

Adam Smith 2317631

ACCA5037

by Adam Smith

Submission date: 10-May-2026 02:38PM (UTC+0100)

Submission ID: 282932009

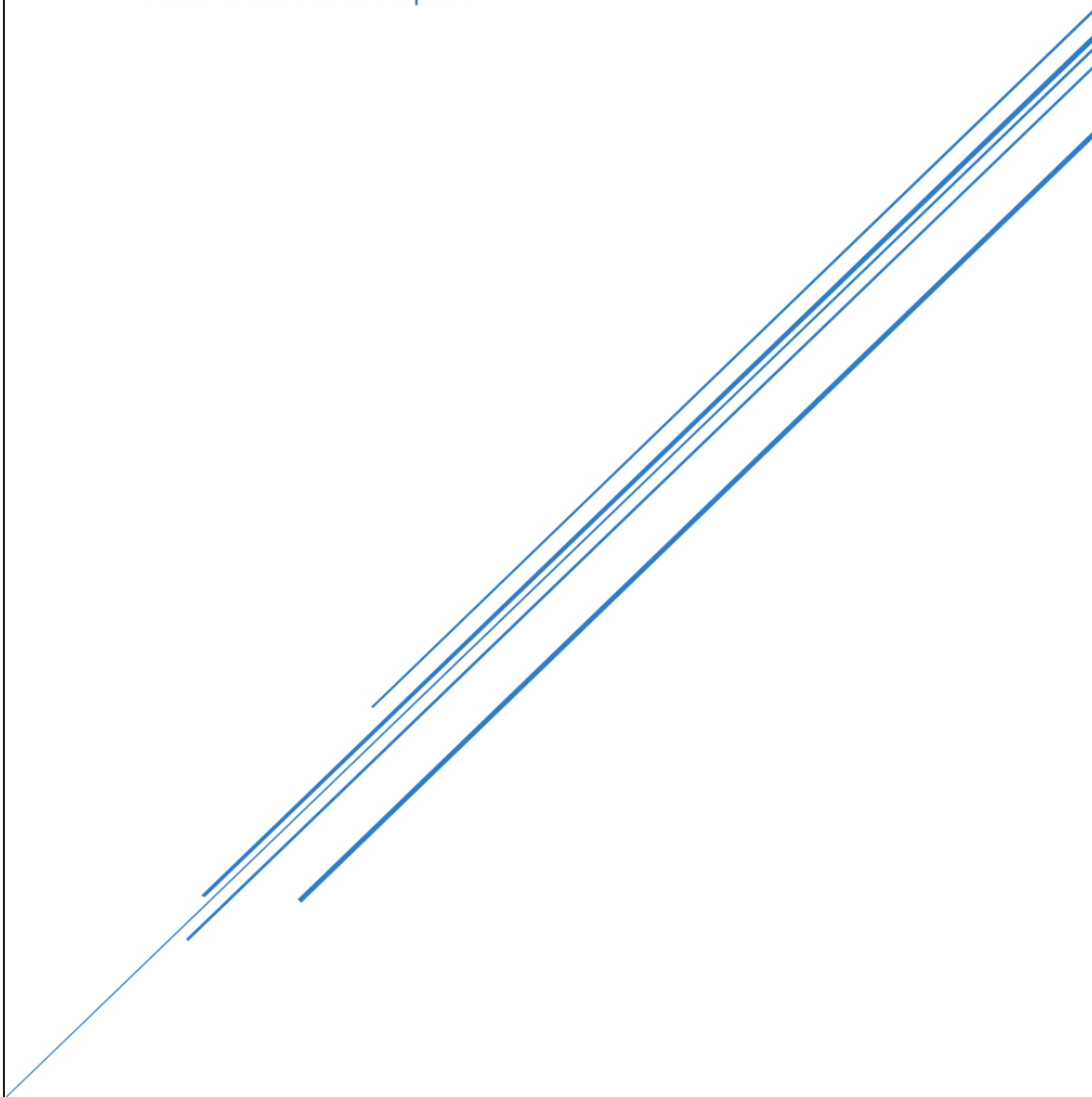
File name: 50077_Adam_Smith_Adam_Smith_2317631_ACCA5037_734550_97760746.docx (11.31M)

Word count: 4101

Character count: 22516

Geographical Information Systems (ACCB5027)

Task 2: Research Report



Adam Smith (2317631)
Digital Degree Apprenticeship

Contents

Abstract	2
Introduction.....	3
Methodology.....	3
Data Processing	3
Thematic Maps.....	9
Outcome	18
Findings from thematic maps.....	18
Analysis.....	19
Principles of thematic mapping and data visualisation.....	19
Challenges and limitations encountered	19
Conclusion	20
Suggestions for further analysis	20
References	20



Abstract

This report provides an analysis of socio-economic healthcare conditions as a result of historical employment in Wales produced using Geographical Information Systems (GIS) to provide thematic maps. The assignment comprises several datasets, from health care data, historic mining sites and working age averages. Thematic maps were created to allow for the visualisation of hospital admission rates, respiratory death rates and their proximity to mine locations, along with areas of health deprivation. The expected correlation of increased admission and death rates with mine sites was not realised, with analysis of the data posing more questions and avenues of further exploration than it definitively answers. This analysis utilised both thematic and heat maps to allow for the identification of trends such as clustering and how these fit with health data across those locations.

Introduction

Geographical Information Systems (GIS) are computer-based tools used to capture, store, analyse, and visualise spatial data linked to specific locations [1]. In this assignment, GIS is applied to examine data from across Wales to explore whether historic mining activity is associated with higher levels of respiratory disease in the local communities. Spatial data such as LSOA boundaries which will indicate health data will be cross referenced with the Welsh Index Multiple Deprivation and spot data such as mine locations to assess any trends. By enabling the spatial representation of multiple datasets, GIS supports the identification of geographic patterns and potential correlations. This approach provides a robust basis for assessing spatial inequalities in health and for informing recommendations on healthcare provision in areas with a history of mining which may still influence population health.

The assessment utilises the following datasets:

LSOA Boundaries 2021 (Data map Wales) [2]

Minescan 2025 (Data map Wales) [3]

Working Age percentage (Health map Wales) [4]

WIMD (Data map Wales) [5]

Respiratory diseases (Health map Wales) [6]

LSOA (2021) to Local Authority District (April 2023) Best Fit Lookup [7]



Methodology

Data Processing

As per the UK GDPR, data processing is governed by seven core principles: lawfulness, fairness and parency; purpose limitation; data minimisation; accuracy; storage limitation; integrity and confidentiality; and accountability [8].

Lawfulness, fairness and transparency – Processing must have a lawful basis (consent), be fair (not misleading) and be transparent (people must understand what the data is used for).

Purpose Limitation – Data must be collected for a specific and legitimate purpose. It cannot be reused for other purposes.

Data minimisation – Data collected must be adequate and should be restricted to what is relevant and necessary

Accuracy – Data must be accurate and kept up to date, with inaccurate data correct or deleted as soon as possible.

Storage Limitation – Data should not be kept longer than needed. Long term use is acceptable under certain circumstances.

Integrity and confidentiality – Data must be protected against unauthorised access, accidental loss and damage or destruction

Accountability – organisations must take responsibility for compliance and be able to demonstrate it.

Adhering to these principles is essential not only for meeting legal and ethical obligations, but also for ensuring best practice and maintaining professionalism throughout the project. It is essential that all data is sourced from reliable and credible sources. To this purpose, datasets for this project were obtained from governmental bodies, ensuring that the information used is publicly available, accredited, and supported by high-quality metadata. These sources provide clear documentation on how each dataset was collected, its intended purpose, and any associated audit trails, including publication dates, updates, amendments, and the availability of more recent versions.

Data Acquisition

Data acquisition also revolves around some core principles – Define purpose; Source from authoritative, credible providers; Ensure data is relevant and appropriate; Check accuracy and validity; Assess completeness; Verify metadata; Ensure ethical and legal compliance; Check data timeliness and currency; Ensure data compatibility [9].

The purpose for our project is clearly defined – to provide an analysis of respiratory health conditions in historic mining areas in Wales. Data has been collected from Welsh government – a credible source. The data collected has been restricted to only that which is required for this assessment. Accuracy, Validity and Completeness will be assessed in the data cleansing stage. Metadata is clearly labelled and detailed on Welsh government sources, along with their collection methods and compliance aspects with GDPR etc. data was checked for compatibility, ensuring required fields are present, along with matching or compatible boundaries

Datasets

Datasets were sourced from Data Map Wales and Health Map Wales, detailed below;

Lower Super Output Area Boundaries – 2021

This data was required to identify the geographical area that the Office for Nation Statistics (ONS) split up the country of Wales into. The ONS split up the country to allow for consistent comparable reporting of small area data across the UK. As our report is restricted to the country of Wales and following the principles of Data Processing and Data Acquisition, we have only obtained data for Wales, not the UK. This data not only splits up the geographical areas but also allows for joining between data layers as they all have a unique LSOA code for each specific area. 2 specific columns required are LSOA code and LSOA name, allowing for the joining of data sets.

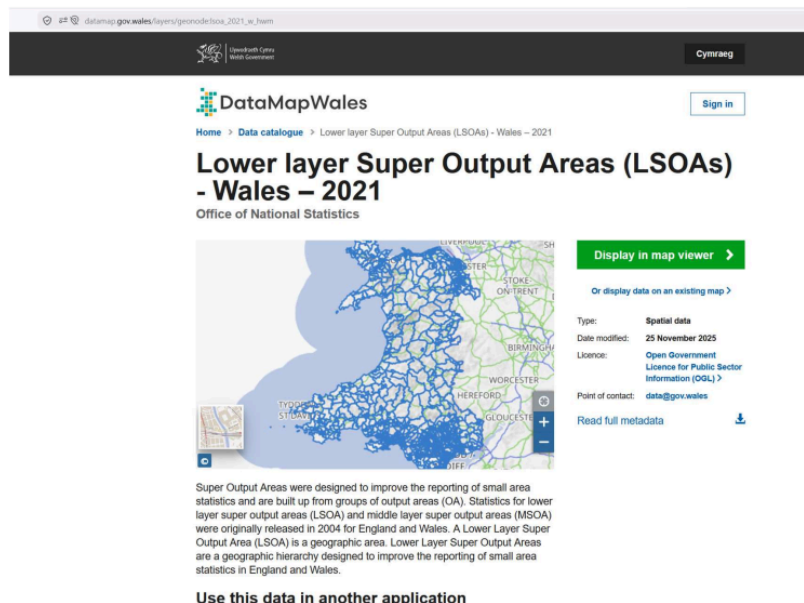


Figure 1 – Data Map Wales source for 2021 LSOA boundaries

The screenshot shows a spreadsheet with the following columns: FID, Isoa21cd, Isoa21nm, Isoa21nmw, Isoa21nmwenglish, Isoa21nmwelsh, id, and geom. The rows contain LSOA data for various areas in Wales, including Amwlch, Bodorgan, Brynllw, Cadnant, Cefni, Cwm Cadnant, Cwm Cadnant, Cyngar, Gwynyll, Gwynyll, Gwynyll, Tref Cybi, and Kingsland. The 'Isoa21cd' and 'Isoa21nm' columns are highlighted in green.

FID	Isoa21cd	Isoa21nm	Isoa21nmw	Isoa21nmwenglish	Isoa21nmwelsh	id	geom
1	W01000003	Isle of Anglesey 001A	Ynys Môn 001A	Amwlch Port	Amwlch Port	1	DATA TRUNCATED POLYGON ((24...
2	W01000004	Isle of Anglesey 001B	Ynys Môn 001B	Amwlch Rural	Amwlch Wledig	2	POLYGON ((243706.8388000002 393...
3	W01000004	Isle of Anglesey 001B	Ynys Môn 001B	Amwlch Rural	Amwlch Wledig	3	DATA TRUNCATED POLYGON ((24...
4	W01000005	Isle of Anglesey 005A	Ynys Môn 005A	Beaumaris	Beaumaris	4	DATA TRUNCATED POLYGON ((23...
5	W01000006	Isle of Anglesey 006A	Ynys Môn 006A	Bodffordd	Bodffordd	5	DATA TRUNCATED POLYGON ((24...
6	W01000007	Isle of Anglesey 009A	Ynys Môn 009A	Bodorgan	Bodorgan	6	POLYGON ((245336.09399999958 37...
7	W01000007	Isle of Anglesey 009A	Ynys Môn 009A	Bodorgan	Bodorgan	7	POLYGON ((243182.61299999999 369...
8	W01000007	Isle of Anglesey 009A	Ynys Môn 009A	Bodorgan	Bodorgan	8	POLYGON ((235929.09690000004 369...
9	W01000007	Isle of Anglesey 009A	Ynys Môn 009A	Bodorgan	Bodorgan	9	DATA TRUNCATED POLYGON ((24...
10	W01000008	Isle of Anglesey 008A	Ynys Môn 008A	Braint	Braint	10	DATA TRUNCATED POLYGON ((23...
11	W01000009	Isle of Anglesey 007C	Ynys Môn 007C	Bryngwran	Bryngwran	11	DATA TRUNCATED POLYGON ((23...
12	W01000010	Isle of Anglesey 002A	Ynys Môn 002A	Brynlleg	Brynlleg	12	DATA TRUNCATED POLYGON ((23...
13	W01000011	Isle of Anglesey 008B	Ynys Môn 008B	Cadnant (Isle of Anglesey)	Cadnant (Ynys Môn)	13	POLYGON ((256128.90290000001 372...
14	W01000011	Isle of Anglesey 008B	Ynys Môn 008B	Cadnant (Isle of Anglesey)	Cadnant (Ynys Môn)	14	POLYGON ((256069.29630000007 37...
15	W01000011	Isle of Anglesey 008B	Ynys Môn 008B	Cadnant (Isle of Anglesey)	Cadnant (Ynys Môn)	15	POLYGON ((254871.39910000004 37...
16	W01000011	Isle of Anglesey 008B	Ynys Môn 008B	Cadnant (Isle of Anglesey)	Cadnant (Ynys Môn)	16	DATA TRUNCATED POLYGON ((23...
17	W01000012	Isle of Anglesey 008B	Ynys Môn 008B	Cefni	Cefni	17	POLYGON ((246430.23599999957 37...
18	W01000013	Isle of Anglesey 005B	Ynys Môn 005B	Cwm Cadnant	Cwm Cadnant	18	POLYGON ((256537.19830000029 37...
19	W01000013	Isle of Anglesey 005B	Ynys Môn 005B	Cwm Cadnant	Cwm Cadnant	19	POLYGON ((256471.79849999957 37...
20	W01000013	Isle of Anglesey 005B	Ynys Môn 005B	Cwm Cadnant	Cwm Cadnant	20	DATA TRUNCATED POLYGON ((23...
21	W01000014	Isle of Anglesey 006C	Ynys Môn 006C	Cyngar	Cyngar	21	DATA TRUNCATED POLYGON ((24...
22	W01000015	Isle of Anglesey 008C	Ynys Môn 008C	Gwynyll	Gwynyll	22	POLYGON ((254465.50299999956 37...
23	W01000015	Isle of Anglesey 008C	Ynys Môn 008C	Gwynyll	Gwynyll	23	POLYGON ((254421.00399999972 37...
24	W01000015	Isle of Anglesey 008C	Ynys Môn 008C	Gwynyll	Gwynyll	24	POLYGON ((254461.24199999962 37...
25	W01000016	Isle of Anglesey 003A	Ynys Môn 003A	Tref Cybi	Tref Cybi	25	POLYGON ((224982.99990000017 38...
26	W01000017	Isle of Anglesey 003B	Ynys Môn 003B	Kingsland	Kingsland	26	POLYGON ((224936.21700000018 38...

Figure 2 – LSOA boundaries with LSOA code and LSOA name columns highlighted

Minescan

This data was obtained from the Data map Wales, detailing the location of all historic mines in Wales represented as spot data. This data does not contain any LSOA codes, only exact coordinates for the location of the mine. The dataset also includes the type of mine, detailing if it is metalliferous, and if so, which metals were mined. This opens the possibilities of further investigation for specific respiratory diseases only affected by certain materials.

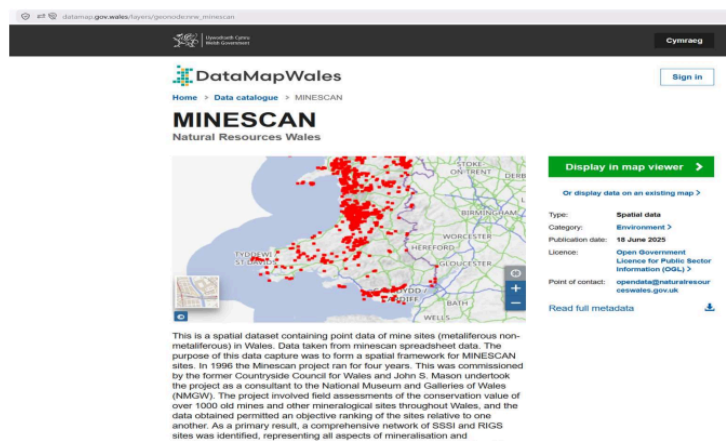


Figure 3 – Minescan obtained from Data Map Wales

ID	Name	Type	Locality	Region	Pb	Cu	Zn	Ag	...
1	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
2	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
3	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
4	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
5	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
6	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
7	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
8	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
9	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
10	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
11	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
12	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
13	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
14	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
15	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
16	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
17	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
18	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
19	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...
20	Albion Works	Metalliferous	Albion	Central Wales	100	50	20	10	...

Figure 4 – Minescan data, detailing coordinates and metalliferous elements mined

Working Age percentage

This data was obtained from Health map Wales. Data is provided as a percentage for 3 split age groups – 0-15 (pre working age), 16-64 (working age) and 65+ (retirees) - provided in 3 separate datasets. The dataset provides percentages from 2015 through to 2024, however, it is missing LSOA codes, instead only featuring the LSOA names.

Indicator	Date	Area	Value
% of people: 65+ years (retirement age)	2015	Isle of Anglesey 001A	21.5
% of people: 65+ years (retirement age)	2016	Isle of Anglesey 001A	21.6
% of people: 65+ years (retirement age)	2017	Isle of Anglesey 001A	22.1
% of people: 65+ years (retirement age)	2018	Isle of Anglesey 001A	22.1
% of people: 65+ years (retirement age)	2019	Isle of Anglesey 001A	22.7
% of people: 65+ years (retirement age)	2020	Isle of Anglesey 001A	22.7
% of people: 65+ years (retirement age)	2021	Isle of Anglesey 001A	22.7
% of people: 65+ years (retirement age)	2022	Isle of Anglesey 001A	23.3
% of people: 65+ years (retirement age)	2023	Isle of Anglesey 001A	23.2
% of people: 65+ years (retirement age)	2024	Isle of Anglesey 001A	24.1
% of people: 65+ years (retirement age)	2015	Isle of Anglesey 001B	31
% of people: 65+ years (retirement age)	2016	Isle of Anglesey 001B	31.6

Figure 5 – percentage of people 65+, 16-64 and 0-16

Welsh Index of Multiple Deprivation (WIMD) Health 2019

The Welsh Index of Multiple Deprivation (WIMD) is a measure of deprivation put in place by the Welsh government. There are several datasets available each with a differing specific focus, along with an aggregated overall score of deprivation for specific areas, split down by individual LSOA areas.

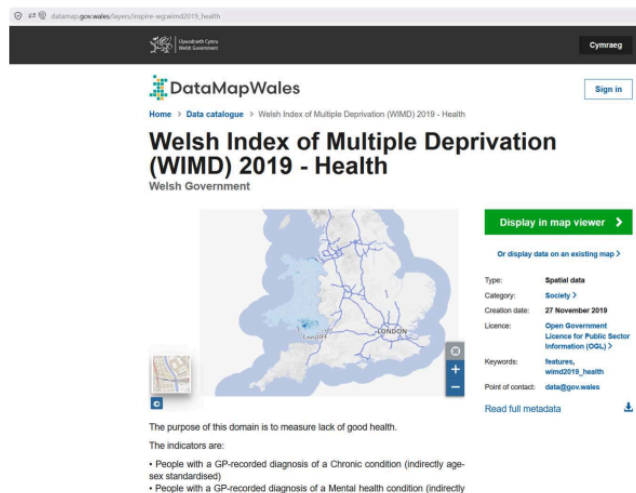


Figure 6 – WIMD 2019 - Health

Respiratory diseases (Health map Wales)

This dataset was obtained from Health map Wales, comprising of several separate files measuring – All respiratory death rates age standardised per 100,000 population Females, Males and All; All respiratory hospital admission rates age standardised per 100,000 population, Females, Males and All.

The screenshot shows a Microsoft Excel spreadsheet titled 'All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Age'. The spreadsheet has columns A, B, C, and D. Column A contains the indicator name, Column B contains the date, Column C contains the area, and Column D contains the value. The data is organized by region, with 'Isle of Anglesey' and 'Gwynedd' listed separately. The values range from 119.2 to 170.3.

Indicator	Date	Area	Value
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2015	Isle of Anglesey	170.3
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2016	Isle of Anglesey	133.9
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2017	Isle of Anglesey	132.6
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2018	Isle of Anglesey	152.7
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2019	Isle of Anglesey	142.3
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2020	Isle of Anglesey	123.5
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2021	Isle of Anglesey	109.7
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2022	Isle of Anglesey	129.4
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2023	Isle of Anglesey	146.7
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2024	Isle of Anglesey	128.9
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2015	Gwynedd	134.4
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2016	Gwynedd	136.1
All Respiratory Diseases, Death Rates (Age-standardised) per 100k pop, Persons, All Ages	2017	Gwynedd	119.2

Figure 7 – All Respiratory Diseases, Death rates

Data Cleansing

All data underwent a systematic process to detect any errors, either correcting them or removing them and identify any null or missing entries. The cleaning process ensures data is accurate, consistent and ready for analysis. As Data Map Wales and Health Map Wales are governmental bodies with robust quality assurance procedures in place no errors were encountered in the raw data. Upon searching the metadata for Minescan there is a clear warning for incorrect information in some of the locations. Once loading the data into QGIS it is quickly apparent which of the mine locations were incorrect. Mines were located and their unique mine number logged, then removed from dataset.

Another issue identified was missing fields / additional fields required. A number of the datasets were obtained with LSOA names only, not LSOA codes. As a result, for the data to be usable and joins possible, the LSOA codes must be added.

Figure 10 – Corrected mine location data

Get & Transform Data			
E7			
A	B	C	D
Indicator	Date	Area	Value
% of people: 0 - 15 years	2015	Isle of Anglesey 001A	17.3
% of people: 0 - 15 years	2016	Isle of Anglesey 001A	17.6
% of people: 0 - 15 years	2017	Isle of Anglesey 001A	18
% of people: 0 - 15 years	2018	Isle of Anglesey 001A	18.3
% of people: 0 - 15 years	2019	Isle of Anglesey 001A	17.8
% of people: 0 - 15 years	2020	Isle of Anglesey 001A	17.7
% of people: 0 - 15 years	2021	Isle of Anglesey 001A	17.7
% of people: 0 - 15 years	2022	Isle of Anglesey 001A	16.9
% of people: 0 - 15 years	2023	Isle of Anglesey 001A	16.5
% of people: 0 - 15 years	2024	Isle of Anglesey 001A	16.1
% of people: 0 - 15 years	2015	Isle of Anglesey 001B	15.7
% of people: 0 - 15 years	2016	Isle of Anglesey 001B	15.1

Figure 11 – All Age data is provided with an LSOA name but not an LSOA code to accompany it

Data Manipulation and transformation

At this stage the data sets were prepared ready to be joined in QGIS software. The Age datasets and Respiratory health datasets were both obtained without LSOA codes, meaning that the data required processing to add these, enabling them to be joined, as documented in figures 13 and 14 below.

A	B
Isa21cd	Isa21nm
W01000003	Isle of Anglesey 001A
W01000004	Isle of Anglesey 001B
W01000004	Isle of Anglesey 001B
W01000005	Isle of Anglesey 005A
W01000006	Isle of Anglesey 006A
W01000007	Isle of Anglesey 009A
W01000007	Isle of Anglesey 009A
W01000007	Isle of Anglesey 009A
W01000007	Isle of Anglesey 009A
W01000008	Isle of Anglesey 008A
W01000009	Isle of Anglesey 007C
W01000010	Isle of Anglesey 002A
W01000011	Isle of Anglesey 008B
W01000011	Isle of Anglesey 008B
W01000011	Isle of Anglesey 008B
W01000011	Isle of Anglesey 008B
W01000012	Isle of Anglesey 006B
W01000013	Isle of Anglesey 005B

Figure 12 – List of all LSOA names and corresponding LSOA codes

A	B	C	D	E
Indicator	Date	LSOA code	Area	Value
% of people: 16 - 64 years (working age)	2015	W01000003	Isle of Anglesey 001A	61.2
% of people: 16 - 64 years (working age)	2016	W01000003	Isle of Anglesey 001A	60.8
% of people: 16 - 64 years (working age)	2017	W01000003	Isle of Anglesey 001A	59.8
% of people: 16 - 64 years (working age)	2018	W01000003	Isle of Anglesey 001A	59.6
% of people: 16 - 64 years (working age)	2019	W01000003	Isle of Anglesey 001A	59.4
% of people: 16 - 64 years (working age)	2020	W01000003	Isle of Anglesey 001A	59.6
% of people: 16 - 64 years (working age)	2021	W01000003	Isle of Anglesey 001A	59.6
% of people: 16 - 64 years (working age)	2022	W01000003	Isle of Anglesey 001A	59.9
% of people: 16 - 64 years (working age)	2023	W01000003	Isle of Anglesey 001A	60.3
% of people: 16 - 64 years (working age)	2024	W01000003	Isle of Anglesey 001A	59.9
% of people: 16 - 64 years (working age)	2015	W01000004	Isle of Anglesey 001B	53.2
% of people: 16 - 64 years (working age)	2016	W01000004	Isle of Anglesey 001B	53.2
% of people: 16 - 64 years (working age)	2017	W01000004	Isle of Anglesey 001B	53.2
% of people: 16 - 64 years (working age)	2018	W01000004	Isle of Anglesey 001B	53
% of people: 16 - 64 years (working age)	2019	W01000004	Isle of Anglesey 001B	52.9

