University rankings, economy, and human development

Andrea Valle & Adrian Imfeld

September 3, 2023

Contents

1 Introduction

In an increasingly interconnected and competitive global landscape, education has become a cornerstone of individual opportunity and national success. Universities serve as incubators for talent, research, and innovation, often becoming significant drivers of economic and human development. This data science project aims to investigate the relationships between university rankings, economic development, and human development by leveraging a multi-dimensional combined dataset spanning multiple years and countries.

University rankings, produced annually by several organizations, consider a range of metrics including research output, faculty qualifications, and global reputation. Economic development, frequently measured in Gross Domestic Product (GDP) per capita, encapsulates a nation's economic health and standard of living. Human Development Index (HDI), on the other hand, offers a composite score of life expectancy, education, gender equality, and income indices to provide a fuller understanding of human well-being.

The objective of this project is to unearth the hidden patterns and correlations among these variables over time, across a broad cross-section of countries. This project aims to reveal whether higher-ranked universities contribute directly or indirectly to economic and human development. The results of this research could have implications for educational policy, international development, and global talent migration, potentially serving as a guide to leverage education as a tool for holistic national advancement.

However, it's crucial to acknowledge the limitations of this small research project. One of the most significant limitations is that the project may identify correlations but cannot infer causation. In other words, even if the data suggests that countries with more and higher-ranked universities tend to have higher GDP per capita and better HDI scores, this does not prove that the universities themselves are the cause of these improvements. On the contrary, richer countries may simply be able to invest more into universities and education. Numerous confounding variables, such as government policies, cultural factors, and natural resources, could also contribute to these outcomes. The synergistic effects between these variables are complex and may involve lagging indicators or feedback loops that are beyond the scope of this project. Therefore, while the correlations unearthed may serve as a starting point for more targeted research, they should not be used as a basis for causal claims without further investigation.

2 Data preparation

2.1 Datasets

2.1.1 Academic ranking of world universities (Shanghai ranking)

According to their website, the Academic Ranking of World Universities (ARWU) is recognized as the precursor of global university rankings and the most trustworthy one. ARWU presents the world's top 1000 research universities annually based on transparent methodology and objective third-party data.

- Name: Academic Ranking of World Universities (ARWU) - Shanghai Rankings from 2005-2018
- Website: https://www.shanghairanking.com/
- Dataset: https://www.kaggle.com/datasets/joebeachcapital/shanghai-world-university-ranking

The dataset contains a row per university and year, assigning a rank to each university based on various indicators:

- Quality of Education, with Alumni and Award indicators (10% and 20% of the final mark)
- Quality of Faculty, with HiCi and N&S indicators (20% and 20% of the final mark)
- Research Output, with PUB indicator (20% of the final mark)
- Per Capita Performance, with PCP indicator (10% of the final mark)

More on the methodology and definitions of the indicators can be found on their methodology website. The structure of the dataset looks like this (indicators and ISO country codes are ommitted):

World.rank	Rank	Year	University	Country
201-302	201	2009	University of Bochum	Germany
303-401	303	2009	The University of Texas Medical Branch at Galveston	United States
151-200	151	2010	Florida State University	United States
201-300	201	2010	Lancaster University	United Kingdom
101-150	101	2010	University of Groningen	Netherlands
301-400	301	2010	University of Southern Denmark	Denmark
16	16	2010	University of Washington	United States
151-200	151	2010	University of Hamburg	Germany
301-400	301	2010	Eotvos Lorand University	Hungary
71	71	2010	Ecole Normale Superieure - Paris	France

2.1.2 Economic dataset

todo

2.1.3 Human Development Index (HDI)

The Human Development Index (HDI) was created to emphasize that people and their capabilities should be the ultimate criteria for assessing the development of a country, not economic growth alone. The HDI is a summary measure of average achievement in key dimensions of human development: a long and healthy life, being knowledgeable and having a decent standard of living. The health dimension is assessed by life expectancy at birth, the education dimension is measured by mean of years of schooling for adults aged 25 years and more and expected years of schooling for children of school entering age. The standard of living dimension is measured by gross national income per capita.

