SEASONAL ANALYSIS OF AQI TRENDS IN MUMBAI

Sakshi Taparia (2017MT10748) - Serial Number 503
Anchit Tandon (2017MT10772) - Serial Number 511
Gauri Gupta (2017MT60207) - Serial Number 512
Sharut Gupta (2017MT60250) - Serial Number 513
Harshit Khanna (2017MT60774) - Serial Number 516

April 22, 2019

1 THEORY & INTRODUCTION

Awareness of daily levels of air pollution is important to the people, especially those who suffer from illnesses aggravated by contaminated air. Further, the success of attempts to improve air quality depends on the support from the citizens of the nation who are well-informed about the local and national air pollution problems and the progress of mitigation efforts. In order to achieve this, the concept of **AIR QUALITY INDEX (AQI)** has been introduced by the government as a simple yet effective means of communicating the air quality to public.

1.1 Definition of Air Quality Index

AQI is defined as an overall scheme that **transforms weighted values of individual air pollutants into a single number or set of numbers** that is widely used for air quality communication and decision making in many countries.

1.2 Understanding Air Quality Index

The proposed index has **six categories** with an elegant colour scheme. (CPCB, 2014)

Good (0-50)	Satisfactory (51-100)	Moderately polluted (101-200)	Poor (201-300)	Very poor (301-400)	Severe (> 401)

AQI values are easy to comprehend. The higher the value, the greater the level of air pollution and the greater the health concern. An AQI value of 100 generally corresponds to the national air quality standard for the pollutant, which is the level that Environmental Protection Agency (EPA) has set to protect public health. Further, the health concerns associated with each of the six divisions mentioned above are summarized below. (AirNow, 2016)

Air Quality Index Levels of Health Concern	Numerical Value	Meaning
Good	0 to 50	Air quality is considered satisfactory, and air pollution poses little or no risk.
Moderate	51 to 100	Air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people who are unusually sensitive to air pollution.
Unhealthy for Sensitive Groups	101 to 150	Members of sensitive groups may experience health effects. The general public is not likely to be affected.
Unhealthy	151 to 200	Everyone may begin to experience health effects; members of sensitive groups may experience more serious health effects.
Very Unhealthy	201 to 300	Health warnings of emergency conditions. The entire population is more likely to be affected.
Hazardous	301 to 500	Health alert: everyone may experience more serious health effects.

1.3 Applications of Air Quality Index

Ott (1978) has listed the following six objectives that are served by AQI. (CPCB, 2014)

1.3.1 Resource Allocation

AQI values are used to assist administrators in allocating funds and determining priorities. This enables evaluation of trade-offs involved in alternative air pollution control strategies.

1.3.2 Ranking of Locations

AQI values also assist in comparing air quality conditions at different locations/cities. This helps in pointing out areas and frequencies of potential hazards.

1.3.3 Enforcement of Standards

Another application is to determine extent to which the legislative standards and existing criteria are being adhered. It also helps in identifying faulty standards and inadequate monitoring programs.

1.3.4 Trend Analysis

These values are also used to determine change in air quality (degradation or improvement) which have occurred over a specified period. This enables forecasting of air quality and planning of pollution control measures.

1.3.5 Public Information

An important aspect of such policies is to inform the general public about environmental conditions. It's useful for people who suffer from illness aggravated or caused by air pollution as it enables them to modify their daily activities at times when they are informed of high pollution levels.

1.3.6 Scientific Research

Finally, these values act as means for reducing a large set of data to a comprehensible form that gives better insight to researchers while conducting a study of some environmental phenomena. This enables more objective determination of the contribution of individual pollutants and sources to overall air quality.

2 METHODOLOGY & CALCULATIONS

2.1 Calculation of Air Quality Index

The air quality index or AQI is calculated as a piece-wise linear function of all the pollutant concentrations in the air (CPCB, 2014).

Primarily, two steps are involved in formulating an AQI: (i) formation of sub-indices (for each pollutant) and (ii) aggregation of sub-indices.

The Sub-index function represents the relationship between pollutant concentration X_i and corresponding sub-index I_i. Typically, the $I - X$ relationship is represented as follows -

$$I = \alpha * X + \beta$$

Where, $\alpha =$ slope of the line, $\beta =$ intercept at $X = 0$.

The general equation for the sub-index I_i for a given pollutant concentration (C_p) as based on the **linear segmented principle** is calculated as:

$$I_i = [(I_{HI} - I_{LO})/(B_{HI} - B_{LO}) * (C_p - B_{LO})] + I_{LO}$$

where,
$B_{HI} =$ Breakpoint concentration greater or equal to given concentration.
$B_{LO} =$ Breakpoint concentration smaller or equal to given concentration.
$I_{HI} =$ AQI value corresponding to B_{HI}.
$I_{LO} =$ AQI value corresponding to B_{LO}.
$I_p =$ Pollutant concentration.

Once the sub-indices are formed, they are combined or aggregated in a simple additive form or weighted additive form using either of the following methods -

- Weighted Additive Form
- Root-Mean-Square Form
- Root-Sum-Power Form (non-linear aggregation form)
- Min or Max Operator

India uses the **MAX OPERATOR**. Hence, the pollutant concentration of the most dominant species is reported as the final AQI value.

However, the calculation of AQI requires availability of data for a **minimum of three pollutants out of which one should necessarily be either PM10 or PM2.5** (CPCB, 2014).

2.2 Data Collection

We have downloaded the official data for **BANDRA STATION, MUMBAI** available at the **CENTRAL CONTROL ROOM FOR AIR QUALITY MANAGEMENT** website (CCR, 2019). The data has been downloaded in tabular format taking 24 1-hour average values per day.

Since the **1-hour average values of** $PM_{2.5}$, PM_{10}, SO_2, NO_2, O_3 **were predominantly zero from 01/01/2018 to 06/05/2018, we have rejected the data** during that period. In order to compensate for the lost data and carry out a complete seasonal analysis, we have **taken the data from 01/01/2019 to 19/04/2019 instead.** Hence, our analysis is done on the hourly averages for all the 6 parameters stated above from **7TH MAY 2018 TO 19TH APRIL 2019**.

2.3 Code for Calculation

The following code has been used to generate the pollutant concentrations. As per the values given in the break-point table shown below, we have **used the 1-hour average data to calculate the 24-hour averages and 8-hour averages** per day. In order to calculate the 8-hour averages, we have divided a day into 23 time slots and then taken the average of 23 values.

The maximum of all the pollutant concentrations is reported as the final AQI for that day. We have **also taken the average of all the AQI values reported over a month** in order to calculate the monthly AQI values for our seasonal analysis.

AQI Category (Range)	PM_{10} (24hr)	$PM_{2.5}$ (24hr)	NO_2 (24hr)	O_3 (8hr)	CO (8hr)	SO_2 (24hr)
Good (0–50)	0–50	0–30	0–40	0–50	0–1.0	0–40
Satisfactory (51–100)	51–100	31–60	41–80	51–100	1.1–2.0	41–80
Moderately polluted (101–200)	101–250	61–90	81–180	101–168	2.1–10	81–380
Poor (201–300)	251–350	91–120	181–280	169–208	10–17	381–800
Very poor (301–400)	351–430	121–250	281–400	209–748	17–34	801–1600
Severe (401–500)	430+	250+	400+	748+	34+	1600+

The code is written in **python** and the input files to our code have been provided in **.csv format**. The break-point values stored in the arrays for the purpose of calculation have been taken from the **break-point value table** shown above.

Listing 1: Python Code for AQI calculations

```
#!/usr/bin/env python3
# -- coding: utf-8 --
"""
Created on Sat Apr 20 00:16:50 2019

@author: Sharut Gupta, Anchit Tandon
"""

import matplotlib.pyplot as plt
import csv
import numpy as np

row_count=0
with open("data1.csv", 'r') as f:
```

