PYTHON

```python
from random import shuffle

import pandas as pd

import matplotlib.pyplot as plt

from sklearn.model_selection import train_test_split

df=pd.read_csv("pollution.csv")

print(df)

AQI_value=df["AQI"]

AQI_value.plot(kind='hist')

df.plot(x="year",y="AQI",kind="scatter")


train, test=train_test_split(df,test_size=0.2,random_state=33,shuffle=True)

print(test)

print(train)
```
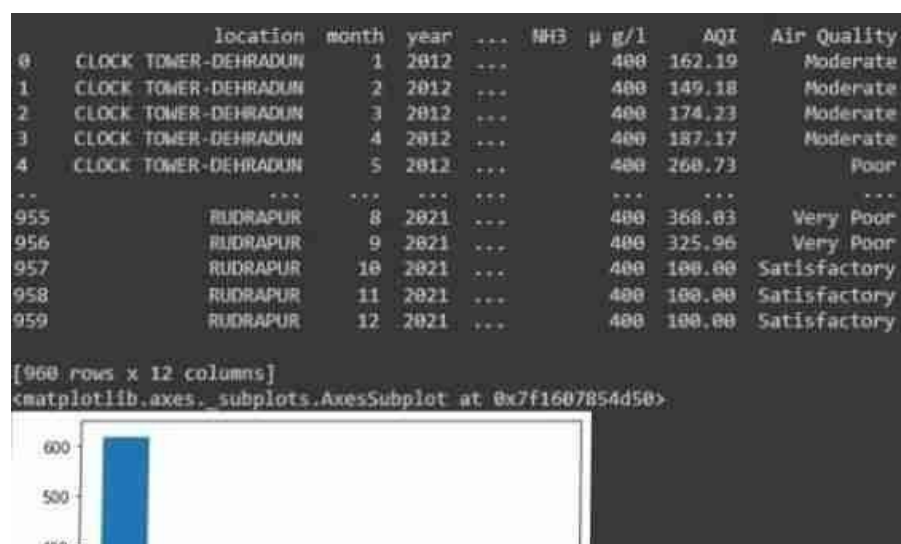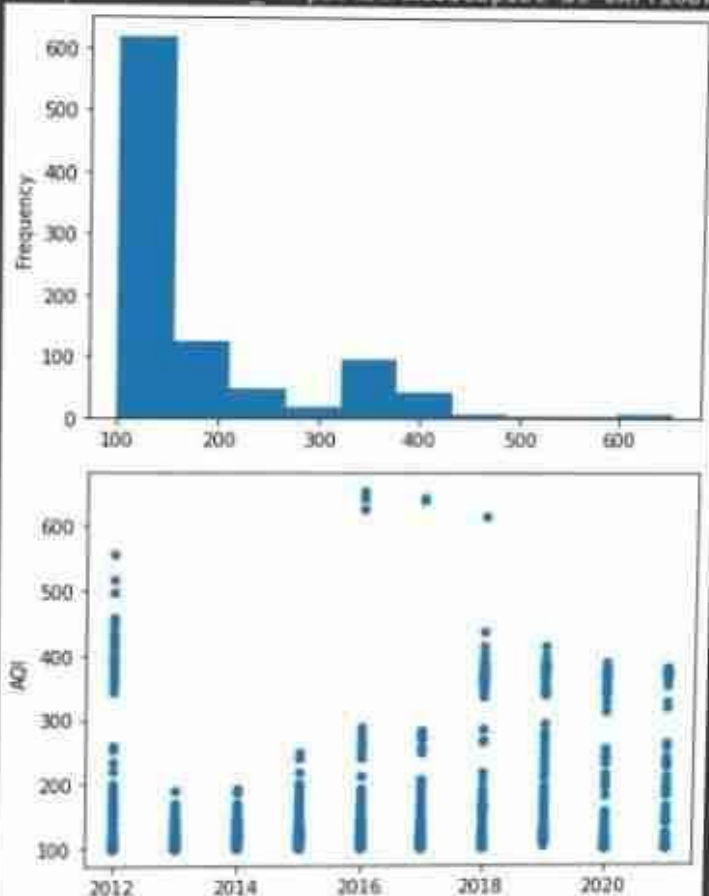
```
                    location  month  year  ...  NH3  µ g/l      AQI  Air Quality
0     CLOCK TOWER-DEHRADUN        1  2012  ...         400  162.19     Moderate
1     CLOCK TOWER-DEHRADUN        2  2012  ...         400  149.18     Moderate
2     CLOCK TOWER-DEHRADUN        3  2012  ...         400  174.23     Moderate
3     CLOCK TOWER-DEHRADUN        4  2012  ...         400  187.17     Moderate
4     CLOCK TOWER-DEHRADUN        5  2012  ...         400  260.73         Poor
..              ...            ...   ...   ...         ...    ...           ...
955               RUDRAPUR        8  2021  ...         400  368.03    Very Poor
956               RUDRAPUR        9  2021  ...         400  325.96    Very Poor
957               RUDRAPUR       10  2021  ...         400  100.00  Satisfactory
958               RUDRAPUR       11  2021  ...         400  100.00  Satisfactory
959               RUDRAPUR       12  2021  ...         400  100.00  Satisfactory

[960 rows x 12 columns]
<matplotlib.axes._subplots.AxesSubplot at 0x7f1607854d50>
```

```
                location  month  year  ...  NH3  μ g/l     AQI    Air Quality
860                RUDRAPUR     9  2013  ...            400  102.49        Moderate
756                 KASHIPUR     1  2015  ...            400  142.72        Moderate
479  RISHIKESH-NAGARNIGAM    12  2021  ...            400  100.00    Satisfactory
841                RUDRAPUR     2  2012  ...            400  156.77        Moderate
684                HALDWANI     1  2019  ...            400  349.28       Very Poor
..                    ...   ...   ...  ...            ...     ...            ...
834                 KASHIPUR     7  2021  ...            400  370.07       Very Poor
554         SIDCUL-HARIDWAR     3  2018  ...            400  109.99        Moderate
446  RISHIKESH-NAGARNIGAM     3  2019  ...            400  121.11        Moderate
767                 KASHIPUR    12  2015  ...            400  125.12        Moderate
896                RUDRAPUR     9  2016  ...            400  100.00    Satisfactory

[192 rows x 12 columns]
                location  month  year  ...  NH3  μ g/l     AQI    Air Quality
858                RUDRAPUR     7  2013  ...            400  100.00    Satisfactory
771                 KASHIPUR     4  2016  ...            400  131.37        Moderate
234  RAIPUR ROAD DEHRADUN     7  2021  ...            400  145.47        Moderate
781                 KASHIPUR     2  2017  ...            400  114.97        Moderate
149  RAIPUR ROAD DEHRADUN     6  2014  ...            400  135.86        Moderate
```