**Project background & context**

**Measuring Air Quality**

Levels of various air borne pollutants such as Nitrogen Monoxide (NO), Nitrogen Dioxide (NO2) and particulate matter (also called particle pollution) are all major contributors to the measure of overall air quality.

For instance, NO2 is measured using micrograms in each cubic metre of air (㎍/m3). A microgram (㎍) is one millionth of a gram. A concentration of 1 ㎍/m3 means that one cubic metre of air contains one microgram of pollutant.

To protect our health, the UK Government sets two air quality objectives for NO2 in their [Air Quality Strategy](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/69336/pb12654-air-quality-strategy-vol1-070712.pdf)

1. The hourly objective, which is the concentration of NO2 in the air, averaged over a period of one hour.
2. The annual objective, which is the concentration of NO2 in the air, averaged over a period of a year.

The following table shows the colour encoding and the levels for Objective 1 above, the mean hourly ratio, adopted in the UK.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Index** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** |
| **Band** | **Low** | **Low** | **Low** | **Moderate** | **Moderate** | **Moderate** | **High** | **High** | **High** | **Very High** |
| **㎍/mÂ³** | **0-67** | **68-134** | **135-200** | **201-267** | **268-334** | **335-400** | **401-467** | **468-534** | **535-600** | **601 or more** |

Further details of colour encodings and health warnings can be found at the [DEFRA Site](https://uk-air.defra.gov.uk/air-pollution/daqi).

**The Input Data**

The following ZIP file provides data ranging from 1993 to 22 October 2023 taken from 19 monitoring stations in and around Bristol.

Download & save the data file:  [Air\_Quality\_Continous.zip](https://fetstudy.uwe.ac.uk/~p-chatterjee/2023-24/dmf/Air_Quality_Continuous.zip) (23.2 Mb)

Create a directory (folder) called “data” on your working machine and unzip the file there to **Air\_Quality\_Continuous.csv** (112 Mb).

Monitors may suffer downtime and may become defunct, so the data isn’t always complete for all stations.

Shown here is the first 8 lines of the file (cropped):

A screen shot of a computer

Description automatically generated

Note the following:

**There are 19 stations (monitors):**  
188 => 'AURN Bristol Centre', 51.4572041156,-2.58564914143  
203 => 'Brislington Depot', 51.4417471802,-2.55995583224  
206 => 'Rupert Street', 51.4554331987,-2.59626237324  
209 => 'IKEA M32', 51.4752847609,-2.56207998299  
213 => 'Old Market', 51.4560189999,-2.58348949026  
215 => 'Parson Street School', xxxxx.432675707,-2.60495665673  
228 => 'Temple Meads Station', 51.4488837041,-2.58447776241  
270 => 'Wells Road', 51.4278638883,-2.56374153315  
271 => 'Trailer Portway P&R', 51.4899934596,-2.68877856929x  
375 => 'Newfoundland Road Police Station', 51.4606738207,-2.58225341824  
395 => "Shiner's Garage", 51.4577930324,-2.56271419977  
452 => 'AURN St Pauls', 51.4628294172,-2.58454081635  
447 => 'Bath Road', 51.4425372726,-2.57137536073  
459 => 'Cheltenham Road \ Station Road', 51.4689385901,-2.5927241667  
463 => 'Fishponds Road', 51.4780449714,-2.53523027459  
481 => 'CREATE Centre Roof', 51.447213417,-2.62247405516  
500 => 'Temple Way', 51.4579497132 ,-2.5839890903  
501 => 'Colston Avenue', 51.4552693827,-2.59664882855  
672 => 'Marlborough Street', 51.4591419717,-2.59543271836

These monitors are spread across the four City of Bristol constituencies represented by the following Members of Parliament (MP's):

* Bristol East - Kerry McCarthy (MP);
* Bristol Northwest - Darren Jones (MP);
* Bristol South - Karin Smyth (MP); &
* Bristol West - Thangam Debbonaire (MP).

Each line represents one reading from a specific detector. Detectors take one reading every hour. If you examine the file using a programming editor, (Notepad++ can handle the job), you can see that the first row gives headers and there are another 1603492 (1.60 million+) rows (lines). There are 19 data items (columns) per line.

The schema for data (what each field represents) is given below:

| **measure** | **desc** | **unit** |
| --- | --- | --- |
| Date Time | Date and time of measurement | datetime |
| SiteID | Site ID for the station | integer |
| NOx | Concentration of oxides of nitrogen | ㎍/m3 |
| NO2 | Concentration of nitrogen dioxide | ㎍/m3 |
| NO | Concentration of nitric oxide | ㎍/m3 |
| PM10 | Concentration of particulate matter <10 micron diameter | ㎍/m3 |
| O3 | Concentration of ozone Concentration of non - volatile particulate matter <10 micron diameter | ㎍/m3 |
| Temperature | Air temperature | °C |
| ObjectID | Object (?) | Integer |
| ObjectID2 | Object (?) | Integer |
| NVPM10 | Concentration of non - volatile particulate matter <10 micron diameter | ㎍/m3 |
| VPM10 | Concentration of volatile particulate matter <10 micron diameter | ㎎/m3 |
| NVPM2.5 | Concentration of non volatile particulate matter <2.5 micron diameter | ㎍/m3 |
| PM2.5 | Concentration of particulate matter <2.5 micron diameter | ㎍/m3 |
| VPM2.5 | Concentration of volatile particulate matter <2.5 micron diameter | ㎍/m3 |
| CO | Concentration of carbon monoxide | ㎎/m3 |
| RH | Relative Humidity | % |
| Pressure | Air Pressure | mbar |
| SO2 | Concentration of sulphur dioxide | ㎍/m3 |