# Submission 2

We have chosen Human Development Index (HDI) as the development indicator for the second assignment. As described popularly,
The Human Development Index is a statistic composite index of life expectancy, education, and per capita income indicators, which are used to rank countries into four tiers of human development.
To understand it better, see the attached image,
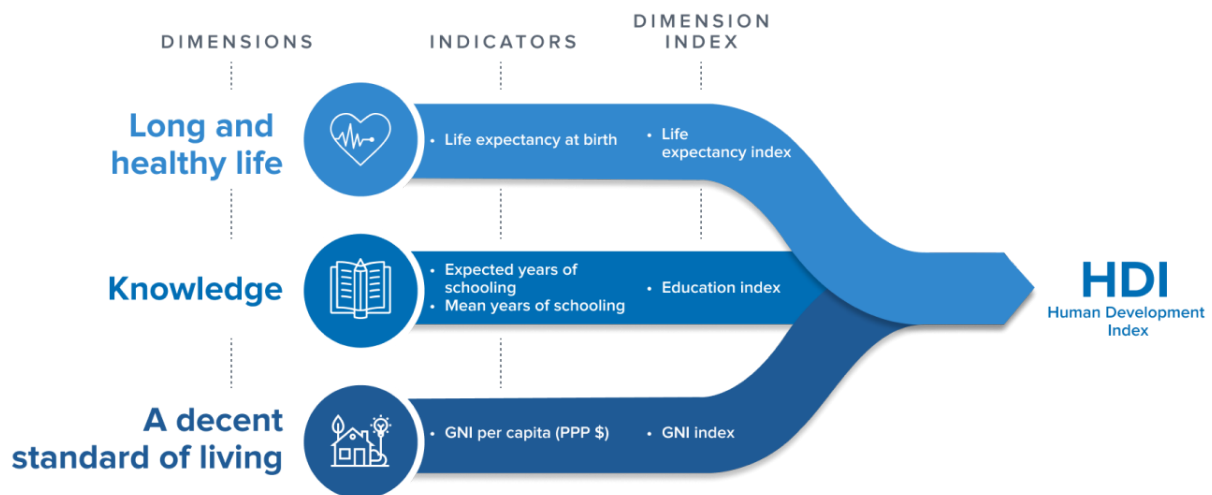
We have taken three indicators namely,
- GNI per capita to determine a decent standard of living
- Life expectancy at birth to determine the long and healthy life
- School enrollment ratio to determine the knowledge

All these variables contribute to computing the Human Development Index and thus are the **independent variables**.
Since HDI is a relatively new development indicator, World Bank Data could only provide data from the year 1990 to 2020. Thus we trained the model during these years.

For the regression, the independent variables were:
X_1 - Year
X_2 - GNI Per Capita (in US$)
X_3 - Life Expectancy (in Years)
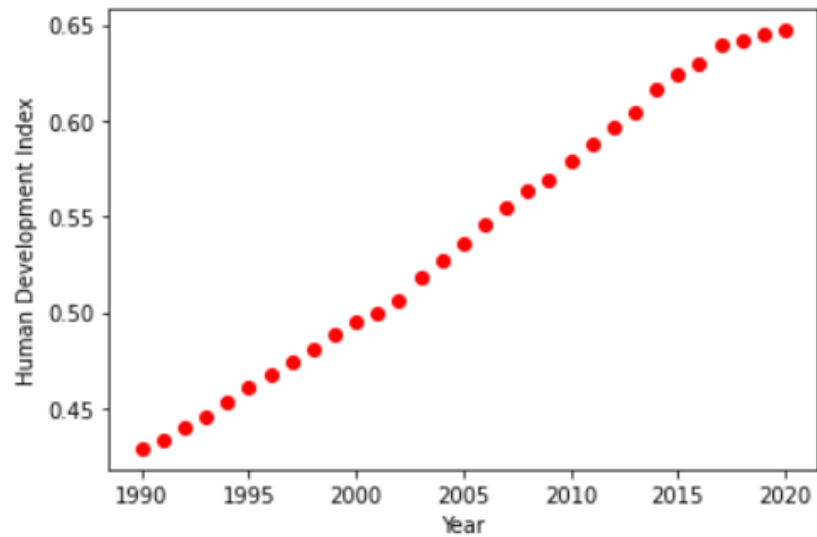X_4 - Gross Enrollment Ratio.
And the dependent variable Y is HDI.

We imported train_test_split from sklearn to divide the data into training data(80%) and testing data(20%).

Further, we used three regressors to train our model,

- Linear Regressor (Accuracy - 99.4358%)
- Decision Tree Regressor (Accuracy - 90.6253%)
- Random Forest Regressor (Accuracy - 93.9737%)

Graph plotted:

<u>Using Matplotlib (Also included in Colab file)</u>

<u>Using Excel</u>