Name: Abu Syed Sujon

Topic: Assignment on Linear Regression

Task-01: Compute the slope (M) and Y - Intercept (c).

To compute the slope (M) and y - intercept (c) using ordinary least squares Linear Regression, we use the following formulas:

→ Slope (M):

$$M = \frac{N(\Sigma xy) - (\Sigma x)(\Sigma y)}{N(\Sigma x^2) - (\Sigma x)^2}$$

→ Y - Intercept (c):

$$c = \frac{(\Sigma y) - M(\Sigma x)}{N}$$

Where:

- N is the number of data points.
- Σxy is the sum of the products of each pair of x and y.
- Σx and Σy are the sums of x values and y values, respectively.
- Σx^2 is the sum of each x value squared.

Now,

compute these values with the given dataset:

- Weight (x) : 2, 4, 5, 3, 6, 5, 7
- Price (y) : 35, 60, 20, 50, 50, 55, 60

1. Calculate the Sums:

* $\Sigma x = 2+4+5+3+6+5+7 = 32$
* $\Sigma y = 35+60+20+50+50+55+60 = 330$
* $\Sigma xy = (2\times35)+(4\times60)+(5\times20)+(3\times50)+(6\times50)$
  $+(5\times55)+(7\times60)$
  $= 1555$

* $\Sigma x^2 = 2^2+4^2+5^2+3^2+6^2+5^2+7^2 = 164$
* $N = 7$

2. Compute M and C:

* $M = \dfrac{(7\times1555)-(32\times330)}{(7\times164)-32^2} = \dfrac{10885-10560}{1148-1024}$
  $= \dfrac{325}{124}$
  $= 2.621$ (approx)

* $C = \dfrac{330-(2.621\times32)}{7} = \dfrac{330-83.872}{7}$
  $= \dfrac{246.128}{7}$
  $= 35.161$

So, the equation of the line is approximately:

$$y = 2.621x + 35.161 \quad \text{--- ①}$$

Now, Predict the Price for a vegetable weight of 6

Using the equation no ①;

$$y = (2.621 \times 6) + 35.161$$
$$= 15.726 + 35.161$$
$$= 50.887$$

Task-02:

Compute the Residuals.

The residual for each data point is the difference between the observed value and the predicted value.

The formula for each residual is:

$$\text{Residual} = y_i - (Mx_i + C)$$

Let's compute the residuals for each data point:

1. Weight = 2:
   - Predicted price $= (2.621 \times 2) + 35.161 = 40.403$
   - Residual $= 35 - 40.403 = -5.403$

(P.T.O)

2. Weight = 4:

* Predicted price = $(2.621 \times 4) + 35.161 = 45.645$

* Residual = $60 - 45.645 = 14.355$

3. Weight = 5:

* predicted price = $(2.621 \times 5) + 35.161 = 48.266$

* Residual = $20 - 48.266 = -28.266$

4. weight = 3:

* Predicted price = $(2.621 \times 3) + 35.161 = 43.024$

* Residual = $50 - 43.024 = 6.976$

5. Weight = 6:

* Predicted price = 50.887 (calculated before)

* Residual = $50 - 50.887 = -0.887$

6. Weight = 5:

* Predicted price = 48.266 (same as above)

* Residual = $55 - 48.266 = 6.734$

7. Weight = 7:

* Predicted price = (6.621 × 7) + 35.161 = 53.508
* Residual = 60 − 53.508 = 6.492

SO, the residual for each data point are approximately:

* Weight		Value
2	−	−5.403
4	−	14.355
5	−	−28.266
3	−	6.976
6	−	−0.887
5	−	6.734
7	−	6.492

Task-03:

Compute the Mean Squared Error (MSE) and Mean Absolute Error (MAE)

1. MSE is the average of the squared residuals:

$$MSE = \frac{1}{N} \sum (Y_i - \hat{Y}_i)^2$$

(P.T.O)

$$MSE = \frac{1}{7}\left[(-5.403)^2 + (14.355)^2 + (-28.266)^2 + (6.976)^2 + (-0.887)^2 + (6.734)^2 + (6.492)^2\right]$$

$$= \frac{1}{7}\left[29.217 + 206.090 + 798.976 + 48.663 + 0.787 + 45.343 + 42.156\right]$$

$$= \frac{1}{7}\left[1170.232\right]$$

$$= 167.176$$

2. <u>MAE</u>

MAE is the average of the absolute residuals:

$$MAE = \frac{1}{N} \sum |Y_i - \hat{Y}_i|$$

$$MAE = \frac{1}{7}\left[|-5.403| + |14.355| + |-28.266| + |6.976| + |-0.887| + |6.734| + |6.492|\right]$$

$$= \frac{1}{7}\left[5.403 + 14.355 + 28.266 + 6.976 + 0.887 + 6.734 + 6.492\right]$$

$$= \frac{1}{7}\left[69.113\right]$$

$$= 9.873$$