# Why the Long Face? A Statistical Comparison of Additional Variables to the United Nations World Happiness Report

### *by Marieth Coetzer, Reedley College*

### *Mentor: Dr. John Heathcote*

## Abstract

In 2008, the government of Bhutan announced that, in addition to the traditional GDP measurement, it would adopt a happiness index with which to measure its citizens' wellbeing. This decision inspired the United Nations to release the first World Happiness Report in 2012 with the objective of reexamining the origins of well-being. They identified key factors such as social support and generosity, using survey results gathered by the Gallup World Poll. In a time of escalating debates about policy-making and its relationship to the human pursuit of happiness, a quantitative investigation into happiness is helpful for making data-based decisions. The goal of this study was to correlate additional variables, such as national service, family size, and divorce rates with social support in the World Happiness Report's top twenty countries. The resulting enhanced social insight can serve to inform governments, businesses, and educational institutions to make improved decisions during this crucial time.

This study compared social support to additional variables retrieved from official government statistics. Correlation was established using Pearson's r for interval variables and point biserial correlation for dichotomous variables.

## Why the Long Face?

Why the Long Face started as an exploration of using the World Happiness Reports published by the United Nations to try and understand what additional variables affect the Key Factors that the Reports identified. The Key Factors have been carefully defined by the World Happiness Report authors as important aspects that contribute to people's happiness. These were identified to be gross domestic product, social support, healthy life expectancy, freedom to make life choices, generosity, and perceptions of corruption. This study focused on social support as it is one of the most important factors, but also one of the hardest to measure.

Social support for each country was determined using a simple yes/no survey. Around one thousand people every year were asked the question, "if you were in trouble, do you have relatives or friends you can count on to help you whenever you need them, or not?" If a person answered "yes," a score of one was assigned to their answer. If they answered "no," their score was assigned a zero. The average of all of these answers were taken to determine the social score of a country.

The goal of this research was to find a mathematical correlation between the social scores and other factors such as national service, family size, and divorce rates. These mathematical correlations will lead to deeper understanding about what affects social support in society and how certain systems can serve to include or isolate individuals. This research is particularly important in the time of COVID-19, where social support is becoming more important than ever.

**Methodology**

The methodology was to visit the World Happiness Report website and download the 2019 Online Data Spreadsheet. The social support scores for the top twenty countries were taken and most of the tests were run using data from 2017. For certain variables where data was missing, the tests were run from the years 2015-2018. When certain social support scores were missing (such as Iceland omitting social support data from 2018), the tests were run regardless. The missing information is categorized below in the endnotes. For tests run over several years, the average of Pearson's r was taken.

Pearson's r is a statistical calculation of the correlation between two variables. It is measured on a scale from negative one to positive one. The closer Pearson's r falls to either one, the stronger the correlation. If Pearson's r falls close to zero, there is no relationship. Examples of this can be seen below in Figure 1. This project's objective was to find a value below -0.5 or above 0.5, which would be considered a "moderate" correlation. Pearson's r was calculated using Excel.
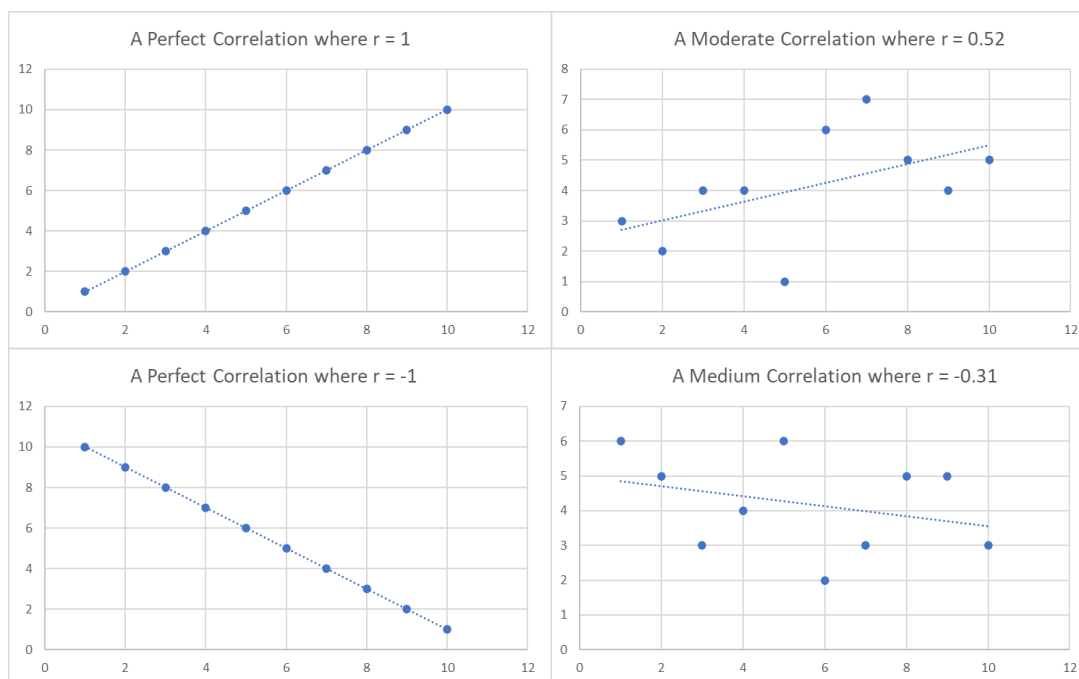


