DATA VISUALIZATION

PROJECT

World Alcohol Consumption

Thi Bich Ngoc Hoang - 28496337

Thoa0006@student.monash.edu.au

I. Introduction

Alcoholic drinks are some of the most favourite beverages in the world, an indispensable part of formal events, social affairs and the like. It is observed that nation has different drinking behaviours and a general favourite type of alcoholic beverage, which are heavily influenced by either geographical, economic and/or cultural factors. As a result, the visualisation project plans to aim to people who developed an interest in alcoholic beverages around the world, as well as the general public, who are curious about how culture influences the alcohol usage, and/or the drinking frequency differs around of the world.

By utilising and combining datasets given by renowned organisations such as the World Health Organization and the World Data, the project is aiming to deliver how can different countries having different drinking behaviours, identifying the patterns and explaining the causes and effects of this pattern.

The data file, named "merge.csv" having an approximate of 3000 rows with 14 columns, which integrates from the sources aforementioned. The name, type and description of the columns are as in table 1 below:

Column	Data type	Description
country	string	Country name
code	string	Country code, following the ISO3 standard
year	numeric	Year of data
prod	numeric	Country productivity (unit: USD/person)
gdp_per_cap	numeric	GDP per capita (unit: USD)
life_exp	numeric	Life expectancy (unit: year)
total	numeric	Total alcohol consumption (unit: litre per person per year)
beer	numeric	Total beer consumption (unit: litre per person per year)
wine	numeric	Total wine consumption (unit: litre per person per year)
spirit	numeric	Total spirit consumption (unit: litre per person per year)
other	numeric	Total other alcoholic drink consumption (unit: litre per person per year)
pc_m_dis	numeric	Percentage of males having alcohol disorder (%)
pc_fm_dis	numeric	Percentage of females having alcohol disorder (%)
de_per_mil	numeric	Death from alcohol use disorder (unit: number of death per million people)

Table 1: Metadata of the final dataset

II. Design

To ensure the message from the visualisation can be effectively delivered to target audience, the Five Design Sheet Methodology is applied in the project to ensure the best visualisations are chosen. This section is going to describe how the final data dashboard is brainstormed and designed.

Sheet 1: Brainstorming

Brainstorming ideas are written down in the first sheet, which tries to draw out as much possible ideas as possible. The first sheet comprises of trivial ideas about how each component of the visualisation dashboard can be organised, designed, developed and integrated. The following figure below states the ideas, target messages and audience type, grouping and categorise the ideas based on the message purposes, combining possible solutions.
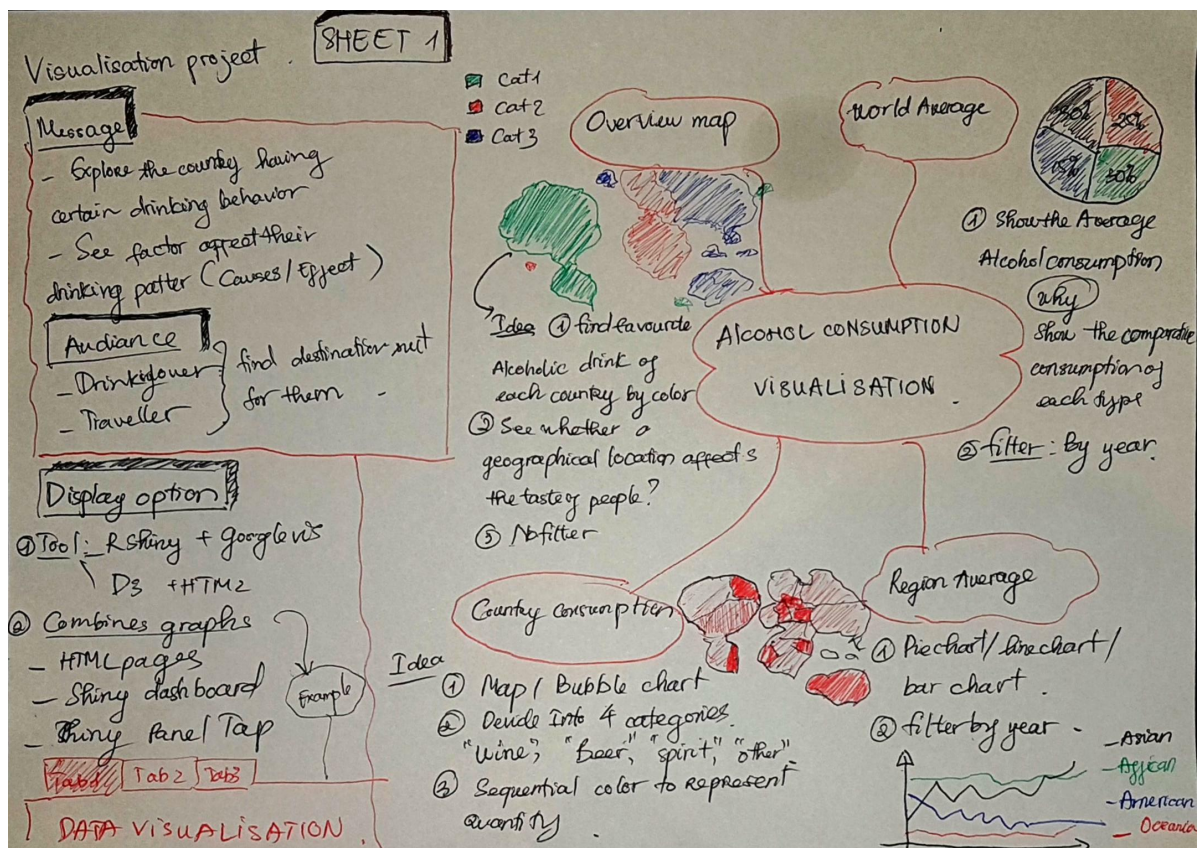

