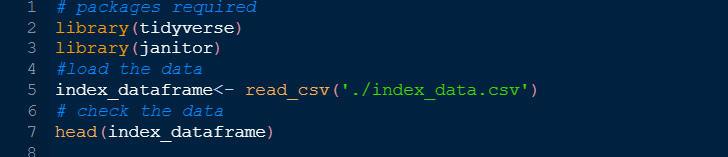
**AN ANALYSIS ON DATA STORYTELLING USING THE LIFE INDEX DATA**

**Introduction**

Data visualization is the art of presenting data in a visually appealing and meaningful way. While there are many ways to visualize data, storytelling is one of the most powerful methods to convey insights to an audience. Storytelling in data visualization involves combining data, visuals, and narrative to create a compelling and engaging story that delivers insights and impacts decision-making. Success stories in data visualization demonstrate the power of effective storytelling to communicate complex data insights in a way that is easily understood by non-technical audiences. Whether it's a chart, graph, or interactive dashboard, the success of data visualization ultimately depends on the ability of the visualization to convey the intended message to its audience. In this context, the combination of data and storytelling is increasingly being recognized as a powerful tool for businesses, researchers, and individuals to achieve success. By effectively combining data and storytelling, organizations can create visualizations that not only inform but also inspire action and drive results. In this article, we will explore the importance of storytelling in data visualization and examine some of the success stories that illustrate its impact. In this case the data to use will be the life index data taken from the Organization for Economic Cooperation and Development.

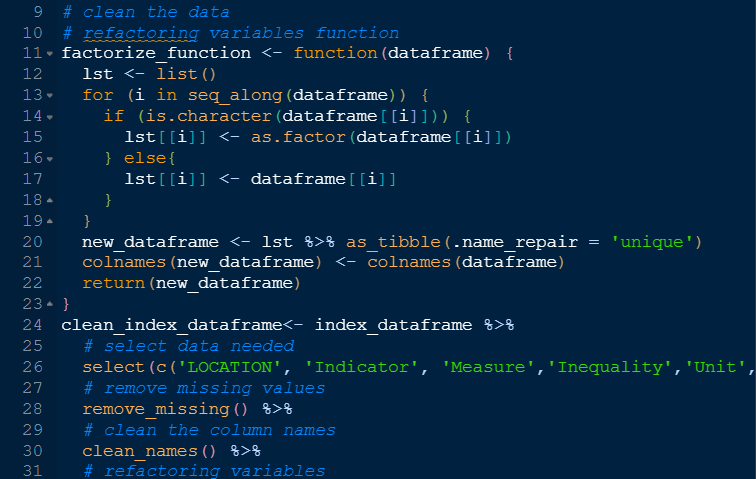
**Loading the data and packages required**

We will load this file since it’s in csv format with read\_csv() function found in the tidyverse’s readr package therefore we will also load the package tidyverse. This is shown below how it was done.



**Cleaning the data**

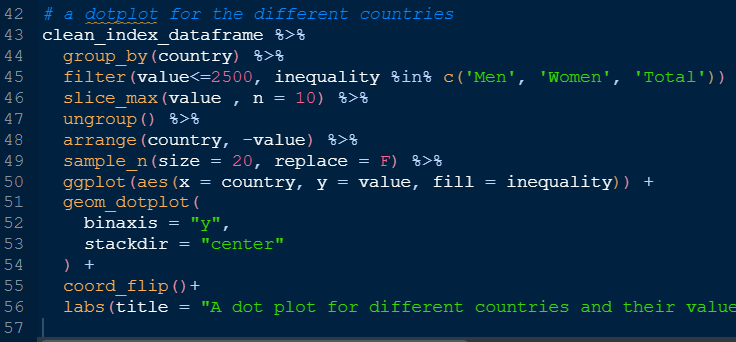
Data cleaning, which involves locating and fixing mistakes, inconsistencies, and missing values in a dataset, is a crucial stage in the data analysis process. For cleaning up data, R offers a number of functions and packages, including tidyr, dplyr, and stringr. Managing missing data, recoding variables, merging datasets, and other tasks can all be accomplished with these tools. Analysts can acquire insights into the underlying patterns and trends in the data as well as verify the accuracy and dependability of their outputs by cleaning the data. In this sense, data cleansing is an essential step in the data analysis process and should be carried out meticulously. We then need to perform this process before proceeding to the analysis. Data cleaning was performed and displayed below.

