

```

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sea
import sklearn
from sklearn.ensemble import ExtraTreesClassifier
from sklearn.model_selection import GridSearchCV
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import RandomForestClassifier
from sklearn.naive_bayes import GaussianNB
from sklearn.linear_model import LogisticRegression
from sklearn.svm import SVC
from sklearn.model_selection import cross_val_score
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score

import pandas as pd
data = pd.read_csv("C:/Users/TAWFEEQ/Desktop/wine dataset.csv")

data.head()

data.describe()

```

	fixed acidity	volatile acidity	citric acid	residual sugar \
count	6487.000000	6489.000000	6494.000000	6495.000000
mean	7.216579	0.339691	0.318722	5.444326
std	1.296750	0.164649	0.145265	4.758125
min	3.800000	0.080000	0.000000	0.600000
25%	6.400000	0.230000	0.250000	1.800000
50%	7.000000	0.290000	0.310000	3.000000
75%	7.700000	0.400000	0.390000	8.100000
max	15.900000	1.580000	1.660000	65.800000

	chlorides	free sulfur dioxide	total sulfur dioxide
density \			
count	6495.000000	6497.000000	6497.000000
mean	0.056042	30.525319	115.744574
std	0.035036	17.749400	56.521855
min	0.009000	1.000000	6.000000
25%	0.038000	17.000000	77.000000
50%	0.047000	29.000000	118.000000
75%	0.065000	41.000000	156.000000

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

data.info()

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6497 entries, 0 to 6496
Data columns (total 13 columns):
type                6497 non-null object
fixed acidity       6487 non-null float64
volatile acidity    6489 non-null float64
citric acid         6494 non-null float64
residual sugar      6495 non-null float64
chlorides           6495 non-null float64
free sulfur dioxide 6497 non-null float64
total sulfur dioxide 6497 non-null float64
density             6497 non-null float64
pH                 6488 non-null float64
sulphates           6493 non-null float64
alcohol             6497 non-null float64
quality             6497 non-null int64
dtypes: float64(11), int64(1), object(1)
memory usage: 659.9+ KB
```

data.isnull().sum()

type	0
fixed acidity	10
volatile acidity	8
citric acid	3
residual sugar	2
chlorides	2
free sulfur dioxide	0
total sulfur dioxide	0
density	0
pH	9
sulphates	4
alcohol	0
quality	0
dtype: int64	

```
data['fixed acidity'].value_counts()
```

6.80	354
6.60	326
6.40	305
7.00	282
6.90	279
7.20	273
6.70	264
7.10	257
6.50	242
7.40	238
7.30	222
6.20	210
6.30	201
7.60	199
6.00	197
7.50	174
6.10	169
7.80	146
7.70	141
5.80	125
8.00	122
7.90	116
5.90	111
8.20	101
8.30	92
5.70	90
5.60	85
8.10	82
8.40	61
8.60	52
...	
12.60	4
12.20	4
12.40	4
13.00	3
11.70	3
11.00	3
13.30	3
4.40	3
13.20	3
6.15	2
15.50	2
12.90	2
4.20	2
15.60	2
7.15	2
13.70	2
15.00	2

4.60	2
13.40	1
13.50	1
3.80	1
4.50	1
12.10	1
14.00	1
3.90	1
14.30	1
15.90	1
13.80	1
14.20	1
6.45	1

Name: fixed acidity, Length: 106, dtype: int64

```
mean_fa = data['fixed acidity'].mean()
data['fixed acidity'].fillna(mean_fa, inplace = True)
data['fixed acidity'].isnull().sum()
```

0

```
data['volatile acidity'].value_counts()
```

0.280	286
0.240	265
0.260	255
0.250	238
0.220	235
0.270	232
0.230	221
0.200	217
0.300	214
0.320	205
0.210	197
0.180	187
0.310	178
0.290	176
0.190	171
0.340	164
0.330	154
0.160	143
0.360	142
0.170	140
0.350	107
0.380	98
0.400	96
0.390	95
0.370	89
0.150	88
0.410	87

0.430	78
0.440	69
0.420	67

...

0.890	1
0.135	1
0.090	1
1.115	1
1.010	1
0.215	1
0.355	1
0.955	1
1.240	1
1.070	1
0.930	1
1.580	1
1.025	1
0.920	1
0.085	1
1.035	1
0.975	1
0.845	1
1.100	1
0.905	1
0.950	1
0.825	1
0.175	1
0.895	1
0.405	1
0.805	1
1.130	1
1.185	1
0.565	1
0.865	1

Name: volatile acidity, Length: 187, dtype: int64

```
mean_va = data['volatile acidity'].mean()
data['volatile acidity'].fillna(mean_va, inplace = True)
data['volatile acidity'].isnull().sum()
```

0

```
data['citric acid'].value_counts()
```

0.30	337
0.28	301
0.32	289
0.49	283
0.26	257
0.34	249

0.29	244
0.27	236
0.24	232
0.31	229
0.33	208
0.36	197
0.25	163
0.37	153
0.35	150
0.00	150
0.40	146
0.38	136
0.22	131
0.39	129
0.42	124
0.23	108
0.21	99
0.41	98
0.20	95
0.44	86
0.18	71
0.46	70
0.19	69
0.45	68
	...
0.66	21
0.68	18
0.11	16
0.60	15
0.62	15
0.65	15
0.64	15
0.59	14
0.63	14
0.61	11
0.71	10
0.69	9
0.67	9
0.73	8
0.72	6
1.00	6
0.70	5
0.78	3
0.76	3
0.79	3
0.80	2
0.91	2
0.81	2
0.82	2

```

0.88      1
1.66      1
0.75      1
0.86      1
0.99      1
1.23      1
Name: citric acid, Length: 89, dtype: int64

mean_ca = data['citric acid'].mean()
data['citric acid'].fillna(mean_ca, inplace = True)
data['citric acid'].isnull().sum()

0

data['residual sugar'].value_counts()

2.00      235
1.80      228
1.60      223
1.40      219
1.20      195
2.20      187
2.10      179
1.90      176
1.70      175
1.50      171
1.30      152
2.30      150
1.10      146
2.40      127
2.50      124
2.60      112
1.00       93
2.80       85
2.70       77
2.90       49
4.60       46
5.00       44
7.80       43
3.20       43
3.00       42
4.80       41
6.30       41
0.90       41
7.40       40
6.40       37
...
15.25      1
19.90      1
8.55       1

```

5.55	1
65.80	1
18.50	1
15.15	1
20.40	1
6.35	1
11.45	1
18.40	1
7.95	1
17.35	1
20.15	1
3.65	1
8.45	1
19.10	1
16.45	1
9.05	1
16.55	1
19.60	1
7.45	1
22.60	1
6.95	1
8.65	1
14.15	1
14.05	1
3.15	1
8.95	1
20.30	1

Name: residual sugar, Length: 316, dtype: int64

```
mean_rs = data['residual sugar'].mean()
data['residual sugar'].fillna(mean_rs, inplace = True)
data['residual sugar'].isnull().sum()
```

0

```
data['chlorides'].value_counts()
```

0.044	206
0.036	200
0.042	187
0.046	185
0.050	182
0.048	182
0.040	182
0.047	175
0.045	174
0.034	169
0.038	168
0.039	161
0.037	160


```
0.041    151
0.043    142
0.049    141
0.053    135
0.035    130
0.033    119
0.051    116
0.052    114
0.054    112
0.032    109
0.030    108
0.031    107
0.056     97
0.028     85
0.029     81
0.057     78
0.080     70
...
0.143     1
0.290     1
0.149     1
0.267     1
0.209     1
0.464     1
0.403     1
0.236     1
0.211     1
0.204     1
0.156     1
0.190     1
0.130     1
0.240     1
0.222     1
0.212     1
0.301     1
0.113     1
0.263     1
0.611     1
0.413     1
0.009     1
0.239     1
0.387     1
0.343     1
0.213     1
0.165     1
0.150     1
0.144     1
0.125     1
Name: chlorides, Length: 214, dtype: int64
```

```
mean_c = data['chlorides'].mean()
data['chlorides'].fillna(mean_c, inplace = True)
data['chlorides'].isnull().sum()
```

```
0
```

```
data['pH'].value_counts()
```

3.16	200
3.14	193
3.22	185
3.20	176
3.19	170
3.15	170
3.18	168
3.24	160
3.12	154
3.10	154
3.17	151
3.30	150
3.26	149
3.08	147
3.23	147
3.25	140
3.36	139
3.11	135
3.21	131
3.32	131
3.13	130
3.28	129
3.29	126
3.06	124
3.27	123
3.34	122
3.31	118
3.04	107
3.38	106
3.09	103
...	
3.63	9
3.68	7
3.69	6
3.72	6
3.62	5
3.65	4
3.67	4
2.83	4
3.71	4
3.74	3
2.79	3

3.75	3
2.80	3
3.77	2
3.80	2
2.74	2
3.70	2
3.78	2
3.76	2
3.90	2
3.64	2
4.01	2
2.72	1
3.85	1
3.79	1
3.81	1
2.77	1
2.82	1
2.84	1
3.82	1

Name: pH, Length: 108, dtype: int64

```
mean_pH = data['pH'].mean()
data['pH'].fillna(mean_pH, inplace = True)
data['pH'].isnull().sum()
```

0

```
data['sulphates'].value_counts()
```

0.50	275
0.46	243
0.54	234
0.44	232
0.38	214
0.48	208
0.52	203
0.49	197
0.47	191
0.45	190
0.42	186
0.53	186
0.40	172
0.43	169
0.56	168
0.58	167
0.51	166
0.39	157
0.60	157
0.55	152
0.59	148

```
0.41    139
0.57    138
0.37    131
0.62    129
0.36    120
0.64    113
0.63     96
0.61     90
0.35     85
```

```
...
1.05     3
1.28     2
1.10     2
1.13     2
1.01     2
1.03     2
1.95     2
1.00     2
1.04     2
1.07     2
1.14     2
1.34     1
2.00     1
1.16     1
1.61     1
1.15     1
1.22     1
1.31     1
1.09     1
0.23     1
0.22     1
1.11     1
1.20     1
1.98     1
1.26     1
1.59     1
1.12     1
1.33     1
1.56     1
1.62     1
```

```
Name: sulphates, Length: 111, dtype: int64
```

```
mean_s = data['sulphates'].mean()
data['sulphates'].fillna(mean_s, inplace = True)
data['sulphates'].isnull().sum()
```

```
0
```

```
data.isnull().sum()
```

```

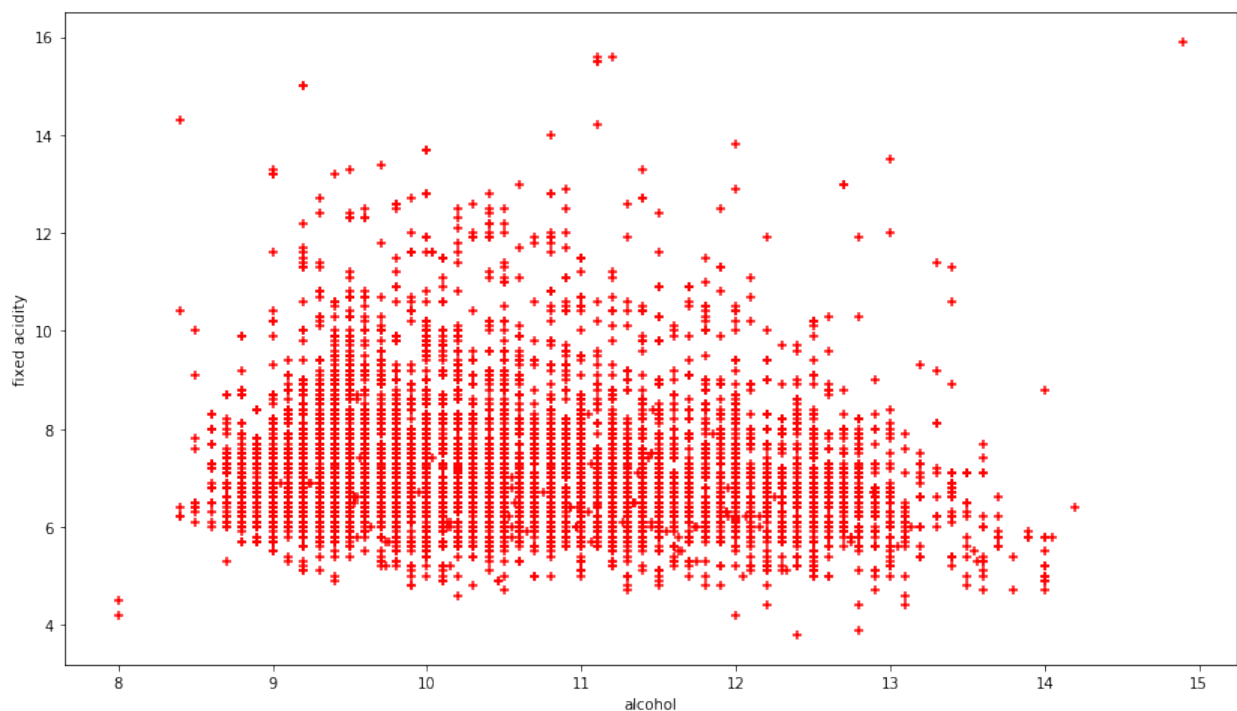
type          0
fixed acidity  0
volatile acidity  0
citric acid    0
residual sugar  0
chlorides      0
free sulfur dioxide  0
total sulfur dioxide  0
density       0
pH            0
sulphates     0
alcohol       0
quality       0
dtype: int64

'''
Visualization of the data
x - alcohol
y - fixed acidity
'''

plt.figure(figsize=(14, 8))
plt.scatter(x = 'alcohol', y = 'fixed acidity', data = data, marker =
'+', c='r')
plt.xlabel('alcohol')
plt.ylabel('fixed acidity')
plt.show()

