Happiness Report

Mariia Kucheiko Hendrik Schulze Bröring Georg Daniel Tsambasis

July 22, 2020

1 Abstract

Suicide is a serious global public health issue, which takes close to 800 000 lives every year, more than malaria, breast cancer, or war and homicide. Therefore we dedicate this paper to studying the phenomenon of suicide, in particular which social groups are more susceptible to suicide and which factors are closely interconnected to suicidal behaviour. We show that elderly people tend to commit more suicides than the youngest part of the population. Males are generally more likely to commit suicides than females. The regression analysis, standalone correlation-coefficients and correlation-coefficient matrices show that suicides significantly correlate with depressions and substance misuse, as well as with the World Happiness Report rankings.

2 Introduction

Suicide is a serious global public health issue. It is among the top twenty leading causes of death worldwide, with more deaths due to suicide than to malaria, breast cancer, or war and homicide. Close to 800 000 people die by suicide every year. The reduction of suicide mortality has been prioritized by the World Health Organization (WHO) as a global target and included as an indicator in the United Nations Sustainable Development Goals (SDGs), the WHO 13th General Programme of Work 2019-2023 and the WHO Mental Health Action Plan 2013-2030. A comprehensive and coordinated response to suicide is critical to ensure that this phenomenon does not continue to cost lives and affect millions of people through the loss of loved ones or their suicide attempts. [1]

Organizations supporting people in acute hardship, such as crisis hotlines, are rescuing lives on a day-to-day basis. But since no durable prevention is possible without a deep understanding of the underlying issue, we consider it equally important to study the phenomenon of suicide in details and from different perspectives.

The tragedy of every person is unique. There is a great variety of reasons for suicide thoughts, which often do stand alone, but rather form a complex bundle of socio-economic burdens, which can easily overwhelm a person in need. Similarly diverse are the circumstances of suicides: for some it is an impulsive act of despair, for others its seemed to be the only way out of the situation which has been haunting them for years.

Nevertheless this complexity should not stop us from studying the phenomenon, since there seem to be factors, that come along with many suicide cases. This is giving us an indication that they are worth investigating. What they are and how strong the interconnection is- these are the major questions which guided us during this research.

Since the scope of this study is limited, we needed to set priorities and focus on several factors. In the first stage, we investigated the demographic aspect, i.e. we analyzed, which population groups are most susceptible to suicide. We found it important in order to gain an overview of the underlying phenomenon. Further, since we assumed that there might be regional differences, we took the geographical factor into consideration. The interesting findings of this stage are presented in section 4.1. Next, we wanted to verify the common believe that mental diseases, in particular depressions, and substance misuses often accompany suicides. We discuss the results of this analysis in the section 4.3 and 4.4. Furthermore, we compare suicide

rates, cases of depression and substance misuse with rankings from the World Happiness Report in the section 4.5. In the final section we summarize our results and give a resume, as well as an outlook.

3 Methods

Two major data sources, which have been used for this study, are the suicide data set and The World Happiness Report. The years we took into consideration were 2015-2017. Unfortunately, within the given time, we weren't able to find a more up-to-date suicide dataset of a comparable quality. The examined data showed a stable development without significant outliers. Therefore, we are confident to use this data for the purpose of our research question.

The suicide data set originates from The Global Burden of Disease Study 2017 (GBD 2017). It is coordinated by the Institute for Health Metrics and Evaluation (IHME) and estimated the burden of diseases, injuries, and risk factors for 195 countries and territories, and at the subnational level for a subset of countries. [2] In particular, we have used the age-sex-specific data for suicide rates, mental diseases and substances abuse. GDB also provides population estimates, which we have used to customize the age groups according to the needs of our research.

The World Happiness Report is a publication of the Sustainable Development Solutions Network, powered by data from the Gallup World Poll. It is a landmark survey of the state of happiness of 156 countries. [3] In particular, the report provides happiness score per country and a possible explanation for it. The scores represent an average country citizen's answer to the question "how would you rate your own current happiness on a 0 to 10 scale". With the help of a correlation analysis the authors estimated the extent to which each of six factors - levels of GDP, life expectancy, generosity, social support, freedom, and corruption - contribute to making happiness evaluations higher[4].

With this data available, we have proceeded as follows. First, we merged the population dataset with the dataset containing absolute suicide numbers, based on the location, age category and sex. Then we combined multiple disjoint age groups into 3 base categories: "under 25", "25 to 49" and "50+". This categorization better reflected our needs than the predefined age groups, since they divide the world's population into groups of comparable size. At this point we had the data needed for the demographic analysis. We hence visualize the data from different points of view. We will present and discuss the results in the section 4.1.

Furthermore, we wanted to be able to compare suicides with cases of depression and substance misuse respectively. Firstly, we had to clean all three datasets by reducing them to only the features necessary for comparison. Secondly, we merged the suicide dataset with the depression and substance misuse dataset respectively, so that only one dataframe with the necessary information was left. As a third step, we wanted to make comparisons on a yearly basis. We therefore created a dataframe for each year by filtering by the feature "year". At this point we had three tables for the suicide-depression-comparison and three for the substance-misuse-depression-comparison. We then calculated the correlation coefficient, the respective p-values and drew a regression line to assess whether there is a relationship between the examined features. You will find the explanation of the terms correlation coefficient, p-value and regression line alongside a discussion of the results in the section 4.3 and section 4.4.

We compared suicide rates, cases of depression and substance misuse with rankings from the World Happiness Report, using correlation matrices and filtering for significance. Afterwards we calculated these correlation matrices for each World Bank region and compared the significant factors. The results can be found in section 4.5, with global results being presented in subsection: 4.5.1, and regional results in subsection 4.5.2.

4 Results

4.1 Demographics

Figure 1 provides the first insights into the demographics of suicide. Here the y-axis represents the suicide rate, i.e. number of suicides committed by people of a particular age group and particular gender per 100.000

people of the same age group and gender.

