MULTICOLLINEARITY

- Multicollinearity occurs when there are linear relationships between the independent variables.
- This causes issues in regression analysis because it does not follow the assumption of independence among predictors.

Causes of Multicollinearity:

1. Correlation among predictor variables:

    * Predictor variable = independent variable
    * In regression model, predictor variables exhibit highly correlated with each other. In this situation, predictor variable is used to find the outcome variable so Multicollinearity occurs.

2. Over parameterization of the model: Too many predictor variables

3. Data Collection Issues: Problem in the data collection process can cause Multicollinearity.

Detection of Multicollinearity:

1. Correlation Matrices:

* This involves calculating the correlation coefficients between pair of predictor variables.
* High correlation coefficients (close to -1 to +1) indicate strong linear relationship between the variables causes Multicollinearity.

2. Variance Inflation Factor (VIF):

* $VIF = 1/ 1- R_i^2$ R – correlation coefficient
* If VIF = 1, it is not correlated, between 1 to 5 (some uses 10) = moderately correlated, More than 5 = highly correlated.