

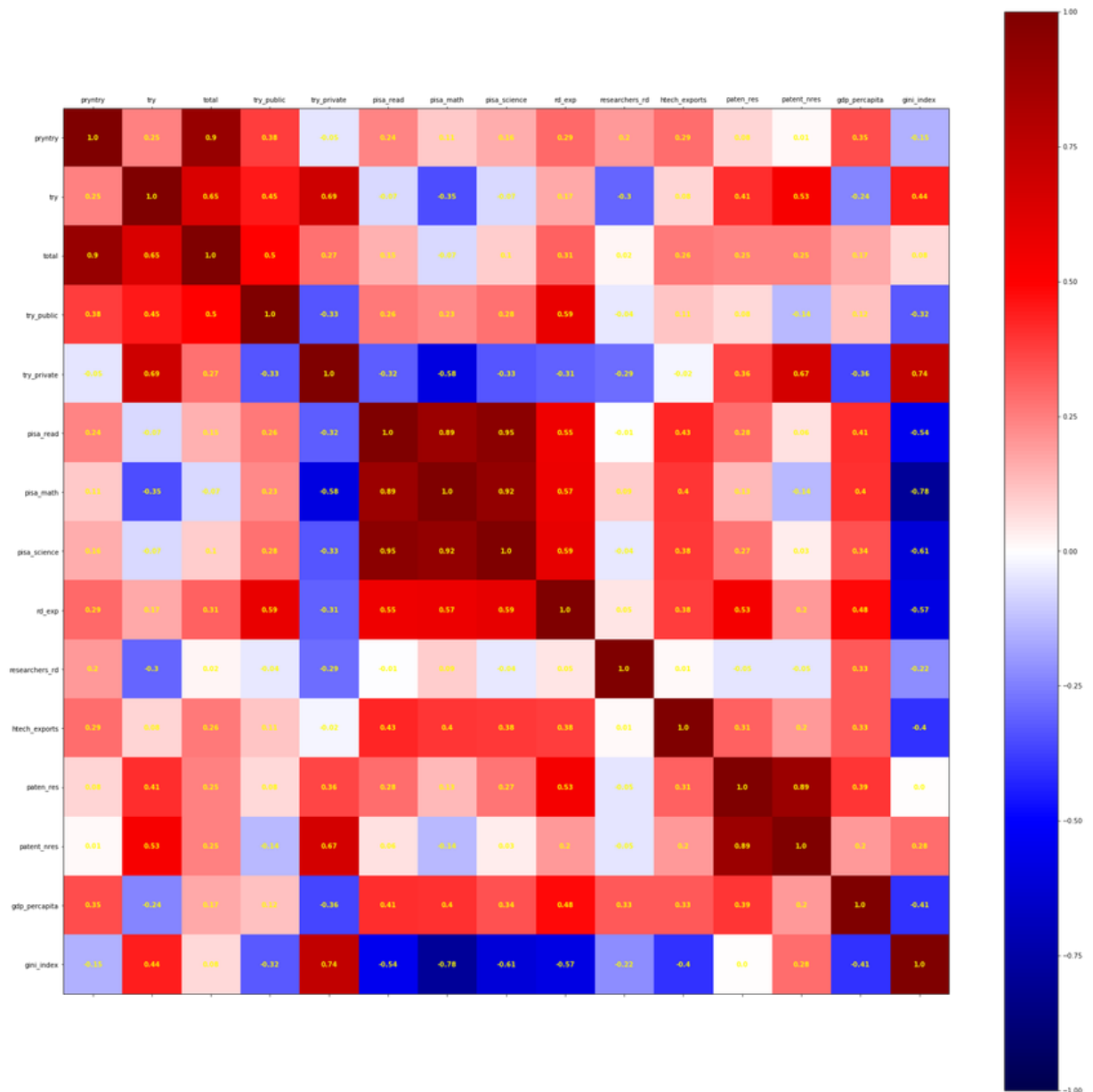
CS 210 Project Blog

MENU

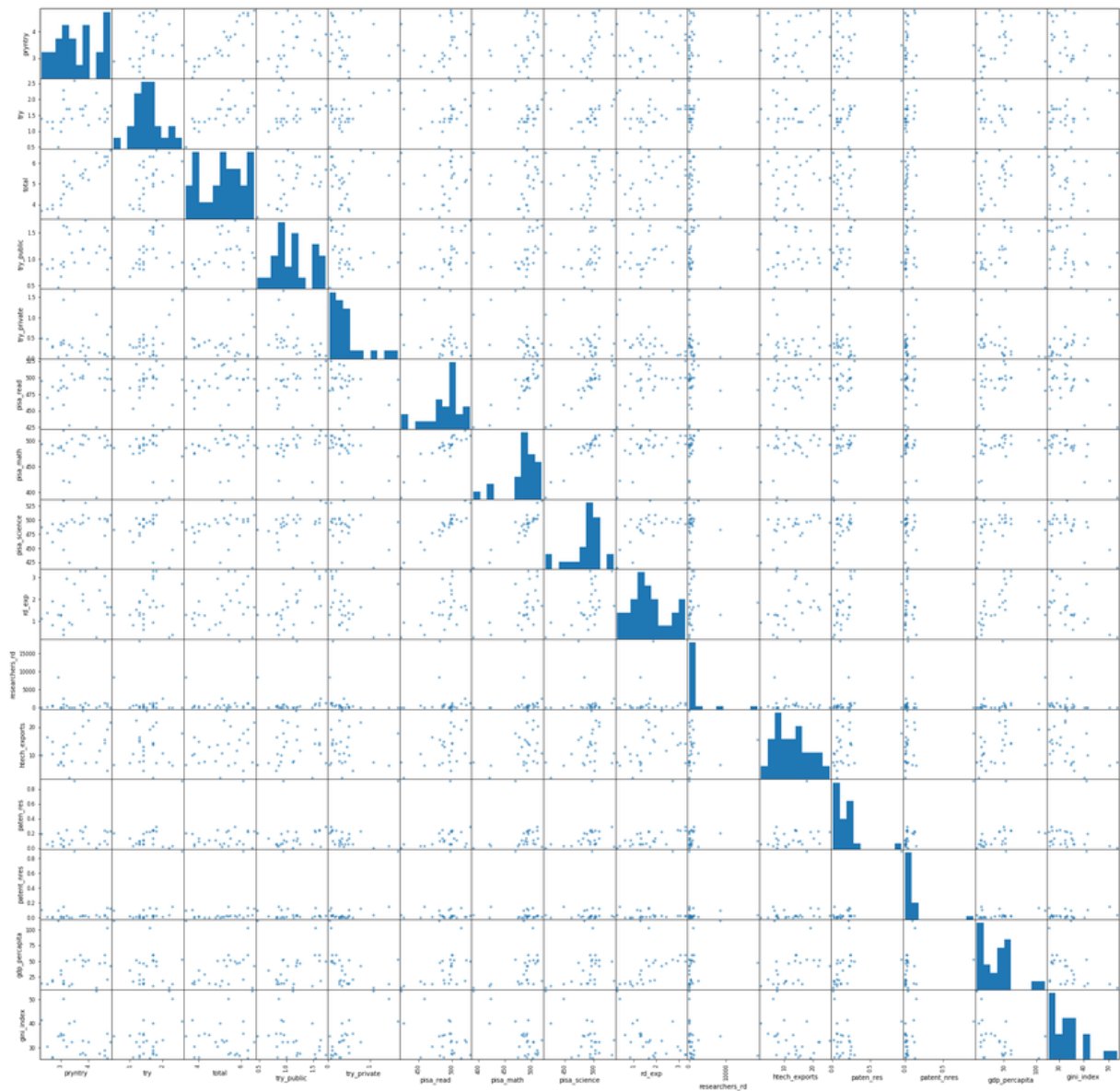
Correlation Matrix & Scatter Plot Matrix

Code : [Statistics-Plottings.pdf](#)

Correlation Matrix:



Scatter Plot Matrix:



March 24, 2017 / [Leave a comment](#)

Filtered Missing Data

After merging all data, I had 41 countries with their data. However, in some of the countries, I had some missing data. Therefore, I excluded those countries.

