CEEW_assignment_Samarth_Narula

December 5, 2023

1 CEEW Assignment

Using Excel, Python, Jupyter Notebook, and Markdown

Submitted by: Samarth Narula (M2022DS043)

2 Part 1

Q1) Download the daily and annual PM10 concentrations of Delhi for FY (fiscal year) 2021-2022 from this link. Your task is to analyse and present the findings from the data in a coherent write-up (Max word limit 1000 words). Below are some of the guiding questions that you could consider for the analysis: - Summarise statistics from the data. - Show the spatial distribution of air quality stations on a map. - Do they have missing data? How is it impacting the city's measurement? - Do you observe any seasonality in the data? Elaborate on the possible reasons - What are some measures you would recommend to improve air quality in Delhi

2.0.1 1. Introduction

2.0.2 Analysis of PM10 Concentrations in Delhi for Fiscal Year 2021-2022

This report presents a comprehensive analysis of PM10 (Particulate Matter 10 microns) concentrations in Delhi for the fiscal year 2021-2022. Utilizing data from various air quality monitoring stations across the city, the analysis aims to provide insights into the air quality trends, spatial distribution of pollution, the impact of missing data, and seasonal variations in PM10 levels. Additionally, the report suggests measures to improve air quality in Delhi. The study employs a range of data analytical techniques, including statistical summaries, mapping, and seasonal trend analysis, to derive meaningful conclusions about Delhi's air quality.

2.0.3 2. Data loading and Pre-processing

- Details
  - Load data using pandas
  - Perform data cleaning
  - Handling missing values

[4]:
```
import pandas as pd
#Load the data
data_delhi_pm10 = pd.read_excel("Delhi_PM10_Daily_Data.xlsx")
#Display some of the starting data
```

```
data_delhi_pm10.head()
```

```
[4]:   S.No             From Date               To Date Alipur, Delhi - DPCC  \
0  S.No             From Date               To Date          PM10(ug/m3)
1     1  01-Apr-2021 - 00:00  02-Apr-2021 - 00:00               209.75
2     2  02-Apr-2021 - 00:00  03-Apr-2021 - 00:00               199.88
3     3  03-Apr-2021 - 00:00  04-Apr-2021 - 00:00               183.83
4     4  04-Apr-2021 - 00:00  05-Apr-2021 - 00:00               209.71

  Anand Vihar, Delhi - DPCC Ashok Vihar, Delhi - DPCC Aya Nagar, Delhi - IMD  \
0              PM10(ug/m3)              PM10(ug/m3)            PM10(ug/m3)
1                   241.08                   209.68                 158.37
2                    182.8                   174.79                 146.93
3                   189.39                   182.27                  106.6
4                   187.58                   180.22                 145.36

  Bawana, Delhi - DPCC Burari Crossing, Delhi - IMD  \
0          PM10(ug/m3)                  PM10(ug/m3)
1               314.67                          NaN
2               248.58                          NaN
3               221.77                          NaN
4               281.98                          NaN

  CRRI Mathura Road, Delhi - IMD  … Pusa, Delhi - DPCC Pusa, Delhi - IMD  \
0                    PM10(ug/m3)  …        PM10(ug/m3)       PM10(ug/m3)
1                         183.55  …             238.64               NaN
2                         190.37  …             217.67               NaN
3                         223.05  …                180               NaN
4                         241.47  …              217.9               NaN

  R K Puram, Delhi - DPCC Rohini, Delhi - DPCC Shadipur, Delhi - CPCB  \
0             PM10(ug/m3)          PM10(ug/m3)            PM10(ug/m3)
1                  204.71               209.48                 183.57
2                  219.59               196.25                 145.34
3                   204.6               174.72                 182.15
4                  213.32               234.62                 233.35

  Sirifort, Delhi - CPCB Sonia Vihar, Delhi - DPCC  \
0            PM10(ug/m3)               PM10(ug/m3)
1                 192.84                    228.59
2                 173.11                    199.05
3                 172.98                    182.11
4                  161.5                    211.51

  Sri Aurobindo Marg, Delhi - DPCC Vivek Vihar, Delhi - DPCC  \
0                      PM10(ug/m3)               PM10(ug/m3)
1                           163.92                    237.12
```

```
2                        121.23                    181.2
3                         81.67                   206.36
4                        106.52                   187.91

  Wazirpur, Delhi - DPCC
0            PM10(ug/m3)
1                    NaN
2                    NaN
3                    NaN
4                    NaN

[5 rows x 43 columns]
```

(a) Here the first row is largerly redundant and might interfere with the data analysis since it contains non-numeric values

[13]:
```python
data_cleaned = data_delhi_pm10.iloc[1:].copy()
#data_cleaned.head()
```

(b) Next, convert 'From Date' and 'To Date' to datetime format and PM10 Concentration to numeric for better readability and data analysis:

[34]:
```python
data_cleaned['From Date'] = pd.to_datetime(data_cleaned['From Date'],format='%d-%b-%Y - %H:%M')
data_cleaned['To Date'] = pd.to_datetime(data_cleaned['To Date'], format='%d-%b-%Y - %H:%M')
#data_cleaned.head()
#Convert PM10 concentration columns to numeric
pm10_columns = data_cleaned.columns[3:]
for col in pm10_columns:
    data_cleaned[col] = pd.to_numeric(data_cleaned[col], errors='coerce')
data_cleaned.head()
```

[34]:
```
  S.No  From Date    To Date  Alipur, Delhi - DPCC  Anand Vihar, Delhi - DPCC  \
1    1 2021-04-01 2021-04-02                209.75                     241.08
2    2 2021-04-02 2021-04-03                199.88                     182.80
3    3 2021-04-03 2021-04-04                183.83                     189.39
4    4 2021-04-04 2021-04-05                209.71                     187.58
5    5 2021-04-05 2021-04-06                216.77                     290.15

   Ashok Vihar, Delhi - DPCC  Aya Nagar, Delhi - IMD  Bawana, Delhi - DPCC  \
1                     209.68                  158.37                314.67
2                     174.79                  146.93                248.58
3                     182.27                  106.60                221.77
4                     180.22                  145.36                281.98
5                     236.39                  201.95                309.48

   Burari Crossing, Delhi - IMD  CRRI Mathura Road, Delhi - IMD  …  \
```

```
1                         NaN                         183.55  ...
2                         NaN                         190.37  ...
3                         NaN                         223.05  ...
4                         NaN                         241.47  ...
5                         NaN                         290.61  ...

   Pusa, Delhi - DPCC  Pusa, Delhi - IMD  R K Puram, Delhi - DPCC  \
1              238.64                NaN                   204.71
2              217.67                NaN                   219.59
3              180.00                NaN                   204.60
4              217.90                NaN                   213.32
5              181.71                NaN                   287.79

   Rohini, Delhi - DPCC  Shadipur, Delhi - CPCB  Sirifort, Delhi - CPCB  \
1                209.48                  183.57                  192.84
2                196.25                  145.34                  173.11
3                174.72                  182.15                  172.98
4                234.62                  233.35                  161.50
5                224.46                  244.42                  183.89

   Sonia Vihar, Delhi - DPCC  Sri Aurobindo Marg, Delhi - DPCC  \
1                     228.59                            163.92
2                     199.05                            121.23
3                     182.11                             81.67
4                     211.51                            106.52
5                     253.76                            144.89

   Vivek Vihar, Delhi - DPCC  Wazirpur, Delhi - DPCC
1                     237.12                     NaN
2                     181.20                     NaN
3                     206.36                     NaN
4                     187.91                     NaN
5                     267.46                     NaN

[5 rows x 43 columns]
```

2.0.4 3. Handling missing data

Table 1: shows the count of missing values for each air quality monitoring station in Delhi: `Note:0 represents no missing value`

