

Data Analysis & Visualization of Alcohol Consumption in The world

The data consists of 5 variables Country, beer_servings, sprite_servings, wine_servings and Total liters of pure alcohol consumed. The objective of the study was to find out which countries have the highest alcohol, wines and sprites consumption. Further analysis is detailed below

Introduction

Research scientists utilise exploratory data analysis to examine and explore data sets and describe their key properties, frequently with the use of data visualisation techniques. It assists data scientists in identifying patterns, detecting anomalies, testing hypotheses, and validating assumptions by discovering how to effectively modify data sources to retrieve the information they need. It also assists scientists that ensuring that the data they provide are accurate and suitable to any targeted commercial consequences and objectives (Tukey, 1977). Stakeholders are also aided by verifying that they are asking the appropriate questions. Standard deviations, categorical variables, and confidence intervals may all be calculated using EDA. When EDA is finished and insights have been gained, its features, such as machine learning, may be employed for more advanced data analysis or modelling.

Software: Python statistical software(Pandas module) was used to perform the explanatory data analysis where modules such as Pandas, Matplotlib, Seaborn among others were used.

Data: Alcohol consumption dataset was used and it was obtained from the attached link below [Dear Mona Followup: Where Do People Drink The Most Beer, Wine, And Sprites? | FiveThirtyEight](#)

The key objective of the study was to find out in which countries of the world there are high levels of beer, wine, and sprites consumption in the world.

EXPLORATORY DATA ANALYSIS

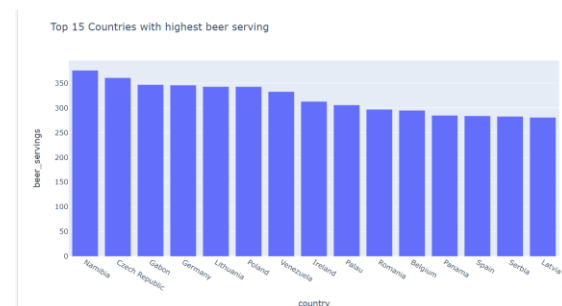
The dataset contains five variables and 193 observations that include Country, Beer_servings, sprite_servings, wines_servings, and total_liters_of pure alcohol.

SUMMARY STATISTICS

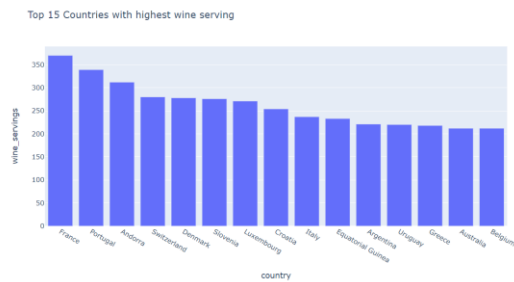
| | beer_servings | spirit_servings | wine_servings | total_liters_of_pure_alcohol |
|-------|---------------|-----------------|---------------|------------------------------|
| count | 193.000000 | 193.000000 | 193.000000 | 193.000000 |
| mean | 106.160622 | 80.994819 | 49.450777 | 4.717098 |
| std | 101.143103 | 88.284312 | 79.697598 | 3.773298 |
| min | 0.000000 | 0.000000 | 0.000000 | 0.000000 |
| 25% | 20.000000 | 4.000000 | 1.000000 | 1.300000 |
| 50% | 76.000000 | 56.000000 | 8.000000 | 4.200000 |
| 75% | 188.000000 | 128.000000 | 59.000000 | 7.200000 |
| max | 376.000000 | 438.000000 | 370.000000 | 14.400000 |

The summary statistics for the four continuous variables is as per the attached table which shows that the mean beer servings are 106 with a standard deviation of 101. The minimum beer serving is Nil with a maximum being 376 servings. The average sprite servings are 80 with a standard deviation of 88. Finally, the average wine serving is 49 with a standard deviation of 79.69. From the summary statistics table, we can conclude that the global consumption of beer is high compared to sprites and wines.

To prepare our data for analysis we checked for missing values in our dataset which confirmed that there are no missing values in the dataset. Also checked was the existence of duplicated values in the dataset but there was none.



The figure above shows top 15 countries in beer consumptions. As expected Namibia tops the list followed by Czech Republic.



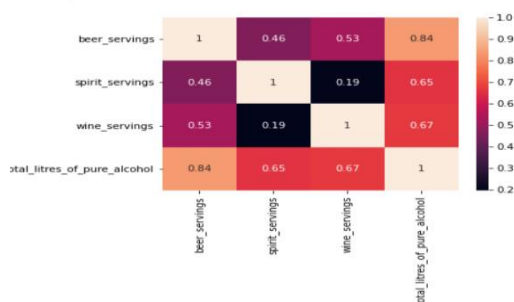
Also the top 15 countries with highest wines consumption shows France tops the list followed by Portugal.

CORRELATION ANALYSIS

The correlation between beer servings and wine servings is 52.7% while that of beer servings and spirit servings is 45.8%. The attached table and figure show the correlations between individual variables.

}}:

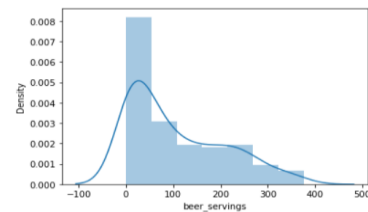
| | beer_servings | spirit_servings | wine_servings | total_litres_of_pure_alcohol |
|------------------------------|---------------|-----------------|---------------|------------------------------|
| beer_servings | 1.000000 | 0.458819 | 0.527172 | 0.835839 |
| spirit_servings | 0.458819 | 1.000000 | 0.194797 | 0.654968 |
| wine_servings | 0.527172 | 0.194797 | 1.000000 | 0.667598 |
| total_litres_of_pure_alcohol | 0.835839 | 0.654968 | 0.667598 | 1.000000 |



The heat map above shows the correlations between the variables.

The distribution of the variables is assessed by plotting a distribution plot which shows that most of the variables are skewed to the right implying that most of the beer servings observations are between 0 and 200 while that of sprite servings is between 0 and 150 servings. This is evidenced by the attached figures.

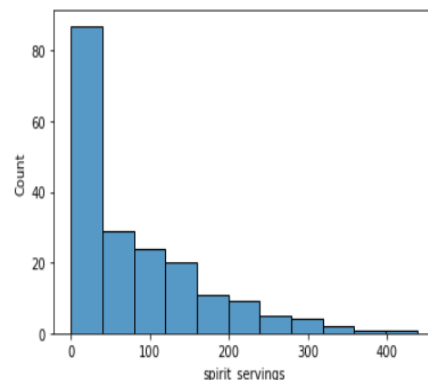
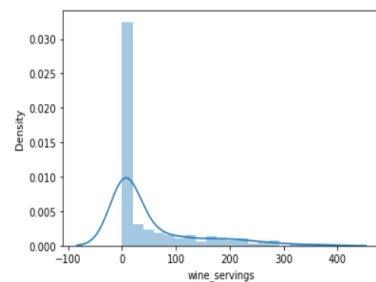
Out[10]: <AxesSubplot:xlabel='beer_servings', ylabel='Density'>



The figure above shows that the number of beer servings are skewed to the right with the majority of countries having beer consumption ranging from 0 to 250.

In [12]: sns.distplot(mydata['wine_servings'])

Out[12]: <AxesSubplot:xlabel='wine_servings', ylabel='Density'>



The Histogram above shows the distribution of the spirit servings as they are skewed to the right with majority observations ranging between 0 and 200.

The country with the highest number of beer servings is Namibia while the countries with the minimum are Afghanistan, Bangladesh, and the Cook Islands.

Also, the Countries with the highest number of spirit servings are Grenada while the countries with the lowest spirit intake are Afghanistan, Algeria, and Bangladesh.

Finally, the countries with the highest number of wine servings are France and the lowest are Afghanistan, Bangladesh, and Bhutan.

Conclusions

Some of the findings corresponded to stereotypes. The French drink more wine than any other nation, with an average of 370 glasses per person per year, comparing to 84 glasses in the US. However, keep in mind that these are only per capita figures, and they do not account for all of the abstainers. France's drinking population's wine consumption will rise even more. Namibia has the world's highest per

capita beer consumption, with each person consuming 376 12-ounce cans of beer. The average in Germany is 346 cans, while in the United States it is 249 cans. We can also conclude from the analysis that people in Islamic countries conformed to stereotypes about being the least likely to drink alcohol..

References

- Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2, pp. 131-160).
- Morgenthaler, S. (2009). Exploratory data analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(1), 33-44