- Name: Human Development Index (HDI)
- Website: https://hdr.undp.org/data-center/human-development-index
- Dataset: https://hdr.undp.org/sites/default/files/2021-22_HDR/HDR21-22_Composite_indices_complete_time_series.csv
- Metadata: https://hdr.undp.org/sites/default/files/2021-22_HDR/HDR21-22_Composite_indices_metadata.xlsx

The dataset is in wide format, containing a row for each of the 206 countries and a column for each indicator-year combination. The time span ranges from 1990-2021 for most indicators, resulting in 1008 columns total. The structure looks like this (only a selected few indicator-year columns are shown):

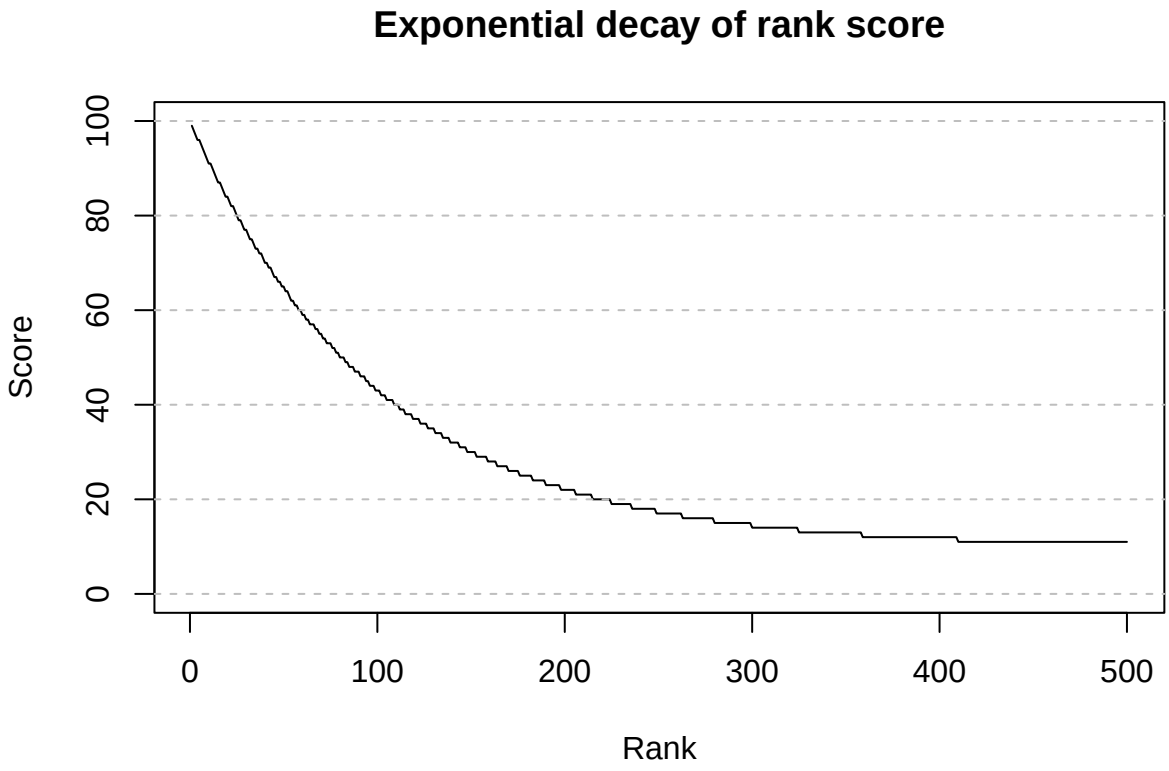
iso3	country	hdicode	hdi_1990	hdi_2021	le_1990	le_2021
AFG	Afghanistan	Low	0.273	0.478	45.9672	61.9824
AGO	Angola	Medium	NA	0.586	41.8933	61.6434
ALB	Albania	High	0.647	0.796	73.1439	76.4626
AND	Andorra	Very High	NA	0.858	78.4063	80.3684
ARE	United Arab Emirates	Very High	0.728	0.911	71.9004	78.7104
ARG	Argentina	Very High	0.723	0.842	71.7837	75.3899
ARM	Armenia	High	0.656	0.759	68.8209	72.0431
ATG	Antigua and Barbuda	High	NA	0.788	73.4922	78.4968

