

Red Wine Quality Prediction

The dataset is taken from <https://www.kaggle.com/uciml/red-wine-quality-cortez-et-al-2009>

IMPORTING THE LIBRARIES

```
In [2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

LOAD THE DATASET

```
In [3]: dataset=pd.read_csv('C:/Users/HP/Desktop/coursera/project/Red_wine/red_
wine.csv')
```

EDA

View the size of the dataset

```
In [4]: dataset.shape
```

```
Out[4]: (1599, 12)
```

View the columns of the dataset

In [5]: `dataset.columns`

Out[5]: Index(['fixed acidity', 'volatile acidity', 'citric acid', 'residual sugar', 'chlorides', 'free sulfur dioxide', 'total sulfur dioxide', 'density', 'pH', 'sulphates', 'alcohol', 'quality'], dtype='object')

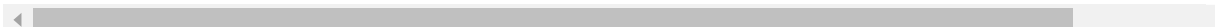
View few rows of the dataset

In [6]: `dataset`

Out[6]:

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol
0	7.4	0.700	0.00	1.9	0.076	11.0	34.0	0.99780	3.51	0.56	
1	7.8	0.880	0.00	2.6	0.098	25.0	67.0	0.99680	3.20	0.68	
2	7.8	0.760	0.04	2.3	0.092	15.0	54.0	0.99700	3.26	0.65	
3	11.2	0.280	0.56	1.9	0.075	17.0	60.0	0.99800	3.16	0.58	
4	7.4	0.700	0.00	1.9	0.076	11.0	34.0	0.99780	3.51	0.56	
...
1594	6.2	0.600	0.08	2.0	0.090	32.0	44.0	0.99490	3.45	0.58	
1595	5.9	0.550	0.10	2.2	0.062	39.0	51.0	0.99512	3.52	0.76	
1596	6.3	0.510	0.13	2.3	0.076	29.0	40.0	0.99574	3.42	0.75	
1597	5.9	0.645	0.12	2.0	0.075	32.0	44.0	0.99547	3.57	0.71	
1598	6.0	0.310	0.47	3.6	0.067	18.0	42.0	0.99549	3.39	0.66	

1599 rows × 12 columns



Get statistical info about the data

In [7]: `dataset.describe()`

Out[7]:

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide
count	1599.000000	1599.000000	1599.000000	1599.000000	1599.000000	1599.000000	1599.000000
mean	8.319637	0.527821	0.270976	2.538806	0.087467	15.874922	46.46779
std	1.741096	0.179060	0.194801	1.409928	0.047065	10.460157	32.89532
min	4.600000	0.120000	0.000000	0.900000	0.012000	1.000000	6.00000
25%	7.100000	0.390000	0.090000	1.900000	0.070000	7.000000	22.00000
50%	7.900000	0.520000	0.260000	2.200000	0.079000	14.000000	38.00000
75%	9.200000	0.640000	0.420000	2.600000	0.090000	21.000000	62.00000
max	15.900000	1.580000	1.000000	15.500000	0.611000	72.000000	289.00000

Get the information about the variables

In [8]: `dataset.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1599 entries, 0 to 1598
Data columns (total 12 columns):
fixed acidity      1599 non-null float64
volatile acidity   1599 non-null float64
citric acid        1599 non-null float64
residual sugar     1599 non-null float64
chlorides          1599 non-null float64
free sulfur dioxide 1599 non-null float64
total sulfur dioxide 1599 non-null float64
density            1599 non-null float64
pH                 1599 non-null float64
sulphates          1599 non-null float64
```

```
alcohol          1599 non-null float64
quality          1599 non-null int64
dtypes: float64(11), int64(1)
memory usage: 150.0 KB
```

Check for null values

```
In [9]: dataset.isnull().sum()
```

```
Out[9]: fixed acidity      0
        volatile acidity  0
        citric acid       0
        residual sugar    0
        chlorides         0
        free sulfur dioxide 0
        total sulfur dioxide 0
        density           0
        pH                0
        sulphates         0
        alcohol           0
        quality           0
dtype: int64
```

```
In [10]: dataset['quality'].unique()
```

```
Out[10]: array([5, 6, 7, 4, 8, 3], dtype=int64)
```

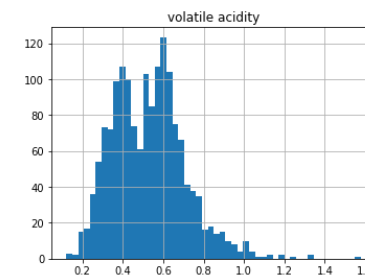
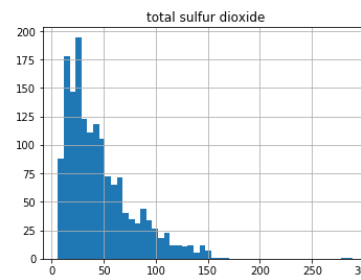
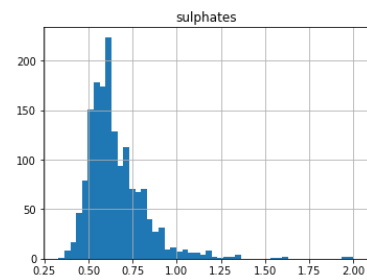
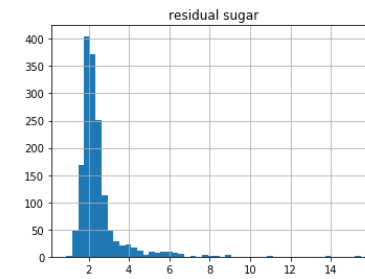
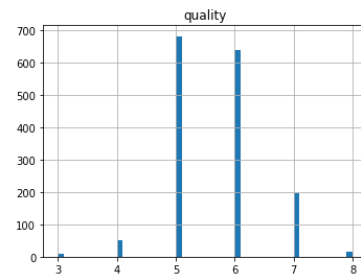
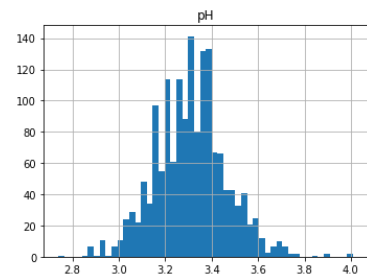
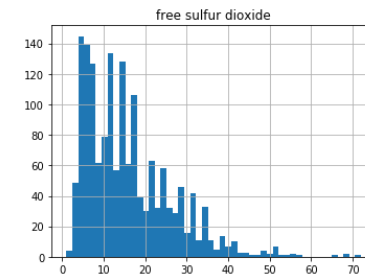
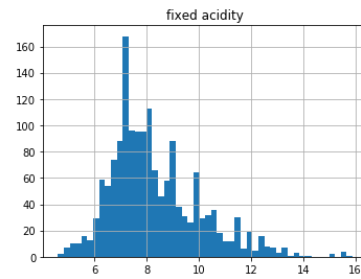
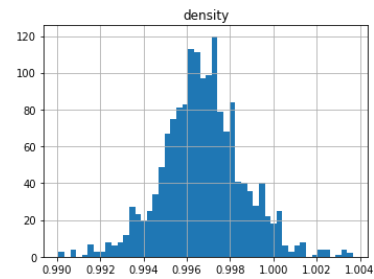
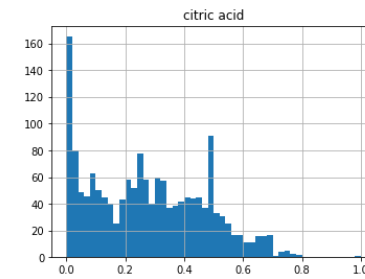
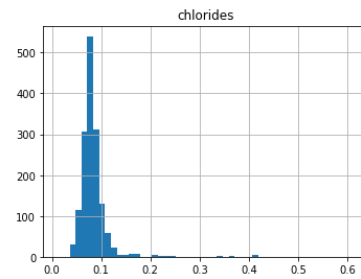
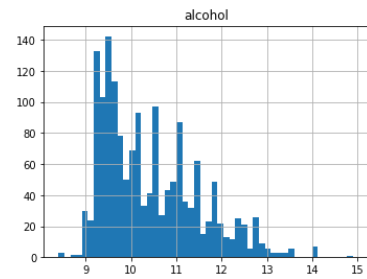
No missing data

PLOT FEW GRAPHS

```
In [11]: dataset.hist(bins=50,figsize=(20,20))
```

```
Out[11]: array([[<matplotlib.axes._subplots.AxesSubplot object at 0x000001F833B0
EA88>,
                 <matplotlib.axes._subplots.AxesSubplot object at 0x000001F83607
```

```
B608>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001F8360B
0F08>],
    [<matplotlib.axes._subplots.AxesSubplot object at 0x000001F8360E
D048>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001F83612
6108>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001F83615
A248>],
    [<matplotlib.axes._subplots.AxesSubplot object at 0x000001F83619
5348>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001F8361D
3F48>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001F8361D
B048>],
    [<matplotlib.axes._subplots.AxesSubplot object at 0x000001F83621
3248>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001F83627
7708>,
    <matplotlib.axes._subplots.AxesSubplot object at 0x000001F8362A
E808>]],
    dtype=object)
```



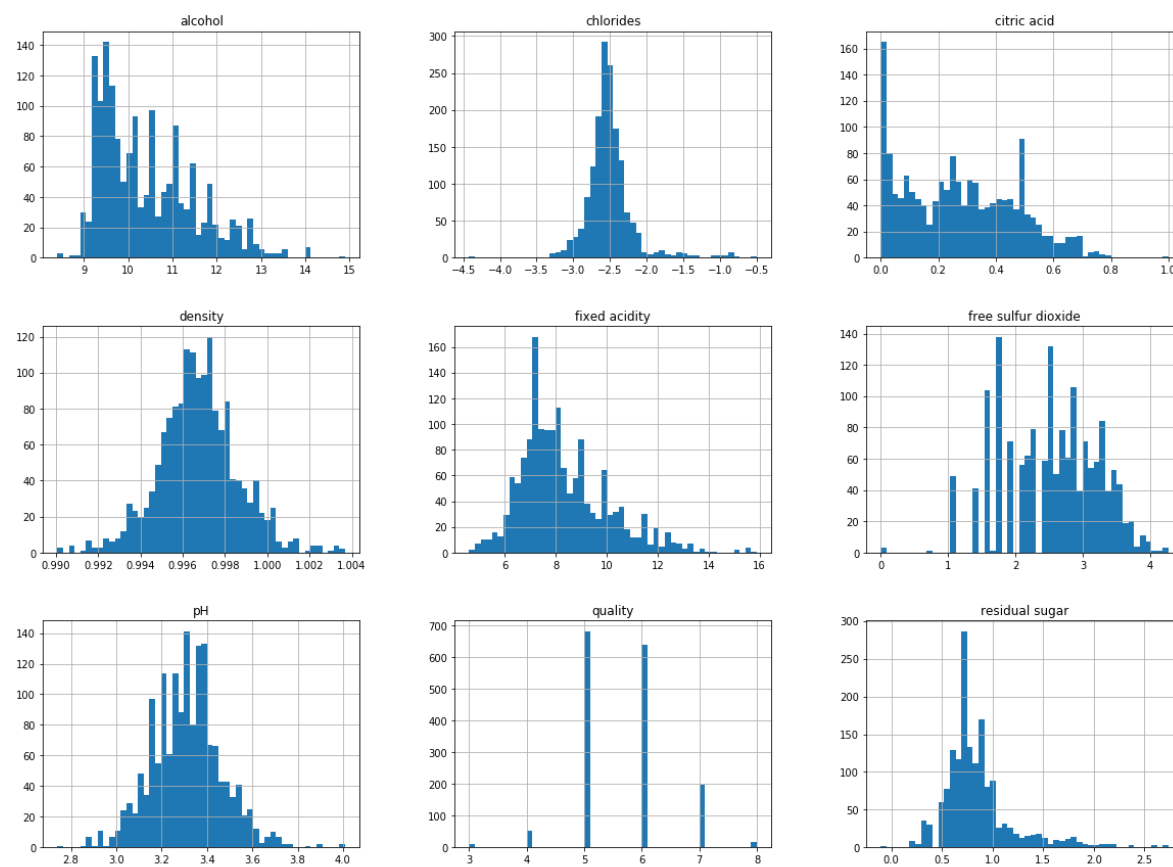
Residual Sugar, Chlorides, Free Sulphur Dioxide, Total Sulphur Dioxide, and Sulphates are highly right skewed. Apply logarithm

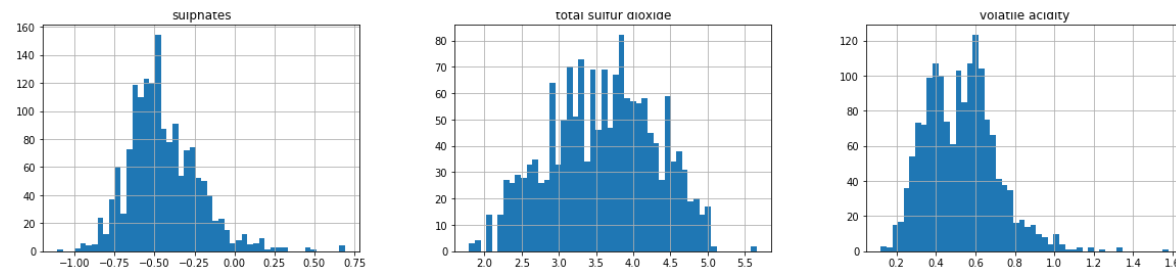
```
In [12]: def log_transform(col):  
         return np.log(col[0])  
  
dataset['residual sugar'] = dataset[['residual sugar']].apply(log_trans  
form, axis=1)  
dataset['chlorides'] = dataset[['chlorides']].apply(log_transform, axis  
=1)  
dataset['free sulfur dioxide'] = dataset[['free sulfur dioxide']].apply  
(log_transform, axis=1)  
dataset['total sulfur dioxide'] = dataset[['total sulfur dioxide']].app  
ly(log_transform, axis=1)  
dataset['sulphates'] = dataset[['sulphates']].apply(log_transform, axis  
=1)
```

```
In [13]: dataset.hist(bins=50,figsize=(20,20))
```

```
Out[13]: array([[<matplotlib.axes._subplots.AxesSubplot object at 0x000001F835  
DAB5C8>,  
               <matplotlib.axes._subplots.AxesSubplot object at 0x000001F837  
52AF08>,  
               <matplotlib.axes._subplots.AxesSubplot object at 0x000001F836  
A0BB48>],  
              [<matplotlib.axes._subplots.AxesSubplot object at 0x000001F836  
A23548>,  
               <matplotlib.axes._subplots.AxesSubplot object at 0x000001F836  
A3BF48>,  
               <matplotlib.axes._subplots.AxesSubplot object at 0x000001F836  
A71A08>],
```

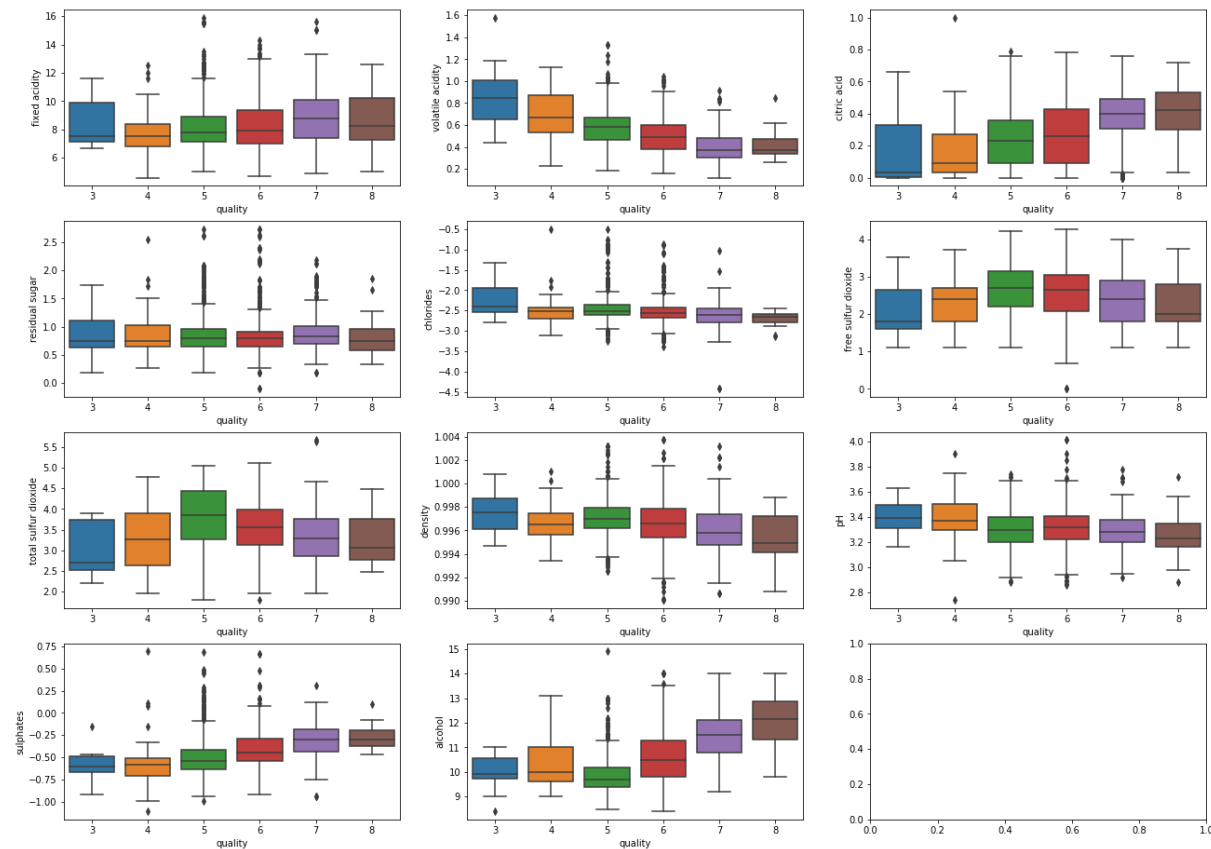
```
[<matplotlib.axes._subplots.AxesSubplot object at 0x000001F836
AABA88>,
 <matplotlib.axes._subplots.AxesSubplot object at 0x000001F836
AE2B88>,
 <matplotlib.axes._subplots.AxesSubplot object at 0x000001F836
AEF788>],
 [<matplotlib.axes._subplots.AxesSubplot object at 0x000001F836
B26988>,
 <matplotlib.axes._subplots.AxesSubplot object at 0x000001F836
B8CEC8>,
 <matplotlib.axes._subplots.AxesSubplot object at 0x000001F836
BC4FC8>]],
 dtype=object)
```





Check for outliers

```
In [14]: fig, ax1 = plt.subplots(4,3, figsize=(22,16))
          k = 0
          columns = list(dataset.columns)
          for i in range(4):
              for j in range(3):
                  if k != 11:
                      sns.boxplot(dataset['quality'], dataset[columns[k]], ax = a
x1[i][j])
                      k += 1
          plt.show()
```

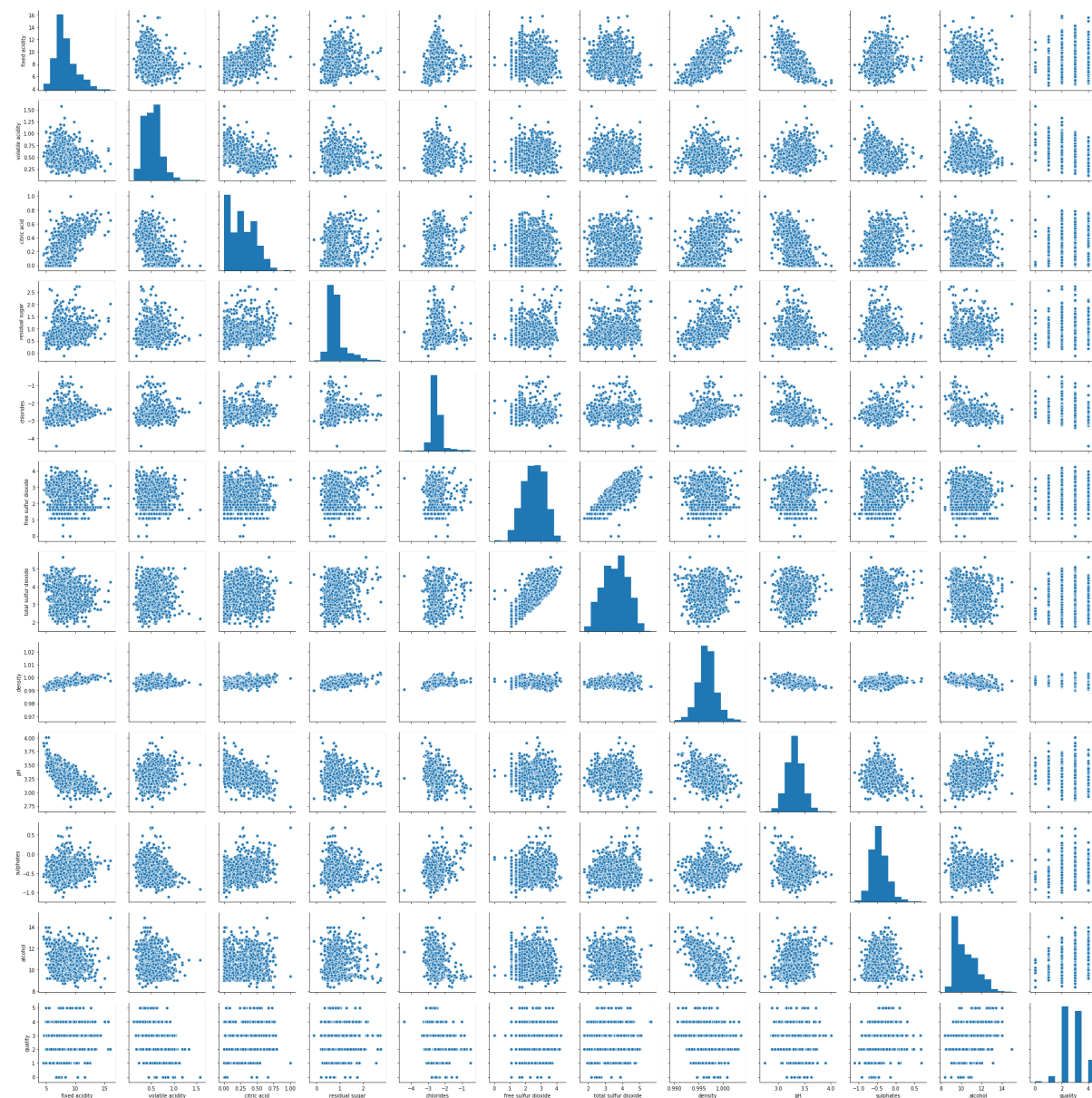


```
In [15]: def scale_outputs(col):
          return col[0] - 3

dataset['quality'] = dataset[['quality']].apply(scale_outputs, axis=1)
```

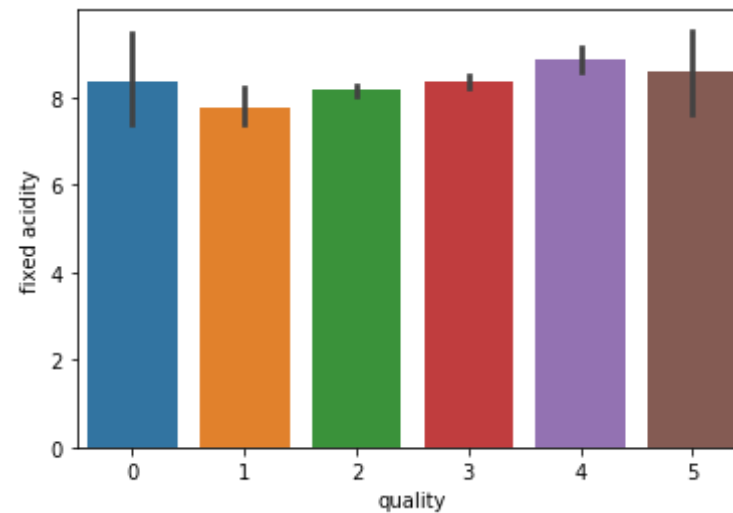
```
In [16]: sns.pairplot(dataset)
```

```
Out[16]: <seaborn.axisgrid.PairGrid at 0x1f837ed0dc8>
```



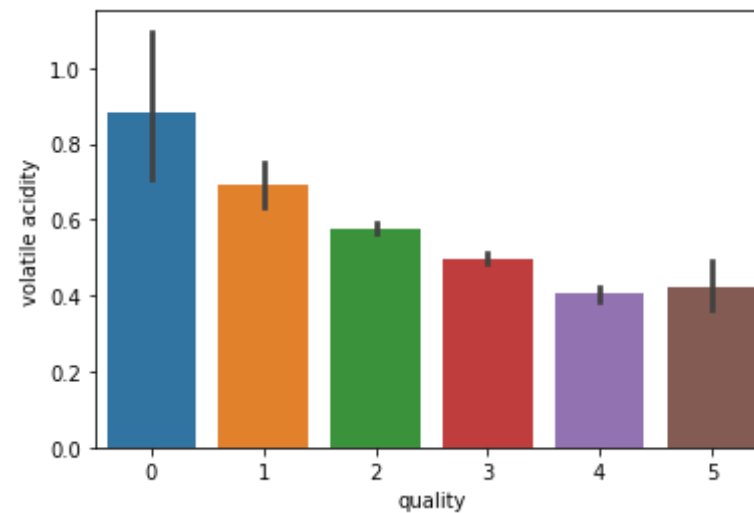
```
In [17]: sns.barplot(x='quality',y='fixed acidity',data=dataset)
```

```
Out[17]: <matplotlib.axes._subplots.AxesSubplot at 0x1f83f6e3248>
```



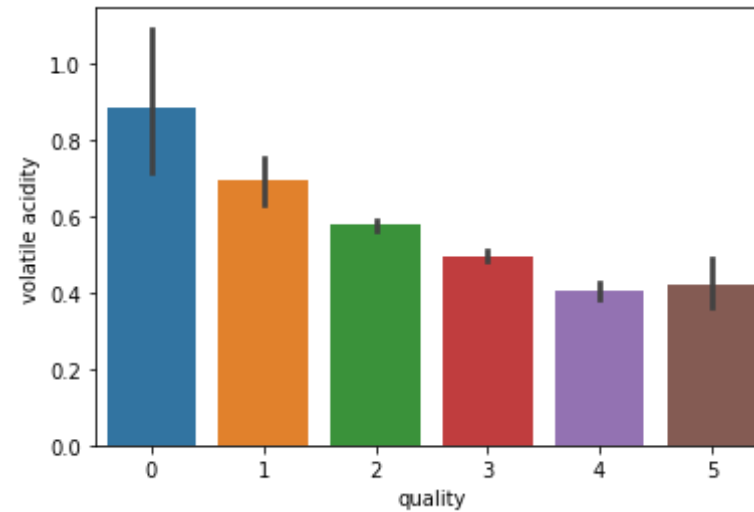
```
In [18]: sns.barplot(x='quality',y='volatile acidity',data=dataset)
```

```
Out[18]: <matplotlib.axes._subplots.AxesSubplot at 0x1f840701408>
```



```
In [19]: sns.barplot(x='quality',y='volatile acidity',data=dataset)
```

Out[19]: <matplotlib.axes._subplots.AxesSubplot at 0x1f84196c948>

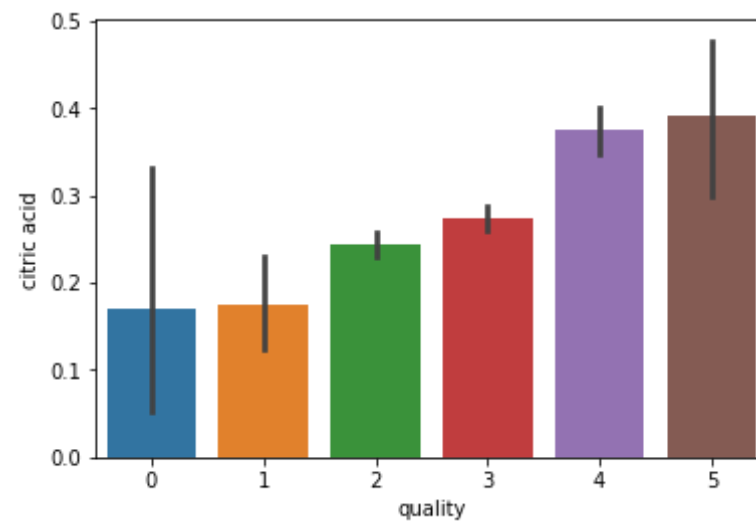


In [20]: `dataset.columns`

Out[20]: Index(['fixed acidity', 'volatile acidity', 'citric acid', 'residual sugar',
'chlorides', 'free sulfur dioxide', 'total sulfur dioxide', 'density',
'pH', 'sulphates', 'alcohol', 'quality'],
dtype='object')

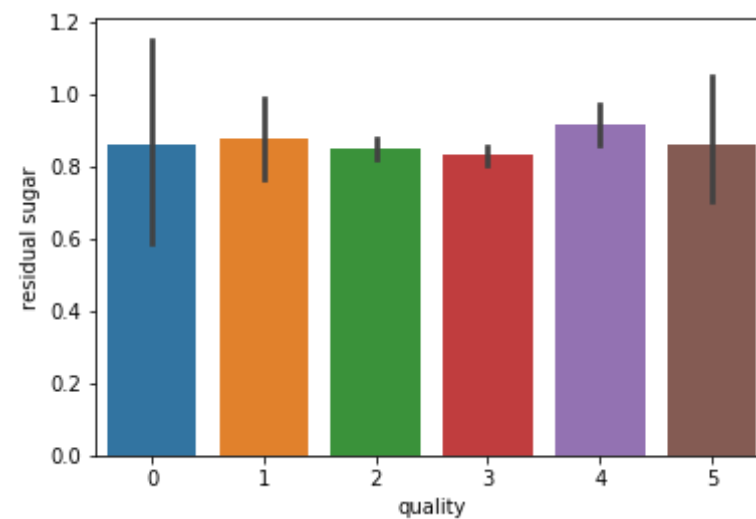
In [21]: `sns.barplot(x='quality',y='citric acid',data=dataset)`

Out[21]: <matplotlib.axes._subplots.AxesSubplot at 0x1f840777988>



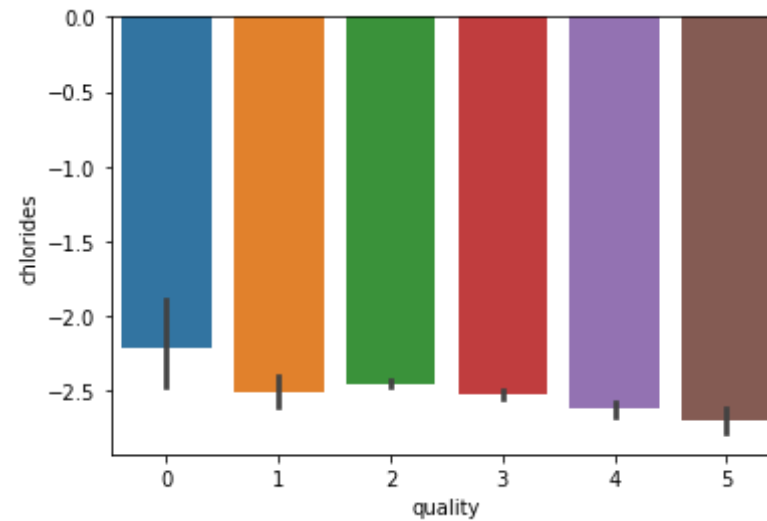
```
In [22]: sns.barplot(x='quality',y='residual sugar',data=dataset)
```

```
Out[22]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841a4de48>
```



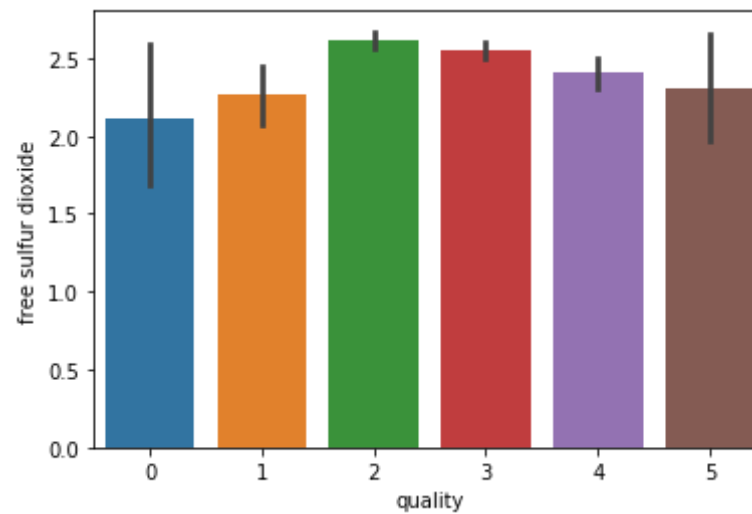
```
In [23]: sns.barplot(x='quality',y='chlorides',data=dataset)
```

Out[23]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841ad1708>



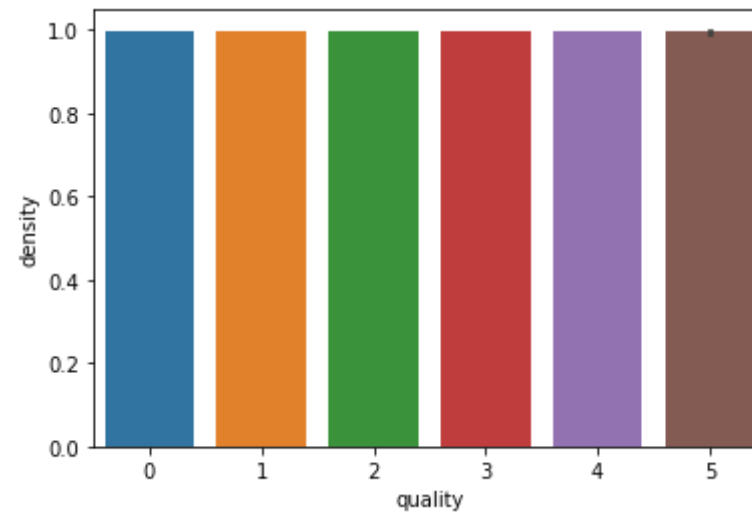
In [24]: `sns.barplot(x='quality',y='free sulfur dioxide',data=dataset)`

Out[24]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841b4a988>



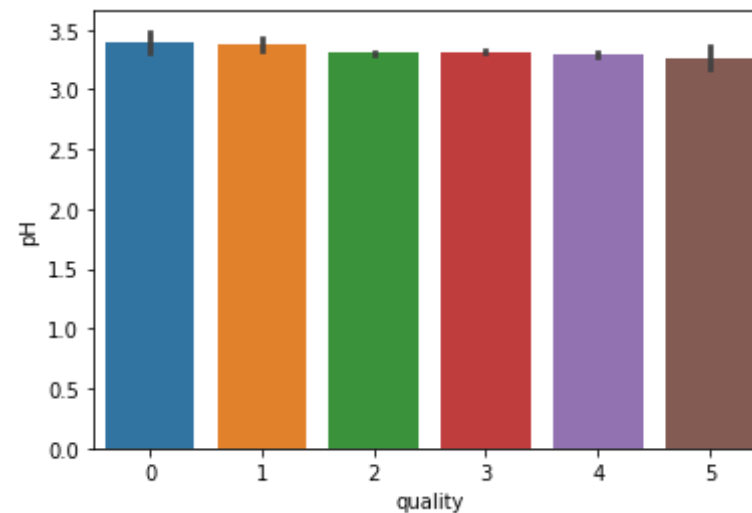
```
In [25]: sns.barplot(x='quality',y='density',data=dataset)
```

```
Out[25]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841a30588>
```



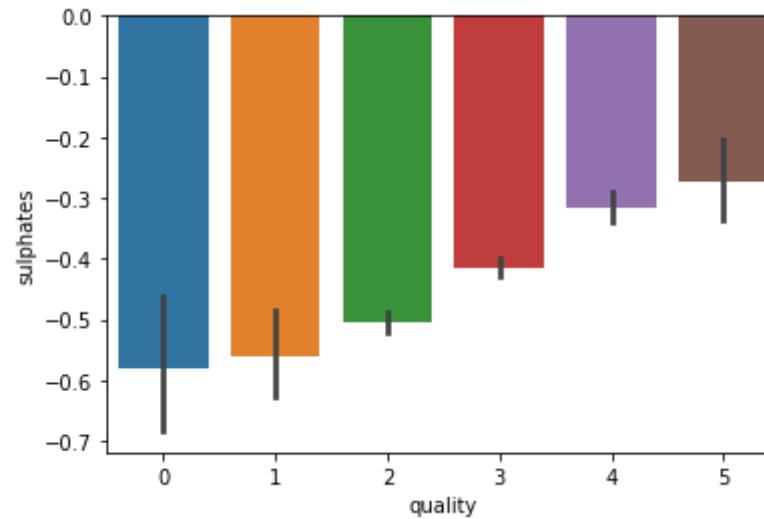
```
In [26]: sns.barplot(x='quality',y='pH',data=dataset)
```

```
Out[26]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841c3d888>
```



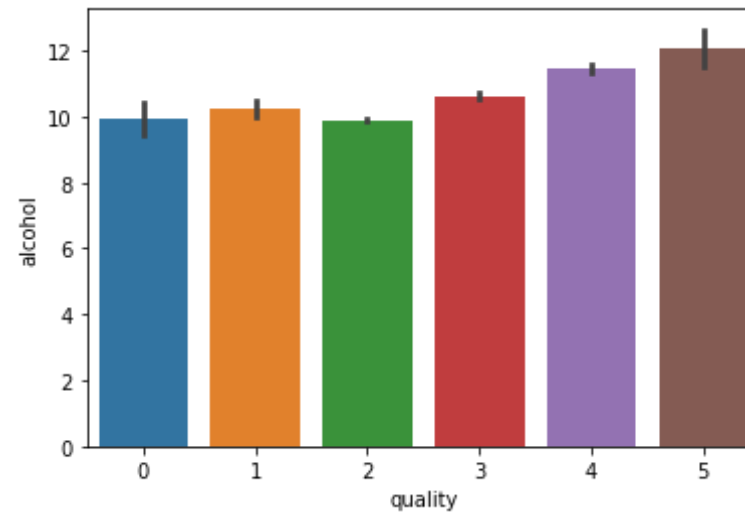

```
In [27]: sns.barplot(x='quality', y='sulphates', data=dataset)
```

```
Out[27]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841cc1988>
```



```
In [28]: sns.barplot(x='quality', y='alcohol', data=dataset)
```

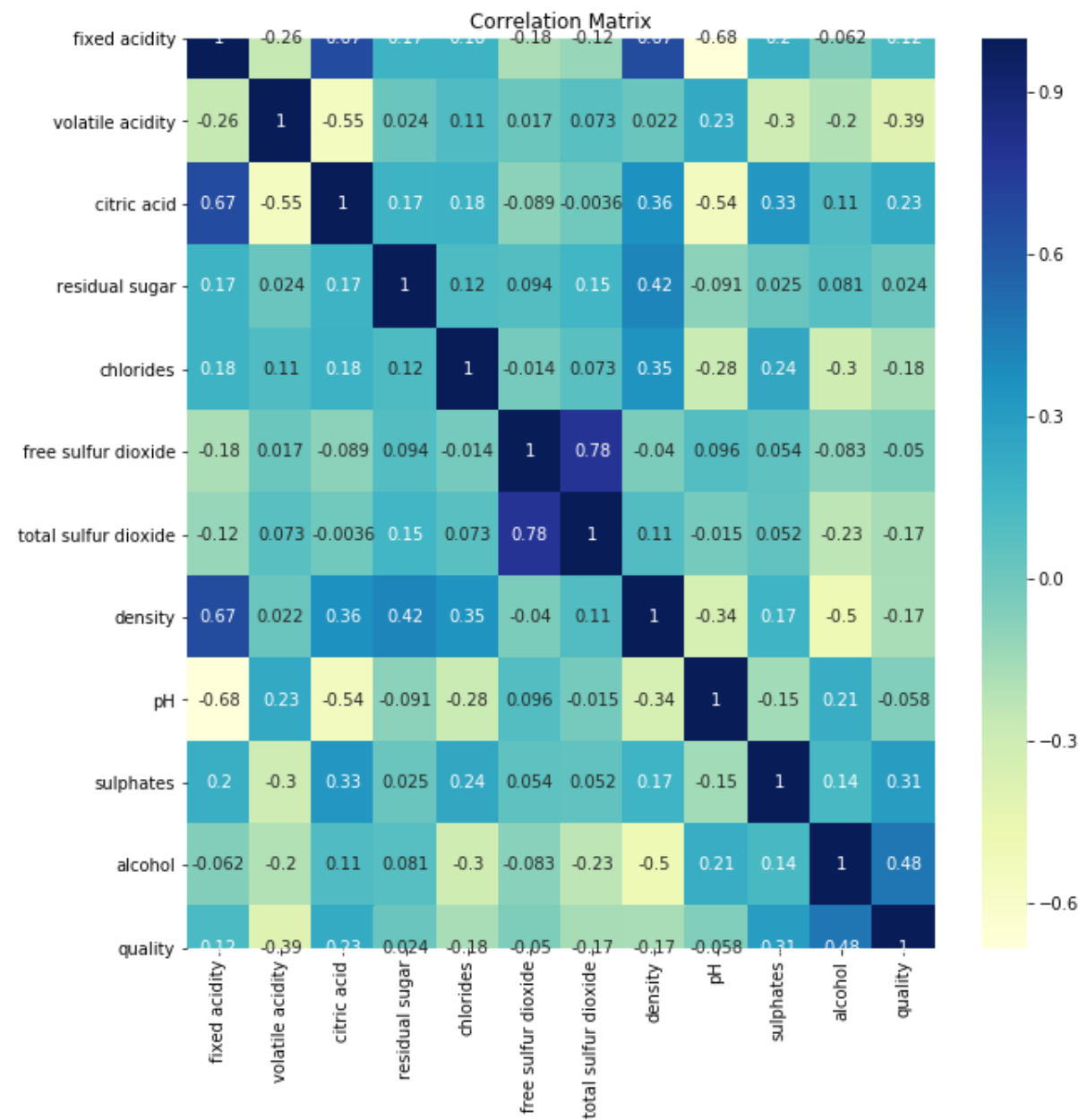
```
Out[28]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841d435c8>
```



CORRELATION MATRIX

```
In [29]: corr = dataset.corr()  
plt.figure(figsize=(10, 10))  
plt.title('Correlation Matrix')  
sns.heatmap(corr, cmap='YlGnBu', annot=True)
```

```
Out[29]: <matplotlib.axes._subplots.AxesSubplot at 0x1f841c23c88>
```



SPLITTING THE DATASET INTO FEATURES AND LABELS

```
In [31]: X=dataset.drop('quality',axis=1)
y=dataset['quality']
```

```
In [32]: X
```

```
Out[32]:
```

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates
0	7.4	0.700	0.00	0.641854	-2.577022	2.397895	3.526361	0.99780	3.51	-0.579818
1	7.8	0.880	0.00	0.955511	-2.322788	3.218876	4.204693	0.99680	3.20	-0.385662
2	7.8	0.760	0.04	0.832909	-2.385967	2.708050	3.988984	0.99700	3.26	-0.430783
3	11.2	0.280	0.56	0.641854	-2.590267	2.833213	4.094345	0.99800	3.16	-0.544727
4	7.4	0.700	0.00	0.641854	-2.577022	2.397895	3.526361	0.99780	3.51	-0.579818
...
1594	6.2	0.600	0.08	0.693147	-2.407946	3.465736	3.784190	0.99490	3.45	-0.544727
1595	5.9	0.550	0.10	0.788457	-2.780621	3.663562	3.931826	0.99512	3.52	-0.274437
1596	6.3	0.510	0.13	0.832909	-2.577022	3.367296	3.688879	0.99574	3.42	-0.287682
1597	5.9	0.645	0.12	0.693147	-2.590267	3.465736	3.784190	0.99547	3.57	-0.342490
1598	6.0	0.310	0.47	1.280934	-2.703063	2.890372	3.737670	0.99549	3.39	-0.415515

1599 rows × 11 columns



```
In [33]: y=pd.DataFrame(y)
y
```

```
Out[33]:
```

quality

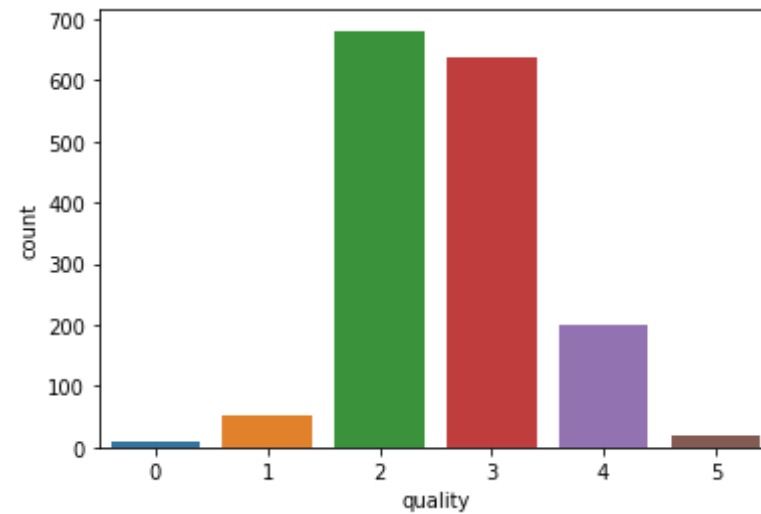
quality	
0	2
1	2
2	2
3	3
4	2
...	...
1594	2
1595	3
1596	3
1597	2
1598	3

1599 rows × 1 columns

Checking the distribution of the labels

```
In [34]: sns.countplot(y.quality)
```

```
Out[34]: <matplotlib.axes._subplots.AxesSubplot at 0x1f84220dac8>
```



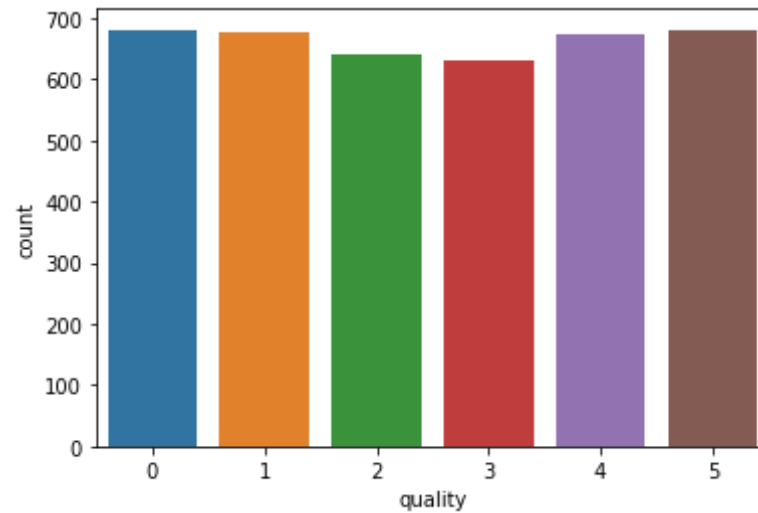
We find that the dataset is an unbalanced dataset

Using SMOTETomek to make the dataset balanced

```
In [ ]: from imblearn.combine import SMOTETomek  
sampler = SMOTETomek()  
X,y = sampler.fit_sample(X,y)
```

```
In [49]: sns.countplot(y.quality)
```

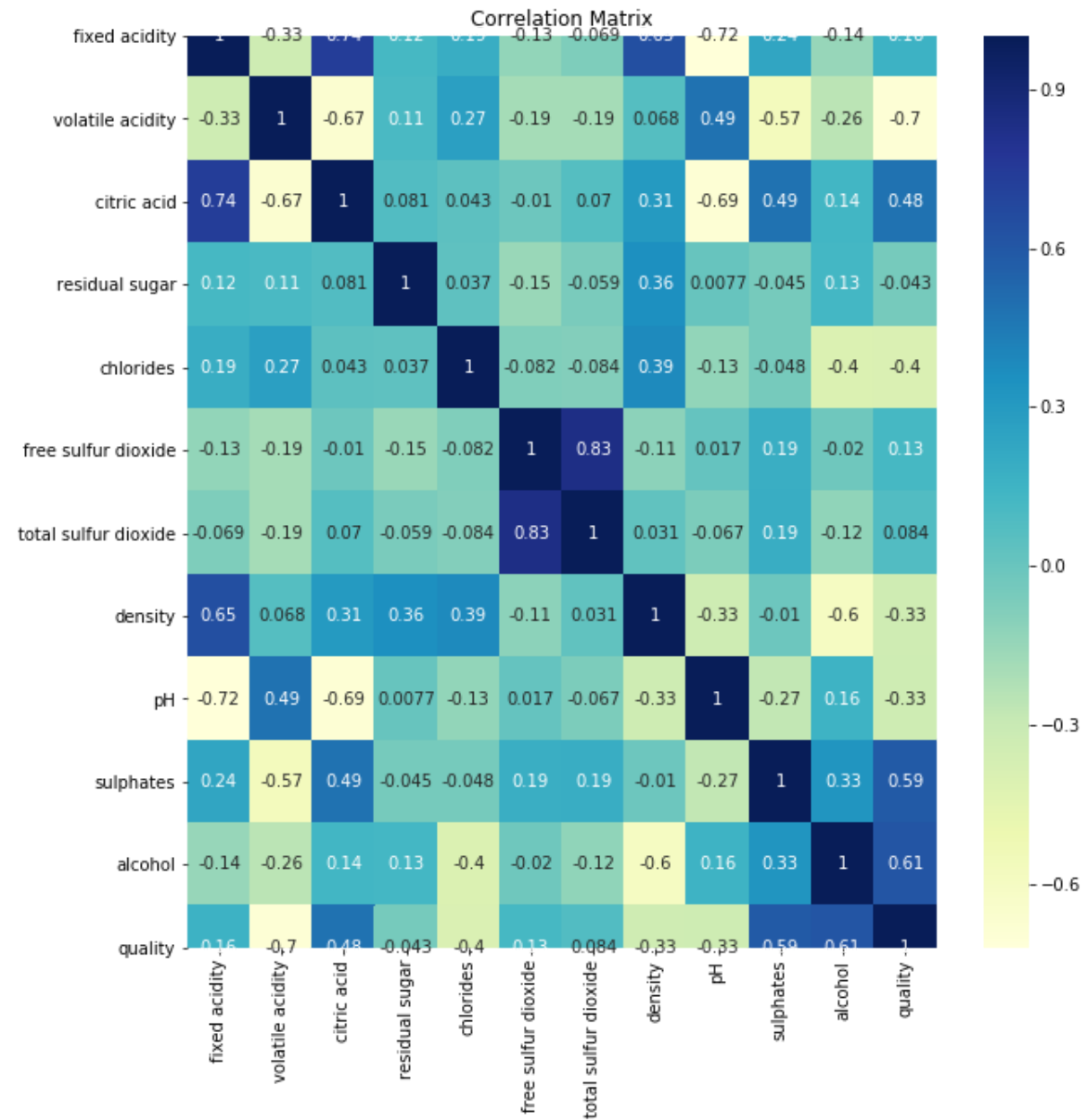
```
Out[49]: <matplotlib.axes._subplots.AxesSubplot at 0x1f84500d3c8>
```



Correlation matrix

```
In [50]: total_data = pd.concat([X, y], axis=1)
         corr = total_data.corr()
         plt.figure(figsize=(10, 10))
         plt.title('Correlation Matrix')
         sns.heatmap(corr, cmap='YlGnBu', annot=True)
```

```
Out[50]: <matplotlib.axes._subplots.AxesSubplot at 0x1f84316ff88>
```



MODEL

```
In [51]: from xgboost import XGBClassifier
from pprint import pprint
from sklearn.preprocessing import MinMaxScaler
from sklearn.model_selection import train_test_split
from sklearn.model_selection import StratifiedKFold
from sklearn.model_selection import RandomizedSearchCV
from sklearn.model_selection import GridSearchCV
from sklearn.ensemble import RandomForestClassifier
from sklearn.ensemble import ExtraTreesClassifier
from sklearn.model_selection import StratifiedKFold
from sklearn.metrics import accuracy_score
```

K FOLD CROSS VALIDATION AND HYPER PARAMETER TUNNING

```
In [52]: kfold = StratifiedKFold(n_splits=5, random_state=42, shuffle=True)

#cross validation.
def cross_validate(model, train, labels):
    scores = []
    best_accuracy = 0
    print(train.shape)

    for train_ind, val_ind in kfold.split(train, labels):

        t_data = train.loc[train_ind]
        t_labels = labels.loc[train_ind]
        v_data = train.loc[val_ind]
        v_labels = labels.loc[val_ind]

        model.fit(t_data, t_labels)
        preds = model.predict(v_data)
        score = accuracy_score(v_labels, preds)
        scores.append(score)
```

```

        if score > best_accuracy:
            best_accuracy = score
            best_model = model

    print('Accuracy : ' + str(round(sum(scores)/len(scores), 2)))
    return best_model, best_accuracy

#parameter tuning.
def tune(model, params, train, labels):
    search = RandomizedSearchCV(model, params, n_iter=20, cv=6, random_
state=21)
    best_model = search.fit(train, labels)
    pprint(best_model.best_estimator_.get_params())
    return best_model

```

RANDOM FOREST

```

In [53]: params = {
    'n_estimators' : [125, 150, 175, 200],
    'max_depth' : [6, 7, 8],
    'max_features' : [4, 5, 6, 7],
    'bootstrap' : [True],
    'min_samples_leaf' : [2, 3, 4]
}

model = RandomForestClassifier()

best_model = tune(model, params, X, y)
rfc = RandomForestClassifier(**best_model.best_estimator_.get_params())
final_model, best_accuracy = cross_validate(rfc, X, y)
print('Accuracy of the best model : ' + str(round(best_accuracy, 2)))

```

```

C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
1d array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)

```

```
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
```

```

estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n samples,), fo

```

```
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
```

```
ld array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model selection\ valida
```

```
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
```



```
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
```



```

estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n samples,), fo

```

```
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a
```

```
ld array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model selection\ valida
```

```
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
```

```
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
```

```

estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n samples,), fo

```



```
ld array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida  
tion.py:531: DataConversionWarning: A column-vector y was passed when a  
1d array was expected. Please change the shape of y to (n_samples,), fo  
r example using ravel().  
    estimator.fit(X_train, y_train, **fit_params)  
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model selection\ valida
```



```
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_valida
tion.py:531: DataConversionWarning: A column-vector y was passed when a
ld array was expected. Please change the shape of y to (n_samples,), fo
r example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
```

```
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_validation.py:531: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    estimator.fit(X_train, y_train, **fit_params)
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\model_selection\_search.py:765: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
    self.best_estimator_.fit(X, y, **fit_params)
```

```
{'bootstrap': True,
 'ccp_alpha': 0.0,
 'class_weight': None,
 'criterion': 'gini',
 'max_depth': 8,
 'max_features': 5,
 'max_leaf_nodes': None,
 'max_samples': None,
 'min_impurity_decrease': 0.0,
 'min_impurity_split': None,
 'min_samples_leaf': 4,
 'min_samples_split': 2,
 'min_weight_fraction_leaf': 0.0,
 'n_estimators': 175,
 'n_jobs': None,
 'oob_score': False,
 'random_state': None,
 'verbose': 0,
 'warm_start': False}
(3980, 11)
```

```
C:\Users\HP\Anaconda3\lib\site-packages\ipykernel_launcher.py:16: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using r
```

```

avel().
    app.launch_new_instance()
C:\Users\HP\Anaconda3\lib\site-packages\ipykernel_launcher.py:16: DataC
onversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples,), for example using r
avel().
    app.launch_new_instance()
C:\Users\HP\Anaconda3\lib\site-packages\ipykernel_launcher.py:16: DataC
onversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples,), for example using r
avel().
    app.launch_new_instance()
C:\Users\HP\Anaconda3\lib\site-packages\ipykernel_launcher.py:16: DataC
onversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples,), for example using r
avel().
    app.launch_new_instance()
C:\Users\HP\Anaconda3\lib\site-packages\ipykernel_launcher.py:16: DataC
onversionWarning: A column-vector y was passed when a 1d array was expe
cted. Please change the shape of y to (n_samples,), for example using r
avel().
    app.launch_new_instance()

```

Accuracy : 0.8

Accuracy of the best model : 0.83

XG-BOOST

```

In [55]: params = {
    'eta' : [0.1, 0.2, 0.3, 0.4],
    'max_depth' : [4, 5, 6, 7],
    'verbosity' : [1],
    'subsample' : [0.5, 0.75, 1],
    'n_estimators' : [75, 100, 125, 150],
    'min_child_weight' : [2, 3, 4, 5],
    'objective' : ['multi:softmax'],
    'num_class' : [6]

```

```

}

model = XGBClassifier()

best_model = tune(model, params, X, y)
xgb = XGBClassifier(**best_model.best_estimator_.get_params())
final_model, best_accuracy = cross_validate(xgb, X, y)
print('Accuracy of the best model : ' + str(round(best_accuracy, 2)))

```

C:\Users\HP\Anaconda3\lib\site-packages\sklearn\utils\validation.py:72:
DataConversionWarning: A column-vector y was passed when a 1d array was
expected. Please change the shape of y to (n_samples,), for example us
ing ravel().
return f(**kwargs)

```

{'base_score': 0.5,
 'booster': 'gbtree',
 'colsample_bylevel': 1,
 'colsample_bynode': 1,
 'colsample_bytree': 1,
 'eta': 0.3,
 'gamma': 0,
 'gpu_id': -1,
 'importance_type': 'gain',
 'interaction_constraints': '',
 'learning_rate': 0.300000012,
 'max_delta_step': 0,
 'max_depth': 7,
 'min_child_weight': 2,
 'missing': nan,
 'monotone_constraints': '()',
 'n_estimators': 150,
 'n_jobs': 0,
 'num_class': 6,
 'num_parallel_tree': 1,
 'objective': 'multi:softprob',
 'random_state': 0,
 'reg_alpha': 0,
 'reg_lambda': 1,
 'scale_pos_weight': None,

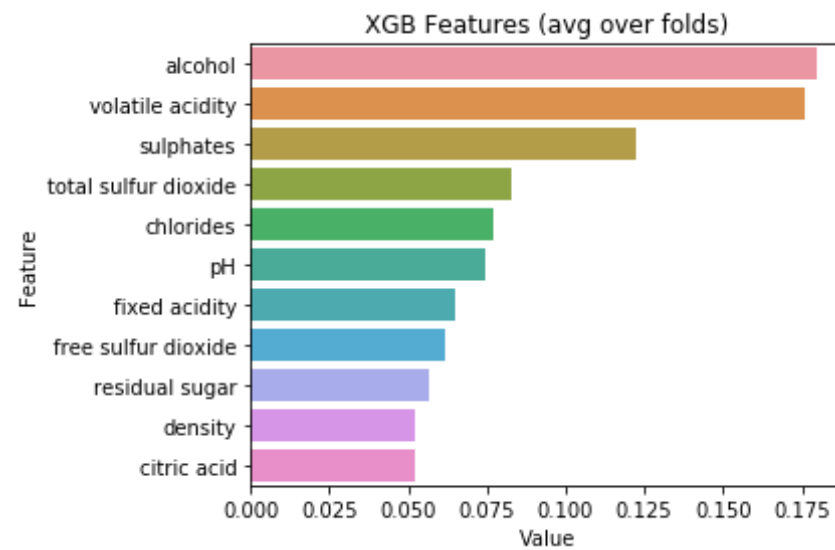
```

```
'subsample': 0.75,  
'tree_method': 'exact',  
'validate_parameters': 1,  
'verbosity': 1}  
(3980, 11)
```

```
C:\Users\HP\Anaconda3\lib\site-packages\sklearn\utils\validation.py:72:  
DataConversionWarning: A column-vector y was passed when a 1d array was  
expected. Please change the shape of y to (n_samples, ), for example us  
ing ravel().  
    return f(**kwargs)
```

```
Accuracy : 0.89  
Accuracy of the best model : 0.9
```

```
In [57]: feature_imp = pd.DataFrame(sorted(zip(final_model.feature_importances_,  
X.columns)), columns=['Value', 'Feature'])  
  
sns.barplot(x="Value", y="Feature", data=feature_imp.sort_values(by="Va  
lue", ascending=False))  
plt.title('XGB Features (avg over folds)')  
plt.tight_layout()  
plt.show()
```



In []: