World Happiness, Suicide and Homicide/RStudio

**Author**

Kathrine Kvist Jensen

**Abstract.**

In this project, I examine the connection between the happiness of the country, based on data from the World Happiness Report[[1]](#footnote-1), and the tendencies in suicide and homicide, based on data from Our World in Data[[2]](#footnote-2). In order to do that, I use different types of software, including Microsoft Excel, Rstudio and OpenRefine.

The project shows, that from a small sample size view, there is a big difference in homicide rates, depending on which end of the happiness rank a country is, but that the difference in suicide rates is small. This could mean that there is a connection between homicide rates and the happiness of a county, but no connection between the suicide rates and the happiness of a country.

**Keywords:**

*Happiness; Suicide; Homicide;*

### 1 Introduction / Goal

Suicide and homicide are something that happens all around the world. However, the number of cases varies widely. But what is the reason behind the number of cases in each country? Can the county’s condition in form of economy, safety etc., have an effect on the tendencies in suicide and/or homicide? Can the country’s overall happiness or unhappiness have an effect on these numbers?

It is the last question that I am interested in researching. My motivation for doing this, is to explore the explanations for the rates of suicide and homicide, and hopefully (after more research), find a way to lower the rates.

My first thoughts starting this project, was that there was an obvious link between the happiness of a country (or lack thereof) and the rates of suicide/homicide. Therefore, I chose to look at the World Happiness Report and examine if there is a link, or if I have to look in another direction.

### 2 Problems and Background / Context

With the dataset from Our World in Data one can also explore the connection between suicide/homicide and the rising normality of discussing mental health more openly. This project and the other possibilities within this subject, could be a starting point of a discussion on how to get happier people with lower suicide and homicide rates. To further expand the research, one could also look at why the countries rise or fall in rank on the World Happiness Report; Is it because of events in a country (war, terror, disease, politics) or is it the other countries that rise or fall that affects.

With this project, I hope to start a bigger debate than the one currently going on, in a way that more everyday people or students, can come aware of the problems of suicide/homicide rates. With that more people could contribute to find a solution to make more people in the world happy and safe. To broaden the discussion, the use of digital tools and methods are necessary to make the research available all around the world.

### Software Framework

For this project, I have used my 7-year-old HP Spectre x360 Convertible 13-ac0XX, 8 Gb RAM, which runs Windows 10 operating system. I worked in the desktop version of R (4.0.3) and RStudio (3.167). Furthermore, I used Microsoft Excel from Microsoft Office Home and Student 2016.

### 4 Data Acquisition and Processing

**Sources of data used in this paper:**

Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2013. World Happiness Report 2013. New York: UN Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2013/>

Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2015. World Happiness Report 2015. New York: Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2015/>

Helliwell, J., Layard, R., & Sachs, J. (2016). World Happiness Report 2016, Update (Vol. I). New York: Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2016/>

Helliwell, J., Layard, R., & Sachs, J. (2017). World Happiness Report 2017, New York: Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2017/>

Roser, Max and Hannah Ritchie (2013) - "Homicides". *Published online at OurWorldInData.org.* Retrieved from: <https://ourworldindata.org/homicides>

Guide to RStudio, Data Carpentry: <https://datacarpentry.org/r-socialsci/>

**Details of data extraction, filtering and preparation.**

For the execution of this project, I intend to make a graph of each year, with the 10 happiest and the 10 unhappiest countries. Then I will see if there is a link between the happiness of the countries and the suicide/homicide rates.

To do this, I started by looking at the World Happiness Report to find the happiest and unhappiest countries for each year. The rankings can be found in the reports itself and as an appendix. In the report the ranking can be found at approximately the same place in each report. If one wishes to find the rankings in the appendix, the name of the file changes year after year, ex. in the World Happiness Report 2015, the data file is called “*Statistical appendix*”. In the World Happiness Report 2017, the data file is called “*Final data for Figures in Chapter 2*”. The difference between the two, is that the appendix file is more detailed or sometimes even an excel file that is easy to work with. In the reports themselves, there is just a picture with detailed descriptions.