Figure 1: Sheet 1 – Brainstorming idea

Figure 1 illustrates the idea of sketching in the beginning of the project. From the figure, it is noted that the project is kicked off with clear messages and target audiences – telling a story to alcohol and travelling enthusiasts about different drinking patterns of in regions around the world and assist them in finding places which are in line with their favourite alcoholic drink.

The idea is trying to generate a 'flow' which help users understand the message from a general picture (what is favourite alcoholic drink of each country) to a more detailed approach (what is the drinking pattern of countries/regions over a period of time). Finally, it aims at giving them freedom of exploring data by providing interactions, such as filtering and other elements to assist them in seeing the data.

The sheet includes possible options for designing graphs, including choropleth map, pie chart, bubble chart and line chart. Several filter elements also listed such as time, country filter, drink type or region filters. Furthermore, suggestions in terms of visualisation tools/framework are also listed including R Shiny, HTML or D3.

Sheets 2, 3 and 4: Sketching alternative options for visualisations

From the initial idea of the dashboard formed by using sheet 1, the next three sheets are sketching the three most feasible options of the alcohol consumption visualisation dashboard. The following illustrations and explanations emphasise on the highlights of these three sheets.
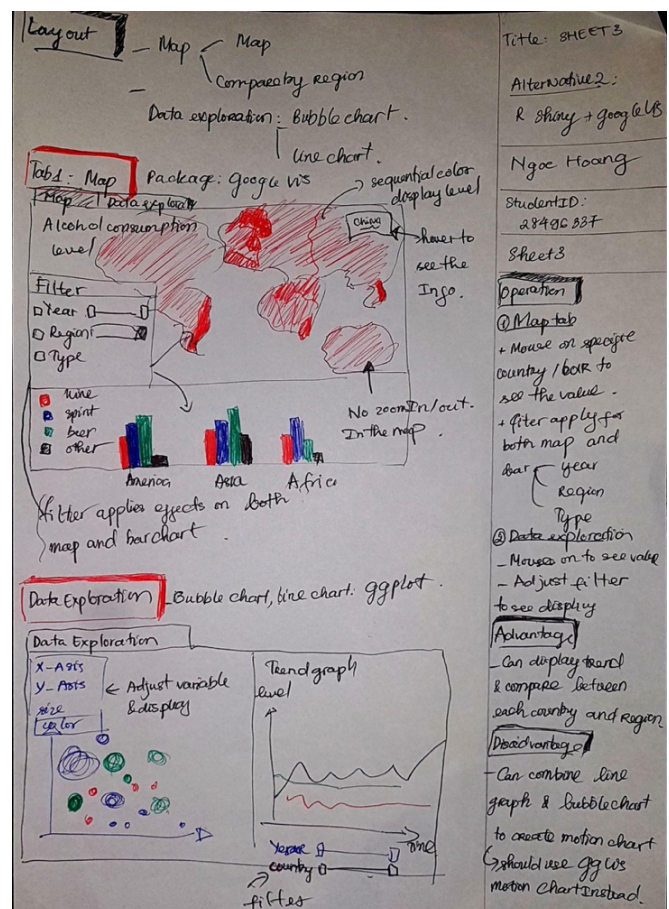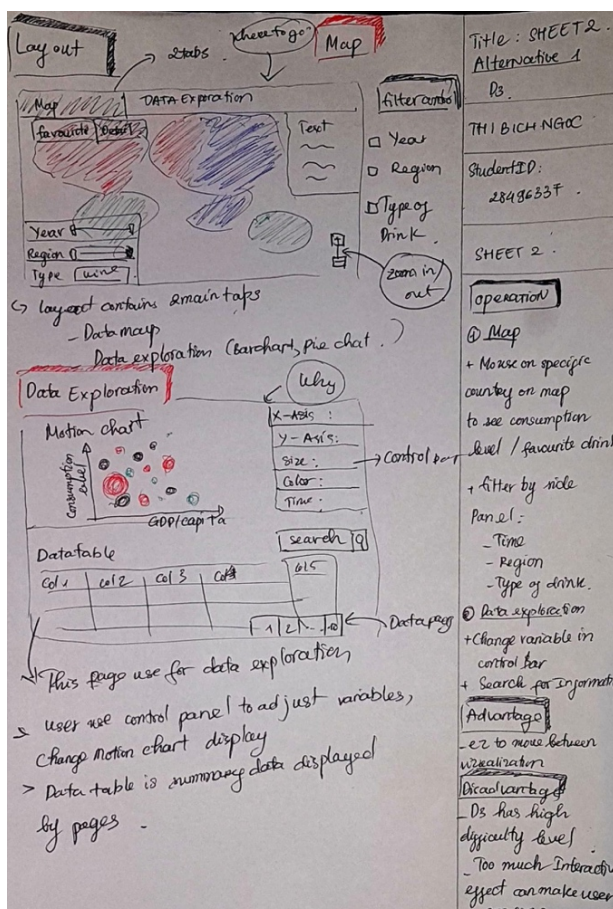
Figures 2 and 3: Sheets 2 and 3 – The first and second alternatives

