



# Analysis of the commercial retail, hairdressing, Cosmetology, Beauty sectors in Chicago

IBM CAPSTONE STUDY

CLARENCE MGUTSHINI

26/08/2021

# BACKGROUND

- ▶ Chicago is considered a world- class city with a highly diversified economy. The economic activities consist of financial services, commerce, industry, education, hairdressing, cosmetology, beauty, telecommunications and transportation.
- ▶ Chicago city is a vibrant city to consider for careful and targeted development of either individual stores or shopping centres for commercial retail, hairdressing, cosmetology and beauty.

# BUSINESS PROBLEM AND NEEDS

- ▶ There is a need to find the areas in Chicago that are highly competitive in terms of commercial retail, hairdressing, cosmetology and beauty stores in the city of Chicago and to identify areas with a low number of these stores in order to identify areas for potential development.
- ▶ Property developers of new shopping centres consisting of commercial retail, hairdressing, cosmetology and beauty stores will be interested in the location information of areas with high concentrations of these stores for the purpose of competitive analysis and to establish the key reasons and factors for these areas being desirable for development of these stores.
- ▶ Furthermore, market researchers and strategists will be interested in areas with a low and high concentration of these stores for the purpose of conducting research into opportunities of possible growth in these under- represented areas.

# DATA ACQUISITION: WEB-SCRAPPING

- ▶ The data for the list of the neighbourhoods in Chicago was web-scraped from the Wikipedia page with the url given by [https://en.wikipedia.org/wiki/List\\_of\\_neighborhoods\\_in\\_Chicago](https://en.wikipedia.org/wiki/List_of_neighborhoods_in_Chicago) .
- ▶ The beautiful soup package in python was used for web-scraping the table data. There were 247 Neighbourhoods in Chicago

[5]:

	Neighborhood	Community area
0	Albany Park	Albany Park
1	Altgeld Gardens	Riverdale
2	Andersonville	Edgewater
3	Arcadia Terrace	West Ridge
4	Archer Heights	Archer Heights
...	...	...
242	Wildwood	Forest Glen
243	Woodlawn	Woodlawn
244	Wrightwood	Ashburn
245	Wrightwood Neighbors	Lincoln Park
246	Wrigleyville	Lake View

247 rows × 2 columns

# DATA ACQUISITION: LOCATION DATA

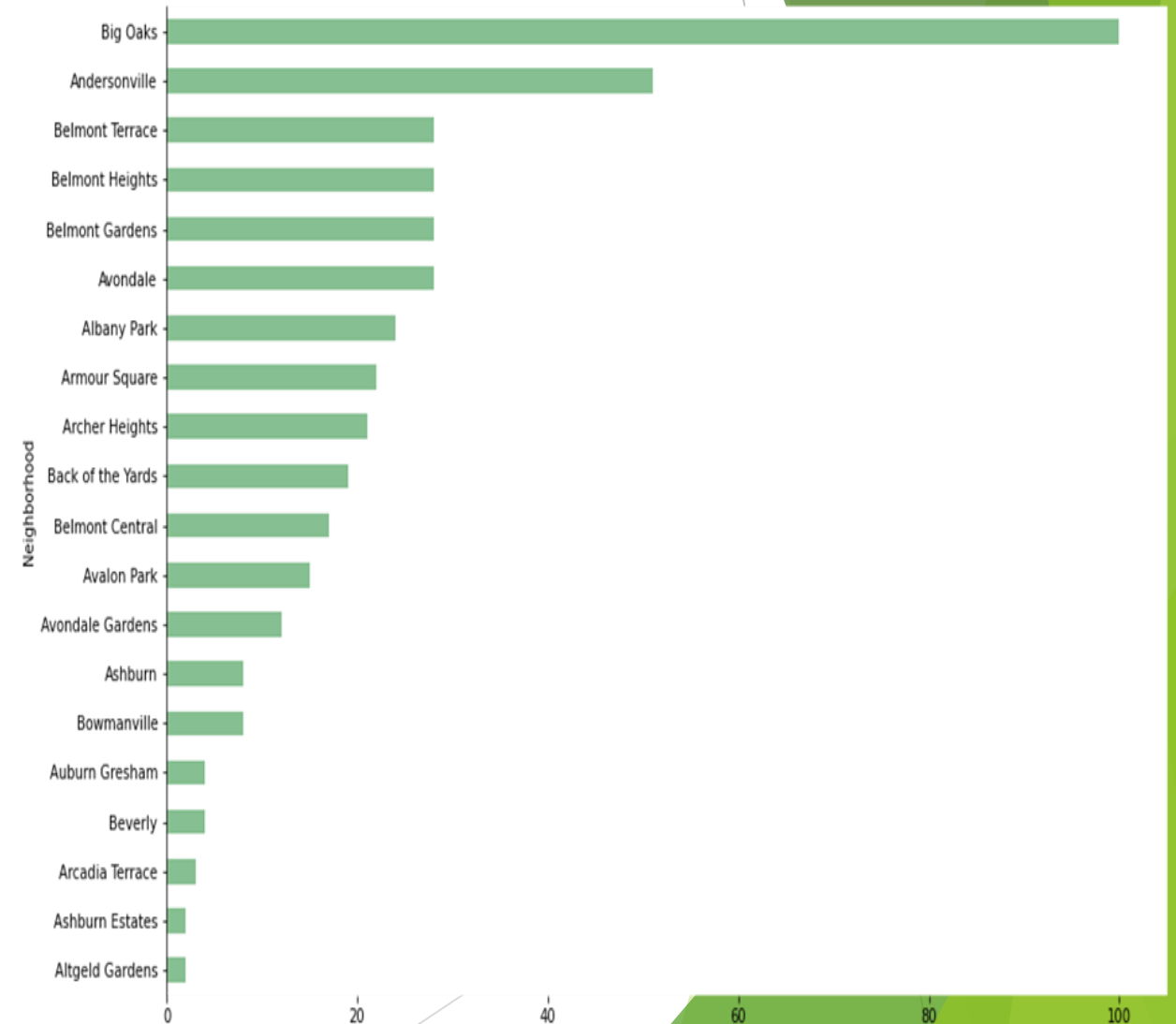
- Foursquare location data through its API, is used to acquire the 100 most common venues in the city of Chicago and to get the venues near each neighbourhood. The limit of the API was set to 100 venues and radius was set to 500 metres.
- The geocoder package was then used for scrapping the co-ordinate data of all the neighbourhoods in Chicago. Finally, the data was combined in a single data-frame which is given by the output table below:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Albany Park	41.96829	-87.72338	Lawrence Fish Market	41.968280	-87.726250	Seafood Restaurant
1	Albany Park	41.96829	-87.72338	Chicago Kalbi Korean BBQ	41.968314	-87.722771	Korean Restaurant
2	Albany Park	41.96829	-87.72338	Starbucks	41.968911	-87.728817	Coffee Shop
3	Albany Park	41.96829	-87.72338	El Gallo Bravo #6	41.968324	-87.721338	Mexican Restaurant
4	Albany Park	41.96829	-87.72338	Rojo Gusano	41.968425	-87.724549	Taco Place
...	...	...	...	...	...	...	...
6340	Wrigleyville	41.94725	-87.65320	CTA Bus Stop 1825	41.944617	-87.654612	Bus Station
6341	Wrigleyville	41.94725	-87.65320	Chase ATM	41.944031	-87.653679	ATM
6342	Wrigleyville	41.94725	-87.65320	CTA Bus Stop 5361	41.947426	-87.647445	Bus Station
6343	Wrigleyville	41.94725	-87.65320	CTA #152 Bus Stop at Addison & Racine	41.947506	-87.658993	Bus Station
6344	Wrigleyville	41.94725	-87.65320	Newport Harbor	41.944195	-87.657616	Harbor / Marina

6345 rows × 7 columns

# Exploratory Data Analysis

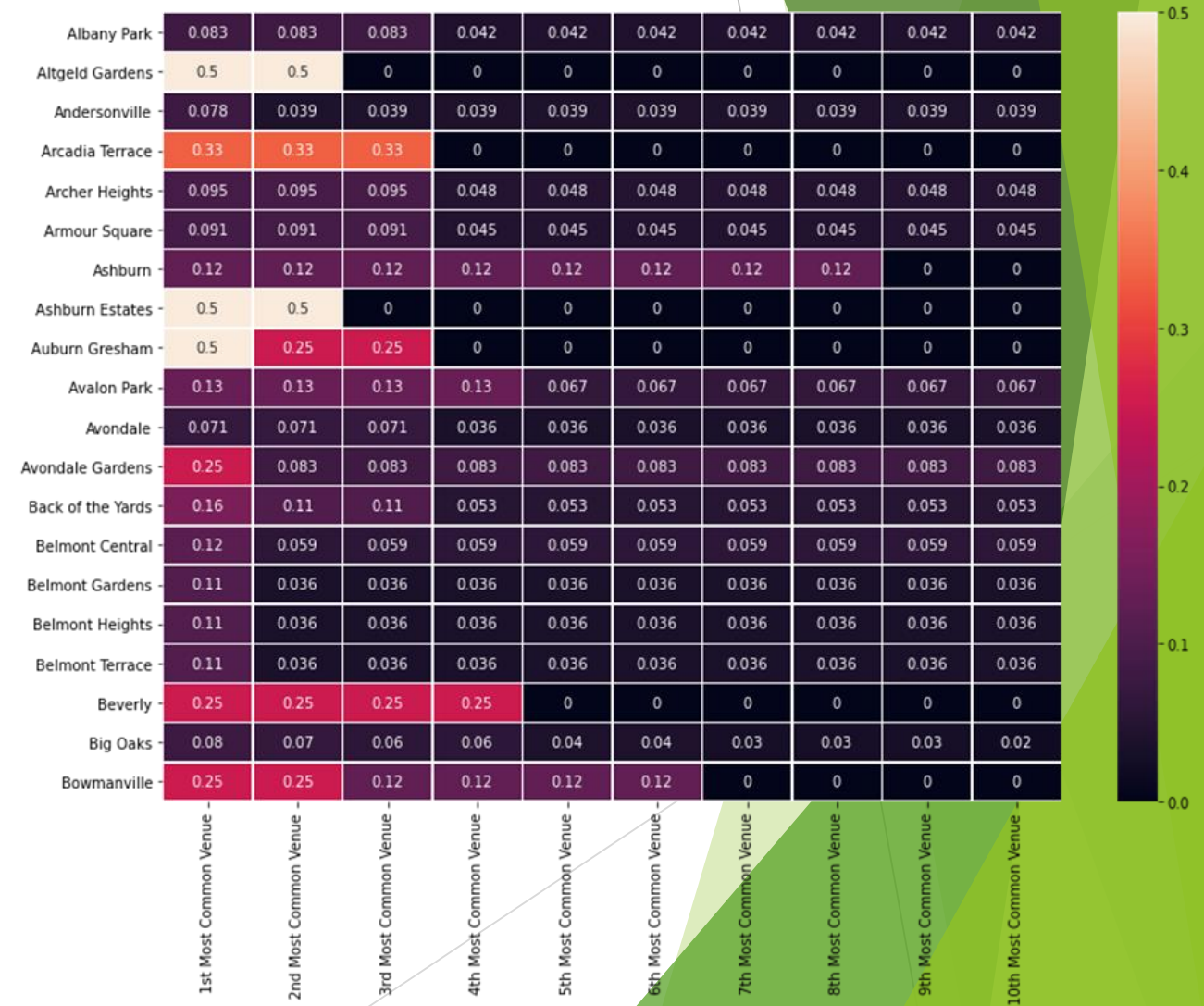
- ▶ The top 20 neighbourhoods in Chicago with the highest number of venue categories were considered for exploration.
- ▶ The Big Oaks neighbourhood area had the highest number of venue categories with 100 venue categories followed by Andersonville with 51 venue categories and Belmont Terrace with 28 venue categories





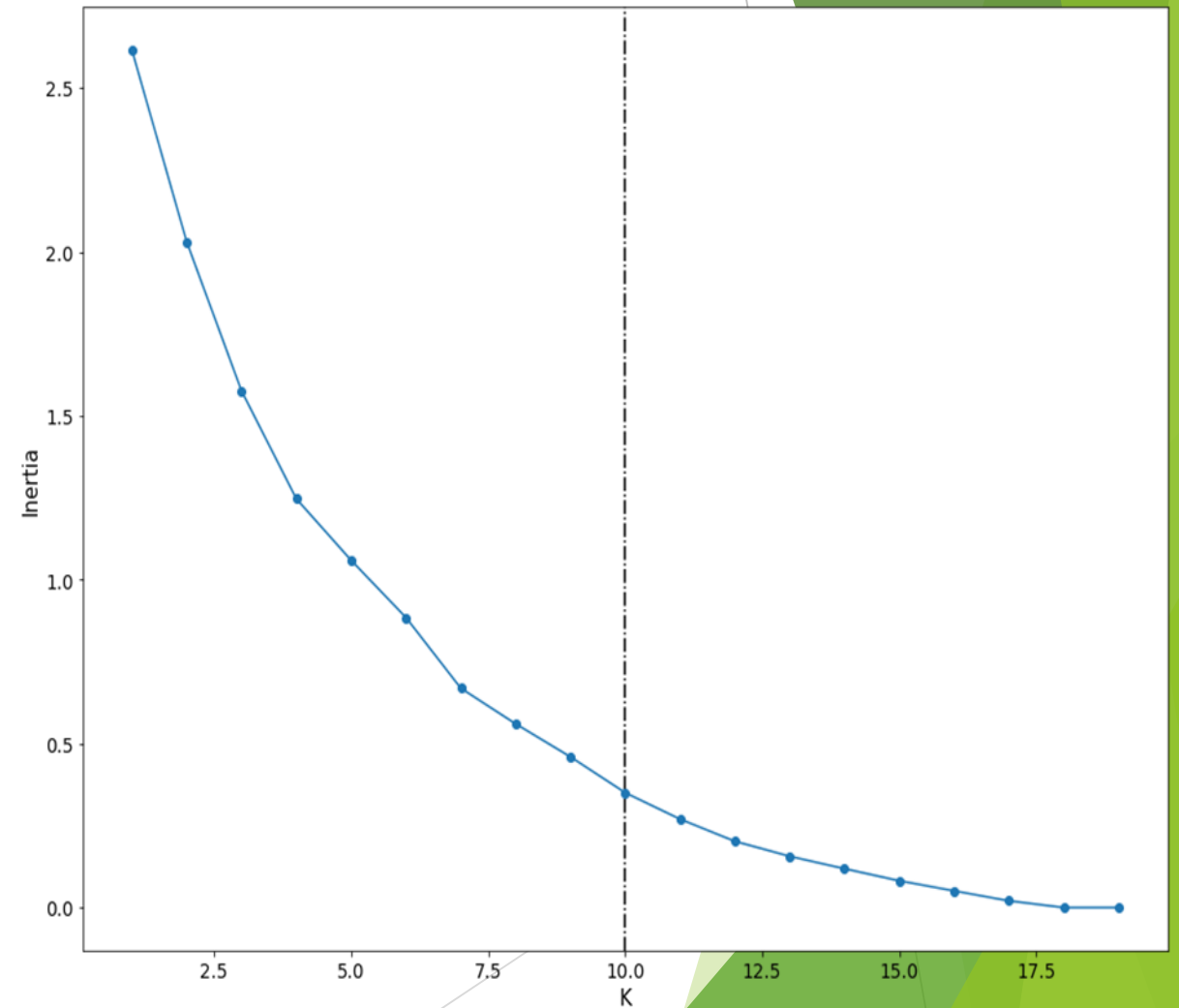
# Exploratory Data Analysis: Heat Map of Distribution of Common Venues

- The heat-map confirms that the distribution of the venue categories was skewed towards to the most commonly occurring venues.



# Cluster Analysis

- K-means analysis was used as the unsupervised classification technique
- The number of clusters used was 10 based on an elbow chart





# Results: Clusters with hair salons, commercial retail stores and cosmetic stores

- The neighbourhoods of Belmont Terrace, Belmont heights, Belmont Gardens had the highest concentration of these stores.

