EXPLORATORY DATA ANALYSIS

WORLD BANK GLOBAL EDUCATION

By:
Neel Naik
Prasad Khadatkar
Piyush Nirwan
Shivank Shukla
Saransh Jain

Introduction

- Education is very important for **personal development** as well as for the **socio-economic development** of our country.
- **Education is important** for living life happily.
- It **empowers and motivates** our mind to conceive the good thoughts and ideas and grow our knowledge day by day.

Questions???

- What are the parameter which **effect Education** on broader level ?
- Is there any correlation between **amount of money spent** by Government ?
- Why some countries having **good Education system** and literate labour Force ?
- Does Education **really affect society** if it affect, then How?

EDA APPROACH
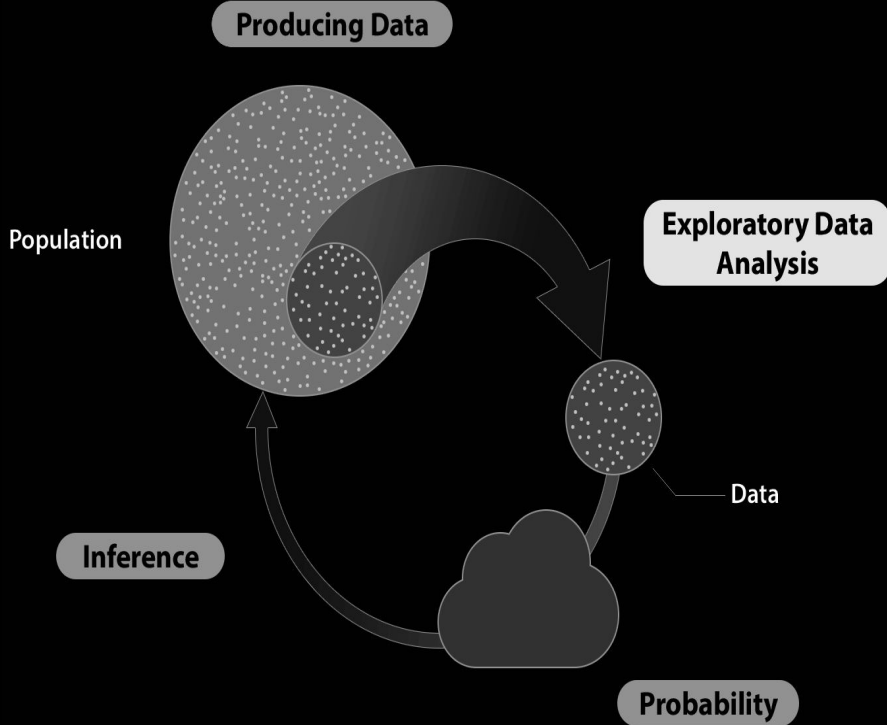
- **Importing Required Libraries**

- **Loading Datasets**

- **Data Understanding**

- **Data Cleaning**

- **Data Analysis**

- **Data Visualization**

- **Inference**

Understanding Dataset

- **Tables**
- **Columns**
- **Data Quality**
- **Data Selection**
- **Column analysis**
- **Table analysis**

We understood that...

- Indicators for various parameters of education and economy.
- Values of parameters from **1970 - 2015**
- Discount the columns that are **not useful**.
- **Classes** w.r.t economy
- Correlation b/w spending on education and its outcome
- **Teachers to student** ratio.
- Trend change in importance of education

Data Understanding

main_data.shape							
(886930, 70)							

main_data.head(7)

	Country Name	Country Code	Indicator Name	Indicator Code	1970	1971	1972
0	Arab World	ARB	Adjusted net enrolment rate, lower secondary, ...	UIS.NERA.2	NaN	NaN	NaN
1	Arab World	ARB	Adjusted net enrolment rate, lower secondary,	UIS.NERA.2.F	NaN	NaN	NaN

main_data.describe()

	1970	1971	1972	1973	1974
count	7.228800e+04	3.553700e+04	3.561900e+04	3.554500e+04	3.573000e+04
mean	1.974772e+09	4.253638e+09	4.592365e+09	5.105006e+09	5.401493e+09
std	1.211687e+11	1.804814e+11	1.914083e+11	2.059170e+11	2.112150e+11
min	-1.435564e+00	-1.594625e+00	-3.056522e+00	-4.032582e+00	-4.213563e+00
25%	8.900000e-01	8.853210e+00	9.240920e+00	9.595200e+00	9.861595e+00
50%	6.317724e+00	6.316240e+01	6.655139e+01	6.969595e+01	7.087760e+01
75%	6.251250e+01	5.655200e+04	5.863650e+04	6.202900e+04	6.383675e+04
max	1.903929e+13	1.986457e+13	2.100916e+13	2.238367e+13	2.282991e+13

Using Basic attributes and methods on dataframe for having data insight.

Data Cleaning

- Calculating % of **null values**.
- Getting Info. for the columns.
- Dropping columns having null values above **90%**.
- Creating **sub-Data frames** according to the required info. With the provided Indicators.

