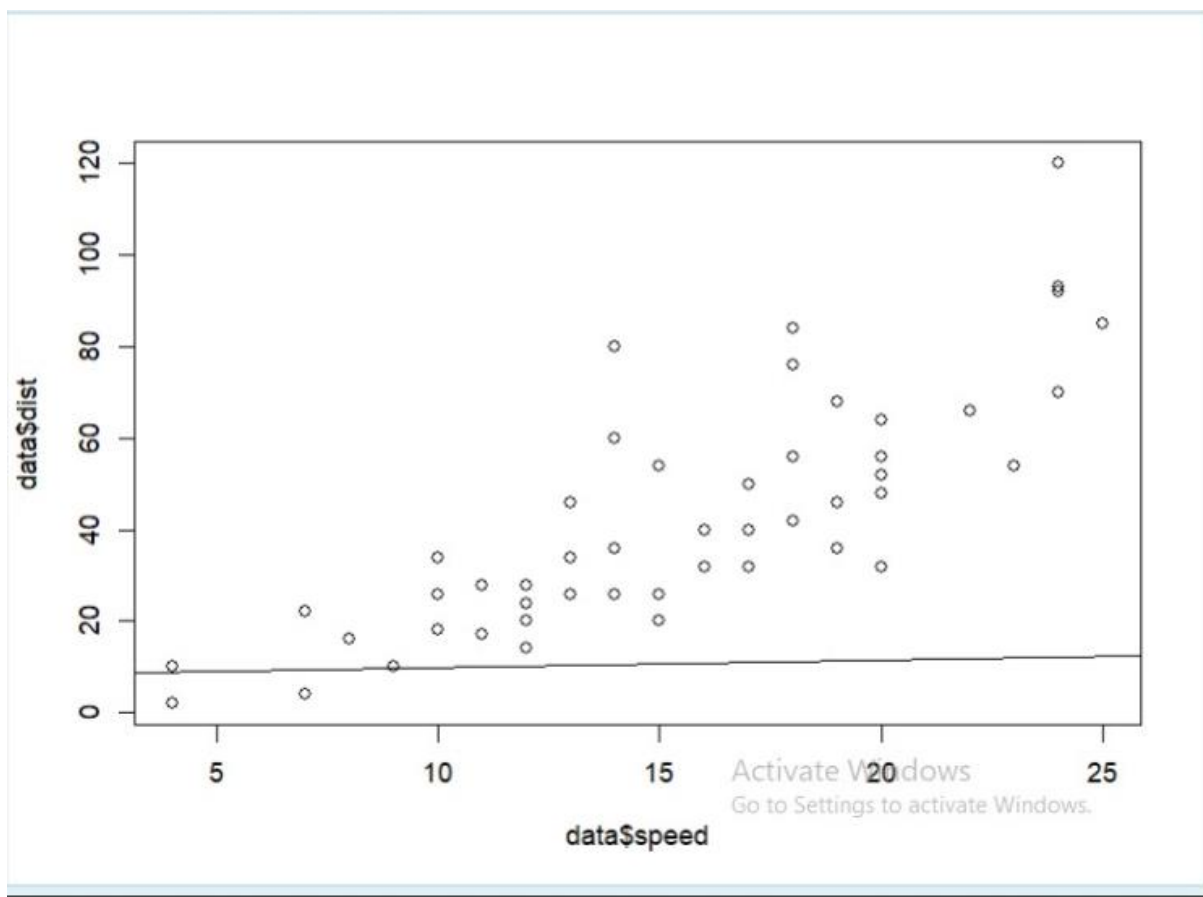
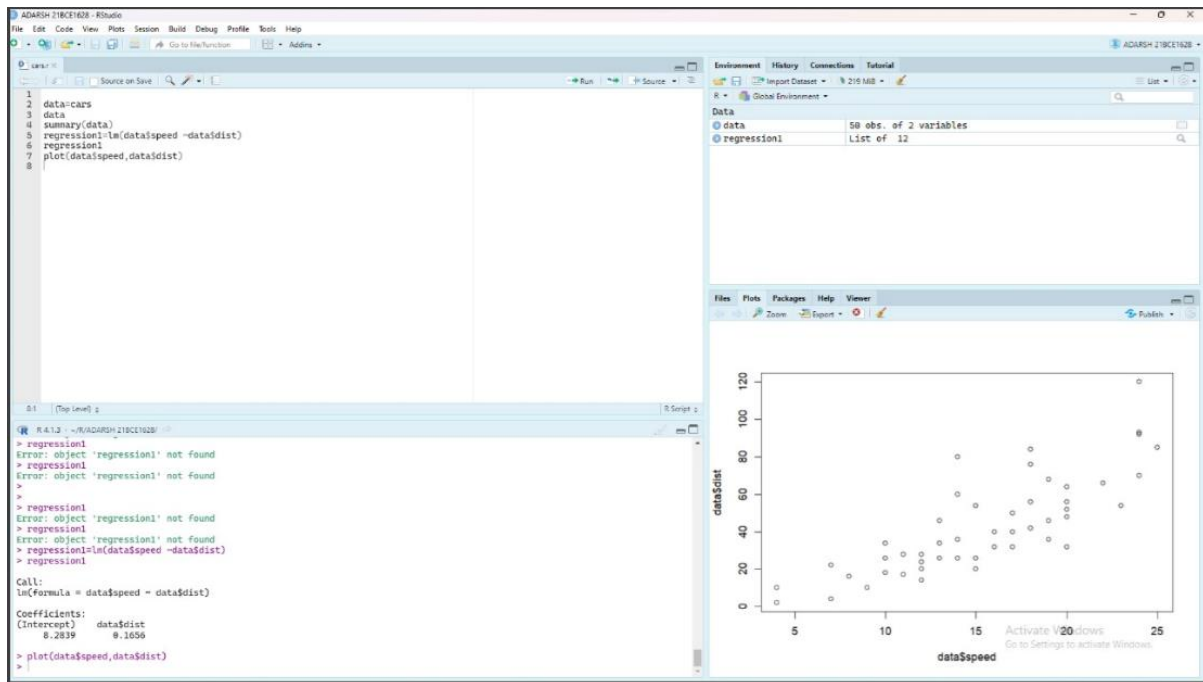
OUTPUT

