LIFE EXPECTANCY

Life expectancy is the **number of years** an individual can **expect to live**.

The data in this chosen dataset was collected from the Global Health Observatory (GHO) data repository under the World Health Organization (WHO) and from United Nation website.

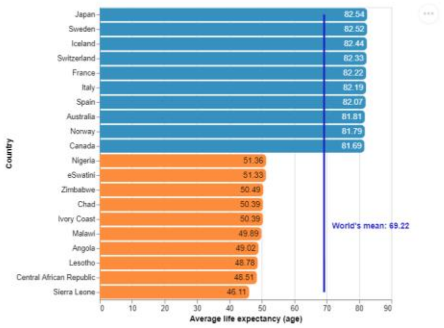
This dataset contains data from year **2000 to 2015** for **193 countries** for further analysis.

All the predicting variables in this dataset had been divided into several broad categories, such as **Immunization** related factors, **mortality** factors, **economical** factors, and also **social** factors.

1 COUNTRIES WITH THE 10 HIGHEST AND LOWEST AVERAGE LIFE EXPECTANCY RESPECTIVELY ACROSS THE WORLD FROM 2000 TO 2015

To begin with, let us take a look at the top and bottom 10 countries across the world based on their average life expectancy over the 16 years.

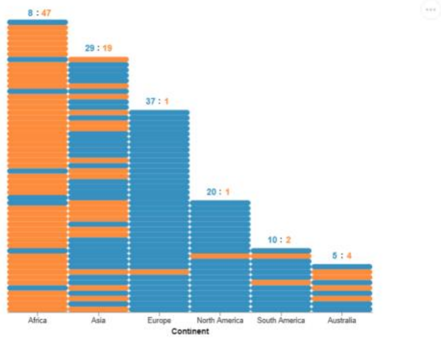
Japan had been the champion in the world by having the highest average life expectancy, whereas **Sierra Leone** had been staying at the bottom one by having the lowest average life expectancy, among all other countries.



2 COUNTRIES WITH AVERAGE LIFE EXPECTANCY ABOVE AND BELOW THE WORLD'S MEAN LIFE EXPECTANCY ACROSS DIFFERENT CONTINENTS

How were the countries distributed across different continents based on their status of average life expectancy, that is, whether below or above the world's mean life expectancy?

It is interesting to note that both **Europe** and **North America** had an extremely **large portion** of those above-mean countries, whereas **Africa** had a very **large portion** of those below-mean countries.

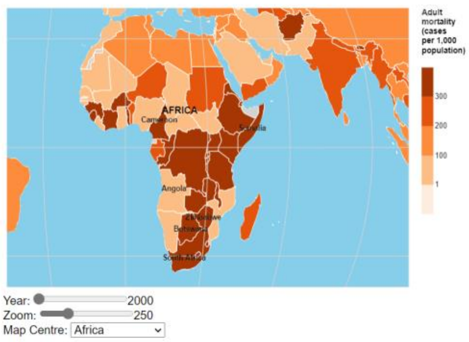


3 ADULT MORTALITY RATE IN DIFFERENT COUNTRIES ACROSS THE WORLD FROM 2000 TO 2015

We first inspect the effect of adult mortality rate on the average life expectancy across different countries.

We see that the countries with relatively **higher** adult mortality rate were predominantly in the **Africa** continent, such as **Cameroon**, **Somalia**, **Angola**, **Zimbabwe**, **Botswana** and **South Arfica**.

Therefore, it shows that **high** adult mortality rate was one of the major causes of a country's **low average life expectancy**.

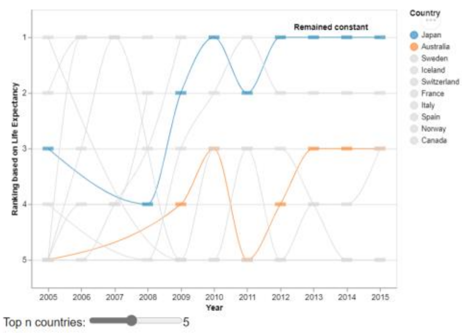


4 RANKING AMONG THE TOP 10 COUNTRIES WITH THE HIGHEST AVERAGE LIFE EXPECTANCY OVER YEARS

Next, we might be interested in seeing the ranking among different countries based on their average life expectancy over years.

We can observe that the average life expectancy of Japan had kept unchanged as the **highest** among all other countries for **4 consecutive years** from 2012 to 2015.

Also, we discover that the ranking of Australia over years showed a **similar pattern** as that of Japan.

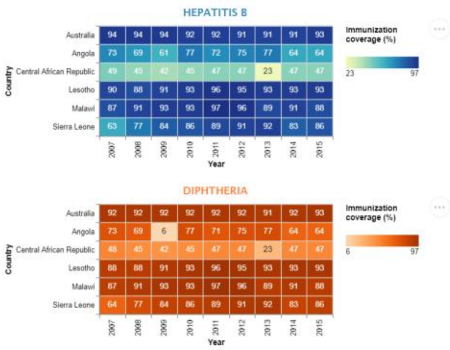


5 HEPATITIS B AND DIPHTHERIA IMMUNIZATION COVERAGE IN DIFFERENT COUNTRIES OVER YEARS

The low average life expectancy in **Angola** and **Central African Republic** (as shown in Section 1) could be explained by their **low Hepatitis B and Diphtheria immunization coverage**.

Online resources have been found to support the reasons behind **outliers** such as 6 and 23.

In **2009**, there was a serious flood disaster happening in **Angola** *(Wikipedia, n.d.)*, plus a critical political conflict happened between **Angola** and Democratic Republic of Congo *(British Broadcasting Corporation [BBC], 2018)*. Meanwhile, there also happened the **Central African Republic conflict** in **2013** *(Wikipedia, n.d.)*. These incidents could potentially be the reasons leading to such **low immunization coverage** in both countries.

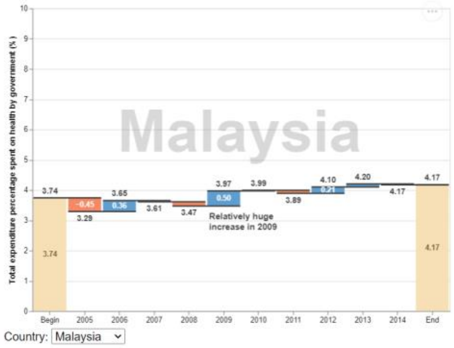


6 TOTAL EXPENDITURE PERCENTAGE SPENT ON HEALTH SECTORS BY GOVERNMENT OVER YEARS

From the comparison between **Australia** and **Malaysia**, we could infer that government's contribution is important in improving a country's average life expectancy.

We see that the total expenditure percentage spent by **Australia**'s government was **significantly higher** than that of **Malaysia**'s. This difference could potentially explain the reason why Malaysia was **not** in the Top 10 Ranking Board (as shown in Section 4).

Nevertheless, in overall there was a **slight improvement** on the total expenditure percentage spent on health by the government of both countries from 2005 to 2014.

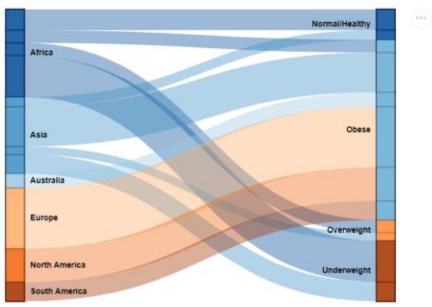


7 DISTRIBUTION OF COUNTRIES IN BETWEEN DIFFERENT CONTINENTS AND AVERAGE HEALTH STATUSES BASED ON BMI

Question may arise such as does BMI status affect the life expectancy of an individual?

Before that, let us have an overview on the distribution of countries based on their average BMI status across different continents.

We can clearly note that the average BMI status of all **European, North** and **South American** countries appeaerd to be **Obese**. It tells that for every person in those countries, we may expect that he/she was obese.

