

The Detroit Project

Table of Contents

I. INTRODUCTION: THE BUSINESS PROBLEM.....	1
II. THE DATA.....	2
III. THE METHODOLOGY	4
IV. RESULTS AND ANALYSIS.....	7
V. DISCUSSION	14
VI. CONCLUSION	14

I. INTRODUCTION: THE BUSINESS PROBLEM

The Business Problem

Like many cities around the United States and the world, Detroit was profoundly impacted economically by the financial crisis of 2008. In addition to this global shock, Detroit was already in the process of an economic decline over the preceding years and decades. The Detroit of world renown was an automotive manufacturing superpower, but its peak may be traced back as far as the early 1950's when the city was the fourth largest city in the United States and home to 1.8 million people.¹ Since this peak, people and businesses have left Detroit for a variety of reasons, drastically dropping the total population down to a current level of 670,000.² The city even experienced a major bankruptcy in 2013.³

In the intervening years since the 2008 financial crisis, the city of Detroit has been the subject of several programs to revitalize the city and grow the business community. These programs were initiated by a variety of prominent organizations and individuals, but despite their tremendous efforts, attracting businesses and people to a city with the economic struggles of Detroit is quite difficult. Much work still needs to be done to return Detroit to its former vibrancy.

Project Goal

The goal of this particular project is to leverage Foursquare API business data for Detroit and New York City to identify neighborhoods in Detroit that are similar to neighborhoods in New York City. The similarity of neighborhoods will be determined by the types of businesses in the neighborhoods, as found in the Foursquare database. Using the "similar neighborhood" results along with some basic demographic information (such as house value and median income), community and business leaders in Detroit can attract businesses and people from specific New York City neighborhoods to their respective "similar" Detroit neighborhoods.

¹ https://en.wikipedia.org/wiki/Demographic_history_of_Detroit

² <https://en.wikipedia.org/wiki/Detroit>

³ Ibid

Target Audience - Interested Parties and Stakeholders

The target audience for this project are the community, political and business leaders of Detroit who are actively engaged in trying to attract new businesses and people to the city. These leaders would be interested in the results of this project because it can narrow their geographic focus to specific neighborhoods in New York and also give them some good information to convince people and businesses to make the move to specific Detroit neighborhoods. New York City is one of the most prominent business centers in the world and home to many talented people that would bolster the future of any city. Due to the population congestion and high cost of living (among other things), people and businesses are moving out of New York City to other cities⁴, including Detroit⁵. If Detroit leaders can attract more of those business and people, everyone in the city stands to benefit from new economic activity.

II. THE DATA

Geolocation Data

Neighborhood names and latitude/longitude pairs for Detroit were scraped from https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Detroit, and the corresponding information for New York were obtained from https://geo.nyu.edu/catalog/nyu_2451_34572. For Detroit, some of the neighborhoods listed by Wikipedia lacked geolocation data and were removed. This is fine for the purposes of this project, as most of those neighborhoods seemed relatively minor.

Detroit data frame head

	City	Borough	Neighborhood	Latitude	Longitude
0	Detroit	Downtown	Broadway Avenue Historic District	42.33500	-83.04611
1	Detroit	Downtown	Campus Martius Park	42.33139	-83.04667
2	Detroit	Downtown	Capitol Park Historic District	42.33278	-83.04944
3	Detroit	Downtown	Detroit Financial District	42.32954	-83.04734
4	Detroit	Downtown	Grand Circus Park Historic District	42.33611	-83.05056

New York data frame head

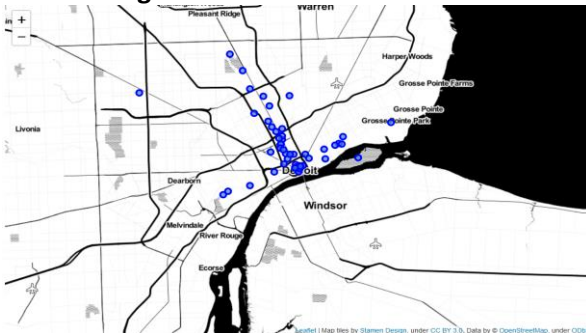
	City	Borough	Neighborhood	Latitude	Longitude
0	New York	Bronx	Wakefield	40.894705	-73.847201
1	New York	Bronx	Co-op City	40.874294	-73.829939
2	New York	Bronx	Eastchester	40.887556	-73.827806
3	New York	Bronx	Fieldston	40.895437	-73.905643
4	New York	Bronx	Riverdale	40.890834	-73.912585

⁴ Kelly, Jack, "New Yorkers Are Leaving The City In Drove: Here's Why They're Moving And Where They're Going", Forbes.com, dated Sept. 5, 2019. <https://www.forbes.com/sites/jackkelly/2019/09/05/new-yorkers-are-leaving-the-city-in-drove-heres-why-theyre-moving-and-where-theyre-going/#32201ed741ac>

⁵ Anders, Melissa, "Why moving to Detroit from New York was the 'best decision' for SPLT", Model D, dated Oct. 24, 2016. <https://www.modeldmedia.com/features/splt-detroit-102416.aspx>

In total, this project analyzed 49 neighborhoods in Detroit and 306 neighborhoods in New York City.

