Applied Data Science Capstone

Final Presentation

Mohammad Owais Kerney

Agenda

Project Background

1

Data

2

Methodology

3

Results

4

bluewolf

1. Project Background

Background

Singapore and Hong Kong are rival cities yet very similar

Migration between these cities for work or better opportunities is common

Challenges

Migrating families need a source to compare districts in the two cities to find a suitable destination when moving between these places

Solution

A comparison of the similarities between the districts of Hong Kong and Singapore as a cluster analysis

2. Data

Source

- Wikipedia and Geohack for districts and their location coordinates

- Foursquare data for venues and facilities of these two cities

A Quick Look

	Name	Category	Latitude	Longitude	City	District	District Latitude	District Longitude
0	Immigration Department Travel Documents Issuin...	Government Building	22.286864	114.155123	Hong Kong	Central and Western	22.28666	114.15497
1	Harbour Building (海港政府大樓)	Government Building	22.286592	114.154997	Hong Kong	Central and Western	22.28666	114.15497
2	Marine Department 海事處	Government Building	22.286571	114.154943	Hong Kong	Central and Western	22.28666	114.15497
3	Infinitus Plaza (無限極廣場)	Shopping Mall	22.286463	114.153689	Hong Kong	Central and Western	22.28666	114.15497
4	Fitness First Platinum	Gym / Fitness Center	22.285881	114.154117	Hong Kong	Central and Western	22.28666	114.15497

	Name	Category	Latitude	Longitude	City	District	District Latitude	District Longitude
0	Arcade Money Changers	Currency Exchange	1.283997	103.851197	Singapore	District 01 - Raffles Place, Marina, Cecil	1.28372	103.851239
1	Clifford Centre	Office	1.283816	103.852219	Singapore	District 01 - Raffles Place, Marina, Cecil	1.28372	103.851239
2	City Money Changers	Financial or Legal Service	1.283475	103.850161	Singapore	District 01 - Raffles Place, Marina, Cecil	1.28372	103.851239
3	Ocean Financial Centre	Building	1.283260	103.851521	Singapore	District 01 - Raffles Place, Marina, Cecil	1.28372	103.851239
4	In Touch Physio	Medical Center	1.283503	103.851163	Singapore	District 01 - Raffles Place, Marina, Cecil	1.28372	103.851239

bluewolf

3. Methodology (Pre-Analysis)

Steps

A Quick Look

Next Up

Clustering

	District	ATM	Accessories Store	Acupuncturist	Adult Education Center	Advertising Agency	African Restaurant	Airport	Airport Gate	Airport Lounge	...	Waterfront	Well	Wine Bar	Wine Shop	Winery	Wings Joint
0	Central and Western	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
1	Central and Western	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
2	Central and Western	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
3	Central and Western	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
4	Central and Western	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0

One Hot Encoding

Grouping and Finding Top Venues

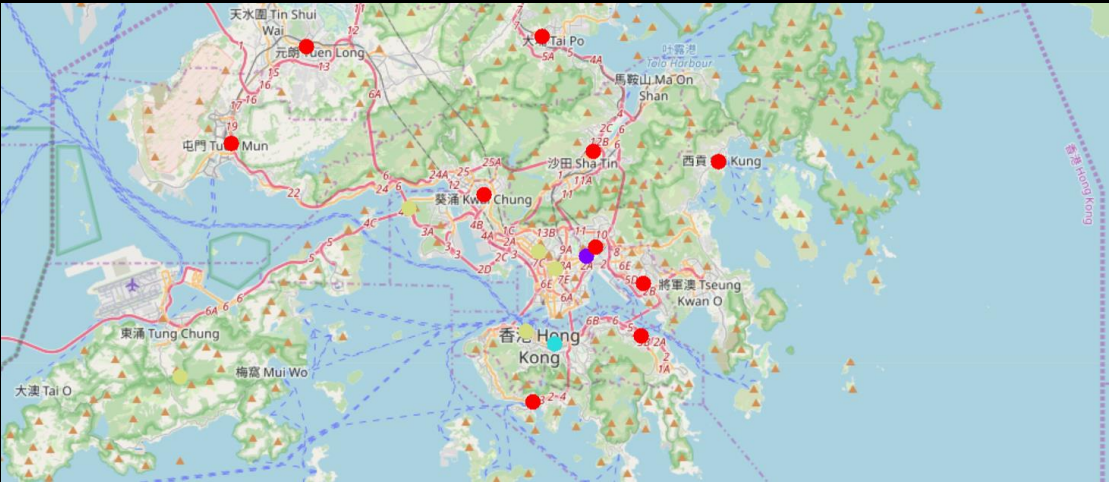|	District	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Central and Western	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space	Lawyer	Dentist's Office	Gym / Fitness Center	Coffee Shop
1	District 01 - Raffles Place, Marina, Cecil	Office	Bakery	Bank	Building	Café	Salad Place	Health & Beauty Service	Medical Center	Japanese Restaurant	Asian Restaurant
2	District 02 - Tanjong Pagar, Chinatown	Japanese Restaurant	Office	Café	Coffee Shop	Building	Sandwich Place	Asian Restaurant	Chinese Restaurant	Cosmetics Shop	Park
3	District 03 - Tiong Bahru, Alexandra, Queenstown	College Classroom	Building	Coffee Shop	Chinese Restaurant	Residential Building (Apartment / Condo)	Asian Restaurant	Food Court	Dentist's Office	Medical Center	Nail Salon
4	District 04 - Mount Faber, Telok Blangah, Harb...	Clothing Store	Chinese Restaurant	Office	Cosmetics Shop	Jewelry Store	Food Court	Boutique	Café	Sporting Goods Shop	Housing Development

3. Methodology (Clustering)

Steps

Use k-means clustering with k=4

Grouping all results

A Quick Look

	Name	Category	Latitude	Longitude	City	District	District Latitude	District 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
0	Immigration Department Travel Documents Issuin...	Government Building	22.286864	114.155123	Hong Kong	Central and Western	22.28666	114.15497	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space
1	Harbour Building (海港政府大樓)	Government Building	22.286592	114.154997	Hong Kong	Central and Western	22.28666	114.15497	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space
2	Marine Department 海事處	Government Building	22.286571	114.154943	Hong Kong	Central and Western	22.28666	114.15497	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space
3	Infinitus Plaza (無限極廣場)	Shopping Mall	22.286463	114.153689	Hong Kong	Central and Western	22.28666	114.15497	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space
4	Fitness First Platinum	Gym / Fitness Center	22.285881	114.154117	Hong Kong	Central and Western	22.28666	114.15497	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space

bluewolf

4. Results

Findings

Two giant clusters with 18 and 24 clusters each highlighting the similarity of the two cities

The biggest cluster is shown here

A Quick Look

	Category	District	District Latitude	District 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	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	
0	Government Building	Central and Western	22.286660	114.154970	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space	Lawyer	Dentist's Office	Gym / Fitness Center	
1	Government Building	Central and Western	22.286660	114.154970	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space	Lawyer	Dentist's Office	Gym / Fitness Center	
2	Government Building	Central and Western	22.286660	114.154970	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space	Lawyer	Dentist's Office	Gym / Fitness Center	
3	Shopping Mall	Central and	22.286660	114.154970	3	Building	Government Building	Doctor's Office	Office	Bus Stop	Coworking Space	Lawyer	Dentist's Office	Gym / Fitness	

Discussion

The model provides a good starting point for migrating families to find suitable districts to move to

Further enhancements could include considering crime rate, population and health data

The End