

**Swinburne University Of Technology**

# *School of Science, Computing, and Engineering Technologies*

**ASSIGNMENT AND PROJECT COVER SHEET**

Unit Code: COS30045 Unit Title: DATA VISUALISATION

Assignment number and title: VISUALISATION ASSIGNMENT Due date:  **20 September 2024**

Lab Group: Tutor: Lecturer: FAIZAL ALIAS

Family name: KING Identity no: J23039507

Other names: JOSHUA SANJAY

## To be completed if this is an INDIVIDUAL ASSIGNMENT

I declare that this assignment is my individual work. I have not worked collaboratively nor have I copied from any other student’s work or from any other source except where due acknowledgment is made explicitly in the text, nor has any part been written for me by another person.

Signature: JSK

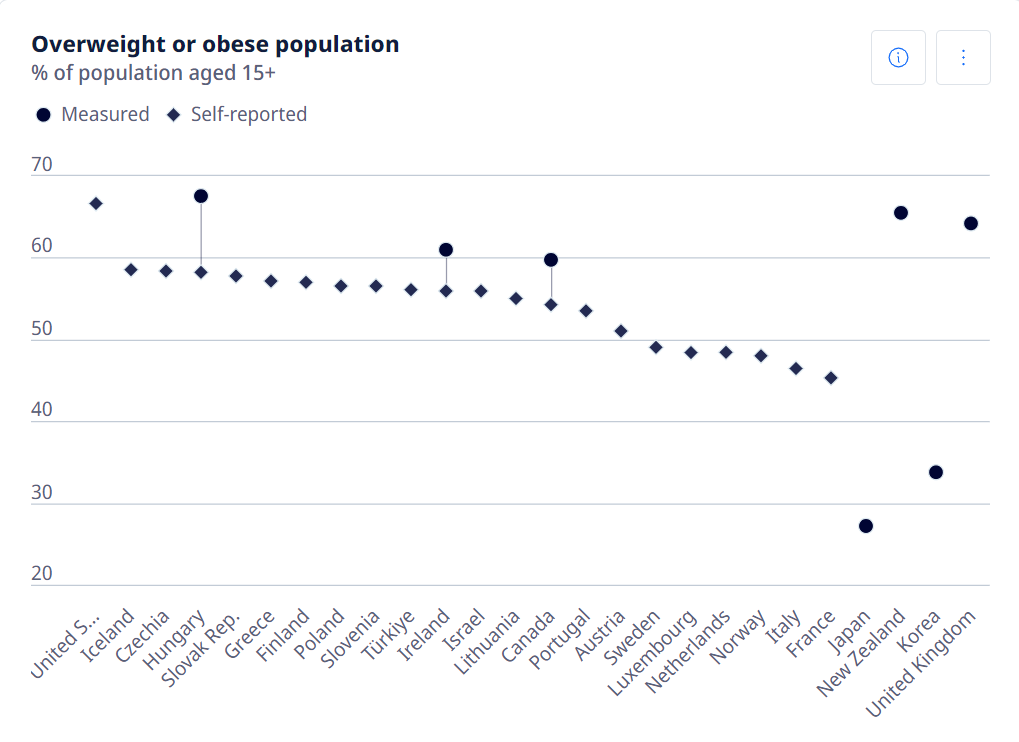
**Introduction**

OECD is an international organisation that collects data with the purpose of improving the welfare of people across the globe. The data collected covers important aspects of raising the international standards for the healthcare of individuals such as: health systems, health outcomes, health expenditure, and related health indicators across OECD countries.

Good visualisation is crucial for industries and policymakers as it represents complex data into simpler terms. An accurate presentation of data will assist stakeholders to make well-informed decisions and would also be able to identify inefficiencies found in the data presentation and allocate their resources accordingly where output would be least costly, whether in terms of production or ideology. Having data presented well can also enhance communication among a group as they would be able to understand better compared to hearing from someone or reading from a long and complex report.

The purpose of this report is to identify and analyse three data visualisations of OECD health statistics and critique the visualisations according to good data visualisation principles and guidelines.

