

HarvardX - PH125.9x Data Science: Capstone

PISA Project

Julie COLLEONI

2025-06-05

Contents

1	Introduction	2
1.1	Dataset Overview	3
1.2	Unique school ID creation	4
2	Methods and Data Analysis	4
2.1	Data Preparation	4
2.1.1	Database explanation	4
2.1.2	Data cleaning	7
2.2	Database exploration and visualization	16
2.2.1	Evolution of the PISA score by Region	16
2.2.2	GDP and Global Education: Mapping PISA Participation by Income Level	18
2.2.3	Educational Mobility Analysis: ESCS Bottom 25% to Top 25% Performance	19
2.2.4	PISA Performance in Lower-Income Countries: The Socioeconomic Achievement Gap Across Time	21
2.2.5	PISA Performance in Upper-Income Countries: The Socioeconomic Achievement Gap Across Time	23
2.2.6	PISA Performance in High Income Countries: The Socioeconomic Achievement Gap Across Time	26
2.2.7	Educational investment and educational outcomes for Lower-middle Income	30
2.2.8	Educational investment and educational outcomes for Upper-middle Income	32
2.2.9	Educational investment and educational outcomes for High Income	34
2.2.10	Economic Prosperity vs Educational Achievement	37
2.2.11	Student-Teacher Ratio vs Educational Achievement	40
2.2.12	Public vs Private school performance for Lower Middle Income Countries	44
2.2.13	Public vs Private school performance for Upper Middle Income Countries	47
2.2.14	Public vs Private school performance for High Income Countries	49
2.2.15	The gender performance gap for the Lower Middle Income Countries	52
2.2.16	Mathematics Gender Gap Analysis: Lower-Middle Income Countries in 2022	55

2.2.17	The gender performance gap for the Upper Middle Income Countries	57
2.2.18	Mathematics Gender Gap Analysis: Upper-Middle Income Countries in 2022	62
2.2.19	The gender performance gap for the High Income Countries	64
2.2.20	Mathematics Gender Gap Analysis: High Income Countries in 2022	69
2.2.21	UNICEF Gender Parity Index vs PISA Gender Gap Analysis	72
2.2.22	UNICEF Gender Parity Index vs PISA Mathematics Gender Gap Analysis	75
3	Modelling	76
3.1	Introduction to Modeling	76
3.2	Performance metric: RMSE	76
3.3	Train/Test Split Methodology	77
3.4	Understanding the academic performance correlations	78
3.5	Naive base line approach	80
3.6	Predictive Model Benchmarking	80
4	Results	81
5	Conclusion	81
6	Comments on the research project and references	82

1 Introduction

This report documents a predictive modeling project conducted as part of the HarvardX program in Data Science.

Education is widely recognized as a key driver of social and economic development. The Programme for International Student Assessment (PISA), led by the OECD, evaluates educational systems worldwide by testing the skills and knowledge of 15-year-old students. Differences in school performance, both within and between countries, reflect not only national education policies, but also local context, resources, and individual backgrounds.

This project addresses the predicability of academic performance:

- **Student-level success probability:** Can the likelihood of an individual student achieving a high score be estimated, based on national trends, the context of their school, family background, and personal characteristics?

This project explores the determinants of educational success using the **LearningTower dataset**, which compiles international data from the **OECD PISA studies at the school and student levels** and provides an opportunity to investigate these questions.

The research work on PISA survey data through LearningTower dataset was made possible thanks to the great work of the MIT team:

Wang K, Yacobellis P, Siregar E, Romanes S, Fitter K, Dalla Riva G, Cook D, Tierney N, Dingorkar P, Sai Subramanian S, Chen G (2024). learningtower: OECD PISA Datasets from 2000-2022 in an Easy-to-Use Format. R package version 1.1.0, <https://github.com/kevinwang09/learningtower>, <https://kevinwang09.github.io/learningtower/>.

In addition to the core PISA LearningTower datasets, several auxiliary databases are integrated to enrich the contextual and explanatory power of the analysis. These include:

- **UNESCO School Gender Enrollment Index** Source: UNESCO, World Bank Data 360 Description: This dataset contains enrollment indices by gender for secondary and upper secondary education, sourced from UNESCO and curated for this project. It provides an overview of the relative school participation rates for males and females, allowing for the analysis of gender parity and educational equality across countries and over time.
- **UNESCO Government Expenditure on Education** Source: UNESCO, World Bank Data 360 Description: This table reports the percentage of total government expenditure allocated to education for each country. It is based on UNESCO data and provides insight into national education priorities and funding trends, which are key contextual factors in the analysis of educational outcomes.
- **ILO/World Bank Country Groupings (Region, Subregion, and Income Level)** Source: ILOSTAT Country Groupings Description: This classification table groups countries by ILO region, subregion, and World Bank income classification (low, lower-middle, upper-middle, and high income). It enables regional and socioeconomic stratification of educational indicators, supporting comparative and equity-focused analysis.
- **GDP per capita (PPP) /the OECD and Eurostat databases** Source: World Bank Data 360 Description: A country's gross domestic product (GDP) at purchasing power parity (PPP) per capita is the PPP value of all final goods and services produced within an economy in a given year, divided by the average (or mid-year) population for the same year. This is similar to nominal GDP per capita but adjusted for the cost of living in each country.

By merging these sources, the project enables **multilevel analysis of educational performance, accounting for gender parity, resource allocation, regional context, and national socioeconomic status.**

The objectives of the present analysis are as follows:

1. **To estimate, for each student, the probability of “success”,** by incorporating information about the student's school, family environment, and individual profile.
2. **To identify and interpret the key drivers of student success,** and to suggest targeted policy recommendations.

The methodology combines data preparation, exploratory visualization, and predictive modeling, with a particular emphasis on interpretability. Special attention is given to disentangling the respective roles of country, school, family, and individual factors in educational outcomes, and to quantifying the relative importance of these factors using machine learning techniques.

By leveraging these analytical approaches, the project aims to deliver both accurate predictive tools and actionable insights for educators and policymakers.

1.1 Dataset Overview

The **LearningTower dataset** compiles detailed information on schools and students across multiple countries, primarily based on data from the **OECD PISA studies**. This dataset offers a rich set of variables, including academic results, school resources, student demographics, and various contextual indicators. For this project, the most recent publicly available data will be used. The dataset can be directly downloaded from the LearningTower GitHub repository.

```
knitr::opts_chunk$set(echo=FALSE)

# Load necessary packages
if (!require(tidyverse))
  install.packages("tidyverse", repos = "https://cloud.r-project.org/")

## # A tibble: 6 x 2
##   country country_name
##   <chr>    <chr>
```

```
## 1 AZE      Azerbaijan
## 2 ARG      Argentina
## 3 AUS      Australia
## 4 AUT      Austria
## 5 BEL      Belgium
## 6 BRA      Brazil
```

year	Number of Students	Number of Schools
2000	127236	8526
2003	276165	10274
2006	398750	14365
2009	515958	18641
2012	480174	18139
2015	519334	17908
2018	612004	21903
2022	613744	21629

Verification of the columns name of the different tables

```
## [1] "country"          "country_name"      "school_level"
## [4] "year"             "gender_parity_index"

## [1] "country"          "country_name"
## [3] "year"             "percentage_education_on_gdp"

## [1] "country"          "country_name"
## [3] "ilo_region"       "ilo_subregion"
## [5] "ilo_subregion_detail" "world_bank_income_group"
## [7] "ilo_member"

## [1] "country"          "country_name"      "GDP_per_PPP_2022"
```

1.2 Unique school ID creation

To prevent data leakage and to ensure a fair evaluation, the test set was constructed in such a way that all key identifiers present in the test set (for instance, school IDs for the school dataset or school IDs within the student dataset) are also present in the training set. Any observations in the test set referencing identifiers not found in the training set were returned to the training set.

For any analyses or predictions at the school level, the corresponding schools can be filtered directly from the main school dataset based on the unique school identifiers present in the student test set. This approach guarantees consistency across student- and school-level analyses and avoids potential issues related to unseen entities.

2 Methods and Data Analysis

2.1 Data Preparation

2.1.1 Database explanation

Explanation for Sampling Weights (stu_wgt) in PISA

Analyses In PISA datasets, not all sampled students are equally representative of the national student population. To account for differences in selection probability, sampling weights must be used in all analyses. For instance, if students from small schools are intentionally oversampled in a given country, unweighted statistics would disproportionately reflect the characteristics of those students, leading to biased results. The use of sampling weights corrects for such over- or under-representation, ensuring unbiased and population-representative estimates.

The final student weight, stored in the variable `stu_wgt` in the international PISA files, reflects each student's probability of selection and relevant adjustments. The sum of these weights estimates the total number of 15-year-old students enrolled in school in the target country. As a result, countries with larger populations will naturally contribute more to the aggregated results than smaller countries.

It is essential to incorporate the `stu_wgt` variable in all analyses, including any calculations, summaries, and modeling, to accurately reflect the survey design and avoid biased conclusions. For example, when calculating means or regression coefficients, weighted methods should be used so that each student's contribution is proportional to their sampling weight.

Note:

Using student weights designed for one subject (for instance, reading) to analyze results for another subject (for instance, mathematics or science) can introduce bias, especially due to oversampling of students in special education. In such cases, correction factors must be applied to avoid under- or overestimation, as detailed in the PISA technical documentation.

In summary, the `stu_wgt` variable is critical for producing valid and representative results and must always be considered in all analyses using PISA data.

Explanation of the ESCS Index (Economic, Social and Cultural Status Index) in PISA

The ESCS index, developed for the PISA assessment, is a **composite measure that summarizes multiple facets of a student's social, economic, and cultural background**. It was designed to facilitate cross-country comparisons of equity in educational outcomes.

Structure:

The ESCS index is constructed using **three main components**:

- 1. ISEI – Parental Occupational Status:**

The highest International Socio-Economic Index of Occupational Status among the student's parents, based on the type of occupation (using ISCO/CITP classification).

- 2. PARED – Parental Education:**

The highest educational attainment of the parents, measured in years of schooling and codified using the International Standard Classification of Education (ISCED/CITE).

- 3. HOMEPOS – Home Possessions:**

An index summarizing the student's access to resources at home, itself built from several sub-components:

- **WEALTH:** Material possessions such as a private room, internet connection, dishwasher, DVD player, number of telephones, televisions, computers, and cars.
- **HEDRES:** Educational resources, including a desk/table for study, a quiet place to work, a computer for schoolwork, educational software, technical books, and a dictionary.
- **CULTPOSS:** Cultural possessions, such as classical literature, poetry collections, and works of art.
- **Number of books at home:** As a separate indicator.

Calculation method:

All these components are **combined through a factor analysis**, resulting in a single, standardized ESCS score for each student. The relative weights of each component are determined internationally and then applied to all students, regardless of country.

Purpose and interpretation:

- The ESCS is widely used as a benchmark for analyzing equity and the impact of socio-economic background on student performance.
- It provides a **unidimensional summary** of socio-economic and cultural status, facilitating comparisons across countries and time.
- However, the aggregation of diverse indicators means that nuances in the individual contributions of each component may be lost. The importance of each factor can vary significantly between countries and educational systems.

Limitations:

- The ESCS relies on student-reported data, which can be subject to reporting errors or non-response (especially for parental education or occupation).
- Some items may have different social significance or predictive power across countries. For example, owning a dishwasher may correlate positively with achievement in some countries, but not in others.

In summary:

The ESCS index provides a standardized measure of students' economic, social, and cultural background in PISA, integrating parental occupation and education, as well as material, educational, and cultural resources at home. It is a powerful tool for analyzing educational equity but should be interpreted with an understanding of its components and limitations.

```
## Rows: 3,543,365
## Columns: 23
## $ year      <int> 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000~
## $ country   <fct> ALB, ALB, ALB, ALB, ALB, ALB, ALB, ALB, ALB, ALB, ALB, ALB~
## $ school_id <chr> "1001", "1001", "1001", "1001", "1001", "1001", "1001", "1001", "1~
## $ student_id <int> 1, 3, 6, 8, 11, 12, 17, 20, 21, 22, 24, 27, 28, 29, 30, 32~
## $ mother_educ <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ father_educ <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ gender     <fct> female, female, male, female, female, female, male, female~
## $ computer   <fct> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ internet   <fct> no, no, no, no, no, no, yes, no, no, no, no, no, no, no, n~
## $ math       <dbl> 324.35, NA, NA, 235.79, NA, 290.74, NA, 266.31, NA, NA, 29~
## $ read       <dbl> 397.87, 368.41, 294.17, 241.49, 287.16, 307.84, 181.73, 33~
## $ science    <dbl> 345.66, 385.83, 327.94, 341.09, 307.15, 277.04, 279.19, 46~
## $ stu_wgt     <dbl> 2.1600, 2.1600, 2.1600, 2.1600, 2.1600, 2.1600, 2.1600, 2.~
## $ desk       <fct> yes, yes, yes, yes, yes, yes, NA, yes, yes, yes, yes, yes, ~
## $ room       <fct> no, yes, yes, yes, yes, yes, yes, no, yes, yes, yes, yes, ~
## $ dishwasher <fct> no, no, no, no, no, yes, no, yes, no, no, no, no, no, no, ~
## $ television <fct> 1, 2, 2, 1, 2, 1, 1, 1, 1, 2, 1, 1, 1, 2, 1, 1, 3+, 2, ~
## $ computer_n <fct> 3+, 0, 0, 1, 0, NA, 0, 0, NA, 0, 0, 0, 0, 0, 0, 0, 0, 1~
## $ car        <fct> 1, 0, 0, 0, 1, 3+, 0, NA, NA, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0~
## $ book       <fct> 11-50, 1-10, 1-10, 11-50, 11-50, 11-50, 1-10, 1-10, 51-100~
## $ wealth     <dbl> -0.60, -1.84, -1.46, -1.46, -1.13, -0.40, -1.84, -1.57, -1~
## $ escs       <dbl> 0.10575583, -1.42404458, -1.30668386, -0.46246393, -1.2816~
## $ school_uid <chr> "ALB_2000_1001", "ALB_2000_1001", "ALB_2000_1001", "ALB_20~

## Rows: 131,385
## Columns: 14
## $ year      <int> 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2000, 2~
## $ country   <chr> "ALB", "ALB", "ALB", "ALB", "ALB", "ALB", "ALB", "ALB", "ALB", ~
## $ school_id <chr> "1001", "1004", "1005", "1010", "1013", "1017", "1020", ~
```

```
## $ fund_gov      <dbl> 100, 98, 91, 100, 0, 95, 100, 20, 100, 100, 99, 98, 100~
## $ fund_fees     <dbl> 0, 1, 5, 0, 50, 2, 0, 80, 0, 0, 0, 0, 1, 0, 0, NA, 2~
## $ fund_donation <dbl> 0, 1, 2, 0, 30, 3, 0, 0, 0, 0, 1, 2, 0, 2, 0, 0, NA, 8,~
## $ enrol_boys    <dbl> 1191, 334, 403, 114, 250, 771, 103, 510, 558, 591, 470,~
## $ enrol_girls   <dbl> 1176, 479, 600, 201, 248, 626, 162, 741, 736, 873, 837,~
## $ stratio       <dbl> 23.67, 24.64, NA, 22.50, 26.92, 25.40, 8.15, 23.17, 24.~
## $ public_private <fct> public, public, public, public, private, public, public~
## $ staff_shortage <dbl> 0.60, -0.95, -0.17, 1.87, -0.95, 0.27, -0.17, 0.89, -0.~
## $ sch_wgt       <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1~
## $ school_size    <dbl> 2367, 813, 1003, 315, 498, 1397, 265, 1251, 1294, 1464,~
## $ school_uid     <chr> "ALB_2000_1001", "ALB_2000_1004", "ALB_2000_1005", "ALB~
```

2.1.2 Data cleaning

Key Structure and Uniqueness Verification

Before any analysis, it is important to ensure that all key identifiers are unique where required, and that no missing keys exist. This step guarantees integrity for all subsequent merges and analyses.

```
## [1] 0
```

```
## [1] 0
```

Assessment and Handling of Missing Data:

Identifying missing data helps to determine whether to impute, remove, or otherwise handle gaps to avoid introducing bias or losing critical information.

For critical variables (for instance `escs`, `stu_wgt`, performance scores), it is important to consider imputation or filtering.

For non-critical variables, it is possible to consider imputation with a placeholder or removal.

Table 2: Percentage of Missing Values (student)

	Variable	Percent_Missing
dishwasher	dishwasher	51.27
wealth	wealth	26.37
desk	desk	19.52
computer_n	computer_n	11.68
car	car	11.54
father_educ	father_educ	11.06
television	television	10.82
mother_educ	mother_educ	8.85
computer	computer	6.55
room	room	3.44
book	book	3.42
internet	internet	3.18
escs	escs	2.52
math	math	2.30
read	read	1.32
science	science	0.15
gender	gender	0.05
year	year	0.00
country	country	0.00
school_id	school_id	0.00

	Variable	Percent_Missing
student_id	student_id	0.00
stu_wgt	stu_wgt	0.00
school_uid	school_uid	0.00

Table 3: Percentage of Missing Values (school)

	Variable	Percent_Missing
fund_donation	fund_donation	22.41
fund_fees	fund_fees	19.65
stratio	stratio	17.13
sch_wgt	sch_wgt	16.59
fund_gov	fund_gov	12.17
enrol_boys	enrol_boys	11.90
enrol_girls	enrol_girls	11.90
school_size	school_size	11.47
staff_shortage	staff_shortage	10.13
public_private	public_private	8.10
year	year	0.00
country	country	0.00
school_id	school_id	0.00
school_uid	school_uid	0.00

An observation can be made that in the **student_train** dataset there are several critical variables that need filtering such as: **escs**, **math**, **read**, **science** and the gender variables. No critical variables in the **school_train** dataset need to be filtered.

Table 4: Percentage of Missing Values (student)

	Variable	Percent_Missing
dishwasher	dishwasher	51.00
wealth	wealth	25.99
desk	desk	18.72
computer_n	computer_n	10.44
car	car	10.31
television	television	9.57
father_educ	father_educ	7.04
mother_educ	mother_educ	4.74
computer	computer	2.81
book	book	1.87
room	room	1.81
internet	internet	1.51
year	year	0.00
country	country	0.00
school_id	school_id	0.00
student_id	student_id	0.00
gender	gender	0.00
math	math	0.00
read	read	0.00
science	science	0.00
stu_wgt	stu_wgt	0.00
escs	escs	0.00
school_uid	school_uid	0.00

Table 5: Summary Statistics for Math, Reading, and Science Scores

Statistic	Math	Reading	Science
Min.	-53.45	-79.04	-169.93
1st Qu.	387.67	386.75	392.18
Median	461.59	464.28	466.44
Mean	463.08	461.31	467.96
3rd Qu.	537.01	538.39	542.53
Max.	1022.21	1083.51	940.31

Table 6: Summary Statistics for Math, Reading, and Science Scores

Statistic	Math	Reading	Science
Min.	52.38	50.25	51.04
1st Qu.	387.70	386.78	392.23
Median	461.61	464.29	466.44
Mean	463.12	461.36	468.00
3rd Qu.	537.03	538.41	542.55
Max.	962.23	938.68	940.31

Variable	Percent_Missing
----------	-----------------

The lines with the critical missing data have been cleaned.

Variable Type Harmonization:

Ensuring that categorical variables have consistent types and representations (for instance, converting factors to characters) avoids errors in modeling and plotting.

Outlier Detection and Filtering:

Detect and handle outliers, especially for test scores, to avoid undue influence from erroneous or extreme values.

After identifying the potential presence of outliers in the test score distributions, the next step is to clean the dataset by removing extreme or implausible values. Since PISA scores should logically fall within a defined range, any math, reading, or science scores outside the interval 50 to 1000 are considered invalid or likely to result from data entry errors.

By filtering out these values, we ensure that our analysis is not skewed by erroneous or unrealistic data points. This cleaning step helps maintain the integrity and reliability of all subsequent statistical analyses and modeling.

Standardization of Categorical Levels:

Ensuring that all categorical variables (for instance school type) have standardized values (no typos, capitalization mismatches, or ambiguous entries).

Table 7: Number of Schools by Type

School Type	Count
1	3959
2	15276
private	18791
public	82711
NA	10648

Table 8: Number of Schools with Inconsistent ‘public_private’ Label and Government Funding

Check	Number.of.inconsistent.cases
‘private’ with fund_gov > 80	5091
‘public’ with fund_gov < 50	5740

Correct the public_private column based on fund_gov:

To improve the classification of school sector, we recoded the public_private variable based on the percentage of government funding (fund_gov). Schools for which government funding exceeded 80% were classified as “public,” while those with less than 50% government funding were classified as “private.” For schools with government funding between 50% and 80%, the original public_private label was retained. This approach allows for a more objective and consistent distinction between public and private schools in the dataset.

Table 9: Number of Schools with Inconsistent ‘public_private’ Label and Government Funding

Check	Number.of.inconsistent.cases
‘private’ with fund_gov > 80	0
‘public’ with fund_gov < 50	0

Average Performance score for each student:

Creation of a new variable; the average performance score for each student that may enhance predictive power.

Incorporation of Sampling Weights:

All analyses (descriptive statistics, modeling, etc.) should incorporate the stu_wgt sampling weight to produce unbiased, population-representative results.

```
## [1] 452.4942
```

Correct the country code:

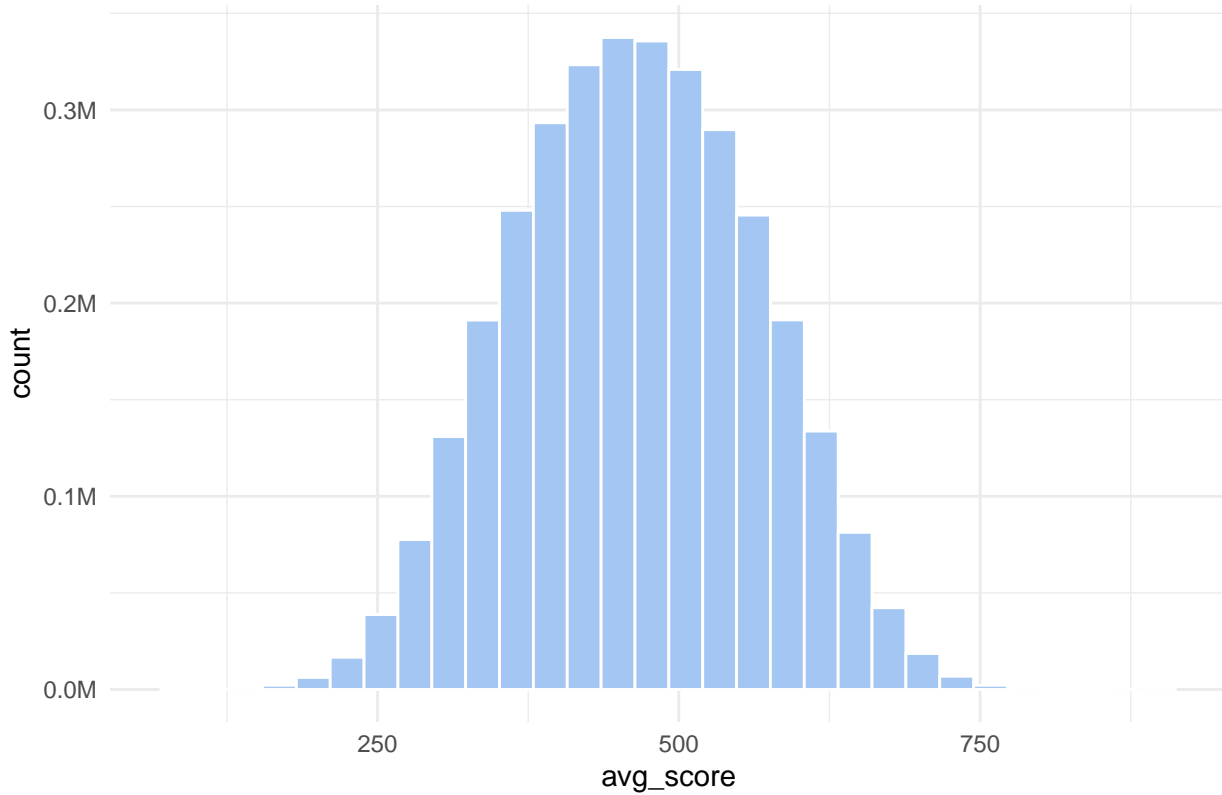
To ensure consistency in country identification across both the student and school datasets, we first standardize all country codes by converting them to uppercase characters. This step helps to avoid issues caused by variations in letter casing (for example, “usa” vs. “USA”). Next, we systematically recode country codes using a predefined mapping. This mapping aligns alternative or legacy codes with their official ISO3 codes (for example, converting “TAP” to “TWN” for Taiwan, “QES” to “ESP” for Spain, and so on). The recoding is applied to both datasets using the recode() function, which replaces each value in the country column according to the specified mapping.

After recoding, the remaining unique country codes are checked in the different datasets, and any codes not included in the mapping are identified. This harmonization process is essential for accurate merging, aggregation, and analysis at the country level.

```
## # A tibble: 0 x 2
## # i 2 variables: country <chr>, n <int>
```

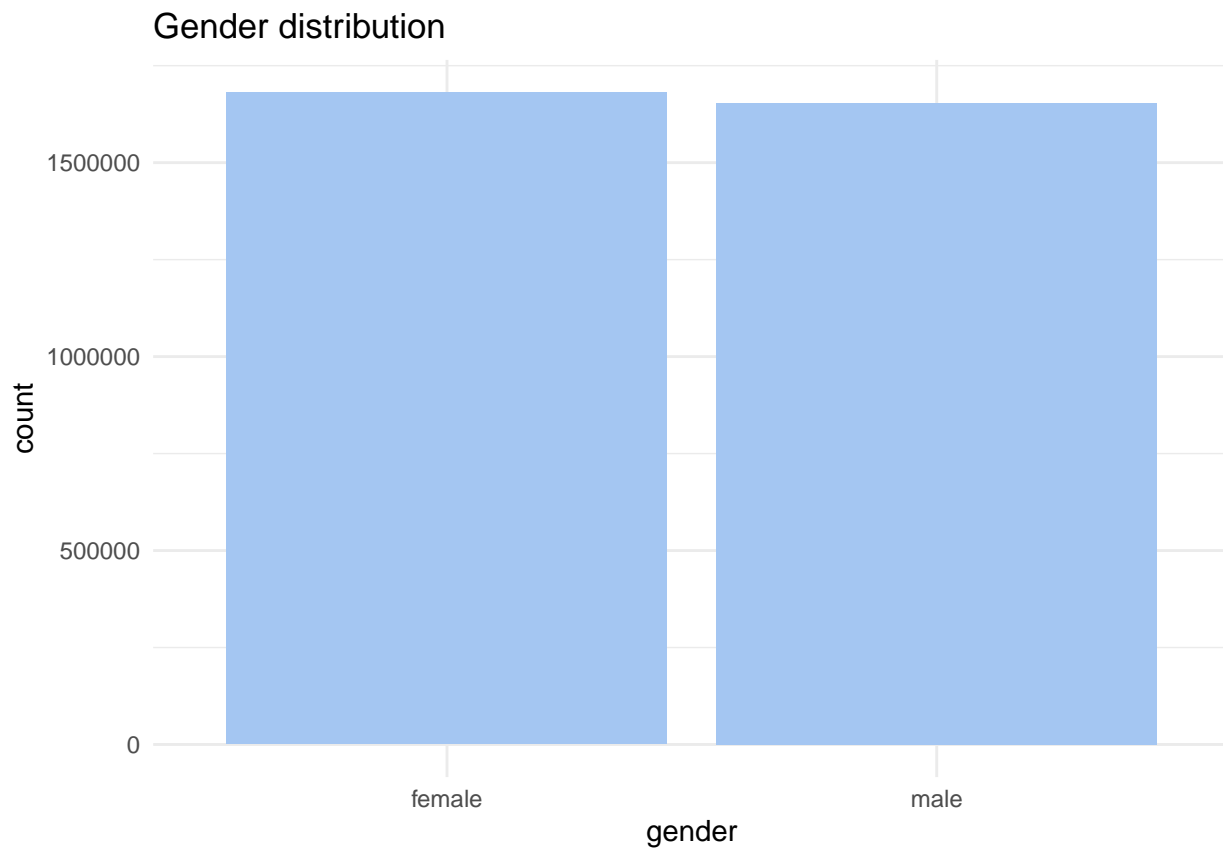
Distribution of the average PISA score :

Distribution of the average score of math, read and science in PISA the st



The histogram displays the distribution of the average PISA scores (math, reading, and science) for all students in the dataset, with the number of students expressed in millions. The distribution appears to be approximately normal (bell-shaped), centered around a mean score of roughly 450–500 points. Most students scored between 350 and 650, with very few students achieving extremely low or extremely high average scores.

