**Linear Regression Model**

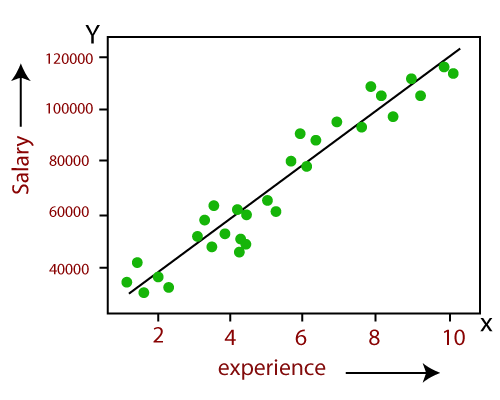
Linear regression models are used to show or predict the relationship between two [variables or factors](https://www.thebalancesmb.com/quantitative-research-methods-using-cross-tabs-2297160). The factor that is being predicted is called thedependent variable. The factors that are used to predict the value of the dependent variable are called the independent variables. In [this simple model](https://www.thebalancesmb.com/quantitative-research-using-structural-equation-modeling-2297146), a straight line approximates the relationship between the dependent variable and the independent variable.

A linear regression line has an equation of the form **Y = a + bX**, where X is the explanatory variable and Y is the dependent variable. Linear regression quantifies the relationship between one or more predictor variable(s) and one outcome variable. Linear regression is commonly used for predictive analysis and modelling. For example, it can be used to quantify the relative impacts of age, gender, and diet (the predictor variables) on height (the outcome variable).  Linear regression is also known as multiple regression*,* multivariate regression*,* ordinary least squares (OLS)*,* andregression*.*

When two or more independent variables are used in regression analysis, the model is no longer a simple linear one. This is known as *multiple regression*.

Linear-regression models are relatively simple and provide an easy-to-interpret mathematical formula that can generate predictions. Linear regression can be applied to various areas in business and academic study.

The output for Linear regression should only be the continuous values such as price, age, salary, etc. The relationship between the dependent variable and independent variable can be shown in below image:



In above image the dependent variable is on Y-axis (salary) and independent variable is on x-axis(experience)

Linear Regression is one of the most simple Machine learning algorithm that comes under Supervised learning technique and used for solving regression problems. Linear regression is used to predict the continuous dependent variable using a given set of independent variables.

Before attempting to fit a linear model to observed data, a modeler should first determine whether or not there is a relationship between the variables of interest. This does not necessarily imply that one variable causes the other (for example, higher SAT scores do not cause higher college grades), but that there is some significant association between the two variables. A [scatterplot](http://www.stat.yale.edu/Courses/1997-98/101/scatter.htm) can be a helpful tool in determining the strength of the relationship between two variables. If there appears to be no association between the proposed explanatory and dependent variables (i.e., the scatterplot does not indicate any increasing or decreasing trends), then fitting a linear regression model to the data probably will not provide a useful model. A valuable numerical measure of association between two variables is the [correlation coefficient](http://www.stat.yale.edu/Courses/1997-98/101/correl.htm), which is a value between -1 and 1 indicating the strength of the association of the observed data for the two variables.