Term Project - Impact of COVID-19 Lockdowns on the Air Quality Index

Vaibhav Ramakrishnan

2022-12-09

1) Introduction

Countries began developing various strategies to combat COVID-19's spread in early 2020. One of the most visible methods was the implementation of nationwide lockdowns and the prohibition of international travel. While this initially slowed the spread of COVID-19, it was not economically feasible in the long run. However, it did demonstrate to the world how the strict lockdowns benefited the environment, as air quality improved as a result of the factories being closed and fewer vehicles on the roads, leading to reduced transportation and emissions, among other factors. The first case of COVID-19 in India was reported on January 30, 2020, in the state of Kerala. The infected person was an international student in Wuhan, China, who had returned to India for vacation. As the number of reported cases increased across the country to around 500, the Indian government decided to impose a 21-day nationwide lockdown on March 25, 2020. As a result, major industrial activities were halted, public transportation was halted, people were confined to their homes, social distance was created, and so on. The government also extended the lockdown by 14 days. As the lockdown wore on, the government relaxed some restrictions, such as allowing goods traffic by transport trucks and allowing people stranded in various parts of the country to travel by road. Finally, on May 31, 2020, the final Phase 4 of the lockdown came to an end.

During this period of restrictions, many positive changes in the environment were observed, including a significant improvement in air quality (AQI). To determine the true impact of the lockdowns on the pollution index, the Air Quality Index (AQI) data was used to estimate the difference in air quality. By comparing Iran's air quality to major cities such as Paris, Wuhan, Madrid, and Rome, Yazdani et al. (2021) contextualised the findings. The data was compared over a five-year period. The study discovered that by drastically reducing human activity, pollutant levels in all cities significantly decreased. Many major cities in major countries face air pollution issues, and the government has implemented various strategies such as setting an emission standard for vehicles, encouraging public transportation and promoting the use of fans instead of air conditioners, as well as banning firecrackers during major festivals, but most of them have not made a significant difference. Delhi and Mumbai are among the most polluted cities in the world, with doctors claiming that the toxic air in Delhi is equivalent to smoking 10 to 15 cigarettes per day - Pinto (2020). (*How Breathing in Indian Cities is Equivalent to Smoking Packs of Cigarettes in One Day* (29 January,2020), from https://weather.com/en-IN/india/pollution/news/2020-01-29-how-breathing-indian-cities-equivalent-smoking-packs-cigarettes-day).

The pandemic has resulted in self-isolation and lockdowns all over the world. We are now nearly three years into the pandemic, but the origins of the virus remain unknown. Vaccines have been a huge help, but the sheer number of vaccines available has caused some confusion. While many doctors and health experts have conducted research to develop effective anti-virus strategies, many environmental experts believe that strict quarantine measures implemented around the world have given nature time to heal. Benchrif et al. (2021) calculate the impact of lockdowns on 21 cities with populations greater than one million people worldwide and discovered a reduction in daily NO_2 (Nitrogen Dioxide) concentrations ranging from 3 to 58% in all of them. The current study plans to conduct a statistical analysis of the difference in concentrations of major pollutants that comprise the Air Quality Index (AQI) in Delhi and Mumbai in 2020. In India, Lokhandwala and Gautam (2020) conducted a similar analysis using an Analysis of Variance (ANOVA) over various pollutants and discovered that not only the air quality but also the river quality had improved across the country. Fu, Purvis-Roberts, and Williams (2020) measured and found the difference in AQI in 20 major cities across six continents where the virus was most severe. The AQI (pollutant concentrations) fell precipitously. Their analysis formed the basis for the current study. They conducted their research using ANOVA, Boxplots, and Tukey statistics. Among all cities in their study, Delhi had the greatest reduction in AQI pollutant concentrations. When we look at the AQI statistics of countries, most of them saw the greatest reductions in Nitrogen Dioxide and Carbon Monoxide due to restrictions on automobile use, which is the primary source of these pollutants. PM (particulate matter) concentrations have also decreased significantly as a result of reduced road traffic. Tripathi (2021) conducted their research in four major Indian cities: Delhi, Mumbai, Kolkata, and Hyderabad, and found a significant drop in AQI levels in all of them. Their thorough examination spanned four phases of the lockdown, revealing a gradual increase in pollutant levels as restrictions were lifted. Python programming and its libraries were also used to create the graphs. This demonstrates to prospective researchers how to incorporate programming into their studies or research. Gupta et al. (2022) conducted a thorough analysis on a timeline of different phases of lockdown periods as well as before and after the lockdown periods (Pre-Lockdown, 1st Lockdown (March 25 - April 14), 2nd Lockdown (April 15 - May 3), 3rd Lockdown (May 4 - May 17), 4th Lockdown (May 18 - May 31), Unlock Phase 1 (June 1 - June 30), Unlock Phase 2 (June 1 (July 1 – July 31). In addition to the AQI, their research focused on the river water quality index. Their findings eventually concluded that the lockdown was a blessing in disguise, as it could help governments maintain air and water quality by lowering pollution levels while also keeping the economy running. For a detailed analysis of Delhi's AQI, see Mahato, Pal, and Ghosh (2020), which concluded a reduction of approximately 57% to 33% in PM_{10} and $PM_{2.5}$ concentrations, respectively. The most intriguing finding was that after just one day of lockdown (March 25th), there is a 40% improvement in air quality. While on the fourth day of the lockdown (March 27th) the PM_{10} and $PM_{2.5}$ concentrations returned to acceptable levels, and the National Air Quality Index had dropped by about 51%.

Interestingly, a National Geographic article by Gardiner (2020) (*Pollution made COVID-19 worse. Now, lockdowns are clearing the air*, National Geographic,https://www.nationalgeographic.com/science/article/pollution-made-the-pandemic-worse-but-lockdowns-clean-the-sky) reported the findings of Harvard researchers who studied the PM 2.5 concentrations in the blood of COVID-19 fatalities and came to the conclusion that the death rate was indeed 15% higher in countries with higher levels of PM 2.5 in the air.PM (particulate matter) are extremely dangerous minute particles in the air that include smoke, sulphate, and metals. Carbon monoxide CO is released into the atmosphere as a result of vehicle emissions and the combustion of fossil fuels. Power plants, fuel combustion, vehicle exhaust, and other sources emit NO_2, which is composed of nitrogen and oxygen. Furthermore, when considering lockdowns as an alternative strategy to reduce air pollution, the policy implications must be thoroughly considered, as well as whether restricting only a few vehicles from driving during these restrictions has a significant impact or not (Wang, Liu, and Zheng (2021)). According to Schaefer, Tuitjer, and Levin-Keitel (2021), a study that discusses the decrease in demand for public transportation in Hannover (Germany), people switched to bikes and personal cars out of fear. The study by Chang, Meyerhoefer, and Yang (2021) discusses how increasing or decreasing fares in public transits promotes the use of public transportation, which can help transportation authorities increase ridership and reduce air pollution because metros and subways reduce the combustion of fossil fuels.