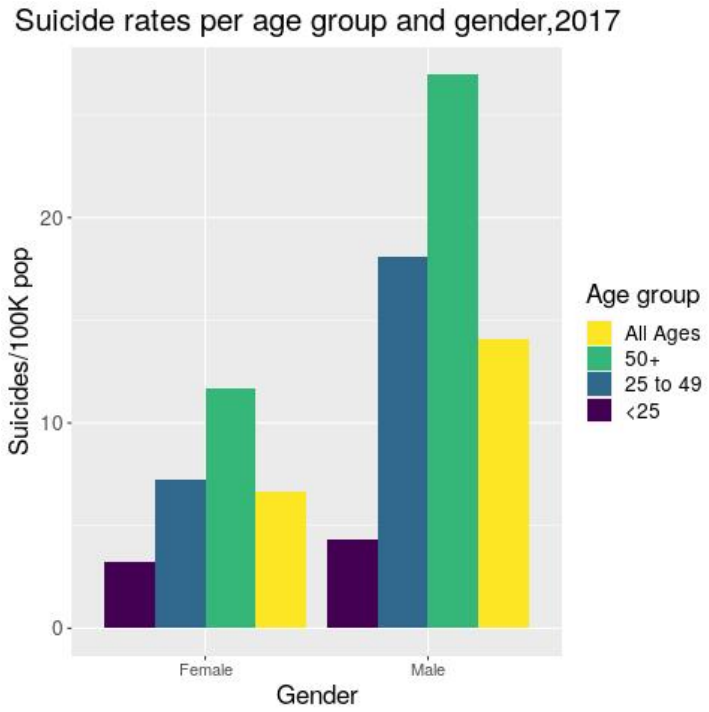

Figure 1: Suicide rates per age and gender, 2017

The key findings from this visualization are following:

- In each age group males are more susceptible to suicide than females are.
- Elder age group, independent of gender, have higher suicide rates.
- Extremely high is the suicide rate of elderly males.

This view gives us an indication that societies with larger proportion of elderly people might have higher overall suicide rates. In order to validate this assumption we consider following two figures. Indeed, the detailed view provided by figure 2 is conform with the findings from the figure 1: the suicide rates for young people is comparably low in all the regions. South Asia has the highest suicide rates among people in the age under 25, followed by North America and Europe & Central Asia.

Besides that, figure 2 confirms that elderly people commit (in some regions significantly) more suicides than people from the younger age categories. The only exception is Middle East & North Africa, where middle aged people are more susceptible to suicide that the elderly, but the difference is insignificant.

Figure 3 shows that the age structure of societies differs significantly among the regions. This information helps to interpret the relations between the height of the bars representing different age groups within a region and the region's average (the "All ages" bar) in the Figure 2. E.g., at first glance it seems strange that Sub-Saharan Africa has such an outstanding suicide rate among elderly people and comparably low overall suicide rates. But it becomes clear after taking a look at this region's population structure: people of age 50+ represent an extremely low portion of the the overall population.

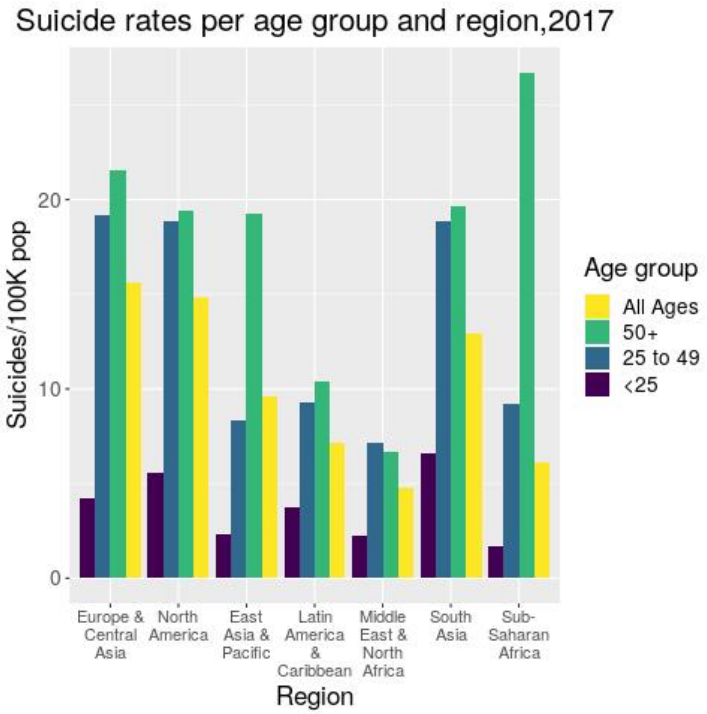

Figure 2: Suicide rates per age and region, 2017

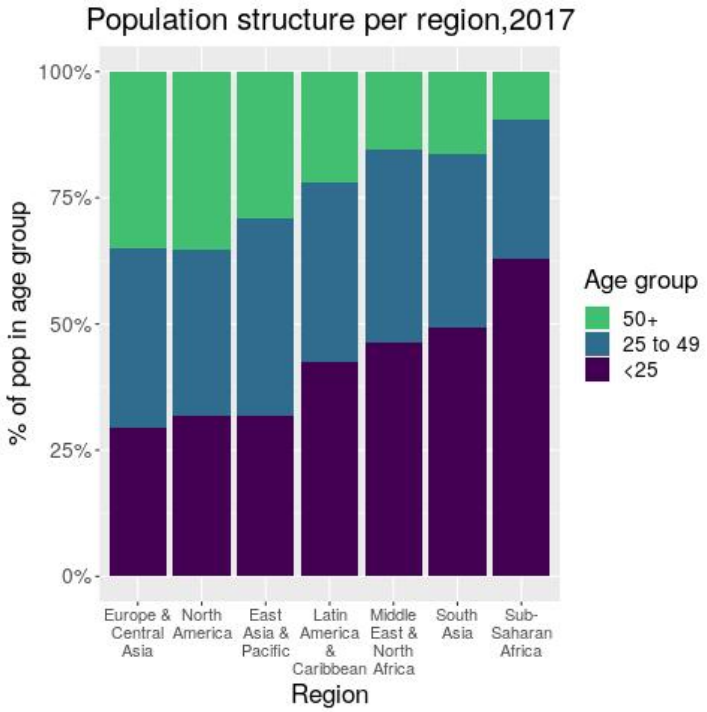

Figure 3: Population structure per region, 2017

Another question could be: why South Asia, having approximately the same suicide rates for the both eldest age groups and higher rate for the youngest age group, still has a lower average suicide rate than North America. Again, figure 3 provides the answer: South Asia has a significantly younger population than North America, hence the younger age group gets more weight in the average suicide rate. And, as we already learned, youngest age group has lower suicide rate than the elder ones.

