

# **Capstone Project - The Battle of Neighborhoods**

**Location Analysis for Opening a New Restaurant in  
Cleveland**

Chia-Hung Kuo  
May 2020

# Outline

- Introduction
- Analytic Approach
- Data requirements
- Results
- Conclusion and Suggestions

# Introduction

- **Business understanding**

- **Introduction:**

- Great Cleveland area is a fast-growing metropolitan area in northeast Ohio
    - 3.5 million residents and approximately 2.5 million visitors annually
    - Great place to start a new restaurant

- **Project objective:**

- Find the location for starting a successful restaurant

- **Targeted Audience**

- People wants to start a new restaurant in the area

- **The problem**

- In a go-to neighborhood of dining out and have regular customers

- **Clear question**

- How to find the optimal location for the restaurant?



## **Analytic approach**

- Analyzing the income, population, and similar business in the borough to obtain the insight of the neighborhood
- Using K-Means clustering to find out the groups of neighborhoods that seems to behave similarly to each other in the scale of borough level, and then narrow down to zip code area

# Data requirements

- Foursquare API for the listing of greater Cleveland neighborhoods information with clustered venues
- Ranking every Ohio city, county for median family income - Census Snapshot by Data Central ([https://www.cleveland.com/datacentral/2017/12/ranking\\_every\\_ohio\\_city\\_county\\_2.html](https://www.cleveland.com/datacentral/2017/12/ranking_every_ohio_city_county_2.html))
- Great Cleveand Area Population Estimates by Cuyahoga County Planning Commission (<https://www.countyplanning.us/resources/census-data/population-estimates/>)
- US Zip Code Latitude and Longitude by Open Data Soft (<https://public.opendatasoft.com/explore/dataset/us-zip-code-latitude-and-longitude/table/?refine.state=OH>)
- ZIP Code Database by United States Zip Codes (<https://www.unitedstateszipcodes.org/zip-code-database/>)
- Libraries: Pandas, Urllib, Geopy, Requests, Scikit Learn, and Folium

# Data requirements

- Data Cleaning

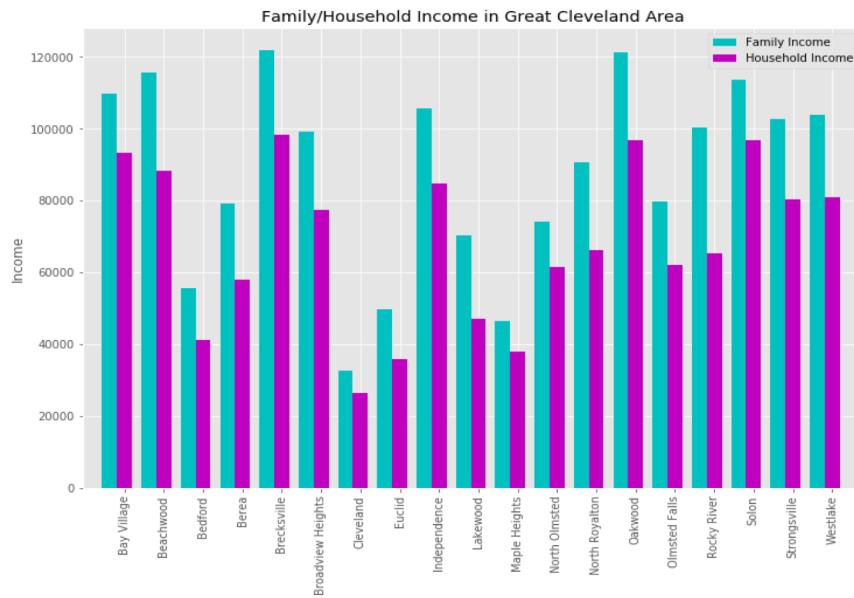
	Zip	City	State	Latitude	Longitude	Timezone	Daylight savings time flag	geopoint
0	43984	New Rumley	OH	40.296490	-81.102502	-5	1	40.29649, -81.102502
1	43681	Toledo	OH	41.686778	-83.439430	-5	1	41.686778, -83.43943
2	43733	Derwent	OH	39.923616	-81.542965	-5	1	39.923616, -81.542965
3	43334	Marengo	OH	40.399648	-82.807830	-5	1	40.399648, -82.80783
4	45841	Jenera	OH	40.881217	-83.731990	-5	1	40.881217, -83.73199
	county	state	zip	Neighborhood	Latitude	Longitude		
18504	Licking County	OH	43001.0	Alexandria	40.08	-82.61		
18505	Franklin County	OH	43002.0	Amlin	40.07	-83.18		
18506	Delaware County	OH	43003.0	Ashley	40.40	-82.95		
18507	Franklin County	OH	43004.0	Blacklick	40.02	-82.80		
18508	Knox County	OH	43005.0	Bladensburg	40.29	-82.28		
	zip	Neighborhood	Latitude	Longitude				
19056	44101	Cleveland	41.49	-81.67				
19057	44102	Cleveland	41.48	-81.74				
19058	44103	Cleveland	41.52	-81.64				
19059	44104	Cleveland	41.48	-81.63				
19060	44105	Cleveland	41.45	-81.63				

	Neighborhood	Family Income	Household Income	zip	Latitude	Longitude	2018
0	Bay Village	109673	93220	44140	41.48	-81.92	15295
1	Beachwood	115536	88287	44122	41.47	-81.50	11658
2	Bedford	55684	41285	44146	41.39	-81.53	12561
3	Berea	79306	57896	44017	41.36	-81.86	18655
4	Brecksville	121974	98345	44141	41.30	-81.61	13632
5	Broadview Heights	99286	77480	44147	41.32	-81.68	19202

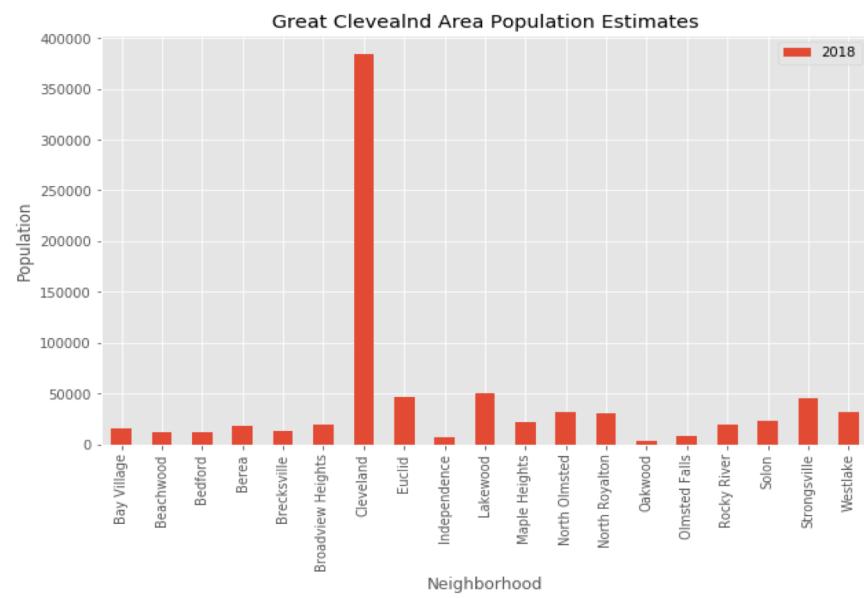
# Data requirements

- Data Understanding & Visualization

Rank	City	2012-2016 Median family income	2007-2011 Median family income	2012-2016 Median household income	2007-2011 Median household income
0	1 New Albany	214850	202243	191375	182321
1	2 Indian Hill	211795	\$250,000+	205221	243500
2	3 Pepper Pike	188765	171547	164471	151250
3	4 Powell	146442	151066	132917	142083
4	5 Dublin	144005	143619	125540	121431

