

EDS MINI PROJECT

GROUP MEMBERS:

- 1) 845: ACHYUT GODE
- 2) 851: KANCHAN SANAP
- 3) 854: SAWALE MANAS
- 4) 855: SUMEGH SHANKADARWAR

CODE:

```
import pandas as pd
wine=pd.read_csv("winequalityN.csv")

# 1) print all records of dataset
print(wine)

# 2) print types of wine
print(wine['type'])

# 3) print mean of fixed acidity
print("Mean of fixed acidity is:",wine['fixed acidity'].mean())

# 4)print median of residual sugar
print("Median of residual sugar is:",wine['residual sugar'].median())

# 5)print maximum % of alcohol
print("Max % of alcohol:",wine['alcohol'].max())

# 6)print aggregation of citric acid present in wine(groupby)
```

```
print("The aggregation is :",wine.groupby('citric acid').agg(['min']))
```

7)print correlation

```
print("The correlation is:",wine.corr())
```

8)print type of wine in uppercase

```
wine['type'].str.upper()
```

9)print reviews of wine in lowercase

```
wine['reviews'].str.lower()
```

10)print description of the dataset

```
print("The description of the dataset is:",wine.describe())
```

11)print fixed acidity of 7.0

```
neutral=wine.groupby('fixed acidity').get_group(7.0)
```

```
print("The wine with fixed acidity 7 is:", neutral)
```

```
d=neutral.count()
```

```
print (neutral)
```

12) Display record of first 20 wines

```
print("The record of first 20 wines is:",wine.iloc[1:20])
```

13) Display missing value in chlorides

```
miss=wine['chlorides']
```

```
print("The missing values in chlorides is:",miss.isnull())
```

14) Display maximum volatile acidity in wine

```
print(" maximum volatile acidity in wine:",wine['volatile acidity'].max())
```

15) Display records of wine with quality 9

```
a=wine['quality']  
b=wine.groupby('quality').get_group(9.0)  
print("The records of wine with quality 9:",b)
```

16) Display minimum of sulphates

```
print(" minimum of sulphates in wine:",wine['sulphates'].min())
```

17) drouping of missing wine

```
print(wine.dropna())
```

18) Display varience of pH in wine

```
j=wine['pH']  
k=j.var()  
print("variance of pH in wine:",k)
```

19) Display mean of density

```
print("Mean of density is:",wine['density'].mean())
```

20) Display median of sulphates

```
print("Median of sulphates is:",wine['sulphates'].median())
```

#SCATTER

```
import matplotlib.pyplot as plt  
import numpy as np
```

```
import pandas as pd
```

```
df1=pd.read_csv('KANCHAN PROJECT.csv')\
```

```
plt.title("QUALITY VS TYPE")
```

```
plt.xlabel("QUALITY")
```

```
plt.ylabel("TYPE")
```

```
plt.scatter(df1['quality'],df1['type'])
```

```
plt.show()
```

#BAR GRAPH

```
import pandas as pd
```

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
df=pd.read_csv('KANCHAN PROJECT.csv')
```

```
fa=np.array(df['fixed acidity'])
```

```
va=np.array(df['volatile acidity'])
```

```
ca=np.array(df['citric acid'])
```

```
rs=np.array(df['residual sugar'])
```

```
cl=np.array(df['chlorides'])
```

```
fsd=np.array(df['free sulfur dioxide'])
```

```
tsd=np.array(df['total sulfur dioxide'])
```

```
den=np.array(df['density'])
```

```
p=np.array(df['pH'])
```

```
sul=np.array(df['sulphates'])
```

```
alc=np.array(df['alcohol'])
ql=np.array(df['quality'])
re=np.array(df['reviews'])
plt.xlabel("FIXED ACIDITY")
plt.ylabel("FREE SULFUR DIOXIDE")
plt.bar(fa,fsd)
plt.show()
```

#HISTOGRAM

```
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np

df=pd.read_csv('KANCHAN PROJECT.csv')
ql=np.array(df['quality'])
x=np.random.normal(ql)

plt.ylabel("QUALITY")
plt.hist(x)
plt.show()
```

#LINE GRAPH

```
import matplotlib.pyplot as plt
import numpy as np
p=np.array(df['pH'])
ypoints = np.array(p)
plt.ylabel('pH')
plt.plot(ypoints,color= 'r')
plt.show()
```

#SUBPLOT

