

Go, No Go & What Way...

**A minimal viable product report
on opening a Sushi or Greek
themed restaurant in the
Durham Ontario region**

Background

Oshawa is a city in Southern Ontario, Canada that sits on the Lake Ontario; it's considered part of the Greater Toronto Area (GTA) as it sits 60 kilometers east of Downtown Toronto (according to Google Maps). It is commonly viewed as the eastern book-end of the GTA (while the right book-end stretches out to Oakville and Mississauga regions).

There is a lot of debate about where the GTA ends on the West side, but not much in the East side. Oshawa's roots are tied to the automobile industry, specifically the Canadian division of General Motors (GM), known as General Motors Canada ... but in 2019 they announced a closure of their manufacturing operations in the region

The city of Oshawa has been trying to transform itself over the last decade into a higher-technical location away from 'blue collar' manufacturing. The advent of the University of Ontario Institute of Technology (UOIT), alongside Durham college, give it a unique 'drawing' of talent pools and the ability to create skilled staff back to surrounding business, For example, UOIT has the most sophisticated wind tunnel research infrastructure in Canada (<https://ace.ontariotechu.ca/>); they have one of the world's leader health informatics researchers (Dr. Carolyn McGregor) - <http://hir.uoit.ca/cms/?q=node/6>; and more.



For this reason (among others), the city of Oshawa was experienced significant growth in the recent years. In 2017, Oshawa was considered "*One of the fastest growing economies...*" (<https://www.thestar.com/business/2017/08/03/one-of-canadas-fastest-growing-economies-is-in-oshawa.html>).

In 2018, Oshawa was called "*the fastest growing economy...*" (<https://www.durhamradionews.com/archives/113040>) despite overall all Canadian growth slowing down. In other words, Oshawa has bucked the trend despite the manufacturing (blue collar) job losses.

**"Oshawa will have the fastest
growing economy this year..."**
Conference Board of Canada (2018)

The population and economic growth to Oshawa comes from a living population that draws beyond the traditional boundaries of the city of Oshawa. This area is known "Durham Region" (its more proper name is "The Regional Municipality of Durham"). This region is vast (approximately 2,500 square kilometers) and its government operations are headquartered in Whitby, Ontario.

Because of the vastness of this area, we have been approached by a food and beverage (F&B) venture capital (VC) firm to see if it would make good business sense to launch restaurant operations for either Sushi or Greek style food in Oshawa (and potentially a seed for a chain), or don't enter either market at all; however, we will also consider any region (also referred to as neighborhood into this paper) as well.

Our main data will come from Foursquare reviews which should flush out the various competitors and business districts in this vast area. We will use LAT / LONG data to sub-divide the region, and even pull in exogenous data (like rental activity) to try and gauge the viability of an area of our new investment.

This aim of this research is to create a minimal viable product (MVP) and analysis to not just give some intuition on the aforementioned investment objective, but also lay the ground work to evolve this framework to apply to restaurant selection anywhere. With more budget and time, we would have brought in a richer set of data (other review sites such as Yelp; weather data; health inspection reports, event calendars; and more) that would lead to even richer analysis.



The Data

We will use social ratings (which will also allow us to count the number of restaurants) from Foursquare (a search & discovery services that provides search recommendations of places to go near a user's current location based on users' previous browsing history and check-in history) for both Greek and Sushi style restaurants in the various neighborhoods of the Durham region with a focus on Oshawa. This should allow us to see if the region is over-serviced by these styles of foods, and even if so, are there any good restaurants.

When we call the APIs for data, we actually decided to use a limited 8km radius (8,000 meters). This seems small based on the size of Oshawa and the Durham regions. For substantial analysis we will need more funding from the VC to pay for the planned API calls to the Foursquare data service. In other words, we need more data, and we will have to pay for it. (That said, this report, its methodology, and results should make it a no brainer for our VC firm to fund the account to expand the data set upon which we can perform analysis).

Creating regions for data analysis

Because of the vastness of the Durham region we wanted to set up a number of sub-regions (in this case 8) with a set radius for which we can flush out the existing Sushi and Greek restaurants. While our final data analysis will be limited due to the API calls, this framework to expand the analysis across the vast Durham region is already done ... hence these regions.

We created our regions using Python and latitudinal and longitudinal data from Google Maps and ended up with these regions:

	Neighborhoods	Latitude	Longitude
0	Scugog	44.154083	-78.864576
1	Uxbridge	44.113606	-79.219505
2	Ajax	43.836971	-79.020873
3	Pickering	43.826172	-79.104141
4	Whitby	43.855884	-78.934945
5	Oshawa	43.869820	-78.838396
6	Clarington	43.920869	-78.614387
7	Brock	44.452814	-79.162786

Using this LAT / LONG data we generated a "REGIONS" map (many refer to these regions as *neighborhoods*, but we prefer to use the term regions because neighborhoods are often thought of with a radius defined by city blocks which is too granular for this analysis.

Preprocessing the data for restaurants in a region

When the review data came in from our Foursquare API call, and we organized it via a region's radius, we found that that would get duplicates names; obviously, a restaurant can (and most likely will) have multiple reviews. Since we are first looking to plot the Greek and Sushi restaurants on a map to look at penetration rates, ignoring if they are any good or not, we removed duplicate entries so we could count the restaurants. This output of this data was saved into a CSV file and we made a data frame too.

