

WORLD DEVELOPMENT INDICATORS

DATASCI 200
UNIVERSITY OF CALIFORNIA, BERKELEY

SECTION 2, 2PM PT
JORDAN ANDERSEN, COURTNEY CHEN, MAIA KENNEDY

01

BACKGROUND

- About the dataset
- Audience
- Why it matters
- Our research question

02

THE DATA PROCESS

- Data Cleaning
- Variable Selection
- Visualizations
- Challenges & Mitigations

03

RESEARCH FINDINGS

- Answers to our key questions
- Challenges & Mitigations
- Visualizations

04

CLOSING

- Conclusions
- Lessons Learned for the research & data process
- Questions
- Thank You

CONTEXT

About the Data

The World Bank is a global development organization that finds solutions to multifaceted social development issues. The indicators are a repository of key development indicators collected by internationally recognized sources to form one of the most comprehensive and accurate global development databases.

Our Research Question:

Is primary school enrollment linked to labor force participation and unemployment in low-, middle-, and high-income countries across genders?

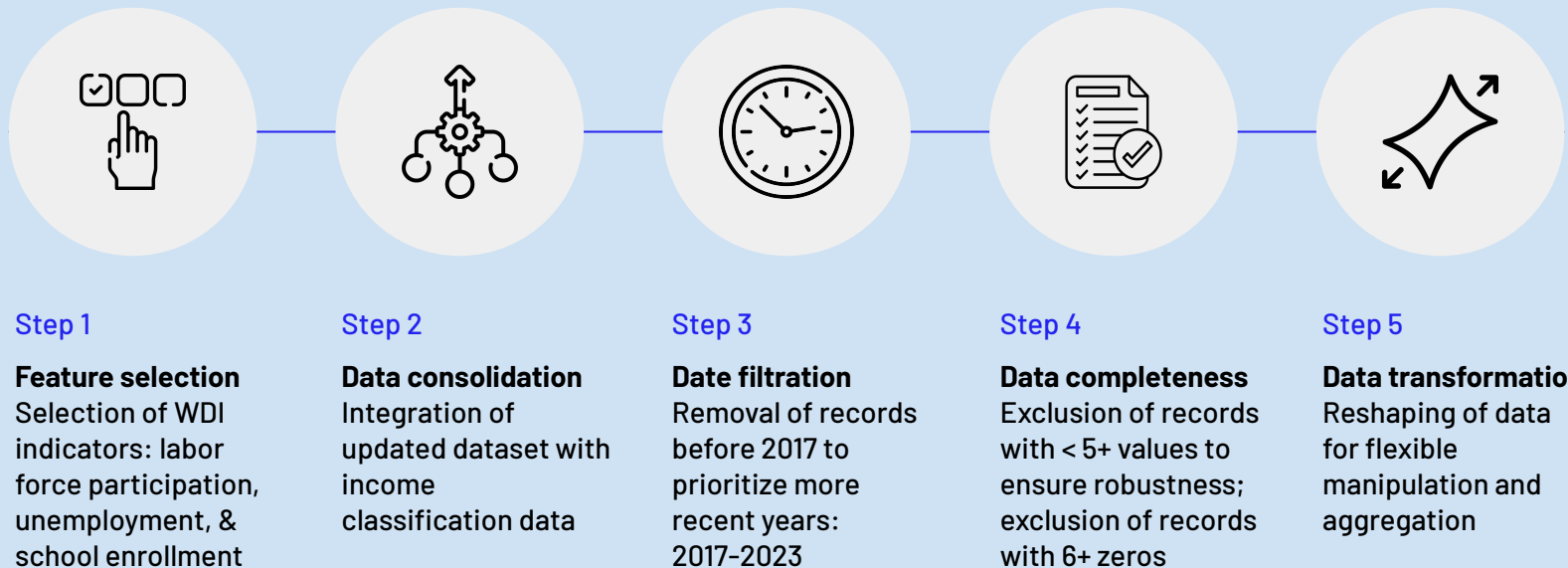
Why it matters:

- Highlighting the Interdependence of Key Development Indicators
- Addressing Gender Disparities Across Income Levels
- Informing Policy and Program Design Through Nuanced Analysis



PREPARING THE DATA

The data process aims to reduce inconsistencies, focus on relevant trends, and create a robust foundation for meaningful insights.



SELECTING THE VARIABLES

Education

1 | School enrollment, primary and secondary, male and female (gross %)

Why: Highest count of countries with school enrollment data

2 | School Enrollment Gender Parity Index (GPI)

Why: Initial observations & research question was more focused on the gap between genders in enrollment

Labor Force & Unemployment

1 | Labor force participation rate, male and female (% of population ages 15+)

Why: Insights into a country's workforce and economic health

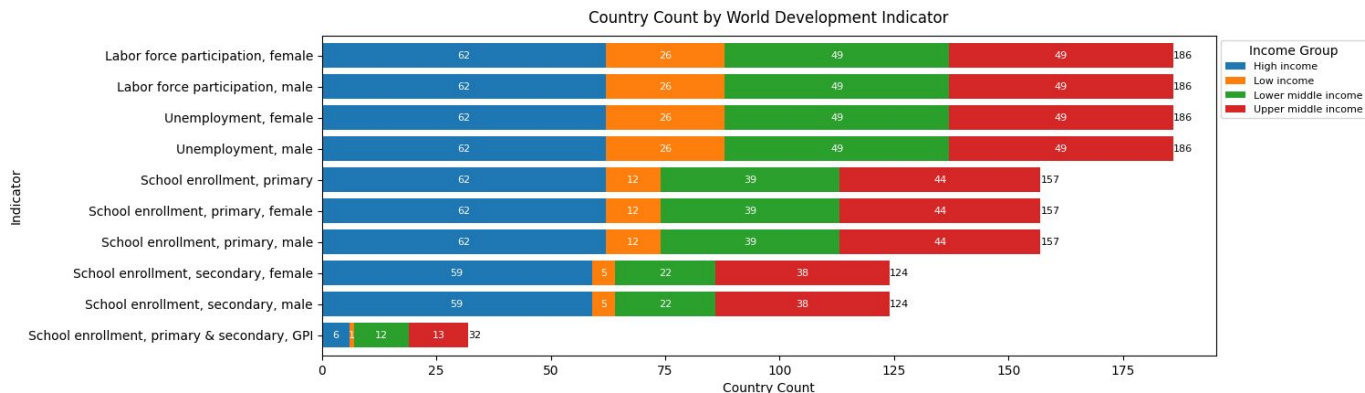
2 | Unemployment, male and female (% of labor force)

Why: To explore the intersection between unemployment and education access

Income Groupings

The World Bank's classification of economies into four income groups—low, lower-middle, upper-middle, and high income—using gross national income (GNI) per capita in USD:

- Low: $\leq \$1,145$
- Lower-middle: $\$1,146 - \$4,515$
- Upper-middle: $\$4,516 - \$14,005$
- High: $> \$14,005$



EDUCATION

Primary & secondary school enrollment¹

% gross, calculated as total enrollments / population of corresponding age group

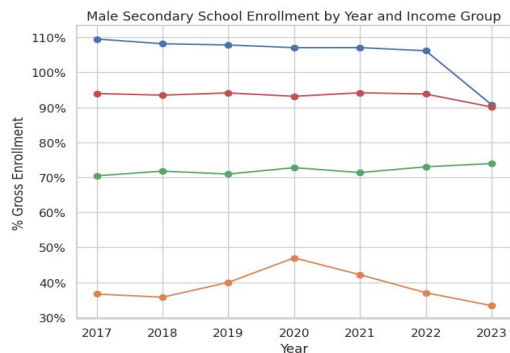
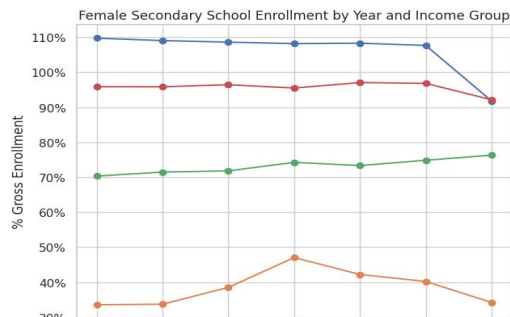
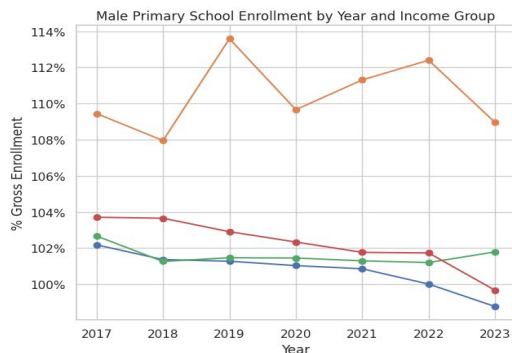
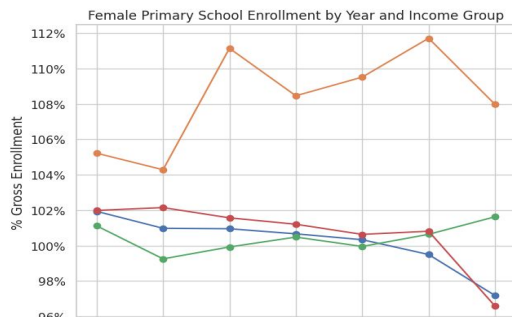
High income
Low Income
Lower Middle Income
Upper Middle Income

Simplified Data Process

Grouped
variables

Extracted
Means

Plotted
Data



Findings

Primary school enrollment is strong in low-income countries, but secondary school rates remain low.

