

Mini project S18389

Sunera Hettiarachchi

28 January 2023

1 Introduction

This project aims to analyse and discover the trends in air quality across London city based on data collected at 36 air monitoring sites located in London from 01/01/2022 to 31/12/2023. The dataset - "london local data 2022" provides information about hourly measurement of the air pollutants NO₂, NO₂, NO_x, NO, O₃, SO₂, PM₁₀, PM_{2.5} In each site in each month. The dataset includes following variables.

- site - Name of the site
- code - Identification code of the site
- date - date that the data are collected
- nox - Amount of Nox
- no2 - Amount of No2
- no - Amount of No
- pm10 - Amount of Pm10
- o3 - Amount of o3
- pm25 - Amount of Pm25
- so2 - Amount of So2

This analyse aims to understand the facts like the air qualities of each site in each month, which air pollutant has the highest impact, what are the sites that can be vulnerable to be exposed, what are the sites that has the least air pollution, Which air pollutant has the highest effect, ect..

This data set consist of lot of missing values. Since a data set cannot be analysed with null values each null value has been replaced with 0. Also the "date" column has been divided into three new columns called "year", "date" and "time".

2 Literature review

Air pollution stands as a prominent environmental concern of considerable significance, has been extensively studied through various lenses. This literature review synthesizes findings from three key studies addressing different facets of air pollution.

One study focused on how people walking along a major road in Central London are exposed to harmful particles and gases like (PM_{2.5}), reflectance loss, ultrafine particle counts, and carbon monoxide (CO). This study conducted over a 12-day period, emphasized that the time of day and where people walk can affect how much pollution they breathe in.

Another study delved into the impact of a congestion charging scheme (CCS) introduced in central London in 2003, primarily exploring pollutant concentrations. Although inconclusive findings emerged from roadside monitors, background monitors suggested a decrease in nitric oxide (NO) and increases in nitrogen dioxide (NO₂) and ozone (O₃), both within and outside CCS hours. However, attributing these changes solely to the CCS was challenging due to concurrent traffic and emissions interventions.

Furthermore, a study investigated atypical weather patterns and their influence on mortality estimates related to sulfur dioxide (SO₂), CO, and smoke in Greater London. Colder, less windy days associated with higher pollutant levels were identified, with a notable increase in excess deaths during prolonged cold weather. The study emphasized the importance of considering weather patterns in mortality estimations, indicating potential associations with pollutants such as smoke, possibly acting as a surrogate for PM₁₀.

By putting these studies together, we can learn more about how air pollution affects people and what we can do about it. This review is all about sharing what we found in these studies to help us understand air pollution better.

3 Analysing the Dataset

3.1 Total Amount of Air Pollutants in Each Site in London

site	Code	Total amount of air pollutants	Percentage
City of London - Upper Thames Street	CT8	251.0	0.00077
Wandsworth - Putney	WA9	222050.8	0.68644
Ealing - Acton Vale	EI3	290426.5	0.89782
Wandsworth - Lavender Hill (Clapham Jct)	WAC	367405.8	1.13579
Lewisham - Deptford	LW5	442325.9	1.36740
Wandsworth - Putney High Street	WA7	524368.1	1.62103
Richmond Upon Thames - Barnes Wetlands	RI2	534336.7	1.65184
Westminster - Ebury Street (Grosvenor)	GV1	539722.3	1.66849
Richmond Upon Thames - Castelnau	RI1	596237.4	1.84320
City of London - Sir John Cass School	CT3	624412.9	1.93030
Westminster - Strand (Northbank BID)	NB1	653406.9	2.01994
Lewisham - New Cross	LW2	685837.4	2.12019
Wandsworth - Battersea	WAA	712895.8	2.20384
Westminster - Duke Street (Grosvenor)	GV2	714365.0	2.20838
Greenwich - Blackheath	GR7	859664.3	2.65756
Brent - John Keble Primary School	BT6	876809.3	2.71056
Westminster - Cavendish Square	WMC	916085.9	2.83198
Brent - ARK Franklin Primary Academy	BT8	946047.5	2.92460
Wandsworth - Putney High Street Facade	WA8	973624.6	3.00098
Southwark - Tower Bridge Road	SK8	978156.6	3.02387
Southwark - Elephant and Castle	SK6	1022606.7	3.16128
Westminster - Elizabeth Bridge	WMD	1064271.0	3.29008
Hounslow Gunnersbury	HS8	1109840.6	3.43095
Hounslow Chiswick	HS4	1149342.0	3.55307
Westminster - Oxford Street	WM6	1154628.1	3.56941
Lambeth - Bondway Interchange	LB5	1163040.4	3.59542
City of London - Beech Street	CT4	1196247.3	3.69807
Hackney - Old Street	HK6	1274900.9	3.94122
Westminster - Oxford Street East	WMB	1296355.5	4.00755
Ealing - Western Avenue	EI1	1322689.2	4.08895
Lewisham - Loampit Vale	LW4	1711118.6	5.28974
City of London - Walbrook Wharf	CT6	1755353.6	5.42649
Ealing - Hanger Lane Gyratory	EA6	2313841.5	7.15300
Lambeth - Brixton Road	LB4	2355163.5	7.28074

