

Data Oriented Programming Paradigms, Exercise 2, Group 26

Reema George Dass 12144026, Simon Glatzer 12019265,
Harald Hubinger 09426524, Valentin Schweitzer 51829840

December 2022

1 Questions to answer and how to answer them

1. In all countries across the world today, how many children finish primary school?
We'd rephrase to percentages per country and would predict 2022 unless it's meant to find a timepoint with complete data. We'll treat "primary education" as defined in [1].
2. How does this differ for low-, medium- and high-income countries?
We plan to use GNI per capita, as defined in [2], as a classifier according to the World Bank.
3. Is it possible to identify objectively reasons for them not finishing school?
We plan to look at a subset of these factors: Resources (public/private financing, teacher/student ratio), gender equality, urbanisation, democracy index, development status (life expectancy, health, overall education rates...)
4. How well can predictions be made for the figures 5 and 10 years in the future?
Predictions could be made from historical data of the target variable in other countries. If this is not feasible, predictions of explanatory variables from questions 2 and 3 could be of use.

2 Data sets

Primarily, we are going to use data provided by the OurWorldInData page analysis on primary education [3] and the financial background [4], which is based on data accumulated by the Unesco Insitute for Statistics [5] via the World Bank. Additionally, we are going to supplement this data using specific data sets taken from the World Bank data bank on education statistics [6] itself. For question 3 we will also use data taken from the United Nations regarding the global gender inequality [7].

3 Project Plan

1. Nov 30 – Dec 12: project setup, brainstorming & initial analysis and data collection
2. Dec 13 – 23: define target structure, explore, preprocess & align data, collect ideas for modeling.
3. Dec 27 – Jan 06: model, fit & validate. If necessary, iterate (back).
4. Jan 09 – Jan 13: prepare final presentation & management summary, prepare submission
5. Jan 16 – deadline: buffer time

Work to be organized in individual tasks / groups of two; min 1 meeting/week for all (online) with task assignments; central task list and clearing point for all data; project manager (Harald Hubinger) takes care of time structure and task management. We will start with 2. mostly in parallel according to common guidelines.

References

- [1] *International Standard Classification of Education (ISCED) 2011*. UNESCO Institute for Statistics, 2012. DOI: 10.15220/978-92-9189-123-8-en. [Online]. Available: <http://dx.doi.org/10.15220/978-92-9189-123-8-en>.
- [2] W. Bank, *The world bank atlas method - detailed methodology*. [Online]. Available: <https://datahelpdesk.worldbank.org/knowledgebase/articles/378832-what-is-the-world-bank-atlas-method>.
- [3] M. Roser and E. Ortiz-Ospina. “Primary and secondary education.” (2013), [Online]. Available: <https://ourworldindata.org/primary-and-secondary-education>.
- [4] M. Roser and E. Ortiz-Ospina. “Financing education.” (2016), [Online]. Available: <https://ourworldindata.org/financing-education>.
- [5] “Unesco institute for statistics.” (2022), [Online]. Available: <https://uis.unesco.org/> (visited on 12/07/2022).
- [6] “World bank databank for education statistics.” (2022), [Online]. Available: <https://databank.worldbank.org/reports.aspx?source=education-statistics-~-all-indicators> (visited on 12/07/2022).
- [7] “United nations development programme - gender inequality index.” (2022), [Online]. Available: <https://hdr.undp.org/data-center/thematic-composite-indices/gender-inequality-index#/indicies/GII> (visited on 12/07/2022).