```
import matplotlib.pyplot as plt
import numpy as np
```

#PLOT 1

```
plt.xlabel("CHLORIDES")
plt.ylabel("RESIDUAL SUGAR")
```

```
x=np.array([0.04,0.08,0.05])
y=np.array([1.1,1.6,1.5])
```

```
plt.subplot(1,2,1)
plt.plot(x,y)
```

```
#PLOT 2
```

```
plt.xlabel("FIXED ACIDITY")  
plt.ylabel("VOLATILE ACIDITY")  
x=np.array([6.3,7.2,7.4])  
y=np.array([0.3,0.6,0.1])
```

```
plt.subplot(1,2,2)  
plt.plot(x,y)
```

```
plt.show()
```

OUTPUT:

type	fixed acidity	volatile acidity	citric acid	residual sugar	\
0	white	7.0	0.270	0.36	20.7
1	white	6.3	0.300	0.34	1.6
2	white	8.1	0.280	0.40	6.9
3	white	7.2	0.230	0.32	8.5
4	white	7.2	0.230	0.32	8.5
...
6492	red	6.2	0.600	0.08	2.0
6493	red	5.9	0.550	0.10	2.2
6494	red	6.3	0.510	0.13	2.3
6495	red	5.9	0.645	0.12	2.0
6496	red	6.0	0.310	0.47	3.6

	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH
\					
0	0.045	45.0	170.0	1.00100	3.00
1	0.049	14.0	132.0	0.99400	3.30
2	0.050	30.0	97.0	0.99510	3.26
3	0.058	47.0	186.0	0.99560	3.19
4	0.058	47.0	186.0	0.99560	3.19
...
6492	0.090	32.0	44.0	0.99490	3.45
6493	0.062	39.0	51.0	0.99512	3.52
6494	0.076	29.0	40.0	0.99574	3.42
6495	0.075	32.0	44.0	0.99547	3.57
6496	0.067	18.0	42.0	0.99549	3.39

	sulphates	alcohol	quality	reviews
0	0.45	8.8	6	BAD
1	0.49	9.5	6	VERY BAD
2	0.44	10.1	6	GOOD
3	0.40	9.9	6	VERY GOOD
4	0.40	9.9	6	AVERAGE
...
6492	0.58	10.5	5	BAD
6493	NaN	11.2	6	VERY BAD
6494	0.75	11.0	6	GOOD
6495	0.71	10.2	5	VERY GOOD
6496	0.66	11.0	6	AVERAGE

[6497 rows x 14 columns]

0	white
1	white
2	white

3 white
4 white

...
6492 red
6493 red
6494 red
6495 red
6496 red

Name: type, Length: 6497, dtype: object

Mean of fixed acidity is: 7.2165793124710795

Median of residual sugar is: 3.0

Max % of alcohol: 14.9

The aggregation is : type fixed acidity volatile acidity resi
dual sugar chlorides \

	min	min	min	min	min
citric acid					
0.00	red	4.6	0.26	0.8	0.020
0.01	red	5.0	0.38	1.5	0.017
0.02	red	3.8	0.31	0.8	0.031
0.03	red	5.3	0.35	1.2	0.043
0.04	red	5.0	0.35	0.8	0.031
...
0.91	white	6.3	0.30	8.2	0.034
0.99	white	6.6	0.19	1.2	0.122
1.00	red	7.2	0.21	1.1	0.030
1.23	white	7.6	0.25	4.6	0.035
1.66	white	7.4	0.20	2.1	0.022

free sulfur dioxide total sulfur dioxide density pH sulphat
es \

	min	min	min	min	m
in					
citric acid					
0.00	3.0	8.0	0.98918	3.10	0.
34					
0.01	3.0	13.0	0.98940	3.15	0.
30					
0.02	4.0	11.0	0.98960	3.10	0.
32					
0.03	4.0	8.0	0.99080	3.02	0.
33					
0.04	4.0	9.0	0.98940	3.14	0.
32					
...
..					
0.91	50.0	199.0	0.99394	3.39	0.
49					

