

Name:- Atharva Ingale
Registration No.:- 21BCE10275
Assignment 4

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score

# Task 1: Load the dataset
print("Task 1: Load the dataset\n\n\n\n")
missing_values = ["", "NA", "N/A", "NaN"]
winequality = pd.read_csv('A:\\\\VIT Bhopal\\\\AI and ML with smartbridge
google\\\\winequality-red.csv', na_values=missing_values)
print(winequality)

# Find and handle missing values
missing_values = winequality.isnull().sum()
print("\n\nMissing Values:")
print(missing_values)

# Replace missing values with column means
winequality.fillna(winequality.mean(), inplace=True)
print("\nMissing Values After Handling:")
print(winequality.isnull().sum(), "\n\n")

# Find and handle outliers
def find_and_handle_outliers(column):
    Q1 = winequality[column].quantile(0.25)
    Q3 = winequality[column].quantile(0.75)
    IQR = Q3 - Q1
    lower_bound = Q1 - 1.5 * IQR
    upper_bound = Q3 + 1.5 * IQR

    outliers = (winequality[column] < lower_bound) |
(winequality[column] > upper_bound)

    if outliers.sum() > 0:
        print(f"Outliers in {column}:")
        print(winequality[outliers])
        print("\n")

    winequality.loc[outliers, column] = winequality[column].mean()

# Apply outlier handling to numeric columns
numeric_columns =
```

```
winequality.select_dtypes(include=[np.number]).columns
for column in numeric_columns:
    find_and_handle_outliers(column)

# Visualize outliers
plt.figure(figsize=(12, 8))
sns.boxplot(data=winequality[numeric_columns])
plt.title("Boxplot of Numeric Features (Outliers Handled)")
plt.xticks(rotation=45)
plt.show()
```

Task 1: Load the dataset

	fixed acidity	volatile acidity	citric acid	residual sugar
0	7.4	0.700	0.00	1.9
0.076				
1	7.8	0.880	0.00	2.6
0.098				
2	7.8	0.760	0.04	2.3
0.092				
3	11.2	0.280	0.56	1.9
0.075				
4	7.4	0.700	0.00	1.9
0.076				
...
...				
1594	6.2	0.600	0.08	2.0
0.090				
1595	5.9	0.550	0.10	2.2
0.062				
1596	6.3	0.510	0.13	2.3
0.076				
1597	5.9	0.645	0.12	2.0
0.075				
1598	6.0	0.310	0.47	3.6
0.067				

	free sulfur dioxide	total sulfur dioxide	density	pH
0	11.0	34.0	0.99780	3.51
0.56				
1	25.0	67.0	0.99680	3.20
0.68				
2	15.0	54.0	0.99700	3.26
0.65				
3	17.0	60.0	0.99800	3.16

0.58					
4	11.0	34.0	0.99780	3.51	
0.56					
...	
...					
1594	32.0	44.0	0.99490	3.45	
0.58					
1595	39.0	51.0	0.99512	3.52	
0.76					
1596	29.0	40.0	0.99574	3.42	
0.75					
1597	32.0	44.0	0.99547	3.57	
0.71					
1598	18.0	42.0	0.99549	3.39	
0.66					

	alcohol	quality
0	9.4	5
1	9.8	5
2	9.8	5
3	9.8	6
4	9.4	5
...
1594	10.5	5
1595	11.2	6
1596	11.0	6
1597	10.2	5
1598	11.0	6

[1599 rows x 12 columns]

Missing Values:

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

Missing Values After Handling:

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

```

Outliers in fixed acidity:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
205	12.8	0.300	0.74	2.6
0.095				
206	12.8	0.300	0.74	2.6
0.095				
243	15.0	0.210	0.44	2.2
0.075				
244	15.0	0.210	0.44	2.2
0.075				
264	12.5	0.560	0.49	2.4
0.064				
294	13.3	0.340	0.52	3.2
0.094				
328	13.4	0.270	0.62	2.6
0.082				
338	12.4	0.490	0.58	3.0
0.103				
339	12.5	0.280	0.54	2.3
0.082				
347	13.8	0.490	0.67	3.0
0.093				
353	13.5	0.530	0.79	4.8
0.120				
359	12.6	0.380	0.66	2.6
0.088				
363	12.5	0.460	0.63	2.0
0.071				
364	12.8	0.615	0.66	5.8
0.083				
366	12.8	0.615	0.66	5.8
0.083				
374	14.0	0.410	0.63	3.8
0.089				
381	13.7	0.415	0.68	2.9

0.085				
391	13.7	0.415	0.68	2.9
0.085				
394	12.7	0.600	0.65	2.3
0.063				
409	12.5	0.460	0.49	4.5
0.070				
429	12.8	0.840	0.63	2.4
0.088				
440	12.6	0.310	0.72	2.2
0.072				
442	15.6	0.685	0.76	3.7
0.100				
446	12.5	0.380	0.60	2.6
0.081				
470	13.0	0.320	0.65	2.6
0.093				
472	12.5	0.370	0.55	2.6
0.083				
509	13.3	0.290	0.75	2.8
0.084				
510	12.4	0.420	0.49	4.6
0.073				
516	12.5	0.600	0.49	4.3
0.100				
538	12.9	0.350	0.49	5.8
0.066				
544	14.3	0.310	0.74	1.8
0.075				
548	12.4	0.350	0.49	2.6
0.079				
554	15.5	0.645	0.49	4.2
0.095				
555	15.5	0.645	0.49	4.2
0.095				
557	15.6	0.645	0.49	4.2
0.095				
559	13.0	0.470	0.49	4.3
0.085				
560	12.7	0.600	0.49	2.8
0.075				
564	13.0	0.470	0.49	4.3
0.085				
565	12.7	0.600	0.49	2.8
0.075				
596	12.4	0.400	0.51	2.0
0.059				
599	12.7	0.590	0.45	2.3
0.082				

601	13.2	0.460	0.52	2.2
0.071				
603	13.2	0.460	0.52	2.2
0.071				
611	13.2	0.380	0.55	2.7
0.081				
652	15.9	0.360	0.65	7.5
0.096				
680	13.3	0.430	0.58	1.9
0.070				
811	12.9	0.500	0.55	2.8
0.072				
814	12.6	0.410	0.54	2.8
0.103				
1224	12.6	0.390	0.49	2.5
0.080				

	free sulfur dioxide	total sulfur dioxide	density	pH
sulphates \				
205	9.0	28.0	0.99940	3.20
0.77				
206	9.0	28.0	0.99940	3.20
0.77				
243	10.0	24.0	1.00005	3.07
0.84				
244	10.0	24.0	1.00005	3.07
0.84				
264	5.0	27.0	0.99990	3.08
0.87				
294	17.0	53.0	1.00140	3.05
0.81				
328	6.0	21.0	1.00020	3.16
0.67				
338	28.0	99.0	1.00080	3.16
1.00				
339	12.0	29.0	0.99970	3.11
1.36				
347	6.0	15.0	0.99860	3.02
0.93				
353	23.0	77.0	1.00180	3.18
0.77				
359	10.0	41.0	1.00100	3.17
0.68				
363	6.0	15.0	0.99880	2.99
0.87				
364	7.0	42.0	1.00220	3.07
0.73				
366	7.0	42.0	1.00220	3.07
0.73				

374	6.0	47.0	1.00140	3.01
0.81				
381	17.0	43.0	1.00140	3.06
0.80				
391	17.0	43.0	1.00140	3.06
0.80				
394	6.0	25.0	0.99970	3.03
0.57				
409	26.0	49.0	0.99810	3.05
0.57				
429	13.0	35.0	0.99970	3.10
0.60				
440	6.0	29.0	0.99870	2.88
0.82				
442	6.0	43.0	1.00320	2.95
0.68				
446	31.0	72.0	0.99960	3.10
0.73				
470	15.0	47.0	0.99960	3.05
0.61				
472	25.0	68.0	0.99950	3.15
0.82				
509	23.0	43.0	0.99860	3.04
0.68				
510	19.0	43.0	0.99780	3.02
0.61				
516	5.0	14.0	1.00100	3.25
0.74				
538	5.0	35.0	1.00140	3.20
0.66				
544	6.0	15.0	1.00080	2.86
0.79				
548	27.0	69.0	0.99940	3.12
0.75				
554	10.0	23.0	1.00315	2.92
0.74				
555	10.0	23.0	1.00315	2.92
0.74				
557	10.0	23.0	1.00315	2.92
0.74				
559	6.0	47.0	1.00210	3.30
0.68				
560	5.0	19.0	0.99940	3.14
0.57				
564	6.0	47.0	1.00210	3.30
0.68				
565	5.0	19.0	0.99940	3.14
0.57				
596	6.0	24.0	0.99940	3.04