Cluster Labels		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
9	0	Avalon Park	Boutique	Fast Food Restaurant	Burger Joint	Pizza Place	Diner	Cajun / Creole Restaurant	BBQ Joint	Food	Grocery Store	ATM
16	0	Belmont Terrace	Salon / Barbershop	Cosmetics Shop	River	Clothing Store	Grocery Store	Other Nightlife	Golf Course	Automotive Shop	Gift Shop	Bank
15	0	Belmont Heights	Salon / Barbershop	Cosmetics Shop	River	Clothing Store	Grocery Store	Other Nightlife	Golf Course	Automotive Shop	Gift Shop	Bank
14	0	Belmont Gardens	Salon / Barbershop	Cosmetics Shop	River	Clothing Store	Grocery Store	Other Nightlife	Golf Course	Automotive Shop	Gift Shop	Bank
1	1	Allgeld Gardens	Park	Fast Food Restaurant	Zoo Exhibit	Falafel Restaurant	Electronics Store	Elementary School	Empanada Restaurant	Escape Room	Ethiopian Restaurant	Event Space

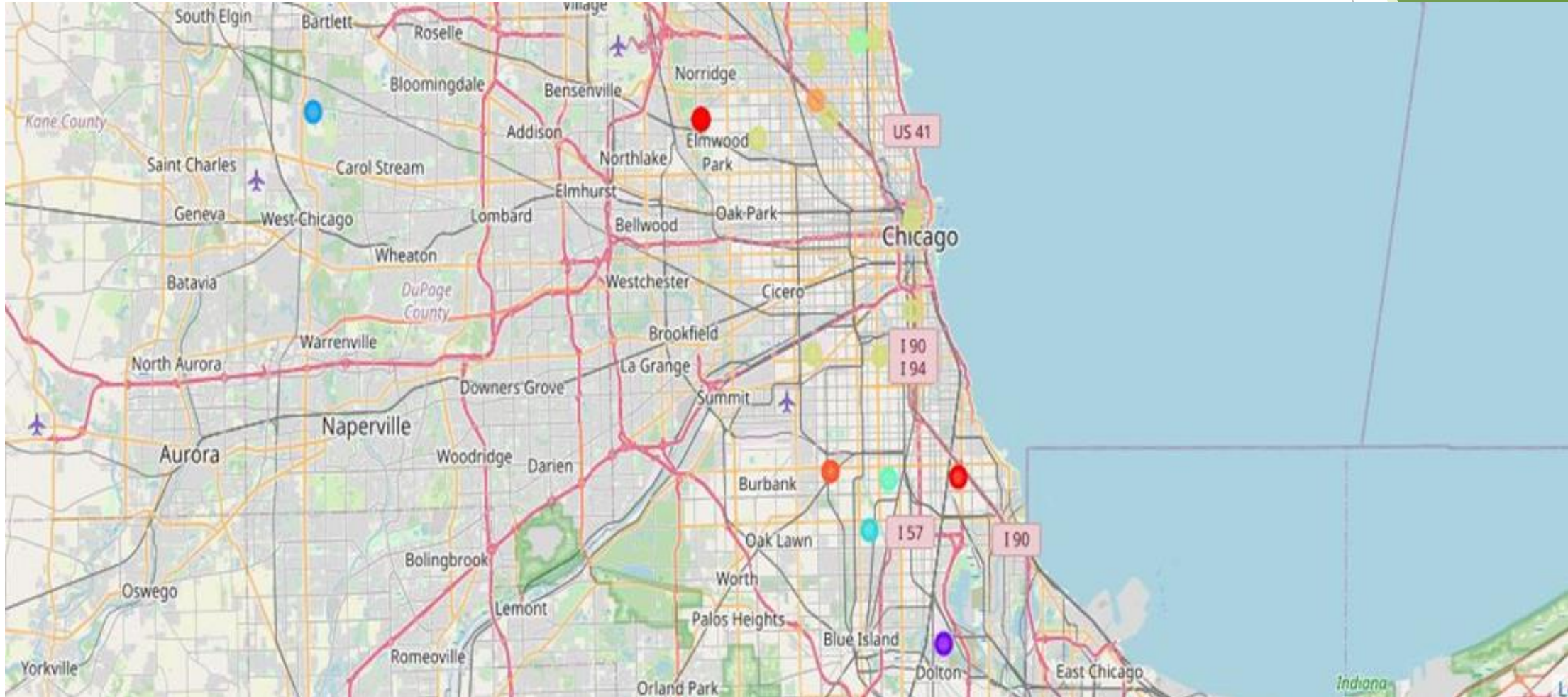
# Clusters with fewest hair salons, commercial retail stores and cosmetic stores

- These clusters consisted mainly of the restaurants, fast food outlets, coffee shops, bars and diners.

