

# **Internship Project on**

**“World Bank Global Education Analysis”**

**(Exploratory Data Analysis)**

**AlmaBetter Edutech Private Limited**

**Bengaluru – 560025, India**

## **Presented by:**

**Ajay Thakur**

**Roll No. 2016kuec2026**

**Electronics and Communication Engineering  
Indian Institute of Information Technology Kota**

## **Submitted to:**

**AlmaBetter Edutech Private Limited**

**Bengaluru - 560025, India**

## **Guided by:**

**Mr. Vikash Srivastava**

**Co-founder & CPTO**

**AlmaBetter Edutech Private Limited  
Bengaluru - 560025, India**

## Points for discussion:

- Introduction
- Problem statement
- Objective of the project
- Data summary
- The number of missing values (NaN) of different dataframes
- The number of countries according to their regions and income groups
- The population rate of different countries in their income groups during 2005-2014
- The percentage ratio of developing and developed countries during 2000-2030 (based on GDP per capita)
- The GDP per capita of top 3 and bottom 3 developing countries (including India) and developed countries (including Australia) in 2020
- Adjusted Net Enrolment Rate (AER) of India in different educational levels during 2000-2030
- The population rate of India between different age groups in different educational levels during 2000-2030 (Based on Barro-Lee dataset)
- Capital Expenditure (CapEx) of India in different types of educational institutions during 2000-2030
- The correlation between different factors affecting economic growth of India during 2000-2030
- Gross Enrolment Ratio (GER) of certain countries in different educational levels in 2010
- The number of countries participating in the mean score of students in different scales of PISA according to their income groups during 2000-2015
- The mean score of students in different scales of PISA during 2000-2015
- The mean score of top 10 and bottom 10 countries in different scales of PISA in 2015
- The number of countries participating in the mean score of fourth grade students in different scales of TIMSS according to their income groups during 2000-2015
- The mean score of fourth grade students in different scales of TIMSS during 2000-2015
- The mean score of top 10 and bottom 10 countries in different scales of TIMSS in 2015
- Conclusion

## **Introduction:**

The World Bank Group is the largest financier of education in the developing world. The World Bank EdStats All Indicator Query holds over 4,000 internationally comparable indicators that describe education access, progression, completion, literacy, teachers, population, and expenditures. The indicators cover the education cycle from pre-primary to vocational and tertiary education and also holds learning outcome data from international and regional learning assessments (e.g. PISA, TIMSS, PIRLS), equity data from household surveys, and projection/attainment data.

In this project, We are going to explore and analyse the data to identify variation of indicators across the globe, which countries are more alike and different.

## **Problem statement:**

The data provided by the World Bank EdStats (Education Statistics) are in unformatted manner. The data is corrupted, duplicate and sometimes irrelevant because it's a piled-up data coming from different countries. For doing the analysis on the data, it should be in well organised format.

Various datasets that need to be cleaned are as follows – EdStatsCountry, EdStatsData, EdStatsSeries, EdStatsFootNote, EdStatsCountry\_Series

## **Objective of the project:**

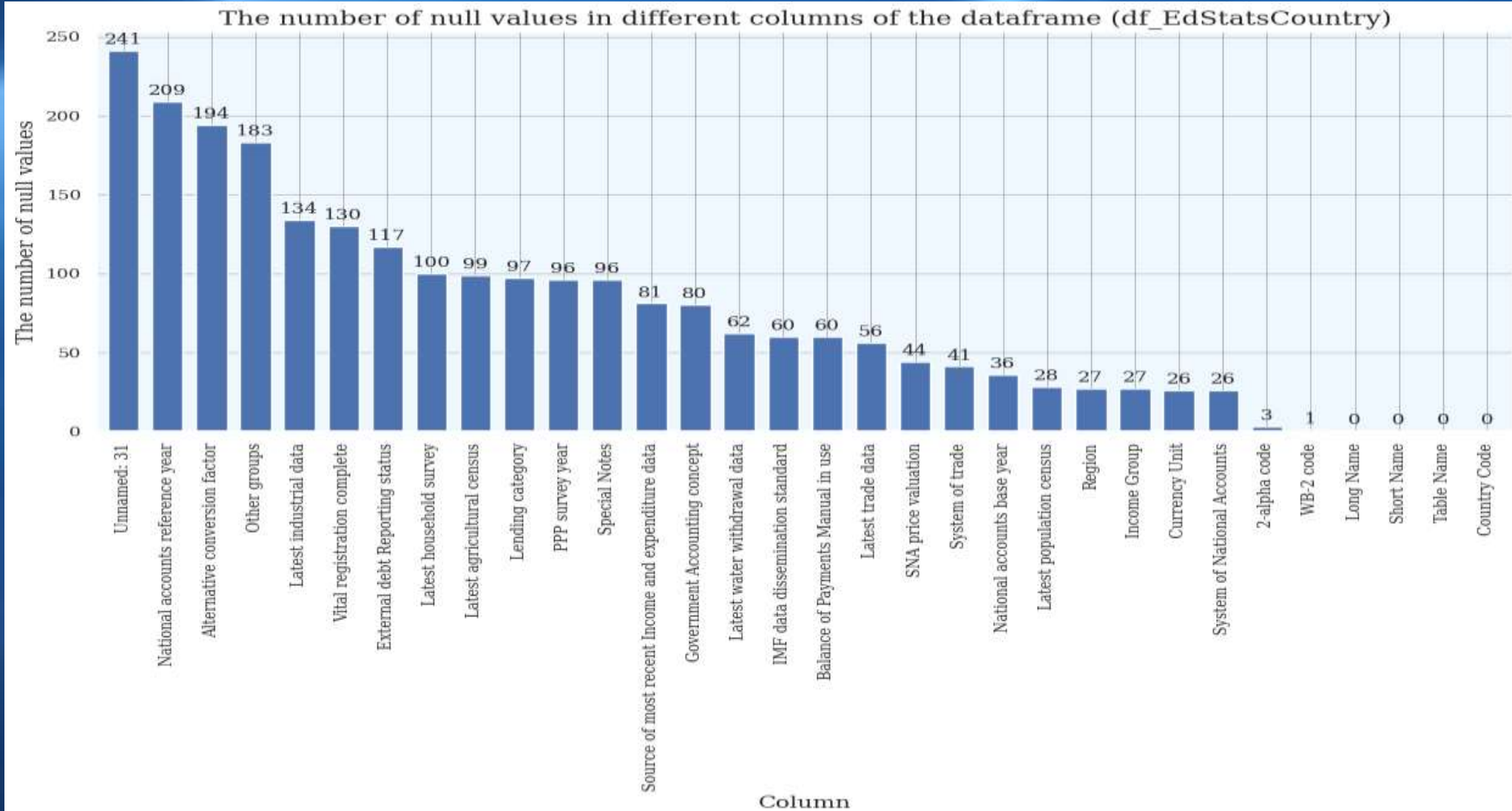
The main objective of the analysis is to obtain the meaning full information and facts from the given huge datasets as shown above, by cleaning the datasets, doing a proper analysis and visualization and plotting the useful information into different graph and charts so that the trend and relationship between the various indicators on which the analysis is done can be understand easily.

This will also help the World Bank group in taking the future decision based on the past events.

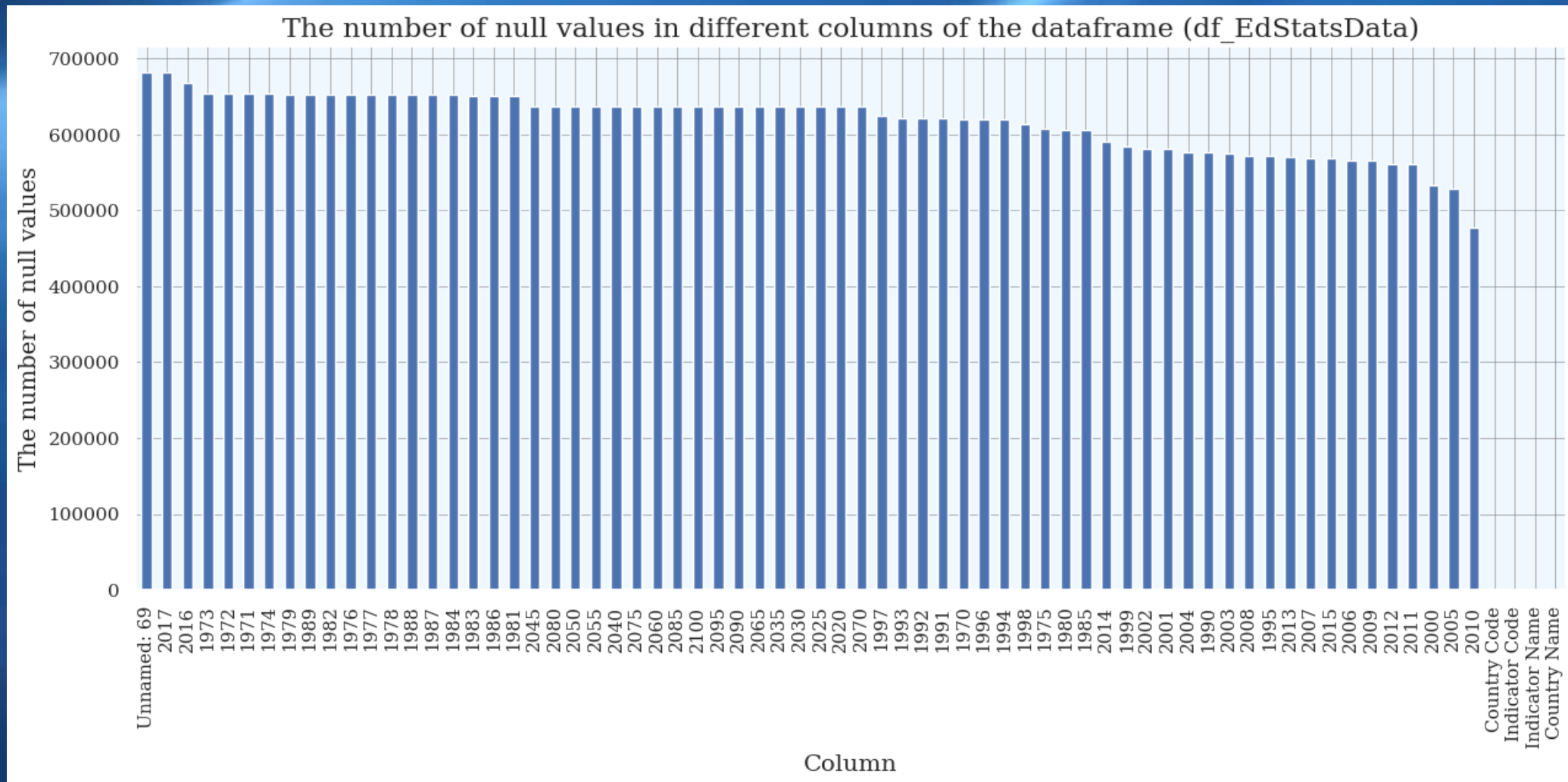
## Data summary: View all the datasets (one by one)

- **EdStatsCountry:** We observe that it contains list of all 241 countries that are present in the data. It has many columns such as - Country Code, Short Name, Table Name, Long Name, 2-alpha code, Special Notes, Region, Income Group, ..., Unnamed: 31.
- **EdStatsData:** We observe that it contains many columns such as - Country Name, Country Code, Indicator Name, Indicator Code, 1970, 1971, 1972, ..., 2060, after that 2065, 2070, 2075, ... (following same manner) till 2100, Unnamed: 69. Column 2010 has the most amount of data with 242442 (non - null) values, so we will prefer it as base year.
- **EdStatsSeries:** We observe that it contains many columns such as – Series Code, Topic, Indicator Name, Short definition, ..., Base Period, ..., Unnamed: 20. It contains list of all 3665 indicators.
- **EdStatsFootNote:** We observe that it contains just 5 columns – CountryCode, SeriesCode, Year, DESCRIPTION, Unnamed: 4. It contains the estimations and uncertainty bounds for each year - looks like some years are missing.
- **EdStatsCountry Series:** We observe that it contains just 4 columns - CountryCode, SeriesCode, DESCRIPTION, Unnamed: 3. It contains indicators and data sources for certain countries.

After analysing data of all the datasets, we have decided to work with two datasets – EdStatsCountry and EdStatsData.

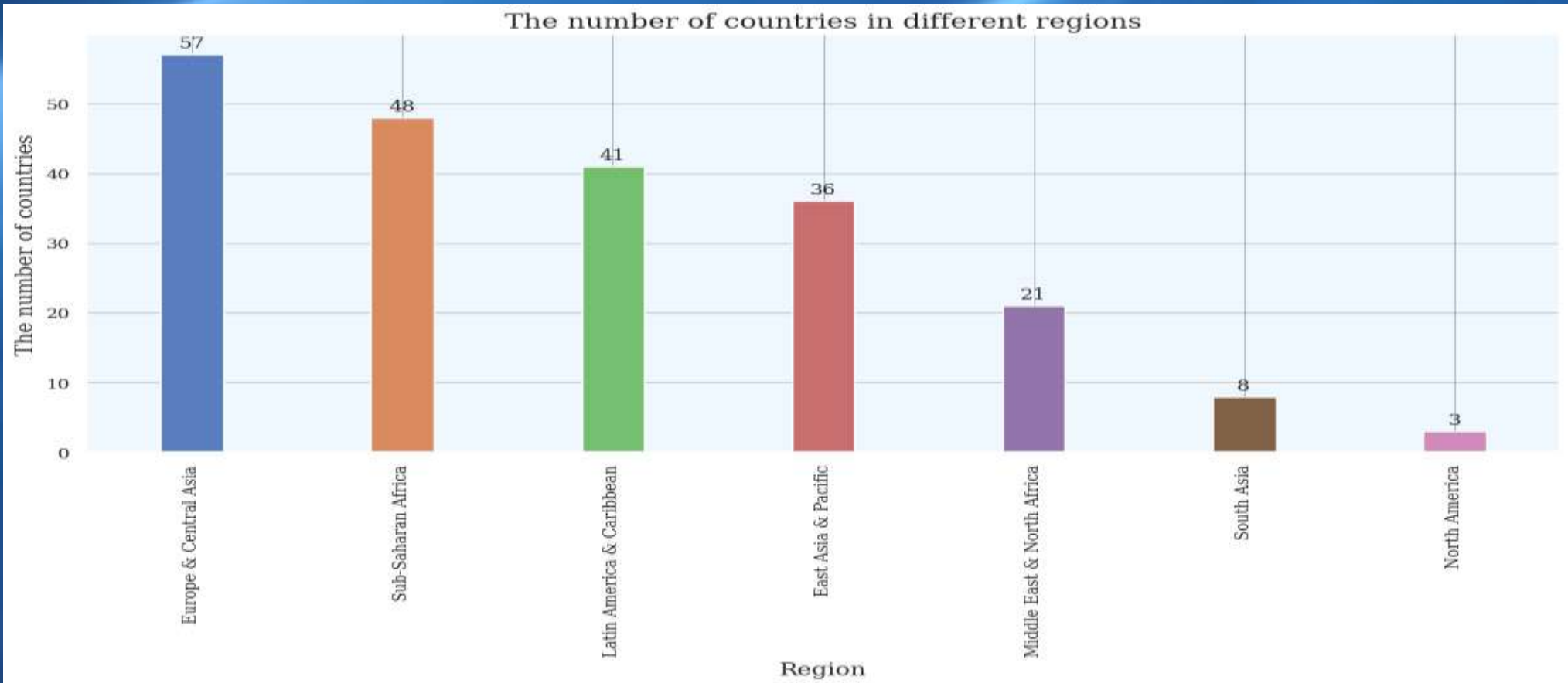


**Observation:** We observe that the columns of the dataframe (df\_EdStatsCountry) which have the maximum number of missing values (NaN) are – Unnamed: 31, National accounts reference year, Alternative conversion factor, Other groups, Latest industrial data, Vital registration complete and some others also.



