Capstone Project

Ideal location for Indian Restaurant in Dublin

Introduction/Business Problem

- Stakeholders have hired me to investigate of into which neighbourhood in the city of Dublin is the best to open Indian Restaurant. As a data scientist, I'll be developing a model to get some recommendation on which are to most ideal locations to invest.
- I will be taking a city (Toronto), which is similar to Dublin in many ways. Toronto is also more developed in some categories, which Dublin is hoping to do in coming years. So these are one of the reasons investors are interested in opening a restaurant.
- With the growing diverse population, a Indian Restaurant investment will be ideal setup if I am able to pin-point which are the best Neighbourhood to open this venture.

Data

- Web scrapping to obtain the data that is in the table of postal codes from https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	Rouge ,Malvern
1	M1C	Scarborough	Highland Creek ,Rouge Hill ,Port Union
2	M1E	Scarborough	Guildwood ,Morningside ,West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae
5	M1J	Scarborough	Scarborough Village
6	M1K	Scarborough	East Birchmount Park ,Ionview ,Kennedy Park
7	M1L	Scarborough	Clairlea ,Golden Mile ,Oakridge
8	M1M	Scarborough	Cliffcrest ,Cliffside ,Scarborough Village West
9	M1N	Scarborough	Birch Cliff ,Cliffside West

- Get the latitude and longitude for each postcodes from http://cocl.us/Geospatial_data
- Combine both datasets to contains Latitude and longitude for neighbourhoods.

	PostalCode	Borough	Neighbourhood	Latitude	Longitude
0	M1B	Scarborough	Rouge ,Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek ,Rouge Hill ,Port Union	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

- Downloaded Dublin town names with longitude and latitude from https://www.irelandtownslist.com

name	eircode	latitude	longitude	postal_town
Sutton South	D13	53.37021	-6.09392	Dublin
Swords	K67	53.45575	-6.21975	Swords
Swords Demesne	K67	53.45626	-6.21751	Swords
Swords Glebe	K67	53.45750	-6.22520	Swords
Tallaght	D24	53.28881	-6.35571	Dublin

- Using Foursquare API, we will get nearby venues to each neighbourhood and will also group it by ‘Indian Restaurant’.

	Neighborhood	American Restaurant	Amphitheater	Aquarium	Art Gallery	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	Auto Dealership	BBQ Joint	...	Toy / Game Store	T
0	Adelaide ,King ,Richmond	1	0	1	2	1	0	0	0	0	...	0	0
1	Berczy Park	1	0	1	2	0	0	1	0	0	...	0	0
2	Brockton ,Exhibition Place ,Parkdale Village	1	0	0	2	2	2	0	0	1	...	0	0
3	Business Reply Mail Processing Centre 969 East...	3	0	0	0	0	1	1	0	1	...	1	0
4	CN Tower ,Bathurst Quay ,Island airport ,Harbo...	1	0	1	2	0	0	0	0	0	...	0	0

5 rows × 177 columns

```
toronto_grouped[['Neighborhood','Indian Restaurant']]
```

	Neighborhood	Indian Restaurant
0	Adelaide ,King ,Richmond	0
1	Berczy Park	0
2	Brockton ,Exhibition Place ,Parkdale Village	0
3	Business Reply Mail Processing Centre 969 East...	1
4	CN Tower ,Bathurst Quay ,Island airport ,Harbo...	0
5	Cabbagetown ,St. James Town	0

- Create a model that uses number of venues in neighbourhoods except Indian restaurants inputs and number of Indian restaurant as outputs. We'll be building predictive model with SVR algorithm.

- Then use Dublin neighbourhood as target to predict 10 ideal locations to start Indian Restaurant

Methodology

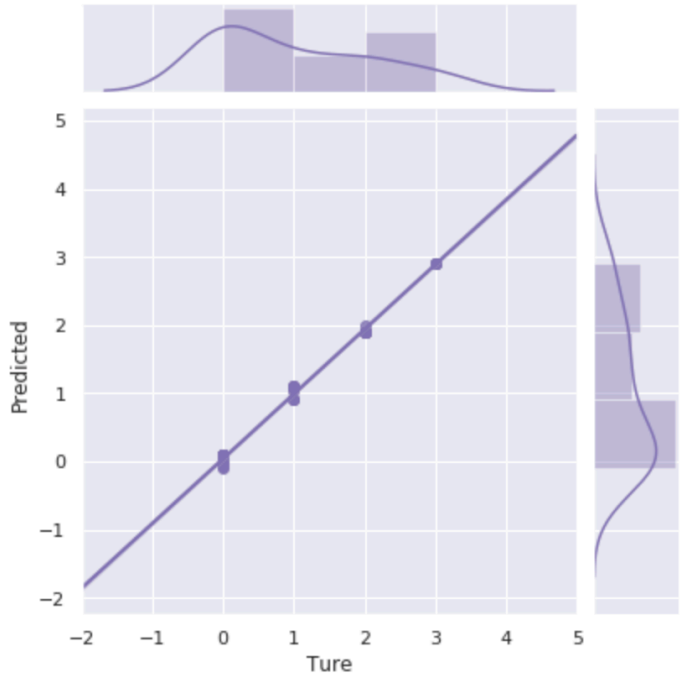
Used the following steps:

- Learn how other venues around the Indian restaurant in a big city such as Toronto.
- Build a model that can pride number of Indian Restaurants given the environmental venues information.
- Predict number of Indian restaurant in Dublin.
- Compare current existing numbers and recommend ideal locations.

Build Model

- Create a model that uses number of venues in neighbourhoods except Indian restaurants inputs and number of Indian restaurant as outputs. We'll be building predictive model with SVR algorithm.

- Train the dataset the plot prediction from the model and true values



- Use foursquare API and get venue information for Dublin towns similar to Toronto

	Neighborhood	Airport	Airport Lounge	Airport Service	American Restaurant	Aquarium	Arcade	Art Gallery	Art Museum	Arts & Crafts Store	...	Waterfront	Whisky Bar
0	Abbeyville	1	1	0	1	0	0	0	0	0	...	0	0
1	Abbotstown	0	0	0	1	0	0	0	0	0	...	0	0
2	Adamstown	1	0	0	0	0	0	0	0	0	...	0	0
3	Aderrig	1	0	0	0	0	0	0	0	0	...	0	0
4	Aghfarrell	0	0	0	0	0	0	0	0	0	...	0	0

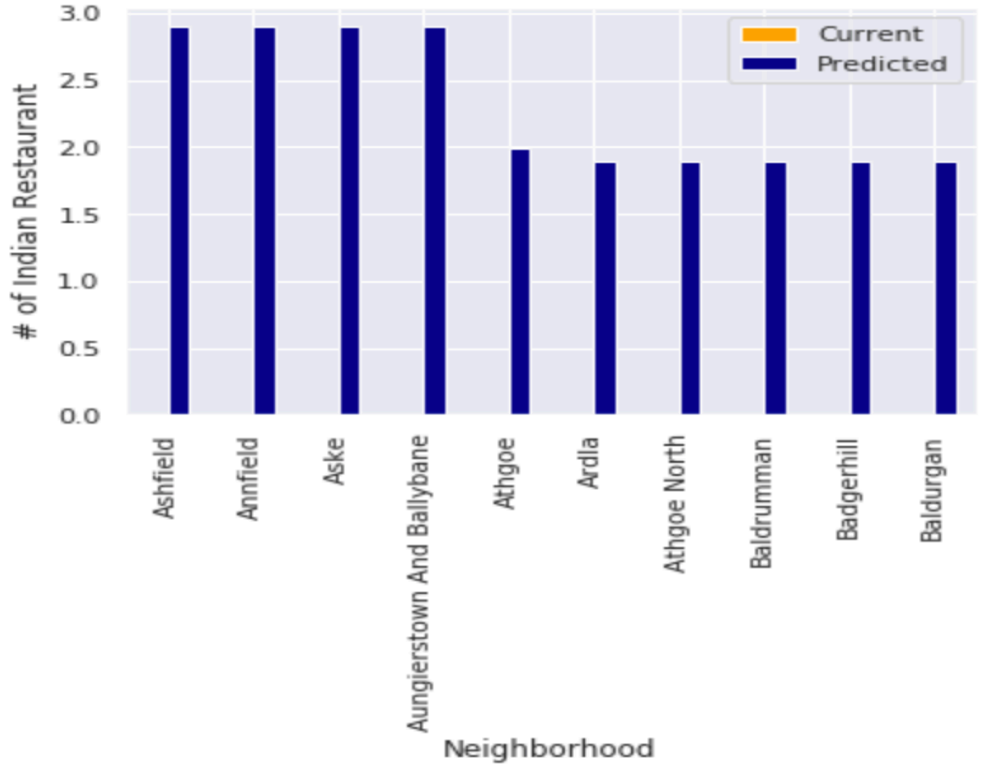
5 rows × 223 columns

```
dublin_grouped[['Neighborhood','Indian Restaurant']]
```

	Neighborhood	Indian Restaurant
0	Abbeyville	2
1	Abbotstown	1
2	Adamstown	0
3	Aderrig	0
4	Aghfarrell	0
5	Allagour	1

Predicted locations using the model

- Using this model, it provided best 10 ideal location in the city of Dublin for our investors to start an Indian Restaurants.



Conclusion

- Using Foursquare API we extracted venue locations for both cities
- SVR algorithm was used to build the model
- The model was able to get the investors the top ideal locations for starting an Indian Restaurant.