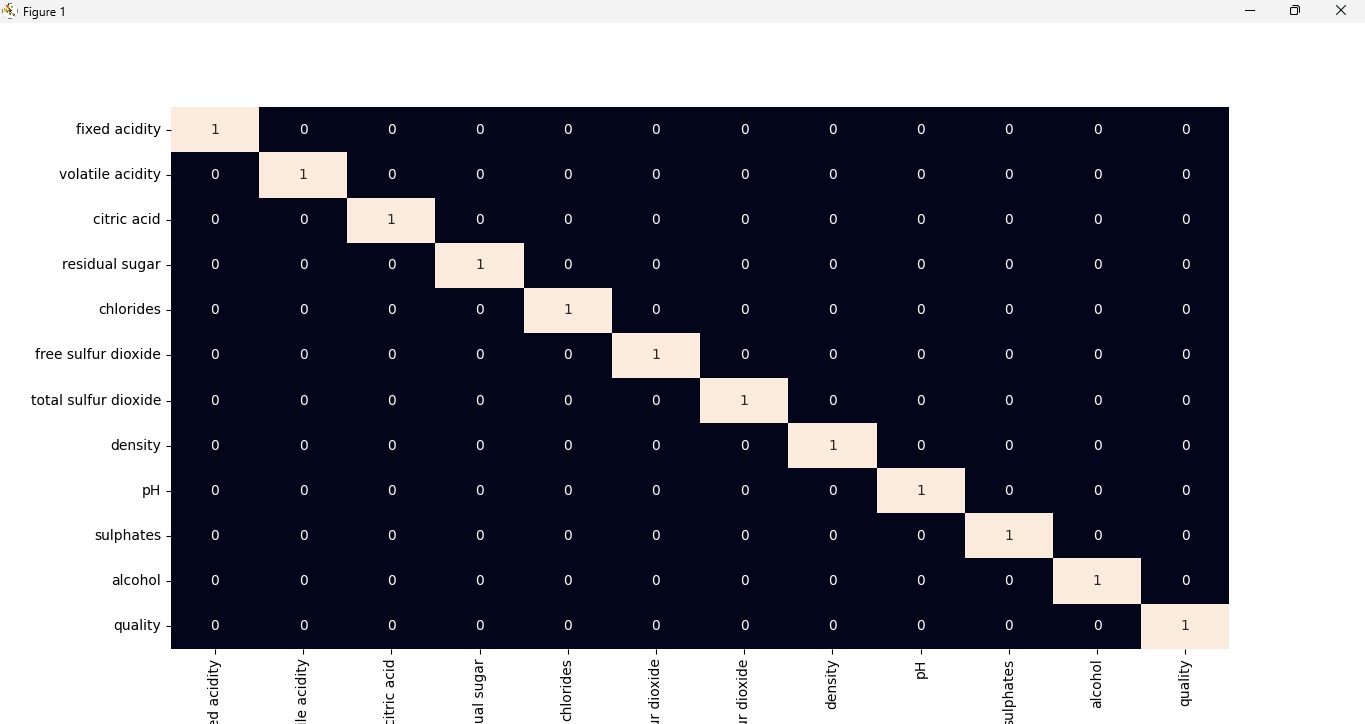
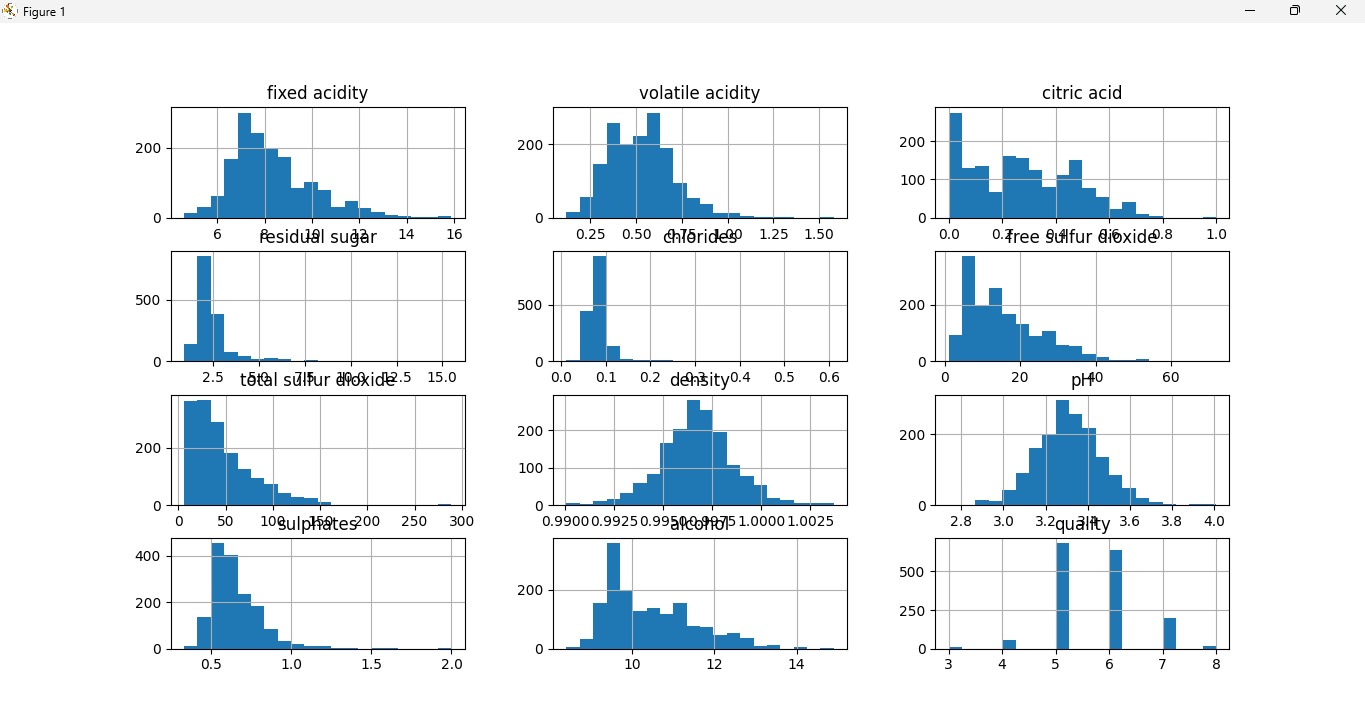
**README FILE**

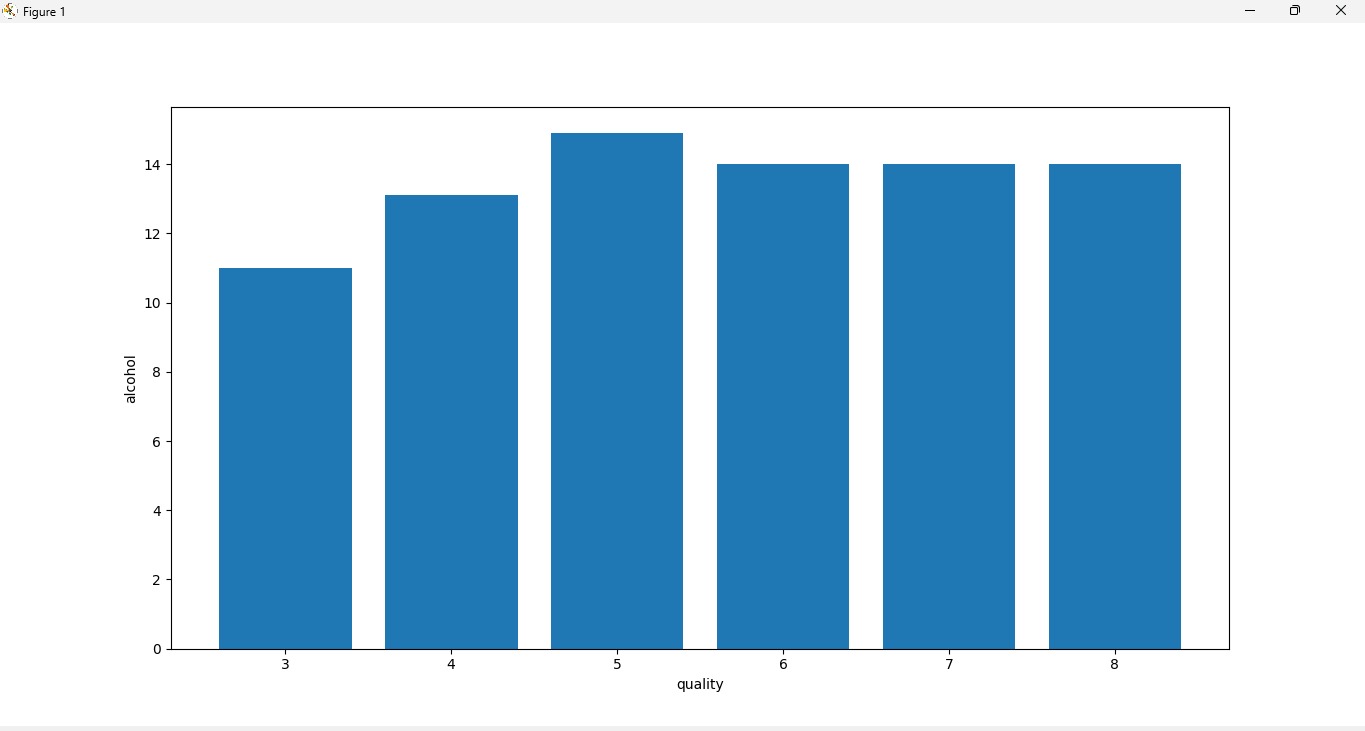
**WINE-1**

1.**Installed libraries**

* Pandas is a useful library in data handling.
* Numpy library used for working with arrays.
* Seaborn/Matplotlib are used for data visualization purposes.
* Sklearn – This module contains multiple libraries having pre-implemented functions to perform tasks from data preprocessing to model development and evaluation.
* XGBoost – This contains the eXtreme Gradient Boosting machine learning algorithm which is one of the algorithms which helps us to achieve high accuracy on predictions.
* import numpy as np

2.winequality.csv is the dataset for the file wine.py - write the path as per the file downloaded on our system.

3. After running this code your output should look like this.



**WINE-2**

1. We used Elbow Method to determine the number of clusters
2. PCA1 & PCA2- Principal Component Analysis is a technique that transforms high-dimension data into lower-dimension while retaining as much information as possible. It is used to interpret and visualize data. The number of variables decreases, which simplifies further analysis.
3. Effect of PCA1 & PCA2 on Clusters: If we really want to reduce the size of the dataset, the best number of principal components is much less than the number of variables in the original dataset.
4. After running this code your output should look like this.

