



1. Dataset Link: <https://www.kaggle.com/datasets/uciml/red-wine-quality-cortez-et-al-2009> Download the above dataset and perform the following operations:

- A. Fill the missing values with the respective mean.
- B. Create a new Column named **avg\_acidity** which takes the mean of the columns **fixed acidity** and **volatile acidity**.
- C. Delete the columns **fixed acidity** and **volatile acidity**.
- D. Drop duplicate rows
- E. Print first 20 rows and last 6 columns.

2. Create a dictionary with keys as Age, Gender and Qualification. Age should be between 20 and 50 with steps as 5, Gender should be M or F and make the list using random module. Same goes for qualification -> B-Tech, BSc or BA. Make a Dataframe out of it and print number of unique values for Gender and Qualification column.

3. Given an array of strings words, return the words that can be typed using letters of the alphabet on only one row of American keyboard.

In the American keyboard:

- the first row consists of the characters "qwertyuiop",
- the second row consists of the characters "asdfghjkl", and
- the third row consists of the characters "zxcvbnm".

```
Input: words = ["Hello", "Alaska", "Dad", "Peace"]  
Output: ["Alaska", "Dad"]
```

4. Given two integer arrays nums1 and nums2, return an array of their intersection. Each element in the result must be unique and you may return the result in any order.

```
Input: nums1 = [1,2,2,1], nums2 = [2,2]  
Output: [2]
```

5. For the given dataset, perform all the EDA techniques known with its inference. And train any ML model of your choice and show its respective evaluation metrics. Preferably train 3 models and compare them based on the metrics. If it has the case of class imbalance, apply SMOTE.

Dataset Link: <https://www.kaggle.com/datasets/parisrohan/credit-score-classification>