Suicide Rates

Data Science Basics Project

Group 8 (P. Bucher, F. Meyer, L. Schuler, M. Thalmann)

December 8, 2019

Abstract

The project *Suicide Rates* aims at answering the question, whether or not there is a correlation between the GDP[1] per capita and the number of suicides in a country. The underlying data sets not only contains yearly GDP and suicide rate data, but also breaks down the suicide rate by generation and sex, and further provides the HDI[2] and population size of the respective year, which makes it possible to correlate suicide rates to other socio-economic parameters as well.

TODO: what has been found out?

Research Question

Based on the information available we assume there to be a correlation between GDP per capita and the number of suicides in a country.

Related Questions

1. There should also be different suicide rates depending on generation.

2. The countries can also be compared with each other.

---

[1]Gross Domestic Product
[2]Human Development Index

Contents

1 Data

The underlying data (see `master.csv`) set has been provided by Kaggle[3] and provides data on suicide rates of 102 countries[4] from 1985 to 2016.

Figure 1 *Available data per Country and Year* (p. 4) shows the data set is all but complete. The dark tiles show years for which data of a country is available, light tiles represent missing data. Some countries have very few data points, whereas others have data for most of the years. The less yearly data is available, the harder an assessment of the development for a single country becomes.

The same data is summarized in figure 2 *Histogram Showing the Number of Measurements per Country* (p. 5), which quantifies the number of observations that are available per country.

Figure 3 *Numbers of Years with Suicide Data Provided per Country* (p. 6) provides summarized insight into invidivual countries.

---

[3] `https://www.kaggle.com/russellyates88/suicide-rates-overview-1985-to-2016/download`
[4] Executing `cut -d, -f1 master.csv | sort | uniq | wc -l` on a Unix shell counts the number of countries in the CSV file.

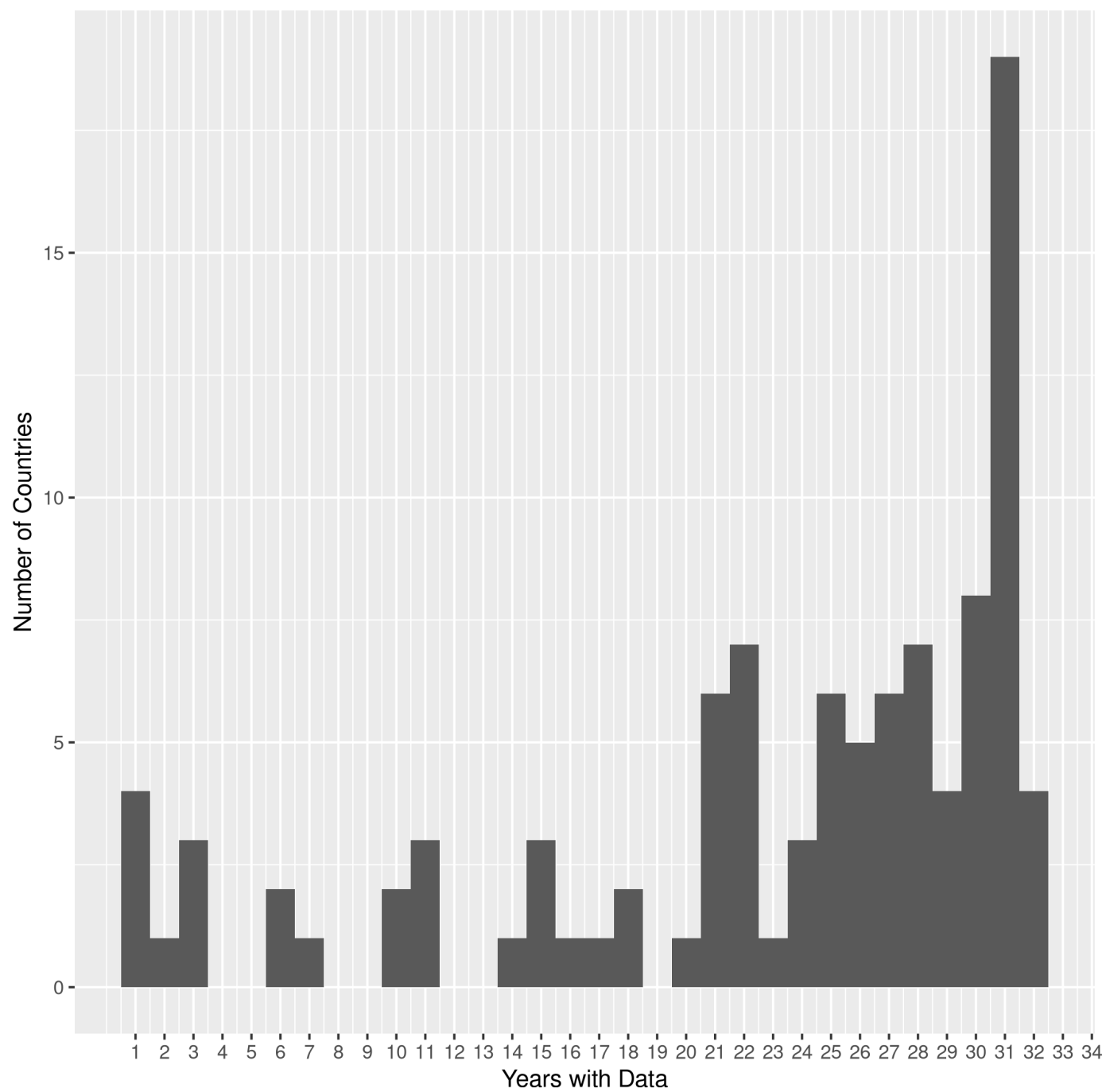Figure 1: Available data per Country and Year

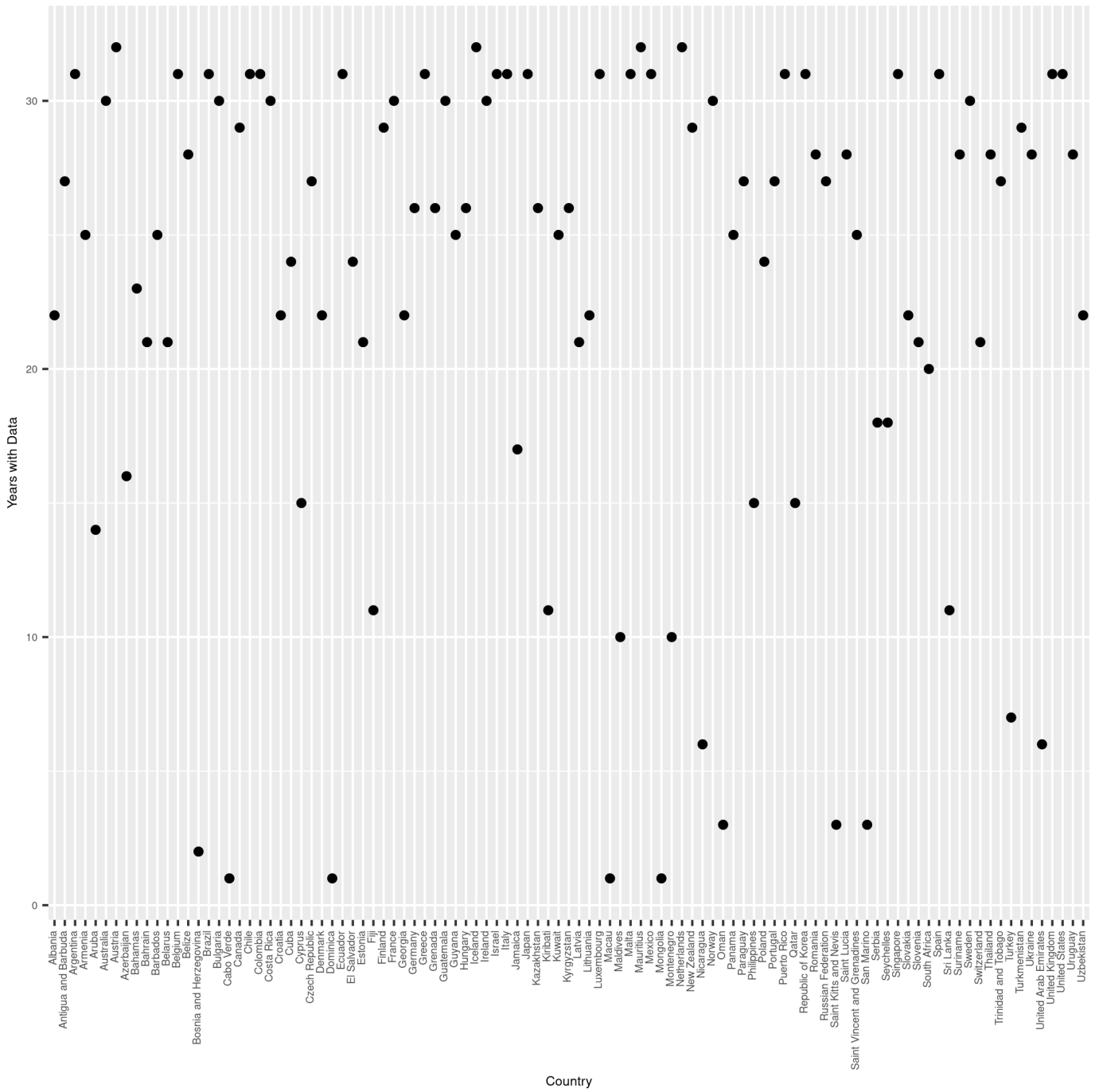Figure 2: Histogram Showing the Number of Measurements per Country

