The Restaurateurs Dilemma – by Dee Trainham

Introduction:

A restaurateur with ten years' experience in the trade has two restaurant locations serving general faire at competitive prices in the cities of Seattle Washington and Vancouver British Columbia. The Seattle location is in the neighborhood of Roosevelt in the Northern part of the city while the Vancouver location is in the neighborhood of Kitsilano in the Western part of the city. Currently they both receive average levels of patronage with each making a profit of 2,000 dollars weekly. Both the Seattle and Vancouver restaurant locations have now been open for a full three years which means that the restaurateur now has over 300,000 dollars from each location to reinvest into the business. The restaurateur is really thinking about the future and how to grow the business and sees three possible options:

Option 1: Create a new location in a new neighborhood in either city
Option 2: Change the menu from general faire to something more suited to neighborhood tastes
Option 3: relocate the restaurant to a neighborhood with other similarly priced faire
< or a combination of the options above>

To decide which of the options is best with an eye to the future, the restauranteur consults both amateur and professional restaurant reviews. Both the Vancouver Sun and Seattle Times newspapers publish a best cheap eats list for their respective cities each year.

The restauranteur decides to compile these best cheap eats lists together to determine what can be learned about why a business might be reviewed by a professional or amateur reviewer and apply this to the decision making process about the best option for the restaurants in Seattle and Vancouver.

Data Section:

Data was gathered using the Foursquare API and from periodicals. The periodicals used were the Vancouver Sun and the Seattle Times. Both newspapers publish and annual list of best cheap eats in their respective cities as rated by professional reviewers. Additionally foursquare publishes a list of the top 15 cheap eats in each city. These will be used to provide general information to the restaurateur about where reviews are clustered in each city giving them information to possibly theme their restaurant to attract reviewers.

Links to the Seattle Times, Vancouver Sun and Foursquare restaurant lists can be found below:

https://www.seattletimes.com/life/food-drink/top-18-new-cheap-eats-of-the-year-the-seattle-area-food-and-the-stories-behind-it/

https://vancouversun.com/life/food/local-food-reviews/cheap-eats-in-metro-vancouver-dine-in-style-without-breaking-the-bank-with-map

https://foursquare.com/top-places/vancouver/best-places-affordable-prices

https://foursquare.com/top-places/seattle/best-places-affordable-prices

Following the cleaning of the address data in the linked articles above, the geopy geocoder and folium will be used to generate a map of Seattle and Vancouver to show where restaurant reviews are geographically distributed.

The foursquare api will then be used to determine the distribution of venues throughout each city as well as an examination of the local neighborhoods of Kitsilano and Roosevelt so as to determine what competition currently exists for the restaurateur. The restauranteur will also use the API to query the neighborhoods where the best cheap eats are located to gauge what the most popular types of cuisine are and the frequency distribution of different types of cuisine in these neighborhoods. Using the resulting data the restauranteur will carefully consider options for the future of the business.

From observing the basic restaurant data in a list format we can begin to make a few observations about the sample size and that Vancouver is slightly overrepresented compared to Seattle. To not let this larger sample size unduly influence the results each city must be looked at independently. This requires the construction of two python notebooks (one for each city), these notebooks will:
-   map the city neighborhoods;
-   map the restaurant reviews to their respective neighborhoods using the geopy geocoder and folium
-   use the Foursquare API to examine the neighborhoods where the restaurant reviews are located as well as the neighborhoods where the restauranteur current has restaurants in Seattle (neighborhood of Roosevelt) and Vancouver (neighborhood of Kitsilano)
-   determine where the reviews cluster, why a restaurant might be reviewed and how this will influence the possible options for the future of each of the two current restaurant locations.
These maps will then be analyzed and annotated to display relevant information building a case to support the best of three given options to the restaurateur.

Methodology:

For this project I thought that good option would be to explore two different cities and understand the importance of geography in data gathering and decision making. I first started with basic exploration of what questions could be asked that could be solved using foursquare data while foursquare data does contain reviews of many items from bus stops to yoga studios, I thought it would be best to stick with food venues as these are the most common type of venue.

Now that I knew that I wanted to examine food venues using the foursquare API, I hit upon the idea of looking at restaurant reviews in Seattle and Vancouver where the top venues in terms of affordability and quality were given, as it happened foursquare also listed the top 15 cheap eats in each city. From these lists I geocoded the addresses in python and created a map for Seattle and Vancouver that I would use to compare with the foursquare API data.

Using the foursquare API I managed to collect all venue data for both Seattle and Vancouver I then segmented this data by neighbourhood to get and venue count and frequency distribution for each type of venue. I then removed all venues that were not food related so that I could make comparisons between neighborhoods.

Now that I had the data segmented by neighbourhood I could use the iloc function in python to further segment out the two neighborhoods where the restaurateur has venue locations in Roosevelt, Seattle and Kitsilano, Vancouver. I then mapped the neighborhoods using folium to determine any patterns in the geographical distribution of venues across the neighborhoods. I also examined the frequency distribution of venues in each neighbourhood and found that coffee shops were the most common in each neighbourhood as well as across each city. This became the first piece of information that I could use to inform the restaurateur about making a decision on their venues.

After examining the Roosevelt and Kitsilano neighborhoods I conducted a k-means cluster analysis on all neighborhoods and venues with a default of five clusters. This demonstrated the diversity and distribution of food venues across each respective city. I then examined the cluster membership and provided this information as advice to the restaurateur about venue diversity.

Lastly I examined the neighbourhood location of restaurant reviews to determine if there was any relevant pattern and found a single neighbourhood in both Seattle and Vancouver which accounted for over 30 percent of reviews. Using the folium package in python I then generated a map of all venues in Downtown Vancouver and Capitol Hill, Seattle to compare with the map of reviewed venues that the restaurateur cold use to determine how they might want to change their business model in order to get a review.

Results:

From the data analysis I was able to determine a frequency distribution for food venues across both Seattle and Vancouver as well as determine neighbourhood diversity of venues using k-means clustering. The most interesting finding was how prevalent coffee venues were in both cities not dispelling but instead confirming a long held stereotype of both cities. These findings helped me to provide detailed advice to the restaurateur on making a determination as to the best option for the future of their business that I will discuss in the section below.

Discussion:

The restaurateur originally gave three possible options for the future of their locations in Seattle and Vancouver:

Option 1: Create a new location in a new neighborhood in either city
Option 2: Change the menu from general faire to something more suited to neighborhood tastes
Option 3: relocate the restaurant to a neighborhood with other similarly priced faire
< or a combination of the options above>

I will discuss viability of each of these options in turn.

Option 1: Create a new location in a new neighborhood in either city:
Based on the data we found the restaurateur could consider this option for Vancouver due to the concentration of venues and reviews in the Downtown area of the city; however as Seattle has a more even distribution of venues this does not seem like a good option there.

Option 2: Change the menu from general faire to something more suited to neighborhood tastes:
This is a viable option for both the Seattle and Vancouver locations so long as the restaurateur doesn't make their venue a coffee shop (or offer sushi as in Vancouver).

Option 3: relocate the restaurant to a neighborhood with other similarly priced faire:
In the case of Vancouver this may be a good option as the current location in Kitsilano is relatively unique being the only neighborhood in its cluster. Perhaps a move to cluster 0 which Downtown is also a member would be a good option. Roosevelt clusters with many neighborhoods by comparison so moving the venue would not produce any tangible benefit.

From the data collected and imputed the best recommendation that could be made to the restaurateur was to consider a new location (possibly in Downtown) for Vancouver and to potentially specialise the menu for the Seattle location provided the menu did not center on coffee.

Conclusion:

This was a challenging and rewarding activity in that I learned new approaches to data analysis and web scraping that were unknown to me previously. This helped me to also consider how to best

approach a data analysis project in the future by planning/mapping out the project to determine the best use of resources and what problem could be solved/answered given the data.

