

 New York City schoolbus

Photo by [Jannis Lucas \(https://unsplash.com/@jannis\\_lucas\)](https://unsplash.com/@jannis_lucas) on [Unsplash \(https://unsplash.com\)](https://unsplash.com).

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 tasked with answering three key questions about New York City (NYC) public school SAT performance:

## Which schools produce the highest math scores?

- Specifically, which schools have an average math SAT score of at least 80%?
- Save the results as a pandas DataFrame called `best_math_schools`.

## Who are the top 10 schools based on average results across reading, math, and writing?

- Save the results as a pandas DataFrame called `top_10_schools`.

## Which NYC borough has the largest standard deviation for SAT results?

- Save the results as a pandas DataFrame called `largest_std_dev`.

```
In [22]: # Start coding here...
import pandas as pd
data = pd.read_csv("schools.csv")
data.head()
```

	school_name	borough	building_code	average_math	average_reading	average_writing	percent_tested
0	New Explorations into Science, Technology and ...	Manhattan	M022	657	601	601	NaN
1	Essex Street Academy	Manhattan	M445	395	411	387	78.9
2	Lower Manhattan Arts Academy	Manhattan	M445	418	428	415	65.1
3	High School for Dual Language and Asian Studies	Manhattan	M445	613	453	463	95.9
4	Henry Street School for International Studies	Manhattan	M056	410	406	381	59.7

```
In [23]: df = pd.DataFrame(data, columns = ["school_name", "average_math"])
```

```
In [24]: best_math_schools = df[df["average_math"] >= 0.8 * max(df.average_math)].sort_values("average_math", ascend:
```

```
In [25]: best_math_schools
```

	<b>school_name</b>	<b>average_math</b>
88	Stuyvesant High School	754
170	Bronx High School of Science	714
93	Staten Island Technical High School	711
365	Queens High School for the Sciences at York Co...	701
68	High School for Mathematics, Science, and Engi...	683
280	Brooklyn Technical High School	682
333	Townsend Harris High School	680
174	High School of American Studies at Lehman College	669
0	New Explorations into Science, Technology and ...	657
45	Eleanor Roosevelt High School	641
5	Bard High School Early College	634
213	Baccalaureate School for Global Education	633
204	Bard High School Early College Queens	631
237	Brooklyn Latin School	625
3	High School for Dual Language and Asian Studies	613

```
In [26]: data["total_SAT"] = data["average_math"] + data["average_reading"] + data["average_writing"]
```

```
In [33]: top_10_schools = pd.DataFrame(data, columns = ["school_name", "total_SAT"]).sort_values("total_SAT", ascend:
```

```
In [35]: boroughs = data.groupby("borough")["total_SAT"].agg(["count", "mean", "std"]).round(2)
        largest_std_dev = boroughs[boroughs["std"] == boroughs["std"].max()].reset_index()
        largest_std_dev.rename(columns={"count": "num_schools", "mean": "average_SAT", "std": "std_SAT"}, inplace=True)
```

```
In [36]: boroughs
```

	count	mean	std
borough			
Bronx	98	1202.72	150.39
Brooklyn	109	1230.26	154.87
Manhattan	89	1340.13	230.29
Queens	69	1345.48	195.25
Staten Island	10	1439.00	222.30

```
In [37]: largest_std_dev
```

	borough	num_schools	average_SAT	std_SAT
0	Manhattan	89	1340.13	230.29