Information Visualisation Project (COMP3045 UNMC)

**Analysing Data with Tableau System**

Student Name: Loh Jin Xian

Student ID: 016763

# Hypotheses

## 1st dataset:

1. There are more males than females that committed suicides between the year 1984 and 2018.
2. The number of millennials that committed suicides are the highest compared to other generations.
3. The highest suicide age range is between 20 – 30 years old.
4. Country that has a lower GDP (Gross Domestic Product) tends to have a higher suicide rate.

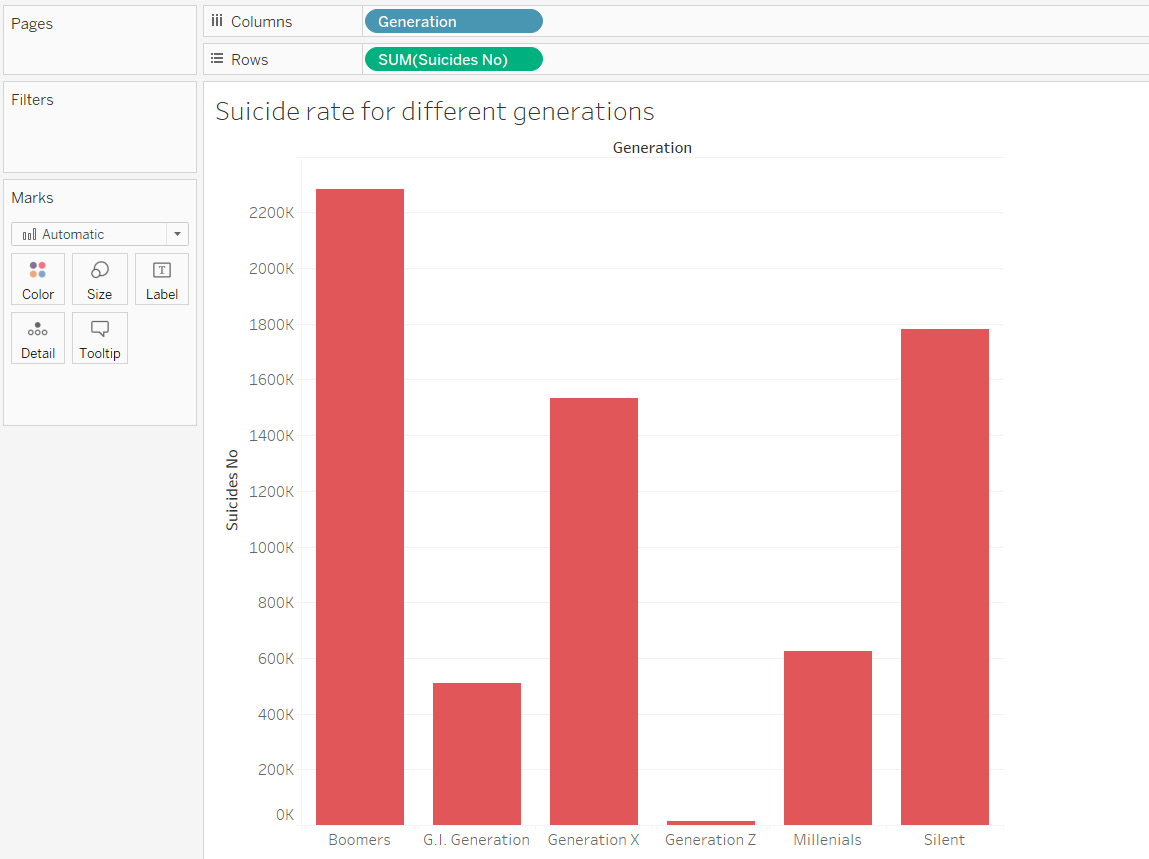
## 2nd dataset:

1. People that are generous are generally happier than less generous people.
2. Country that has a higher GDP (Gross Domestic Product) has a higher happiness score.
3. Asian countries have higher happiness score than other countries.
4. People that are happy tends to live longer than their counterpart.

# Data Analysis

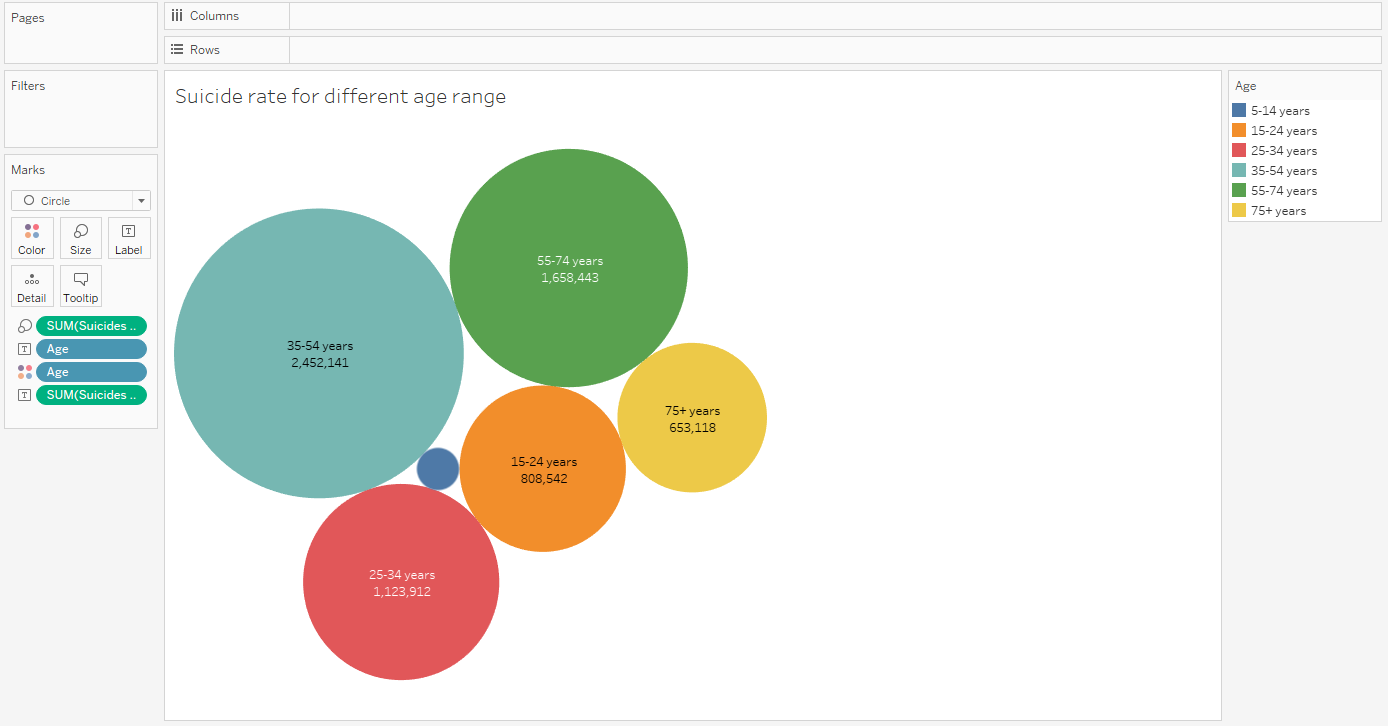
## 1st dataset:

