Final Project Data Memo

Alexandra Chang

10/22/2020

Data Source

For my final project, I chose to do an exploratory data analysis on the Civil Rights Data Collection (CRDC) — a biennial study conducted by the Office for Civil Rights within the U.S. Department of Education — for the 2017-2018 school year. This survey collects "a variety of information including student enrollment and educational programs and services, most of which is disaggregated by race/ethnicity, sex, limited English proficiency, and disability."

For this project, I will be using a copy of the public-use data file that is published on the Department of Education's website, which can be found at the link below in the *Works Cited* section of this memo. This data file contains the entire 2017-2018 CRDC in CSV format. I specifically will be using the "Gifted and Talented" dataset.

Why This Data

I chose to analyze this data because I am passionate about social justice and education. As an Ethnic Studies major, educational inequities, such as the opportunity gap and the racialization of academic success, are issues of particular interest to me. The primary reason that I chose this data was to shed light on an issue that I am passionate about.

Within this large data file, I found a variable of particular interest to me. I would like to investigate how Gifted and Talented programs across the United States are divided by race and by sex. I hypothesize that students of color will be underrepresented in these "gifted" programs, while white students will be disproportionately represented in these programs due to racial biases. Additionally, I believe that there will be more boys than girls in Gifted and Talented programs. I am also considering investigating how these disparities differ across different regions of the United States.

Potential Data Issues

While this data is in a relatively tidy format, I will need to merge multiple datasets in order to explore the relationships that I want to investigate for this project. Additionally, I will need to clean the data. Currently, there are 97,632 entries — which each represent a school in the United States — but not all schools have Gifted and Talented programs. Thus, I will need to first filter out the schools without Gifted and Talented programs by referencing the attached codebook. I am excited to use the skills that I have learned in this class, however, on a challenging project for an issue that I am passionate about. I believe another shortcoming of this data is the perpetuation of the gender binary by collecting data by sex. Discrimination against gender minorities is a particularly pressing issue among youth in education that causes stress and, in turn, underperformance.

Additional Information

For my timeline, I am hoping to meet with Professor Coburn during class or during Office Hours this week to discuss my final project and receive feedback on my data memo. Additionally, I am hoping to complete a

rough draft of my final project by Thursday, November 5, so I can receive feedback on my visualizations during Office Hours.

Works Cited

Civil Rights Data Collection (CRDC) for the 2017-18 school year. (2020, October 15). Retrieved October 23, 2020, from https://www2.ed.gov/about/offices/list/ocr/docs/crdc-2017-18.html