Figure 3: Numbers of Years with Suicide Data Provided per Country
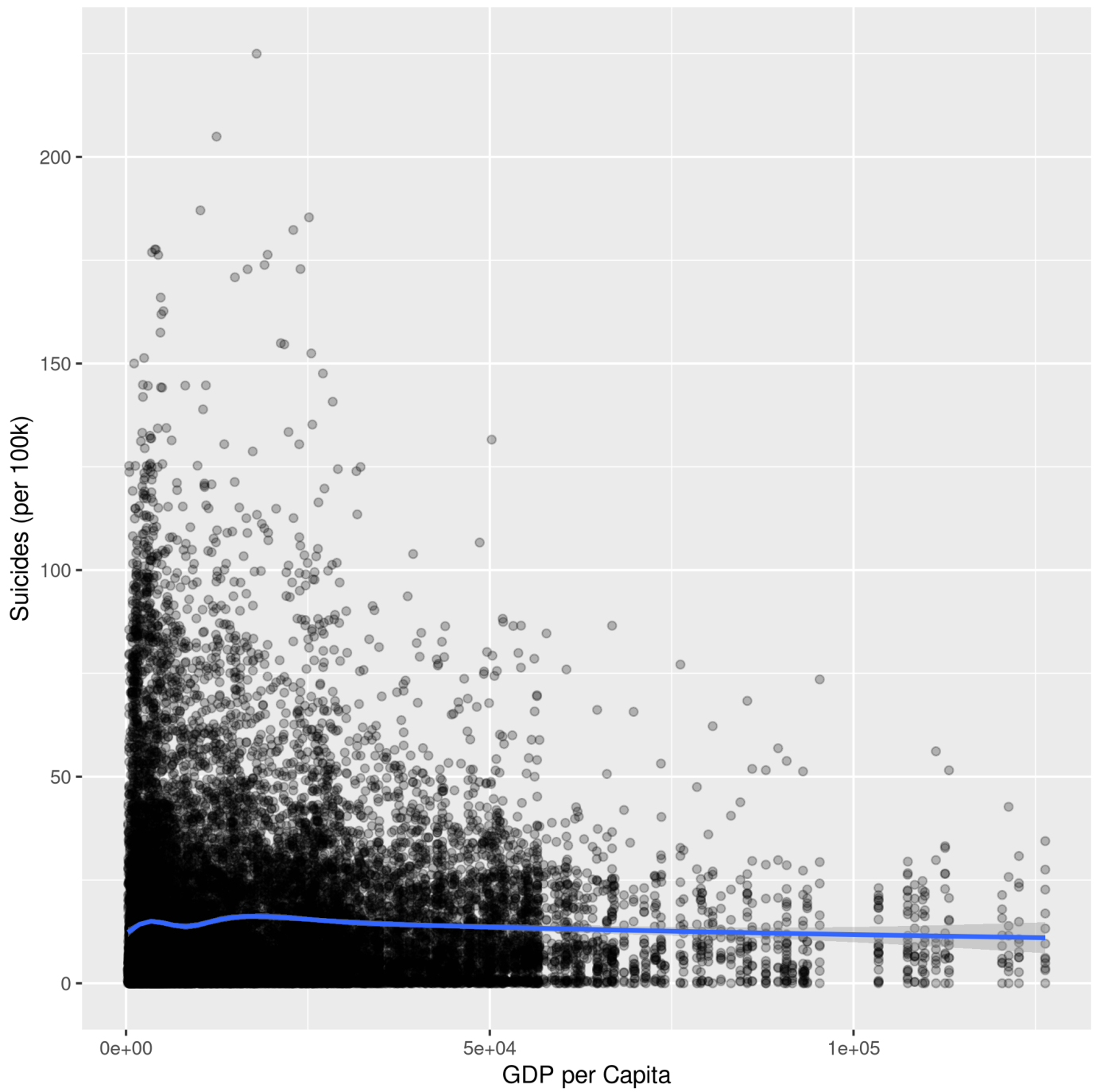
2 Analysis

Figure 4: Suicide Rate per 100k Inhabitants with Regression Curve

List of Figures