Sinéad Dunne

s\_dunne1@hotmail.com

UCD Certificate in Introductory Data Analytics

A picture containing clipart

Description automatically generated

global consumption

for wine, beer & spirits

Project Report

# GitHub URL: <https://github.com/sdunne1/UCDPA_SineadDunne>

# Abstract

My data analysis project investigates data figures published in 2010 by the Global Information System on Alcohol and Health (WHO GISAH), and manipulated by Mona Chalabi (data editor at Guardian U.S.), as part of her 2014 article for website FiveThirtyEight: “Where Do People Drink the Most Beer, Wine and Spirits?”

WHO originally published this data, in total litres of pure alcohol consumed per person in each country. However, for the purpose of this article, the average alcohol content and average serving size for each beverage was taken and those numbers converted into standard serving sizes. The data produced shows units of wine by glass, units of beer by cans and units of spirits by shots, all consumed per person in each listed country. This data was retrieved from Kaggle as referenced below and is the dataset I have used for my data analysis report.

# Introduction

Ireland as a small country on the outskirts of Europe is widely renowned as being the top country in the world, if not one of them, for the highest level of alcohol consumption. I wanted to explore alcohol stereotypes, using real data figures published. The data description is outlined in the Abstract of this report.

There are some limitations with this dataset that should be highlighted. As it was published over ten years ago, it can give us good insights. However, they will not be the most up to date. The dataset only deals with three alcohol types: Wine, Beer and Spirits. The dataset only contains 193 countries of the total 195 in the world. For some countries listed the serving size figures are set to 0. These are mostly for countries with strong religious influences. However, I believe there to be an inaccuracy for these rows, as there may still be a small minority in these countries who consume alcohol. In which case, these figures will not be represented.

With more time, I would have liked to extend my project to explore the latest figures available from Global Information System on Alcohol and Health (GISAH). This would allow me to demonstrate more current and up to date global insights. It would also facilitate data visualisations of a ten-year comparison. Demonstrating increases and reductions in total litres of alcohol, and the types of alcohol by country / continent, linking in possible trends that might explain them. For example, campaigns to increase sales for alcohol types or brands.

# Dataset

The data is a csv file which contains 5 column headers: Country, Beer\_Servings, Wine\_Servings, Spirit\_Servings, total\_litres\_of\_pure\_alcohol, and 193 rows of data. The source of the data is [www.kaggle.com](http://www.kaggle.com), and the target URL is listed in ‘References’ of this report.

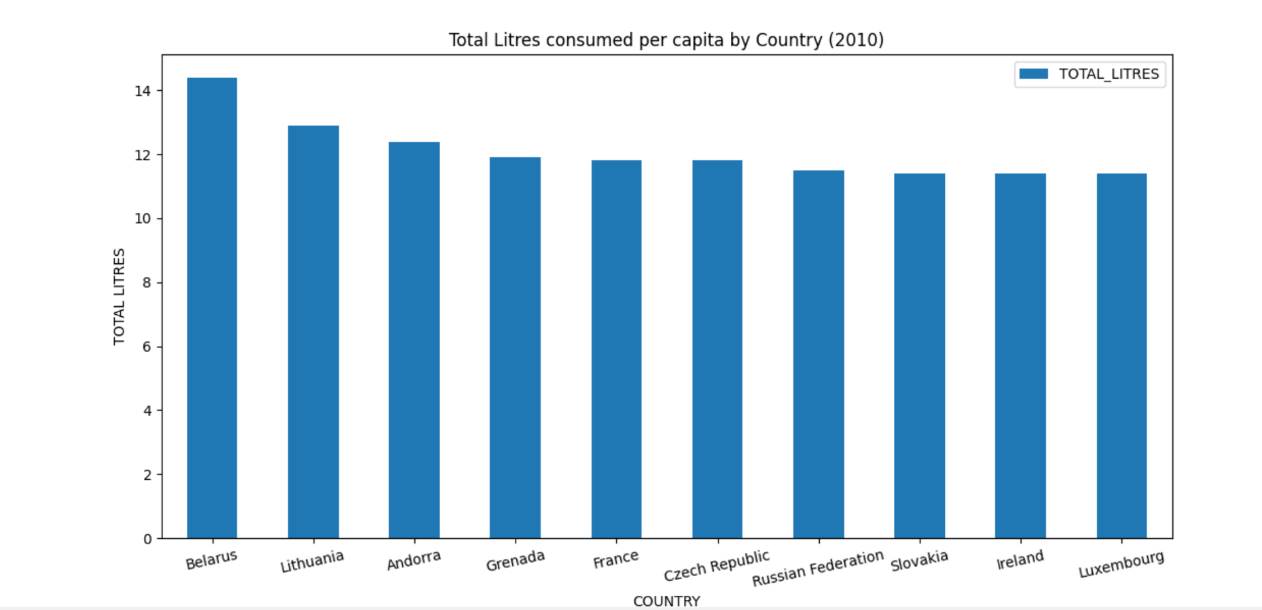
I chose this dataset as it was in line with the area that I wanted to review. It was also a good size for me to work with, being new to data analytics. The dataset contained several data types - categorical and numerical, which I could use to make good data visualisations. Further, I was able to build on the data provided by creating my own dataset to include ‘Continent’. This was done by merging an additional dataframe and allowed me to aggregate the data by continent.

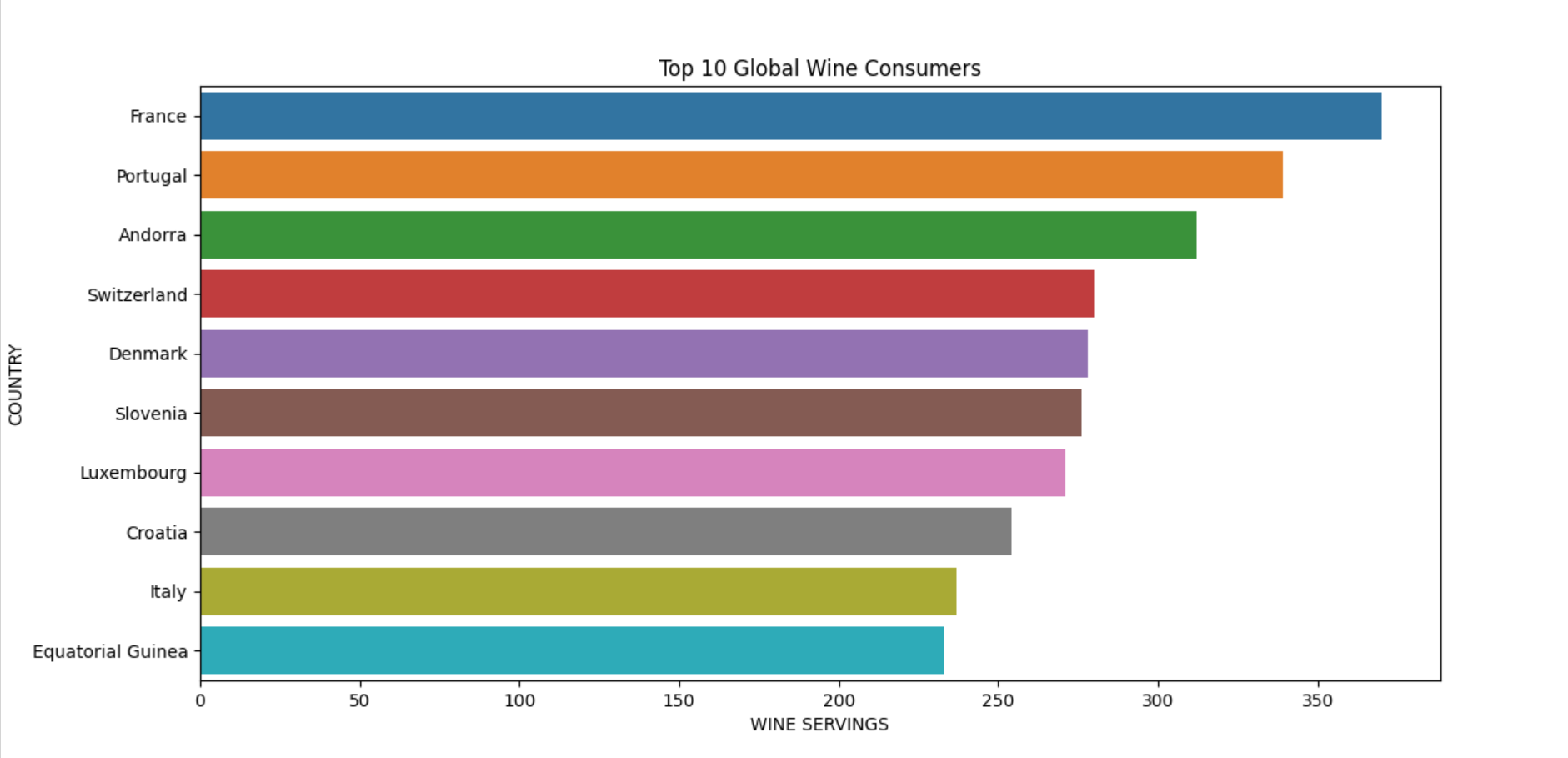
# Implementation Process

<<Describe your entire process in detail>>

# Results

<<Include the charts and describe them>>





# Insights

<<Point out at least 5 insights in bullet points>>

1. France stereotype of wine
2. Russian and eastern European stereotype of spirit drinkers
3. Ireland biggest drinkers in the world? Population versus the highest level – find these and make some notes?
4. Beer drinkers ?
5. Type of alcohol mostly consumed by highest total litres of alcohol.
6. Something from the scatterplot diagram

# References

|  |  |
| --- | --- |
| **KAGGLE** | <https://www.kaggle.com/codebreaker619/alcohol-comsumption-around-the-world> |
| **WORLD HEALTH ORGANISATION** | https://www.who.int/data/gho/data/indicators/indicator-details/GHO/alcohol-recorded-per-capita-(15-)-consumption-(in-litres-of-pure-alcohol) |
| **DATACAMP** | <https://learn.datacamp.com/custom-tracks/custom-certificate-in-introductory-data-analytics-6859d143-be7b-49e0-ad09-007f207ad33c> |
| **FIVETHIRTYEIGHT** | https://fivethirtyeight.com/features/dear-mona-followup-where-do-people-drink-the-most-beer-wine-and-spirits/ |
| **PYTHON DOCS** | <https://docs.python.org/3/> |
| **PANDAS DOCS** | <https://pandas.pydata.org/docs/> |
| **SHANELYNN.IE** | https://www.shanelynn.ie |
| **STACKOVERFLOW** | https://stackoverflow.com/questions |

Visualisations:

* Top 10 countries for total volume (completed)
* Most popular of each drink in each continent; of the three listed
* Scatter plot for sum of servings for each alcohol by continent

To Do:

* LOC
* LISTS
* Looping / iterrows
* Complete report