

High School Socioeconomic Neighborhood Status and CS1 Performance

Jennifer Alexandra Thompson
Virginia Tech
Blacksburg, VA, USA
alexthompson06@vt.edu

Margaret Ellis
Virginia Tech
Blacksburg, VA, USA
maellis1@vt.edu

Sara Hooshangi
Virginia Tech
Blacksburg, VA, USA
shoosh@vt.edu

ABSTRACT

CS1 student success rates are a longstanding issue in the computer science community. Indicators of performance prior to CS1 continue to be investigated in research, especially concerning prior programming and math courses taken at the high school level. This study aims to take a look at students' high school socioeconomic neighborhood status and determines whether there is a correlation to CS1 performance. Specifically, we examine the Area Deprivation Index (ADI) of the high schools that CS1 students attended and the passing rates in CS1 based on the socioeconomic status of these high schools. The goal is to compare the performance of students from socioeconomic disadvantaged high schools to students from advantaged high schools. In this research, we find that students from the top 15% high schools ADI percentile pass CS1 at a higher rate with a significant difference.

1 BACKGROUND/METHOD

In computer science education, there is a large body of literature concerning the connection of high school preparation and success in CS1. Studies have shown both programming and math courses from high school can be predictors of performance in CS1 [1]. Our research aims to investigate the impact of a high school's socioeconomic neighborhood on a student's CS1 performance.

Our method is to match each student's high school to a socioeconomic index measure with the goal of determining if there is a correlation between a student's high school socioeconomic status and the student's performance in CS1. The socioeconomic index measure used in this study is the Area Deprivation Index (ADI). ADI measures the socioeconomic status on a national percentile of a United States' census block group neighborhood on seventeen different socioeconomic factors [2, 3].

We use overall final CS1 grade as the performance metric to determine the passing rate of the students. We define the passing CS1 group as the group of students who successfully pass CS1 with a sufficient grade for enrollment in CS2. Alternatively, the non-passing group are those students who receive an insufficient grade for enrollment in CS2, or those who withdraw from the course prior to the end of the term.

2 DATA

Our data consists of seven years of CS1 student records from Virginia Tech, including their course information, final grades, demographic information, high school information, and their majors. From our data set, 4863 students attended a high school in the United States, and this group will be the population studied in our work. Overall, the passing rate of this population is 75.8%.

3 RESULTS

In our research, we separated our student population by ADI percentile splits in order to compare how students from a more advantaged high school region perform relative to students from a less advantaged high school region. We examined students with a high school ADI in the top 50%, top 15% and bottom 15% of the national percentiles to compare how these groups performed against the rest of the population. As shown in Table 1, we found that students from the top 15% of advantaged high school neighborhoods pass CS1 statistically more often than those from the bottom 85%. Additionally, students in the top 50% of advantaged high school neighborhoods passed CS1 statistically more frequently. From these results, we found a preliminary correlation between high school socioeconomic advantage and CS1 passing rates.

Table 1: Passing rates of CS1 students based on high school neighborhood ADI percentile splits

ADI Splits (top %, bottom %)	Passing Rate of Top Split	Passing Rate of Bottom Split	<i>p-value</i>
(50, 50)	76.32%	73.23%	0.038*
(15, 85)	77.57%	74.46%	0.006**
(85, 15)	75.77%	77.18%	0.344

The future work of this research will further explore passing rates between ADI groups, as well as the use an additional socioeconomic index, Social Deprivation Index (SDI). SDI is measured over census tracts and counties, which will help us compare our findings with a larger region surrounding the students' high schools. Additionally, we plan to investigate whether students' demographic information and majors have any correlation to CS1 performance.

REFERENCES

- [1] Susan Bergin and Ronan Reilly. 2005. Programming: Factors That Influence Success. *SIGCSE Bull.* 37, 1 (feb 2005), 411–415. <https://doi.org/10.1145/1047124.1047480>
- [2] Amy J.H. Kind and William R. Buckingham. 2018. Making Neighborhood Disadvantage Metrics Accessible — The Neighborhood Atlas. *New England Journal of Medicine* 378, 26 (2018), 2456–2458. <https://doi.org/10.1056/NEJMp1802313>
- [3] University of Wisconsin School of Medicine and Public Health. 2022. *Area Deprivation Index*. Retrieved October, 2022 from <https://www.neighborhoodatlas.medicine.wisc.edu/>

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
SIGCSE 2023, March 15–18, 2023, Toronto, ON, Canada
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9433-8/23/03.
<https://doi.org/10.1145/3545947.3576294>