Team Anaconda

02/20/2021

Project 1 Summary and Analyses

Contents

[Happiness Data Report 2](#_Toc64624221)

[What factors play a role in the happiness score of a country? 2](#_Toc64624222)

[1. Unemployment Analysis 2](#_Toc64624223)

[2. Government health spend vs. Happiness 4](#_Toc64624224)

[3. Social Support vs Happiness 5](#_Toc64624225)

[4. Government Analysis 8](#_Toc64624226)

[5. Weather Analysis 8](#_Toc64624227)

[What is the development of a country or region’s happiness over time? 11](#_Toc64624228)

[What differences can be found between regions of the world? 12](#_Toc64624229)

[References 13](#_Toc64624230)

# Happiness Data Report

The first iteration of the World Happiness Report was published in 2012 in support of the United Nations High Level Meeting on Happiness and Well-Being. It was created in response to the 2011 Resolution of the UN General Assembly, which invited countries to measure the happiness of their people and use this information when making public policy decisions. The main purpose of this report was to bring internationally comparable data, based on the Gallup World Poll, and a common scientific understanding to the global stage. In doing so, the report was successful in establishing a shared understanding of subjective well-being and methods to measure it. Subsequent reports built on this foundation and made it so that the World Happiness Reports are a key guide in the progress towards sustainable development.

The happiness scores within these reports are based on responses to the Gallup World Poll that has been surveying countries since 2005 and now represents almost all of the world’s population. The questions on this poll adhere to the guidelines outlined in the Organisation for Economic Co-operation and Development (OECD) *Guidelines on Measuring Subjective Well-being*. The recommendation is for there to be a central measure of life satisfaction on a Cantril ladder. This question would essentially ask respondents to evaluate their current lives on a 0 to 10 scale, with 0 representing the worst possible life and 10 the best life for them.[[1]](#footnote-1) Additionally, the guidelines recommend a series of questions about life meaning or purpose that would complement the primary measure. The responses to these secondary questions could differ by region due to cultural and other differences across the globe. The Gallup World Poll fulfills this requirement by including a series of experiential questions primarily focused on the previous day.

The World Happiness Report also estimates how six key variables can explain the happiness scores for a given country. The six factors are GDP per capita, healthy years of life expectancy, social support (as measured by having someone to count on in times of trouble), trust (as measured by a perceived absence of corruption in government and business), perceived freedom to make life decisions, and generosity (as measured by recent donations, adjusted for differences in income).[[2]](#footnote-2) Each of the six variables tries to highlight a different aspect of life.

# What factors play a role in the happiness score of a country?

## Unemployment Analysis

This analysis served to study the correlation between the overall unemployment rate in each country and its happiness score. The analysis was based primarily on two different datasets – one measuring happiness and the other unemployment. Both datasets contained data over the same time frame, 2015 to 2019. The hypothesis being that there is an inverse correlation between unemployment rate and the happiness score.

The datasets came from two different sources and did not contain the same countries. Specifically, the happiness dataset from Kaggle contained information for 141 countries. The unemployment dataset, on the other hand, came from the World Economic Outlook (WEO) database (International Monetary Fund) and included over 200 records. After cleaning and merging these datasets, the final dataset only included 88 records. The merge was conducted using the names of the countries and it is very likely that differences in the ways in which they were spelled in each dataset affected the result.

Nevertheless, analysis showed a moderate correlation between unemployment rates and happiness scores (Fig. 1). The r-value of -0.38 demonstrated that there is a negative moderate correlation, in other words, a higher unemployment rate is likely to result in a lower happiness score. The overall correlation calculation is affected by the sample size. Of the 88 records, about six were potential outliers and were excluded from the above calculation. When included, these outliers bump the overall r-value to -0.40.

Another factor that could potentially impact these values is the access of information. Was it easier for happier countries to present information? Were countries with higher employment rates more likely to share information? There is no way to definitively answer this question, however, it is likely that these factors are important. For example, there was no available data on the countries with the lowest happiness scores. Most of these countries had suffered from political turmoil during the timeframe covered in the dataset and it is likely that information was not easily available. Chart, scatter chart

Description automatically generated

Figure 1

## Government health spend vs. Happiness

This aspect of our project serves to evaluate countries expenditures for Healthcare and their relationship to Happiness scores. To begin this part of the project, data was gathered about I located data on how much each country spends on healthcare, as well as data about populations.

I then divided a countries total spending by population to determine an expenditure per person. I converted this into a csv file. I combined this information with the World Happiness Report Data Index csv file into 1 dataframe. I then cleaned the merged data and renamed any columns to make the dataframe fields more readable. I ran a regression analysis to determine correlation between healthcare spending per person and happiness score. (Fig. 2)

I found a strong correlation between the two. It made me curious if only highly populated areas spent the most on healthcare so I did an analysis just between population and happiness. (Fig. 3)

What I found was there was no correlation between happiness and population, so my first finding, spend on health per person, told the real picture.

Chart, scatter chart

Description automatically generated

Figure 2

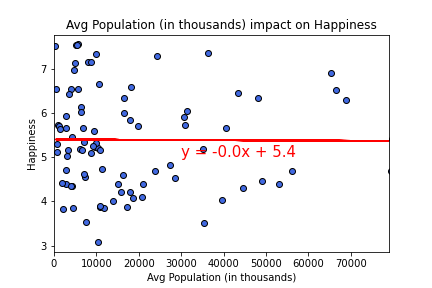


Figure 3

## Social Support vs Happiness

This analysis was to evaluate the impact of the strength of social support on the average happiness score of a country. The hypothesis being that there is a correlation between social support score and the happiness index. To begin this project, I combined the World Happiness Report Data Index csv files, and cleaned and renamed the combined file. I ran a regression analysis to determine correlation between average scores for social support and happiness. I found a strong positive relationship between the two with a r-square 0.78. (Fig. 4)

I analyzed the data from 2015 to 2019 and noticed that there was a big depression happened in terms of social support in 2016. It was the year the Syrian refugee crisis happened, and the spread of Zika virus and the Zika fever epidemic caused worldwide alarm. I guess these significant events had an impact on people and a fear did spread all over the world; as a result, people had lost their confidence level about they do have someone to count on if they were in trouble. But after that period there was a huge improvement in their confidence level about the social support.

Development of a country’s happiness index and social support score over time was evaluated among the top 10 countries and among the bottom 10 countries. Iceland is always on the top position in terms of social support scores. In the happiness score ranking it was in the second position in 2015, but there is a gradual decrease in happiness index and in 2019 it was placed down in the fourth position. In case of Finland, the happiness score is gradually increasing, and it became the top one country in 2019; in terms of social support score it was in the sixth position in 2015 but it improved the score and became the second highest country in the list of social support score. (Fig. 5)

