PREMIUM PROJECT

Exploring NYC Public School Test Result Scores

Use data manipulation and summary statistics to analyze test scores across New York City's public schools!

Project Description

Every year, school test results impact the college admissions fate of millions of students.

In this project, you will use standardized test performance data from NYC's public schools to identify the schools with top math results, look at how performance varies by borough, and find the city's top ten performing schools!

Every year, American high school students take SATs, which are standardized tests intended to measure literacy, numeracy, and writing skills. There are three sections - reading, math, and writing, each with a **maximum score of 800 points**. These tests are extremely important for students and colleges, as they play a pivotal role in the admissions process.

Analyzing the performance of schools is important for a variety of stakeholders, including policy and education professionals, researchers, government, and even parents considering which school their children should attend.

You have been provided with a dataset called `schools.csv`, which is previewed below.

You have been tasked with answering three key questions about New York City (NYC) public school SAT performance.

Project Instructions

Which NYC schools have the best math results?

- The best math results are at least 80% of the ***maximum possible score of 800*** for math.
- Save your results in a pandas DataFrame called `best_math_schools`, including `"school_name"` and `"average_math"` columns, sorted by `"average_math"` in descending order.

What are the top 10 performing schools based on the combined SAT scores?

- Save your results as a pandas DataFrame called `top_10_schools` containing the `"school_name"` and a new column named `"total_SAT"`, with results ordered by `"total_SAT"` in descending order.

Which single borough has the largest standard deviation in the combined SAT score?

- Save your results as a pandas DataFrame called `largest_std_dev`.
- The DataFrame should contain one row, with:
    - `"borough"` - the name of the NYC borough with the largest standard deviation of `"total_SAT"`.
    - `"num_schools"` - the number of schools in the borough.
    - `"average_SAT"` - the mean of `"total_SAT"`.
    - `"std_SAT"` - the standard deviation of `"total_SAT"`.
- Round all numeric values to two decimal places.