iso3	country	hdicode	hdi_1990	hdi_2021	le_1990	le_2021
AUS	Australia	Very High	0.865	0.951	77.0443	84.5265
AUT	Austria	Very High	0.825	0.916	75.7396	81.5797

2.2 Preprocessing

2.2.1 Univerity rank aggregation

In order to correlate the university rankings with economic and human development over time, we aggregated the university rankings per country and year. First, we assigned a score per university based on its rank according to an exponential decay model. The idea is to score top universities a lot higher and introduce an exponential decay in score for lower ranks. This approach was inspired by webometrics, which orders countries by the number of universities in top categories first.



Once every university was assigned a score per year based on its rank, the scores were summed up and the universities counted per country and year. The aggregation resulted in the following structure, ready to be joined with the economic and HDI datasets:

ISO3.CODE	Country	Year	Points	Count
USA	United States	2005	6448	168
USA	United States	2006	6366	167
USA	United States	2007	6373	166
USA	United States	2008	6289	159
USA	United States	2009	6186	152

2.2.2 Economic dataset

todo

2.2.3 HDI dataset

From the Human Development Index (HDI) dataset, the timelines of the following indicators of interest were picked:

- Human Development Index – a summary of a long and healthy life, access to knowledge and a decent standard of living.
- Life expectancy at birth
- Expected years of schooling
- Mean years of schooling
- Gross National Income (GNI) per capita
- Gender Development Index (GDI) – measures gender inequalities in achievement in three basic dimensions of human development:
  - health, measured by female and male life expectancy at birth
  - education, measured by female and male expected years of schooling for children and female and male mean years of schooling for adults ages 25 years and older
  - command over economic resources, measured by female and male estimated earned income
- Gender Inequality Index (GII) – reflects gender-based disadvantage in three dimensions for as many countries as data of reasonable quality allow:
  - reproductive health
  - empowerment
  - labourmarket

For details on the indicators, see the technical notes.

The dataset was transformed into a suitable format, containing a row per country and year, and columns for the indicators. After joining the transformed HDI dataset with the previously prepared university ranking data, world region and continent information was joined as well as country population data from the world bank dataset. This is the structure of the final dataset (some indicators are omitted):