Data Cleaning

Country Name	0.000000
Country Code	0.000000
Indicator Name	0.000000
Indicator Code	0.000000

2010	72.665036
2011	83.537370
2012	83.396209
2013	84.496071
2014	87.170464
2015	85.223411
2016	98.144160
2017	99.983877
2020	94.200670
2025	94.200670
2030	94.200670
2035	94.200670
2040	94.200670

Keeping **non null** values

Removing columns with
Null data more than **90%**

Data Analysis

Analysis based on the following:

- **Income group**

- **Region**

- **Big Economies of the world**

GDP of world on map
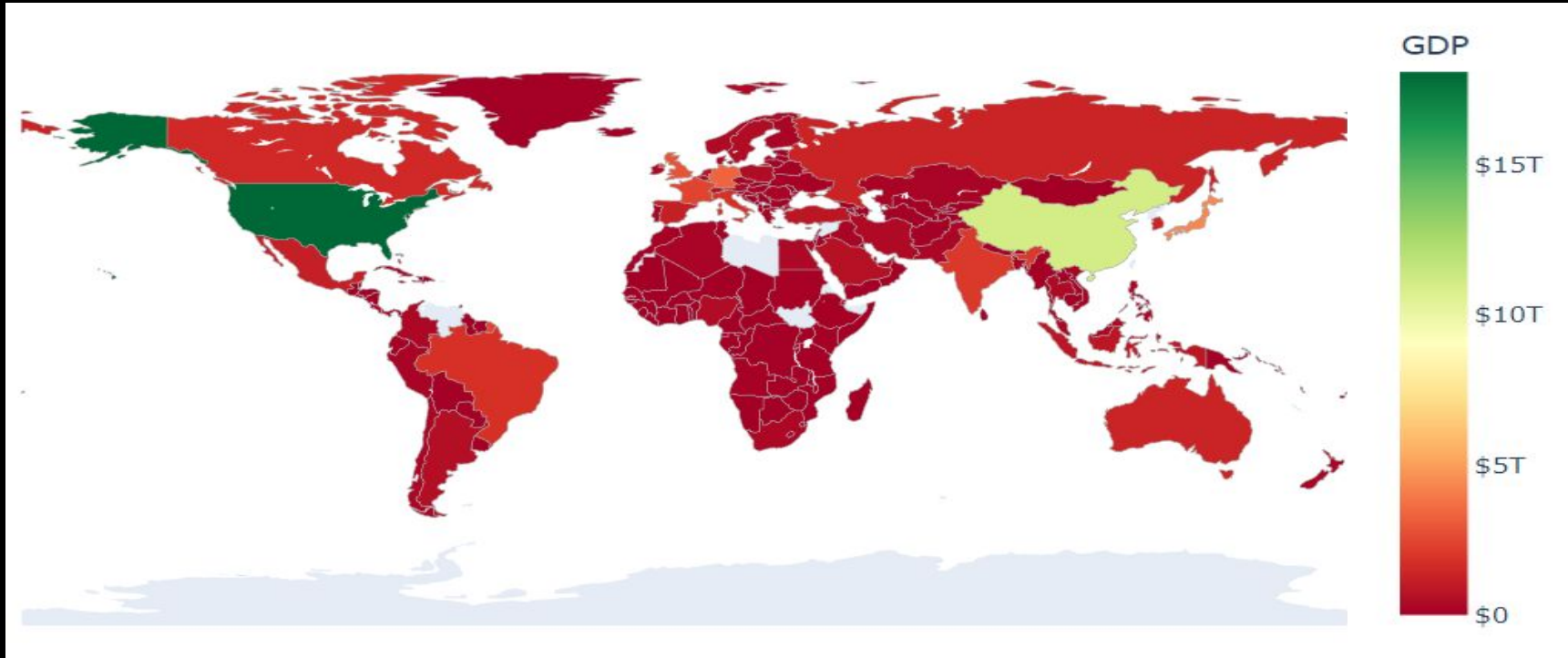
Inference: Most of the world having GDP **less than $ 3T** economy, GDP of **INDIA is around $ 3T.**

Countries above mean gdp of world Mean_GDP = $374979491861.97833
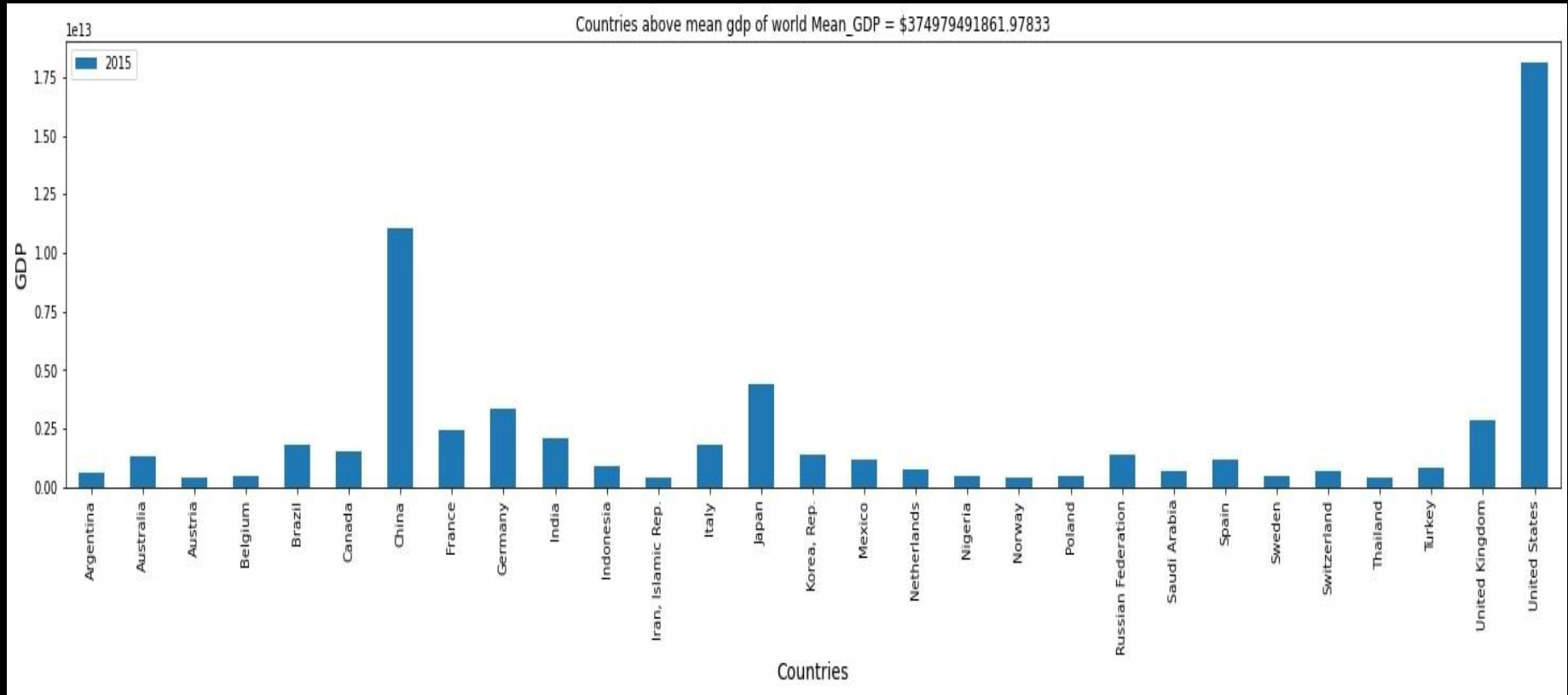
Countries are having **GDP** above **mean GDP** and **mean per capita GDP** of world, which gives us macro and micro perspective.

