Capstone Project-1

(Exploratory Data Analysis)

World Bank Global Education Analysis

Team Member

Er. Mudita Garg(mudita.garg.99@gmail.com),
Pankaj Kumar Yadav (pankajkumaryadav1500420@gmail.com),
Harshada Gore(harshada1442@gmail.com)

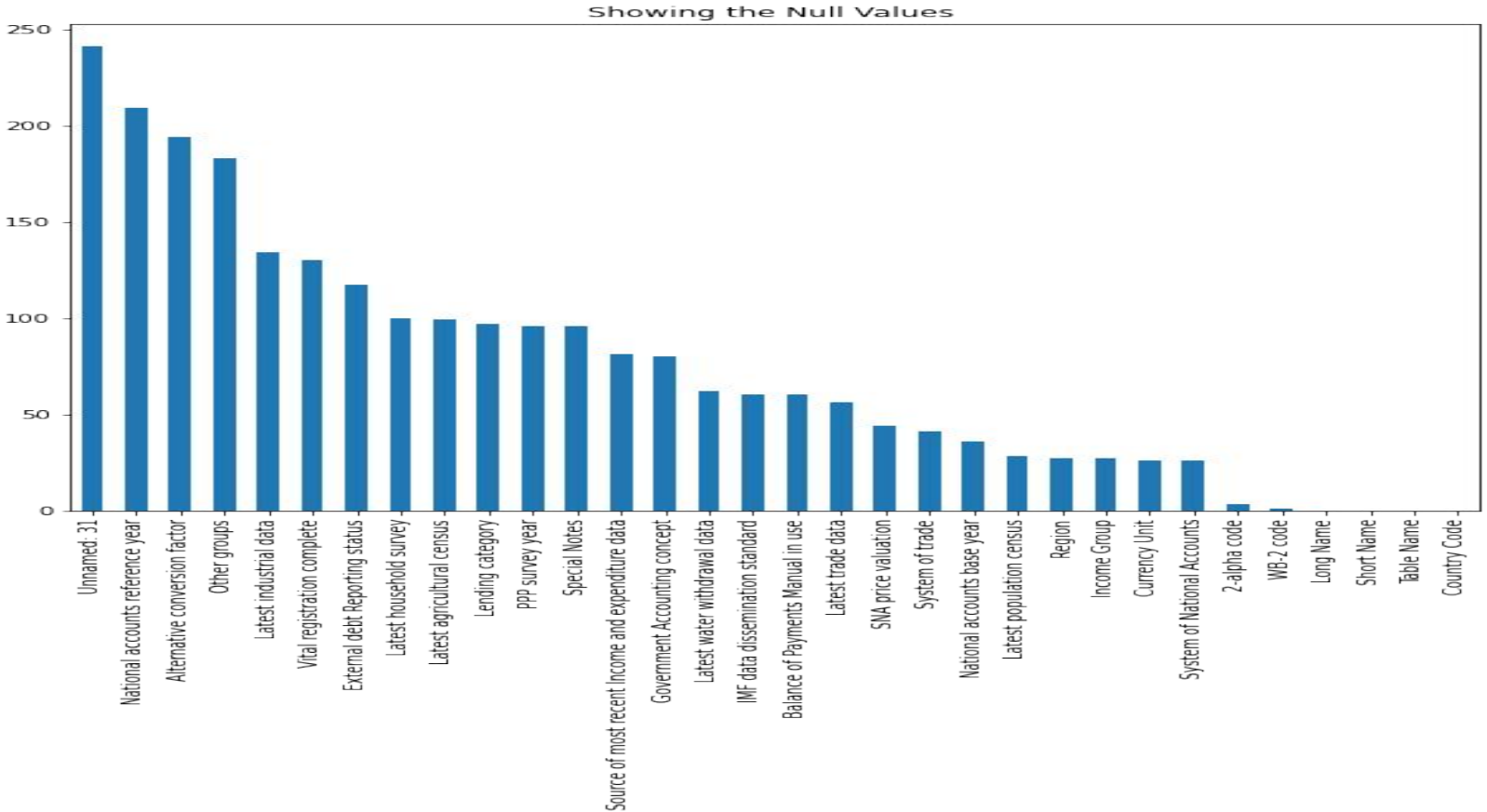
Points for Discussion :

- Data Summary
- Showing Null Values
- No. of country in Region and Various Income Group.
- Population growth % of various countries categories by Income Group.
- Percentage ratio of "Developed Countries" and "Developing countries“
- *GDP (Gross domestic product) of 3 top and bottom countries of "Developed Countries" and "Developing countries“*
- Deep analysis of education scenario in India
- Analysis of Adjusted net enrolment rate calculated
- Find the correlation between factors on affecting the GDP of India
- Comparing India with neighboring countries - Pakistan, Nepal, Bhutan, China, Sri Lanka in Gross Enrolment Ratio **(2010)**
- Variation in performance of Females and Males in PISA and TIMSS
- Conclusion

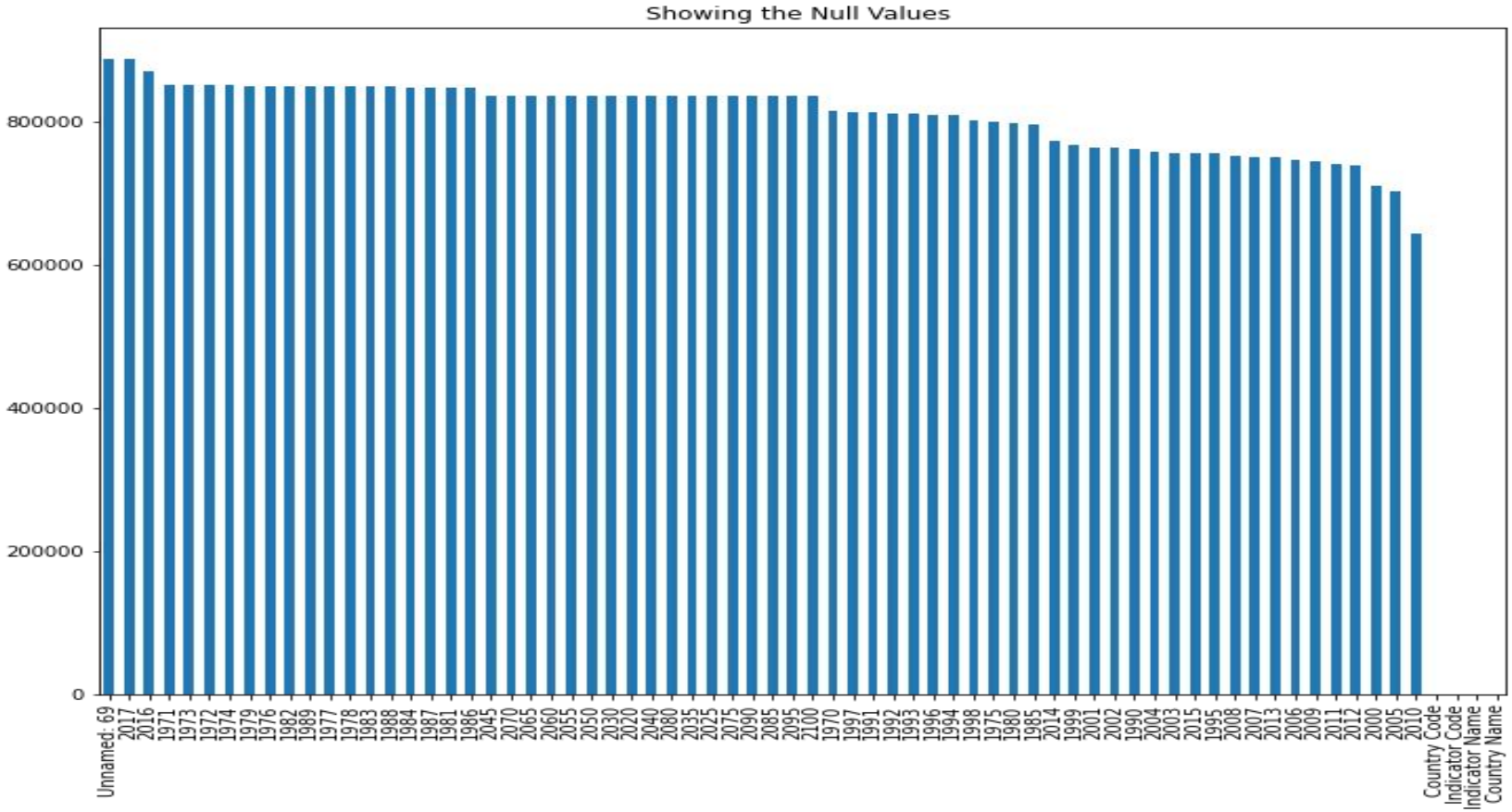
Data Summary :