Appendicies:

The following are the maps produced from the data analysis using folium and are discussed in some detail in the presentation pdf.
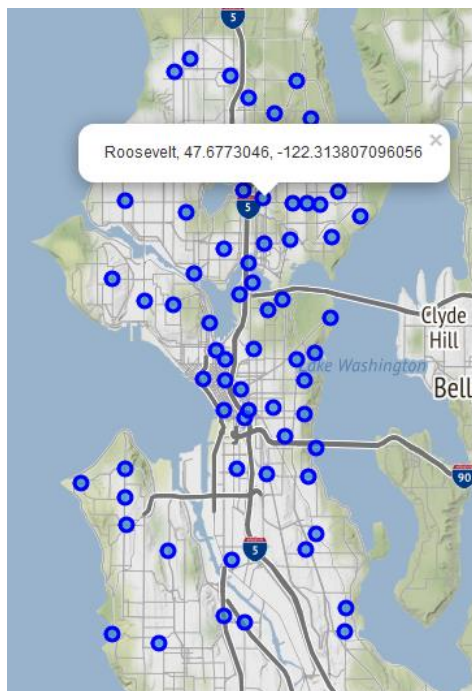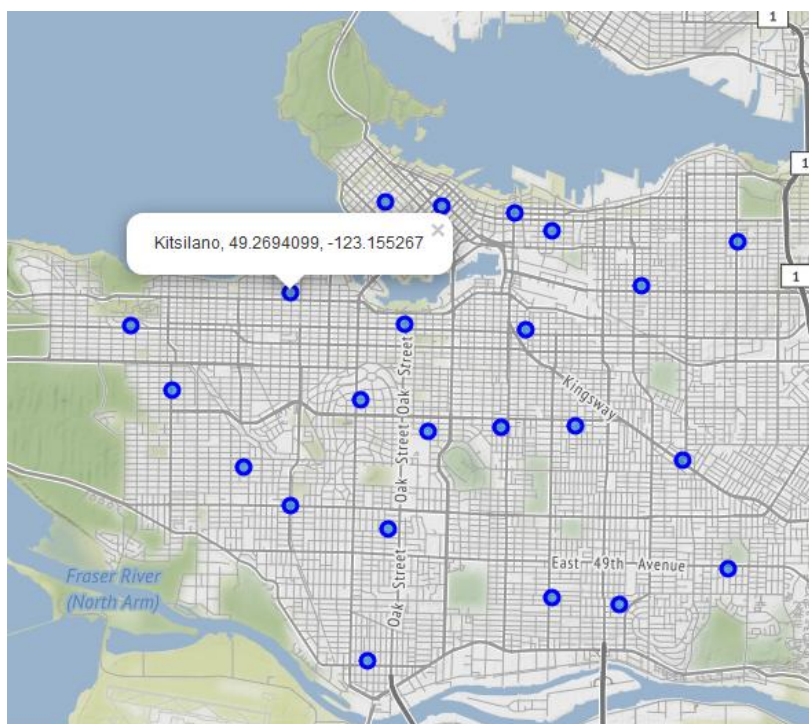