Figure 1: Pearson's r Examples

The data for additional variables were found using the following resources:

- Association of Religion Data Archives (ARDA): https://www.thearda.com/
- PrisonStudies.org: https://www.prisonstudies.org/
- Organization for Economic Co-operation and Development (OECD): https://data.oecd.org/
- World Data Bank: https://data.worldbank.org/
- World Factbook from the CIA: https://www.cia.gov/library/publications/resources/the-world-factbook/
- World Health Organization (WHO): https://www.who.int/hiv/data/en/
- United Nations: http://data.un.org/

**Findings**

The findings were categorized into four categories: "weak," "medium," "intermediate," and "moderate." There was also an additional category called "complicated" for tests that yielded multiple results.

The "weak" results are summarized below in Table 1. Although Pearson's r is shown to be both positive and negative in these results, it is important to keep in mind that the values are so close to zero that any positive or negative correlations are negligible.

| "Weak" Results: Between -0.3 and 0.3 | | | |
|---|---|---|---|
| **Additional variable** | **Data Retrieved from** | **Year or Years** | **Pearson's r** |
| Age of Population | World Data Bank | 2017 | -0.017865976[1] |
| Alcohol Consumption | WHO | 2014-2018 | 0.122212642 |
| Dependents in Population | World Data Bank | 2017 | 0.180019744 |
| Divorce Rate | OECD | 2017 | -0.184475947[2] |
| Domestic Care: Men | World Data Bank | 2008-2018 | 0.229971561[3] |
| Domestic Care: Women | World Data Bank | 2008-2018 | -0.009668428[3] |
| Family Responsibility | OECD | 2015-2016 | -0.032758353[4] |
| Fertility: Number of Children | OECD | 2017 | -0.0755171164 |
| Homelessness | OECD | 2015-2018 | 0.148163994[5] |
| National Service | CIA World Factbook | 2015-2018 | 0.16890405 |
| Suicide Rate | WHO | 2015-2016 | 0.177893832 |
| Unemployment Rate | OECD | 2018 | -0.122930926 |
| Rural Population | World Data Bank | 2018 | -0.142905784 |
| Urban Population | World Data Bank | 2018 | 0.142905784 |
| Weekly Hours Worked | OECD | 2017 | -0.242064006[6] |

Table 1: "Weak" Results

Although most of the tests run ended up in the "weak" category, they should not be immediately discarded. Instead, they serve to provide a scope as to what constitutes as "weak" and act as a type of process of elimination. Any variables that ended up in this category did not require any additional tests to be run.

Table 2 below summarizes the results that fit in the "medium" category. While Pearson's r isn't significant, there are some observations that can be drawn from these results.

| "Medium" Results: Between (-0.3 and -0.4) or (0.3 and 0.4) | | | |
|---|---|---|---|
| **Additional variable** | **Data Retrieved from** | **Year or Years** | **Pearson's r** |
| Access to the Internet | World Data Bank | 2017 | 0.377763765 |
| Household Size | OECD | 2015 | -0.380755221 |
| Incarceration Rate | PrisonStudies.org | 2015-2018 | -0.390495864 |
| Density of Population | UN | 2017 | -0.345657064 |
| Visiting with Friends | OECD | 2015-2017 | -0.33314989[7] |

Table 2: "Medium" Results

One of the most unexpected results in this category was the positive correlation between social support and the percentage of population with access to the internet. This could be explained by speculating that when individuals are online, they are able to keep in touch with family and friends or even meet new acquaintances online. Another surprising result was that the more time adults spend visiting with friends once a week outside of work, the less the social support appears to be in a country.

