

The background features abstract, overlapping green geometric shapes in various shades of green, creating a modern and dynamic visual effect. These shapes are primarily located on the left and right sides of the slide, framing the central text.

Toronto Restaurant

Capstone Project

Business Problem

- ▶ The client would like to open a new restaurant.
- ▶ This restaurant would be situated in the financial district of Toronto.
- ▶ They are specifically targeting employees' lunches near the "First Canadian Place" building, which is located in the heart of the financial district. They believe there is room for another restaurant / take-away.
 - ▶ What is already available?
 - ▶ What restaurant is recommended?
 - ▶ Should it be low-end or high-end?

Data gathering

Sources

- ▶ Foursquare API
- ▶ Self generated excel sheet with restaurant prices

Data for the following restaurant cuisines were collected

- ▶ Thai
- ▶ Steak
- ▶ Italian
- ▶ Sandwich
- ▶ Sushi
- ▶ French
- ▶ Mexican

Retrieve all restaurant cuisines within 1.5km from target market

Combine the data and clean it

Collect the average price per restaurant

Combine the 2 tables (restaurants with average price)

Analyze the final table

Methodology

Collect all cuisines data

Collect average price for data

Combine the data

Show range in prices

Show frequency of each type

Show a cluster of each restaurant type

Predict possible prices for cuisines to estimate business possibilities



Data issues and cleaning

Searches were made regarding each cuisine type how this did not necessarily only bring back food type places.

The 11 rows of data consisted of the following erroneous data

- ❑ Embassy / Consulate (e.g. Italian Consulate)
- ❑ Residential Building (Apartment / Condo)
- ❑ Neighborhood
- ❑ Gym / Fitness Center
- ❑ Clothing Store
- ❑ Government Building