Figure 1 - Seattle Neighborhoods
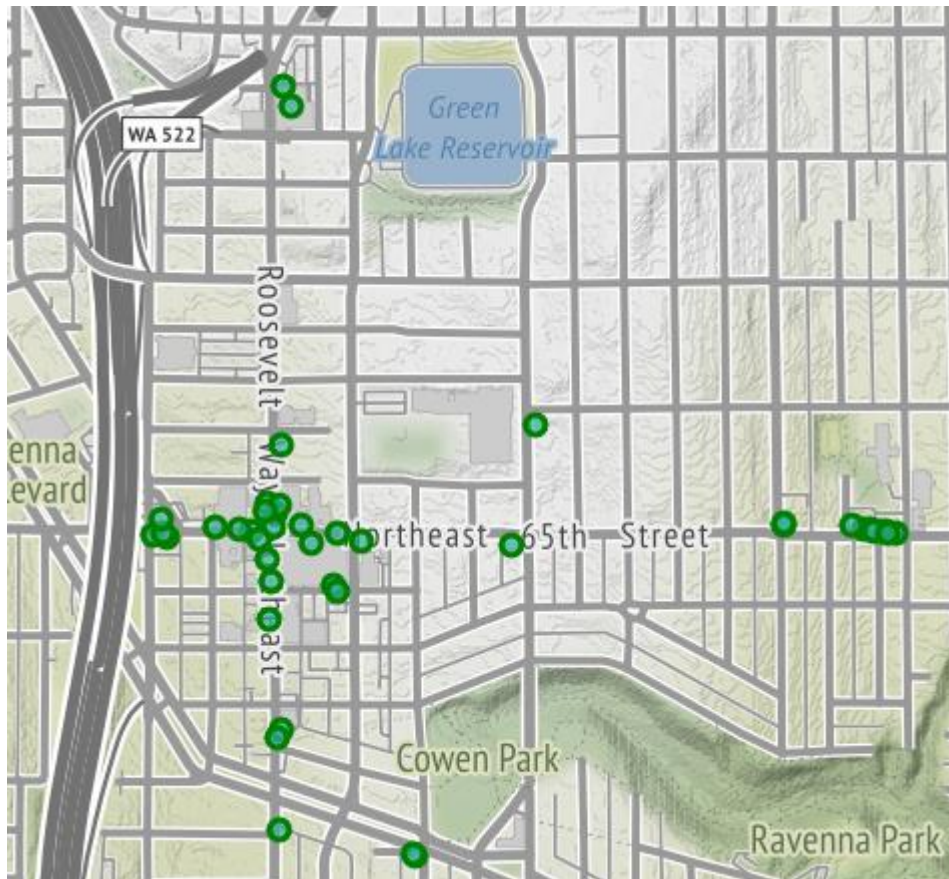

Figure 2 - Vancouver Neighborhoods

Figure 3 - Roosevelt Venues

Coffee Shop	5
Bar	3
Vegetarian / Vegan Restaurant	3
Burger Joint	3
Grocery Store	3
Vietnamese Restaurant	2
Mexican Restaurant	2
Southern / Soul Food Restaurant	2
Pub	2
Pizza Place	2
Bakery	2
Indian Restaurant	2
Thai Restaurant	1
Sushi Restaurant	1
Bagel Shop	1
Italian Restaurant	1
Gourmet Shop	1
German Restaurant	1
Farmers Market	1
Creperie	1
Café	1
Breakfast Spot	1
Greek Restaurant	1

Figure 4 - Count of Roosevelt Venues
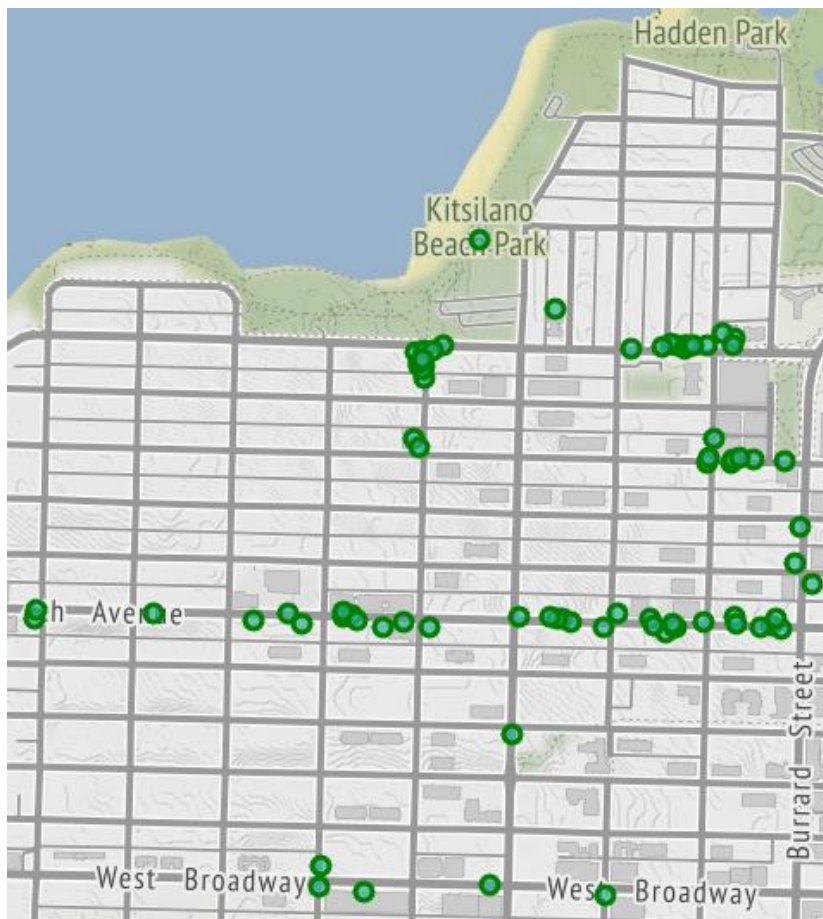
Figure 5 - Kitsilano Venues

Coffee Shop	11
Sushi Restaurant	5
Restaurant	5
Bakery	4
Burger Joint	3
Café	3
Pizza Place	3
Japanese Restaurant	3
American Restaurant	2
Grocery Store	2
Thai Restaurant	2
Tea Room	2
Seafood Restaurant	2
Noodle House	2
Ice Cream Shop	2
Italian Restaurant	2
Vegetarian / Vegan Restaurant	2
French Restaurant	2
Food Truck	2
Ramen Restaurant	1
Bagel Shop	1
Bar	1
Breakfast Spot	1

Figure 6 - Count of Kitsilano Venues
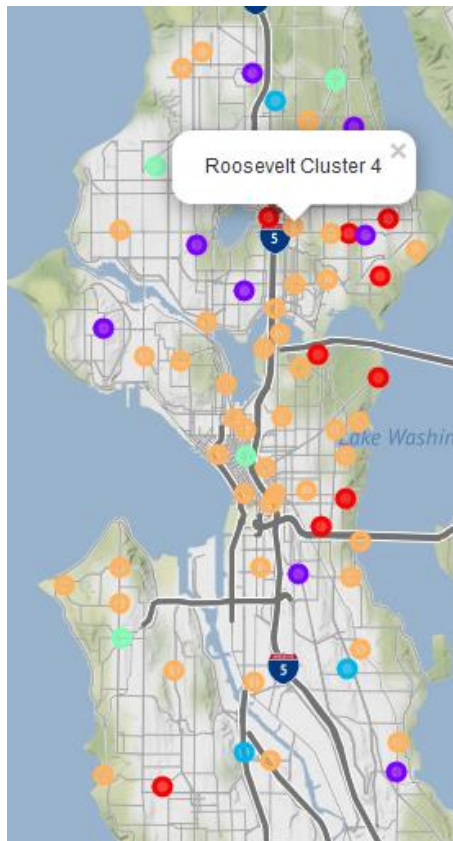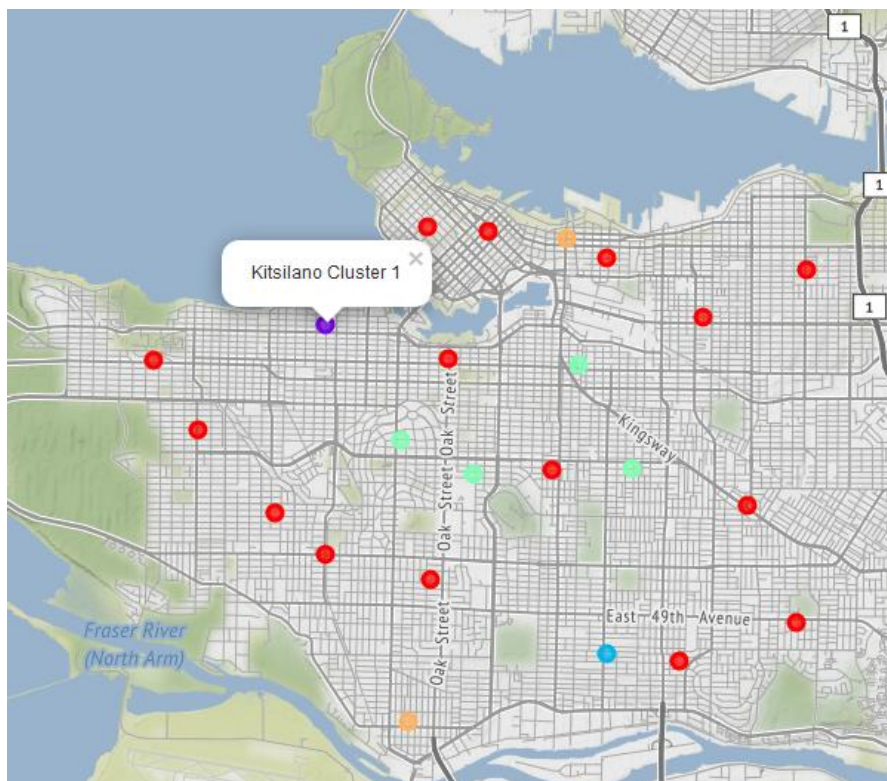
Figure 7 - Seattle Clusters

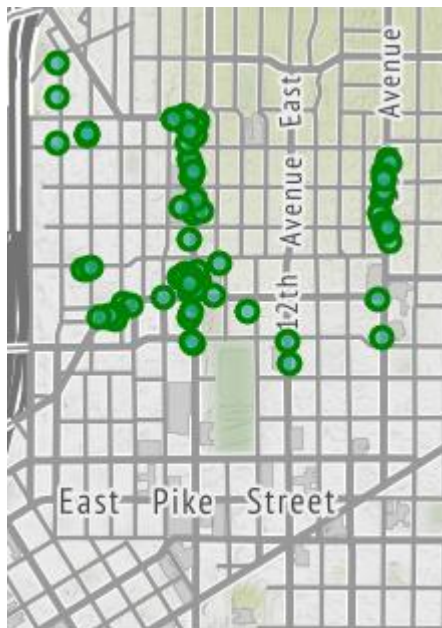**Figure 8 - Vancouver Clusters**

Figure 9 - Capital Hill Venues

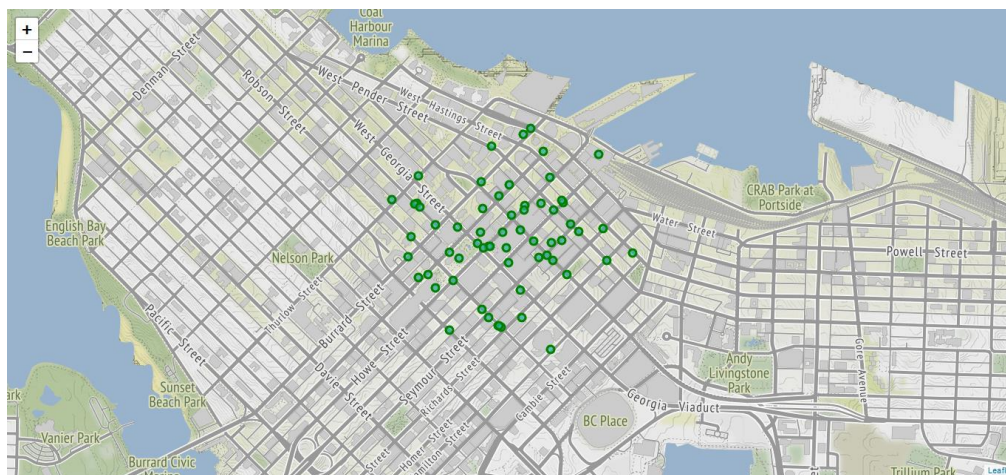

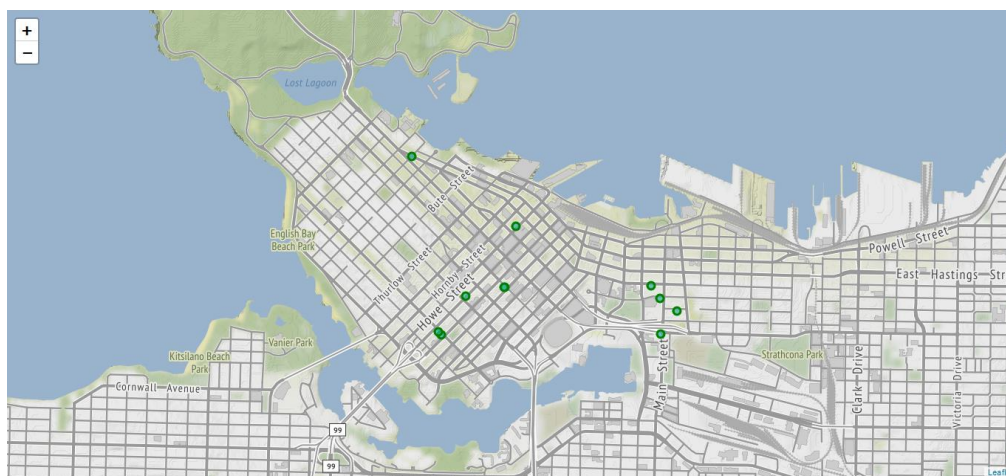**Figure 10 - Capital Hill Reviews**

Figure 11 - Downtown Vancouver Venues

Figure 12 - Downtown Vancouver Reviews