

The background features abstract, overlapping green geometric shapes, primarily triangles and polygons, in various shades of green, creating a modern and dynamic visual effect.

IBM & Coursera

Applied Data Science Capstone

Lindsay D
November 2020

Problem

- ▶ This research project aims to provide a recommendation for the style and location of a new restaurant in Calgary, Canada.
- ▶ There are many factors to consider when creating this recommendation such as: the number of **existing restaurants in Calgary areas and their popularity**, the popularity of **different types of cuisines** offered in Calgary, and information on the customers being served such as **age and income**.



Data Utilized

- ▶ The following list outlines data sources utilized for this analysis:
 - ▶ Calgary neighborhood data from *Wikipedia*:

Postal Code	Borough	Neighborhood	Latitude	Longitude
T1A	Medicine Hat	Central Medicine Hat	50.036460	-110.679250
T2A	Calgary	Penbrooke Meadows, Marlborough	51.049680	-113.964320
T3A	Calgary	Dalhousie, Edgemont, Hamptons, Hidden Valley	51.126060	-114.143158
T4A	Airdrie	East Airdrie	51.272450	-113.986980
T5A	Edmonton	West Clareview, East Londonderry	53.5899	-113.4413
T6A	Edmonton	North Capilano	53.5483	-113.408
T7A	Drayton Valley	Not assigned	53.2165	-114.9893
T8A	Sherwood Park	West Sherwood Park	53.519	-113.3216

Data Utilized

- ▶ The following list outlines data sources utilized for this analysis:

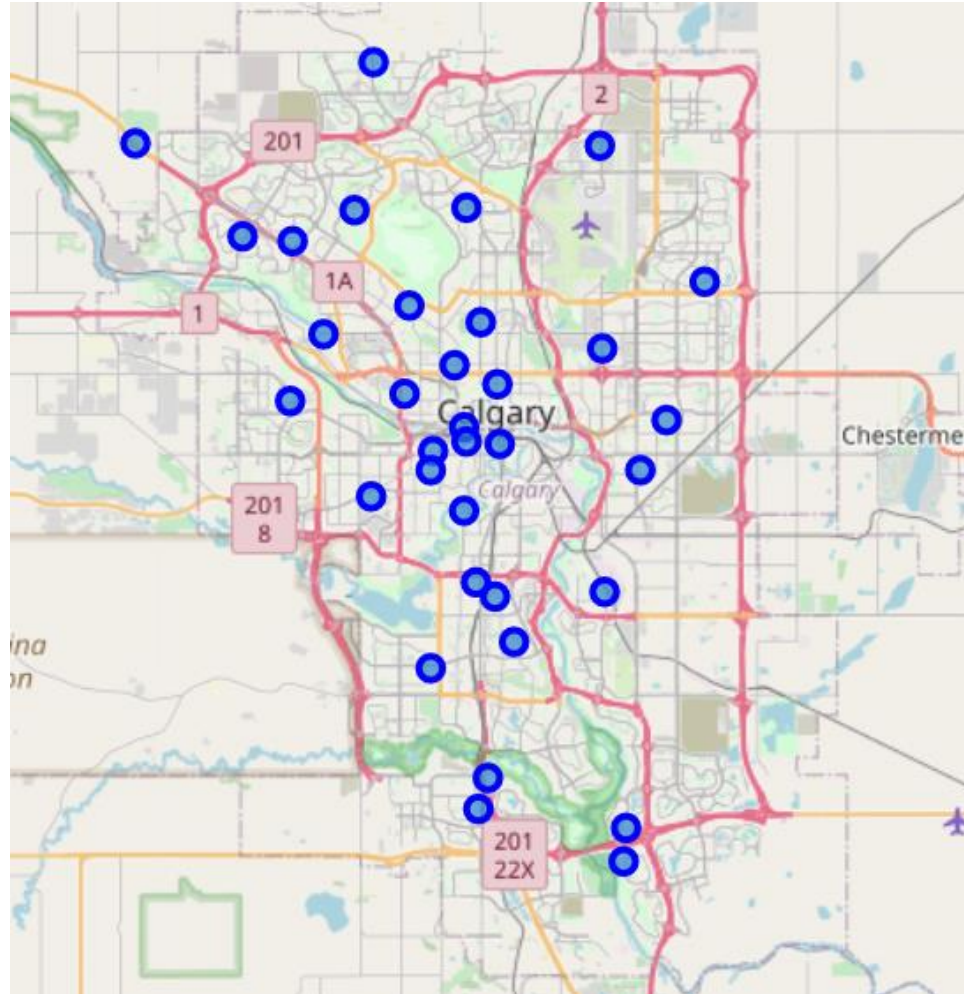
- ▶ Calgary census data from *Great-news.ca*:

Community	Median Household Income	Median Age	City Quadrant	Median Home Sale Price
Abbeydale	\$55,345	34	SE	\$305,000
Acadia	\$46,089	42	SE	\$447,000
Albert Park / Radisson Heights	\$38,019	37	SE	\$349,900
Altadore	\$53,786	37	SW	\$925,000
Applewood Park	\$65,724	33	SE	\$380,000

- ▶ Venue information from FourSquare location services

Methodology

- Step 1: Neighborhoods in Calgary were visualized on a map to provide a base-level understanding of the city



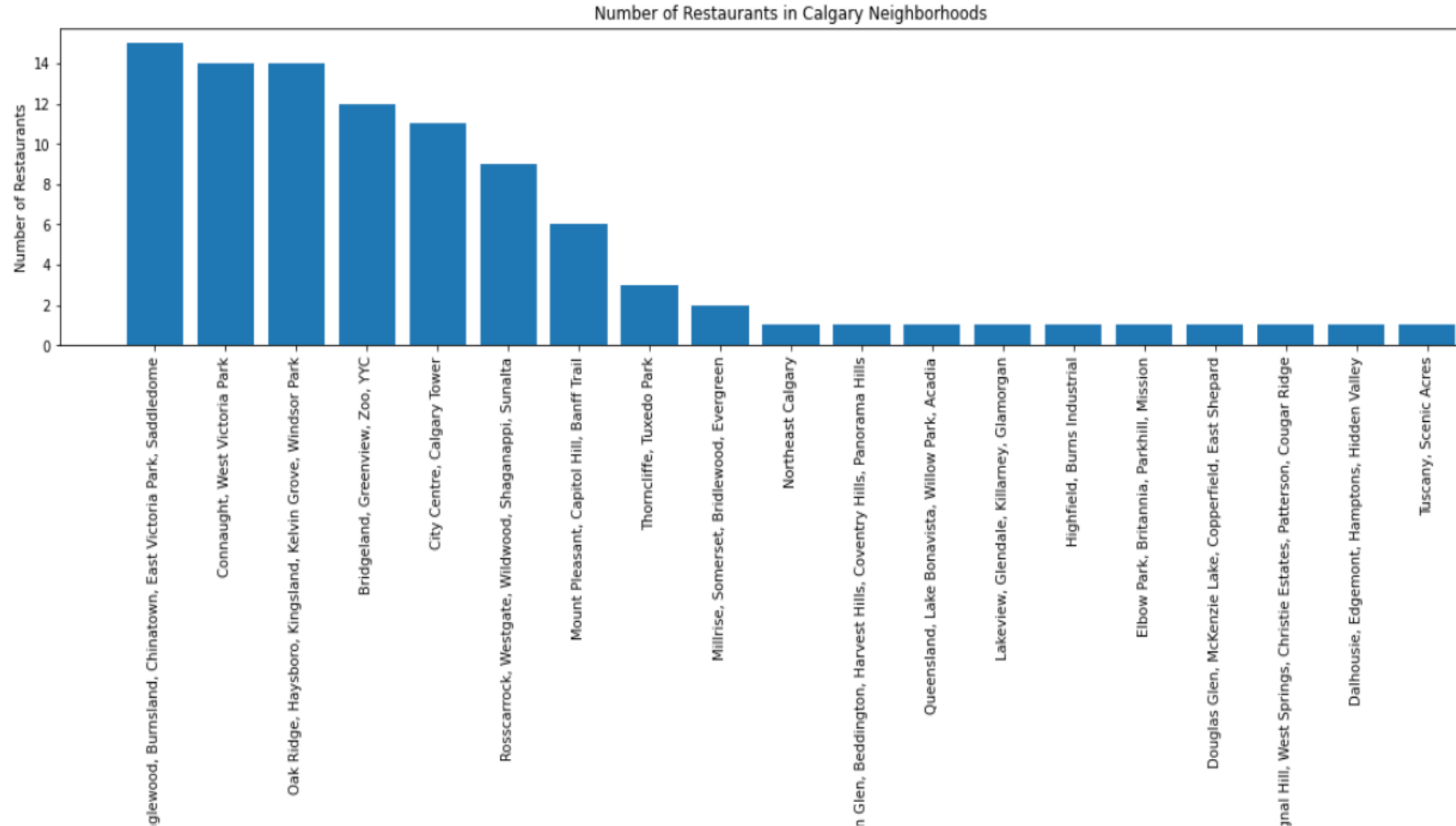
Methodology

- Step 2: The FourSquare API was called upon to retrieve information on venues local to each of these neighborhoods.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
1	Dalhousie, Edgemont, Hamptons, Hidden Valley	51.12606	-114.143158	Petro-Canada	51.128068	-114.138057	Gas Station
2	Dalhousie, Edgemont, Hamptons, Hidden Valley	51.12606	-114.143158	Edgemont City	51.126473	-114.138997	Asian Restaurant
3	Dalhousie, Edgemont, Hamptons, Hidden Valley	51.12606	-114.143158	Friends Cappuccino Bar & Bake Shop	51.126370	-114.138676	Café
4	Dalhousie, Edgemont, Hamptons, Hidden Valley	51.12606	-114.143158	Mac's	51.128309	-114.137902	Convenience Store
5	Forest Lawn, Dover, Erin Woods	51.03180	-113.978600	Bonasera Pizza And Sports Bar	51.029893	-113.982543	Bar

