MAT5007 – Applied Statistical Methods

Embedded Lab – R Statistical Software

FALL SEMESTER – 20222023
L25+L26 SLOT

E-RECORD
Experiment No.: 2

Submitted By

Rajat Singh
Reg. no: 22MCA0139

MCA– I Year
SITE

DEPARTMENT OF MATHEMATICS
SCHOOL OF ADVANCED SCIENCES
VELLORE INSTITUTE OF TECHNOLOGY
VELLORE – 632 014
TAMIL NADU
INDIA

Date : 04-12-2022

Experiment 1: The sales-alldata.csv data describes the revenue and weather in a region during the first quarter of 2020. Fit a linear regression model with REVENUE as the response variable and TEMPERATURE as an explanatory variable. Fit the Linear Regression and interpret your result. Next, include SUN_HOURS (A “Sun-Hour” is “1000 watts of energy shining on 1 square meter of surface for 1 hour”.) as one more independent variable. Fit the Multiple Regression and interpret your result.

Data after importing:

```
> data
         DATE REVENUE   SUN_HOURS PRECIPITATION SNOW_DEPTH TEMPERATURE
1  2020-01-01    7637  0.00000000           0.0       0.00         0.0
2  2020-01-02    9276  0.37250000           0.0       0.00         5.8
3  2020-01-03   11170  0.26361111           0.0       0.00         5.4
4  2020-01-04   11863  3.54861111           0.0       0.00         2.4
5  2020-01-05   10880  2.51250000           2.4       0.00         0.1
6  2020-01-06    6702  0.09333333           0.2       0.00         4.8
7  2020-01-07    8652  0.00000000           0.0       0.00         5.5
8  2020-01-08    8346  0.16166667           0.0       0.00         6.5
9  2020-01-09    6543  0.31944444           0.6       0.00         3.9
10 2020-01-10    8115  1.58694444           0.0       0.00         1.0
11 2020-01-11    7728  0.00000000           0.2       0.00         1.2
12 2020-01-12   10649  0.44194444           0.0       0.00         5.9
13 2020-01-13    6787  2.42083333           0.0       0.00         2.6
14 2020-01-14    4555  0.00000000           7.4       0.00         3.8
15 2020-01-15    5885  0.00000000           0.0       0.00         7.8
16 2020-01-16   10127  0.85583333           0.0       0.00         6.1
17 2020-01-17    8893  0.00000000           0.0       0.00         4.8
18 2020-01-18   12520  0.05500000           0.0       0.00         4.8
19 2020-01-19   11860  6.18555556           0.0       0.00         3.1
20 2020-01-20    8515  0.77305556           0.0       0.00         3.8
21 2020-01-21    8129  2.09777778           0.0       0.00         6.8
22 2020-01-22   10405  3.55250000           0.0       0.00         3.2
23 2020-01-23    6672  2.02694444           0.0       0.00         1.5
24 2020-01-24   12300  4.97361111           0.0       0.00         5.1
25 2020-01-25   10651  4.45861111           0.0       0.00         2.1
26 2020-01-26   11882  0.00000000           0.0       0.00         5.8
27 2020-01-27    9397  0.00000000           0.6       0.00         4.8
```

```
> x=data$TEMPERATURE
> x
 [1]  0.0  5.8  5.4  2.4  0.1  4.8  5.5  6.5  3.9  1.0  1.2  5.9  2.6  3.8  7.8  6.1  4.8
[18]  4.8  3.1  3.8  6.8  3.2  1.5  5.1  2.1  5.8  4.8  3.4  3.3  2.4  3.0  6.1  3.8  0.5
[35] -1.3  0.5  0.4  0.1  2.4  4.5  5.7  4.0  3.2  1.0  0.8  3.8  7.1  6.6  6.6  4.5  2.8
[52]  5.7  6.5  5.2  2.3  1.4  0.2 -1.3 -1.3  1.1  4.2  4.4  3.7  2.1  2.1  2.0  1.9  4.5
[69]  6.8  5.0  5.6  4.9  2.1 -0.9  2.0  4.1  5.7  6.3  5.5  2.1  0.0  0.1  2.8  4.2  7.0
[86]  6.5  6.2  5.3  1.3  0.1  3.4
> y=data$REVENUE
> y
 [1]  7637  9276 11170 11863 10880  6702  8652  8346  6543  8115  7728 10649  6787  4555
[15]  5885 10127  8893 12520 11860  8515  8129 10405  6672 12300 10651 11882  9397  5174
[29]  4436  6202  9949 10192 13904 11208  8578  6638  7093  8187 10099  8160  3797  8734
[43]  4355  8452  7367  9339  5427  7022  8417  8760  4936 11273  7046 10074  9114  4493
[57]  4992  3789  7480  8358  7269  6489  3586  2570  5059  9087 11570 10476  8977  4509
[71]  5837  3559  7489  8138  8876  4800  3836  2750  7715  4364  9393  7830  4858  5140
[85]  7043  8662  9196  8267  8237  4249  8291
```

```
> fit=lm(y~x)
> fit

Call:
lm(formula = y ~ x)

Coefficients:
(Intercept)            x
    7543.44        66.28
```

Equation of Linear Regression: Y = 66.28 X + 7543.44

```
> w=data$SUN_HOURS
> w
 [1]  0.00000000  0.37250000  0.26361111  3.54861111  2.51250000  0.09333333  0.00000000
 [8]  0.16166667  0.31944444  1.58694444  0.00000000  0.44194444  2.42083333  0.00000000
[15]  0.00000000  0.85583333  0.00000000  0.05500000  6.18555556  0.77305556  2.09777778
[22]  3.55250000  2.02694444  4.97361111  4.45861111  0.00000000  0.00000000  0.00000000
[29]  0.00000000  0.00000000  0.00000000  1.50750000  6.70888889  6.85888889  0.09138889
[36]  3.96861111  3.26555556  1.15611111  0.00000000  1.14805556  0.23944444  0.69166667
[43]  1.69333333  8.53333333  6.66250000  0.01416667  0.00000000  1.55083333  2.58805556
[50]  7.62638889  3.16527778  7.49055556  0.20055556  6.78666667  8.94694444  2.53583333
[57]  0.51944444  0.00000000  6.55861111  2.25750000  0.91555556  6.01888889  0.00000000
[64]  0.00000000  0.04666667  9.34805556  2.85055556  2.89916667  3.60500000  0.72888889
[71]  0.13555556  0.66138889  5.24972222 11.19111111  5.43888889  6.09472222  0.59083333
[78]  1.32916667 11.57500000  3.21694444 10.37333333 11.13527778  6.43416667  5.35000000
[85]  6.95083333 11.44222222  9.69944444  3.41638889 11.69916667  7.36500000 10.23666667
```

```
> input=data.frame(y,x,w)
> input
      y    x           w
1   7637  0.0  0.00000000
2   9276  5.8  0.37250000
3  11170  5.4  0.26361111
4  11863  2.4  3.54861111
5  10880  0.1  2.51250000
6   6702  4.8  0.09333333
7   8652  5.5  0.00000000
8   8346  6.5  0.16166667
9   6543  3.9  0.31944444
10  8115  1.0  1.58694444
11  7728  1.2  0.00000000
12 10649  5.9  0.44194444
13  6787  2.6  2.42083333
14  4555  3.8  0.00000000
15  5885  7.8  0.00000000
16 10127  6.1  0.85583333
17  8893  4.8  0.00000000
18 12520  4.8  0.05500000
19 11860  3.1  6.18555556
```

```
> mulreg=lm(y~x+w,data=input)
> mulreg

Call:
lm(formula = y ~ x + w, data = input)

Coefficients:
(Intercept)            x            w
     6788.7        122.1        173.0
```

Equation For Multiple Regression: Y = 122.1 X + 173 W + 6788.7