1. Abstract

For this project, I intend to investigate negative impacts of economic growth on forest areas at the national level. Using countries’ GDP and Forest Area data, I am able to look at trends of economic growth and forest area changes. To further assist my study, I group countries by their income levels. And I calculate a Deforestation-GDP statistic and use ggplot to identify best and worst countries in each group. Hopefully, worst countries can draw lessons from best countries to protect their forests while improving their economies. Further more, to query about which continents have biggest deforestation problems, I create a world heat map. We can clearly see Latin America, Africa, and South Asia fare worst in term of deforestation. Finally, I run a linear regression to project the worst country’s forest areas for the next two years using World Bank’s projected GDP data. It intends to answer the question what will happen if we do not reverse the trend for the nation.

1. Introduction

This analysis focuses on the relationship between economic growth and deforestation. Economic growth is essential to improve people’s standard of living. On the other hand, forests play a very important role in protecting our environment. They absorb greenhouse gases and provide habitat for for millions of species. Some countries are sacrificing their forests in order to grow their economies. This project tries to identify the best and worst countries at different income levels in term of forest area changes with respect to GDP growth. The worst countries can draw some lessons from the best countries to grow their economies without having grieving impacts on their forests. In addition, this project tried to identify the worst continents in term of deforestation. Furthermore, the project also tries to project the forest area of the worst country in the world in term of deforestation for next two years given its projected GDP data.

1. Background

In order to perform this analysis to answer guiding questions, I explored the World Bank’s datasets. I was looking at Environment and Economy & Growth sections. Missing-value is a big issue for many datasets. Finally I chose Forest area (sq km) and GDP(current US$), which are two relatively clean datasets. In order to group countries by income levels, I also download the countries’ meta data, which do not have API available. So I downloaded a csv instead. I wanted to show negative impacts of deforestation on species. But the Mammal species(threatened) dataset has lots of missing values to be useful. So I gave up on the idea.

1. Method

3.1 Workflow

Download a csv of projected GDP

Simple plots to explore trends of GDP and Forest area changes

Calculate a statistic to measure forest area changes with respect to GDP growth and identify best 3 and worst 3 countries at every income level

Create plots to look at trends of GDP and forest area changes for those countries

