



UNIVERSITÉ  
**BISHOP'S**  
UNIVERSITY

**Bishop's University**

**CS-503 Data Visualization**

**Dr. Mohammed Ayoub Alaoui Mhamdi**

**Final Project**

Considering the 2015 Human Development report dataset, we try to answer the following two questions:

- 1- Which Region was more developed on HDI between 2011 and 2013?
- 2- What has a bigger effect on HDI, Gross National income, or Education?

With each question, we try to answer them with one python Plotly figure. We need to clarify some points about data wrangling we did on this dataset:

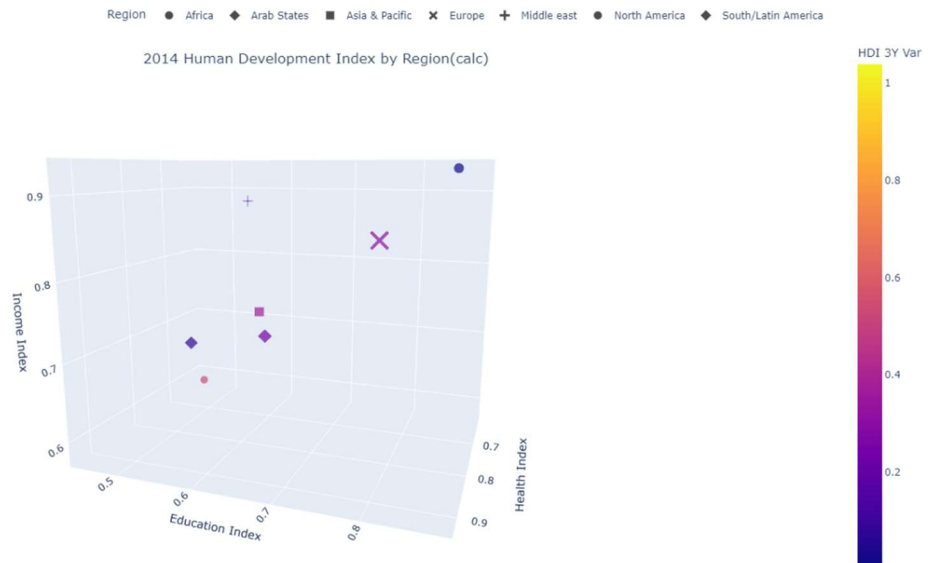
- We did some classification helped by Wikipedia to extract countries regions and assign them to the dataset
- We calculated HDI (Human development index) according to each region classification (formulas provided in [index\\_calculation.pdf](#))
- We choose `historical_index.csv` and `human_development.csv` as main datasets.

The Human Development Index (HDI) is a summary of measurement that represents countries' population well-being. It is a geometric mean of normalized indices that involved Health, decent standard living, and education. With the previous 3D graph, we tried to accomplish a summary of the 2014 Human Development Index by its 3 main measures, segregated by the world's regions and colored by its 3 previous year variations, giving the visualization a present and past look and feel.

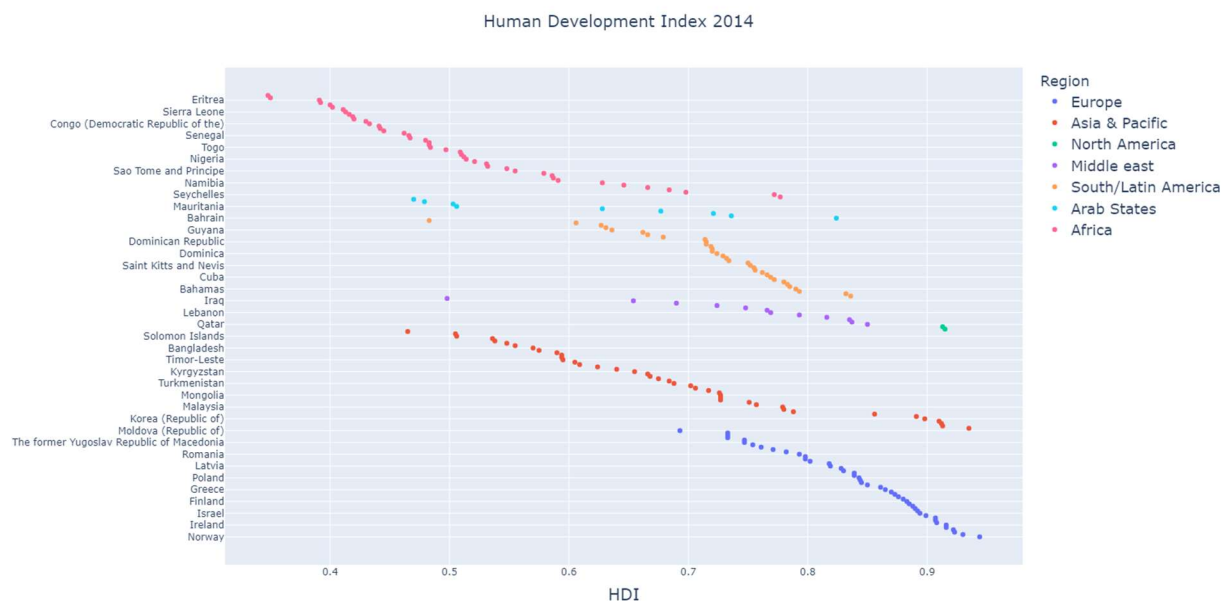
A big challenge with this dataset was to understand how each index is related to its detailed data because we did calculate the region HDI. The dataset already had on its detailed data, the overall HDI, and its region HDI calculated, so, we replicated what was done for each country and transposed the result for each region accordingly to their countries.

