

STATISTIC FOR DATA SCIENCE Part 2

Import Library

In [1]:

import pandas as np

Import data

In [2]:

dataset_red = pd.read_csv(r"D:\SEMESTER 4\IS411 Data Modelling\LAB\Bahan Modul 6\winequality-red.csv", delimiter=',')
dataset_red.head(5)

Out[2]:

| | fixed acidity | volatile acidity | citric acid | residual sugar | chlorides | free sulfur dioxide | total sulfur dioxide | density | pH | sulphates | alcohol | quality |
|---|---------------|------------------|-------------|----------------|-----------|---------------------|----------------------|---------|------|-----------|---------|---------|
| 0 | 7.4 | 0.70 | 0.00 | 1.9 | 0.076 | 11.0 | 34.0 | 0.9978 | 3.51 | 0.56 | 9.4 | 5 |
| 1 | 7.8 | 0.88 | 0.00 | 2.6 | 0.098 | 25.0 | 67.0 | 0.9968 | 3.20 | 0.68 | 9.8 | 5 |
| 2 | 7.8 | 0.76 | 0.04 | 2.3 | 0.092 | 15.0 | 54.0 | 0.9970 | 3.26 | 0.65 | 9.8 | 5 |
| 3 | 11.2 | 0.28 | 0.56 | 1.9 | 0.075 | 17.0 | 60.0 | 0.9980 | 3.16 | 0.58 | 9.8 | 6 |
| 4 | 7.4 | 0.70 | 0.00 | 1.9 | 0.076 | 11.0 | 34.0 | 0.9978 | 3.51 | 0.56 | 9.4 | 5 |

In [3]:

dataset_red.info()
dataset_red.shape

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1599 entries, 0 to 1598
Data columns (total 12 columns):
Column Non-Null Count Dtype ---
0 fixed acidity 1599 non-null float64
1 volatile acidity 1599 non-null float64
2 citric acid 1599 non-null float64
3 residual sugar 1599 non-null float64
4 chlorides 1599 non-null float64
5 free sulfur dioxide 1599 non-null float64
6 total sulfur dioxide 1599 non-null float64
7 density 1599 non-null float64
8 pH 1599 non-null float64
9 sulphates 1599 non-null float64
10 alcohol 1599 non-null float64
11 quality 1599 non-null int64
dtypes: float64(11), int64(1)
memory usage: 158.0 KB
(1599, 12)

In [4]:

dataset_red.describe()

Out[4]:

| | fixed acidity | volatile acidity | citric acid | residual sugar | chlorides | free sulfur dioxide | total sulfur dioxide | density | pH | sulphates | alcohol | quality |
|-------|---------------|------------------|-------------|----------------|-------------|---------------------|----------------------|-------------|-------------|-------------|-------------|-------------|
| count | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 |
| mean | 8.319637 | 0.527821 | 0.270976 | 2.538806 | 0.087467 | 15.874922 | 46.467792 | 0.996747 | 3.311113 | 0.658149 | 10.422983 | 5.636023 |
| std | 1.741096 | 0.179060 | 0.194801 | 1.409928 | 0.047065 | 10.460157 | 32.895324 | 0.001887 | 0.154386 | 0.169507 | 1.065668 | 0.807569 |
| min | 4.600000 | 0.120000 | 0.000000 | 0.900000 | 0.012000 | 1.000000 | 6.000000 | 0.990070 | 2.740000 | 0.330000 | 8.400000 | 3.000000 |
| 25% | 7.100000 | 0.390000 | 0.050000 | 1.900000 | 0.070000 | 7.000000 | 22.000000 | 0.995600 | 3.210000 | 0.550000 | 9.500000 | 5.000000 |
| 50% | 7.900000 | 0.520000 | 0.260000 | 2.200000 | 0.079000 | 14.000000 | 38.000000 | 0.996750 | 3.310000 | 0.620000 | 10.200000 | 6.000000 |
| 75% | 9.200000 | 0.640000 | 0.420000 | 2.600000 | 0.090000 | 21.000000 | 62.000000 | 0.997835 | 3.400000 | 0.730000 | 11.100000 | 6.000000 |
| max | 15.900000 | 1.580000 | 1.000000 | 15.500000 | 0.611000 | 72.000000 | 289.000000 | 1.003690 | 4.010000 | 2.000000 | 14.900000 | 8.000000 |

Menampilkan Statistik Data

Menampilkan Statistik Data

Mean

Mean, digunakan untuk menghitung rata-rata.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|---------------|----------|------------------|----------|-------------|----------|----------------|----------|-----------|----------|---------------------|-----------|----------------------|-----------|---------|----------|----|----------|-----------|----------|---------|-----------|---------|----------|--------|---------|
| In [5]: | <pre>#Menampilkan hasil mean dari salah satu kolom print('mean dari fixed acidity : ', dataset_red['fixed acidity'].mean())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | mean dari fixed acidity : 8.319637273295838 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [6]: | <pre>#Menampilkan mean dari keseluruhan data print(dataset_red.mean())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>8.319637</td></tr><tr><td>volatile acidity</td><td>0.527821</td></tr><tr><td>citric acid</td><td>0.270976</td></tr><tr><td>residual sugar</td><td>2.538806</td></tr><tr><td>chlorides</td><td>0.087467</td></tr><tr><td>free sulfur dioxide</td><td>15.874922</td></tr><tr><td>total sulfur dioxide</td><td>46.467792</td></tr><tr><td>density</td><td>0.996747</td></tr><tr><td>pH</td><td>3.311113</td></tr><tr><td>sulphates</td><td>0.658149</td></tr><tr><td>alcohol</td><td>10.422983</td></tr><tr><td>quality</td><td>5.636023</td></tr><tr><td>dtype:</td><td>float64</td></tr></tbody></table> | fixed acidity | 8.319637 | volatile acidity | 0.527821 | citric acid | 0.270976 | residual sugar | 2.538806 | chlorides | 0.087467 | free sulfur dioxide | 15.874922 | total sulfur dioxide | 46.467792 | density | 0.996747 | pH | 3.311113 | sulphates | 0.658149 | alcohol | 10.422983 | quality | 5.636023 | dtype: | float64 |
| fixed acidity | 8.319637 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.527821 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.270976 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 2.538806 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.087467 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 15.874922 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 46.467792 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.996747 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 3.311113 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.658149 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 10.422983 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 5.636023 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| dtype: | float64 | | | | | | | | | | | | | | | | | | | | | | | | | | |

Median

Median, digunakan untuk menemukan nilai tengah dan menentukan bagian atas atau bawah dari kumpulan data, menggunakan function median().

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|---------------|----------|------------------|----------|-------------|----------|----------------|----------|-----------|----------|---------------------|-----------|----------------------|-----------|---------|----------|------|----------|-----------|----------|---------|-----------|---------|----------|--------|---------|
| In [7]: | <pre>#Menampilkan median salah satu kolom print('median dari fixed acidity : ', dataset_red['fixed acidity'].median())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | median dari fixed acidity : 7.9 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [8]: | <pre>#Menampilkan median semua data print(dataset_red.median())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>7.900000</td></tr><tr><td>volatile acidity</td><td>0.520000</td></tr><tr><td>citric acid</td><td>0.260000</td></tr><tr><td>residual sugar</td><td>2.200000</td></tr><tr><td>chlorides</td><td>0.079000</td></tr><tr><td>free sulfur dioxide</td><td>14.000000</td></tr><tr><td>total sulfur dioxide</td><td>38.000000</td></tr><tr><td>density</td><td>0.996750</td></tr><tr><td>pH</td><td>3.310000</td></tr><tr><td>sulphates</td><td>0.620000</td></tr><tr><td>alcohol</td><td>10.200000</td></tr><tr><td>quality</td><td>6.000000</td></tr><tr><td>dtype:</td><td>float64</td></tr></tbody></table> | fixed acidity | 7.900000 | volatile acidity | 0.520000 | citric acid | 0.260000 | residual sugar | 2.200000 | chlorides | 0.079000 | free sulfur dioxide | 14.000000 | total sulfur dioxide | 38.000000 | density | 0.996750 | pH | 3.310000 | sulphates | 0.620000 | alcohol | 10.200000 | quality | 6.000000 | dtype: | float64 |
| fixed acidity | 7.900000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.520000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.260000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 2.200000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.079000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 14.000000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 38.000000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.996750 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 3.310000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.620000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 10.200000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 6.000000 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| dtype: | float64 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [9]: | <pre>#Jika data median yang ingin dicari adalah per row, dapat menggunakan attribut axis = 1 print(dataset_red.median(axis=1))</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>0</td><td>2.785</td></tr><tr><td>1</td><td>2.980</td></tr><tr><td>2</td><td>2.780</td></tr><tr><td>3</td><td>2.530</td></tr><tr><td>4</td><td>2.705</td></tr><tr><td>...</td><td>...</td></tr><tr><td>1594</td><td>2.725</td></tr><tr><td>1595</td><td>2.860</td></tr><tr><td>1596</td><td>2.860</td></tr><tr><td>1597</td><td>2.785</td></tr><tr><td>1598</td><td>3.495</td></tr></tbody></table> Length: 1599, dtype: float64 | 0 | 2.785 | 1 | 2.980 | 2 | 2.780 | 3 | 2.530 | 4 | 2.705 | ... | ... | 1594 | 2.725 | 1595 | 2.860 | 1596 | 2.860 | 1597 | 2.785 | 1598 | 3.495 | | | | |
| 0 | 2.785 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 2.980 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | 2.780 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 2.530 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | 2.705 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ... | ... | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1594 | 2.725 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1595 | 2.860 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1596 | 2.860 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1597 | 2.785 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1598 | 3.495 | | | | | | | | | | | | | | | | | | | | | | | | | | |

Modus

Mode, digunakan untuk menghitung data yang paling banyak muncul

1596

2.860

1597

2.785

1598

3.495

Length: 1599, dtype: float64

Modus

Mode, digunakan untuk menghitung data yang paling banyak muncul

In [10]:

#Menampilkan modus dari salah satu kolom

print('mode dari fixed acidity : ', dataset_red['fixed acidity'].mode())

mode dari fixed acidity : 0 7.2

Name: fixed acidity, dtype: float64

In [11]:

#Menampilkan modus semua data

print(dataset_red.mode())

| | | | | | |
|---------------------|----------------------|-------------|----------------|-----------|------|
| fixed acidity | volatile acidity | citric acid | residual sugar | chlorides | \ |
| 0 | 7.2 | 0.6 | 0.0 | 2.0 | 0.68 |
| free sulfur dioxide | total sulfur dioxide | density | pH | sulphates | \ |
| 0 | 6.0 | 28.0 | 0.9972 | 3.3 | 0.6 |
| alcohol | quality | | | | |
| 0 | 9.5 | 5 | | | |

Standard Deviation

digunakan untuk menghitung banyaknya variasi data terhadap mean dari data tersebut.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|--|---------------|----------|------------------|----------|-------------|----------|----------------|----------|-----------|----------|---------------------|-----------|----------------------|-----------|---------|----------|----|----------|-----------|----------|---------|----------|---------|----------|--------|---------|
| In [12]: | <pre>#Menampilkan Standard Deviation salah satu kolom print('Standard deviation dari fixed acidity : ', dataset_red['fixed acidity'].std())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Standard deviation dari fixed acidity : 1.7410963181277086 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [13]: | <pre>#Menampilkan Standard Deviation semua data print(dataset_red.std())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>1.741096</td></tr><tr><td>volatile acidity</td><td>0.179060</td></tr><tr><td>citric acid</td><td>0.194801</td></tr><tr><td>residual sugar</td><td>1.409928</td></tr><tr><td>chlorides</td><td>0.047065</td></tr><tr><td>free sulfur dioxide</td><td>10.460157</td></tr><tr><td>total sulfur dioxide</td><td>32.895324</td></tr><tr><td>density</td><td>0.001887</td></tr><tr><td>pH</td><td>0.154386</td></tr><tr><td>sulphates</td><td>0.169507</td></tr><tr><td>alcohol</td><td>1.065668</td></tr><tr><td>quality</td><td>0.807569</td></tr><tr><td>dtype:</td><td>float64</td></tr></tbody></table> | fixed acidity | 1.741096 | volatile acidity | 0.179060 | citric acid | 0.194801 | residual sugar | 1.409928 | chlorides | 0.047065 | free sulfur dioxide | 10.460157 | total sulfur dioxide | 32.895324 | density | 0.001887 | pH | 0.154386 | sulphates | 0.169507 | alcohol | 1.065668 | quality | 0.807569 | dtype: | float64 |
| fixed acidity | 1.741096 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.179060 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.194801 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 1.409928 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.047065 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 10.460157 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 32.895324 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.001887 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 0.154386 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.169507 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 1.065668 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 0.807569 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| dtype: | float64 | | | | | | | | | | | | | | | | | | | | | | | | | | |

Variance

digunakan untuk menghitung penyebaran data. Dan merupakan kuadrat dari standard deviasi, dan menjadi kovarian, atau hubungan dari salah satu variable ke variable lainnya.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|---------------|----------|------------------|----------|-------------|----------|----------------|----------|-----------|----------|---------------------|------------|----------------------|-------------|---------|----------|----|----------|-----------|----------|---------|----------|---------|----------|--------|---------|
| In [14]: | <pre>#Menampilkan Variance Salah satu kolom print('variance dari fixed acidity : ', dataset_red['fixed acidity'].var())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | variance dari fixed acidity : 3.031416388997835 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [15]: | <pre>#Menampilkan Variance semua data print(dataset_red.var())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>3.031416</td></tr><tr><td>volatile acidity</td><td>0.032802</td></tr><tr><td>citric acid</td><td>0.037947</td></tr><tr><td>residual sugar</td><td>1.987897</td></tr><tr><td>chlorides</td><td>0.002215</td></tr><tr><td>free sulfur dioxide</td><td>109.414804</td></tr><tr><td>total sulfur dioxide</td><td>1082.102373</td></tr><tr><td>density</td><td>0.000004</td></tr><tr><td>pH</td><td>0.022835</td></tr><tr><td>sulphates</td><td>0.020723</td></tr><tr><td>alcohol</td><td>1.135647</td></tr><tr><td>quality</td><td>0.652168</td></tr><tr><td>dtype:</td><td>float64</td></tr></tbody></table> | fixed acidity | 3.031416 | volatile acidity | 0.032802 | citric acid | 0.037947 | residual sugar | 1.987897 | chlorides | 0.002215 | free sulfur dioxide | 109.414804 | total sulfur dioxide | 1082.102373 | density | 0.000004 | pH | 0.022835 | sulphates | 0.020723 | alcohol | 1.135647 | quality | 0.652168 | dtype: | float64 |
| fixed acidity | 3.031416 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.032802 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.037947 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 1.987897 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.002215 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 109.414804 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 1082.102373 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.000004 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 0.022835 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.020723 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 1.135647 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 0.652168 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| dtype: | float64 | | | | | | | | | | | | | | | | | | | | | | | | | | |

Quartile

adaiah persentase 25%, 50%, dan 75% yang membagi data menjadi 4 bagian dengan rata.

| | |
|----------|--|
| In [16]: | <pre>#Menampilkan Q1-Q4 print('Q1 dari fixed acidity : ',np.quantile(dataset_red['fixed acidity'], 0.25)) print('Q2 dari fixed acidity : ',np.quantile(dataset_red['fixed acidity'], 0.5)) print('Q3 dari fixed acidity : ',np.quantile(dataset_red['fixed acidity'], 0.75)) print('Q4 dari fixed acidity : ',np.quantile(dataset_red['fixed acidity'], 0.95))</pre> |
| | Q1 dari fixed acidity : 7.1 Q2 dari fixed acidity : 7.9 Q3 dari fixed acidity : 9.2 |
| | Q2 atau median dari fixed acidity : 7.9 Q3 atau median dari fixed acidity : 9.2 |

Interquartile Range(IQR)

dapat dihitung dengan rumus sederhana Q3 - Q1 ataupun bisa menggunakan library scipy

Menggunakan rumus Q3-Q1 : 2.6999999999999996
Menggunakan bantuan library scipy.stats 2.6999999999999996

In [18]:

dataset_red.describe(include='all')

Out[18]:

| | fixed acidity | volatile acidity | citric acid | residual sugar | chlorides | free sulfur dioxide | total sulfur dioxide | density | pH | sulphates | alcohol | quality |
|-------|---------------|------------------|-------------|----------------|-------------|---------------------|----------------------|-------------|-------------|-------------|-------------|-------------|
| count | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 | 1599.000000 |
| mean | 8.319637 | 0.527821 | 0.270976 | 2.538806 | 0.087467 | 15.874922 | 46.467792 | 0.996747 | 3.311113 | 0.658149 | 10.422983 | 5.636023 |
| std | 1.741096 | 0.179060 | 0.194801 | 1.409928 | 0.047065 | 10.460157 | 32.895324 | 0.001887 | 0.154386 | 0.169507 | 1.065668 | 0.807569 |
| min | 4.600000 | 0.120000 | 0.000000 | 0.900000 | 0.012000 | 1.000000 | 6.000000 | 0.990070 | 2.740000 | 0.330000 | 8.400000 | 3.000000 |
| 25% | 7.100000 | 0.390000 | 0.050000 | 1.900000 | 0.070000 | 7.000000 | 22.000000 | 0.995600 | 3.210000 | 0.550000 | 9.500000 | 5.000000 |
| 50% | 7.900000 | 0.520000 | 0.260000 | 2.200000 | 0.079000 | 14.000000 | 38.000000 | 0.996750 | 3.310000 | 0.620000 | 10.200000 | 6.000000 |
| 75% | 9.200000 | 0.640000 | 0.420000 | 2.600000 | 0.090000 | 21.000000 | 62.000000 | 0.997835 | 3.400000 | 0.730000 | 11.100000 | 6.000000 |
| max | 15.900000 | 1.580000 | 1.000000 | 15.500000 | 0.611000 | 72.000000 | 289.000000 | 1.003690 | 4.010000 | 2.000000 | 14.900000 | 8.000000 |

In [19]:

#install statistic for the first time

In [20]:

import statistics as st

print('Mean dari Fixed Acidity : ', st.mean(dataset_red['fixed acidity']))
print('Q1 dari Fixed Acidity : ', st.median_low(dataset_red['fixed acidity']))
print('Median dari Fixed Acidity : ', st.median(dataset_red['fixed acidity']))
print('Q3 dari Fixed Acidity : ', st.median_high(dataset_red['fixed acidity']))
print('Modus dari Fixed Acidity : ', st.mode(dataset_red['fixed acidity']))
print('Standard Deviation dari Fixed Acidity : ', st.stdev(dataset_red['fixed acidity']))
print('Variance dari Fixed Acidity : ', st.variance(dataset_red['fixed acidity']))

Mean dari Fixed Acidity : 8.31963727329581
Q1 dari Fixed Acidity : 7.9
Median dari Fixed Acidity : 7.9
Q3 dari Fixed Acidity : 9.2
Modus dari Fixed Acidity : 7.2
Standard Deviation dari Fixed Acidity : 1.7410963181276955
Variance dari Fixed Acidity : 3.031416388997817

Challenge

```
print('Variable dari Fixed Acidity : ', st.variance(dataset_red['fixed acidity']))
```

```
Mean dari Fixed Acidity : 8.3196327329581
Q1 dari Fixed Acidity : 7.9
Median dari Fixed Acidity : 7.9
Q3 dari Fixed Acidity : 7.9
Modus dari Fixed Acidity : 7.2
Standard deviation dari Fixed Acidity : 1.7410963181276955
Variance dari Fixed Acidity : 3.03416388997817
```

Challenge

```
In [21]: #importing whitewine dataset
```

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|--|---------------|----------|------------------|----------|-------------|----------|----------------|----------|-----------|----------|---------------------|-----------|----------------------|------------|---------|----------|----|----------|-----------|----------|---------|-----------|---------|----------|--------|---------|
| In [23]: | <pre>#Menampilkan hasil mean dari salah satu kolom print('mean dari density : ', dataset_white['density'].mean())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | mean dari density : 0.9940273764801896 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [24]: | <pre>#Menampilkan mean dari keseluruhan data print(dataset_white.mean())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>0.654788</td></tr><tr><td>volatile acidity</td><td>0.278241</td></tr><tr><td>citric acid</td><td>0.334192</td></tr><tr><td>residual sugar</td><td>6.391415</td></tr><tr><td>chlorides</td><td>0.045772</td></tr><tr><td>free sulfur dioxide</td><td>35.308085</td></tr><tr><td>total sulfur dioxide</td><td>138.360657</td></tr><tr><td>density</td><td>0.994027</td></tr><tr><td>pH</td><td>3.188267</td></tr><tr><td>sulphates</td><td>0.489847</td></tr><tr><td>alcohol</td><td>10.514267</td></tr><tr><td>quality</td><td>5.877909</td></tr><tr><td>dtype:</td><td>float64</td></tr></tbody></table> | fixed acidity | 0.654788 | volatile acidity | 0.278241 | citric acid | 0.334192 | residual sugar | 6.391415 | chlorides | 0.045772 | free sulfur dioxide | 35.308085 | total sulfur dioxide | 138.360657 | density | 0.994027 | pH | 3.188267 | sulphates | 0.489847 | alcohol | 10.514267 | quality | 5.877909 | dtype: | float64 |
| fixed acidity | 0.654788 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.278241 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.334192 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 6.391415 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.045772 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 35.308085 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 138.360657 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.994027 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 3.188267 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.489847 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 10.514267 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 5.877909 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| dtype: | float64 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Median | | | | | | | | | | | | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|---------------|----------|------------------|----------|-------------|----------|----------------|-----------|-----------|----------|---------------------|------------|----------------------|-------------|---------|----------|------|----------|-----------|----------|---------|----------|---------|-----------|---------|----------|--------|---------|
| In [25]: | <pre>#Menampilkan median salah satu kolom print('median dari density : ', dataset_white['density'].median())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | median dari density : 0.99374 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [26]: | <pre>#Menampilkan median semua data print(dataset_white.median())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>0.643868</td></tr><tr><td>volatile acidity</td><td>0.160795</td></tr><tr><td>citric acid</td><td>0.014646</td></tr><tr><td>residual sugar</td><td>25.725770</td></tr><tr><td>chlorides</td><td>0.008477</td></tr><tr><td>free sulfur dioxide</td><td>289.242729</td></tr><tr><td>total sulfur dioxide</td><td>1806.085491</td></tr><tr><td>density</td><td>0.990899</td></tr><tr><td>pH</td><td>0.022801</td></tr><tr><td>sulphates</td><td>0.013925</td></tr><tr><td>alcohol</td><td>0.470000</td></tr><tr><td>alcohol</td><td>10.400000</td></tr><tr><td>quality</td><td>6.000000</td></tr><tr><td>dtype:</td><td>float64</td></tr></tbody></table> | fixed acidity | 0.643868 | volatile acidity | 0.160795 | citric acid | 0.014646 | residual sugar | 25.725770 | chlorides | 0.008477 | free sulfur dioxide | 289.242729 | total sulfur dioxide | 1806.085491 | density | 0.990899 | pH | 0.022801 | sulphates | 0.013925 | alcohol | 0.470000 | alcohol | 10.400000 | quality | 6.000000 | dtype: | float64 |
| fixed acidity | 0.643868 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.160795 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.014646 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 25.725770 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.008477 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 289.242729 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 1806.085491 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.990899 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 0.022801 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.013925 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 0.470000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 10.400000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 6.000000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| dtype: | float64 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [27]: | <pre>#Jika data median yang ingin dicari adalah per row, dapat menggunakan attribut axis = 1 print(dataset_white.median(axis=1))</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>0</td><td>4.509000</td></tr><tr><td>1</td><td>2.450000</td></tr><tr><td>2</td><td>4.630000</td></tr><tr><td>3</td><td>4.595000</td></tr><tr><td>4</td><td>4.595000</td></tr><tr><td>...</td><td>...</td></tr><tr><td>4893</td><td>2.450000</td></tr><tr><td>4894</td><td>4.075000</td></tr><tr><td>4895</td><td>2.095000</td></tr><tr><td>4896</td><td>2.220000</td></tr><tr><td>4897</td><td>2.124705</td></tr></tbody></table> Length: 4898, dtype: float64 | 0 | 4.509000 | 1 | 2.450000 | 2 | 4.630000 | 3 | 4.595000 | 4 | 4.595000 | ... | ... | 4893 | 2.450000 | 4894 | 4.075000 | 4895 | 2.095000 | 4896 | 2.220000 | 4897 | 2.124705 | | | | | | |
| 0 | 4.509000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 1 | 2.450000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 2 | 4.630000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | 4.595000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4 | 4.595000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| ... | ... | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4893 | 2.450000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4894 | 4.075000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4895 | 2.095000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4896 | 2.220000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 4897 | 2.124705 | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Modus

| | | | | | |
|----------|-------------------------------|-----------|---------------|---------|--|
| | 8 | ph | 4898 non-null | float64 | |
| | 9 | sulphates | 4898 non-null | float64 | |
| | 10 | alcohol | 4898 non-null | float64 | |
| | 11 | quality | 4898 non-null | int64 | |
| | dtypes: float64(11), int64(1) | | | | |
| | memory usage: 469.5 KB | | | | |
| Out[22]: | (4898, 12) | | | | |

Mean

```
In [23]: #Menampilkan hasil mean dari salah satu kolom
print('mean dari density : ', dataset_white['density'].mean())
```

Standard Deviation

| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|--|---------------|----------|------------------|----------|-------------|----------|----------------|----------|-----------|----------|---------------------|-----------|----------------------|-----------|---------|----------|----|----------|-----------|----------|---------|----------|---------|----------|--------|---------|
| In [30]: | <pre>#Menampilkan Standard Deviation salah satu kolom print('Standard deviation dari density : ', dataset_white['density'].std())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Standard deviation dari density : 0.0029909069169369354 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| In [31]: | <pre>#Menampilkan Standard Deviation semua data print(dataset_white.std())</pre> | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>0.843868</td></tr><tr><td>volatile acidity</td><td>0.160795</td></tr><tr><td>citric acid</td><td>0.121620</td></tr><tr><td>residual sugar</td><td>5.072058</td></tr><tr><td>chlorides</td><td>0.021848</td></tr><tr><td>free sulfur dioxide</td><td>17.007137</td></tr><tr><td>total sulfur dioxide</td><td>42.498085</td></tr><tr><td>density</td><td>0.002991</td></tr><tr><td>pH</td><td>0.151001</td></tr><tr><td>sulphates</td><td>0.114126</td></tr><tr><td>alcohol</td><td>1.230621</td></tr><tr><td>quality</td><td>0.085039</td></tr><tr><td>dtype:</td><td>float64</td></tr></tbody></table> | fixed acidity | 0.843868 | volatile acidity | 0.160795 | citric acid | 0.121620 | residual sugar | 5.072058 | chlorides | 0.021848 | free sulfur dioxide | 17.007137 | total sulfur dioxide | 42.498085 | density | 0.002991 | pH | 0.151001 | sulphates | 0.114126 | alcohol | 1.230621 | quality | 0.085039 | dtype: | float64 |
| fixed acidity | 0.843868 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.160795 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.121620 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 5.072058 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.021848 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 17.007137 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 42.498085 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.002991 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 0.151001 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.114126 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 1.230621 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 0.085039 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| dtype: | float64 | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Variance | | | | | | | | | | | | | | | | | | | | | | | | | | |

| | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|---------------|----------|------------------|----------|-------------|----------|----------------|-----------|-----------|----------|---------------------|------------|----------------------|-------------|---------|----------|----|----------|-----------|----------|---------|----------|---------|----------|
| In [32]: | <pre>#Menampilkan Variance Salah satu kolom print('variance dari density : ', dataset_white['density'].var())</pre> | | | | | | | | | | | | | | | | | | | | | | | | |
| | variance dari density : 8.945524185781283e-06 | | | | | | | | | | | | | | | | | | | | | | | | |
| In [33]: | <pre>#Menampilkan Variance semua data print(dataset_white.var())</pre> | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tbody><tr><td>fixed acidity</td><td>0.712114</td></tr><tr><td>volatile acidity</td><td>0.030160</td></tr><tr><td>citric acid</td><td>0.014646</td></tr><tr><td>residual sugar</td><td>25.725770</td></tr><tr><td>chlorides</td><td>0.008477</td></tr><tr><td>free sulfur dioxide</td><td>289.242729</td></tr><tr><td>total sulfur dioxide</td><td>1806.085491</td></tr><tr><td>density</td><td>0.000899</td></tr><tr><td>pH</td><td>0.022801</td></tr><tr><td>sulphates</td><td>0.013925</td></tr><tr><td>alcohol</td><td>1.514427</td></tr><tr><td>quality</td><td>0.784356</td></</tr></tbody></table> | fixed acidity | 0.712114 | volatile acidity | 0.030160 | citric acid | 0.014646 | residual sugar | 25.725770 | chlorides | 0.008477 | free sulfur dioxide | 289.242729 | total sulfur dioxide | 1806.085491 | density | 0.000899 | pH | 0.022801 | sulphates | 0.013925 | alcohol | 1.514427 | quality | 0.784356 |
| fixed acidity | 0.712114 | | | | | | | | | | | | | | | | | | | | | | | | |
| volatile acidity | 0.030160 | | | | | | | | | | | | | | | | | | | | | | | | |
| citric acid | 0.014646 | | | | | | | | | | | | | | | | | | | | | | | | |
| residual sugar | 25.725770 | | | | | | | | | | | | | | | | | | | | | | | | |
| chlorides | 0.008477 | | | | | | | | | | | | | | | | | | | | | | | | |
| free sulfur dioxide | 289.242729 | | | | | | | | | | | | | | | | | | | | | | | | |
| total sulfur dioxide | 1806.085491 | | | | | | | | | | | | | | | | | | | | | | | | |
| density | 0.000899 | | | | | | | | | | | | | | | | | | | | | | | | |
| pH | 0.022801 | | | | | | | | | | | | | | | | | | | | | | | | |
| sulphates | 0.013925 | | | | | | | | | | | | | | | | | | | | | | | | |
| alcohol | 1.514427 | | | | | | | | | | | | | | | | | | | | | | | | |
| quality | 0.784356 | | | | | | | | | | | | | | | | | | | | | | | | |