Countries above mean per capita gdp of world = $ 14829.991228734547
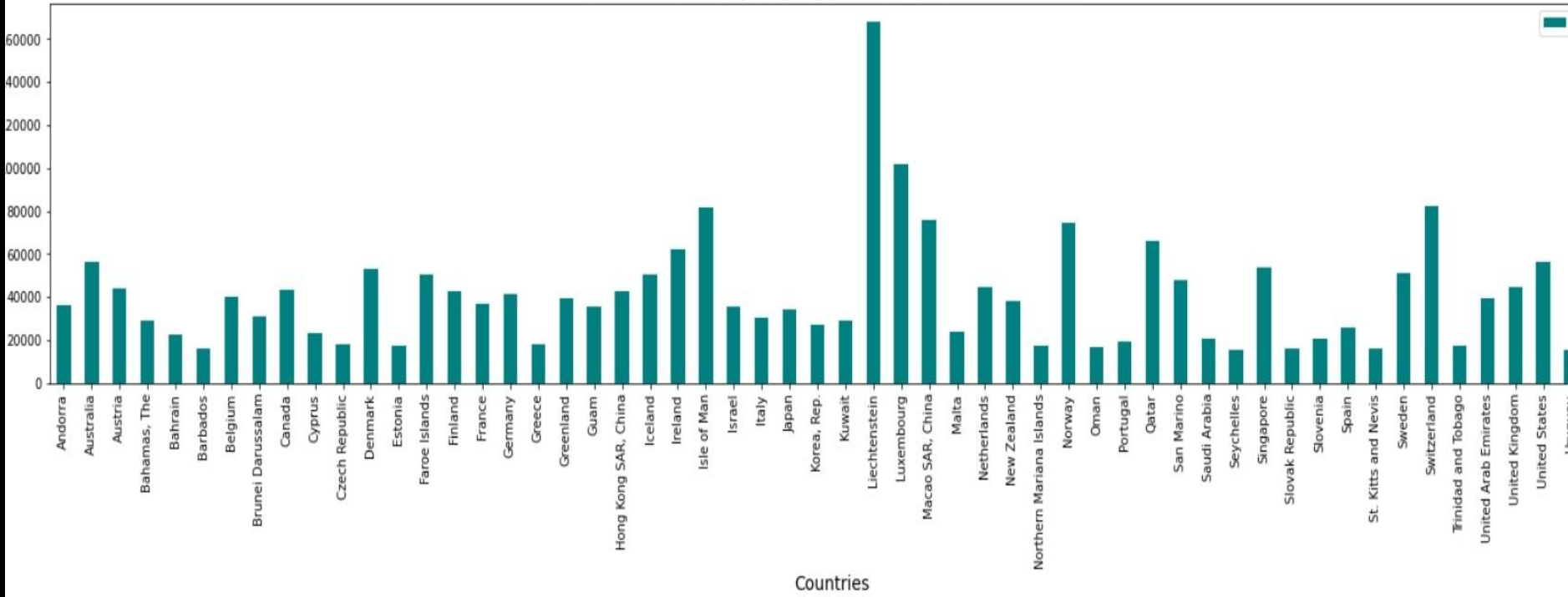
Inference: India's per capita GDP is below the mean per capita GDP of world despite of the fact we are **5th largest** economy of world. Which affects India's education and population growth.
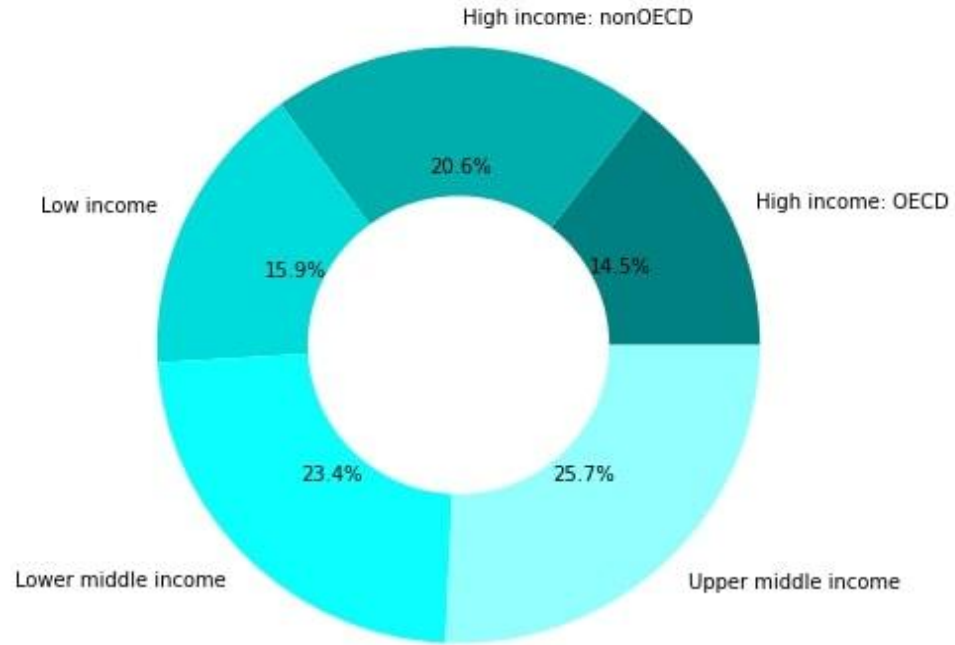
World Economy class distribution

Low Income = 1, Lower Middle Income = 2, Upper Middle Income = 3, High income = 4

Inference: **Europe and Northern countries** lies within **High Income Group** while the **Southern and asian** countries belongs to **upper middle** income and **lower middle** income.
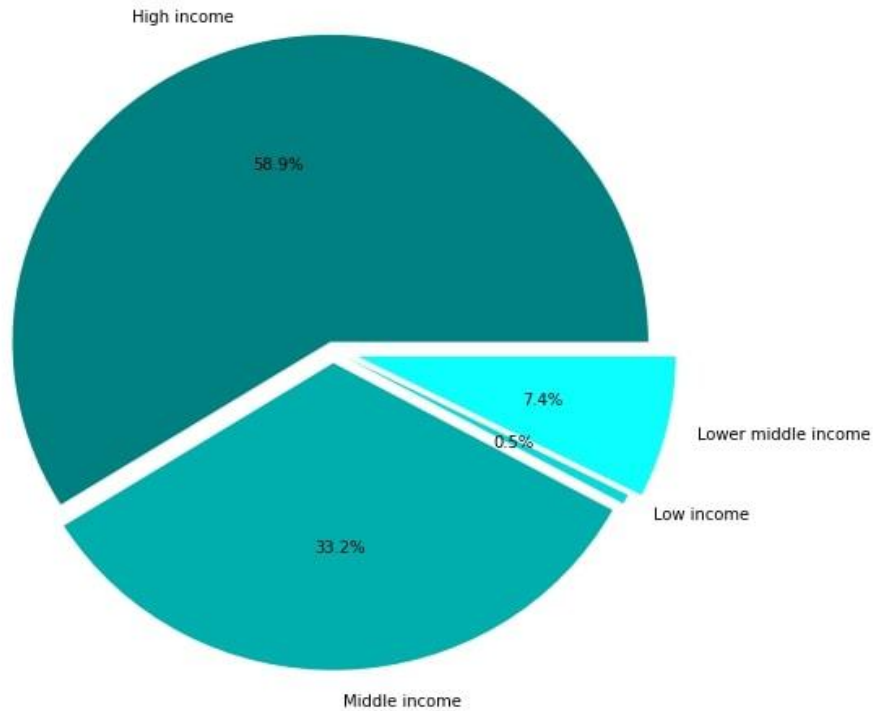
**Percentage share of Income group in GDP of World**

