

The perfect location for a coffee roastery in Canada

By Daniel Helwig

Introduction

A friend of mine is looking to open a coffee roastery in Canada and wants to know which city between Toronto, Montreal and Vancouver would be the best to start the business. For this we will evaluate the number of coffee shops that are available in these cities. The logic is that the more coffee shops are available the more opportunities would be available to sell the roasted coffee beans. This would mean a larger target group and probably a higher revenue. The target audience is the friend who wants to open his business and ensure that he has enough prospective customers. He should get a list of which customers he could provide coffee beans to and why a specific city would be best.

Data

First the locations of the different cities have to be found so that coffee shops in the vicinity of the city centre can be found. This will be done via **geopy**. The locations will be defined as the longitude and latitude. The radius in which the search will be done is chosen as **5km** from the city centre. This data will be extracted from the **foursquare** database. We will then implement the results of all cities in one **dataframe**. To check that the coffee shops are in the correct area we will show the positions on a **folium** map. This will be the basis for the evaluation. This data will still have to be looked into for consistency and relevance and thus might have to be cleaned in the evaluation stage.

Methodology

With this data we will have a look at which city has the highest density of coffee shops. First of all we will filter for pure coffee shops or similar. This means that no restaurants, sushi places, etc. will be included in the dataframe.

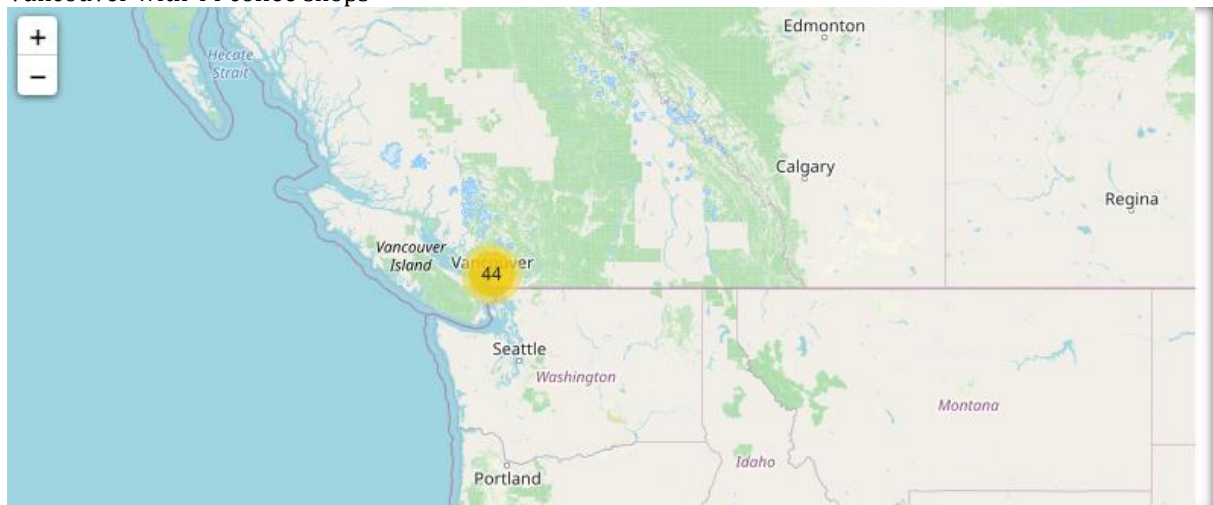
After cleaning the data we will analyze the amount of coffee shops near the centre of the city using a folium map. A simple map showing the amount will suffice in showing which city has the highest density. This will lead to the city selection.

Considering that the coffee has to be distributed we will go one step further and cluster the city into n amounts. For the cluster we will consider that per week each coffee shop has to receive fresh coffee once. The coffee is distributed via bike and 8 coffee shops can be supplied per day as a first estimate. The clusters should show the area the bike has to supply.

