National College of Ireland

Project Submission Sheet – 2020/2021

Student Name:	Omkar Ratnoji Tawade		
Student ID:	19232136		
Programme:	MSCDAD_A	**Year:**	2020-21
Module:	Database and Analytics Programming		
Lecturer:	Anu Sahni		
Submission Due Date:	08-01-2021		
Project Title:	Terminal Assignment		
Word Count:	5333 words		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	Omkar Ratnoji Tawade
Date:	08-01-2021

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

National College of Ireland

Project Submission Sheet – 2020/2021

Student Name:	Michael Dunne		
Student ID:	15420892		
Programme:	MSCDAD_A	**Year:**	2020-21
Module:	Database and Analytics Programming		
Lecturer:	Anu Sahni		
Submission Due Date:	08-01-2021		
Project Title:	Terminal Assignment		
Word Count:	5333 words		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	Michael Dunne
Date:	08-01-2021

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

National College of Ireland

Project Submission Sheet – 2020/2021

Student Name:	Vivek Kumar		
Student ID:	19201885		
Programme:	MSCDAD_A	**Year:**	2020-21
Module:	Database and Analytics Programming		
Lecturer:	Anu Sahni		
Submission Due Date:	08-01-2021		
Project Title:	Terminal Assignment		
Word Count:	5333 words		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	Vivek Kumar
Date:	08-01-2021

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

National College of Ireland

National College of Ireland

Project Submission Sheet – 2020/2021

Student Name:	Suraj Karyamapudi		
Student ID:	19232632		
Programme:	MSCDAD_A	**Year:**	2020-21
Module:	Database and Analytics Programming		
Lecturer:	Anu Sahni		
Submission Due Date:	08-01-2021		
Project Title:	Terminal Assignment		
Word Count:	5333 words		

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
ALL internet material must be referenced in the references section. Students are encouraged to use the Harvard Referencing Standard supplied by the Library. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action. Students may be required to undergo a viva (oral examination) if there is suspicion about the validity of their submitted work.

Signature:	Suraj Karyamapudi
Date:	08-01-2021

PLEASE READ THE FOLLOWING INSTRUCTIONS:

1. Please attach a completed copy of this sheet to each project (including multiple copies).
2. Projects should be submitted to your Programme Coordinator.
3. **You must ensure that you retain a HARD COPY of ALL projects**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. Please do not bind projects or place in covers unless specifically requested.
4. You must ensure that all projects are submitted to your Programme Coordinator on or before the required submission date. **Late submissions will incur penalties.**
5. All projects must be submitted and passed in order to successfully complete the year. **Any project/assignment not submitted will be marked as a fail.**

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Database and Analytics Programming

AIR QUALITY INDEX ANALYSIS

Michael Dunne
Data Analytics
National College of Ireland
Dublin, Ireland
x15420892@student.ncirl.ie

Omkar Tawade
Data Analytics
National College of Ireland
Dublin, Ireland
x19232136@student.ncirl.ie

Suraj Karyamapudi
Data Analytics
National College of Ireland
Dublin, Ireland
x19232632@student.ncirl.ie

Vivek Kumar
Data Analytics
National College of Ireland
Dublin, Ireland
x19201885@student.ncirl.ie

***Abstract*—This document is a report done for our group project, group D for the module of Database and Analytics Programming for Semester 1 at National College of Ireland. The world we live in has faced drastic changes over the past decade due to the overwhelming rate that climate change has affected the earth and the people who inhabit this planet. Major companies seem to forget the long-term effects that manufacturing, and the burning of fuel can do for the sole purpose of gaining multi-millions in profit. One of the biggest ways in which we can see the effects of us mistreating our planet is by analyzing the air pollution levels in major cities. The goal of this project is to highlight the growing concern climate activists have by illustrating the rise in the air pollution levels in major cities by parsing the semi-structured dataset JSON/XML using Python to Mongo DB and then storing the structured data to Postgres for analytics purpose.**

Keywords *- Air Pollution, JSON, XML, Air Quality Index, Python, MongoDB, Postgres.*

I. INTRODUCTION

Our aim of this project is to identify and illustrate the air quality and air pollution levels in the populated regions of India or surrounding southeast Asia using Python as our programming language to parse JSON/XML data which consists of air pollutants such as SO2, NO2, RSPM_PM10 and SPM along with their values. We will store this semi-structured data in Mongo DB and then we will fetch this data back to python to convert data as structured data. We will then store this structured data in the Postgres database and with the help of Python, we will fetch data from Postgres and will visualize our structured data. In the world we live in now, air pollution ranks as one of the main growing concerns for countries and is in fact ranks sixth as the most dangerous killer in South Asia. We want to see from our analysis ourselves the majorly affected areas that we can confirm need immediate action from government officials and manufacturing companies operating in that region. It is integral that we start to use data analysis more often to shape a brighter future for the next generations by looking at the data such as this to develop strategies to combat air pollution and climate change. We hope to see if data analysis can paint a clear picture of what we can physically see.

Figure 1. Shows the air pollution affecting people in today's world.

For this analysis to work, we need to know exactly what we are looking at in terms of the figures recorded for air pollution in these cities. Air quality is measured by eight pollutants which are as follows PM10, PM2.5, Ozone, Sulphur dioxide, nitrogen dioxide, carbon monoxide, lead, and ammonia. These are all the components measured to grade the overall air quality (AQI). This is then ranked by the range in total ranging from good to severe, the figure below shows a table measuring the air quality.

AQI Category (Range)	PM_{10} 24-hr	$PM_{2.5}$ 24-hr	NO_2 24-hr	O_3 8-hr	CO 8-hr (mg/m^3)	SO_2 24-hr	NH_3 24-hr	Pb 24-hr
Good (0-50)	0-50	0-30	0-40	0-50	0-1.0	0-40	0-200	0-0.5
Satisfactory (51-100)	51-100	31-60	41-80	51-100	1.1-2.0	41-80	201-400	0.6 –1.0
Moderate (101-200)	101-250	61-90	81-180	101-168	2.1- 10	81-380	401-800	1.1-2.0
Poor (201-300)	251-350	91-120	181-280	169-208	10.1-17	381-800	801-1200	2.1-3.0
Very poor (301-400)	351-430	121-250	281-400	209-748*	17.1-34	801-1600	1201-1800	3.1-3.5
Severe (401-500)	430 +	250+	400+	748+*	34+	1600+	1800+	3.5+

Figure 2. Air Quality Index Table

It must be taken into account that all the eight pollutants might not be taken at all the test locations, this is the problem of course if the data is not available then how can a decision or an analysis be implemented. For this at least three pollutants out of the eight must be calculated, PM2.5 and PM10 are the most essential of the pollutants needed. The minimum time span in which data can be considered is 16 hours.

Over this report, we hope to answer the question, what is the worst and most affected cities when it comes to air quality and air pollution? Can we see a significant difference in the type of emissions in each city? Can we use python and other software as a means to store and work on this data and also can we show in a clear and concise way an illustration of how bad these rates out to the average person let's say a politician who could potentially work locally in one of these cities that should be very concerned if provided with this analysis?

II. Related Work

In their research, Humaib Nasir1, Kirti Goyal2, and Dolonchapa Prabhakar [1] propose research in which they came up with details about the condition of air quality in different cities of India and what are the reasons and effects of air pollution. In their research, they find out that from the last few years the increase in air pollution in India is alarming because air pollution contains harmful pollutants that are also used by people for breathing. An increase in industries, manufacturing units, and no. of vehicles has resulted in too much collection of pollutants in the air which makes air pollution a national emergency in many cities of India.

In their research, Kanchan, Pramila Goyal and Amit Kumar Gorail [2] found out that no universally accepted method can determine the true level of AQI (Air Quality Index) So they reviewed all the big air quality measurements methods invented and used in the world. They conclude that when they apply AQI on a dataset of air quality, they found big differences in index values and pollutants.

Snehal Sirsikar, Priya Karemore presented research [3] in which they highlight some technologies which are used for air pollution monitoring and how effective these technologies are. They conclude that monitoring air pollution with wireless sensors has several benefits over the traditional environment.

In another research, Georgina Santos, Megan Mathias and Sarah Quarmby [4] gives a review on air quality strategies and technologies. It focused on road transport in cities. They conclude that air pollution has a noteworthy effect on respiratory and cardiovascular diseases and mortality. They also conclude that well-made, powerful policy inventions can greatly minimize air pollution.

Jinfang Sun, Zhichao Zhou, Jing Huang, and Guoxing Li presented research [5] in which they investigated the association between air pollution and children. In their research, they summarize that household air pollution exposure and children's lung function should be paid more attention.

Gowtham Sarella, and Dr. Anjali. K. Khambete [6] accounts for the study of atmospheric air in Vapi city's Air Quality Index. SO2, NO2, PM10, and PM2.5 were selected for 2013 (March) to 2014 (February) at 4 different sites in Vapi. They conclude that PM10 is a critical pollutant at all selected locations.

Envis Centre CPCB presented research [7] in which they conclude that the air of city Delhi is very polluted in which 72% of pollution is caused by vehicles only. This research also concludes that the concentration of SO2 in Delhi is under control and to minimize the NO2 pollution NO, O3 and hydrocarbons need to be targeted.

In another research Ackchai Sirikijpanichkul, Madhumita Iyengar, Professor Luis Ferreira [8] provide an analysis of the most appropriate methods for analyzing the air quality impact. It deals with the main problems coming in recognizing and classify air quality effects; the methodologies that are adopted for recognizing the effects of air pollution on health; and the evaluations proposed by studies in the US, Australia, and Europe.

Badri Nath, Srinivas Devarakonda, Ruilin Liu, Hongzhang Liu, Liviu Iftode, Parveen Sevusu [9] introduce a vehicle-based mobile approach for quantifying air quality in sync. They proposed two models that are cheap and one of the models can be brought into play on public transportation and another one is a personal sensing device.

In another research, M. Lopes, C. Borrego1, P. Cascão, J. H. Amorim, N. Chrysoulakis, J. Martins, V. Rodrigues, A.I. Miranda & R. Tavares [10] assess the influence on air quality due to different city PA, URBAIR was applied to selected areas in Athens of Greece, Helsinki of Finland, and Gliwice of Poland, to evaluate traffic-related outflow of gases and bring on pollutant concentration of different air pollutants, in an hourly basis for the entire year of 2008.

III. Methodology

We have selected air quality index data for this group project. The measurement of air quality is based on eight pollutants namely, Particulate Matter (PM10), Particulate Matter (PM2.5), Nitrogen Dioxide (NO2), Sulphur Dioxide (SO2), Carbon Monoxide (CO), Ozone (O3), Ammonia (NH3), and Lead (Pb). We have downloaded the JSON/XML of different states of India [11]. Central Pollution Control Board collected the data from 2011 and 2015 and based on this they started a campaign of National Clean Air Program. We have downloaded the data of cities such as Assam, Gujarat, Kerala, and West Bengal. We will analyze which state needs the implementation of such a program with help of plots and figures. The first process of our project was to download JSON/XML data of the respective state and we need to upload this data in JSON format on mongo dB. But JSON file we downloaded was not having an ideal JSON structured, so we needed to first structure our data as ideal JSON data which would get properly uploaded on the mongo dB. Whereas XML data needed to be converted into JSON format data before uploading the data on mongo dB.


```
{"fields":[{"id":"a","label":"Stn Code","type":"string"},{"id":"b","label":"Sampling
Date","type":"string"},{"id":"c","label":"State","type":"string"},
{"id":"d","label":"City\/Town\/Village\/Area","type":"string"},{"id":"e","label":"Location
of Monitoring Station","type":"string"},{"id":"f","label":"Agency","type":"string"},
{"id":"g","label":"Type of Location","type":"string"},
{"id":"h","label":"SO2","type":"string"},{"id":"i","label":"NO2","type":"string"},
{"id":"j","label":"RSPM\/PM10","type":"string"},
{"id":"k","label":"SPM","type":"string"}],"data":[["386","1\/1\/2010","West
Bengal","Asansol","Asansol Municipal Corporation, Asansol","West Bengal State Pollution
Control Board","Industrial Area","10","80.1","252","525"],["386","1\/3\/2010","West
Bengal","Asansol","Asansol Municipal Corporation, Asansol","West Bengal State Pollution
Control Board","Industrial Area","10.6","84.1","248","516"],["386","1\/8\/2010","West
```


Figure 3. West Bengal Air Quality Index JSON data.


```
{"fields":[{"id":"a","label":"Stn Code","type":"string"},{"id":"b","label":"Sampling
Date","type":"string"},{"id":"c","label":"State","type":"string"},
{"id":"d","label":"City\/Town\/Village\/Area","type":"string"},{"id":"e","label":"Location
of Monitoring Station","type":"string"},{"id":"f","label":"Agency","type":"string"},
{"id":"g","label":"Type of Location","type":"string"},
{"id":"h","label":"SO2","type":"string"},{"id":"i","label":"NO2","type":"string"},
{"id":"j","label":"RSPM\/PM10","type":"string"},
{"id":"k","label":"SPM","type":"string"}],"data":[["386","1\/1\/2010","West
Bengal","Asansol","Asansol Municipal Corporation, Asansol","West Bengal State Pollution
Control Board","Industrial Area","10","80.1","252","525"],["386","1\/3\/2010","West
Bengal","Asansol","Asansol Municipal Corporation, Asansol","West Bengal State Pollution
Control Board","Industrial Area","10.6","84.1","248","516"],["386","1\/8\/2010","West
```


Figure 3 shows that data is not stored as a key-value pair. The keys are stored in one array and data is stored in a different

array. So, with the help of python, we structured this data as key-value pair which will help us to upload the data on mongo dB correctly. We parsed the XML file with help of python, and it was converted into JSON data, and the structure of data was key-value pair but here the problem was that each record or data was stored in different dictionaries. So, we performed some data processing steps and converted the data into an ideal JSON format.


```
{'results': [{'STN_CODE': '542', 'SAMPLING_DATE': '010214', 'STATE': 'Assam',
'CITYTOWNVILLAGEAREA': 'Bongaigaon', 'LOCATION_OF_MONITORING_STATION': 'Campu
s of Oil India Ltd. PS6 Bongaigaon Chirang', 'AGENCY': 'Assam State Pollution
Control Board', 'TYPE_OF_LOCATION': 'Residential Rural and other Areas', 'SO
2': '6', 'NO2': '13', 'RSPMPM10': '56', 'PM_2_5': 'NA'}, {'STN_CODE': '542',
'SAMPLING_DATE': '010414', 'STATE': 'Assam', 'CITYTOWNVILLAGEAREA': 'Bongaiga
on', 'LOCATION_OF_MONITORING_STATION': 'Campus of Oil India Ltd. PS6 Bongaiga
```


Figure 4. Assam Air quality Index JSON data after processing.

To upload this data on mongo dB, we have created a remote machine on the virtual box. We have installed Mongo DB on that machine. Python is a very flexible and multipurpose language because it has helped us to migrate and communicate with the data stored in databases such as mongo dB, Postgres, and as well as for analytics and parsing purposes. So, using python we migrated the data to mongo dB successfully.

Figure 5. JSON data stored in mongo dB on the remote machine.

Using the collection name, we can access the data inside the database. Now, we will import this semi-structured data back to python. Before proceeding to the analysis, we will need to structure our data. In our case, "sampling_date" needed to be cleaned. Also, we dropped a few columns because we believed that they will not be helpful in our analysis. While importing data from mongo dB into python we have stored data in the dataframe. So, after executing all necessary data processing steps we have saved the dataframe as a CSV file.

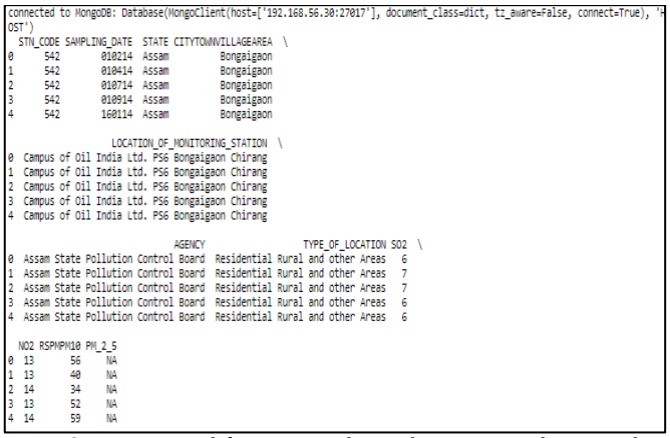


Figure 6. Data imported from mongo dB on the remote machine into the python.

A	B	C	D	E	F	G	H
STN_COD	SAMPLING_DATE	STATE	CITYTOWNVILLAGEAREA	LOCATION_OF_MONITORING_STATION	SO2	NO2	RSPM/PM10
542	1/2/2014	Assam	Bongaigaon	Campus of Oil India Ltd. PS6 Bongaigaon (	6	13	56
542	1/4/2014	Assam	Bongaigaon	Campus of Oil India Ltd. PS6 Bongaigaon (	7	13	40
542	1/7/2014	Assam	Bongaigaon	Campus of Oil India Ltd. PS6 Bongaigaon (	7	14	34
542	1/9/2014	Assam	Bongaigaon	Campus of Oil India Ltd. PS6 Bongaigaon (	6	13	52
542	16/01/2014	Assam	Bongaigaon	Campus of Oil India Ltd. PS6 Bongaigaon (	6	14	59

Figure 7. Data stored in CSV after processing.

After this process, we had to export clubbed data from all four-CSV files produced by each member into the Postgres. We have used python to club the data from all four CSV's. We just imported the data from all four CSV's into different dataframe and then appended the dataframe one by one. We stored this merged data as a CSV file for exporting data into Postgres. We have installed Postgres on our remote machine. The advantage of storing the data on the remote machine is reducing data redundancy, data integrity, data security, backup and recovery, and data consistency. The first step is to create the database in Postgres. After establishing the connection and creating the database, we will need to write a query to create a table in the dataset.

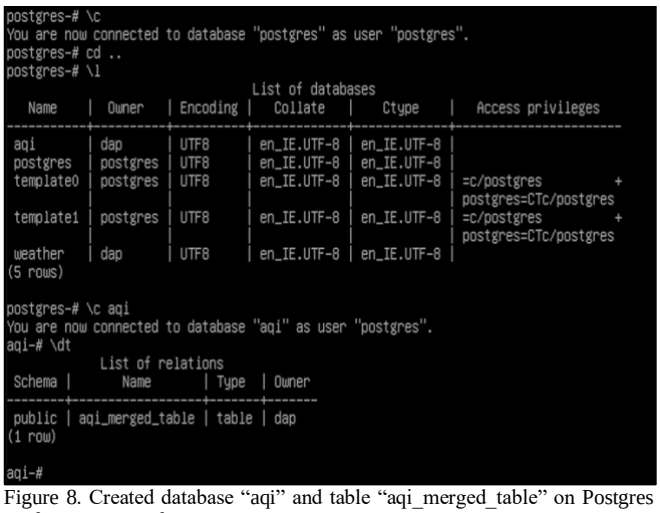
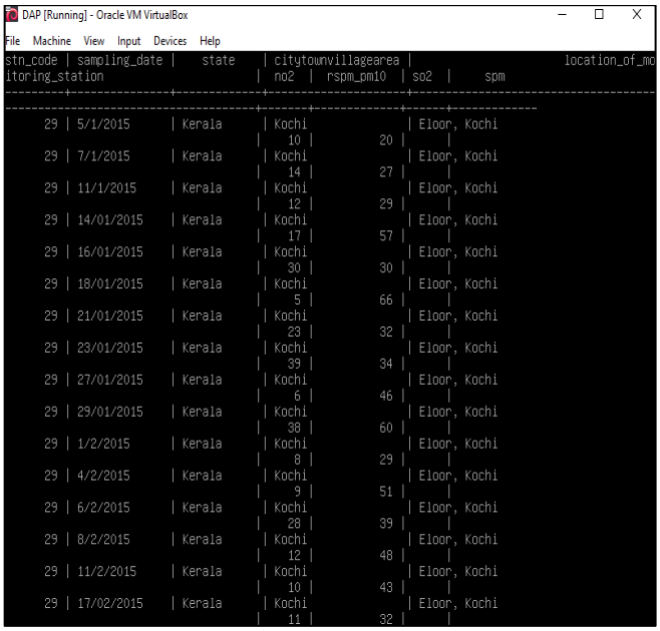


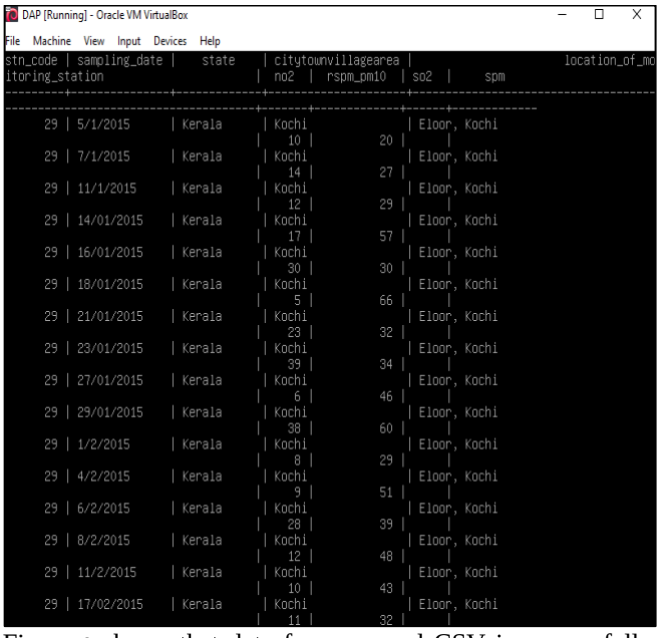
Figure 8. Created database "aqi" and table "aqi_merged_table" on Postgres on the remote machine.

After creating the table, we will export the data stored in our merged CSV to the Postgres table using the PostgreSQL query. Python also helped to communicate with the Postgres database using PostgreSQL queries.



stn_code	sampling_date	state	citytounvillagearea			location_of_monitoring_station
			no2	rspm_pm10	so2	spm
29	5/1/2015	Kerala	Kochi			Eloor, Kochi
			10	20		
29	7/1/2015	Kerala	Kochi			Eloor, Kochi
			14	27		
29	11/1/2015	Kerala	Kochi			Eloor, Kochi
			12	29		
29	14/01/2015	Kerala	Kochi			Eloor, Kochi
			17	57		
29	16/01/2015	Kerala	Kochi			Eloor, Kochi
			30	30		
29	18/01/2015	Kerala	Kochi			Eloor, Kochi
			5	66		
29	21/01/2015	Kerala	Kochi			Eloor, Kochi
			23	32		
29	23/01/2015	Kerala	Kochi			Eloor, Kochi
			39	34		
29	27/01/2015	Kerala	Kochi			Eloor, Kochi
			6	46		
29	29/01/2015	Kerala	Kochi			Eloor, Kochi
			38	60		
29	1/2/2015	Kerala	Kochi			Eloor, Kochi
			8	29		
29	4/2/2015	Kerala	Kochi			Eloor, Kochi
			9	51		
29	6/2/2015	Kerala	Kochi			Eloor, Kochi
			28	39		
29	8/2/2015	Kerala	Kochi			Eloor, Kochi
			12	48		
29	11/2/2015	Kerala	Kochi			Eloor, Kochi
			10	43		
29	17/02/2015	Kerala	Kochi			Eloor, Kochi
			11	32		

Figure 9. Data exported in aqi_merged_table in Postgres on the remote machine using PostgreSQL query written in python.



stn_code	sampling_date	state	citytounvillagearea			location_of_monitoring_station
			no2	rspm_pm10	so2	spm
29	5/1/2015	Kerala	Kochi			Eloor, Kochi
			10	20		
29	7/1/2015	Kerala	Kochi			Eloor, Kochi
			14	27		
29	11/1/2015	Kerala	Kochi			Eloor, Kochi
			12	29		
29	14/01/2015	Kerala	Kochi			Eloor, Kochi
			17	57		
29	16/01/2015	Kerala	Kochi			Eloor, Kochi
			30	30		
29	18/01/2015	Kerala	Kochi			Eloor, Kochi
			5	66		
29	21/01/2015	Kerala	Kochi			Eloor, Kochi
			23	32		
29	23/01/2015	Kerala	Kochi			Eloor, Kochi
			39	34		
29	27/01/2015	Kerala	Kochi			Eloor, Kochi
			6	46		
29	29/01/2015	Kerala	Kochi			Eloor, Kochi
			38	60		
29	1/2/2015	Kerala	Kochi			Eloor, Kochi
			8	29		
29	4/2/2015	Kerala	Kochi			Eloor, Kochi
			9	51		
29	6/2/2015	Kerala	Kochi			Eloor, Kochi
			28	39		
29	8/2/2015	Kerala	Kochi			Eloor, Kochi
			12	48		
29	11/2/2015	Kerala	Kochi			Eloor, Kochi
			10	43		
29	17/02/2015	Kerala	Kochi			Eloor, Kochi
			11	32		

Figure 9 shows that data from merged CSV is successfully exported to Postgres on the remote machine. We have executed a few commands of Postgres shell to verify whether data is stored correctly or not. After this process, we can start with our visualization process. To begin with the visualization process, we will need to fetch the data from Postgres using PostgreSQL query. Again, python helped us to write a PostgreSQL query which is sent to Postgres to fetch all data in the python. The idea behind the whole process is that we can easily update the data in the database and also, we can easily retrieve the structured data for analysis purposes. For visualizations, we have used a few visualization libraries such as matplotlib, plotly, and seaborn.

IV. Results

With the data we procured and the analysis we carried upon it, we were able to produce several plots in accordance with respective states and places, and each of them is explained one after the other in detail.

A. Nitrogen dioxide (NO_2):

The reason for Nitrogen dioxide emission into the environment is from automobiles and coal mining. By comparing all the graphs carefully we can find the trend that although the other states are having a rise and fall in the values of NO_2, Gujarat and Assam are not only having low values but also almost uniform in a year, this is because the majority of NO_2 from these states is generated from their power plants only and in the states of Kerala and West Bengal there is a huge release of NO_2 because they have an enormous amount of small and large power projects which keep generating all over the year and also the vehicle's exhaust adds to it resulting in the kind of pattern we observe.

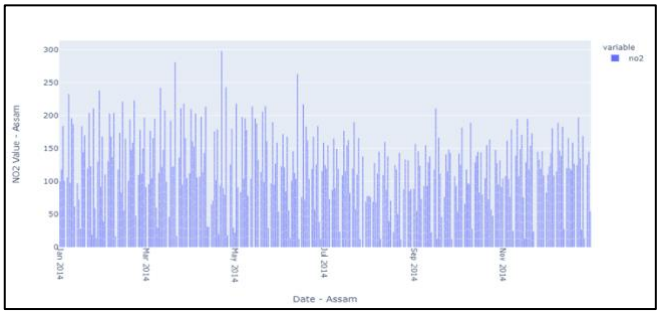


Figure 10. NO_2 value vs Date – Assam.

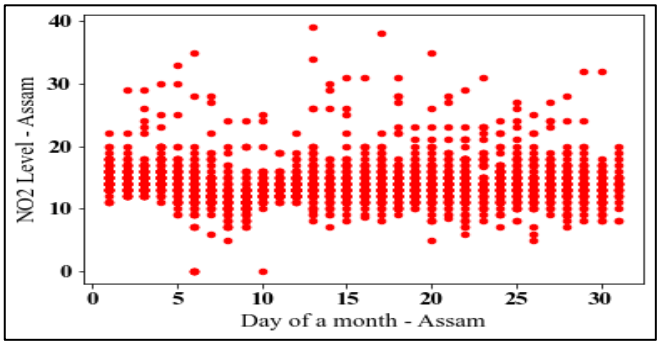


Figure 11. NO_2 level vs Day of a month – Assam.

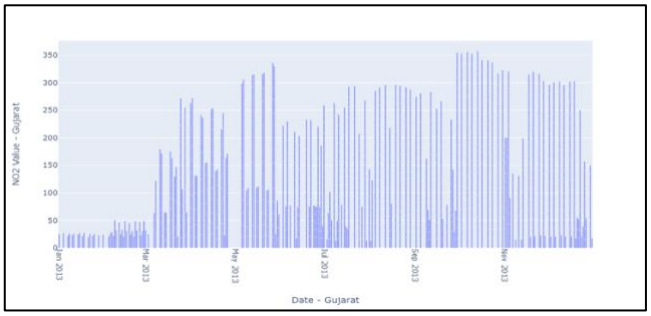


Figure 12. NO_2 value vs Date – Gujarat.

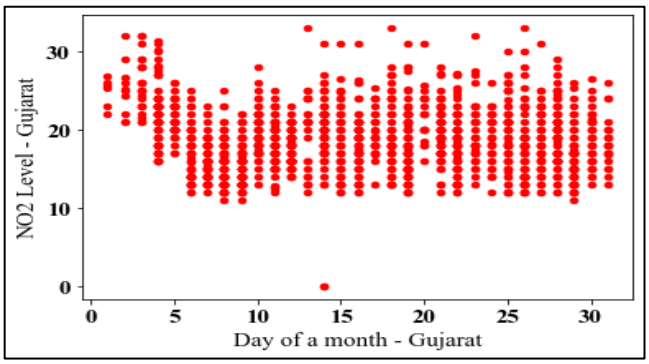


Figure 13. NO_2 level vs Day of a month – Gujarat.

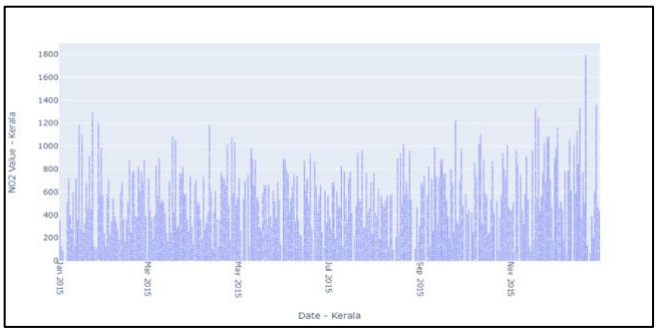


Figure 14. NO_2 value vs Date – Kerala.

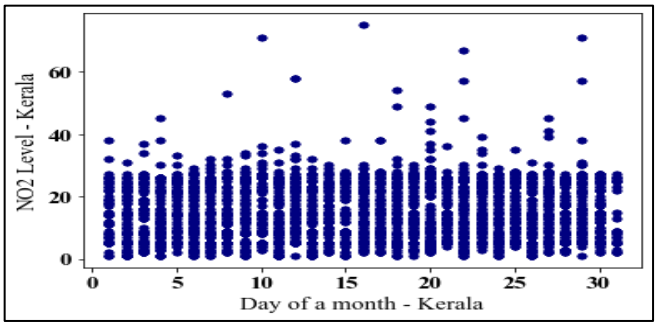


Figure 15. NO_2 level vs Day of a month – Kerala.

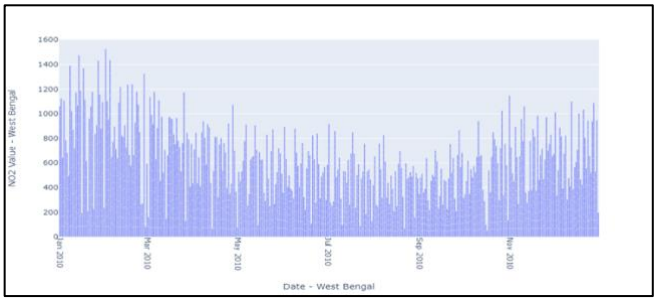


Figure 16. NO_2 value vs Date – West Bengal.

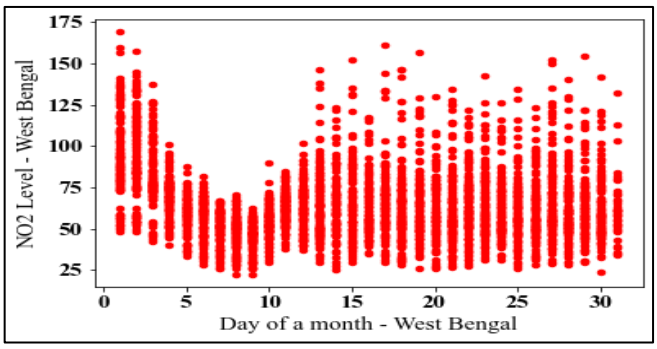


Figure 17. NO_2 level vs Day of a month – West Bengal.

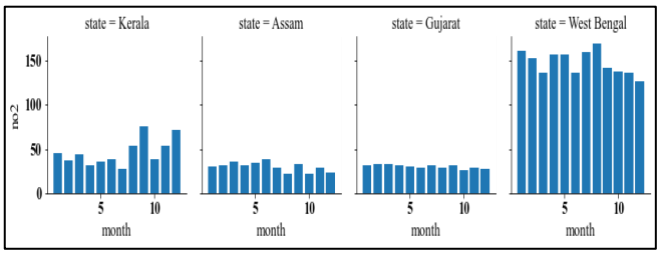


Figure 18. NO_2 values of states in all the months together.

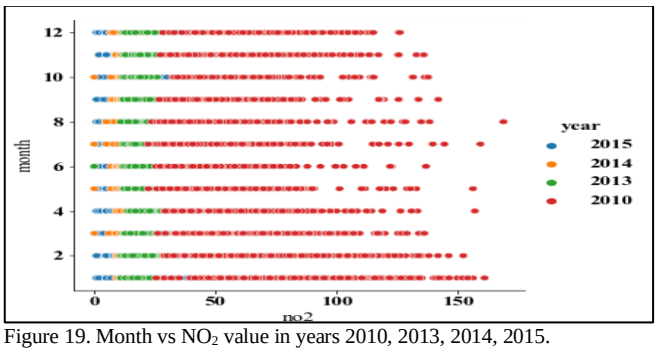


Figure 19. Month vs NO_2 value in years 2010, 2013, 2014, 2015.

B. *Sulphur dioxide (SO_2):*

Sulphur dioxide is enormously emitted from the vehicles and burning of fossil fuels. The graphs and plots above give detailed emission reports of all the states except the state of Kerala as it isn't available within the database. If considering the remaining states and their emission pattern we can clearly see that emissions are major in the developed states like West Bengal and Gujarat compared to relatively less developed Assam.

The below figure depicts the Sulphur dioxide levels in the state of Assam in a year.

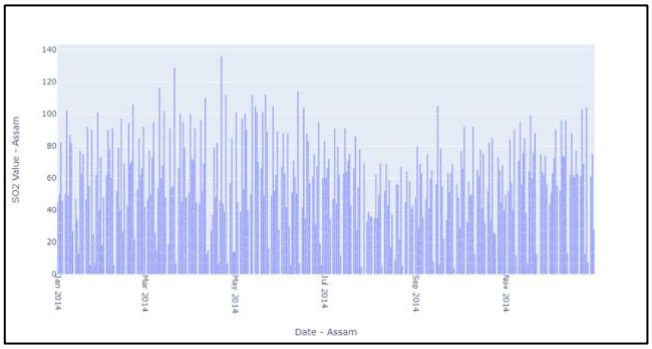


Figure 20. SO_2 value vs Date – Assam.

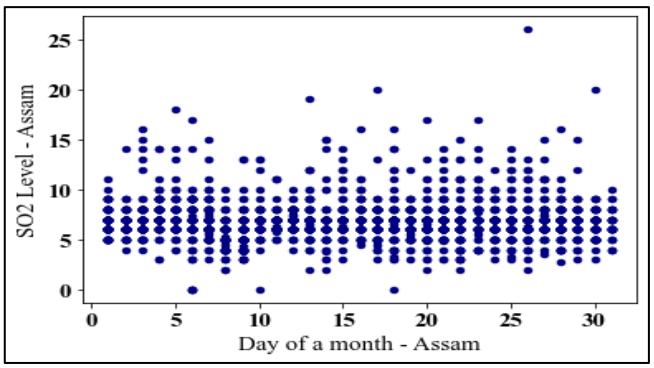


Figure 21. SO_2 level value vs Day of a month – Assam.

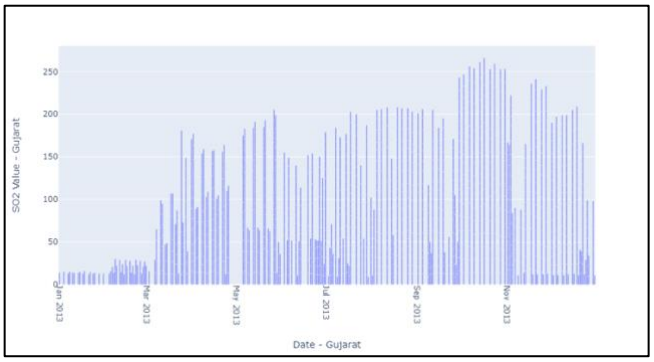


Figure 22. SO_2 value vs Date – Gujarat.

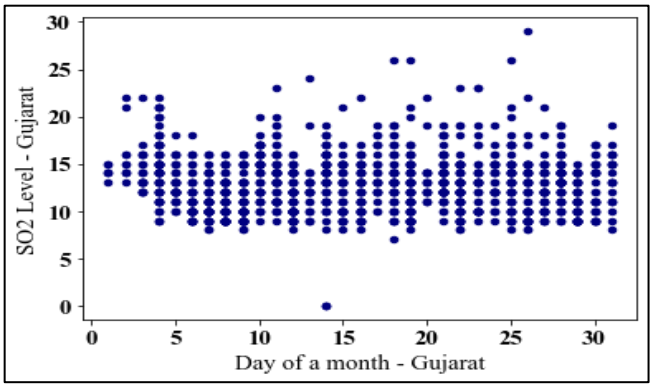


Figure 23. SO_2 level value vs Day of a month – Gujarat.

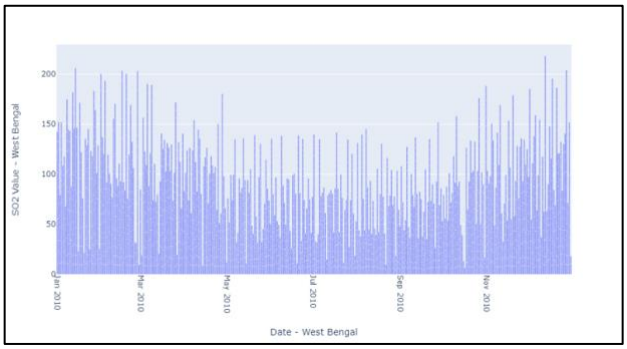


Figure 24. SO_2 value vs Date – West Bengal.

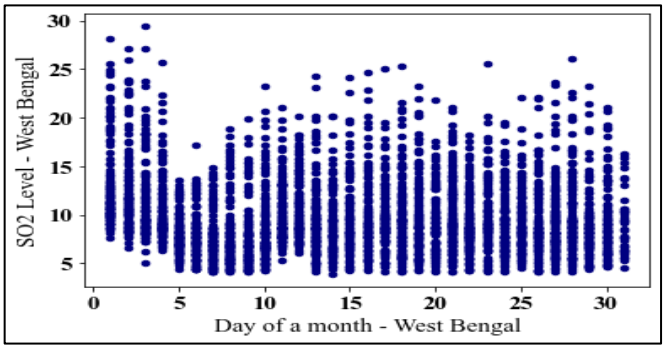


Figure 25. SO_2 level value vs Day of a month – West Bengal

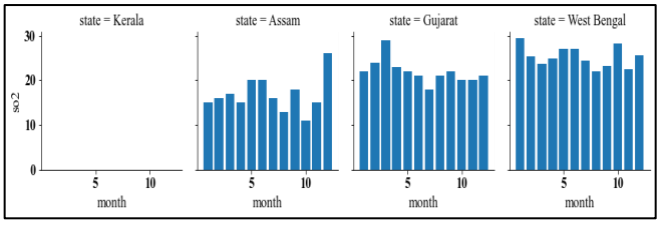


Figure 26. SO_2 value vs Months – Kerala, Assam, Gujarat, West Bengal.

C. Suspended Particulate Matter (SPM):

Suspended Particulate Matter (SPM) is very finely divided particles that are disbursed into the environment during the combustion of fossil fuels. The graphs of Assam and Kerala aren't available as their relevant data wasn't present in the database. But we can here understand clearly from the graphs and plots of the states of West Bengal and Gujarat that the SPM values are completely dependent on the level of urbanization of a particular state. We can see that West Bengal has produced gigantic amounts of SPM compared to Gujarat.

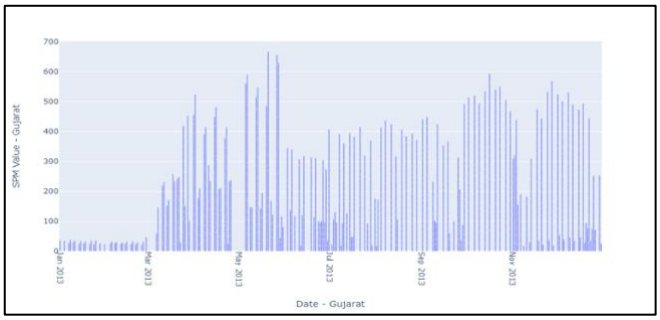


Figure 27.SPM value vs Date – Gujarat

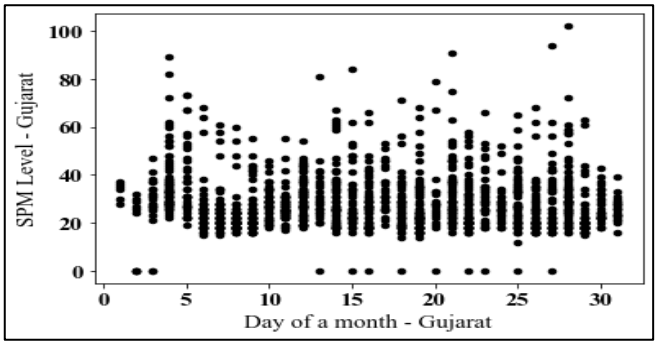


Figure 28.SPM Level vs Day of a month – Gujarat

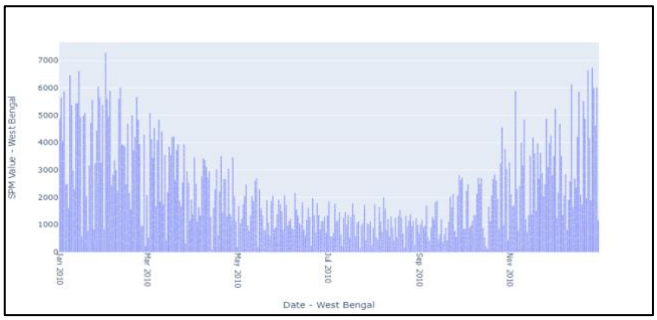


Figure 29. SPM value vs Date – West Bengal.

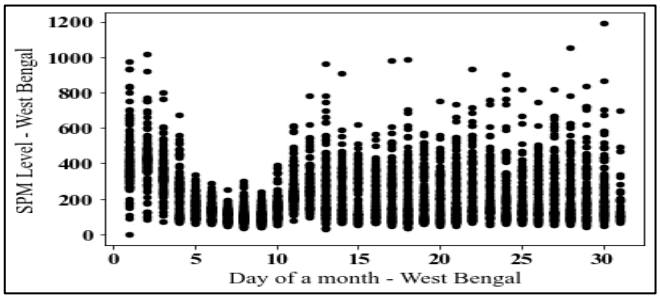


Figure 30. SPM Level vs Day of a month – West Bengal.

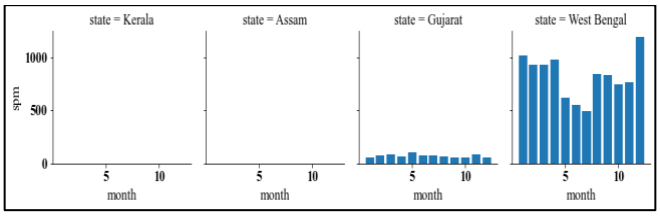


Fig.22: SPM value vs Months – Kerala, Assam, Gujarat, West Bengal

D. RSPM (Respirable Suspended Particulate Matter) / PM10:

Respirable Suspended Particulate Matter (RSPM) or PM10 are very minute and fine particles that are measured in micrometres and are usually inhaled by human beings and get deposited in the lungs and airways causing severe health conditions. They are very hazardous to inhale. They are produced by increased vehicular movement and huge traffic blockades as that increase fuel consumption at a rapid pace resulting in more PM10 released into the environment.

We can see a significant amount of PM10 being produced in a much-developed state like West Bengal compared to other states, which explains the increased vehicular usage.

Interestingly there is a spike of PM10 particles during the end and start of the year in every state. This can be explained by the festive season that usually starts in late October with an Indian festival Dussehra and continues with festivals like Diwali, and followed by Christmas and New Year till early January and the major part of the celebrations are usually involved by firecrackers. These firecrackers contribute to the maximum release of PM10 particles during that season.

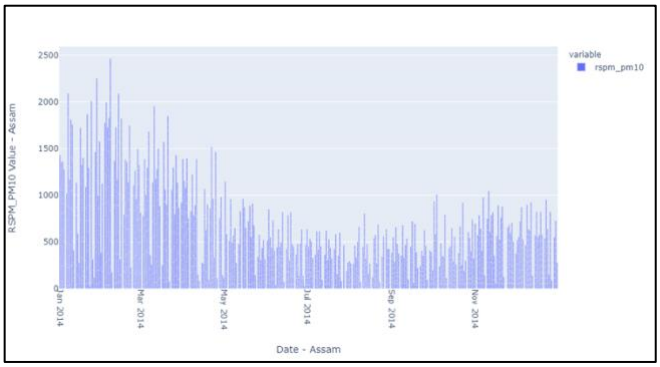


Figure 31.RSPM/PM10 value vs Date – Assam.

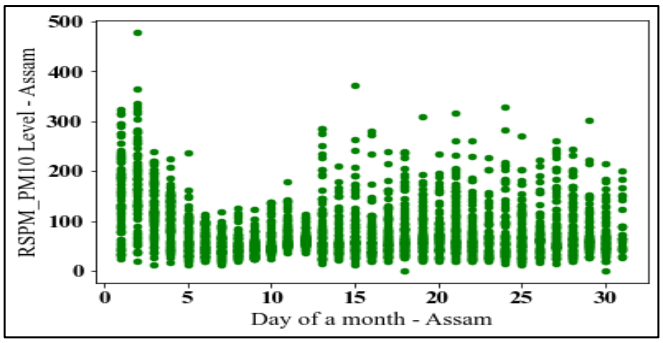


Figure 32.RSPM/PM10 values vs Days of month – Assam

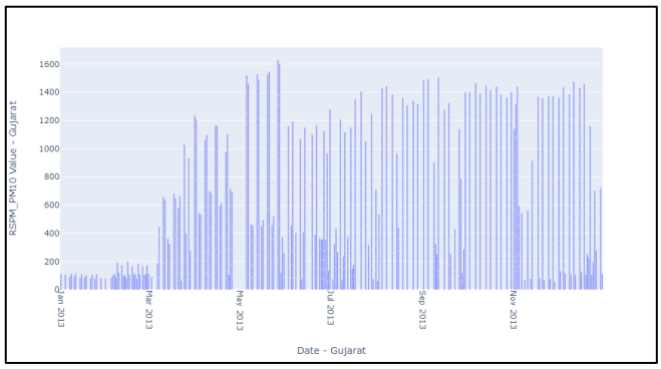


Figure 33.RSPM/PM10 value vs Date – Gujarat

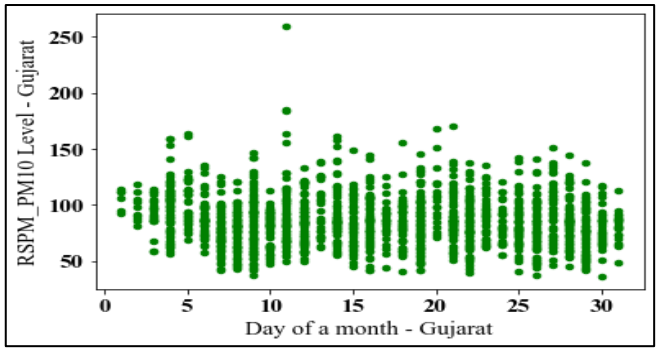


Figure 34.RSPM/PM10 value vs Day of a month – Gujarat

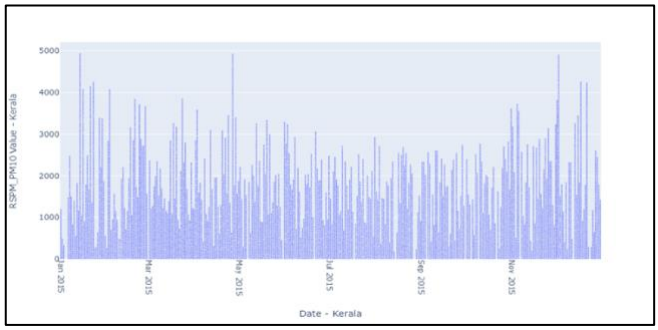


Figure 35.RSPM/PM10 value vs Date – Kerala

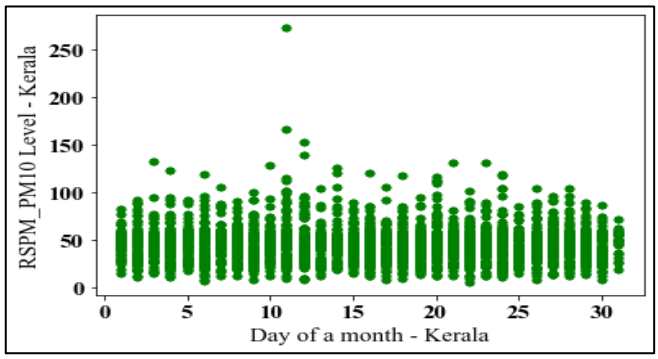


Figure 36.RSPM/PM10 value vs Day of a month – Kerala.

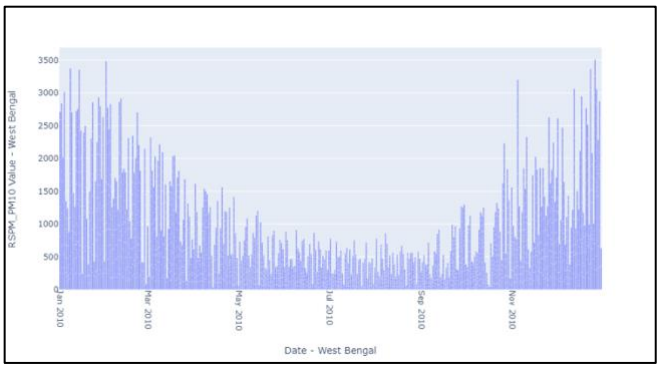


Figure 37.RSPM/PM10 value vs Date – West Bengal

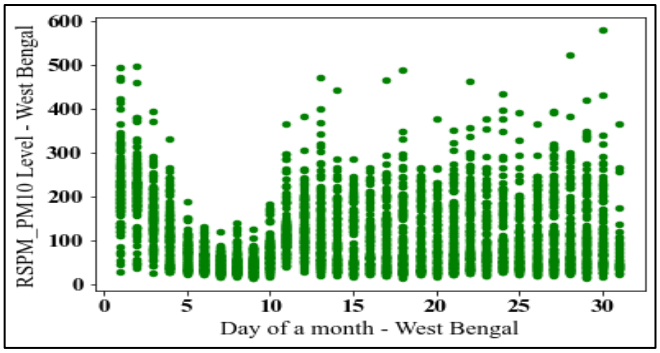


Figure 38.RSPM/PM10 value vs Day of a month – West Bengal

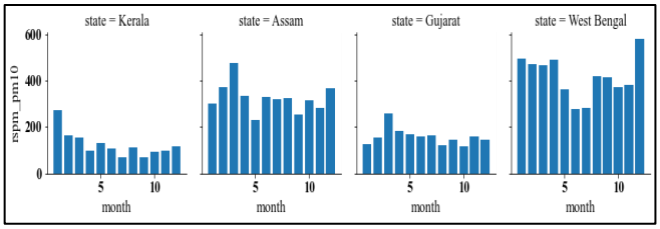


Figure 39.RSPM/PM10 values vs Months – Kerala, Assam, Gujarat, West Bengal.

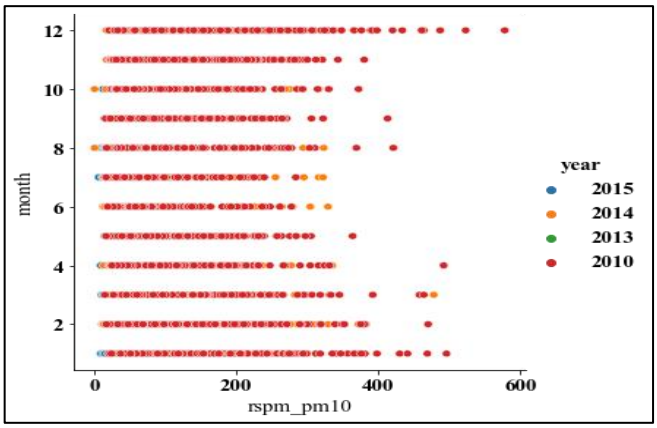


Figure 40.Months vs RSPM value for years 2010, 2013, 2014, 2015

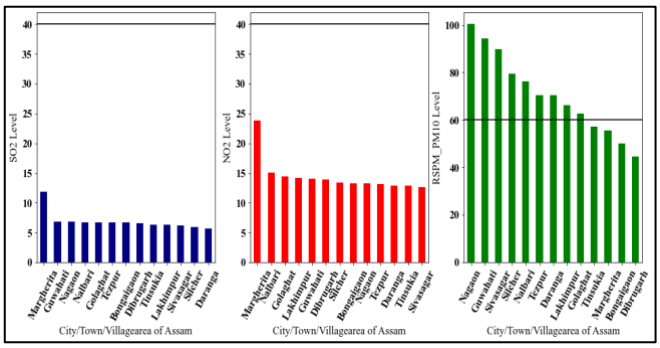


Figure 41.SO_2, NO_2, and RSPM/PM10 level value vs places – Assam.

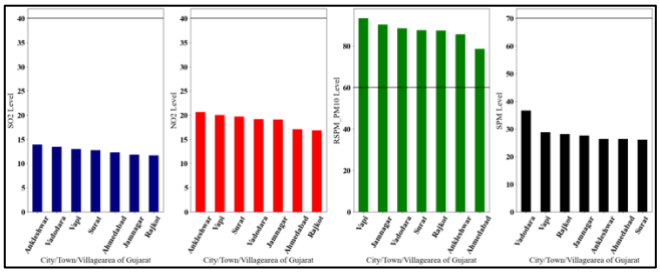


Figure 42.SO_2, NO_2, and RSPM/PM10 level value vs places – Gujarat.

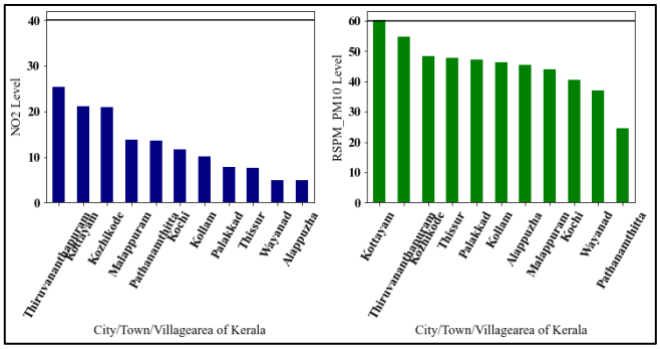


Figure 43. NO_2, RSPM/PM10 level value vs places – Kerala.

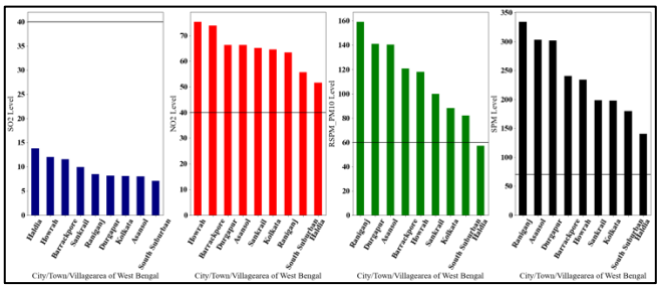


Figure 44. SO_2, NO_2 RSPM/PM10 value vs Places – West Bengal.

As viewed in the above graphs depicted the pollution levels of all the states compared together we can have a clear view that the urbanized regions of the state are having high levels of pollution and the developed states are producing high levels of pollutants compared to the underdeveloped states.

V. Conclusion and Future Work

There are several ways to fight pollution in the future course – one answer to fighting pollution might also lie be in the vision of the world's current richest person 'Elon Musk' – by going completely electric and reducing the dependency on fossil fuels. The complete electrification of every mode of travel and cutting down on the fossil fuels which on combustion produces harmful chemicals like NO_2, SO_2, PM10 which in response builds harmful environmental conditions are to be completely removed if not immediately but at least

in phases. Going electric and the production of electricity from renewable resources like wind, solar and wave power might be a solution for the greener planet and a greener and sustainable generation. Afforestation can also be a possible solution for reducing the enormously increasing pollutants. As long as we save and conserve our planet called Earth, we shall have no reason to make a base on planet Mars like the visionary Elon Musk hopes to make.

REFERENCES

[1] Department of Civil engineering, Lovely Professional University, Phagwara – 144411, Punjab, India. Review of Air Quality Monitoring: Case Study of India. November 2016.

[2] Department of Civil and Environmental Engineering, Birla Institute of Technology, Mesra, Ranchi-835215. A Review on Air Quality Indexing System. June 2015.

[3] International Journal of Advanced Research in Computer and Communication Engineering. Review Paper on Air Pollution Monitoring system. January 2015.

[4] Wales Centre for Public Policy, Cardiff University, Cardiff, Wales CF10 3BG, UK. Air Quality Strategies and Technologies: A Rapid Review of the International Evidence. May 2019.

[5] Office of Epidemiology, Chinese Center for Disease Control and Prevention, Beijing 102206, China. A Bibliometric Analysis of the Impacts of Air Pollution on Children. February 2020.

[6] International Journal for Innovative Research in Science & Technology. Ambient Air Quality Analysis using Air Quality Index – A Case Study of Vapi. March 2015.

[7] ENVIS Centre CPCB. Air Pollution in Delhi An analysis. 2016.

[8] School of Urban Development Faculty of Built Environment and Engineering Queensland University of Technology (QUT) Brisbane, Australi. Valuing Air Quality Impacts of Transportation: A Review of Literature. June 2006.

[9] Department of Computer Science Rutgers University Piscataway, NJ 08854-809. Real-time Air Quality Monitoring Through Mobile Sensing in the Metropolitan Area. 2013.

[10] Department of Computer Science Rutgers University Piscataway, NJ 08854-8091. Impact of urban planning alternatives on air quality: URBAIR model application.

[11] Air Quality Data. Accesed on December 28, 2020 [Online]. Available: https://data.gov.in/search/site?query=air+quality&filter%5Bfield_group_name%3Aname%5D=Air+Quality