**WAYS TO HANDLE MULTICOLLINEARITY**

Remove one or more of the independent variables from the model, use a different statistical method, such as ridge regression or LASSO regression, to fix multicollinearity, one can remove one of the highly correlated variables, combine them into a single variable, or use a dimensionality reduction technique such as principal component analysis to reduce the number of variables while retaining most of the information.

**Variance inflation factor**

* Fit a Multiple Linear Regression Model: Start by fitting a multiple linear regression model using all predictor variable

​

* Check for High VIF Values: Examine the VIF values for each predictor variable. Typically, a VIF value above 10 or 5 indicates a high degree of multicollinearity.
* If high VIF values are found, consider the following approaches to address multicollinearity:
* Drop predictor variables with high VIF values from the model, especially those with the highest VIF values. This reduces multicollinearity as these variables are likely highly correlated with other predictors.
* Sometimes, highly correlated variables can be combined into composite variables. For instance, if two variables are highly correlated, you could create a new variable that represents the average or sum of the two.
* Use regularization techniques such as Ridge Regression or Lasso Regression, which introduce a penalty term to the regression equation, helping to reduce the impact of multicollinearity.

**Lasso regression**

* In the context of multicollinearity, this can effectively eliminate redundant predictors from the model, as highly correlated predictors tend to have similar coefficients.
* When one of these predictors has its coefficient forced to zero by the lasso penalty, the other correlated predictors can compensate for its influence, resulting in a more stable and interpretable model.

**Principal Component Analysis (PCA):**

* Principal Component Analysis (PCA) can indirectly handle multicollinearity by transforming the original correlated variables into a set of uncorrelated variables called principal components.
* When multicollinearity exists among the original variables, the principal components extracted by PCA are orthogonal to each other, meaning they are linearly independent and do not exhibit multicollinearity.
* In the context of Principal Component Analysis (PCA), a principal component is a linear combination of the original variables in a dataset.
* These components are derived in such a way that they are orthogonal to each other, meaning they are uncorrelated.
* Each principal component captures a certain amount of variance in the data, with the first component capturing the maximum variance, the second component capturing the maximum remaining variance orthogonal to the first, and so on.