

Using Curated Datasets

London Air Quality Network

Royal
Geographical
Society
with IBG

Advancing geography
and geographical learning



<https://www.londonair.org.uk>

What is London Air Quality Network?

Established in 1993, London Air Quality Network is an independent body that collates and assesses air pollution data across London. London Air is run by the Environmental Research Group at King's College London who then use the data to inform environmental and health policymakers. They also use the data to help local authorities fulfil their obligations to review air quality regularly under the 1995 Environment Act.



How do I access the data?

The above link is for the homepage of London Air Quality Network. To access particular data, one should visit the **Tools** menu and choose the **Data Download** option. This takes you to a part of the website where one can search for data for a particular monitoring station (of which there are over 120 in active use across all 32 London boroughs) or analyse data by searching for a **species** of pollutant (listed on left). Historical data is also archived and available from sites of monitoring stations that are no longer active.

Species of pollutants available

Carbon monoxide (CO)

Measured in ppm (parts per million) or mg/m^3 . Toxic in quantities above 35ppm.

Nitrogen dioxide (NO₂)

Measured in ppb or $\mu\text{g}/\text{m}^3$. Causes irritation to respiratory system at 10ppm.

Nitric oxide (NO)

Measured in ppb or $\mu\text{g}/\text{m}^3$. Dangerous if continuous exposure to levels of 25ppm or more.

Ozone (O₃)

Measured in ppb or $\mu\text{g}/\text{m}^3$. Irritates respiratory system.

PM^{2.5} particulates (fine particulate matter)

Measured in mg/m^3 . Causes irritation to eyes, nose, throat and lungs.

PM¹⁰ particulates (coarse particulate matter)

Measured in $\mu\text{g}/\text{m}^3$. Causes irritation to eyes, nose, throat and lungs.

Sulphur dioxide (SO₂)

Measured in ppb or $\mu\text{g}/\text{m}^3$. Irritates respiratory system and causes respiratory diseases.

Benzene (C₆H₆)

Measured in ppb or $\mu\text{g}/\text{m}^3$. Linked to birth abnormalities and higher cancer risk.

pollutant (listed on left). Historical data is also archived and available from sites of monitoring stations that are no longer active.

Once a filter has been applied, the screen that follows will give you the various options of data that are then available. Not all species of pollutants have data collected about them at all monitoring sites and the London Air site automatically filters the data available to you based on this. You can also select the **time period** you are interested in and the frequency of the readings provided. This is known as the **averaging period** and can be plotted in 15 minute, hourly, 8 hourly or 24 hourly chunks. You also have the option of simply seeing the daily mean for that species. Selecting the **plot graph** button presents you with a line graph of the result. If multiple species or multiple monitoring sites are chosen, a key is given to allow for the distinction between the different factors. At the bottom of the graph is the option to download the data in CSV format, which would allow for different levels of data analysis to be possible.



How can I use this in my teaching?

The study of pollutants forms an integral part of any environmental geography teaching as well as information central to any discussion of the challenges of living in an increasingly urban-centred world. The impact that human behaviour has on these spaces can also be explored through a study of health and the associated links to people's socio-economic background.

To be able to put data to different forms or species of pollutant spatially moves students away from more generic and simple interpretations of pollution and its causes and instead asks them to question more deeply how a city such as London might try to manage this very real and very dangerous problem. Older students may wish to examine the role of air pollution monitoring in wider levels of governance and how London Air's work fits into wider monitoring at EU level and for the World Health Organisation.

Students can also question the manner in which air pollution data is collected and whether the current geographical spread of air quality monitoring stations is sufficient or varied enough to draw meaningful conclusions about pollution across the UK capital. Students can investigate how the roadside air quality monitoring stations work and why a significant level of the data is only partially collected with regards to different pollutant species.



Curriculum Links

This curated dataset links to a number of parts of the National Curriculum and is relevant to GCSE and A Level Specifications.