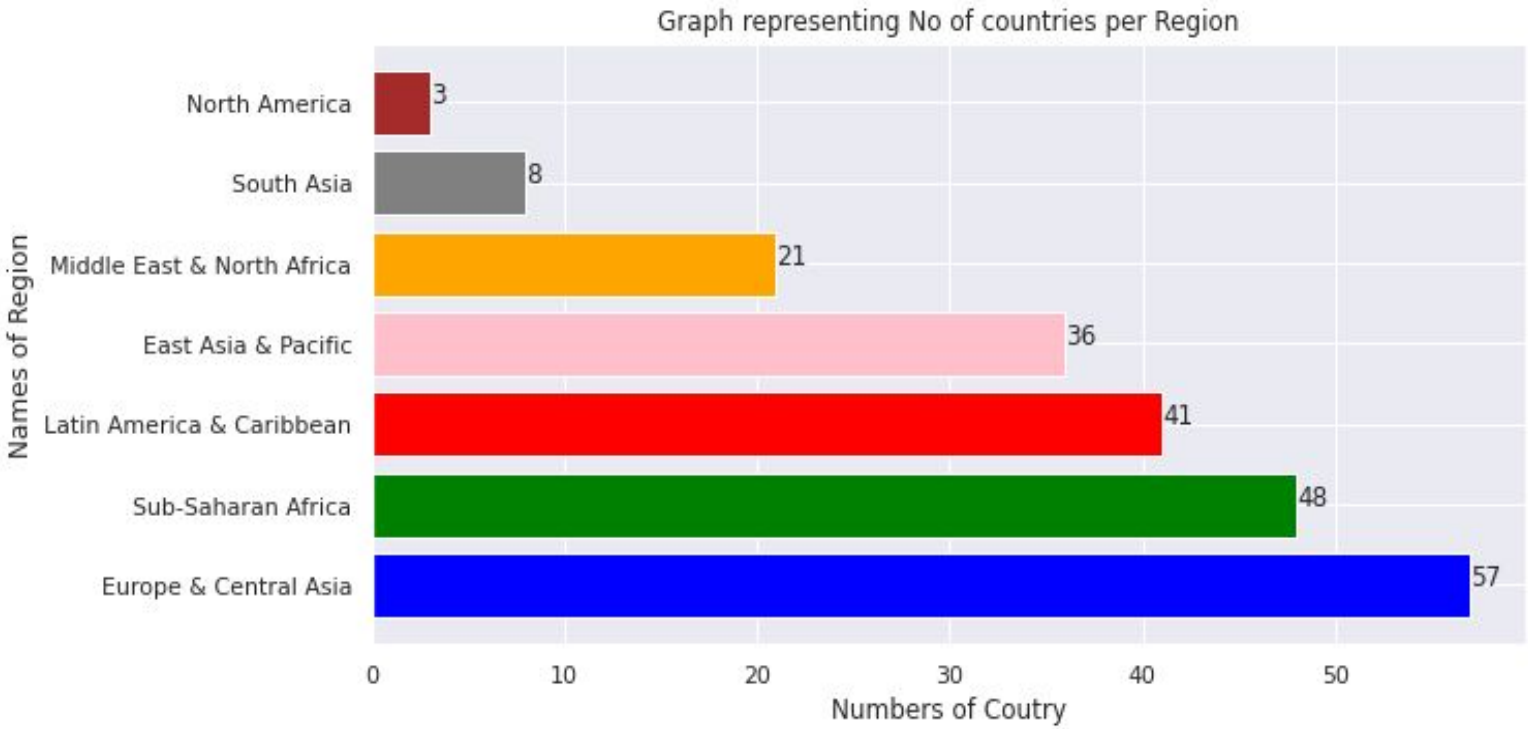
- **"EdStatsCountry.csv" data :** contains list of all countries that are present in the data (total of 241 countries) - along with other features specific to the country like - Region, Income Group specific to the country etc.
- **"EdStatsData.csv" data :** contains each country, with list of indicators (3665 unique indicators); contains measurement value for each indicator from years 1970 to 2017; from 2020 to 2100 - contains projections
- **"EdStatsSeries.csv" data :** contains list of all indicators and the definition of each indicator
- **"EdStatsFootNote.csv" data :** It contains the estimations and uncertainty bounds for each year - looks like some years are missing
- **"EdStatsCountry-Series" data :** It contains indicators and data sources for certain countries

- "EdStatsCountry.csv" data

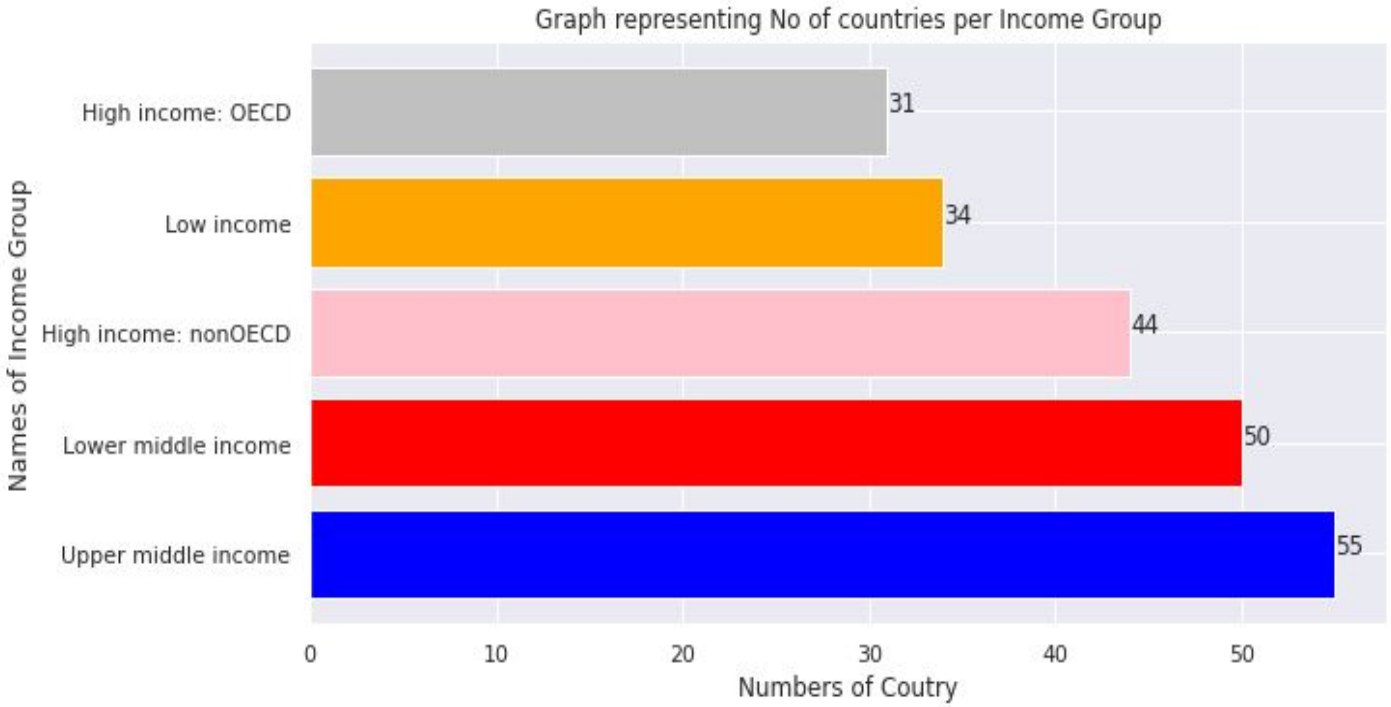


- "EdStatsData.csv" data

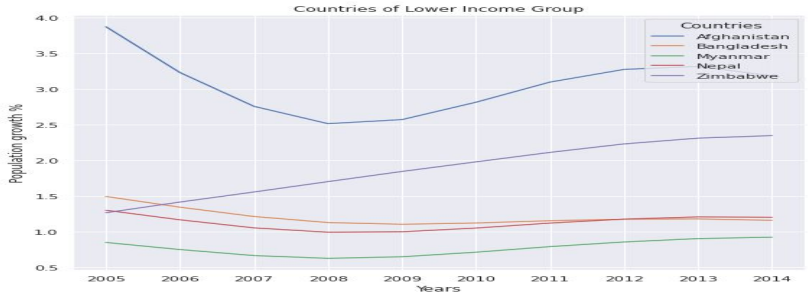


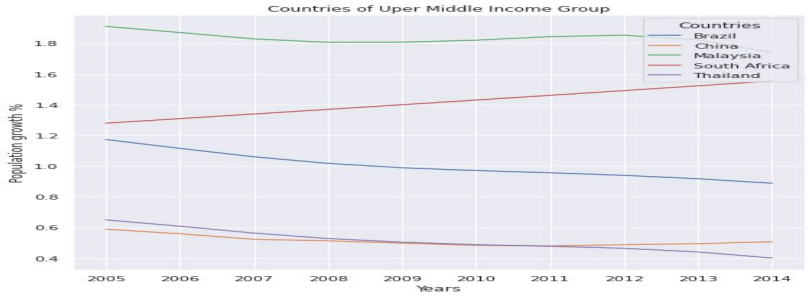


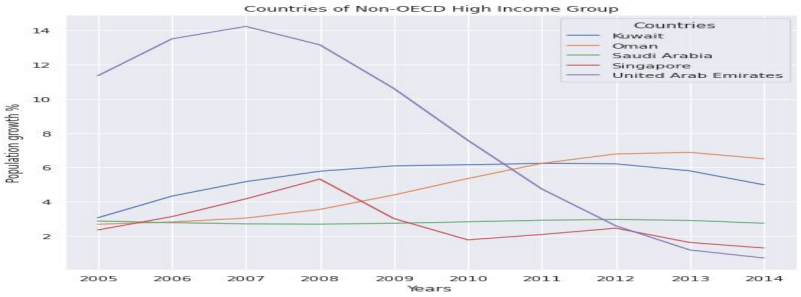
- In this section the countries, which are part of the World Bank, have been counted based on the region. Europe and central Asia have the highest participation (57 countries).

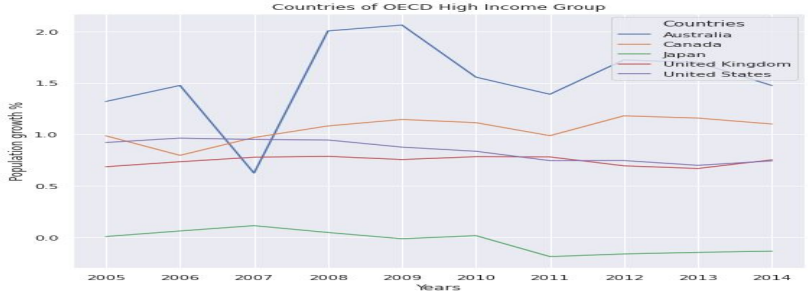


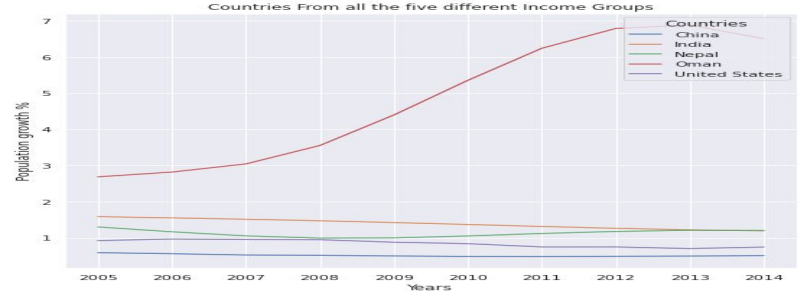
- Next task is to count the number of countries falling in various income groups.
- Here we can easily found that how many countries belongs from which 'Income Group'.



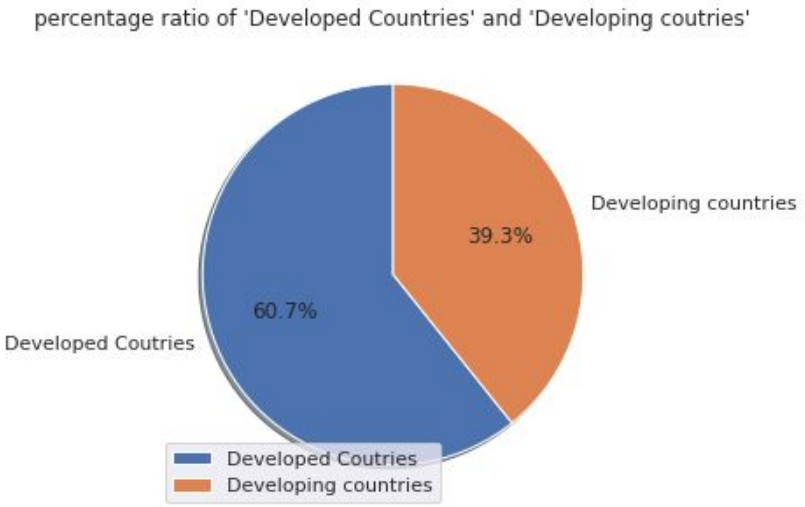








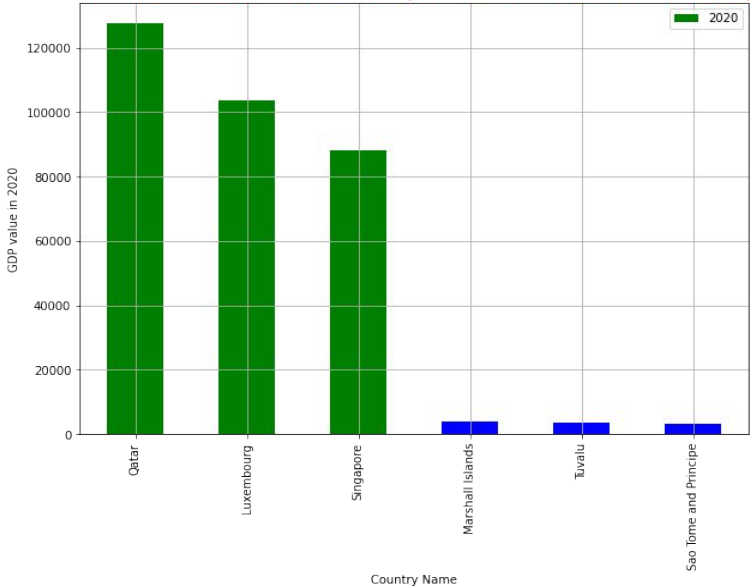
- Based on the income group, last five years data of percentage population growth has been analyzed for selected/renowned countries. Countries fall in ‘Lower and Lower-middle Income’ needs to worry about their population growth (example: Afghanistan, Zimbabwe).



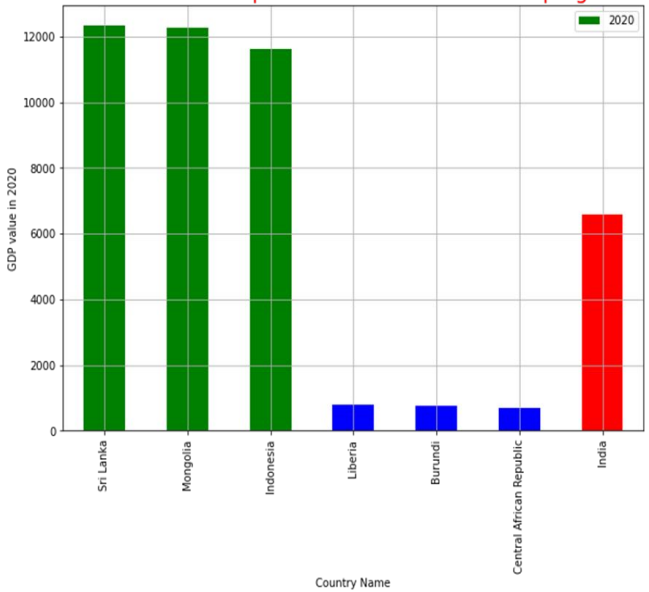
- According to World Bank Global Education 60.7% countries across in the world are Developed Countries

- According to World Bank Global Education 39.3% countries across in the world are Developing Countries

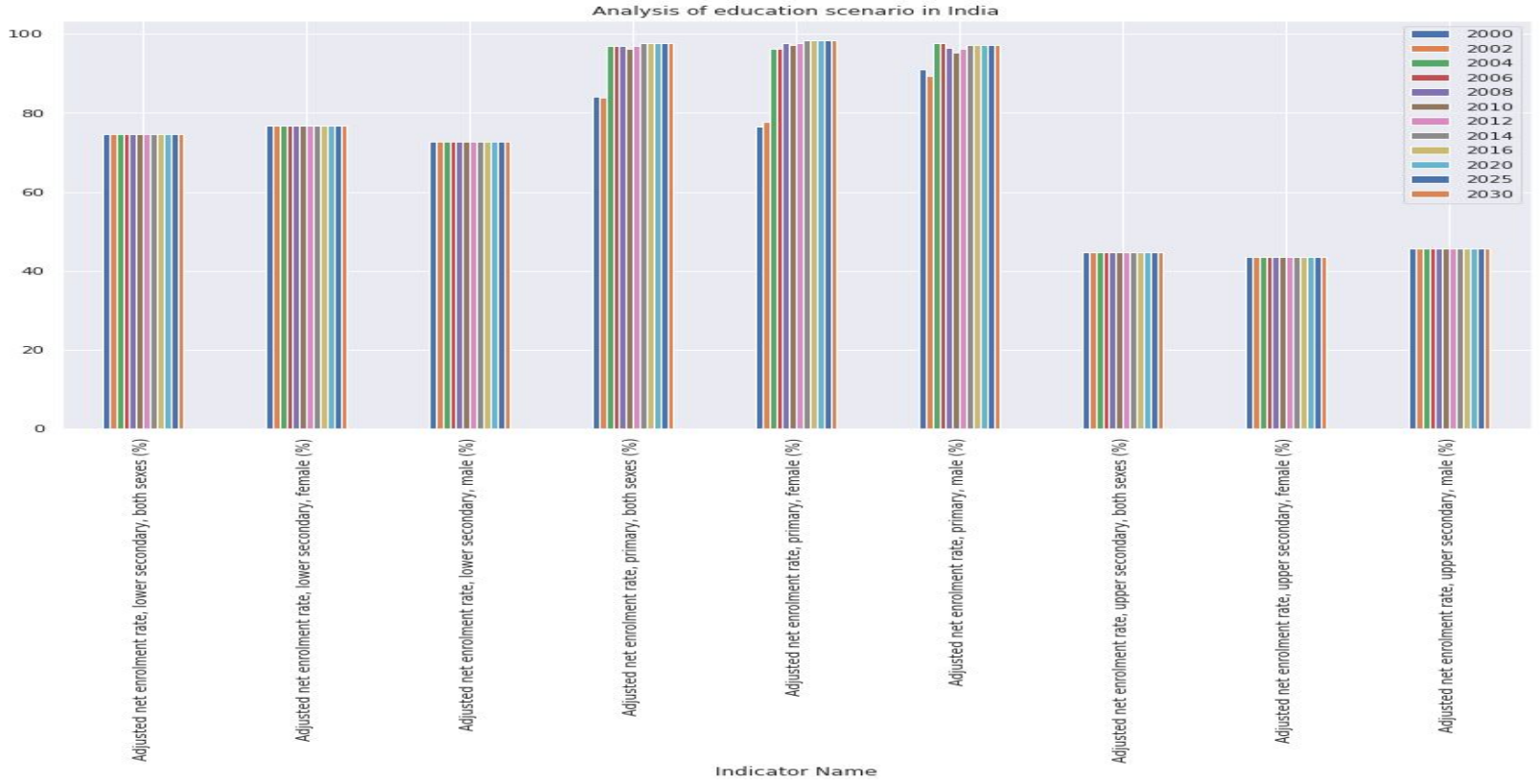
Top and Bottom three GDP values comparision for 2020 for Developed nations



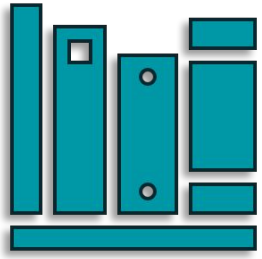
Top and Bottom three GDP values comparision for 2020 for developing countries, including India