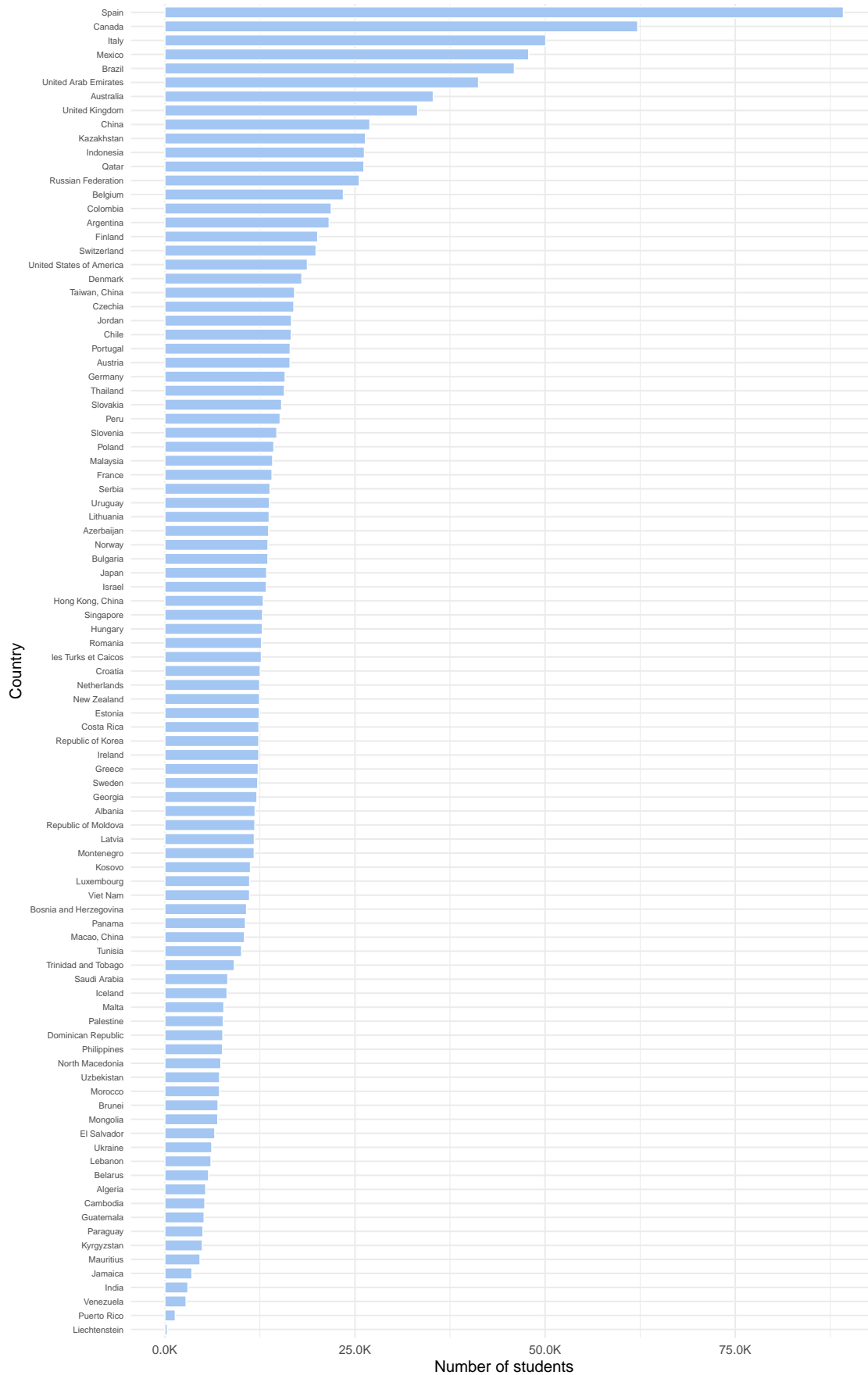
This suggests that the majority of students cluster around the international average, while only a small minority achieve very high or very low scores. The symmetry of the distribution indicates that there is no significant skewness towards higher or lower performance. This is consistent with what would be expected from large-scale educational assessments such as PISA.



There is approximately the same number of female and male which is positive for the robustness of the following studies.

Distribution of the students participating for each countries :

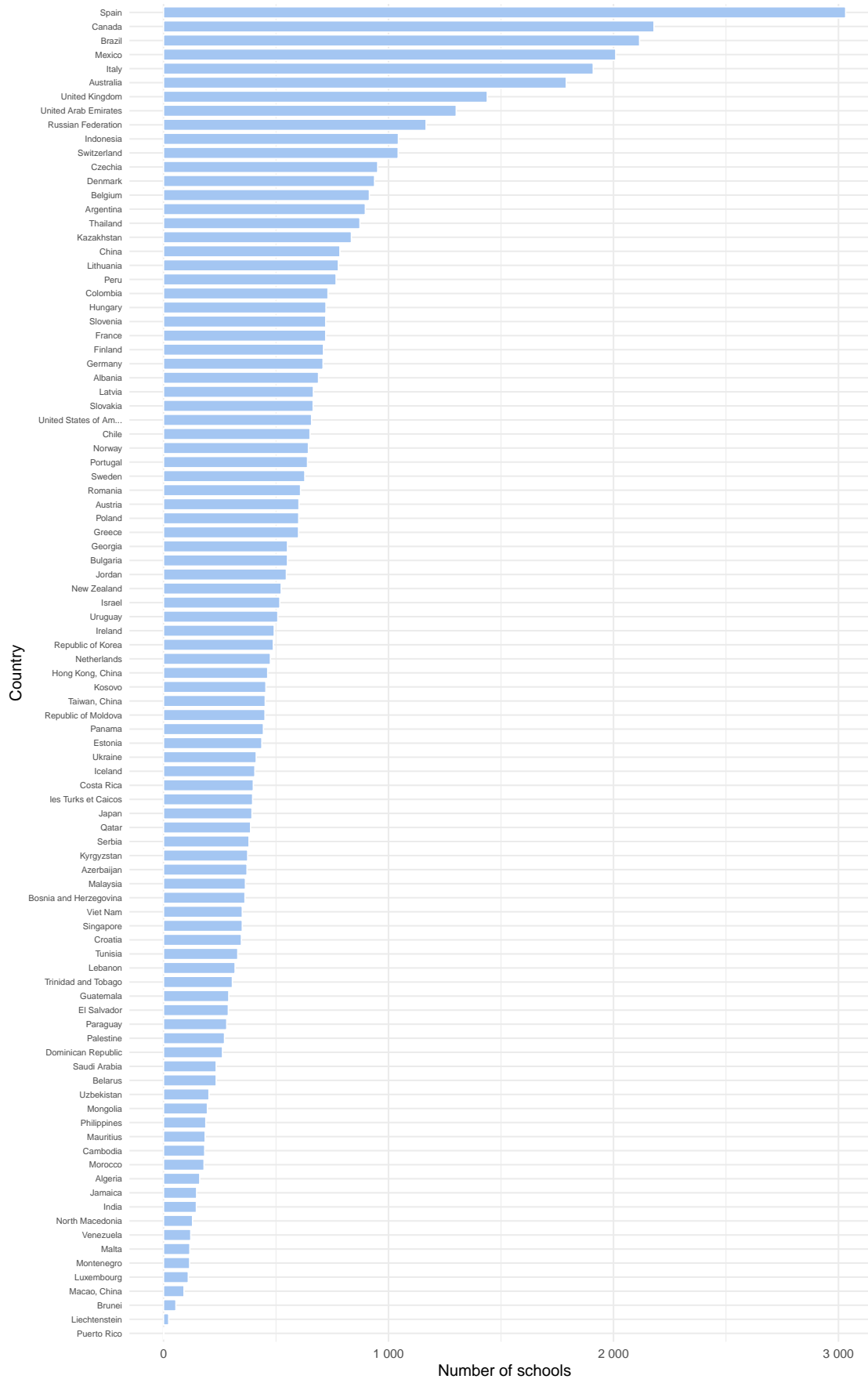
Number of students who participated in PISA by country



This bar chart presents the number of students who participated in the PISA assessment by country. There is significant variation in student participation between countries. Spain, Canada, Italy, Mexico, and Brazil have the highest numbers of students represented, each with tens of thousands of participants. Conversely, several countries, often smaller nations or those with more limited participation—have fewer than 1,000 students included.

Such disparities in participation can reflect differences in country population size, education system organization, sampling methodologies, and engagement with the PISA study. Countries with larger student samples are likely to yield more statistically robust and representative insights, while those with small samples may experience greater sampling error and less generalizability. This variation should be considered when interpreting and comparing student performance across countries.

Number of schools participating in PISA by country



This horizontal bar chart shows the number of schools participating in the PISA assessment by country. There is substantial variation in school participation across countries. Spain, Canada, Brazil, and Mexico have the largest numbers of participating schools, each with over 2,000 schools represented. In contrast, many countries, especially smaller nations or those with lower participation rates—are represented by fewer than 100 schools.

This wide range may reflect differences in country population size, education system structure, sampling strategies, or levels of engagement with the PISA program. Countries with more participating schools are likely to have more diverse student populations represented, whereas those with fewer schools may have less statistical power and greater sampling variability. These differences should be considered when comparing average student outcomes across countries.

2.2 Database exploration and visualization

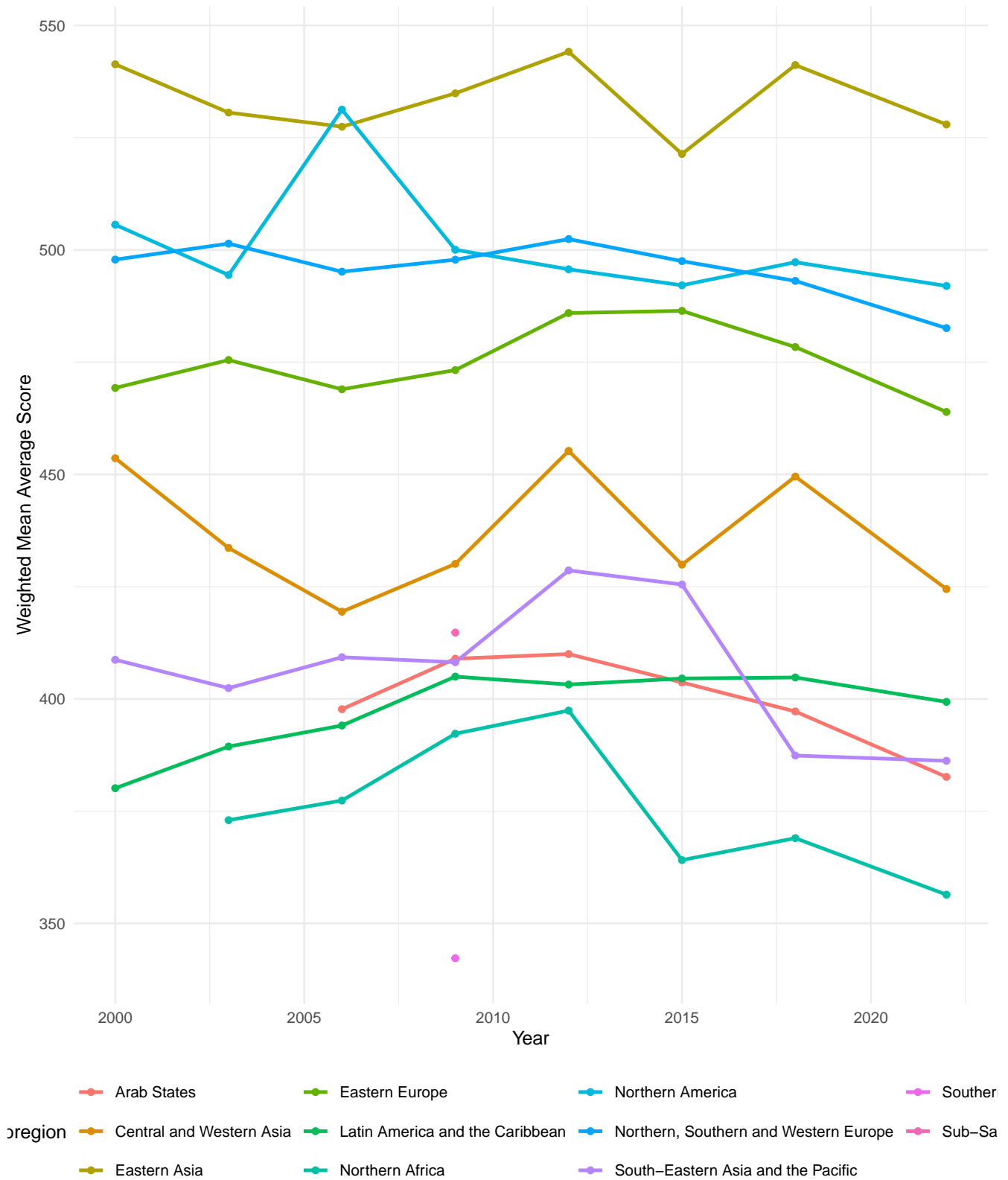
2.2.1 Evolution of the PISA score by Region

To examine regional trends in student achievement, a weighted mean of average student scores was calculated by year and by ILO subregion.

The dataset was constructed by merging student-level performance data with regional classifications based on country codes. Weighted means were computed using each student’s sampling weight (`stu_wgt`) to ensure representativeness.

The resulting summary table was visualized using a line plot, with each subregion represented by a distinct color. This visualization highlights the evolution of average student performance across different world regions over the years covered by the PISA assessments.

Weighted Mean of Average Student Score by Year and Subregion



The graph reveals notable differences in average student performance across ILO subregions over time.

Certain regions, such as Eastern Europe and Northern, Southern and Western Europe, consistently exhibit higher

weighted mean scores, while regions like Northern Africa and Latin America and the Caribbean tend to display lower averages.

Most regions show relative stability in performance trends, though some fluctuations are observed in specific years, which may reflect educational reforms, socio-economic changes, or variations in sample composition.

Overall, the persistence of performance gaps between regions underscores the importance of considering both regional context and policy interventions in global educational analysis.

Table 10: Change in Weighted Mean Average Score by Region (2009–2022)

Region	Weighted Mean 2009	Weighted Mean 2022	Change (%)
Eastern Asia	535	528	-1.3
Central and Western Asia	430	424	-1.4
Latin America and the Caribbean	405	399	-1.5
Northern America	500	492	-1.6
Eastern Europe	473	464	-1.9
Northern, Southern and Western Europe	498	483	-3.0
South-Eastern Asia and the Pacific	408	386	-5.4
Arab States	409	383	-6.4
Northern Africa	392	356	-9.2

2.2.2 GDP and Global Education: Mapping PISA Participation by Income Level

To better understand the global reach and representativeness of PISA assessments, this analysis examines the relationship between countries' economic status and their participation in international educational evaluations. The methodology involved cross-referencing **the complete list of PISA participating countries** with the **World Bank's country classifications** across all income groups.

For each World Bank income category, three key metrics were calculated: the **total number of countries** classified within that income group, **the number of countries participating in PISA**, and the resulting **coverage percentage** - representing what proportion of each economic tier is represented in PISA data.

This approach reveals important patterns about which economies engage in international educational assessment and highlights potential gaps in global educational monitoring.

The findings clearly demonstrate that PISA participation is heavily skewed toward wealthier nations, which aligns with the program's origins and institutional framework under the OECD. The data reveals a stark participation divide across income levels:

High-income countries show the strongest engagement, with 55.6% coverage (50 out of 90 countries participating). This represents the most comprehensive representation within any income group, reflecting both the financial capacity and institutional priorities of developed economies.

Table 11: PISA Country Coverage by World Bank Income Group

Income group	WB countries	PISA countries	Coverage (%)
Low income	28	0	0.0
Lower-middle income	57	15	26.3
Upper-middle income	59	30	50.8
High income	90	50	55.6

Note:

WB Countries: World Bank Countries classification

Upper-middle income countries demonstrate moderate participation at 50.8% coverage (30 out of 59 countries), indicating that many emerging economies prioritize international educational benchmarking as part of their development strategies.

Lower-middle income countries show limited participation with only 26.3% coverage (15 out of 57 countries), suggesting significant barriers to participation among developing economies.

Most strikingly, **low-income countries** have zero participation in PISA assessments, representing a complete absence of the world's poorest economies from this influential international educational evaluation.

This distribution pattern is understandable given PISA's OECD heritage and the substantial financial and administrative resources required for participation.

However, it also means that global educational discourse and policy recommendations are primarily informed by data from non-poor countries, potentially limiting the relevance of findings for the world's most economically disadvantaged education systems.

2.2.3 Educational Mobility Analysis: ESCS Bottom 25% to Top 25% Performance

This analysis examines educational mobility by measuring the proportion of students from the bottom 25% of the Economic, Social and Cultural Status (ESCS) index who achieve scores in the top 25% of academic performance. This indicator serves as a crucial measure of educational equity, as it captures the extent to which socioeconomically disadvantaged students can overcome their initial disadvantages to achieve high academic outcomes.

The data is stratified by World Bank income classifications (High income, Upper-middle income, and Lower-middle income countries) and spans from 2003 to 2022, excluding the year 2000 due to a insufficient number of students in certain groups. This longitudinal approach allows us to observe trends in educational mobility across different economic contexts and time periods.

Table 12: Educational Mobility: ESCS Bottom Quartile to Top Quartile PISA

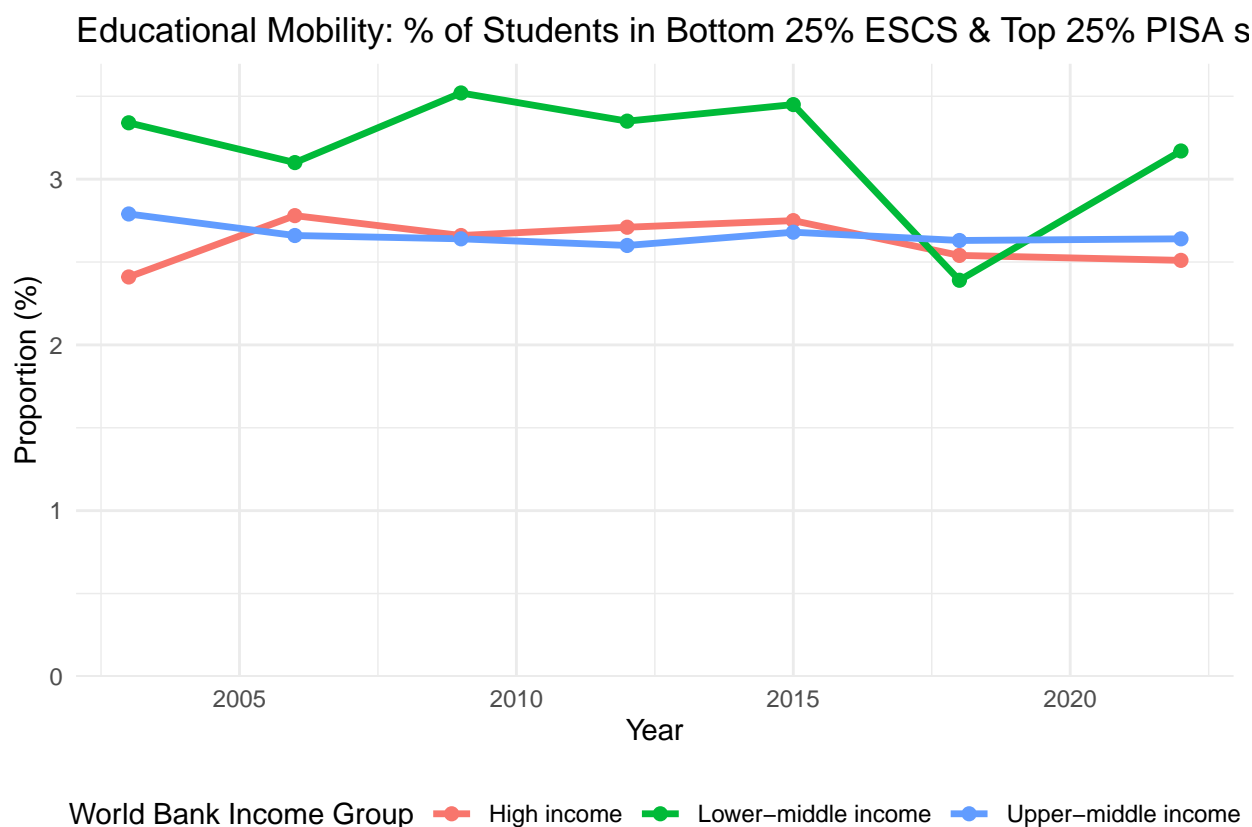
Year	Income Group	Total Students	Students BottomESCS TopPISA	Proportion Perc.	Mean Score Top Quartile
2003	High income	202 040	4 875	2.41	617
2003	Lower-middle income	15 434	515	3.34	467
2003	Upper-middle income	54 560	1 521	2.79	523
2006	High income	277 102	7 696	2.78	606
2006	Lower-middle income	21 092	654	3.10	455
2006	Upper-middle income	90 875	2 418	2.66	517
2009	High income	329 741	8 779	2.66	601
2009	Lower-middle income	19 498	686	3.52	452
2009	Upper-middle income	160 428	4 232	2.64	519
2012	High income	319 644	8 667	2.71	606
2012	Lower-middle income	14 904	499	3.35	522
2012	Upper-middle income	133 803	3 482	2.60	524
2015	High income	346 710	9 524	2.75	602
2015	Lower-middle income	27 368	945	3.45	502
2015	Upper-middle income	129 270	3 459	2.68	524
2018	High income	323 378	8 216	2.54	596
2018	Lower-middle income	37 366	893	2.39	497
2018	Upper-middle income	196 209	5 158	2.63	535
2022	High income	348 686	8 741	2.51	595
2022	Lower-middle income	70 771	2 244	3.17	466
2022	Upper-middle income	168 721	4 457	2.64	504

The comprehensive table reveals several important patterns across income groups:

Lower-Middle Income Countries: This group demonstrates the highest educational mobility rates, with proportions ranging from 2.39% to 3.52%. The peak occurs in 2009 (3.52%), followed by fluctuations including a notable drop in 2018 (2.39%) and recovery by 2022 (3.17%). This pattern suggests that countries in this income bracket may have more fluid social structures or educational systems that are more effective at promoting mobility among disadvantaged students.

Upper-Middle Income Countries: These nations show intermediate mobility rates (2.60% to 2.79%), with relatively stable performance over time. The consistency of these figures suggests a middle ground between the other two groups, possibly reflecting moderate levels of both resources and social stratification.

High-Income Countries: These nations consistently show the lowest rates of educational mobility, with proportions ranging from 2.41% to 2.78%. This paradoxical finding suggests that despite greater overall resources, high-income countries may have more entrenched socioeconomic stratification in their educational systems. The relatively stable trend over time (slight decline from 2015 onwards) indicates persistent inequality despite economic prosperity.



Despite the variations observed across income groups and time periods, a fundamental limitation emerges from this analysis: **the proportion of students from the bottom 25% ESCS who achieve top 25% PISA performance remains consistently low across all regions, fluctuating only between 2.39% and 3.52%.**

This narrow range reveals that **no income group has truly succeeded in enabling substantial educational mobility through their education systems.**

Even the “highest performing” lower-middle income countries at their peak (3.52% in 2009) still leave over 96% of the most disadvantaged students unable to reach top academic performance levels. This suggests that while relative differences between income groups may be statistically meaningful, the absolute levels of educational mobility remain **universally limited** across all economic contexts.

Implications: These findings indicate that educational systems worldwide, regardless of their economic development level, face similar structural challenges in overcoming socioeconomic disadvantage. The persistent ceiling

of approximately 3.5% suggests that **educational mobility remains an exceptional rather than systematic outcome**, highlighting the need for more fundamental reforms to achieve meaningful equity in educational opportunities.

The visual representation of these data reveals several striking patterns:

Divergent Trajectories: The most remarkable feature is the dramatic divergence that occurs around 2018. While high-income and upper-middle income countries maintain relatively stable, low levels of educational mobility, lower-middle income countries experience significant volatility.

Stability vs. Volatility: High-income and upper-middle income countries show remarkable stability across the time period, suggesting established educational systems with consistent (though limited) mobility outcomes. In contrast, lower-middle income countries exhibit much greater variability, which could reflect ongoing educational reforms, economic instability, or different measurement sensitivities.

Counter-Intuitive Findings: The data challenges conventional assumptions about the relationship between national wealth and educational equity. The fact that lower-middle income countries generally outperform high-income countries in educational mobility suggests that factors beyond simple resource availability; such as social structure, educational policy, or cultural values, play crucial roles in determining educational outcomes for disadvantaged students.

These findings have significant implications for educational policy and our understanding of equity. They suggest that educational mobility is not simply a function of national wealth, but rather depends on complex interactions between social structures, educational policies, and cultural factors. The superior performance of lower-middle income countries in promoting educational mobility may offer valuable lessons for higher-income nations seeking to improve equity in their educational systems.

2.2.4 PISA Performance in Lower-Income Countries: The Socioeconomic Achievement Gap Across Time

This analysis examines PISA performance trends for the countries classified as lower-middle income countries by the World Bank. This analysis gives a specific focus on socioeconomic disparities within educational systems. The visualization presents weighted mean scores across three assessment cycles (2003, 2018, and 2022) for countries participating in PISA while classified as lower-middle income economies.

