

ACCB5027 GIS - DS2306723

by David Sheppard

Submission date: 09-May-2026 10:11PM (UTC+0100)

Submission ID: 282903479

File name: 49805_David_Sheppard_ACCB5027_GIS_-_DS2306723_734550_42141056.docx (36.96M)

Word count: 3425

Character count: 21485

2026

ACCB5027: Geographic Information Systems

GEODEMOGRAPHIC ANALYSIS USING GIS SOFTWARE

DAVID SHEPPARD (2306723)

UWTSD | Swansea, UK

Table of Contents

1. Abstract	2
2. Introduction	3
3. Methodology.....	3
3.1 Data definitions.....	3
3.1.1 Data acquisition.....	3
3.1.2 Data collection	4
3.1.3 Data cleaning.....	10
3.1.4 Data manipulation and transformation	14
3.2 Thematic maps	15
3.2.1 Joining and preparing data for thematic mapping	16
3.2.2 Creation of thematic maps	18
3.2.3 Creating a thematic map for air quality	21
3.2.4 Final outputs	22
4. Outcomes	29
4.1 Findings from the thematic maps.....	29
4.2 Spatial query insights	30
5. Analysis	30
5.1 Principles of thematic mapping and visualisation	30
5.2 Challenges and limitations	30
6. Conclusions.....	31
7. Bibliography.....	32



1. Abstract

This study investigates the spatial relationships between historic coal mining activity, socio-economic deprivation, and respiratory health outcomes across Wales using Geographic Information Systems (GIS). A combination of spatial and attribute datasets, sourced from Welsh Government and UK platforms, was integrated at Lower Super Output Area (LSOA) and County scales. Data was cleaned, standardised, and transformed to ensure compatibility and reliability prior to analysis, with thematic mapping the primary method for visualising patterns and trends.

A series of thematic maps and spatial queries were produced to explore income, employment, and health deprivation alongside population age structure, air quality, and coalmine locations. The results indicate that former coal mining areas, particularly in Northeast and South Wales, are associated with higher levels of deprivation and poorer respiratory health outcomes compared to other regions. While improvements in air quality and respiratory health were observed between 2015 and 2024, higher baseline rates in historic mining areas suggest that the impacts of industrial legacy persist. Demographic patterns further highlight differences in population structure near former coalfields, indicating the need for cautious interpretation and further investigation.

The study demonstrates the effectiveness of GIS and thematic mapping in communicating complex geodemographic and health relationships, while acknowledging limitations related to data. Recommendations for further analysis include proximity-based health assessments, extended temporal analysis, incorporation of additional environmental variables, and the development of interactive GIS outputs. Overall, the project illustrates the value of GIS as a decision-support tool for understanding and addressing spatial health inequalities linked to industrial legacy.

2. Introduction

Geographic Information Systems (GIS) provide a robust framework for analysing geodemographic and health data, enabling the identification of spatial patterns, relationships, and potential hotspot areas. This project uses GIS software to investigate spatial associations between historic mining activity in Wales and respiratory health outcomes.

The study focuses on Wales, using national datasets and consistent administrative boundaries to ensure reliable spatial analysis. Lower Super Output Areas (LSOAs) are employed as the primary spatial unit, offering sufficient geographic detail for analysing deprivation and health patterns while remaining comparable across datasets.

The main objective of the project is to explore deprivation alongside the location of coalmines and any potential relationship with respiratory-related hospital admissions and mortality. The analysis combines deprivation indicators, demographic characteristics, and the spatial distribution of historic mines to examine whether areas of higher deprivation and mining hotspots correspond with poorer respiratory health outcomes. The scope of the study is influenced by data access constraints, including restrictions on granular health data, incomplete historic mine records, and practical time limitations.

3. Methodology

3.1 Data definitions

A Geographic Information System (GIS) integrates spatial (location-based) data with attribute data to analyse how social, environmental, and health patterns vary across space. Spatial data defines geographic location and relationships, while attribute data provide descriptive information such as socioeconomic, demographic, and health characteristics. When combined with contextual and temporal information, GIS enables effective spatial analysis, visualisation, and evidence-based interpretation [1].

3.1.1 Data acquisition



The reliability of a GIS analysis is strongly dependent on data quality and source credibility. For this project, datasets were obtained from government and recognised open-data platforms to ensure accuracy, consistency, and transparency. Metadata

accompanying each dataset was used to assess data origins, collection methods, assumptions, and limitations.

Data acquisition was closely followed by data cleaning, transformation, and standardisation to ensure compatibility across datasets. Key quality considerations included completeness, timeliness, positional accuracy, attribute accuracy, and temporal relevance. Ensuring consistent formats and units was essential for integrating multiple datasets and minimising analytical error [2].

Ethical and legal considerations also informed data selection and handling. All data were managed in accordance with data protection regulations, including GDPR, with restricted access to sensitive health information shaping the scale of analysis. Ultimately, all datasets were selected based on fitness for purpose, ensuring they directly supported the analytical objectives of the GIS project. [2].

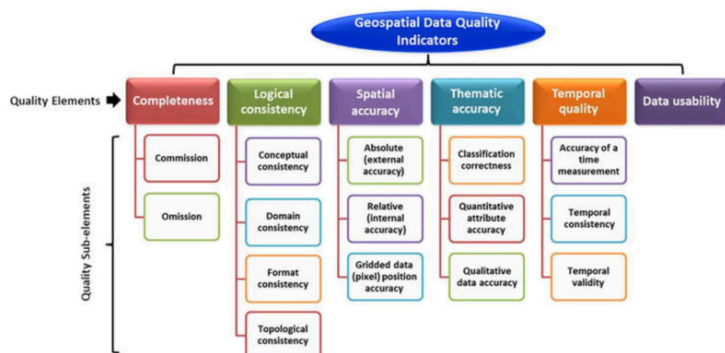


Figure 1: Geospatial data quality indicators [2]

3.1.2 Data collection

This project utilised a combination of spatial (vector) and non-spatial (attribute) datasets to analyse socio-economic deprivation, historic mining activity, and respiratory health outcomes across Wales.

Spatial Data

LSOA Boundaries (Wales) [3]

Polygon vector (shapefile) data defining Lower Super Output Areas across Wales. This dataset formed the core spatial framework of the project, providing a consistent

geographic unit to which deprivation, demographic, and health data were joined for mapping and analysis (Figure 2 and Figure 3).



Figure 2: LSOA boundary data download [3]

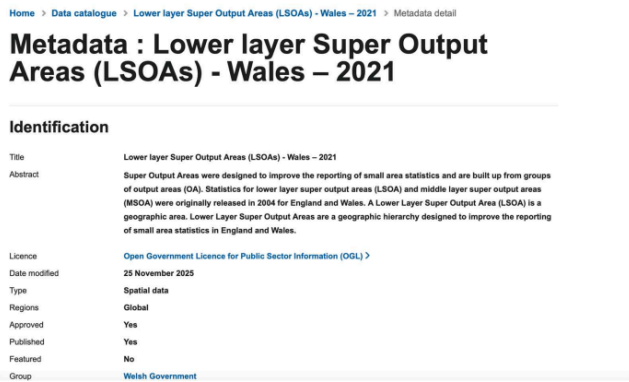


Figure 3: LSOA Metadata [3]

Welsh Index of Multiple Deprivation (WIMD) 2019 [4]

Shapefile datasets representing employment, income, and health deprivation at LSOA level. These indicators were used to identify spatial patterns of socio-economic vulnerability and to support comparative analysis with respiratory health outcomes (Figure 4 and Figure 5).

Welsh Index of Multiple Deprivation (WIMD) 2019

Welsh Government

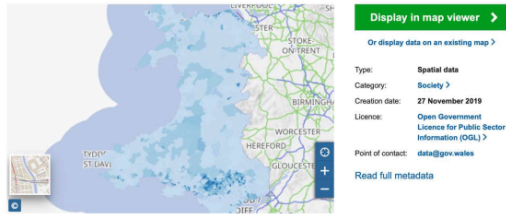


Figure 4: WIMD data source [4]

Identification

Title	Welsh Index of Multiple Deprivation (WIMD) 2019
Abstract	WIMD is the Welsh Government's official measure of relative deprivation for small areas in Wales. It is designed to identify those small areas where there are the highest concentrations of several different types of deprivation. Deprivation is the lack of access to opportunities and resources which we might expect in our society. This can be in terms of material goods or the ability of an individual to participate in the normal social life of the community. An index is a group of separate measurements which are combined into a single number. WIMD is currently made up of eight separate domains (or types) of deprivation: - Income - Employment - Health - Education - Access to Services - Housing - Community Safety - Physical Environment

Each of the domains include several indicators of deprivation. WIMD ranks all small areas in Wales from 1 (most

Figure 5: WIMD Metadata [4]

Historic Mine Locations [5]

Shapefile data representing the recorded locations of historic mines across Wales. This dataset supported analysis of the spatial distribution of mining activity and its potential association with deprivation and respiratory health patterns (Figure 6).