The "intermediate" results, although not quite breaking through the 0.5 barrier hoped for, were close enough to significant to warrant special attention and their own separate category. These are summarized below in Table 3.

| "Intermediate" Results: Between (-0.4 and -0.5) or (0.4 and 0.5) | | | |
|---|---|---|---|
| **Additional variable** | **Data Retrieved from** | **Year or Years** | **Pearson's r** |
| Adolescents out of School | World Data Bank | 2017 | -0.423008662[8] |
| Females in the Workforce | World Data Bank | 2015-2018 | 0.432069099 |
| Intentional Homicide | World Data Bank | 2016 | -0.486567085 |

Table 3: "Intermediate" Results

As can be seen from these results, there is a negative correlation between social support and adolescents out of school along with intentional homicide.  On the other hand, there is a positive correlation between social support and percentage of females in the workforce.

While most of the tests run were very straight forward with one clear result, there were a few compilations of data that needed to be categorized as one test but yielded multiple results. These will be expanded on below.

The first test that yielded multiple results was the case of education, which is summarized below in Table 4. The data used was the percentage of every country's population that attained a certain degree in 2018. The social support scores were taken from the years 2017-2018.

| Education | |
|---|---|
| **Additional variable** | **Pearson's r** |
| Bachelor's or equivalent education | 0.442712951 |
| Master's or equivalent education | -0.011070067 |
| Doctoral or equivalent education | 0.059112417 |

Table 4: Education

These results show that the more students gain bachelor degrees in countries, the more social support increases. This could perhaps be because of the sense of camaraderie that is built between students as they work towards getting the same degree.

One of the most unforeseen results was found in some data on the OECD website that broke down the proportion of children by the residence and marital status of their parents in the household. Some of this data was quite old. However, this old data is very valuable to test some of the possible long-term effects of children in living situations. If a 10-year-old child lives with both parents in 2005, that means the possible effect of that living situation might come to light if he or she partakes in the World Happiness Survey ten years later at the age of twenty. As a result, the test was run taking living situation data from 2005-2009 and compared to social support data from 2015-2018. The results are summarized below in Figure 2.
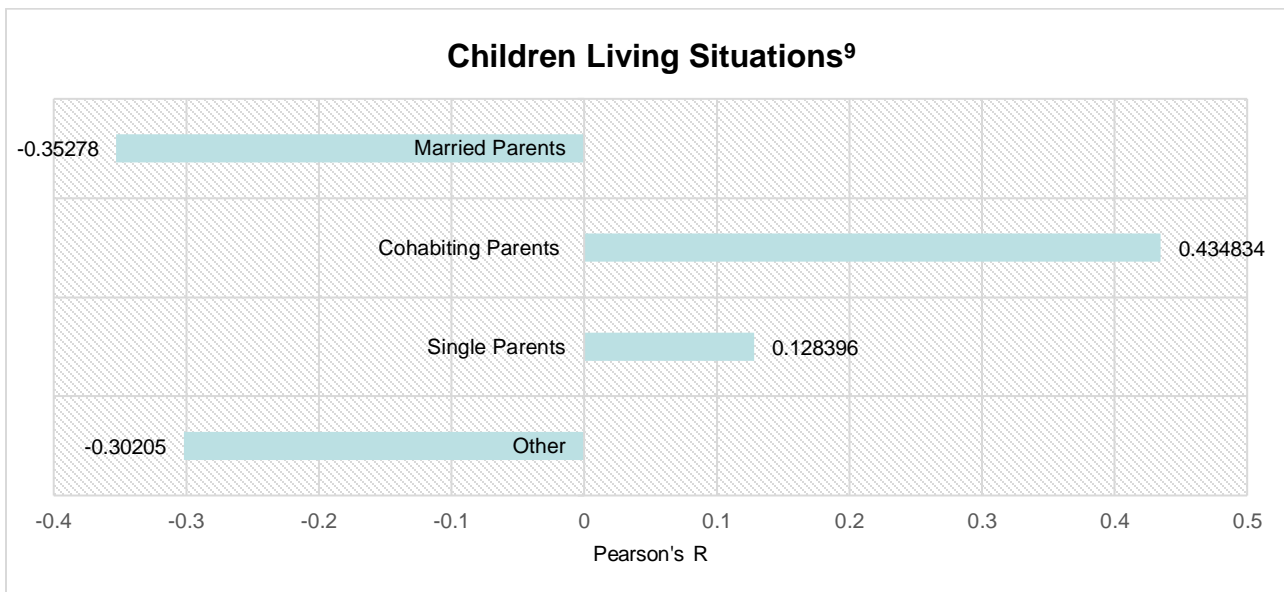


