The Data Professional, Assignment 4 - End of module assignment

As discussed recently, Transport for Wales (TfW) is the public body, responsible for transport and travelling in Wales. This includes personal and public methods of transport such as driving, walking, taking the bus, etc. TfW aim to achieve high-quality, safe, integrated, affordable, and accessible travel for everyone. Infrastructure is one of the main pillars for economic growth, mainly by providing job mobility. Especially with the investment of £800 into a new fleet of trains in the North of Wales (Transport for Wales 2023). In addition, focus for TfW has also been towards providing environmental benefits and encouraging more sustainable methods of transportation. For example, walking and biking schemes to reduce congestion and heavy traffic, especially in urban areas of Wales. However, economic benefits and environmental benefits usually act as a trade-off, therefore making it difficult to achieve both. The most recent challenge for TfW has been the impact of COVID-19, which has resulted in a fall in demand for public transport, due to fears of safety from passengers, particularly in the older generation who are more vulnerable.

As discussed in my Data Analytics Report, we'll be using the 42 statistical tables provided by The National Survey for Wales (2014), which show confidence intervals and coefficients of variations for their results. I have extracted the data I needed and saved it into individual comma separated value (CSV) files, by using Microsoft Excel. Manipulation of the original tables occurred by merging tables of a relationship, and sorting by highest to lowest, and vice versa. Some results were merged for 'Very' and 'Fairly'. For example, 'Very Safe' and 'Fairly Safe' were merged to get a result for 'Safe'. After the data tables were manipulated, the CSV files were then uploaded to my GitHub repository, github.com/cfinal15/cfinal15.github.io, making them accessible for anyone to view. For my visualisations, I used a combination of different python libraries to help with my coding. These libraries are NumPy, Pandas and Plotly. All libraries were downloaded through my local machines pip. The coding then took place in the web-based computing platform Jupyter Notebook. With Plotly, I will be using its functions of Bar, Line, and Pie for creating bar and pie charts, and line graphs. The Jupyter Notebook file, that will be attached to this assignment, shows the code I executed to produce the graphs within this report.

From my previous assignment, Data Analytics Report, I outlined the data analysis methods that I would be using:

- Maps for comparing deprivation scores and seeing if there was any correlation with other variables, such as satisfaction scores.
- Dummy variables such urban or rural, to see if comparing variables will provide reasoning for difference in feeling safe whilst using the Welsh Transport System.

I will therefore be looking to prove the points from above by a series of bar and pie charts, and line graphs. From figure 1 we can see a bar chart shows the mean scores of 'Satisfaction with State of Transport System' by area. The graph shows the areas that are above and below the average, which is labelled 'Total' on the x-axis.

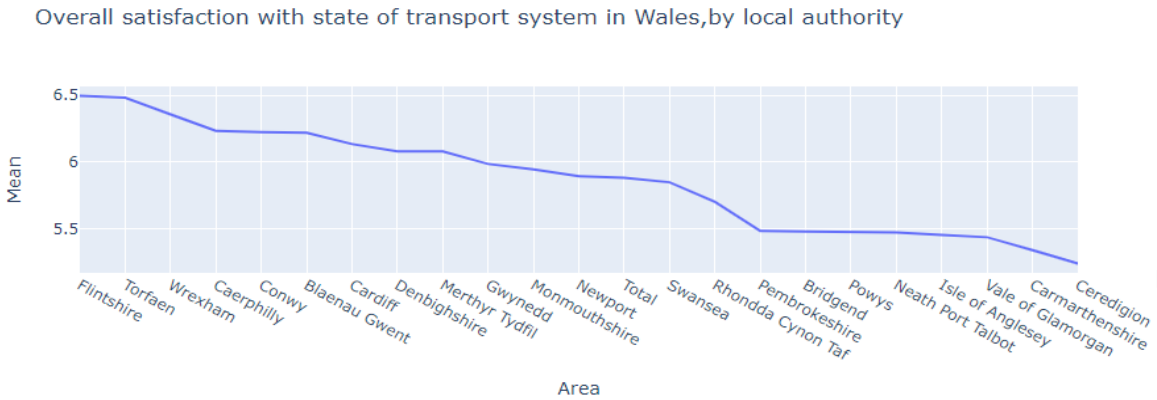

Figure 1

The average mean score for 'Satisfaction with State of Transport System' was 5.9 (1 d.p). The areas of above that average was Flintshire, Torfaen, Wrexham, Caerphilly, Conwy, Blaenau Gwent, Cardiff, Denbighshire, Merthyr Tydfil, Gwynedd, Monmouthshire, and Newport. The areas below average were Swansea, Rhondda Cyon Taf, Pembrokeshire, Bridgend, Powys, Neath Port Talbot, Isle of Anglesey, Vale of Glamorgan, Carmarthenshire, and Ceredigion. From the map provided by Mappr (2023), figure 2 shows all the districts of Wales, and the Southwest of Wales is where satisfaction levels with transport was the lowest. The regions of Swansea, Carmarthenshire, Pembrokeshire, and Ceredigion, all being below average and part of the Southwest of Wales.

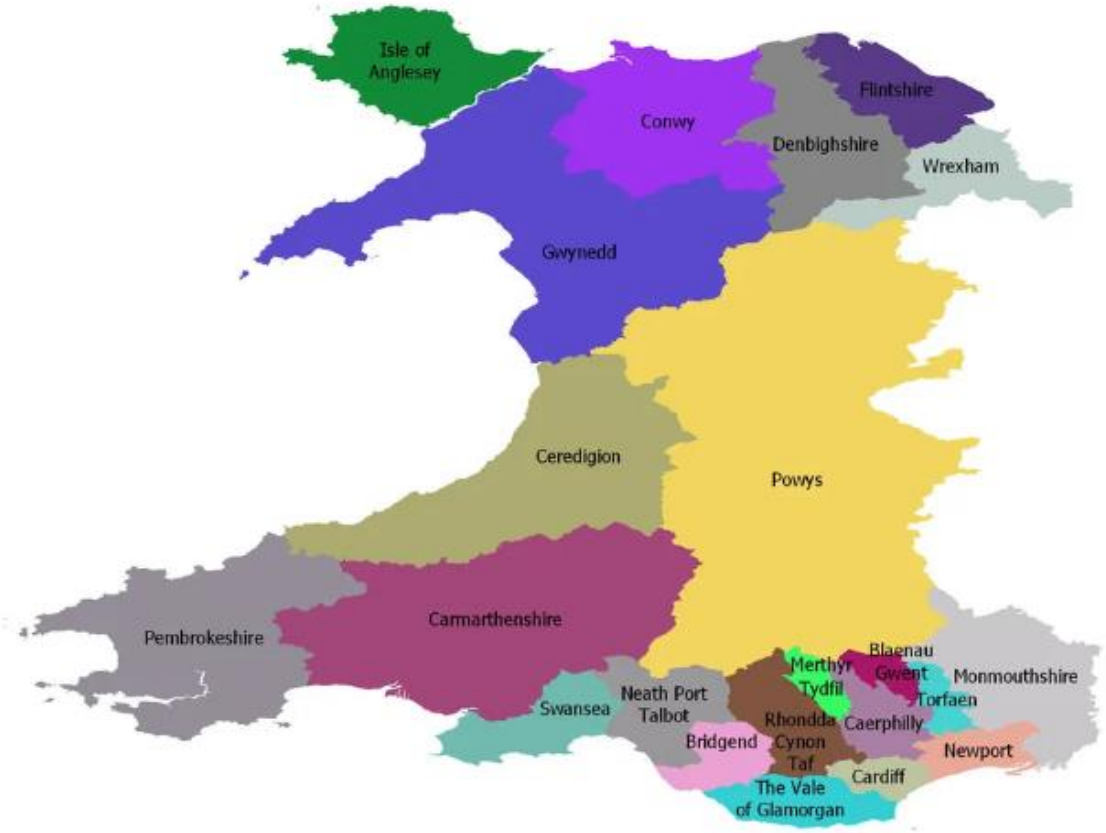

Figure 2

In the statistical tables reported by Transport for Wales, IMD scores are used. Indices of Multiple Deprivation (IMD) is a measure for evaluating an areas deprivation, taking into account variables such as income, employment, education, health, and crime. When we compare the results in figure 1 with StatsWales (2019) and The Welsh Index of Multiple Deprivation in Figure 3, we can see those areas of high deprivation, like the Southwest/South Wales, have lower satisfaction with the transport system in Wales. Showing a correlation between deprivation and satisfaction with transport.