### Millennials have the highest suicide rates



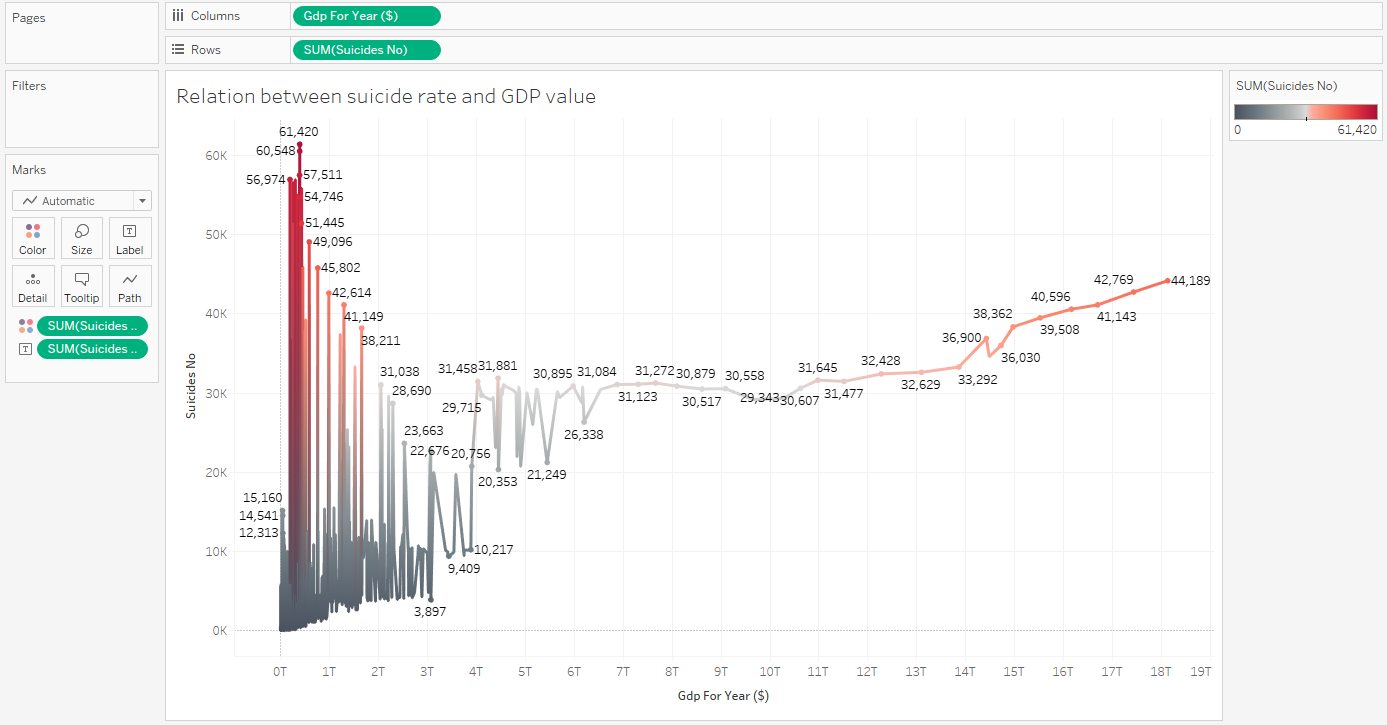
Based on the bar chart that has been generated, millennials are not the highest suicide rate generation. The generation that has the highest suicide rate is generation X. A bar chart is used to visualize this data as it is straightforward and clean.

### The highest suicide rate is between 20 – 30 years’ old



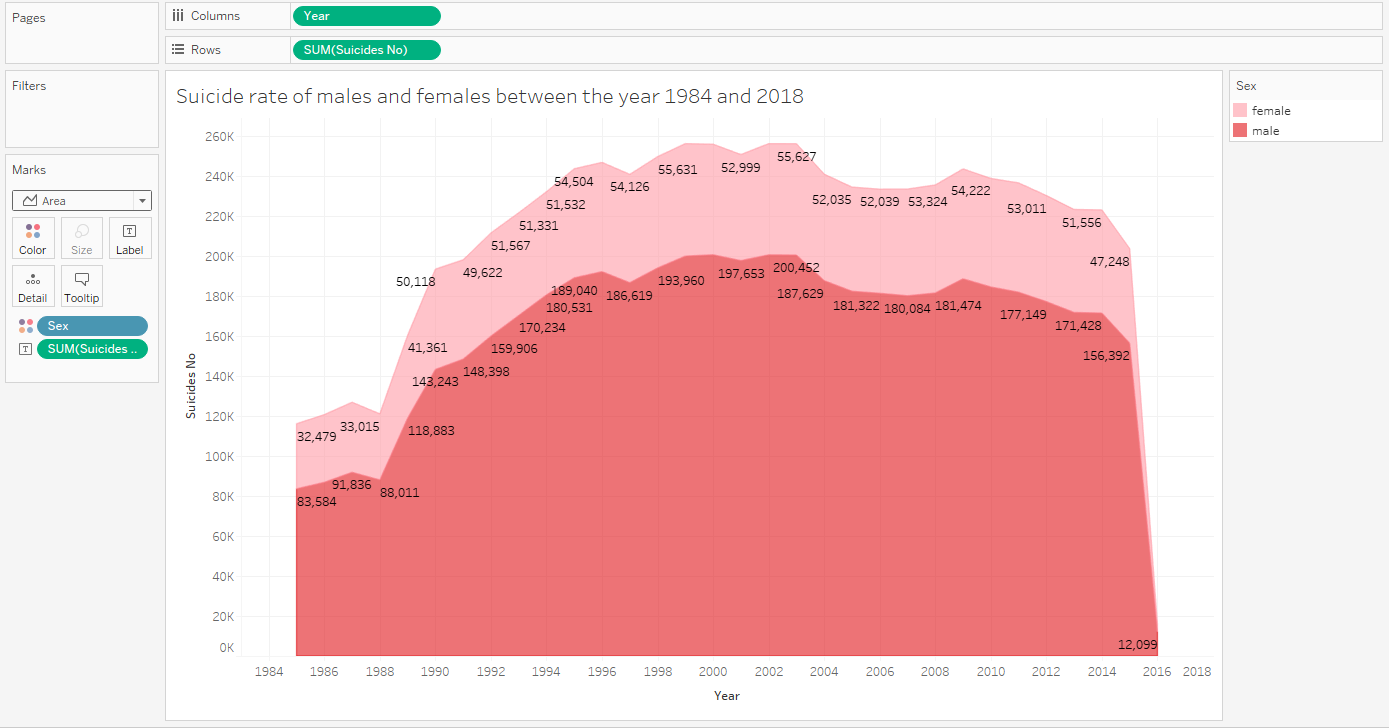
Based on the packed bubble visualization, the highest suicide rate is between 35 – 54 years old. There are many reasons that contribute to this finding, one of them might be the stress that they faced in their adulthood like stress from work or stress from family. Packed bubble visualization is used to give a quick overview on which bubbles are the biggest.

### Lower GDP of country lead to higher suicide rate



Based on the lines graph, the hypothesis is accepted as lower GDP countries will tend to have a higher suicide rate compared to countries that have higher GDP. This visualization technique is chosen as it gives a clean and simplistic look to the data without sacrificing any forms. However, there tends to be some over plotting on the left side of the graph, which could make the visualisation more complicated. This is solved by having a colour indicator to simplify the data.

### Suicide rate of females and males



Based on this area chart graph, there are definitely more males that committed suicides than females for the past 30 years. This area graph visualization can give both the total and separated numbers of suicide rate in 1 simple colour coded graph which is also pleasing to the eyes.

