Colorado Public School Academic Performance

A Study of Financing versus Student Performance

Jarryd Allison
CSPB 4502: Data Mining
University of Colorado
Boulder, CO 80309
allisonj@colorado.edu.com

Problem Statement

Public education in the United States has come under increased scrutiny of late, due to consistently poor trends in academic performance. This has hit the Denver metropolitan area as well, with numerous articles bemoaning the state of public education[1,2].

The most recent data from US News for Denver area schools shows that while the operating budget grows, considerably over the past decade[3], test scores continue to decline or at least fail to improve. Proficiency scores for reading hover in the low 30s to upper 40 percentages, but math remains stuck below 30% proficiency across all age groups[4].

The combination of low test scores but accompanied with a growing budget seems to be counterintuitive, and demands closer inspection to better understand the various relationships that exist within the available data. For example, Denver does allocate funding differently based on perceived and measured needs of students[5]. Could a focused look at financing across Denver public schools reveal additional measures to better allocate funding, or provide links between funding and increased performance? The aim of this research topic is to answer this critical question.

CCS Concepts

Social and professional topics → Professional Topics

Keywords

Education, finance, data mining, analysis

ACM Reference format:

Jarryd Allison. 2024. Colorado Public School Academic Performance: A Study of Financing versus Student Performance. In *. ACM, New York, NY, USA, 3 pages.* https://doi.org/10.1145/XXXX

Literature survey

There exists a large number of articles and research that describe the issues that Denver Public Schools have with regards to test scores and low performance[1,2,6], and financial data exists to describe the growth in the budget[3], but very little reporting and analysis exists to justify increases, at least from what appears to be publicly available at the time of the writing of this paper. Aside from the sources listed, no analysis exists that intends to identify patterns and provide a holistic view of the budget versus student performance, and possibly help inform public policy to level set expectations, understand expenditures, and identify programs that may increase performance and educational outcomes in Denver public schools. There do exist compilation reports describing funding and performance, though these fail to describe significant patterns and a deeper analysis regarding the data[9].

Proposed Work

The proposed work involves a careful process of cleaning, preprocessing, and integration of multiple datasets to better understand the problem set. This differs from the literature survey conducted simply because the depth of analysis is lacking. The literature survey indeed studies the test performance dataset, and also arrays anecdotal evidence atop it to determine whether performance is good or bad. But the literature review is severely lacking in objective data mining techniques to truly develop a deeper understanding, through pattern exploration and analysis. Additionally, layering in financial data also appears to be completely ignored by the majority of the literature reviewed for this project.

The following is the proposed work to complete this project, including data cleaning, data preprocessing, and data integration.

1. **Data Cleaning**

Much of the data currently exists in excel documents, specific to individual school years, with omissions and various discrepancies for multiple attributes across the data sets. The data cleaning step will infer missing data, using either the mean, median, to apply the central tendency, or from determining the most likely values from surrounding data points. Because the dataset is at times inconsistent, this step will also involve identifying discrepancies, and developing measures to process the data to derive meaningful insights relevant to the research topic.

2. **Data Preprocessing**

Data preprocessing will first involve data reduction, given the size of the dataset and the scope of the project. Multiple attributes simply will not apply, and will be removed. Duplicates are not an anticipated issue, as the data is structured by schools in the Denver area, meaning that duplicates are highly unlikely, but will be screened to ensure duplicates do not exist. Sampling and clustering will be attempted, both to better refine the processes developed in this research, as well as to determine if significant patterns emerge. Finally, data normalization will be used to identify relevant trends amongst various factors, such as test scores across subjects, aggregate pass/fail percentages, and so on.

3. **Data Integration**

The dataset identified for school performance is clean, well structured, and relatively complete. However, integrating this data with financial data will be a comprehensive step in the work required to align finance information alongside the academic performance dataset. To do this, a web scraper will likely need to be developed in order to collect the finite information required for this research.