Key Stage Three: An understanding, through the study of place, of the key processes in urbanisation and the use of natural resources.

An understanding of how human processes change environments.

GCSE: An understanding of the interactions between people and environments as well as the change in processes over space.

A knowledge and understanding of the UK's geography, including its environmental challenges.

A detailed study of issues related to the sustainable use of resources and the effects of rapid urbanisation.

An examination of one major city in an economically advanced country, and the challenges arising from urban change.

A Level: A knowledge of the pathways and processes which controls the cycling of carbon, including fossil fuel combustion.

An understanding of how interactions between different scales are fundamental to an understanding of global governance.

The following specifications make particular reference to air pollution and its management:

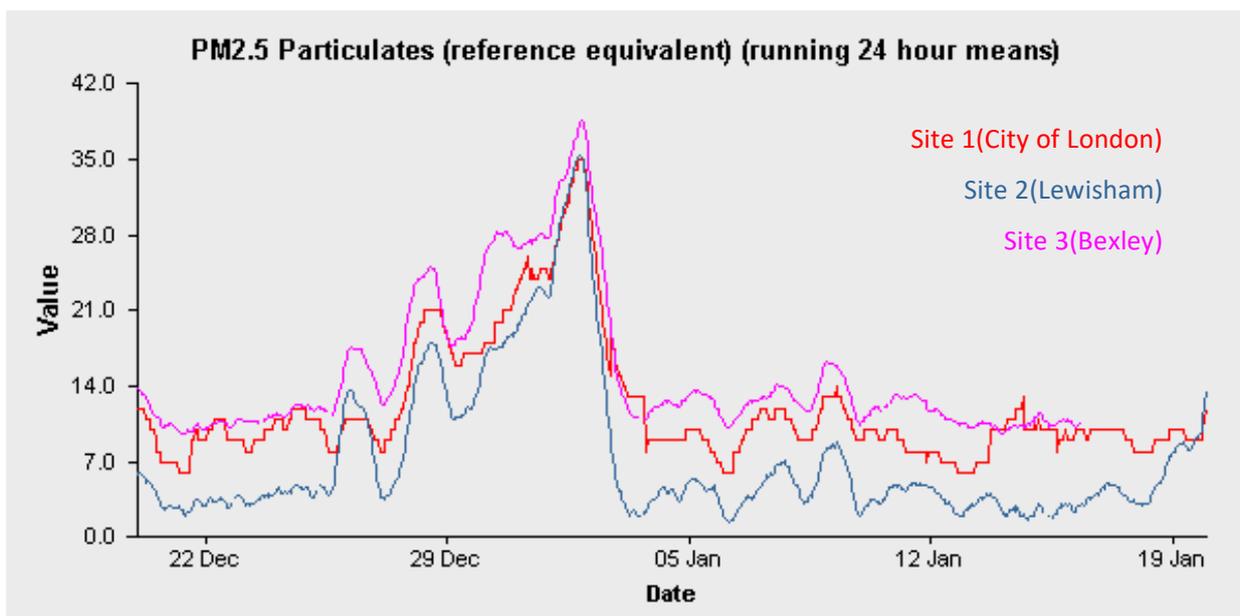
GCSE:		A Level:	
Edexcel A	Edexcel B	AQA	Edexcel
Cambridge iGCSE		Eduqas	OCR



An example data walk-through

A student is wishing to study how pollutant levels vary across London. They have an understanding that the more dense the living and working conditions, as well as the more dense the road network, the greater the likely volume of air pollution being produced from the combustion engines of vehicles. The student also understands how the built environment may theoretically have an effect on the concentration and accumulation of air pollution. Therefore, the student wishes to study a site in the centre of London, another on the furthest outskirts of the city and a third in between these two sites.

