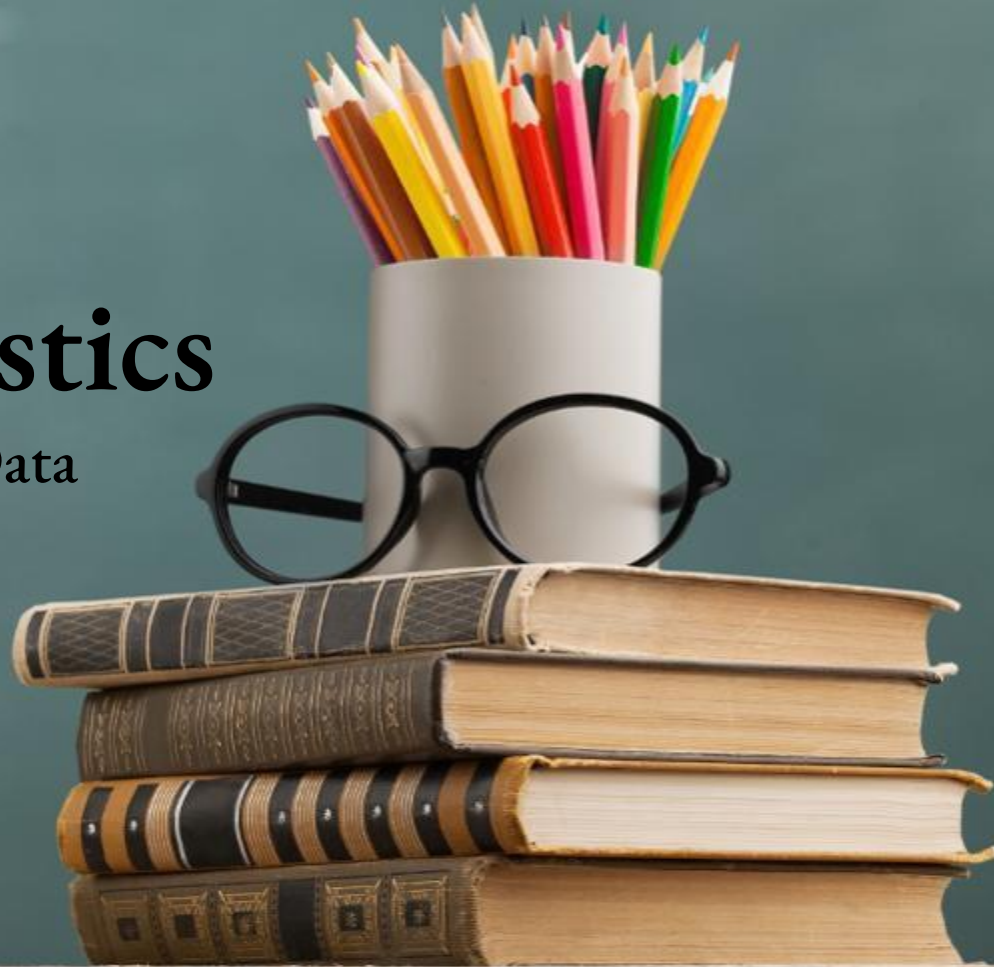


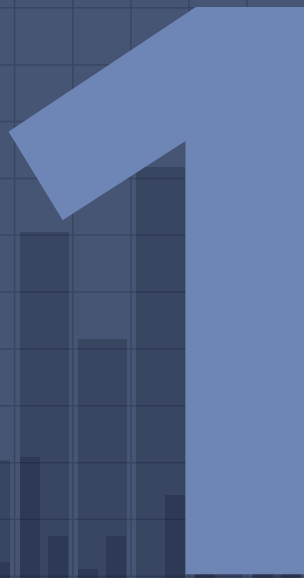
Education Statistics

From World Bank Open Data

By Nikhil Sharma, Biao (Frank)
Wang, Edward Shiang, Nadine
Tabucol



Introduction



Education

- What it is
 - The process of facilitating learning, or the acquisition of knowledge, skills, values, beliefs, and habits (from Wikipedia)
- Why we decided to analyze it
 - Vital for a person's development
 - Perceived differences in education for the wealthy and impoverished
 - Investigate the relationship between country income and education level

Data Description

- ▣ Where we got it from
 - Kaggle: Education Statistics
- ▣ Data collected by The World Bank
 - Africa Dataset
 - Barro-lee
- ▣ Contents of the data
 - Hundreds of internationally comparable indicators for education
 - Ranging from 1970 to 2100