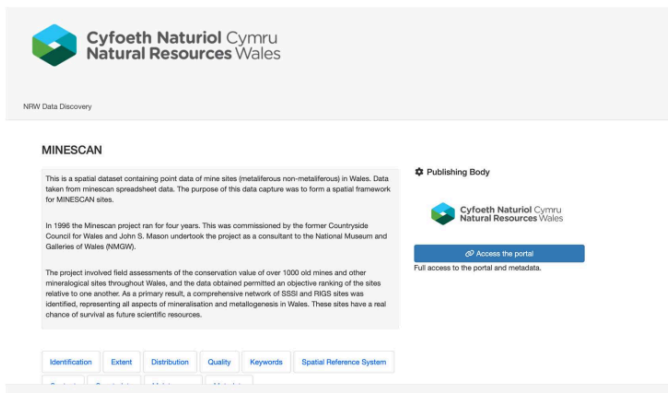


Figure 6: Mine location data [5]

Attribute (Tabular) Data

Population Age Profiles (0–15, 16–64, 65+) [6]

LSOA-level demographic data describing age structure, used to assess the potential influence of population composition on observed health outcomes (Figure 7 and Figure 8).

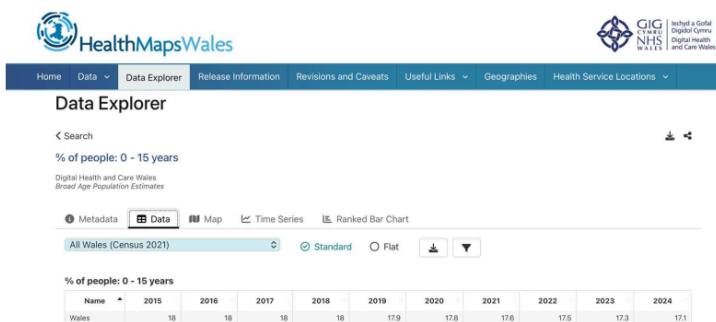


Figure 7: Population data - ages 0-15 [6]

Temporal
Census 2021-based geographies Wales, Local Health Board, Local Authority, MSOA and LSOA: 2015 - 2024.
Creator
Digital Health and Care Wales
Publisher
Digital Health and Care Wales
Methodology
Percentage is calculated as number of people in age group divided by total number of people in population
Clinical Codes
Not applicable for this measure
Last Published
5th March 2026
Additional Information
• Please note that the latest Health Maps Wales publication uses the rebased ONS population estimates following the 2021 Census. This has resulted in changes to the population estimates for 2012 - 2021. Please see www.ons.gov.uk for further information.
Numerator
Number of people in group
Denominator
Total number of people in population
Limits
Not applicable for this measure

Figure 8: Metadata population [6]

Respiratory Disease Hospital Admissions and Mortality Rates [7]

County-level data reporting respiratory-related hospital admissions and deaths by gender, used as indicators of respiratory health outcomes, classified using ICD-10 coding standards (Figure 9 and Figure 10).

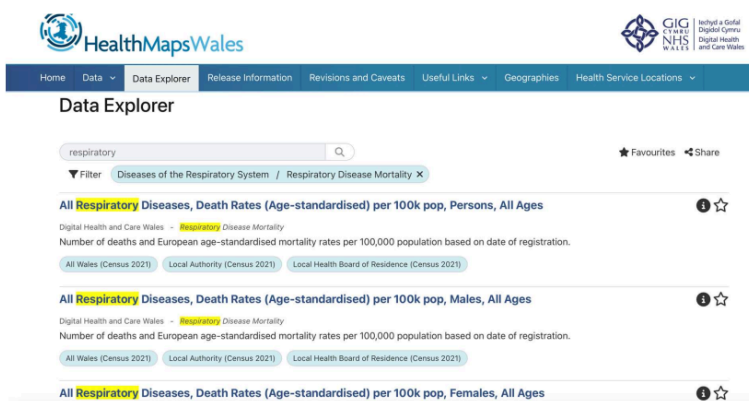


Figure 9: Respiratory disease data [7]

Additional Information

- There have been two ICD-10 coding changes for deaths covering the period published. The first change affects deaths registered from 1st January 2011 and the second affects deaths registered from 1st January 2014. Comparability ratios supplied by Public Health England (where available) have been used to adjust the number of deaths registered prior to 2011 for the first change and prior to 2014 for the second change. Where comparability ratios have been applied, the figures displayed are adjusted rates and unadjusted registered deaths (numerator). Please note, the following mortality data is unadjusted due to the absence of available comparability ratios at the time of publication: Heart Failure, All Heart Disease, Pedestrian Injury, Assault, Road Traffic Accidents, Avoidable Mortality. Please be aware of this when interpreting trends. These figures may be subject to revision if such comparability ratios become available. Further information on the changes to code-of-death coding can be found here: www.ons.gov.uk/ons/guide-method/user-guidance/health-and-life-events/Changes-to-cause-of-death-coding-in-England-and-Wales/index.html

Numerator

Annual District Deaths Extract (ADDE); Office for National Statistics (ONS), number of registered deaths.

Denominator

- Resident population, the denominators having been derived from mid-year estimates supplied by the Office for National Statistics (ONS).
- Please note that the latest Health Maps Wales publication uses the rebased ONS population estimates following the 2021 Census which has resulted in changes to the population estimates for 2012 – 2021. Please see www.ons.gov.uk for further information.

Limits

95% upper and lower confidence limits are shown. These were calculated using the method proposed by Dobson in the 1991 paper "Confidence Intervals for weighted sums of Poisson parameters" Statistics in Medicine 1991;10:457-462

Revisions Information

From 1 April 2019, health service provision for residents of Bridgend local authority moved from Abertawe Bro Morgannwg to Cwm Taf. For more information see the joint statement from Cwm Taf and Abertawe Bro Morgannwg University Health Boards here www.cwmtafbromorgannwg.wales/welsh-government-announce-decision-on-bridgend-boundary-change. The health board names have changed with Cwm Taf University Health Board becoming Cwm Taf Morgannwg University Health Board and Abertawe Bro Morgannwg University Health Board becoming Swansea Bay University Health Board. The data published is based on the new Health Board boundaries.

Figure 10: Respiratory disease metadata [7]

Air Quality (PM2.5) [8]

Data measuring fine particulate matter concentrations, included as a contextual environmental indicator relevant to respiratory health (Figure 11).

Department for Environment, Food & Rural Affairs

UK AIR

Air Information Resource

Search on UK Air

Home

Air Pollution

Data

Monitoring Networks

Library

Science & Research

LAQM

Home > Data Archive > Modelled air quality data > Modelled background pollution data

Modelled background pollution data

Introduction

Background pollution maps at 1x1 km resolution are modelled each year under Defra's Modelling of Ambient Air Quality (MAAQ) contract. These maps are used to provide policy support for Defra and to fulfil the UK's reporting obligations to Europe. They are also made publicly available for other research applications such as university studies, epidemiology and public health research, nature and conservation studies.

These maps should not be mistaken for the Local Authority-specific maps which are alternative maps based on the same model results as presented here but which provide source-sector splits and projections to future years by Local Authority to aid the Local Air Quality Management process.

Please acknowledge Defra as the source for this data if using the maps for your work.

Data Archive

PM_{2.5} Targets (PERT and AMCT)

Methodology for calculating progress towards the PM_{2.5} targets

Data Selector

Annual and Exceedance Statistics

Data availability

Figure 11: DEFRA air quality data [8]

Assessed but Excluded Data

Several datasets were reviewed but excluded due to limitations in spatial resolution, data suppression, or relevance to the study objectives. These included alcohol-related hospital admissions and mortality data, and address-based datasets requiring geocoding with limited analytical value at the study scale.

Overall, the final dataset contained a combination of polygon and point vector layers with associated LSOA and county level attribute data, enabling relevant spatial analysis.

3.1.3 Data cleaning

Once data sources were collected, a systematic review of file contents was undertaken. This step was essential, as unresolved formatting or value issues are likely to cause problems during GIS implementation. Several data-cleaning tasks were completed within .xlsx files prior to import, helping to optimise GIS performance and ensure compliance with principles of accuracy, relevance, and usability.

3.1.3.1 *Population data cleanse*

Population data was available at LSOA level; however, age group information was distributed across three separate Excel files (Figure 12). Importing these individually would have introduced unnecessary complexity and clutter into the GIS project. To address this, a single consolidated file was created containing all three age bands (0–15, 16–64, and 65+), limited to the year 2024 to provide a consistent basis for analysis (Figure 13).

Care was taken to ensure compatibility with the spatial boundary data. The Area field used in the population dataset was aligned exactly with the LSOA name field in the LSOA boundary layer, allowing for accurate joins within QGIS.

