**Dr. Alok kumar**

**Correlation Vs Regression**

Correlation and Regression are the two analysis based on multivariate distribution. A multivariate distribution is described as a distribution of multiple variables. **Correlation** is described as the analysis which lets us know the association or the absence of the relationship between two variables ‘x’ and ‘y’. On the other end, **Regression** analysis, predicts the value of the dependent variable based on the known value of the independent variable, assuming that average mathematical relationship between two or more variables.

The difference between correlation and regression is one of the commonly asked questions in interviews. Moreover, many people suffer ambiguity in understanding these two.

| **Basis for Comparison** | **Correlation** | **Regression** |
| --- | --- | --- |
| Meaning | Correlation is a statistical measure which determines co-relationship or association of two variables. | Regression describes how an independent variable is numerically related to the dependent variable. |
| Usage | To represent linear relationship between two variables. | To fit a best line and estimate one variable on the basis of another variable. |
| Dependent and Independent variables | No difference | Both variables are different. |
| Indicates | Correlation coefficient indicates the extent to which two variables move together. | Regression indicates the impact of a unit change in the known variable (x) on the estimated variable (y). |
| Objective | To find a numerical value expressing the relationship between variables. | To estimate values of random variable on the basis of the values of fixed variable. |



In summary, correlation and regression have many similarities and some important differences. Regression is primarily used to build models/equations to predict a key response, Y, from a set of predictor (X) variables. Correlation is primarily used to quickly and concisely summarize the direction and strength of the relationships between a set of 2 or more numeric variables.