0.60					
599	11.0	22.0	1.00000	3.00	
0.70					
601	12.0	35.0	1.00060	3.10	
0.56					
603	12.0	35.0	1.00060	3.10	
0.56					
611	5.0	16.0	1.00060	2.98	
0.54					
652	22.0	71.0	0.99760	2.98	
0.84					
680	15.0	40.0	1.00040	3.06	
0.49					
811	7.0	24.0	1.00012	3.09	
0.68					
814	19.0	41.0	0.99939	3.21	
0.76					
1224	8.0	20.0	0.99920	3.07	
0.82					

	alcohol	quality
205	10.8	7
206	10.8	7
243	9.2	7
244	9.2	7
264	10.9	5
294	9.5	6
328	9.7	6
338	11.5	6
339	9.8	7
347	12.0	6
353	13.0	5
359	9.8	6
363	10.2	5
364	10.0	7
366	10.0	7
374	10.8	6
381	10.0	6
391	10.0	6
394	9.9	5
409	9.6	4
429	10.4	6
440	9.8	8
442	11.2	7
446	10.5	5
470	10.6	5
472	10.4	6
509	11.4	7
510	9.5	5

516	11.9	6
538	12.0	7
544	8.4	6
548	10.4	6
554	11.1	5
555	11.1	5
557	11.1	5
559	12.7	6
560	11.4	5
564	12.7	6
565	11.4	5
596	9.3	6
599	9.3	6
601	9.0	6
603	9.0	6
611	9.4	5
652	14.9	5
680	9.0	5
811	10.9	6
814	11.3	6
1224	10.3	6

Outliers in volatile acidity:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
38	5.7	1.130	0.09	1.50
0.172				
94	5.0	1.020	0.04	1.40
0.045				
120	7.3	1.070	0.09	1.70
0.178				
126	8.2	1.330	0.00	1.70
0.081				
127	8.1	1.330	0.00	1.80
0.082				
134	7.9	1.040	0.05	2.20
0.084				
199	6.9	1.090	0.06	2.10
0.061				
553	5.0	1.040	0.24	1.60
0.050				
672	9.8	1.240	0.34	2.00
0.079				
690	7.4	1.185	0.00	4.25
0.097				
700	10.6	1.020	0.43	2.90
0.076				
705	8.4	1.035	0.15	6.00

0.073				
710	10.6	1.025	0.43	2.80
0.080				
724	7.5	1.115	0.10	3.10
0.086				
899	8.3	1.020	0.02	3.40
0.084				
1261	6.3	1.020	0.00	2.00
0.083				
1299	7.6	1.580	0.00	2.10
0.137				
1312	8.0	1.180	0.21	1.90
0.083				
1467	6.7	1.040	0.08	2.30
0.067				

	free sulfur dioxide	total sulfur dioxide	density	pH
sulphates \				
38	7.0	19.0	0.99400	3.50
0.48				
94	41.0	85.0	0.99380	3.75
0.48				
120	10.0	89.0	0.99620	3.30
0.57				
126	3.0	12.0	0.99640	3.53
0.49				
127	3.0	12.0	0.99640	3.54
0.48				
134	13.0	29.0	0.99590	3.22
0.55				
199	12.0	31.0	0.99480	3.51
0.43				
553	32.0	96.0	0.99340	3.74
0.62				
672	32.0	151.0	0.99800	3.15
0.53				
690	5.0	14.0	0.99660	3.63
0.54				
700	26.0	88.0	0.99840	3.08
0.57				
705	11.0	54.0	0.99900	3.37
0.49				
710	21.0	84.0	0.99850	3.06
0.57				
724	5.0	12.0	0.99580	3.54
0.60				
899	6.0	11.0	0.99892	3.48
0.49				
1261	17.0	24.0	0.99437	3.59

0.55				
1299	5.0	9.0	0.99476	3.50
0.40				
1312	14.0	41.0	0.99532	3.34
0.47				
1467	19.0	32.0	0.99648	3.52
0.57				

	alcohol	quality
38	9.8	4
94	10.5	4
120	9.0	5
126	10.9	5
127	10.9	5
134	9.9	6
199	11.4	4
553	11.5	5
672	9.5	5
690	10.7	3
700	10.1	6
705	9.9	5
710	10.1	5
724	11.2	4
899	11.0	3
1261	11.2	4
1299	10.9	3
1312	10.5	5
1467	11.0	4

Outliers in citric acid:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
151	9.2	0.52	1.0	3.4
0.61				

	free sulfur dioxide	total sulfur dioxide	density	pH
sulphates \				
151	32.0	69.0	0.9996	2.74
2.0				

	alcohol	quality
151	9.4	4

Outliers in residual sugar:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
9	7.5	0.500	0.36	6.1
0.071				

11	7.5	0.500	0.36	6.1
0.071				
14	8.9	0.620	0.18	3.8
0.176				
15	8.9	0.620	0.19	3.9
0.170				
18	7.4	0.590	0.08	4.4
0.086				
...
...				
1552	6.3	0.680	0.01	3.7
0.103				
1558	6.9	0.630	0.33	6.7
0.235				
1574	5.6	0.310	0.78	13.9
0.074				
1577	6.2	0.700	0.15	5.1
0.076				
1589	6.6	0.725	0.20	7.8
0.073				
free sulfur dioxide total sulfur dioxide density pH				
sulphates \				
9	17.0	102.0	0.99780	3.35
0.80				
11	17.0	102.0	0.99780	3.35
0.80				
14	52.0	145.0	0.99860	3.16
0.88				
15	51.0	148.0	0.99860	3.17
0.93				
18	6.0	29.0	0.99740	3.38
0.50				
...
...				
1552	32.0	54.0	0.99586	3.51
0.66				
1558	66.0	115.0	0.99787	3.22
0.56				
1574	23.0	92.0	0.99677	3.39
0.48				
1577	13.0	27.0	0.99622	3.54
0.60				
1589	29.0	79.0	0.99770	3.29
0.54				
alcohol quality				
9	10.5	5		
11	10.5	5		
14	9.2	5		

15	9.2	5
18	9.0	4
...
1552	11.3	6
1558	9.5	5
1574	10.5	6
1577	11.9	6
1589	9.2	5

[155 rows x 12 columns]

Outliers in chlorides:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
14	8.9	0.620000	0.18	2.538806
0.176				
15	8.9	0.620000	0.19	2.538806
0.170				
17	8.1	0.560000	0.28	1.700000
0.368				
19	7.9	0.320000	0.51	1.800000
0.341				
38	5.7	0.527821	0.09	1.500000
0.172				
...
...				
1476	9.9	0.500000	0.50	2.538806
0.205				
1490	7.1	0.220000	0.49	1.800000
0.039				
1558	6.9	0.630000	0.33	2.538806
0.235				
1570	6.4	0.360000	0.53	2.200000
0.230				
1571	6.4	0.380000	0.14	2.200000
0.038				

	free sulfur dioxide	total sulfur dioxide	density	pH
sulphates \				
14	52.0	145.0	0.99860	3.16
0.88				
15	51.0	148.0	0.99860	3.17
0.93				
17	16.0	56.0	0.99680	3.11
1.28				
19	17.0	56.0	0.99690	3.04
1.08				
38	7.0	19.0	0.99400	3.50
0.48				

...
1476	48.0	82.0	1.00242	3.16
0.75				
1490	8.0	18.0	0.99344	3.39
0.56				
1558	66.0	115.0	0.99787	3.22
0.56				
1570	19.0	35.0	0.99340	3.37
0.93				
1571	15.0	25.0	0.99514	3.44
0.65				

	alcohol	quality
14	9.2	5
15	9.2	5
17	9.3	5
19	9.2	6
38	9.8	4
...
1476	8.8	5
1490	12.4	6
1558	9.5	5
1570	12.4	6
1571	11.1	6

[112 rows x 12 columns]

Outliers in free sulfur dioxide:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
14	8.9	0.620	0.18	2.538806
0.087467				
15	8.9	0.620	0.19	2.538806
0.087467				
57	7.5	0.630	0.12	2.538806
0.111000				
396	6.6	0.735	0.02	2.538806
0.087467				
400	6.6	0.735	0.02	2.538806
0.087467				
497	7.2	0.340	0.32	2.500000
0.090000				
522	8.2	0.390	0.49	2.300000
0.099000				
584	11.8	0.330	0.49	3.400000
0.093000				
634	7.9	0.350	0.21	1.900000
0.073000				

free sulfur dioxide sulphates \	total sulfur dioxide	density	pH
14 0.88	52.0	145.0	0.99860 3.16
15 0.93	51.0	148.0	0.99860 3.17

57	50.0	110.0	0.99830	3.26
0.77				
396	68.0	124.0	0.99940	3.47
0.53				
400	68.0	124.0	0.99940	3.47
0.53				
497	43.0	113.0	0.99660	3.32
0.79				
522	47.0	133.0	0.99790	3.38
0.99				
584	54.0	80.0	1.00020	3.30
0.76				
634	46.0	102.0	0.99640	3.27
0.58				
678	45.0	87.0	0.99830	3.48
0.53				
925	53.0	77.0	0.99604	3.47
0.87				
926	52.0	73.0	0.99786	3.47
0.90				
982	51.0	70.0	0.99418	3.34
0.82				
1075	45.0	67.0	0.99769	3.44
0.86				
1131	57.0	135.0	0.99341	3.32
0.44				
1154	50.0	63.0	0.99544	3.59
0.68				
1156	45.0	88.0	0.99524	3.33
0.76				
1175	48.0	59.0	0.99541	3.61
0.70				
1217	43.0	74.0	0.99408	3.23
0.81				
1231	48.0	90.0	0.99621	3.38
0.62				
1244	72.0	160.0	0.99721	3.33
0.54				
1256	43.0	60.0	0.99499	3.10
0.42				
1295	51.0	77.5	0.99558	3.20
0.45				
1296	51.0	77.5	0.99558	3.20
0.45				
1358	52.0	98.0	0.99736	3.28
0.50				
1434	55.0	95.0	1.00369	3.18
0.77				
1435	55.0	95.0	1.00369	3.18

0.77					
1474	48.0	82.0	1.00242	3.16	
0.75					
1476	48.0	82.0	1.00242	3.16	
0.75					
1558	66.0	115.0	0.99787	3.22	
0.56					

	alcohol	quality
14	9.2	5
15	9.2	5
57	9.4	5
396	9.9	5
400	9.9	5
497	11.1	5
522	9.8	5
584	10.7	7
634	9.5	5
678	10.0	5
925	11.0	7
926	10.2	6
982	12.9	6
1075	10.2	7
1131	9.5	5
1154	11.4	6
1156	11.8	7
1175	11.5	6
1217	12.0	6
1231	10.8	5
1244	10.3	6
1256	9.2	5
1295	9.5	5
1296	9.5	5
1358	9.5	5
1434	9.0	6
1435	9.0	6
1474	8.8	5
1476	8.8	5
1558	9.5	5

Outliers in total sulfur dioxide:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
14	8.9	0.620000	0.18	2.538806
0.087467				
15	8.9	0.620000	0.19	2.538806
0.087467				
86	8.6	0.490000	0.28	1.900000
0.110000				

88	9.3	0.390000	0.44	2.100000
0.107000				
90	7.9	0.520000	0.26	1.900000
0.079000				
91	8.6	0.490000	0.28	1.900000
0.110000				
92	8.6	0.490000	0.29	2.000000
0.110000				
109	8.1	0.785000	0.52	2.000000
0.087467				
130	8.0	0.745000	0.56	2.000000
0.118000				
145	8.1	0.670000	0.55	1.800000
0.117000				
154	7.1	0.430000	0.42	2.538806
0.070000				
155	7.1	0.430000	0.42	2.538806
0.071000				
156	7.1	0.430000	0.42	2.538806
0.070000				
157	7.1	0.430000	0.42	2.538806
0.071000				
188	7.9	0.500000	0.33	2.000000
0.084000				
189	7.9	0.490000	0.32	1.900000
0.082000				
190	8.2	0.500000	0.35	2.900000
0.077000				
192	6.8	0.630000	0.12	2.538806
0.099000				
201	8.8	0.370000	0.48	2.100000
0.097000				
219	7.8	0.530000	0.33	2.400000
0.080000				
313	8.6	0.470000	0.30	3.000000
0.076000				
354	6.1	0.210000	0.40	1.400000
0.066000				
396	6.6	0.735000	0.02	2.538806
0.087467				
400	6.6	0.735000	0.02	2.538806
0.087467				
415	8.6	0.725000	0.24	2.538806
0.117000				
417	7.0	0.580000	0.12	1.900000
0.091000				
463	8.1	0.660000	0.70	2.200000
0.098000				
515	8.5	0.655000	0.49	2.538806

0.087467				
522	8.2	0.390000	0.49	2.300000
0.099000				
523	9.3	0.400000	0.49	2.500000
0.085000				
591	6.6	0.390000	0.49	1.700000
0.070000				
636	9.6	0.880000	0.28	2.400000
0.086000				
637	9.5	0.885000	0.27	2.300000
0.084000				
649	6.7	0.420000	0.27	2.538806
0.068000				
651	9.8	0.880000	0.25	2.500000
0.104000				
672	9.8	0.527821	0.34	2.000000
0.079000				
684	9.8	0.980000	0.32	2.300000
0.078000				
694	9.0	0.470000	0.31	2.700000
0.084000				
723	7.1	0.310000	0.30	2.200000
0.053000				
741	9.2	0.530000	0.24	2.600000
0.078000				
771	9.4	0.685000	0.26	2.400000
0.082000				
772	9.5	0.570000	0.27	2.300000
0.082000				
791	8.8	0.640000	0.17	2.900000
0.084000				
1079	7.9	0.300000	0.68	2.538806
0.050000				
1081	7.9	0.300000	0.68	2.538806
0.050000				
1131	5.9	0.190000	0.21	1.700000
0.045000				
1244	5.9	0.290000	0.25	2.538806
0.067000				
1400	7.9	0.690000	0.21	2.100000
0.080000				
1401	7.9	0.690000	0.21	2.100000
0.080000				
1419	7.7	0.640000	0.21	2.200000
0.077000				
1493	7.7	0.540000	0.26	1.900000
0.089000				
1496	7.7	0.540000	0.26	1.900000
0.089000				

1559 0.080000	7.8	0.600000	0.26	2.000000
1560 0.080000	7.8	0.600000	0.26	2.000000
1561 0.080000	7.8	0.600000	0.26	2.000000

	free sulfur dioxide	total sulfur dioxide	density	pH
free sulphates \				
14 0.88	15.874922	145.0	0.99860	3.16
15 0.93	15.874922	148.0	0.99860	3.17
86 1.95	20.000000	136.0	0.99720	2.93
88 1.22	34.000000	125.0	0.99780	3.14
90 0.54	42.000000	140.0	0.99640	3.23
91 1.95	20.000000	136.0	0.99720	2.93
92 1.98	19.000000	133.0	0.99720	2.93
109 0.69	37.000000	153.0	0.99690	3.21
130 0.66	30.000000	134.0	0.99680	3.24
145 0.62	32.000000	141.0	0.99680	3.17
154 0.72	29.000000	129.0	0.99730	3.42
155 0.71	28.000000	128.0	0.99730	3.42
156 0.72	29.000000	129.0	0.99730	3.42
157 0.71	28.000000	128.0	0.99730	3.42
188 0.55	15.000000	143.0	0.99680	3.20
189 0.55	17.000000	144.0	0.99680	3.20
190 0.62	21.000000	127.0	0.99760	3.23
192 0.61	16.000000	126.0	0.99690	3.28
201 1.03	39.000000	145.0	0.99750	3.04
219 0.60	24.000000	144.0	0.99655	3.30

313	30.000000	135.0	0.99760	3.30
0.53				
354	40.500000	165.0	0.99120	3.25
0.59				
396	15.874922	124.0	0.99940	3.47
0.53				
400	15.874922	124.0	0.99940	3.47
0.53				
415	31.000000	134.0	1.00140	3.32
1.07				
417	34.000000	124.0	0.99560	3.44
0.48				
463	25.000000	129.0	0.99720	3.08
0.53				
515	34.000000	151.0	1.00100	3.31
1.14				
522	15.874922	133.0	0.99790	3.38
0.99				
523	38.000000	142.0	0.99780	3.22
0.55				
591	23.000000	149.0	0.99220	3.12
0.50				
636	30.000000	147.0	0.99790	3.24
0.53				
637	31.000000	145.0	0.99780	3.24
0.53				
649	24.000000	148.0	0.99480	3.16
0.57				
651	35.000000	155.0	1.00100	3.41
0.67				
672	32.000000	151.0	0.99800	3.15
0.53				
684	35.000000	152.0	0.99800	3.25
0.48				
694	24.000000	125.0	0.99840	3.31
0.61				
723	36.000000	127.0	0.99650	2.94
1.62				
741	28.000000	139.0	0.99788	3.21
0.57				
771	23.000000	143.0	0.99780	3.28
0.55				
772	23.000000	144.0	0.99782	3.27
0.55				
791	25.000000	130.0	0.99818	3.23
0.54				
1079	37.500000	278.0	0.99316	3.01
0.51				
1081	37.500000	289.0	0.99316	3.01

0.51					
1131	15.874922	135.0	0.99341	3.32	
0.44					
1244	15.874922	160.0	0.99721	3.33	
0.54					
1400	33.000000	141.0	0.99620	3.25	
0.51					
1401	33.000000	141.0	0.99620	3.25	
0.51					
1419	32.000000	133.0	0.99560	3.27	
0.45					
1493	23.000000	147.0	0.99636	3.26	
0.59					
1496	23.000000	147.0	0.99636	3.26	
0.59					
1559	31.000000	131.0	0.99622	3.21	
0.52					
1560	31.000000	131.0	0.99622	3.21	
0.52					
1561	31.000000	131.0	0.99622	3.21	
0.52					

	alcohol	quality
14	9.2	5
15	9.2	5
86	9.9	6
88	9.5	5
90	9.5	5
91	9.9	6
92	9.8	5
109	9.3	5
130	9.4	5
145	9.4	5
154	10.5	5
155	10.5	5
156	10.5	5
157	10.5	5
188	9.5	5
189	9.5	5
190	9.4	5
192	9.5	5
201	9.3	5
219	9.5	5
313	9.4	5
354	11.9	6
396	9.9	5
400	9.9	5
415	9.3	5
417	10.5	5

463	9.0	5
515	9.3	5
522	9.8	5
523	9.4	5
591	11.5	6
636	9.4	5
637	9.4	5
649	11.3	6
651	11.2	5
672	9.5	5
684	9.4	5
694	9.4	5
723	9.5	5
741	9.5	5
771	9.4	5
772	9.4	5
791	9.6	5
1079	12.3	7
1081	12.3	7
1131	9.5	5
1244	10.3	6
1400	9.9	5
1401	9.9	5
1419	9.9	5
1493	9.7	5
1496	9.7	5
1559	9.9	5
1560	9.9	5
1561	9.9	5

Outliers in density:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
142	5.200000	0.340	0.00	1.800000
0.050000				
144	5.200000	0.340	0.00	1.800000
0.050000				
294	8.319637	0.340	0.52	3.200000
0.094000				
324	10.000000	0.490	0.20	2.538806
0.071000				
325	10.000000	0.490	0.20	2.538806
0.071000				
353	8.319637	0.530	0.79	2.538806
0.087467				
354	6.100000	0.210	0.40	1.400000
0.066000				
364	8.319637	0.615	0.66	2.538806
0.083000				

366	8.319637	0.615	0.66	2.538806
0.083000				
374	8.319637	0.410	0.63	2.538806
0.089000				
381	8.319637	0.415	0.68	2.900000
0.085000				
391	8.319637	0.415	0.68	2.900000
0.085000				
415	8.600000	0.725	0.24	2.538806
0.117000				
442	8.319637	0.685	0.76	2.538806
0.100000				
480	10.600000	0.280	0.39	2.538806
0.069000				
538	8.319637	0.350	0.49	2.538806
0.066000				
554	8.319637	0.645	0.49	2.538806
0.095000				
555	8.319637	0.645	0.49	2.538806
0.095000				
557	8.319637	0.645	0.49	2.538806
0.095000				
559	8.319637	0.470	0.49	2.538806
0.085000				
564	8.319637	0.470	0.49	2.538806
0.085000				
588	5.000000	0.420	0.24	2.000000
0.060000				
591	6.600000	0.390	0.49	1.700000
0.070000				
608	10.100000	0.650	0.37	2.538806
0.110000				
695	5.100000	0.470	0.02	1.300000
0.087467				
821	4.900000	0.420	0.00	2.100000
0.048000				
836	6.700000	0.280	0.28	2.400000
0.087467				
837	6.700000	0.280	0.28	2.400000
0.087467				
889	10.700000	0.900	0.34	2.538806
0.112000				
999	6.400000	0.690	0.00	1.650000
0.055000				
1017	8.000000	0.180	0.37	0.900000
0.049000				
1018	8.000000	0.180	0.37	0.900000
0.049000				
1114	5.000000	0.400	0.50	2.538806

0.046000				
1122	6.300000	0.470	0.00	1.400000
0.055000				
1126	5.800000	0.290	0.26	1.700000
0.063000				
1228	5.100000	0.420	0.00	1.800000
0.044000				
1269	5.500000	0.490	0.03	1.800000
0.044000				
1270	5.000000	0.380	0.01	1.600000
0.048000				
1298	5.700000	0.600	0.00	1.400000
0.063000				
1434	10.200000	0.540	0.37	2.538806
0.087467				
1435	10.200000	0.540	0.37	2.538806
0.087467				
1474	9.900000	0.500	0.50	2.538806
0.087467				
1475	5.300000	0.470	0.11	2.200000
0.048000				
1476	9.900000	0.500	0.50	2.538806
0.087467				
1477	5.300000	0.470	0.11	2.200000
0.048000				

	free sulfur dioxide	total sulfur dioxide	density	pH
sulphates \				
142	27.000000	63.000000	0.99160	3.68
0.79				
144	27.000000	63.000000	0.99160	3.68
0.79				
294	17.000000	53.000000	1.00140	3.05
0.81				
324	13.000000	50.000000	1.00150	3.16
0.69				
325	13.000000	50.000000	1.00150	3.16
0.69				
353	23.000000	77.000000	1.00180	3.18
0.77				
354	40.500000	46.467792	0.99120	3.25
0.59				
364	7.000000	42.000000	1.00220	3.07
0.73				
366	7.000000	42.000000	1.00220	3.07
0.73				
374	6.000000	47.000000	1.00140	3.01
0.81				
381	17.000000	43.000000	1.00140	3.06

0.80				
391	17.000000	43.000000	1.00140	3.06
0.80				
415	31.000000	46.467792	1.00140	3.32
1.07				
442	6.000000	43.000000	1.00320	2.95
0.68				
480	6.000000	23.000000	1.00260	3.12
0.66				
538	5.000000	35.000000	1.00140	3.20
0.66				
554	10.000000	23.000000	1.00315	2.92
0.74				
555	10.000000	23.000000	1.00315	2.92
0.74				
557	10.000000	23.000000	1.00315	2.92
0.74				
559	6.000000	47.000000	1.00210	3.30
0.68				
564	6.000000	47.000000	1.00210	3.30
0.68				
588	19.000000	50.000000	0.99170	3.72
0.74				
591	23.000000	46.467792	0.99220	3.12
0.50				
608	11.000000	65.000000	1.00260	3.32
0.64				
695	18.000000	44.000000	0.99210	3.90
0.62				
821	16.000000	42.000000	0.99154	3.71
0.74				
836	36.000000	100.000000	0.99064	3.26
0.39				
837	36.000000	100.000000	0.99064	3.26
0.39				
889	23.000000	99.000000	1.00289	3.22
0.68				
999	7.000000	12.000000	0.99162	3.47
0.53				
1017	36.000000	109.000000	0.99007	2.89
0.44				
1018	36.000000	109.000000	0.99007	2.89
0.44				
1114	29.000000	80.000000	0.99020	3.49
0.66				
1122	27.000000	33.000000	0.99220	3.45
0.48				
1126	3.000000	11.000000	0.99150	3.39
0.54				

1228	18.000000	88.000000	0.99157	3.68
0.73				
1269	28.000000	87.000000	0.99080	3.50
0.82				
1270	26.000000	60.000000	0.99084	3.70
0.75				
1298	11.000000	18.000000	0.99191	3.45
0.56				
1434	15.874922	95.000000	1.00369	3.18
0.77				
1435	15.874922	95.000000	1.00369	3.18
0.77				
1474	15.874922	82.000000	1.00242	3.16
0.75				
1475	16.000000	89.000000	0.99182	3.54
0.88				
1476	15.874922	82.000000	1.00242	3.16
0.75				
1477	16.000000	89.000000	0.99182	3.54
0.88				

	alcohol	quality
142	14.000000	6
144	14.000000	6
294	9.500000	6
324	9.200000	6
325	9.200000	6
353	13.000000	5
354	11.900000	6
364	10.000000	7
366	10.000000	7
374	10.800000	6
381	10.000000	6
391	10.000000	6
415	9.300000	5
442	11.200000	7
480	9.200000	5
538	12.000000	7
554	11.100000	5
555	11.100000	5
557	11.100000	5
559	12.700000	6
564	12.700000	6
588	14.000000	8
591	11.500000	6
608	10.400000	6
695	12.800000	6
821	14.000000	7
836	11.700000	7

837	11.700000	7
889	9.300000	5
999	12.900000	6
1017	12.700000	6
1018	12.700000	6
1114	13.600000	6
1122	12.300000	6
1126	13.500000	6
1228	13.600000	7
1269	14.000000	8
1270	14.000000	6
1298	12.200000	6
1434	9.000000	6
1435	9.000000	6
1474	8.800000	5
1475	13.566667	7
1476	8.800000	5
1477	13.600000	7

Outliers in pH:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
45	4.600000	0.520000	0.150000	2.100000
0.054000				
94	5.000000	0.527821	0.040000	1.400000
0.045000				
95	4.700000	0.600000	0.170000	2.300000
0.058000				
151	9.200000	0.520000	0.270976	3.400000
0.087467				
268	6.900000	0.540000	0.040000	3.000000
0.077000				
276	6.900000	0.540000	0.040000	3.000000
0.077000				
440	8.319637	0.310000	0.720000	2.200000
0.072000				
544	8.319637	0.310000	0.740000	1.800000
0.075000				
553	5.000000	0.527821	0.240000	1.600000
0.050000				
554	8.319637	0.645000	0.490000	2.538806
0.095000				
555	8.319637	0.645000	0.490000	2.538806
0.095000				
557	8.319637	0.645000	0.490000	2.538806
0.095000				
588	5.000000	0.420000	0.240000	2.000000
0.060000				

614	9.200000	0.755000	0.180000	2.200000
0.087467				
650	10.700000	0.430000	0.390000	2.200000
0.106000				
656	10.700000	0.430000	0.390000	2.200000
0.106000				
657	12.000000	0.500000	0.590000	1.400000
0.073000				
695	5.100000	0.470000	0.020000	1.300000
0.087467				
821	4.900000	0.420000	0.000000	2.100000
0.048000				
930	6.600000	0.610000	0.010000	1.900000
0.080000				
934	6.600000	0.610000	0.010000	1.900000
0.080000				
996	5.600000	0.660000	0.000000	2.200000
0.087000				
997	5.600000	0.660000	0.000000	2.200000
0.087000				
1017	8.000000	0.180000	0.370000	0.900000
0.049000				
1018	8.000000	0.180000	0.370000	0.900000
0.049000				
1111	5.400000	0.420000	0.270000	2.000000
0.092000				
1270	5.000000	0.380000	0.010000	1.600000
0.048000				
1300	5.200000	0.645000	0.000000	2.150000
0.080000				
1316	5.400000	0.740000	0.000000	1.200000
0.041000				
1319	9.100000	0.760000	0.680000	1.700000
0.087467				
1321	5.000000	0.740000	0.000000	1.200000
0.041000				
1377	5.200000	0.490000	0.260000	2.300000
0.090000				
1470	10.000000	0.690000	0.110000	1.400000
0.084000				
1488	5.600000	0.540000	0.040000	1.700000
0.049000				
1491	5.600000	0.540000	0.040000	1.700000
0.049000				
free sulfur dioxide total sulfur dioxide density pH				
sulphates \				
45	8.0	65.0	0.993400	3.90
0.56				

94	41.0	85.0	0.993800	3.75
0.48				
95	17.0	106.0	0.993200	3.85
0.60				
151	32.0	69.0	0.999600	2.74
2.00				
268	7.0	27.0	0.998700	3.69
0.91				
276	7.0	27.0	0.998700	3.69
0.91				
440	6.0	29.0	0.998700	2.88
0.82				
544	6.0	15.0	1.000800	2.86
0.79				
553	32.0	96.0	0.993400	3.74
0.62				
554	10.0	23.0	0.996747	2.92
0.74				
555	10.0	23.0	0.996747	2.92
0.74				
557	10.0	23.0	0.996747	2.92
0.74				
588	19.0	50.0	0.996747	3.72
0.74				
614	10.0	103.0	0.996900	2.87
1.36				
650	8.0	32.0	0.998600	2.89
0.50				
656	8.0	32.0	0.998600	2.89
0.50				
657	23.0	42.0	0.998000	2.92
0.68				
695	18.0	44.0	0.996747	3.90
0.62				
821	16.0	42.0	0.996747	3.71
0.74				
930	8.0	25.0	0.997460	3.69
0.73				
934	8.0	25.0	0.997460	3.69
0.73				
996	3.0	11.0	0.993780	3.71
0.63				
997	3.0	11.0	0.993780	3.71
0.63				
1017	36.0	109.0	0.996747	2.89
0.44				
1018	36.0	109.0	0.996747	2.89
0.44				
1111	23.0	55.0	0.994710	3.78

0.64					
1270	26.0	60.0	0.996747	3.70	
0.75					
1300	15.0	28.0	0.994440	3.78	
0.61					
1316	16.0	46.0	0.992580	4.01	
0.59					
1319	18.0	64.0	0.996520	2.90	
1.33					
1321	16.0	46.0	0.992580	4.01	
0.59					
1377	23.0	74.0	0.995300	3.71	
0.62					
1470	8.0	24.0	0.995780	2.88	
0.47					
1488	5.0	13.0	0.994200	3.72	
0.58					
1491	5.0	13.0	0.994200	3.72	
0.58					

	alcohol	quality
45	13.1	4
94	10.5	4
95	12.9	6
151	9.4	4
268	9.4	6
276	9.4	6
440	9.8	8
544	8.4	6
553	11.5	5
554	11.1	5
555	11.1	5
557	11.1	5
588	14.0	8
614	10.2	6
650	9.6	5
656	9.6	5
657	10.5	7
695	12.8	6
821	14.0	7
930	10.5	5
934	10.5	5
996	12.8	7
997	12.8	7
1017	12.7	6
1018	12.7	6
1111	12.3	7
1270	14.0	6
1300	12.5	6

1316	12.5	6
1319	9.1	6
1321	12.5	6
1377	12.2	6
1470	9.7	5
1488	11.4	5
1491	11.4	5

Outliers in sulphates:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
13	7.800000	0.610	0.290000	1.600000
0.114000				
17	8.100000	0.560	0.280000	1.700000
0.087467				
19	7.900000	0.320	0.510000	1.800000
0.087467				
43	8.100000	0.660	0.220000	2.200000
0.069000				
79	8.300000	0.625	0.200000	1.500000
0.080000				
81	7.800000	0.430	0.700000	1.900000
0.087467				
83	7.300000	0.670	0.260000	1.800000
0.087467				
86	8.600000	0.490	0.280000	1.900000
0.110000				
88	9.300000	0.390	0.440000	2.100000
0.107000				
91	8.600000	0.490	0.280000	1.900000
0.110000				
92	8.600000	0.490	0.290000	2.000000
0.110000				
106	7.800000	0.410	0.680000	1.700000
0.087467				
151	9.200000	0.520	0.270976	3.400000
0.087467				
161	7.600000	0.680	0.020000	1.300000
0.072000				
169	7.500000	0.705	0.240000	1.800000
0.087467				
181	8.900000	0.610	0.490000	2.000000
0.087467				
201	8.800000	0.370	0.480000	2.100000
0.097000				
226	8.900000	0.590	0.500000	2.000000
0.087467				
240	8.900000	0.635	0.370000	1.700000

0.087467				
258	7.700000	0.410	0.760000	1.800000
0.087467				
281	7.700000	0.270	0.680000	3.500000
0.087467				
338	8.319637	0.490	0.580000	3.000000
0.103000				
339	8.319637	0.280	0.540000	2.300000
0.082000				
340	12.200000	0.340	0.500000	2.400000
0.066000				
369	9.400000	0.270	0.530000	2.400000
0.074000				
372	9.100000	0.280	0.480000	1.800000
0.067000				
376	11.500000	0.450	0.500000	3.000000
0.078000				
377	9.400000	0.270	0.530000	2.400000
0.074000				
415	8.600000	0.725	0.240000	2.538806
0.117000				
451	8.400000	0.370	0.530000	1.800000
0.087467				
477	10.400000	0.240	0.490000	1.800000
0.075000				
482	10.600000	0.360	0.590000	2.200000
0.087467				
483	10.600000	0.360	0.600000	2.200000
0.087467				
503	10.500000	0.260	0.470000	1.900000
0.078000				
504	10.500000	0.240	0.420000	1.800000
0.077000				
506	10.400000	0.240	0.460000	1.800000
0.075000				
515	8.500000	0.655	0.490000	2.538806
0.087467				
586	11.100000	0.310	0.490000	2.700000
0.094000				
614	9.200000	0.755	0.180000	2.200000
0.087467				
639	8.900000	0.290	0.350000	1.900000
0.067000				
689	8.100000	0.380	0.480000	1.800000
0.087467				
692	8.600000	0.490	0.510000	2.000000
0.087467				
723	7.100000	0.310	0.300000	2.200000
0.053000				