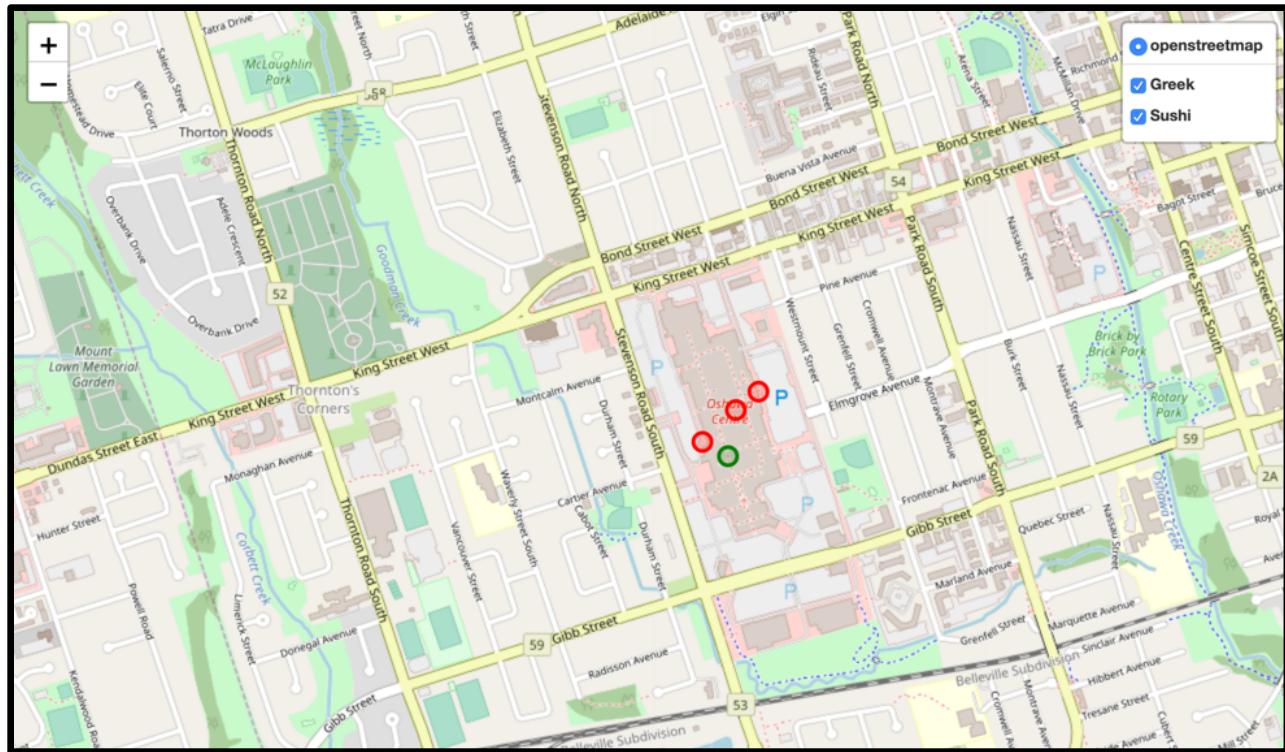
Data helper functions

To aid in our analysis we created a number of data helper functions that would could easily call to get all sorts of information – such as the actual printed reviews, and so on. We did this to make it easier to retrieve data since we are looking at both Greek and Sushi restaurants, we simply call the function.

The Methodology

We used a combination of qualitative and quantitative analysis in our report.

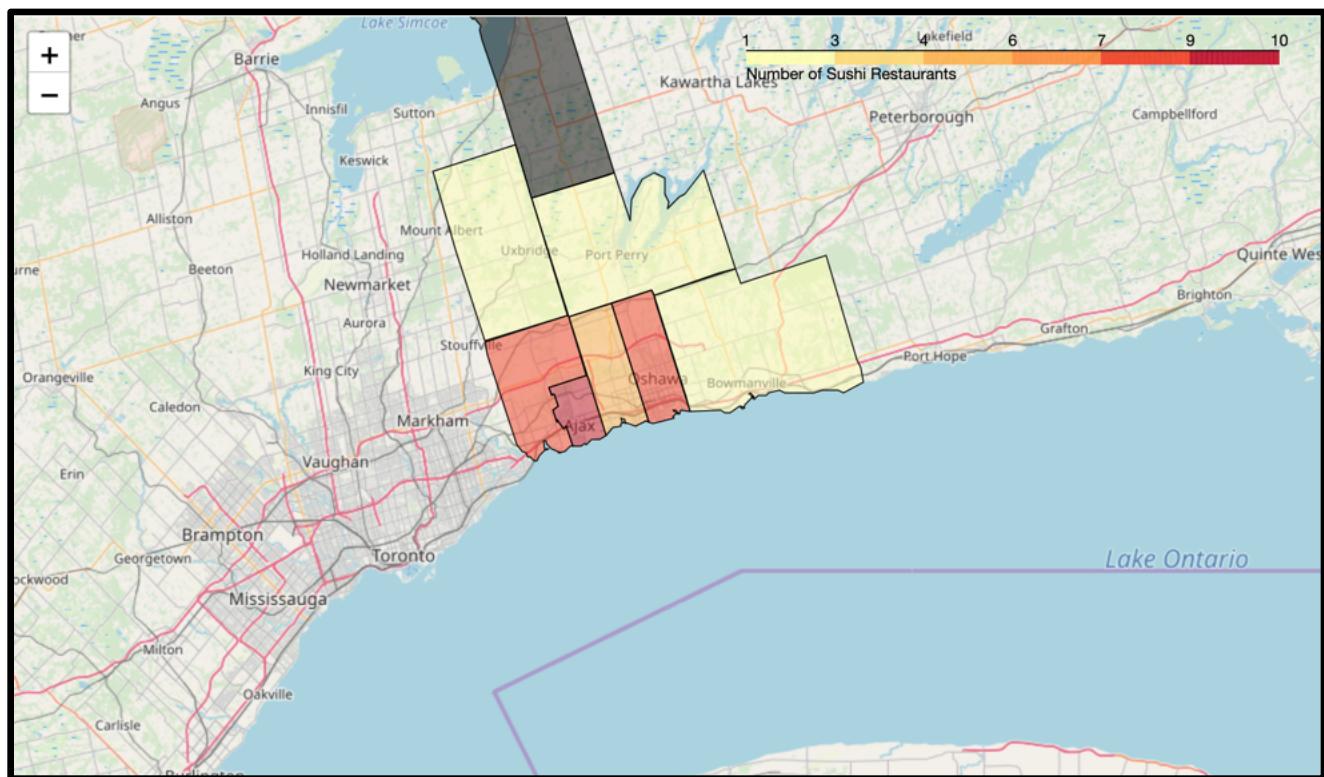
We start out with a visual inspection of the data. We created an interactive map using Folium so our sponsors could browse the various Durham regions. You can zoom in and out of sections of the region BUT also filter for Greek or Sushi restaurants. We felt it important that our VC sponsors be to quickly generate their own hypothesis (you will see there are some of the Durham region that are indeed super-saturated with Sushi restaurants); of course, data will lead the way to a final decision. For example, it could be the case the dense cluster of Sushi restaurants near the Oshawa Center all receive very poor reviews and therefore still present opportunity - that type of conclusion or 'risk' can only be accomplished via quantitative analysis and data.



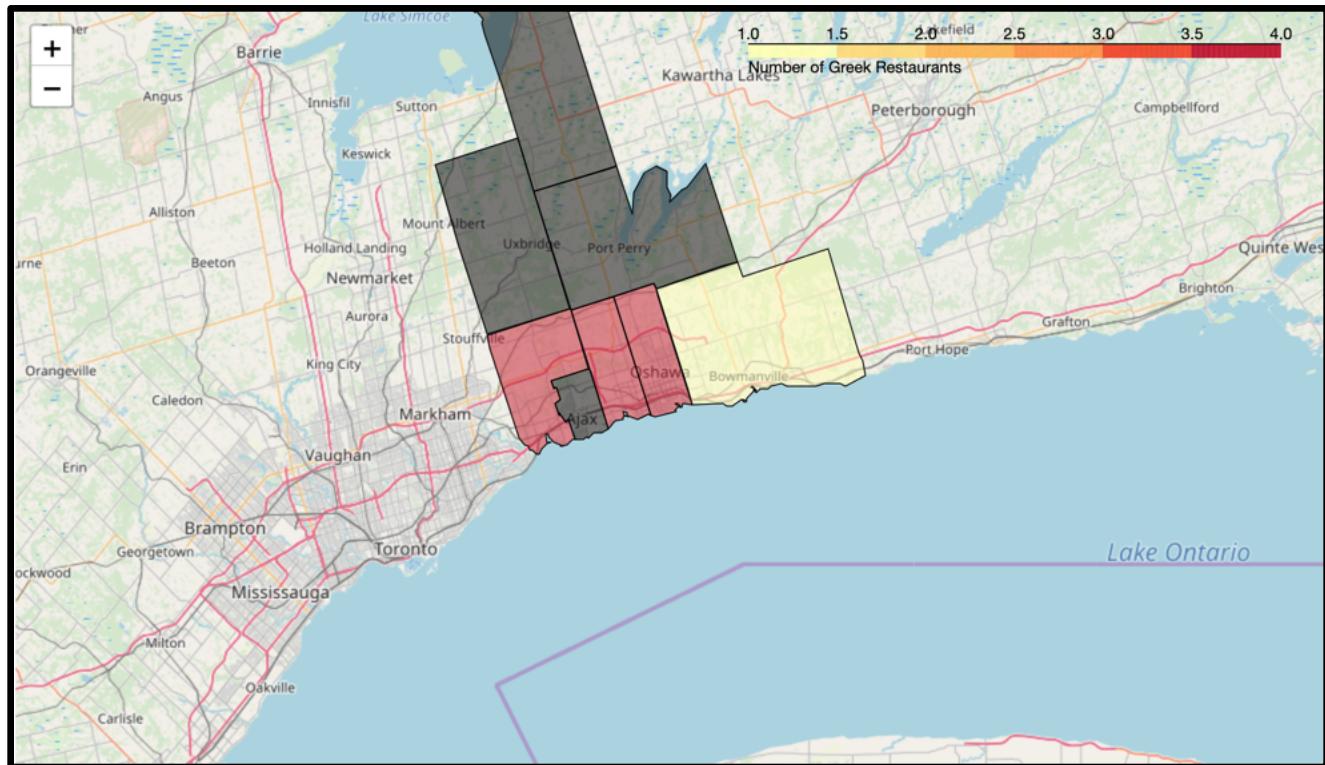
Our VC board was comprised of sponsors who own the different styles of restaurants they fund. There is a VP of Sushi Restaurants and a VP of Greek Restaurants. Of course, they have to make a joint decision, but each has domain knowledge.

To make it easier (and less noisy) we thought it beneficial to generate temperature maps that show the penetration of each style of restaurant across the region; for this we used a Choropleth class to show this 'temperature' scale. (These maps are still interactive from a zoom perspective.)

Sushi Restaurant Penetration in the Durham Region



Greek Restaurant Penetration in the Durham Region



The number of restaurants in a certain Durham region is one indicator where a potential investment may occur, but as previously mentioned, we also want to look at the quality of those locations. After all, if there are 2 Greek food restaurants next door to each other, and both got very low reviews, it stands to reason that there is opportunity here because if both restaurants are going concerns, it means they make enough money to stay open (indicating demand) but the low reviews indicate the supply of the demand is not high quality.

We found that there are not a lot of Greek restaurant ratings that come to us via the Foursquare API in the Durham regions (and none in Oshawa for that matter). We're not sure if this is because these restaurants are run by traditional immigrants who have been in the business for years, or some other factors.

A cursory look at Yelp reviews found over 300 reviews:

Best Greek Restaurants in Regional Municipality of Durham, ON

Showing 1-30 of 388

All Filters \$ \$\$ \$\$\$ \$\$\$\$ Open Now

That said, some are not “pure” Greek restaurants, but they do serve Greek food (and are likely owned by Greeks which would likely indicate the food is authentic). This study did not pull in this data, but it may be a consideration for a deeper study to derive more meaningful results.

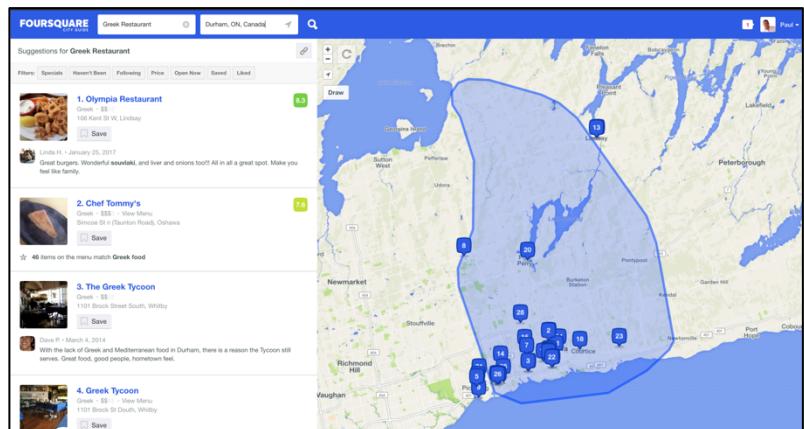
We found the same problem (although not as severe) with Sushi restaurants. The number of Sushi restaurant reviews coming from Foursquare compared to a cursory Yelp! search suggest more investigation is needed.

