BIOS 611 Project Report

Global Indicator Data Analysis

Brooke Felsheim

1. Introduction

The future existence of humankind is dependent on our ability to live sustainably. As human populations rise along with greenhouse gas emissions, deforestation rates, and generation of waste, we will continue to deplete natural resources, disrupt ecosystems, and increase global temperatures, leading to an unsustainable future. Because of this, it is critical to study environmental indicators to assess the current state and trajectory of the environment.

For my BIOS 611 project, I chose to analyze global environmental indicator data along with global economic and happiness indicator data. My goal was to assess recent environmental trends of countries around the world and to see how these trends might correspond with the state of the economy and measured levels of happiness within the countries.

2. Source data description

There were three types of source data sets used for this analysis: environmental indicator data, economic indicator data, and happiness indicator data. Each data type contains quantitative indicator measures by country and year.

Environmental indicator data

The environmental indicator source data come from the United Nations Statistics Division (UNSD) / United Nations Environment Programme (UNEP) Questionairre on Environment Statistics. The data were downloaded via Kaggle here (last updated June 5, 2021). Multiple types of environmental indicator data were used in this analysis and fall under the categories of air and climate, biodiversity, energy, forest, inland water resources, land and agriculture, natural disasters, and waste. Environmental indicator data are available within the year range 1990-2020.

Economic indicator data

The economic indicator source data come from the UNSD Human Development Report and were downloaded via Kaggle here (last updated August 11, 2020). The primary measure of economic activity used for this analysis was gross domestic product (GDP) by country. Economic indicator data are available within the year range 1990-2018.

Happiness indicator data

The happiness indicator data come from the World Happiness Report published by the Sustainable Development Solutions Network. The data were downloaded via Kaggle here (last updated November 26, 2019). Each country is given a "happiness score" (0 to 10) that is based on life evaluation survey responses. Happiness indicator data are available within the year range 2015-2019.

3. Results

Exploration of indicator trends within countries

The first goal of my analysis was to explore trends of indicator data within individual countries. To achieve this goal, I created an interactive R shiny app that plots many different types indicator data over time for 190 different countries. The country of interest can first be selected via a drop-down menu in the app. For the selected country, thirteen different types of plots are generated:

- Environmental indicator plots
  - Greenhouse gas emissions by type over time
  - Greenhouse gas emissions by sector
  - Energy supply per capita over time
  - Renewable energy production percentage over time
  - Forest area over time
  - Precipitation over time
  - Natural disaster occurrences over time
  - Natural disaster deaths over time
  - Hazardous waste by type over time
  - Municipal waste recycled over time
- Economic indicator plots
  - Gross domestic product per capita over time
  - Gross national income by gender over time
- Happiness indicator plots
  - Happiness score over time

The plots displayed in the shiny app can give insight into the level and ways that a country may be negatively affecting the environment, the status of a country's economy, and the estimated happiness level of a country's citizens over time.

As an example, we can look at all of the indicator plots generated for Sweden in the shiny app. From this data, we can see that Sweden's greenhouse gas emissions have been decreasing over time, and that most of these greenhouse gas emissions come from energy use. Correspondingly, energy supply per capita has been decreasing over time and the total percentage renewable energy production increasing over time. The total forest area by year in Sweden increased from 1990-2000, but decreased from 2000-2020. While the total precipitation fluctuates year by year in Sweden, the indicator plot shows a general trend of increased precipitation since 1990. Additionally, Sweden has had very few recent natural disasters, treats/disposes of approximately half of its hazardous waste, and has been increasing the percentage of municipal waste it recycles. We can also see that Sweden's GDP has been steadily rising over time, and while the national income has been rising as well, it remains higher for men than women. Furthermore, Sweden's happiness score has only fluctuated by less than 0.1 out of 10 from 2015-19.

We can also notice some other notable trends from the indicator plots generated by the shiny app. In Brazil, for example, where the Amazon rainforest is located, forest area has been decreasing over time. From

1990-2020, Brazil has lost approximately 10 million hectares of forest area. Additionally, we can see that the gross national income has been increasing at a higher rate for men than for women over time in India. Furthermore, in Japan, we can see that the number of meteorological natural disasters (hazards caused by short-lived extreme weather and atmospheric conditions such as extreme temperature or storms) has been steadily increasing over time. This could be due to the effects of global warming. We also see a large number (approximately 20,000) of deaths from geophysical natural disasters (hazards originating from solid earth such as earthquakes or wildfires) between 2010-2019 in Japan. This can be explained by the 9.0 magnitude 2011 Tōhoku earthquake and tsunami.

