Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

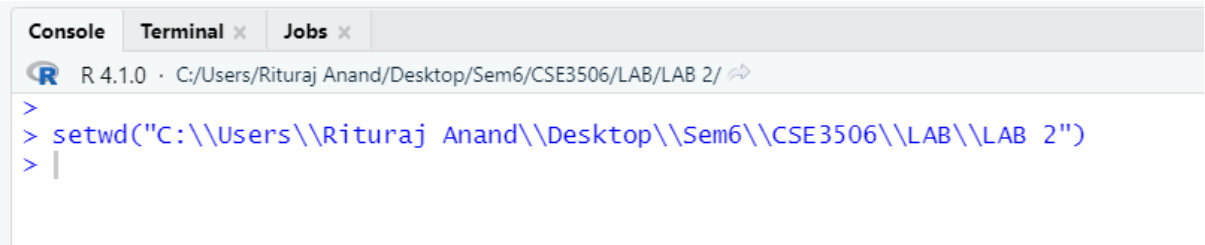
Programme	:	**B.Tech – ECE and ECM**	Semester	:	**Win 2022**
Course	:	**Essentials of Data Analytics Lab**	Code	:	**CSE3506**
Faculty	:	**Gobinath N**	Slot	:	**L51 + L52**

Multiple Linear Regression_Ex.01

Basic Commands:

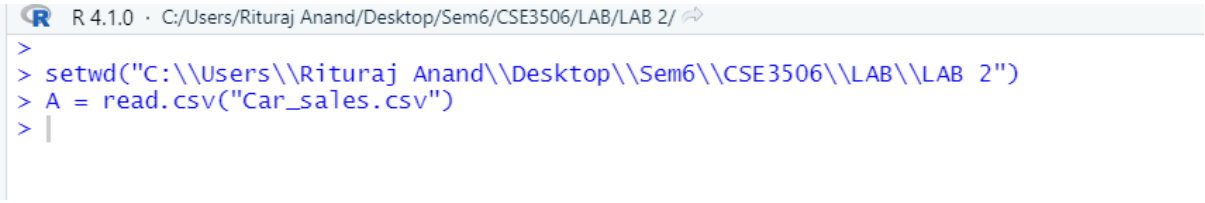
#Setting the working directory:

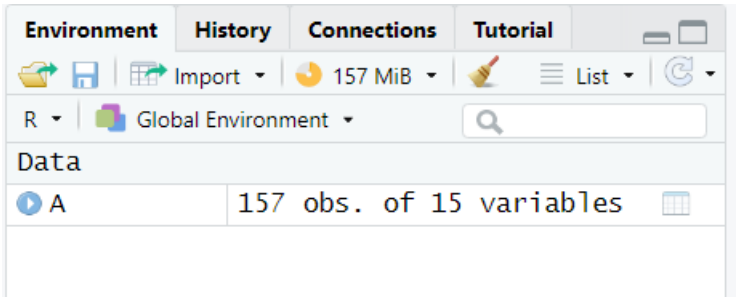
```
setwd("C:\\Users\\Rituraj Anand\\Desktop\\Sem6\\CSE3506\\LAB\\LAB 2")
```