Best sushi Restaurants in Regional Municipality of Durham, ON

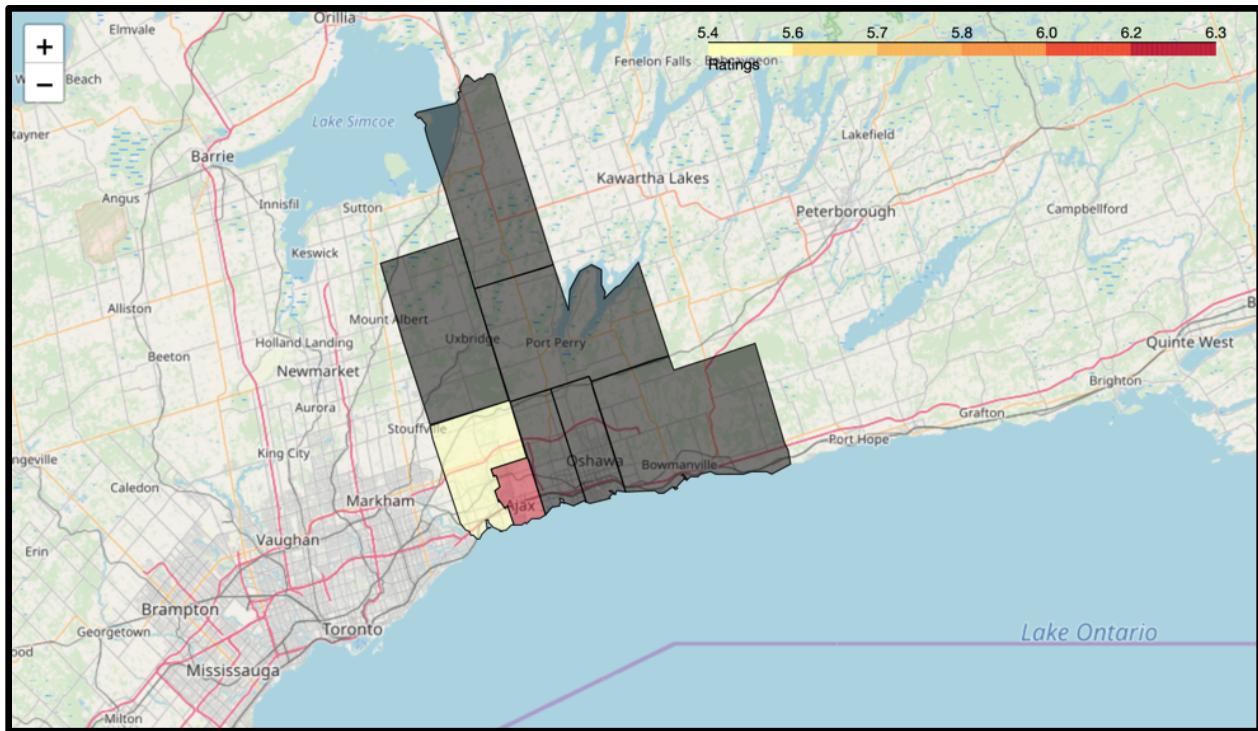
Showing 1-30 of 1153

All Filters \$ \$\$ \$\$\$ \$\$\$\$ Open Now Reservations Good For Happy Hour Full Bar

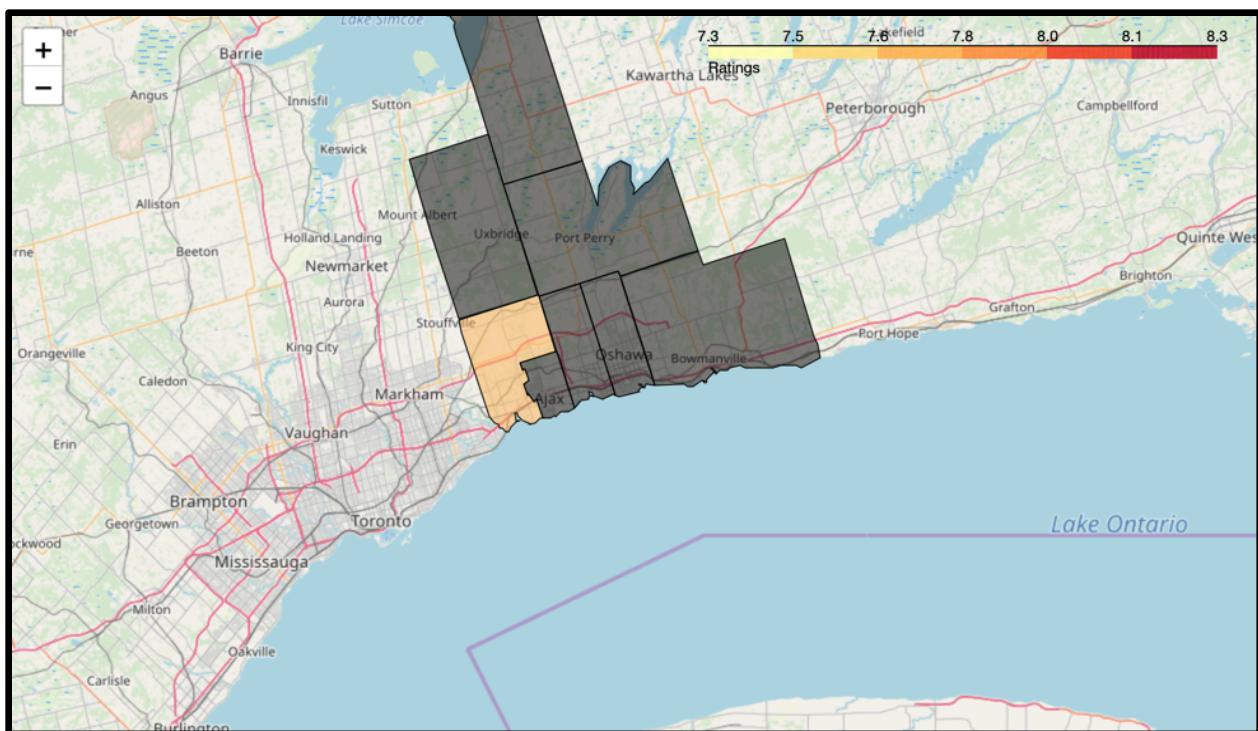
We did find the Foursquare API to be rather unreliable, returning different results upon different iterations. Sometimes it would return nothing, other times a subset. Indeed, we believe the ‘free’ version of the Foursquare API is limiting our research (we used Foursquare as directed by the VC board) because a search using the Foursquare “City Guide” web site shows many more reviews than being returned by the API. Nevertheless, we present our findings in this report.



Sushi Restaurant Ratings



Greek Restaurant Ratings



We decided to look at the top business categories in each of the regions/neighborhoods. This would help us understand if there was a high penetration of other restaurants (after all, if there are eight restaurants competing for dining dollars spent, even if they are not in our targeted food categories, there is still a finite amount of dollars that will be spent on dining out) and if there are other business surrounding the area that might feed clientele to our restaurants (for example, a soccer field or a beach could drive hungry teams and beach-goers to our establishment).

Region Neighborhood	Top Business Categories by Review
Whitby	American Restaurant Breakfast Spot Park Restaurant Arcade
Uxbridge	Coffee Shop Farm Fast Food Restaurant Golf Course Grocery Store
Scugog	Food Truck Grocery Store Harbor / Marina Casino Clothing Store
Pickering	Beach Fast Food Restaurant Indian Restaurant Mexican Restaurant Park

Region Neighborhood

Top Business Categories by Review

Oshawa

Burger Joint
Coffee Shop
Restaurant
Sandwich Place
Steakhouse

Clarington

Coffee Shop
Pharmacy
Fast Food Restaurant
Sandwich Place
Grocery Store

Brock

Bank
Beer Store
Canal Lock
Sandwich Place
Supermarket

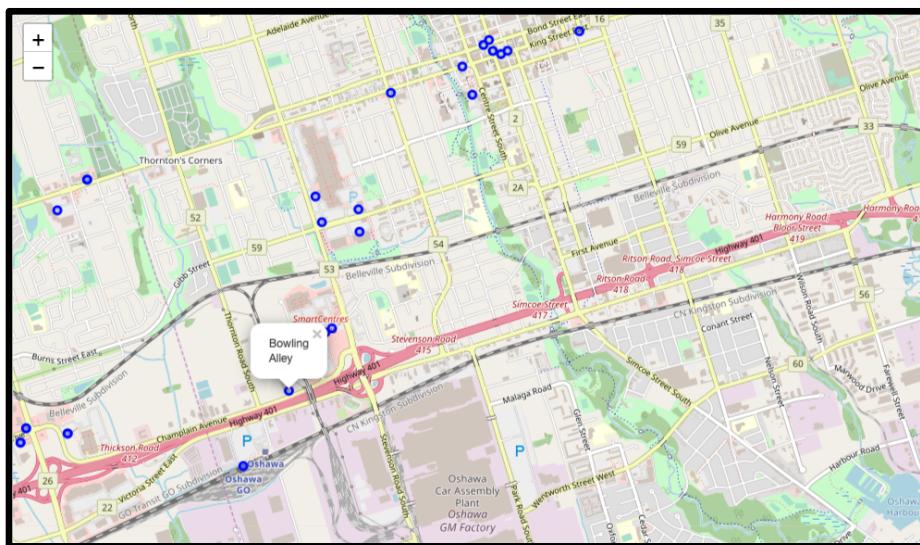
Ajax

Park
Coffee Shop
Gym
Afghan Restaurant
Bookstore

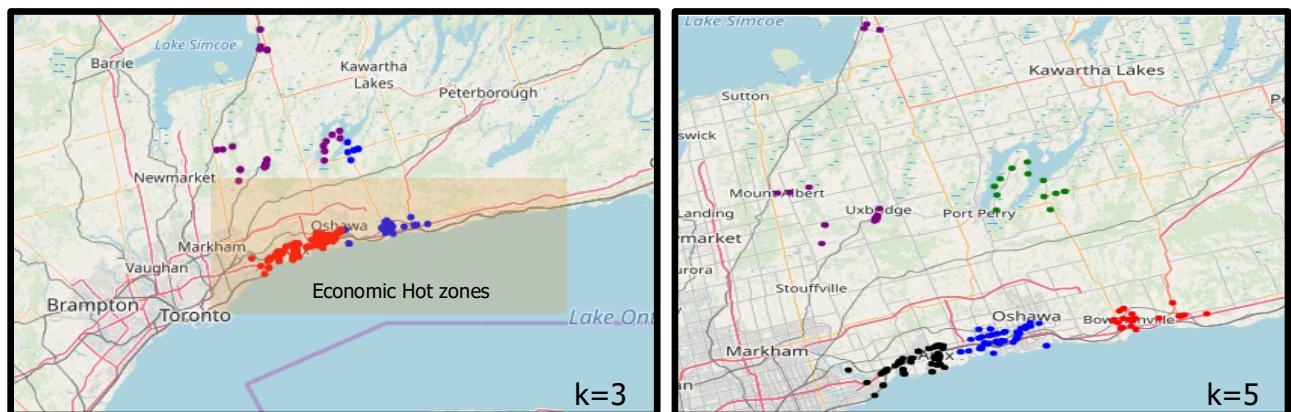
Top Business Areas: k-means Clustering

k-means clustering aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. We decided to use this algorithm to group together business that exists in the area as clusters so we could see opportunities for placement of a potential store based on the kinds of business traffic that surround the area. We did this because as we saw in the previous maps, there are no Greek or Sushi restaurants (or reviews for that matter) in the upper most northern part of the region and that might be for good reason – perhaps there is no pulling factors of a population with disposable income to that area.

Once again, we used Folium to visual the results – this was done with $k=5$ (5 clusters).



We experimented with various values of k and settled on $k=3$ because at $k = 5$ there were clusters that obviously were not too feasible (while there were no competing Greek restaurants at all, and very few Sushi ones, the intensity of business in that area was very sparse).



