SQL Analysis of SAT Data from New York City's Public Schools

This analysis examines the SAT performance of students in New York City's public schools, providing insights into trends, achievement gaps and factors influencing test outcomes.

MU by Mahesh Ushir

Introduction

In the following task, we will delve into SAT data from New York City's public schools.

This information encompasses reading, math and writing test sections, each with a maximum score of 800 points. Given the pivotal role of SATs in the college admissions process, the comprehensive analysis of schools' performance is crucial for policy and education professionals, government bodies, researchers and parents making decisions about their children's educational paths.

Our database contains single table :

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

School Data

```
QUERY : SELECT *
        FROM SCHOOLS;
```

SAMPLE

school_name	borough	building_code	average_math	average_reading	average_writing	percent_tested
New Explorations into Science, Technology and ...	Manhattan	M022	657	601	601	
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
Urban Assembly Academy of Government and Law	Manhattan	M445	389	395	381	80.8
Marta Valle High School	Manhattan	M025	438	413	394	35.6
University Neighborhood High School	Manhattan	M446	437	355	352	69.9
New Design High School	Manhattan	M445	381	396	372	73.7
Pace High School	Manhattan	M131	430	435	427	87.8
High School for Health Professions and Human S...	Manhattan	M475	452	445	430	86.9
High School for Language and Diplomacy	Manhattan	M460	446	433	411	70.2

Find the missing values

It appears that the first school in our database did not provide data in the **percent_tested** column, signifying that the percentage of students tested was not reported. We need to identify how many schools have missing data in this column. This will enable us to gauge the extent of the missing data issue in New York.

Furthermore, we will calculate the total number of schools in the database to gain a comprehensive understanding of the dataset's scope.

QUERY :

```
SELECT
    COUNT(school_name) - COUNT(percent_tested) AS num_tested_missing,
    COUNT(school_name) AS num_schools
FROM schools;
```

OUTPUT :

	num_tested_missing	num_schools
▶	20	375

Schools by building code

Out of all the rows in the database, 20 schools are identified to have missing data for the **percent_tested** column, representing approximately 5% of the total dataset entries.

Another observation made from displaying the initial ten rows of the database is the presence of repeated building_code values, indicating that multiple schools might be situated at the same location.

To further explore this, we aim to determine the count of unique school locations stored in our database.

```
QUERY : SELECT

            COUNT(DISTINCT building_code) AS num_school_buildings

        FROM schools;
```

OUTPUT :

	num_school_buildings
▶	233

Best schools for math

Out of the total 375 schools, only 233 (62%) have a unique **building_code** associated with them. Moving forward to our analysis of school performance, we will approach the evaluation of each school's performance individually, rather than consolidating them by **building_code**.

To commence our analysis, we will identify all schools with an average math score of at least 80% (out of 800).

QUERY : SELECT

```
        school_name,

        average_math

        FROM schools

        WHERE average_math >= 640

        ORDER BY average_math DESC;
```

OUTPUT :

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 Col...	701
High School for Mathematics, Science, and Engi...	683
Brooklyn Technical High School	682
Townsend Harris High School	680
High School of American Studies at Lehman Coll...	669
New Explorations into Science, Technology and ...	657
Eleanor Roosevelt High School	641

Lowest reading score

The analysis reveals that there are merely ten public schools in New York City with an average math score reaching or exceeding 640 out of 800. Shifting focus to the opposite end of the scale, we will extract the singular lowest reading score from the dataset.

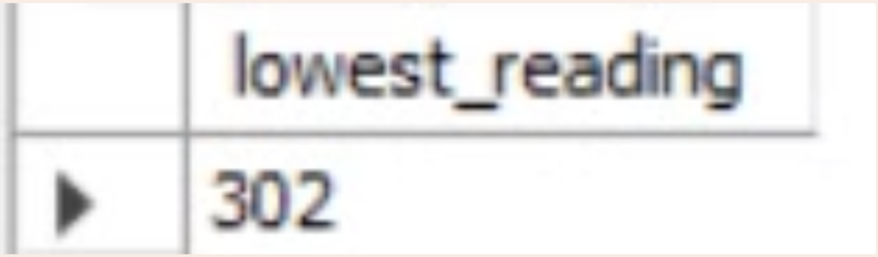
This approach will solely involve identifying the lowest reading score without associating it with any specific school to maintain anonymity.

QUERY : SELECT

```
        MIN(average_reading) AS lowest_reading

    FROM schools;
```

OUTPUT :



	lowest_reading
▶	302

Best writing school

The data indicates that the lowest average reading score across schools in New York City falls below 40% of the total available points.

Moving forward, we will now identify the school with the highest average writing score.

QUERY : SELECT

```
        school_name,

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

OUTPUT :

	school_name	max_writing
▶	Stuyvesant High School	693

Top 10 schools

Stuyvesant High School achieved an extraordinary average writing score of 693, which is quite remarkable. This exceptional performance in writing was complemented by Stuyvesant High School's top math score, further solidifying its reputation as one of the premier schools in New York.

To identify other schools that excel across the board, we will assess their performance in reading, writing, and math to pinpoint additional standout institutions.

QUERY :

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

OUTPUT :

	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

Ranking boroughs

Let's construct a query to analyze the performance of schools based on New York City boroughs.

This query will calculate the total number of schools and the average SAT scores for each borough.

QUERY :

```
SELECT

    borough,

    COUNT(school_name) AS num_schools,

    (SUM(average_math) + SUM(average_reading) + SUM(average_writing))/COUNT(school_name) AS average_borough_sat

    FROM schools

    GROUP BY borough

    ORDER BY average_borough_sat DESC;
```

OUTPUT :

	borough	num_schools	average_borough_sat
▶	Staten Island	10	1439.0000
	Queens	69	1345.4783
	Manhattan	89	1340.1348
	Brooklyn	109	1230.2569
	Bronx	98	1202.7245

Brooklyn numbers

Upon analysis, it appears that, on average, schools in Staten Island achieve higher scores across all three categories. Nonetheless, the number of schools in Staten Island is limited, with only 10 schools, in contrast to an average of 91 schools in the other four boroughs.

For our final database query, we will narrow our focus to Brooklyn, which comprises 109 schools. Our objective is to identify the five schools with the highest performance in math.

QUERY : SELECT

```
        school_name,

        average_math
        FROM schools
        WHERE borough = 'Brooklyn'
        ORDER BY average_math DESC LIMIT 5;
```

OUTPUT :

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

Observations

1. **Average Scores**:
    - The average scores for Math, Reading, and Writing are around 433, 424, and 418 respectively. These scores provide an indication of the overall performance of students in these subjects.
2. **Percentage of Students Tested**:
    - The average percentage of students tested is approximately 65%, with a minimum of 18.5% and a maximum of 100%. This indicates the extent of student participation in the assessments.
3. **Score Variability**:
    - The standard deviations for the scores are relatively small, indicating that the scores are clustered around the mean. This suggests that there may not be significant variability in scores among the students.
4. **Data Completeness**:
    - It is worth noting that there are 20 missing values in the '**percent_tested**' column, as the count for this column is 355 instead of 375.
5. **Comparison Between Subjects**:
    - The average Math score is slightly higher than the average Reading and Writing scores. This could indicate that students, on average, perform slightly better in Math compared to Reading and Writing.

Recommendation based on top 10 schools

Here is the list of the top 10 schools based on their average total scores:

1. Stuyvesant High School
2. Bronx High School of Science
3. Staten Island Technical High School
4. High School of American Studies at Lehman College
5. Townsend Harris High School
6. Queens High School for the Sciences at York College
7. Bard High School Early College
8. Brooklyn Technical High School
9. Eleanor Roosevelt High School
10. High School for Mathematics, Science, and Engineering at City College

Recommendations:

1. **Further Success**: These top-performing schools should continue to focus on maintaining high academic standards and providing quality education.
2. **Challenges**: Schools facing challenges should identify areas for improvement and implement targeted strategies to enhance student performance.
3. **Best Practices Sharing**: Schools with high performance can share best practices with others to promote academic excellence across the educational system.
4. **Continuous Improvement**: All schools should strive for continuous improvement in teaching practices, student support and overall educational quality.
5. **Recognition**: Recognize and celebrate the achievements of both students and educators to foster a culture of learning and excellence.

Recommendation based on bottom 10 schools

Here is a list of the bottom 10 schools based on their average total scores:

1. Pan American International High School at Monroe
2. Multicultural High School
3. International High School at Prospect Heights
4. Pan American International High School
5. Kingsbridge International High School
6. International Community High School
7. Bronx International High School
8. Manhattan Academy for Arts and Language
9. W. H. Maxwell Career and Technical Education High School
10. High School of Language and Innovation

These schools may benefit from the following suggestions for improvement:

1. **Identify Weak Areas**: Schools should analyze student performance data to identify weak subject areas and topics in which students may need additional support.
2. **Enhanced Support**: Provide targeted support and resources for students who are struggling academically, such as tutoring programs or study groups.
3. **Professional Development**: Offer professional development opportunities for teachers to enhance their instructional techniques and support student learning effectively.
4. **Parental Involvement**: Encourage parental involvement in the education process to create a supportive learning environment at home.
5. **Curriculum Review**: Evaluate the curriculum to ensure alignment with academic standards and address any gaps in content coverage.
6. **Student Engagement**: Implement strategies to increase student engagement in the learning process, such as interactive lessons and project-based learning.

Strategies for Improving SAT Preparedness

1

Early Intervention

Introducing SAT prep programs in middle school can help students build a strong academic foundation and develop test-taking skills.

2

Targeted Curriculum

Aligning high school coursework with the content and format of the SAT can better prepare students for the exam.

3

Expanding Access

Providing free or low-cost SAT prep resources, including practice tests and tutoring, can make test preparation more accessible.

Implications for College Admissions and Equity

1 Admissions Considerations

As colleges and universities increasingly emphasize holistic admissions, SAT scores may carry less weight, helping to level the playing field.

2 Diversity and Inclusion

Addressing achievement gaps and providing equitable access to test preparation can foster greater diversity and representation in higher education.

3 Rethinking Assessments

The limitations of standardized tests in measuring student potential has sparked discussions around alternative admissions criteria that better capture individual strengths.

Role of Tutoring and Test Prep Programs

Targeted Skill-Building

Tutoring and test prep programs can help students develop the specific skills and strategies needed to excel on the SAT.

Improved Confidence

Regular practice and guidance from experienced instructors can boost students' test-taking confidence and reduce anxiety.

Measurable Gains

Studies show that students who participate in comprehensive test prep programs can see significant improvements in their SAT scores.

Recommendations for Policymakers and Educators

Increase Funding Equity	Ensure high-need schools receive adequate resources to support student achievement
Expand Access to Test Prep	Provide free or low-cost SAT preparation programs, including tutoring and practice tests
Strengthen Early Interventions	Introduce SAT readiness initiatives in middle schools to build a strong academic foundation
Reevaluate Admissions Criteria	Consider alternative assessment methods that better capture student potential and promote equity