

## Assignments-04 Linear Regressions assignment manually.

### Task 01:

Your objective is to manually compute the slope (M) and y-intercept (C) using Ordinary Least Squares Linear Regression. Once determined, apply these values to predict the price when the vegetable weight is 6.

Mean of (x):

$$\begin{aligned} \bar{x} &= (2 + 4 + 5 + 3 + 6 + 5 + 7) / 7 \\ &= 32 / 7 \\ &= 4.57 \end{aligned}$$

Mean of (y):

$$\begin{aligned} \bar{y} &= (35+60+20+50+50+55+60) / 7 \\ &= 330 / 7 \\ &= 47.14 \end{aligned}$$

### Calculate the slope (M):

$$\begin{aligned} (2-4.57)(35-47.14) &= (-2.57)(-12.14) = 31.27 \\ (4-4.57)(60-47.14) &= (-0.57)(12.86) = -7.33 \\ (5-4.57)(20-47.14) &= (0.43)(-27.14) = -11.68 \\ (3-4.57)(50-47.14) &= (-1.57)(2.86) = 4.49 \\ (6-4.57)(50-47.14) &= (1.43)(2.86) = 4.09 \\ (5-4.57)(55-47.14) &= (0.43)(7.86) = 3.39 \\ (7-4.57)(60-47.14) &= (2.43)(12.86) = 31.27 \end{aligned}$$

---

$$\text{Total} = 46.92$$

$$\begin{aligned} &(2-4.57)^2 + (4-4.57)^2 + (5-4.57)^2 + (3-4.57)^2 + (6-4.57)^2 + (5-4.57)^2 + (7-4.57)^2 \\ &= (-2.57)^2 + (-0.57)^2 + (0.43)^2 + (-1.57)^2 + (1.43)^2 + (0.43)^2 + (2.43)^2 \\ &= 6.60 + 0.32 + 0.18 + 2.46 + 2.04 + 0.18 + 5.90 \\ &= 17.68 \end{aligned}$$

Now, calculate this slope is:

$$\begin{aligned} M &= (46.92 / 17.68) \\ &= 2.65 \end{aligned}$$

y-intercept (C):

$$\begin{aligned}C &= \bar{y} - M \cdot \bar{x} \\&= 47.14 - 2.65 \cdot 4.57 \\&= 47.14 - 12.13 \\&= 35.01\end{aligned}$$

Predict price for weight of 6:

$$\begin{aligned}y &= Mx + c \\&= 2.65 \cdot 6 + 35.01 \\&= 15.90 + 35.01 \\&= 50.91\end{aligned}$$

### Task 02:

**Compute the residuals for each data point.**

1 নং টাস্ক হতে পাবি M and c value.

The slope is  $M = 2.65$  and the intercept is  $C = 35.01$ .

রিগ্রেশনের সমীকরণটা: এই সমীকরণ দিয়ে প্রাইসটা প্রেডিকশন করব

$$Y_i = 2.65 \cdot x + 35.01$$

Weight (xi)	Actual price (y)	Predicted price (yi)	Residual (y - yi)
2	35	$2.65 \cdot 2 + 35.01 = 40.31$	$35 - 40.31 = -5.31$
4	60	45.61	$60 - 45.61 = 14.39$
5	20	48.26	$20 - 48.26 = -28.26$
3	50	42.96	$50 - 42.96 = 7.04$
6	50	50.91	$50 - 50.91 = -0.91$
5	55	48.26	$55 - 48.26 = 6.74$
7	60	53.56	$60 - 53.56 = 6.44$

**Task 03: Calculate both the Mean Squared Error (MSE) and Mean Absolute Error (MAE).**

1. Mean Squared Error (MSE):

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

MSE = mean squared error  
 $n$  = number of data points  
 $Y_i$  = observed values  
 $\hat{Y}_i$  = predicted values

২ নং হতে পাই রেসিডুয়ালগুলো,,

Residuals = -5.31, 14.39, -28.26, 7.04, -0.91, 6.74, 6.44

Number of data point ( $n = 7$ )

$$(-5.31)^2 = 28.21$$

$$(14.39)^2 = 207.03$$

$$(-28.26)^2 = 798.70$$

$$(7.04)^2 = 49.56$$

$$(-0.91)^2 = 0.83$$

$$(6.74)^2 = 45.43$$

$$(6.44)^2 = 41.47$$

$$(7.04)^2 = 49.56$$

$$\begin{aligned} \text{Sum} &= 28.21 + 207.03 + 798.70 + 49.56 + 0.83 + 45.43 + 41.47 \\ &= 1171.23 \end{aligned}$$

$$\text{Now, MSE} = 1171.23 / 7$$

$$= 167.32$$

2. Mean Absolute Error (MAE):

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

**Absolute values of residuals:**

$$\begin{aligned} |-5.31| &= 5.31 \\ |14.39| &= 14.39 \\ |-28.26| &= 28.26 \\ |7.04| &= 7.04 \\ |-0.91| &= 0.91 \\ |6.74| &= 6.74 \\ |6.44| &= 6.44 \end{aligned}$$

Sum of absolute residuals:

$$5.31 + 14.39 + 28.26 + 7.04 + 0.91 + 6.74 + 6.44 = 69.09$$

$$\begin{aligned} \text{MAE} &= 69.09 / 7 \\ &= 9.83 \end{aligned}$$

**Final Task: Generate an Excel file for the given dataset. Utilize Python for all the calculations.**

A1					Weight
	A	B	C	D	E
1	Weight	Actual Price	Predicted Price	Residual	
2	2	35	40.31	-5.31	
3	4	60	45.61	14.39	
4	5	20	48.26	-28.26	
5	3	50	42.96	7.04	
6	6	50	50.91	-0.91	
7	5	55	48.26	6.74	
8	7	60	53.56	6.44	
9					