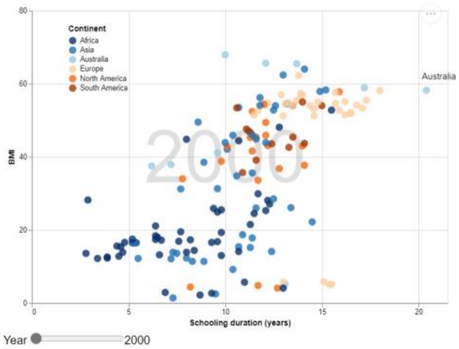


8 INTERACTIVE PLOT SHOWING RELATIONSHIP BETWEEN SCHOOLING DURATION WITH BMI ACROSS DIFFERENT CONTINENTS AND TRENDS OVER TIME

According to the scatter plot shown below, there seemed to have **moderate to strong positive correlation** between average schooling duration with average BMI. This statement could be supported by *Abrams, Cohen, Rai, & Rehkopf (2013)*.

From 2000 to 2015, the average schooling duration worldwide had been gradually **increasing**. Meanwhile, **Australia** had constantly be the country that had the **highest average schooling duration** over the 16 years.

Another interesting finding is that European countries had relatively **higher** average schooling duration, whereas African countries had relatively **lower** average schooling duration among all.

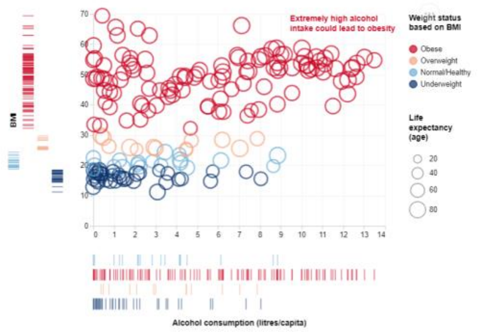


9 ASSOCIATION BETWEEN ALCOHOL PER CAPITA CONSUMPTION, BMI AND LIFE EXPECTANCY

We find out that BMI was **slightly or moderately associated** with alcohol consumption. However, from the bubble plot, it was visible that when the alcohol per capita consumption was **too high** (more than 10 litres/capita), then the BMI status always showed an Obese.

Other than that, we realize that both alcohol consumption and BMI **did not** seem to affect the life expectancy, as the the big and small-sized bubbles were scattered everywhere and **did not** really gather at any position in the plot. So, **no clear pattern** was displayed.

From the dash plot, we learn that a **large proportion** of the population worldwide was obese.

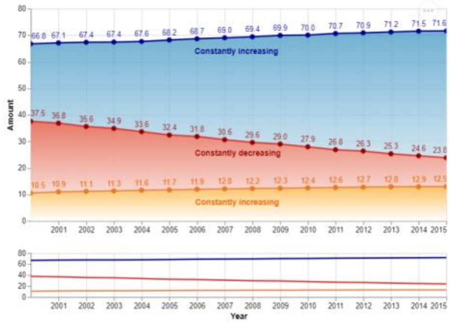


10 CHANGES IN THE WORLD'S AVERAGE LIFE EXPECTANCY, INFANT DEATH RATE AND SCHOOLING DURATION FROM 2000 TO 2015

Last but not least, we might also be curious about the correlation between infant death rate and schooling duration with the average life expectancy.

As indicated below, from 2000 to 2015, the **world's average** life expectancy and average schooling duration were constantly **improving**. Following this, we discover that **lower** infant death rate and **higher** schooling duration could lead to **higher average** life expectancy.

In conclusion, since we know that high schooling duration brought to high BMI (as shown in Section 8) and high average life expectancy, hence we could probably deduce that **obesity** was **not** a culprit of **low life expectancy**.



BIBLIOGRAPHY

Abrams, B., Cohen, A. K., Rai, M., & Rehkopf, D. H. (2013). Educational attainment and obesity: A systematic review. *Obesity Reviews, 14*(12), 989-1005. doi: 10.1111/obr.12062

British Broadcasting Corporation. (2018, March 7). Angola profile – Timeline. Retrieved from https://www.bbc.com/news/world-africa-13037271

Kumarrajarshi. (2017). *Life expectancy (WHO)* [Data file]. Retrieved from https://www.kaggle.com/datasets/kumarajarshi/life-expectancy-who

Wikipedia. (n.d.). Central African Republic conflict (2013-2014). Retrieved from https://en.wikipedia.org/wiki/Central_African_Republic_conflict_(2013%E2%80%932014)

Wikipedia. (n.d.). 2009 Angola, Namibia and Zambia floods. Retrieved from https://en.wikipedia.org/wiki/2009_Angola,_Namibia_and_Zambia_floods#:~:text=The%202009%20Angola%2C%20Namibia%20and%20Zambia%20floods%20was,two%20regions%20of%20Angola%20and%20part%20of%20Botswana

Visualisation by Chong Ming Sheng (Student ID: 32202792) for FIT3179 Data Visualisation 2, Monash University Malaysia.
The dataset chosen is retrieved from Kaggle (Kumarrajarshi, 2017), and the data has been preprocessed properly before use.

URL of Visualisation

https://mingsheng2002.github.io/3179-Data-Visualisation-2/

Domain

My visualisation mainly focuses on average life expectancy of population from different countries from 2000 to 2015, and also the potential factors affecting their life expectancy.

Why

My visualisation shows the effects of certain predicting factors have on the average life expectancy of population based on data collected in the past few years. This visualisation can be used to predict the life expectancy of an individual based on his/her current residing country, living lifestyles and health status. Therefore, perhaps extra precautions could be taken by individuals and government to improve the population's health level in order to prolong their life expectancy.

Who

My visualisation aims to show to the worldwide population about some possible factors that may influence their life expectancy according to the past data.

What

This table dataset was retrieved from Kaggle website (Kumarrajarshi, 2017). It includes information of different countries based on certain aspects.

Data had been cleaned and preprocessed properly via Jupyter Notebook to remove outliers and impute missing values where necessary. Data transformation had been carried out where appropriate to display certain chart(s).

TopoJSON files used in the map idiom were retrieved from the map example given by Liu (2022).

Why and How

Figure 1: Choropleth map.

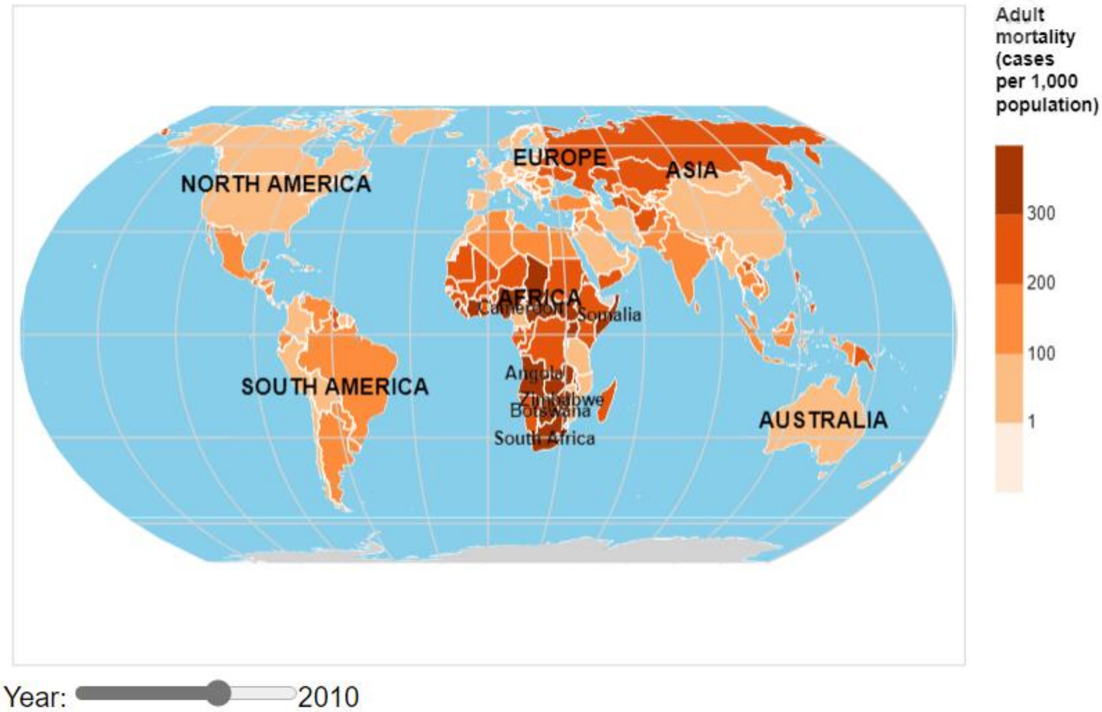


Figure 2: First special feature of choropleth map.

