PROJECT: EXPLORING NYC PUBLIC SCHOOL TEST RESULT SCORES

Photo by Jannis Lucas on Unsplash.

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.

```
# Re-run this cell
import pandas as pd

# Read in the data
schools = pd.read_csv("schools.csv")

# Preview the data
schools.head()


# Start coding here...
# Add as many cells as you like...
best_math_schools = schools[ schools["average_math"] >= 0.8*800 ]
best_math_schools =
best_math_schools[["school_name","average_math"]].sort_values("average_math",ascendi
ng = False)
print(best_math_schools)
```

```
                                        school_name  average_math
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
```

```
#top 10 performing schools based on the sat combined scores:
schools["total_SAT"] = schools["average_math"] + schools["average_reading"] +
schools["average_writing"]
top_schools = schools[["school_name","total_SAT"]].sort_values("total_SAT",ascending
= False)
top_10_schools = top_schools.head(10)
```

```
#Which single borough has the largest standard deviation in the combined SAT score
std_SAT = schools.groupby("borough")["total_SAT"].std()
avg_SAT = schools.groupby("borough")["total_SAT"].mean()
num_schools = schools.groupby("borough")["school_name"].count()

df = pd.DataFrame({
    "num_schools": num_schools,
    "average_SAT": avg_SAT,
    "std_SAT": std_SAT,
    })
largest_std_dev = df.reset_index().sort_values("std_SAT",ascending =
False).round(2).head(1)

print(df)
```

```
               num_schools  average_SAT     std_SAT
borough
Bronx                   98  1202.724490  150.393901
Brooklyn               109  1230.256881  154.868427
Manhattan               89  1340.134831  230.294140
Queens                  69  1345.478261  195.252899
Staten Island           10  1439.000000  222.303596
```