Determining ecological effects on national happiness using model selection

Shiyao Li, Amanda Short, Qianyi Sun, and Michael Wade

DS-5610-01

Data Science Institute, Vanderbilt University, 2414 Highland Ave, Nashville TN 37212

Submitted 13 December 2019

Abstract – National happiness is an important characteristic for governments and NGOs as increased international mobility allows for more individual choice in country of residence. Improving culturally broad elements of national happiness, however, is difficult. Here, we seek to understand the specific relationships between ecological impact and happiness at the national level. National happiness was shown to increase with carbon footprint, though this relationship was strongly driven by gross domestic product (GDP). Countries with weaker economies experience a steeper rise in national happiness with increases in carbon footprint, suggesting that investing in industrial economies in these countries can result in improved national happiness.

I. Introduction

The World Happiness Report, produced annually by the United Nations Sustainable Development Solutions Network, is an aggregation of survey results detailing national happiness for more than 150 countries worldwide[1]. Typically, 1,000 people from each country are surveyed each year. Survey responses are provided by the Gallup World Poll, in which participants are asked to evaluate certain aspects of their lives. This evaluation includes six metrics: economic production, social support, life expectancy, freedom, absence of corruption, and generosity. In order to evaluate each of these aspects, participants are instructed to envision a scale (ladder) from 0 to 10, with higher values signifying better quality of life in that particular category. Serving as a sort of benchmark in the data is the concept of "dystopia," which represents the equivalent of a country that has the minimum possible score in each of the six areas. This country does not actually exist, but is rather a criterion against which all other countries are compared.

Many of the components of national happiness are representative of broad cultural, political, or economic characteristics of a nation. For example, "Freedom to make life choices" is a national average of responses to the question "Are you satisfied or dissatisfied with your freedom to choose what you do with your life?"[2]. For national leaders (or those in a position to enact large-scale change) seeking to improve the happiness of their countries, this information is not wholly useful, even if there is a genuine desire for improvement.

[1] Helliwell, J, Layard, R, & Sachs, J. 2019. World Happiness Report 2019, New York: Sustainable Development Solutions Network.
[2] Helliwell, J, Layard, R, & Sachs, J. 2017. Statistical Appendix for "The social foundations of world happiness". Sustainable Development Solutions Network.

Especially as the economic and physical constraints on human mobility continue to decrease, there is a real and pressing need for each country to offer improved quality of life in order to remain a significant contributor in an ever-globalizing world[3]. With that in mind, it is valuable to identify other metrics that might be more practically adjusted at the national level and result in improvements to national happiness.

There is demonstrated international interest in the human population's environmental impact. Human effects on the Earth's ecology are now both well understood, but still fiercely debated. The number of annual scholarly publications on the subject of climate change and global warming have increased by roughly 500% since the year 2000[4]. Over a similar time period, population-weighted negative affect ("the average frequency of worry, sadness and anger on the previous day") has increased by nearly 25% globally[5]. It stands to reason that there is potentially some relationship between our overall perception of human effects on climate and the frequency of negative affect. Furthermore, it is possible that the perception of each nation's climate mindedness has an effect on that country's national happiness.

In order to determine the existence of such a relationship, we compared national happiness with national ecological footprint. These data were provided by the 2016 Global Ecological Footprint report, and are publicly available[6]. A total of 21 variables, primarily numeric indicators of ecological wellbeing, are included in this report. Footprint-specific metrics comprise a substantial portion of the ecological data, including national usage of "six productive surface areas: cropland, grazing land, fishing grounds, [urban land], forest area, and carbon demand on land." Also included in the data are biocapacity metrics, which refer to the total usable resources at the national level. If footprint exceeds the corresponding biocapacity limit, the country is operating at an ecological deficit, which is unsustainable over time. Conversely, operating with a lower footprint than biocapacity represents an ecological reserve.

Our hypothesis is relatively simple: national happiness shares a relationship with ecological wellbeing. Specifically, though, we predict that the relationship will be different for relatively developed versus undeveloped (or underdeveloped) countries.

II. Methods

All data, including the World Happiness Report (www.kaggle.com/unsdsn/world-happiness) and Global Ecological Footprint (www.kaggle.com/footprintnetwork/ecological-footprint) were retrieved online. We performed all analyses and visualization using R version 3.6.1 (www.cran.r-project.org) and RStudio version 1.2.1335 (www.rstudio.com/products/rstudio). All of the relevant code for this project, including the two necessary datasets in full, are available on GitHub (github.com/dsi-explore/eda-project-happiness-report).

---

[3] Koslowski, R. 2011. Global mobility regimes: A conceptual framework. Global Mobility Regimes. New York: Palgrave Macmillan.
[4] McSweeny, R. 2015. Analysis: The most 'cited' climate change papers. CarbonBrief.org. https://www.carbonbrief.org/analysis-the-most-cited-climate-change-papers.
[5] World Happiness Report 2019.
[6] www.kaggle.com/footprintnetwork/ecological-footprint

Data Exploration – National Happiness

First, we sought to understand geographic trends in national happiness. To do so, we compared average national happiness across regions and continents, as well as at the national level. Additionally, we performed k-means cluster analysis to determine which countries are similar to one another on bases aside from geography and to identify the specific components of national happiness score that each cluster of countries demonstrated. We then generated a map of countries distinguished by clusters.

Data Exploration – Ecological Data

We also explored the distribution of ecological resources around the world. This included the larger measures such as total ecological footprint, biocapacity, and reserve or deficit. We then merged this dataset with the happiness data by country in order to determine any existing correlations between happiness and ecology. Again, we performed k-means cluster analysis to identify similar countries in terms of ecological characteristics.