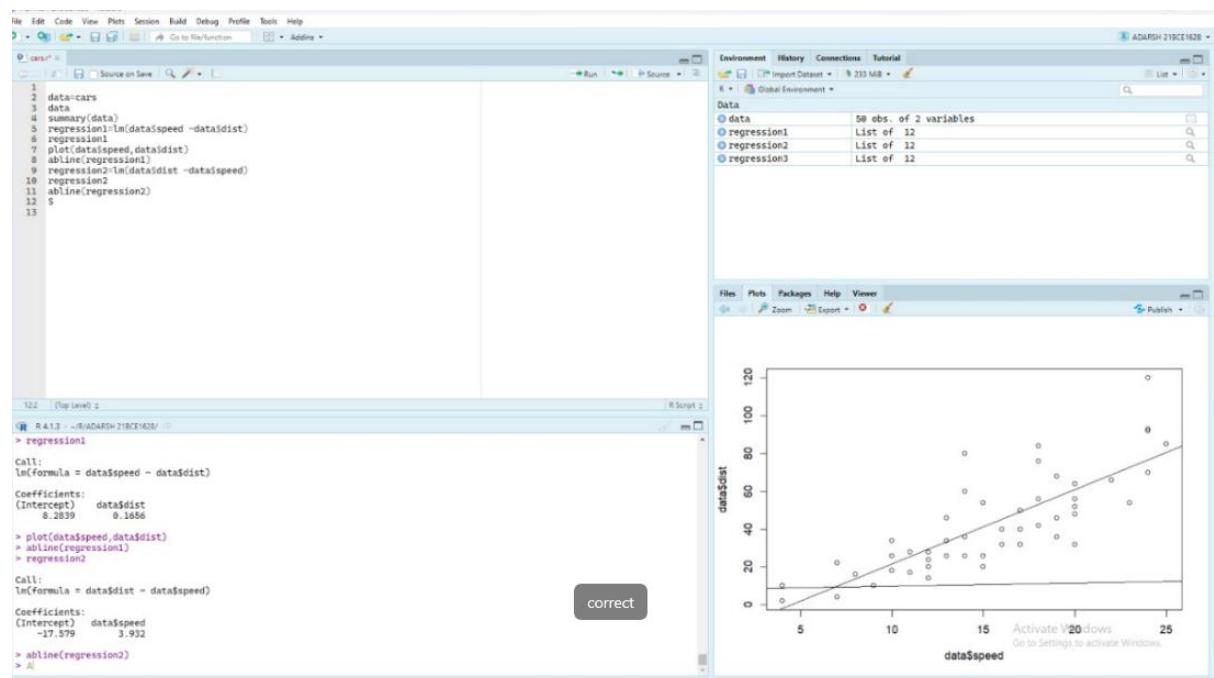
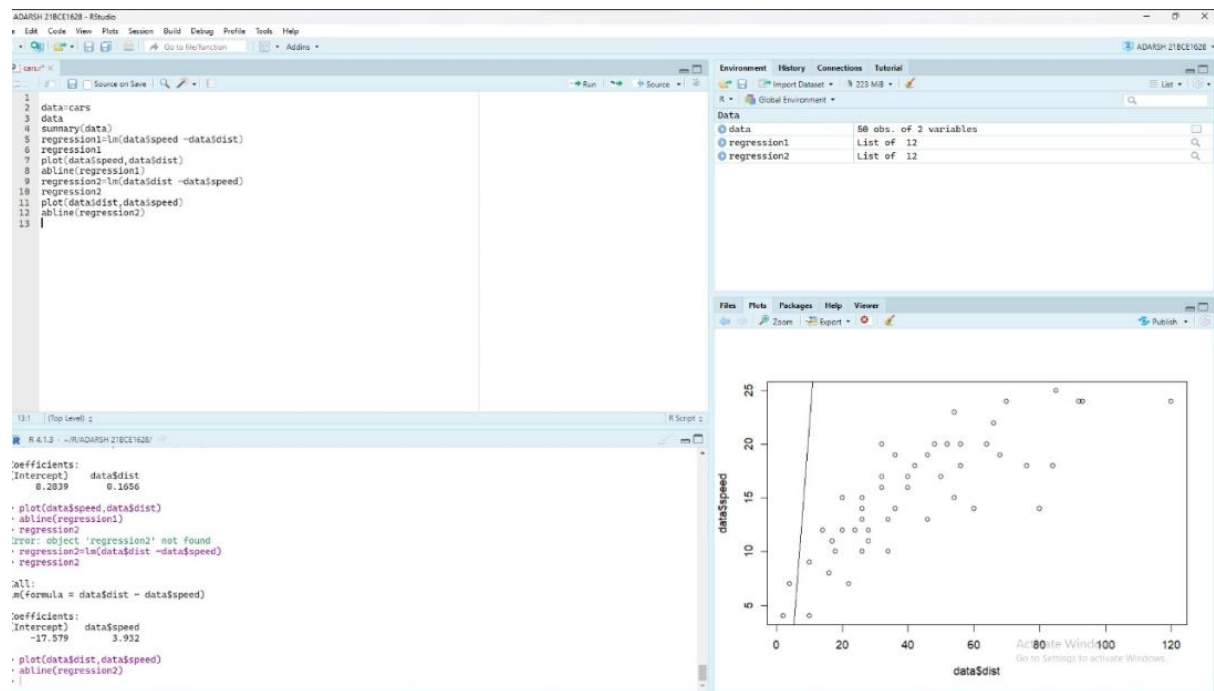




LAB-3**Applying correlation and simple linear regression model to real data set; computing and interpreting the coefficient of determination**

```
data=cars
data
summary(data)
regression1=lm(data$speed ~data$dist)
regression1
plot(data$speed,data$dist)
abline(regression1)
regression2=lm(data$dist ~data$speed)
regression2
plot(data$dist,data$speed)
abline(regression2)
```





CODE

```
weight = c(15,26,27,25,25.5,27,32,18,22,20,26,24)
```

```
weight
```

```
bmi = c(13.35,16.12,16.74,16,13.59,15.73,15.65,13.85,16.07,12.8,13.65,14.42)
```

```
bmi
```

```
cor(weight,bmi)
```

```
Regression1 = lm(bmi~weight)
```

