Vidyavardhini's College of Engineering and Technology

Department of Artificial Intelligence & Data Science

AY: 2023-24

Class:	TE	**Semester:**	VI
Course Code:	CSL601	**Course Name:**	Data Analytics and Visualization

Name of Student:	Parth Manoj Raut
Roll No.:	44
Experiment No.:	3
Title of the Experiment:	Implement R Program for Multiple Linear Regression.
Date of Performance:	
Date of Submission:	

Evaluation

Performance Indicator	Max. Marks	Marks Obtained
Performance	5	
Understanding	5	
Journal work and timely submission	10	
Total	20	

Performance Indicator	Exceed Expectations (EE)	Meet Expectations (ME)	Below Expectations (BE)
Performance	4-5	2-3	1
Understanding	4-5	2-3	1
Journal work and timely submission	8-10	5-8	1-4

Checked by

Name of Faculty : Ms Bhavika Gharat

Signature :

Date :

Experiment No - 3

Aim :- Implement R Program for Multiple Linear Regression.

Objective:- To understand the use of Multiple linear regression techniques by implementing a predefined dataset of R Studio.

Description-

Multiple linear regression is the extension of linear regression in the relationship between more than two variables. In simple linear regression, we have one predictor and one response variable. But in multiple regressions, we have more than one predictor variable and one response variable.

There is the following general mathematical equation for multiple regression -

$$y=b_0+b_1*x_1+b_2*x_2+b_3*x_3+\cdots b_n*x_n$$

Here,

- y is a response variable.
- **bO, bl, b2...bn** are the coefficients.
- **xl, x2, ...xn** are the predictor variables.

In R, we create the regression model with the help of the **Im()** function. The model will determine the value of the coefficients with the help of the input data. We can predict the value of the response variable for the set of predictor variables using these coefficients.

There is the following syntax of lm() function in multiple regression

1. lm(y x l+x2+x3 ...., data)

Before proceeding further, we first create our data for multiple regression. We will use the "mtcars" dataset present in the R environment. The main task of the model is to create the relationship between the "mpg" as a response variable with "wt", "disp" and "hp" as predictor variables.

For this purpose, we will create a subset of these variables from the "mtcars" dataset.

1. data<-mtcars[,c("mpg","wt","disp","hp")]
2. print(head(input)

Program-

```
input <- mtcars[,c("mpg" ,"disp","hp","wt")]
#mtcars file is already present in R
print(head(input))
input <- mtcars[,c("mpg" ,"disp","hp","wt")]

# Create the relationship model.
model <- lm(mpg  disp+hp+wt, data = input)

# Show the model.
print(model)

# Get the Intercept and coefficients as vector elements.
cat("# # # # The Coefficient Values # # # ","\n")

a <- coef(model)[l ]
print( a)

Xdisp <-
coef(model)[2] Xhp <-
coef(model)[3] Xwt <-
coef(model)[4]

print(Xdisp)
print(Xhp)
print(Xwt)
```

Output-

```
> input<-mtcars[,c("mpg","disp","hp","wt")]
> print(head(input))
                   mpg disp  hp    wt
Mazda RX4         21.0  160 110 2.620
Mazda RX4 Wag     21.0  160 110 2.875
Datsun 710        22.8  108  93 2.320
Hornet 4 Drive    21.4  258 110 3.215
Hornet Sportabout 18.7  360 175 3.440
Valiant           18.1  225 105 3.460
> input<-mtcars[,c("mpg","disp","hp","wt")]
> model<-lm(mpg~disp+hp+wt, data=input)
> print(model)

Call:
lm(formula = mpg ~ disp + hp + wt, data = input)

Coefficients:
(Intercept)         disp           hp           wt
  37.105505    -0.000937    -0.031157    -3.800891

 > cat("#### The Coefficient Values ###-","/n")
 #### The Coefficient Values ###- /n
 > a<-coef(model)[1]
 > print(a)
 (Intercept)
    37.10551
 > Xdisp<-coef(model)[2]
 > Xhp<-coef(model)[3]
 > Xwt<-coef(model)[4]
 > print(Xdisp)
         disp
 -0.0009370091
 > print(Xhp)
          hp
 -0.03115655
 > print(Xwt)
        wt
 -3.800891
```

Environment:

Data	
input	32 obs. of 4 variables
model	List of 12
values	
a	Named num 37.1
Xdisp	Named num -0.000937
Xhp	Named num -0.0312
Xwt	Named num -3.8

Conclusion-

1. Equation for multiple linear regression is
   $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n + \varepsilon$

2. When there is only one dependent variable and multiple independent variable then this types of regression is known as
   Multiple Linear Regression

3. How to check inbuilt dataset in R studio?
   data(package = "datasets")