[15]:
```
#Count the number of missing values in each column
missing_data_count = data_cleaned.isnull().sum()
#Display the missing data
print(f"Missing Data Count: ")
print(missing_data_count)
```

```
Missing Data Count:
S.No                                                0
From Date                                           0
To Date                                             0
Alipur, Delhi - DPCC                                1
Anand Vihar, Delhi - DPCC                           3
Ashok Vihar, Delhi - DPCC                           0
Aya Nagar, Delhi - IMD                              4
Bawana, Delhi - DPCC                                0
Burari Crossing, Delhi - IMD                      208
CRRI Mathura Road, Delhi - IMD                     60
Chandni Chowk, Delhi - IITM                        11
DTU, Delhi - CPCB                                   6
Dr. Karni Singh Shooting Range, Delhi - DPCC        0
Dwarka-Sector 8, Delhi - DPCC                       0
IGI Airport (T3), Delhi - IMD                     105
IHBAS, Dilshad Garden, Delhi - CPCB                 1
ITO, Delhi - CPCB                                 155
Jahangirpuri, Delhi - DPCC                          0
Jawaharlal Nehru Stadium, Delhi - DPCC              0
Lodhi Road, Delhi - IITM                           39
Lodhi Road, Delhi - IMD                            14
Major Dhyan Chand National Stadium, Delhi - DPCC    1
Mandir Marg, Delhi - DPCC                           0
Mundka, Delhi - DPCC                                1
NSIT Dwarka, Delhi - CPCB                           0
Najafgarh, Delhi - DPCC                             5
Narela, Delhi - DPCC                                1
Nehru Nagar, Delhi - DPCC                           0
New Moti Bagh, Delhi - MHUA                       365
North Campus, DU, Delhi - IMD                      16
Okhla Phase-2, Delhi - DPCC                         0
Patparganj, Delhi - DPCC                           20
Punjabi Bagh, Delhi - DPCC                          0
Pusa, Delhi - DPCC                                  1
Pusa, Delhi - IMD                                  62
R K Puram, Delhi - DPCC                             3
Rohini, Delhi - DPCC                                0
Shadipur, Delhi - CPCB                              0
Sirifort, Delhi - CPCB                             16
Sonia Vihar, Delhi - DPCC                           8
Sri Aurobindo Marg, Delhi - DPCC                    2
Vivek Vihar, Delhi - DPCC                           0
Wazirpur, Delhi - DPCC                             71
dtype: int64
```

Interpretation of missing column values:

1. Station-wise Missing Data: Each line in the output (e.g., Alipur, Delhi - DPCC 1, Anand Vihar, Delhi - DPCC 3) shows the number of days for which the PM10 data is missing at each station. For example, Alipur, Delhi - DPCC has 1 missing value, meaning there's one day's data missing from this station in the dataset.
2. Significant Missing Data: Some stations, like Burari Crossing, Delhi - IMD, and New Moti Bagh, Delhi - MHUA, have a large number of missing values (208 and 365 in this case), indicating that a considerable portion of data is not available for these locations. This could impact any analysis that requires a complete or near-complete dataset for every station.
3. Reliable Data Sources: Stations with zero missing values, such as Ashok Vihar, Delhi - DPCC, Bawana, Delhi - DPCC, Dr. Karni Singh Shooting Range, Delhi - DPCC, and others, can be considered reliable sources of data for the time period in question. These stations have consistently reported data without any gaps, which suggests that their monitoring equipment was operational and the data collection process was uninterrupted during this period.
4. Robustness in Analysis: For any analysis where completeness of data is crucial, such as time-series analysis, trend analysis, or seasonal variation studies, these stations provide a robust dataset. Since they don't have missing values, their data can be used with more confidence, knowing that it represents an uninterrupted record of PM10 concentrations over time.

[16]:
```python
# Counting missing values in each row
missing_data_per_row = data_cleaned.isnull().sum(axis=1)

# Displaying the count of missing values in the first few rows
print("Missing Data per Row in First Few Rows:")
print(missing_data_per_row.head())
```

```
Missing Data per Row in First Few Rows:
1    5
2    5
3    5
4    6
5    5
dtype: int64
```

Interpretation of missing row values:

1. Consistent Number of Missing Values per Day: The count of missing values per row (representing each day) is quite consistent, with most days missing data from 5 or 6 stations. This consistency suggests that a specific subset of stations consistently failed to report data on these days.
2. Potential Data Reporting Issues: The regular occurrence of missing data across similar numbers of stations each day could indicate systematic issues with data collection or reporting at certain stations. This might be due to technical issues, maintenance activities, or other operational challenges at these particular monitoring locations.

Impact on City's Air Quality Measurements:

1. Accuracy of City-Wide Air Quality Index (AQI): An accurate AQI requires data from multiple stations to represent different parts of the city. Missing data from key stations could lead to an AQI that doesn't fully capture the variability of air quality across Delhi.
2. Policy and Health Implications: Reliable air quality data is crucial for formulating public health advisories and environmental policies. Incomplete data might lead to less informed decisions or delayed responses to pollution events.
3. Temporal Analysis Challenges: For time-sensitive analyses, like identifying pollution trends or correlating air quality with specific events (like festivals, weather changes), missing data can create significant challenges. It may lead to biased conclusions if the missing data corresponds to critical periods.

2.0.5 4. Summary Statistics

The following table shows the summary statistics of PM10 concentrations for each station. One can look at the various measures and understand what they indicate about air quality. Here are some key points to consider from the data:

Mean Concentration Levels: The mean (average) PM10 levels at each station provide a general idea of the air quality. Higher mean values indicate higher levels of particulate matter at those stations, suggesting poorer air quality.

Variability of PM10 Levels: The standard deviation (std) gives an idea of how much the PM10 levels vary over time. A high standard deviation means that the PM10 levels fluctuate significantly, which could be due to various factors like weather changes, local activities, or seasonal variations.

Extremes in Air Quality: The minimum and maximum values show the range of PM10 levels. Extremely high maximum values might indicate pollution events or specific days with very poor air quality.

Distribution of PM10 Levels: The 25th, 50th (median), and 75th percentiles help you understand the distribution of PM10 levels. For example, if the 75th percentile is very high compared to the median, it suggests that there are periods with exceptionally high pollution levels.

Table 2: Descriptive statistics for PM10 concentrations in each station:

[29]:
```
# Selecting only the columns with PM10 data (excluding 'S.No', 'From Date', 'To␣
↪Date')
pm10_data_only = data_cleaned.iloc[:, 3:]

# Calculating descriptive statistics for PM10 concentration data
descriptive_stats_pm10 = pm10_data_only.describe()

# Round the numbers for better readability and transpose for easier viewing
descriptive_stats_full = descriptive_stats_pm10.transpose()
# Apply rounding after transposing to retain all summary statistics
descriptive_stats_full_rounded = descriptive_stats_full.round(2)

#Remove width constraint
pd.options.display.width = None
```

```python
pd.options.display.max_colwidth = 200

# Selecting key statistics
key_stats = descriptive_stats_full_rounded[['mean', 'std', 'min', '50%', 'max']]

# Displaying the key statistics
print("Formatted Descriptive Statistics for PM10 Concentrations in each Station␣
↪in Delhi:")
print(key_stats)
```

```
Formatted Descriptive Statistics for PM10 Concentrations in each Station in
Delhi:
                                                    mean     std    min  \
Alipur, Delhi - DPCC                              193.64  118.89  19.37
Anand Vihar, Delhi - DPCC                         288.47  163.22  45.44
Ashok Vihar, Delhi - DPCC                         207.55  127.54  20.94
Aya Nagar, Delhi - IMD                            142.51   89.72   9.77
Bawana, Delhi - DPCC                              236.62  138.83  32.63
Burari Crossing, Delhi - IMD                      242.94   91.99  42.78
CRRI Mathura Road, Delhi - IMD                    206.35  124.99  14.78
Chandni Chowk, Delhi - IITM                       270.31  134.66  49.26
DTU, Delhi - CPCB                                 185.60  103.96  15.72
Dr. Karni Singh Shooting Range, Delhi - DPCC      173.30  103.00  17.28
Dwarka-Sector 8, Delhi - DPCC                     239.86  131.76  19.79
IGI Airport (T3), Delhi - IMD                     174.32  102.59  17.65
IHBAS, Dilshad Garden, Delhi - CPCB               199.92  118.20  36.39
ITO, Delhi - CPCB                                 175.53   73.61  16.00
Jahangirpuri, Delhi - DPCC                        244.72  144.06  30.85
Jawaharlal Nehru Stadium, Delhi - DPCC            180.32  106.27  16.09
Lodhi Road, Delhi - IITM                          142.34   91.04   2.38
Lodhi Road, Delhi - IMD                           163.11   90.51  18.74
Major Dhyan Chand National Stadium, Delhi - DPCC  185.23  107.53  25.51
Mandir Marg, Delhi - DPCC                         159.54   88.70  29.59
Mundka, Delhi - DPCC                              241.23  135.07  22.25
NSIT Dwarka, Delhi - CPCB                         215.25  114.23  31.92
Najafgarh, Delhi - DPCC                           150.15   88.82  11.81
Narela, Delhi - DPCC                              230.04  124.29  39.14
Nehru Nagar, Delhi - DPCC                         196.93  125.36  24.18
New Moti Bagh, Delhi - MHUA                          NaN     NaN    NaN
North Campus, DU, Delhi - IMD                     182.47  111.79  23.80
Okhla Phase-2, Delhi - DPCC                       210.72  119.38  26.65
Patparganj, Delhi - DPCC                          194.03  110.12  24.91
Punjabi Bagh, Delhi - DPCC                        203.08  122.93  27.10
Pusa, Delhi - DPCC                                197.76  113.84  19.80
Pusa, Delhi - IMD                                 155.85   99.30  26.40
R K Puram, Delhi - DPCC                           185.58   98.18  16.31
Rohini, Delhi - DPCC                              218.99  134.82  23.93
Shadipur, Delhi - CPCB                            229.35  134.03  40.27
```