Download Forest and GDP, and meta datasets

Read data from MYSQL and join meta data and main datasets

Upload datasets to MYSQL

Clean datasets and handle missing values

Create a deforestation world map

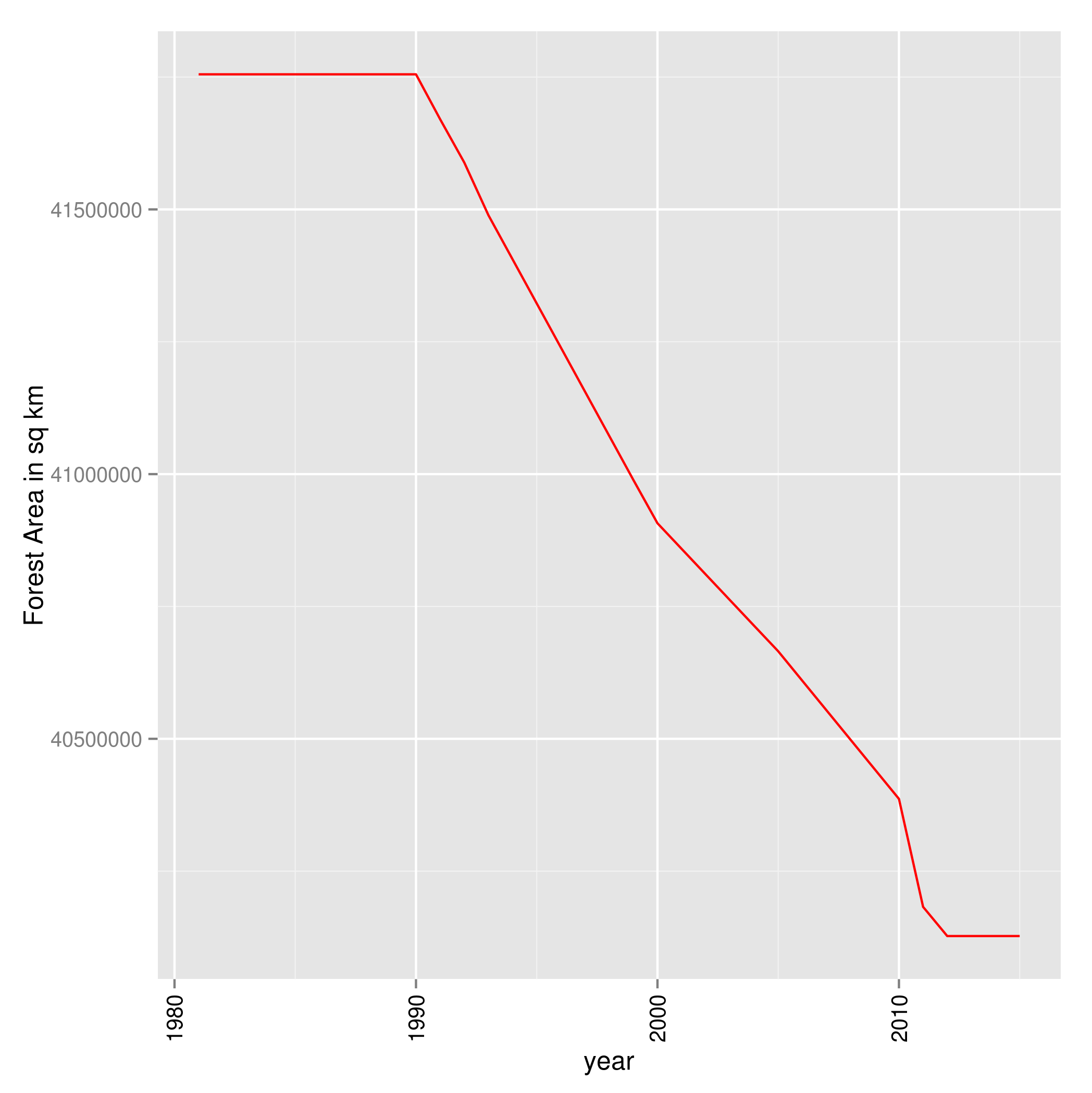
Run linear regression to project its forest area

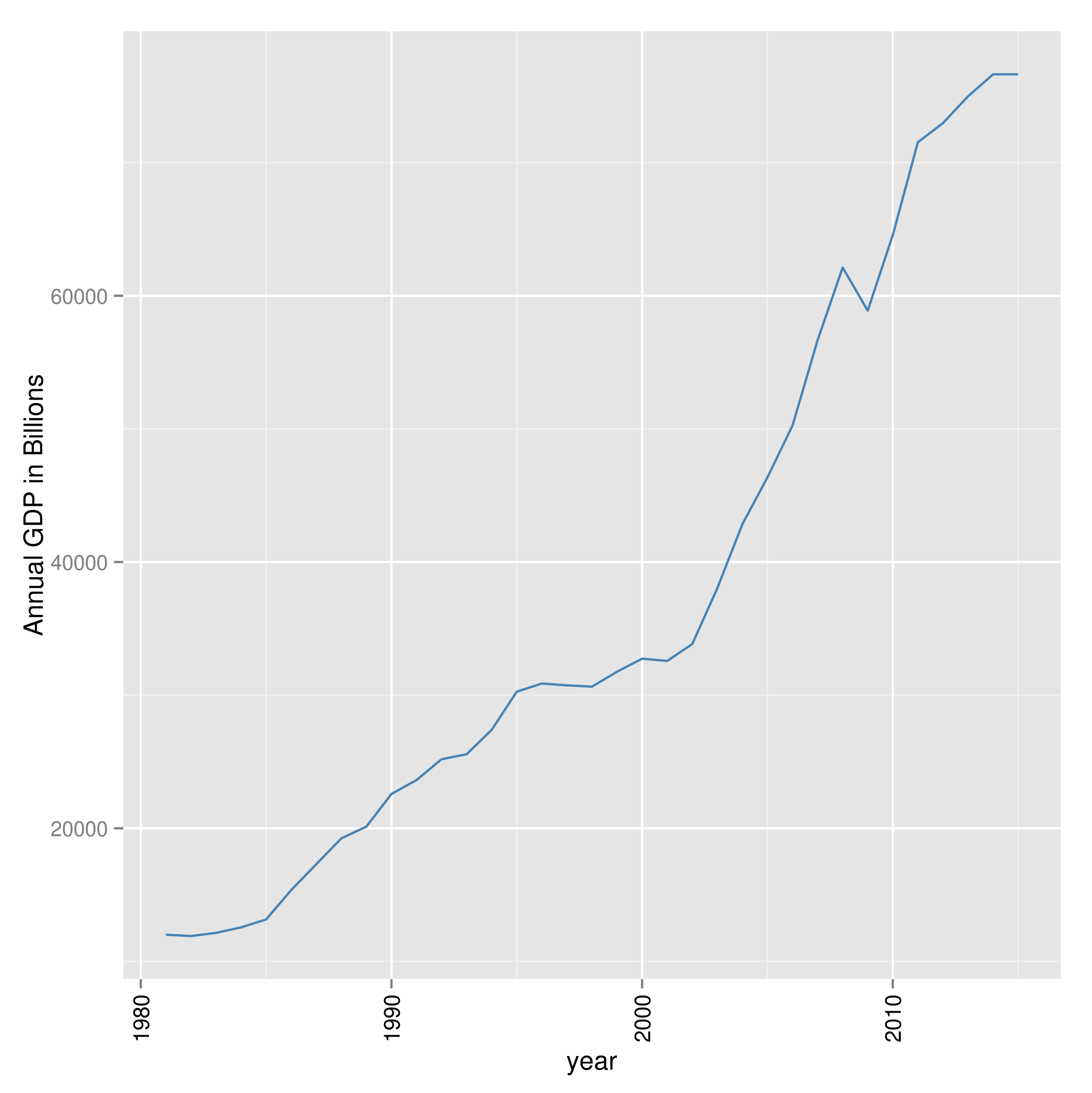
Identify the worst country in term of deforestation

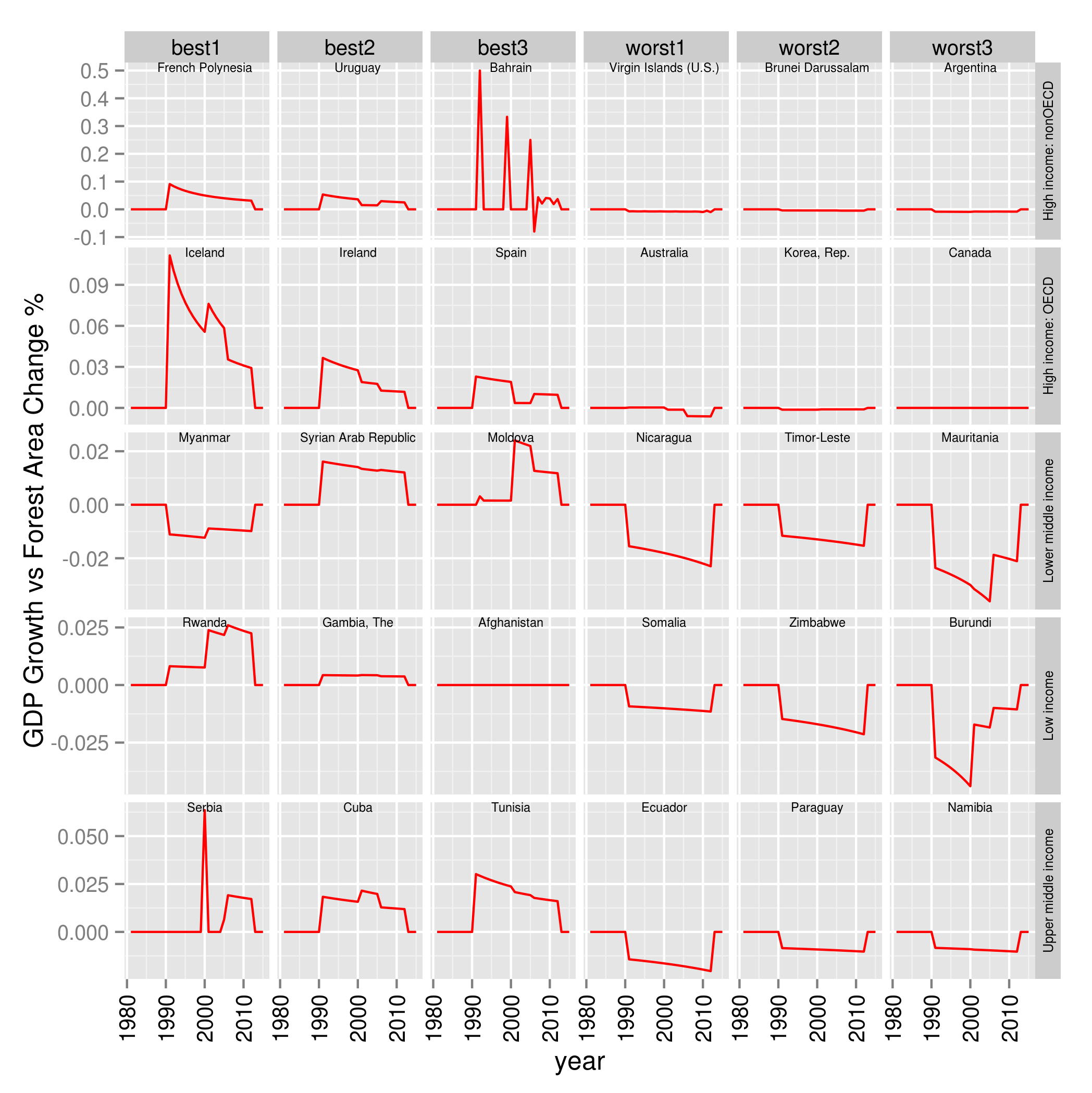
* 1. Project structure

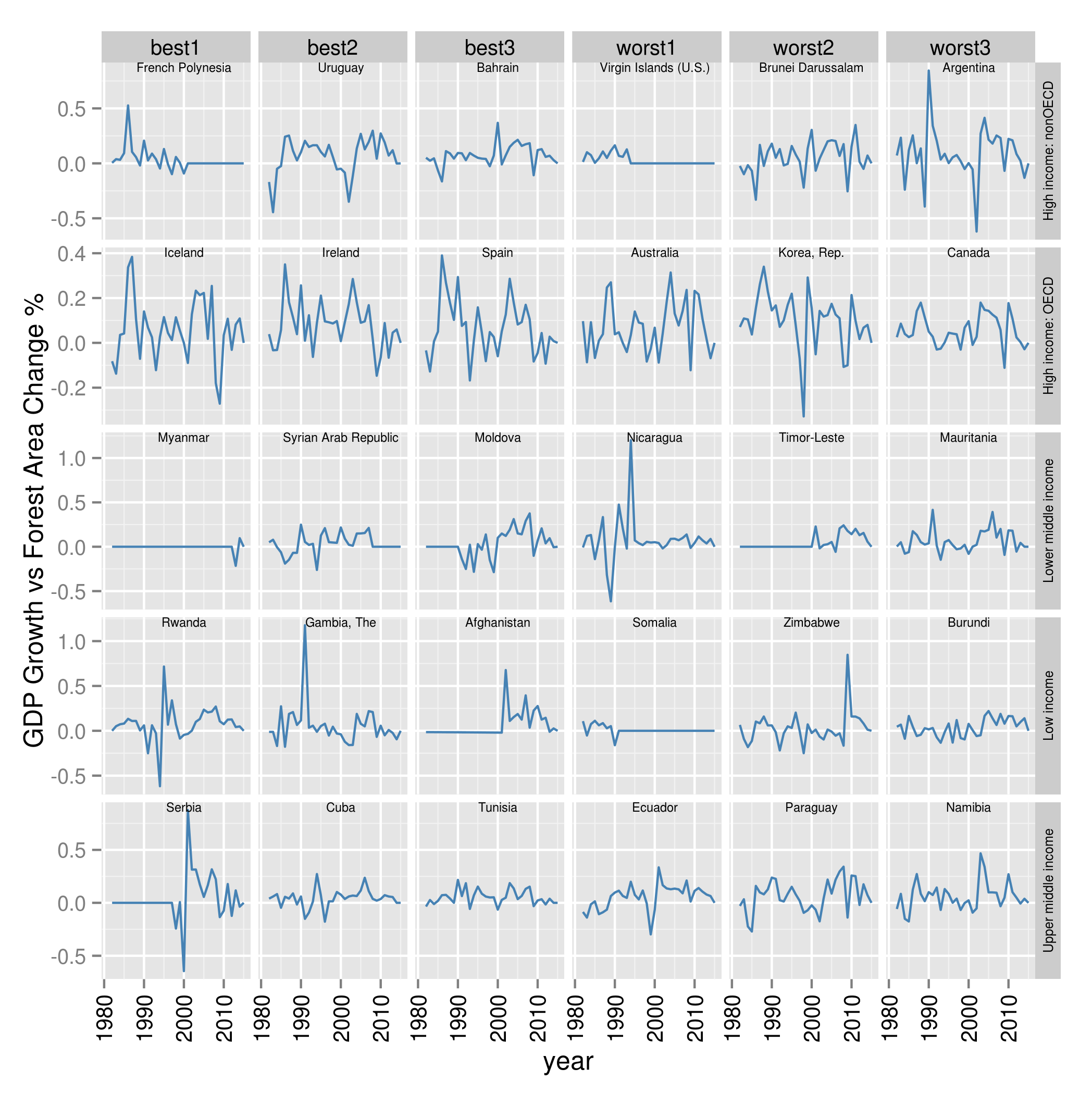
The project analyzes the relationship between World Bank’s GDP and Forest Area datasets. In order to group countries by income level, additional country meta data are downloaded. The project also develops a linear regression model to project the worst country, in term of deforestation, ‘s forest area for next two years using its projected GDP path.

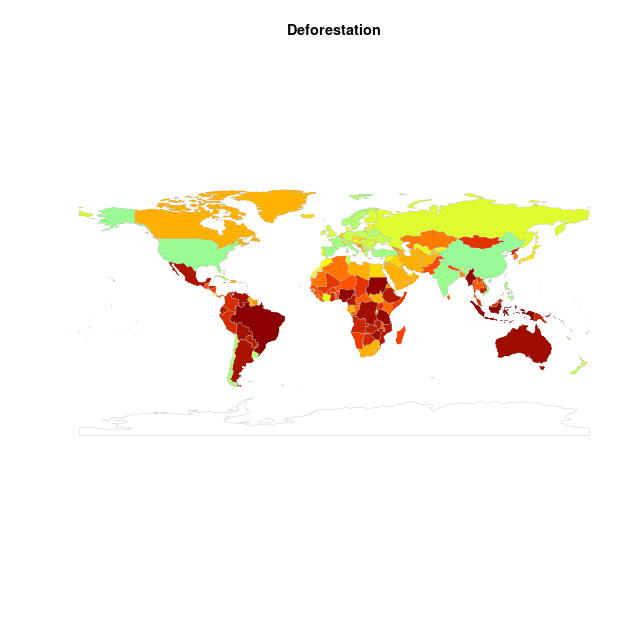
* 1. Figures and Tables











1. Discussion

The project groups countries by their income levels and identify the best and the worst countries in term of forest area changes with respect to GDP growth. The worst countries can draw lessons from the best countries to protect their forests while promoting economic growth. The projects demonstrate the ability to utilize both Python and R, the two most popular data science tools, to analyze data and built models.

6 Conclusion

As we can see from plots, for the low-income countries, the African countries occupy the worst spots. But those countries can learn from the experience of Rwanda which is the best country and also an African country to protect their forests while improving standard of living of their citizens. That demonstrate the main focus of this project. Forest protection and economic growth are both important for a nation. They do not have to have an inverse relationship. But if the worst countries do not learn from best countries, there will be serious consequence in term of damaging their environments. Brazil’s projected forest area for the next two year shows impacts of its economic growth on its forests. Many species may become extinct due to inaction.