Each alternative can be summarised as follows:

First alternative (sheet 2)

- **Expected tool:** D3
- **Main layout:** 2 pages. The first page is data map, the second page is data exploration by users.

The first page uses map data created by D3 to visualise the favourite alcoholic drinks of each country, as well as the consumption level of each type of drink. User can switch between 2 views by using navigations at the top left of the page. This page will be controlled by filters (year, type of drink and region). It is noted that D3 map can provide with zooming options.

The second page will contain 2 elements – motion chart and data table. The GoogleVis package can be used to generate motion chart. Users can adjust data by using built-in filters and change display of the motion chart (size, colour, x–axis, y-axis). The table chart is the second element, which tries to give detail information of the data table. User can search for any information by using search buttons.

- **Operations:**

  - Hover action: User can hover over map/motion chart to see the detailed information of this element.
  - Zooming: User can zoom in/out the map to see the general or detail picture of drinking pattern.
  - Filter action: User can change the display of the graph by applying filter by time, regions and type of drink.

- **Advantages:**

  - This design is user-friendly with many interactive effects. Users can easily to move between tabs.
  - D3 can provide scientist the freedom to create any type of chart.
  - The visualization does not provide comparison with the world level or region level.

- **Disadvantages:**

  - D3 requires high level of difficulty.
  - Too much interactive effects sometimes make users confused.

Second alternative (sheet 3)

- **Expected tool:** R

- **Main layout:** 2 pages. The first page is the data map, second page is data exploration by users.

Unlike the first design, the first page will provide the alcohol consumption level. This is due to the limitation of the R package (GoogleVis), which cannot provide choropleth map with categorical data. The map will use the sequential colour to display the level of alcohol consumption. There is no filter in this R map. The filters could be by year, type of drink and region.

The second page contains 2 elements – bubble chart and line chart which is visualized by the GoogleVis package. With bubble chart, users can change the display data by using filters in terms of size, x-axis, y-axis. Line chart provide trend view to user. The filter could be by region or by country.

- **Operations:**

  o Hover action: User can hover on map, line or bubble charts to reveal the detailed information of this element.

  o Filter action: User can change the display of the graph by applying filters by time, regions, type of drink, etc.

- **Advantages:**

  o Easy to build with R packages such as Plotly, GoogleVis, ggplot.

  o This design is simple and user-friendly with few complicated interactions.

- **Disadvantages:**

  o Limitations in zooming.

  o This design cannot provide categorical map in terms of favourite drink.

  o The visualization does not provide comparison with the world level or region level.

