QTM 350 Final Project – School Enrollment Trends Analysis

Da Eun Kim (2483332) Alix Morales (2594298)
Katie Park (2563640) Yoonsuh Park (2499926)

2024-12-09

Table of contents

Introduction

Education is a cornerstone of sustainable development, playing a pivotal role in shaping economic and social outcomes. School enrollment rates, which measure access to education at primary, secondary, and tertiary levels, serve as critical indicators of a country's progress toward achieving universal education. This project explores global patterns and disparities in school enrollment using data from the World Bank's World Development Indicators (WDI)

database. By examining trends over time, identifying leading and lagging countries, and analyzing regional differences, this study sheds light on the successes and persistent challenges in global education systems.

Focusing on three key indicators—primary enrollment (% gross), secondary enrollment (% gross), and tertiary enrollment (% gross)—this analysis investigates the relationships between education levels and their evolution over time. Additionally, it evaluates how regions and countries vary in their educational performance, highlighting areas for policy intervention. This study aims to provide actionable insights into the state of global education, contributing to the broader discourse on achieving equitable and inclusive access to quality education worldwide.

Data Description

This dataset is a long-format representation of World Development Indicators (WDI), capturing socio-economic and developmental statistics across multiple countries and regions. It includes 664 entries spanning a range of indicators, such as school enrollment rates at different educational levels. The data consists of seven columns: `country_name`, `country_code`, `series_name` (indicator name), `series_code` (indicator code), `year`, `value` (observed statistic), and `region` (geographical classification). Each row corresponds to a specific indicator for a country in a given year, making it suitable for time-series analysis and cross-country comparisons. The dataset provides insights into regional and global trends, supporting research on education, economic development, and social progress.

Primary Enrollment (% gross): Indicator code SE.PRM.ENRR. Secondary Enrollment (% gross): Indicator code SE.SEC.ENRR. Tertiary Enrollment (% gross): Indicator code SE.TER.ENRR. Countries and Regions: Data is available for countries worldwide, categorized into regions such as “South Asia,” “Sub-Saharan Africa,” etc. Years: The dataset includes time-series data, enabling the analysis of enrollment trends over time. Variables: The dataset contains numerical values for enrollment rates (value) along with metadata such as country name, region, and year. The dataset has been cleaned for consistency: no missing values were observed in key columns, and all columns have appropriate data types.

Data Analysis

```
import pandas as pd
wdi_data = pd.read_csv("wdi_data_with_region.csv")
```

Global Trends in School Enrollment Over Time

This analysis examines the global trends in gross enrollment rates across primary, secondary, and tertiary education levels over time. By grouping data by year and indicator, we calculate global averages and plot these trends.

```
#1. Global Trends in School Enrollment Over Time
import matplotlib.pyplot as plt

# Extract relevant indicators for analysis

relevant_indicators = ['SE.PRM.ENRR', 'SE.SEC.ENRR', 'SE.TER.ENRR']

filtered_data = wdi_data[wdi_data['series_code'].isin(relevant_indicators)]

# Group by year and calculate global averages for each indicator

trend_data = filtered_data.groupby(['year', 'series_name'])['value'].mean().reset_index()

# Plot trends over time

plt.figure(figsize=(12, 6))

for indicator in trend_data['series_name'].unique():
    indicator_data = trend_data[trend_data['series_name'] == indicator]
    plt.plot(indicator_data['year'], indicator_data['value'], label=indicator)

plt.title('Global Trends in School Enrollment Over Time', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Enrollment Rate (% gross)', fontsize=12)
plt.legend(title='Education Level')
plt.grid(True)
plt.show()
```

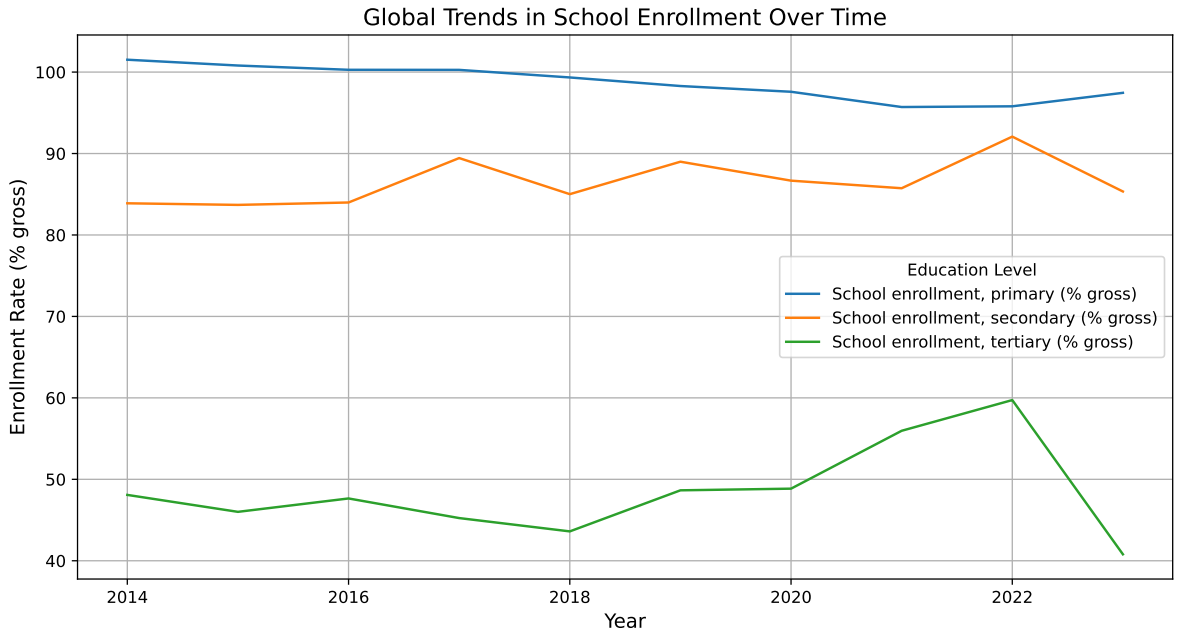

Figure 1: Trends in global school enrollment rates over time, showing gross enrollment rates for primary, secondary, and tertiary education. Primary enrollment rates are consistently high, while secondary and tertiary levels show gradual increases, reflecting progress in global education access.

Regional Comparison of Mean Enrollment Rates by Indicator

This analysis compares average enrollment rates by education level (primary, secondary, tertiary) across different regions. By calculating mean rates per region, the bar chart highlights disparities in education access.

```
#2. Regional Comparison of Mean Enrollment Rates by Indicator

indicator_region_averages = wdi_data.groupby(['series_name', 'region'])['value'].mean().unsta
indicator_region_averages.plot(kind='bar', figsize=(12, 8))
plt.title("Average Enrollment Rates by Indicator and Region", fontsize=14)
plt.xlabel("Education Indicator", fontsize=12)
plt.ylabel("Mean Enrollment Rate (%)", fontsize=12)
plt.xticks(rotation=45, ha='right')
plt.legend(title='Region', bbox_to_anchor=(1.05, 1), loc='upper left')
plt.tight_layout()
plt.show()
```

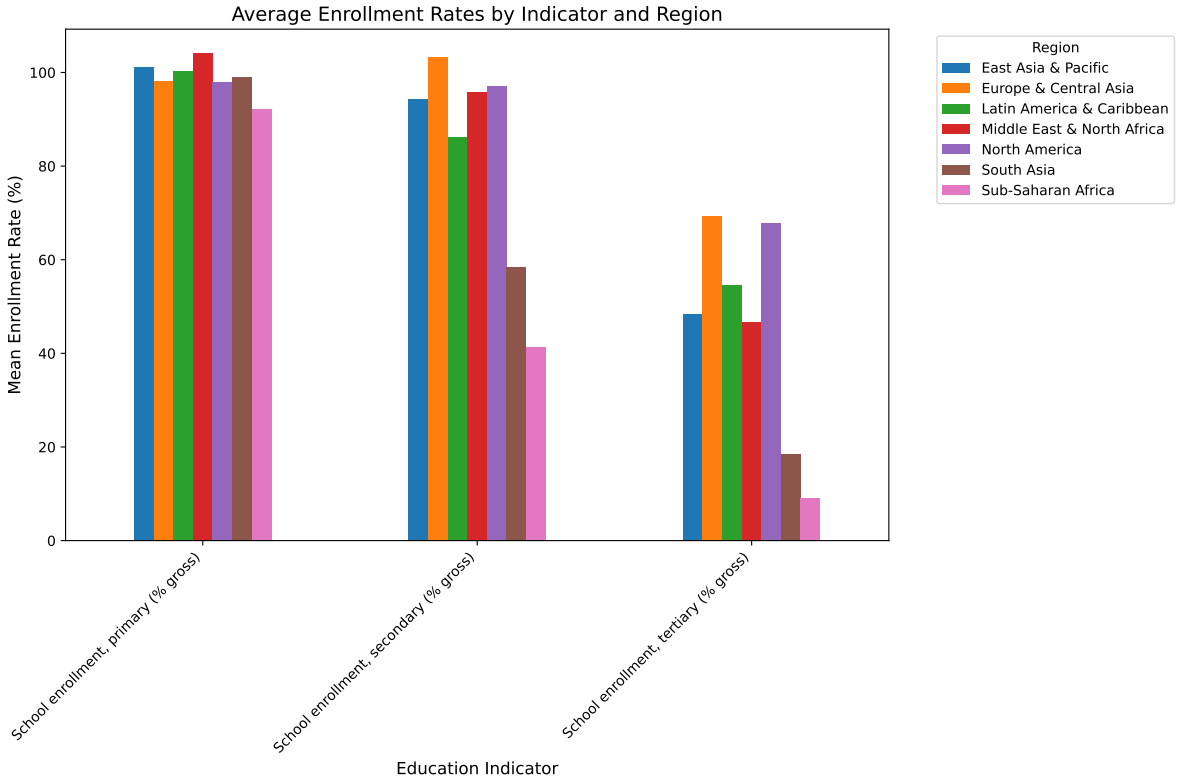

Figure 2: Average school enrollment rates by region and education level, illustrating significant disparities. Developed regions exhibit higher enrollment rates across all levels, while Sub-Saharan Africa and South Asia lag behind, especially in tertiary education.

Correlation Between Primary, Secondary, and Tertiary Enrollment

This analysis explores the relationship between enrollment rates at primary, secondary, and tertiary levels. A heatmap visualizes Pearson correlation coefficients.

```
#3. Correlation Between Primary, Secondary and Tertiary Indictors

# Filter data to include only relevant indicators
filtered_data = wdi_data[wdi_data['series_name'].isin([
    'School enrollment, primary (% gross)',
    'School enrollment, secondary (% gross)',
    'School enrollment, tertiary (% gross)'
])]

# Pivot data to have indicators as columns
```

```
pivoted_data = filtered_data.pivot_table(
    index=['country_name', 'year'],
    columns='series_name',
    values='value'
).dropna()

# Rename columns for easier reference
pivoted_data.columns = ['Primary Enrollment', 'Secondary Enrollment', 'Tertiary Enrollment']

# Calculate Pearson correlation coefficients
correlation_matrix = pivoted_data.corr()

import seaborn as sns

# Plotting the correlation heatmap
plt.figure(figsize=(8, 6))
sns.heatmap(correlation_matrix, annot=True, fmt=".2f", cmap="coolwarm", cbar=True)
plt.title("Correlation Heatmap: Primary, Secondary, and Tertiary Enrollment")
plt.show()
```