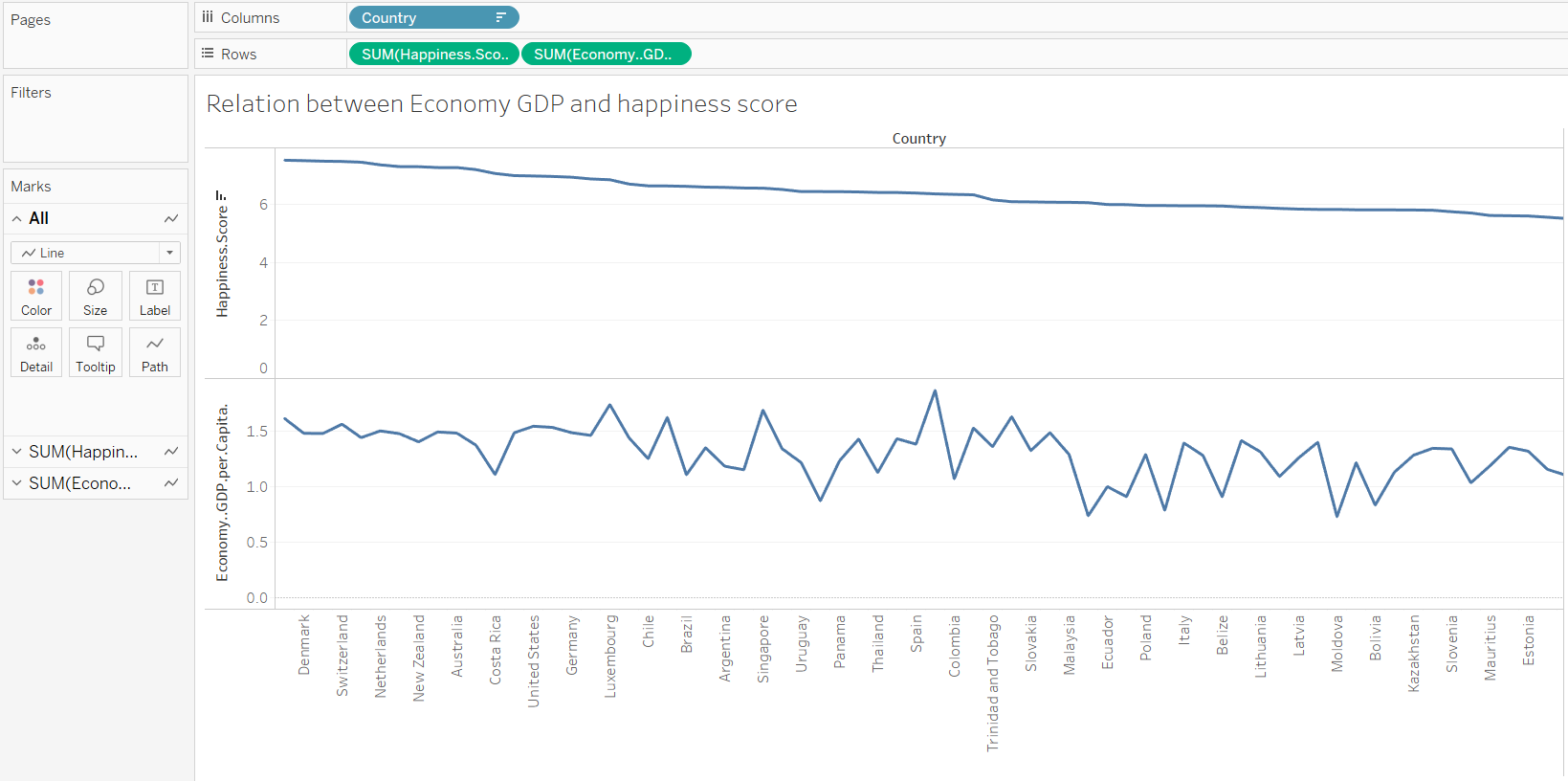
## 2nd dataset:

### Generous people are happier

### 

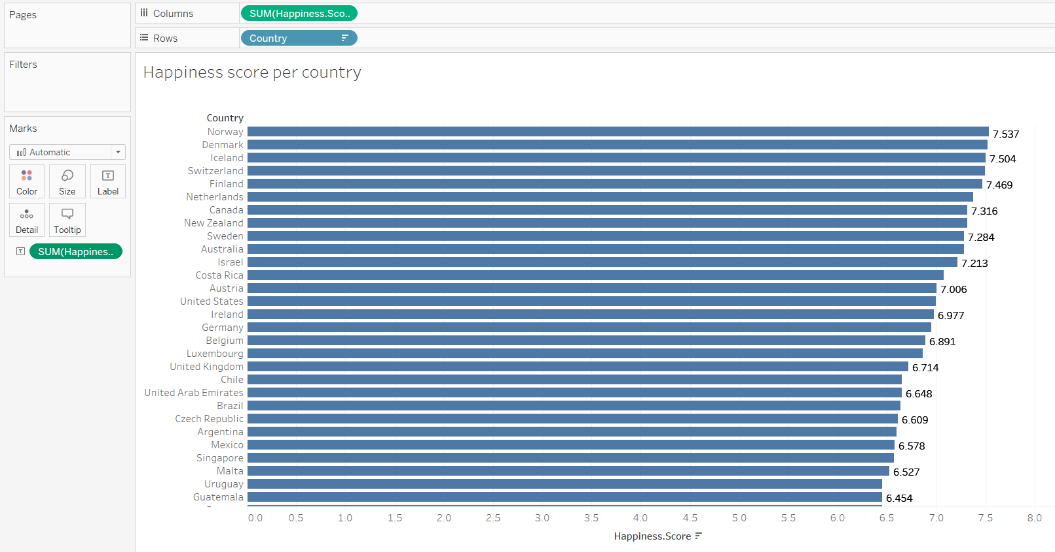
Based on heat map graph, although generous people tend to have a higher happiness score. In some countries, generous people don’t have a high happiness score, which lead to inconsistent relation between generosity and happiness score. Heat map graph can provide both locations and numbers visualization fitted in a map diagram.

### Higher GDP (Gross Domestic Product) leads to higher happiness



According to the lines graph, there is no direct relation between GDP and happiness score. This line graph visualization can gives a clear comparison between different data in a non-obtrusive way.

### Asian countries have higher happiness score



### Based on the bar graph, the top 10 highest happiness countries are not from Asia. Therefore, people living in Asian countries like China, Japan etc are not the happiest.

### Happy people live longer

### 

Based on the horizontal bars, happiness score and health life expectancy don’t have a direct relation too. As seen in the graph, the country that has the highest happiness score doesn’t really increase the life expectancy of the people. Bar chart visualisation technique can also be used in this way for comparison purposes.

# Tableau system review

## Advantages:

* Can accept a variety of file formats as dataset (e.g. Excel, JSON, CSV file)
* Simplified method in generating a visualisation graph by using the ‘Show Me’ function which automatically suggest suitable graph format based on the data that you want to visualize
* Can have many rows and columns for a dataset
* Tableau has a filter feature that can instantly show or hide the necessary data. Which improve productivity

## Disadvantages:

* The quality of the visualization graph is largely dependent on the quality of the dataset. Low quality dataset will generate a bad or unusable graph. For example, complicated dataset will generate a graph that is over plotting.
* Tableau might sometimes have inconsistency problem in showing the dataset. For example. 2 datasets that have the same format will be shown by Tableau in a different way, which can be quite confusing to the users sometimes.
* Tableau doesn’t provide a dark theme user interface, which can annoy some users that work at late night hours. White theme Tableau background is too bright to the eyes in a dim environment.
* Given the functionalities that tableau provides, it might be too overwhelming to beginners or someone that is not computer literate. Therefore, simple guides that are explained in images instead of words can be included for reference.

# Conclusion

Tableau is definitely one of the best software in the market right now that make databases and spreadsheets understandable to ordinary people. All functionalities that are included in the software aims to simplify the traditional process of visualizing complex data and modernize it to something that can keep up to the pace of technology. In reviewing this particular software, we have just scratched the surface of what Tableau is capable of and we hope that Tableau updates its software more frequently in hopes of solving the disadvantages as soon as possible.

# References

1. https://public.tableau.com/en-us/s/resources
2. 1st dataset: <https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016>
3. 2nd dataset: https://www.kaggle.com/unsdsn/world-happiness