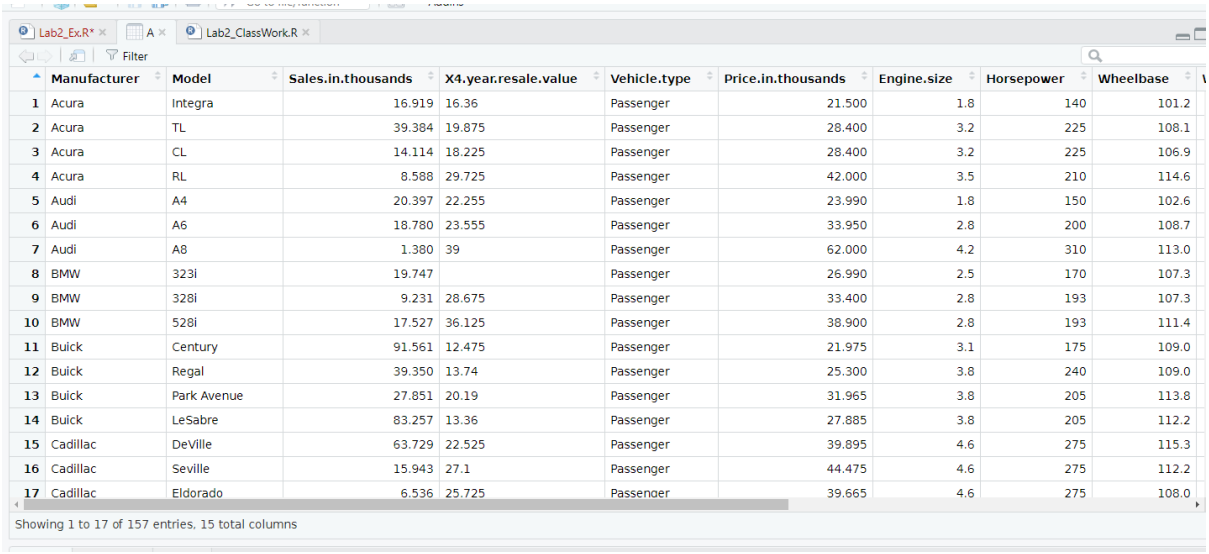


#Reading the csv File:

```
A = read.csv("Car_sales.csv")
```







Lab2_Ex.R* | A | Lab2_ClassWork.R

Filter

	Manufacturer	Model	Sales.in.thousands	X4.year.resale.value	Vehicle.type	Price.in.thousands	Engine.size	Horsepower	Wheelbase
1	Acura	Integra	16.919	16.36	Passenger	21.500	1.8	140	101.2
2	Acura	TL	39.384	19.875	Passenger	28.400	3.2	225	108.1
3	Acura	CL	14.114	18.225	Passenger	28.400	3.2	225	106.9
4	Acura	RL	8.588	29.725	Passenger	42.000	3.5	210	114.6
5	Audi	A4	20.397	22.255	Passenger	23.990	1.8	150	102.6
6	Audi	A6	18.780	23.555	Passenger	33.950	2.8	200	108.7
7	Audi	A8	1.380	39	Passenger	62.000	4.2	310	113.0
8	BMW	323i	19.747		Passenger	26.990	2.5	170	107.3
9	BMW	328i	9.231	28.675	Passenger	33.400	2.8	193	107.3
10	BMW	528i	17.527	36.125	Passenger	38.900	2.8	193	111.4
11	Buick	Century	91.561	12.475	Passenger	21.975	3.1	175	109.0
12	Buick	Regal	39.350	13.74	Passenger	25.300	3.8	240	109.0
13	Buick	Park Avenue	27.851	20.19	Passenger	31.965	3.8	205	113.8
14	Buick	LeSabre	83.257	13.36	Passenger	27.885	3.8	205	112.2
15	Cadillac	DeVille	63.729	22.525	Passenger	39.895	4.6	275	115.3
16	Cadillac	Seville	15.943	27.1	Passenger	44.475	4.6	275	112.2
17	Cadillac	Eldorado	6.536	25.725	Passenger	39.665	4.6	275	108.0

