Clustering Culinary Neighborhoods in Vancouver, British Columbia

Warren Lee, March 2021

Introducing the Presenter

Pomona College Graduate

Claremont, California;
studied Economics and Mathematics

Economic Consultant

Supported Economics PhDs, performing
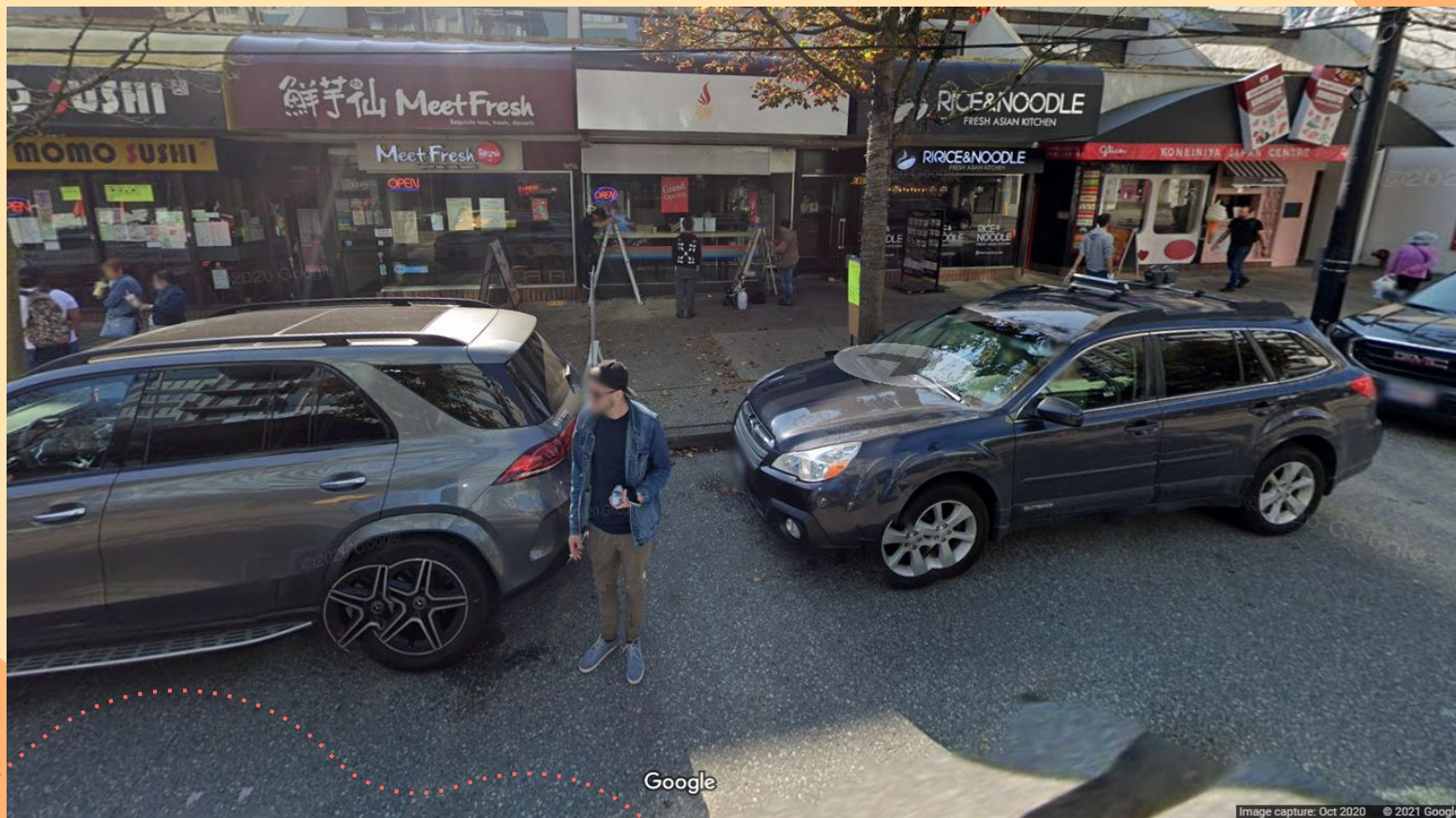economic/statistical analysis for lawyers

Vancouver, BC Native

Vancouver, BC:

- **Most populous city in the province**
- **One of the most ethnically diverse cities in Canada**
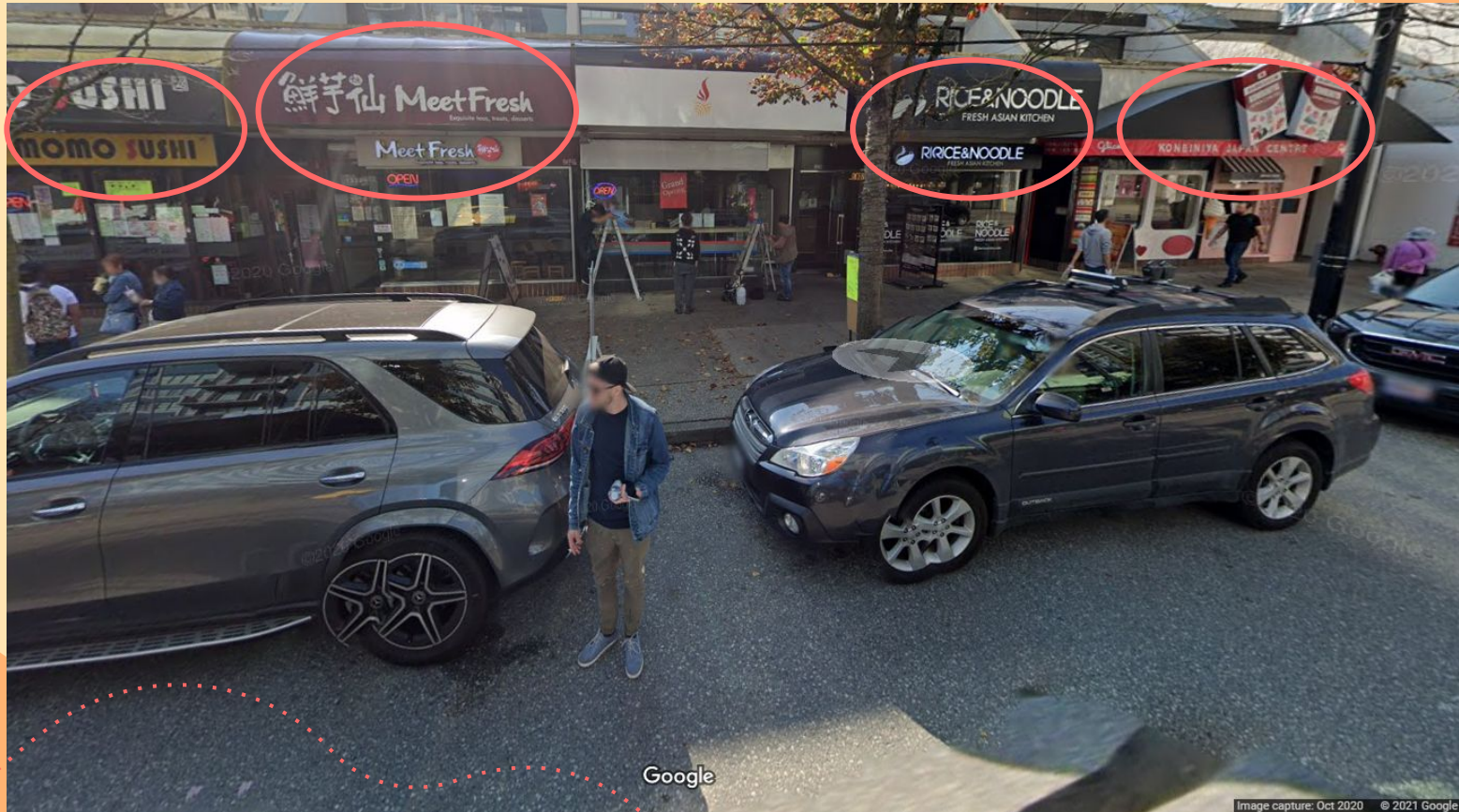- **Touted as "Canada's Dine-Out Capital"**