Figure 13 – Age data with the added field for LSOA code

Minescan utilises the periodic table symbol for all metals mined – to make this easier to read for a layman I have changed these to the actual metals. Data can be manipulated from within the QGIS software, the field calculator allows for a nested script to be entered to change the periodic table symbols to the metals name.

minesite_n	minesite_a	region	ngr	geology	metals_sol	ref_	grid_ref	x_coord	y_coord
Abbey Consols	Florida	Central Wales	SN743661	L. Pal. Sed.	Pb, Zn, Ag	1d, 2f	274300 266100	274300.0000000000000000	266100.0000000000000000
Aber Glaslyn		Snowdonia	SH594462	L. Pal. Sed/Ig	Cu		259400 346200	259400.0000000000000000	346200.0000000000000000
Aber Hirnant		Arans & Berwyns	SH957324	L. Pal. Sed	-	2a	295700 332400	295700.0000000000000000	332400.0000000000000000
Aber Mine		Snowdonia	SH666728	L. Pal. Sed	Fe		266600 372800	266600.0000000000000000	372800.0000000000000000
Aber Rhuddant		Central Wales	SN767770	L. Pal. Sed	-	2f	276700 277000	276700.0000000000000000	277000.0000000000000000
Aberdaunant		Central Wales	SN906865	L. Pal. Sed	Pb, Zn, Ag	1c, 2e	290600 286500	290600.0000000000000000	286500.0000000000000000
Aberduna		NE Wales	SI206618	Carb Lst	Pb, Ag	1b	320600 361800	320600.0000000000000000	361800.0000000000000000
Aberffwrdd		Central Wales	SN688788	L. Pal. Sed	Pb, Cu	1d, 2f	268800 278800	268800.0000000000000000	278800.0000000000000000
Abercynon area (Isuall)		NE Wales	SH936782	Carb Lst	-		293600 378200	293600.0000000000000000	378200.0000000000000000

Figure 14 – Minescan raw data – Metals and period table symbols

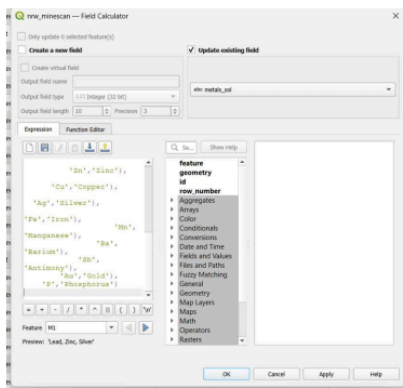


Figure 15 – Field calculator within QGIS software, used to rename chemical symbols to metal names

Thematic Maps

‘A thematic map is a map designed to show the spatial distribution of a single subject or variable—such as population density, disease rates, climate, or economic activity—using visual techniques like colour, shading, or symbols to highlight patterns and relationships across geographic space’[10]. To create a thematic map several layers must be integrated, allowing for the pulling of data from separate areas to cover things such as mine locations, health rates, deprivation rates and working ages.

Joining and preparing data for thematic mapping

The separate datasets require a universal common denominator for them to be linked. In this instance, it requires the addition of a .csv detailing all LSOA codes, along with the corresponding local authority they operate under. This process is completed by first loading the shapefile detailing the boundaries of the LSOA's. The local authority data is stored in as a .csv which is brought in as a comma separated value layer. Data is then joined on the common field, in this instance it is joined on the LSOA21code and lsoa21cd fields.

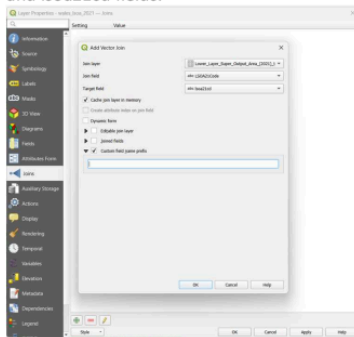


Figure 16 – Join layer

	id	loc1	loc2	name	JOAG Name English	JOAG Name Welsh	whorlity District 2	whorlity District 2	city District 2B	Object
1	W0100194	Pembrokeshire	W0300000449	Pembrokeshire	St Benfro 112E	W060000009	Pembrokeshire	St Benfro		3416
2	W0100191	Sweeney DSSG	W0300000449	Sweeney DSSG	Sweeney DSSG	Abertawe 105S	W060000011	Sweeney	Abertawe	3478
3	W0101134	Vale of Glamor...	W0300000449	Vale of Glamor...	Vale of Glamor...	Bro Morgannwg...	W060000014	Vale of Glamor...	Bro Morgannwg	3485
4	W0101265	Rhonda Cynon	W0300000449	Rhonda Cynon	Rhonda Cynon	Rhonda Cynon	W060000016	Rhonda Cynon	Rhonda Cynon	3484
5	W0101256	Rhonda Cynon	W0300000449	Rhonda Cynon	Rhonda Cynon	Rhonda Cynon	W060000016	Rhonda Cynon	Rhonda Cynon	3484
6	W01000264	Ffintshire 012A	W0300000449	Ffintshire 012A	Ffintshire 012A	Sy Ffint 012A	W060000005	Ffintshire	Sy Ffint	3480
7	W01000898	Neath Port Talb...	W0300000449	Neath Port Talb...	Neath Port Talb...	Castell-nedd 012A	W060000012	Neath Port Talbot	Castell-nedd	3466
8	W0101565	Momonthouth...	W0300000449	Momonthouth...	Momonthouth...	St Fynwy 006D	W060000021	Momonthouth	St Fynwy	3520

Figure 17 – Attribute table now containing all fields

Creating a thematic map for Respiratory health data

Once the above preparation is completed, allowing for a base map with geographical outlines for LSOA's which also allows for an attribute with local authority data, the respiratory data can be added. This data is only available at local authority level as an average, however; the .csv provided does give 10 years of averages. Data can either be prepared and filtered prior to the layer being added by only leaving the chosen year, or a filter can be applied once the layer is included. As a filter gives the best option and allows for changing this is the method shown below.

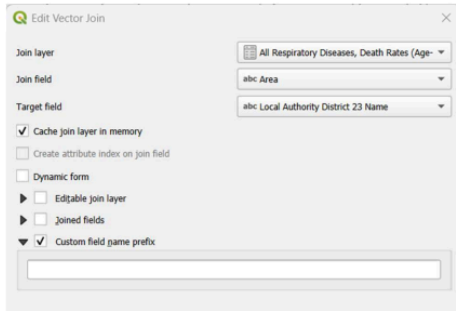


Figure 18 – Respiratory data join.

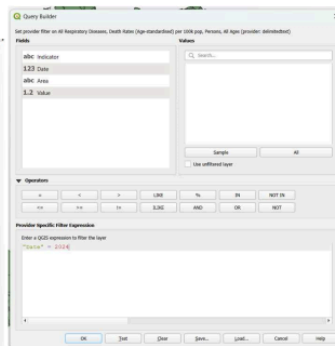


Figure 19 – Filter applied to Respiratory data layer

The layer is joined by the same method, this time matching the data on the Area field which matches with the local authority district. A filter is then applied restricting the output to the chosen year of 2024. Once a filter is applied, the data can be displayed. In this instance a graduated approach is most applicable. It applies a scale to the data, allowing you to display concentration by tone of colour.

