Suicide Rates Overview 1985 to 2015

DSC 520, Summer 2019, Final Project

*Gourav Verma**

Aug 8 2019

Abstract

According to the WHO Suicides organization, 800,000 committed suicide in 2018. This means every 40 seconds a person dies by suicide. This number is fortunately dropping. In this project I am trying to explore the evolution of suicide rate using dataset extracted from Kaggle. This compiled dataset pulled from four other datasets linked by time and place, and was built to find signals correlated to increased suicide rates among different cohorts globally, across the socio-economic spectrum.Majority of the data used in this analysis is obtained from World Health Organization.

Structure of cleaned dataset

Data cleaning notes

- 7 countries removed(<= 3 years of data total)
- 2016 data was removed(few countries data was present and those did had data missing)
- HDI variable was removed due to more than 50% data was missing.
- Generation and Age variable looks similar. We will find our in analysis later, if Generation variable produces different outcome.
- Created new variable 'continent' using countrycode package.
- Tdying data includes changing format of variables and arranging the values to ordinal.

```
head(suicide)
```

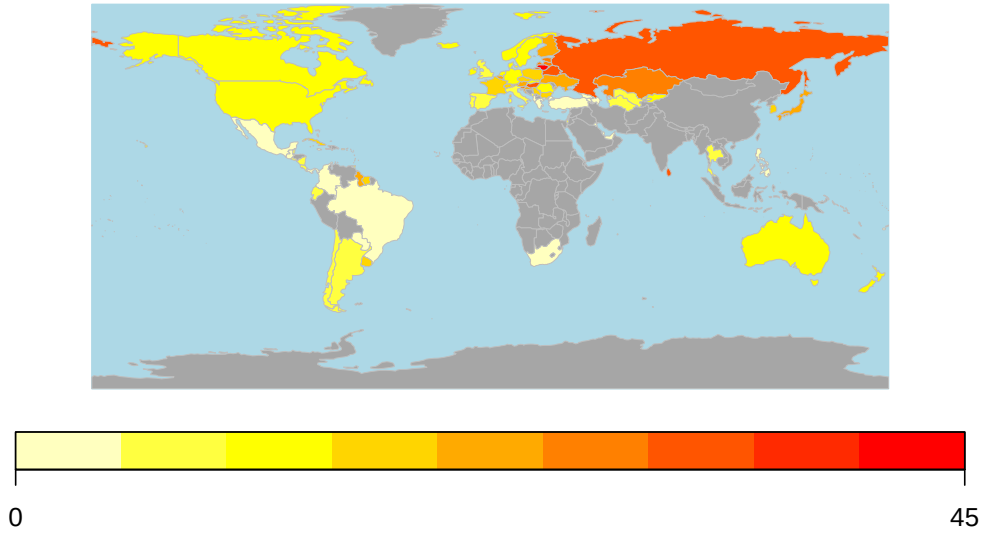
```
##   country year    sex   age suicides_no population  gdp_for_year
## 1 Albania 1987   male 15-24          21     312900 2,156,624,900
## 2 Albania 1987   male 35-54          16     308000 2,156,624,900
## 3 Albania 1987 female 15-24          14     289700 2,156,624,900
## 4 Albania 1987   male   75+           1      21800 2,156,624,900
## 5 Albania 1987   male 25-34           9     274300 2,156,624,900
## 6 Albania 1987 female   75+           1      35600 2,156,624,900
##   gdp_per_capita      generation continent
## 1            796    Generation X    Europe
## 2            796          Silent    Europe
## 3            796    Generation X    Europe
## 4            796 G.I. Generation    Europe
## 5            796         Boomers    Europe
## 6            796 G.I. Generation    Europe
```

---

*Bellevue University, Nebraska

Geographical heat map

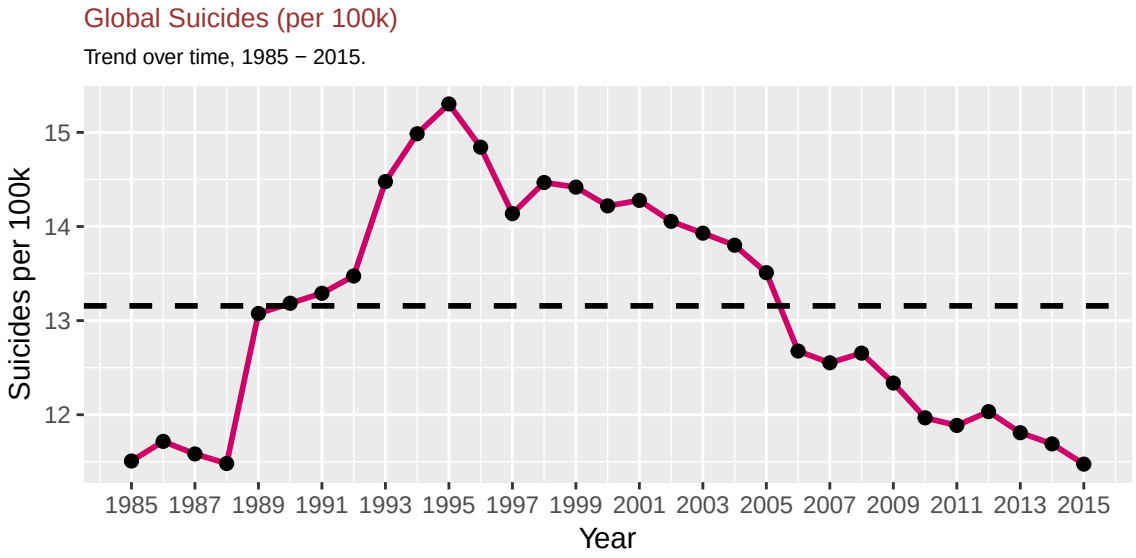
Below geographical heat map shows suicide rate for timeframe(1985-2015). **Grey area** shows lack of information for those countries, specially in **Africa and Asia**. TO get complete understanding of global trend we have limited data, but, we will continue with the data available.



Global Analysis

Below chart shows global trend of suicide rate per 100k population over the years. The dashed line is the **Global average suicide rate** from 1985-2015: **13.5 deaths** (per 100k, per year).

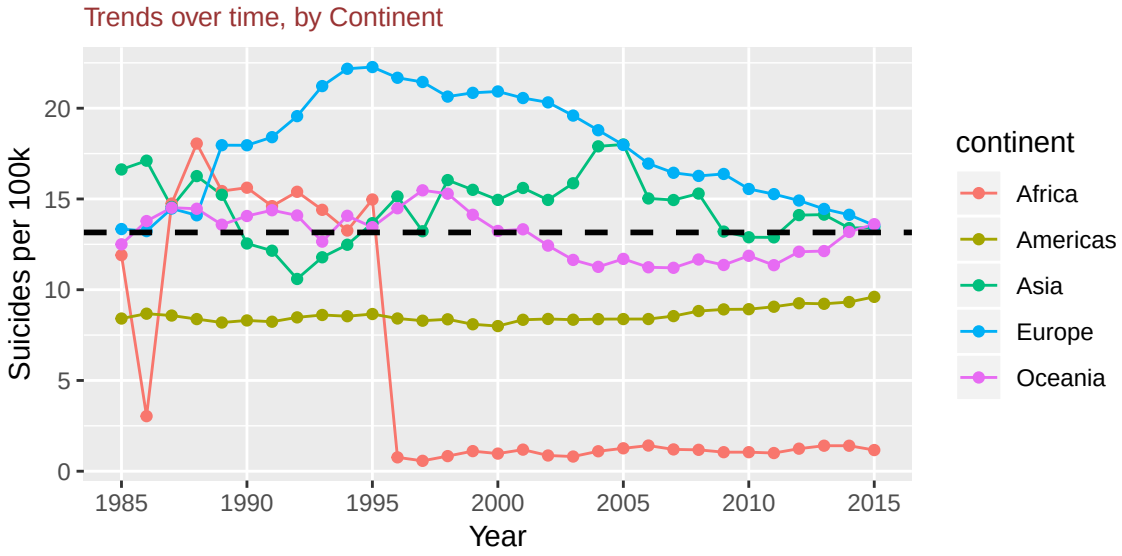
- Peak Suicide rate was **15.3** deaths per 100k in **1995**.
- Rates are now steadly decreasing and was **11.5** in **2015**.



Continent Trend

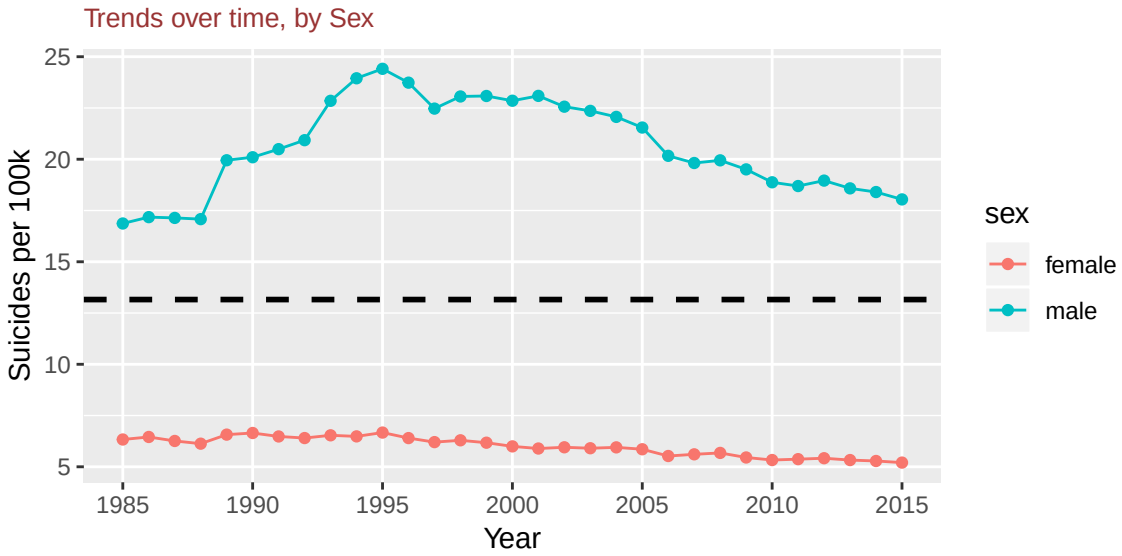
Below graph shows suicide rate variance over the time for different continents.

- **Europe has highest rate overall**, but steady decrease observed after 1995.
- Europe rate is similar to Asia and Oceania in 2015.
- **Oceania and American** trend is increasing over the time, which is concerning.



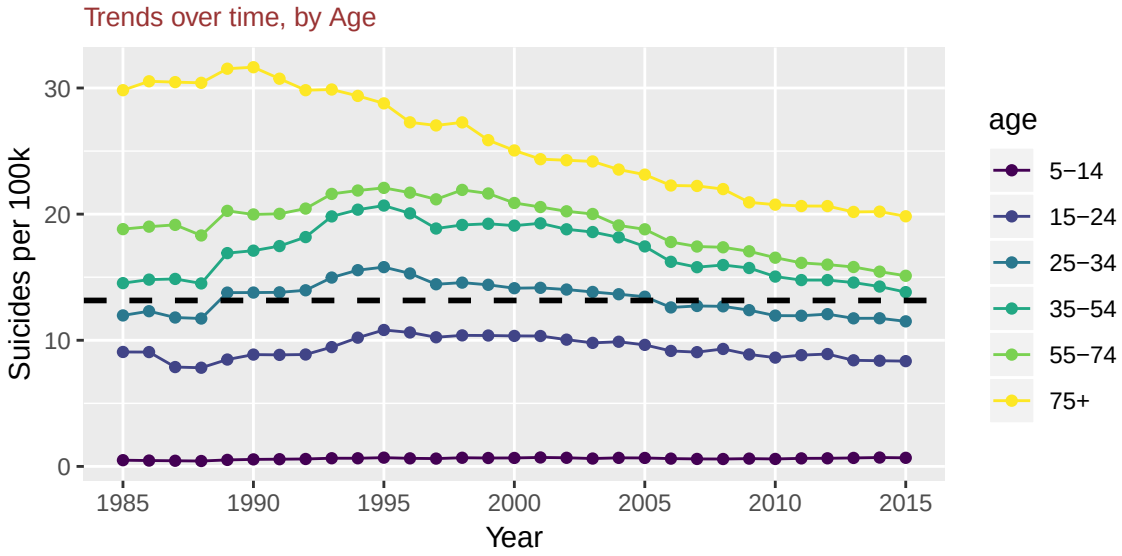
Trend by Sex

- Globally, the rate of suicide for men has been ~**3.5x higher for men**.
- Difference between male and female suicide rate might be affected by sex ratio.



Trend by Age

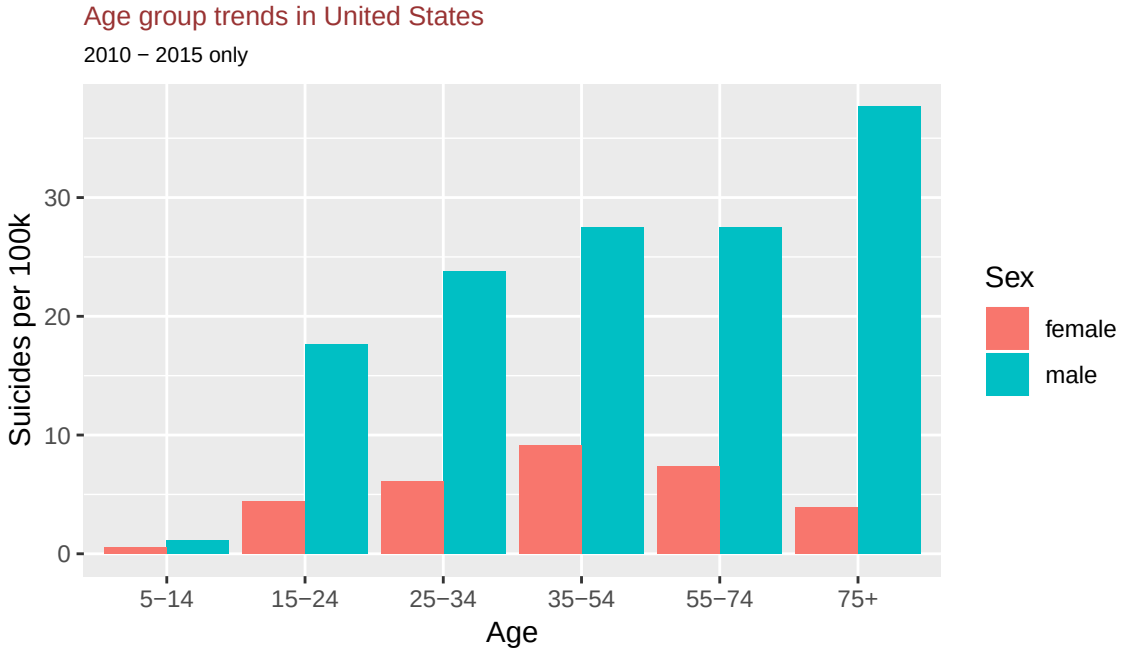
- Suicide rate **increases with age**.
- Rate of suicide is **highest** for age group **75+**.
- Rate for age group 5-14 appears static.



United States Trend Analysis

Below plot shows data from 2010 to 2015 of United States. Bar chart is plotted to differentiate between male and female suicide rates.