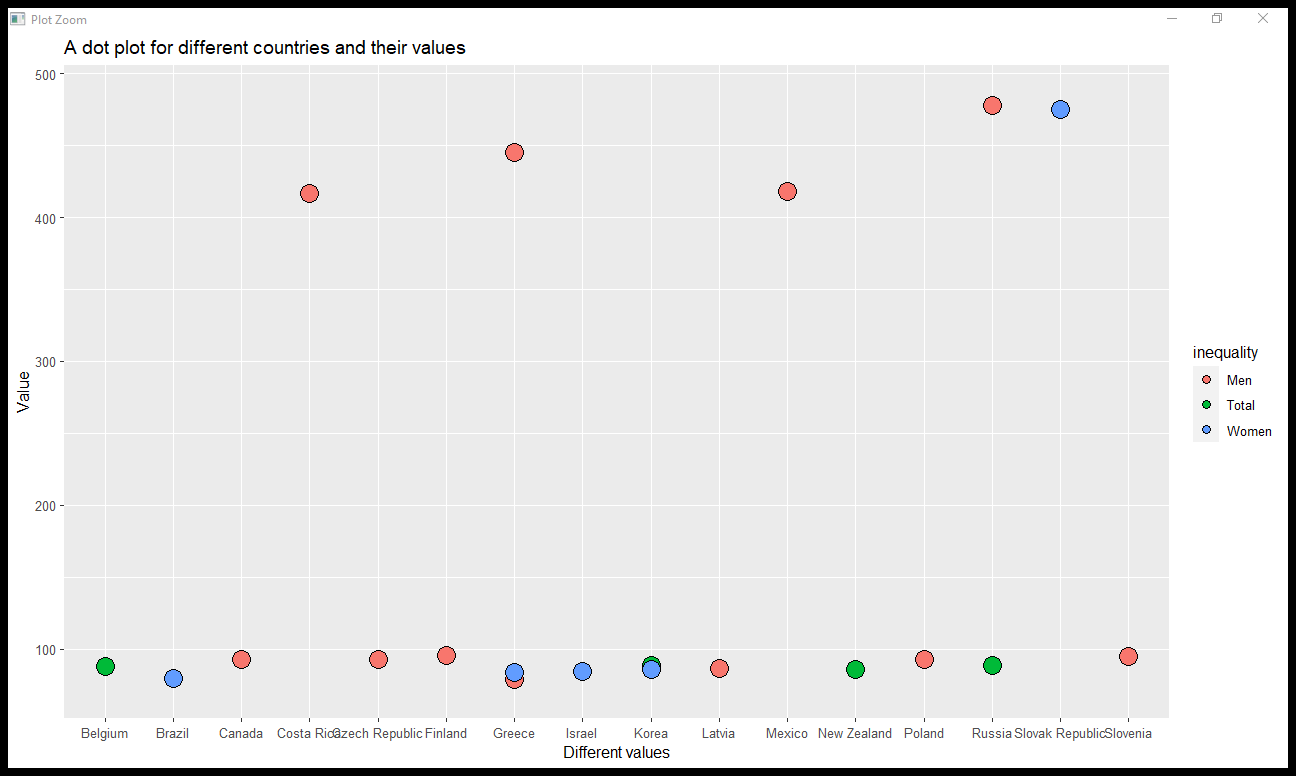


After cleaning the data a new data was created which will be used for the analysis.

**Analysis and data visualization**

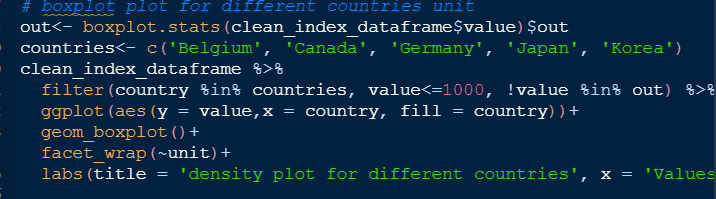
Let's begin by creating a visual representation that will help us understand what storytelling is and how we can present it using a data visualization. First plot will be a dot plot that will be created from 2 categorical variables. A dot plot can be used to show the distribution of one categorical variable over the levels of another categorical variable. In this plot, each dot represents an observation, and the dots are grouped by the categories of the second variable. The dot plot is useful for displaying both the frequency and distribution of the categories. In our case the variables from the data used will be different countries and their values colored with the levels of inequality in the data. This code is displayed below.