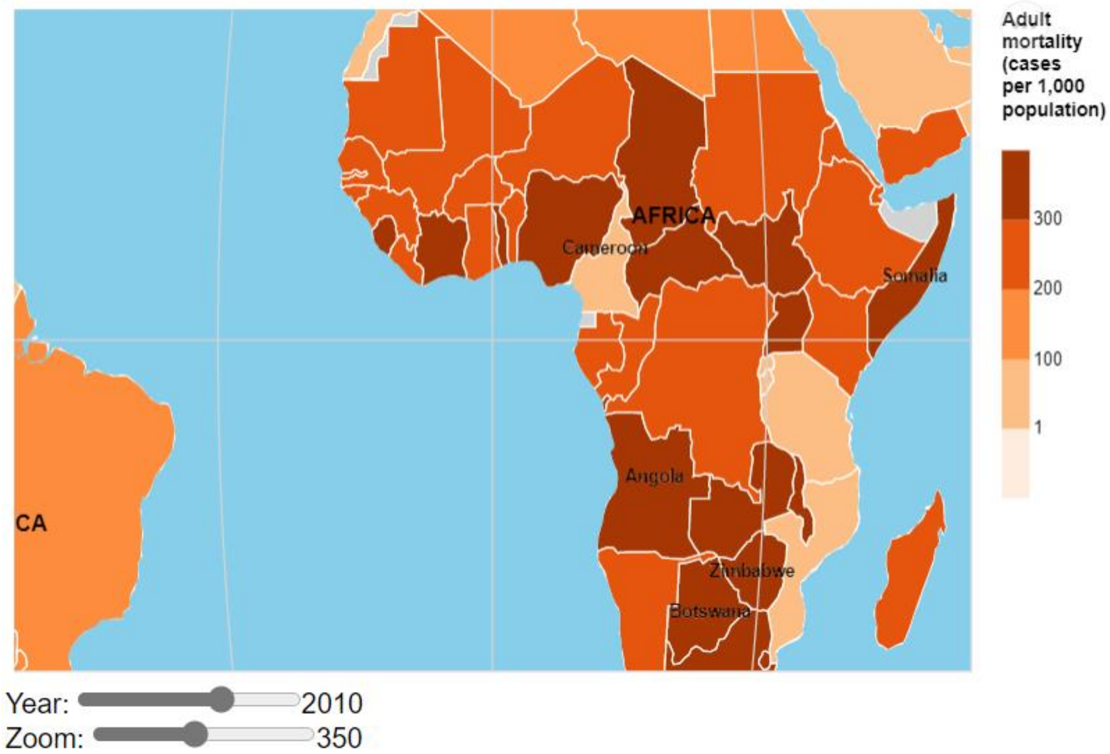


Figure 3: Second special feature of choropleth map.

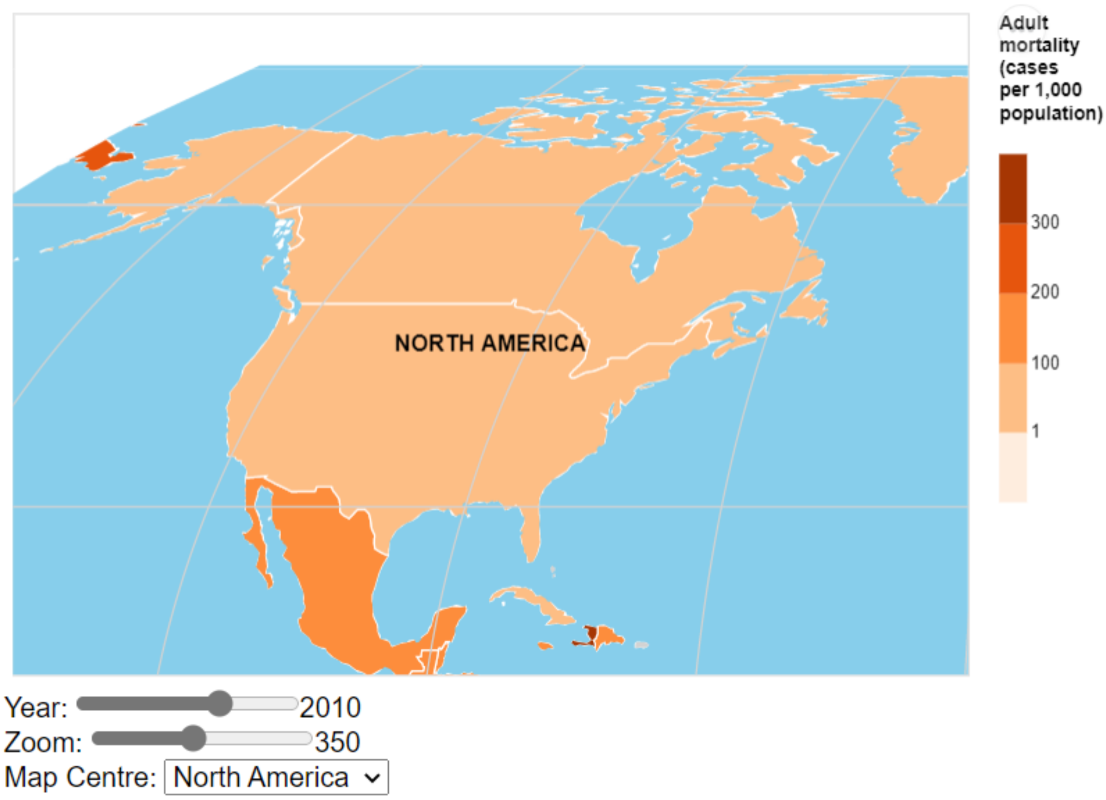


Figure 4: Third special feature of choropleth map.

Created with reference to “Week 8 studio activity: Create maps with Vega-Lite” (n.d.), idiom in Figure 1 allows readers to have an overview on the adult mortality rate across different countries more intuitively. Readers can compare the adult mortality rate in different countries easily at a glance.

Mark used for this choropleth map is geographical areas, whereas channel used is colour luminance to encode the adult mortality rate in different countries.

Based on Figure 2, users can select to visualise the adult mortality rate in a specific year via slider. Based on Figure 3, zoomable map allows users to visualise the adult mortality rate in smaller countries in detail via slider. Based on Figure 4, users can also zoom into a specific continent to focus on particular countries via selection menu.

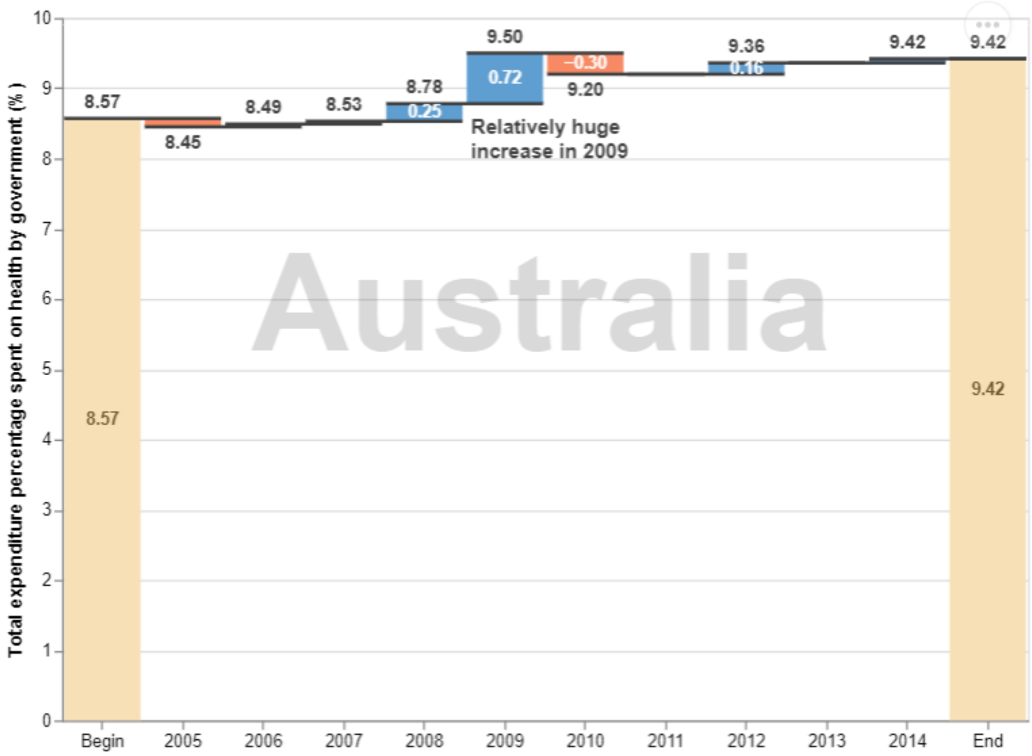


Figure 5: Waterfall chart.