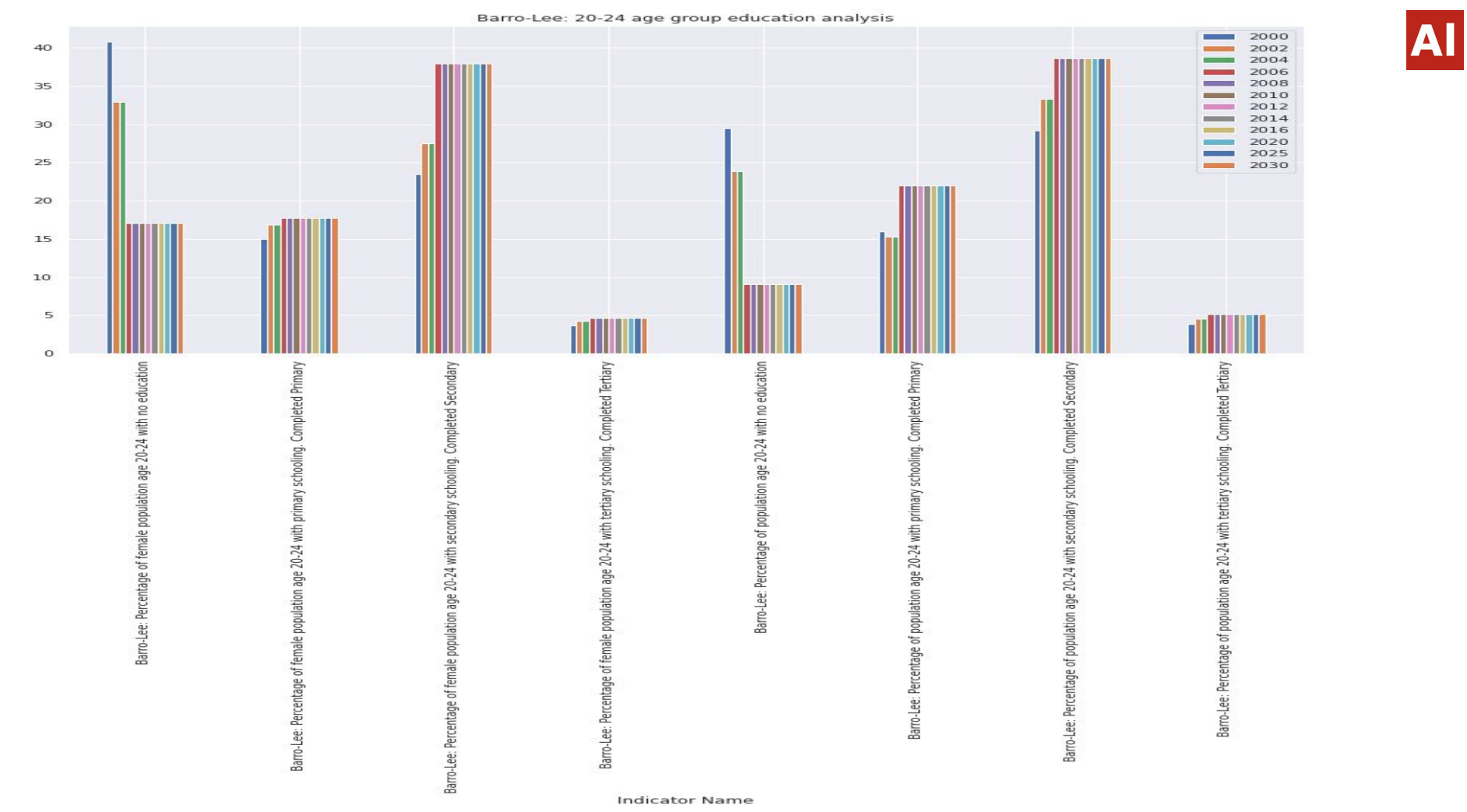
- India is a part of developing nations.
- GDP analysis for both developed and developing sections has been made to show the highest and lowest GDP countries in both sections. Qatar (exceeding 120000) has the highest GDP in 2020(developed nation).

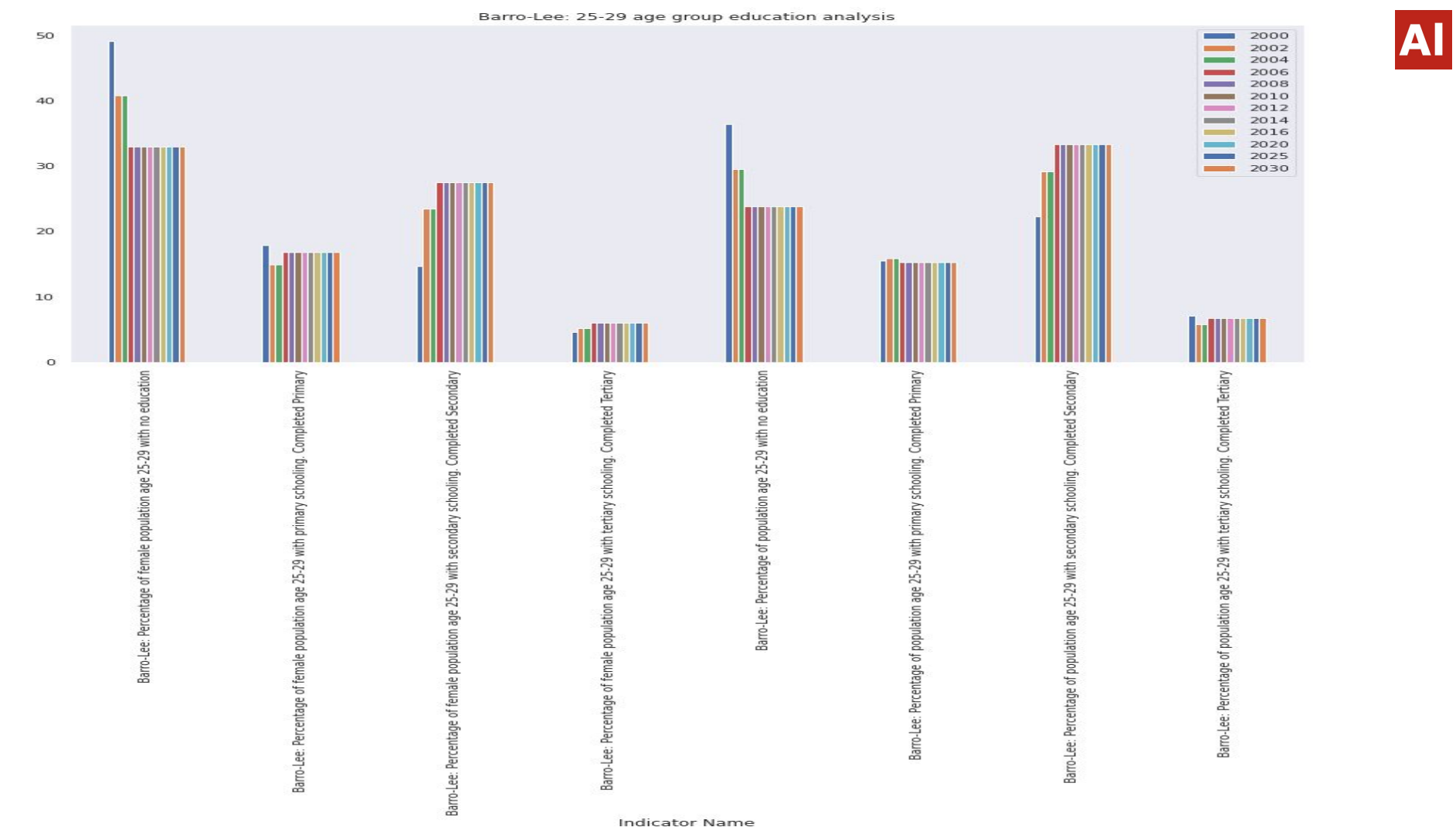


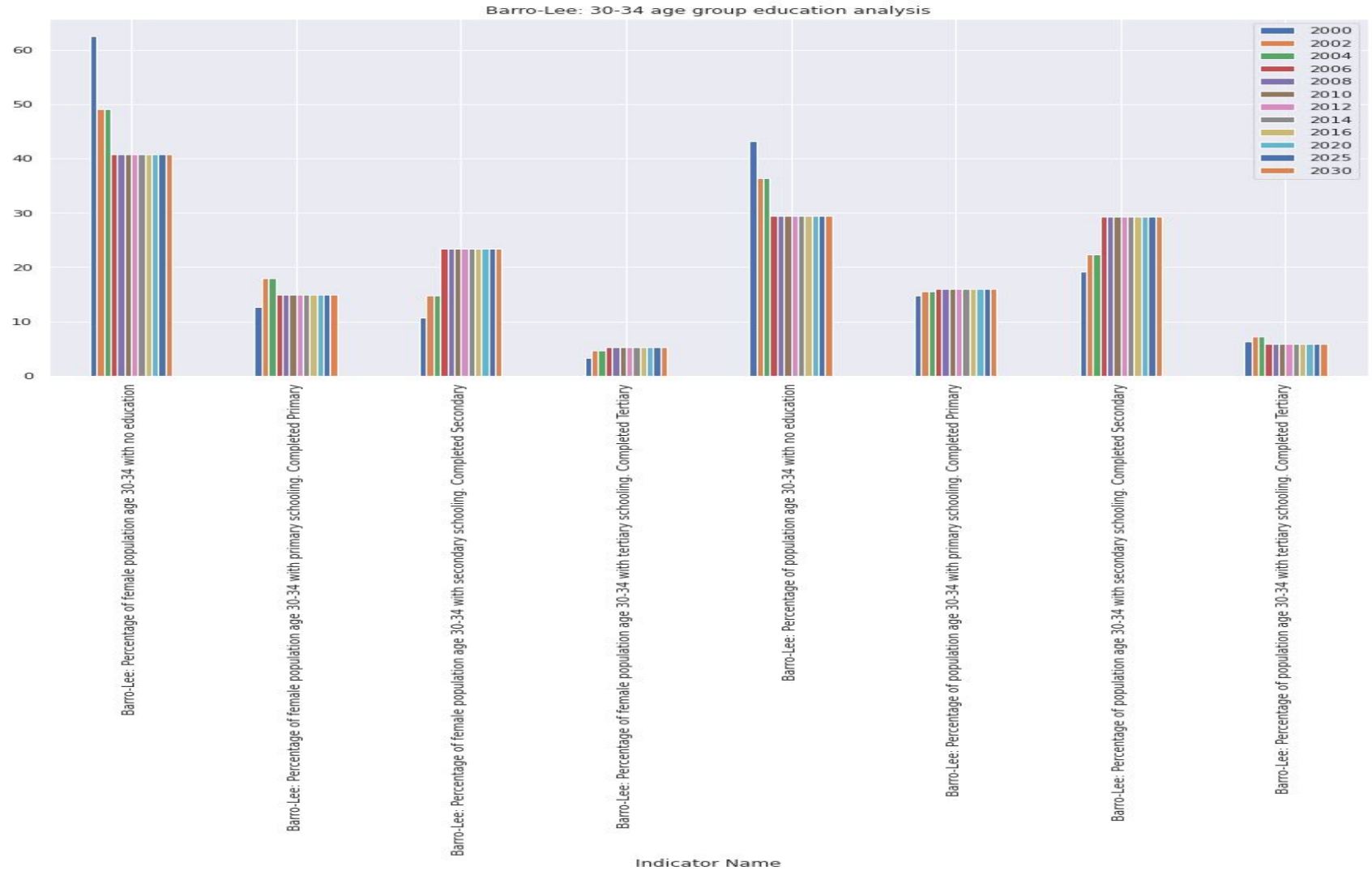
- Adjusted net enrolment rate calculated by dividing the number of children in the official primary school age who are enrolled in primary or secondary education by the population of the same age group and multiplying by 100.