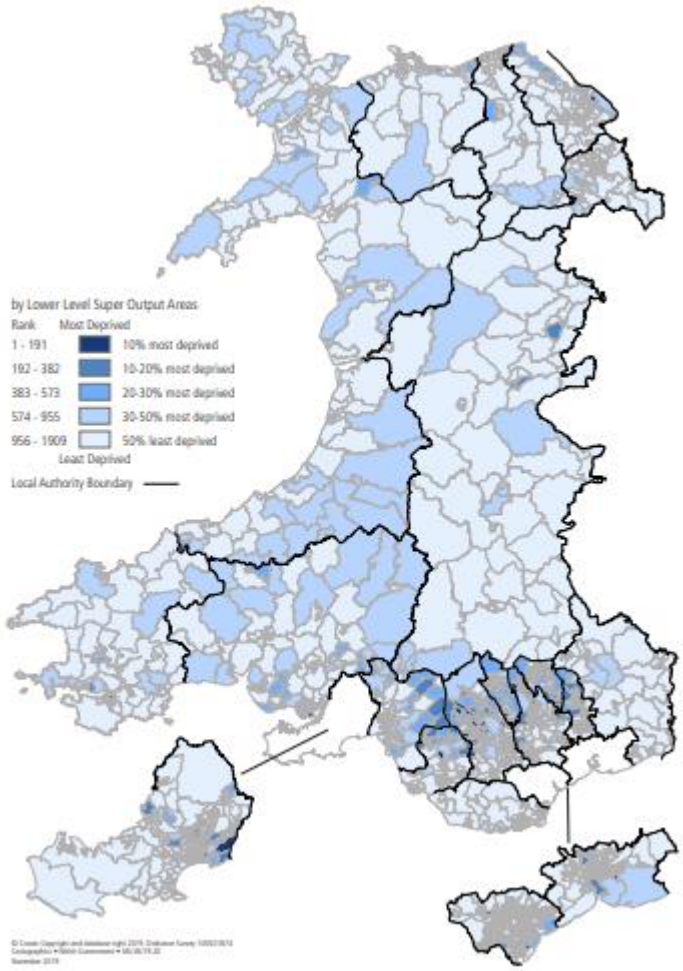

Figure 3

I continued my investigation into IMD scores to see if there were patterns amongst the statistical tables provided by TfW. In figure 4, I modelled deprivation scores on the x-axis, against the percentage score for how safe respondents felt whilst using transport in the dark. We can see a correlation with more deprived areas not feeling as safe travelling in the dark, as those in less deprived areas. A similar correlation as seen in figure 1, where more deprived areas were less satisfied with the transport system in Wales.