Detroit Neighborhoods



New York City Neighborhoods



Foursquare API Data

The Foursquare API provides a variety of data about venues visited by Foursquare app users.⁶ Using the geolocation data for the neighborhoods, the Foursquare API was queried to provide information on venues within 500 meters of latitude and longitude coordinates of the neighborhoods, with a limit of 100 venues per neighborhood. The data that used for clustering analysis was the category of each venue. As an example, the following are a few of the venues returned for the Marble Hill neighborhood in New York:

	name	categories	lat	lng
0	Arturo's	Pizza Place	40.874412	-73.910271
1	Bikram Yoga	Yoga Studio	40.876844	-73.906204
2	Tibbett Diner	Diner	40.880404	-73.908937
3	Starbucks	Coffee Shop	40.877531	-73.905582
4	Astral Fitness & Wellness Center	Gym	40.876705	-73.906372

Between the Detroit and New York neighborhoods, there are 436 venue categories which can be used to cluster the neighborhoods. The K-means clustering is based on the frequency of venue types in each neighborhood on the belief that a successful business in a New York neighborhood can also be successful in a Detroit neighborhood of the same cluster.

Demographic Data

In addition to the Foursquare-based clustering analysis, demographic data on the selected cluster neighborhoods was obtained via zip codes. First, openstreetmap.org's Nominatim API was used to convert latitude and longitude data into zip code data. Once each of the selected neighborhoods had a zip code associated with it, the zip code information was used to obtain census-type data from CDX Technologies at <https://www.cdxtech.com/tools/demographicdata/>. The returned data file includes dozens of demographic data fields for each zip code, but only the following attributes were analyzed: 'Median Age', 'Median Age Female', 'Median Age Male', 'Persons Per Households', 'Household Median

⁶ Available at <https://foursquare.com/>

Income’ and ‘Median House Value’. The following is an example of what the demographic data looks like:

	City	Borough	Neighborhood	Cluster Labels	ZipCode	Median Age	Median Age Female	Median Age Male	Persons Per Household	Household Median Income	Median House Value
0	Detroit	Midtown	Cass-Davenport Historic District	23.0	48201	30.0	29.4	30.7	1.75	15168.0	263600.0
1	Detroit	Midtown	Cultural Center Historic District	23.0	48202	36.7	36.7	36.7	1.91	26164.0	92100.0
2	Detroit	Midtown	East Ferry Avenue Historic District	23.0	48202	36.7	36.7	36.7	1.91	26164.0	92100.0
3	Detroit	Midtown	Midtown Woodward Historic District	23.0	48201	30.0	29.4	30.7	1.75	15168.0	263600.0
4	Detroit	Midtown	University-Cultural Center	23.0	48201	30.0	29.4	30.7	1.75	15168.0	263600.0

III. THE METHODOLOGY

Overview

This project and its analysis entailed (i) collecting latitude and longitude data on neighborhoods in both Detroit and New York City, (ii) using that geographical data to collect information through the Foursquare API on business types in those neighborhoods and (iii) using a K-means clustering algorithm to cluster the neighborhoods based on their respective business types. K-means clustering is an algorithm that takes a set of data points and a specified number of clusters and partitions those data points into that number of clusters. The algorithm attempts to minimize the squared “distance” from the cluster data points to the center of the cluster. In a simple version in 2-D space, the distance could be physical distance on a geographical map. In this case, the “distance” is along the Foursquare “Venue Category” space, which has 436 dimensions. The algorithm doesn’t guarantee an optimal solution, but it should yield clusters where the neighborhoods have a substantially similar “Venue Category” nature.⁷

After clusters containing both Detroit and New York neighborhoods were identified, additional demographic information was obtained on those neighborhoods using zip code data. The combined data was then used to analyze the clusters.

