Swinburne University of Technology

COS30045

DATA VISUALISATION

Assignment 1 (Week 6)

Analysis of OECD Health Statistic Visualisations

Student Name: Chong Hong Liang

Student ID: 105215045

Section: C2

Table of Contents

1.0 Introduction

1.1 Report introduction

In today's technology world, data visualisation has been closely integrated into our lives as it helps us simplify communicating complex information. We often use it to help us understand and make decisions as it gives more informative data from raw data.

Data visualisation helps us translate raw data into graphical representations, allowing us to quickly discover the trends, patterns, and insights that could be ignored. It has been stated that this tool is essential in several fields that come with statistical data, especially in public health. For example, I will use the public health data provided by the Organisation for Economic Co-operation and Development (OECD) in this report. This will inform us about crucial policy decisions and public health initiatives.

1.2 Objective of the report

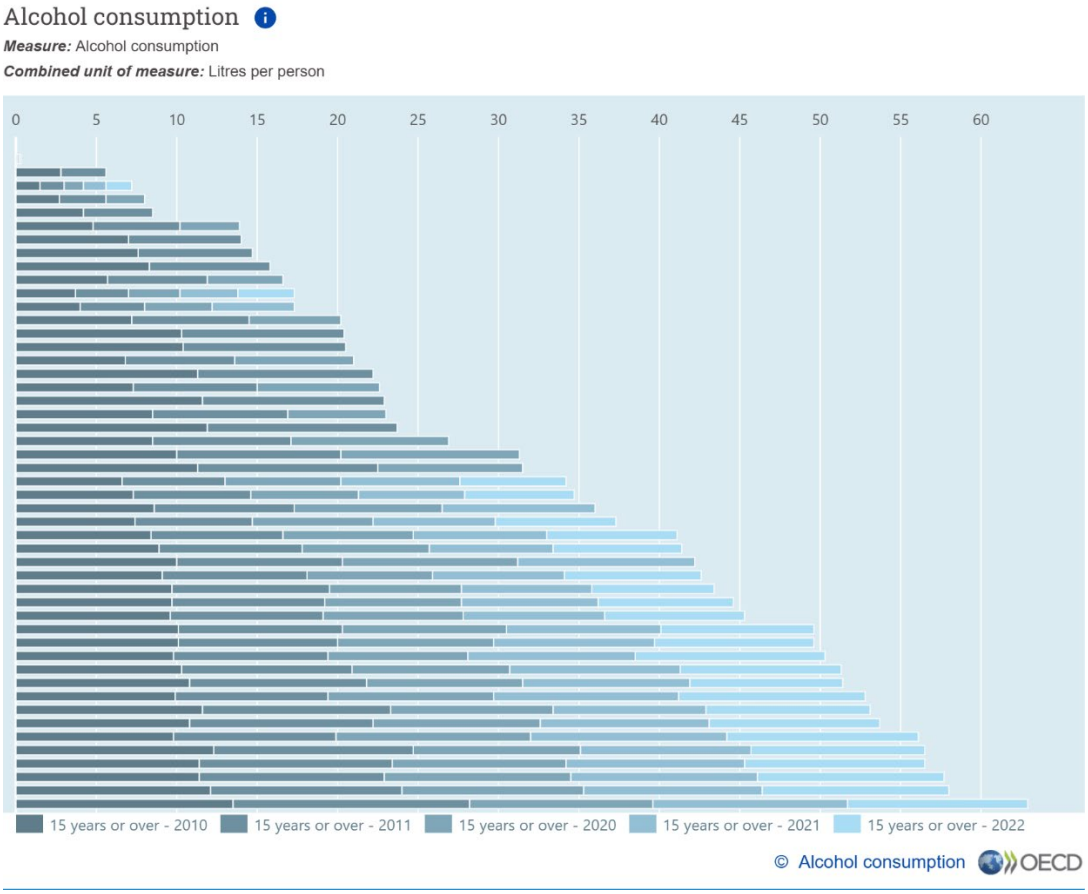
This report aims to identify three visualisations using OECD health statistics by critiquing the visualisations according to data visualisation principles and guidelines we learned in lecture class. This report will critique the visualisation, mainly focusing on Tufte’s integrity principles and visualisation design principles. The are a total of three visualisations have been selected, which are Alcohol consumption, use of vaping products and absence from work due to illness.

The visualisation will be critiqued based on these requirements: clarity, accuracy, design, efficiency, and ability to communicate the underlying data. Additionally, this report will also point out areas for improvement in these visualisations to improve data representation.

