

Applied Data Science Capstone:

Identify New York Public High Schools with the Highest SAT Test Scores

Table of contents

- Introduction: Business Problem
- Data
- Methodology
- Analysis
- Results and Discussion
- Conclusion

Introduction: Business Problem

The purpose of this project is to identify which public high school's within the five boroughs that make up the New York City area have the highest overall SAT scores, while isolating the public high school's that received the highest score for each section (math, reading, writing) of the SAT test and through k-mean clustering locate which borough has the highest performing public high schools.

This report will be targeted to people who are looking to relocate or trying to apply to New York City area public high schools with the highest performing SAT scores.

Data

Concentrating on the stated business problem, factors that will influence the location of high schools are:

The SAT score averages that have been provided by the College Board in each of the five boroughs creating the New York City area during the years 2014-2015.

The high school data that has been compiled and published by the New York City Department of Education.

Following data sources will be needed to extract/generate the required information:

Part 1: Preprocessing the average SAT 2014-2015 school year SAT test scores data from Kaggle. Crafting a dataframe illustrating the average SAT scores for every accredited high school in the New York City, five borough area.

Part 2: Generate a map of the public high schools with the highest average total SAT test scores and three main sections of the test within the five boroughs area of New York City via folium.

Part 3: Identify the closest vendors to the public high schools with the highest SAT test scores with the 5 borough New York City area, via FourSquare API and geocoding.

Part 1: Preprocessing real world data from Kaggle illustrating the average SAT scores for every accredited high school in the New York City, five borough area.

About this file: New York City SAT test scores for the 2014-2015 school year. borough: Common name for individual cities that together make up the New York City area.

Data set URL: <https://www.kaggle.com/nycopendata/high-schools> (<https://www.kaggle.com/nycopendata/high-schools>)

```
In [1]: import requests # library to handle requests
import pandas as pd # library for data analysis
import numpy as np # library to handle data in a vectorized manner
import random # library for random number generation
from bs4 import BeautifulSoup # library for web scrapping

#!conda install -c conda-forge geocoder --yes
import geocoder

#!conda install -c conda-forge geopy --yes
from geopy.geocoders import Nominatim # module to convert an address into latitude and longitude values

# libraries for displaying images
from IPython.display import Image
from IPython.core.display import HTML

# transforming json file into a pandas dataframe library
from pandas.io.json import json_normalize

#!conda install -c conda-forge folium=0.5.0 --yes
import folium # plotting library

print('done')
```

done

```
In [2]: # Foursquare API
CLIENT_ID = 'R01LINGO2WC45KLRLKT3ZHU2QENAO2IPRK2N2ELOHRNK4P3K' # your Foursquare ID
CLIENT_SECRET = '4JT1TWRMXMPLX5IOKNBAFU3L3ARXK4D5JJDPFK1CLRZM2ZVW' # your Foursquare Secret

VERSION = '20180604'
LIMIT = 30

print('Your credentials:')
print('CLIENT_ID: ' + CLIENT_ID)
print('CLIENT_SECRET: ' + CLIENT_SECRET)
```

Your credentials:

CLIENT_ID: R01LINGO2WC45KLRLKT3ZHU2QENAO2IPRK2N2ELOHRNK4P3K

CLIENT_SECRET: 4JT1TWRMXMPLX5IOKNBAFU3L3ARXK4D5JJDPFK1CLRZM2ZVW

```
In [3]: # Read in the data
df = pd.read_csv("NY_SAT_Scores.csv")
```

```
In [4]: # View the top rows of the dataset
df.head()
```

Out[4]:

	School ID	School Name	Borough	Building Code	Street Address	City	State	Zip Code	Latitude	Longitude
0	02M260	Clinton School Writers and Artists	Manhattan	M933	425 West 33rd Street	Manhattan	NY	10001	40.75321	-73.99
1	06M211	Inwood Early College for Health and Informatio...	Manhattan	M052	650 Academy Street	Manhattan	NY	10002	40.86605	-73.91
2	01M539	New Explorations into Science, Technology and ...	Manhattan	M022	111 Columbia Street	Manhattan	NY	10002	40.71873	-73.91
3	02M294	Essex Street Academy	Manhattan	M445	350 Grand Street	Manhattan	NY	10002	40.71687	-73.91
4	02M308	Lower Manhattan Arts Academy	Manhattan	M445	350 Grand Street	Manhattan	NY	10002	40.71687	-73.91

