```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [3]: df=pd.read_csv(r"C:\Users\Admin\Downloads\11_winequality-red - 11_winequality-r
df

Out[3]:

fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	free sulfur dioxide	total sulfur dioxide	density	рН	sulphates	alcol
7.4	0.700	0.00	1.9	0.076	11.0	34.0	0.99780	3.51	0.56	
7.8	0.880	0.00	2.6	0.098	25.0	67.0	0.99680	3.20	0.68	
7.8	0.760	0.04	2.3	0.092	15.0	54.0	0.99700	3.26	0.65	
11.2	0.280	0.56	1.9	0.075	17.0	60.0	0.99800	3.16	0.58	
7.4	0.700	0.00	1.9	0.076	11.0	34.0	0.99780	3.51	0.56	
6.2	0.600	80.0	2.0	0.090	32.0	44.0	0.99490	3.45	0.58	1
5.9	0.550	0.10	2.2	0.062	39.0	51.0	0.99512	3.52	0.76	1
6.3	0.510	0.13	2.3	0.076	29.0	40.0	0.99574	3.42	0.75	1
5.9	0.645	0.12	2.0	0.075	32.0	44.0	0.99547	3.57	0.71	1
6.0	0.310	0.47	3.6	0.067	18.0	42.0	0.99549	3.39	0.66	1
	7.4 7.8 7.8 11.2 7.4 6.2 5.9 6.3 5.9	acidity acidity 7.4 0.700 7.8 0.880 7.8 0.760 11.2 0.280 7.4 0.700 6.2 0.600 5.9 0.550 6.3 0.510 5.9 0.645	acidity acidity acid 7.4 0.700 0.00 7.8 0.880 0.00 7.8 0.760 0.04 11.2 0.280 0.56 7.4 0.700 0.00 6.2 0.600 0.08 5.9 0.550 0.10 6.3 0.510 0.13 5.9 0.645 0.12	acidity acidity acid sugar 7.4 0.700 0.00 1.9 7.8 0.880 0.00 2.6 7.8 0.760 0.04 2.3 11.2 0.280 0.56 1.9 7.4 0.700 0.00 1.9 6.2 0.600 0.08 2.0 5.9 0.550 0.10 2.2 6.3 0.510 0.13 2.3 5.9 0.645 0.12 2.0	acidity acidity acid sugar chlorides 7.4 0.700 0.00 1.9 0.076 7.8 0.880 0.00 2.6 0.098 7.8 0.760 0.04 2.3 0.092 11.2 0.280 0.56 1.9 0.075 7.4 0.700 0.00 1.9 0.076 6.2 0.600 0.08 2.0 0.090 5.9 0.550 0.10 2.2 0.062 6.3 0.510 0.13 2.3 0.076 5.9 0.645 0.12 2.0 0.075	fixed acidity volatile acidity citric acid residual sugar chlorides sulfur dioxide 7.4 0.700 0.00 1.9 0.076 11.0 7.8 0.880 0.00 2.6 0.098 25.0 7.8 0.760 0.04 2.3 0.092 15.0 11.2 0.280 0.56 1.9 0.075 17.0 7.4 0.700 0.00 1.9 0.076 11.0 6.2 0.600 0.08 2.0 0.090 32.0 5.9 0.550 0.10 2.2 0.062 39.0 6.3 0.510 0.13 2.3 0.076 29.0 5.9 0.645 0.12 2.0 0.075 32.0	fixed acidity volatile acidity citric acid residual sugar chlorides sulfur dioxide sulfur dioxide 7.4 0.700 0.00 1.9 0.076 11.0 34.0 7.8 0.880 0.00 2.6 0.098 25.0 67.0 7.8 0.760 0.04 2.3 0.092 15.0 54.0 11.2 0.280 0.56 1.9 0.075 17.0 60.0 7.4 0.700 0.00 1.9 0.076 11.0 34.0 6.2 0.600 0.08 2.0 0.090 32.0 44.0 5.9 0.550 0.10 2.2 0.062 39.0 51.0 6.3 0.510 0.13 2.3 0.076 29.0 40.0 5.9 0.645 0.12 2.0 0.075 32.0 44.0	fixed acidity volatile acidity citric acid residual sugar