## 1 Problem analysis

### 1 Introduction to poverty

The world has committed itself to intensive targets for reducing poverty by 2030. The goal to reduce poverty and its related side effects are specified in Sustainable Development Goal (SDG) number one of the total seventeen SDG's set up in the 2030 Agenda for Sustainable Development. All seventeen goals and quantitative targets are meant to stimulate action over the next fifteen years in areas of critical importance for humanity and the planet [2].

### 1.1 UN Targets

The negative effects of poverty are enormous. Poverty and related effects slow down development within whole countries. Certain groups are disproportionately represented and face additional constraints, such as limited access to resources and capabilities in escaping poverty. These include women, persons with disabilities, children, and indigenous peoples.

* Target 1: By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than $1.25 a day.
* Target 2: By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions.
* Target 3: Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable.
* Target 4: By 2030, ensure that all men and women, in particular the poor and the vulnerable, have equal rights to economic resources.
* Target 5: By 2030, build the resilience of the poor and those in vulnerable situations.

### 1.2 The role of data in battling poverty

To mobilise resources to overcome inequalities between developed and developing countries data is needed. Data is widely used within decision-making and forms the raw material for accountability of policies. Today, in the private sector, analysis of big data is common for example with consumer profiling and personalised services. Similar techniques are being adopted to gain real-time insights into people’s wellbeing and to target aid interventions to vulnerable groups [2]. Sources of data, if applied responsibly, can enable more evidence-based decision-making, hereby contributing to making progress on the Sustainable Development Goals (SDGs) [1].

This is where a problem lies for analysing the poverty problem. Within the World Development Indicators, a lot of data is missing on various indicators on poverty. According to the IEG, reaching the UN targets faces many obstacles, among one of them is the lack of regularly high-quality data. Of the 139 developing countries, 22 have no data for the past three decades, some 30 countries have not had a survey in the past decade, and 20 countries only have had one survey since 2000 [2].

What is even more striking, is that one third of these countries are in Sub-Saharan Africa, where poverty is most frequent. Furthermore, from the countries that are reporting new surveys, eighty percent are from middle-income economies or higher. Also, additions to the survey database have been declining in recent years and the number of surveys within five years of a given date are declining rapidly [2].

One of the problems of building and maintaining a usable poverty database are financial, it would cost around 30 million/year to continuously perform surveys in 77 lower income countries. But these costs are not the main problem. There is also an enormous lack of political will within the government to proper store poverty data. Political will covers a multitude of obstacles from a refusal to collect data to unwillingness to release data for analysis. Also, many surveys on poverty known to World Bank staff are not available to other analysts or even to the citizens of the country [3].

### 1.3 Goal of this report

The success of the Sustainable Development Goals will be judged on what the data is telling us about their current state. For SDG 1, this will be mainly done in achieving the headline indicator of extreme poverty reduction, which is the share the population in a country living under the extreme poverty line of 1.90 a day. Without better poverty data, collected more frequently, and open to all, we will never know if that goal can be achieved. This can be done by collecting data, but it would take a lot of time to gather all this data and process it. Therefore, the main goal of this report is constructing a predictive model on basis of indicators that are related to poverty. We also try to get it an insight in the current state of poverty around the world the determine if our model is predicting accurate enough.

### 1.4 Research questions

Our main research question twofold, the main part is about constructing a model that predicts poverty for various countries around the world. Besides that, we try to explore the current state of poverty and the distribution of wealth around the world. Our main research question is:

*How can we predict the prevalence of extreme poverty within countries based on a MCMC model and what is the development of poverty around the world since 1990?*

**To answer this research, several sub questions are defined:**

* Question 1: How did poverty in both absolute and multi-level develop from 1990 – 2015 within regions and individual countries?
* Question 2: How is wealth distributed within countries with the highest poverty headcount?
* Question 3: Which indicators resemble good estimators for our poverty model?
* Question 4: What are the parameters of our MCMC model to estimate the prevalence of poverty for countries with missing data?
* Question 5: What is the accuracy of our predictive model?

### 1.5 Methodology

Within this report, we try to gain more insight behind the data of poverty around the world. First, we will clean and analyse the data by selecting the relevant indicators. By means of critical analysis and interpretation of figures and numbers, we try to find relations in the data to obtain useful indicators for our model. Also, relevant primary research on the topic of poverty will be included to compare our primary research findings.

R-studio is used as a tool to perform the data analysis within this research. R is widely used as a programming language for the visualisation of data. It contains additional packages that are very powerful for data analysis and visualisation purposes. Within our computations in R, several additional packages have been used. For example, Plotly has been used to make our detailed graphics. In Plotly visualisations are displayed in an efficient way and many possibilities to visualize the data are present. We also use several scripts provided within this course to estimate our model.

Resources

[1] <https://www.globalgoals.org/1-no-poverty>

[2] <https://sdgcompass.org/sdgs/sdg-1>

[3] <https://ourworldindata.org/extreme-poverty>

[4] <http://pubdocs.worldbank.org/en/109701443800596288/PRN03Oct2015TwinGoals.pdf>

[5] <https://opendatawatch.com/blog/ending-poverty-needs-better-data/>

[6] <http://ieg.worldbankgroup.org/sites/default/files/Data/reports/poverty_focus_cp_1.pdf>

[7] 7<https://qz.com/africa/1428639/world-banks-measure-of-poverty-is-flawed/>

<http://www.un.org/en/sections/issues-depth/big-data-sustainable-development/index.html>