Third alternative (sheet 4)

- **Expected tool:** Combination of R, HTML, CSS and JS

- **Main layout:** 3 pages. First page is favourite map, second page is alcohol consumption level and third page is data exploration by users.
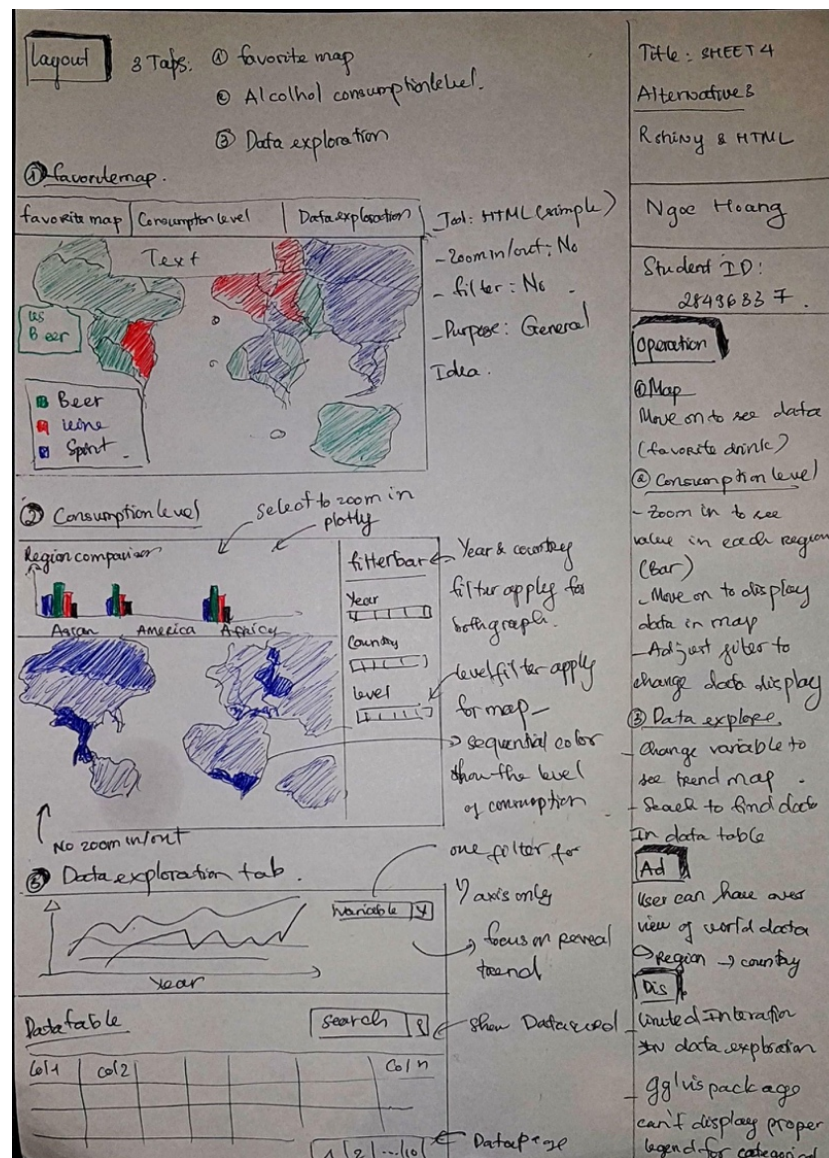
Figure 4: Sheet 4 – The third alternative

Unlike the first two designs, the first page will solely provide the favourite map, which is created by a HTML page with the support of the GoogleVis package. Each colour used in the map represents a different alcoholic drink. It is noted that the source code of GoogleVis does not support categorical data map, hence the colours cannot explain the information. There is no filter or zoom in/out action in this page. Only hover is applied.

The second page is alcohol consumption level map. Similar to the second design, it is a choropleth map with sequential colour indicate the high/low level of consumption. The filters could be by year, type of drink and region.

The third page contains 2 elements – a line chart to illustrates the consumption. Trend and a data table for detailed information. Users can use the search action to find a specific data

- **Operations:**

  - Hover action: User can hover on map (both favourite map and consumption level map), line chart or bubble chart to reveal the detailed information of this element.

  - Filter action: User can change the display of the graph by applying filters by time, regions, type of drink, etc.

- **Advantages:**

  - User can experience a journey from general picture with favourite map, then go into details with consumption level map.

  - This design is simple and user-friendly with not many complicated interactions.

- **Disadvantages:**

  - Limitation in zoom in/out action

  - Limitation with GoogleVis package, the colour legend for favourite map cannot represent the data legend properly.

  - The visualization does not provide comparison with the world level or region level

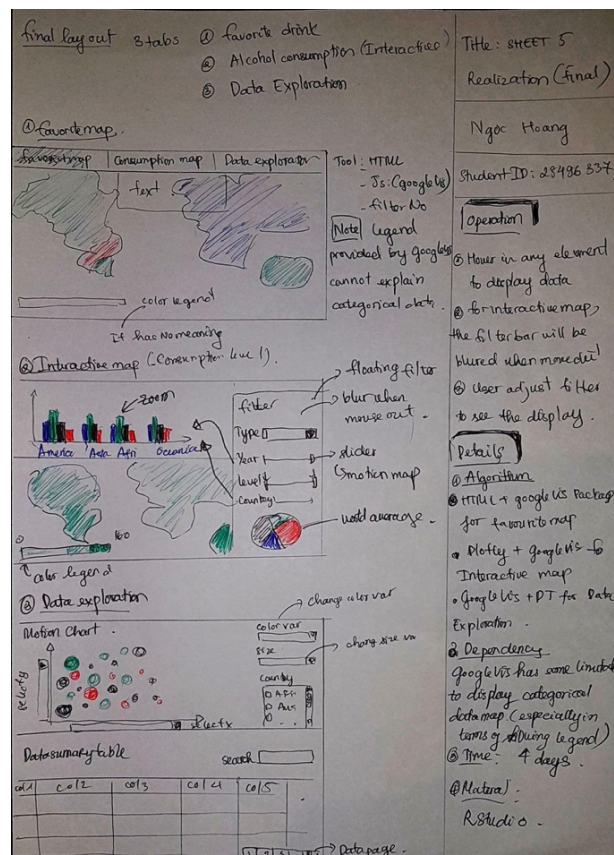## Sheet 5: Realization (Final Design)

Figure 5: Sheet 5 – Final design

After comparing and combining sheets 2, 3 and 4, the final design for the data is illustrated in sheet 5.

- **Expected tool:** Combination of R, HTML, CSS and JS

- **Main layout:** 3 pages. The first page is favourite map, the second page is alcohol consumption level and third page is the data exploration page.

    o **First page:** Favourite map - introduction

The first page can be considered as the introduction of the story. This is created by combining HTML, JavaScript with the integration of the GoogleVis package. It displays the favourite drink in each country by using 4 different colours. User can hover on the map to see the favourite drink of each country.

    o **Second page:** Alcohol consumption – messages in a country level

The first part is bar graph, providing an average region alcohol consumption level, created by using the Plotly package. It provides zooming actions and hover action. The second part is choropleth map for alcohol consumption level. The sequential colour indicates the level of consumption. It allows hovering action. The third part is a pie chart showing the average world consumption level. This is attached with the filter panel.

A floating filter panel is created for this design, in which the year filter can be applied for 3 charts, the country filter can be applied for bar chart and map, and the drink type filter can be applied for the map. The filter panel is designed with the support of a CSS file, which allows adjusting the opacity of the panel. For example, when the cursor is out, the opacity of the panel will be reduced, which help users to focus on other elements.

    o **Third page:** Data exploration – Tool for self-exploration

The last page is the most detailed page, which gives users freedom to explore data. It is not limited with only the alcohol consumption level. Users can freely choose any element to visualise in the motion chart to explore some possible cause and effect of the certain alcohol consumption level.

Another part of the page is the data table. User can use search button to find any data.

- **Choice of colour:** Apply the same colour when visualize a certain drink type.

- **Operations:**

  o Hover action: User can hover on map, line chart or bubble chart to reveal the detailed information of this element.

  o Filter action: User can change the display of the graph by applying filters by time, regions, type of drink, etc.

  o Search for data using search button

- **Dependencies:**

  o It is still limited by the original display of packages (GoogleVis). Therefore, the colour legend for the favourite map still cannot displayed properly.

  o Environment: R shiny

- **Expected time to complete:** 4-5 days.

III.   Implementation

First page:

The first page is the favourite alcoholic beverage map. The reason this display is chosen is trying to give the overview of the favourite drink by nations to user before giving them more details messages.

