The Battle of the Neighborhoods

By: Lu Yang

Jan 3, 2020

Table of Contents

- Introduction
- Data Acquisition and Cleaning
- Exploratory Data Analysis
- Data Visualization and Analysis
- Conclusion
- Future Work

Introduction

- Immigrants to the US
- Chinese immigrants increasing
- Queens borough with numerous Asian people
- New immigrants need help and directions
- Government helps new immigrants

Open a new Chinese restaurant? Where?

Problem Definition

- Where to open a new Chinese restaurant in the neighborhoods of Queens, New York City

Target Audience

- New immigrants
- Government funded multicultural associations

Data Acquisition and Cleaning – Data Source

- Venues categories from Foursquare
- Population of neighborhoods of Queens borough
  (https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoodpop.htm)

Data Acquisition and Cleaning – Data Cleaning

- Venues categories from Foursquare

(1)

	Neighborhood	Accessories Store	Afghan Restaurant	Airport Terminal	American Restaurant	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Arts & Entertainment	Asian Restaurant	Athletics & Sports	Automotive Shop	BBQ Joint	Bagel Shop
0	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(2)

	Neighborhood	Afghan Restaurant	American Restaurant	Arepa Restaurant	Argentinian Restaurant	Asian Restaurant	Brazilian Restaurant	Cajun / Creole Restaurant	Cantonese Restaurant	Caribbean Restaurant	Chinese Restaurant	Colombian Restaurant	Comfort Food Restaurant	Cuban Restaurant
0	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Astoria	0	0	0	0	0	1	0	0	0	0	0	0	0
2	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Astoria	0	0	0	0	0	0	0	0	0	0	0	0	0

Data Acquisition and Cleaning – Data Cleaning

- Venues categories from Foursquare

(3)

Neighborhood	Cantonese Restaurant	Chinese Restaurant	Dim Sum Restaurant	Dumpling Restaurant	Filipino Restaurant	Hotpot Restaurant	Indonesian Restaurant	Japanese Restaurant	Korean Restaurant	Malay Restaurant	Shanghai Restaurant	Szechuan Restaurant	Taiwanese Restaurant	Vietnamese Restaurant
Arverne	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Astoria	0	1	0	0	0	0	0	2	1	0	0	0	0	0
Astoria Heights	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Auburndale	0	0	0	0	0	0	0	0	1	0	0	0	0	0
Bay Terrace	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bayside	0	1	0	0	0	0	0	0	1	0	1	0	0	1
Bayswater	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Beechhurst	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Bellaire	0	1	0	0	0	0	0	0	0	0	0	0	0	0

Data Acquisition and Cleaning – Data Cleaning

- Venues categories from Foursquare and population from Department of Health

(4)

Neighborhood	Cantonese Restaurant	Chinese Restaurant	Dim Sum Restaurant	Dumpling Restaurant	Filipino Restaurant	Hotpot Restaurant	Indonesian Restaurant	Japanese Restaurant	Korean Restaurant	Malay Restaurant	Shanghai Restaurant	Szechuan Restaurant	Taiwanese Restaurant	Vietnamese Restaurant	Population
Astoria, Astoria Height, Long Island City	0	4	0	0	0	0	0	2	1	0	0	0	0	0	180171
Jackson Heights, Corona	0	0	0	0	0	0	0	0	0	0	0	0	0	0	189150
Flushing, Murray Hill, Whitestone	1	3	1	1	0	6	0	1	26	1	1	1	0	0	258315
Bayside, Douglaston, Little Neck	1	6	0	0	0	0	0	1	4	0	1	0	0	2	121623
Queens Village, Cambria Heights, Rosedale	0	2	0	0	0	0	0	0	0	0	0	0	0	0	195531
Briarwood, Fresh Meadows & Hillcrest	0	2	1	0	0	0	0	0	0	0	0	0	0	0	159617
Elmhurst, East Elmhurst	0	4	0	0	0	1	1	0	0	1	0	0	0	2	150886
Forest Hills, Rego Park	0	2	0	0	0	0	0	1	0	0	1	0	0	0	117027
Sunnyside, Woodside	0	4	0	0	4	0	0	1	1	0	0	0	0	0	138673
Ridgewood, Glendale, Middle Village	0	3	0	0	0	0	0	0	1	0	0	0	0	0	176336
Richmond Hill, Woodhaven	0	0	0	0	0	0	0	1	0	0	0	0	0	0	147541
Jamaica, Hollis, St. Albans	0	3	0	0	0	0	0	1	0	0	0	0	0	0	228533
Howard Beach, Ozone Park	0	2	0	0	0	0	0	1	0	0	0	0	0	0	127565
Far Rockaway, Breezy Point, Broad Channel	0	3	0	0	0	0	0	0	0	0	0	0	0	0	119043

Exploratory Data Analysis

- We would hope to choose one neighborhood
  (1) Big population
  (2) Heavily populated by Chinese/east Asian people
  (3) Relatively small number of existing Chinese restaurants in neighborhoods satisfying (2)
- How do Queens's neighborhoods look like?

Exploratory Data Analysis

- Hmm…we need more knowledge. The wiki of demographics of Queens says:

  Numerous Asians live in:
    - Flushing
    - Little Neck
    - Bayside
    - Fresh Meadows
    - Jamaica Estates
    - Elmhurst
   - Woodside
    - Richmond Hill
    - Ozone Park

Exploratory Data Analysis

- Tweak the dataset again

Neighborhood	Cantonese Restaurant	Chinese Restaurant	Dim Sum Restaurant	Dumpling Restaurant	Filipino Restaurant	Hotpot Restaurant	Indonesian Restaurant	Japanese Restaurant	Korean Restaurant	Malay Restaurant	Shanghai Restaurant	Szechuan Restaurant	Taiwanese Restaurant	Vietnamese Restaurant	Population	Total Restaurant
Astoria, Astoria Height, Long Island City	0	4	0	0	0	0	0	2	1	0	0	0	0	0	180171	7
Jackson Heights, Corona	0	0	0	0	0	0	0	0	0	0	0	0	0	0	189150	0
Flushing, Murray Hill, Whitestone	1	3	1	1	0	6	0	1	26	1	1	1	0	0	258315	42
Bayside, Douglaston, Little Neck	1	6	0	0	0	0	0	1	4	0	1	0	0	2	121623	15
Queens Village, Cambria Heights, Rosedale	0	2	0	0	0	0	0	0	0	0	0	0	0	0	195531	2
Briarwood, Fresh Meadows & Hillcrest	0	2	1	0	0	0	0	0	0	0	0	0	0	0	159617	3
Elmhurst, East Elmhurst	0	4	0	0	0	1	1	0	0	1	0	0	0	2	150886	9
Forest Hills, Rego Park	0	2	0	0	0	0	0	1	0	0	1	0	0	0	117027	4
Sunnyside, Woodside	0	4	0	0	4	0	0	1	1	0	0	0	0	0	138673	10
Ridgewood, Glendale, Middle Village	0	3	0	0	0	0	0	0	1	0	0	0	0	0	176336	4
Richmond Hill, Woodhaven	0	0	0	0	0	0	0	1	0	0	0	0	0	0	147541	1
Jamaica, Hollis, St. Albans	0	3	0	0	0	0	0	1	0	0	0	0	0	0	228533	4
Howard Beach, Ozone Park	0	2	0	0	0	0	0	1	0	0	0	0	0	0	127565	3
Far Rockaway, Breezy Point, Broad Channel	0	3	0	0	0	0	0	0	0	0	0	0	0	0	119043	3

Exploratory Data Analysis

- Still need to add one more column to the dataset

Neighborhood	Cantonese Restaurant	Chinese Restaurant	Dim Sum Restaurant	Dumpling Restaurant	Filipino Restaurant	Hotpot Restaurant	Indonesian Restaurant	Japanese Restaurant	Korean Restaurant	Malay Restaurant	Shanghai Restaurant	Szechuan Restaurant	Taiwanese Restaurant	Vietnamese Restaurant	Population	Total Restaurant	Chinese/Total
Astoria, Astoria Height, Long Island City	0	4	0	0	0	0	0	2	1	0	0	0	0	0	180171	7	0.57
Jackson Heights, Corona	0	0	0	0	0	0	0	0	0	0	0	0	0	0	189150	0	0.00
Flushing, Murray Hill, Whitestone	1	3	1	1	0	6	0	1	26	1	1	1	0	0	258315	42	0.07
Bayside, Douglaston, Little Neck	1	6	0	0	0	0	0	1	4	0	1	0	0	2	121623	15	0.40
Queens Village, Cambria Heights, Rosedale	0	2	0	0	0	0	0	0	0	0	0	0	0	0	195531	2	1.00
Briarwood, Fresh Meadows & Hillcrest	0	2	1	0	0	0	0	0	0	0	0	0	0	0	159617	3	0.67
Elmhurst, East Elmhurst	0	4	0	0	0	1	1	0	0	1	0	0	0	2	150886	9	0.44
Forest Hills, Rego Park	0	2	0	0	0	0	0	1	0	0	1	0	0	0	117027	4	0.50

Data Visualization and Analysis

Number of Chinese Restaurants and Population

Data Visualization and Analysis

Number of east Asian Restaurants and Population
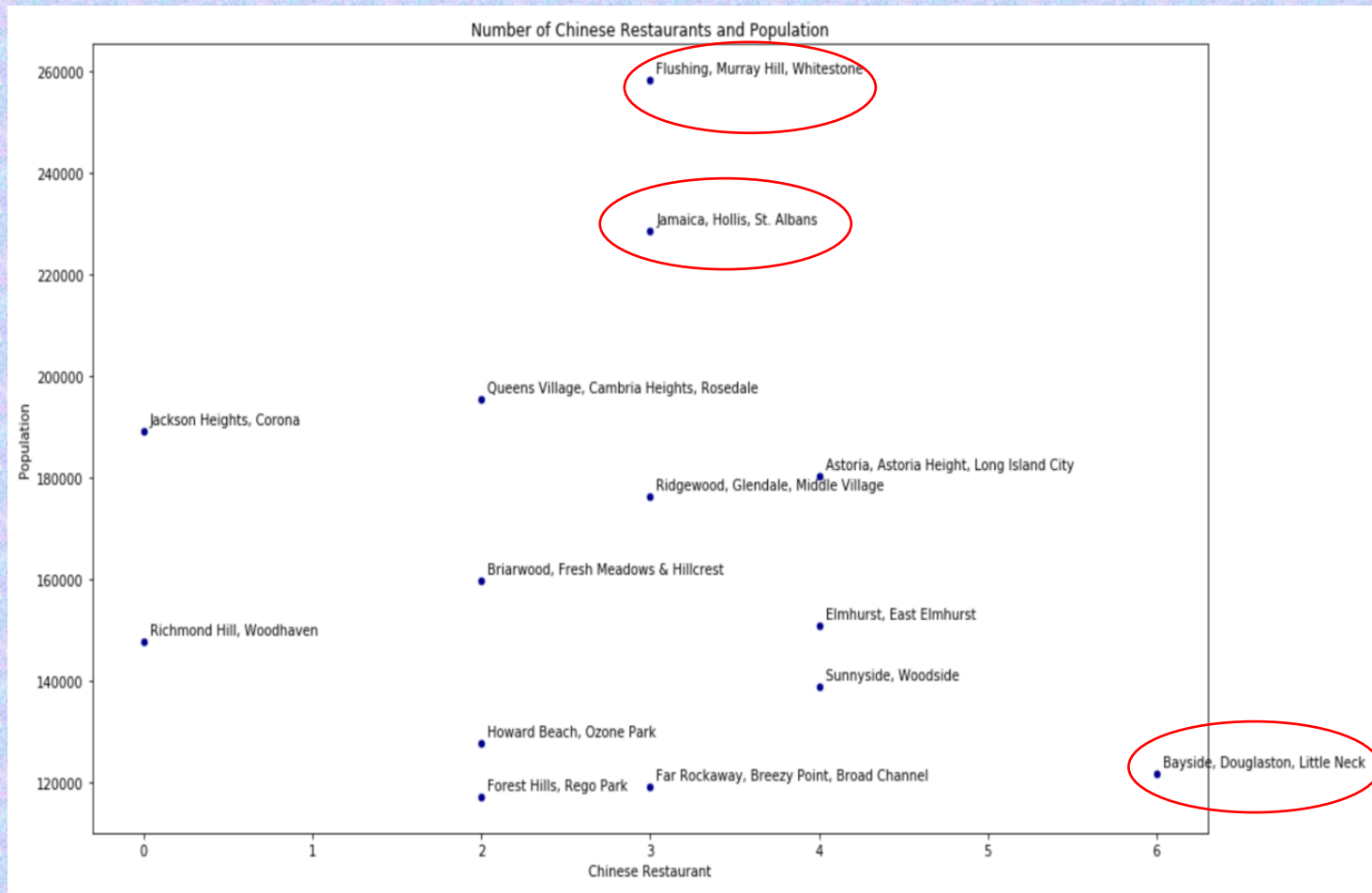
Data Visualization and Analysis

