

# Global Education Analysis: Secondary School Age by Region

**Author:** Stefan Voeten

**Date:** May 2025

**Tools used:** Google BigQuery

**Dataset source:** World Bank via Google Cloud Public Datasets

---

## Project Goal

The aim of this project was to determine how many people around the world were of official age for secondary education in 2015, broken down by global region. This insight can be useful for educational policy, regional planning, or global demographic research.

---

## Research Question

*In 2015, how many people were of the official age for secondary education, grouped by region of the world?*

---

## Data Sources Used

- `bigquery-public-data.world_bank_intl_education.international_education`
- `bigquery-public-data.world_bank_intl_education.country_summary`

These tables contain education indicators per country and country metadata, including region assignments.

---

## SQL Query Used

```
1 SELECT
2     summary.region,
3     SUM(edu.value) AS secondary_edu_population
4 FROM
5     `bigquery-public-data.world_bank_intl_education.international_education` AS edu
6 JOIN
7     `bigquery-public-data.world_bank_intl_education.country_summary` AS summary
8 ON
9     edu.country_code = summary.country_code
10 WHERE
11     summary.region IS NOT NULL
12     AND edu.indicator_name = 'Population of the official age for secondary education, both sexes (number)'
13     AND edu.year = 2015
14 GROUP BY
15     summary.region
16 ORDER BY
17     secondary_edu_population DESC
```

## Query Breakdown

- **JOIN:** Combines education data with country metadata based on country code.
- **WHERE:** Filters only rows for the year 2015, correct indicator, and with non-null region.
- **GROUP BY:** Aggregates values by region.
- **ORDER BY:** Sorts the output by population in descending order.

## Final Output (Results)

### Query results

Job information		Results	Chart	JSON
Row	region ▾	secondary_edu_popu		
1	South Asia	237541684.0		
2	East Asia & Pacific	172016129.0		
3	Sub-Saharan Africa	135639085.0		
4	Europe & Central Asia	70181959.0		
5	Latin America & Caribbean	67937467.0		
6	Middle East & North Africa	44318682.0		
7	North America	27003321.0		

(Source: World Bank data in BigQuery)

## Key Learnings

- I practiced writing multi-table SQL queries using JOIN, WHERE, GROUP BY, and ORDER BY.
  - I learned how to filter based on specific indicators and years within large datasets.
  - I gained experience translating a real-world question into a structured SQL solution.
  - I recognized how clean, grouped results require understanding how rows should aggregate (per region vs. per country).
-