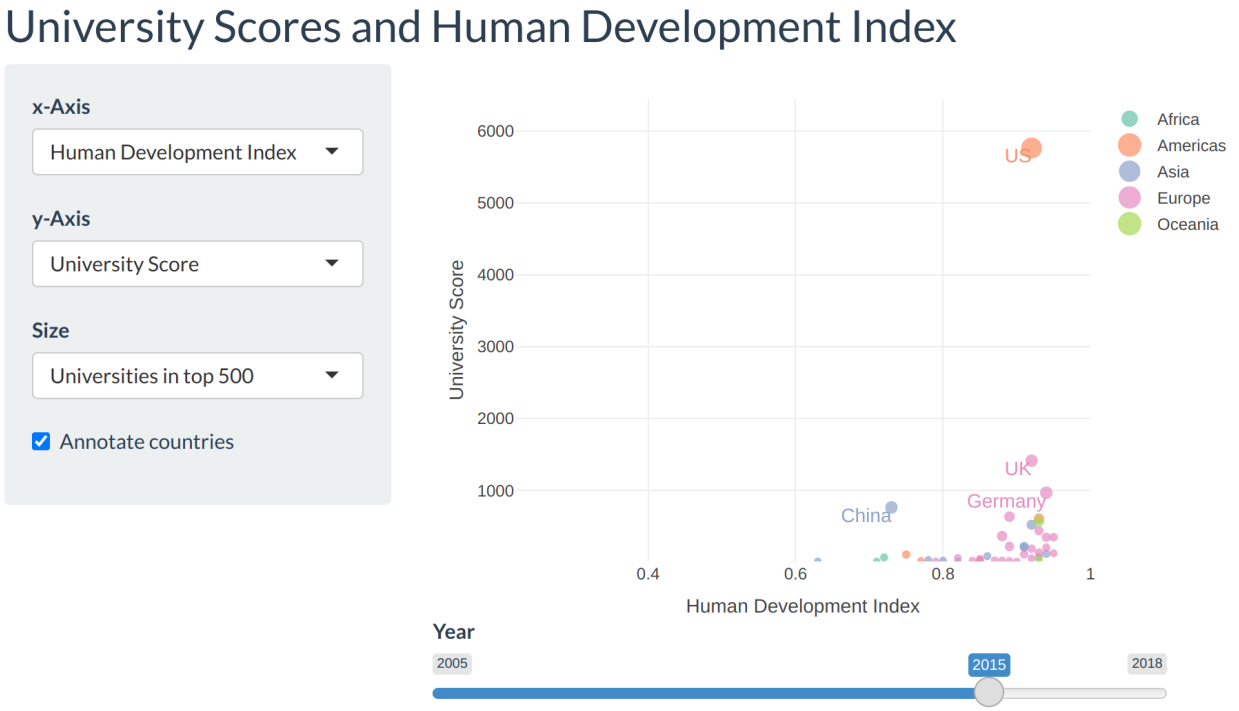
country	continent	year	population	hdi	le	eys	mys	gnipc	uni_score	uni_count
Germany	Europe	2005	82469422	0.91	79.26	16.34	13.70	44193.77	1059	40
Germany	Europe	2006	82376451	0.92	79.59	16.39	13.60	46045.61	1047	40
Germany	Europe	2007	82266372	0.92	79.74	16.44	13.68	47424.30	1049	41
Germany	Europe	2008	82110097	0.92	79.83	16.48	13.76	47413.82	1053	40
Germany	Europe	2009	81902307	0.92	79.92	16.53	13.83	46040.43	1056	40

<–{r child = '03_Economy_analysis.Rmd'} –>

3 Human development

3.1 Interactive app

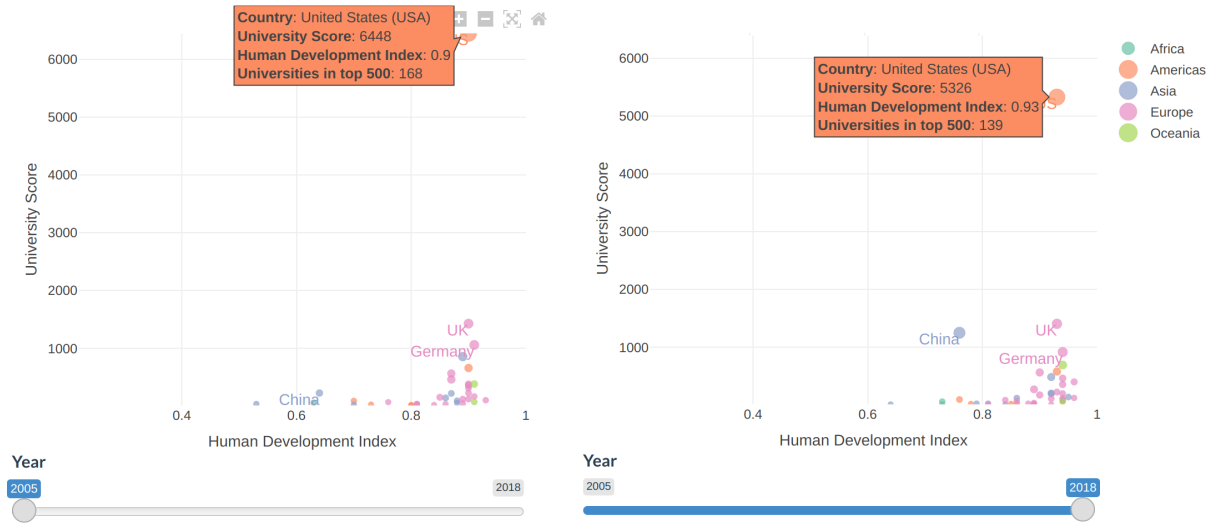
In order to explore the relationships between the university rankings and the HDI indicators, an interactive R shiny app was built and published here: https://aimfeld.shinyapps.io/uni-ranks. The source code and data is available on Github.