Lastly, each visualisation is divided into three sections: a visualisation introduction, a detailed analysis and critique and problem & enhancements to optimise the data representation. The goal is to ensure that these visualisations will fulfil the fundamental design principles, allowing the public to access the data more effectively and accessible.

2.0 Visualisations 1: Alcohol consumption

2.1 Visualisations 1: Alcohol Consumptions



Note. (2024). Alcohol Consumptions. OECD Data Explorer.

2.1.1 Visualisation Introduction

Alcohol consumption is a severe issue in this society as it keeps increasing over time. This will lead the individual to an unhealthy lifestyle. Alcoholic drinks have caused millions of people to die prematurely, based on research. In addition, the core causes of many serious illnesses are unhealthy lifestyle choices like cigarette smoking, excessive drinking, not exercising, and being overweight. Hence, public health strategies often focus on reducing harmful alcohol use to improve overall health outcomes and benefit by reducing the burden of disease in public.

2.1.2 Detailed analysis and critique

This stacked row chart represents data on alcohol consumption from the OECD Health Statistics, which focuses on people aged 15 years or older. This chart measured alcohol consumption based on litres per person and provided insights into consumption patterns across various countries over several years. This chart illustrates drinking habits over time by comparing alcohol consumption levels from 2010, 2011, 2020, 2021 and 2022.

Besides, this dataset contains over 2802 data points, which will provide accurate trends results in alcohol consumption trends. The last update to this data was made in July 2024, ensuring that it reflects the most current trends in alcohol consumption globally.

In this visualisation, the horizontal bars show alcohol consumption levels for each year, providing a clear representation of how consumption patterns have changed across different regions.

2.1.3 Problem & Enhancements

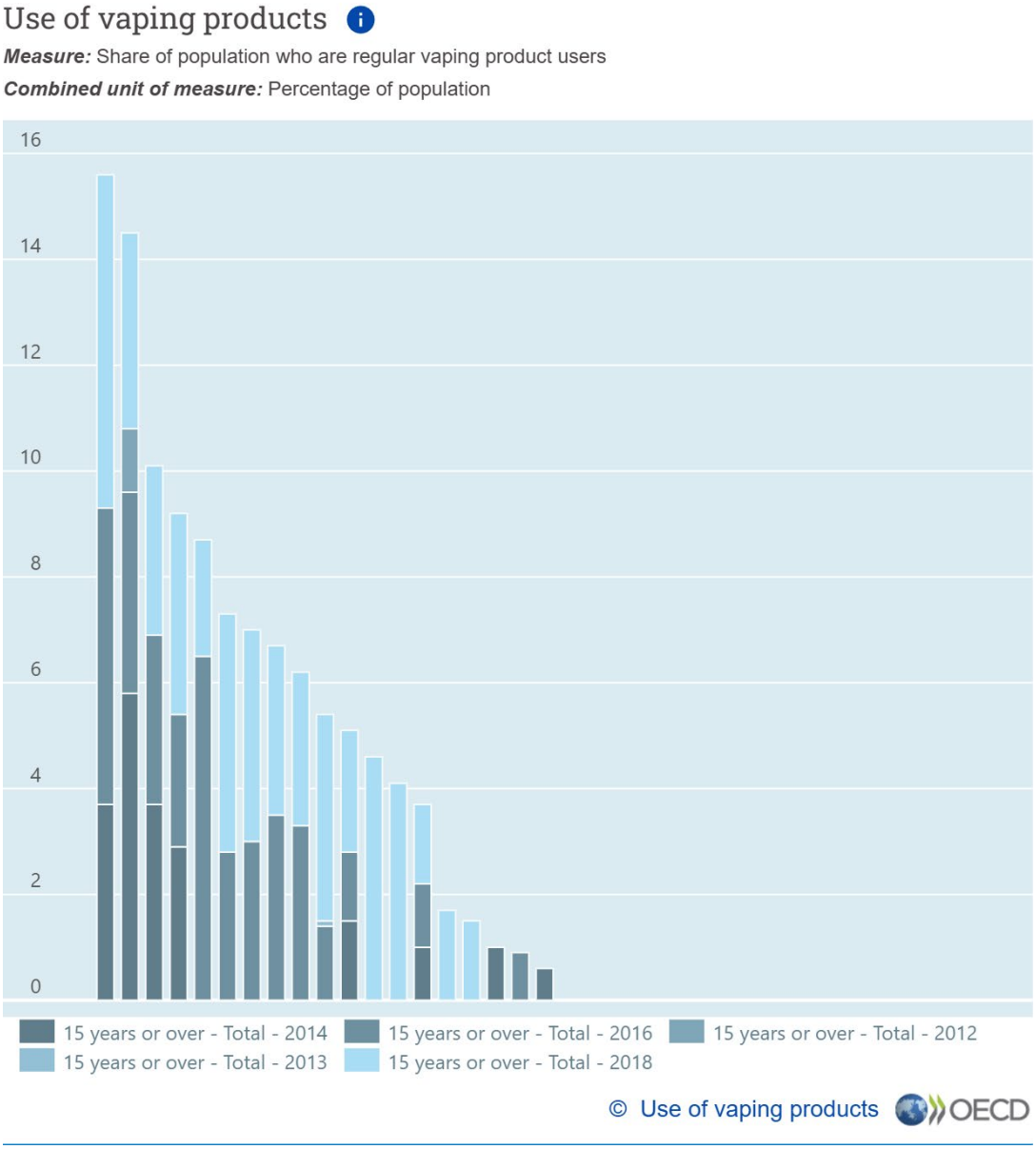
The visual hierarchy is weak on the first chart as it doesn't guide the viewer's eye naturally across the years in an intuitive way. Hence, we can implement "Alignment on this chart, which uses Tufte's Principle of Integrity and visualisation design principles. For instance, the chart data can align in the middle using "15 years or over – 2020". This will provide a more compact view without excessive spacing between bars, potentially allowing for quicker comparisons of values. (Michael, 2016)