High income countries exhibit the highest secondary school enrollment, followed by upper middle-, lower middle-, and low.

Percentages > 100% is due to the inclusion of over-aged and under-aged students because of early or late school entrance and grade repetition.

Challenges: Data scarcity in school enrollment variables of interest, primarily due to limited reporting and technological access in low-income countries.

¹ Note: Ranges of y-axes differ to highlight patterns, differences, and trends.

LABOR FORCE & UNEMPLOYMENT¹

Labor Force²

Proportion of the population ages 15+ that is economically active who supply labor for the production of goods and services

Unemployment

Unemployment refers to the share of the labor force that is without work but available for and seeking employment.²

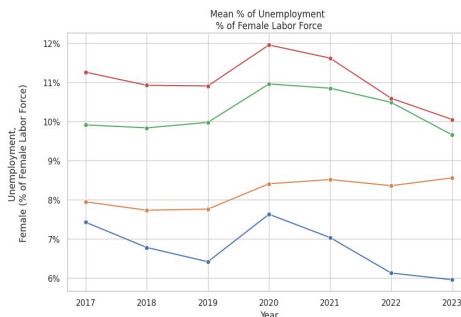
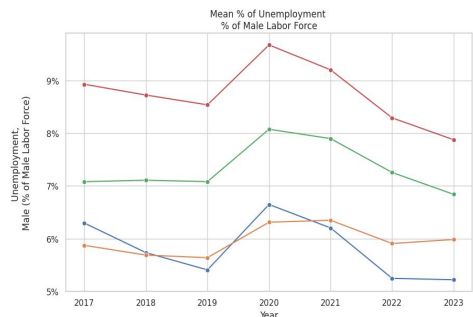
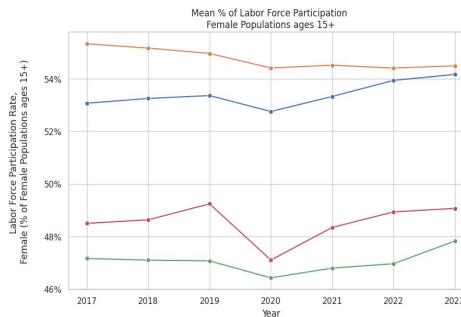
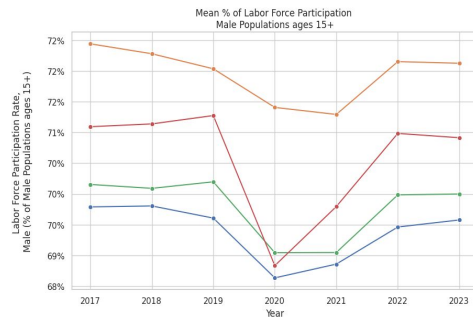
- High income
- Low Income
- Lower Middle Income
- Upper Middle Income

Simplified Data Process

Grouped
variables

Extracted
Means

Plotted
Data



Findings

In all income country groupings, males typically have less unemployment than females

High-income countries have the lowest male labor force compared to other income groups.

Females in high-income countries represent a smaller share of the unemployed, possibly due to the requirement of "actively seeking work," which may exclude those from more privileged backgrounds

1) Note: Ranges of y-axes differ to highlight patterns, differences, and trends.

2) Definitions of labor force and unemployment differ by country.



Indicates same trends for
both male and female groups

OUR RESEARCH CONCLUSIONS

Is primary school enrollment linked to labor force participation and unemployment in low-, middle-, and high-income countries across genders?

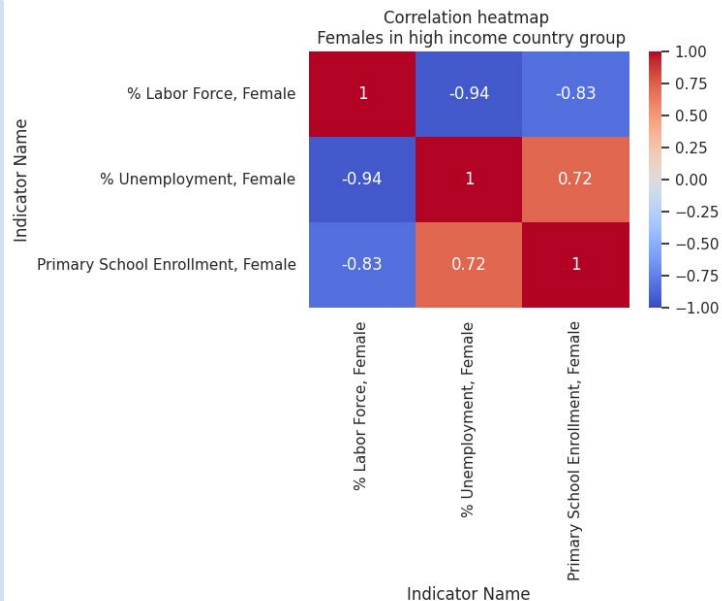
Approach: We selected the relevant variables and conducted a **correlation analysis** between male and female data, measuring the strength of their relationships (from -1 to +1) and taking the means of each of the income countries within each group.

Disclaimer: This analysis only shows **correlation** and is no way making **causal claims on an outcome variable**.

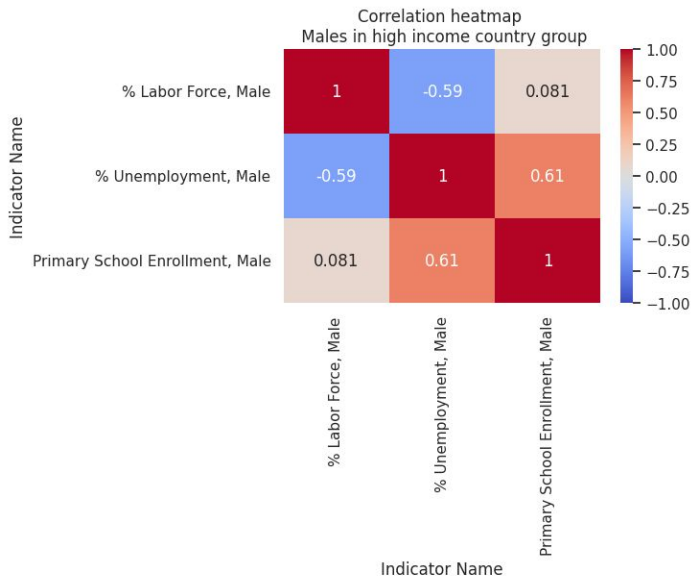


Income Group	Gender	Higher primary school enrollment in s correlated with...
High Income	Female	<ul style="list-style-type: none">• Lower labor force participation• Higher unemployment
	Male	<ul style="list-style-type: none">• Nearly the same labor force participation• Higher unemployment
Low Income	Female	<ul style="list-style-type: none">• Lower labor force participation• Higher unemployment
	Male	<ul style="list-style-type: none">• Lower (minimal) labor force participation• Nearly the same unemployment
Lower-Middle-Income	Female	<ul style="list-style-type: none">• Higher labor force participation• Lower unemployment
	Male	<ul style="list-style-type: none">• Higher labor force participation• Lower unemployment
Upper-Middle-Income	Female	<ul style="list-style-type: none">• Lower labor force participation• Higher unemployment
	Male	<ul style="list-style-type: none">• Higher labor force participation• Higher unemployment

Correlation Heatmap for Females in High Income Groups



Correlation Heatmap for Males in High Income Groups



OUR RESEARCH CONCLUSIONS

Conclusion: The relationship between primary school enrollment and labor force participation varies by income level, with higher enrollment linked to higher unemployment for upper-middle and high income groups for both males and females. The relationship between enrollment and labor force participation is negative for most female groups and neutral to positive for most male groups. School enrollment generally has a stronger correlation with labor force participation for females than males, regardless of direction. These trends highlight the complex interaction between education, labor markets, and economic development across different income groups.

On data cleaning

- Handling empty data through filtering and dropping
- Reshaped data to simplify trend analysis over time
- Ensured timeliness and data validity through date filtration

On variable selection

- Observed data availability by examining value counts upfront
- Considered practical applications of variable examination

On defining our research

- Think big first, then scope down
- Refine research question after selecting variables

On visualizations

- Consider readability for the audience
 - Simplify axes and titles
 - Standardize colors to reinforce consistency
-

QUESTIONS

SECTION 2, 2PM PT:
JORDAN ANDERSEN, COURTNEY CHEN, MAIA KENNEDY