Sirifort, Delhi - CPCB	193.31	110.06	19.80
Sonia Vihar, Delhi - DPCC	214.97	124.22	13.00
Sri Aurobindo Marg, Delhi - DPCC	141.55	87.86	13.11
Vivek Vihar, Delhi - DPCC	220.20	135.68	37.94
Wazirpur, Delhi - DPCC	259.48	146.52	42.50

	50%	max
Alipur, Delhi - DPCC	176.26	563.33
Anand Vihar, Delhi - DPCC	265.62	727.53
Ashok Vihar, Delhi - DPCC	184.54	661.59
Aya Nagar, Delhi - IMD	126.75	553.51
Bawana, Delhi - DPCC	222.76	679.54
Burari Crossing, Delhi - IMD	234.38	513.33
CRRI Mathura Road, Delhi - IMD	198.87	609.60
Chandni Chowk, Delhi - IITM	249.60	790.18
DTU, Delhi - CPCB	173.13	492.19
Dr. Karni Singh Shooting Range, Delhi - DPCC	159.78	570.09
Dwarka-Sector 8, Delhi - DPCC	224.11	613.00
IGI Airport (T3), Delhi - IMD	158.48	512.38
IHBAS, Dilshad Garden, Delhi - CPCB	177.70	638.88
ITO, Delhi - CPCB	168.19	448.62
Jahangirpuri, Delhi - DPCC	224.54	759.47
Jawaharlal Nehru Stadium, Delhi - DPCC	167.83	559.43
Lodhi Road, Delhi - IITM	122.36	704.60
Lodhi Road, Delhi - IMD	149.63	525.69
Major Dhyan Chand National Stadium, Delhi - DPCC	167.50	564.20
Mandir Marg, Delhi - DPCC	142.54	506.83
Mundka, Delhi - DPCC	229.18	660.04
NSIT Dwarka, Delhi - CPCB	208.88	576.30
Najafgarh, Delhi - DPCC	136.65	427.14
Narela, Delhi - DPCC	201.71	611.25
Nehru Nagar, Delhi - DPCC	172.38	656.25
New Moti Bagh, Delhi - MHUA	NaN	NaN
North Campus, DU, Delhi - IMD	158.73	610.72
Okhla Phase-2, Delhi - DPCC	192.81	640.14
Patparganj, Delhi - DPCC	177.53	586.38
Punjabi Bagh, Delhi - DPCC	178.64	640.77
Pusa, Delhi - DPCC	178.61	602.59
Pusa, Delhi - IMD	136.71	508.23
R K Puram, Delhi - DPCC	178.10	522.26
Rohini, Delhi - DPCC	195.53	665.35
Shadipur, Delhi - CPCB	203.95	644.37
Sirifort, Delhi - CPCB	176.76	651.04
Sonia Vihar, Delhi - DPCC	194.44	670.19
Sri Aurobindo Marg, Delhi - DPCC	126.90	489.46
Vivek Vihar, Delhi - DPCC	191.41	674.34
Wazirpur, Delhi - DPCC	221.35	793.65

[1]:
```
pip install matplotlib seaborn
```

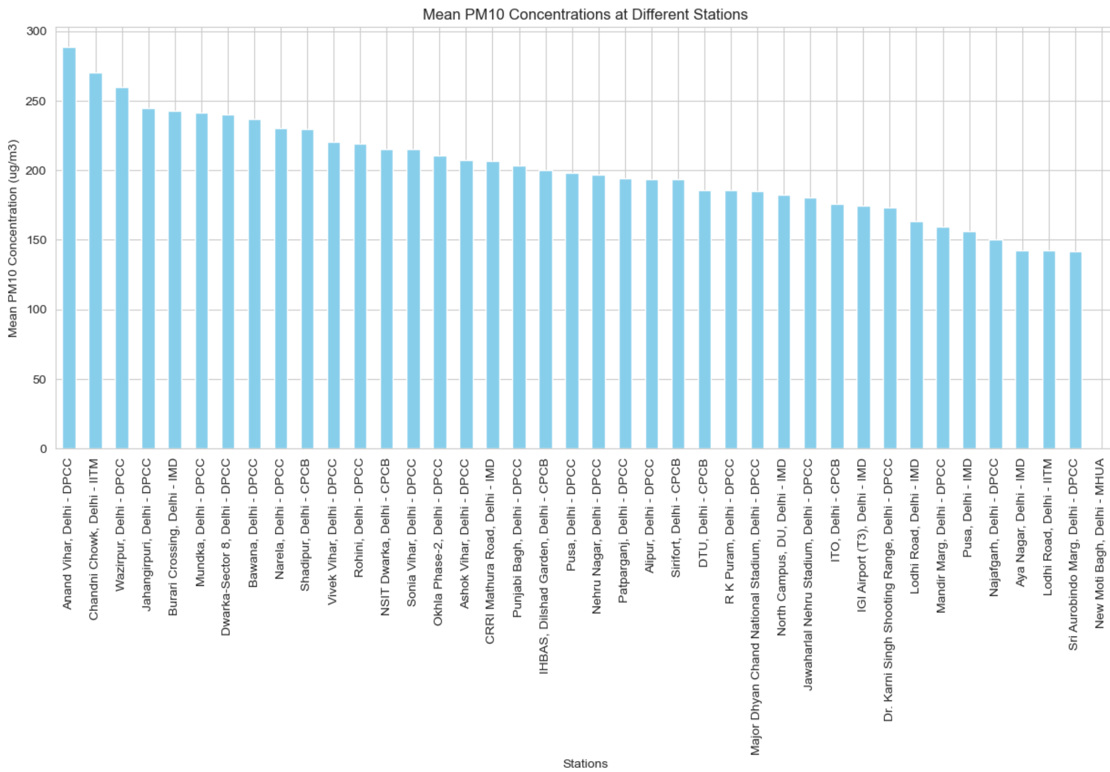
```
Requirement already satisfied: matplotlib in c:\users\samar\anaconda3\lib\site-
packages (3.7.2)
Requirement already satisfied: seaborn in c:\users\samar\anaconda3\lib\site-
packages (0.12.2)
Requirement already satisfied: contourpy>=1.0.1 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (1.0.5)
Requirement already satisfied: cycler>=0.10 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: numpy>=1.20 in c:\users\samar\anaconda3\lib\site-
packages (from matplotlib) (1.24.3)
Requirement already satisfied: packaging>=20.0 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (23.1)
Requirement already satisfied: pillow>=6.2.0 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (10.0.1)
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in
c:\users\samar\anaconda3\lib\site-packages (from matplotlib) (2.8.2)
Requirement already satisfied: pandas>=0.25 in
c:\users\samar\anaconda3\lib\site-packages (from seaborn) (2.0.3)
Requirement already satisfied: pytz>=2020.1 in
c:\users\samar\anaconda3\lib\site-packages (from pandas>=0.25->seaborn)
(2023.3.post1)
Requirement already satisfied: tzdata>=2022.1 in
c:\users\samar\anaconda3\lib\site-packages (from pandas>=0.25->seaborn) (2023.3)
Requirement already satisfied: six>=1.5 in c:\users\samar\anaconda3\lib\site-
packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

2.0.6 (a) Graphical representation of station-wise average

Fig 1: Bar Graph showing mean PM10 concentrations at stations in Delhi

[31]:
```python
import matplotlib.pyplot as plt
import seaborn as sns