5 rows × 22 columns



```
In [5]: # Dropping columns that are not relevant to the problem
df.drop(['Start Time', 'End Time', 'Phone Number', 'Percent White', 'Percent Black', 'Percent Hispanic', 'Percent Asian', 'Percent Tested'], axis = 1, inplace = True)
df.head()
```

Out[5]:

	School ID	School Name	Borough	Building Code	Street Address	City	State	Zip Code	Latitude	Longitude
0	02M260	Clinton School Writers and Artists	Manhattan	M933	425 West 33rd Street	Manhattan	NY	10001	40.75321	-73.99642
1	06M211	Inwood Early College for Health and Information Science	Manhattan	M052	650 Academy Street	Manhattan	NY	10002	40.86605	-73.99642
2	01M539	New Explorations into Science, Technology and ...	Manhattan	M022	111 Columbia Street	Manhattan	NY	10002	40.71873	-73.99642
3	02M294	Essex Street Academy	Manhattan	M445	350 Grand Street	Manhattan	NY	10002	40.71687	-73.99642
4	02M308	Lower Manhattan Arts Academy	Manhattan	M445	350 Grand Street	Manhattan	NY	10002	40.71687	-73.99642



```
In [6]: # re-name column
df.rename(columns={'School Name': 'School'}, inplace=True)
```

```
In [7]: # insert column with average total SAT scores
df.insert(14, "Average Score (SAT Total)", "0", allow_duplicates=False)
df['Average Score (SAT Total)'] = df.iloc[:, -4:].sum(axis=1)
```

In [8]: `df.head()`

Out[8]:

	School ID	School	Borough	Building Code	Street Address	City	State	Zip Code	Latitude	Longi
0	02M260	Clinton School Writers and Artists	Manhattan	M933	425 West 33rd Street	Manhattan	NY	10001	40.75321	-73.9
1	06M211	Inwood Early College for Health and Informatio...	Manhattan	M052	650 Academy Street	Manhattan	NY	10002	40.86605	-73.9
2	01M539	New Explorations into Science, Technology and ...	Manhattan	M022	111 Columbia Street	Manhattan	NY	10002	40.71873	-73.9
3	02M294	Essex Street Academy	Manhattan	M445	350 Grand Street	Manhattan	NY	10002	40.71687	-73.9
4	02M308	Lower Manhattan Arts Academy	Manhattan	M445	350 Grand Street	Manhattan	NY	10002	40.71687	-73.9

Methodology

The methodology in this project consists of two sections, exploratory data analysis and machine learning modeling.

Exploratory Data Analysis:

Visualize the areas of the five borough that posses the highest public high schools with the average total SAT test scores. Identify boroughs that have the highest public high schools with within each section of the SAT test.

Modelling:

To assist families/students locate public high schools with the highest average SAT test scores. Then identifying the closest vendors to the public high schools FourSquare data. The vendor clustering will enable families/students to be aware of the closest local vendors for further evaluation of not only the public high school, but surrounding area as well.

Part II - Generate a map of the public high schools with the highest average total SAT test scores and three main sections of the test within the five boroughs area of New York City via folium.

```
In [9]: # import libraries

# use the inline backend to generate the plots within the browser
%matplotlib inline

import matplotlib as mpl
import matplotlib.pyplot as plt

mpl.style.use('ggplot') # optional: for ggplot-like style

# check for latest version of Matplotlib
print('Matplotlib version: ', mpl.__version__) # >= 2.0.0

Matplotlib version: 3.1.1
```

```
In [10]: # Matplotlib and associated plotting modules
import numpy as np

import matplotlib.cm as cm
import matplotlib.colors as colors

import pandas as pd
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json

import requests
from pandas.io.json import json_normalize

from sklearn.cluster import KMeans

print('done')
```

done

```
In [11]: place = 'New York City, New York'

geolocator=Nominatim(user_agent='can_explorer')
location = geolocator.geocode(place)
latitude=location.latitude
longitude=location.longitude
print('The coordinates of New York City are {}, {}'.format(latitude, longitude))
```

The coordinates of New York City are 40.7127281, -74.0060152.