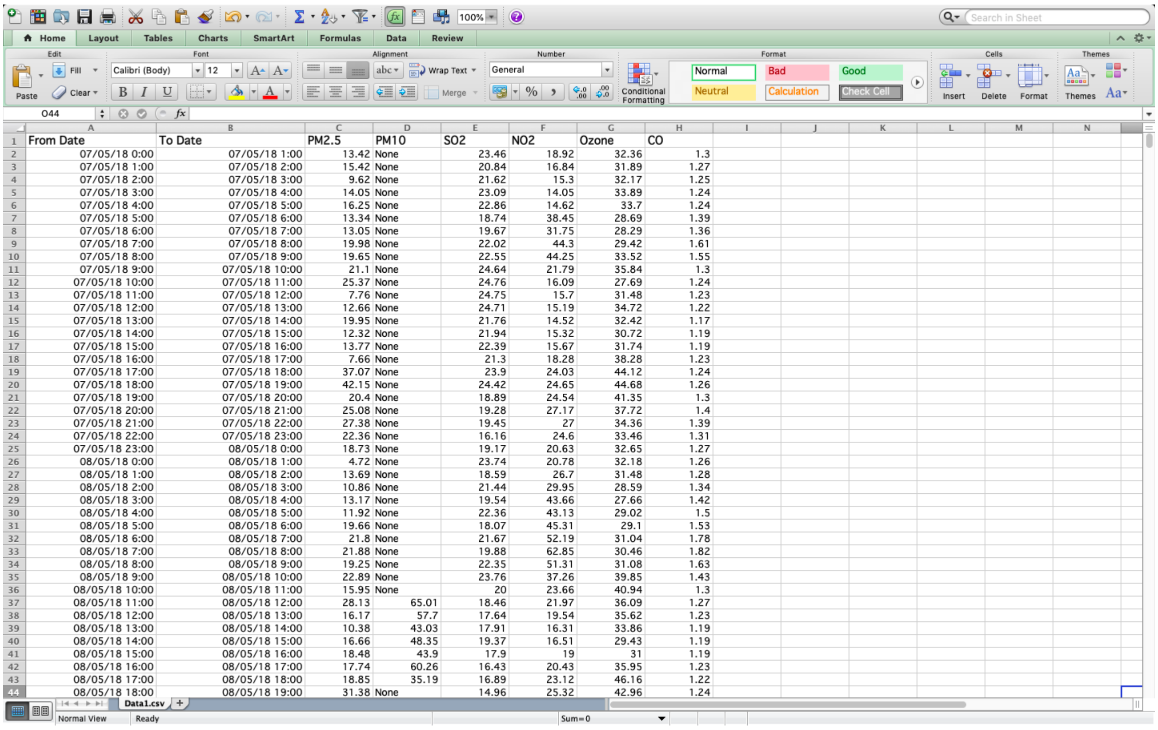
```
    for line in f:
        row_count += 1
row_count = row_count - 1
first_row = True
X = np.zeros((row_count, 6))
with open("data1.csv", 'r') as train_file:
    csv_reader = csv.reader(train_file, delimiter=',')
    i = 0
    for row in csv_reader:
        if first_row:
            first_row = False
            continue
        row = [w.replace('None', 'nan') for w in row]
        X[i] = row[2:]
        i=i+1

PM25 = X[:, 0]
PM10 = X[:, 1]
SO2 = X[:, 2]
NO2 = X[:, 3]
O3 = X[:, 4]
CO = X[:, 5]

#Eight hour average defined
def EightHrAvg(X):
    Y = np.zeros((int(np.size(X, 0)/8), 1))
    for i in range(0, np.size(X, 0), 8):
        Y[int(i/8)] = np.nanmean(X[i:i+8])
    return Y

#Twenty four hour average defined
def TwentyFourHrAvg(X):
    flag=1
    Y = np.zeros((int(np.size(X, 0)/24), 1))
    for i in range(0, np.size(X, 0), 24):
        if (int(i/24)==474 and flag==1):
            Y[int(i/24)] = np.nanmean(X[i:i+1])
            flag=0
        elif (flag==1):
            Y[int(i/24)] = np.nanmean(X[i:i+24])
    return Y

#How to calculate AQI
def AQI(X, vals):
    aqi = [50, 100, 200, 300, 400]
    aqi_vals = np.zeros((np.size(X, 0)))

    for i in range(np.size(X, 0)):
        prev_val = 0
        prev_aqi = 0
        for j in range(np.size(vals, 0)):
            if X[i] <= vals[j]:
                aqi_vals[i] = prev_aqi +
                (X[i] - prev_val) / (vals[j] - prev_val) * (aqi[j]-prev_aqi)
                break
            prev_aqi = aqi[j]-1
            prev_val = vals[j]
```

```python
    return aqi_vals

#calculate 24-hour average for PM25, PM10, SO2, NO2
PM25_avg = TwentyFourHrAvg(PM25)
PM10_avg = TwentyFourHrAvg(PM10)
SO2_avg = TwentyFourHrAvg(SO2)
NO2_avg = TwentyFourHrAvg(NO2)

#calculate 8-hour average for O3, CO
O3_avg = EightHrAvg(O3)
CO_avg = EightHrAvg(CO)

#calculate the AQI sub-indices
AQI_PM25 = AQI(PM25_avg, [30, 60, 90, 120, 250])
AQI_PM10 = AQI(PM10_avg, [50, 100, 250, 350, 430])
AQI_SO2 = AQI(SO2_avg, [40, 80, 380, 800, 1600])
AQI_NO2 = AQI(NO2_avg, [40, 80, 180, 280, 400])
AQI_O3_temp = AQI(O3_avg, [50, 100, 168, 208, 748])
AQI_CO_temp = AQI(CO_avg, [1, 2, 10, 17, 34])

AQI_O3 = np.zeros((int(np.size(AQI_O3_temp, 0)/3)))
AQI_CO = np.zeros((int(np.size(AQI_CO_temp, 0)/3)))

#find the max out of the 8 hour averages for O3 and CO
for i in range(0, np.size(AQI_O3, 0), 3):
    AQI_O3[int(i/3)] = np.max(AQI_O3_temp[i:i+2])
    AQI_CO[int(i/3)] = np.max(AQI_CO_temp[i:i+2])

#calculate the AQI for each day
AQI = np.zeros((np.size(AQI_PM10, 0)))
for i in range(np.size(AQI_PM10, 0)):
    AQI[i] = np.max([AQI_PM25[i], AQI_PM10[i], AQI_SO2[i], AQI_NO2[i], AQI_O3[i], AQI_CO[i]])

with open('my_data.csv', 'a') as csvFile:
        row = ['PM2.5','PM10','SO2','NO2','O3','CO']
        writer = csv.writer(csvFile)
        writer.writerow(row)
        for i in range(len(AQI_PM25)):
            my = [AQI_PM25[i], AQI_PM10[i], AQI_SO2[i], AQI_NO2[i], AQI_O3[i], AQI_CO[i]]
            writer.writerow(my)
        csvFile.close()

months = [25, 30, 31,  31, 30, 31, 30, 31,31,28,31,19]
monthlyAQI = np.zeros(12)

days = 0
for i in range(np.size(months, 0)):
    X = AQI[days:days+months[i]]
    a = np.nonzero(X)
    monthlyAQI[i] = np.mean(X[a])
    days = days + months[i]

print(monthlyAQI)
print(np.mean(monthlyAQI))

#Plot graph of monthly data points
```

```python
x = np.array(['May18','June18','July18','Aug18','Sept18','Oct18','Nov18','Dec18',
'Jan19','Feb19','March19','April19'])
xa = np.array([0,1,2,3,4,5,6,7,8,9,10,11])
plt.xticks(np.arange(12), x)
y= list(monthlyAQI)
print(x)
print(y)
plt.xlabel('Months')
plt.ylabel('Monthly_AQI')
for i in range(len(x)):
    plt.plot(x[i:i+2],y[i:i+2],'ro-')
plt.show()
```

3 RESULTS OBTAINED

A screen shot tabulating the input file given to our code has been attached below for reference.

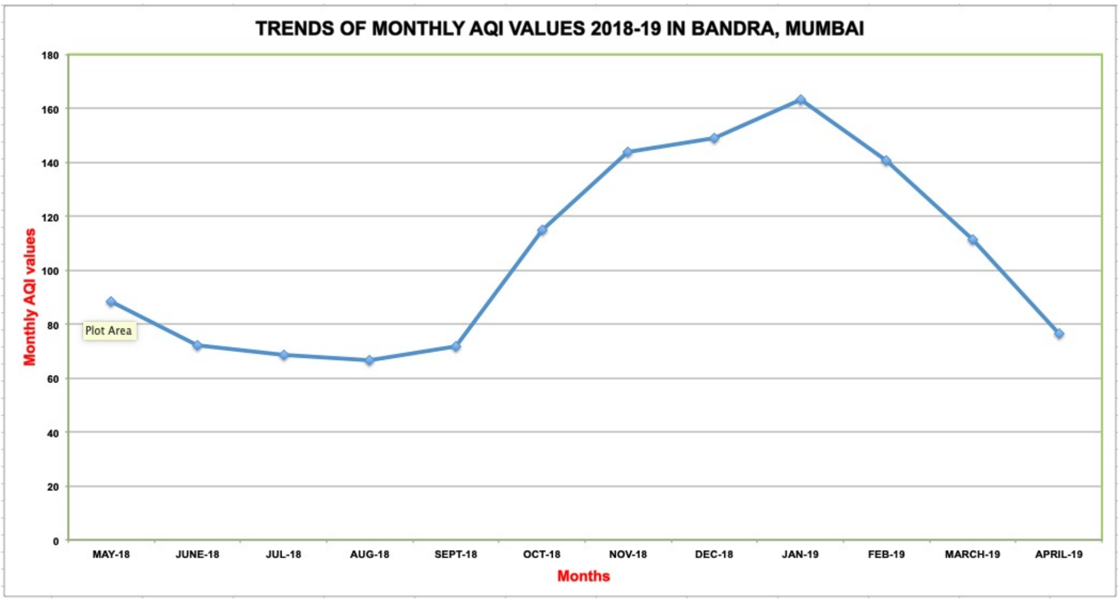


	A	B	C	D	E	F	G	H
1	From Date	To Date	PM2.5	PM10	SO2	NO2	Ozone	CO
2	07/05/18 0:00	07/05/18 1:00	13.42	None	23.46	18.92	32.36	1.3
3	07/05/18 1:00	07/05/18 2:00	15.42	None	20.84	16.84	31.89	1.27
4	07/05/18 2:00	07/05/18 3:00	9.62	None	21.62	15.3	32.17	1.25
5	07/05/18 3:00	07/05/18 4:00	14.05	None	23.09	14.05	33.89	1.24
6	07/05/18 4:00	07/05/18 5:00	16.25	None	22.86	14.62	33.7	1.24
7	07/05/18 5:00	07/05/18 6:00	13.34	None	18.74	38.45	28.69	1.39
8	07/05/18 6:00	07/05/18 7:00	13.05	None	19.67	31.75	28.29	1.36
9	07/05/18 7:00	07/05/18 8:00	19.98	None	22.02	44.3	29.42	1.61
10	07/05/18 8:00	07/05/18 9:00	19.65	None	22.55	44.25	33.52	1.55
11	07/05/18 9:00	07/05/18 10:00	21.1	None	24.64	21.79	35.84	1.3
12	07/05/18 10:00	07/05/18 11:00	25.37	None	24.76	16.09	27.69	1.24
13	07/05/18 11:00	07/05/18 12:00	7.76	None	24.75	15.7	31.48	1.23
14	07/05/18 12:00	07/05/18 13:00	12.66	None	24.71	15.19	34.72	1.22
15	07/05/18 13:00	07/05/18 14:00	19.95	None	21.76	14.52	32.42	1.17
16	07/05/18 14:00	07/05/18 15:00	12.32	None	21.94	15.32	30.72	1.19
17	07/05/18 15:00	07/05/18 16:00	13.77	None	22.39	15.67	31.74	1.19
18	07/05/18 16:00	07/05/18 17:00	7.66	None	21.3	18.28	38.28	1.23
19	07/05/18 17:00	07/05/18 18:00	37.07	None	23.9	24.03	44.12	1.24
20	07/05/18 18:00	07/05/18 19:00	42.15	None	24.42	24.65	44.68	1.26
21	07/05/18 19:00	07/05/18 20:00	20.4	None	18.89	24.54	41.35	1.3
22	07/05/18 20:00	07/05/18 21:00	25.08	None	19.28	27.17	37.72	1.4
23	07/05/18 21:00	07/05/18 22:00	27.38	None	19.45	27	34.36	1.39
24	07/05/18 22:00	07/05/18 23:00	22.36	None	16.16	24.6	33.46	1.31
25	07/05/18 23:00	08/05/18 0:00	18.73	None	19.17	20.63	32.65	1.27
26	08/05/18 0:00	08/05/18 1:00	4.72	None	23.74	20.78	32.18	1.26
27	08/05/18 1:00	08/05/18 2:00	13.69	None	18.59	26.7	31.48	1.28
28	08/05/18 2:00	08/05/18 3:00	10.86	None	21.44	29.95	28.59	1.34
29	08/05/18 3:00	08/05/18 4:00	13.17	None	19.54	43.66	27.66	1.42
30	08/05/18 4:00	08/05/18 5:00	11.92	None	22.36	43.13	29.02	1.5
31	08/05/18 5:00	08/05/18 6:00	19.66	None	18.07	45.31	29.1	1.53
32	08/05/18 6:00	08/05/18 7:00	21.8	None	21.67	52.19	31.04	1.78
33	08/05/18 7:00	08/05/18 8:00	21.88	None	19.88	62.85	30.46	1.82
34	08/05/18 8:00	08/05/18 9:00	19.25	None	22.35	51.31	31.08	1.63
35	08/05/18 9:00	08/05/18 10:00	22.89	None	23.76	37.26	39.85	1.43
36	08/05/18 10:00	08/05/18 11:00	15.95	None	20	23.66	40.94	1.3
37	08/05/18 11:00	08/05/18 12:00	28.13	65.01	18.46	21.97	36.09	1.27
38	08/05/18 12:00	08/05/18 13:00	16.17	57.7	17.64	19.54	35.62	1.23
39	08/05/18 13:00	08/05/18 14:00	10.38	43.03	17.91	16.31	33.86	1.19
40	08/05/18 14:00	08/05/18 15:00	16.66	48.35	19.37	16.51	29.43	1.19
41	08/05/18 15:00	08/05/18 16:00	18.48	43.9	17.9	19	31	1.19
42	08/05/18 16:00	08/05/18 17:00	17.74	60.26	16.43	20.43	35.95	1.23
43	08/05/18 17:00	08/05/18 18:00	18.85	35.19	16.89	23.12	46.16	1.22
44	08/05/18 18:00	08/05/18 19:00	31.38	None	14.96	25.32	42.96	1.24

The output values that were generated from our code were used for the purpose of analysis and plotting of graphs. We have shown **two plots** in the analysis section -

- **Trends of monthly AQI values** - The monthly AQI averages obtained from May 2018 to April 2019 have been summarized in Table 2 & 3. These values have been used to obtain a bar-graph and line-graph in order to carry out **seasonal analysis of AQI variations** across the year.
- **Variation of pollutant concentration** - This plot contains the daily pollutant concentrations calculated for all the 6 parameters. This plot has been used for the **analysis of the most dominant species** contributing to the AQI value on daily basis. We can also infer other trends like least dominant species and variation of concentrations of all the species with time.

4 ANALYSIS & CONCLUSION

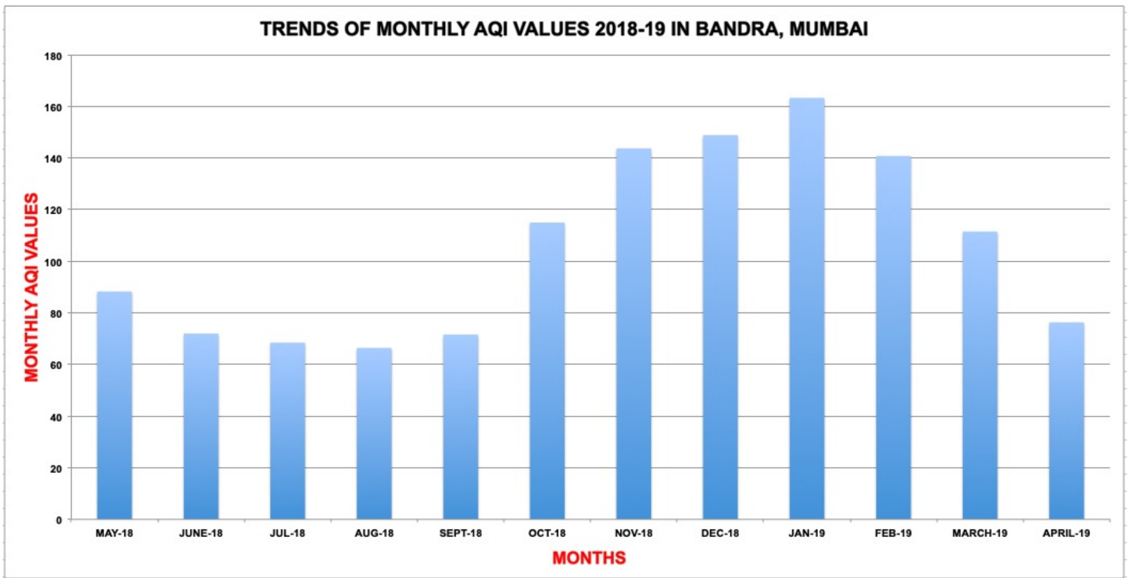


Month	May 18	June 18	July 18	Aug 18	Sep 18	Oct 18	Nov 18	Dec 18
Average AQI	88.29	72.01	68.47	66.43	71.62	115.01	143.74	148.87

Table 1: Average Monthy AQI values for 2018

Month	Jan 19	Feb 19	Mar 19	Apr 19
Average AQI	163.30	140.78	111.51	76.33

Table 2: Average Monthy AQI values for 2019



Seasonal Analysis

Mumbai is the capital of Maharashtra state with a population of approximately 12 million people. It has tropical hot and humid climate with hot summers, moderately cold winters, severe monsoons and large diurnal variations in the wind speed.

After analysis of the monthly AQI data from the months of May 2018 to April 2019 it was observed that the values were significantly **high from Nov 2018 to Feb 2019 (mostly during winter season)** wherein highest value of average AQI was observed in January 2019 (AQI = 163.306). Also, the study showed that the **AQI values were marginally low from July 2018 to September 2018 (largely comprising of the monsoon season)** wherein the lowest average AQI value was obtained in the month of August 2018 (AQI = 66.429).

The average AQI values of Mumbai air fall under 3 different categories of air quality as follows -

- **Moderate (AQI = 51-100)**

  This comprises of the **months of May 2018 to September 2018 along with April 2019.** Such levels may cause minor breathing discomfort to sensitive people.

- **Unhealthy for sensitive groups (AQI = 101-150)**

  This comprises of months from **October 2018 to December 2018 and from February 2019 to March 2019.** This may cause breathing discomfort to the people with lung disease such as asthma and discomfort to people with heart disease, children and older adults.

- **Unhealthy (AQI = 151-200)**

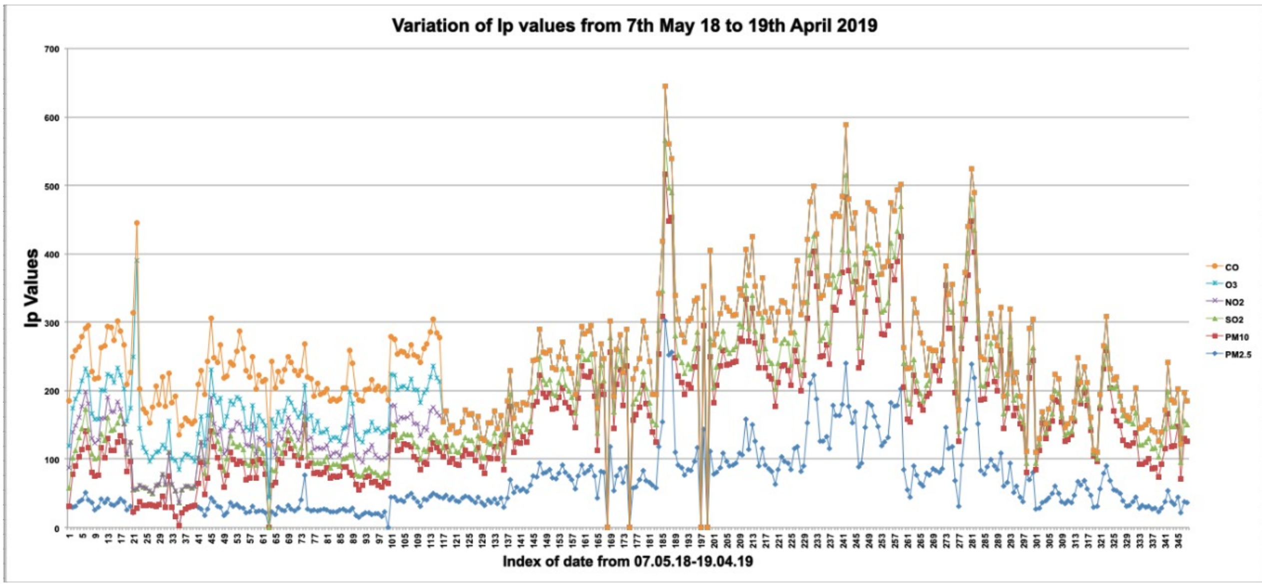
  This category comprises of **only January 2019** which is the peak winter season, corresponding to the highest AQI value found among all months. These elevated levels may cause breathing discomfort to people on prolonged exposure and discomfort to people with heart disease with short exposure.

A few facts that are supported and strengthened by our analysis are -

- Activities that **majorly cause air pollution tend to increase as the winter season approaches.** For example, people start burning more wood, coal and biomass to heat their homes, factories and workplaces. This contributes to increase in SO_2 levels and particulate matter in the air during winters. Moreover, cool air becomes inactive over the city horizon around this time. This way, the pollution gets trapped, close to the ground significantly raising the AQI levels of nearby areas.

- AQI values of Mumbai lie mostly between 100-300 and have very rarely shot up to the severe zone. This trend contrasts with landlocked cities like Delhi where AQI values shot upto 400 more often. This is due to the coastal regions leading to the phenomena of land and sea breeze. The **wind speed in Mumbai is higher, which disperses majority of the pollutants** and lowers the AQI.

- Infact, on rare days when AQI in Mumbai was crossed 400, it was observed that temperatures were not low but owing to the high humidity and negligible wind speed, the pollutants were not dispersed from the city's air leading to haze formation during early morning and evening.



Dominant Species Variability and Analysis

- **Most Dominant Species**

  It can be clearly observed from given graph that **CO is the dominant pollutant** in the AQI calculations for each month with the maximum concentration observed during January 2019 making January the month with the least air quality. Further our analysis showed that high concentrations of CO were found during the winter season- fom December 2018 to February 2019.

  Once reason behind this trend is that Mumbai, being a metropolitan city, has witnessed a tremendous increase in the number of vehicles and industries during recent times. CO being a primary component of vehicle exhaust fumes and industrial air exhaust, has thus contributed towards a significant depreciation in air quality.

- **Least Dominant Species**

  We also concluded that $PM_{2.5}$ **had considerably low concentrations** during most the months showing some rise in its concentration during winters i.e. months of January 2019 to March 2019.

  Possible reason for the same is that $PM_{2.5}$ is primarily caused by dust, spores and pollen. These components are scarce in Mumbai air. Additionally, the concentration of $PM_{2.5}$ is further pulled down due to the settling of the dust particles on sea and water bodies thus making the above observations coherent with the given facts.

- **Other major observations**

  Analysis further suggest that ozone was the second dominant pollutant seen during May 2018 to August 2018. However, during the later months, SO_2 become the second dominant pollutant having concentrations close to ozone. The variability of SO_2 concentrations was more in winters and less in rainy season. Its concentration was relatively stable in the summer season.

References

AirNow. 2016. “Air Quality Index (AQI) Basics.”.
URL: *https://airnow.gov/index.cfm?action=aqibasics.aqi*

CCR. 2019. “Central Control Room for Air Quality Management - All India.”.
URL: *https://app.cpcbccr.com/ccr//caaqm-dashboard-all/caaqm-landing/data*

CPCB. 2014. “National Air Quality Index.”.
URL: *http://www.indiaenvironmentportal.org.in/files/file/Air%20Quality%20Index.pdf*