In order to investigate, whether the fact that suicide rates are higher among males than among females holds for all regions, we consider the figure 4. Indeed, it is true, moreover, in all the regions, except South and East Asia, the ratio is more than 2:1.

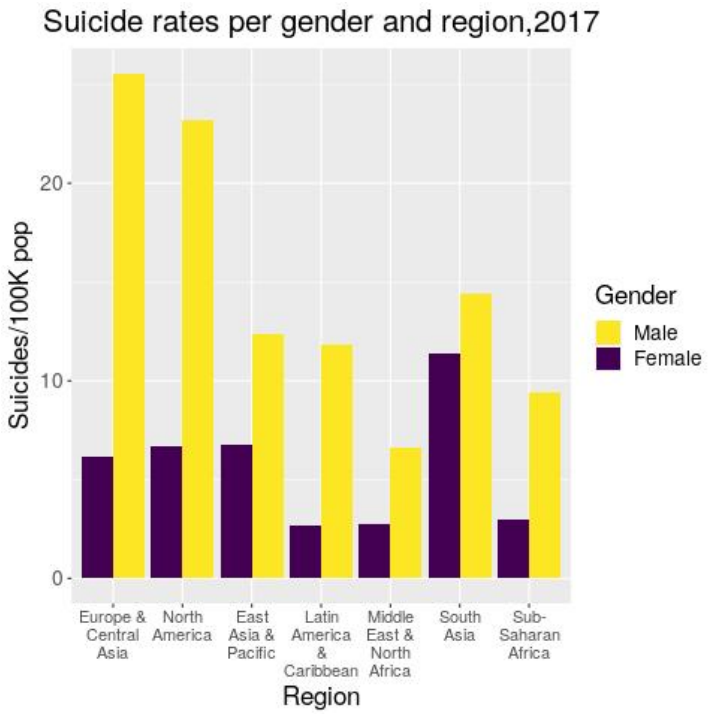

Figure 4: Suicide rates per age and gender, 2017

4.2 Explanation of statistical terms

For those who aren't familiar with statistics yet, we will give a brief explanation of the aforementioned statistical terms.

The correlation coefficient describes a statistical relationship of two values x and y. The correlation coefficient R is a real number between -1 and 1. -1 indicates a very strong negative correlation between x and y. 0 indicates no correlation. 1 indicates a strong positive correlation. A negative correlation can be described as a relation between the values x and y, in which a high x indicates a low y and vice versa. A positive correlation can be described as a relation between the values x and y, in which a high x indicates a high y.

The p-value describes whether or not a correlation is significant. You can try to perceive the p-value as the value that tells you whether or not your correlation coefficient would withstand a scientific statistical assessment. If your p-value is lower than 0.05 it will withstand. If it is higher or equal 0.05 it will not.[5]

A linear regression has the attempt to model a relationship between two variables. It does so by fitting a linear equation to the data that has been observed. In this context fitting can be understood as the process of finding a linear model, that minimizes the overall distance from itself to the observed data points.[6]

4.3 Depression

In order to answer the question whether or not there is a relation between the suicide rate and the depression rate we calculated the correlation coefficient, the respective p-value and a regression line between these two features. While the suicide rate has been introduced as the the number of suicides per 100,000 inhabitants, the depression rate hasn't been introduced yet. The measure that accounts for the depression rate is the so called point-prevalance. In this case point-prevalance can be described as the number of people who were diagnosed with depression per 100,000 inhabitants at a specific point in time.
Now as the terms have been described and explained we can take a closer look into the figure 5. The variable R describes the correlation coefficient and the variable p describes the respective p-value. With a value of 0.52 there is a moderate, significant relation between suicide rate and depression. This relation helps to support the believe that depression might be an influential factor on whether or not people commit suicide.

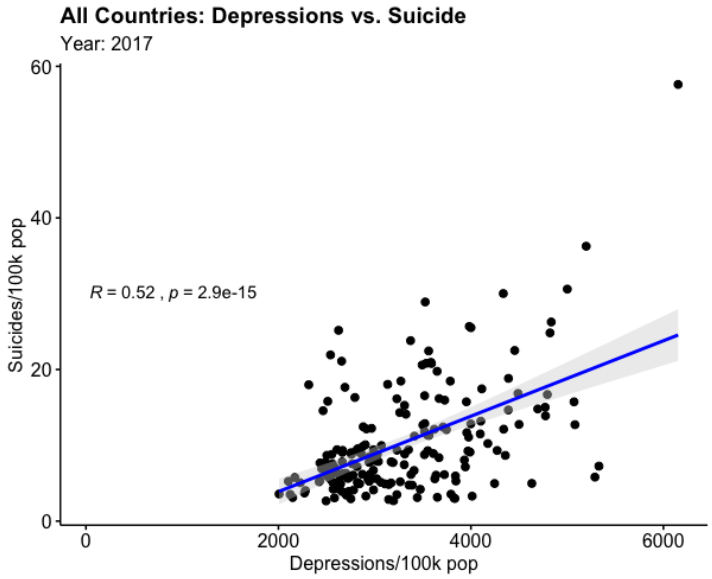

Figure 5: Correlation of Depression rate and suicide rate, 2017

4.4 Substance misuse

Another factor which might have a statistical relationship with the suicide rate is the rate of people with a substance abuse disorder. The measure that accounts for rate of people with a substance abuse disorder is point-prevalance and has been introduced in the section 4.3. Substances include alcohol, cannabis, cocaine and opioids amongst others. We recognize that in figure 6 there is a moderate, significant correlation of 0.47. That isn't very surprising given the fact that there is also a moderate correlation between the depression rate and the rate of people with a substance misuse disorder as one can see from figure 7.