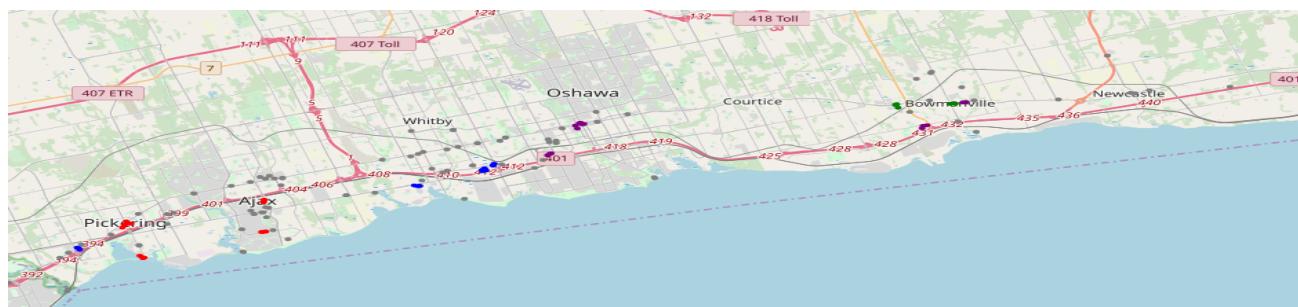
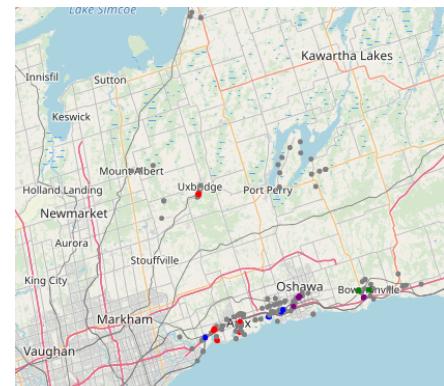
Top Business Areas: Density-based Spatial Clustering of Applications with Noise

We decided to try a different algorithm above and beyond k-means to look at our clustering data: Density-based Spatial Clustering of Applications with Noise (DBSCAN). We chose this algorithm because in our research we noted that in 2014, DBSCAN was awarded the "Test of Time" award which is given to algorithms that have received substantial attention in not just theory, but practice too at leading data mining conferences. This gave us assurances that while we were not using the most advanced latest and greatest algorithms, we were using a sure fire trusted and proven approach to our data analysis.

Quite simply, DBSCAN is one of the most common clustering algorithms and also most cited in scientific literature. DBSCAN is a density-based clustering non-parametric algorithm (meaning you don't get to choose the parameters); when given a set of points in some space (like our regions), it groups together points that are closely packed together (points with many nearby neighbors), marking as outliers points that lie alone in low-density regions (whose nearest neighbors are too far away).

DBSCAN takes in parameters where we can apply our learned investment values from business to the algorithm to help it make decisions on the clustering (as opposed to the manner in which k-means works). Specifically, we supply the *eps* parameter to define a local radius for our expanding clusters (you can think of this as a step size) ... this ensures that DBSCAN will never take a step larger than this parameter: we settled on a value of 0.225. In addition, we specify the *min_samples* parameter to make the number of data points in a radius matter (in other words, if the region wasn't that busy it didn't become a region).

In fact it was this analysis that helped us settle on a k-means clustering of 3 in the previous section; you can see in our results our business best practices inspired results showcase that the northern part of the Durham region shouldn't not be considered at all – it's too sparse for significant activity. On the other hand, the southern part of the Durham region clearly has a lot more significant activity centers and while some of the surrounding areas are not 'lit' up we can see all kinds of opportunities here.

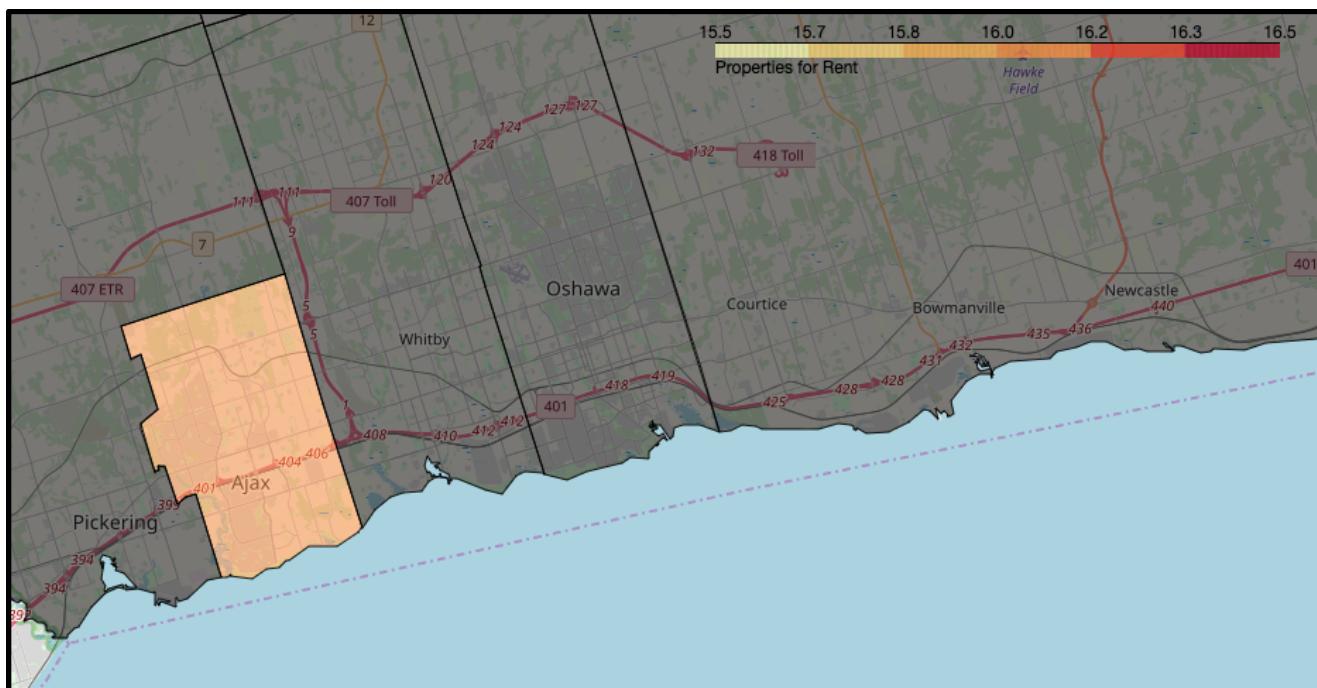


Rental Properties by Region

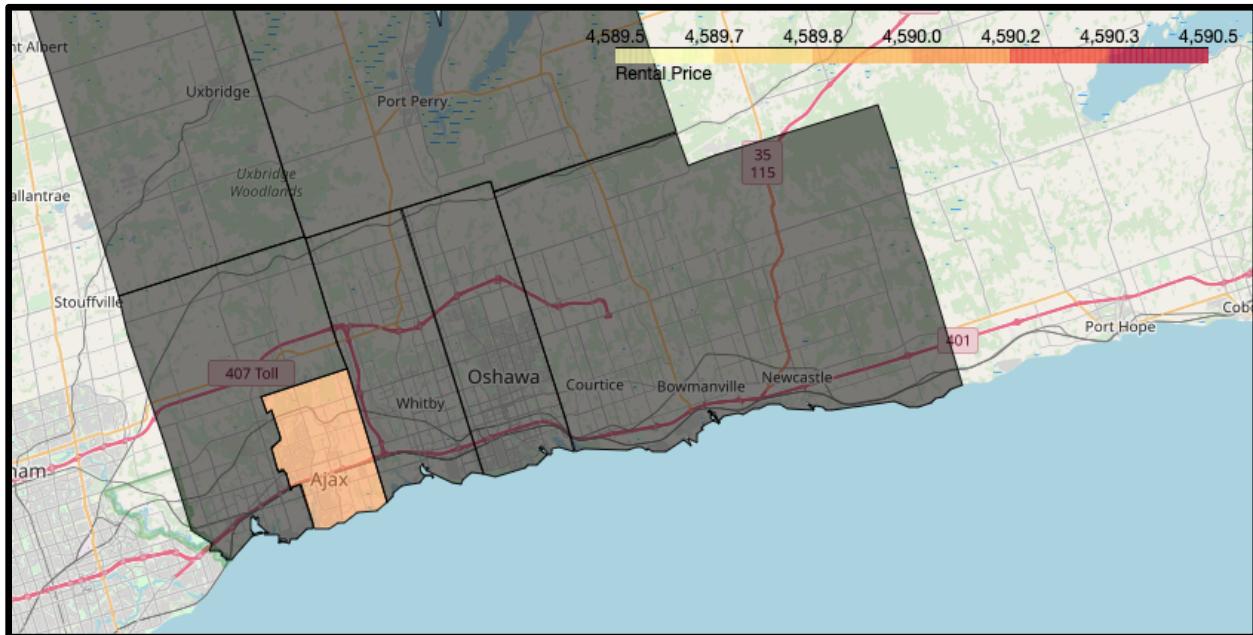
We could have gone on forever adding more and more data sets (we could have looked at weather for example on top of the other variables we talked about earlier) but since we wanted to provide a proof of concept for this particular decision (that could be extensible to others) we decided to deliver a minimal viable product (MVP).

What that said we decided to quickly look properties available for rent (posted on spacelist.ca) and use the Beautiful Soup algorithms to pull in and visualize rental listings. At this point we are pretty sure that that “Swiss Cheese” opportunity lies in the Ajax region, but observing vacancies for rentals will give us an indication of the activity and churn rate of the population in the area.

Spacelist.ca does not provide a “Durham” search region, so we searched by our defined regions (for example, Ajax) with a 15km radius. The rental activity (using a Chlorepeth visualization) suggest that the Ajax region did not have an alarming rate of vacancies (such that it would indicate it’s a place everyone is leaving) and showcased a natural churn (which presents new opportunities to reach out to new-comers to the area).

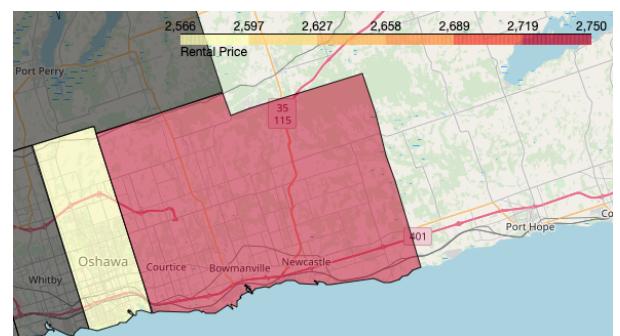
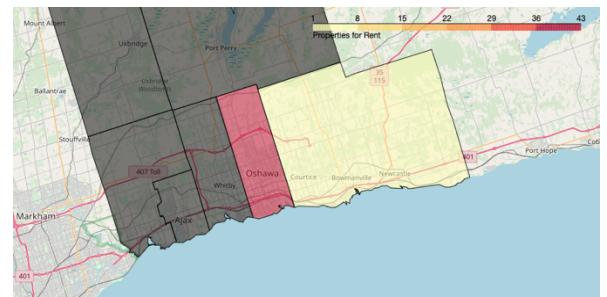


When we looked at where actual rental prices sat on a scale, it was not at the top end (note these numbers and index and not actual rental dollar values).



We also looked at the Oshawa and Clarington regions and found some 'interesting' information. For example, we found that Oshawa was experiencing a significant amount of rental listing activity suggesting that perhaps there is a large movement out of the area ... which might be another factor for us to consider as to why Oshawa didn't end up being the right choice. WE speculated that prices were going up in Oshawa driving more people to the east towards Clarington.

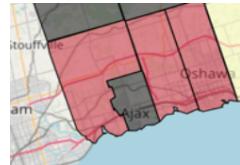
More so, when we looked at the rental costs in Oshawa's neighboring region, Clarington, the rental prices in Clarington are more expensive than Oshawa which would lend support to our hypothesis that rental activity in Oshawa is due to people looking to leave the area; which would be due to the closing of the General Motors plant that we talked about – but it still doesn't explain where those folks are moving to (higher rents that are further east of Toronto which violently go against pricing norms).



As a final test we wanted to look at rental information for the Whitby neighborhood as it was to the west of the Oshawa neighborhood (which should mean more expensive rent) and we were curious as to the rental activity in that area to see if our suspicions could be strengthened (although not fully proven) as to what was going on in Oshawa. Here we found that there was more rental activity than Ajax (25% more than) but much less activity compared to Oshawa (which had 115% more rental activity than Oshawa). Whitby was more expensive than Oshawa ... as was to be expected.