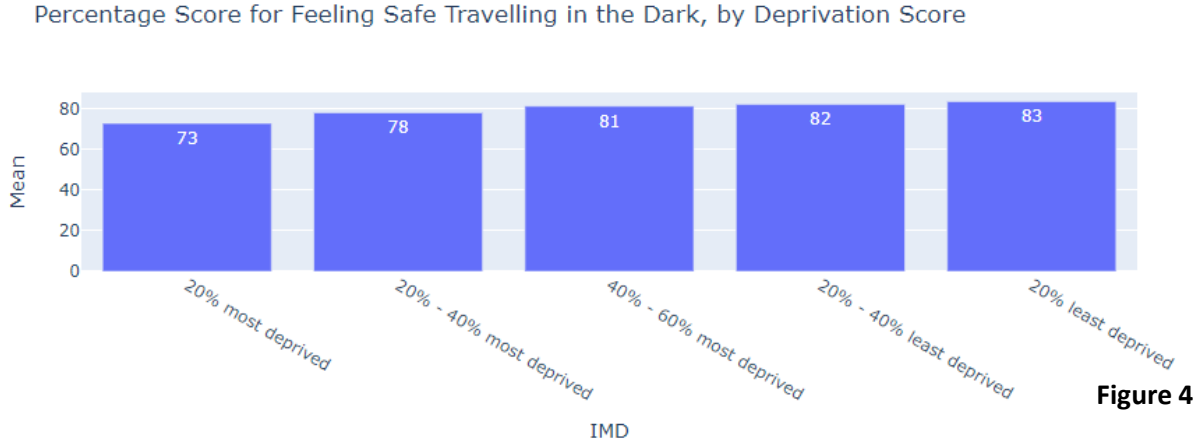

Figure 4

A causality for their being a correlation between deprivation and safety whilst travelling in the dark/satisfaction with transport system, could be down to whether individuals have access to a car. The statistical tables and figure 5, show that 78.9% of survey respondents have access to a car. Using the table that shows the access to a car, by deprivation score we can see a divergence along the deprivation scale. This is illustrated by figure 6, as the least deprived an individual becomes, the more likely they are to have access to a car. The pattern in divergence is similar to that seen in figure 1 and figure 4. Having access to a car provides extra security when travelling in the dark, instead of reliance on public transport like trains and buses, which explains why higher deprivation areas didn't feel as safe compared to less deprived areas. Access to a personal car also brings reliability and time advantage compared to other transportation methods, giving another explanation as to why less deprived areas are more satisfied with their transport system, as explained by figure 1.

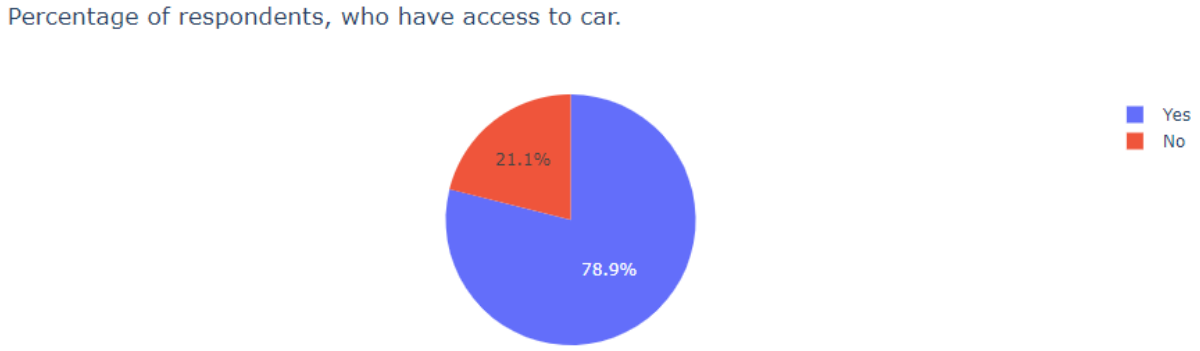

Figure 5

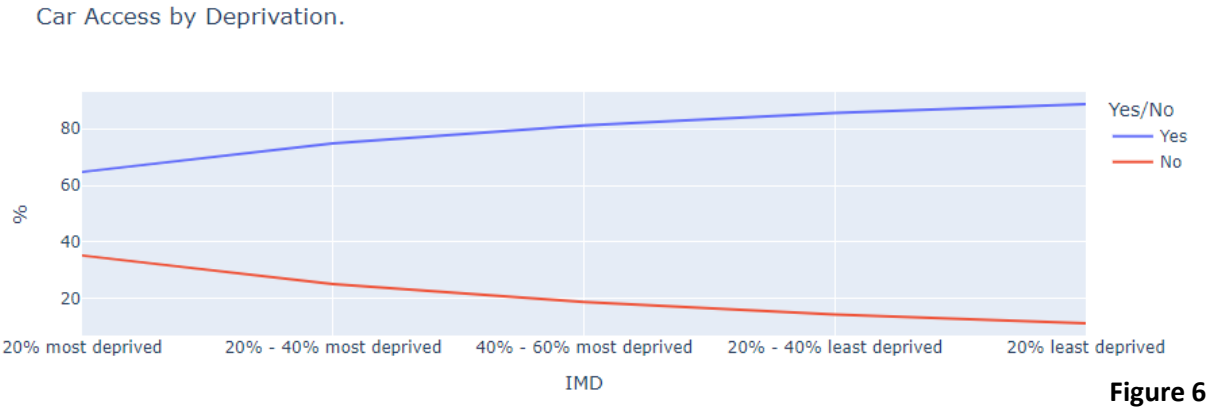

Figure 6

Based on the evidence above, TfW focus should be towards improving the transport system that is used predominantly by those who don't have access to a car. Despite the statistical tables being from 2013/14, Tfw has invested heavily into public transport recently. In 2019, £800 million was invested into a new fleet of trains for TfW named 'Happy Valley', connecting the North of Wales with Liverpool (Transport for Wales 2023). However, from the evidence above, to see improvements in satisfaction in the transport system, focus should be towards regions in the South and Southwest. As identified in figure 1, areas like Carmarthenshire and Ceredigion have the lowest satisfaction with the transport system. Furthermore, transport system investment in the South and Southwest, should go towards improving safety as well. More research is needed to understand the specific needs into why deprived areas, as shown in figure 4, don't feel safe when travelling in the dark, compared to less deprived areas. London Travel Watch (2022) conducted research and found that "frequent and reliable public transport is critical if people are to feel safe when travelling, particularly after dark", which gives added to proof to our earlier evidence of why satisfaction in transport and travelling in

the dark had similar correlations. More investment into streetlights was also a big contributor to making transport users feel safer in London, especially amongst women.

Also discussed in my previous Data Analytics Report, I mentioned about modelling the dummy variable for urban/rural for showing comparisons between variables. From figure 7, you can see the satisfaction scores for the transport system, between urban and rural regions in Wales. There is a slight difference in satisfaction between the two groups, with people in Urban Areas being more satisfied with Transport.

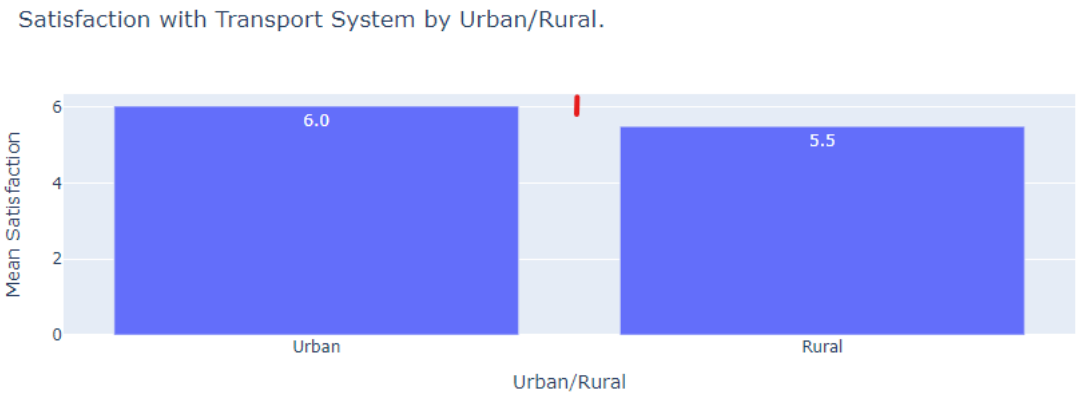

Figure 7

I then investigated the summary data for ease of access getting to your GP and Hospital based off general health. This is illustrated in figure 8, where it shows that health plays a big role on whether you can access services easily. However, you are at a greater disadvantage if you are to be living in a Rural location and needing hospital attention. Those in bad or very bad health find it 10-15% more difficult to get transport to hospital treatment over GP treatment. This is reflected in figure 7, with rural residents not being as satisfied with the transport system, compared with users in urban areas, who are likely to have better infrastructure to get to hospital more accessibly.

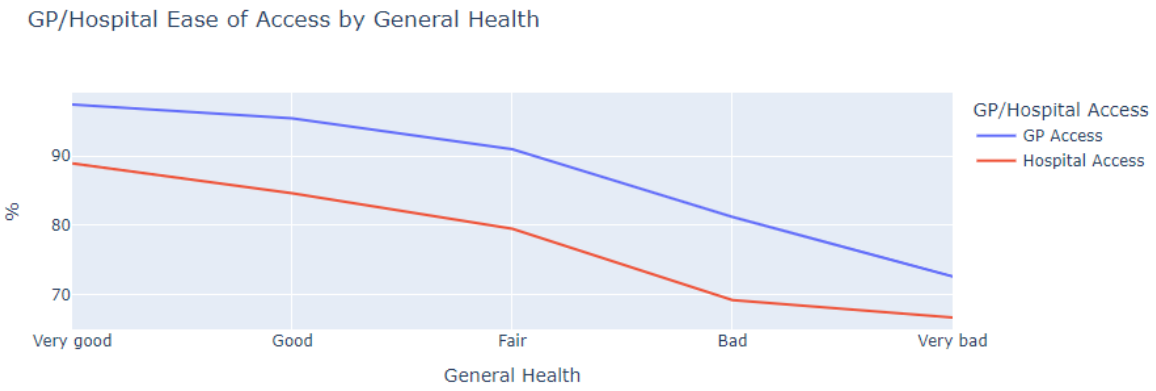

Figure 8

With the evidence from figures 7 and 8, Transport for Wales needs to focus on rural areas of Wales, in making sure that everyone has access to the same facilities, without the transport system acting as a barrier for gaining medical attention. This will help to alleviate deprivation scores in Wales, with Health being a main contributor to the Indices of Multiple Deprivation. Consequently, the satisfaction in the transport system will improve for rural areas, by preventing a skew in results by those whose general health is considered ‘Bad’ or ‘Very Bad’. I believe this ties in with my previous evidence with

