THE BATTLE OF THE NEIGHBOURHOOD - COURSERA IBM CAPSTONE PROJECT

BACKGROUND

Restaurants are a notoriously difficult business to own or operate. Not only is it a relatively capital and employee-intensive business, restaurants are also highly regulated, low margin and in most cases have a plethora of competition to deal with. There are more dining establishments and more diners now than during any other time in history; the National Restaurant Association reports over 1 million restaurant locations in the U.S. alone, a particular boon considering more than half the American population visits at least one of them each week. To fill a niche and remain relevant, prospective, as well as established, restaurants have to hedge their bets with well-rounded and well-directed dataset. Data science provides valuable insights regarding market trends and evolving consumer lifestyles so that restaurateurs can better address and meet public demand.

I.BUSINESS PROBLEM

No single restaurant is ever going to appeal to everyone. Some people like quiet, intimate settings; others prefer boisterous ones. Some people want to bring their kids along; others want to dine alone. The preferences are as varied as the possible offerings, with specific generational cohorts preferring one thing, "people who like Indian food" preferring another and everyone else liking a million other things in between. My client, a successful Filipino restaurant chain in Philippines is looking to expand operation into North America through New York (NYC).

They are interested in building in an area that meets the following criteria:

- Above average economy. (GDP)
- Average to above average total population.
- Average to above average density. (People/sqkm)
- Average to above average Asian ethnic population.
- Neighborhood with above average Rating Filipino restaurant.

II.DATA REQUIREMENTS AND RESOURCES

The necessary information needed by the Client will come from the following sources:

Part 1: NYC Population & Demographic Characteristics

Data source: https://en.wikipedia.org/wiki/New_York_City ;https://en.wikipedia.org/wiki/Demographics_of_New_York_City.

Web scraping techniques was used to get NYC's population density and demographics data from Wikipedia.

NYC Population

New York City's five boroughs								
Jurisdiction		Population	Gross Domestic Product		Land area		Density	
Borough	*County*	*Estimate (2018)*[150]	*billions (US$)*[151]	*per capita (US$)*	*square miles*	*square km*	*persons / sq. mi*	*persons / km²*
The Bronx	Bronx	1,432,132	42.695	29,200	42.10	109.04	34,653	13,231
Brooklyn	Kings	2,582,830	91.559	34,600	70.82	183.42	37,137	14,649
Manhattan	New York	1,628,701	600.244	360,900	22.83	59.13	72,033	27,826
Queens	Queens	2,278,906	93.310	39,600	108.53	281.09	21,460	8,354
Staten Island	Richmond	476,179	14.514	30,300	58.37	151.18	8,112	3,132
City of New York		**8,398,748**	**842.343**	**97,700**	**302.64**	**783.83**	**28,188**	**10,947**
State of New York		19,745,289	1,701.399	85,700	47,214	122,284	416.4	159

Sources:[152] *and see individual borough articles*

Demographic Characteristics

Jurisdiction	Population 2000 census	% white	% black or African American	% Asian	% Other	% mixed race	% Hispanic/ Latino of any race		% Catholic	% not affiliated	% Jewish	% Protestant	Estimate of % not reporting
		Race					Ethnicity		Religious groups				
Brooklyn	2,465,326	41.2	36.4	7.5	10.6	4.3	19.8		37	4	15	8	33
Queens	2,229,379	44.1	20.0	17.6	12.3	6.1	25.0		29	37	11	5	15
Manhattan	1,537,195	54.4	17.4	9.4	14.7	4.1	27.2		37	11	20	9	19
Bronx	1,332,650	29.9	35.6	3.0	25.7	5.8	48.4		44	14	6	5	29
Staten Island	443,728	77.6	9.7	5.7	4.3	2.7	12.1		60	11	8	5	14
NYC Total	8,008,278	44.7	26.6	9.8	14.0	4.9	27.0		37	17	13	6	24
NY State	18,976,457	67.9	15.9	5.5	7.5	3.1	15.1		42	20	9	10	16
USA	281,421,906	75.1	12.3	3.6	6.5	2.4	12.5		22	37	2	23	12

Source: 2000 Census[63]

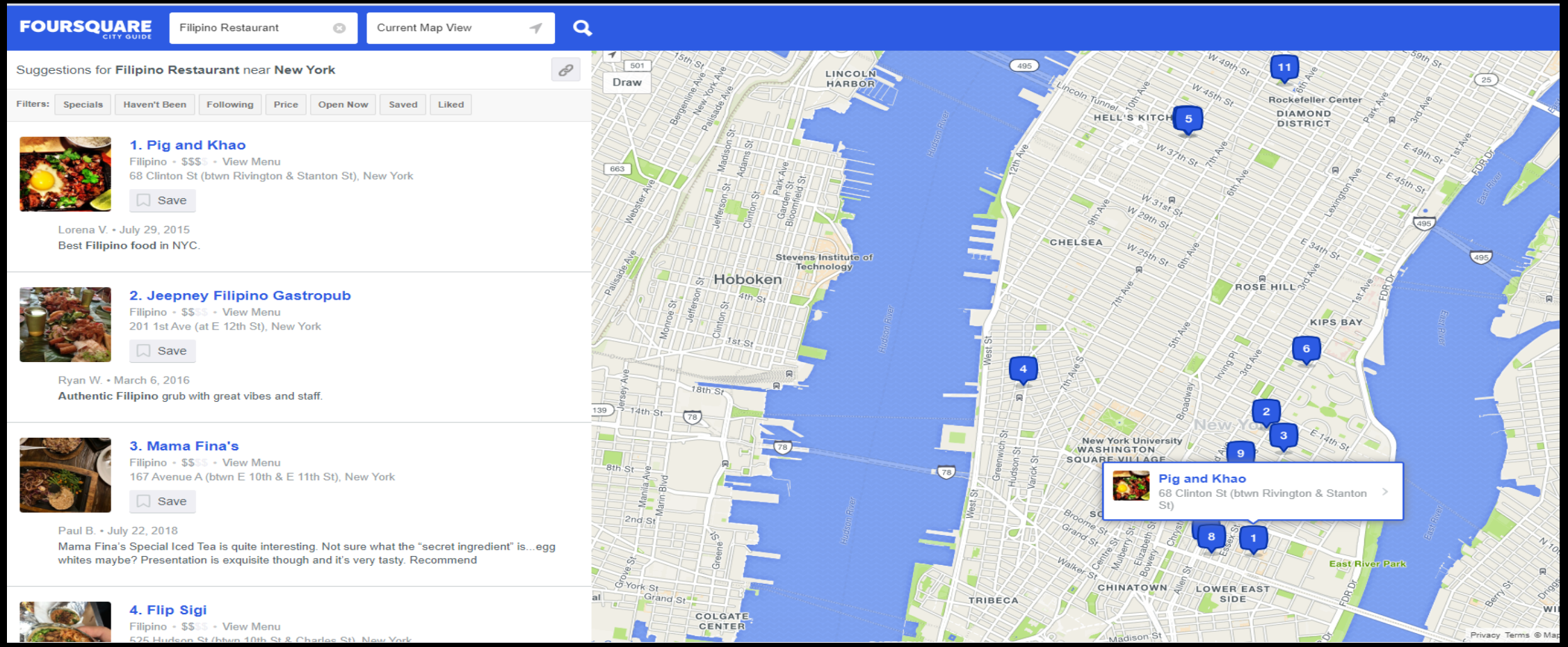
American Indian, Native Alaskan, Native Hawaiian, and Pacific Islander make up 2.9% of the population of NYC, and have been included with "Other".

Source for religious groups: ARDA[64]

II.DATA REQUIREMENTS AND RESOURCES

Part 2: Who are the competitors in that location?

Data source: https://cocl.us/new_york_dataset and Foursquare API. These dataset was used to explore various neighborhoods and each Filipino restaurants venues in the neighborhood.



III.METHODOLOGY

In order to establish the targeted neighborhood(s), we will explore the demographics of the neighborhoods in the city of New York by segmenting the data and conducting descriptive analysis using Panda. Additional data will be extracted by web scraping and API will be used to generate data.

Data Group 1 (Population and Demographic Data)

1. Data was pulled from Wikipedia by web scraping and using python library beautiful soup.
2. Pandas was used to transform the data to data frame.

(1)

```
In [5]: #web scrapping data table from wikipedia using bs4
response_obj = requests.get('https://en.wikipedia.org/wiki/New_York_City').text
soup = BeautifulSoup(response_obj,'lxml')
Neighborhoods_NYC_Table = soup.find('table', {'class':'wikitable sortable'})

In [10]: #checking website table from wikipedia
rows = Neighborhoods_NYC_Table.select("tbody > tr")[3:8]

boroughs = []
for row in rows:
    borough = {}
    tds = row.select('td')
    borough["borough"] = tds[0].text.strip()
    borough["county"] = tds[1].text.strip()
    borough["population"] = float(tds[2].text.strip().replace(",",""))
    borough["gdp_billions"] = float(tds[3].text.strip().replace(",",""))
    borough["gdp_per_capita"] = float(tds[4].text.strip().replace(",",""))
    borough["land_sqm"] = float(tds[5].text.strip().replace(",",""))
    borough["land_sqkm"] = float(tds[6].text.strip().replace(",",""))
    borough["persons_sqm"] = float(tds[7].text.strip().replace(",",""))
    borough["persons_sqkm"] = float(tds[8].text.strip().replace(",",""))

    boroughs.append(borough)

print(boroughs)

df = pd.DataFrame(boroughs, columns=["borough","county", "population", "gdp_per_capita", "persons_sqkm"])
df
```

```
[{'borough': 'The Bronx', 'county': 'Bronx', 'population': 1432132.0, 'gdp_billions': 42.695, 'gdp_per_capita': 29200.0, 'land_
sqm': 42.1, 'land_sqkm': 109.04, 'persons_sqm': 34653.0, 'persons_sqkm': 13231.0}, {'borough': 'Brooklyn', 'county': 'Kings',
'population': 2582830.0, 'gdp_billions': 91.559, 'gdp_per_capita': 34600.0, 'land_sqm': 70.82, 'land_sqkm': 183.42, 'persons_sq
m': 37137.0, 'persons_sqkm': 14649.0}, {'borough': 'Manhattan', 'county': 'New York', 'population': 1628701.0, 'gdp_billions':
600.244, 'gdp_per_capita': 360900.0, 'land_sqm': 22.83, 'land_sqkm': 59.13, 'persons_sqm': 72033.0, 'persons_sqkm': 27826.0},
{'borough': 'Queens', 'county': 'Queens', 'population': 2278906.0, 'gdp_billions': 93.31, 'gdp_per_capita': 39600.0, 'land_sq
m': 108.53, 'land_sqkm': 281.09, 'persons_sqm': 21460.0, 'persons_sqkm': 8354.0}, {'borough': 'Staten Island', 'county': 'Richm
ond', 'population': 476179.0, 'gdp_billions': 14.514, 'gdp_per_capita': 30300.0, 'land_sqm': 58.37, 'land_sqkm': 151.18, 'perso
ns_sqm': 8112.0, 'persons_sqkm': 3132.0}]
```

(2)

	borough	county	population	gdp_per_capita	persons_sqkm
0	The Bronx	Bronx	1432132.0	29200.0	13231.0
1	Brooklyn	Kings	2582830.0	34600.0	14649.0
2	Manhattan	New York	1628701.0	360900.0	27826.0
3	Queens	Queens	2278906.0	39600.0	8354.0
4	Staten Island	Richmond	476179.0	30300.0	3132.0