754	7.800000	0.480	0.680000	1.700000
0.087467				
795	10.800000	0.890	0.300000	2.600000
0.087467				
852	8.000000	0.420	0.320000	2.500000
0.080000				
1051	8.500000	0.460	0.590000	1.400000
0.087467				
1158	6.700000	0.410	0.430000	2.800000
0.076000				
1165	8.500000	0.440	0.500000	1.900000
0.087467				
1260	8.600000	0.635	0.680000	1.800000
0.087467				
1288	7.000000	0.600	0.300000	2.538806
0.068000				
1289	7.000000	0.600	0.300000	2.538806
0.068000				
1319	9.100000	0.760	0.680000	1.700000
0.087467				
1367	6.900000	0.540	0.300000	2.200000
0.088000				
1370	8.700000	0.780	0.510000	1.700000
0.087467				
1371	7.500000	0.580	0.560000	3.100000
0.087467				
1372	8.700000	0.780	0.510000	1.700000
0.087467				
1403	7.200000	0.330	0.330000	1.700000
0.061000				
1408	8.100000	0.290	0.360000	2.200000
0.048000				
	free sulfur dioxide	total sulfur dioxide	density	pH \
13	9.0	29.000000	0.997400	3.260000
17	16.0	56.000000	0.996800	3.110000
19	17.0	56.000000	0.996900	3.040000
43	9.0	23.000000	0.996800	3.300000
79	27.0	119.000000	0.997200	3.160000
81	22.0	67.000000	0.997400	3.130000
83	16.0	51.000000	0.996900	3.160000
86	20.0	46.467792	0.997200	2.930000
88	34.0	46.467792	0.997800	3.140000
91	20.0	46.467792	0.997200	2.930000
92	19.0	46.467792	0.997200	2.930000
106	18.0	69.000000	0.997300	3.080000
151	32.0	69.000000	0.999600	3.311113
161	9.0	20.000000	0.996500	3.170000
169	15.0	63.000000	0.996400	3.000000
181	23.0	110.000000	0.997200	3.120000

201	39.0	46.467792	0.997500	3.040000
226	27.0	81.000000	0.996400	3.040000
240	5.0	62.000000	0.997100	3.000000
258	8.0	45.000000	0.996800	3.060000
281	5.0	10.000000	0.997200	3.250000
338	28.0	99.000000	1.000800	3.160000
339	12.0	29.000000	0.999700	3.110000
340	10.0	21.000000	1.000000	3.120000
369	6.0	18.000000	0.996200	3.200000
372	26.0	46.000000	0.996700	3.320000
376	19.0	47.000000	1.000300	3.260000
377	6.0	18.000000	0.996200	3.200000
415	31.0	46.467792	0.996747	3.320000
451	9.0	26.000000	0.997900	3.060000
477	6.0	20.000000	0.997700	3.180000
482	6.0	18.000000	0.998600	3.040000
483	7.0	18.000000	0.998600	3.040000
503	6.0	24.000000	0.997600	3.180000
504	6.0	22.000000	0.997600	3.210000
506	6.0	21.000000	0.997600	3.250000
515	34.0	46.467792	1.001000	3.310000
586	16.0	47.000000	0.998600	3.120000
614	10.0	103.000000	0.996900	3.311113
639	25.0	57.000000	0.997000	3.180000
689	5.0	17.000000	0.997600	3.300000
692	16.0	62.000000	0.997900	3.030000
723	36.0	46.467792	0.996500	2.940000
754	14.0	32.000000	0.996560	3.090000
795	7.0	60.000000	0.997860	2.990000
852	26.0	122.000000	0.998010	3.220000
1051	16.0	45.000000	0.997020	3.030000
1158	22.0	54.000000	0.995720	3.420000
1165	15.0	38.000000	0.996340	3.010000
1260	19.0	56.000000	0.996320	3.020000
1288	20.0	110.000000	0.999140	3.300000
1289	20.0	110.000000	0.999140	3.300000
1319	18.0	64.000000	0.996520	3.311113
1367	9.0	105.000000	0.997250	3.250000
1370	12.0	66.000000	0.996230	3.000000
1371	5.0	14.000000	0.994760	3.210000
1372	12.0	66.000000	0.996230	3.000000
1403	3.0	13.000000	0.996000	3.230000
1408	35.0	53.000000	0.995000	3.270000

	sulphates	alcohol	quality
13	1.56	9.1	5
17	1.28	9.3	5
19	1.08	9.2	6
43	1.20	10.3	5

79	1.12	9.1	4
81	1.28	9.4	5
83	1.14	9.4	5
86	1.95	9.9	6
88	1.22	9.5	5
91	1.95	9.9	6
92	1.98	9.8	5
106	1.31	9.3	5
151	2.00	9.4	4
161	1.08	9.2	4
169	1.59	9.5	5
181	1.02	9.3	5
201	1.03	9.3	5
226	1.61	9.5	6
240	1.09	9.3	5
258	1.26	9.4	5
281	1.08	9.9	7
338	1.00	11.5	6
339	1.36	9.8	7
340	1.18	9.2	6
369	1.13	12.0	7
372	1.04	10.6	6
376	1.11	11.0	6
377	1.13	12.0	7
415	1.07	9.3	5
451	1.06	9.1	6
477	1.06	11.0	6
482	1.05	9.4	5
483	1.06	9.4	5
503	1.04	10.9	7
504	1.05	10.8	7
506	1.02	10.8	7
515	1.14	9.3	5
586	1.02	10.6	7
614	1.36	10.2	6
639	1.36	10.3	6
689	1.05	9.4	5
692	1.17	9.0	5
723	1.62	9.5	5
754	1.06	9.1	6
795	1.18	10.2	5
852	1.07	9.7	5
1051	1.34	9.2	5
1158	1.16	10.6	6
1165	1.10	9.4	5
1260	1.15	9.3	5
1288	1.17	10.2	5
1289	1.17	10.2	5
1319	1.33	9.1	6

1367	1.18	10.5	6
1370	1.17	9.2	5
1371	1.03	11.6	6
1372	1.17	9.2	5
1403	1.10	10.0	8
1408	1.01	12.4	7

Outliers in alcohol:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
142	5.200000	0.34	0.00	1.800000
0.050				
144	5.200000	0.34	0.00	1.800000
0.050				
467	8.800000	0.46	0.45	2.600000
0.065				
588	5.000000	0.42	0.24	2.000000
0.060				
652	8.319637	0.36	0.65	2.538806
0.096				
821	4.900000	0.42	0.00	2.100000
0.048				
1114	5.000000	0.40	0.50	2.538806
0.046				
1132	7.400000	0.36	0.34	1.800000
0.075				
1228	5.100000	0.42	0.00	1.800000
0.044				
1269	5.500000	0.49	0.03	1.800000
0.044				
1270	5.000000	0.38	0.01	1.600000
0.048				
1475	5.300000	0.47	0.11	2.200000
0.048				
1477	5.300000	0.47	0.11	2.200000
0.048				

	free sulfur dioxide	total sulfur dioxide	density	pH \
142	27.0	63.0	0.996747	3.680000
144	27.0	63.0	0.996747	3.680000
467	7.0	18.0	0.994700	3.320000
588	19.0	50.0	0.996747	3.311113
652	22.0	71.0	0.997600	2.980000
821	16.0	42.0	0.996747	3.311113
1114	29.0	80.0	0.996747	3.490000
1132	18.0	38.0	0.993300	3.380000
1228	18.0	88.0	0.996747	3.680000
1269	28.0	87.0	0.996747	3.500000
1270	26.0	60.0	0.996747	3.311113

1475	16.0	89.0	0.996747	3.540000
1477	16.0	89.0	0.996747	3.540000

	sulphates	alcohol	quality
142	0.79	14.000000	6
144	0.79	14.000000	6
467	0.79	14.000000	6
588	0.74	14.000000	8
652	0.84	14.900000	5
821	0.74	14.000000	7
1114	0.66	13.600000	6
1132	0.88	13.600000	7
1228	0.73	13.600000	7
1269	0.82	14.000000	8
1270	0.75	14.000000	6
1475	0.88	13.566667	7
1477	0.88	13.600000	7