## Which Region was more developed on HDI between 2011 and 2013?

An important aspect to mention is that we are grouping all countries into their region ethnicity, that way we understood was better to show 3 years difference and how these regions improved between 2012 and 2014 included by color gradient, showing that the best improvement was made by Africa, Asia & Pacific, and Europe were the regions that best improved between 2011 and 2014. North America, which only includes Canada and the United States, did not have any big changes on these 3 years time-lapse but were the best HDI on this time.



In contrast with the previous chart, this is more detailed on countries and less on how the index is calculated. We are focusing here on how each country contributes to its region and how dense or spare are each country on its region which was a factor that was not presented before. Norway, Australia, and Switzerland are the three biggest HDI in our dataset.

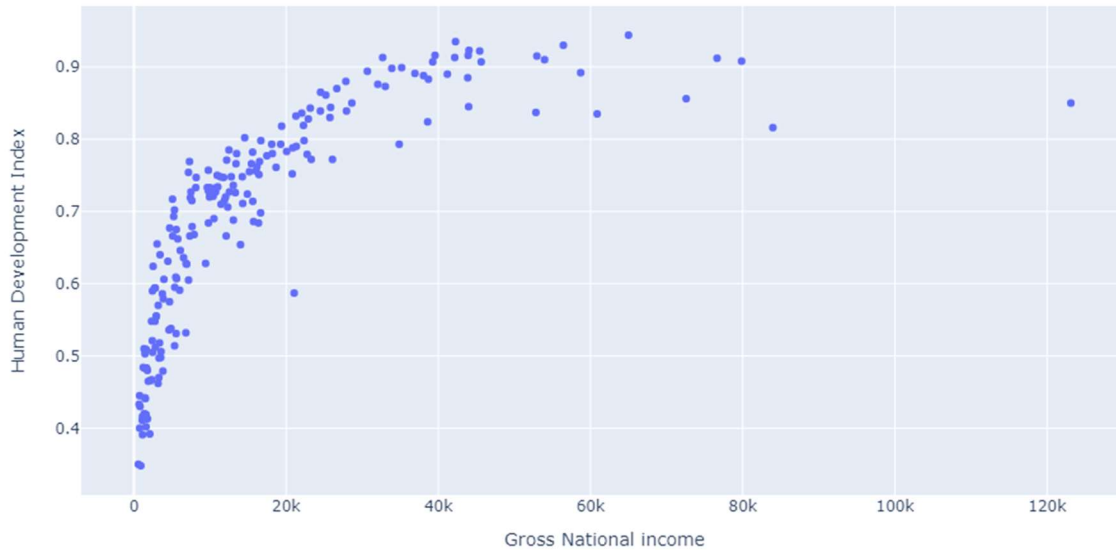


## What has a bigger effect on HDI, Gross National income, or Education?

The graphs above showed that there was a clear disparity between the HDI for each country. We wanted to know what were the variables that had the most effect on HDI. To do so we first created a correlation graph between the **Gross National income and the HDI** and after a correlation graph between **Education and the HDI** for each country. To do so, we created a scatter plot. Scatter plots are useful when trying to confirm a correlation between two variables. We added each variable on an axis.

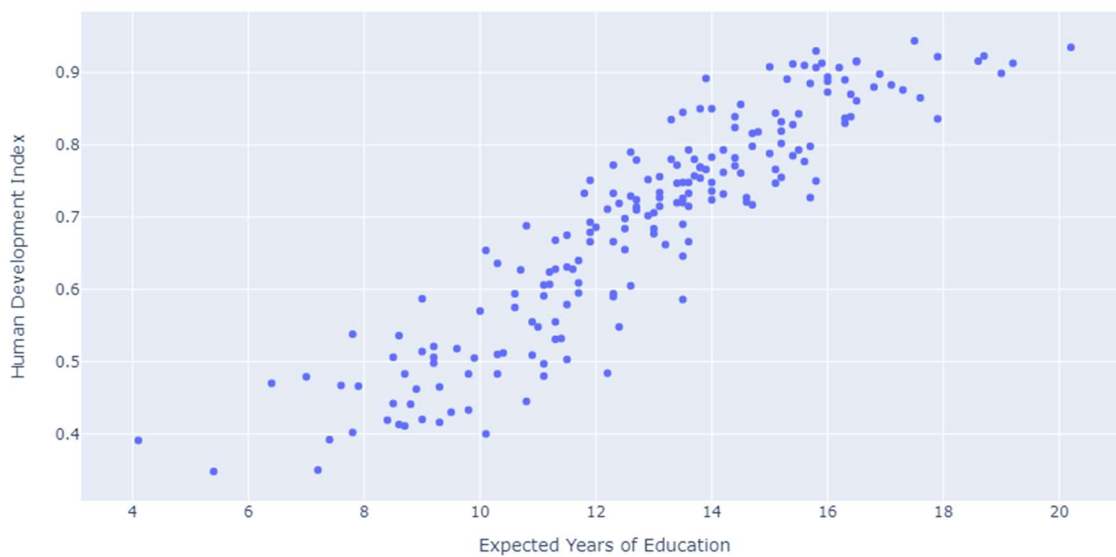
1)

Relationship between Gross National income and Human Development Index



2)

Relationship between Expected Years of Education and Human Development Index



We can see for both graphs if the value in the X-axis increases or decreases then the corresponding value of Y also increases or decreases. This means both plots have a positive slope if we fit a line through the points. Hence, both have a positive correlation. Moreover, we can observe in graph 1 that the rate of change of the value of Y for change of X is much higher than in graph 2. Hence, **Gross National income** has a stronger positive correlation with HDI compared to **Education**.