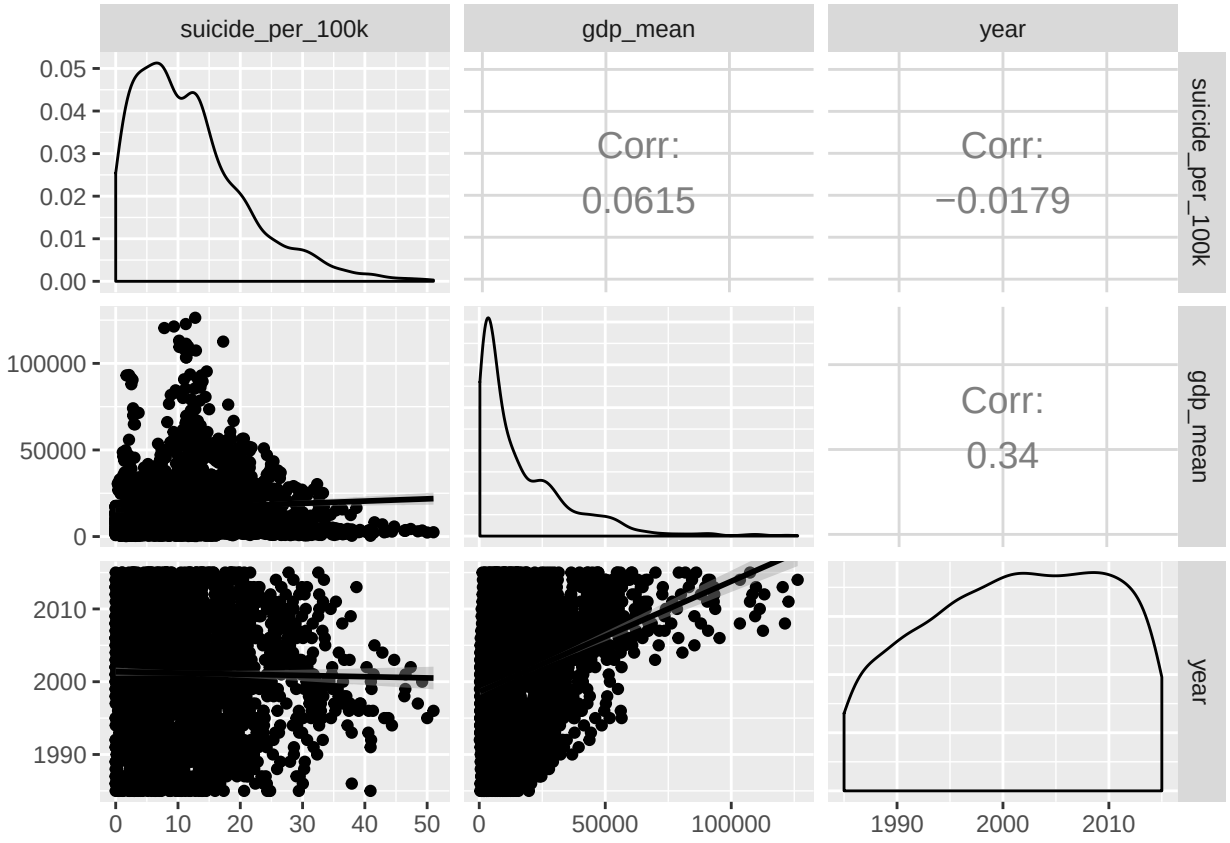
- US suicide rates for men continues to **increase with age**.
- **Females** in US of age group **35-54** has higher suicide rate of **9 per 100k**.



How Suicide Rate is related to GDP of a Country?

Below is the correlation plot between GDP mean, Suicide rate per 100k and year, grouped by each country each year.

- It is evident that GDP of a country increases by year. Have strong correlation of **0.34**.
- Correlation between suicide rate per 100k and GDP mean for acountry per year is very week, i.e. **0.06**.
- We can say that we can't bulid an accurate predictive model for suicide rate based on GDP of a country.



Conclusion

Suicide depends on many factors and can be caused in different scenarios. This analysis was done keeping suicide prevention as a motive. With the available data, we saw several trends. Most striking was to find high suicide rates among people with age group 75+. It's good that the global suicide rate is coming down. Globally Europe seems to have high suicide rates over the years, which is also coming down and became equivalent to Asia and America in 2015. On the other hand, America and Oceania have concerning trend of increasing suicide rate.There is huge difference between suicide rate of male and of female. Male seems to have 3.5x time suicide rate of female. In the end, we also saw that there is a week relation between the richness of a country and its suicide rate.

References-

- GitHub. (2019). vincentarelbundock/countrycode. [online] Available at: https://github.com/vincentarelbundock/countrycode [Accessed 8 Aug. 2019].
- Rusty. "Suicide Rates Overview 1985 to 2016." Kaggle, 1 Dec. 2018, www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016
- United Nations Development Program. (2018). Human development index (HDI). Retrieved from http://hdr.undp.org/en/indicators/137506
- World Bank. (2018). World development indicators: GDP (current US$) by country:1985 to 2016. Retrieved from http://databank.worldbank.org/data/source/world-development-indicators#
- [Szamil]. (2017). Suicide in the Twenty-First Century [dataset]. Retrieved from https://www.kaggle.com/szamil/suicide-in-the-twenty-first-century/notebook
- World Health Organization. (2018). Suicide prevention. Retrieved from http://www.who.int/mental_health/suicide-prevention/en/