

UP-STAT 2022

Introduction

Here, we worked with CRDC(Civil Rights Data Collection) data. The Civil Rights Data Collection (CRDC) of the United States Department of Education (ED) collects data on the nation's public school districts and elementary and secondary schools' commitment to offer equal educational opportunity. Since 1968, the CRDC has gathered statistics on public schools across the country, including student enrollment and educational programs and services data disaggregated by race/ethnicity, sex, limited English proficiency, and disability.

The CRDC for 2013-14 is a universe survey. All data supplied is self-reported and certified by school districts.

Detailed Analysis

We developed different Bar plots and Line plots to visualize the hidden patterns of the dataset. So we first read csv file using the `read_csv` function of pandas library and used encoding as `cp-1252`. We selected specific features from the dataset and then performed the preprocessing operations to make data more efficient and easy for further usage. We first plot the line graph in which we plot the number of students per district. So from the graph we can see that New York public schools have the highest number of students.

Then we displayed the number of male and female students based on different races who participated in the GED competition. So from the graph we can see that almost all the students from different races have participated in GED.

After that we plotted the bar plot regarding the high school equivalency credential received by students of different races. And from the graph we can see that students coming from different races have similar trends.

Then we used another file for further visualization and found some pattern for the dataset. So first we plotted the bar graph of students who enrolled for preschool. Therefore we see that white people have the highest number of students who enrolled whereas Native Hawaiian/Pacific Islander people have the smallest number of students who enrolled.

Then we can see that the number of the students who are limited english proficient is less for Native Hawaiian/Pacific Islander and the number of hispanic students is very high.

After this, we can see that the number of students who have disabilities served under IDEA is very less for Native Hawaiian/Pacific Islander and the number of white students is very high.