0.99	45.0	129.0	0.99360	3.09	0.
31					
1.00	28.0	69.0	0.99310	2.74	0.
38					
1.23	51.0	294.0	0.99018	3.03	0.
43					
1.66	34.0	113.0	0.99165	3.26	0.
55					

	alcohol	quality	reviews
	min	min	min
citric acid			
0.00	8.8	3	AVERAGE
0.01	9.5	4	AVERAGE
0.02	9.0	3	AVERAGE
0.03	9.0	4	AVERAGE
0.04	9.2	4	AVERAGE
...
0.91	11.7	6	BAD
0.99	8.7	6	VERY GOOD
1.00	9.2	4	AVERAGE
1.23	13.1	6	GOOD
1.66	12.2	6	VERY BAD

[89 rows x 13 columns]

The correlation is:

		fixed acidity	volatile acidity
citric acid \			
fixed acidity	1.000000	0.220172	0.323736
volatile acidity	0.220172	1.000000	-0.378061
citric acid	0.323736	-0.378061	1.000000
residual sugar	-0.112319	-0.196702	0.142486
chlorides	0.298421	0.377167	0.039315
free sulfur dioxide	-0.283317	-0.353230	0.133437
total sulfur dioxide	-0.329747	-0.414928	0.195218
density	0.459204	0.271193	0.096320
pH	-0.251814	0.260660	-0.328689
sulphates	0.300380	0.225476	0.057613
alcohol	-0.095603	-0.038248	-0.010433
quality	-0.077031	-0.265953	0.085706

	residual sugar	chlorides	free sulfur dioxide \
fixed acidity	-0.112319	0.298421	-0.283317
volatile acidity	-0.196702	0.377167	-0.353230
citric acid	0.142486	0.039315	0.133437
residual sugar	1.000000	-0.128902	0.403439
chlorides	-0.128902	1.000000	-0.195042
free sulfur dioxide	0.403439	-0.195042	1.000000

total sulfur dioxide	0.495820	-0.279580	0.720934
density	0.552498	0.362594	0.025717
pH	-0.267050	0.044806	-0.145191
sulphates	-0.185745	0.395332	-0.188489
alcohol	-0.359706	-0.256861	-0.179838
quality	-0.036825	-0.200886	0.055463

	total sulfur dioxide	density	pH	sulphates
\				
fixed acidity	-0.329747	0.459204	-0.251814	0.300380
volatile acidity	-0.414928	0.271193	0.260660	0.225476
citric acid	0.195218	0.096320	-0.328689	0.057613
residual sugar	0.495820	0.552498	-0.267050	-0.185745
chlorides	-0.279580	0.362594	0.044806	0.395332
free sulfur dioxide	0.720934	0.025717	-0.145191	-0.188489
total sulfur dioxide	1.000000	0.032395	-0.237687	-0.275381
density	0.032395	1.000000	0.011920	0.259454
pH	-0.237687	0.011920	1.000000	0.191248
sulphates	-0.275381	0.259454	0.191248	1.000000
alcohol	-0.265740	-0.686745	0.121002	-0.003261
quality	-0.041385	-0.305858	0.019366	0.038729

	alcohol	quality
fixed acidity	-0.095603	-0.077031
volatile acidity	-0.038248	-0.265953
citric acid	-0.010433	0.085706
residual sugar	-0.359706	-0.036825
chlorides	-0.256861	-0.200886
free sulfur dioxide	-0.179838	0.055463
total sulfur dioxide	-0.265740	-0.041385
density	-0.686745	-0.305858
pH	0.121002	0.019366
sulphates	-0.003261	0.038729