# Set the aesthetic style of the plots
sns.set_style("whitegrid")

# Calculate mean PM10 levels for each station
mean_pm10_levels = pm10_data_only.mean()

# Plotting
```

```python
plt.figure(figsize=(15, 6))
mean_pm10_levels.sort_values(ascending=False).plot(kind='bar', color='skyblue')
plt.title('Mean PM10 Concentrations at Different Stations')
plt.xlabel('Stations')
plt.ylabel('Mean PM10 Concentration (ug/m3)')
plt.xticks(rotation=90)  # Rotates station names for better visibility
plt.show()
```



2.0.7 (b) Graphical representation of the city-wide average

Note: This gives the average PM10 concentration across all stations for each day. It gives a daily city-wide average, which can then be further analyzed to understand daily variations or trends.

Table 3: Few rows of daily city-wide average

[49]:
```python
# Calculate the mean across all stations for each day
city_daily_avg = pm10_data_only.mean(axis=1)

# Add this as a new column to the dataframe
data_cleaned['Citywide Daily Average'] = city_daily_avg

# Now, you can view the first few rows to see the daily city-wide average
```

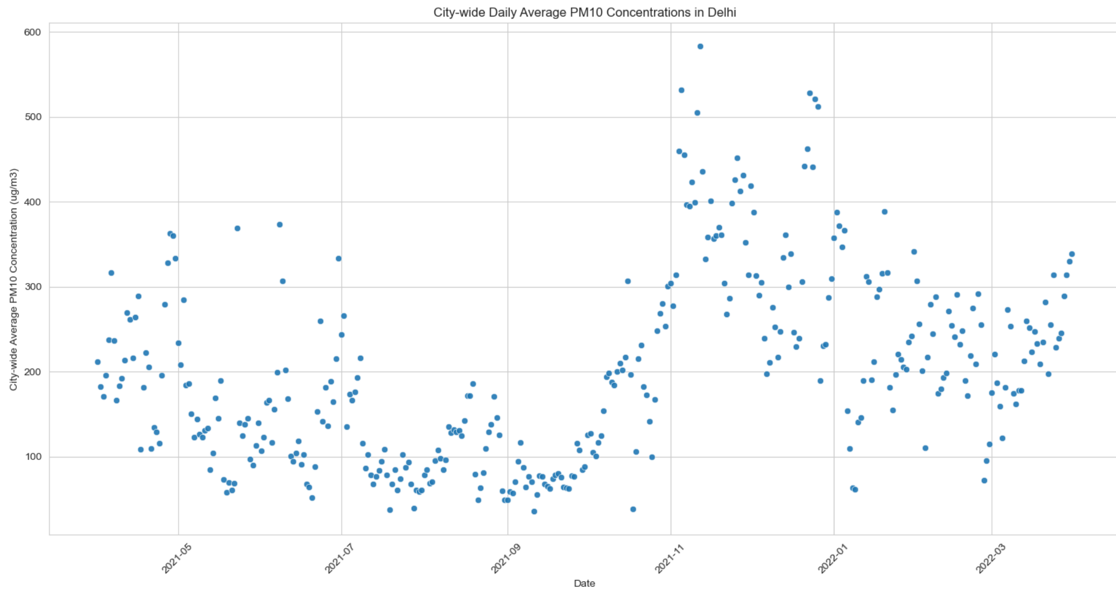
```
print(data_cleaned[['From Date', 'Citywide Daily Average']].head())
```

```
   From Date  Citywide Daily Average
1 2021-04-01              212.208000
2 2021-04-02              182.344571
3 2021-04-03              171.100571
4 2021-04-04              196.231471
5 2021-04-05              237.381714
```

Fig 2. Scatter Plot showing city-wide average PM10 concentration

[51]:

```
import matplotlib.pyplot as plt
import seaborn as sns

# Set the aesthetic style of the plots
sns.set_style("whitegrid")

# Create a DataFrame for plotting that includes the date and the city-wide␣
↪daily average
plot_data = data_cleaned[['From Date', 'Citywide Daily Average']]

# Plotting
plt.figure(figsize=(15, 8))
sns.scatterplot(data=plot_data, x='From Date', y='Citywide Daily Average',␣
↪alpha=0.9)

plt.title('City-wide Daily Average PM10 Concentrations in Delhi')
plt.xlabel('Date')
plt.ylabel('City-wide Average PM10 Concentration (ug/m3)')
plt.xticks(rotation=45)  # Rotates the dates for better visibility
plt.tight_layout()  # Adjusts the plot to ensure everything fits without␣
↪overlapping
plt.show()
```



Interpretation of city-wide average PM10 concentration:

1. Fluctuating Air Quality: The scatter plot may show significant day-to-day fluctuations in PM10 levels, suggesting that air quality in Delhi varies considerably. Such variation could be influenced by daily changes in traffic patterns, industrial activities, or weather conditions.

2. Seasonal Highs and Lows: Since data spans different seasons, the higher PM10 concentrations in cooler months, is possibly due to increased heating and stagnant air which traps pollutants. Conversely, lower concentrations might be observed during monsoon months due to rain washing away particulates.

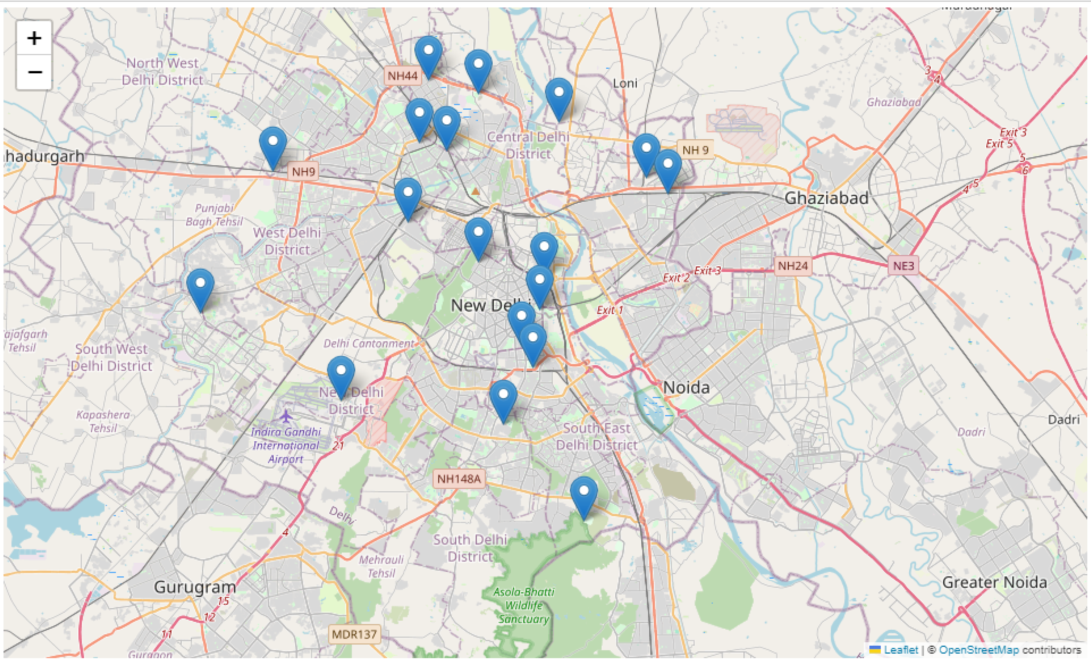
[47]:
```
pip install folium
```

```
Requirement already satisfied: folium in c:\users\samar\anaconda3\lib\site-
packages (0.15.1)Note: you may need to restart the kernel to use updated
packages.

Requirement already satisfied: branca>=0.6.0 in
c:\users\samar\anaconda3\lib\site-packages (from folium) (0.7.0)
Requirement already satisfied: jinja2>=2.9 in c:\users\samar\anaconda3\lib\site-
packages (from folium) (3.1.2)
Requirement already satisfied: numpy in c:\users\samar\anaconda3\lib\site-
packages (from folium) (1.24.3)
Requirement already satisfied: requests in c:\users\samar\anaconda3\lib\site-
packages (from folium) (2.31.0)
Requirement already satisfied: xyzservices in c:\users\samar\anaconda3\lib\site-
packages (from folium) (2022.9.0)
Requirement already satisfied: MarkupSafe>=2.0 in
c:\users\samar\anaconda3\lib\site-packages (from jinja2>=2.9->folium) (2.1.1)
```

```
Requirement already satisfied: charset-normalizer<4,>=2 in
c:\users\samar\anaconda3\lib\site-packages (from requests->folium) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in
c:\users\samar\anaconda3\lib\site-packages (from requests->folium) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in
c:\users\samar\anaconda3\lib\site-packages (from requests->folium) (1.26.16)
Requirement already satisfied: certifi>=2017.4.17 in
c:\users\samar\anaconda3\lib\site-packages (from requests->folium) (2023.11.17)
```

2.0.8 5. Spatial distribution of air quality stations on the Delhi Map

Fig 3: Map: Air quality stations in Delhi Note: Map created using folium, includes interactivity. I am adding a screenshot of the map as well for visibility in a



pdf

[46]:
```python
import pandas as pd
import folium

# Creating a DataFrame with station names and their coordinates
stations_data = pd.DataFrame({
    'Station': ['Ashok Vihar, Delhi - DPCC', 'Burari Crossing, Delhi - IMD',
                'Dr. Karni Singh Shooting Range, Delhi - DPCC', 'IGI Airport␣
 ↪(T3), Delhi - IMD',
                'IHBAS, Dilshad Garden, Delhi - CPCB', 'ITO, Delhi - CPCB',
                'Jahangirpuri, Delhi - DPCC', 'Jawaharlal Nehru Stadium, Delhi␣
 ↪- DPCC',
                'Lodhi Road, Delhi - IMD', 'Major Dhyan Chand National Stadium,␣
 ↪Delhi - DPCC',
                'Mandir Marg, Delhi - DPCC', 'Mundka, Delhi - DPCC',
```

```
                'Najafgarh, Delhi - DPCC', 'North Campus, DU, Delhi - IMD',
                'Sirifort, Delhi - CPCB', 'Sonia Vihar, Delhi - DPCC',
                'Vivek Vihar, Delhi - DPCC', 'Wazirpur, Delhi - DPCC'],
    'Latitude': [28.695381, 28.7256504, 28.498571, 28.5627763, 28.6811736, 28.
↪628624,
                 28.732820, 28.580280, 28.5918245, 28.611281, 28.636429, 28.
↪684678,
                 28.60909, 28.6573814, 28.5504249, 28.710508, 28.672342, 28.
↪699793],
    'Longitude': [77.181665, 77.2011573, 77.264840, 77.1180053, 77.3025234, 77.
↪241060,
                  77.170633, 77.233829, 77.2273074, 77.237738, 77.201067, 77.
↪076574,
                  77.0325413, 77.1585447, 77.2159377, 77.249485, 77.315260, 77.
↪165453]
})

stations_data
# Create a map centered around an average location in Delhi
delhi_map = folium.Map(location=[28.7041, 77.1025], zoom_start=11)

# Add markers for each air quality station
for idx, row in stations_data.iterrows():
    folium.Marker(
        location=[row['Latitude'], row['Longitude']],
        popup=row['Station']
    ).add_to(delhi_map)

# Display the map
delhi_map
```

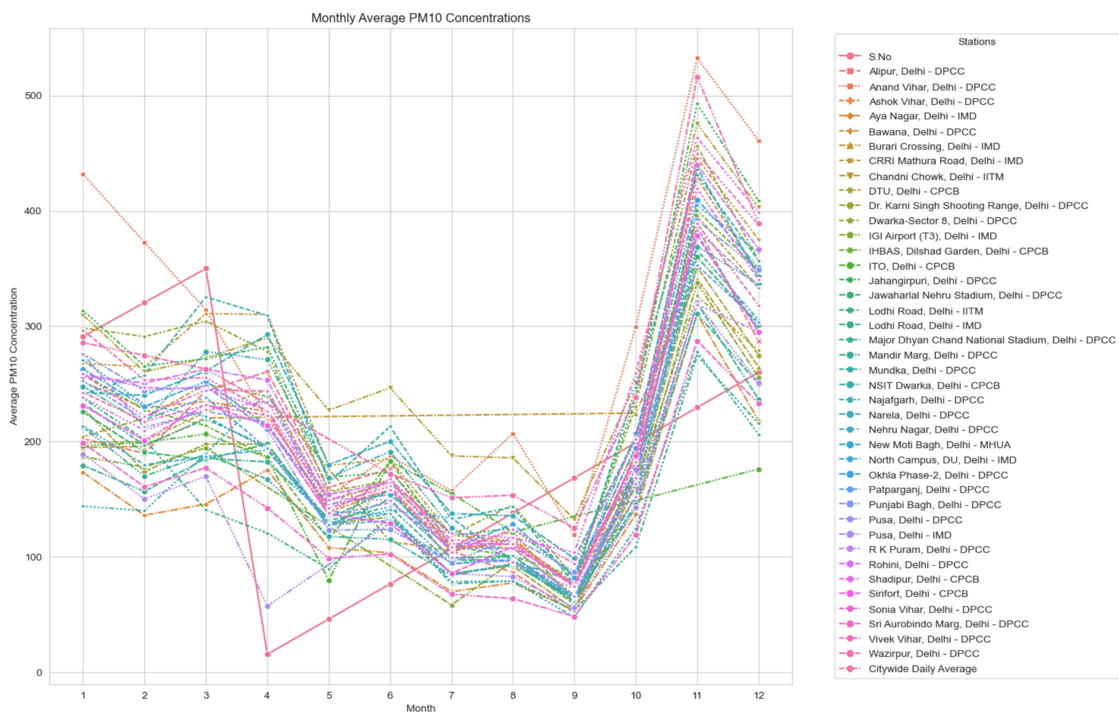
[46]: <folium.folium.Map at 0x1e5ed286190>

2.0.9 6. Seasonality Analysis

Fig 4: Line chart showing seasonality in data

[55]:
```
import matplotlib.pyplot as plt
import seaborn as sns

# Extract month from the 'From Date' column
data_cleaned['Month'] = data_cleaned['From Date'].dt.month

# Calculate average PM10 for each month
monthly_avg_pm10 = data_cleaned.groupby('Month').mean()

# Plotting
plt.figure(figsize=(15, 10))
```

```python
sns.lineplot(data=monthly_avg_pm10, markers=True)
plt.title('Monthly Average PM10 Concentrations')
plt.xlabel('Month')
plt.ylabel('Average PM10 Concentration')
plt.xticks(range(1, 13))  # Set x-axis ticks to represent each month
# Moving the legend outside the plot
plt.legend(title='Stations', bbox_to_anchor=(1.05, 1), loc='upper left')

plt.tight_layout()  # Adjust layout
plt.show()
```



Interpretation of seaosnal analysis:

1. Drop in PM10 in September: The decrease in PM10 levels around September may correspond to the end of the monsoon season in Delhi. The increased rainfall during monsoon typically helps in settling down airborne particles, leading to improved air quality.

2. Rapid Increase Post-Monsoon: The rapid increase in PM10 concentrations post-September could be attributed to factors like the onset of cooler weather, reduced rainfall, and agricultural stubble burning practices in surrounding regions, which are prevalent during October and November.

3. Peak in November: The peak in PM10 levels in November might coincide with Diwali and other festivals, where the use of fireworks significantly contributes to air pollution. Additionally, cooler temperatures and lower wind speeds in winter can trap pollutants closer to the

ground, exacerbating air quality issues.

4. Gradual Decrease Post-Winter: The subsequent decrease after November could be due to changing meteorological conditions, like increased wind speeds, which help disperse pollutants.

5. Regarding specific stations, stations like Anand Vihar, ITO, and Ashok Vihar, known for their high traffic and industrial activities, show more pronounced seasonal variations in PM10 levels, reflecting the impact of urban activities on air quality.

2.0.10 7. Recommendations for improving air quality in Delhi

- **Advanced Air Quality Monitoring:** Implement a city-wide network of advanced sensors for real-time air quality monitoring. Use data analytics to identify pollution hotspots and patterns.
- **Green Infrastructure:** Expand urban green spaces and rooftop gardens to absorb pollutants. Encourage tree plantation drives, especially in industrial and high-traffic areas.
- **Public Transport Enhancement:** Invest in clean public transportation options, like electric buses and metro services, to reduce vehicular emissions.
- **Strict Industrial Emission Controls:** Enforce stringent emission standards for industries and implement regular audits.
- **Public Awareness Campaigns:** Educate the public about the health impacts of air pollution and promote environment-friendly practices.
- **Policy Interventions During High Pollution Periods:** Implement policies like odd-even vehicle schemes during peak pollution months. Restrict activities like open burning and construction on high pollution days.
- **Encourage Renewable Energy:** Promote solar and wind energy to reduce reliance on fossil fuels.

2.0.11 8. Conclusion

Improving Delhi's air quality requires a multifaceted strategy combining technology, public policy, community involvement, and sustainable practices. By addressing the issue from multiple angles, the city can make significant strides towards cleaner air, ultimately leading to a healthier environment and improved quality of life for its residents.

2.0.12 9. References

1. Central Pollution Control Board (CPCB). (2021). *National Air Quality Standards*. https://cpcb.nic.in/
2. World Health Organization (WHO). (2021). *Air Quality Guidelines*. https://www.who.int/
3. The Energy and Resources Institute (TERI). (2021). *Strategies for Combating Air Pollution*. https://www.teriin.org/
4. Greenpeace India. (2020). *Renewable Energy and Air Quality*. https://www.greenpeace.org/india/en/
5. Delhi Metro Rail Corporation (DMRC). (2022). *Public Transport and Pollution Reduction*. https://www.delhimetrorail.com/

3 Part 2

Q2) The Indian government designated cities grappling with bad air quality as 'Non-Attainment Cities' (NACs) a few years ago. Do such designations exist in other countries, too? What parameters are employed to categorise cities based on their 'attainment'/'non-attainment' status in these countries? Your answer must include a comparative analysis of at least three countries, highlighting the framework they have adopted to classify their cities based on air pollution levels

3.0.1 1. Introduction

3.0.2 Non-Attainment Cities (NAC's) in India:

The concept of Non-Attainment Cities (NACs) in India is primarily framed within the National Clean Air Programme (NCAP), which was launched by the Ministry of Environment, Forest, and Climate Change (MoEF&CC) in January 2019. The program aims to tackle air pollution in a comprehensive manner across the country, with specific emphasis on cities that have persistently poor air quality. Here's a detailed breakdown of the criteria and implementation strategies for NACs in India:

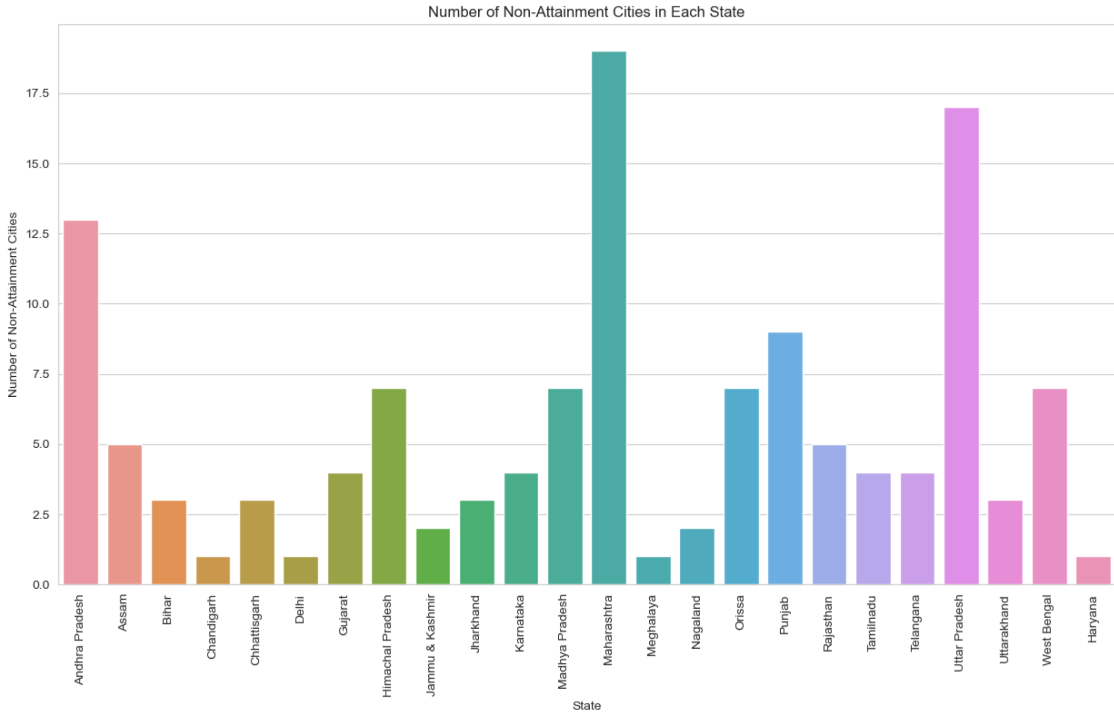
1. **Criteria for Non-Attainment Cities:** A city is designated as a Non-Attainment City if it consistently shows air quality worse than the National Ambient Air Quality Standards (NAAQS). The identification of NACs is based on an exceedance of the NAAQS for five consecutive years.
2. **Number of Non-Attainment Cities:** Initially, the NCAP targeted 102 cities, which later expanded to include 20 more cities, and currently encompasses 132 cities across India. These cities have exceeded the national ambient standards for pollutants like PM2.5, PM10, or NO2.
3. **Air Quality Monitoring and Pollutants:** In India, the NAAQS lists 12 major pollutants, but monitoring does not happen for all these pollutants in every city. The focus is primarily on particulate matter and nitrogen dioxide levels.
4. **Goals of the NCAP:** The NCAP aims for a 20 to 30% reduction in Particulate Matter (PM) concentration by 2024, compared to 2017 levels. This target was later revised to aim for a 40% reduction by 2026.
5. **Implementation Strategies:**
    - Expansion of air quality monitoring networks, including both manual and continuous monitoring stations.
    - Initiation of pilot projects to assess alternative ambient monitoring technologies like low-cost sensors and satellite-based monitoring.
    - Development of city-specific air action plans for the reduction of PM pollution levels.

3.0.3 Distribution of Non-Attainment Cities Across Indian States: A Measure of Air Quality Challenges:

Fig 1: Bar chart

[64]:
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```python
# Data from the PDF
data = {
    'State': ['Andhra Pradesh', 'Assam', 'Bihar', 'Chandigarh', 'Chhattisgarh', 'Delhi', 'Gujarat', 'Himachal Pradesh', 'Jammu & Kashmir', 'Jharkhand', 'Karnataka', 'Madhya Pradesh', 'Maharashtra', 'Meghalaya', 'Nagaland', 'Orissa', 'Punjab', 'Rajasthan', 'Tamilnadu', 'Telangana', 'Uttar Pradesh', 'Uttarakhand', 'West Bengal', 'Haryana'],
    'Number of Cities': [13, 5, 3, 1, 3, 1, 4, 7, 2, 3, 4, 7, 19, 1, 2, 7, 9, 5, 4, 4, 17, 3, 7, 1]
}

df = pd.DataFrame(data)
# Set the aesthetic style of the plots
sns.set_style("whitegrid")

# Create a bar plot
plt.figure(figsize=(15, 8))
barplot = sns.barplot(x='State', y='Number of Cities', data=df)

# Add labels and title
plt.xlabel('State')
plt.ylabel('Number of Non-Attainment Cities')
plt.title('Number of Non-Attainment Cities in Each State')
plt.xticks(rotation=90)  # Rotate labels for better readability

# Show the plot
plt.show()
```



Analytical Interpretations of the bar chart: **High Concentration in Urbanized States:** States with a higher number of non-attainment cities, like Maharashtra and Uttar Pradesh, are typically more urbanized and industrialized. This correlation suggests a strong link between urbanization, industrial activities, and air quality issues.

Regional Variations: The variation in the number of non-attainment cities across states points to regional disparities in air quality. States like West Bengal and Tamilnadu, despite being highly populated, show a relatively lower count of non-attainment cities, indicating possible differences in industrial activities, vehicular density, or environmental policies.

Policy Implications: States with a higher count of non-attainment cities may need more stringent air quality management policies. This could include stricter emission norms, promotion of cleaner fuels, and investment in public transport.

Health and Environmental Concerns: The presence of non-attainment cities in a state highlights potential health risks for its residents due to poor air quality. It also raises concerns about environmental degradation that could have broader implications for climate change.

Need for Targeted Action: The data suggests a need for targeted action in states with the most non-attainment cities. This could involve localized strategies to address specific sources of pollution, such as industrial emissions, vehicular pollution, or construction dust.

Role of Public Awareness: Raising public awareness about the sources and impacts of air pollution is crucial. Educating citizens about the importance of air quality can drive public demand for cleaner air and support for environmental policies.

Comparative Analysis for Policy Learning: States with fewer non-attainment cities can serve as models for effective air quality management. Comparative analysis can help identify best practices that could be replicated in states facing greater challenges.

3.0.4 2. Global Comparison:

(a) United States: The US Environmental Protection Agency (EPA) designates areas as attainment, nonattainment, or unclassifiable for each of the six criteria pollutants: ozone, particulate matter (PM), carbon monoxide, nitrogen dioxide, sulfur dioxide, and lead. An attainment area meets the national ambient air quality standards (NAAQS) for a given pollutant, while a nonattainment area does not. An unclassifiable area has insufficient data to determine its status. The EPA also assigns classifications to nonattainment areas based on the severity of their pollution problem, such as marginal, moderate, serious, severe, or extreme for ozone, and moderate or serious for PM. These classifications determine the deadlines and the control measures required for the areas to attain the NAAQS.

(b) China: The Chinese Ministry of Ecology and Environment (MEE) divides the country into five key regions for air pollution control: Beijing-Tianjin-Hebei and surrounding areas, Yangtze River Delta, Fenwei Plain, Chengdu-Chongqing, and Pearl River Delta. Within each region, the MEE sets annual targets for the average concentration of PM2.5, PM10, sulfur dioxide, nitrogen dioxide, and ozone for each city or prefecture. The MEE also evaluates the performance of each city or prefecture based on the attainment rate, which is the percentage of days in a year that meet the national ambient air quality standards (NAAQS). The MEE ranks the cities or prefectures into four categories: excellent, good, fair, or poor. The MEE also publishes monthly and quarterly reports on the air quality status and trends of each region and city or prefecture.

(c) European Union: The European Commission (EC) monitors and assesses the air quality situation in the member states based on the ambient air quality directives, which set limit values and target values for the concentration of several pollutants, such as PM, nitrogen dioxide, sulfur dioxide, ozone, lead, benzene, carbon monoxide, and arsenic. A limit value is legally binding and must not be exceeded, while a target value is an objective to be attained by a certain date. The EC classifies the air quality zones in each member state as compliant or non-compliant with the limit values or the target values for each pollutant. The EC also requires the member states to report the air quality data and the air quality plans to the European Environment Agency (EEA), which publishes annual reports on the air quality status and trends in Europe.

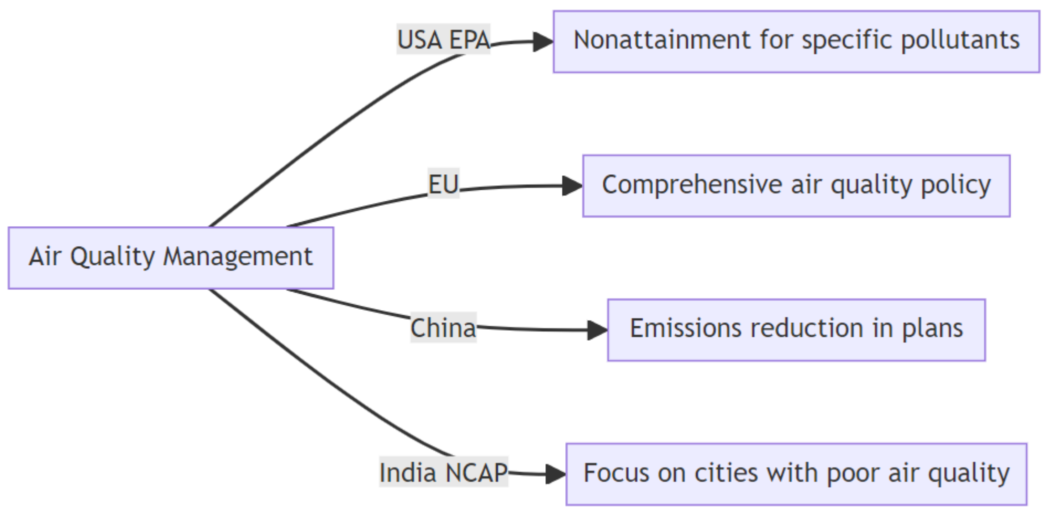


Fig 2: Comparative Air quality management approaches:

3.0.5 3. Comparison with India

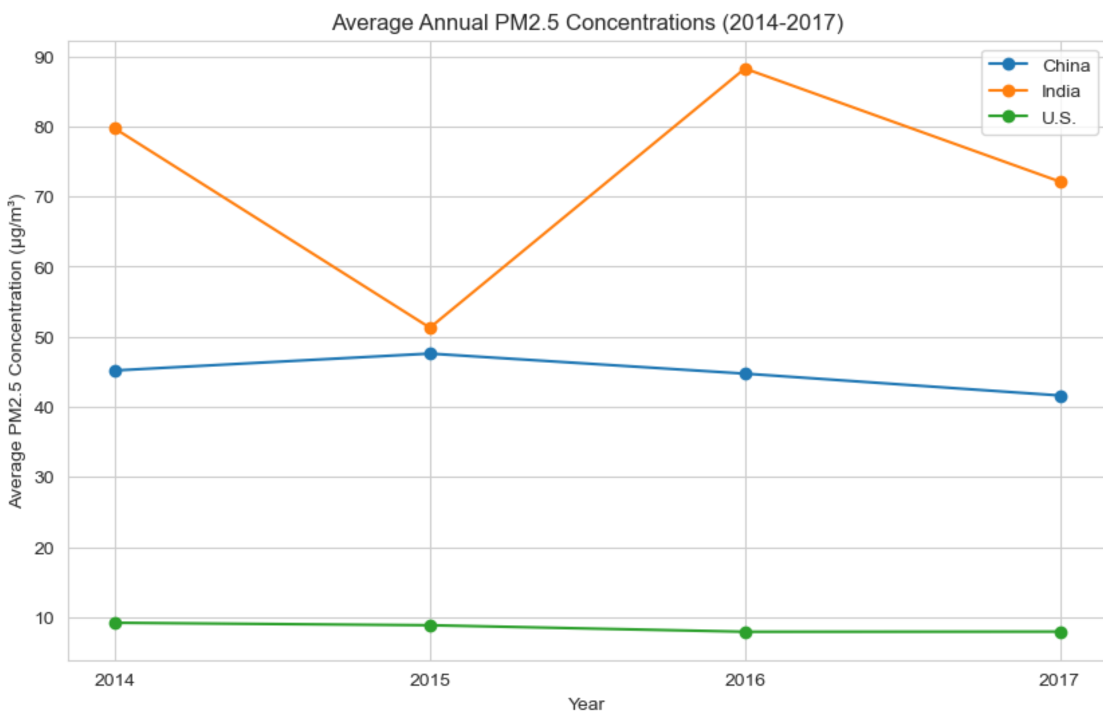
India's NCAP focuses on cities that consistently show air quality worse than the NAAQS for five consecutive years, targeting a reduction in particulate matter concentration. The U.S. designates nonattainment areas based on specific pollutants and compliance with NAAQS. The EU's policy doesn't specifically categorize cities but focuses on overarching air quality standards and pollution reduction across the region. China has integrated air quality targets into its development plans, requiring specific cities to develop air quality attainment plans.

3.0.6 (a) Graphical representation

Fig 3: Line Chart: Comparison of Average Annual PM2.5 Concentrations Overview of the Line Chart X-Axis (Horizontal): Represents the years from 2014 to 2017. Y-Axis (Vertical): Shows the average PM2.5 concentration in micrograms per cubic meter (g/m³). Lines: There are three lines, each representing one of the countries (China, India, and the U.S.). China's Line: Marked with a specific color and a 'o' marker for each year. India's Line: Similarly marked but in a different color. U.S. Line: Also distinctively marked.

[62]:
```
import matplotlib.pyplot as plt

# Actual Data
years = ['2014', '2015', '2016', '2017']
china_mean = [45.19, 47.60, 44.74, 41.62]
india_mean = [79.70, 51.29, 88.26, 72.13]
us_mean = [9.20, 8.85, 7.92, 7.94]

# Plotting
plt.figure(figsize=(10, 6))
plt.plot(years, china_mean, marker='o', label='China')
```

```python
plt.plot(years, india_mean, marker='o', label='India')
plt.plot(years, us_mean, marker='o', label='U.S.')

# Adding titles and labels
plt.title('Average Annual PM2.5 Concentrations (2014-2017)')
plt.xlabel('Year')
plt.ylabel('Average PM2.5 Concentration ( g/m³)')
plt.legend()

# Show plot
plt.show()
```



Interpretation of the line chart: Trends Over Years:

- China: Shows a slight decrease in PM2.5 levels over the years, indicating some improvement in air quality.
- India: Exhibits fluctuation, with a significant drop in 2015, followed by an increase in 2016, and a decrease again in 2017. This suggests variability in air quality control measures or external factors influencing air quality.
- U.S.: Consistently low PM2.5 levels compared to China and India, with a slight downward trend, indicating relatively better air quality conditions.

Comparative Analysis:

- The U.S. consistently maintains lower PM2.5 levels, indicating better air quality management or lower baseline pollution levels.
- China and India show higher PM2.5 concentrations, but China's measures seem to be gradually improving air quality.
- India's air quality shows more variability, which might require more consistent or aggressive pollution control strategies.

Public Health and Policy Implications:

- The high PM2.5 levels in India and China could be associated with higher health risks for the population, necessitating urgent air quality management policies.
- The U.S.'s lower levels are indicative of effective air quality standards and enforcement but still require continuous monitoring and improvement.

Environmental Factors:

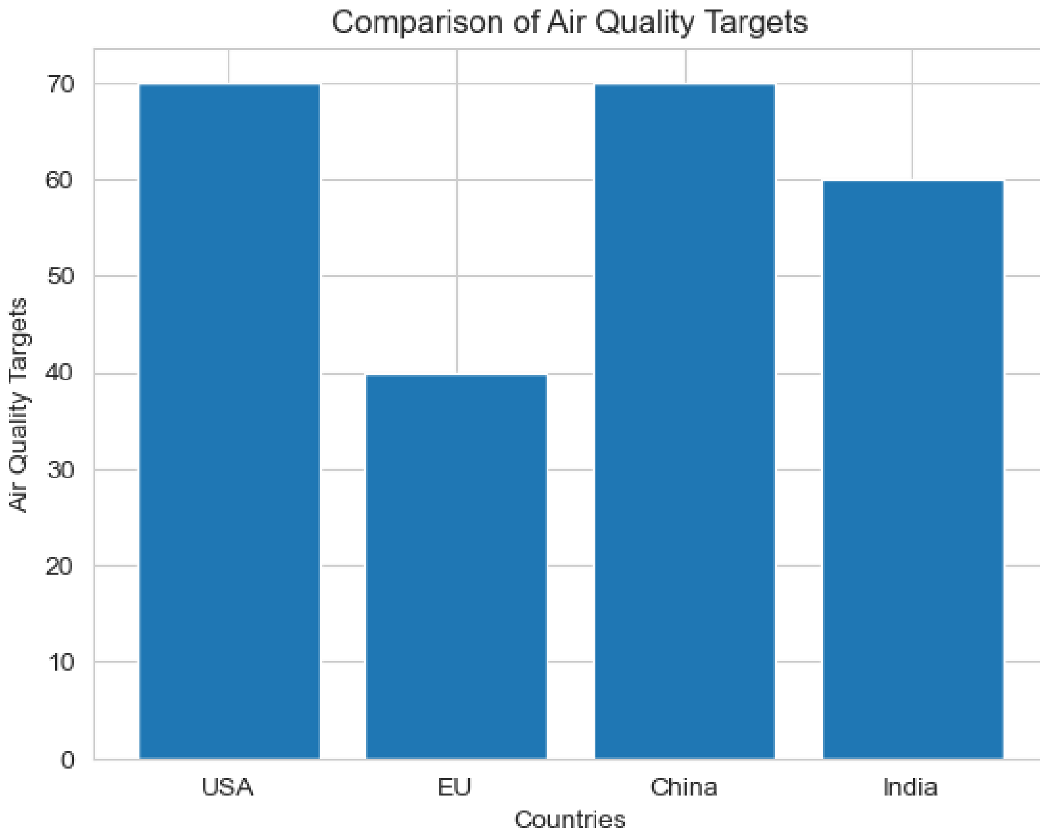
- External factors like industrial activities, vehicular emissions, geographical and meteorological conditions might also play a significant role in these trends.
- Seasonal variations, such as winter heating in China or crop burning in India, might also influence these numbers.

3.0.7 (b) Hypothetical Bar Chart of AQI Targets

`Note: Since different countries use different estimates for targetting, there is ambiguity in estimates of AQI targets for different countries`

Fig 4: Bar Chart for comparaing AQI targets of select countries

[63]:
```python
import matplotlib.pyplot as plt

# Sample data
countries = ['USA', 'EU', 'China', 'India']
values = [70, 40, 70, 60]

plt.bar(countries, values)
plt.xlabel('Countries')
plt.ylabel('Air Quality Targets')
plt.title('Comparison of Air Quality Targets')
plt.show()
```



3.0.8 4. Perspectives and Recommendations

The comparative analysis of air quality management frameworks across different countries, including India, reveals diverse approaches yet a common goal: to mitigate air pollution and safeguard public health and the environment. Countries like the United States, China, and members of the European Union have established their unique criteria and strategies, reflecting their environmental priorities and socio-economic contexts.

Recommendations:

Adopting Best Practices: India can benefit from adopting best practices from these countries. For instance, the stringent air quality standards and robust monitoring systems of the European Union can serve as a model for enhancing India's air quality monitoring infrastructure.

Public Participation and Awareness: The success of air quality management in the United States, partly attributed to public participation, suggests that India should also engage its citizens more actively in air quality management. This can be achieved through awareness campaigns and inclusive policymaking.

Technology Integration: Learning from China's use of advanced technology for air quality monitoring and pollution control, India could explore similar technological solutions, including satellite-based monitoring and AI-driven predictive analysis, to strengthen its air quality management.

Policy Harmonization: Harmonizing state and central policies, akin to the EU's cohesive framework, could ensure a more unified and effective approach to air quality management across Indian states.

Economic Instruments: Implementing economic instruments such as pollution taxes or cap-and-trade systems, as seen in some of these countries, could incentivize industries to adopt cleaner technologies and practices.

3.0.9 5. Conclusion

The global comparison underscores the complexity and urgency of tackling air pollution, a challenge that transcends national boundaries. While India's initiative to designate Non-Attainment Cities is a commendable step towards addressing air pollution, insights from international practices highlight opportunities for further enhancement. By integrating global best practices, leveraging technology, and fostering public engagement, India can strengthen its air quality management framework. This not only aligns with the country's environmental objectives but also contributes to global efforts in combating air pollution and climate change. These recommendations could guide India's journey towards cleaner air and a healthier future.

3.0.10 References

1. Standards for Air Quality Indices in Different Countries - AQI
2. National Ambient Air Quality Standards (NAAQS) - NCBI
3. Press Information Bureau Government of India - NCAP
4. US EPA - Nonattainment Areas for Criteria Pollutants
5. Centre for Energy, Environment & Water (CEEW)
6. National Clean Air Programme (NCAP) Tracker
7. Air Pollution Data of India 2020-2023 - Kaggle

[ ]: