**SIMPLE LINEAR REGRESSION**

This a type of regression that has only one independent variable and has linear association with the dependent variable. The regression equation has only one explanatory variable and is expressed in linear form as follows:

Y = A + βX

The above equation is purely mathematical in nature because it assumes exactness. However, in econometrics, we don’t believe that all the factors that cause variation in the dependent variable are included in the model, hence we include the error term (Є or U) to represent other factors that are not included in the model. Therefore, we state a new model as follows:

Y = A + βX +Є

Where

Y = The Dependent variable also called outcome variable or Regressand

A= Constant

β = Coefficient of the independent variable and X is the independent variable also called explanatory variable, predictor or regressor among others

X = the Independent variable

Є = Error term

**Illustration**

The income and expenditure of a particular household for the second half of year 2019 is given below:

|  |  |
| --- | --- |
| **Income ($)** | **Expenditure ($)** |
| **1** | **42** |
| **3** | **50** |
| **10** | **75** |
| **16** | **100** |
| **26** | **150** |
| **36** | **200** |

**Required:**

1. Estimate the model
2. Predict the value of expenditure when income is $70

**Solution**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **X** | **Y** | **X2** | **Y2** | **XY** |
| 1 | 42 | 1 | 1764 | 42 |
| 3 | 50 | 9 | 2500 | 150 |
| 10 | 75 | 100 | 5625 | 750 |
| 16 | 100 | 256 | 10,000 | 1,600 |
| 26 | 150 | 676 | 22,500 | 3,900 |
| 36 | 200 | 1,296 | 40,000 | 7,200 |
| **92** | **617** | **2338** | **82,389** | **13, 642** |
| **∑X** | **∑Y** | **∑X2** | **∑Y2** | **∑XY** |

β = n(∑XY) - (∑X) (∑)

 n∑X2 - (∑X)2

= 6(13642) -92(617)

 6(2338) – (92)2

 = 81,852 – 56,764

 = 14,028- 8464

 = 25,088

 = 5,564

 = 4.51

A = Ȳ -β x̅

Ȳ = ∑Y = 617 = 102.83

 N 6

x̅ = ∑X = 92 = 15.333 = 15.3

 N 6

A = 102.83 - 4.51(15.3)

 = 102.83 - 69.003

 = 33.827

 = $33. 83

Therefore, our model is estimated below:

 **ŷ**  = 33.83 + 4.51X + Є

This means that when income is zero, that is when the household earn zero income, the constant expenditure for the household is **$33.83**

To predict the value of expenditure when the household income is $70, the new value of income which is $70 is substituted in our model as follows:

**ŷ**  = 33.83 + 4.51X

 = 33.83 + 4.15(70)

 = 33.83 + 290.5

 = **$324. 33**

**MULTIPLE LINEAR REGRESSION**

Multiple linear regression is a type of linear regression that has 2 or more independent variables and each independent variable has a linear relationship with the dependent variable.

**The following table represents the income, expenditure and savings of allowance of Delight Venture in 2019**

|  |  |  |  |
| --- | --- | --- | --- |
| **Period** | **Income** | **Other Expenditure** | **Food expenditure** |
| January, 2019 | $200 | $150 | $50 |
| February, 2019 | $245 | $180 | $65 |
| March, 2019 | $ 275 | $200 | $75 |
| April, 2019 | $295 | $215 | $80 |
| May, 2019 | $315 | $250 | $65 |
| June, 2019 | $350 | $270 | $80 |
| July, 2019 | $390 | $300 | $90 |
| August, 2019 | $415 | $330 | $85 |
| September, 2019 | $450 | $350 | $100 |
| October, 2019 | $480 | $370 | $110 |
| November, 2019 | $510 | $385 | $125 |
| December, 2019 | $600 | $420 | $180 |

**Required:**

1. **Calculate the constant food expenditure of Delight Venture per month**
2. **Examine the effect of income and other expenditure on food expenditure**
3. **Predict the food expenditure when the income and other expenditure increase to $1,000 and $800 respectively in June, 2020**