Table 1: **Total Amount of Air Pollutants in Each Site in London**

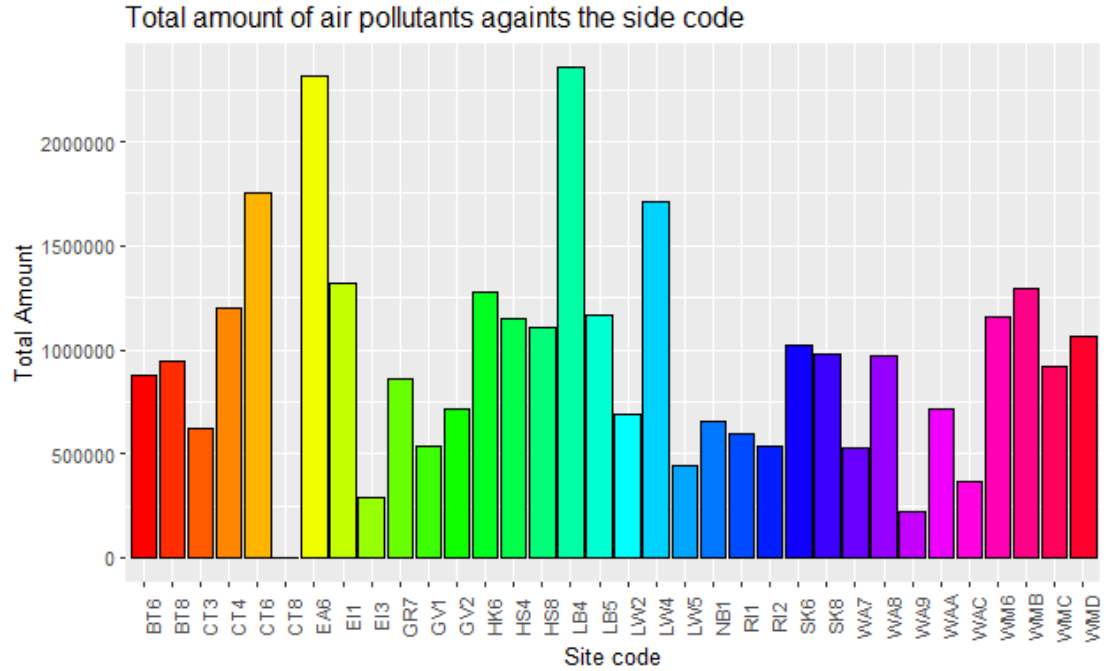


Figure 1: Graph of Total Amount of Air Pollutants in Each Site

This bar graph gives us the information about the total amount of air pollutants in each site. So according to the bar chart the site “Lambeth- Brixton Road” (LB4) has the highest amount of air pollution which is 2355163.5 where the site “Ealing- Hanger Lane Gyrotory” (EA6) has the 2nd highest amount which is 2313841.5. This implies that the air quality of these 2 sites are not good. So we cannot recommend these two sites to people for living as their air pollution is too high.

Also some sites such as “City of London - Walbrook Wharf”(CT6) and “Lewisham – Loampit Vale” (LW4) has considerable amount of air pollutants. So it is doubtful to recommend these areas to people for living.

Further we can see that the site “City of London – Upper thames street” (CT8) has the lowest amount of air pollution which is 251. This implies that the air quality of this area is way better when compared with the other areas. So we can highly recommend this site to people for living.

Also some areas like “Ealing - Western Avenue” (EI1), “Lewisham – Deptford” (LW5), “Wandsworth – Putney” (WA9), “Wandsworth - Lavender Hill (Clapham Jct)” (WAC) has comparatively low amount of air pollutants. So these areas have comparatively better air quality than the other areas. So these areas also can be recommended.