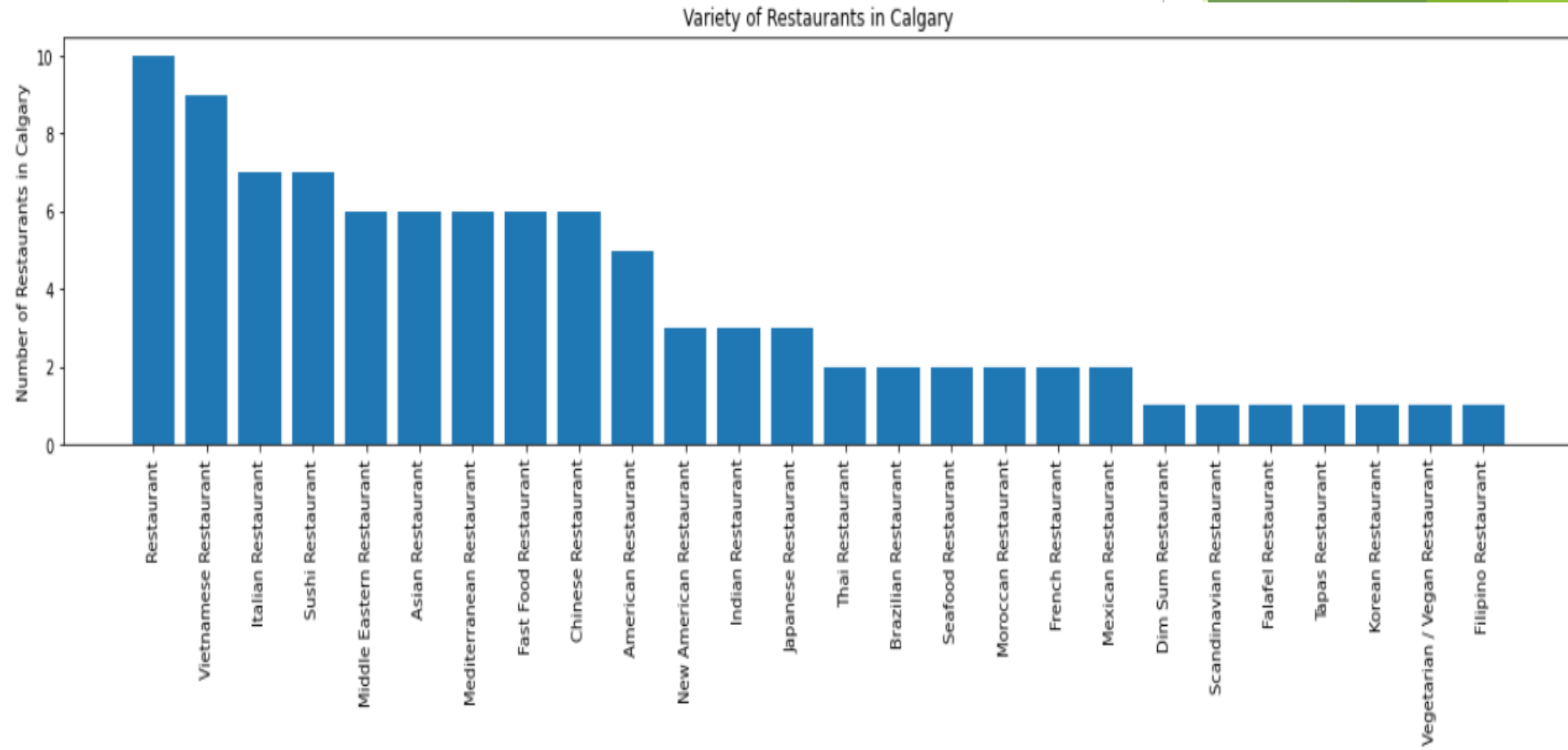
Methodology

- Step 3: This dataset was filtered out to only display restaurant data. The number of restaurants in each neighborhood was evaluated..



Methodology

- ▶ ..as well as the different types of cuisines offered at these restaurants.



Methodology

- Step 4: Venue types were sorted among neighborhoods in accordance with their occurrence frequency, to determine the popularity of each type of cuisine

```
----Bridgeland, Greenvview, Zoo, YYC----  
      venue  freq  
0    Vietnamese Restaurant 0.19  
1           Restaurant 0.11  
2         Sushi Restaurant 0.11  
3 Mediterranean Restaurant 0.11  
4         Chinese Restaurant 0.11
```

Methodology

- Step 4: Venue types were sorted among neighborhoods in accordance with their occurrence frequency, to determine the popularity of each type of cuisine

```
----Bridgeland, Greenview, Zoo, YYC----
      venue  freq
0    Vietnamese Restaurant 0.19
1           Restaurant 0.11
2         Sushi Restaurant 0.11
3 Mediterranean Restaurant 0.11
4         Chinese Restaurant 0.11
```

