In [1]:
```
import pandas as pd
import numpy as np
```

```
---------------------------------------------------------------------------
NameError                                 Traceback (most recent call las
t)
Cell In[1], line 3
      1 import pandas as pd
      2 import numpy as np
----> 3 cd

NameError: name 'cd' is not defined
```

In [2]:
```
cd
```

```
C:\Users\cvr
```

In [6]:
```
import pandas as pd
import numpy as np
hd=pd.read_csv('C:\coe 67b6\wine_data.csv')

hd
```

Out[6]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_dioxide
0	11.6	0.580	0.66	2.20	0.074	10.0
1	10.4	0.610	0.49	2.10	0.200	5.0
2	7.4	1.185	0.00	4.25	0.097	5.0
3	10.4	0.440	0.42	1.50	0.145	34.0
4	8.3	1.020	0.02	3.40	0.084	6.0
...	...	...	...	...	...	...
20995	9.7	1.020	0.91	50.00	0.412	114.6
20996	10.2	0.610	0.88	53.80	0.250	62.4
20997	13.4	0.460	1.04	52.10	0.449	63.0
20998	6.6	1.030	1.09	25.30	0.138	179.8
20999	9.3	0.930	1.32	33.60	0.412	128.7

21000 rows × 12 columns

In [7]:
```
import pandas as pd
import numpy as np
hd=pd.read_csv('C:\coe 67b6\wine_data.csv')
hd.head()
hd.tail()
```

Out[7]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_dioxide
20995	9.7	1.02	0.91	50.0	0.412	114.6
20996	10.2	0.61	0.88	53.8	0.250	62.4
20997	13.4	0.46	1.04	52.1	0.449	63.0
20998	6.6	1.03	1.09	25.3	0.138	179.8
20999	9.3	0.93	1.32	33.6	0.412	128.7

In [8]:
```
import pandas as pd
import numpy as np
hd=pd.read_csv('C:\coe 67b6\wine_data.csv')
hd.shape
```

Out[8]:
```
(21000, 12)
```

In [9]:
```
import pandas as pd
import numpy as np
hd=pd.read_csv('C:\coe 67b6\wine_data.csv')
hd.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21000 entries, 0 to 20999
Data columns (total 12 columns):
 #   Column                Non-Null Count  Dtype  
---  ------                --------------  -----  
 0   fixed_acidity         21000 non-null  float64
 1   volatile_acidity      21000 non-null  float64
 2   citric_acid           21000 non-null  float64
 3   residual_sugar        21000 non-null  float64
 4   chlorides             21000 non-null  float64
 5   free_sulfur_dioxide   21000 non-null  float64
 6   total_sulfur_dioxide  21000 non-null  float64
 7   density               21000 non-null  float64
 8   pH                    21000 non-null  float64
 9   sulphates             21000 non-null  float64
 10  alcohol               21000 non-null  float64
 11  quality               21000 non-null  int64  
dtypes: float64(11), int64(1)
memory usage: 1.9 MB
```

In [10]:
```
hd.columns
```

Out[10]:
```
Index(['fixed_acidity', 'volatile_acidity', 'citric_acid', 'residual_suga
r',
       'chlorides', 'free_sulfur_dioxide', 'total_sulfur_dioxide', 'densit
y',
       'pH', 'sulphates', 'alcohol', 'quality'],
      dtype='object')
```

In [11]:
```
hd.isnull()
```

Out[11]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_dioxide
0	False	False	False	False	False	False
1	False	False	False	False	False	False
2	False	False	False	False	False	False
3	False	False	False	False	False	False
4	False	False	False	False	False	False
...	...	...	...	...	...	...
20995	False	False	False	False	False	False
20996	False	False	False	False	False	False
20997	False	False	False	False	False	False
20998	False	False	False	False	False	False
20999	False	False	False	False	False	False

21000 rows × 12 columns

In [12]:
```
hd.isnull().sum()
```

Out[12]:
```
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 [13]:
```
hd.iloc[0]
```

Out[13]:
```
fixed_acidity           11.6000
volatile_acidity         0.5800
citric_acid              0.6600
residual_sugar           2.2000
chlorides                0.0740
free_sulfur_dioxide     10.0000
total_sulfur_dioxide    47.0000
density                  1.0008
pH                       3.2500
sulphates                0.5700
alcohol                  9.0000
quality                  3.0000
Name: 0, dtype: float64
```

In [15]:
```
hd.pH.unique()
```

Out[15]:
```
array([3.25, 3.16, 3.63, 3.38, 3.48, 3.5 , 3.32, 3.31, 3.4 , 3.55, 3.02,
       3.53, 3.23, 3.24, 2.89, 2.87, 3.2 , 3.04, 3.14, 2.93, 3.05, 3.42,
       2.9 , 3.37, 3.44, 3.01, 3.03, 2.98, 3.08, 3.15, 3.1 , 3.58, 2.96,
       2.91, 2.85, 2.99, 3.21, 3.39, 3.11, 3.34, 3.47, 2.78, 2.94, 3.19,
       3.09, 3.18, 3.17, 2.95, 3.45, 3.29, 3.65, 3.22, 3.13, 3.51, 3.26,
       3.56, 2.92, 3.07, 3.33, 2.86, 3.28, 3.12, 3.41, 2.97, 3.  , 3.3 ,
       3.46, 3.27, 2.88, 2.84, 3.06, 3.35, 3.36, 3.62, 2.81, 3.6 , 3.57,
       2.8 , 3.43, 3.52, 3.59, 2.83, 3.68, 2.79, 3.61, 3.54, 3.49, 2.82,
       3.72, 3.9 , 3.75, 2.74, 3.66, 2.75, 3.64, 2.77, 3.71, 3.67, 3.74,
       3.69, 3.77, 3.79, 3.81, 3.85, 3.7 , 3.78, 4.01, 2.72, 3.8 , 3.76,
       3.82, 3.73])
```

In [18]:
```
hd.drop_duplicates()
```

Out[18]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_dioxide
0	11.6	0.580	0.66	2.20	0.074	10.0
1	10.4	0.610	0.49	2.10	0.200	5.0
2	7.4	1.185	0.00	4.25	0.097	5.0
3	10.4	0.440	0.42	1.50	0.145	34.0
4	8.3	1.020	0.02	3.40	0.084	6.0
...	...	...	...	...	...	...
20985	10.3	1.080	0.53	50.80	0.290	84.2
20987	11.4	0.900	0.64	51.20	0.417	197.1
20988	11.8	1.100	0.84	46.50	0.255	198.4
20993	13.0	0.580	1.22	52.80	0.247	93.0
20994	12.8	0.850	1.12	44.90	0.188	215.3

14940 rows × 12 columns

In [19]:
```
hd.describe()
```

Out[19]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_
count	21000.000000	21000.000000	21000.000000	21000.000000	21000.000000	21000.
mean	9.797079	0.774796	0.793870	31.289348	0.200245	129.
std	2.413919	0.365015	0.384833	19.015391	0.124933	77.
min	3.800000	0.080000	0.000000	0.600000	0.009000	1.
25%	7.600000	0.430000	0.410000	9.800000	0.072000	45.
50%	10.000000	0.830000	0.870000	37.600000	0.205000	145.
75%	11.800000	1.080000	1.110000	46.800000	0.298000	194.
max	15.900000	1.580000	1.660000	65.800000	0.611000	289.

In [22]:
```
hd.rename(columns={'free_sulfur_dioxide':'free_so2'})
```

Out[22]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_so2	total_sulf
0	11.6	0.580	0.66	2.20	0.074	10.0	
1	10.4	0.610	0.49	2.10	0.200	5.0	
2	7.4	1.185	0.00	4.25	0.097	5.0	
3	10.4	0.440	0.42	1.50	0.145	34.0	
4	8.3	1.020	0.02	3.40	0.084	6.0	
...	...	...	...	...	...	...	
20995	9.7	1.020	0.91	50.00	0.412	114.6	
20996	10.2	0.610	0.88	53.80	0.250	62.4	
20997	13.4	0.460	1.04	52.10	0.449	63.0	
20998	6.6	1.030	1.09	25.30	0.138	179.8	
20999	9.3	0.930	1.32	33.60	0.412	128.7	

21000 rows × 12 columns

In [23]:
```
avg=hd['pH'].mean()
print(avg)
```

```
3.1587119047619048
```

In [27]:
```
avg=hd['pH'].median()
print(avg)
```

```
3.15
```

In [29]:
```
avg=hd['pH'].std()
print(avg)
```

```
0.17137100174822836
```

In [30]:
```
hd.corr()
```

Out[30]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_s
fixed_acidity	1.000000	0.562166	0.596388	0.653793	0.601889	
volatile_acidity	0.562166	1.000000	0.634551	0.730159	0.622915	
citric_acid	0.596388	0.634551	1.000000	0.765075	0.660209	
residual_sugar	0.653793	0.730159	0.765075	1.000000	0.747546	
chlorides	0.601889	0.622915	0.660209	0.747546	1.000000	
free_sulfur_dioxide	0.653679	0.746446	0.700625	0.781699	0.647345	
total_sulfur_dioxide	0.506634	0.622015	0.636895	0.748704	0.606934	
density	0.625238	0.673260	0.740276	0.787859	0.715181	
pH	-0.254357	-0.213949	-0.185421	-0.254473	-0.113158	
sulphates	0.579008	0.660118	0.645684	0.679813	0.603356	
alcohol	0.167870	0.375957	0.358110	0.349538	0.278020	
quality	0.037545	0.020286	0.052341	0.049734	0.052905	

In [35]:
```
hd.fillna(hd.mean())
```

Out[35]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_dioxide
0	11.6	0.580	0.66	2.20	0.074	10.0
1	10.4	0.610	0.49	2.10	0.200	5.0
2	7.4	1.185	0.00	4.25	0.097	5.0
3	10.4	0.440	0.42	1.50	0.145	34.0
4	8.3	1.020	0.02	3.40	0.084	6.0
...	...	...	...	...	...	...
20995	9.7	1.020	0.91	50.00	0.412	114.6
20996	10.2	0.610	0.88	53.80	0.250	62.4
20997	13.4	0.460	1.04	52.10	0.449	63.0
20998	6.6	1.030	1.09	25.30	0.138	179.8
20999	9.3	0.930	1.32	33.60	0.412	128.7

21000 rows × 12 columns

In [36]:
```
hd.index
```

Out[36]:
```
RangeIndex(start=0, stop=21000, step=1)
```

In [37]:
```
hd.to_numpy()
```

Out[37]:
```
array([[11.6  ,  0.58 ,  0.66 , ...,  0.57 ,  9.   ,  3.   ],
       [10.4  ,  0.61 ,  0.49 , ...,  0.63 ,  8.4  ,  3.   ],
       [ 7.4  ,  1.185,  0.   , ...,  0.54 , 10.7  ,  3.   ],
       ...,
       [13.4  ,  0.46 ,  1.04 , ...,  1.76 ,  9.3  ,  9.   ],
       [ 6.6  ,  1.03 ,  1.09 , ...,  1.54 , 12.9  ,  9.   ],
       [ 9.3  ,  0.93 ,  1.32 , ...,  1.42 , 13.   ,  9.   ]])
```

In [38]:
```
hd.iloc[0:2,:]
```

Out[38]:

	fixed_acidity	volatile_acidity	citric_acid	residual_sugar	chlorides	free_sulfur_dioxide	tota
0	11.6	0.58	0.66	2.2	0.074	10.0	
1	10.4	0.61	0.49	2.1	0.200	5.0	

In [39]:
```
hd.T
```

Out[39]:

	0	1	2	3	4	5	6	
fixed_acidity	11.6000	10.4000	7.4000	10.40000	8.30000	7.60000	6.80000	7.300
volatile_acidity	0.5800	0.6100	1.1850	0.44000	1.02000	1.58000	0.81500	0.980
citric_acid	0.6600	0.4900	0.0000	0.42000	0.02000	0.00000	0.00000	0.050
residual_sugar	2.2000	2.1000	4.2500	1.50000	3.40000	2.10000	1.20000	2.100
chlorides	0.0740	0.2000	0.0970	0.14500	0.08400	0.13700	0.26700	0.061
free_sulfur_dioxide	10.0000	5.0000	5.0000	34.00000	6.00000	5.00000	16.00000	20.000
total_sulfur_dioxide	47.0000	16.0000	14.0000	48.00000	11.00000	9.00000	29.00000	49.000
density	1.0008	0.9994	0.9966	0.99832	0.99892	0.99476	0.99471	0.997
pH	3.2500	3.1600	3.6300	3.38000	3.48000	3.50000	3.32000	3.310
sulphates	0.5700	0.6300	0.5400	0.86000	0.49000	0.40000	0.51000	0.550
alcohol	9.0000	8.4000	10.7000	9.90000	11.00000	10.90000	9.80000	9.700
quality	3.0000	3.0000	3.0000	3.00000	3.00000	3.00000	3.00000	3.000

12 rows × 21000 columns

In [40]:
```
hd.agg(lambda x: np.mean(x) * 5.6)
```

Out[40]:
```
fixed_acidity             54.863640
volatile_acidity           4.338859
citric_acid                4.445669
residual_sugar           175.220347
chlorides                  1.121373
free_sulfur_dioxide      724.877067
total_sulfur_dioxide    1282.449067
density                    5.655845
pH                        17.688787
sulphates                  5.715592
alcohol                   63.233608
quality                   33.600000
dtype: float64
```

In [ ]: