DATA ANALYTICS REPORT

Module: The Data Professional

An overview of citizens' issues and views on the Welsh public transport system

Student number: ra22895

Contents

Introduction

According to the Welsh Government (2019) "Everyone should have the same freedom, dignity, choice and control, at home, work, in education and in the community". Having access to local amenities is a vital part of this and transport links are integral to this vision. Public transport in Wales is a complex picture due to both geography and the population spread. Wales is a small country with an area of 20779 km^2 and an estimated current population of 3267500 (2022). Almost half of the population live in South-East Wales (mainly Cardiff) which makes up just 14% of the land area. Most of mid-Wales, South-West Wales and approx. half of North Wales is classed as rural with agriculture being the primary land use (78% of land). Much of North and Mid-Wales is mountainous with transport links being predominantly East-West allowing access to ports and leisure facilities along the coastline.

According to Sustrans (2022) between 30% and 60% of the Welsh population currently face transport poverty – defined as spending 10% or more of household income on running a car. They also highlight the fact that the 23% of the population who do not have a car have faced rail, coach and bus ticket prices increasing between 34% and 74% in the last 10 years (Sustrans, 2022).

Rural communities are particularly affected by a lack of public transport and the services that do run may not meet the needs of the population. Services that don't run in the evenings or weekends mean that shift workers are left with few choices. The disabled are also disproportionately affected by a lack of a regular and reliable service (Equality and Human Rights Commission, 2020).

One measure used to assess access to services is the Welsh Index of Multiple Deprivation (WIMD). This is "designed to identify those small areas of Wales that are most deprived" (Welsh Government, 2022). In the case of the map below the services considered were:

- Pharmacy
- Food shop
- General Practitioner
- Post Office
- Primary School
- Public Library
- Sports Facility
- Secondary School
- Broadband availability
- Petrol Station

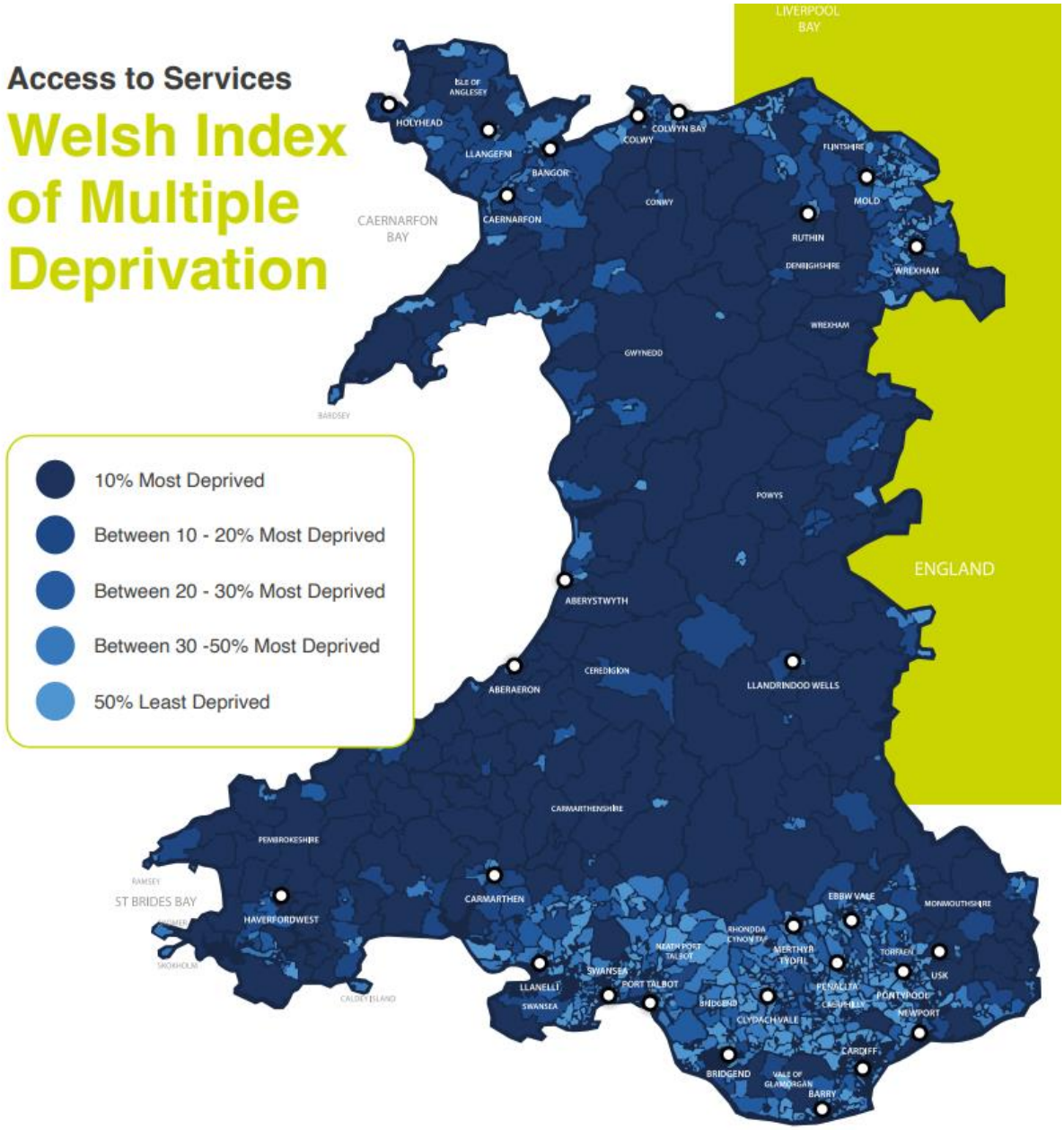
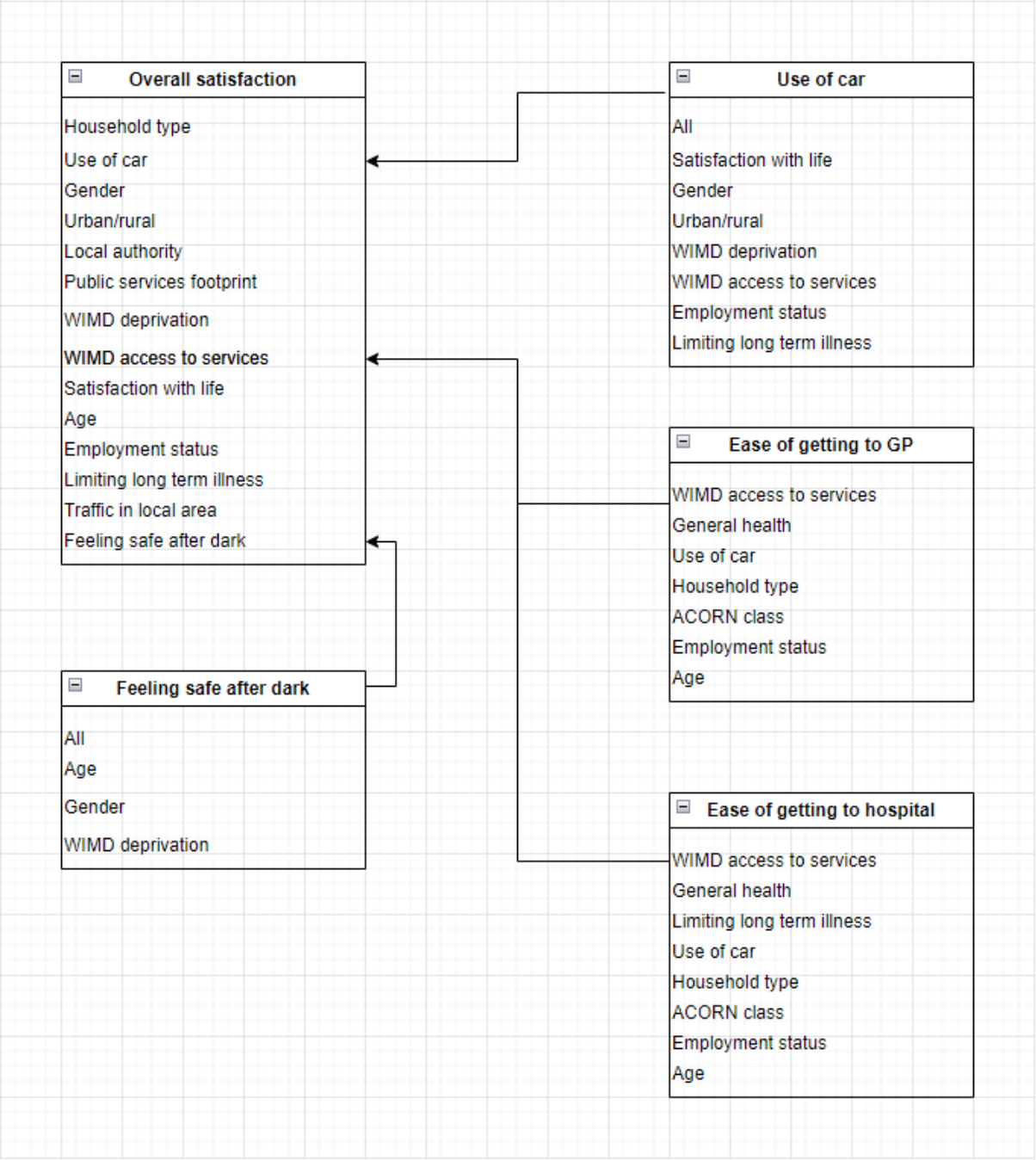


Figure 1: Access to services, Welsh Index of Multiple Deprivation (Sustrans, 2022)

The dataset

The National Survey for Wales, 2013-2014 – Transport surveyed people for their opinions on various aspects of travel in Wales. The survey results were grouped into five main areas: Overall satisfaction with the state of the transport system in Wales; Having the use of a car; Ease of getting to and from a GP surgery; Ease of getting to and from a hospital; and Feeling of safety travelling by public transport after dark. Each of these areas was broken down by different population classifications such as age, gender, employment status etc.

Along with WIMD, the ACORN classification was also used. This groups the population by various markers of wealth.



There are some limitations with the data. Only the results of the survey are available, with access to the questionnaire, quality report, technical report and publications is not currently possible. Without the raw data being available it must be assumed that there were no missed opportunities to explore trends in the raw data.

Sample sizes have been rounded to the nearest 100 and, where multiple questions have been used to create a single table, a representative sample size has been given. As it is not indicated which tables have the representative sample size this means that the true sample sizes for any question are not known. This may cause a slight bias in the results due to the possible effect of the rounding of the sample size.

Tables 1 – 14 and 16 have the results as a number but no unit or scale is given. It seems likely that this is a scale of 1 – 10 but this is supposition. It is important to establish the scale used as a different scale could place more significance on small differences in values

Twenty-three of the tables have values that are not precise estimates, with nine of the tables having at least one value that is not reliable – having a coefficient of variation that is greater than or equal to 20. The coefficient of variation (CV) is the standard deviation divided by the mean of the sample so is an indicator of the measure of dispersion of the sample. Hence, caution must be used in drawing inferences from these values.

In several tables the stated mean value is different to the calculated mean value. This appears to be due to rounding errors in the preparation of the tables, but this may affect the overall result by a small margin.

There appears to be some inconsistencies in the way people are grouped for some measures – when looking at households the term pensioners is used but when age groups are used as a measure the people of pensionable age fall into two separate groups (65 – 74 and 75+). This may well have an impact on the results as people in the 65-74 age group are likely to be fitter and healthier than the older group so may have very different transport requirements. In the same way, the term 'limiting long term illness' is not defined. There is no indication if this is something that may need more frequent minor medical care or less frequent urgent care. These two scenarios will have a very different impact on a patient's transport requirements.

Methodology

Methodology is defined as "a system of methods used in a particular area of study or activity" (Oxford languages, 2010). This report used secondary survey data that was supplied as pre-prepared tables using MS Excel. The survey data used coefficient of variation (CV) to ensure robustness and areas where there was doubt about precision were highlighted. A 95% confidence interval was used for statistical analysis and sample sizes were rounded to the nearest 100 where possible.

In order to leverage insights from this data MS Power BI will be used. This is the method of choice due to familiarity with and access to the package.

The data analysis will focus on the areas directly relating to the transport system by comparing overall satisfaction with public transport across different demographics (Tables 1 – 14). In addition, the analysis will look for trends in ease of access to essential services such as the GP or hospital (Tables 23 – 38).

Data representation

According to Kirk (2019) data visualisation is "The visual representation and presentation of data to facilitate understanding". Not only must the person creating the report understand the data, but it must clearly convey insights to the user, or it has not achieved its aim. There are a wide number of both static and dynamic chart types available to aid with data visualisation, with many being built into common packages such as MS Excel as well as specific libraries for use with, for example, Python. Care must be taken to choose a chart type that is easy to understand so knowledge of the target audience is essential.

To represent the data in this report I propose to create a dynamic dashboard using bar charts that can be sliced by different attributes. The dashboard will have three charts, Overall satisfaction, Ease of getting to and from GP surgery and Ease of getting to and from hospital.

The y-axis will be either percentage satisfaction or the numerical values quoted in the tables. The x-axes will have the attribute that the data was sliced by. Slicers, linked to all the charts, will be used so selecting an attribute will show relevant values on each chart. This will allow the comparison of different survey areas for the same attribute from a single click. By making it a dynamic dashboard the amount of data viewed at any one time can be easily controlled and viewers will not be faced with multiple static charts to pick through.

References

Equality and Human Rights Commission. (2020) Accessible public transport for older and disabled people in Wales. Available from: https://www.equalityhumanrights.com/sites/default/files/accessible-publictransport-for-older-and-disabled-people-in-wales [Accessed 13 November 2022]

Kirk, A. (2019) *Data visualisation; a handbook for data driven design*. SAGE

Oxford languages (2010) *Oxford dictionary of English*. Oxford: Oxford University Press

Sustrans (2022) Making the connection: Why Wales must act now to tackle transport poverty and ensure access for everyone. Available from: https://www.sustrans.org.uk/media/10425/transportpovertypaper-sustrans_eng.pdf [Accessed 13th November 2022]

Welsh Government. (2019) Action on Disability: The Right to Independent Living. Available from: https://gov.wales/sites/default/files/publications/2019-09/action-on-disabilitythe-right-to-independent-living-framework-and-action-plan [Accessed 13 November 2022]

Welsh Government. (2022) Welsh Index of Multiple Deprivation. Available from: https://gov.wales/welsh-index-multiple-deprivation [Accessed 13th November 2022]