Because the R package GoogleVis does not support the categorical map, an HTML page is used to create the first page. By using JavaScript to link the favourite data by nation (which is calculated by R after considering the timestamp from 2000 to 2013), we can have the following view for first page.
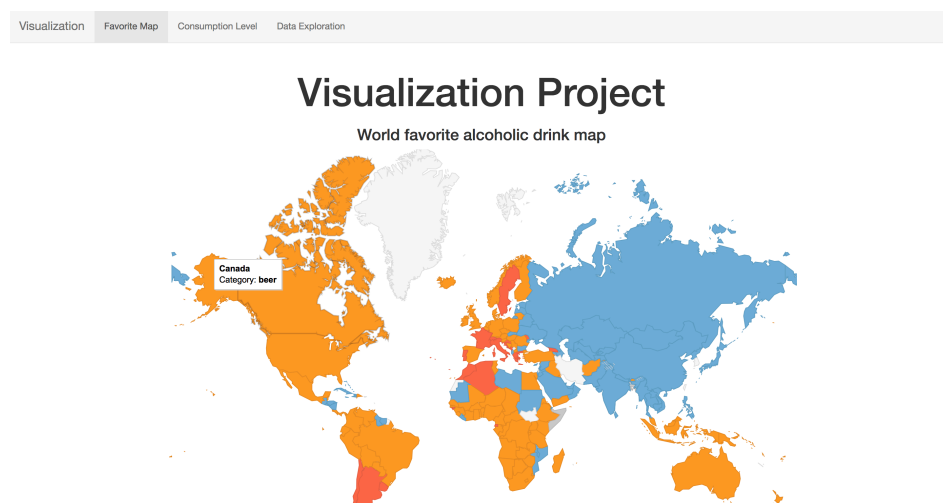

Figure 6: First page – Favourite Map

Second page:

The second page combines bar region consumption graph, a pie world consumption map and an alcohol consumption level map. The reason this display is chosen is trying to give the comparison between countries, regions and world average level.

The bar chart is created by using Plotly package while GoogleVis packages applies for the interactive Map and the pie chart.

The panel with filters is connected with each element by using R code. Year filter can affect 3 charts while country filter applies on both bar chart and map, and the type filter apply for the map only.

A CSS file is used to add more effects on the filter panel. In fact, when user mouse in or mouse out of the filter panel, the opacity will be changed, which allows user focus on other elements after adjusting variables.
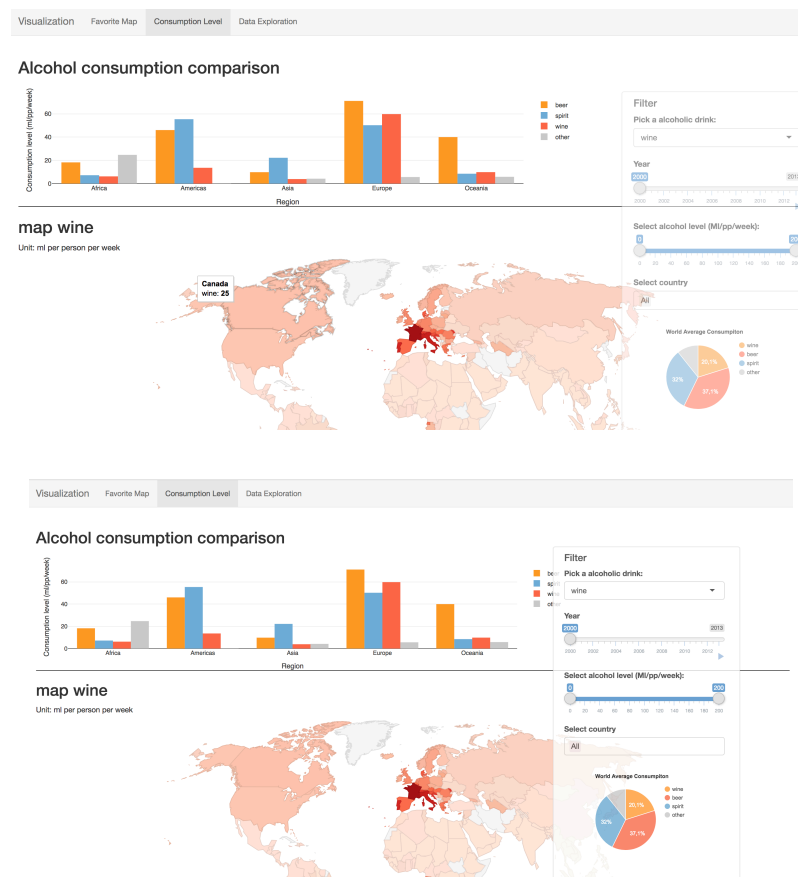

Figure 7: Page 2 – when mouse in (top picture) and move out (bottom picture) of filter

The final page gives the most detailed data level by using Motion chart and data table. The reason is there are a lot of causes and effects related to the alcohol consumption. Causes may be the GDP (economic factor) or effects can be the life expectancy level or the number of death caused by alcohol. Therefore, motion chart can be a powerful tool to let user explore the data by themselves.

The motion chart is created by using GoogleVis package in R, and the data table created by using DT package. With the built-in design of 2 packages, user can adjust the size, colour, x-axis variable, y-axis variable in motion chart as well as search for any data they need.

I.   User Guide

Please run the App externally to see all the visualization!!!
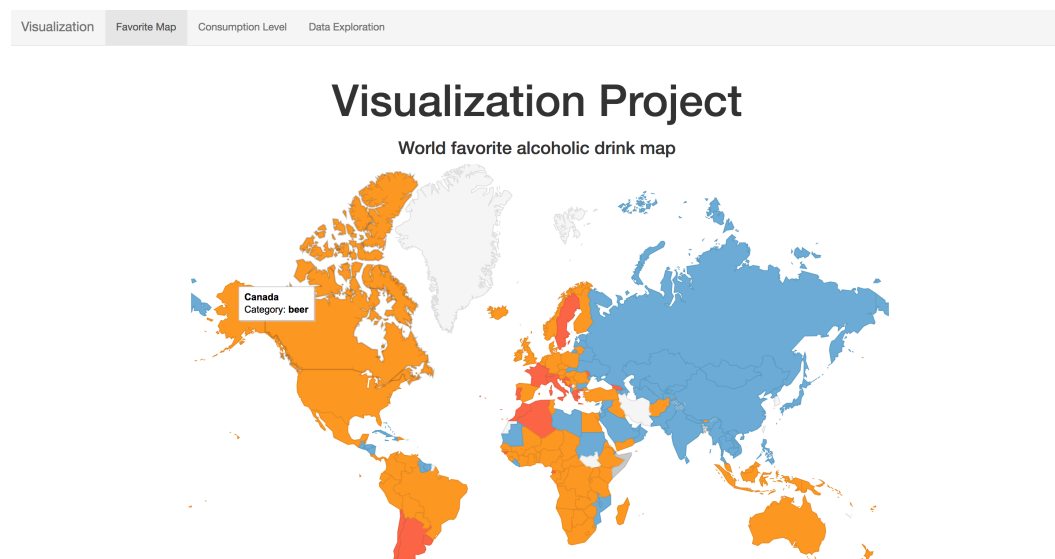
Favourite drinks map

Figure 8: First page – Favourite Map

To explore the page, user just needs to focus on hover action. When hover on the chart, the tooltip will pop up with name of the country and the favourite alcoholic drink over the period 2000 – 2013.
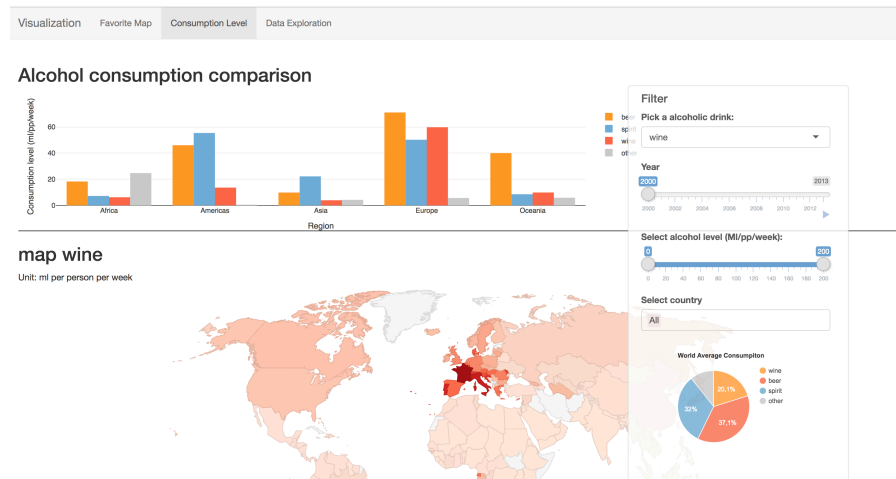
Consumption level page

Figure 9: Second page – Alcohol consumption

Adjust filter:

The filter bar can be blur or focus by changing the opacity. It can be moved to any places of the page by dragging the panel, which can help user to see the data display more clearly.

