



Checkpoint I: Project Proposal

Group: 14

Date: 2021/10/02

Domain

Explore the relation between state-level investment in education with academic success and grades, providing a comparison between states and its evolution throughout the years

Dataset

In this project we will be using the U.S Education Dataset from the Unification Project. This dataset was obtained by aggregating data from the U.S. Census Bureau and the National Center for Education Statistics (NCES). It contains information aggregated by state and year from 1998 to 2016 about revenue, expenditure on instruction, support services and capital outlay and NAEP grades for mathematics and reading for both fourth grade and eight grade students. The dataset was taken from [Kaggle](#).

We will also be using data from the National Governors Association that was aggregated by an investigator from the University of Pennsylvania in a dataset which includes information regarding the governing parties in each state throughout the history of the United States of America. This dataset was taken from [OpenICPSR](#).

In order to obtain per capita values for the information regarding revenue and expenditure we will be using a dataset from the [U.S. Census Bureau website](#) that contains every state's population in each decade.

Example Questions

In this section we provide an overview and quick description on some of the questions we will answer in this project.

- How did changes in investment in education impact grades and student enrolment?
The goal of this question is to understand if a larger investment in education is correlated with better results.
- Does state revenue impact academic success in a better way than state investment in education?
We are assuming that higher state revenue has some influence in parents' investment in their children's education, allowing state investment to be less necessary without having a negative impact in their academic success.
- Which type of investment (instruction, support services, capital outlay) has the greatest impact in enrolment and academic success?
It's important to identify the areas where a state should invest heavily, and areas with investment gaps.
- Is there any correlation between the party that governs a state and its investment in education?

We aim to analyse whether the political party of the elected officials has any impact in their willingness to invest in education. To answer this question, we will have to merge information from two different datasets: the U.S Education dataset and the National Governors Association dataset.

- Is there a nationwide trend regarding student ethnicity and their academic success?

With this we aim to understand if there is any correlation between ethnicity and academic success.

Data Sample

Some examples of data that show that the above is possible and adequate.

```
(from "states_all_extended.csv")
state; year; total_expenditure; ...
California; 2016; 85320133.0; 26058021.0; ...
California; 2015; 78365958.0; 26058021.0; ...
California; 2014; 72506810.0; 26058021.0; ...

(from "states_all_extended.csv")
state; year; ...; instruction_exp; support_services_exp; capital_outlay_exp; ...
California; 2015; ...; 35539425; 21594573; 6124103; ...
Kentucky; 2015; ...; 3807748; 2417226; 600859; ...

(from "states_all_extended.csv")
state; year; ...; math_8_average_score; reading_8_average_score; ...
California; 2015; ...; 275.324; 258.768; ...
Kentucky; 2015; ...; 277.947; 268.200; ...

(from "states_all_extended.csv")
state; year; ...; reading_8_african; reading_8_asian; reading_8_hispanic; ...
California; 2015; ...; 274; 282; 268; ...
Kentucky; 2015; ...; 262; 252; 270; ...

(from "states_all_extended.csv")
state; year; ...; state_revenue; local_revenue; total_expenditure; ...;
math_8_average_score; reading_8_average_score; ...
California; 2013; ...; 35141208.0; 25891045.0; 69727119.0; ...; 234.0; 213.0;
...
Colorado; 2013; ...; 3693829.0; 4436773.0; 8699125.0; ...; 247.0; 227.0; ...

(from "united_states_governors_1775_2020.csv")
governor; state; time_in_office; party; year
Pete Wilson; California; 1991 - 1999; Republican; 1992
Pete Wilson; California; 1991 - 1999; Republican; 1993
Gray Davis; California; 1999 - 2003; Democrat; 1999

(from "apportionment.csv")
name; ...; year; resident_population; ...
Alabama; ...; 2010; 4,779,736; ...
Alabama; ...; 2020; 5,024,279; ...
```