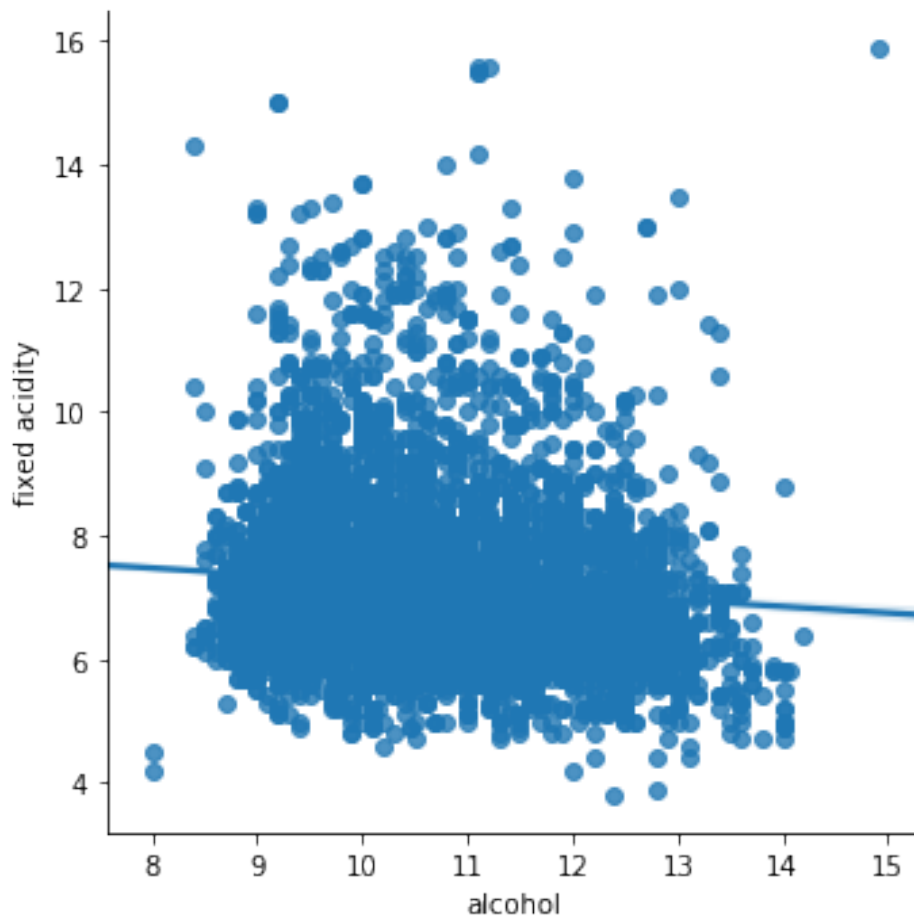
<matplotlib.text.Text at 0x2b39aef7ef0>

```



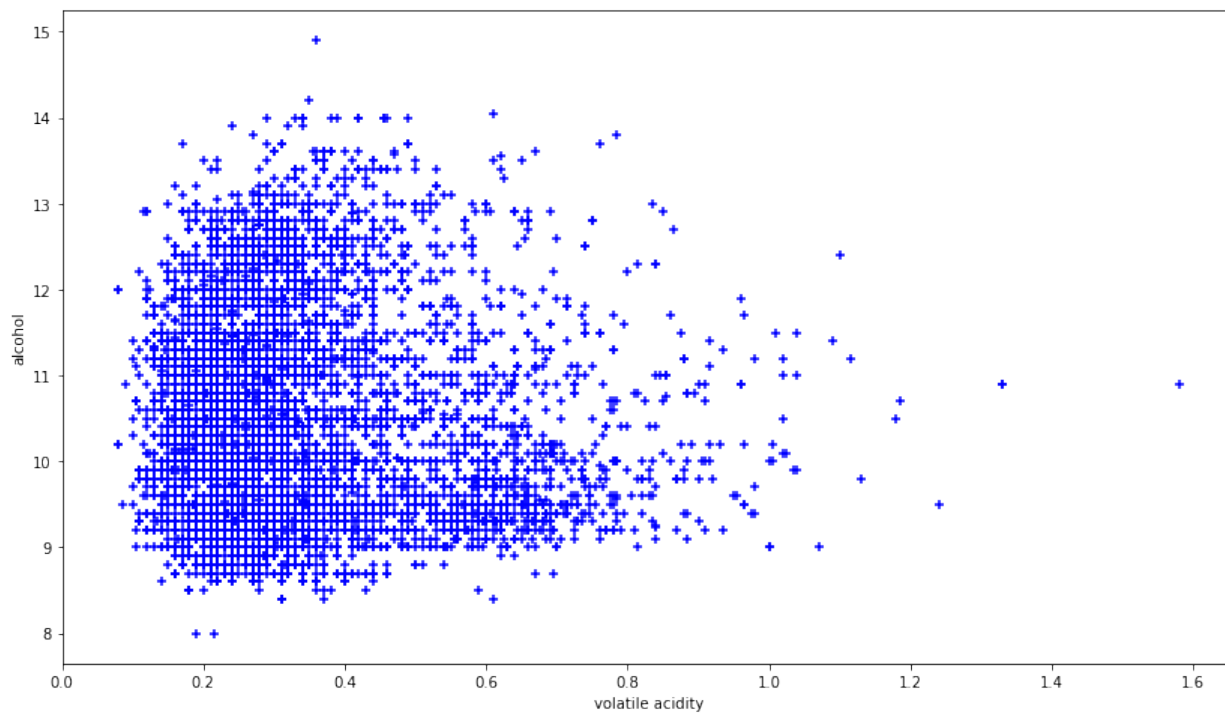
```
sea.lmplot(x='alcohol', y='fixed acidity', data = data)
#plt.show()
```

<seaborn.axisgrid.FacetGrid at 0x2b394faca58>



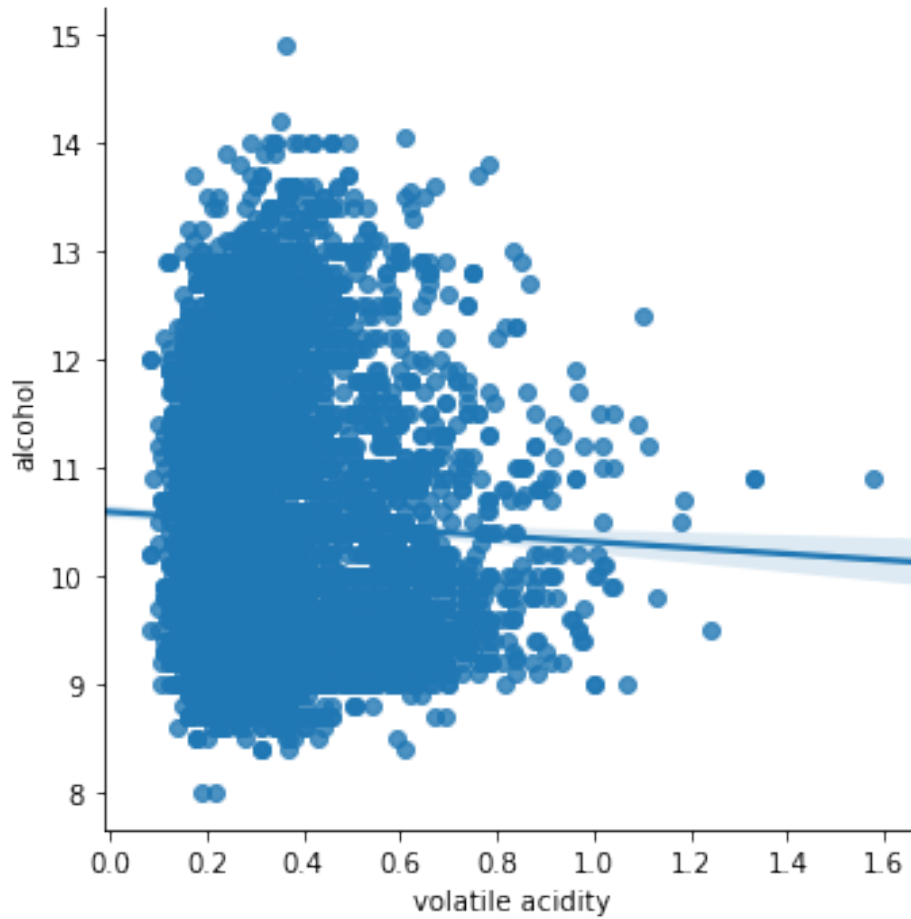
```
plt.figure(figsize=(14,8))
plt.scatter(x='volatile acidity', y='alcohol', data = data, marker
= '+', c = 'b')
plt.xlabel('volatile acidity')
plt.ylabel('alcohol')
#plt.show()
```

<matplotlib.text.Text at 0x2b39b25ac88>



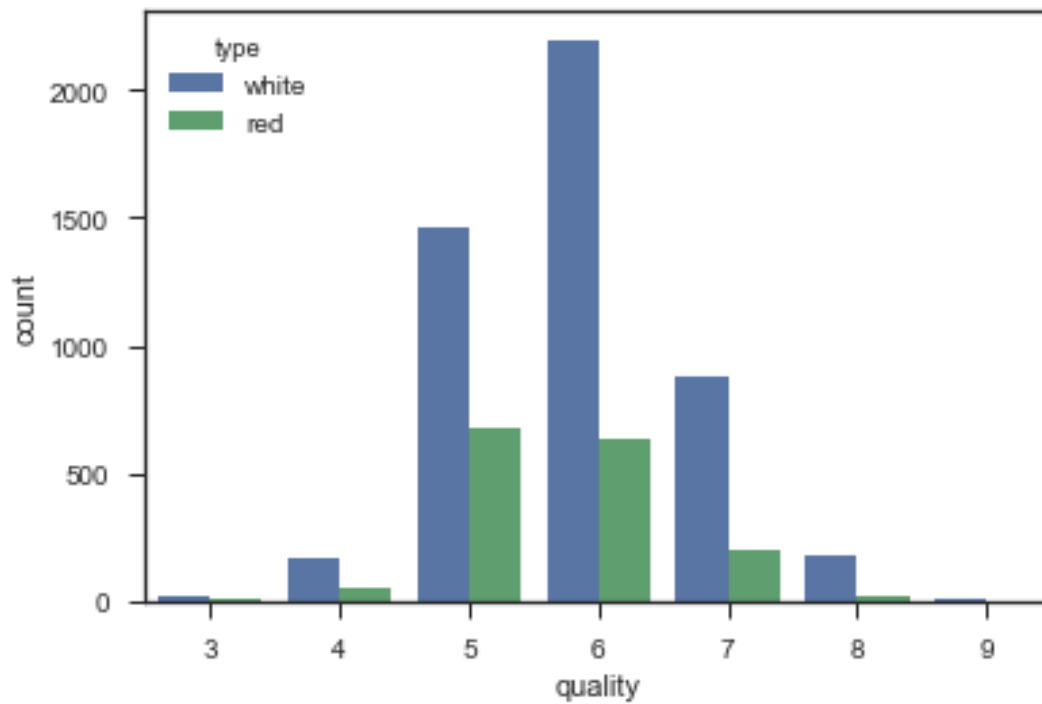
```
sea.lmplot(x='volatile acidity', y='alcohol', data = data)  
#plt.plot()
```

```
<seaborn.axisgrid.FacetGrid at 0x2b39b2e7ef0>
```

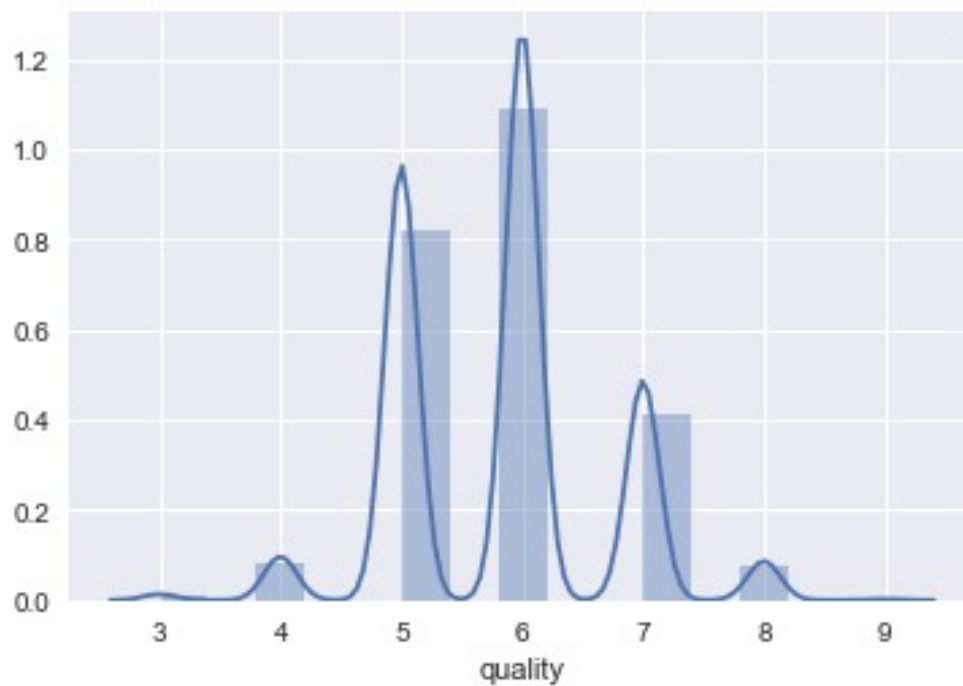


```
sea.set(style = 'ticks')
sea.countplot(data['quality'], hue = 'type', data = data)
#plt.show()

<matplotlib.axes._subplots.AxesSubplot at 0x2b39b25a828>
```

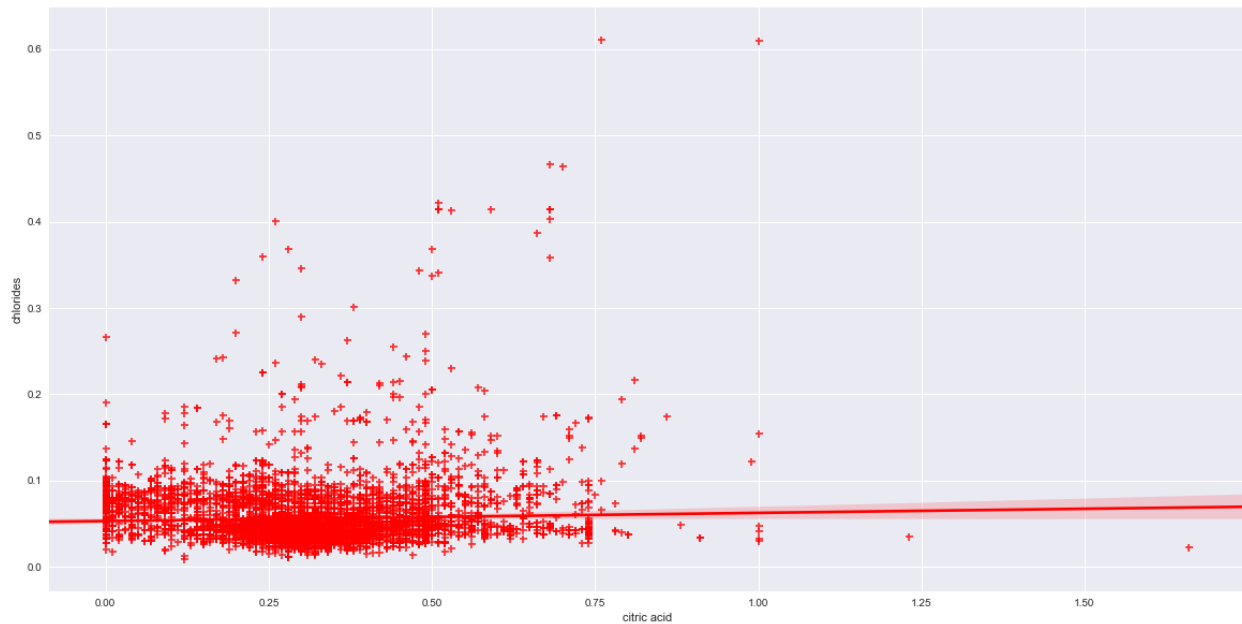



```
sea.set()  
sea.distplot(data['quality'], bins = 15)  
#plt.show()  
<matplotlib.axes._subplots.AxesSubplot at 0x2b39b724320>
```



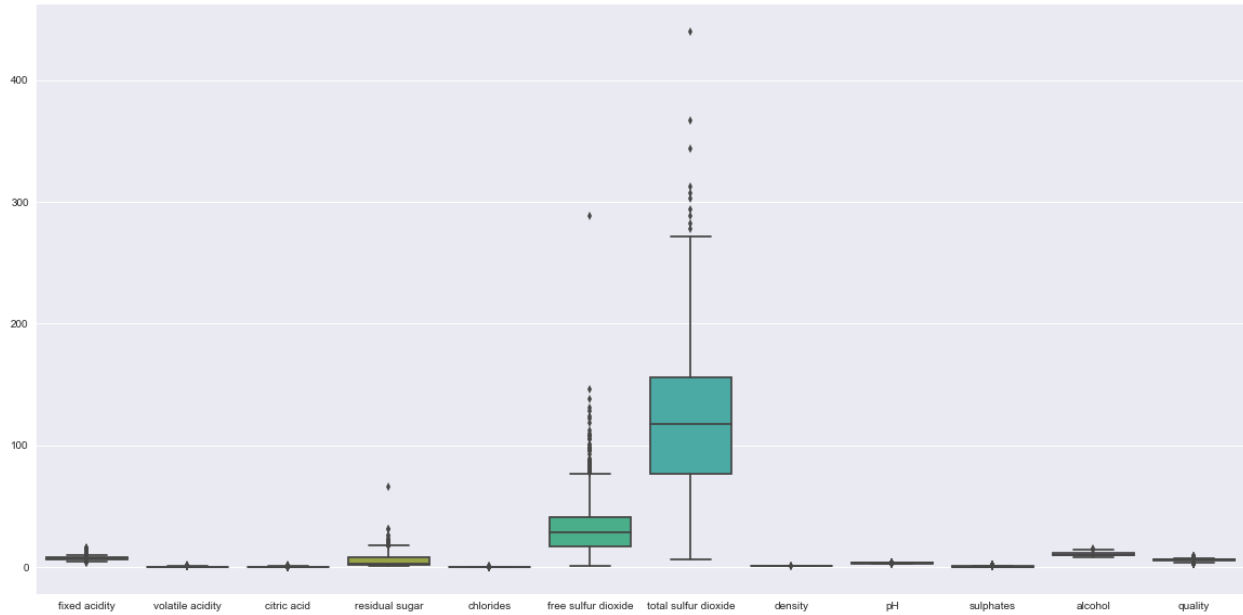
```
plt.figure(figsize = (20, 10))
sea.regplot(x = 'citric acid', y = 'chlorides', data = data, marker =
'+', color = 'r')
#plt.show()
```

<matplotlib.axes._subplots.AxesSubplot at 0x2b39b873d68>