shows deprivation impacts satisfaction with the transport system, and that deprived areas are usually rural areas often in the South/Southwest of Wales.

To conclude, having access to a car was a causality for deprived areas being less satisfied with the transport system, and not feeling as safe whilst travelling in dark times. With deprivation being more concentrated in the South/Southwest of Wales, investment should be directed more into these areas to improve satisfaction and safety, especially in transport that excludes personal car. I also investigated the skewness in data for the ease of access to GPs/hospitals, based of respondent's health. Rural communities were struggling with accessing hospitals considerably over urban communities, especially if your health was 'Bad' or 'Very Bad'. This linked back into satisfaction with the transport system in Wales, with rural areas feeling more 'left out', especially the South/Southwest of Wales. Making public transport more reliable so that it's easier to access hospitals for rural patients, will not only boost satisfaction with the Welsh transport system, but also reduce deprivation by improving the health outcomes of individuals. A big limitation of implementing such solutions is that you won't be able to see benefits in the short-term. Adaptations to infrastructure take time and the impacts will be seen in the long-term. The building of the 'Happy Valley' fleet in 2019, is an example of where the North regions of Wales, won't see the economic benefit of the infrastructure, until 5 to 10 years after it was completed.

If I was to improve my analysis, having access to the raw dataset would provide more insight, than only be restricted to the analysing summary tables provided by TfW. One advantage would be that the survey results could be used on a heat map, to shows alterations between the districts of Wales, perhaps by using software packages such as ArcGIS. The raw dataset will provide the opportunity to create custom regressions and histograms which will help identify more correlations and casualties amongst different variables. The statistical summary tables limited the ability to compare variables. It would've been good to explore in more depth how different age groups found transport in Wales to be. As the summary tables showed coefficient of variation (CV) to be lower for the younger age groups such as, 16–24-year-olds. As shown by figure 9 below, this would suggest a bigger sample size of younger people is needed to get a higher CV, and therefore more reliable results. The implementation of incentives, such as vouchers could be used to increase the survey size of the younger generation.

Ease of getting to and from GP surgery [(a)]

	Very easy			Fairly easy			Fairly difficult			Very difficult		
	%	Lower CI	Upper CI	%	Lower CI	Upper CI	%	Lower CI	Upper CI	%	Lower CI	Upper CI
16-24	72	67	76	25	20	29	3	2	5	-	-	-
25-44	73	71	75	23	21	25	3	2	4	1	1	1
45-64	68	66	70	27	25	29	4	3	4	1	1	2
65-74	62	59	64	31	28	33	5	4	6	3	2	3
75 and over	46	43	49	36	34	39	12	10	14	6	4	7
Total	66	65	67	27	26	28	5	4	5	2	2	2

e' was excluded from all Transport tables recording "Ease of getting to and from the GP surgery".
similar analysis for the StatsWales Health tables, resulting in a small difference between the table totals.

team

Key:

	Estimate is precise
	Estimate is reasonably precise
	Estimate is considered acceptable
	Estimate is not reliable
-	Value is suppressed due to small cell size

Figure 9

References

StatsWales (2019) Welsh Index of Multiple Deprivation, Available at: https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation/WIMD-maps-2019 [Accessed: 23rd July 2023]

London Travel Watch (2022) Personal Security on London's Transport Network, Available at: https://www.londontravelwatch.org.uk/news/new-research-highlights-the-importance-of-maintaining-frequent-night-time-tubes-trains-and-buses-as-nearly-half-of-women-say-they-avoid-travelling-at-certain-times/ [Accessed: 23rd July 2023]

Maprr (2023) Available at: https://www.mappr.co/counties/wales/ [Accessed: 23rd July 2023]

Transport for Wales (2023) Homepage. Available at: https://tfw.wales/about-us [Accessed: 23rd July 2023]