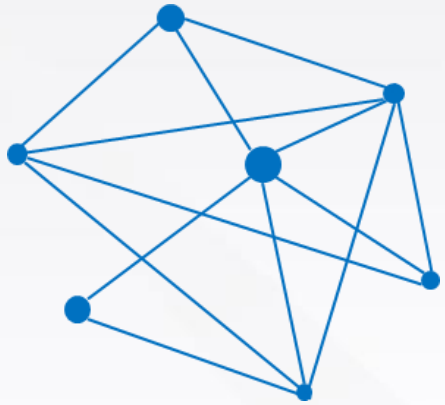
Compare the neighborhoods of New York and Toranto



Based on k-means method

CHEN, Erdi

2020/4/11

CONTENTS

Description

CONTENTS

Background:

New York City and Toronto are very diverse and are the financial capitals of their respective countries.

Suppose you work in Toronto and you're happy with the neighborhood you live in, but you've got a job offer in New York, and you hope to find a neighborhood in New York with a high degree of similarity, and the results of this project will provide that possibility.

CONTENTS

Data Description:

All the data could be found on the Internet. The details as below:

1. **obtain the data of Borough and Neighborhood in Toranto via 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'**
2. **obtain the data of Latitude and Longitude in Toranto via geojson file.**
3. **obtain the data of Borough,Neighborhood Latitude and Longitude in New York via geojson file.**
4. **obtain the data of venues via Foursquare location data**

CONTENTS

- Introduction
- Data Description
- Methodology
- Result

I used BeautifulSoup library to get the neighborhoods in Toranto

Import Data of Toranto

Creat dataframe and obtain the table from website

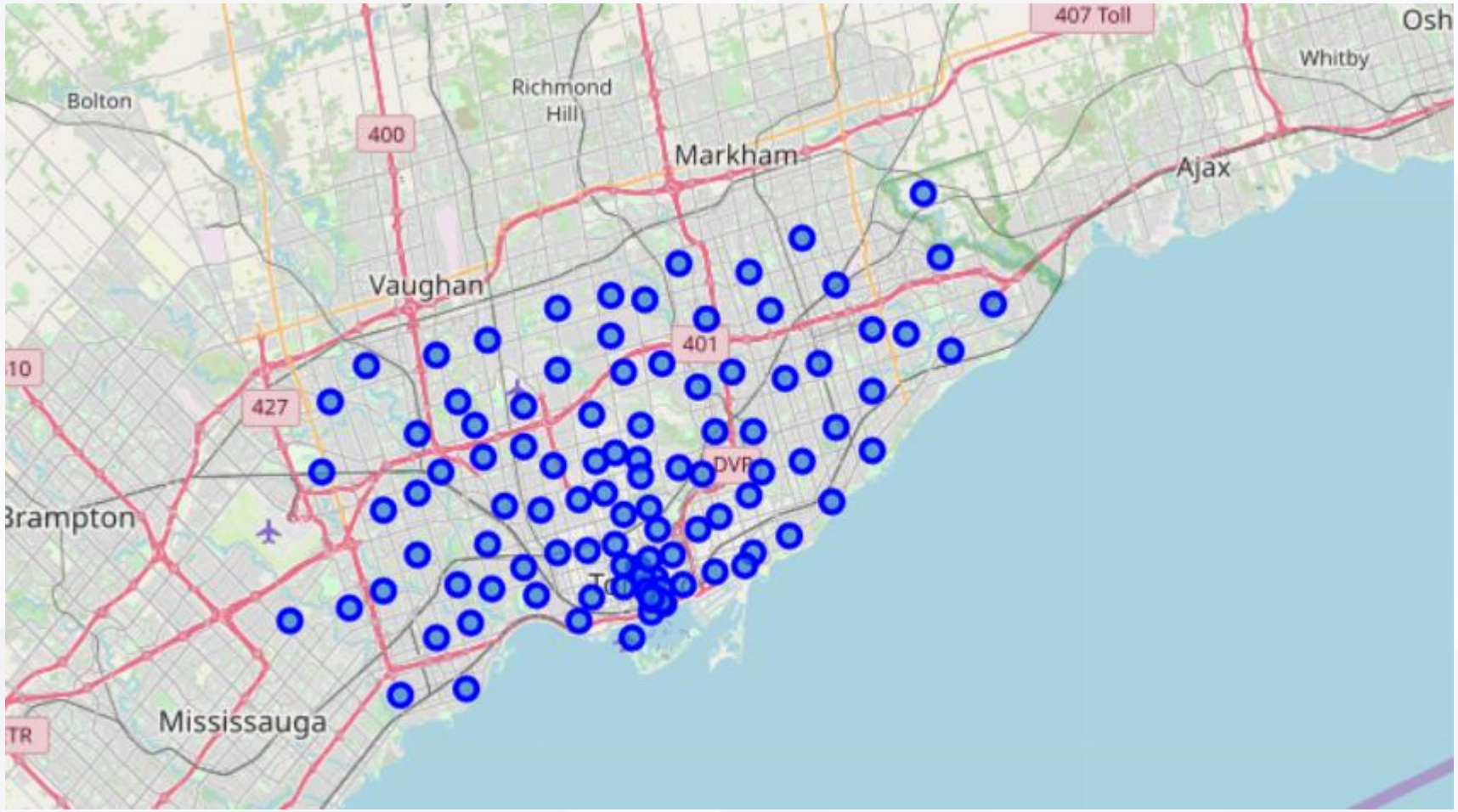
```
url = 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'
res = requests.get(url)
soup = BeautifulSoup(res.text, 'lxml')
tables = soup.select('table')[0]
df_1 = pd.read_html(tables.prettify(),header = 0)[0]
df_1.to_csv('Neighborhoodtable_Toranto.csv')
```

```
df_1.head()
```

	Postal code	Borough	Neighborhood
0	M1A	Not assigned	NaN
1	M2A	Not assigned	NaN
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park / Harbourfront

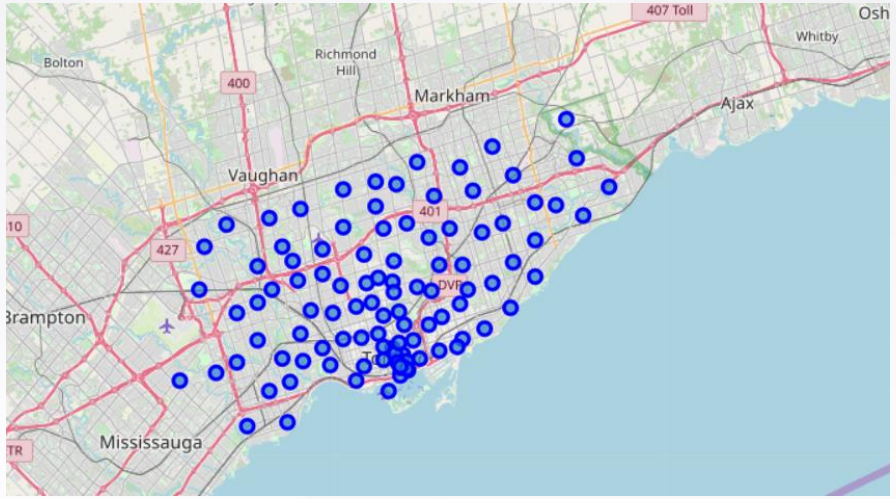
CONTENTS

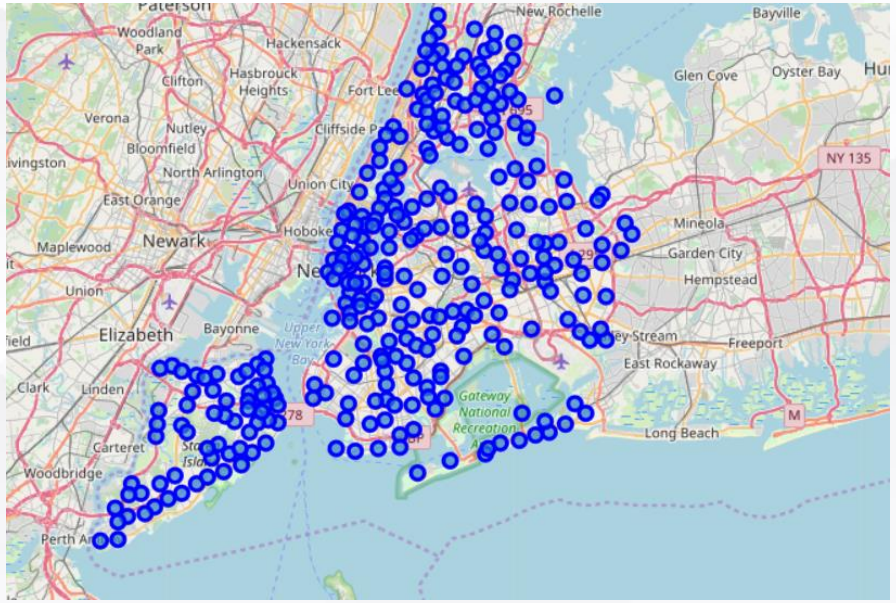
Visualization with Folium library



CONTENTS

Visualization with Folium library





CONTENTS

Merge the data with Latitude and Longitude

	Borough	Neighborhood	Latitude	Longitude
0	North York	Parkwoods	43.753259	-79.329656
1	North York	Victoria Village	43.725882	-79.315572
2	Downtown Toronto	Regent Park , Harbourfront	43.654260	-79.360636
3	North York	Lawrence Manor , Lawrence Heights	43.718518	-79.464763
4	Downtown Toronto	Queen's Park , Ontario Provincial Government	43.662301	-79.389494

```
newyork_neighborhoods.head()
```

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

CONTENTS

Merge the data with Latitude and Longitude

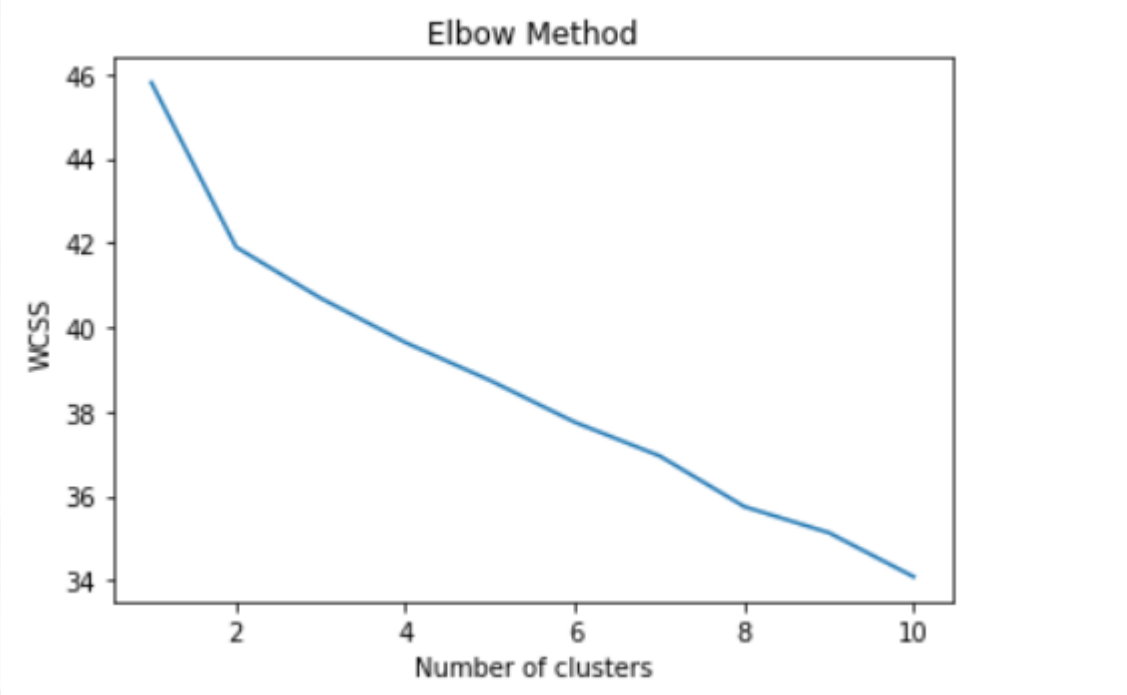
	Borough	Neighborhood	Latitude	Longitude
0	North York	Parkwoods	43.753259	-79.329656
1	North York	Victoria Village	43.725882	-79.315572
2	Downtown Toronto	Regent Park , Harbourfront	43.654260	-79.360636
3	North York	Lawrence Manor , Lawrence Heights	43.718518	-79.464763
4	Downtown Toronto	Queen's Park , Ontario Provincial Government	43.662301	-79.389494

```
newyork_neighborhoods.head()
```

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

CONTENTS

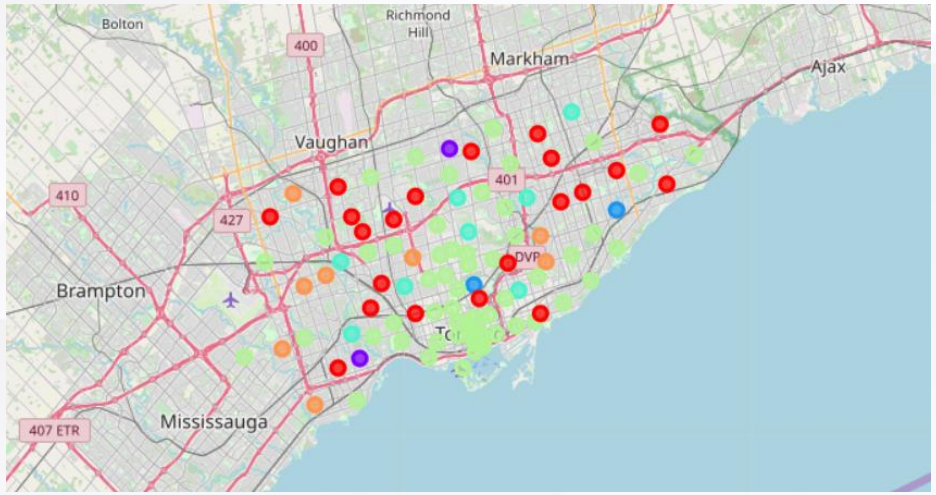
Choose the k value with Elbow Method



CONTENTS

- Introduction
- Data Description
- Methodology
- Result

Visualization with result



Thanks