

 data_report.md

Data Description

Data Description

As we need to analyze the restaurant distribution and explore various venues, the neighborhoods in the city of Ottawa need to be collected. First, we scrap the neighborhoods through Wikipedia page. Before we proceed to the next step, we need to clean up the obtained data. Then, we need acquire the corresponding coordinates of each neighborhood. The coordinates information will help us obtain more useful information for this project. After these steps, the data are organized in a format like the following:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	Kennedy Park, Ionview, East Birchmount Park	43.727929	-79.262029
7	M1L	Scarborough	Golden Mile, Clairlea, Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffside, Cliffcrest, Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848

Data Feature

To explore more information of Ottawa, we use Foursquare api to explore the venues in each neighborhood of Ottawa. For each neighborhood, we collect the venue name and the venue category. The samples of collected data is shown below:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Blackburn Hamlet / Pine View / Sheffield Glen	45.424293	-75.598500	Royal Oak	45.427109	-75.601272	Pub
1	Blackburn Hamlet / Pine View / Sheffield Glen	45.424293	-75.598500	Pineview Golf course	45.423598	-75.598886	Golf Course
2	Blackburn Hamlet / Pine View / Sheffield Glen	45.424293	-75.598500	Pineview Park	45.423603	-75.598869	Park
3	Blackburn Hamlet / Pine View / Sheffield Glen	45.424293	-75.598500	Circle K	45.426962	-75.600981	Convenience Store
4	Blackburn Hamlet / Pine View / Sheffield Glen	45.424293	-75.598500	Pineview Pizza	45.427197	-75.601278	Pizza Place
5	Britannia /Whitehaven / Bayshore / Pinecrest	45.361676	-75.784503	Dairy Queen	45.363599	-75.790115	Ice Cream Shop
6	Britannia /Whitehaven / Bayshore / Pinecrest	45.361676	-75.784503	Biagio's Italian Kitchen	45.364543	-75.788269	Italian Restaurant
7	Britannia /Whitehaven / Bayshore / Pinecrest	45.361676	-75.784503	Wendy's	45.363739	-75.786848	Fast Food Restaurant
8	Britannia /Whitehaven / Bayshore / Pinecrest	45.361676	-75.784503	Subway	45.364652	-75.785944	Sandwich Place
9	Britannia /Whitehaven / Bayshore / Pinecrest	45.361676	-75.784503	Shoppers Drug Mart	45.362917	-75.790231	Pharmacy

Data usage

The collected data will be used in the following way:

1. I will explore the distribution of the restaurant and analyze the influence of the existing restaurants.
2. Then, we show the influence area of each restaurant. According to this information, we can choose the locations which are not covered by the current restaurant as candidates.
3. I will cluster these restaurants and analyze the restaurant category of each cluster. Then, according to the ranked restaurant category, the best restaurant type in each cluster can be selected.