The user interface allows for selecting an indicator for the x-axis, y-axis, and marker size. Development over time can be observed by dragging the year slider. Finally, some countries of interest annotations can be toggled on and off. Since the university ranking data only spans the years 2005-2018, the time range is adjusted dynamically when selecting the indicators.

The app is inspired by the famous Gapminder animations, presented by Hans Rosling in a very entertaining TED talk.

3.2 University scores and human development

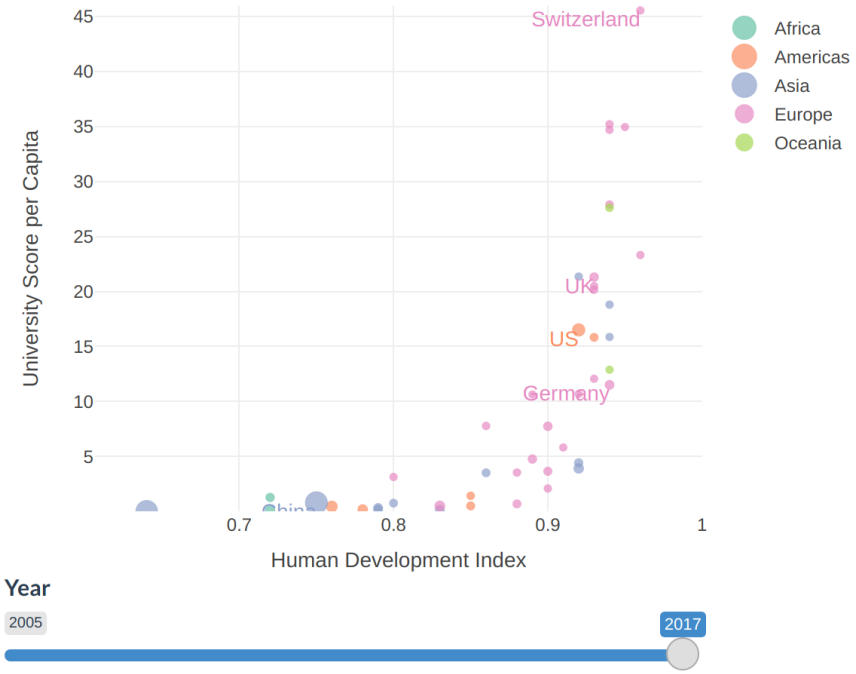


In terms of university scores, the US was dominating with 168 universities in the top 500 in 2005. Over the years, China has been catching up in number and ranking of top 500 universities as well as in human development. The decline in combined university scores of the US from 2005 to 2018 may be explained by