**Solutions**

**Step 1.**

**Find the correlation between two variables. Let’s take the two independent variables first.**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Period** | **Y** | **X1** | **X2** | **Y2** | **X12** | **X22** | **X1X2** | **X1Y** | **X2Y** |
|  | **$** | **$** | **$** | **$** | **$** | **$** | **$** | **$** | **$** |
| January | 50 | 200 | 150 | 2,500 | 40,000 | 22,500 | 30,000 | 10,000 | 7,500 |
| February | 65 | 245 | 180 | 4,225 | 60,025 | 32,400 | 44,100 | 15,925 | 11,700 |
| March | 75 | 275 | 200 | 5,625 | 75,625 | 40,000 | 55,000 | 20,625 | 15,000 |
| April | 80 | 295 | 215 | 6,400 | 87,025 | 46,225 | 63,425 | 23,600 | 17,200 |
| May | 65 | 315 | 250 | 4,225 | 99,225 | 62,500 | 78,750 | 20,475 | 16,250 |
| June | 80 | 350 | 270 | 6,400 | 122,500 | 72,900 | 94,500 | 28,000 | 21,600 |
| July | 90 | 390 | 300 | 8,100 | 152,100 | 90,000 | 117,000 | 35,100 | 27,000 |
| August | 85 | 415 | 330 | 7,225 | 172,225 | 108,900 | 136,950 | 35,275 | 28,050 |
| September | 100 | 450 | 350 | 10,000 | 202,500 | 122,500 | 157,500 | 45,000 | 35,000 |
| October | 110 | 480 | 370 | 12,100 | 230,400 | 136,900 | 177,600 | 52,800 | 40,700 |
| November | 125 | 510 | 385 | 15,625 | 260,100 | 148,225 | 196,350 | 63,750 | 48,125 |
| December | 180 | 600 | 420 | 32400 | 360,000 | 176,400 | 252,000 | 108,000 | 75,600 |
|  | **1,105** | **4,535** | **3,420** | **114,825** | **1,861,725** | **1,059,450** | **1,403,175** | **458,550** | **343,725** |
|  | **∑Y** | **∑X1** | **∑X2** | **∑Y2** | **∑X12** | **∑X22** | **∑X1X2** | **∑X1Y** | **∑X2Y** |

**Let the correlation between the two independent variables be represented by rx1x2**

**rx1x2 =** n∑x1x2 -∑x1∑x2

√[n∑x12- (∑X1)2] [n∑x22 –(∑x2)2

= 12(1,403,175) -4535x3420

√[12x1861,725-43252] [12x1,059,450-34202]

= 16,838,100 -15,509,700

√[22,340,700-18,705,625] [12,713,400-11,696,400]

= 1,328,400

√[3,635,075] [1,017,000]

= 1328,400

∑3,696,871,275,000

=1,328,400

1,922,724.96

**rx1x2** = 0.6909

Let the correlation between the dependent variable and the first independent variable(X1) be represented by ryx1

ryx1 =n∑yx1 -∑y∑x1

√ [n∑y2- (∑y)2] [n∑x12 –(∑x1)2

= 12(458,550) -1,105x4535

√[12x114,825-1,1052] [12x1,861,725-45352]

= 5,502,600 -5,011,175

√ [1,377,900-1,221,025] [22,340,700-20,566,225]

= 491,425

√[156,875] [1,774,475]

= 491,425

 √278,370,765,625

491,425

527,608.53

ryx1 = 0.9314

Let ryx2 represent the correlation between the dependent variable(y) and the second independent variable(X2)

ryx2 = =n∑yx2 -∑y∑x2

√ [n∑y2- (∑y)2] [n∑x22 –(∑x2)2

= 12(343,725) -1,105x3420

√[12x114,825-1,1052] [12x1,059,450-34202]

4,124,700- 3,779,100

√[1,377,900- 1,221,025] [12,713,400-11,696,400]

345,600

√ [156,875] [1,017,000]

345,600

√159,541,875,000

345,600

399,426.93

ryx2 = 0.8652

**Step 2: Calculate the mean and standard deviation**

**Mean**

 X̅1= ∑x1 = 4535 = **377.91**

 N 12

 X̅2= ∑x2 = 3430 = **285.83**

 N 12

 **Ȳ**= ∑Y = 1,105 =**92.08**

 N 12

**Standard Deviation**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **y** | **y- ȳ** | **(y- ȳ )2** | **X1** | **X1-** X̅1 | **(X1-** X̅1**)2** | **X2** | **X2-** X̅2 | **(X2-** X̅2**)2** |
| 50 | -42.08 | 1,770.72 | 200 | -177.91 | 31,651.97 | 150 | -135.83 | 18,449.79 |
| 65 | -27.08 | 733.32 | 245 | -132.91 | 17,665.07 | 180 | -105.83 | 11,199.99 |
| 75 | -17.08 | 291.73 | 275 | -102.91 | 10,590.47 | 200 | -85.83 | 7,366.79 |
| 80 | -12.08 | 145.93 | 295 | -82.91 | 6,874.07 | 215 | -70.83 | 5016.89 |
| 65 | -27.08 | 733.32 | 315 | -62.91 | 3,957.67 | 250 | -35.83 | 1,283.79 |
| 80 | -12.08 | 145.93 | 350 | -27.91 | 778.97 | 270 | -15.83 | 250.59 |
| 90 | -2.08 | 4.33 | 390 | 12.09 | 146.17 | 300 | 14.17 | 200.79 |
| 85 | -7.08 | 50.13 | 415 | 37.09 | 1,375.67 | 330 | 44.17 | 1,950.99 |
| 100 | 7.92 | 62.73 | 450 | 72.09 | 5,196.97 | 350 | 64.17 | 4,177.79 |
| 110 | 17.92 | 321.13 | 480 | 102.91 | 10,506.08 | 370 | 84.17 | 7,084.59 |
| 125 | 32.92 | 1,083.73 | 510 | 132.09 | 17,447.77 | 385 | 99.17 | 9,834.69 |
| 180 | 87.92 | 7729. 93 | 600 | 222.09 | 49,323.97 | 420 | 134.17 | 18,001.59 |
|  |  | **13,072.93** |  |  | **155,514.85** |  |  | **84,818.28** |

**SD of X1 =√∑(X1 -X1)2 = √**155,514.85 = √155,514.85

 N-1 12-1 11

= √14,137.71 = **118.902**

**SD of X2 =√∑(X2 -X2)2 = √**84,818,28 = √84,818.28

 N-1 12-1 11

= √7,710.75 = 87.81

**SD of Y =√∑(Y-Y)2 = √**13,072.93 = √13,072.93

 N-1 12-1 11

= √13,072.93 = **114.34**

**MEAN SD**

**Ȳ** = 92.08 114.34

 X̅1 = 377.91 118.90

X̅2 = 285.83 87.81

R = √[(ryx1)2 +(ryx2)2 ] – 2(ryx1\*ryx2\*rx1x2)

 1-(rx1x2)2

= √[ (0.9314)2 +(0.8652)2] -2(0.9314\*0.8652\*0.6909)

1. (0.6909)2

= √[ (0.8675)+(0.7486)] -2(0.9314\*0.8652\*0.6909)

1-(0.4773)

= √[ 1.6161] –[1.1135]

0.5227

= √[ 1.6161] –[1.1135]

0.5227

 = √0.5026

 0.5227

 = √0.9615

 = 0.980561

R = **0. 9806**

This implies a positive strong correlation between savings, income and expenditure. It means the higher the income, the higher the expenditure and the higher the savings

R2 = 0.96157

 = 0.9616

This result indicates that 96.16% of factors that affect savings are included in our model while the remaining 3.84% are not included in our model.

 **Making Predictions: Multiple Regression Model**

**Y = a + β1x1 + β2x2 +Є**

**Where:**

Y **=** Food expenditure

A = constant (the value of y where X1 and X2 are zero)

β 1 = The coefficient of income

β2 =The coefficient of other expenditure

X1 = Income

X2 = Other expenditure

Є = Error term

**Step 3: Calculate β1 and β2**

**β1 =** **[ryx1-(ryx2\*rx1x2)] (SDY)**

1. **(rx1x2)2 SDX1**

0.9314 – (0.8652\*0.6909) 114.34

1 - (0.6909)2 118.90

0.9314 – 0.5978 0.9616

1- 0.4773

 0.3336 \* 0.9616

 0.5227

 = 0.6382 \* 0.9616

 = **0.6136**

**β2 = [ryx2-(ryx1\*rx1x2)] (SDY)**

1. **(rx1x2)2 SDX2**

 = [ 0.8652 – (0.9314\* 0.6909)] (114.34)

1. (0.6909)2  87.81

= 0.8652 - 0.6435 (1.3021)

1. 0.4773

0.222 (1.3021)

0.5227

 0.4247 (1.3021)

1. **5530**

**a = Ȳ - β1**X̅1 **- β2**X̅2

a = 92.08 -( 0.6136\*377.91) – (0.5530 \*285.83)

a = 92.08 – (231.8856) - (158. 06)

a = 92.08 - 231.8856 – 158.0639

a = 92.08 - 389.9499

**a = -297.8699**

**Therefore, the constant food expenditure per month is ₦-297.8699**

**The final model is given below:**

**Y** = a + β1x1 + β2x2 +Є

Y = - 297. 6699 + 0.6136(Income) + 0.5530( Other expenditure) +Є

The amount of food expenditure when income increases to 1,000 and other expenditure increases to 800 is predicted below:

**Y** = a + β1x1 + β2x2 +Є

Y = - 297.8699 +0.6136\*1,000 + 0.5530\*800

Y = -297.8699 + 613.6 + 442.4

**Y = ₦758.1301**