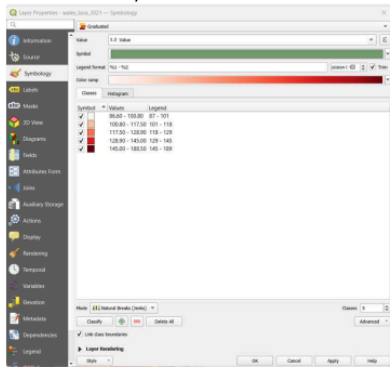


Figure 20 – graduated colour applied to data

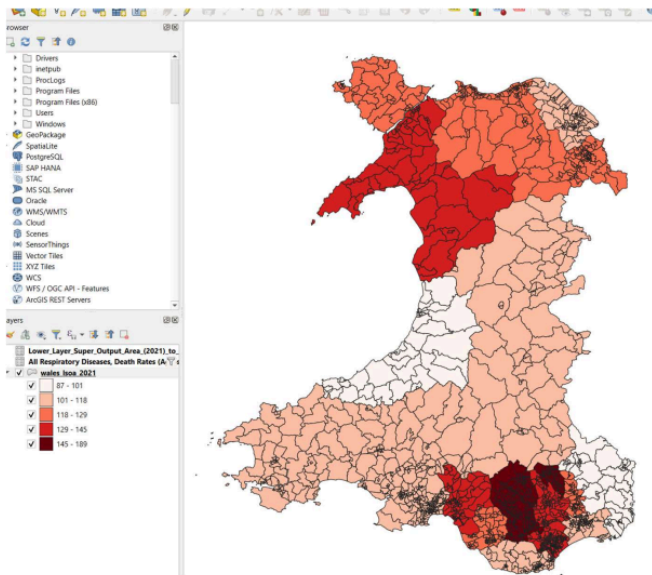


Figure 21 – Detailed choropleth map for respiratory death rate average per local authority in 2024

The same process is followed for creating a choropleth for hospital admissions – this time the process fails at the stage for adding graduation to the map. Research identifies the cause as commas utilised as separators in the data. This means QGIS reads the data as text, not numeric. This was rectified from within QGIS, opening the attribute table and utilising the field calculator function to create a new field which contains the data without commas.

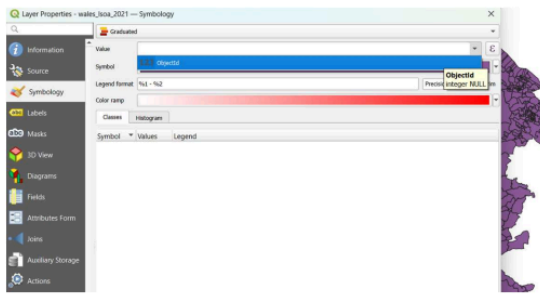


Figure 22 – Error not allowing for graduation

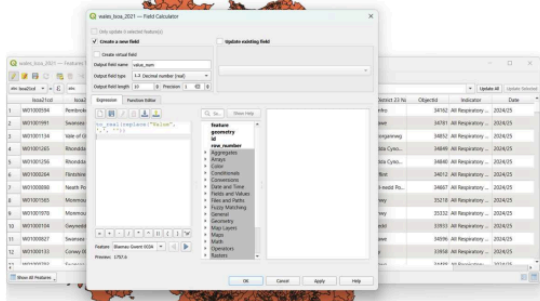


Figure 23 – QGIS field calculator to create new field

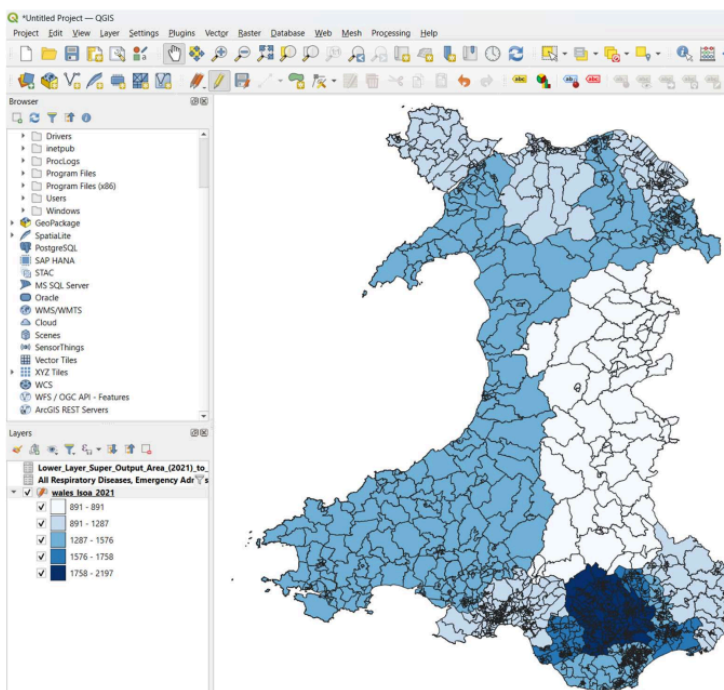


Figure 24 - Detailed choropleth map for respiratory hospital admission average per local authority in 2024

Graduated thematic maps for percentage of population working

The same method as above is applied to produce 2 choropleth maps for population ages, one for working age (16-65), and another for retirees (65+). These choropleth maps are a good example of how data can be misleading, both use graduation, but the key identifies that the scale has changed, meaning the display is completely different.

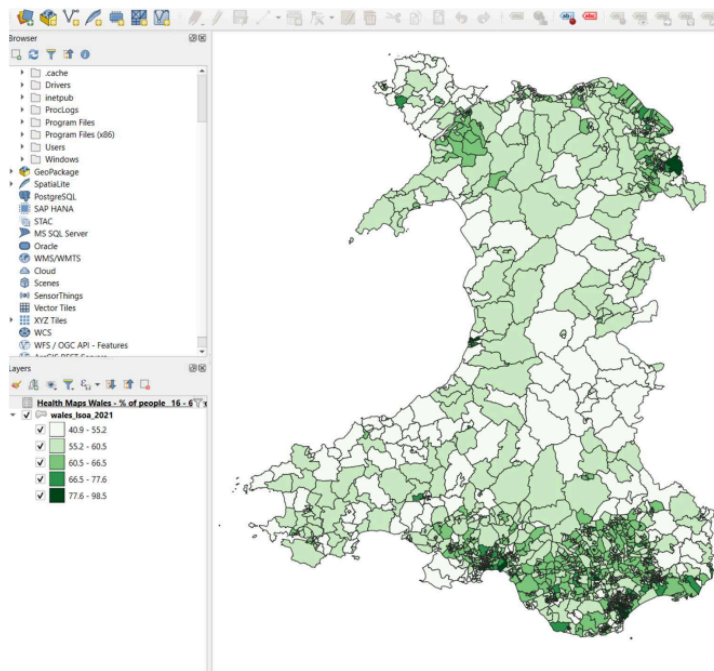


Figure 25 – percentage of people working Age graduated

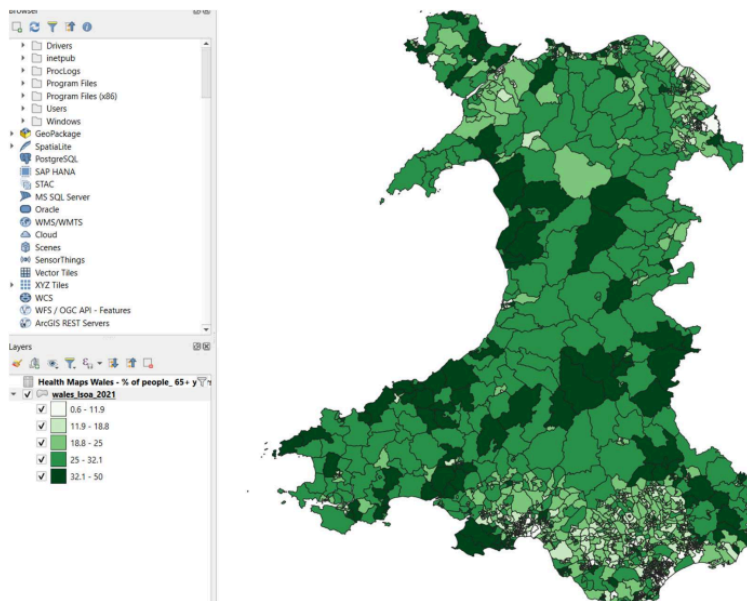


Figure 26 – percentage of people retired graduated

A choropleth was created with 2 variations of graduation, showing both the working age and WIMD health rating for the area. This did not work well, the data was too granular, needing to be zoomed in to a much greater level, along with poor visibility, it was not appropriate for this projected.