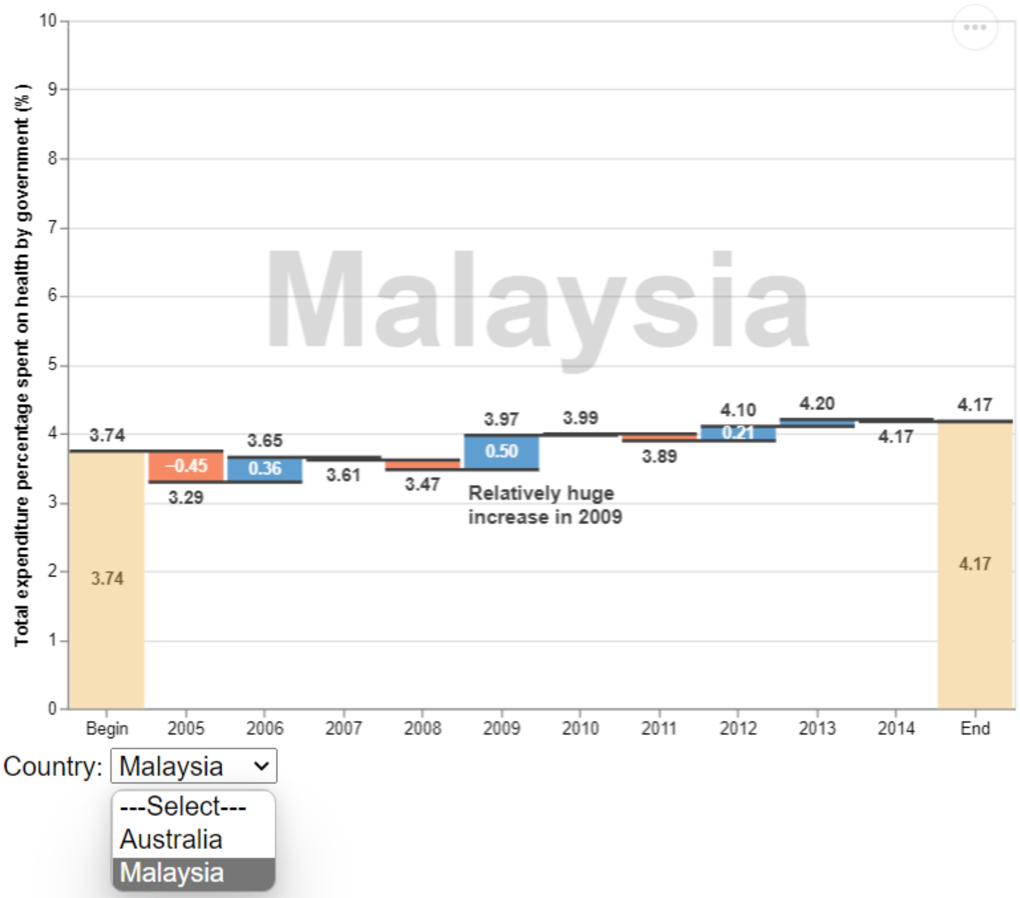


Figure 6: Special feature of waterfall chart.

Created with reference to “Waterfall chart of monthly profit and loss” (n.d.), idiom in Figure 5 clearly displays the changes in total expenditure percentage spent on health by government over years. Rather than showing absolute data values, the difference in percentages between every two adjacent years is shown to the readers so that they can obtain more information from the chart.

Mark used for this waterfall chart is lines, whereas channels used are horizontal and vertical positions, vertical length and colour hue, which respectively encode quantitative attribute value along y-axis and year along x-axis, total expenditure percentage spent, and changes in percentage (either increase or decrease).

Based on Figure 6, users can select to visualise the total expenditure percentage spent on health by either Australia’s or Malaysia’s government over years via selection menu. This feature was purposely implemented to avoid overlapping of data to be displayed in the waterfall chart.

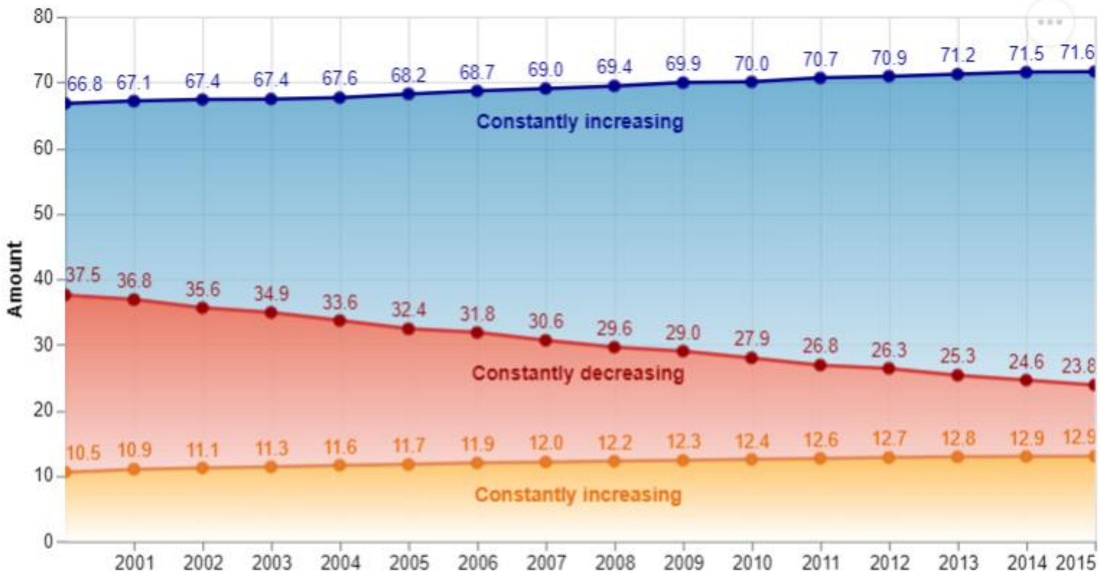


Figure 7: Area chart.

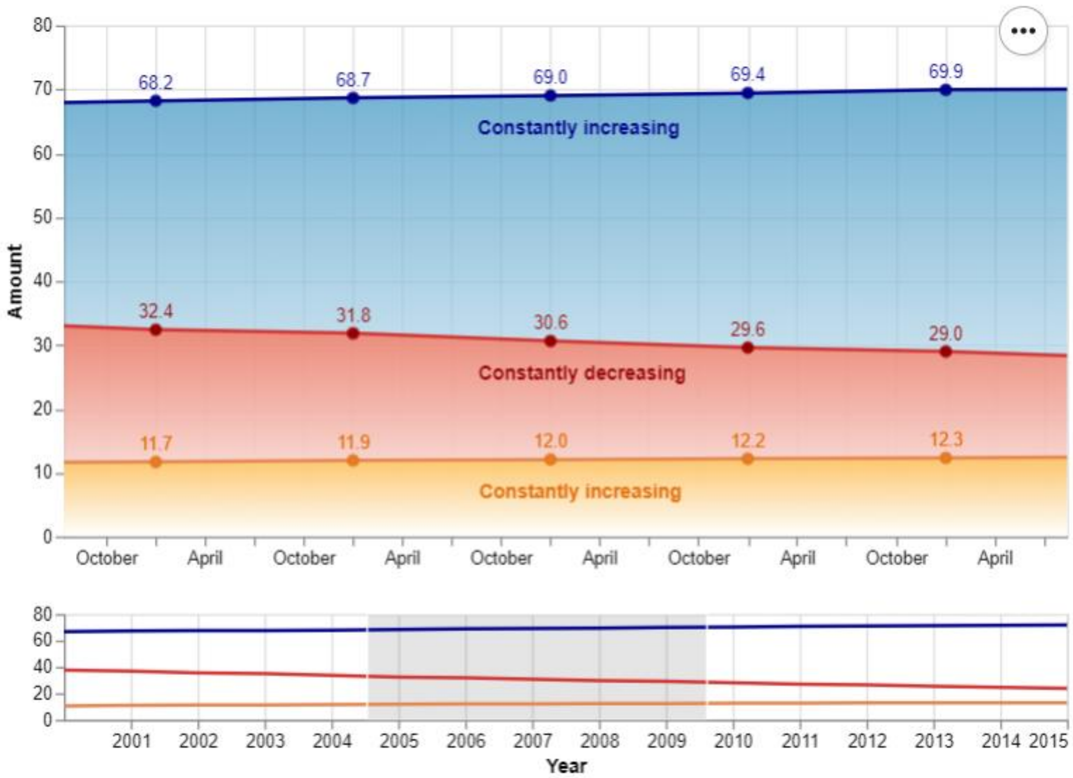


Figure 8: Special feature of area chart.