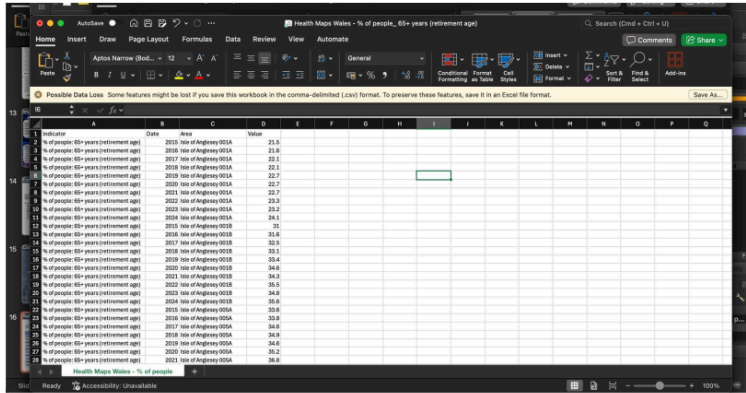


Figure 12: % LSOA population aged 65+

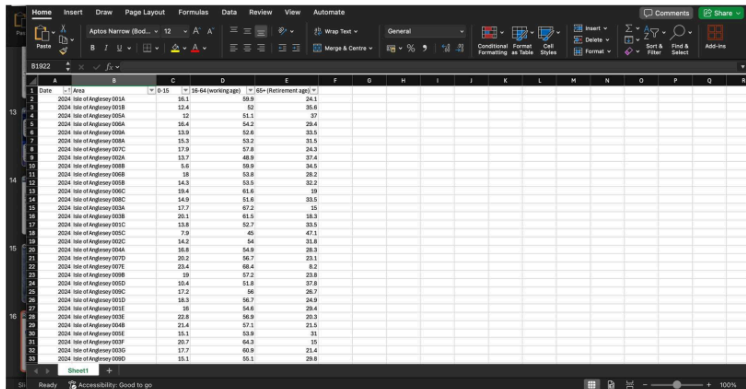


Figure 13: Combined population groups

3.1.3.2 Respiratory data cleanse

Respiratory hospital admission and mortality data was available only at County level and distributed across multiple files by year and category. To support effective analysis and reduce project complexity, these datasets were combined into a single file (Figure 14).

As part of this process, only the years 2015 and 2024 were retained to enable temporal comparison. County-level admission and mortality rates were aligned by year and gender, and additional calculations were carried out to derive percentage change between the two years. Due to later challenges encountered within QGIS, all

numeric fields were explicitly formatted as numbers during this stage to minimise import errors (Figure 15).

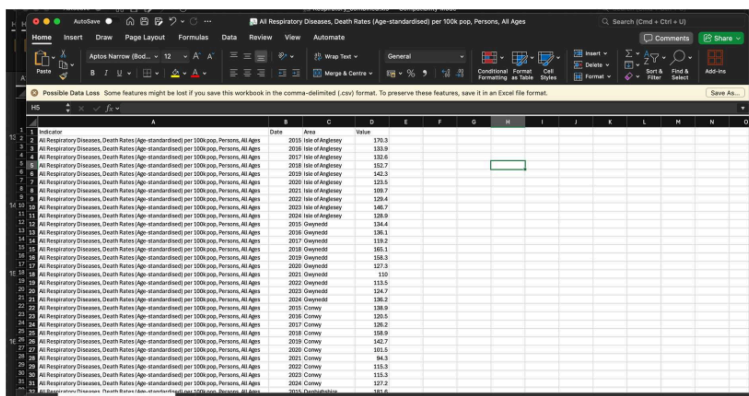


Figure 14: Death rates (persons) dataset

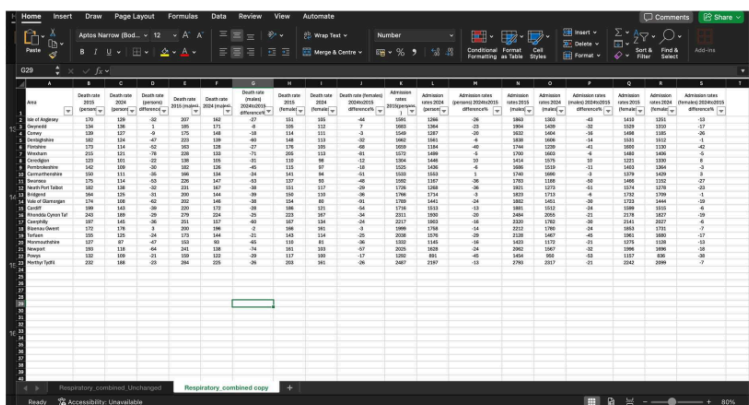


Figure 15: Single respiratory file

3.1.3.3 Air quality data cleanse

Air quality data (PM_{2.5}) was available at County level, with separate files for each year and combined coverage for England and Wales. In line with the respiratory analysis, 2015 and 2024 were selected as years of interest. The 2015 dataset required additional correction, as county names were inconsistent and Local Authority codes were missing. English counties were also removed prior to import (Figure 16).

A manual lookup process was applied in Excel to align the 2015 data with the 2024 dataset using consistent Welsh county naming. These data were consolidated into a single file, and percentage change in PM2.5 concentrations between 2015 and 2024 was calculated to support later analysis (Figure 17).

LA code	LA name	PM2.5 2015	PM2.5 2024	Difference
W06000001	Blanaugor	9.8204	1.214	7.6064
W06000002	Blanaugor	9.8204	1.214	7.6064
W06000003	Blanaugor	9.8204	1.214	7.6064
W06000004	Blanaugor	9.8204	1.214	7.6064
W06000005	Blanaugor	9.8204	1.214	7.6064
W06000006	Blanaugor	9.8204	1.214	7.6064
W06000007	Blanaugor	9.8204	1.214	7.6064
W06000008	Blanaugor	9.8204	1.214	7.6064
W06000009	Blanaugor	9.8204	1.214	7.6064
W06000010	Blanaugor	9.8204	1.214	7.6064
W06000011	Blanaugor	9.8204	1.214	7.6064
W06000012	Blanaugor	9.8204	1.214	7.6064
W06000013	Blanaugor	9.8204	1.214	7.6064
W06000014	Blanaugor	9.8204	1.214	7.6064
W06000015	Blanaugor	9.8204	1.214	7.6064
W06000016	Blanaugor	9.8204	1.214	7.6064
W06000017	Blanaugor	9.8204	1.214	7.6064
W06000018	Blanaugor	9.8204	1.214	7.6064
W06000019	Blanaugor	9.8204	1.214	7.6064
W06000020	Blanaugor	9.8204	1.214	7.6064
W06000021	Blanaugor	9.8204	1.214	7.6064
W06000022	Blanaugor	9.8204	1.214	7.6064
W06000023	Blanaugor	9.8204	1.214	7.6064
W06000024	Blanaugor	9.8204	1.214	7.6064
W06000025	Blanaugor	9.8204	1.214	7.6064
W06000026	Blanaugor	9.8204	1.214	7.6064
W06000027	Blanaugor	9.8204	1.214	7.6064
W06000028	Blanaugor	9.8204	1.214	7.6064
W06000029	Blanaugor	9.8204	1.214	7.6064
W06000030	Blanaugor	9.8204	1.214	7.6064
W06000031	Blanaugor	9.8204	1.214	7.6064
W06000032	Blanaugor	9.8204	1.214	7.6064
W06000033	Blanaugor	9.8204	1.214	7.6064
W06000034	Blanaugor	9.8204	1.214	7.6064
W06000035	Blanaugor	9.8204	1.214	7.6064
W06000036	Blanaugor	9.8204	1.214	7.6064
W06000037	Blanaugor	9.8204	1.214	7.6064
W06000038	Blanaugor	9.8204	1.214	7.6064
W06000039	Blanaugor	9.8204	1.214	7.6064
W06000040	Blanaugor	9.8204	1.214	7.6064
W06000041	Blanaugor	9.8204	1.214	7.6064
W06000042	Blanaugor	9.8204	1.214	7.6064
W06000043	Blanaugor	9.8204	1.214	7.6064
W06000044	Blanaugor	9.8204	1.214	7.6064
W06000045	Blanaugor	9.8204	1.214	7.6064
W06000046	Blanaugor	9.8204	1.214	7.6064
W06000047	Blanaugor	9.8204	1.214	7.6064
W06000048	Blanaugor	9.8204	1.214	7.6064
W06000049	Blanaugor	9.8204	1.214	7.6064
W06000050	Blanaugor	9.8204	1.214	7.6064