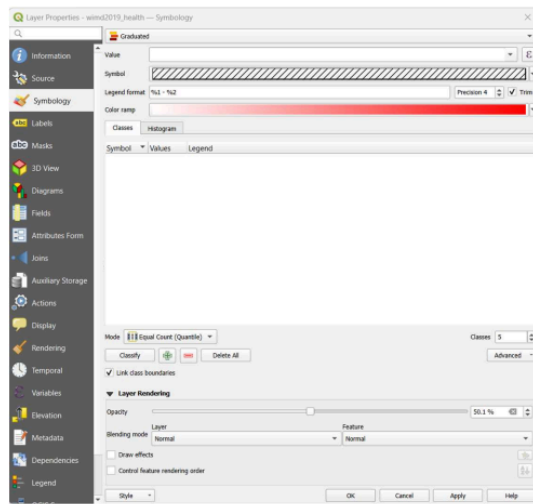


Figure 27 – Setting cross hatch for bivariate graduated map

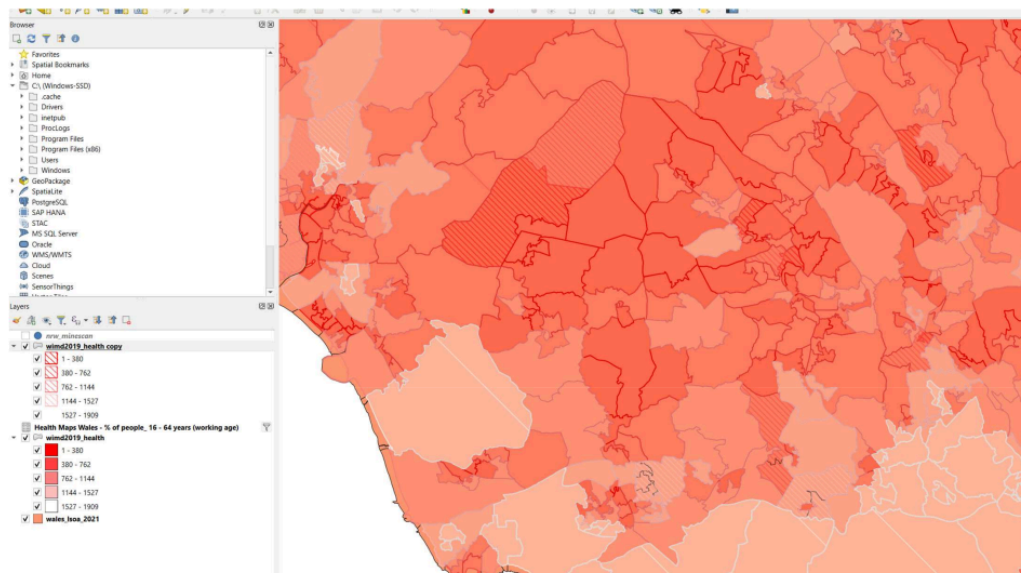


Figure 28 – Bivariate graduated choropleth

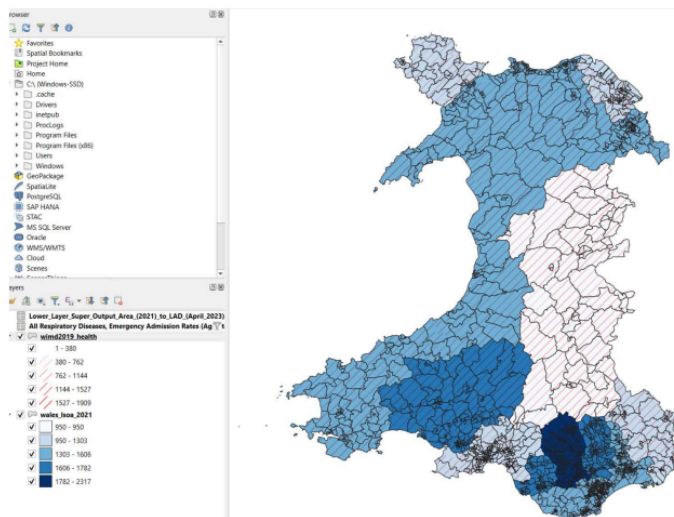


Figure 29 - Bivariate graduated map showing hospital admission rates against WIMD Health rating

Graduated thematic map for WIMD health rating

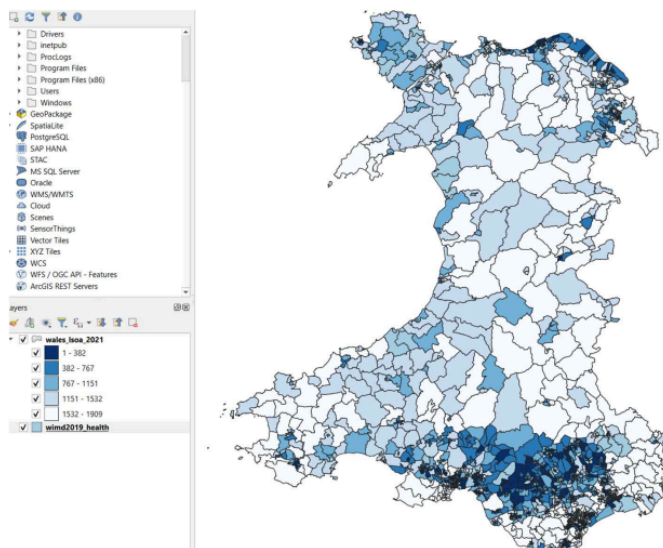
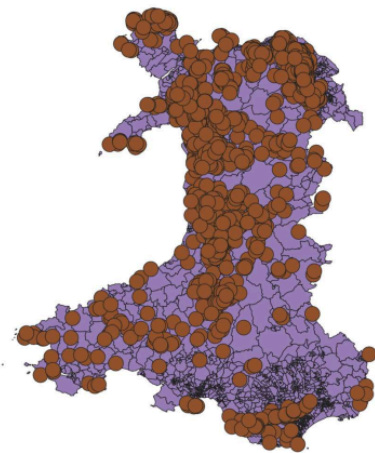


Figure 30 – WIMD – Health choropleth

Insertion of Mine location data

As covered in the data cleansing stage, the original Minescan data did feature some errors which were noted in the metadata provided with the data. This data was removed from within the QGIS software. The metadata identified the location of the errors and a manual search allowed for the identification of 8 mines which were then deleted from the attribute table by simply removing the entire row for the unique mine number. Once the data was added a buffer zone was applied to mine locations, an initial size of 4km was chosen as a historically accurate average is stated as 3-5km as concluded by early UK National Travel Surveys [11].



12

Figure 31 – Mine locations with a 4km buffer zone

Another method of detailing the mine density is via heatmap. A heatmap was created to check the difference between the 2 graphs. This is accessed via the processing toolbox as displayed below. As the heatmap overpowered the other elements, the project continued with the point data.

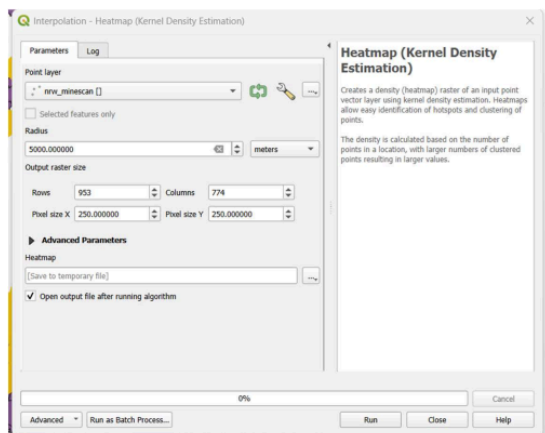


Figure 32 – Heatmap raster creator

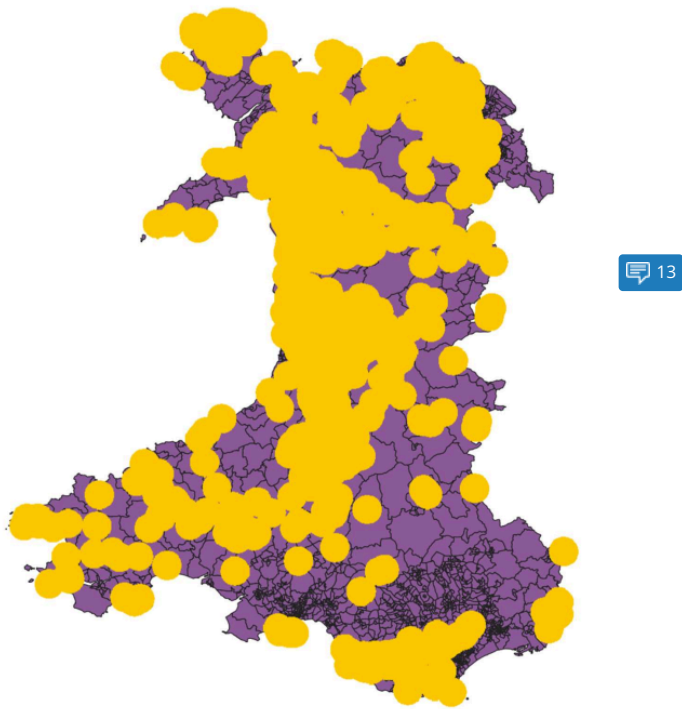


Figure 33 – Heat map created for mine location

Final Outputs

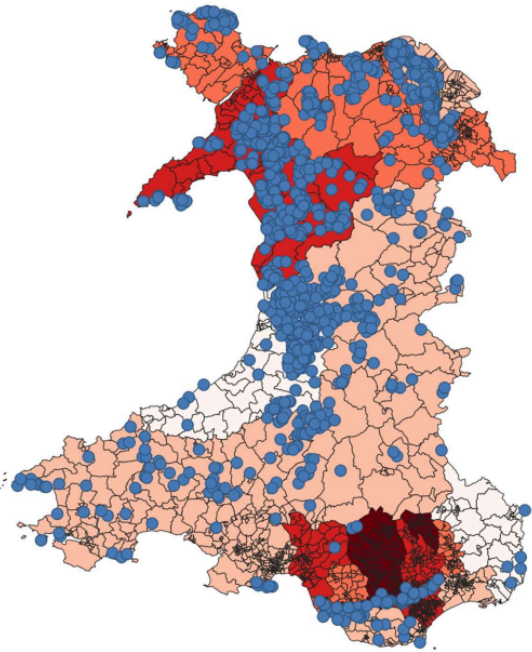


Figure 34 – Graduated Respiratory Death Rates with detailed mine locations

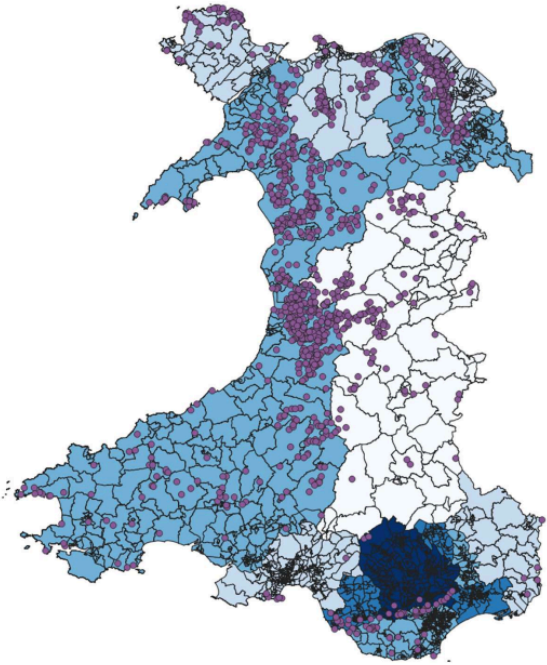


Figure 35 – Graduated Respiratory hospital admission rates including mine locations

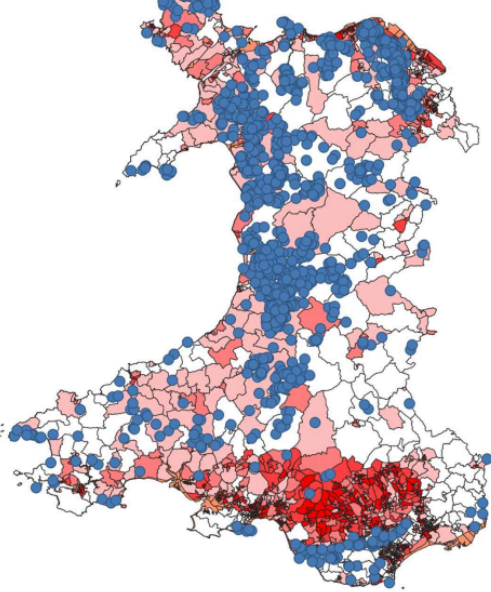


Figure 36 – WIMD – Health with overlaid mine locations

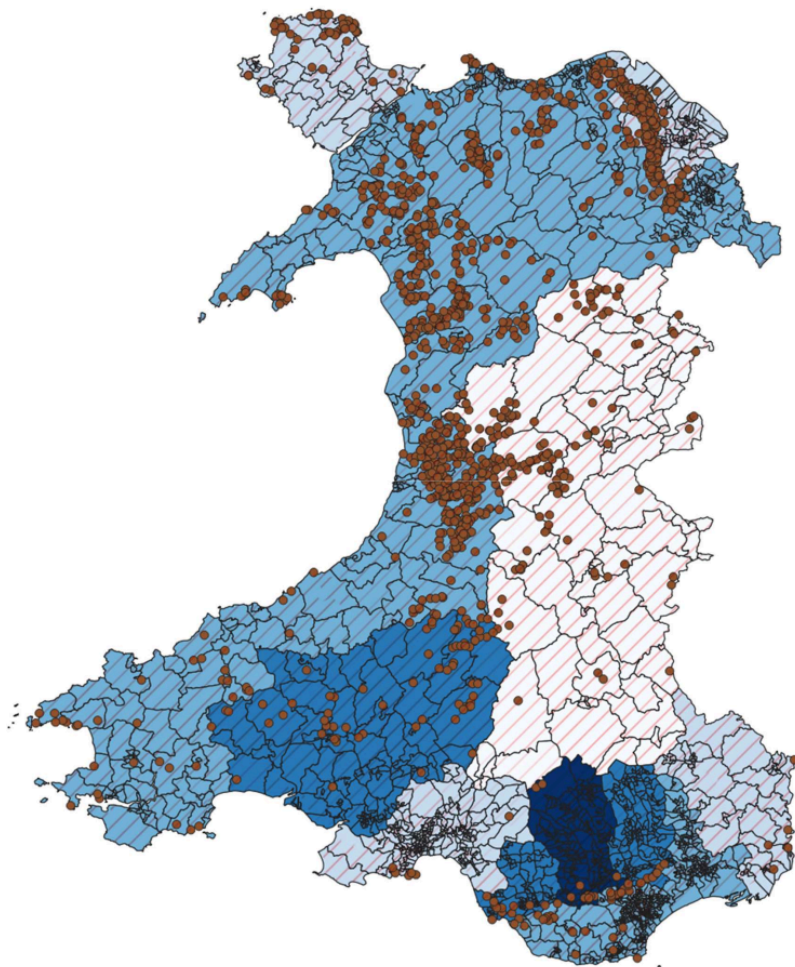


Figure 37 – Bivariate graduated WIMD Health ranking against respiratory hospital admission rates with mine locations added

Outcome

Findings from thematic maps

Mine locations: Wales has a large number of mines spread across the entire country, while individual mines can and do occur there are also large clusters of mines at the same locations.

Death and hospital admission rates follow a rough correlation – as would be expected from the data collections, with an expectation of someone experiencing respiratory issues being seen several times by health professions before expiring.

The highest death rates are located in 2 clusters, one in the north and 1 in the south. Whilst the cluster in the north does match with the highest mine density in the country - the south does not, although it is located at a cluster with lower mine density.

The highest hospital admission rates which are located in the south of the country, whilst in the same vicinity of a cluster of mines, does not have the same mine density as in the north. This would support there being another reason for increased hospital admission rates – calling for further investigation.

Historically mining communities are reported as areas of deprivation, however, this isn't supported by the data. The WIMD supports the south as being the most deprived which follows a cluster of mines, but mine density is significantly higher in the north of the country which scores lower in the WIMD, meaning it does not track with mine location.

Analysis

Principles of thematic mapping and data visualisation

Clear Theme – A thematic map should have a singular, clear theme. This should be utilised to bring the intended data to the forefront of what you intend to display to the audience.

Appropriate visuals – Themes should be chosen to accentuate a desired element, multiple colours, differing schemes with shading and graduation of colours etc should not distort or confuse the audience, leaving them struggling to understand the message.

Appropriate map selection – This project utilises choropleth maps, however, there are also dot density, proportional symbol, heat, flow and graduated symbol maps. They each represent data in a different method. The map choice should support the message you are delivering to the audience.

Layering – Thematic maps allow for multiple layers to be added allowing for differing datasets to be added to provide a focus for the data you are presenting.

Challenges and limitations encountered

Challenges

Data integration: The Minescan dataset consisted solely of point features with coordinate geometry, and contained no LSOA or MSOA identifiers. Because of this, it was not possible to perform a direct attribute join with administrative boundary layers. Attempts to add joinable fields outside QGIS were unsuccessful, as the dataset lacked any inherent geographic codes that could be matched to LSOA boundaries. Without a shared key field, QGIS correctly reported that there was “nothing to join on,” meaning the Minescan points could only be spatially related to LSOAs through spatial operations (e.g., spatial join, point-in-polygon), not attribute joins.