Created with reference to “Area chart with gradient” (n.d.), idiom in Figure 7 depicts the changes in average life expectancy, infant death rate and schooling duration over years. Meanwhile, it shows the correlation between these three factors.

Marks used for this area chart are areas, lines and points, whereas channels used are colour hue and colour luminance, which respectively encode the three different factors and also the quantitative values of factors displayed in the chart.

Based on Figure 8, readers can visualise data in a certain interval (year range) at one time via the brushing feature. Instead of showing overview of data, more detailed information could be observed clearly by limiting a certain time interval for the data to be visualised.

Design Layout

Apparently, my visualisation is split into six rows and two columns. Whitespaces were properly utilised to create borders and proximity for my visualisation, by visually separating different charts so that readers could easily group relevant elements together visually. This could minimise the number of sight lines and make the entire visualisation balanced and symmetrical.

Design Colour

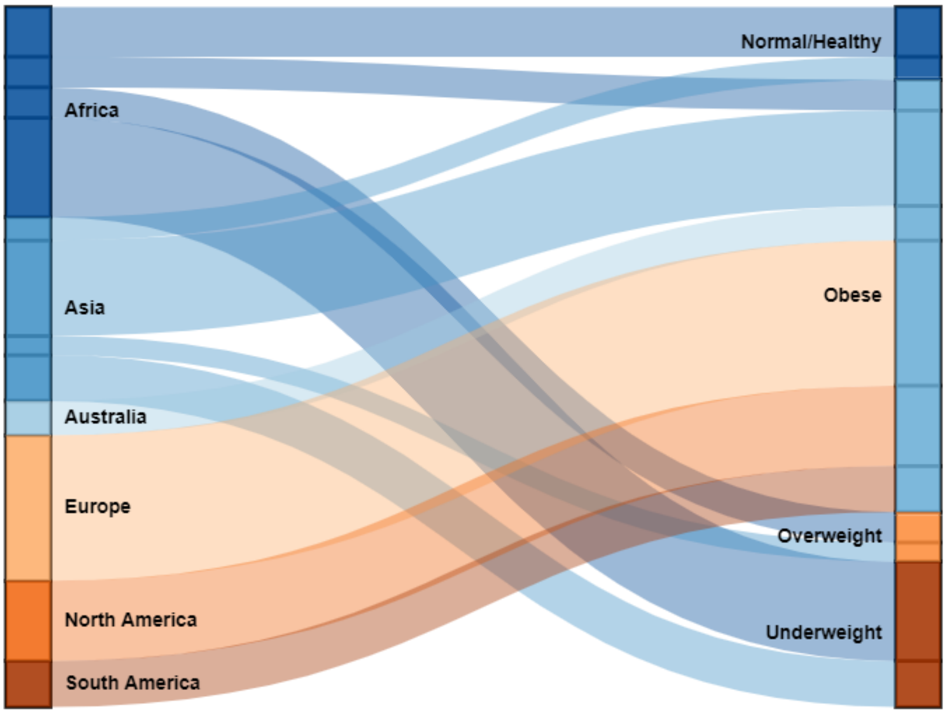


Figure 9: Consistent use of colours (red, orange, blue and pale blue).

I selected the colour-blindness palette consisting of only red, orange, blue and pale blue that are visually friendly for colour blind people. This is important so that they would not face any difficulties in differentiating the colours applied in my entire visualisation.

Consistently, I had applied these colours to every chart and annotation to intuitively encode specific categorical attributes in my entire visualisation where applicable and appropriate, for example as shown in the idiom in Figure 9 which was created with reference to Astrakhan (2018).

Design Figure-ground

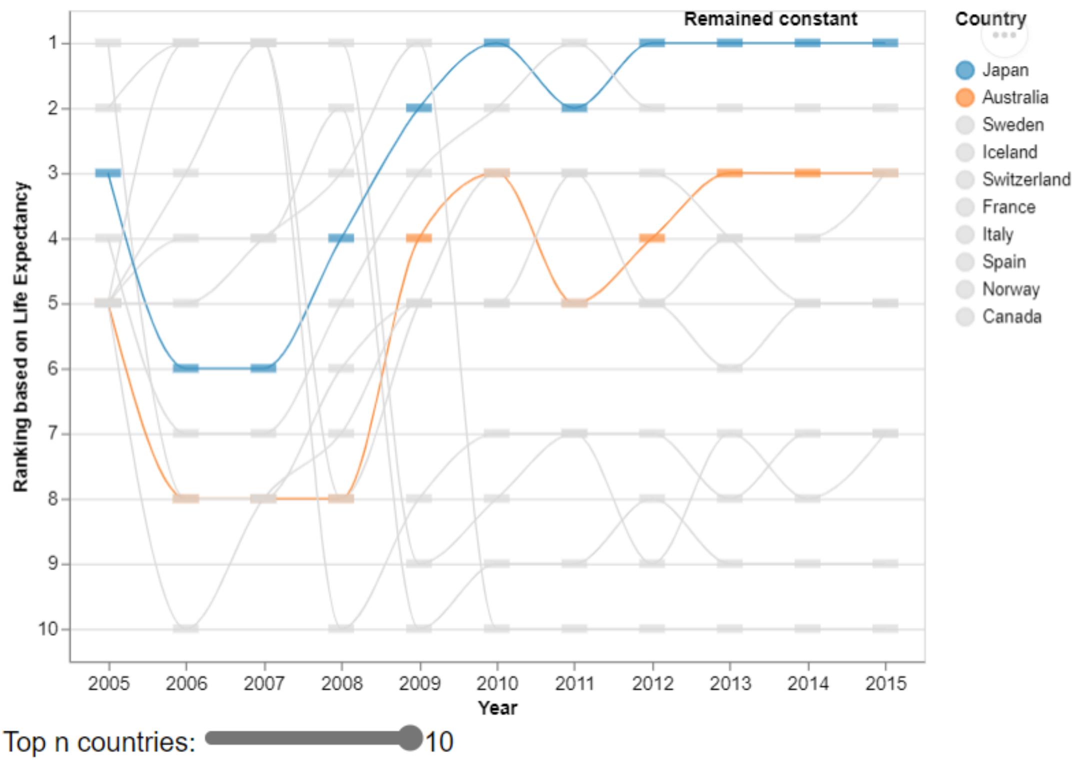


Figure 10: Less important data was encoded with desaturated colour.

Figure 11: Visual hierarchy was created via application of typography.

Created with reference to Borges (2021), idiom in Figure 10 shows that more saturated colours like blue and orange were applied to depict more important data, which is Japan and Australia in this case. Desaturated colour like light gray was applied to encode unimportant data. This ensures that important data could look prominent to the readers.

Additionally, as shown in Figure 11, the visualisation title and chart titles used bigger, bolder and capitalized texts. Meanwhile, bolder and coloured texts were used in captions to highlight important keywords.

Design Typography

Life expectancy is the **number of years** an individual can **expect to live**.

The data in this chosen dataset was collected from the Global Health Observatory (GHO) data repository under the World Health Organization (WHO) and from United Nation website.

This dataset contains data from year **2000 to 2015** for **193 countries** for further analysis.

All the predicting variables in this dataset had been divided into several broad categories, such as **Immunization** related factors, **mortality** factors, **economical** factors, and also **social** factors.

Figure 12: Captions in two text blocks.

My visualisation used altogether four typefaces, which are Lucida Sans, Josefin Sans, Montserrat and Arial. Without serifs, these typefaces are highly readable. Except for Arial, the first three typefaces mentioned belong to non-standard typefaces and hence they might look interesting to readers.