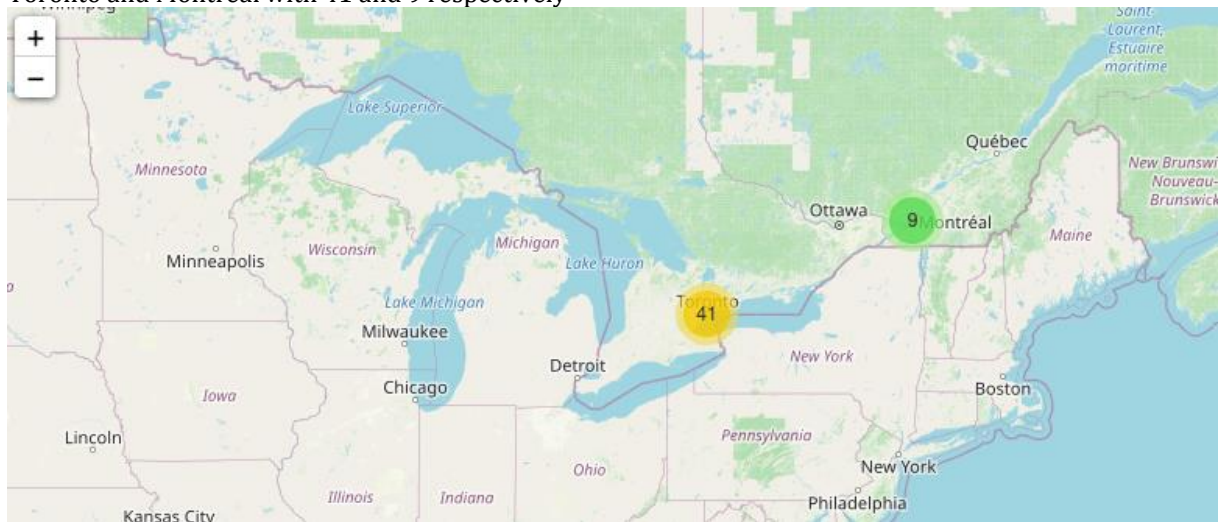
Results

The three different locations give a dataframe with 2850 datapoints. These are then evaluated to show the following image with the following amount of coffee shops. The top shows Vancouver the bottom Toronto and Montréal.

Vancouver with 44 coffee shops

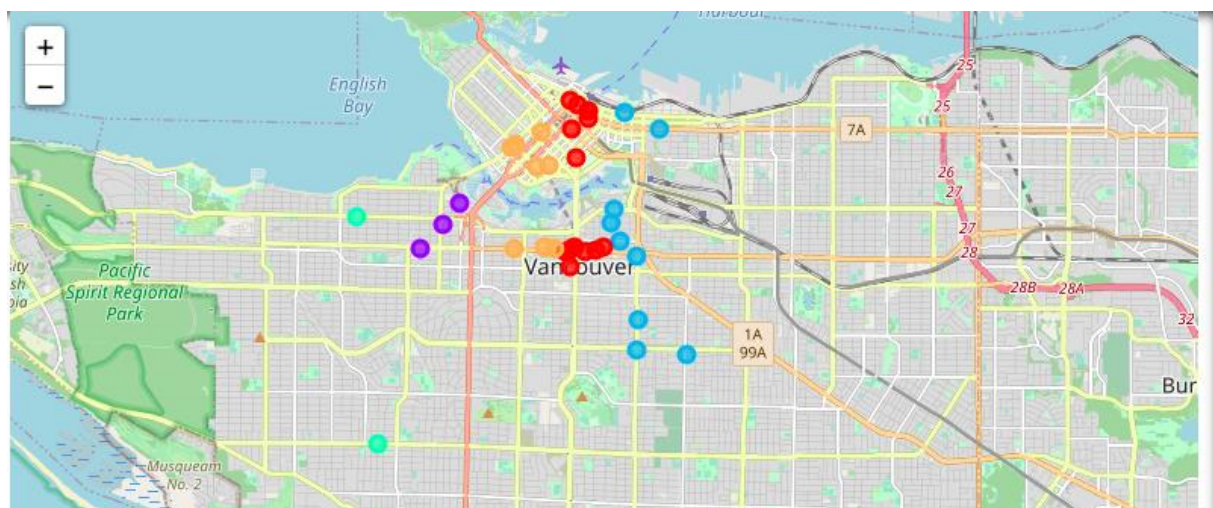


Toronto and Montréal with 41 and 9 respectively



Although close the clear winner is Vancouver.

For the distribution of the coffee to the coffee shops we will cluster them into 5 areas so that the coffee can be distributed via bike. This is shown in the following map.



Discussion

The final map shows how the coffee could be distributed in the city of Vancouver. Vancouver was chosen as the basis for the coffee roastery as it had the greatest density of coffee shops in a 5km radius of the city centre compared to the cities Toronto and Montréal. From the analysis this would have the highest chance of succeeding as the assumption is that the higher amount of coffee shops the higher the amount of potential for the need of coffee beans from a coffee roastery.

The analysis does however have a few weaknesses in that a few points have to be evaluated in greater detail. The coffee shops that were evaluated are from a single database and maybe not all coffee shops were captured. For this other databases should be evaluated as well. A point that has to be considered is that Montreal shows far less coffee shops in comparison to the other two cities. The reasons for this should be looked into as well and maybe the location of the city centre moved to find an area where the coffee shop density is higher.

The analysis focussed on three cities in Canada. It could be possible that further cities might have a higher density of coffee shops. This would however have to be discussed if a different city is an option.

Furthermore it should be noted that the coffee shops probably already have their own coffee roastery distributor that supplies the coffee beans. This is difficult to evaluate but a first step would be to look into what coffee shops are part of a franchise and how many coffee roasteries are in the general area.

Conclusion

The analysis gave a first insight into which city is most probable to have a good customer base for a coffee roastery. Further analysis should be done to see where the coffee roastery could be physically based, i.e. in the city centre or further outside the city. This would have a large effect on the running costs.

As the names of the coffee shops are available a survey should be done to see if there is a need for a different supplier of coffee beans. This is the proposed next step in finding the acceptance by the potential customers