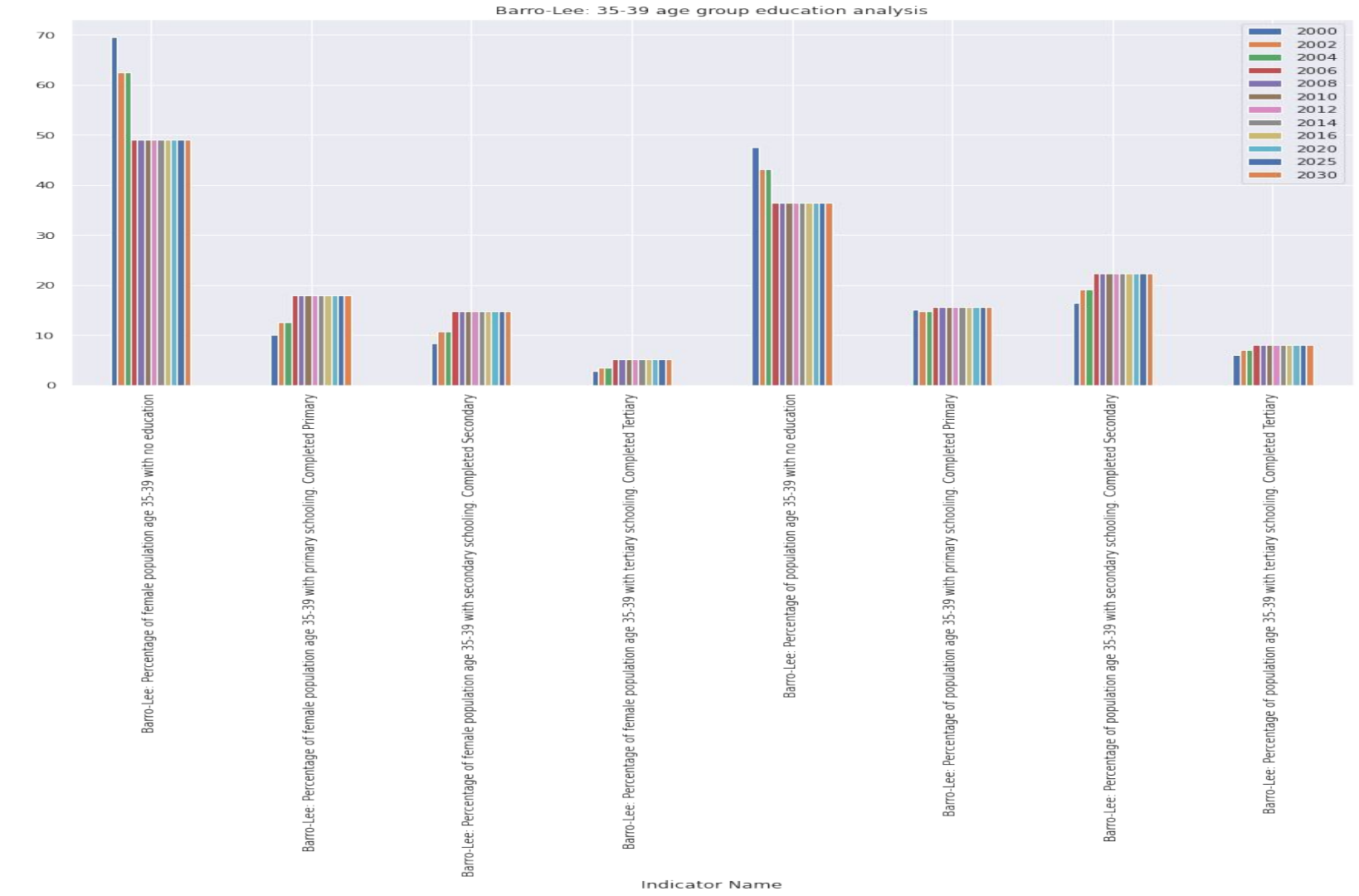


- ✔ Now, lets focus on the **Education of Young Generation (Age: 20 to 40)** for even years from the past two decades
- ✔ Barro-Lee: Age Group Education Analysis

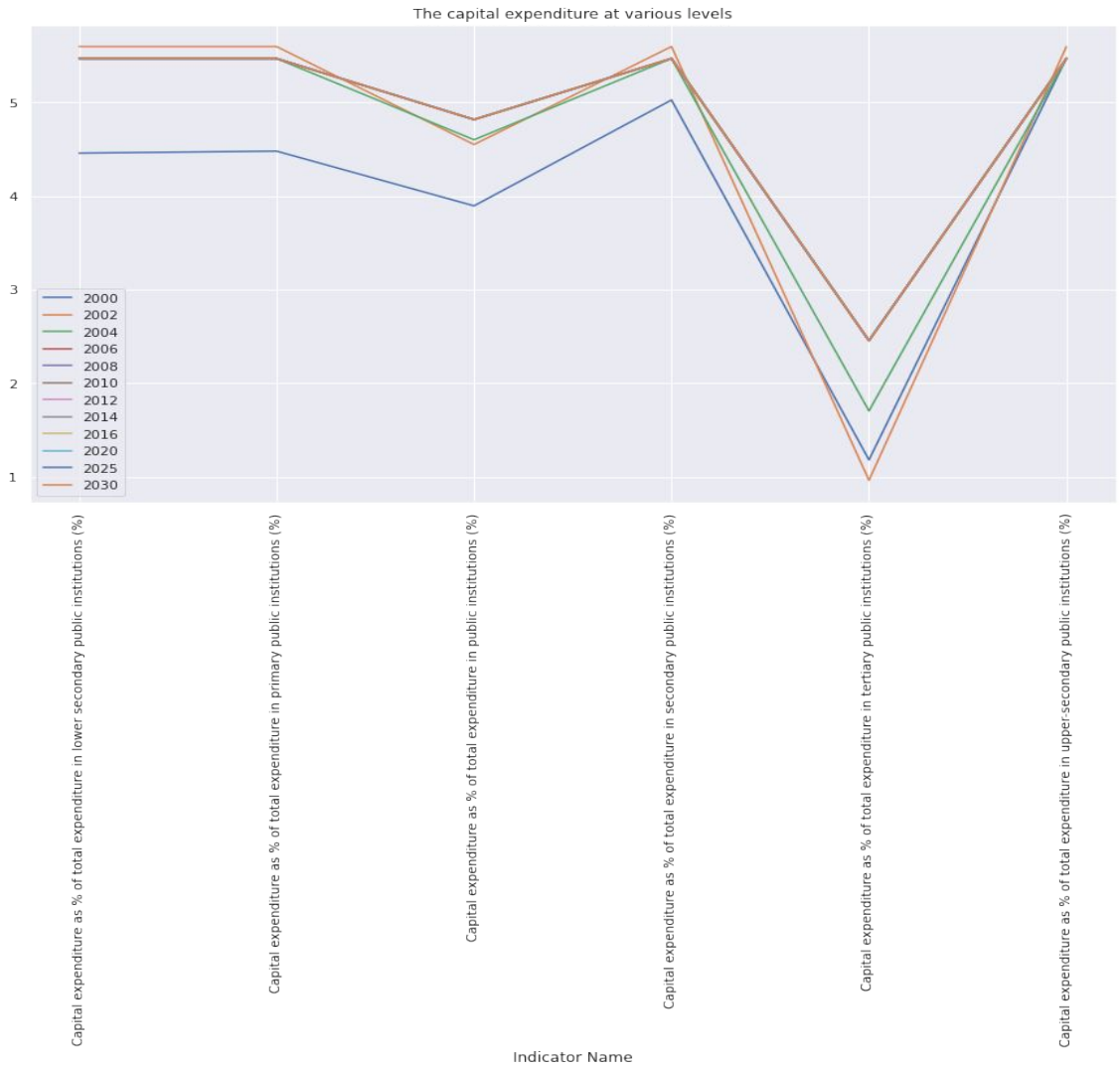

Barro-Lee: 20-24 age group education analysis
2000
2002
2004
2006
2008
2010
2012
2014
2016
2020
2025
2030
0
5
10
15
20
25
30
35
40
Barro-Lee: Percentage of female population age 20-24 with no education
Barro-Lee: Percentage of female population age 20-24 with primary schooling. Completed Primary
Barro-Lee: Percentage of female population age 20-24 with secondary schooling. Completed Secondary
Barro-Lee: Percentage of female population age 20-24 with tertiary schooling. Completed Tertiary
Barro-Lee: Percentage of population age 20-24 with no education
Barro-Lee: Percentage of population age 20-24 with primary schooling. Completed Primary
Barro-Lee: Percentage of population age 20-24 with secondary schooling. Completed Secondary
Barro-Lee: Percentage of population age 20-24 with tertiary schooling. Completed Tertiary
Indicator Name


Barro-Lee: 25-29 age group education analysis
50
40
30
20
10
0
2000
2002
2004
2006
2008
2010
2012
2014
2016
2020
2025
2030
Barro-Lee: Percentage of female population age 25-29 with no education
Barro-Lee: Percentage of female population age 25-29 with primary schooling. Completed Primary
Barro-Lee: Percentage of female population age 25-29 with secondary schooling. Completed Secondary
Barro-Lee: Percentage of female population age 25-29 with tertiary schooling. Completed Tertiary
Barro-Lee: Percentage of population age 25-29 with no education
Barro-Lee: Percentage of population age 25-29 with primary schooling. Completed Primary
Barro-Lee: Percentage of population age 25-29 with secondary schooling. Completed Secondary
Barro-Lee: Percentage of population age 25-29 with tertiary schooling. Completed Tertiary
Indicator Name