**Visualisation 1 – Obesity or obese population rate**



Analysis: The scatter plot chart shows the trend of obesity rates across the OECD countries aged 15 and above, indicating United States with the highest obesity rate and Japan for the lowest. Tufte’s principle of graphical excellence is evident as the chart presents itself well as the title of the chart reads the main topic of the research with the appropriate sizing and the font boldness, and below provides the context of the research that shows that the research is aimed at individuals aged 15 and above; the legend is also provided as there are different shapes in the chart, circle, which is measured, and diamond, which is self-reported (Tufte,1983). Having a clear title, legend and labels of the countries aligns with the principles of having good data visualisation as it provides clarity to readers. The chart also uses a consistent scale with every country which aligns with the Few’s principles of data visualisation (Few,2012).

Critique: For comparing purposes, it is recommended to use a bar graph as Few mentioned that scatter plot charts are typically used to show trends and for this specific case study, we are observing the obesity rates across the OECD countries (Few,2012). A rotated bar-chart would provide a more straightforward comparison between measured and self-reported rates for each country. By switching to a bar chart, the countries would be listed clearly along the x-axis, and comparisons would be easier to grasp for the reader.

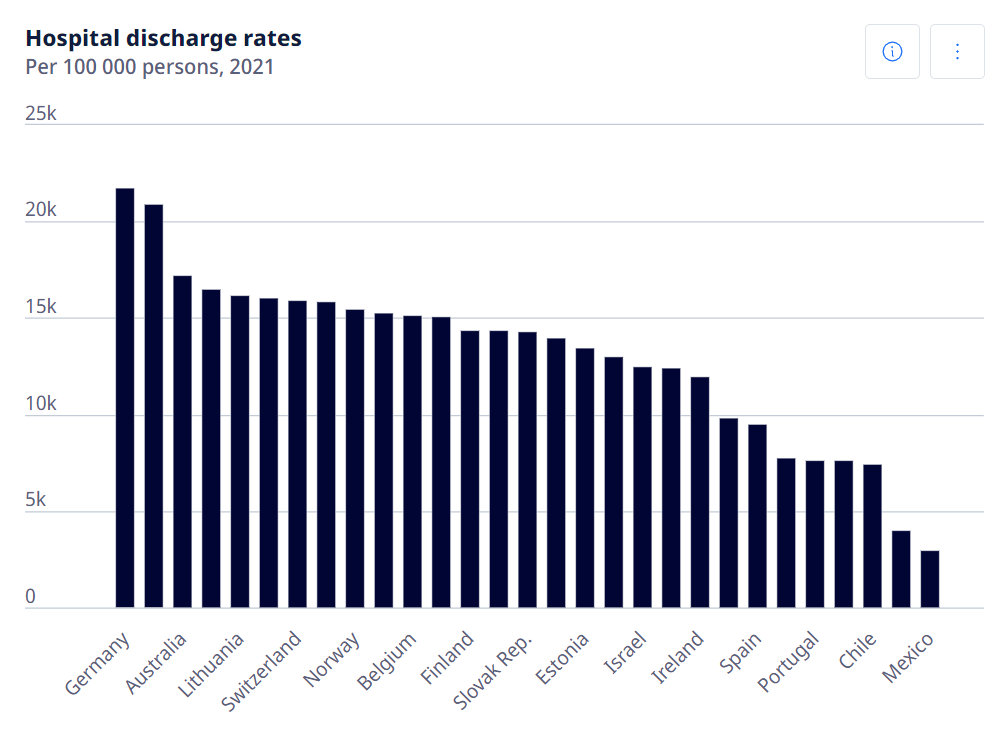
The current graph provided can be confusing to readers as there are self-reported statistics and measured in the same graph. According to Stephen Few’s principles on clarity in visualisations, readers should be able to interpret the data quickly and accurately. Readers would question the integrity of the data visualisation as they are unsure if the self-reported statistics is from a reliable source. Furthermore, as you observe the graph, you can see countries like Hungary have a different rate of self-reported statistic and measured rate. This difference of rates in the same country would easily confuse readers and not trust the data visualisation.

The x-axis is congested with OEBD countries, making it difficult for readers to locate data for each country. This violates Edward Tufte’s principle of maximizing data-to-ink ratio by overloading the visual with information.

The flaws listed goes against the principles of good data visualisation as data integrity is questionable, and the graph doesn’t provide simplicity, making it confusing for readers to understand the data.

