

Champton_ACCB5027

by Christopher Hampton

Submission date: 12-Jan-2024 02:49PM (UTC+0000)

Submission ID: 221760989

File name: 32525_Christopher_Hampton_Champton_ACCB5027_545816_1665895231.docx (1.36M)

Word count: 3049

Character count: 16080

Trinity Saint David University of Wales

Geographical Information Systems (ACCB5027)

Author:

Chris Hampton

2116517@student.uwstd.ac.uk

Tutor:

Beverley Holland

beverley.holland@uwtsd.ac.uk

12/01/2024

Contents

Table of Figures.....	2
1. Introduction	4
1.1 Hypothesis.....	4
1.2 Abstract.....	4
2. Research and Data Collation	4
2.1 Research.....	4
2.1 Data Collection.....	5
2.1.1 Mapping Data Collection.....	5
2.1.2 Downloading Mapping Data	6
2.2.1 Visual Data Collection	7
2.2.2 Manipulating Data	8
3. Geographical Information System Build	9
3.1 Importing Map Files	9
3.2 Plotting Data	10
3.2.1 Thematic Mapping	12
4. Analysis	14
4.1 Thematic Analysis	14
4.2 Point Analysis	15
5. Conclusion.....	17
5.1 Further Steps.....	17
5.2 Final Conclusion	17
References	17

Table of Figures

Figure 1 - MSOA Data Selection	6
Figure 2 - Boundary Data File Formats	6
Figure 3 - SHAPE File Boundary Data Download.....	6
Figure 4 - GP Surgery and Branch Download.....	7
Figure 5 - Infuse Data Variety	7
Figure 6 - Infuse Data	7
Figure 7 - MSOA to LSOA example	8
Figure 8 – Shapefiles	9
Figure 9 - Map Projection	9
Figure 10 - Primary Care Boundaries	10
Figure 11 - Boundary Data	10
Figure 12 - Importing Data	10
Figure 13 - Data Types	11

Figure 14 - Geocoding 1	11
Figure 15 - Geocoding 2	12
Figure 16 - SQL Selection	12
Figure 17 - SQL Creating Subset	12
Figure 18 - Step 2 Thematic Mapping	12
Figure 19 - Step 1 Thematic Mapping	12
Figure 20 - Specify Join.....	13
Figure 21 - Temporary Column	13
Figure 22 - Thematic Map	13
Figure 23 - MSOA	14
Figure 24 - Health Board	14
Figure 25 - Primary Care Clusters.....	15
Figure 26 - GP Branches and Sites	15
Figure 27 - Isolate Boundaries	15
Figure 28 - Joining selections back into map.	16
Figure 29 - Boundary Comparison	16
Figure 30 - Comparison	16

1. Introduction

1.1 Hypothesis

Is there correlation between deprivation and access to health services in Wales?

1.2 Abstract

There are many factors that can affect the health of an individual but can certain factors that are usually beyond the control of the subject contribute to the health status of a population. These factors can range from socio-economic elements to the geographic location of where someone lives.

By researching and building a geographical information system the hypothesis of: *Is there correlation between deprivation and access to health services in Wales?* can be tested to prove or disprove whether there is correlation.

2. Research and Data Collation

2.1 Research

Deprivation can be described as a basic lack of necessity within a society that should be considered the norm.

According to Wilkinson et al. there are many inequalities that can be associated with the prevalence of deprivation in an area [1]. A common inequality is income. Direct correlation can be attributed to the health of a population depending on the affluence of the environment. There are many different causes that affluence can have an influence on health, one possibility is the reliance on state-funded health care. As an example, a person that has a significantly higher income is more than likely to self-fund their healthcare, through private practitioners, than someone who is on a lower or moderate income.

The country of Wales has a unique geographic landscape in that the north and the south coasts of the country are highly populated with relatively good transport links and roadways. However, a high proportion of the area between these areas is rural and semi-rural. This can be challenging for healthcare providers to provide high quality and accessible care that caters for the population.

The journal article titled '*Understanding shortages of sufficient health care in rural areas*' [2] details analysis from numerous papers on the subject and concluded that healthcare shortages can fall into the five following categories:

- **Provider Shortages**
Healthcare organisations can suffer from lack of resources and skills when recruiting for positions to provide care in rural areas.
- **Maldistribution**
Where there are rural locations there is usually the same prevalence of conditions but lack the same volume as a more populated area. This directly affects funding for care within these areas.
- **Quality Deficiencies**
The quality deficiencies can both be resource personnel and healthcare facilities.
- **Access Limitations**
The access to healthcare settings and facilities can be difficult for the population within a rural area. This can be attributed to transport links.

- **Inefficient Utilisation of Health Care Services**

Often facilities and services are underutilised, this creates a huge waste of resources that can be used elsewhere.

Many of the points raised above can be analysed using a Geographical Information System (GIS). The basic idea of a GIS is to plot data of something – somewhere [3]. More and more data is generated everyday – the idea of being able to plot data and reference it to a location is where the power of a GIS can be extrapolated. Not only to determine the relationship between an area and a piece of data or information but to compare one location to another along with their corresponding information [4].

The offerings of mapping software have improved vastly over the years and being able to build a GIS can be as simple or as complex as the user requires. Many businesses use mapping technology to deliver their business goals, whether it's a delivery company using a GIS to map routes for its drivers or a health board or trust making decisions on where to locate services. This location intelligence can be imbedded into the organisations Business Intelligence tools – such as ARCGIS being available in Microsoft's PowerBI.

There is also stand-alone GIS technology, and the software being used in this research assignment, such as MapInfo Professional. MapInfo was initially founded in the 1980s and has been through numerous iterations to become a robust mapping software tool [5].

2.1 Data Collection

2.1.1 Mapping Data Collection

MapInfo Professional can handle a collection of different file formats to produce intuitive mapping solutions. MapInfo's native file format is called 'TAB'. A 'TAB' file is a geospatial vector file specifically created for MapInfo software. This file type is used as a link between different files and holds meta-data from the datasets involved.

A high proportion of data services, due to the popularity of MapInfo, enable a user to download TAB file formats from their repository.

One such data service is UK Data Service. This service hosts a large amount of data that can be used for research and development. One key area they produce data on is Census Data. Census Data is an important aspect for a lot of organisations, such as Government, Local Authorities and Healthcare organisations [6].

The geographical landscapes needed for the analysis include:

- **Middle Super Output Areas (MSOA)**

MSOAs are made up of a small number of Lower Super Output Areas. Usually up to 6,000 households and have a population density of up to 15,000 [7].

- **Welsh Health Board Boundary Data**

The health boards in Wales are split into seven areas. Although there are smaller healthcare providers, such as Velindre Cancer Trust and the Welsh Ambulance Service, the seven are classed as the major health boards in Wales.

- **General Practitioner Surgery Data**

Compared to the MSOA and Health Board data, the GP Surgery data isn't boundary data. It is, however, point data. It highlights a location, or point, on a map.

2.1.2 Downloading Mapping Data

Using the UK Data Service Portal – boundary data can be downloaded in various formats. [8]

The data selection service allows the user to download specific data that meets their needs. In this example in figure 1 the Welsh Middle Super Output Area boundary data is being selected. [9]

The file format is also available for the user to select, depending on the needs and software used. In Figure 2 the options show that the data can be downloaded in SHAPE, KML, MAPINFO and CSV formats.

For this case, as SHAPE file format is selected. A SHAPE file (.shp), is developed by an organisation called 'EVRI'. It has been devised to use in most mapping software solutions and provides interoperability between software types.

The screenshot shows the 'What and Where?' and 'Format' sections of the UK Data Service Portal. Under 'What and Where?', the 'Wales' region is selected, and the '2011 and later' time period is chosen. The 'Format' section shows a list of available datasets, with 'Wales Middle Layer Super Output Areas, 2011' selected. Below this, a list of 'Areas of Interest' is shown, including 'Blaenau Gwent', 'Bridgend', 'Cardiff', 'Cardiff and Vale of Glamorgan', 'Ceredigion', and 'Conwy'. The 'Search Summary' section indicates the target geography is 'Wales Middle Layer Super Output Areas, 2011' and the data format is 'SHAPE'. The 'Archive Method' is set to 'Zip'.

Figure 1 - MSOA Data Selection

N.B. Throughout the building of the GIS it was noted that the underlying data to be represented was collected in 2019, so the MSOA Boundary Data is the 2011 version rather than 2021 to prevent any missing boundaries that have been added.

The other file formats that are available is KML – which is used for applications such as Google Earth. And the CSV data file has the raw data of the boundaries but doesn't include and mapping capabilities.

The screenshot shows the 'Export Format' section of the UK Data Service Portal. The 'Export Format' dropdown menu is open, showing the following options: 'SHAPE (ESRI Shapefile format)', 'KML (Keyhole Markup Language)', 'MAPINFO (native format)', and 'CSV (Comma Separated Value - attributes only)'. The 'SHAPE (ESRI Shapefile format)' option is selected.

Figure 2 - Boundary Data File Formats

Once the data type is selected, in this case a SHAPE file, a folder containing various files is downloaded. The main .shp file is used in MapInfo to determine the boundaries and the other files are subsidiaries to enable the polygons to be built.

- TermsAndConditions
- wales_msoa_2011_clipped.dbf
- wales_msoa_2011_clipped.prj
- wales_msoa_2011_clipped.shp
- wales_msoa_2011_clipped.shx

Figure 3 - SHAPE File Boundary Data Download

For the GP Surgery Data – this is available from Data Map Wales [10]. Data Map Wales is open-source data from the Welsh Government that allows the public access to data resources regarding the Welsh population. The data can be downloaded in various formats, the same as UK Data Service.

Again, for this purpose a shapefile has been used.

Spatial data download

Figure 4 - GP Surgery and Branch Download

2.2.1 Visual Data Collection

There are several different sources to use to create the data layer for the geographic mapping software. One data platform is called Infuse. Infuse uses aggregate data from the 2001 and 2011 Census to allow users to explore different aspects of the population [11].

The data can be broken down to different granularities as shown below in figure 5.

- ☐ Wales
 - ☒ Local Authorities (22 areas)
 - ☐ Wards and Electoral Divisions (868 areas)
 - ☒ Middle Super Output Areas and Intermediate Zones (410 areas)
 - ☐ Lower Super Output Areas and Data Zones (1909 areas)
 - ☐ Output Areas and Small Areas (10036 areas)
 - ☐ Merging Wards and Electoral Divisions (25 areas)
 - ☐ Workplace Zone Layer (2710 areas)

Figure 6 - Infuse Data

Figure 5 - Infuse Data Variety

The figure above shows a selection of the topics available in InFuse. In total there is 404 topics or combination of topics to select.

However, for this assignment the Welsh Index of Multiple Deprivation (WIMD) will be used to determine the different levels of deprivation subject matters within Wales.

The WIMD is a measure of deprivation within small communities within Wales. These are broken down into Lower Super Output Areas. There are eight subject matters that are ranked between 1 (most deprived) to 1909 (least deprived) [12].

The eight subject matters are:

- Income
- Employment
- Health
- Education
- Access to Services
- Community Safety
- Physical Environment
- Housing

The ranking of these subjects is then used to create an overall WIMD rank.

Data Source: <https://stats.wales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation/WIMD-2019/welshindexofmultipledeprivation2019-by-rank-decileandquintile-lowerlayersuperoutputarea>

2.2.2 Manipulating Data

To clean up the data to make it usable with the GIS, some data manipulation and cleansing needs to take place.

The data is split into LSOA level. There are 1909 LSOAs in Wales and it was deemed to granular to create any meaningful analysis from such a large sample. There are four to five LSOAs to each MSOA and a mapping exercise was needed to aggregate up the subject matters to MSOAs.

By using a VLOOKUP in Microsoft Excel each LSOA was matched up to their corresponding MSOA. The index scores were then added up and ranked from one to 410 to give an overall MSOA index score.

For example, the MSOA **Blaenau Gwent 001** has 5 LSOAs (Beaufort 1, Beaufort 2, Beaufort 3, Rassau 1, Rassau 2).

MSOA	LSOA Name	Sum of WIMD
Blaenau Gwent 001	Beaufort 1	595
	Beaufort 2	235
	Beaufort 3	1462
	Rassau 1	1103
	Rassau 2	211
Grand Total		3606

Figure 7 - MSOA to LSOA example

The sum of WIMD is added together to give an overall score of 3606 for **Blaenau Gwent 001**.

This was repeated for the Primary Care clusters and Health Boards.

3. Geographical Information System Build

Once the boundary, point and analytical data has been gathered and cleansed it can be pulled into the MapInfo software.

3.1 Importing Map Files

Each shapefile that was downloaded previously can be seen in the file location below.

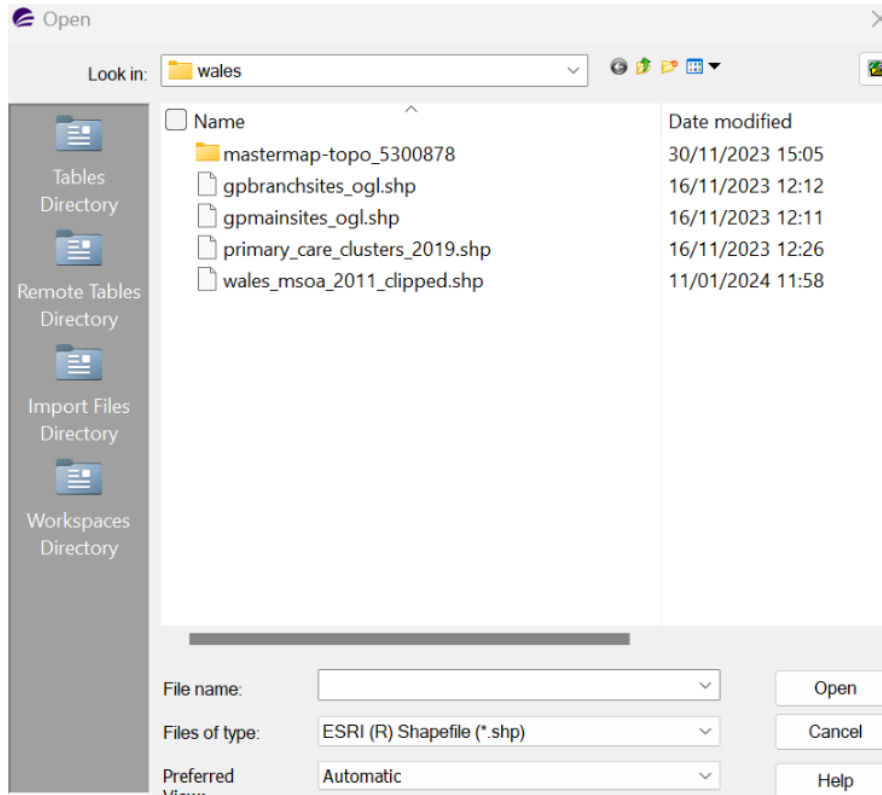


Figure 8 – Shapefiles

These are imported into MapInfo.

MapInfo allows the user to choose the projection of the boundary data. The projection is used to locate each polygon within the software. This is to anticipate the curvature of the map and apply it to a flat surface, such as a computer scene. Most data that is downloaded contains the projection within the files meta-data and usually the user does not need to change anything.

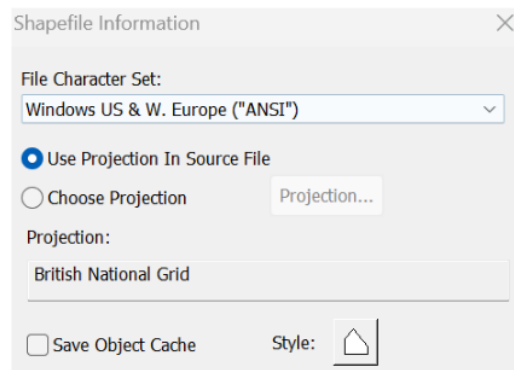


Figure 9 - Map Projection

Once the maps have been imported, they will show in MapInfo.

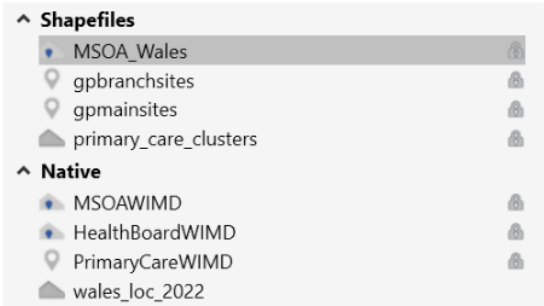


Figure 11 - Boundary Data

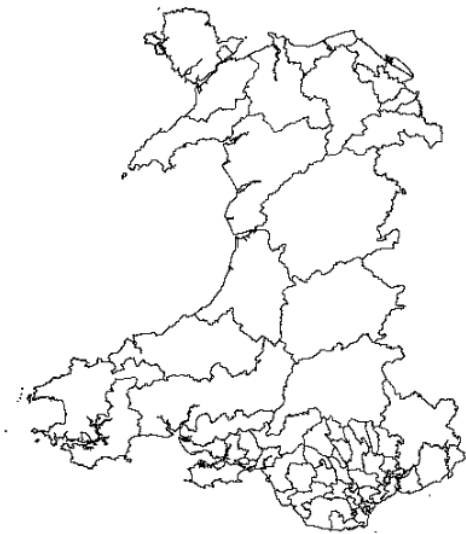


Figure 10 - Primary Care Boundaries

3.2 Plotting Data

To plot data to use either for point data or thematic data, data gathered previously also be imported. By selecting on Table>Open>Table each one of the data sources is selected.

Depending on the data, the top row usually contains the titles for each field, especially when importing from Excel. The tick box is selected to ensure the correct data is selected.

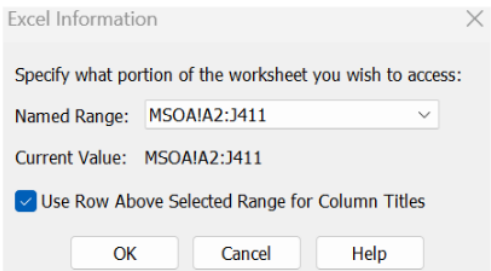


Figure 12 - Importing Data

The data that is imported must match the schema for MapInfo to plot the data. By ensuring that each field is a Float, as per below, MapInfo can plot the data and the map.

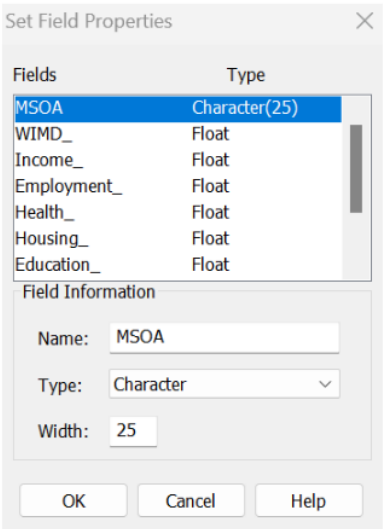


Figure 13 - Data Types

Once the data has been imported, the dataset must be geocoded to its corresponding boundary map. For example, MSOA Data will need to be geocoded to the MSOA boundary map.

