
Educate to Success

Isaac Duerr, Arjun Iyer



Data Sources

Educational Indicators Data
- World Bank

<https://databank.worldbank.org/reports.aspx?source=2&Topic=4>

Human Development Index Data
- United Nations Development Program

<https://hdr.undp.org/data-center>

Why this project?

We want to help set goals for what worldwide education should tend to.

This must start with being able to find a useful indicator that helps us put countries in comparison and how it correlates with educational indicators.

One of the stand out indicators we can look at is Human Development Index (HDI), which helps us do this.



HDI: What and Why?

$$HDI = (I_{Health} * I_{Education} * I_{Income})^{1/3}$$

Dimension	Indicator	Minimum	Maximum
Health	Life expectancy (years)	20	85
Education	Expected years of schooling (years)	0	18
	Mean years of schooling (years)	0	15
Standard of living	Gross national income per capita (2011 PPP \$)	100	75,000

- HDI is a summary primarily about average achievements in three main aspects of human development such like health, knowledge(education) and standard of living.
- Our thesis predicts that HDI serves as a good reference point to compare countries and shows correlation with multiple educational indicators (excluding expected years of schooling since it is already considered in calculating HDI).



Project thesis

Human Development Index correlates with multiple educational indicators.

Note: If true this shall provide us correlations that will help us set worldwide education indicator goals by looking at well performing countries in terms of HDI.



Process



Deducing Key Indicators

Fitting a multiple regression model to depict what factors play an important role..

Analysis

Exploring trends in the specified educational indicators to understand the relationships..



Conclusions

Utilizing specific indicators and HDI to narrow down to certain goals.

Looking at data richness

- We started by researching key education indicators.
 - Many indicators however had several NA values.
 - We, then sorted by Indicators with the least NA values so that the model can fit better and the plots can form better.
 - With a combination of the above two we picked indicators that stood out and narrowed it down to 15-20 indicators.
 - In the next step we fit a multiple regression model to narrow down to fewer indicators.

Fitting a Multiple Regression to narrow down factors

- Eliminating variables with high correlation.
- Eliminating or mitigating unscaled variables.
- Narrowing the model to factors significant at 0.05 level at the minimum.

```
Call:
lm(formula = y ~ ., data = df1)

Residuals:
    Min      1Q  Median      3Q     Max 
-0.19414 -0.03493  0.00361  0.03714  0.15548 

Coefficients:
                                         Estimate Std. Error t value
(Intercept)                         8.492e-01  2.779e-02 30.556
`Cumulative drop-out rate to the last grade of primary education, both sexes (%)` -1.702e-03  5.976e-04 -2.848
`Enrolment in tertiary education per 100,000 inhabitants, both sexes`          2.801e-05  5.174e-06  5.414
`Expenditure on education as % of total government expenditure (%)`           -4.305e-03  1.168e-03 -3.687
`Pupil-teacher ratio in primary education (headcount basis)`                  -5.707e-03  7.145e-04 -7.988
Pr(>|t|)                                < 2e-16 ***
                                         ***

(Intercept)                         0.005134 **
`Cumulative drop-out rate to the last grade of primary education, both sexes (%)` 0.000334 ***
`Enrolment in tertiary education per 100,000 inhabitants, both sexes`            2.93e-07 ***
`Expenditure on education as % of total government expenditure (%)`             6.89e-13 ***
`Pupil-teacher ratio in primary education (headcount basis)`                   ---

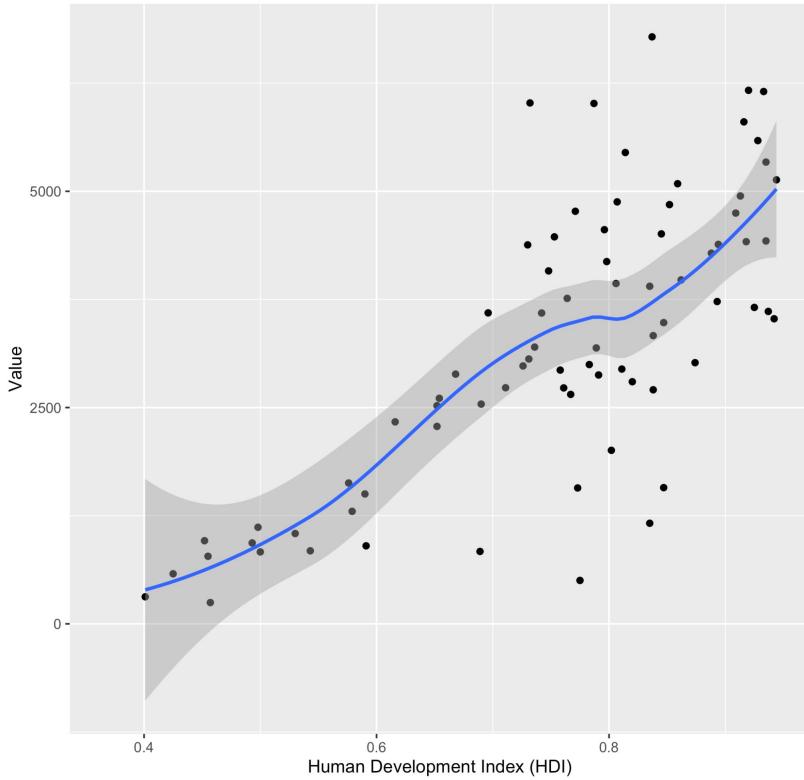
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05929 on 128 degrees of freedom
(71 observations deleted due to missingness)
Multiple R-squared:  0.8624,    Adjusted R-squared:  0.8581
F-statistic: 200.6 on 4 and 128 DF,  p-value: < 2.2e-16
```

Enrolment in tertiary education per 100,000 inhabitants, both sexes

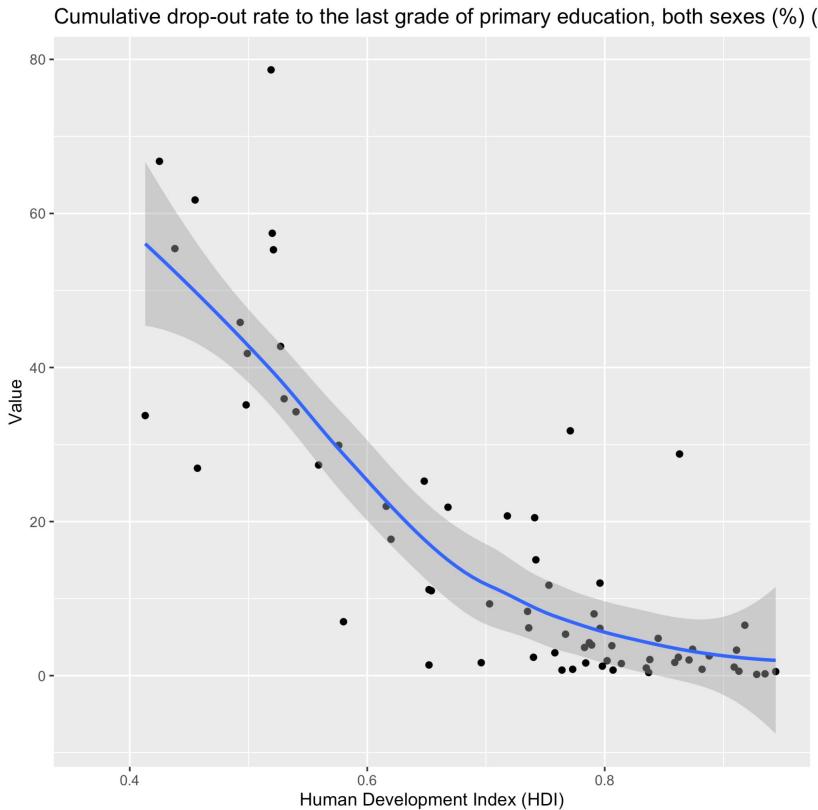
O1

Enrolment in tertiary education per 100,000 inhabitants, both sexes (2014)



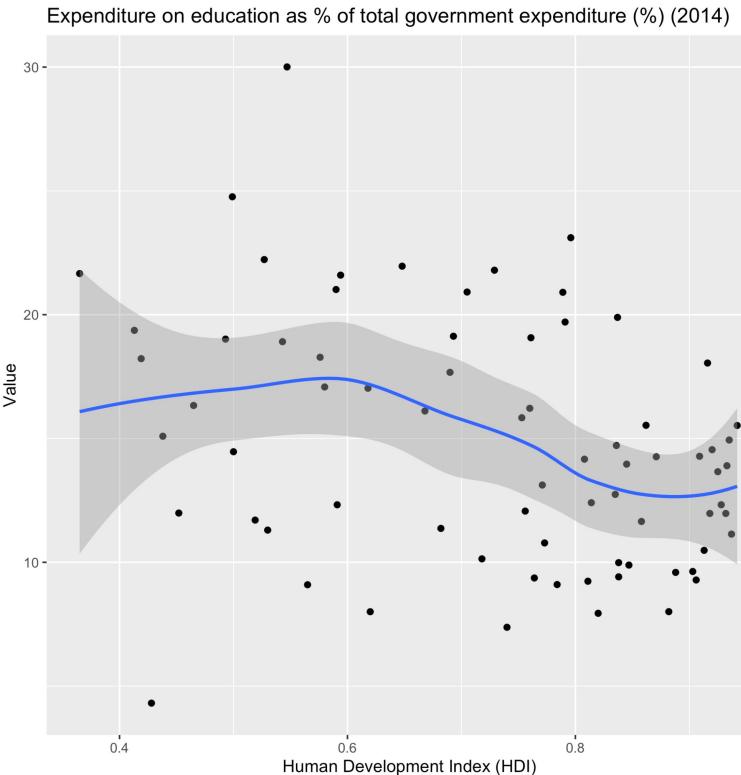
Cumulative drop-out rate to the last grade of primary education, both sexes (%)