Step 1 – Get Foursquare API Venue Information for each Neighborhood

In the first step, the data frame containing each neighborhood and its respective latitude and longitude coordinates was used to make an “explore” query to the Foursquare API. The Foursquare API returned venues within 500 meters of the listed coordinates, up to a limit of 100 venues per neighborhood. In total, 11,841 venues were returned. Below is a snapshot of the head of the resulting data frame:

⁷ K-means algorithm description from https://en.wikipedia.org/wiki/K-means_clustering

	City	Borough	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Detroit	Downtown	Broadway Avenue Historic District	42.335	-83.04611	The Belt	42.334386	-83.046087	Art Gallery
1	Detroit	Downtown	Broadway Avenue Historic District	42.335	-83.04611	Vault of Midnight	42.334135	-83.046621	Comic Shop
2	Detroit	Downtown	Broadway Avenue Historic District	42.335	-83.04611	The Standby	42.334439	-83.046009	New American Restaurant
3	Detroit	Downtown	Broadway Avenue Historic District	42.335	-83.04611	Madcap Coffee	42.334530	-83.048220	Coffee Shop
4	Detroit	Downtown	Broadway Avenue Historic District	42.335	-83.04611	Vicente's Cuban Cuisine	42.334436	-83.047193	Cuban Restaurant

Step 2 – Create Dummy Variable Data Frame for All Unique Venue Types

In the second step, each unique “Venue Category” value was turned into a column in a data frame that is essentially a matrix of dummy variables for each venue/neighborhood combination. Next, the data frame is grouped by neighborhood, with each venue category having a value equal to the mean of all the dummy variables in that category for that neighborhood. For example, if a neighborhood had 100 venues returned from Foursquare and 20 of them were in the “Arcade” category, that category would have a value of 0.2. This matrix of means are what was used to do the K-means clustering analysis. A snapshot of the data frame looks like this:

	City	Borough	Neighborhood	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	Airport Terminal	American Restaurant	Antique Shop	Arcade	Arepa Restaurant	Argentinian Restaurant	Art Gallery
0	Detroit	Downtown	Broadway Avenue Historic District	0.0	0.0	0.0	0.0	0.0	0.030000	0.0	0.0	0.0	0.0	0.020000
1	Detroit	Downtown	Campus Martius Park	0.0	0.0	0.0	0.0	0.0	0.050000	0.0	0.0	0.0	0.0	0.020000
2	Detroit	Downtown	Capitol Park Historic District	0.0	0.0	0.0	0.0	0.0	0.050000	0.0	0.0	0.0	0.0	0.010000
3	Detroit	Downtown	Detroit Financial District	0.0	0.0	0.0	0.0	0.0	0.040000	0.0	0.0	0.0	0.0	0.000000
4	Detroit	Downtown	Grand Circus Park Historic District	0.0	0.0	0.0	0.0	0.0	0.026667	0.0	0.0	0.0	0.0	0.013333

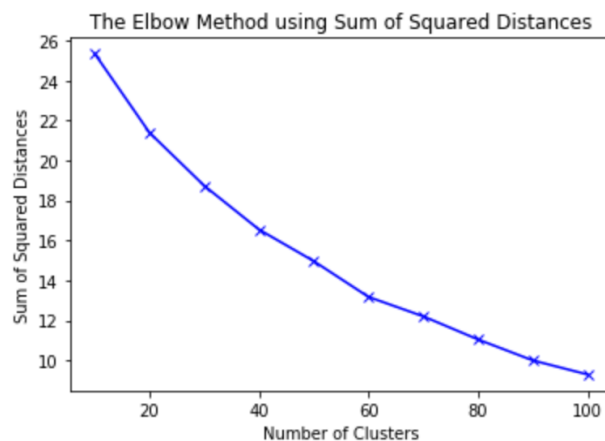
Step 3 – Identify Top 10 Venues of Each Neighborhood

Before doing the clustering analysis, we can use the matrix of venue category means to identify the top 10 venues for each neighborhood, based on frequency of venue type. This information will be used in the “Results and Analysis” section. The resulting data frame head looks like this:

	City	Borough	Neighborhood	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	Detroit	Downtown	Broadway Avenue Historic District	Coffee Shop	Lounge	Hotel	Bar	Burger Joint	Pizza Place	Park	American Restaurant	Cocktail Bar	New American Restaurant
1	Detroit	Downtown	Campus Martius Park	Coffee Shop	Cocktail Bar	American Restaurant	Bar	Steakhouse	New American Restaurant	Lounge	Sandwich Place	Hotel	Restaurant
2	Detroit	Downtown	Capitol Park Historic District	Coffee Shop	Cocktail Bar	American Restaurant	Hotel	Restaurant	New American Restaurant	Sandwich Place	Lounge	Thai Restaurant	Steakhouse
3	Detroit	Downtown	Detroit Financial District	Coffee Shop	Sandwich Place	Restaurant	Hotel	Café	American Restaurant	Steakhouse	New American Restaurant	Pizza Place	Diner
4	Detroit	Downtown	Grand Circus Park Historic District	Coffee Shop	Hotel	Bar	Lounge	New American Restaurant	Jazz Club	Park	Cocktail Bar	Sports Bar	Music Venue