Solution: Firstly, addressing the issue of complexity in the current graph, a rotated bar-graph would be used so that readers would be able to easily locate data for a specific country and each bar-graph would be labelled with precise percentage to make ensure the principle of simplicity is applied and making it easier for comparison purposes. Using a rotated bar-graph can also improve clarity of the x-axis as countries would no longer be congested. In addition to apply the principle of simplicity, self-reported and measured statistics should be in separate graphs to reduce confusion. Separating the two types of statistics would improve data integrity and readers would be able to understand the graphs more clearly.

**Visualisation 2 – Hospital discharge rates**

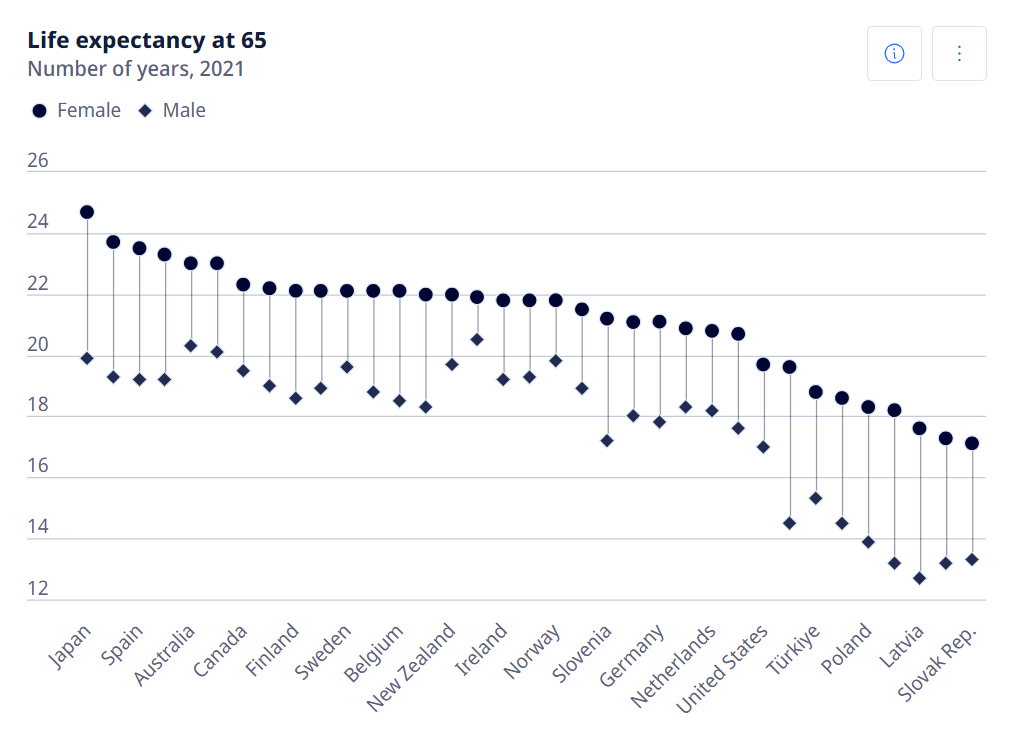
****

Analysis: The bar-graph above displays the hospital discharge rates across OECD countries, showing that Germany is most quick to discharge patients and Mexico is the slowest to discharge patients. A bar-graph is an appropriate data visualisation for this case study as readers can make comparisons much simpler across countries and identify which country has the highest or lowest discharge rates. The bar graph follows Tufte’s principle of graphical clarity, the data visualisation has its case study title clearly stated at the top of the graph with appropriate sizing and bold font, accompanied with a caption below showing how the hospital discharge rate is accounted for and the year that the data is taken from (Tufte, 1983). In addition, the graph also provides consistent scales across the data, ensuring consistency and data integrity. With an appropriate data visualisation that readers can understand, they can more likely understand the grasp the comparisons between the highest and the lowest discharge rates. In this case study, readers would be able to conclude that Germany’s healthcare is more efficient than Mexico’s.

Critique: There could be more improvements in terms of clarity according to the principles of a good data visualisation. Currently, it requires mental effort for the readers to locate the data of each country. For example, tracking Switzerland’s data can take some time as the X axis is congested with countries. This leads to difficulties in locating specific countries, violating Tufte’s principle of reducing non-data ink and maximizing the data-to-ink ratio (Tufte, 1983). Furthermore, the cognitive load for the reader is unnecessarily increased, which goes against Few’s recommendation for simplicity and clarity (Few, 2012). In addition, this can also confuse readers and lead to them drawing up false conclusions.