Energy consumption trends across countries

While the R shiny app displayed various indicator trends over time for individual countries, I also wanted to look at the indicator trends over time between countries. I chose to focus on energy consumption trends for this analysis. I was interested in what countries were consuming the greatest amounts of energy (both in total and per capita), as well as how much of this energy was renewable.

First, I looked at energy consumption trends of the ten countries with the highest total levels of energy consumption between 1990-2017 (Figure 1). Countries were ranked by their highest year of energy consumption in petajoules over the 27 year period. For the top ten energy-consuming countries, I then plotted the total energy consumption by year (Figure 1a) along with the percent of total energy consumption that is renewable by year (Figure 1b). From these plots, we can see that there are two countries with a much higher level of energy consumption than any other country: the United States and China. The United States started as the highest energy-consuming country in 1990, and the amount of energy consumed per year has remained approximately steady since then. Meanwhile, we see that China's energy consumption has greatly increased from 1990 to 2017, surpassing the United States in 2009 as the highest energy-consuming country. While the United States and China are by far the most energy-consuming countries, less than 25% of the energy they consume is renewable. We can see that of the ten countries with the highest total energy consumption, only Brazil and Canada source more than 50% of their total energy from renewable sources.

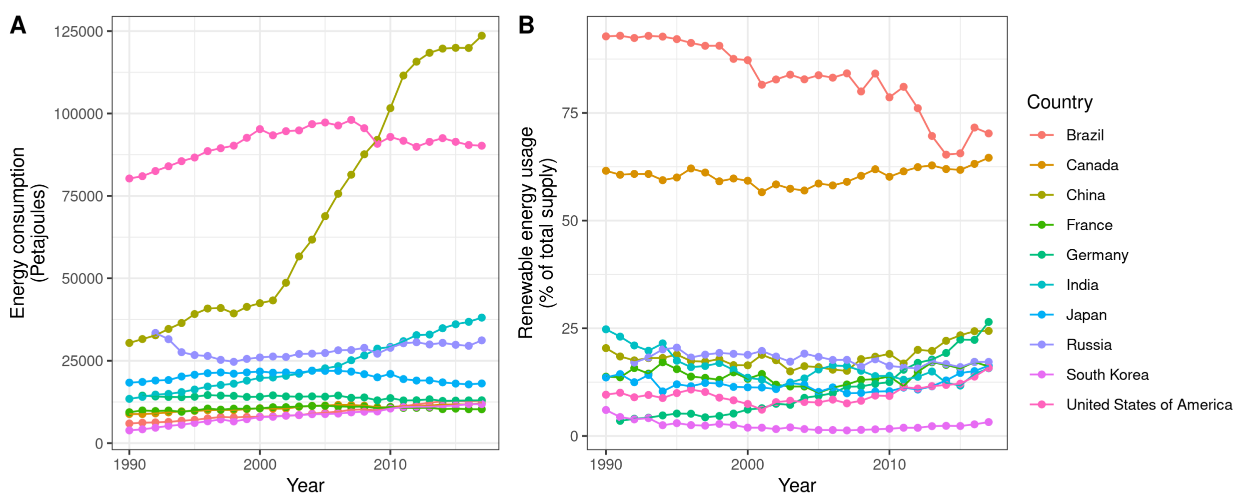

Figure 1: Energy consumption trends for top ten energy-consuming countries. **(a)** Total energy consumption by country over time. **(b)** Percent of total energy consumed that is renewable by country over time.

Next, I looked at energy consumption trends of the ten countries with the highest levels of energy consumption per capita between 1990-2017 (Figure 2). Countries were ranked by their highest year of energy consumption in gigajoules per capita over the 27 year period. Interestingly, none of the top ten energy-consuming countries per capita overlapped with the top ten total energy-consuming countries. For the top ten energy-consuming countires per capita, I then plotted the energy consumption per capita by year (Figure 2a) along with the percent of total energy consumption that is renewable by year (Figure 2b). We can see from these plots that the two countries with the highest energy consumption per capita as of 2017 are Iceland and Qatar. Iceland's energy consumption per capita has been steadily increasing over time, while Qatar's energy consumption per capita experienced a sharp increase followed by a general decrease over time. We also see that almost all of Iceland's energy comes from renewable sources, which makes sense as it is home to many volcanoes and hot springs that can be used for geothermal energy. On the other hand, we see that almost none of Qatar's energy comes from renewable sources, which also makes sense as it is home to some of the world's largest natural gas and oil reserves.

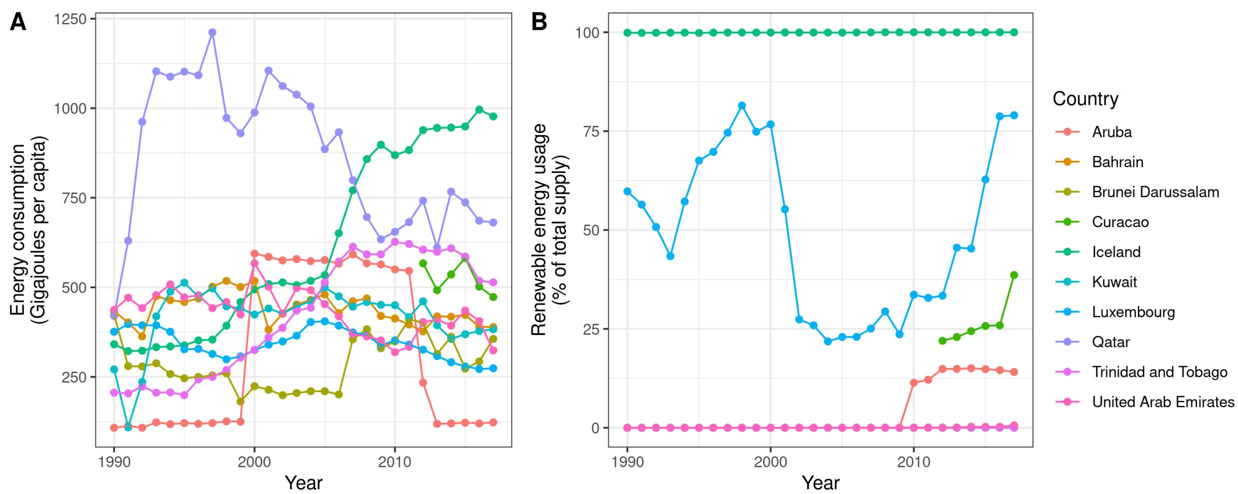

Figure 2: Energy trends for top ten energy-consuming countries per capita. **(a)** Energy consumption per capita by country over time. **(b)** Percent of total energy consumed that is renewable by country over time.

Exploration of trends between indicators

```
## Importance of components:
##                           PC1    PC2    PC3    PC4     PC5    PC6
## Standard deviation     1.4154 1.2902 0.9436 0.8999 0.66219 0.4395
## Proportion of Variance 0.3339 0.2774 0.1484 0.1350 0.07308 0.0322
## Cumulative Proportion  0.3339 0.6113 0.7597 0.8947 0.96780 1.0000
```

Prediction of happiness level from environmental indicator data

```
## glmnet
##
## 86 samples
##  6 predictor
##  2 classes: 'Low', 'High'
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 77, 77, 78, 77, 78, 78, ...
## Resampling results across tuning parameters:
##
##   alpha  lambda        Accuracy   Kappa
##   0.10   0.0005588992  0.7805556  0.5610976
```

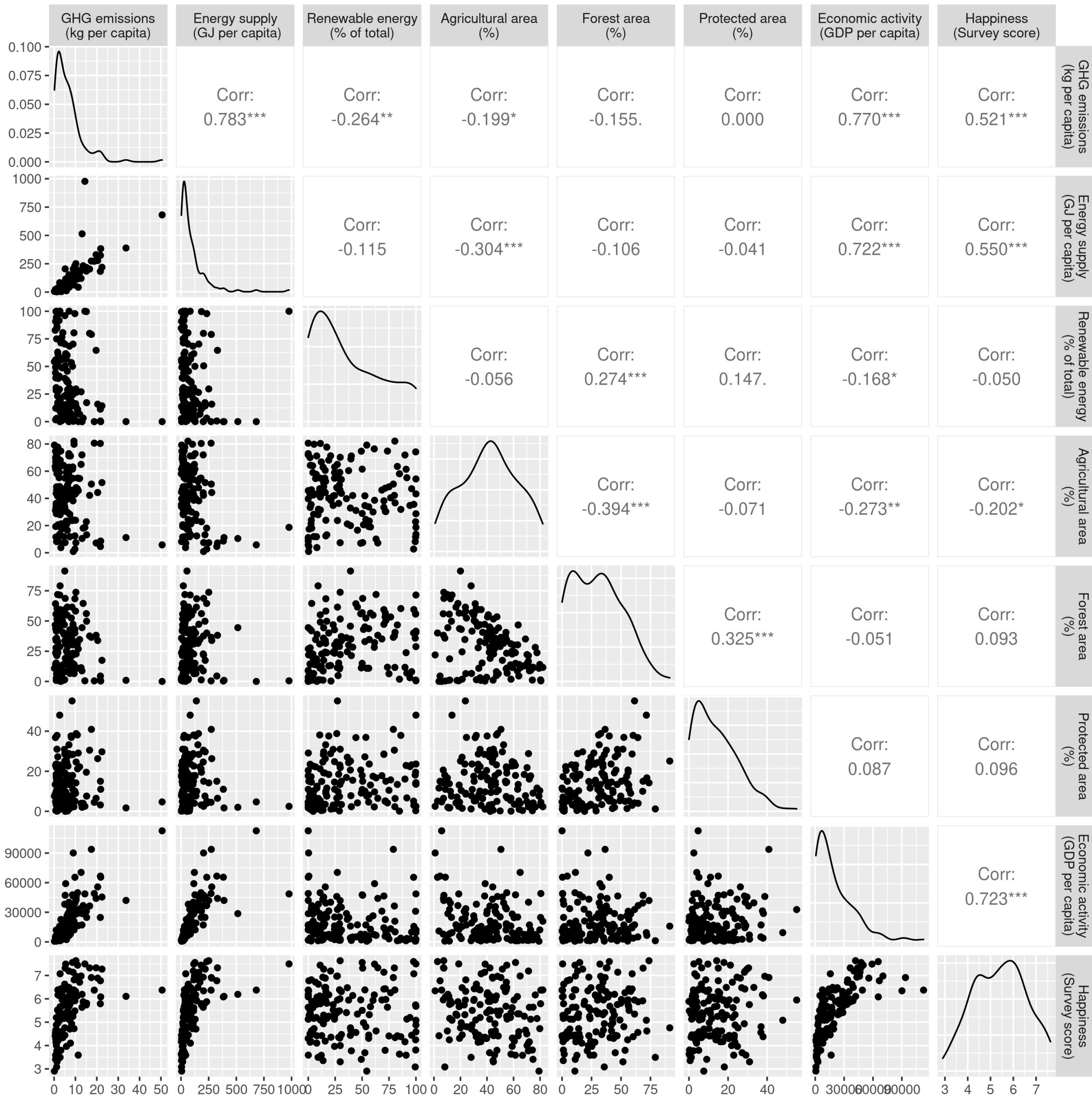

Figure 3: Paired indicators

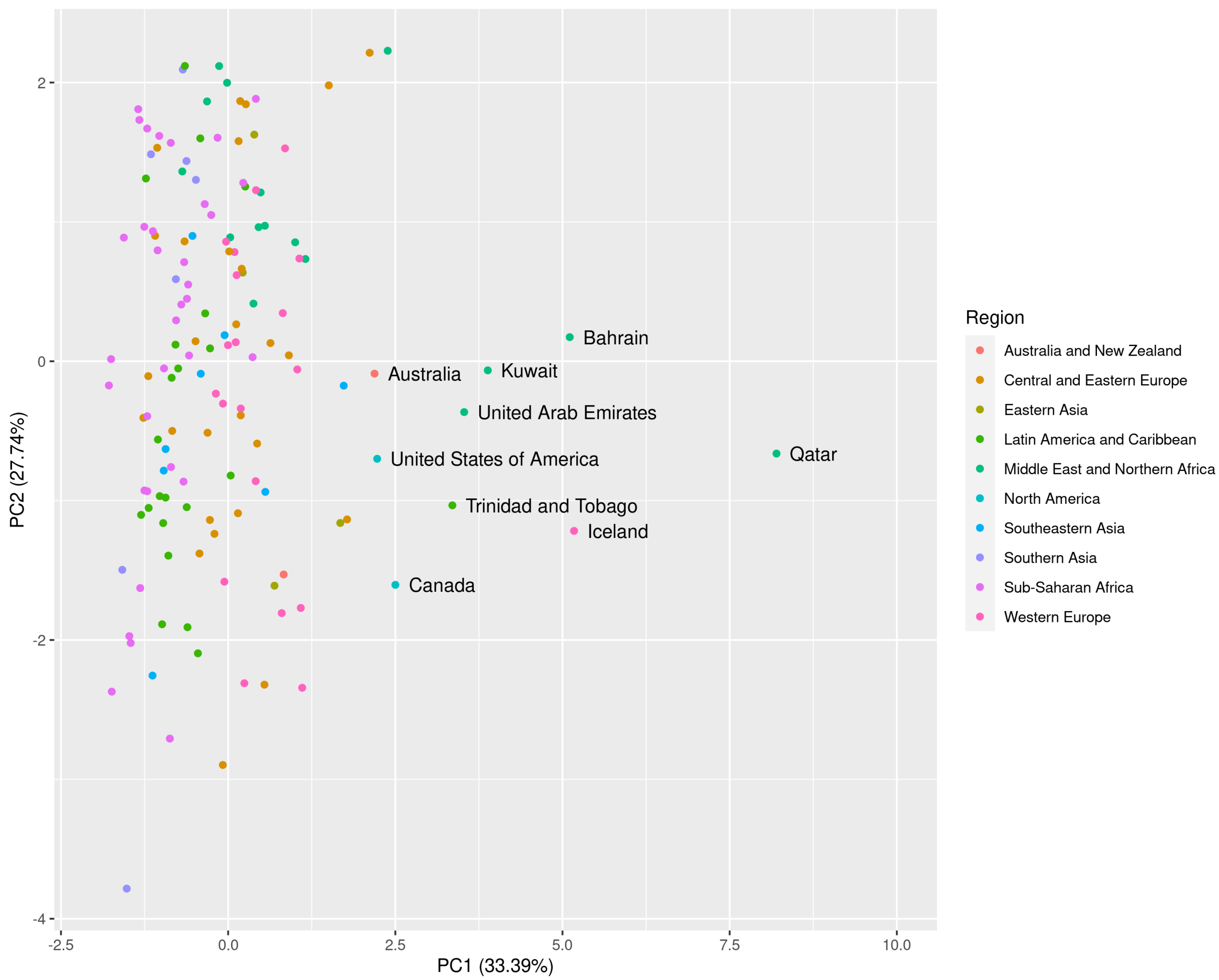

Figure 4: Environmental indicator PCA

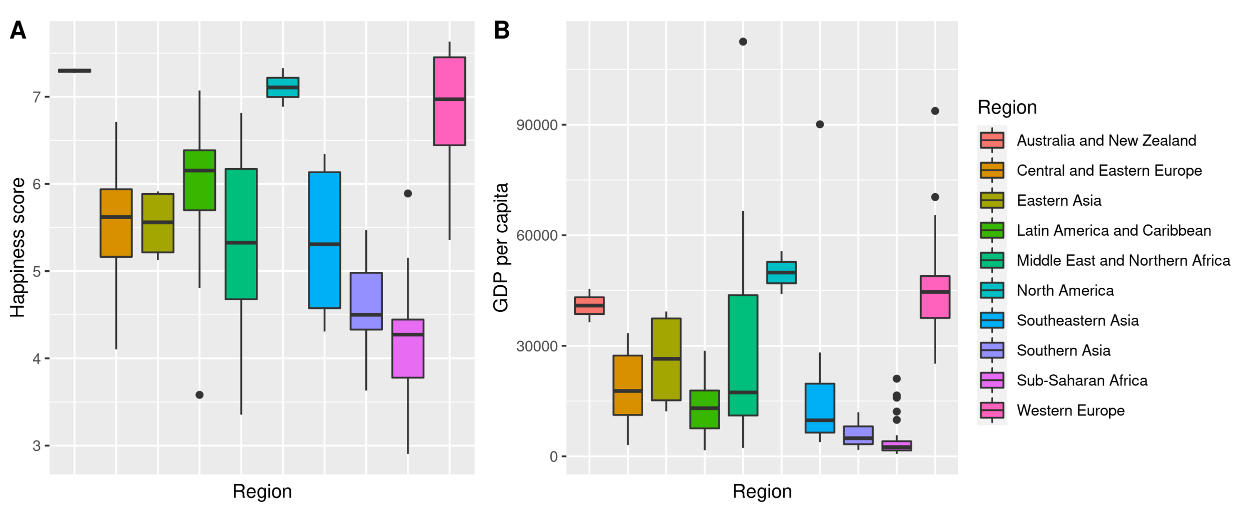

Figure 5: Region boxplots

```
##   0.10   0.0055889918  0.7594444  0.5202439
##   0.10   0.0558899181  0.7594444  0.5202439
##   0.55   0.0005588992  0.7805556  0.5610976
##   0.55   0.0055889918  0.7594444  0.5202439
##   0.55   0.0558899181  0.7783333  0.5603833
##   1.00   0.0005588992  0.7805556  0.5610976
##   1.00   0.0055889918  0.7694444  0.5402439
##   1.00   0.0558899181  0.7758333  0.5553833
##
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were alpha = 0.1 and lambda = 0.0005588992.
```

```
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                                   s1
## (Intercept)                -4.02755295
## GHG_per_capita_emissions    0.29352022
## Energy_per_capita           0.02136925
## Renewable_energy_percent    0.01427311
## Agricultural_area_percent   0.01430855
## Forest_area_percent         0.01609668
## Protected_area_percent     -0.03236558
```

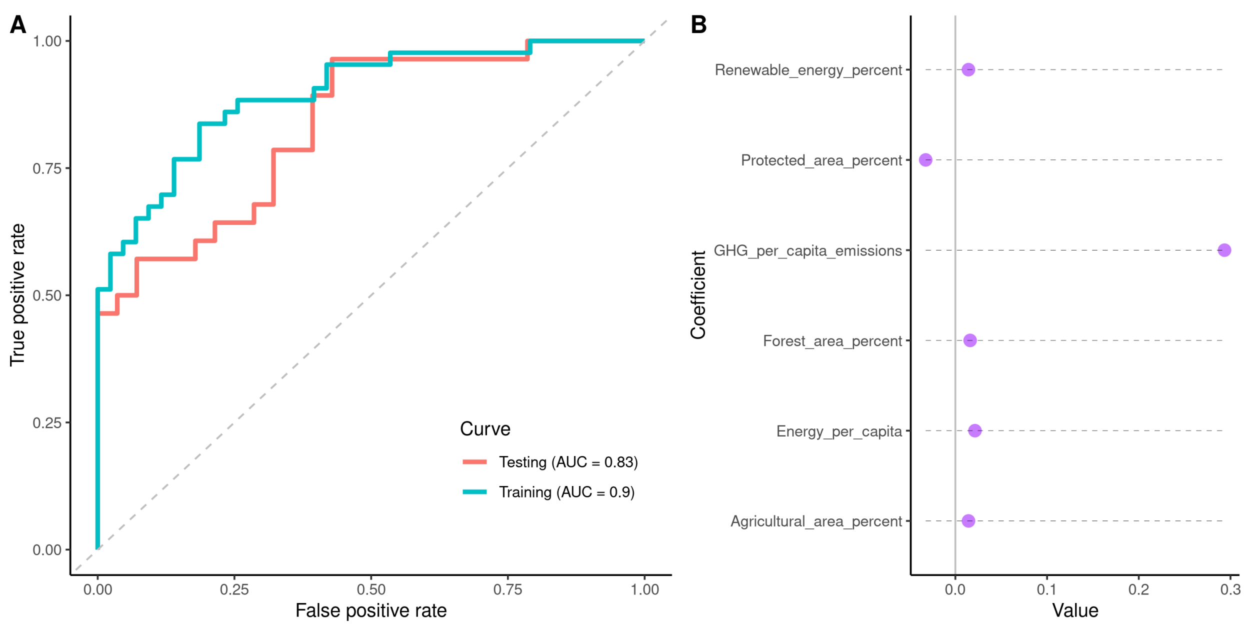

Figure 6: Happiness predictor

Prediction of GDP level from environmental indicator data

```
## glmnet
##
## 86 samples
##  6 predictor
##  2 classes: 'Low', 'High'
```

```
##
## No pre-processing
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 77, 77, 78, 77, 78, 78, ...
## Resampling results across tuning parameters:
##
##   alpha  lambda        Accuracy   Kappa
##   0.10   0.0006950761  0.9305556  0.8599719
##   0.10   0.0069507607  0.8844444  0.7680206
##   0.10   0.0695076065  0.9094444  0.8180206
##   0.55   0.0006950761  0.9305556  0.8599719
##   0.55   0.0069507607  0.8944444  0.7880206
##   0.55   0.0695076065  0.9094444  0.8180206
##   1.00   0.0006950761  0.9208333  0.8399719
##   1.00   0.0069507607  0.9319444  0.8630206
##   1.00   0.0695076065  0.8969444  0.7930206
##
## Accuracy was used to select the optimal model using the largest value.
## The final values used for the model were alpha = 1 and lambda = 0.006950761.
```

```
## 7 x 1 sparse Matrix of class "dgCMatrix"
##                                     s1
## (Intercept)                -4.746593788
## GHG_per_capita_emissions    0.240806557
## Energy_per_capita           0.066914612
## Renewable_energy_percent   -0.012684575
## Agricultural_area_percent  -0.005305846
## Forest_area_percent         0.011387377
## Protected_area_percent      0.022893270
```

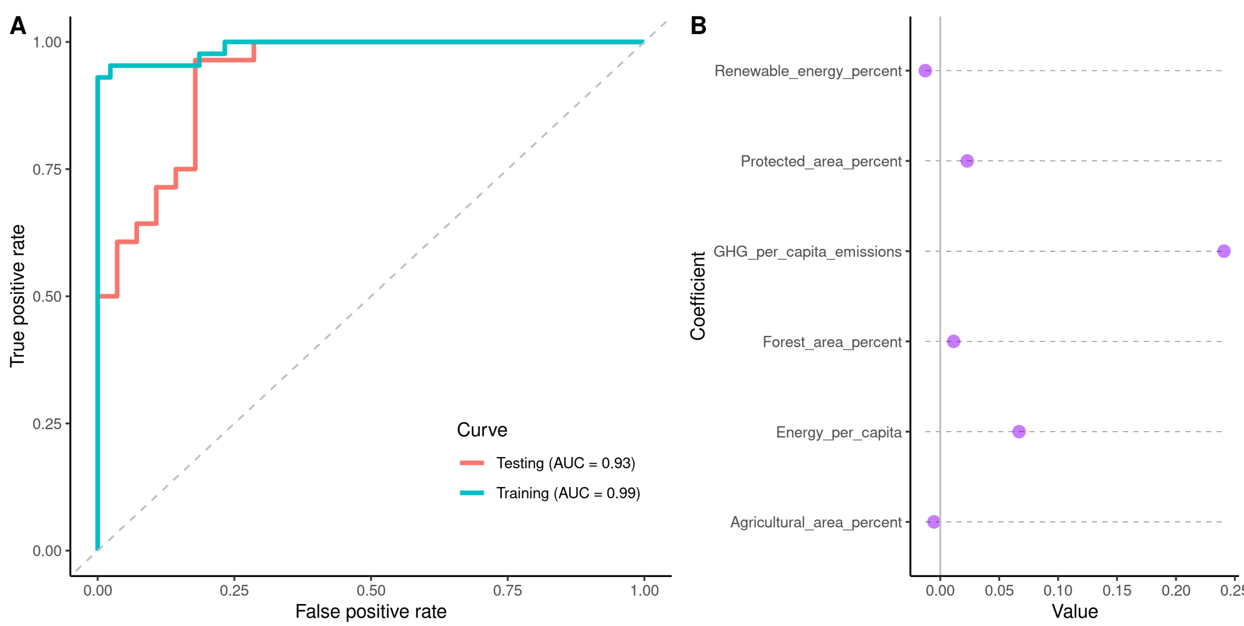

Figure 7: GDP predictor

4. Conclusions

5. Further exploration