- Firstly, user can choose a drink of interest by select from the dropdown list.
- Change the year slider to see the data change over each year. This filter applies for bar chart, pie chart and the map
- User can choose any country or select the whole world map by using the select country filter. If user does not choose the "All" option, the list of countries chosen will be highlighted on the map, and the bar graph will display the regions where these countries belong to.
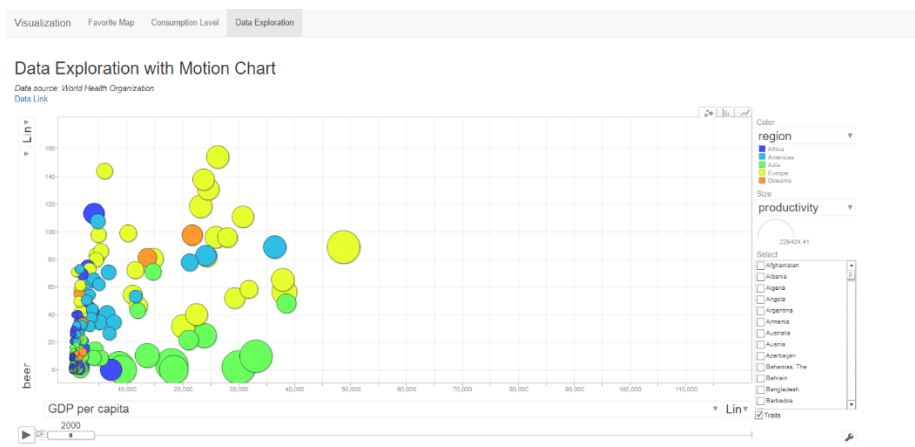
Data Exploration page

Figure 10: Motion Chart

To explore the possible causes and effect of the different alcohol consumption, user can select and adjust any element of the motion chart.

- Pick up the colour: using the top right dropdown colour variable

- Pick up the size: using the dropdown size variable

- Highlight and track the change of a country by ticking in the check box on the right hand side.

- Select x-axis and y-axis variable: click to the arrow symbol on the x-axis and y-axis to change the variable

- Click the play button to see the change in data

Data Summary Table

Data source: World Health Organization
Data Link
Show 10 entries   Search:

	code	country	year	productivity	GDP per capita	Life expectancy	total	beer	spirit	wine	other	favorite	Alcohol disorder % in Males	Alcohol disorder % in Females	Death caused by alcohol (per million people)	region
2	AFG	Afghanistan	2008	5253.99	373.36	60.243	0.58	0.19	0	0.38	0	0.75	0.45	7.93	Asia	spirit
3	AFG	Afghanistan	2005	4465.17	250.29	58.5	0.38	0.19	0	0.19	0	0.74	0.45	8	Asia	other
4	AFG	Afghanistan	2007	5105.54	369.84	59.694	0.38	0.19	0	0	0.19	0.75	0.45	7.96	Asia	other
5	AFG	Afghanistan	2006	4634.25	272.56	59.11	0.58	0.19	0	0.19	0.19	0.75	0.45	7.99	Asia	other
6	AFG	Afghanistan	2013	6947.21	638.61	62.494	0.19	0.19	0	0	0	0.76	0.45	7.74	Asia	beer
7	AFG	Afghanistan	2002	4314.05	187.85	56.637	0.19	0	0	0	0	0.74	0.45	7.88	Asia	other
8	AFG	Afghanistan	2004	4284.07	219.14	57.875	0.38	0.38	0	0	0	0.74	0.45	7.97	Asia	beer
9	AFG	Afghanistan	2012	7163.71	669.01	62.086	0.19	0.19	0	0	0	0.76	0.45	7.75	Asia	beer
10	AFG	Afghanistan	2001	2836.92	117.41	56.044	0	0	0	0	0	0.74	0.45	7.85	Asia	other
11	AFG	Afghanistan	2009	6087.76	445.89	60.754	0.19	0.19	0	0	0	0.75	0.45	7.88	Asia	beer

Showing 1 to 10 of 2,400 entries   Previous 1 2 3 4 5 … 240 Next

Figure 11: Data Summary Table

To use the data table, user can type the information in the search bar. Users can also change the display option by adjusting "show entries" option on the left-hand side.

II.   Conclusion

After the visualization project, a dashboard is built to tell a story of alcohol consumption around the world to alcoholic and travelling enthusiasts. From this project, I learnt how to integrate R Shiny with CSS, HTML and JavaScript to create interactive visualization.

However, due to time limit, there are some limitations which need to improve for further development. One of the hind sight is the colour legend of the categorical map (favourite map). This legend can confuse the users as it does not tell them properly the meaning of the colour used in the map. It can be fixed by using D3 and JavaScript to create customize colour legend and tooltips. In addition, the layout of the pages, in some extent, is not optimal. It could be still improved if more CSS

styles are applied. These limitations could be solved by practicing how to use CSS, JavaScript and D3 in the future.

References

Global Health Data Exchange. (n.d.). *GBD Results TRool*. Retrieved from Global Health Data Exchange: http://ghdx.healthdata.org/gbd-results-tool

JSFiddle. (n.d.). Retrieved from JSFiddle: https://jsfiddle.net/LLknt4cc/1/

Ritchie, H., & Roser, M. (2018, April). *Alcohol consumption*. Retrieved from Our World in Data: https://ourworldindata.org/alcohol-consumption