Model Selection

In order to effectively describe the relationship between national happiness score and the merged ecological data, it was first necessary to remove any variables whose relationship with happiness was not significant. Initially, we attempted a principal components analysis (PCA) approach to arrange the many ecological variables into fewer dimensions and identify those variables that could be removed. We determined that four principal components would be optimal for the ecological footprint data. However, after considering the goals of our analysis, namely to identify specific ecological traits that share a relationship with national happiness, we determined that a simple model selection would be sufficient in isolating the variable(s) that most strongly factored in with national happiness.

We then chose to apply model selection by Akaike Information Criterion (AIC). AIC operates by searching all combinations of input parameters and rewarding those with (a) better least-squares fit against the response variable and (b) fewer parameters included. The relative strength of each model is indicated in part by its ΔAICc value, or the difference in AIC values between each model and the best model. The primary limitation of AIC is that it cannot determine the overall strength of a model, only the relative strength of each combination of parameters included in the test. Still, we expected that among our several ecological variables, several shared strong relationships with national happiness score.

In our case, the response variable was national happiness score, and the input parameters included all of the data collected by the Global Footprint Network's 2016 survey.

Ecological Effects on National Happiness

From the ecological data, we identified carbon footprint as the strongest predictor of national happiness in 2016. However, we also determined that this relationship was strongly driven by the high correlation between carbon footprint and gross domestic product (GDP). Therefore, we

separated the countries into three equal groups based on GDP and compared the relationship of carbon footprint and national happiness between GDP groups.

III. Results

National Happiness

Across all three years, national happiness scores ranged from a minimum of 2.84 to a maximum of 7.62. First, we explored how these scores are distributed geographically. The dataset groups each country into continents and more specific regions. The distributions of happiness scores by region are illustrated in Figure 1[i]. In this case, each country's data across all years is presented as one mean value. Australia and New Zealand demonstrate the highest median happiness score and very little variation due to such a small sample size. Southern Asia and Sub Saharan Africa have the lowest median happiness scores, with the former demonstrating a relatively narrow distribution.

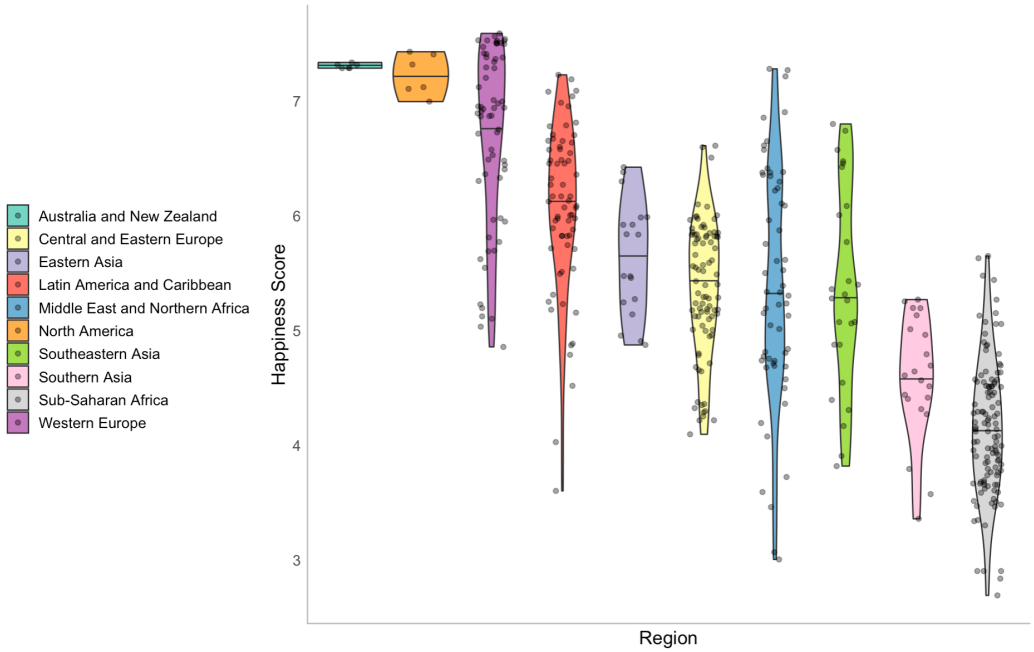


Figure 1. National happiness score by region.

While it is useful to look at each region separately, it is also helpful to combine the data into continents and look at the happiness scores less granularly. Here, we group national happiness scores by continent for each of the three years. Australia has the highest mean happiness, and Africa the lowest (Table 1)[ii].

	National happiness score (mean)		
Continent	***2015***	***2016***	***2017***
Australia	7.29	7.32	7.30
South America	6.36	6.25	6.10
North America	6.15	6.08	6.03
Europe	6.04	6.06	6.10
Asia	5.23	5.23	5.28
Africa	4.46	4.45	4.24

Table 1. Mean national happiness score by continent.

These average values are informative on a broad scale, but do not reflect the degree of variation between countries within each continent. In Figure 2[iii], happiness scores by continent are again reported. While the overall continental happiness medians did not change significantly over the three-year period, there are noticeable differences in the distributions within each continent. For example, North America has fewer countries near the minimum happiness score in 2017 than in 2016 or 2015. These countries near the lower extreme seem to be improving in overall happiness. Alternatively, the distribution of countries within Africa appears to be trending toward a more uniform shape around the median. In other words, African countries that were previously at the upper extremes of happiness score within that continent are trending towards lower happiness scores.