02



Expenditure on education as % of total government expenditure (%)

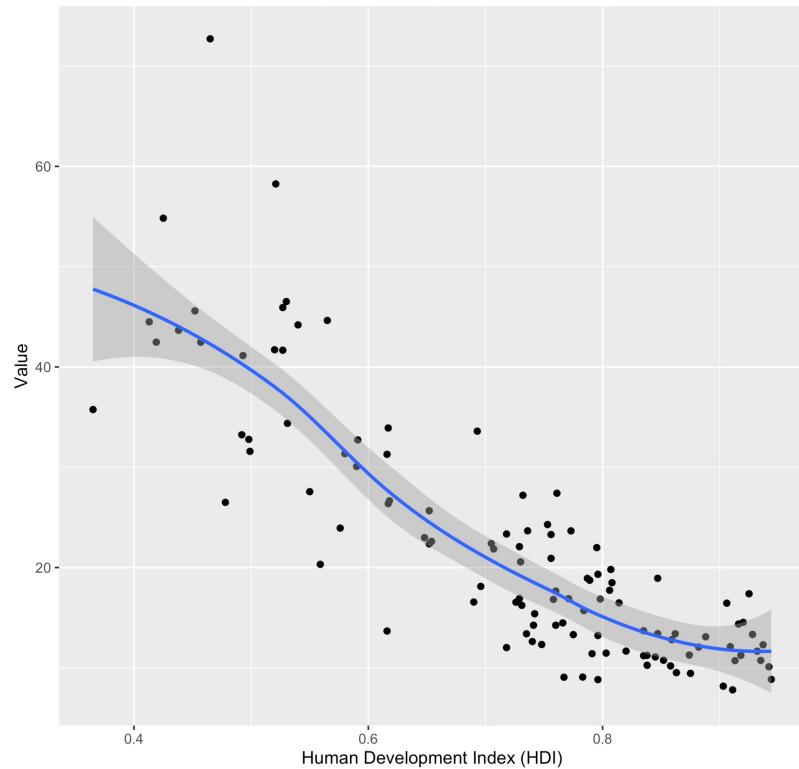
03



Pupil-teacher ratio in primary education (headcount basis)

04

Pupil-teacher ratio in primary education (headcount basis) (2014)



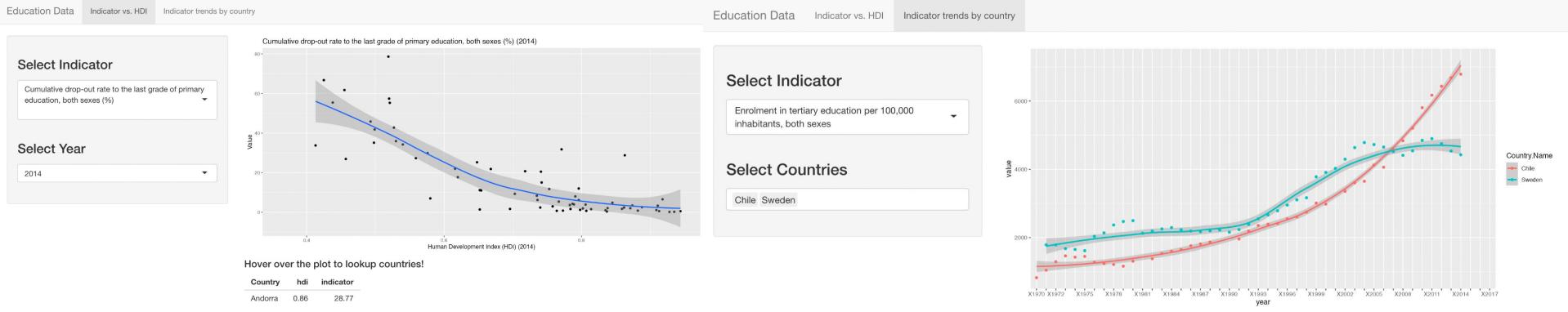


Gaining insight



Shiny!

We use our shiny app to visualise this data further and arrive at some goals.



Goals to set for development

Encouraging enrolment to tertiary education

Being consistent in sending younger generations through tertiary education(College level education)



Increasing government expenditure on education

Countries with consistent high HDIs can be seen modifying expenditure on education over time adjusting to their birth rate.

Reducing cumulative drop-out rate to the last year of primary education

Maintaining continuity and completion of primary education.

Reducing Pupil-teacher ratio

Lower pupil-teacher ratios lead to more individual attentiveness. This trend has been observed in countries with high HDIs.



Thank you.