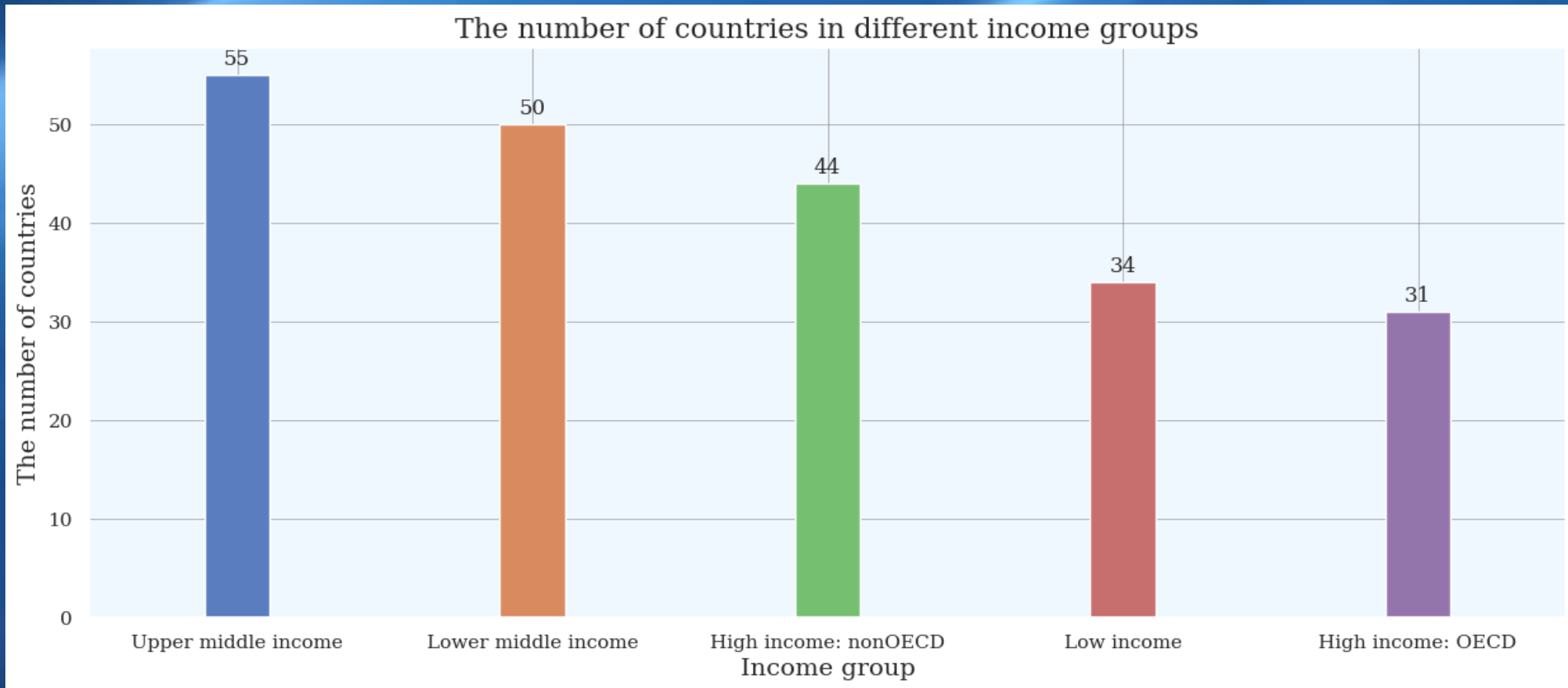
**Observation:** We observe that the columns of the dataframe (df\_EdStatsData) which have the maximum number of missing values (NaN) are - Unnamed: 69, 2017, 2016, 1971, 1973, 1972, ..., and some other year columns also. Most of the year columns are having missing values (NaN). We will fill these missing values by using bfill() and ffill() function.





**Observations:** The number of countries according to their regions

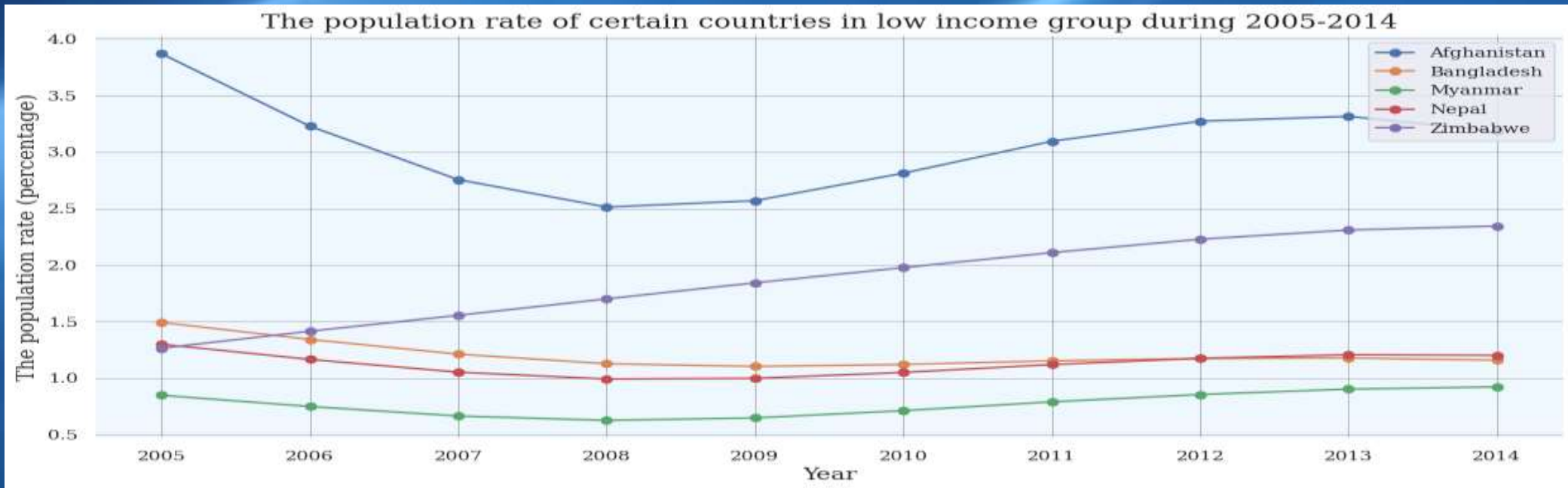
- We observe that Europe & Central Asia region has maximum countries and North America region has minimum countries.
- We observe that Sub-Saharan Africa region has 48 countries, Latin America & Caribbean region has 41 countries, East Asia & Pacific region has 36 countries, Middle East & North Africa region has 21 countries and South Asia region has 8 countries.



**Observations:** The number of countries according to their income groups

- We observe that upper middle income group has maximum countries and high income: OECD group has minimum countries.
- We observe that lower middle income group has 50 countries, high income: nonOECD group has 44 countries and low income group has 34 countries.



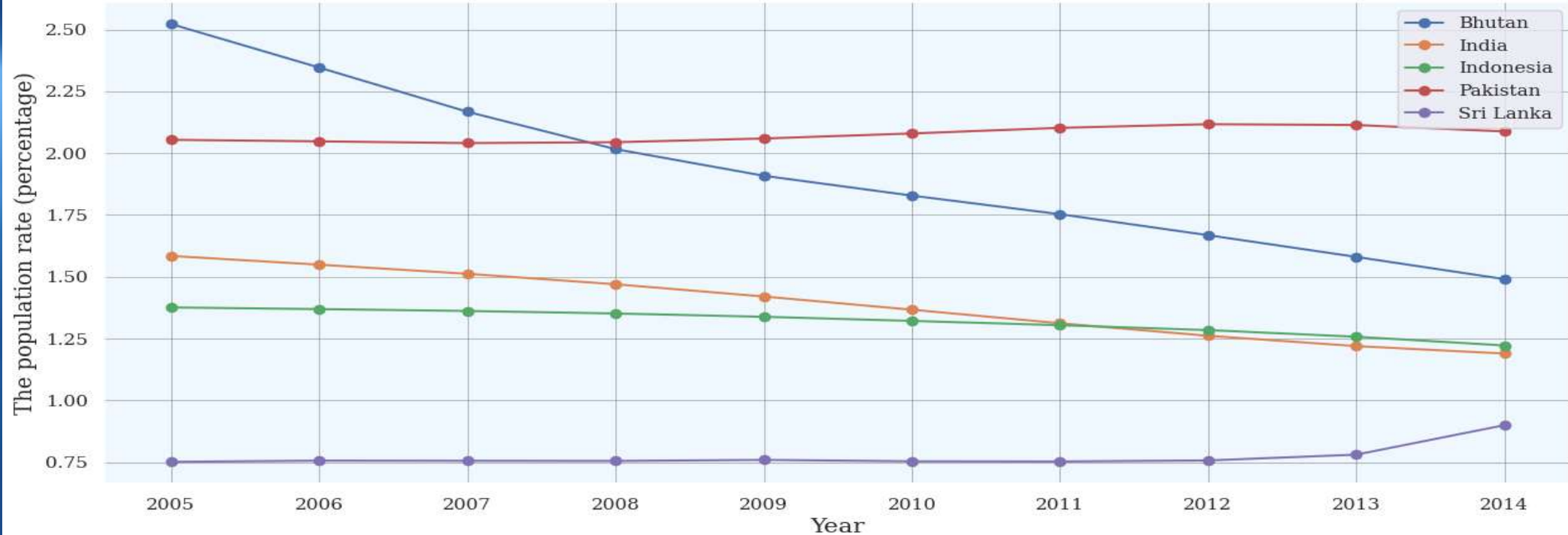


**Population Rate** - It is the rate of change of population size, for a given country, territory or geographic area during a specified period. It expresses the ratio between the annual increase in the population size and the total population for that year, usually multiplied by 100. The annual increase in the population size is defined as a sum of differences: the difference between births less deaths and the difference between immigrants less emigrants, in a given country, territory or geographic area at a given year.

**Observations:** The population rate of certain countries in low income group during 2005-2014

- **Afghanistan:** We observe a decline during 2005-2008 and incline during 2008-2012. It becomes constant during 2012-2013 and falls during 2013-2014.
- **Bangladesh:** We observe a decline during 2005-2008 and become constant during 2008-2010. It inclines during 2010-2011 and becomes stable during 2011-2014.
- **Myanmar:** We observe a decline during 2005-2006 and become constant during 2006-2007. It falls during 2007-2008 and becomes stable during 2008-2009. It inclines during 2009-2012 and again becomes constant during 2012-2014.
- **Nepal:** We observe a decline during 2005-2008 and become constant during 2008-2010. It inclines during 2010-2012 and becomes stable during 2012-2014.
- **Zimbabwe:** We observe an incline during 2005-2013 and become constant during 2013-2014.

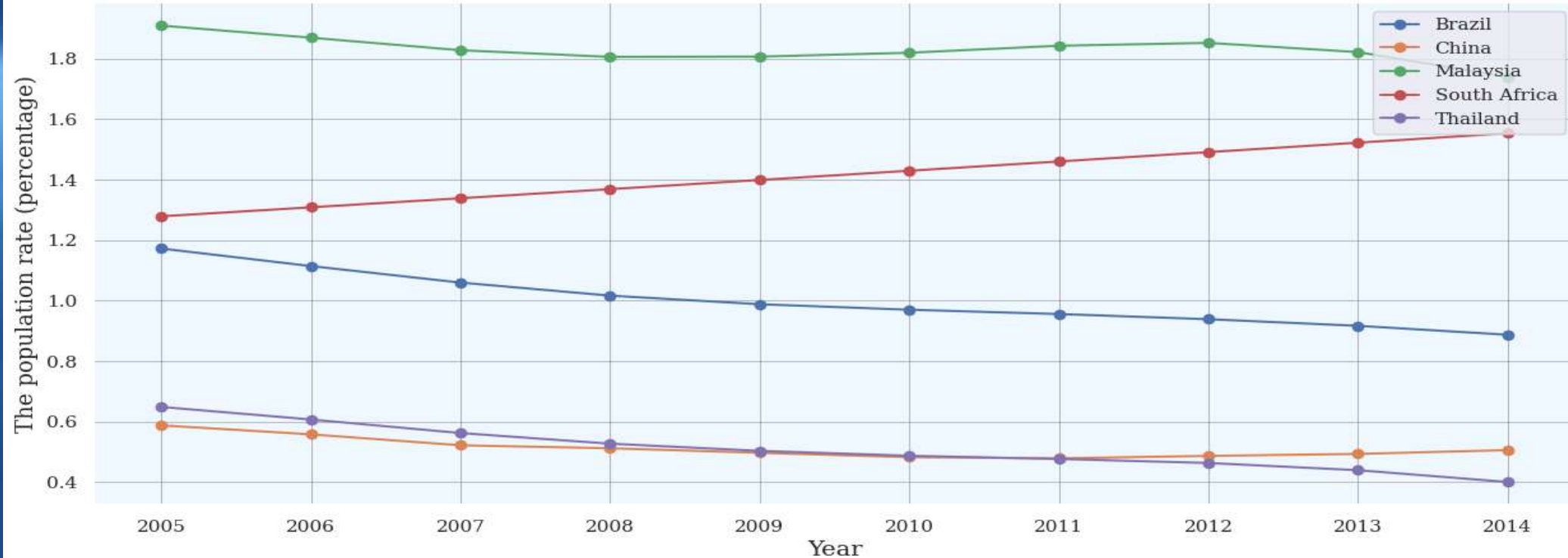
The population rate of certain countries in lower middle income group during 2005-2014



**Observations:** The population rate of certain countries in lower middle income group during 2005-2014

- **Bhutan:** We observe a decline during 2005-2010 and become constant during 2010-2011. It falls during 2011-2014.
- **India:** We observe a decline during 2005-2006 and become constant during 2006-2008. It falls during 2008-2009 and becomes stable during 2009-2010. It goes down during 2010-2011 and again becomes constant during 2011-2012. It decreases during 2012-2013 and again becomes stable during 2013-2014.
- **Indonesia:** We observe a constant trend during 2005-2008 and decline during 2008-2009. It becomes stable during 2009-2013 and falls during 2013-2014.
- **Pakistan:** We observe a decline during 2005-2006 and become constant during 2006-2008. It inclines during 2008-2009 and becomes stable during 2009-2014.
- **Sri Lanka:** We observe a constant trend during 2005-2013 and incline during 2013-2014.

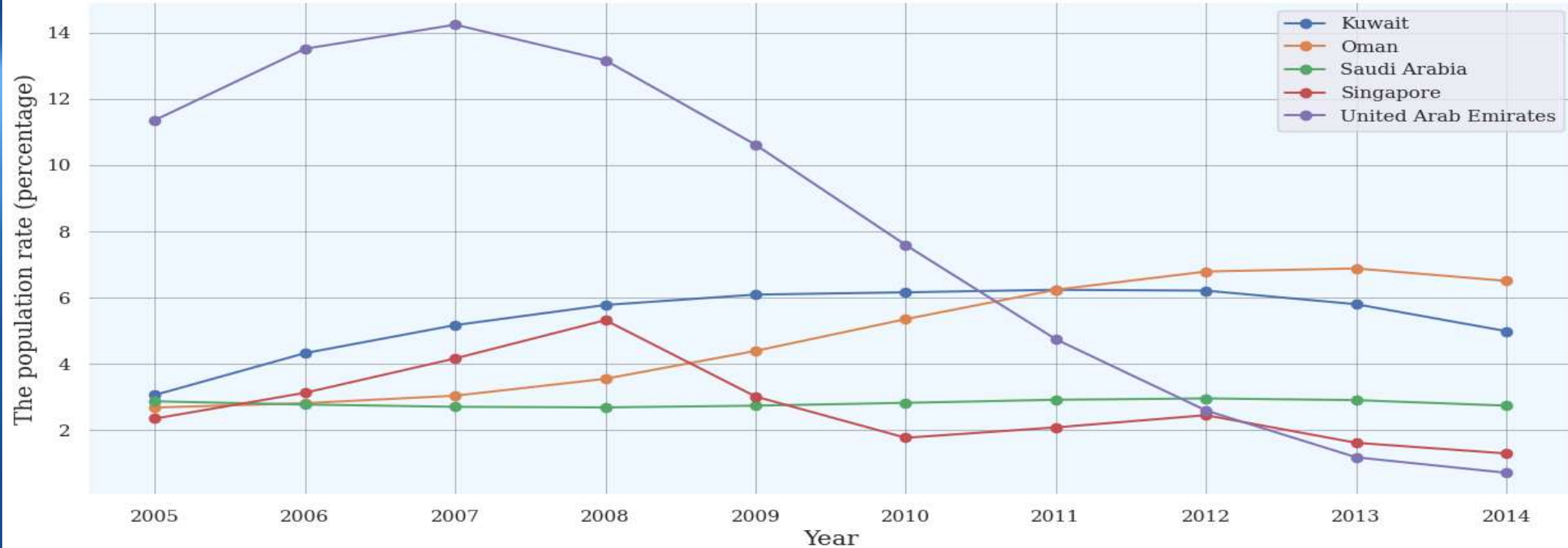
The population rate of certain countries in upper middle income group during 2005-2014



**Observations:** The population rate of certain countries in upper middle income group during 2005-2014

- **Brazil:** We observe a decline during 2005-2006 and becomes constant during 2006-2007. It falls during 2007-2008 and becomes stable during 2008-2011. It goes down during 2011-2012 and again becomes constant during 2012-2014.
- **China:** We observe a constant trend during 2005-2006 and decline during 2006-2007. It becomes stable during 2007-2014.
- **Malaysia:** We observe a constant trend during 2005-2006 and decline during 2006-2007. It becomes stable during 2007-2011 and inclines during 2011-2012. It falls during 2012-2014.
- **South Africa:** We observe a constant trend during 2005-2007 and incline during 2007-2008. It becomes stable during 2008-2010 and climbs up during 2010-2011. Once again it becomes stable during 2011-2013 and goes up during 2013-2014.
- **Thailand:** We observe a constant trend during 2005-2007 and decline during 2007-2008. It becomes stable during 2008-2012 and falls during 2012-2013. Once again it becomes constant during 2013-2014.