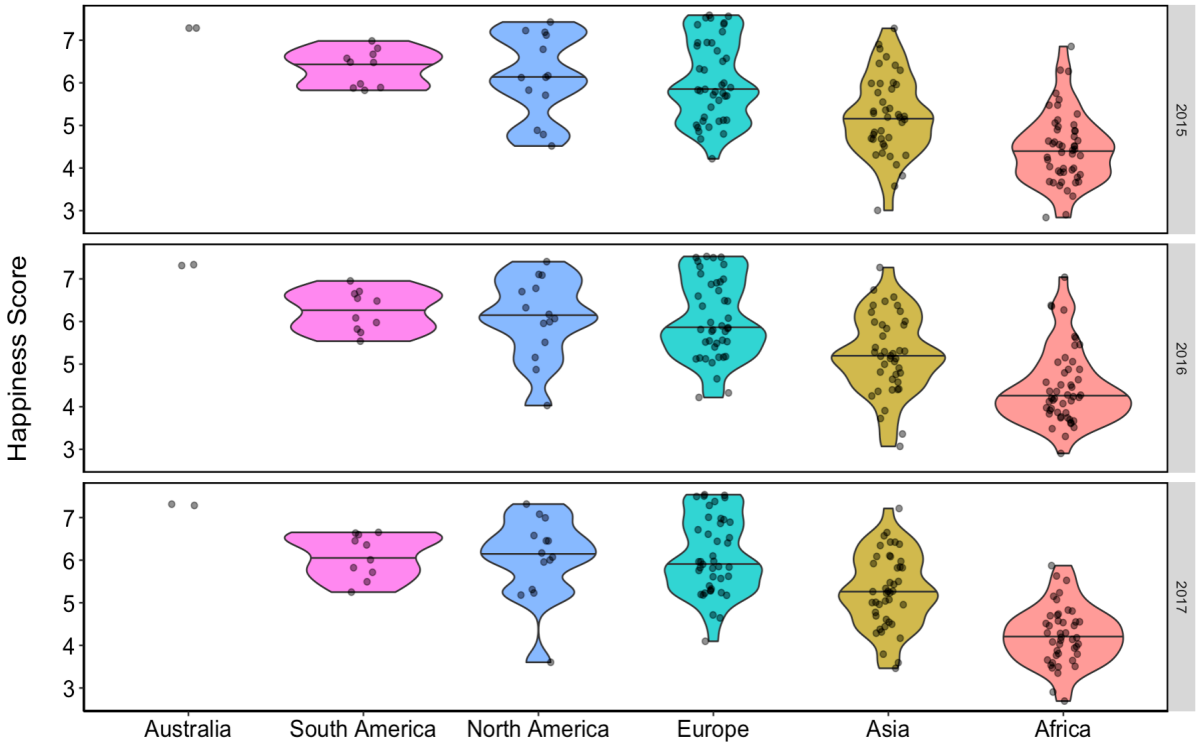


Figure 2. Annual national happiness score by continent.

We then took a different approach to visualizing national happiness across all countries by mapping national happiness scores. In general, the geographic trends that appear in Figure 4 are

also evident in Figure 5 below. Figure 3[iv] contains data from 2017, the year with the most national happiness scores reported. The red to green scale is continuous from the minimum happiness score (2.69) to the highest (7.54). Countries in the Western Hemisphere (North America, South America, Western Africa, and Western Europe) are generally happier than those in the East (Asia and Eastern Africa), although Australia and New Zealand were among the happiest countries in the world in 2017.

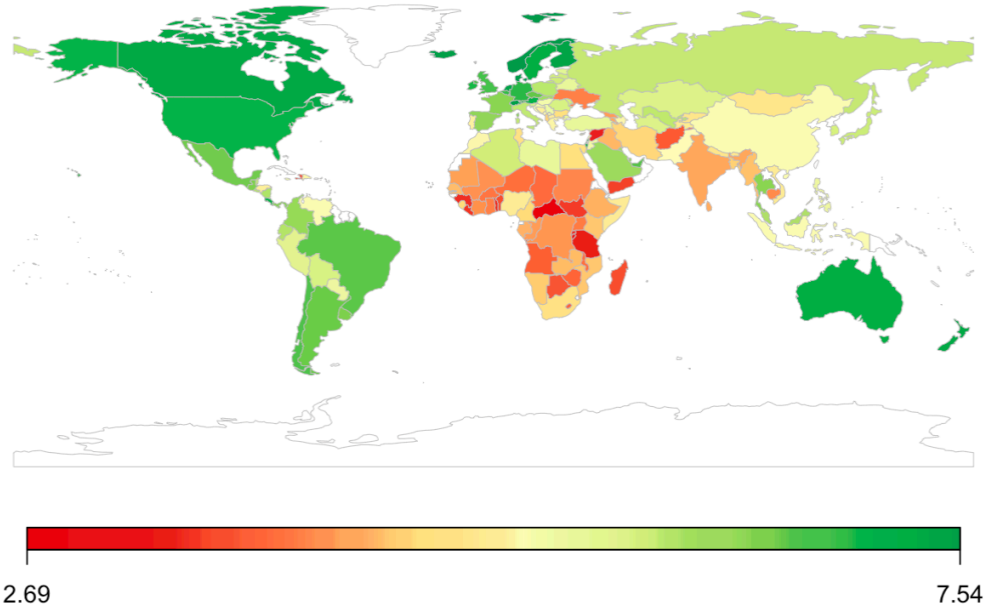


Figure 3. National happiness scores (2017).

For the national happiness data, within-sum-of-squares (wss) was used to identify the optimal number of clusters (Figure 4)[v]. These clusters are visualized (Figure 5a)[vi], along with the specific happiness components of each cluster (Figure 5b)[vii].