Figure 2: Children Living Situations

As can be seen from these results, the living situation of single parents seems to make very little difference in social support ten years later. However, there is a strange difference between married parents and cohabiting parents. Married parents have a "medium" negative correlation while cohabitating parents have an "intermediate" positive correlation.

Another test with unforeseen results was regarding time use. A collection of survey data found on the OECD website asked people how they spend their time each day. When the test was run

using social support from 2008-2018, there was a positive correlation between unpaid work and social support as can be seen below in Table 5.

| Time Use[10] | |
|---|---|
| **Additional variable** | **Pearson's r** |
| Paid work or Study | -0.250707753 |
| Unpaid Work | 0.400195825 |
| Personal Care | -0.251611238 |
| Leisure | 0.131712307 |
| Other | 0.188853573 |

Table 5: Time Use

What was this unpaid work and why did it show a positive correlation that was close to the "moderate" value of 0.5? OECD's data bank was explored further and the breakdown of unpaid work was found. A second test was run using the same social support years and the results can be seen in Table 6.

| Unpaid Work[10] | |
|---|---|
| **Additional variable** | **Pearson's r** |
| Routine Housework | -0.017503912 |
| Shopping | 0.214520475 |
| Caring for Household Members | 0.395667597 |
| Child Care | 0.247870424 |
| Adult Care | -0.011645845 |
| Care for Non-Household Members | 0.245464 |
| Volunteering | 0.066999512 |
| Travel Related to Household Activities | 0.045518 |
| Other Unpaid | 0.355234 |

Table 6: Unpaid Work

What this implies is that as far as time use is concerned, individuals using their time to care for household members leads to stronger social support.

Although there were plenty of exciting results already found, no result had broken through the "moderate" barrier of 0.5. That changed, however, upon the discovery of the Association of Religion Data Archives. The data was from around the year 2015 and the test was run with social support from the years 2015-2018 due to the unlikelihood that a religious population would change drastically over time.

For the religious population four categories were chosen: total religious, Christian population, Muslim population, and non-religious or unknown. These categories were chosen because there was enough data to run the test for Christianity and Islam beliefs. Many countries didn't report other religions such as Judaism and simply put them under the "other" category, meaning those additional

beliefs were unable to be included in the test. The results are summarized below in Table 7 and graphically displayed on Figure 3.
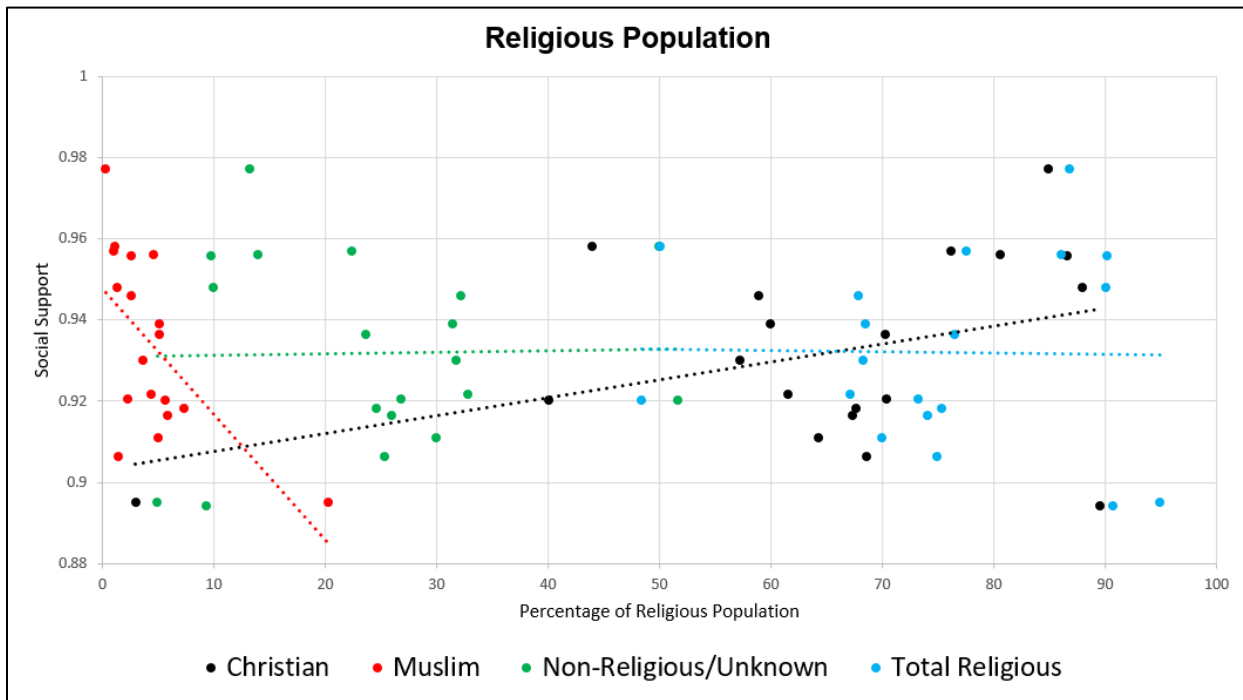


Figure 3: Religious Population

| Religious Population[11] | | | | |
|---|---|---|---|---|
| **Additional variable** | Christian | Muslim | Non-religious or unknown | Total Religious |
| **Pearson's r** | 0.308489413 | -0.514511766 | 0.037227387 | -0.037089601 |

Table 7: Religious Population

The most distinct result was the -0.51 value, showing a negative correlation between social support and percentage of Muslims in a country. Why was this so? One theory was that it was a matter of immigration. As can be seen from Figure 3, the top twenty countries in the World Happiness Report all have a Muslim population of less than 25%. Immigrants may have difficulty building up a social network, which would undoubtedly affect social support.

Along with this possibility, there was also data available to extend the list to the top forty countries in the World Happiness Report. By adding these twenty countries, the test would include regions where people of the Muslim faith make up a majority of the population. The results of this test are summarized below in Table 8 and graphically displayed on Figure 4.
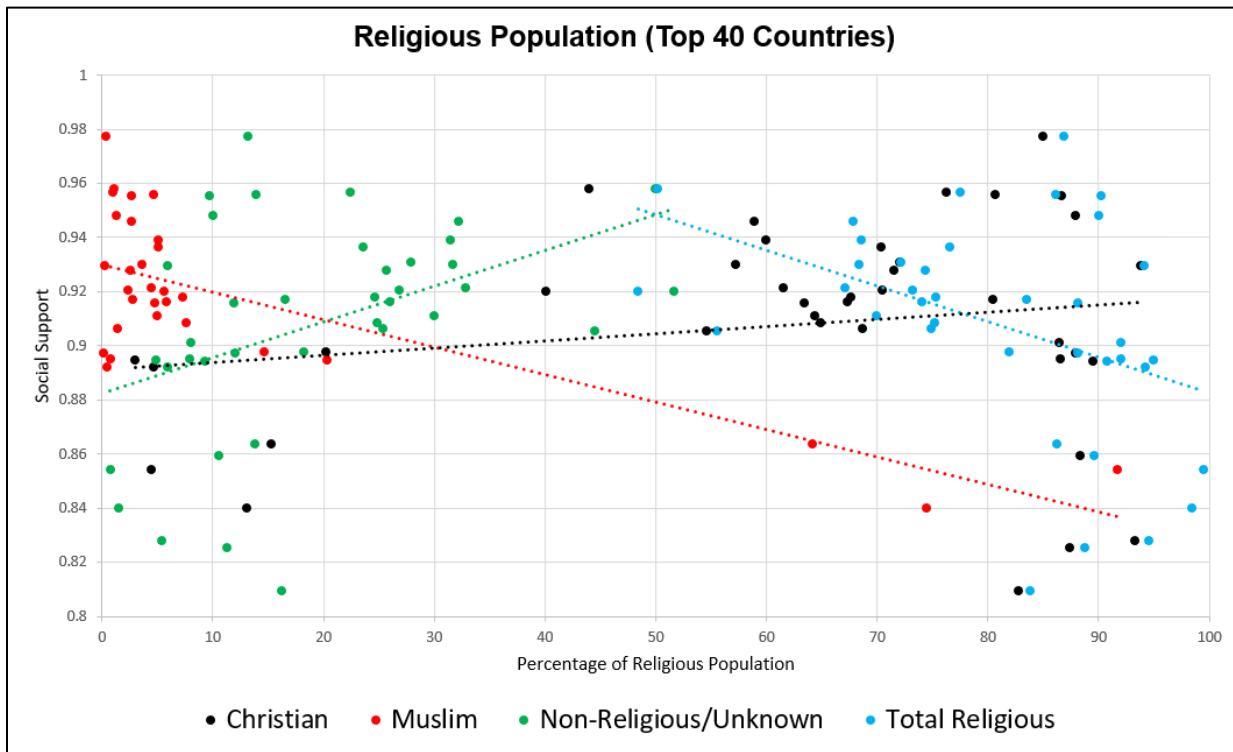
Figure 4: Religious Population (Top 40 Countries)

| Religious Population (Top 40 Countries)[12] | | | | |
|---|---|---|---|---|
| **Additional variable** | Christian | Muslim | Non-religious or unknown | Total Religious |
| **Pearson's r** | 0.14814249 | -0.689329213 | 0.413332086 | -0.413808363 |

Table 8: Religious Population (Top 40 Countries)

The three countries where people of the Muslim faith made up over 60% of the country's population were Bahrain, Saudi Arabia, and the United Arab Emirates. As can be seen from Table 8, the negative correlation became even stronger, making the result even more difficult to explain. However, as can be seen from Figure 4, those three countries ended up being outliers, which can really skew the results in statistics. The outliers were removed and a final test was run on religion. The results can be seen in Table 9 and are graphically displayed on Figure 5.
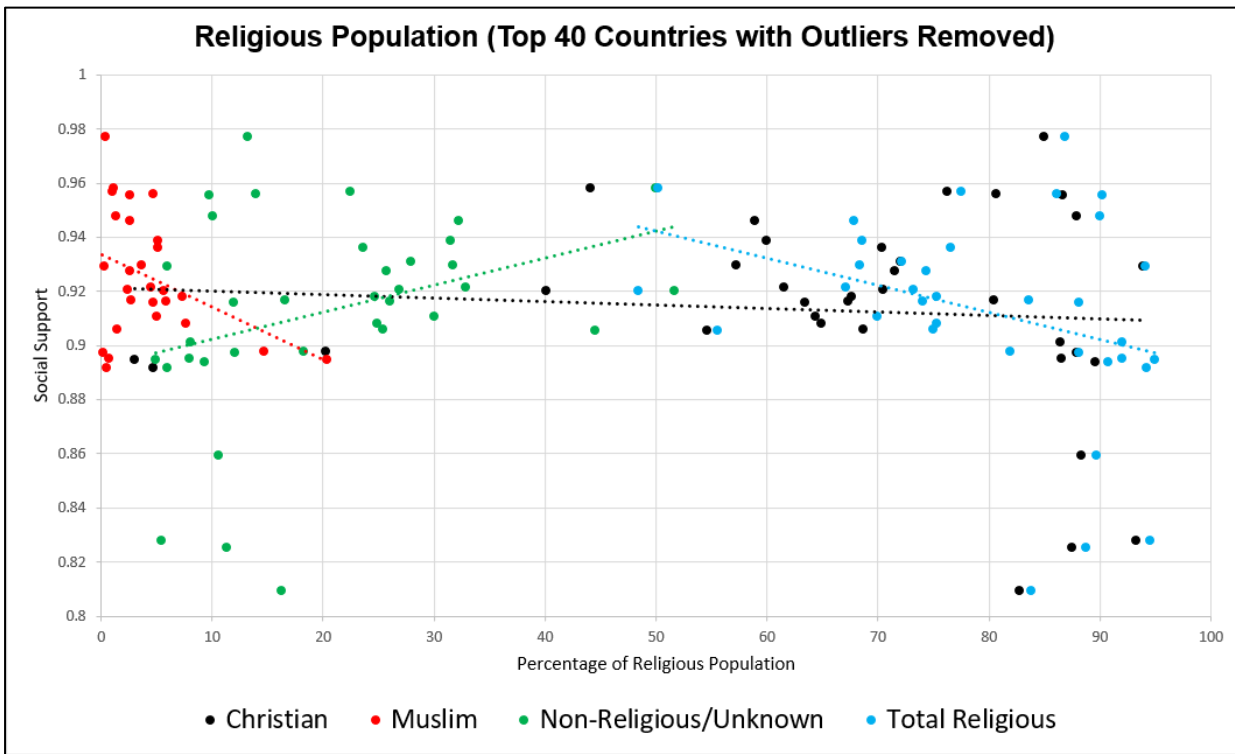
Figure 5: Religious Population (Top 40 Countries with Outliers Removed)

| Religious Population (top 40 countries with Outliers Removed)[12] | | | | |
|---|---|---|---|---|
| **Additional variable** | Christian | Muslim | Non-religious or unknown | Total Religious |
| **Pearson's r** | -0.088406875 | -0.318714855 | 0.325073485 | -0.325655603 |

Table 9: Religious Population Outliers Excluded

As can be seen from Table 9, removing the three outlier countries reduced the original value by half, putting the results in the "medium" category. This showed that only three countries had made a significant difference.

An alternative line of investigation to pursue was that of immigration. To start with, variables regarding worldwide immigration were selected and run over the years 2015-2018.  The averages of the results are shown below in Table 10.

| Worldwide Immigration | | | |
|---|---|---|---|
| **Additional variable** | **Data Retrieved from** | **Year or Years** | **Pearson's r** |
| Average Number of Immigrants | UN | 2015-2018 | -0.043573129 |
| Migrant Percentage | World Data Bank | 2015-2018 | -0.077588328 |
| Refugees by Population | World Data Bank | 2015-2018 | -0.296886769 |
| Refugee Percentage | World Data Bank | 2015-2018 | -0.04227956 |

Table 10: Worldwide Immigration

All of these results fell below the -0.30 mark which put them in the "weak" category. The only variable that came close to being in the "medium" category was refugees by population which does

come to -0.30 when rounded. However, this doesn't come close to the original -0.51 that was found in the first religion test.

To further explore the concept of immigration, information was found on the OECD website which showed data from immigrants based on continent of birth. The tests were run and the results are summarized below in Figure 6. Although most of the results were understandably negative, most of them tapered off around -0.35, which again doesn't provide an explanation of the -0.51 value as found in the first religion test.

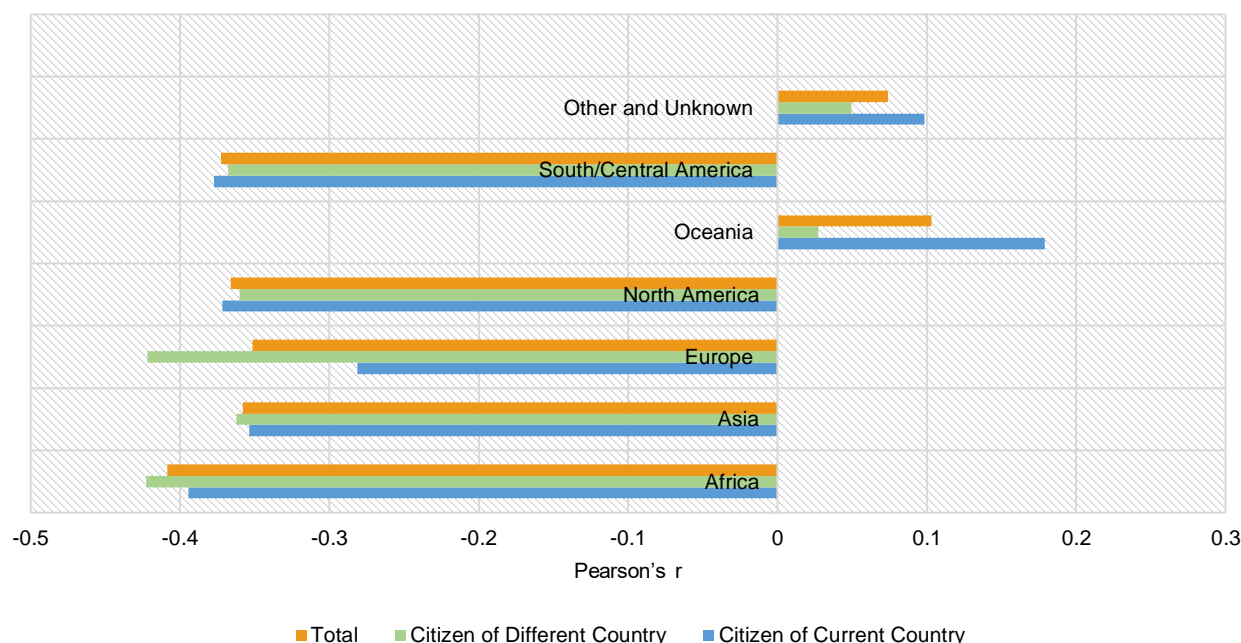## Immigration from Continent of Birth[13]

Figure 6: Immigration from Continent of Birth

A third theory is that this result could be a question of culture as different people might interpret the original question differently. However, only three countries reported detailed census data that broke down the ethnicities in a country so the test remains unrun.

**Conclusion**

In the course of this research, 27 tests were run. As is the case with a lot of research, the results lead to more questions rather than providing concrete answers.

In some ways, the stronger negative results are even more significant than the positive ones as they show the gaps in society's social support and expose vulnerable individuals. The only question that remains concerning these results is do these factors change social support or are they a symptom of it? This question can be applied to every test result. Are more women in the workforce causing stronger social support or are they a symptom of it? Are children dropping out of school because they don't have social support or are they the origin of the lower social support? Are more homicides occurring due to low social support or are they causing it?

These are all important questions to be asked. But given this research, they can now be focused, precise, and informed questions instead of just speculations.

Furthermore, these questions serve as an important reminder that humans don't exist in isolation. Every decision an individual makes affects not only their circle of friends and family, but also their community. Trends in individual decisions affect entire countries and even the world as a whole. Although Pearson's r is not able to account for every individual's decisions, it is able to point towards certain trends that can be observed on a worldwide scale. As these trends transcend across borders, they can serve to have either a positive or negative effect on social support across the globe.

This research has already shown which factors have a strong effect on social support and which factors do not. It has reinforced the author's belief that a better understanding of social support can lead to better system of social support, which in turn, will lead to a better, healthier, and happier society.

**References**

Helliwell, J. F., Layard, R., & Sachs, J. (2012). World Happiness Report 2012. Retrieved from
https://worldhappiness.report/ed/2012/

Helliwell, J. F., Layard, R., & Sachs, J. (2019). World Happiness Report 2019. Retrieved from
https://worldhappiness.report/ed/2019/

Phillips, B. C., & Horowitz, J. E. (2013). Maximizing Data Use: A Focus on the Completion
Agenda. *New Directions for Community Colleges*, *2013*(164), 17–25.
https://doi.org/10.1002/cc.20077

Schnorr-Baecker, S. (2017). Statistical monitoring systems to inform policy decision-making, and new
data sources. *Statistical Journal of the IAOS*, *33*(2), 407–421. https://doi.org/10.3233/SJI-
160324

Stearns, P. N. (2019). A Happy History? *Historian*, *81*(4), 613–626. https://doi.org/10.1111/hisn.13246

**Endnotes**

1. The percentage of total population was taken for the age groups 0-14, 15-64, and 65 or above. The result for group 0-14 was 0.195632904, for group 15-64 was -0.184175705, and for group 65 or above was -0.065055126.
2. Six countries had outdated statistics ranging from 2008-2016.
3. Iceland, Israel, and the Czech Republic excluded due to lack of data.
4. Switzerland, Canada, and Costa Rica excluded due to lack of data.
5. Switzerland, Costa Rica, the United Kingdom, and Belgium excluded due to lack of data.
6. Austria and the United States excluded due to lack of data.
7. Switzerland, Canada, Costa Rica, and the United States excluded due to lack of data.
8. Israel excluded due to lack of data.
9. Costa Rica and Israel excluded from all categories due to lack of data. New Zealand and Australia excluded from "married" and "cohabiting" parents categories due to lack of data. New Zealand excluded from the "other" category due to lack of data. Some data missing for certain years as well.
10. Iceland, Switzerland, Costa Rica, Israel, and Czech Republic excluded. Data for time use from the years 2001-2018.
11. Czech Republic excluded due to unknown being above 15%. Costa Rica did not report Muslim population.
12. Czech Republic excluded due to unknown being above and 15%. Qatar excluded due to lack of social support data. Costa Rica, Mexico, Chile, Guatemala, Brazil, Uruguay, El Salvador, and Slovakia did not report Muslim population.
13. Iceland, New Zealand, Costa Rica, Israel, the United Kingdom, and Germany excluded due to lack of data.