chlorides sulfur dioxide sulfur dioxide density 7.4 0.700 0.00 1.9 0.076 11.0 34.0 0.99780 7.8 0.880 0.00 2.6 0.098 25.0 67.0 0.99680 7.8 0.760 0.04 2.3 0.092 15.0 54.0 0.99700 11.2 0.280 0.56 1.9 0.075 17.0 60.0 0.99800 7.4 0.700 0.00 1.9 0.076 11.0 34.0 0.99780 60.0 0.99800 7.4 0.700 0.00 1.9 0.076 11.0 34.0 0.99780 6.2 0.600 0.08 2.0 0.090 32.0 44.0 0.99512	fixed acidity volatile acidity citric acidity residual sugar chlorides sulfur dioxide sulfur dioxide density pH 7.4 0.700 0.00 1.9 0.076 11.0 34.0 0.99780 3.51 7.8 0.880 0.00 2.6 0.098 25.0 67.0 0.99680 3.20 7.8 0.760 0.04 2.3 0.092 15.0 54.0 0.99700 3.26 11.2 0.280 0.56 1.9 0.075 17.0 60.0 0.99800 3.16 7.4 0.700 0.00 1.9 0.076 11.0 34.0 0.99780 3.51	Tixed acidity volatile acidity citric acidity residual sugar chlorides sulfur dioxide sulfur dioxide density pH sulphates 7.4 0.700 0.00 1.9 0.076 11.0 34.0 0.99780 3.51 0.56 7.8 0.880 0.00 2.6 0.098 25.0 67.0 0.99680 3.20 0.68 7.8 0.760 0.04 2.3 0.092 15.0 54.0 0.99700 3.26 0.65 11.2 0.280 0.56 1.9 0.075 17.0 60.0 0.99800 3.16 0.58 7.4 0.700 0.00 1.9 0.076 11.0 34.0 0.99780 3.51 0.56 <

1599 rows × 12 columns

In [4]: df.info()

<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	1599 non-null	float64
7	density	1599 non-null	float64
8	рН	1599 non-null	float64
9	sulphates	1599 non-null	float64
10	alcohol	1599 non-null	float64
11	quality	1599 non-null	int64
		>	

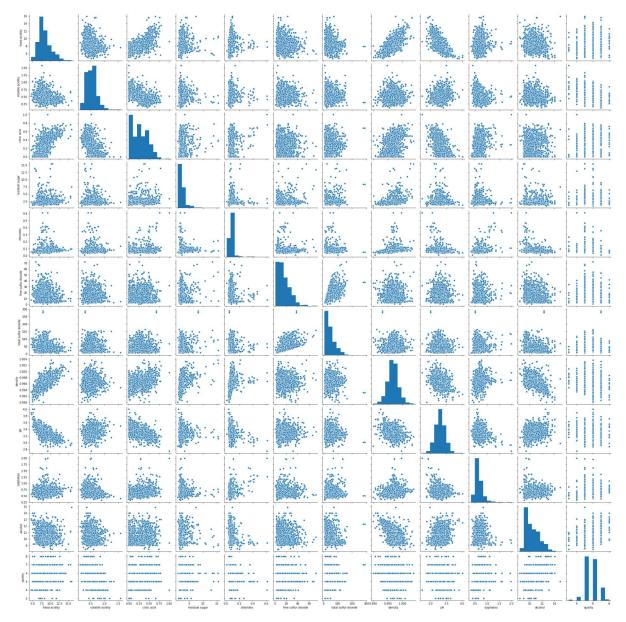
dtypes: float64(11), int64(1)

memory usage: 150.0 KB

In [5]: df.columns

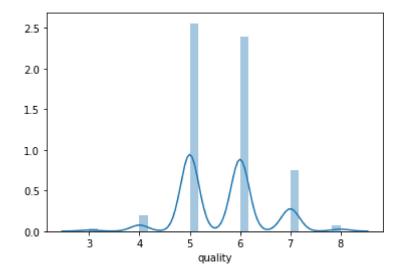
In [6]: sns.pairplot(df)

Out[6]: <seaborn.axisgrid.PairGrid at 0x26b207cf400>



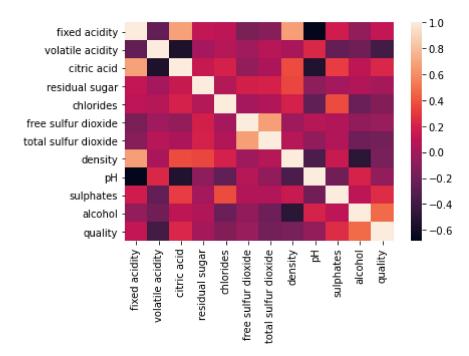
```
In [7]: | sns.distplot(df['quality'])
```

Out[7]: <matplotlib.axes._subplots.AxesSubplot at 0x26b26ca3f70>



In [8]: sns.heatmap(df.corr())

Out[8]: <matplotlib.axes._subplots.AxesSubplot at 0x26b27fe3910>



```
In [10]: from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.3)
```

```
In [11]: from sklearn.linear_model import LinearRegression
lr= LinearRegression()
lr.fit(x_train,y_train)
```

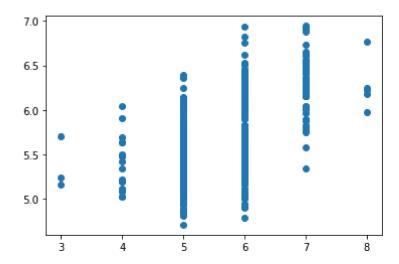
Out[11]: LinearRegression()

In [12]: print(lr.intercept_)

31.513398155038633

In [13]: prediction= lr.predict(x_test)
plt.scatter(y_test,prediction)

Out[13]: <matplotlib.collections.PathCollection at 0x26b28434cd0>



In [14]: print(lr.score(x_test,y_test))

0.3441083808514308

In []: