Singapore Planning Areas

Are they created equal?

Background

Singapore is a small country and is divided into several planning areas.

Are these planning areas created equal?

- Classify and explore planning areas for someone who's on the hunt for a new home

- Differences and similarities between planning areas can be identified to help in house hunting

Data Sources

OneMap SG

Provides planning areas and population details for each planning area.

FourSquare API

Provides venues details for each venue area.

Methodology

1. **Data Collection**

   Data is extracted from OneMap and Foursquare API separately and merged for wrangling.

2. **Data Wrangling**

   Excluded the following data;

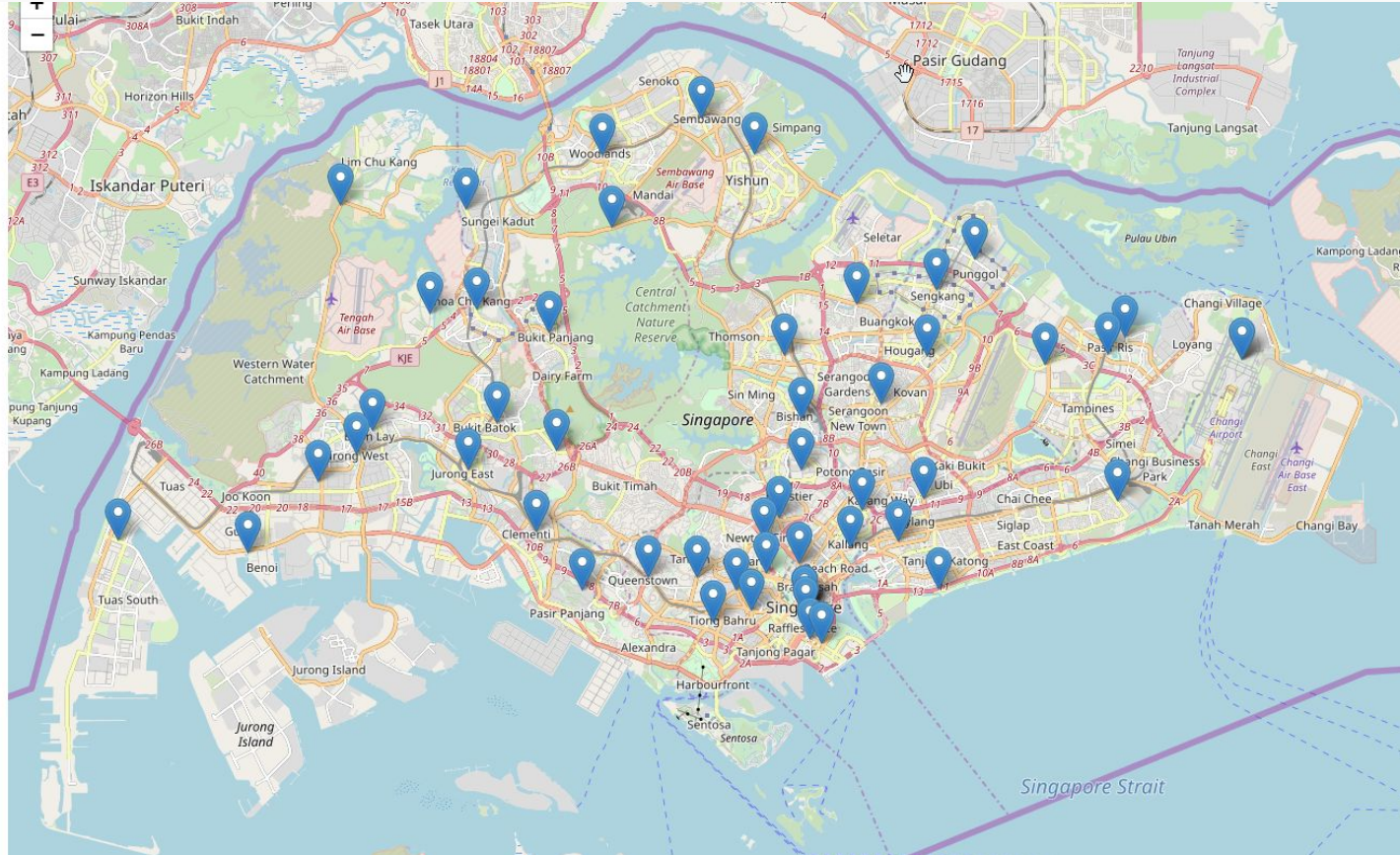   > Area with zero(0) population

   > Neighborhood venue category

   > Building venue category

3. **Clustering & Analysis**

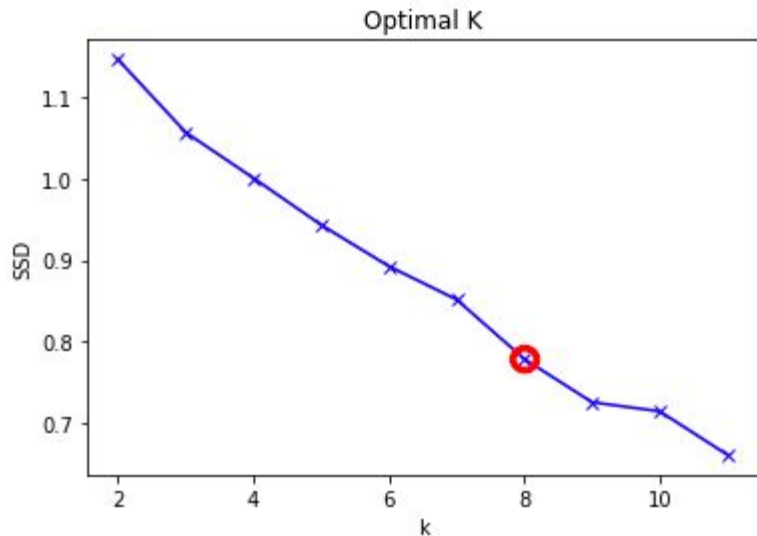   Perform clustering and analyze results.

Planning Areas

There are a total of 55 planning areas in Singapore.

After data wrangling, 43 planning areas are left.

Clustering

KMeans clustering is used cluster planning areas based on venue categories as features.
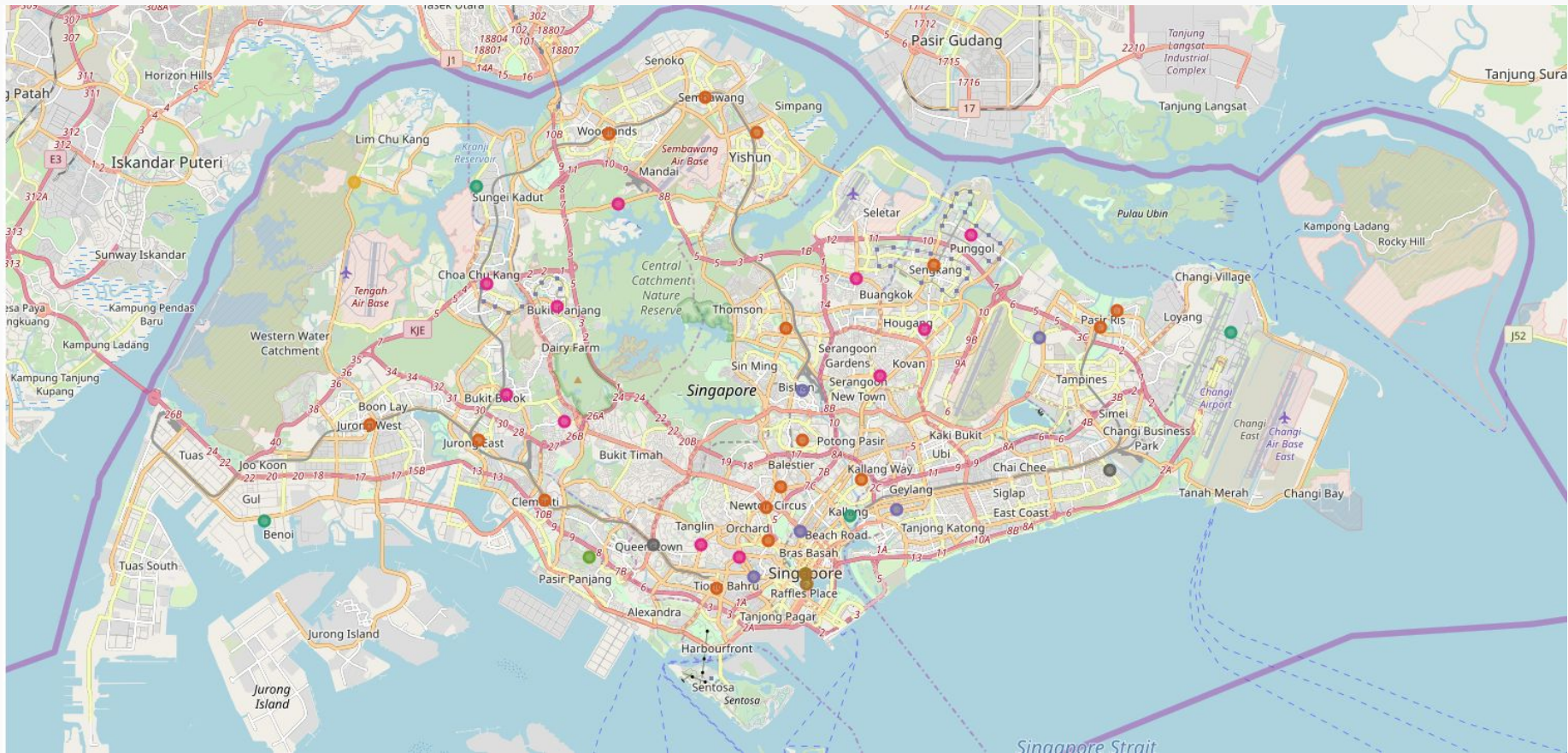
K-Value

Optimal K

Elbow method is used to determine the optimal k-value.

Several executions of elbow method were performed as each run gives different results.

In the end, "8" is found to be the most common value.

Results

Results Analysis - 1

Most Suitable Residential Areas

- **Cluster 5**
  Areas in this cluster have a good mix of amenities and places to eat out.

- **Cluster 3**
  Foodies' heavens, these areas are a good choice when it comes to filling up your tummy.

- **Cluster 4**
  Residential areas in this area could be in the higher price range due to their proximity to hotels and entertainment venues.

Results Analysis - 2

Cluster 5

Areas in this cluster have a good mix of amenities and places to eat out.

```
merged_df.loc[merged_df['Cluster Labels'] == 4, merged_df.columns[[0] + [1] + list(range(3, merged_df.shape[1]))]]
```

	Planning Area	population	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
7	HOUGANG	223010	Bus Station	Parking	Salon / Barbershop	Student Center	Lounge	Field	Metro Station	Asian Restaurant	Residential Building (Apartment / Condo)	Moving Target
10	MANDAI	2090	Exhibit	Ice Cream Shop	Trail	Other Great Outdoors	Office	Electronics Store	Asian Restaurant	Meeting Room	Park	Coffee Shop
11	BUKIT PANJANG	71290	Residential Building (Apartment / Condo)	Parking	Bus Station	College Administrative Building	Office	Park	Housing Development	Government Building	Gym	Café
12	BUKIT TIMAH	41450	Café	Residential Building (Apartment / Condo)	Seafood Restaurant	Doctor's Office	Spa	Bus Line	Dentist's Office	Pet Store	Bakery	Bike Shop
17	BUKIT BATOK	73550	Doctor's Office	Residential Building (Apartment / Condo)	Furniture / Home Store	Spa	Parking	Bus Station	Bank	Salon / Barbershop	Bus Line	Bakery
19	CHOA CHU KANG	94070	Bus Station	Student Center	Karaoke Bar	Bakery	Food Truck	Other Nightlife	Gas Station	Jewelry Store	Nightclub	Salon / Barbershop

Results Analysis - 3

Cluster 3

Foodies' heavens, these areas are a good choice when it comes to filling up your tummy.

```
merged_df.loc[merged_df['Cluster Labels'] == 2, merged_df.columns[[0] + [1] + list(range(3, merged_df.shape[1]))]]
```

	Planning Area	population	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	MARINE PARADE	25050	Noodle House	Asian Restaurant	Bakery	Chinese Restaurant	Residential Building (Apartment / Condo)	Snack Place	Coffee Shop	Dessert Shop	Café	BBQ Joint
3	JURONG WEST	266720	Japanese Restaurant	Bakery	Sporting Goods Shop	Asian Restaurant	Shoe Store	Fast Food Restaurant	Cosmetics Shop	Dessert Shop	Medical Center	Food Court
5	NOVENA	23700	Sporting Goods Shop	Office	Café	Outdoor Supply Store	Ice Cream Shop	Japanese Restaurant	Dance Studio	Thai Restaurant	Basketball Court	Juice Bar
6	PASIR RIS	74440	Bank	Dessert Shop	Bakery	Cosmetics Shop	Japanese Restaurant	Café	Sushi Restaurant	Food Court	Fast Food Restaurant	Convenience Store
8	JURONG EAST	40910	Japanese Restaurant	Cosmetics Shop	Dessert Shop	Pharmacy	Bakery	Café	Chinese Restaurant	Juice Bar	Korean Restaurant	Snack Place
13	ANG MO KIO	165710	Chinese Restaurant	Residential Building (Apartment / Condo)	Karaoke Bar	Noodle House	Medical Center	Dessert Shop	Doctor's Office	Fast Food Restaurant	Asian Restaurant	Steakhouse

Results Analysis - 4

Cluster 4

Residential areas in this area could be in the higher price range due to their proximity to hotels and entertainment venues.

```python
merged_df.loc[merged_df['Cluster Labels'] == 3, merged_df.columns[[0] + [1] + list(range(3, merged_df.shape[1]))]]
```

	Planning Area	population	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
5	BISHAN	42900	Coffee Shop	Food Court	Residential Building (Apartment / Condo)	Playground	College Classroom	Seafood Restaurant	Fried Chicken Joint	Housing Development	Post Office	Other Great Outdoors
1	GEYLANG	55250	Hotel	Temple	Coffee Shop	Residential Building (Apartment / Condo)	Chinese Restaurant	Motel	Gas Station	Korean Restaurant	Strip Club	Food Court
3	TAMPINES	131330	Temple	Coffee Shop	Bus Station	Bus Stop	Café	Office	Asian Restaurant	Martial Arts Dojo	Athletics & Sports	Spiritual Center
0	OUTRAM	20030	Hotel	Nightclub	Hotel Bar	Buffet	Event Space	Pool	Meeting Room	Bus Station	Lounge	Office
4	ROCHOR	13450	Hotel	Bar	Chinese Restaurant	Asian Restaurant	Coffee Shop	Hostel	Turkish Restaurant	Boutique	Food Court	Korean Restaurant

Closing Comments

Limitation

Due to the fact that venues are extracted based on central coordinates of planning areas, some venues may not have been extracted. This may adversely impact cluster results.

Improvements

Limitations stated above may be overcome by using GeoJSON data that indicates the boundaries of each planning area. Based on the maximum radius, venues around each coordinate can be retrieved to further improve the cluster results.

Thank you.

Stephen