alcohol	1.000000	0.444319
quality	0.444319	1.000000

The description of the dataset is:

	fixed acidity	volatile acidity
citric acid		
residual sugar		
\		
count	6487.000000	6489.000000
mean	7.216579	0.339691
std	1.296750	0.164649
min	3.800000	0.080000
25%	6.400000	0.230000
50%	7.000000	0.290000
75%	7.700000	0.400000
max	15.900000	1.580000

	chlorides	free sulfur dioxide	total sulfur dioxide	density
\				
count	6495.000000	6497.000000	6497.000000	6497.000000
mean	0.056042	30.525319	115.744574	0.994697
std	0.035036	17.749400	56.521855	0.002999
min	0.009000	1.000000	6.000000	0.987110
25%	0.038000	17.000000	77.000000	0.992340
50%	0.047000	29.000000	118.000000	0.994890
75%	0.065000	41.000000	156.000000	0.996990
max	0.611000	289.000000	440.000000	1.038980

	pH	sulphates	alcohol	quality
count	6488.000000	6493.000000	6497.000000	6497.000000
mean	3.218395	0.531215	10.491801	5.818378
std	0.160748	0.148814	1.192712	0.873255
min	2.720000	0.220000	8.000000	3.000000
25%	3.110000	0.430000	9.500000	5.000000
50%	3.210000	0.510000	10.300000	6.000000
75%	3.320000	0.600000	11.300000	6.000000
max	4.010000	2.000000	14.900000	9.000000

The wine with fixed acidity 7 is: type fixed acidity volatile acid

ity citric acid residual sugar \

0	white	7.0	0.27	0.36	20.7
7	white	7.0	0.27	0.36	20.7
25	white	7.0	0.25	0.32	9.0
27	white	7.0	0.28	0.39	8.7
37	white	7.0	0.33	0.32	1.2
...
6393	red	7.0	0.43	0.02	1.9
6433	red	7.0	0.55	0.13	2.2
6444	red	7.0	0.57	0.02	2.0
6453	red	7.0	0.56	0.17	1.7
6466	red	7.0	0.56	0.13	1.6

	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH
\					
0	0.045	45.0	170.0	1.00100	3.00
7	0.045	45.0	170.0	1.00100	3.00
25	0.046	56.0	245.0	0.99550	3.25
27	0.051	32.0	141.0	0.99610	3.38
37	0.053	38.0	138.0	0.99060	3.13
...
6393	0.080	15.0	28.0	0.99492	3.35
6433	0.075	15.0	35.0	0.99590	3.36
6444	0.072	17.0	26.0	0.99575	3.36
6453	0.065	15.0	24.0	0.99514	3.44
6466	0.077	25.0	42.0	0.99629	3.34

	sulphates	alcohol	quality	reviews
0	0.45	8.80	6	BAD
7	0.45	8.80	6	VERY BAD
25	0.50	10.40	6	VERY BAD
27	0.53	10.50	6	VERY GOOD
37	0.28	11.20	6	VERY BAD
...
6393	0.81	10.60	6	VERY GOOD
6433	0.59	9.70	6	VERY BAD
6444	0.61	10.20	5	BAD
6453	0.68	10.55	7	VERY GOOD
6466	0.59	9.20	5	AVERAGE

[282 rows x 14 columns]

	type	fixed acidity	volatile acidity	citric acid	residual sugar
\					
0	white	7.0	0.27	0.36	20.7
7	white	7.0	0.27	0.36	20.7
25	white	7.0	0.25	0.32	9.0
27	white	7.0	0.28	0.39	8.7
37	white	7.0	0.33	0.32	1.2
...
6393	red	7.0	0.43	0.02	1.9
6433	red	7.0	0.55	0.13	2.2
6444	red	7.0	0.57	0.02	2.0
