Simple Linear Regression

Thursday, 25 July 2024 1:25 PM

[Machine Learning Tutorial Python - 2: Linear Regression Single Variable]

Home prices in Monroe Twp, NJ (USA)

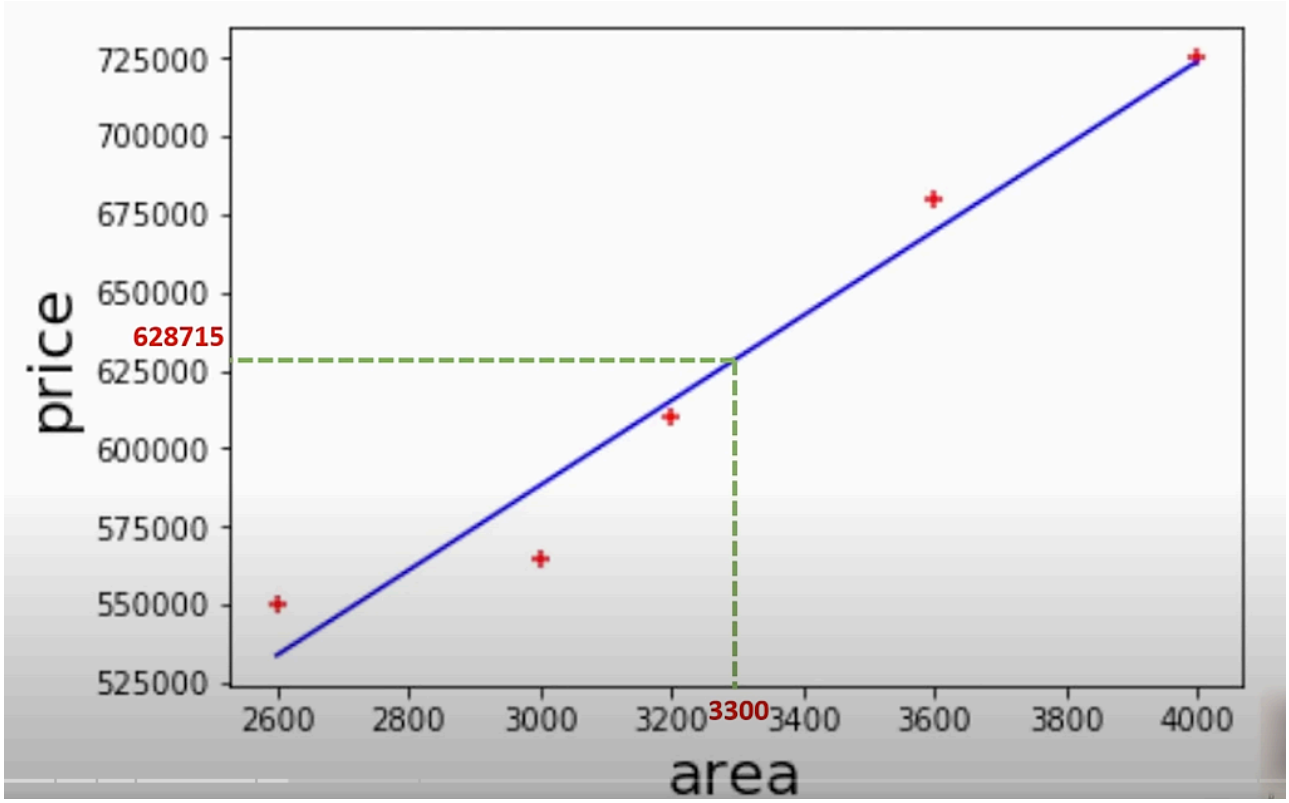
area	price
2600	550000
3000	565000
3200	610000
3600	680000
4000	725000

Given these home prices find out prices of homes whose area is,

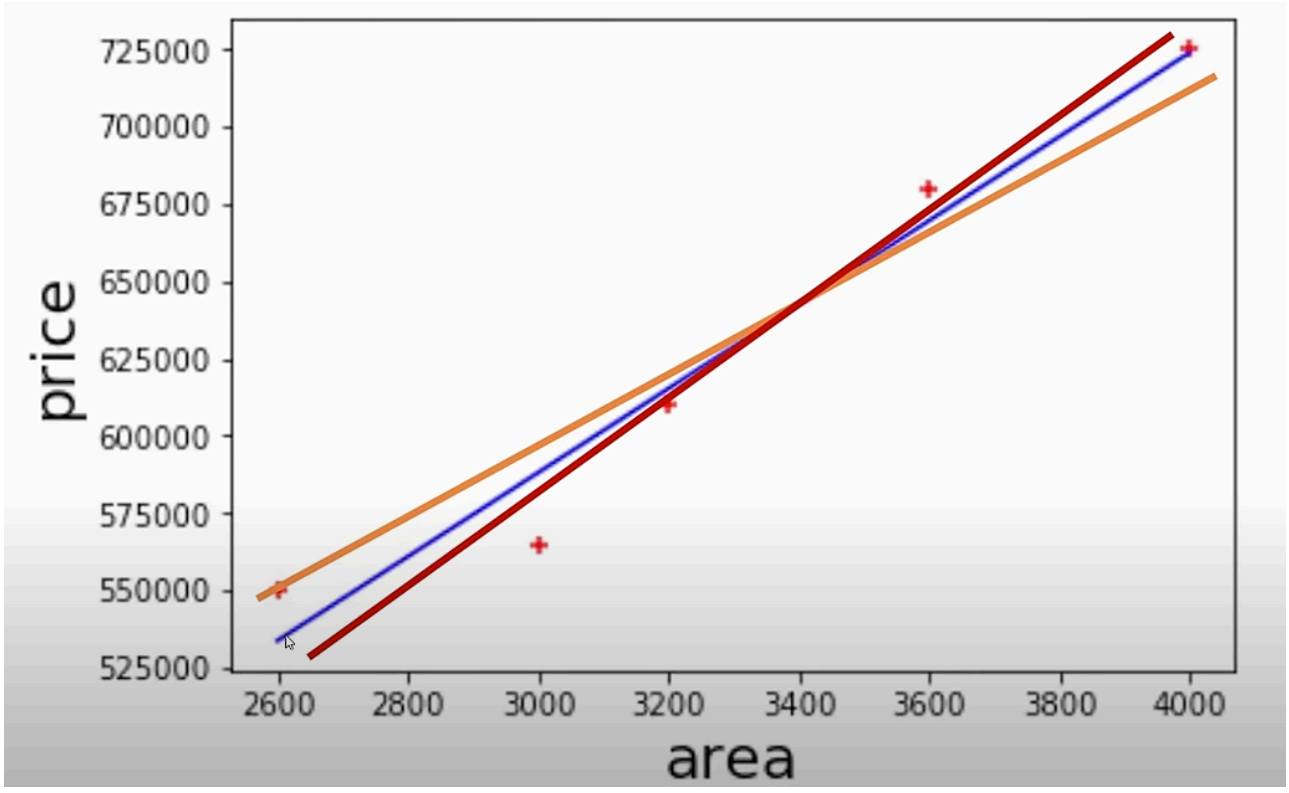
3300 square feet
5000 square feet

In linerare regression we plot a linear line through all the available data points that best fits all the points.
This line then can be used to predict prices for other house sizes.

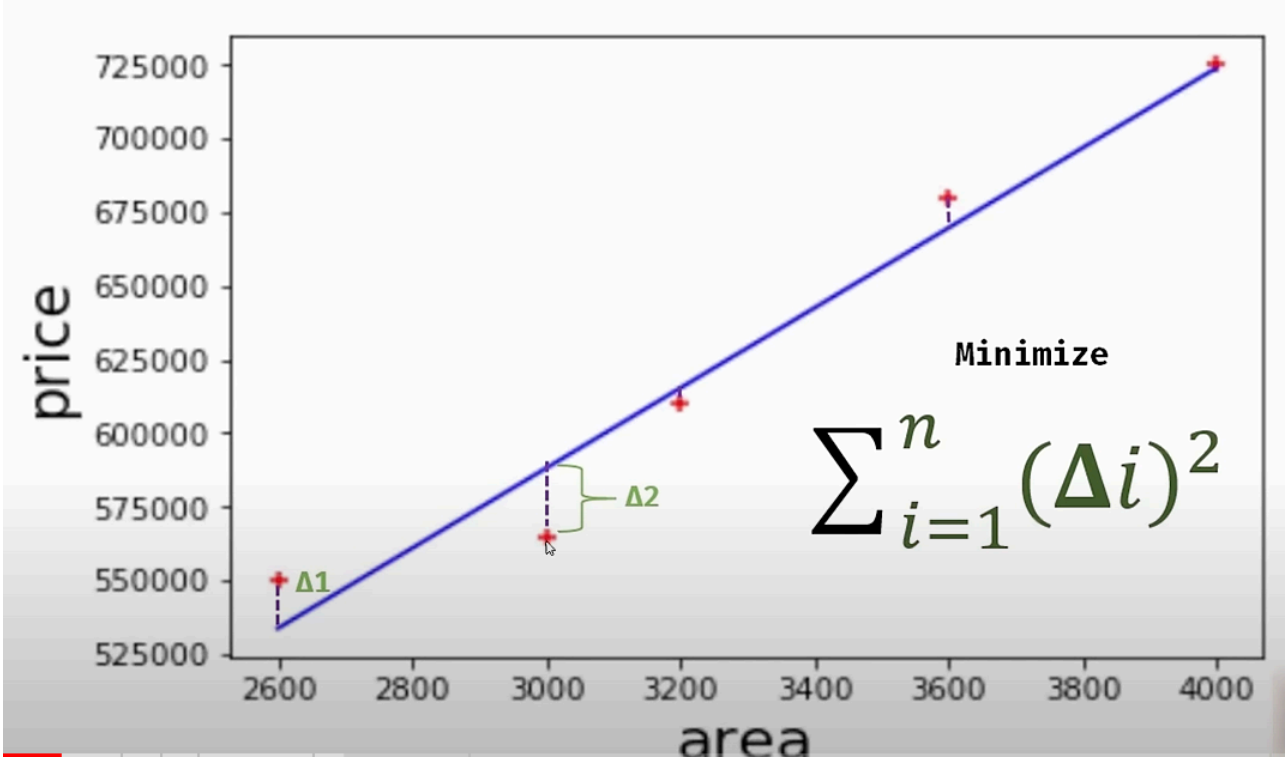
Linear line depicts the relationship between the dependednt and ondependent variables.
In this case price is

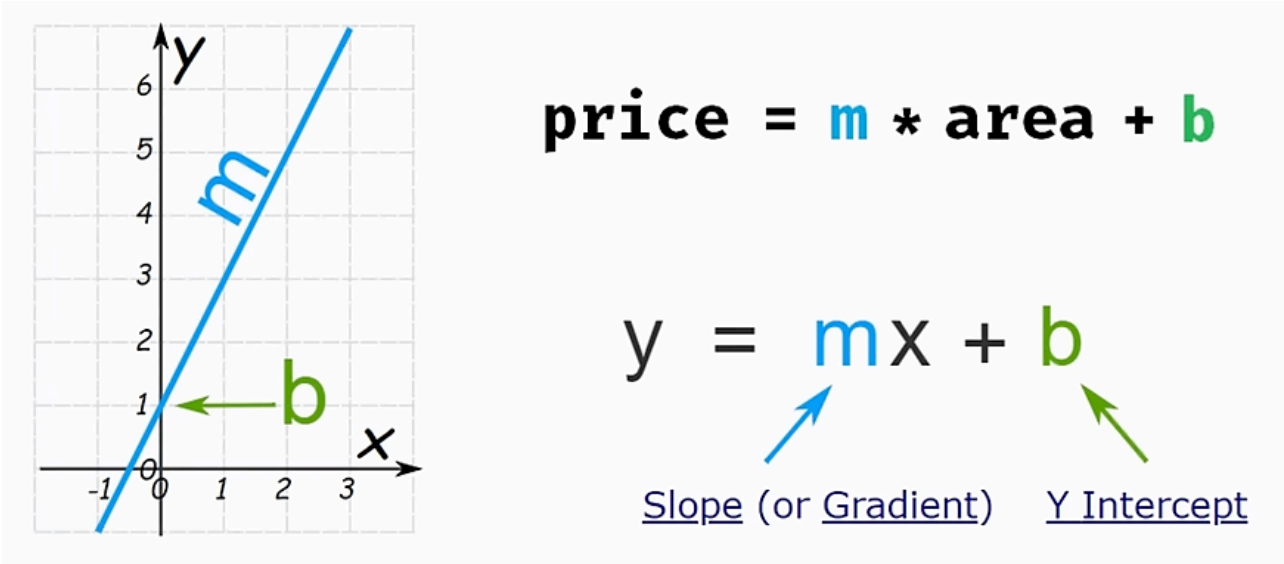


There can be ,multiple lines drawn through the available points but **we choose the one that best fits based on calculation of delta between the actual point and the line and then choose the line that has the lowest aggregated error.**



We get the error, square it and then minimize it, the line that has the minimum of this squared error is chosen as the final line and it would have a equation.

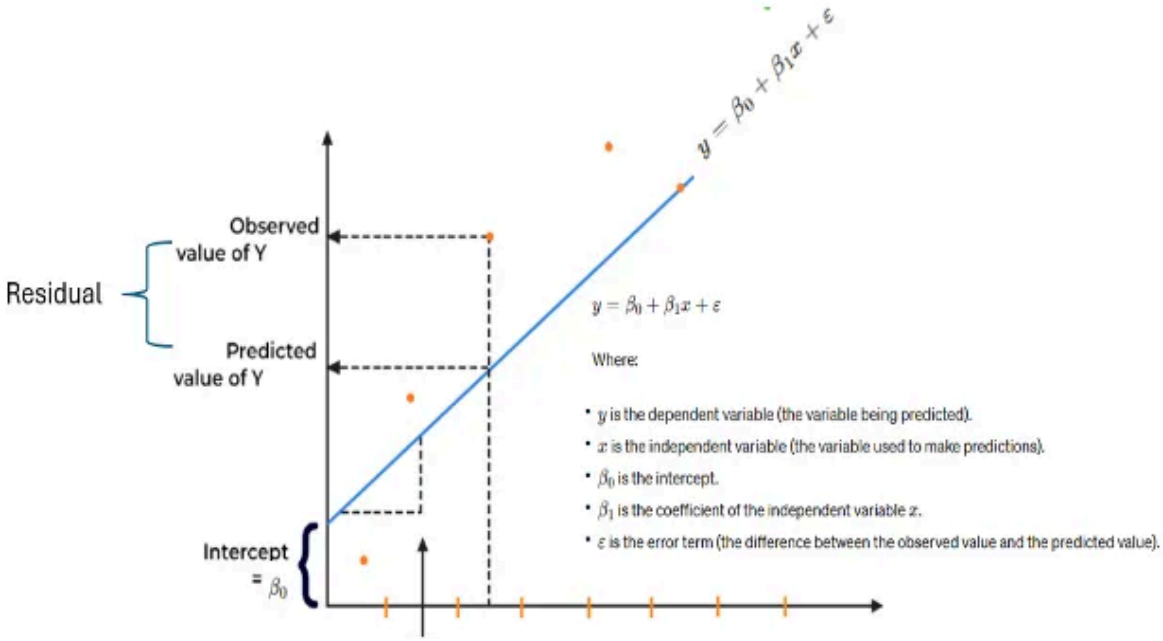




Linear equation -
Area is independent variable
Price is dependent as price is computed from area.

B represents intercept, meaning the point at which the line will touch the y axis means y coordinate of (0,y) point.
Say (0,-1) - so intercept is -1

https://www.dataversity.net/understanding-linear-regression-intercepts-in-plain-language/



Technically, the intercept in the linear regression model can be positive, negative, or even zero.

1. **Positive Intercept:** If the intercept in the regression model is positive, it means that the predicted value of the dependent variable (Y) when the independent variable (X) is zero is positive. This implies that the regression line crosses the y-axis above the zero value.
2. **Negative Intercept:** Conversely, if the intercept in a linear regression model is negative, it means that the predicted value of Y when X is zero is negative. In this case, the regression line crosses the y-axis below the zero value.
3. **Zero Intercept:** If the intercept in a regression model is zero, it implies that the regression line passes through the origin (0,0) on the graph. This means that the predicted value of the dependent variable is zero when all independent variables are also zero. In other words, there is no additional constant term in the regression equation. This situation is extremely rate and very theoretical.

Basically, you deal with negative or positive intercepts, and when you come across the negative intercept you deal with the negative intercept the same way as you would deal with a positive intercept. But in practical terms, a negative intercept may or may not make sense depending on the context of the data being analyzed. For example, if you are analyzing the day's temperature (X) and sales of ice cream (Y), a negative intercept would not be meaningful since it is impossible to have negative sales. However, in other domains such as financial analysis, a negative intercept could make sense.

M represents slope means , the angle at which line is tilted.
Slop is also called coefficient


Dependent variable
Independent variable