Datasets

The datasets identified for the scope of this research are as follows:

1. **Colorado Measures of Academic Success (CMAS) - Mathematics, English Language Arts, Science and Social Studies Data and Results (2014 - 2023)**

The CMAS performance dataset[7] is compiled and provided by the Colorado Department of Education annually. It is provided in multiple formats, to include .csv format, and is available for download.

2. **Colorado Department of Education Financial Transparency Office**

The Colorado Department of Education also offers Financial Transparency, both in the form of

an online tool[8] as well as comprehensive quarterly financial statements[3]. These datasets are not as well structured as the performance data, and will need to be heavily processed in order to correctly align them with the performance dataset described above.

Evaluation Methods

While the body of literature is relatively small, the Common Sense Institute report[9] compiles a list of generalized information regarding performance and finances across Colorado schools. Additionally, historic[10] and more current research[11] seems to be mixed, with previous studies finding no correlation between funding and performance, but more focused studies finding a strong correlation. To date, there does not seem to be any Denver-focused study on this topic. However, the existing research can help measure and provide additional techniques to better understand the relationships that may exist.

Tools

To conduct this research, the following tools will be utilized:

1. Jupyter notebooks will be heavily utilized to run various data mining techniques and produce images to support the findings.
2. Python will be used extensively for data mining, as it is adept at formatting csvs to run advanced data mining techniques against the underlying data.
3. Numpy/pandas will be used in python, as these libraries contain multiple powerful features to simplify the data processing steps.
4. Javascript may be used to develop a web scraper in order to compile financial data for the research.
5. Looker studio will be utilized as well to produce high quality data visualizations.

Milestones

July 14: Complete financial data cleaning, preprocessing, and integration.

July 21: Complete all data cleaning, preprocess, and integration.

July 28: First draft of final report complete.

August 9: Final report complete.

REFERENCES

[1] Melanie Asmar, 2022. More Colorado schools and districts earn low state ratings. https://www.chalkbeat.org/colorado/2022/9/8/23343341/colorado-school-performance-framework-ratings-2022/.

[2] Alan Gottlieb, 2022. Performance plummets in Denver's schools. https://gazette.com/opinion/denver-columns/column-performance-plummets-in-denver-s-schools/article_54f8631a-b4dd-11ec-ac74-c36f0357135b.html

[3] Colorado Financial Services Reports, accessed 2024 for years 2014-2024. https://financialservices.dpsk12.org/o/financialservices/page/financial-transparency

[4] US News. 2024. Overview of Denver Public Schools. https://www.usnews.com/education/k12/colorado/districts/school-district-no-1-in-the-county-of-denver-and-state-of-c-112125

[5] Melanie Asmar, 2022. The $3,500-per-student difference between two Denver schools. https://www.chalkbeat.org/colorado/2022/4/28/23045997/denver-student-based-budgeting-smith-carson-elementary/.

[6] Ari Armstrong, 2023. Denver Public Schools celebrates lowering achievement bar. https://pagetwo.completecolorado.com/2023/09/26/armstrong-denver-public-schools-celebrates-lowering-the-bar/

[7] CMAS - Mathematics, English Language Arts, Science and Social Studies Data and Results, 2024. https://www.cde.state.co.us/assessment/cmas-dataandresults

[8] Financial Transparency for Colorado Schools, 2024. https://www.cde.state.co.us/schoolview/financialtransparency/homepage

[9] Common Sense Institute, 2022. Dollars and Data: A look at PK-12 Funding and Performance in Colorado. https://commonsenseinstituteco.org/dollars-and-data-a-look-at-pk-12-funding-and-performance-in-colorado/

[10] Lawrence O. Picus, 1995. Does Money Matter in Education? A Policymaker's Guide. https://nces.ed.gov/pubs97/web/97536-2.asp

[11] C. Kirabo Jackson, Rucker C. Johnson, Claudia Persico, 2015. The effects of school spending on Educational and Economic Outcomes: Evidence from School Finance Reforms. https://www.nber.org/system/files/working_papers/w20847/w20847.pdf