

CAPSTONE PROJECT

INTRODUCTION/BUSINESS PROBLEM

New York city data will analyse by this project. According to the number of check-ins, we will find the most visited commercial shop, there could be potential business opportunity but some neighbourhoods that are lacking the selected type of shop we try to find it.

TARGET AUDIENCE

people who want to do business in commercial shops who have no idea

DATA SECTION

The data comes from Ding qi Yang from the following link

<https://sites.google.com/site/yangdingqi/home/foursquare-dataset>. It contains 227,428 check-ins in New York city. The data contains a file in csv format. Each file contains 8 columns, which are:

User ID (anonymized)

Venue ID (Foursquare)

Venue category ID (Foursquare)

Venue category name (Foursquare)

Latitude

Longitude

Time zone offset in minutes (The offset in minutes between when this check-in occurred and the same time in UTC)

UTC time

APPLICATION

We will find the most visited type of shop (commercial) according to the number of check-ins given in the data, then we will try to find neighbourhoods that has none of this type of shop.

Examples are for 2000 venues, and the red dot is the centre neighbourhood which has the greatest number of Bars between selected coordinates. We did find two neighbourhoods that are closest to it having none Bars within 4 kilometres.

Out[6]:

	VenueID	CategoryName	Visitor Count	Latitude	Longitude
0	49bbd6c0f964a520f4531fe3	Arts & Crafts Store	7	40.719810375488535	-74.00258103213994
1	4a43c0aef964a520c6a61fe3	Bridge	37	40.60679958140643	-74.04416981025437
2	4c5cc7b485a1e21e00d35711	Home (private)	1	40.716161684843215	-73.88307005845945
3	4bc7086715a7ef3bef9878da	Medical Center	1	40.7451638	-73.982518775
4	4cf2c5321d18a143951b5cec	Food Truck	4	40.74010382743943	-73.98965835571289

```
[[('Train Station', 943), ('Park', 778), ('Airport', 769), ('Bar', 756), ('Subway', 587), ('Coffee Shop', 447), ('Gym / Fitness Center', 447), ('Food & Drink Shop', 426), ('Neighborhood', 362), ('Plaza', 342), ('Stadium', 339), ('Bridge', 272), ('Office', 264), ('Department Store', 240), ('Mall', 238), ('Burger Joint', 206), ('American Restaurant', 202), ('Road', 201), ('Bus Station', 196), ('Hotel', 184), ('Other Great Outdoors', 178), ('Music Venue', 166), ('Home (private)', 158), ('Mexican Restaurant', 154), ('Electronics Store', 137), ('Ferry', 126), ('College Academic Building', 116), ('Sandwich Place', 115), ('BBQ Joint', 109), ('Bookstore', 105), ('Building', 100), ('Medical Center', 94), ('University', 94), ('Clothing Store', 89), ('Drugstore / Pharmacy', 83), ('Beach', 72), ('Government Building', 70), ('Convention Center', 70), ('Sporting Goods Shop', 68), ('Bakery', 68), ('Fast Food Restaurant', 59), ('Chinese Restaurant', 59), ('Theater', 57), ('Deli / Bodega', 55), ('Movie Theater', 53), ('Food Truck', 51), ('Sushi Restaurant', 50), ('Pizza Place', 47), ('General Entertainment', 47), ('Ice Cream Shop', 46), ('Bank', 45), ('Miscellaneous Shop', 41), ('Light Rail', 40), ('Church', 38), ('Concert Hall', 38), ('French Restaurant', 36), ('Seafood Restaurant', 35), ('Fried Chicken Joint', 34), ('Residential Building (Apartment / Condo)', 33), ('Italian Restaurant', 33), ('Comedy Club', 33), ('Diner', 30), ('Toy / Game Store', 29), ('Vegetarian / Vegan Restaurant', 27), ('Café', 26), ('Community College', 26), ('Scenic Lookout', 24), ('Donut Shop', 23), ('German Restaurant', 23), ('Bowling Alley', 22), ('Beer Garden', 22), ('Gastropub', 19), ('Candy Store', 19), ('Bagel Shop', 19), ('Cuban Restaurant', 18), ('School', 18), ('Cupcake Shop', 18), ('Breakfast Spot', 18), ('General Travel', 17), ('Salon / Barber Shop', 16), ('General College & University', 16), ('Hardware Store', 16), ('Japanese Restaurant', 16), ('Latin American Restaurant', 15), ('Athletic & Sport', 15), ('Restaurant', 13), ('Spanish Restaurant', 12), ('Spa / Massage', 12), ('Middle Eastern Restaurant', 12), ('Malaysian Restaurant', 12), ('Record Shop', 12), ('Library', 12), ('Student Center', 11), ('Wings Joint', 11), ('High School', 11), ('Arts & Crafts Store', 10), ('Laundry Service', 10), ('Gas Station / Garage', 10), ('Asian Restaurant', 9), ('Burrito Place', 9), ('Parking', 9), ('Harbor / Marina', 9), ('Thai Restaurant', 9), ('Playground', 8), ('Campground', 8), ('Salad Place', 8), ('Event Space', 8), ('Pool Hall', 8), ('Ramen / Noodle House', 8), ('Automotive Shop', 7), ('Convenience Store', 7), ('Tea Room', 7), ('Post Office', 7), ('Indian Restaurant', 7), ('River', 7), ('Thrift / Vintage Store', 7), ('Paper / Office Supplies Store', 6), ('Cosmetics Shop', 6), ('Dessert Shop', 6), ('Museum', 6), ('Pet Store', 6), ('College & University', 6), ('Brazilian Restaurant', 6), ('Tanning Salon', 5), ('Bike Shop', 5), ('Art Gallery', 5), ('Arts & Entertainment', 5), ('Food', 5), ('Southern / Soul Food Restaurant', 4), ('Smoke Shop', 4), ('College Theater', 4), ('Snack Place', 4), ('Furniture / Home Store', 4), ('Sculpture Garden', 4), ('Mobile Phone Shop', 3), ('Tattoo Parlor', 3), ('Other Nightlife', 3), ('Moving Target', 3), ('Caribbean Restaurant', 3), ('Video Game Store', 3), ('Arcade', 3), ('Steakhouse', 3), ('Greek Restaurant', 3), ('Soup Place', 2), ('Music Store', 2), ('Dumpling Restaurant', 2), ('Performing Arts Venue', 2), ('Camera Store', 2), ('Housing Development', 2), ('Synagogue', 2), ('Mediterranean Restaurant', 2), ('African Restaurant', 2), ('Taxi', 2), ('Professional & Other Places', 2), ('Taco Place', 2), ('Jewelry Store', 1), ('Animal Shelter', 1), ('Factory', 1), ('Cemetery', 1), ('Medical School', 1), ('Pool', 1), ('Garden', 1), ('Hot Dog Joint', 1), ('Outdoors & Recreation', 1), ('Sorority House', 1), ('Casino', 1), ('Temple', 1), ('Historic Site', 1), ('Rest Area', 1), ('Bar', 1)]

'Bar' is the most visited commercial category according to given data.
```

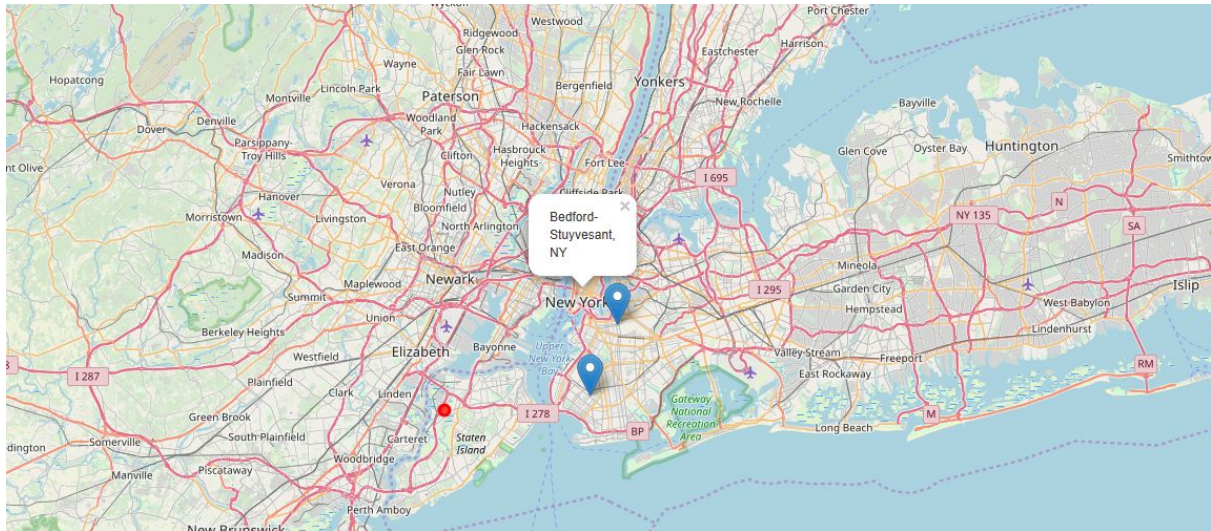
Then we check coordinates and count how many numbers of bars are there (2000 trials)

Coordinates with number of Bar shops within 4 kilometers according to 2000 venues.

```
('40.60613336268842', '-74.17904376983643') : 2
('40.719810375488535', '-74.00258103213994') : 0
('40.60679958140643', '-74.04416981025437') : 0
('40.716161684843215', '-73.88307005845945') : 0
('40.69042711809854', '-73.95468677509598') : 0
('40.751591431346306', '-73.9741214009634') : 0
('40.60613336268842', '-74.17904376983643') : 2
```

Find two neighbourhoods that are closest to the coordinate which has the greatest number of the specific shop type but lacking that within 4 kilometres.

- Bensonhurst
- Bedford-Stuyvesant



Results and Conclusions

Our Sample contains about 2000 venues, more than 10 coordinates with most visited shop type within four kilometres has no bar. We pin down into the map two closest neighbourhoods which has the greatest number of the specific shop type but lacking that within 4 kilometres i.e., Bensonhurst and Bedford-Stuyvesant with the help foursquare database (API). Further research is needed for the data because day to day the data might get updated. The above map can be used for future entrepreneurs.