Figure 4

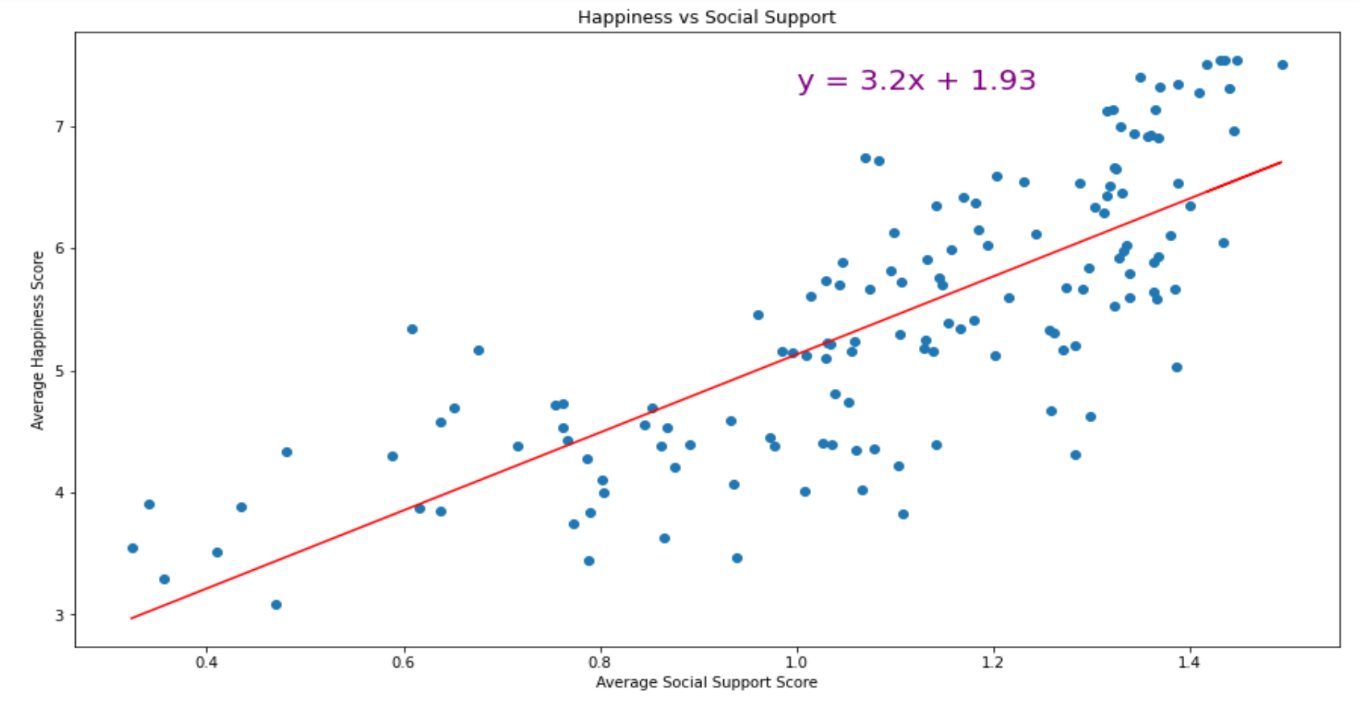
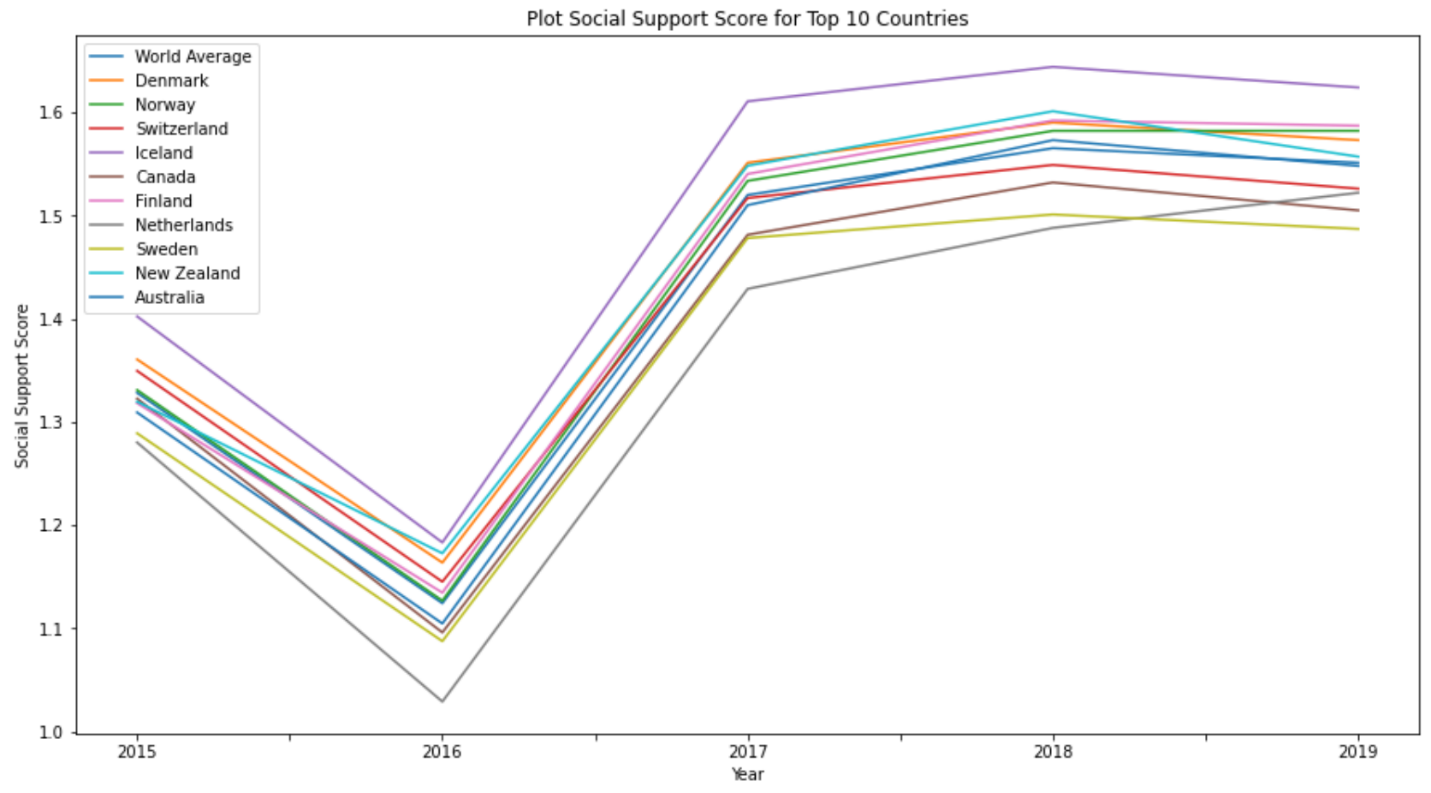


