BDA Final Project - Evaluation of the relationship between mean income and suicide rate.

Anonymous

Contents

Introduction

Motivation

An oft-wondered and oft-debated question for people, for as long as the concept of money has existed, is whether having money makes a person happier. While happiness is a tricky thing to quantify, extreme sadness can readily be quantified by looking at the suicide rate for a particular group of people. This brings us to the rather macabre question we have chosen to examine for this project: is there a relationship between average income and the suicide rates for groups of people?

The problem and modelling idea

The relationship between the mean income and the suicide rate was modeled as a linear regression problem, and this was done in three different contexts, so that we can try to get some kind of statistically sound answer to the question posed above. The first context is for the 9 different broad groups of professions in the United Kingdom, sourced from the government database for the year 2019. The second is for same, but for 22 groups of professions in the United States. The final context is an international one, the suicide rates were seen for the 93 countries of the world with a population >100k, and the mean incomes for those countries found for the year 2019. In all three cases, the mean income was taken in dollars and standardized by the cost of living for the corresponding country, in order to be able to club such different data together.

Prior art

The US data was examined by the website registerednursing.org, and a single linear trend was fitted with no statistical information or information about the methods. Reviewing past academic publication reveals that similar studies were done with the data in previous years for South Korea [1] and Denmark [2]. The former calculated the suicide hazard ratios for each income bracket to a certain confidence level, while the latter included more factors in the analysis such as age and gender, transforming the problem.

The data

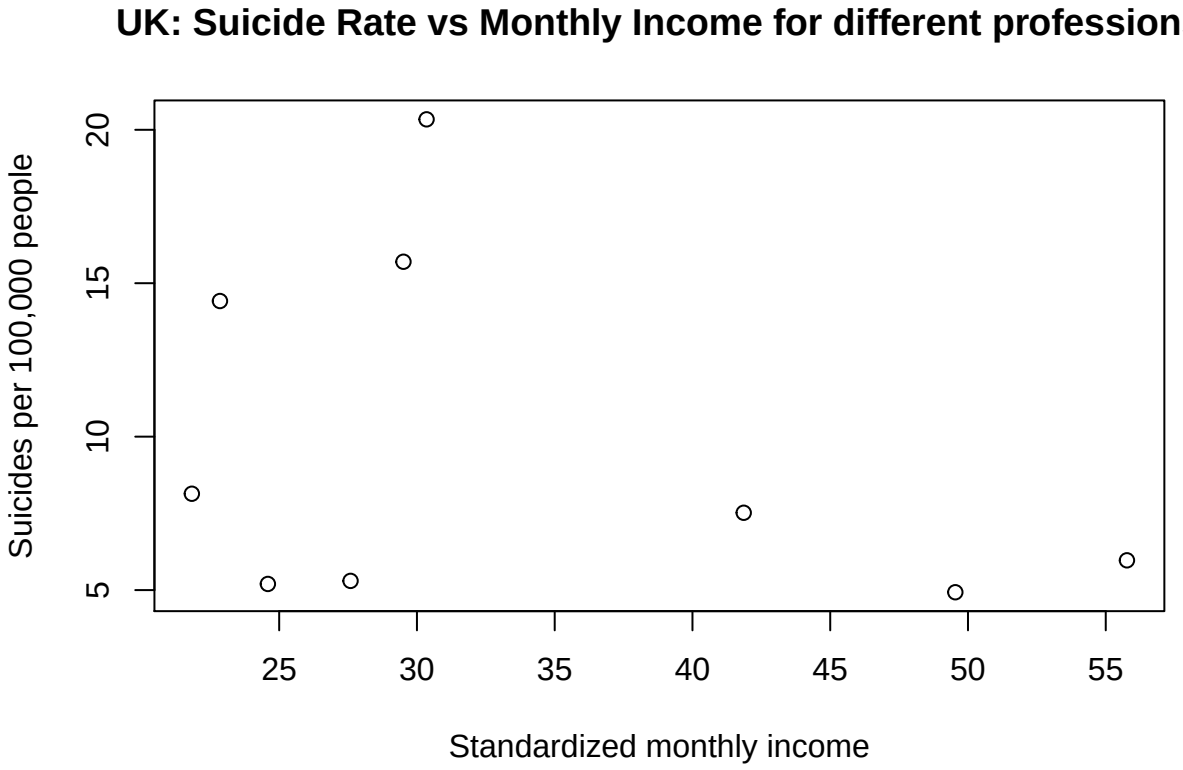
Our data was not retrieved from some common source, but rather from different sources and they were adjusted so that they were coherent together.