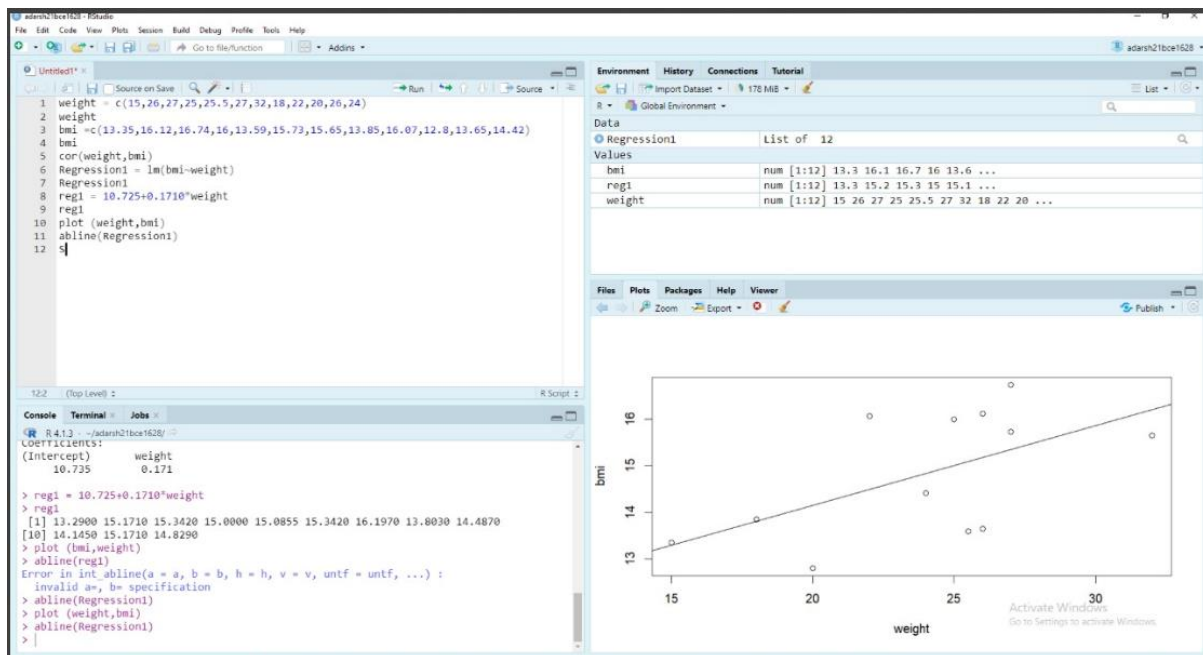
```
Regression1
```

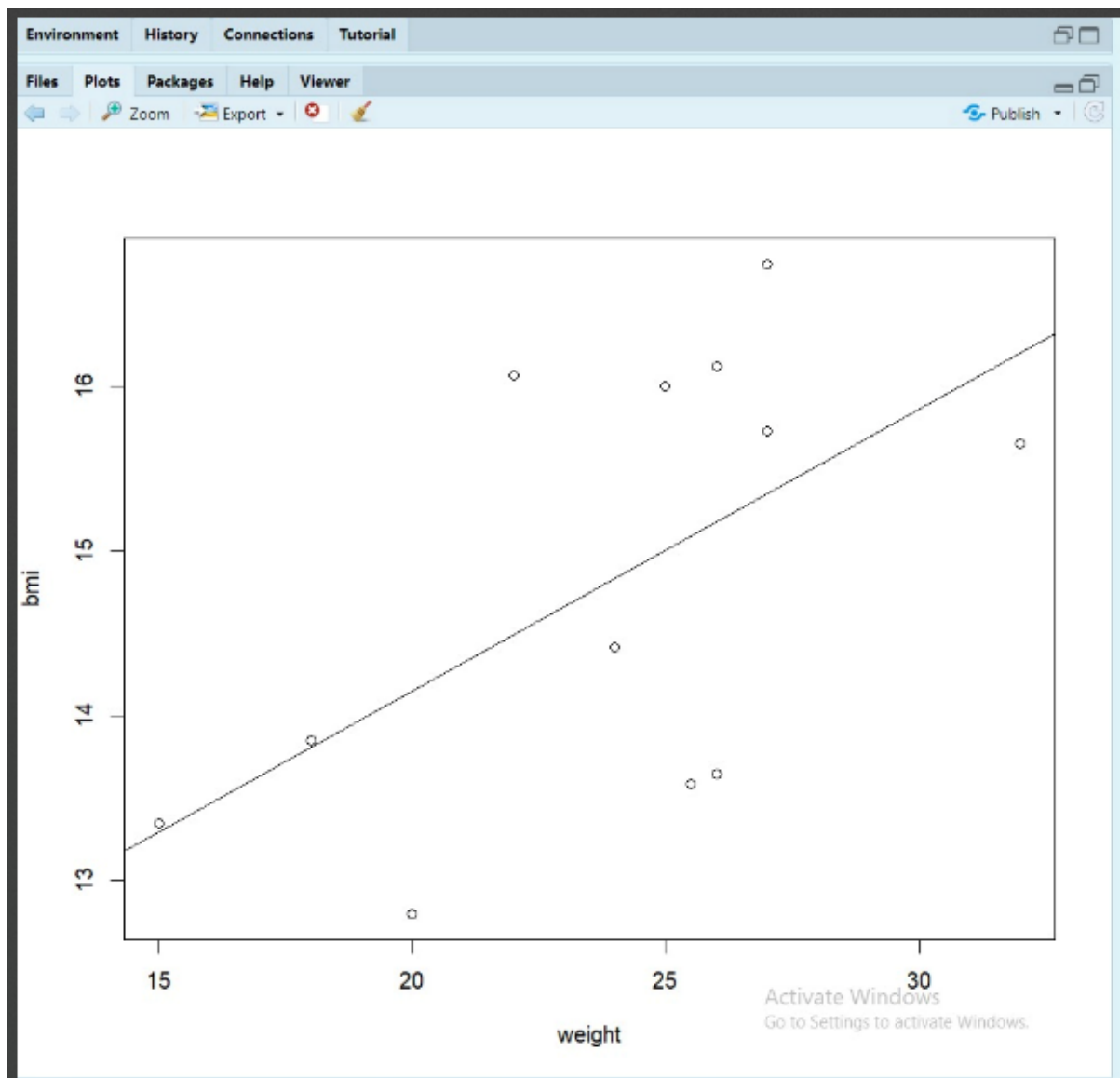
```
reg1 = 10.725+0.1710*weight
```

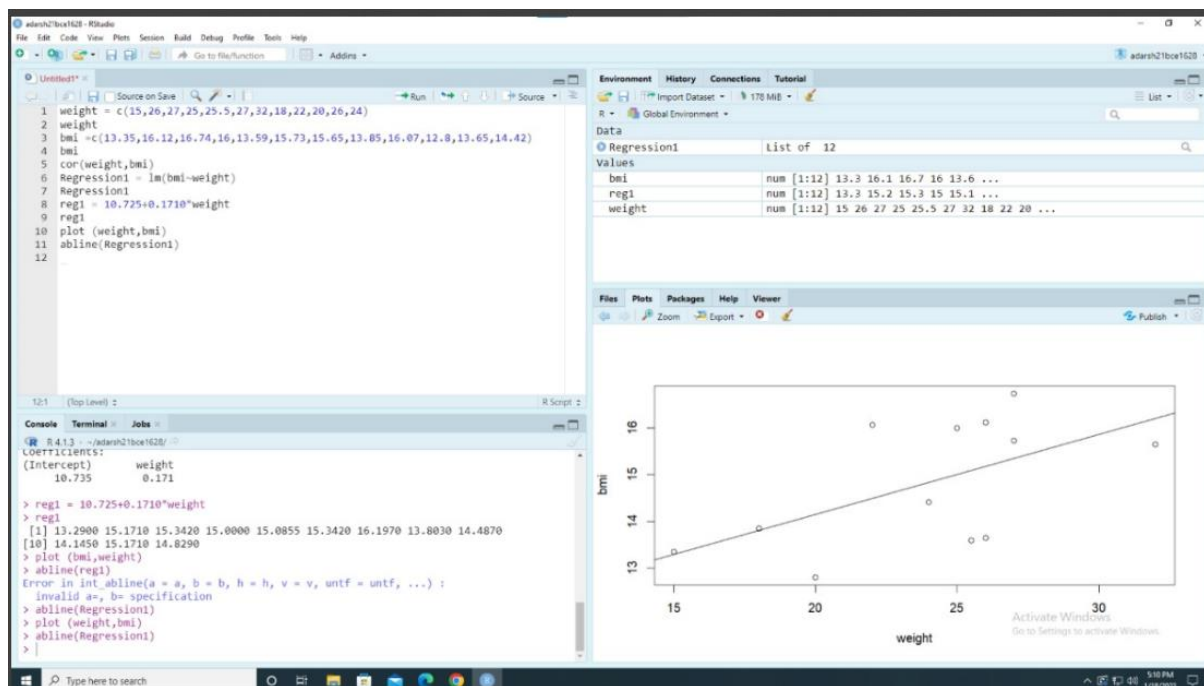
```
reg1
```

```
plot (weight,bmi)
```

```
abline(Regression1)
```







