EXERCISE SIMPLE LINEAR REGRESSION

Given the data in the following table:

Online Store	Advertising Dollars	Monthly Sales
1	1.7	368
2	1.5	340
3	2.8	665
4	5.0	954
5	1.3	331
6	2.2	556
7	1.3	376
8	1.9	412
9	4.5	846
10	2.0	392
11	3.4	603
12	2.9	355
13	3.8	538
14	4.6	709
15	3.9	461

Calculate:

- **b0 and b1 values of the regression line:**

$$b_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad b_0 = \bar{y} - b_1\bar{x}$$

Online Store	Advertising Dollars	Monthly Sales
1	1.7	368
2	1.5	340
3	2.8	665
4	5.0	954
5	1.3	331
6	2.2	556
7	1.3	376
8	1.9	412
9	4.5	846
10	2.0	392
11	3.4	603
12	2.9	355
13	3.8	538
14	4.6	709
15	3.9	461
Mean:	**2.9**	**527.07**

$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$y_i - \bar{y}$	$(y_i - \bar{y})^2$	$(x_i - \bar{x})(y_i - \bar{y})$
-1.2	1.3	-159.07	25302.2	183.4568889
-1.4	1.8	-187.07	34993.9	253.1635556
-0.1	0.0	137.93	19025.6	-7.356444444
2.1	4.6	426.93	182272.1	916.4835556
-1.6	2.4	-196.07	38442.1	304.5568889
-0.7	0.4	28.93	837.1	-18.90311111
-1.6	2.4	-151.07	22821.1	234.6568889
-1.0	0.9	-115.07	13240.3	109.6968889
1.6	2.7	318.93	101718.5	525.1768889
-0.9	0.7	-135.07	18243.0	115.2568889
0.5	0.3	75.93	5765.9	41.51022222
0.0	0.0	-172.07	29606.9	-8.029777778
0.9	0.9	10.93	119.5	10.35022222
1.7	3.1	181.93	33099.7	317.7768889
1.0	1.1	-66.07	4364.8	-69.14977778
Sum:	**22.7**		**529852.9**	**2908.646667**

$$b_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{2908.646667}{22.7} = 128.04$$

$$b_0 = \bar{y} - b_1\bar{x} = 527.07 - 128.04 \cdot 2.9 = 161.74$$

- **Estimated values for the monthly sales and residual errors:**

Online Store	Advertising Dollars	Monthly Sales	Monthly Sales (predicition)	Residuals
1	1.7	368	379.41	-11.41
2	1.5	340	353.80	-13.80
3	2.8	665	520.25	144.75
4	5.0	954	801.94	152.06
5	1.3	331	328.19	2.81
6	2.2	556	443.43	112.57
7	1.3	376	328.19	47.81
8	1.9	412	405.02	6.98
9	4.5	846	737.92	108.08
10	2.0	392	417.82	-25.82
11	3.4	603	597.08	5.92
12	2.9	355	533.06	-178.06
13	3.8	538	648.29	-110.29
14	4.6	709	750.72	-41.72
15	3.9	461	661.10	-200.10
Sum:				**-0.21**

- **Pearson's correlation coefficient:**

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} = \frac{2908.646667}{\sqrt{22.7 \cdot 529852.9}} = 0.8384$$