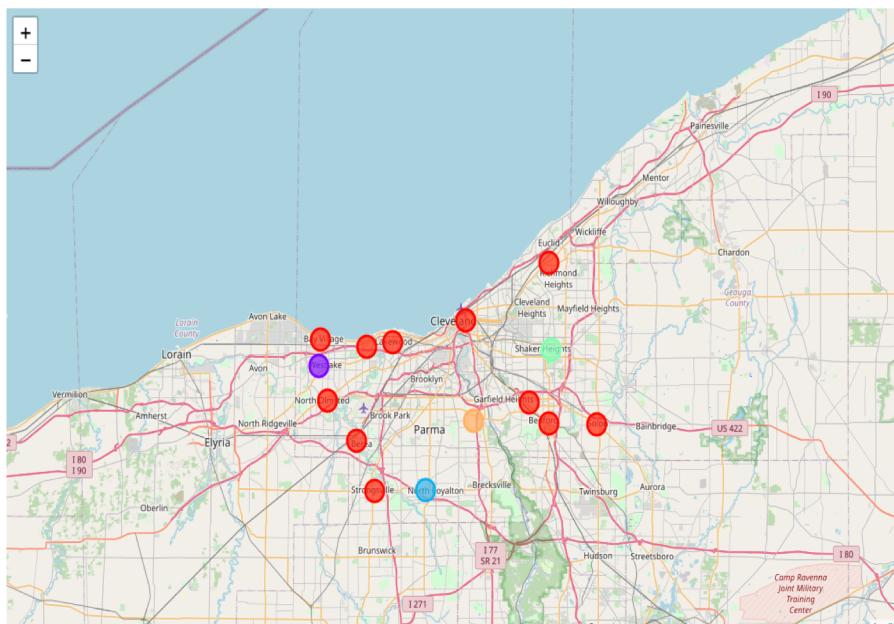


Geography	census	Estimate Base	[ 5 ] :									
			2010	2011	2012	2013	2014	2015	2016	2017	2018	
0 Bay Village	15651		15651	15629	15544	15491	15484	15470	15428	15378	15343	15295
1 Beachwood	11953		11927	11909	11844	11814	11803	11785	11743	11703	11677	11658
2 Bedford	13074		13074	13052	12966	12909	12879	12831	12767	12704	12626	12561
3 Bedford Heights	10751		10757	10740	10692	10654	10729	10725	10680	10633	10579	10534
4 Bentleyville	864		864	863	864	859	859	857	855	854	853	851



# Results

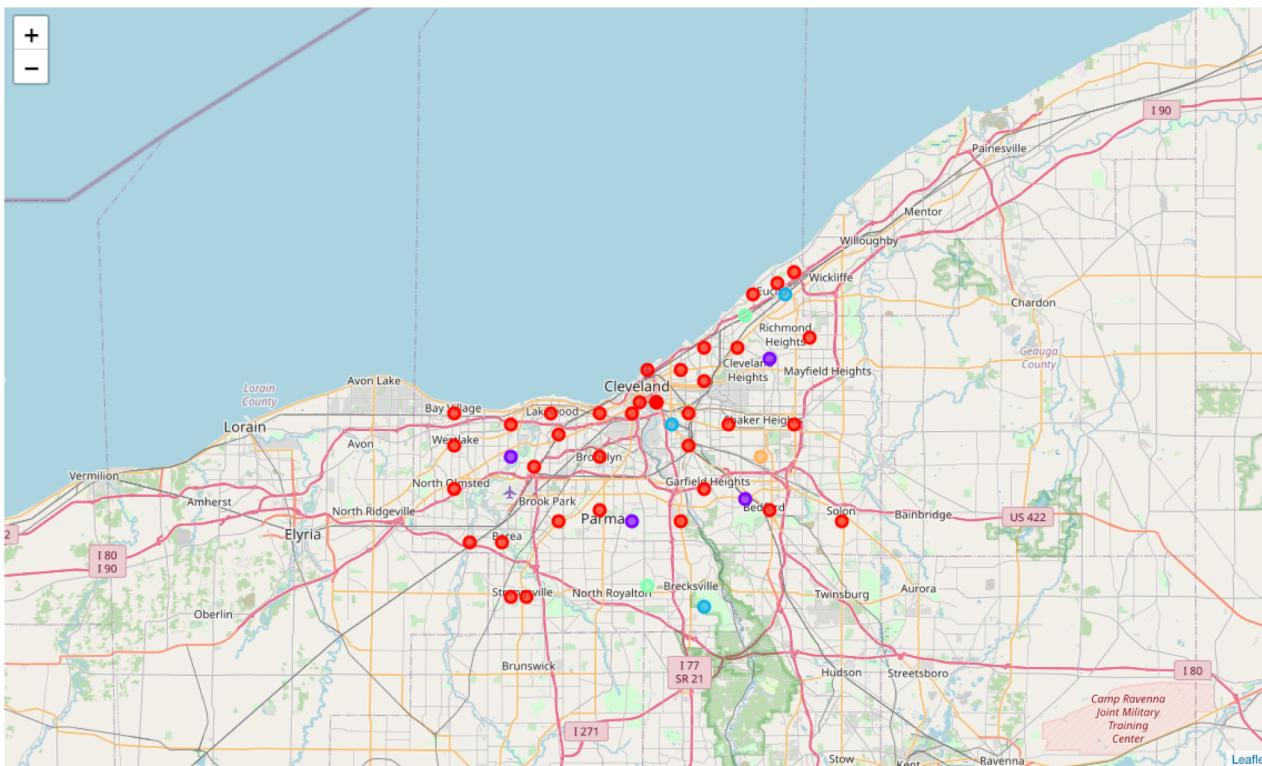
- Map of Cleveland and K-Means Clusters (K = 5, Cluster of Cities)



Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0 Bay Village	41.484193	-81.926580	2.0	Gym / Fitness Center	Liquor Store	Park	Sandwich Place	Mobile Phone Shop
1 Beachwood	41.472881	-81.522090	4.0	Intersection	Golf Course	Women's Store	Football Stadium	Dim Sum Restaurant
2 Bedford	41.389371	-81.527090	2.0	Bar	Home Service	Pharmacy	Video Store	Supermarket
3 Berea	41.369950	-81.862590	2.0	Pet Store	Fast Food Restaurant	Pharmacy	Vietnamese Restaurant	Discount Store
6 Cleveland	41.505744	-81.672797	2.0	Grocery Store	Gas Station	Bowling Alley	Shopping Plaza	Sandwich Place
7 Euclid	41.570344	-81.527260	2.0	American Restaurant	Discount Store	Sandwich Place	Bar	Fast Food Restaurant
8 Independence	41.391753	-81.658160	2.0	Salad Place	Financial or Legal Service	Hotel	Mobile Phone Shop	Dim Sum Restaurant
9 Lakewood	41.480881	-81.800360	2.0	Sandwich Place	Mobile Phone Shop	Salon / Barbershop	Bakery	Bank
10 Maple Heights	41.412653	-81.560670	2.0	Hardware Store	Sandwich Place	Discount Store	Basketball Court	Farmers Market
11 North Olmsted	41.415097	-81.914360	2.0	Furniture / Home Store	Mobile Phone Shop	American Restaurant	Bank	Shoe Store

# Results

- Map of Cleveland and K-Means Clusters (Clusters of ZIP Code Areas)



# Results

- Cluster of ZIP Code Areas

zip	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 44140	Coffee Shop	Gym / Fitness Center	Diner	Pizza Place	Pharmacy	New American Restaurant	Italian Restaurant	Ice Cream Shop	Grocery Store	Furniture / Home Store
1 44122	American Restaurant	Cycle Studio	Steakhouse	Park	Cosmetics Shop	Cupcake Shop	Convenience Store	Dance Studio	Falafel Restaurant	Diner
2 44146	Bar	Bowling Alley	Pharmacy	Shipping Store	Lawyer	Café	Supermarket	Pizza Place	Breakfast Spot	Video Store
3 44017	Lake	Scenic Lookout	Snack Place	Discount Store	Falafel Restaurant	Entertainment Service	Dry Cleaner	Donut Shop	Dive Bar	Yoga Studio
6 44129	Coffee Shop	Business Service	Yoga Studio	Farm	Cosmetics Shop	Cupcake Shop	Cycle Studio	Dance Studio	Department Store	Diner

zip	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
8 44134	Park	Construction & Landscaping	Farm	Convenience Store	Cosmetics Shop	Cupcake Shop	Cycle Studio	Dance Studio	Department Store	Diner
39 44126	Park	Beer Garden	Discount Store	Falafel Restaurant	Entertainment Service	Dry Cleaner	Donut Shop	Dive Bar	Yoga Studio	Farm
45 44121	Park	Furniture / Home Store	Construction & Landscaping	Convenience Store	Cosmetics Shop	Cupcake Shop	Cycle Studio	Dance Studio	Farm	Department Store
53 44137	Construction & Landscaping	Concert Hall	Convenience Store	Cosmetics Shop	Cupcake Shop	Cycle Studio	Dance Studio	Department Store	Diner	Discount Store

zip	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
4 44141	Trail	Stables	Yoga Studio	Diner	Entertainment Service	Dry Cleaner	Donut Shop	Dive Bar	Discount Store	Department Store
25 44127	Bus Station	Pizza Place	Intersection	Bike Rental / Bike Share	Clothing Store	Trail	Grocery Store	Cycle Studio	Dance Studio	Department Store
50 44117	Bus Station	Yoga Studio	Farm	Convenience Store	Cosmetics Shop	Cupcake Shop	Cycle Studio	Dance Studio	Department Store	Diner

# Conclusion and Suggestions

- Based on the analysis, we recommended the locations that are near the downtown Cleveland cluster
- The demographics, accessibility, and labor costs, are also the important factors for opening a new restaurant.
- The next step will be thinking about the other factors such as the operating cost and so on