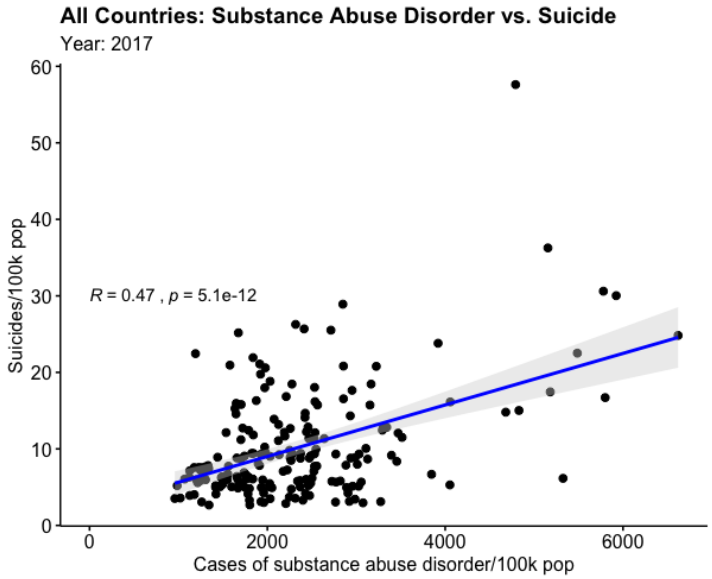

Figure 6: Correlation of Depression rate and suicide rate, 2017

4.5 Suicides,Disorders & World Happiness Report Rankings

To investigate the correlations between suicides, disorders, and the World Happiness Report (WHR), correlation matrices were built.
The parameters for the WHR Rankings are:

- Happiness Score
- Log GDP per capita
- Social support
- Healthy life expectancy at birth
- Freedom to make life choices
- Generosity
- Perceptions of corruption
- Confidence in national government
- Democratic Quality

The parameters from suicides and disorders are:

- Depression/100k pop
- Suicides/100k pop
- Substance Misuse/100k pop

For the given timeframe, we chose to investigate correlations between suicides and WHR rankings as well as suicide and depression/substance misuse. This report leaves out the assessment of the correlations between disorders and WHR rankings, as well as the correlation of suicide and disorders other than depression/substance misuse.

4.5.1 Global: Suicides,Disorders & World Happiness Report Rankings

When taking a look at the global correlations between suicides and WHR rankings in figure 7, we can see that with p-value < 0.05, there are positive correlations for the ranking factors *Happiness Score*, *Log GDP per capita*, *Social Support*, *Healthy life expectancy at birth*, and *Democratic Quality* and a negative correlation for ranking factor *Confidence in national government*.

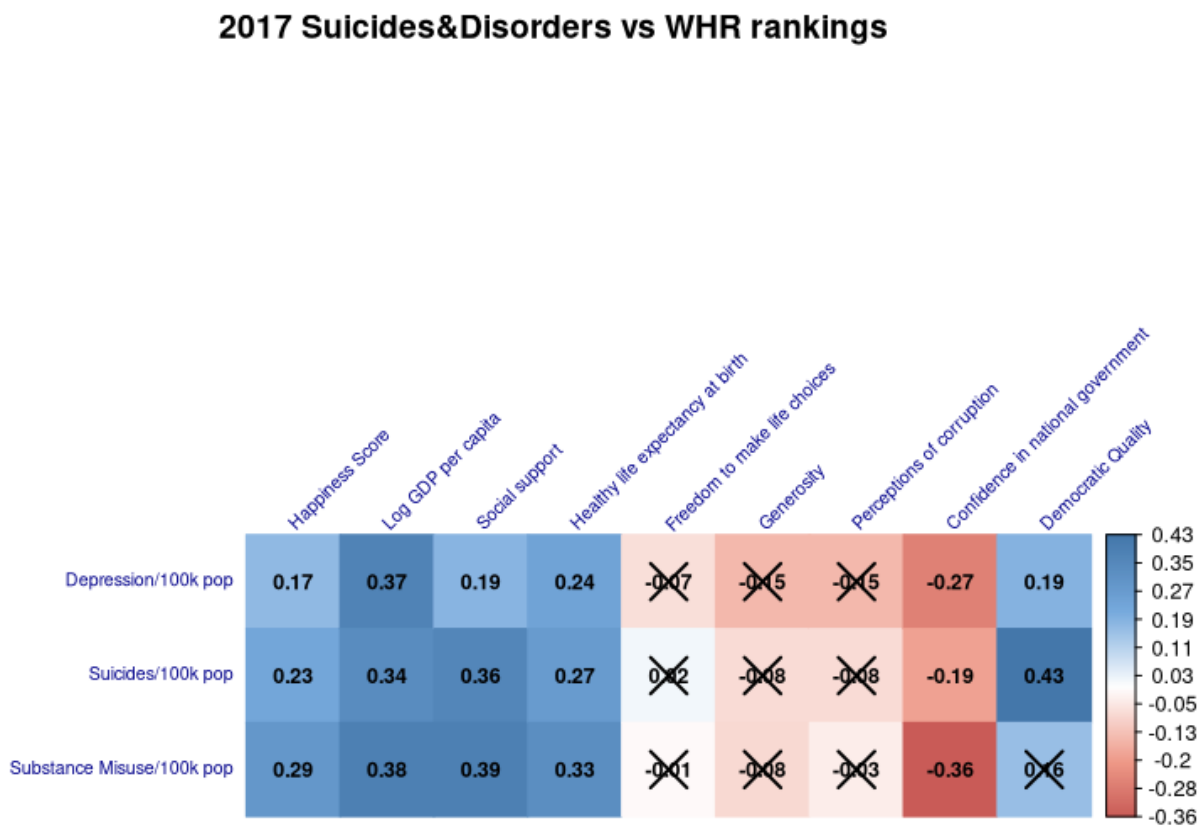

Figure 7: Suicides&Disorders vs WHR rankings, 2017

4.5.2 Regional: Suicides,Disorders & World Happiness Report Rankings

In order to investigate regional differences between suicide rates and WHR rankings, we assembled correlation matrices for each WB region and their respective WHR rankings.
Figures 8 to 13 show the correlation matrices. Correlations with a p-value > 0.05 have been blanked out.

- Figure 8 shows East Asia & Pacific, positive correlations for the ranking factors *Log GDP per capita*,

Healthy life expectancy at birth, and *Democratic Quality*, negative correlations for the ranking factors *Freedom to make life choices* and *Confidence in national government*.

- Figure 9 shows Europe & Central Asia, positive correlation for the ranking factor *Social Support*.
- Figure 10 shows Latin America & Caribbean, positive correlation for the ranking factor *Democratic Quality*.
- Figure 11 shows Middle East & North Africa, no significant correlations.
- Figure 12 shows South Asia, positive correlations for the ranking factors *Log GDP per capita*, and *Democratic Quality*.
- Figure 13 shows Sub-Saharan Africa, negative correlation for the ranking factor *Generosity*.

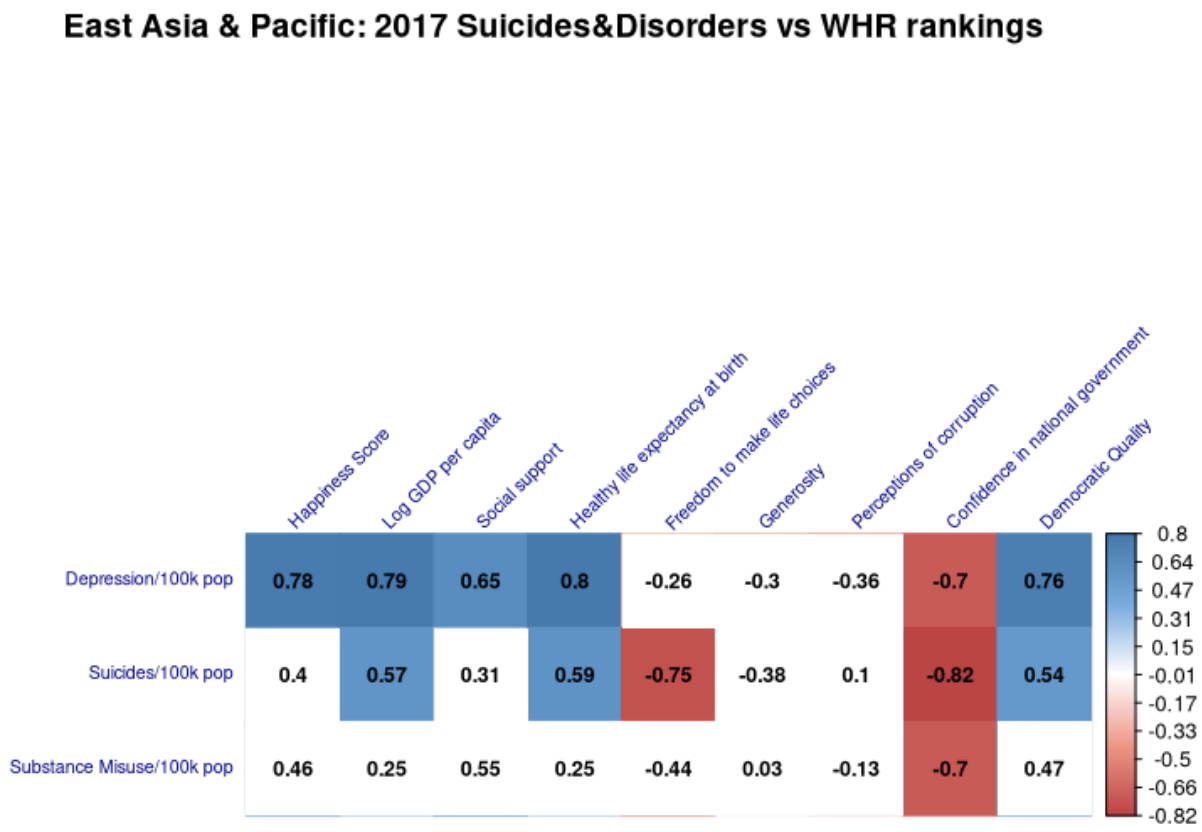

Figure 8: East Asia & Pacific: Suicides&Disorders vs WHR rankings, 2017

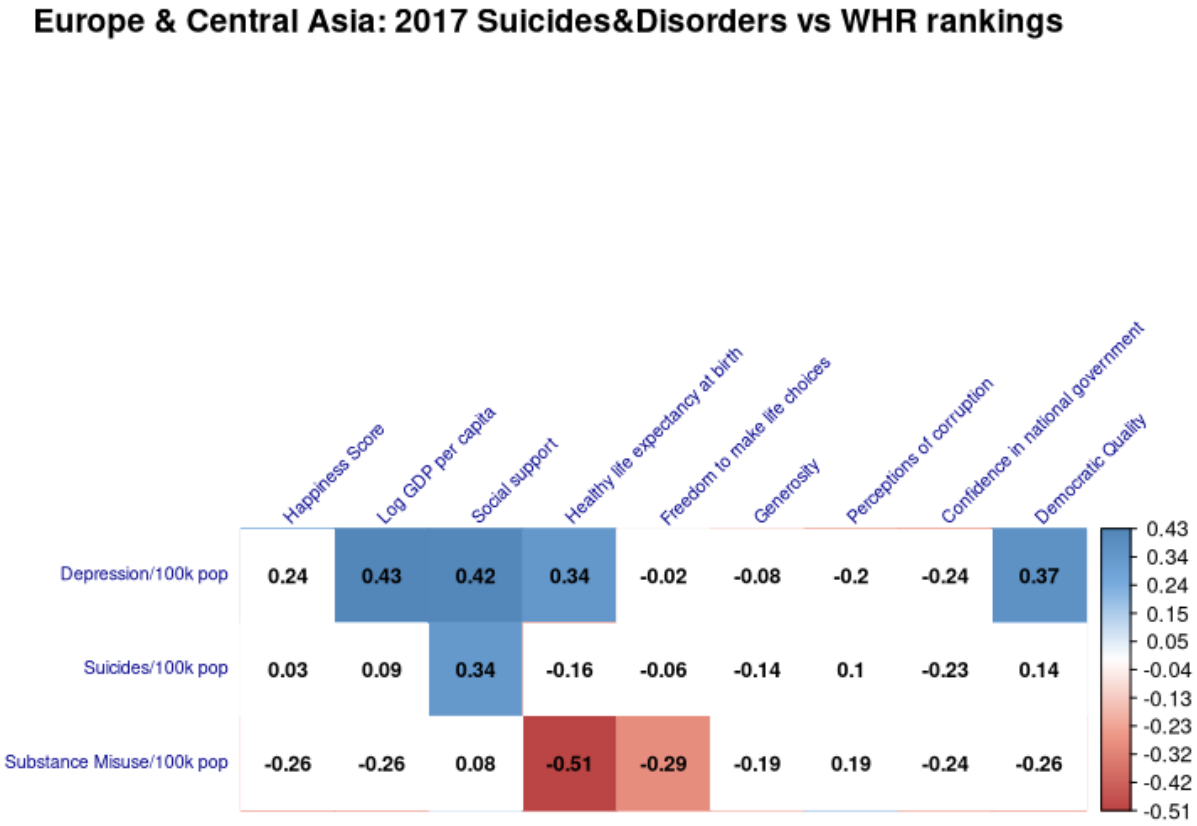

Figure 9: Europe & Central Asia: Suicides&Disorders vs WHR rankings, 2017

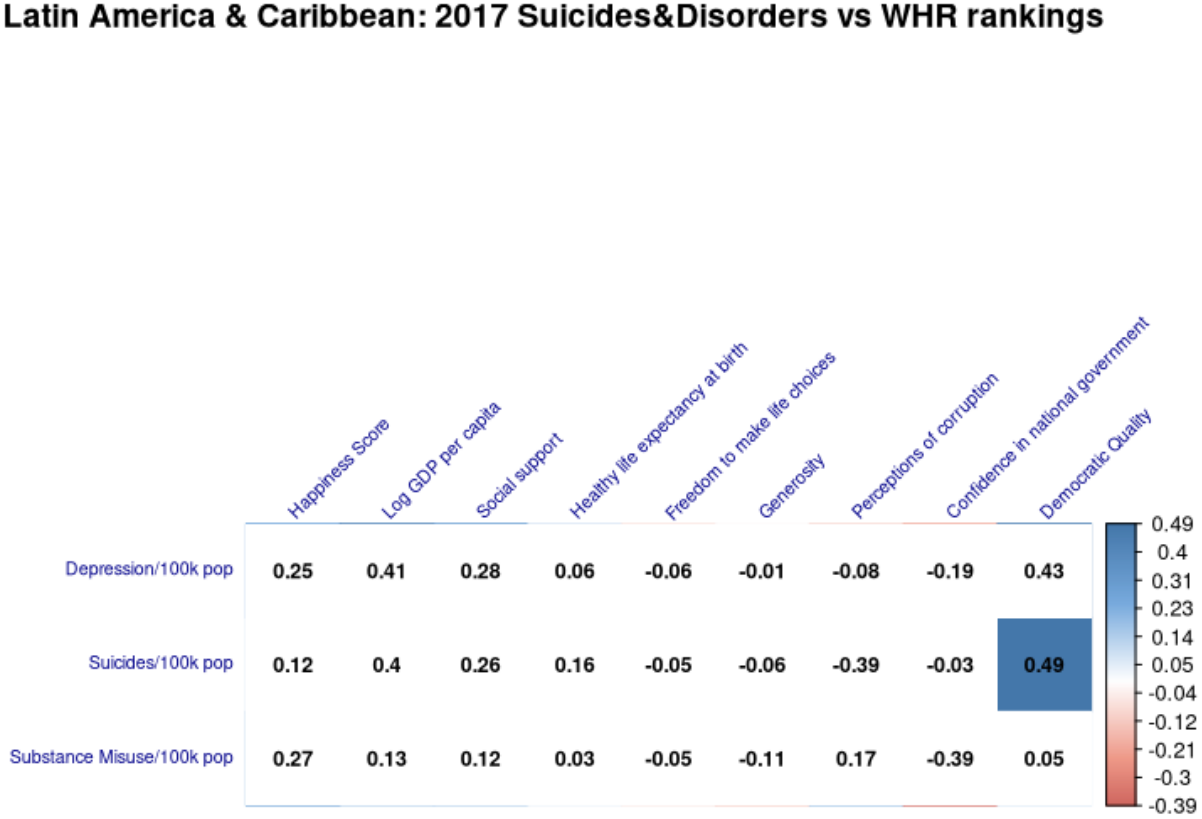

Figure 10: Latin America & Caribbean: Suicides&Disorders vs WHR rankings, 2017

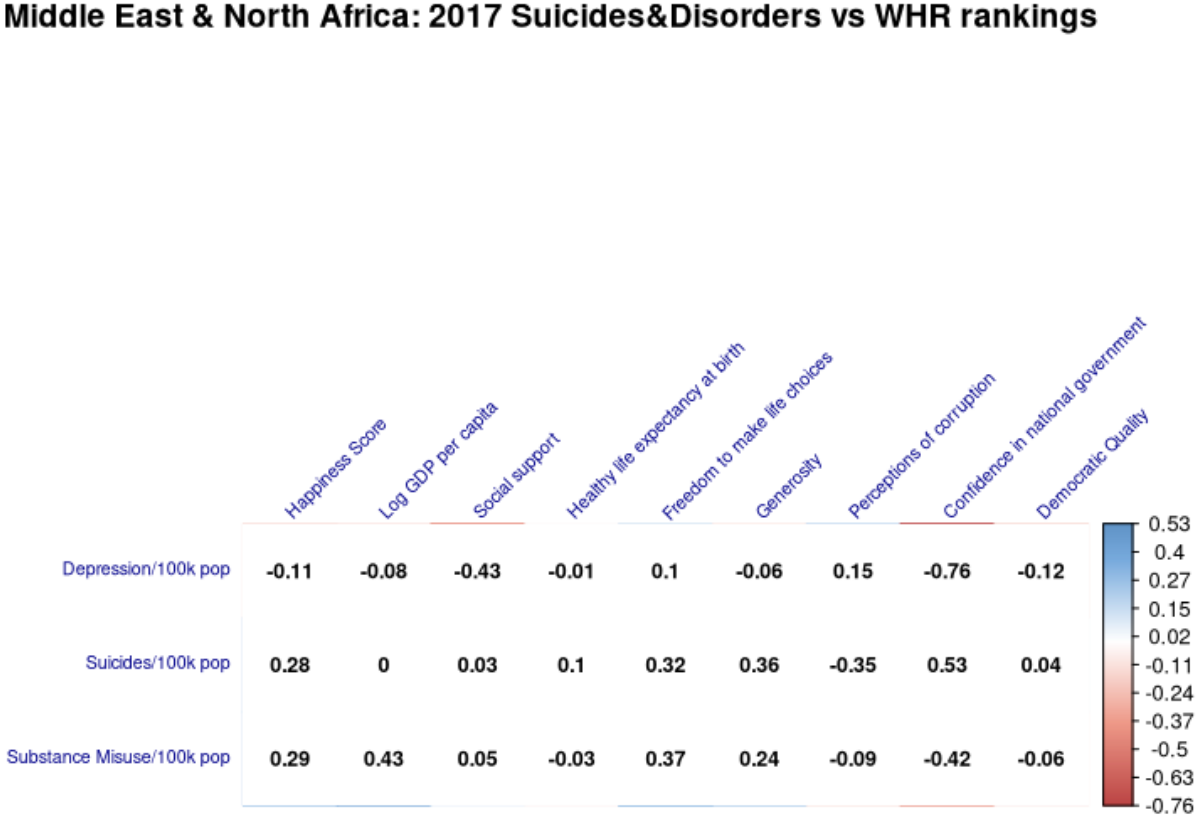

Figure 11: Middle East & North Africa: Suicides&Disorders vs WHR rankings, 2017

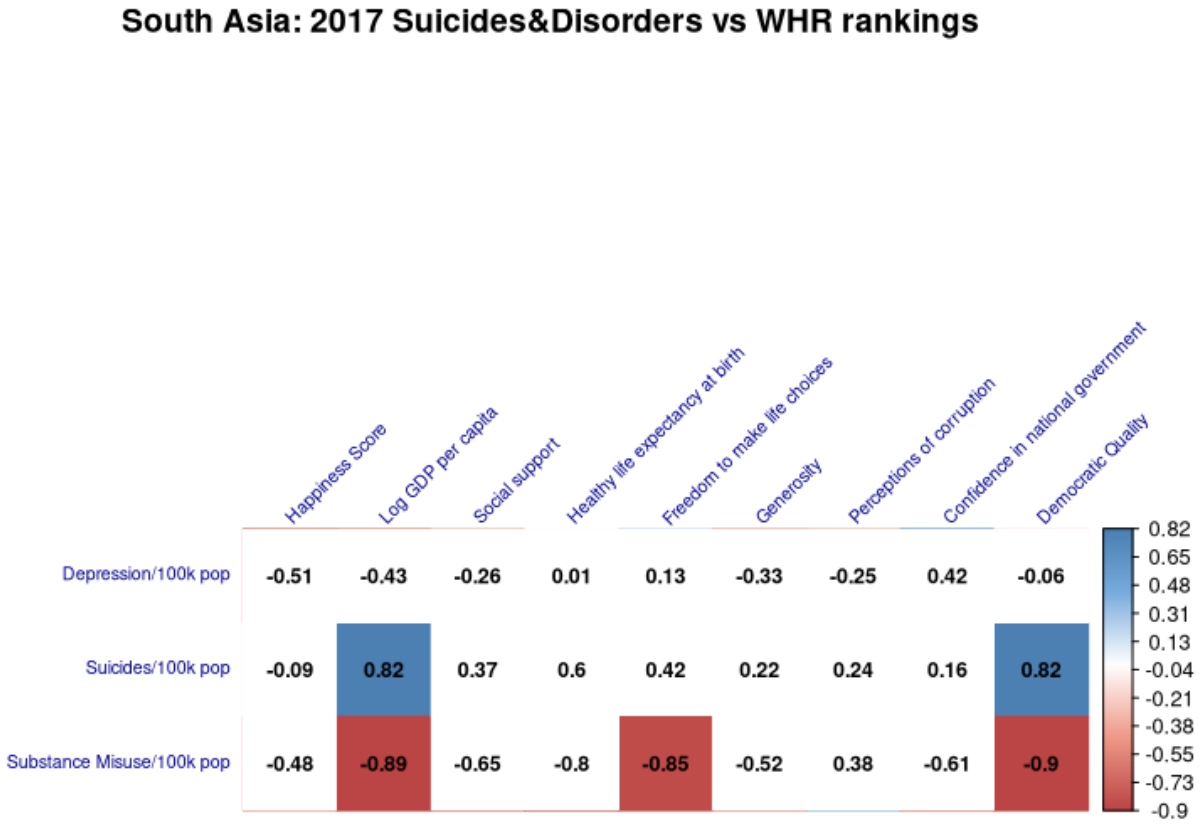

Figure 12: South Asia: Suicides&Disorders vs WHR rankings, 2017

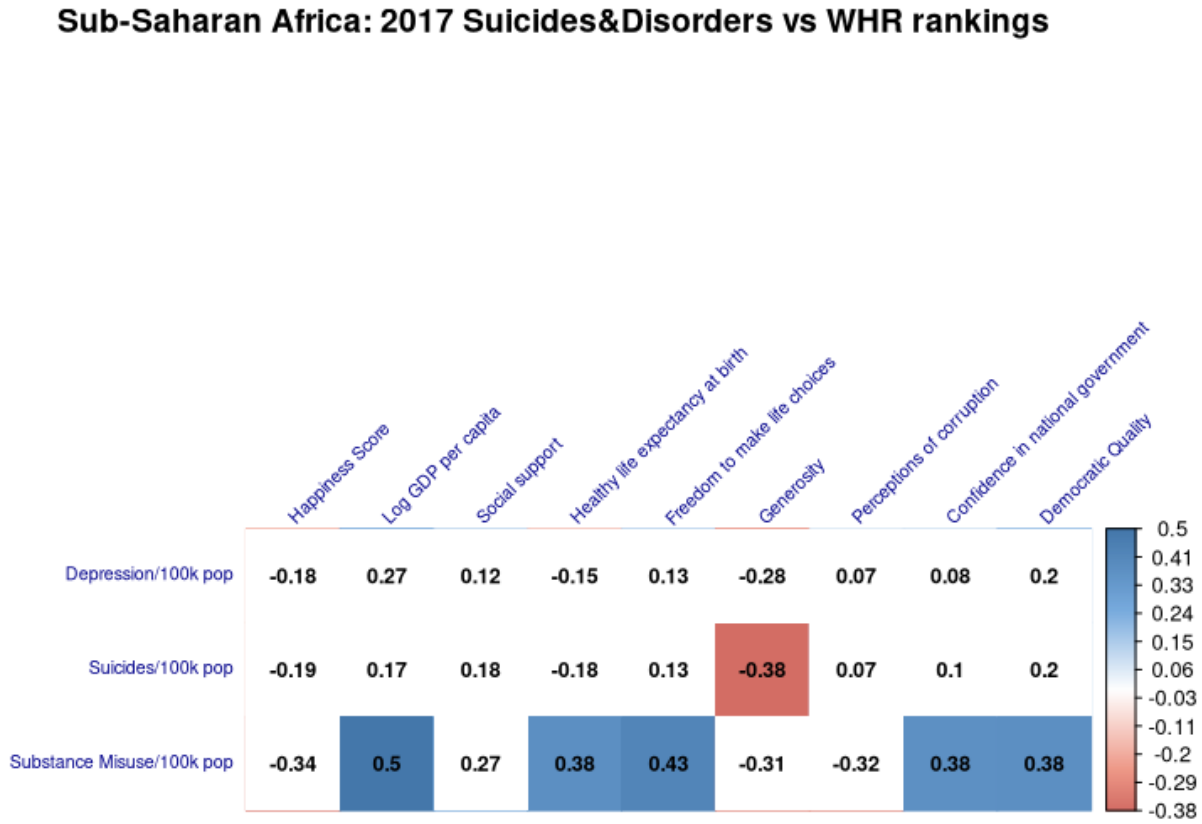

Figure 13: Sub-Saharan Africa: Suicides&Disorders vs WHR rankings, 2017

5 Discussion

Summing up the findings from our research, we can state, that even though the distribution of suicide rates differ between regions, still there are some commonalities. Such as, elderly people tend to commit more suicides than the youngest part of the population. For 5 out of 7 regions, they are closely followed by the middle aged people. Concerning the interconnection between gender and suicide rate, there are strong evidences that males are much more susceptible to suicide than females in all the regions. It has also been shown that suicide is strongly related to mental diseases such as depression or substance abuse disorder.

