

# ANALYZING CHANGE IN GENDER EQUALITY WORLDWIDE

Presented by Team 3B

Yulong Gong, Peter Mankiewich,  
Ruchika Venkateswaran, Phoenix Wang, Yangyang Zhou

# AGENDA

- Project Overview
- About the Dataset
- Gender Equality in the Education System
- Gender Equality in the Labor Market
- Exploring Relationships Between Education and Employment by Gender
- Conclusion
- Challenges and Future Work

# PROJECT OVERVIEW

## ■ Objective

- Our goal is to analyze the fluctuations in gender equality in different countries between 2000 to 2015 across education and employment indicators in high, upper middle, lower middle and low income segment groups.

## ■ Data Source

- The raw dataset is taken from The World Bank and can be found at  
<https://datacatalog.worldbank.org/dataset/world-development-indicators>

# ABOUT THE DATASET

## ■ Overview

- The dataset contains over **5.65 million** rows with **1344** annual indicators of development from **214** countries around the world dated between 1960 to 2015.

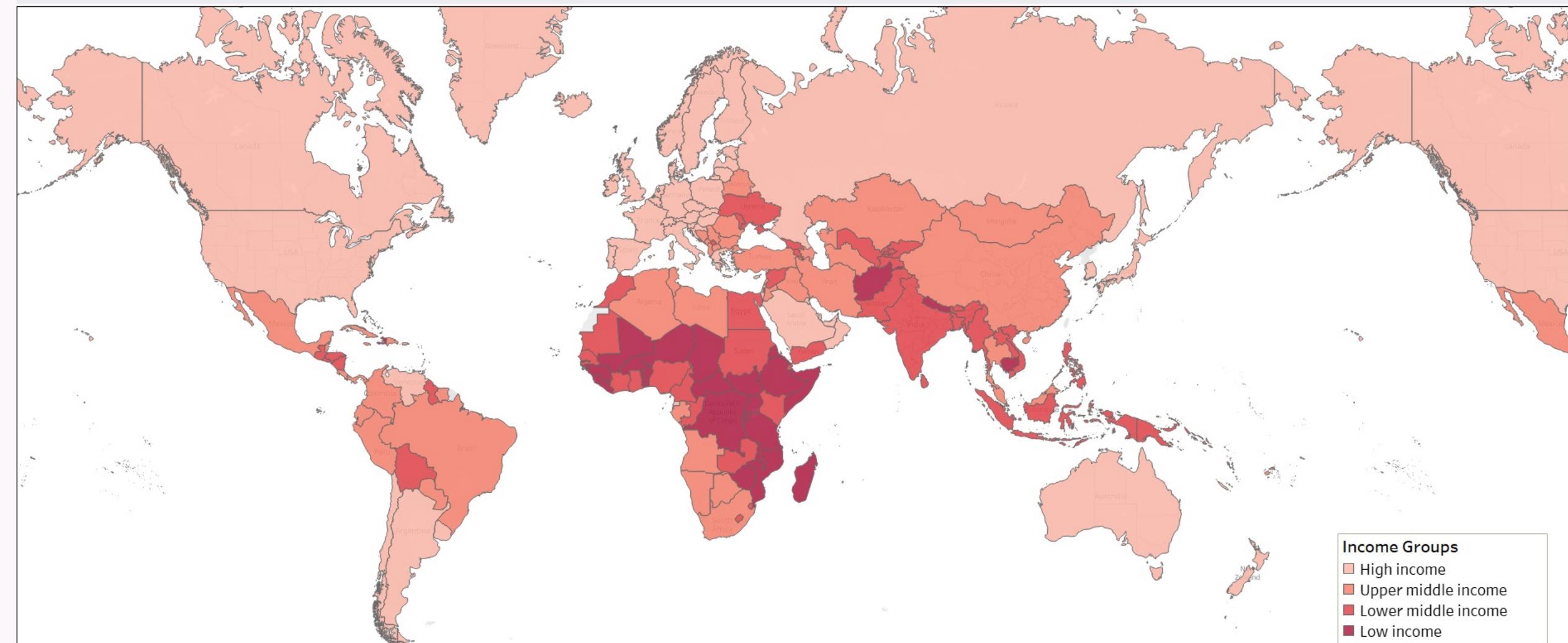
## ■ Indicators

- The indicators are categorized into five groups:  
**① Economic    ② Education    ③ Employment    ④ Health    ⑤ Life Expectancy**

## ■ Countries

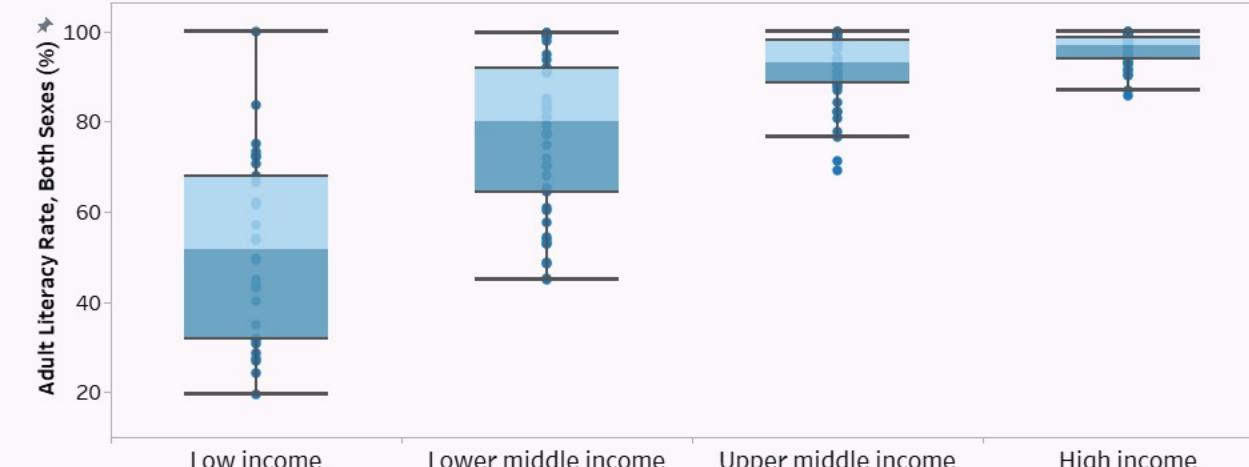
- The countries are classified according to their GNI per capita:  
**① High Income    ② Upper Middle Income    ③ Lower Middle Income    ④ Low Income**

