


MARKET EXPANSION USING DATA SCIENCE

by Pratham Sharma


Data

The data that will be used for this business use-case is location data for neighborhoods in Ontario, Canada. This data will be collected through Wikipedia, along with the Foursquare API.

The Wikipedia page “Postal codes in Canada” gives details and list of all postal codes in Canada. Since we only need to focus on the province of Ontario, only postal codes within Ontario will be considered. These postal codes begin with the letters K, L, M, N and P.



Canadian postal codes



NL	NS	PE	NB	QC				ON				MB	SK	AB	BC	NU/NT	YT
A	B	C	E	G	H	J	K	L	M	N	P	R	S	T	V	X	Y

Therefore, data for all neighborhoods in Ontario, including their names, respective boroughs and postal codes will be collected from Wikipedia by web scraping.

Further, data about all different venues in each neighborhood will be collected using the Foursquare Places API, using latitude and longitude values, which will be collected using geospatial data.

The final data will include the following:

1. Neighborhood name
2. Borough name
3. Postal Code
4. Latitude and longitude values
5. Venues in each neighborhood

This data will be analyzed to provide a solution to the discussed business problem.