IBM APPLIED DATA SCIENCE CAPSTONE

THE ADVENTURE OF SIMULATING WHERE TO OPEN A NEW COFFEE SHOP, CONSIDERING THE INFLUENCE OF REGIONS WITH TRANSPORT STATIONS

GUSTAVO SOUSA

GUSTAVOMPSOUSA@GMAIL.COM

LINKEDIN.COM/GUSTAVOMPSOUSA

INTRODUCTION

* A RESEARCH TO A FICTIONAL COFFEE STORE

* NEW YORK AND TORONTO

* NEAR TRANSPORT STATIONS.

BUSINESS PROBLEM

* DEFINING THE PLACE TO BEGIN A NEW BUSINESS ACTIVITY IS CRITICAL

* VARIABLES: ECONOMIC, CULTURAL, SOCIAL

* TWO POPULAR KINDS OF VENUES: COFFEE SHOPS AND TRANSPORT STATIONS

* ADVERTISING CAMPAIGN ON TRANSPORT STATIONS

DATA

* DATASETS OF NEIGHBORHOODS (NAME, POSTAL CODE)

* LIST OF VENUES DISTRIBUTED BY QUANTITATIVE RELEVANCE, GROUPED BY NEIGHBORHOOD

* GEOPY: PYTHON CLIENT FOR GEOCODING WEB SERVICE - OBTAINING THE COORDINATES BY POSTAL CODE

* FOURSQUARE: SERVICE PROVIDES ACCESS TO A HUGE AMOUNT OF DATA RELATED TO POINTS OF INTEREST, NAMED BY VENUES, IN ALMOST EVERY CITY AROUND THE GLOBE

DATA

* SAMPLES OF NEIGHBORHOOD DATA, POPULATED WITH COORDINATES

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585
5	Bronx	Kingsbridge	40.881687	-73.902818
6	Manhattan	Marble Hill	40.876551	-73.910660
7	Bronx	Woodlawn	40.898273	-73.867315
8	Bronx	Norwood	40.877224	-73.879391
9	Bronx	Williamsbridge	40.881039	-73.857446

|	Postal Code	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 W...	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff / Cliffside West	43.692657	-79.264848

New York

Toronto

METHODOLOGY

* THE FIRST STEP: VISUALIZE THE FULL DISTRIBUTION OF NEIGHBORHOODS
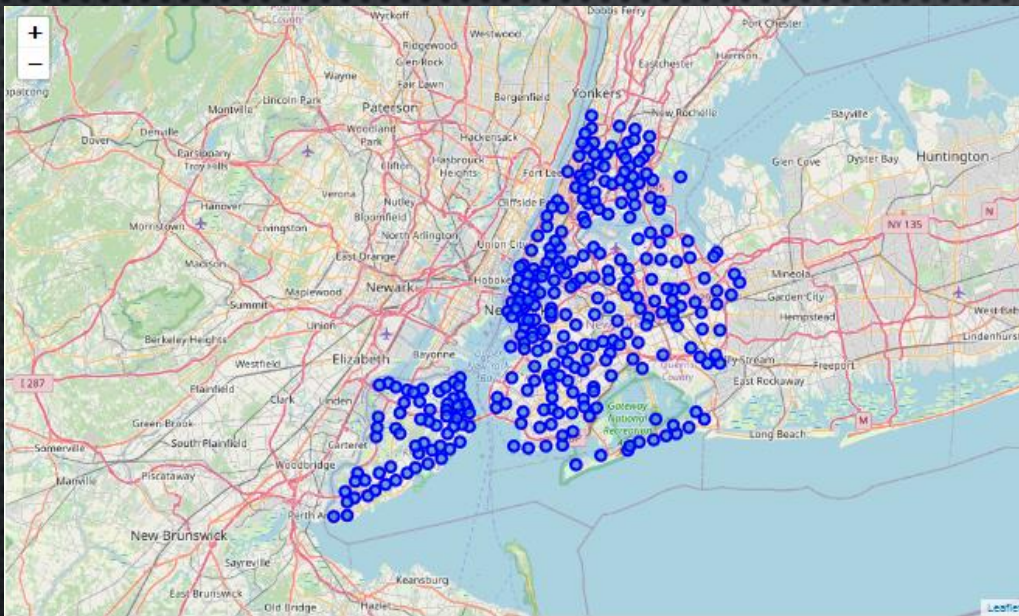
New York
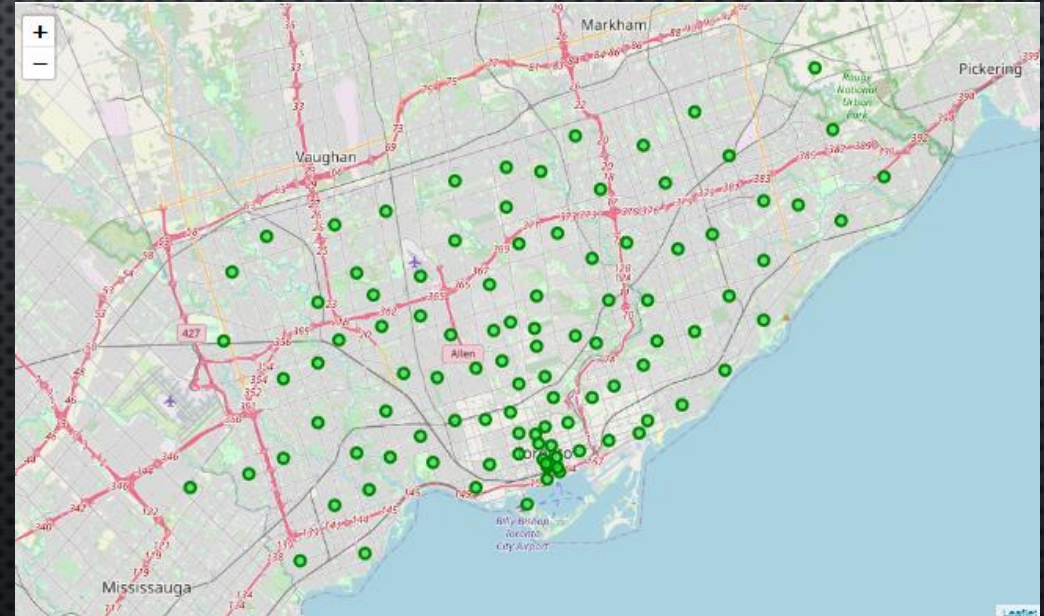
Toronto

METHODOLOGY

* RETRIEVE THE PLACES OF INTEREST GROUPED BY NEIGHBORHOOD (FOURSQUARE + PYTHON)

* NORMALIZATION: REAL AND COMPARATIVE VIEW OF THE MOST REPRESENTATIVE VENUES WITHIN EACH NEIGHBORHOOD

* RANK THE MOST FREQUENT PLACES OF INTEREST IN EACH NEIGHBORHOOD (TEN MOST FREQUENT IN DESCENDING ORDER)

METHODOLOGY

* RESULTING DATASETS OF FIVE MOST FREQUENT VENUES, GROUPED BY NEIGHBORHOOD

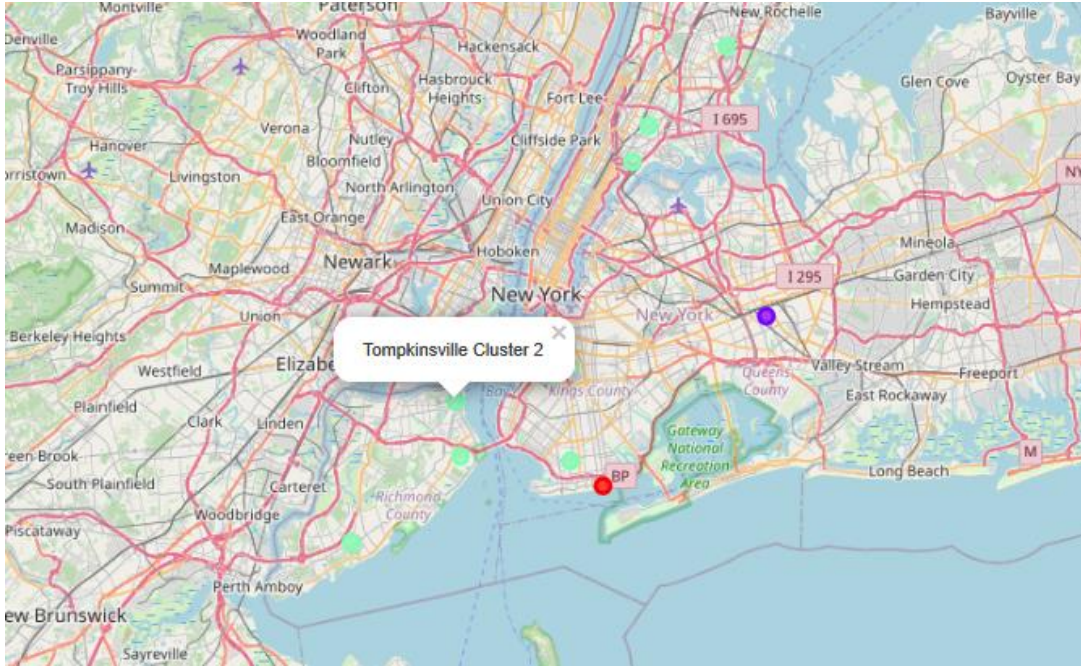|	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Allerton	Pizza Place	Deli / Bodega	Cosmetics Shop	Department Store	Supermarket
1	Annadale	Pub	Restaurant	Diner	Pizza Place	Train Station
2	Arden Heights	Pizza Place	Pharmacy	Pool	Playground	Coffee Shop
3	Arlington	Bus Stop	Intersection	American Restaurant	Deli / Bodega	Grocery Store
4	Arrochar	Bus Stop	Italian Restaurant	Deli / Bodega	Cosmetics Shop	Athletics & Sports
5	Arverne	Surf Spot	Sandwich Place	Metro Station	Bus Stop	Playground
6	Astoria	Middle Eastern Restaurant	Bar	Pizza Place	Mediterranean Restaurant	Greek Restaurant
7	Astoria Heights	Chinese Restaurant	Plaza	Laundromat	Bakery	Hostel
8	Auburndale	Furniture / Home Store	Supermarket	Pet Store	Pharmacy	Toy / Game Store
9	Bath Beach	Chinese Restaurant	Pharmacy	Gas Station	Sushi Restaurant	Italian Restaurant

New York

	Neighborhood	1st Most Common	2nd Most Common	3rd Most Common	4th Most Common	5th Most Common
0	Agincourt	Latin American Restaurant	Skating Rink	Lounge	Breakfast Spot	Women's Store
1	Alderwood / Long Branch	Pizza Place	Pharmacy	Skating Rink	Pool	Sandwich Place
2	Bathurst Manor / Wilson Heights / Downsview No...	Coffee Shop	Bank	Gift Shop	Pizza Place	Sandwich Place
3	Bayview Village	Japanese Restaurant	Chinese Restaurant	Bank	Café	Electronics Store
4	Bedford Park / Lawrence Manor East	Sandwich Place	Italian Restaurant	Restaurant	Coffee Shop	Women's Store
5	Berczy Park	Coffee Shop	Cocktail Bar	Bakery	Cheese Shop	Café
6	Birch Cliff / Cliffside West	College Stadium	Skating Rink	General Entertainment	Café	Empanada Restaurant
7	Brockton / Parkdale Village / Exhibition Place	Café	Breakfast Spot	Coffee Shop	Nightclub	Burrito Place
8	Business reply mail Processing CentrE	Gym / Fitness Center	Spa	Auto Workshop	Brewery	Burrito Place
9	CN Tower / King and Spadina / Railway Lands / ...	Airport Service	Airport Lounge	Airport Terminal	Sculpture Garden	Harbor / Marina

Toronto

METHODOLOGY

* FILTERING, STEP1: KEEP ONLY THE NEIGHBORHOODS WITH TRANSPORT STATIONS (BUS, TRAIN, METRO) ON THE TOP 10 VENUES

* FILTERING, STEP 2: MAINTAIN ONLY THE NEIGHBORHOODS WITH CAFETERIA-LIKE VENUES (CAFÉ, COFFEE SHOPS, DONUT, BAKERY) ON THE TOP 10 VENUES

Neighborhood		1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Annadale	Pizza Place	Bakery	Park	Diner	Sushi Restaurant	Bagel Shop	Train Station	Restaurant	American Restaurant	Sports Bar
1	Claremont Village	Grocery Store	Chinese Restaurant	Pizza Place	Bus Station	Food	Liquor Store	Caribbean Restaurant	Gym	Bakery	Discount Store
2	Eastchester	Caribbean Restaurant	Deli / Bodega	Diner	Bowling Alley	Donut Shop	Metro Station	Bakery	Seafood Restaurant	Fast Food Restaurant	Pizza Place
3	Grasmere	Bus Stop	Bakery	Deli / Bodega	Restaurant	Park	Grocery Store	Pharmacy	Bank	Bagel Shop	Home Service
4	Gravesend	Italian Restaurant	Lounge	Pizza Place	Bus Station	Bakery	Furniture / Home Store	Donut Shop	Sporting Goods Shop	Record Shop	Men's Store
5	Manhattan Beach	Bus Stop	Café	Sandwich Place	Beach	Food	Ice Cream Shop	Playground	Women's Store	Exhibit	Eye Doctor
6	Mott Haven	Gym	Spanish Restaurant	Donut Shop	Pizza Place	Latin American Restaurant	Metro Station	Burger Joint	Bakery	Peruvian Restaurant	Electronics Store
7	South Jamaica	Bus Station	Vegetarian / Vegan Restaurant	Bakery	Supermarket	Caribbean Restaurant	Grocery Store	Sandwich Place	Field	Exhibit	Eye Doctor
8	Tompkinsville	Thrift / Vintage Store	Brewery	Rock Club	Bus Stop	Spanish Restaurant	Supermarket	Café	Caribbean Restaurant	Mexican Restaurant	Gastropub

RESULTS

* DATASET AND CLUSTERED VISUALIZATION FOR *NEW YORK* (8 NEIGHBORHOODS)
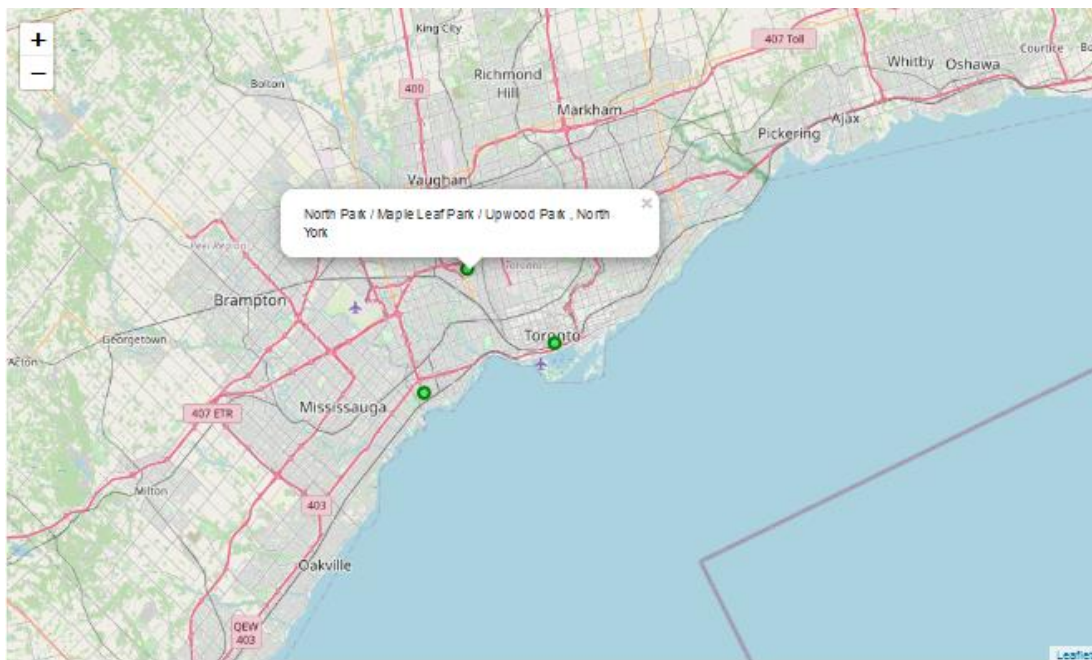
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Alderwood / Long Branch	Pizza Place	Bakery	Park	Diner	Sushi Restaurant	Bagel Shop	Train Station	Restaurant	American Restaurant	Sports Bar
1	North Park / Maple Leaf Park / Upwood Park	Grocery Store	Chinese Restaurant	Pizza Place	Bus Station	Food	Liquor Store	Caribbean Restaurant	Gym	Bakery	Discount Store
2	Toronto Dominion Centre / Design Exchange	Caribbean Restaurant	Deli / Bodega	Diner	Bowling Alley	Donut Shop	Metro Station	Bakery	Seafood Restaurant	Fast Food Restaurant	Pizza Place

RESULTS

* DATASET AND VISUALIZATION FOR *TORONTO* (3 NEIGHBORHOODS) – NO PRACTICAL NEED FOR CLUSTERIZATION

DISCUSSIONS

* TWO APPARENTLY INDEPENDENT DATABASES, WHEN ASSERTIVELY RELATED, CAN BRING NUMEROUS DIFFERENT ANALYSES

* SUBJECTIVE OPPORTUNITY ASSESSMENT

* AFTER ALL THE CLEANING, PREPARATION AND ANALYSIS, ARE WE STILL SUBJECTED TO ASSUMPTIONS AND INFERENCES WHEN STARTING A BUSINESS?

Q & A?

Thank you!