United Kingdom data

The data for both the suicide rate (per 100,000 people) and the mean weekly income (which was converted into monthly income, into dollars, and standardized by the cost of living) were both retrieved for the same 9 profession groups from UK Office for National Statistics (ons.gov.uk). The 9 occupation groups selected are

1. Managers, directors and senior officials
2. Professional occupations
3. Associate professional and technical occupations
4. Administrative and secretarial occupations
5. Skilled trades occupations
6. Caring, leisure and other service occupations
7. Sales and customer service occupations
8. Process, plant and machine operatives
9. Unskilled occupations / elementary operations

```
ukdata =read.csv('UKdata.csv');
plot(ukdata$wagepercol, ukdata$suicides100k, xlab='Standardized monthly income', ylab = 'Suicides per 1
```

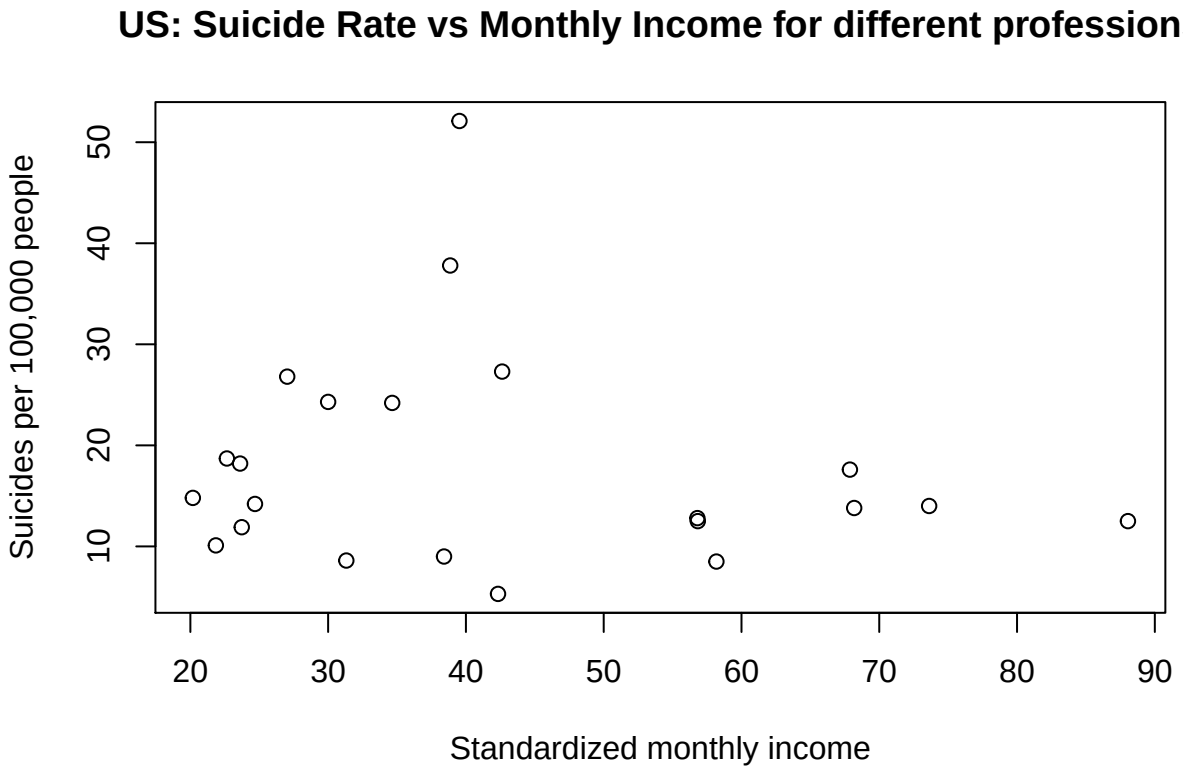


United States data

The data for the suicide rate (per 100,000 people) for 22 broad profession groups was retrieved from the Maerican Centre for Disease Control (CDC), and the mean monthly income (which was standardized by the cost of living) was retrieved for the same 22 profession groups from US Bureau of Labor Statistics (BLS). The 22 occupation groups selected are

1	Management occupations	12	Protective service occupations
2	Business and financial operations occupations	13	Food preparation and serving related occupations
3	Computer and mathematical occupations	14	Building and grounds cleaning and maintenance occupations
4	Architecture and engineering occupations	15	Personal care and service occupations
5	Life, physical, and social science occupations	16	Sales and related occupations
6	Community and social service occupations	17	Office and administrative support occupations
7	Legal occupations	18	Farming, fishing, and forestry occupations
8	Educational instruction and library occupations	19	Construction and extraction occupations
9	Arts, design, entertainment, sports, and media occupations	20	Installation, maintenance, and repair occupations
10	Healthcare practitioners and technical occupations	21	Production occupations
11	Healthcare support occupations	22	Transportation and material moving occupations