The results of the regional analysis show, that the significance of correlation between suicides and WHR rankings varies regionally. Comparing global and regional correlation matrices, the significant WHR ranking factors are similar, with exception of WB region *Sub-Saharan Africa*, where factor *Generosity* had a significant negative correlation with suicides.

Of course there are certain limitations to this paper that should be discussed. This paper doesn't use

advanced statistical methods. Therefore, we cannot predict or prevent suicides based solely on our findings, but they definitely support transparency about the issue. Besides that the data lacks transparency and up-to-dateness. In our opinion the data lacks transparency because we cannot say with 100 percent certainty that the data is unbiased. There are autocratic countries such as China who's statements about numbers of suicides or depressions might be questionable. We decided to put our faith in the sources as they are, since the experts composing this dataset are in a far better position to judge whether or not these numbers are trustworthy. A lack of up-to-dateness is due to the fact that the data is obviously not from 2020 or 2019. Nonetheless we were able to explain why it is reasonable to use our data.

Lastly, this paper is not able to show the exact causes of suicide. This is not only due to limitations of time, resources and academic experience but also because this report shows strong indications that there is no such thing as "the major cause" of suicide. We were able to show that mental health, suicide and socio-economic factors influence each other. Depicting the mutual influences in detail would exceed the goal of this report. Instead, we would like to use this report as a call to address the very present issue of suicide by encouraging every reader to go out there and inform themselves about suicide. It might save lives.

References

[1] W. H. Organization, "Suicide in the world: global health estimates," technical documents, 2019.

[2] http://ghdx.healthdata.org/gbd-2017.

[3] https://worldhappiness.report/ed/2020/.

[4] https://worldhappiness.report/faq/.

[5] https://www.dummies.com/education/math/statistics/how-to-calculate-a-regression-line.

[6] http://www.stat.yale.edu/Courses/1997-98/101/linreg.htm.