Solution: A rotated bar graph would offer a clearer presentation, reducing clutter and improving readability. Implementing appropriate spacing and rotating the graph will align with the principle of graphical excellence as outlined by Tufte (1983) and a[[1]](#endnote-1)s Few advocates, a well-spaced bar graph can significantly reduce the cognitive effort required by the viewer, making comparisons between countries more straightforward (Few, 2012).

**Visualisation 3 – Life expectancy at 65**



Analysis: The scatter plot diagram exhibits the difference of life expectancy at 65 between male and female. Directly from the diagram, it is noticeable that women from every OECD country have a higher life expectancy compared to men. The visualization adheres to key principles of clarity in effective data presentation (Few, 2006) as the title is stated at the top with appropriate font size and boldness, and below the title is a caption giving context of the year which the data is extracted from. A legend is provided along with different shapes, circle which is for female, and diamond which is for males. Readers are able to that women in general across the OECD countries live longer, as the data visualisation provided is clear and concise.

Critique: Although giving different shapes for each gender helps distinguish the data, readers would require looking at the legend beforehand to understand the graph, which would require additional mental effort to gauge the data visualisation given (Tufte, 2001). Besides that, it is difficult for readers to be able to read the data of each country as the x axis is clustered together. For example, getting data for Finland and Sweden can be confusing as the countries are all clustered together.

Solution: To make it simpler for readers to understand the graph, adding default colours for genders such as pink and blue, and adding the female and male icons in replacement of the shapes in the current graph, so that readers will be able to immediately gauge the data without needing a legend. In addition, to solve the x axis being clustered, increasing spacings between each country listed or slightly reducing the font size would be sufficient for clarity as this aligns with the principle of reducing unnecessary complexity in visual design, as emphasized by Tufte's data-ink ratio (Tufte, 2001).

**Conclusion**

It is important to apply principles of good data visualisations when presenting complex data, such as those collected by the OECD, so that industries and policymakers would be able to interpret data precisely and make well informed decision based on the data visualisation provided. The three visualisations selected in the report had applied some principles but still fell short in several areas, which would limit the effectiveness of the data presentation.

The first visualisation showing obesity rates, uses Tufte’s principle of graphical excellence by using clear labels and a consistent scale. However, placing self-reported and measured statistics in the same graph causes confusion to readers as there were few countries that had data for both measured and self-reported statistic. To improve the visualisation, separating the two types of statistics and placing the data on a rotated graph can simplify comparisons and cause less confusion to readers.

The second visualisation illustrating hospital discharge rates uses a bar graph which is good for comparison, but the congested x axis makes it difficult to locate data of a specific country. By using a rotated bar graph and improving spacing, the visualisation would be more reading and would align with Tufte’s principles of reducing non-data ink and Few’s principles of simplicity.

The third visualisation which compares life expectancy at 65 between men and women presents the comparison well but would benefit from using colour-coded gender indicators and increasing spacing between countries to make it easier to read. These adjustments would align with Few’s principles of reducing cognitive effort.

By applying these principles from experts like Tufte and Few, visualisations would be more effective and concise

## To be completed if this is a GROUP ASSIGNMENT

We declare that this is a group assignment and that no part of this submission has been copied from any other student's work or from any other source except where due acknowledgment is made explicitly in the text, nor has any part been written for us by another person.

ID Number Name Signature

Marker's comments:

Total Mark:

**Extension certification:**

This assignment has been given an extension and is now due on

Signature of Convener: Date: /2023

1. **References**

   Few, S. (2012). *Show Me the Numbers: Designing Tables and Graphs to Enlighten*. 2nd ed. Analytics Press.

   Tufte, E. (1983). *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

   Few, S. (2006). *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media.

   Tufte, E. R. (2001). *The Visual Display of Quantitative Information, 2nd edition*. Graphics Press.

   OECD (2021). Obesity or obese population rates. OECD Health Statistics

   OECD (2021). Life expectancy at 65 by sex. OECD Health Statistics

   OECD (2021). Hospital discharge rates. OECD Health Statistics [↑](#endnote-ref-1)