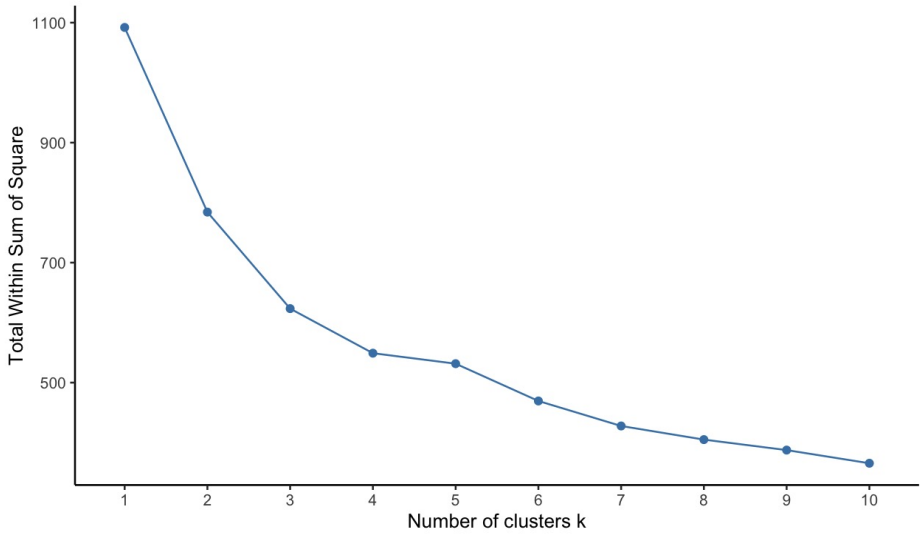


Figure 4. Total within sum of squares "elbow" plot for happiness data.

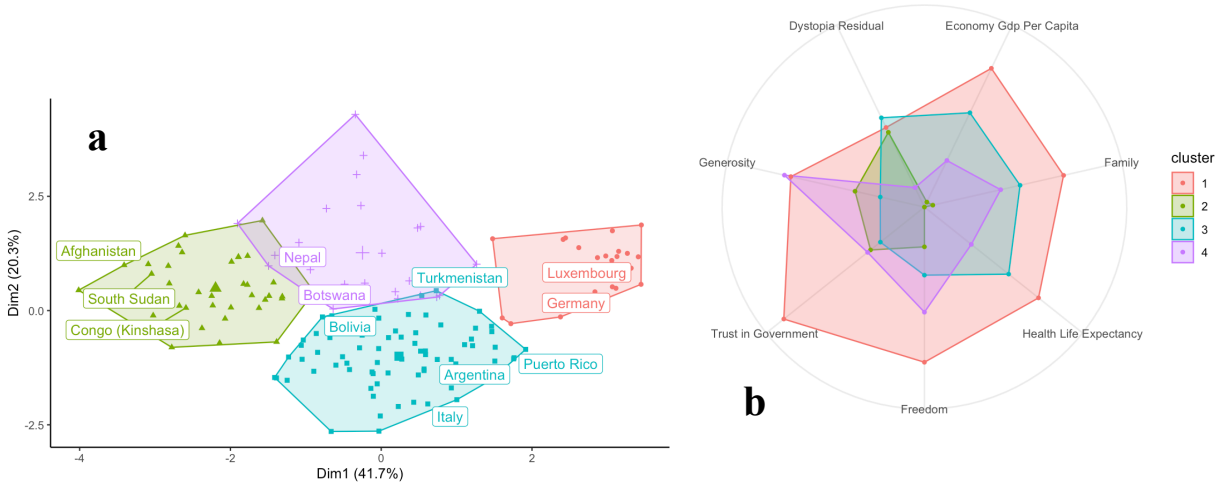


Figure 5. (a) Clustering of countries by k-means and components of national happiness. (b) Happiness components of each cluster.

Many of the factors of happiness present in this dataset are challenging to manipulate on a large scale. In an effort to find other factors that would correlate with happiness and be more easily controlled or managed, we joined the World Happiness Report data with data from the 2016 Global Ecological Survey. This dataset contains information on emissions, footprint, and other environmental factors at the national level.

Ecological footprint

Total ecological footprint measures a country's use of six different "productive surface areas." These include "cropland, grazing land, fishing grounds, built-up (or urban) land, forest area, and carbon demand on land." The Africa and Asia-Pacific regions have the lowest median ecological footprint, but there is a significant amount of variation among countries within each region (Figure 6)[viii]. North America has the highest median ecological footprint of any region, but only three countries are contained within that region. Luxembourg, the country with the highest total ecological footprint, lies within the European Union region. Luxembourg is a relatively small country, but its environmental performance is historically among the least sustainable in the world[7]. Aruba, Qatar, and Australia had the second, third, and fourth highest average ecological footprint, respectively, from 2015 – 2017.

[7] Environmental Performance Reviews: Luxembourg (2010). Organisation for Economic Cooperation and Development (OECD).

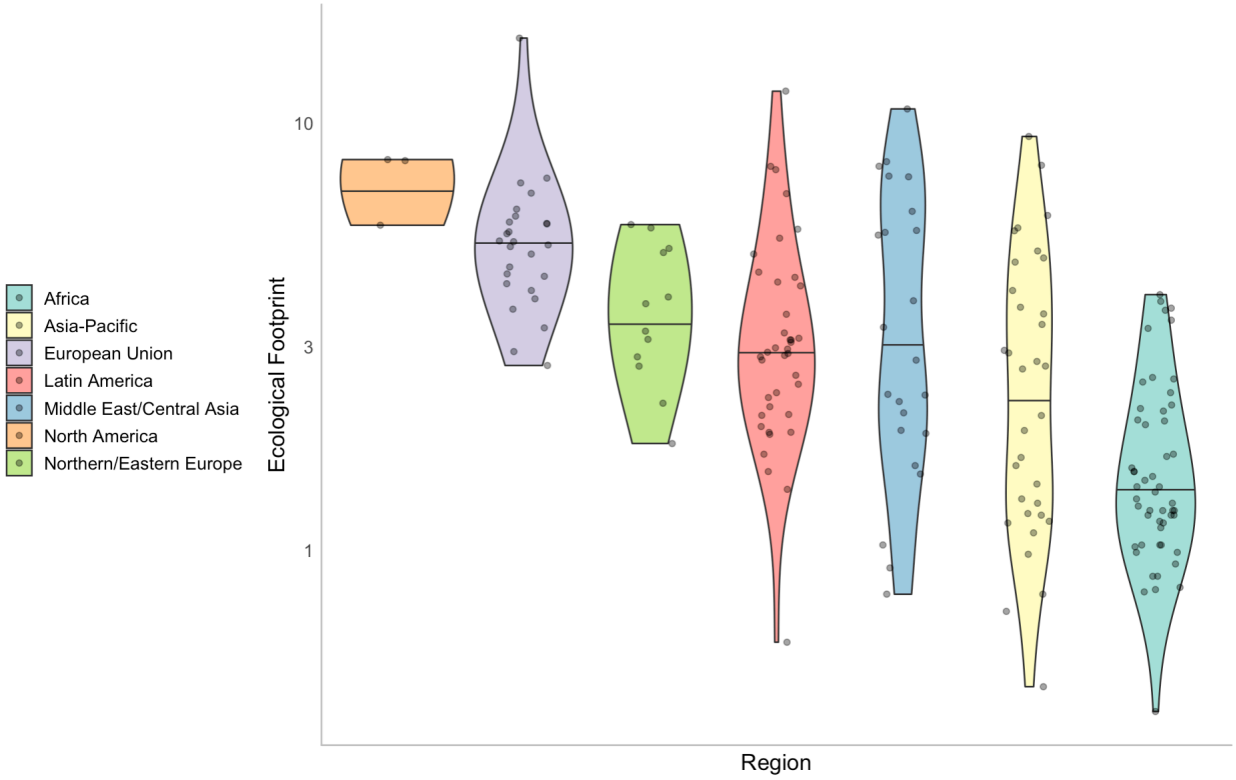


Figure 6. Total ecological footprint by region

Model Selection

We selected all of the variables present in the top five models as those to include in subsequent analyses: carbon footprint, cropland, fish footprint, urban land, Earths required, forest footprint, and total ecological footprint. Notably, fish footprint and urban land appeared in each of the top models, suggesting that these are relatively strong predictors of national happiness. The results of our AIC are presented in Table 2[ix] below.

Model	Parameters	ΔAICc
1	Carbon footprint, cropland, fish footprint, urban land	0.00
2	Earths required, fish footprint, forest footprint, urban land	0.66
3	Fish footprint, forest footprint, total ecological footprint, urban land	0.66
4	Fish footprint, total ecological footprint, urban land	1.06
5	Earths required, fish footprint, urban land	1.07

Table 2. AIC results for ecological predictors of national happiness score.

Our focus is to identify specific ecological factors that best predict national happiness. Because of the cumulative nature of total ecological footprint and Earths required (they are summary measures based on smaller components), we chose to omit those from our analyses.

With the five parameters identified, we then explored the relationships between each parameter and our output variable: national happiness score. First, we used a correlogram to easily visualize these differences (Figure 7)[x]. There are relatively strong, positive correlations between happiness score and carbon footprint (correlation = 0.61), as well as urban land and forest footprint (correlation = 0.54) and happiness score and urban land (correlation = 0.51).

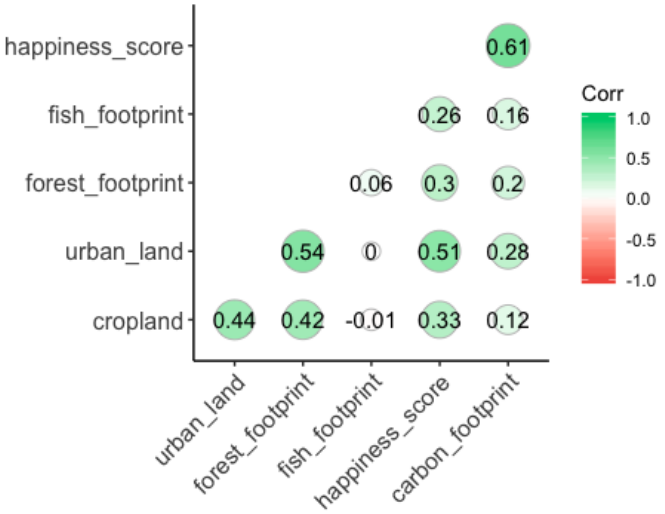


Figure 7. Correlations between ecological variables and national happiness score.

For the ecological footprint data, within-sum-of-squares (wss) was used to identify the optimal number of clusters (Figure 8)[xi]. These clusters (Figure 9a)[xii], as well as the specific components of each (Figure 9b)[xiii], are presented below. Two clusters contain countries with high values in just one specific group. The green cluster contains countries with a very high fish footprint only. The red cluster contains countries with a very high carbon footprint only. The blue cluster, which contains the vast majority of countries, have very low values in all five categories.

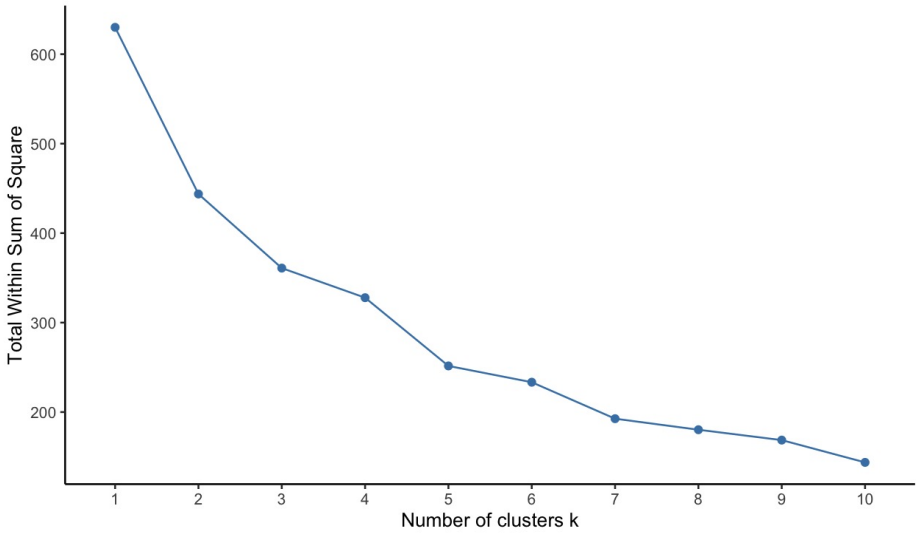


Figure 8. Total within sum-of-squares "elbow" plot for ecological data.

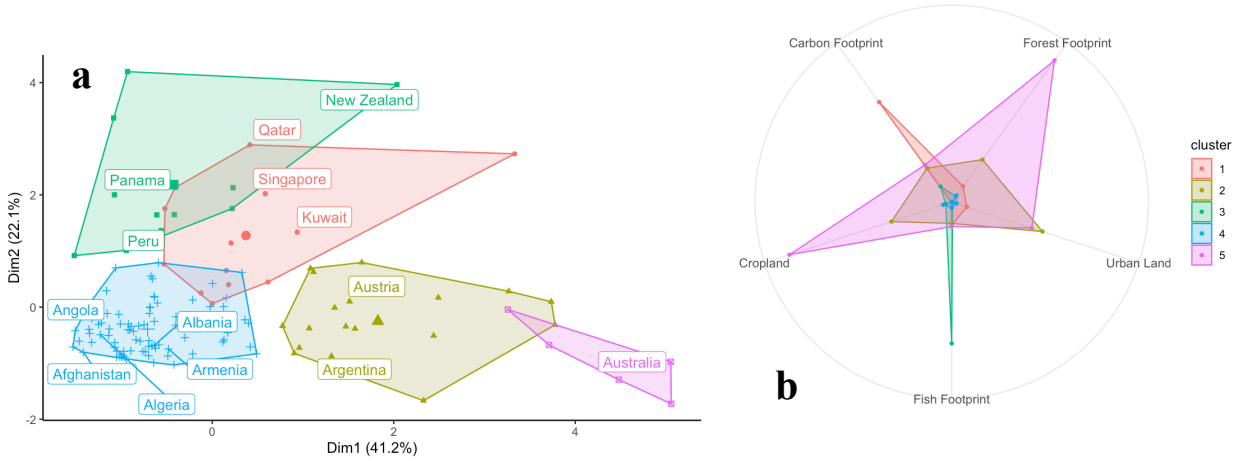


Figure 9. (a) Clustering of countries by k-means and selected ecological variables. (b) Ecological components of each cluster.

Regional trends in GDP per capita (Figure 10a) are similar to those in carbon footprint (Figure 10b), with the notable exception of Northern and Eastern Europe. African countries have the lowest median GDP per capita and carbon footprint, but the distribution is quite wide. North American countries have the highest median GDP per capita and carbon footprint, and the small number of countries in this region contributes to a very narrow distribution in GDP and no visible variation in carbon footprint.

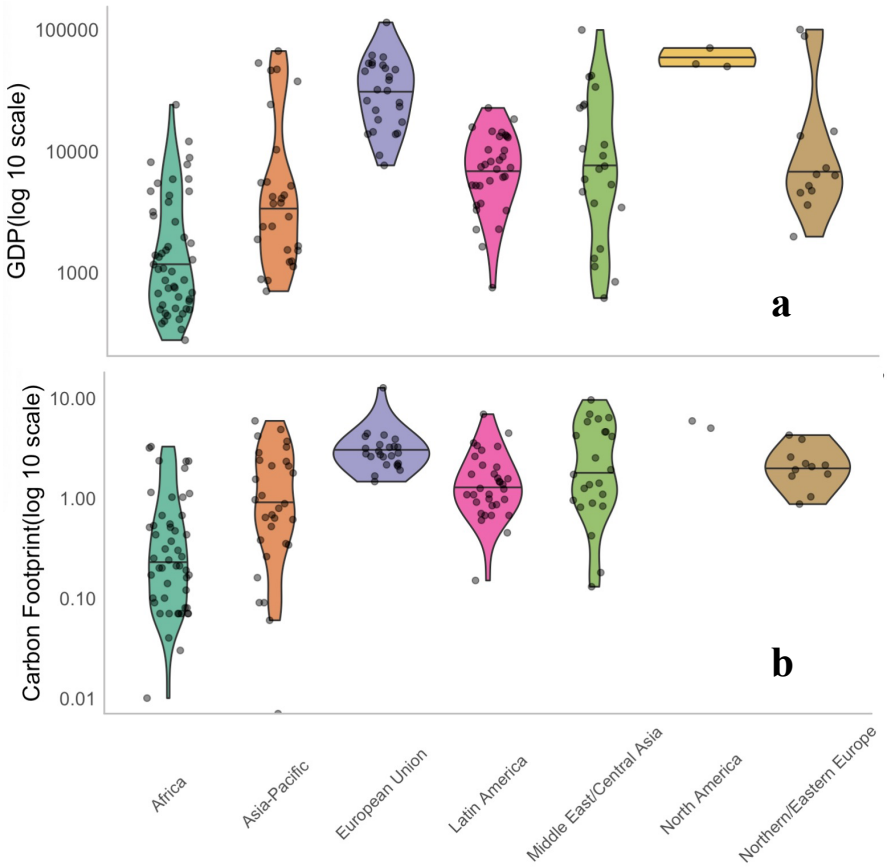


Figure 10. (a) GDP per capita within each continent. (b) Total carbon footprint within each continent.

National per-capita GDP and national carbon footprint indeed share a strong positive relationship, with a correlation of 0.82 (Figure 11). There is considerable variation in both carbon footprint and GDP per capita between countries, with a minority of countries exhibiting relatively high values in both categories.

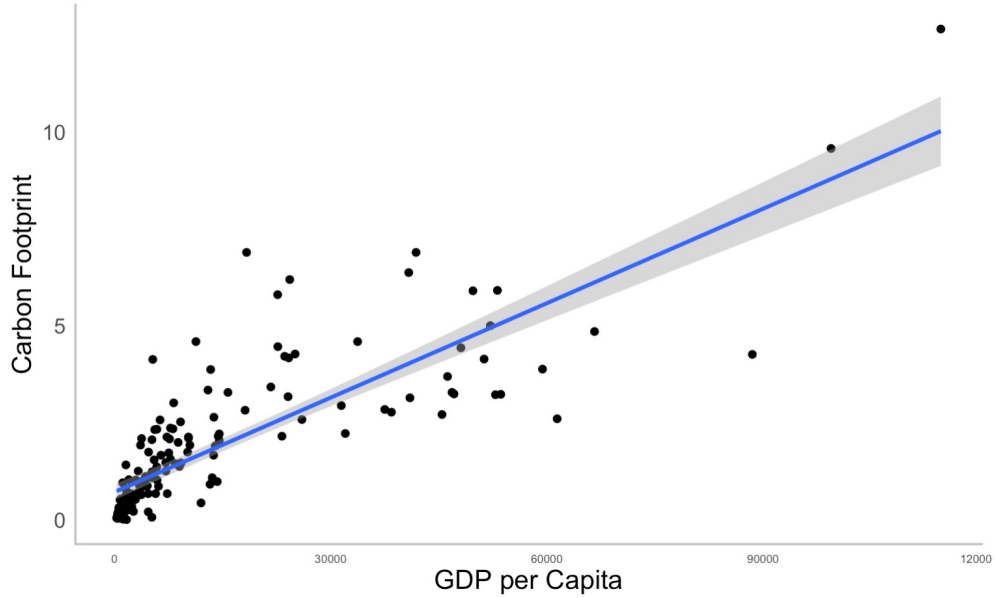
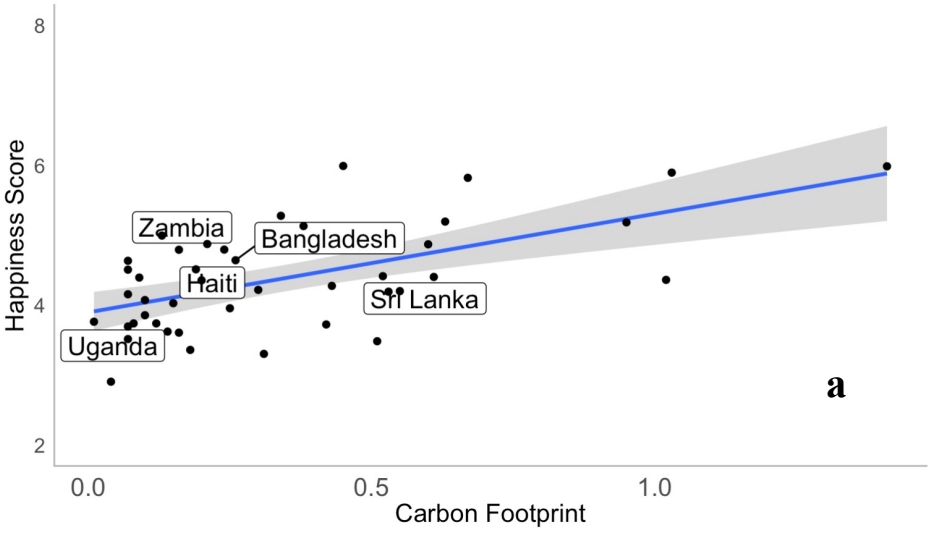


Figure 11. Total carbon footprint vs GDP per capita.

Ecological Effects on National Happiness

Countries were arranged into three classes based on GDP quantiles: the lowest third (Figure 12a), middle third (Figure 12b), and highest third (Figure 12c). While a positive relationship between national happiness score and carbon footprint is observed across GDP classes, the relationship was significantly stronger for countries among the weakest economies ($R^2 = 0.35$) than for average economies ($R^2 = 0.02$) or the strongest economies ($R^2 = 0.04$). In general, countries within the highest GDP class have higher national happiness scores and carbon footprints than those with lower GDPs.



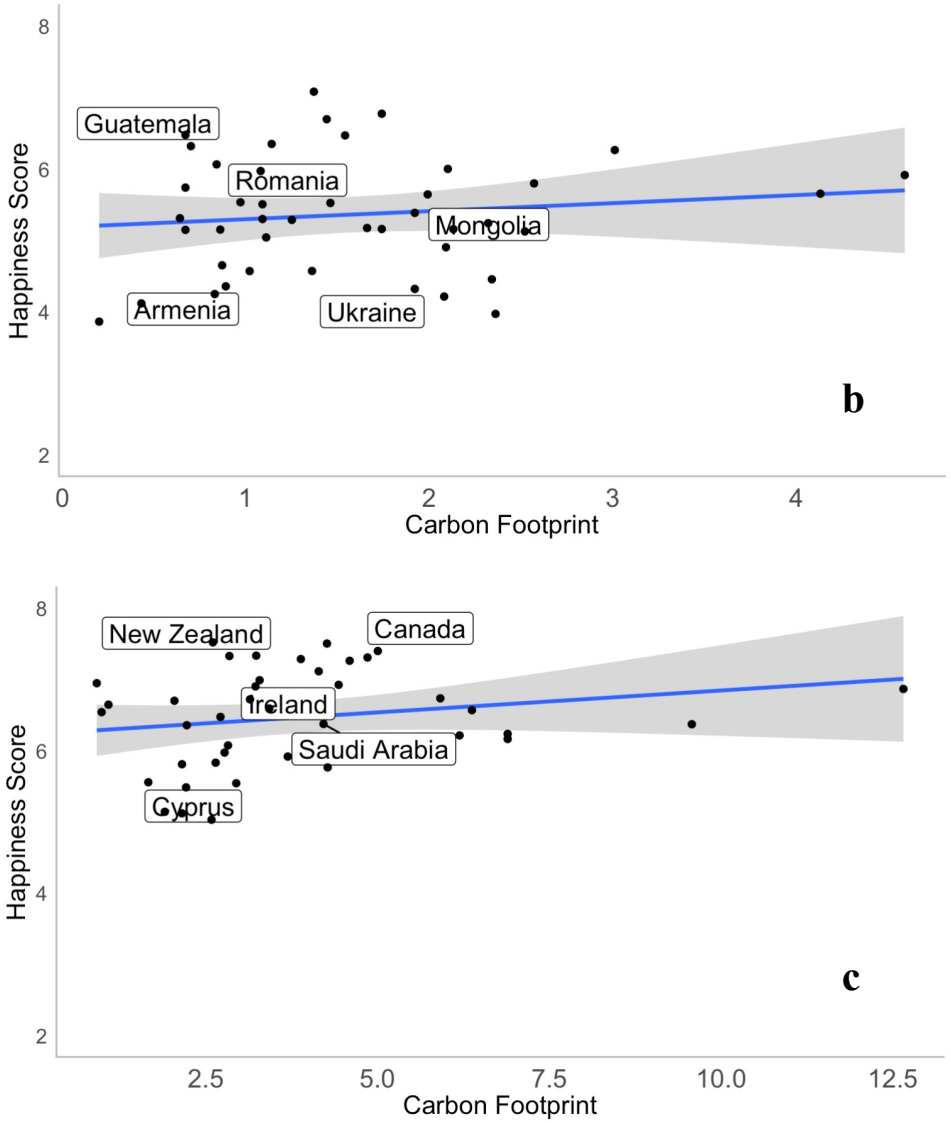


Figure 12. (a) National happiness score vs carbon footprint in countries among lowest third of global GDP, (b) middle third of global GDP, and (c) highest third of global GDP.

Our hypothesis, that national happiness shares a relationship with ecological wellbeing and that the relationship is different for relatively developed versus undeveloped (or underdeveloped) countries, is confirmed. Changes in carbon footprint result in more significant variation to national happiness in weaker economies than stronger economies.

IV. Conclusion

Continental trends in national happiness do not vary significantly over brief spans of time, but national variation within each continent is observed. Australia and New Zealand maintain the highest national happiness among regions, as those few countries within this region feature near the upper bound of global national happiness. Countries in Africa and Asia are consistently among the unhappiest nationally, and each display a trend of distributing more closely around the continental median over time.

North American countries exhibit the highest median total ecological footprint of any region, and each North American country maintains a higher total ecological footprint than the median of any other region on Earth. There is a clear difference in medians between regions in terms of total ecological footprint, but the distributions are widespread in nearly all regions.

Countries with a higher carbon footprint tend to have higher national happiness. However, a key driver of this relationship is the effect of GDP on carbon footprint and therefore happiness. Per-capita GDP and national carbon footprint are highly correlated, suggesting that countries with stronger economies are also producing more carbon, as expected. Countries with the weakest GDPs exhibit a much stronger positive relationship between carbon footprint and national happiness, though some positive trend is observed across GDP classes.

There are certain ethical implications that accompany these results. Increasing carbon footprint (by way of enhancing industrial activity or output) will likely have negative impacts on several factors, including environmental or human health. Doing so for the sake of increasing national happiness should not be done without thorough understanding of potential environmental impact. Considering the many alternatives to high-emissions industries available today, countries should look to specific avenues such as widening internet access to boost gross domestic product without substantially affecting carbon output.

There are two countries globally, Bhutan (South Asia) and Suriname (South America), which have reportedly achieved carbon neutrality. As of this decade, the former has actually achieved a carbon-negative status[8,9]. In both cases, this can be attributed to organized efforts by indigenous peoples to protect the vast natural forest reserves in each country. These forest reserves are capable of capturing and storing more carbon than their relatively small populations emit on an annual basis, providing the possibility of carbon neutrality. It is not feasible in many countries, specifically those without relatively large forest land within which to sequester carbon. More than two dozen other countries, primarily within Europe, have pledged to follow suit under a specific timeline, though even making such a pledge requires a relatively unique combination of circumstances[10].

Changes in ecological footprint may result in changes to national happiness. However, while there is a clear international interest in methods for improving environmental impact, the relationship between carbon footprint and national happiness is still positive. Counteracting this relationship to improve both environmental impact and national happiness often requires a rare combination of natural resources, state policy, and wealth. Therefore, while there is a demonstrable relationship between ecological footprint and national happiness, other routes for improving the latter are likely more practical for most of the world's countries.

---

[8] Aravindh, MA & Giri, GP. 2016. An overview of the solar energy utilization in Bhutan. Concurrent Advances in Mechanical Engineering 2(2):1-7.
[9] Banerjee A & Bandopadhyay R. 2016. Biodiversity hotspot of Bhutan and its sustainability. Current Science 110(4):521-527.
[10] Flagg JA. 2014. Aiming for zero: What makes nations adopt carbon neutral pledges? Environmental Sociology 1(3):202-212.

Methodology for tables and figures located at the following:

[i] master/final_project.Rmd, chunk 2

[ii] master/final_project.Rmd, chunk 6

[iii] master/final_project.Rmd, chunk 5

[iv] master/final_project.Rmd, chunk 7

[v] master/cluster_for_happiness.Rmd, chunk 5

[vi] master/cluster_for_happiness.Rmd, chunk 17

[vii] master/cluster_for_happiness.Rmd, chunk 11

[viii] master/EDA_footprint.Rmd, chunk 18

[ix] master/aic_footprint.Rmd, chunk 4

[x] master/aic_footprint.Rmd, chunk 6

[xi] master/cluster_for_footprint, chunk 6

[xii] master/cluster_for_footprint, chunk 17

[xiii] master/cluster_for_footprint, chunk 12