

Where should I open a bakery in NE Ohio?

Introduction:

A fellow baker friend asked me which location would be best to open a bakery nearby? I do not know the answer, but we wondered if having certain businesses and the population around the area would influence the business model. Finding out if there are business type patterns of successful bakeries that we know of around the nearby neighborhoods, we may be able to assess other areas without bakeries to be suitable to open a bakery

Data Acquisition, Cleaning, and Joining:

Acquiring list of cities and townships in Ohio:

First off, we obtained a list of cities in Ohio through [Wikipedia](#). The list was scraped through Rvest. Because this list only contained cities and not towns, we went to the Summit County website and obtained the list of towns part of Summit County ([link](#)). Because Foursquare API query uses latitude and longitude of a location, we used OpenStreetMap to get the latitude and longitude of the list.

Successful Bakeries we want to mimic:

Koko Bakery, Tiffany's Bakery, and Peace Love Doughnut shops were selected to be the bakeries we want to mimic due to the success we see in them. Latitude and longitude were obtained manually and joined into the existing data frame. A "zzz" prefix was added to discern the bakeries of interest and the rest of the neighborhood/cities.

Financial hypothesis:

Several articles such as [this](#), indicate that millennials and gen z are more likely to purchase baked goods than other populations. We will estimate that the majority of bakery customers will be above 18 years old and below 65 years ago. Assuming that, at baseline, we will sell at a minimum of 100 baked goods with a profit of \$2 per baked goods, these are to be catered towards certain hotels, offices, etc. Additional customers will depend on the existing population, assuming that returning customers will be about 0.1% of the population aged above 18 and below 65 years old. Expenses that are not included such as the rental; the number of employees will differ according to the location.

Methodology:

Business pattern assessment:

With the above data frame, we used FourSquare API to explore nearby venue categories within 1000 meters of the latitude and longitude provided per location. Then the venue categories were transformed into one-hot-encoding. The variables were then normalized and sorted into the top 10 venues found. K

nearest neighbor technique was then utilized to assess patterns of the bakeries we are interested in and also the other locations to assess similarity. After that, the cluster of interest will then be filtered. To avoid huge competition, if a bakery is listed as one of the top 10 businesses in the city/town, we will filter that city/town off. We also sorted the distance closer to home and filtered off anything that is more than 31 miles away from home. We then chose the top 5 areas that we want to look at for opening a bakery

Population demographics/Rental data:

Population demographic data were obtained through census.gov quick facts <https://www.census.gov/quickfacts/fact/table/OH/PST045219> Total population estimate for 2019, percentage of age <5, <18 and >65, median income were obtained for all the cities/towns of interest. Total ages 18 through 65 were calculated by 100% subtracted by the percentage of age <5, age <18, and age >65. Rental estimate was obtained through scouting prices at loop.net. The estimated square footage required is 900 sq feet. Hence data obtained were based on cost per square footage per year. Any unobtainable data will remain blank.

Cost analysis:

This bakery is estimated to open five days a week, with a baseline selling of 100 baked goods per day. It is hypothesized that 0.1% of the total population of age 18 through 65 would come to the bakery to buy at least 1 baked good per week, in addition to our baseline 100 baked goods per day. Assuming that each baked good would cost \$1 to make/burn (any defect that cannot be sold), another \$1 for hiring employees per baked good (approximately \$12.5/hour rate per employee), and we're selling for \$4 per baked good. With regards to the rental, assuming that any increase in 1 baked good above the baseline 100 baked goods per day would require additional square footage of rent, the rental would increase as the total population increases. Total net revenue would be total profit subtracted by adjusted rental.

Results

We observed that Koko bakery and Peace, Love Doughnut shop belong to cluster 2. After the filter, 17 locations of interest (Table 1) were sorted ascending of the nearest city to our home. At a glance, the pattern of these locations appears to have top businesses that attract local and out-of-state people, such as ice cream shop, park, zoo, restaurants etc.