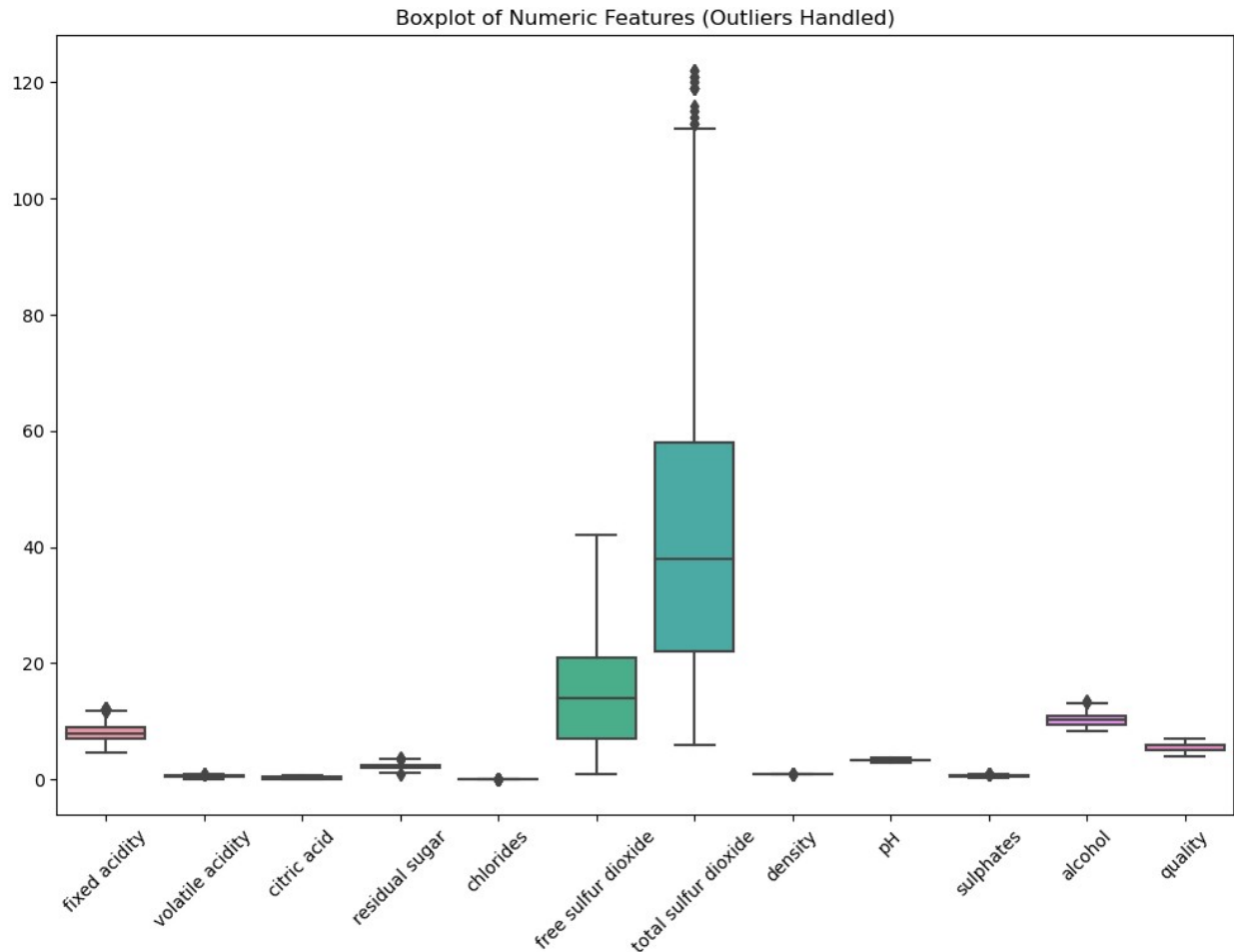
Outliers in quality:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
267	7.900000	0.350000	0.46	3.600000
0.078000				
278	10.300000	0.320000	0.45	2.538806
0.073000				
390	5.600000	0.850000	0.05	1.400000
0.045000				
440	8.319637	0.310000	0.72	2.200000
0.072000				
455	11.300000	0.620000	0.67	2.538806
0.086000				
459	11.600000	0.580000	0.66	2.200000
0.074000				
481	9.400000	0.300000	0.56	2.800000
0.080000				
495	10.700000	0.350000	0.53	2.600000
0.070000				
498	10.700000	0.350000	0.53	2.600000
0.070000				
517	10.400000	0.610000	0.49	2.100000
0.087467				
588	5.000000	0.420000	0.24	2.000000
0.060000				
690	7.400000	0.527821	0.00	2.538806
0.097000				
828	7.800000	0.570000	0.09	2.300000
0.065000				
832	10.400000	0.440000	0.42	1.500000
0.087467				

899	8.300000	0.527821	0.02	3.400000
0.084000				
1061	9.100000	0.400000	0.50	1.800000
0.071000				
1090	10.000000	0.260000	0.54	1.900000
0.083000				
1120	7.900000	0.540000	0.34	2.500000
0.076000				
1202	8.600000	0.420000	0.39	1.800000
0.068000				
1269	5.500000	0.490000	0.03	1.800000
0.044000				
1299	7.600000	0.527821	0.00	2.100000
0.087467				
1374	6.800000	0.815000	0.00	1.200000
0.087467				
1403	7.200000	0.330000	0.33	1.700000
0.061000				
1449	7.200000	0.380000	0.31	2.000000
0.056000				
1469	7.300000	0.980000	0.05	2.100000
0.061000				
1478	7.100000	0.875000	0.05	2.538806
0.082000				
1505	6.700000	0.760000	0.02	1.800000
0.078000				
1549	7.400000	0.360000	0.30	1.800000
0.074000				
	free sulfur dioxide	total sulfur dioxide	density	pH \
267	15.0	37.0	0.997300	3.350000
278	5.0	13.0	0.997600	3.230000
390	12.0	88.0	0.992400	3.560000
440	6.0	29.0	0.998700	3.311113
455	6.0	19.0	0.998800	3.220000
459	10.0	47.0	1.000800	3.250000
481	6.0	17.0	0.996400	3.150000
495	5.0	16.0	0.997200	3.150000
498	5.0	16.0	0.997200	3.150000
517	5.0	16.0	0.999400	3.160000
588	19.0	50.0	0.996747	3.311113
690	5.0	14.0	0.996600	3.630000
828	34.0	45.0	0.994170	3.460000
832	34.0	48.0	0.998320	3.380000
899	6.0	11.0	0.998920	3.480000
1061	7.0	16.0	0.994620	3.210000
1090	42.0	74.0	0.994510	2.980000
1120	8.0	17.0	0.992350	3.200000
1202	6.0	12.0	0.995160	3.350000

1269	28.0	87.0	0.996747	3.500000
1299	5.0	9.0	0.994760	3.500000
1374	16.0	29.0	0.994710	3.320000
1403	3.0	13.0	0.996000	3.230000
1449	15.0	29.0	0.994720	3.230000
1469	20.0	49.0	0.997050	3.310000
1478	3.0	14.0	0.998080	3.400000
1505	6.0	12.0	0.996000	3.550000
1549	17.0	24.0	0.994190	3.240000

	sulphates	alcohol	quality
267	0.860000	12.800000	8
278	0.820000	12.600000	8
390	0.820000	12.900000	8
440	0.820000	9.800000	8
455	0.690000	13.400000	8
459	0.570000	9.000000	3
481	0.920000	11.700000	8
495	0.650000	11.000000	8
498	0.650000	11.000000	8
517	0.630000	8.400000	3
588	0.740000	10.422983	8
690	0.540000	10.700000	3
828	0.740000	12.700000	8
832	0.860000	9.900000	3
899	0.490000	11.000000	3
1061	0.690000	12.500000	8
1090	0.630000	11.800000	8
1120	0.720000	13.100000	8
1202	0.690000	11.700000	8
1269	0.820000	10.422983	8
1299	0.400000	10.900000	3
1374	0.510000	9.800000	3
1403	0.658149	10.000000	8
1449	0.760000	11.300000	8
1469	0.550000	9.700000	3
1478	0.520000	10.200000	3
1505	0.630000	9.950000	3
1549	0.700000	11.400000	8



```
# Task 2: Data preprocessing including visualization
print("Task 2: Data preprocessing including visualization\n\n\n\n")

# Summary statistics
print(winequality.describe())

# Correlation heatmap
correlation_matrix = winequality.corr()
plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap="coolwarm")
plt.title("Correlation Heatmap")
plt.show()
```