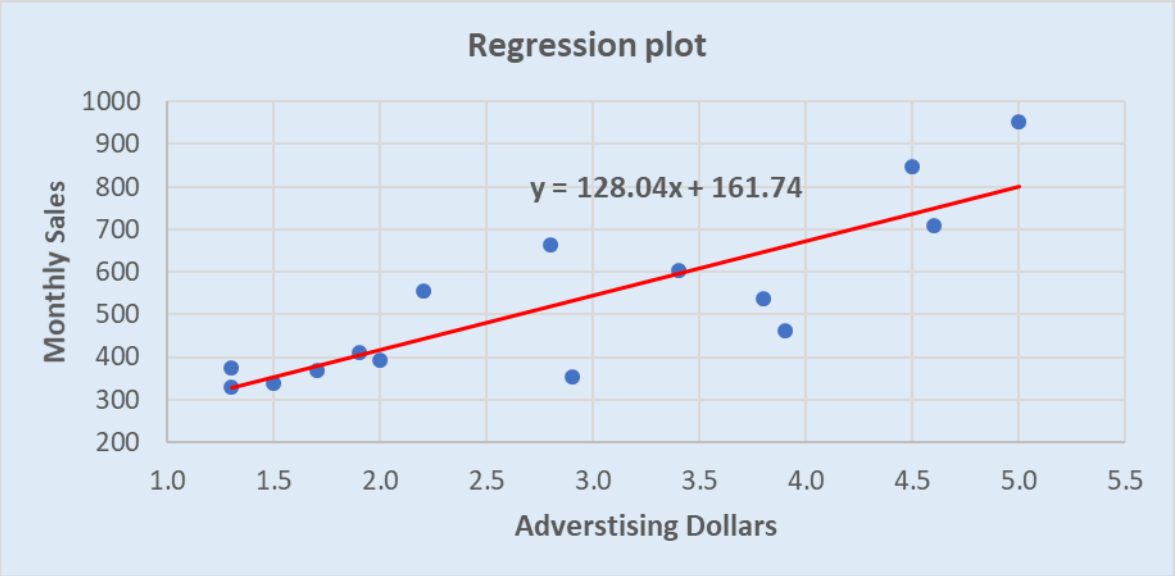
- **R^2:**

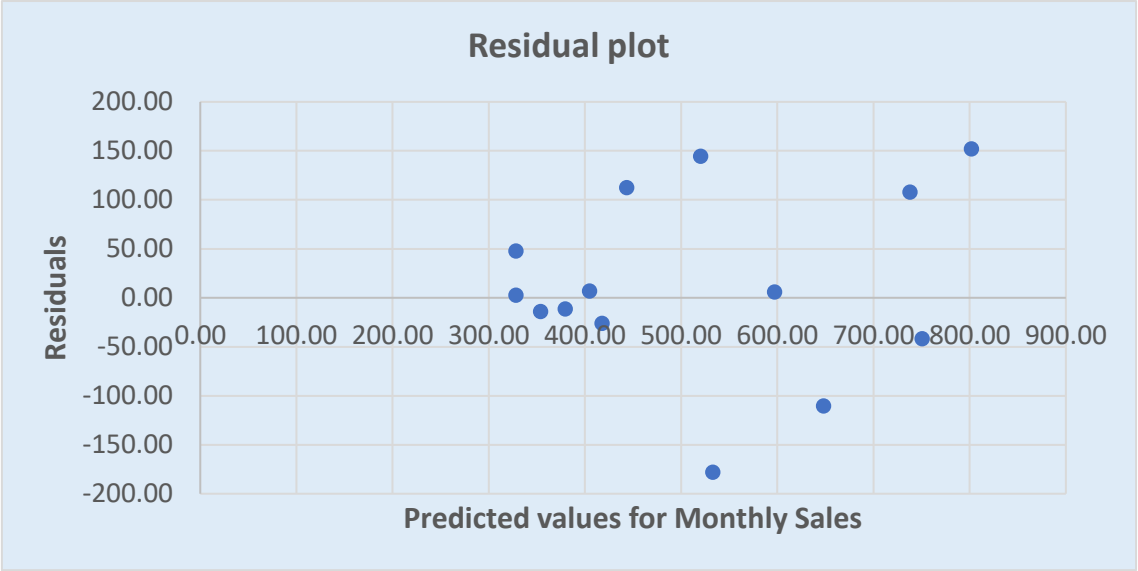
$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = 1 - \frac{157440.1}{529852.9} = 0.7029$$

- **MSE:**

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \frac{157440.1}{15} = 10496.01$$

Represent the regression plot and the residual plot:





How can we interpret the regression line?

The slope shows that the average monthly sales increase by $128.04 for each one-unit increase in advertising dollars.