The table below shows the number of grams of carbohydrates, X and the number of Calories, Y of six different foods. Find linear regression equation for this data set and plot the best fit line.

Carbihydrades (X)	8	9.5	10	6	7	4
Calories (Y)	12	138	147	88	108	62

The linear regression is a supervised machine learning algorithm which calculates the linear relationship between the dependent variable output Y and independent features input X by fitting a linear equation .

Following equation is used to calculate the best fit line:

$$Y = a * X + b$$

Where ,
w is the intercept of line
b is the slope called as linear regression coefficient .

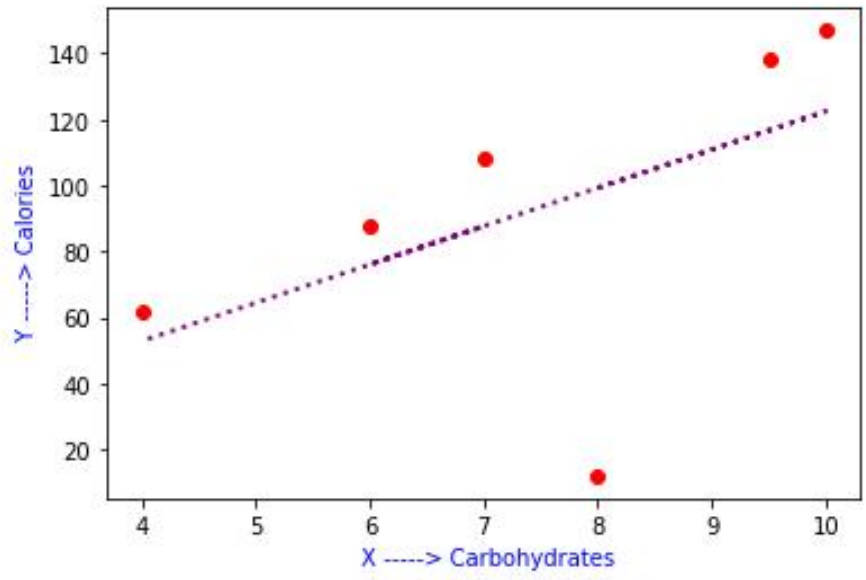
$$a = (Y - b) / X$$

`

$$b = Y - (a * X)$$

Plot the line for dependent variables & independent variable values on X-Y graph, refer following table and graph . The given data-set and relationship between X & Y variables is given below in table and graph :

Carbohydrates (X)	8	9.5	10	6	7	4
Calories (Y)	12	138	147	88	108	62



The different values for weights or the coefficient of lines (w, b) gives a different line of regression.

y	=	a	+	b	+	x
prediction		Bias		feature weight		Feature value
		calculated from training				

In this example, six dependent variables and six independent features are given, the regression model that relies on six features can be written as follows:

$$y' = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6$$

Now, calculate the best values for a and b to find the best fit line and calculate the cost function. cost function is used to estimate the values of the coefficient for the best fit line. Cost function optimizes the regression coefficients and measures linear regression model performance. linear regression, used to find the best fit line that means the error between predicted values and actual values should be minimized. The best fit line will have the minimum error.

x	**y**	X*y	x^2	y^2	$\bar{y}$
8	**12**	96	64	144	99.19
9.5	**138**	1311	90.25	19044	116.605
10	**147**	1470	100	21609	122.41
6	**88**	528	36	7744	75.97
7	**108**	756	49	11664	87.58
4	**62**	248	16	3844	52.75
44.5	555	4409	355.25	64049	
Slope: b = 11.61					
intercept: a = 6.39					

Calculate the following values before you can calculate a regression line:

Mean of x values $\bar{x} = \sum x/n$

Mean of y values - $\bar{Y} = \sum y/n$

Calculate slope b for given values using following equation,

$$b = \frac{n*(\sum x*y - \sum x*\sum y}{n*\sum x^2 - \sum y^2}$$

Calculate intercept a using following equation,

$$a = \bar{y} - b*\bar{x}$$

Following is the linear regression equation for the data set:

$$y = 6.39 + 11.61*y$$

The corresponding plot the linear regression line equation is given below.