## GNI PER CAPITA BY INCOME GROUP

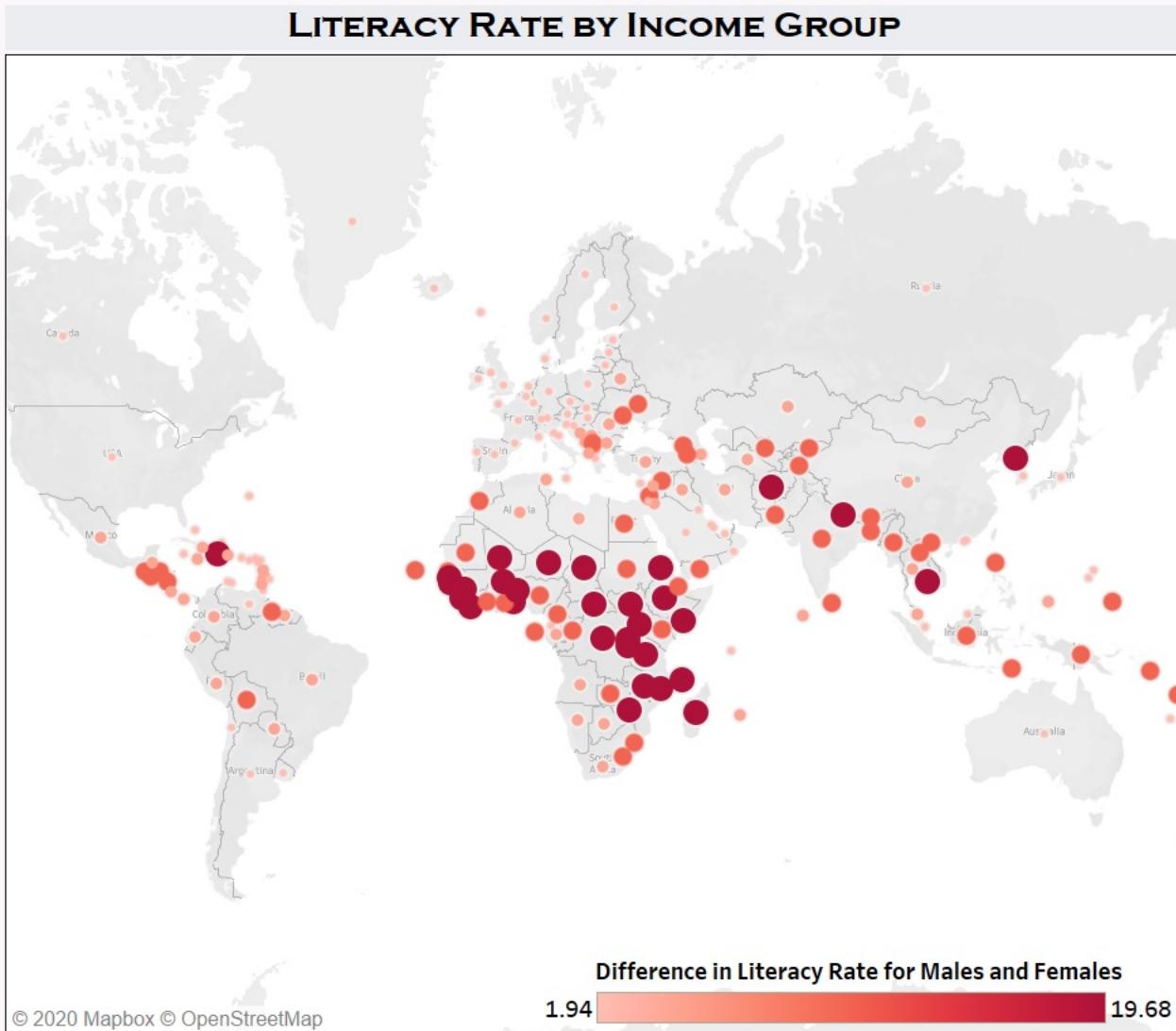


# ANALYTICAL REPORT ON EDUCATION

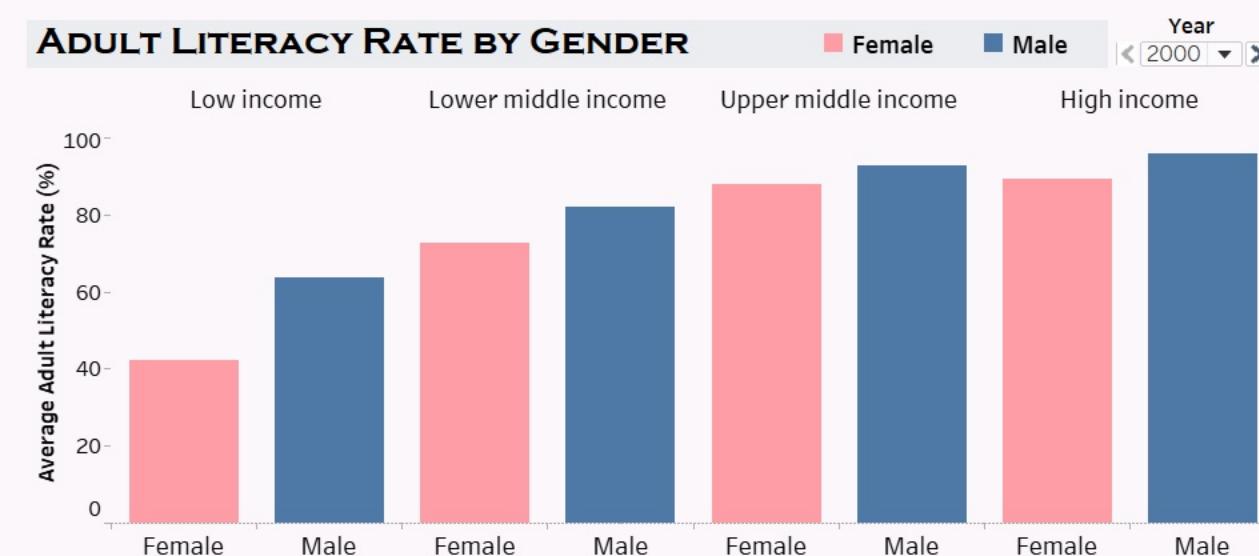