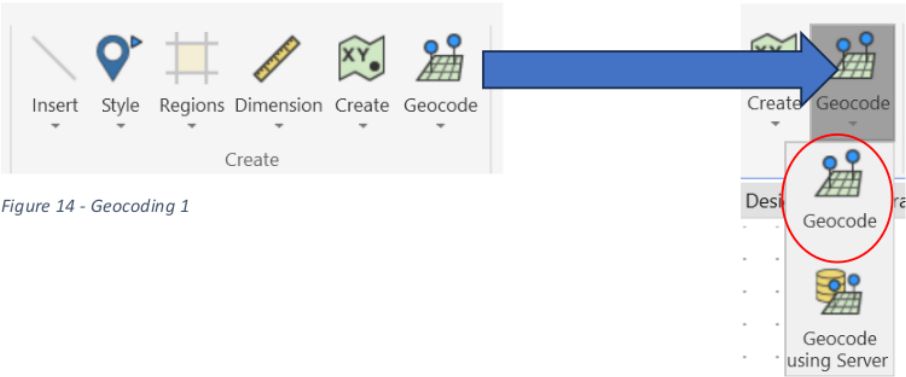


Figure 14 - Geocoding 1

In the figure below the data being imported, MSOAWIMD, must align to the MSOA boundary data. Ensuring the names of the corresponding columns align. For example, column 'MSOA' in 'MSOAWIMD' must match the values of 'name' in 'wales_msoa_2021'.

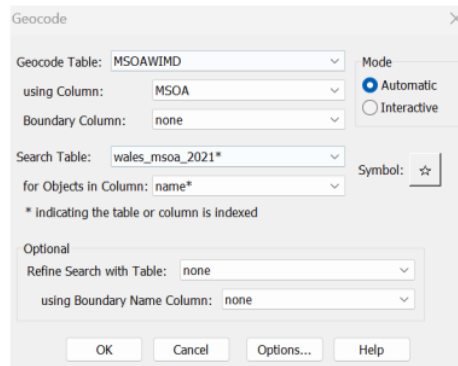


Figure 15 - Geocoding 2

Creating a subset of the data with only the data you need; SQL functionality can be used.

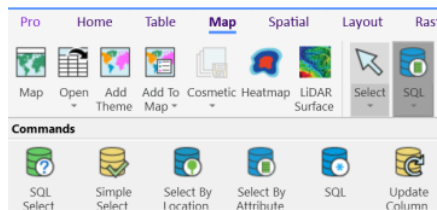


Figure 16 - SQL Selection

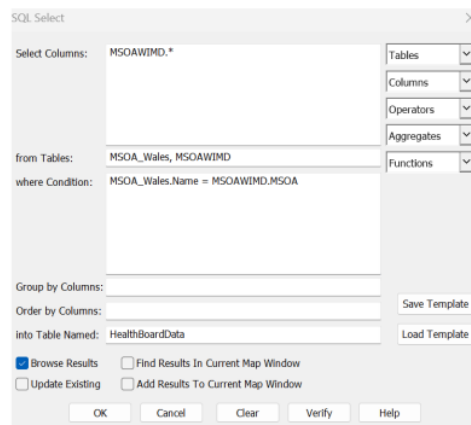


Figure 17 - SQL Creating Subset

The figures above show the creation of a new table called 'HealthBoardData'. This function joins the tables to only include matching results.

3.2.1 Thematic Mapping

Thematic Mapping can be used to highlight a certain boundary or to plot points within the boundary.

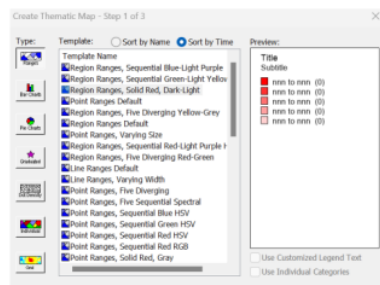


Figure 19 - Step 1 Thematic Mapping

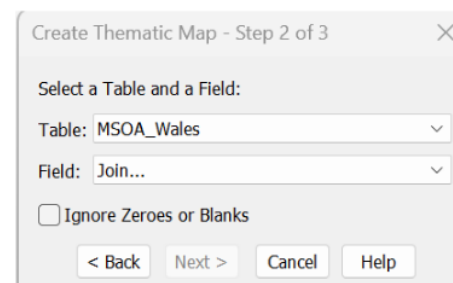


Figure 18 - Step 2 Thematic Mapping

A New temporary column is then added to the boundary data. This needs to match the base table column.

Update Column for Thematic

Table to Update: MSOA_Wales

Column to Update: Add New Temporary Column

Get Value From Table: MSOADData Join...

Calculate: Value

of: WIMD_

OK Cancel Clear Help

Figure 21 - Temporary Column

Specify Join

Join

☒ where Name from table MSOA_Wales

matches MSOA from table MSOADData

☐ where object from table MSOADData

is within object from table MSOA_Wales

OK Cancel Help

Figure 20 - Specify Join

Once these steps are applied the map will then highlight the corresponding boundaries with the correct data from the data tables.

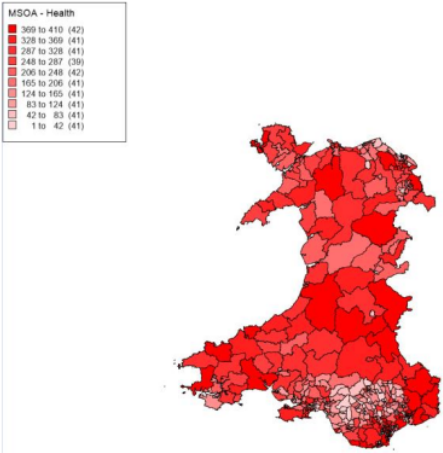


Figure 22 - Thematic Map

The same principle can be applied to points data. The same steps will be taken.

4. Analysis

4.1 Thematic Analysis

To analyse the data, a side-by-side comparison can be made to make assumptions around the data. The first set of maps show a breakdown of MSOA with the corresponding MSOA Data applied on top. The first image shows the overall WIMD index score, then the health index score and finally the Access to Services index score.

Showing the maps with this granularity highlights a key point. The areas highlighted show the area circled in black has a lighter shade of blue in the first map and lighter shade of green in the second map indicate a low WIMD and health score. However, the access to services shown in the third map shows a deeper red colour, representing a higher access to services. This goes against the hypothesis of higher deprivation indicates and lower access to services.

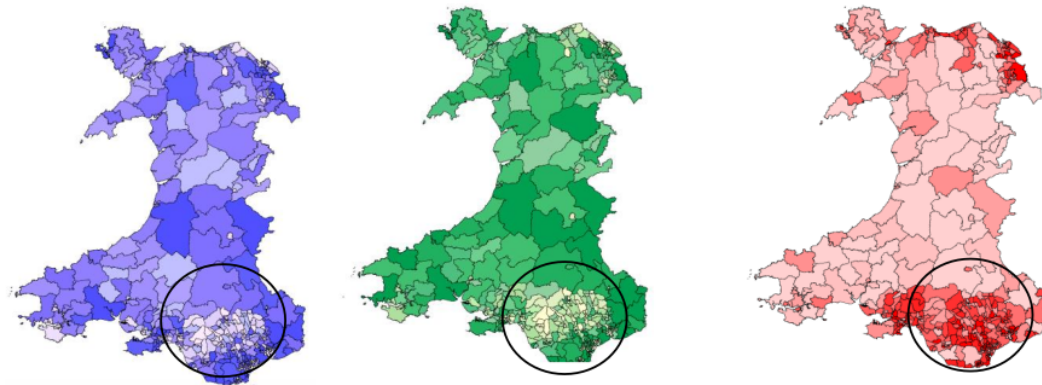


Figure 23 - MSOA

The next set of maps display the breakdown by health board. Unlike the MSOA maps above, each map shows a similar pattern across all three. The reinforces the hypothesis.

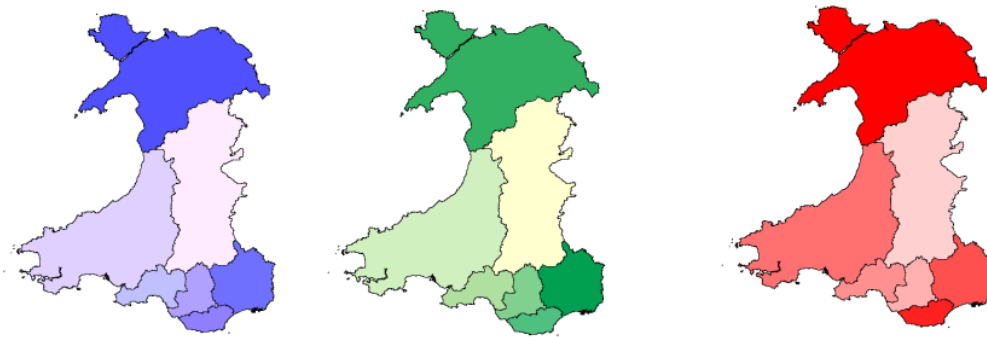


Figure 24 - Health Board

The third set of maps shows the breakdown by Primary Care Cluster. Again, like the MSOA maps there shows a discrepancy between each of the maps. The first, WIMD index score map, shows light colours within the centre, whereas the second and third show darker colouring. This proves the hypothesis to be false.

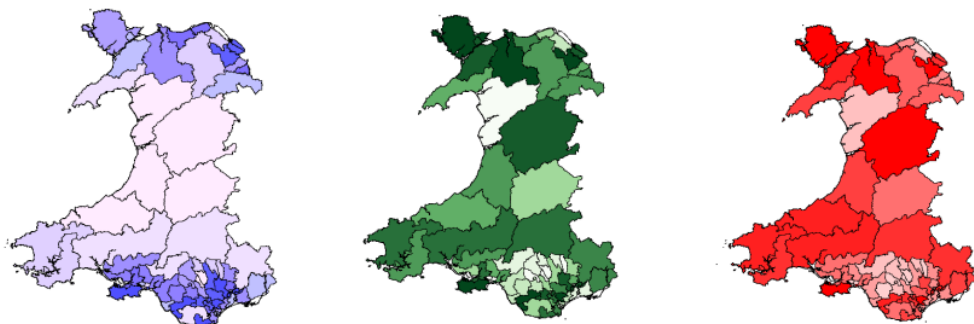


Figure 25 - Primary Care Clusters

4.2 Point Analysis

Further to the thematic mapping above, point data analysis is carried out. The below maps show the GP surgeries and the GP Branch sites mapped over the Access to services thematic. To complete the analysis a separation of the top cluster and bottom cluster within the Access to Services ranking is to be made and compare them side by side.

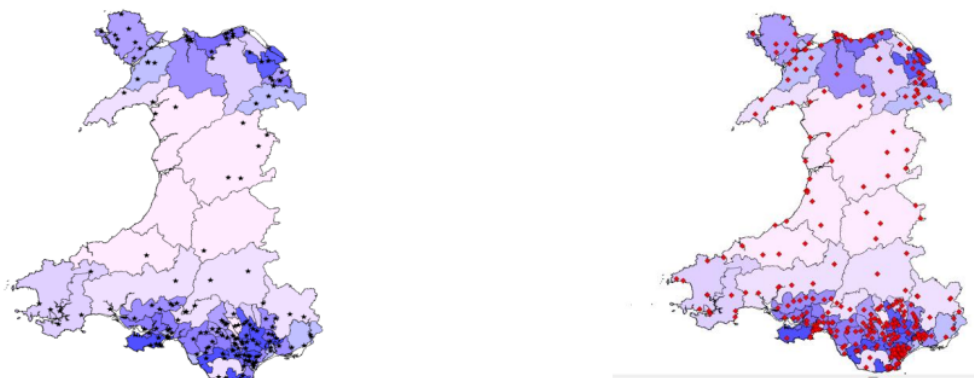


Figure 26 - GP Branches and Sites

To select the top and bottom cluster two SQL statements can be made. By limited the selection to 1 using the LIMIT function the order of the dataset is then determined by ascending and descending the data, as shown below.

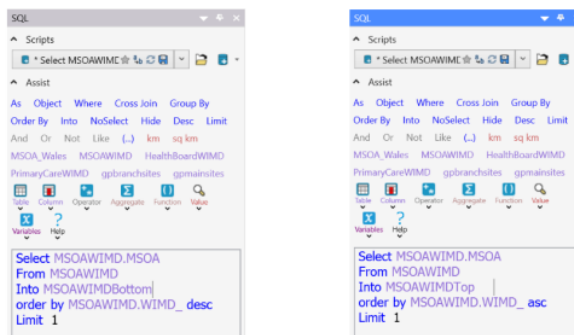


Figure 27 - Isolate Boundaries

By joining the data back into the maps will select the two polygons that are required..

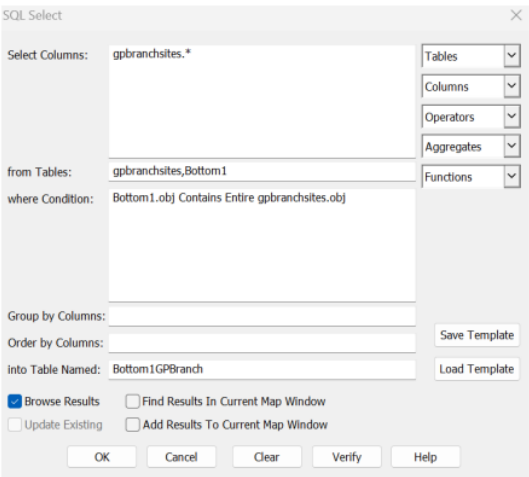


Figure 28 - Joining selections back into map.

Showing the scale of each map reflects the land mass between each boundary. Side-by-side there is a vast difference between each one.



Figure 29 - Boundary Comparison

Scaling each vector to similar sizes an obvious difference can be seen.

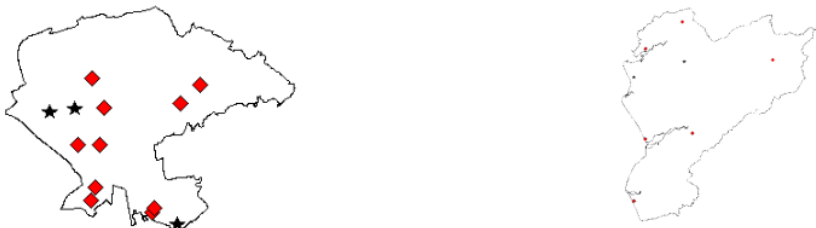


Figure 30 - Comparison

There are clearly more GP practices and branches in the higher WIMD scoring boundary, thus proving that the hypothesis can be true.

5. Conclusion

5.1 Further Steps

There is much more than can be achieved by expanding on the analysis above. By adding in further data can also help to prove or disprove the hypothesis. By including major road networks and transport links can enhance the detail in the research but also create a better understanding of the areas involved.

5.2 Final Conclusion

As can be seen from the analysis above it is possible to determine whether an area with high deprivation can result in a lack of services. However, it cannot be proven against some types of data. By understanding the landscape by using the mapping capabilities of this kind of software will inform users to make decisions that can benefit the population.

References

- [1] R. G. W. a. K. E. Pickett, "The problems of relative deprivation: Why some societies do better than others," *Social Science & Medicine*, vol. 65, pp. 1965-1978, 2007.
- [2] I. W. a. S. Gurtner, "Understanding shortages of sufficient health care in rural areas," *Health Policy*, vol. 118, 2014.
- [3] R. A. M. C. D. L. Peter A. Burrough, *Principles of Geographical Information Systems*, OUP Oxford, 2015.
- [4] R. L. Church, "Geographical information systems and location science," *Computers & Operations Research*, vol. 29, pp. 541-562, 2002.
- [5] W. contributors, "MapInfo Corporation --- {Wikipedia}{,} The Free Encyclopedia," Wikipedia, 2022. [Online]. Available: https://en.wikipedia.org/w/index.php?title=MapInfo_Corporation&oldid=1103208090. [Accessed 01 2024].
- [6] W. a. M. M. osgen, "Census data mining—an application," in *Proceedings of the 6th European conference on principles of data mining and knowledge discovery*, 2002, pp. 65--79.
- [7] O. f. N. Statistics, "Census 2021 geographies," Office for National Statistics, 2024. [Online]. Available: <https://www.ons.gov.uk/methodology/geography/ukgeographies/censusgeographies/census2021geographies#middle-layer-super-output-areas-msoas>. [Accessed 01 2024].
- [8] U. D. Service, "Features in MapInfo TAB format," Uk Data Service, 2024. [Online]. Available: <https://statistics.ukdataservice.ac.uk/dataset/2011-census-geography-boundaries-middle-layer-super-output-areas-and-intermediate-zones/resource/839b6031-8f3f-4988-9753-65dae72769a9>. [Accessed 01 2024].
- [9] U. D. Service, "Boundary Data Selector," Uk Data Service, 2024. [Online]. Available: <https://borders.ukdataservice.ac.uk/bds.html>. [Accessed 11 2023].

- [1 W. Government, "GP surgeries: Branch Sites (Open Government Licensed Version)," Data Map
0] Wales, 2023. [Online]. Available:
https://datamap.gov.wales/layers/geonode:gpbranchsites_ogl. [Accessed 11 2023].
- [1 Infuse, "InFuse 2011 Population Statistics," Infuse, 2023. [Online]. Available:
1] <https://infuse2011gf.ukdataservice.ac.uk/infusewizgeo.aspx>. [Accessed 11 2023].
- [1 W. Government, "Welsh Index of Multiple Deprivation," Welsh Government, 2022. [Online].
2] Available: <https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation>. [Accessed 11 2023].

FINAL GRADE

74/100

GENERAL COMMENTS

Clear statement of hypothesis followed by a good factual abstract introduction.

Excellent knowledge and understanding demonstrated with reference to deprivation and its impact on the health of the nation and a good overview of the geographical landscape of Wales. I got understanding of file types and data cleansing required for GIS has been demonstrated.

A good comparison of Thematic maps used within your analysis where you had side-by-side visuals however, no legend was provided or key for the map data therefore it was difficult to interpret the colour schemes. With such a lot of great data captured I do feel that your report could have been more visual perhaps more time spent on the analysis and little less on how you obtained the data and how mapping for worked. You provided a good point for reflection on land mass sizes but again however some more imagery with colour codes would have helped in your explanation.

A well written report throughout with some good conclusions and references to future work.

Good participation in Group work assessment with good understanding of the data capture requirements and the data available for such research.

20/25 18/25 29/40 7/10 total 74%

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18