Referring to Figure 11 previously, Josefin Sans makes the visualisation title the most stand out by giving it extra bold. Captions written in Montserrat are easily distinguishable from the chart titles written in Lucida Sans. Chart annotations were written in Arial to be consistent with the axis and legend titles in charts.

Based on Figure 12, dividing the captions into two text blocks could improve their readability, as each line of texts consists of not more than ten words.

Besides, to reduce sight lines, captions below chart titles were put in horizontal text layout.

Design Storytelling

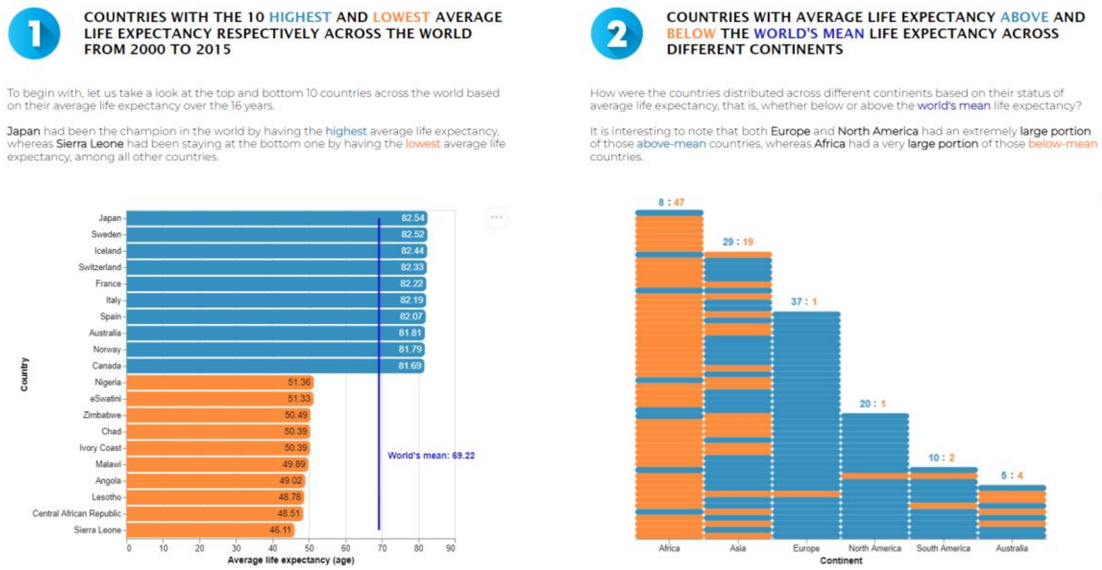


Figure 13: Flow of charts was indicated by number icons.

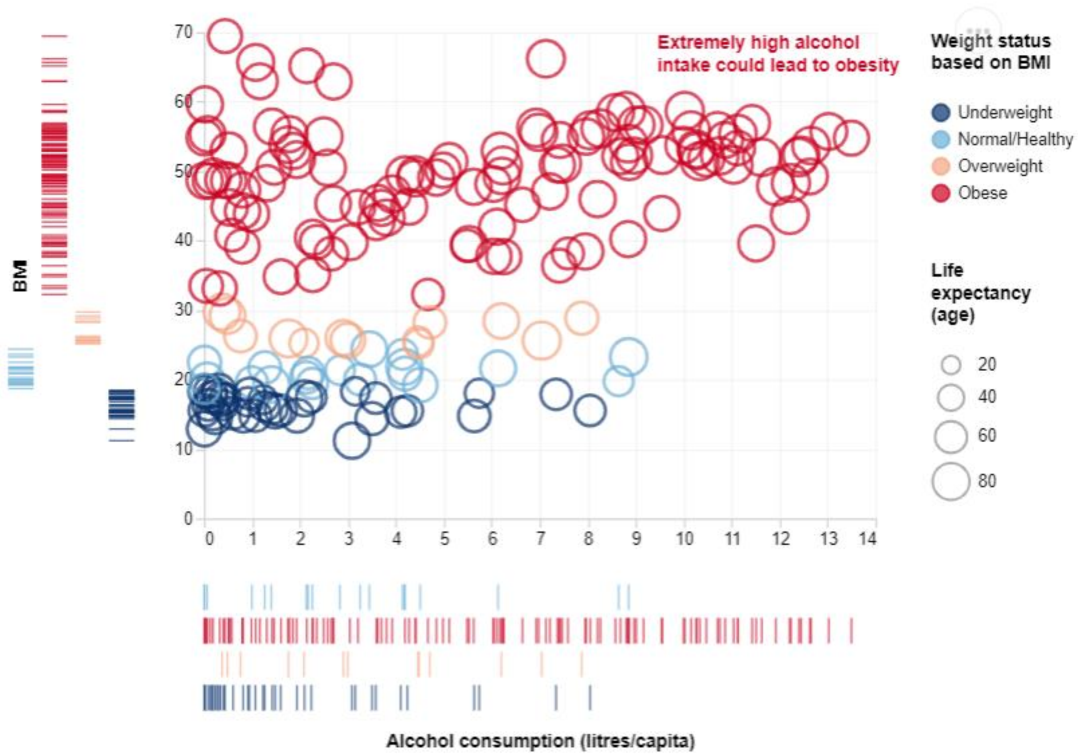


Figure 14: Message was conveyed via chart annotation.

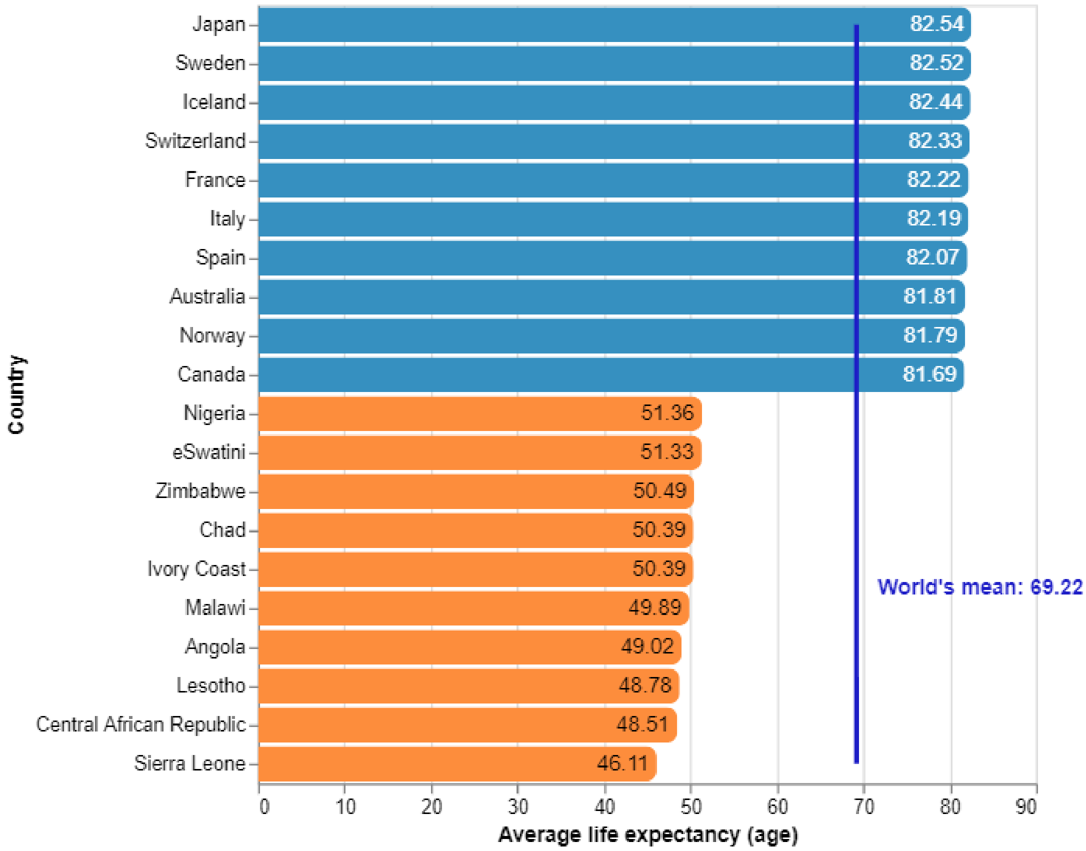


Figure 15: Data displayed was sorted in descending order.