## ADULT LITERACY RATE DISTRIBUTUON BY INCOME GROUP



## LITERACY RATE BY INCOME GROUP

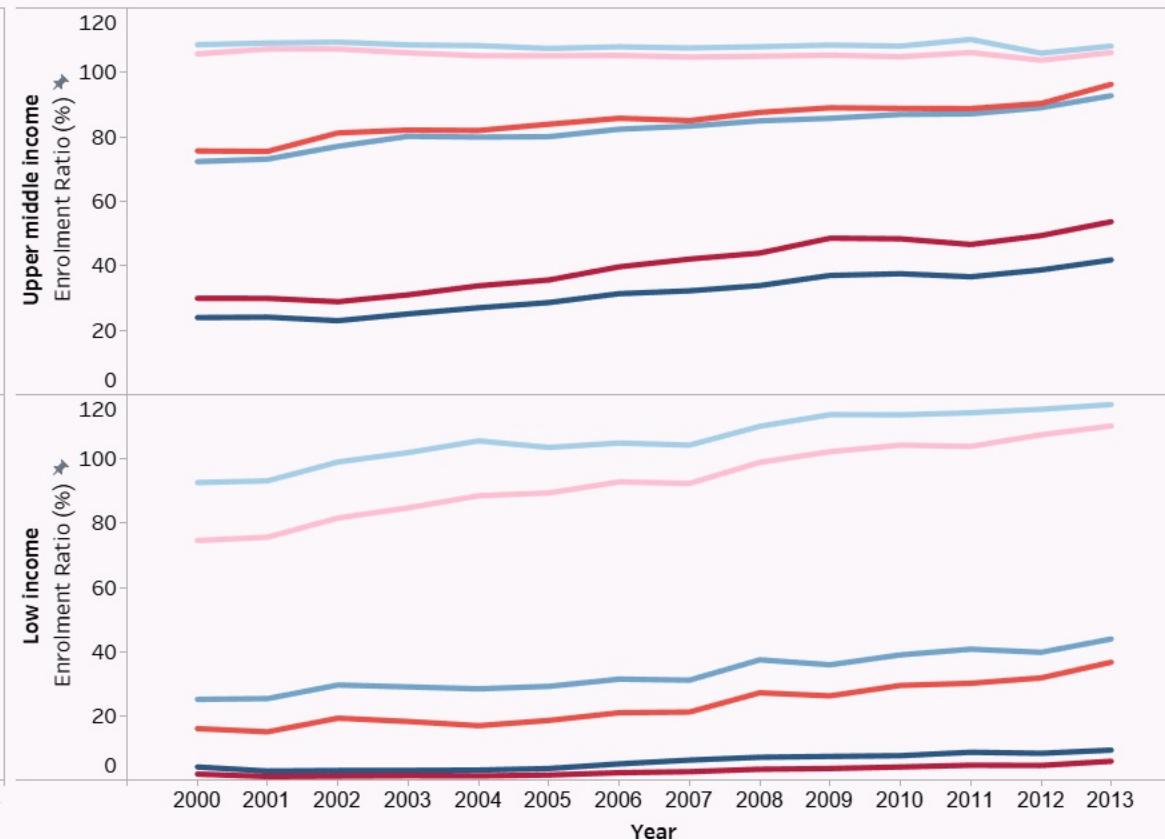
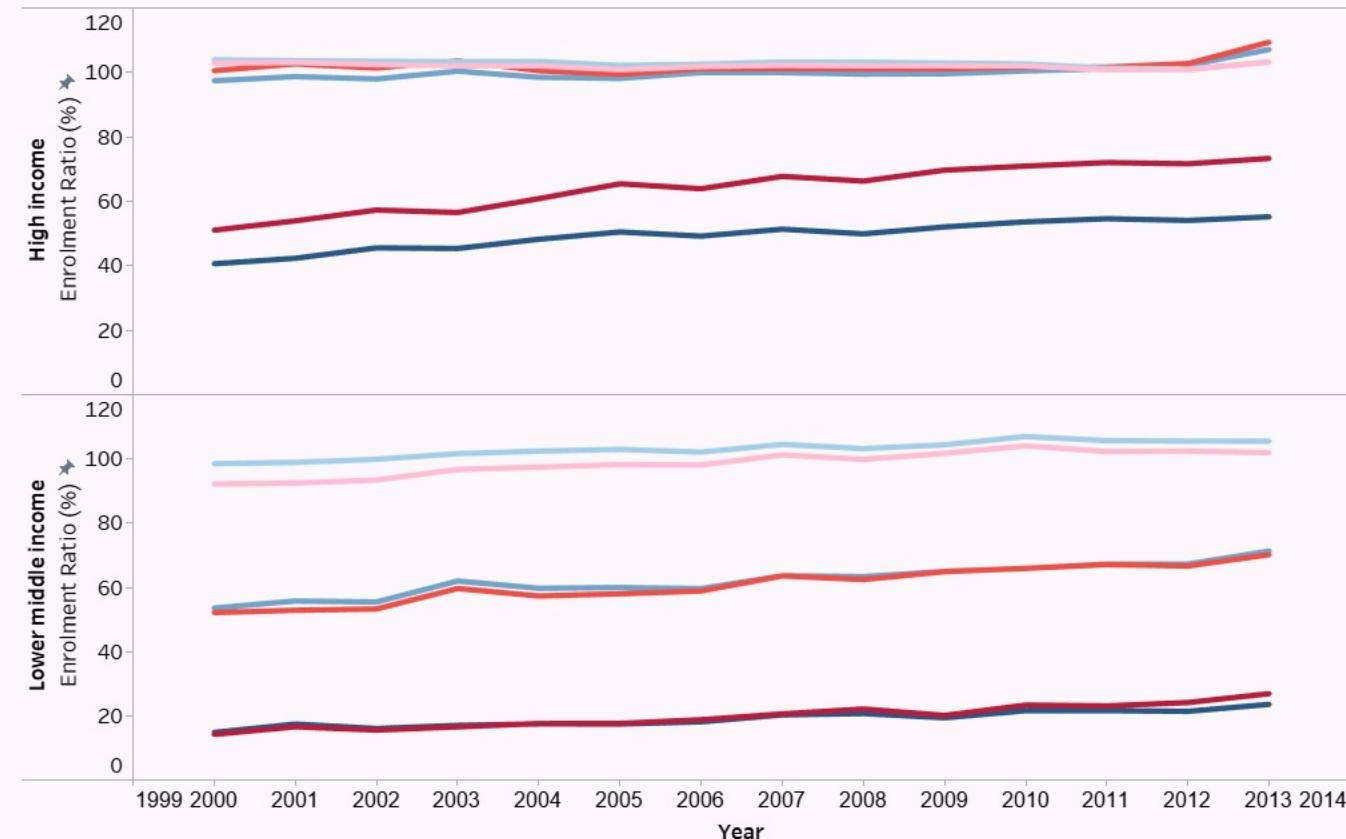


## ADULT LITERACY RATE BY GENDER



# ANALYTICAL REPORT ON EDUCATION

## GROSS ENROLMENT RATIO BY INCOME GROUP AND EDUCATION LEVEL (PRIMARY, SECONDARY, TERTIARY)



The gross enrollment ratio is the ratio of the number of students enrolled in a particular level of education over the number of students eligible for that education level. This measure may be over 100% if students are enrolled in an education level for which they may not formally qualify. For example, students who are repeating a grade or who enrolled late.

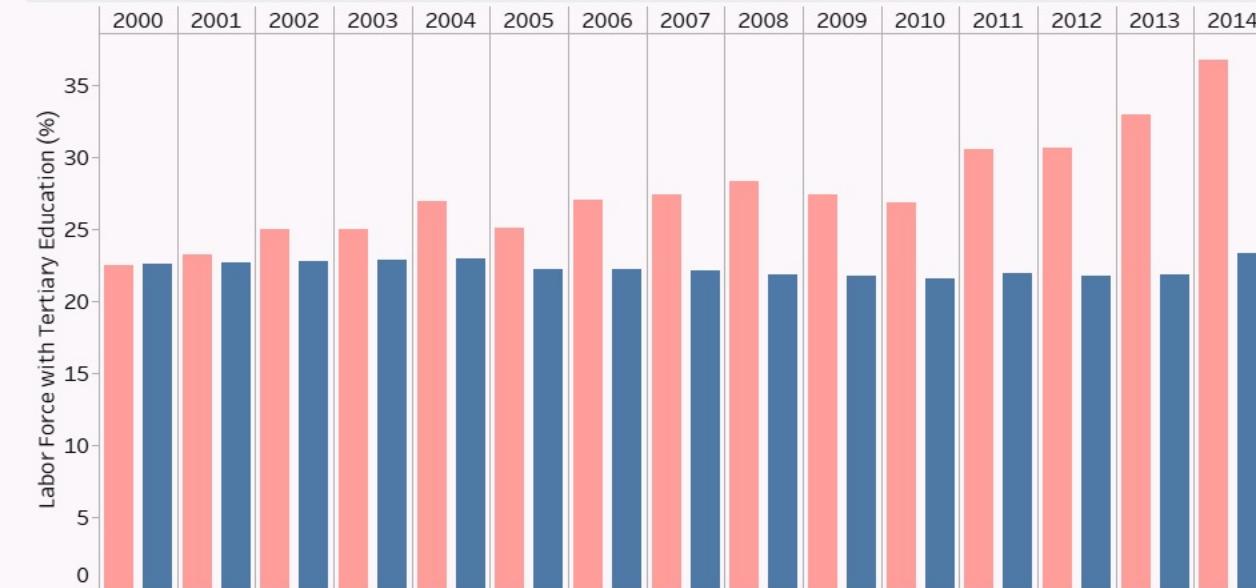
- Primary, Female
- Primary, Male
- Secondary, Female
- Secondary, Male
- Tertiary, Female
- Tertiary, Male

# ANALYTICAL REPORT ON EMPLOYMENT

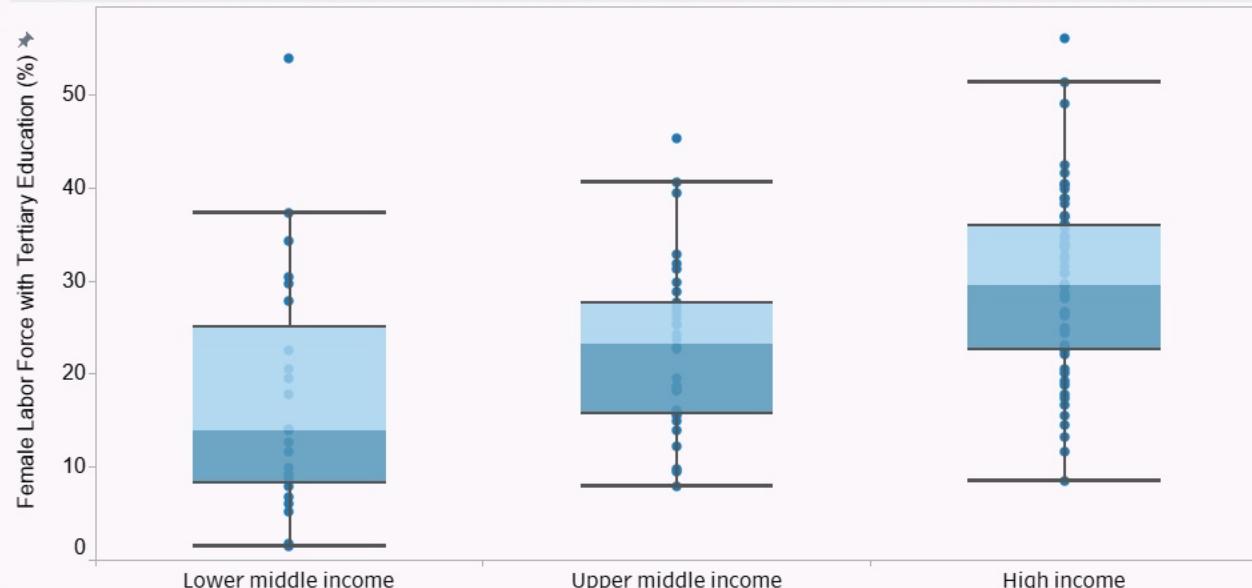
## CHILD LABOR RATE (7-14 YEARS OLD) OVER TIME



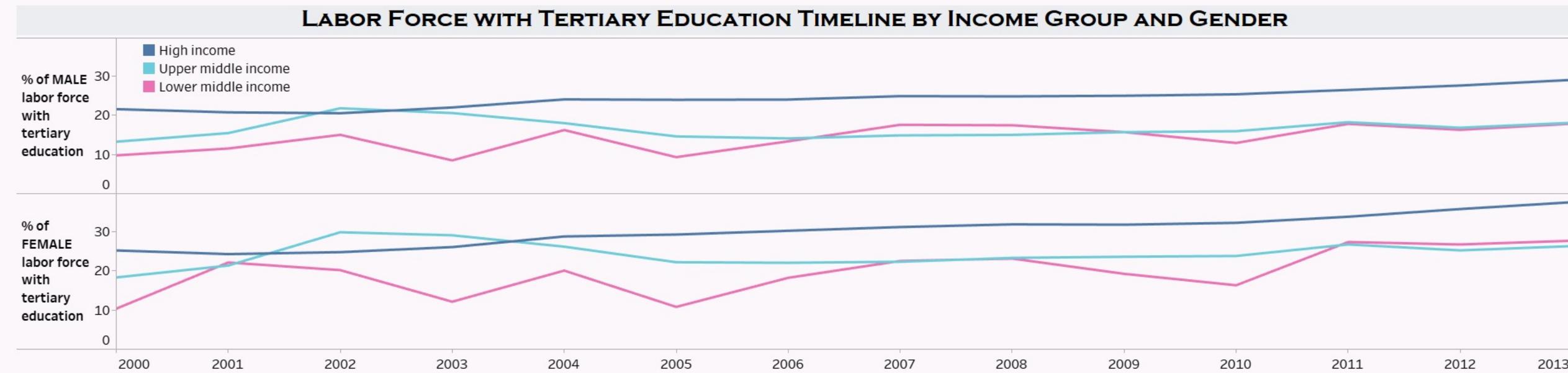
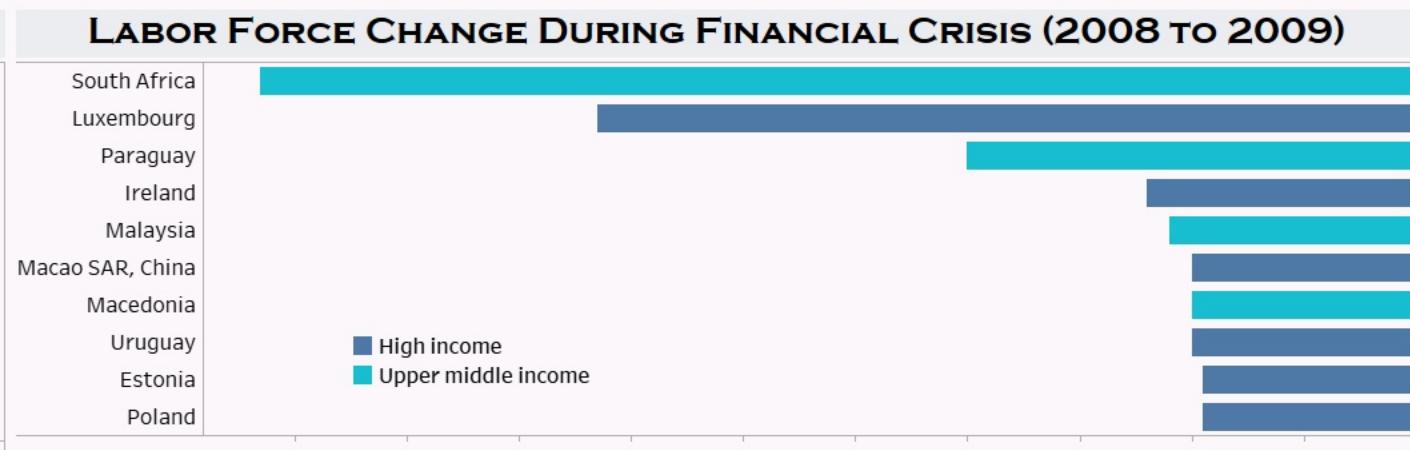
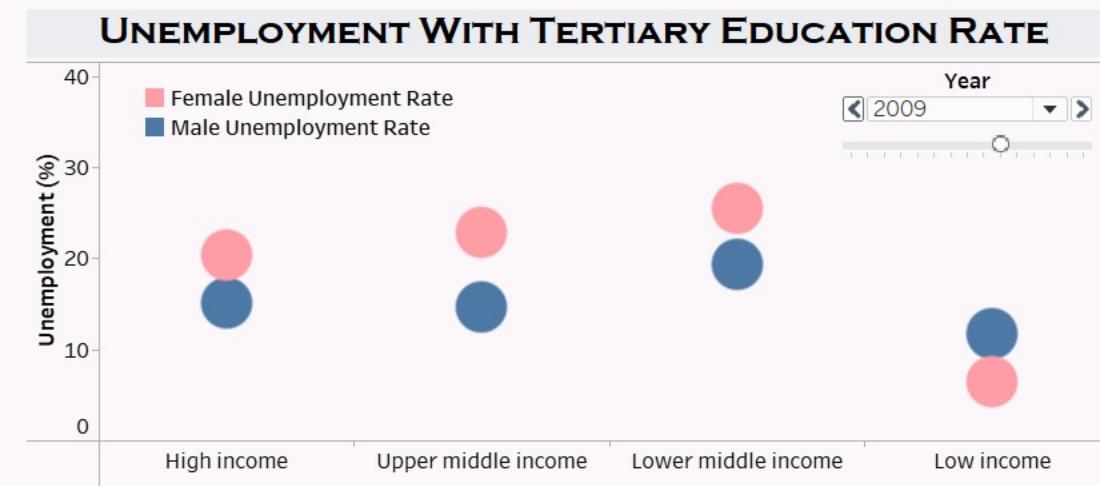
## LABOR FORCE WITH TERTIARY EDUCATION TIMELINE BY GENDER