LAB 4

Applying multiple linear regression model to real dataset; computing and interpreting the multiple coefficients of determination

CODE

```

Y=c(110,80,70,120,150,90,70,120)
Y
x1=c(30,40,20,50,60,40,20,60)
x1
x2=c(11,10,7,15,19,12,8,14)
x2
# linera regression model of y on x1 and x2
regmodel=lm(Y~x1+x2)
regmodel
#summary of regresiion
summary(regmodel)
#install.packages("scatterplot3d")
library(scatterplot3d)
scatterplot3d(Y,x1,x2)

```

```

R 4.1.3 ~ /adarsh21bce1628/
>
>
> Y=c(110,80,70,120,150,90,70,120)
> Y
[1] 110 80 70 120 150 90 70 120
> x1=c(30,40,20,50,60,40,20,60)
> x1
[1] 30 40 20 50 60 40 20 60
> x2=c(11,10,7,15,19,12,8,14)
> x2
[1] 11 10 7 15 19 12 8 14
> regmodel=lm(Y1-x1+x2)
Error in stats::model.frame(formula = Y1 - x1 + x2, drop.unused.levels = TRUE) :
  object 'Y1' not found
> regmodel

Call:
lm(formula = y ~ x1 + x2)

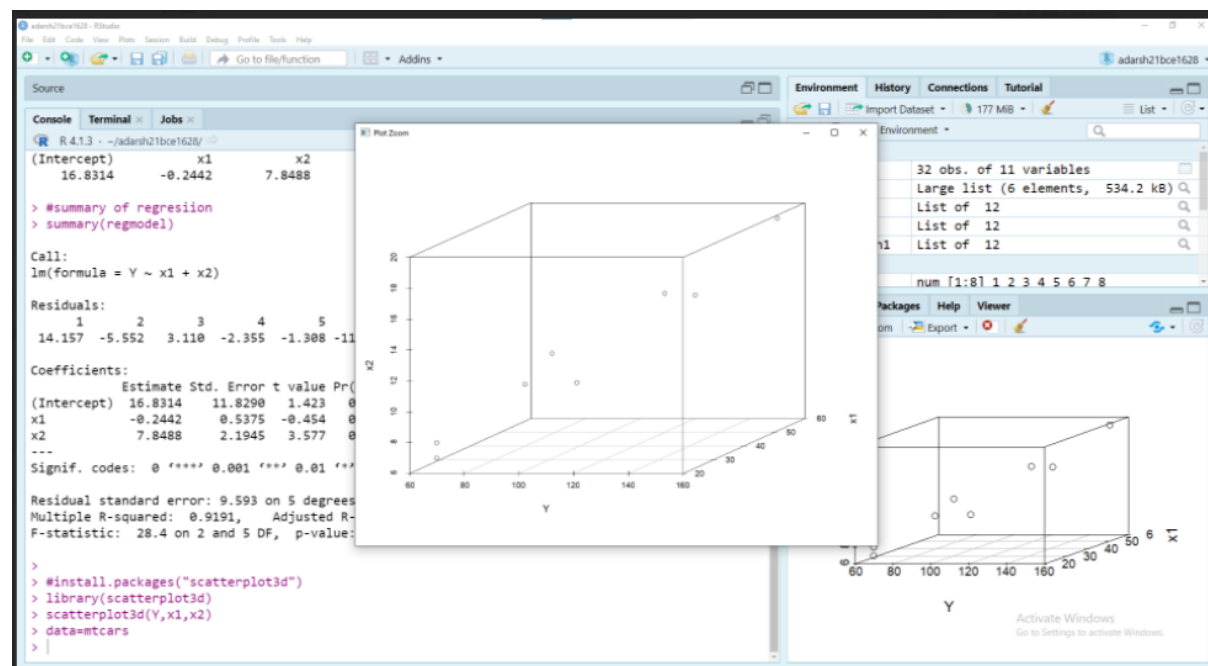
Coefficients:
(Intercept)          x1          x2
    16.8314     -0.2442     7.8488

> regmodel=lm(Y-x1+x2)
Error in formula.default(object, env = baseenv()) : invalid formula
> regmodel

Call:
lm(formula = y ~ x1 + x2)

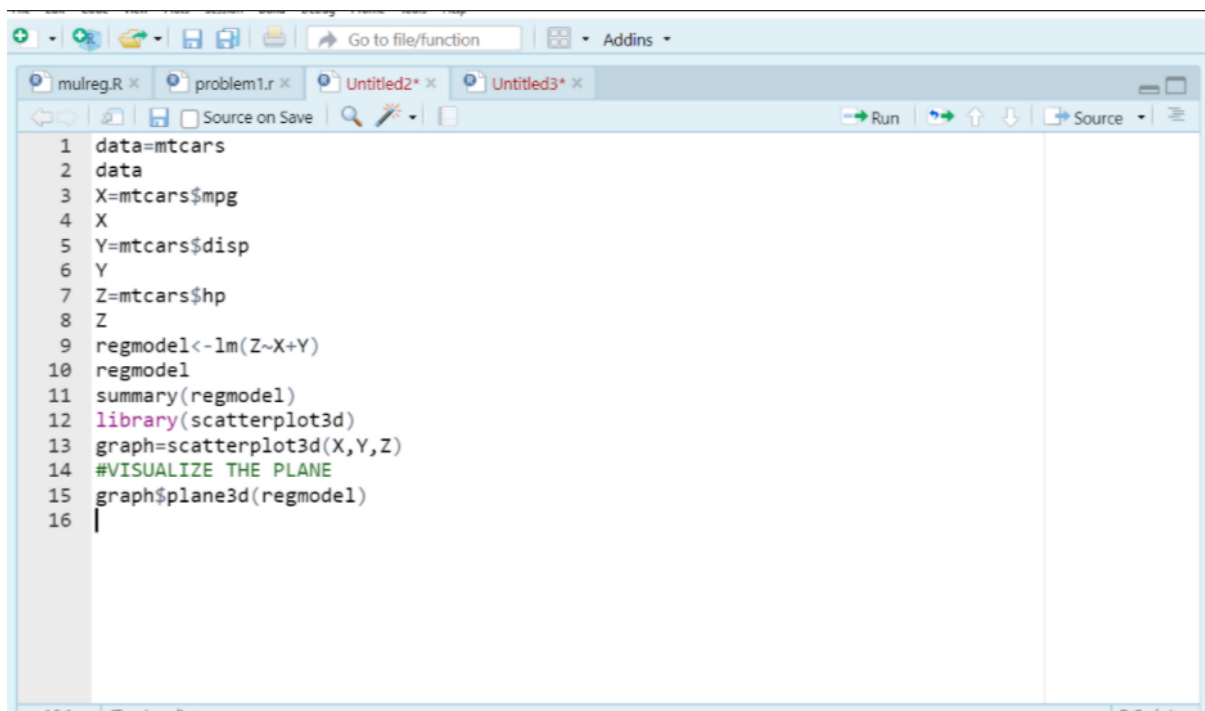
Coefficients:

```



CODE

```
data=mtcars
data
X=mtcars$mpg
X
Y=mtcars$displ
Y
Z=mtcars$hp
Z
regmodel<-lm(Z~X+Y)
regmodel
summary(regmodel)
library(scatterplot3d)
graph=scatterplot3d(X,Y,Z)
#VISUALIZE THE PLANE
graph$plane3d(regmodel)
```



```
1 data=mtcars
2 data
3 X=mtcars$mpg
4 X
5 Y=mtcars$displ
6 Y
7 Z=mtcars$hp
8 Z
9 regmodel<-lm(Z~X+Y)
10 regmodel
11 summary(regmodel)
12 library(scatterplot3d)
13 graph=scatterplot3d(X,Y,Z)
14 #VISUALIZE THE PLANE
15 graph$plane3d(regmodel)
16
```

| | mpg | cyl | disp | hp | drat | wt | qsec | vs | am | gear | carb |
|---------------------|------|-----|-------|-----|------|-------|-------|----|----|------|------|
| Mazda RX4 | 21.0 | 6 | 160.0 | 110 | 3.90 | 2.620 | 16.46 | 0 | 1 | 4 | 4 |
| Mazda RX4 Wag | 21.0 | 6 | 160.0 | 110 | 3.90 | 2.875 | 17.02 | 0 | 1 | 4 | 4 |
| Datsun 710 | 22.8 | 4 | 108.0 | 93 | 3.85 | 2.320 | 18.61 | 1 | 1 | 4 | 1 |
| Hornet 4 Drive | 21.4 | 6 | 258.0 | 110 | 3.08 | 3.215 | 19.44 | 1 | 0 | 3 | 1 |
| Hornet Sportabout | 18.7 | 8 | 360.0 | 175 | 3.15 | 3.440 | 17.02 | 0 | 0 | 3 | 2 |
| Valiant | 18.1 | 6 | 225.0 | 105 | 2.76 | 3.460 | 20.22 | 1 | 0 | 3 | 1 |
| Duster 360 | 14.3 | 8 | 360.0 | 245 | 3.21 | 3.570 | 15.84 | 0 | 0 | 3 | 4 |
| Merc 240D | 24.4 | 4 | 146.7 | 62 | 3.69 | 3.190 | 20.00 | 1 | 0 | 4 | 2 |
| Merc 230 | 22.8 | 4 | 140.8 | 95 | 3.92 | 3.150 | 22.90 | 1 | 0 | 4 | 2 |
| Merc 280 | 19.2 | 6 | 167.6 | 123 | 3.92 | 3.440 | 18.30 | 1 | 0 | 4 | 4 |
| Merc 280C | 17.8 | 6 | 167.6 | 123 | 3.92 | 3.440 | 18.90 | 1 | 0 | 4 | 4 |
| Merc 450SE | 16.4 | 8 | 275.8 | 180 | 3.07 | 4.070 | 17.40 | 0 | 0 | 3 | 3 |
| Merc 450SL | 17.3 | 8 | 275.8 | 180 | 3.07 | 3.730 | 17.60 | 0 | 0 | 3 | 3 |
| Merc 450SLC | 15.2 | 8 | 275.8 | 180 | 3.07 | 3.780 | 18.00 | 0 | 0 | 3 | 3 |
| Cadillac Fleetwood | 10.4 | 8 | 472.0 | 205 | 2.93 | 5.250 | 17.98 | 0 | 0 | 3 | 4 |
| Lincoln Continental | 10.4 | 8 | 460.0 | 215 | 3.00 | 5.424 | 17.82 | 0 | 0 | 3 | 4 |
| Chrysler Imperial | 14.7 | 8 | 440.0 | 230 | 3.23 | 5.345 | 17.42 | 0 | 0 | 3 | 4 |
| Fiat 128 | 32.4 | 4 | 78.7 | 66 | 4.08 | 2.200 | 19.47 | 1 | 1 | 4 | 1 |
| Honda Civic | 30.4 | 4 | 75.7 | 52 | 4.93 | 1.615 | 18.52 | 1 | 1 | 4 | 2 |
| Toyota Corolla | 33.9 | 4 | 71.1 | 65 | 4.22 | 1.835 | 19.90 | 1 | 1 | 4 | 1 |
| Toyota Corona | 21.5 | 4 | 120.1 | 97 | 3.70 | 2.465 | 20.01 | 1 | 0 | 3 | 1 |
| Dodge Challenger | 15.5 | 8 | 318.0 | 150 | 2.76 | 3.520 | 16.87 | 0 | 0 | 3 | 2 |
| AMC Javelin | 15.2 | 8 | 304.0 | 150 | 3.15 | 3.435 | 17.30 | 0 | 0 | 3 | 2 |
| Camaro Z28 | 13.3 | 8 | 350.0 | 245 | 3.73 | 3.840 | 15.41 | 0 | 0 | 3 | 4 |
| Pontiac Firebird | 19.2 | 8 | 400.0 | 175 | 3.08 | 3.845 | 17.05 | 0 | 0 | 3 | 2 |
| Fiat X1-9 | 27.3 | 4 | 79.0 | 66 | 4.08 | 1.935 | 18.90 | 1 | 1 | 4 | 1 |
| Porsche 914-2 | 26.0 | 4 | 120.3 | 91 | 4.43 | 2.140 | 16.70 | 0 | 1 | 5 | 2 |
| Lotus Europa | 30.4 | 4 | 95.1 | 113 | 3.77 | 1.513 | 16.90 | 1 | 1 | 5 | 2 |
| Ford Pantera L | 15.8 | 8 | 351.0 | 264 | 4.22 | 3.170 | 14.50 | 0 | 1 | 5 | 4 |
| Ferrari Dino | 19.7 | 6 | 145.0 | 175 | 3.62 | 2.770 | 15.50 | 0 | 1 | 5 | 6 |
| Maserati Bora | 15.0 | 8 | 301.0 | 335 | 3.54 | 3.570 | 14.60 | 0 | 1 | 5 | 8 |

```

R 4.1.3 ~ jadashiz162828y
Lotus Europa      30.4  4  95.1 113 3.77 1.513 16.90  1  1  5  2
Ford Pantera L    15.8  8 351.0 264 4.22 3.170 14.50  0  1  5  4
Ferrari Dino      19.7  6 145.0 175 3.62 2.770 15.50  0  1  5  6
Maserati Bora     15.0  8 301.0 335 3.54 3.570 14.60  0  1  5  8
Volvo 142E        21.4  4 121.0 109 4.11 2.780 18.60  1  1  4  2
> X=mtcars$mpg
> X
[1] 21.0 21.0 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 17.8 16.4 17.3 15.2 10.4 10.4 14.7 32.4
[19] 30.4 33.9 21.5 15.5 15.2 13.3 19.2 27.3 26.0 30.4 15.8 19.7 15.0 21.4
> Y=mtcars$disp
> Y
[1] 160.0 160.0 108.0 258.0 360.0 225.0 360.0 146.7 140.8 167.6 167.6 275.8 275.8 275.8 472.0
[16] 460.0 440.0 78.7 75.7 71.1 120.1 318.0 304.0 350.0 400.0 79.0 120.3 95.1 351.0 145.0
[31] 301.0 121.0
> Z=mtcars$hp
> Z
[1] 110 110 93 110 175 105 245 62 95 123 123 180 180 180 205 215 230 66 52 65 97 150
[23] 150 245 175 66 91 113 264 175 335 109
> regmodel<-lm(Z~X+Y)
> regmodel

Call:
lm(formula = Z ~ X + Y)

Coefficients:
(Intercept)          X          Y
    172.2204    -4.2732     0.2614

> summary(regmodel)