Besides, this chart creates "chart junk" and unutilised white space, leading to a data-to-ink ratio issue. Hence, we can remove unnecessary elements like the background using data-to-ink ratio

visualisation design principles to provide a minimal chart version.(Raj, 2022) Furthermore, after implementing the first enhancements, we continue implementing this principle, as the modified one will utilise the white space. Hence, it also improved the clarity of the data.

Lastly, after implementing alignment and the data-to-ink design principle, the modified stack row bar chart will better compare candidates' data without scanning as far across the screen. While compared to the original chart, it will now be easier to draw such comparisons without scrolling or zooming.

2.2 Visualisations 2: Use of vaping products



Note. (2024). Use of vaping products. OECD Data Explorer.

2.2.1 Visualisation Introduction

Vaping products have become a significant public health issue in recent years due to their increasing popularity. Based on the chart, the share of the population who are regular vaping product users in Australia in 2019 is 3.2, while in 2022, it is 5.7, which has been increasing rapidly. The use of vaping products trends has been growing, especially among younger populations, because it is often

marketed as a safer alternative to smoking. However, the long-term health effects remain uncertain.

2.2.2 Detailed analysis and critique

This stacked column chart represents data on the use of vaping products among individuals aged 15 years or older from the OECD Health Statistics. The use of vaping products is measured in the share of the population who are regular users of vaping products, with the unit of measure being the percentage of the population. This chart also illustrates the number of vaping product users between 2012 and 2018.

Besides, this dataset contains over 756 data points, providing valuable information about the trends in vaping habits and offering a breakdown by age and sex. Furthermore, this chart also visualises how the share of the population using vaping products has fluctuated or grown over the years. Lastly, the last update to this dataset was in July 2024, ensuring the data reflects the most current trends in global vaping product use.