## DISTRIBUTION OF FEMALE LABOR FORCE WITH TERTIARY EDUCATION

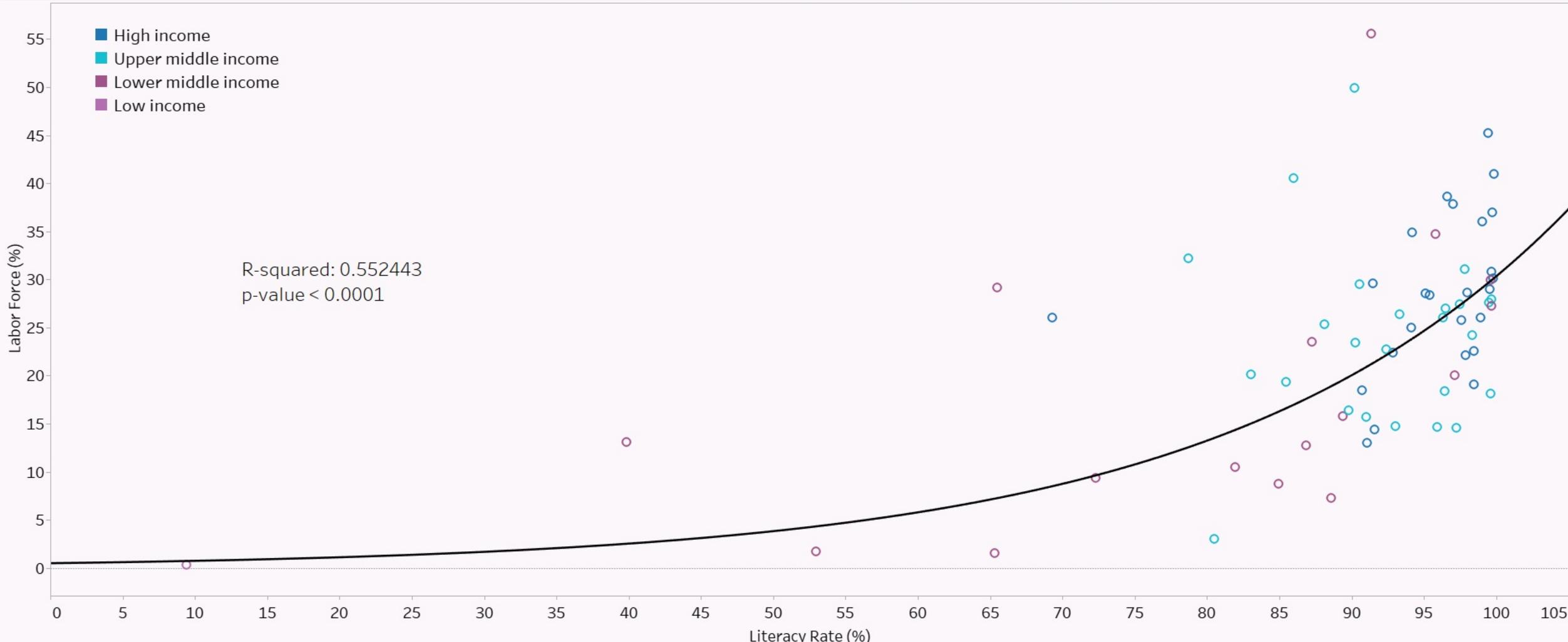


# ANALYTICAL REPORT ON EMPLOYMENT



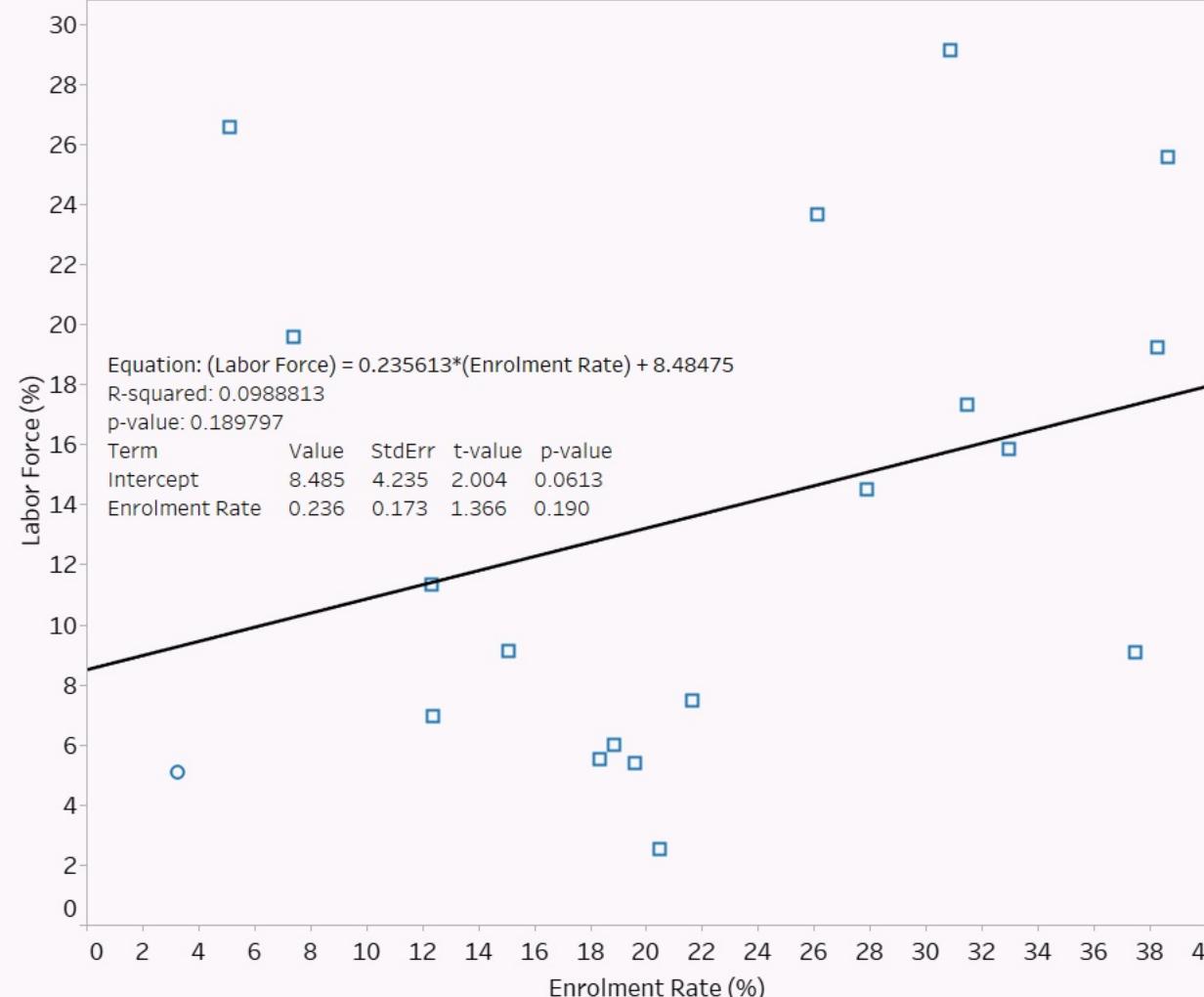
# RELATIONSHIP BETWEEN EDUCATION AND EMPLOYMENT