In [12]: *## Sort the highest average test scores in each section to view the location within the 5 boroughs*

```
dftotal= df.sort_values(['Average Score (SAT Total)'], ascending = False, axis = 0, inplace = True )
```

```
dftop10total= df.sort_values(['Average Score (SAT Total)'], ascending = False, axis = 0, inplace = True )
dftop10total = df.head(10)
dftop10total
```

Out[12]:

	School ID	School	Borough	Building Code	Street Address	City	State	Zip Code	Latitude	L
105	02M475	Stuyvesant High School	Manhattan	M477	345 Chambers Street	Manhattan	NY	10282	40.71775	-
110	31R605	Staten Island Technical High School	Staten Island	R440	485 Clawson Street	Staten Island	NY	10306	40.56791	-
203	10X445	Bronx High School of Science	Bronx	X445	75 West 205th Street	Bronx	NY	10468	40.87706	-
208	10X696	High School of American Studies at Lehman College	Bronx	X905	2925 Goulden Avenue	Bronx	NY	10468	40.87126	-
385	25Q525	Townsend Harris High School	Queens	Q515	149-11 Melbourne Avenue	Flushing	NY	11367	40.73441	-
424	28Q687	Queens High School for the Sciences at York Co...	Queens	Q774	94-50 159th Street	Jamaica	NY	11433	40.70100	-
7	01M696	Bard High School Early College	Manhattan	M097	525 East Houston Street	Manhattan	NY	10002	40.71896	-
326	13K430	Brooklyn Technical High School	Brooklyn	K430	29 Ft Greene Place	Brooklyn	NY	11217	40.68811	-
55	02M416	Eleanor Roosevelt High School	Manhattan	M855	411 East 76th Street	Manhattan	NY	10021	40.77012	-
80	05M692	High School for Mathematics, Science, and Engi...	Manhattan	M812	240 Convent Avenue	Manhattan	NY	10031	40.82112	-

```

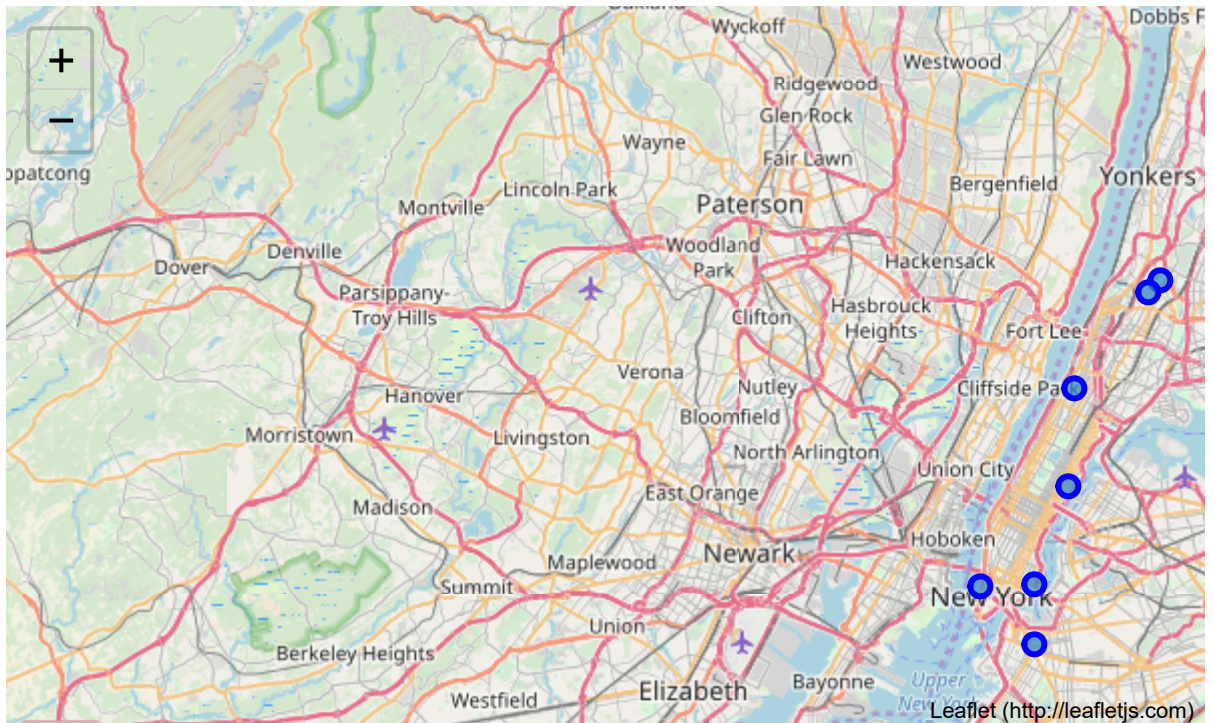
In [13]: NYMapTotal = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, school in zip(dftop10total['Latitude'], dftop10total['Longitude'], dftop10total['Borough'], dftop10total['School']):
    label = '{}', {}'.format(borough, school)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(NYMapTotal)