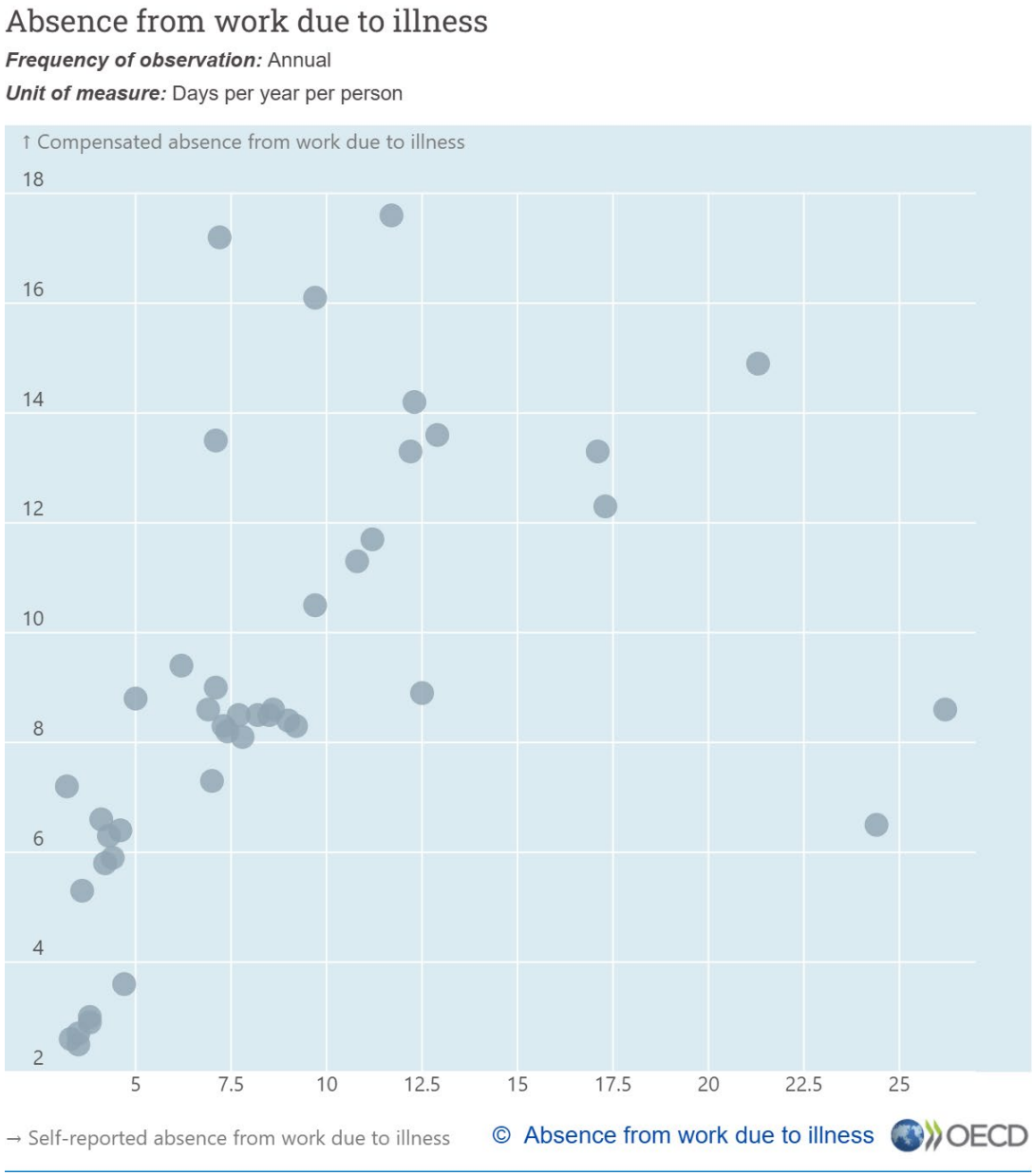
2.2.3 Problem & Enhancements

This chart needs to adjust and redesign each colour contrast on each chart bar as it can cause visual confusion. (Disney, 2024). Moreover, the shades of blue and grey are very similar in tone, making it difficult to differentiate between the various years represented. We can use different colours for each identical data to highlight and create more differentiation between the critical data. Hence, using a more distinct or contrasting colour would improve readability.

Besides, there is a missing label on this chart. The labelling is missing on the X and Y axes, which failed to provide clarity and

accuracy to the users. The x-axis is not labelled with the countries or categories being represented, making it hard for the reader to understand the data context without knowing what these bars represent. Hence, adding simple and straightforward labels would improve this. On the other hand, the Y-axis is also not labelled with the numbers being represented. While the chart title mentions the percentage of the population using vaping products, adding a label such as “Percentage of Population (%)” to the Y-axis will make it clear to viewers without them having to infer the meaning. Hence, this will avoid confusion and improve the overall understanding of the chart.

2.3 Visualisations 3: Absence from work due to illness



Note. (2024). Absence from work due to illness. OECD Data Explorer.

2.3.1 Visualisation Introduction

Absence from work due to illness significantly impacts productivity and public health. This chart compares data on self-reported and compensated absences from work across multiple countries. Based on the chart, the country with the highest value on self-reported absence from work due to illness is Lithuania, with a value of 26.2. On the

other hand, the country with the highest compensated absence from work due to illness is awarded Germany with a value of 17.6. This dataset highlights potential differences in workplace health policies and practices by reflecting how different countries report and compensate for work absences due to illness. Lastly, absences from work are an essential indicator of the general health of the workforce and are regularly tracked by countries.

2.3.2 Detailed analysis and critique

This scatter plot chart represents data on absences from work due to illness, using a unit of measure of days per year per person. The X-axis on this chart represents reported absences, and the Y-axis represents compensated absences, compared across various countries.

There are 997 unfiltered data points, each point on the chart representing a country. Besides, data is collected through an annual frequency of observation. Importantly, this dataset excludes maternity leave and only focuses on absences due to illness. The data was last updated in July 2024, ensuring the data reflects the most recent information available.

