Project (SQL): Analyzing NYC Public School Test Result Scores

by Jenn Lai

Introduction

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.

In this notebook, we will take a look at data on SATs across public schools in New York City.

Our database contains a single table:

schools

column	type	description
school_name	varchar	Name of school
borough	varchar	Borough that the school is located in
building_code	varchar	Code for the building
average_math	int	Average math score for SATs
average_reading	int	Average reading score for SATs
average_writing	int	Average writing score for SATs
percent_tested	numeric	Percentage of students completing SATs

1. Inspecting the data

```sql
SELECT
  school_name,
  MAX(average_writing) AS max_writing
FROM schools
GROUP BY school_name
ORDER BY max_writing DESC
LIMIT 1;
```

Result :

school_name	borough	building_code	average_math	average_reading	average_writing	percent_tested
New Explorations into Science, Technology and Math High School	Manhattan	M022	657	601	601	None
Essex Street Academy	Manhattan	M445	395	411	387	78.9
Lower Manhattan Arts Academy	Manhattan	M445	418	428	415	65.1
High School for Dual Language and Asian Studies	Manhattan	M445	613	453	463	95.9
Henry Street School for International Studies	Manhattan	M056	410	406	381	59.7
Bard High School Early College	Manhattan	M097	634	641	639	70.8

*It looks like the first school in our database had no data in the percent_tested column!

2. Finding missing values

Count rows with percent_tested missing and total number of schools

```sql
SELECT (SELECT
          COUNT(*) AS num_tested_missing
        FROM schools
        WHERE percent_tested IS NULL),
        COUNT(DISTINCT school_name) AS num_schools
FROM schools;
```

Result :

num_tested_missing	num_schools
20	375

3. Schools by building code

Count the number of unique building_code values

```sql
SELECT
  COUNT(DISTINCT building_code) AS num_school_buildings
FROM schools;
```

Result :

num_school_buildings
233

*Out of 375 schools, only 233 (62%) have a unique `building_code`

4. Best schools for math

let's find all schools with an average math score of at least 80% (out of 800)

```sql
SELECT
  school_name,
  average_math
FROM schools
WHERE average_math >= 640
ORDER BY average_math DESC;
```

Result :

school_name	average_math
Stuyvesant High School	754
Bronx High School of Science	714
Staten Island Technical High School	711
Queens High School for the Sciences at York College	701
High School for Mathematics, Science, and Engineering at City College	683
Brooklyn Technical High School	682
Townsend Harris High School	680
High School of American Studies at Lehman College	669
New Explorations into Science, Technology and Math High School	657
Eleanor Roosevelt High School	641

* there are only ten public schools in New York City with an average math score of at least 640

5. Lowest reading score

```sql
SELECT
  MIN(average_reading) AS lowest_reading
FROM schools;
```

Result :

lowest_reading
302

6. Best writing school

```sql
SELECT
  school_name,
  MAX(average_writing) AS max_writing
FROM schools
GROUP BY school_name
ORDER BY max_writing DESC
LIMIT 1;
```

Result :

school_name	max_writing
Stuyvesant High School	693

7. Top 10 schools

```sql
SELECT
  school_name,
  average_math + average_reading + average_writing AS average_sat
FROM schools
GROUP BY school_name
ORDER BY average_sat DESC
LIMIT 10;
```

Result :

school_name	average_sat
Stuyvesant High School	2144
Staten Island Technical High School	2041
Bronx High School of Science	2041
High School of American Studies at Lehman College	2013
Townsend Harris High School	1981
Queens High School for the Sciences at York College	1947
Bard High School Early College	1914
Brooklyn Technical High School	1896
Eleanor Roosevelt High School	1889
High School for Mathematics, Science, and Engineering at City College	1889

8. Ranking boroughs

Calculates the number of schools and the average SAT score per borough

```sql
SELECT
  borough,
  COUNT(school_name) AS num_schools,
  SUM(average_math + average_reading + average_writing) / COUNT(school_name) AS average_borough_sat
FROM schools
GROUP BY borough
ORDER BY average_borough_sat DESC;
```

Result :

borough	num_schools	average_borough_sat
Staten Island	10	1439
Queens	69	1345
Manhattan	89	1340
Brooklyn	109	1230
Bronx	98	1202

9. Brooklyn numbers

Let's focus on Brooklyn, which has 109 schools. We wish to find the top five schools for math performance.

```
SELECT
  school_name,
  average_math
FROM schools
WHERE borough = 'Brooklyn'
GROUP BY school_name
ORDER BY average_math DESC
LIMIT 5;
```

Result :

school_name	average_math
Brooklyn Technical High School	682
Brooklyn Latin School	625
Leon M. Goldstein High School for the Sciences	563
Millennium Brooklyn High School	553
Midwood High School	550