Change in Mean Scores by Social Background (Lower–middle income countries)

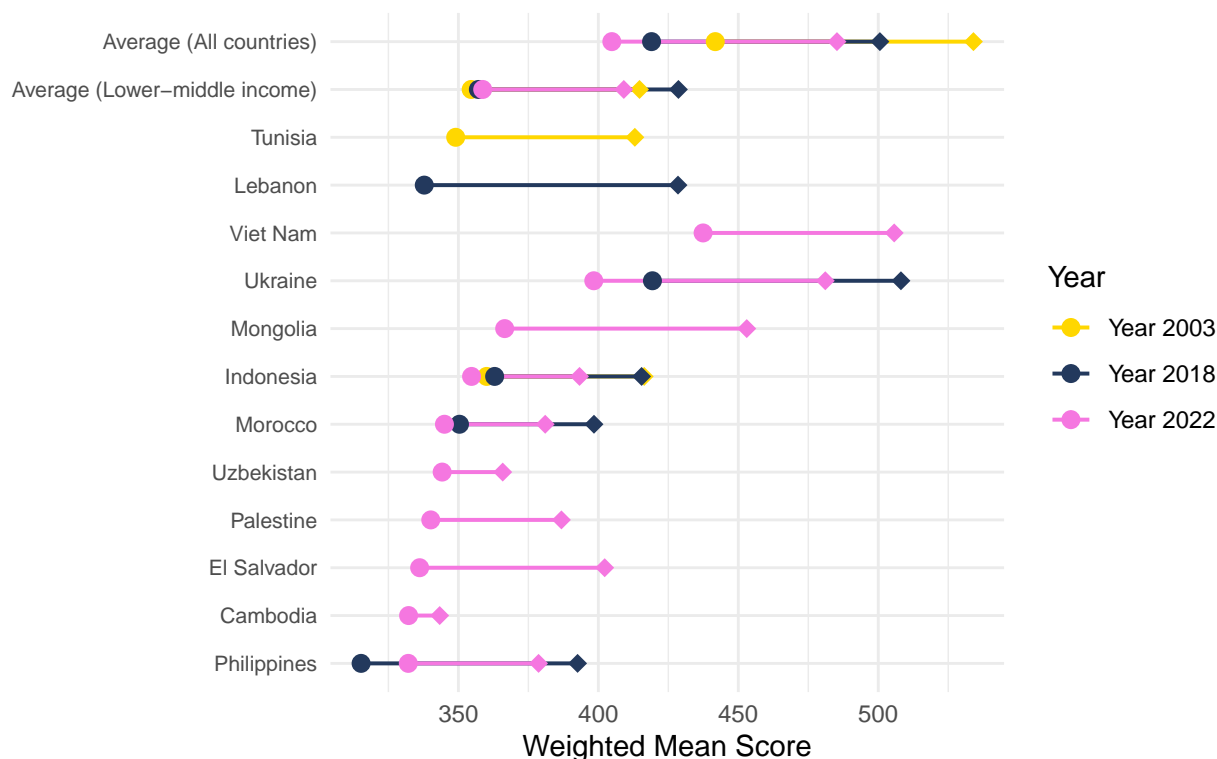


Table 13: Difference in Mean Scores between Top and Bottom ESCS Quartiles (Most Recent Year, Lower-middle Income Countries)

Country	Q1 Mean Score (Bottom 25% ESCS)	Q4 Mean Score (Top 25% ESCS)	Difference (Q4 - Q1)	Year
Lebanon	338	428	91	2018
Mongolia	367	453	86	2022
Ukraine	398	481	83	2022
Viet Nam	437	506	68	2022
El Salvador	336	402	66	2022
Tunisia	349	413	64	2003
Philippines	332	379	47	2022
Palestine	340	387	47	2022
Indonesia	355	393	39	2022
Morocco	345	381	36	2022
Uzbekistan	344	366	22	2022
Cambodia	332	343	11	2022

The data reveals dramatic differences in educational equity across lower-middle income countries, with socioeconomic gaps ranging from as narrow as 11 points in **Cambodia** to as wide as 91 points in **Lebanon**.

Lebanon stands out with the largest achievement gap (91 points) between the bottom and top quartiles, indicating severe educational inequality despite being classified in the same income category as more equitable systems. **Mongolia** and **Ukraine** also show substantial gaps of 86 and 83 points respectively, suggesting that higher-performing students significantly outpace their disadvantaged peers.

In contrast, **Cambodia** demonstrates remarkable equity with only an 11-point gap, followed by **Uzbekistan** (22

points) and **Morocco** (36 points). These countries appear to have more compressed achievement distributions, though this may reflect both lower overall performance levels and different educational system structures.

Vietnam presents an interesting case with a moderate 68-point gap ; while achieving relatively high absolute performance levels, it still maintains reasonable equity compared to countries like **Lebanon** or **Mongolia**. This suggests **Vietnam** has managed to raise overall achievement without creating excessive disparities. The **Philippines** and **Occupied Palestinian Territory** both show 47-point gaps, representing a middle ground in this group. These moderate gaps suggest some inequality but not at the extreme levels seen in **Lebanon** or **Mongolia**.

This analysis demonstrates that within the lower-middle income classification, educational equity varies tremendously, with some countries achieving relatively compressed achievement distributions while others exhibit gaps comparable to more stratified educational systems.

2.2.5 PISA Performance in Upper-Income Countries: The Socioeconomic Achievement Gap Across Time

The below analysis focuses on three key time points: 2003, 2018, and 2022. The data represents weighted mean scores by ESCS quartiles (bottom 25% vs top 25%). The achievement gaps are calculated as the difference between top and bottom quartile performance. The countries included here are based on World Bank Upper-middle income classification.

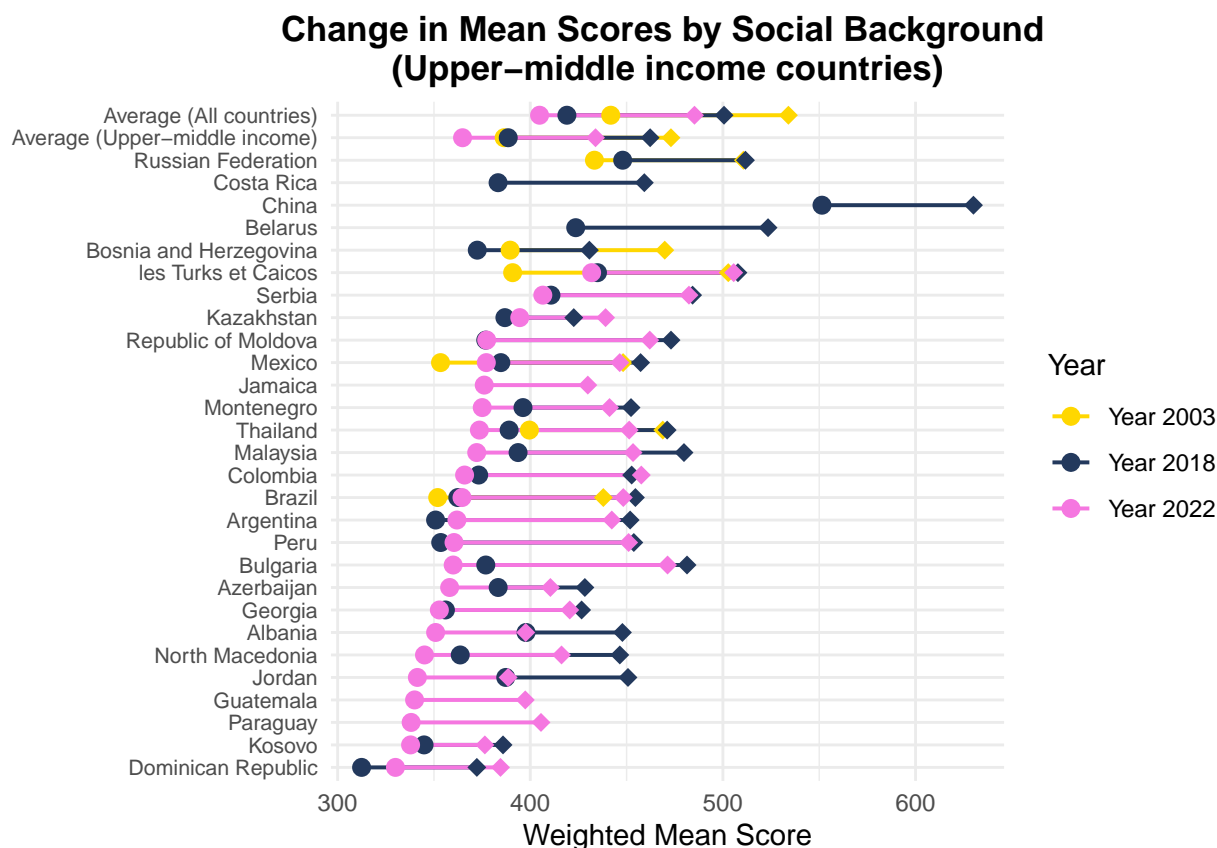


Table 14: Difference in Mean Scores between Top and Bottom ESCS Quartiles (Most Recent Year, Upper-middle Income Countries)

Country	Q1 Mean Score (Bottom 25% ESCS)	Q4 Mean Score (Top 25% ESCS)	Difference (Q4 - Q1)	Year
Bulgaria	360	471	111	2022
Belarus	424	523	100	2018

Country	Q1 Mean Score (Bottom 25% ESCS)	Q4 Mean Score (Top 25% ESCS)	Difference (Q4 - Q1)	Year
Colombia	366	458	92	2022
Peru	360	451	91	2022
Republic of Moldova	377	462	85	2022
Brazil	365	448	84	2022
Malaysia	372	453	81	2022
Argentina	362	442	80	2022
China	551	630	79	2018
Thailand	374	451	78	2022
Costa Rica	383	459	76	2018
Serbia	407	482	76	2022
les Turks et Caicos	432	506	74	2022
North Macedonia	345	416	71	2022
Mexico	377	446	69	2022
Georgia	353	420	68	2022
Paraguay	338	406	67	2022
Montenegro	375	441	66	2022
Russian Federation	448	512	64	2018
Bosnia and Herzegovina	372	431	58	2018
Guatemala	340	397	57	2022
Dominican Republic	330	385	55	2022
Jamaica	376	430	54	2022
Azerbaijan	358	410	52	2022
Albania	351	398	47	2022
Jordan	341	389	47	2022
Kazakhstan	395	439	45	2022
Kosovo	338	376	39	2022

Persistent Achievement Gaps:

Upper-middle income countries demonstrate substantial and persistent achievement gaps between students from different socioeconomic backgrounds. The performance differences between the bottom 25% and top 25% of students by ESCS (Economic, Social and Cultural Status) range from approximately 40 to over 110 points across countries and measurement years.

Temporal Evolution:

The data reveals varying patterns of change across the three time points (2003, 2018, and 2022). Many countries show notable shifts in performance levels and gap sizes, suggesting that educational systems in these nations are experiencing dynamic changes rather than stagnation.

Country-Specific Observations:

Several countries show incomplete data across the three time periods. Countries such as **Costa Rica**, **China**, and **Belarus** have limited participation in certain PISA cycles, affecting longitudinal comparisons.

Highest Performing Countries:

China leads in absolute performance with the top 25% ESCS students scoring 630 points, though with a substantial 79-point gap. **Turks and Caicos Islands** shows exceptional performance for advantaged students (506 points) but with a concerning 74-point gap. **Russian Federation** (2018 data) demonstrates strong performance with top students at 512 points and a moderate 64-point gap. **Belarus** (2018 data) achieves high performance (523 points for top 25%) but maintains a significant 100-point gap.

Most Equitable Systems:

Kosovo displays the smallest achievement gap at just 39 points, though both groups perform at lower absolute levels (338-376 points). **Kazakhstan** shows strong equity with a 45-point gap and reasonable performance levels (395-

439 points). **Jordan** and **Albania** demonstrate relatively small gaps (47 points each) with moderate performance levels.

These countries prove that more equitable outcomes are achievable, though often at lower absolute performance levels.

Largest Achievement Gaps:

Bulgaria exhibits the widest gap at 111 points (360-471 points), indicating severe inequality despite decent top-performer achievement. **Belarus** shows a 100-point gap, combining high performance with significant inequality. **Colombia**(92-point gap) and **Peru** (91-point gap) demonstrate substantial inequities in South American contexts.

These large gaps suggest deeply embedded structural inequalities that persist despite national economic development.

Countries Showing Improvement:

Mexico demonstrates substantial progress across both socioeconomic groups from 2003 to 2022, with particularly notable gains for disadvantaged students. **Brazil** exhibits modest improvements, particularly for the bottom 25% of students. **Argentina** shows signs of convergence between groups over time, suggesting improving equity without sacrificing overall performance.

Countries with Declining Performance:

Several countries, including **Montenegro**, **Malaysia** and **Thailand**, show concerning declines in performance for both groups between 2018 and 2022.

Stability Patterns:

Countries like **Serbia** and **Georgia** show remarkable consistency across time periods, maintaining both their performance levels and gap sizes. This stability could indicate either effective policy consistency or systemic rigidity.

Group Average Performance: The upper-middle income country average shows:

Top 25% ESCS students: range from 376 points (Kosovo) to 630 points (China), with most countries between 400-500 points. **Bottom 25% ESCS students:** range from 330 points (Dominican Republic) to 551 points (China), with most countries between 340-400 points. **Achievement gaps:** range from 39 points (Kosovo) to 111 points (Bulgaria), with an average gap of approximately 67 points across all countries. **The median gap** is around 69 points, indicating that most countries cluster around this moderate inequality level.

Critical Insights and Policy Implications:

The Middle-Income Challenge: Upper-middle income countries face a distinctive educational challenge; having sufficient economic resources to support quality education while struggling with persistent inequality. The data suggests that economic development alone is insufficient to ensure equitable educational outcomes.

Successful Models: Countries like **Russia**, **China** and **Costa Rica** demonstrate that upper-middle income status can support both relatively high achievement and manageable equity gaps, providing potential models for policy learning. We can question whether for China and Russia the schools that have participated in PISA are representative of the true level of education within the country or only in the wealthiest regions.

Dynamic vs. Static Systems: This data reveals that many upper-middle income countries are experiencing significant changes in their educational outcomes, both positive and negative. This dynamism suggests that policy interventions can have measurable impacts.

Resource Allocation Implications: The variation in both absolute performance and equity outcomes within this income group suggests that how resources are allocated and policies are implemented may be more critical than the absolute level of national wealth in determining educational outcomes.

Future Trajectory Concerns: The mixed patterns of improvement and decline across countries highlight the need for sustained policy attention to both excellence and equity in educational systems, as gains can be reversed without consistent effort.

2.2.6 PISA Performance in High Income Countries: The Socioeconomic Achievement Gap Across Time

The below analysis focuses on three key time points: 2003, 2018, and 2022. The data represents weighted mean scores by ESCS quartiles (bottom 25% vs top 25%). The achievement gaps are calculated as the difference between top and bottom quartile performance. The countries included here are based on World Bank High income classification.

Change in Mean Scores by Social Background (High income countries)

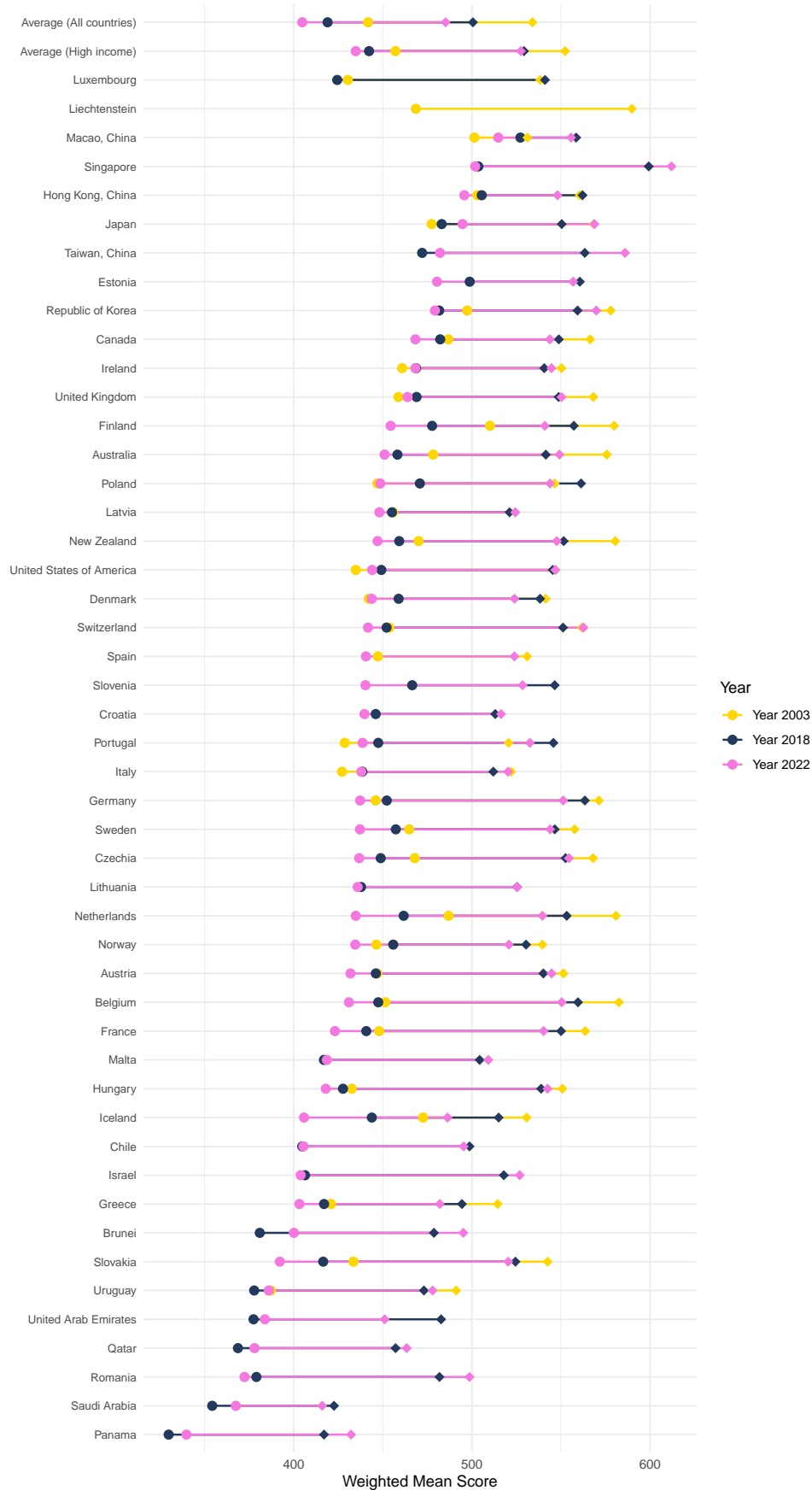


Table 15: Difference in Mean Scores between Top and Bottom ESCS Quartiles illustrates the educational equity gap (Most Recent Year, High Income Countries)

Country	Q1 Mean Score (Bottom 25% ESCS)	Q4 Mean Score (Top 25% ESCS)	Difference (Q4 - Q1)	Year
Slovakia	392	520	128	2022
Romania	372	499	126	2022
Hungary	418	542	125	2022
Israel	404	527	123	2022
Switzerland	442	563	121	2022
Liechtenstein	469	590	121	2003
Belgium	431	550	119	2022
Czechia	437	554	118	2022
France	423	540	117	2022
Luxembourg	424	541	117	2018
Germany	437	551	114	2022
Austria	432	545	113	2022
Singapore	502	612	110	2022
Sweden	437	544	107	2022
Netherlands	435	540	105	2022
Taiwan, China	482	586	104	2022
United States of America	444	547	103	2022
New Zealand	447	548	101	2022
Australia	451	549	98	2022
Brunei	400	495	95	2022
Poland	449	544	95	2022
Portugal	439	533	94	2022
Panama	340	432	92	2022
Uruguay	386	478	92	2022
Republic of Korea	479	570	91	2022
Malta	419	509	91	2022
Chile	405	495	90	2022
Lithuania	436	525	89	2022
Slovenia	440	528	88	2022
Finland	454	541	87	2022
United Kingdom	464	550	86	2022
Norway	435	521	86	2022
Qatar	378	463	85	2022
Spain	441	524	83	2022
Italy	438	520	82	2022
Iceland	406	486	81	2022
Denmark	444	524	80	2022
Greece	403	482	79	2022
Croatia	440	516	77	2022
Estonia	480	557	76	2022
Ireland	468	545	76	2022
Latvia	448	524	76	2022
Canada	468	544	75	2022
Japan	495	569	74	2022
United Arab Emirates	384	451	67	2022
Hong Kong, China	496	548	52	2022
Saudi Arabia	368	416	48	2022
Macao, China	515	556	41	2022

Overall Patterns and Characteristics

Substantial Achievement Gaps:

High-income countries demonstrate some of the largest achievement gaps globally between students from different socioeconomic backgrounds. The performance differences between the bottom 25% and top 25% of students by ESCS (Economic, Social and Cultural Status) range dramatically from 41 points to 128 points, with most countries showing gaps between 80-120 points.

Economic Resources vs. Educational Equity:

Paradoxically, despite having the greatest economic resources, high-income countries often struggle more with educational equity than their lower-income counterparts. This suggests that wealth alone does not guarantee equitable educational outcomes and may, in some cases, exacerbate existing inequalities.

Country-Specific Performance Analysis

Highest Performing Systems:

Singapore leads in absolute performance with top 25% ESCS students scoring 612 points, though maintaining a substantial 110-point gap. **Liechtenstein** (2003 data) shows exceptional performance for advantaged students (590 points) with a concerning 121-point gap. **Taiwan** demonstrates strong performance with top students at 586 points and a moderate 104-point gap. **Republic of Korea** achieves high performance (570 points for top 25%) with a relatively manageable 91-point gap.

Most Equitable High-Income Systems:

Macao displays the smallest achievement gap at just 41 points while maintaining high absolute performance (515-556 points), representing an exceptional model of high achievement with equity. **Saudi Arabia** shows strong equity with a 48-point gap, though at lower absolute performance levels (368-416 points). **Hong Kong** demonstrates excellent equity with only a 52-point gap while achieving very high performance (496-548 points). **United Arab Emirates** maintains a moderate 67-point gap with reasonable performance levels.

Largest Achievement Gaps:

Slovakia exhibits the widest gap at 128 points (392-520 points), indicating severe inequality despite moderate performance levels. **Romania** shows a 126-point gap, combining moderate performance with significant inequality. **Hungary** demonstrates a 125-point gap (418-542 points), suggesting deeply embedded educational stratification. **Israel** maintains a 123-point gap, indicating substantial inequities despite strong top-performer achievement.

Regional and System-Type Patterns

East Asian Excellence with Variation:

East Asian systems show remarkable diversity: **Macao** and **Hong Kong** achieve both high performance and equity, while **Taiwan** maintains high performance with moderate equity. **Japan** demonstrates strong performance (495-569 points) with a reasonable 74-point gap. **Republic of Korea** balances high achievement with relatively manageable inequality.

European Stratification:

Many European countries, particularly in Central and Eastern Europe (**Slovakia**, **Romania**, **Hungary**), show concerning levels of educational inequality. Western European nations like **Germany** (114-point gap) and **France** (117-point gap) also struggle with significant achievement gaps. Nordic countries show good results: While **Sweden** shows larger disparities (107-point gap), the other Nordic countries such as **Finland**, **Normay**, **Iceland** and **Denmark** maintain moderate equity (87-80 gap) and above average performance of the high income countries except for Iceland.

Anglo-Saxon Countries:

Australia (98-point gap), **New Zealand** (101-point gap), and **Canada** (75-point gap) demonstrate relatively moderate inequality levels. **United States** shows a 103-point gap, indicating persistent educational stratification. **United Kingdom** maintains an 86-point gap with strong absolute performance.

Temporal Evolution Patterns

Recent Performance Decline (2018-2022):

A concerning trend emerges across many high-income countries showing declining performance between 2018 and 2022 for both socioeconomic groups. This decline affects both advantaged and disadvantaged students, suggesting systemic challenges rather than equity-specific issues. The COVID-19 pandemic period coincides with this decline, potentially indicating disrupted learning processes and educational delivery systems. Countries that previously showed strong performance trajectories have experienced notable setbacks, raising questions about the resilience of high-income educational systems.

Persistent Inequality:

Many high-income countries show remarkable stability in their achievement gaps over time, suggesting entrenched systemic inequalities. Countries like **Switzerland, Hungary, Germany, France, and Belgium** maintain consistently large gaps across measurement periods

Performance Volatility:

Some countries show fluctuating patterns, indicating that educational equity can be influenced by policy changes and economic conditions. The consistency of gaps in many countries suggests that structural factors may be more influential than cyclical policy changes.

System Resilience Concerns:

The widespread 2018-2022 decline raises questions about the resilience and adaptability of high-income educational systems. Some countries show greater volatility in their performance trajectories, while others maintain more stable (though declining) patterns. The consistency of achievement gaps even during periods of overall decline suggests that structural inequalities are deeply embedded in these systems.

Comparative Analysis

Performance Ranges:

Top 25% ESCS students: range from 416 points (Saudi Arabia) to 612 points (Singapore), with most countries between 520-570 points **Bottom 25% ESCS students:** range from 340 points (Panama) to 515 points (Macao), with most countries between 400-470 points **Achievement gaps:** range from 41 points (Macao) to 128 points (Slovakia), with a median gap of approximately 92 points

High-Income Country Average Performance:

The high-income group demonstrates higher absolute performance than other income groups but also larger average achievement gaps, confirming the equity-excellence tension in wealthy nations.

Critical Insights and Policy Implications

The Equity Paradox: High-income countries face a fundamental challenge—their economic advantages do not automatically translate to educational equity. In fact, some of the largest achievement gaps globally are found in the wealthiest nations.

Successful Equity Models: Macao, Hong Kong, and Japan demonstrate that high-income status can support both strong performance and reasonable equity, providing valuable models for policy learning.

Resource vs. Structure: The variation in outcomes among high-income countries suggests that how educational systems are structured and how resources are distributed may be more critical than absolute resource levels.

Policy Intervention Potential: The existence of high-performing, equitable systems within the high-income group proves that wealth-based educational stratification is not inevitable, indicating that targeted policy interventions can address inequality while maintaining excellence.

Long-term Sustainability Concerns: The persistence of large achievement gaps in many wealthy nations raises questions about social cohesion and the long-term sustainability of educational advantages, particularly in increasingly diverse societies.

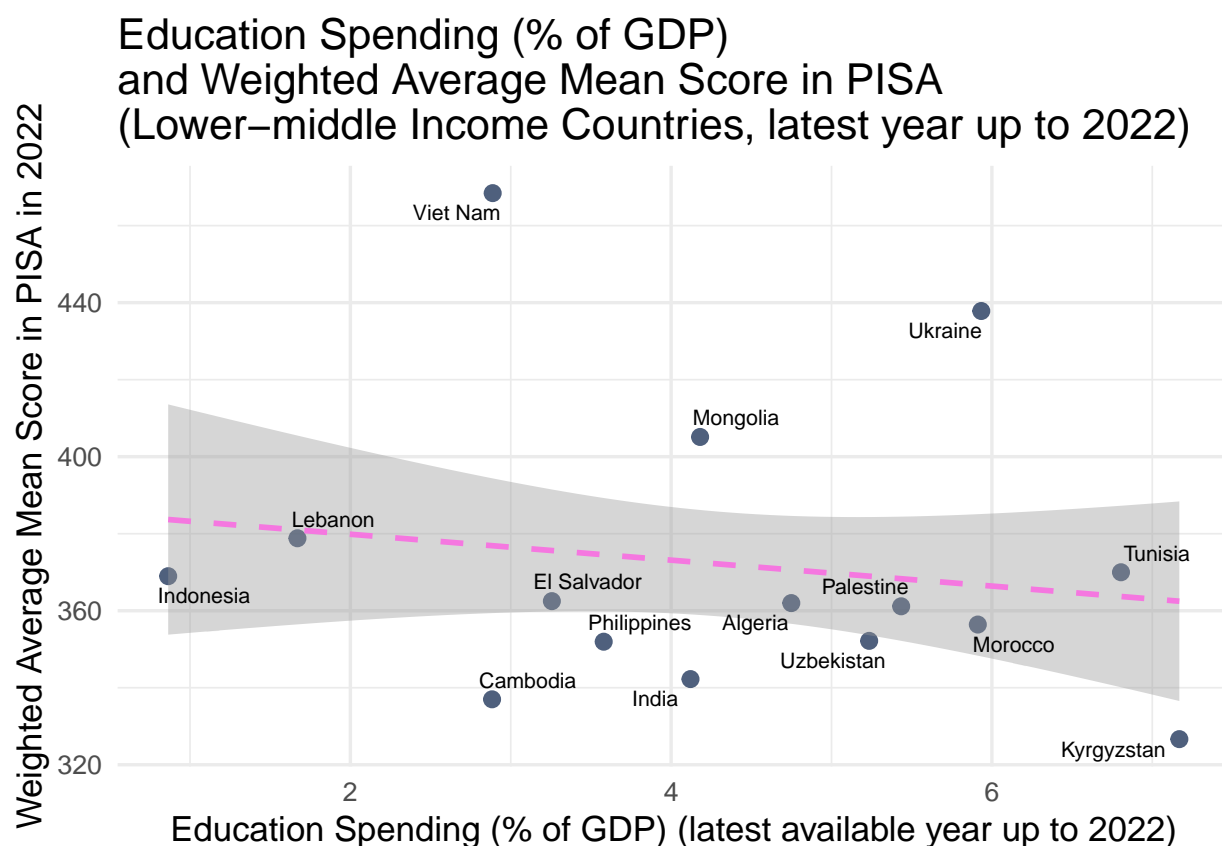
2.2.7 Educational investment and educational outcomes for Lower-middle Income

This analysis employs a scatter plot design to examine the relationship between **educational investment** (measured as education spending as a percentage of GDP from UNESCO) and **educational outcomes** (weighted

average mean PISA scores) across lower-middle income countries in 2022. The methodology includes several important considerations:

Measurement Approach: The use of education spending as a percentage of GDP provides a standardized metric that accounts for economic scale differences between countries, making cross-national comparisons more meaningful than absolute spending figures. The weighted average mean PISA score offers a comprehensive measure of educational achievement across multiple domains.

Sample Selection: Focusing exclusively on lower-middle income countries creates a controlled comparison group, reducing the confounding effects of vastly different economic development levels that might obscure the spending-achievement relationship.



The correlation analysis reveals a **weak positive relationship between education spending and PISA performance among lower-middle income countries**, with several notable patterns and outliers.

Limited Correlation Strength: The relatively flat trend line and wide confidence interval indicate that education spending as a percentage of GDP explains only a modest portion of the variation in educational achievement. This suggests that other factors beyond financial investment play crucial roles in determining educational outcomes.

Exceptional Efficiency Cases: **Vietnam** emerges as the most remarkable example, achieving the highest PISA scores (468) with one of the lowest spending rates (2.89% of GDP). This exceptional performance-to-investment ratio suggests highly efficient resource utilization and effective educational policies that maximize learning outcomes per dollar invested.

High Investment, High Performance: **Ukraine** demonstrates strong performance (438) with the highest spending level (5.93% of GDP), representing a different but successful approach that combines substantial investment with effective implementation.

High Investment, Moderate Returns: **Tunisia** and **Kyrgyzstan** invest a high percentage in GDP (superior to 6%) but have deceiving results. **Morocco** shows concerning efficiency patterns, spending 5.91% of GDP—nearly as much as Ukraine—but achieving significantly lower scores (356). This suggests potential inefficiencies in resource allocation or educational system effectiveness that warrant investigation.

Low Investment Outliers: **Indonesia** presents the most extreme efficiency challenge, achieving only moderate scores (369.0) with the lowest spending rate (0.86% of GDP). While the low investment partly explains the performance, the efficiency gap with Vietnam is striking.

Investment-Performance Paradoxes: **Mongolia** (405 points, 4.18% spending) significantly outperforms countries with similar or higher investment levels like **Uzbekistan** (352.2 points, 5.23% spending), highlighting substantial differences in educational system effectiveness.

Clustering Patterns: Most countries cluster in the 350-370 score range despite varying spending levels from 2.9% to 5.9% of GDP, suggesting potential systemic constraints that limit the impact of additional financial investment alone.

Policy Implications: The data strongly suggests that educational spending efficiency varies dramatically across countries. **Vietnam's** model deserves particular study, as it demonstrates that strategic resource deployment can yield world-class results even with modest investment levels. Countries like **Morocco** and **Uzbekistan** may need comprehensive educational system reforms before additional funding can translate into proportional achievement gains.

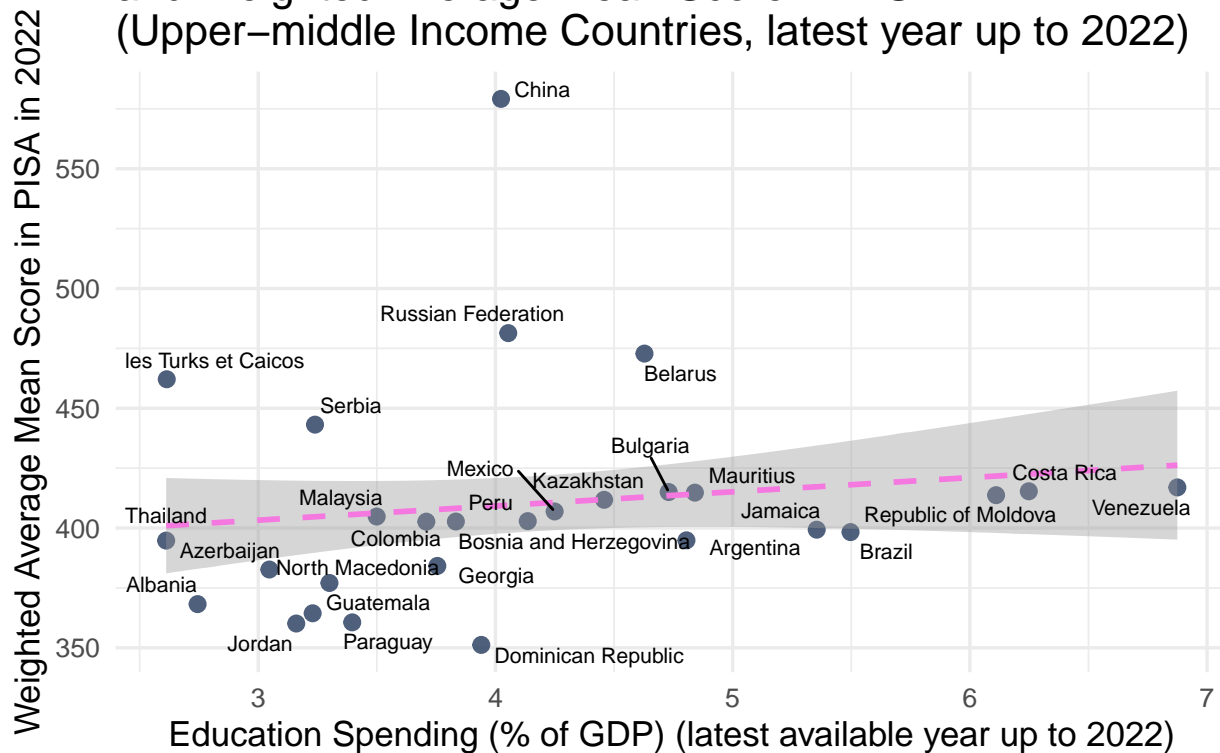
2.2.8 Educational investment and educational outcomes for Upper-middle Income

This analysis employs a scatter plot design to examine the relationship between **educational investment** (measured as education spending as a percentage of GDP from UNESCO) and **educational outcomes** (weighted average mean PISA scores) across Higher-middle income countries in 2022. The methodology includes several important considerations:

Measurement Approach: The use of education spending as a percentage of GDP provides a standardized metric that accounts for economic scale differences between countries, making cross-national comparisons more meaningful than absolute spending figures. The weighted average mean PISA score offers a comprehensive measure of educational achievement across multiple domains.

Sample Selection: Focusing exclusively on Upper-middle income countries creates a controlled comparison group, reducing the confounding effects of vastly different economic development levels that might obscure the spending-achievement relationship.

Education Spending (% of GDP) and Weighted Average Mean Score in PISA (Upper–middle Income Countries, latest year up to 2022)



The correlation analysis reveals a **weak positive relationship** between education spending and PISA performance among upper-middle income countries, with notable patterns of efficiency and outlier performance:

Limited Overall Correlation: The relatively flat trend line and wide confidence interval indicate that education spending explains only a modest portion of achievement variation. This suggests that factors beyond financial investment—such as system efficiency, teacher quality, and policy implementation—play crucial roles in determining outcomes.

Exceptional Efficiency Leaders: **Turks and Caicos Islands** (462 points, 2.61% GDP) and **Serbia** (443 points, 3.24% GDP) emerge as remarkable efficiency exemplars, achieving the highest PISA scores with among the lowest spending rates in this group. This suggests highly effective resource utilization and educational system optimization.

High Investment, Strong Returns: **The Republic of Moldova**, **Venezuela**, **Costa Rica** demonstrate that substantial investment (6.11% GDP) can yield good results, representing the highest spending level in the sample with correspondingly solid performance.

Investment-Performance Disconnects: **Jamaica** and **Brazil** present concerning efficiency patterns, spending more than 5.35% of GDP—among the highest rates—while achieving moderate scores (below 400 points). This suggests potential resource allocation inefficiencies or systemic challenges that prevent investment from translating into proportional achievement gains.

Moderate Efficiency Cases: Countries like **Kazakhstan** (412 points, 4.46% GDP) and **Malaysia** (405 points, 3.50% GDP) show reasonable but not exceptional efficiency, achieving solid performance with moderate investment levels.

Underperforming Despite Resources: Several countries including **Dominican Republic** (351 points, 3.94% GDP), **Jordan** (360 points, 3.16% GDP), and **Paraguay** (361 points, 3.40% GDP) show concerning patterns where moderate-to-substantial investments yield relatively low achievement outcomes.

Performance Clustering: Most countries cluster in the 350-410 score range despite spending variations from 2.6% to 6.1% of GDP, suggesting potential systemic constraints or ceiling effects that limit the impact of additional financial investment alone.

Policy Implications:

The data demonstrates dramatic efficiency variations across upper-middle income countries. **Turks and Caicos Islands'** and **Serbia's** models merit detailed study as they achieve exceptional results with relatively modest investments.

Countries like **Jamaica** and **Dominican Republic** may require comprehensive educational system reforms and improved resource allocation strategies before additional funding can generate proportional achievement improvements. The weak overall correlation emphasizes that strategic implementation and system effectiveness matter more than spending levels alone.

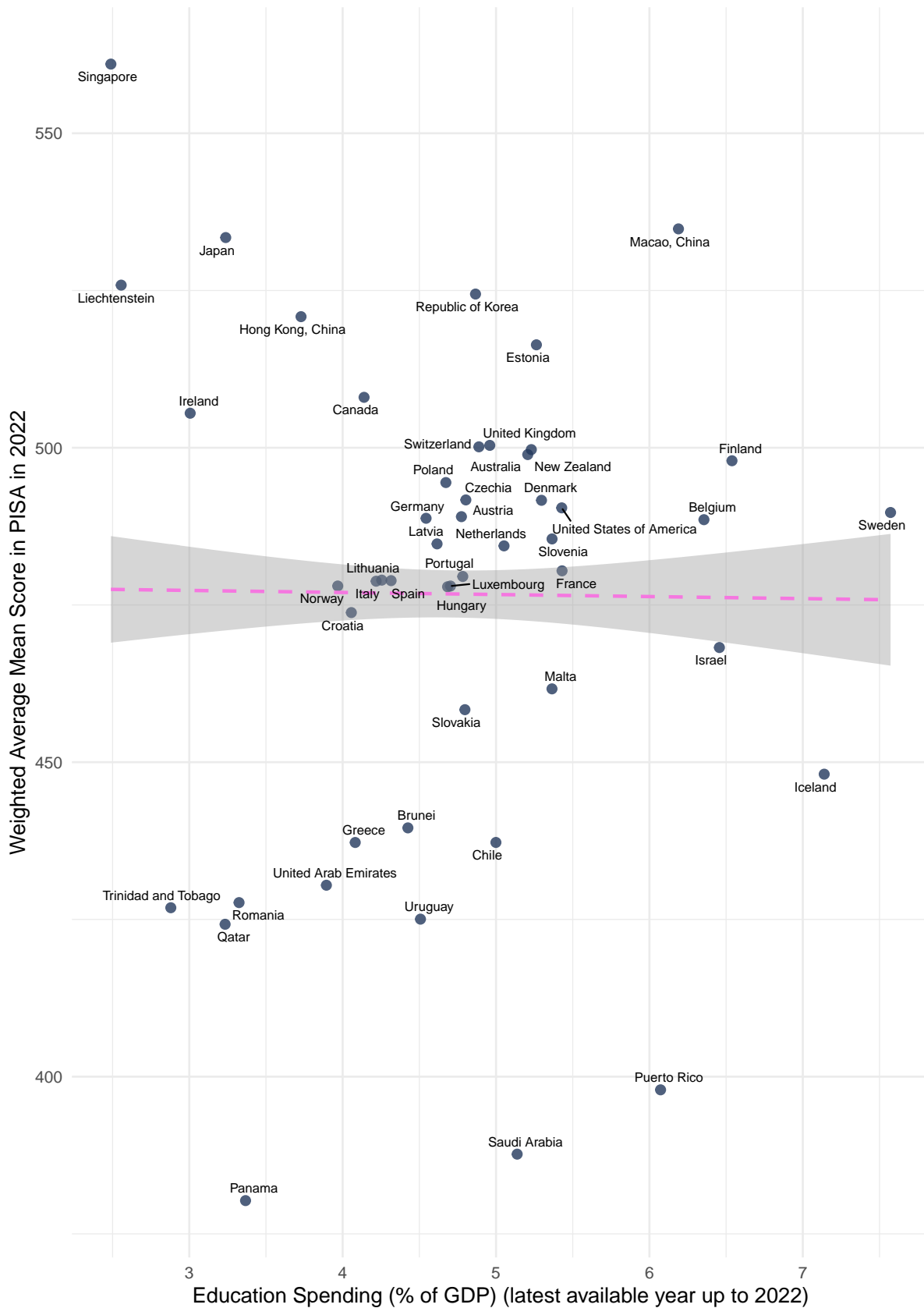
2.2.9 Educational investment and educational outcomes for High Income

This analysis employs a scatter plot design to examine the relationship between **educational investment** (measured as education spending as a percentage of GDP from UNESCO) and **educational outcomes** (weighted average mean PISA scores) across High income countries in 2022. The methodology includes several important considerations:

Measurement Approach: The use of education spending as a percentage of GDP provides a standardized metric that accounts for economic scale differences between countries, making cross-national comparisons more meaningful than absolute spending figures. The weighted average mean PISA score offers a comprehensive measure of educational achievement across multiple domains.

Sample Selection: Focusing exclusively on High income countries creates a controlled comparison group, reducing the confounding effects of vastly different economic development levels that might obscure the spending-achievement relationship.

Education Spending (% of GDP) and Weighted Average Mean Score in PISA (High Income Countries, latest year up to 2022)



This analysis examines the relationship between education expenditure (as % of GDP) and PISA performance across **45 high-income countries**, revealing a **complex and counterintuitive pattern** that challenges conventional assumptions about education investment and outcomes.

Weak Correlation Between Spending and Performance:

The scatter plot reveals a notably weak positive correlation between education spending and PISA scores, suggesting that simply increasing education budgets does not guarantee improved student outcomes. The wide confidence interval (gray shaded area) indicates substantial variation around any potential trend line.

High-Performing Outliers with Low Spending:

Several countries achieve exceptional PISA results with relatively modest education spending:

Singapore (561 PISA score, 2.5% GDP): Highest performer globally with lowest spending rate **Japan** (533 PISA score, 3.2% GDP): Second-highest performer with below-average spending **Hong Kong** (521 PISA score, 3.7% GDP): Third-highest performer with low-moderate spending **Macao** (535 PISA score, 6.2% GDP): High performer with moderate-high spending

These cases suggest that educational efficiency and system design may matter more than raw investment levels.

High-Spending Underperformers:

Conversely, several countries invest heavily in education but achieve mediocre PISA results:

Iceland (448 PISA score, 7.1% GDP): Highest spender with below-average performance **Belgium** (489 PISA score, 6.4% GDP): High spender with moderate performance **Sweden** (490 PISA score, 7.6% GDP): High spender with good but not exceptional performance

This pattern indicates potential inefficiencies in resource allocation or systemic challenges that money alone cannot resolve.

The “Efficiency Frontier” Concept:

Countries like **Singapore**, **Japan**, and **South Korea** appear to operate on an “efficiency frontier,” maximizing educational outcomes per unit of investment. This suggests these systems have optimized factors such as:

- Teacher quality and training
- Curriculum design and implementation
- Educational technology integration
- Student motivation and cultural attitudes toward learning

Regional and Cultural Patterns:

East Asian Excellence:

The top-performing countries (**Singapore**, **Macao**, **Japan**, **South Korea**, **Hong Kong**) are predominantly East Asian, suggesting cultural, pedagogical, or systemic factors beyond spending levels.

European Variation:

European countries show wide variation in both spending (2.6% to 7.6% GDP) and performance (428 to 525 PISA points), indicating diverse approaches and effectiveness across the continent.

Policy Implications:

Beyond Budgetary Solutions The weak spending-performance correlation suggests that education policy should focus on:

- System efficiency rather than just increased funding
- Teacher quality improvements through better recruitment, training, and retention
- Curriculum optimization and pedagogical innovation

- Student engagement and motivation strategies

Learning from High-Efficiency Systems:

Policymakers should study the practices of high-performing, low-spending countries to identify transferable strategies:

- Singapore's teacher development programs
- Japan's classroom management techniques
- South Korea's integration of technology and traditional pedagogy

Diminishing Returns Consideration:

Countries already spending above 5-6% of GDP on education may face diminishing returns and should focus on optimization rather than expansion of budgets.

2.2.10 Economic Prosperity vs Educational Achievement

This analysis examines the relationship between economic prosperity (GDP per capita PPP) and educational achievement (PISA 2022 scores) across 77 participating countries and territories.

The data reveals a moderate positive correlation with notable exceptions that challenge assumptions about the relationship between national wealth and educational performance.

This scatter plot illustrates the relationship between GDP per Capita (PPP, 2022, log scale) on the x-axis and an unlabeled vertical axis. The data points represent various countries, showing a strong positive correlation. A shaded gray area indicates the confidence interval, and a dashed pink line represents the trend. The countries are labeled with their names, and the x-axis is marked with values 10,000, 30,000, and 100,000.

Country	GDP per Capita (PPP, 2022, log scale)
Singapore	~120,000
Japan	~40,000
Republic of Korea	~35,000
Macao, China	~45,000
Hong Kong, China	~40,000
Estonia	~35,000
United Kingdom	~40,000
Canada	~45,000
New Zealand	~35,000
Finland	~40,000
Australia	~45,000
Czechia	~35,000
Poland	~35,000
Slovenia	~35,000
Germany	~40,000
Sweden	~45,000
Switzerland	~50,000
Denmark	~45,000
Austria	~45,000
United States of America	~50,000
France	~45,000
Netherlands	~45,000
Norway	~50,000
Italy	~40,000
Spain	~35,000
Portugal	~35,000
Hungary	~35,000
Croatia	~35,000
Israel	~40,000
Slovakia	~35,000
Malta	~40,000
Iceland	~45,000
Brunei	~50,000
United Arab Emirates	~45,000
Qatar	~100,000
Romania	~35,000
Greece	~35,000
Chile	~30,000
Uruguay	~25,000
Bulgaria	~30,000
Kazakhstan	~35,000
Malaysia	~35,000
Argentina	~30,000
Panama	~35,000
Georgia	~25,000
Azerbaijan	~25,000
North Macedonia	~25,000
Albania	~25,000
Indonesia	~20,000
Guatemala	~15,000
El Salvador	~15,000
Paraguay	~15,000
Dominican Republic	~20,000
Philippines	~10,000
Uzbekistan	~10,000
El Salvador	~10,000
Jordan	~10,000
Morocco	~10,000
Palestine	~5,000
Cambodia	~5,000
Viet Nam	~15,000

The logarithmic scatter plot reveals a moderate positive correlation between GDP per capita (PPP) and PISA performance, with the pink trend line showing that economic prosperity generally associates with better educa-

tional outcomes. However, the wide confidence interval indicates substantial variation, with numerous countries significantly over- or under-performing relative to their economic status.

Exceptional High-Achievers Across Economic Spectrum

Several countries demonstrate outstanding educational performance regardless of their economic positioning:

High-Income High-Achievers:

Singapore (561 PISA, \$132,469 GDP): Global leader in both metrics **Ireland** (505 PISA, \$125,625 GDP): Strong performance with highest GDP per capita **Switzerland** (500 PISA, \$82,203 GDP): Consistent high performance

Middle-Income Educational Champions:

Estonia (516 PISA, \$43,690 GDP): Exceptional performance relative to economic status **Japan** (533 PISA, \$44,937 GDP): Second-highest PISA score with moderate GDP **South Korea** (524 PISA, \$49,778 GDP): Third-highest performance **Macao** (535 PISA, \$60,090 GDP): Highest performer after Singapore

Economic Prosperity Without Educational Excellence

Several wealthy nations underperform relative to their economic capacity:

Qatar (424 PISA, \$114,740 GDP): High wealth, moderate educational performance **Brunei** (440 PISA, \$76,357 GDP): Significant underperformance given economic resources **UAE** (430 PISA, \$68,868 GDP): Below-average education despite high GDP **Norway** (478 PISA, \$90,839 GDP): Surprisingly modest performance given wealth

Remarkable Low-Income Achievers

Some countries with limited economic resources achieve impressive educational outcomes:

Vietnam (468 PISA, \$12,930 GDP): Outstanding performance given economic constraints **Albania** (368 PISA, \$17,113 GDP): Reasonable performance for income level **Ukraine** (438 PISA, \$13,787 GDP): Strong performance despite economic challenges

Regional and Development Patterns

East Asian Excellence Confirmed: The top educational performers (**Singapore, Macao, Japan, South Korea, Hong Kong**) reinforce East Asian educational superiority across different income levels.

Resource-Rich Underperformers: Several oil-rich nations (**Qatar, UAE, Saudi Arabia, Brunei**) demonstrate that natural resource wealth doesn't automatically translate to educational excellence.

European Consistency: Most European nations cluster around the trend line, showing relatively predictable relationships between wealth and educational performance.

Latin American Challenges: Most Latin American countries fall below the trend line, indicating systematic educational challenges despite varying economic conditions.

Policy Implications

1. Economic Development Alone Insufficient for Educational Excellence

The moderate correlation demonstrates that while economic prosperity provides important resources for education, it does not guarantee superior outcomes. Countries must strategically invest in educational systems, teacher quality, and pedagogical approaches.

2. Learning from Cross-Income Success Models

High-Efficiency Developing Nations: Countries like Vietnam demonstrate that effective educational systems can be built without extensive resources through:

- Focused investment in teacher training and development
- Emphasis on fundamental skills and rigorous curricula

- Strong cultural emphasis on educational achievement
- Efficient resource allocation

Resource-Rich Optimization: Wealthy underperformers (**Qatar, UAE, Brunei**) should examine why their economic advantages haven't translated to educational success, potentially focusing on:

- Educational system design and governance
- Teacher recruitment and retention strategies
- Student motivation and engagement programs
- Long-term educational planning beyond infrastructure investment

3. **The “Development Trap”:** Consideration Some middle-income countries may face a “development trap” where economic growth doesn't automatically improve educational outcomes, requiring targeted educational reforms alongside economic development.
4. **Regional Learning Networks:** The clustering of high-performing East Asian nations suggests value in regional educational cooperation and knowledge sharing, while underperforming regions could benefit from systematic study of successful educational models.

Limitations and Considerations:

Causality Questions: The correlation doesn't establish whether economic prosperity drives educational achievement or vice versa.

Cultural Factors: PISA performance may reflect cultural attitudes toward education and testing that aren't directly related to economic development.

Historical Context: Some countries may be experiencing educational improvements or declines not captured in cross-sectional data.

System Maturity: Educational systems require time to develop, and recent economic growth may not yet have impacted educational outcomes.

Natural Resource Dependence: Resource-rich economies may face specific challenges in developing human capital

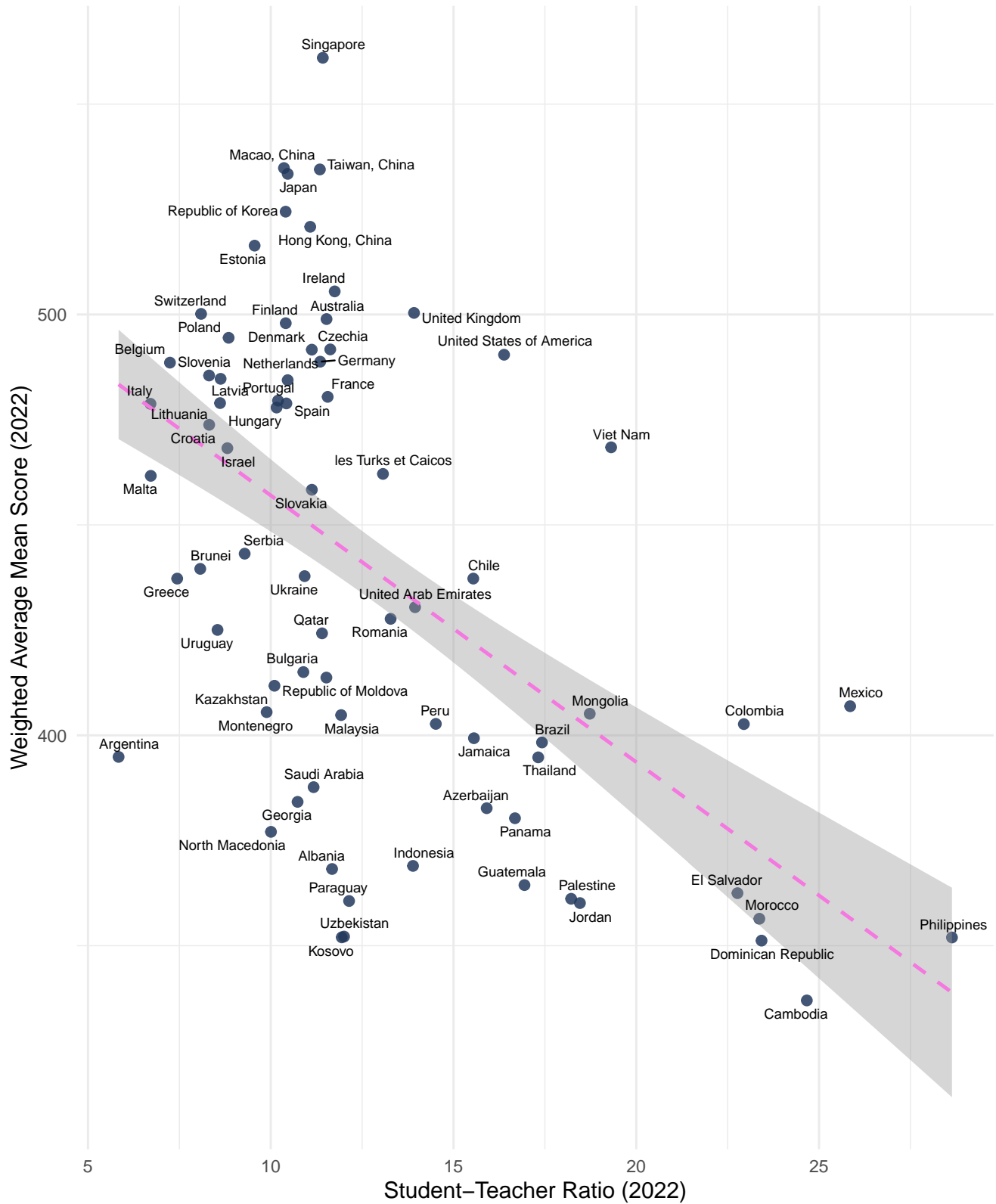
This analysis reveals that while economic prosperity and educational achievement are positively correlated, the relationship is far from deterministic. The most successful educational systems appear to optimize factors beyond pure economic resources, including cultural attitudes, pedagogical approaches, and system design. Countries at all income levels can achieve educational excellence through strategic focus on educational quality rather than relying solely on economic development to drive educational improvement.

2.2.11 Student-Teacher Ratio vs Educational Achievement

This analysis examines the relationship between student-teacher ratios and educational achievement (PISA 2022 scores) across 74 participating countries and territories.

The data reveals a strong negative correlation, demonstrating that smaller class sizes generally associate with better educational outcomes, though notable exceptions highlight the complexity of educational effectiveness.

Weighted Average Mean Score (2022) vs Student–Teacher Ratio (2022)



1. Strong Negative Correlation with Clear Trend

The scatter plot shows a pronounced negative correlation between student-teacher ratios and PISA performance, with the pink trend line clearly indicating that countries with lower student-teacher ratios (smaller class sizes) tend

to achieve higher PISA scores. The confidence interval suggests this relationship is statistically robust across most of the range.

2. Low Student-Teacher Ratio Champions

Countries achieving exceptional educational performance with favorable student-teacher ratios:

Ultra-Low Ratios (5-8 students per teacher):

Italy (479 PISA, 6.7 ratio): Strong performance with excellent teacher availability **Malta** (462 PISA, 6.7 ratio): Good performance with very low ratios **Belgium** (489 PISA, 7.2 ratio): High performance with favorable staffing **Greece** (437 PISA, 7.4 ratio): Moderate performance despite excellent ratios

Optimal Range (8-10 students per teacher):

Switzerland (500 PISA, 8.1 ratio): Excellent performance with optimal staffing **Slovenia** (485 PISA, 8.3 ratio): Strong performance **Poland** (494 PISA, 8.8 ratio): High achievement with good ratios **Estonia** (516 PISA, 9.6 ratio): Outstanding performance

3. High-Performing Systems with Moderate Ratios

Several top-performing countries achieve excellence despite moderate student-teacher ratios:

Singapore (561 PISA, 11.4 ratio): Global leader with moderate staffing levels **Macao** (535 PISA, 10.4 ratio): Second-highest performer **Taiwan** (534 PISA, 11.3 ratio): Exceptional achievement **Japan** (533 PISA, 10.5 ratio): Top-tier performance **South Korea** (524 PISA, 10.4 ratio): Elite educational system **Hong Kong** (521 PISA, 11.1 ratio): Outstanding results

4. High Ratio Underperformers

Countries with high student-teacher ratios facing educational challenges: Extreme Ratios (25+ students per teacher):

Philippines (352 PISA, 28.6 ratio): Severe staffing constraints **Mexico** (407 PISA, 25.9 ratio): Large classes limiting effectiveness **Cambodia** (337 PISA, 24.7 ratio): Resource constraints evident

High Ratios (20-25 students per teacher):

Morocco (356 PISA, 23.4 ratio): Limited teacher availability **Colombia** (403 PISA, 22.9 ratio): Staffing challenges **Dominican Republic** (351 PISA, 23.4 ratio): Resource limitations

5. Notable Exceptions and Outliers

Vietnam Anomaly (468 PISA, 19.3 ratio): Remarkable achievement despite high student-teacher ratios, suggesting exceptional teacher quality or educational system efficiency. **USA Performance** (490 PISA, 16.4 ratio): Good performance despite relatively high ratios for a developed nation. **Resource-Rich Underperformers:** Countries like UAE (430 PISA, 13.9 ratio) and Qatar (424 PISA, 11.4 ratio) show that moderate ratios don't guarantee success without system optimization.

Regional and Developmental Patterns

European Excellence in Staffing

Most European countries cluster in the favorable 7-12 student-teacher ratio range, with generally strong PISA performance, suggesting systematic investment in teacher availability.

East Asian Efficiency

Top East Asian performers (Singapore, Macao, Taiwan, Japan, South Korea, Hong Kong) achieve exceptional results with moderate ratios (10-11.5), indicating highly efficient educational systems that maximize teacher effectiveness.

Latin American Challenges

Most Latin American countries show high student-teacher ratios (15-26) with correspondingly lower PISA scores, indicating systematic under-investment in teaching staff.

Developing Nation Constraints: Many lower-income countries face severe staffing constraints (ratios above 20), directly correlating with educational performance challenges.

Policy Implications

1. Teacher Availability as Foundation

The strong correlation suggests that ensuring adequate teacher availability is fundamental to educational success. Countries should prioritize:

- Teacher recruitment and retention programs
- Competitive compensation packages to attract quality educators
- Investment in teacher training infrastructure
- Long-term workforce planning for educational staffing

2. The “Efficiency Threshold” Concept

East Asian examples suggest there may be an “efficiency threshold” around 10-12 students per teacher where exceptional systems can achieve outstanding results through:

- Superior teacher quality and training
- Optimized pedagogical approaches
- Strong educational culture and student motivation
- Effective classroom management techniques

3. Beyond Simple Ratios

Vietnam’s exceptional performance despite high ratios demonstrates that student-teacher ratios, while important, are not deterministic. Other critical factors include:

- Teacher quality and preparation
- Curriculum design and implementation
- Educational leadership and system management
- Cultural attitudes toward education

4. Targeted Investment Strategies

For High-Ratio Countries: Priority should be systematic teacher recruitment and training programs to reduce class sizes while maintaining quality. **For Low-Ratio Countries:** Focus should shift to teacher quality improvement and pedagogical optimization rather than further ratio reduction.

Limitations and Considerations

Causality Direction: While correlation is strong, the direction of causality may be complex; better educational systems may both attract more teachers and produce better outcomes.

Quality vs. Quantity: Student-teacher ratios don’t capture teacher quality, training, or effectiveness.

System Context: Different educational approaches may optimize different ratio ranges.

Resource Allocation: Some countries may achieve better results by investing in teacher quality rather than quantity.

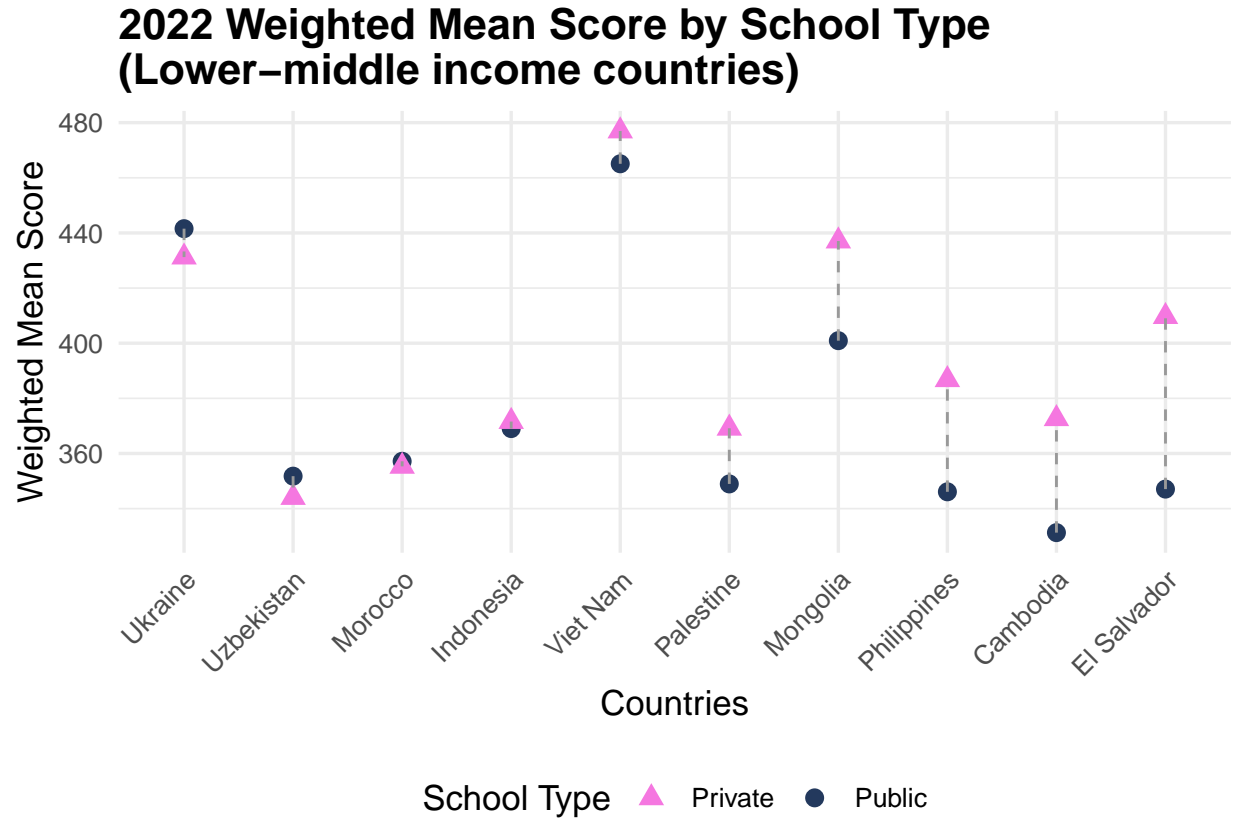
Cultural Factors: Educational effectiveness may depend on cultural attitudes that influence how class size impacts learning.

This analysis demonstrates a strong relationship between student-teacher ratios and educational achievement, with smaller class sizes generally correlating with better PISA performance. However, the most successful educational systems appear to optimize both teacher availability and teacher effectiveness.

While adequate staffing is clearly important, countries like Vietnam and the East Asian high-performers show that exceptional educational outcomes are possible through systematic focus on teacher quality, pedagogical excellence, and educational system optimization, even within resource constraints. The key lesson is that teacher availability provides the foundation, but educational excellence requires comprehensive system design and implementation.

2.2.12 Public vs Private school performance for Lower Middle Income Countries

This analysis examines the performance gap between private and public schools in lower-middle income countries participating in PISA 2022. The data reveals significant variation in the public-private achievement gap, with some countries showing substantial private school advantages while others demonstrate superior public school performance, challenging common assumptions about private education superiority in developing contexts.



1. Varied Private School Advantage Patterns

The data shows three distinct patterns in public-private school performance across lower-middle income countries:

Strong Private School Advantage (40+ point gaps):

El Salvador (+63 points): Largest private advantage with private schools scoring 410 vs public 347 **Cambodia** (+41 points): Significant private school superiority (373 vs 331) **Philippines** (+41 points): Notable private advantage (387 vs 346) **Mongolia** (+36 points): Substantial gap favoring private schools (437 vs 401)

Moderate Private Advantage (10-25 point gaps):

Occupied Palestinian Territory (+20 points): Moderate private school advantage (369 vs 349) **Vietnam** (+12 points): Small but consistent private advantage (477 vs 465)

Public School Superiority (negative gaps):

Ukraine (-10 points): Public schools outperform private (442 vs 431) **Uzbekistan** (-8 points): Public advantage (352 vs 344) **Morocco** (-2 points): Minimal public advantage (357 vs 355)

Minimal Difference:

Indonesia (+3 points): Virtually no performance gap (372 vs 369)

2. Vietnam as Educational Excellence Model

Vietnam stands out as the highest-performing lower-middle income country with strong results in both sectors:

Overall Excellence: Significantly outperforming all other countries in both public (465) and private (477) schools
Sector Balance: Small private advantage (12 points) suggests both sectors function effectively. **System Effectiveness:** Public schools in Vietnam (465) outperform private schools in most other countries.

3. El Salvador's Dual System Challenge

El Salvador presents the most dramatic public-private divide:

Extreme Gap: 63-point difference represents nearly two years of learning. **Private Performance:** Private schools achieve respectable 410 score. **Public Struggles:** Public schools severely underperform at 347, indicating systemic challenges.

4. High-Performing Public Systems

Several countries demonstrate effective public education systems:

Ukraine: Public schools (442) outperform private schools, suggesting strong state education investment. **Uzbekistan:** Public system provides better outcomes than private alternatives. **Morocco:** Minimal gap indicates relatively balanced system performance.

5. System Equity Implications

- High Private Advantage Countries likely face:
- Significant educational inequality based on family income
- Brain drain from public to private sector teachers
- Under-investment in public education infrastructure
- Social stratification through education access

Balanced Systems suggest:

- More equitable educational opportunities
- Effective public education governance
- Better resource distribution across sectors

Regional and Developmental Context

Southeast Asian Patterns

Vietnam: Demonstrates that lower-middle income countries can achieve educational excellence **Philippines** and **Cambodia:** Show typical developing country patterns with significant private advantages **Indonesia:** Unique balance suggests effective public-private integration

Post-Soviet Space

Ukraine and **Uzbekistan:** Both show public sector strength, possibly reflecting Soviet educational legacy Strong state education systems despite economic constraints

Latin American Context

El Salvador: Represents common Latin American pattern of significant public-private gaps Suggests need for targeted public education investment

Policy Implications

1. Context-Specific Strategies Required

For High Private Advantage Countries (El Salvador, Cambodia, Philippines, Mongolia):

Public Sector Investment: Urgent need to improve public school resources, teacher training, and infrastructure

Teacher Quality Programs: Focus on recruiting and retaining quality teachers in public schools

Equity Initiatives: Policies to reduce educational inequality between sectors

For Balanced Systems (Vietnam, Indonesia, Morocco):

System Optimization: Continue strengthening both sectors while maintaining balance **Best Practice Sharing:** These countries can serve as models for other developing nations

For Public-Advantage Countries (Ukraine, Uzbekistan):

System Preservation: Maintain public sector strengths while supporting private sector development

Resource Efficiency: Optimize public education investments to sustain performance

2. Learning from Vietnam's Success

Vietnam's exceptional performance across both sectors suggests:

Teacher Development: Systematic investment in teacher quality and training

Curriculum Standards: Rigorous academic standards applied consistently

Cultural Factors: Strong emphasis on educational achievement

System Governance: Effective oversight of both public and private sectors

3. Addressing Educational Inequality

Countries with large private advantages need:

Targeted Public Investment: Focus resources on improving public school quality.

Teacher Mobility Programs: Initiatives to share expertise between sectors.

Regulatory Frameworks: Ensure private schools contribute to overall system improvement.

Scholarship Programs: Provide access to quality education regardless of family income.

Limitations and Considerations

Sample Size Effects: Some countries may have small private school samples affecting reliability.

Selection Bias: Private school students may come from more advantaged backgrounds.

Resource Variations: Private schools may have access to better resources independent of teaching quality.

Cultural Context: Different countries may have varying cultural attitudes toward public vs private education.

System Maturity: Some private education sectors may be relatively new and still developing.

This analysis reveals that the relationship between school sector and educational achievement in lower-middle income countries is far from uniform. While several countries show significant private school advantages, others demonstrate that effective public education systems can achieve superior or equivalent results.

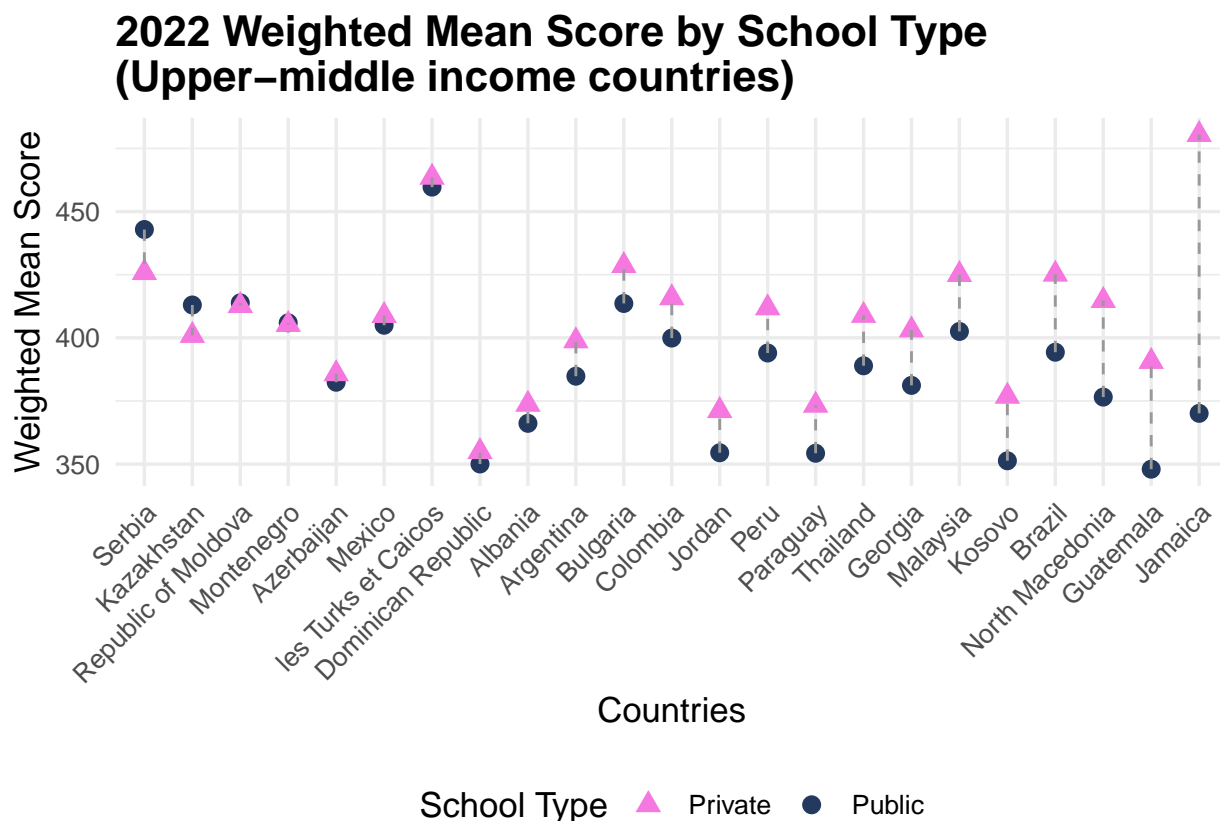
Vietnam's exceptional performance in both sectors proves that lower-middle income countries can achieve educational excellence through systematic investment and effective governance.

The key insight is that sector type (public vs private) is less important than system quality, teacher effectiveness, and educational governance. Countries should focus on improving educational quality across all sectors rather than assuming private education is inherently superior.

2.2.13 Public vs Private school performance for Upper Middle Income Countries

This analysis examines the performance gap between private and public schools across 23 upper-middle income countries in PISA 2022.

The data reveals significant variation in the private-public school performance differential, ranging from a 17-point advantage for public schools to a remarkable 110-point advantage for private schools.



Performance Distribution Patterns

The upper-middle income countries demonstrate a more pronounced private school advantage compared to other income groups. Of the 23 countries analyzed, 19 countries (83%) show private schools outperforming public schools, while only 4 countries exhibit the opposite pattern.

Countries Where Public Schools Outperform Private Schools

Four countries buck the global trend of private school superiority:

- **Serbia** shows the largest public school advantage (-17.1 points), suggesting effective public education policies.
- **Kazakhstan** demonstrates a moderate public school lead (-12.1 points).
- **Republic of Moldova** and **Montenegro** show minimal differences (-1.1 and -0.7 points respectively), indicating near-parity between sectors.

Countries with Moderate Private School Advantages

Several countries exhibit small to moderate private school advantages (0-20 points):

- **Azerbaijan, Mexico, and Turkey** show minimal gaps (3.4-3.9 points).
- **Dominican Republic through Thailand** demonstrate moderate advantages (4.8-19.9 points).

This group suggests relatively balanced educational systems with manageable inequality between sectors.

Countries with Significant Private School Advantages

Eight countries display substantial private school advantages (20-45 points):

- **Georgia, Malaysia, Kosovo, and Brazil** show gaps of 20-31 points.
- **North Macedonia and Guatemala** exhibit larger disparities (38-43 points).

These gaps indicate significant educational inequality within national systems.

Exceptional Case: Jamaica

Jamaica stands as a complete outlier with a 110.4-point private school advantage, representing the most extreme educational inequality in the dataset. This gap suggests fundamental systemic differences between private and public education quality in the country.

Regional and Economic Context

Economic Development Correlation

The upper-middle income group shows interesting patterns related to economic transition:

- Countries with recent economic growth (Kazakhstan, Azerbaijan) tend to have smaller gaps.
- Traditional middle-income countries (Argentina, Colombia, Brazil) show moderate to large private advantages.

This suggests that educational inequality may persist even as countries develop economically.

Geographic Patterns

Eastern Europe and Central Asia: Mixed results, with Serbia favoring public schools while Bulgaria and North Macedonia show private advantages.

Latin America: Consistently shows private school advantages, from moderate (Mexico) to extreme (Jamaica).

Southeast Asia: Thailand and Malaysia demonstrate significant private advantages, reflecting regional educational stratification.

Educational Policy Implications

Systems Requiring Attention

Countries with large private-public gaps (over 25 points) may need comprehensive education reform focusing on:

- Public school resource allocation
- Teacher training and retention
- Infrastructure development
- Quality assurance mechanisms

Successful Public Systems

Serbia and **Kazakhstan** provide models for effective public education in middle-income contexts, suggesting that well-managed public systems can compete with or exceed private sector performance.

Equity Concerns

The prevalence of private school advantages in this income group highlights the risk of educational stratification during economic development phases, potentially undermining social mobility and equality of opportunity.

Upper-middle income countries face a critical juncture in educational development. While economic growth provides resources for educational improvement, the data suggests that benefits often accrue disproportionately to private institutions.

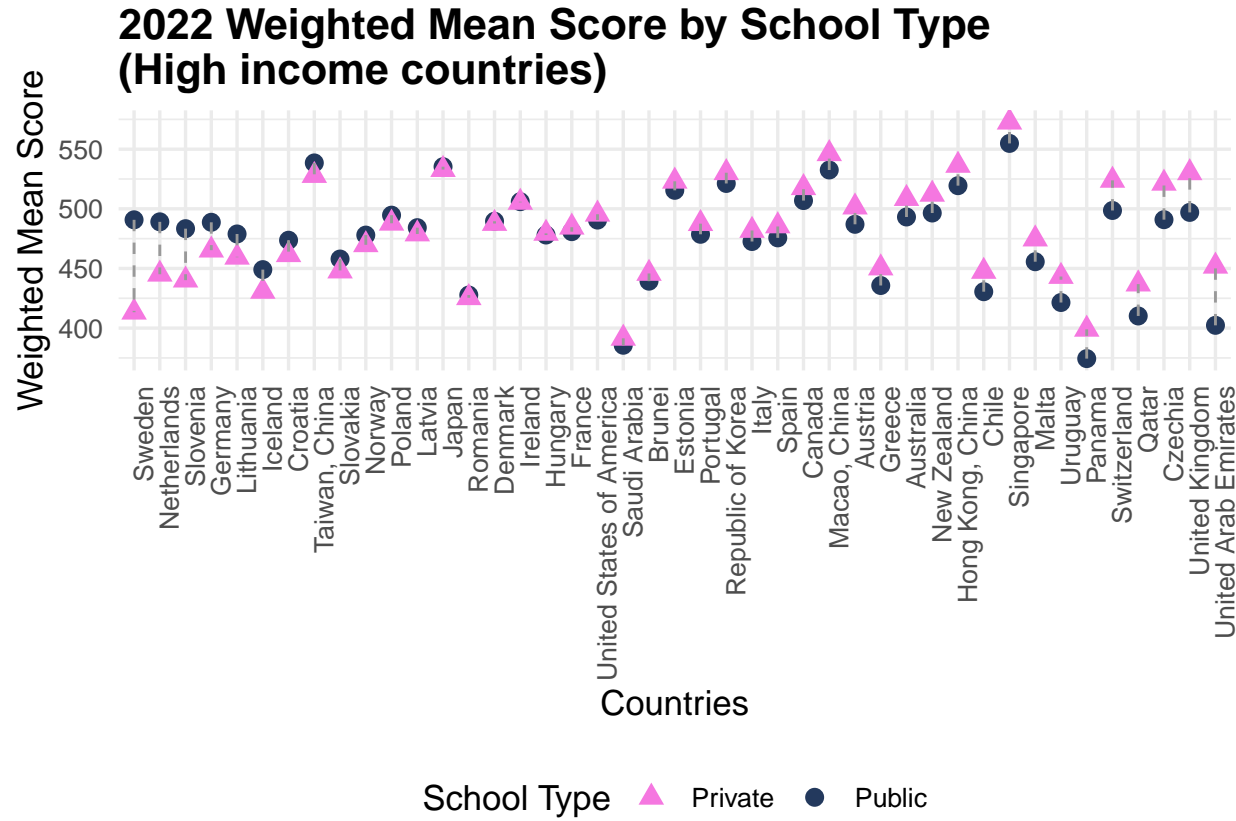
Countries like Serbia demonstrate that strong public systems are achievable, while extreme cases like Jamaica highlight the urgent need for comprehensive educational reform to prevent educational apartheid.

Policymakers in this income group must prioritize public education investment to ensure that economic development translates into equitable educational opportunities for all students.

2.2.14 Public vs Private school performance for High Income Countries

This analysis examines the performance gap between private and public schools across 43 high-income countries in PISA 2022.

Unlike lower-income groups, high-income countries show a more balanced distribution of outcomes, with 16 countries (37%) favoring public schools and 27 countries (63%) showing private school advantages. The performance gaps are generally smaller than in other income groups, reflecting more mature and equitable educational systems.



Performance Distribution Patterns

High-income countries demonstrate the most balanced private-public performance relationship among all income groups. The data reveals three distinct clusters: countries with strong public school advantages, those with near-parity, and nations with moderate to significant private school benefits.

Countries Where Public Schools Significantly Outperform Private Schools

Strong Public School Systems (gaps > 20 points):

- **Sweden** leads with the largest public advantage (-77.5 points), demonstrating exceptional public education effectiveness.
- **Netherlands** (-43.9 points) and **Slovenia** (-43.1 points) show substantial public school superiority.
- **Germany** (-23.4 points) rounds out this high-performing public sector group.

Moderate Public School Advantages (10-20 points):

- **Lithuania, Iceland, Croatia, Taiwan, and Slovakia** show consistent public school benefits.

These countries demonstrate well-funded, effectively managed public systems.

Minimal Public School Advantages (0-10 points):

- **Norway, Poland, Latvia, Japan, Romania, Denmark, and Ireland** show small public advantages or near-parity.

This group represents highly equitable educational systems with minimal sector-based disparities.

Countries with Private School Advantages

Minimal Private Advantages (0-10 points):

- **Hungary through Spain** show small private school benefits (1.4-9.9 points).

This includes major economies like **France, USA, and South Korea**, indicating relatively equitable systems.

Moderate Private Advantages (10-20 points):

- **Canada, Macao, Austria, Greece, Australia, New Zealand, Hong Kong, Chile, and Singapore** demonstrate moderate gaps.

Even Singapore, despite its 17.8-point private advantage, maintains high overall performance in both sectors.

Significant Private Advantages (20+ points):

- **Switzerland, Qatar, Czechia, and the UK** show substantial private school benefits (25-33 points)
- **UAE** stands out with the largest private advantage (49.5 points) among high-income countries

Regional and Development Patterns

Nordic Model Excellence

Scandinavian countries (Sweden, Iceland, Norway, Denmark) predominantly favor public schools, reflecting their comprehensive welfare state approach to education. Sweden's exceptional 77-point public advantage exemplifies this model's effectiveness.

Continental European Balance

Western and Central Europe shows mixed patterns:

- **German-speaking countries** vary significantly: Germany favors public schools while Switzerland and Austria favor private schools.
- **Southern Europe** (Italy, Spain, Greece) shows moderate private advantages.
- **Eastern European EU members** generally maintain competitive public systems.

Anglo-Saxon Private Tendencies

English-speaking countries consistently show private school advantages:

- **Australia, New Zealand, UK, and Canada** demonstrate moderate private benefits.
- **USA** shows a surprisingly small gap (4.9 points), indicating relatively balanced systems.

This pattern reflects historical private school traditions in these societies.

Asian High Performers

East Asian countries show interesting variation:

- **Japan** achieves near-parity with slight public advantage.
- **South Korea** shows small private advantage.
- **Singapore** and **Hong Kong** maintain moderate private advantages while achieving exceptional overall performance.
- **Taiwan** favors public schools, bucking regional trends.

Gulf States Pattern

Oil-rich nations (UAE, Qatar, Saudi Arabia) show consistent private school advantages, likely reflecting rapid educational development and reliance on international private institutions.

Educational System Maturity Indicators

Equity Achievement

High-income countries demonstrate significantly smaller performance gaps compared to other income groups, suggesting that economic development enables more equitable educational systems regardless of public-private distinctions.

Quality Convergence

Many high-income countries achieve high performance in both sectors, indicating that competition and regulation have elevated standards across the board.

Policy Implications

For Countries with Large Public Advantages:

- Sweden and Netherlands provide models for exceptional public education.
- These systems suggest that well-funded, professionally managed public education can outperform private alternatives

For Countries with Large Private Advantages:

- UAE, UK, and Czechia may need to examine public sector investment and management.
- These gaps suggest potential equity concerns even in wealthy nations.

For Balanced Systems:

- Countries achieving near-parity (Ireland, Denmark, Japan) demonstrate optimal educational equity.
- These systems provide models for minimizing socioeconomic educational stratification.

Global Competitiveness Context

High-income countries occupy most top positions in global PISA rankings, regardless of public-private performance gaps. Countries like **Singapore, Japan, and Taiwan** achieve world-leading results while maintaining different sector balance approaches, suggesting multiple pathways to educational excellence.

High-income countries demonstrate that economic development enables educational equity across sectors. The prevalence of strong public systems (particularly in Nordic countries) challenges assumptions about private school superiority, while countries with private advantages often maintain high public performance standards.

The key insight is that in mature, well-funded educational systems, sector type matters less than overall system quality, professional management, and equitable resource allocation. Countries like Sweden prove that exceptional public education is achievable, while nations like Singapore show that private advantages need not compromise overall system excellence.

Policymakers in high-income countries should focus on system-wide quality improvement rather than sector competition, using successful models like those in Scandinavia and East Asia to inform comprehensive educational strategies.

2.2.15 The gender performance gap for the Lower Middle Income Countries

This analysis examines the gender performance gap across 10 lower-middle income countries participating in PISA 2022.

The data reveals a consistent pattern where female students outperform male students in all countries analyzed, with gaps ranging from minimal (0.5 points) to substantial (31 points).

This finding challenges traditional assumptions about gender performance in mathematics and highlights important educational equity considerations.



Performance Gap Categories

Minimal Gender Gaps (0-5 points):

- **Vietnam** (0.5 points): Achieves near-perfect gender parity while maintaining the highest overall performance (468 points for both genders).

El Salvador (1.3 points): Shows minimal gender disparity with very small female advantage.

- **Uzbekistan** (5.4 points): Maintains relatively balanced performance despite overall lower scores.

Moderate Gender Gaps (6-15 points):

- **Ukraine** (6.6 points): Demonstrates moderate female advantage with high overall performance (441 female, 435 male).
- **Cambodia** (9.9 points): Shows growing gender disparity with lower overall achievement.
- **Morocco** (10.9 points): Exhibits double-digit female advantage.
- **Indonesia** (12.8 points): Demonstrates significant female superiority.
- **Mongolia** (15.3 points): Shows substantial gender gap while maintaining moderate overall performance.

Large Gender Gaps (15+ points):

- **Philippines** (20.0 points): Exhibits substantial female advantage with concerning overall low performance.
- **Occupied Palestinian Territory** (31.0 points): Shows the largest gender gap in the group, with females significantly outperforming males. This could be explained by the current geopolitical instability requiring.

Performance Level Analysis

High-Performing Countries with Gender Equity

Vietnam stands out as the exemplary model, achieving both the highest overall performance and near-perfect gender parity. This demonstrates that educational excellence and gender equity are achievable simultaneously in lower-middle income contexts.

Ukraine combines strong overall performance with manageable gender gaps, suggesting effective educational policies prior to recent disruptions.

Moderate Performers with Growing Gaps

Mongolia and Indonesia show concerning patterns where moderate overall performance coexists with significant gender disparities, suggesting systemic issues affecting male student achievement.

Lower Performers with Substantial Gaps

Philippines and Occupied Palestinian Territory exhibit the most troubling combination of low overall performance and large gender gaps, indicating multiple educational challenges requiring comprehensive intervention.

Regional and Cultural Patterns

Southeast Asian Excellence

Vietnam demonstrates that Southeast Asian educational approaches can achieve both high performance and gender equity, serving as a regional model.

Indonesia and Philippines show contrasting outcomes within the same region, suggesting that similar cultural contexts can produce different gender equity results based on educational policies and implementation.

Post-Soviet Transition Success

Ukraine and Uzbekistan represent post-Soviet educational systems with relatively strong female performance, possibly reflecting historical emphasis on gender equality in education.

Middle Eastern Complexities

Morocco and Occupied Palestinian Territory show varying degrees of female advantage, challenging stereotypes about gender in Middle Eastern education while highlighting the need for targeted male student support.

Latin American Variation

El Salvador achieves near-parity, suggesting effective gender-inclusive educational approaches in Central America.

Educational System Implications

Factors Contributing to Female Advantage

Educational Engagement: The consistent female advantage may reflect higher educational motivation and engagement among girls in these contexts, possibly driven by perceptions of education as a pathway to social mobility.

Cultural Shifts: Changing societal attitudes toward female education may be creating environments where girls outperform traditional expectations.

Teaching Methods: Educational approaches in these countries may be more conducive to female learning styles, or girls may be adapting better to current assessment formats.

Male Student Achievement Concerns

Systemic Underperformance: The universal male underachievement raises concerns about educational approaches that may not be engaging male students effectively.

Social and Economic Factors: Economic pressures may disproportionately affect male students through early workforce entry or different educational expectations.

Educational Policies: Current educational strategies may inadvertently favor female students or fail to address specific challenges facing male learners.

Policy Recommendations

For Countries with Large Gaps (Philippines, Palestinian Territory)

Comprehensive Gender Analysis: Conduct detailed studies to understand root causes of large gender disparities.

Male Student Support Programs: Develop targeted interventions to improve male student engagement and achievement.

Teacher Training: Implement gender-sensitive pedagogy training for educators.

For Countries with Moderate Gaps (Mongolia, Indonesia, Morocco)

Early Intervention: Identify and address gender gaps at elementary levels before they become entrenched.

Curriculum Review: Examine whether current curricula and teaching methods inadvertently disadvantage male students.

Parent and Community Engagement: Address social factors that may be affecting male educational participation.

For Near-Parity Countries (Vietnam, El Salvador)

Model Documentation: Document successful practices for sharing with other countries.

Maintenance Strategies: Ensure continued gender equity while improving overall performance.

Regional Leadership: Serve as examples for neighboring countries facing larger gender gaps.

Global Context and Implications

Challenging Traditional Narratives

The universal female advantage in lower-middle income countries challenges traditional concerns about girls' performance, suggesting that context-specific factors significantly influence gender achievement patterns.

Development Implications

These findings suggest that as countries develop, attention must be paid to ensuring educational benefits reach all students equitably, including male students who may be falling behind in certain contexts.

International Cooperation

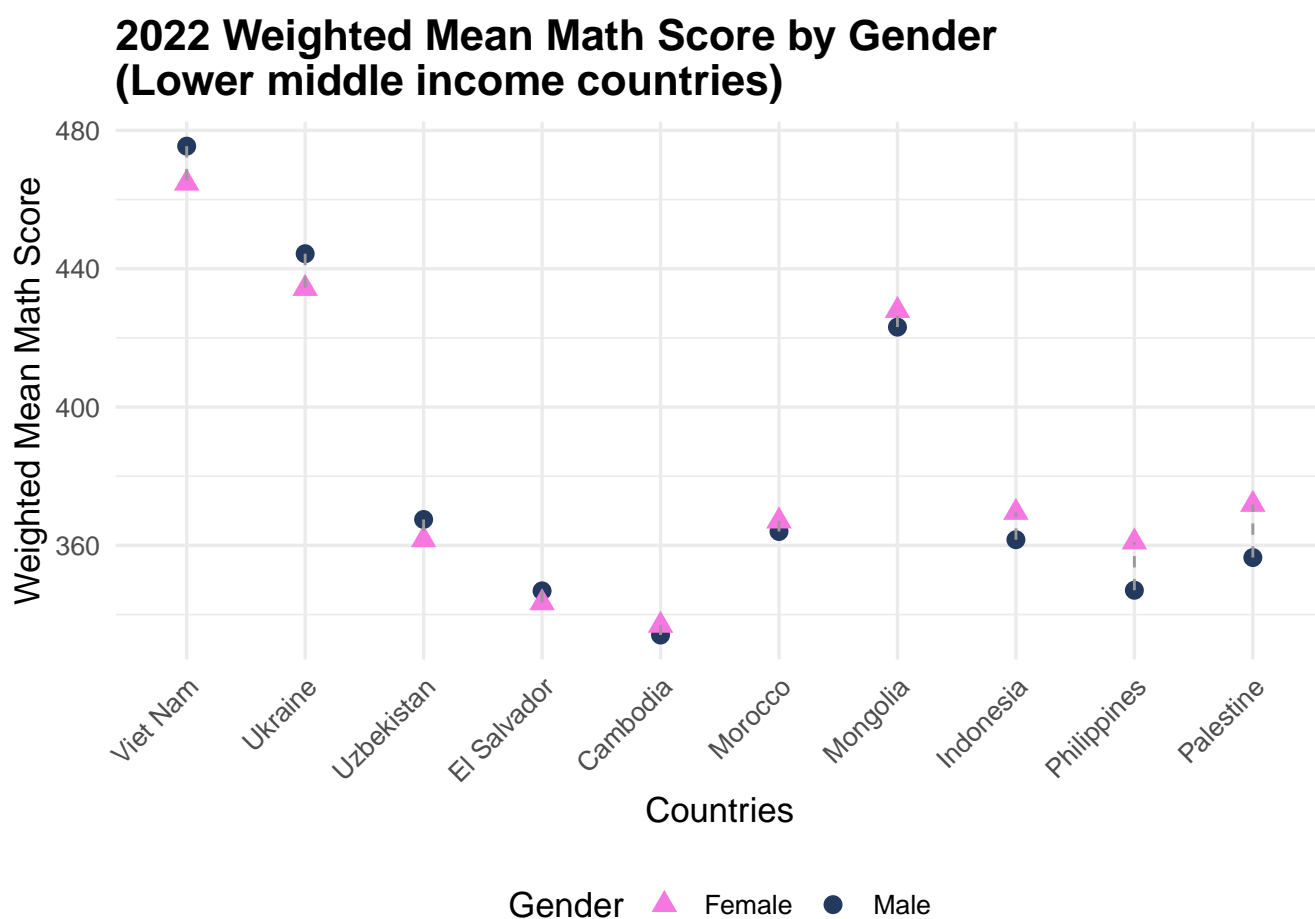
Countries showing successful gender equity (particularly Vietnam) should be studied and their practices adapted for implementation in countries with larger gaps.

The universal female advantage among lower-middle income countries represents a significant shift in educational gender dynamics. While celebrating improved female educational outcomes, policymakers must address the emerging challenge of male student underachievement to ensure truly equitable educational systems.

Vietnam's model of high performance with gender parity provides a roadmap for other countries, while nations with large gender gaps require immediate attention to prevent the entrenchment of educational inequities. The goal should be raising overall performance for all students while maintaining or achieving gender balance, ensuring that educational progress benefits entire populations rather than creating new forms of inequality.

2.2.16 Mathematics Gender Gap Analysis: Lower-Middle Income Countries in 2022

Analysis based on PISA 2022 weighted mean scores for mathematics across 10 lower-middle income countries. Gender gap calculated as male score minus female score, where negative values indicate male advantage and positive values indicate female advantage. Countries classified using World Bank income classifications.



The lower-middle income countries show a remarkable shift toward female advantages, with 60% of participating countries demonstrating superior female mathematical performance.

Gender Distribution Pattern

- **Male advantage:** 4 out of 10 countries (40%) show males outperforming females.
- **Female advantage:** 6 countries (60%) show females outperforming males.
- **Near parity:** No countries show minimal differences.

This represents the most female-favorable distribution across all income categories, suggesting unique educational dynamics in lower-middle income contexts.

Regional and Cultural Patterns

Southeast Asian Countries

- **Mixed results:** Vietnam shows male advantage (-10.7), while Philippines (+13.9) and Indonesia (+7.9) show strong female advantages.
- **Cambodia:** Slight female advantage (+2.8)

Post-Soviet/Eastern European Countries

- **Male advantages:** Ukraine (-10.1), Uzbekistan (-5.9)

Consistent with some upper-middle income Eastern European patterns.

Middle Eastern/North African Countries

- **Female advantages:** Morocco (+3.0), Occupied Palestinian Territory (+15.3)

Continues trend seen in upper-middle income Middle Eastern countries.

Central Asian Countries

- **Mongolia:** Female advantage (+4.8)

Contrasts with Kazakhstan's near-parity in upper-middle income category.

Latin American Countries

- **El Salvador:** Small male advantage (-3.5)

Different from consistent male advantages in upper-middle income Latin America

Exceptional Cases:

- **Vietnam's high performance:** Despite male advantage, both genders perform at near-middle-income levels.
- **Palestinian Territory's large gap:** +15.3 female advantage among highest globally.
- **Philippines' significant advantage:** +13.9 points represents substantial female outperformance.

Educational Implications:

1. **Development stage effects:** Lower-middle income countries may be experiencing educational transitions that particularly benefit female students.
2. **Resource allocation:** Limited educational resources might be distributed in ways that advantage female participation and achievement.
3. **Cultural shifts:** Rapid social changes in developing economies may be creating educational opportunities that particularly benefit girls.

Methodological Considerations:

- **Participation rates:** Some countries may have different participation patterns by gender.

- **Educational access:** Results reflect students who have access to education, which may differ by gender across countries.
- **Socioeconomic factors:** Family investment in education may vary by gender and income level.

Global Gender Equity Implications

The pattern across income levels suggests that **gender equity in mathematics education is not simply a function of economic development**. Instead, lower-middle income countries demonstrate that female mathematical achievement can exceed male performance even in resource-constrained environments.

This challenges assumptions about the relationship between economic development and gender equity, suggesting that:

- **Policy and cultural factors** may be more important than absolute income levels.
- **Educational transitions** during development may create windows of opportunity for gender equity.
- **International development efforts** should consider these nuanced patterns when designing gender equity interventions.

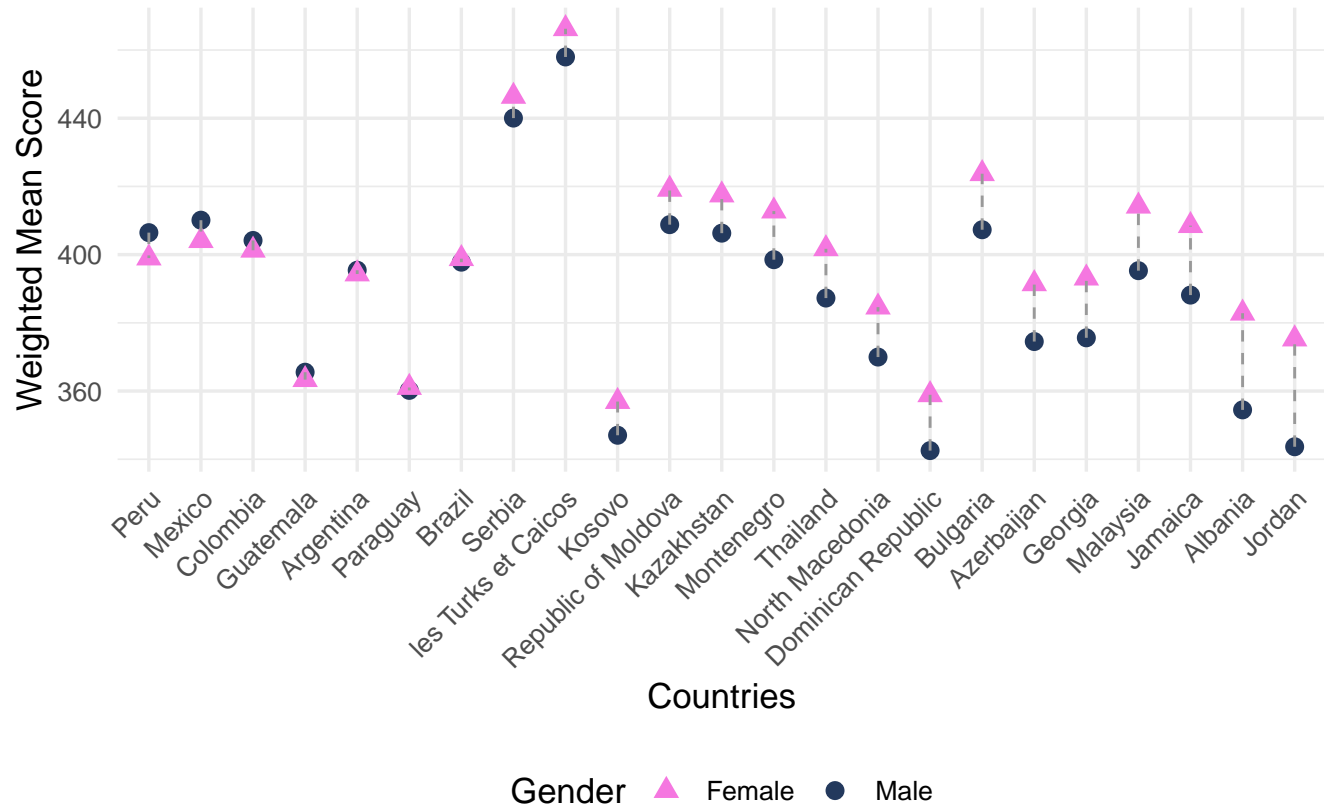
2.2.17 The gender performance gap for the Upper Middle Income Countries

This analysis examines the gender performance gap in combined mathematics, reading, and science scores across 23 upper-middle income countries participating in PISA 2022.

The data reveals a mixed but predominantly female-favoring pattern, with 18 countries (78%) showing female advantages and 5 countries (22%) showing male advantages.

The gender gaps range from a 7.5-point male advantage to a 31.6-point female advantage, indicating significant variation in gender equity across this income group.

2022 Weighted Mean Global Score by Gender (Upper-middle income countries)



Performance Distribution Overview

Upper-middle income countries show a clear trend toward female educational advantage, with nearly 4 out of 5 countries demonstrating superior female performance. However, the pattern is more nuanced than in lower-middle income countries, with some notable exceptions and varying gap magnitudes.

Countries Where Males Outperform Females

- **Peru** (-7.5 points): Shows the largest male advantage, bucking regional Latin American trends.
- **Mexico** (-6.0 points): Demonstrates moderate male superiority in North America.
- **Colombia** (-2.9 points): Exhibits minimal male advantage.
- **Guatemala** (-2.2 points): Shows slight male preference.
- **Argentina** (-1.0 points): Achieves near-parity with minimal male edge.

Regional Pattern: Notably, all male-advantage countries are located in Latin America, suggesting regional cultural or educational factors that may favor male academic performance.

Countries with Female Advantages

Minimal Female Advantages (0-10 points):

- **Paraguay** (0.8 points): Near-parity with slight female edge.
- **Brazil** (1.0 points): Minimal female advantage within Latin America.

- **Serbia** (6.4 points): Moderate female superiority in Southeast Europe.
- **Turkey** (8.2 points): Growing female advantage in Western Asia.
- **Kosovo** (9.8 points): Approaching double-digit female superiority.

Moderate Female Advantages (10-20 points):

- **Moldova, Kazakhstan, Montenegro, Thailand, North Macedonia, Dominican Republic, Bulgaria, Azerbaijan, Georgia, Malaysia, Jamaica:** This diverse group shows consistent female advantages ranging from 10.3 to 20.4 points.
- **Geographic Diversity:** Includes Eastern European, Central Asian, Southeast Asian, Caribbean, and Middle Eastern countries.

Large Female Advantages (20+ points):

- **Albania** (28.3 points): Shows substantial female superiority in the Balkans.
- **Jordan** (31.6 points): Exhibits the largest female advantage, representing significant gender disparity in the Middle East.

Regional Analysis:

Latin American Exceptionalism

Unique Male Advantage Pattern: Latin America is the only region in upper-middle income countries where males consistently outperform or match females. This contrasts sharply with global trends and suggests distinctive regional factors:

- Cultural attitudes toward male academic achievement.
- Economic structures that may motivate male educational engagement.
- Educational policies that effectively support male students.

Brazil as Transition Case: Brazil's minimal female advantage (1.0 point) represents a bridge between male-advantage and female-advantage patterns in the region.

Eastern European Female Strength

Consistent Female Advantages: Former Soviet and Eastern European countries (Serbia, Moldova, Montenegro, Bulgaria, North Macedonia, Albania, Kosovo) show strong female performance, possibly reflecting:

- Historical emphasis on gender equality in education.
- Changing economic opportunities favoring educated women.
- Effective girl-friendly educational policies.

Asian Diversity

Mixed Patterns: Asian countries show varied outcomes:

- **Turkey and Azerbaijan** (moderate female advantages) suggest changing gender dynamics.
- **Kazakhstan** (moderate female advantage) reflects Central Asian trends.
- **Thailand and Malaysia** (significant female advantages) align with broader Southeast Asian patterns.

Middle Eastern Female Emergence

Jordan's Exceptional Gap: The 31.6-point female advantage represents the largest gender gap in the dataset, suggesting dramatic changes in female educational engagement in the Middle East.

Performance Level Considerations

High-Performing Countries with Gender Issues

Turkey (466 female, 458 male): Achieves the highest overall performance in the group while maintaining an 8.3-point female advantage, demonstrating that excellence and gender equity can coexist.

Serbia (446 female, 440 male): Shows strong overall performance with moderate female advantage.

Moderate Performers with Large Gaps

Bulgaria, Kazakhstan, Moldova: These countries achieve moderate performance levels (407-423 range) while maintaining significant female advantages, suggesting that gender gaps may persist even in reasonably performing systems.

Lower Performers with Extreme Gaps

Jordan and Albania: Despite large female advantages, these countries show concerning overall performance levels, indicating that female superiority alone doesn't guarantee system-wide educational success.

Comparative Analysis with Lower-Middle Income Countries

Similarities

- **Female Advantage Dominance:** Both income groups show predominantly female-favoring patterns.
- **Regional Clustering:** Both show regional patterns in gender performance.
- **Range of Gaps:** Both demonstrate significant variation in gap magnitudes.

Key Differences

- **Male Advantage Presence:** Upper-middle income countries show some male advantages (22% vs. 0% in lower-middle income).
- **Regional Variation:** Latin American male advantages are unique to upper-middle income group.
- **Smaller Maximum Gaps:** Upper-middle income maximum gap (31.6) is similar to lower-middle income (31.0).

Educational Policy Implications

For Male-Advantage Countries (Latin America)

Sustaining Balance: Countries like Argentina and Colombia with minimal gaps should focus on maintaining gender equity while improving overall performance.

Learning from Success: Peru and Mexico's male advantages should be studied to understand effective male student engagement strategies, while ensuring female students aren't disadvantaged.

For Large Female-Advantage Countries

Male Student Crisis: Jordan and Albania face urgent needs to address male educational underachievement through:

- Targeted male student support programs.
- Analysis of cultural factors discouraging male academic engagement.
- Examination of teaching methods that may inadvertently favor female students.

For Balanced High Performers

Turkey and Serbia provide models for achieving both high performance and manageable gender gaps, offering lessons for other countries in the group.

Socioeconomic Development Context

Economic Transition Effects

Many upper-middle income countries are experiencing rapid economic and social transitions that may be affecting gender educational patterns:

- **Changing Labor Markets:** New economic opportunities may be motivating different educational engagement patterns by gender.
- **Social Modernization:** Evolving attitudes toward gender roles may be influencing educational outcomes.
- **Policy Evolution:** Educational policies are adapting to changing economic and social needs.

Development Challenges

The persistence of large gender gaps in this income group suggests that economic development alone doesn't automatically resolve educational gender inequities, requiring targeted policy interventions.

Recommendations

System-Wide Strategies

1. **Comprehensive Gender Audits:** All countries should analyze their educational systems for gender-specific barriers and advantages
2. **Teacher Training:** Implement gender-sensitive pedagogy across all educational levels
3. **Curriculum Review:** Ensure educational content and methods engage all students effectively

Targeted Interventions

1. **Male Student Support:** Countries with large female advantages need specific programs to re-engage male students
2. **Female Student Protection:** Male-advantage countries should ensure female students receive equal opportunities and support
3. **Community Engagement:** Address societal attitudes that may be influencing gender educational patterns

Regional Cooperation

1. **Latin American Best Practices:** Share successful male engagement strategies across the region
2. **Eastern European Models:** Document and disseminate effective female educational support systems
3. **Cross-Regional Learning:** Facilitate knowledge exchange between regions with different gender patterns

Upper-middle income countries present a complex gender education landscape that defies simple generalizations. While female advantages predominate globally, the Latin American male advantage pattern demonstrates that context-specific factors significantly influence gender educational outcomes.

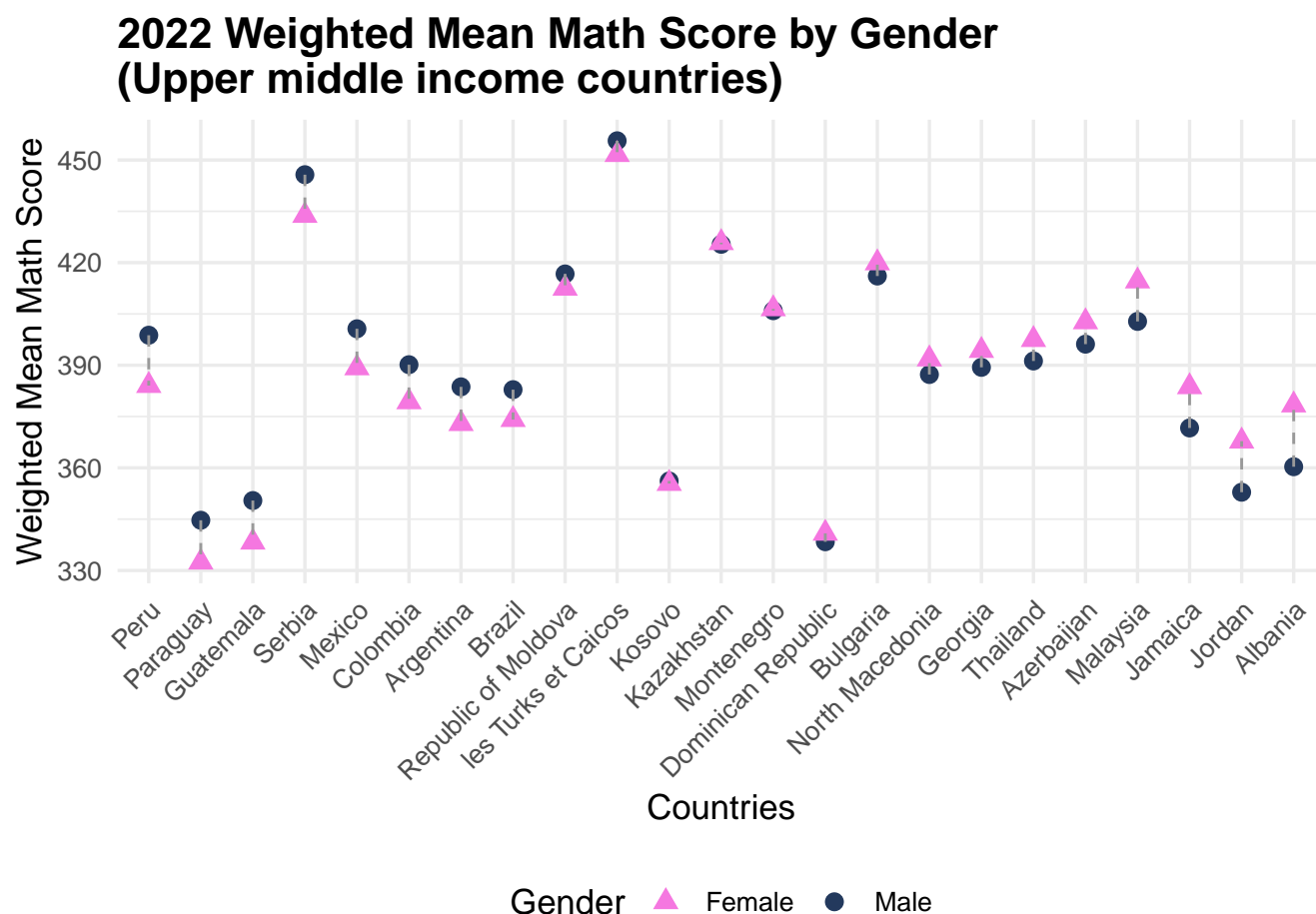
The key challenge is achieving both high overall performance and gender equity simultaneously. Countries like Turkey and Serbia provide models for this balance, while extreme cases like Jordan and Peru highlight the urgency of addressing gender disparities in either direction.