- Step 5: Cluster Machine Learning algorithms were applied to this dataset, to gain and understanding of the attributes that classify popular restaurants in the city.

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	6th Most Common Venue
Dalhousie, Edgemont, Hamptons, Hidden Valley	51.12606	-114.143158	2	Thai Restaurant	Asian Restaurant	Middle Eastern Restaurant	Indian Restaurant	Vietnamese Restaurant	Japanese Restaurant
Forest Lawn, Dover, Erin Woods	51.03180	-113.978600	3	Vietnamese Restaurant	Indian Restaurant	Mediterranean Restaurant	Japanese Restaurant	Asian Restaurant	Brazilian Restaurant
Montgomery, Bowness, Silver Springs, Greenwood	51.08090	-114.161600	1	Mexican Restaurant	Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Japanese Restaurant	Asian Restaurant
Lynnwood Ridge, Ogden, Foothills	50.98780	-114.000100	0	Vietnamese Restaurant	Asian Restaurant	Sushi Restaurant	Seafood Restaurant	Japanese Restaurant	Brazilian Restaurant

Methodology

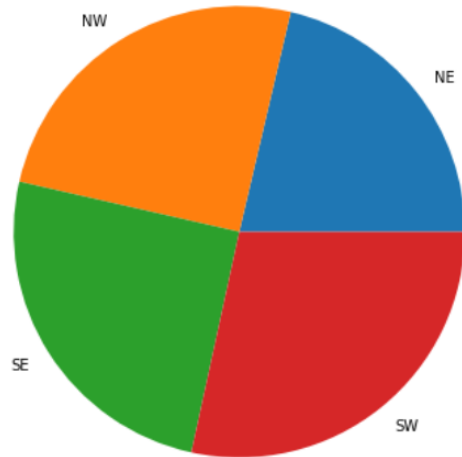
- Step 6: Calgary Census data was analyzed, to understand other factors that may affect the residents willingness to eat at a restaurant.