3.2 Total Amount of Air Pollutants and percentages in Each Month in Year 2022

month	Total amount of air pollutants	percentage
01	4596983	14.21
02	2338595	7.22
03	3261107	10.08
04	2310568	7.14
05	2226913	6.88
06	2011924	6.21
07	2061997	6.37
08	2176580	6.72
09	2591667	8.01
10	2831317	8.75
11	2579491	7.97
12	3360687	10.38

Table 2: Total Amount of Air Pollutants and percentages in Each Month

Pie Chart For Total amount of air pollutant for each month

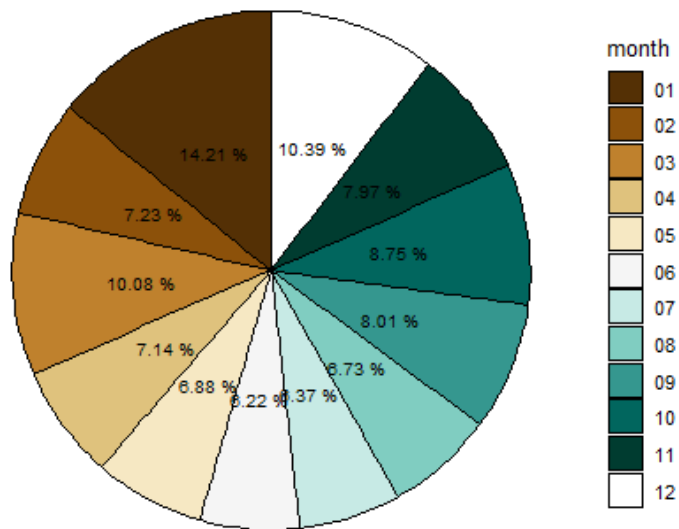


Figure 2: Pie chart of air pollutant for each month

This pie graph shows the percentages of total air pollutants in each month. According to the chart January has the highest amount of air pollution which is 14.21 of the total air pollution. Also March And December has higher percentages of air pollutants contributing 10.08 and 10.38 respectively. June has the lowest amount which is 6.21 of the total air pollution. Which means June is the month having the best air quality. According to the pie chart we can see that there is not a huge difference in the amount of air pollutants in the middle time of the year. But at the beginning of the year and at the end of the year the propotion of air pollutants are way more high than the other time periods of the year.

3.3 Total Amount of Air Pollutants in London in Year 202

Air Pollutant	Total Amount
nox	15212981.0
no	4683212.0
no2	8029464.0
pm10	3259356.0
o3	959034.6
pm25	177702.6
s02	26079.7