First, the student went to the **Monitoring Menu** on the main site and select **Monitoring Site**. This shows a map of all the monitoring sites in use by London Air. Here there is also the option to filter the sites on the map so that it only shows sites that monitor particular species of pollutant. The student was most interested in fine particulates so selected **PM2.5 particulates**. From the map they could then select three sites that fitted the geographical spread they were looking for: **City of London - Farringdon Street** in the centre of London, **Lewisham - Deptford** which is further out and **Bexley - Slade Green** which is on the outskirts of the monitored area. Using the **Data Download Tool** the students selected the **PM2.5 particulates** filter again and in the first to third site, selected each of the sites they had chosen. The student only wished to view data for a one month period and selected **20th December 2019 to 20th January 2020**. As they wanted to simply see the changes in the general trend over that time period they chose the **Running 24 hour** averaging period. They then plotted the graph.



From the graph, the student could see that actually the site further from the centre of London (Bexley - Slade Green) almost always had the highest concentration of PM2.5 particulate, followed by City of London - Farringdon Street. By downloading the data, the student could get a more accurate picture of the actual numbers involved. The maximum PM2.5 particulate level recorded between 20th December 2019 and 20th January 2020 was $38.6 \mu\text{g}/\text{m}^3$ (Bexley - Slade Green on 1st January) while the minimum recorded was $1.5 \mu\text{g}/\text{m}^3$ (Lewisham - Deptford on 6th January). The World Health Organisation suggests that PM^{2.5} particulate levels in healthy cities should reach no more than $25 \mu\text{g}/\text{m}^3$. From the data, the student recognised that at least one of the three London sites they were studying had exceeded that level on five days out of the month under investigation.



Suggested delivery activities

Urban Transects

Students can create geographically meaningful transects through London on a map and plot, showing how air pollution might change along those transects. The transects could take a traditional spatial account of the city and move from the centre to the outskirts. However, students could also devise their own forms of transect - lines that travel through areas of increasing or decreasing levels of wealth, for example, or similar socio-economic statuses such as general health indicators, using secondary data research to populate a data table. With the help of GIS packages such as ArcGIS, students would be able to plot the two factors on the same digital map. This would help them to see if there is relationship between air pollution levels and other factors, leading them to further think about possible causes for relationships.



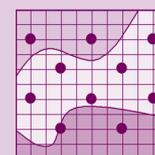
Living Graphs

With the use of long term secondary weather data for London, students can compare pollution level graphs over a long time scale with graphs that show the weather conditions (such as wind speed and temperature) in those locations. Students may be able to find a pattern between these two sets of data. Using forthcoming weather forecasts for a coming week, students can then try to estimate how pollution levels in London might change in that time to coincide with the weather conditions. A separate group of students might investigate social factors and whether other variables (such as day of the week or school holiday periods) affect how much pollution is produced. Students can create a 'living graph' for a pollution week in London with different annotations coming off the graph showing possible reasons for the changes they see.



Critique and Tweak

Students can examine the spread of air pollution monitoring stations across London. They may be able to identify monitoring hot spots (where there appears to be a proliferation of monitoring stations) and monitoring dead zones (where there appear to be few or none). Students can assess the extent to which the spread of air monitoring sites is fair and then question how the position of the monitoring sites may affect the nature of the air pollution data that London Air is able to collect. Students can think about sampling strategies that London Air might deploy if they were to start their monitoring scheme from scratch and the extent to which a stratified sampling framework may be needed to really understand the nature of air pollution in a city like London.



Interquartile Ranges

The Interquartile Range (IQR) of some data is a measure of the extent of its 'spread' when placed in numerical order. Calculating the interquartile range removes the data which is most outlying, (that in the first and fourth quartile, which are at each extreme of the range) and so is thought of as a truer representation of the spread of the data than the raw data alone.

Students can study the interquartile range for different pollutants at one site or for a single pollutant at multiple sites or in multiple time frames. They can hypothesise which sets of data are likely to have the larger spread and then carry out the test to find out if they are correct. A guide to interquartile ranges appears on the next page.

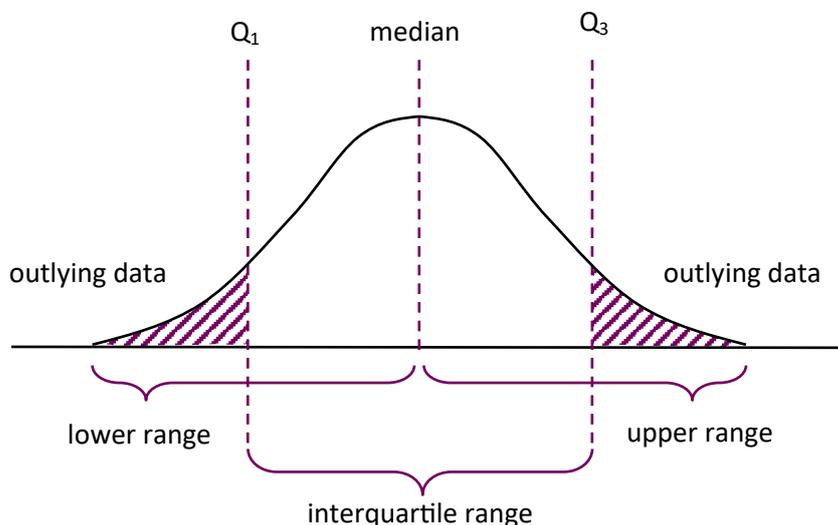




A Guide to Calculating Interquartile Ranges using Pollutant Data

The interquartile range for a set of data is the amount of spread or dispersion there is when the data is placed in numerical order. However, it is more than just the difference between the top and bottom values as it only looks at the data that is found around the median value. This means that outlying data (that in each of the two extremes) is removed from the measurement.

Q_1 (the lower quartile) is the median value of the lower range while Q_3 (the upper quartile) is the median value of the upper range. Values either side of these quartiles are outlying data and not included in the calculation of the interquartile range. Using the data from the London Air Quality Network, students can compare the interquartile range for pollutants recorded every hour on a weekday (where one might expect there to be a large spread of values to account for rush hours) to those of a weekend (where one might expect the spread of values to be smaller) for a particular location (in this case, City of London - Farringdon Street).



Worked example:

The relevant data was extracted from the London Air Quality Network site and placed in numerical order.

PM ^{2.5} hourly readings ($\mu\text{g}/\text{m}^3$)																								
Sun 19th Jan 2020	5	6	6	6	6	7	7	7	7	7	7	7	8	9	11	11	12	13	13	14	16	24	31	38
Mon 20th Jan 2020	20	21	21	23	23	23	23	23	24	25	25	25	25	26	26	26	27	27	28	29	29	31	33	38

The median value (in this case, as there are 24 values, the value at the 12.5th **position**) for each day was then calculated:

Sunday: median value = 7.5

Monday: median value = 25

Using the following equation, the **position** of the lower and upper quartiles was calculated:

$$Q_1 \text{ or } Q_3 = \frac{n + 1}{2} \qquad Q_1 \text{ or } Q_3 = \frac{13}{2} \qquad = 6.5^{\text{th}} \text{ position}$$

Sunday: lower quartile value = 7 median value = 7.5 upper quartile value = 13

Monday: lower quartile value = 23 median value = 25 upper quartile value = 27.5

The interquartile range (IQR) is calculated by subtracting the lower quartile value from the upper quartile value:

Sunday: lower quartile value = 7 upper quartile value = 13 Therefore IQR = 6

Monday: lower quartile value = 23 upper quartile value = 27.5 Therefore IQR = 4.5

Therefore, on these two days in London, the weekday (Monday) actually had a smaller spread of data than that of the weekend day (Sunday). This shows that there was more variety in pollutant levels on Sunday than on Monday.