# ST 306 Mini Project

Tharindu Kariyawasam

January 2024

# 1 Introduction

Air pollution is a critical issue not only for London but also for the whole world. Several causes contribute to air pollution including industrialization, urbanization, and motor vehicle emissions. Polluted air has led to severe health issues. Therefore governments and health organizations are taking various initiatives to prevent air pollution. Particulate matter, Carbon monoxide, Ozone, Nitrogen dioxide, and Sulfur dioxide are key air pollutants.

In this project, we are going to perform an exploratory data analysis to discover the trends in air quality across London City from 01/01/2022 to 31/12/2023. We have used data obtained from 36 air monitoring sites located in London.

## 1.1 Data

We have used two data sets for the project. "london_local_data_2022" which contains hourly measurements of pollutants and "london_local_sites" which contains data about the monitoring sites. Both data sets are in csv format.

In the "london_local_data_2022" file it has recorded the name of the site and code of the site which measurements are taken, date and time correspond to measurements and hourly measurements of $NO_2$, $NO_x$, NO, $O_3$, $SO_2$, $PM_{10}$, $PM_{2.5}$. In the "london_local_sites" file it has recorded data under the following variables.

- code - identification code of the site
- site - name of the site
- latitude - latitude of the site
- longitude - longitude of the site
- parameter name - name of the substance measured

# 2 Literature Review

Analyzing air quality timely is vital in several aspects. Governments need to make policies to protect the air. People also need to have an idea of how the chemical composition of the air surrounding them has formed. Industries also have to consider air quality when controlling waste emissions. Therefore universities, research institutions, government agencies, and even individuals have done analyses on various scales about the air quality.

Luis A. Gil-Alana, OlaOluwa S. Yaya, and Nieves Carmona-Gonzalez have published a paper Gil-Alana, Yaya, and Carmona-Gonzalez (2020) on Air Quality in London: Evidence of Persistence, Seasonality, and Trends. There they

have investigated issues such as persistence, seasonality, and time trends. They have used a dataset that consists of roadside and background air quality for seven standard pollutants: nitric oxide (NO), nitrogen dioxide ($NO_2$), oxides of nitrogen ($NO_x$), ozone ($O_3$), particulate matter ($PM_{10}$ and $PM_{2.5}$) and sulfur dioxide ($SO_2$). They have observed a large degree of heterogeneity across pollutants and a persistent behavior based on a long memory pattern. The findings in the paper have served as a guide to air pollution management and European Union (EU) policymakers.

Mackline Ninsiima, Alex Ndyabakira, Sarah Zalwango, Richard Migisha, Daniel Kadobera, Claire Biribawa, Lilian Bulage, Alex Riolexus Ario, Julie R. Harris, Daniel, Okello Ayen have published a paper Ninsiima et al. (2023) on Spatio–temporal trends of air quality, Kampala City, Uganda, 2020–2022. They have analyses about particulate matter ($PM_{2.5}$).

Karn Vohra, Eloise A. Marais, Shannen Suckra, Louisa Kramer, William J. Bloss, Ravi Sahu, Abhishek Gaur, Sachchida N. Tripathi, Martin Van Damme, Lieven Clarisse, and Pierre-F. Coheur have published a paper Vohra et al. (2021) on Long-term trends in air quality in major cities in the UK and India.

In this medium article, Anuebunwa (2022) authors have performed an Exploratory Data Analysis of U.S. Pollution Data.

# 3 Results and Discussion

The analysis was conducted based on data collected from 36 air monitoring sites located in London. The following map 1 indicates the locations of monitoring sites.


Figure 1: Monitoring Centers

One main objective of this EDA is to find out how pollutant levels change throughout the year. First, consider how pollutant levels varied in Brent area.

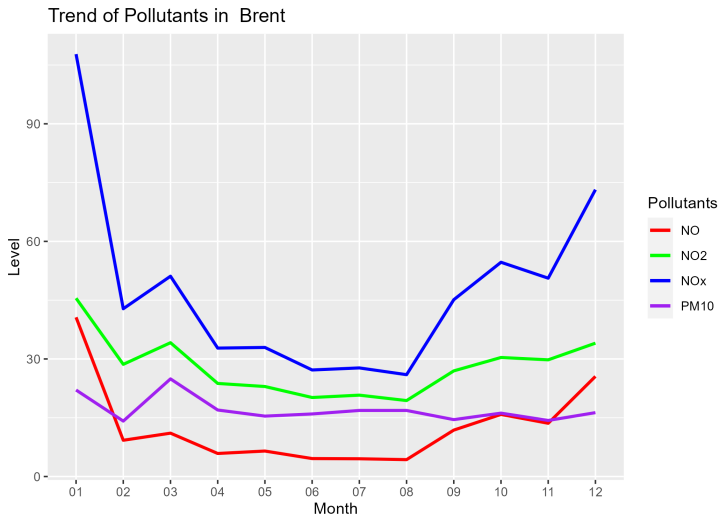


Figure 2: Pollutant Levels in Brent

According to the graph 2 $NO_x$ is the pollutant with the highest concentration NO is the pollutant with the lowest concentration. For $NO_x$, $NO_2$, and NO January is the month with the highest concentration and December is the month with the second highest concentration. August is the month with the least concentration. For $PM_{10}$ March is the month with the highest concentration and February is the month with the lowest concentration.

As for the area Brent we can visualize how pollutant levels vary throughout the year for other areas as well.

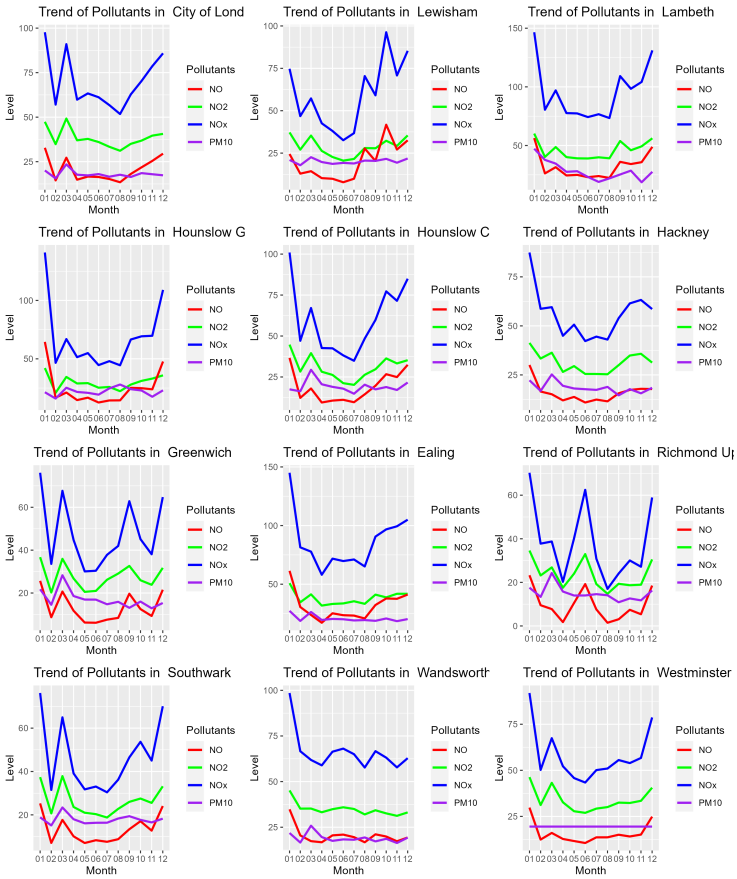


Figure 3: Pollutant Levels in London

From the above set of plots 3, we can notice that $NO_x$ is the pollutant with the highest level and NO, $PM_{10}$ is the pollutant with the lowest level. Similar for the area Brent January and December are the months with the most and second most polluted air.

Now let's consider the concentration of different pollutants across the areas.

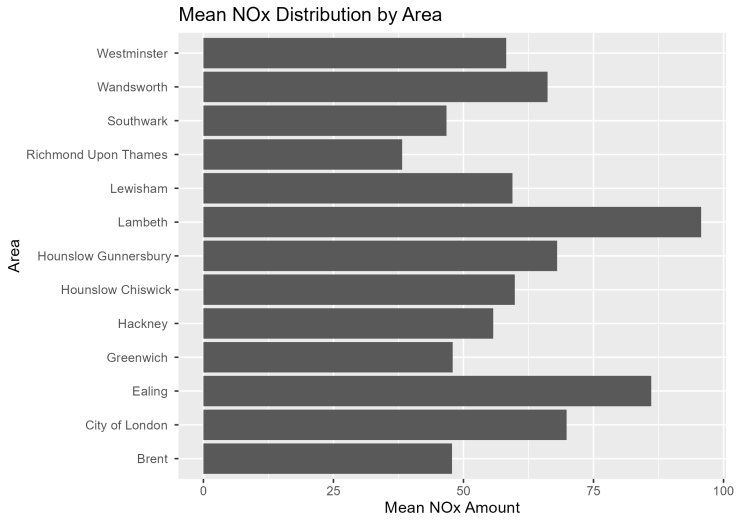


Figure 4: Level of $NO_x$ across the areas

According to the above graph, we can see that Lambeth is the area with the highest level of $NO_x$ and Richmond Upon Thames is the area with the lowest level of $NO_x$.

Similarly, we obtain plots for other areas also.

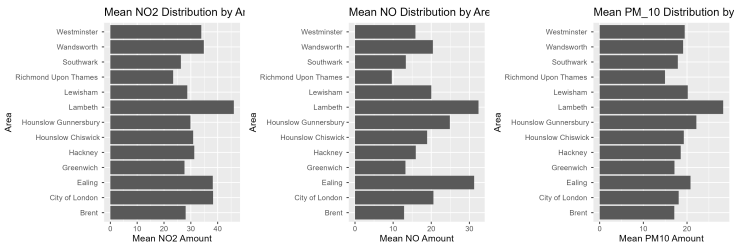


Figure 5: Pollutants distributions across the areas

For $NO_2$, NO, and $PM_{10}$ Lambeth is the area with the highest level and Richmond Upon Thames is the area with the lowest level.

For $O_3$ data is available only for three areas, Hackney, Richmond Upon Thames, and Southwark. The below graph shows the distribution of $O_3$ among those areas. For $PM_{2.5}$ data is available only for two areas, Hackney, and Westminster.

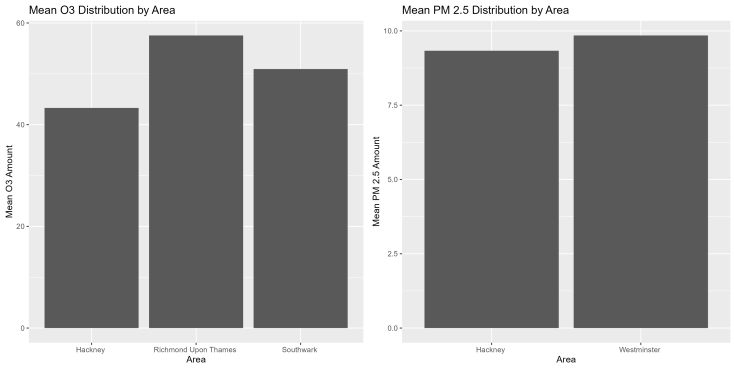


Figure 6: $O_3$ and $PM_{2.5}$ distributions across the areas

Let's consider hourly measurements of pollutants.

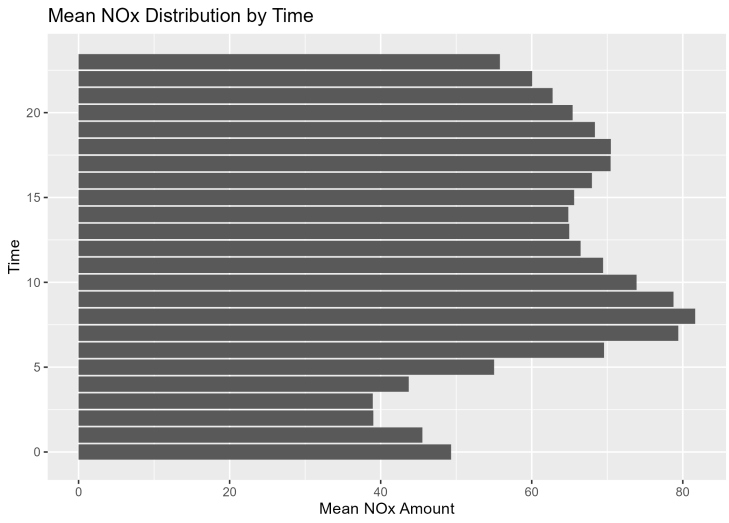


Figure 7: Hourly Distribution $NO_x$

According to the above graph 7 8.00 a.m. is the hour with the highest $NO_x$ levels. 2.00 a.m. and 3.00 a.m. are the hours with the lowest levels.

For the other pollutants also hourly measurements can be plotted.

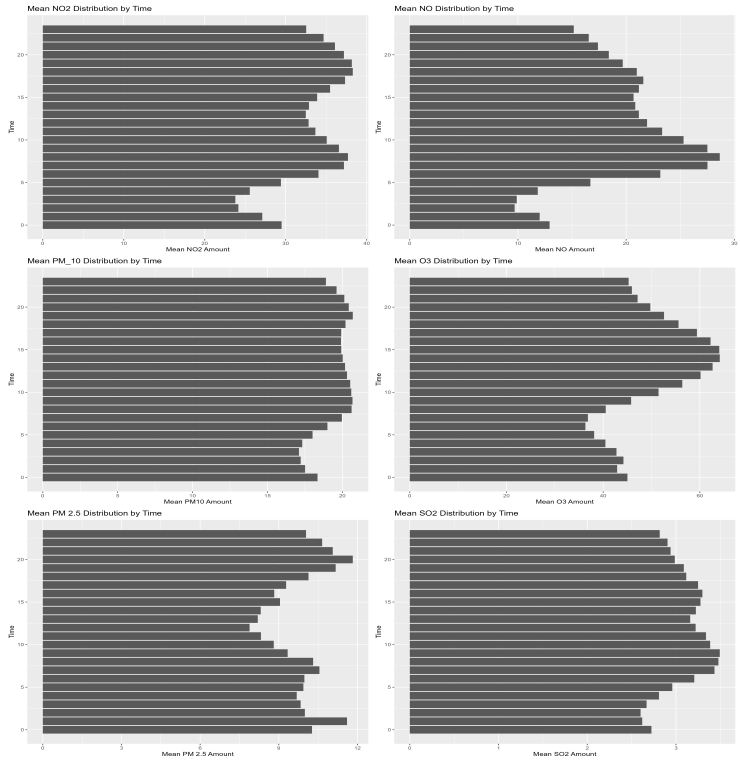


Figure 8: Hourly Distribution Pollutants

For $NO_2$, NO, $PM_{10}$, and $SO_2$ highest values were recorded around 8.00 a.m. The lowest values were recorded around 2.00 a.m. and 3.00 a.m. For $O_3$ highest value was recorded around 2.00 p.m. and 3.00 p.m. Lowest values were recorded at 6.00 a.m. and 7.00 a.m. For $PM_{2.5}$ highest values were recorded around 1.00 a.m. and 8.00 p.m. Lowest values were recorded around noon.

Let's examine the correlation between pollutant levels.

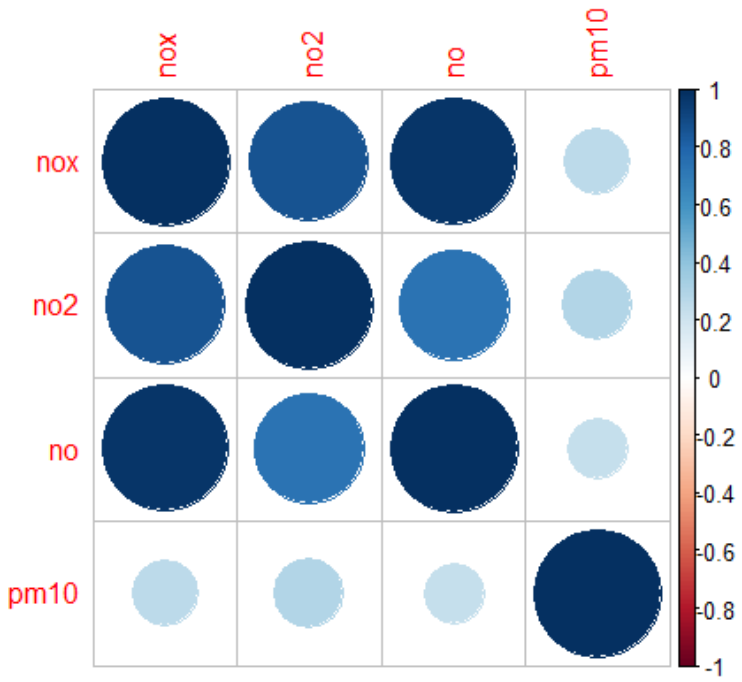


Figure 9: Correlations Between Pollutants

By the above graph, 9 we can notice that there is a perfect positive correlation between $NO_x$ and NO. Also strong positive correlation between $NO_2$ and $NO_x$, moderate correlation between $NO_2$ and NO. $PM_{10}$ has a weak positive correlation with other pollutants.

# 4 Conclusion

By the above graphs, we can conclude that Richmond Upon Thames has a comparingly good air composition. As this area borders the River Thames it has more eco-friendly surroundings. It is one of the richest boroughs in London in terms of the total area of green space. Therefore air quality is maintained at a good level by the natural purification process done by plants.

The City of London and Westminster are the most polluted areas in London due to heavy traffic and industrial activities.

When we consider the months January and December are the months with rising pollutant levels. It is mainly due to Winter weather conditions. Energy

consumption is high in those months. Usage of heating systems is increased in homes and buildings. Some areas use wood stoves and fireplaces therefore wood burning and biomass combustion also get increased.

When it comes to hourly trends around 8.00 a.m. is the time when pollutant levels go high. It is mainly because that is the morning rush hour. As people are commuting to work and school traffic emissions is high at that time.

# Bibliography

Anuebunwa, C. (2022). Exploratory data analysis of u.s. pollution data.

Gil-Alana, L. A., Yaya, O. S., & Carmona-Gonzalez, N. (2020). Air quality in london: Evidence of persistence, seasonality and trends.

Ninsiima, M., Ndyabakira, A., Zalwango, S., Migisha, R., Kadobera, D., Biribawa, C., ... Ayen, O. (2023). Spatio–temporal trends of air quality, kampala city, uganda, 2020–2022.

Vohra, K., Marais, E. A., Suckra, S., Kramer, L., Bloss, W. J., Sahu, R., ... Coheur, P.-F. (2021). Long-term trends in air quality in major cities in the uk and india:a view from space.