# Analyzing the Effect of Mobility Restrictions on Air Pollution During COVID-19

**Aman Kumar[1], Ankita Singh[2]**

Amity School of Engineering & Technology, Amity University, Patna, Bihar, India
Orcid ID: 0000-0002-6302-8145
Email ID: *aman.science.2000[at]gmail.com*

Amity School of Engineering and Technology, Amity University, Patna, Bihar, India
Orcid ID: 0000-0003-3215-1789
Email ID: *ankitaasingh21[at]gmail.com*

**Abstract:** *With advancement in computing and the field of science, analysis has become faster than ever. An enormous amount of real-time data is collected and stored with the help of IoT. Data science has helped in the integration of those complex data and reconstructed them into new knowledge. The augmentation of Data Science during times of COVID-19 is unprecedented. From methods to analyze the spread of disease to the evaluation of government policies into action, Data Science had a significant contribution in understanding COVID-19. Various state and national governments enforced strict lockdowns and mobility restrictions during the times of COVID. The enforced stratagems helped to contain the spread of the virus and, at the same time, also proved to be advantageous for the environment. This paper is an effort to study the effects of mobility restriction and lockdown over the pollution figures in India. In this study, pollutants were taken for the calculation of the Air Quality Index. Based on AQI, parameters and time frames where pollution levels are significantly feebler are determined. Succeeding, on observing patterns between these parameters and mobility restriction, a model is developed. This model shows the effect of mobility restriction on pollution, testing the covariates used for this study.*

**Keywords:** Data Sciences, Air Pollution, COVID-19, Lockdowns, Government Policies, Environmental Impact

## 1. Introduction

For the past two years, the world has been under the impact of the SARS-CoV-2. Originating in Wuhan province, China, this virus has affected and taken the lives of millions of people. Mobility restrictions and residential stay had been the only methods to be safe from this pandemic [1,2]. For that purpose, the governments across geographies implemented strict mobility restrictions by enforcing nationwide lockdowns. Lockdown and mobility restriction measures had been economically agitative; yet, they highlighted a brighter side too. The restrictions put in place helped in a contained virus spread, the decline in the death rates, and favorable environmental effects [3,4].

Because of mobility restriction and lockdown, many industries cease to pause their work. Thus, the concentration of defilements that create air pollution; had been substantially reduced [5]. With the use of State Models and concepts of Statistics, the pollution concentration across geographies is evaluated, compared with mobility restrictions put into place and compared to analyze the available correlations [6]. Data science is one of the emerging fields of science that extract knowledge and insights from data (structured or unstructured). It involves the use of a scientific method, processes, algorithm, and system. Such knowledge and insights from data apply to a broad range of application domains. These studies, with these tools, could help the governments in the analysis and iteration of the existing policies on pollution control.

## 2. Literature Survey

Data science can endeavor a lot in terms of deep knowledge of the natural environment and informing mitigation, analysis and adoption of policies. IoT and sensors have the potential to monitor and access the real-time data of any geographical area across the world. Cloud computing can help us to store data and do real-time data analysis. So, data science with modern technologies is a silver bullet to understand any change in pollution and respond to it [7].

### a) Data collection and availability
This research finds the use of various data references on air pollution levels across India, with the time-series data for mobility restrictions during COVID-19, for pattern analysis. The related data obtained from Kaggle and ACAPS community reports; is loaded to an ETL process for further wrangling to build statistical inferences and graphs for analysis. The following are the collection of datasets used for this study:
- ACAPS COVID-19 Government Measures Data [8]
- Google Community Mobility Report [9]
- Air Quality Data in India (2015 - 2020) [10]

### b) Data Selection
The datasets used for this research study are as under:
- ACAPS COVID-19 Government Measures dataset, filtered to collect results only for India. The data was plotted with air pollution numbers to carry out a comparative study between the two.
- Google's Community Mobility Report dataset, fetched via API Calls. Plots on mobility restrictions put into place across India were collected.
- The Air Quality Data in India (2015 - 2020); obtained from Kaggle wrangled - to make it fit for modelling purposes.

## 3. Methodology

The research presented here involves using various tools for implementing the code for mathematical and statistical analysis. This section of the paper covers: the tools used, mathematical equations put into place for calculating air pollution, and the methodology of hypothesis testing for the effect of lockdowns on pollution control.