Street view of Robson St. from Google Maps

Sushi

Taiwanese Desserts

Thai

Japanese Crepes

Street view of Robson St. from Google Maps

Motivation

1. Are certain types of cuisines more represented in some areas than in others?
2. Can we group areas in Vancouver based on the most popular cuisine?

Motivation

1. Are certain types of cuisines more represented in some areas than in others?
2. Can we group areas in Vancouver based on the most popular cuisine?
   - ➤ Identify areas of dining when traveling
   - ➤ Mark areas of entry for restauranteurs of a particular cuisine

Data Collection

- Collected data from Yelp on ~4,000 restaurants located within Vancouver's local area boundary
- Features analyzed:
  - Geographic location
  - Cuisine type from Yelp-defined restaurant tags
  - Star Rating

Sample data:

Name	Address	Latitude	Longitude	Cuisines	Stars
Dinesty Dumpling House	1719 Robson Street	49.290818	-123.134391	Chinese; Dumplings; Bubble Tea	4.00

Identifying Culinary Neighborhoods

- Municipal neighborhoods fail to capture culinary neighborhoods

Identifying Culinary Neighborhoods

- Municipal neighborhoods fail to capture culinary neighborhoods
  - Does not capture clusters occurring at boundaries

Identifying Culinary Neighborhoods

- Municipal neighborhoods fail to capture culinary neighborhoods
  - Does not capture clusters occurring at boundaries
  - Does not identify smaller groups of businesses within areas that expand large distances
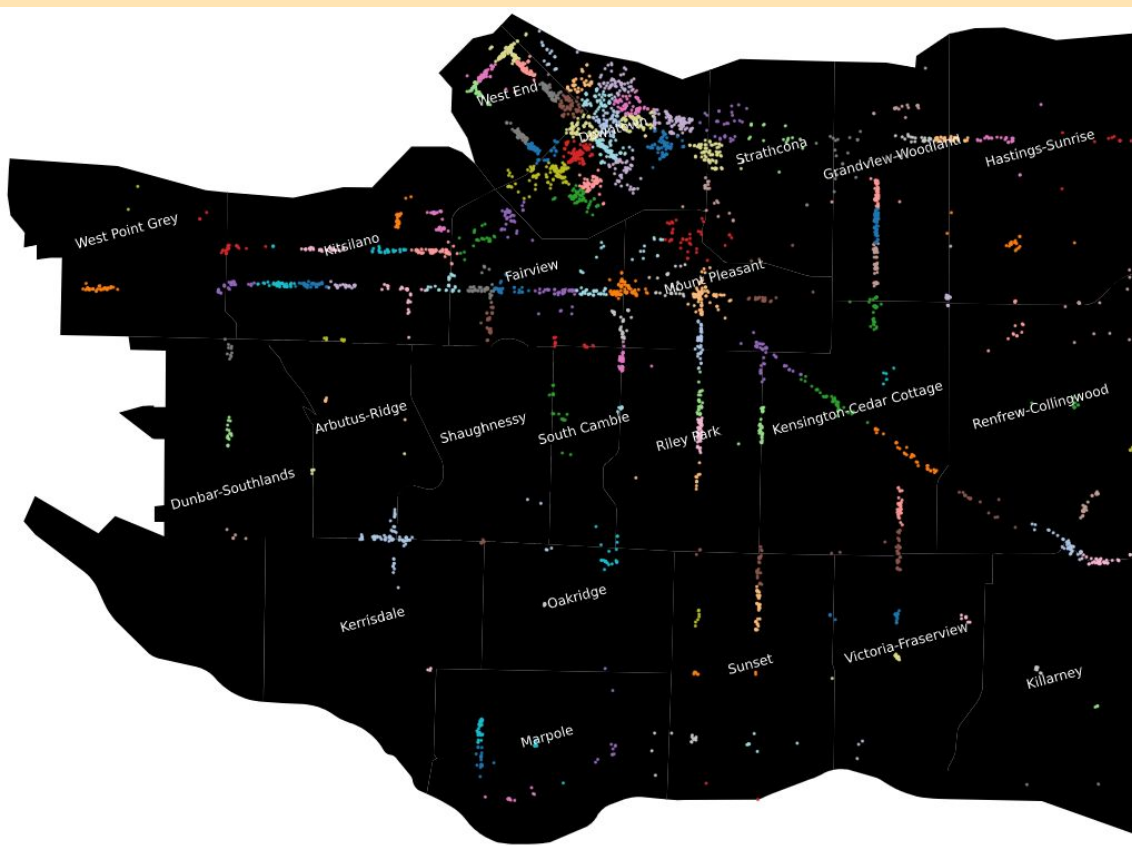
Identifying Culinary Neighborhoods

- Municipal neighborhoods fail to capture culinary neighborhoods
  - Does not capture clusters occurring at boundaries
  - Does not identify smaller groups of businesses within areas that expand large distances
- Applied **KMeans Clustering** to break up Data into ~120 Culinary Neighborhoods
  - Cluster restaurants based on **location**
  - **Shape of clusters** simulates areas that visitors might walk around

Data Modeling Process

1

2

3

Identify Counts of Cuisine Types in Culinary Neighborhoods

Model Cuisine Groupings Based on Occurence of Cuisine Types

Cluster Culinary Neighborhoods based on Grouping Weights

Neighbor -hood #	Cuisine Type			
	Sushi	Cafe	Korean	Pizza
1	16	0	20	0
2	17	0	0	0
3	0	4	0	6
...	...	...	...	...

