**DATA 601: Group Project Report**

Air Quality Across Countries in COVID-19

**Prepared by:**

Group 4

**Members:**

**Monica Chandramurthy, 30191289 Saketh Ram Mamillapalli, 30187315**

## Table of Contents

Introduction Guiding Question

1. Data Collection
2. Data Preprocessing
   1. Setting Air Quality Index Scale and Color Legend
   2. Calculating AQI by US-EPA 2016 standard for each Pollutant

# Answering guiding questions

* 1. Guiding Questions 1: How is AQI distributed across countries during COVID-19 (2019-Q4 to 2022- Q4)?
  2. Guiding Questions 2: Which countries experienced an increase in AQI during the sampled period, and what were the highest levels of pollution observed by country and city?
  3. Guiding Questions 3: What factors influence AQI, and how do different pollutants correlate with AQI?
  4. Guiding Questions 4: Is there a relationship between AQI and COVID-19 cases in each country during the pandemic?

Conclusion

Future Scope and Recommendations References

# INTRODUCTION:

Air pollution is a major environmental issue that can significantly impact human health. [1]. A variety of pollutants cause it, including particulate matter (PM2.5 and PM10), nitrogen dioxide (NO2), sulphur dioxide (SO2), carbon monoxide (CO), and ozone (O3). [i] These pollutants are monitored by government agencies using the air quality index (AQI), which communicates the level of air pollution to the public. Exposure to high levels of air pollution can lead to respiratory and cardiovascular diseases, headaches, and eye irritation.

This project aims to demonstrate air quality's effect on health, particularly regarding COVID-19. We aim to raise awareness about the impact of air pollution and key pollutants affecting the AQI by analyzing the AQI across seven significant countries, including Brazil, India, Italy, France, the United States, Canada, and China, during the COVID-19 pandemic (2019-Q4 to 2022-Q4). This study aims to examine whether there was a trend in AQI affecting the number of COVID-19 cases in each country, which can contribute to a better understanding of the relationship between human activities and air pollution[i]

# GUIDING QUESTIONS:

1. How is AQI distributed across countries during COVID (2019-Q4 to 2022-Q4)?
2. Which countries experienced a rise in AQI during the sampled duration, and which country and city had the highest pollution level?
3. What factors influenced AQI, and what is the correlation between different pollutants and AQI?
4. Was there a relationship between AQI and COVID-19 cases in each country during the pandemic?

# DATA COLLECTION:

The data for this project were collected from two sources. The first source was the Air Quality Open Data Platform [ii], which provided AQI data from 2019-Q4 to 2022-Q4 in nine CSV files. The dataset offers average (median) values for each major city based on data from several stations, including air pollutant species such as PM2.5, PM10, and ozone, as well as meteorological data such as wind and temperature. All air pollutant species are converted to the US EPA standard, and all dates are UTC-based [ii]. The second source is Kaggle, where COVID-19 data were collected from the World Health Organization's website, specifically the daily total cases, daily new cases, active cases, total daily deaths, and daily new deaths for each country [iii]. Figure 1 and Figure 2 show the snapshots of the raw data for the AQI and COVID-19 datasets, respectively.

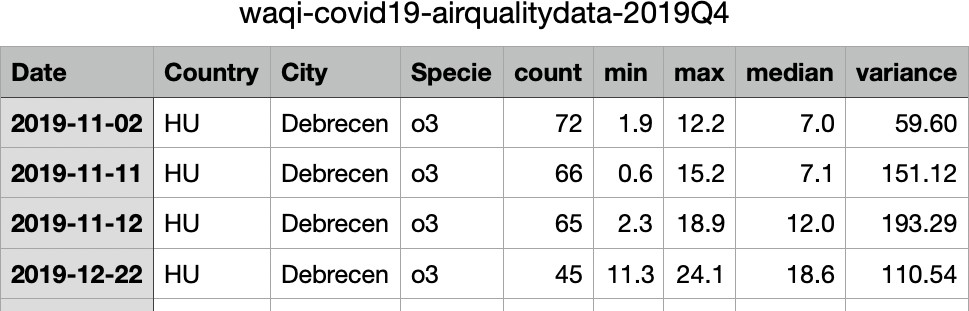


Figure 1: Raw AQI data for the period of 2019-Q4

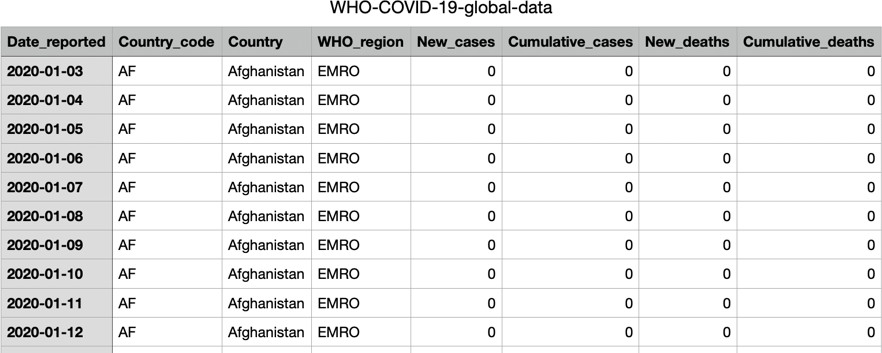


Figure 2: Raw COVID-19 data until 2021

# DATA PREPROCESSING:

The following steps were taken to preprocess the data:

1. The AQI data for the 2019 Q4-2021 Q4 (9csv files) was imported from Air Quality Open Data Platform [ii] using the pandas’ library in Python and stored in data frames.
2. The data imported from different files were merged, resulting in a combined dataset with a total of 5,240,359 rows and nine columns, as shown in Figure 3.

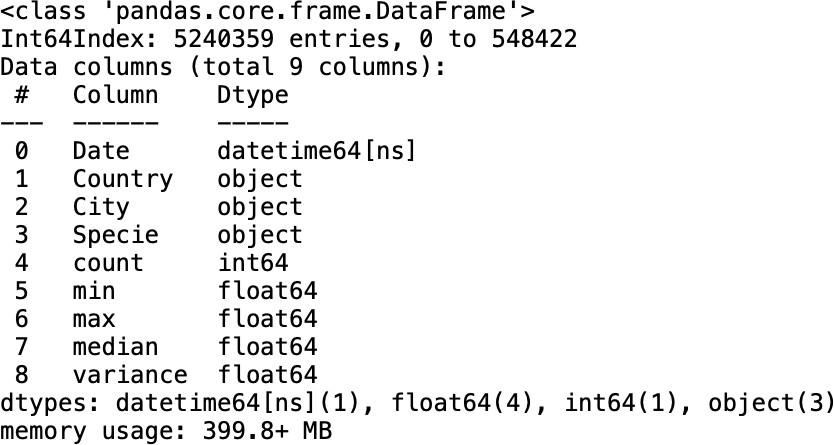


Figure 3: Air Quality Dataset after merging

1. Columns not needed for analysis, such as minimum, maximum, and variance, were dropped. Only the date, country, city, specie, and median value were retained for calculating the AQI and conducting further analysis.
2. The dataset was checked for missing values, and we found that it was clean without any null or missing values. Figure 4 shows the summary statistics for each column in the dataset.

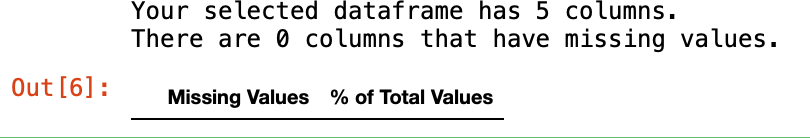


Figure 4: Missing values

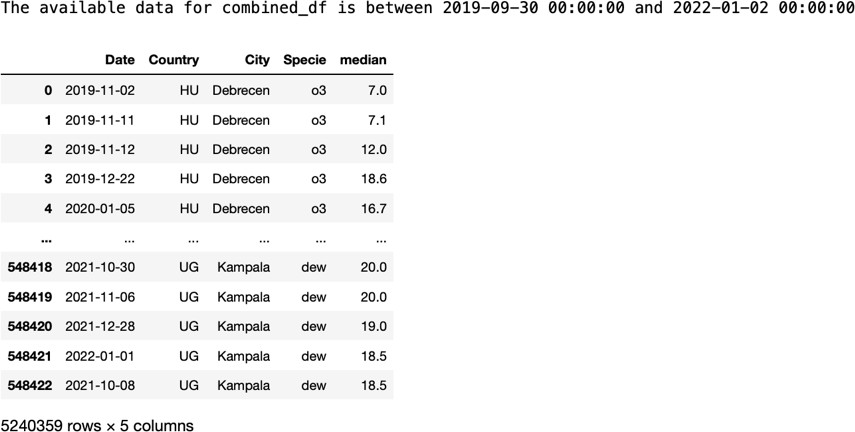


Figure 5: AQI data after performing the above Data preprocessing and Data Wrangling

1. The AQI for each pollutant was calculated based on the US-EPA 2016 standard using the following formulas in section 2.2.

## Setting Air Quality Index Scale and Color Legend

The table in Figure 6 presents the Air Quality Index (AQI) scale, defined by the US EPA standards, including AQI ranges, corresponding health effects, and colour codes for each field. The table has four columns: AQI, Remark, Color Code, and Possible Health Effects. The AQI column displays the ranges of AQI values from 0 to 500. We used the same scale and colour palette based on this table to calculate and represent the AQI throughout our project. [iv]

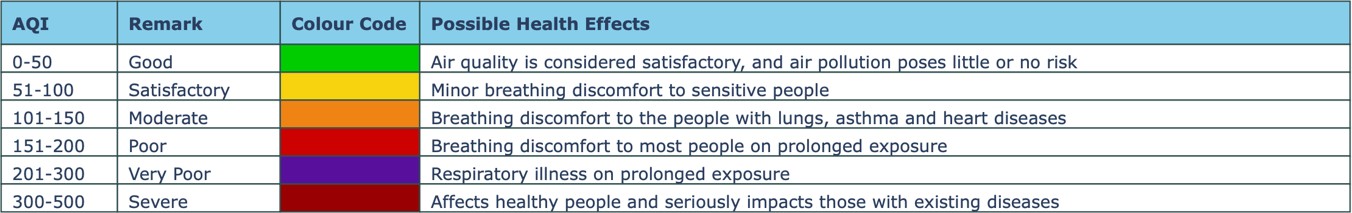


Figure 6: Air Quality Index Scale and Color Legend

## Calculating AQI by US-EPA 2016 standard for each Pollutant

The datasets include different types of air pollutants, which can be classified into particulate matter (PM2.5 and PM10), nitrogen oxides (NO2), sulphur dioxide (SO2), carbon monoxide (CO), and ozone (O3) [v]. Figure 7 displays the concentration of different pollutants across countries.

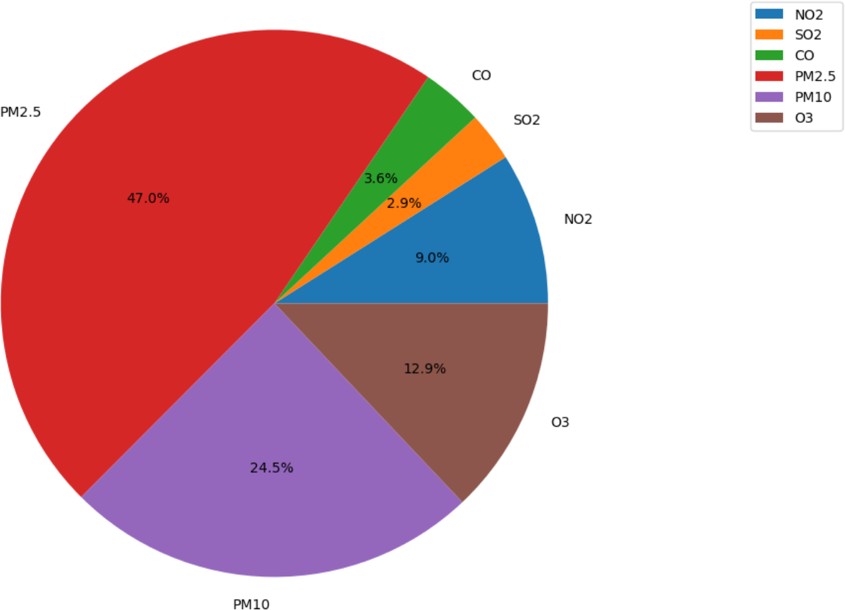


Figure 7: Concentration on different pollutants across the countries To understand the pollutants better, here are some key points:

* Particulate matter is a mixture of solids and liquids in the air, including carbon and chemicals. The most harmful particles are PM2.5 and PM10.
* Nitrogen oxides are gases that can cause respiratory problems.
* Sulphur dioxide is a colourless gas formed by burning fuel containing sulphur, contributing to air pollution and respiratory problems.
* Carbon monoxide is a highly poisonous gas produced by burning fuel, which can be deadly at high levels.
* Ozone is a highly irritating gas that forms when nitrogen oxides and volatile organic compounds react in sunlight, causing respiratory problems.

AQI (Air Quality Index) is a numerical index used to report daily air quality. It provides a standardized way to assess and communicate air quality, with higher values indicating poorer air quality and a greater risk to public health. AQI is calculated for each pollutant using the US-EPA 2016 standard, with formulas based on the pollutant's concentration in micrograms per cubic meter. [v]

The AQI for each pollutant is calculated using the following formulas based on the US-EPA 2016 standard [v]:

1. PM2.5 AQI = (500/50) \* (Cp / 12.0) ^ 0.333
2. PM10 AQI = (500/50) \* (Cp / 54.0) ^ 0.333
3. SO2 AQI = (500/75) \* (Cp / 75.0) ^ 0.667
4. NO2 AQI = (500/100) \* (Cp / 100.0) ^ 0.667
5. O3 AQI = (500/100) \* (Cp / 0.070) ^ 0.667
6. CO AQI = (500/50) \* (Cp / 9.0) ^ 0.667

Where Cp is the concentration of the pollutant in micrograms per cubic meter (µg/m3), these formulas are used to calculate the AQI for each pollutant based on its concentration in the air.

The source of these formulas is the US EPA Air Quality Index - A Guide to Air Quality and Your Health (2018). [vi]

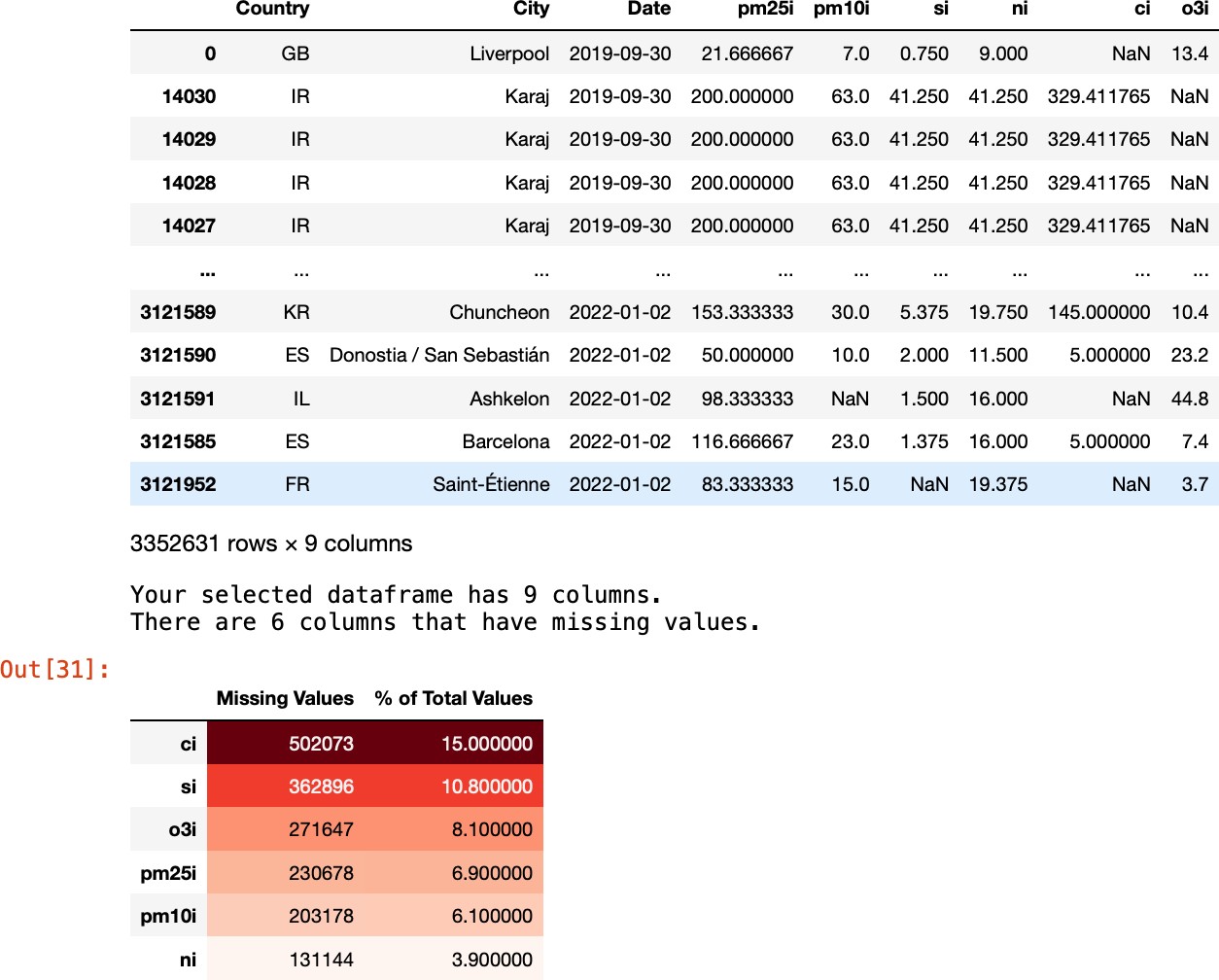


Figure 8. AQI Data after calculating pollutant levels and merging the data frame

It can be observed that some pollutants did not have an entry, resulting in missing values when restructuring the data by pollutant levels, as shown above. Therefore, we replaced these missing values in the abovementioned columns to avoid future errors.

We then calculated the AQI and the AQI bucket that the country belonged to, as shown below:



Figure 9. AQI Data after AQI levels and its respective bucket.

# ANSWERING GUIDING QUESTIONS:

## Guiding Question 1: How is AQI spread across different Countries during COVID (2020-2021) (2019-Quarter 4 to 2021-Quarter 1)?

This question explores how the AQI spread across different countries during the COVID pandemic from 2019 Q4 to 2021 Q4. This will help in understanding how the air quality has been impacted during the COVID period in different countries.

An interactive choropleth map using the Plotly Express library is created to visualize the AQI across countries over time. The map is colour-coded based on the AQI values, with a fixed scale like the one defined by the US-EPA 2016 standard. The animation\_frame parameter is used to animate the map over time, and the locationmode parameter is set to "country names" to map country names to their respective locations.

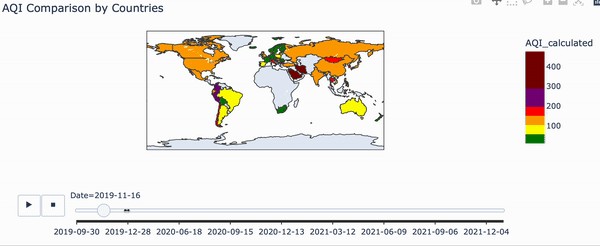


Figure 9. Interactive Map on AQI across the countries over sampled time-

### Observations:

The AQI values in several countries appear to be above the US-EPA standard limit of 100 during the sample period, indicating high pollution levels in these regions. However, there are a few countries where AQI remained within the safe range throughout the observed period, like Australia and France.

It is interesting to note that the AQI levels in some countries appear to vary significantly over time. For instance, AQI values in India and China are consistently high throughout the observed period. At the same time, in some other countries, such as the United States, there are noticeable variations in the AQI values over time.

## Further data wrangling was conducted to plot maps. The following steps were taken:

* First, a data frame was created to store all cities and countries.
* The geopy. geocoders - Nominatim library function was used to obtain the coordinates for each city.
* A lambda function was used to separate the latitude and longitude values and store them in the dataset.
* After the calculations, missing values were checked to ensure no null values were present. Any invalid values were replaced as necessary.

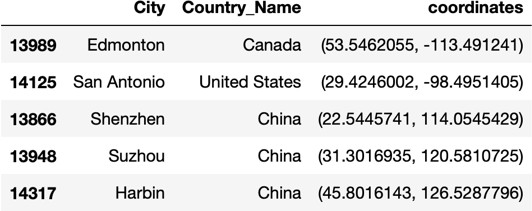


Figure 10. Coordinates calculated by cities in a data frame

## Visualizing AQI across cities of major countries using a bubble map:

We used the Plotly Express library to create an interactive bubble map visualizing the AQI (Air Quality Index) across various cities of significant countries over time. Compared to the previous visualization of AQI across countries over time using a choropleth map, the bubble map visualizes the AQI values for individual cities. In contrast, the choropleth map only provides an average value for the entire country. This allows for a more granular view of the AQI levels and can help identify areas within a country with exceptionally high or low levels of air pollution.

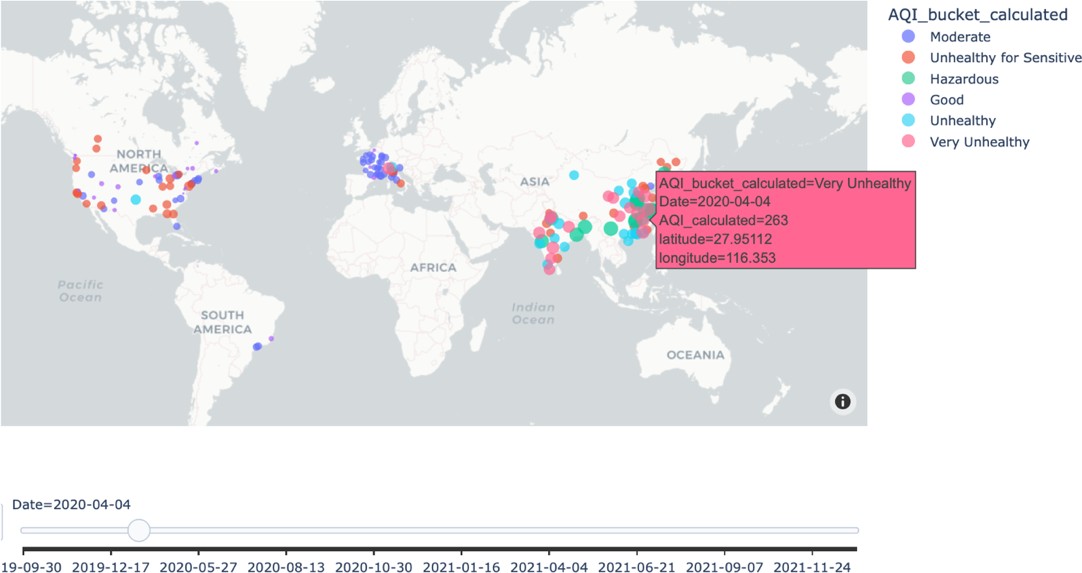


Figure 12. AQI across the cities over sampled time

### Observations:

The bubbles in several cities appear to be above the safe AQI limit of 100 during the sampled period, indicating high pollution levels in these regions. However, there are also many cities where AQI remained within the safe range throughout the observed period.

The AQI levels in some cities appear to vary significantly over time. For instance, the AQI values in Delhi, India, and Beijing, China, are consistently high throughout over 250 in the observed period. At the same

time, in some other cities, such as Los Angeles, USA, there are noticeable variations in the AQI values over time.

Overall, the bubble map provides an excellent visual representation of the AQI values at the city level, highlighting the significant differences across different cities over time. This can help us identify regions that require immediate attention to control pollution levels and ensure the environment's safety and public health.

## Visualizing AQI in 2020 and 2021 only for the major countries (Brazil, China, India, the United States, France, Italy, and Canada).

We visualized the AQI data for major countries across 2020 and 2021 years to see which countries experienced higher levels of air pollution and how the levels changed over time. The countries selected are India (IN), Brazil(BR), USA(US), China(CN), France(FR), Italy(IT), Canada(CN).

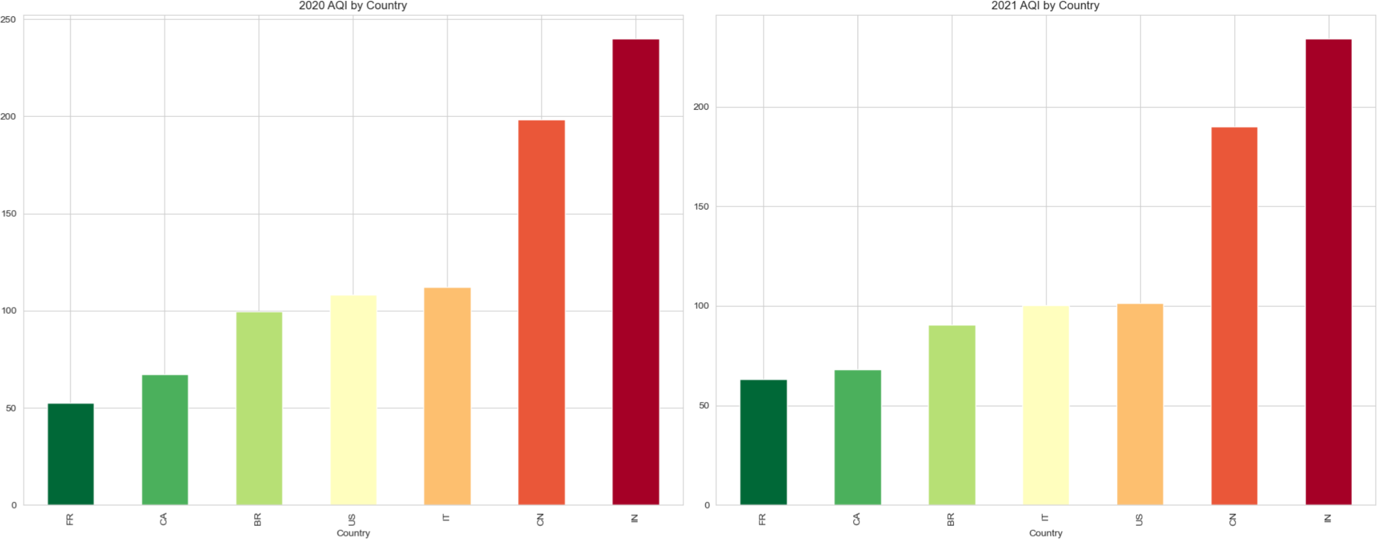


Figure 13. Average AQI across Major Countries in 2020 and 2021, respectively

### Observations:

* Overall, the bar charts suggest that significant countries across the world experienced high levels of air pollution in 2020 and 2021, with India having an annual average AQI of over 240 and 234, respectively, and China with an average AQI of over 198 and 190 respectively, consistently ranking among the most polluted countries.
* USA had 58,733 fires (5th least) and burned 7,139,713 acres (10th least), which is 121.56 acres burned/fire (8th most) in 2021, which led to the release of pm2.5 pollutant leading to a rise in AQI levels [viii]. Italy also had several major wildfires in the summer of 2021. [xi]
* The charts also highlight the need for continued efforts to reduce air pollution and improve air quality in these and other countries worldwide.

## Guiding Question -2: Which countries saw a rise in AQI in the sampled duration?

To identify which countries saw a rise in AQI in the sampled duration, we plotted the trend in AQI for each country over time, taking the moving average of the AQI value. The tested course was from 2019

to 2021, and the major countries considered were Brazil, China, India, the United States, France, Italy, and Canada.

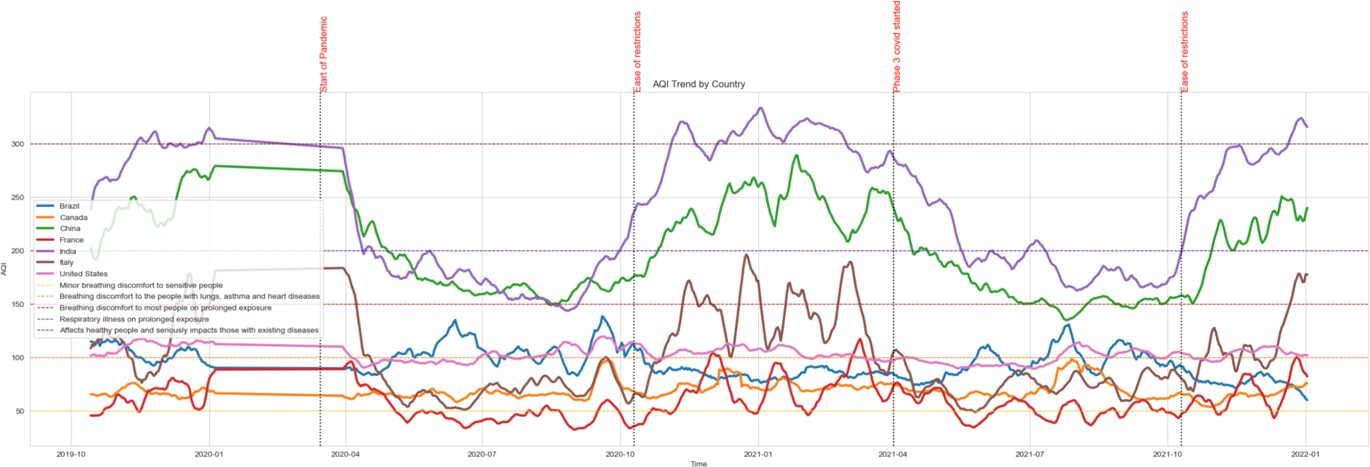


Figure 14. The trend in AQI for Major Countries in sampled duration

### Observations:

* The above graph visualizes the trend in AQI over time for the major countries of Brazil, China, India, the United States, France, Italy, and Canada. The AQI is calculated as the moving average of the AQI values over a window size of 15.
* The graph also includes lines representing the different levels of AQI and vertical lines indicating key events, such as the start of the pandemic and easing restrictions. [xii]
* On March 15, 2020, COVID-19 was declared a pandemic by the World Health Organization, leading to regional lockdowns in April and May. However, the graph shows that most countries experienced a decrease in AQI around this time, likely due to reduced human activity and emissions.
* However, as the pandemic progressed and restrictions were lifted, many countries saw a rise in AQI around October 2020.
* The Delta variant's discovery in mid-2021 led to another round of regional lockdowns, causing AQI levels to dip in mid-2021. India saw the most significant increase in AQI, followed by China at the end of 2021. [xi]
* IQAir’s 2020 World Air Quality Report said human-related emissions from industry and transport fell during lockdowns, and 65% of global cities analyzed experienced better air quality in 2020 compared to 2019. Some 84% of nations polled reported air quality improvements overall. [xii]
* The above fact also supports the trend in figure 14 that the COVID-19 pandemic and the associated restrictions significantly impacted AQI levels in some countries.

## Most Polluted Country and City:

We analyzed the Air Quality Index (AQI) data to identify the most polluted city and country based on the average AQI.

First, they grouped the AQI data by city and year to get the average AQI of each town in each year. After that was sorted, the data was to find the city with the highest AQI, which is identified as the most polluted city. Then, the average AQI of this city is also calculated and printed.

We then grouped the AQI data by country to get the average AQI for each country and sorted the data to find the country with the highest AQI. Finally, the most polluted country is identified, and the average AQI of this country is calculated and printed.

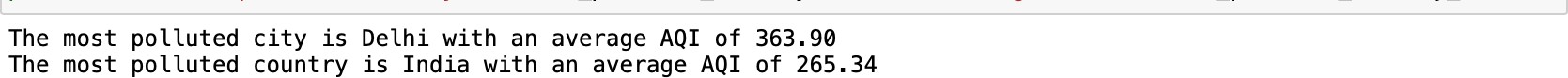


Figure 15. Most Polluted cities and country

## Most Polluted Country:

To visualize the most polluted country geographically, we use the Folium library. We plot the data for the most polluted country on a Folium map centred on that country. To differentiate AQI levels, we define a function that maps the AQI levels to colours and adds a marker for each location in the AQI data on the map. The AQI level determines the marker's size and colour, and the latitude and longitude coordinates determine the location of each marker in the data. [xiv]

The resulting map can identify areas with high pollution levels and target pollution control measures. This visualization helps highlight the locations with the highest AQI levels in the most polluted country.

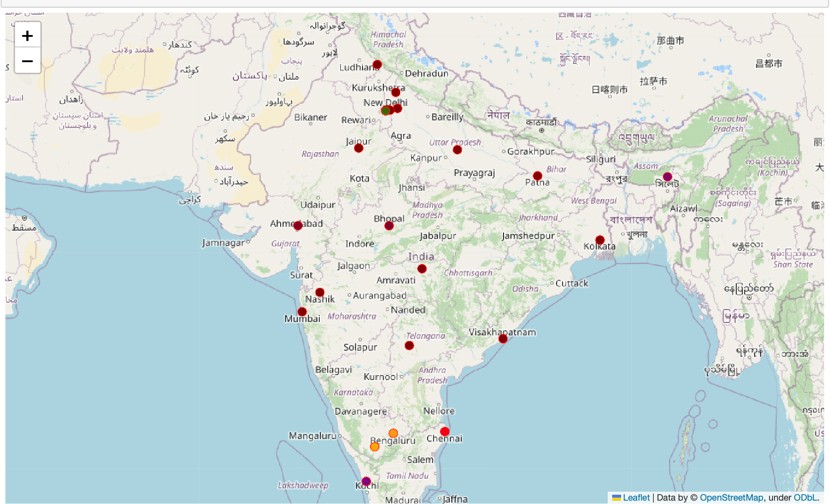


Figure 16. Most Polluted Country

### Observations:

* As seen in figure 16, most of the major cities in India have AQI levels over 150, indicating the air is hazardous to breathe, i.e., exposure to this air would cause serious respiratory diseases.
* In India, approximately 59% of the air pollution can be attributed to Dust & Construction, with Waste Burning being the second largest contributor. Crafting activities are prevalent in

urban areas, whereas Waste Burning primarily occurs in rural regions that are involved in agriculture. [xix].

* According to a study published in The Lancet, it is estimated that in 2019, air pollution was responsible for around 1.67 million deaths in India. This number may fluctuate annually and could be affected by various factors such as weather patterns, industrial activities, and policies to reduce air pollution. [xx]

## Most Polluted City in each country:

Visualizing AQI across cities in each country, most to least polluted:

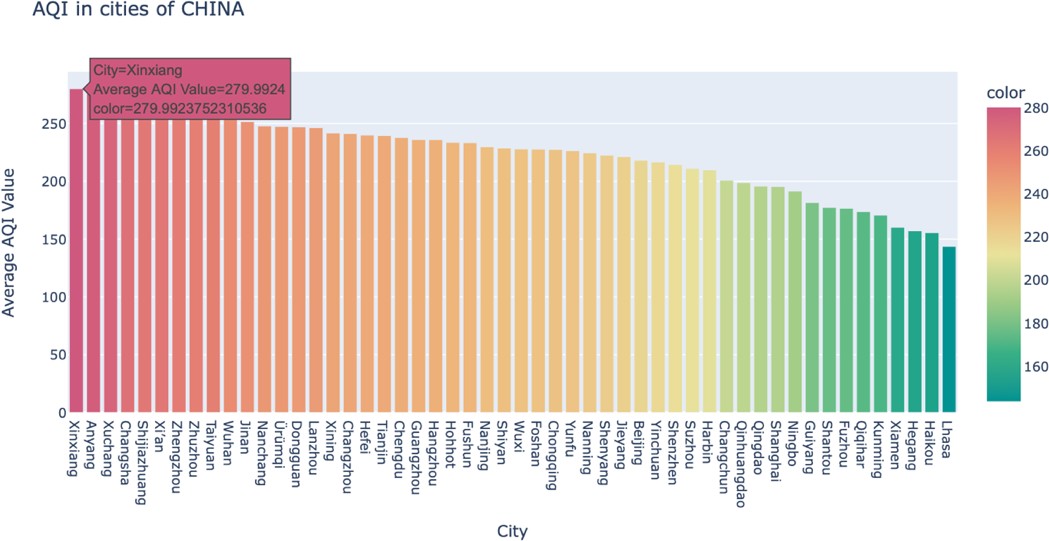


Figure 17. AQI in cities in China

Observation: Xinjiang is the most polluted city in China, with an average AQI of 279.9924

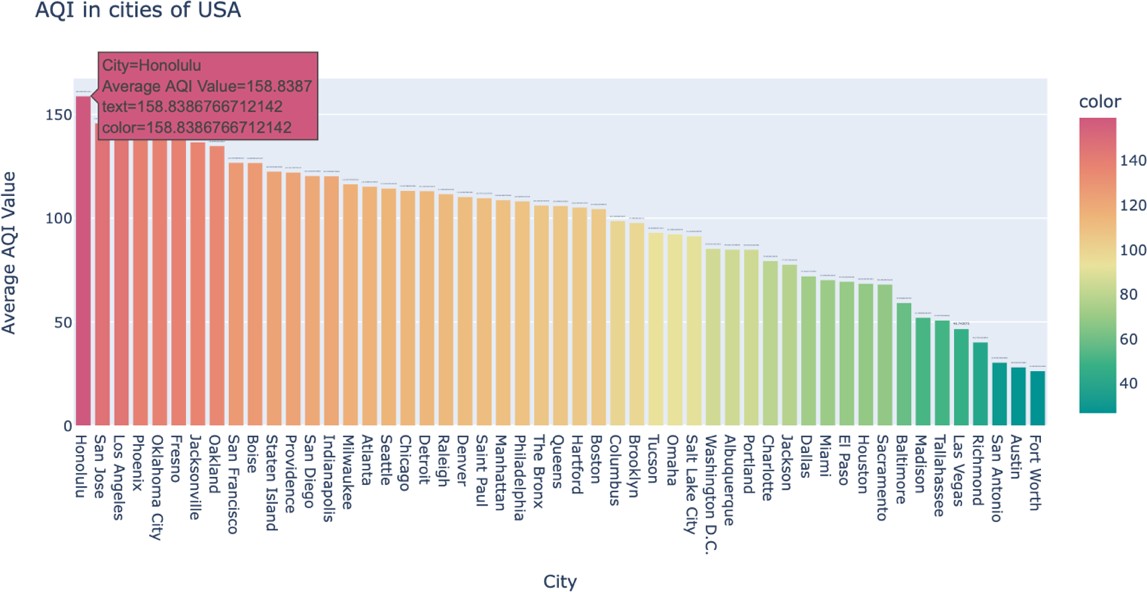


Figure 18. AQI in cities of the USA

Observation: Honolulu is the most polluted in the USA, with an average AQI of 158.8387.

Honulu has large fires which occur almost annually on Oahu's Waiʻanae coast, impacting predominantly Native Hawaiian and Pacific Islander communities. Also contributing to high AQI levels. [xxi]

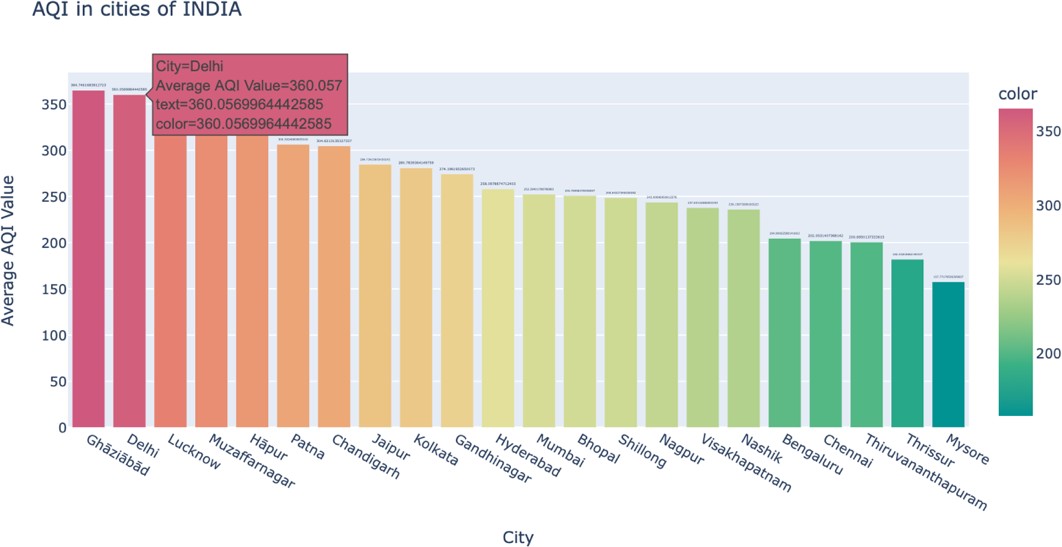


Figure 19. AQI in cities of India

Observation: Delhi is the most polluted city in China, with an average AQI of 360.057

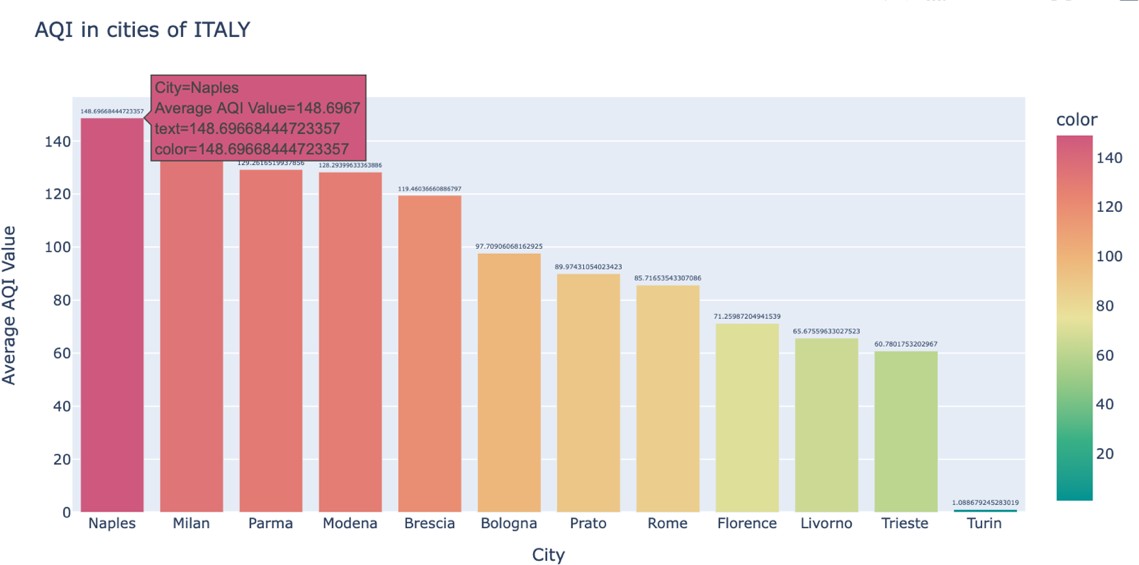


Figure 20. AQI in cities of Italy

Observation: Naples is the most polluted city in Italy, with an average AQI of 148.6967

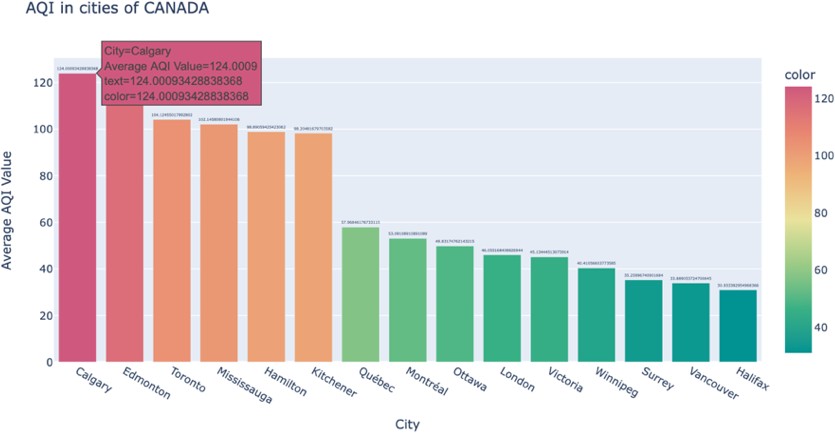


Figure 21. AQI in cities of Canada

Observation: Calgary is the most polluted city in Canada, with an average AQI of 124.0009.

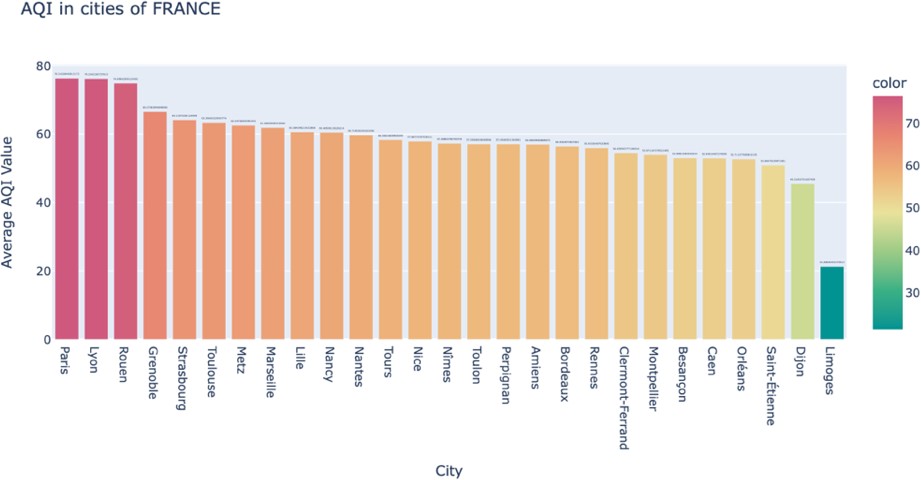


Figure 22. AQI in cities in China

Observation: Paris is the most polluted city in France, with an average AQI of 78.5441

## The trend in AQI – Taking Moving Average in Most Polluted City: Delhi

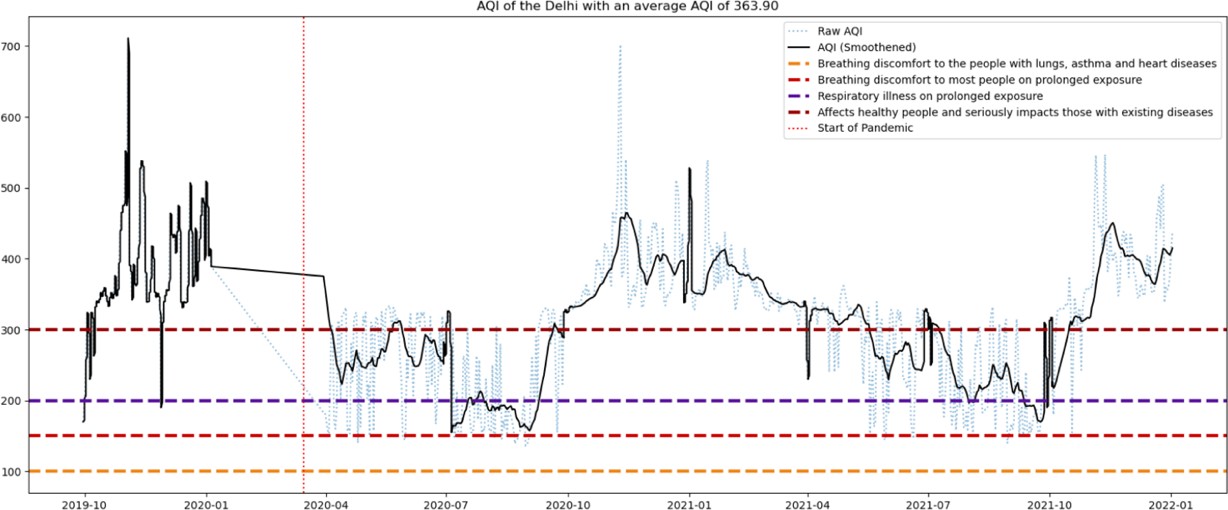


Figure 23: AQI Trend in Delhi (10/2019-12/2021)

From the data, we can observe the following trends:

* In general, the AQI of Delhi is highest in the winter months of December, January, and February and lowest in the summer months of June, July, and August.
* In October 2020, Delhi experienced a sharp surge in air pollution, with PM2.5 levels peaking at

829.2 micrograms per cubic meter (ug/m3) and the U.S. Air Quality Index recording 1005, which surpasses the index's maximum limit of 500. As a result, the graph shows a significant spike in air pollution during that time. [xvi]

* The AQI levels in March 2020 dropped significantly in Delhi, possibly due to the nationwide lockdown that was in place in India during that time compared to 2019, which can also be observed in Figure 23 and Figure 24.



Figure 24: Delhi in 04/2019 vs Delhi in 04/2020 – during lockdown [source: [https://w](http://www.bbc.com/news/science-environment-57149747)ww.bbc.[com/ne](http://www.bbc.com/news/science-environment-57149747)w[s/science-environment-57149747]](http://www.bbc.com/news/science-environment-57149747)

* There was a spike in AQI in September 2020, possibly due to the ease of restrictions on COVID.
* The AQI in October is consistently high across the years, possibly due to the increase in crop burning and the Diwali festival in India. [xviii]
* Overall, the data shows that the AQI of Delhi is highly dependent on various factors such as weather conditions, seasonal changes, environmental policies, and human activities. It also highlights the need for constant monitoring and measures to control pollution levels in the city.

## Guiding Question 3: Which parameter affects the AQI while establishing the correlation between different pollutants and AQI?

This question aims to identify the factors that influenced AQI, pollutants such as NO2, SO2, CO, O3, PM2.5, and PM10. The question also looks at the correlation between different pollutants and AQI, which can provide insights into which pollutants had the greatest impact on AQI.

To find the correlation between pollutants and AQI, we can use the corr method in Pandas. This method calculates the Pearson correlation coefficient between columns in a data frame. The correlation coefficient is between -1 and 1, representing the strength of the linear relationship between the two variables.

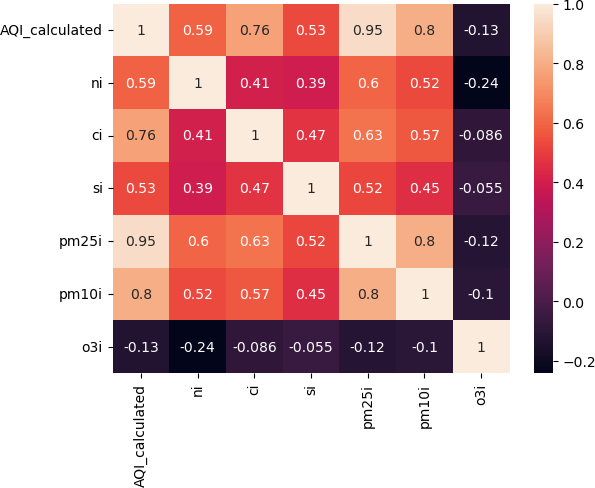


Figure 25: Correlation matrix of AQI vs different pollutants.

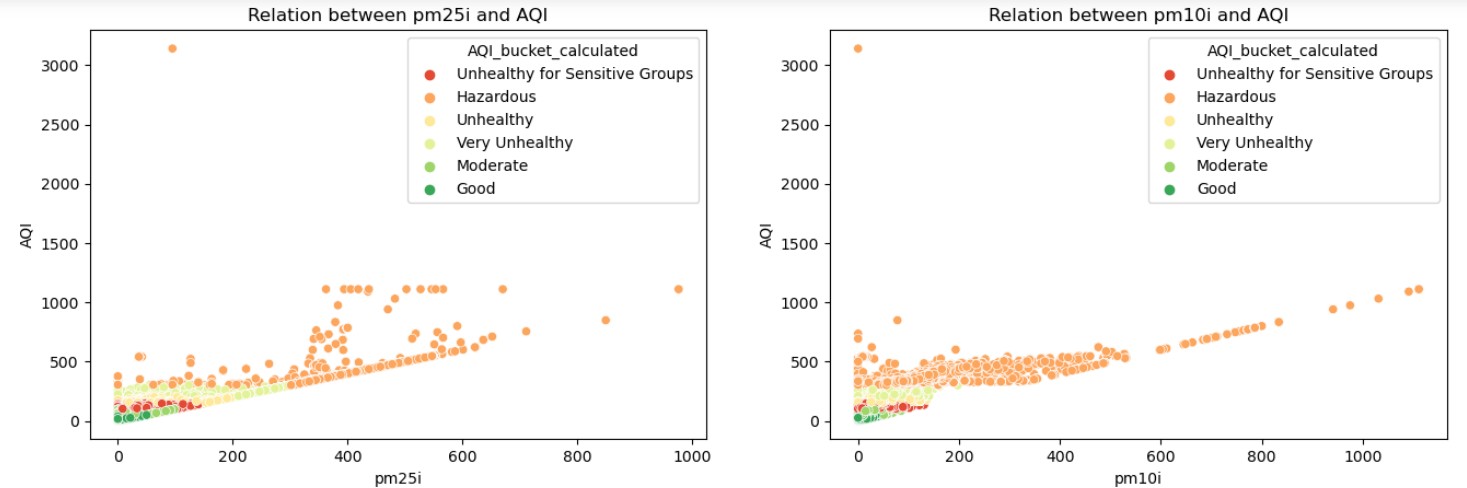
In the matrix shown in figure 25, AQI is highly correlated with PM2.5 (0.950752), PM10 (0.801915), and CO (0.764433), with moderate correlation with SO2 (0.589289) and a weak correlation with NO2 (0.525227) and O3 (-0.125582).

The high correlation between AQI and PM2.5, PM10, and CO suggests that these pollutants are the main contributors to air pollution. In contrast, the weak correlation with NO2 and O3 indicates that they have a lesser impact.

## Visualizing the relationship between AQI and Pollutants using scatter plots

Using a scatterplot matrix, we can further visualize the relationship between air quality variables and the AQI. The scatterplot matrix shows the scatterplots between each variable and the AQI, with the points coloured by AQI level.

The pollutants considered are pm25i(PM2.5), pm10i(PM10), ni(NO2), ci(CO), si(SI), and o3i(O3).



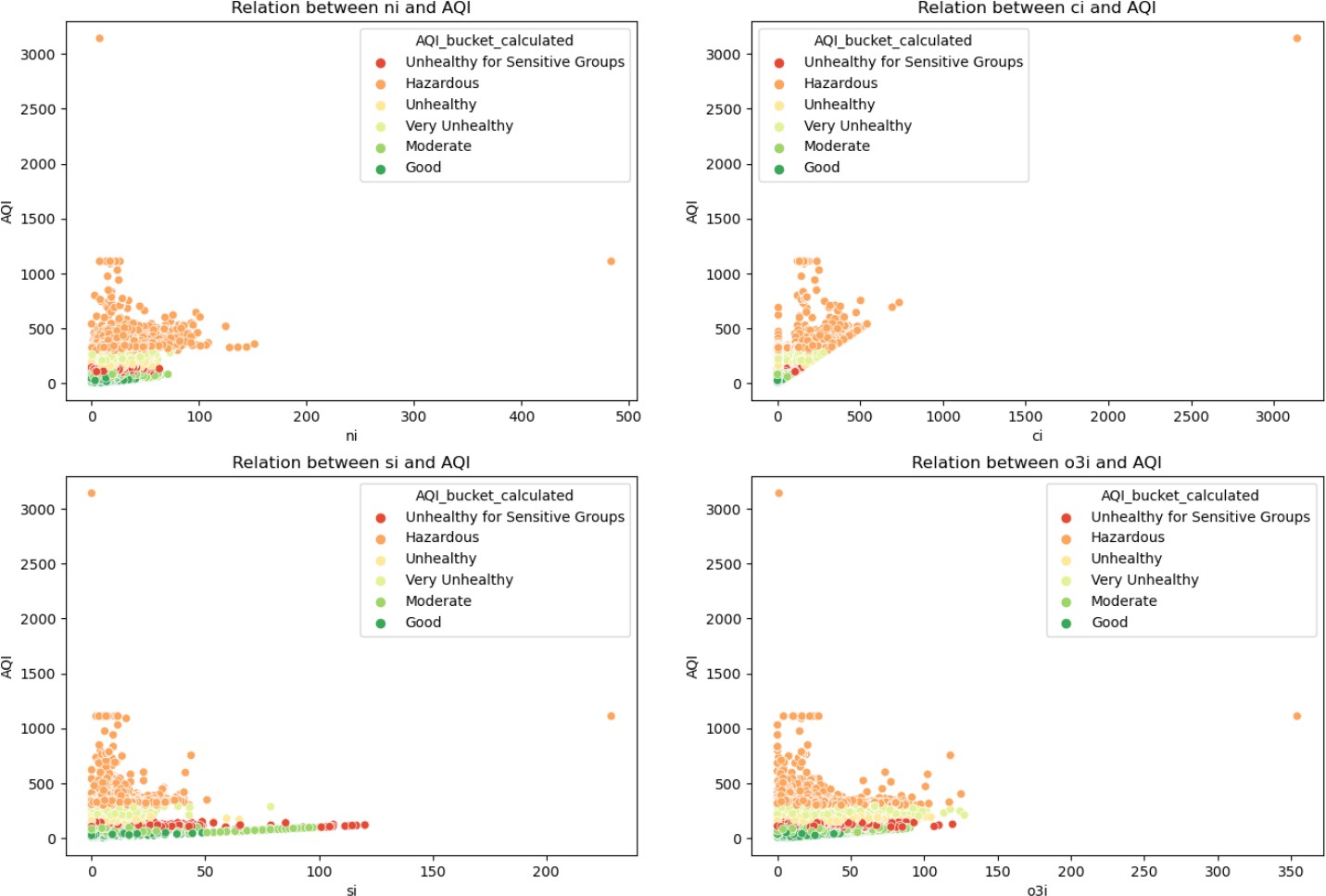


Figure 26: Scatter plot of AQI vs different pollutants.

From the scatterplot matrix, we can observe a positive correlation between the AQI and all variables except O3, which has a negative correlation. The scatterplots also show that as the values of these variables increase, the AQI also tends to increase. The negative correlation between O3 and AQI indicates that higher ozone levels are associated with lower AQI values. This suggests that the factors contributing to higher ozone levels may differ from those contributing to higher levels of other air pollutants, such as particulate matter.

These visualizations suggest that PM2.5, or Particulate Matter of size 2.5mm and PM10 or Particulate Matter of size 10mm have the most significant impact on the AQI for the sampled data. This insight can be helpful while designing targeted pollution control measures.

## I was visualizing the relationship distribution of Pollutants across countries.

The below graphs show how pollutants are distributed across each country:

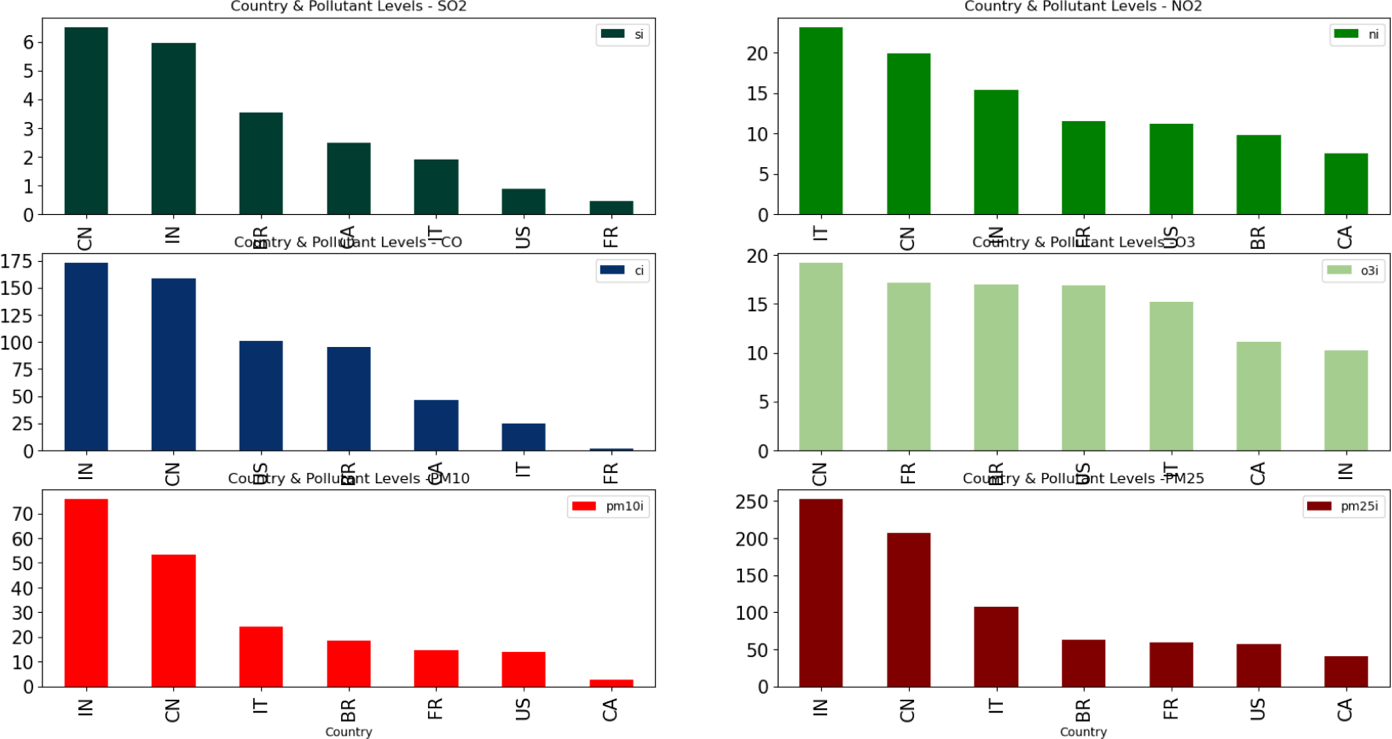


Figure 27: Distribution of pollutants in each country

Observation:

* Over 50 over cent of this pollution comes from industry **which contributes to PM 2.5 and PM10—followed** by 27 percent from vehicles **which contributes to CO.**
* Italy had multiple fires, which contributed to the spike in NO2 levels.
* China has chemical pollutant outlets which impact Ozone and add to SO2 levels the most.

## Guiding Question 4: Analyze the impact of AQI on COVID in each country.

This question investigates whether there is a relationship between AQI and COVID-19 cases in each country during the pandemic. The question seeks to establish whether AQI affected the number of COVID-19 cases in each country, which can help better understand the impact of air pollution on public health during the pandemic. [xxii]

Data Source:

We fetch the COVID data from a CSV file from the Data Source: [iii]. The Data has Columns:

* Date: designates the date of observation of the row's data in YYYY-MM-DD format.
* Country: designates the Country in which the row's data was observed.
* cumulative\_total\_cases: sets the cumulative number of confirmed cases as of the row's date for the row's country.
* daily\_new\_cases: sets the daily new number of confirmed cases on the row's date for the row's country.
* active\_cases: designates the number of active patients (i.e., confirmed cases that still didn't recover nor died) on the row's date for the row's country.
* cumulative\_total\_deaths: designates the cumulative number of confirmed deaths as of the row's date for the row's country.
* daily\_new\_deaths: sets the daily new number of confirmed deaths on the row's date for the row's country.

## Data Wrangling on the data source:

* The 'date' column was converted to DateTime format and separated Month-Year columns for further analysis and visualizations.
* Replace some country names to ensure consistency in the dataset.
* Then, merge the COVID data with AQI data on the country name and date columns using the Pandas merge() function and stores the result in merged\_covid\_aqi\_data.
* The merged dataset contains both COVID and AQI data for the countries and dates where both are available.

## Verifying if there is any correlation between cumulative cases of Covid and AQI levels in the country

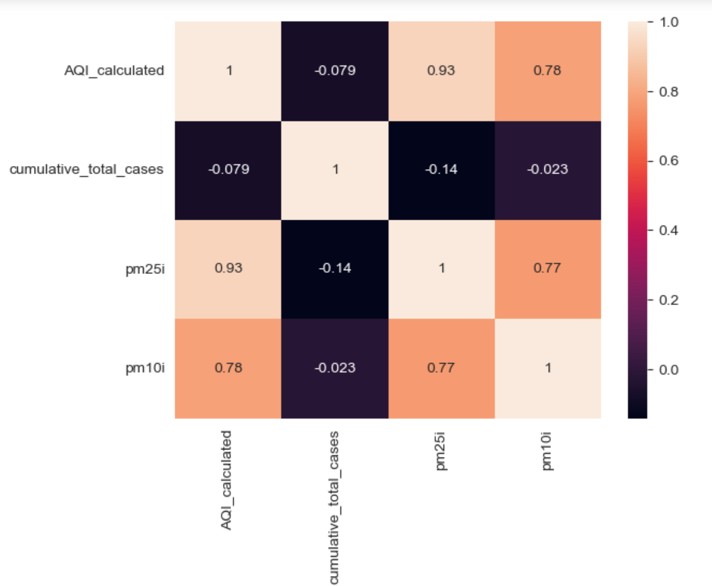


Figure 28: Correlation between AQI level and cumulative covid cases

This correlation suggests a weak negative correlation between AQI levels and cumulative COVID cases. This means that the AQI levels go down, as cumulative COVID cases slightly go down, but the relationship is not strong.

However, correlation does not necessarily imply causation, so it is important to consider other factors influencing AQI levels and COVID cases. Further analysis and interpretation of the data are needed to draw definitive conclusions about the relationship between AQI and COVID.

## Visualizing an animated scatter plot to show the relationship between air quality (measured by AQI):

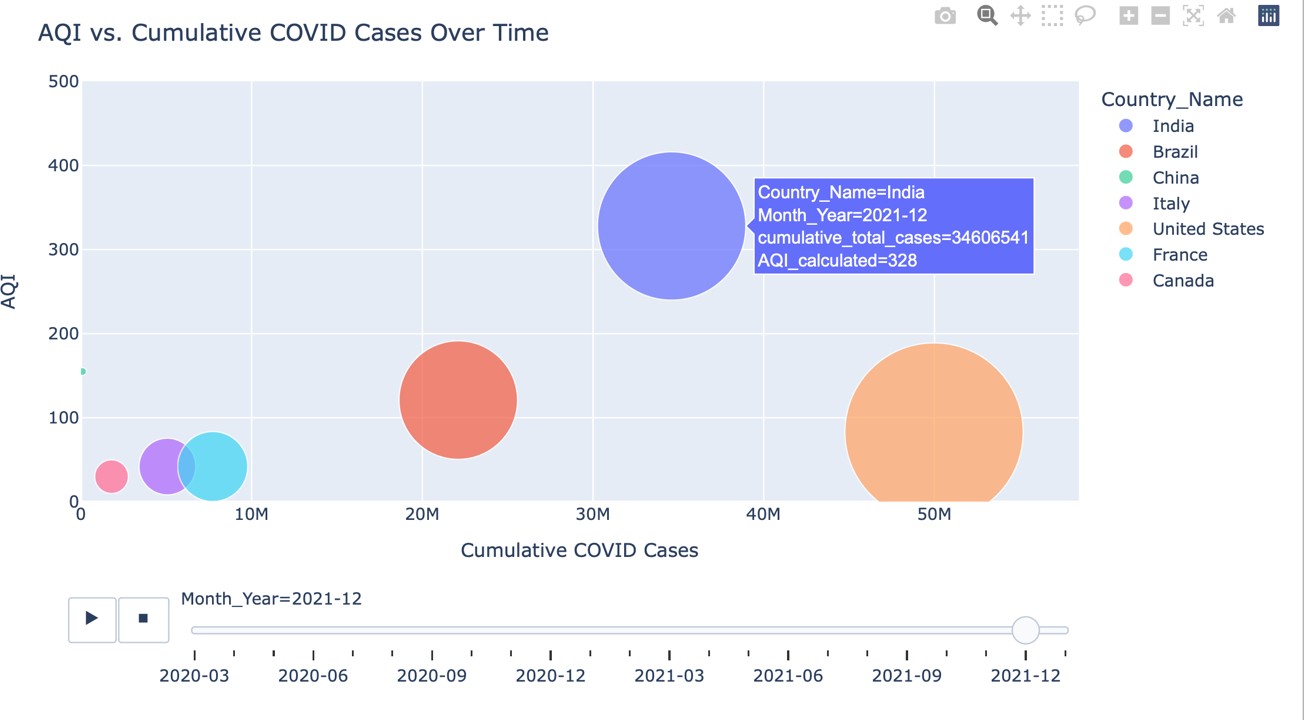
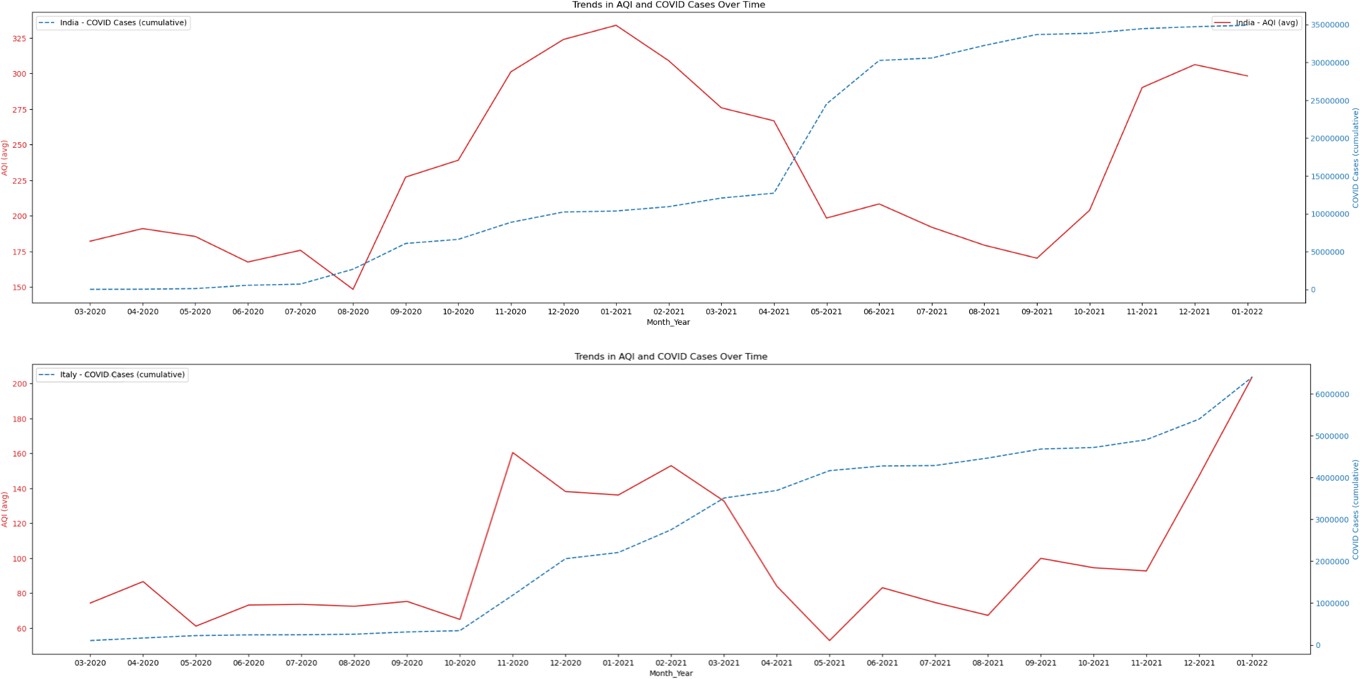


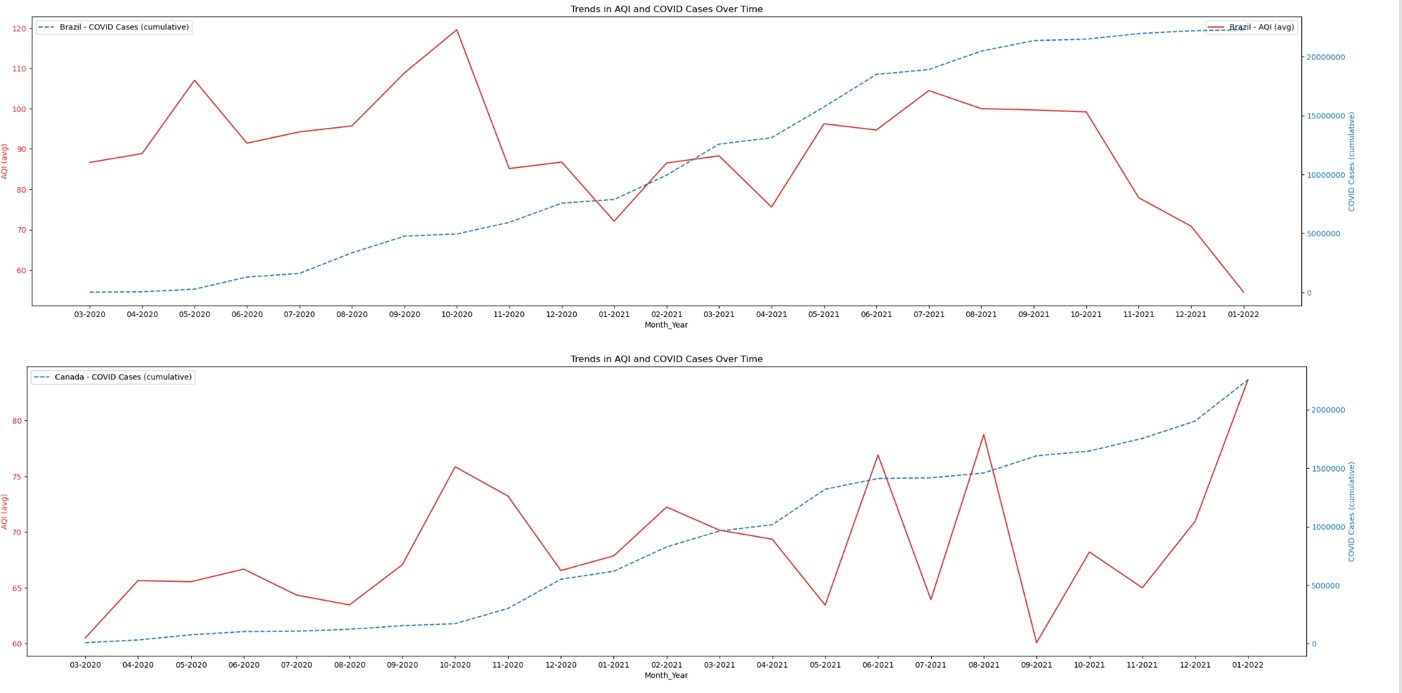
Figure 29: Scatter plot AQI level vs cumulative covid cases over time

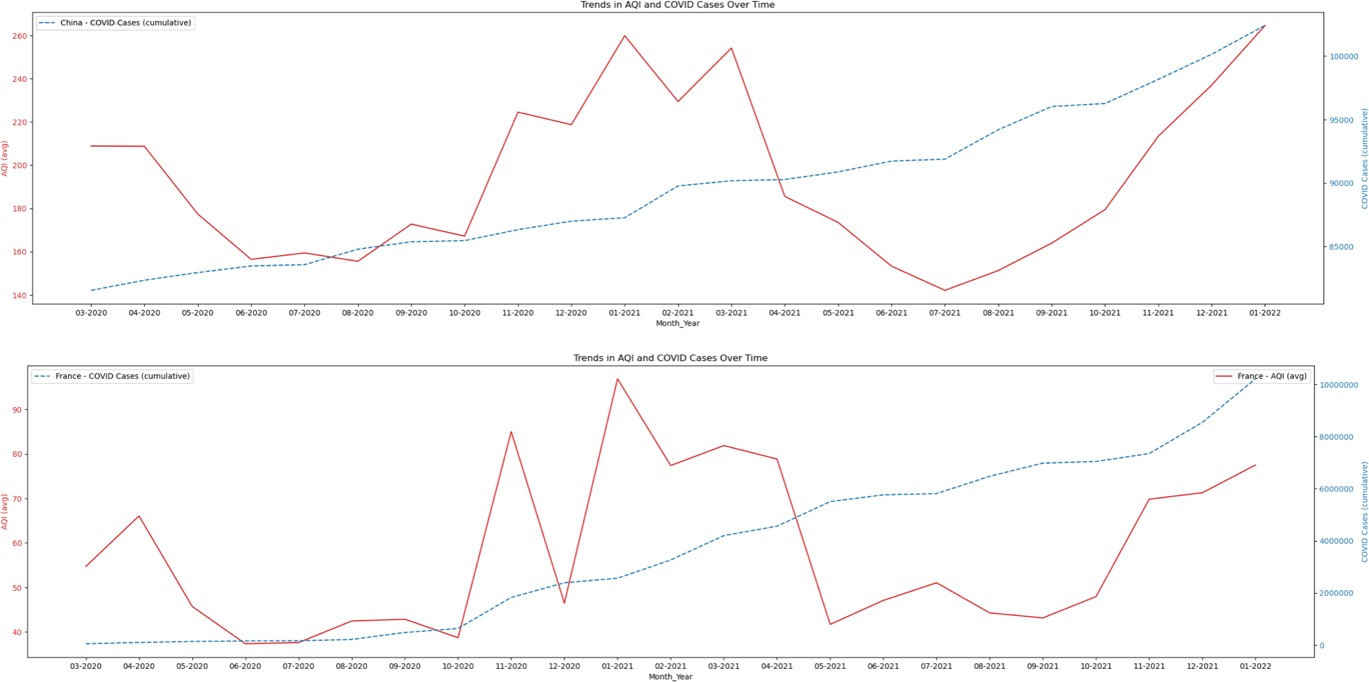
The animated scatter plot above shows the relationship between air quality (measured by AQI) and the cumulative number of COVID-19 cases for various countries over time. Each point represents a country, with the size of the point indicating the number of cumulative COVID-19 cases in that country. The colour of the point represents the country name.

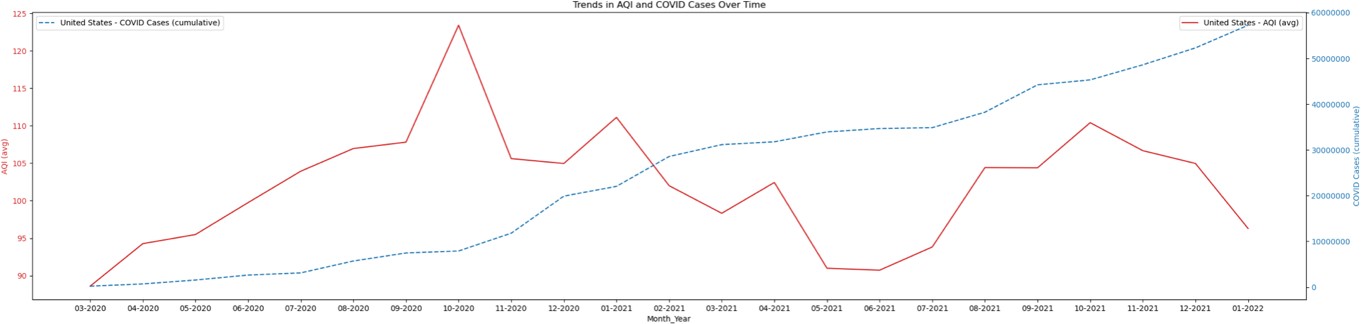
The plot reveals a negative relationship between air quality and the number of COVID-19 cases. Initially, most countries had relatively low AQI values and few COVID-19 issues. However, as time progresses, some countries experience a sharp increase in COVID-19 cases despite having low AQI (USA and Brazil, as seen in Figure 29, while India ranks second in COVID cases despite having the highest AQI.

## Visualizing a line plot to show the trend in air quality (measured by AQI) and Cumulative cases of COVID over time for all countries, the y-axis on the left side shows the AQI level. The y-axis on the right side shows cumulative cases.









The plot suggests a complex relationship between air quality and COVID-19 transmission.

Conclusion

In conclusion, analyzing air quality data from multiple countries has provided valuable insights into the factors affecting air quality and the AQI. We found that the levels of pollutants such as PM2.5, PM10, NO2, SO2, and O3 significantly impact the AQI, with PM2.5 being the most critical factor contributing to poor air quality. We also identified a correlation between the COVID-19 period and air quality, with lockdowns and travel restrictions positively impacting air quality.

The visualizations created for this analysis have clarified the relationships between the variables and the AQI. The animated scatter plot of AQI vs. cumulative COVID cases over time has provided an excellent representation of the relationship between these two variables. This plot clearly shows that there is a relationship between COVID-19 issues and air quality and that air quality has improved during periods of lockdowns and travel restrictions.

Despite being the world's biggest environmental killer, only 1% of global development aid addresses air pollution. Disturbingly, governments gave 20% more in funding to fossil fuel projects which caused the breakdown.[xxiv]. We hope our analysis sheds light on the importance of bringing down the AQI levels.

Overall, this analysis can be helpful for policymakers and environmental organizations to develop targeted measures to improve air quality and reduce pollution levels in the most affected areas. In addition, the insights gained from this analysis can be used to develop policies and strategies to improve air quality, particularly in urban areas most affected by poor air quality.

## Future Scope and Recommendations:

While air pollution may worsen respiratory illnesses and increase the severity of COVID-19 symptoms, other factors, such as population density, public health measures, and socioeconomic status, may also play a role. Therefore, it is necessary to conduct further research to understand better the interrelationship between air quality and COVID-19 transmission, as well as to identify effective strategies for reducing the impact of both. In addition, more extensive investigation, such as controlled experiments or multivariate analysis, will be required to establish a causal relationship.

# LINK TO CODE:

https://github.com/monicacs5830/Data\_Science\_Projects/blob/main/DATA\_601\_Final\_Project\_Report.i pynb

# REFERENCES:

1. World Health Organization. (n.d.). Air Pollution. World Health Organization. Retrieved February 16, 2023, from [https://w](http://www.who.int/health-topics/air-pollution#tab%3Dtab_1)ww.who.i[nt/health-topics/air-pollution#tab=tab\_1](http://www.who.int/health-topics/air-pollution#tab%3Dtab_1)
2. AQI DATA SOURCE: The World Air Quality Index project (no date) Covid-19 Worldwide Air Quality Data, aqicn.org. Available at: <http://aqicn.org/data-platform/covid19/> (Accessed: February 16, 2023).
3. COVID DATA SOURCE: [https://w](http://www.kaggle.com/datasets/josephassaker/covid19-global-)ww.kaggl[e.com/datasets/josephassaker/covid19-global-](http://www.kaggle.com/datasets/josephassaker/covid19-global-) dataset
4. The World Air Quality Index project (no date) Air Quality index scale and colour

legend, aqicn.org. Available at: https://aqicn.org/scale/] (Accessed: February 16, 2023).

1. (no date) EPA. Environmental Protection Agency. Available at: [https://w](http://www.epa.gov/outdoor-air-quality-data/air-quality-index-report)ww.epa.g[ov/outdoor-air-quality-data/air-quality-index-report](http://www.epa.gov/outdoor-air-quality-data/air-quality-index-report) (Accessed: February 16, 2023).
2. Rohanrao (2020) Calculating Aqi (Air Quality Index) tutorial, Kaggle. Kaggle. Available at: [https://w](http://www.kaggle.com/code/rohanrao/calculating-aqi-air-quality-index-tutorial)ww.kaggl[e.com/code/rohanrao/calculating-aqi-air-quality-index-tutorial](http://www.kaggle.com/code/rohanrao/calculating-aqi-air-quality-index-tutorial) (Accessed: February 16, 2023).
3. https://nominatim.org
4. [https://w](http://www.ncei.noaa.gov/access/monitoring/monthly-report/fire/202113)ww.ncei.[noaa.gov/access/monitoring/monthly-report/fire/202113](http://www.ncei.noaa.gov/access/monitoring/monthly-report/fire/202113)
5. https://plotly.com
6. https://en.wikipedia.org/wiki/2021\_Italy\_wildfires
7. [https://w](http://www.geeksforgeeks.org/python-imdbpy-getting-the-country-codes-of-the-series/)ww.geek[sforg](http://www.geeksforgeeks.org/python-imdbpy-getting-the-country-codes-of-the-series/)e[eks.org/python-imdbpy-getting-the-country-codes-of-the-series/](http://www.geeksforgeeks.org/python-imdbpy-getting-the-country-codes-of-the-series/)
8. [https://w](http://www.iqair.com/world-air-quality-report)ww.iqai[r.com/worl](http://www.iqair.com/world-air-quality-report)d[-air-quality-report](http://www.iqair.com/world-air-quality-report)
9. https://en.wikipedia.org/wiki/Timeline\_of\_the\_COVID-19\_pandemic
10. [https://w](http://www.geeksforgeeks.org/get-the-city-state-and-country-names-from-latitude-and-)ww.geek[sforg](http://www.geeksforgeeks.org/get-the-city-state-and-country-names-from-latitude-and-)e[eks.org/get-the-city-state-and-country-names-from-latitude-and-](http://www.geeksforgeeks.org/get-the-city-state-and-country-names-from-latitude-and-) longitude-using-python/
11. https://en.wikipedia.org/wiki/Air\_pollution\_in\_India
12. [https://w](http://www.iqair.com/newsroom/no-let-up-in-air-pollution-this-diwali-delhi-air-quality-)ww.iqai[r.com/ne](http://www.iqair.com/newsroom/no-let-up-in-air-pollution-this-diwali-delhi-air-quality-)w[sroom/no-let-up-in-air-pollution-this-diwali-delhi-air-quality-](http://www.iqair.com/newsroom/no-let-up-in-air-pollution-this-diwali-delhi-air-quality-) worsens
13. https://impakter.com/air-pollution-leading-environmental-killer/
14. [https://w](http://www.bbc.com/news/science-environment-57149747)ww.bbc.[com/ne](http://www.bbc.com/news/science-environment-57149747)w[s/science-environment-57149747](http://www.bbc.com/news/science-environment-57149747)
15. https://en.wikipedia.org/wiki/Air\_pollution\_in\_India
16. [https://w](http://www.thelancet.com/journals/lansea/article/PIIS2772-3682(22)00097-X/fulltext)ww.thel[ancet.com/journals/lansea/articl](http://www.thelancet.com/journals/lansea/article/PIIS2772-3682(22)00097-X/fulltext)e[/PIIS2772-3682(22)00097-X/fulltext](http://www.thelancet.com/journals/lansea/article/PIIS2772-3682(22)00097-X/fulltext)
17. [https://w](http://www.hawaii.edu/news/2021/08/23/pacific-islands-wildfires-the-hill-editorial/)ww.hawai[i.edu/news/2021/08/23/pacific-islands-wildfires-the-hill-editorial/](http://www.hawaii.edu/news/2021/08/23/pacific-islands-wildfires-the-hill-editorial/)]
18. https://projects.iq.harvard.edu/covid-pm/home
19. [https://w](http://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-)ww.who.i[nt/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-](http://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-) health
20. https://impakter.com/air-pollution-leading-environmental-killer/
21. https://indianexpress.com/article/cities/pune/new-study-breathing-polluted-air-can-make- children-fat-increase-risk-of-developing-asthma-7482481/
22. Sullivan, Frank, et al. “The Effect of Air-Pollution and Weather Exposure on Mortality and Hospital Admission and Implications for Further Research: A Systematic Scoping Review.” PLoS One, vol. 15, no. 10, Public Library of Science, Oct. 2020, p. e0241415.
23. What Are Health Issues Associated with Rodent Infestations? https://rodentsolutioninc.com/blog/what-health-issues-are-associated-with-rodent- infestations/
24. Exposure to high levels of air pollution can make children overweight .... https://indianexpress.com/article/cities/pune/new-study-breathing-polluted-air-can-make- children-fat-increase-risk-of-developing-asthma-7482481/
25. Akwah, Emmanuela. “Air Pollution Exposure about Guard Duty at Tidworth Camp: A Cross- Sectional Study.” PLoS One, vol. 16, no. 9, Public Library of Science, Sept. 2021, p. e0258070.
26. Lednicky, John, et al. “Earliest Detection to Date of SARS-CoV-2 in Florida: Identification Together with Influenza Virus on the Main Entry Door of a University Building, February 2020.” PLoS One, vol. 16, no. 1, Public Library of Science, Jan. 2021, p. e0245352.
27. Ansari, Rabiya, and Jennifer Landin. “Coverage of Climate Change in Introductory Biology Textbooks, 1970–2019.” PLoS One, vol. 17, no. 12, Public Library of Science, Dec. 2022, p. e0278532.
28. For diagnostic hemoglobin A1c testing, many expert organizations cite .... https://uwpqzfj.erdosell.pl/hba1c-normal-degeri.html.
29. Create Amazing Animated Graphs in Python with this 2-Liner Code. [https://w](http://www.analyticsvidhya.com/blog/2021/04/animated-bar-graph-data-science-project/)ww.analy[ticsvidhya.com/blog/2021/04/animated-bar-graph-data-science-project/](http://www.analyticsvidhya.com/blog/2021/04/animated-bar-graph-data-science-project/)
30. Correlation Coefficients: Positive, Negative, & Zero - Investopedia. [https://w](http://www.investopedia.com/ask/answers/032515/what-does-it-mean-if-correlation-)ww.inves[topedia.com/ask/answers/032515/what-does-it-mean-if-correlation-](http://www.investopedia.com/ask/answers/032515/what-does-it-mean-if-correlation-) coefficient-positive-negative-or-zero.asp
31. Nolch, Guy. “Newborn Thyroid Activity Linked to Academic Struggles.” Australasian Science, vol. 37, no. 7, Control Publications Pty Ltd, Sept. 2016, p. 7.
32. Pandemic lockdowns improved air quality in 84% of countries ... - CNN. [https://w](http://www.cnn.com/2021/03/16/health/world-air-quality-report-intl-hnk-scn/index.html)ww.cnn.[com/2](http://www.cnn.com/2021/03/16/health/world-air-quality-report-intl-hnk-scn/index.html)0[21/03/16/health/world-air-quality-report-intl-hnk-scn/index.html](http://www.cnn.com/2021/03/16/health/world-air-quality-report-intl-hnk-scn/index.html)
33. Wunsch, Hannah, et al. “The Impact of the Organization of High-Dependency Care on Acute Hospital Mortality and Patient Flow for Critically Ill Patients.” American Journal of Respiratory and Critical Care Medicine, vol. 191, no. 2, American Thoracic Society, Jan. 2015,

p. 186.

1. Exposure to high levels of air pollution can make children overweight .... https://indianexpress.com/article/cities/pune/new-study-breathing-polluted-air-can-make- children-fat-increase-risk-of-developing-asthma-7482481/