In final version, I have 26 countries.

Code for filtering: [FilterMissingData.pdf](#)

Data in .pdf format: [AllDataFull.pdf](#)

March 24, 2017 / [Leave a comment](#)

Merged Data

I merged all data.

I excluded public and private spending on education since there were inconsistencies between these 2 data and the other data I have.

In the resulting data I have 41 countries which were originally in OECD data. For World Bank data, I also used only these countries.

While merging,

- for education data, I used the same-year data
- for PISA scores, I used the latest data
- for other data, I used the same-year data if possible; then I first looked at the following year's data and then the previous year's data

In conclusion, I combined all data and exported that to a .csv file. From now on, I will use this data.

Code for merging all data: [MergeData.pdf](#)

Data in .pdf format: [AllData.pdf](#)

March 23, 2017 / [Leave a comment](#)

Code for Histograms

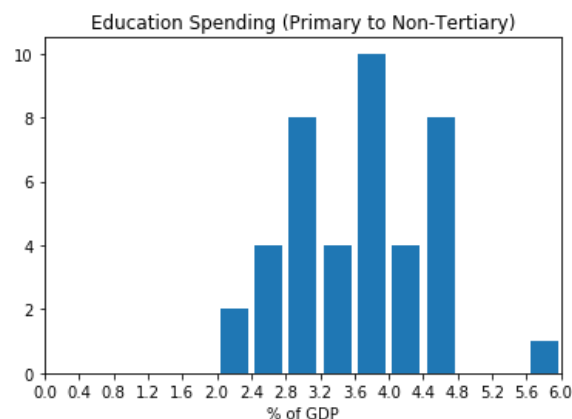
Jupyter notebook in pdf format is below:

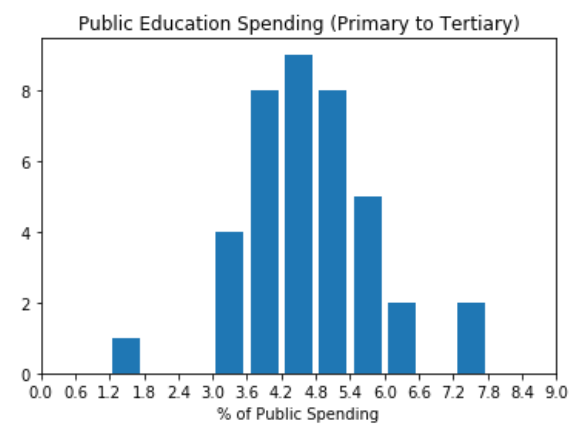
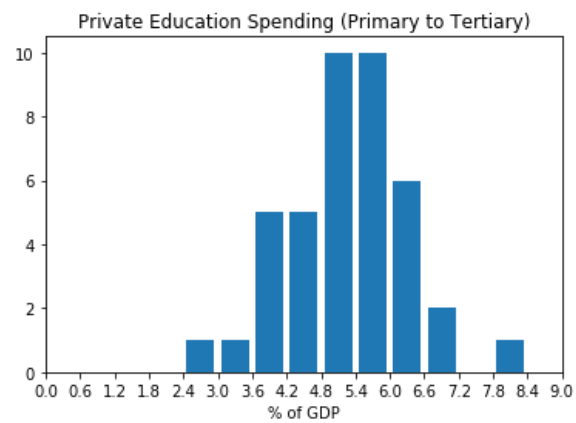
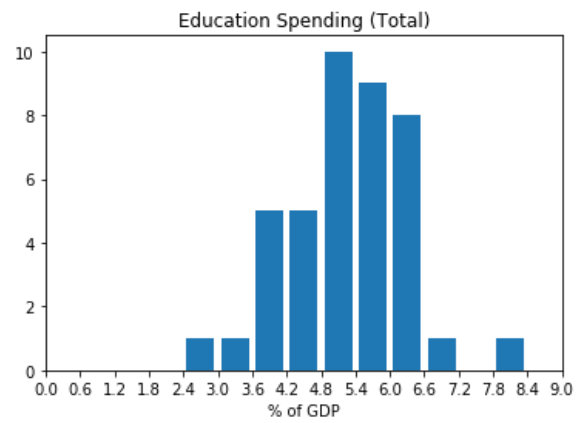
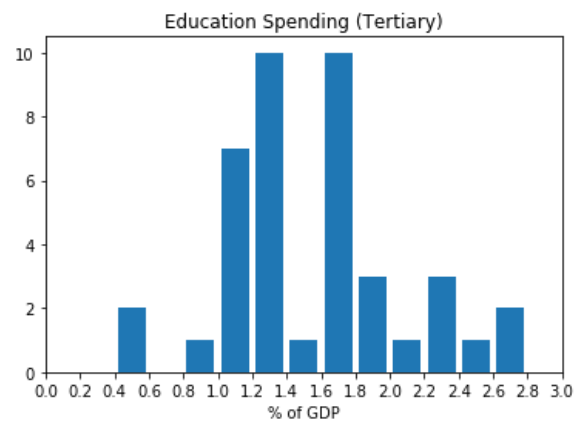
[Histograms.pdf](#)