Based on Figure 13, the flow of charts in my visualisation was indicated by the number icons, so that readers could be guided through my visualisation by reading and viewing the charts according to the particular order.

Furthermore, according to Figure 14 in which the idiom was created with reference to g3o2 (2017), there was annotation provided to attractively convey message to readers and guide them through the annotated visualisation.

Lastly, data displayed in some charts in my visualisation was sorted in a specific order. One example was shown in Figure 15, in which the idiom was created with reference to “Calculate difference from average” (n.d.). Intuitively, readers could be guided to read and view the data from top to bottom, in order to know about the ranking of countries by their average life expectancy.

Bibliography

Area chart with gradient. (n.d.). Retrieved from https://vega.github.io/vega-lite/examples/area_gradient.html

Astrakhan, Y. (2018, March 30). Sankey visualisation with Vega in Kibana 6.2 [Web blog post]. Retrieved from https://www.elastic.co/blog/sankey-visualization-with-vega-in-kibana

Borges, J. L. M. (2021, July 9). Vega-Lite to create a rank-plot for car registrations in Portugal [Web blog post]. Retrieved from https://observablehq.com/@jlmborges/rank-plot_vega-lite

Calculate difference from average. (n.d.). Retrieved from https://vega.github.io/vega-lite/examples/joinaggregate_mean_difference.html

g3o2. (2017). Dot-dash plot. Retrieved from https://bl.ocks.org/g3o2/bd4362574137061c243a2994ba648fb8

Liu, J. (2022, September 22). Homework week 9 ocean issue [Monash University Ed Forum post]. Retrieved from https://edstem.org/au/courses/8853/discussion/1034079?answer=2312029

Waterfall chart of monthly profit and loss. (n.d.). Retrieved from https://vega.github.io/vega-lite/examples/waterfall_chart.html

Week 8 studio activity: Create maps with Vega-Lite. (n.d.). *FIT3179 Week 8 Lab* (pp 19-27). Melbourne, Australia: Monash University.

Appendix

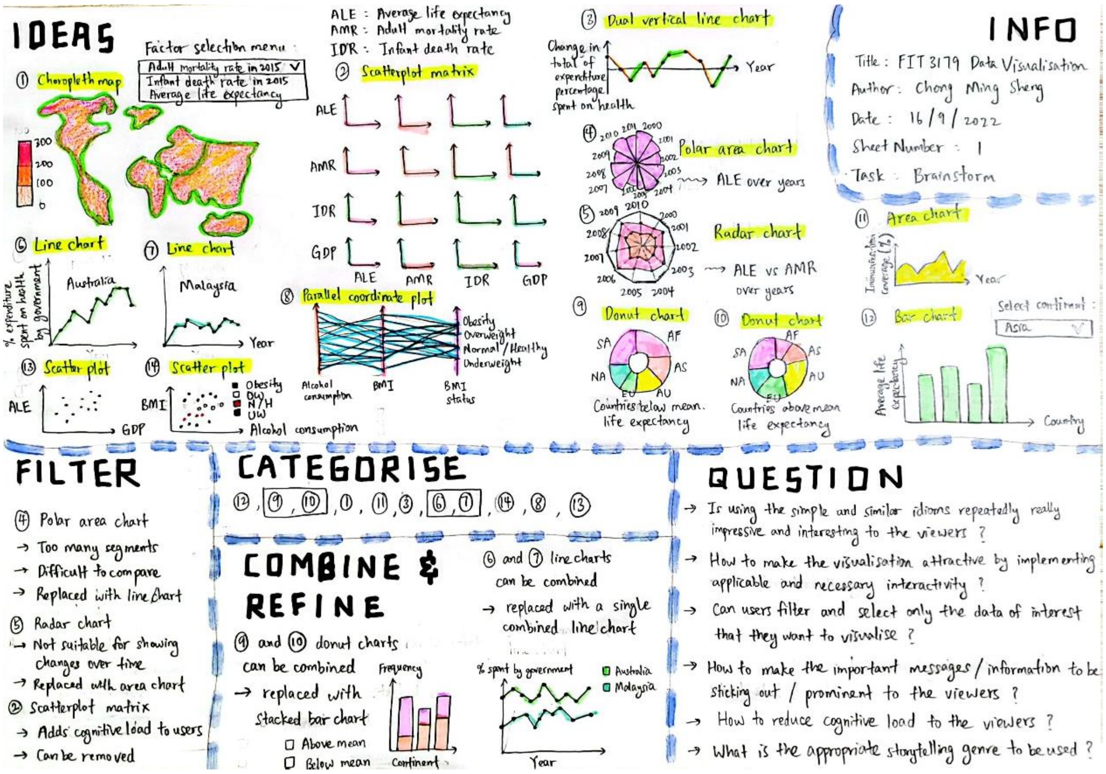


Figure 16: Sheet 1 of 5DS Methodology (Brainstorm).

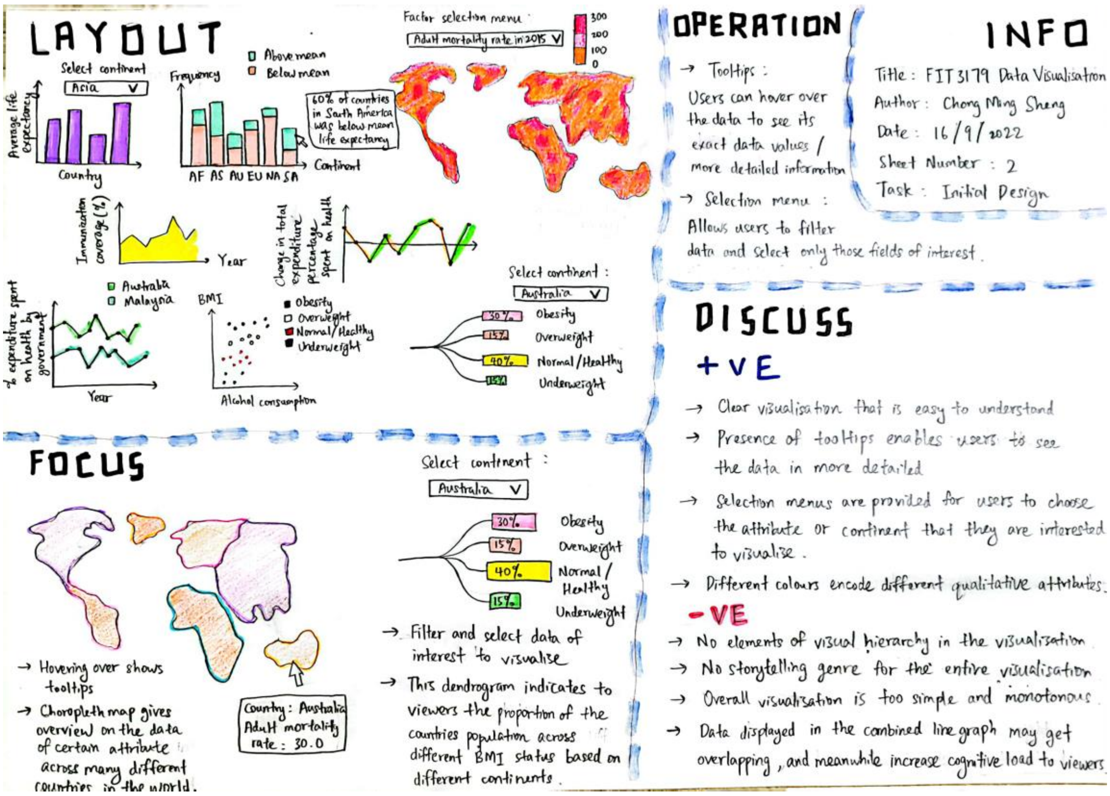


Figure 17: Sheet 2 of 5DS Methodology (Initial design I).

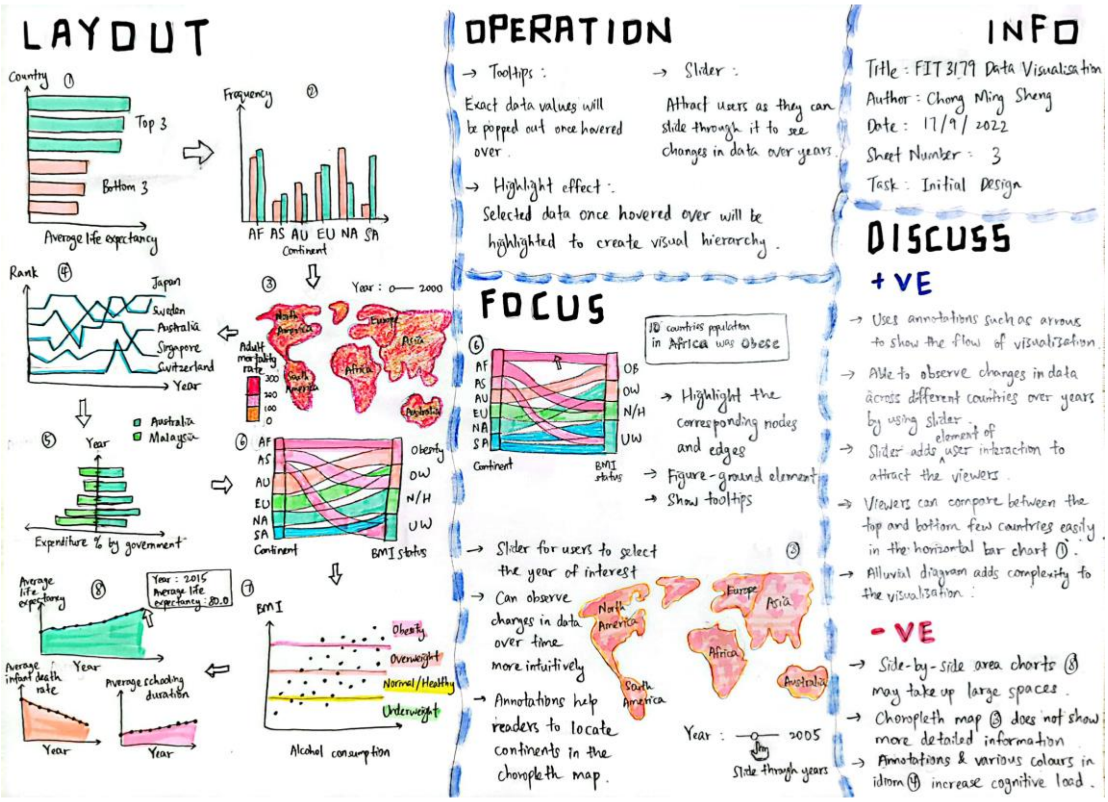


Figure 18: Sheet 3 of 5DS Methodology (Initial design II).

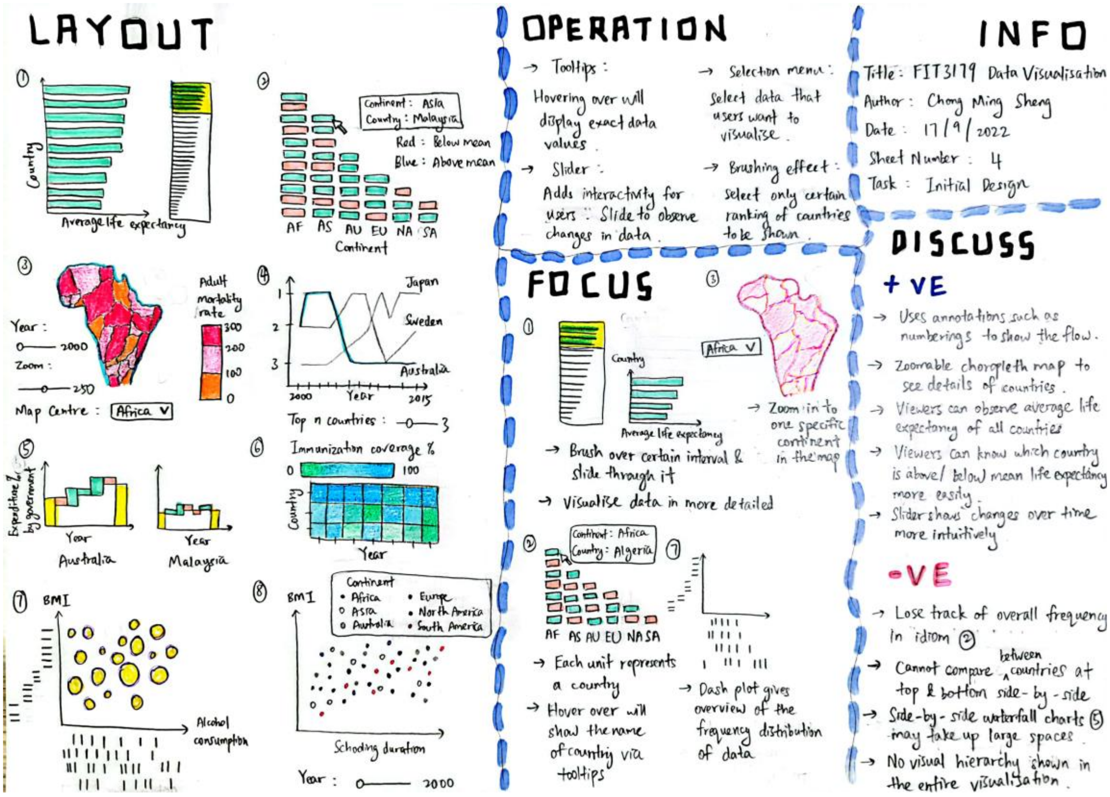


Figure 19: Sheet 4 of 5DS Methodology (Initial design III).

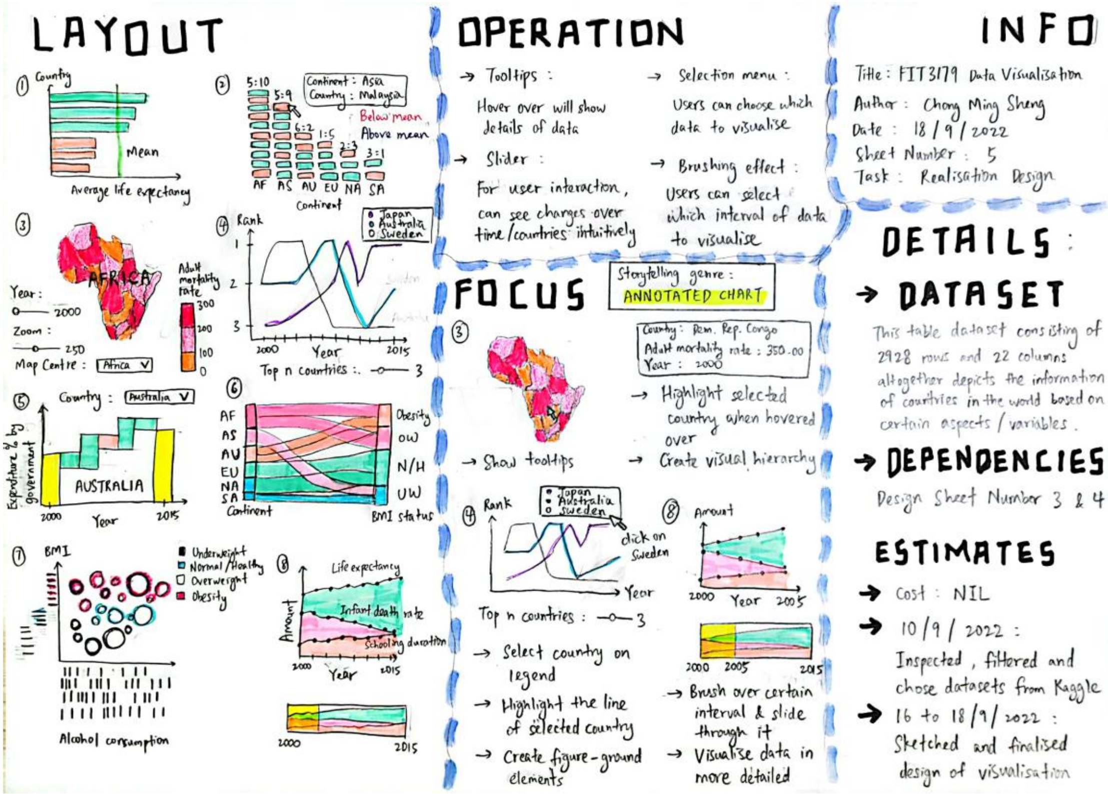


Figure 20: Sheet 5 of 5DS Methodology (Realisation design).