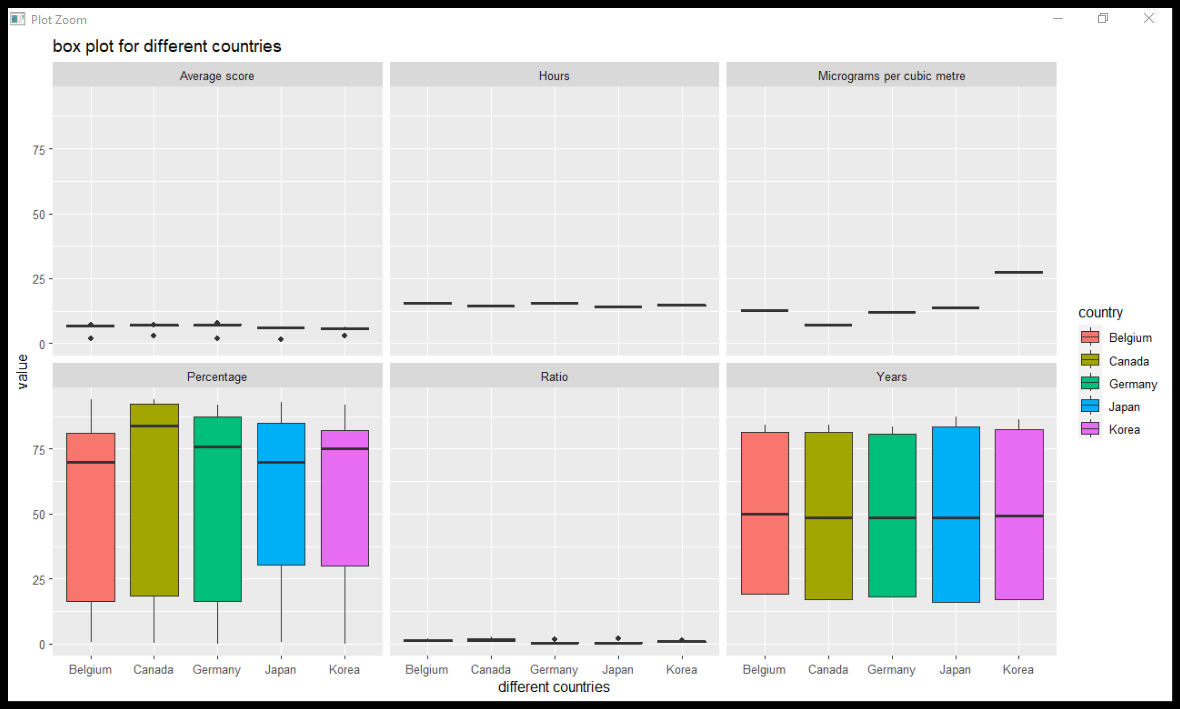




In our plot above we filtered and selected 20 countries for plotting. Using this data visualization, we can create a success story from it. The plot is displaying points of 20 different countries with their values created on the y-axis while the countries on the x-axis. We see that the country with highest value is Russia followed by the Slovak Republic. Although these countries have high value, say income earned per year, they have flaws, that is, colored with the levels of inequality, we see that Russia country has high level of inequality for men meaning they are biased to a male gender while the Slovak Republic are biased to women due to the color she has. This is an example showing how we can create a story from a data, that is we can turn the data visualization into a storytelling tool to be used to provide information to users.

To clarify that data visualization can be used as a storytelling tool we create another plot from the data. This time using a density plot facetted on different units, we create a data visualization for the analysis. This is shown below.





In the plot above we created a boxplot which has country variable on the x-axis and value on the y-axis. We first filtered out any outliers and selected 5 countries to be displayed in the plot with the values less than 1000. The data visualization gives a success story that is, we see that the boxplot is used to compare continuous variable across different unit categories. We see the data has some values well distributed in the percentage unit level and year unit level, this means that the majority of the data values are higher, with fewer extreme values in the upper range. A high boxplot is also indicative of a more concentrated or tight distribution, with less variability in the data. Other levels have less or no distribution of data and this represents a dataset with a lower median and lower quartile. This means that the majority of the data values are lower, with fewer extreme values in the lower range. A low boxplot is indicative of a less concentrated or more spread-out distribution, with more variability in the data.

**Conclusion**

In conclusion, data storytelling is essential in communicating insights and trends in data analysis. Success stories can be created using data visualization techniques, such as density plots and dot plots, to showcase important findings from data analysis. The life index data from OECD provides a good example of how these visualization techniques can be used to identify patterns and trends. The success of data storytelling lies in the ability to effectively communicate insights in a visually appealing and easy-to-understand way.