

Diplomas and Disadvantage: Mapping U.S. Colleges and Deep Disadvantage Metrics

Annamarie Warnke

November 9, 2024

1 Project Description, Background, and Motivation

Colleges are unequally distributed across the United States; there is a high concentration of colleges along the East half of the country and the West Coast, but sparse options in the Midwest. There is even variation within regions and states. Though more students than ever before are traveling across the country to attend university thanks to air travel, over two-thirds of college students stay within 50 miles of home [1]. Therefore, having a college in close proximity can be a deciding factor for whether a student pursues higher education, and lack of nearby colleges can compound on other community disadvantages. One metric to understand the disadvantages that communities face is the Index of Deep Disadvantage, which combines health indicators, poverty metrics, and social mobility data [2]. To better understand the potential relationship between college proximity and community deep disadvantage, this project aims to map the distribution of colleges and universities across the United States in relation to metrics that signal deep disadvantage.

The idea for this project came from my internship this summer at MDRC, a nonprofit doing education and social policy research. Their Postsecondary Education policy area had a Lunch and Learn event earlier this year about community disadvantages, and they were interested in mapping their portfolio of partners on the Index of Deep Disadvantage. The code that I wrote during that internship is property of MDRC, so I want to rewrite the code in Python, apply it to subsets of U.S. colleges instead of just one list of schools, and add more user controls. This will allow me to showcase the work I did this summer without violating my agreement with MDRC.

2 Data Description

The U.S. colleges data comes from the Integrated Postsecondary Education Data System (IPEDS). It contains data for every postsecondary institution in the United States. I pulled data on institution names, state, longitude, latitude, highest degree level offered, public status, Historically Black College or University (HBCU) status, tribal college status, and number of students receiving each type of degree in 2023.

The pulled college data is in a comma-separated values (CSV) format. It contains 5,994 postsecondary institutions. Nearly half of these schools offer at least a four-year degree, and about one-third are public. Exactly 100 colleges are classified as HBCUs, and 35 colleges are classified as tribal colleges. There is no universal definition of community colleges, but generally they are public colleges that mostly offer certificates and associate's degrees. After calculating the percent of students who received a certificate or associate's degree at each school in 2023 and looking through the schools around various cutoff values, I determined the best definition of community colleges for this project would be public institutions where at least 90% of graduating students receive a certificate or associate's degree. With this definition, there are 1,227 community colleges in the dataset.

The deep disadvantage data comes from the University of Michigan's Understanding Communities of Deep Disadvantage page (see [2] for link), and it is downloadable as an Excel workbook. There are 45 variables, including basic demographic data and measures of community well-being. These variables are available for 3,120 counties (the U.S. has between 3,007 and 3,244 counties, depending on how you define a county [3]) and the 500 largest cities in the U.S.

3 Progress and Next Steps

Thus far, I have cleaned the data and created a map using plotly. The code for cleaning can be found in the `collectAndClean` function in `create_map.py`, and the code for creating a basic map can be found in the `createMap` function in `create_map.py`. `createMap` will continue to be updated as I add options for metrics other than the index of deep disadvantage rankings.

My next step is to use Dash to add more controls for the user such as toggling the subset of colleges, changing the disadvantage metric, and potentially filtering by state.

References

- [1] Nick Hillman. How Far Do Students Travel for College? *The Institute for College Access & Success*. October 2023. https://ticas.org/wp-content/uploads/2023/11/Hillman-Geography-of-Opportunity-Brief-2_2023.pdf
- [2] University of Michigan Poverty Solutions. *Understanding Communities of Deep Disadvantage*. <https://poverty.umich.edu/projects/understanding-communities-of-deep-disadvantage/>
- [3] Wikipedia. *List of United States Counties and County Equivalents*. https://en.wikipedia.org/wiki/List_of_United_States_counties_and_county_equivalents