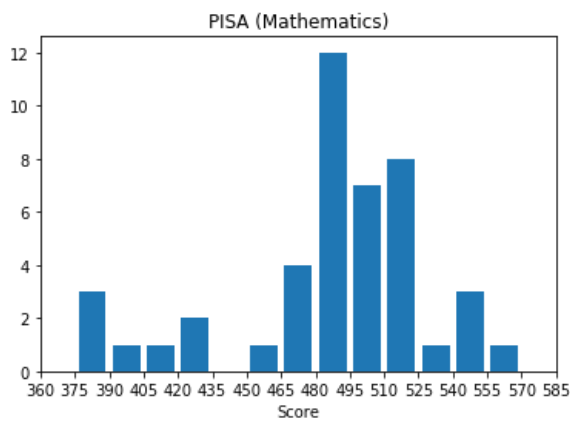
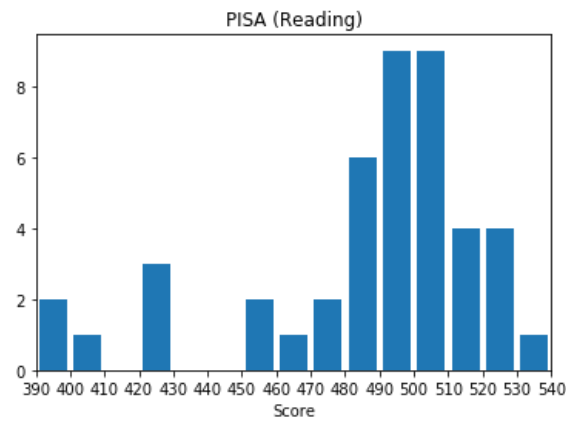
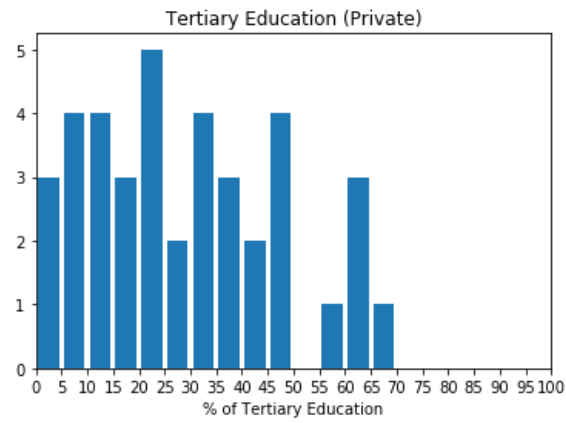
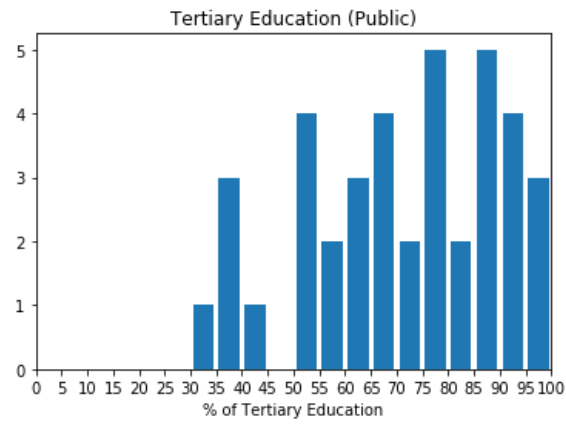
March 18, 2017 / [Leave a comment](#)