Community	Newsletter Name	Median Household Income	Median Age	Population 2014	Dwellings 2014	City Quadrant	Median Home Sale Price
Abbeydale	-	\$55,345	34	6071	2031	SE	\$305,000
Acadia	Acadia	\$46,089	42	10969	5067	SE	\$447,000
Albert Park / Radisson Heights	-	\$38,019	37	6529	2936	SE	\$349,900
Altadore	The Source	\$53,786	37	9518	4537	SW	\$925,000
Applewood Park	-	\$65,724	33	6864	2228	SE	\$380,000
...
Willow Park	Your Willow Park Maple Ridge	\$63,588	45	5440	2284	SE	\$705,000
Windsor Park	The Elbow Scene	\$39,425	37	4417	2453	SW	\$675,000
Winston Heights/Mountview	Community Matters	\$41,065	42	3845	1899	NE	\$625,000

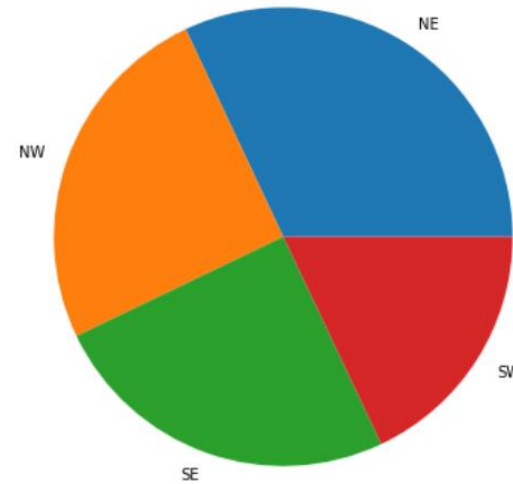
Methodology

- Step 7: The census data was sorted and visualized:

Median Household Income

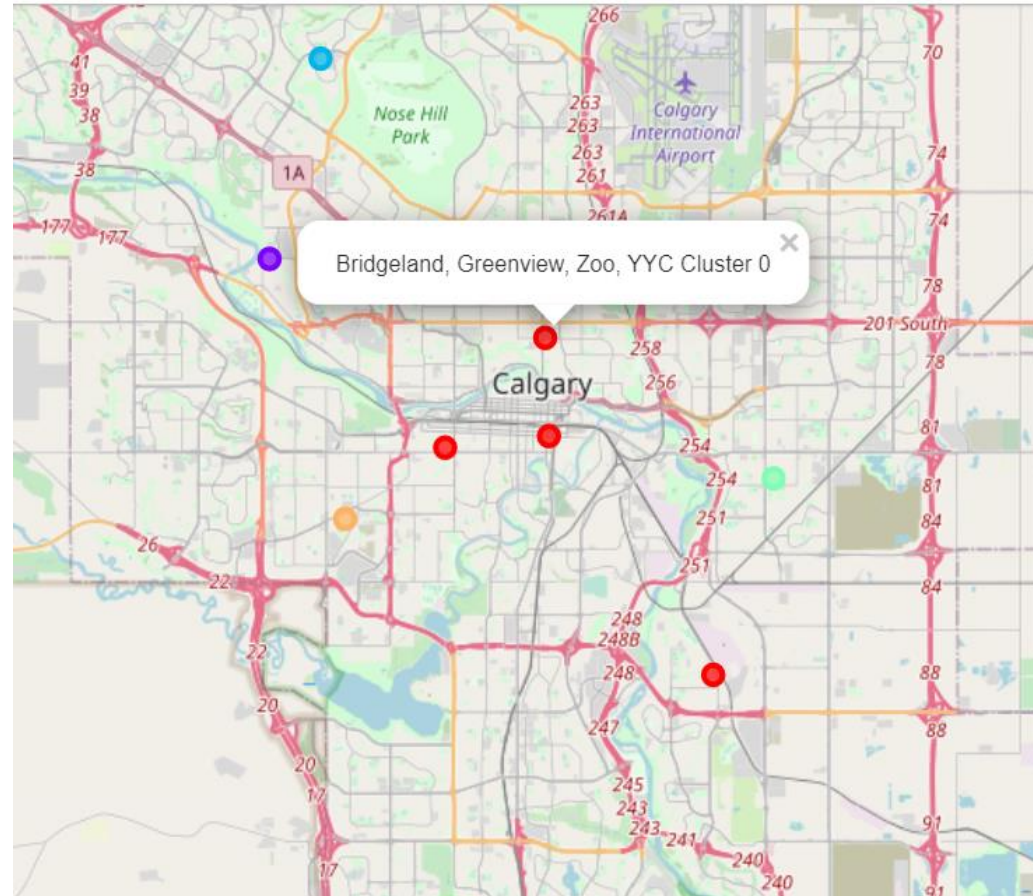


Resident Population



Results and Discussion

- With a visual of the Clustered venue data, it is clear that Category 0-type restaurants are prevalent in the city of Calgary. The visual also highlights that there are not many popular restaurant venues in the North East quadrant of the city.