```
sea.set()
plt.figure(figsize = (20, 10))
sea.boxplot(data = data)
#plt.show()
```

<matplotlib.axes._subplots.AxesSubplot at 0x2b39b878b70>



```
low = data['free sulfur dioxide'].mean() - 3*data['free sulfur
dioxide'].std()
high = data['free sulfur dioxide'].mean() + 3*data['free sulfur
dioxide'].std()
print('low = ', low)
print('high = ', high)
```

```
low = -22.722879937833156
high = 83.77351869418224
```

```
data_copy = data[(data['free sulfur dioxide'] > low) & (data['free
sulfur dioxide'] < high)]
print(data_copy)
```

	type	fixed acidity	volatile acidity	citric acid	residual
sugar \					
0	white	7.000000	0.270000	0.36	
20.70					
1	white	6.300000	0.300000	0.34	
1.60					
2	white	8.100000	0.280000	0.40	
6.90					
3	white	7.200000	0.230000	0.32	
8.50					
4	white	7.200000	0.230000	0.32	
8.50					
5	white	8.100000	0.280000	0.40	
6.90					
6	white	6.200000	0.320000	0.16	
7.00					
7	white	7.000000	0.270000	0.36	

20.70				
8	white	6.300000	0.300000	0.34
1.60				
9	white	8.100000	0.220000	0.43
1.50				
10	white	8.100000	0.270000	0.41
1.45				
11	white	8.600000	0.230000	0.40
4.20				
12	white	7.900000	0.180000	0.37
1.20				
13	white	6.600000	0.160000	0.40
1.50				
14	white	8.300000	0.420000	0.62
19.25				
15	white	6.600000	0.170000	0.38
1.50				
16	white	6.300000	0.480000	0.04
1.10				
17	white	7.216579	0.660000	0.48
1.20				
18	white	7.400000	0.340000	0.42
1.10				
19	white	6.500000	0.310000	0.14
7.50				
20	white	6.200000	0.660000	0.48
1.20				
21	white	6.400000	0.310000	0.38
2.90				
22	white	6.800000	0.260000	0.42
1.70				
23	white	7.600000	0.670000	0.14
1.50				
24	white	6.600000	0.270000	0.41
1.30				
25	white	7.000000	0.250000	0.32
9.00				
26	white	6.900000	0.240000	0.35
1.00				
27	white	7.000000	0.280000	0.39
8.70				
28	white	7.400000	0.270000	0.48
1.10				
29	white	7.200000	0.320000	0.36
2.00				
...
...				
6467	red	6.200000	0.510000	0.14
1.90				

6468	red	6.400000	0.360000	0.53
2.20				
6469	red	6.400000	0.380000	0.14
2.20				
6470	red	7.300000	0.690000	0.32
2.20				
6471	red	6.000000	0.580000	0.20
2.40				
6472	red	5.600000	0.310000	0.78
13.90				
6473	red	7.500000	0.520000	0.40
2.20				
6474	red	8.000000	0.300000	0.63
1.60				
6475	red	6.200000	0.700000	0.15
5.10				
6476	red	6.800000	0.670000	0.15
1.80				
6477	red	6.200000	0.560000	0.09
1.70				
6478	red	7.400000	0.350000	0.33
2.40				
6479	red	6.200000	0.560000	0.09
1.70				
6480	red	6.100000	0.715000	0.10
2.60				
6481	red	6.200000	0.460000	0.29
2.10				
6482	red	6.700000	0.320000	0.44
2.40				
6483	red	7.200000	0.390000	0.44
2.60				
6484	red	7.500000	0.310000	0.41
2.40				
6485	red	5.800000	0.610000	0.11
1.80				
6486	red	7.200000	0.339691	0.33
2.50				
6487	red	6.600000	0.725000	0.20
7.80				
6488	red	6.300000	0.550000	0.15
1.80				
6489	red	5.400000	0.740000	0.09
1.70				
6490	red	6.300000	0.510000	0.13
2.30				
6491	red	6.800000	0.620000	0.08
1.90				
6492	red	6.200000	0.600000	0.08

2.00				
6493	red	5.900000	0.550000	0.10
2.20				
6494	red	6.300000	0.510000	0.13
2.30				
6495	red	5.900000	0.645000	0.12
2.00				
6496	red	6.000000	0.310000	0.47
3.60				
	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				
5	0.050	30.0	97.0	0.99510
3.26				
6	0.045	30.0	136.0	0.99490
3.18				
7	0.045	45.0	170.0	1.00100
3.00				
8	0.049	14.0	132.0	0.99400
3.30				
9	0.044	28.0	129.0	0.99380
3.22				
10	0.033	11.0	63.0	0.99080
2.99				
11	0.035	17.0	109.0	0.99470
3.14				
12	0.040	16.0	75.0	0.99200
3.18				
13	0.044	48.0	143.0	0.99120
3.54				
14	0.040	41.0	172.0	1.00020
2.98				
15	0.032	28.0	112.0	0.99140
3.25				
16	0.046	30.0	99.0	0.99280
3.24				
17	0.029	29.0	75.0	0.98920
3.33				
18	0.033	17.0	171.0	0.99170

3.12				
19	0.044	34.0	133.0	0.99550
3.22				
20	0.029	29.0	75.0	0.98920
3.33				
21	0.038	19.0	102.0	0.99120
3.17				
22	0.049	41.0	122.0	0.99300
3.47				
23	0.074	25.0	168.0	0.99370
3.05				
24	0.052	16.0	142.0	0.99510
3.42				
25	0.046	56.0	245.0	0.99550
3.25				
26	0.052	35.0	146.0	0.99300
3.45				
27	0.051	32.0	141.0	0.99610
3.38				
28	0.047	17.0	132.0	0.99140
3.19				
29	0.033	37.0	114.0	0.99060
3.10				
...
...				
6467	0.056	15.0	34.0	0.99396
3.48				
6468	0.230	19.0	35.0	0.99340
3.37				
6469	0.038	15.0	25.0	0.99514
3.44				
6470	0.069	35.0	104.0	0.99632
3.33				
6471	0.075	15.0	50.0	0.99467
3.58				
6472	0.074	23.0	92.0	0.99677
3.39				
6473	0.060	12.0	20.0	0.99474
3.26				
6474	0.081	16.0	29.0	0.99588
3.30				
6475	0.076	13.0	27.0	0.99622
3.54				
6476	0.118	13.0	20.0	0.99540
3.42				
6477	0.053	24.0	32.0	0.99402
3.54				
6478	0.068	9.0	26.0	0.99470
3.36				

6479	0.053	24.0	32.0	0.99402
3.54				
6480	0.053	13.0	27.0	0.99362
3.57				
6481	0.074	32.0	98.0	0.99578
3.33				
6482	0.061	24.0	34.0	0.99484
3.29				
6483	0.066	22.0	48.0	0.99494
3.30				
6484	0.065	34.0	60.0	0.99492
3.34				
6485	0.066	18.0	28.0	0.99483
3.55				
6486	0.068	34.0	102.0	0.99414
3.27				
6487	0.073	29.0	79.0	0.99770
3.29				
6488	0.077	26.0	35.0	0.99314
3.32				
6489	0.089	16.0	26.0	0.99402
3.67				
6490	0.076	29.0	40.0	0.99574
3.42				
6491	0.068	28.0	38.0	0.99651
3.42				
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	
0	0.450000	8.8	6	
1	0.490000	9.5	6	
2	0.440000	10.1	6	
3	0.400000	9.9	6	
4	0.400000	9.9	6	
5	0.440000	10.1	6	
6	0.470000	9.6	6	
7	0.450000	8.8	6	
8	0.490000	9.5	6	
9	0.450000	11.0	6	
10	0.560000	12.0	5	

11	0.530000	9.7	5
12	0.630000	10.8	5
13	0.520000	12.4	7
14	0.670000	9.7	5
15	0.550000	11.4	7
16	0.360000	9.6	6
17	0.390000	12.8	8
18	0.530000	11.3	6
19	0.500000	9.5	5
20	0.390000	12.8	8
21	0.350000	11.0	7
22	0.480000	10.5	8
23	0.510000	9.3	5
24	0.470000	10.0	6
25	0.500000	10.4	6
26	0.440000	10.0	6
27	0.530000	10.5	6
28	0.490000	11.6	6
29	0.710000	12.3	7
...
6467	0.570000	11.5	6
6468	0.930000	12.4	6
6469	0.650000	11.1	6
6470	0.510000	9.5	5
6471	0.670000	12.5	6
6472	0.480000	10.5	6
6473	0.640000	11.8	6
6474	0.780000	10.8	6
6475	0.600000	11.9	6
6476	0.670000	11.3	6
6477	0.600000	11.3	5
6478	0.600000	11.9	6
6479	0.600000	11.3	5
6480	0.500000	11.9	5
6481	0.620000	9.8	5
6482	0.800000	11.6	7
6483	0.840000	11.5	6
6484	0.850000	11.4	6
6485	0.660000	10.9	6
6486	0.780000	12.8	6
6487	0.540000	9.2	5
6488	0.820000	11.6	6
6489	0.560000	11.6	6
6490	0.750000	11.0	6
6491	0.820000	9.5	6
6492	0.580000	10.5	5
6493	0.531215	11.2	6
6494	0.750000	11.0	6
6495	0.710000	10.2	5
6496	0.660000	11.0	6

```

[6461 rows x 13 columns]
data.shape[0]
6497
data_copy.shape[0]
6461
difference = data.shape[0] - data_copy.shape[0]
difference
36
low = data_copy['total sulfur dioxide'].mean() - 3*data_copy['total sulfur dioxide'].std()
high = data_copy['total sulfur dioxide'].mean() + 3*data_copy['total sulfur dioxide'].std()
print('low = ', low)
print('high = ', high)
low = -53.15243132839596
high = 283.65436601342924
data_copy2 = data_copy[(data_copy['total sulfur dioxide'] > low) & (data_copy['total sulfur dioxide'] < high)]
print(data_copy2)

```

	type	fixed acidity	volatile acidity	citric acid	residual
sugar \					
0	white	7.000000	0.270000	0.36	
20.70					
1	white	6.300000	0.300000	0.34	
1.60					
2	white	8.100000	0.280000	0.40	
6.90					
3	white	7.200000	0.230000	0.32	
8.50					
4	white	7.200000	0.230000	0.32	
8.50					
5	white	8.100000	0.280000	0.40	
6.90					
6	white	6.200000	0.320000	0.16	
7.00					
7	white	7.000000	0.270000	0.36	
20.70					
8	white	6.300000	0.300000	0.34	
1.60					
9	white	8.100000	0.220000	0.43	
1.50					

10	white	8.100000	0.270000	0.41
1.45				
11	white	8.600000	0.230000	0.40
4.20				
12	white	7.900000	0.180000	0.37
1.20				
13	white	6.600000	0.160000	0.40
1.50				
14	white	8.300000	0.420000	0.62
19.25				
15	white	6.600000	0.170000	0.38
1.50				
16	white	6.300000	0.480000	0.04
1.10				
17	white	7.216579	0.660000	0.48
1.20				
18	white	7.400000	0.340000	0.42
1.10				
19	white	6.500000	0.310000	0.14
7.50				
20	white	6.200000	0.660000	0.48
1.20				
21	white	6.400000	0.310000	0.38
2.90				
22	white	6.800000	0.260000	0.42
1.70				
23	white	7.600000	0.670000	0.14
1.50				
24	white	6.600000	0.270000	0.41
1.30				
25	white	7.000000	0.250000	0.32
9.00				
26	white	6.900000	0.240000	0.35
1.00				
27	white	7.000000	0.280000	0.39
8.70				
28	white	7.400000	0.270000	0.48
1.10				
29	white	7.200000	0.320000	0.36
2.00				
...
...				
6467	red	6.200000	0.510000	0.14
1.90				
6468	red	6.400000	0.360000	0.53
2.20				
6469	red	6.400000	0.380000	0.14
2.20				
6470	red	7.300000	0.690000	0.32

2.20				
6471	red	6.000000	0.580000	0.20
2.40				
6472	red	5.600000	0.310000	0.78
13.90				
6473	red	7.500000	0.520000	0.40
2.20				
6474	red	8.000000	0.300000	0.63
1.60				
6475	red	6.200000	0.700000	0.15
5.10				
6476	red	6.800000	0.670000	0.15
1.80				
6477	red	6.200000	0.560000	0.09
1.70				
6478	red	7.400000	0.350000	0.33
2.40				
6479	red	6.200000	0.560000	0.09
1.70				
6480	red	6.100000	0.715000	0.10
2.60				
6481	red	6.200000	0.460000	0.29
2.10				
6482	red	6.700000	0.320000	0.44
2.40				
6483	red	7.200000	0.390000	0.44
2.60				
6484	red	7.500000	0.310000	0.41
2.40				
6485	red	5.800000	0.610000	0.11
1.80				
6486	red	7.200000	0.339691	0.33
2.50				
6487	red	6.600000	0.725000	0.20
7.80				
6488	red	6.300000	0.550000	0.15
1.80				
6489	red	5.400000	0.740000	0.09
1.70				
6490	red	6.300000	0.510000	0.13
2.30				
6491	red	6.800000	0.620000	0.08
1.90				
6492	red	6.200000	0.600000	0.08
2.00				
6493	red	5.900000	0.550000	0.10
2.20				
6494	red	6.300000	0.510000	0.13
2.30				

6495	red	5.900000	0.645000	0.12
2.00				
6496	red	6.000000	0.310000	0.47
3.60				
	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				
5	0.050	30.0	97.0	0.99510
3.26				
6	0.045	30.0	136.0	0.99490
3.18				
7	0.045	45.0	170.0	1.00100
3.00				
8	0.049	14.0	132.0	0.99400
3.30				
9	0.044	28.0	129.0	0.99380
3.22				
10	0.033	11.0	63.0	0.99080
2.99				
11	0.035	17.0	109.0	0.99470
3.14				
12	0.040	16.0	75.0	0.99200
3.18				
13	0.044	48.0	143.0	0.99120
3.54				
14	0.040	41.0	172.0	1.00020
2.98				
15	0.032	28.0	112.0	0.99140
3.25				
16	0.046	30.0	99.0	0.99280
3.24				
17	0.029	29.0	75.0	0.98920
3.33				
18	0.033	17.0	171.0	0.99170
3.12				
19	0.044	34.0	133.0	0.99550
3.22				
20	0.029	29.0	75.0	0.98920
3.33				

21	0.038	19.0	102.0	0.99120
3.17				
22	0.049	41.0	122.0	0.99300
3.47				
23	0.074	25.0	168.0	0.99370
3.05				
24	0.052	16.0	142.0	0.99510
3.42				
25	0.046	56.0	245.0	0.99550
3.25				
26	0.052	35.0	146.0	0.99300
3.45				
27	0.051	32.0	141.0	0.99610
3.38				
28	0.047	17.0	132.0	0.99140
3.19				
29	0.033	37.0	114.0	0.99060
3.10				
...
...				
6467	0.056	15.0	34.0	0.99396
3.48				
6468	0.230	19.0	35.0	0.99340
3.37				
6469	0.038	15.0	25.0	0.99514
3.44				
6470	0.069	35.0	104.0	0.99632
3.33				
6471	0.075	15.0	50.0	0.99467
3.58				
6472	0.074	23.0	92.0	0.99677
3.39				
6473	0.060	12.0	20.0	0.99474
3.26				
6474	0.081	16.0	29.0	0.99588
3.30				
6475	0.076	13.0	27.0	0.99622
3.54				
6476	0.118	13.0	20.0	0.99540
3.42				
6477	0.053	24.0	32.0	0.99402
3.54				
6478	0.068	9.0	26.0	0.99470
3.36				
6479	0.053	24.0	32.0	0.99402
3.54				
6480	0.053	13.0	27.0	0.99362
3.57				
6481	0.074	32.0	98.0	0.99578

3.33				
6482	0.061	24.0	34.0	0.99484
3.29				
6483	0.066	22.0	48.0	0.99494
3.30				
6484	0.065	34.0	60.0	0.99492
3.34				
6485	0.066	18.0	28.0	0.99483
3.55				
6486	0.068	34.0	102.0	0.99414
3.27				
6487	0.073	29.0	79.0	0.99770
3.29				
6488	0.077	26.0	35.0	0.99314
3.32				
6489	0.089	16.0	26.0	0.99402
3.67				
6490	0.076	29.0	40.0	0.99574
3.42				
6491	0.068	28.0	38.0	0.99651
3.42				
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
0	0.450000	8.8	6
1	0.490000	9.5	6
2	0.440000	10.1	6
3	0.400000	9.9	6
4	0.400000	9.9	6
5	0.440000	10.1	6
6	0.470000	9.6	6
7	0.450000	8.8	6
8	0.490000	9.5	6
9	0.450000	11.0	6
10	0.560000	12.0	5
11	0.530000	9.7	5
12	0.630000	10.8	5
13	0.520000	12.4	7
14	0.670000	9.7	5
15	0.550000	11.4	7

16	0.360000	9.6	6
17	0.390000	12.8	8
18	0.530000	11.3	6
19	0.500000	9.5	5
20	0.390000	12.8	8
21	0.350000	11.0	7
22	0.480000	10.5	8
23	0.510000	9.3	5
24	0.470000	10.0	6
25	0.500000	10.4	6
26	0.440000	10.0	6
27	0.530000	10.5	6
28	0.490000	11.6	6
29	0.710000	12.3	7
...
6467	0.570000	11.5	6
6468	0.930000	12.4	6
6469	0.650000	11.1	6
6470	0.510000	9.5	5
6471	0.670000	12.5	6
6472	0.480000	10.5	6
6473	0.640000	11.8	6
6474	0.780000	10.8	6
6475	0.600000	11.9	6
6476	0.670000	11.3	6
6477	0.600000	11.3	5
6478	0.600000	11.9	6
6479	0.600000	11.3	5
6480	0.500000	11.9	5
6481	0.620000	9.8	5
6482	0.800000	11.6	7
6483	0.840000	11.5	6
6484	0.850000	11.4	6
6485	0.660000	10.9	6
6486	0.780000	12.8	6
6487	0.540000	9.2	5
6488	0.820000	11.6	6
6489	0.560000	11.6	6
6490	0.750000	11.0	6
6491	0.820000	9.5	6
6492	0.580000	10.5	5
6493	0.531215	11.2	6
6494	0.750000	11.0	6
6495	0.710000	10.2	5
6496	0.660000	11.0	6

[6456 rows x 13 columns]

data_copy.shape[0] - data_copy2.shape[0]

5