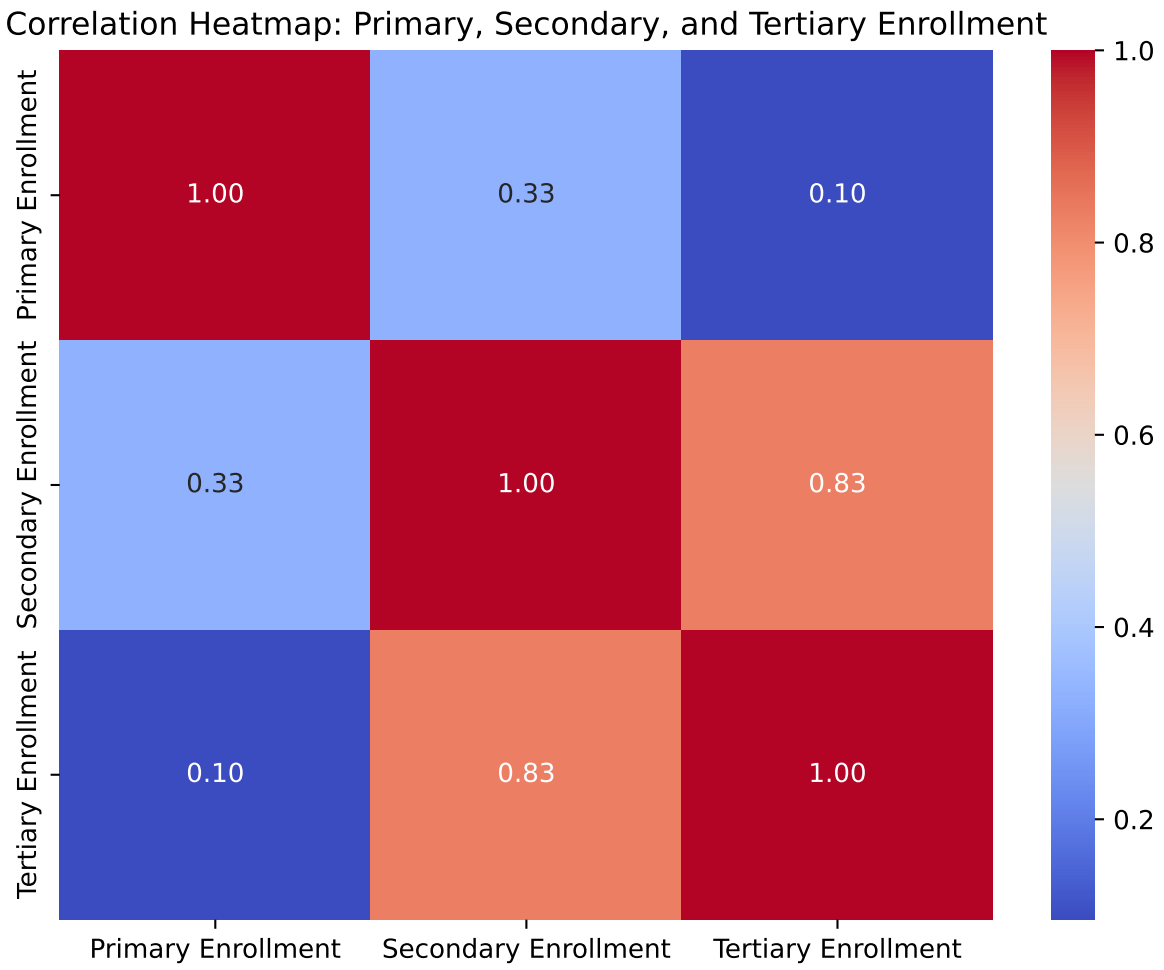

Figure 3: Correlation heatmap of primary, secondary, and tertiary school enrollment rates. The strong correlations indicate a cascading effect, where higher primary enrollment positively influences subsequent levels of education.

Results and Discussion

Global Trends in School Enrollment Over Time

- Primary Enrollment: Typically the highest across all years, reflecting widespread prioritization of basic education globally.
- Secondary Enrollment: Generally lower than primary but shows an upward trend, indicating improvements in access to education beyond primary school.

- Tertiary Enrollment: The lowest but with significant growth over the years, suggesting increasing opportunities for higher education globally.

These trends highlight global efforts to enhance education access, with significant progress at all levels, particularly in tertiary education.

Regional Comparison of Mean Enrollment Rates by Indicator

- Developed Regions: Tend to have high enrollment rates across all levels, indicating well-established education systems.
- Developing Regions: Display gaps, especially in secondary and tertiary levels, pointing to challenges like economic barriers and lack of infrastructure.
- Notable Gaps: Sub-Saharan Africa and South Asia show significantly lower tertiary enrollment rates, emphasizing regional inequalities.

Such comparisons underline the need for targeted policies to address disparities and improve access in underperforming regions.

Correlation Between Primary, Secondary, and Tertiary Enrollment

- Primary vs. Secondary: Strong positive correlation, reflecting that countries with high primary enrollment are likely to achieve better secondary enrollment.
- Secondary vs. Tertiary: Moderate to strong correlation, highlighting that success at secondary education often translates into higher tertiary participation.
- Primary vs. Tertiary: Weak to moderate correlation, suggesting that barriers between these levels may be distinct.

These findings emphasize the interconnectedness of educational levels, where improvements in one level positively influence others, though with diminishing returns as the education level increases.

Conclusion

In conclusion, the analysis of global school enrollment trends provides valuable insights into the progress and challenges in achieving universal education. Primary education exhibits consistently high enrollment rates, reflecting global commitments to foundational education access. However, the disparities in secondary and tertiary enrollment highlight persistent inequalities, particularly in developing regions such as Sub-Saharan Africa and South Asia. These disparities emphasize the need for targeted policy interventions, such as improving education infrastructure, addressing economic barriers, and enhancing teacher training.

Moreover, the positive correlation between primary, secondary, and tertiary enrollment underscores the interconnected nature of education levels. Policies promoting smooth transitions between these stages are essential for long-term educational success. The upward trend in tertiary enrollment is promising, indicating growing opportunities for higher education. However, achieving equitable access remains a challenge, requiring sustained efforts from governments, international organizations, and communities.

By shedding light on these patterns, this study underscores the importance of data-driven approaches to education policymaking. Future research could expand on this analysis by exploring the impact of socioeconomic factors, gender disparities, and digital access on enrollment trends, offering a more holistic view of global education systems.