The population rate of certain countries in high income: nonOECD group during 2005-2014

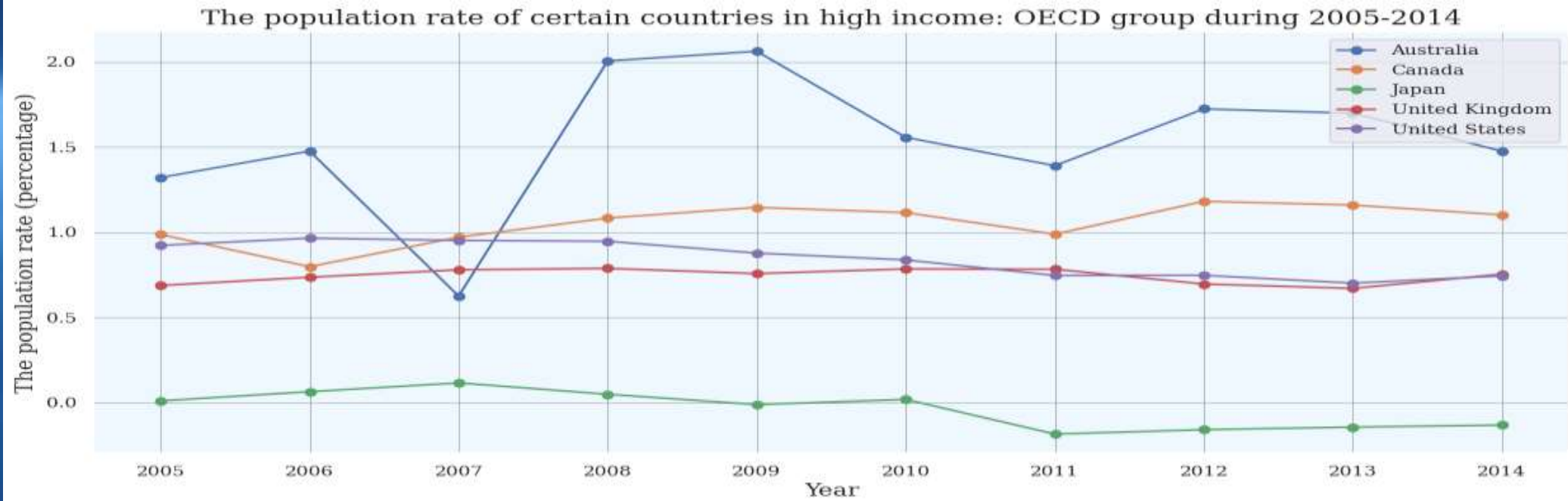


**High Income: nonOECD Group** - The income group, which belong to the countries that have high income but not a member of the Organisation of Economic Co-operation and Development according to the World Bank, such a group is called High Income: nonOECD Group.

**Observations:** The population rate of certain countries in high income: nonOECD group during 2005-2014

- **Kuwait:** We observe an incline during 2005-2010 and become constant during 2010-2012. It declines during 2012-2014.
- **Oman:** We observe an incline during 2005-2013 and decline during 2013-2014.
- **Saudi Arabia:** We observe a decline during 2005-2007 and become constant during 2007-2009. It inclines during 2009-2012 and falls during 2012-2014.
- **Singapore:** We observe an incline during 2005-2008 and decline during 2008-2010. It goes up during 2010-2012 and falls during 2012-2014.
- **United Arab Emirates:** We observe an incline during 2005-2007 and decline during 2007-2014.

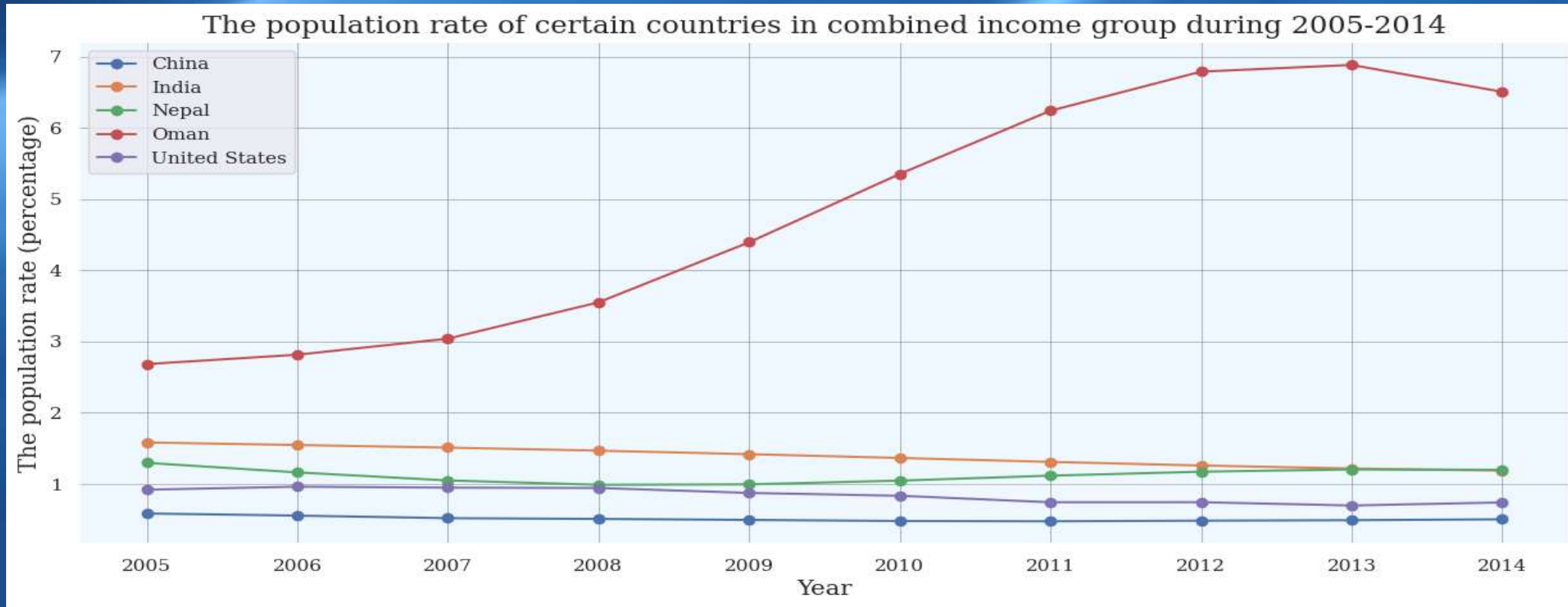




**High Income: OECD Group** - The income group, which belong to the countries that are both high income and members of the Organisation of Economic Co-operation and Development according to the World Bank, such a group is called High Income: OECD Group.

**Observations:** The population rate of certain countries in high income: OECD group during 2005-2014

- **Australia:** We observe an incline during 2005-2006 and decline during 2006-2007. It climbs up during 2007-2009 and falls during 2009-2011. It increases during 2011-2012 and becomes constant during 2012-2013. It goes down during 2013-2014.
- **Canada:** We observe a decline during 2005-2006 and incline during 2006-2008. It becomes constant during 2008-2010 and falls during 2010-2011. It goes up during 2011-2012 and becomes stable during 2012-2013. It goes down during 2013-2014.
- **Japan:** We observe an incline during 2005-2006 and become constant during 2006-2007. It declines during 2007-2011 and becomes stable during 2011-2012. It goes up during 2012-2013 and again becomes constant during 2013-2014.
- **United Kingdom:** We observe a constant trend during 2005-2006 and incline during 2006-2007. It becomes stable during 2007-2011 and declines during 2011-2012. Once again it becomes constant during 2012-2013 and goes up during 2013-2014.
- **United States:** We observe an incline during 2005-2006 and becomes constant during 2006-2007. It declines during 2007-2008 and becomes stable during 2008-2009. It falls during 2009-2011 and again becomes constant during 2011-2014.

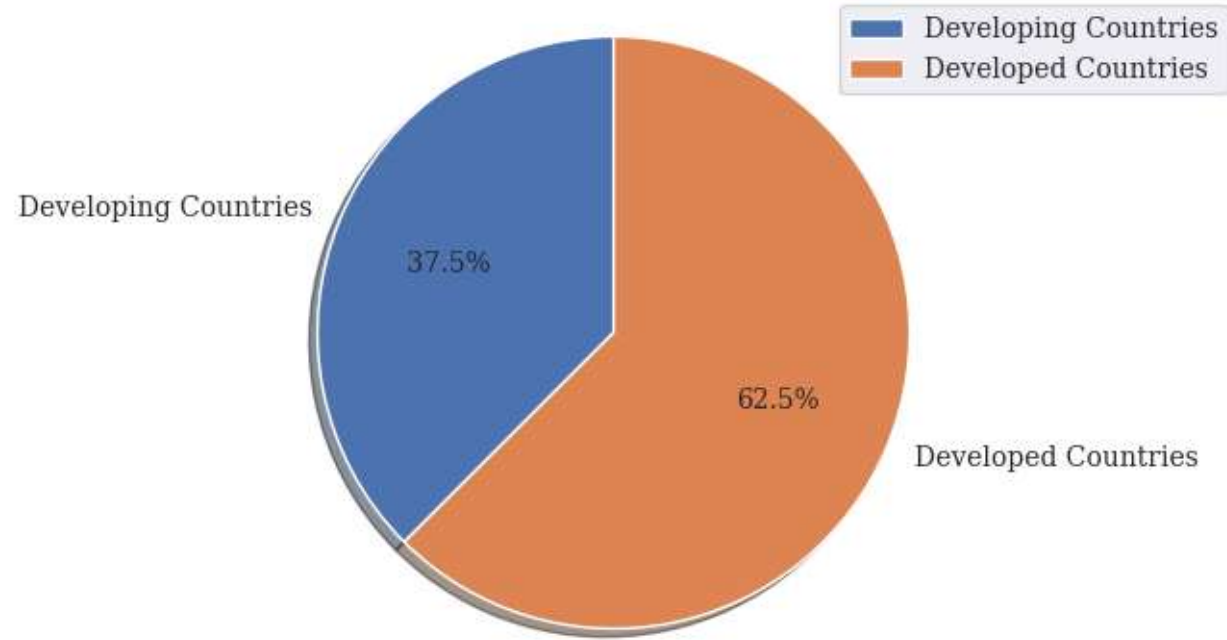


**Observations:** The population rate of certain countries in combined income group during 2005-2014

- **China:** We observe a constant trend during 2005-2006 and decline during 2006-2007. It becomes stable during 2007-2014.
- **India:** We observe a decline during 2005-2006 and become constant during 2006-2008. It falls during 2008-2009 and becomes stable during 2009-2010. It goes down during 2010-2012 and again becomes constant during 2012-2014.
- **Nepal:** We observe a decline during 2005-2008 and become constant during 2008-2010. It inclines during 2010-2012 and becomes stable during 2012-2014.
- **Oman:** We observe an incline during 2005-2013 and decline during 2013-2014.
- **United States:** We observe an incline during 2005-2006 and become constant during 2006-2007. It declines during 2007-2008 and becomes stable during 2008-2009. It falls during 2009-2011 and again becomes constant during 2011-2014.

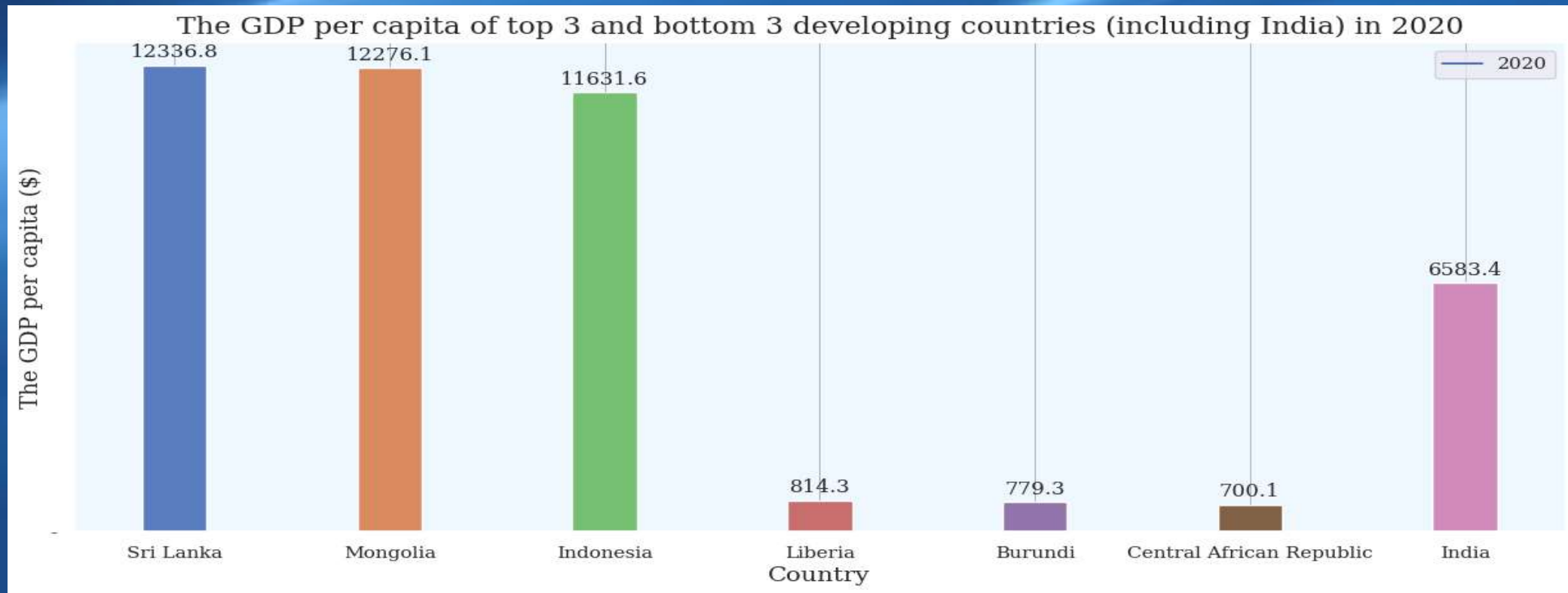


The percentage ratio of developing and developed countries



**Observations:** The percentage ratio of developing and developed countries during 2000-2030 (based on GDP per capita)

- The percentage ratio of developing countries is 37.5%
- The percentage ratio of developed countries is 62.5%
- The percentage ratio of developing countries is less than the percentage ratio of developed countries.



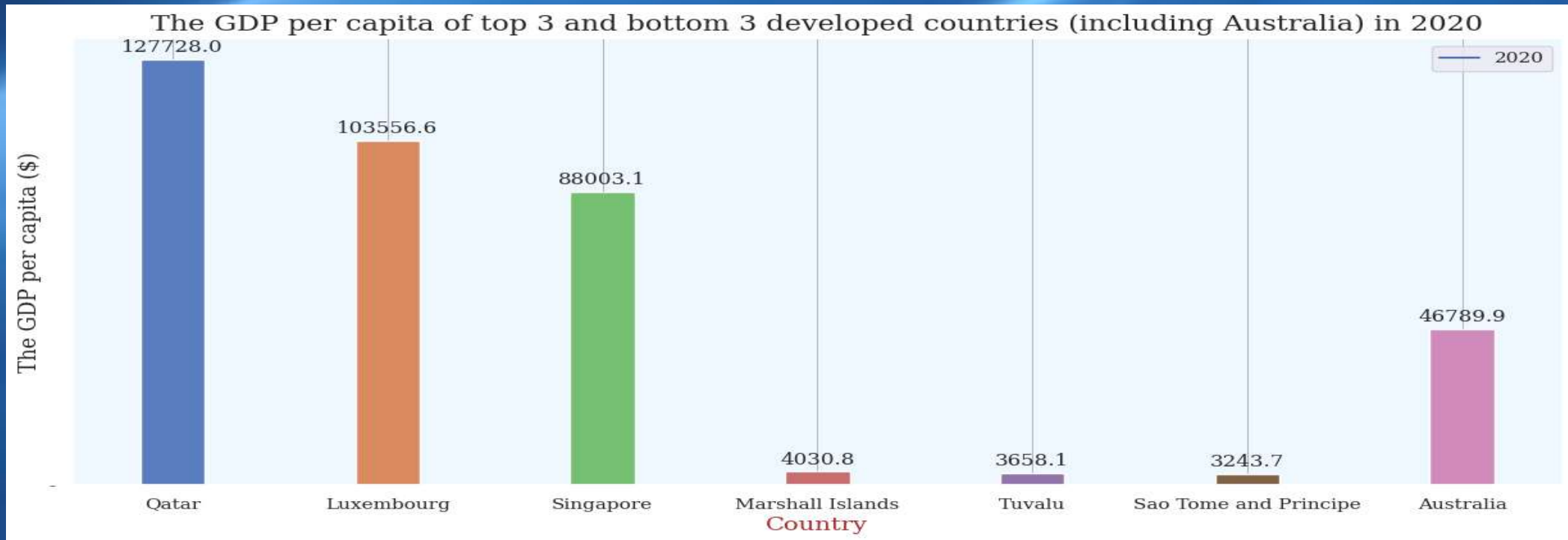
**Observations:** The GDP per capita of top 3 and bottom 3 developing countries (including India) in 2020

**GDP per capita of top 3 developing countries**

- We observe that Sri Lanka has the highest GDP per capita and Indonesia has the lowest GDP per capita in top 3 developing countries in 2020.
- We observe that out of top 3 developing countries, all 3 are Asian countries.

**GDP per capita of bottom 3 developing countries**

- We observe that Liberia has the highest GDP per capita and Central African Republic has the lowest GDP per capita in bottom 3 developing countries (including India) in 2020.
- We observe that out of bottom 3 developing countries (including India), all 3 are African countries and India is the only Asian country.



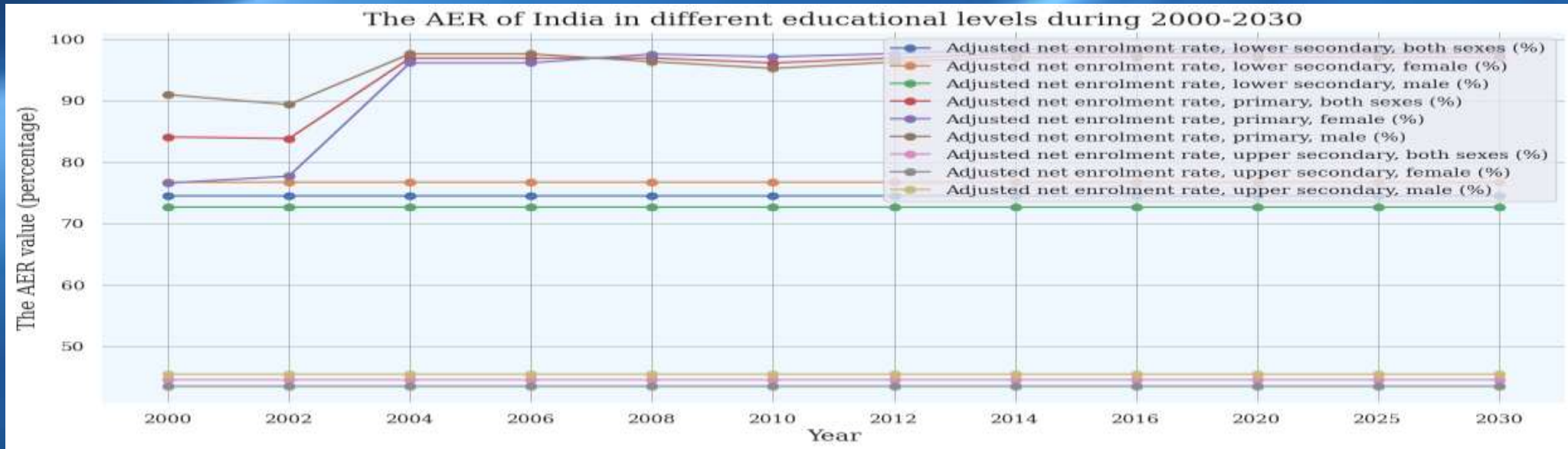
**Observations:** The GDP per capita of top 3 and bottom 3 developed countries (including Australia) in 2020

#### **GDP per capita of top 3 developed countries**

- We observe that Qatar has the highest GDP per capita and Singapore has the lowest GDP per capita in top 3 developed countries in 2020.
- We observe that out of top 3 countries, Qatar is Middle Eastern country, Luxembourg is European country and Singapore is the only Asian country.

#### **GDP per capita of bottom 3 developed countries**

- We observe that Marshall Islands has the highest GDP per capita and Sao Tome and Principe has the lowest GDP per capita in bottom 3 developed countries (including Australia) in 2020.
- We observe that out of bottom 3 countries (including Australia), 3 are Asian countries and Sao Tome and Principe is the only African country.

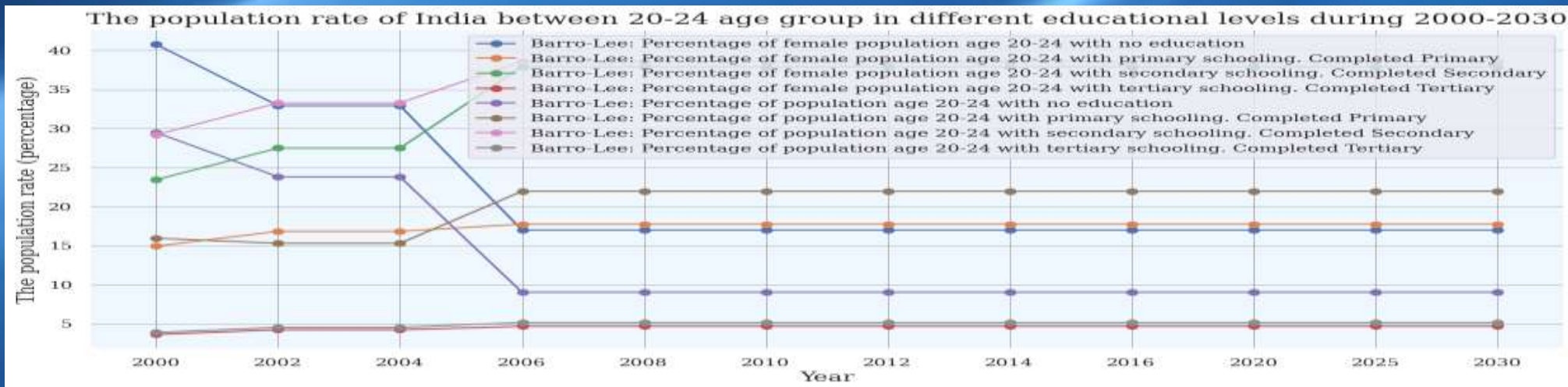


**Adjusted Net Enrolment Rate (AER)** - Adjusted net enrolment rate in primary education is calculated by dividing the number of children in the official primary school age who are enrolled in primary or secondary education by the population of the same age group and multiplying by 100.

**Observations:** The AER of India in different educational levels during 2000-2030

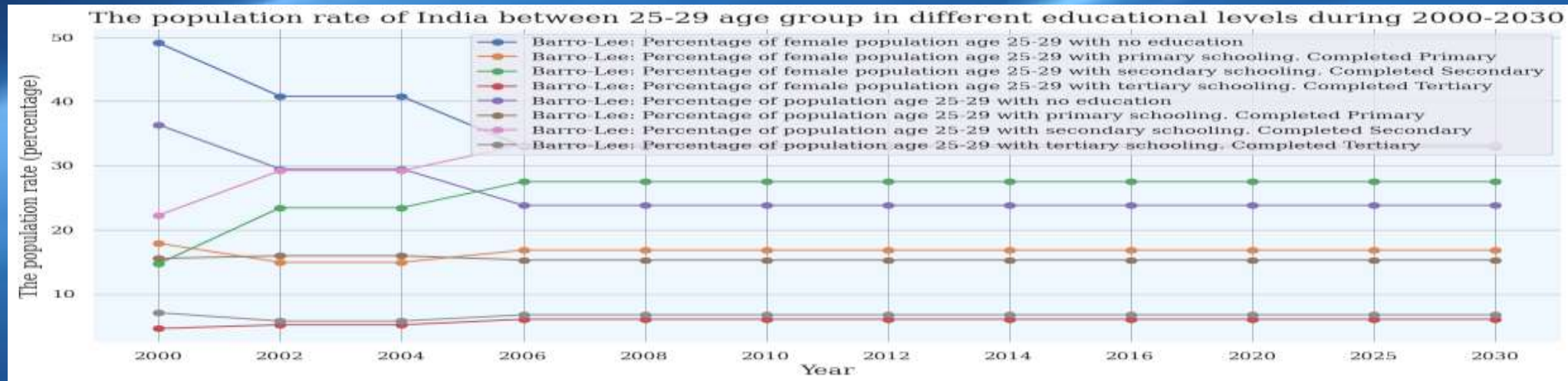
- **Lower secondary educational level (both sexes):** We observe a constant trend during 2000-2030.
- **Lower secondary educational level (female):** We observe a constant trend during 2000-2030.
- **Lower secondary educational level (male):** We observe a constant trend during 2000-2030.
- **Primary educational level (both sexes):** We observe a decline during 2000-2002 and incline during 2002-2004. It becomes constant during 2004-2008 and falls during 2008-2010. It goes up during 2010-2014 and becomes stable during 2014-2030.
- **Primary educational level (female):** We observe an incline during 2000-2004 and become constant during 2004-2006. It increases during 2006-2008 and declines during 2008-2010. It goes up during 2010-2014 and becomes stable during 2014-2030.
- **Primary educational level (male):** We observe an decline during 2000-2002 and incline during 2002-2004. It becomes constant during 2004-2006 and falls during 2006-2010. It goes up during 2010-2014 and becomes stable during 2014-2030.
- **Upper secondary educational level (both sexes):** We observe a constant trend during 2000-2030.
- **Upper secondary educational level (female):** We observe a constant trend during 2000-2030.
- **Upper secondary educational level (male):** We observe a constant trend during 2000-2030.





**Observations:** The population rate of India between 20-24 age group in different educational levels during 2000-2030 (Based on Barro-Lee dataset)

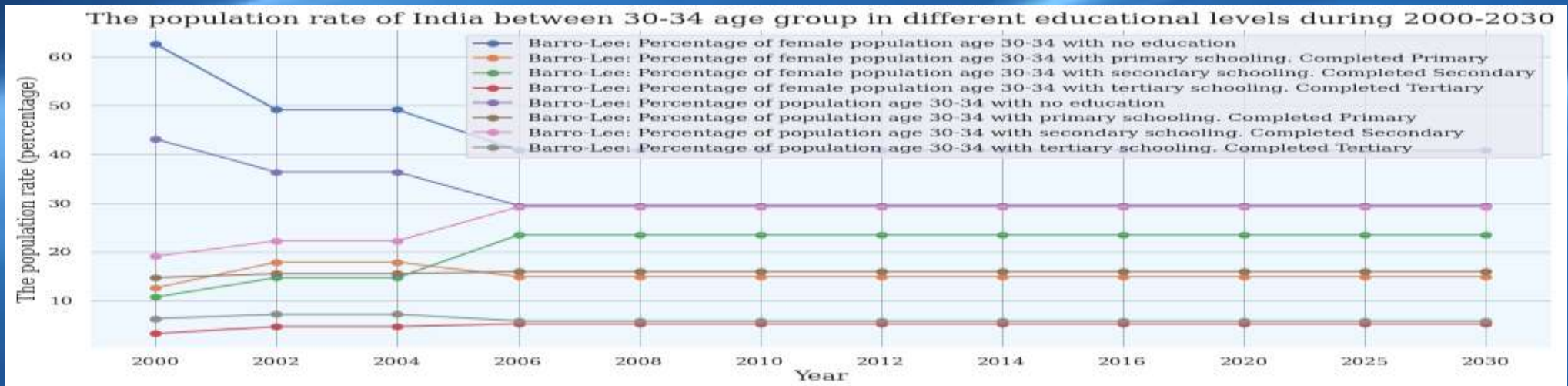
- **No educational level (female):** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **No educational level (both sexes):** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (both sexes):** We observe a decline during 2000-2002 and become constant during 2002-2004. It inclines during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (both sexes):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (both sexes):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.



**Observations:** The population rate of India between 25-29 age group in different educational levels during 2000-2030 (Based on Barro-Lee dataset)

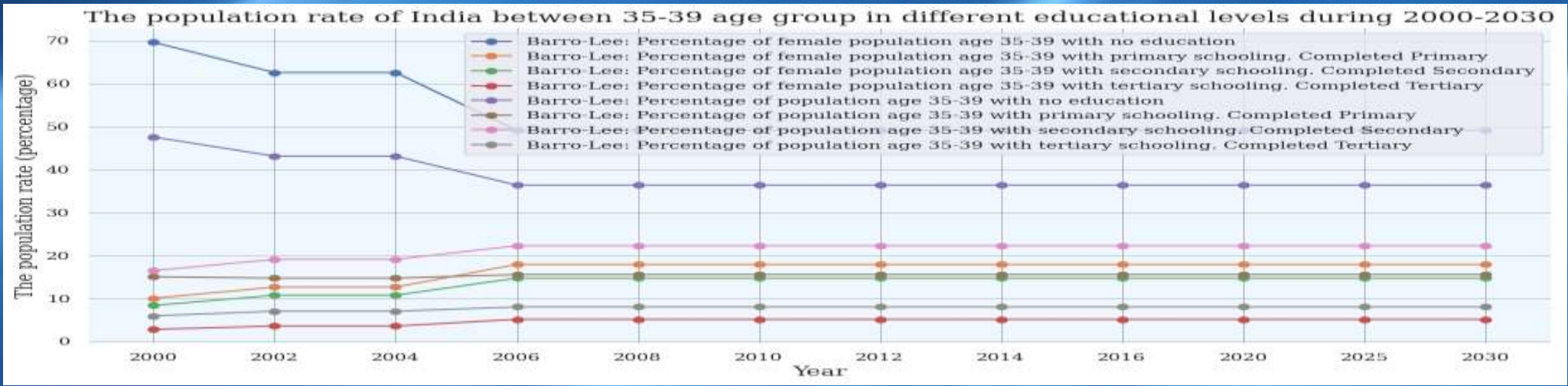
- **No educational level (female):** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (female):** We observe a decline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **No educational level (both sexes):** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (both sexes):** We observe an incline during 2000-2002 and becomes constant during 2002-2004. It declines during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (both sexes):** We observe an incline during 2000-2002 and becomes constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (both sexes):** We observe a decline during 2000-2002 and become constant during 2002-2004. It inclines during 2004-2006 and becomes stable during 2006-2030.