- Cluster 0
- Cluster 1
- Cluster 2
- Cluster 3
- Cluster 4

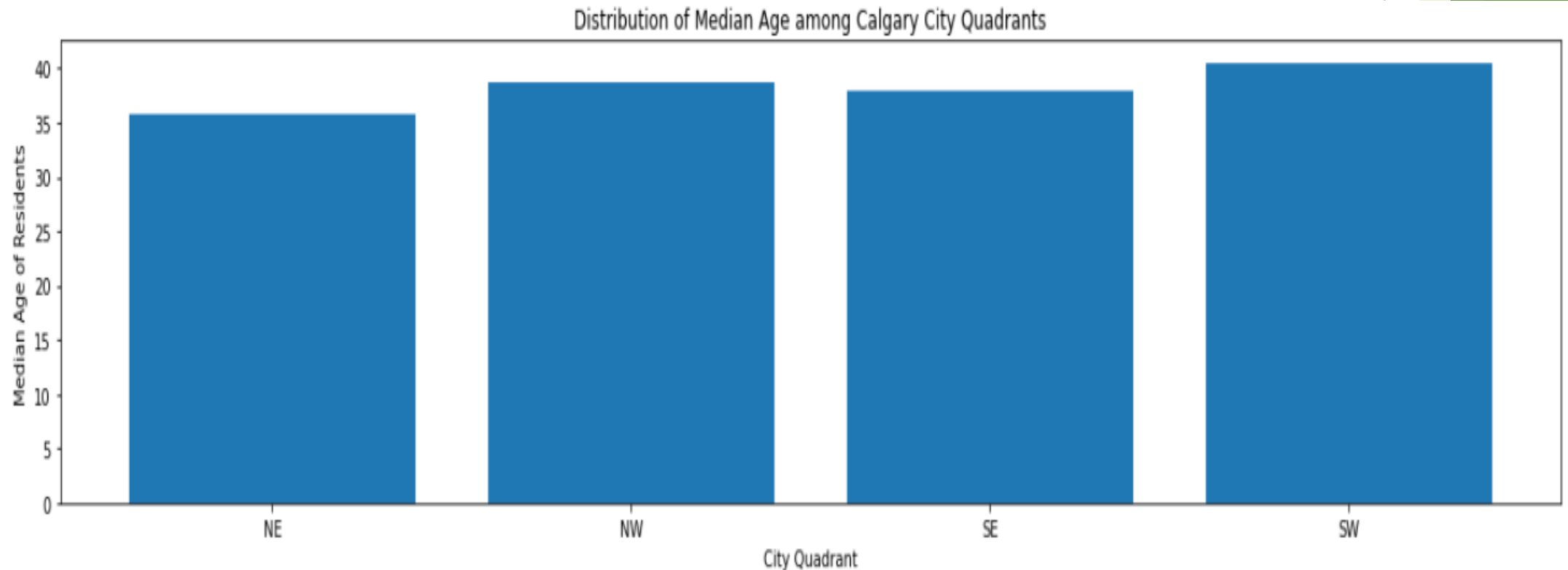
Results and Discussion

- Cluster 0 venues are thought to represent venues of Eastern Asian cuisine, including Vietnamese, Asian, Chinese cuisine, as well as sushi and seafood options. The table below displays common restaurant venues in Cluster 0 neighborhoods.

	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
19	Lynnwood Ridge, Ogden, Foothills Industrial, G...	Vietnamese Restaurant	Asian Restaurant	Sushi Restaurant	Seafood Restaurant	Japanese Restaurant	Brazilian Restaurant	Chinese Restaurant	Dim Sum Restaurant	Falafel Restaurant	Fast Food Restaurant
20	Rosscarrock, Westgate, Wildwood, Shaganappi, S...	Restaurant	Italian Restaurant	American Restaurant	Asian Restaurant	Fast Food Restaurant	New American Restaurant	Middle Eastern Restaurant	Vegetarian / Vegan Restaurant	Chinese Restaurant	Dim Sum Restaurant
28	Bridgeland, Greenview, Zoo, YYC	Vietnamese Restaurant	Chinese Restaurant	Sushi Restaurant	Mediterranean Restaurant	Restaurant	Moroccan Restaurant	Fast Food Restaurant	Italian Restaurant	American Restaurant	Japanese Restaurant
37	Inglewood, Burnside, Chinatown, East Victoria...	Sushi Restaurant	Restaurant	Mediterranean Restaurant	Asian Restaurant	Brazilian Restaurant	French Restaurant	Vietnamese Restaurant	Chinese Restaurant	Fast Food Restaurant	Filipino Restaurant

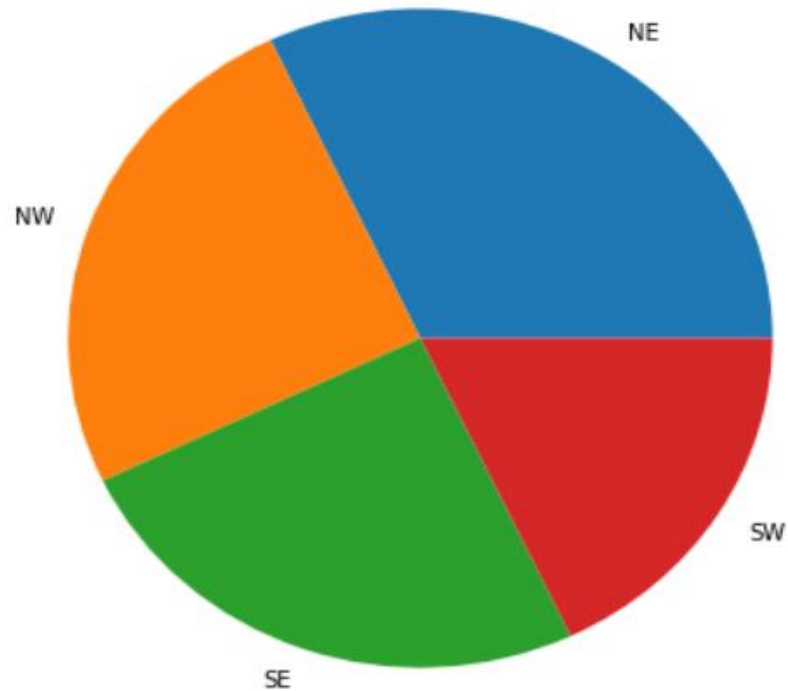
Results and Discussion

- It is not thought that Median Household Income or Median Age of Residents have a significant impact on this analysis, as there is not a major contrast in these variables among the different quadrants of the city.



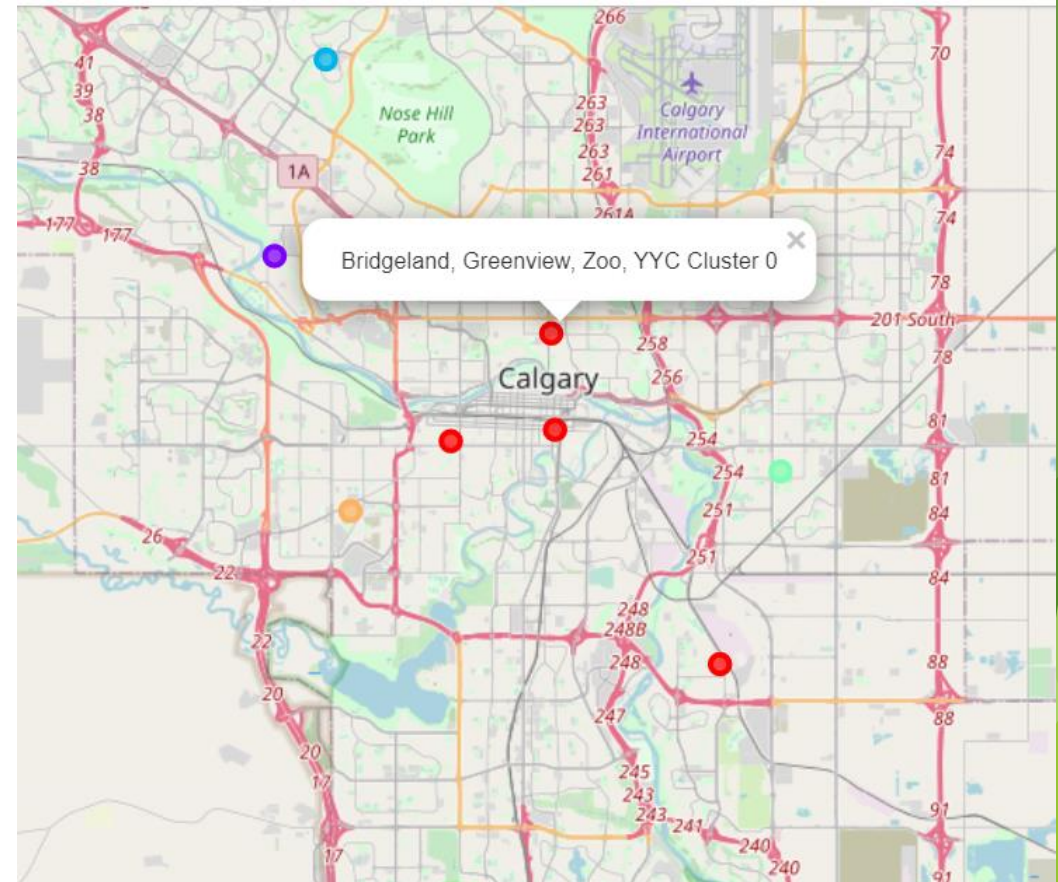
Results and Discussion

- The Census data demonstrated that there are measurably more residents in North Eastern Calgary neighborhoods than other neighborhoods



Results and Discussion

- It is also evident from map visuals produced that there is a lack of restaurants currently in the North Eastern neighborhoods:



Conclusion and Recommendation

- ▶ The Wikipedia Calgary neighborhood dataset, Calgary census dataset, and Foursquare API venue dataset provided a number of insights into Calgary's current population and appetite to eat out at restaurants.
- ▶ The existing restaurants and their locations were evaluated and compared, and the most popular restaurant venue types were discovered.
- ▶ Through visual data analysis and clustering machine learning techniques, the user concluded that there is a strong business opportunity to open a restaurant of East-Asian cuisine in any Northeastern Calgary neighborhood.

References

- ▶ [1]: Wikipedia: “*List of postal codes of Canda: T*”, Table 2: “*Alberta - 157 FSAs*”. Retrieved: Nov 20, 2020. URL: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_T
- ▶ [2]: Great News: “*Calgary Community Demographics*”. Retrieved Nov 23, 2020. URL: <https://great-news.ca/demographics/>