Best neighbourhood for an Indian Restaurant

Capstone Project

Points to consider

- Neighbourhood's that reflect a dining preference for Indian cuisine

- Identifying all the "Food" categories for Indian Restaurant's

- Household Incomes

- A preference for a centrally located neighbourhood within the city

Data needed for Business Problem

- Postcodes, latitude and longitude data scrapped from https://www.geonames.org/postalcode-search.html?q=auckland&country=NZ

- Household income data was taken from New Zealand's most recent Census (2018). These were downloaded as a.csv file from https://www.stats.govt.nz/

- Household income data required additional information. Latitude and longitude data was used to acquired "Postcodes" using https://nominatim.org/ . Missing Neighbourhood names were scrapped from https://www.nzpost.co.nz/personal/sending-within-nz/how-to-address-mail/postcodes/postcode-directory as a pdf and converted to a .csv file.

- Distances via the most direct route (road travel) were calculated from https://www.google.co.nz/maps

Data Cleaning

- A total of 202 neighbourhood's and 199 postcodes were identified for Auckland City.

- Latitudes and longitudes from the median household incomes data were used to acquire "postcodes" by reverse geocoding with https://nominatim.org/ . A data frame of additional postcodes and neighbourhood's was merged to provide missing neighbourhood data.

- The household incomes data frame was amended to included distances from the Central Business District (CBD) to each neighbourhood of interest. These were presented as 2 features:
  - Direct line between 2 points (latitude and longitude)
  - The most direct driving route (Google maps).

K-means Cluster 2 Neighbourhood's & Top Venues

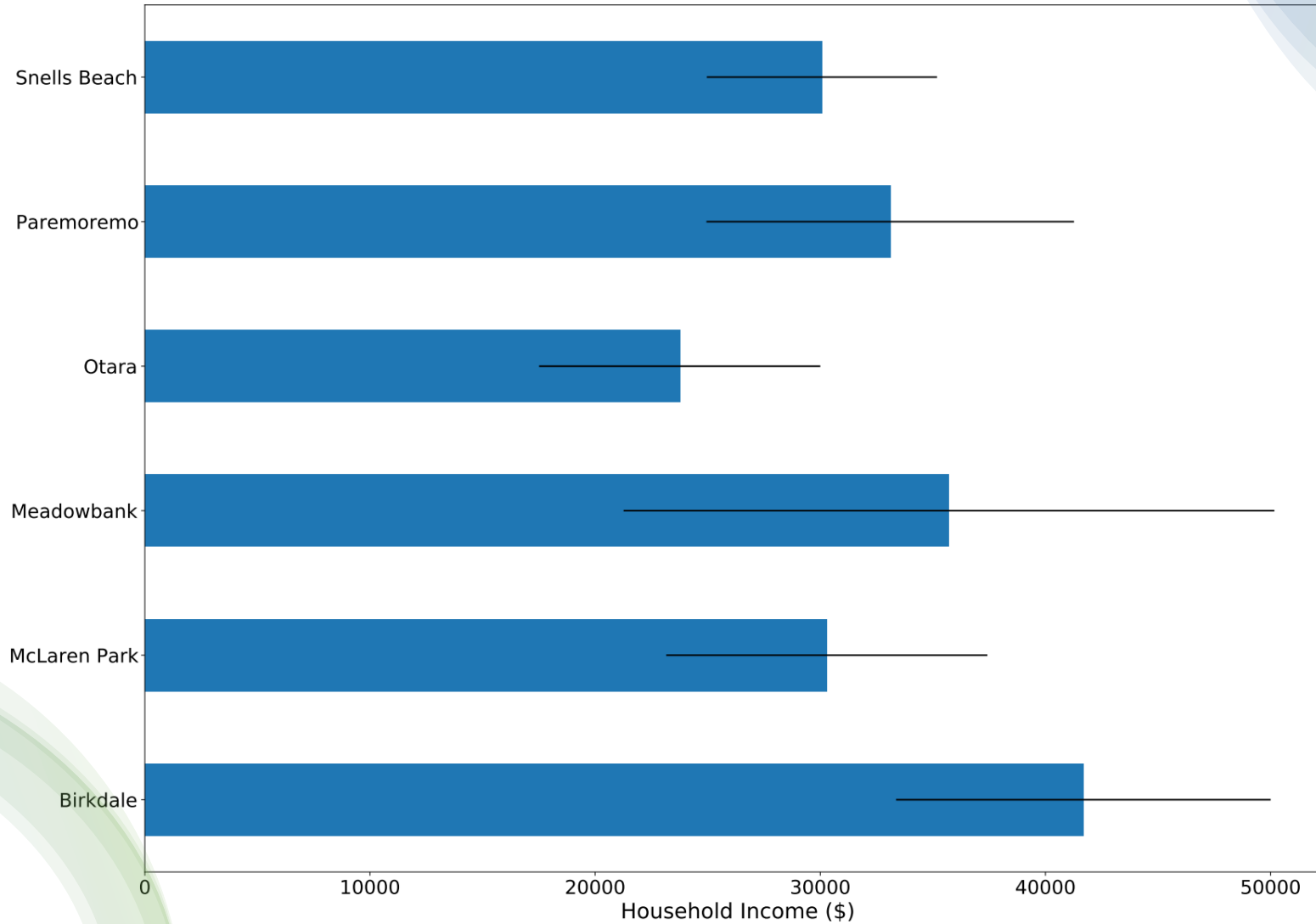| Neighbourhoods	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
McLaren Park	Indian Restaurant	Malay Restaurant	Dim Sum Restaurant	German Restaurant	Gastropub	Fried Chicken Joint	French Restaurant	Food Court	Fish & Chips Shop	Filipino Restaurant
Birkdale	Indian Restaurant	Bakery	Fast Food Restaurant	Yakitori Restaurant	Dim Sum Restaurant	German Restaurant	Gastropub	Fried Chicken Joint	French Restaurant	Food Court
Paremoremo	Indian Restaurant	Yakitori Restaurant	Dim Sum Restaurant	German Restaurant	Gastropub	Fried Chicken Joint	French Restaurant	Food Court	Fish & Chips Shop	Filipino Restaurant
Snells Beach	Indian Restaurant	Restaurant	Deli / Bodega	German Restaurant	Gastropub	Fried Chicken Joint	French Restaurant	Food Court	Fish & Chips Shop	Filipino Restaurant
Meadowbank	Indian Restaurant	Café	Yakitori Restaurant	Dim Sum Restaurant	German Restaurant	Gastropub	Fried Chicken Joint	French Restaurant	Food Court	Fish & Chips Shop
Otara	Indian Restaurant	Yakitori Restaurant	Dim Sum Restaurant	German Restaurant	Gastropub	Fried Chicken Joint	French Restaurant	Food Court	Fish & Chips Shop	Filipino Restaurant

Table of results from K-means clustering, showing the cluster selected for further examination

Household Income of Cluster 2 Neighbourhood's

Neighbourhood	count	mean	std	min	25%	50%	75%	max
Birkdale	5124.0	41683.606557	8318.946846	19800.0	36400.0	41800.0	47800.0	61200.0
McLaren Park	30178.0	30289.240506	7132.470791	10300.0	24400.0	29900.0	36000.0	50500.0
Meadowbank	5796.0	35715.527950	14451.863287	13500.0	22400.0	35200.0	45500.0	72200.0
Otara	15703.0	23755.951092	6244.625565	11000.0	18900.0	22200.0	28400.0	58800.0
Paremoremo	2142.0	33105.555556	8162.100243	11700.0	26900.0	32700.0	39300.0	57700.0
Snells Beach	1984.0	30071.875000	5111.953432	21300.0	26800.0	29650.0	33275.0	45600.0

Table of basic statistics showing, means, standard deviations, minimum & maximum values, and quartiles for the six neighbourhood's of Cluster 2 .

Income's (mean ± Std Dev) for Cluster 2 Neighbourhood's

Distances for Neighbourhood's in Cluster 2

Neighbourhood	Direct Distance (Km's) [1]	Distance via road (Km's) [2]
McLaren Park	11.7	18.9
Birkdale	8.9	12.6
Paremoremo	14.1	25.3
Meadowbank	9.2	10.1
Otara	17.2	18
Snells Beach	46.6	66.6

Table of distances for the six neighbourhood's of Cluster 2 .

[1]   Distance calculated as a direct line between latitude and longitude points.

[2]   Distance calculated by most direct driving route, as determined by Googlemaps.com

Conclusions

- A cluster of six neighbourhood's that might potentially meet the needs of the client were found using a K-means algorithm.

- Birkdale and Meadowbank had higher median household, with Otara at the lower end.

- Birkdale and Meadowbank were the most central neighbourhood's with Snell's beach representing the most isolated location.

- Further information beyond the scope of this report would assist the client decide on the best location to open her first restaurant , such as commercial property availability and location, cost of property purchase or lease.