Figure 16: Air quality raw data

LA code	LA name	PM2.5 2015	PM2.5 2024	Difference
W06000001	Blanaugor	9.8204	1.214	7.6064
W06000002	Blanaugor	9.8204	1.214	7.6064
W06000003	Blanaugor	9.8204	1.214	7.6064
W06000004	Blanaugor	9.8204	1.214	7.6064
W06000005	Blanaugor	9.8204	1.214	7.6064
W06000006	Blanaugor	9.8204	1.214	7.6064
W06000007	Blanaugor	9.8204	1.214	7.6064
W06000008	Blanaugor	9.8204	1.214	7.6064
W06000009	Blanaugor	9.8204	1.214	7.6064
W06000010	Blanaugor	9.8204	1.214	7.6064
W06000011	Blanaugor	9.8204	1.214	7.6064
W06000012	Blanaugor	9.8204	1.214	7.6064
W06000013	Blanaugor	9.8204	1.214	7.6064
W06000014	Blanaugor	9.8204	1.214	7.6064
W06000015	Blanaugor	9.8204	1.214	7.6064
W06000016	Blanaugor	9.8204	1.214	7.6064
W06000017	Blanaugor	9.8204	1.214	7.6064
W06000018	Blanaugor	9.8204	1.214	7.6064
W06000019	Blanaugor	9.8204	1.214	7.6064
W06000020	Blanaugor	9.8204	1.214	7.6064
W06000021	Blanaugor	9.8204	1.214	7.6064
W06000022	Blanaugor	9.8204	1.214	7.6064
W06000023	Blanaugor	9.8204	1.214	7.6064
W06000024	Blanaugor	9.8204	1.214	7.6064
W06000025	Blanaugor	9.8204	1.214	7.6064
W06000026	Blanaugor	9.8204	1.214	7.6064
W06000027	Blanaugor	9.8204	1.214	7.6064
W06000028	Blanaugor	9.8204	1.214	7.6064
W06000029	Blanaugor	9.8204	1.214	7.6064
W06000030	Blanaugor	9.8204	1.214	7.6064
W06000031	Blanaugor	9.8204	1.214	7.6064
W06000032	Blanaugor	9.8204	1.214	7.6064
W06000033	Blanaugor	9.8204	1.214	7.6064
W06000034	Blanaugor	9.8204	1.214	7.6064
W06000035	Blanaugor	9.8204	1.214	7.6064
W06000036	Blanaugor	9.8204	1.214	7.6064
W06000037	Blanaugor	9.8204	1.214	7.6064
W06000038	Blanaugor	9.8204	1.214	7.6064
W06000039	Blanaugor	9.8204	1.214	7.6064
W06000040	Blanaugor	9.8204	1.214	7.6064
W06000041	Blanaugor	9.8204	1.214	7.6064
W06000042	Blanaugor	9.8204	1.214	7.6064
W06000043	Blanaugor	9.8204	1.214	7.6064
W06000044	Blanaugor	9.8204	1.214	7.6064
W06000045	Blanaugor	9.8204	1.214	7.6064
W06000046	Blanaugor	9.8204	1.214	7.6064
W06000047	Blanaugor	9.8204	1.214	7.6064
W06000048	Blanaugor	9.8204	1.214	7.6064
W06000049	Blanaugor	9.8204	1.214	7.6064
W06000050	Blanaugor	9.8204	1.214	7.6064

Figure 17: PM2.5 data 2015 and 2024

3.1.4 Data manipulation and transformation

Additional data manipulation and transformation were undertaken within QGIS.

3.1.4.1 Mining data: Identifying coal mines

The historic mine location dataset contained several anomalies, including incomplete records and limited classification of mine type. As distinguishing coal mines was essential to this study, a best-estimate approach was adopted. The geology attribute was identified as the most appropriate field for differentiating mine types, and an SQL query was used to create a new classification field within the attribute table (Figure 18 and Figure 19).

```
CASE
  WHEN "geology" = 'L. Pal. Sed.' THEN 'Lower Paleozoic Sedimentary'
  WHEN "geology" = 'L. Pal. Sed/Ig' THEN 'Lower Paleozoic Mixed'
  WHEN "geology" = 'Carb Lst' THEN 'Carboniferous Limestone'
  ELSE 'Other/Unknown'
END
```

Figure 18: SQL script to identify mine types

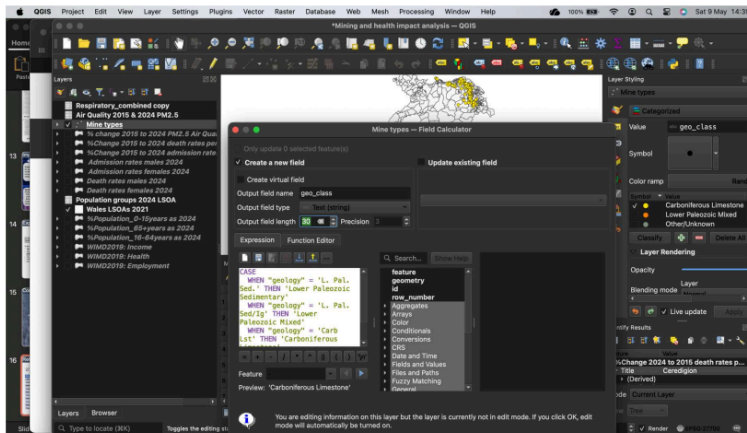


Figure 19: Adding calculated field to mine data

A sense check was performed by comparing the resulting coal mine distribution with known historic coalfield regions in Wales, notably the South Wales Valleys and North-east Wales (Figure 20). While this approach does not provide definitive

classification, it offers a reasonable and transparent approximation, which is acknowledged as a limitation in subsequent analysis. This manipulation was carried out in GIS due to constraints associated with directly editing the original shapefile.

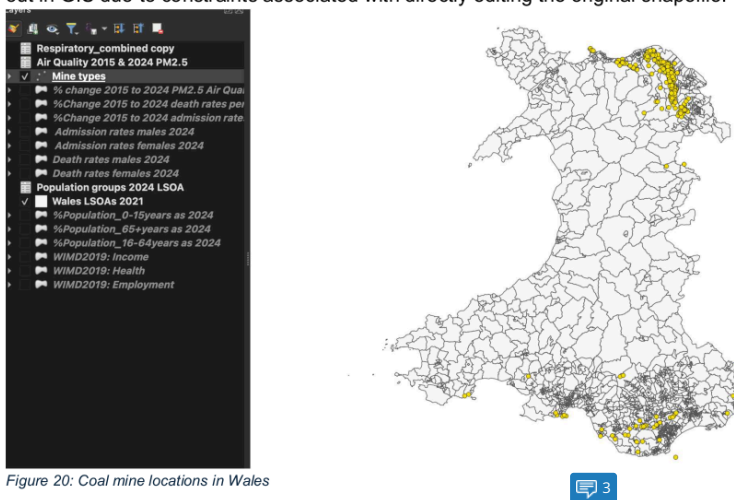


Figure 20: Coal mine locations in Wales

3.1.4.2 Data format issues

During data import, particularly from CSV files, some numeric fields were incorrectly interpreted as strings. Considerable time was spent attempting to resolve this within QGIS by creating new fields and converting values using SQL expressions such as `to_real`. Ultimately, a more reliable solution was to correct data types at source by adjusting formats in Excel and saving the files as .xlsx prior to import.

3.2 Thematic maps

Thematic maps are maps designed to display the spatial distribution of a specific variable or theme, such as social, economic, or health characteristics, allowing geographic patterns and relationships to be explored visually rather than through numerical tables [1].

The use of thematic mapping enables powerful visualisation of complex datasets, allowing patterns and trends to be identified quickly when maps are designed appropriately. In this project, thematic maps were central to exploring spatial

relationships between deprivation, historic mining activity, population characteristics, and respiratory health outcomes.

3.2.1 Joining and preparing data for thematic mapping

A critical step in preparing data for thematic mapping is ensuring that spatial and non-spatial datasets can be integrated accurately. All datasets must share a consistent coordinate reference system, as differing global reference systems can introduce spatial misalignment due to variations in projection and earth curvature. To ensure spatial consistency, all layers in this project were standardised to EPSG:27700 (OSGB 1936 / British National Grid) (Figure 21).

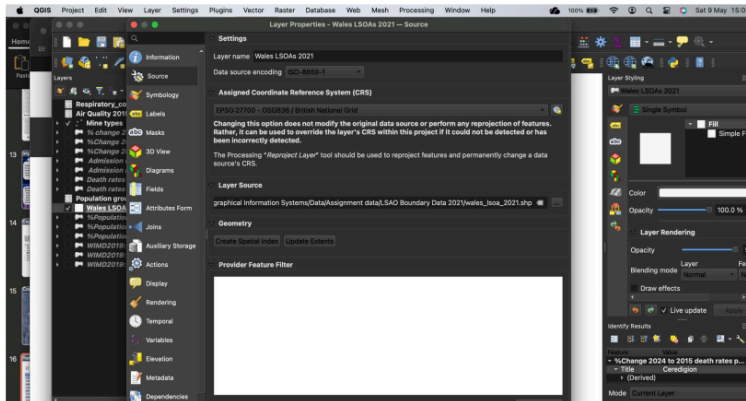


Figure 21: Project Coordinate Reference System (CRS)

3.2.1.1 LSOA to population group join

Population data was joined to the LSOA boundary layer using a common identifier, linking the Area field in the population dataset to the Isoa21nm field in the spatial boundary data. This attribute join enabled population characteristics to be mapped accurately within each LSOA polygon (Figure 22).

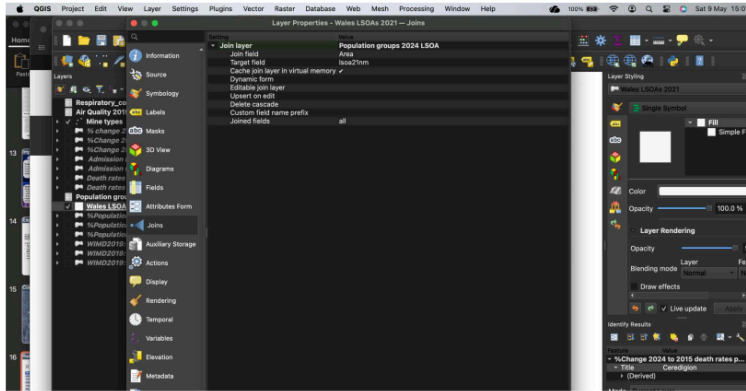


Figure 22: LSOA boundary and population join

3.2.1.2 Creating a dissolved layer to enable County joins

As only LSOA boundaries were available initially, a dissolved layer was created using the Dissolve tool in QGIS to aggregate LSOAs into County-level polygons (Figure 23). This derived County layer allowed joins to datasets available only at County scale, such as respiratory mortality and admission rates (Figure 24).

