**Dr. Alok Kumar**

**Regression**



**Regression** is stepping back or returning to the average value. The term was first used by British biometrician Sir Francis Galton in the later part of the 19th century in connection with some studies he made on estimating the extent to which the stature of sons of tall parents reverts or regresses back to the mean stature of the population. He studied the relationship between the heights of about one thousand fathers and sons and published the results in a paper *‘Regression towards Mediocrity in Heriditary Stature’.*

Regression is a statistical method used in finance, investing, and other disciplines that attempts to determine the strength and character of the relationship between one dependent variable (usually denoted by Y) and a series of other variables (known as independent variables).

Regression helps investment and financial managers to value assets and understand the relationships between variables, such as [commodity prices](https://www.investopedia.com/terms/c/commodity.asp) and the stocks of businesses dealing in those commodities.

Sir Francis Galton identified some interesting features of his study as follows:-

1. The tall fathers have tall sons and short fathers have short sons.
2. The average height of the sons of group of tall fathers is less than that of the fathers and the average height of the sons of a group of short fathers is more than that of the fathers.

In other words, Galton’s studies revealed that the off springs of abnormally tall or short parents tend to revert or step back to the average height of the population, a phenomenon which Galton described as regression to mediocrity.

He concluded that if the average height of a certain group of fathers is ‘a’ cms. Above (below) the general average height then average height of their sons will be (a\*r) cms, above (below) the general average height where r is the correlation coefficient between the height of the given group of fathers and their sons. In this case correlation is positive and since [r] <= 1 we have a\*r<=a. This supports the result in (ii) above.

But today the word regression as used in Statistics has a much wider perspective without any reference to biometry.

**Types of Regression** :-



[Linear regression](https://en.wikipedia.org/wiki/Linear_regression) is a statistical model that examines the linear relationship between two (Simple Linear Regression ) or more (Multiple Linear Regression) variables — a dependent variable and independent variable(s). Linear relationship basically means that when one (or more) independent variables increases (or decreases), the dependent variable increases (or decreases) too:



a linear relationship can be positive (independent variable goes up, dependent variable goes up) or negative (independent variable goes up, dependent variable goes down).