LINEAR REGRESSION

Weight = [2,4,5,3,6,5,7]

Price = [35,60,20,50,50,55,60]

Here, No. of observations, N = 7

SLOPE(M) & y-INTERCEPT(C)

$$Slope, M = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2}$$

Calculation of $\bar{x}$ (mean of x):

$$\bar{x} = \frac{\sum x}{N} = \frac{2 + 4 + 5 + 3 + 6 + 5 + 7}{7} = \frac{32}{7} = 4.571428571$$

Calculation of $\bar{\bar{y}}$ (mean of y):

$$\bar{\bar{y}} = \frac{\sum y}{N} = \frac{35 + 60 + 20 + 50 + 50 + 55 + 60}{7} = \frac{330}{7} = 47.14285714$$

Calculation of $\sum(x - \bar{x})(y - \bar{y})$:

$$\begin{aligned}(2 - 4.571428571)(35 - 47.14285714) &= (-2.571428571)(-12.14285714) \\ &= 31.22448978\end{aligned}$$

$$\begin{aligned}(4 - 4.571428571)(60 - 47.14285714) &= (-0.571428571)(12.85714286) \\ &= -7.346938772\end{aligned}$$

$$\begin{aligned}(5 - 4.571428571)(20 - 47.14285714) &= (0.428571429)(-27.14285714) \\ &= -11.63265307\end{aligned}$$

$$(3-4.571428571)(50-47.14285714) = (-1.571428571)(2.85714286)$$
$$= -4.489795922$$

$$(6-4.571428571)(50-47.14285714) = (1.428571429)(2.85714286)$$
$$= 4.081632658$$

$$(5-4.571428571)(55-47.14285714) = (0.428571429)(7.85714286)$$
$$= 3.367346943$$

$$(7-4.571428571)(60-47.14285714) = (2.428571429)(12.85714286)$$
$$= 31.22448981$$

$$\sum(x-\bar{x})(y-\bar{y}) = 31.22448978 - 7.346938772 - 11.63265307 - 4.489795922 +$$
$$4.081632658 + 3.367346943 + 31.22448981$$
$$= 46.42857143$$

Calculation of $\sum(x-\bar{x})^2$:

$(2-4.571428571)^2 = 6.612244896$

$(4-4.571428571)^2 = 0.3265306118$

$(5-4.571428571)^2 = 0.1836734698$

$(3-4.571428571)^2 = 2.469387754$

$(6-4.571428571)^2 = 2.040816328$

$(5-4.571428571)^2 = 0.1836734698$

$(7-4.571428571)^2 = 5.897959186$

$$\sum(x-\bar{x})^2 = 6.612244896 + 0.3265306118 + 0.1836734698 + 2.46938775 +$$
$$2.040816328 + 0.1836734698 + 5.897959186$$
$$= 17.71428572$$

$$\boldsymbol{Slope, M} = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sum(x - \bar{x})^2}$$

$$= \frac{46.42857143}{17.71428572}$$

$$= 2.620967741$$

$$\boldsymbol{Intercept, C} = \bar{y} - m\,\bar{x}$$

$$= 47.14285714 - (2.620967741 \times 4.571428571)$$

$$= 35.16129033$$

Predicted price for the vegetable weight 6,

$$\boldsymbol{y = mx + c = 2.620967741 \times 6 + 35.16129033 = 50.88709678}$$

RESIDUAL

Residuals for each data point,

$$Residual = Observed\ Value - Predicted\ Value = y - \hat{y}$$

$For\ x\ =\ 2,\ \hat{y} = 2.620967741 \times 2 + 35.16129033 = 40.40322581$

$\boldsymbol{Residual = 35 - 40.40322581 = -5.40322581}$

$For\ x\ =\ 4,\ \hat{y} = 2.620967741 \times 4 + 35.16129033 = 45.64516129$

$\boldsymbol{Residual = 60 - 45.64516129 = 14.35483871}$

$For\ x\ =\ 5,\ \hat{y} = 2.620967741 \times 5 + 35.16129033 = 48.26612904$

$\boldsymbol{Residual = 20 - 48.26612904 = -28.26612904}$

$For\ x\ = 3,\ \hat{y} = 2.620967741 \times 5 + 35.16129033 = 43.02419355$

$\boldsymbol{Residual = 50 - 43.02419355 = 6.97580645}$

$For\ x\ = 6,\ \hat{y} = 2.620967741 \times 6 + 35.16129033 = 50.88709678$

$\boldsymbol{Residual = 50 - 50.88709678 = -0.88709678}$

$For\ x\ =\ 5,\ \hat{y} = 2.620967741 \times 5 + 35.16129033 = 48.26612904$

$\boldsymbol{Residual = 55 - 48.26612904 = 6.73387096}$

$For\ x\ =\ 7,\ \hat{y} = 2.620967741 \times 7 + 35.16129033 = 53.50806452$

$\boldsymbol{Residual = 60 - 53.50806452 = 6.491935483}$

MEAN SQUARED ERROR (MSE):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \widehat{y_i})^2$$

Calculation for MSE,

$$\boldsymbol{MSE} = \frac{1}{7}((-5.40322581)^2 + (14.35483871)^2 + (-28.26612904)^2 + (6.97580645)^2 + (-0.88709678)^2 + (6.73387096)^2 + |(6.491935483)^2)$$

$$= \frac{1}{7} \times 1171.169355$$

=167.3099079

Root Mean Squared Error, $RMSE = \sqrt{MSE} = \sqrt{167.3099079} = 12.93483312$

MEAN ABSOLUTE ERROR (MAE):

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - \hat{y}_i|$$

Calculation of MAE,

$$\boldsymbol{MAE} = \frac{1}{7}(|-5.40322581| + |14.35483871| + |-28.26612904| + |6.97580645| + |-0.88709678| + |6.73387096| + |6.491935483|)$$

$$= \frac{1}{7} \times 69.11290323$$

$$=9.87327189$$