I have chosen to use the data from the reports themselves, as I found it the best option in terms of what I needed the data for. In this report, I will focus more on the data from the suicide/homicide rates, than the variables used to determine each country’s placement on the World Happiness Report. Furthermore, data pictured in the reports (see appendix) is consistent throughout all the report, whereas the data from the appendix change label names ex. Most reports are ranked by a a so called “ladder score”, in 2017 they change the name to “happiness score”.

In the report one can read more detailed explanations on what variables the “ladder scores” are based on, and what subjects have contributed to happiness or unhappiness for the different years. The ladder score is based on the data from the previous three years, so that the report from ex. 2015 is based on data from the years 2012-2014. Furthermore, the rankings are explained by, among other subjects; GDP per capita, social support, healthy life expectancy, freedom to make life choices, generosity, perceptions of corruption, etc.

The World Happiness Report was first made in 2012, and the data from Our World in Data ends in 2017, therefore, it was an obvious choice to only use the data from 2012-2017. However, the World Happiness Report rankings were updated to include more data from 2012 to 2013, and I have therefore chosen to use data starting from 2013.

As mentioned before, one annual ranking is based off of the three previous years, however I have chosen to look at the data from Our World in Data that matches the year the specific World Happiness Report ranking is published, ex. The happiness ranking from 2015 will match the suicide/homicide data from 2015, even though the happiness ranking is based of the years 2012-2014)

I wanted to make graphs of these data, to make it easier to compare the data, to examine if there could be found a link between happiness and the suicide/homicide rates.

For the making of the graphs, and to get a better overview, I wanted to see the data in RStudio and find a way that made the data more manageable. I did this, by starting with loading the CVS file with the suicide/homicide data into RStudio. From here, I could open the data in a table that was easier to read. In this table, there is a search-bar, to make it easy to sort in the 56,685 entries.

With this search bar, I could find and sort out the data I needed. I then noted all the row numbers matching the data for the countries and years I need.

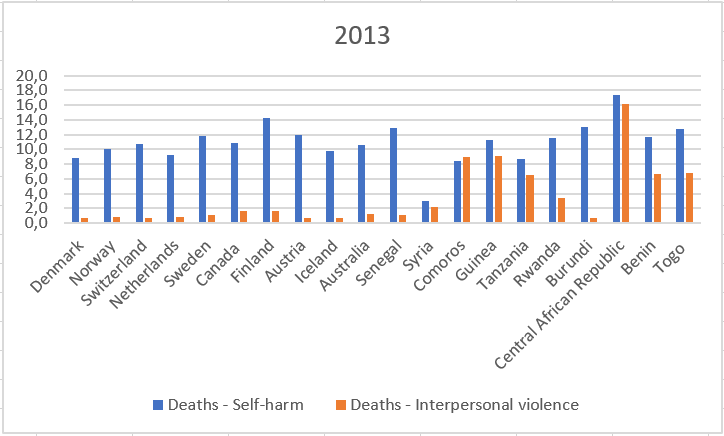
### 5 Empirical Results (Implementation for technical projects)

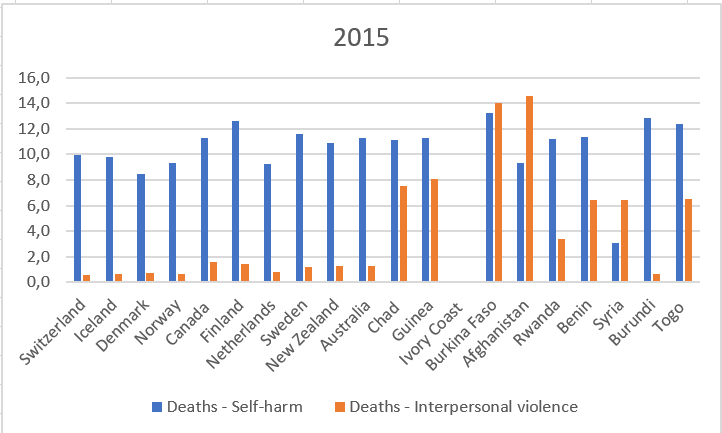
When doing the research as described in the previous section, I stumbled upon many problems.