6453	red	7.0	0.56	0.17	1.7
6466	red	7.0	0.56	0.13	1.6

	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH
\					
0	0.045	45.0	170.0	1.00100	3.00
7	0.045	45.0	170.0	1.00100	3.00
25	0.046	56.0	245.0	0.99550	3.25
27	0.051	32.0	141.0	0.99610	3.38
37	0.053	38.0	138.0	0.99060	3.13
...
6393	0.080	15.0	28.0	0.99492	3.35
6433	0.075	15.0	35.0	0.99590	3.36
6444	0.072	17.0	26.0	0.99575	3.36
6453	0.065	15.0	24.0	0.99514	3.44
6466	0.077	25.0	42.0	0.99629	3.34

	sulphates	alcohol	quality	reviews
0	0.45	8.80	6	BAD
7	0.45	8.80	6	VERY BAD
25	0.50	10.40	6	VERY BAD

27	0.53	10.50	6	VERY GOOD
37	0.28	11.20	6	VERY BAD
...
6393	0.81	10.60	6	VERY GOOD
6433	0.59	9.70	6	VERY BAD
6444	0.61	10.20	5	BAD
6453	0.68	10.55	7	VERY GOOD
6466	0.59	9.20	5	AVERAGE

[282 rows x 14 columns]

The record of first 20 wines is:

	type	fixed acidity	volatile acidity
--	------	---------------	------------------

citric acid	residual sugar \			
1	white	6.3	0.30	0.34
2	white	8.1	0.28	0.40
3	white	7.2	0.23	0.32
4	white	7.2	0.23	0.32
5	white	8.1	0.28	0.40
6	white	6.2	0.32	0.16
7	white	7.0	0.27	0.36
8	white	6.3	0.30	0.34
9	white	8.1	0.22	0.43
10	white	8.1	0.27	0.41
11	white	8.6	0.23	0.40
12	white	7.9	0.18	0.37
13	white	6.6	0.16	0.40
14	white	8.3	0.42	0.62
15	white	6.6	0.17	0.38
16	white	6.3	0.48	0.04
17	white	NaN	0.66	0.48
18	white	7.4	0.34	0.42
19	white	6.5	0.31	0.14

	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH \
1	0.049	14.0	132.0	0.9940	3.30
2	0.050	30.0	97.0	0.9951	3.26
3	0.058	47.0	186.0	0.9956	3.19
4	0.058	47.0	186.0	0.9956	3.19
5	0.050	30.0	97.0	0.9951	3.26
6	0.045	30.0	136.0	0.9949	3.18
7	0.045	45.0	170.0	1.0010	3.00
8	0.049	14.0	132.0	0.9940	3.30
9	0.044	28.0	129.0	0.9938	3.22
10	0.033	11.0	63.0	0.9908	2.99
11	0.035	17.0	109.0	0.9947	3.14
12	0.040	16.0	75.0	0.9920	3.18
13	0.044	48.0	143.0	0.9912	3.54
14	0.040	41.0	172.0	1.0002	2.98

15	0.032	28.0	112.0	0.9914	3.25
16	0.046	30.0	99.0	0.9928	3.24
17	0.029	29.0	75.0	0.9892	3.33
18	0.033	17.0	171.0	0.9917	3.12
19	0.044	34.0	133.0	0.9955	3.22

	sulphates	alcohol	quality	reviews
1	0.49	9.5	6	VERY BAD
2	0.44	10.1	6	GOOD
3	0.40	9.9	6	VERY GOOD
4	0.40	9.9	6	AVERAGE
5	0.44	10.1	6	EXCELLENT
6	0.47	9.6	6	BAD
7	0.45	8.8	6	VERY BAD
8	0.49	9.5	6	GOOD
9	0.45	11.0	6	VERY GOOD
10	0.56	12.0	5	AVERAGE
11	0.53	9.7	5	EXCELLENT
12	0.63	10.8	5	BAD
13	0.52	12.4	7	VERY BAD
14	0.67	9.7	5	GOOD
15	0.55	11.4	7	VERY GOOD
16	0.36	9.6	6	AVERAGE
17	0.39	12.8	8	EXCELLENT
18	0.53	11.3	6	BAD
19	0.50	9.5	5	VERY BAD

The missing values in chlorides is: 0 False