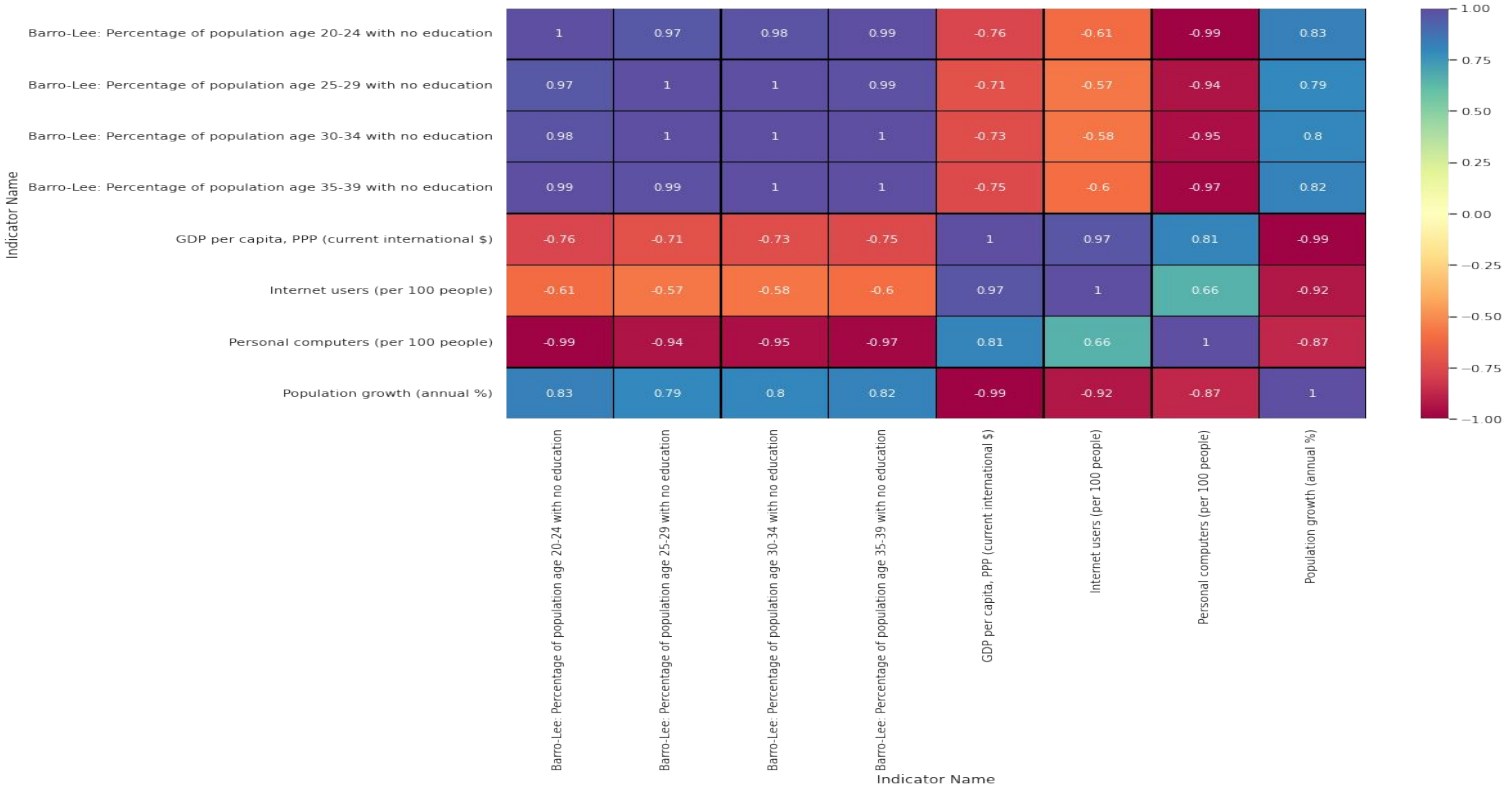

Barro-Lee: 30-34 age group education analysis
2000
2002
2004
2006
2008
2010
2012
2014
2016
2020
2025
2030
0
10
20
30
40
50
60
Barro-Lee: Percentage of female population age 30-34 with no education
Barro-Lee: Percentage of female population age 30-34 with primary schooling. Completed Primary
Barro-Lee: Percentage of female population age 30-34 with secondary schooling. Completed Secondary
Barro-Lee: Percentage of female population age 30-34 with tertiary schooling. Completed Tertiary
Barro-Lee: Percentage of population age 30-34 with no education
Barro-Lee: Percentage of population age 30-34 with primary schooling. Completed Primary
Barro-Lee: Percentage of population age 30-34 with secondary schooling. Completed Secondary
Barro-Lee: Percentage of population age 30-34 with tertiary schooling. Completed Tertiary
Indicator Name


Barro-Lee: 35-39 age group education analysis
70
60
50
40
30
20
10
0
2000
2002
2004
2006
2008
2010
2012
2014
2016
2020
2025
2030
Barro-Lee: Percentage of female population age 35-39 with no education
Barro-Lee: Percentage of female population age 35-39 with primary schooling Completed Primary
Barro-Lee: Percentage of female population age 35-39 with secondary schooling Completed Secondary
Barro-Lee: Percentage of female population age 35-39 with tertiary schooling Completed Tertiary
Barro-Lee: Percentage of population age 35-39 with no education
Barro-Lee: Percentage of population age 35-39 with primary schooling Completed Primary
Barro-Lee: Percentage of population age 35-39 with secondary schooling Completed Secondary
Barro-Lee: Percentage of population age 35-39 with tertiary schooling Completed Tertiary
Indicator Name

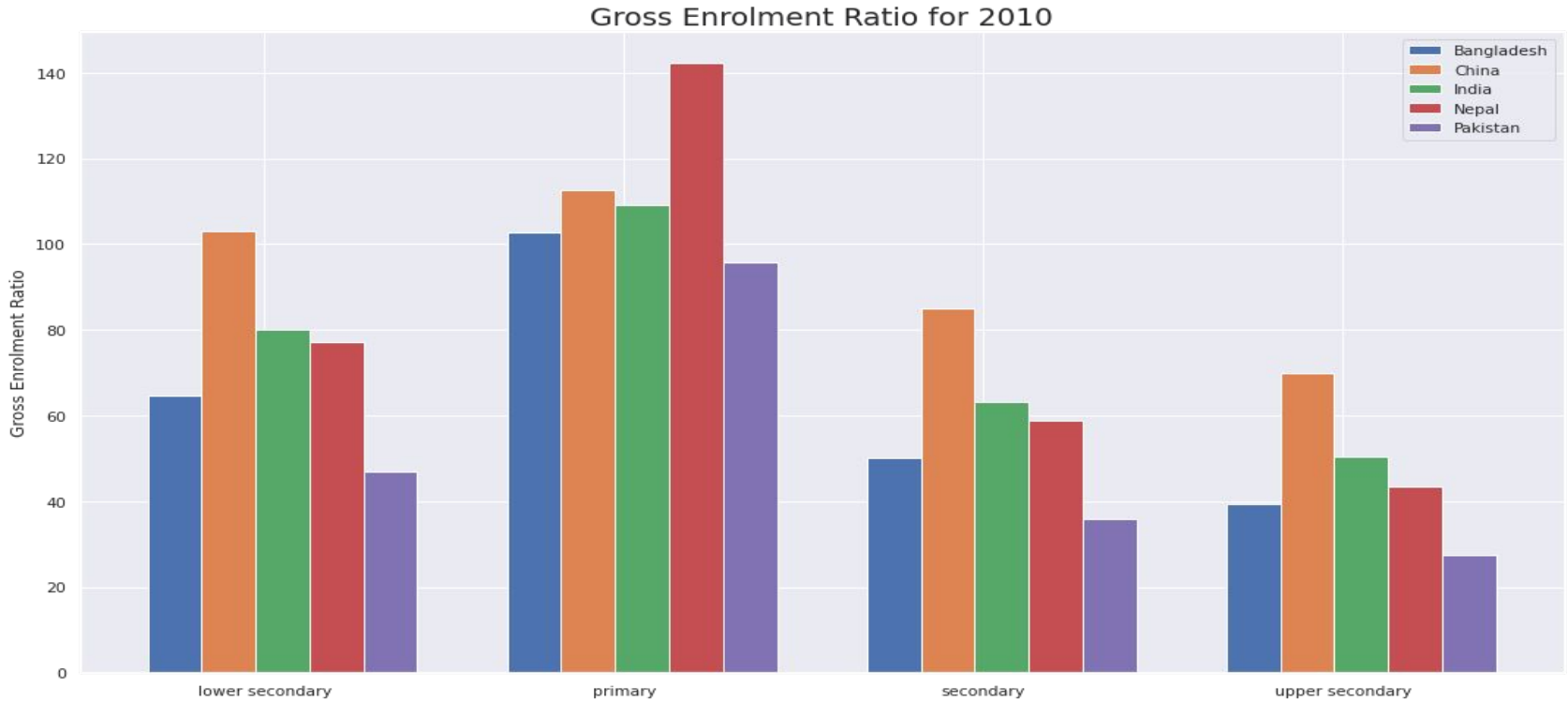
- GDP analysis is also done by observing the factors influencing it directly and indirectly.
- Expenditure analysis on these various categories is also performed to gather insights with respect to the peer countries. Capital Expenditure trends have also been compared in the proceeding sections . Analysis of Adjusted net enrolment rate calculated on the Barro-Lee Dataset is performed on the following age groups.
- Barro-Lee: 20-24 age group education analysis
- Barro-Lee: 25-29 age group education analysis
- Barro-Lee: 30-34 age group education analysis
- Barro-Lee: 35-39 age group education analysis.



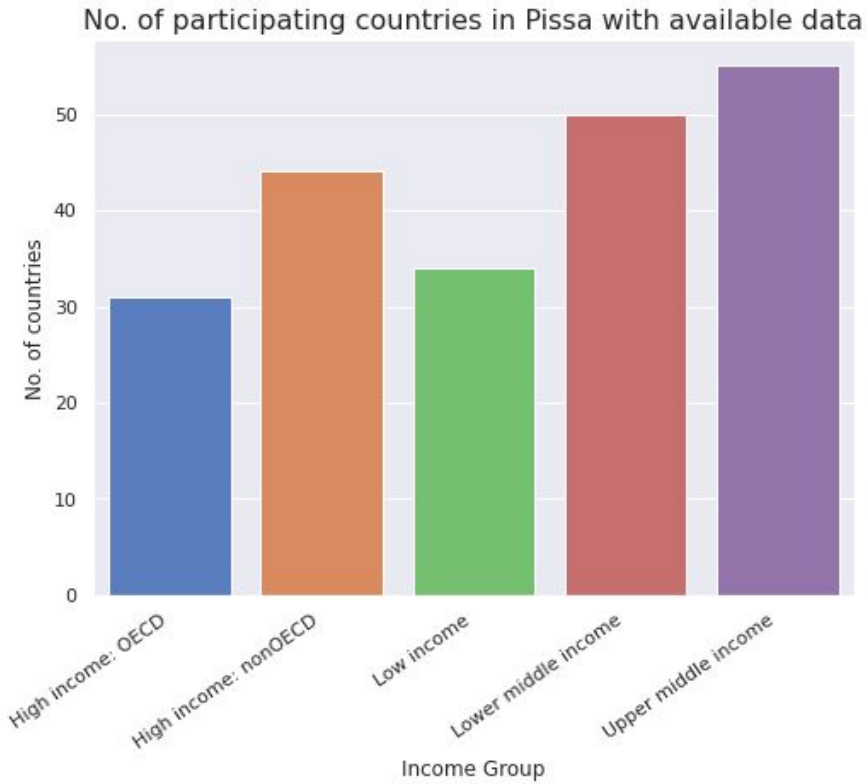
- Capital expenditure on public institutes at various levels show that it is very less for the tertiary section.