```

```

Call:
lm(formula = Z ~ X + Y)

```

```

Console Terminal Jobs
R 4.1.3 - ~/adarsh21bce1628/

Coefficients:
(Intercept)          X              Y
    172.2204     -4.2732      0.2614

> summary(regmodel)

Call:
lm(formula = Z ~ X + Y)

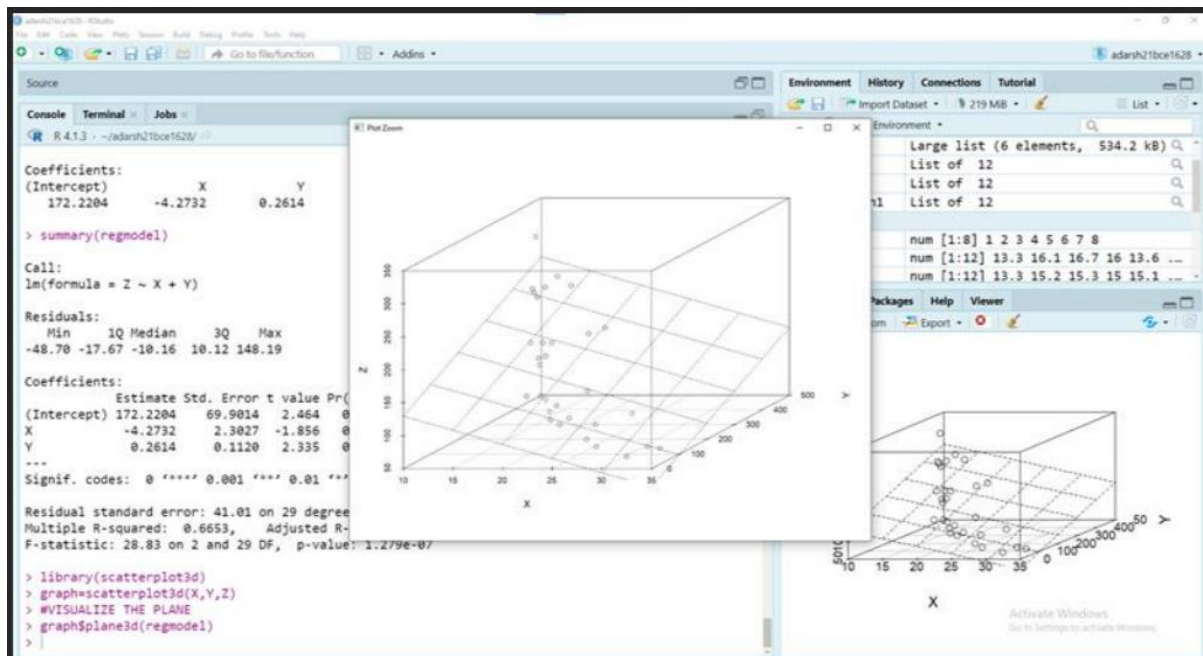
Residuals:
    Min       1Q   Median       3Q      Max
-48.70 -17.67 -10.16   10.12  148.19

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  172.2204    69.9014   2.464  0.0199 +
X            -4.2732     2.3027  -1.856  0.0737 .
Y              0.2614     0.1120   2.335  0.0267 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 41.01 on 29 degrees of freedom
Multiple R-squared:  0.6653,    Adjusted R-squared:  0.6423
F-statistic: 28.83 on 2 and 29 DF,  p-value: 1.279e-07

> library(scatterplot3d)
> graph=scatterplot3d(X,Y,Z)
> #VISUALIZE THE PLANE
> graph$plane3d(regmodel)
>

```



LAB – 5

Fitting the probability distributions: Binomial distribution

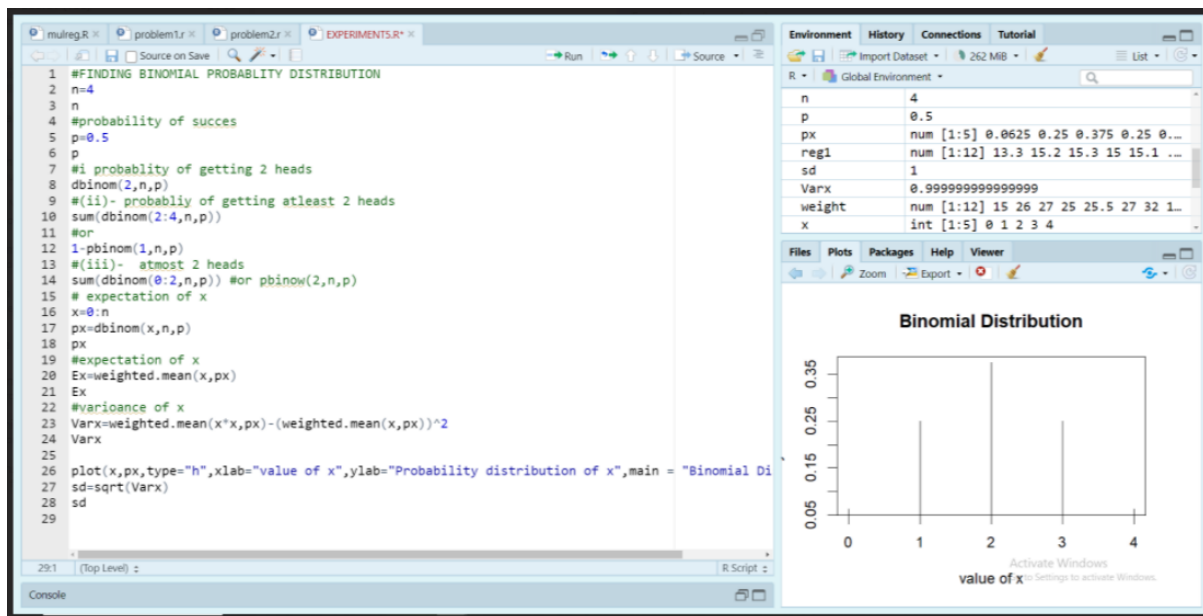
CODE

```
#FINDING BINOMIAL PROBABILITY DISTRIBUTION
n=4
n
#probability of succes
p=0.5
p
#i probability of getting 2 heads
dbinom(2,n,p)
#(ii)- probabliiy of getting atleast 2 heads
sum(dbinom(2:4,n,p))
#or
1-pbinom(1,n,p)
#(iii)- atmost 2 heads
sum(dbinom(0:2,n,p)) #or pbinow(2,n,p)
# expectation of x
x=0:n
px=dbinom(x,n,p)
px
#expectation of x
Ex=weighted.mean(x,px)
Ex
#varioance of x
```