### a) Tools Used
- Google Collaboratory Notebook on Python
- Microsoft Excel for Data Analysis
- SAS University Edition for EDA
- Tableau 2020& Flourish for data visualizations

### b) Implementation of the Model
With the data from the Air Quality Data in India (2015 - 2020) dataset, the pollution levels in India are fetched and studied. The dataset contains daily reported values across multiple stations for pollutants: Nitrogen Oxides (NO, $NO_X$), Sulphur Dioxide ($SO_2$), Carbon Monoxide (CO), Ammonia ($NH_3$), Ozone ($O_3$), BTX (Benzene, Toluene and Xylene) and Particulate Matter 2.5 & 10).

Time Series analysis of these values provides a holistic view of the situation of air pollution in India. With the help of mathematical and statistical techniques, a comparative study of air pollution pre and post COVID-19 lockdowns is studied.

By merging the contents for ACAPS Covid-19 Government Measures Dataset and Google's Community Mobility Reports, a parallel study of mobility restrictions and air pollution is studied. The study holds potential benefit in providing an idea over the positive effects of lockdowns on the environment. It could also help in assessing and framing government policies in combating pollution.

### c) Mathematical Equations to calculate AQI
Air Quality Index is an effective measure of pollution for a specific region as it considers various pollutant numbers. The AQI numbers are a generalized and normalized representation of pollutants. The Air Quality Data in India (2015 - 2020) dataset contains data from 24 cities of India, with individual pollutant concentrations reported between 2015-2020. As a generalization for this, the AQI is calculated based on the data available for this research. The following table and predefined equations provide a method to do the same.

### Factors to consider while calculation of AQI:
The AQI calculation methodology adopted by Central Control Room for Air Quality Management: Delhi NCR [11] used for this study; states the following:
- Out of all the available pollutants: pM 2.5, pM 10, $SO_2$, $NO_X$, CO, $NH_3$, and $O_3$; at least three contaminants must be present with one of them must be either pM 2.5, pM 10 [10].
- Concentration for all the pollutants has measurements in $\mu g/m^3$, except CO, which has units in $mg/m^3$.
- The 24-hour average calculated values for pM 2.5, pM 10, $SO_2$, $NO_X$, and $NH_3$, and an 8-hour average calculated values for CO and $O_3$ pollutants implied for calculating the AQI.

### AQI Sub Index Calculation Table for pollutants
With the help of Tables 1-7, for any catered pollutant, its AQI sub-index could be assessed. The AQI for a particular place is:

$$Max\ Sub\text{-}index\ (pM2.5\&10, SO_2, NO_X, CO, NH_3, O_3) \quad (i)$$

**Table 1:** AQI Sub-index calculation for pM 2.5

| Value of PM 2.5 (x) | PM2.5 AQI SubIndex Formula |
|---|---|
| x < 30 | x * 50 / 30 |
| 30 < x <= 60 | 50 + (x - 30) * 50 / 30 |
| 60 < x <= 90 | 100 + (x - 60) * 100 / 30 |
| 90< x <= 120 | 200 + (x - 90) * 100 / 30 |
| 120 < x <=250 | 300 + (x - 120) * 100 / 130 |
| x >250 | 400 + (x - 250) * 100 / 130 |

**Table 2:** AQI Sub-index calculation for pM 10

| Value of PM 10 (x) | PM10 AQI Sub Index Formula |
|---|---|
| x <= 100 | x |
| 100 < x <= 250 | 100 + (x - 100) * 100 / 150 |
| 250 < x <= 350 | 200 + (x - 250) |
| 350 < x <= 430 | 300 + (x - 350) * 100 / 80 |
| x > 430 | 400 + (x - 430) * 100 / 80 |

**Table 3:** AQI Sub-index calculation for Ozone

| Value of $O_3$ (x) | $O_3$ AQI Sub Index Formula |
|---|---|
| x <= 50 | x * 50 / 50 |
| 50 < x <=100 | 50 + (x - 50) * 50 / 50 |
| 100 < x <= 168 | 100 + (x - 100) * 100 / 68 |
| 168 < x <= 208 | 200 + (x - 168) * 100 / 40 |
| 208 < x <= 748 | 300 + (x - 208) * 100 / 539 |
| x > 748 | 400 + (x - 400) * 100 / 539 |

**Table 4:** AQI Sub-index calculation for Carbon Monoxide (CO)

| Value of CO (x) | CO AQI Sub Index Formula |
|---|---|
| x <= 1 | x * 50 / 1 |
| 1 < x <= 2 | 50 + (x - 1) * 50 / 1 |
| 2 < x <= 10 | 100 + (x - 2) * 100 / 8 |
| 10 < x <= 17 | 200 + (x - 10) * 100 / 7 |
| 17 < x <= 34 | 300 + (x - 17) * 100 / 17 |
| x > 34 | 400 + (x - 34) * 100 / 17 |

**Table 5:** AQI Sub-index calculation for Sulphur Dioxide ($SO_2$)

| Value of $SO_2$ (x) | $SO_2$ AQI Sub Index Formula |
|---|---|
| x <= 40 | x * 50 / 40 |
| 40 < x <= 80 | 50 + (x - 40) * 50 / 40 |
| 80 < x <= 380 | 100 + (x - 80) * 100 / 300 |
| 380 < x <= 800 | 200 + (x - 380) * 100 / 420 |
| 800 < x <= 1600 | 300 + (x - 800) * 100 / 800 |
| x > 1600 | 400 + (x - 1600) * 100 / 800 |

**Table 6:** AQI Sub-index calculation for NOx

| Value of NOx (x) | NOx AQI Sub Index Formula |
|---|---|
| x <= 40 | x * 50 / 40 |
| 40 < x <= 80 | 50 + (x - 40) * 50 / 40 |
| 80 < x <= 180 | 100 + (x - 80) * 100 / 100 |
| 180 < x <= 280 | 200 + (x - 180) * 100 / 100 |
| 280 < x <= 400 | 300 + (x - 280) * 100 / 120 |
| x > 400 | 400 + (x - 400) * 100 / 120 |

**Table 7:** AQI Sub-index calculation for $NH_3$(Ammonia)

| Value of $NH_3$ (x) | $NH_3$ AQI Sub Index Formula |
|---|---|
| x <= 200 | x * 50 / 200 |
| 200 < x <= 400 | 50 + (x - 200) * 50 / 200 |
| 400 < x <= 800 | 100 + (x - 400) * 100 / 400 |
| 800 < x <= 1200 | 200 + (x - 800) * 100 / 400 |
| 1200 < x <= 1800 | 300 + (x - 1200) * 100 / 600 |
| x > 1800 | 400 + (x - 1800) * 100 / 600 |

## 4. Results & Analysis

### a) Calculation of Air Quality Index from data

With the help of formulas in (Tables 1-7), AQI for every observation on a particular date and station is calculated. The Air Quality Data in India (2015 - 2020) dataset was fetched from Kaggle in its crude form and was loaded to a Python Notebook on Google Collaboratory for EDA purposes. The data holds a set of observations between January 01, 2015, to July 1, 2020, for 24 cities of India. Table 8 highlights the sample view of the dataset. AQI is the standardized value as a pollution metric; hence, the data fetched was converted to Sub-indexes for every pollutant (Fig.1)

**Table 8:** Data from Air Quality Data in India (2015-2020) dataset

| City | Date | PM 2.5 | PM10 | NOx | NH3 | CO | SO2 | O3 |
|---|---|---|---|---|---|---|---|---|
| Mumbai | 22-06-2020 | 33.2 | 42.5 | 27.1 | 13.7 | 0.7 | 13.7 | 34.9 |
| Mumbai | 23-06-2020 | 25.4 | 34.1 | 19.9 | 13.1 | 0.5 | 10.4 | 43.3 |
| Mumbai | 24-06-2020 | 34.4 | 23.4 | 13.1 | 14.5 | 0.6 | 10.9 | 35.1 |
| Mumbai | 25-06-2020 | 13.5 | 21.6 | 13.1 | 12.3 | 0.4 | 8.2 | 29.4 |
| Mumbai | 26-06-2020 | 7.6 | 23.3 | 17.2 | 11.2 | 0.5 | 6.9 | 19.9 |
| Mumbai | 27-06-2020 | 15.0 | 25.1 | 19.5 | 12.5 | 0.5 | 8.6 | 23.3 |
| Mumbai | 28-06-2020 | 24.4 | 26.1 | 16.5 | 12.0 | 0.5 | 12.7 | 30.1 |
| Mumbai | 29-06-2020 | 22.9 | 29.5 | 18.3 | 10.7 | 0.5 | 8.4 | 31.0 |



**Figure 1:** Python function to convert pM 2.5 values into sub-indexes for Air Quality Index calculation

The code snippet in Fig.1 converts the pM 2.5 values into sub-indexes based on conversion formulas in Table 1. The same method was carried out for all the pollutants. Using formula 1, with constraints mentioned in C.1, the overall AQI was calculated. The results generated were stored as a new variable and appended to the existing data frame.

### b) Binning the AQI Obtained on basis of guidelines

Based on methodologies adopted by Central Control Room for Air Quality Management: Delhi NCR, the AQI could be categorized into the following distinct categories:

- **Good:** AQI (0-50) – Minimal to no impact on humans
- **Satisfactory:** AQI (51-100) – Minor discomfort to sensitive groups
- **Moderate:** AQI (101-200) – Breathing discomfort to people with heart and lung diseases.
- **Poor:** AQI (201-300) – Breathing discomfort on prolonged exposure.
- **Very Poor:** AQI (301-400) – Respiratory illness on prolonged exposure.
- **Severe:** AQI (>401) – Effects even healthy people.

The generated values of AQI (in Step IVA)were binned and categorized under the categories mentioned above. It helped to obtain a picture of the severity of pollution across cities on a time-series basis. The generated variable was added to the dataset to make it ready for analysis purposes.

### c) Patterns in AQI Index in India

Using the dataset developed, analysis on AQI patterns over India could be made easily. Fig. 2 plots the average AQI numbers every quarter for India. Based on the figure, the following inferences could be drawn:

- The AQI values are highest in the Q4 (Winter months) of the year.
- The AQI values are lowest in Q3 (Monsoon).
- A sharp increase in AQI is visible between Q3–Q4-Q1.

A seasonal pattern in AQI is evident across the years. Poor AQI is visible in India across the winter months over the years. Also, the sheer decline in overall AQI values is visible between 2019 Q4 and 2020 Q1-Q3 (Denoted as a vertical yellow line in Fig.2). This was the time in India when lockdowns were being imposed.
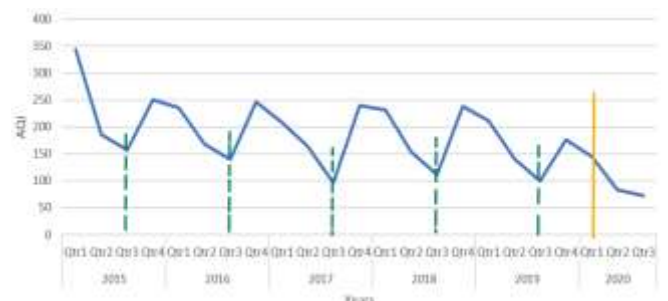


**Figure 2:** Average AQI Patterns in India between 2015 – 2020

### d) Specific Pollution Patterns in India



**Figure 3:** Monthly variation of various pollutants in India

With the help of power pivoting methods obtainable in Excel, the data could be wrangled and plotted on Tableau for analysis. Fig. 3 follows a similar scheme and plots the average pollutant concentration; recorded monthly over the

years. A similar pattern in the individual pollutants is observed which, was visible earlier in the AQI indices. The following are the deductions drawn from Fig.3:

- PM2.5 and PM10 pollution show a seasonal effect, with concentrations being more distinguished in winter months than summertime.
- Presence of a definite leaning that the pollution level in India falls in July and August. A possible answer to this might lie in the fact that monsoon season sets in during these months.

The median values of pollutant concentration for the year 2020 are less compared to other years, giving us discernment of a pollution reduction lately, possibly due to lockdowns put into place. For the most prominent pollutants, between the months January- May, generally a decline of:

- **41.1%** in AQI
- **36.7%** in PM10 &**52.33%** in PM 2.5 is visible.

### e) Effect on Lockdown in lowering the air pollution

Due to a very high surge in active COVID-19 cases and as a precautionary initial step, the government in India imposed a strict lockdown starting March 25th 2020. With the complete shutdown of services, the air pollution levels dropped significantly. Using the ACAPS Covid-19 Measures Dataset & Google Mobility Reports fetched via API calls, the effective study of the percentage of restrictions could be evaluated. The same is presented in Fig.4
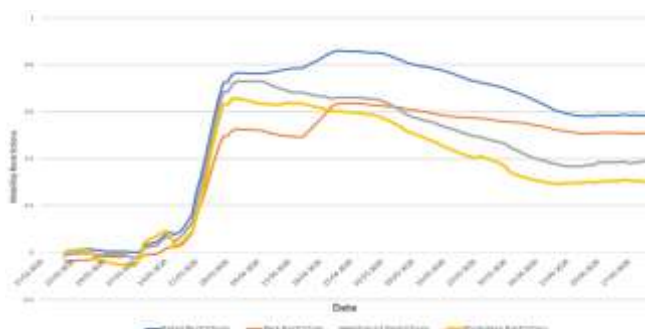


**Figure 4:** Mobility Reduction across categories for 2020 in India (7-day moving average with exponential smoothing) Following inferences could be drawn from Fig.4:

- Restrictions in – Retail activities, Park visits, transit (buses/airports/trains), and workplaces were seen post 14th March 2020.
- On 25th April peaks in restrictions were seen. Close to 85% of people stopped visiting shops, 66% fewer people travelled and offices were running only with 40% capacity.

The trend of restrictions with AQI could be compared by plotting the restrictions put into place with the AQI for that time frame. The same is represented in Fig. 5
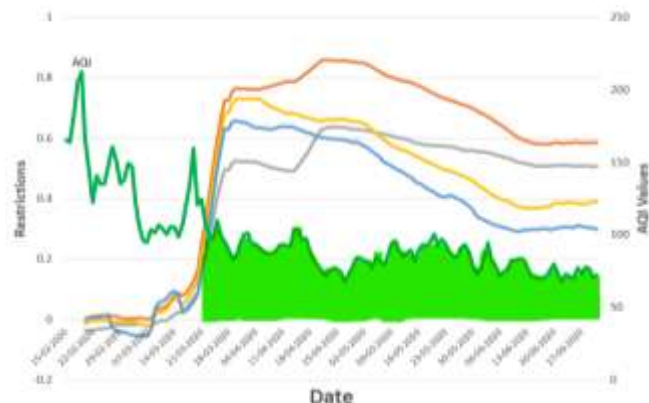


**Figure 5:** Mobility Restrictions vs AQI across India post lockdown



**Figure 6:** AQI across major cities in India (Red Line highlights the date on which lockdown was imposed)

From Fig. 5, as the lockdowns were imposed in India, a steep decrease in the AQI index was observed. The decline was as sharp as **61.2%** in comparison to **41.1%** that were observed earlier due to seasonal variations. This could infer a statistical relationship of AQI with lockdowns and restrictions being imposed. Using the similar approaches in Fig.6, the AQI for the major cities of India were plotted. The following were the inference that can be drawn from the graphs:

- The red vertical line highlights the exact date on which the lockdown came into effect in India.

- The majority of the Indian cities seem to be having a dangerously high level of pollution level during the initial months of the year,
- There appears to be a rapid decline after 25th March 2020 in AQI levels across India. The AQI during this phase fell under the category – **Good.**

**f) Hypothesis Testing for the Results**

**Table 9:** Hypothesis Testing for the results

| Values | AQI |
|---|---|
| Mean | 97.42703067 |
| Variance | 772.7954691 |
| Observations | 138 |
| Hypothesized Mean Difference | 0 |
| df | 137 |
| t Stat | 40.93990532 |
| P(T<=t) one-tail | 5.15737E-79 |
| t Critical one-tail | 1.65605208 |
| P(T<=t) two-tail | 1.03147E-78 |
| t Critical two-tail | 1.977431212 |

By using the T-Tailed Test for 2 Samples assuming unequal variances, we could do the hypothesis testing for our study. The following two hypotheses were built:
- $H_0$: No significant association between variables. AQI doesn't have a relationship with restrictions.
- $H_1$: Association present between the variables.

On running the T-Test the p-value is 1.03147E-78. Since:

*t Stat > t Critical two-tail.* **(40.93990532 > 1.977431212)**, we can safely reject our null hypothesis. This statistically approves for a significant relationship between the AQI and mobility restrictions. With the statistical inferences drawn above, we can say that the data of different years (2020 and previous years) have a comparable difference between their mean and variance values. Hence, the statistical testing made above provides a significant conclusion to draw out any "by chance possibility". This concludes that pollution decrease in the year 2022 is not because of natural causes but because of the lockdown measures imposed in the very same year.

## 5. Conclusions

This study effectively evaluates the mobility restrictions put in place by governments and their association with lowering the AQI levels to reduce air pollution. The AQI sub-index equations developed considers the pollution parameters and evaluates the overall AQI for the particular region. Using this modelling approach, the AQI was studied between 2015-2020 in India. The trends in AQI highlighted pollution seasonality. Q4 and Q1, the winter months in India have the highest pollution index with Q3 (Monsoon) being the lowest. A similar trend was observed for the specific individual pollutants as well. With the help of the ACAPS Covid-19 Government Measures dataset and Google's Mobility reports, the effect of lockdowns on AQI was studied statistically. Via API calls over Python, the data was fetched from Google on mobility restrictions in India during the start of the COVID-19 pandemic. Using methods of exponential smoothing and moving average (7-day period)

the parameters of mobility were studied with AQI to observe patterns. The inferences drawn from the analysis highlighted a drop of **61.2%** in AQI numbers, which in earlier years was only **41.1%** due to seasonality. The conclusion was tested statistically later with a t-test where the null hypothesis of lack of association between AQI and mobility restrictions were rejected. The study hence provided a significant light on the good effects of restrictions that were put into place. Similar to the studies made here, statistical modelling could be carried out across geographies to understand pollution control measures on a better note. With much greater diversity in the data of social interventions adopted to cover widespread geographies, a better analysis could be carried out for this study. The methods as such can help policymakers to regulate restrictions accordingly during the peak air pollution season for a country, helping in better decision making over environmental sustainability policies.

## 6. Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## 7. Acknowledgements

## References

[1] Coronavirus Disease (COVID-19) (2020), World Health Organization. [Online]. Accessed: March, 2021. Retrieved from : https://www.who.int/emergencies/diseases/novel-coronavirus-2019

[2] Coronavirus Resource Center (2020), Johns Hopkins University & Medicine. [Online]. Accessed: March, 2021. Retrieved from : https://coronavirus.jhu.edu/map.html

[3] A. C. Euloge Mouvoh, A. Bouchnita and A. Jebrane, "A contact-structured SEIR model to assess the impact of lockdown measures on the spread of COVID-19 in Morocco's population," 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), 2020, pp. 1-4, doi: 10.1109/ICECOCS50124.2020.9314462.

[4] S. Arora, A. Kumar and S. Sambhav, "Analysing the Effect of Gender on Mortality of COVID-19 Patients through Cox-Proportional Hazard Model," 2021 International Conference on Intelligent Technologies (CONIT), 2021, pp. 1-5, doi: 10.1109/CONIT51480.2021.9498331.

[5] C. Zhan, W. Jiang, J. Li, H. Xu and W. Sha, "Impact of COVID-19 Lockdown on Human Activity and Air Quality in China," 2020 IEEE International Symposium on Product Compliance Engineering-Asia (ISPCE-CN),

2020, pp. 1-5, doi: 10.1109/ISPCE-CN51288.2020.9321817.

[6] S. K. Sunori, P. B. Negi, S. Maurya, P. Juneja, A. Rana and Bhawana, "K-Means Clustering of Ambient Air Quality Data of Uttarakhand, India during Lockdown Period of Covid-19 Pandemic," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 1254-1259, doi: 10.1109/ICICT50816.2021.9358627.

[7] Blair Gordon S., Henrys Peter, Leeson Amber, Watkins John, Eastoe Emma, Jarvis Susan, Young Paul J., "Data Science of the Natural Environment: A Research Roadmap," Frontiers in Environmental Science, 2019, doi: 10.3389/fenvs.2019.00121

[8] COVID-19 Government Measures Dataset (2020), ACAPS. [Online]. Accessed: December, 2020. Retrieved from: https://www.acaps.org/covid-19-government-measures-dataset

[9] Google Community Mobility Reports (2020), Google. [Online]. Accessed: December, 2020. Retrieved from : https://www.google.com/covid19/mobility/

[10] Air Quality Data in India (2015 - 2020), Kaggle [Online]. Accessed: June, 2020. Retrieved from : https://www.kaggle.com/rohanrao/air-quality-data-in-india

[11] How AQI is calculated - Central Control Room for Air Quality Management : Delhi NCR [Online]. Accessed: June, 2020. Retrieved from : https://app.cpcbccr.com/ccr/

## Author Profile

**Mr. Aman Kumar** is currently a student at Amity School of Engineering and Technology, Amity University, Patna (India). He is also associated with Indian Institute of Technology, Madras (India) as a student at Computer Science Department. His research interests lie in the field of Data Sciences, Big Data, Machine Learning and Cloud Computing.

**Ms. Ankita Singh** is currently a final year student at Amity School of Engineering and Technology, Amity University, Patna (India), pursuing her Bachelors in Technology Her research interests lie in the field of Cloud Computing, Data Sciences, Machine Learning and Computer Networks.