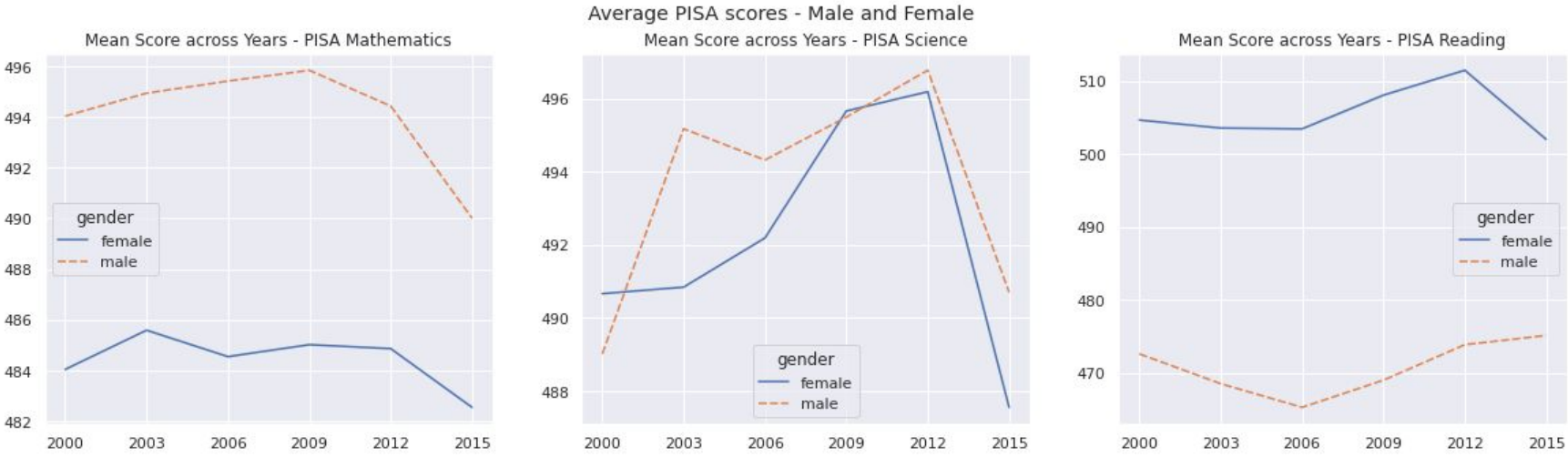
- Heatmap shows GDP is inversely proportional to ‘no education’ and population growth and its directly proportional to the technology.



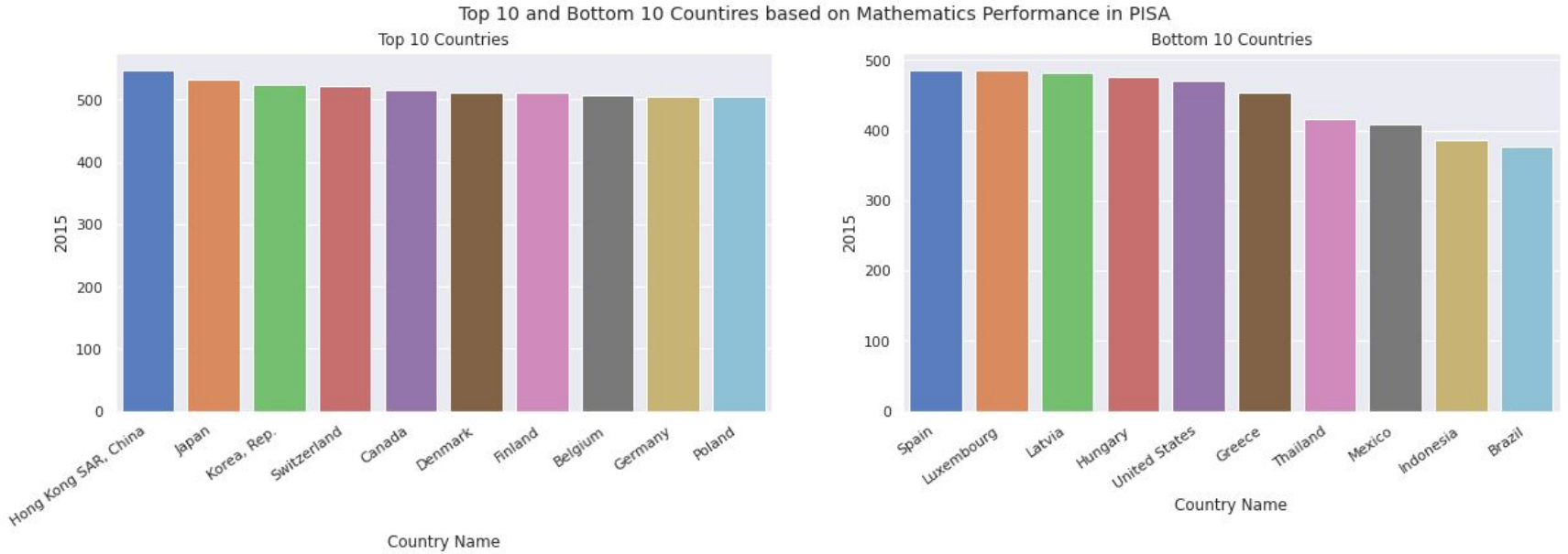
- Gross Enrollment Ratio for Bangladesh, China, Nepal, India and Pakistan has been plotted for 2010. It shows that all the countries, especially Nepal, are focusing on Primary Education.



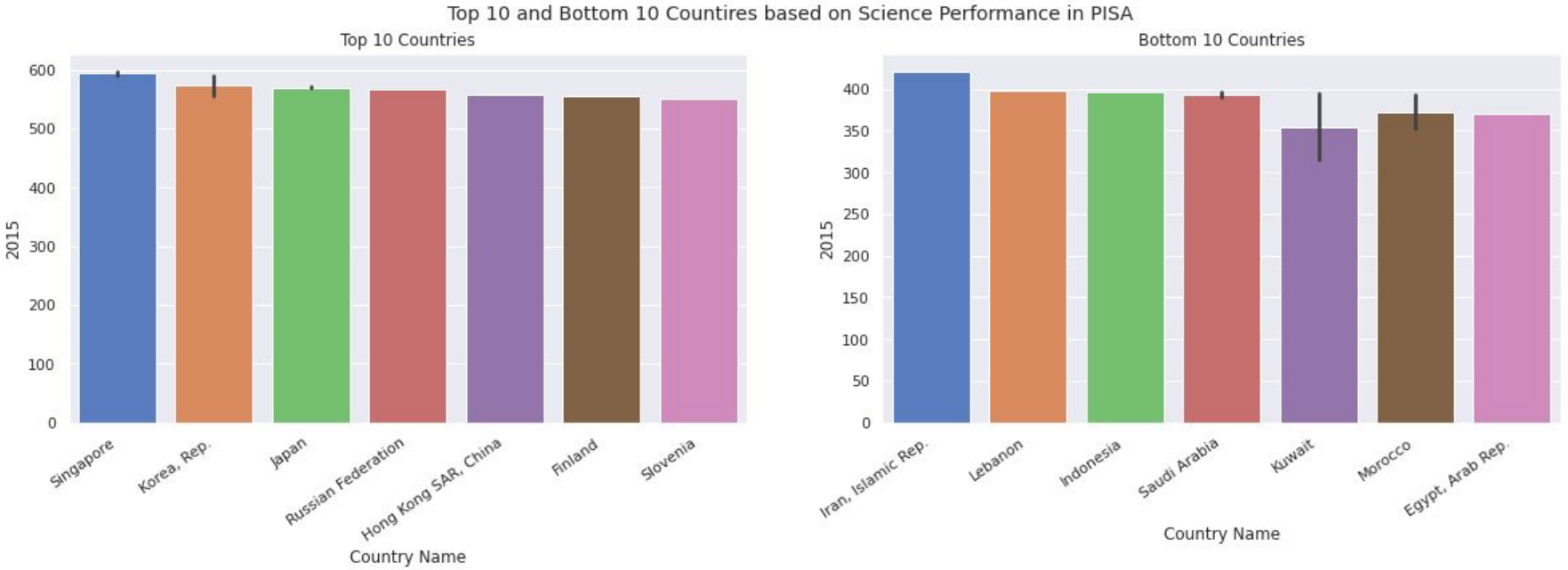
- The **Program for International Student Assessment** (PISA) is an international assessment that measures 15-year-old students' reading, mathematics, and science literacy every 3 years.



