Capstone Project

Advises for Dinner Takeout and Grocery Shopping During the Pandemic Period

Mei Ding   May 2020

Introduction

During Covid 19, many restaurants and markets are closed.

Where to order dinner while maintaining low risk of exposure to coronavirus?

Where to shop for grocery safely?

We will build a data clustering model to provide advises and guidance.

Data

1. Coronavirus cases
        csv file, uploaded to the data asset on IBM Watson

2. Geospatial data for 70 boroughs in Bergen County, NJ
        This data is retrieved using the geocode library.

3. Venue information
        FourSquare API https://foursquare.com/developers/api

4. Distance between venue and my home
        Calculated using pyproj geod library.

Methodology

K-Means clustering
> Used to cluster restaurants and supermarkets based on # of coronavirus cases, venue rating, and distance to my home

Z-score normalization
> Normalize data before applying K-Means clustering

Folium map
> Visually present the clustering and segmentation. The plot empowers a user to quickly choose a venue.

Data Analysis and Data Modeling Process

Step 1. Data collection and wrangling

Data includes coronavirus cases, geospatial data for 70 boroughs, restaurant and supermarket rating and open status.

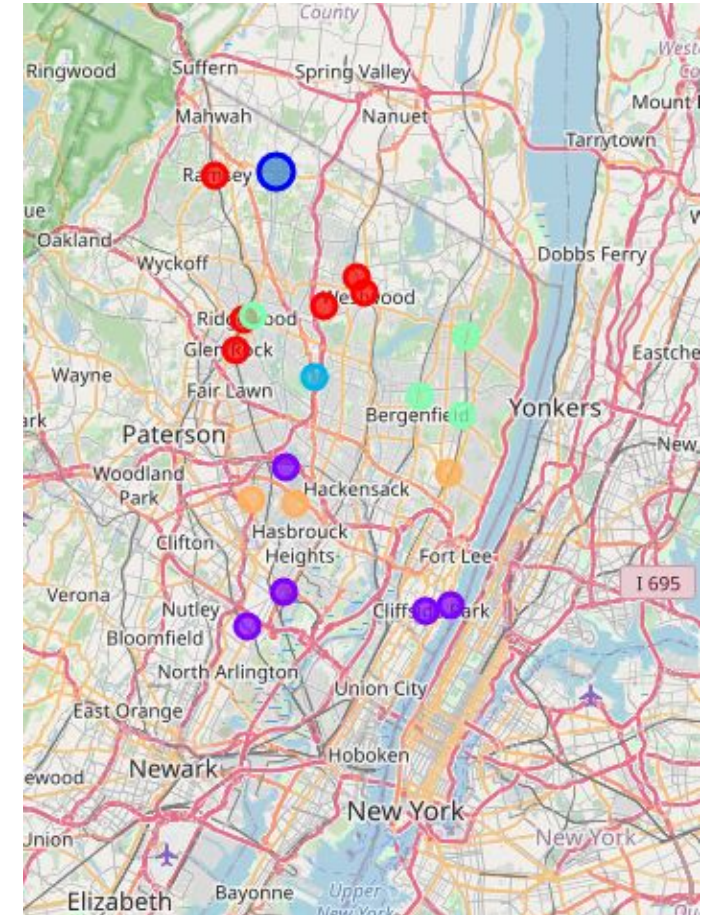Step 2. Build data clustering model for supermarkets

Model is built based on # of coronavirus cases, venue rating, and distance to my home.

Step 3. Build data clustering model for restaurants

Model is built based on # of coronavirus cases, venue rating, and distance to my home.
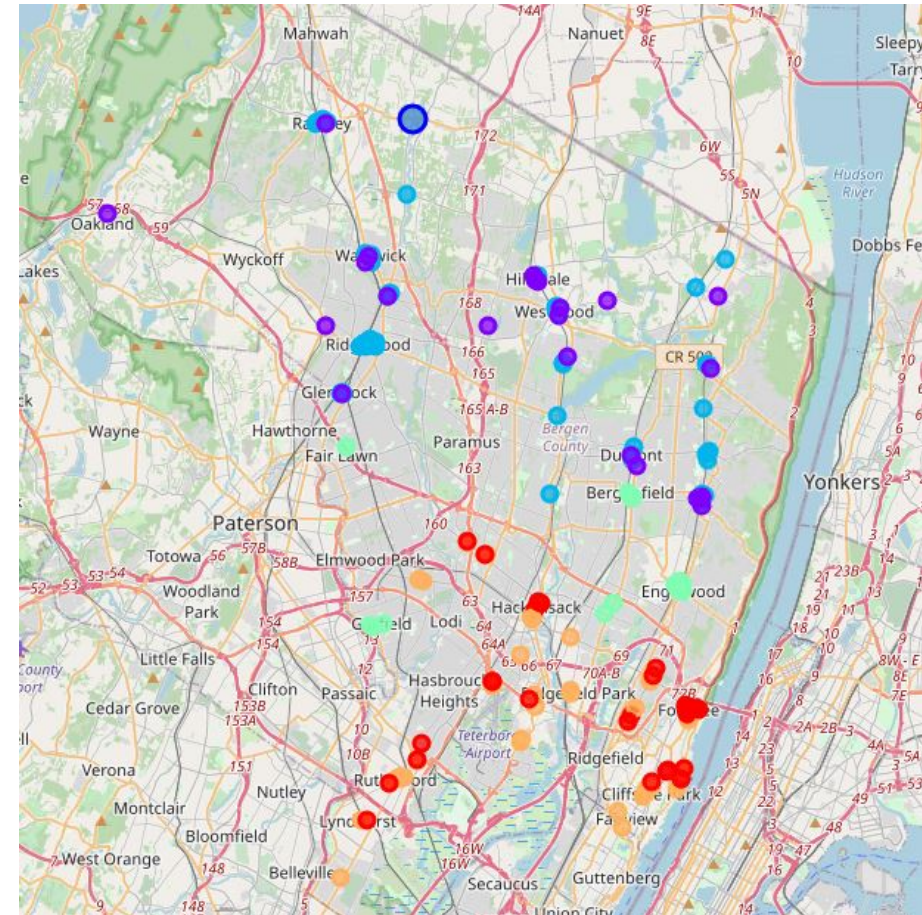
Supermarket Clustering

	Venue	Venue Category	Borough	Coronavirus Positive Cases	Rating	Distance	Venue Latitude	Venue Longitude	Open Status
1	Stop & Shop	Supermarket	Closter	72	5.9	15097	40.969031	-73.962986	Closed until 7:30 AM tomorrow
17	Super Foodtown	Supermarket	Washington Twp.	89	7.3	15097	40.986631	-74.064956	Closed until 7:00 AM tomorrow
7	Kilroy's Wonder Market	Grocery Store	Glen Rock	99	7.6	15093	40.962113	-74.127409	Closed until 8:00 AM tomorrow
8	Kings Food Markets	Grocery Store	Hillsdale	101	7.6	15099	41.001234	-74.041039	Closed until 7:00 AM tomorrow
12	Ramsey Farmer's Market	Farmers Market	Ramsey	109	8.8	15103	41.056289	-74.142212	Closed until 9:00 AM Sunday
4	Whole Foods Market	Grocery Store	Edgewater	119	8.5	15082	40.824706	-73.973828	Closed until 9:00 AM tomorrow
3	Aldi's	Supermarket	East Rutherford	123	7.2	15080	40.831393	-74.092718	Closed until 9:00 AM tomorrow
16	Stop & Shop	Supermarket	Tenafly	141	5.7	15093	40.926974	-73.965214	Closed until 7:30 AM tomorrow
18	Trader Joe's	Grocery Store	Westwood	177	8.3	15098	40.992950	-74.035835	Closed until 10:00 AM tomorrow
13	Whole Foods Market	Grocery Store	Ridgewood	262	7.9	15095	40.978418	-74.122593	Closed until 9:00 AM tomorrow
14	Super Stop & Shop	Supermarket	Ridgewood	262	5.8	15095	40.980982	-74.115379	Closed until 7:30 AM tomorrow
15	Acme	Supermarket	Saddle Brook	267	7.1	15087	40.898949	-74.092395	Closed until 7:00 AM tomorrow
2	Super Stop & Shop	Supermarket	Dumont	278	6.8	15093	40.937412	-73.996043	Closed until 7:30 AM tomorrow
10	ShopRite	Supermarket	Lyndhurst	331	6.9	15078	40.812360	-74.118955	Open until Midnight
0	Food Bazaar Supermarket	Supermarket	Cliffside Park	415	8.0	15081	40.821489	-73.991550	Open until 10:30 PM
9	ShopRite of Lodi	Grocery Store	Lodi	574	7.3	15085	40.879279	-74.084812	Open until 11:59 PM
6	Price Rite Of Garfield	Grocery Store	Garfield	669	6.9	15085	40.881520	-74.116071	Closed until 8:00 AM tomorrow
5	ShopRite of Englewood	Grocery Store	Englewood	690	7.2	15090	40.896316	-73.975366	Open
11	Trader Joe's	Grocery Store	Paramus	831	9.2	15093	40.948220	-74.070867	Closed until 9:00 AM tomorrow

Restaurant Clustering

	Venue	Venue Category	Cluster Labels	Borough	Coronavirus Positive Cases	Rating	Distance	Venue Latitude	Venue Longitude	Open Status
0	Chapala Grill	Mexican Restaurant	3	Bergenfield	726	8.4	15092	40.926046	-73.996719	Closed until 11:30 AM tomorrow
1	Empanada Mania	Empanada Restaurant	3	Bergenfield	726	7.7	15092	40.925710	-73.996562	Closed until Noon tomorrow
2	Nick's Pizza	Pizza Place	3	Bergenfield	726	7.5	15092	40.927700	-73.998180	Closed until 11:00 AM tomorrow
3	Cusinera	Asian Restaurant	3	Bergenfield	726	7.4	15092	40.925610	-73.996678	Closed until 8:00 AM tomorrow
4	Bamboo Grill	Asian Restaurant	3	Bergenfield	726	6.7	15092	40.925985	-73.996648	Closed until 11:00 AM tomorrow

Conclusion

1. Built data clustering and segmentation model to find the venues with good ratings, low coronavirus cases, and short driving distance.

2. Models can be very useful in advising people as to where to go for dinner takeout and grocery shopping while keeping less exposure to coronavirus.

3. Models could be easily adapted to any location/district in the US.