Composition of world wrt to Income group

Number of kid not enroll to primary level
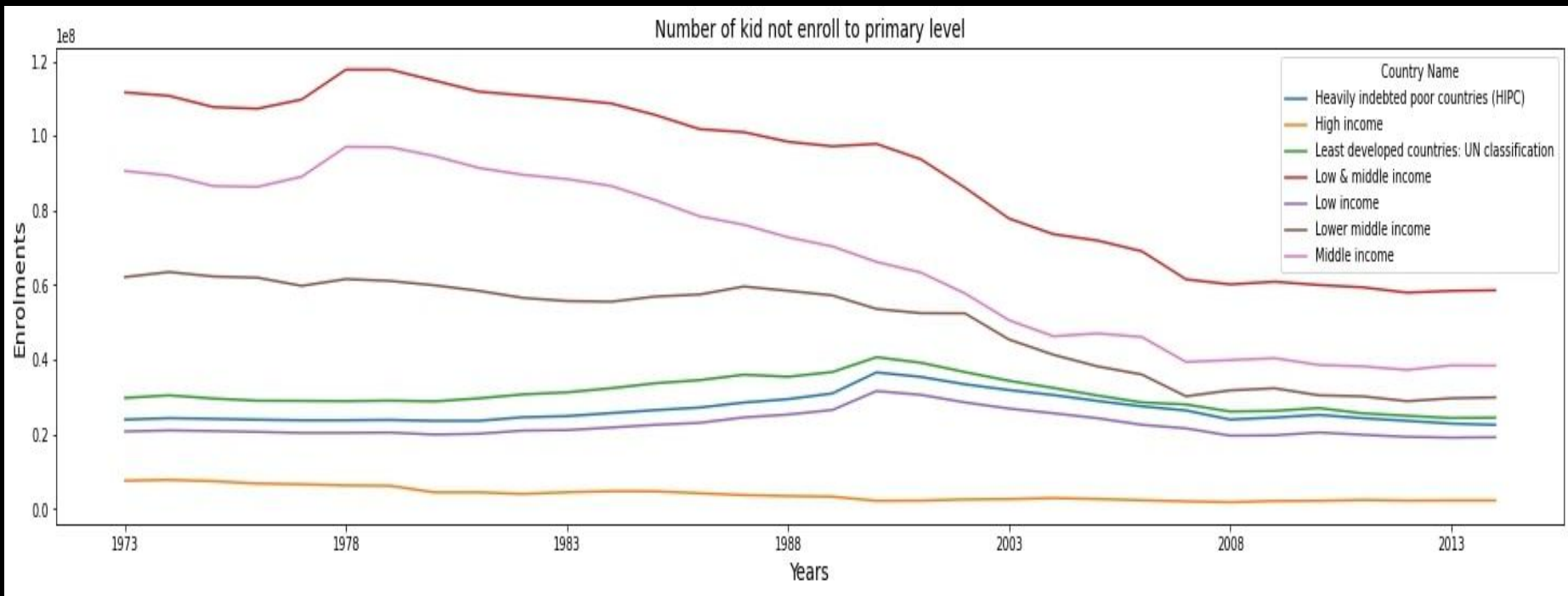
Inference: Number of kids not getting **primary education** is **decreasing** over the years, and it is **lowest** for **High income** group. The more economical stable the income group is the more they are educated. **Education** is **directly proportional** to the **money** we earn and vise-e-versa.

Female Enrolment of various level of schooling for different income group
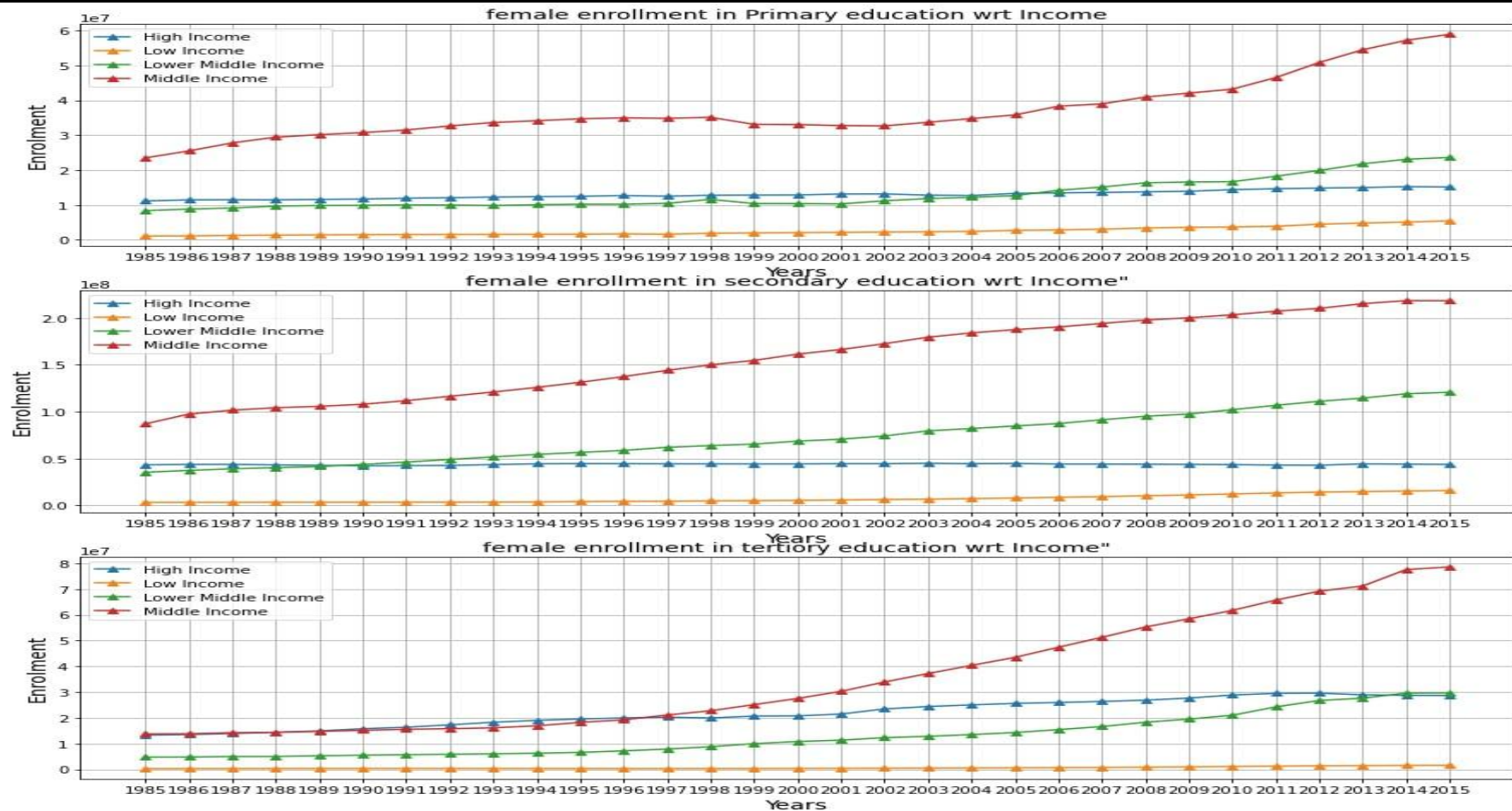
Population growth among different income groups



Income Based Population growth
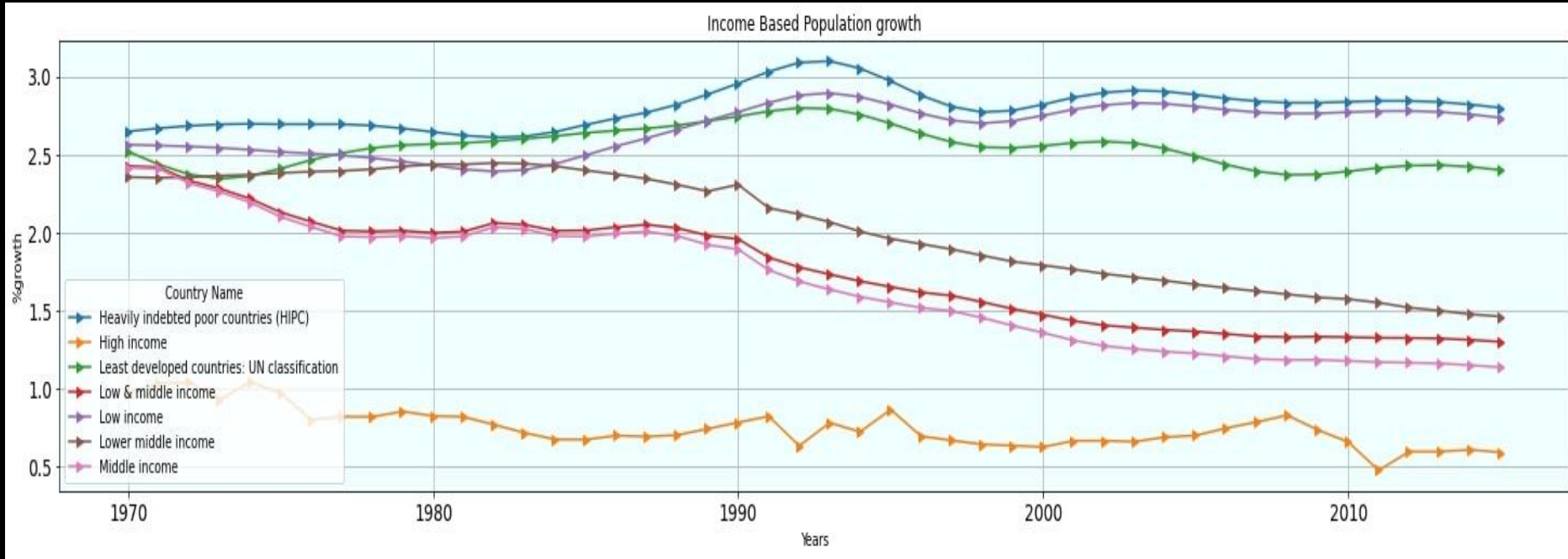
Inference: With the **Increase in education** the population growth shows a **decreasing trend** over the years, which is a **positive trend** and the population growth is almost stagnant for High income Countries and it is highest for Heavily indebted countries, again it is **proved that economy, education and Social condition are highly correlated.**

Teacher to pupil ratio, a important parameter for quality education

Pupil to teacher ration in primary education
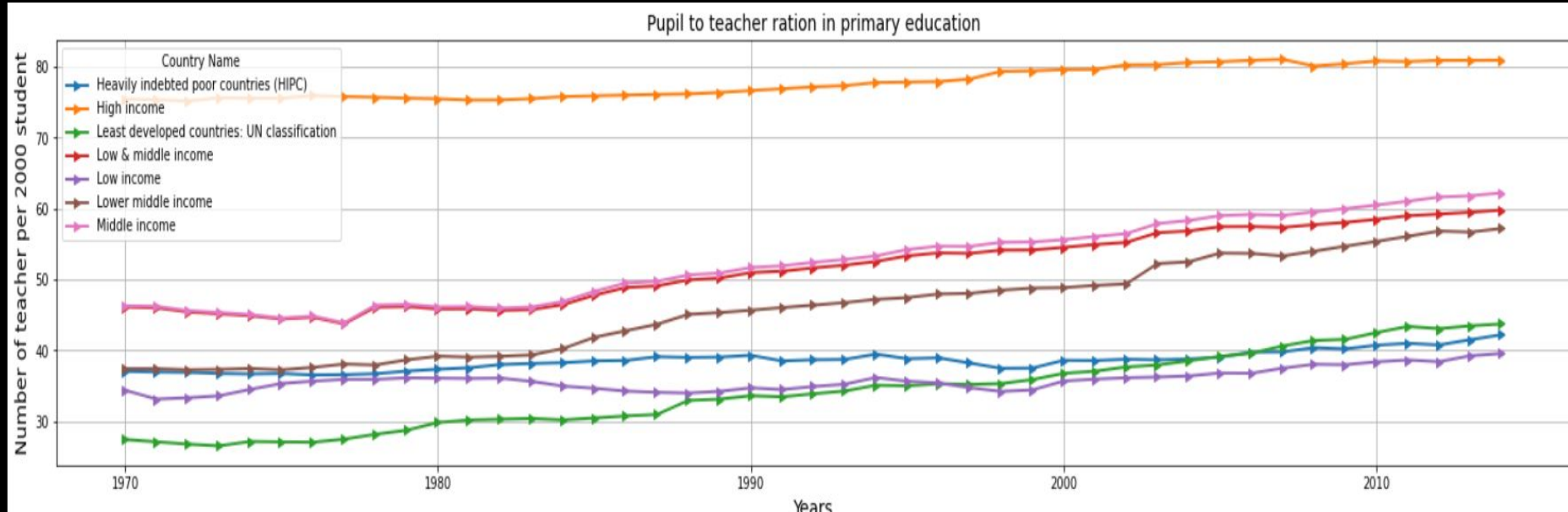
Inference: High Income Countries are able to provide quality education because the **ratio of teacher is over 2000 students** is maximum there, and it is not as per the suggested no. by world bank in Lower income countries and least developed countries, but the **trend is increasing** which is positive outcome.

Education spending by big Economies as a % of GDP

Education spending by big Economies as a percetage of GDP
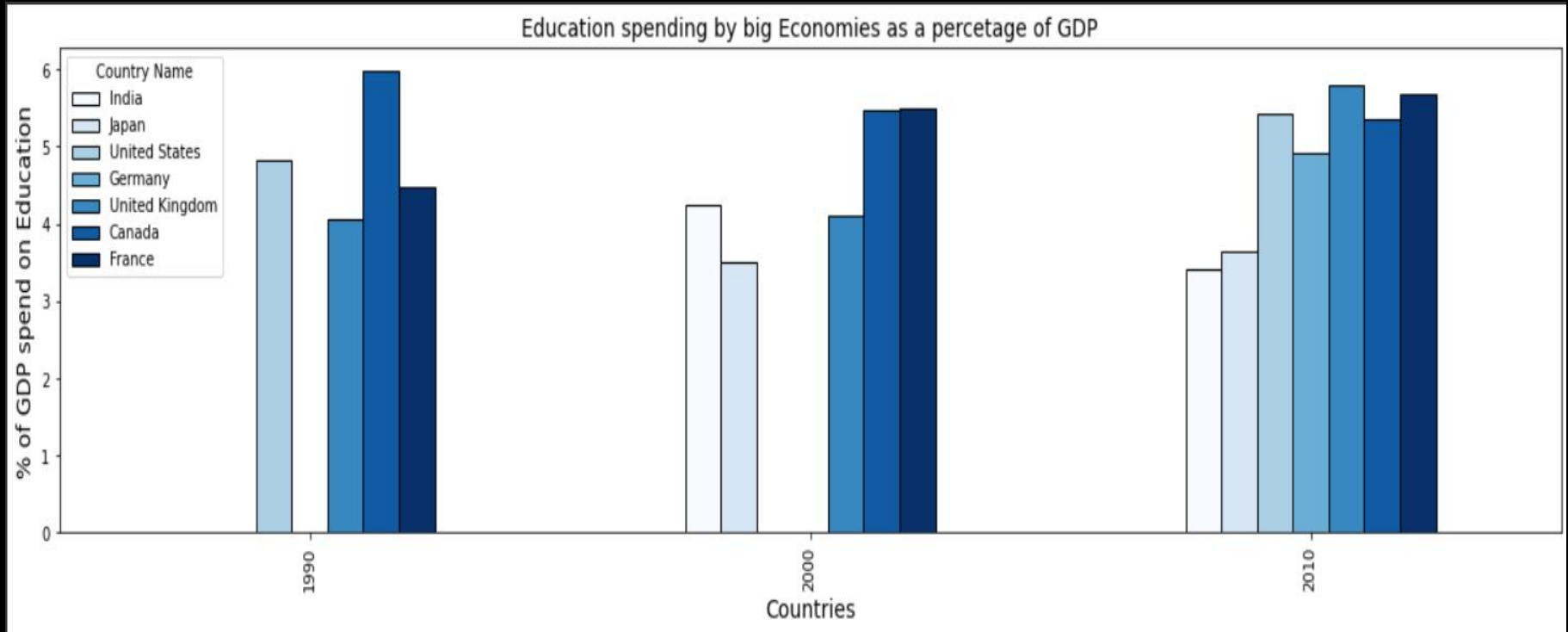
Inference: The big economies spent a **big percentage** of GDP **on Education**, hence Education is directly proportional to Development and **growth of Nation.**

Labour Force from Big economies

Labour workforce of big 7
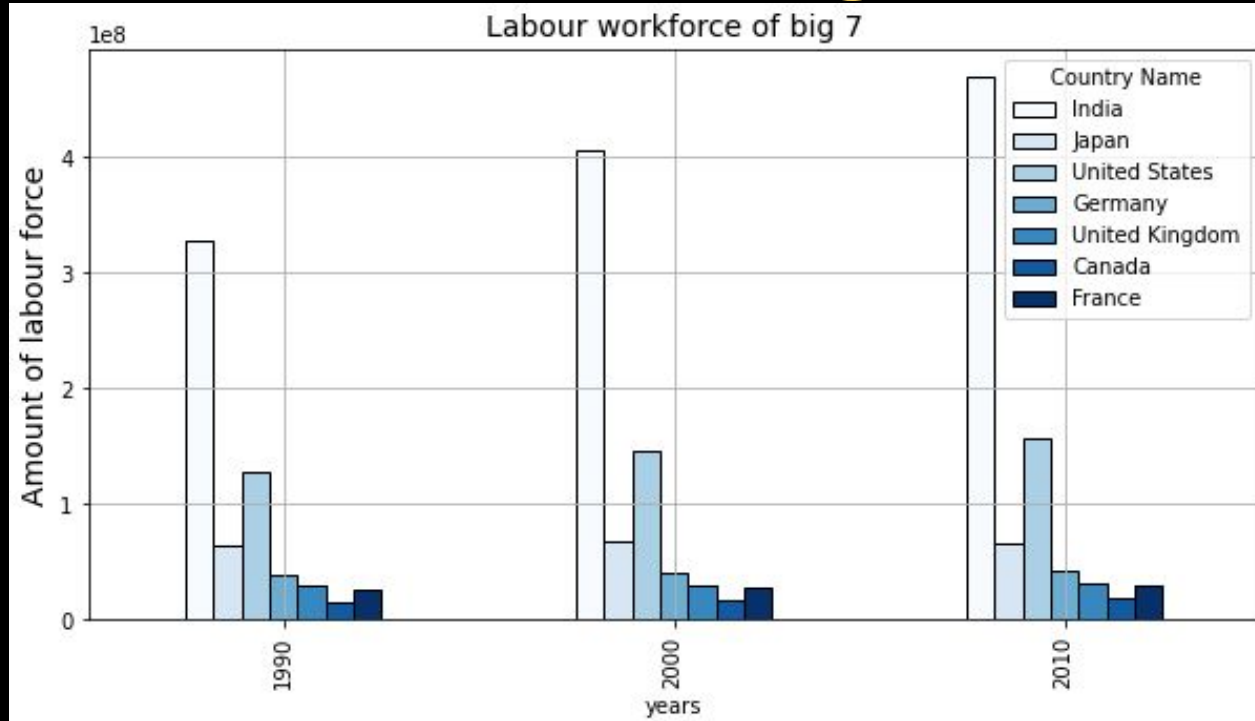
Inference:- The **maximum work force** is available in **India** within World's largest economies but **per capita income** is even **less than mean** per capita income of world which also influence lifestyle of Citizens.

Correlation between GDP, Unemployment and illiterate labour
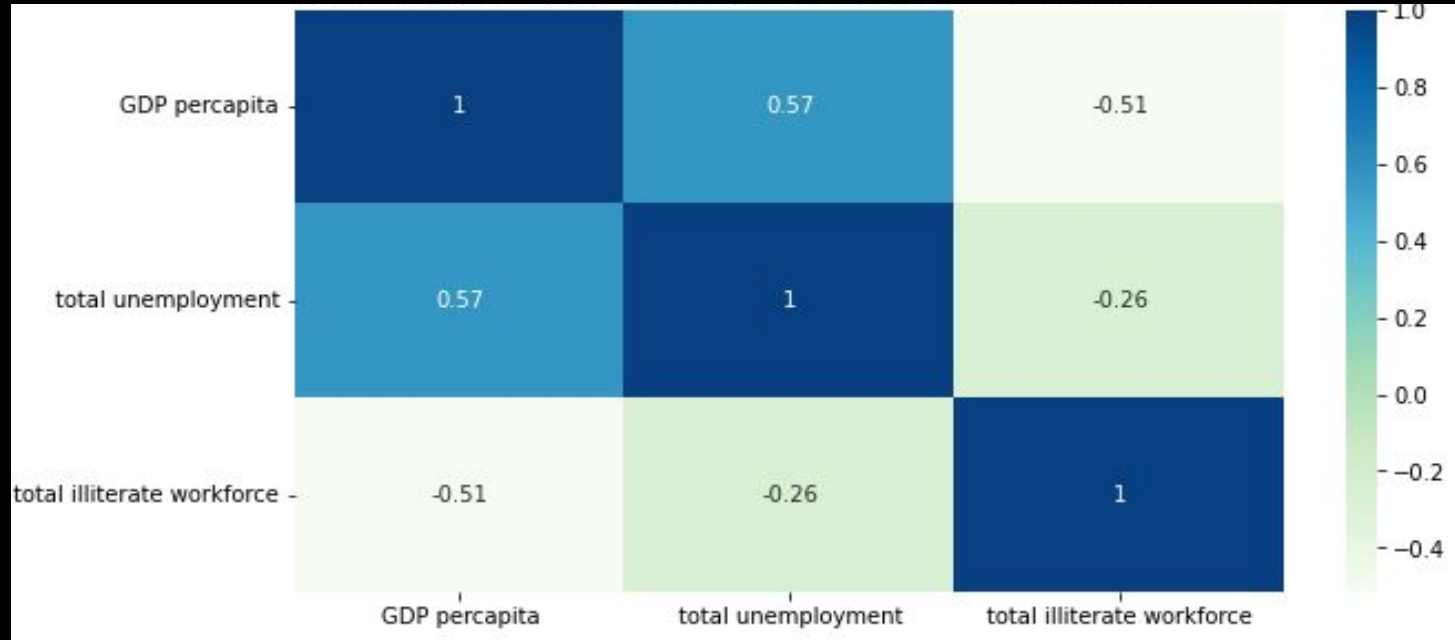
Inference: A skilled labour force is the building block of a nation but skills also requires education, which is directly or indirectly related to per capita GDP of that country. We can correlate GDP with various parameters. As the **per Capita GDP decreases unemployment increase shows a negative correlation.**

CONCLUSION

1. The **economic condition** of a country and region highly **influence the education** of that country or region.
2. An **educated labour force** always paved path to the **growth** and development of nation.
3. Increasing the **spending on Education**, Awareness for the enrollment on every level, **Increase teacher to pupil ratio** and awareness toward the **population growth.**

SUGGESTION

1. **Awareness programmes** should be incorporated in order to have a sustainable population growth
2. There should be more **spending on education** as a percentage of GDP
3. we must be focused towards **education of children** and re-education of the labor force.
4. we should follow the **norms suggested by the world bank** for Teacher to Pupil ratio.
5. **Better data collection .**