ST306 - Mini Project

S18847

January 30, 2024

1 Introduction

In this study, we're looking at the air quality in London by checking data from 36 monitoring sites from January 1, 2022, to December 31, 2023. Our dataset, "london local data 2022," gives us hourly info on important air pollutants like NO2, NOx, NO, O3, SO2, PM10, and PM2.5, although sometimes we're missing data. The "london local sites" dataset adds important details for each site, like ID codes, names, and what substances are measured. Even though there are studies on London's air, our research helps by filling in gaps and giving a full look at air pollution trends.

Our study matters because it could guide health plans, city development, and environmental rules by showing where and when air pollution changes. We want to find patterns and high-pollution areas, making sure our findings are trustworthy by dealing with any missing data. To reach our goals, we use a strong plan that fixes missing info using stats. We also look at where the monitoring sites are to see if pollution changes in different places.

As we check the data, we're finding interesting patterns in pollution levels at different sites. Looking at how pollution changes over time helps us understand more about what affects air quality. Checking how different pollutants are connected gives us a better picture of what's going on. When we look at our results, we see some sites have more pollution all the time, telling us where we might need to act fast. These findings add to what we already know about London's air quality and help leaders make smart choices.

In closing, this study digs deep into London's air quality, dealing with missing data and giving useful insights. Our ideas for future research include watching pollution closely, helping places with lots of pollution, and planning cities in ways that keep the air clean. By working together, we hope to make London a healthier and greener place for everyone.

2 Literature Review

London is the largest and most populated city in the UK and has a history of severe pollution events such as the notorious 1952 London smog that precipitated an increased awareness of air quality issues within the city. Investigating the trends, seasonality and cyclic patterns exhibited by air pollutants gives insights into their sources and properties, and is important for health considerations and policy development. Bigi and Harrison (2010) analysed 13 years of hourly data from a central urban background site in London for particulate matter and gas phase pollutants in terms of long-term trends, annual, weekly and diurnal cycles. The analysis showed generally declining trends for all the pollutants considered, with the exception of O3 which exhibited a steady increase over the period. Clear seasonal variations were observed, with CO, NO, NO2 and SO2 showing a summer and winter maximum and a pattern associated with traffic emissions (for CO, NO and NO2). O3 showed a maximum in May and a minimum in winter, and the particle number count was at a minimum in August and a maximum in winter. Colette et al. (2011) investigated air quality trends in Europe over the past decade by looking at pollutants such as NO2, O3 and PM10 from urban background, suburban background and rural background stations. They observed a general decline in NO2 for the majority of the monitoring stations, with a slight increase of O3 observed (especially at urban sites) due to a decrease in NOx emissions. PM10 levels declined over the decade in UK and Germany. Analysis of 18 years of data from Fresno (California) using time series and multiple linear regression models showed that the concentrations of NOx, EC and ammonium nitrate had halved since 2000, but the PM2.5 levels had not declined significantly (de Foy and Schauer, 2019). Similarly, in Los

Angeles, an assessment of the effectiveness of regulations to reduce tail pipe emission was undertaken by investigating the trend in PM2.5 mass concentration and chemical species concentrations for the period 2005–2015 (Altuwayjiri et al., 2021). The study reported an overall significant downward trend in mass concentration of EC and OC (major contributors to the PM2.5 mass concentration). Data from 18 sites for the period 1999–2016 analysed for Seoul also showed a decrease in the long-term measurements of PM10 due to a reduction in the local source contribution, and an increase in O3 from local secondary production, with NO2 and SO2 not showing significant trends (Seo et al., 2018). This study also looked at short-term variability in pollutant concentrations, and was able to associate high PM10 and primary gaseous pollutant concentrations with migratory high-pressure systems that enhance regional transport and local accumulation during warmer periods.

This study examined data generated by different monitoring networks at the roadside London, Marylebone Road supersite, one of the most investigated roadside locations in Europe, with data from the North Kensington and the Westminster background sites utilized as required for comparison. Changes in the short and long-term trends of regulated pollutants, geographic source apportionment and roadside contribution increments were investigated. A similar analysis was also conducted for unregulated pollutant metrics (i.e. eBC, heavy metals, hydrocarbon and particle numbers). The results are consider as likely to be representative of heavily trafficked roadside locations across Europe, as over the relevant period, the UK was subject to EU Directives which applied also across many other countries in the region. This is especially true of vehicle emission standards which are applied to vehicles across Europe, even outside of the EU. Consequently, although there will be differences due to local source emissions and differing meteorology, there are broadly similar air pollution climates across Europe, especially in relation to road vehicle emissions.

3 Result And Discussion

3.1 Distribution Across the NOX,NO2.NO,PM(10),O3,PM(25),SO2

In this section we consider about the following criteria.

- **Concentration range** The x-axis represents the concentration levels of Gas, and the y-axis displays the frequency or count of observations within each concentration range. Peaks in the histogram indicate the most common concentration levels.
- **Central Tendency** Look for a central tendency or the main peak in the histogram. This point reflects the typical or average concentration of Considered variable in the data set.
- **Skewness** Assess the shape of the histogram. Symmetric,Positive and Negative
- **Outliers** Identify any outliers or unusual concentration levels that appear as isolated bars away from the main peak. These outliers may indicate specific events or locations with extreme.
- **Spread of data** The width of the bars and the overall spread of the histogram indicate the variability. A wider spread suggests greater variability across observations.

Variable	Typical Or Average concentration	Skewness	Outliers	Variability
NOx	central peak around 40 units	The positive skew indicates that the majority of monitoring sites have lower NOx concentrations, contributing to a rightward tail.	No significant outliers are apparent, as there are no isolated bars far from the main concentration range.	Higher Variability
NO_2	central peak around 40 units	The data appears relatively symmetric, suggesting a moderate spread of NO_2 concentrations.	No significant outliers are apparent	Lower dispersion data than NOx.
NO	central peak around 0 units	The positive skew indicates that the majority of monitoring sites have lower NO concentrations, contributing to a rightward tail.	No significant outliers are apparent	Lower dispersion data than NOx and higher variability than NO_2
PM_10	central peak around 12.5 units	The positive skew indicates that the majority of monitoring sites have lower $PM_(10)$ concentrations, contributing to a rightward tail.	No significant outliers are apparent	Lower dispersion data than NO_2.
O_3	central peak around 40 units	The positive skew indicates that the majority of monitoring sites have lower NO concentrations, contributing to a rightward tail.	No significant outliers are apparent	Lower dispersion data than NOx and higher variability than PM_10 And lower variability than NO_2
$PM_(2_5)$	central peak around 0 units	The positive skew indicates that the majority of monitoring sites have lower $PM_(2_5)$ concentrations, contributing to a rightward tail.	No significant outliers are apparent	Lower dispersion than $PM(_10)$.
SO_2	central peak around 3 units	The data appears relatively symmetric, suggesting a moderate spread of SO_2 concentrations.	No significant outliers are apparent	Lowest dispersion data considering other variable .

Table 1: Table of Distribution Across the NO_X,NO_2.NO,$PM_(10)$,O_3,$PM_(25)$,SO_2

3.2 Monthly Analysis Across the NOX,NO2.NO,PM(10),O3,PM(25),SO2

NO$_2$ vs Months

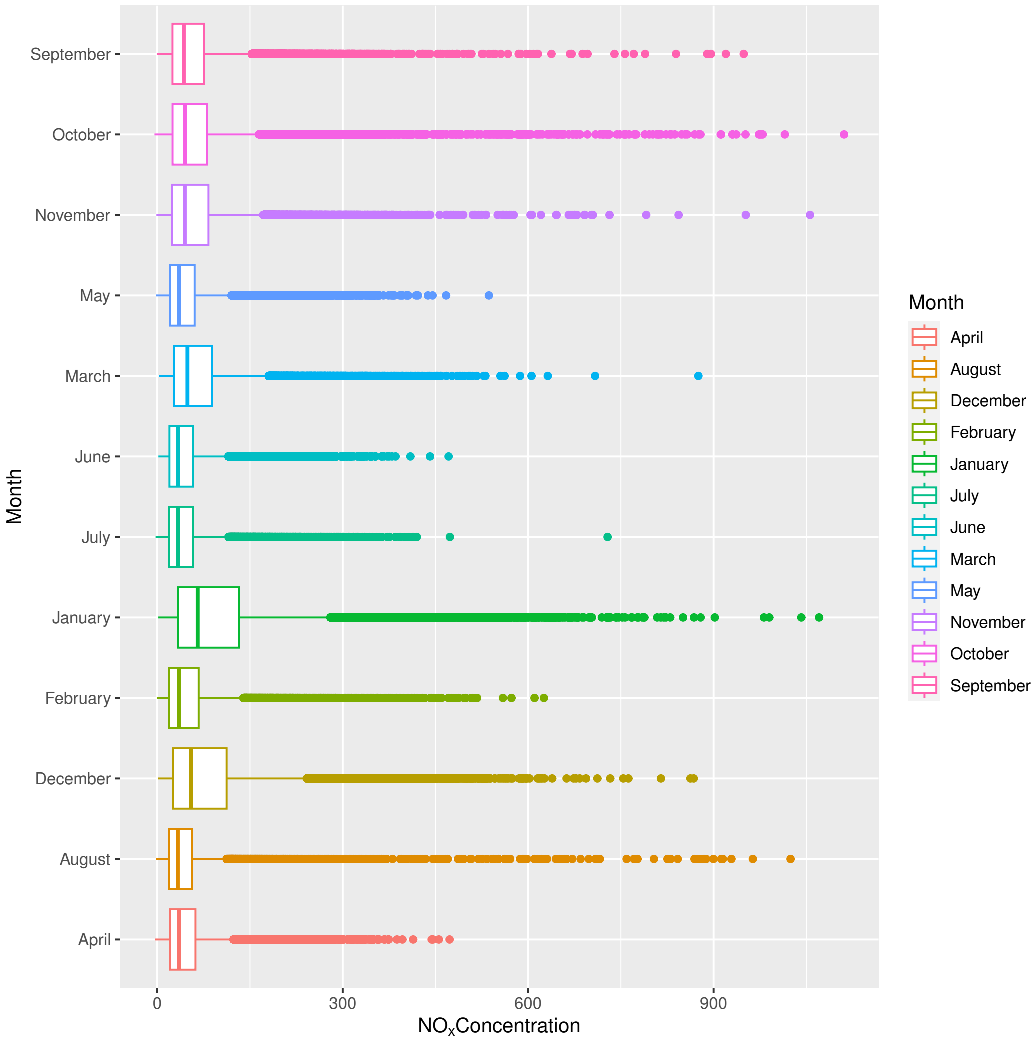

Figure 1: NO levels vs Months