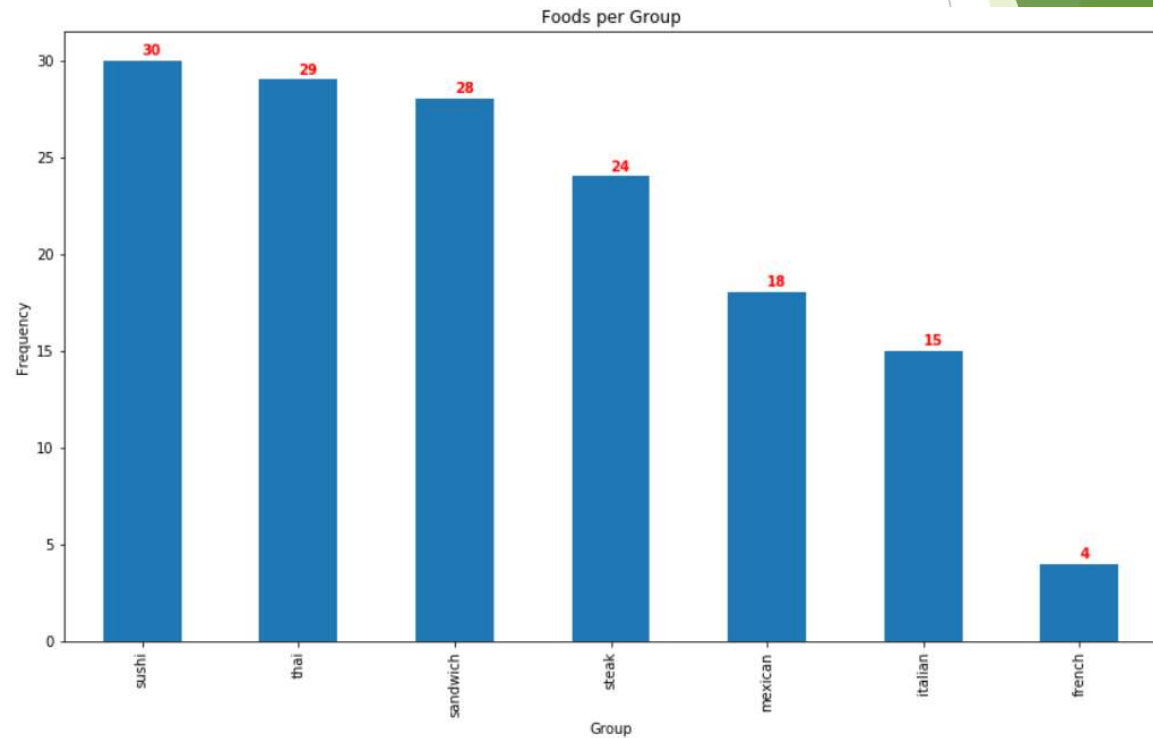
These were removed from the dataframe

Results

159 rows were found within 1.5km
After data cleaning, 148 rows were left

Frequency of each cuisine

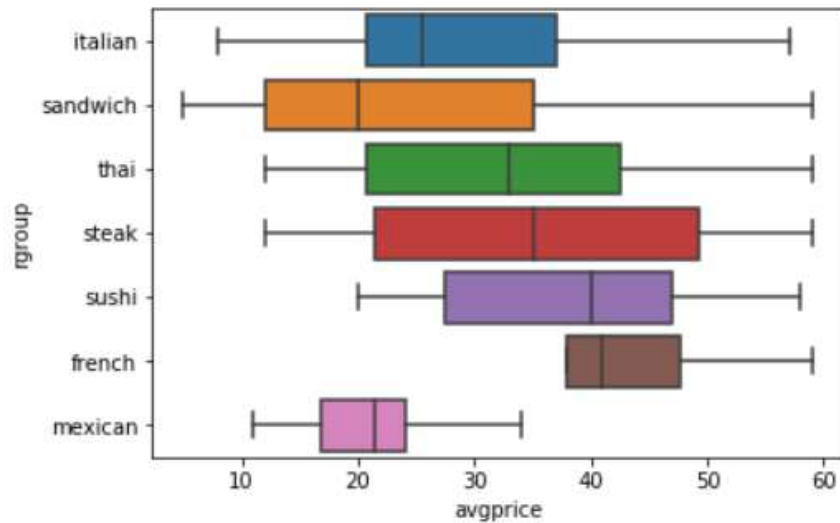
- ❑ Sushi(>30) exceeded the limit
- ❑ Thai(29) was at the limit
- ❑ French(4) was the lowest



Results

Box Plot Analysis of Average Prices

- ▶ Range of each cuisine
 - ▶ Italian had the widest range with very cheap and very expensive
 - ▶ Sandwiches had the bulk in the cheap area
 - ▶ French ranged only in expensive ranges

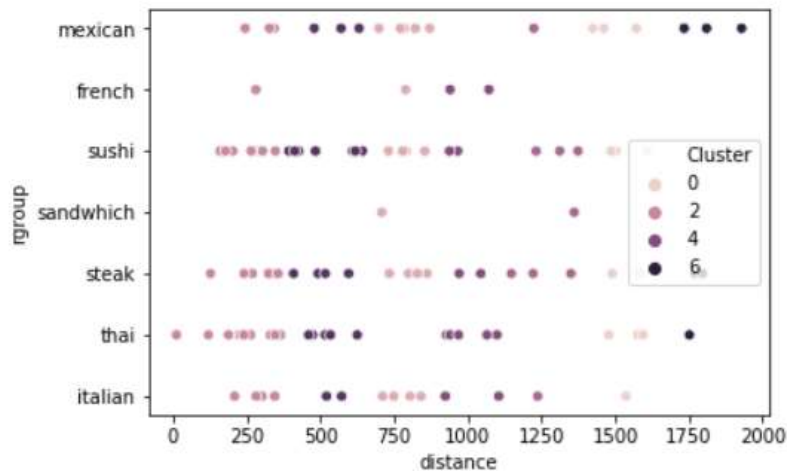


Results

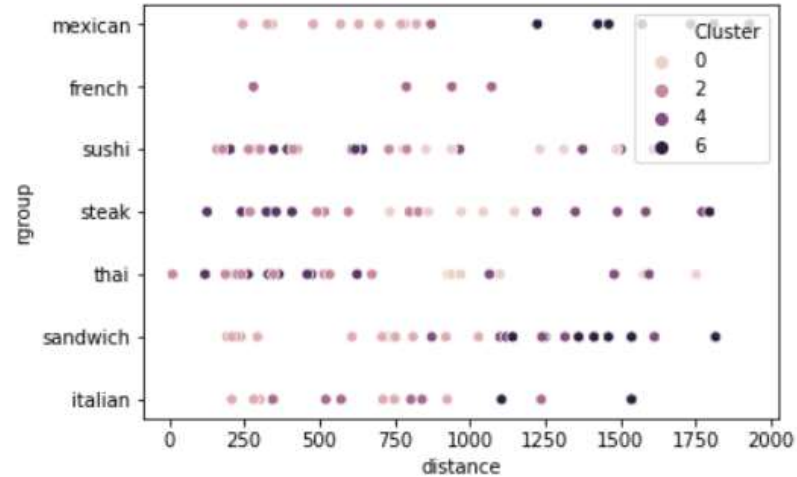
- ▶ Using a clustering analysis in the K-means analysis the data was first normalized to get an appropriate view.
- ▶ Without normalization the distance field overwhelmed all other statistics

Pre Normalization

(Clusters very bound around distance field only)



With Normalization



Conclusion

Sushi is the most saturated with more than 30 restaurants therefore one can either

- 1) go with Thai (2nd highest) with an average price of \$33 or
- 2) go for the least saturated market (French) with only 4 restaurants, average price of \$ 45

French might be better as there's less competition and a higher price attached

	rgroup	avgprice
0	french	44.750000
1	italian	29.750000
2	mexican	21.500000
3	sandwich	24.862069
4	steak	35.041667
5	sushi	38.111111
6	thai	32.785714

	Cluster	avgprice
0	0	24.882353
1	1	19.607143
2	2	47.592593
3	3	43.538462
4	4	48.555556
5	5	22.652174
6	6	16.650000