Spatial and Temporal variation of Air Quality during the COVID lockdown.

A case study of the 3 most polluted states of the U.S. during COVID.

by

Himanshu Raj
20d180015

Environmental Science and Engineering Department
Indian Institute of Technology, Bombay
Term-Project, Spring,'22

Abstract

COVID and the methods to combat COVID have affected humanity in a drastic manner. Due to lockdown (stay-at-home) there are numerous changes in the environment. But one can easily draw a 'fuzzy-hypothesis' that air pollution must decrease due to shut-down of daily commuting, transport system, industries and so on. This should be made as a general trend, since to decrease the air-pollution, the lockdown must be imposed in a planned and in-synchronised way, for a considerable duration of time, without any natural calamities like storms, wildfires etc. occurring.

In this project, we have tried to study the air-quality of three most polluted states of the U.S.A. And, we will study at the county-level also. We have included California, Pennsylvania and Texas. We have not limited our study to just COVID-lockdown, but invoked some of the natural calamities and environmental injustice issues that are prevailing there. Also, we have tried to incorporate the statistics models like-Box-plot, Time-series Analysis—along with the maps for a deeper insight.

As a pollutant parameter we have restricted ourselves to only *Particulate Matters* only, because the menace caused by it, are relatively higher and tangible than any other parameters.

Acknowledgements

I would like to thank my instructors Prof. Subhankar Karmakar and Prof. Srinidhi Balasubramanian for their constant guidance and support during the course of my project work. They have provided me, excellent insights in Geographical Information Systems, which was pivotal during the work—especially the lectures given by Prof. Subhankar Karmakar on Spatial Data and Interpolation Techniques Prof. Srinidhi B. on Data Structures and Geo-Referencing. These lectures were simply outstanding.

I would also like to thank my TAs Kaustav Mondal, Ravi Ranjan and Avik K. Sam for their constant assistance. Especially, Kaustav tried to left no stone unturned while delivering the lab-lectures. The friendly nature of all the Teaching team and their style of teaching at the very fundamental level has helped me a lot in building my concepts to its deepest root.

Further, I would like to thank all the guest lecturers—Prof Ravindra Dhiman, Anokha Shilin, Dr. Varun Goel, Dr. Mehebub Sahana, V. Jyothiprakash and Mousumi Ghosh.

Contents

Introduction

The inception of COVID occurs amidst Nov 2019 with a small scale. The first large cluster appearing in Wuhan, China in Dec 2019, from China's open-air "Wet Market". Hypotheses are there, which tells that COVID may have as a bio-weapon in a lab in China, but, with a very less confidence. Up to Nov 10, 2020, one year since its outbreak, it has infected 250 million people and with a death toll up to 5 million.

There is a two-way relationship between COVID and air-pollution with some correlation with sociodemographic indices like age, health status etc. Due to this unprecedented situation, where there is a dramatic loss of human life, worldwide —Numerous countries has imposed what it's called as "LOCKDOWN", which is a non-pharmaceutical intervention to combat a disease.

It is colloquially known as stay-home, curfew, quarantine etc., similar to measures to intervene plagues.

Data and Methodology

The first and foremost step of any GIS project is Data Acquisition. We have done majorly it through US EPA. Other sources are mentioned in *Bibliography.* PM Concentration Data obtained was in .csv format, on a daily basis, from US EPA Daily Data. Then the graphs were plotted in MATLAB, using plotting function. MATLAB was chosen over Excel because it is more user-friendly, and redundancy is automatically removed. Box-plot is chosen for depicting spatial variations and time-series analysis is done using stem-plot.

Nevertheless, we have incorporated a variety of heatmaps with a story woven around it. All maps are furnished in QGIS only. The use of any other platform is avoided. All the maps can be found in my GitHub Repository, in modules. The link for the same is attached herewith. Heatmap is chosen over other variety of maps—because it represents the features in an evocative manner. And for a quantitative representation, we have used graphs and charts. Also, in time-series analysis, the daily data is converted to monthly basis before plotting the time-series. Because, daily data would be too big to handle and track.

The shapefiles were downloaded from various source viz., ArcGIS, TIGER, and other sources like Pennsylvania Spatial Data Access (PASDA) etc.

USA

State: California

The inception if COVID in CA took place on Jan 26, 2020.The early confirmed cases were persons who had been to China, recently then. The U.S. dept. of state was successful in evacuating 195 persons from Wuhan, China. Further, 345 more citizens were evacuated from Hubei Province of China to two military bases in California. On March 4, 2020, a state of emergency was declared and of Feb 24,2021, it remained in effect.

Here, we can easily see from the Heatmap (Figure 3.2), that out of 58 counties, Los Angeles is the most polluted. Los Angeles is also called as "*Yang-na (Spanish)"* which means—*the valley of smoke.* People there, are completely relying on automobiles for daily commuting. This has some correlation to its geography. Due to this, Los Angeles has a high level of pollution. Vehicular and airplane exhausts, shipping and manufacturing industries contributes significantly to LA pollution. The pollution can be experience in the form of Smog.

PM 2.5

Spatial Analysis:

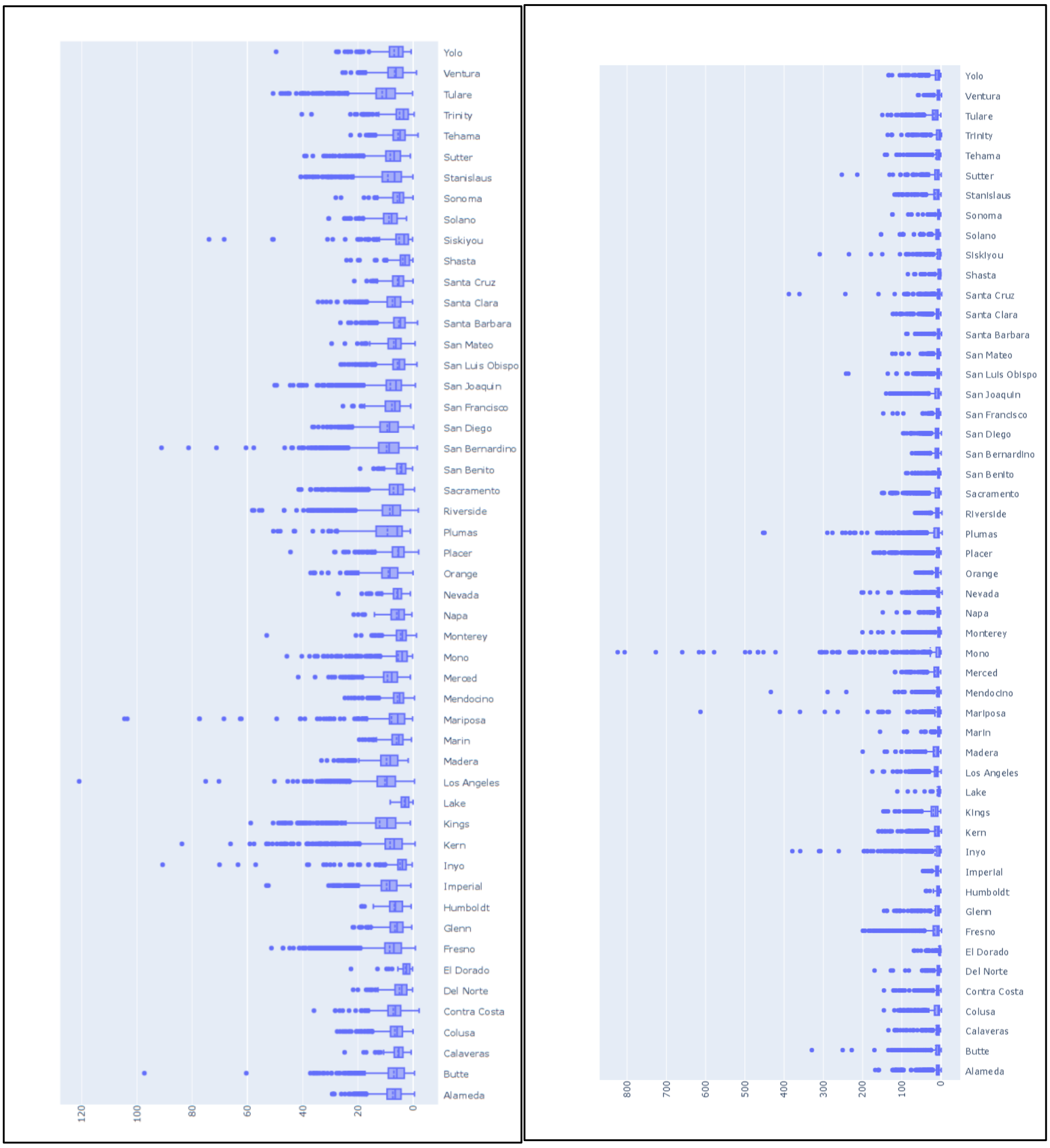


Figure 3.1: Box-plot showing PM2.5 Comparison for 2019(left) and 2020(right) for different counties.

Clearly, we can see the anomaly that the pollutant concentration has increased from 2019 to 2020—substantially. Mono county, located in east central portion of the U.S. of California, was highly affected by wildfires. Wildfires occurred here from Feb 2020 to Dec 2020, and it was the most ferocious, astonishing and horrifying wildfires recorded till date. As a result, it was affected a lot by particulate matters. Mono encountered the highest level of PM2.5, around 800+. Also, the counties had a higher level of pollutants. The daily average for the year 2019 was 7.741 µg/m^3 and for the year 2020 it was 12.474 µg/m^3.

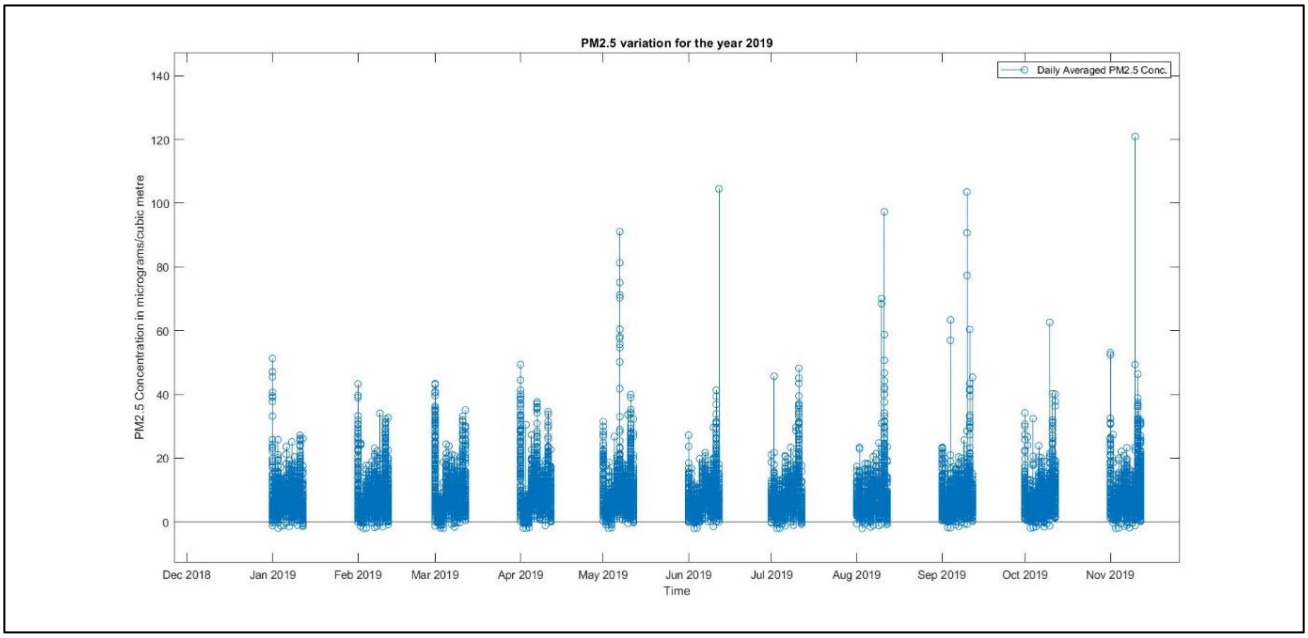
This contradicts the prevailing fuzzy hypothesis that—With the Lockdown, pollution level decreases.

Figure 3.2: Variation of PM2.5 for year 2019(a) and 2020(b) using heatmap.

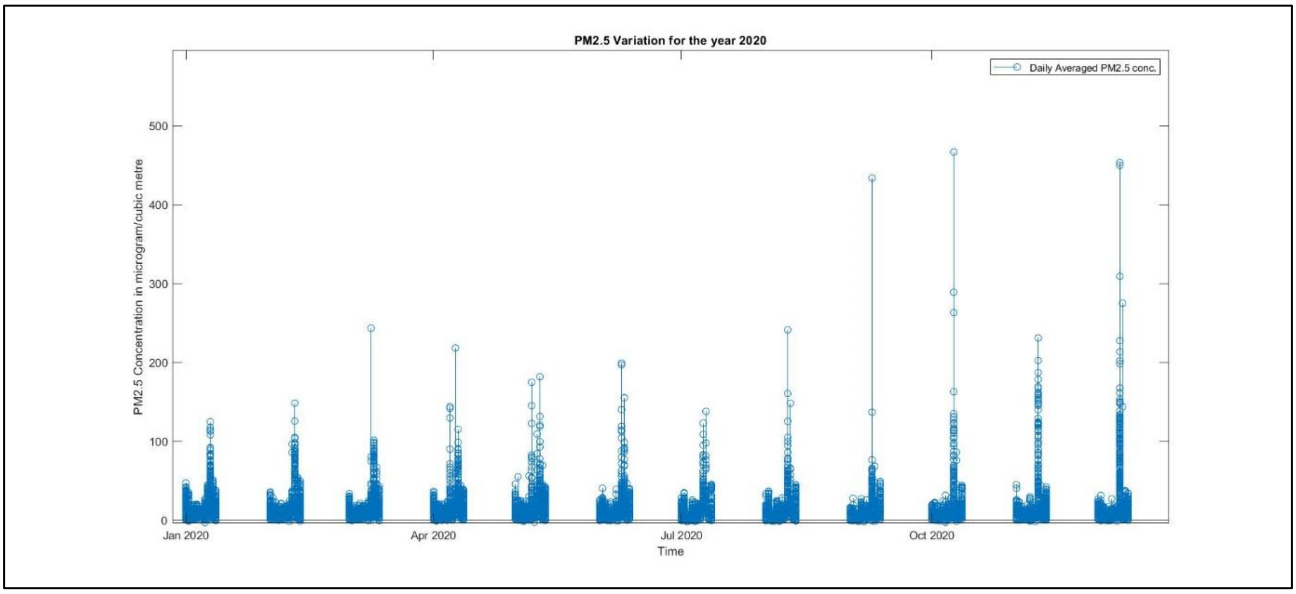
Time-series Analysis:

The scale of the below stem-plot shows all the difference. This has achieved by stacking all the 55 counties together in a group for all the 12 months. Then the data is plotted by importing the .csv file to MATLAB.

(a) Stem-plot showing the variation of PM2.5 with time for 2019.



(b) Stem-plot showing the variation of PM2.5 with time for 2020.



Initially, when lockdown was imposed, there is a decrease in the PM2.5 conc. during the months—January to July, the wildfire was in its dormant state but as the wildfire grew more ferocious in September with the heatwaves—intensified by thunderstorms on August 16-17 due to the moisture from remnants of Tropical Storm Fausto, the pollution level rose significantly. With the highest average on October. The wildfires have added more vexation to Covid, making the citizens to live with the "COVID-FIRE Conundrum".

PM 10

Spatial Analysis:

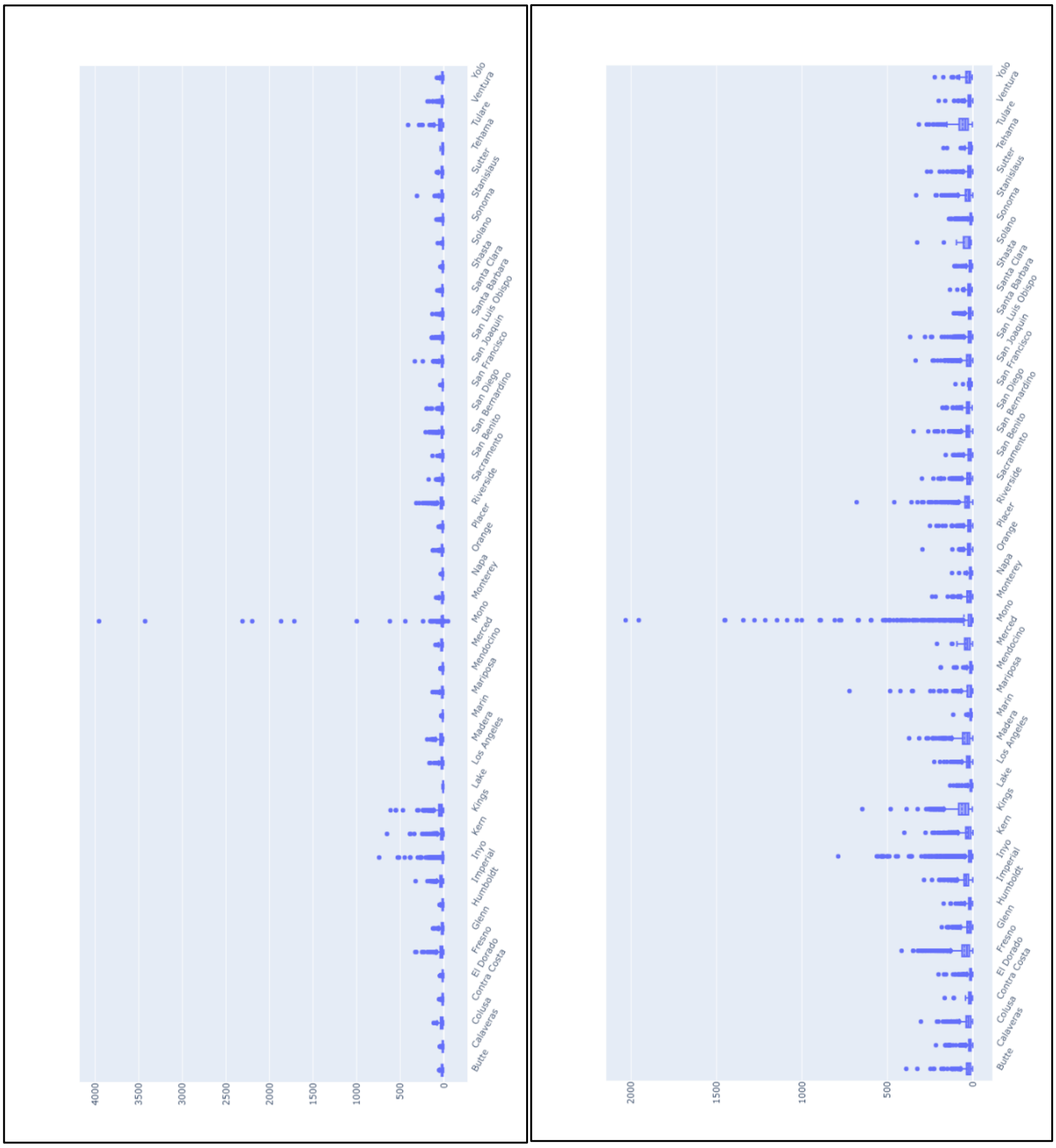


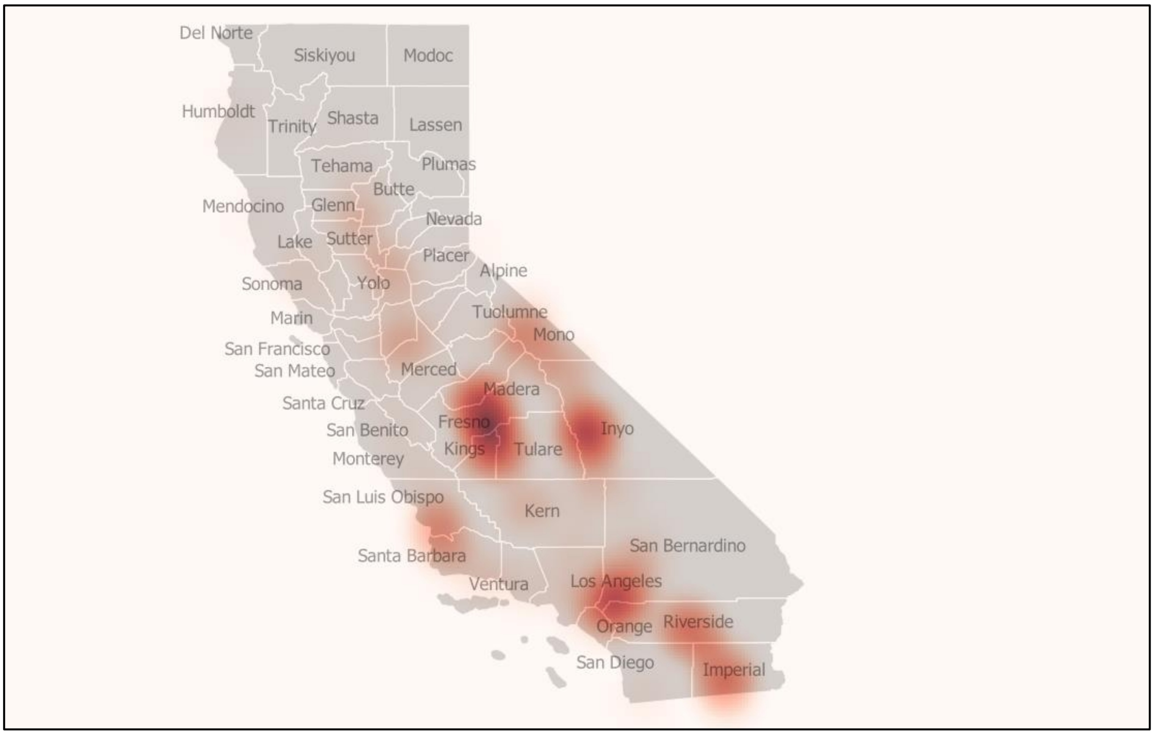
Figure 3.3: Box-plot showing PM10 Comparison for 2019(left) and 2020(right) for different counties.

From the plot, one can easily draw that maximum conc. of PM10 has decreased from the year 2019 to 2020. This doesn't comply with PM2.5 variation trend. This is because, the coarser particles, like PM10, are mainly generated from mechanical operations such as construction activities and agriculture—only a small percentage is present in wildfires (Vicente et. al. 2013, Groβ et. al. 2013).

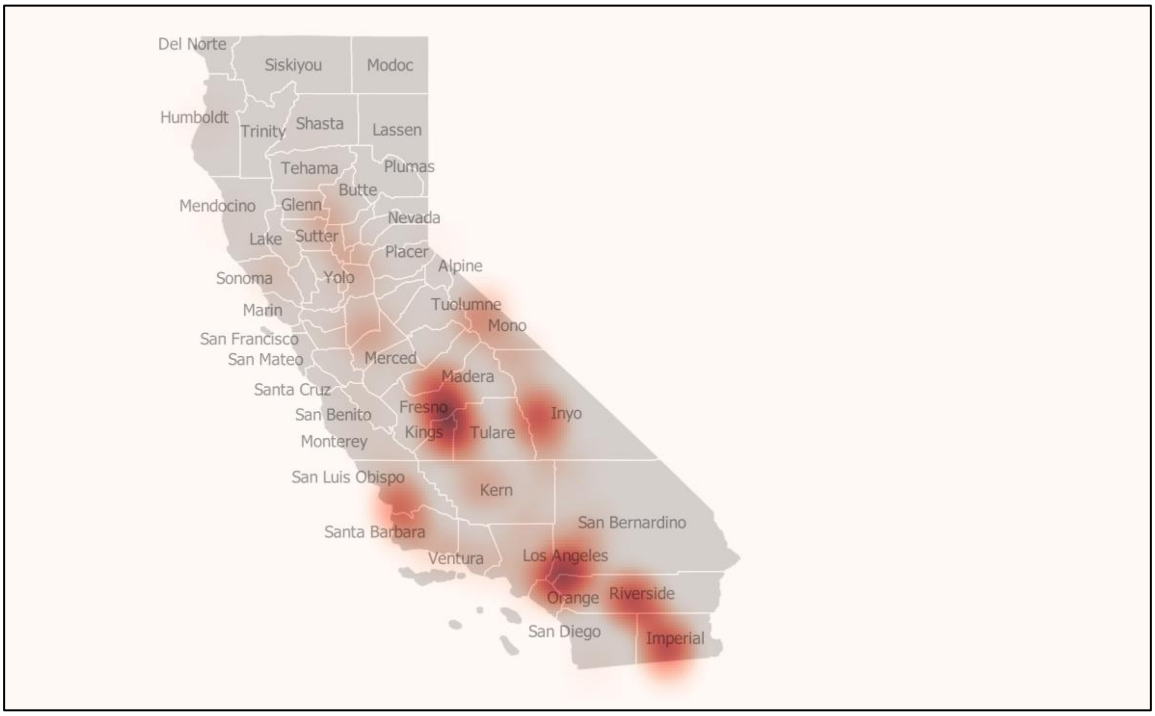
Whereas PM2.5 are the major constituent of wildfires (around 90%). That is why we see a considerable increase in its concentration. But the average value of PM10 has increased, because persistence of the wildfire around the year. Again, Mono county has highest concentration of PM10. The daily average for the year 2019 was 23.340 $\mu g/m^3$ and for the year 2020 it was 31.544 $\mu g/m^3$.

Figure 3.4: Variation of PM10 for year 2019(a) and 2020(b) using heatmap.

(a)



(b)

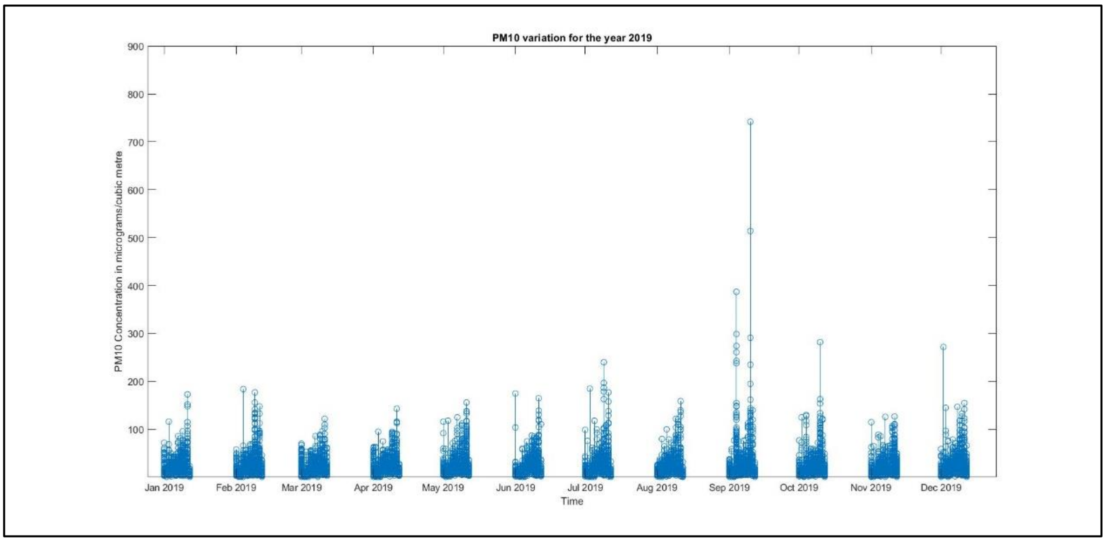


From the above Fig. we could see that there is just a minute difference in the conc. Of PM10. The 'burn-marks' is just intensified. This may be accounted to the increase in the average value of PM10 concentartion. So, in a nutshell, we can say that, wildfire shows more dramatic effect in terms of PM2.5 conc. as compared to PM10 conc. and COVID lockdown has failed to show its effect.

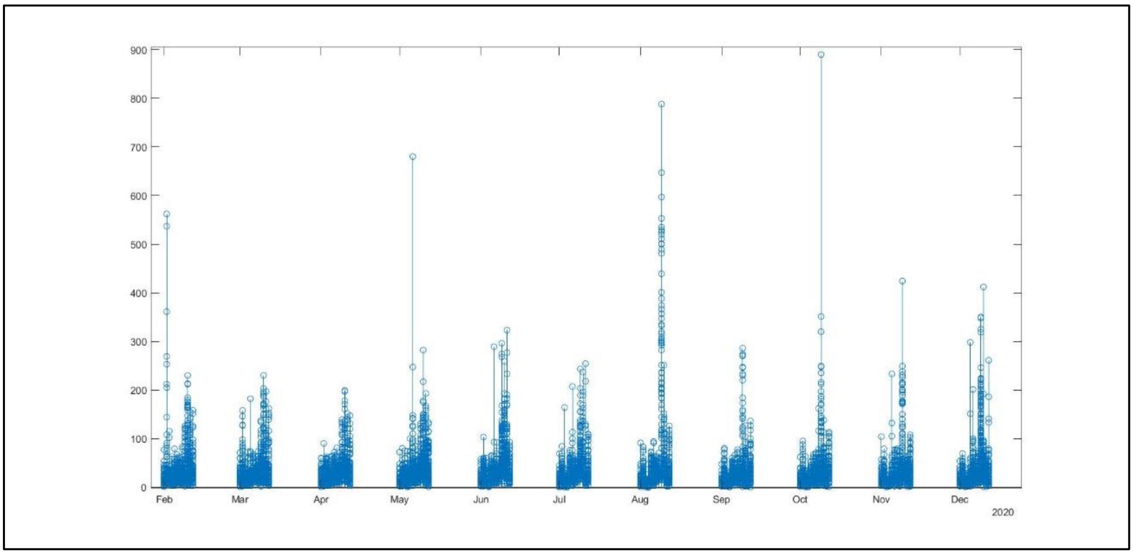
Time-series Analysis:

This has achieved by stacking all the 55 counties together in a group for all the 12 months. Then the data is plotted by importing the .csv file to MATLAB. Clearly, the average value of PM10 has increased from 2019 to 2020. The maximum value of PM10 recorded was in October months, when wildfire was in its peak.

(a) Stem-plot showing the variation of PM10 with time for 2019.



(b) Stem-plot showing the variation of PM10 with time for 2020.



Pennsylvania

A mid-Atlantic north-eastern industrial state of the U.S., famous for its large-scale mushroom production. It was polluted since world war 2nd. But, after rapid development, industries began to expand from the countryside to the downtown areas, pollution started affecting the daily lives. Premature deaths are common. Around 4,800 premature deaths were recorded in the year 2018 due to air-pollution. The culprit is not just the industrial and power-plant emissions but commercial and residential too. The most polluted counties are Philadelphia, Delaware, Allegheny (*Pittsburgh, urban-centre*).

COVID has played a significant role in changing people lives. Air-pollution has decreased by a markable amount due to the "stay-at-home" order.

The first "stay-at-home" order was deployed on Mar 23rd, 2020—for Allegheny, Bucks, Chester, Delaware, Monroe, Montgomery and Philadelphia counties. Then followed by Erie (Mar 24th 2020), Lehigh, Northampton (Mar 25th 2020), Lancaster and York (Mar 27th 2020).

It was altogether, a short-period lockdown—causing a decrease in pollutant conc. by around 10%.

The state and federal govt. both are working extensively to combat Climate Change and air-pollution so as to not let a situation happen like the deadly 1948 Donora-smog.

Figure x.x: The killer Donora-smog.

PM 2.5

Spatial Analysis:

The below box-plot clearly reveals the decrease in the maximum value of PM2.5 conc. from the year 2019 to 2020. Also, it proves that Allegheny County has that highest level of PM2.5 conc. This is because Pittsburgh city has a large number of power-plants and industries which is the main source of pollution in Allegheny.

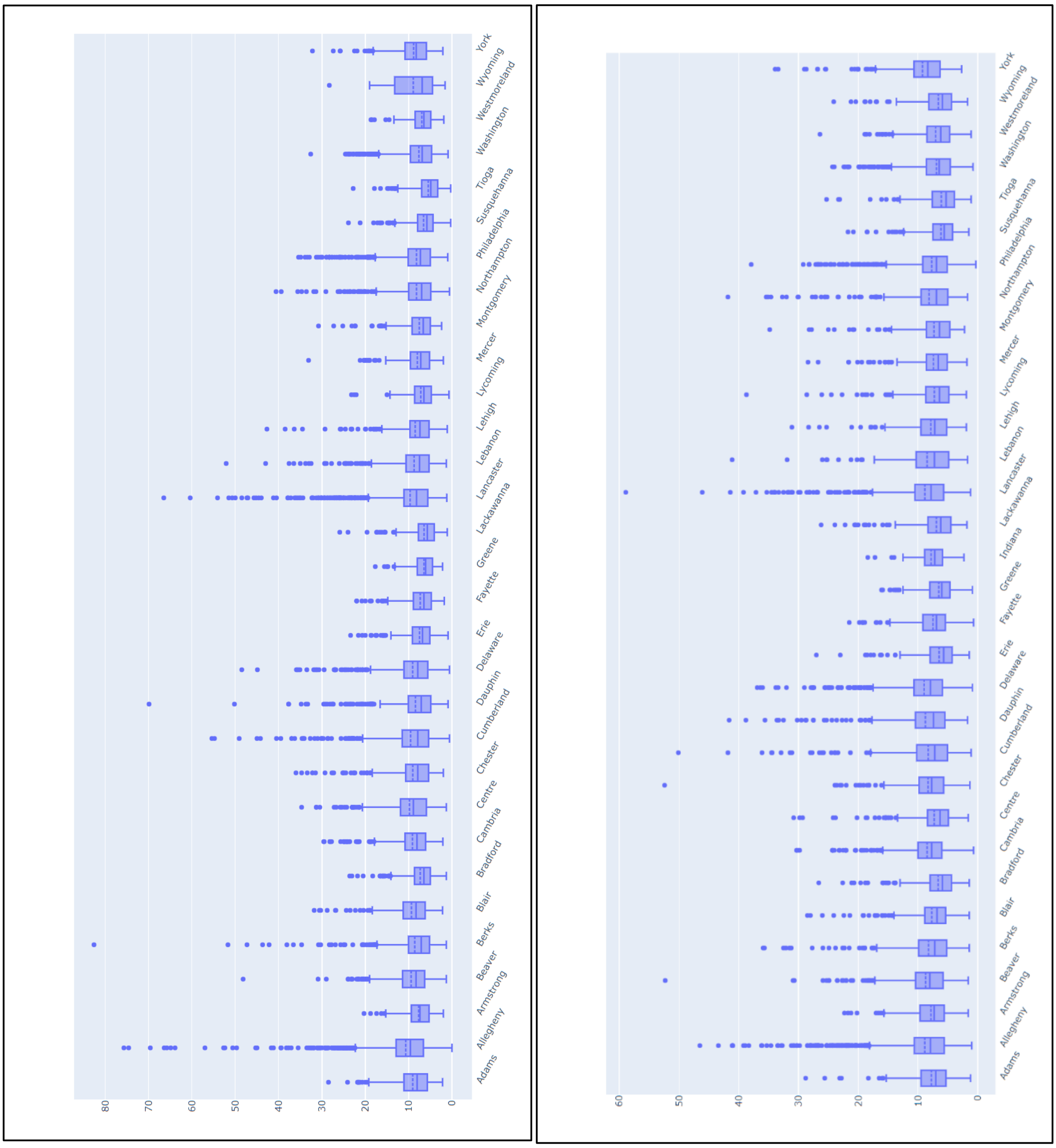
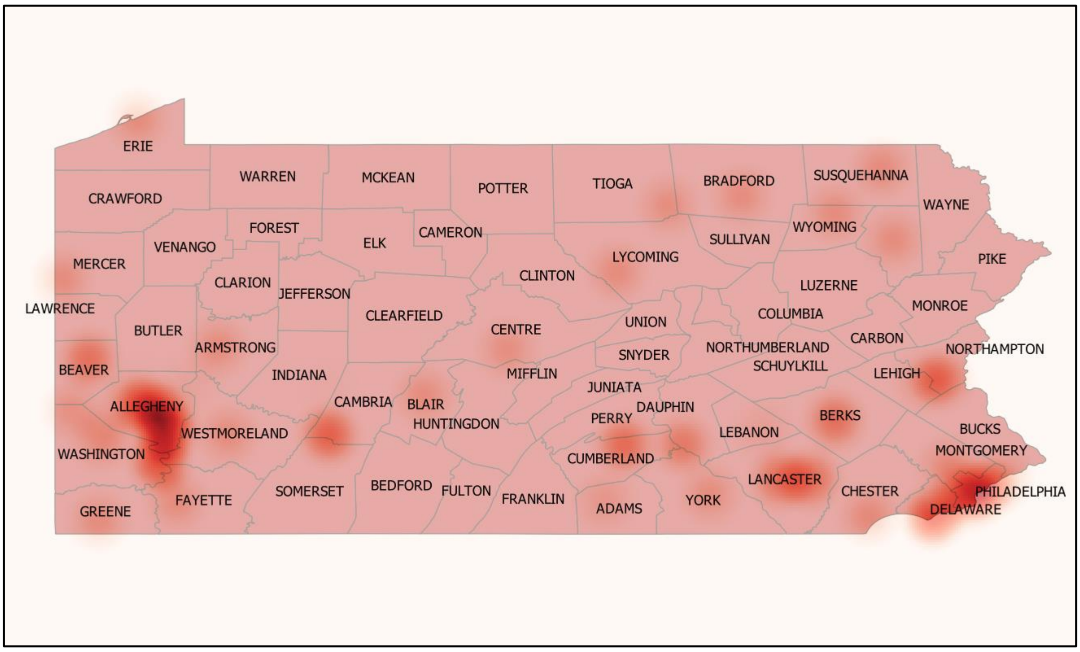


Figure 3.5: Box-plot showing PM2.5 Comparison for 2019(left) and 2020(right) for different counties.

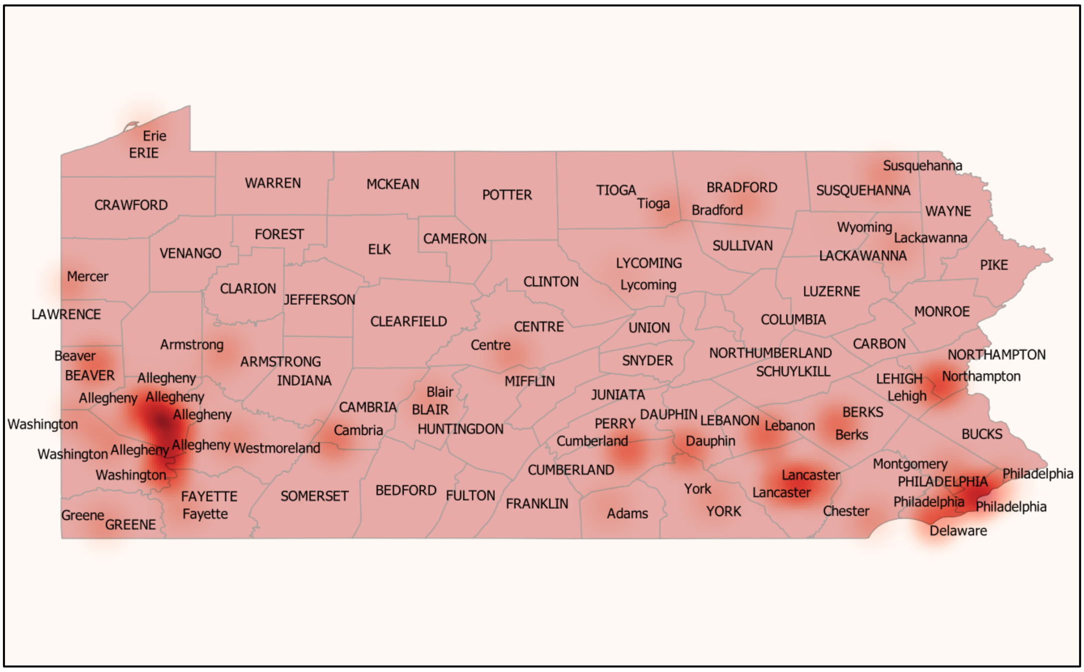
From the map below, we can see that the intensity of the burn-marks is decreased, in the counties like Cambria, Cumberland, Berks by an abstract amount. This can be accounted to the shorter lockdown period and partial and at variance with different counties.

Figure 3.6: Variation of PM2.5 for year 2019(a) and 2020(b) using heatmap.

(a)



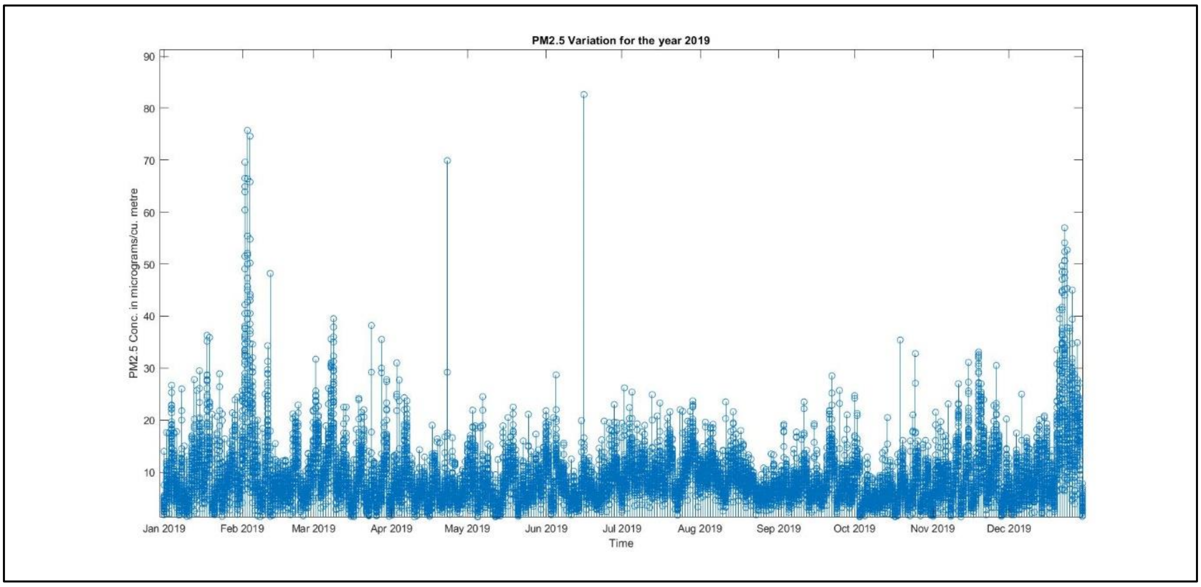
(b)



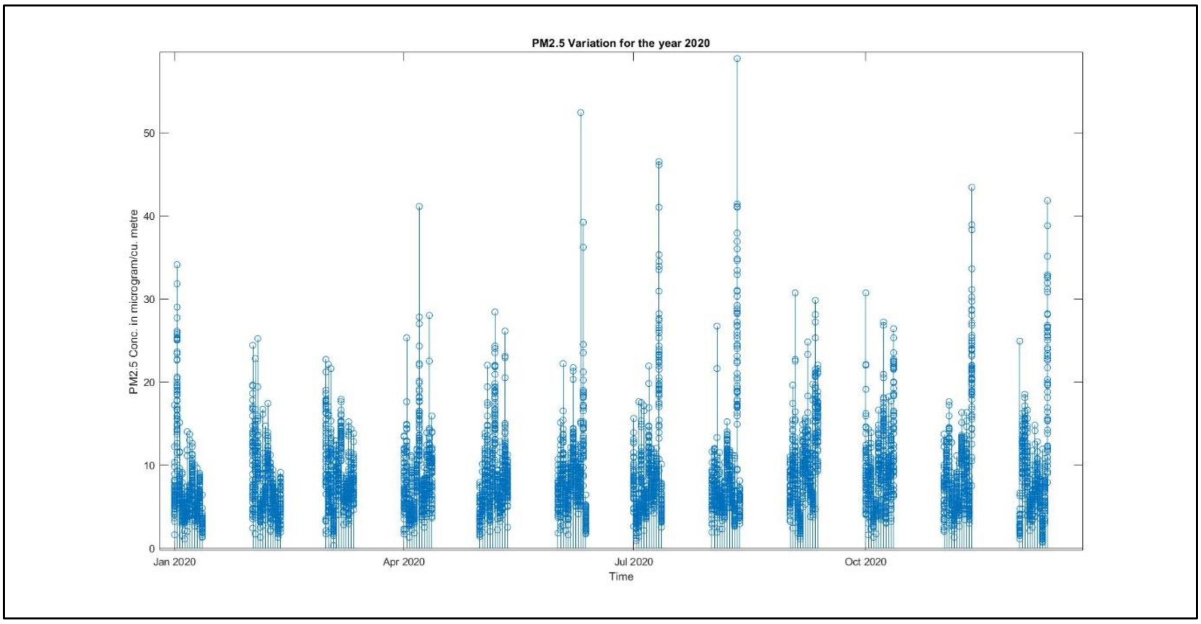
Time-series Analysis:

This has achieved by stacking all the 67 counties together in a group for all the 12 months. Then the data is plotted by importing the .csv file to MATLAB. Clearly, the average value of PM2.5 has increased from 2019 to 2020.

(a) Stem-plot showing the variation of PM2.5 with time for 2019.



(b) Stem-plot showing the variation of PM2.5 with time for 2020.

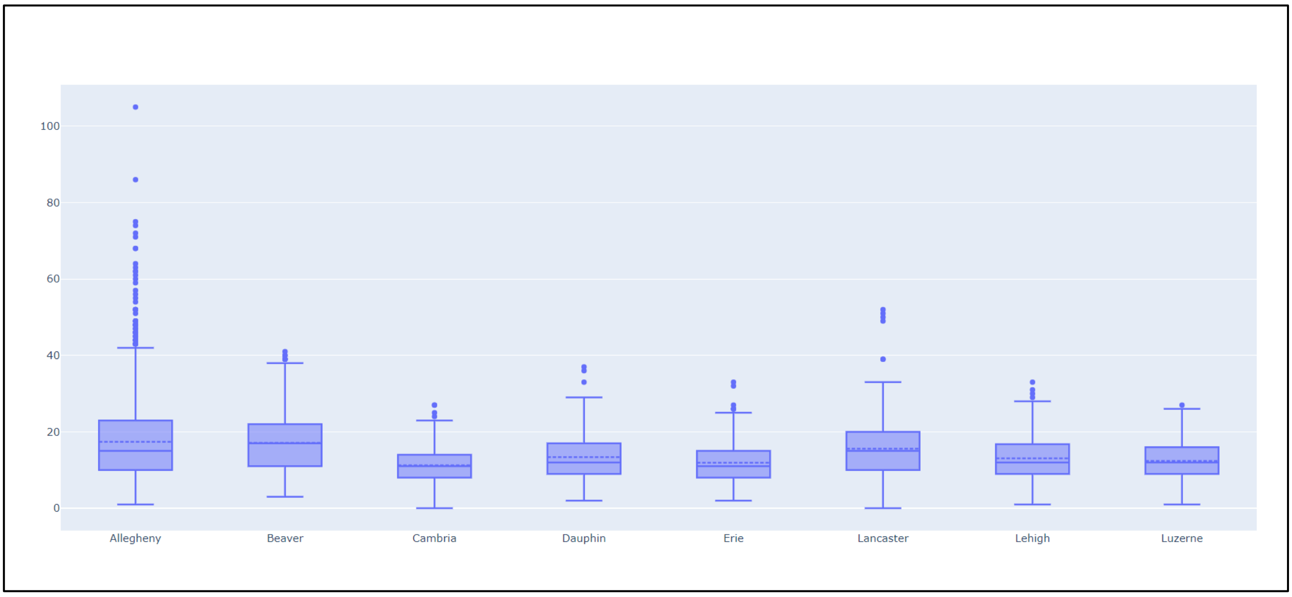


Clearly, we can see that the maximum value of PM2.5 conc. has decreased from 2019 to 2020. Also, the average value of PM2.5 conc. decreased from 8.687 µg/m^3 to 7.929 µg/m^3, which is easily observable from the downward-shift in the crammed region of the graph.

PM 10

Spatial Analysis:

Clearly, we can see that the maximum value of the PM10 conc. has decreased from the year 2019 to 2020. For PM10, Allegheny County bagged the top position, having the upper outlier at around 120. Also, the average value has decreased by around 6%, due to the closure of construction and agricultural activities.



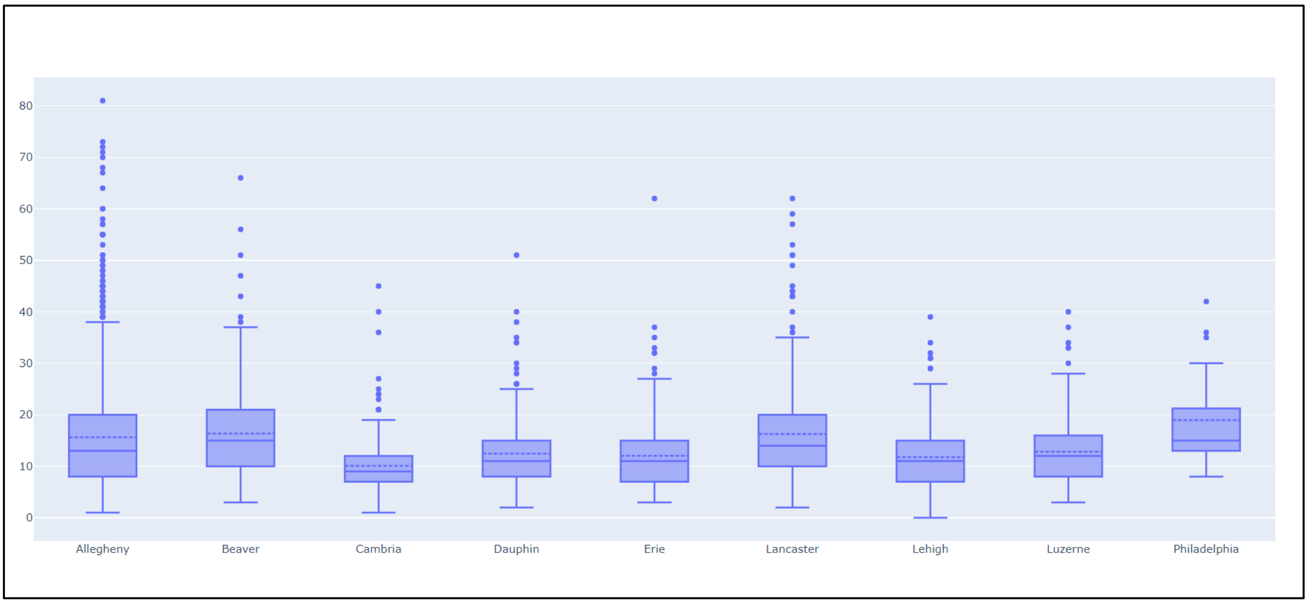
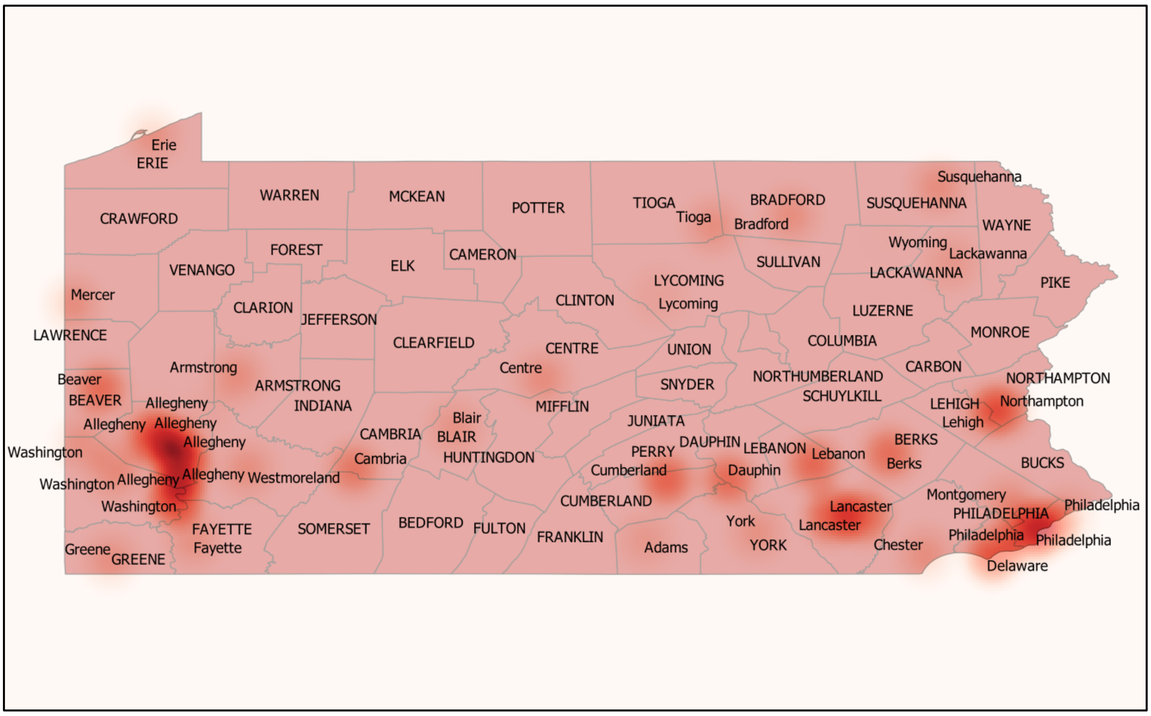


Figure 3.7: Box-plot showing PM10 Comparison for 2019(above) and 2020(below) for different counties.

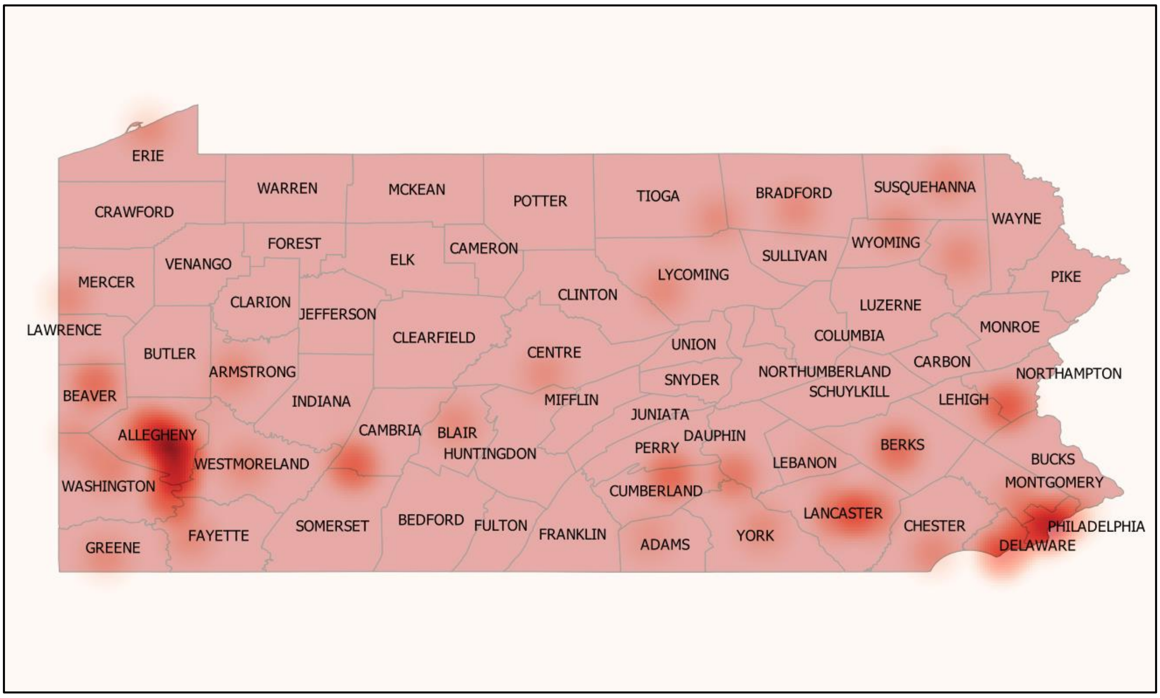
From the map below, we can see that the intensity of the burn-marks is decreased by an abstract amount. This can be accounted to the shorter lockdown period and partial and at variance with different counties.

Figure 3.8: Variation of PM10 for year 2019(a) and 2020(b) using heatmap

(a)



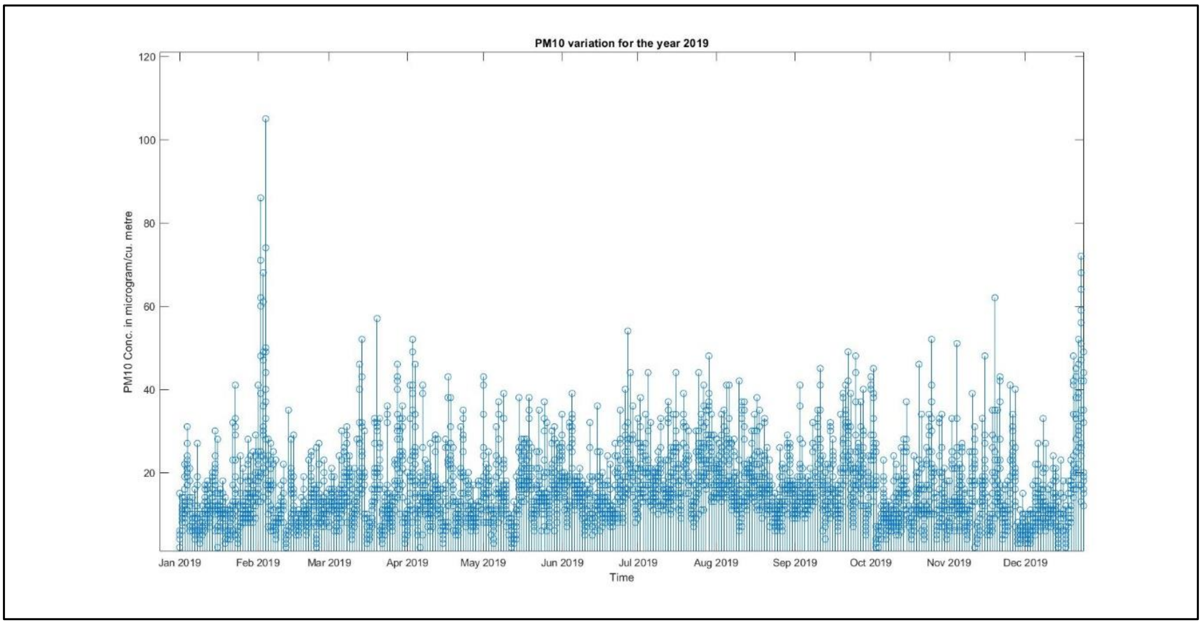
(b)



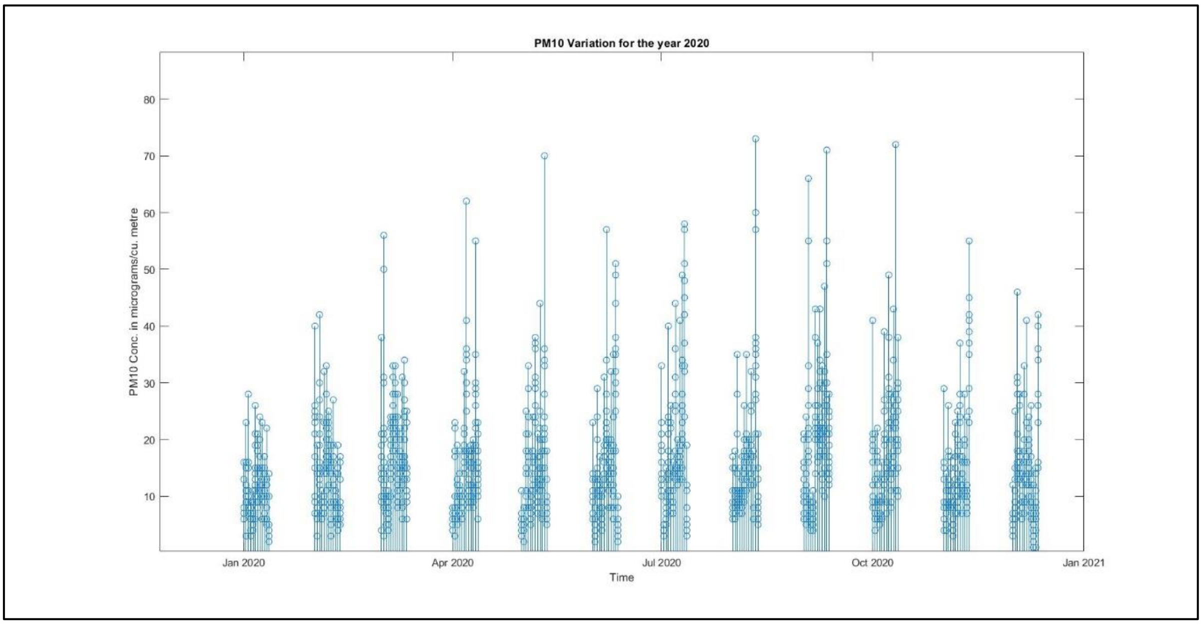
Time-series Analysis:

This has achieved by stacking all the 67 counties together in a group for all the 12 months. Then the data is plotted by importing the .csv file to MATLAB. Clearly, the average value of PM10 has increased from 2019 to 2020.

(a) Stem-plot showing the variation of PM10 with time for 2019



(b) Stem-plot showing the variation of PM10 with time for 2019



Legibly, we can see that the peak and the average value is decreased from 15.378 µg/m^3 to 14.310 µg/m^3.

<u>Texas</u>

A state located in the South-central region of the U.S., is the second largest state after California (both by area and by population). It is also called as “Lone Star State”, and so is its seal and flag. It is 10% larger than France. It is the home-state of the famous U.S. President—George W. Bush.

But, this prominent state of United States suffers very badly from the particulate matter pollution and suffers a pollution with a racial-discrimination basis (Dallas) and coloured are dying to breath clean air,

The “Shingle-Mountain” conflict began around 2020, December, when Marsha Jackson (a coloured woman) began protesting against a mountain of Shingles were piled just behind her house. Shingles are a source of PM10 and PM2.5 with particles of glass and cements and sometimes metals. She was constantly facing health related issues and her health was deteriorating day by day. Finally, after a long protest the state had to clean that area.

Generally, the garbage is dumped to the area where black and brown communities live. Michael Waters, a Pastor and Activist said, *“the life expectancy of the people here depends upon their zip codes here”.* This can be attributed to the white supremacy—existing till date. According to Washington Post, the life expectancy of the people living in Floral Farms (where brown community live) is around 71 years while, in University Park (where white live), it’s around 84 years. It reflects a nation-wide phenomenon. There are hundreds of Marsha Jackson due to hundreds of Shingle Mountains-in the west who are suffering constantly because of Environmental Injustice.

But when it comes about COVID lockdown, there was no as such strict lockdown. There was a partial closure of activities and places such as recreational places, schools etc. There was an executive order from Governor Abbott, on Mar 31st 2020, which amends social distancing policies, minimising the activities that are susceptible to the spread of COVID, till May 4th 2020. So, the lockdown doesn’t play a key role here. And, as a general hypothesis, the pollution level would not undergo drastic change.

Rather, the data reveals that there is increase in the PM2.5 level by 3%, because of the continuation of daily commuting, industrial activities, petrochemical industry near Houston, and refineries.

Figure 3.9 (a)

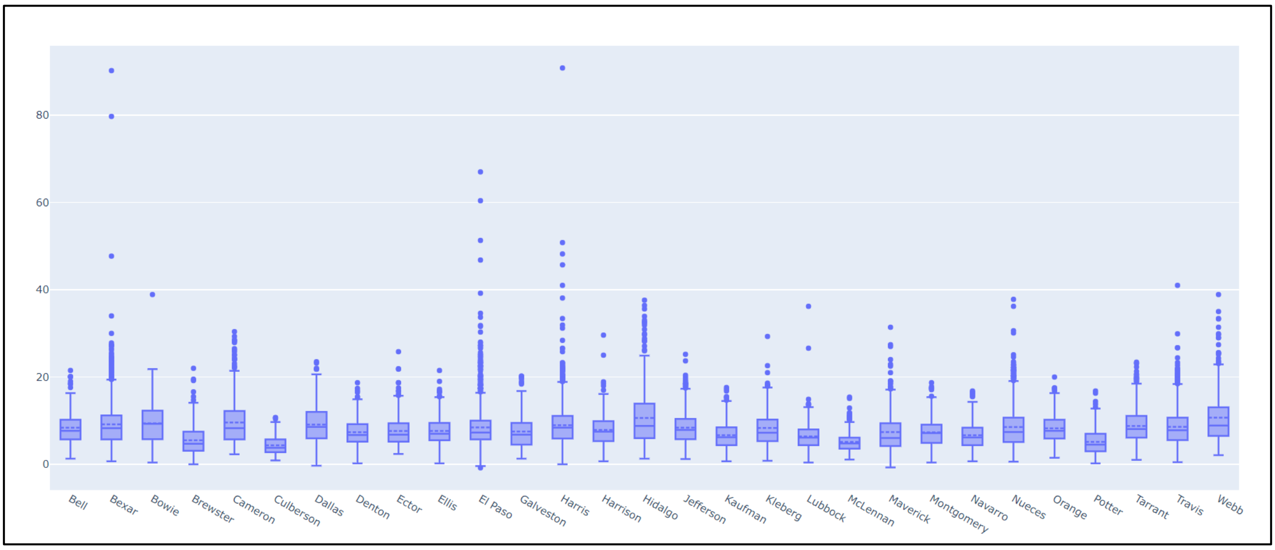
Figure 3.9 (b)

Figure 3: Protest against Texas Govt. against piling us disposed shingles in residential area.

PM 2.5

Spatial Analysis:

Clearly, we can see that the maximum value of the PM2.5 conc. has increased from the year 2019 to 2020.



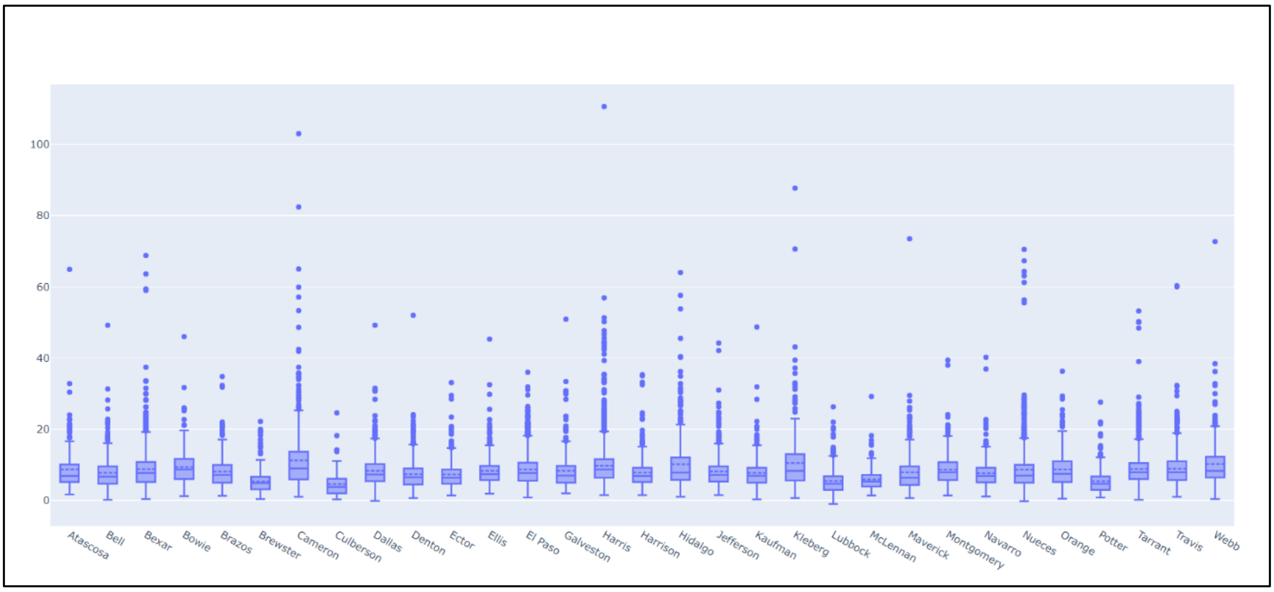
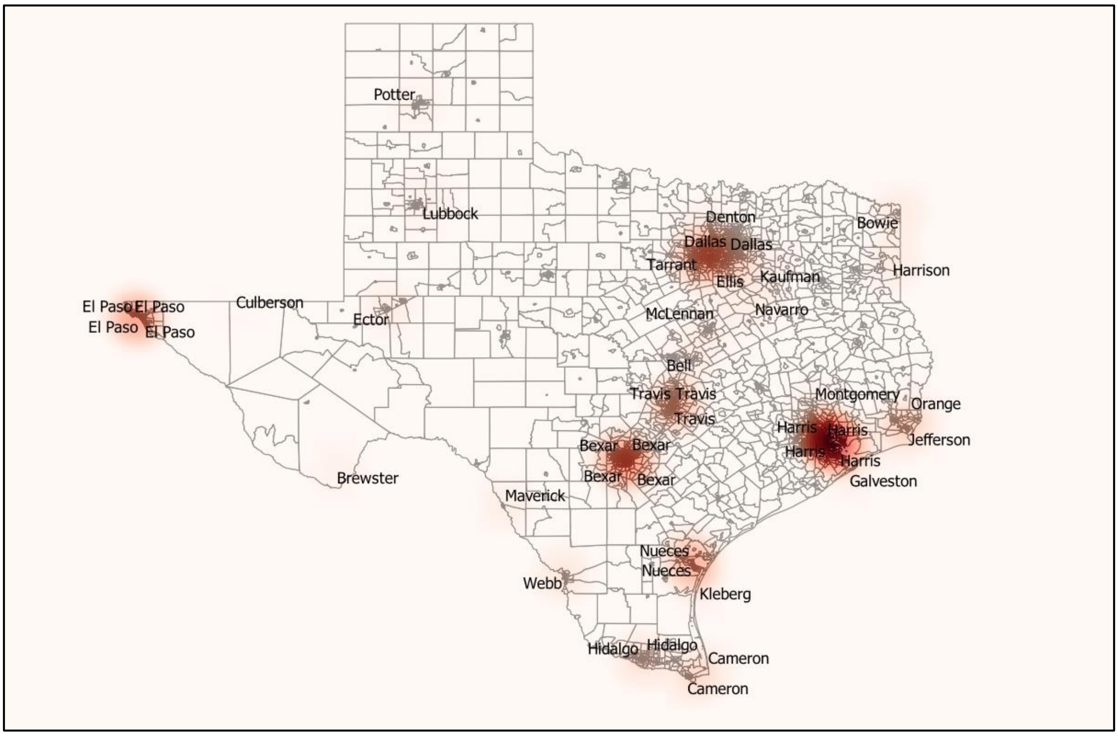


Figure 3.10: Box-plot showing PM2.5 Comparison for 2019(above) and 2020(below) for different counties.

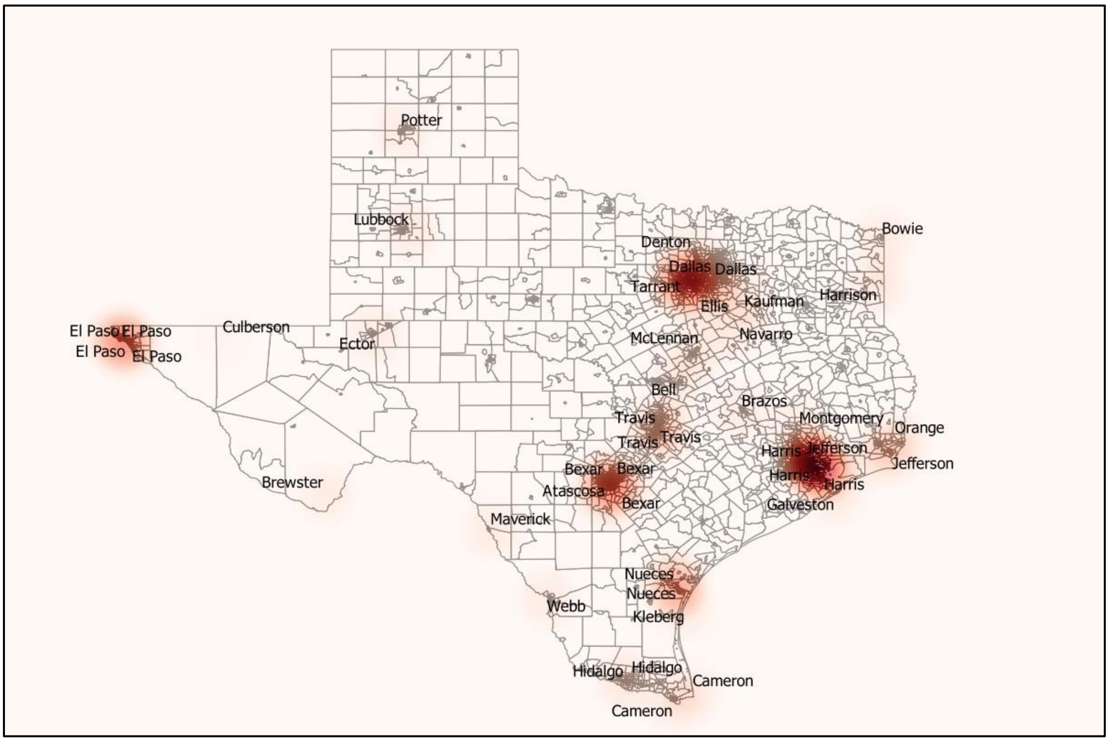
There is no observable change inferred from the heatmap below.

Figure 3.11: Variation of PM2.5 for year 2019(a) and 2020(b) using heatmap.

(a)



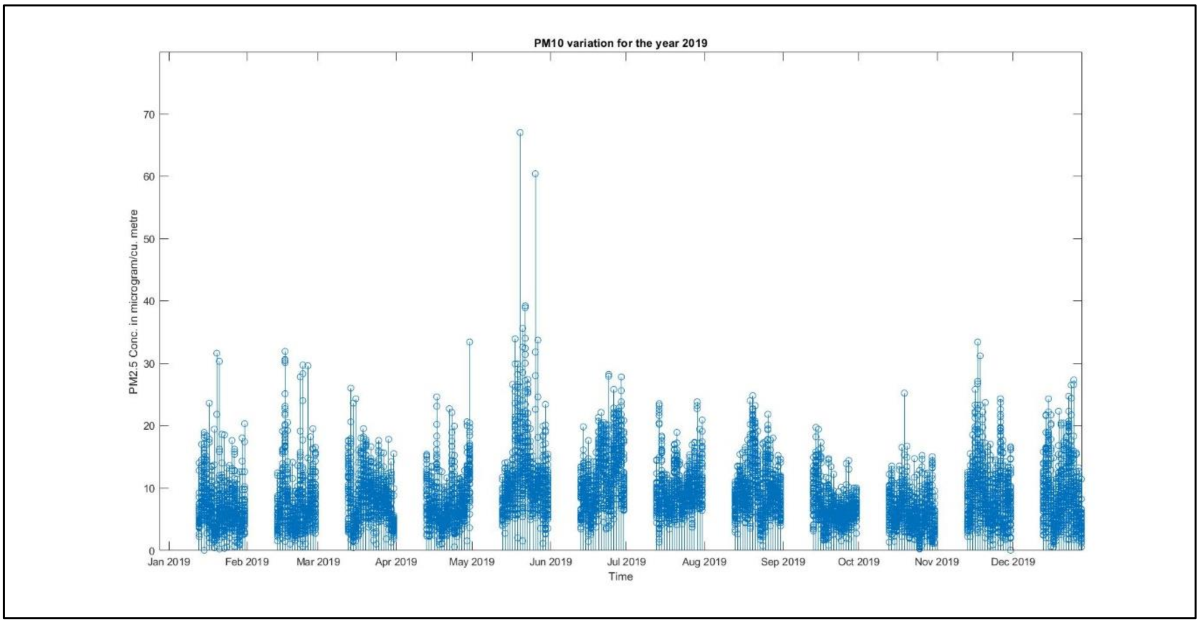
(b)



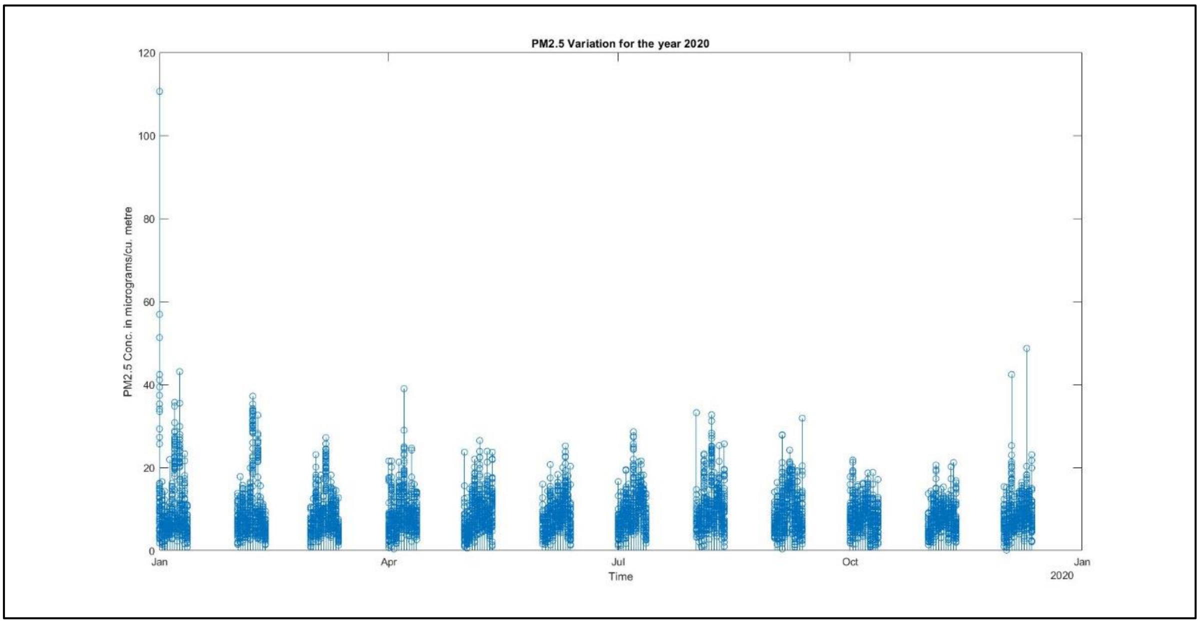
Time-series Analysis:

The time series analysis shows that the average value remained the same for both the years.

(a) Stem-plot showing the variation of PM2.5 with time for 2019.



(b) Stem-plot showing the variation of PM2.5 with time for 2020.



PM10

Clearly, we can see that the maximum value of the PM10 conc. has increased from the year 2019 to 2020, because of unsynchronised lockdown periods in different counties.

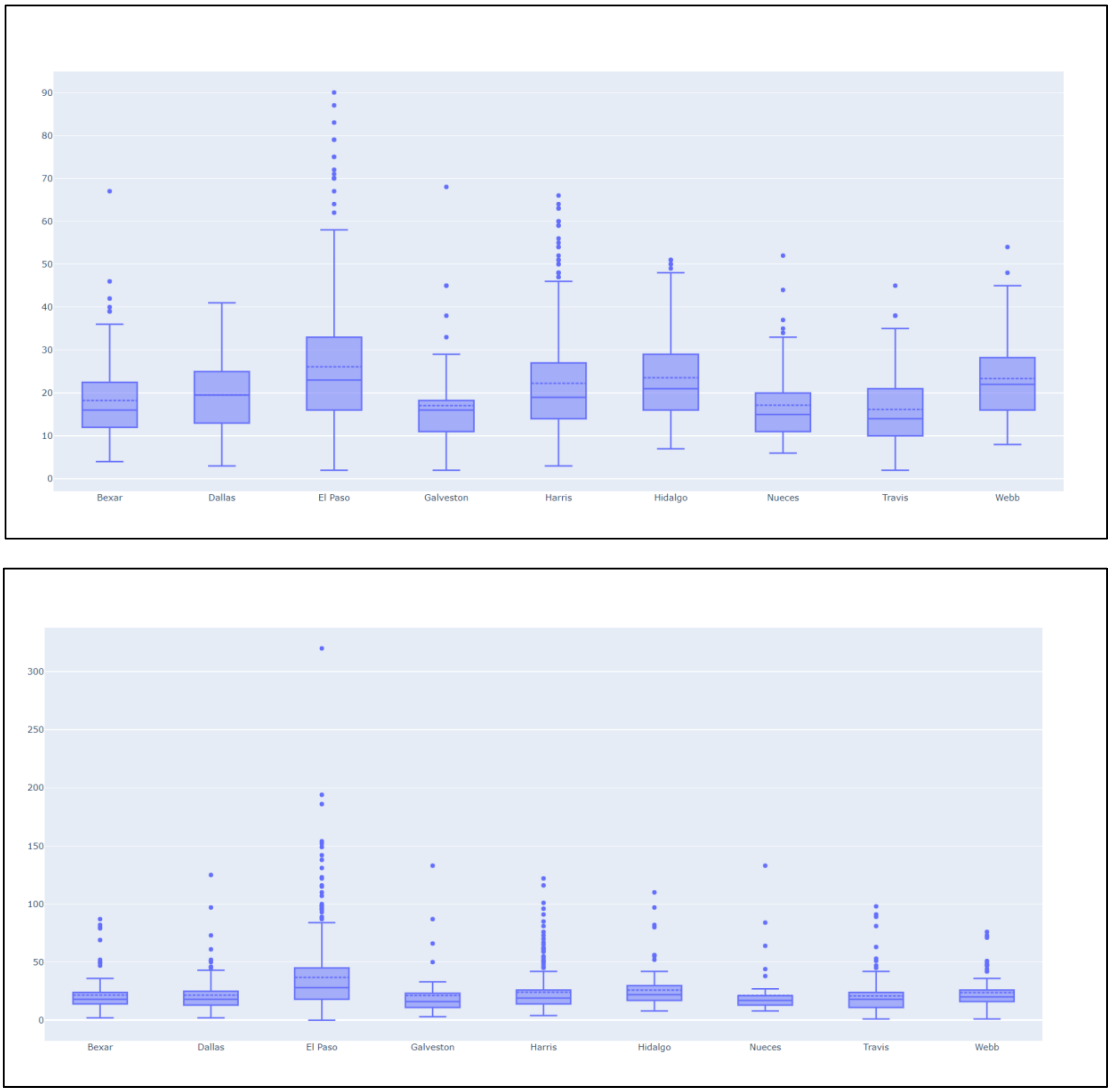
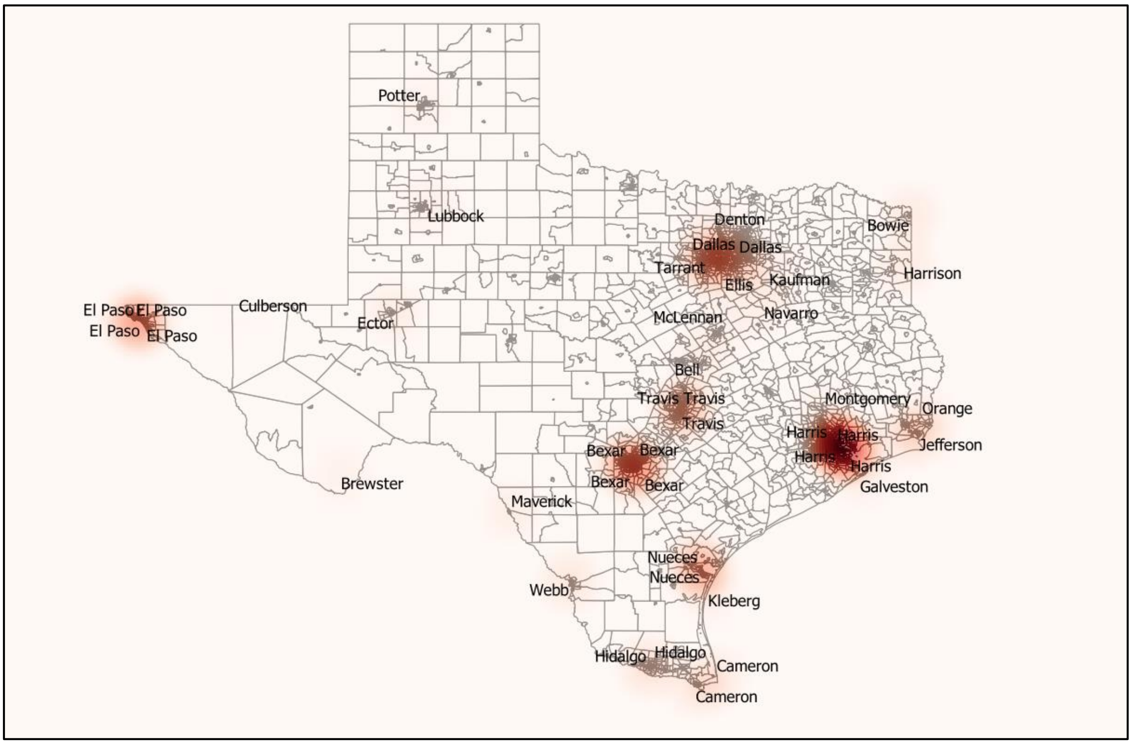


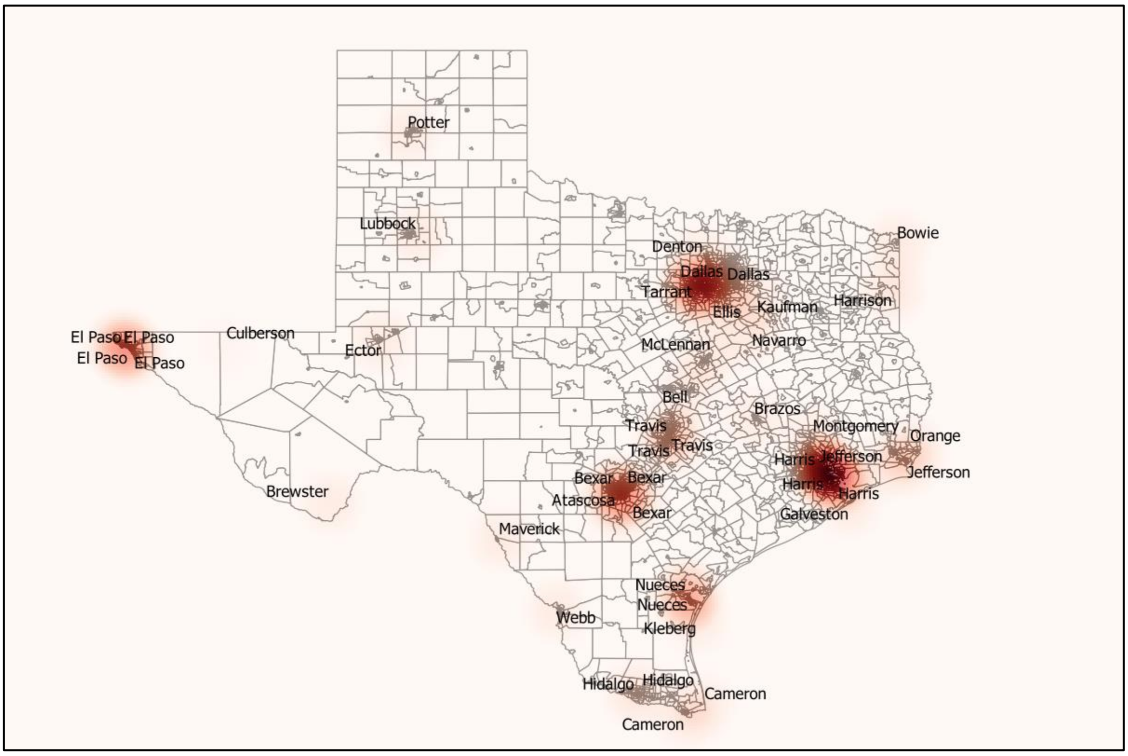
Figure 3.12: Box-plot showing PM10 Comparison for 2019(above) and 2020(below) for different counties.

Figure 3.13: Variation of PM10 for year 2019(a) and 2020(b) using heatmap

(a)



(b)

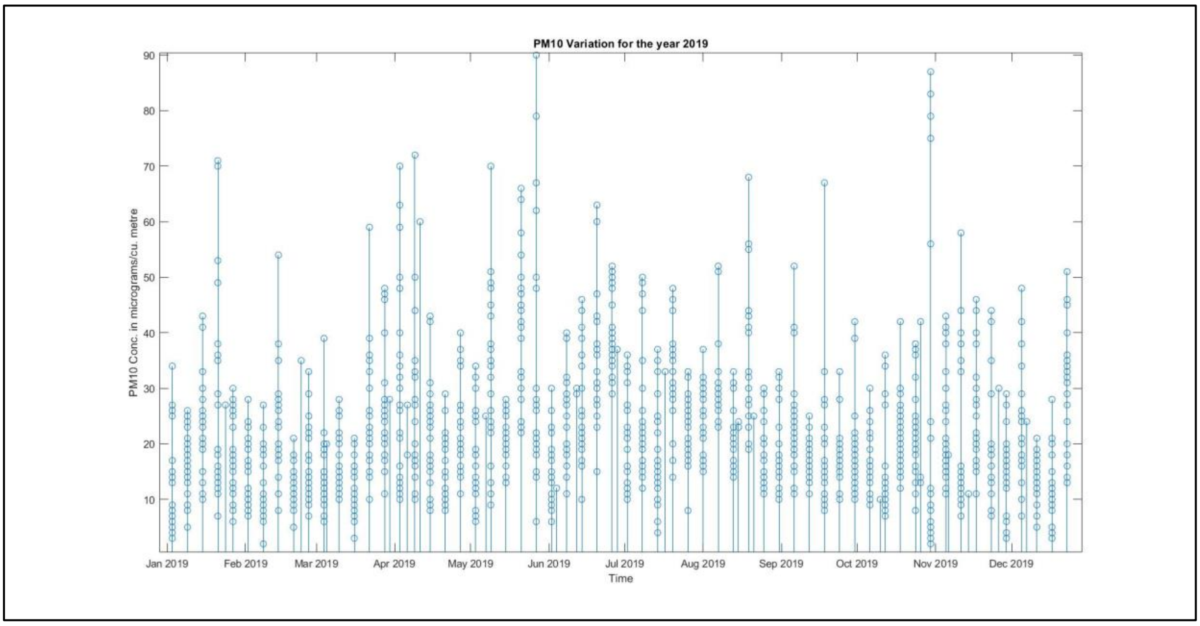


We can barely see any difference in the two maps. COVID lockdown does not always decrease the pollution levels—if not imposed in synchronisation.

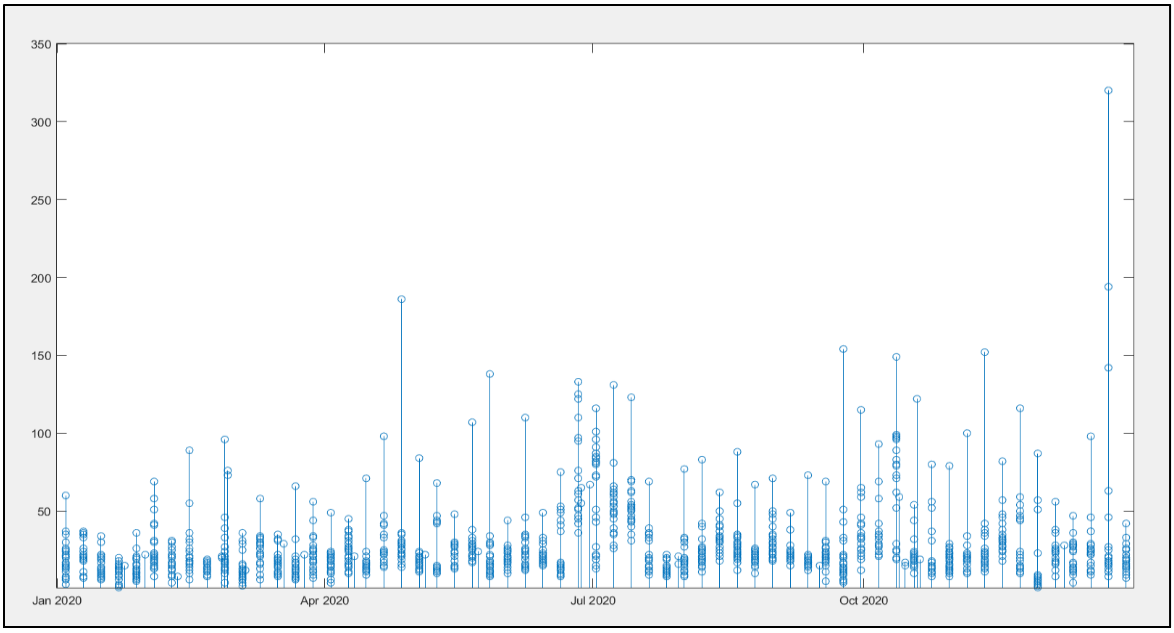
Time-series Analysis:

We can see an increase in the average and the maximum value of PM10 concentration from 2019 to 2020.

(a) Stem-plot showing the variation of PM10 with time for 2019.



(c) Stem-plot showing the variation of PM10 with time for 2020.



Conclusion

COVID lockdown has numerous impacts on our lives and also the environment. The pertaining fuzzy-hypothesis that—air pollution must decrease due to the lockdown is proved wrong here. If we incorporate the effects of other parameters like—synchronization of lockdown, natural calamities, location of industries and power-plants from the control area, we would find intimidating results.

In California, pollution has increased by a huge amount due to the wildfire. Counties such as -San Bernardino, Modoc, remained unaffected by lockdown as well as fire.

When it comes about Pennsylvania, due to the proper and synchronised lockdown, particulate pollution is decreased, but only by a smaller amount—6%. Still, the counties, like Allegheny, which are highly polluted since decades, remained nearly unaffected.

For Texas, the data reveals that there is increase in the PM level by 3%, because of the unsynchronised and short-duration lockdown continuation of daily commuting, industrial activities, petrochemical industries and refineries.

Limitations and Future works

Heatmap doesn't always tells the variation, rigorously. So, we have tried to incorporate interpolation methods to our project. But, since our data has around 57000 control points, it wasn't able to handle such a big data. For a i5 processor, it is not possible in an Open-source software. So, we have kept it for our future study.

This was our limitation, otherwise we would have shown IDW and kriging which would also produce a trend surface.

Some interpolation techniques, which can be performed in MATLAB, requires a proficiency in programming languages like Python and a deeper insight of Statistics. So, with the course of our study we will implement in our works, further.

Bibliography

[1] Examining Effects of the COVID-19 National Lockdown on Ambient Air Quality Across Urban India-*Navinya et. al.*

[2] Air quality index variation before and after the onset of COVID-19 pandemic: a comprehensive study on 87 capital, industrial and polluted cities of the world — *Mohammad Sarmadi1 , Sajjad Rahimi1, Mina Rezaei , Daryoush Sanaei and Mostafa Dianatinasab*

[3] www.epa.gov

[4] www.washingtonpost.com

[5] Download Daily Data | US EPA

[6] www.pnas.org

[7] www.theguardian.com

[8] Earthdata (nasa.gov)

[9] www.pasda.psu.edu

[10] www.cen.acs.org

[11] Mapping PM2.5 Air Pollution in Texas—*Zhipeng Xing*

[12] www.catalog.data.gov

[13] www.austintexas.gov

[14] Vicente et. al. 2013

[15] Groβ et. al. 2013

[20] www.census.gov

[21] www.wikipedia.com

[22] https://in.mathworks.com/learn/tutorials/matlab-onramp.html

[23] Documentation for QGIS 3.22 — QGIS Documentation documentation

[24] Principles of Geographic Information Systems—*Peter A. Burrough*

[25] The ArcGIS® Book

[26] GIS by *Kang-Tsung-Chang*

[27] Air pollution (who.int)

Appendix

Links to the self-furnished QGIS maps-

https://github.com/himanshuraj-iitb/ES216_TermProject