Showing 1 to 17 of 157 entries, 15 total columns

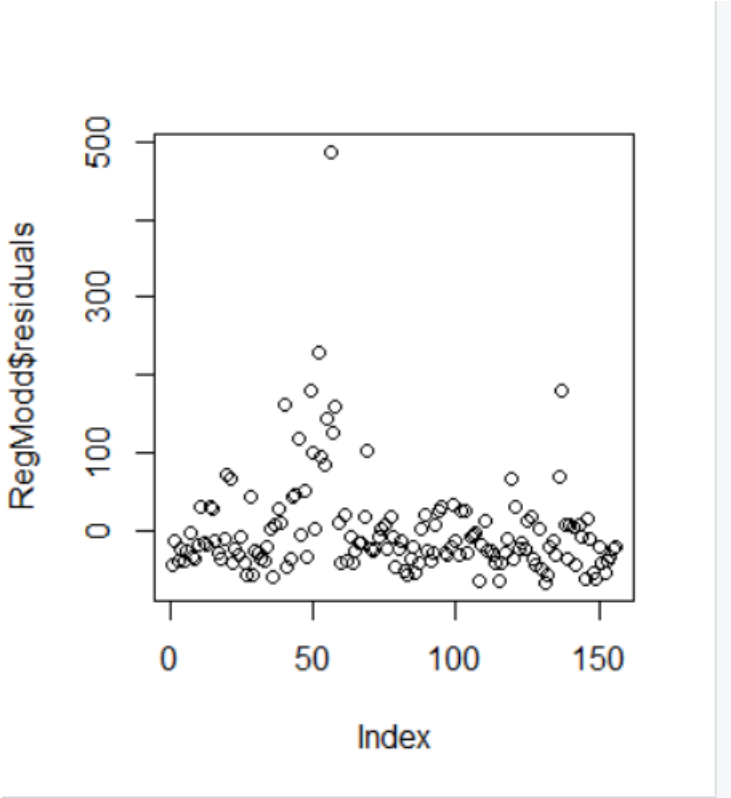
#Checking Linear Regression between Sales and Price of cars from the Dataset

```
RegModd=lm(Sales.in.thousands~Price.in.thousands,A)  #Y_Var~X_Var

summary(RegModd)

attributes(RegModd)
```

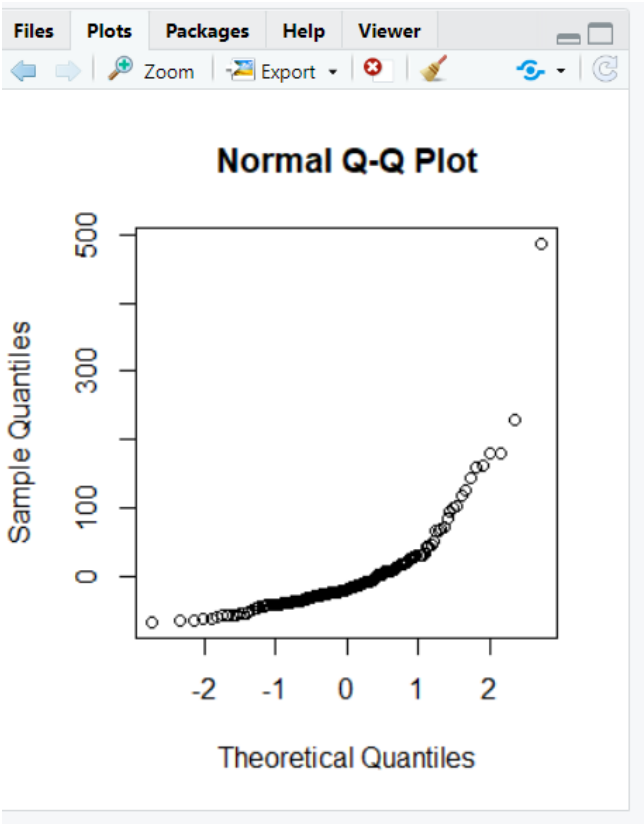
```
R 4.1.0 · C:/Users/Rituraj Anand/Desktop/Sem6/CSE3506/LAB/LAB 2/
>
> setwd("C:\\Users\\Rituraj Anand\\Desktop\\Sem6\\CSE3506\\LAB\\LAB 2")
> A = read.csv("Car_sales.csv")
> View(A)
> #Linear Regression
> RegModd=lm(Sales.in.thousands~Price.in.thousands,A)  #Y_Var~X_Var
> View(RegModd)
> summary(RegModd)

Call:
lm(formula = Sales.in.thousands ~ Price.in.thousands, data = A)

Residuals:
   Min     1Q Median     3Q    Max
-66.82 -35.09 -19.43  10.61 486.89

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)        92.8170    11.3093   8.207 8.45e-14 ***
Price.in.thousands -1.4535     0.3662  -3.970  0.00011 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 65.21 on 154 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.09282,   Adjusted R-squared:  0.08693
F-statistic: 15.76 on 1 and 154 DF,  p-value: 0.0001102
```

Plotting using qqnorm:

```
plot(RegModd$residuals)
```




```
qqnorm(RegModd$residuals)
```



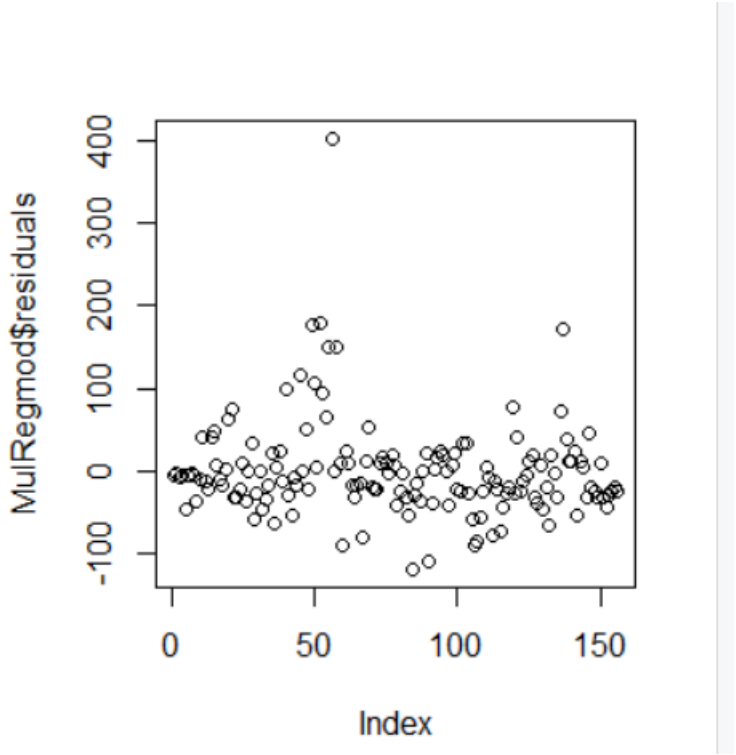
#Checking Multiple Regression b/w Sales, Price and Fuel Efficieny

```
MulRegmod=lm(Sales.in.thousands~Price.in.thousands+Fuel.efficiency,A)

summary(MulRegmod)

attributes(MulRegmod)
```