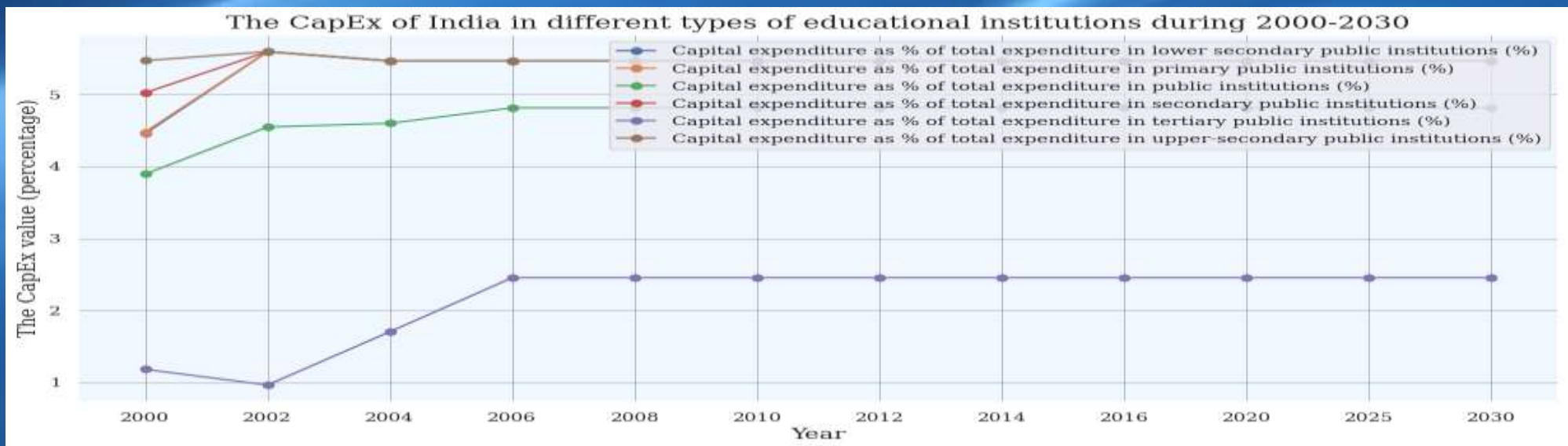
**Observations:** The population rate of India between 30-34 age group in different educational levels during 2000-2030 (Based on Barro-Lee dataset)

- **No educational level (female):** We observe a decline during 2000-2002 and become constant during 2004-2006. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It declines during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **No educational level (both sexes):** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (both sexes):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (both sexes):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (both sexes):** We observe an incline during 2000-2002 and become constant during 2002-2004. It declines during 2004-2006 and becomes stable during 2006-2030.



**Observations:** The population rate of India between 35-39 age group in different educational levels during 2000-2030 (Based on Barro-Lee dataset)

- **No educational level (female):** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (female):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **No educational level (both sexes):** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Primary educational level (both sexes):** We observe a decline during 2000-2002 and become constant during 2002-2004. It inclines during 2004-2006 and becomes stable during 2006-2030.
- **Secondary educational level (both sexes):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.
- **Tertiary educational level (both sexes):** We observe an incline during 2000-2002 and become constant during 2002-2004. It goes up during 2004-2006 and becomes stable during 2006-2030.

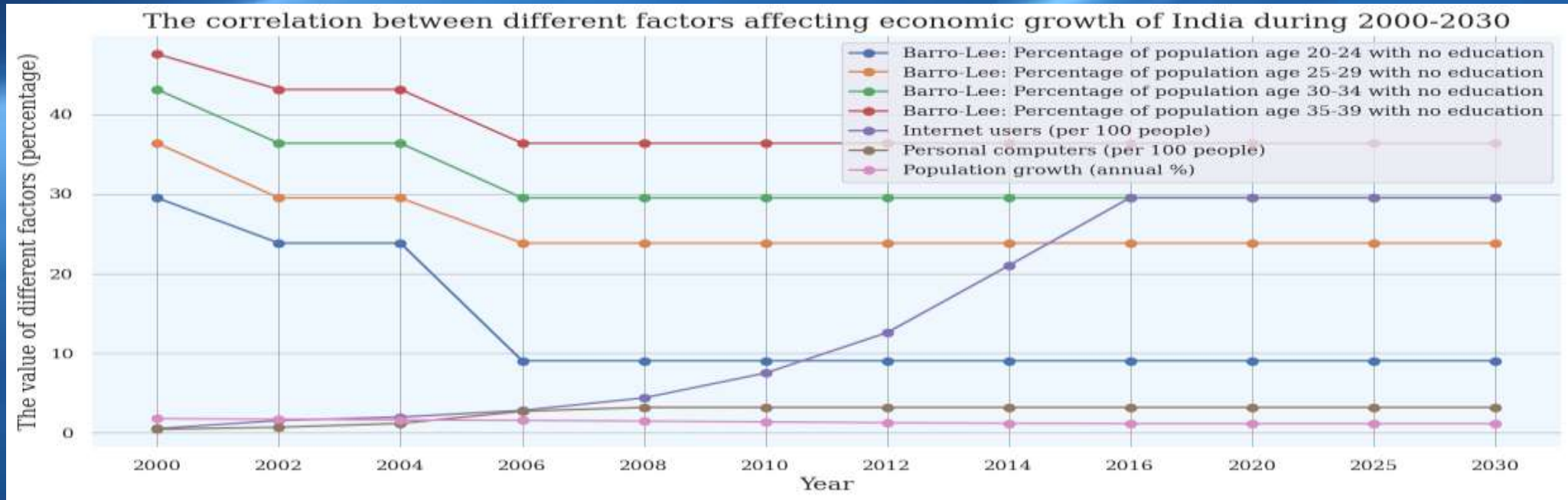


**Capital Expenditure (CapEx)** - Expenditure for education goods or assets that yield benefits for a period of more than one year. It includes expenditure for construction, renovation and major repairs of buildings and the purchase of heavy equipment or vehicles. It represents the value of assets acquired or created – i.e. the amount of capital formation – during the year in which the expenditure occurs.

**Observations:** The CapEx of India in different types of educational institutions during 2000-2030

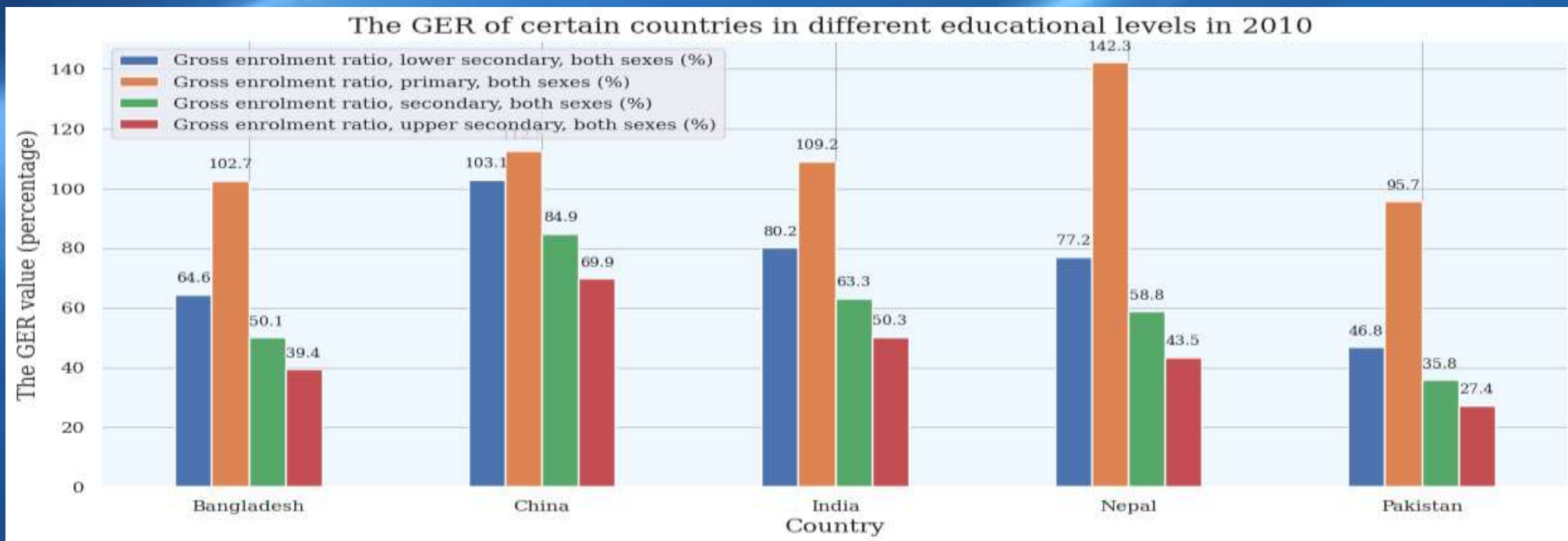
- **Lower secondary public educational institutions:** We observe an incline during 2000-2002 and decline during 2002-2004. It becomes constant during 2004-2030.
- **Primary public educational institutions:** We observe an incline during 2000-2002 and decline during 2002-2004. It becomes constant during 2004-2030.
- **Public educational institutions:** We observe an incline during 2000-2006 and become constant during 2006-2030.
- **Secondary public educational institutions:** We observe an incline during 2000-2002 and decline during 2002-2004. It becomes constant during 2004-2030.
- **Tertiary public educational institutions:** We observe a decline during 2000-2002 and incline during 2002-2006. It becomes constant during 2006-2030.
- **Upper-secondary public educational institutions:** We observe an incline during 2000-2002 and decline during 2002-2004. It becomes constant during 2004-2030.





**Observations:** The correlation between different factors affecting economic growth of India during 2000-2030

- **Population rate between 20-24 age group in no educational level:** We observe a decline during 2000-2002 and become constant during 2002-2004. It falls during 2004-2006 and becomes stable during 2006-2030.
- **Population rate between 25-29 age group in no educational level:** We observe a decline during 2000-2002 and become constant during 2002-2004. It declines during 2004-2006 and becomes stable during 2006-2030.
- **Population rate between 30-34 age group in no educational level:** We observe a decline during 2000-2002 and become constant during 2002-2004. It declines during 2004-2006 and becomes stable during 2006-2030.
- **Population rate between 35-39 age group in no educational level:** We observe a decline during 2000-2002 and become constant during 2002-2004. It declines during 2004-2006 and becomes stable during 2006-2030.
- **Internet users (per 100 people):** We observe an incline during 2000-2016 and become constant during 2016-2030.
- **Personal computers (per 100 people):** We observe an incline during 2000-2008 and become constant during 2008-2030.
- **Population rate:** We observe a decline during 2000-2006 and become constant during 2006-2008. It falls during 2008-2016 and becomes constant during 2016-2030.

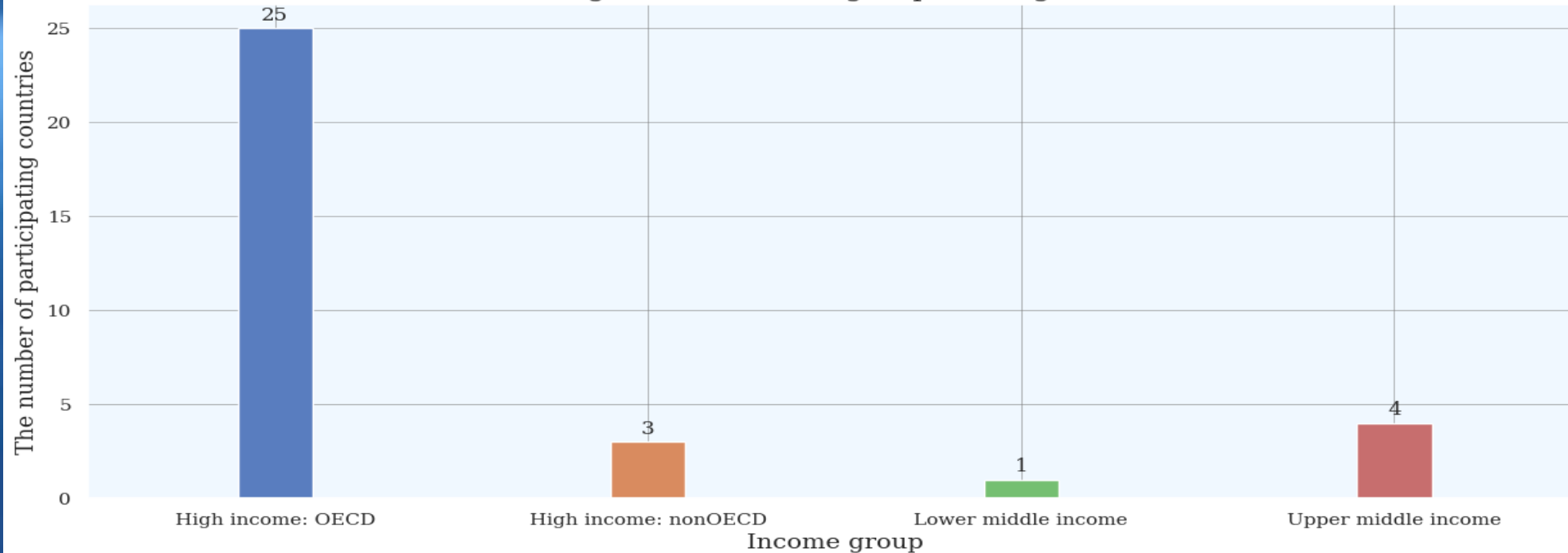


**Gross Enrolment Ratio (GER)** - Gross enrolment ratio (GER) is a statistical measure used in the education sector, and formerly by the UN in its Education Index, to determine the number of students enrolled in school at several different grade levels (like elementary, middle school and high school), and use it to show the ratio of the number of students who live in that country to those who qualify for the particular grade level.

**Observations:** The GER of certain countries in different educational levels in 2010

- **Lower secondary educational level (both sexes):** We observe an increase from Bangladesh to China and decrease from China to Pakistan.
- **Primary educational level (both sexes):** We observe an increase from Bangladesh to China and decrease from China to India. Again it increases from India to Nepal and reduces from Nepal to Pakistan.
- **Secondary educational level (both sexes):** We observe an increase from Bangladesh to China and decrease from China to Pakistan.
- **Upper secondary educational level (both sexes):** We observe an increase from Bangladesh to China and decrease from China to Pakistan.

The number of countries participating in the mean score of students in different scales of PISA according to their income groups during 2000-2015

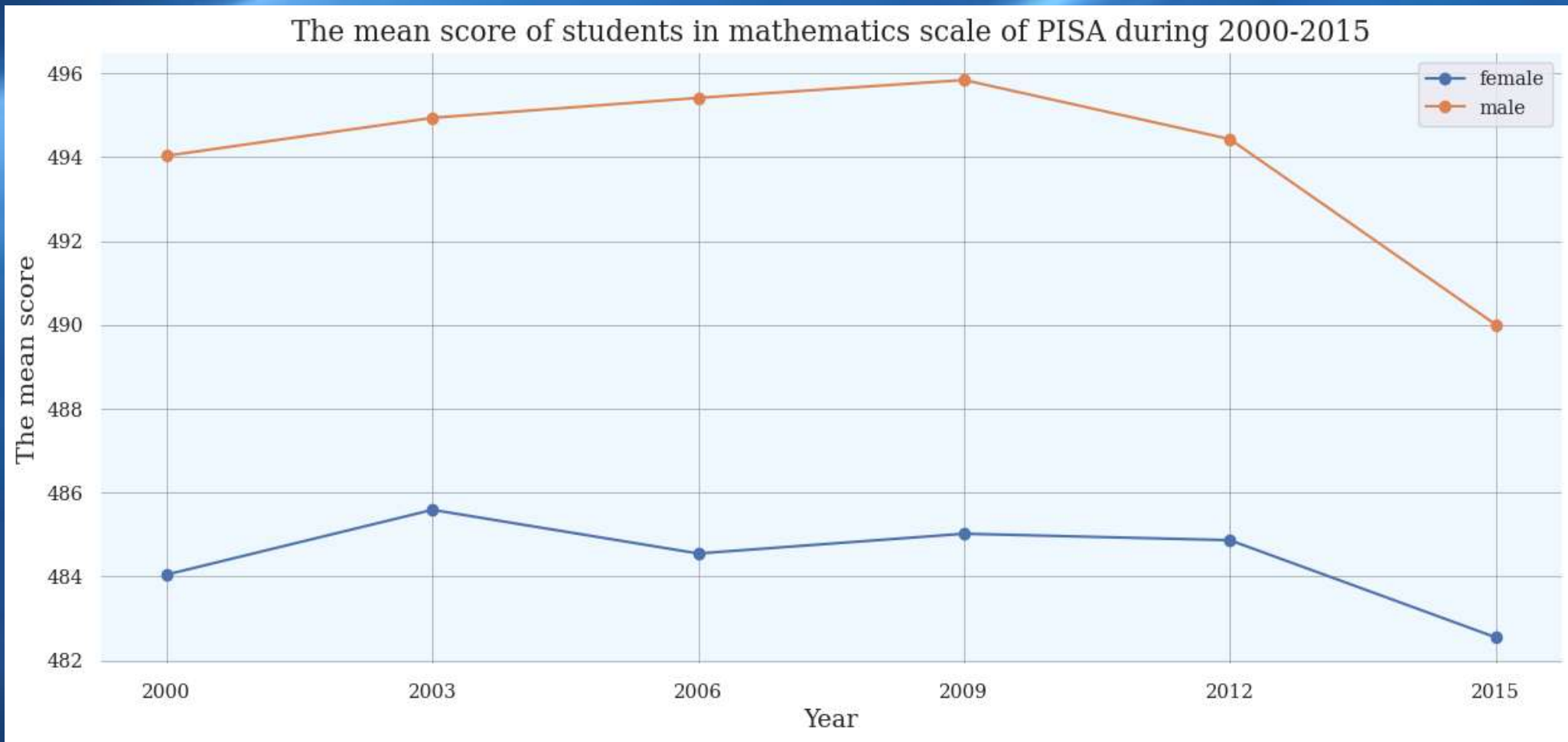


**PISA** - The Program for International Student Assessment (PISA) is an international assessment that measures 15-year-old students' reading, mathematics, and science literacy every 3 years. First conducted in 2000, the major domain of study rotates between reading, mathematics, and science in each cycle.

**Observations:** The number of countries participating in the mean score of students in different scales of PISA according to their income groups during 2000-2015

- We observe that high income: OECD group has maximum participating countries and lower middle income group has minimum participating countries in the mean score of students in different scales of PISA during 2000-2015.
- We observe that upper middle income group has 4 participating countries and high income: nonOECD group has 3 participating countries in the mean score of students in different scales of PISA during 2000-2015.

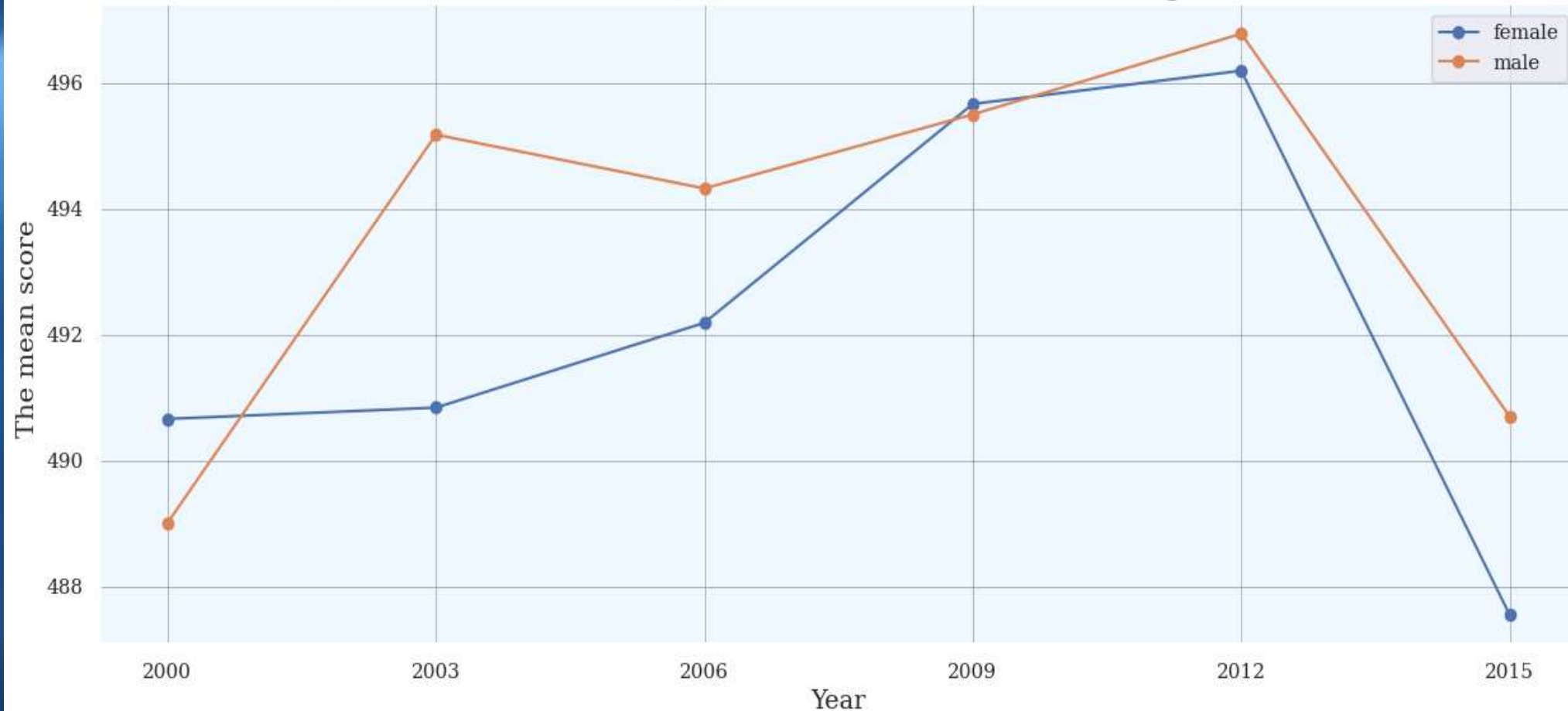




**Observations:** The mean score of students in mathematics scale of PISA during 2000-2015

- **Students (female):** We observe an incline during 2000-2003 and decline during 2003-2006. It goes up during 2006-2009 and falls during 2009-2015.
- **Students (male):** We observe an incline during 2000-2009 and decline during 2009-2015.

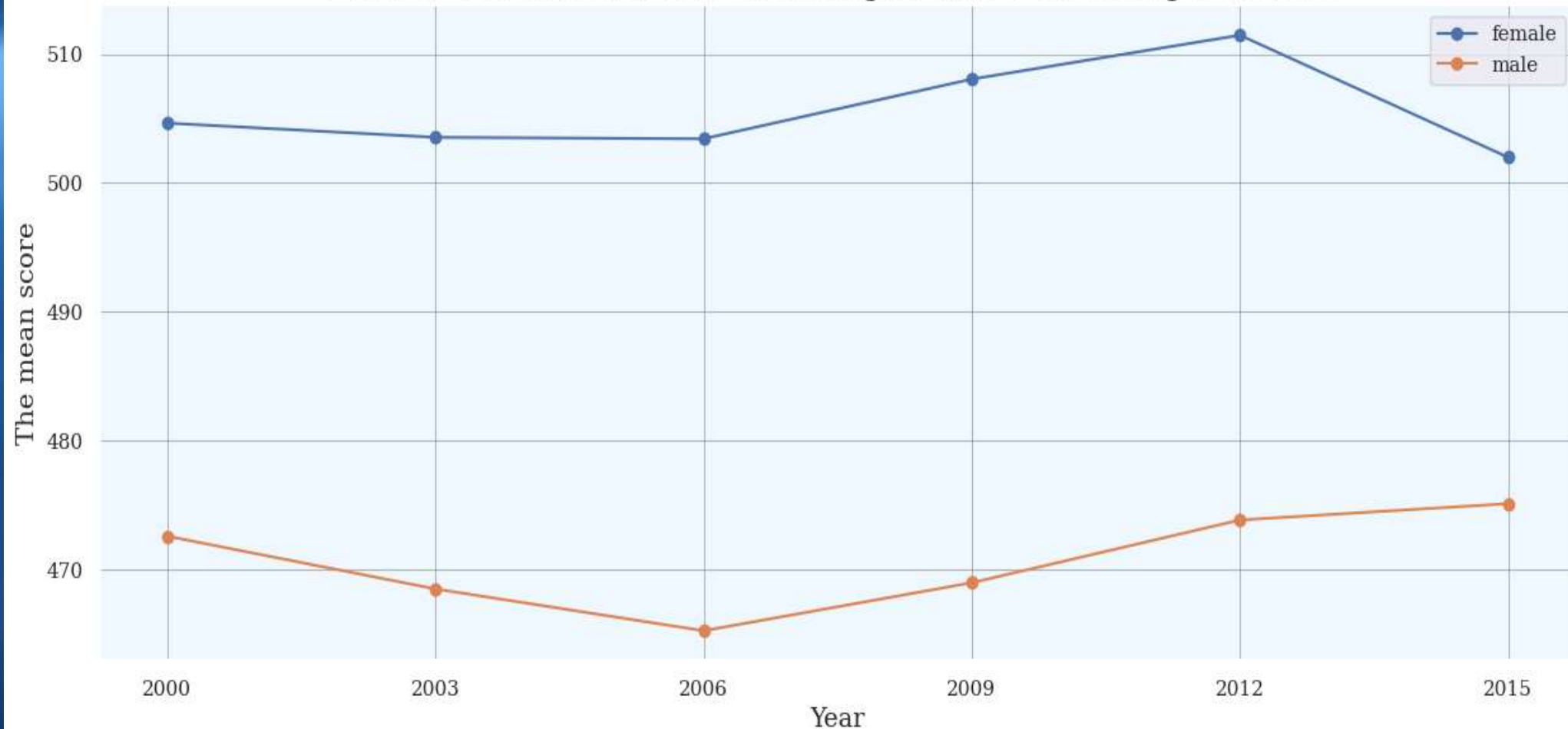
The mean score of students in science scale of PISA during 2000-2015



**Observations:** The mean score of students in science scale of PISA during 2000-2015

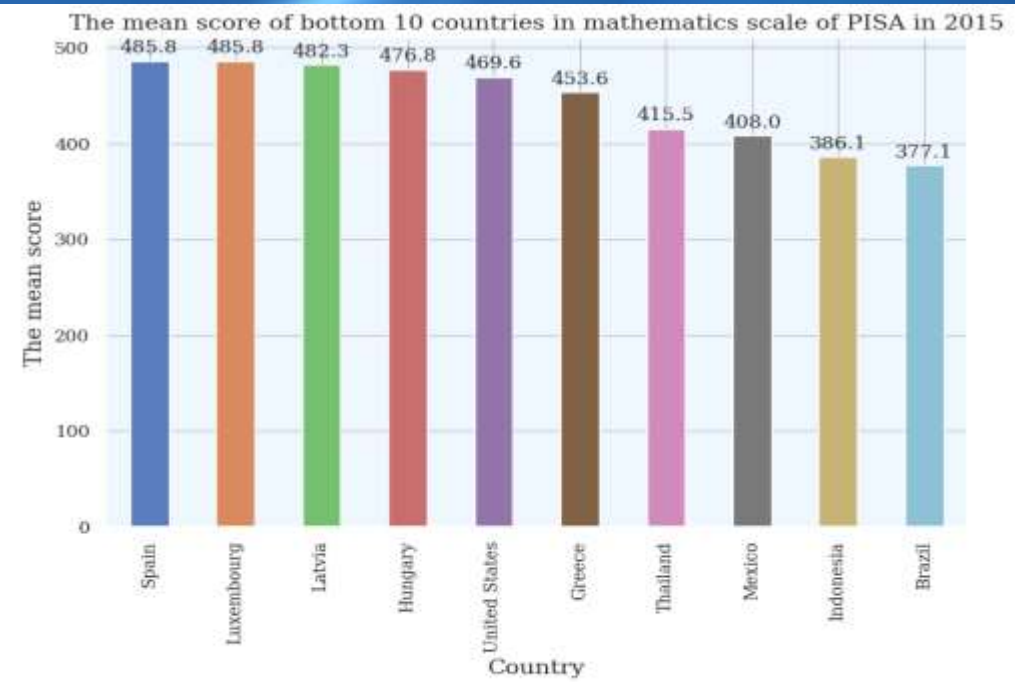
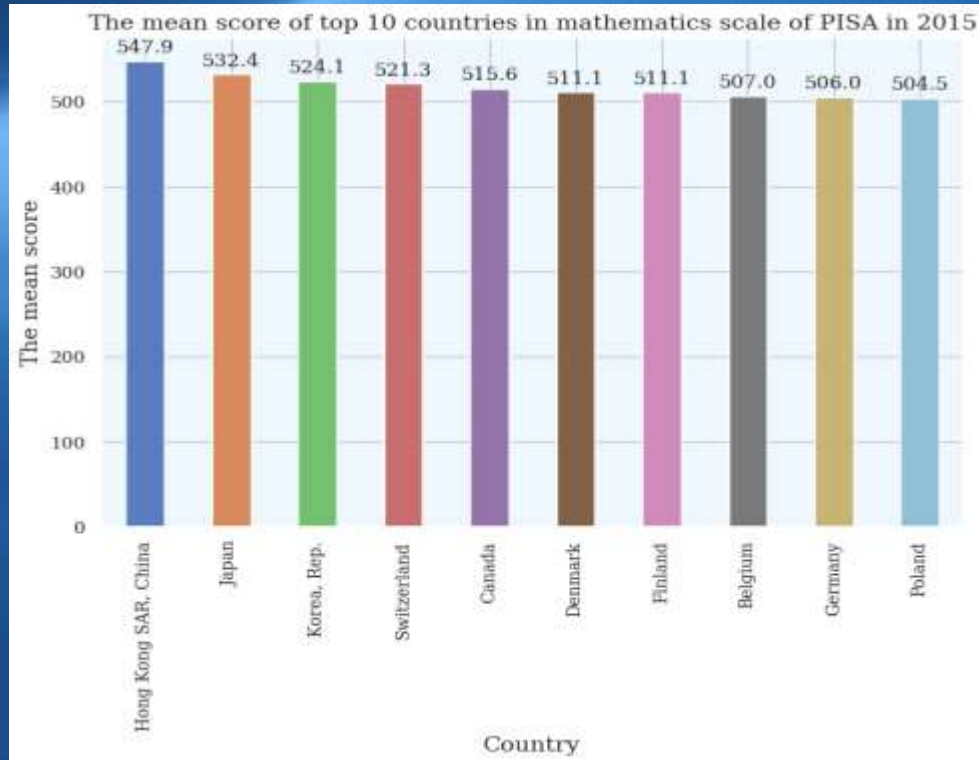
- **Students (female):** We observe an incline during 2000-2012 and decline during 2012-2015.
- **Students (male):** We observe an incline during 2000-2003 and decline during 2003-2006. It goes up during 2006-2012 and falls during 2012-2015.

The mean score of students in reading scale of PISA during 2000-2015



**Observations:** The mean score of students in reading scale of PISA during 2000-2015

- **Students (female):** We observe a decline during 2000-2006 and incline during 2006-2012. It falls during 2012-2015.
- **Students (male):** We observe a decline during 2000-2006 and incline during 2006-2015.



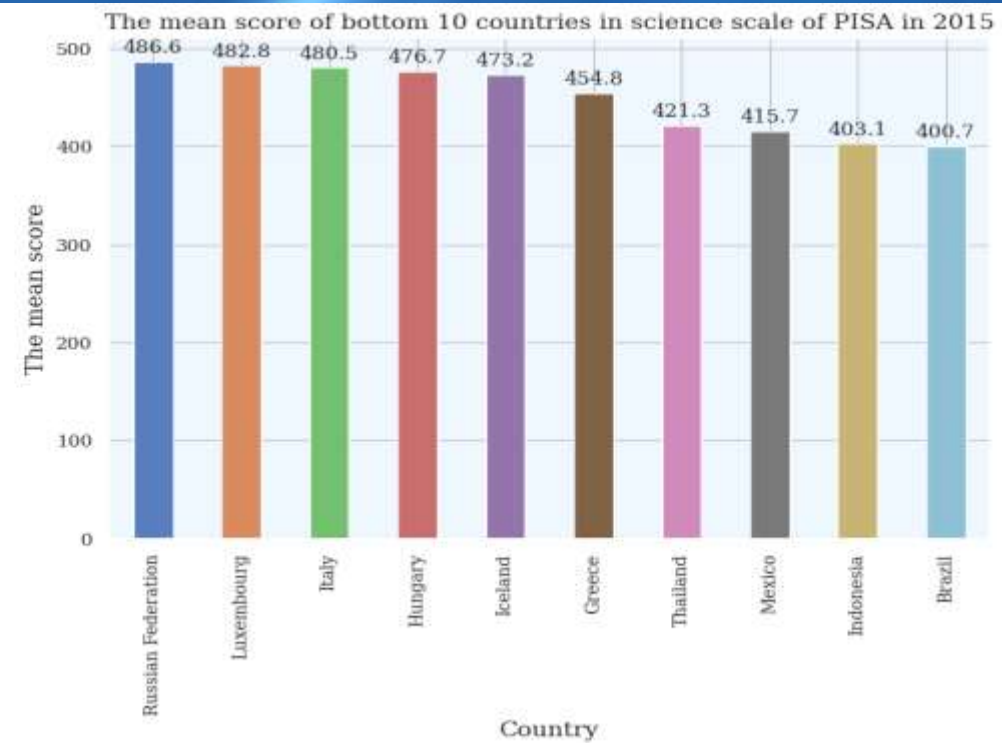
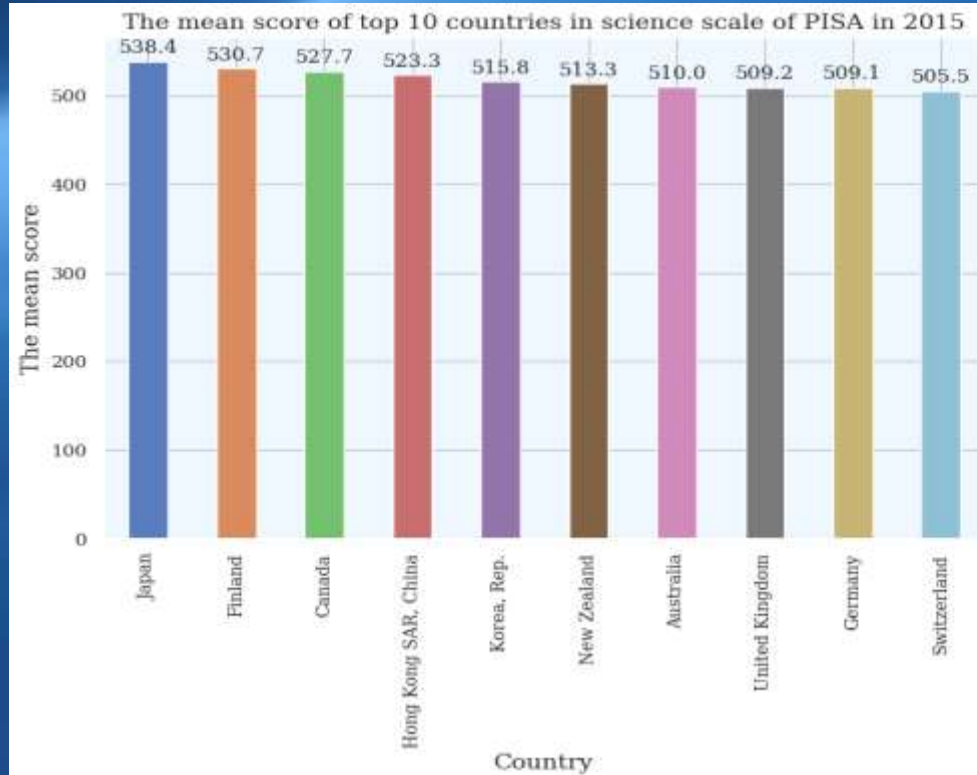
**Observations:** The mean score of top 10 and bottom 10 countries in mathematics scale of PISA in 2015

#### **The mean score of top 10 countries in mathematics scale of PISA in 2015**

- We observe that Hong Kong SAR, China has the highest mean score and Poland has the lowest mean score among all the top 10 countries in the mathematics scale of PISA in 2015.
- We observe that out of top 10 countries, 6 are European countries, 3 are Asian countries and Canada is the only North American country.

#### **The mean score of bottom 10 countries in mathematics scale of PISA in 2015**

- We observe that Spain has the highest mean score and Brazil has the lowest mean score among all the bottom 10 countries in the mathematics scale of PISA in 2015.
- We observe that out of bottom 10 countries, 5 are European countries, 2 are Asian countries, 2 are Latin American countries and United States is the only North American country.



**Observations:** The mean score of top 10 and bottom 10 countries in science scale of PISA in 2015

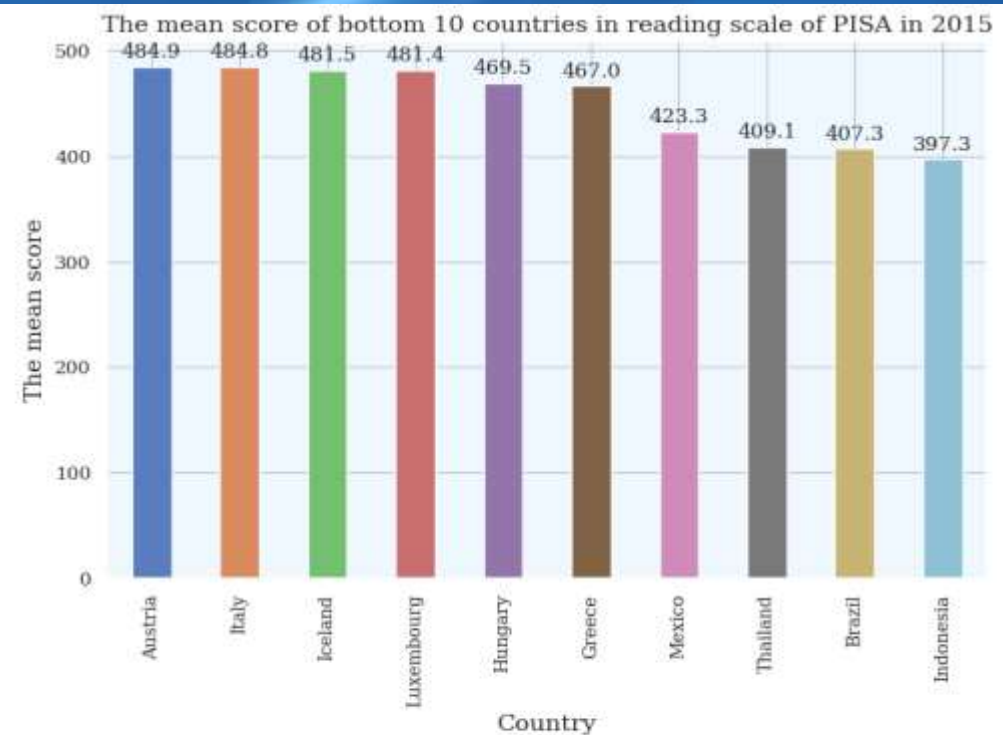
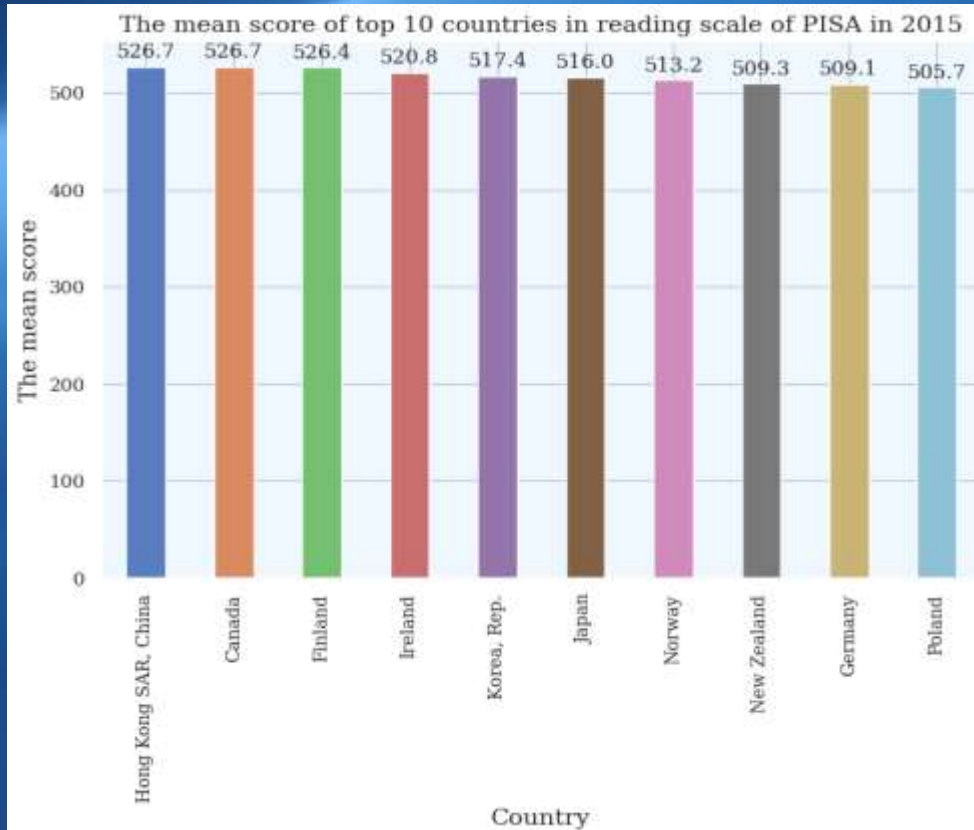
#### **The mean score of top 10 countries in science scale of PISA in 2015**

- We observe that Japan has the highest mean score and Switzerland has the lowest mean score among all the top 10 countries in the science scale of PISA in 2015.
- We observe that out of top 10 countries, 5 are Asian countries, 3 are European countries and 2 are North American countries.

#### **The mean score of bottom 10 countries in science scale of PISA in 2015**

- We observe that Russian Federation has the highest mean score and Brazil has the lowest mean score among all the bottom 10 countries in the science scale of PISA in 2015.
- We observe that out of bottom 10 countries, 6 are European countries, 2 are Asian countries and 2 are Latin American countries.





**Observations:** The mean score of top 10 and bottom 10 countries in reading scale of PISA in 2015

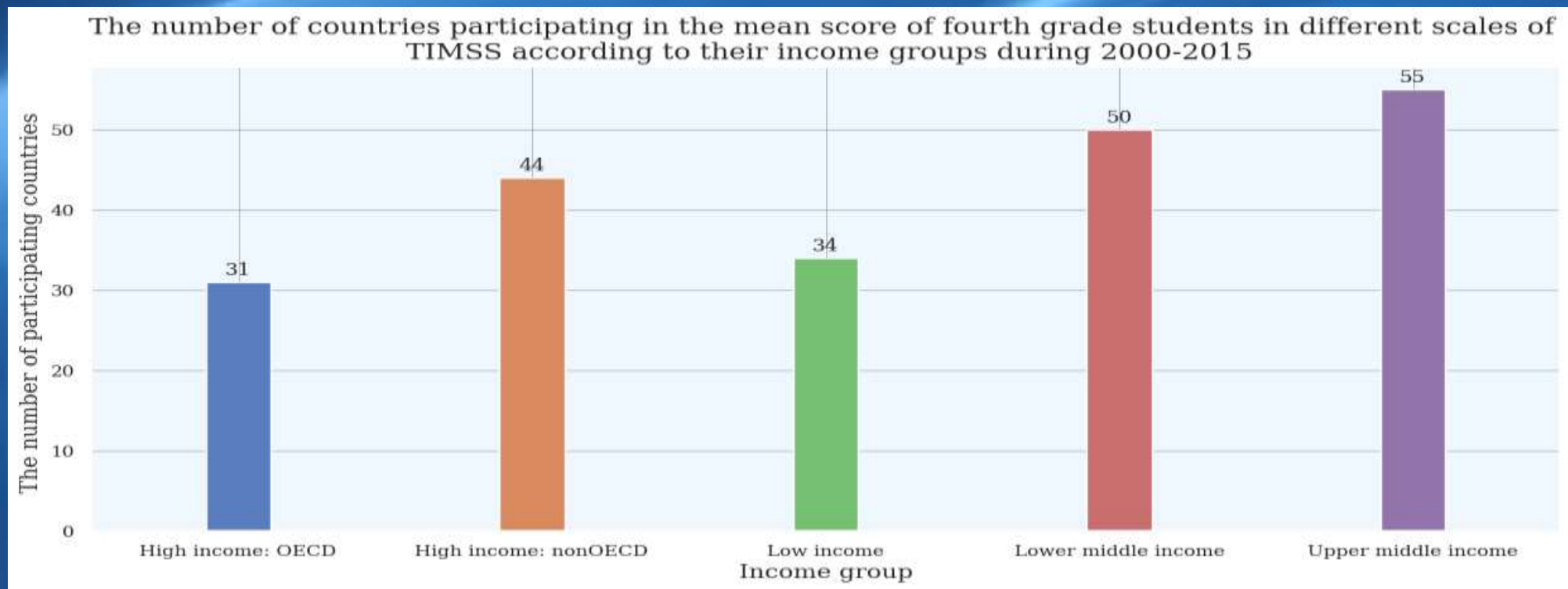
#### **The mean score of top 10 countries in reading scale of PISA in 2015**

- We observe that Hong Kong SAR, China has the highest mean score and Poland has the lowest mean score among all the top 10 countries in the reading scale of PISA in 2015.
- We observe that out of top 10 countries, 5 are European countries, 4 are Asian countries and Canada is the only North American country.

#### **The mean score of bottom 10 countries in reading scale of PISA in 2015**

- We observe that Austria has the highest mean score and Indonesia has the lowest mean score among all the bottom 10 countries in the reading scale of PISA in 2015.
- We observe that out of bottom 10 countries, 6 are European countries, 2 are Asian countries and 2 are Latin American countries.

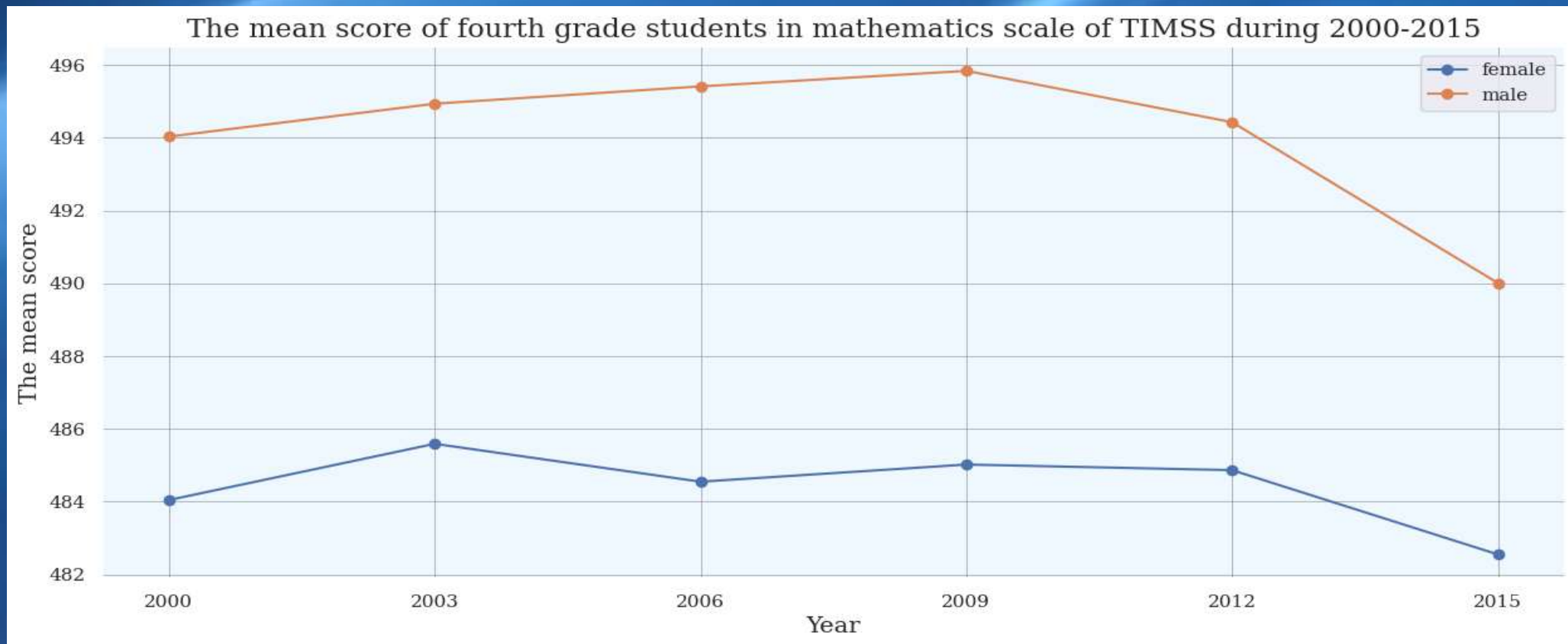




**TIMSS** - The Trends in International Mathematics and Science Study (TIMSS) is an assessment of the mathematics and science knowledge of fourth (or fifth) and eighth (or ninth) grade learners in various countries around the world.

**Observations:** The number of countries participating in the mean score of fourth grade students in different scales of TIMSS according to their income groups during 2000-2015

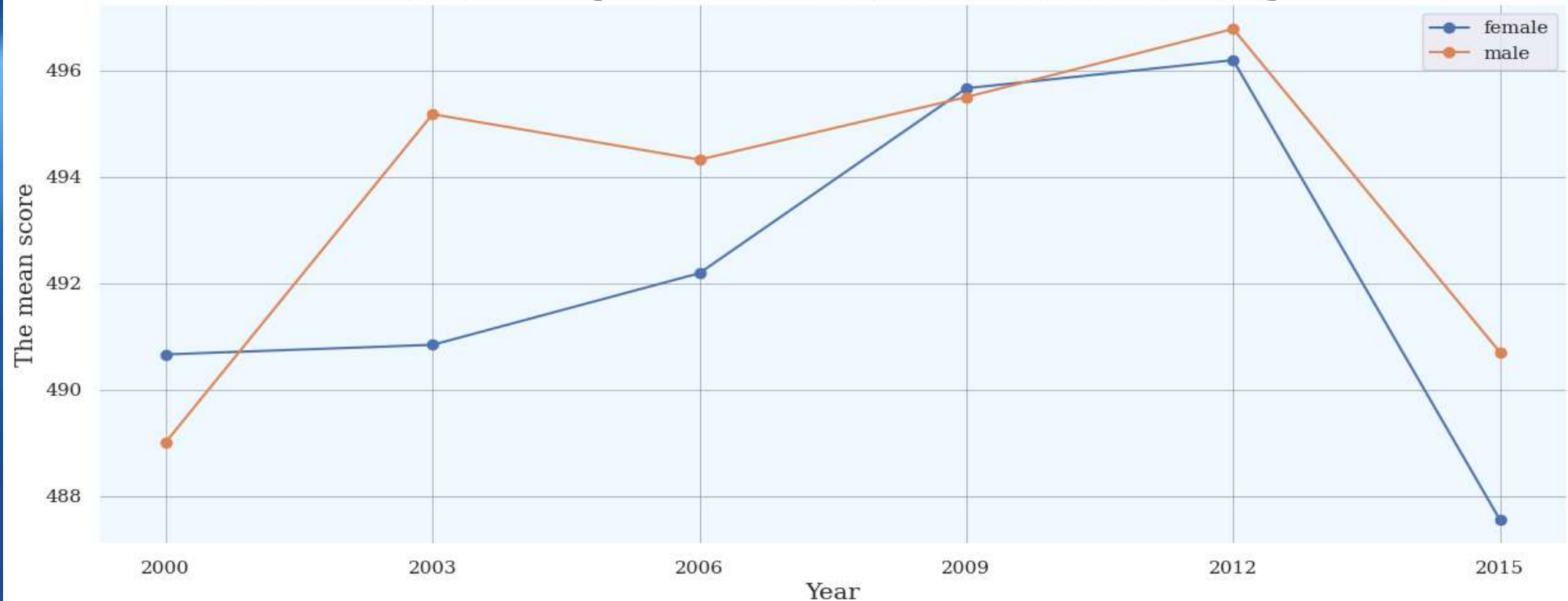
- We observe that upper middle income group has maximum participating countries and high income: OECD group has minimum participating countries in the mean score of fourth grade students in different scales of TIMSS during 2000-2015.
- We observe that lower middle income group has 50 participating countries, high income: nonOECD group has 44 participating countries and low income group has 34 participating countries in the mean score of fourth grade students in different scales of TIMSS during 2000-2015.



**Observations:** The mean score of fourth grade students in mathematics scale of TIMSS during 2000-2015

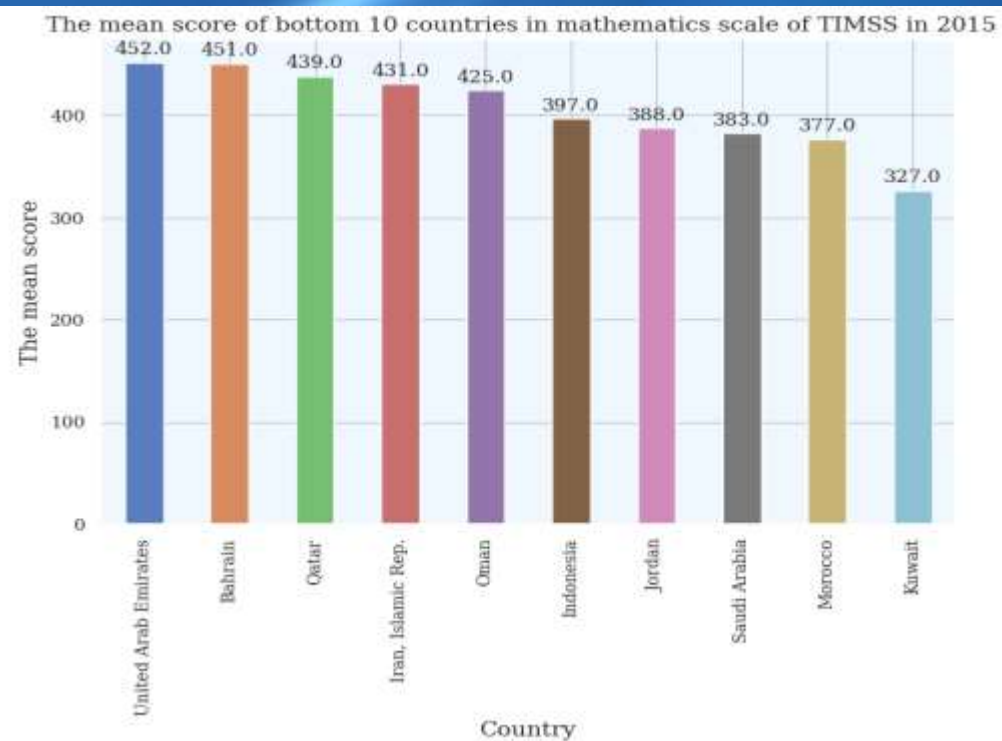
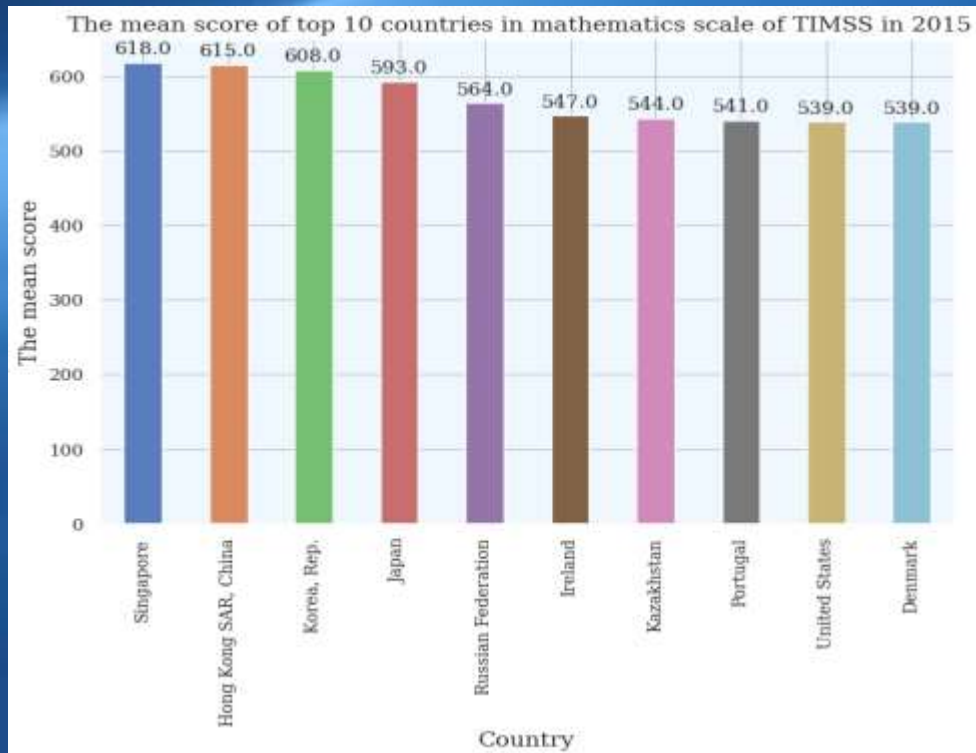
- **Students (female):** We observe an incline during 2000-2003 and decline during 2003-2006. It goes up during 2006-2009 and falls during 2009-2015.
- **Students (male):** We observe an incline during 2000-2009 and decline during 2009-2015.

The mean score of fourth grade students in science scale of TIMSS during 2000-2015



**Observations:** The mean score of fourth grade students in science scale of TIMSS during 2000-2015

- **Students (female):** We observe an incline during 2000-2012 and decline during 2012-2015.
- **Students (male):** We observe an incline during 2000-2003 and decline during 2003-2006. It goes up during 2006-2012 and falls during 2012-2015.



**Observations:** The mean score of top 10 and bottom 10 countries in mathematics scale of TIMSS in 2015

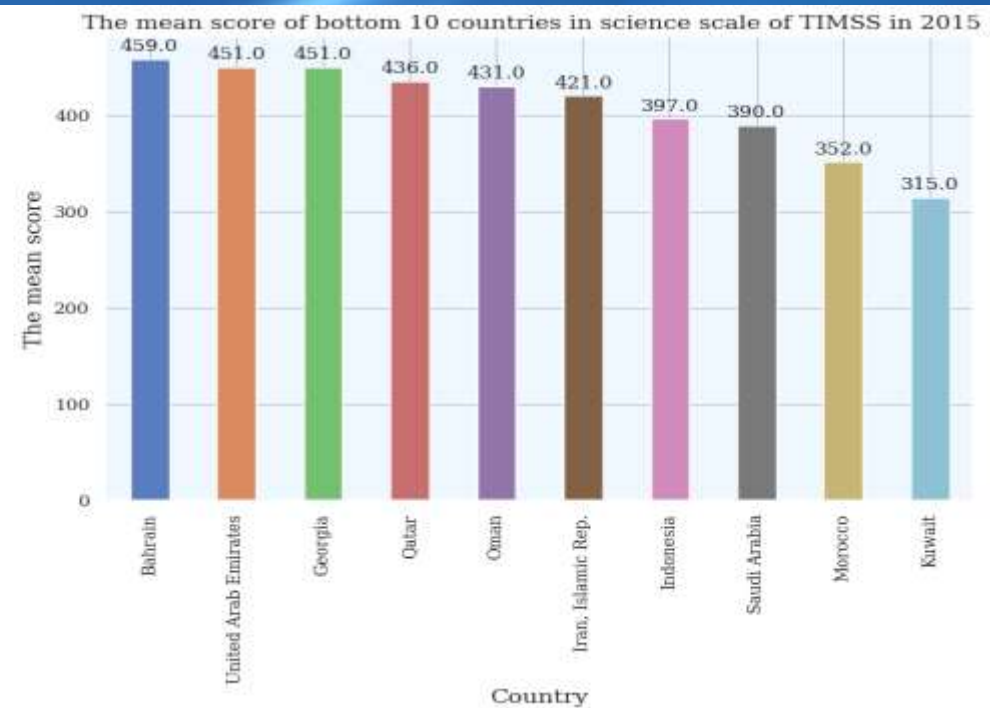
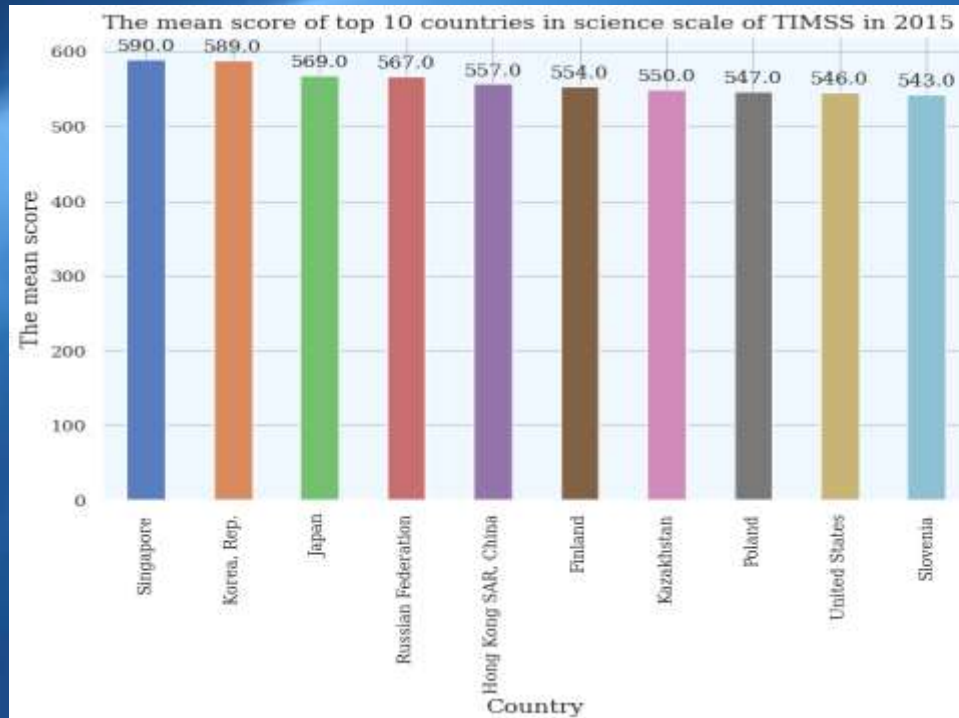
#### **The mean score of top 10 countries in mathematics scale of TIMSS in 2015**

- We observe that Singapore has the highest mean score and Denmark has the lowest mean score among all the top 10 countries in the mathematics scale of TIMSS in 2015.
- We observe that out of top 10 countries, 5 are Asian countries, 4 are European countries and United states is the only North American country.

#### **The mean score of bottom 10 countries in mathematics scale of TIMSS in 2015**

- We observe that United Arab Emirates has the highest mean score and Kuwait has the lowest mean score among all the bottom 10 countries in the mathematics scale of TIMSS in 2015.
- We observe that out of bottom 10 countries, 7 are Middle Eastern countries, 2 are Arab States countries and Indonesia is the only Asian country.





**Observations:** The mean score of top 10 and bottom 10 countries in science scale of TIMSS in 2015

#### **The mean score of top 10 countries in science scale of TIMSS in 2015**

- We observe that Singapore has the highest mean score and Slovenia has the lowest mean score among all the top 10 countries in the science scale of TIMSS in 2015.
- We observe that out of top 10 countries, 5 are Asian countries, 4 are European countries and United States is the only North American country.

#### **The mean score of bottom 10 countries in science scale of TIMSS in 2015**

- We observe that Bahrain has the highest mean score and Kuwait has the lowest mean score among all the bottom 10 countries in the science scale of TIMSS in 2015.
- We observe that out of bottom 10 countries, 6 are Middle Eastern countries, 2 are Arab States countries, Georgia is European country and Indonesia is the only Asian country.

## Conclusion:

- This study shows that most of the countries, whether developing and developed, are part of the World Bank, that's why the "Region" with a high number of countries has high involvement.
- The population growth analysis, based on income, indicates that the country's government of law, middle wages should implement some laws to control the population growth.
- GDP bar shows that developed countries have GDP exceeding 120000 and developing countries have GDP exceeding 12000. This indicates a very vast difference in terms of economy of the country.
- Educational analysis of the youth shows that literacy rate is increasing over the years but still females are lagging behind, compared to the total population literacy. Government and individuals should pay attention to education.
- Literacy rate and technological advancement, influence the GDP of the country, should be increased.
- PISA and TIMSS analysis shows that it's quite different for males and females in the mathematics scale and reading scale.

Thank You