	Cluster Labels	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
7	3	Ashburn Estates	Gym	Mexican Restaurant	Farm	Electronics Store	Elementary School	Empanada Restaurant	Escape Room	Ethiopian Restaurant	Event Space	Exhibit
17	4	Beverly	Flower Shop	Platform	Park	Coffee Shop	Falafel Restaurant	Electronics Store	Elementary School	Empanada Restaurant	Escape Room	Ethiopian Restaurant
10	7	Avondale	Chinese Restaurant	Food Truck	Park	Discount Store	Gym	Grocery Store	Diner	Korean Restaurant	Road	Rental Car Location
12	7	Back of the Yards	Grocery Store	Brewery	Mexican Restaurant	Farm	Video Store	Bar	Optical Shop	Gas Station	Discount Store	Shoe Store
13	7	Belmont Central	Mexican Restaurant	Ice Cream Shop	Liquor Store	Gas Station	Automotive Shop	Chinese Restaurant	Grocery Store	Latin American Restaurant	Donut Shop	Butcher
5	7	Armour Square	Bar	Coffee Shop	American Restaurant	Park	Historic Site	Gas Station	Chinese Restaurant	Beer Garden	Restaurant	Clothing Store
4	7	Archer Heights	Mobile Phone Shop	Grocery Store	Mexican Restaurant	Optical Shop	Bar	Candy Store	Gas Station	Sandwich Place	Taco Place	Chinese Restaurant
2	7	Andersonville	Coffee Shop	Bar	Lounge	Deli / Bodega	Grocery Store	Bakery	Italian Restaurant	Café	Middle Eastern Restaurant	Miscellaneous Shop
18	7	Big Oaks	Hotel	Sandwich Place	Theater	Coffee Shop	American Restaurant	Mexican Restaurant	Donut Shop	Salad Place	Mediterranean Restaurant	Snack Place
0	7	Albany Park	Dive Bar	Bank	Mexican Restaurant	Grocery Store	Taco Place	Fried Chicken Joint	Financial or Legal Service	Fast Food Restaurant	Seafood Restaurant	Mobile Phone Shop
11	8	Avondale Gardens	Bus Station	Sandwich Place	Italian Restaurant	Thai Restaurant	Lounge	Empanada Restaurant	Gas Station	Park	BBQ Joint	Convenience Store
6	9	Ashburn	Automotive Shop	Nightclub	Cosmetics Shop	Snack Place	Bar	Martial Arts School	Light Rail Station	Bus Station	Escape Room	Ethiopian Restaurant