```
low = data_copy2['residual sugar'].mean() - 3*data_copy2['residual  
sugar'].std()  
high = data_copy2['residual sugar'].mean() + 3*data_copy2['residual  
sugar'].std()  
print('low = ', low)  
print('high = ', high)
```

```
low = -8.835863630876144  
high = 19.712440041288158
```

```
data_copy3 = data_copy2[(data_copy2['residual sugar'] > low) &  
(data_copy2['residual sugar'] < high)]  
print(data_copy3)
```

	type	fixed acidity	volatile acidity	citric acid	residual sugar \
1	white	6.300000	0.300000	0.34	1.60
2	white	8.100000	0.280000	0.40	6.90
3	white	7.200000	0.230000	0.32	8.50
4	white	7.200000	0.230000	0.32	8.50
5	white	8.100000	0.280000	0.40	6.90
6	white	6.200000	0.320000	0.16	7.00
8	white	6.300000	0.300000	0.34	1.60
9	white	8.100000	0.220000	0.43	1.50
10	white	8.100000	0.270000	0.41	1.45
11	white	8.600000	0.230000	0.40	4.20
12	white	7.900000	0.180000	0.37	1.20
13	white	6.600000	0.160000	0.40	1.50
14	white	8.300000	0.420000	0.62	19.25
15	white	6.600000	0.170000	0.38	1.50
16	white	6.300000	0.480000	0.04	1.10
17	white	7.216579	0.660000	0.48	1.20

18	white	7.400000	0.340000	0.42
1.10				
19	white	6.500000	0.310000	0.14
7.50				
20	white	6.200000	0.660000	0.48
1.20				
21	white	6.400000	0.310000	0.38
2.90				
22	white	6.800000	0.260000	0.42
1.70				
23	white	7.600000	0.670000	0.14
1.50				
24	white	6.600000	0.270000	0.41
1.30				
25	white	7.000000	0.250000	0.32
9.00				
26	white	6.900000	0.240000	0.35
1.00				
27	white	7.000000	0.280000	0.39
8.70				
28	white	7.400000	0.270000	0.48
1.10				
29	white	7.200000	0.320000	0.36
2.00				
30	white	8.500000	0.240000	0.39
10.40				
31	white	8.300000	0.140000	0.34
1.10				
...
...				
6467	red	6.200000	0.510000	0.14
1.90				
6468	red	6.400000	0.360000	0.53
2.20				
6469	red	6.400000	0.380000	0.14
2.20				
6470	red	7.300000	0.690000	0.32
2.20				
6471	red	6.000000	0.580000	0.20
2.40				
6472	red	5.600000	0.310000	0.78
13.90				
6473	red	7.500000	0.520000	0.40
2.20				
6474	red	8.000000	0.300000	0.63
1.60				
6475	red	6.200000	0.700000	0.15
5.10				
6476	red	6.800000	0.670000	0.15

1.80				
6477	red	6.200000	0.560000	0.09
1.70				
6478	red	7.400000	0.350000	0.33
2.40				
6479	red	6.200000	0.560000	0.09
1.70				
6480	red	6.100000	0.715000	0.10
2.60				
6481	red	6.200000	0.460000	0.29
2.10				
6482	red	6.700000	0.320000	0.44
2.40				
6483	red	7.200000	0.390000	0.44
2.60				
6484	red	7.500000	0.310000	0.41
2.40				
6485	red	5.800000	0.610000	0.11
1.80				
6486	red	7.200000	0.339691	0.33
2.50				
6487	red	6.600000	0.725000	0.20
7.80				
6488	red	6.300000	0.550000	0.15
1.80				
6489	red	5.400000	0.740000	0.09
1.70				
6490	red	6.300000	0.510000	0.13
2.30				
6491	red	6.800000	0.620000	0.08
1.90				
6492	red	6.200000	0.600000	0.08
2.00				
6493	red	5.900000	0.550000	0.10
2.20				
6494	red	6.300000	0.510000	0.13
2.30				
6495	red	5.900000	0.645000	0.12
2.00				
6496	red	6.000000	0.310000	0.47
3.60				
	chlorides	free sulfur dioxide	total sulfur dioxide	density
pH \				
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				
5	0.050	30.0	97.0	0.99510
3.26				
6	0.045	30.0	136.0	0.99490
3.18				
8	0.049	14.0	132.0	0.99400
3.30				
9	0.044	28.0	129.0	0.99380
3.22				
10	0.033	11.0	63.0	0.99080
2.99				
11	0.035	17.0	109.0	0.99470
3.14				
12	0.040	16.0	75.0	0.99200
3.18				
13	0.044	48.0	143.0	0.99120
3.54				
14	0.040	41.0	172.0	1.00020
2.98				
15	0.032	28.0	112.0	0.99140
3.25				
16	0.046	30.0	99.0	0.99280
3.24				
17	0.029	29.0	75.0	0.98920
3.33				
18	0.033	17.0	171.0	0.99170
3.12				
19	0.044	34.0	133.0	0.99550
3.22				
20	0.029	29.0	75.0	0.98920
3.33				
21	0.038	19.0	102.0	0.99120
3.17				
22	0.049	41.0	122.0	0.99300
3.47				
23	0.074	25.0	168.0	0.99370
3.05				
24	0.052	16.0	142.0	0.99510
3.42				
25	0.046	56.0	245.0	0.99550
3.25				
26	0.052	35.0	146.0	0.99300
3.45				
27	0.051	32.0	141.0	0.99610
3.38				
28	0.047	17.0	132.0	0.99140
3.19				

29	0.033	37.0	114.0	0.99060
3.10				
30	0.044	20.0	142.0	0.99740
3.20				
31	0.042	7.0	47.0	0.99340
3.47				
...
...				
6467	0.056	15.0	34.0	0.99396
3.48				
6468	0.230	19.0	35.0	0.99340
3.37				
6469	0.038	15.0	25.0	0.99514
3.44				
6470	0.069	35.0	104.0	0.99632
3.33				
6471	0.075	15.0	50.0	0.99467
3.58				
6472	0.074	23.0	92.0	0.99677
3.39				
6473	0.060	12.0	20.0	0.99474
3.26				
6474	0.081	16.0	29.0	0.99588
3.30				
6475	0.076	13.0	27.0	0.99622
3.54				
6476	0.118	13.0	20.0	0.99540
3.42				
6477	0.053	24.0	32.0	0.99402
3.54				
6478	0.068	9.0	26.0	0.99470
3.36				
6479	0.053	24.0	32.0	0.99402
3.54				
6480	0.053	13.0	27.0	0.99362
3.57				
6481	0.074	32.0	98.0	0.99578
3.33				
6482	0.061	24.0	34.0	0.99484
3.29				
6483	0.066	22.0	48.0	0.99494
3.30				
6484	0.065	34.0	60.0	0.99492
3.34				
6485	0.066	18.0	28.0	0.99483
3.55				
6486	0.068	34.0	102.0	0.99414
3.27				
6487	0.073	29.0	79.0	0.99770

3.29				
6488	0.077	26.0	35.0	0.99314
3.32				
6489	0.089	16.0	26.0	0.99402
3.67				
6490	0.076	29.0	40.0	0.99574
3.42				
6491	0.068	28.0	38.0	0.99651
3.42				
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
1	0.490000	9.5	6
2	0.440000	10.1	6
3	0.400000	9.9	6
4	0.400000	9.9	6
5	0.440000	10.1	6
6	0.470000	9.6	6
8	0.490000	9.5	6
9	0.450000	11.0	6
10	0.560000	12.0	5
11	0.530000	9.7	5
12	0.630000	10.8	5
13	0.520000	12.4	7
14	0.670000	9.7	5
15	0.550000	11.4	7
16	0.360000	9.6	6
17	0.390000	12.8	8
18	0.530000	11.3	6
19	0.500000	9.5	5
20	0.390000	12.8	8
21	0.350000	11.0	7
22	0.480000	10.5	8
23	0.510000	9.3	5
24	0.470000	10.0	6
25	0.500000	10.4	6
26	0.440000	10.0	6
27	0.530000	10.5	6
28	0.490000	11.6	6
29	0.710000	12.3	7

30	0.530000	10.0	6
31	0.400000	10.2	6
...
6467	0.570000	11.5	6
6468	0.930000	12.4	6
6469	0.650000	11.1	6
6470	0.510000	9.5	5
6471	0.670000	12.5	6
6472	0.480000	10.5	6
6473	0.640000	11.8	6
6474	0.780000	10.8	6
6475	0.600000	11.9	6
6476	0.670000	11.3	6
6477	0.600000	11.3	5
6478	0.600000	11.9	6
6479	0.600000	11.3	5
6480	0.500000	11.9	5
6481	0.620000	9.8	5
6482	0.800000	11.6	7
6483	0.840000	11.5	6
6484	0.850000	11.4	6
6485	0.660000	10.9	6
6486	0.780000	12.8	6
6487	0.540000	9.2	5
6488	0.820000	11.6	6
6489	0.560000	11.6	6
6490	0.750000	11.0	6
6491	0.820000	9.5	6
6492	0.580000	10.5	5
6493	0.531215	11.2	6
6494	0.750000	11.0	6
6495	0.710000	10.2	5
6496	0.660000	11.0	6

[6430 rows x 13 columns]

data_copy2.shape[0] - data_copy3.shape[0]

26

data_copy3.isnull().sum()

type	0
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

```

```
tmp = pd.get_dummies(data_copy3['type'], drop_first = True)
```

```
data_copy3 = pd.concat([data_copy3, tmp], axis = 1)
```

```
data_copy3.drop('type', axis = 1, inplace = True)
```

```
print(data_copy3)
```

```
print('-----')
```

```
data_copy3.head()
```

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
1	6.300000	0.300000	0.34	1.60
0.049				
2	8.100000	0.280000	0.40	6.90
0.050				
3	7.200000	0.230000	0.32	8.50
0.058				
4	7.200000	0.230000	0.32	8.50
0.058				
5	8.100000	0.280000	0.40	6.90
0.050				
6	6.200000	0.320000	0.16	7.00
0.045				
8	6.300000	0.300000	0.34	1.60
0.049				
9	8.100000	0.220000	0.43	1.50
0.044				
10	8.100000	0.270000	0.41	1.45
0.033				
11	8.600000	0.230000	0.40	4.20
0.035				
12	7.900000	0.180000	0.37	1.20
0.040				
13	6.600000	0.160000	0.40	1.50
0.044				
14	8.300000	0.420000	0.62	19.25
0.040				
15	6.600000	0.170000	0.38	1.50
0.032				
16	6.300000	0.480000	0.04	1.10
0.046				
17	7.216579	0.660000	0.48	1.20
0.029				
18	7.400000	0.340000	0.42	1.10

0.033				
19	6.500000	0.310000	0.14	7.50
0.044				
20	6.200000	0.660000	0.48	1.20
0.029				
21	6.400000	0.310000	0.38	2.90
0.038				
22	6.800000	0.260000	0.42	1.70
0.049				
23	7.600000	0.670000	0.14	1.50
0.074				
24	6.600000	0.270000	0.41	1.30
0.052				
25	7.000000	0.250000	0.32	9.00
0.046				
26	6.900000	0.240000	0.35	1.00
0.052				
27	7.000000	0.280000	0.39	8.70
0.051				
28	7.400000	0.270000	0.48	1.10
0.047				
29	7.200000	0.320000	0.36	2.00
0.033				
30	8.500000	0.240000	0.39	10.40
0.044				
31	8.300000	0.140000	0.34	1.10
0.042				
...
...				
6467	6.200000	0.510000	0.14	1.90
0.056				
6468	6.400000	0.360000	0.53	2.20
0.230				
6469	6.400000	0.380000	0.14	2.20
0.038				
6470	7.300000	0.690000	0.32	2.20
0.069				
6471	6.000000	0.580000	0.20	2.40
0.075				
6472	5.600000	0.310000	0.78	13.90
0.074				
6473	7.500000	0.520000	0.40	2.20
0.060				
6474	8.000000	0.300000	0.63	1.60
0.081				
6475	6.200000	0.700000	0.15	5.10
0.076				
6476	6.800000	0.670000	0.15	1.80
0.118				