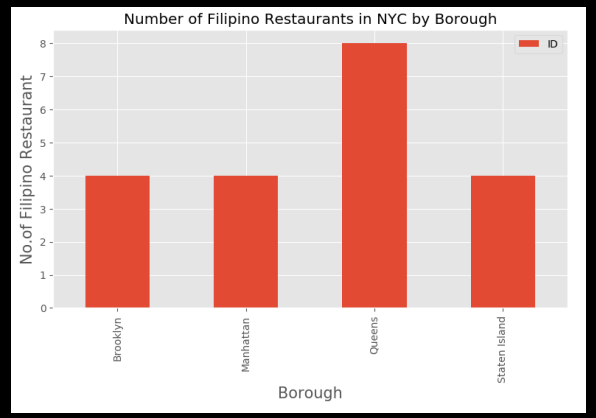
	jurisdiction	%_white	%_black_or_african_amercian	%_Asian	%_other	%_mixed_race	%_hispanic_latino_of_other_race
0	Queens	44.1	20.0	17.6	12.3	6.1	25.0
1	Manhattan	54.4	17.4	9.4	14.7	4.1	27.2
2	Bronx	29.9	35.6	3.0	25.7	5.8	48.4
3	Staten Island	77.6	9.7	5.7	4.3	2.7	12.1
4	NYC Total	44.7	26.6	9.8	14.0	4.9	27.0

III.METHODOLOGY

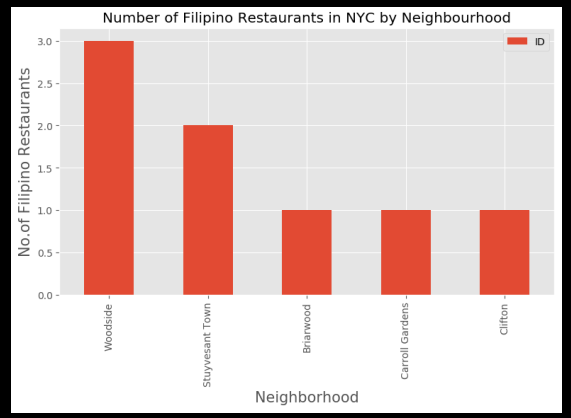
Data Group 2 (NYC Geographical Neighbourhood Data)

1. Retrieved Borough, Neighbourhood and Coordinates data from https://cocl.us/new_york_dataset , transformed to data frame and plot it using Bar chart.
2. Used Foursquare API to retrieve the number of Filipino restaurant in NYC Borough.
3. Retrieved the number of Filipino restaurants in NYC by Neighboorhood.
4. Called again to get the Neighbourhood's Name, Likes, Rating and Tips.
5. Analyzed the Data Group 1/Group 2 and plot the final Map.
6. Validated Findings/Observations and Hypothesis.
7. Constructed Conclusion and Hypothesis.

(1)



(2)



(3)

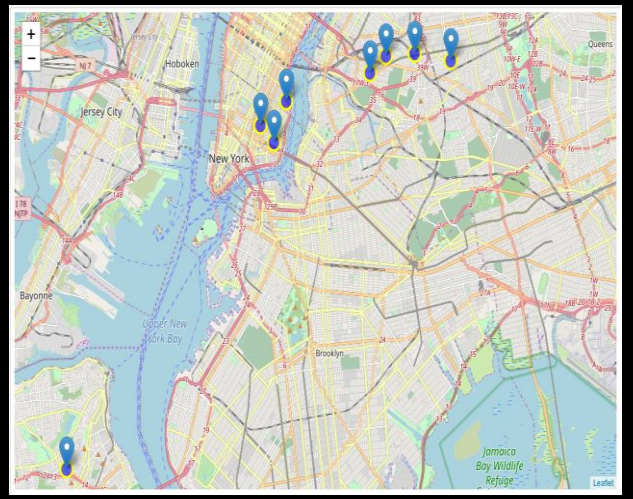
	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
0	Brooklyn	Flatbush	4afa2432f964a520081722e3	Purple Yam	137	7.9	76
1	Brooklyn	Carroll Gardens	591e55677247502fb9514269	Fob Restaurant	55	7.8	28
2	Brooklyn	Prospect Park South	4afa2432f964a520081722e3	Purple Yam	137	7.9	76
3	Manhattan	Lower East Side	50588bc652626048 3ab2860e	Pig and Khao	1081	9.1	301
4	Queens	Woodside	4b5232def964a520456f27e3	Renee's Kitchenette & Grille	60	8.6	20

	Neighborhood	Average Rating
8	Noho	9.100000
7	Lower East Side	9.100000
15	Sunnyside Gardens	8.900000
14	Sunnyside	8.900000
13	Stuyvesant Town	8.350000
16	Woodside	8.166667
4	Elmhurst	8.000000
5	Flatbush	7.900000
10	Prospect Park South	7.900000
3	Ditmas Park	7.900000

(4)

	Borough	Neighborhood	Latitude	Longitude	Average Rating
1	Manhattan	Lower East Side	40.717807	-73.980890	9.100000
2	Manhattan	Noho	40.723259	-73.988434	9.100000
4	Queens	Sunnyside	40.740176	-73.926916	8.900000
5	Staten Island	Sunnyside	40.612760	-74.097126	8.900000
6	Queens	Sunnyside Gardens	40.745652	-73.918193	8.900000
3	Manhattan	Stuyvesant Town	40.731000	-73.974052	8.350000
7	Queens	Woodside	40.746349	-73.901842	8.166667
0	Queens	Elmhurst	40.744049	-73.881656	8.000000

(5)



V.DISCUSSIONS AND RESULTS

Part 1: NYC Population & Demographic Characteristics

- **Findings and Observations**
  1. Brooklyn is the most populous Borough but has the lowest gdp_per_capita among the Boroughs.
  2. Queen is the second most populous Borough and has higher gdp_per_capita than Brooklyn but smaller population density. Highest in-terms of asian ethnic minority population.
  3. Manhattan is the second asian ethnic minority population, first in gdp_per_capita and population density. High person per km means people are living in a high rise building like appartments, hotels and condominiums.

- **Hypothesis 1**

  Manhattan could be the best place to start a Filipino Restaurant given that;
  1. Good Economy (highest gdp_per_capita).
  2. High Population density (person per km can be relate to people who are living in a high rise building
  3. Second in Asian ethnic minority population.

	jurisdiction	%_white	%_black_or_african_amercian	%_Asian	%_other	%_mixed_race	%_hispanic_latino_of_other_race
0	Queens	44.1	20.0	17.6	12.3	6.1	25.0
1	Manhattan	54.4	17.4	9.4	14.7	4.1	27.2
2	Bronx	29.9	35.6	3.0	25.7	5.8	48.4
3	Staten Island	77.6	9.7	5.7	4.3	2.7	12.1
4	NYC Total	44.7	26.6	9.8	14.0	4.9	27.0

V.DISCUSSIONS AND RESULTS

Part 2: Where is the best location and Who are the competitors in that location?

- **Findings and Observations**

1. Lower East Side, Noho, Stuyvesant in Manhattan are some of the best neighborhoods for Filipino cuisine.
2. Staten Island Neighborhood has the lowest rated Filipino Restaurants in NYC.
3. Manhattan Borough is the best place to stay if you prefer Filipino Cuisine.
4. Bronx Borough does not have Filipino Restaurant at the moment.
5. Pig and Khao Restaurant will be the tough competitor. It has 2 branches in Manhattan (Lower East Side and Noho). Stuyvesant Town has Mama Fina's Restaurant which second behind Pig and Khao.

- **Hypothesis 2**

Manhattan could be the best place to start a Filipino Restaurant given that;

1. It has the best neigborhoods for Filipino cusine.
2. It has the top rating Filipino Restaurants.
3. It is the best place to stay if you prefer Filipino Cuisine.

VI.CONCLUSION AND RECOMMENDATION

1. Lower East Side or Noho in Manhattan would be the best choice to start a restaurant given that;
    - It is the third most populous urban area in New York City (NYC).
    - It has a population density of 27,826 people per square km, and confirmed highest of any borough in the United States.
    - It has a good economy (gdp_per_capita).
    - It has some of the top rated Filipino restaurants located in that area
    - It has the second highest Asian ethnic minority population in NYC.
    - Tough competitors are there but people looks for other choices.

2. There are very few Filipino Restaurants around Manhattan and the competition is very weak therefore, this the best time to start a Filipino Restaurant in the neighbourhood of Manhattan.