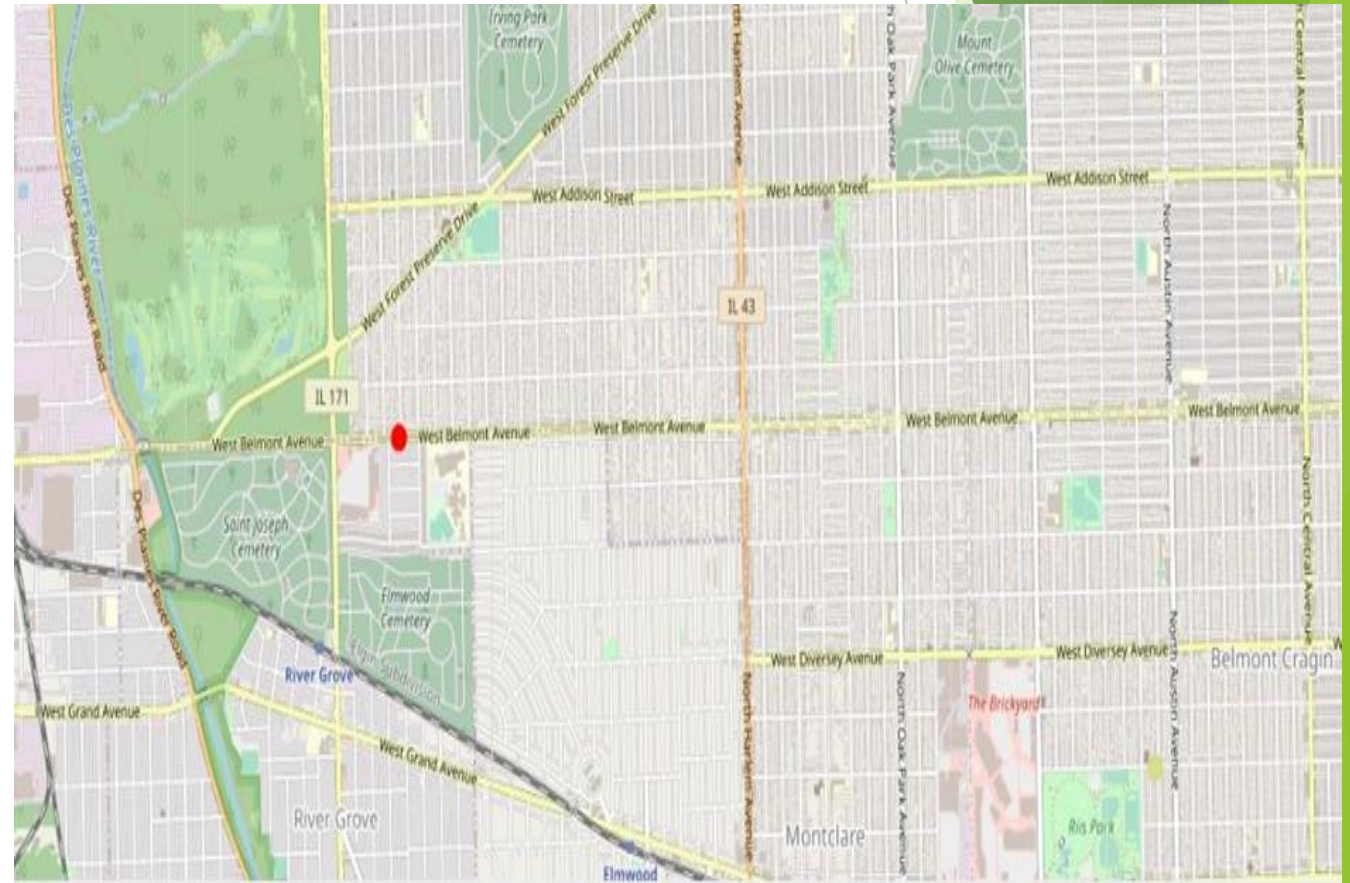


Clusters of the 20 neighbourhoods with the most number of venue categories.



# Market Research Targeting

- ▶ The neighborhoods in Belmont that had a significant commercial retail, cosmetic and beauty venues were in a single cluster.
- ▶ This area is suitable for conducting market research to determine which factors affect the desirability of commercial retail, hairdressing, cosmetic and beauty stores.



# Recommendations: Property development and Competitive Analysis

- ▶ The insights from the market research can be used to determine other suitable areas to develop commercial retail, hairdressing, cosmetic and beauty stores in the city of Chicago
- ▶ The cluster analysis also indicates which areas have high competition and which areas have low concentration of these stores and can offer growth opportunities

# Limitations and Recommendation of use of K-means clustering technique

- ▶ K-means the number of clusters to be pre-specified and the final cluster assignments are determined by the initial assignment of the clusters.
- ▶ The elbow chart was used to determine the number of clusters but the selection of the appropriate number of clusters on the chart is some what subjective and not too precise.
- ▶ Therefore, it is recommended that the k-means clustering method is repeated many times with the same number of clusters and with different number of clusters to determine whether the clusters established in this case study are robust



# Conclusion

- ▶ Large world cities like Chicago will continue to be expand and grow in the future. The classification of the neighborhoods with similar economic activities in Chicago using machine learning can be achieved.
- ▶ Cluster analysis can be used in the planning for development of infrastructure and for finding opportunities for growth in commercial stores, cosmetics, hairdressing, and beauty economic activities whilst avoiding high competition.