Data Preview

Country Name	Country Code	Indicator Name	Indicator Code	1970	1971	1972	1973	1974	1975	1976	1977
Arab World	ARB	Adjusted net enrolment rate, lower secondary, both sexes (%)	UIS.NERA.2								
Arab World	ARB	Adjusted net enrolment rate, lower secondary, female (%)	UIS.NERA.2.F								
Arab World	ARB	Adjusted net enrolment rate, lower secondary, gender parity index (GPI)	UIS.NERA.2.GPI								
Arab World	ARB	Adjusted net enrolment rate, lower secondary, male (%)	UIS.NERA.2.M								
Arab World	ARB	Adjusted net enrolment rate, primary, both sexes (%)	SE.PRM.TENR	54.82212	54.89414	56.20944	57.26711	57.99114	59.36554	60.99996	61.92268
Arab World	ARB	Adjusted net enrolment rate, primary, female (%)	SE.PRM.TENR.FE	43.3511	43.31815	44.6407	45.84572	46.4495	48.36389	50.04619	51.24528
Arab World	ARB	Adjusted net enrolment rate, primary, gender parity index (GPI)	UIS.NERA.1.GPI	0.65857	0.6564	0.66329	0.67204	0.67261	0.69176	0.69995	0.71014
Arab World	ARB	Adjusted net enrolment rate, primary, male (%)	SE.PRM.TENR.MA	65.82623	65.99358	67.30186	68.21908	69.05901	69.91455	71.49951	72.16206
Arab World	ARB	Adjusted net enrolment rate, upper secondary, both sexes (%)	UIS.NERA.3								
Arab World	ARB	Adjusted net enrolment rate, upper secondary, female (%)	UIS.NERA.3.F								
Arab World	ARB	Adjusted net enrolment rate, upper secondary, gender parity index (GPI)	UIS.NERA.3.GPI								
Arab World	ARB	Adjusted net enrolment rate, upper secondary, male (%)	UIS.NERA.3.M								
Arab World	ARB	Adjusted net intake rate to Grade 1 of primary education, both sexes (%)	UIS.NIRA.1	52.44892	52.48975	52.63559	53.327	54.18465	54.86563	58.70163	58.87696
Arab World	ARB	Adjusted net intake rate to Grade 1 of primary education, female (%)	UIS.NIRA.1.F	44.34249	44.25571	44.44614	45.27044	46.17181	47.06202	50.17876	51.17889
Arab World	ARB	Adjusted net intake rate to Grade 1 of primary education, gender parity index (GPI)	UIS.NIRA.1.GPI	0.73634	0.73288	0.73473	0.74142	0.74621	0.75474	0.75024	0.7723
Arab World	ARB	Adjusted net intake rate to Grade 1 of primary education, male (%)	UIS.NIRA.1.M	60.22044	60.38639	60.49314	61.05912	61.8748	62.35518	66.88373	66.26824
Arab World	ARB	Adult illiterate population, 15+ years, % female	UIS.LPP.AG15T99								
Arab World	ARB	Adult illiterate population, 15+ years, both sexes (number)	UIS.LP.AG15T99								
Arab World	ARB	Adult illiterate population, 15+ years, female (number)	UIS.LP.AG15T99.F								
Arab World	ARB	Adult illiterate population, 15+ years, male (number)	UIS.LP.AG15T99.M								
Arab World	ARB	Adult literacy rate, population 15+ years, both sexes (%)	SE.ADT.LITR.ZS								
Arab World	ARB	Adult literacy rate, population 15+ years, female (%)	SE.ADT.LITR.FE.ZS								
Arab World	ARB	Adult literacy rate, population 15+ years, gender parity index (GPI)	UIS.LR.AG15T99.GPI								
Arab World	ARB	Adult literacy rate, population 15+ years, male (%)	SE.ADT.LITR.MA.ZS								
Arab World	ARB	Africa Dataset: Average number of grades per multigrade class in primary schools	UIS.AFR.AGRADMG.1.PU								
Arab World	ARB	Africa Dataset: Average number of pupils per mathematics textbook in Grade 1	UIS.AFR.PBR.1.G1.PU.MATH								
Arab World	ARB	Africa Dataset: Average number of pupils per mathematics textbook in Grade 2	UIS.AFR.PBR.1.G2.PU.MATH								
Arab World	ARB	Africa Dataset: Average number of pupils per mathematics textbook in Grade 3	UIS.AFR.PBR.1.G3.PU.MATH								
Arab World	ARB	Africa Dataset: Average number of pupils per mathematics textbook in Grade 4	UIS.AFR.PBR.1.G4.PU.MATH								
Arab World	ARB	Africa Dataset: Average number of pupils per mathematics textbook in Grade 5	UIS.AFR.PBR.1.G5.PU.MATH								

Analysis

2

Data Cleaning

- ▣ Data was very messy!
- ▣ Eliminated columns with missing values
- ▣ Filtered down to countries where high and low income were demarcated
- ▣ Took out columns directly relating to country prosperity (GDP, GNI, etc.)
- ▣ Dummy coded wealthy regions as 1, non-wealthy as 0

New Regions

- East Asia and Pacific
- East Asia and Pacific (excluding high income)
- Latin America and Caribbean
- Latin America and Caribbean (excluding high income)
- Europe and Central Asia
- Europe and Central Asia (excluding high income)
- Sub-Saharan Africa
- Sub-Saharan Africa (excluding high income)
- Middle East and North Africa
- Middle East and North Africa (excluding high income)

New Data Preview

Country Code	Country Name	Year	Enrollment in Tertiary Education	Income
EAP	East Asia & Pacific (excluding high income)	2014	59809028	0
EAS	East Asia & Pacific	2014	69097800	1
ECA	Europe & Central Asia (excluding high income)	2014	18188350	0
ECS	Europe & Central Asia	2014	37693412	1

Feature Importance and Modeling

- ▣ Goal: classify whether a country is wealthy or not based on education statistics
- ▣ Univariate Feature Selection from *scikit-learn* in Python
- ▣ Selected $k=10$ best features via repeated F-tests
- ▣ Random Forest

$k=10$ Important Features

- Enrollment in tertiary education, all programs, both sexes
- Enrollment in tertiary education, all programs, female
- Enrollment in tertiary education, ISCED 8 programs, both sexes
- Enrollment in tertiary education, ISCED 8 programs, female
- Enrollment in upper secondary education, female
- Enrollment in upper secondary general, both sexes
- Enrollment in upper secondary general, female
- Percentage of all students in tertiary education enrolled in ISCED 8, both sexes
- Teachers in secondary education, both sexes
- Teachers in secondary education, female

Variable Importance, Random Forest

Important Variable #1	Important Variable #2	Important Variable #3	Important Variable #4	Important Variable #5
Enrollment in Upper Secondary, General, Both Sexes	Percentage of All Students in Tertiary Education	Enrollment in Upper Secondary, General, Female	Enrollment in Tertiary Education, Both Sexes	Enrollment in Upper Secondary, Female

- Criteria for Variable Importance: Mean Decrease in Gini Importance
- Accuracy: 73.33%

Discussion

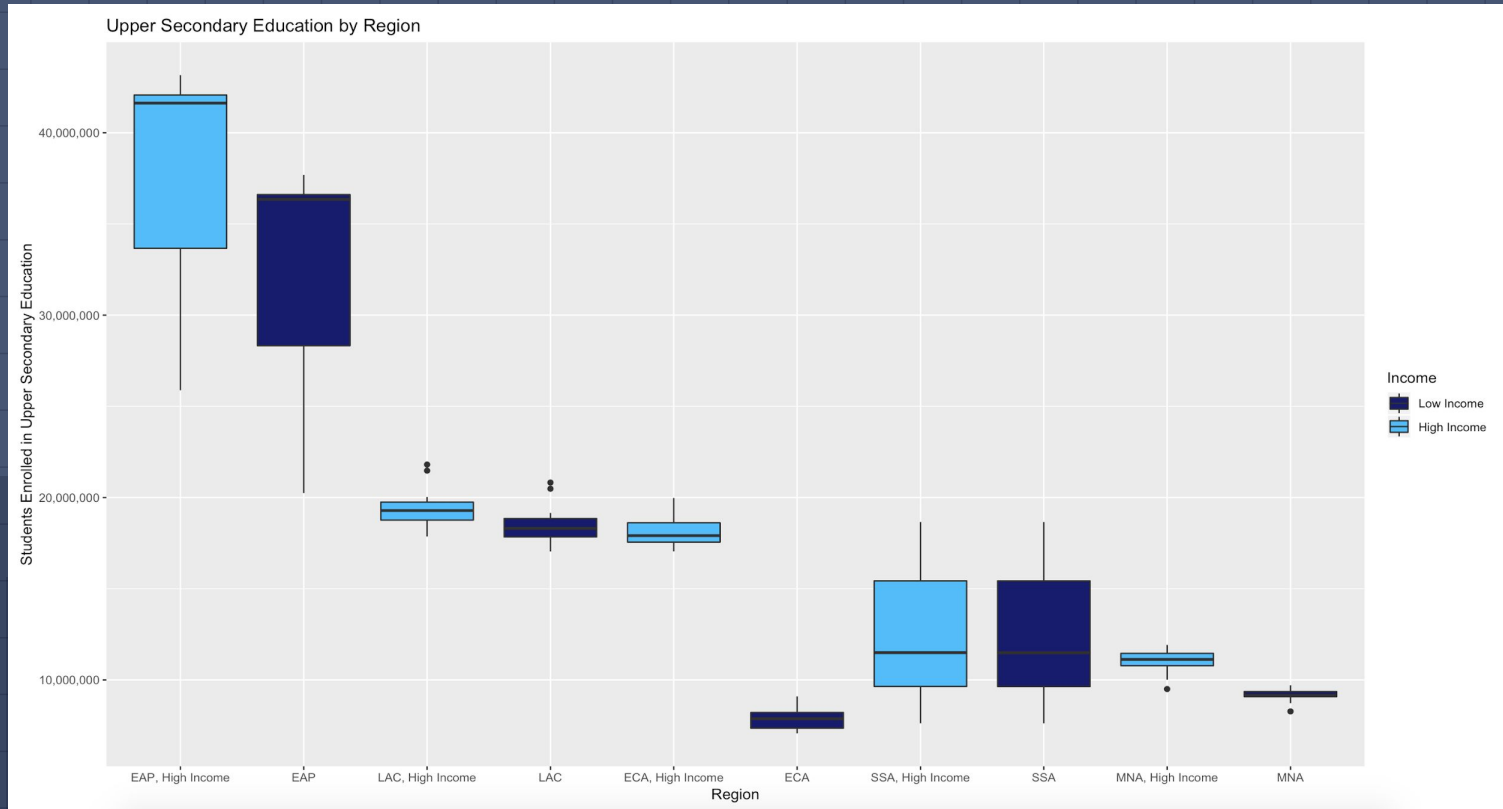
A large, light blue number 3 is positioned on the right side of the slide. The background is a dark blue grid with a silhouette of a bar chart at the bottom. The word 'Discussion' is written in white in the top left corner.

3

General Findings

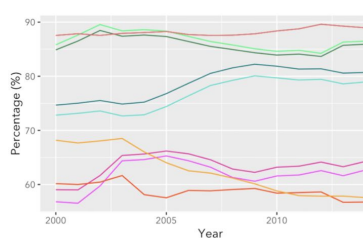
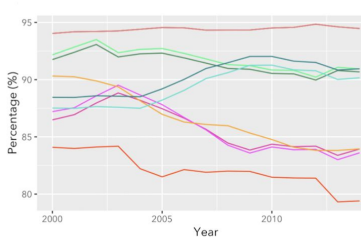
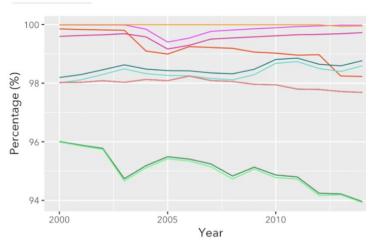
- ▣ **Secondary and Tertiary Education** are the best indicators of prosperity
 - Some countries may have limited access to high schools and colleges
- ▣ **Educated Females** also impact prosperity
 - Vital for historically marginalized people to be educated!

Upper Secondary Education by Region

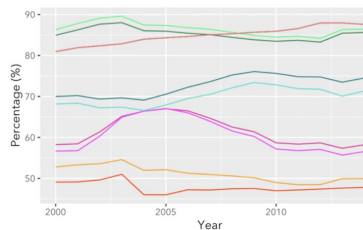
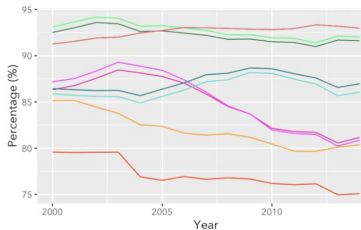
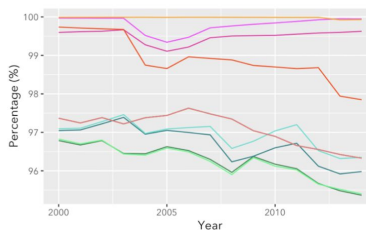


Percentage of Students Enrolled In General Programs By Gender and Education Level (%)

Female



Male



Lower Secondary
Education

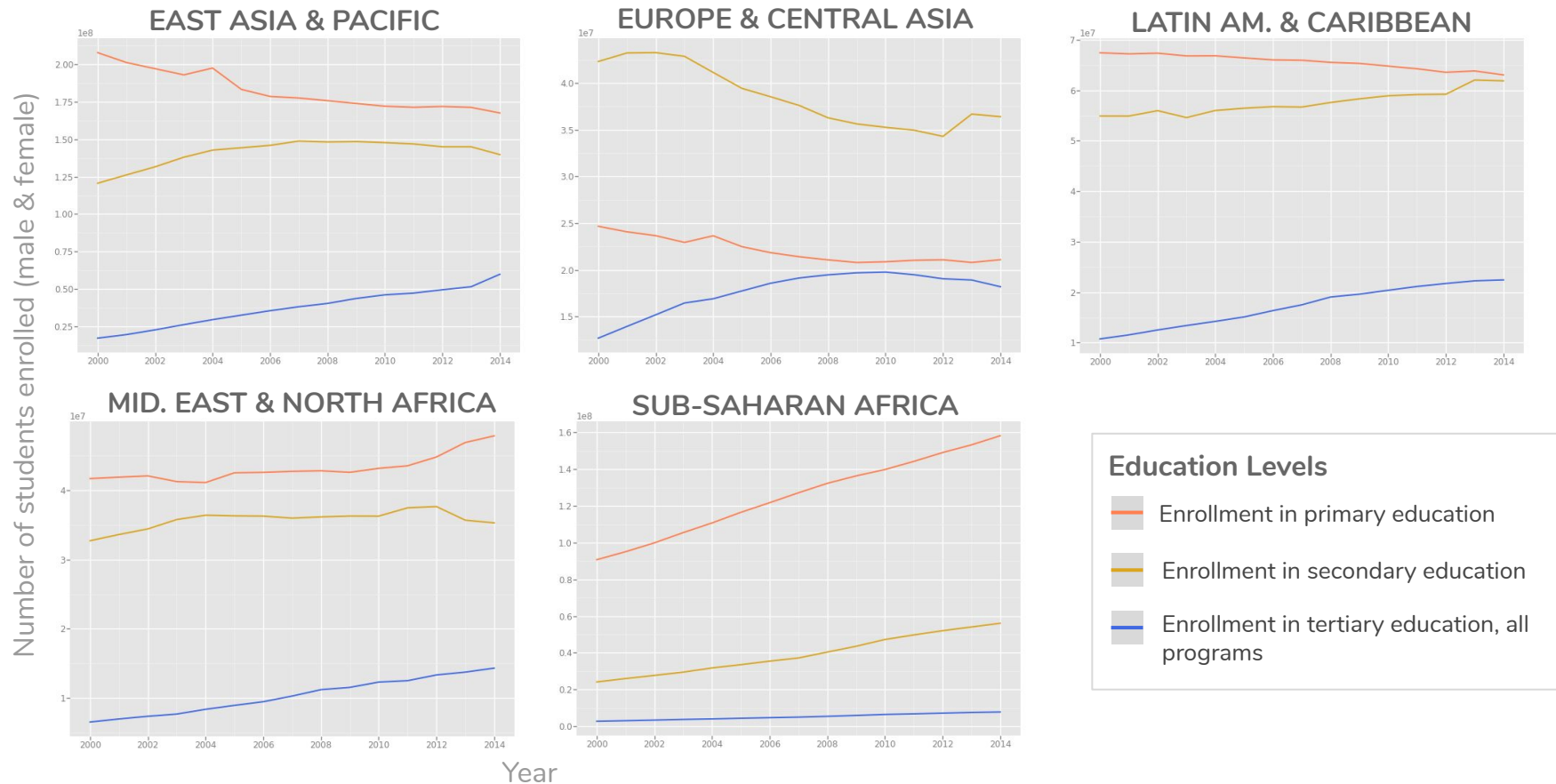
Secondary
Education

Upper Secondary
Education

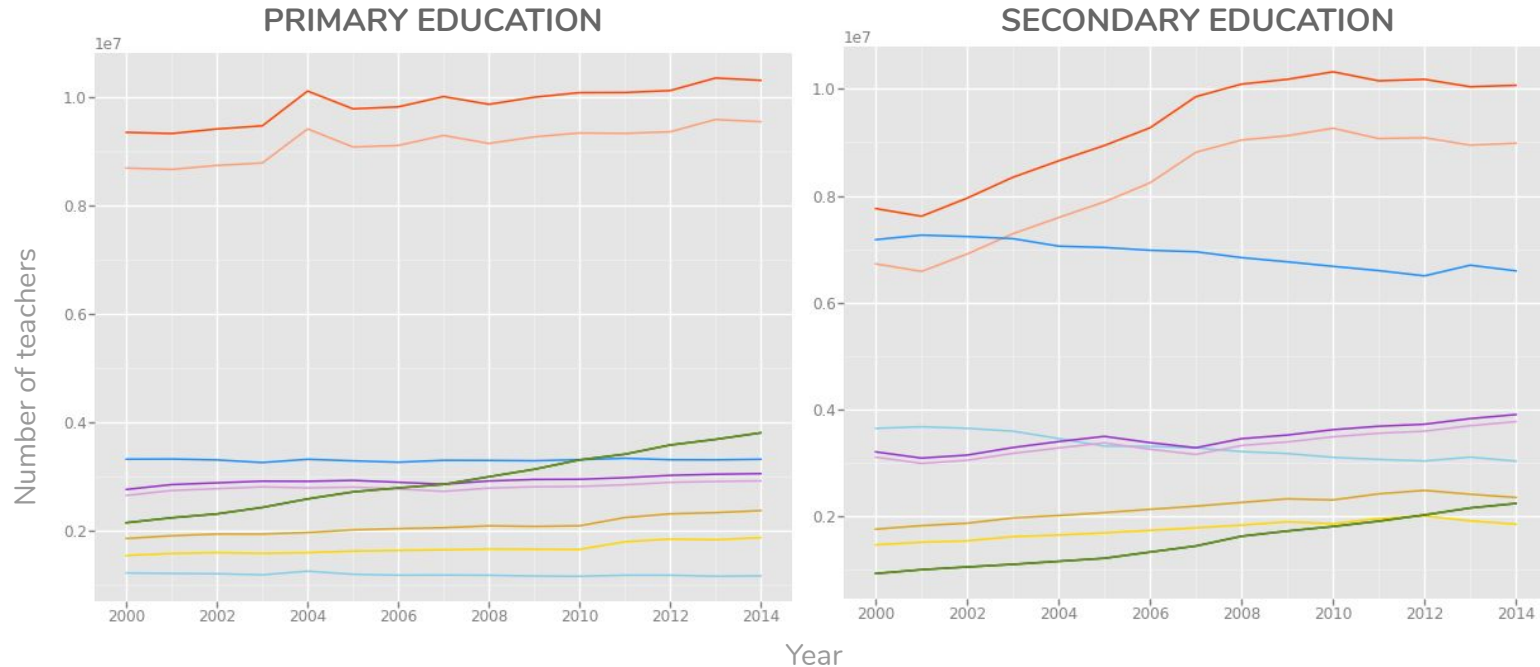
World Region Specification

- East Asia & Pacific
- East Asia & Pacific (excluding high income)
- Europe & Central Asia
- Europe & Central Asia (excluding high income)
- Latin America & Caribbean
- Latin America & Caribbean (excluding high income)
- Middle East & North Africa
- Middle East & North Africa (excluding high income)
- Sub-Saharan Africa
- Sub-Saharan Africa (excluding high income)

Enrollment rates across education levels, per region, excluding high-income (2000-2014)



Teachers in primary and secondary education, by region (2000-2014)



World Region Specification

East Asia & Pacific
... (excluding high-income)

Latin America & Caribbean
... (excluding high-income)

Europe & Central Asia
... (excluding high-income)


Middle East & North Africa
... (excluding high-income)

Sub-Saharan Africa
... (excluding high-income)

Further Studies



Further Studies

- Dig into the dataset and use more countries!
 - Make the analysis more granular
 - Try out some more models to see if they agree with each other
 - Do some external research on specific education factors (e.g., number of colleges in certain regions)
 - Use unsupervised methods and see if they match up with actual prosperity levels
- 
- A decorative background graphic at the bottom of the slide. It features a line chart with white circular markers connected by a thin white line, showing an overall upward trend with some fluctuations. Below the line chart is a bar chart with numerous vertical bars of varying heights, rendered in a light blue-grey color. The entire graphic is set against a dark blue background with a subtle grid pattern.