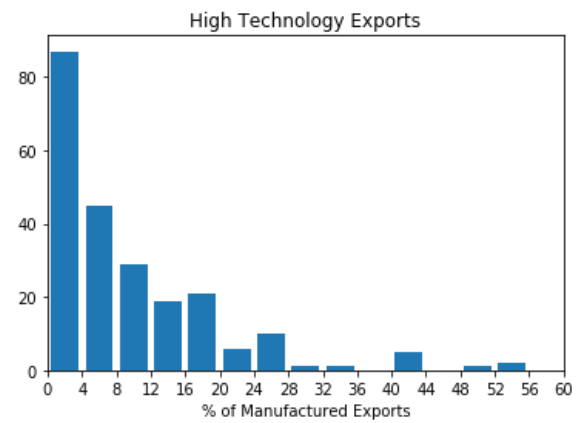
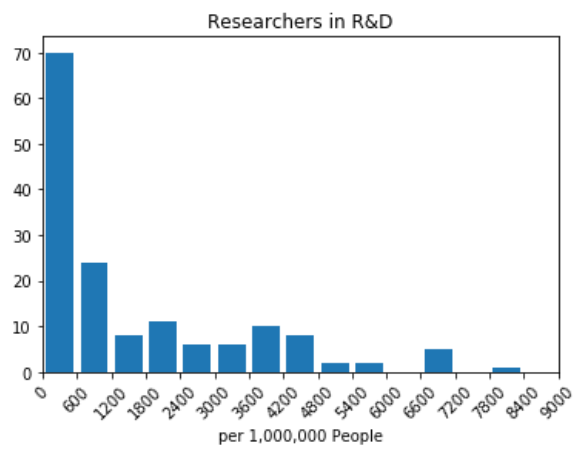
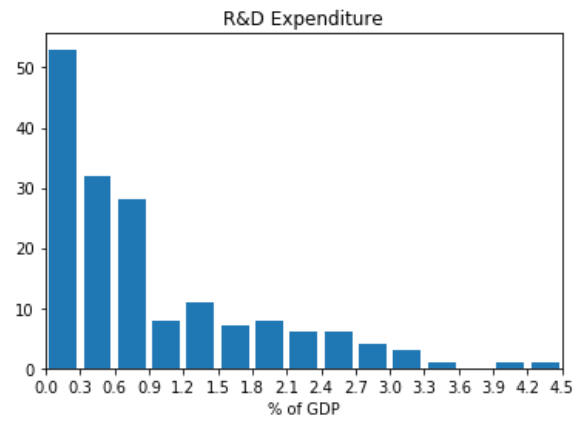
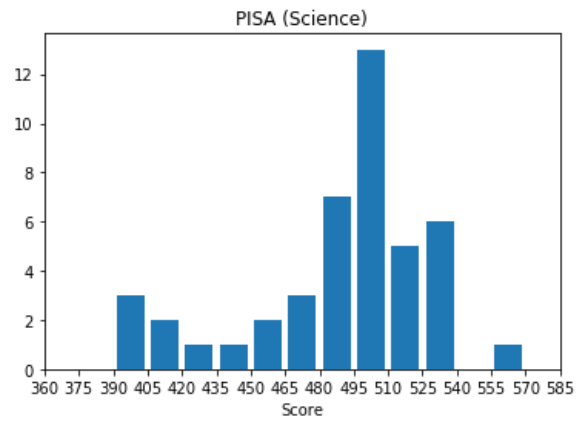
Histograms

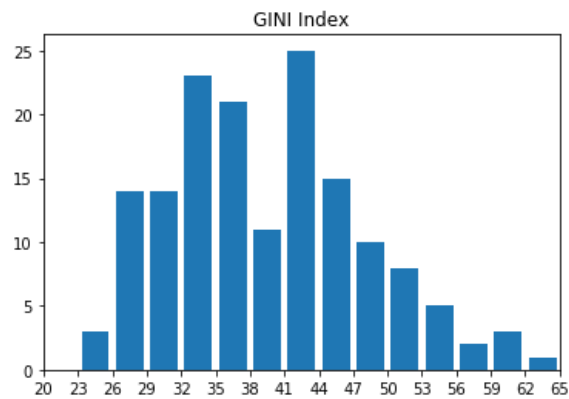
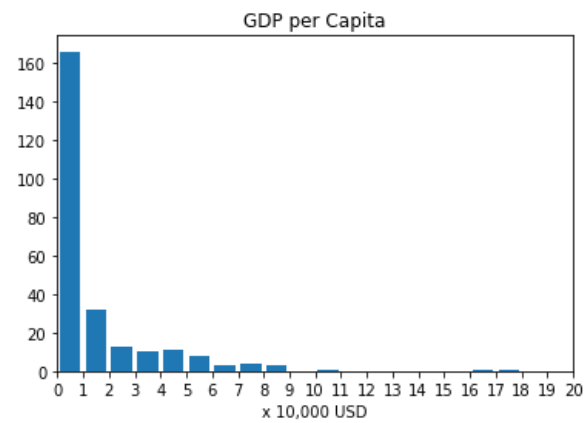
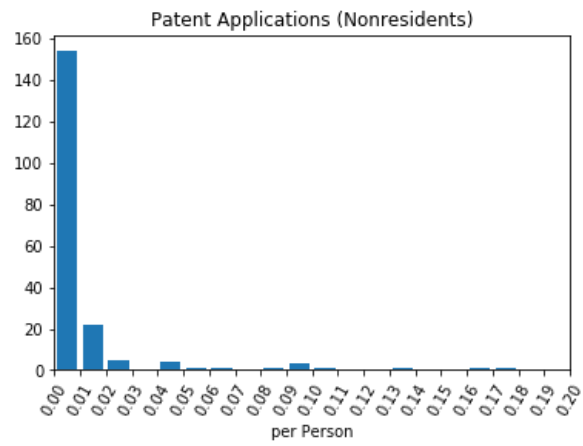
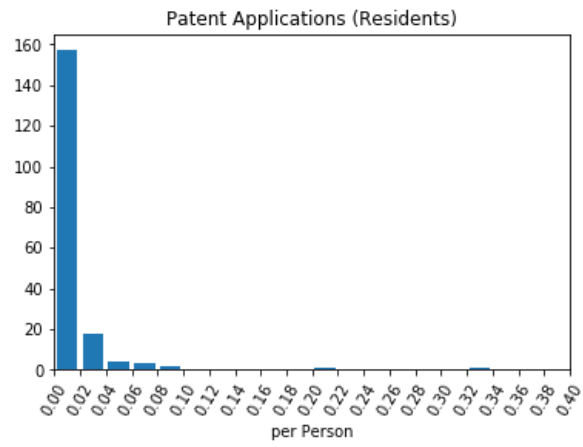
Histograms of each data are below:











Data Downloaded

All data are downloaded in .csv format.

- Data downloaded from OECD contains only last available data.
- Data downloaded from World Bank contains all data.

March 18, 2017 / [Leave a comment](#)

Project Proposal (v3)

Performance Analysis of Education According to Various Factors

Education industry is a very large industry according to both size and importance. Governments and/or people spend very large amount of money for this industry, and in return both society and the people who took education benefits from it.

However, there are lots of discussions on the approach to the education. Some people think that the state should provide education to its citizens, whereas others think that private sector can provide better and more efficient service. Also, there is another discussion about the consequences of education. Is it beneficial to the society or to the person more?

In this project, I will analyse the spendings on education both public and private, the distribution of these spendings over education levels (pre-university, university), and the consequences related to the education such as test scores or its effects on society and economy. Then, I will try to find some correlations about all these factors and consequences.

(According to these correlations, I may develop a model to classify countries in different groups such as developed and developing countries according to their spendings on education and some other indicators.)

Below there are links to some data that may be useful on this project:

Education Spending

<https://data.oecd.org/eduresource/education-spending.htm>

Private Spending on Education:

<https://data.oecd.org/eduresource/private-spending-on-education.htm>

Public Spending on Education:

<https://data.oecd.org/eduresource/public-spending-on-education.htm>

Spending on Tertiary Education:

<https://data.oecd.org/eduresource/spending-on-tertiary-education.htm>

PISA Scores:

<https://data.oecd.org/pisa/reading-performance-pisa.htm>

<https://data.oecd.org/pisa/mathematics-performance-pisa.htm>

<https://data.oecd.org/pisa/science-performance-pisa.htm>

R&D Expenditures:

[http://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?
end=2013&name__desc=false&start=1996&view=chart](http://data.worldbank.org/indicator/GB.XPD.RSDV.GD.ZS?end=2013&name__desc=false&start=1996&view=chart)

Researchers in R&D:

<http://data.worldbank.org/indicator/SP.POP.SCIE.RD.P6?view=chart>

High Technology Exports:

<http://data.worldbank.org/indicator/TX.VAL.TECH.MF.ZS?view=chart>

Patent Applications (Residents):

<http://data.worldbank.org/indicator/IP.PAT.RESD?view=chart>

Patent Applications (Nonresidents):

<http://data.worldbank.org/indicator/IP.PAT.NRES?view=chart>

Population:

<http://data.worldbank.org/indicator/SP.POP.TOTL?view=chart>

GDP Per Capita:

<http://data.worldbank.org/indicator/NY.GDP.PCAP.CD>

GINI Index:

<http://data.worldbank.org/indicator/SI.POV.GINI?view=chart>

March 18, 2017 / [Leave a comment](#)

Search ...



RECENT POSTS

- [Reports](#)
May 19, 2017
- [ZeroR Update](#)
May 19, 2017
- [Equal Weight DTCs](#)
May 14, 2017
- [Plan for last phase](#)
May 12, 2017
- [Logistic Regressions for Percentages](#)
May 7, 2017
- [Decision Trees for Percentages](#)
May 7, 2017
- [Processing Data for Percentages](#)
May 7, 2017
- [Confusion Matrices](#)
May 3, 2017
- [Comparison of Different Models](#)
April 29, 2017
- [Logistic Regressions](#)
April 29, 2017
- [SVMs](#)
April 29, 2017
- [Decision Trees](#)
April 29, 2017
- [Linear Regressions](#)
April 29, 2017
- [Getting and Processing Data](#)
April 29, 2017
- [GitHub Integration](#)
April 26, 2017
- [Part 1 Overview](#)
April 24, 2017
- [Planning for the rest of the project](#)
April 24, 2017
- [Regression \(New\)](#)
April 14, 2017
- [Cluster](#)
April 13, 2017
- [Regressions and Some Plots](#)
April 9, 2017
- [Minimized Data](#)
April 8, 2017
- [Correlation Matrix & Scatter Plot Matrix](#)
March 24, 2017
- [Filtered Missing Data](#)
March 24, 2017

- [Merged Data](#)

March 23, 2017

- [Code for Histograms](#)

March 18, 2017

- [Histograms](#)

March 18, 2017

- [Data Downloaded](#)

March 18, 2017

- [Project Proposal \(v3\)](#)

March 18, 2017

- [Project Proposal \(v2\)](#)

March 2, 2017

- [Project Proposal](#)

March 1, 2017

 Follow ...