```

NYMapTotal

Out[13]:



In [14]: *## Sort the highest average test scores in each section to view the location within the 5 boroughs*

```
dftop10Math= df.sort_values(['Average Score (SAT Math)'], ascending = False, axis = 0, inplace = True )
```

```
dftop10Math = df.head(10)
dftop10Math
```

Out[14]:

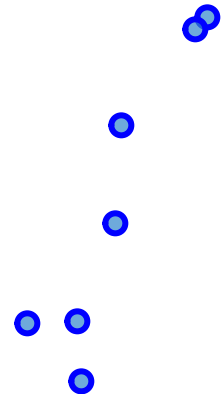
	School ID	School	Borough	Building Code	Street Address	City	State	Zip Code	Latitude	L
105	02M475	Stuyvesant High School	Manhattan	M477	345 Chambers Street	Manhattan	NY	10282	40.71775	-
203	10X445	Bronx High School of Science	Bronx	X445	75 West 205th Street	Bronx	NY	10468	40.87706	-
110	31R605	Staten Island Technical High School	Staten Island	R440	485 Clawson Street	Staten Island	NY	10306	40.56791	-
424	28Q687	Queens High School for the Sciences at York Co...	Queens	Q774	94-50 159th Street	Jamaica	NY	11433	40.70100	-
80	05M692	High School for Mathematics, Science, and Engi...	Manhattan	M812	240 Convent Avenue	Manhattan	NY	10031	40.82112	-
326	13K430	Brooklyn Technical High School	Brooklyn	K430	29 Ft Greene Place	Brooklyn	NY	11217	40.68811	-
385	25Q525	Townsend Harris High School	Queens	Q515	149-11 Melbourne Avenue	Flushing	NY	11367	40.73441	-
208	10X696	High School of American Studies at Lehman College	Bronx	X905	2925 Goulden Avenue	Bronx	NY	10468	40.87126	-
2	01M539	New Explorations into Science, Technology and ...	Manhattan	M022	111 Columbia Street	Manhattan	NY	10002	40.71873	-
55	02M416	Eleanor Roosevelt High School	Manhattan	M855	411 East 76th Street	Manhattan	NY	10021	40.77012	-

```
In [15]: NYmapMath = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, school in zip(dftop10Math['Latitude'], dftop10Math['Longitude'], dftop10Math['Borough'], dftop10Math['School']):
    label = '{} , {}'.format(borough, school)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(NYmapMath)
```

NYmapMath

Out[15]:



Leaflet (<http://leafletjs.com>)

In [16]: *## Sort the highest average test scores in each section to view the location within the 5 boroughs*

```
dftop10Reading= df.sort_values(['Average Score (SAT Reading)'], ascending = False, axis = 0, inplace = True )

dftop10Reading = df.head(10)
dftop10Reading
```

Out[16]:

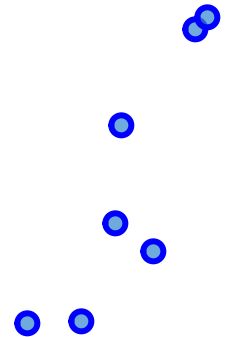
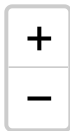
	School ID	School	Borough	Building Code	Street Address	City	State	Zip Code	Latitude	Longitude
105	02M475	Stuyvesant High School	Manhattan	M477	345 Chambers Street	Manhattan	NY	10282	40.71775	-74.0042
208	10X696	High School of American Studies at Lehman College	Bronx	X905	2925 Goulden Avenue	Bronx	NY	10468	40.87126	-89.1911
110	31R605	Staten Island Technical High School	Staten Island	R440	485 Clawson Street	Staten Island	NY	10306	40.56791	-74.1795
203	10X445	Bronx High School of Science	Bronx	X445	75 West 205th Street	Bronx	NY	10468	40.87706	-89.1911
7	01M696	Bard High School Early College	Manhattan	M097	525 East Houston Street	Manhattan	NY	10002	40.71896	-74.0042
385	25Q525	Townsend Harris High School	Queens	Q515	149-11 Melbourne Avenue	Flushing	NY	11367	40.73441	-73.8294
424	28Q687	Queens High School for the Sciences at York College	Queens	Q774	94-50 159th Street	Jamaica	NY	11433	40.70100	-73.8294
251	30Q580	Baccalaureate School for Global Education	Queens	Q798	34-12 36th Avenue	Astoria	NY	11106	40.75517	-73.8294
55	02M416	Eleanor Roosevelt High School	Manhattan	M855	411 East 76th Street	Manhattan	NY	10021	40.77012	-74.0042
80	05M692	High School for Mathematics, Science, and Engineering	Manhattan	M812	240 Convent Avenue	Manhattan	NY	10031	40.82112	-89.1911

```
In [17]: NYMapReading = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, school in zip(dftop10Reading['Latitude'], dftop10Reading['Longitude'], dftop10Reading['Borough'], dftop10Reading['School']):
    label = '{} , {}'.format(borough, school)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(NYMapReading)
```

NYMapReading

Out[17]:



Leaflet (<http://leafletjs.com>)

In [18]: *## Sort the highest average test scores in each section to view the location within the 5 boroughs*

```
dftop10Writing= df.sort_values(['Average Score (SAT Writing)'], ascending = False, axis = 0, inplace = True )

dftop10Writing = df.head(10)
dftop10Writing
```

Out[18]:

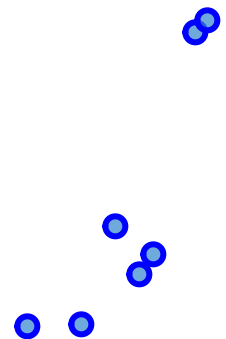
	School ID	School	Borough	Building Code	Street Address	City	State	Zip Code	Latitude	Longitude
105	02M475	Stuyvesant High School	Manhattan	M477	345 Chambers Street	Manhattan	NY	10282	40.71775	-74.00421
208	10X696	High School of American Studies at Lehman College	Bronx	X905	2925 Goulden Avenue	Bronx	NY	10468	40.87126	-73.91781
110	31R605	Staten Island Technical High School	Staten Island	R440	485 Clawson Street	Staten Island	NY	10306	40.56791	-74.18009
203	10X445	Bronx High School of Science	Bronx	X445	75 West 205th Street	Bronx	NY	10468	40.87706	-73.91781
385	25Q525	Townsend Harris High School	Queens	Q515	149-11 Melbourne Avenue	Flushing	NY	11367	40.73441	-73.82941
7	01M696	Bard High School Early College	Manhattan	M097	525 East Houston Street	Manhattan	NY	10002	40.71896	-73.99789
55	02M416	Eleanor Roosevelt High School	Manhattan	M855	411 East 76th Street	Manhattan	NY	10021	40.77012	-73.95691
251	30Q580	Baccalaureate School for Global Education	Queens	Q798	34-12 36th Avenue	Astoria	NY	11106	40.75517	-73.82941
424	28Q687	Queens High School for the Sciences at York College	Queens	Q774	94-50 159th Street	Jamaica	NY	11433	40.70100	-73.82941
241	24Q299	Bard High School Early College Queens	Queens	Q735	30-20 Thomson Avenue	Long Island City	NY	11101	40.74509	-73.95691

```
In [19]: NYmapWriting = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, school in zip(dftop10Writing['Latitude'], dftop10Writing['Longitude'], dftop10Writing['Borough'], dftop10Writing['School']):
    label = '{} , {}'.format(borough, school)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(NYmapWriting)

NYmapWriting
```

Out[19]:

Leaflet (<http://leafletjs.com>)

Part III: Identify the closest vendors to the public high schools with the highest SAT test scores with the 5 borough New York City area, via FourSquare API and geocoding.

```
In [20]: def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]["groups"][0]["items"]

        # return only relevant information for each nearby venue
        venues_list.append([
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item
in venue_list])
    nearby_venues.columns = ['School',
                            'School Latitude',
                            'School Longitude',
                            'Venue',
                            'Venue Latitude',
                            'Venue Longitude',
                            'Venue Category']

    return(nearby_venues)
```

```
In [21]: NYvenues = getNearbyVenues(names=dftop10total['School'],
                                     latitudes=dftop10total['Latitude'],
                                     longitudes=dftop10total['Longitude'],
                                     )
```

Stuyvesant High School
Staten Island Technical High School
Bronx High School of Science
High School of American Studies at Lehman College
Townsend Harris High School
Queens High School for the Sciences at York College
Bard High School Early College
Brooklyn Technical High School
Eleanor Roosevelt High School
High School for Mathematics, Science, and Engineering at City College

```
In [22]: NYvenues.head()
```

Out[22]:

	School	School Latitude	School Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Stuyvesant High School	40.71775	-74.01405	Teardrop Park	40.716843	-74.015576	Park
1	Stuyvesant High School	40.71775	-74.01405	Asphalt Green Battery Park City	40.715924	-74.014673	Athletics & Sports
2	Stuyvesant High School	40.71775	-74.01405	North End Ave Dog Run	40.716876	-74.016149	Dog Run
3	Stuyvesant High School	40.71775	-74.01405	SoulCycle TriBeCa	40.716178	-74.013007	Cycle Studio
4	Stuyvesant High School	40.71775	-74.01405	Nelson A. Rockefeller Park	40.717095	-74.016716	Park


```
In [23]: # one hot encoding
schoolonehot = pd.get_dummies(NYvenues[['Venue']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
schoolonehot['School'] = NYvenues['School']

fixed_columns = [schoolonehot.columns[-1]] + list(schoolonehot.columns[:-1])
schoolonehot = schoolonehot[fixed_columns]

schoolonehot.head()
```

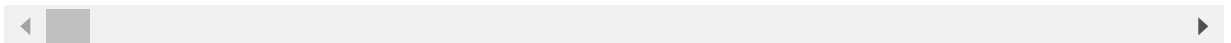
Out[23]:

	School	150th And Jewel Bus Stop	2nd Ave Deli	2nd Floor Bar & Essen	52st Street Food Truck	APEX - Lehman College	Agata & Valentina UES Market	Anita's Roti Shop	Applebee's Grill + Bar	Art Gallery @Lehman College	A.
0	Stuyvesant High School	0	0	0	0	0	0	0	0	0	
1	Stuyvesant High School	0	0	0	0	0	0	0	0	0	
2	Stuyvesant High School	0	0	0	0	0	0	0	0	0	
3	Stuyvesant High School	0	0	0	0	0	0	0	0	0	
4	Stuyvesant High School	0	0	0	0	0	0	0	0	0	

```
In [24]: school_grouped = schoolonehot.groupby('School').mean().reset_index()
school_grouped
```

Out[24]:

	School	150th And Jewel Bus Stop	2nd Ave Deli	2nd Floor Bar & Essen	52st Street Food Truck	APEX - Lehman College	Agata & Valentina UES Market	Anita's Roti Shop	Applebee Grill + B
0	Bard High School Early College	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
1	Bronx High School of Science	0.000000	0.000000	0.000000	0.000000	0.076923	0.000000	0.000000	0.000000
2	Brooklyn Technical High School	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
3	Eleanor Roosevelt High School	0.000000	0.033333	0.033333	0.000000	0.000000	0.033333	0.000000	0.000000
4	High School for Mathematics, Science, and Engi...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
5	High School of American Studies at Lehman College	0.000000	0.000000	0.000000	0.033333	0.000000	0.000000	0.000000	0.000000
6	Queens High School for the Sciences at York Co...	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.033333	0.033333
7	Staten Island Technical High School	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
8	Stuyvesant High School	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
9	Townsend Harris High School	0.058824	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000



```
In [25]: close_venues = 5

for stuff in school_grouped['School']:
    print("-----"+stuff+"-----")
    temp = school_grouped[school_grouped['School'] == stuff].T.reset_index()
    temp.columns = ['venue', 'dist']
    temp = temp.iloc[1:]
    temp['dist'] = temp['dist'].astype(float)
    temp = temp.round({'dist': 2})
    print(temp.sort_values('dist', ascending=False).reset_index(drop=True).head(close_venues))
    print('\n')
```

----Bard High School Early College----

	venue	dist
0	East River Park Fields	0.14
1	Bedlam	0.14
2	East River Park	0.14
3	East River Running Path	0.14
4	Brian Watkins Tennis Center	0.14

----Bronx High School of Science----

	venue	dist
0	Fine Fare Supermarket	0.08
1	Dunkin'	0.08
2	Jerome's Pizza	0.08
3	Blink Fitness	0.08
4	Choi Yuan - Chinese Restaurant	0.08

----Brooklyn Technical High School----

	venue	dist
0	Melt Massage	0.03
1	The Center For Fiction	0.03
2	Las Santas	0.03
3	Theatre For A New Audience at Polonsky Shakesp...	0.03
4	Brooklyn Academy of Music (BAM)	0.03

----Eleanor Roosevelt High School----

	venue	dist
0	Pilates Reforming New York - E. 76th St	0.03
1	Jones Wood Foundry	0.03
2	Boqueria	0.03
3	The Sweet Shop NYC	0.03
4	Orwasher's Bakery	0.03

----High School for Mathematics, Science, and Engineering at City College----

	venue	dist
0	The Edge Harlem	0.03
1	Sugar Hill Cafe	0.03
2	The Harlem Scrabble Room	0.03
3	The Grange Bar & Eatery	0.03
4	Cafe One	0.03

----High School of American Studies at Lehman College----

	venue	dist
0	Little Ceasars	0.03
1	CTown Supermarkets	0.03
2	Cơm Tấm Ninh Kiều	0.03
3	Duane Reade	0.03
4	Dunkin'	0.03

----Queens High School for the Sciences at York College----

	venue	dist
0	Starbucks	0.07

1	Modell's Sporting Goods	0.03
2	Taco Bell / Pizza Hut	0.03
3	K&G	0.03
4	PP Boy Japan Teriyaki & Sushi	0.03

----Staten Island Technical High School----

	venue	dist
0	Chase Bank	0.07
1	CVS pharmacy	0.07
2	Starbucks	0.03
3	New China Buffet	0.03
4	Chipotle Mexican Grill	0.03

----Stuyvesant High School----

	venue	dist
0	Polarn O. Pyret	0.03
1	Barnes & Noble	0.03
2	Tribeca Performing Arts Center	0.03
3	Tribeca Greenmarket	0.03
4	Nelson A. Rockefeller Park	0.03

----Townsend Harris High School----

	venue	dist
0	150th And Jewel Bus Stop	0.06
1	Bakhtar Halal Kabab Main Street	0.06
2	Fitzgerald Gym	0.06
3	S & M Pharmacy Inc	0.06
4	Starbucks	0.06

Results and Discussion

Results:

Stuyvesant High School located in Manhattan was the top performing average SAT test score in all sections of the SAT test. The same ten high schools were listed for each section of test and subsequently, had the highest total average test scores as well. Those schools are: Stuyvesant High School, Staten Island Technical High School, Bronx High School of Science, High School of American Studies at Lehman College, Townsend Harris High School, Queens High School for the Sciences at York College, Bard High School Early College, Brooklyn Technical High School, Eleanor Roosevelt High School, High School for Mathematics, Science, and Engineering at City College.

Discussion:

The aim of this project is to help families and students identify public high schools within the New York City, five boroughs area with the highest average SAT test scores. xpaths can chose the neighborhoods to which they want to relocate based on the most common SAT test scores are used help colleges interpret students' overall academic performance in relation to the national applicant pool. For families and students that are planning for college admission after high school, the public high schools with the total highest test scores are of interest. Additionally, if a student is weak in one section of the test, a public high school that has high average test scores within this academic disipline is of intersest as well. Lastly, identifying the closest local vendors to the highest total average test score public high school assisst family with planning for after school activities, pick-up/drop-off locations, and an overall general idea of what is within walking distance of the school.

Conclusion:

This project aids families and students alike with a better understanding of public high schools with high academic achievements with respect to the SAT test scores. Venues near the public high schools aslo allow families and students to find local, walking distance places near the public high schools for planning purposes and neighborhood associations. Reconizing that this project was singularly focused on SAT test scores, future projects could use this data as a platform for more indepth analysis of public high schools located within the New York City, five boroughs area.

In []: