Data Collection, Preprocessing and exploratory data analysis

By DHARSHINI M

Air Quality Data (from year 2015 - 2024)

This dataset shows the air quality from 2015 to 2024. The data tracks harmful pollutants, Air Quality Index (AQI), and how clean or polluted the air is at both the city and station levels. You can use it to analyze trends, find pollution hotspots, or study seasonal air changes. Context Air is what keeps humans alive. Monitoring it and understanding its quality is of immense importance to our well-being. The dataset contains air quality data and AQI (Air Quality Index) at hourly and daily level of various stations across multiple cities in India.

Dataset source: https://www.kaggle.com/datasets/ankushpanday1/air-quality-data-in-india-2015-2024/data

In [1]:
```
import pandas as pd
```

In [2]:
```
city_day_data=pd.read_csv('/content/city_day.csv')
city_hour_data=pd.read_csv('/content/city_hour.csv')
station_day_data=pd.read_csv('/content/station_day.csv')
station_hour_data=pd.read_csv('/content/station_hour.csv')
stations_data=pd.read_csv('/content/stations.csv')
```

In [3]:
```
def print_first_5_columns(data):
  for i in data:
    print(data.head())
    break
```

```
print_first_5_columns(city_hour_data)
```

```
       City             Datetime  PM2.5   PM10     NO   NO2    NOx   NH3  \
0      Delhi  2015-01-01 00:00:00  258.0  340.6  191.0  13.4  104.1  16.2
1     Mumbai  2015-01-01 00:00:00  120.1   47.9  165.3  57.9  139.2  14.4
2    Chennai  2015-01-01 00:00:00  130.1  375.0   21.9  23.7  205.6  33.6
3    Kolkata  2015-01-01 00:00:00  189.1  174.3  139.7  58.9  103.7  19.2
4  Bangalore  2015-01-01 00:00:00  357.8   48.8  121.2  83.3    8.3  11.9

     CO   SO2     O3  Benzene  Toluene  Xylene    AQI AQI_Bucket
0  1.15  39.8   70.4    11.64    10.23    1.95  411.5   Moderate
1  0.76   3.3   23.2    11.35    11.38    1.53  134.8       Poor
2  3.50  63.0  138.9    16.30    24.07    0.56  329.6       Poor
3  0.81   3.3   14.6     5.12     1.42    2.35  351.9       Good
4  6.31  67.4  195.9    15.12    16.80    8.95   23.7   Moderate
```

In [4]:
```
print_first_5_columns(city_day_data)
```

```
       City    Datetime  PM2.5   PM10     NO   NO2    NOx   NH3    CO   SO2  \
0      Delhi  2015-01-01  153.3  241.7  182.9  33.0   81.3  38.5  1.87  64.5
1     Mumbai  2015-01-01   70.5  312.7  195.0  42.0  122.5  31.5  7.22  83.8
2    Chennai  2015-01-01  174.1  275.4   56.2  68.8  230.9  28.5  8.56  60.8
3    Kolkata  2015-01-01  477.2  543.9   14.1  76.4  225.9  45.6  2.41  42.1
4  Bangalore  2015-01-01  171.6  117.7  123.3  12.4   61.9  49.7  1.26  79.7

      O3  Benzene  Toluene  Xylene    AQI    AQI_Bucket
0   83.6    18.93    20.81    8.32  204.5        Severe
1  108.0     2.01    19.41    2.86   60.9  Satisfactory
2   43.9    19.07    10.19    9.63  486.5        Severe
3  171.1     9.31    11.65    9.39  174.4     Very Poor
4  164.3     6.04    12.74    9.59  489.7          Good
```

In [5]:
```
print_first_5_columns(station_hour_data)
```

```
      City             Datetime     Station  PM2.5   PM10     NO    NO2  \
0    Delhi  2015-01-01 00:00:00  Station_D1  243.1  193.5  182.4   58.8
1    Delhi  2015-01-01 00:00:00  Station_D2  476.7  504.0  172.4   80.7
2   Mumbai  2015-01-01 00:00:00  Station_M1  463.9   92.7  142.5   21.7
3   Mumbai  2015-01-01 00:00:00  Station_M2  447.1  545.1   85.5  124.6
4  Chennai  2015-01-01 00:00:00  Station_C1  117.6  390.6  149.5  115.4

     NOx   NH3    CO   SO2     O3  Benzene  Toluene  Xylene    AQI  \
0  100.9  30.4  3.08  24.3  166.7    12.66    18.42    1.91  301.5
1  114.4  30.3  7.96  48.1   72.9     8.35    26.53    1.66  243.1
2  130.2  11.3  4.89  99.5  166.5    15.96    24.41    7.88   14.7
3   66.6  21.3  3.92  61.6  109.3    19.34     6.09    5.37  471.7
4  222.4  25.2  4.85  15.1  198.3     0.20     6.49    1.69  312.3

     AQI_Bucket
0          Good
1  Satisfactory
2        Severe
3     Very Poor
4  Satisfactory
```

In [6]:
```
print_first_5_columns(station_day_data)
```

```
      City    Datetime     Station  PM2.5   PM10     NO   NO2    NOx   NH3  \
0    Delhi  2015-01-01  Station_D1   58.3  223.2  126.6  85.5  207.5  18.4
1    Delhi  2015-01-01  Station_D2  222.5  541.4  198.6  20.8    0.6  25.5
2   Mumbai  2015-01-01  Station_M1   36.6  160.8  164.2  18.6   94.8   8.9
3   Mumbai  2015-01-01  Station_M2  368.8  526.2  140.6  74.4  153.2  15.2
4  Chennai  2015-01-01  Station_C1  188.6   88.8   87.5  58.0    6.1  48.3

     CO   SO2     O3  Benzene  Toluene  Xylene    AQI AQI_Bucket
0  0.25  42.7  184.9     8.03    12.72    4.10  266.4   Moderate
1  6.21  95.0   76.2    17.69    13.85    2.79  185.3       Poor
2  6.84  70.7  195.4     5.53    26.93    2.64    5.9   Moderate
3  4.03  61.0  164.2     1.62     9.35    3.27  176.7       Poor
4  9.66  46.2   17.6     2.31     7.55    9.52  171.9       Good
```

In [7]:
```
print_first_5_columns(stations_data)
```

```
      City     Station
0    Delhi  Station_D1
1    Delhi  Station_D2
2   Mumbai  Station_M1
3   Mumbai  Station_M2
4  Chennai  Station_C1
```

In [8]:
```
def print_info(data):
  for i in data:
    print(data.info())
    break
print_info(city_day_data)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18265 entries, 0 to 18264
Data columns (total 16 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   City        18265 non-null  object 
 1   Datetime    18265 non-null  object 
 2   PM2.5       18265 non-null  float64
 3   PM10        18265 non-null  float64
 4   NO          18265 non-null  float64
 5   NO2         18265 non-null  float64
 6   NOx         18265 non-null  float64
 7   NH3         18265 non-null  float64
 8   CO          18265 non-null  float64
 9   SO2         18265 non-null  float64
 10  O3          18265 non-null  float64
 11  Benzene     18265 non-null  float64
 12  Toluene     18265 non-null  float64
 13  Xylene      18265 non-null  float64
 14  AQI         18265 non-null  float64
 15  AQI_Bucket  18265 non-null  object 
dtypes: float64(13), object(3)
memory usage: 2.2+ MB
None
```

In [9]:
```
print_info(city_hour_data)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 438245 entries, 0 to 438244
Data columns (total 16 columns):
 #   Column      Non-Null Count   Dtype  
---  ------      --------------   -----  
 0   City        438245 non-null  object 
 1   Datetime    438245 non-null  object 
 2   PM2.5       438245 non-null  float64
 3   PM10        438245 non-null  float64
 4   NO          438245 non-null  float64
 5   NO2         438245 non-null  float64
 6   NOx         438245 non-null  float64
 7   NH3         438245 non-null  float64
 8   CO          438245 non-null  float64
 9   SO2         438245 non-null  float64
 10  O3          438245 non-null  float64
 11  Benzene     438245 non-null  float64
 12  Toluene     438245 non-null  float64
 13  Xylene      438245 non-null  float64
 14  AQI         438245 non-null  float64
 15  AQI_Bucket  438245 non-null  object 
dtypes: float64(13), object(3)
memory usage: 53.5+ MB
None
```

In [10]:
```
print_info(station_hour_data)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7210 entries, 0 to 7209
Data columns (total 17 columns):
 #   Column      Non-Null Count  Dtype
---  ------      --------------  -----
 0   City        7210 non-null   object
 1   Datetime    7210 non-null   object
 2   Station     7210 non-null   object
 3   PM2.5       7210 non-null   float64
 4   PM10        7210 non-null   float64
 5   NO          7210 non-null   float64
 6   NO2         7210 non-null   float64
 7   NOx         7210 non-null   float64
 8   NH3         7210 non-null   float64
 9   CO          7210 non-null   float64
 10  SO2         7210 non-null   float64
 11  O3          7210 non-null   float64
 12  Benzene     7210 non-null   float64
 13  Toluene     7210 non-null   float64
 14  Xylene      7210 non-null   float64
 15  AQI         7210 non-null   float64
 16  AQI_Bucket  7210 non-null   object
dtypes: float64(13), object(4)
memory usage: 957.7+ KB
None
```

In [11]:
```
print_info(station_day_data)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 36530 entries, 0 to 36529
Data columns (total 17 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   City        36530 non-null  object 
 1   Datetime    36530 non-null  object 
 2   Station     36530 non-null  object 
 3   PM2.5       36530 non-null  float64
 4   PM10        36530 non-null  float64
 5   NO          36530 non-null  float64
 6   NO2         36530 non-null  float64
 7   NOx         36530 non-null  float64
 8   NH3         36530 non-null  float64
 9   CO          36530 non-null  float64
 10  SO2         36530 non-null  float64
 11  O3          36530 non-null  float64
 12  Benzene     36530 non-null  float64
 13  Toluene     36530 non-null  float64
 14  Xylene      36530 non-null  float64
 15  AQI         36530 non-null  float64
 16  AQI_Bucket  36530 non-null  object 
dtypes: float64(13), object(4)
memory usage: 4.7+ MB
None
```

In [12]:
```
print_info(stations_data)
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 2 columns):
 #   Column   Non-Null Count  Dtype 
---  ------   --------------  ----- 
 0   City     10 non-null     object
 1   Station  10 non-null     object
dtypes: object(2)
memory usage: 292.0+ bytes
None
```

In [13]:
```
def describe_data(data):
  for i in data:
    print(data.describe())
```

```
    break

describe_data(city_day_data)
```

```
              PM2.5          PM10            NO           NO2           NOx  \
count  18265.000000  18265.000000  18265.000000  18265.000000  18265.000000
mean     250.597695    299.442491    100.481035     75.415916    125.964079
std      144.460292    173.479906     57.774795     43.460066     72.403893
min        0.000000      0.000000      0.000000      0.000000      0.000000
25%      125.700000    150.100000     50.600000     37.700000     63.100000
50%      251.000000    300.300000    100.200000     76.000000    126.200000
75%      376.200000    450.000000    151.000000    113.200000    188.900000
max      499.900000    600.000000    200.000000    150.000000    250.000000

                NH3            CO           SO2            O3       Benzene  \
count  18265.000000  18265.000000  18265.000000  18265.000000  18265.000000
mean      25.065042      5.002451     49.835839    100.406740     10.070033
std       14.452019      2.889439     28.988739     57.591436      5.785282
min        0.000000      0.000000      0.000000      0.000000      0.000000
25%       12.600000      2.490000     24.400000     50.600000      5.080000
50%       25.300000      5.000000     49.900000    100.700000     10.080000
75%       37.600000      7.510000     75.100000    150.400000     15.110000
max       50.000000     10.000000    100.000000    200.000000     20.000000

             Toluene        Xylene           AQI
count  18265.000000  18265.000000  18265.000000
mean      15.063365      4.996341    251.111382
std        8.619433      2.899152    144.502626
min        0.000000      0.000000      0.000000
25%        7.640000      2.510000    125.400000
50%       15.130000      4.960000    251.200000
75%       22.500000      7.530000    376.400000
max       30.000000     10.000000    500.000000
```

In [14]:
```
describe_data(city_hour_data)
```

```
              PM2.5           PM10             NO            NO2  \
count  438245.000000  438245.000000  438245.000000  438245.000000
mean      249.418096     299.639617     100.077829      75.071440
std       144.360171     173.399412      57.647433      43.258969
min         0.000000       0.000000       0.000000       0.000000
25%       124.400000     149.000000      50.200000      37.700000
50%       249.100000     299.400000     100.000000      75.100000
75%       374.300000     450.200000     150.000000     112.500000
max       500.000000     600.000000     200.000000     150.000000

                NOx            NH3             CO            SO2  \
count  438245.000000  438245.000000  438245.000000  438245.000000
mean      125.039725      25.021188       5.002301      50.088662
std        72.180907      14.420788       2.885777      28.887099
min         0.000000       0.000000       0.000000       0.000000
25%        62.400000      12.500000       2.500000      25.100000
50%       125.200000      25.000000       5.000000      50.100000
75%       187.600000      37.500000       7.500000      75.200000
max       250.000000      50.000000      10.000000     100.000000

                O3        Benzene        Toluene         Xylene  \
count  438245.000000  438245.000000  438245.000000  438245.000000
mean       99.982519       9.993992      14.995036       4.994013
std        57.759809       5.772458       8.653210       2.889434
min         0.000000       0.000000       0.000000       0.000000
25%        50.000000       5.000000       7.500000       2.490000
50%        99.900000       9.980000      14.990000       4.990000
75%       150.000000      14.990000      22.470000       7.490000
max       200.000000      20.000000      30.000000      10.000000

                AQI
count  438245.000000
mean      250.095557
std       144.129039
min         0.000000
25%       125.400000
50%       250.000000
75%       374.900000
max       500.000000
```

In [15]:
```
describe_data(station_day_data)
```

```
              PM2.5          PM10            NO           NO2           NOx  \
count  36530.000000  36530.000000  36530.000000  36530.000000  36530.000000
mean     251.062562    299.846315     99.872672     75.116006    124.656512
std      144.238278    172.750276     57.685717     43.199727     72.264938
min        0.000000      0.000000      0.000000      0.000000      0.000000
25%      126.400000    149.600000     50.200000     38.000000     62.025000
50%      252.600000    298.400000     99.500000     75.300000    124.400000
75%      376.200000    450.000000    149.900000    112.200000    186.800000
max      500.000000    600.000000    200.000000    150.000000    250.000000

                NH3            CO           SO2            O3       Benzene  \
count  36530.000000  36530.000000  36530.000000  36530.000000  36530.000000
mean      25.075160      4.974892     49.852111    100.294615      9.991377
std       14.368798      2.888939     28.883560     57.874168      5.784603
min        0.000000      0.000000      0.000000      0.000000      0.000000
25%       12.600000      2.470000     24.900000     50.300000      4.950000
50%       25.200000      4.960000     49.800000    100.300000      9.975000
75%       37.500000      7.480000     74.700000    150.600000     15.030000
max       50.000000     10.000000    100.000000    200.000000     20.000000

             Toluene        Xylene           AQI
count  36530.000000  36530.000000  36530.000000
mean      15.112225      5.012123    250.291005
std        8.667173      2.886312    144.192988
min        0.000000      0.000000      0.000000
25%        7.650000      2.510000    125.400000
50%       15.120000      5.030000    250.800000
75%       22.650000      7.510000    374.700000
max       30.000000     10.000000    500.000000
```

In [16]:
```
describe_data(station_hour_data)
```

```
              PM2.5         PM10           NO          NO2          NOx  \
count  7210.000000  7210.000000  7210.000000  7210.000000  7210.000000
mean    251.119501   296.500374   101.430180    75.853135   125.027365
std     143.121895   173.714070    57.441876    43.202037    72.166481
min       0.000000     0.000000     0.000000     0.000000     0.000000
25%     128.400000   145.200000    52.400000    38.425000    62.125000
50%     250.850000   294.050000   102.750000    75.600000   124.400000
75%     376.275000   445.875000   151.575000   113.700000   188.575000
max     499.700000   599.900000   200.000000   150.000000   250.000000

                NH3           CO          SO2           O3      Benzene  \
count  7210.000000  7210.000000  7210.000000  7210.000000  7210.000000
mean     24.859764     5.027320    50.044535   100.186602    10.018950
std      14.483497     2.878153    29.153902    57.770853     5.775424
min       0.000000     0.000000     0.000000     0.000000     0.000000
25%      12.325000     2.540000    24.500000    50.600000     4.972500
50%      24.800000     5.050000    50.200000    99.000000    10.020000
75%      37.500000     7.520000    75.300000   150.800000    15.030000
max      50.000000    10.000000   100.000000   199.900000    20.000000

           Toluene       Xylene          AQI
count  7210.000000  7210.000000  7210.000000
mean     14.966846     5.019162   249.699168
std       8.611854     2.883486   143.926543
min       0.000000     0.010000     0.000000
25%       7.592500     2.520000   127.100000
50%      14.970000     4.990000   250.200000
75%      22.300000     7.510000   372.975000
max      30.000000    10.000000   500.000000
```

In [17]:
```
describe_data(stations_data)
```

```
         City     Station
count      10          10
unique      5          10
top     Delhi  Station_D1
freq        2           1
```

In [18]:
```
def check_null_values(data):
  for i in data:
    return data.isnull().sum()
```

```
check_null_values(city_day_data)
```

Out[18]:

	0
City	0
Datetime	0
PM2.5	0
PM10	0
NO	0
NO2	0
NOx	0
NH3	0
CO	0
SO2	0
O3	0
Benzene	0
Toluene	0
Xylene	0
AQI	0
AQI_Bucket	0

dtype: int64

In [19]:
```
check_null_values(city_hour_data)
```

Out[19]:

	0
City	0
Datetime	0
PM2.5	0
PM10	0
NO	0
NO2	0
NOx	0
NH3	0
CO	0
SO2	0
O3	0
Benzene	0
Toluene	0
Xylene	0
AQI	0
AQI_Bucket	0

dtype: int64

In [20]:
```
check_null_values(station_day_data)
```

Out[20]:

	0
City	0
Datetime	0
Station	0
PM2.5	0
PM10	0
NO	0
NO2	0
NOx	0
NH3	0
CO	0
SO2	0
O3	0
Benzene	0
Toluene	0
Xylene	0
AQI	0
AQI_Bucket	0

dtype: int64

In [21]:
```
check_null_values(station_hour_data)
```

Out[21]:

	0
City	0
Datetime	0
Station	0
PM2.5	0
PM10	0
NO	0
NO2	0
NOx	0
NH3	0
CO	0
SO2	0
O3	0
Benzene	0
Toluene	0
Xylene	0
AQI	0
AQI_Bucket	0

dtype: int64

In [22]:
```
check_null_values(stations_data)
```

Out[22]:

	0
City	0
Station	0

dtype: int64

In [23]:
```
def check_for_duplicates(data):
  for i in data:
    data.duplicated().sum()

print(check_for_duplicates(city_hour_data))
```

None

In [24]:
```
print(check_for_duplicates(city_day_data))
```

None

In [25]:
```
print(check_for_duplicates(station_day_data))
```

None

In [26]:
```
print(check_for_duplicates(station_hour_data))
```

None

In [27]:
```
print(check_for_duplicates(stations_data))
```

None

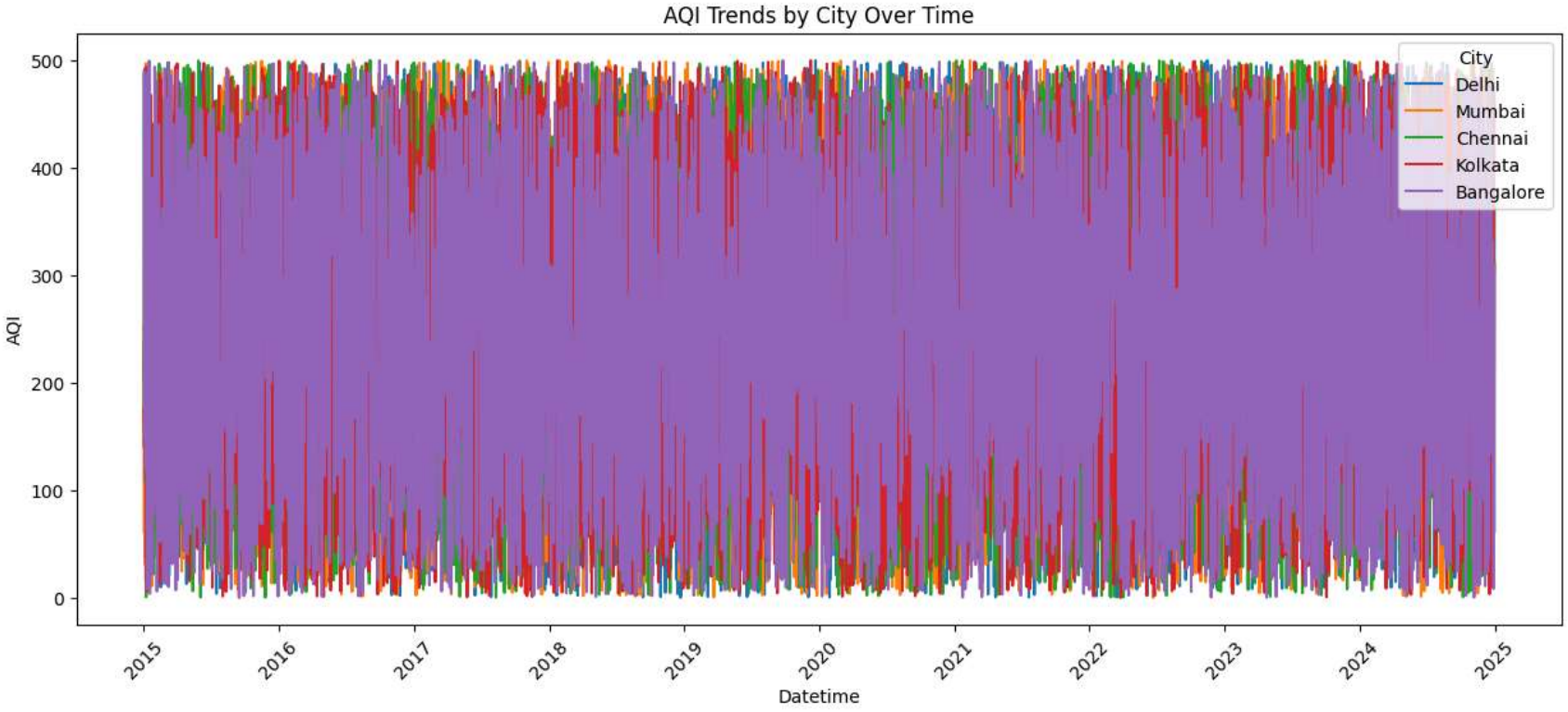
Visualizations

Temporal Analysis

In [41]:
```
import matplotlib.pyplot as plt
import seaborn as sns
# Convert datetime columns
for df in [city_day_data, city_hour_data, station_day_data, station_hour_data]:
    df['Datetime'] = pd.to_datetime(df['Datetime'])
```

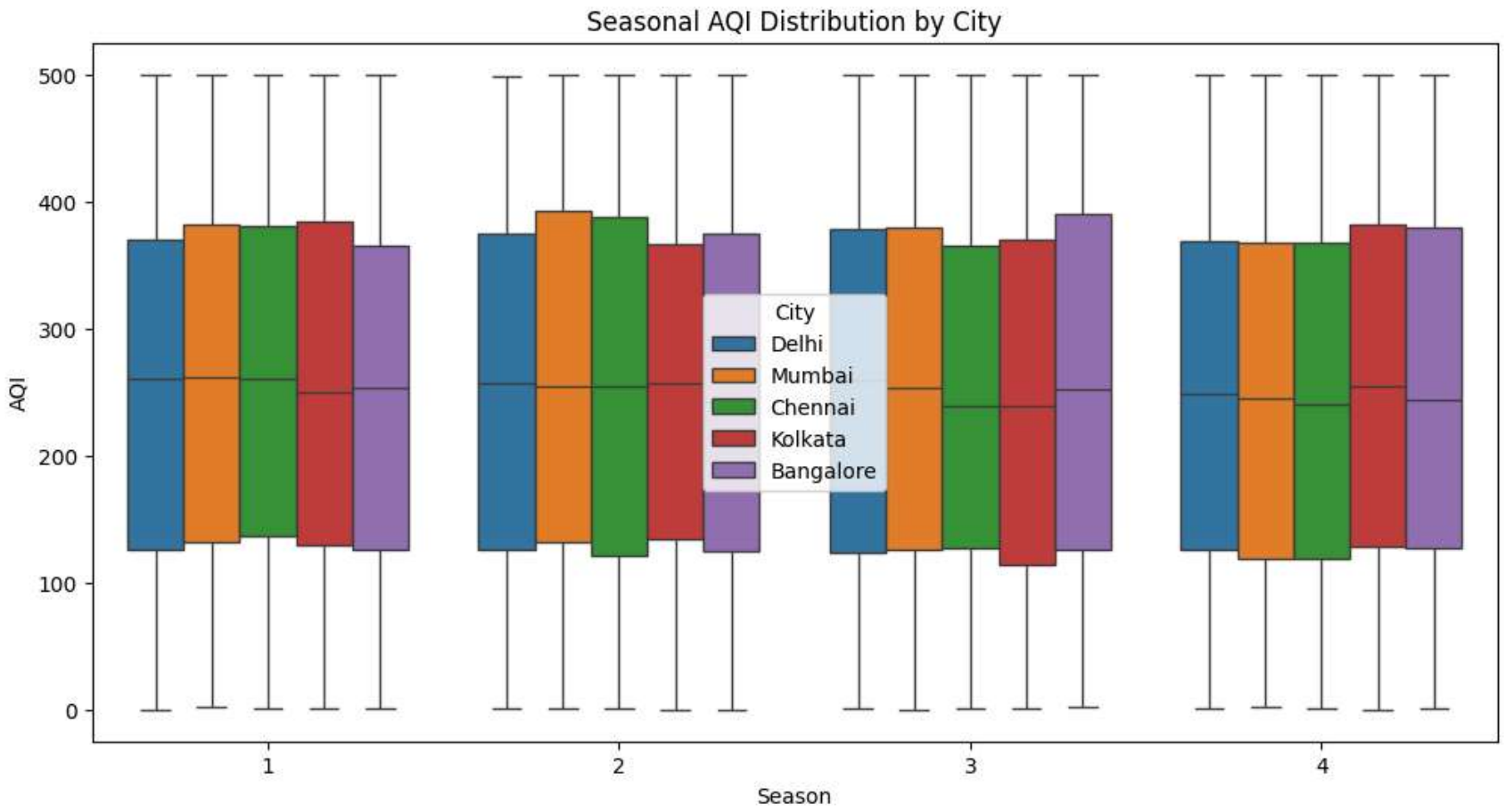
```
    # Extract time components
    df['Year'] = df['Datetime'].dt.year
    df['Month'] = df['Datetime'].dt.month
    df['Day'] = df['Datetime'].dt.day
    df['Hour'] = df['Datetime'].dt.hour  # For hourly data
    df['DayOfWeek'] = df['Datetime'].dt.dayofweek
    df['Season'] = df['Month'] % 12 // 3 + 1

# Plot AQI trends over time
plt.figure(figsize=(15, 6))
sns.lineplot(data=city_day_data, x='Datetime', y='AQI', hue='City')
plt.title('AQI Trends by City Over Time')
plt.xticks(rotation=45)
plt.show()
```

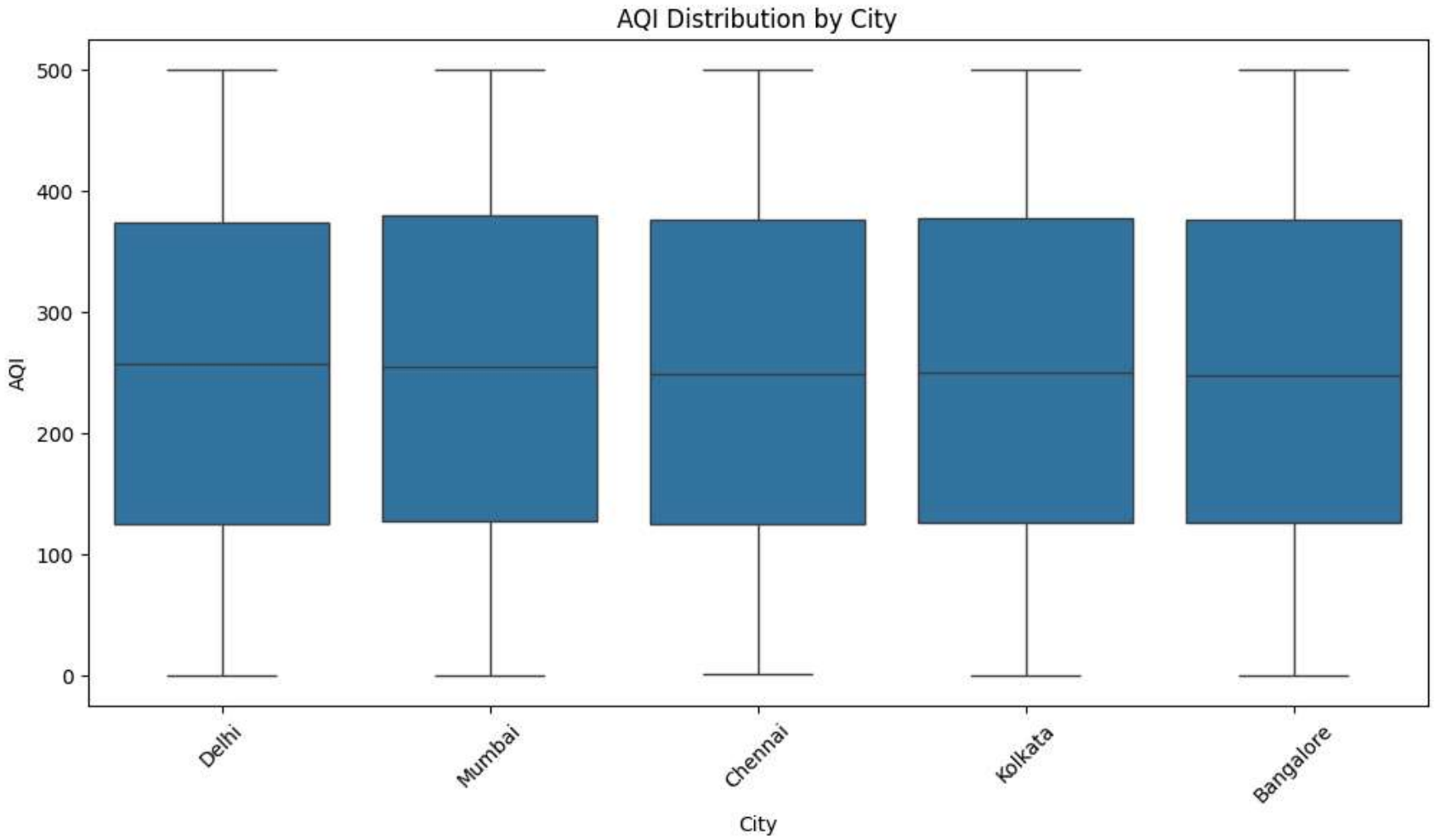


In [42]:
```
# Seasonal analysis
plt.figure(figsize=(12, 6))
sns.boxplot(data=city_day_data, x='Season', y='AQI', hue='City')
```

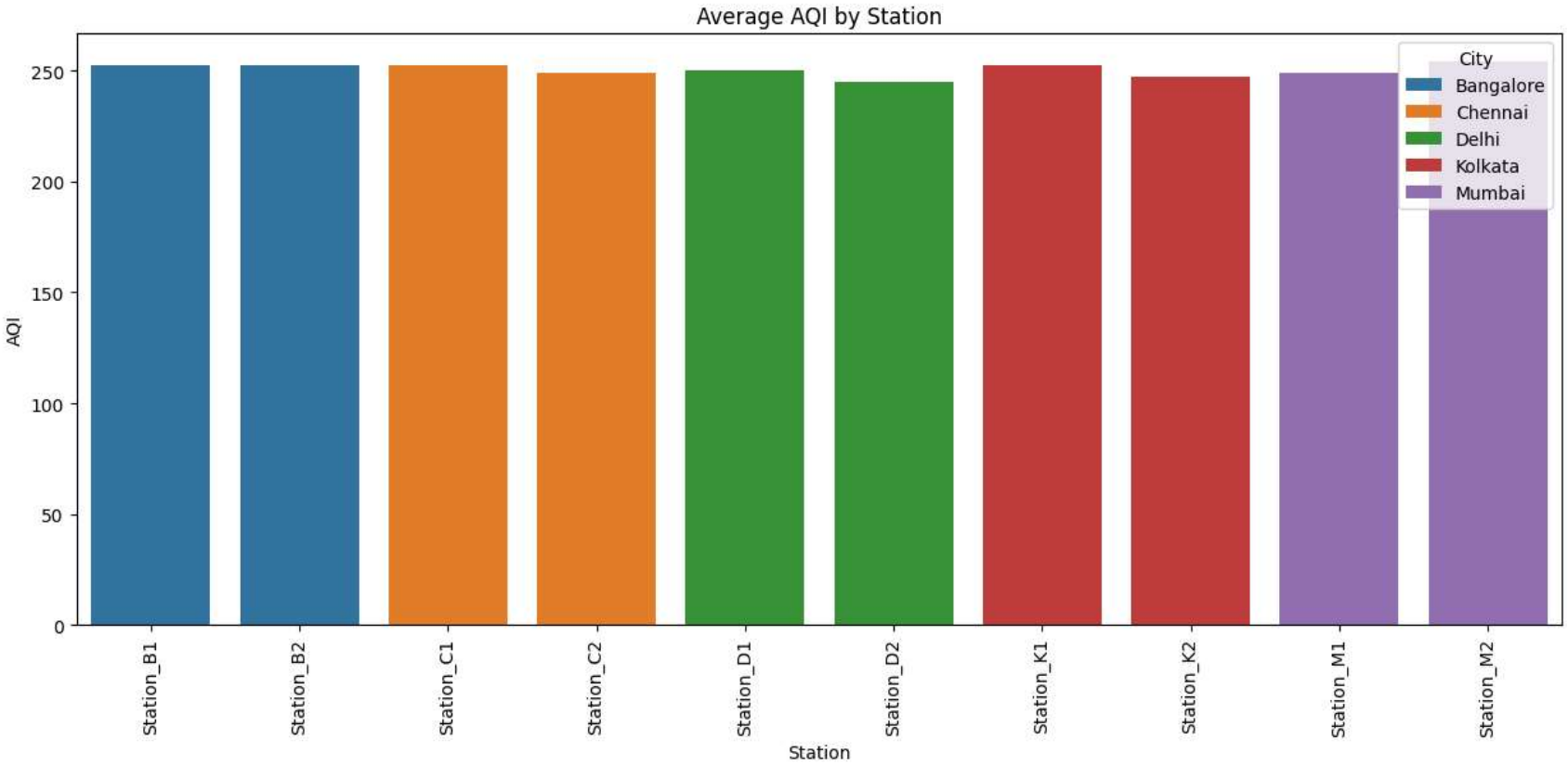
```
plt.title('Seasonal AQI Distribution by City')
plt.show()
```



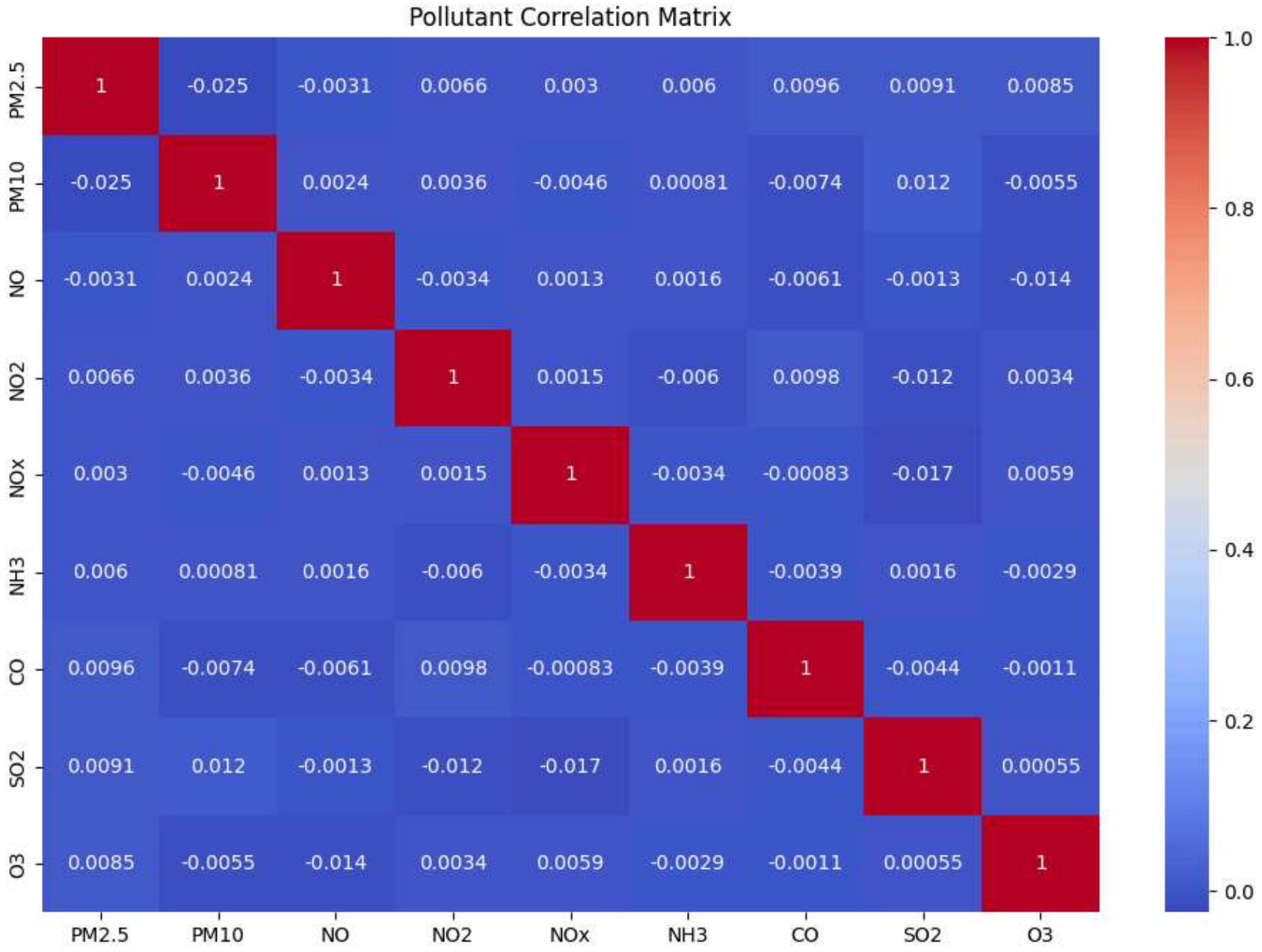
In [40]:
```
# City-wise comparison
plt.figure(figsize=(12, 6))
sns.boxplot(data=city_day_data, x='City', y='AQI')
plt.title('AQI Distribution by City')
plt.xticks(rotation=45)
plt.show()
```



In [39]:
```
# Station analysis (merge station data with measurements)
station_aqi = station_day_data.groupby(['City', 'Station'])['AQI'].mean().reset_index()
plt.figure(figsize=(15, 6))
sns.barplot(data=station_aqi, x='Station', y='AQI', hue='City')
plt.title('Average AQI by Station')
plt.xticks(rotation=90)
plt.show()
```

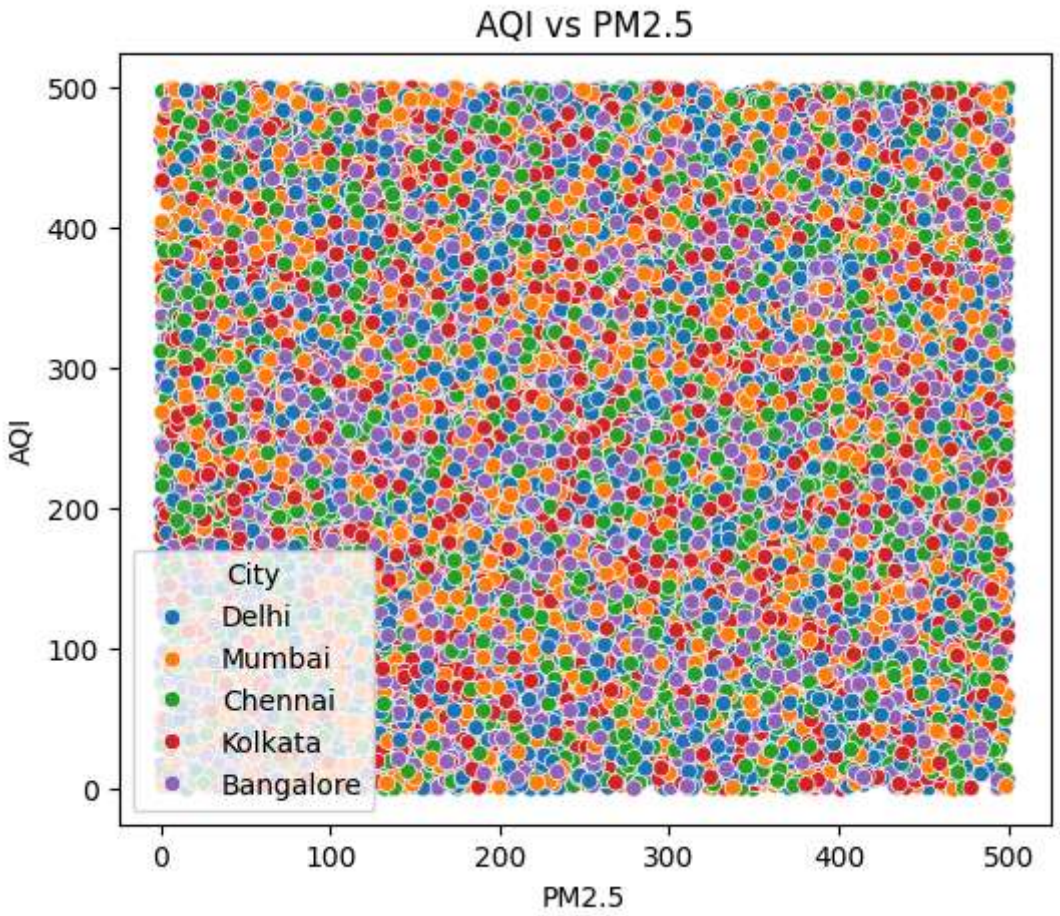


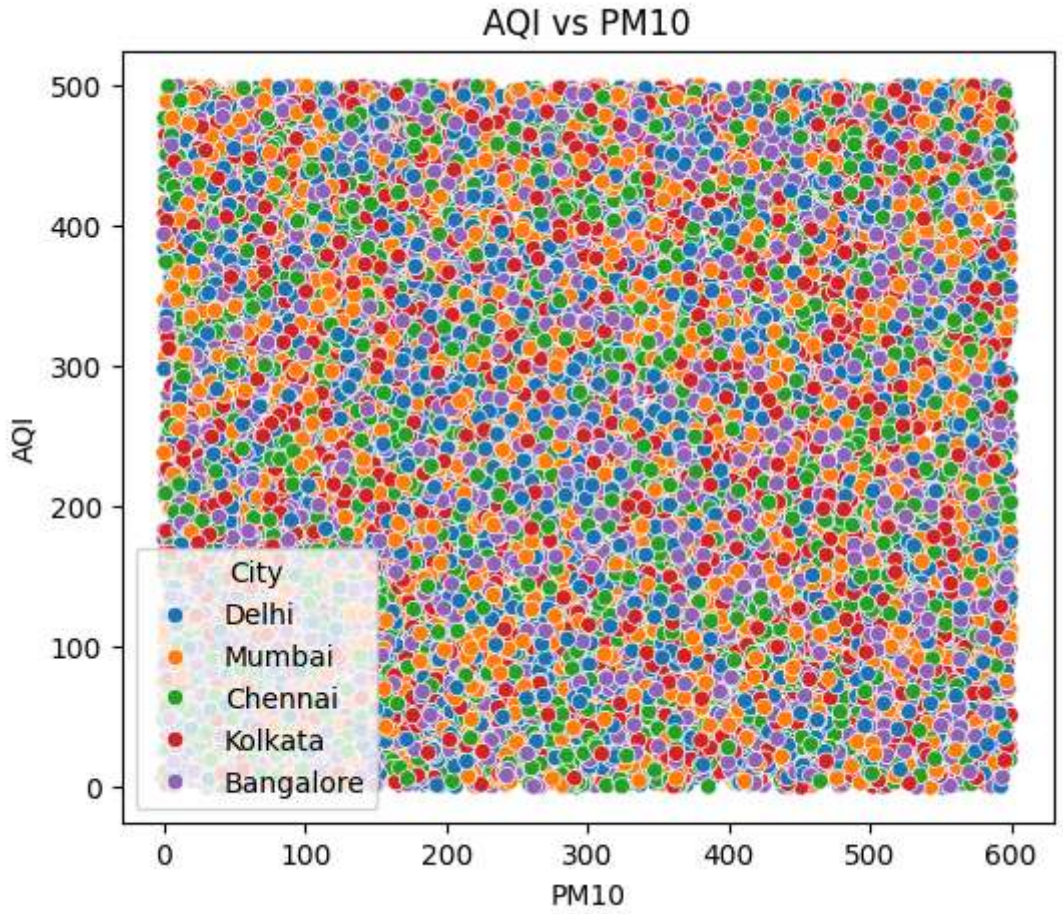
In [38]:
```
# Correlation matrix for pollutants
pollutants = ['PM2.5', 'PM10', 'NO', 'NO2', 'NOx', 'NH3', 'CO', 'SO2', 'O3']
plt.figure(figsize=(12, 8))
sns.heatmap(city_day_data[pollutants].corr(), annot=True, cmap='coolwarm')
plt.title('Pollutant Correlation Matrix')
plt.show()
```



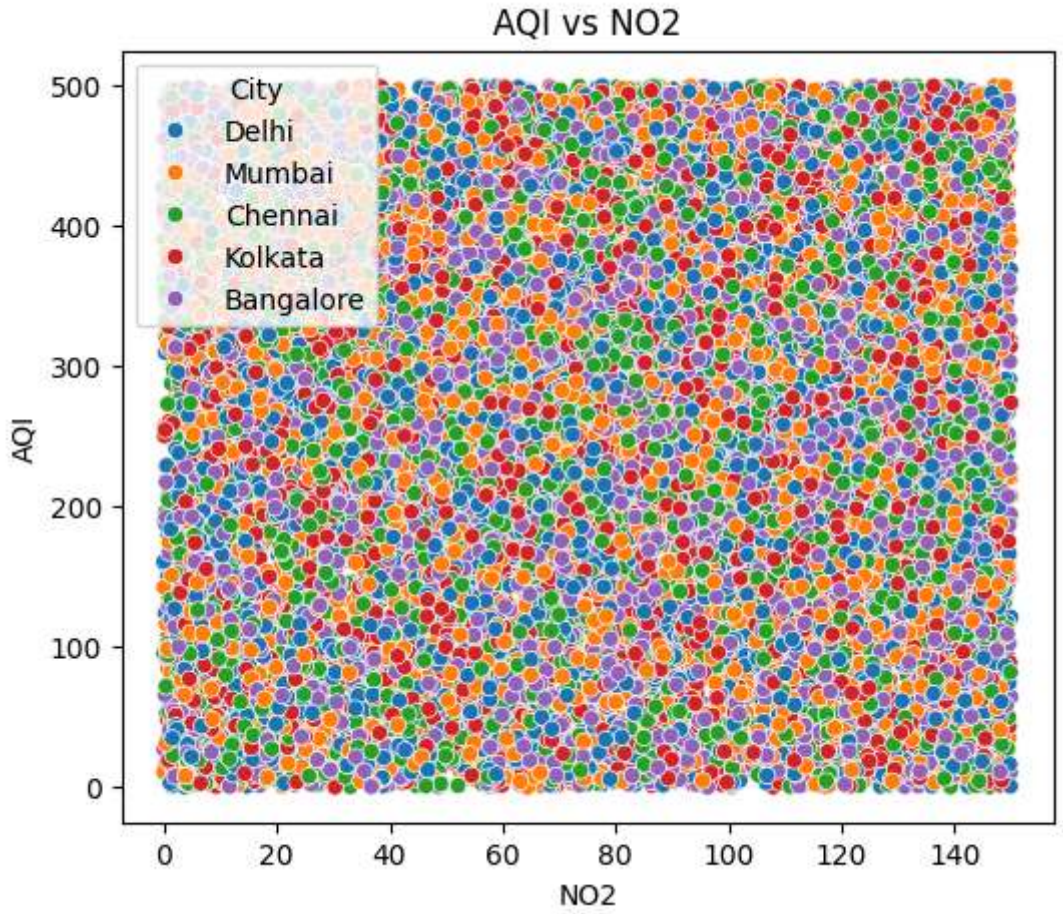
In [37]:
```
# Pollutant contribution to AQI
for pollutant in pollutants:
    plt.figure(figsize=(6, 5))
```

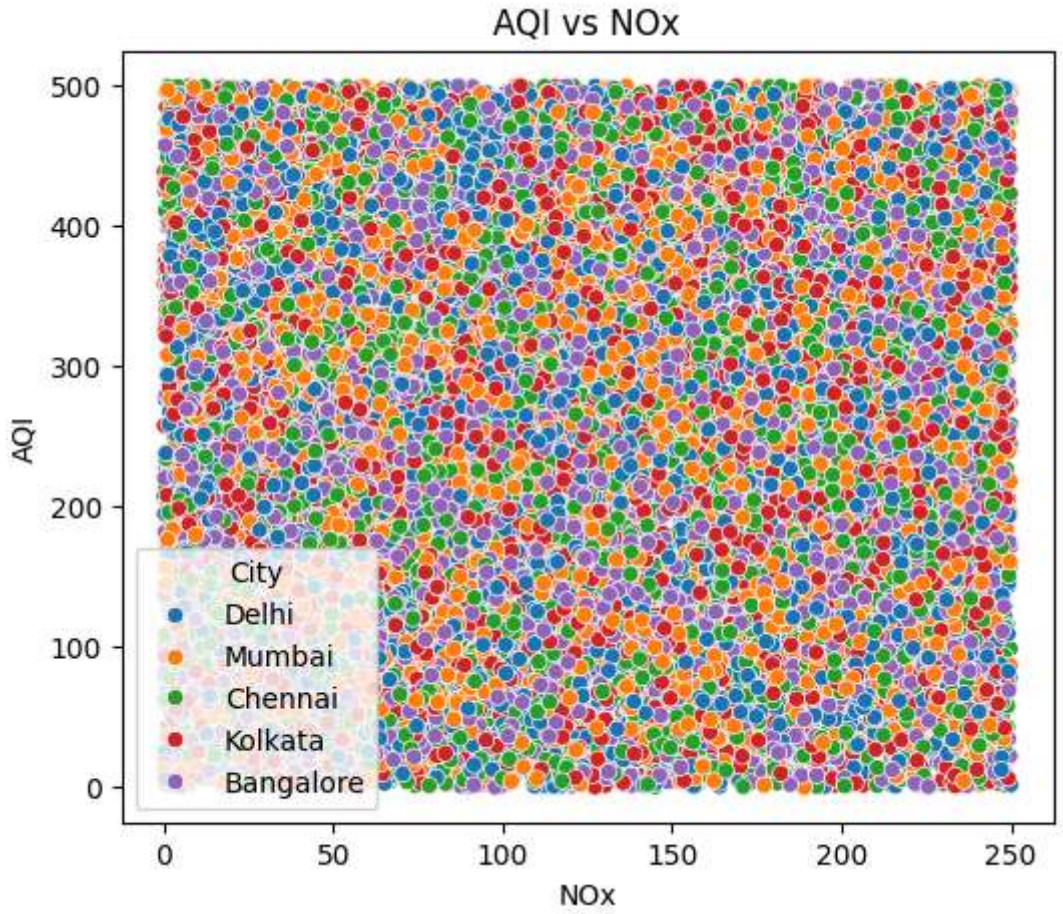
```
    sns.scatterplot(data=city_day_data, x=pollutant, y='AQI', hue='City')
    plt.title(f'AQI vs {pollutant}')
    plt.show()
```

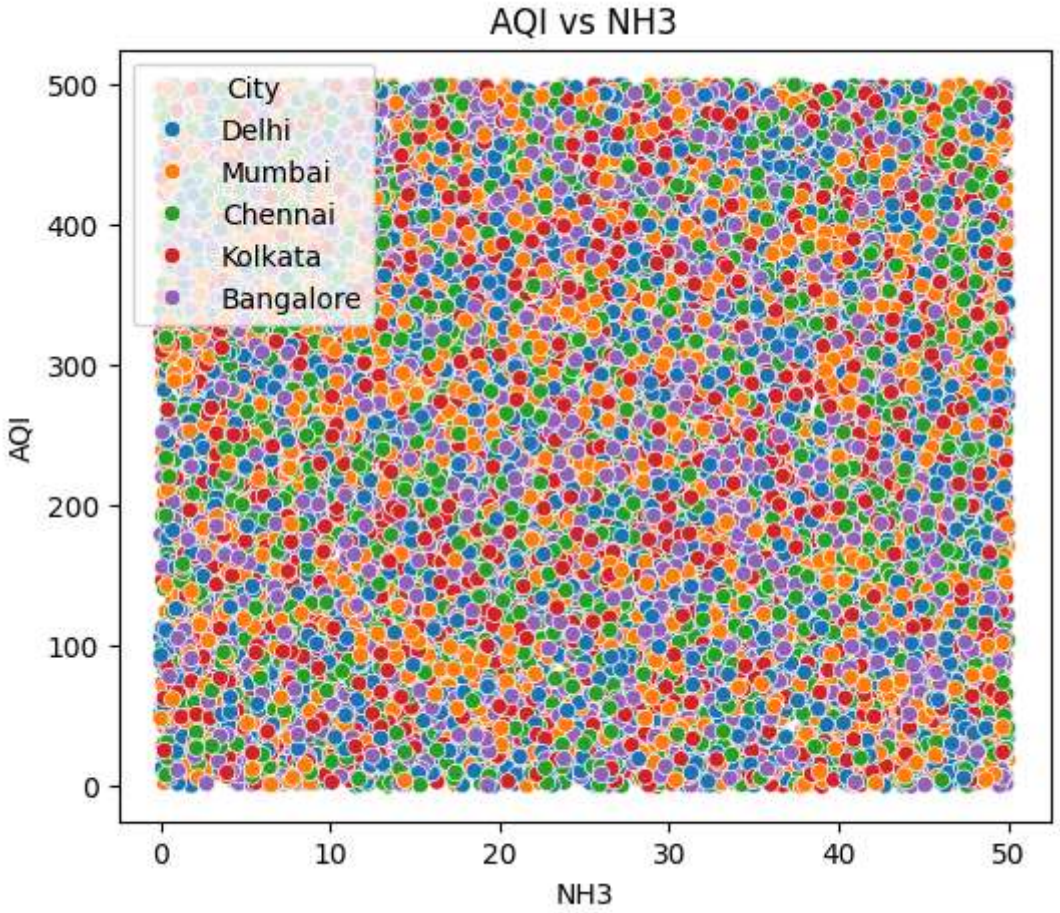


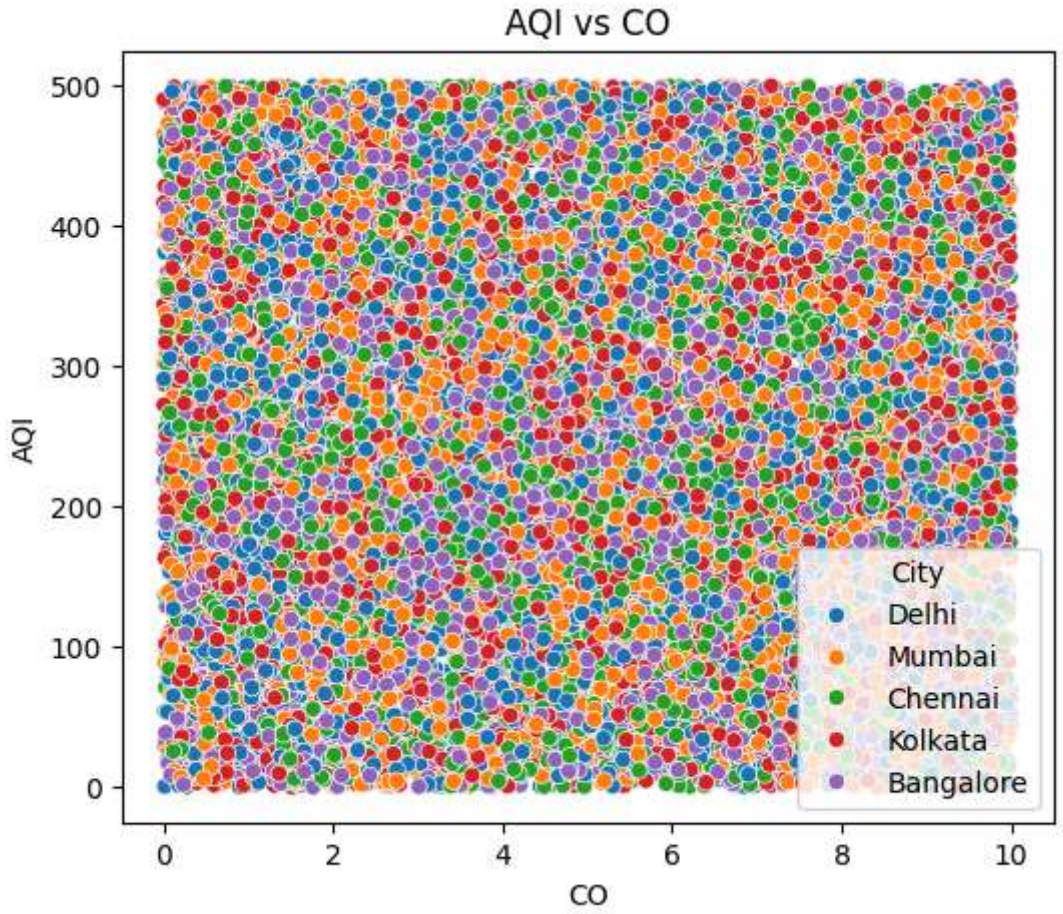

AQI vs PM10
AQI
PM10
0
100
200
300
400
500
600
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore

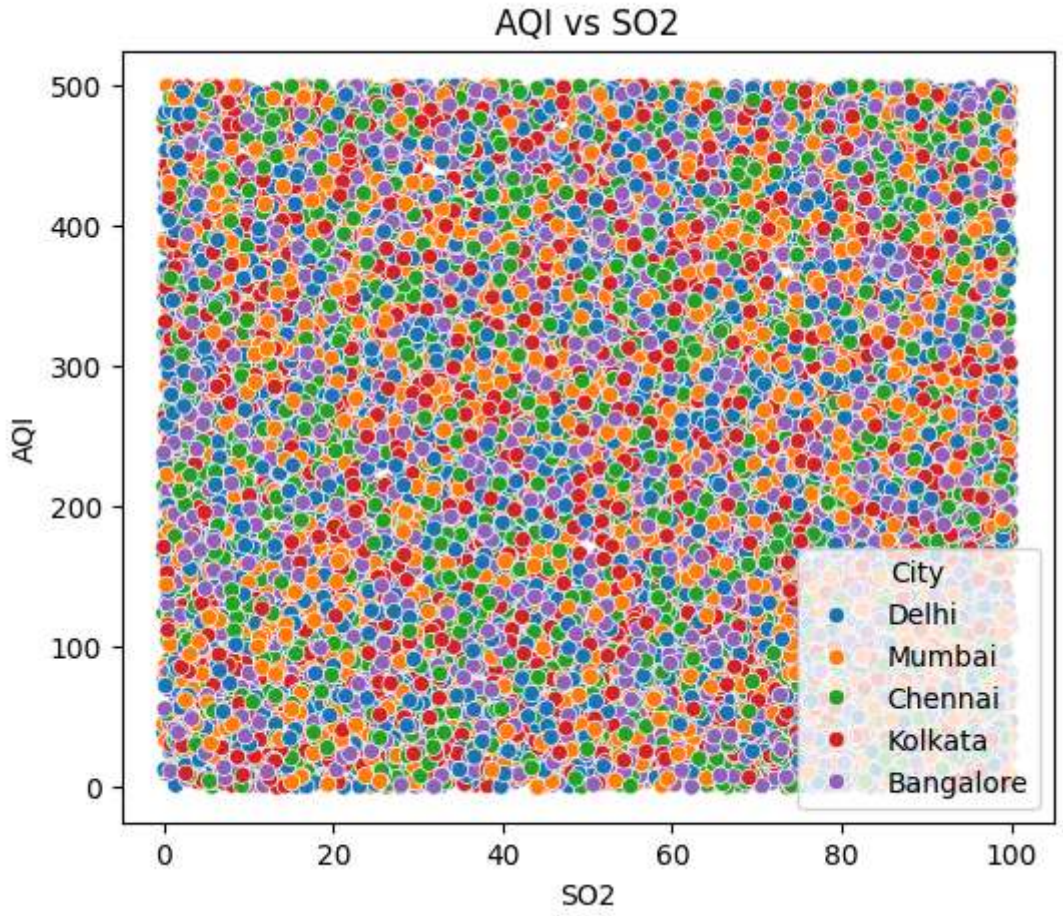

AQI vs NO
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore
AQI
NO
0
100
200
300
400
500
0
25
50
75
100
125
150
175
200

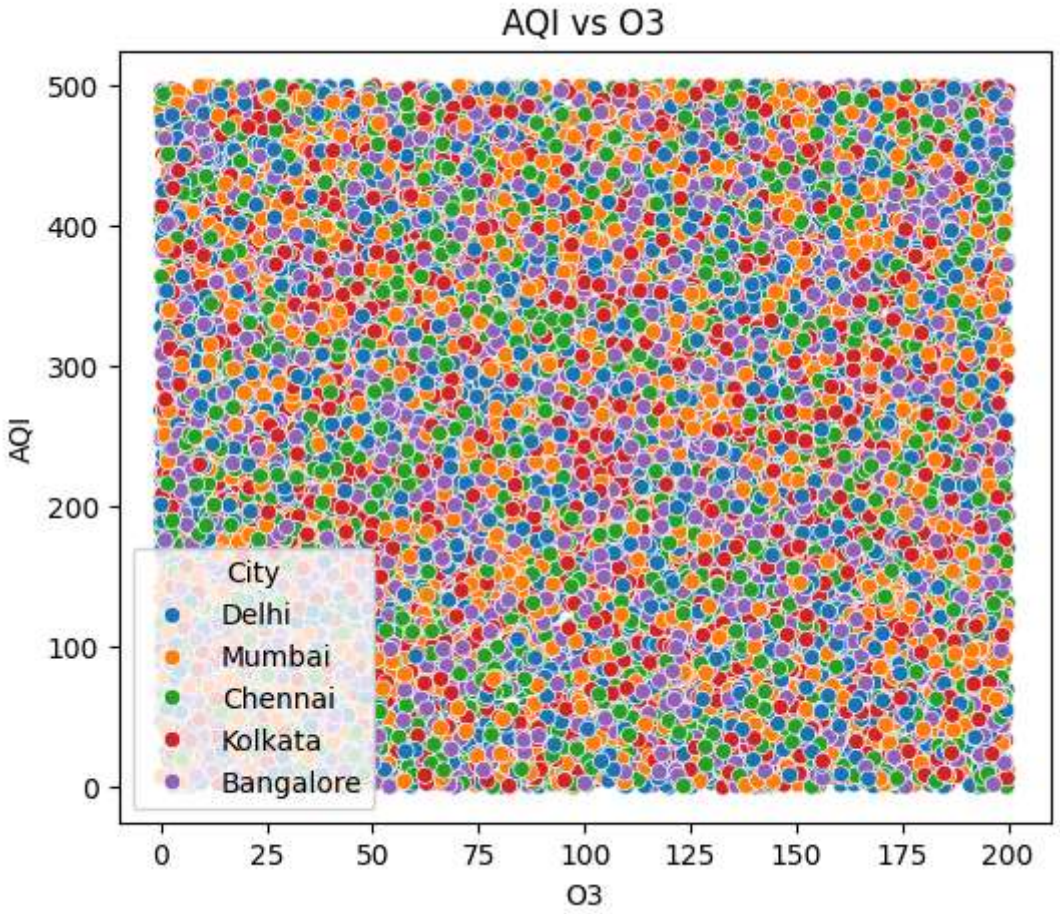

AQI vs NO2
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore
AQI
NO2
0
100
200
300
400
500
0
20
40
60
80
100
120
140


AQI vs NOx
AQI
NOx
0
100
200
300
400
500
0
50
100
150
200
250
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore

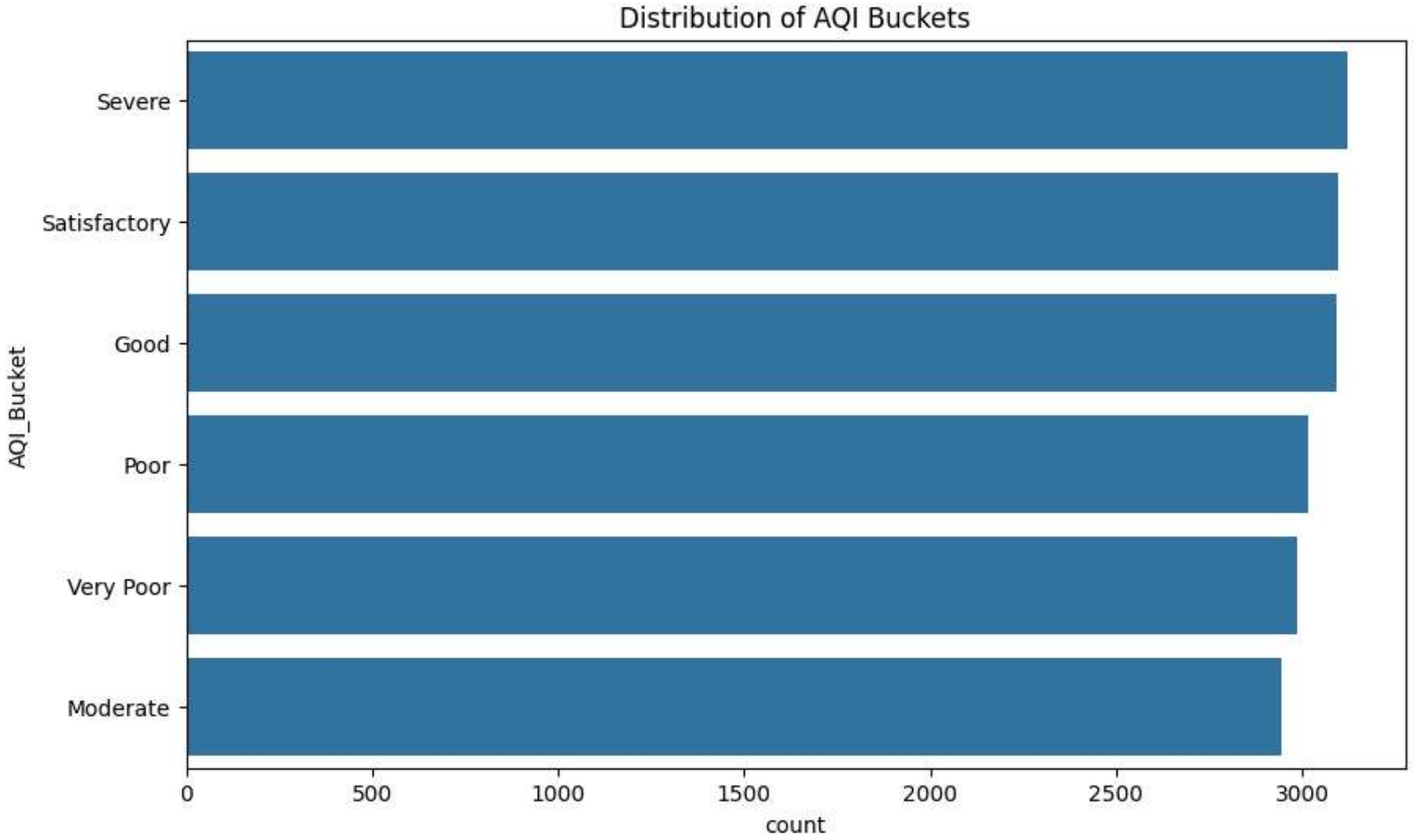

AQI vs NH3
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore
AQI
NH3
0
100
200
300
400
500
0
10
20
30
40
50


AQI vs CO
AQI
CO
0
100
200
300
400
500
0
2
4
6
8
10
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore


AQI vs SO2
AQI
SO2
0
100
200
300
400
500
0
20
40
60
80
100
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore

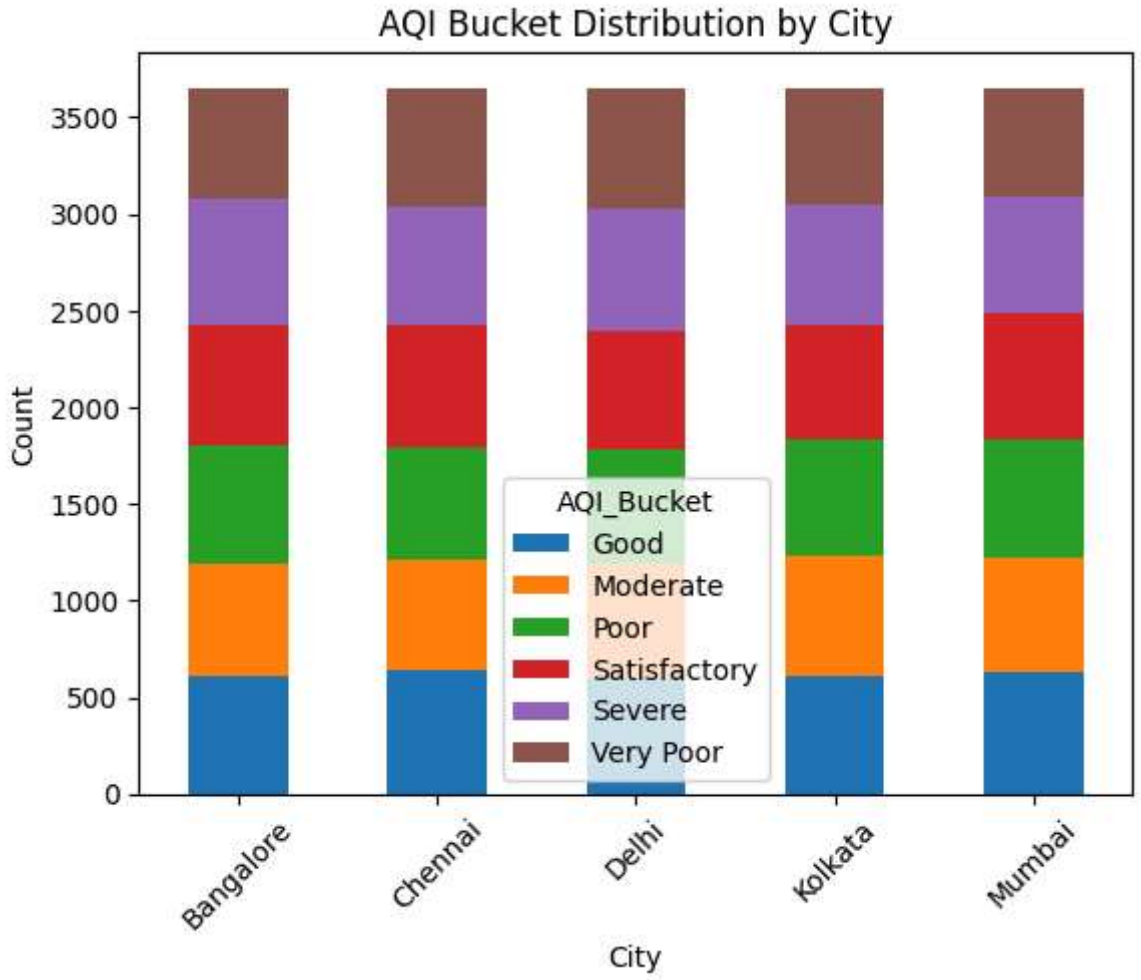


In [33]:
```
# Distribution of AQI buckets
plt.figure(figsize=(10, 6))
sns.countplot(data=city_day_data, y='AQI_Bucket', order=city_day_data['AQI_Bucket'].value_counts().index)
plt.title('Distribution of AQI Buckets')
plt.show()
```

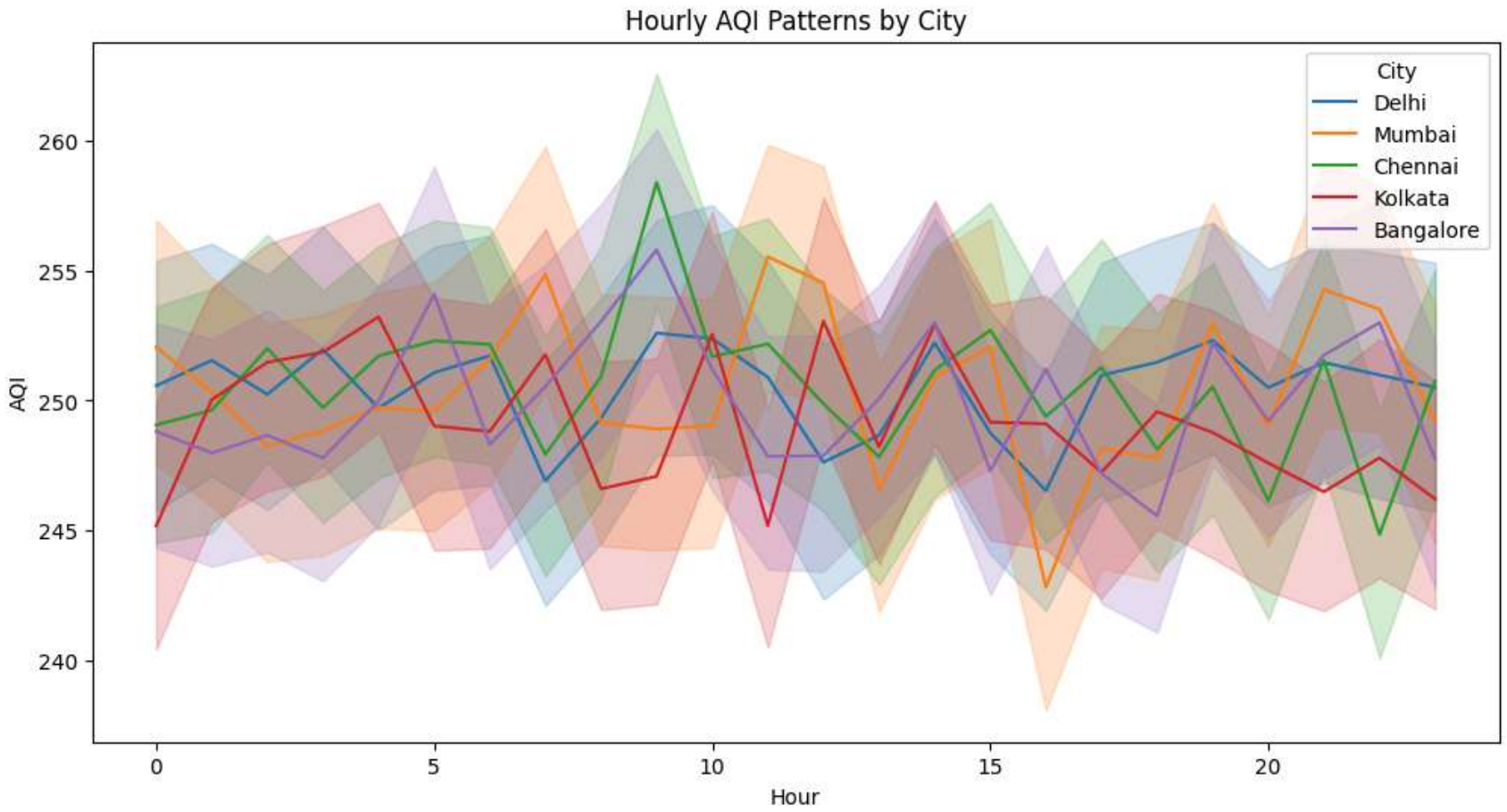


In [34]:
```
# AQI bucket by city
plt.figure(figsize=(12, 6))
pd.crosstab(city_day_data['City'], city_day_data['AQI_Bucket']).plot(kind='bar', stacked=True)
plt.title('AQI Bucket Distribution by City')
plt.ylabel('Count')
plt.xticks(rotation=45)
plt.show()
```

```
<Figure size 1200x600 with 0 Axes>
```



In [32]:
```
# Hourly patterns
plt.figure(figsize=(12, 6))
sns.lineplot(data=city_hour_data, x='Hour', y='AQI', hue='City')
plt.title('Hourly AQI Patterns by City')
plt.show()
```


Hourly AQI Patterns by City
City
Delhi
Mumbai
Chennai
Kolkata
Bangalore
AQI
260
255
250
245
240
0
5
10
15
20
Hour