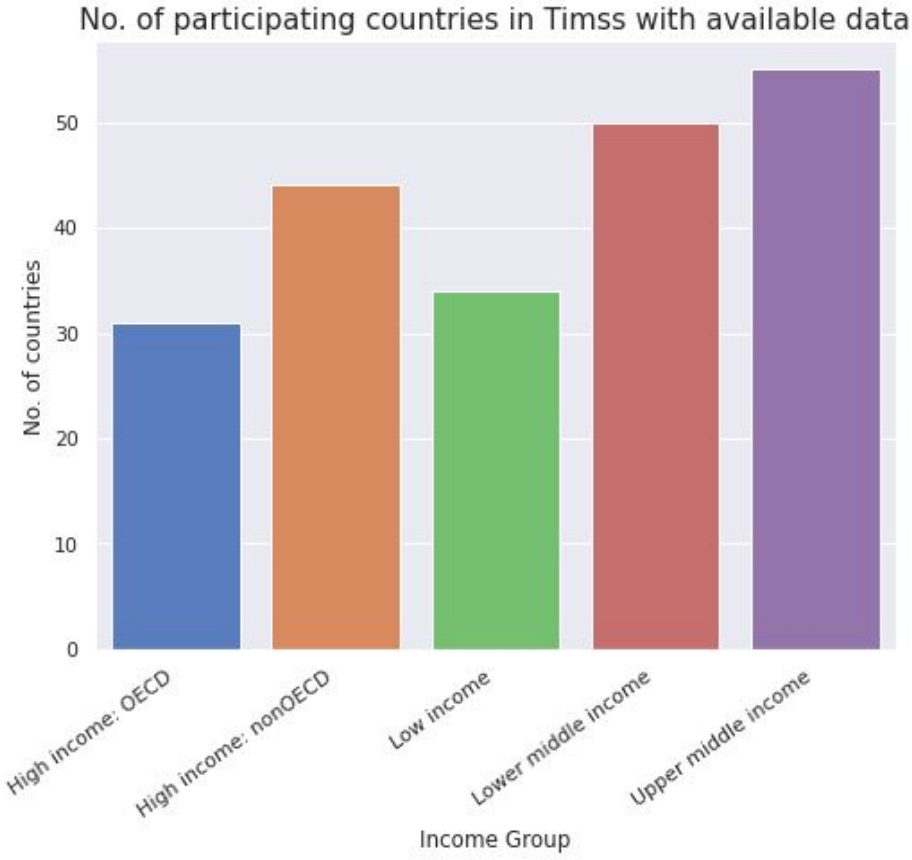
- PISA Mathematics: Male at higher side and female at lower side, shows the vast difference between them.
- PISA Science: Male and female performed equally same, but going lower in 2015.
- PISA Reading : Female at higher side and male at lower side, shows the vast difference between them.



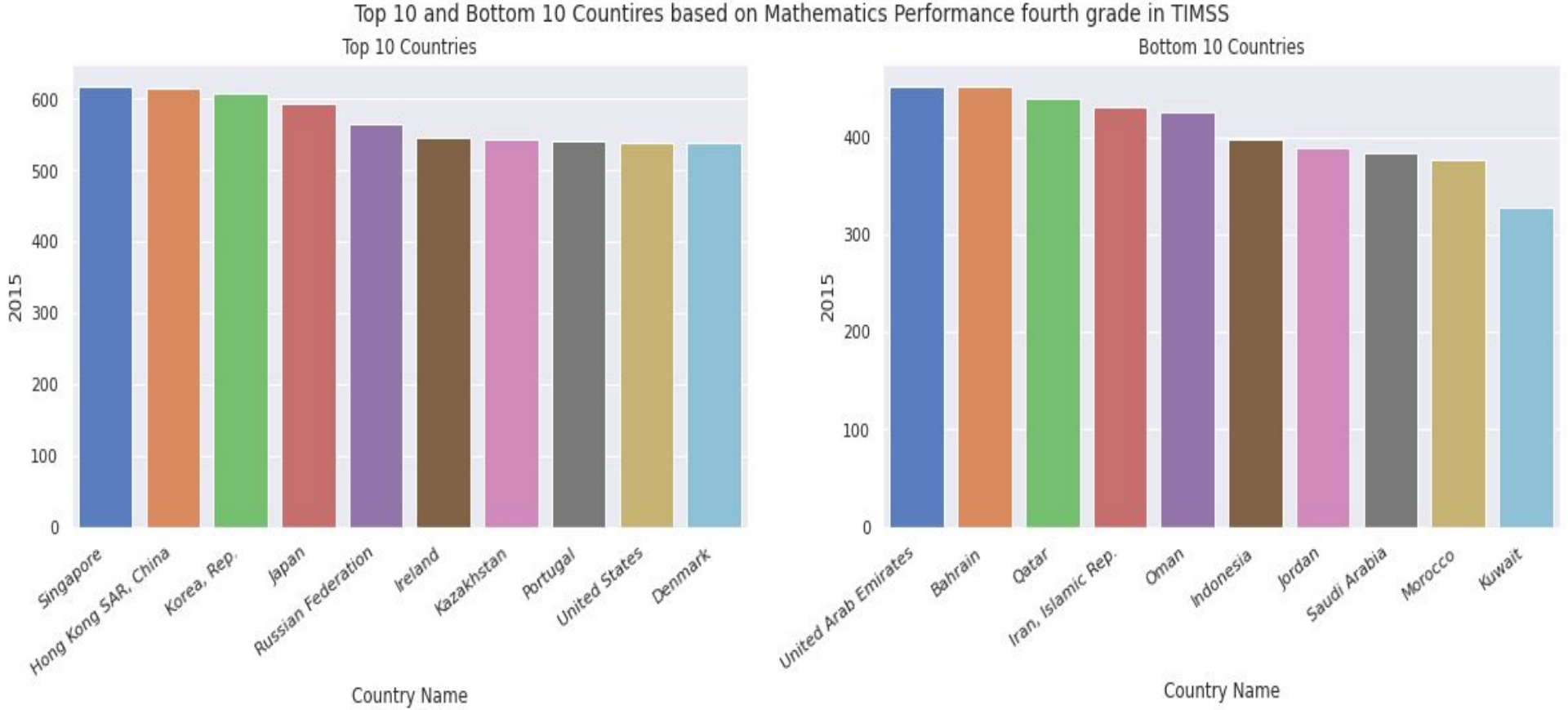
- **Mathematics Performance PISA Analysis** for top and bottom ten countries is done in Mathematics and Science. Hong Kong SAR, China is at the top and Brazil is at the bottom.

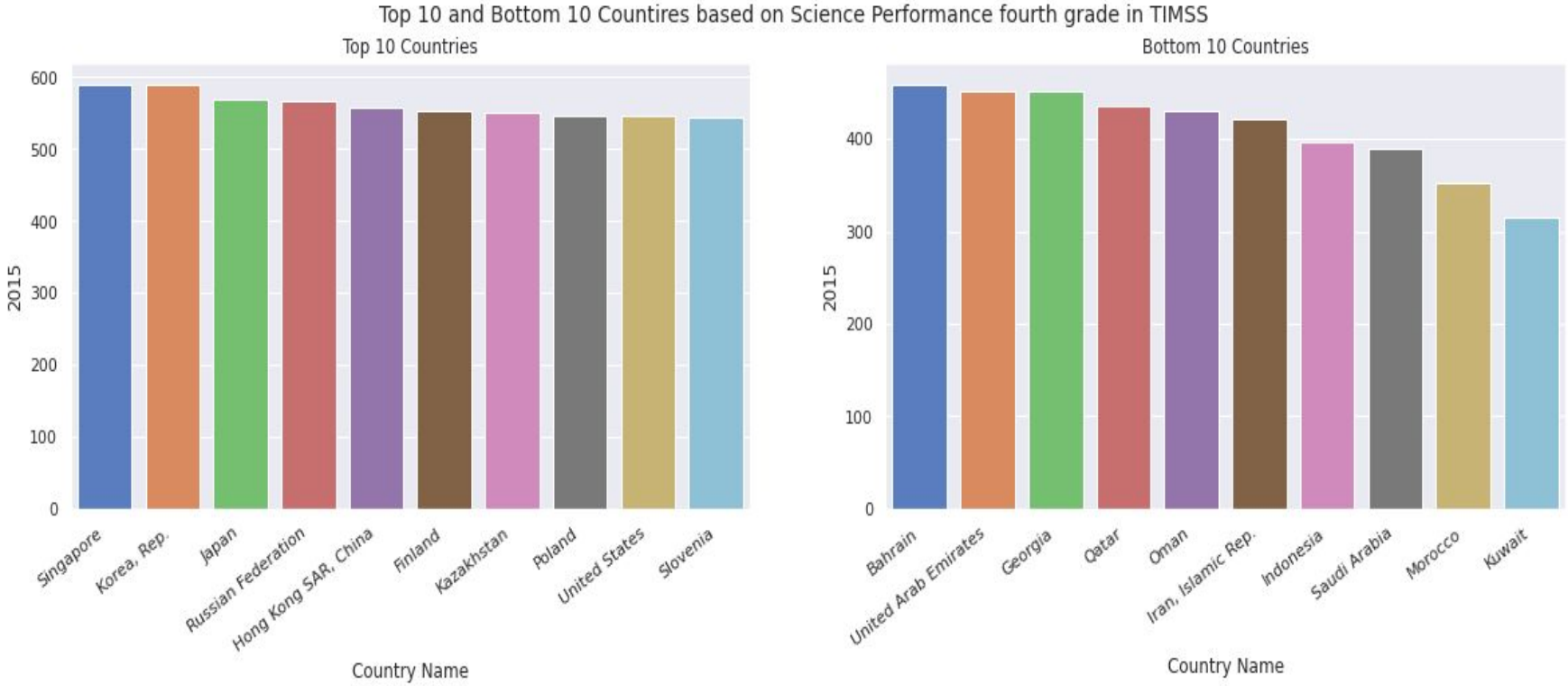


- **Science Performance PISA Analysis** for top and bottom ten countries is done in Mathematics and Science. Singapore is at the top and Egypt, Arab Rep. is at the bottom.



- **Trends in International Mathematics and Science Study** (TIMSS)


Top 10 and Bottom 10 Countires based on Mathematics Performance fourth grade in TIMSS
Top 10 Countries
2015
0
100
200
300
400
500
600
Singapore
Hong Kong SAR, China
Korea, Rep.
Japan
Russian Federation
Ireland
Kazakhstan
Portugal
United States
Denmark
Country Name
Bottom 10 Countries
2015
0
100
200
300
400
United Arab Emirates
Bahrain
Qatar
Iran, Islamic Rep.
Oman
Indonesia
Jordan
Saudi Arabia
Morocco
Kuwait
Country Name



- **TIMSS Analysis** for top and bottom ten countries is done in **Mathematics** and **Science**. Singapore is at the top and Kuwait is at the bottom.

Conclusion :

a) This study shows that most of the countries, whether developed and developing, are part of the World Bank, that's why the region with a high number of countries has high involvement.

b) The population growth analysis, based on income, indicates that the country's government of low, middle wages should implement some laws to control the population growth.

c) GDP bars show that highly developed countries have GDP exceeding 120000 and top developing countries have GDP exceeding 12000. This indicates a very vast difference in terms of economics of the country.

d) Education analysis of the youth shows that literacy rate is increasing over the years but still females are lacking behind compared to the total population literacy. Government and individuals should pay attention to education.

e) Literacy rate and technical advancement, directly influencing the GDP of the Country, should be increased.

f) PISA and TIMSS analysis show that it's quite different for male and females in the Mathematics and Reading field.

THANKS
FOR WATCHING