WINE QUALITY ANALYSIS

(CORIZO PROJECT)

LIBRARIES USED:

OpenCV - pip install opencv-contrib-python
Numpy - pip install numpy
Pandas - pip install pandas
Matplot library- pip install matplotlib
Scipy- pip install Scipio
PIL- Python imaging library
Skimage- pip install scikit-image
Seaborn- pip install seaborne

FILE USED: 1613779-Red_wine_.csv

IN THIS I WILL SHOW ALL THE QUALITY ANALYSIS OF THE WINE WHICH INCLUDES VOLATILITY, CITRIC ACID, ACIDITY, ETC.

CODE WITH OUTPUTS IN JUPYTER NOTEBOOKS:

Wine Quality Analysis (corizo project)

All the work here is done by PRIYANSH CHHABRA AND ANANYA BISHT

[68]
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```
Python

[69]
```
data = pd.read_csv("/Users/priyansh/Downloads/1613779-Red_wine_.csv")
data.head()
```
Python

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
0	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5.0
1	7.8	0.88	0.00	2.6	0.098	25.0	67.0	0.9968	3.20	0.68	9.8	5.0
2	7.8	0.76	0.04	2.3	0.092	15.0	54.0	0.9970	3.26	0.65	9.8	5.0
3	11.2	0.28	0.56	1.9	0.075	17.0	60.0	0.9980	3.16	0.58	9.8	6.0
4	7.4	0.70	0.00	1.9	0.076	11.0	34.0	0.9978	3.51	0.56	9.4	5.0

[70]
```
data.info()
```
Python

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1599 entries, 0 to 1598
Data columns (total 12 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   fixed acidity         1599 non-null   float64
 1   volatile acidity      1599 non-null   float64
 2   citric acid           1599 non-null   float64
 3   residual sugar        1599 non-null   float64
 4   chlorides             1599 non-null   float64
 5   free sulfur dioxide   1599 non-null   float64
 6   total sulfur dioxide  1598 non-null   float64
 7   density               1599 non-null   float64
 8   pH                    1598 non-null   float64
 9   sulphates             1599 non-null   float64
 10  alcohol               1599 non-null   float64
 11  quality               1598 non-null   float64
dtypes: float64(12)
memory usage: 150.0 KB
```

```
[71]
data = data.iloc[:,1:]
print(data.head())
```

```
   volatile acidity  citric acid  residual sugar  chlorides  \
0              0.70         0.00             1.9      0.076
1              0.88         0.00             2.6      0.098
2              0.76         0.04             2.3      0.092
3              0.28         0.56             1.9      0.075
4              0.70         0.00             1.9      0.076

   free sulfur dioxide  total sulfur dioxide  density    pH  sulphates  \
0                 11.0                  34.0   0.9978  3.51       0.56
1                 25.0                  67.0   0.9968  3.20       0.68
2                 15.0                  54.0   0.9970  3.26       0.65
3                 17.0                  60.0   0.9980  3.16       0.58
4                 11.0                  34.0   0.9978  3.51       0.56

   alcohol  quality
0      9.4      5.0
1      9.8      5.0
2      9.8      5.0
3      9.8      6.0
4      9.4      5.0
```

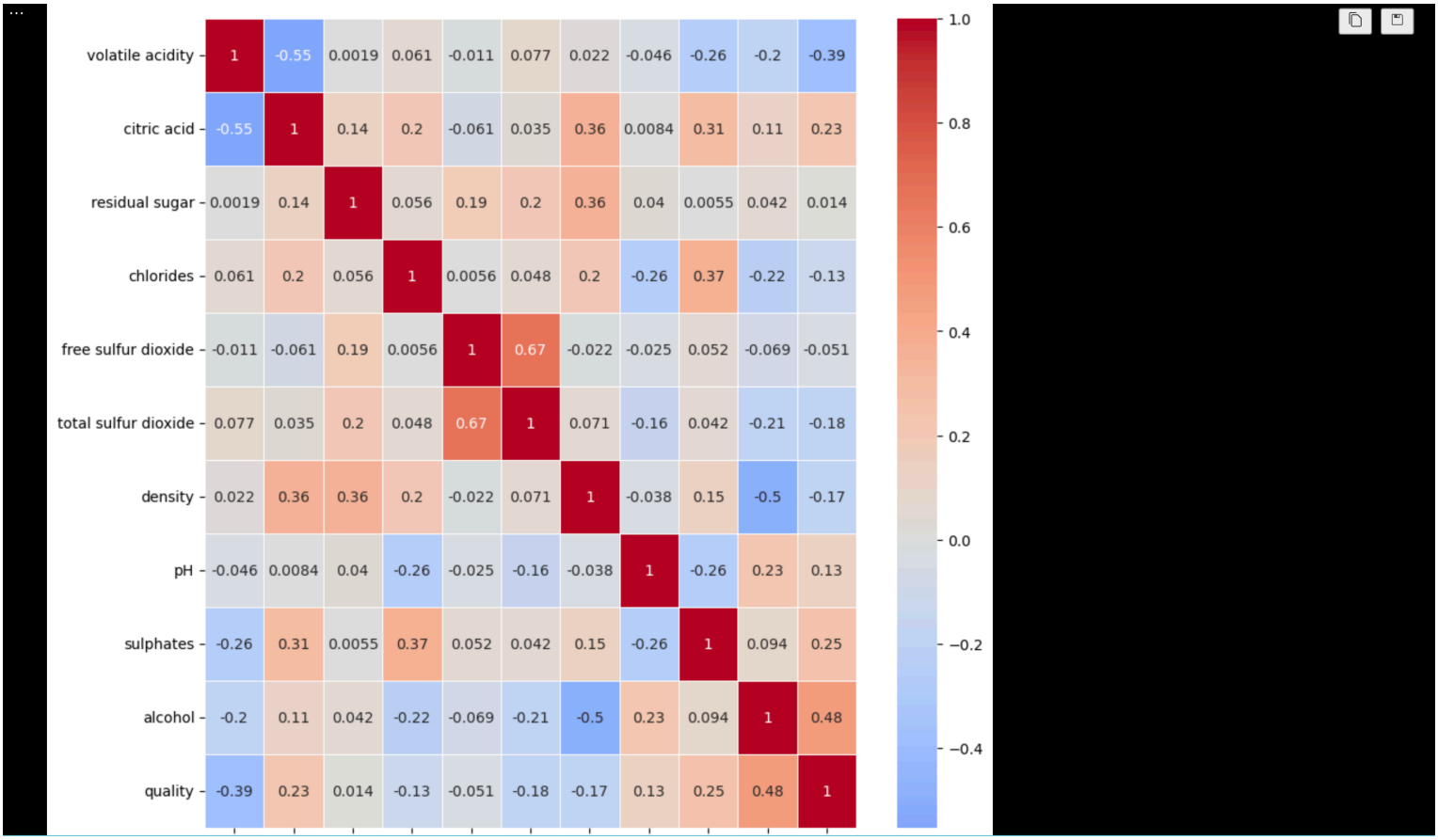
```
[72]
data.describe()
```

	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	alcohol	quality
count	1599.000000	1599.000000	1599.000000	1599.000000	1599.000000	1598.000000	1599.000000	1598.000000	1599.000000	1599.000000	1598.000000
mean	0.527821	0.270976	2.538806	0.087467	15.874922	46.433041	0.996747	3.498586	0.658149	10.422983	5.636421
std	0.179060	0.194801	1.409928	0.047065	10.460157	32.876249	0.001887	0.080346	0.169507	1.065668	0.807665
min	0.120000	0.000000	0.900000	0.012000	1.000000	6.000000	0.990070	2.740000	0.330000	8.400000	3.000000
25%	0.390000	0.090000	1.900000	0.070000	7.000000	22.000000	0.995600	3.520000	0.550000	9.500000	5.000000
50%	0.520000	0.260000	2.200000	0.079000	14.000000	38.000000	0.996750	3.520000	0.620000	10.200000	6.000000
75%	0.640000	0.420000	2.600000	0.090000	21.000000	62.000000	0.997835	3.520000	0.730000	11.100000	6.000000
max	1.580000	1.000000	15.500000	0.611000	72.000000	289.000000	1.003690	3.900000	2.000000	14.900000	8.000000

```
[73]
data.corr()['quality']
```

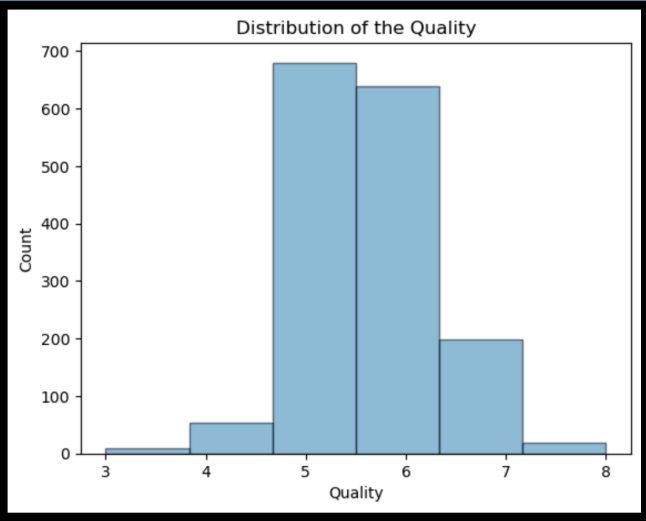
```
volatile acidity       -0.390258
citric acid             0.225867
residual sugar          0.013756
chlorides              -0.129011
free sulfur dioxide    -0.050899
total sulfur dioxide   -0.184699
density                -0.174741
pH                      0.133961
sulphates               0.251118
alcohol                 0.475943
quality                 1.000000
Name: quality, dtype: float64
```

```
[74]
plt.figure(figsize=(10,10))
sns.heatmap(data.corr(),annot=True,linewidth=0.5,center=0,cmap='coolwarm')
plt.show()
```


volatile acidity 1 -0.55 0.0019 0.061 -0.011 0.077 0.022 -0.046 -0.26 -0.2 -0.39
citric acid -0.55 1 0.14 0.2 -0.061 0.035 0.36 0.0084 0.31 0.11 0.23
residual sugar 0.0019 0.14 1 0.056 0.19 0.2 0.36 0.04 0.0055 0.042 0.014
chlorides 0.061 0.2 0.056 1 0.0056 0.048 0.2 -0.26 0.37 -0.22 -0.13
free sulfur dioxide -0.011 -0.061 0.19 0.0056 1 0.67 -0.022 -0.025 0.052 -0.069 -0.051
total sulfur dioxide 0.077 0.035 0.2 0.048 0.67 1 0.071 -0.16 0.042 -0.21 -0.18
density 0.022 0.36 0.36 0.2 -0.022 0.071 1 -0.038 0.15 -0.5 -0.17
pH -0.046 0.0084 0.04 -0.26 -0.025 -0.16 -0.038 1 -0.26 0.23 0.13
sulphates -0.26 0.31 0.0055 0.37 0.052 0.042 0.15 -0.26 1 0.094 0.25
alcohol -0.2 0.11 0.042 -0.22 -0.069 -0.21 -0.5 0.23 0.094 1 0.48
quality -0.39 0.23 0.014 -0.13 -0.051 -0.18 -0.17 0.13 0.25 0.48 1
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4

```
plt.hist(data.quality,bins=6,alpha=0.5,histtype='bar',ec='black')
plt.title('Distribution of the Quality')
plt.xlabel('Quality')
plt.ylabel('Count')
plt.show()
```

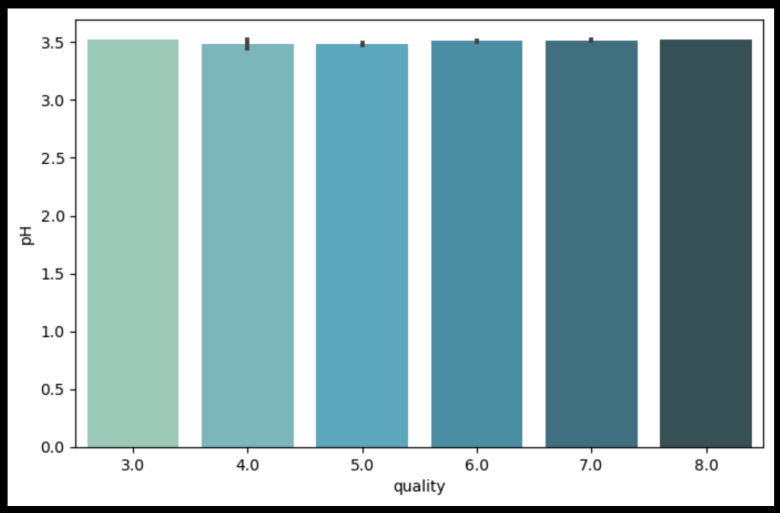
[75] Python




```
plt.figure(figsize=(8,5))
sns.barplot(data['quality'],data['pH'],palette="GnBu_d")
plt.show()
```

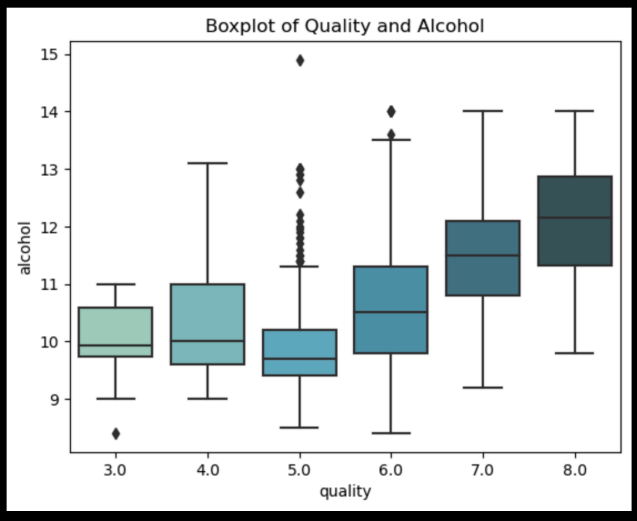
[76] Python

```
/Users/priyansh/opt/anaconda3/lib/python3.9/site-packages/seaborn/_decorators.py:36: FutureWarning: Pass the following variables as keyword args: x, y. From version 0.12, the only
valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```



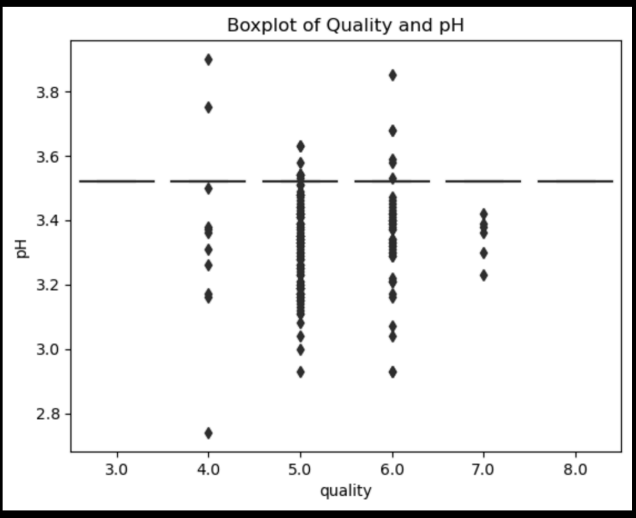

```
ax = sns.boxplot(x='quality',y='alcohol',data=data,palette='GnBu_d')
plt.title("Boxplot of Quality and Alcohol")
plt.show()
```

[78] Python



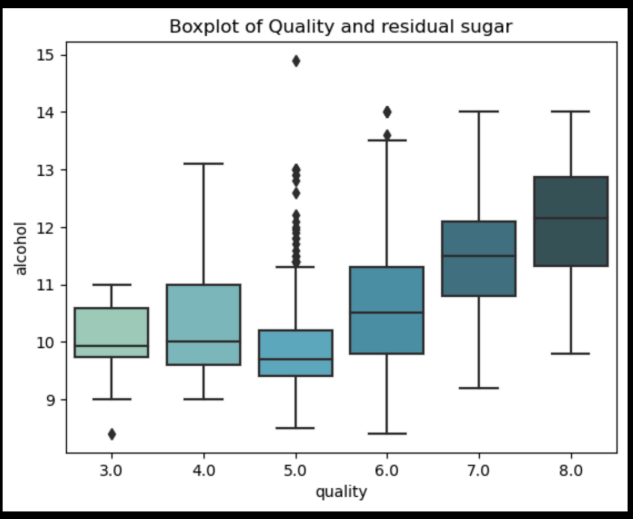
```
sns.boxplot(x='quality',y='pH',data=data,palette='GnBu_d')
plt.title("Boxplot of Quality and pH")
plt.show()
```

[77] Python



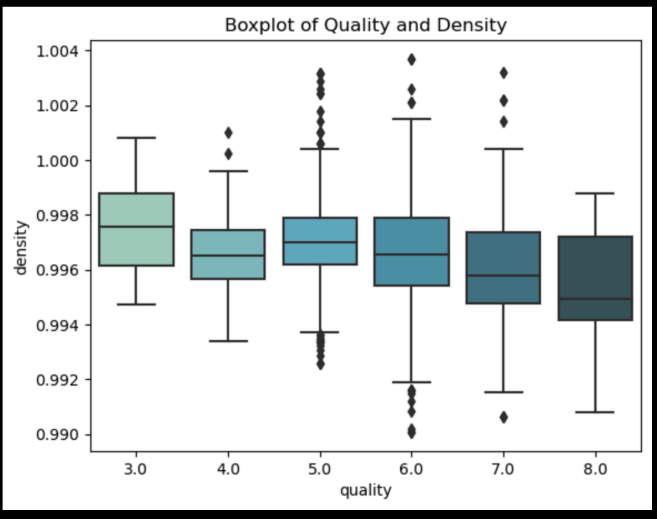

```
py=sns.boxplot(x="quality",y="alcohol",data=data,palette="GnBu_d")
plt.title("Boxplot of Quality and residual sugar")
plt.show()
```

[79] Python



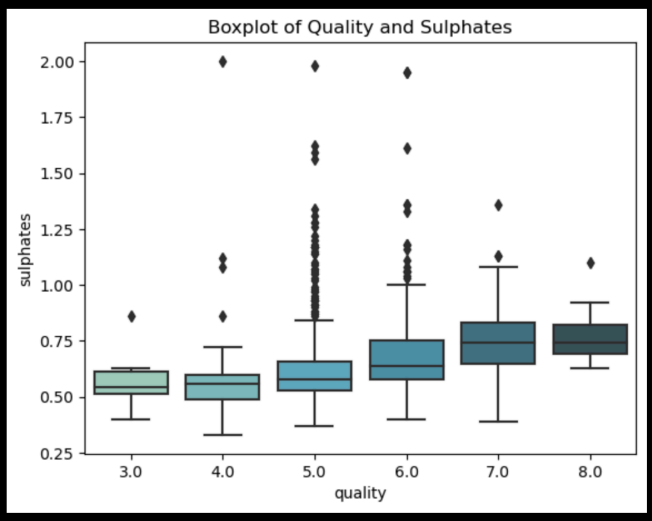

```
sns.boxplot(x="quality",y="density",data=data,palette="GnBu_d")
plt.title("Boxplot of Quality and Density")
plt.show()
```

[80] Python



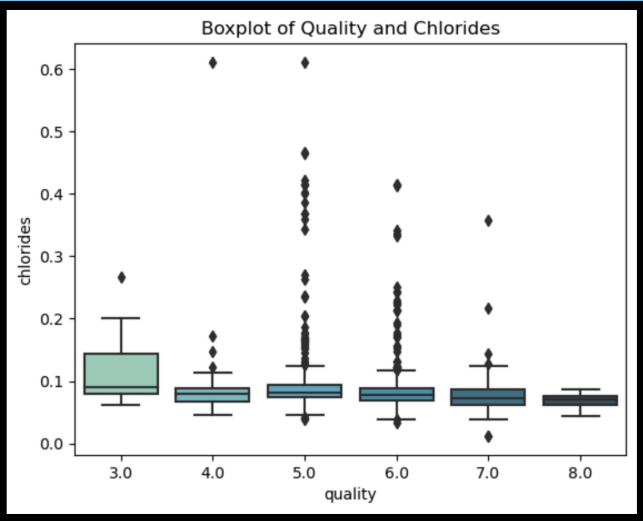
```
sns.boxplot(x="quality",y="sulphates",data=data,palette="GnBu_d")
plt.title("Boxplot of Quality and Sulphates")
plt.show()
```

[81] Python



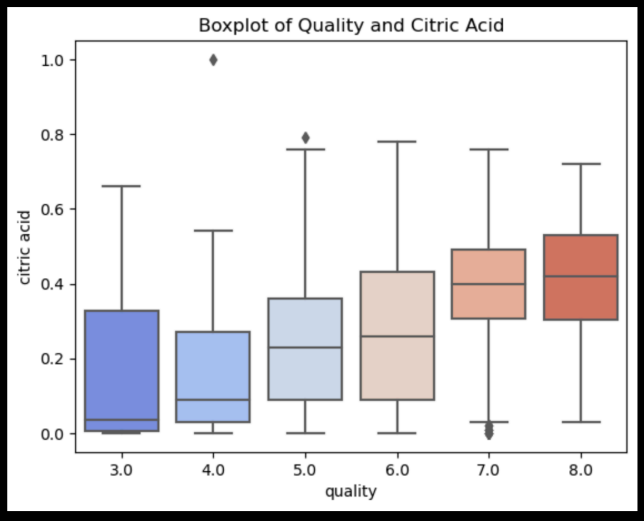

```
sns.boxplot(x="quality",y="chlorides",data=data,palette="GnBu_d")
plt.title("Boxplot of Quality and Chlorides")
plt.show()
```

[82] Python




```
sns.boxplot(x="quality",y="citric acid",data=data,palette="coolwarm")
plt.title("Boxplot of Quality and Citric Acid")
plt.show()
```

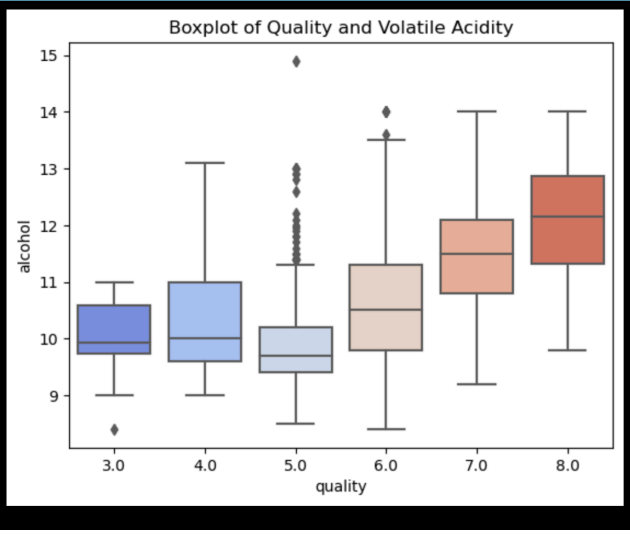
[83] Python



```
sns.boxplot(x="quality",y="alcohol",data=data,palette="coolwarm")
plt.title("Boxplot of Quality and Volatile Acidity")
plt.show()
```

[84]

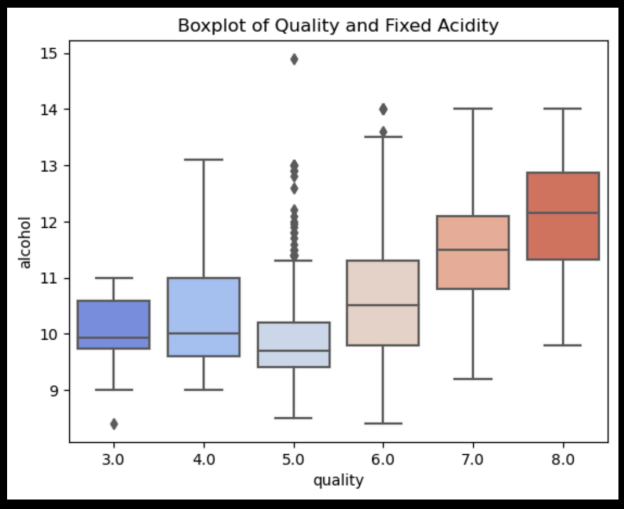
Python




```
sns.boxplot(x="quality",y="alcohol",data=data,palette="coolwarm")
plt.title("Boxplot of Quality and Fixed Acidity")
plt.show()
```

[85]

Python



THANK YOU!

SUBMITTED BY:

PRIYANSH CHHABRA
&
ANANYA BISHT