## RELATIONSHIP BETWEEN LITERACY RATE AND LABOR FORCE

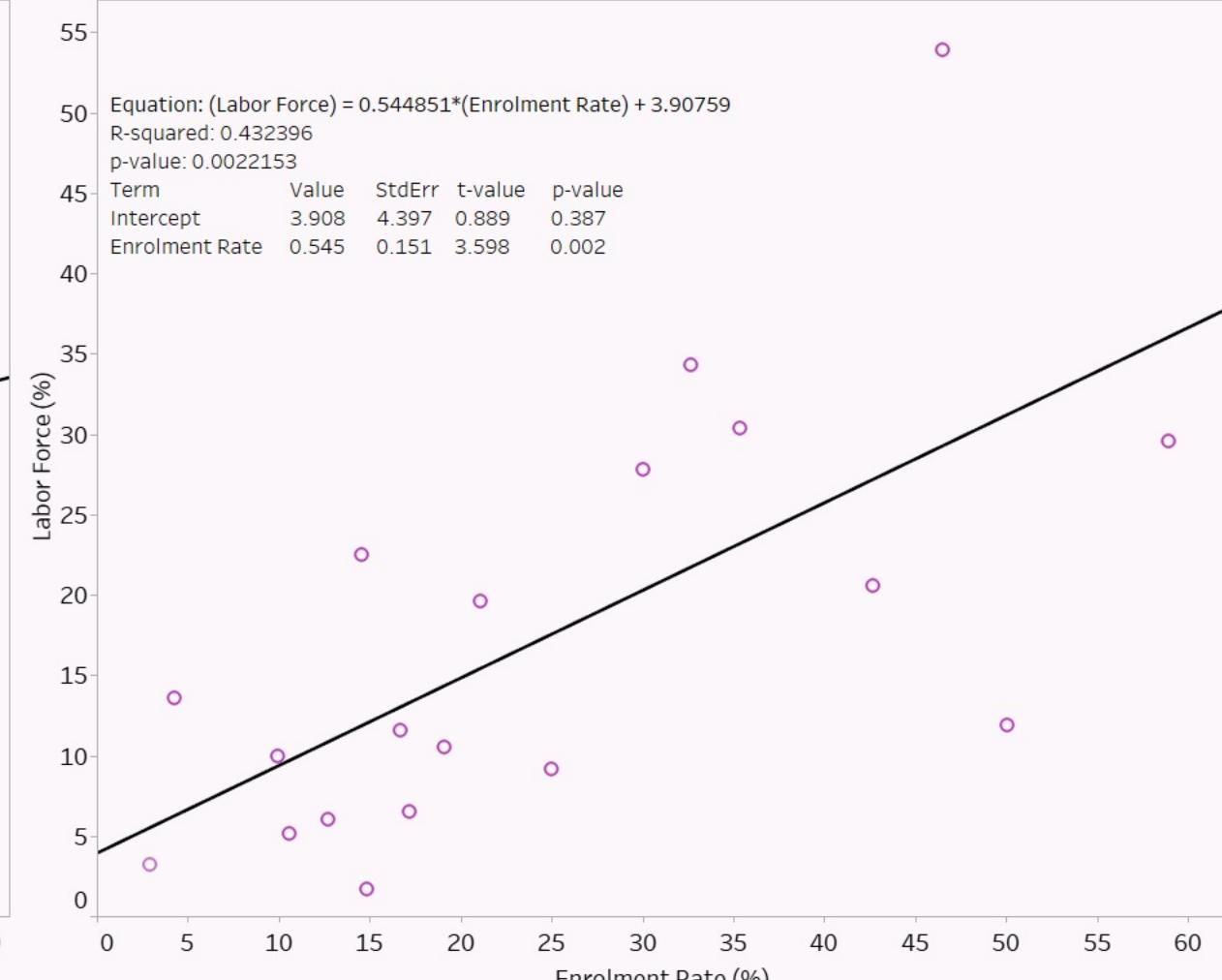


# RELATIONSHIP BETWEEN EDUCATION AND EMPLOYMENT (LOWER INCOME GROUP)

## MALE LABOR FORCE AND ENROLMENT RATE



## FEMALE LABOR FORCE AND ENROLMENT RATE



# CONCLUSION

- There is higher variation in indicators for low income countries.
  - The distribution of literacy rates is wider for low income groups.
- Gender inequality is more prevalent in low income countries.
- Child labor is higher among males which implies that there are higher expectations for males to enter the workforce than females, even at a young age.
- Labor force for high income and upper middle income countries were most affected during the 2008 financial crisis.

## CHALLENGES FACED

- Data format
  - The indicators were stacked in one column, making it difficult to use them without writing custom SQL queries.
- Feature selection
  - There are a vast number of indicators in the dataset which made feature selection more difficult for our linear regression model.
  - E.g., predicting male labor force using enrolment rate did not generate a strong model

## FUTURE WORK

- Gathering additional region-specific data and observations for the low-income group
- Feature engineering
  - Identify the best model to select most important features for predicting male and female labor force
- Predictive analysis
  - Use a classification model to predict income groups for countries for future years
- Create pivoted data frames for specific use cases