Daniel Gong
Capstone Project
SpringBoard Data Science

California Public School Analysis

Problem

This project aims to discover if increased state funding for public / charter school districts over time improves performance metrics such as enrollment, graduation rate, and dropout rate. Since state funding is tied to metrics described above, an increase could theoretically kickstart a feedback loop. This research also aims to uncover the effects of funding changes on non-traditional forms of educational institution such as virtual schools.

Clientele

Clientele will consist legislators and education professionals at a state level who are positioned to influence funding policies. Based upon patterns derived from this analysis, the client may choose to alter their distribution decisions to maximize ROI across different education institutions in the state of California. It will also allow for a segmented targeting approach to funding that is justified by data.

On the local level, this analysis could better inform grant giving nonprofits to channel their funding as well as local schools to better understand the merits and demerits of their larger institutional landscape.

Data

This project will analyze California Department of Education's (CDE) compilation of public and charter school data (enrollment, dropout, outcomes) and, SACS funding reports dating back to 2009. All datasets are hosted on a public domain.

Sample Data:

Hypothesis

Consistent increase in funding in public and charter school districts will produce a positive feedback loop that desired metrics such as increases in enrollment and decreased dropouts. Due to the innovative nature of virtual schools that allow for low overhead costs, this positive effect should produce a larger effect on these schools / districts.

Approach

Exploratory Analysis / Predictive Model Development

I will be identifying schools where funding stayed relatively level since 2009 as the control group. Multiple regression will be used to determine statistical significance of funding increases with variables of interest (enrollment, dropout, graduation, etc.). The processing will be done with python to yield a model that can be trained on a group of test data. Analysis will prove / disprove our hypothesis. Finally, code and data will be used to derive insights for deliverables.

Granger causality test will be employed to determine causal relationship between funding changes and school performance.

Deliverables

1. Code and data, graphics/visualizations
2. Paper with Executive Summary
3. Presentation