Table 3: Total Amount of Air Pollutants in London in Year 2022

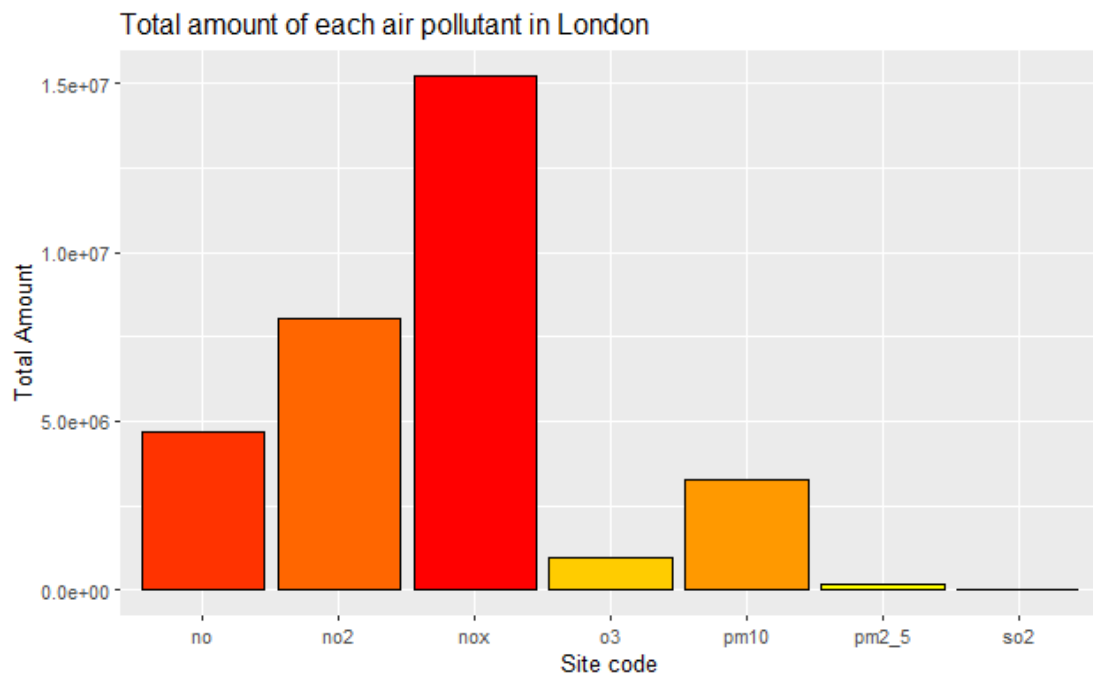


Figure 3: Graph of nox vs month

This bar graph shows the total amount of each pollutant in the London area. So according to the bar chart nox has the highest amount which is 15212981. Also no2 has a considerable amount of 8029464 as well. From the air pollutants considered so2 has the lowest amount which is 26079.7. Further pm2.5 and o3 also has a comparatively lower amount which is 177702.6 and 959034.6 respectively. The air pollutants no and pm10 have a moderate effect to the total air pollution in the London City. So we can recommend that it's better to identify the activities that produce the air pollutants which have higher effects to the total pollution and check whether there is any chance to reduce the amount of them.

3.4 Nox amount vs month

According to the graph4 among the air pollutants considered Nox is the most impactful air pollutant in the London city. So let's consider the effect of nox in each site here.

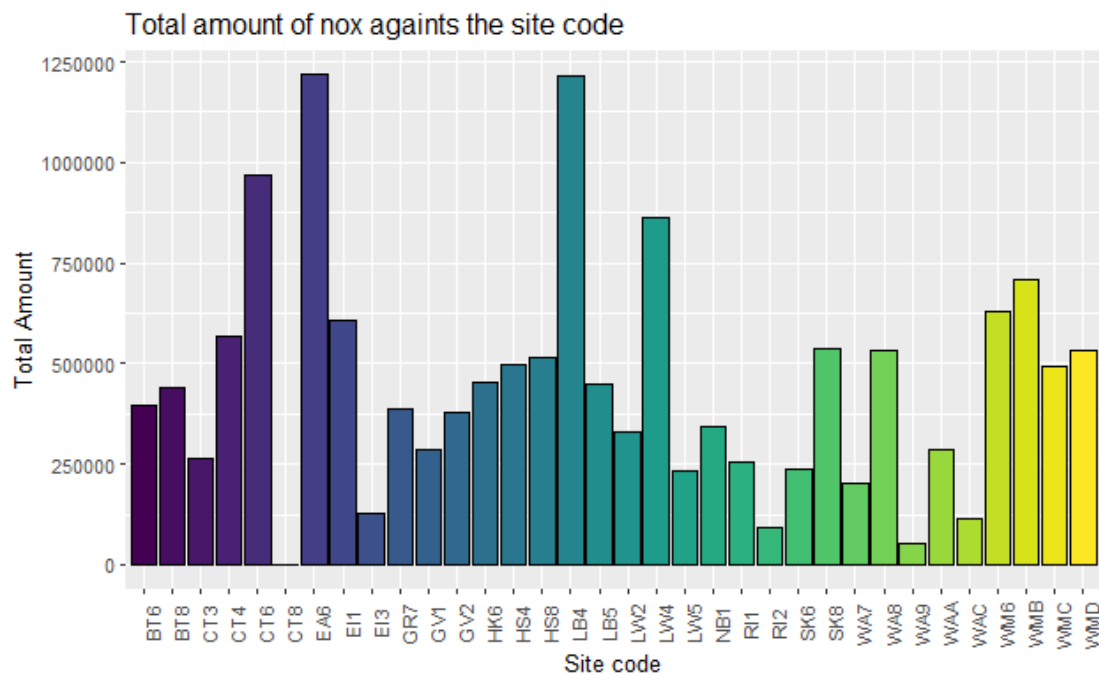


Figure 4: Graph of nox vs month

So according to the above graph Ealing - Hanger Lane Gyratory is the site which has the highest effect from Nox which is 1217362.4. Also Lambeth - Brixton Road is highly affected by Nox as well which is slightly less than Ealing - Hanger Lane Gyratory. Surprisingly City of London - Upper Thames Street is not affected by Nox.

