

In the year 2015, the Every Student Succeeds Act (ESSA) was signed into law, as a replacement for the No Child Left Behind (NCLB) act. ESSA reauthorized the Elementary and Secondary Education Act (ESEA) of 1965, which shifted accountability for school performance back to the individual states, rather than the federal standards outlined by the NCLB. After ESSA was passed into law, each state was then mandated to submit a plan to address specific components of the law, which entailed providing equitable access to education, as well as high standards and accountability, and to decrease the gap among student groups, including the most vulnerable populations. In response, Iowa implemented its own consolidated ESSA Plan, and a part of that accountability included having their schools' performances available publicly.

For this data analysis, data was collected from the Iowa Department of Education's website, specifically data containing information regarding the 'Iowa School Performance Profiles'. From that site, a user can select the Reporting Year, School District(s), Individual School(s), or filter by Specific Measure(s). The Reporting Year selected was the 2022-2023 Academic Year, all School Districts, and Individual Schools were selected and then the file was downloaded as a CSV file. Prior to importing the CSV file into a Jupyter Notebook and using a Python Kernel, the dataset was accessed using Microsoft Excel to have an initial view of the data and to confirm that the file was not corrupted.

During that initial view, it was clear that the set contained more than 50 columns and over 1000 rows with each row corresponding to a different school from the Iowa Public School System. In the 'Final Project - Data Exploration' Jupyter Notebook, the CSV file was imported and read as a DataFrame after Importing Pandas library. In that Data Exploration, it was discovered that there were a total of 71 columns and 1295 rows in the original dataset. There were a total of 16 columns that had missing values. The column names were intuitive and concise.

The next step of this project was cleaning the original data-frame, and the steps taken were documented in the 'Final Project - Data Cleaning' Jupyter Notebook. Several modules/packages were imported initially, including 'Pandas' once again, and 'Numpy'. First, the columns that had missing values and would not contain information valuable for data analysis were removed. Next, a value of '0' was added to columns containing an Index_Score for a particular ethnic group or subpopulation. They were left blank in the original documentation when the population size was not large enough to allow for a score to be assigned. Removing the rows with Missing Values would eliminate nearly all of the schools, however, which is why the value of '0' was added instead as a placeholder. Finally, the last two columns with missing values, 'Targeted_Subgroup' and 'School_Index_Source', had intuitive values added as placeholders. After missing values were addressed, the data-frame was consolidated to include only percentages of subgroups rather than counts (as if the count was too small, there would be the Index Score of '0' that was just assigned), and several other columns were also removed that were identifying information of the school (i.e. address, website, phone number, fax, etc.), ultimately bringing the total number of columns down to 40 and keeping the original number of rows at 1295, so all schools' data could be included in analysis.

There were four primary questions that were answered using Data Analysis. Two of those questions, 'How do schools with a greater percentage of low-socioeconomic status students perform compared to those with a smaller percentage (using FRL as a measure)?' and 'How do schools with a greater percentage of students having an IEP Plan perform compared to those with a smaller percentage?' both served similar purposes in that they both were used as a means of determining what factors may contribute to a School's Index Score. Through the use of a Violin Plot, after importing both Matplotlib and Seaborn packages, it was determined that both a high FRL measure and a higher student population with IEP's were correlated with schools that had lower Index Scores. In both cases, an initial analysis was done to determine the quartiles and median of the data, so that when grouped, there would be roughly a uniform distribution of the numbers of schools of each. However, it was also determined using the .corr() function (as well as being apparent using the respective plots) that higher rates of poverty had a greater impact on a school's Index Score. Therefore, in addition to allocating additional support and funding to those schools who have a lower School Index Score, finding additional ways to provide support for students & families from those districts who have a greater percentage of families with lower socioeconomic status, would also contribute to greater equity and education access, which is one of the primary missions of ESSA. For further study, an additional subgroup that should likely be considered would be the schools with a higher percentage of 'English Learners', as those students

would likely struggle in understanding material outside of their native language and thus impact the school's index score as well.

The next Data Question was, 'What is the distribution of school performance ratings in the state (i.e. "Exceptional", "Needs Improvement", etc.)?'. The purpose of this question was to have an overview of the performance of Schools in the Iowa Public School system. Using both a Bargraph and Pie Chart, it was determined that only 13.6% of schools went beyond the minimum threshold of achieving and exceeded the standard sought while about 15% of schools fell below the standard required and thus an Intervention was put into place. For future studies, it would be helpful to compare with at least a few other states to determine how Iowa's School Performance distribution compares. Despite the fact that ESSA brought the power and responsibility back to states to have their respective school systems achieve success, having a general model/baseline in place to determine how a state's school system is performing relative to other states' would be beneficial as well to see from a federal funding standpoint which states would benefit most from additional funding allocations. Finally, the last question asked was to determine, 'Is there variation across school levels (elementary, middle, high school) in how they perform on the school index when in the same school district?'. Using a Faceted Histogram along with a KDE curve and labeled Median line of the data, it was determined that despite the fact that there are significantly more Elementary Schools than Middle/High Schools in the state of Iowa, their performance, using School Standard Index as a measure, were fairly consistent. For further studies, taking a more granular look at specific Learning Measures and students' performances across different grade levels (i.e. 1st vs 2nd vs 3rd) while still keeping the school district consistent would also demonstrate if there are specific age groups/grade levels where students struggle to achieve the standard set and what could be implemented/changed in the curriculum to make those changes.

In conclusion, it was discovered that poverty and prevalence of students with learning disabilities still do contribute to 'gaps' in terms of student achievement in the state of Iowa during the 2022-2023 Academic Year. However, the prevalence of students with IEPs impact a school's performance less than does having a higher rate of poverty. Therefore, in addition to providing greater support through ESSA, it is also imperative that families in those districts receive additional relief at home as well, and it would be beneficial for state legislators to push for additional federal funding to assist families in those districts as well, as improving the schools alone will not resolve the underlying issues those families face. Using the data from the third question, as well as further studies, a minimum of 15% of schools should be 'targeted' for ESSA support, along with extra funding for those school districts that have greater percentages of 'at risk' families, identified using the information gleaned from the first two questions. School performance is not dictated by academic level, and is instead a greater product of some of the other factors cited above.