Results

Our results suggest that we are leaning to a **recommendation of Greek food** in the area. We believe the **"Swiss Cheese" opportunity in the Ajax region** could be a **possibly overlooked entry point because our data suggested that there are no Greek restaurants in that area.**



When we looked at the reviews of restaurants in the area (please see the "Discussion" section for considerations for our findings) we see that there aren't any Greek restaurant reviews in this "Swiss Cheese" opportunity area (for obvious reasons, our data suggests there are no restaurants) but there are very little reviews for Greek restaurants in the surrounding regions where we know Greek restaurants exist. This finding suggests a strong opportunity in our "Swiss Cheese" neighborhood and we believe a strong social media campaign could really make people aware of our establishment. It's clear **other restaurants do not have robust social programs.**



Another appealing characteristic of the "Swiss Cheese" opportunity is the top-reviewed surrounding business. If your recall from the table above, the top reviewed businesses (which is an indicator of their popularity and the types of traffic nearby) are: *Park, Coffee Shop, Gym, Afghan Restaurant, Bookstore*. That represents but **a single food establishment in this area is in the top 5**. In contrast, food eateries in the areas for which we are interested (our business intensity map have encouraged us to abandon any thoughts of setting up a going concern in the regions of Uxbridge, Brock, and Scugog) have **3-4x more food establishments in their surrounding neighborhoods!** (Note that Casinos, by law, must have food establishments, and many people go to Casinos for food; today's Supermarkets (Farm Boy, Metro, +++) have full meals available that serve food from all genres including the ones we are investigating. You could argue Coffee Shops serve food, but they are generally not at the same price point as where we plan to compete – they are a processed food low margin business - and wouldn't serve our genre of food.)

Finally, we found that it would appear people are leaving the Oshawa region and it's not related to rental pricing because they are moving to higher rental indexes.

Region Neighborhood	Top Business Categories by Review	
Whitby	American Restaurant Breakfast Spot Park Restaurant Arcade	3 of 5 are food establishments
Uxbridge	Coffee Shop Farm Fast Food Restaurant Golf Course Grocery Store	2* of 5 are food establishments
Scugog	Food Truck Grocery Store Harbor / Marina Casino Clothing Store	2* of 5 are food establishments
Pickering	Beach Fast Food Restaurant Indian Restaurant Mexican Restaurant Park	3 of 5 are food establishments

Region Neighborhood	Top Business Categories by Review	
Oshawa	Burger Joint Coffee Shop Restaurant Sandwich Place Steakhouse	4 of 5 are food establishments
Clarington	Coffee Shop Pharmacy Fast Food Restaurant Sandwich Place Grocery Store	3* of 5 are food establishments
Brock	Bank Beer Store Canal Lock Sandwich Place Supermarket	3* of 5 are food establishments
Ajax	Park Coffee Shop Gym Afghan Restaurant Bookstore	1 of 5 are food establishments

Discussion

Many of the discussion points have been pointed out in the methodology. While we have results, we feel that more data (consistent data) is needed and that would require the purchase of a fully funded Foursquare API account. We feel that the analysis could also be enriched with Yelp! data so we recommend an investigation into the process and fees associated with that review corpus as it seems to be more utilized than Foursquare in the Durham Region.

Another factor that came up with our results that we could not answer is "What if a business exists and yet there are no Foursquare reviews for it?" then it would not show up in our data set as potentials. This is a blind spot to the analysis we have, and we believe expanding the data (the purchasable Foursquare API package, Yelp, and additional information such as the Better Business Bureau of Durham) can help us identify potential competitive locations.

Furthermore, we believe while we have showcased a prototype of the analysis of a Greek or Sushi restaurant there is more data that could help us understand where to place such an establishment and how competitors really are. For example, health inspection records are public domain; a restaurant might even have a strong review, but if we notice a number of sanitation violations we believe that ultimately the market will react, and we could even use that in some "Gorilla" social media attacks when we enter the marketplace (perhaps not super nice ... but it will be very effective).

We did make an assumption that Coffee Shops were not food shops; and while this is changing drastically (Tim Hortons serves sandwiches for example), as mentioned above, we held this assumption constant for all regions and these types of establishments are likely to serve food as a lower price point (due to their low margin business and speed of delivery), it won't be tasty (it's all pre-processed), and they are unlikely to serve our genre of food (there is no Sushi at Tim Hortons – that would be gross).

It is fair to note that the Ajax region's single instance of Top 5 reviewed business was *Afghan Restaurant* ... this would be a direct competitor to a Greek restaurant as the food is very similar. That said, we believe (as unfortunate as it is) there are general biases and concern towards this domain of restaurants and that a Greek restaurant is more 'neutral' in terms of

acceptance. We don't like those kinds of intuitions, but they are to be considered (we recommend further investigation).

The rental data was quite interesting and indeed pushed us towards Ajax; however, it should be noted that we were only allowed to access a single (albeit long) page of rental listings. These rental listing were however random (not sorted by size or rent costs) but we did not due significant statistical diligence on sample data set; we recommend re-running this data with a license for unlimited access ... that said, when we run the experiment over and over again with different listings, the numbers were very close.

Conclusion

We are recommending further investigation of a **Greek Restaurant** in the neighborhood of **Ajax**. We have **eliminated the northern neighborhoods of the Durham region because they do not have a lot of business activity**.

While the initial direction to us from the Venture Board was to favor Oshawa (or start there) we have **excluded** that region as well. The rental data suggested people are leaving the area; rents to the west (expected) and east (**completely not expected**) were higher! What's more, Oshawa had a whopping 4x the occurrence of good establishments in the Top 5 Foursquare reviewed business compared to Ajax that had a single mention.

After looking at exclusions areas for lack of business intensity ... we excluded Pickering and Whitby because they had more food competitors in the Top 5 business reviews (meaning that they have loyal followings that take time to write reviews), and they did not exhibit the "Swiss Cheese" opportunity detailed earlier in this report.

Our data has also shown there are many Sushi restaurants, so we feel this market is saturated; we also found that there are a limited number of Greek restaurants. We believe with a focus on health and nutrition, the "The Mediterranean Diet" (one of the Top 3 diets in the marketplace) can land us some

new customers, a strong social program and take market share away from existing Greek establishment who we believe lack the expertise to drive today's consumers to their brand.

Finally, as a casual observation, the 401 (Canada's longest and busiest highway, in fact, this highway is one of the busiest in North America) starts to jam up in and around Ajax (401 WEST). During the summer months we believe we could tap this 'frustration' of brutal traffic jams as an opportunity by locating close to the 401 (which is within the "Swiss Cheese" zone and offer dynamic pricing coupons; we could cheaply and easily acquire traffic data and location ... as the highway slows more and more, our discounts aren't as aggressive as when it's flowing freely.

FINAL RECOMMENDATION.

OPEN A GREEK RESTAURANT IN AJAX ONTARIO.