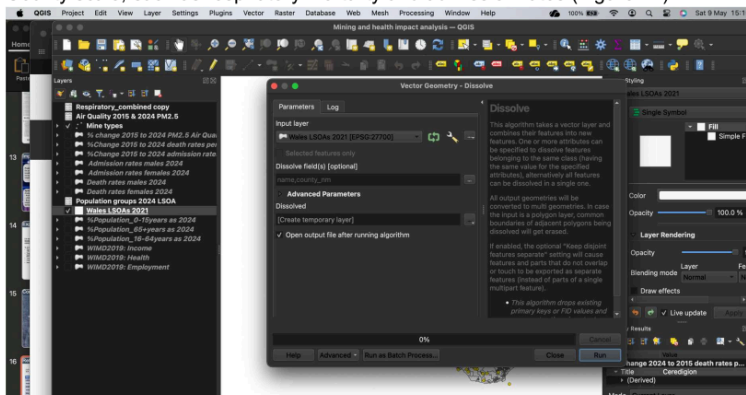


Figure 23: Creating a hidden layer containing Counties, derived from LSOA

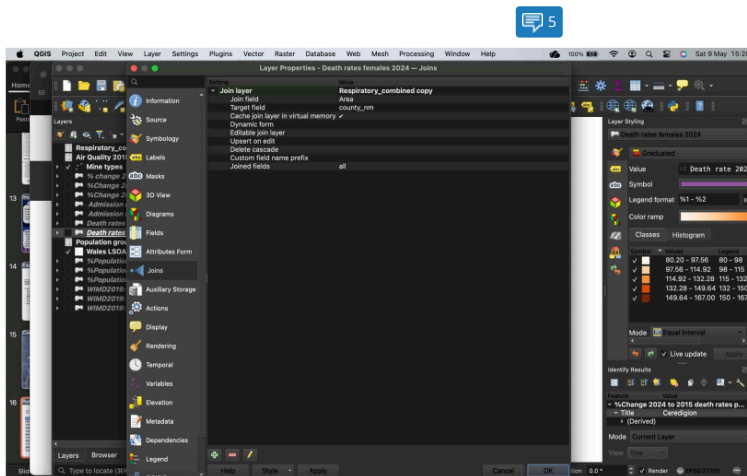


Figure 24: Joining respiratory data to aggregated County data

3.2.2 Creation of thematic maps

A series of thematic maps were produced to support analysis across different datasets:

- LSOA Boundary Map: A simple reference map displaying LSOA boundaries across Wales (Figure 25).

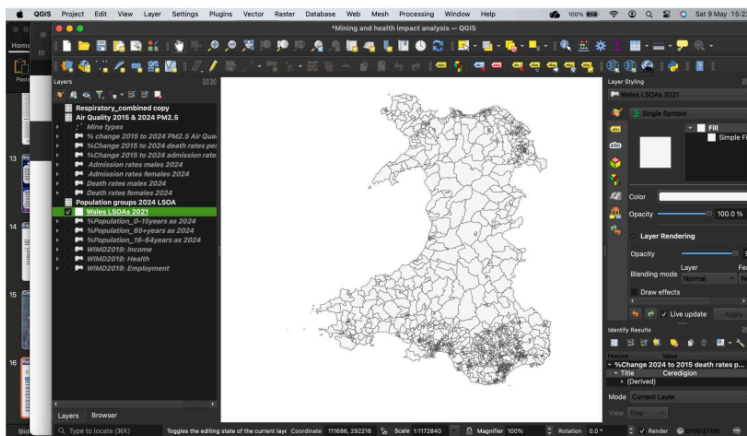


Figure 25: LSOA boundaries

- Deprivation Decile Maps: Graduated choropleth maps using distinct colour ramps for employment, income, and health deprivation, with decile classifications highlighting deprivation hotspots where 1 represents the highest deprivation (Figure 26-28).

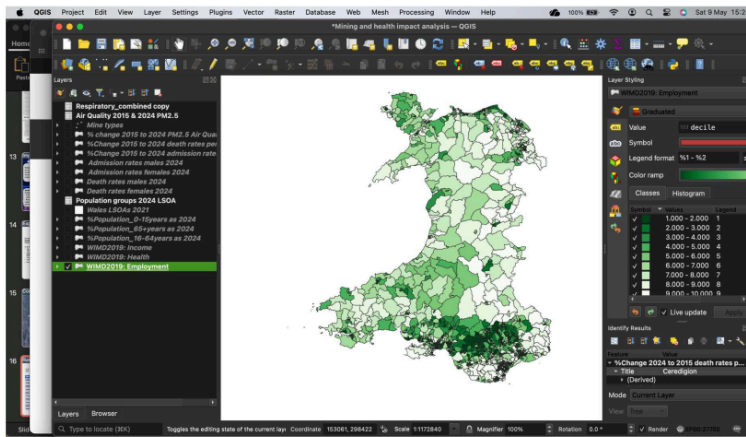


Figure 26: Employment deprivation

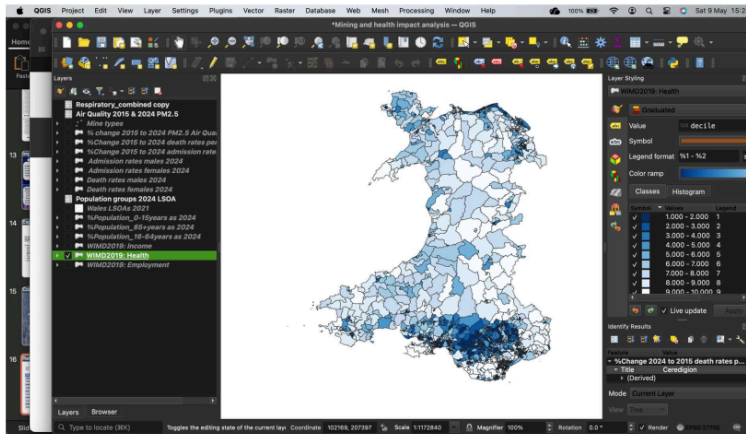


Figure 27: Health deprivation

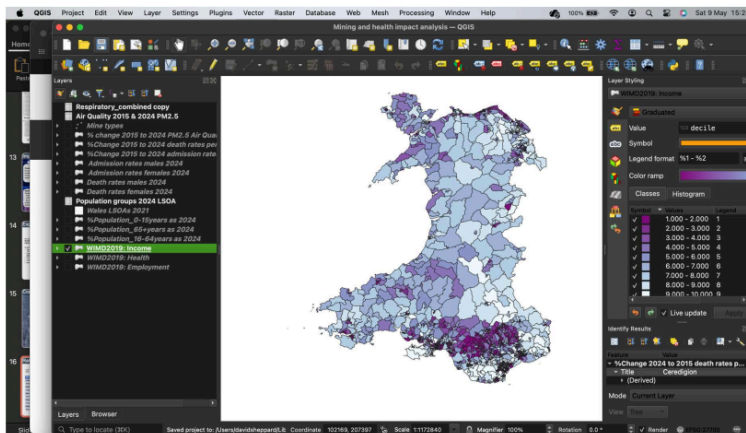


Figure 28: Income deprivation

- Respiratory Illness Maps: County-level graduated maps showing respiratory-related death and admission rates, with county labels added for clarity (Figure 29).

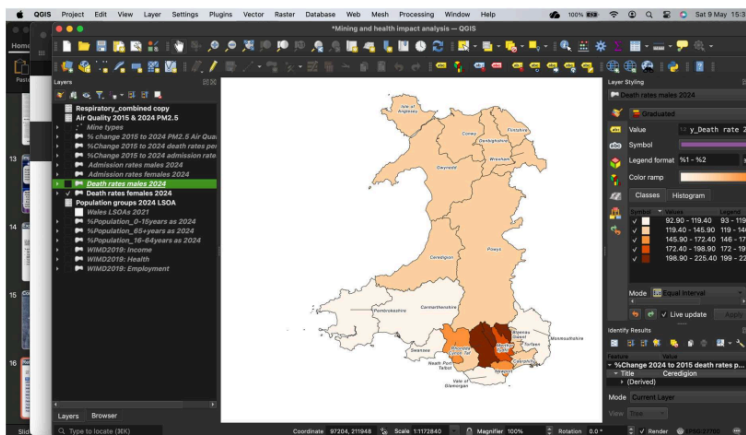


Figure 29: Respiratory illness death rates 2024

- Population Age Structure Maps: Graduated symbology was applied using consistent classification across age groups to enable comparison (Figure 30).