Table 1: Cities in Cluster 2 and its Top 10 Businesses

City	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
Cleveland	Hotel	Brewery	Coffee Shop	Theater	Deli / Bodega	Music Venue	Park	Steakhouse	Performing Arts Venue	Sports Bar
Parma	Ice Cream Shop	Pizza Place	Grocery Store	Donut Shop	Coffee Shop	Bar	Steakhouse	Sports Bar	Diner	Sandwich Place
Brooklyn	Ice Cream Shop	Zoo Exhibit	Coffee Shop	Park	Pizza Place	Brewery	Donut Shop	Diner	Deli / Bodega	Burger Joint
Richfield Township	Trail	Park	Hotel	Scenic Lookout	Bar	Farmers Market	Ski Area	American Restaurant	Sandwich Place	Rest Area
Boston Township	Trail	Park	Scenic Lookout	Hotel	Rest Area	Bar	Ski Area	Farmers Market	Sandwich Place	Campground
Lakewood	Bar	Pizza Place	Park	Coffee Shop	Grocery Store	Mexican Restaurant	Ice Cream Shop	Sandwich Place	Gym / Fitness Center	Gastropub
Rocky River	Pizza Place	Park	Bar	American Restaurant	Coffee Shop	Gym / Fitness Center	Mexican Restaurant	Grocery Store	Ice Cream Shop	New American Restaurant
Euclid	Pizza Place	American Restaurant	Bar	Ice Cream Shop	Gym / Fitness Center	Sandwich Place	Pharmacy	Diner	Chinese Restaurant	Italian Restaurant
Richmond Heights	Pizza Place	Park	American Restaurant	Grocery Store	Bar	Italian Restaurant	Sandwich Place	Gym / Fitness Center	Ice Cream Shop	Coffee Shop
Cuyahoga Falls	American Restaurant	Bar	Park	Brewery	Ice Cream Shop	Mexican Restaurant	Italian Restaurant	Café	Diner	Breakfast Spot
Willowick	Bar	American Restaurant	Pizza Place	Diner	Sandwich Place	Italian Restaurant	Park	Grocery Store	Coffee Shop	Restaurant
Wickliffe	Bar	Pizza Place	Italian Restaurant	American Restaurant	Diner	Gym / Fitness Center	Sandwich Place	Park	Restaurant	Ice Cream Shop
Bay Village	Grocery Store	Park	Ice Cream Shop	American Restaurant	Bar	Coffee Shop	Italian Restaurant	Pizza Place	Sandwich Place	Seafood Restaurant
Eastlake	Bar	Pizza Place	Diner	Sandwich Place	Park	Ice Cream Shop	American Restaurant	Italian Restaurant	Grocery Store	Coffee Shop
Willoughby Hills	Pizza Place	Bar	Park	Ice Cream Shop	American Restaurant	Italian Restaurant	Gym / Fitness Center	Sandwich Place	Restaurant	Coffee Shop
Willoughby	Bar	Pizza Place	Park	Sandwich Place	American Restaurant	Diner	Mexican Restaurant	Ice Cream Shop	Italian Restaurant	Coffee Shop
Kent	Coffee Shop	Park	Pizza Place	American Restaurant	Mexican Restaurant	Bar	Ice Cream Shop	Sandwich Place	Pharmacy	Dessert Shop

When further data such as total population, percentage of age 18 through 65, median income, estimated rent per square footage per year were merged into the existing data frame, we found Richfield and Boston Township to lack further details of percentage of the population between age 18 and 65. Also, the total population of those towns was relatively small, 5424 and 1664, respectively. These places were dropped due to the lack of data and also likely non-feasible for sustainable business. Rocky river township was also dropped as well as we were unable to estimate the rental cost. When profit analysis was conducted and sorted descending in order of percentage of age 18 through 65, followed by median income, followed by profit, and then by total population, the final rank order was produced (Table 2).

Table 2: Cities, Population, Rent and Profit Analysis

City	Total Population	% Population of Age 18 through 65	Median Income	Rent (\$ per month)	Profit
Kent	29646	71.0	32993	1507.0162	2661.373
Lakewood	49678	64.8	53290	1510.7304	2746.800
Willoughby Hills	9553	62.5	59328	1126.4927	2921.272
Brooklyn	10646	60.3	48552	901.2839	3150.072
Eastlake	18042	59.0	54698	902.1290	3183.029
Parma	78103	56.8	57120	908.8725	3446.028
Wickliffe	12744	56.6	61105	751.2022	3306.503
Cuyahoga Falls	49106	56.6	57101	754.6323	3467.720
Willoughby	22977	56.2	60332	902.5826	3200.722
Euclid	46550	56.0	38242	754.3447	3454.199
Richmond Heights	10342	55.8	51505	750.9618	3295.205
Cleveland	381009	55.6	30967	1963.2671	3731.461
Willowick	14105	55.2	62016	676.1679	3386.120
Bay Village	15194	49.7	103582	1051.7620	3008.649

The top 5 locations to open a bakery are Kent, Lakewood, Willoughby Hills, Brooklyn and Eastlake. At a glance, Willoughby Hills total population appears to be the smallest among all; we decided to drop this and replace it with another city, Parma. Of these cities/towns, the range of total population is 10646 to 78103, percentage of age 16 to 65 is 56.8 to 71%, the median income of \$32993 to 57120, rental per month \$902.12 to 1507 and profit of \$2661 to 3446 per month.

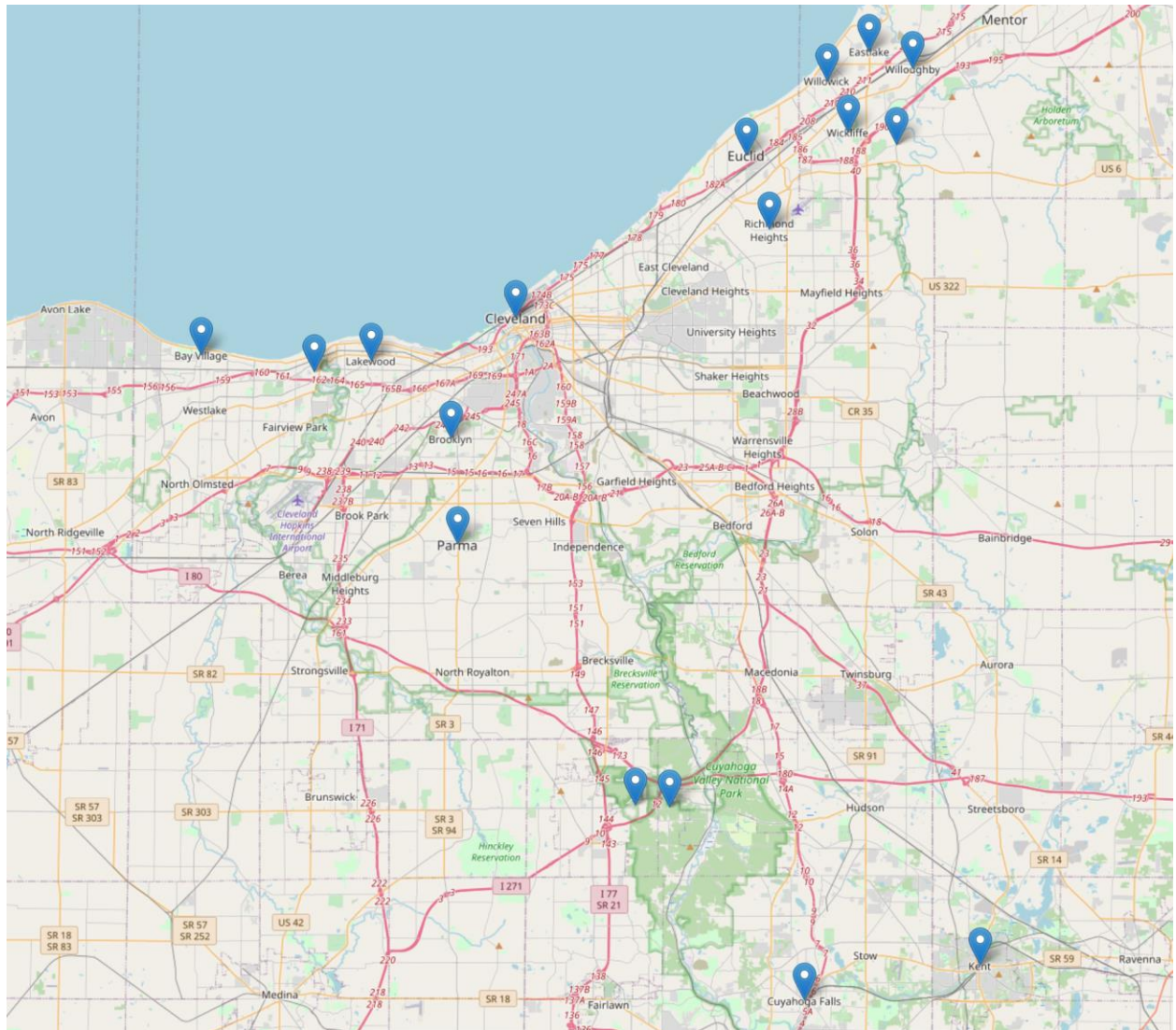


Image 1. Locations of interest on map

Discussion:

The top 5 locations best mimic the famous bakery we like are Kent, Lakewood, Brooklyn, Eastlake, and Parma. However, Brooklyn and Eastlake populations are smaller, so Kent, Lakewood, and Parma would be better choices. Although Kent appears to be the top mainly because of the high percentage of the population aged 18 through 65, the median income seems to be the lowest among Lakewood and Parma, as did the profit margin. Kent is known for the diversity of population given the University of Kent nearby, which brings a good mix of people. It is an attractive location to open a bakery because of this. However, rental appears to be high; profit seems to be low. Unless there is a considerable discount or incentive program, the gain may not be as lucrative as other cities. That being said, it may be easier to employ someone because of the university setting.

The following locations of interest would be Lakewood and Parma. When compared with the total population, median income, rental cost, and profit, it seems like Parma is a better place than Lakewood. The bonus is that it is closer to home than Lakewood. We have also heard great things about Parma being a lovely suburban location for business. Looking closer at the top 10 businesses at Parma would be Ice Cream shop, pizza place, Grocery store, donut shop, Coffee shop, Bar, Steakhouse, Sports bar, Diner, Sandwich place. When looked at Google Map, there are a few bakeries there already, mainly Strudel, which could be a possibility that there are more Austrian descent population there. I do wonder if an Asian bakery would be attractive in this location given the diversity and also a good median income level. Rent is on the lower side as well.

Limitation:

The limitation of this research is mainly from the data that is available through free FourSquare API. We could dive deeper into the number of bakeries in the area and also including tips as well. However, I think having a big picture might be better. Another limitation of this research is that it is based heavily on hypotheses and assumptions of baked goods number production. Most baked goods do not follow a constant, linear profit/production due to seasonality (season, school vacations, etc.). Lastly, we could have also included the nearby towns in nearby counties, but we only chose Summit County; this could have further broadened our outlook of other nearby areas

Conclusion:

Parma appears to be an excellent location for a bakery. It has a great neighborhood, good median income, sustainable customers, and most importantly, good profit!