Figure 5

The bottom 10 countries based on the happiness score was analyzed. As shown in the graph, Syria has a lower score in social support, especially after 2016, but its happiness index is not the lowest. This is easily explained by the conflict the country has been. (Fig. 6)

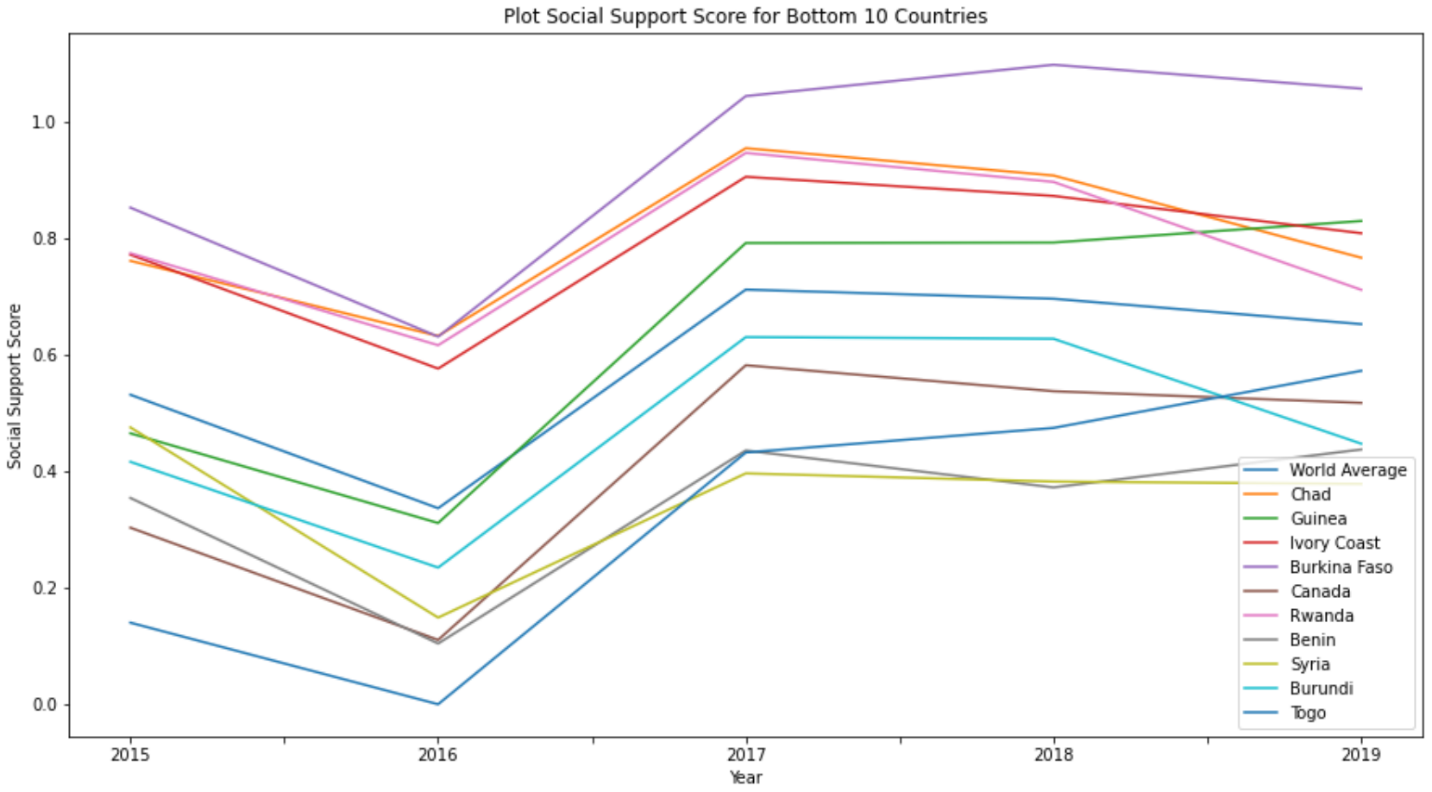
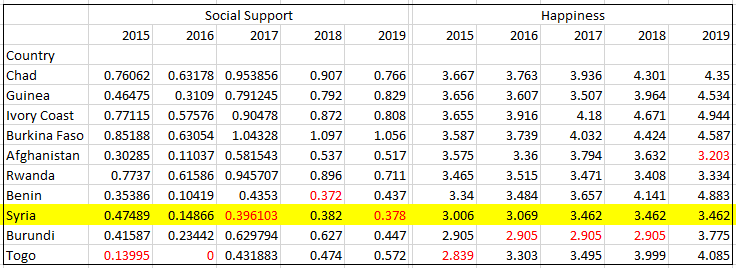


