## **USL - Day 4 - Case Study**

## **About Dataset:**

## This dataset is adapted from the Wine Data Set from [Wine - UCI Machine Learning Repository](https://archive.ics.uci.edu/ml/datasets/wine) by removing the information about the types of wine for unsupervised learning.

The following descriptions are adapted from the UCI webpage:

These data are the results of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The analysis determined the quantities of 13 constituents found in each of the three types of wines.

**The attributes are:**

* Alcohol
* Malic acid
* Ash
* Alcalinity of ash
* Magnesium
* Total phenols
* Flavanoids
* Nonflavanoid phenols
* Proanthocyanins
* Color intensity
* Hue
* OD280/OD315 of diluted wines
* Proline

**Problem Statement:**

The problem statement is to analyze the data to create a cluster of wines based on their chemical constituents and predict the cluster labels obtained from the clustering algorithms using the supervised Learning classification techniques.

## 

## **Steps to Perform:**

## Research about the variables and understand them

## Analyze and prepare the data that can be best suitable to use for clustering-based algorithms that you have learned in the USL course

## Once you have prepared the data, use unsupervised techniques like clustering algorithms and dimensionality deduction if required to find out the optional number of clustering with the less number of variables by ensuring that minimal information is lost

## Finally, Use supervised Learning to predict the cluster labels obtained from the final clustering algorithm. Check the model performance using Recall, Precision, and F1 score

* Write the insights that you have derived from the case study