Air quality will undoubtedly improve due to the reduction in construction and industrial activity during the lockdowns, but will this difference be significant enough? The purpose of this study is to see if there was a significant difference in air quality before and after lockdowns in two major Indian cities, Delhi and Mumbai, and if lockdowns can be used as a future strategy to combat air pollution.

2) Method

2.1) Estimation of Air Quality Index

The Air Quality Index helps the governments across the globe to communicate with people, as to how polluted the air is and the kinds of precautions they need to take if the air category is Severe. Each pollutant's baseline is used to calculate the Air Quality Index (AQI). The major pollutants considered when calculating AQI are $PM2.5, PM10, NO, NO_2, NOX, CO, SO2, NH_3$ and O_3 and their average is taken, with the highest value for each pollutant considered. $PM10, NO_2\ and\ CO$ would be considered and evaluated for the current study. The categories of AQI along with their health implications can be defined as below in Table 1 –

Table 1: CPCB AQI Calculation Framework

Category of AQI	**Range**	**Health Implications**
Good	AQI of 0-50	Air quality is deemed adequate, and air pollution presents minimal or no risk.
Satisfactory	AQI of 51-100	Air quality is acceptable, some pollutants may pose a moderate health risk to few.
Moderately Polluted	AQI of 101-200	Members of vulnerable groups may suffer health consequences. The general public is unlikely to be harmed.
Poor	AQI of 201-300	People might notice health effects; members of vulnerable groups may notice severe health consequences.
Extremely Poor	AQI of 301-400	Health alerts for emergency situations. The general public is much more inclined to be affected.
Severe	AQI of 401-500	Everyone is at risk of experiencing more serious health consequences.

The New Delhi, India based Central Pollution Control Board (CPCB 2015) calculation framework was used to assess the AQI in this study.

2.2) Area of Study

The study has primarily focused on two populous cities of India - Delhi and Mumbai. According to the United Nations, as of 2018, Delhi is the second-largest mega city of the world, having a population of over 16.78 million. Delhi is located in Northern India and has dual status as a state and city. Mumbai is also known as the financial capital of India. Mumbai is one of India's most populous cities and ranked the 17th most populous city globally with a population of over 20 million. Both of these places are known as car-congested cities and both cities' state governments imposed strict restrictions and shut down all industrial activities. The Indian lockdown was announced in late March and lasted until May, after which a phased lifting of restrictions began. Once, the restrictions were lifted, the human activites resumed in both the cities, which resulted in significant rise in the levels of pollutants.

2.3) Dataset Description

The .CSV datasets for this study were obtained from the World's Air Quality Index (WAQI) project, which is an online public dataset, which created datasets after the virus's outbreak covering the pandemic phases, for COVID-19-related research - WAQIProjectTeam (2022) (https://aqicn.org/data-platform/register/) [Datset]. The WAQI project provides the option to choose from different areas and/or AQI stations from a state. The cities' data sources are the Air headquarters in R.K Puram in Delhi and Worli in Mumbai. These two residential colonies were chosen for their central location in their respective states as well as their low population density. By filtering the dataset and selecting the year 2020 for analysis, the dataset for both cities Delhi and Mumbai contained approximately 360-365 samples from various months in 2020. The information for these two cities was obtained from (CPCB - India Central Pollution Control Board). The data required enough samples to make a reasonable comparison across months for the analysis. The .CSV file had 6 major pollutants' (SO_2, NO_2, $PM10$, $PM25$, O_3, CO) concentration reports for most days. This gave the confidence to move forward with the analysis and produce reliable results. $PM10$ and $PM25$ particulate concentrations were measured in µg/m3, and other pollutants such as Sulphur dioxide (SO_2), Nitrogen dioxide (NO_2), and Carbon monoxide (CO) were measured in parts per million (ppm). For the current study, we will focus on primary pollutants NO_2 and CO, and particulate matter $PM10$ concentrations.

2.4) Approach and statistical analysis used to explore project goals

R programming (RStudio) is used to conduct our statistical analysis and visualise our results using boxplots, as shown in the Result section below. Boxplots are an excellent tool for comparing a numerical variable (on the y-axis) representing pollutant concentration in our study and a categorical variable (on the x-axis) representing months of the year 2020. A boxplot, in essence, can show the variation in concentration levels across months. The current study shows a decrease in the concentration levels of all three pollutants NO_2, CO, and $PM10$ in both Delhi and Mumbai during the months of March, April, and May 2020 when compared to other months of the year.

Tukey's Honest Significant Difference Test function (Tukey's HSD) is used in the study to examine the difference in mean concentration of the respective pollutants between pairwise months. Tukey's HSD test determines the greatest value of the difference between two pairs. If the p-value for a pair of months is 0.05, we can conclude that the difference between them is significant (assuming our significance level is 0.05). We will mostly compare the months of April and May 2020 to months with few restrictions, such as November and December 2020.

Furthermore, using the linear model function *lm()* in R, we were able to summarise our findings and conclude that the months of April and May had the lowest levels of $NO2$, CO, and $PM10$ concentrations when compared to other months. A linear model will aid in the description of a continuous response variable (pollutant concentration) as a function of a predictor variable (month / date of year). Similar to Tukey's HSD, the output gives the individual concentration (intercept value) of different months. This allows us to determine which months had the highest and lowest concentrations of various pollutants.

Finally, to draw meaningful conclusions from the statistical analysis, the study employs the *summary()* function over the ANOVA model (due to the study's use of a categorical variable *Month* and a continuous variable *Pollutant concentration*). The output shows the intercept values for the 12 months of 2020, demonstrating that the month with the highest intercept value has the highest concentration of the relevant pollutant.

3) Result

Only the the major pollutants that contribute to the AQI are considered in the current study: nitrogen dioxide (NO_2), particulate matter ($PM10$), and carbon monoxide (CO). From late March 2020 to May 2020, India experienced its first strict lockdown, which was gradually lifted. To answer the main question, whether the lockdown has a significant effect on air quality, or whether there is a significant improvement or difference in air quality if the lockdown is not in place, we examine the months of March, April, and May 2020. By comparing pollutant concentrations in two major cities, Delhi and Mumbai, both of which are among the most polluted in the world, thus can givet a clear picture and justifiable results. The research question *If lockdowns can be used as a strategy to combat air pollution* would be answered if there is a significant difference in pollutant levels between these three months and the rest of the year.

The boxplots in Figures below, show the decline in the concentration levels of all three pollutants during the months of March, April, and May 2020 as compared to other months of the year.

3.1) Analysis for Delhi

```
delaqi <- read.csv("r.k.-puram, delhi, delhi, india-air-quality.csv",
                   header=TRUE) #Read the excel file for Delhi AQI data

#Keep the months in order, by using factor function with levels argument
delaqi$date <- factor(delaqi$date, levels = c("Jan", "Feb", "Mar","Apr",
                                              "May","Jun","Jul","Aug",
                                              "Sep","Oct","Nov","Dec"))
```

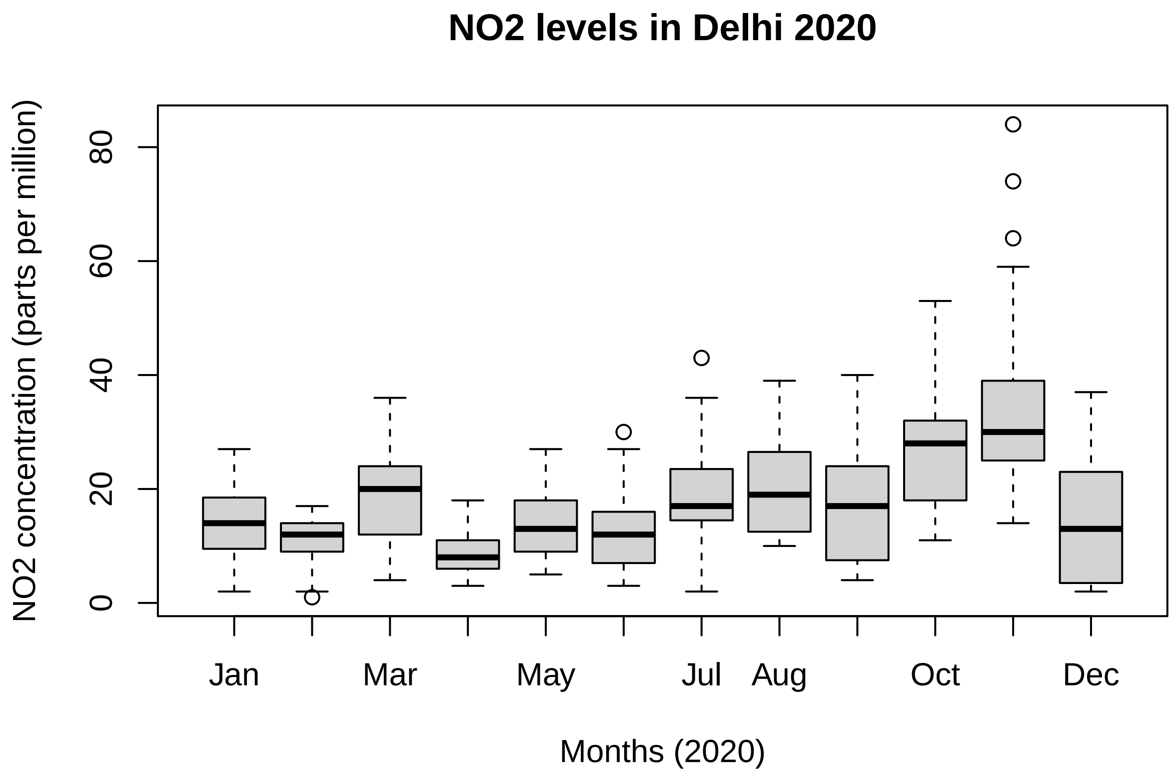

Figure 1: NO_2 concentration in Delhi during 2020

The above Figure-1 clearly depicts the difference in the NO_2 concentration levels between the months of April and November. Since the lockdown was gradually lifted in stages after the month of May, concentration levels have steadily increased, with a peak at the end of the year in November.

This can be statistically verified by using the TukeyHSD function as below which shows the pairwise concentration comparison between months in 2020 with April being lowest. Tukey's Honest Significant Difference Test showed if we were to compare the months of April (month of full lockdown) and November (month with almost no restrictions), there was a significant difference between the two months (with p value of < 0.05).

```
#TukeyHSD function for NO2
#by passing ANOVA function aov to our linear model
#continuous variable NO2  and categorical variable date
TukeyHSD(aov(lm(no2 ~ date, data = delaqi)))
```

A similar linear model function as well can be used with the NO_2 concentrations being the continuous variable and the *date* or month of the year being the categorical variable, as shown in below R code. The output clearly shows the month of April with the lowest concentration of NO_2 in comparison to other months in the year 2020.

```r
#summary() function to produce result summary
#of linear model of our data
summary(lm(no2 ~ date, data = delaqi))
```

```
##
## Call:
## lm(formula = no2 ~ date, data = delaqi)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -21.433  -6.419  -0.667   4.935  48.567
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  14.4194     1.6409   8.788  < 2e-16 ***
## dateFeb      -3.3394     2.4559  -1.360  0.17480
## dateMar       4.4839     2.3206   1.932  0.05415 .
## dateApr      -5.6194     2.3398  -2.402  0.01685 *
## dateMay      -0.8387     2.3206  -0.361  0.71800
## dateJun      -1.6607     2.3602  -0.704  0.48214
## dateJul       4.8065     2.3206   2.071  0.03908 *
## dateAug       6.3871     2.3206   2.752  0.00623 **
## dateSep       2.2473     2.4050   0.934  0.35073
## dateOct      11.8140     2.3398   5.049  7.2e-07 ***
## dateNov      21.0140     2.3398   8.981  < 2e-16 ***
## dateDec       0.6452     2.3206   0.278  0.78117
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.136 on 345 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.3754, Adjusted R-squared:  0.3555
## F-statistic: 18.85 on 11 and 345 DF,  p-value: < 2.2e-16
```

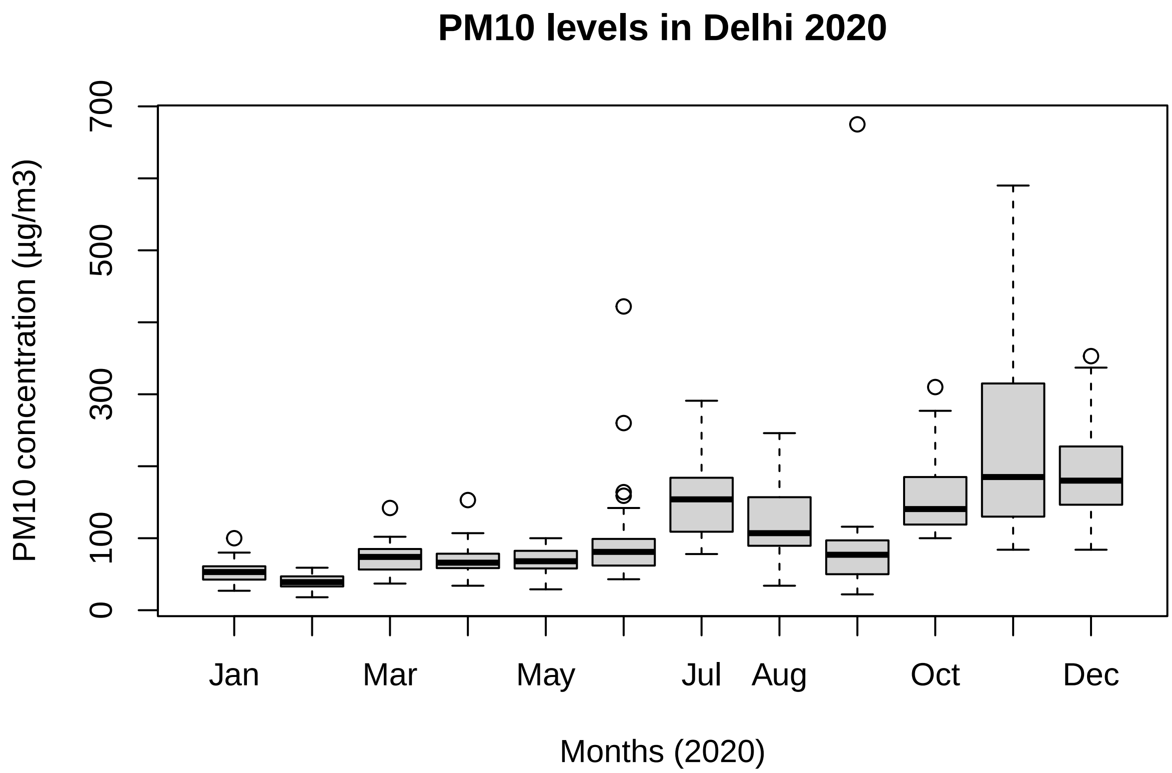

Figure 2: $PM10$ concentration in Delhi during 2020

The above Figure-2 clearly depicts the difference in the $PM10$ concentration levels between the months of April and November.

This can be statistically verified by using the TukeyHSD function as below which shows the pairwise concentration comparison between months in 2020. Tukey's Honest Significant Difference Test showed if we were to compare the months of April (month of full lockdown) and November (month with almost no restrictions), there was a significant difference between the two months (with p value of < 0.05).

```
#TukeyHSD function for PM10
#by passing ANOVA function aov to our linear model
#continuous variable PM10 and categorical variable date
TukeyHSD(aov(lm(pm10 ~ date, data = delaqi)))
```

A similar linear model function can be used as well,with $PM10$ concentrations being the continuous variable and the *date* or month of the year being the categorical variable, as shown in below R code. The output clearly shows that during the lockdown period (March, April and May 2020) the concentration of $PM10$ is relatively stable and low in comparison to post unlock period (June - December 2020).

```r
#summary() function to produce result summary
#of linear model of our data
summary(lm(pm10 ~ date, data = delaqi))
```

```
##
## Call:
## lm(formula = pm10 ~ date, data = delaqi)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -147.60  -29.00   -6.58   15.15  582.00
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    53.32      11.87   4.492 9.63e-06 ***
## dateFeb       -13.48      17.57  -0.767  0.44373
## dateMar        20.26      16.79   1.207  0.22833
## dateApr        16.82      17.23   0.976  0.32965
## dateMay        16.42      16.79   0.978  0.32870
## dateJun        47.58      16.93   2.811  0.00522 **
## dateJul       100.68      16.79   5.998 5.06e-09 ***
## dateAug        70.39      16.79   4.193 3.50e-05 ***
## dateSep        39.68      17.40   2.281  0.02318 *
## dateOct       105.61      16.93   6.240 1.29e-09 ***
## dateNov       178.28      16.93  10.533  < 2e-16 ***
## dateDec       141.10      16.79   8.405 1.13e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 66.09 on 345 degrees of freedom
##   (6 observations deleted due to missingness)
## Multiple R-squared:  0.4366, Adjusted R-squared:  0.4186
## F-statistic:  24.3 on 11 and 345 DF,  p-value: < 2.2e-16
```

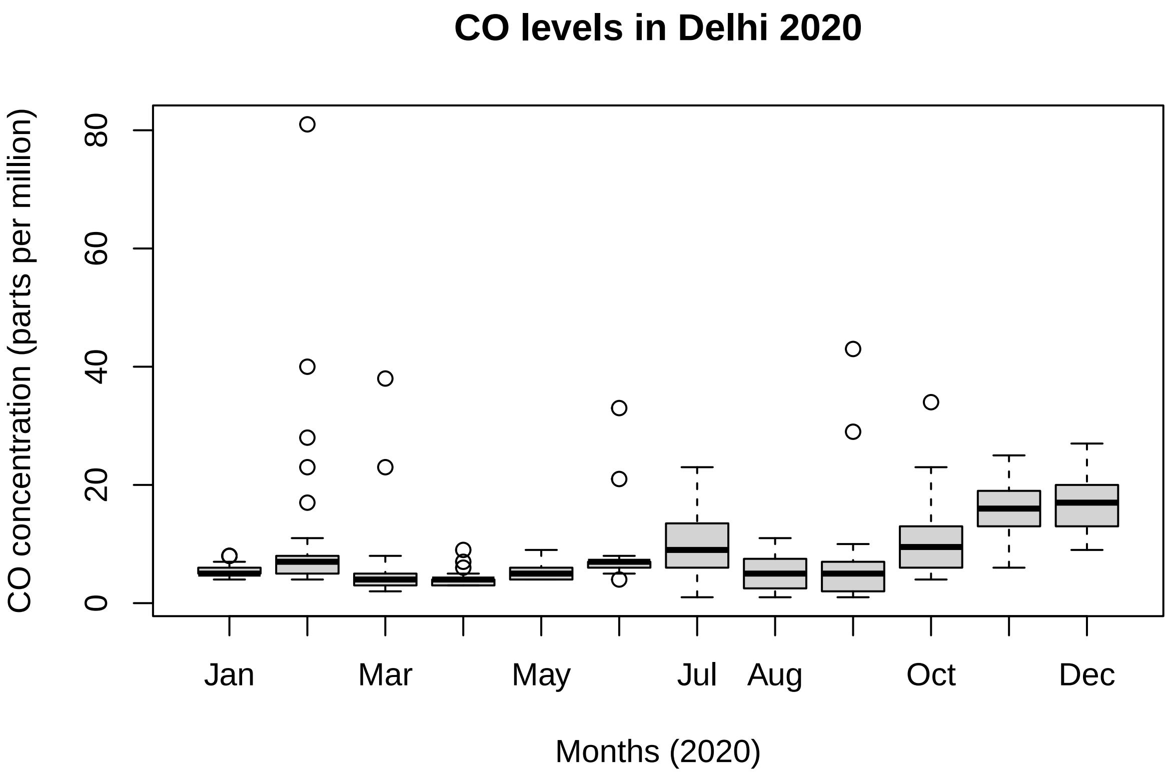

Figure 3: *CO* concentration in Delhi during 2020

The above Figure-3 clearly depicts the difference in the *CO* concentration levels between the months of April and November.

This can be statistically verified again by using the TukeyHSD function as below which shows the pairwise concentration comparison between months in 2020. Tukey's Honest Significant Difference Test showed if we were to compare the months of April (month of full lockdown) and November (month with almost no restrictions), there was a significant difference between the two months (with p value of < 0.05).

```
#TukeyHSD function for CO
#by passing ANOVA function aov to our linear model
#continuous variable CO and categorical variable date
TukeyHSD(aov(lm(co ~ date, data = delaqi)))
```

A similar linear model function can be used as well,with *CO* concentrations being the continuous variable and the *date* or month of the year being the categorical variable, as shown in below R code. The output clearly shows the month of April with the lowest concentration of *CO* in comparison to other months in the year 2020.

```r
#summary() function to produce result summary
#of linear model of our data
summary(lm(co ~ date, data = delaqi))
```

```
##
## Call:
## lm(formula = co ~ date, data = delaqi)
##
## Residuals:
##     Min      1Q Median      3Q     Max
## -9.733 -2.900 -0.900  0.862 68.654
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   5.5806     1.1820   4.721 3.43e-06 ***
## dateFeb       6.7655     1.7501   3.866 0.000133 ***
## dateMar       0.5675     1.7324   0.328 0.743431
## dateApr      -1.4427     1.7002  -0.849 0.396721
## dateMay      -0.1613     1.6716  -0.096 0.923190
## dateJun       2.3194     1.6855   1.376 0.169702
## dateJul       4.1613     1.6716   2.489 0.013273 *
## dateAug      -0.5161     1.6716  -0.309 0.757691
## dateSep       1.4994     1.7691   0.848 0.397289
## dateOct       5.5860     1.6855   3.314 0.001018 **
## dateNov      10.1527     1.6855   6.024 4.43e-09 ***
## dateDec      11.0000     1.6716   6.581 1.77e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 6.581 on 340 degrees of freedom
##   (11 observations deleted due to missingness)
## Multiple R-squared:  0.2847, Adjusted R-squared:  0.2616
## F-statistic:  12.3 on 11 and 340 DF,  p-value: < 2.2e-16
```

3.2) Analysis for Mumbai

```r
mumaqi <- read.csv("worli,-mumbai, india-air-quality.csv",header=TRUE)
#Read the excel file for Mumbai AQI data

#Keep the months in order, by using factor function with levels argument
mumaqi$date <- factor(mumaqi$date, levels = c("Jan", "Feb", "Mar","Apr",
                                              "May","Jun","Jul","Aug",
                                              "Sep","Oct","Nov","Dec"))
```

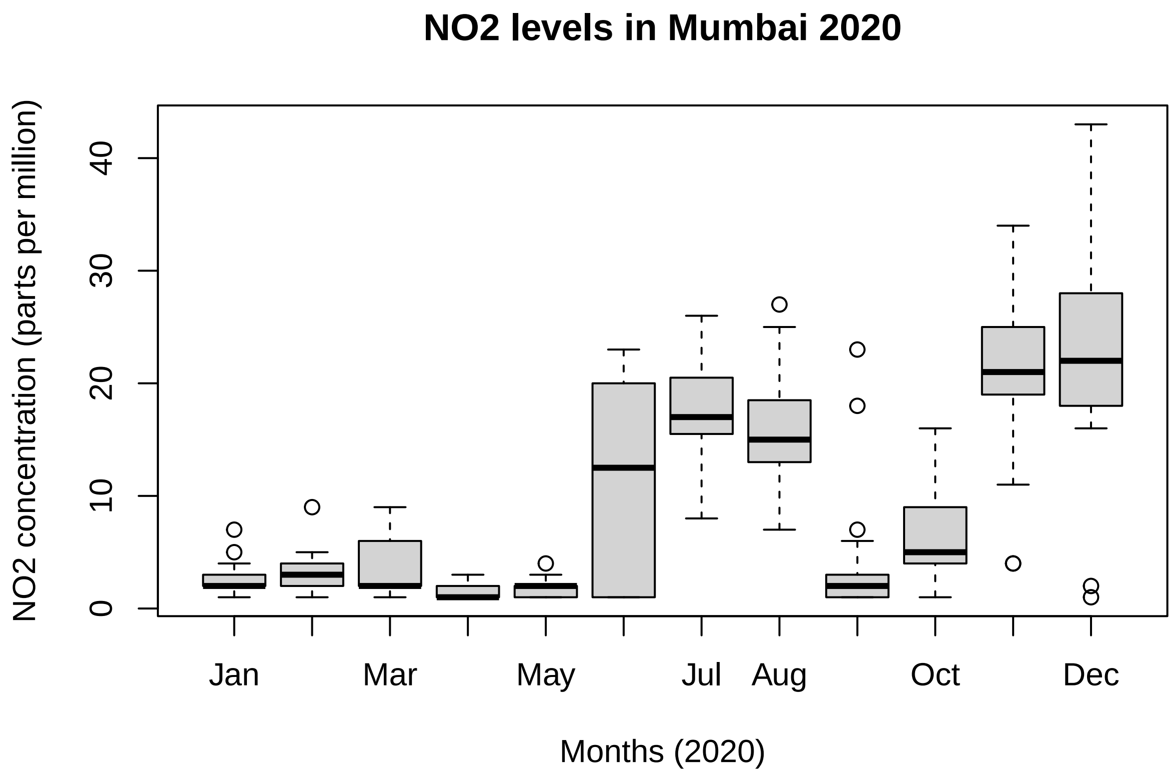

Figure 4: NO_2 concentration in Mumbai during 2020

The above Figure-4 clearly depicts the difference in the NO_2 concentration levels between the months of April and November.

This can be statistically verified by using the TukeyHSD function as below which shows the pairwise concentration comparison between months in 2020 with April being lowest. Tukey's Honest Significant Difference Test showed if we were to compare the months of April (month of full lockdown) and November (month with almost no restrictions), there was a significant difference between the two months (with p value of < 0.05).

```
#TukeyHSD function for NO2
#by passing ANOVA function aov to our linear model
#continuous variable NO2 and categorical variable date
TukeyHSD(aov(lm(no2 ~ date, data = mumaqi)))
```

A similar linear model function can be used as well,with the NO_2 concentrations being the continuous variable and the *date* or month of the year being the categorical variable, as shown in below R code. The output clearly shows the month of April with the lowest concentration of NO_2 in comparison to other months in the year 2020.

```r
#summary() function to produce result summary
#of linear model of our data
summary(lm(no2 ~ date, data = mumaqi))
```

```
##
## Call:
## lm(formula = no2 ~ date, data = mumaqi)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -21.7742  -1.7241  -0.5172   1.3704  20.2258
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.6296     0.9434   2.788  0.00561 **
## dateFeb       0.4418     1.3221   0.334  0.73847
## dateMar       0.8037     1.3003   0.618  0.53694
## dateApr      -1.1124     1.3109  -0.849  0.39672
## dateMay      -0.8232     1.2904  -0.638  0.52394
## dateJun       8.7550     1.3469   6.500 2.84e-10 ***
## dateJul      15.1768     1.2904  11.762  < 2e-16 ***
## dateAug      13.3704     1.2904  10.362  < 2e-16 ***
## dateSep       1.0945     1.3109   0.835  0.40434
## dateOct       4.0478     1.2904   3.137  0.00186 **
## dateNov      18.3370     1.3003  14.102  < 2e-16 ***
## dateDec      20.1446     1.2904  15.612  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.902 on 342 degrees of freedom
##   (11 observations deleted due to missingness)
## Multiple R-squared:  0.7224, Adjusted R-squared:  0.7134
## F-statistic: 80.89 on 11 and 342 DF,  p-value: < 2.2e-16
```

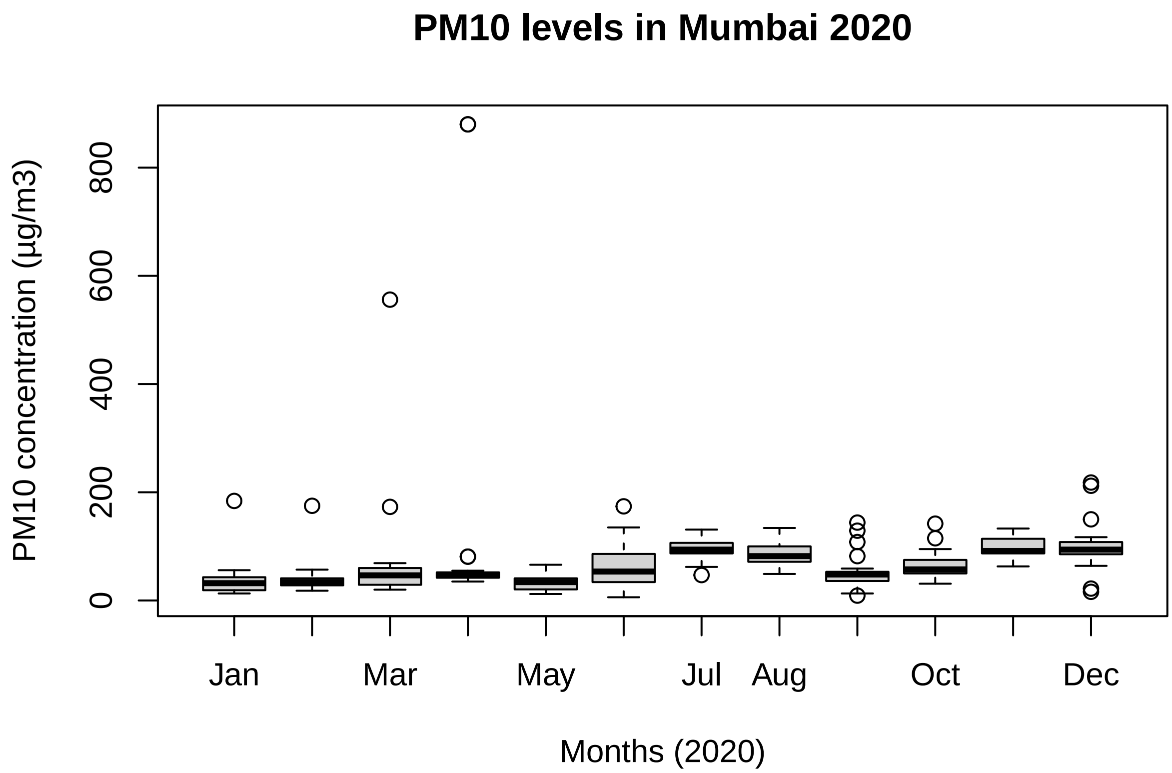

Figure 5: $PM10$ concentration in Mumbai during 2020

The above Figure-5 clearly depicts the difference in the $PM10$ concentration levels between the months of April and November.

This can be statistically verified by using the TukeyHSD function as below which shows the pairwise concentration comparison between months in 2020. Tukey's Honest Significant Difference Test showed if we were to compare the months of April (month of full lockdown) and December (month with almost no restrictions), there was a significant difference between the two months (with p value of < 0.05).

```
#TukeyHSD function for PM10
#by passing ANOVA function aov to our linear model
#continuous variable PM10 and categorical variable date
TukeyHSD(aov(lm(pm10 ~ date, data = mumaqi)))
```

A similar linear model function can be used as well,with the $PM10$ concentrations being the continuous variable and the *date* or month of the year being the categorical variable, as shown in below R code. The output clearly shows that during the lockdown period the month of May 2020 showed lowest concentration of $PM10$ in the whole year.

```r
#summary() function to produce result summary
#of linear model of our data
summary(lm(pm10 ~ date, data = mumaqi))
```

```
##
## Call:
## lm(formula = pm10 ~ date, data = mumaqi)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -83.77 -18.47  -6.61   7.20 774.28
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   36.621     13.398   2.733 0.006598 **
## dateFeb        2.379     19.296   0.123 0.901937
## dateMar       29.986     19.117   1.569 0.117665
## dateApr       69.103     18.948   3.647 0.000307 ***
## dateMay       -3.717     18.640  -0.199 0.842042
## dateJun       26.146     18.790   1.392 0.164975
## dateJul       57.992     18.640   3.111 0.002020 **
## dateAug       49.315     18.640   2.646 0.008530 **
## dateSep       16.231     19.296   0.841 0.400840
## dateOct       26.313     18.790   1.400 0.162306
## dateNov       61.179     18.790   3.256 0.001243 **
## dateDec       63.154     18.640   3.388 0.000786 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 72.15 on 342 degrees of freedom
##   (11 observations deleted due to missingness)
## Multiple R-squared:  0.1129, Adjusted R-squared:  0.08434
## F-statistic: 3.956 on 11 and 342 DF,  p-value: 1.981e-05
```

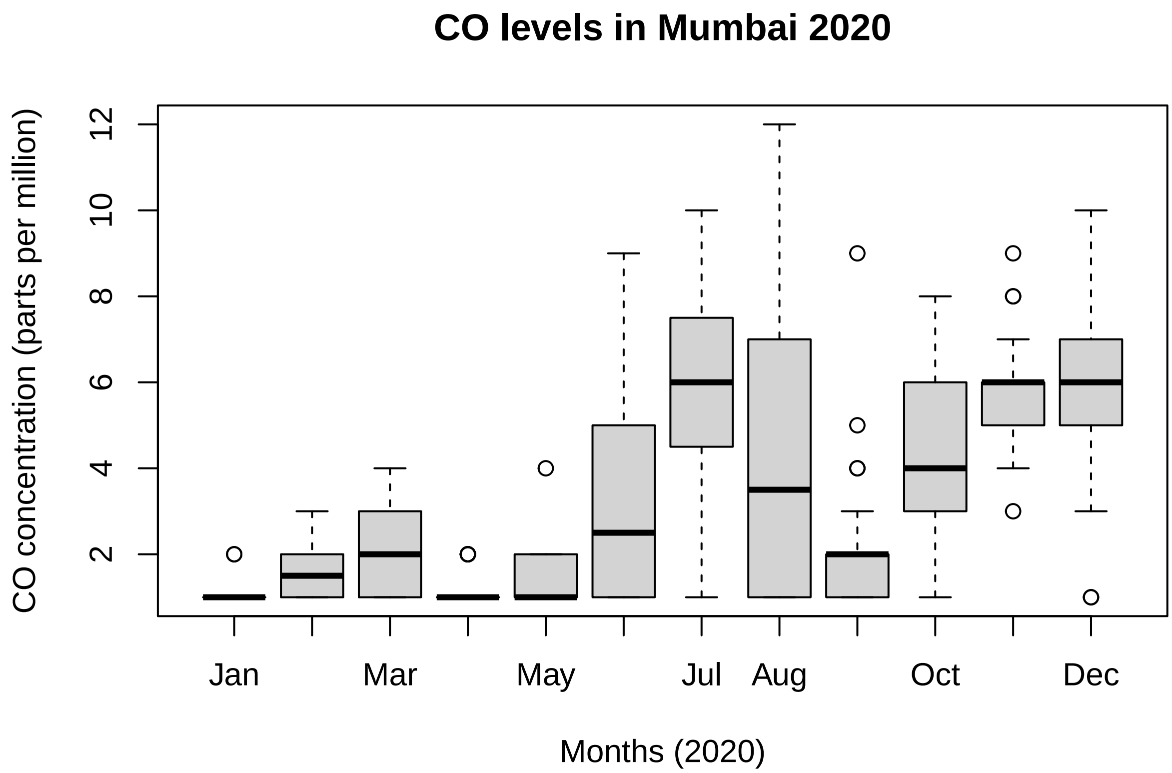

Figure 6: CO concentration in Mumbai during 2020

The above Figure-6 clearly depicts the difference in the CO concentration levels between the months of April and November.

This can be statistically verified again by using the TukeyHSD function as below which shows the pairwise concentration comparison between months in 2020. Tukey's Honest Significant Difference Test showed if we were to compare the months of April (month of full lockdown) and November (month with almost no restrictions), there was a significant difference between the two months (with p value of < 0.05).

```
#TukeyHSD function for CO
#by passing ANOVA function aov to our linear model
#continuous variable CO and categorical variable date
TukeyHSD(aov(lm(co ~ date, data = mumaqi)))
```

A similar linear model function can be used as well, with the CO concentrations being the continuous variable and the *date* or month of the year being the categorical variable, as shown in below R code. The output clearly shows the month of April with the lowest concentration of CO in comparison to other months in the year 2020.

```r
#summary() function to produce result summary
#of linear model of our data
summary(lm(co ~ date, data = mumaqi))
```

```
##
## Call:
## lm(formula = co ~ date, data = mumaqi)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.7419 -1.1034 -0.1667  0.8333  7.8333
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.22222    0.62062   1.969  0.04982 *
## dateFeb      0.37778    0.85546   0.442  0.65909
## dateMar      0.88123    0.71042   1.240  0.21578
## dateApr     -0.05556    0.70761  -0.079  0.93747
## dateMay      0.13262    0.70498   0.188  0.85091
## dateJun      2.04444    0.70761   2.889  0.00414 **
## dateJul      4.51971    0.70498   6.411 5.53e-10 ***
## dateAug      2.94444    0.70761   4.161 4.13e-05 ***
## dateSep      1.03865    0.73204   1.419  0.15698
## dateOct      3.16487    0.70498   4.489 1.02e-05 ***
## dateNov      4.47778    0.70761   6.328 8.92e-10 ***
## dateDec      4.48746    0.70498   6.365 7.20e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.862 on 303 degrees of freedom
##   (50 observations deleted due to missingness)
## Multiple R-squared:  0.4807, Adjusted R-squared:  0.4619
## F-statistic:  25.5 on 11 and 303 DF,  p-value: < 2.2e-16
```

4) Discussion

The current study attempted to investigate the impact of COVID-19 lockdowns on the Air Quality Index in Delhi and Mumbai in March, April, and May of early 2020. The main pollutants in this study were NO_2, CO, and $PM10$. The findings clearly demonstrate a significant improvement in air quality during lockdowns in both cities, owing to a decrease in pollutant concentrations. In addition, after the lockdown restrictions were lifted, the concentrations of NO_2, CO, AND $PM10$ began to rise in both cities. This helped us conclude that industrial pollutants and vehicular emissions are the major sources of pollution in these cities. Significant reductions in air pollution, on the other hand, can be achieved through careful planning and phased restrictions on industrial activities, vehicle movement, and traffic. (Mahato, Pal, and Ghosh (2020)). This helped us answer our question about whether lockdowns should be considered as an alternative strategy for dealing with air pollution in the future because it requires the government to consider economic implications.

The discovery in current study is related to the work of Benchrif et al. (2021), who investigated the impact of lockdowns on the Air Quality Index in 21 cities around the world with populations greater than one million people. They also concluded that PM concentrations in 2020 would be lower than they were prior to the pandemic.@wang2021air, on the other hand, discovered that the reductions in NO_2 concentrations were more significant than the reductions in PM concentrations. This holds true in the current research as well, where the difference in PM levels is slightly smaller than the difference in NO_2 levels. The current research found a significant increase in the levels of these pollutants after the month of May, which is particularly significant given that Delhi and Mumbai are two of India's largest cities and major industrial centres. This demonstrates how geography can influence concentration levels in different cities around the world, depending on factors such as the number and type of industries in a city, population density, vehicular emissions, biomass burning, thermal power generation, and others (Benchrif et al. (2021)).

In comparison to the findings of Tripathi (2021)'s study, which shows an increase in pollutant levels in cities such as Delhi and Kolkata following a lockdown, the findings are similar for this research. This demonstrated that major urban cities with dense populations and a diverse range of industries will show the most dramatic improvement and results in these types of analyses. Yazdani et al. (2021) contrasted Tehran's AQI levels with those of four other major cities: Paris, Wuhan, Rome, and Madrid. Each comparison proved significant, and air quality improved significantly in all cities. Wuhan, the most densely populated city studied and home to a diverse range of industries, experienced the greatest reduction in AQI, with a 22% improvement.

Mahato, Pal, and Ghosh (2020) talk about significant improvements in air quality in parts of Delhi where energy footprints were high and lockdown guidelines were followed. Their study also discuss briefly about how a 2-4 day lockdown can be implemented to combat air pollution, but the economic implications must be thoroughly investigated. This relates to current study findings about the importance of planning for a systemic approach to tackling air pollution in megacities with diverse economies and dense populations that include migrants, such as Delhi or Mumbai. Countries such as Japan and South Korea, on the other hand, did not implement strict lockdowns and instead relied on infected individuals' contact tracing, intensive testing, and social distancing, which worked well for them. In contrast to other European countries, Germany placed a strong emphasis on social distance. People in Berlin, for example, could only leave their house if they were alone or with another member of the same household. Similarly, Australia limited public gatherings to two people and advised residents to stay at home and go out only for necessary activities. This further demonstrated that, when compared to other countries, Berlin, Germany, and Sydney, Australia experienced the least reduction in NO_2 levels (Fu, Purvis-Roberts, and Williams (2020)). This discovery adds another dimension to the research, concluding that a ban on industrial activities and a reduction in vehicular movements are the most effective methods of combating air pollution.

Overall, the environmental impact of COVID-19 lockdowns and restrictions can be viewed as a blessing in disguise, as described by Lokhandwala and Gautam (2020), who highlighted how nature recovered in various parts of India after 1-2 months of restrictions. Lockdowns and halts in industrial and transportation operations have assisted in identifying the major sources of air pollution. Chang, Meyerhoefer, and Yang (2021) discuss how public transportation demand in Taiwan has decreased since the pandemic began and the government imposed restrictions. When the public is afraid of contracting and becoming infected while using public transportation, such as subways and buses, effective public transportation policies are required. Raising or lowering fares during peak or off-peak hours to smooth demand throughout the day allows public transportation to return to normal levels, further reducing air pollution. Schaefer, Tuitjer, and Levin-Keitel (2021) conducted their research in the city of Hannover, Germany, and discovered that people avoided public transportation due to environmental concerns. These studies produce interesting results that can help governments develop systematic plans to combat not only air pollution, but also water pollution (Gupta et al. (2022)), by controlling industrial activities and reducing private transportation, while also taking economic implications into account.

References

Benchrif, Abdelfettah, Ali Wheida, Mounia Tahri, Ramiz M Shubbar, and Biplab Biswas. 2021. “Air Quality During Three Covid-19 Lockdown Phases: AQI, Pm2. 5 and No2 Assessment in Cities with More Than 1 Million Inhabitants.” *Sustainable Cities and Society* 74: 103170.

Chang, Hung-Hao, Chad D Meyerhoefer, and Feng-An Yang. 2021. “COVID-19 Prevention, Air Pollution and Transportation Patterns in the Absence of a Lockdown.” *Journal of Environmental Management* 298: 113522.

Fu, Franck, Kathleen L Purvis-Roberts, and Branwen Williams. 2020. “Impact of the COVID-19 Pandemic Lockdown on Air Pollution in 20 Major Cities Around the World.” *Atmosphere* 11 (11): 1189.

Gardiner, Beth. 2020. “Pollution Made COVID-19 Worse. Now, Lockdowns Are Clearing the Air.” April.

Gupta, Vedika, Nikita Jain, Deepali Virmani, Senthilkumar Mohan, Ali Ahmadian, and Massimiliano Ferrara. 2022. “Air and Water Health: Industrial Footprints of COVID-19 Imposed Lockdown.” *Arabian Journal of Geosciences* 15 (8): 1–18.

Lokhandwala, Snehal, and Pratibha Gautam. 2020. “Indirect Impact of COVID-19 on Environment: A Brief Study in Indian Context.” *Environmental Research* 188: 109807.

Mahato, Susanta, Swades Pal, and Krishna Gopal Ghosh. 2020. “Effect of Lockdown Amid COVID-19 Pandemic on Air Quality of the Megacity Delhi, India.” *Science of the Total Environment* 730: 139086.

Pinto, Deekshith Nevil. 2020. “How Breathing in Indian Cities Is Equivalent to Smoking Packs of Cigarettes in One Day.” *The Weather Channel*. https://weather.com/en-IN/india/pollution/news/2020-01-29-how-breathing-indian-cities-equivalent-smoking-packs-cigarettes-day.

Schaefer, Kerstin J, Leonie Tuitjer, and Meike Levin-Keitel. 2021. “Transport Disrupted–Substituting Public Transport by Bike or Car Under Covid 19.” *Transportation Research Part A: Policy and Practice* 153: 202–17.

Tripathi, Ayush. 2021. “Air Pollution in Four Indian Cities During the Covid-19 Pandemic.” *International Journal of Environmental Studies* 78 (4): 696–717.

Wang, Meichang, Feng Liu, and Meina Zheng. 2021. “Air Quality Improvement from COVID-19 Lockdown: Evidence from China.” *Air Quality, Atmosphere & Health* 14 (4): 591–604.

WAQIProjectTeam, World Air Quality Index (WAQI). 2022. “Air Quality Historical Data Platform.” https://aqicn.org/data-platform/register/.

Yazdani, Mohsen, Zeynab Baboli, Heidar Maleki, Yaser Tahmasebi Birgani, Maryam Zahiri, Seyede Saba Heydari Chaharmahal, Mahdis Goudarzi, et al. 2021. “Contrasting Iran’s Air Quality Improvement During COVID-19 with Other Global Cities.” *Journal of Environmental Health Science and Engineering* 19 (2): 1801–6.