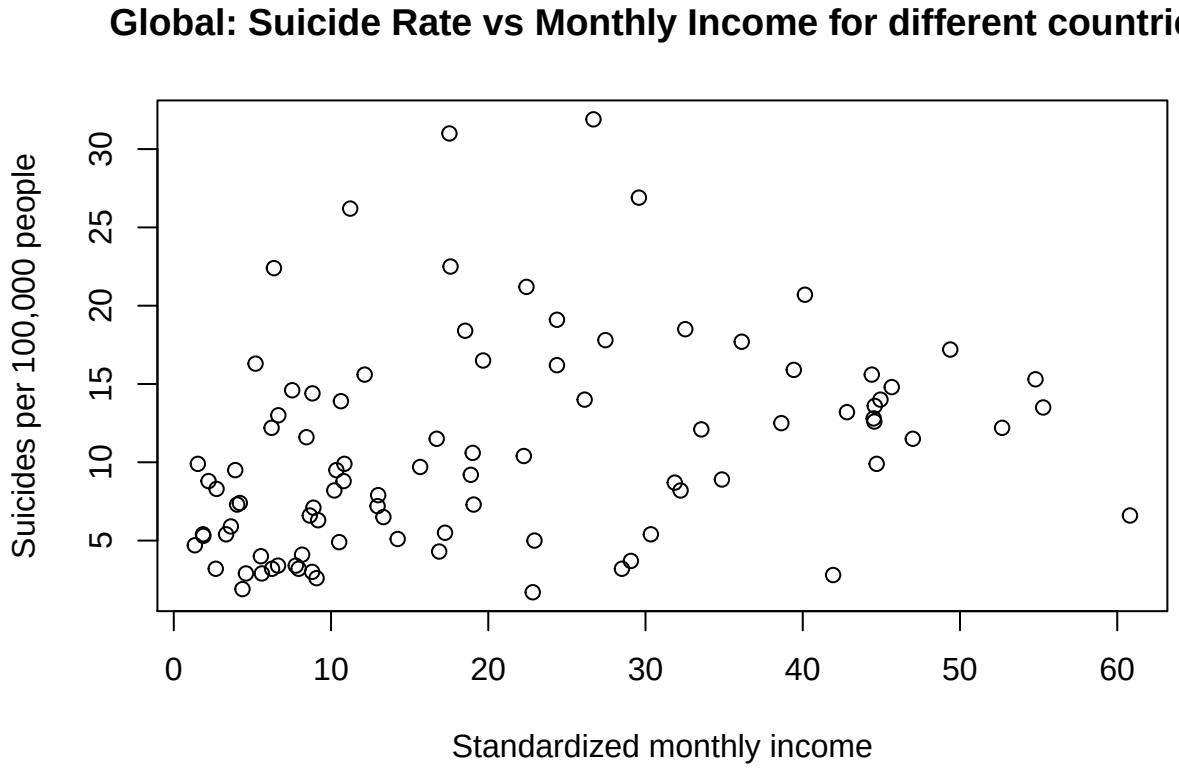
```
usdata =read.csv('USdata.csv');
plot(usdata$wagepercol, usdata$suicides100k, xlab='Standardized monthly income', ylab = 'Suicides per 1
```



Global data

The data for both the suicide rate (per 100,000 people) and the mean monthly income (which was converted into dollars, and standardized by the cost of living of each country) were both retrieved for the 93 countries which have a population greater than 100,000 from the World Population Review website (worldpopulationreview.com). This is quite a different context from the profession-wise data for the countries above, as it is instead countrywise. The standardizing by the cost of living of each country, which was also obtained from the World Population Review website, was our attempt to render the comparison as a fair one.

```
countrydata =read.csv('countrydata.csv');
plot(countrydata$wagepercol, countrydata$suicides100k, xlab='Standardized monthly income', ylab = 'Suici
```



This is the most complicated context in which this relationship is examined. It must be noted that there are numerous factors contributing to the suicide rates of different countries besides the income standardized by cost of living (Which is similar to the Purchasing Power of the country). In addition to the level of poverty, there are cultural factors (more emphasis on personal happiness, varying sense of familial responsibilities), geographical factors (failing agricultural produce, larger periods of darkness) and several possible other factors (political unrest, war). Therefore the problem in this context is quite simplified, and on cursory viewing of the figure above, it actually appears that there is positive correlation between wealth and income, which on first thought might seem unexpected.

Methodology

References:

[1] Qin, Ping, Esben Agerbo, and Preben Bo Mortensen. "Suicide risk in relation to socioeconomic, demographic, psychiatric, and familial factors: a national register–based study of all suicides in Denmark, 1981–1997." American journal of psychiatry 160.4 (2003): 765-772. [2] Lee, Sang-Uk, et al. "Suicide rates across income levels: retrospective cohort data on 1 million participants collected between 2003 and 2013 in South Korea." Journal of epidemiology 27.6 (2017): 258-264.