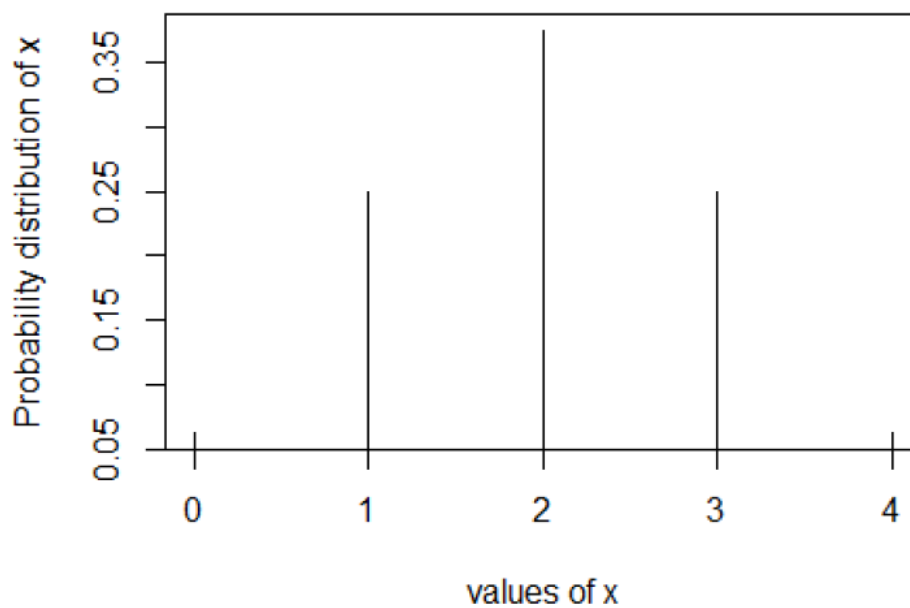
$$\text{Var}x = \text{weighted.mean}(x * x, px) - (\text{weighted.mean}(x, px))^2$$

$$\text{Var}x$$

$$\text{plot}(x, px, \text{type} = "h", \text{xlab} = "value of x", \text{ylab} = "Probability distribution of x", \text{main} = "Binomial Distribution")$$



Binomial distribution



ALL EXPERIMENT R CODE

#Experiment-1: Understanding data types, importing and exporting data

#Generate data

#Class Code

#Generate data

1:10

#Assign variable name to the value

X = 10

Combine numeric values into vector

c(1,2,5)

#Arithmetic operations on vectors

a = c(1, 3, 5, 7)

b = c(2, 4, 6, 8)

#Addition

a+b

#Subtraction

a-b

#Constant multiplication

5*a

#Product

a*b

#Division

a/b

#String values

X = as.character(5.2)

X

#Concatenation of strings

paste("Baa", "Baa", "Black", "Sheep")

#Practice code

#Generate data

1:10

#Assign variable name to the value

X <- 5

Combine numeric values into vector

c(12,43,14,54,100)

c

#Arithmetic operations on vectors

a = c(12, 43, 11, 14, 9, 10)

b = c(59, 110, 2, 19, 6, 39)

#Addition

a+b

#Subtraction

a-b

#Constant multiplication

10*a

#Product

a*b

#Division

a/b

#String values

X = as.character(12.9)

X

#Concatenation of strings

paste("Hello", "World")

#-----

Experiment-2: Plotting and visualizing data

#Class Question

#Put data into vectors

empid = c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15)

empid

age = c(30,37,45,32,50,60,35,32,34,43,32,30,43,50,60)

age

gender = c(0,1,0,1,1,1,0,0,1,0,0,1,1,0,0)

gender

status = c(1,1,2,2,1,1,1,2,2,1,2,1,2,1,2)

status

#Create dataframe

```
empinfo = data.frame(empid, age, gender, status)
```

empinfo

labeling character to numeric

```
empinfo$gender = factor(empinfo$gender, labels = c("male", "female"))
```

empinfo\$gender

```
empinfo$status = factor(empinfo$status, labels = c("staff", "faculty"))
```

empinfo\$status

empinfo

extract male data

```
male = subset(empinfo, empinfo$gender == "male")
```

male

extract female data

```
female = subset(empinfo, empinfo$gender == "female")
```

female

summary statistics for empinfo data

```
summary(empinfo)
```

```
summary(male)
```

```
summary(female)
```

```
summary(age)
```

#creating table (one-way)

```
table1 = table(empinfo$gender)
```

table1

```
table2 = table(empinfo$status)
```

table2

#creating table (two-way)

table3 = table(empinfo\$gender, empinfo\$status)

table3

#Graphical representation

```
plot(empinfo$age, type="l", main="Age of employees", xlab="empid", ylab="Age in years", col =  
"blue")  
pie(table1)  
barplot(table3, beside=T, xlim=c(1,15), ylim=c(0,5), col=c("blue", "red"))  
legend("topright", legend=rownames(table3), fill=c('blue', 'red'), bty="n")  
boxplot(empinfo$age~empinfo$status, col=c('red', 'blue'))
```

pie(table1)

```
barplot(table3, beside=T, xlim=c(1,15), ylim=c(0,5), col=c("blue", "red"))  
legend("topright", legend=rownames(table3), fill=c('blue', 'red'), bty="n")
```

```
boxplot(empinfo$age~empinfo$status, col=c('red', 'blue'))
```

#Practice question

#Put data into dataframe

name = c('p1', 'p2', 'p3', 'p4', 'p5', 'p6', 'p7', 'p8', 'p9', 'p10')

age = c(20, 20, 21, 20, 21, 22, 21, 21, 20, 22)

math = c(100, 90, 95, 98, 92, 99, 95, 100, 93, 90)

phy = c(90, 100, 91, 92, 98, 92, 98, 99, 100, 93)

chem = c(93, 89, 99, 92, 94, 95, 99, 85,)

Create dataframe for student marks

studMarks = data.frame(name, age, math, phy, chem)

studMarks

summary(studMarks)

Plot math marks of students

```
plot(studMarks$math, type="l", main="Marks in math", xlab="Student", ylab="Marks", col = "blue")
```

```
#creating table (one-way)
```

```
tableAge = table(studMarks$age)
```

```
tableAge
```

```
#Pie chart for age
```

```
pie(tableAge)
```

```
tablePhyMarks = table(studMarks$phy)
```

```
tablePhyMarks
```

```
#Bar plot for physics marks
```

```
barplot(tablePhyMarks, beside=T, xlim=c(0,10), ylim=c(0,5), col=c("blue", "red"))
```

```
#Box plot for math marks and age
```

```
boxplot(studMarks$math~studMarks$age, col=c('red', 'blue'))
```

```
#-----
```

```
#Experiment-3: Apply correlation and simple linear regression model to real data
```

```
#Class question
```

```
#Problem-1
```

```
#Import the inbuilt dataset cars
```

```
data=cars
```

```
data
```

```
#Summary of data set
```

```
summary(data)
```

```
#Variance of speed
```

```
v1=var(data$speed)
```

```
v1
```



```
#variance of "dist"
```

```
v2=var(data$dist)
```

```
v2
```

```
#covariance between speed and dist
```

```
covariance=cov(data$speed,data$dist)
```

```
covariance
```

```
#Standard deviation
```

```
stdDevSpeed=sd(data$speed)
```

```
stdDevSpeed
```

```
stdDevDist = sd(data$dist)
```

```
stdDevDist
```

```
#correlation coefficient
```

```
corr = covariance/(stdDevSpeed*stdDevDist)
```

```
corr
```

```
#or
```

```
corr2 = cor(data$speed, data$dist)
```

```
corr2
```

```
cor.test(data$speed, data$dist)
```

```
cor.test(data$speed, data$dist, method="pearson")
```

```
cor.test(data$speed, data$dist, method="spearman")
```

```
regression1 = lm(data$speed~data$dist)
```

```
regression1
```

```
#Plot regression2
```

```
plot(data$speed, data$dist)
```

```
abline(regression1)
```

```
regression2 = lm(data$dist~data$speed)
```

```
regression2
```

```
#Plot regression1
```

```
plot(data$dist, data$speed)
```

```
abline(regression2)
```

```
#Problem-2
```

```
weight = c(15, 26, 27, 25, 25.5, 27, 32, 18, 22, 20, 26, 24)
```

```
weight
```

```
bmi = c(13.35, 16.12, 16.74, 16, 13.59, 15.73, 15.65, 13.85, 16.07, 12.8, 13.65, 14.42);
```

```
bmi
```

```
#Apply correlation
```

```
cor(weight, bmi)
```

```
#Find regression
```

```
Regression1 = lm(bmi~weight)
```

```
Regression1
```

```
reg1 = lm(bmi~weight)
```

```
reg1 = 10.735+0.1710*weight
```

```
reg1
```

```
#Plot regression
```

```
plot(weight, bmi)
```

```
abline(Regression1)
```

```
#Practice problem
```

```
#Vector for age of husbands
```

```
ageHusbands = c(23,27,28,29,30,31,33,35,36,39)
```

```
#Vector for age of wives
```

```
ageWives = c(18,22,23,24,25,26,28,29,30,32)
```

```
#Create dataframe
```

```
data = data.frame(ageHusbands, ageWives)
```

```
#Summary of data set
```

```
summary(data)
```

```
#Variance of husbands' age
```

```
v1=var(data$ageHusbands)
```

```
v1
```

```
#variance of wives' age
```

```
v2=var(data$ageWives)
```

```
v2
```

```
#covariance between husbands' age and wives' age
```

```
covariance=cov(data$ageHusbands,data$ageWives)
```

```
covariance
```

```
#Standard deviation
```

```
stdDevHusbAge=sd(data$ageHusbands)
```

```
stdDevHusbAge
```

```
stdDevWivesAge = sd(data$ageWives)
```

stdDevWivesAge

#correlation coefficient

corr = covariance/(stdDevHusbAge*stdDevWivesAge)

corr

#or

corr2 = cor(data\$ageHusbands, data\$ageWives)

corr2 = cor(data)

cor.test(data\$ageHusbands, data\$ageWives)

cor.test(data\$ageHusbands, data\$ageWives, method="pearson")

cor.test(data\$ageHusbands, data\$ageWives, method="spearman")

regression1 = lm(data\$ageHusbands~data\$ageWives)

regression1

#Plot regression1

plot(data\$ageHusbands, data\$ageWives)

abline(regression1)

regression2 = lm(data\$ageWives~data\$ageHusbands)

regression2

#Plot regression2

plot(data\$ageWives, data\$ageHusbands)

abline(regression2)

#-----

#Experiment-4: Apply multiple linear regression model to real data set

#Class question

#Problem-1

Input the variables

Y = c(110, 80, 70, 120, 150, 90, 70, 120)

Y

X1 = c(30, 40, 20, 50, 60, 40, 20, 60)

X1

X2 = c(11, 10, 7, 15, 19, 12, 8, 14)

X2

Linear regression model of Y on X1 and X2\

RegModel = lm(Y~X1+X2) RegModel

Summary of the data

summary(RegModel)

install.packages("scatterplot3d")

library(scatterplot3d)

plot the data set

scatterplot3d(Y,X1,X2)

#Problem-2

data = mtcars

data

X = mtcars\$mpg

X

Y = mtcars\$disp

Y

Z = mtcars\$hp

Z

```
RegModel <- lm(Z~X+Y)
```

```
RegModel
```

```
summary(RegModel)
```

```
graph = scatterplot3d(X,Y,Z)
```

```
# Visualize the plane
```

```
graph$plane3d(RegModel)
```

```
#Practice question
```

```
Y = c(240,236,290,274,301,316,300,296,267,276,288,261)
```

```
X1 = c(25,31,45,60,65,72,80,84,75,60,50,38)
```

```
X2 = c(24,21,24,25,25,26,25,25,24,25,25,25)
```

```
# Linear regression model of Y on X1 and X2
```

```
RegModel = lm(Y~X1+X2)
```

```
RegModel
```

```
# Summary of the data
```

```
summary(RegModel)
```

```
library(scatterplot3d)
```

```
# plot the data set scatterplot3d(Y,X1,X2)
```

```
graph$plane3d(RegModel)
```

```
#-----
```

```
#Experiment-5: Binomial Distribution
```

```
#Class Question
```

```
#Four coins are tossed simultaneously. What is the probability of getting
```

```
 #(i) 2 heads
```

```
 #(ii) atleast 2 heads
```

```
 #(iii) atmost 2 heads
```

#(iv) Expectation of x

#(v) Variance of x

#(vi) Visualize the probability distribution

number of trials

n = 4

n

probability of success

p = 0.5

p

(i) probability of getting exactly two heads

dbinom(2,n,p)

(ii) probability of getting at least two heads

sum(dbinom(2:4, n, p))

or

1 - pbinom(1, n, p)

(iii) probability of getting at most two heads

sum(dbinom(0:2,n,p))

(iv) Expectation of x

x = 0:n

px = dbinom(x, n, p)

px

Ex = weighted.mean(x, px)

Ex

(v) variance of x

```
varx = weighted.mean(x*x, px) - (weighted.mean(x, px))^2
```

```
varx
```

```
std <- sqrt(varx)
```

```
std
```

```
# (vi) visualize probability distribution
```

```
plot(x, px, type = "h", xlab = "values of x", ylab = "Probability distribution of x", main =  
      "Binomial distribution")
```

```
#Practice question
```

```
#The incidence of an occupational disease in an industry is such that the workers have a 20% chance  
#of suffering from it. What is the probability that out of 6 workers chosen at random 4 or more will  
suffer  
#disease?
```

```
# number of trials
```

```
n = 6
```

```
n
```

```
# probability of success
```

```
p = 0.2
```

```
p
```

```
sum(dbinom(4:6, n, p))
```

```
x = 0:n
```

```
px = dbinom(x, n, p)
```

```
px
```

```
# visualize probability distribution
```

```
plot(x, px, type = "h", xlab = "values of x", ylab = "Probability distribution of x", main =  
      "Binomial distribution")
```