2.3.3 Problem & Enhancements

Some modifications can be made to the point size to improve the understanding and clarity from the first sight. We can modify the point size dynamically based on the third value. This will give the chart more dimensionality without adding clutter. For instance, it will provide more context and depth if point size represents another variable like population or economic impact. (Atlassian, n.d.)

Implementing dynamic point size in the chart can facilitate a more engaging user experience. The interactive features will encourage

viewers to explore the data further. However, it is also important to ensure the point size scaling is intuitive and clearly explained in the chart legend. This will help maintain the data representation's integrity and avoid potential confusion.

Besides, highlighting by making the point darker and country name labelling directly can be done on the chart on the top 3 countries that provide the highest compensated absence from work due to illness and the top 3 countries with the highest self-reported absence from work due to illness. This will be a more effective way of improving the clarity of the chart data to the viewer as it allows viewers to interpret without the clutter of too many data points.

3.0 Conclusion

This report discusses three visualisations from the OECD Health Statistics, focusing on alcohol consumption, the use of vaping products, and the absence of work due to illness. The analysis applied the requirements in clarity, accuracy, design, efficiency, and ability to communicate the underlying data. By using Tufte's integrity principles and implementing visual design improvements, such as better chart alignment and removing unnecessary elements, the suggested enhancements aim to improve the data's accessibility and readability for public health decision-makers.

These enhancements not only ensure the effectiveness of the data communication but also provide accuracy. Lastly, this report aims to achieve valid practice visualisation principles to ensure their usefulness in real-world applications.

4.0 References

Atlassian. (n.d.). *Mastering scatter plots: Visualize data correlations*.

https://www.atlassian.com/data/charts/what-is-a-scatter-plot

Atlassian. (n.d.-a). *A complete guide to bubble charts*.

https://www.atlassian.com/data/charts/bubble-chart-complete-guide

Disney, A. (2024, April 2). *Color theory for Data Visualization*. Cambridge Intelligence.

https://cambridge-intelligence.com/choosing-colors-for-your-data-visualization/

Michael. (2016, October 23). *Political Dataviz: Who lies more – a comparison (Robert Mann) – updated*. Michael Sandberg's Data Visualization Blog. https://datavizblog.com/2016/07/24/political-dataviz-who-lies-more-a-comparison-robert-mann/

Raj, A. (2022, November 10). *Data-ink ratio explained with example*. Code Conquest.

https://www.codeconquest.com/blog/data-ink-ratio-explained-with-example/#:~:text=To%20maximize%20the%20data%2Dink%20ratio%2C%20Edward%20Tufte%20suggested%20two,be%20deleted%20from%20the%20visualization.

OECD (Organisation for Economic Co-operation and Development). (2024, July 12).

Alcohol consumption. Data.https://data-explorer.oecd.org/vis?fs%5B0%5D=Topic%2C0%7CHealth%23HEA%23&pg=0&fc=Topic&bp=true&snb=79&df%5Bds%5D=dsDisseminateFinalDMZ&df%5Bid%5D=DSD_HEALTH_LVNG%40DF_HEALTH_LVNG_AC&df%5Bag%5D=OECD.ELS.HD&df%5Bvs%5D=1.0&dq=.A.....&pd=2010%2C&to%5BTIME_PERIOD%5D=false&ly%5Bcl%5D=TIME_PERIOD&ly%5Brs%5D=AGE&ly%5Brw%5D=REF_AREA&vw=sr

OECD (Organisation for Economic Co-operation and Development). (2024, July 12).

Use of vaping products. Data. https://data-explorer.oecd.org/vis?df%5Bds%5D=dsDisseminateFinalDMZ&df%5Bid%5D=DSD_HEALTH_LVNG%40DF_HEALTH_LVNG_VP&df%5Bag%5D=OECD.ELS.HD&df%5Bvs%5D=1.0&dq=.A.....&pd=2012%2C&to%5BTIME_PERIOD%5D=false&ly%5Bcl%5D=TIME_PERIOD&ly%5Brs%5D=SEX%2CAGE&ly%5Brw%5D=REF_AREA&vw=ov

OECD (Organisation for Economic Co-operation and Development). (2024, July 23).

Absence from work due to illness. Data. https://data-explorer.oecd.org/vis?df%5Bds%5D=dsDisseminateFinalDMZ&df%5Bid%5D=DSD_HEALTH_STAT%40DF_AWDI&df%5Bag%5D=OECD.ELS.HD&df%5Bvs%5D=1.0&dq=.A.....&pd=2010%2C&to%5BTIME_PERIOD%5D=false&vw=sp