Data Modeling Process

1

2

3

Identify Counts of Cuisine Types in Culinary Neighborhoods

Model Cuisine Groupings Based on Occurence of Cuisine Types

Cluster Culinary Neighborhoods based on Grouping Weights

Neighbor -hood #	Cuisine Type			
	Sushi	Cafe	Korean	Pizza
1	16	0	20	0
2	17	0	0	0
3	0	4	0	6
...	...	...	...	...

Cuisine Type	Cuisine Distribution		
	Group 1	Group 2	...
Sushi	6.3	0.2	...
Cafe	0.7	4.1	...
Korean	2.5	0.1	...
Pizza	0.3	3.2	...
...	...	...	...

NMF
Topic Modeling

Data Modeling Process

1

Identify Counts of Cuisine Types in Culinary Neighborhoods

	Cuisine Type			
Neighbor-hood #	Sushi	Cafe	Korean	Pizza
1	16	0	20	0
2	17	0	0	0
3	0	4	0	6
...	...	...	...	...

2

Model Cuisine Groupings Based on Occurence of Cuisine Types

Cuisine Type	Cuisine Distribution		
	Group1	Group 2	...
Sushi	6.3	0.2	...
Cafe	0.7	4.1	...
Korean	2.5	0.1	...
Pizza	0.3	3.2	...
...	...	...	...

3

Cluster Culinary Neighborhoods based on Grouping Weights

NMF
Topic Modeling

Data Modeling Process

1 — **Identify Counts of Cuisine Types in Culinary Neighborhoods**

2 — **Model Cuisine Groupings Based on Occurence of Cuisine Types**

3 — **Cluster Culinary Neighborhoods based on Grouping Weights**

Neighbor-hood #	Cuisine Type			
	Sushi	Cafe	Korean	Pizza
1	16	0	20	0
2	17	0	0	0
3	0	4	0	6
...	...	...	...	...

Cuisine Type	Cuisine Distribution		
	Group1	Group 2	...
Sushi	6.3	0.2	...
Cafe	0.7	4.1	...
Korean	2.5	0.1	...
Pizza	0.3	3.2	...
...	...	...	...

NMF
Topic Modeling

Data Modeling Process

1

Identify Counts of Cuisine Types in Culinary Neighborhoods

	Cuisine Type			
Neighbor-hood #	Sushi	Cafe	Korean	Pizza
1	16	0	20	0
2	17	0	0	0
3	0	4	0	6
...	...	...	...	...

2

Model Cuisine Groupings Based on Occurence of Cuisine Types

Cuisine Type	Cuisine Distribution		
	Group 1	Group 2	...
Sushi	6.3	0.2	...
Cafe	0.7	4.1	...
Korean	2.5	0.1	...
Pizza	0.3	3.2	...
...	...	...	...

3

Cluster Culinary Neighborhoods based on Grouping Weights

Neighbor-hood #	Group Weights		
	Group 1	Group 2	...
1	2.8	0.1	...
2	2.3	0.5	...
3	0.4	1.8	...
...	...	...	...

NMF
Topic Modeling

Data Modeling Process

1 Identify Counts of Cuisine Types in Culinary Neighborhoods

2 Model Cuisine Groupings Based on Occurence of Cuisine Types

3 Cluster Culinary Neighborhoods based on Grouping Weights

Neighbor-hood #	Cuisine Type			
	Sushi	Cafe	Korean	Pizza
1	16	0	20	0
2	17	0	0	0
3	0	4	0	6
...	...	...	...	...

Cuisine Type	Cuisine Distribution		
	Group 1	Group 2	...
Sushi	6.3	0.2	...
Cafe	0.7	4.1	...
Korean	2.5	0.1	...
Pizza	0.3	3.2	...
...	...	...	...

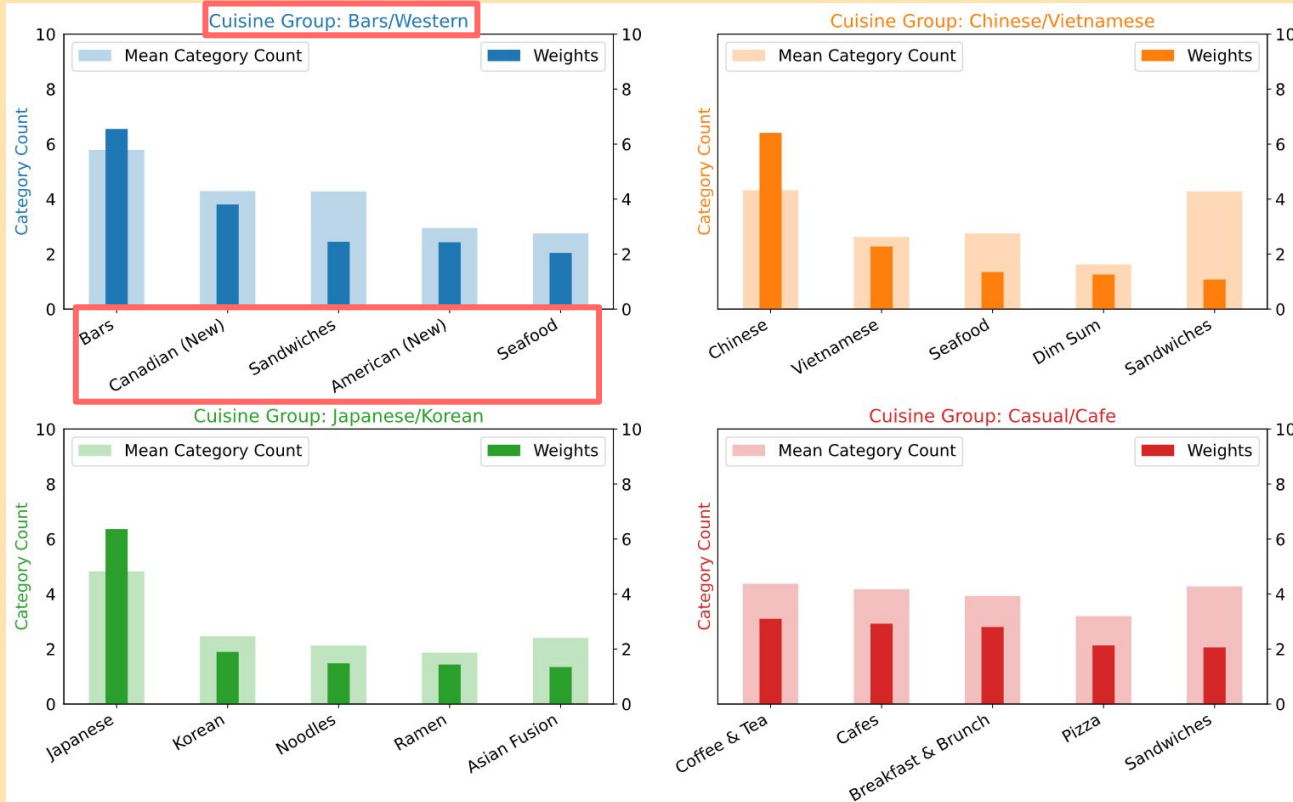| Neighbor-hood #	Group Weights		
	Group 1	Group 2	...
1	2.8	0.1	...
2	2.3	0.5	...
3	0.4	1.8	...
...	...	...	...

Cluster of Cuisine Group

NMF Topic Modeling

KMeans Clustering

Modeling reveals four distinct ethnic cuisine groups
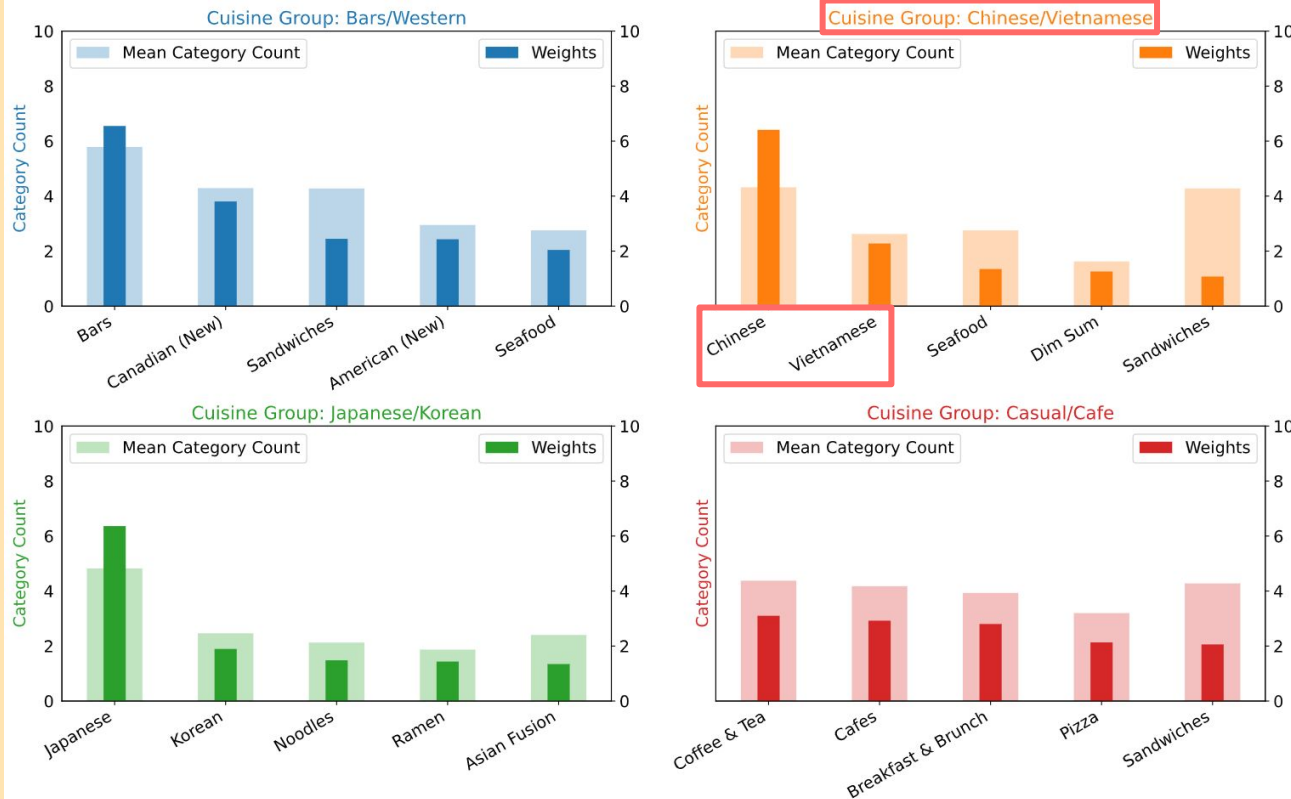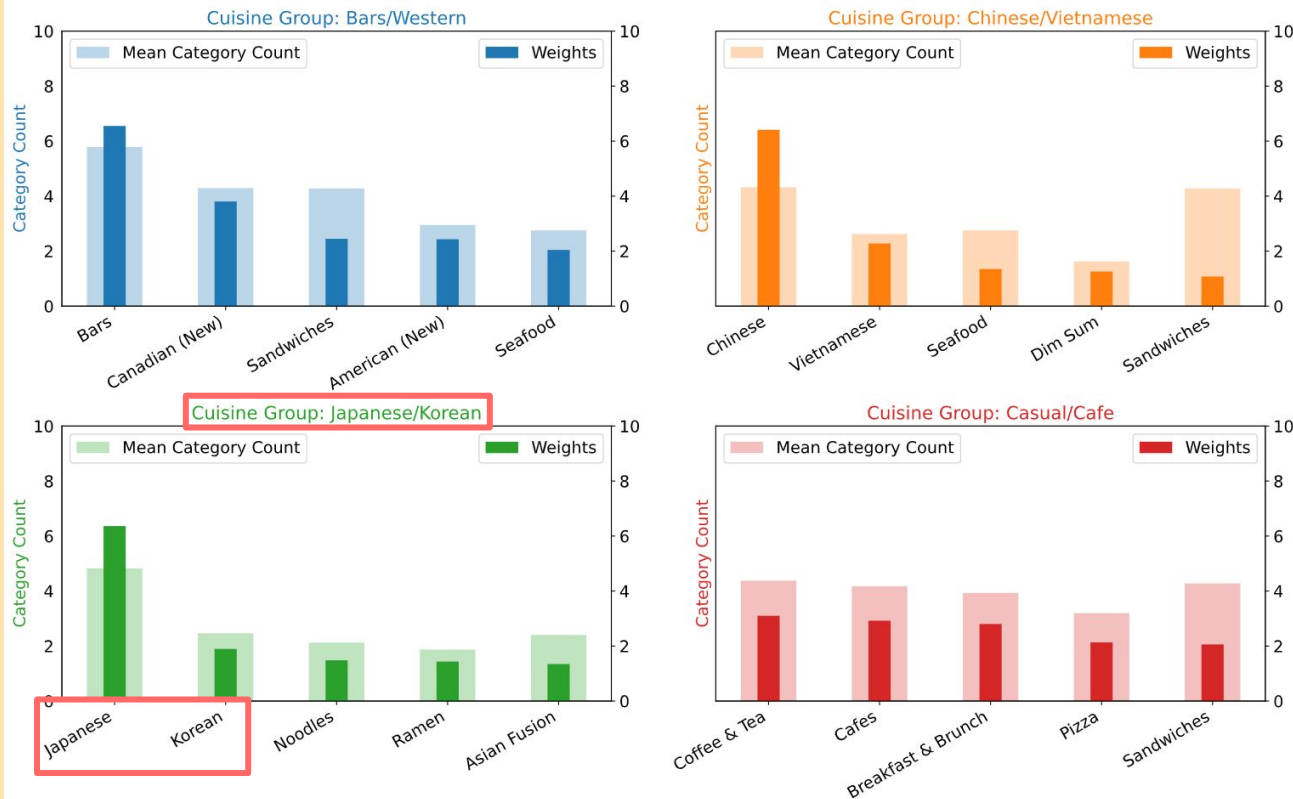## Three cuisine groups weighted heavily towards one cuisine

Cuisine Group: Bars/Western
- Mean Category Count
- Weights
- Categories: Bars, Canadian (New), Sandwiches, American (New), Seafood

Cuisine Group: Chinese/Vietnamese
- Mean Category Count
- Weights
- Categories: Chinese, Vietnamese, Seafood, Dim Sum, Sandwiches

Cuisine Group: Japanese/Korean
- Mean Category Count
- Weights
- Categories: Japanese, Korean, Noodles, Ramen, Asian Fusion

Cuisine Group: Casual/Cafe
- Mean Category Count
- Weights
- Categories: Coffee & Tea, Cafes, Breakfast & Brunch, Pizza, Sandwiches
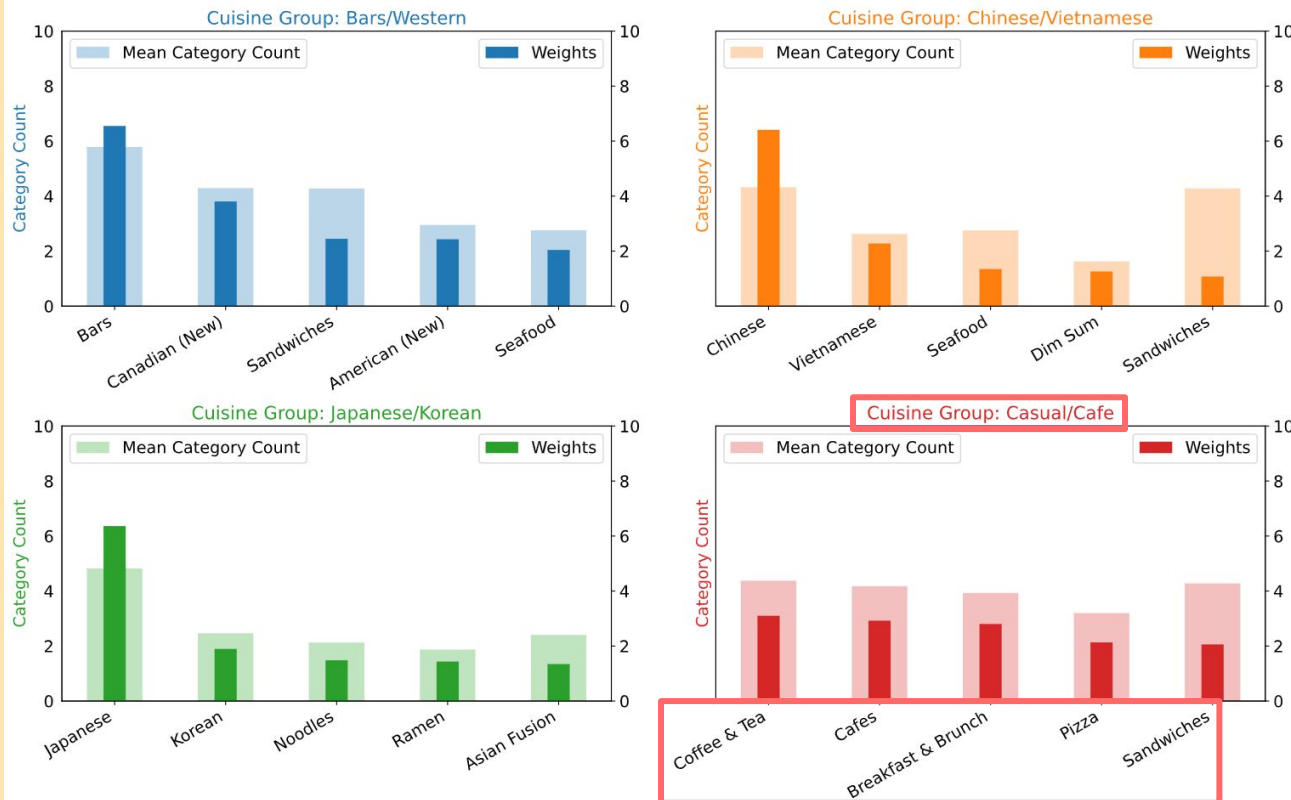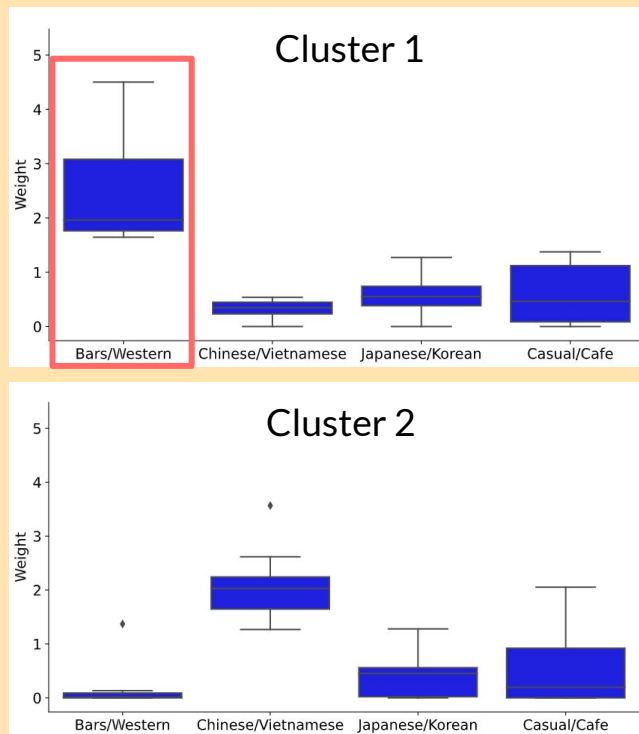
Modeling reveals four distinct ethnic cuisine groups
Three cuisine groups weighted heavily towards one cuisine
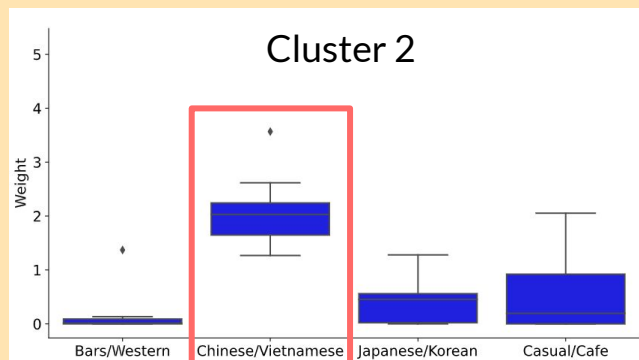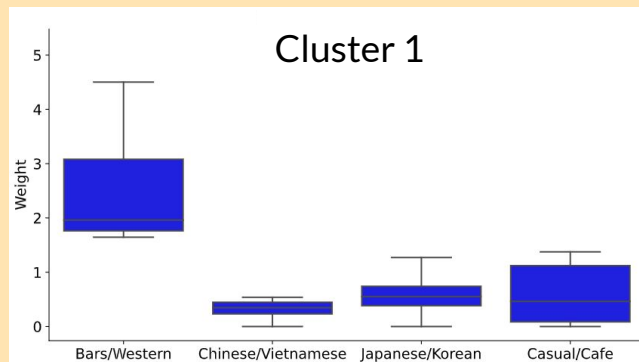
Modeling reveals four distinct ethnic cuisine groups
Three cuisine groups weighted heavily towards one cuisine

Modeling reveals four distinct ethnic cuisine groups
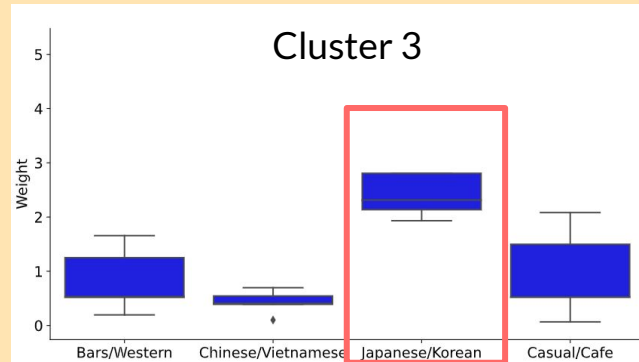## Three cuisine groups weighted heavily towards one cuisine

Cuisine Group: Bars/Western
- Mean Category Count
- Weights
- Bars, Canadian (New), Sandwiches, American (New), Seafood

Cuisine Group: Chinese/Vietnamese
- Mean Category Count
- Weights
- Chinese, Vietnamese, Seafood, Dim Sum, Sandwiches

Cuisine Group: Japanese/Korean
- Mean Category Count
- Weights
- Japanese, Korean, Noodles, Ramen, Asian Fusion

Cuisine Group: Casual/Cafe
- Mean Category Count
- Weights
- Coffee & Tea, Cafes, Breakfast & Brunch, Pizza, Sandwiches

Modeling reveals four distinct ethnic cuisine groups
Three cuisine groups weighted heavily towards one cuisine

Cuisine Group: Bars/Western
- Mean Category Count
- Weights
- Category Count
- Bars, Canadian (New), Sandwiches, American (New), Seafood

Cuisine Group: Chinese/Vietnamese
- Mean Category Count
- Weights
- Category Count
- Chinese, Vietnamese, Seafood, Dim Sum, Sandwiches

Cuisine Group: Japanese/Korean
- Mean Category Count
- Weights
- Category Count
- Japanese, Korean, Noodles, Ramen, Asian Fusion

Cuisine Group: Casual/Cafe
- Mean Category Count
- Weights
- Category Count
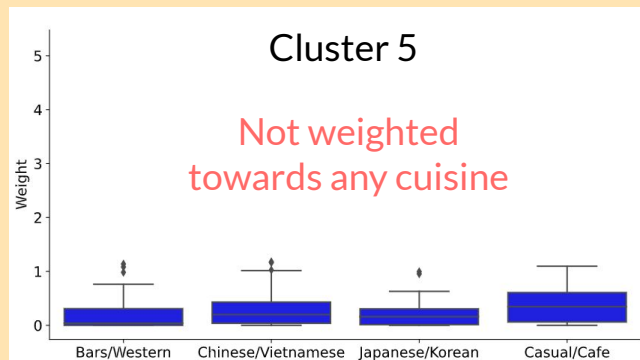- Coffee & Tea, Cafes, Breakfast & Brunch, Pizza, Sandwiches

k=5 clusters show that restaurant areas can be organized by cuisine groups

k=5 clusters show that restaurant areas can be organized by cuisine groups

k=5 clusters show that restaurant areas can be organized by cuisine groups

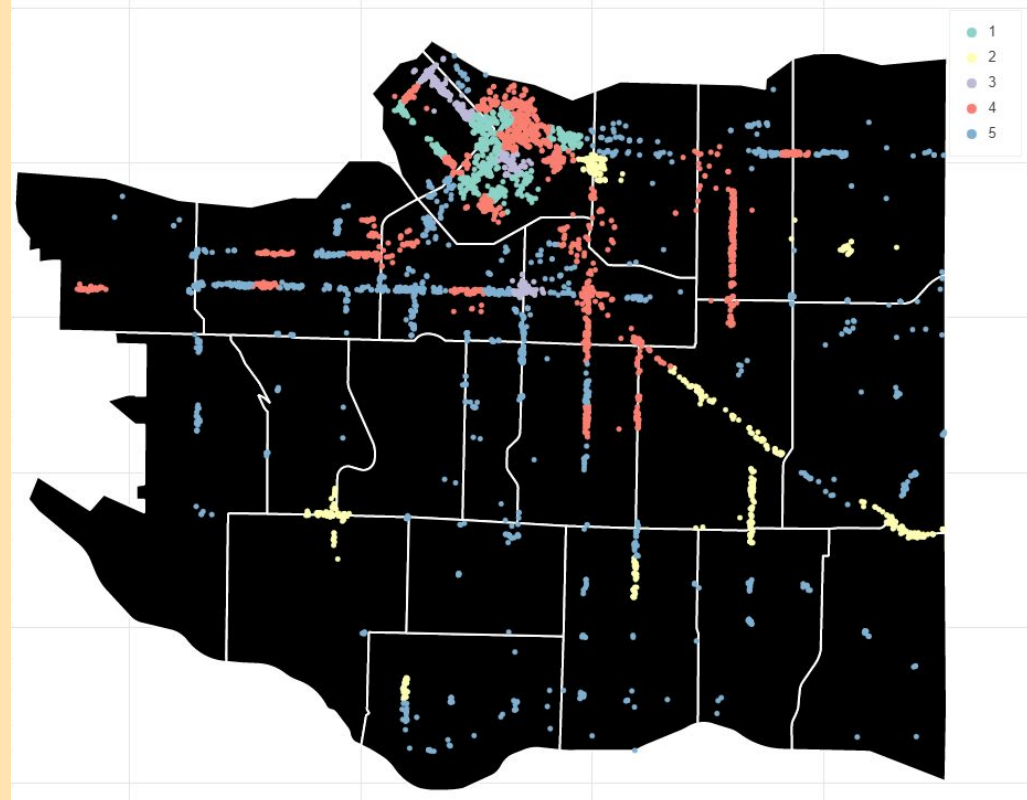# k=5 clusters show that restaurant areas can be organized by cuisine groups

k=5 clusters show that restaurant areas can be organized by cuisine groups

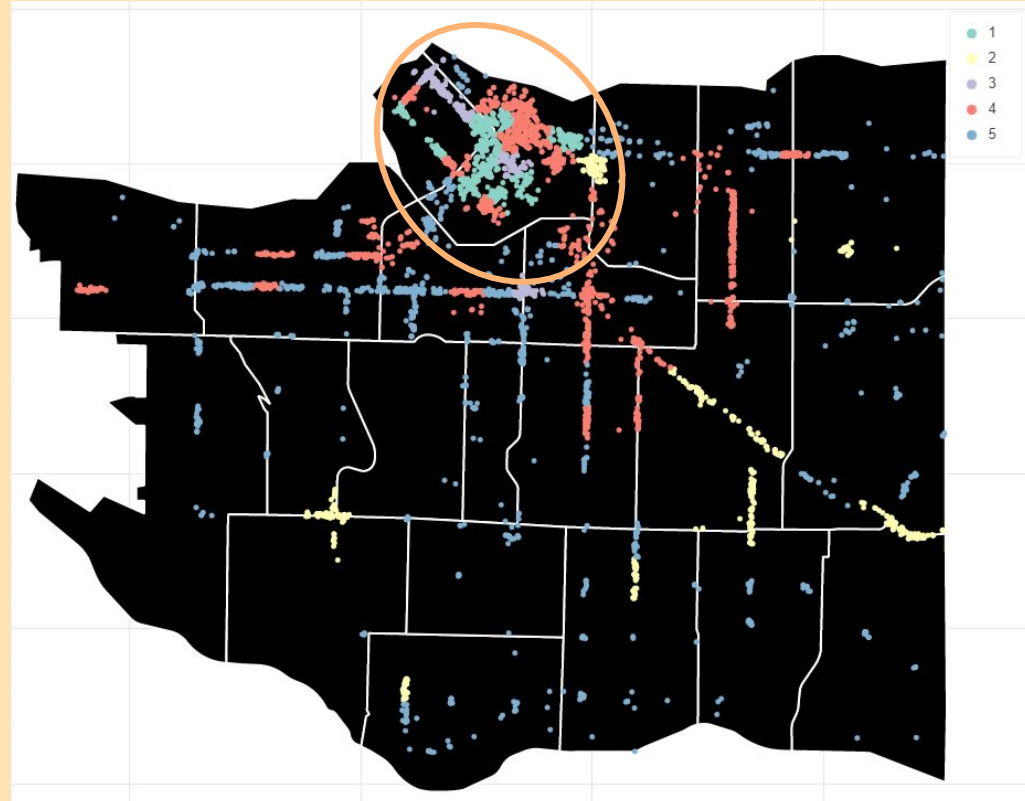# Clustering of Restaurant Areas Reveal Locations Saturated by Cuisine Group

- Legend:
  - Cluster 1: Weighted towards Bars/Western
  - Cluster 2: Weighted towards Chinese/Vietnamese
  - Cluster 3: Weighted towards Japanese/Korean
  - Cluster 4: Weighted towards Casual/Cafes
  - Cluster 5: Not weighted towards any cuisine

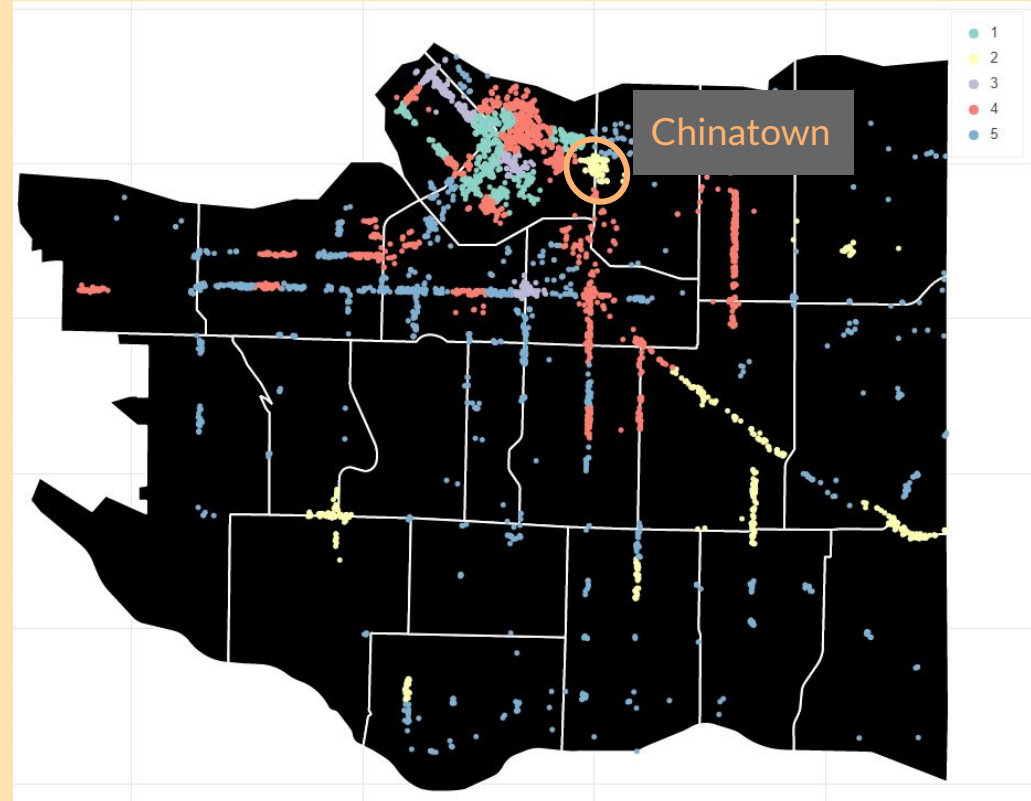# Clustering of Restaurant Areas Reveal Locations Saturated by Cuisine Group

- Legend:
  - Cluster 1: Weighted towards Bars/Western
  - Cluster 2: Weighted towards Chinese/Vietnamese
  - Cluster 3: Weighted towards Japanese/Korean
  - Cluster 4: Weighted towards Casual/Cafes
  - Cluster 5: Not weighted towards any cuisine