```

1      False
2      False
3      False
4      False
...
6492   False
6493   False
6494   False
6495   False
6496   False

```

Name: chlorides, Length: 6497, dtype: bool

maximum volatile acidity in wine: 1.58

The records of wine with quality 9:

	type	fixed acidity	volatile ac
774	white	9.1	0.27
820	white	6.6	0.36
827	white	7.4	0.24
876	white	6.9	0.36
1605	white	7.1	0.26

	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH
\					
774	0.035	28.0	124.0	0.99700	3.20
820	0.021	24.0	85.0	0.98965	3.41
827	0.031	27.0	139.0	0.99055	3.28
876	0.018	57.0	119.0	0.98980	3.28
1605	0.032	31.0	113.0	0.99030	3.37

	sulphates	alcohol	quality	reviews
774	0.46	10.4	9	BAD
820	0.61	12.4	9	AVERAGE
827	0.48	12.5	9	EXCELLENT
876	0.36	12.7	9	BAD
1605	0.42	12.9	9	VERY GOOD

minimum of sulphates in wine: 0.22

	type	fixed acidity	volatile acidity	citric acid	residual sugar
\					
0	white	7.0	0.270	0.36	20.7
1	white	6.3	0.300	0.34	1.6
2	white	8.1	0.280	0.40	6.9
3	white	7.2	0.230	0.32	8.5
4	white	7.2	0.230	0.32	8.5
...
6491	red	6.8	0.620	0.08	1.9
6492	red	6.2	0.600	0.08	2.0
6494	red	6.3	0.510	0.13	2.3
6495	red	5.9	0.645	0.12	2.0
6496	red	6.0	0.310	0.47	3.6

	chlorides	free sulfur dioxide	total sulfur dioxide	density	pH
\					
0	0.045	45.0	170.0	1.00100	3.00
1	0.049	14.0	132.0	0.99400	3.30
2	0.050	30.0	97.0	0.99510	3.26
3	0.058	47.0	186.0	0.99560	3.19
4	0.058	47.0	186.0	0.99560	3.19
...
6491	0.068	28.0	38.0	0.99651	3.42
6492	0.090	32.0	44.0	0.99490	3.45
6494	0.076	29.0	40.0	0.99574	3.42
6495	0.075	32.0	44.0	0.99547	3.57
6496	0.067	18.0	42.0	0.99549	3.39

	sulphates	alcohol	quality	reviews
0	0.45	8.8	6	BAD
1	0.49	9.5	6	VERY BAD
2	0.44	10.1	6	GOOD

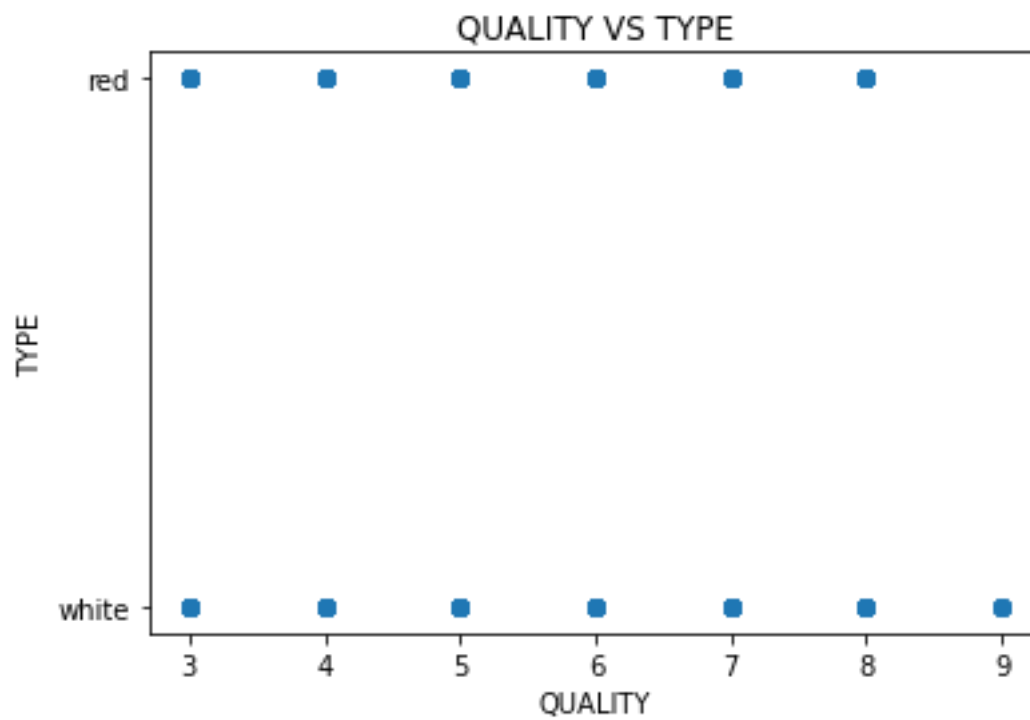
3	0.40	9.9	6	VERY GOOD
4	0.40	9.9	6	AVERAGE
...
6491	0.82	9.5	6	EXCELLENT
6492	0.58	10.5	5	BAD
6494	0.75	11.0	6	GOOD
6495	0.71	10.2	5	VERY GOOD
6496	0.66	11.0	6	AVERAGE

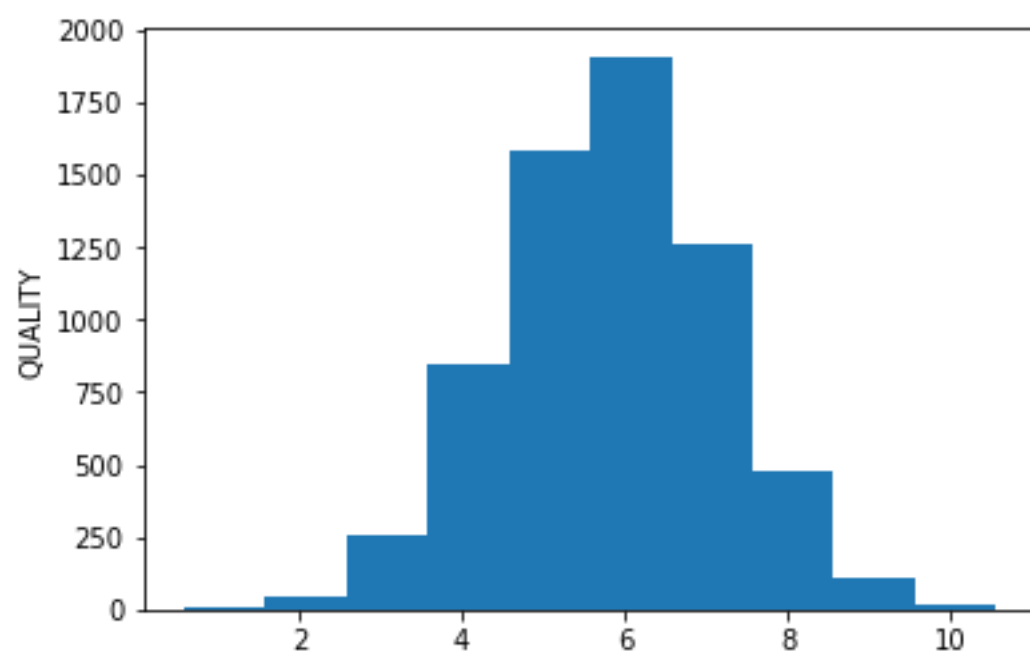
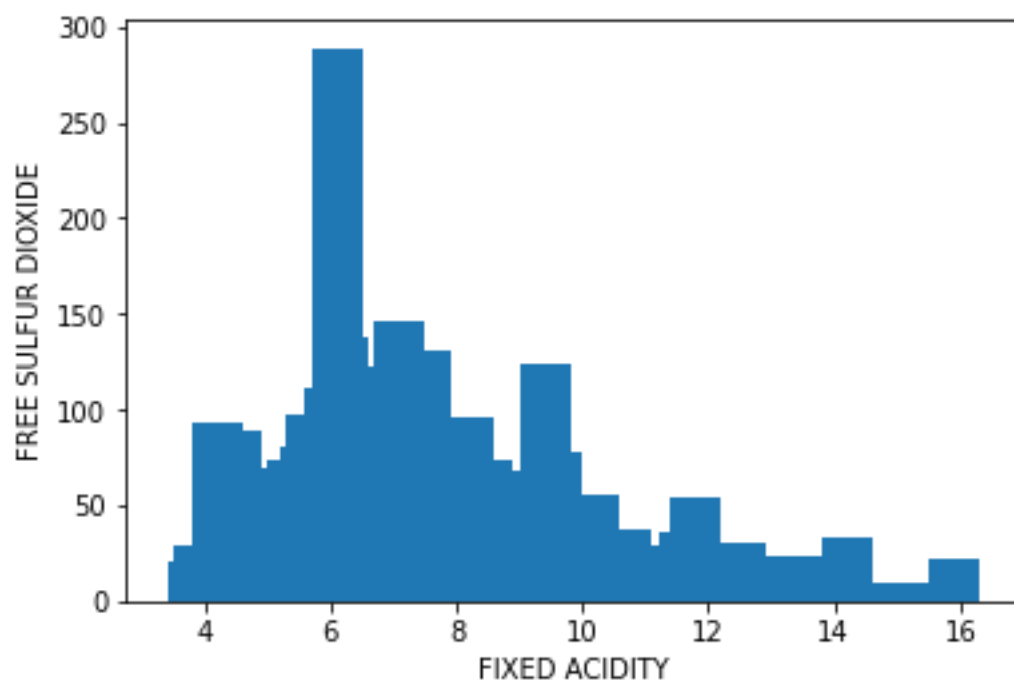
[6463 rows x 14 columns]

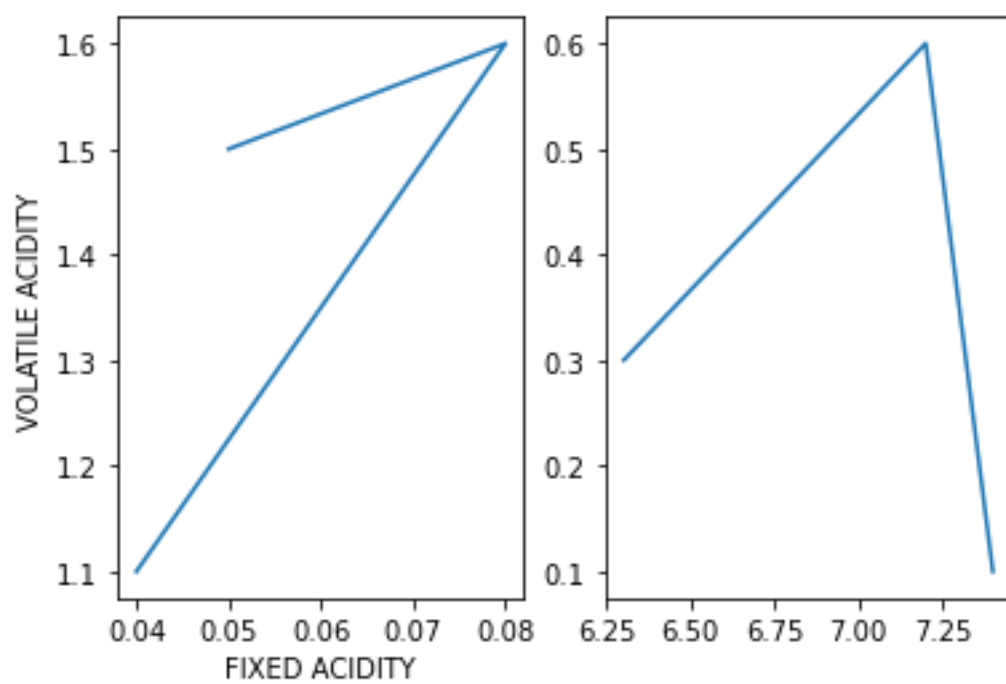
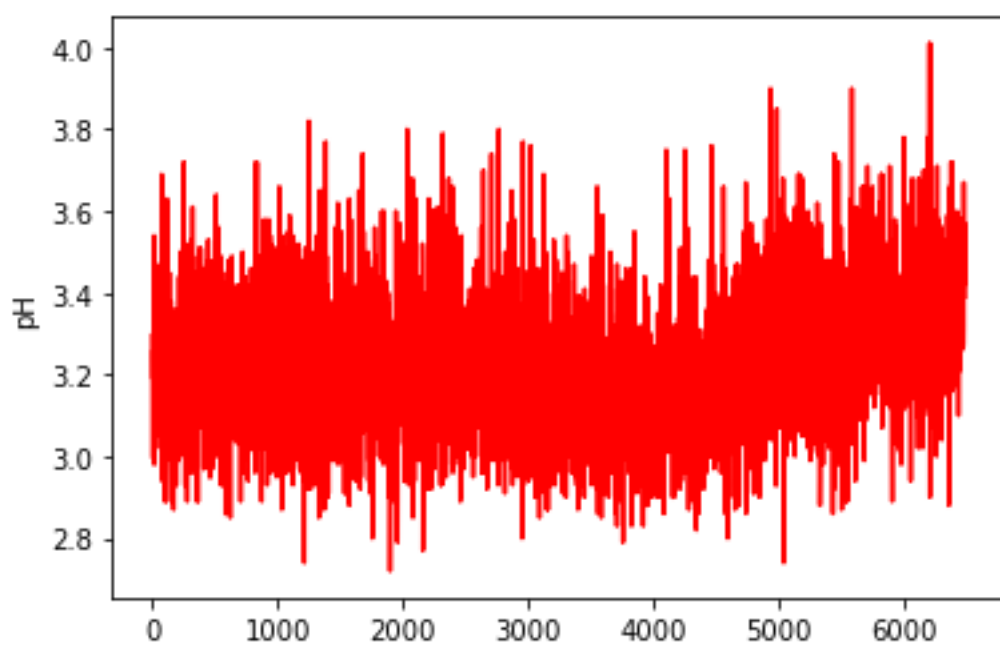
variance of pH in wine: 0.025840018058976632

Mean of density is: 0.9946966338309937

Median of sulphates is: 0.51







**THANK
YOU**