site	code	Nox Amount
City of London - Upper Thames Street	CT8	0
Wandsworth - Putney	WA9	52280.0
Ealing - Acton Vale	EI3	129099.8
Wandsworth - Lavender Hill (Clapham Jct)	WAC	114150.0
Lewisham - Deptford	LW5	232732.1
Wandsworth - Putney High Street	WA7	203009.5
Richmond Upon Thames - Barnes Wetlands	RI2	94210.9
Westminster - Ebury Street (Grosvenor)	GV1	285298.0
Richmond Upon Thames - Castelnau	RI1	254088.3
City of London - Sir John Cass School	CT3	264457.8
Westminster - Strand (Northbank BID)	NB1	344383.8
Lewisham - New Cross	LW2	331326.4
Wandsworth - Battersea	WAA	287618.7
Westminster - Duke Street (Grosvenor)	GV2	376660.0
Greenwich - Blackheath	GR7	387277.6
Brent - John Keble Primary School	BT6	394464.3
Westminster - Cavendish Square	WMC	494311.6
Brent - ARK Franklin Primary Academy	BT8	438862.4
Wandsworth - Putney High Street Facade	WA8	532045.4
Southwark - Tower Bridge Road	SK8	537185.8
Southwark - Elephant and Castle	SK6	238127.1
Westminster - Elizabeth Bridge	WMD	531886.0
Hounslow Gunnersbury	HS8	514355.8
Hounslow Chiswick	HS4	496943.4
Westminster - Oxford Street	WM6	630550.4
Lambeth - Bondway Interchange	LB5	449608.9
City of London - Beech Street	CT4	566653.3
Hackney - Old Street	HK6	452667.5
Westminster - Oxford Street East	WMB	708121.2
Ealing - Western Avenue	EI1	606755.2
Lewisham - Loampit Vale	LW4	863372.4
City of London - Walbrook Wharf	CT6	969462.0
Ealing - Hanger Lane Gyratory	EA6	1217362.4
Lambeth - Brixton Road	LB4	1213652.7

Table 4: Nox amount vs month

3.5 Average of the each air pollutant in each site

Code	Nox	No2	No	Pm10	o3	Pm25	So2
BT6	45.03	27.4	11.49	16.16	0	0	0
BT8	50.09	28.47	14.10	15.31	0	0	0
CT3	30.18	22.54	4.98	13.55	0	0	0
CT4	64.68	39.81	16.22	15.83	0	0	0
CT6	110.66	50.41	39.29	0	0	0	0
CT8	0	0	0	14.46	0	0	0
EA6	138.96	51.20	57.24	16.72	0	0	0
EI1	69.26	34.89	22.41	24.41	0	0	0
EI3	14.73	9.79	3.22	5.39	0	0	0
GR7	44.20	25.67	12.08	16.16	0	0	0
GV1	32.56	22.65	6.39	0	0	0	0
GV2	42.99	30.21	8.33	0	0	0	0
HK6	51.67	29.13	14.69	15.42	33.60	0.99	0
HS4	56.91	29.29	18.0	17.58	0	9.82	0
HS8	58.71	24.95	22.01	21.0	0	0	0
LB4	138.84	62.57	49.54	18.18	0	0	0
LB5	51.32	28.66	14.77	35.02	0	0	2.97
LW2	37.82	21.2	10.71	8.55	0	0	0
LW4	98.55	37.75	39.65	19.36	0	0	0

3.6 Average of the each air pollutant in each month

Month	Nox	No2	No	Pm10	o3	Pm25	So2
January	93.41941	41.34034	33.96563	14.635235	2.753780	1.0211225	0.09902656
February	49.12588	27.61302	14.03037	10.293610	3.802647	0.4849838	0.10562320
March	59.95566	35.39443	16.06954	16.860472	3.447654	0.9985337	0.09822010
April	42.95529	26.35800	10.78881	11.516204	4.886250	0.6547980	0.08675926
May	41.16936	23.68809	11.36439	10.612554	3.186965	0.5524120	0.02012620
June	37.98728	22.10423	10.31972	10.408287	3.464743	0.3199916	0.07268519
July	37.69970	21.91110	10.29343	9.445890	4.073316	0.5730186	0.07063356
August	39.37705	22.31225	11.31407	9.965661	5.353148	0.4157706	0.09389866
September	51.54078	27.20160	15.87162	9.878910	3.953531	0.5494529	0.08101010
October	55.9261	27.07763	18.80923	11.130661	1.912296	0.3633920	0.10037879
November	53.09075	27.17107	16.89533	9.395593	1.387900	0.5429293	0.08084596
December	68.13152	30.86570	24.30425	10.938869	1.651116	0.875851	0.11362822

