

Introduction  
regression analysis  
refers to a tool that



**ASSIGN  
BUSTER**

## **Introduction**

Regression analysis refers to a tool that is used in statistics to establish a relationship between variables, two or more.

The aforementioned variables are quantitative, and they include the explanatory variable, also known as the independent variable, and the dependent variable. This relationship between the aforementioned variables is usually represented in form of an equation, or it can even be represented graphically. The function of regression analysis is thus to establish a relationship between the variables using information available about the explanatory variable, and predict values of the dependent variable for decision making. It is therefore apparent that regression analysis is a very useful forecasting tool.

## **Types of regression**

There are two main types of regression. These are simple regression and multiple regression. In both simple and multiple regression, a regression model is constructed which is presumed to follow a certain distribution. Therefore, regression models can be developed with linear, exponential, or even logarithmic equations.

## **Simple regression**

This is the type of regression in which there is only one dependent variable and one independent variable. That is, the regression relationship is based on only two variables. In this case, the explanatory variable is the only unknown in the right side of the regression equation. An example of a linear regression model for simple regression is:  $\hat{Y} = b_0 + b_1X$  Where Y is the <https://assignbuster.com/introduction-regression-analysis-refers-to-a-tool-that/>

dependent variable,  $X$  is the explanatory variable, and  $b_0$  and  $b_1$  are constants.

The constant  $b_0$  and  $b_1$  are determined by using historical values of  $X$  and  $Y$ . This gives them values, and therefore, future values of  $Y$  can be determined if one has the value of  $X$ . To exemplify the applicability of simple regression, consider a case in which population is assumed to have a linear relationship with time. In this case, population can be represented by the variable  $Y$ , and time can be represented by the variable  $X$ . Historical values of population can be used to get values for the constants, making predictions of future population possible. If the stated historical values are plotted,  $b_0$  is given by the  $y$ -intercept of the graph, while the constant  $b_1$  is given by the gradient of the resultant line.

## **Multiple regression**

Multiple regression is the regression in which there is more than one explanatory variable.

An example of a linear regression model for multiple regression is:  $Y = b_0 + b_1X_1 + b_2X_2$  Where  $Y$  is the dependent variable,  $X_1$  and  $X_2$  are explanatory variables/predictor variables, and  $b_0$ ,  $b_1$  and  $b_2$  are constants. The aforementioned constants are determined as in the case of linear regression above, although in multiple regression, multiple predictor variables are used. An example is a case in which population in a given year is assumed to be predicted by time, and number of births in the previous year. In such a case,  $Y$  will be the population, and either  $X_1$  or  $X_2$  will represent time or number of births in the previous year. Such a variable, like the year, that make data to

change, are known as dummy variables. Another example of a dummy variable is gender. Standard deviation, which is denoted by  $S_{yx}$ , is calculated using the historical values of  $Y$  and  $X$ , and it represents the error of the regression equation, which is measured with a specific confidence level.

Theoretically, it is wise to take large samples as they have less error.  $R^2$ , known as the true coefficient, is used to check the variability of the sample data with the regression line.

### **Other applications of regression**

Regression can be applied in a myriad of other situations. For instance, it is very common in businesses, where business indicators are predicted using business drivers. A good example of this is a case in which the sales of a business enterprise can be predicted using the amount of money invested in advertising. This is just a simple example.

A real situation will have the sales volume being predicted by multiple variables by the use of the principles of multiple regression.