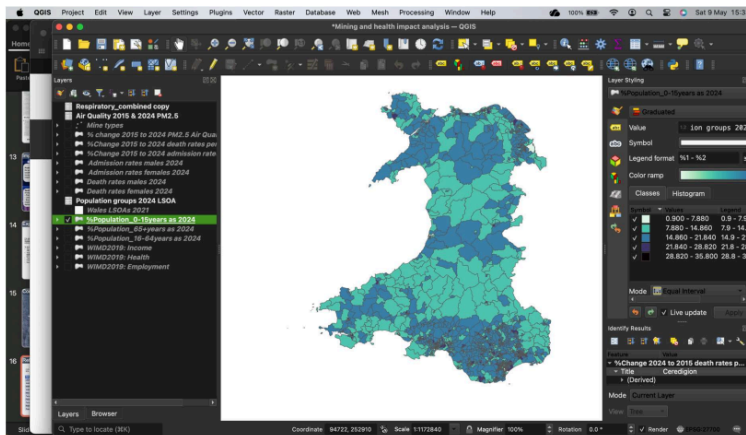


Figure 30: % of population aged 0-15 years

3.2.3 Creating a thematic map for air quality

- Air Quality Maps: PM2.5 concentration change was visualised at County level using a darker colour scheme to emphasise poorer air quality (Figure 31).

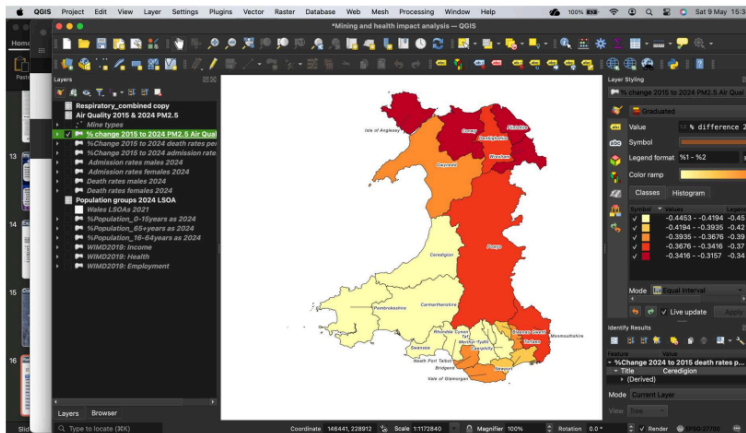


Figure 31: Air quality by County

- Mine Location Overlay: Coal mine locations were added as point layers over thematic maps, with non-coal mines filtered out to maintain focus on project objectives (Figure 32).

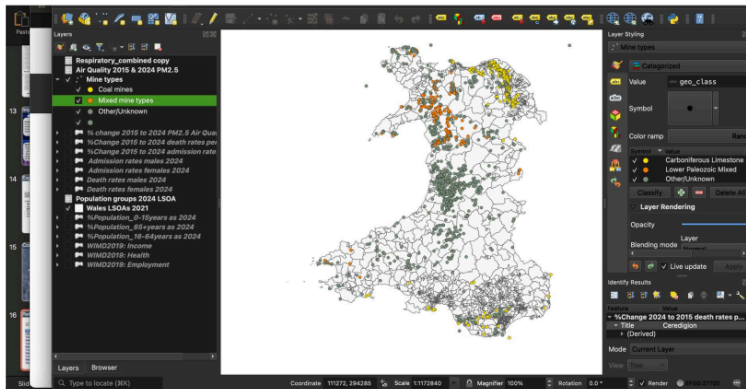


Figure 32: All mine locations

3.2.4 Final outputs

Using the Print Layout functionality in QGIS, a total of fourteen thematic maps were produced. Each layout included consistent titles, legends, symbology, and a fixed map scale to ensure visual comparability. Map themes were used to lock layer visibility and styling, ensuring consistency across outputs and producing clear, publication-ready visuals for analysis and interpretation.

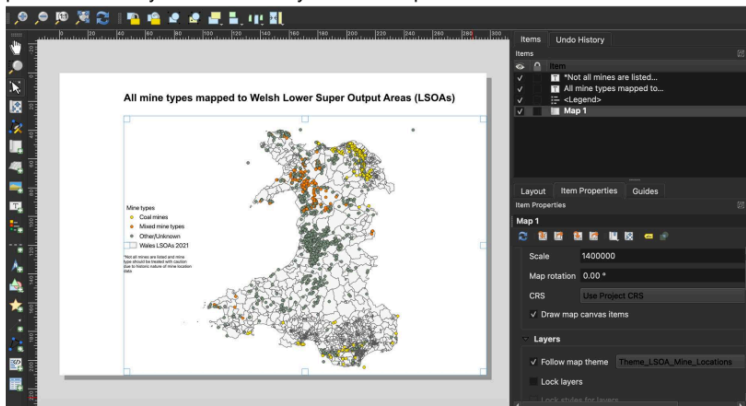


Figure 33: All mines aligned with LSOAs, with visible toolbar

Other final visualisations, as intended for end user, with consistent titles and legend.

Health deprivation deciles and coal mine locations

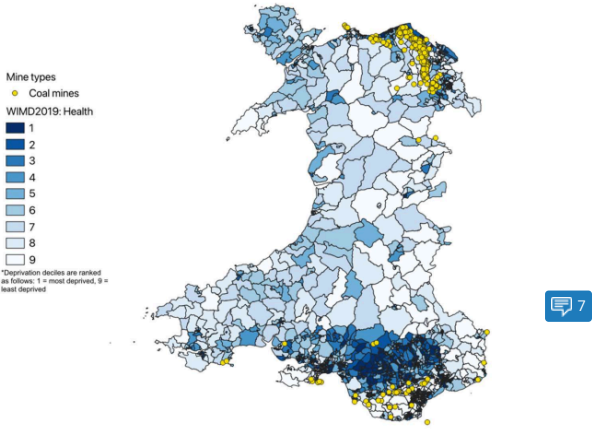


Figure 34: Health deprivation and coal mines

Employment deprivation deciles and coal mine locations

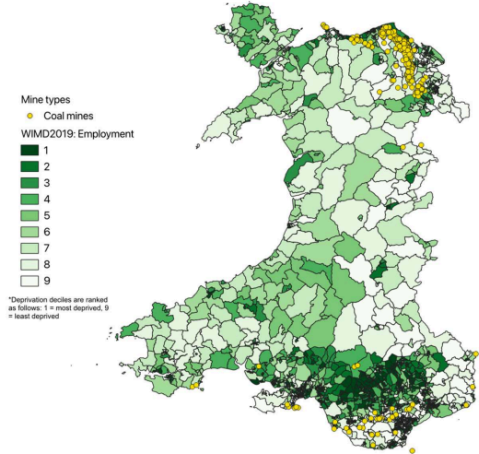


Figure 35: Employment deprivation and coal mines

Income deprivation deciles and coal mine locations

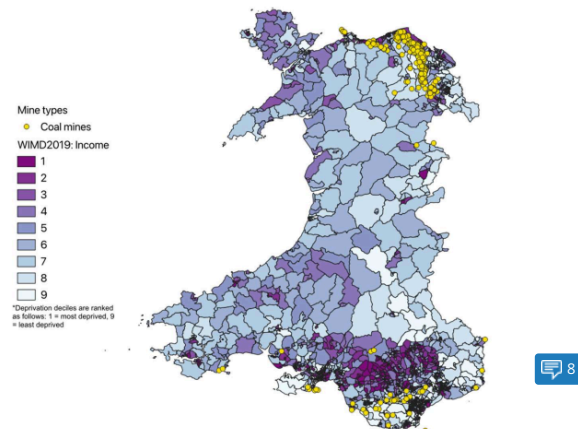


Figure 36: Income deprivation and coal mines

Percentage of population aged 0-15 years by Lower Super Output Area (LSOA) 2024

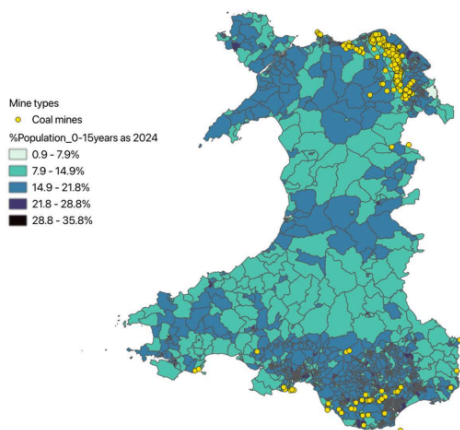


Figure 37: Population aged 0-15 years by LSOA 2024

Percentage of population aged 16-64 years by Lower Super Output Area (LSOA) 2024

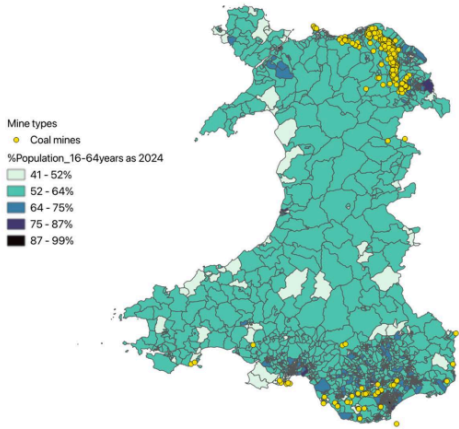
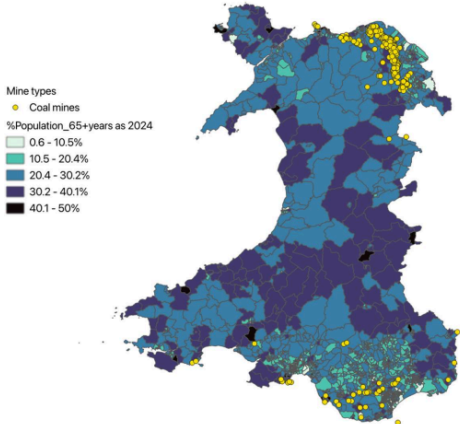