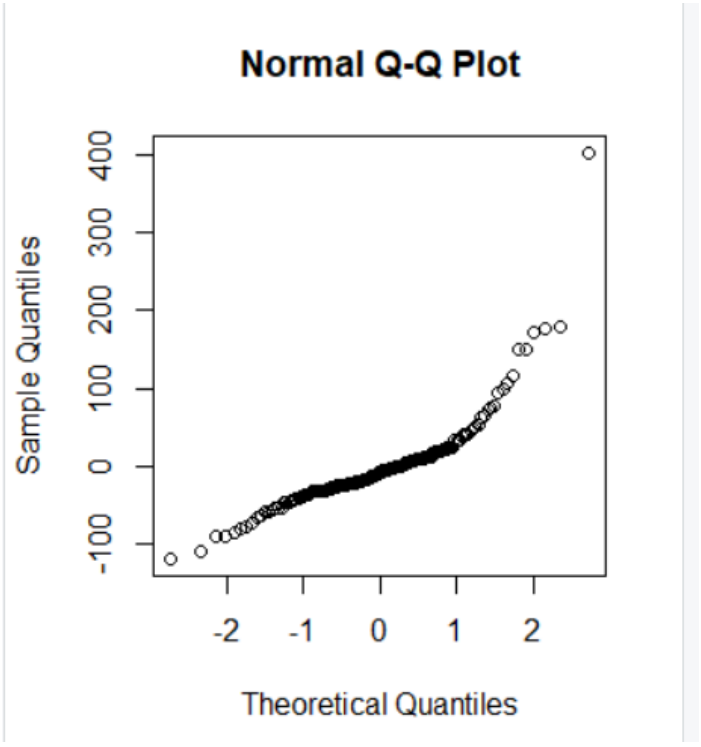
```
Call:
lm(formula = Sales.in.thousands ~ Price.in.thousands + Fuel.efficiency,
    data = A)

Residuals:
    Min      1Q  Median      3Q     Max
-119.34  -29.69   -7.48   12.61  403.00

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)         96.2528    46.1793   2.084  0.03903 *
Price.in.thousands  -1.4130     0.4694  -3.010  0.00312 **
Fuel.efficiency15  -18.8186    53.9961  -0.349  0.72800
Fuel.efficiency16   14.7868    58.5403   0.253  0.80097
Fuel.efficiency17   60.0959    57.9371   1.037  0.30148
Fuel.efficiency18   79.3689    53.0907   1.495  0.13727
Fuel.efficiency19   45.6292    51.8347   0.880  0.38028
Fuel.efficiency20    4.3330    54.4859   0.080  0.93673
Fuel.efficiency21   -5.5191    48.5586  -0.114  0.90968
Fuel.efficiency22  -23.6298    48.3969  -0.488  0.62617
Fuel.efficiency23   -7.5149    48.0833  -0.156  0.87604
Fuel.efficiency24   -2.0931    47.5842  -0.044  0.96498
Fuel.efficiency25  -13.6867    46.7755  -0.293  0.77028
Fuel.efficiency26  -33.7991    48.4634  -0.697  0.48675
Fuel.efficiency27    5.6343    47.7956   0.118  0.90634
Fuel.efficiency28  -44.3212    57.9797  -0.764  0.44596
Fuel.efficiency29  -24.0084    63.6429  -0.377  0.70659
Fuel.efficiency30   -8.6596    53.2745  -0.163  0.87112
Fuel.efficiency31  -50.4471    58.0959  -0.868  0.38676
Fuel.efficiency32  121.6387    77.8582   1.562  0.12057
Fuel.efficiency33   -8.4885    55.1763  -0.154  0.87796
Fuel.efficiency45  -61.3487    77.9848  -0.787  0.43286
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 63.44 on 134 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.2529,    Adjusted R-squared:  0.1358
F-statistic:  2.16 on 21 and 134 DF,  p-value: 0.004544

> attributes(MulRegmod)
$names
 [1] "coefficients"  "residuals"     "effects"       "rank"          "fitted.values" "assign"        "qr"            "df.residual"
 [9] "na.action"     "contrasts"     "xlevels"       "call"          "terms"         "model"

$class
[1] "lm"

> plot(MulRegmod$residuals)
> qqnorm(MulRegmod$residuals)
>
```

#Plotting:

```
plot(MulRegmod$residuals)
```




```
qqnorm(MulRegmod$residuals)
```



Inference:

For linear Regression Model we got:

Residual standard error: 65.21 on 154 degrees of freedom

Multiple R-squared: 0.09282, Adjusted R-squared: 0.08693

F-statistic: 15.76 on 1 and 154 DF, p-value: 0.0001102

For Multiple Regression Model we got:

We got Residual standard error: 63.44 on 134 degrees of freedom

Multiple R-squared: 0.2529, Adjusted R-squared: 0.1358

F-statistic: 2.16 on 21 and 134 DF, p-value: 0.004544

Result:

Hence, we calculated the linear and multiple regression between various attributes of Car_Sales Dataset.