Step 4 – Use Elbow Method to Identify an Acceptable Number of Clusters

The elbow method was used to determine the optimal number of clusters, however, a true optimal number was not identified. Rather, the sum of squared distances between the data points and the cluster centers (the “inertia”) continued to decrease in an approximately linear fashion without a clear inflection point. The elbow method is a heuristic algorithm that charts the inertia of each K-means cluster set so that a person can visually identify an inflection point where additional clusters decrease the inertia far less than previous increases in the number of clusters.⁸ The method gets its name from the elbow shape it is supposed to have at the inflection point, which does not always occur. The elbow method in this project resulted in the following:



The inertia of the algorithm will continue to decrease to 0 until the number of clusters equals the number of data points, but that is not the goal. The goal is to get meaningful clusters with related data points. Accordingly, the number of clusters was set to 100, which yielded 3 moderately sized clusters that each contained both Detroit and New York neighborhoods.

⁸ Elbow method information from <https://www.geeksforgeeks.org/elbow-method-for-optimal-value-of-k-in-kmeans/>

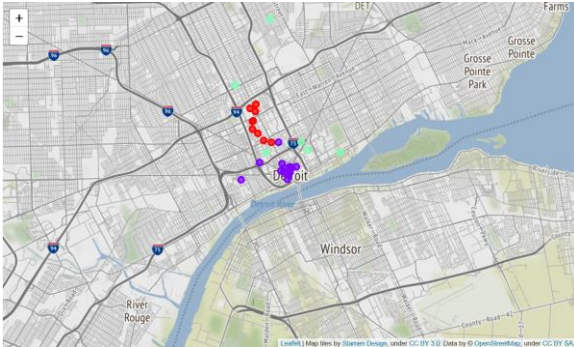
IV. RESULTS AND ANALYSIS

Overall Cluster Results and Maps

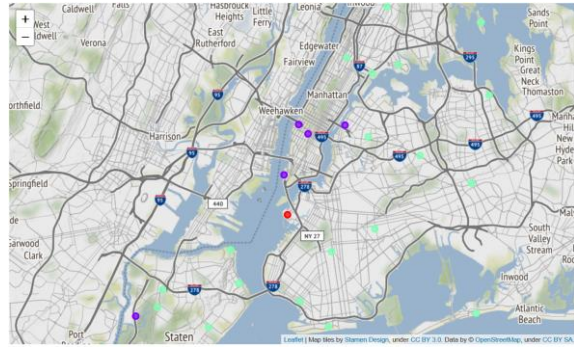
The results of the K-means cluster analysis yielded 3 clusters that contained both Detroit and New York neighborhoods. These are clusters numbered 23, 44 and 78, which are the focus of this report. First, let's visualize these clusters on a map:

- Cluster 23
- Cluster 44
- Cluster 78

Detroit Neighborhood Clusters



New York Neighborhood Clusters



Cluster 23 Analysis – A “Social” / “Residential” Cluster

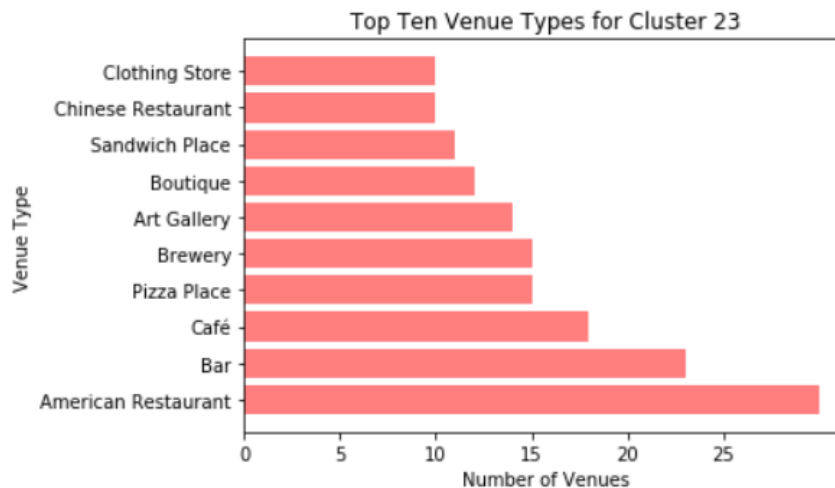
Cluster 23 consists of the following New York and Detroit neighbors:

New York neighborhoods to target:

	City	Borough	Neighborhood
9	New York	Brooklyn	Red Hook

</

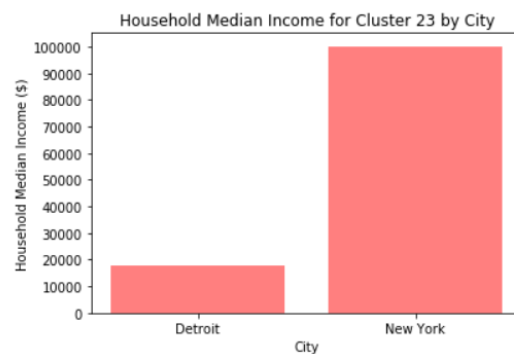
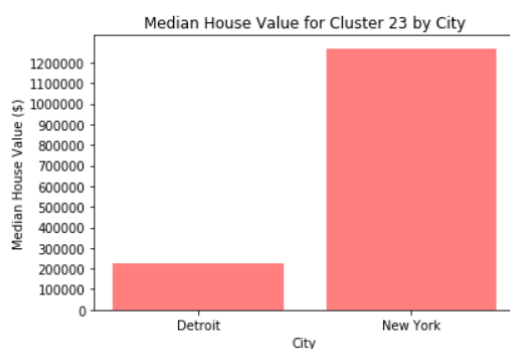
In order to evaluate what kind of neighborhoods make up this cluster, let's look at the top ten venue types.



The top 10 venue types for Cluster 23 primarily consist of eating establishments, bars, clothing stores and art galleries. From a qualitative perspective, this appears to be a residential kind of area with an active social scene. One can imagine people living and working in these areas, walking around and shopping during the days, and going out to restaurants and bars at night.

A demographic summary of the Detroit and New York portions of the cluster are as follows:

	Cluster Labels	Median Age	Median Age Female	Median Age Male	Persons Per Household	Household Median Income	Median House Value
City							
Detroit	23.0	31.49	31.02	32.03	1.79	17611.56	225488.89
New York	23.0	35.90	35.90	35.90	2.41	100213.00	1268600.00



It's clear from the summary table above that the Cluster 23 New York profile is a little older and with a larger household than in Detroit. The demographics also show that the income and house value in New York is approximately 5.6 times greater than in Detroit.

This information can be viewed many ways, but one perspective for people looking to attract business to Detroit is that Cluster 23 New Yorkers who move to Detroit will experience a lower cost of living for themselves and a lower cost of labor for their businesses.


The following are a randomly selected list of Cluster 23 New York businesses that could be targeted to move to a Cluster 23 neighborhood in Detroit. These are only examples, not recommendations, and some may not be appropriate to move to Detroit or may already exist in Detroit (such as a chain store). In addition, targeted businesses do not need to come from the “top 10” business types in a cluster. Any business that thrives in a cluster could be targeted to move another neighborhood in the same cluster.

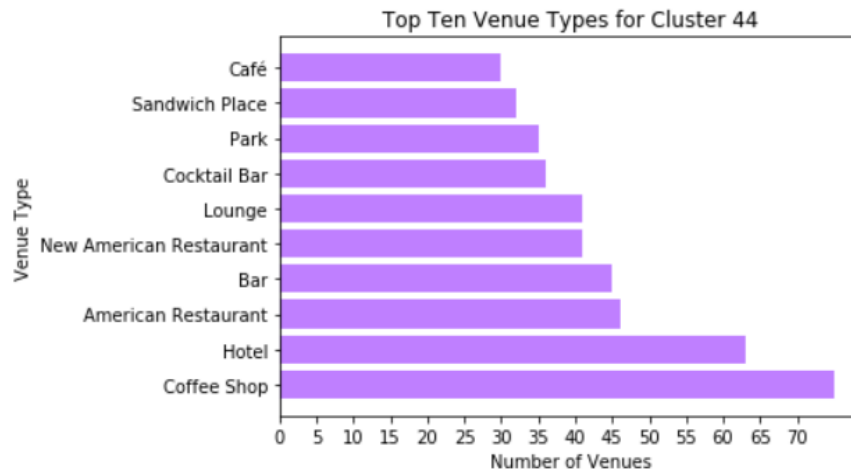
Cluster 23 New York Target Examples:

	City	Borough	Neighborhood	Venue	Venue Category
43	New York	Brooklyn	Red Hook	Kevin's	Seafood Restaurant
10	New York	Brooklyn	Red Hook	Pioneer Works	Event Space
20	New York	Brooklyn	Red Hook	Record Shop	Record Shop
32	New York	Brooklyn	Red Hook	Red Hook Labs	Art Gallery
3	New York	Brooklyn	Red Hook	Court Street Grocers Hero Shop	Sandwich Place
2	New York	Brooklyn	Red Hook	Baked	Bakery
18	New York	Brooklyn	Red Hook	Sixpoint Brewery	Brewery
23	New York	Brooklyn	Red Hook	The Black Flamingo	Café
25	New York	Brooklyn	Red Hook	WOODEN SLEEPERS	Clothing Store
33	New York	Brooklyn	Red Hook	Reed Park	Seafood Restaurant

Cluster 44 Analysis – A “Night Out” Cluster

Let's take a look at the Cluster 44 neighborhoods and top 10 venue types:

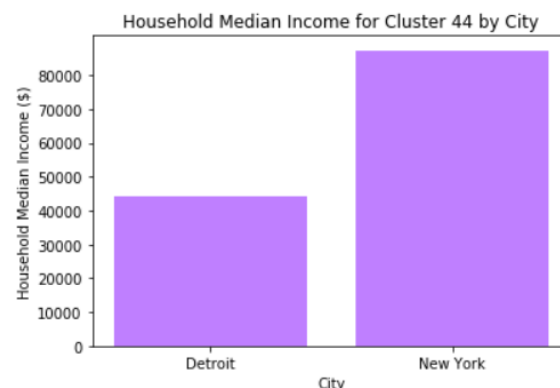
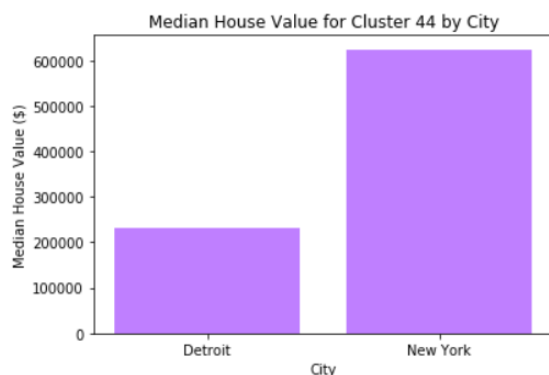
New York neighborhoods to target:				Detroit neighborhoods the New Yorkers can move to:			
	City	Borough	Neighborhood		City	Borough	Neighborhood
23	New York	Staten Island	Travis	10	Detroit	Downtown	Broadway Avenue Historic District
24	New York	Manhattan	Midtown South	11	Detroit	Downtown	Campus Martius Park
25	New York	Manhattan	Hudson Yards	12	Detroit	Downtown	Capitol Park Historic District
26	New York	Queens	Queensbridge	13	Detroit	Downtown	Detroit Financial District
				14	Detroit	Downtown	Grand Circus Park Historic District
				15	Detroit	Downtown	Lower Woodward Avenue Historic District
				16	Detroit	Downtown	Monroe Avenue Commercial Buildings
				17	Detroit	Downtown	Park Avenue Historic District
				18	Detroit	Downtown	Randolph Street Commercial Buildings Historic ...
				19	Detroit	Downtown	Washington Boulevard Historic District
				20	Detroit	Midtown	Brush Park / Woodward East
				21	Detroit	Southwest	Corktown



Based on the top 10 venue types, Cluster 44 looks like a group of neighborhoods where people go to spend a weekend or have a night out. Most of the establishments are bars and restaurants, with hotels and parks as well. These are places people may go for recreation typically. These neighborhoods may have other significant features that don't show up in the activities of Foursquare users.

Let's see what the demographics of this cluster look like.

	Cluster Labels	Median Age	Median Age Female	Median Age Male	Persons Per Household	Household Median Income	Median House Value
City							
Detroit	44.0	32.23	32.92	32.25	1.53	44430.67	230966.67
New York	44.0	36.10	35.72	36.80	2.18	87393.50	624125.00



As we Cluster 23, Cluster 44 shows that the New York part of the cluster is older with a larger household. However, in Cluster 44 the median level of income for Detroit is $\frac{1}{2}$ that of New York while the Detroit house value is $\frac{1}{3}$ rd that of New York. The community leaders in Detroit can argue that Cluster 44 New Yorkers who move to Detroit will not only get a lower cost of living and cost of labor, but will also spend less of their income on housing.

Cluster 44 New York Target Examples:

	City	Borough	Neighborhood	Venue	Venue Category
157	New York	Manhattan	Hudson Yards	Uncle Jack's Steakhouse	Steakhouse
173	New York	Queens	Queensbridge	Queensboro Bridge Pedestrian & Bike Path	Scenic Lookout
166	New York	Manhattan	Hudson Yards	George's	Burger Joint
19	New York	Manhattan	Midtown South	Samwon Garden BBQ	Korean Restaurant
74	New York	Manhattan	Midtown South	Made Nice	Restaurant
116	New York	Manhattan	Hudson Yards	Baryshnikov Arts Center	Theater
67	New York	Manhattan	Midtown South	Five Senses	Korean Restaurant
167	New York	Manhattan	Hudson Yards	Jake's	American Restaurant
18	New York	Manhattan	Midtown South	Louis Vuitton	Boutique
101	New York	Manhattan	Midtown South	Take 31	Korean Restaurant

Cluster 78 Analysis – A “Residential” Cluster

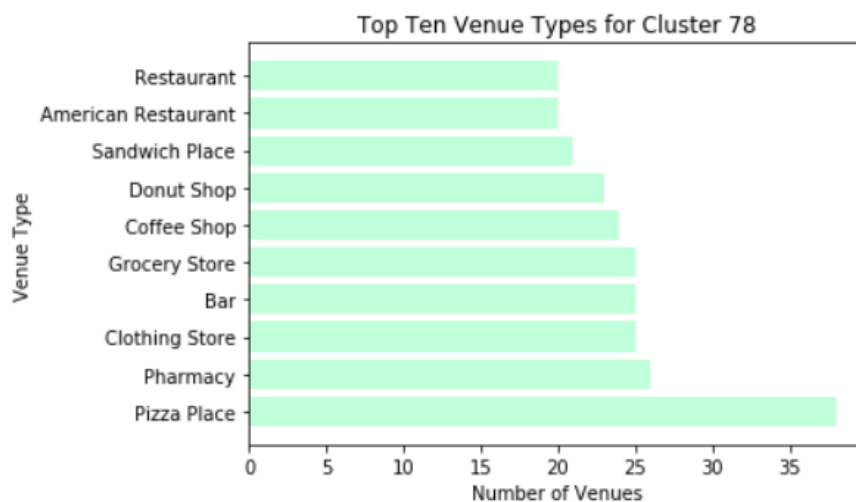
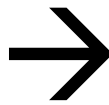
Let’s take a look at the Cluster 78 neighborhoods and top 10 venue types:

New York neighborhoods to target:

	City	Borough	Neighborhood
33	New York	Manhattan	Marble Hill
34	New York	Bronx	City Island
35	New York	Bronx	Fordham
36	New York	Bronx	Port Morris
37	New York	Bronx	Hunts Point
38	New York	Brooklyn	Brighton Beach
39	New York	Brooklyn	Brownsville
40	New York	Brooklyn	Bath Beach
41	New York	Manhattan	Central Harlem
42	New York	Queens	Rego Park
43	New York	Queens	Jamaica Center
44	New York	Queens	Bay Terrace
46	New York	Staten Island	St. George
47	New York	Staten Island	Rosebank
48	New York	Staten Island	New Springville
49	New York	Brooklyn	Ditmas Park
50	New York	Staten Island	Grant City
51	New York	Staten Island	Charleston
52	New York	Staten Island	Bulls Head
53	New York	Queens	Sunnyside Gardens

Detroit neighborhoods the New Yorkers can move to:

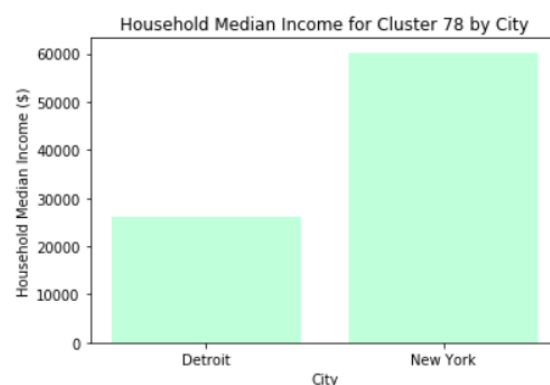
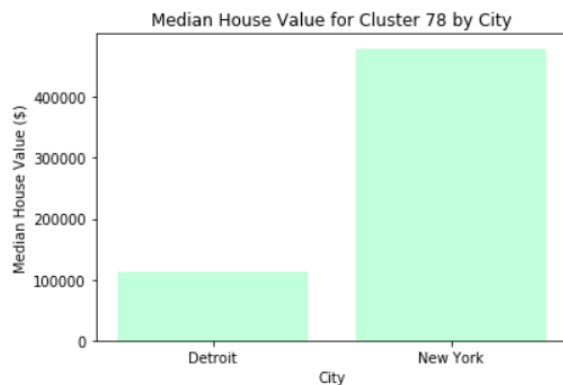
	City	Borough	Neighborhood
27	Detroit	Midtown	Cass Park Historic District
28	Detroit	New Center	New Center
29	Detroit	East Upper	Hamtramck
30	Detroit	East Central	Eastern Market
31	Detroit	East Lower	East Jefferson Avenue Residential District
32	Detroit	East Lower	Lafayette Park / Mies van der Rohe Residential...