Density of Chinese Restaurants and Population
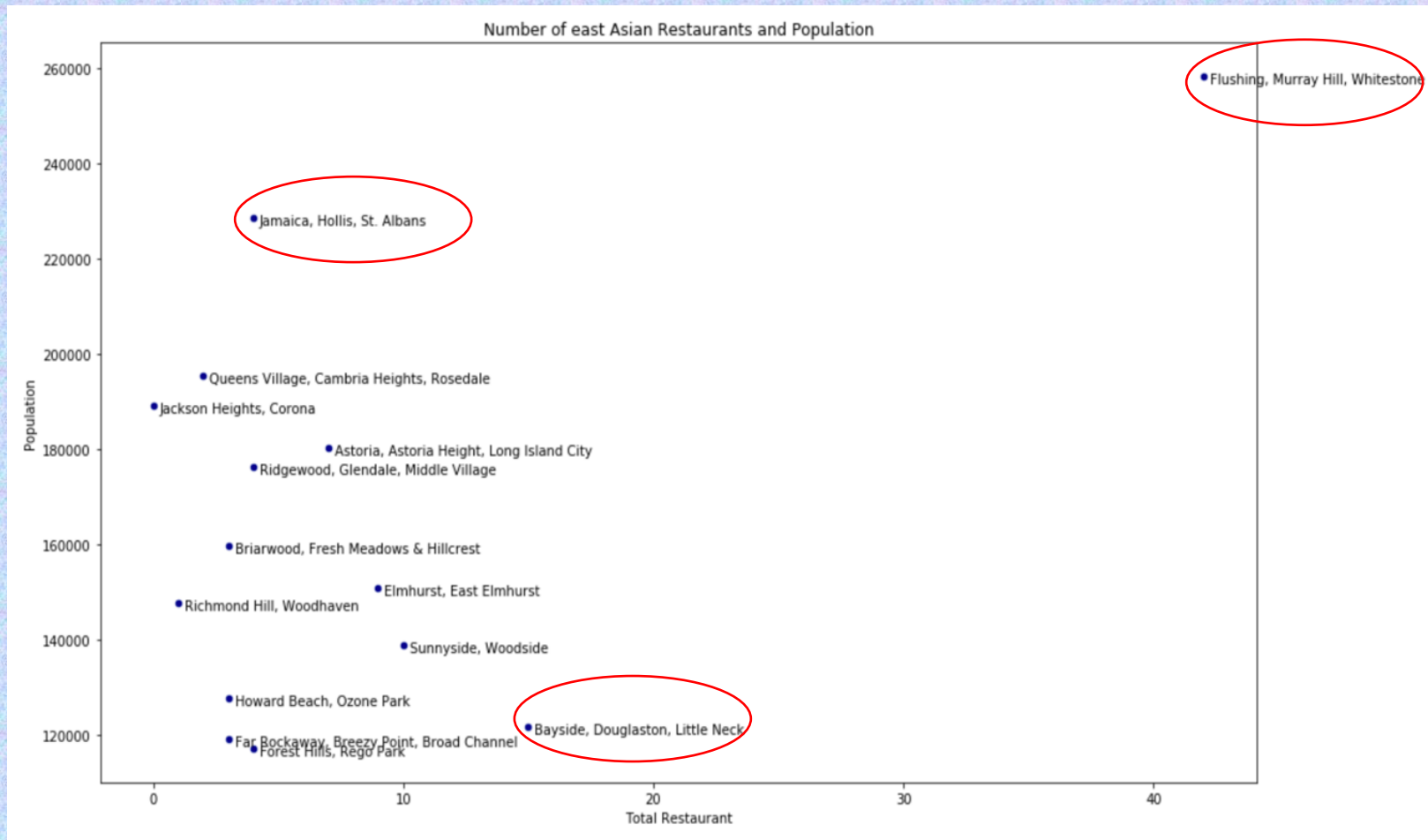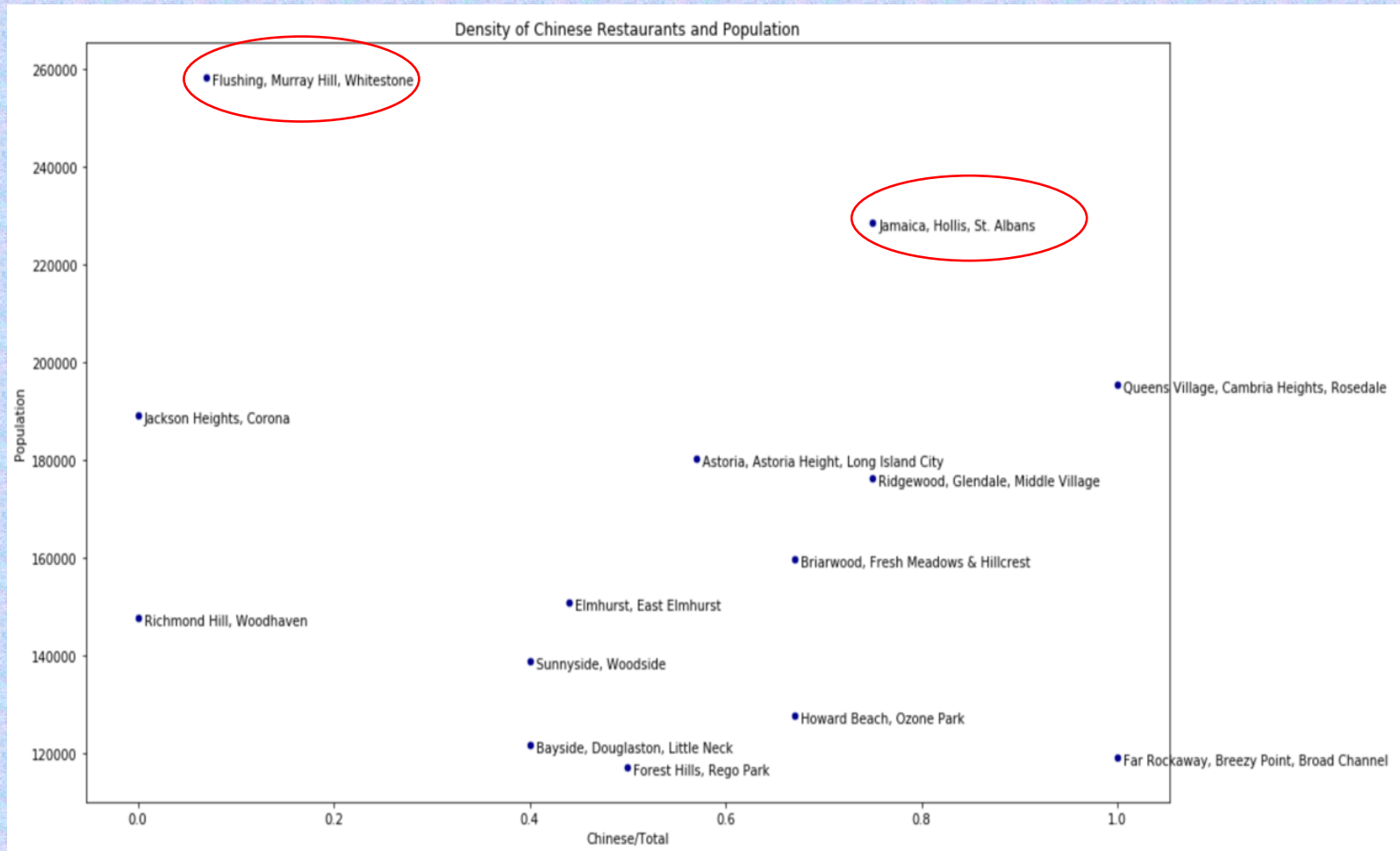
Conclusion

- Data from two data sources
- Data cleaning, exploration, and visualization
- Data analysis using scattered plots
- The neighborhood 'Flushing, Murray Hill, and Whitestone' is recommended

Future Work

- Collect more data
  - Revenue of both the Chinese and east Asian restaurants
  - Numbers of Chinese population in Queens's neighborhoods
  - Demographics of visitors to Queens borough
- More data understanding, data cleaning, model training, data visualization, and analysis