Table 6: Average Of the each air pollutant in each month

3.7 Information about some special cases

Air Pollutant	Site Name	Amount
Nox	City of London - Walbrook Wharf	110.669178
	Ealing - Hanger Lane Gyratory	138.968310
	Lambeth – Brixton Road	138.544829
	Lewisham – Loampit Vale	98.558493
No2	City of London - Walbrook Wharf	50.417306
	Ealing - Hanger Lane Gyratory	51.200685
	Lambeth – Brixton Road	62.576473
No	Ealing - Hanger Lane Gyratory	57.240422
	Lambeth – Brixton Road	49.545137
pm10	Lambeth - Bondway Interchange	35.021119
o3	Hackney – Old Street	33.60518
	Southwark – Elephant and Castle	49.71543
	Richmond Upon Thames-Barnes Wetlands	26.15822
pm25	Hounslow Chiswick	9.8218049
	Westminster - Elizabeth Bridge	9.5013699
so2	Lambeth - Bondway Interchange	2.977135

Table 7: Some special cases that has to be considered

The Table 7 is about some special cases that has to be considered. So above table provides some important information about the sites that each pollutant is appeared in large amounts when compared to the other sites. For instance the pollutant so2 is only appeared in Lambeth - Bondway Interchange. So we need to find what are the activities that caused so2 to be appeared in this area and check whether we can reduce the amount of so2 in that area.

4 Conclusions

Using the above data analysis we can conclude that the “City of London - Upper Thames Street” (CT8) is the site with the least amount of air pollution in London in the year 2022. Hence this site has the best air quality from the 34 sites considered here. pm10 is the only air pollutant occurred in this area. So if we can find what are the activities that occur pm10 in this area then we can try to reduce the amount of pm10 occurred here and make “City of London - Upper Thames Street” a air pollutant free area. Also some areas like “Ealing - Western Avenue” (EI1), “Lewisham – Deptford” (LW5), “Wandsworth – Putney” (WA9), “Wandsworth - Lavender Hill (Clapham Jct)” (WAC) has comparatively low amount of air pollutants. So these area have comparatively better air quality than the other areas.

When considering the sites with highest air pollution we have clear evidence to say that the site “Lambeth- Brixton Road” (LB4) had the highest amount of air pollution which was 2355163.5. So this site has the worst air quality among the 34 sites considered. Also it is noticeable that the site “Ealing- Hanger Lane Gyratory” (EA6) was also a highly air polluted area whose total air pollution is 2313841.5. So this was slightly less than “Lambeth- Brixton Road” (LB4). So definitely we need to take some action to reduce the air pollution in these areas.

Also using above data analysis we can conclude that January had the worst air quality which is 14.21 from the total air pollution. Also considerable amount of air pollution recorded in the months March and December which are 10.08 and 10.38 from the total air pollution respectively. Further June has the best air quality which is 6.21 of the total air pollution. It is noticeable that there was not a huge difference in the amount of air pollutants in the middle time of the year. But at the beginning of the year and at the end of the year the proportion of air pollutants are way more high than the other time periods of the year.

Moreover Highest nox having site was “Ealing - Western Avenue” (EI1) while the highest no2 having site was “Lambeth- Brixton Road” (LB4). Also it is noticeable that “Ealing- Hanger Lane Gyratory” (EA6) has the highest amount of No. Further “Lambeth - Bondway Interchange” has the highest amount of pm10. Moreover the highest amount of O3 having site was Southwark - Elephant and Castle while the highest amount of pm2-5 having site was Hounslow Gunnersbury. Also the highest amount of so2 having site was Lambeth - Bondway Interchange.

5 Discussion

The data set "london local data 2022" is a huge data set with 289069 observations. So we may face some difficulties when analysing a huge data set like this. Moreover it is noticeable that each and every observation has at least one missing value. So we cannot analyse this data set without cleaning. So as a solution every NA value was replaced by 0.

Sometimes analysing the data set using tables is hard to understand. So as an alternative we can use graphs to get a better view and understanding about the data set. Here we have used graphs like bargraph and pie graph. Further we can use graphs like histograms, boxplots, scatterplots, frequency polygons etc. as alternatives.

6 References