

School_District_Analysis

Project Overview

A City School System chief data scientist has given the following tasks to complete the school analysis for the district to finalize the upcoming budget funding to the schools:

- Create the district summary Dataframe.
- Create the school summary Dataframe.
- Find the top 5 and bottom 5 performing schools.
- Calculate the average math score received by students in each grade level at each school.
- Calculate the average reading score received by students in each grade level at each school.
- Calculate the school performance based on the spending per student.
- Calculate the school performance based on the size of the school.
- Calculate the school performance based on the type of school.

Resouces

- Data Source : schools_complete.csv , students_complete.csv
- Software : Python 3.7.6, Pandas, Jupyter Notebook

Summary

After creating the district summary and school summary data-frames, the analysis of the district schools shows that:

- The top 5 performing schools are charter type whereas the bottom 5 performing schools are district type. The “Thomas High School” is the second amongst the top 5 performing schools. One of the reasons for this is that the charter type schools have lesser number of students than the district type schools.
- The school performance based on the spending per student shows that the overall passing percentage is highest for the schools having spending per student < \$584. The schools whose spending per student is \$645-675 have the lowest overall percentage.
- The school performance based on the size of the school shows that the small and medium size schools have the high overall passing percentage than the large size schools.
- The school performance based on the type of school shows that the charter type of schools has the higher overall passing percentage than the district type schools.

Challenge Overview

The grades of the ninth graders at Thomas High School have been changed. The Challenge was to:

- Correct the students' names so there are no professional prefixes or suffixes.
- Replace the reading and math scores for ninth graders at Thomas High School with **NaN**.

- Merge the clean student data with the school dataset.
- Keep all other data associated with the ninth-grade students and Thomas High School intact.
- Recreate the district and school summary DataFrames.
- Recalculate the top 5 and bottom 5 performing schools.
- Recalculate the average math score received by students in each grade level at each school.
- Recalculate the average reading score received by students in each grade level at each school.
- Recalculate the school performance based on the spending per student.
- Recalculate the school performance based on the size of the school.
- Recalculate the school performance based on the type of school.

Challenge Summary

After recreating the district summary, school summary data-frames and all the other respective dataframes, the analysis of the district schools shows that:

- The new district summary shows that there is around **1% (fall)** difference in math, reading and overall percentage by changing the math and reading scores of the 9th grade students of the Thomas High School by **NaN**. But there is no effect on the average math and reading scores.
- The new school summary shows that the position of Thomas High School is **eighth** as compared to the **second** position in the top performing schools previously. The top 5 performing schools are charter type whereas the bottom 5 performing schools are district type. Also, the passing percentage for math, reading and overall of the Thomas High School has reduced drastically from **around 90% to around 60%**. But there is not much difference in the average scores for math and reading.
- There is **no effect** on the average scores for math and reading by students in each grade level at each school because of the above changes.
- Based on the school spending there is about **7% fall** in the percentage for the math, reading and overall in the spending range **\$630-644** because the per student budget of Thomas High School (**\$638**) comes in this range.
- Based on the school size there is no difference on the average math and reading scores. But there is about **6% fall** in the percentage for the math, reading and overall as compared to the previous dataframe for medium (**1000-2000 students**) size schools. This is due to the fact that the Thomas High School has **1635 students** which comes in the category of medium size schools.
- There is no change in the average math and reading scores based on the school type. But since Thomas High School is **charter type** school, there is a **reduction of 4%** in the math and reading passing percentage. The overall percentage **falls by 3 %** from the previous one.