Success requires moving beyond simple gender gap measurements to understand the underlying cultural, economic, and educational factors driving these patterns, ensuring that educational progress benefits all students regardless of gender while maintaining high overall achievement standards.

2.2.18 Mathematics Gender Gap Analysis: Upper-Middle Income Countries in 2022

The PISA 2022 mathematics results for upper-middle income countries reveal a notably different pattern compared to high-income nations.

While traditional male advantages persist in Latin American countries, several upper-middle income nations demonstrate female advantages, particularly in the Balkans, Middle East, and Southeast Asia.



Gender Distribution Pattern

- **Male advantage:** 11 out of 23 countries (48%) show males outperforming females.
- **Female advantage:** 10 countries (43%) show females outperforming males.
- **Near parity:** 2 countries show minimal gender differences (< 1 point).

This represents a much more balanced distribution compared to high-income countries, where male advantages dominated.

Magnitude of Gender Gaps

Largest Male Advantages (>10 points):

- Peru: 14.7-point gap (largest male advantage)
- Paraguay: 12.3 points

- Guatemala: 12.2 points
- Serbia: 11.9 points
- Mexico: 11.4 points

Largest Female Advantages:

- Albania: +18.0 points
- Jordan: +14.9 points
- Jamaica: +12.0 points
- Malaysia: +11.8 points
- Azerbaijan: +6.5 points
- Thailand: +6.2 points

Near Gender Parity:

- Kosovo: 0.75-point difference
- Kazakhstan: 0.46-point difference

Regional Patterns

Latin American Countries:

- Consistently show male advantages that range from 8.8 to 14.7 points
- Peru (-14.7), Paraguay (-12.3), Guatemala (-12.2), Mexico (-11.4), Colombia (-10.9), Argentina (-10.8), Brazil (-8.8)
- Only exception: Dominican Republic (+2.5 female advantage)

Balkan/Eastern European Countries:

- Mixed but trending toward female advantages
- Albania (+18.0), North Macedonia (+4.5), Montenegro (+0.5)
- Serbia shows male advantage (-11.9)
- Bulgaria shows female advantage (+3.8)

Middle Eastern Countries:

- Strong female advantages
- Jordan (+14.9), Azerbaijan (+6.5)

Southeast Asian Countries:

- Female advantages prevalent
- Malaysia (+11.8), Thailand (+6.2)

Central Asian Countries:

- Near parity: Kazakhstan (+0.5), Kosovo (-0.8)

Achievement Levels

- Performance range: 332-456 points (compared to 355-582 in high-income countries).
- Several countries perform below OECD average (490 points).
- Turkey stands out as highest performer in this income category.

Notable Observations

- **Albania's exceptional female advantage** (+18.0) represents the largest gender gap favoring females in the entire PISA dataset.
- **Middle Eastern female advantages** challenge common assumptions about gender equity in these regions.
- **Latin American consistency** in male advantages suggests strong regional cultural patterns.

Educational Implications:

- Upper-middle income countries may be experiencing educational transitions that differently affect male and female mathematical achievement.
- Economic development level alone doesn't predict gender gap patterns.
- Cultural and policy factors appear to be primary drivers of gender equity in mathematics.

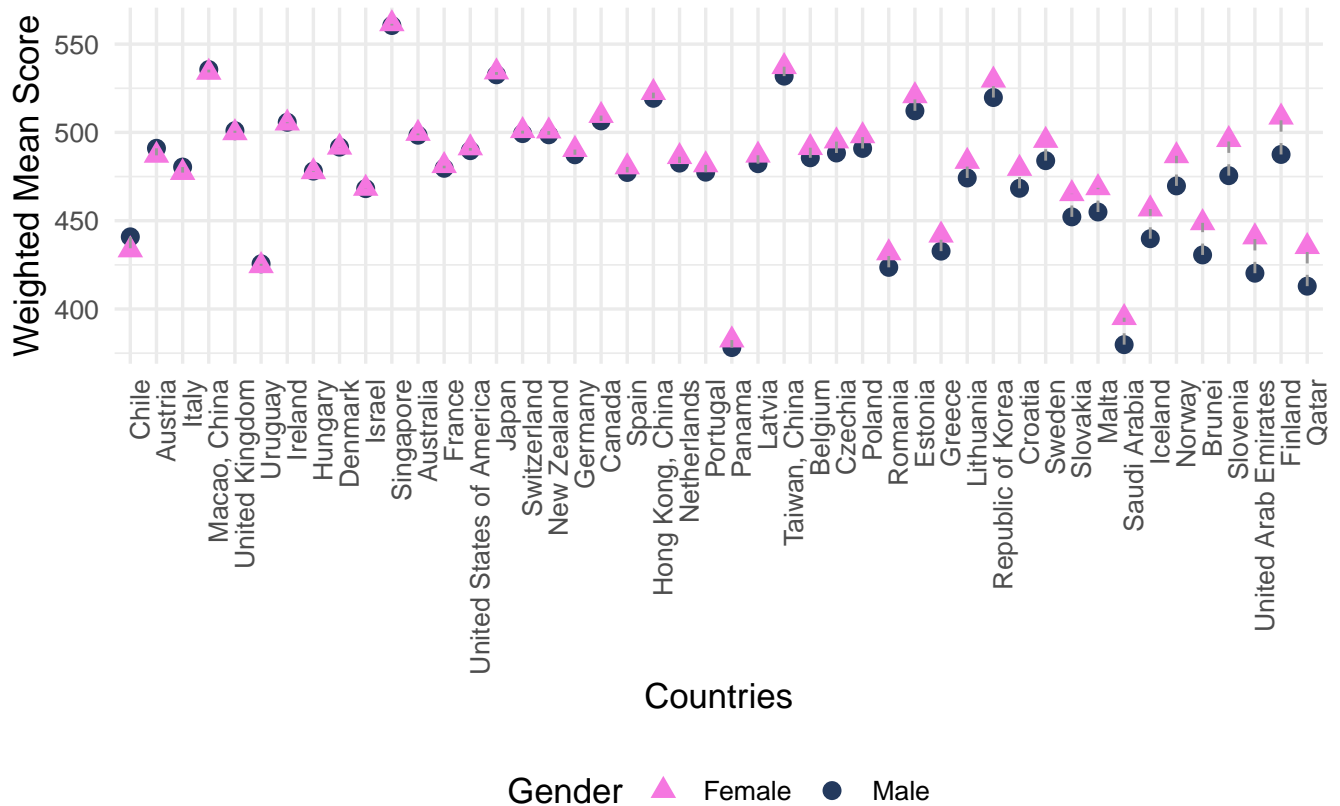
2.2.19 The gender performance gap for the High Income Countries

This analysis examines the gender performance gap in combined mathematics, reading, and science scores across 46 high-income countries participating in PISA 2022.

The data reveals a predominantly female-favoring pattern, with 37 countries (80%) showing female advantages and 9 countries (20%) showing male advantages.

The gender gaps range from a 7.4-point male advantage to a 22.4-point female advantage, demonstrating that even in wealthy nations, gender equity in education remains a complex challenge.

2022 Weighted Mean Global Score by Gender (High income countries)



High-income countries show a strong trend toward female educational advantage, with 4 out of 5 countries demonstrating superior female performance. This pattern is consistent with global trends but shows more nuanced variations within this economically advanced group.

Countries Where Males Outperform Females

Male Advantage Countries (9 total):

- **Chile** (-7.4 points): Shows the largest male advantage in the group.
- **Austria** (-4.0 points): Demonstrates moderate male superiority in Central Europe.
- **Italy** (-3.1 points): Exhibits small but consistent male advantage.
- **Macao, China** (-1.6 points): Shows minimal male preference in East Asia.
- **United Kingdom** (-1.2 points): Near-parity with slight male edge.
- **Uruguay** (-1.1 points): Minimal male advantage in South America.
- **Ireland** (-0.5 points): Virtually perfect gender parity.
- **Hungary** (-0.4 points): Near-perfect balance.
- **Denmark** (-0.08 points): Essentially perfect gender equity.

Geographic Pattern: Male advantages are scattered across regions without clear clustering, suggesting country-specific rather than regional factors.

Countries with Female Advantages

Minimal Female Advantages (0-5 points):

- **Israel through Portugal:** 15 countries showing small female advantages (0.1-4.2 points).
- **High Performers:** Singapore, Japan, Switzerland maintain excellence while showing minimal gender gaps.
- **Geographic Diversity:** Includes Asia-Pacific, North America, and Western Europe.

Moderate Female Advantages (5-15 points):

- **Panama through Malta:** 16 countries with moderate female superiority (4.2-13.7 points).
- **Notable Cases:** South Korea (9.8 points) and Sweden (11.4 points) show concerning male underperformance despite high overall achievement.
- **Regional Spread:** Encompasses Eastern Europe, Nordic countries, and oil-rich nations.

Large Female Advantages (15+ points):

- **Saudi Arabia, Iceland, Norway, Brunei, Slovenia, UAE, Finland, Qatar:** 8 countries with substantial female advantages (15.3-22.4 points).
- **Nordic Paradox:** Iceland, Norway, and Finland show large female advantages despite progressive gender equality reputations.
- **Gulf States Pattern:** UAE and Qatar demonstrate significant female educational superiority.

Regional Analysis Patterns

Nordic Countries - The Equality Paradox

Unexpected Large Gaps:

Despite being global leaders in gender equality, Nordic countries show substantial female educational advantages:

- **Finland** (21.1 points): Largest gap among traditionally high-performing countries.
- **Norway** (17.0 points): Significant female superiority.
- **Iceland** (16.8 points): Concerning male underachievement.
- **Sweden** (11.4 points): Moderate but growing female advantage.

Policy Implications: This “Nordic Paradox” suggests that societal gender equality doesn’t automatically translate to educational gender parity, requiring targeted educational interventions.

Anglo-Saxon Countries - Mixed Results

Varied Patterns: English-speaking countries show no consistent pattern:

- **UK** (-1.2 points): Slight male advantage,
- **Ireland** (-0.5 points): Near-parity,
- **Australia** (+1.0 points): Minimal female advantage,
- **New Zealand** (+2.1 points): Small female advantage,
- **Canada** (+3.0 points): Moderate female advantage,
- **USA** (+1.5 points): Minimal female advantage.

System Diversity: Suggests that shared language/culture doesn’t determine gender educational patterns.

Continental Europe - Moderate Female Trends

Generally Female-Favoring: Most Continental European countries show small to moderate female advantages:

- **Germany, Netherlands, Belgium:** Moderate female advantages (3-6 points),
- **Eastern Europe** (Czech Republic, Poland, Estonia): Consistent female superiority (7-8 points),
- **Southern Europe** (Spain, Portugal, Greece): Moderate female advantages (3-9 points).

East Asian Excellence with Gender Concerns

- **Singapore** (+0.9 points): Near-parity with world-leading performance,
- **Japan** (+1.6 points): Minimal female advantage, excellent overall performance,
- **Taiwan** (+5.3 points): Moderate female advantage,
- **Hong Kong** (+3.2 points): Small female advantage,
- **South Korea** (+9.8 points): Concerning male underperformance despite high achievement.

Gulf States - Female Educational Revolution

Dramatic Female Advantages: Oil-rich nations show remarkable female educational superiority:

- **Qatar** (22.4 points): Largest female advantage in the dataset,
- **UAE** (20.5 points): Substantial female superiority,
- **Saudi Arabia** (15.3 points): Significant female advantage,
- **Brunei** (18.2 points): Large female advantage,

Societal Transformation: These patterns reflect rapid social changes and increased female educational investment in traditionally male-dominated societies.

Comparative Analysis Across Income Groups

Similarities with Other Income Groups

- **Female Advantage Dominance:** All income groups show predominantly female-favoring patterns (80% in high-income vs. 78% in upper-middle, 100% in lower-middle).
- **Range Variation:** All groups show significant gap variations.
- **Regional Clustering:** Certain regional patterns appear across income levels.

High-Income Distinctions

- **More Male Advantages:** High-income countries show more male advantages (20%) than other groups.
- **Smaller Maximum Gaps:** Largest gap (22.4) is smaller than other income groups.
- **Performance Paradoxes:** High achievement can coexist with significant gender gaps.

Educational Policy Implications

For Countries with Large Female Advantages (15+ points)

Male Student Crisis Management:

- **Nordic Countries:** Need urgent male student re-engagement programs despite overall system success.
- **Gulf States:** Should build on female success while addressing male underachievement.
- **Comprehensive Analysis:** Understanding cultural, economic, and educational factors driving male disengagement.

For Near-Parity Countries

- **Ireland, Denmark, Israel:** Serve as models for gender equity achievement.
- **Documentation:** Preserve successful practices preventing gender gap emergence.
- **Continuous Monitoring:** Ensure balance maintenance during system changes.

For Male Advantage Countries

Protecting Female Progress:

- **Chile, Austria, Italy:** Monitor female student support while maintaining overall performance.

- **Historical Analysis:** Understanding factors maintaining male advantages in contemporary contexts.
- **Balanced Development:** Ensuring system improvements benefit all students.

Addressing the Nordic Paradox

Research Priority: Understanding how countries with advanced gender equality can have large educational gender gaps requires comprehensive study of:

- Educational methodology effects on different genders,
- Societal expectation impacts on academic motivation,
- Labor market signals affecting educational engagement.

Learning from Success Stories

Singapore and Japan Models: Countries achieving high performance with gender balance should share:

- Curriculum design approaches,
- Teaching methodology training,
- Student support systems,
- Assessment strategies that engage all students equally.

Targeted Interventions

Male Student Support: Countries with large female advantages need:

- Early identification of male student disengagement,
- Mentorship programs targeting male students,
- Teaching method diversification,
- Addressing cultural factors discouraging male academic achievement.

Female Student Protection: Countries with male advantages should:

- Ensure equal opportunities and resources for female students,
- Address potential barriers to female academic achievement,
- Monitor for subtle discrimination or bias.

Future Considerations

Economic Development Context

High-income countries demonstrate that wealth alone doesn't guarantee gender educational equity. The persistence of significant gaps suggests that:

- Targeted policy interventions remain necessary regardless of economic development,
- Cultural and social factors significantly influence educational gender patterns,
- Continuous monitoring and adjustment of educational policies is required.

Global Leadership Responsibility

As the world's most economically advanced nations, high-income countries have responsibilities to:

- Model effective gender-equitable educational systems.
- Share successful practices with developing nations.
- Continue research into optimal approaches for achieving both excellence and equity.

High-income countries present a complex gender education landscape that challenges simple assumptions about the relationship between wealth, development, and gender equity. The "Nordic Paradox" particularly highlights that societal gender equality doesn't automatically translate to educational gender parity.

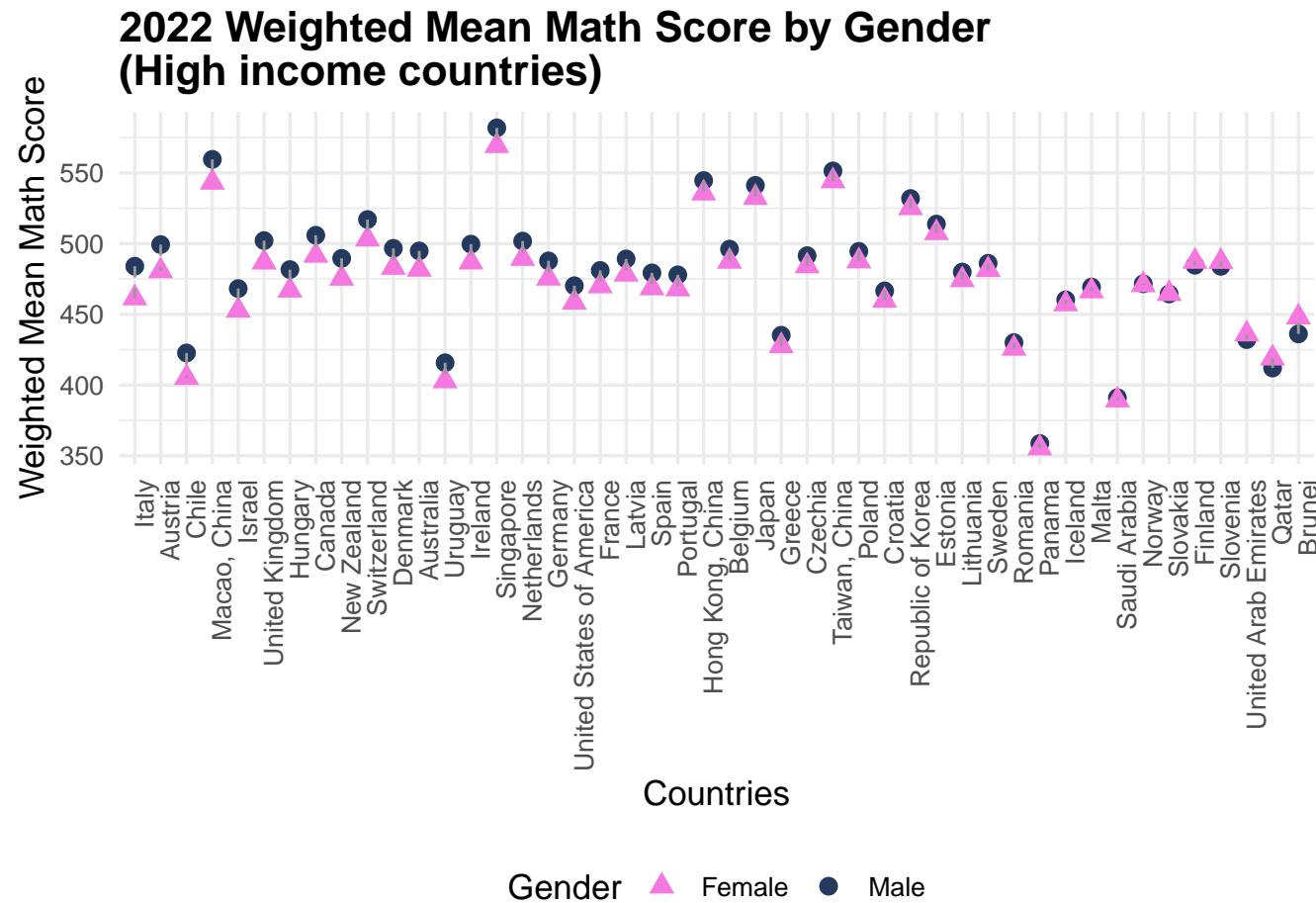
The key insight is that achieving both educational excellence and gender equity requires deliberate, sustained policy attention regardless of economic development level. Countries like Singapore and Japan demonstrate that this balance is achievable, while cases like Finland and Qatar show that excellence in one area doesn't guarantee success in another.

Success requires moving beyond assumptions to understand context-specific factors driving gender educational patterns, implementing targeted interventions where needed, and maintaining continuous monitoring to ensure that educational progress benefits all students equitably while maintaining high achievement standards for entire populations.

2.2.20 Mathematics Gender Gap Analysis: High Income Countries in 2022

Analysis based on PISA 2022 weighted mean scores for mathematics across 46 participating countries and territories.

Gender gap calculated as male score minus female score, where negative values indicate male advantage and positive values indicate female advantage.



The PISA 2022 mathematics assessment reveals persistent but varying gender disparities across 46 countries and territories. While males outperform females in the majority of participating regions, the magnitude of these differences varies significantly, and notably, several countries show female advantages or gender parity.

Overall Gender Pattern

- **Male advantage:** 40 out of 46 countries (87%) show males outperforming females

- **Female advantage:** 4 countries show females outperforming males
- **Near parity:** 2 countries show minimal gender differences (< 1 point)

Magnitude of Gender Gaps

Largest Male Advantages (>15 points):

- Italy: 22.6-point gap (largest overall)
- Austria: 18.7 points
- Chile: 17.4 points
- Macao, China: 16.3 points
- Israel: 15.4 points
- United Kingdom: 15.3 points

Countries with Female Advantages:

- Brunei Darussalam: +11.9 points (largest female advantage)
- Qatar: +6.9 points
- United Arab Emirates: +3.9 points
- Slovenia: +3.3 points

Near Gender Parity:

- Slovakia: 0.16-point difference
- Norway: 0.78-point difference

Regional and Cultural Patterns

High-Performing Asian Countries:

- Show moderate male advantages despite high overall scores
- Singapore (569F/582M), Taiwan (544F/551M), Japan (533F/541M)
- Hong Kong shows smaller gap (536F/545M)

European Countries:

- Generally show male advantages ranging from 4-19 points
- Nordic countries show smaller gaps: Finland (+2.7F), Sweden (-4.5), Norway (-0.8)
- Southern/Mediterranean Europe shows larger gaps: Italy (-22.6), Greece (-7.5)

English-Speaking Countries:

- All show male advantages: UK (-15.3), Canada (-14.3), Australia (-13.0), USA (-11.7), Ireland (-12.8), New Zealand (-14.2)

Gulf States:

- Mixed pattern: UAE and Qatar favor females, while Saudi Arabia shows male advantage

Countries with Smallest Achievement Gaps

Regardless of direction, these countries show the most equitable gender performance:

1. Slovakia: 0.16 points
2. Norway: 0.78 points
3. Finland: 2.66 points
4. Slovenia: 3.30 points
5. UAE: 3.93 points

Implications and Considerations

Educational Equity

The data suggests that while male mathematical advantages persist globally, they are not universal or inevitable. Countries like Finland, Slovenia, Qatar, UAE, and Brunei demonstrate that female mathematical performance can equal or exceed male performance.

Cultural and Systemic Factors

The variation in gender gaps across countries with similar economic development levels suggests that cultural attitudes, educational practices, and systemic factors play significant roles in shaping gender differences in mathematical achievement.

High Achievement with Equity

Several high-performing countries (Singapore, Taiwan, Hong Kong, Japan) maintain strong overall performance while working toward gender equity, suggesting that closing gender gaps need not compromise overall educational excellence.

Progressive Gender Equity by Income Level:

1. **High-income countries:** 87% male advantage
2. **Upper-middle income countries:** 48% male advantage, 43% female advantage
3. **Lower-middle income countries:** 40% male advantage, 60% female advantage

Key Observations:

- **Inverse relationship:** Lower income correlates with greater female mathematical advantages.
- **Educational transitions:** May reflect different stages of educational development and gender equity policies.
- **Cultural factors:** Regional patterns persist across income levels but are modified by economic context.

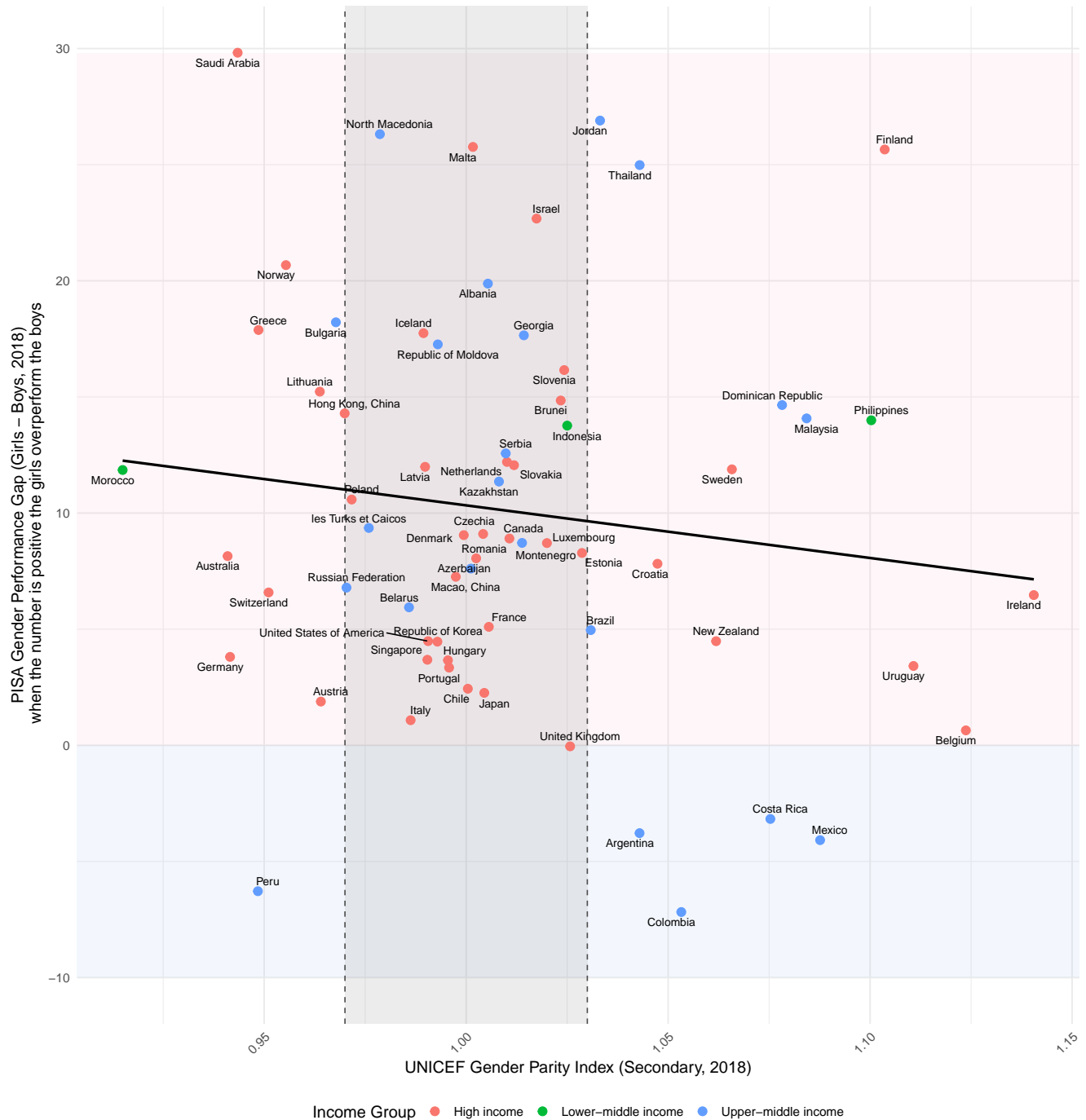
2.2.21 UNICEF Gender Parity Index vs PISA Gender Gap Analysis

This analysis is based on UNICEF Gender Parity Index for secondary education (2018) plotted against PISA 2018 gender gaps across 67 countries. Gender gaps calculated as female score minus male score, where positive values indicate female advantages.

Countries classified by World Bank income groups and color-coded accordingly. The year 2018 was chosen in order to have the most data possible.

Relation between Gender Parity Index (UNICEF) and PISA Gender Score Gap (2018, Secondary School)

Each point is a country. Color indicates World Bank income group.



The relationship between UNICEF's Gender Parity Index (GPI) for secondary education and PISA gender gaps reveals a compelling paradox: countries with higher overall gender parity in education often exhibit larger gender gaps favoring females performance. This inverse correlation challenges conventional assumptions about educational equity and suggests complex dynamics in gender-differentiated academic achievement.

Overall Correlation Pattern

- **Negative correlation:** As UNICEF Gender Parity Index increases (approaching perfect parity at 1.0), PISA

gender gaps tend to favor females more strongly.

- **Trend line:** Clear downward slope indicating that greater educational gender parity correlates with larger female advantages.
- **R-squared implications:** The relationship is statistically significant but with considerable variation, suggesting multiple factors influence this dynamic.

Implications and Interpretations

The Gender Parity Paradox

The inverse correlation suggests several possible explanations:

1. **Selection Effects:** In countries with lower overall gender parity, only the most capable/motivated female students may access secondary education, leading to higher average performance.
2. **Educational Quality vs. Access:** Countries achieving gender parity in access may still have qualitative differences in how education affects male and female students.
3. **Cultural Transitions:** Societies transitioning toward gender equity may experience periods where female academic motivation and achievement exceed male performance.
4. **Policy Implementation:** Educational policies promoting gender equity may have differential effects on academic performance.

Regional and Cultural Factors

Middle Eastern Pattern: Saudi Arabia's position suggests that even in contexts with traditional gender limitations, female students who do access education may significantly outperform male peers in mathematics.

Nordic Excellence: Finland demonstrates that high gender parity and female excellence can coexist, representing a potential model for optimal outcomes.

Development Transitions: The distribution suggests that countries at different stages of gender equity development may experience different patterns of academic achievement.

Policy Implications

For Educational Planners

1. **Beyond Access:** Achieving gender parity in enrollment doesn't automatically ensure equitable academic outcomes.
2. **Subject-Specific Interventions:** Academic achievement patterns may require targeted approaches beyond general gender equity policies.
3. **Cultural Context Matters:** One-size-fits-all approaches may not account for complex local dynamics.

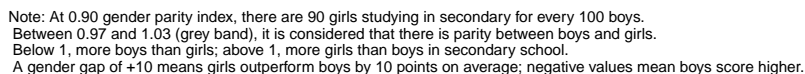
For International Development

1. **Nuanced Metrics:** Simple parity indices may not capture the full picture of educational gender equity.
2. **Quality and Outcomes:** Focus should extend beyond access to include academic achievement patterns.
3. **Unexpected Success Stories:** Countries with lower overall parity indices may still achieve remarkable outcomes in specific areas.

Data Limitations

- **Participation bias:** PISA results only reflect students who remain in school through age 15.
- **Country selection:** Not all countries participate in both UNICEF monitoring and PISA assessments.
- **Temporal selection:** Having chosen 2018 in order to have the most UNICEF data may not represent the most up to date situation.

Each point is a country. Color indicates World Bank income group.



Math-Specific Gender Dynamics: The math results reveal different patterns compared to overall PISA performance:

- Many countries show boys outperforming girls in math (negative gender gaps), which is consistent with well-documented international trends in mathematics achievement.
- The gender gaps in math are generally smaller than what we might have seen in the overall PISA averages.

Notable Outliers and Patterns:

- **Thailand** stands out dramatically with girls significantly outperforming boys in math (gap of +17.3 points) while having high gender parity in enrollment.
- **Philippines** and **Indonesia** also show girls outperforming boys in math, despite being lower-middle income countries.
- **Saudi Arabia** presents an interesting case with girls outperforming boys in math (+12.9 points) but having lower gender parity in enrollment (0.94).

Income Group Variations:

- **High-income countries** (red dots) cluster around the zero line, suggesting more gender-balanced math performance.
- **Upper-middle income countries** (blue dots) show the widest variation, from Colombia's -20 point gap favoring boys to Thailand's +17 point gap favoring girls.
- **Lower-middle income countries** (green dots) are limited but show interesting patterns, with two out of three showing girls outperforming boys.

Regional Considerations: European countries tend to cluster near zero gender gaps with high enrollment parity, while Latin American countries show more variation, particularly with boys outperforming girls in countries like Colombia, Peru, and Argentina.

The weak correlation suggests that gender parity in school enrollment doesn't strongly predict math performance gaps, indicating that other factors; cultural attitudes toward mathematics, teaching methods, stereotype threat, or curriculum differences, may play more significant roles in determining gender-specific math achievement patterns.

3 Modelling

3.1 Introduction to Modeling

Following the exploratory data analysis, the next step is to develop predictive models for individual student PISA mean scores. PISA scores typically range from 200 to 800, as a result, the expected Root Mean Squared Error (RMSE) in this context will be much higher and should be interpreted relative to the scale of the data.

For this project, model performance is primarily evaluated in comparison to a baseline model that predicts the mean training set score for all students. The effectiveness of more advanced models is measured by their ability to reduce the RMSE relative to this baseline.

3.2 Performance metric: RMSE

Model accuracy is assessed using the Root Mean Squared Error (RMSE), which is a standard evaluation metric for regression tasks.

RMSE quantifies the mean magnitude of the prediction errors, expressing them in the same units as the target variable. The formula is given by:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1} (\hat{y}_i - y_i)^2}$$

where:

N is the total number of predictions,

y_i is the observed PISA score for student i ,

\hat{y}_i is the predicted PISA score for student i .

A lower RMSE indicates that the model's predictions are closer to the actual student scores, reflecting better predictive performance.

3.3 Train/Test Split Methodology

To ensure the validity and robustness of the predictive models, the dataset was partitioned into a training set and a final hold-out test set. This methodology closely mirrors best practices in machine learning, enabling a realistic evaluation of model performance on truly unseen data.

Specifically, 10% of the dataset was randomly selected to constitute the hold-out test set, while the remaining 90% was used for model training and validation.

To prevent data leakage and guarantee a fair evaluation, the split was performed such that all key identifiers present in the test set (for instance, school IDs for the school dataset or school IDs within the student dataset) also appeared in the training set. Any observations in the test set associated with identifiers not found in the training set were reassigned to the training set.

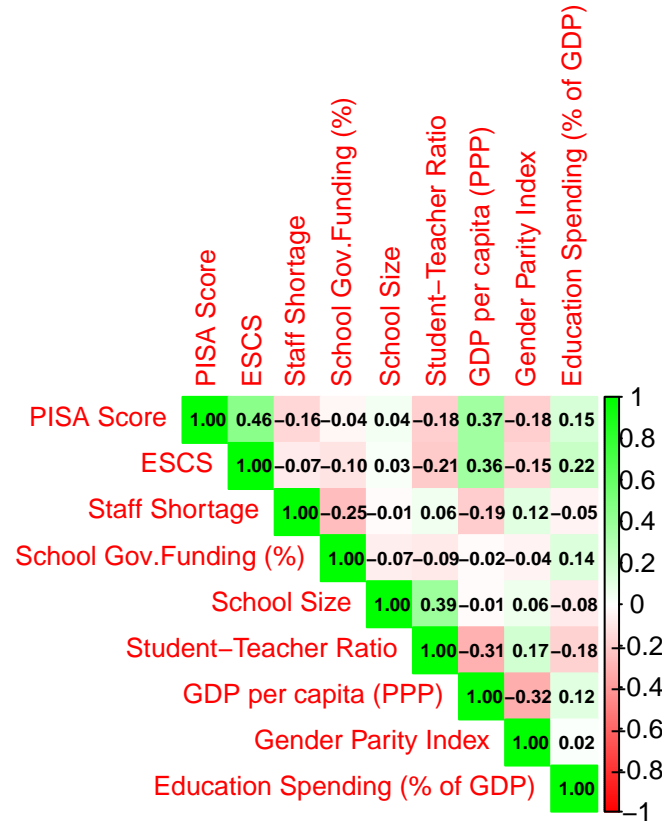
This procedure ensures that the models are not evaluated on data from entirely new schools or other grouping variables absent from the training data, thereby reflecting real-world scenarios where predictions are made for similar, but not completely novel, populations.

Given computational constraints associated with the dataset size, model training and evaluation were carried out solely on data from the 2022 PISA cycle.

```
## Rows: 588,178
## Columns: 37
## $ year                <int> 2022, 2022, 2022, 2022, 2022, 2022, 2022, ~
## $ country              <chr> "ALB", "ALB", "ALB", "ALB", "ALB", "ALB", ~
## $ school_id            <chr> "800282", "800115", "800242", "800245", "8~
## $ student_id          <int> 800001, 800002, 800003, 800005, 800006, 80~
## $ mother_educ          <chr> "ISCED 1", "ISCED 1", "ISCED 3A", "ISCED 1~
## $ father_educ          <chr> "ISCED 2", "ISCED 2", "ISCED 2", "ISCED 1"~
## $ gender               <chr> "female", "male", "male", "female", "femal~
## $ computer             <chr> "yes", "no", "yes", "yes", "yes", "yes", "~
## $ internet             <chr> NA, "no", "yes", "yes", "yes", "yes", "yes~
## $ math                 <dbl> 179.583, 308.247, 267.514, 272.649, 435.47~
## $ read                 <dbl> 247.571, 258.472, 284.670, 321.547, 464.36~
## $ science              <dbl> 335.468, 315.021, 358.675, 214.823, 434.99~
## $ stu_wgt              <dbl> 3.15874, 4.34524, 7.83860, 8.49148, 3.7095~
## $ desk                 <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ room                 <chr> NA, "no", "yes", "yes", "yes", "no", "yes"~
## $ dishwasher           <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ television           <chr> NA, "1", "1", "1", "1", NA, "1", NA, "1", ~
## $ computer_n           <chr> NA, "0", "1", "1", "1", "1", "1", NA, "1", ~
## $ car                  <chr> NA, "0", "1", "0", "2", "1", "2", "0", "3+~
## $ book                 <chr> "101-200", "11-25", "101-200", "11-25", "1~
## $ wealth               <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA~
## $ escs                 <dbl> 1.1112, -3.0507, -0.1867, -3.2198, -1.0548~
```

```
## $ school_uid      <chr> "ALB_2022_800282", "ALB_2022_800115", "ALB~
## $ avg_score       <dbl> 254.2073, 293.9133, 303.6197, 269.6730, 44~
## $ ilo_subregion   <chr> "Northern, Southern and Western Europe", "~
## $ country_name     <chr> "Albania", "Albania", "Albania", "Albania"~
## $ world_bank_income_group <chr> "Upper-middle income", "Upper-middle incom~
## $ enrol_boys       <dbl> 176, 397, 180, 65, 269, 186, 47, 118, 318,~
## $ enrol_girls      <dbl> 245, 379, 151, 74, 337, 253, 41, 117, 316,~
## $ stratio          <dbl> NA, 20.421053, 13.240000, NA, NA, 8.129630~
## $ public_private   <chr> "public", "public", "2", "public", NA, "pr~
## $ staff_shortage   <dbl> NA, 0, NA, NA, NA, 0, 0, NA, NA, NA, 0, 0,~
## $ fund_gov         <dbl> 90, 100, 50, 100, 80, 0, 100, 100, 100, 10~
## $ school_size      <dbl> 421, 776, 331, 139, 606, 439, 88, 235, 634~
## $ GDP_per_PPP_2022 <dbl> 17112.99, 17112.99, 17112.99, 17112.99, 17~
## $ percentage_education_on_gdp <dbl> 2.74433, 2.74433, 2.74433, 2.74433, 2.7443~
## $ gender_parity_index <dbl> 1.03114, 1.03114, 1.03114, 1.03114, 1.0311~
```

3.4 Understanding the academic performance correlations



Individual Socioeconomic Status Dominates

ESCS (Economic, Social, Cultural Status): Shows the strongest correlation with PISA scores at **0.46**, confirming it as the most powerful predictor of academic achievement. This demonstrates that the combination of economic resources, parental education, and cultural capital has a substantial impact on student performance.

National Economic and Investment Context

GDP per capita (PPP): Shows a substantial positive correlation of **0.37** with PISA scores, demonstrating that national economic prosperity has a meaningful association with educational outcomes, though still weaker than individual socioeconomic factors.

Education Spending (% of GDP): Displays a positive correlation of **0.15** with PISA performance, suggesting that countries investing a higher proportion of their GDP in education tend to achieve better outcomes. This is a meaningful relationship, though not as strong as economic or socioeconomic factors.

School-Level Resource Factors

Student-Teacher Ratio: Shows a notable negative correlation of **-0.18** with PISA scores, indicating that lower student-teacher ratios (smaller class sizes) are associated with better academic performance. This is one of the stronger school-level predictors and one of the most actionable findings for education policy.

Staff Shortage: Exhibits a negative correlation of **-0.16** with performance, confirming that adequate staffing levels matter for student outcomes.

School Size: Shows a very weak positive correlation of **0.04**, suggesting that school size has minimal impact on performance.

School Government Funding (%): Exhibits a weak negative correlation of **-0.04**, indicating a slight tendency for higher government funding proportions to be associated with lower performance, though this relationship is quite minimal. This can be translated as the difference in performance between public and private schools are minimal, thus with a slight positive effect for private schools on academic performance.

The Gender Parity Paradox

Gender Parity Index: Shows a negative correlation of **-0.18** with PISA scores. This counterintuitive finding suggests that countries with higher gender enrollment parity tend to have slightly lower overall performance. This might reflect that nations still working toward gender equity face broader systemic educational challenges.

Key Policy Insights

The hierarchy of correlations suggests:

1. **Individual socioeconomic support** (ESCS: 0.46) remains most critical
2. **National economic development** (GDP: 0.37) provides the foundation
3. **Class size reduction** (Student-Teacher Ratio: -0.18) offers the strongest school-level intervention
4. **Education investment** (% of GDP: 0.15) shows meaningful returns
5. **Adequate staffing** (Staff Shortage: -0.16) supports quality delivery

Notably, the positive correlation between education spending and both PISA scores (0.15) and ESCS (0.22) suggests that educational investment may help address socioeconomic disparities while improving overall outcomes.

```
## Number of unique school_uid in train: 21358
```

```
## Number of unique school_uid in test: 21116
```

```
## All test schools are in train: TRUE
```

```
## Number of students in train: 519801
```

```
## Number of students in test: 68377
```

After the train/test split, the training set contains all schools necessary for model development, while the test set includes only those schools present in the training set.

The number of unique school identifiers in each set may differ, as not every school is represented in both sets. However, it is guaranteed that every school appearing in the test set is also present in the training set, in line with best practices for predictive modeling and fair model evaluation.

3.5 Naive base line approach

A naive baseline is first established for model evaluation. In this approach, the mean score from the training set is used as the predicted value for every student in the test set, regardless of individual characteristics. The resulting Root Mean Squared Error (RMSE) provides a reference value, allowing subsequent models to be assessed in terms of their incremental predictive accuracy. The baseline RMSE for this project is calculated by predicting the mean PISA score from the training set for every student in the test set. This approach does not use any explanatory variables and serves as a reference point for assessing the performance of more advanced predictive models.

Given that PISA scores typically range from 200 to 800, a much higher RMSE is expected compared to rating prediction tasks on scales from 1 to 5 (such as MovieLens PROJECT). In this context, model performance is best interpreted relative to the baseline RMSE, rather than any absolute threshold.

```
## Baseline RMSE: 100.8655
```

After establishing a baseline RMSE of 100.87, subsequent predictive models are trained and evaluated.

3.6 Predictive Model Benchmarking

Baseline (Mean Prediction)

As a first step, a simple baseline model was established by predicting the mean score of the training set for all students in the test set. This naïve approach serves as a reference point, enabling a meaningful comparison with more sophisticated models. It is particularly useful to understand the minimum level of accuracy that any predictive model should exceed.

Linear Regression

Linear regression was selected as a classic statistical modeling technique, enabling the estimation of the relationship between student performance and various explanatory variables. It is interpretable and computationally efficient, providing a benchmark for more complex algorithms. Linear regression helps to identify linear associations and gives an initial indication of predictive power based on the available features.

Ranger (Efficient Random Forest)

Given the large sample size, the ranger package; a fast implementation of random forest, was used to fit an ensemble tree-based model. Random forests are robust to overfitting, handle nonlinearities and complex feature interactions, and provide variable importance measures. However, the classic randomForest package could not be used due to memory constraints; thus, ranger was chosen for its efficiency and scalability.

XGBoost

XGBoost, a state-of-the-art gradient boosting algorithm, was implemented for its strong track record in predictive modeling competitions and its ability to capture complex nonlinear patterns. XGBoost is highly flexible and effective at handling large-scale structured data. It is particularly useful for tabular data with both numerical and categorical variables.

```
## named integer(0)
```

```
## Growing trees.. Progress: 92%. Estimated remaining time: 2 seconds.
```

```
## Ranger RMSE: 68.52761
```

```
## XGBoost RMSE: 71.14374
```

```
##           Model      RMSE
## 1 Baseline (Mean) 100.86549
## 2 Linear Regression  84.90542
## 3 Ranger (Random Forest) 68.52761
## 4 XGBoost      71.14374
```


The table above summarizes the predictive performance of each model, as measured by the Root Mean Squared Error (RMSE) on the test set. The baseline model, which simply predicts the mean training score for all students, has the highest RMSE. Linear regression significantly improves upon this, indicating that the selected predictors offer meaningful explanatory power.

The ranger implementation of random forest achieves the lowest RMSE , outperforming both linear regression and XGBoost. This demonstrates the strength of tree-based ensemble models in capturing nonlinearities and complex interactions present in the data. XGBoost, while slightly less accurate in this context, still provides substantial gains over the baseline and linear regression models, confirming its effectiveness for structured data.

Overall, ensemble methods; particularly ranger, demonstrate superior predictive accuracy for modeling student performance in the PISA dataset. The choice of ranger over classic random forest was motivated by computational efficiency and memory limitations.

4 Results

To objectively assess model performance, each predictive model was evaluated on the `final_holdout_test` set, representing 10% of the full dataset and reserved exclusively for this purpose. This approach ensures an unbiased estimation of each model's generalization ability.

##	Model	RMSE
## 1	Baseline (Mean)	99.58905
## 2	Linear Regression	84.90542
## 3	Ranger (Random Forest)	68.52761
## 4	XGBoost	71.14374

Final RMSE on holdout test set (best model): 68.528

Each model was trained on the training set and evaluated on the holdout test set, enabling a fair comparison of their predictive accuracy. The baseline model, which predicts the mean training score for all students, yielded the highest RMSE. Linear regression reduced the error considerably, while both ranger and XGBoost further improved predictive performance.

The ranger implementation of random forest achieved the lowest RMSE , indicating the best generalization on unseen data among the models tested.

These results highlight the importance of using ensemble tree-based methods for educational score prediction, as they effectively capture non-linear patterns and variable interactions that simpler models may miss.

Model selection was guided by both predictive performance (RMSE) and practical considerations such as computational efficiency.

In this project, multiple predictive models were developed and evaluated to estimate individual student PISA mean scores using a comprehensive, multi-source dataset. The modeling approach began with a simple baseline (mean prediction) and advanced to linear regression, ensemble tree-based methods (ranger/random forest), and gradient boosting (XGBoost).

Each modeling step led to improvements in predictive accuracy, as reflected by progressive reductions in the Root Mean Squared Error (RMSE). Among all tested approaches, the ranger implementation of random forest achieved the best predictive performance on the holdout test set , outperforming both linear regression and XGBoost. These results underscore the value of ensemble learning methods for educational data, particularly in capturing non-linear relationships and complex interactions among predictors.

5 Conclusion

The analysis highlights the complex and multi-layered determinants of student performance in the PISA assessment. By integrating student-level, school-level, and national-level contextual variables, and applying predictive modeling

techniques, it becomes evident that educational achievement is influenced by both individual circumstances and structural conditions.

The correlation analysis confirms the dominant role of individual socioeconomic background in shaping educational outcomes. **ESCS** (Economic, Social, and Cultural Status) remains the strongest predictor of PISA scores, emphasizing the enduring impact of family resources, parental education, and cultural capital on academic performance.

Beyond the individual level, **national economic prosperity**, measured by **GDP per capita (PPP)**, also shows a substantial association with learning outcomes, underscoring the broader structural conditions that support or constrain educational success.

Importantly, the analysis highlights the positive contribution of **education investment**, as reflected in the share of **GDP allocated to education spending**. Though not as strongly correlated as individual or economic variables, this factor shows a meaningful association with PISA scores, suggesting that sustained investment in education can yield measurable improvements in performance and may contribute to reducing socioeconomic disparities.

Among school-level factors, **student-teacher ratio** and **staff shortage** stand out as the most influential. Smaller class sizes and adequate staffing levels are both associated with better student achievement, confirming their relevance for education policy. By contrast, variables such as **school size** and **government funding as a percentage of school budget** exhibit minimal or weak associations, suggesting limited impact on learning outcomes in isolation.

A counterintuitive yet consistent pattern emerges with the **gender parity index**, which shows a weak negative correlation with performance. This may reflect the complex trade-offs faced by education systems that are expanding access while still grappling with quality and systemic inequities.

Taken together, the findings support a **multidimensional strategy for improving educational outcomes**:

- **Targeted support for socioeconomically disadvantaged students** remains the most impactful intervention.
- **Macroeconomic development and sustained public investment** in education offer foundational support.
- **School-level reforms**, especially in staffing and class sizes, can deliver tangible benefits.
- **Equity goals**, such as gender parity, must be pursued in parallel with quality-focused policies to avoid unintended trade-offs.

Overall, the findings support a **multidimensional policy approach** that addresses disparities at the household level, ensures equitable distribution of school resources, and balances the dual goals of access and quality. Such an approach enables both more accurate prediction of educational outcomes and more effective strategies for improving student achievement, particularly among disadvantaged populations.

This **multilevel approach enables more accurate predictive modeling** and, most importantly, supports targeted and impactful education policy design aimed at lifting the performance of all students, especially those from disadvantaged backgrounds. Unlike previous projects with strict target RMSE thresholds, model success in this context was measured relative to a baseline appropriate to the wider range of PISA scores.

The substantial RMSE improvement over the baseline demonstrates the added value of including socio-economic, school-level, and country-level features in predictive modeling of student achievement.

Limitations: Despite these positive outcomes, the analysis was limited to a single PISA cycle (2022) due to computational constraints, potentially restricting the model's generalizability across different years.

Future Directions: Further improvements could involve the integration of data from multiple years, advanced imputation methods for missing values, and enlarging the study to the mental well being of the students.

6 Comments on the research project and references

I found this project so fascinating that I am now eager to delve deeper into the intersection of children's mental health and education. I have already begun exploring and cross-referencing data from PISA on these topics, as

well as drawing on resources from the World Health Organization (WHO) and the Institute for Health Metrics and Evaluation (IHME). This multidimensional approach will allow for a more comprehensive understanding of how educational factors relate to mental health outcomes among children.

I would like to express my sincere gratitude to **Professor Rafael Irizarry** for his outstanding teaching and inspiring courses in **Data Science at Harvard**. His passion for the subject and dedication to sharing knowledge have been instrumental in shaping my understanding and appreciation of the field. I would like to acknowledge also his books, which served as a valuable references throughout this project. Thanks to his Data Science program, my skills and understanding of the subject have improved tremendously.

Irizarry, R. A. (2022). *Introduction to Data Science: Data Analysis and Prediction Algorithms with R (Part I)*. Self-published. <https://rafalab.dfci.harvard.edu/dsbook-part-1/>

Irizarry, R. A. (2022). *Introduction to Data Science: Data Analysis and Prediction Algorithms with R (Part II)*. Self-published. <https://rafalab.dfci.harvard.edu/dsbook-part-2/>

I would like also to particularly thank the authors of the learningtower dataset:

Wang K, Yacobellis P, Siregar E, Romanes S, Fitter K, Dalla Riva G, Cook D, Tierney N, Dingorkar P, Sai Subramanian S, Chen G (2024). *learningtower: OECD PISA Datasets from 2000-2022 in an Easy-to-Use Format*. R package version 1.1.0, <https://github.com/kevinwang09/learningtower>, <https://kevinwang09.github.io/learningtower/>.

Throughout this project, I have aimed to provide clear visualizations that facilitate understanding of the data and help identify issues and biases when they arise. I also made an effort to offer step-by-step explanations in R to clarify my approach.

Additionally, I would like to mention that I used ChatGPT to assist with some of the graphs and to check the spelling and the syntaxe of my report. Even with this support, it took me a considerable time to achieve the results for the graph that I wanted, both bug-free and with clear readability. This was especially true for the visualizations in the sections:

2.2.4 PISA Performance in Lower-Income Countries: The Socioeconomic Achievement Gap Across Time

2.2.7 Educational investment and educational outcomes for Lower-middle Income

2.2.21 UNICEF Gender Parity Index vs PISA Gender Gap Analysis

Other references that I found useful and helped me to complete this project:

1. **World Bank Open Data:** The World Bank Open Data platform provides free and comprehensive access to global development data, covering economic, social, demographic, and environmental indicators for virtually every country worldwide. It includes statistics on GDP, poverty, education, health, and more, enabling users to analyze trends, compare countries, and inform policy decisions. The database is widely used by researchers, policymakers, and organizations for cross-national and time-series analysis in fields such as education and socioeconomic development.
2. **UNICEF European Union:** The UNICEF European Union website showcases UNICEF's work in partnership with the EU to advance children's rights and wellbeing in Europe and globally. The site features reports, policy briefs, and data on key topics such as child protection, education, health, and gender equality. It also highlights UNICEF's contributions to evidence-based policymaking, including data and indicators relevant to school enrollment, gender parity, and educational equity—often used in comparative studies and international assessments.
3. **OECD:** The Organisation for Economic Co-operation and Development (OECD) is a leading international organization that provides data, analysis, and policy advice to promote economic and social well-being. Its Education Directorate is responsible for initiatives like PISA (Programme for International Student Assessment), producing influential reports on educational outcomes, equity, and system performance. The OECD's website hosts a wealth of resources, including statistics, country reports, and best practice policy recommendations relevant to education, labor markets, and public policy.
4. **ILO Country Groupings, Concepts and Definitions:** The International Labour Organization (ILO) provides official definitions, classifications, and groupings of countries and regions for use in labor statistics and international reporting. This resource explains how countries are grouped by income, region, or development status according to standardized methodologies. Such classifications ensure data comparability in international labor and social statistics, and are often referenced when segmenting countries in global education and labor market studies.

5. **“Equity and Quality in Education: Supporting Disadvantaged Students and Schools”**: By: OECD (2012) A report focused on reducing inequalities, directly using PISA data and recommendations.
6. **“The Global Testing Culture: Shaping Education Policy, Perceptions, and Practice”**: Edited by: William C. Smith & Alexander W. Wiseman (2016) Explores the effects of global assessments (including PISA) on cultures of testing and learning worldwide.
7. **“Assessing Educational Equity: National and International Perspectives”**: Edited by: Estelle Herbault
8. **“Equity of the European Educational Systems. A Set of Indicators”**: GERES by European Educational Research Journal
9. **“Why Data Collection Matters”** from the book “In Pursuit of Equity in Education”: Gary Orfield, Harvard University published by Springer Dordrecht
10. **“A Theory of Justice”**: Edited by: John Rawls, edited by Oxford University Press, 1971