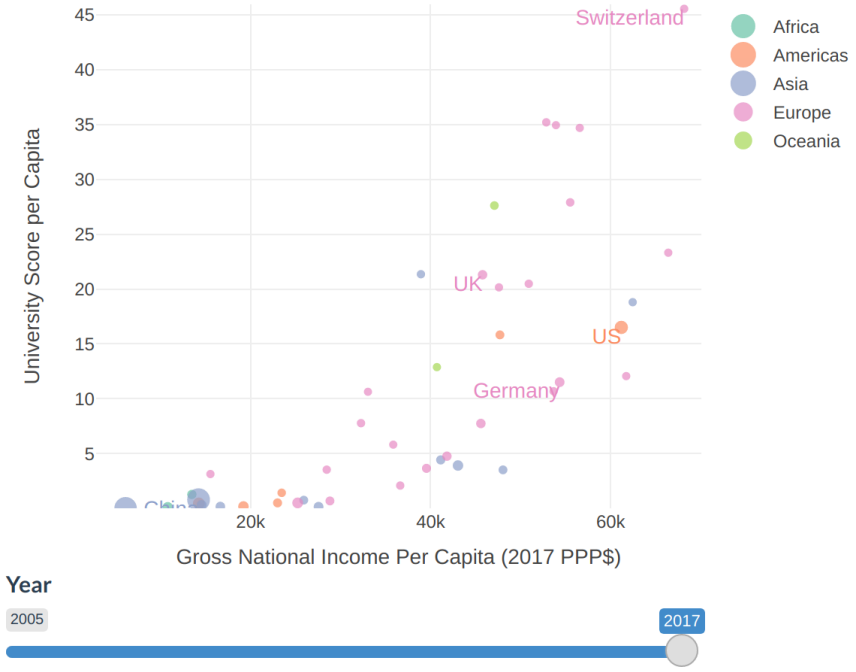
Chinese universities entering the top 500 and displacing some of the lower ranked US universities. However, US dominance remains very strong. In 2023, the top ranks are is still heavily dominated by US universities. In fact, according to the Shanghai rankings 2023, 20 of the top 30 are US universities whereas only 2 are Chinese. These are the top 10 universities in 2023:

World Rank	Institution	All	National/Regional Rank	Total Score	Alumni
1	Harvard University		1	100.0	100.0
2	Stanford University		2	74.8	45.6
3	Massachusetts Institute of Technology (MIT)		3	69.1	72.5
4	University of Cambridge		1	67.9	78.9
5	University of California, Berkeley		4	63.4	65.8
6	Princeton University		5	60.1	61.7
7	University of Oxford		2	59.5	49.9
8	Columbia University		6	55.3	58.5
9	California Institute of Technology		7	54.5	56.1
10	University of Chicago		8	53.8	56.6

The relationship between university scores and HDI is unclear. However, an exponential relationship emerges if we look at university score *per capita*, correcting for country population size. Notably, Switzerland comes out on top due to its small country size and comparatively high number of quality universities.



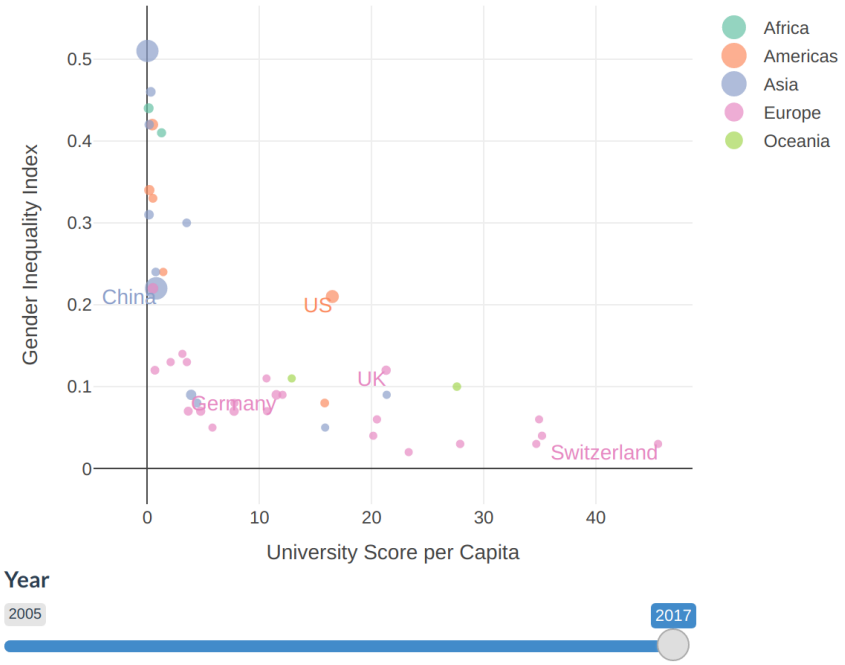
A very similar exponential pattern can be found between Gross National Income (GNI) and university score per capita. This is not surprising, since the GNI is part of the HDI.



3.3 Education and gender inequality

We might hypothesize that better education leads to less gender inequality. Indeed, the higher the university score per capita, the lower the gender inequality index. Asian countries seem to be behind in terms of

gender equality and education, while European countries are doing quite well. Although the situation has been improving in Asia over time, as can be seen interactively by using the year slider.



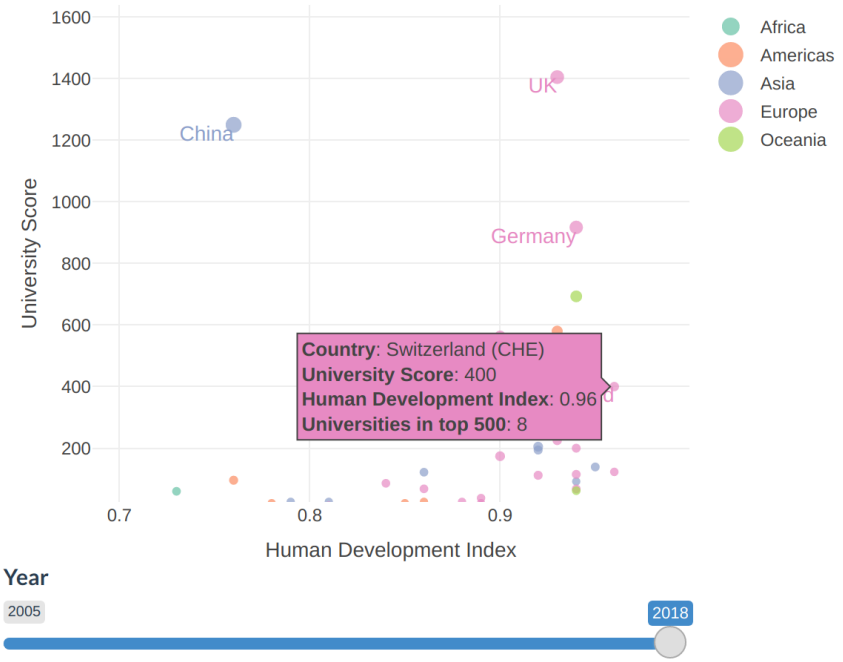
While it's tempting to infer a causal relationship between education and gender equality, both of these variables may be mediated by hidden variables such as economic development or other factors.

3.4 Switzerland

For being a very small country, Switzerland is doing very well in terms of top ranking universities and human development. There are consistently around 8 Swiss universities in the top 500. In 2023, the top Swiss universities include the following, spearheaded by the ETH Zürich with a rank of 20.

World Rank	Institution	Switzerland	National/Regional Rank	Total Score	Alumni
20	ETH Zurich		1	44.1	26.7
49	University of Geneva		2	31.8	37.7
54	Swiss Federal Institute of Technology Lausanne		3	31.3	0.0
59	University of Zurich		4	30.7	0.0
81	University of Basel		5	26.8	21.6
101-150	University of Bern		6-7		11.7
101-150	University of Lausanne		6-7		16.6

With respect to human development, Switzerland is equalled only by Norway with a HDI of 0.96.



4 Conclusion

todo