

DATA REPORT

CAPSTONE PROJECT



Jabin Choi

Coursera – IBM Data Science

12/30/2019

CONTENTS

PROJECT PURPOSE	2
WHAT IS FOURSQUARE API?	2
DATA APPROACH	2

PROJECT PURPOSE

The aim of this project is to clean and analyze features between neighbourhoods by sending out the names of them. The features mentioned here contain the mean of housing price and school rates etc. The result of data analysis helps people who are considering moving to Scarborough in Canada.

WHAT IS FOURSQUARE API?

In this project, dataset related to the average of housing price and school ratings among the venues registered in the app called Foursquare.

Foursquare is a social location service that provides users to explore the location around the world. The app enables users to download it through iPhone, Blackberry or Android devices and then connects their account to other social media accounts.

Foursquare API itself is a RESTful set of addresses to which developers can send requests. Users has nothing to download onto their server and it just allows developers to interact with the Foursquare platform.

The major purpose of using Foursquare API is to gather data source as it has a database of millions of places and the API provides the ability to perform location search and sharing as well as details on a business.

For more details, please refer to the following url:

** Foursquare URL: <https://foursquare.com>

DATA APPROACH

First, data will be loaded from WIKI url:

'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'

The data table looks like the followings:

Out[4]:

	Postalcode	Borough	Neighbourhood
0	0	0	0
1	M1A	Not assigned	Not assigned
2	M2A	Not assigned	Not assigned
3	M3A	North York	Parkwoods
4	M4A	North York	Victoria Village

Data cleaning:

- Drop meaningless contents in the table – meaning ‘Not assigned’ and None values
- Group and re-index the table
- Add two columns called ‘Latitude’ and ‘Longitude’

Once data cleaning is completed, the final data table before analyzing might look as follows:

Out[10]:

	Postalcode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.811525	-79.195517
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.785665	-79.158725
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.765815	-79.175193
3	M1G	Scarborough	Woburn	43.768369	-79.217590
4	M1H	Scarborough	Cedarbrae	43.769688	-79.239440

Data from Foursquare API:

To gain datasets of various venues in different neighborhoods of that specific borough – Scarborough, we need to use ‘Foursquare’ locational information. The information contains venues name, locations, menus and photos.

- Neighbourhoods
- Neighbourhoods Latitude / Longitude
- Venues
- Venue names such as restaurants or stores
- Venue Latitude / Longitude
- Venue Category

In this project, we will extract information that we need such as venue names, categories and latitude / longitude. Once the columns are obtained from through the API, then it might look as follows if having chosen the radius to be 100 meters:

Out[25]:

	name	categories	lat	lng
0	SEPHORA	Cosmetics Shop	43.775017	-79.258109
1	Disney Store	Toy / Game Store	43.775537	-79.256833
2	American Eagle Outfitters	Clothing Store	43.776012	-79.258334
3	DAVIDsTEA	Tea Room	43.776613	-79.258516
4	Hot Topic	Clothing Store	43.775450	-79.257929