Figure 6



Based on the overall evaluation, I want to say that lack of social support seems to affect the happiness index up to a certain level. This analysis suggests that there is a strong correlation between social support and average happiness score of a country, but it is not as I suspected as a linear relationship. Further analysis needed to find what kind of effect social support score has on happiness index, and to see family size and social support has any correlation.

Because of my curiosity, I added generosity score into my analysis to see any effect it has both generosity and social support together on happiness. The analysis shows that generous countries were not automatically happy until social support was added into the picture. Surprisingly, Myanmar has the highest level of generosity score, but its happiness score is below 4.5. It means, people want to make others happy, even if they are not happy themselves.

## Government Analysis

To find the types of governments on the top and bottom ten countries, I used the Database on Political Institutions (DPI) from the World Bank Data Catalogue to extract the types of political systems with in each country during 2015 to 2017. Merged this information with the average happiness per country taken from the "world Happiness Report"

I was interested in looking at what types of governments were the top ten and bottom ten average happiness countries to further understand if this was a contributing factor to the happiness ranking/ score.

I used the bar plot to have a visual of the 3 types of government used in the in the DPI form the world bank database. Described below:

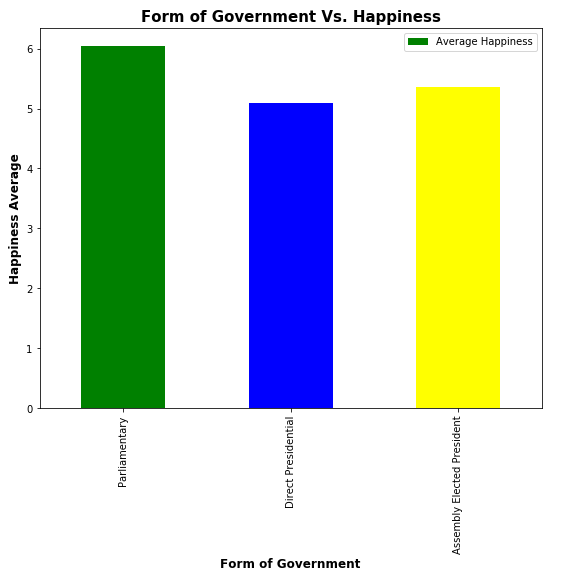


Figure 7

* “0”: Direct Presidential are Countries with presidents who are elected directly or by

an electoral college

* “1”: Assembly Elected President are Countries where the selection of the effective

executive is done by an elected assembly or by an elected but uncommitted

electoral college

* “2”: Parliamentary are Countries in which the legislature elects the chief executive

Findings:

There is a positive relation between the quality of government and the average happiness.

When sorting through the data on the top ten countries ranked by highest average happiness, these countries are all under Parliamentary form of Government. While on the bottom 10 countries 8 out of 10 are Direct Presidential.

## Weather Analysis

To determine the weather impact of a country on their happiness score the happiness data served as the first input. In the jupyter notebook “weather.ipvnb” countries were combined with a list of the nations’ capitals and for each capital, a current weather report was pulled. The question to be answered with this aspect of our project is not what weather influences happiness, but what weather can be expected were we to visit one of the ten on average happiest countries from the World Happiness Report.

The capital locations with their respective countries were processed in preparation to be merged. The city names were then sent out with a request for current weather conditions from an API. The returned data was converted into a new dataframe for further evaluation and exported as a .csv file. Further cleaning of data was completed to prepare for a merge with the happiness data report. Cities lacking weather data (“NaN” in city field) were removed, as were any empty data rows, of which there were none. City weather data was merged back into countries and capitals list, to allow further processing by country instead of by city. Here also, any row that had no value or had an “Nan” fields were removed from the dataset. This dataframe was then again merged with the Average Happiness data file from section 1, which provides average happiness over the four years 2015 – 2019 in one score. This score was used to slice the weather data into a dataframe showing weather for the top 10 happiest countries worldwide. The resulting dataset “happiness\_weather\_top10” was exported as a .csv file for the following visualizations.

Visualizations were prepared to depict any corresponding trend for current weather metrics. It is, of course, winter in most of our regions to temperatures were very low, except for in Oceania, where it is currently Fall. (Fig. 8)

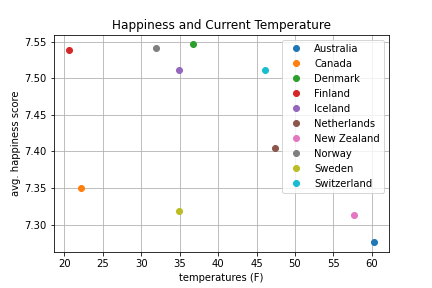


Figure 8

# What is the development of a country or region’s happiness over time?

To answer this question, we decided to focus on the changes in happiness scores for ten countries with the highest average happiness scores and ten countries with the lowest average happiness scores. The goal was to display the apparent differences in trends and overall scores. What we found was that the happiness scores for the top ten countries ranged was always above 7 on a 10 point scale. (Fig. 9)

Chart, line chart

Description automatically generatedOn the other hand, the ten countries with the lowest scores did not have any scores above 4.5 in the same time frame. (Fig. 10)

Figure 9

Chart, line chart

Description automatically generated

Figure 10

# What differences can be found between regions of the world?

In order to answer this question, we needed two pieces of information, the average happiness for each country in the dataset and the region that each country was located in. After cleaning the original happiness datasets from Kaggle we were able to get both these pieces of information. Once the information was available, the data was grouped by region of the world and retrieved the average happiness score for each region.

What we found was that there are clear differences in the average happiness scores by regions of the world. The regions were defined the World Happiness Report 2015.[[3]](#footnote-3) There is regions with the highest score is Australia and New Zealand with an average score of 7.29, closely followed by North America with a score of 7.17. On the other end of the spectrum, we have Sub-Saharan Africa with an average score of 4.19, closely followed by Southern Asia with a score of 4.58. It is also interesting to point out difference in regions that are in close proximity. For example, Western Europe has an average score of 6.79, while Central and Eastern Europe has a score of 5.44. That is a 24.8 % change from the latter to the former. (Fig. 11)

Chart, bar chart

Description automatically generated

Figure 11

# References / Sources

1. Helliwell, John F., Richard Layard, and Jeffrey Sachs, eds. 2015. World Happiness Report 2015. New York: Sustainable Development Solutions Network.
2. “Understanding How Gallup Uses the Cantril Scale.” *Gallup*, <https://news.gallup.com/poll/122453/understanding-gallup-uses-cantril-scale.aspx>.
3. World Happiness Report (base for happiness)  
    <https://www.kaggle.com/unsdsn/world-happiness>World Health Organization Global
4. Expenditure Database (for health care spend comparisons)

<https://apps.who.int/nha/database/Select/Indicators/en>

1. International Monetary Fund (for unemployment data)  
   <https://www.imf.org/en/Publications/SPROLLS/world-economic-outlook-databases#sort=%40imfdate%20descending>
2. World Bank Data Catalogue (government types)  
   <https://documents.worldbank.org/>
3. Weather API (current weather)  
   <https://openweathermap.org/api>

1. (Understanding How Gallup Uses the Cantril Scale, n.d.) [↑](#footnote-ref-1)
2. (Helliwell, Layard, & Sachs, 2015) [↑](#footnote-ref-2)
3. (Helliwell, Layard, & Sachs, 2015) [↑](#footnote-ref-3)