Task 2: Data preprocessing including visualization

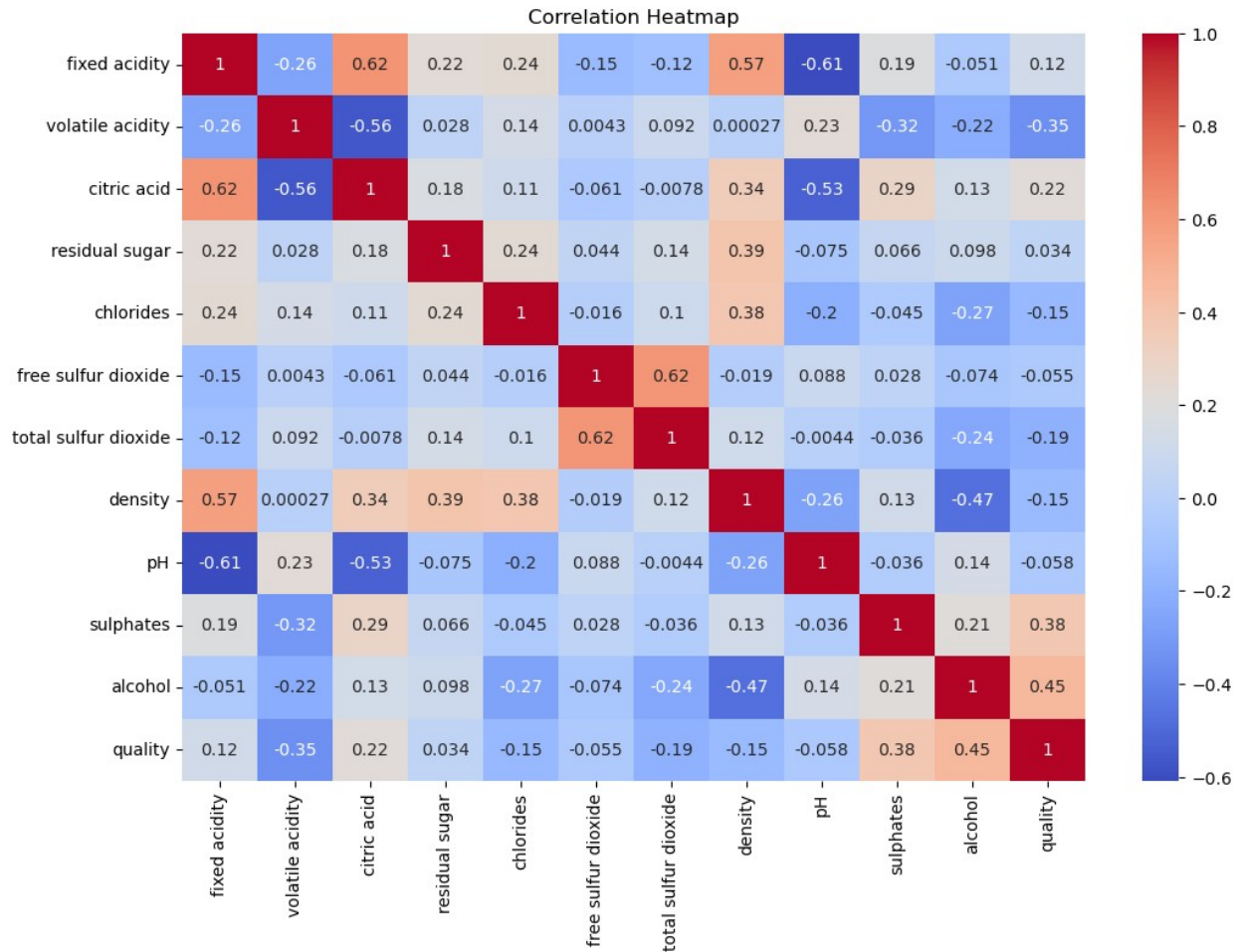
	fixed acidity	volatile acidity	citric acid	residual sugar	\
count	1599.000000	1599.000000	1599.000000	1599.000000	
mean	8.167394	0.52064	0.270520	2.214643	

std	1.490441	0.16559	0.193945	0.439827
min	4.600000	0.12000	0.000000	0.900000
25%	7.100000	0.39000	0.090000	1.900000
50%	7.900000	0.52000	0.260000	2.200000
75%	9.000000	0.63000	0.420000	2.538806
max	12.300000	1.01000	0.790000	3.650000

	chlorides	free sulfur dioxide	total sulfur dioxide
density \			
count	1599.000000	1599.000000	1599.000000
1599.000000			

mean	0.079366	15.199029	43.116778
0.996735			
std	0.014506	9.140046	26.749839
0.001643			
min	0.041000	1.000000	6.000000
0.992350			
25%	0.070000	7.000000	22.000000
0.995680			
50%	0.080000	14.000000	38.000000
0.996747			
75%	0.087467	21.000000	58.000000
0.997800			
max	0.119000	42.000000	122.000000
1.001000			

	pH	sulphates	alcohol	quality
count	1599.000000	1599.000000	1599.000000	1599.000000
mean	3.308824	0.637217	10.394610	5.625897
std	0.138952	0.118780	1.017252	0.738671
min	2.930000	0.330000	8.400000	4.000000
25%	3.210000	0.550000	9.500000	5.000000
50%	3.310000	0.620000	10.200000	6.000000
75%	3.400000	0.710000	11.000000	6.000000
max	3.680000	0.990000	13.500000	7.000000



```
# Task 3: Machine Learning Model building
print("Task 3: Machine Learning Model building\n")

print("simple linear regression model\n\n")

# Split the data into features (X) and target (y)
X = winequality.drop(columns=["quality"])
y = winequality["quality"]

# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size=0.2, random_state=42)

# Create and train the model
model = LinearRegression()
model.fit(X_train, y_train)
print("Coefficients:\n", model.coef_)
print("\nIntercept:", model.intercept_)

# Predict wine quality on the test set
y_pred = model.predict(X_test)
```

Task 3: Machine Learning Model building

simple linear regression model

Coefficients:

```
[ 2.76473387e-02 -7.79557572e-01 -1.97372126e-01  3.08511810e-02  
-8.45068967e-01  2.83769828e-03 -2.62366581e-03 -1.69720757e+01  
-3.05377648e-01  1.53115721e+00  2.47871989e-01]
```

Intercept: 20.298968156503356

Predict wine quality on the test set

```
y_pred = model.predict(X_test)
```

Task 4: Evaluate the model

```
print("Task 4: Evaluate the model\n\n")
```

```
mse = mean_squared_error(y_test, y_pred)
```

```
r2 = r2_score(y_test, y_pred)
```

```
print("Mean Squared Error:", mse)
```

```
print("R-squared:", r2)
```

Task 4: Evaluate the model

Mean Squared Error: 0.3471079906969371

R-squared: 0.3647650475176243

Task 5: Test with random observation

```
print("Task 5: Test with random observation\n\n")
```

```
observations = [
```

```
{
```

```
    'fixed acidity': 7.0,
```

```
    'volatile acidity': 0.6,
```

```
    'citric acid': 0.1,
```

```
    'residual sugar': 2.1,
```

```
    'chlorides': 0.07,
```

```
    'free sulfur dioxide': 25.0,
```

```
    'total sulfur dioxide': 60.0,
```

```
    'density': 0.995,
```

```
    'pH': 3.0,
```

```
    'sulphates': 0.5,
```

```
    'alcohol': 10.0
```

```
},
```

```
{
```

```
    'fixed acidity': 7.2,
```

```
    'volatile acidity': 0.42,
```

```
    'citric acid': 0.24,
```

```

        'residual sugar': 2.5,
        'chlorides': 0.076,
        'free sulfur dioxide': 15.0,
        'total sulfur dioxide': 37.0,
        'density': 0.995,
        'pH': 3.2,
        'sulphates': 0.58,
        'alcohol': 9.0
    },
]

# Create a DataFrame from the list of observations
test_data = pd.DataFrame(observations)
print("Test_data:\n", test_data)

# Predict the quality for each observation
predicted_qualities = model.predict(test_data)
print("\n\n\n\nPredicted Qualities:", predicted_qualities)

```

Task 5: Test with random observation

Test_data:

	fixed acidity	volatile acidity	citric acid	residual sugar
chlorides \				
0	7.0	0.60	0.10	2.1
0.070				
1	7.2	0.42	0.24	2.5
0.076				

	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates
alcohol					
0	25.0	60.0	0.995	3.0	0.50
10.0					
1	15.0	37.0	0.995	3.2	0.58
9.0					

Predicted Qualities: [5.3651332 5.33613338]