Data Visualization Assignment

Delivered as the final project of Data Visualization and Management Course

Dr. Fathima Abubacker

Student Name: Razan Alsulieman

Student ID: 18215161

1. Introduction -

This research project determines the relationship of suicides and unemployment rates of youth people from the age of 15 to 24 years old from 2010 to 2014. The intention of this analysis is to illustrate the association of suicide with unemployment. The data mainly focuses on suicide and unemployment rates based on countries and years.

The coding part which was developed with **Python** is available on GitHub on (https://github.com/RazanAlsulieman/DataVisualizationProject/blob/master/DataVisualAssign.ipynb).

2. Dataset -

Two datasets were obtained from **Kaggle** to reshape and visualize as follows:

- *SuicideRate.csv* **Dataset** [1]

  This dataset contains the rate of suicides in some countries from 1985 to 2016. It has 27820 rows and 12 columns.

- *UnempRate.csv* **Dataset** [2]

  This dataset has 219 rows and 7 columns, which contains the rates of unemployment in some countries from 2010 to 2014.

3. Process -

Three types of packages were used to analyze the data which are: pandas, plotly offline and plotly graph objs. The merged dataset of the two mentioned datasets contains 740 rows with 6 columns. The dataset was imported using *pandas.dataframe.read_csv* for both suicide and unemployment datasets.

- Cleaning Datasets:

  The first step was to view the datasets and clean them from null values using pandas packages *pandas.DataFrame.drop* to drop unnecessary columns and *DataFrame.dropna* to delete null values.

- Reshape DataFrames:

  Reshaping data frames using *pandas.DataFrame.melt, pandas.DataFrame.pivot and pandas.DataFrame.astype* to change columns to rows and their type. As well as, renaming data frames using *pandas.DataFrame.rename* after cleaning the dataset.

- Combine Two Datasets:

The prepared datasets are merged using *pandas.DataFrame.merge*, *pandas.DataFrame.groupby* and *pandas.DataFrame.sort_values* to start the vitualization.

4. **Result -**

After preparing the datasets for visualization, visualizing the rate of suicides by gender was the first bar chart created using *plotly.offline bar chart*, where it interactively presents the rates of each year for both males and females as in **fig.1**. The figure reveals that males has the most suicide rates compared to females. Moreover, it shows that the rates are slightly decreasing in each year.

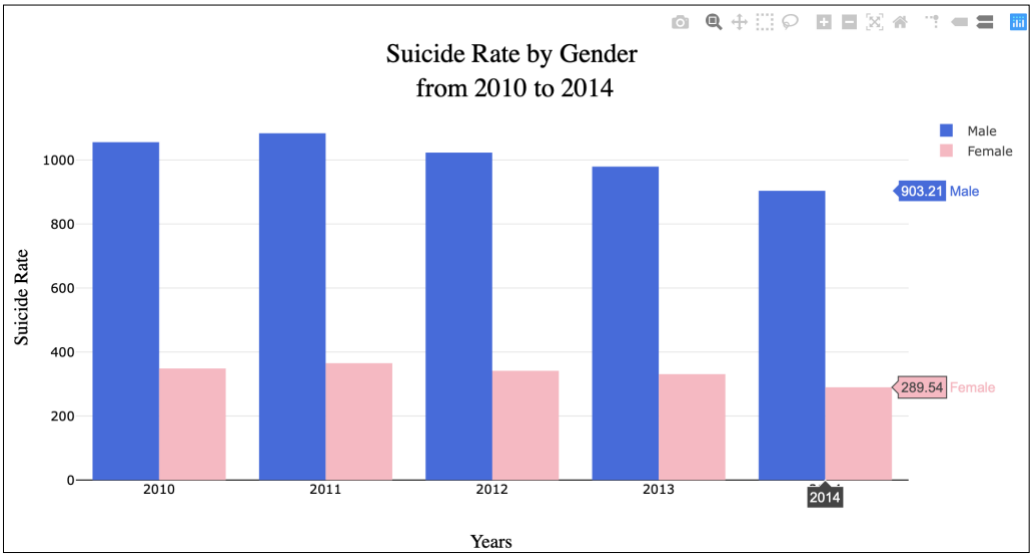

Fig.1: Suicide Rate by Gender.

After that a scattered plot was created to understand the relationship between suicides and unemployment rates using *plotly.graph_objs offline scatter* plot as shown in **fig.2**. The figure proves that if there was an increase in unemployment rate there will be an increase in the suicide rate and vice versa. It might be due to the stress of not finding a job as well as no income for living expenses which leads to suicide. Therefore, it can be concluded that the suicide rate is proportional to the unemployment rate.

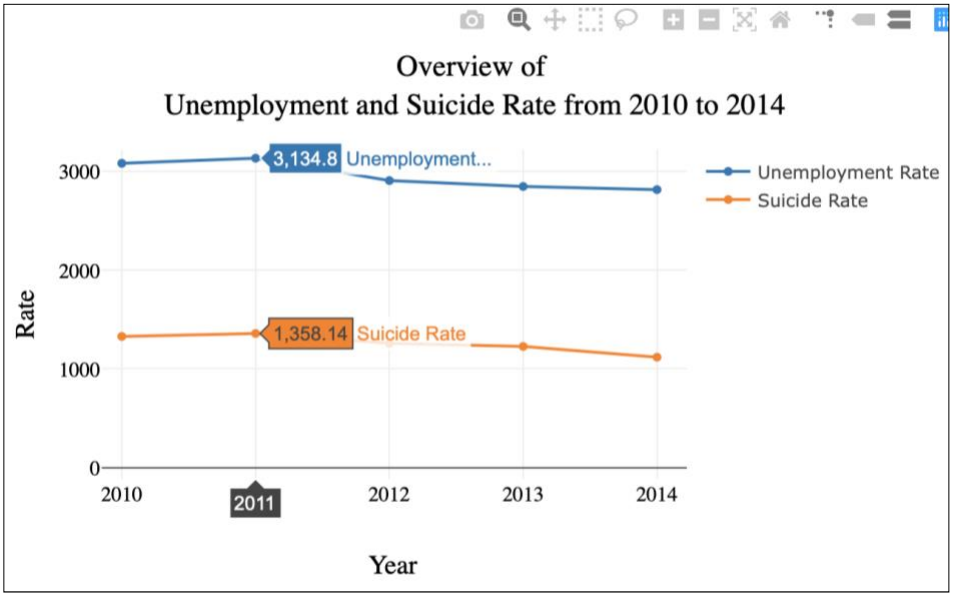

Fig.2: Unemployment and Suicide Rate by Year.

Then the suicide and unemployment rates were visualized in a stacked bar using *plotly.graph_objs.Bar* illustrated in **fig.3** and **fig.4**, where it states the rate for each country in each year from 2010 to 2014. The visualization reveals that most suicides were in Suriname, while Bahrain and South Africa had the least number of suicides. On the other hand, the highest unemployment rate was in South Africa and Spain, while Qatar had the least number of unemployment which is due to their education level and their parents' level of education.

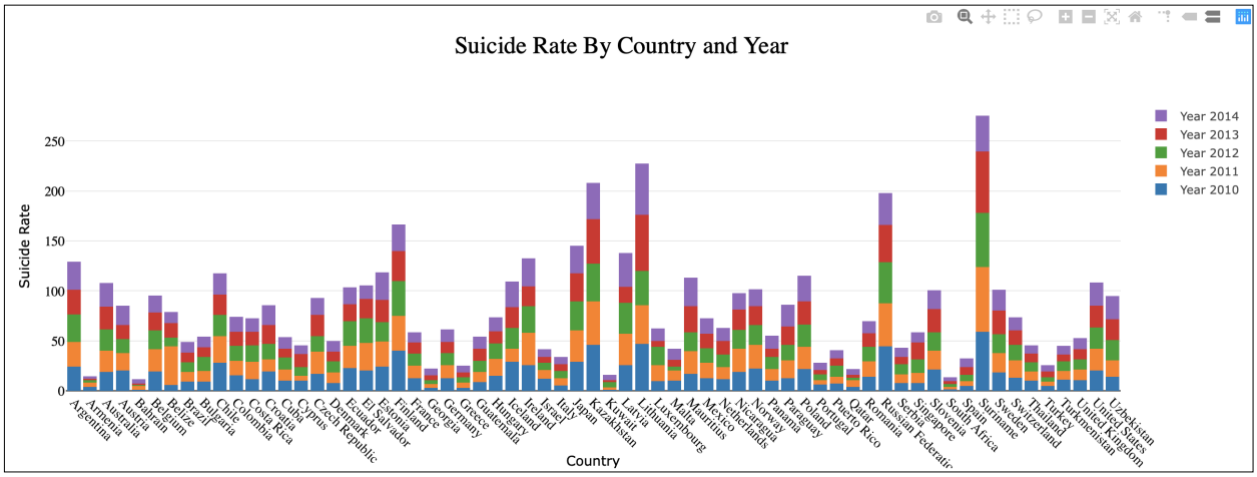

Fig.3: Suicide Rate by Country and Year.

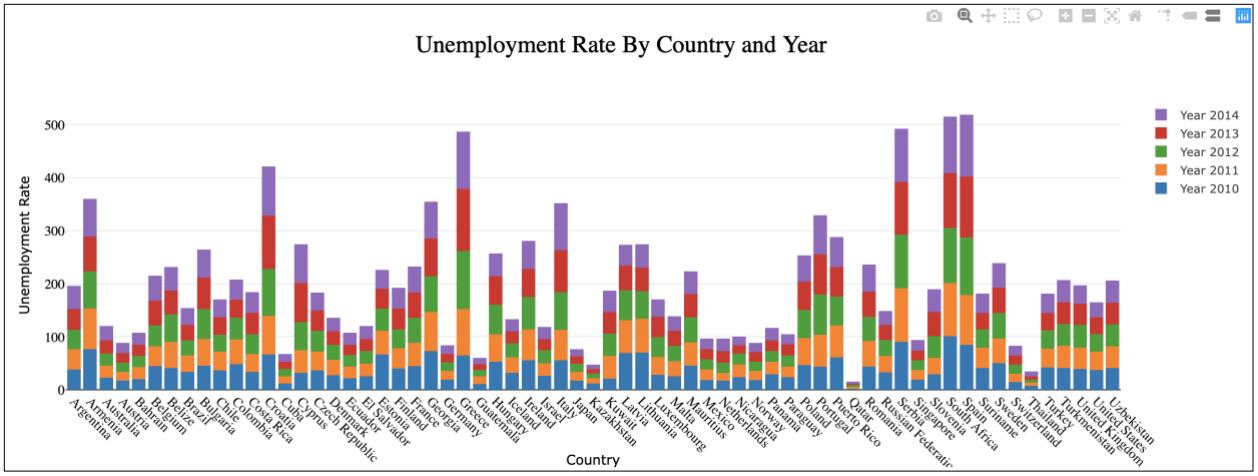

Fig.4: Unemployment Rate by Country and Year.

Visualizing the data on choropleth maps will present the data more clearly. As shown in **fig.5** and **fig.6** the data may reveal some differences regarding the relationship of suicides with unemployment. The map shows that Suriname had the most suicide rates while it did not have high unemployment rate. These contrasts can be due to other reasons and factors other than unemployment such as, their poor economic conditions and the lack of public and ambulant mental healthcare systems [3]. Furthermore, the different association between suicide and unemployment for South Africa with its high unemployment and low suicides, which likewise was affected by other factors such as religious reasons where it acts as a protection factor against attempting suicide [4].

On the other hand, there are some proportional relationships on other countries like Lithuania, Russian Federation, Argentina and Chile where they have high unemployment rate as well as a high suicide rate.

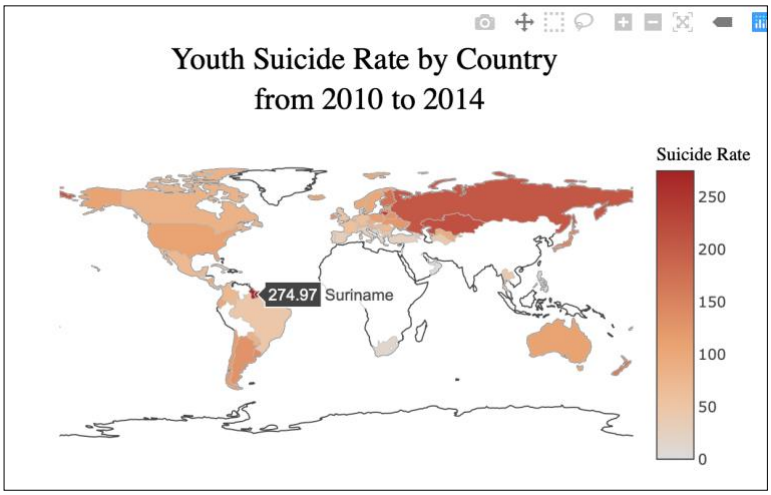

Fig.5: Youth Suicide Rate by Country.

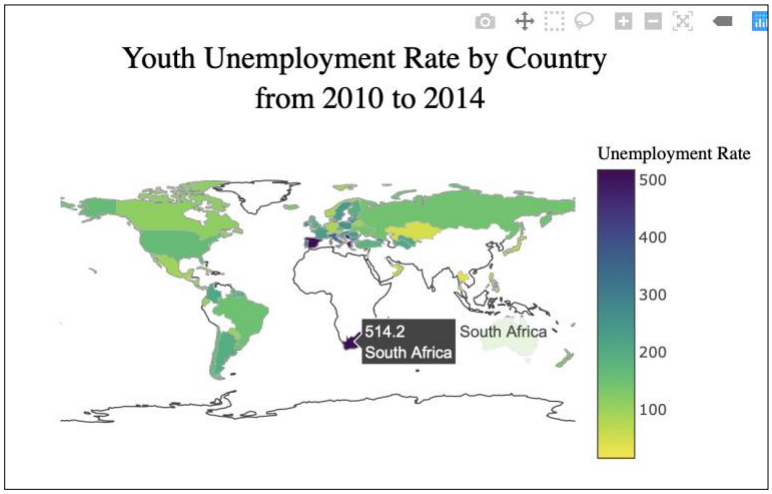

Fig.6: Youth Unemployment Rate by Country.

5. Conclusion –

In conclusion, the over view of the data revealed the proportionality between suicides and unemployment. However, viewing the data more closely shows different association between them due to other factors such as, education level, economic conditions and religious reasons. Interactive graphs were applied such as bar charts to easily read the data, as well as, a scattered plot of the rates to view the changes for each year. Also, a choropleth maps were plotted with an overview to view a better representation of the data for each country.

The functionalities of merging or combining two graphs in a single graph would enhance the visualization's interactivity, however, it was not achieved on this project.

References:

[1] Suicide Rates Overview 1985 to 2016. (2019). Retrieved from https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016#master.csv

[2] World Bank Youth Unemployment Rates. (2019). Retrieved from https://www.kaggle.com/sovannt/world-bank-youth-unemployment#API_ILO_country_YU.csv

[3] Sisask, M., Värnik, A., Kõlves, K., Bertolote, J., Bolhari, J., J Botega, N., … Wasserman, D. (2010). Is Religiosity a Protective Factor Against Attempted Suicide: A Cross-Cultural Case-Control Study. Archives of Suicide Research : Official Journal of the International Academy for Suicide Research, 14, 44–55. https://doi.org/10.1080/13811110903479052

[4] Graafsma, T., Kerkhof, A., Gibson, D., Badloe, R., & Beek, L. M. Van De. (2006). High Rates of Suicide and Attempted Suicide Using Pesticides in Nickerie , Suriname , South America, 27(2004), 77–81. https://doi.org/10.1027/0227-5910.27.2.77