1 NO POVERTY

2 ZERO HUNGER

3 GOOD HEALTH AND WELL-BEING

4 QUALITY EDUCATION

5 GENDER EQUALITY

6 CLEAN WATER AND SANITATION

7 AFFORDABLE AND CLEAN ENERGY

8 DECENT WORK AND ECONOMIC GROWTH
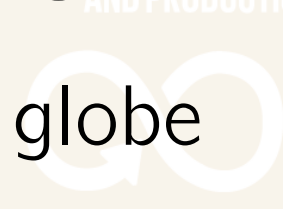
9 INDUSTRY, INNOVATION AND INFRASTRUCTURE

10 REDUCED INEQUALITIES

11 SUSTAINABLE CITIES AND COMMUNITIES

12 RESPONSIBLE CONSUMPTION AND PRODUCTION

The Global Goal #4: QUALITY EDUCATION
A toolkit for researchers and decisionmakers around the globe

13 CLIMATE ACTION

14 LIFE BELOW WATER

15 LIFE ON LAND

16 PEACE, JUSTICE AND STRONG INSTITUTIONS

17 PARTNERSHIPS FOR THE GOALS

THE GLOBAL GOALS
For Sustainable Development

Quality Education a Catalyst for a Better World

- There is probably no greater way to improve the living conditions and well-being of people in any country than quality education.

- Finland is a good global example of a poor developing nation which through quality education rose to its current global position

- That's why it's crucially important to provide tools for researchers and policy-makers to help with making the most impactful decisions to promote Quality Education
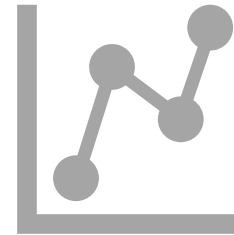
For Whom?

For researchers with only basic Python skills

use the prepared data and ready-made models to conduct your research

For Data Scientists

use our work as a starting point, apply our models for new data and develop the models further
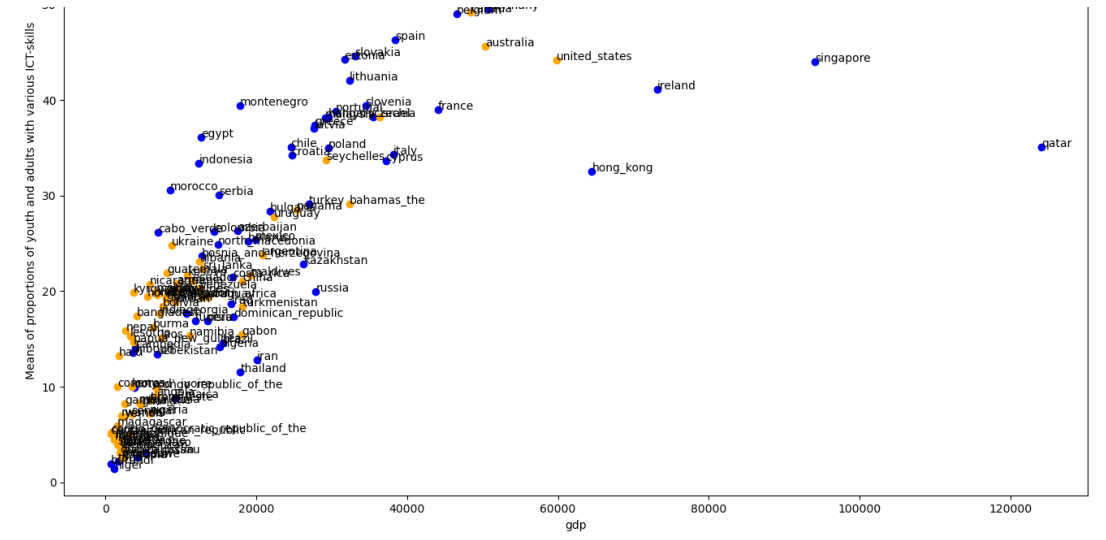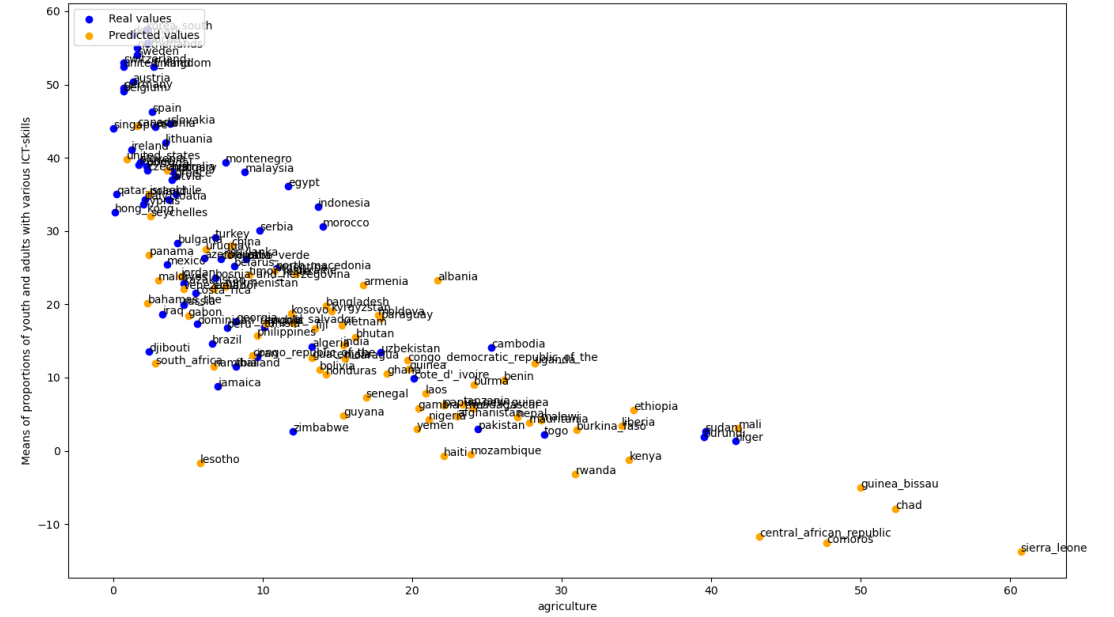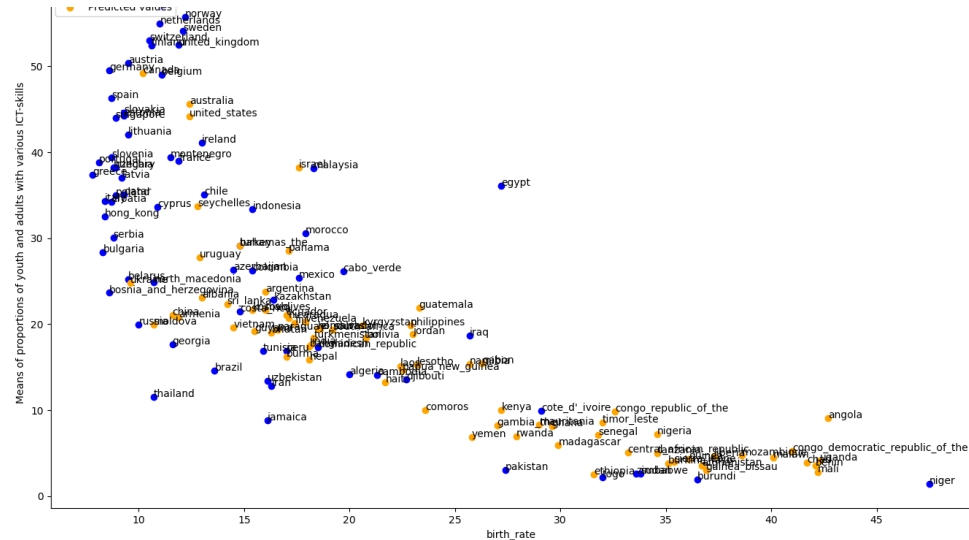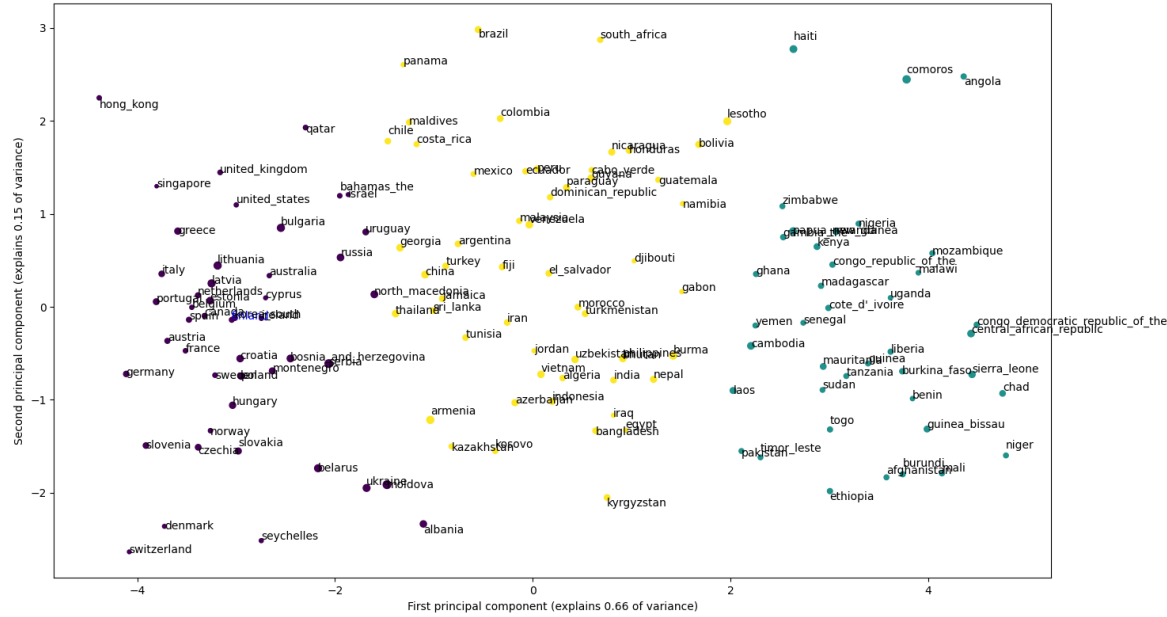
The Package

- Full human readable listing of all SDG4 indicators available from the UN SDG API

- Preprocessed CIA factbook background data from countries across the globe

- A model to use a country's CIA data to predict its missing values in the UN SDG data.

- The user can freely choose both the explanatory variables from the CIA data and the SDG data to use as the target of either Linear Regression or Random Forest Regression.

- Tools for exploratory data analysis are provided, including Principal Component Analysis and k-Means Clustering with user defined number of dimensions and clusters

- Further options to narrow down the list of countries to analyze

- We have used cross-validation to get an idea how accurately the predictive model will perform in practice, typically the Random Forest Regressor outperforms Linear Regression.

- Scatterplots of the real and predicted SDG values as y with each of the chosen explanatory variable as x for the chosen variables

- Clear visualizations make it easy to detect outlier data points. These outliers can be excluded when building the regression model by user's discretion.

- The package will be made available for everyone to use and to contribute on GitHub

Teasers

Value Proposal

- We did the hard work for you and went through the UN SDG API and augmented the data with the CIA Factbook data

- We listed the data attributes available and modified the data to a easily processable format

- You can use our model to see how your country performs in comparison to your bench-mark countries using different general indicators from the CIA factbook data to pinpoint actions with the most impact on key education indicators

- You can conduct your own research using the data and the models offered and even build further on the resources given

Known Issues

- In the UN SDG data there is often a significant number of countries with no data available for the different indicators

- There are some countries that due to political reasons are not statistically treated in the same way in the CIA factbook and the UN SDG data
  – most importantly Taiwan and Palestine

- We did not have resources to fine tune the design of the visualizations at this point as they are made mainly for research purposes only

Thank you for your interest!

Matti Kiviluoto, Juhani Kivimäki, Sebastian Lampinen