Assessing the Impact of COVID-19 Lockdowns on the Air Quality
SEEM 2460 Introduction to Data Science - Final Project Report

1. Abstract

This project aims to show the effectiveness of data visualization and analysis, when targeting real-life issues. Through this paper, we have presented a detailed assessment and a visualization-based analysis of the impact of the COVID-19 caused lockdowns on the world's air quality.

2. Introduction

2.1 Problem Statement and Motivation

Pollution, specifically air pollution is known as a major cause of death and disease globally. According to the World Health Organization (WHO), air pollution accounts for almost four million premature deaths worldwide. Particulate matter like dust, dirt or smoke have the potential to act as carriers for various contagions and thus, they can lead to a rapid spread of diseases over large areas. Further, the global population has been facing harsh lockdowns which aim at halting the spread of the COVID-19 virus. Discussions surrounding how the emissions from fossil fuel

combustions and other pollutant sources have dropped radically in numerous countries, have started surfacing. This has led many to think that there may be a silver lining amongst the dark clouds of the pandemic. This silver lining is a step towards cleaner air and a better environment but it may be a reminder for the world that we need to kick our habit of burning resources.
These factors create a need for us to understand the ongoing situation. Therefore, it is natural to think about the relationship between the lockdowns and air pollution. This might give us some insight as to where we went wrong and how we could contribute to improving the climate. The crux of the paper is to find a correlation between the impact of the COVID-19 virus caused lockdowns and air quality of various major cities.

2.2 Proposed Analysis

In this paper, we have gathered qualitative ordinal data from various sources and have used Tableau and other Python libraries for our data visualization and analytic tasks. By visualizing the daily coronavirus cases over the daily concentration of pollutant particles like PM2.5, PM10, NO_2 , etc. and separating the data into various timelines according to city specific lockdowns, we aim to see trends, patterns and correlation between the pollutant's concentration in the air and the virus cases. This process is done for 4 major cities around the world, namely - London, Madrid, New Delhi and Wuhan. We have carefully selected these cities based on their difference in the severity of lockdowns, varied air quality indexes at different stages of the pandemic. Through this analysis, we gain insights and assess the impact of the lockdowns on a city's air quality.

3. Literature Review

3.1 Air Quality Index and Pollutants

AQI (Air quality Index) is an index that is used for reporting the air quality, pollution levels to masses on a regular basis. Forecasting of air pollution levels is also done under this index. Air quality index is considered important because air quality has a direct impact on the health of the citizens of a country. Therefore, through this index people are able to better understand the affects air quality will be having on their health. AQI is usually calculated for five major pollutants: Sulphur dioxide (SO_2), Nitrogen dioxide (NO_2), Carbon Monoxide (CO), particulate matter (PM2.5, PM10) and Ozone (O_3).
Particulate Matter is formed due to the mixture of aerosols and minute solid particles in the atmosphere.

Table 1. Air Pollutants

Name of Pollutant	Origin/Type/Nature	Size	Sources

PM10 (One of the major pollutants of the atmosphere)	Respirable particulate matter	Upto the size of 10 microns	● Dust Blown off roads and open lands ● Dust from Construction Sites ● Wildfires ● Waste Burning ● Brush Burning ● Combustion of Oil ● Other Industrial Activities
PM2.5	Fine particulate matter	Upto the size of 2.5 microns	● Natural and Man-made (Man-made is considerably higher when compared to natural sources) ● Vehicular Exhausts ● Combustion of Oil, Wood, Gasoline and other Fossil Fuels ● Other Industrial Activities
Carbon Monoxide (CO) (One of the major pollutants of the atmosphere)	● Produced due to incomplete combustion of carbon containing fuels. ● The concentration of CO in the atmosphere however does not change very quickly. This is because it has a long tropospheric lifetime which is affected by environmental factors as well as the topographic factors of a region.	Gas	● Incomplete Combustion in ○ Vehicles ○ Power Generation ○ Households
Nitrogen Dioxide (NO_2)	● Air pollutant ● Levels of NO2 in the atmosphere will be high on regular days in the developed cities around the world. ● NO2 also contributes to the formation of particulate matter and ozone by reacting with existing chemicals in the atmosphere.	Gas	● Vehicles ● Fossil Fuel Combustion ● Metal Refining Processes
Ozone (O_3)	● Air pollutant ● Ground-level ozone is created due to chemical reactions between nitrogen oxides and various organic compounds released because of industrial activity, with sunlight acting as a catalyst.	Gas	● Ozone is sourced from the reactions of Nitrogen oxides

	● Therefore, O3 levels are expected to be high in cities on sunny/summer days and relatively lower during the colder days.		
Sulphur Dioxide (SO_2)	● Gaseous Air pollutant ● Ground-level ozone is created due to chemical reactions between nitrogen oxides and various organic compounds released because of industrial activity, with sunlight acting as a catalyst. ● Therefore, O3 levels are expected to be high in cities on sunny/summer days and relatively lower during the colder days.	Gas	● Processing of materials containing Sulphur ○ Gases ○ Metal Ores ○ Coal ○ Oil ○ And others

3.2 Competitive Study

a) A similar study [6] by an individual named Primož Godec analyzed the pollution data in the city of Ljubljana, Slovenia as well as in Bergamo, Italy. The study focuses on the levels of just 2 pollutants, Nitrogen dioxide (NO_2) and PM10. It was found in this study that the levels of NO_2 in Ljubljana were considerably lower in the months of the lockdown compared to the same months in previous years. This decrease was attributed to considerable decrease in road traffic. On the other hand PM10 levels were higher in the months of lockdown because of an external factor of sand dust from Sahara due to strong winds, along with more direct factors like continuous operation of power generation facilities and certain industries.
What makes our study different from Godec's study is that we focus on SO_2, PM2.5, CO and O3 in addition to PM10 and NO_2. Furthermore, the focus of our study is on four different major cities all around the world, and the impact the lockdowns in those cities had on the air quality of that region.

b) In April, 2020, NASA reported [7] that the aerosol values in the northern regions of the Indian Subcontinent were at an all time low just a week after the implementation of a strict nationwide lockdown in India. The report compared the values of aerosol optical depth over Northern India with the values in the years 2016-2019.
Our study also focuses on the Indian city of New Delhi, however instead of focusing on just aerosol levels (particulate matter), we study the levels of other major air pollutants as well.

4. Data

For the purposes of this project, we required 2 types of data- 1) Daily COVID-19 Cases for the 4 cities 2) Daily Air Quality Indexes (AQIs) of the pollutants in the 4 cities. Along with this we also gathered the information about the lockdown phases in the 4 cities as COVID-19 cases do not have a direct effect on the AQIs, rather it acts as an indirect causal factor as it is responsible for lockdowns which in turn relate to AQIs. The information for the lockdowns was collected from various news sources.

4.1 Sources

In this section, we describe the sources of data for the 4 cities. The AQIs follow the US EPA standard.

Table 2. COVID-19 and AQI Data Sources

Name	**COVID-19 Data**	**AQI Data**
Wuhan, Mainland China	https://github.com/CSSEGISandData/COVID-19	https://aqicn.org/historical (Wuhan)
New Delhi, India	https://github.com/imdevskp/covid-19-india-data/blob/master/complete.csv	https://aqicn.org/historical (Satyawati College, New Delhi)
London, United Kingdom	https://coronavirus.data.gov.uk/	https://aqicn.org/historical (London)
Madrid, Spain	https://cnecovid.isciii.es/covid19/#documentaci%C3%B3n-y-datos	https://aqicn.org/historical (Madrid)

4.2 Data Pre-Processing

For data pre-processing, we use Tableau Prep Builder. We first clean the data by standardizing the date formats. There were no missing values for the COVID-19 data. The missing values for AQI are dealt with using 2 strategies 1) Finding another nearby AQI provider for the missing values, 2) In case there is no other provider, the null values are filled with an average value of the corresponding month. Once the data is cleaned and standardized, we perform an inner-join (using dates) between the COVID-19 and the AQI dataset for each city. The data is then exported as a Tableau Extract and a CSV for further experimentation. The data is now publicly available on our Github.

4.2.1 Structure of processed data

The data has the following structure:

Date	Cumulative Cases	Daily Cases	PM2.5 AQI	PM10 AQI	O3 AQI	NO2 AQI	SO2 AQI	CO AQI

5. Implementation

The visualizations were made using Tableau Desktop. We plot the bar graphs in dual axis to show both the COVID-19 cases and the AQIs of different pollutants with respect to time. We also plot the trend lines with varying degrees of polynomial. Both the axes have been set to continuous instead of discrete to enable the addition of reference bands (representing the lockdown phases) and trend lines. To calculate the correlations we use Pearson's method using Python and Pandas. For visualization of the COVID-19 cities world map, we used Microsoft's Power BI.

6. Data Visualization and Insights

All the graphs/visualizations are presented in the appendix to conserve space. We will reference them when required.

6.1 Wuhan

6.1.1 Lockdown Study

Table 3. Wuhan Lockdown Study

Phase and Duration	**COVID-19 Cases or Spread Rate**	**Government Implementations/ Suspensions**
Phase 1 (23rd January 2020 - 14th March 2020)	> 500 confirmed cases and over 15 deaths reported by the end of January	● Suspension of all public transport, including buses, railways and flights ● Strict ban prohibiting residents of Wuhan from leaving the city ● Gradually, all non-essential companies, schools, cinemas, public gathering areas were ordered to shut down.
Phase 2 (14th March 2020 - 22nd March 2020)	Wuhan remained a high risk area, it consisted of a mixture of low-risk township level divisions and medium risk areas.	● Announced measures to relax the restrictions and lessen controls. Gradually allowed industries and other services to resume work.

		• All inter-province anti-COVID-19 traffic checkpoints and authority assigned areas were removed giving the public more freedom.
Phase 3 (22nd March 2020 - 8th April 2020)	Cases reaching a steady growth rate and not alarmingly exponential	• Wuhan finally loosens its 2 month lockdown. • Resumed all of its transportation by early April. • Implemented the 'Green Code' requirements and residents must completely follow them if they intend to move around or outside the city.

6.1.2 Insights from visualization

We can observe the graphs for Wuhan in Section 8.1 for the AQI of pollutants and COVID-19 cases with respect to time along with the lockdown phases. We observe a decrease in the Carbon Monoxide emissions with the start of lockdown. Furthermore, we notice a dip in NO_2, PM10 and PM2.5 particles. As the lockdown was relaxed in later phases as shown in Table 3, the pollutants started showing an upward trend again. We notice an increase in O_3, it can be attributed to the decrease in NO_2 as stated with reasoning in Table 1.

6.2 New Delhi

6.2.1 Lockdown Study

Table 4. New Delhi Lockdown Study

Phase and Duration	**COVID-19 Cases or Spread Rate**	**Government Implementations/ Suspensions**
Phase 1 (25th March 2020 - 14th April 2020)	> 500 confirmed cases and a death toll of	• Nationwide lockdown. • Nearly all services, industries and factories suspended. Movement restricted from 7pm to 7am.

	approximately 30 people	● Major public transport services like railway, domestic and international air travel were banned.
Phase 2 (14th April 2020 - 3rd May 2020)	Growth rate of the cases had significantly slowed. State governments expressed their opinions regarding the effectiveness of the lockdown and its possible extensions.	● Relaxations were given starting from 20th April where the virus spread was contained by then. ● Classification system introduced. Lockdown areas represented by red, orange and green zones. ● Agricultural businesses, public works programmes, cargo transportation businesses, banks and government centres were allowed to open.
Phase 3 (3rd May 2020 - 17th May 2020)	Cases on the verge of reaching 100,000. Not increasing at an exponential rate.	● Relaxations given to the areas depending on the containment. ● Shops/e-commerce dealing essential goods, personal vehicles, private offices with 33% capacity were allowed for all three zones. ● Non essential movement, major public transport services, educational institutions, cinemas, halls and public gathering areas remained closed.

6.2.2 Insights from visualization

We can observe the graphs for New Delhi in Section 8.2 for the AQI of pollutants and COVID-19 cases with respect to time along with the lockdown phases. We observe a sharp decrease in all the pollutants at the beginning of the first phase of lockdown shown in Table 4. The first phase was the strictest so, it drastically reduced the mobility of vehicles and the opening of industries. The overall AQI was also substantially lower than the previous years. As seen in Wuhan, the AQI started following an upward trend as the lockdown was relaxed and the mobility increased.

6.3 London

6.3.1 Lockdown Study

Table 5. London Lockdown Study

Phase and Duration	**COVID-19 Cases or Spread Rate**	**Government Implementations/ Suspensions**
Phase 1 (24th March 2020 - 30th April 2020)	> 10000 tested cases by end of February. > 2000 positive cases and a death toll of > 250 people	● London began implementing closures and cancellations ● Pubs, restaurants, gyms and other venues prone to public gatherings were closed ● Wide restrictions on the freedom of movement were implemented and people were enforced by law to follow the ‘stay at home’ rule.
Phase 2 (30th April 2020 - 10th May 2020)	> 26000 people died. UK said to have passed the ‘peak of its first coronavirus outbreak’	● ‘Stay at home’ rule was to be exercised as much as possible. ● Work restrictions were eased. Ability to meet another person from another household was also allowed. ● Schools, non-essential retail and other crucial restrictive measures remained closed.
Phase 3 (10th May 2020 - Present)	Cases are rising exponentially	● Government introduced a COVID-19 recovery strategy for the people of the UK and a new 5-level alert system was introduced. ● UK is currently at level 4 - epidemic in general circulation where transmission is high or rising exponentially ● Slowly moving towards level 3 - epidemic in general circulation.

6.3.2 Insights from visualization

We can observe the graphs for London in Section 8.3 for the AQI of pollutants and COVID-19 cases with respect to time along with the lockdown phases. London’s lockdown was not as strict as the ones observed in New Delhi and Wuhan. As a result, we don’t notice a substantial dip in the pollutants. Although, the mobility was reduced but not to the extent noticed in other cities.

Moreover, the late action of the government exacerbated the situation as the lockdown was implemented ~1 week before the peak.

6.4 Madrid

6.4.1 Lockdown Study

Table 6. Madrid Lockdown Study

Phase and Duration	**COVID-19 Cases or Spread Rate**	**Government Implementations/ Suspensions**
Phase 1 (14th March 2020 - 26th April 2020)	Cases rising significantly	● All classes, public gathering areas like cinemas, halls, etc., and other retail activity was suspended ● People were only allowed to leave their homes for essential purposes or for commuting to work. ● Suspension of all non-essential industries and its workers, employees
Phase 2 (26th April 2020 - 10th May 2020)	Cases transmission was high but increasing at a steady rate	● Spain introduced several relaxations ● Adults and children were allowed out for daily walk and exercise, whilst making the rule of wearing face masks mandatory. Public transportation was allowed ● Restaurants and cafes were allowed for people to only collect food in the form of a takeaway delivery service. Some retail services were allowed to reopen.
Phase 3 (10th May 2020 - Present)	Cases are rising exponentially	● Allowed upto 10 people to meet and hold social gatherings. ● Cafes, restaurants and small businesses were now allowed to operate at half capacity whereas other outdoor cultural, religious events were allowed

		• Slowly moving towards level 3 - epidemic in general circulation.

6.4.2 Insights from visualization

We can observe the graphs for Madrid in Section 8.4 for the AQI of pollutants and COVID-19 cases with respect to time along with the lockdown phases. Madrid's lockdown was stricter than that of London but laxer than that of Wuhan and New Delhi. As a result, we see a dip in the pollutants AQI. With the relaxations in Phase 2 and 3 shown in Table 5, the AQIs started showing an upward trend.

6.5 Correlation and Insights

Correlation was found between the different air pollutants and the number of cases in each city.

City Name	**PM2.5**	**PM10**	**O3**	**NO2**	**SO2**	**CO**
Wuhan	-0.205	-0.208	-0.205	-0.135	-0.012	0.083
New Delhi	-0.097	0.191	0.066	-0.125	-0.188	0.061
London	0.239	0.321	0.324	0.409	0.324	0.445
Madrid	-0.135	-0.081	-0.047	-0.041	-0.344	-

7. Conclusion

In our study we have successfully found correlations through the data between the rising daily COVID-19 cases, and the decrease in Air Pollution levels due to the fall in harmful air pollutants that determine the air quality. Our study has given us a visual representation of the positive effect COVID-19 caused lockdowns have had on the air quality around the world. The positive reports issued by the governments and the public have been quantified, and can be clearly seen through the visualisations.

Specifically, we can see that the improvement in air quality in all the cities have been the most distinct and significant during the harshest phases of the lockdowns. Furthermore, it is also clear that wherever the lockdown was not strictly implemented or wherever the lockdown was not followed, the pollution levels did not fall as expected. This further strengthens our hypothesis of an inverse relation between the number of cases and air quality index. This study can be seen as an excellent indicator and can act as an evidence for the next big looming crisis: Climate Change. Whilst learning from experience from the ongoing crisis, humans need to realize the necessity of changing their habits in order to prevent further mishaps in the future.

Application of data science: Numerous data science techniques have been used in this study. These include: Data Visualizations (dual axis bar graphs, heat maps), Statistics (correlations between AQI and number of cases), large data processing.

Data Processes Used: Data collection, Data analysis, Data investigation and Data application. Furthermore, The investigation and visualization of this data provides us with a natural experiment which might help us later to formulate policies to reduce pollution.

8. Appendix

8.1 Wuhan

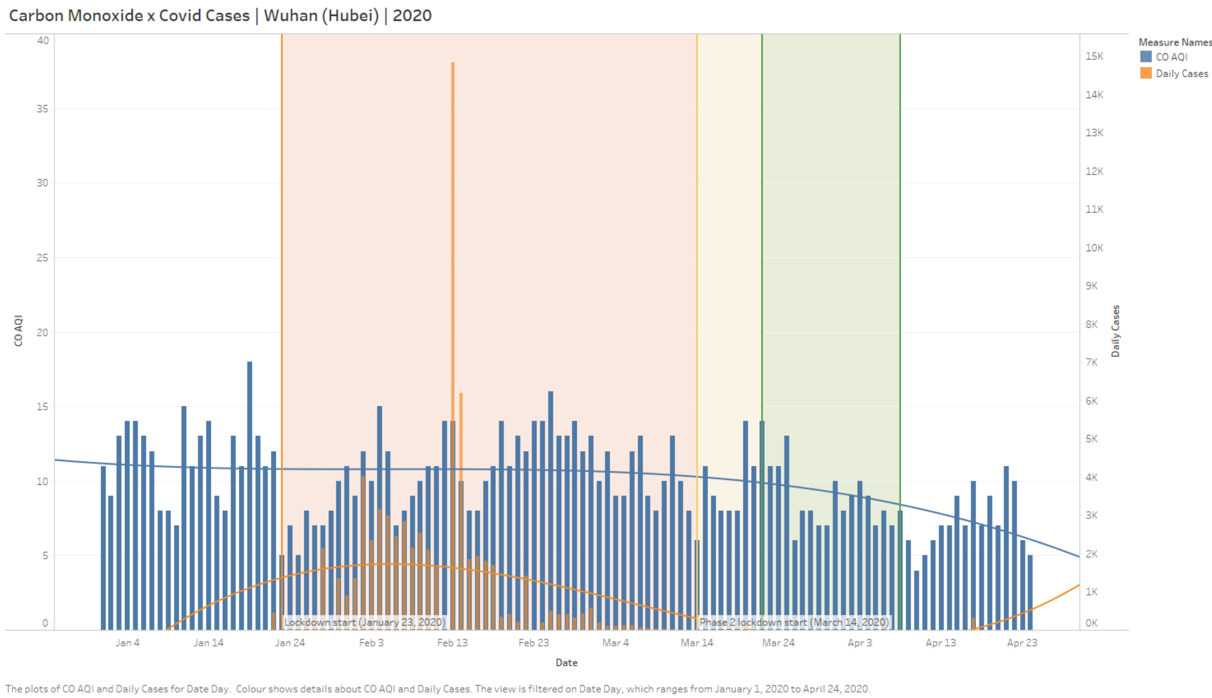


Figure 1.

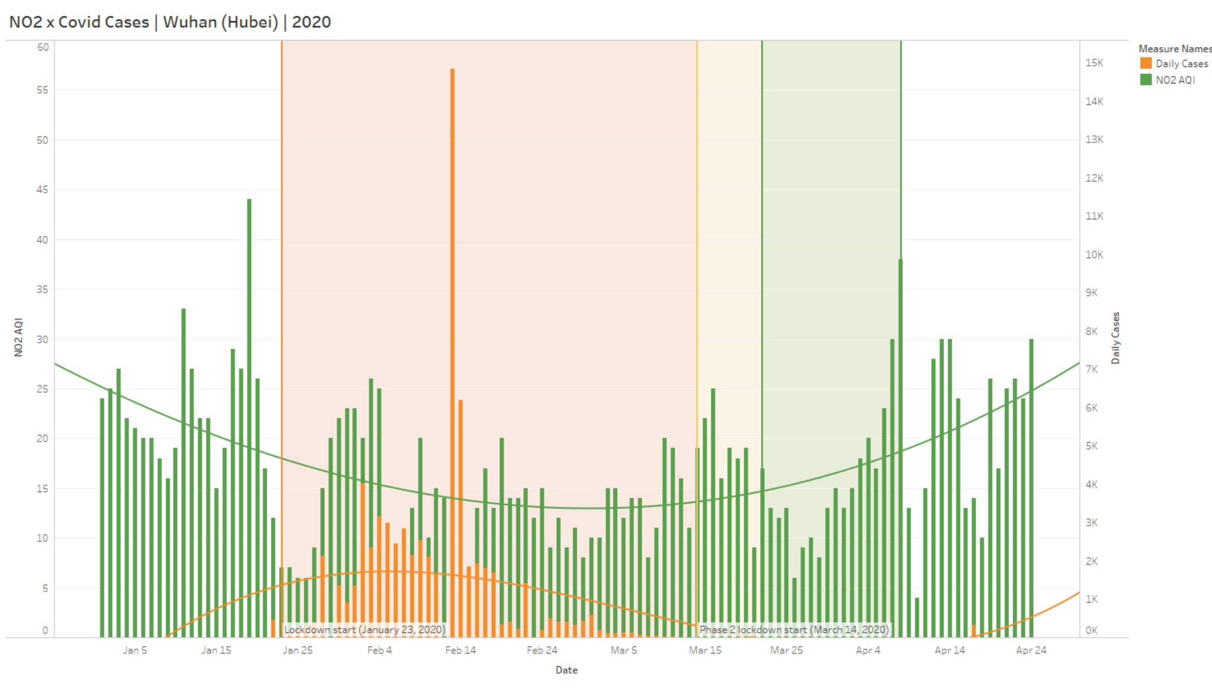


Figure 2.

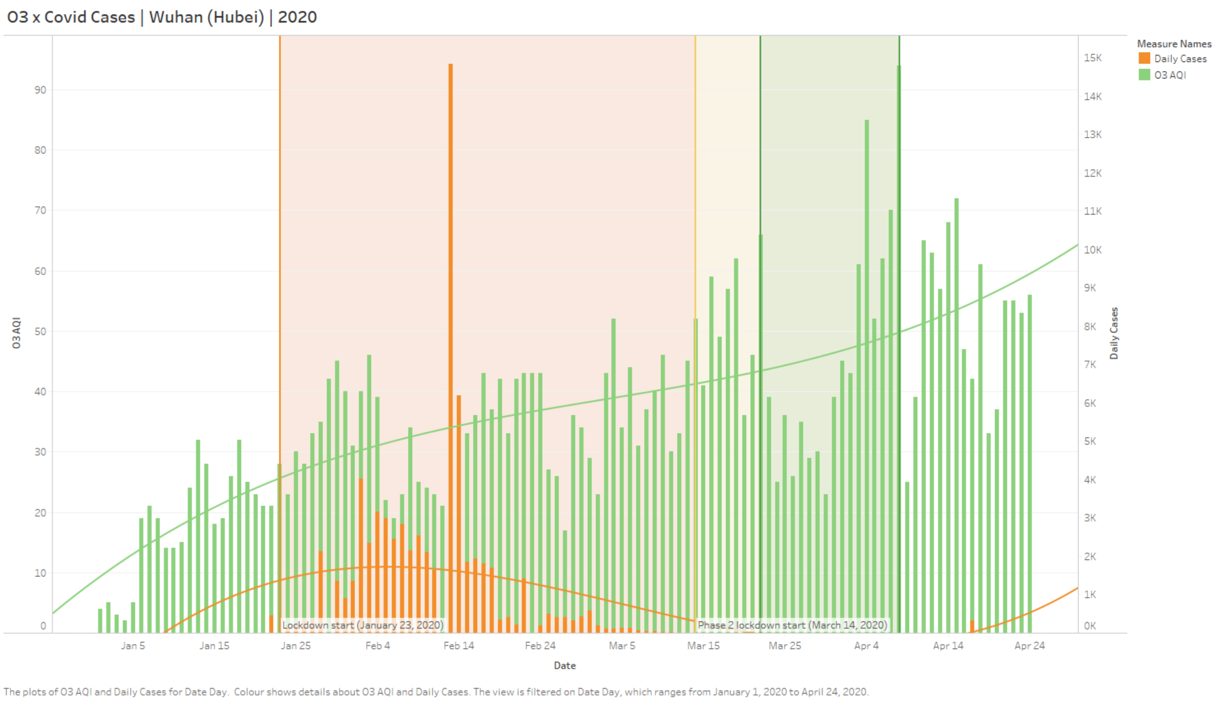


Figure 3.

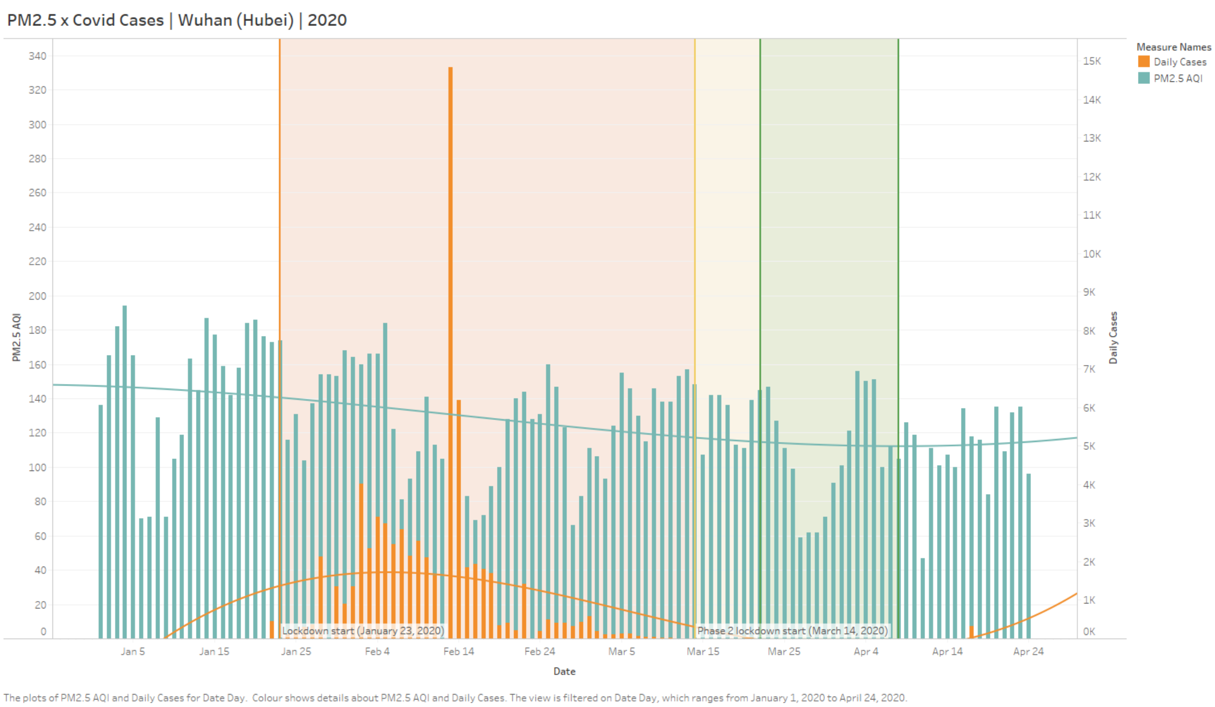


Figure 4.

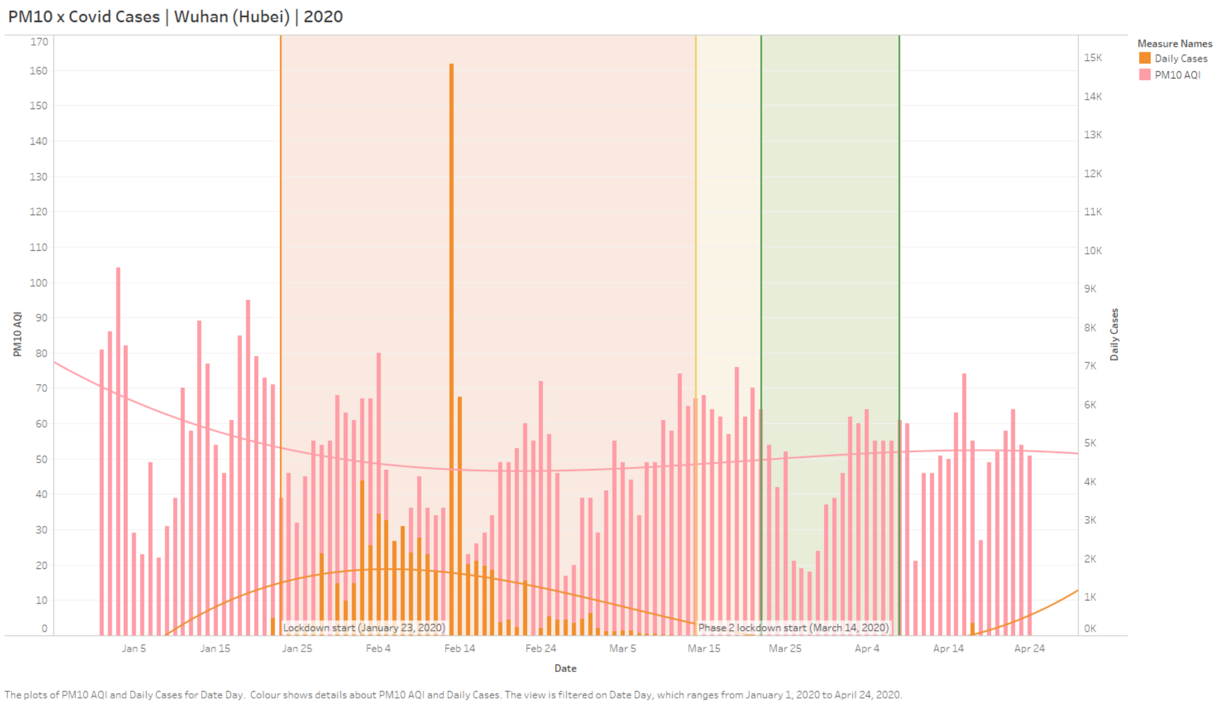


Figure 5.

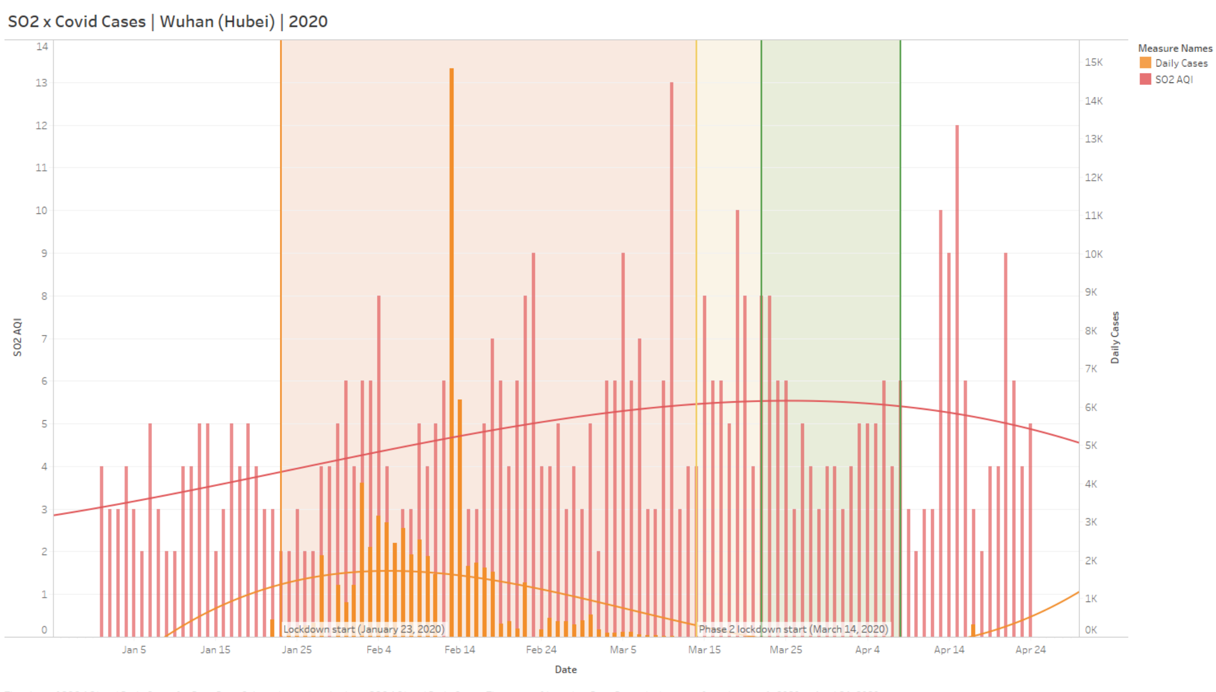


Figure 6.

8.2 New Delhi

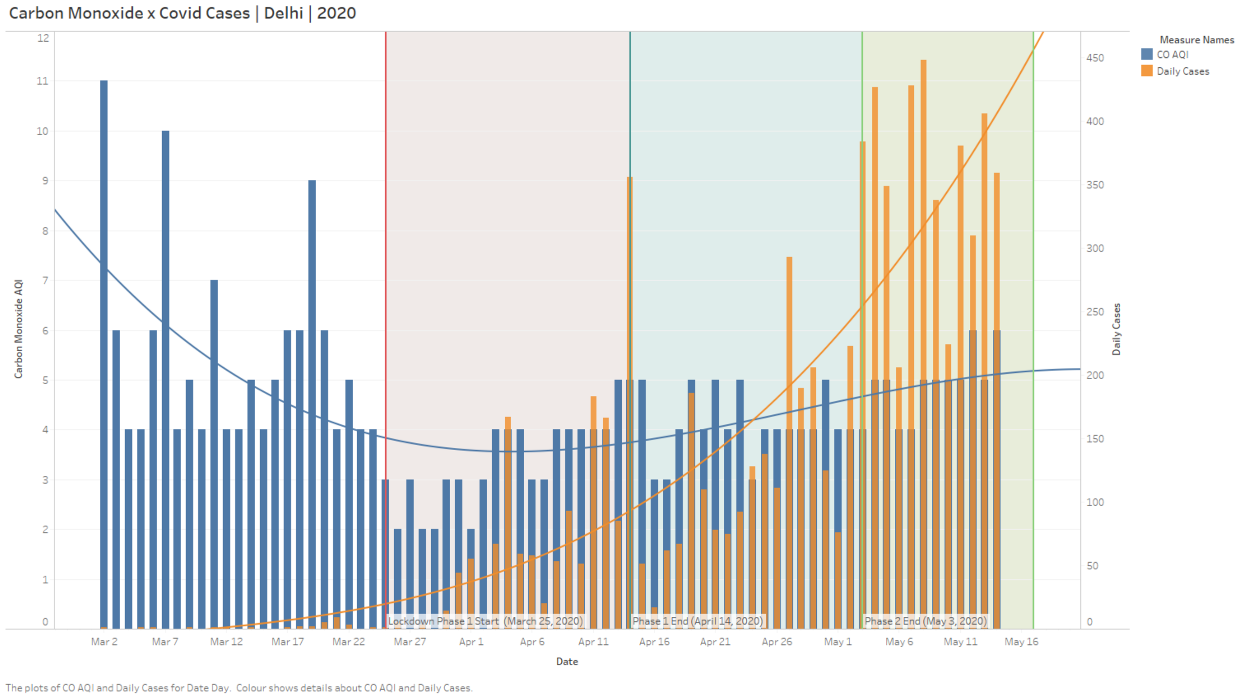


Figure 7.

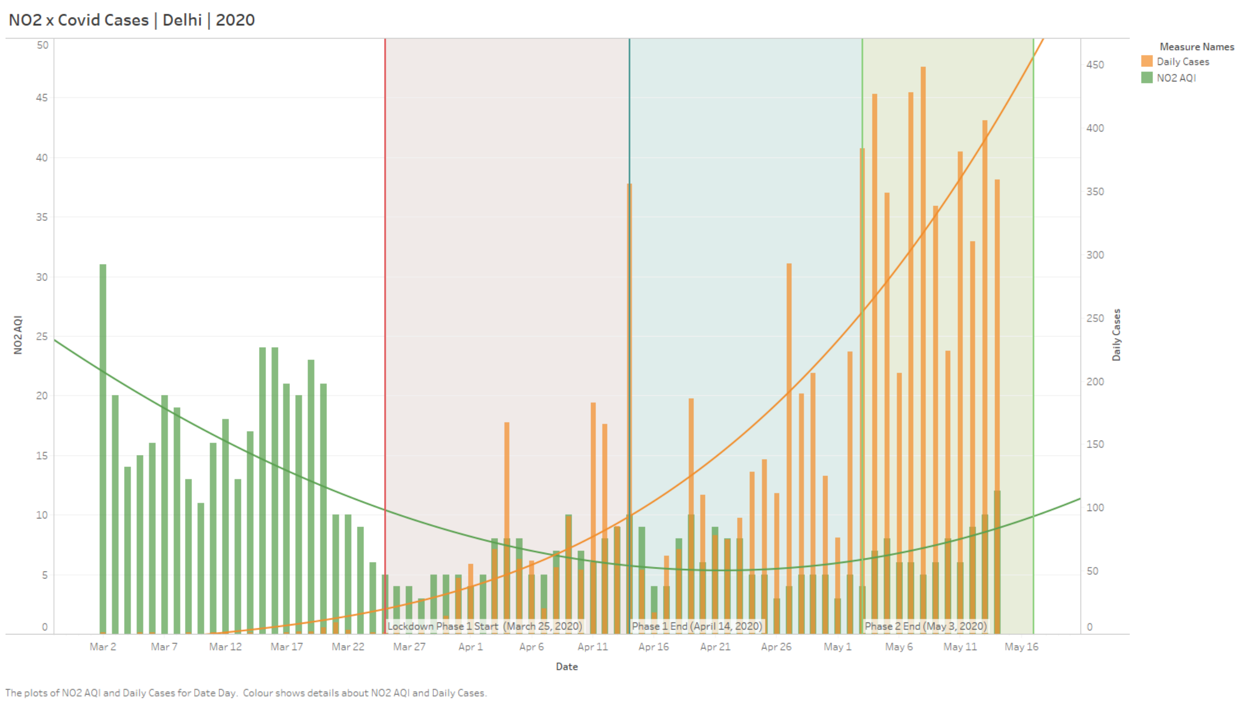


Figure 8.

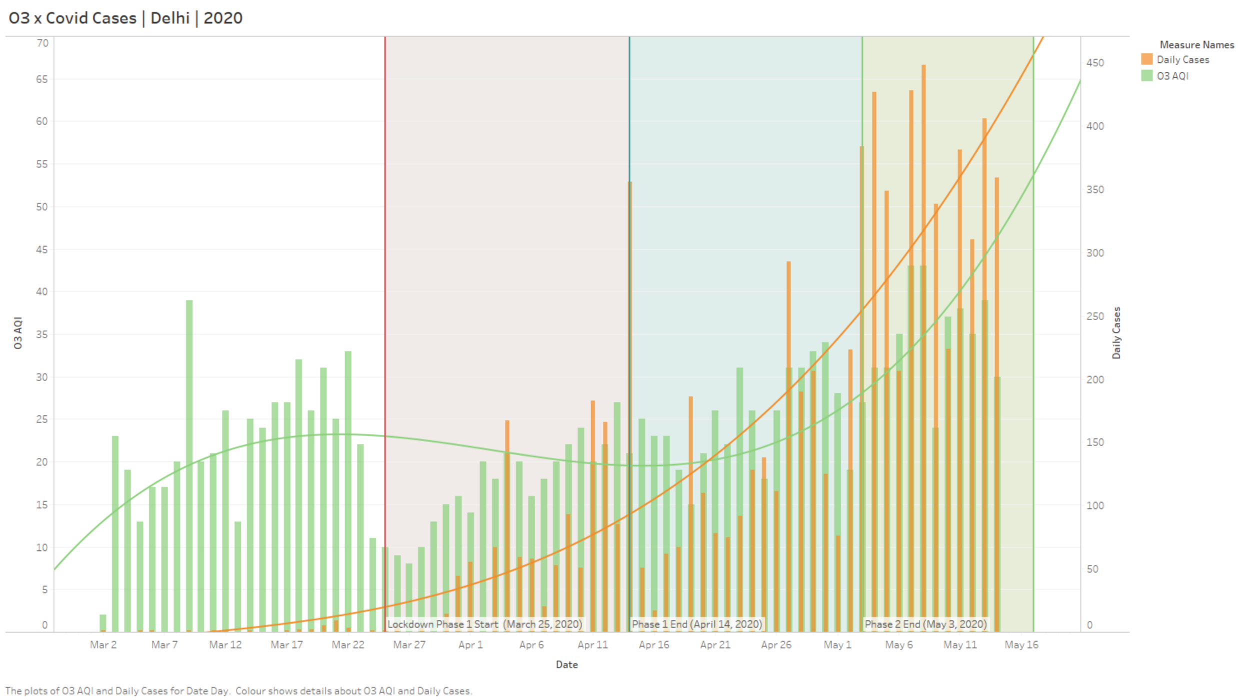


Figure 9.

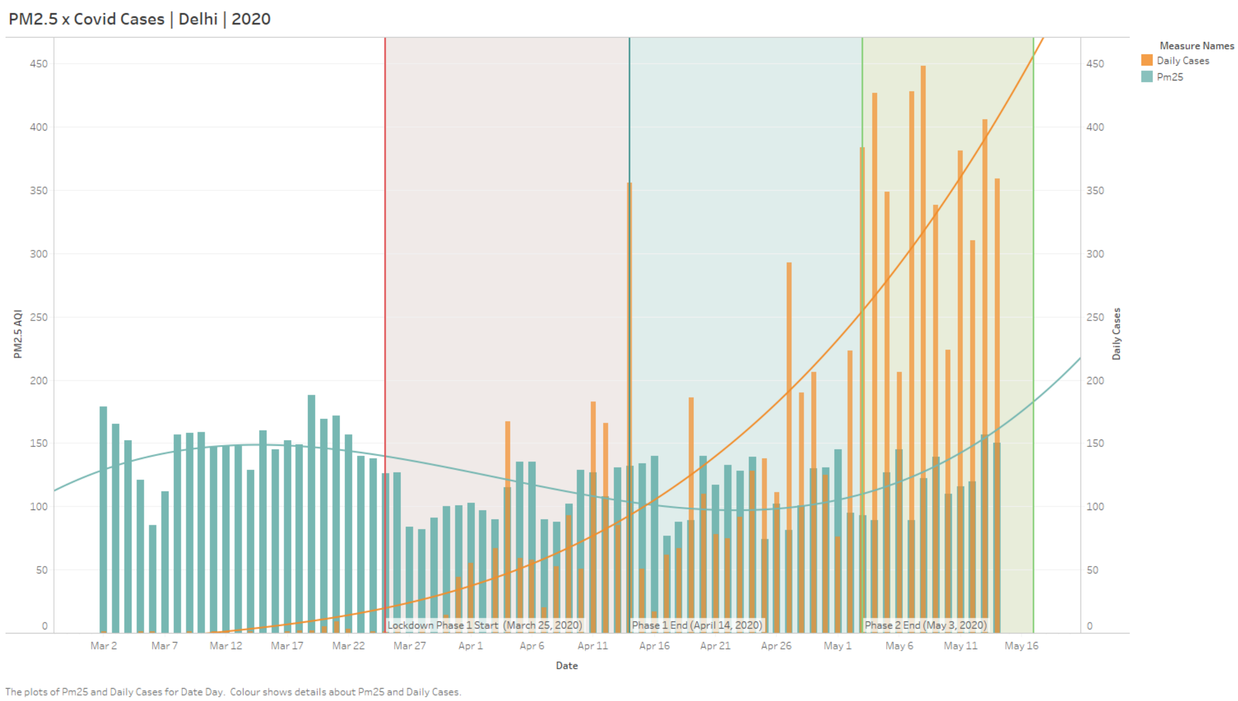
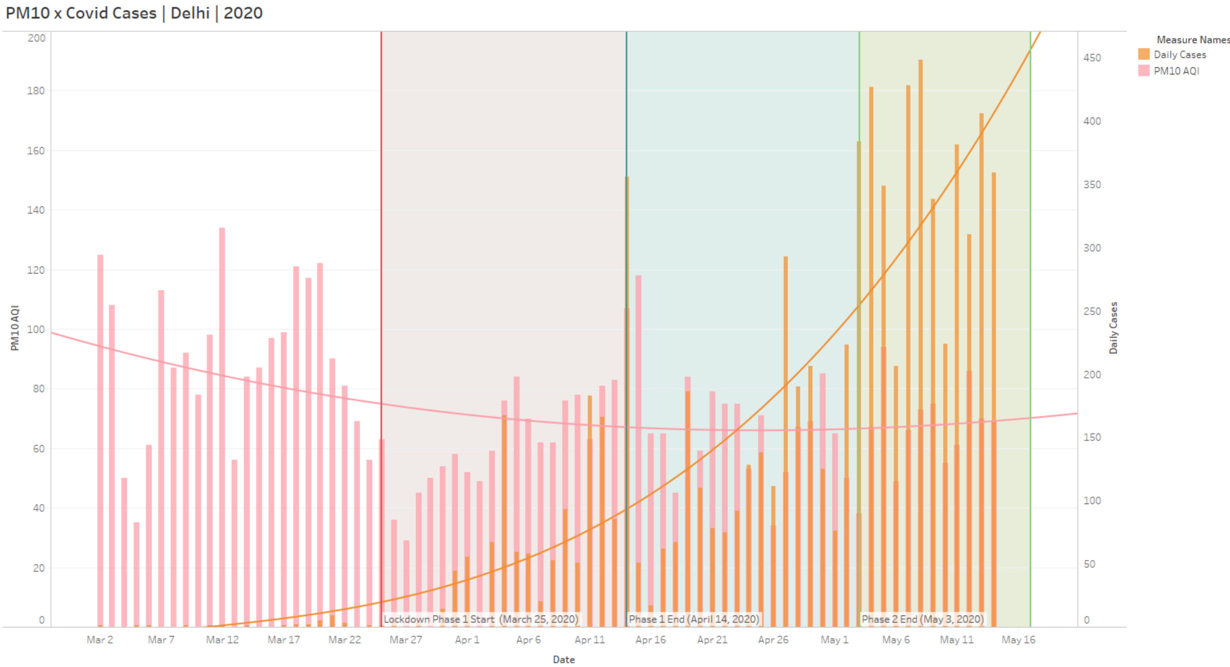
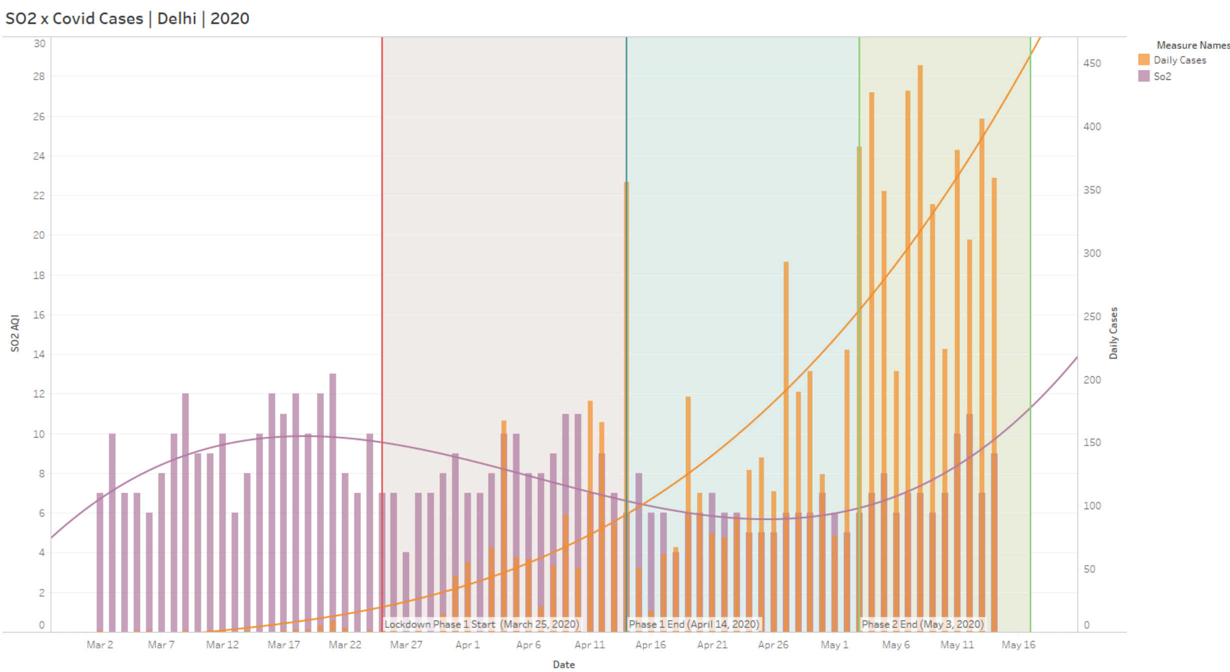


Figure 10.



The plots of PM10 AQI and Daily Cases for Date Day. Colour shows details about PM10 AQI and Daily Cases.

Figure 11.



The plots of So2 and Daily Cases for Date Day. Colour shows details about So2 and Daily Cases.

Figure 12.

8.3 London

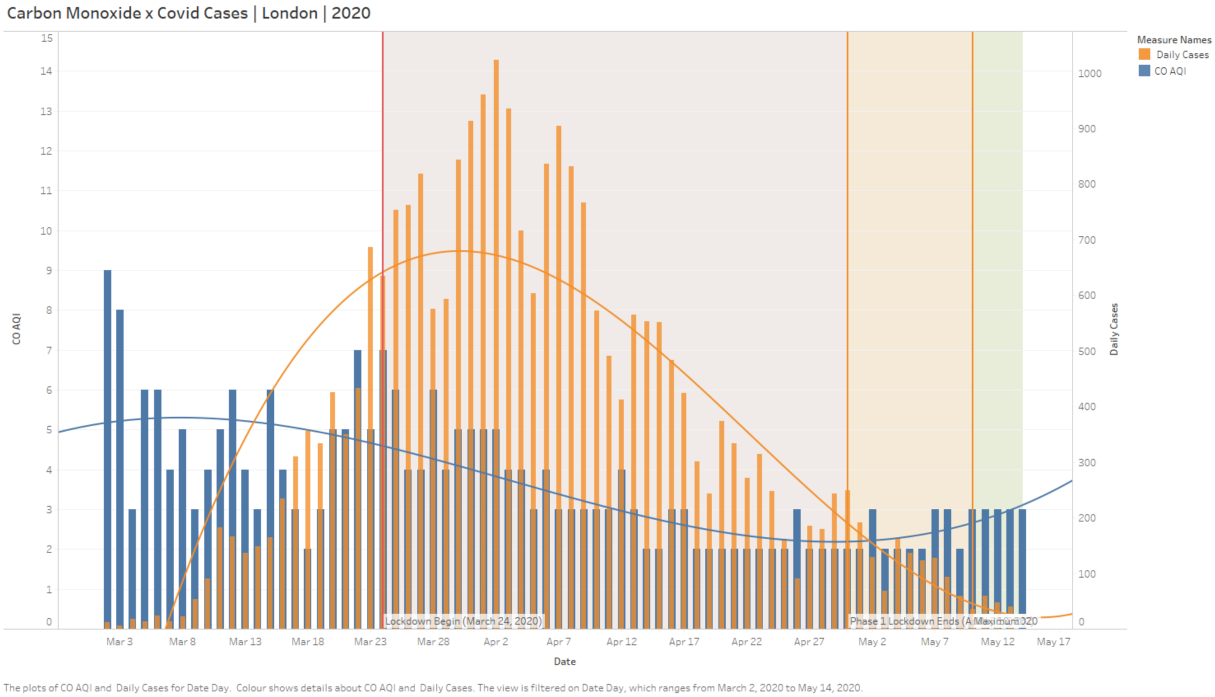


Figure 13.

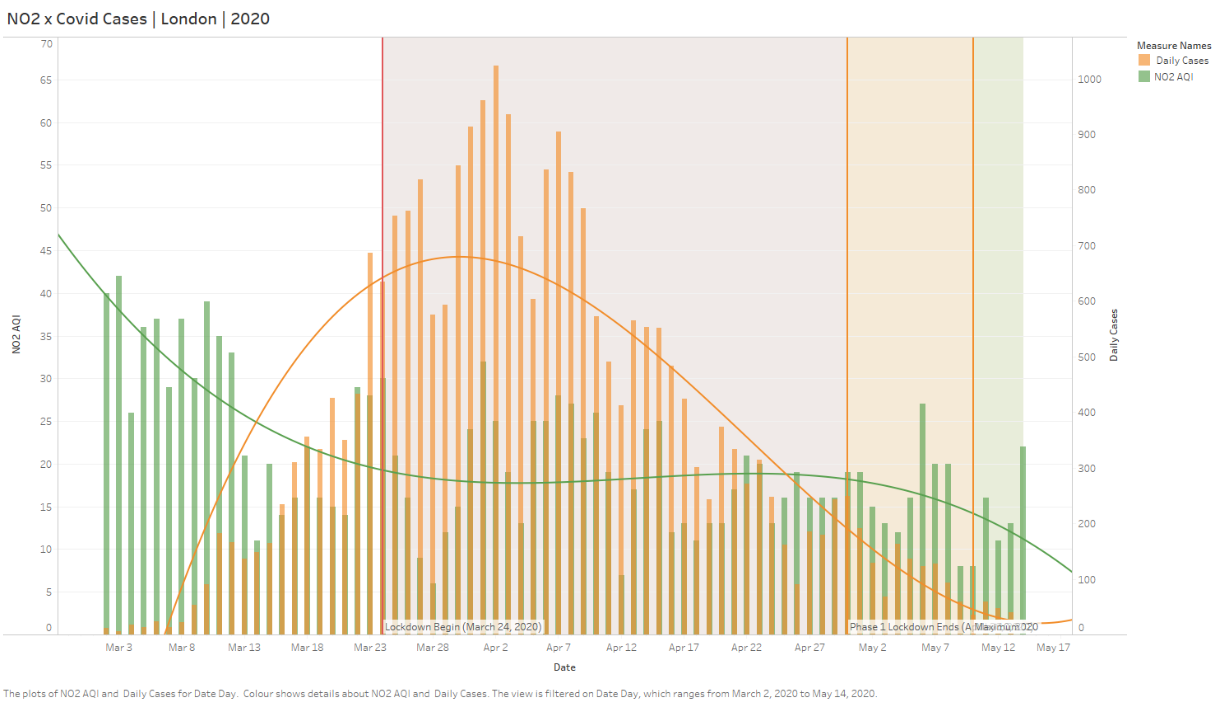


Figure 14.

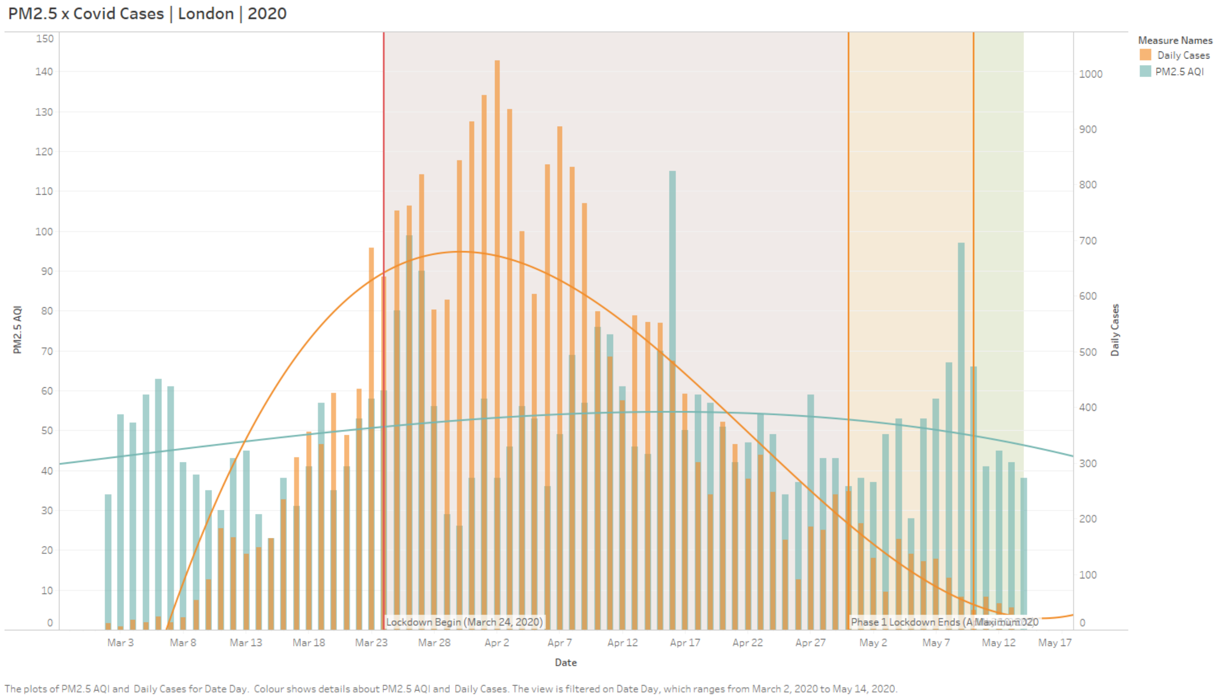


Figure 15.

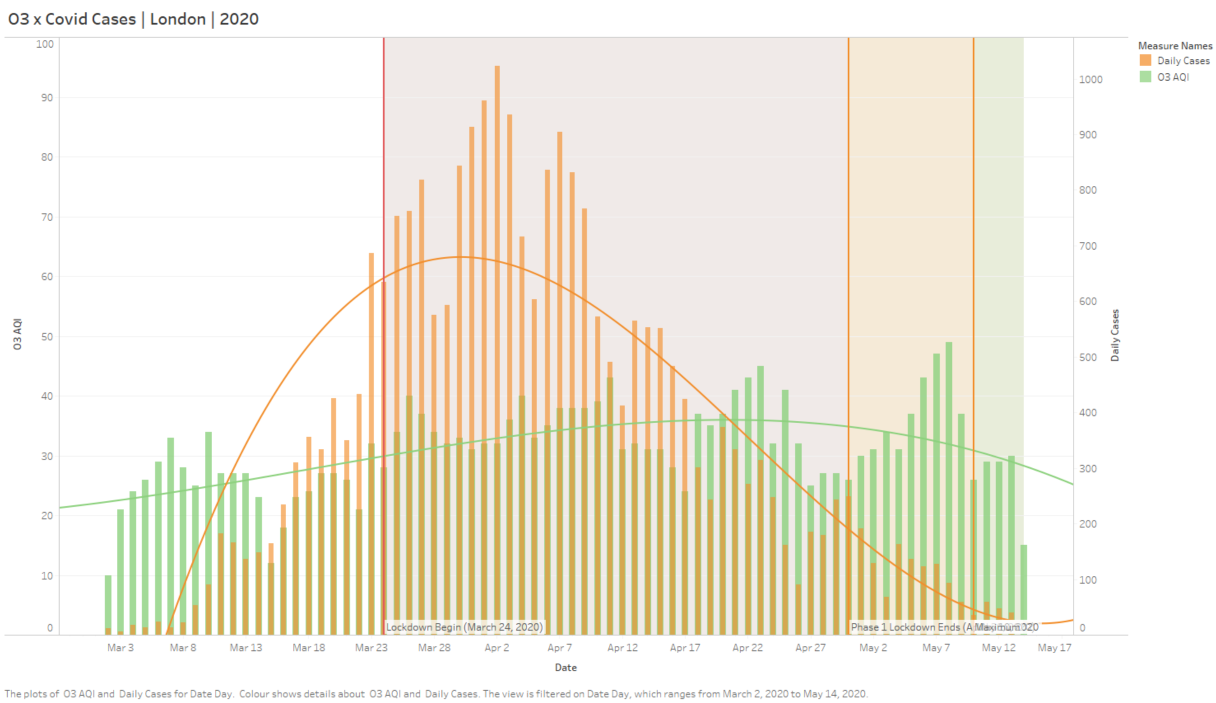


Figure 16.

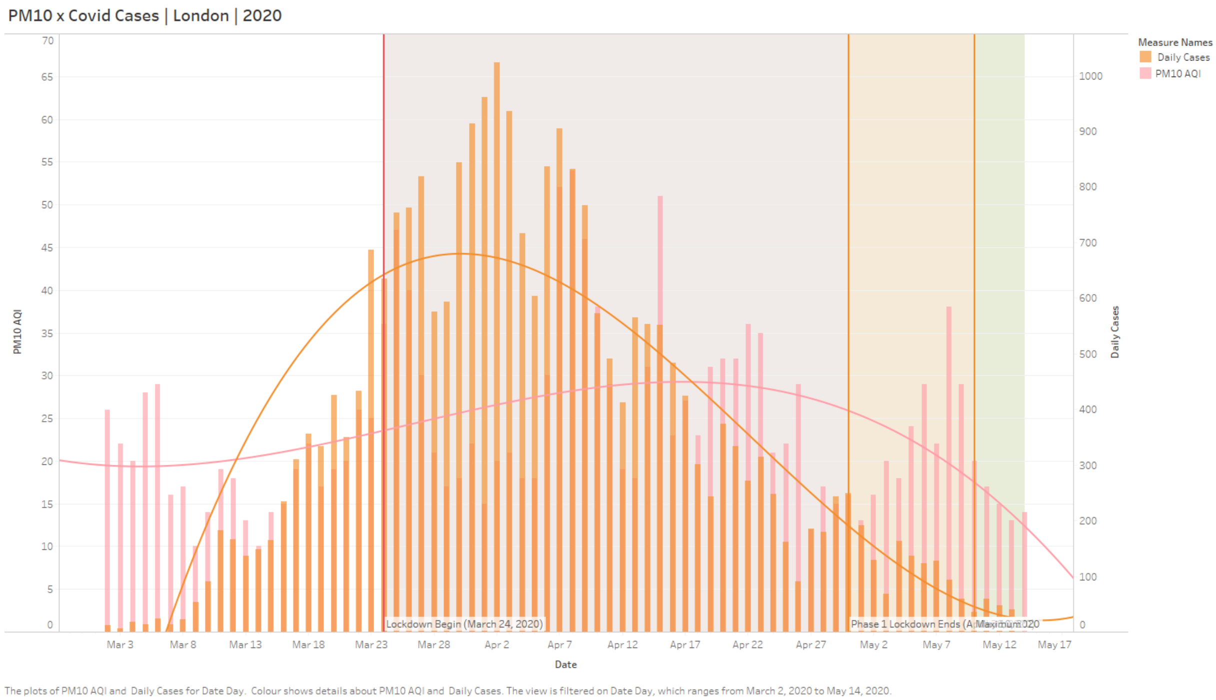


Figure 17.

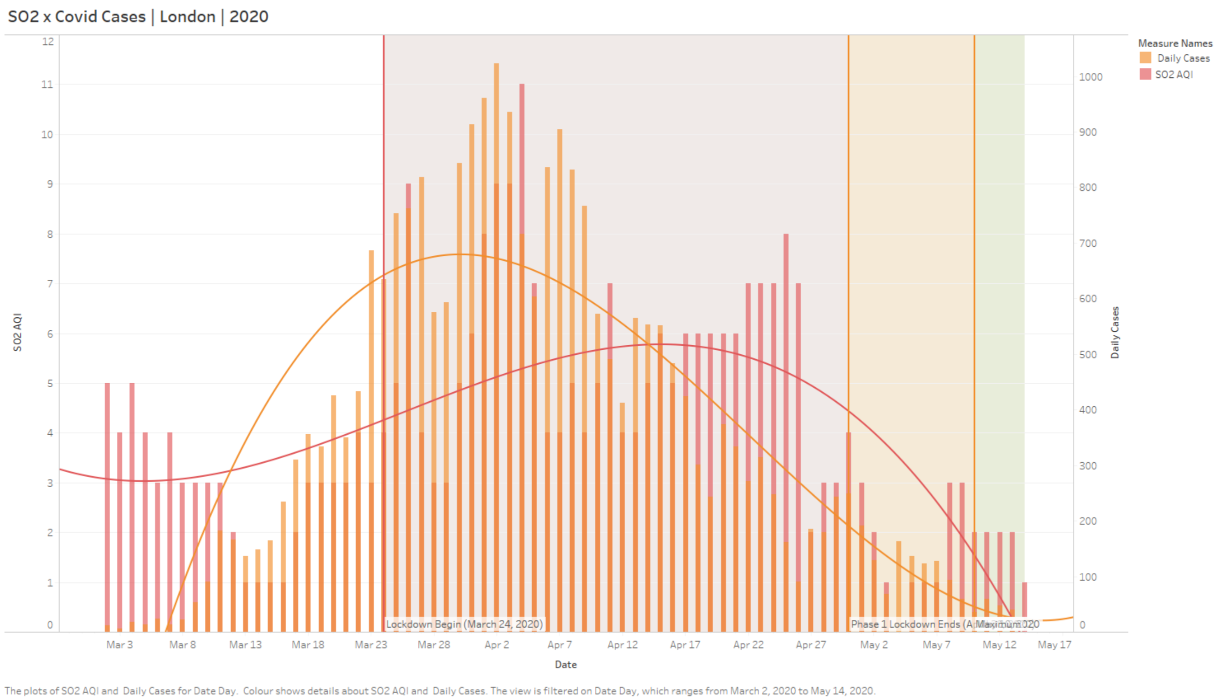


Figure 18.

8.4 Madrid

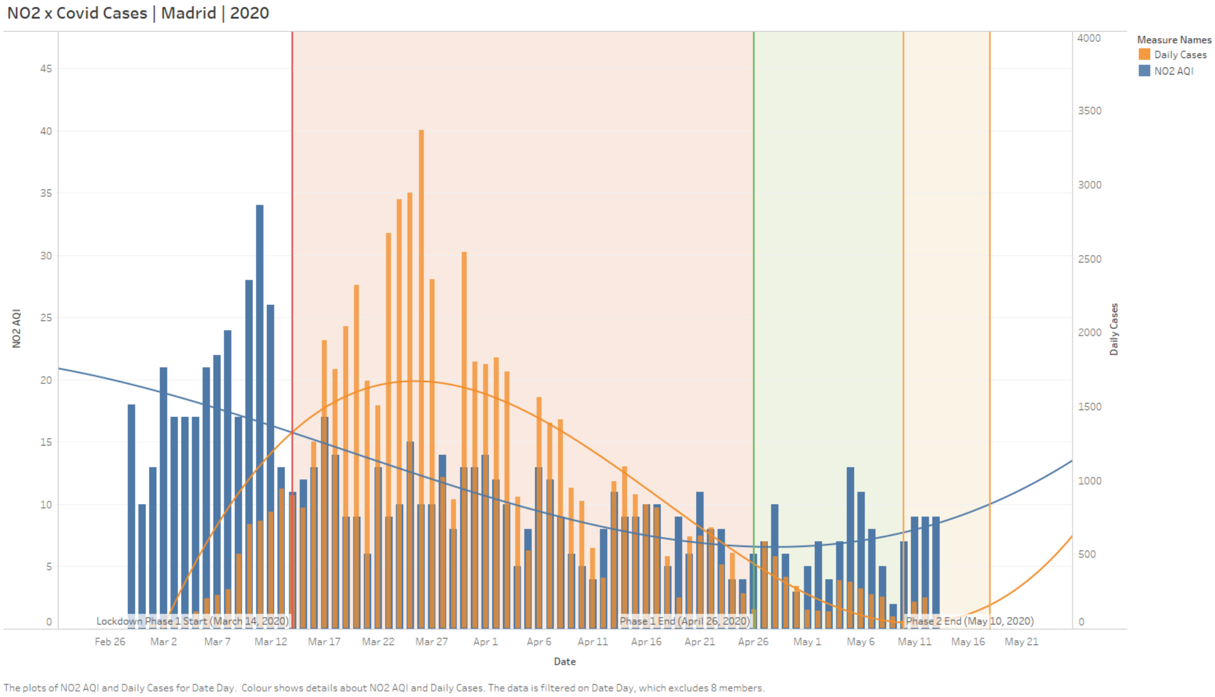


Figure 19.

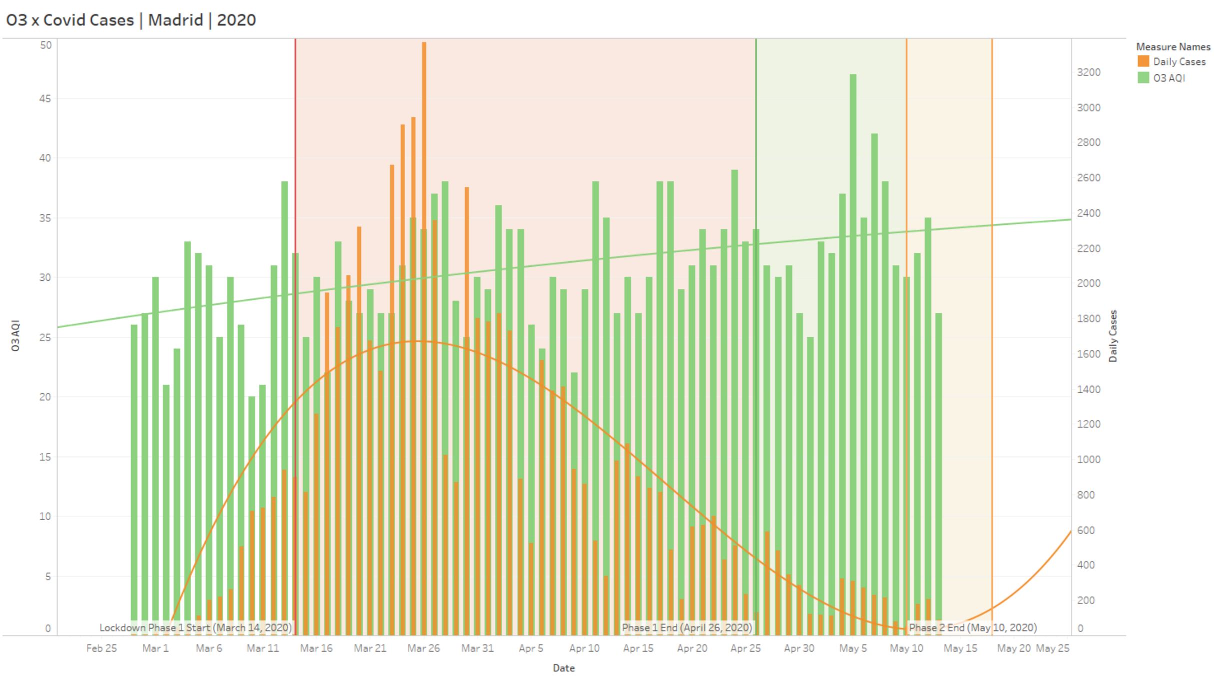


Figure 20.

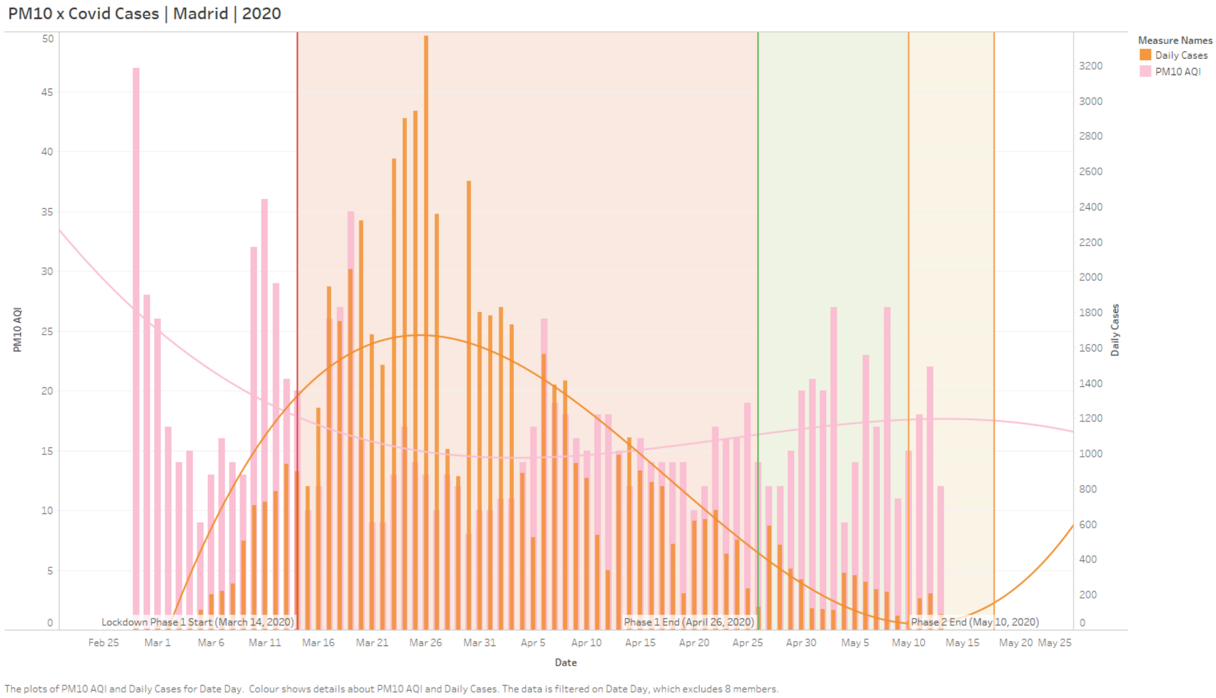


Figure 21.

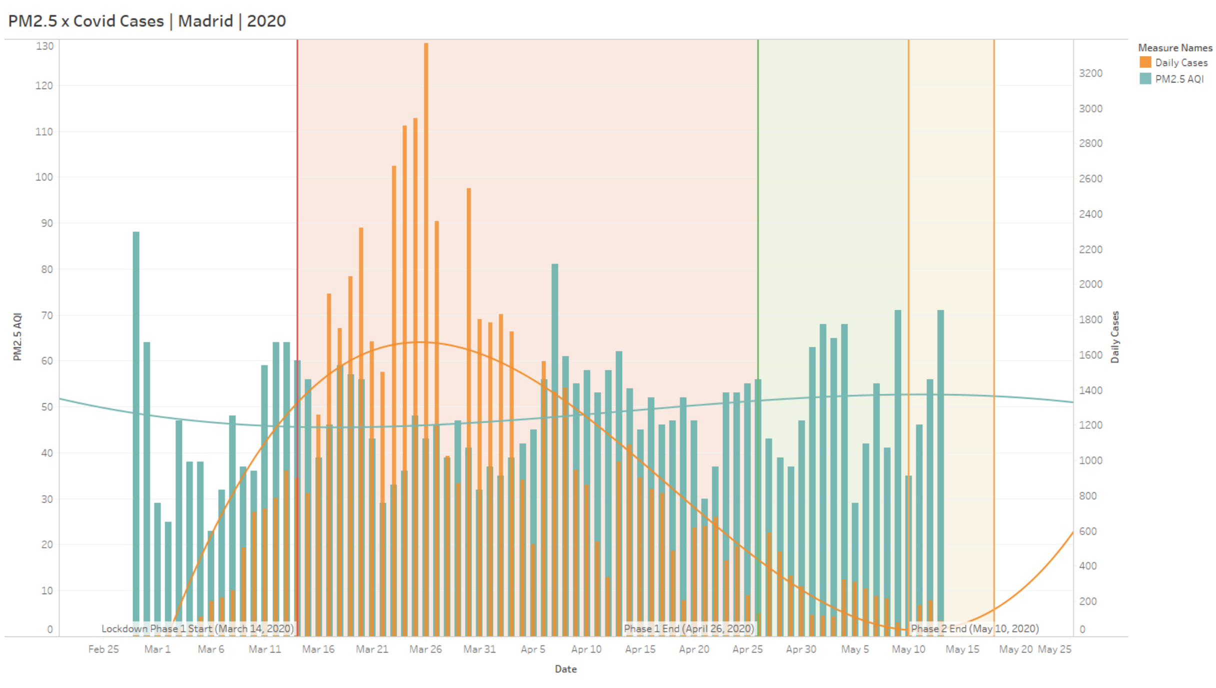


Figure 22.

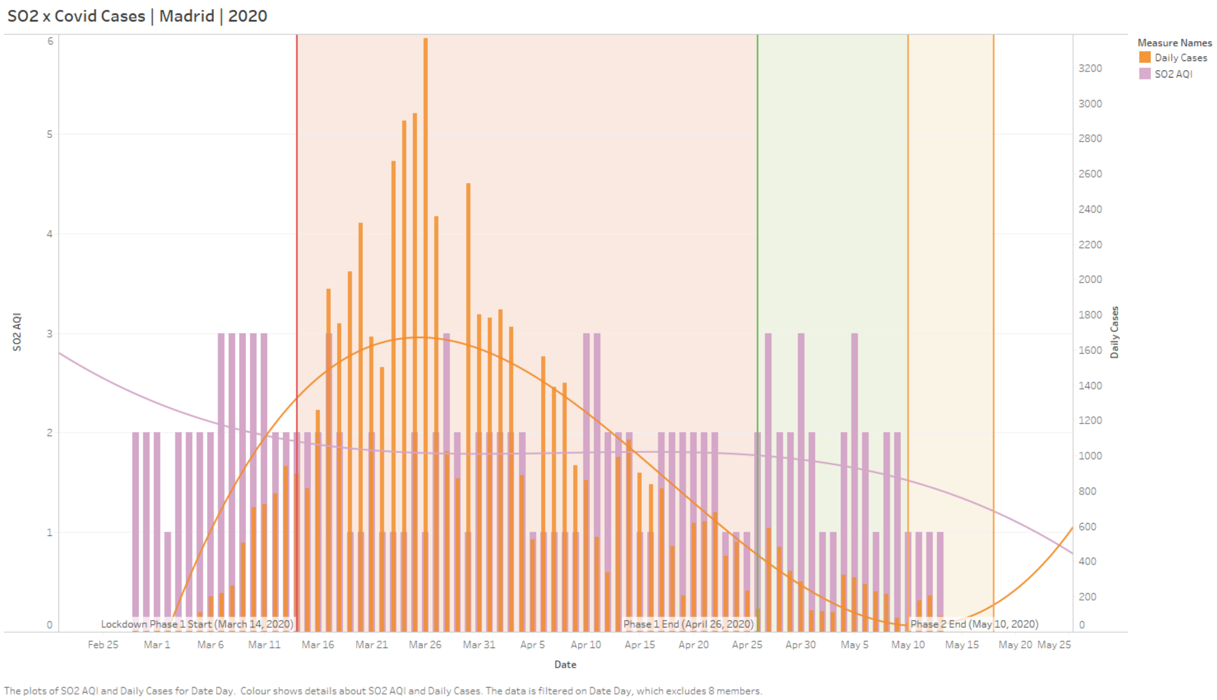


Figure 23.

8.5 Task Distribution:

Mudit Chaudhary: Visualization, Implementation, Data and Data Pre-Processing and Insights
Shreyas Goyal: Literature Review, Data Sources, Insights and Conclusion
Aman Saraff: Introduction, Visualization, Lockdown Study, Insights and Correlation

10. References

[1] Business Standard, “What is Air quality index, AQI Categories, AQI Objectives, AQI News,” *Business Standard*. [Online] [Accessed: 20-May-2020]. Retrieved From: https://www.business-standard.com/about/what-is-air-quality-index.

[2] “Air quality index” (27th March 2020), *Wikipedia* [Online] [Accessed: 20-May-2020]. Retrieved From: https://en.wikipedia.org/wiki/Air_quality_index.

[3] ”Inhalable Particulate Matter and Health (PM2.5 and PM10)”, *California Air Resources Board*. [Online] [Accessed: 20-May-2020]. Retrieved From: https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health.

[4] “Ground-level Ozone Basics”(31st Oct 2018), *EPA* [Online] [Accessed: 20-May-2020]. Retrieved From: https://www.epa.gov/ground-level-ozone-pollution/ground-level-ozone-basics#formation.

[5] “Sulfur dioxide (SO2)”, *Department of Agriculture, Water and the Environment* [Online] [Accessed: 20-May-2020]. Retrieved From: https://www.environment.gov.au/protection/publications/factsheet-sulfur-dioxide-so2.

[6] P. Godec (24th Apr 2020), “Earth: ‘Thanks to COVID-19, I can breathe easier’”, *Medium* [Online] [Accessed: 20-May-2020]. Retrieved From: https://towardsdatascience.com/earth-thanks-to-covid-19-i-can-breathe-easier-ebc2c11337c5.

[7] “Airborne Particle Levels Plummet in Northern India”, *NASA*. [Online] [Accessed: 20-May-2020][Accessed: 20-May-2020]. Retrieved From: https://earthobservatory.nasa.gov/images/146596/airborne-particle-levels-plummet-in-northern-india.

[8] “COVID-19 Pandemic in the United Kingdom”, *Wikipedia* [Online] [Accessed: 21-May-2020]. Retrieved From: https://en.wikipedia.org/wiki/COVID-19_pandemic_in_the_United_Kingdom#January_2020:_first_confirmed_cases

[9] Burgess, M. (21st May 2020), “When will the lockdown end? The UK’s lockdown rules, explained”, *Wired.* [Online] [Accessed: 21-May-2020]. Retrieved From: https://www.wired.co.uk/article/uk-lockdown

[10] “COVID-19 Pandemic Lockdown in India”, *Wikipedia* [Online] [Accessed: 21-May-2020]. Retrieved From: https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdown_in_India

[11] “COVID-19 Pandemic in India”, *Wikipedia* [Online] [Accessed: 21-May-2020]. Retrieved From: https://en.wikipedia.org/wiki/COVID-19_pandemic_in_India

[12] Kaplan, J., Frias, L. and Johnsen, M.M. (17th APr 2020), “A third of the global population is on coronavirus lockdown — here's our constantly updated list of countries and restrictions”, *Business Insider* [Online] [Accessed: 21-May-2020]. Retrieved From: https://www.businessinsider.in/international/news/a-third-of-the-global-population-is-on-coronavirus-lockdown-x2014-hereaposs-our-constantly-updated-list-of-countries-and-restrictions/slidelist/75208623.cms#slideid=75209016

[13] "COVID-19 Pandemic Lockdown in Hubei", *Wikipedia* [Online] [Accessed: 21-May-2020]. Retrieved From: https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdown_in_Hubei

[14] News Desk (19 May 2020), "Lifting of lockdown in Spain – full details of all phases for all regions", *Spain in English* [Online] [Accessed: 21-May-2020]. Retrieved From: https://www.spainenglish.com/2020/05/20/lifting-lockdown-spain-full-details-phases/

[15] News Desk (19 May 2020), "Spain officially in Lockdown from Midnight", *Spain in English* [Online] [Accessed: 21-May-2020]. Retrieved From:https://www.spainenglish.com/2020/03/14/spain-lockdown-coronavirus-covid19/

[16] Chaudhary, M., Goyal, S. and Saraff, A. (27 Apr 2020), "Project Proposal: Assessing the Impact of COVID-19 Lockdowns on the Air Quality", *SEEM2460: Introduction to Data Science, CUHK*.