Figure 38: Population aged 16-64 years by LSOA 2024

Percentage of population aged 65+ years by Lower Super Output Area (LSOA) 2024



Percentage change in admission rates per 100,000 of the population from 2015 to 2024

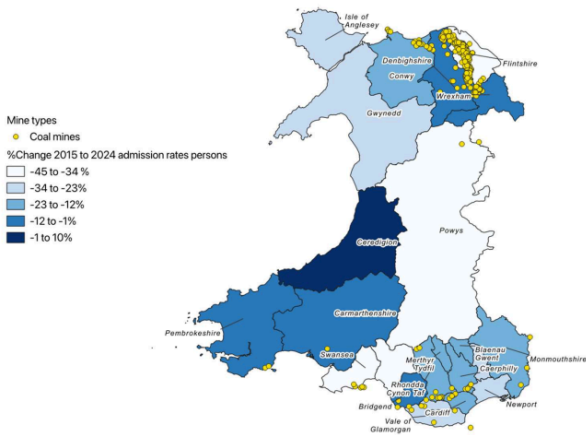


Figure 40: Change in respiratory admission rates 2015 to 2024 (persons)

Respiratory disease related hospital admissions rate per 100,000 of the population 2024 (males)

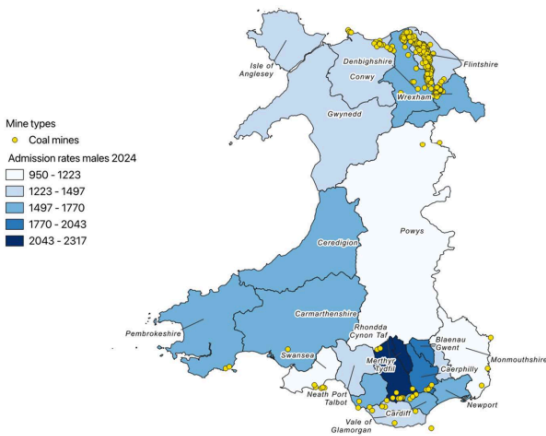


Figure 41: Respiratory admission rates 2024 (males)

Respiratory disease related hospital admissions rate per 100,000 of the population 2024 (females)

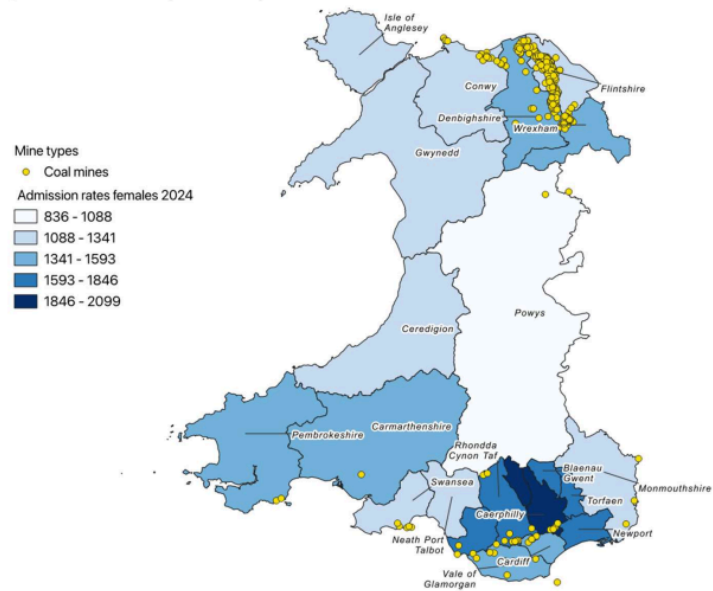


Figure 42: Respiratory admission rates 2024 (females)

Percentage change in death rates per 100,000 of the population from 2015 to 2024 (All genders)

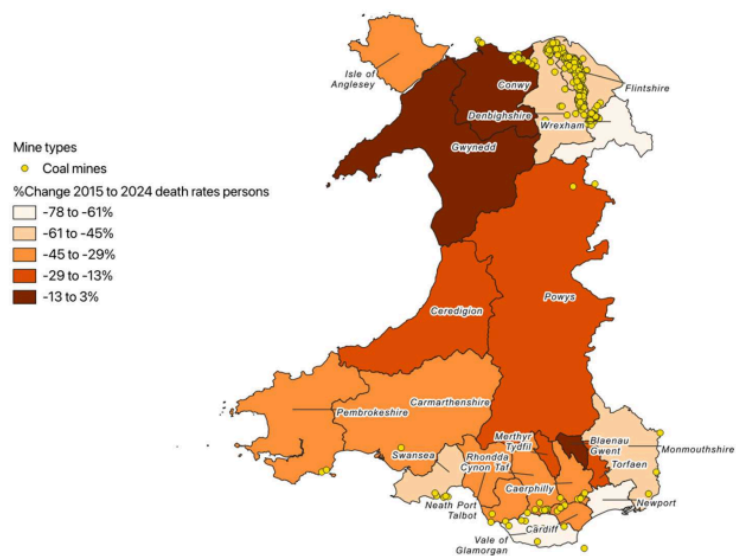


Figure 43: Change in respiratory death rates 2015 to 2024 (persons)

Respiratory disease related death rate per 100,000 of the population 2024 (males)

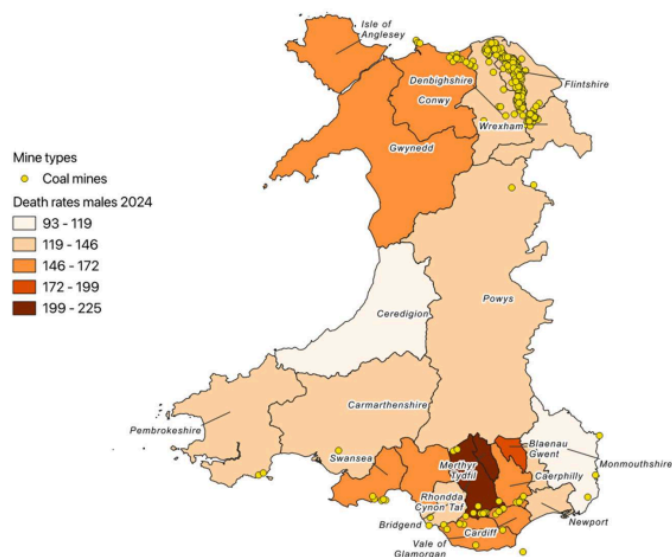


Figure 44: Respiratory death rates 2024 (males)

Respiratory disease related death rate per 100,000 of the population 2024 (females)

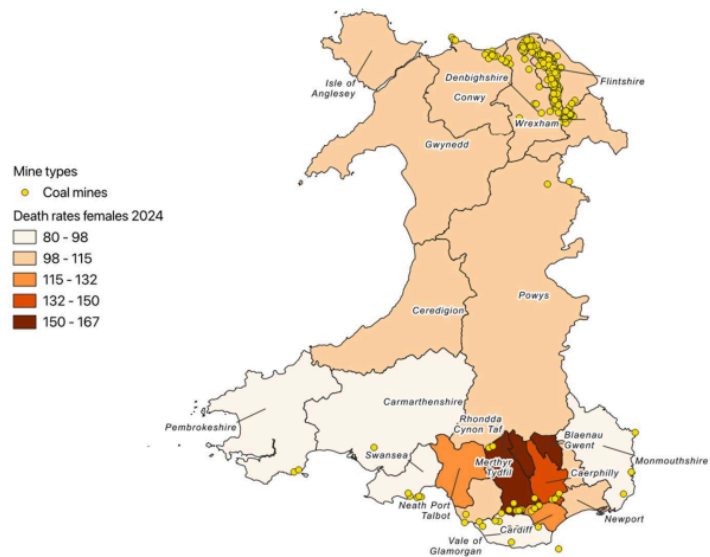


Figure 45: Respiratory death rates 2024 (females)

Percentage change in air quality, measured by PM2.5 particles from 2015 to 2024, by county

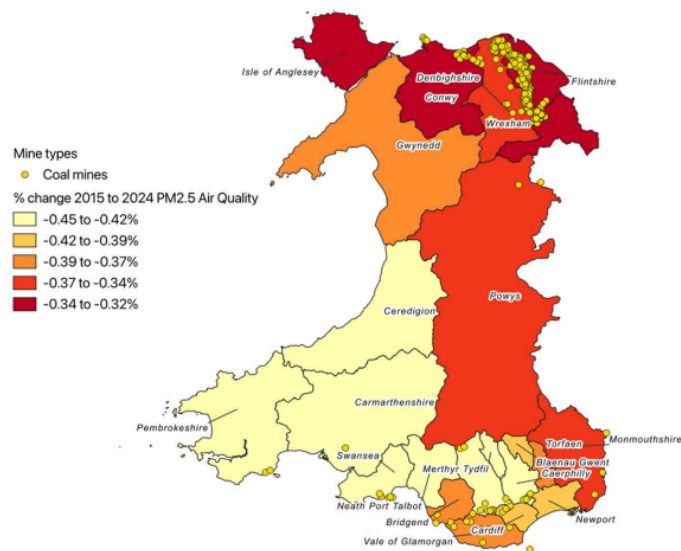


Figure 46: Air quality PM2.5 change 2015 to 2024

4. Outcomes

4.1 Findings from the thematic maps

10

Clear spatial patterns emerged from the analysis, some aligning with expectations and others challenging initial assumptions.

- Coal mine locations are concentrated in South Wales and north-east Wales, consistent with known historic mining regions.
- Areas surrounding South Wales coalfields display higher levels of income, employment, and health deprivation, while deprivation patterns in the north-east are less pronounced. This suggests a stronger legacy effect of mining in South Wales.
- LSOAs near historic coal mines are more heavily populated by working-age residents and children, with proportionally fewer older adults. While this should not be assumed to reflect differences in health outcomes or life expectancy, it highlights an area requiring further demographic investigation.
- Respiratory-related hospital admissions and mortality rates are generally higher in areas associated with historic coal mining. Although reductions between 2015 and 2024 are observed nationally, higher starting rates near coalfields suggest that the health impacts of mining continue to influence

outcomes and health investment opportunities should be explored to improve population health.

- Air quality improvements were greater in South Wales than in the Northeast. When considered alongside deprivation and respiratory health patterns, this suggests that non-air-quality factors may contribute more strongly to ongoing health inequalities in former mining areas.

4.2 Spatial query insights

The application of spatial queries enhanced both analytical depth and interpretive clarity:

- Spatial integration: Attribute and spatial data was joined to LSOA and County boundaries, enabling consistent thematic mapping.
- Geographic filtering: Non-Welsh data was excluded to maintain analytical focus.
- Thematic overlay: The combination of deprivation, demographic, health, air quality, and mining layers revealed spatial relationships and potential correlations.
- Clustering and scale analysis: Hotspots were more clearly identified when analysing patterns at both LSOA and County scales.
- Interpretation and communication: GIS visualisation proved effective in communicating complex spatial relationships in an accessible manner.

5. Analysis



5.1 Principles of thematic mapping and visualisation

Effective thematic mapping prioritises clarity, relevance, and consistency. The use of appropriate colour schemes, standardised symbology, and controlled layering enabled spatial differences to be highlighted without unnecessary visual clutter. The integration of multiple thematic layers allowed relationships between deprivation, health, and environmental factors to be interpreted more intuitively than would be possible using tabular data alone, supporting evidence-based spatial reasoning.

5.2 Challenges and limitations

- Data and import issues: Percentage-formatted fields were sometimes misinterpreted by GIS software, resulting in null values and requiring reformatting.
- Technical limitations: Problems with geocoding, layer alignment, and inconsistent data formats necessitated repeated troubleshooting.

- Integration constraints: Combining multi-source datasets required extensive preprocessing and spatial filtering.
- Time constraints: Debugging and rebuilding limited opportunities for more advanced analysis.
- Data resolution: Some datasets lacked completeness or fine-scale spatial detail.
- Output limitations: Static maps restrict interactivity. GIS software is difficult to master.

6. Conclusions



This project demonstrates the value of Geographic Information Systems (GIS) in exploring complex relationships between environmental, socioeconomic deprivation, and health outcomes. By integrating spatial, demographic, health, and historical mining datasets, the analysis revealed clear geographic patterns across Wales that would be difficult to identify using non-spatial methods alone. In particular, the concentration of historic coal mining activity in South Wales and the Northeast was shown to coincide with areas of higher deprivation and poorer respiratory health outcomes, suggesting that the impacts of the coal mining industry continue to influence communities long after its decline.

Thematic mapping proved effective in communicating these patterns, allowing differences and hotspots to be identified at both LSOA and County scales. While improvements in respiratory health outcomes and air quality were observed over time, the persistence of higher rates in former mining areas highlights ongoing inequalities. At the same time, variations in population age structure and regional air quality trends indicate that these outcomes are likely to be shaped by a range of interacting social, environmental, and demographic factors, rather than mining legacy alone.

Several limitations shaped the scope of the analysis, including restricted access to granular health data, inconsistencies in data formats, and reliance on static visual outputs. These constraints provide clear opportunities for further research. Future analysis could incorporate population-weighted health measures at more granular spatial scales, proximity-based analysis linking mine locations to health outcomes with buffer zones, and extended time-series data to assess long-term trends. The inclusion of additional environmental indicators and the development of interactive GIS outputs would further enhance analytical depth and accessibility.

Overall, this study illustrates how GIS can support evidence-based understanding of spatial health inequalities and provides a foundation for further investigation into the long-term impacts of industrial legacy on public health.

7. Bibliography

- [1] D. I. Heywood, S. Cornelius and S. Carver, An introduction to geographical information systems, 4th ed., Prentice Hall, 2011.  13
- [2] S. A. Selmy, D. E. Kucher, J. Yang and F. J. Garcia-Navarro, "Geospatial Data: Acquisition, Applications, and Challenges," intechopen, 14 October 2024. [Online]. Available: <https://www.intechopen.com/chapters/1192070>. [Accessed 1 April 2026].
- [3] Data Maps Wales, "Lower layer Super Output Areas (LSOAs) - Wales – 2021," Office of National Statistics, 25 November 2025. [Online]. Available: https://datamap.gov.wales/layers/geonode: Isoa_2021_w_hwm. [Accessed 28 March 2026].
- [4] H. M. Wales, "Data Explorer," Digital Health Care Wales (DHCW), 5 March 2026. [Online]. Available: <https://www.healthmapswales.wales.nhs.uk/data-catalog-explorer/?tag=G211>. [Accessed 5 April 2026].
- [5] N. R. Wales, "Minescan," Natural Resources Wales, 22 June 2022. [Online]. Available: https://metadata.naturalresources.wales/geonetwork/srv/api/records/NRW_DS99411#gn-tab-metadataCustom. [Accessed 5 April 2026].
- [6] D. M. Wales, "Data Catalogue," Data Maps Wales, 27 November 2017. [Online]. Available: <https://datamap.gov.wales/search/?limit=20&offset=0>. [Accessed 2026 15 March].
- [7] H. M. Wales, "Data Explorer," DHCW, 5 March 2026. [Online]. Available: https://www.healthmapswales.wales.nhs.uk/data-catalog-explorer/?tag=TDiseasesoftheRespiratorySystem_RespiratoryDiseaseMortality&q=respiratory. [Accessed 11 April 2026].
- [8] "Modelled background pollution data," Department for Environment, Food and Rural Affairs (DEFRA), 2024. [Online]. Available: <https://uk-air.defra.gov.uk/data/pcm-data>. [Accessed 8 May 2026].

FINAL GRADE

64/100

GENERAL COMMENTS

Dear David,

report is technically impressive and the practical GIS work is genuinely strong. However, maps are produced but barely analysed. The literature engagement is also weak for a 25-mark criterion — there is less cartographic or GIS theory discussed.

more detailed comments can be found inside the document

Best Wishes

PAGE 1

PAGE 2



Comment 1 | Transferable Skills(10)

- detailed and well-constructed table of contents with numbered sections and subsections.
-
bibliography is consistently formatted

PAGE 3

PAGE 4



Comment 2 | Knowledge&Understanding(20)

data quality discussion references meaningfully in some of the sections however

exploring how a map is used as a visual information system is very limited

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16



Comment 3



Literature (25)

Discuss what your sources argue rather than just citing them



Comment 4



Literature (25)

sections that should cite literature contain very limited reference

PAGE 17

PAGE 18

PAGE 19



Comment 5



Knowledge&Understanding(20)

decision to dissolve LSOA boundaries into County-level polygons to enable joins with County-scale health data shows genuine understanding of spatial data integration

PAGE 20

PAGE 21

PAGE 22

PAGE 23



Comment 6 | Practical Application (20)

thematic maps

demonstrates sustained practical effort across the full scope of the research question

PAGE 24



Comment 7 | Practical Application (20)

the difference in distrobution of coal mines to be elaborated more clearly

PAGE 25



Comment 8 | Practical Application (20)

Screenshots of data sources, metadata, cleaning steps, and SQL expressions provide genuine transparency about the full workflow

PAGE 26



Comment 9 | Cognitive and Intellectual(25)

Considering age structure as a potential confounding variable in health outcome interpretation shows practical thinking

PAGE 27



Comment 10 | **Transferable Skills(10)**

Outcomes and Analysis sections are structurally weak compared to the methodology



Comment 11 | **Knowledge&Understanding(20)**

Outcomes and analysis sections should have been expanded. each finding should state what map shows and what it means in the context of research question



Comment 12 | **Cognitive and Intellectual(25)**

Future recommendations should be listed with reference to existing studies that have used similar approaches.



Comment 13 | **Literature (25)**

GIS textbook used as a foundational reference, which is appropriate

GRADING FORM: GIS

DAVID SHEPPARD

64

LITERATURE (25)



15

KNOWLEDGE&UNDERSTANDING(20)



17

COGNITIVE AND INTELLECTUAL(25)



19

PRACTICAL APPLICATION (20)



13

TRANSFERABLE SKILLS(10)