6477	6.200000	0.560000	0.09	1.70
0.053				
6478	7.400000	0.350000	0.33	2.40
0.068				
6479	6.200000	0.560000	0.09	1.70
0.053				
6480	6.100000	0.715000	0.10	2.60
0.053				
6481	6.200000	0.460000	0.29	2.10
0.074				
6482	6.700000	0.320000	0.44	2.40
0.061				
6483	7.200000	0.390000	0.44	2.60
0.066				
6484	7.500000	0.310000	0.41	2.40
0.065				
6485	5.800000	0.610000	0.11	1.80
0.066				
6486	7.200000	0.339691	0.33	2.50
0.068				
6487	6.600000	0.725000	0.20	7.80
0.073				
6488	6.300000	0.550000	0.15	1.80
0.077				
6489	5.400000	0.740000	0.09	1.70
0.089				
6490	6.300000	0.510000	0.13	2.30
0.076				
6491	6.800000	0.620000	0.08	1.90
0.068				
6492	6.200000	0.600000	0.08	2.00
0.090				
6493	5.900000	0.550000	0.10	2.20
0.062				
6494	6.300000	0.510000	0.13	2.30
0.076				
6495	5.900000	0.645000	0.12	2.00
0.075				
6496	6.000000	0.310000	0.47	3.60
0.067				
free sulfur dioxide total sulfur dioxide density pH				
1	14.0	132.0	0.99400	3.30
0.490000				
2	30.0	97.0	0.99510	3.26
0.440000				
3	47.0	186.0	0.99560	3.19
0.400000				

4	47.0	186.0	0.99560	3.19
0.400000				
5	30.0	97.0	0.99510	3.26
0.440000				
6	30.0	136.0	0.99490	3.18
0.470000				
8	14.0	132.0	0.99400	3.30
0.490000				
9	28.0	129.0	0.99380	3.22
0.450000				
10	11.0	63.0	0.99080	2.99
0.560000				
11	17.0	109.0	0.99470	3.14
0.530000				
12	16.0	75.0	0.99200	3.18
0.630000				
13	48.0	143.0	0.99120	3.54
0.520000				
14	41.0	172.0	1.00020	2.98
0.670000				
15	28.0	112.0	0.99140	3.25
0.550000				
16	30.0	99.0	0.99280	3.24
0.360000				
17	29.0	75.0	0.98920	3.33
0.390000				
18	17.0	171.0	0.99170	3.12
0.530000				
19	34.0	133.0	0.99550	3.22
0.500000				
20	29.0	75.0	0.98920	3.33
0.390000				
21	19.0	102.0	0.99120	3.17
0.350000				
22	41.0	122.0	0.99300	3.47
0.480000				
23	25.0	168.0	0.99370	3.05
0.510000				
24	16.0	142.0	0.99510	3.42
0.470000				
25	56.0	245.0	0.99550	3.25
0.500000				
26	35.0	146.0	0.99300	3.45
0.440000				
27	32.0	141.0	0.99610	3.38
0.530000				
28	17.0	132.0	0.99140	3.19
0.490000				
29	37.0	114.0	0.99060	3.10

0.710000				
30	20.0	142.0	0.99740	3.20
0.530000				
31	7.0	47.0	0.99340	3.47
0.400000				
...
...				
6467	15.0	34.0	0.99396	3.48
0.570000				
6468	19.0	35.0	0.99340	3.37
0.930000				
6469	15.0	25.0	0.99514	3.44
0.650000				
6470	35.0	104.0	0.99632	3.33
0.510000				
6471	15.0	50.0	0.99467	3.58
0.670000				
6472	23.0	92.0	0.99677	3.39
0.480000				
6473	12.0	20.0	0.99474	3.26
0.640000				
6474	16.0	29.0	0.99588	3.30
0.780000				
6475	13.0	27.0	0.99622	3.54
0.600000				
6476	13.0	20.0	0.99540	3.42
0.670000				
6477	24.0	32.0	0.99402	3.54
0.600000				
6478	9.0	26.0	0.99470	3.36
0.600000				
6479	24.0	32.0	0.99402	3.54
0.600000				
6480	13.0	27.0	0.99362	3.57
0.500000				
6481	32.0	98.0	0.99578	3.33
0.620000				
6482	24.0	34.0	0.99484	3.29
0.800000				
6483	22.0	48.0	0.99494	3.30
0.840000				
6484	34.0	60.0	0.99492	3.34
0.850000				
6485	18.0	28.0	0.99483	3.55
0.660000				
6486	34.0	102.0	0.99414	3.27
0.780000				
6487	29.0	79.0	0.99770	3.29
0.540000				

6488	26.0	35.0	0.99314	3.32
0.820000				
6489	16.0	26.0	0.99402	3.67
0.560000				
6490	29.0	40.0	0.99574	3.42
0.750000				
6491	28.0	38.0	0.99651	3.42
0.820000				
6492	32.0	44.0	0.99490	3.45
0.580000				
6493	39.0	51.0	0.99512	3.52
0.531215				
6494	29.0	40.0	0.99574	3.42
0.750000				
6495	32.0	44.0	0.99547	3.57
0.710000				
6496	18.0	42.0	0.99549	3.39
0.660000				

	alcohol	quality	white
1	9.5	6	1
2	10.1	6	1
3	9.9	6	1
4	9.9	6	1
5	10.1	6	1
6	9.6	6	1
8	9.5	6	1
9	11.0	6	1
10	12.0	5	1
11	9.7	5	1
12	10.8	5	1
13	12.4	7	1
14	9.7	5	1
15	11.4	7	1
16	9.6	6	1
17	12.8	8	1
18	11.3	6	1
19	9.5	5	1
20	12.8	8	1
21	11.0	7	1
22	10.5	8	1
23	9.3	5	1
24	10.0	6	1
25	10.4	6	1
26	10.0	6	1
27	10.5	6	1
28	11.6	6	1
29	12.3	7	1
30	10.0	6	1

31	10.2	6	1
...
6467	11.5	6	0
6468	12.4	6	0
6469	11.1	6	0
6470	9.5	5	0
6471	12.5	6	0
6472	10.5	6	0
6473	11.8	6	0
6474	10.8	6	0
6475	11.9	6	0
6476	11.3	6	0
6477	11.3	5	0
6478	11.9	6	0
6479	11.3	5	0
6480	11.9	5	0
6481	9.8	5	0
6482	11.6	7	0
6483	11.5	6	0
6484	11.4	6	0
6485	10.9	6	0
6486	12.8	6	0
6487	9.2	5	0
6488	11.6	6	0
6489	11.6	6	0
6490	11.0	6	0
6491	9.5	6	0
6492	10.5	5	0
6493	11.2	6	0
6494	11.0	6	0
6495	10.2	5	0
6496	11.0	6	0

[6430 rows x 13 columns]

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
1	6.3	0.30	0.34	1.6
0.049				
2	8.1	0.28	0.40	6.9
0.050				
3	7.2	0.23	0.32	8.5
0.058				
4	7.2	0.23	0.32	8.5
0.058				
5	8.1	0.28	0.40	6.9
0.050				

free sulfur dioxide	total sulfur dioxide	density	pH	sulphates
---------------------	----------------------	---------	----	-----------

\					
1	14.0	132.0	0.9940	3.30	0.49
2	30.0	97.0	0.9951	3.26	0.44
3	47.0	186.0	0.9956	3.19	0.40
4	47.0	186.0	0.9956	3.19	0.40
5	30.0	97.0	0.9951	3.26	0.44

	alcohol	quality	white
1	9.5	6	1
2	10.1	6	1
3	9.9	6	1
4	9.9	6	1
5	10.1	6	1

```
data_copy3.quality.value_counts()
print(data_copy3)
```

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
1	6.300000	0.300000	0.34	1.60
0.049				
2	8.100000	0.280000	0.40	6.90
0.050				
3	7.200000	0.230000	0.32	8.50
0.058				
4	7.200000	0.230000	0.32	8.50
0.058				
5	8.100000	0.280000	0.40	6.90
0.050				
6	6.200000	0.320000	0.16	7.00
0.045				
8	6.300000	0.300000	0.34	1.60
0.049				
9	8.100000	0.220000	0.43	1.50
0.044				
10	8.100000	0.270000	0.41	1.45
0.033				
11	8.600000	0.230000	0.40	4.20
0.035				
12	7.900000	0.180000	0.37	1.20
0.040				
13	6.600000	0.160000	0.40	1.50
0.044				
14	8.300000	0.420000	0.62	19.25
0.040				

15	6.600000	0.170000	0.38	1.50
0.032				
16	6.300000	0.480000	0.04	1.10
0.046				
17	7.216579	0.660000	0.48	1.20
0.029				
18	7.400000	0.340000	0.42	1.10
0.033				
19	6.500000	0.310000	0.14	7.50
0.044				
20	6.200000	0.660000	0.48	1.20
0.029				
21	6.400000	0.310000	0.38	2.90
0.038				
22	6.800000	0.260000	0.42	1.70
0.049				
23	7.600000	0.670000	0.14	1.50
0.074				
24	6.600000	0.270000	0.41	1.30
0.052				
25	7.000000	0.250000	0.32	9.00
0.046				
26	6.900000	0.240000	0.35	1.00
0.052				
27	7.000000	0.280000	0.39	8.70
0.051				
28	7.400000	0.270000	0.48	1.10
0.047				
29	7.200000	0.320000	0.36	2.00
0.033				
30	8.500000	0.240000	0.39	10.40
0.044				
31	8.300000	0.140000	0.34	1.10
0.042				
...
...				
6467	6.200000	0.510000	0.14	1.90
0.056				
6468	6.400000	0.360000	0.53	2.20
0.230				
6469	6.400000	0.380000	0.14	2.20
0.038				
6470	7.300000	0.690000	0.32	2.20
0.069				
6471	6.000000	0.580000	0.20	2.40
0.075				
6472	5.600000	0.310000	0.78	13.90
0.074				
6473	7.500000	0.520000	0.40	2.20

0.060				
6474	8.000000	0.300000	0.63	1.60
0.081				
6475	6.200000	0.700000	0.15	5.10
0.076				
6476	6.800000	0.670000	0.15	1.80
0.118				
6477	6.200000	0.560000	0.09	1.70
0.053				
6478	7.400000	0.350000	0.33	2.40
0.068				
6479	6.200000	0.560000	0.09	1.70
0.053				
6480	6.100000	0.715000	0.10	2.60
0.053				
6481	6.200000	0.460000	0.29	2.10
0.074				
6482	6.700000	0.320000	0.44	2.40
0.061				
6483	7.200000	0.390000	0.44	2.60
0.066				
6484	7.500000	0.310000	0.41	2.40
0.065				
6485	5.800000	0.610000	0.11	1.80
0.066				
6486	7.200000	0.339691	0.33	2.50
0.068				
6487	6.600000	0.725000	0.20	7.80
0.073				
6488	6.300000	0.550000	0.15	1.80
0.077				
6489	5.400000	0.740000	0.09	1.70
0.089				
6490	6.300000	0.510000	0.13	2.30
0.076				
6491	6.800000	0.620000	0.08	1.90
0.068				
6492	6.200000	0.600000	0.08	2.00
0.090				
6493	5.900000	0.550000	0.10	2.20
0.062				
6494	6.300000	0.510000	0.13	2.30
0.076				
6495	5.900000	0.645000	0.12	2.00
0.075				
6496	6.000000	0.310000	0.47	3.60
0.067				
free sulfur dioxide	total sulfur dioxide	density	pH	

sulphates \				
1	14.0	132.0	0.99400	3.30
0.490000				
2	30.0	97.0	0.99510	3.26
0.440000				
3	47.0	186.0	0.99560	3.19
0.400000				
4	47.0	186.0	0.99560	3.19
0.400000				
5	30.0	97.0	0.99510	3.26
0.440000				
6	30.0	136.0	0.99490	3.18
0.470000				
8	14.0	132.0	0.99400	3.30
0.490000				
9	28.0	129.0	0.99380	3.22
0.450000				
10	11.0	63.0	0.99080	2.99
0.560000				
11	17.0	109.0	0.99470	3.14
0.530000				
12	16.0	75.0	0.99200	3.18
0.630000				
13	48.0	143.0	0.99120	3.54
0.520000				
14	41.0	172.0	1.00020	2.98
0.670000				
15	28.0	112.0	0.99140	3.25
0.550000				
16	30.0	99.0	0.99280	3.24
0.360000				
17	29.0	75.0	0.98920	3.33
0.390000				
18	17.0	171.0	0.99170	3.12
0.530000				
19	34.0	133.0	0.99550	3.22
0.500000				
20	29.0	75.0	0.98920	3.33
0.390000				
21	19.0	102.0	0.99120	3.17
0.350000				
22	41.0	122.0	0.99300	3.47
0.480000				
23	25.0	168.0	0.99370	3.05
0.510000				
24	16.0	142.0	0.99510	3.42
0.470000				
25	56.0	245.0	0.99550	3.25
0.500000				

26	35.0	146.0	0.99300	3.45
0.440000				
27	32.0	141.0	0.99610	3.38
0.530000				
28	17.0	132.0	0.99140	3.19
0.490000				
29	37.0	114.0	0.99060	3.10
0.710000				
30	20.0	142.0	0.99740	3.20
0.530000				
31	7.0	47.0	0.99340	3.47
0.400000				
...
...				
6467	15.0	34.0	0.99396	3.48
0.570000				
6468	19.0	35.0	0.99340	3.37
0.930000				
6469	15.0	25.0	0.99514	3.44
0.650000				
6470	35.0	104.0	0.99632	3.33
0.510000				
6471	15.0	50.0	0.99467	3.58
0.670000				
6472	23.0	92.0	0.99677	3.39
0.480000				
6473	12.0	20.0	0.99474	3.26
0.640000				
6474	16.0	29.0	0.99588	3.30
0.780000				
6475	13.0	27.0	0.99622	3.54
0.600000				
6476	13.0	20.0	0.99540	3.42
0.670000				
6477	24.0	32.0	0.99402	3.54
0.600000				
6478	9.0	26.0	0.99470	3.36
0.600000				
6479	24.0	32.0	0.99402	3.54
0.600000				
6480	13.0	27.0	0.99362	3.57
0.500000				
6481	32.0	98.0	0.99578	3.33
0.620000				
6482	24.0	34.0	0.99484	3.29
0.800000				
6483	22.0	48.0	0.99494	3.30
0.840000				
6484	34.0	60.0	0.99492	3.34

0.850000					
6485	18.0	28.0	0.99483	3.55	
0.660000					
6486	34.0	102.0	0.99414	3.27	
0.780000					
6487	29.0	79.0	0.99770	3.29	
0.540000					
6488	26.0	35.0	0.99314	3.32	
0.820000					
6489	16.0	26.0	0.99402	3.67	
0.560000					
6490	29.0	40.0	0.99574	3.42	
0.750000					
6491	28.0	38.0	0.99651	3.42	
0.820000					
6492	32.0	44.0	0.99490	3.45	
0.580000					
6493	39.0	51.0	0.99512	3.52	
0.531215					
6494	29.0	40.0	0.99574	3.42	
0.750000					
6495	32.0	44.0	0.99547	3.57	
0.710000					
6496	18.0	42.0	0.99549	3.39	
0.660000					

	alcohol	quality	white
1	9.5	6	1
2	10.1	6	1
3	9.9	6	1
4	9.9	6	1
5	10.1	6	1
6	9.6	6	1
8	9.5	6	1
9	11.0	6	1
10	12.0	5	1
11	9.7	5	1
12	10.8	5	1
13	12.4	7	1
14	9.7	5	1
15	11.4	7	1
16	9.6	6	1
17	12.8	8	1
18	11.3	6	1
19	9.5	5	1
20	12.8	8	1
21	11.0	7	1
22	10.5	8	1
23	9.3	5	1

24	10.0	6	1
25	10.4	6	1
26	10.0	6	1
27	10.5	6	1
28	11.6	6	1
29	12.3	7	1
30	10.0	6	1
31	10.2	6	1
...
6467	11.5	6	0
6468	12.4	6	0
6469	11.1	6	0
6470	9.5	5	0
6471	12.5	6	0
6472	10.5	6	0
6473	11.8	6	0
6474	10.8	6	0
6475	11.9	6	0
6476	11.3	6	0
6477	11.3	5	0
6478	11.9	6	0
6479	11.3	5	0
6480	11.9	5	0
6481	9.8	5	0
6482	11.6	7	0
6483	11.5	6	0
6484	11.4	6	0
6485	10.9	6	0
6486	12.8	6	0
6487	9.2	5	0
6488	11.6	6	0
6489	11.6	6	0
6490	11.0	6	0
6491	9.5	6	0
6492	10.5	5	0
6493	11.2	6	0
6494	11.0	6	0
6495	10.2	5	0
6496	11.0	6	0

[6430 rows x 13 columns]

```
quality_map = { 3 : 'Low', 4 : 'Low', 5 : 'Medium', 6 : 'Medium', 7 :
'Medium', 8 : 'High', 9 : 'High'}
data_copy3['quality'] = data_copy3['quality'].map(quality_map)
print(data_copy3)
```

```
fixed acidity  volatile acidity  citric acid  residual sugar
chlorides \
```

1	6.300000	0.300000	0.34	1.60
0.049				
2	8.100000	0.280000	0.40	6.90
0.050				
3	7.200000	0.230000	0.32	8.50
0.058				
4	7.200000	0.230000	0.32	8.50
0.058				
5	8.100000	0.280000	0.40	6.90
0.050				
6	6.200000	0.320000	0.16	7.00
0.045				
8	6.300000	0.300000	0.34	1.60
0.049				
9	8.100000	0.220000	0.43	1.50
0.044				
10	8.100000	0.270000	0.41	1.45
0.033				
11	8.600000	0.230000	0.40	4.20
0.035				
12	7.900000	0.180000	0.37	1.20
0.040				
13	6.600000	0.160000	0.40	1.50
0.044				
14	8.300000	0.420000	0.62	19.25
0.040				
15	6.600000	0.170000	0.38	1.50
0.032				
16	6.300000	0.480000	0.04	1.10
0.046				
17	7.216579	0.660000	0.48	1.20
0.029				
18	7.400000	0.340000	0.42	1.10
0.033				
19	6.500000	0.310000	0.14	7.50
0.044				
20	6.200000	0.660000	0.48	1.20
0.029				
21	6.400000	0.310000	0.38	2.90
0.038				
22	6.800000	0.260000	0.42	1.70
0.049				
23	7.600000	0.670000	0.14	1.50
0.074				
24	6.600000	0.270000	0.41	1.30
0.052				
25	7.000000	0.250000	0.32	9.00
0.046				
26	6.900000	0.240000	0.35	1.00

0.052				
27	7.000000	0.280000	0.39	8.70
0.051				
28	7.400000	0.270000	0.48	1.10
0.047				
29	7.200000	0.320000	0.36	2.00
0.033				
30	8.500000	0.240000	0.39	10.40
0.044				
31	8.300000	0.140000	0.34	1.10
0.042				
...
...				
6467	6.200000	0.510000	0.14	1.90
0.056				
6468	6.400000	0.360000	0.53	2.20
0.230				
6469	6.400000	0.380000	0.14	2.20
0.038				
6470	7.300000	0.690000	0.32	2.20
0.069				
6471	6.000000	0.580000	0.20	2.40
0.075				
6472	5.600000	0.310000	0.78	13.90
0.074				
6473	7.500000	0.520000	0.40	2.20
0.060				
6474	8.000000	0.300000	0.63	1.60
0.081				
6475	6.200000	0.700000	0.15	5.10
0.076				
6476	6.800000	0.670000	0.15	1.80
0.118				
6477	6.200000	0.560000	0.09	1.70
0.053				
6478	7.400000	0.350000	0.33	2.40
0.068				
6479	6.200000	0.560000	0.09	1.70
0.053				
6480	6.100000	0.715000	0.10	2.60
0.053				
6481	6.200000	0.460000	0.29	2.10
0.074				
6482	6.700000	0.320000	0.44	2.40
0.061				
6483	7.200000	0.390000	0.44	2.60
0.066				
6484	7.500000	0.310000	0.41	2.40
0.065				

6485	5.800000	0.610000	0.11	1.80
0.066				
6486	7.200000	0.339691	0.33	2.50
0.068				
6487	6.600000	0.725000	0.20	7.80
0.073				
6488	6.300000	0.550000	0.15	1.80
0.077				
6489	5.400000	0.740000	0.09	1.70
0.089				
6490	6.300000	0.510000	0.13	2.30
0.076				
6491	6.800000	0.620000	0.08	1.90
0.068				
6492	6.200000	0.600000	0.08	2.00
0.090				
6493	5.900000	0.550000	0.10	2.20
0.062				
6494	6.300000	0.510000	0.13	2.30
0.076				
6495	5.900000	0.645000	0.12	2.00
0.075				
6496	6.000000	0.310000	0.47	3.60
0.067				

	free sulfur dioxide	total sulfur dioxide	density	pH
free sulphates \				
1	14.0	132.0	0.99400	3.30
0.490000				
2	30.0	97.0	0.99510	3.26
0.440000				
3	47.0	186.0	0.99560	3.19
0.400000				
4	47.0	186.0	0.99560	3.19
0.400000				
5	30.0	97.0	0.99510	3.26
0.440000				
6	30.0	136.0	0.99490	3.18
0.470000				
8	14.0	132.0	0.99400	3.30
0.490000				
9	28.0	129.0	0.99380	3.22
0.450000				
10	11.0	63.0	0.99080	2.99
0.560000				
11	17.0	109.0	0.99470	3.14
0.530000				
12	16.0	75.0	0.99200	3.18
0.630000				

13	48.0	143.0	0.99120	3.54
0.520000				
14	41.0	172.0	1.00020	2.98
0.670000				
15	28.0	112.0	0.99140	3.25
0.550000				
16	30.0	99.0	0.99280	3.24
0.360000				
17	29.0	75.0	0.98920	3.33
0.390000				
18	17.0	171.0	0.99170	3.12
0.530000				
19	34.0	133.0	0.99550	3.22
0.500000				
20	29.0	75.0	0.98920	3.33
0.390000				
21	19.0	102.0	0.99120	3.17
0.350000				
22	41.0	122.0	0.99300	3.47
0.480000				
23	25.0	168.0	0.99370	3.05
0.510000				
24	16.0	142.0	0.99510	3.42
0.470000				
25	56.0	245.0	0.99550	3.25
0.500000				
26	35.0	146.0	0.99300	3.45
0.440000				
27	32.0	141.0	0.99610	3.38
0.530000				
28	17.0	132.0	0.99140	3.19
0.490000				
29	37.0	114.0	0.99060	3.10
0.710000				
30	20.0	142.0	0.99740	3.20
0.530000				
31	7.0	47.0	0.99340	3.47
0.400000				
...
...				
6467	15.0	34.0	0.99396	3.48
0.570000				
6468	19.0	35.0	0.99340	3.37
0.930000				
6469	15.0	25.0	0.99514	3.44
0.650000				
6470	35.0	104.0	0.99632	3.33
0.510000				
6471	15.0	50.0	0.99467	3.58

0.670000					
6472	23.0	92.0	0.99677	3.39	
0.480000					
6473	12.0	20.0	0.99474	3.26	
0.640000					
6474	16.0	29.0	0.99588	3.30	
0.780000					
6475	13.0	27.0	0.99622	3.54	
0.600000					
6476	13.0	20.0	0.99540	3.42	
0.670000					
6477	24.0	32.0	0.99402	3.54	
0.600000					
6478	9.0	26.0	0.99470	3.36	
0.600000					
6479	24.0	32.0	0.99402	3.54	
0.600000					
6480	13.0	27.0	0.99362	3.57	
0.500000					
6481	32.0	98.0	0.99578	3.33	
0.620000					
6482	24.0	34.0	0.99484	3.29	
0.800000					
6483	22.0	48.0	0.99494	3.30	
0.840000					
6484	34.0	60.0	0.99492	3.34	
0.850000					
6485	18.0	28.0	0.99483	3.55	
0.660000					
6486	34.0	102.0	0.99414	3.27	
0.780000					
6487	29.0	79.0	0.99770	3.29	
0.540000					
6488	26.0	35.0	0.99314	3.32	
0.820000					
6489	16.0	26.0	0.99402	3.67	
0.560000					
6490	29.0	40.0	0.99574	3.42	
0.750000					
6491	28.0	38.0	0.99651	3.42	
0.820000					
6492	32.0	44.0	0.99490	3.45	
0.580000					
6493	39.0	51.0	0.99512	3.52	
0.531215					
6494	29.0	40.0	0.99574	3.42	
0.750000					
6495	32.0	44.0	0.99547	3.57	
0.710000					

6496	18.0	42.0	0.99549	3.39
0.660000				

	alcohol	quality	white
1	9.5	Medium	1
2	10.1	Medium	1
3	9.9	Medium	1
4	9.9	Medium	1
5	10.1	Medium	1
6	9.6	Medium	1
8	9.5	Medium	1
9	11.0	Medium	1
10	12.0	Medium	1
11	9.7	Medium	1
12	10.8	Medium	1
13	12.4	Medium	1
14	9.7	Medium	1
15	11.4	Medium	1
16	9.6	Medium	1
17	12.8	High	1
18	11.3	Medium	1
19	9.5	Medium	1
20	12.8	High	1
21	11.0	Medium	1
22	10.5	High	1
23	9.3	Medium	1
24	10.0	Medium	1
25	10.4	Medium	1
26	10.0	Medium	1
27	10.5	Medium	1
28	11.6	Medium	1
29	12.3	Medium	1
30	10.0	Medium	1
31	10.2	Medium	1
...
6467	11.5	Medium	0
6468	12.4	Medium	0
6469	11.1	Medium	0
6470	9.5	Medium	0
6471	12.5	Medium	0
6472	10.5	Medium	0
6473	11.8	Medium	0
6474	10.8	Medium	0
6475	11.9	Medium	0
6476	11.3	Medium	0
6477	11.3	Medium	0
6478	11.9	Medium	0
6479	11.3	Medium	0
6480	11.9	Medium	0

6481	9.8	Medium	0
6482	11.6	Medium	0
6483	11.5	Medium	0
6484	11.4	Medium	0
6485	10.9	Medium	0
6486	12.8	Medium	0
6487	9.2	Medium	0
6488	11.6	Medium	0
6489	11.6	Medium	0
6490	11.0	Medium	0
6491	9.5	Medium	0
6492	10.5	Medium	0
6493	11.2	Medium	0
6494	11.0	Medium	0
6495	10.2	Medium	0
6496	11.0	Medium	0

[6430 rows x 13 columns]

```
data_copy3.quality.value_counts()
```

```
Medium    5997
```

```
Low        239
```

```
High       194
```

```
Name: quality, dtype: int64
```

```
map_quality = {'Low': 0, 'Medium': 1, 'High': 2}
```

```
data_copy3['quality'] = data_copy3['quality'].map(map_quality)
```

```
print(data_copy3)
```

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
1	6.300000	0.300000	0.34	1.60
0.049				
2	8.100000	0.280000	0.40	6.90
0.050				
3	7.200000	0.230000	0.32	8.50
0.058				
4	7.200000	0.230000	0.32	8.50
0.058				
5	8.100000	0.280000	0.40	6.90
0.050				
6	6.200000	0.320000	0.16	7.00
0.045				
8	6.300000	0.300000	0.34	1.60
0.049				
9	8.100000	0.220000	0.43	1.50
0.044				
10	8.100000	0.270000	0.41	1.45
0.033				

11	8.600000	0.230000	0.40	4.20
0.035				
12	7.900000	0.180000	0.37	1.20
0.040				
13	6.600000	0.160000	0.40	1.50
0.044				
14	8.300000	0.420000	0.62	19.25
0.040				
15	6.600000	0.170000	0.38	1.50
0.032				
16	6.300000	0.480000	0.04	1.10
0.046				
17	7.216579	0.660000	0.48	1.20
0.029				
18	7.400000	0.340000	0.42	1.10
0.033				
19	6.500000	0.310000	0.14	7.50
0.044				
20	6.200000	0.660000	0.48	1.20
0.029				
21	6.400000	0.310000	0.38	2.90
0.038				
22	6.800000	0.260000	0.42	1.70
0.049				
23	7.600000	0.670000	0.14	1.50
0.074				
24	6.600000	0.270000	0.41	1.30
0.052				
25	7.000000	0.250000	0.32	9.00
0.046				
26	6.900000	0.240000	0.35	1.00
0.052				
27	7.000000	0.280000	0.39	8.70
0.051				
28	7.400000	0.270000	0.48	1.10
0.047				
29	7.200000	0.320000	0.36	2.00
0.033				
30	8.500000	0.240000	0.39	10.40
0.044				
31	8.300000	0.140000	0.34	1.10
0.042				
...
...				
6467	6.200000	0.510000	0.14	1.90
0.056				
6468	6.400000	0.360000	0.53	2.20
0.230				
6469	6.400000	0.380000	0.14	2.20

0.038				
6470	7.300000	0.690000	0.32	2.20
0.069				
6471	6.000000	0.580000	0.20	2.40
0.075				
6472	5.600000	0.310000	0.78	13.90
0.074				
6473	7.500000	0.520000	0.40	2.20
0.060				
6474	8.000000	0.300000	0.63	1.60
0.081				
6475	6.200000	0.700000	0.15	5.10
0.076				
6476	6.800000	0.670000	0.15	1.80
0.118				
6477	6.200000	0.560000	0.09	1.70
0.053				
6478	7.400000	0.350000	0.33	2.40
0.068				
6479	6.200000	0.560000	0.09	1.70
0.053				
6480	6.100000	0.715000	0.10	2.60
0.053				
6481	6.200000	0.460000	0.29	2.10
0.074				
6482	6.700000	0.320000	0.44	2.40
0.061				
6483	7.200000	0.390000	0.44	2.60
0.066				
6484	7.500000	0.310000	0.41	2.40
0.065				
6485	5.800000	0.610000	0.11	1.80
0.066				
6486	7.200000	0.339691	0.33	2.50
0.068				
6487	6.600000	0.725000	0.20	7.80
0.073				
6488	6.300000	0.550000	0.15	1.80
0.077				
6489	5.400000	0.740000	0.09	1.70
0.089				
6490	6.300000	0.510000	0.13	2.30
0.076				
6491	6.800000	0.620000	0.08	1.90
0.068				
6492	6.200000	0.600000	0.08	2.00
0.090				
6493	5.900000	0.550000	0.10	2.20
0.062				

6494	6.300000	0.510000	0.13	2.30
0.076				
6495	5.900000	0.645000	0.12	2.00
0.075				
6496	6.000000	0.310000	0.47	3.60
0.067				

free sulfur dioxide sulphates \	total sulfur dioxide	density	pH
1	14.0	132.0	0.99400 3.30
0.490000			
2	30.0	97.0	0.99510 3.26
0.440000			
3	47.0	186.0	0.99560 3.19
0.400000			
4	47.0	186.0	0.99560 3.19
0.400000			
5	30.0	97.0	0.99510 3.26
0.440000			
6	30.0	136.0	0.99490 3.18
0.470000			
8	14.0	132.0	0.99400 3.30
0.490000			
9	28.0	129.0	0.99380 3.22
0.450000			
10	11.0	63.0	0.99080 2.99
0.560000			
11	17.0	109.0	0.99470 3.14
0.530000			
12	16.0	75.0	0.99200 3.18
0.630000			
13	48.0	143.0	0.99120 3.54
0.520000			
14	41.0	172.0	1.00020 2.98
0.670000			
15	28.0	112.0	0.99140 3.25
0.550000			
16	30.0	99.0	0.99280 3.24
0.360000			
17	29.0	75.0	0.98920 3.33
0.390000			
18	17.0	171.0	0.99170 3.12
0.530000			
19	34.0	133.0	0.99550 3.22
0.500000			
20	29.0	75.0	0.98920 3.33
0.390000			
21	19.0	102.0	0.99120 3.17
0.350000			

22	41.0	122.0	0.99300	3.47
0.480000				
23	25.0	168.0	0.99370	3.05
0.510000				
24	16.0	142.0	0.99510	3.42
0.470000				
25	56.0	245.0	0.99550	3.25
0.500000				
26	35.0	146.0	0.99300	3.45
0.440000				
27	32.0	141.0	0.99610	3.38
0.530000				
28	17.0	132.0	0.99140	3.19
0.490000				
29	37.0	114.0	0.99060	3.10
0.710000				
30	20.0	142.0	0.99740	3.20
0.530000				
31	7.0	47.0	0.99340	3.47
0.400000				
...
...				
6467	15.0	34.0	0.99396	3.48
0.570000				
6468	19.0	35.0	0.99340	3.37
0.930000				
6469	15.0	25.0	0.99514	3.44
0.650000				
6470	35.0	104.0	0.99632	3.33
0.510000				
6471	15.0	50.0	0.99467	3.58
0.670000				
6472	23.0	92.0	0.99677	3.39
0.480000				
6473	12.0	20.0	0.99474	3.26
0.640000				
6474	16.0	29.0	0.99588	3.30
0.780000				
6475	13.0	27.0	0.99622	3.54
0.600000				
6476	13.0	20.0	0.99540	3.42
0.670000				
6477	24.0	32.0	0.99402	3.54
0.600000				
6478	9.0	26.0	0.99470	3.36
0.600000				
6479	24.0	32.0	0.99402	3.54
0.600000				
6480	13.0	27.0	0.99362	3.57

0.500000					
6481	32.0	98.0	0.99578	3.33	
0.620000					
6482	24.0	34.0	0.99484	3.29	
0.800000					
6483	22.0	48.0	0.99494	3.30	
0.840000					
6484	34.0	60.0	0.99492	3.34	
0.850000					
6485	18.0	28.0	0.99483	3.55	
0.660000					
6486	34.0	102.0	0.99414	3.27	
0.780000					
6487	29.0	79.0	0.99770	3.29	
0.540000					
6488	26.0	35.0	0.99314	3.32	
0.820000					
6489	16.0	26.0	0.99402	3.67	
0.560000					
6490	29.0	40.0	0.99574	3.42	
0.750000					
6491	28.0	38.0	0.99651	3.42	
0.820000					
6492	32.0	44.0	0.99490	3.45	
0.580000					
6493	39.0	51.0	0.99512	3.52	
0.531215					
6494	29.0	40.0	0.99574	3.42	
0.750000					
6495	32.0	44.0	0.99547	3.57	
0.710000					
6496	18.0	42.0	0.99549	3.39	
0.660000					

	alcohol	quality	white
1	9.5	1	1
2	10.1	1	1
3	9.9	1	1
4	9.9	1	1
5	10.1	1	1
6	9.6	1	1
8	9.5	1	1
9	11.0	1	1
10	12.0	1	1
11	9.7	1	1
12	10.8	1	1
13	12.4	1	1
14	9.7	1	1
15	11.4	1	1

16	9.6	1	1
17	12.8	2	1
18	11.3	1	1
19	9.5	1	1
20	12.8	2	1
21	11.0	1	1
22	10.5	2	1
23	9.3	1	1
24	10.0	1	1
25	10.4	1	1
26	10.0	1	1
27	10.5	1	1
28	11.6	1	1
29	12.3	1	1
30	10.0	1	1
31	10.2	1	1
...
6467	11.5	1	0
6468	12.4	1	0
6469	11.1	1	0
6470	9.5	1	0
6471	12.5	1	0
6472	10.5	1	0
6473	11.8	1	0
6474	10.8	1	0
6475	11.9	1	0
6476	11.3	1	0
6477	11.3	1	0
6478	11.9	1	0
6479	11.3	1	0
6480	11.9	1	0
6481	9.8	1	0
6482	11.6	1	0
6483	11.5	1	0
6484	11.4	1	0
6485	10.9	1	0
6486	12.8	1	0
6487	9.2	1	0
6488	11.6	1	0
6489	11.6	1	0
6490	11.0	1	0
6491	9.5	1	0
6492	10.5	1	0
6493	11.2	1	0
6494	11.0	1	0
6495	10.2	1	0
6496	11.0	1	0

[6430 rows x 13 columns]

```

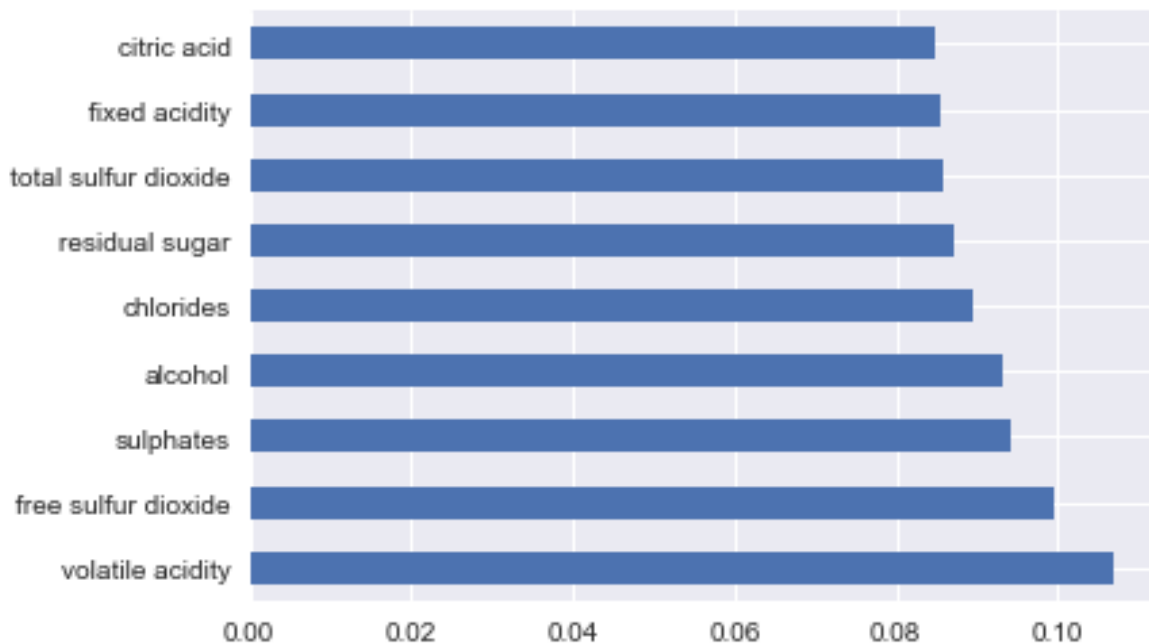
x = data_copy3.drop('quality',axis=1)
y = data_copy3['quality']

model = ExtraTreesClassifier()
model.fit(x, y)
print(model.feature_importances_)

[ 0.08564838  0.10677143  0.08485914  0.0872269   0.08955658
 0.09949098
 0.08592203  0.08216449  0.08405464  0.09414621  0.09330975
 0.00684946]

imp_features = pd.Series(model.feature_importances_, index =
x.columns)
imp_features.nlargest(9).plot(kind = 'barh')
plt.show()

```



```

'''
Using SVC - Support Vector Classifier, Decision Tree Classifier,
Random Forest Classifier,
Gaussian Naive Bayes, and Logistic Regression.
'''
model_params = {
    'svm' : {
        'model': SVC(gamma='auto'),
        'params': {
            'C' : [1,10,20],
            'kernel': ['rbf']
        }
    }
}

```

```

    },
    'decision_tree':{
        'model': DecisionTreeClassifier(),
        'params':{
            'criterion':['entropy','gini'],
            'max_depth':[5,8,9]
        }
    },
    'random_forest':{
        'model': RandomForestClassifier(),
        'params':{
            'n_estimators':[1,5,10],
            'max_depth':[5,8,9]
        }
    },
    'naive_bayes':{
        'model': GaussianNB(),
        'params':{}}
    },
    'logistic_regression' : {
        'model' : LogisticRegression(solver='newton-cg',multi_class =
'multinomial'),
        'params' : {
            "C" : [1,5,10]
        }
    }
}

score = []
for model_name,mp in model_params.items():
    clf = GridSearchCV(mp['model'],mp['params'], cv=8,
return_train_score=False, refit=True)
    clf.fit(x,y)
    score.append({
        'Model' : model_name,
        'Best_Score': clf.best_score_,
        'Best_Params': clf.best_params_
    })

print('The scikit-learn version is {}'.format(sklearn.__version__))

The scikit-learn version is 0.18.2.

data_copy4 = pd.DataFrame(score, columns = ['Model', 'Best_Score',
'Best_Params'])

```

data_copy4

	Model	Best_Score	
Best_Params			
0	svm	0.932193	{'C': 1, 'kernel': 'rbf'}
1	decision_tree	0.921773	{'criterion': 'gini', 'max_depth': 5}
2	random_forest	0.932504	{'max_depth': 5, 'n_estimators': 10}
3	naive_bayes	0.726905	{}
4	logistic_regression	0.932815	{'C': 10}
...			

Using SVM - Support Vector Machine

```
clf_svm = SVC(kernel = 'rbf', C = 1)
scores = cross_val_score(clf_svm, x, y, cv = 8, scoring = 'accuracy')
scores
```

```
array([ 0.93167702,  0.93043478,  0.93159204,  0.93159204,
        0.9340796 ,
        0.93275218,  0.93275218,  0.93266833])
```

```
scores.mean()
```

```
0.93219352133329703
```

```
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size =
0.2, random_state = 0)
print(x_train)
print('-----x_train complete-----')
print(x_test)
print('-----x_test complete-----')
print(y_train)
print('-----y_train complete-----')
print(y_test)
print('-----y_test complete-----')
```

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
2660	6.7	0.330	0.34	6.60
0.067				
3354	5.9	0.200	0.28	12.80
0.038				
1041	7.7	0.275	0.30	1.00
0.039				
3490	7.6	0.310	0.24	1.80
0.037				

2746 0.045	7.0	0.340	0.30	1.80
193 0.055	6.6	0.150	0.34	5.10
5920 0.062	7.0	0.510	0.09	2.10
6339 0.094	7.4	0.785	0.19	5.20
1391 0.029	6.2	0.160	0.47	1.40
5113 0.060	7.0	0.490	0.49	5.60
3934 0.040	7.3	0.270	0.30	1.30
3686 0.087	7.6	0.300	0.37	1.60
2025 0.160	7.5	0.300	0.71	1.30
6363 0.063	6.8	0.590	0.10	1.70
2030 0.034	9.5	0.420	0.41	2.30
2844 0.049	8.0	0.220	0.31	5.60
6018 0.076	7.9	0.540	0.34	2.50
3895 0.021	7.0	0.250	0.33	2.10
3017 0.035	5.9	0.340	0.30	3.80
4787 0.030	3.9	0.225	0.40	4.20
2036 0.046	5.7	0.270	0.32	1.20
3320 0.050	7.5	0.290	0.36	15.70
6055 0.042	5.1	0.510	0.18	2.10
3647 0.037	6.4	0.290	0.24	3.20
3576 0.048	7.4	0.280	0.36	14.60
142 0.039	7.9	0.210	0.40	1.20
1434 0.042	7.4	0.250	0.49	1.10
5373 0.087	9.6	0.680	0.24	2.20
4493	6.8	0.310	0.25	10.50

0.043				
2795	6.4	0.220	0.32	7.20
0.028				
...
...				
2027	6.9	0.320	0.15	8.10
0.046				
101	7.1	0.120	0.32	9.60
0.054				
2520	7.5	0.410	0.23	14.80
0.054				
1889	7.6	0.150	0.40	1.30
0.036				
2065	7.8	0.270	0.28	1.80
0.050				
4917	7.9	0.320	0.51	1.80
0.341				
5138	8.9	0.635	0.37	1.70
0.263				
2183	6.3	0.300	0.24	6.60
0.040				
6103	7.2	0.360	0.46	2.10
0.074				
6283	8.0	0.810	0.25	3.40
0.076				
2930	5.6	0.210	0.40	1.30
0.041				
546	6.3	0.260	0.24	7.20
0.039				
1718	9.3	0.200	0.33	1.70
0.050				
2934	5.3	0.210	0.29	0.70
0.028				
2242	6.1	0.230	0.45	10.60
0.094				
2155	7.8	0.290	0.33	8.75
0.035				
2626	8.0	0.240	0.26	1.70
0.033				
715	6.2	0.210	0.27	1.70
0.038				
3517	6.6	0.290	0.29	1.80
0.036				
5991	9.2	0.310	0.36	2.20
0.079				
5940	8.9	0.500	0.21	2.20
0.088				
4436	6.7	0.280	0.28	4.50
0.051				

1043 0.048	7.5	0.330	0.48	19.45
5893 0.079	7.7	0.600	0.06	2.00
4925 0.106	7.9	0.430	0.21	1.60
4997 0.080	8.1	0.545	0.18	1.90
3306 0.033	6.0	0.320	0.30	1.90
1669 0.036	7.1	0.310	0.38	1.20
2634 0.044	8.7	0.310	0.73	14.35
2764 0.050	7.9	0.160	0.30	7.40

	free sulfur dioxide	total sulfur dioxide	density	pH
sulphates \				
2660 0.48	35.0	156.0	0.99542	3.11
3354 0.57	29.0	132.0	0.99426	3.31
1041 0.56	19.0	75.0	0.99200	3.01
3490 0.44	39.0	150.0	0.99130	3.05
2746 0.45	44.0	142.0	0.99140	2.99
193 0.42	34.0	125.0	0.99420	3.36
5920 0.54	4.0	9.0	0.99584	3.35
6339 0.52	19.0	98.0	0.99713	3.16
1391 0.42	23.0	81.0	0.99000	3.26
5113 0.76	26.0	121.0	0.99740	3.34
3934 0.53	26.0	84.0	0.99222	3.28
3686 0.50	27.0	177.0	0.99438	3.09
2025 0.42	44.0	149.0	0.99480	3.08
6363 0.67	34.0	53.0	0.99580	3.41
2030 0.52	22.0	145.0	0.99510	3.06

2844	24.0	97.0	0.99300	3.10
0.42				
6018	8.0	17.0	0.99235	3.20
0.72				
3895	17.0	76.0	0.99021	3.26
0.45				
3017	57.0	135.0	0.99016	3.09
0.34				
4787	29.0	118.0	0.98900	3.57
0.36				
2036	20.0	155.0	0.99340	3.80
0.41				
3320	29.0	124.0	0.99680	3.06
0.54				
6055	16.0	101.0	0.99240	3.46
0.87				
3647	31.0	95.0	0.98942	2.90
0.66				
3576	35.0	161.0	0.99680	3.14
0.56				
142	38.0	107.0	0.99200	3.21
0.54				
1434	35.0	156.0	0.99170	3.13
0.55				
5373	5.0	28.0	0.99880	3.14
0.60				
4493	30.0	165.0	0.99720	3.36
0.55				
2795	15.0	83.0	0.99300	3.13
0.55				
...
...				
2027	51.0	180.0	0.99580	3.13
0.45				
101	64.0	162.0	0.99620	3.40
0.41				
2520	28.0	174.0	0.99898	3.18
0.49				
1889	24.0	112.0	0.99320	3.14
0.76				
2065	21.0	127.0	0.99340	3.15
0.44				
4917	17.0	56.0	0.99690	3.04
1.08				
5138	5.0	62.0	0.99710	3.00
1.09				
2183	38.0	141.0	0.99500	3.22
0.47				
6103	24.0	44.0	0.99534	3.40

