

 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	K1B	Ottawa	Blackburn Hamlet / Pine View / Sheffield Glen	45.424293	-75.598500
1	K2B	Ottawa	Britannia /Whitehaven / Bayshore / Pinecrest	45.361676	-75.784503
2	K4B	Ottawa	Navan	45.429358	-75.470225
3	K1C	Ottawa	Orleans	45.427085	-75.545548
4	K2C	Ottawa	Queensway / Copeland Park / Central Park / Bel...	45.381048	-75.710576

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

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.