Date: Aug 26, 2023

From UN HDI Stat Analysis (2020 – 2021)

Formula used for determining the indices for the human development index (HDI)

A black text on a white background

Description automatically generated

Education’s minimum value = 0 years

Max value = 18 years

The min value was set because there are still societies that subsist without education. The max value was chosen since 18 years of formal education is enough for one to finish master’s degree in most nations.

Here’s the summary of the three dimensions:

A screenshot of a computer

Description automatically generated

I could analyze the data across different ways. Somehow, there are “flavors” in the education data:

* Raw
* Adjusted for inequality
* Gender differences
* Multifaceted

10:34

I’ll make an analysis of each one

Here’s the metadata of the hdi dataset:



According to this, there are four types:

* HDI (human development index)
* GDI (gender development index)
* IHDI (inequality-adjusted human development index)
* GII (gender inequality index)

I have to query five tables:

1. Overall
2. HDI
3. GDI
4. IHDI
5. GII

Each of them should have the (name of the country, type of country in terms of development, the overall index (e.g., HDI, GDI, etc.), and the educational index)

Kaggle usability rating of the Kaggle data set version in 9.41

10:42 after getting the tables, I’m planning to upload them first to tableau public for viz, gaining insights about the data quickly

After this, proceed with the analysis. Since the dataset is not that huge, I’ll use the pivot tables of excel.

Some of the things I think I need to look for:

* Trends in the world
* Trends in the US
* Which country/ies has the best trend?
* Which continent had it best?
* Which continent had it worst?
* How did the Philippines do?
* Explore it by decade.
* How did the developing countries compare with the developed ones?

Make a cutting-edge dashboard

10:49 I think GNI should also be included. That would be interesting since you’re going to explore the correlation of education and poverty but wouldn’t the project be so big then? Nevertheless, I believe it will help me in expanding my capabilities as a data analyst. So, yeah. Go.

To update,

Each of them should have the (name of the country, type of country in terms of development, the overall index (e.g., HDI, GDI, etc.), the educational index, and the GNP)

According to the metadata, I should get,

HDR Country Name country

Recommended citation of the dataset: “Source: UNDP (United Nations Development Programme). 2022. Human Development Report 2021/2022: Uncertain Times, Unsettled Lives: Shaping our Future in a Transforming World. New York.”

Regarding the UN Developing Regions abbreviations, here’s the list:

|  |  |  |
| --- | --- | --- |
| Region | Abbreviation | Equivalent |
| Africa | AF | All countries in Africa, except Mauritius and Seychelles. |
| Asia and the Pacific | AP | All countries in Asia and the Pacific, except Australia, Japan, New Zealand, and Singapore. |
| Eastern Europe and Central Asia | ECA | All countries in Eastern Europe and Central Asia, except Belarus, Czech Republic, Hungary, Poland, Romania, Russia, Slovakia, and Slovenia. |
| Latin America and the Caribbean | LAC | All countries in Latin America and the Caribbean, except Argentina, Chile, Cuba, and Uruguay. |
| Least developed countries | LDCs | 46 countries that are classified as having the lowest human development in the world. |
| Small island developing states | SIDS | 52 countries that are small islands with high levels of vulnerability to external shocks. |

11:17 I’m thinking if I still have to create new tables or just make queries because SQL is just so fast!

11:22 I found top 10 countries based on HDI already for years 2020 and 2021. Before I query further, I have to clean data first.

It’s worth noting that S. Korea does not have a rating. Weird

13:45 I’m back. I’ll be pasting the important findings on google sheets. I find the sheets environment more welcoming and safe than excel.

14:48 There is too much data for one analysis. I have to concentrate on education only. After doing analyses on the four types, only then should I proceed with the inclusion of the poverty field.

So I’ll be creating tables per type (HDI, GDI, IHDI, GII)

iso3, country, hdicode, region, hdi\_, hdi\_1990

19:48 I ended at I think 16:30. I’m going to continue now

2023/08/27

18:24

“Expected years of schooling is the number of years a child of school entrance age is expected to spend at school, or university, including years spent on repetition. It is the sum of the age-specific enrolment ratios for primary, secondary, post-secondary non-tertiary and tertiary education.”

– The World Bank, UNESCO Institute for Statistics (UIS) <https://databank.worldbank.org/metadataglossary/world-development-indicators/series/SE.SCH.LIFE>

Mean Years of School

Definition: “Average number of completed years of education of a country's population aged 25 years and older, excluding years spent repeating individual grades.”

Interpretation: “In general, the indicator’s value denotes the level of skills and competencies of a country's population, which could be seen as a proxy of both the quantitative and qualitative aspects of the stock of human capital. A relative high value indicates great shares of the adult population according to the highest level of education attained or completed, and reflects a performing educational system.”

Purpose: “The indicator reflects the level of schooling of the population aged 25 years and above, hence the stock and quality of human capital within a country, so as to gauge needs and establish policies for upgrading it. This indicator also informs about the performance of the education system and its accumulated impact on human capital formation.”

<https://uis.unesco.org/en/glossary-term/mean-years-schooling>

18:51 I’m making the education dimension index using the two education indices. According to the UNDP technical notes on hdr, (HUMAN DEVELOPMENT REPORT 2021/2022)

“For the education dimension, equation 1

A black text on a white background

Description automatically generated

is first applied to each of the two indicators, and then the arithmetic mean of the two resulting indices is taken. Using the arithmetic mean of the two education indi- ces allows perfect substitutability between expected years of schooling and mean years of schooling, which seems to be right given that many developing coun- tries have low school attainment among adults but are very eager to achieve universal primary and second- ary school enrolment among school-age children.“

therefore, the education dimension index will just be equal to the arithmetic mean.

20:50 list of nations of particular interest among the indexes checked:

Overall

Hdi highest - Norway

Hdi lowest - Niger

Mys highest – United States of America

Mys lowest - Burkina Faso

Eys highest - Australia

Eys lowest - Niger

Education dimension index highest – Australia

Education dimension index lowest – Niger

Gpnic highest – Liechtenstein

Gpnic lowest – Mozambique

Data as of 2021

Hdi highest - Switzerland

Hdi lowest – South Sudan

Mys highest – Germany

Mys lowest - Burkina Faso

Eys highest - Australia

Eys lowest – South Sudan

Education dimension index highest – Australia

Education dimension index lowest – Niger

Gpnic highest – Liechtenstein

Gpnic lowest – Burundi

2023/09/01 17:02 PM

Using linear regression, I computed the number of years it would take for the Philippine education dimension index to be at par with the leading nation in the world, Australia