Assignment 1

Name: Harsukh Singh Sagri

Roll number: 210428

Problem Statement

Q.) Air pollution measurements available via multi-sensory system are PM10, PM2.5, SO2, NO, NO2, NOx, CO, NH3, O3 and Benzene.

Now, given **dataset** depicts levels of these pollutants at given time intervals spread over a duration of approx. 3 months. On the basis of information provided through dataset, answer the following:

- **(a)Classification:** time series data classify mainly two categories: stock and flow time series. Find out if it identifies and assigns categories to the air pollution data.
- **(b)Curve fitting:** Plot the air pollution data along a curve to study the relationships of variables within the data.
- **(c)Descriptive analysis:** in coal India blasting effect time is 13:45 pm to 14:45 pm, find trends like cycles, or seasonal variation.
- **(d) Explanatory analysis:** The aim of this analysis is to gain insights into the air pollution data and explore the relationships within the data. Additionally, we aim to understand the impact of coal blasting on air pollution in the context of coal India.

- **(e)Forecasting:** Analyse the time series method used for forecasting i.e., which model (AR, MA, ARMA, ARIMA) needs to be used for forecasting?

-->Firstly, we will start by importing all the required libraries for our analysis.

In [1]:
```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.metrics import mean_squared_error
from math import sqrt
from datetime import datetime
```

In [2]:
```
dataSet = pd.read_csv("Data.csv")
```

- `numpy` (imported as `np`): It is a popular numerical computing library in Python.
- `pandas` (imported as `pd`): It offers data structures and functions for efficiently handling and analysing structured data, such as data frames.
- `seaborn` (imported as `sns`): It is a data visualization library built on top of `matplotlib`.
- `matplotlib.pyplot` (imported as `plt`): It is a plotting library in Python that provides a wide range of functions for creating various types of plots, including line plots, scatter plots, bar plots, histograms, etc.
- `sklearn.metrics.mean_squared_error`: It calculates the mean squared error (MSE) between two arrays.
- `math.sqrt`: It is a function from the `math` module in Python. It calculates the square root of a number.

- **`datetime.datetime`**: It is a class from the **`datetime`** module in Python. In this case, it is imported to work with date and time-related operations.

```
dataSet
```

	#	From	To (Interval: 15M)	Singrauli, Surya Kiran Bhawan Dudhichua PM10 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua PM2.5 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NO (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NO2 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NOX (ppb)	Singrauli, Surya Kiran Bhawan Dudhichua CO (mg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua SO2 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NH3 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua Ozone (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua Benzene (µg/m3)
0	1	01-02-2023 00:00	01-02-2023 00:15	95.0	35.0	NaN	90.1	56.2	0.31	NaN	17.7	28.1	0.4
1	2	01-02-2023 00:15	01-02-2023 00:30	95.0	35.0	NaN	88.0	55.1	0.33	NaN	18.3	27.1	0.4
2	3	01-02-2023 00:30	01-02-2023 00:45	95.0	35.0	NaN	87.7	55.2	0.38	NaN	19.7	24.9	0.4
3	4	01-02-2023 00:45	01-02-2023 01:00	122.0	34.0	NaN	88.9	55.7	0.38	NaN	21.3	21.9	0.4
4	5	01-02-2023 01:00	01-02-2023 01:15	122.0	34.0	NaN	90.0	55.8	0.38	NaN	22.3	16.7	0.4

-->In our original dataset, there are many missing values in our data set. Now, we want to estimate the percentage of such missing values and the fill them by using interpolation.

```
columns=dataSet.columns
date = columns[2]
columns=columns[3:]
columns
```

```
Index(['Singrauli, Surya Kiran Bhawan Dudhichua  PM10 (µg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  PM2.5 (µg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  NO (µg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  NO2 (µg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  NOX (ppb)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  CO (mg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  SO2 (µg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  NH3 (µg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  Ozone (µg/m3)',
       'Singrauli, Surya Kiran Bhawan Dudhichua  Benzene (µg/m3)'],
      dtype='object')
```

-->Now, we have eliminated 'Serial number', 'From' and 'To' columns and stored pollutant level columns.

```
In [5]:  for col in columns:
             print((dataSet[col].isna().sum() / len(dataSet[col])) * 100)
```

```
19.45601851851852
2.6157407407407405
15.844907407407408
4.814814814814815
4.8032407407407405
5.7407407407407405
16.79398148148148
3.7731481481481484
5.243055555555555
71.70138888888889
```

-->Now, we have calculated percentage of missing values in each of the columns.

Q1) Updating NA Values By Interpolation (Cubic):

```
In [6]:  for column in columns:
             x = pd.Series(dataSet[column])
             dataSet[column] = x.interpolate(limit_direction='both', kind='cubic')
```

-->Code updates the missing values of the current column in the **dataSet** Data Frame with the interpolated values.

-->Replacing NA values with 0 can distort the representation of the data. Depending on the context, assigning a value of 0 to missing data points can introduce bias or inaccurately portray the actual measurements.

By using cubic interpolation, the missing values are replaced with estimated values based on the surrounding data points, following a cubic interpolation curve. This helps to retain the overall shape and characteristics of the data while filling in the gaps caused by missing values.

```
In [7]: dataSet.describe()
```

Out[7]:

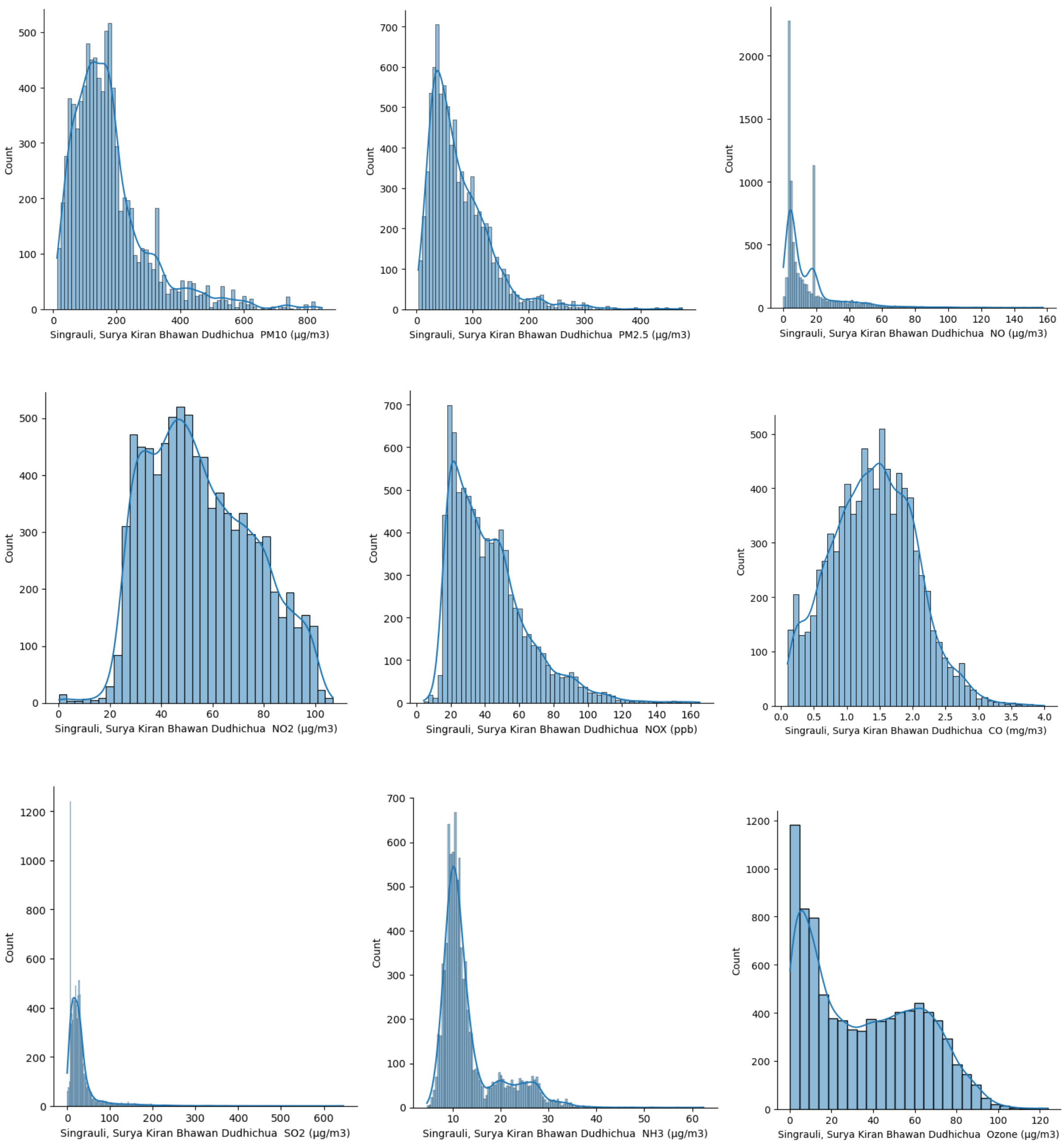
	#	Singrauli, Surya Kiran Bhawan Dudhichua PM10 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua PM2.5 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NO (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NO2 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NOX (ppb)	Singrauli, Surya Kiran Bhawan Dudhichua CO (mg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua SO2 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua NH3 (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua Ozone (µg/m3)	Singrauli, Surya Kiran Bhawan Dudhichua Benzene (µg/m3)
count	8640.000000	8640.000000	8640.000000	8640.000000	8640.000000	8640.000000	8640.000000	8640.000000	8640.000000	8640.000000	8640.000000
mean	4320.500000	177.463079	75.557350	14.940208	55.430689	42.328802	1.401927	31.923270	13.286956	35.193970	0.122002
std	2494.297496	124.773568	54.826850	17.862679	20.205531	22.181225	0.633087	39.028371	6.171697	26.867859	0.063168
min	1.000000	12.000000	3.000000	0.100000	0.200000	4.200000	0.100000	0.100000	4.600000	0.100000	0.100000
25%	2160.750000	97.000000	36.821429	4.000000	39.300000	24.900000	0.950000	12.300000	9.500000	10.300000	0.100000
50%	4320.500000	151.900794	61.000000	7.500000	52.800000	37.500000	1.410000	22.800000	11.000000	31.700000	0.100000
75%	6480.250000	215.000000	101.000000	18.100000	70.700000	53.200000	1.850000	33.400000	14.000000	58.100000	0.100000
max	8640.000000	847.000000	474.000000	157.500000	106.900000	165.200000	4.000000	645.600000	62.400000	123.800000	0.600000

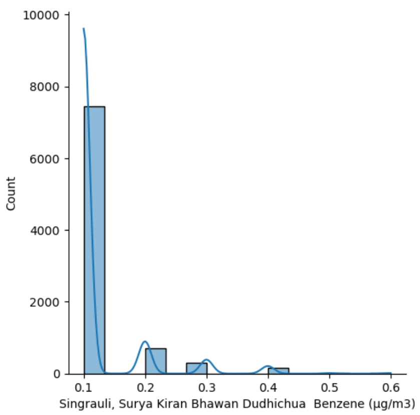
-->Clearly, we have filled all the missing values by using cubic interpolation as can be seen from updated dataset.

Plotting Frequency Distribution Curves:

```
In [8]: for column in columns:
            sns.displot(dataSet[column], kde = True)
            plt.show()
```

-->Using the given code, we have plotted frequency distribution curves of various pollutant levels.





Frequency Distribution Curves

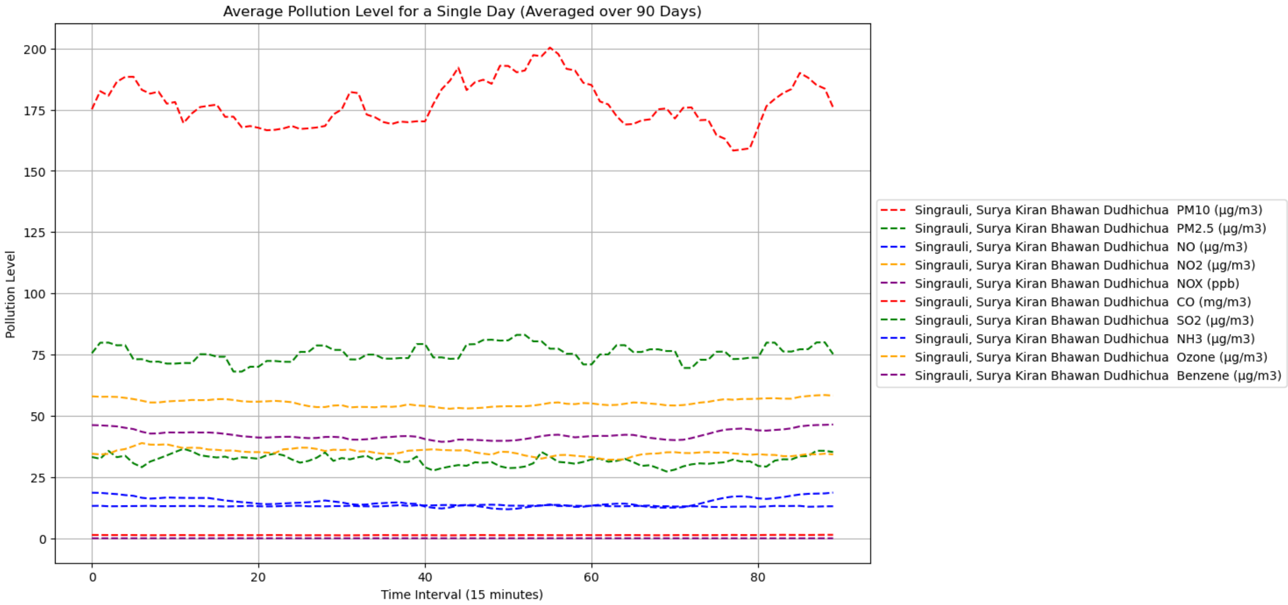
--><u>Classification (Flow Time Series Data) Q(3a),</u>
<u>Explanatory Analysis Q(3d) and</u>
<u>Statistical Inference (QQ Plot and Blasting Time) Q(2):</u>

-->Results from graph plotted below indicate that *blasting time of coal in India is most likely occurring between 13:45 to 14:45 pm.*

```
avg_level = dataSet.groupby(dataSet.index % 90).mean(numeric_only=True)
```

```
plt.figure(figsize=(12, 8))

colors = ['red', 'green', 'blue', 'orange', 'purple']

for i, column in enumerate(columns):
    plt.plot(avg_level[column], label=column, color=colors[i % len(colors)], linestyle='--')

plt.xlabel('Time Interval (15 minutes)')
plt.ylabel('Pollution Level')
plt.title('Average Pollution Level for a Single Day (Averaged over 90 Days)')

plt.legend(loc='center left', bbox_to_anchor=(1, 0.5))

plt.grid(True)
plt.show()
```

Plotted Graph (Flow Time Series):



The plotted graph shows the average pollution levels for different gases during a single day, averaged over the 90-day period.

The pollutants' levels exhibit noticeable increases specifically during and after the time interval of 13:45 to 14:45(somewhere between 50th and 60th interval), suggesting a correlation with the blasting activities of coal in the region.

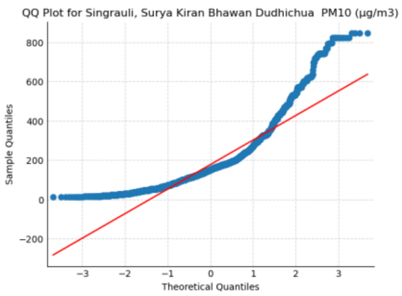
The graph provides a clear visual representation of the heightened pollution levels during and after the suspected blasting time. This observation supports the conclusion that the blasting time of coal in India is most likely occurring between 13:45 to 14:45 pm.

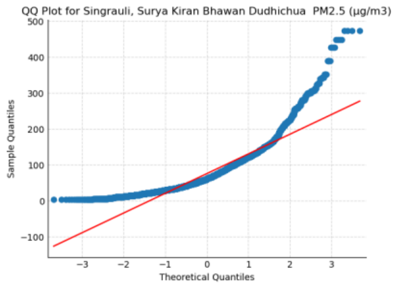
This inference is based on the analysis of average pollution levels and spikes during the specified time interval.

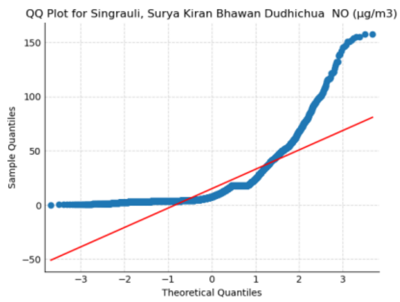
This indicates level of most pollutant gases spikes after blasting time of coal in India starts (i.e 13:45 pm to 14:45 pm).

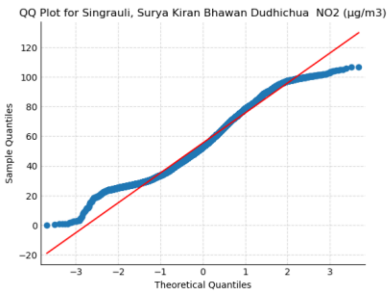
QQ Plots (Statistical Inference Q2):

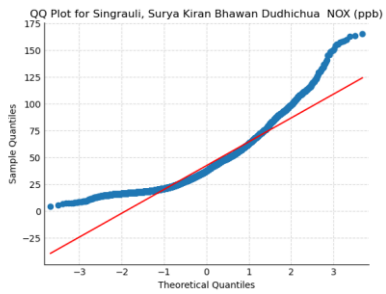
```python
import statsmodels.api as sm
colors = ['red', 'green', 'blue', 'orange', 'purple']  # List of colors for each pollutant

for i, column in enumerate(columns):
    fig, ax = plt.subplots()
    sm.qqplot(dataSet[column], line='s', ax=ax)
    ax.set_title(f"QQ Plot for {column}")
    ax.spines['top'].set_visible(False)  # Remove top spine
    ax.spines['right'].set_visible(False)  # Remove right spine
    ax.grid(True, linestyle='--', alpha=0.5)  # Add grid lines
    ax.tick_params(axis='both', which='both', length=0)  # Remove tick marks
    ax.set_xlabel('Theoretical Quantiles')
    ax.set_ylabel('Sample Quantiles')
    ax.lines[0].set_color(colors[i % len(colors)])  # Set color for QQ plot line
    plt.show()
```

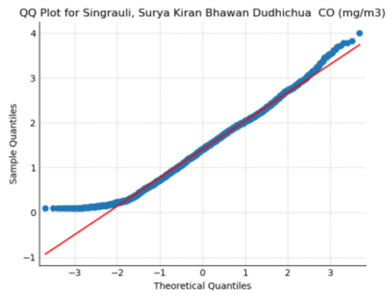


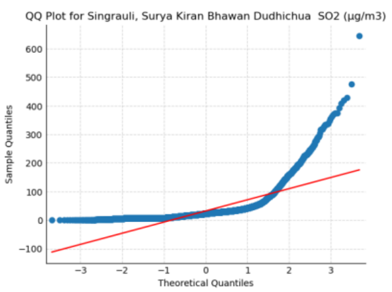


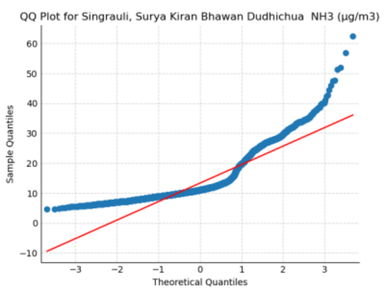


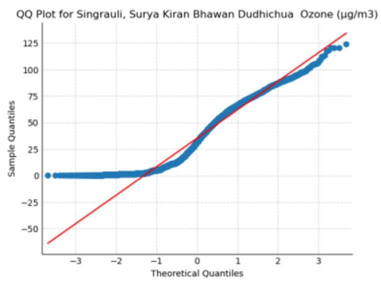


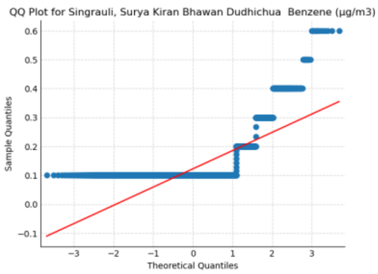












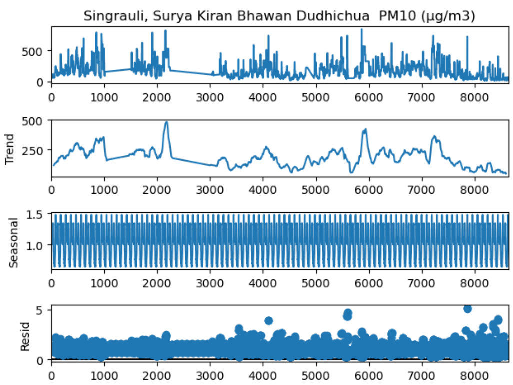
As we can infer from QQ plots, in general, the curves are positively skewed and *none of the curves resembles normal distribution.*

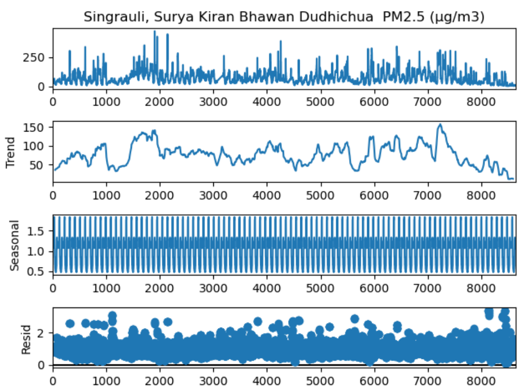
-->Descriptive Analysis (Trend and Seasonality) Q3(c):

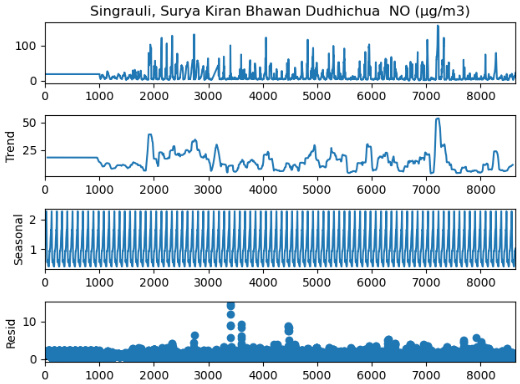
I have used below code to decompose a time series into its trend, seasonal, and residual components.

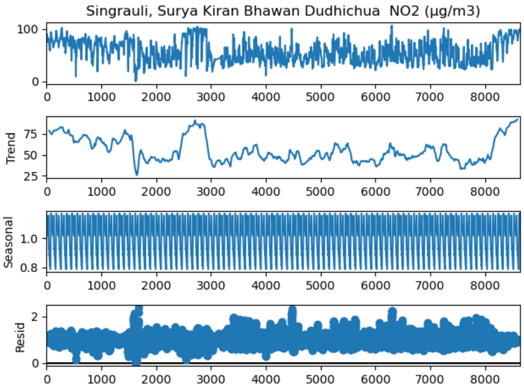
```
from statsmodels.tsa.seasonal import seasonal_decompose
for column in columns:
    decompose_result_mult = seasonal_decompose(dataSet[column], period = 96, model="multiplicative")
    plt.figure(figsize=(30, 50))
    trend = decompose_result_mult.trend
    seasonal = decompose_result_mult.seasonal
    decompose_result_mult.plot()
```

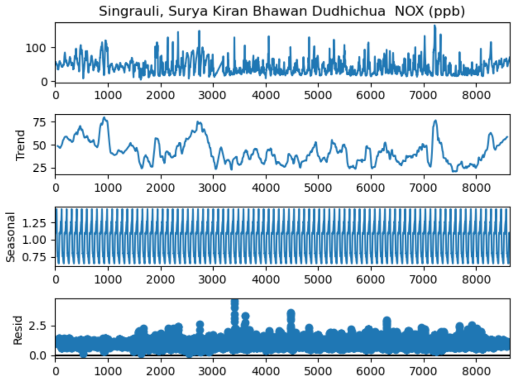
-->Following are the plots of all the 3 components for all the pollutants:

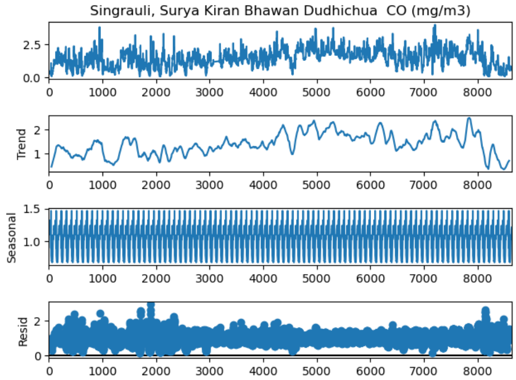


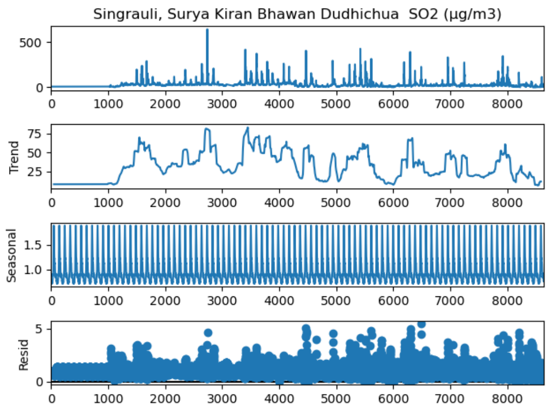


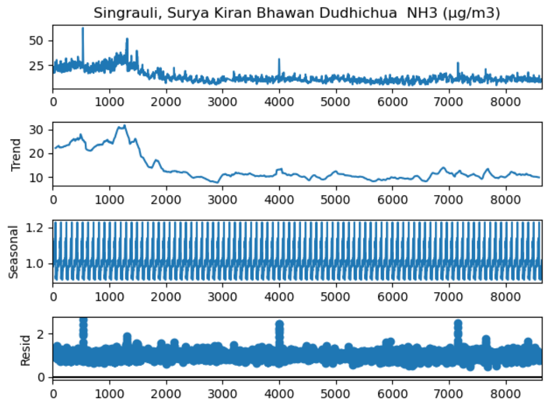

Singrauli, Surya Kiran Bhawan Dudhichua NO (µg/m3)
Trend
Seasonal
Resid

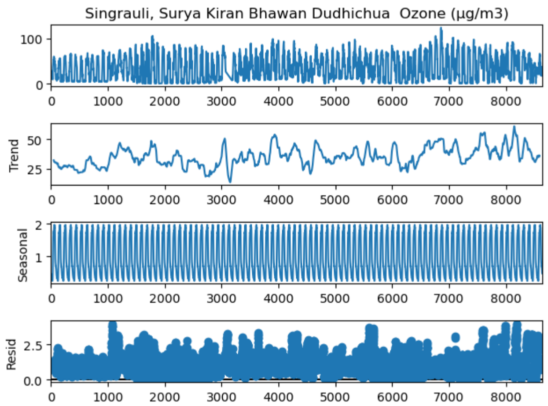

Singrauli, Surya Kiran Bhawan Dudhichua NO2 (µg/m3)
Trend
Seasonal
Resid

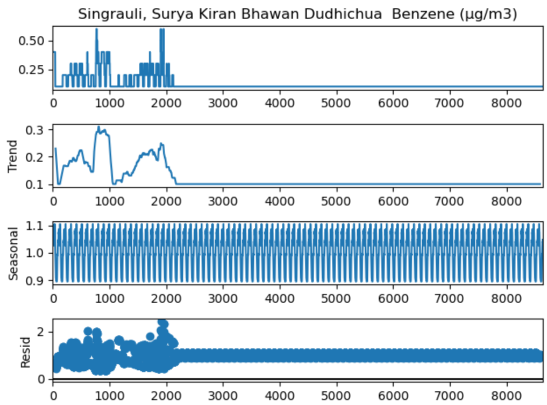

Singrauli, Surya Kiran Bhawan Dudhichua NOX (ppb)
Trend
Seasonal
Resid


Singrauli, Surya Kiran Bhawan Dudhichua CO (mg/m3)
Trend
Seasonal
Resid


Singrauli, Surya Kiran Bhawan Dudhichua SO2 (µg/m3)
Trend
Seasonal
Resid


Singrauli, Surya Kiran Bhawan Dudhichua NH3 (µg/m3)
Trend
Seasonal
Resid


Singrauli, Surya Kiran Bhawan Dudhichua Ozone (µg/m3)
Trend
Seasonal
Resid


Singrauli, Surya Kiran Bhawan Dudhichua Benzene (µg/m3)
Trend
Seasonal
Resid

--> Trend: By looking at these graphs, we can conclude that since trend graph is not continuously increasing or decreasing, *there is no trend in level of pollutant gases.*

-->Seasonality: Since duration in which data is available is 3 months and seasonality graphs repeat in very small duration of time, we can conclude *there is no seasonality in level of pollutant gases.*

--> <u>ADF Test For Seasonality and Trend: Q3(C)</u>

```
from statsmodels.tsa.stattools import adfuller
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
for column in columns:
    result = adfuller(dataSet[column])
    print('ADF Statistic: %f' % result[0])
    print('p-value: %f' % result[1])
    print('No. of Lags : %f' % result[2])
    print('No of Observation used for ADF regression and Critical Value Prediction : %f' % result[3])
    print('Critical Values:')
    for key, value in result[4].items():
        print('\t%s: %.3f' % (key, value))
```

-->In our context of analysing air pollution data and its relationship with coal blasting, the ADF test can help us assess the presence of trend and seasonality in the data.

In the ADF test, the null hypothesis assumes that the time series has a unit root, indicating that it is non-stationary and contains a trend. The alternative hypothesis suggests that the time series is stationary, meaning it does not contain a trend.

Below is the result of code:

```
ADF Statistic: -9.023314
p-value: 0.000000
No. of Lags : 36.000000
No of Observation used for ADF regression and Critical Value Prediction : 8603.000000
Critical Values:
	1%: -3.431
	5%: -2.862
	10%: -2.567
ADF Statistic: -11.159054
p-value: 0.000000
No. of Lags : 36.000000
No of Observation used for ADF regression and Critical Value Prediction : 8603.000000
Critical Values:
	1%: -3.431
	5%: -2.862
	10%: -2.567
ADF Statistic: -14.799960
p-value: 0.000000
No. of Lags : 11.000000
No of Observation used for ADF regression and Critical Value Prediction : 8628.000000
Critical Values:
	1%: -3.431
	5%: -2.862
	10%: -2.567
ADF Statistic: -9.181725
p-value: 0.000000
No. of Lags : 21.000000
No of Observation used for ADF regression and Critical Value Prediction : 8618.000000
Critical Values:
	1%: -3.431
	5%: -2.862
	10%: -2.567
```

```
ADF Statistic: -12.734220
p-value: 0.000000
No. of Lags : 24.000000
No of Observation used for ADF regression and Critical Value Prediction : 8615.000000
Critical Values:
	 1%: -3.431
	 5%: -2.862
	 10%: -2.567
ADF Statistic: -9.976113
p-value: 0.000000
No. of Lags : 8.000000
No of Observation used for ADF regression and Critical Value Prediction : 8631.000000
Critical Values:
	 1%: -3.431
	 5%: -2.862
	 10%: -2.567
ADF Statistic: -14.141881
p-value: 0.000000
No. of Lags : 20.000000
No of Observation used for ADF regression and Critical Value Prediction : 8619.000000
Critical Values:
	 1%: -3.431
	 5%: -2.862
	 10%: -2.567
ADF Statistic: -3.059899
p-value: 0.029667
No. of Lags : 33.000000
No of Observation used for ADF regression and Critical Value Prediction : 8606.000000
Critical Values:
	 1%: -3.431
	 5%: -2.862
	 10%: -2.567
```

```
ADF Statistic: -20.958946
p-value: 0.000000
No. of Lags : 34.000000
No of Observation used for ADF regression and Critical Value Prediction : 8605.000000
Critical Values:
	 1%: -3.431
	 5%: -2.862
	 10%: -2.567
ADF Statistic: -9.037111
p-value: 0.000000
No. of Lags : 37.000000
No of Observation used for ADF regression and Critical Value Prediction : 8602.000000
Critical Values:
	 1%: -3.431
	 5%: -2.862
	 10%: -2.567
```

-->Based on the ADF test results, we can make the following conclusions:

- p-value: The p-value is significantly lower than the significance level (e.g., 0.05), indicating strong evidence against the null hypothesis. By rejecting the null hypothesis, we can conclude that the time series is stationary and does not exhibit a trend.

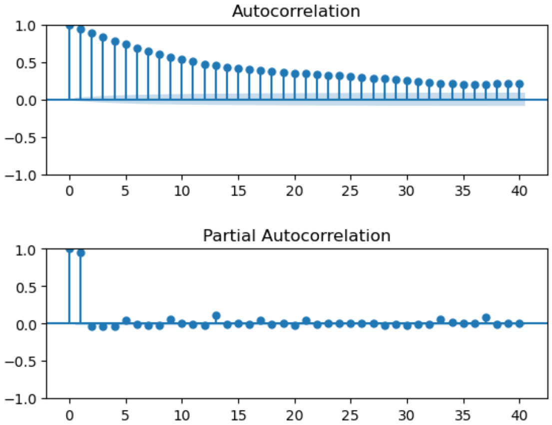
- ADF Statistic: The ADF statistic is significantly negative, lower than the critical values at 1%, 5%, and 10% significance levels. This further supports the rejection of the null hypothesis.

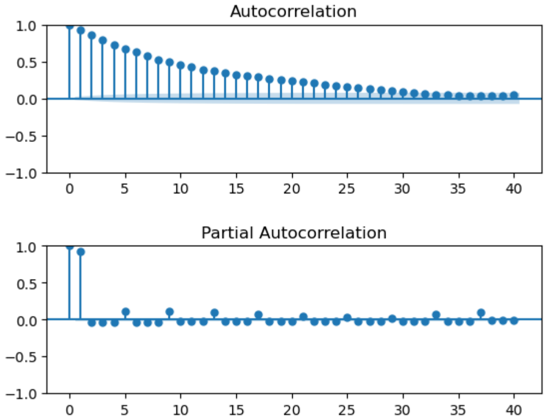
Based on these findings, we can confidently conclude that there is no presence of trend or seasonality in the pollutant levels.

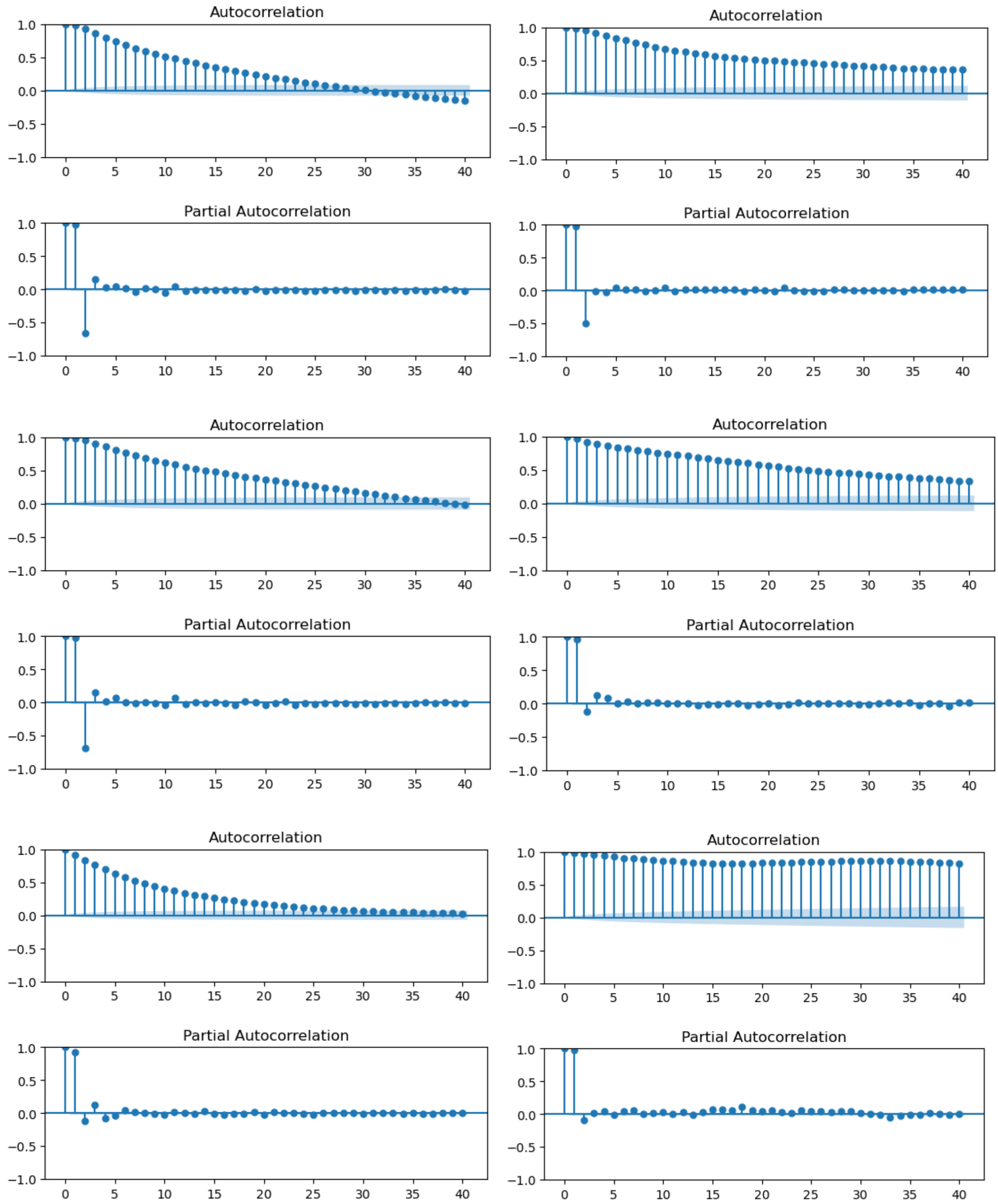
-->Forecasting Q3 (E):

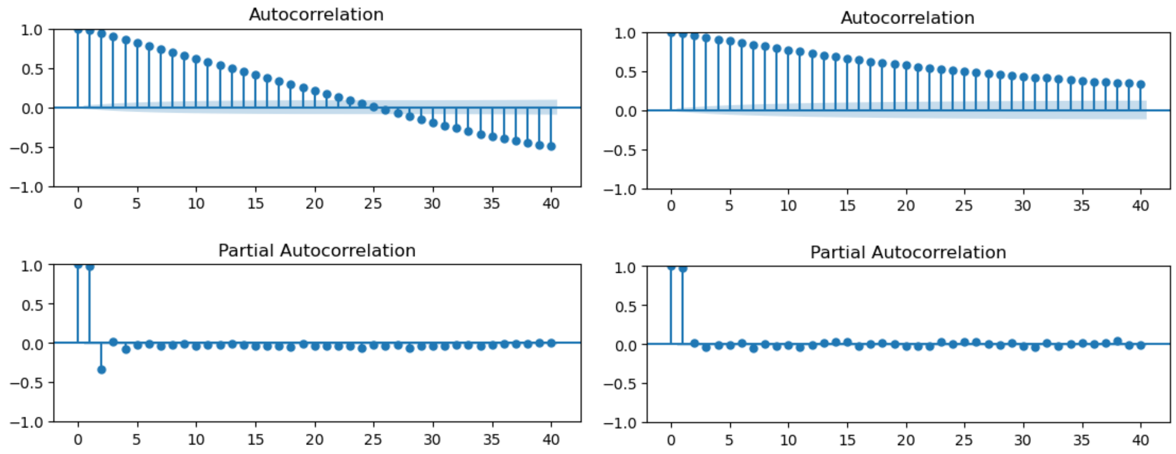
In [15]:
```
for column in columns:
    fig, (ax1, ax2) = plt.subplots(2)
    plot_acf(dataSet[column], ax = ax1)
    plot_pacf(dataSet[column], ax = ax2)
    plt.subplots_adjust(hspace=0.5)
    plt.show()
```

ACF (Autocorrelation) and PACF (Partial Autocorrelation) Plots:






Autocorrelation
Autocorrelation
Partial Autocorrelation
Partial Autocorrelation
Autocorrelation
Autocorrelation
Partial Autocorrelation
Partial Autocorrelation
Autocorrelation
Autocorrelation
Partial Autocorrelation
Partial Autocorrelation



Based on the autocorrelation function (ACF) and partial autocorrelation function (PACF) graphs, we can make the following observations:

- ACF: The ACF graph shows that the autocorrelation values do not decay fully after a certain lag. This indicates the presence of significant autocorrelation in the data. This behaviour is indicative of a time series that follows an autoregressive (AR) model.

- PACF: The PACF graph shows that the partial autocorrelation values decay after a lag of 2 or 3. This pattern is consistent with an **autoregressive model** with a lag of 2 or 3.

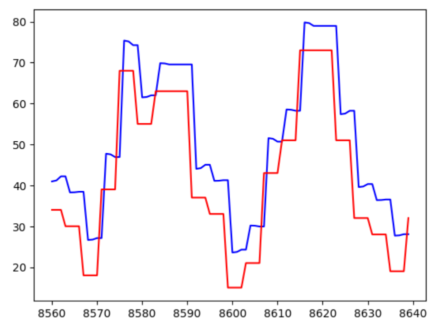
-->Based on these observations, we can infer that *the time series data follows an autoregressive model, specifically an AR (2) or AR (3) model.* The autocorrelation and partial autocorrelation patterns indicate the presence of a memory effect and direct influence of past values on the current value.

-->Autoregressive Model (AR): Stock Time Series Data Q3(A):

```python
from statsmodels.tsa.ar_model import AutoReg
for column in columns:
    data = dataSet[column]
    train_data = data[:-80]
    test_data = data[-80:]
    ar_model = AutoReg(data, lags = 3).fit()
    print(ar_model.summary())
    pred = ar_model.predict(start = len(train_data), end = len(data) - 1, dynamic=False)
    plt.plot(pred, color = "blue")
    plt.plot(test_data, color = "red")
    plt.show()
    rmse = sqrt(mean_squared_error(pred, test_data))
    mean = data.mean()
    print("Mean : %f" % mean)
    print("Root Mean Squared Error : %f" %rmse)
```

PM10

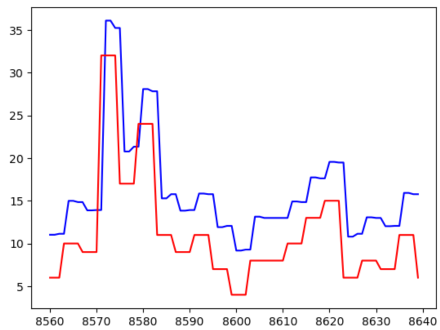
```
                              AutoReg Model Results
==============================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  PM10 (µg/m3)   No. Observations:         8640
Model:                                           AutoReg(3)   Log Likelihood          -44157.239
Method:                                     Conditional MLE   S.D. of innovations         40.190
Date:                                      Tue, 27 Jun 2023   AIC                      88324.479
Time:                                              18:10:56   BIC                      88359.798
Sample:                                                   3   HQIC                     88336.522
                                                       8640
=================================================================================================
                                                      coef    std err          z      P>|z|      [0.025      0.975]
-------------------------------------------------------------------------------------------------
const                                              10.1385      0.766     13.231      0.000       8.637      11.640
Singrauli, Surya Kiran Bhawan Dudhichua  PM10 (µg/m3).L1   0.9816   0.011   91.272   0.000   0.961   1.003
Singrauli, Surya Kiran Bhawan Dudhichua  PM10 (µg/m3).L2  -0.0075   0.015   -0.499   0.618  -0.037   0.022
Singrauli, Surya Kiran Bhawan Dudhichua  PM10 (µg/m3).L3  -0.0312   0.011   -2.905   0.004  -0.052  -0.010
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0660           +0.0000j            1.0660            0.0000
AR.2            4.8645           +0.0000j            4.8645            0.0000
AR.3           -6.1714           +0.0000j            6.1714            0.5000
-----------------------------------------------------------------------------
```




```
Mean : 177.463079
Root Mean Squared Error : 10.852239
```

PM2.5

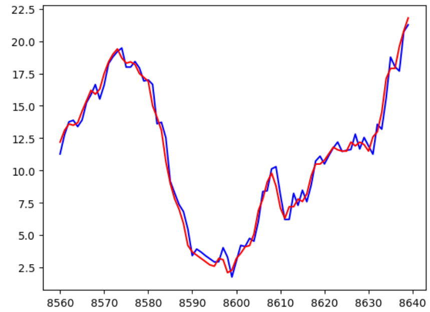
```
                              AutoReg Model Results
==============================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  PM2.5 (µg/m3)   No. Observations:        8640
Model:                                            AutoReg(3)   Log Likelihood         -38110.068
Method:                                      Conditional MLE   S.D. of innovations        19.955
Date:                                       Tue, 27 Jun 2023   AIC                     76230.136
Time:                                               18:10:56   BIC                     76265.455
Sample:                                                    3   HQIC                    76242.179
                                                        8640
================================================================================================
=
                                                          coef    std err          z      P>|z|      [0.025      0.97
5]
------------------------------------------------------------------------------------------------
-
const                                                   5.5796      0.374     14.901      0.000       4.846       6.31
3
Singrauli, Surya Kiran Bhawan Dudhichua  PM2.5 (µg/m3).L1      0.9635      0.011     89.609      0.000       0.942       0.98
5
Singrauli, Surya Kiran Bhawan Dudhichua  PM2.5 (µg/m3).L2 -2.203e-05      0.015     -0.001      0.999      -0.029       0.02
9
Singrauli, Surya Kiran Bhawan Dudhichua  PM2.5 (µg/m3).L3     -0.0374      0.011     -3.477      0.001      -0.058      -0.01
6
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0878           +0.0000j            1.0878            0.0000
AR.2            4.4441           +0.0000j            4.4441            0.0000
AR.3           -5.5325           +0.0000j            5.5325            0.5000
-----------------------------------------------------------------------------
```




```
Mean : 75.557350
Root Mean Squared Error : 6.033443
```

NO

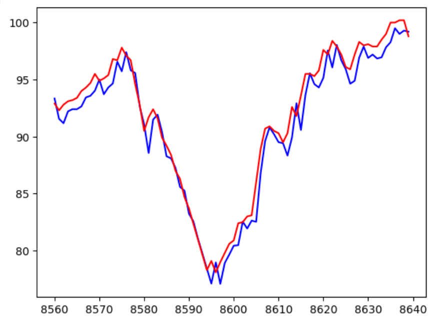
```
                                        AutoReg Model Results
==========================================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  NO (µg/m3)   No. Observations:             8640
Model:                                              AutoReg(3)   Log Likelihood              -21046.714
Method:                                        Conditional MLE   S.D. of innovations              2.767
Date:                                         Tue, 27 Jun 2023   AIC                          42103.428
Time:                                                 18:10:56   BIC                          42138.747
Sample:                                                      3   HQIC                         42115.471
                                                          8640
=================================================================================================================
                                                          coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------------------------
const                                                   0.4641      0.039     11.799      0.000       0.387       0.541
Singrauli, Surya Kiran Bhawan Dudhichua  NO (µg/m3).L1  1.7210      0.011    161.729      0.000       1.700       1.742
Singrauli, Surya Kiran Bhawan Dudhichua  NO (µg/m3).L2 -0.9002      0.019    -47.289      0.000      -0.937      -0.863
Singrauli, Surya Kiran Bhawan Dudhichua  NO (µg/m3).L3  0.1481      0.011     13.916      0.000       0.127       0.169
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0963           +0.0000j            1.0963            0.0000
AR.2            2.2758           +0.0000j            2.2758            0.0000
AR.3            2.7064           +0.0000j            2.7064            0.0000
-----------------------------------------------------------------------------
```




```
Mean : 14.940208
Root Mean Squared Error : 0.724277
```

NO2

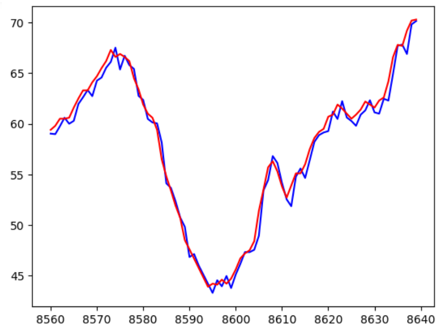
```
                                        AutoReg Model Results
==========================================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  NO2 (µg/m3)   No. Observations:            8640
Model:                                              AutoReg(3)   Log Likelihood              -21774.830
Method:                                        Conditional MLE   S.D. of innovations              3.011
Date:                                         Tue, 27 Jun 2023   AIC                          43559.661
Time:                                                 18:10:57   BIC                          43594.980
Sample:                                                      3   HQIC                         43571.704
                                                          8640
=================================================================================================================
                                                          coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------------------------
const                                                   1.2417      0.096     12.952      0.000       1.054       1.430
Singrauli, Surya Kiran Bhawan Dudhichua  NO2 (µg/m3).L1  1.4797     0.011    137.515      0.000       1.459       1.501
Singrauli, Surya Kiran Bhawan Dudhichua  NO2 (µg/m3).L2 -0.5027     0.018    -27.264      0.000      -0.539      -0.467
Singrauli, Surya Kiran Bhawan Dudhichua  NO2 (µg/m3).L3  0.0007     0.011      0.062      0.950      -0.020       0.022
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0496           +0.0000j            1.0496            0.0000
AR.2            1.9026           +0.0000j            1.9026            0.0000
AR.3          746.5461           +0.0000j          746.5461            0.0000
-----------------------------------------------------------------------------
```




```
Mean : 55.430689
Root Mean Squared Error : 1.292297
```

NOX

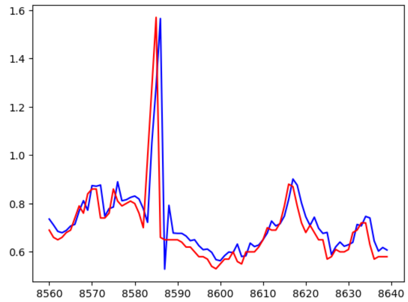
```
                                        AutoReg Model Results
==========================================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  NOX (ppb)   No. Observations:              8640
Model:                                              AutoReg(3)   Log Likelihood              -21068.541
Method:                                        Conditional MLE   S.D. of innovations              2.774
Date:                                         Tue, 27 Jun 2023   AIC                          42147.081
Time:                                                 18:10:57   BIC                          42182.400
Sample:                                                      3   HQIC                         42159.124
                                                          8640
=================================================================================================================
                                                          coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------------------------------------
const                                                   0.9017      0.065     13.767      0.000       0.773       1.030
Singrauli, Surya Kiran Bhawan Dudhichua  NOX (ppb).L1   1.7625      0.011    165.867      0.000       1.742       1.783
Singrauli, Surya Kiran Bhawan Dudhichua  NOX (ppb).L2  -0.9412      0.019    -48.931      0.000      -0.979      -0.904
Singrauli, Surya Kiran Bhawan Dudhichua  NOX (ppb).L3   0.1574      0.011     14.815      0.000       0.137       0.178
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0661           +0.0000j            1.0661            0.0000
AR.2            2.1829           +0.0000j            2.1829            0.0000
AR.3            2.7294           +0.0000j            2.7294            0.0000
-----------------------------------------------------------------------------
```




```
Mean : 42.328802
Root Mean Squared Error : 0.888006
```

CO

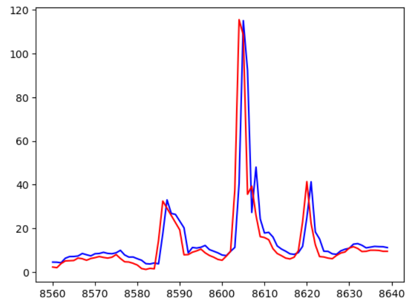
```
                                        AutoReg Model Results
==========================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  CO (mg/m3)   No. Observations:                 8640
Model:                                                AutoReg(3)   Log Likelihood                3269.221
Method:                                          Conditional MLE   S.D. of innovations              0.166
Date:                                           Tue, 27 Jun 2023   AIC                          -6528.441
Time:                                                   18:10:57   BIC                          -6493.122
Sample:                                                        3   HQIC                         -6516.398
                                                            8640
======================================================================================================
                                                    coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------------------------------
const                                             0.0495      0.004     11.221      0.000       0.041       0.058
Singrauli, Surya Kiran Bhawan Dudhichua  CO (mg/m3).L1     1.0969      0.011    102.796      0.000       1.076       1.118
Singrauli, Surya Kiran Bhawan Dudhichua  CO (mg/m3).L2    -0.2610      0.016    -16.667      0.000      -0.292      -0.230
Singrauli, Surya Kiran Bhawan Dudhichua  CO (mg/m3).L3     0.1289      0.011     12.080      0.000       0.108       0.150
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0365           -0.0000j            1.0365           -0.0000
AR.2            0.4944           -2.6909j            2.7360           -0.2211
AR.3            0.4944           +2.6909j            2.7360            0.2211
-----------------------------------------------------------------------------
```




```
Mean : 1.401927
Root Mean Squared Error : 0.123370
```

SO2

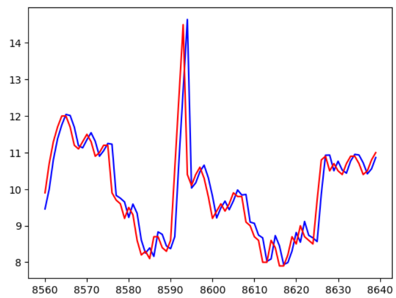
```
                                        AutoReg Model Results
==========================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  SO2 (µg/m3)   No. Observations:                 8640
Model:                                                AutoReg(3)   Log Likelihood              -35623.041
Method:                                          Conditional MLE   S.D. of innovations             14.962
Date:                                           Tue, 27 Jun 2023   AIC                          71256.081
Time:                                                   18:10:58   BIC                          71291.400
Sample:                                                        3   HQIC                         71268.124
                                                            8640
======================================================================================================
                                                    coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------------------------------
const                                             2.4705      0.212     11.661      0.000       2.055       2.886
Singrauli, Surya Kiran Bhawan Dudhichua  SO2 (µg/m3).L1     1.0446      0.011     97.801      0.000       1.024       1.066
Singrauli, Surya Kiran Bhawan Dudhichua  SO2 (µg/m3).L2    -0.2433      0.015    -15.920      0.000      -0.273      -0.213
Singrauli, Surya Kiran Bhawan Dudhichua  SO2 (µg/m3).L3     0.1214      0.011     11.362      0.000       0.100       0.142
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0829           -0.0000j            1.0829           -0.0000
AR.2            0.4609           -2.7197j            2.7585           -0.2233
AR.3            0.4609           +2.7197j            2.7585            0.2233
-----------------------------------------------------------------------------
```




```
Mean : 31.923270
Root Mean Squared Error : 12.248406
```

NH3

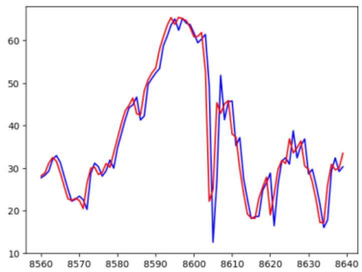
```
                                        AutoReg Model Results
==========================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  NH3 (µg/m3)   No. Observations:                 8640
Model:                                                AutoReg(3)   Log Likelihood              -12623.571
Method:                                          Conditional MLE   S.D. of innovations              1.044
Date:                                           Tue, 27 Jun 2023   AIC                          25257.142
Time:                                                   18:10:58   BIC                          25292.461
Sample:                                                        3   HQIC                         25269.185
                                                            8640
======================================================================================================
                                                    coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------------------------------
const                                             0.2084      0.027      7.766      0.000       0.156       0.261
Singrauli, Surya Kiran Bhawan Dudhichua  NH3 (µg/m3).L1     1.0795      0.011    100.339      0.000       1.058       1.101
Singrauli, Surya Kiran Bhawan Dudhichua  NH3 (µg/m3).L2    -0.1078      0.016     -6.826      0.000      -0.139      -0.077
Singrauli, Surya Kiran Bhawan Dudhichua  NH3 (µg/m3).L3     0.0126      0.011      1.168      0.243      -0.009       0.034
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0175           -0.0000j            1.0175           -0.0000
AR.2            3.7817           -7.9967j            8.8458           -0.1797
AR.3            3.7817           +7.9967j            8.8458            0.1797
-----------------------------------------------------------------------------
```




```
Mean : 13.286956
Root Mean Squared Error : 0.701728
```

Ozone

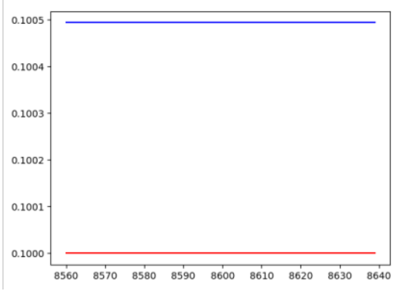
```
                              AutoReg Model Results
==========================================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  Ozone (µg/m3)   No. Observations:          8640
Model:                                          AutoReg(3)   Log Likelihood                -26579.724
Method:                                    Conditional MLE   S.D. of innovations                5.251
Date:                                     Tue, 27 Jun 2023   AIC                            53169.447
Time:                                             18:10:58   BIC                            53204.766
Sample:                                                  3   HQIC                           53181.490
                                                      8640
======================================================================================================================
=
                                                             coef    std err          z      P>|z|      [0.025      0.97
5]
----------------------------------------------------------------------------------------------------------------------
-
const                                                      0.9851      0.094     10.473      0.000       0.801       1.17
0
Singrauli, Surya Kiran Bhawan Dudhichua  Ozone (µg/m3).L1  1.3038      0.011    121.196      0.000       1.283       1.32
5
Singrauli, Surya Kiran Bhawan Dudhichua  Ozone (µg/m3).L2 -0.3546      0.017    -20.543      0.000      -0.388      -0.32
1
Singrauli, Surya Kiran Bhawan Dudhichua  Ozone (µg/m3).L3  0.0229      0.011      2.128      0.033       0.002       0.04
4
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0430           +0.0000j            1.0430            0.0000
AR.2            4.0138           +0.0000j            4.0138            0.0000
AR.3           10.4373           +0.0000j           10.4373            0.0000
-----------------------------------------------------------------------------
```



```
Mean : 35.193970
Root Mean Squared Error : 5.236985
```

Benzene

```
                              AutoReg Model Results
==========================================================================================================
Dep. Variable:     Singrauli, Surya Kiran Bhawan Dudhichua  Benzene (µg/m3)   No. Observations:        8640
Model:                                            AutoReg(3)   Log Likelihood               25044.425
Method:                                      Conditional MLE   S.D. of innovations              0.013
Date:                                       Tue, 27 Jun 2023   AIC                         -50078.850
Time:                                               18:10:58   BIC                         -50043.531
Sample:                                                    3   HQIC                        -50066.807
                                                        8640
======================================================================================================================
===
                                                               coef    std err          z      P>|z|      [0.025       0.9
75]
----------------------------------------------------------------------------------------------------------------------
---
const                                                        0.0029      0.000      9.209      0.000       0.002        0.
004
Singrauli, Surya Kiran Bhawan Dudhichua  Benzene (µg/m3).L1  0.9524      0.011     88.582      0.000       0.931        0.
973
Singrauli, Surya Kiran Bhawan Dudhichua  Benzene (µg/m3).L2  0.0625      0.015      4.210      0.000       0.033        0.
092
Singrauli, Surya Kiran Bhawan Dudhichua  Benzene (µg/m3).L3 -0.0389      0.011     -3.625      0.000      -0.060       -0.
018
                                    Roots
=============================================================================
                  Real          Imaginary           Modulus         Frequency
-----------------------------------------------------------------------------
AR.1            1.0251           +0.0000j            1.0251            0.0000
AR.2           -4.7244           +0.0000j            4.7244            0.5000
AR.3            5.3040           +0.0000j            5.3040            0.0000
-----------------------------------------------------------------------------
```



```
Mean : 0.122002
Root Mean Squared Error : 0.000493
```