Data manipulation: A significant amount of time was spent attempting to edit attribute tables outside QGIS. Through this process, I learned that the .dbf file is the attribute table for shapefiles, but modern versions of Excel can open .dbf files yet cannot save them back in that format. LibreOffice previously supported .dbf editing, but current versions also fail to save reliably due to DBF field-length and decimal-precision limitations. GeoPackage avoids these issues and allows attribute tables to be exported to CSV for editing in Excel. However, once edited, the CSV must be re-imported into QGIS and requires a join field to reconnect it to the spatial layer. Because the Minescan dataset does not contain LSOA or MSOA codes, no join field could be created for this dataset, making attribute-based integration impossible. Additional issues were encountered with data from separate sources using differing standards, such as comma separation in some and not in others.

Data availability: A further challenge was the limited availability of data at LSOA resolution. While some datasets were available at this level, key health indicators were only published at Local Authority scale. This restricts the granularity of analysis and means that any health-related spatial patterns can only be interpreted at a broad administrative level, rather than at the fine-grained neighbourhood scale that LSOAs provide.

Limitations

Data Quality: The Minescan dataset includes a field indicating the type of metal extracted at each site. However, this field is incomplete: many entries are blank, and the dataset does not distinguish between metalliferous and non-metalliferous mines. Wales has a history of coal and slate extraction, neither of which appear in the dataset. The absence of this information restricts the ability to analyse whether different mine types (e.g., metal vs coal vs slate) may have had differing long-term impacts on environmental or respiratory health.

Interaction: QGIS allows for production of static spatial maps, showing mine locations, buffer zones, or LSOA boundaries. These maps provide a clear visual overview and allow users to investigate individual features, but the interaction is limited to that point. Other tools such as Power BI offer a more interactive prospect. Users can filter and explore relationships between variables. This interaction can reveal patterns that are not visible in static GIS images.

Conclusion

Analysis of the geodemographic data produced through GIS allows for several factors to be considered quickly and accurately – it also highlights the shortfalls of data collected, opening up further questions and avenues for further research. The thematic maps revealed that WIMD does not follow mine locations. This can be explained by either mining positions becoming sparser, but also more qualified and so well paid, or that after the closing of many mines in the last century, the communities have now managed to bounce back and become prosperous.

Mine density does not necessarily correlate with increased hospital admission rate or death rates. There is not enough information available at this stage to rule this out or prove that there is no connection. Several areas remain which will need to be explored before this can be confirmed:

Mine size – The data provided the location of mines, it does not include anything about the size of the workings. Large mines can have 2000+ workers but still show as a single data point, whilst a group of small drift mines might have 6-10 workers each but show as several data points.

Mine type / target – different mines will target different minerals and/or metals. The north of Wales has a higher proportion of open or surface mines, with the associated lower risk (for respiratory conditions)

Time – Many mines closed in the 70's. Epidemiological studies suggest miners at this time had an average lifespan of 50-60. Someone of working age at this point in time would now be 65+, meaning they potentially would have already passed. All death rates and admission rates are limited to 2024 and so likely wouldn't include these workers.

Suggestions for further analysis

Further investigation could be completed by way of obtaining more granular data for the datasets utilised, along with adding additional datasets. Currently the datasets provided are kept at MSOA level for public access due to the sensitivity of the information they contain. This data is likely able to be broken down into LSOA level available for qualified individuals. This would allow for much further analysis. In addition to this, were specific conditions to be flagged for specific metalliferous pollutants/contaminants then this could be cross referenced with the mine type to target specific areas to see if they have corresponding increases in the specific health conditions, allowing for an argument to be made towards correlation and causation.

Data utilised is averaged and has some large assumptions included with its usage. The age series is an average, it only has 3 groupings which are roughly, pre working age, working age and retirees – however, there will be outliers with this data. Some will have worked earlier than 16, some will have continued working past retirement age whilst others will have retired early. The dataset is from the 2021 census, whilst the target groupings will have been working in the mining industry approximately 40-60 years prior to this census.

Mine data is location only, further data and analysis including the size of the workings, with the number of employees could potentially change the outlook of the choropleths provided.

References

- [1] D. I. Heywood, S. Cornelius, and S. Carver, *An Introduction to Geographical Information Systems*, 4th ed. Harlow, U.K.: Prentice Hall, 2011.
- [2] Office for National Statistics, *Lower Layer Super Output Areas (LSOAs) – Wales – 2021*. DataMapWales, Welsh Government, Nov. 25, 2025. Accessed: May 2, 2026. [Online]. Available: <https://geoportal.statistics.gov.uk/datasets/ons::lower-layer-super-output-areas-december-2021-boundaries-ew-bsc-v4-2/about>
- [3] Natural Resources Wales, *MINESCAN* (point dataset of metalliferous and non-metalliferous mine sites in Wales), Dataset Publication Date: Jun. 18, 2025. Accessed: May 2, 2026. [Online]. Available: https://metadata.naturalresources.wales/geonetwork/srv/api/records/NRW_DS99411
- [4] Welsh Government, *Health Maps Wales: 2023* (Working Age Percentage indicator). GOV.WALES, Apr. 16, 2024. Accessed: May 2, 2026. [Online]. Available: <https://www.healthmapswales.wales.nhs.uk/>
- [5] Welsh Government, *Welsh Index of Multiple Deprivation (WIMD) 2025*. DataMapWales, Publication date: Nov. 25, 2025. Accessed: May 2, 2026. [Online]. Available: <https://datamap.gov.wales/maps/8160>
- [6] Welsh Government, *Health Maps Wales: 2023* (Respiratory Diseases indicator). GOV.WALES, Released Apr. 16, 2024. Accessed: May 2, 2026. [Online]. Available: <https://www.gov.wales>
- [7] Office for National Statistics, *LSOA (2021) to Local Authority District (April 2023) Best Fit Lookup in England and Wales*. ONS Open Geography Portal, Apr. 2023. Accessed: May 2, 2026. [Online]. Available: <https://geoportal.statistics.gov.uk/datasets/ons::lsoa-2021-to-local-authority-district-april-2023-best-fit-lookup-in-ew/explore>

- [8] U.K. Government, "UK General Data Protection Regulation (UK GDPR), Article 5: Principles relating to processing of personal data," *Legislation.gov.uk*, 2016. Accessed: May 3, 2026. [Online]. Available: www.legislation.gov.uk
- [9] P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, *Geographic Information Systems and Science*, 4th ed. Hoboken, NJ, USA: Wiley, 2015.
- [10] A. H. Robinson, J. L. Morrison, P. C. Muehrcke, A. J. Kimerling, and S. C. Guptill, *Elements of Cartography*, 6th ed. New York, NY, USA: John Wiley & Sons, 1995.
- [11] C. Pooley and J. Turnbull, *Travel, Transport and the Urban Environment*, Manchester, UK: Manchester University Press, 1999.

62/100

PAGE 1

PAGE 2



Comment 1 |  Transferable Skills(10)

Contents page, numbered sections, consistent figure labelling throughout



Comment 2 |  Transferable Skills(10)

Abstract should state key findings or methodology



Comment 3 |  Cognitive and Intellectual(25)

formal research question or hypothesis missing

PAGE 3



Comment 4 |  Literature (25)

- GIS defined with Heywood et al

- brief discussion on principles of thematic mapping section shows awareness of visual variables



Comment 5 |  Literature (25)

- engagement with cartographic semiology could have been shown

-
literature on choropleth design, colour theory or perceptual effectiveness of map types missing



Comment 6 |  Knowledge&Understanding(20)

Six datasets integrated

**Comment 7** | Cognitive and Intellectual(25)

Correctly identifies LSOA granularity as a constraint on what can be concluded

**Comment 8** | Knowledge&Understanding(20)

Buffer zone created but could have been used analytically — no mine-adjacent vs non-adjacent comparison

**Comment 9** | Cognitive and Intellectual(25)

Miner lifespan / census timing point is genuinely insightful — shows independent thinking

**Comment 10** | Cognitive and Intellectual(25)

Additional data suggestions like NRW Contaminated Land Register, PHW ICD-10 coded data, DEFRA air quality could have been explored

**Comment 11** | Practical Application (20)

Data cleansing documented

**Comment 12** | Practical Application (20)

buffer zones, graduated symbology all demonstrated

**Comment 13** | Practical Application (20)

report doesn't discuss what design choices were made in these layouts or why



Comment 14 |  **Practical Application (20)**

print layout / atlas features to be used — map outputs appear to be basic canvas screenshots

PAGE 17

PAGE 18

PAGE 19

PAGE 20

PAGE 21

GRADING FORM: GIS

ADAM SMITH

62

LITERATURE (25)



13

KNOWLEDGE&UNDERSTANDING(20)



14

COGNITIVE AND INTELLECTUAL(25)



16

PRACTICAL APPLICATION (20)



12

TRANSFERABLE SKILLS(10)



7