Cluster 78 looks primarily like a residential cluster of neighborhoods. It includes a variety of quick-stop eating places, as well as pharmacies, clothing stores and grocery stores. This is what you would expect to see in an area that is substantially residential within an urban setting.

Let's take a look at the demographics of the cluster:

	Cluster Labels	Median Age	Median Age Female	Median Age Male	Persons Per Household	Household Median Income	Median House Value
City							
Detroit	78.0	37.98	37.78	38.18	2.06	26107.0	114316.67
New York	78.0	38.77	40.53	36.83	2.65	60261.8	480420.00



Unlike Clusters 23 and 44, the age difference in Cluster 78 is small, with both the median age in New York and Detroit in the late 30's. New York's household is also only slightly larger than Detroit's. The income and house value profile of Cluster 78 is very similar to Cluster 44, with New York incomes approximately twice that of Detroit, but New York house values being more than 4 times greater than in Detroit. As with Cluster 44, community leaders in Detroit can argue that New Yorkers who move will get a lower cost of living and cost of labor, and may also spend less of their income on housing.

Cluster 78 New York Target Examples:

	City	Borough	Neighborhood	Venue	Venue Category
158	New York	Bronx	Hunts Point	the point cafe & grill	Café
695	New York	Queens	Sunnyside Gardens	Ida's Nearabout	Bar
160	New York	Bronx	Hunts Point	H. Weiss Machinery & Supply	Construction & Landscaping
761	New York	Queens	Sunnyside Gardens	Boston Market	American Restaurant
292	New York	Manhattan	Central Harlem	ACP Coffee	Café
309	New York	Manhattan	Central Harlem	Pizza Stop	Pizza Place
304	New York	Manhattan	Central Harlem	Presto Fresh Cafe	Cafeteria
206	New York	Brooklyn	Brighton Beach	Домашняя кухня	Eastern European Restaurant
775	New York	Queens	Sunnyside Gardens	Paddy Duggans	Bar
545	New York	Brooklyn	Ditmas Park	Island Express	Caribbean Restaurant

V. DISCUSSION

The goal of this project was to identify similar neighborhoods between Detroit and New York City, based on the business information provided by the Foursquare API. Using the K-means clustering algorithm described above, three clusters were identified containing both Detroit and New York neighborhoods. Cluster 23 included neighborhoods that were residential but with many venues for social activities. Cluster 44 was characterized by venues a person might go to on a night out or on a weekend. Lastly, Cluster 78 appeared to be populated by purely urban residential neighborhoods.

The demographics between the New York neighborhoods and their Detroit counterparts, showed that New York was slightly older, with much higher levels of income and house values. In a broad sense, this could be translated into Detroit having a lower cost of living and cost of labor. Many New Yorkers may be turned off by Detroit's lower numbers, but some percentage of them may be attracted to them.

This project has identified 25 New York City neighborhoods that community leaders in Detroit can target for attracting new businesses to the city of Detroit. Community leaders could target business owners and other people in these neighborhoods directly, through television ads, through social media, or any number of other ways. Regardless of the method of contact, there is some value in having a narrow focus on where to direct your efforts.

This project, of course, has many limitations. First, it relies heavily on Foursquare API data to do the clustering, which limits business data to those places frequented by Foursquare users. In a broader project, a more robust business dataset would be better, perhaps incorporating non-business data as well. The demographic data used in this project does not have very much depth, and should be viewed primarily as a beginning point for analysis. The K-means clustering algorithm was not necessarily optimized, however, it did appear to deliver a small set of clusters with meaningfully related neighborhoods, so that should be seen as a success. Lastly, the full task of attracting people to a new city is multi-dimensional and well beyond the scope of this project. Other factors, such as tax policy, local politics, school systems, among other things, also matter tremendously. This project is intended to be only a small piece of that puzzle for the interested parties and stakeholders to utilize.

VI. CONCLUSION

Attracting new people and businesses to a city that is shrinking economically is a difficult task that needs to take a multi-faceted approach. One aspect of that approach is to narrow the focus of who you can attract to the city. Broad-based, generic approaches are unlikely to be very successful, particularly in an age where people and business are receiving advertisements and other content that are personalized to them. While this project is small in scope, the hope is that the people tasked with growing the City of Detroit utilize the underlying principles of using sophisticated data modeling and analytics to specifically target new growth. The success or failure of Detroit's growth initiatives will be evident many years in the future, but that future will be determined by what people do now.