Clustering of Restaurant Areas Reveal Locations Saturated by Cuisine Group
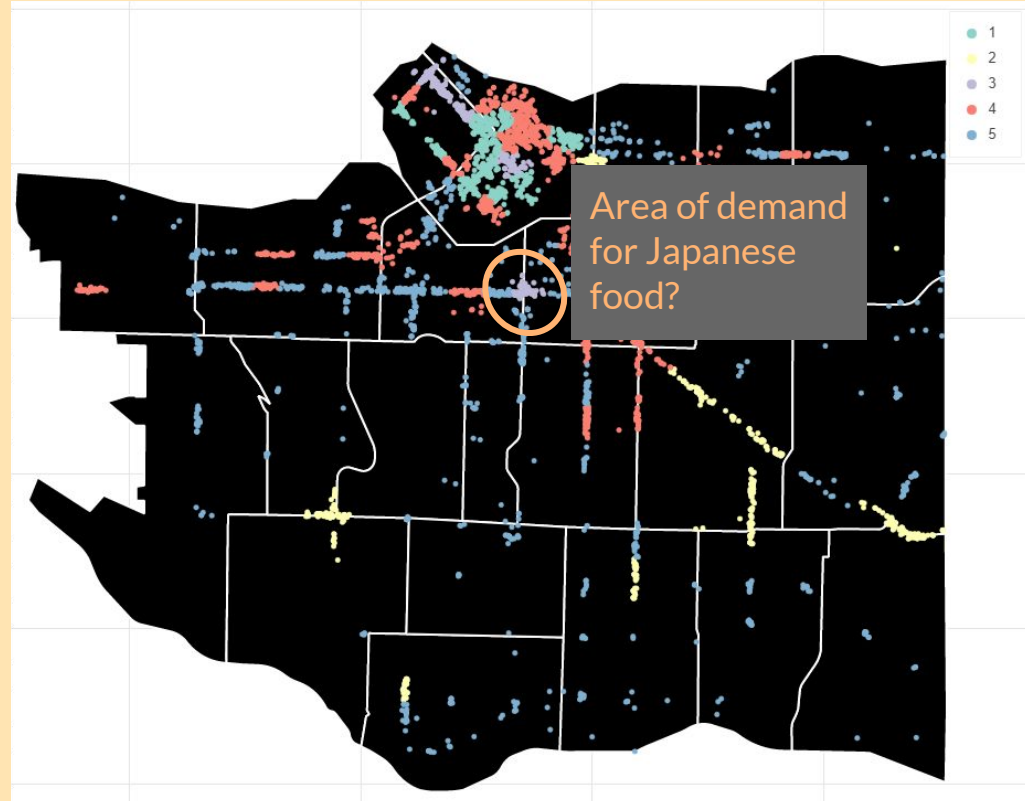
- Legend:
  - Cluster 1: Weighted towards Bars/Western
  - Cluster 2: Weighted towards Chinese/Vietnamese
  - Cluster 3: Weighted towards Japanese/Korean
  - Cluster 4: Weighted towards Casual/Cafes
  - Cluster 5: Not weighted towards any cuisine

Chinatown

Clustering of Restaurant Areas Reveal Locations Saturated by Cuisine Group
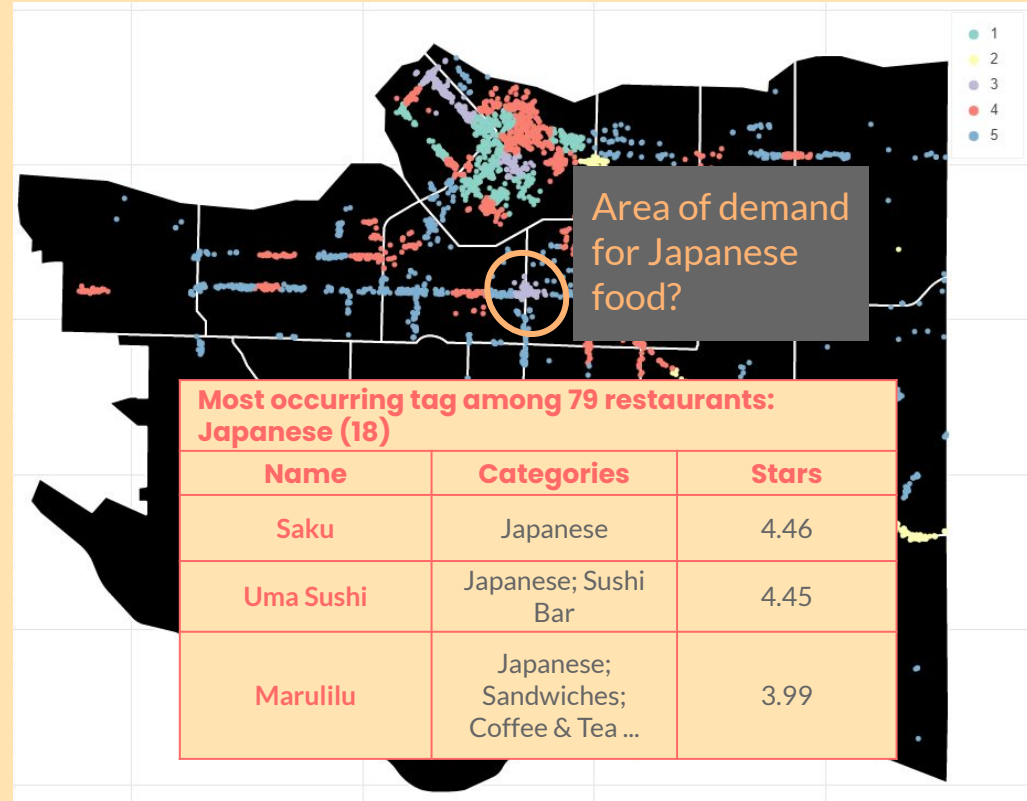
- Legend:
  - **Cluster 1:** Weighted towards Bars/Western
  - **Cluster 2:** Weighted towards Chinese/Vietnamese
  - **Cluster 3:** Weighted towards Japanese/Korean
  - **Cluster 4:** Weighted towards Casual/Cafes
  - **Cluster 5:** Not weighted towards any cuisine

Clustering of Restaurant Areas Reveal Locations Saturated by Cuisine Group

- Legend:
  - Cluster 1: Weighted towards Bars/Western
  - Cluster 2: Weighted towards Chinese/Vietnamese
  - Cluster 3: Weighted towards Japanese/Korean
  - Cluster 4: Weighted towards Casual/Cafes
  - Cluster 5: Not weighted towards any cuisine

Area of demand for Japanese food?

Clustering of Restaurant Areas Reveal Locations Saturated by Cuisine Group

- Legend:
  - Cluster 1: Weighted towards Bars/Western
  - Cluster 2: Weighted towards Chinese/Vietnamese
  - Cluster 3: Weighted towards Japanese/Korean
  - Cluster 4: Weighted towards Casual/Cafes
  - Cluster 5: Not weighted towards any cuisine

Area of demand for Japanese food?

Most occurring tag among 79 restaurants: Japanese (18)

Name	Categories	Stars
Saku	Japanese	4.46
Uma Sushi	Japanese; Sushi Bar	4.45
Marulilu	Japanese; Sandwiches; Coffee & Tea ...	3.99

Conclusion

1. Modeling culinary neighborhoods in Vancouver into four cuisine groups skew towards three cuisines: **Bars**, **Chinese**, and **Japanese**

2. Clustering neighborhoods by cuisine group show areas where **certain ethnic cuisines dominate**

3. Analysis provides a basis for **recommending similar cuisines in multiple areas**

Thanks!

Does anyone have any questions?