**TASK 4 - DIMENSIONALITY REDUCTION**

**Dimensionality reduction** - or dimension reduction, is the transformation of data from a high-dimensional space into a low-dimensional space so that the low-dimensional representation retains some meaningful properties of the original data, ideally close to its intrinsic dimension.

In this task of dimensionality reduction task, we are going to use sparse PCA which is the type of PCA.

**SPARSE PCA** - Sparse principal component analysis is a specialised technique used in statistical analysis and in the analysis of multivariate data sets. It extends the classic method of principal component analysis for the reduction of dimensionality of data by introducing sparsity structures to the input variables.

So, in this task we are going to reduce dimension using sparse PCA on the wine quality dataset attached with this file with jupyter notebook file.

We have done following step: -

1. Importing basic and needed library.
2. Reading the Data set
3. Exploratory Data Analysis
4. Visualization
5. Sparse PCA
6. Standard scaling

In our analysis we have taken few steps in Sparse PCA part we use alpha, n\_component parameter to reduce the dimension of project. We also use Visualization such as pair plot, heatmap.

So, after doing this much of analysis we came to know the file number of columns using only alpha parameter gives us 11 columns but in reduced file we have 9 columns.

Thank you.