Firstly, there is not published a World Happiness Report for the year 2014 and I therefor had to overlook all data regarding that year. This made my sample size smaller and thereby my result more prone to flaws. Secondly, there is NA data in the suicide/homicide data, for the Ivory Coast. This was not that big of a problem, as the Ivory Coast was only ranked in top/bottom 10 one year.

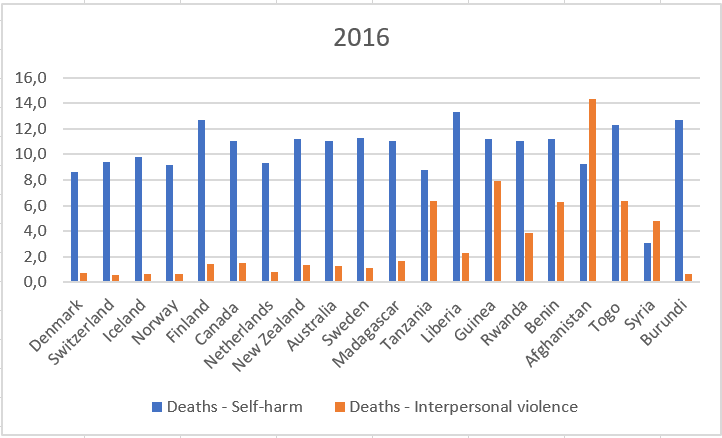
Thirdly, when trying to work with code on RStudio, the program would not recognize the columns and therefore, my plan for graphs did not work. After several hours trying to make it work, using previous assignments and Data Carpentry, with different approaches, still nothing worked.

To get some results, I therefore chose to do my graphs in Microsoft Excel. To do this, I had to transfer all the data from RStudio to Excel, where I could finally get graphs to compare.

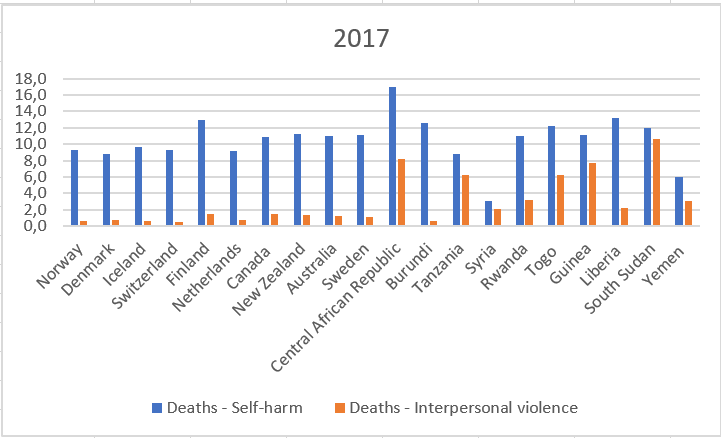
The results I got for 2013 was that the bottom 10 countries had higher homicides rates, though not in an order matching the ranking, with an average rate of 6,2, where as the top 10 countries had an average rate of 1. However, in the suicide rates, there is not that much difference. The top 10 countries have an average rate of 10,8, and the bottom 10 had an average of 11 – only a difference of 0,2.

For 2015 the results had a similar pattern. The homicide rates were highest in the bottom 10 countries, with an average of 6,8 compared to the top 10 countries average of 1. This year, there has been a higher suicide rate in the top 10 countries than in the bottom 10, which is opposite of 2013.

Now, the average rate in the top is 10,5 and 9,6 in the bottom. This is also, a bigger difference than in 2013, as the difference is now 0,9.

In 2016, the homicide rates fell on in the bottom 10 countries, and are an average of

5,5, whereas it has been over 6 the previous years. In the top 10 countries the average stays the same, 1. The suicide rates for this year are now the same, with an average of 10,4.

In 2017 the average homicide rates for the top 10 countries stayed the same, 1, as the bottom 10 fell to an average rate of 5. This continues the pattern of falling homicide rates for the bottom 10 over the years.

The average rate of suicide almost stayed the same for the top 10 countries, this year 10,3, but rose in the bottom 10 to 10,7.

The progress we can see in the years 2013-2017 for the top 10 countries is, that the average suicide rate is declining, while the average homicide rate is constant. For the bottom 10 countries, the average suicide rate fell from 2013 to 2015, but is since then rising, and the average homicide rate rose from 2013 to 2015, but is now falling at a steady pace.

The changes we see in the average homicide rate of the bottom 10 countries, could be caused by the fact that there is a larger shift in the countries on the bottom 10 rankings than in the top 10. The top 10 countries are primarily the same, and the shift in rank is more stable. In the bottom 10 however, there is a bigger shift in which countries are ranked, and their place in the ranking.

Still the question remains, is there a link between a country’s rank in the World Happiness Report and the rates of suicide and homicide?

From this project, it is hard to tell. From the small sample size, we see that there is a big difference in homicide rates, depending on which end of the rank a country is, but that the difference in suicide rates is small. This could mean that there is a connection between homicide rates and the happiness of a county. However, it indicates that there is no connection between the suicide rates and the happiness of a country.

### 6 Critical evaluation

The result did only partly match my expectations, as I expected there to be an oblivious difference between the suicide and homicide rates, depending on the placement on the happiness rank. This was not the case, as the suicide rates were very similar.

I would have liked a bigger sample size, to get a more accurate result, as I find it hard to give a good conclusion to my project. Furthermore, I would have liked the ability to get RStudio to work in the way that I wanted, to spare me some time, and to get better graphs and calculations.

However, considering the many problems I had during this project, I am still somewhat content about the results. This is a problem I would be interested to research further, so I believe this was a good start.

### 7 Conclusions

From this project, I can conclude that my sample size and use of digital tools, were to small to determine whether or not a county’s happiness is connected to tendencies in suicide and/or homicide. I can conclude, that from my small sample size there is a big difference in homicide rates, depending on which end of the happiness rank a country is, but that the difference in suicide rates is small. Which could mean that there is a connection between homicide rates and the happiness of a county, but no connection between the suicide rates and the happiness of a country.

### References

Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2013. World Happiness Report 2013. New York: UN Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2013/>

Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2015. World Happiness Report 2015. New York: Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2015/>

Helliwell, J., Layard, R., & Sachs, J. (2016). World Happiness Report 2016, Update (Vol. I). New York: Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2016/>

Helliwell, J., Layard, R., & Sachs, J. (2017). World Happiness Report 2017, New York: Sustainable Development Solutions Network: <https://worldhappiness.report/ed/2017/>

Roser, Max and Hannah Ritchie (2013) - "Homicides". *Published online at OurWorldInData.org.* Retrieved from: <https://ourworldindata.org/homicides>

Guide to RStudio, Data Carpentry: <https://datacarpentry.org/r-socialsci/>

## B- Required Metadata

*Table 1 – Software metadata*

|  |  |  |
| --- | --- | --- |
| **Nr** | **Software metadata description** |  |
| S1 | Current software version | *R (4.0.3), RStudio (3.167), Microsoft Excel from Microsoft Office Home and Student 2016.* |
| S2 | Permanent link to Github repository where you put your script or R project | [*https://github.com/KathrineKvist/Final\_Project*](https://github.com/KathrineKvist/Final_Project) |
| S3 | Legal Software License | List one of the approved licenses, e.g. Creative Commons 4.0; see Week 6 lecture recordings for more |
| S4 | Computing platform / Operating System | *HP Spectre x360 Convertible 13-ac0XX, 8 Gb RAM, which runs Windows 10 operating system*. |
| S5 | Support email for questions | *201807066@post.au.dk* |

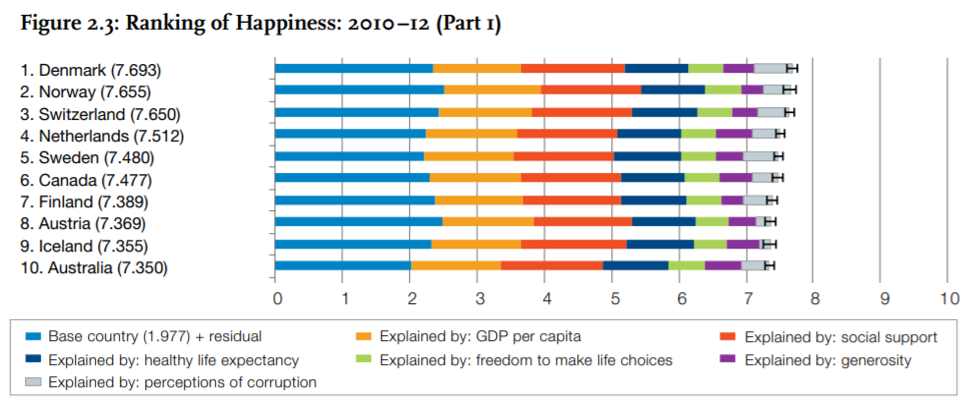
*Table 2 – Data metadata (use the template below or create your own metadata table)*

|  |  |  |
| --- | --- | --- |
| Nr | Metadata description |  |
| D1 | Appendix | Data from World Happiness Report from the years that I am using. The ladder score/happiness score is included in the figures, to make it easier to compare. |
| D2 | [Final Project.xlsx](https://github.com/KathrineKvist/Final_Project/blob/main/Final%20Project.xlsx) | Excel file where I made the graphs for comparisons after RStudio did not cooperate. The only data in this document are the data I have selected to use. |
| D3 | [Final project.Rproj](https://github.com/KathrineKvist/Final_Project/blob/main/Final%20project.Rproj) | My project in RStudio. Here one can read the data I have used and see how I have used them. One can also see the different ways I tried to make a graph with no success. |
| D4 | [suicide-vs-homicide-rate.csv](https://github.com/KathrineKvist/Final_Project/blob/main/suicide-vs-homicide-rate.csv) | The data I downloaded from <https://ourworldindata.org/homicides#the-global-distribution-of-homicides>. The data is sorted by 7 collumns; *Entity, Code, Year, Deaths - Self-harm - Sex: Both - Age: Age-standardized (Rate), Deaths - Interpersonal violence - Sex: Both - Age: Age-standardized (Rate), Population (historical estimates), Continent* |

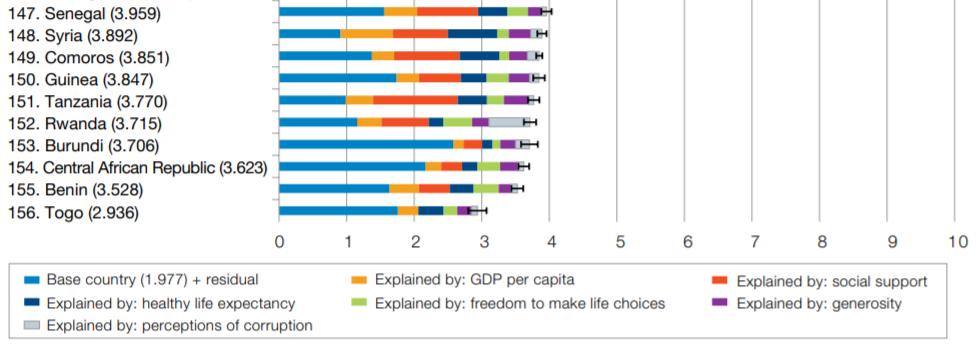
# Appendix

### 2013 (2010-2012):

**Top 10[[3]](#footnote-3):**

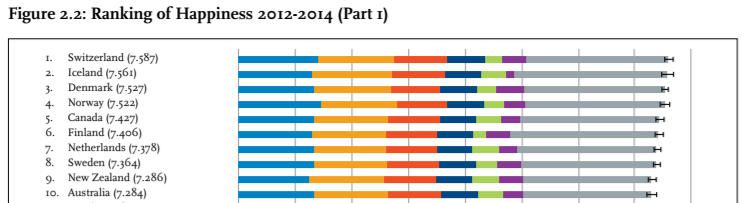


**Bottom 10[[4]](#footnote-4):**

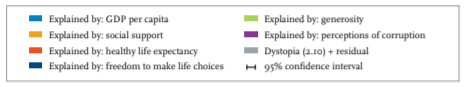


### 2015:

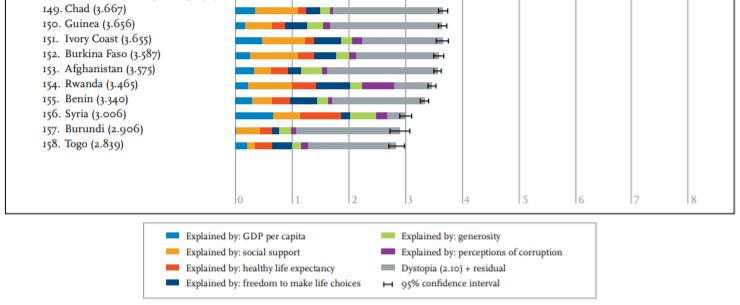
**Top 10[[5]](#footnote-5):**





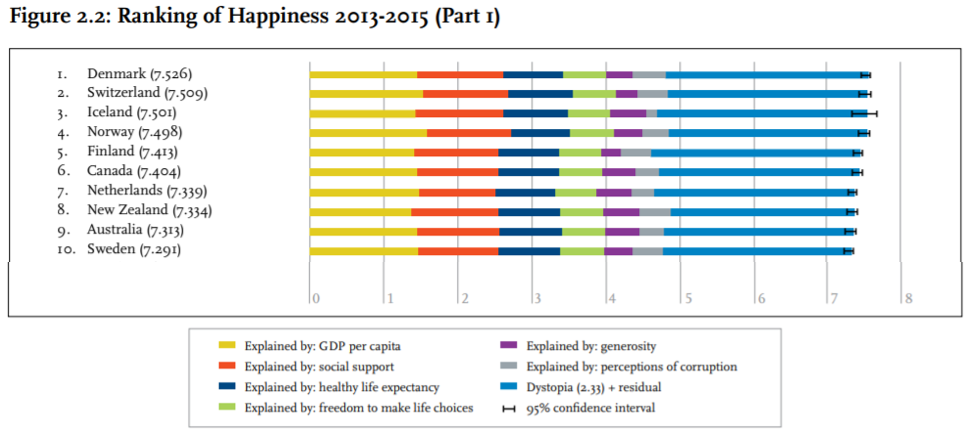


**Bottom 10[[6]](#footnote-6):**

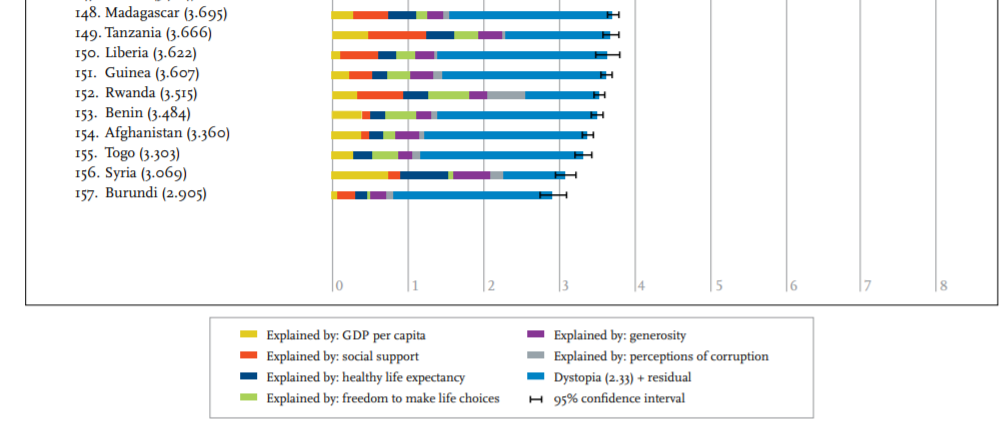


### 2016:

**Top 10[[7]](#footnote-7):**

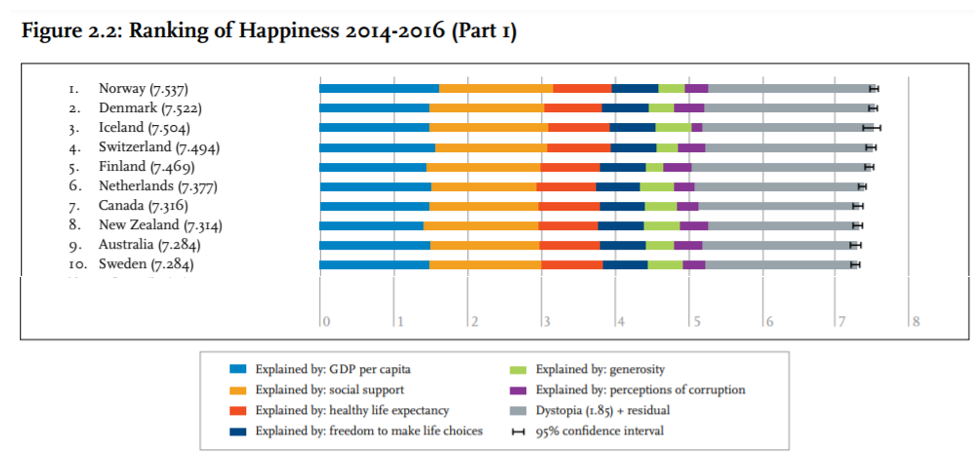


**Bottom 10[[8]](#footnote-8):**

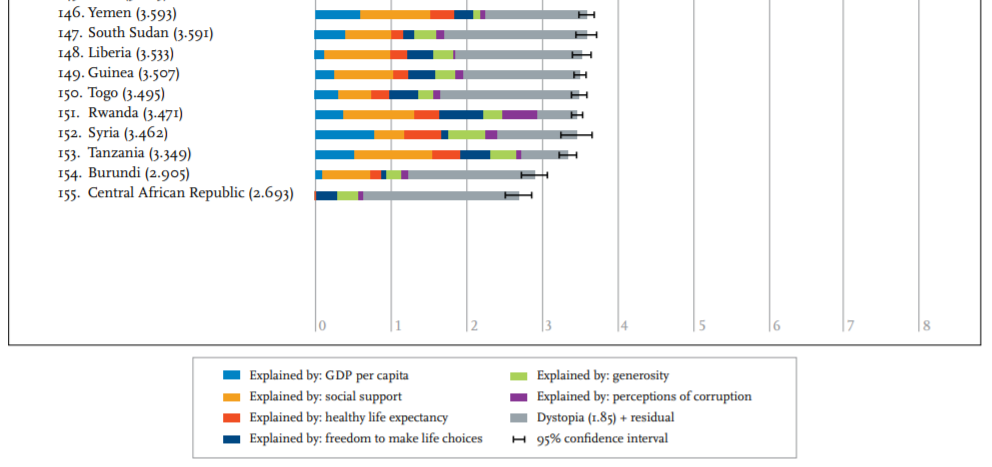


### 2017:

**Top 10[[9]](#footnote-9):**



**Bottom 10[[10]](#footnote-10):**

****

1. https://worldhappiness.report/ [↑](#footnote-ref-1)
2. https://ourworldindata.org/homicides#the-global-distribution-of-homicides [↑](#footnote-ref-2)
3. Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2013. World Happiness Report 2013. New York: UN Sustainable Development Solutions Network, p. 23 [↑](#footnote-ref-3)
4. Helliwell, p. 25 [↑](#footnote-ref-4)
5. Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2015. World Happiness Report 2015. New York: Sustainable Development Solutions Network, p. 16 [↑](#footnote-ref-5)
6. Helliwell, p. 18 [↑](#footnote-ref-6)
7. Helliwell, J., Layard, R., & Sachs, J. (2016). World Happiness Report 2016, Update (Vol. I). New York: Sustainable Development Solutions Network, p. 22 [↑](#footnote-ref-7)
8. Helliwell, p. 24 [↑](#footnote-ref-8)
9. Helliwell, J., Layard, R., & Sachs, J. (2017). World Happiness Report 2017, New York: Sustainable Development Solutions Network, p. 22 [↑](#footnote-ref-9)
10. Helliwell, p. 24 [↑](#footnote-ref-10)