

CARDIFF UNIVERSITY

CMT218 DATA VISUALISATION

Assessment 2 - Data Analysis and Visualisation Creation

C22107414

8 May 2023

Contents

1	Introduction	3
2	Global Alcohol Consumption Trends 2000-2018	3
3	GDP Per Capita and Alcohol Consumption by Country 2000-2018	4
	References	5

1 Introduction

I am writing an online article for [The Guardian Datablog \(2023\)](#) on alcohol consumption trends around the world, for which I need to produce two visualisations based on Kaggle’s “Alcohol Consumption by Country” dataset. According to [Rogers \(2012\)](#), the main goal of The Guardian Datablog is to help the readers understand the data behind news stories. They aim to achieve this by backing up news stories with data and various analyses of it. I am going to implement the same approach and equip my news article with two informative visualisations aimed at The Guardian Datablog’s target audience, which is the general public in the UK and English-speaking audiences worldwide.

2 Global Alcohol Consumption Trends 2000-2018

The first visualisation is the “Global Alcohol Consumption Trends 2000-2018” dynamic choropleth map depicting how alcohol consumption across the world evolved throughout the years 2000, 2005, 2010, 2015 and 2018. The screenshot (figure 1) of the dynamic map is shown below and to access the full version showing alcohol consumption for all years, please click on the following hyperlink <http://www.ggsoroka.com/alc.gif>.

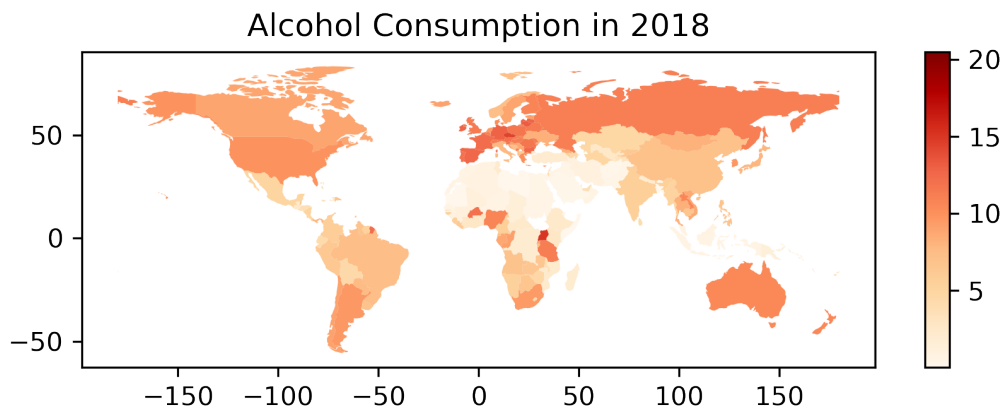


Figure 1: A screenshot of the “Global Alcohol Consumption Trends 2000-2018” dynamic choropleth map showing the year 2018 only. To access the dynamic visualisation, please click on the following hyperlink <http://www.ggsoroka.com/alc.gif>.

The “Global Alcohol Consumption Trends 2000-2018” dynamic map is effective because:

To start with, the visualisation is user-friendly and engaging due to its animated nature and is, therefore, an appropriate choice for the introductory visualisation in the news article. It is eye-catching and attracts the reader’s attention, and as argued by [Lupi \(2018\)](#), aesthetically pleasing visualisations might trigger the viewer’s curiosity about the topic, which can potentially result in more people viewing the visualisation thus leading to the message being spread to a wider audience.

This visualisation effectively communicates the changing alcohol consumption patterns worldwide over the years, using high-contrast colours, such as red and orange, on a continuous colour scale. This approach helps encode the data, making it easier to read and also accessible for readers with colour blindness ([Junk Charts, 2015](#)). When creating data visualisations, it is crucial to consider the needs of the target audience. In the UK, for example, approximately 4.5% of the population is affected by this condition ([Colour Blind Awareness, 2023](#)).

Next, this visualisation provides a general idea of alcohol consumption rates by country without overwhelming the reader. The Guardian Datablog’s target audience is the general public, who may not have the expertise or time to engage with complex graphics and, therefore, the level of complexity of this visualisation is appropriate. However, some of the readers might find it lacking in detail, and I aim

to cover the needs of such readers with a second visualisation. This is an example of why data visualisation is subjective in nature and that selecting an appropriate level of complexity for a visualisation depends on the intended audience (Cairo, 2013).

From a technical standpoint, this choropleth map animation efficiently utilises space by embedding five different graphics into one.

The “Global Alcohol Consumption Trends 2000-2018” map has a number of limitations:

Small regions, such as Slovenia and other European countries, can be difficult to discern due to their size on the map. To address this, a higher-resolution choropleth map could be generated and made available for readers. Additionally, providing a more advanced zoom option would allow readers to view smaller areas in greater detail.

In addition, the alcohol consumption scale on the right of the map is not labelled. To ensure that readers can accurately interpret the visualisation, clarification on that the scale represents litres per capita will have to be included in the article

Finally, some countries, such as Canada, were missing alcohol consumption data for particular years, which resulted in a graphical inconsistency within the visualisation. To address this issue, we could look for other alcohol consumption datasets and extract the required data from them.

3 GDP Per Capita and Alcohol Consumption by Country 2000-2018

Further, I have produced the “GDP per capita and Alcohol Consumption by Country” interactive bubble plot depicting the relationship between the GDP per capita and alcohol consumption in different countries in the years 2000, 2005, 2010, 2015 and 2018. The screenshot of the interactive bubble plot is shown below and to access the full interactive version, please click on the following hyperlink <http://www.ggsoroka.com/visualisation.html>.

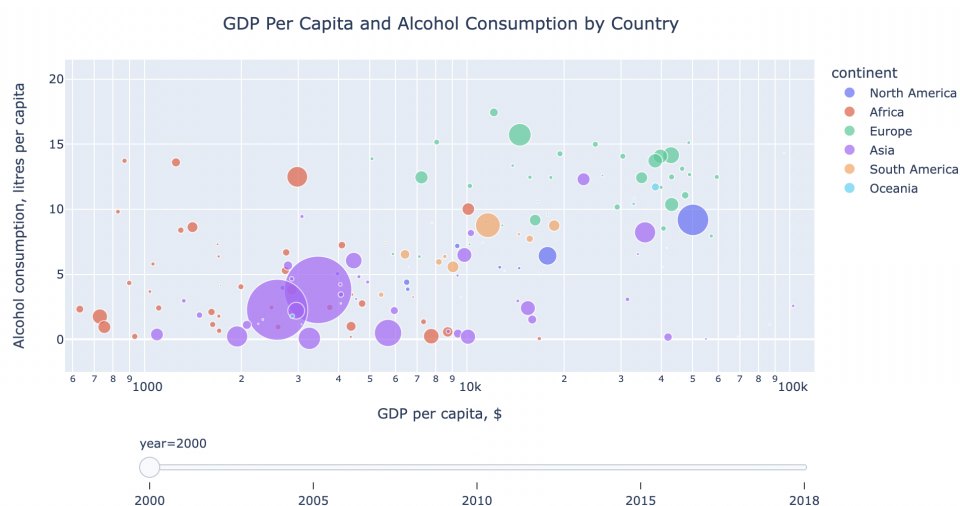


Figure 2: A screenshot of the “GDP per capita and Alcohol Consumption by Country” interactive bubble plot showing the year 2018 only. To access the full interactive visualisation, please click on the following hyperlink <http://www.ggsoroka.com/visualisation.html>.

The “GDP per capita and Alcohol Consumption by Country” plot is effective because:

First, the interactive element of the plot gives the reader control over certain elements of the visualisation, for instance, the reader can select any combination of continents to be displayed to get a better insight into the changes in alcohol consumption between different continents. Not only it can result in

a better understanding of alcohol consumption tendencies but also lead to higher rates of engagement with the visualisation.

The bubbles in the plot do not overlap because the larger bubbles are positioned in the background, making the smaller bubbles clearly visible and effortlessly accessible.

Additionally, the countries and continents are easily distinguishable due to the effective use of colours, which include blue, purple, orange, green, red and cyan. Colours can be used to emphasise significant patterns in the data, allowing the reader to swiftly grasp the information being presented (Evergreen and Emery, 2016). However, this colour palette might not be suitable for readers with colour blindness and could be improved by using colours that are more appropriate for people with such visual impairment.

The “GDP per capita and Alcohol Consumption by Country” plot’s limitation is:

If The Guardian decides to also publish this article in the print version, the visualisation will be reduced to its static form. Without the interactive component, several features, such as clustering alcohol consumption by continent, will be lost. Furthermore, if the print publication is going to be black-and-white, the colours will practically be indistinguishable. This will hinder the extent to which the message on alcohol consumption by country is conveyed to the target audience.

A major limitation applicable to both visualisations is that they only reflect the “Alcohol Consumption by Country” data and do not perform any advanced statistical analysis like regression, cluster analysis or statistical significance tests. This data could be further mined to explore potential relationships between the variables. Another limitation of both visualisations is that they do not inform the reader about the sample used in the dataset, which comprises individuals aged over 15 years old. This sample information is to be included in the article.

Overall, it is worth noting that I have made use of all the available data from the “Alcohol Consumption by Country” dataset when creating the visualisations. All dataset variables, including “Entity” (countries), “Code” (country codes), “Year”, “Total alcohol consumption”, “GDP per capita”, “Population” and “Continent”, were used.

References

- Cairo, A. (2013). *The Functional Art: An Introduction to Information Graphics and Visualization*. Voices That Matter Series. New Riders.
- Colour Blind Awareness (2023). About colour blindness. Online. <https://www.colourblindawareness.org/colour-blindness> Accessed 06 May 2023.
- Evergreen, S. and Emery, A. K. (2016). Data visualization checklist. Online. <https://stephanieevergreen.com/data-visualization-checklist> Accessed 05 May 2023.
- Junk Charts (2015). A data visualization that is invariant to the data. Online. https://junkcharts.typepad.com/junk_charts/2015/10/a-data-visualization-that-is-invariant-to-the-data.html Accessed 06 May 2023.
- Lupi, G. (2018). Beautiful reasons. Online. <https://medium.com/accurat-in-sight/beautiful-reasons-c1c6926ab7d7> Accessed 02 May 2023.
- Rogers, S. (2012). Behind the scenes at the guardian datablog. In Bounegru, L., Chambers, L., and Gray, J., editors, *The Data Journalism Handbook*. O’Reilly Media, Inc. <https://datajournalism.com/read/handbook/one/in-the-newsroom/behind-the-scenes-at-the-guardian-datablog> Accessed 05 May 2023.
- The Guardian Datablog (2023). The Guardian. Online. <https://www.theguardian.com/data> Accessed 05 May 2023.