0.85					
6283	34.0	85.0	0.99668	3.19	
0.42					
2930	81.0	147.0	0.99010	3.22	
0.95					
546	38.0	172.0	0.99580	3.49	
0.64					
1718	28.0	178.0	0.99540	3.16	
0.43					
2934	11.0	66.0	0.99215	3.30	
0.40					
2242	49.0	169.0	0.99699	3.05	
0.54					
2155	33.0	181.0	0.99620	3.11	
0.46					
2626	36.0	136.0	0.99316	3.44	
0.51					
715	41.0	150.0	0.99330	3.49	
0.71					
3517	38.0	102.0	0.98819	3.08	
0.42					
5991	11.0	31.0	0.99615	3.33	
0.86					
5940	21.0	39.0	0.99692	3.33	
0.83					
4436	14.0	92.0	0.99224	3.36	
0.58					
1043	55.0	243.0	1.00100	2.95	
0.40					
5893	19.0	41.0	0.99697	3.39	
0.62					
4925	10.0	37.0	0.99660	3.17	
0.91					
4997	13.0	35.0	0.99720	3.30	
0.59					
3306	41.0	142.0	0.98912	3.29	
0.42					
1669	10.0	124.0	0.99240	3.14	
0.44					
2634	27.0	191.0	1.00013	2.96	
0.88					
2764	58.0	152.0	0.99612	3.12	
0.37					

	alcohol	white
2660	9.300000	1
3354	11.800000	1
1041	10.700000	1
3490	11.800000	1

2746	10.800000	1
193	9.600000	1
5920	10.500000	0
6339	9.566667	0
1391	12.200000	1
5113	10.500000	0
3934	10.700000	1
3686	9.800000	1
2025	8.900000	1
6363	9.700000	0
2030	11.000000	1
2844	10.900000	1
6018	13.100000	0
3895	12.300000	1
3017	12.000000	1
4787	12.800000	1
2036	10.200000	1
3320	10.400000	1
6055	12.900000	0
3647	12.600000	1
3576	10.600000	1
142	10.800000	1
1434	11.300000	1
5373	10.200000	0
4493	10.550000	1
2795	10.900000	1
...
2027	8.900000	1
101	9.400000	1
2520	9.700000	1
1889	10.000000	1
2065	9.900000	1
4917	9.200000	0
5138	9.300000	0
2183	9.500000	1
6103	11.000000	0
6283	9.200000	0
2930	11.600000	1
546	9.700000	1
1718	9.000000	1
2934	9.800000	1
2242	8.800000	1
2155	10.700000	1
2626	10.400000	1
715	10.500000	1
3517	13.700000	1
5991	12.000000	0
5940	11.100000	0
4436	11.900000	1

1043	8.800000	1
5893	10.100000	0
4925	9.500000	0
4997	9.000000	0
3306	12.800000	1
1669	9.900000	1
2634	8.700000	1
2764	9.500000	1

[5144 rows x 12 columns]

-----x_train complete-----

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
1935	8.8	0.340	0.33	9.70
0.036				
1106	5.2	0.240	0.45	3.80
0.027				
2932	6.4	0.250	0.32	11.30
0.038				
743	5.9	0.260	0.25	12.50
0.034				
2230	7.1	0.200	0.36	11.60
0.042				
1584	6.4	0.250	0.74	7.80
0.045				
4583	6.4	0.230	0.37	7.90
0.050				
567	6.0	0.260	0.50	2.20
0.048				
4333	7.3	0.190	0.27	13.90
0.057				
4414	7.1	0.270	0.27	10.40
0.041				
4966	9.3	0.320	0.57	2.00
0.074				
2992	6.6	0.200	0.27	10.90
0.038				
2730	6.8	0.300	0.22	6.20
0.060				
5699	8.6	0.550	0.09	3.30
0.068				
3860	7.3	0.250	0.26	7.20
0.048				
4397	7.4	0.270	0.26	11.80
0.053				
3539	7.2	0.270	0.31	1.20
0.031				
2261	6.5	0.360	0.31	4.10
0.061				

1873	5.8	0.330	0.20	16.05
0.047				
2671	5.9	0.300	0.30	2.00
0.030				
5898	7.5	0.430	0.30	2.20
0.062				
5922	7.7	0.580	0.01	1.80
0.088				
476	7.4	0.280	0.36	1.10
0.028				
1182	7.4	0.220	0.33	2.00
0.045				
4972	9.7	0.320	0.54	2.50
0.094				
4994	6.8	0.775	0.00	3.00
0.102				
638	5.6	0.350	0.14	5.00
0.046				
3456	6.0	0.390	0.13	1.20
0.042				
4182	7.5	0.380	0.56	9.70
0.055				
4315	6.7	0.410	0.24	5.40
0.035				
...
...				
4705	5.7	0.270	0.16	9.00
0.053				
3695	6.4	0.150	0.40	1.30
0.053				
6288	6.0	0.490	0.00	2.30
0.068				
4615	6.1	0.440	0.28	4.25
0.032				
1258	6.9	0.290	0.16	6.80
0.034				
3912	6.5	0.330	0.30	3.80
0.036				
406	5.8	0.280	0.34	4.00
0.031				
2237	6.4	0.340	0.20	14.90
0.060				
5637	9.0	0.690	0.00	2.40
0.088				
1234	7.7	0.260	0.31	1.30
0.043				
972	6.8	0.220	0.35	17.50
0.039				
4719	6.7	0.150	0.32	7.90

0.034				
4080	5.7	0.250	0.27	10.80
0.050				
5876	7.0	0.400	0.32	3.60
0.061				
3303	6.3	0.220	0.34	5.00
0.032				
178	6.0	0.670	0.07	1.20
0.060				
1216	8.2	0.370	0.36	1.00
0.034				
3965	7.3	0.380	0.23	6.50
0.050				
4760	6.6	0.170	0.26	7.40
0.052				
2162	9.9	0.490	0.23	2.40
0.087				
3859	7.3	0.250	0.26	7.20
0.048				
877	6.0	0.280	0.34	1.60
0.119				
4993	4.7	0.600	0.17	2.30
0.058				
5865	8.5	0.660	0.20	2.10
0.097				
1108	6.1	0.290	0.27	1.70
0.024				
2579	6.4	0.280	0.56	1.70
0.156				
5825	8.4	0.670	0.19	2.20
0.093				
5782	8.8	0.610	0.19	4.00
0.094				
4435	6.5	0.290	0.30	9.15
0.051				
5416	10.9	0.210	0.49	2.80
0.088				

	free sulfur dioxide	total sulfur dioxide	density	pH
sulphates \				
1935	46.0	172.0	0.99660	3.080000
0.40				
1106	21.0	128.0	0.99200	3.550000
0.49				
2932	69.0	192.0	0.99573	3.140000
0.50				
743	38.0	152.0	0.99770	3.330000
0.43				
2230	45.0	124.0	0.99700	2.920000

0.59				
1584	52.0	209.0	0.99560	3.210000
0.42				
4583	60.0	150.0	0.99488	2.860000
0.49				
567	59.0	153.0	0.99280	3.080000
0.61				
4333	45.0	155.0	0.99807	2.940000
0.41				
4414	26.0	114.0	0.99335	3.040000
0.52				
4966	27.0	65.0	0.99690	3.280000
0.79				
2992	29.0	130.0	0.99496	3.110000
0.44				
2730	41.0	190.0	0.99858	3.180000
0.51				
5699	8.0	17.0	0.99735	3.230000
0.44				
3860	52.0	207.0	0.99587	3.120000
0.37				
4397	55.0	173.0	0.99699	3.110000
0.60				
3539	27.0	80.0	0.98892	3.030000
0.33				
2261	20.0	134.0	0.99475	3.180000
0.45				
1873	26.0	166.0	0.99760	3.090000
0.46				
2671	38.0	142.0	0.98892	3.410000
0.41				
5898	6.0	12.0	0.99495	3.440000
0.72				
5922	12.0	18.0	0.99568	3.320000
0.56				
476	42.0	105.0	0.98930	2.990000
0.39				
1182	31.0	101.0	0.99310	3.420000
0.55				
4972	28.0	83.0	0.99840	3.280000
0.82				
4994	8.0	23.0	0.99650	3.450000
0.56				
638	48.0	198.0	0.99370	3.300000
0.71				
3456	60.0	172.0	0.99114	3.060000
0.52				
4182	15.0	170.0	0.99605	3.130000
0.65				
4315	33.0	115.0	0.99010	3.120000

0.44				
...
...				
4705	32.0	111.0	0.99474	3.360000
0.37				
3695	61.0	146.0	0.99112	3.170000
0.68				
6288	15.0	33.0	0.99292	3.580000
0.59				
4615	43.0	132.0	0.99160	3.260000
0.47				
1258	65.0	212.0	0.99550	3.080000
0.39				
3912	34.0	88.0	0.99028	3.250000
0.63				
406	40.0	99.0	0.98960	3.390000
0.39				
2237	37.0	162.0	0.99830	3.130000
0.45				
5637	19.0	38.0	0.99900	3.350000
0.60				
1234	47.0	155.0	0.99370	3.420000
0.50				
972	38.0	153.0	0.99940	3.218395
0.42				
4719	17.0	81.0	0.99512	3.290000
0.31				
4080	58.0	116.0	0.99592	3.100000
0.50				
5876	9.0	29.0	0.99416	3.280000
0.49				
3303	36.0	93.0	0.99012	3.270000
0.36				
178	9.0	108.0	0.99310	3.110000
0.35				
1216	17.0	93.0	0.99060	3.040000
0.32				
3965	18.0	102.0	0.99304	3.100000
0.55				
4760	45.0	128.0	0.99388	3.160000
0.37				
2162	19.0	115.0	0.99480	2.770000
0.44				
3859	52.0	207.0	0.99587	3.120000
0.37				
877	33.0	104.0	0.99210	3.190000
0.38				
4993	17.0	106.0	0.99320	3.850000
0.60				

5865	23.0	113.0	0.99733	3.130000
0.48				
1108	13.0	76.0	0.98930	3.210000
0.51				
2579	49.0	106.0	0.99354	3.100000
0.37				
5825	11.0	75.0	0.99736	3.200000
0.59				
5782	30.0	69.0	0.99787	3.220000
0.50				
4435	25.0	166.0	0.99339	3.240000
0.56				
5416	11.0	32.0	0.99720	3.220000
0.68				

	alcohol	white
1935	10.200000	1
1106	11.200000	1
2932	10.200000	1
743	9.400000	1
2230	9.500000	1
1584	9.200000	1
4583	9.300000	1
567	9.800000	1
4333	8.800000	1
4414	11.500000	1
4966	10.700000	0
2992	10.500000	1
2730	9.200000	1
5699	10.000000	0
3860	9.200000	1
4397	9.800000	1
3539	12.700000	1
2261	9.000000	1
1873	8.900000	1
2671	12.900000	1
5898	11.500000	0
5922	10.500000	0
476	12.400000	1
1182	11.400000	1
4972	9.600000	0
4994	10.700000	0
638	10.300000	1
3456	10.600000	1
4182	9.900000	1
4315	12.893333	1
...
4705	10.400000	1
3695	11.000000	1

6288	12.500000	0
4615	11.266667	1
1258	9.000000	1
3912	12.500000	1
406	12.800000	1
2237	9.000000	1
5637	9.300000	0
1234	10.100000	1
972	9.000000	1
4719	10.000000	1
4080	9.800000	1
5876	11.300000	0
3303	13.500000	1
178	8.700000	1
1216	11.700000	1
3965	11.200000	1
4760	10.000000	1
2162	9.400000	1
3859	9.200000	1
877	10.200000	1
4993	12.900000	0
5865	9.200000	0
1108	12.600000	1
2579	9.200000	1
5825	9.200000	0
5782	10.000000	0
4435	11.333333	1
5416	11.700000	0

[1286 rows x 12 columns]

-----x_test complete-----

2660	1
3354	1
1041	1
3490	1
2746	1
193	1
5920	1
6339	1
1391	1
5113	1
3934	1
3686	1
2025	1
6363	1
2030	1
2844	1
6018	2
3895	1

```
3017    1
4787    2
2036    1
3320    1
6055    1
3647    1
3576    1
142     1
1434    1
5373    1
4493    1
2795    2
      ..
2027    1
101     1
2520    1
1889    1
2065    1
4917    1
5138    1
2183    1
6103    1
6283    1
2930    2
546     1
1718    0
2934    1
2242    1
2155    1
2626    1
715     1
3517    1
5991    1
5940    1
4436    1
1043    1
5893    1
4925    1
4997    1
3306    1
1669    1
2634    1
2764    1
Name: quality, Length: 5144, dtype: int64
-----y_train complete-----
1935    1
1106    2
2932    1
743     1
```

2230	1
1584	1
4583	1
567	1
4333	2
4414	1
4966	1
2992	1
2730	1
5699	1
3860	1
4397	1
3539	1
2261	1
1873	1
2671	1
5898	1
5922	1
476	1
1182	1
4972	1
4994	1
638	1
3456	1
4182	1
4315	1
	..
4705	1
3695	1
6288	1
4615	1
1258	1
3912	1
406	1
2237	0
5637	1
1234	1
972	1
4719	1
4080	1
5876	1
3303	1
178	0
1216	2
3965	0
4760	1
2162	1
3859	1
877	1

```

4993    1
5865    1
1108    1
2579    1
5825    0
5782    1
4435    1
5416    1
Name: quality, Length: 1286, dtype: int64
-----y_test complete-----

```

```

clf_svm1 = SVC(kernel='rbf', C = 1)
clf_svm1.fit(x_train, y_train)

```

```

SVC(C=1, cache_size=200, class_weight=None, coef0=0.0,
    decision_function_shape=None, degree=3, gamma='auto', kernel='rbf',
    max_iter=-1, probability=False, random_state=None, shrinking=True,
    tol=0.001, verbose=False)

```

```

y_pred = clf_svm1.predict(x_test)

```

```

accuracy = accuracy_score(y_test, y_pred)
accuracy

```

```

0.93779160186625199

```

```

'''

```

Testing on the split test-data

```

'''

```

```

accuracy_dataframe = pd.DataFrame({"y_test": y_test, "y_pred":
y_pred})
accuracy_dataframe.head(n = 100)

```

	y_pred	y_test
1935	1	1
1106	1	2
2932	1	1
743	1	1
2230	1	1
1584	1	1
4583	1	1
567	1	1
4333	2	2
4414	1	1
4966	1	1
2992	1	1
2730	1	1
5699	1	1
3860	1	1
4397	1	1
3539	1	1

2261	1	1
1873	1	1
2671	1	1
5898	1	1
5922	1	1
476	1	1
1182	1	1
4972	1	1
4994	1	1
638	1	1
3456	1	1
4182	1	1
4315	1	1
...
150	1	1
650	1	1
2863	1	1
2841	1	1
4272	1	1
5485	1	1
2669	1	1
4317	1	1
4026	1	1
5033	1	1
6474	1	1
316	1	1
210	1	1
1713	1	1
4568	1	1
5391	1	1
206	1	1
557	1	1
4053	1	1
890	1	1
5439	1	1
5386	1	1
6317	1	1
1105	1	1
1181	1	1
3039	1	1
4461	1	1
542	1	1
4517	1	1
5953	1	1

[100 rows x 2 columns]