

Picking A Singaporean Neighborhood to Live In

Mao Tian Tan

Introduction

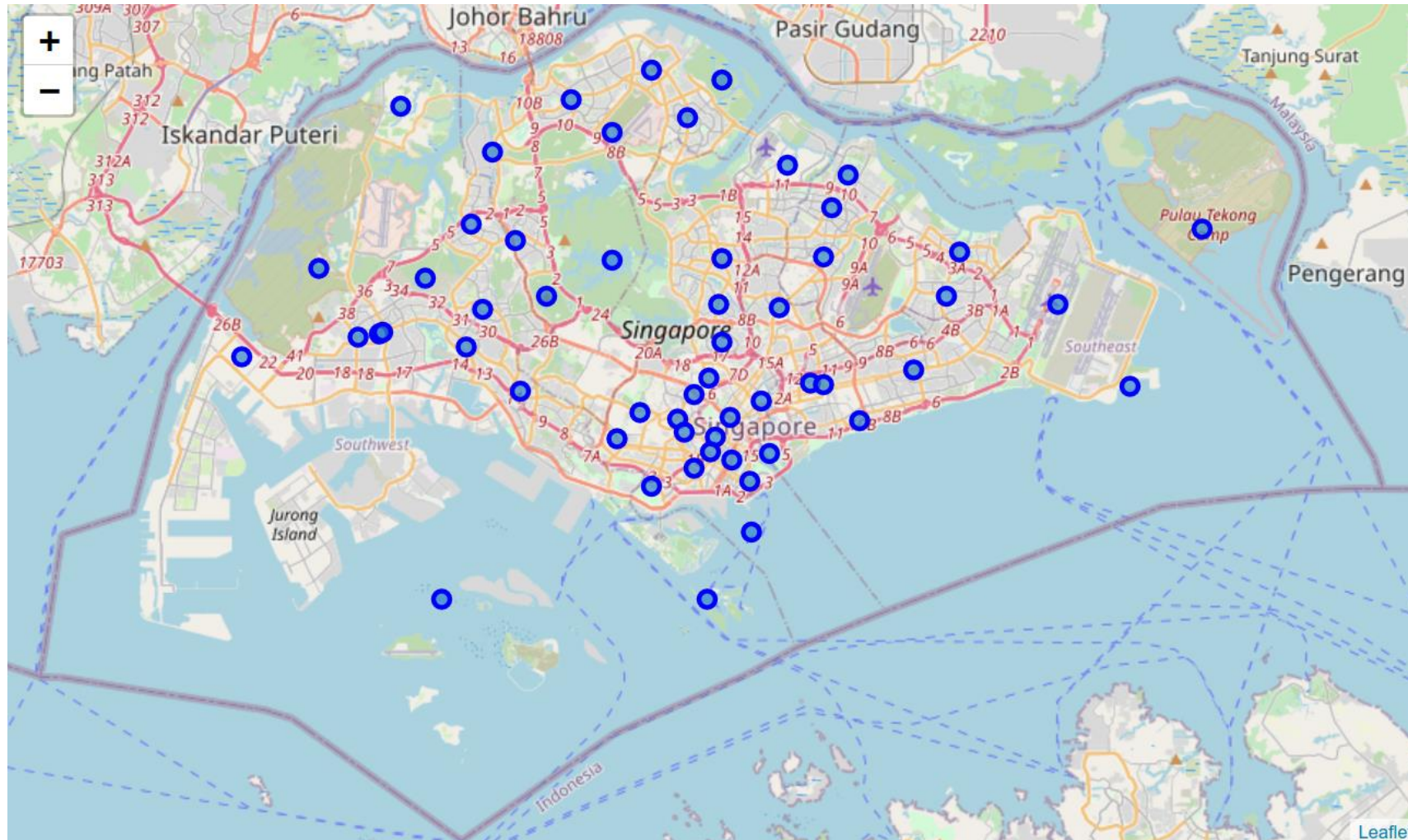
- Large number of neighborhoods in Singapore
- Difficult for foreign tenant to pick neighborhood
- Goal: Narrow down the choices by grouping the neighborhoods into clusters.

Data

- List of 55 planning areas can be scraped from [https://en.wikipedia.org/wiki/Planning Areas of Singapore](https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore)
- Latitudes and longitude of these neighborhoods are obtained from geopy
- Use explore query in Foursquare API to obtain venues with 500 meters of these coordinates.
- Changing the search radius to 2km has little impact on the analysis.
- Create dataframe of venues with the features: neighborhood, neighborhood coordinates, venue name, venue coordinates, venue category

Exploratory Data Analysis

Map of Singapore along with 55 planning areas.



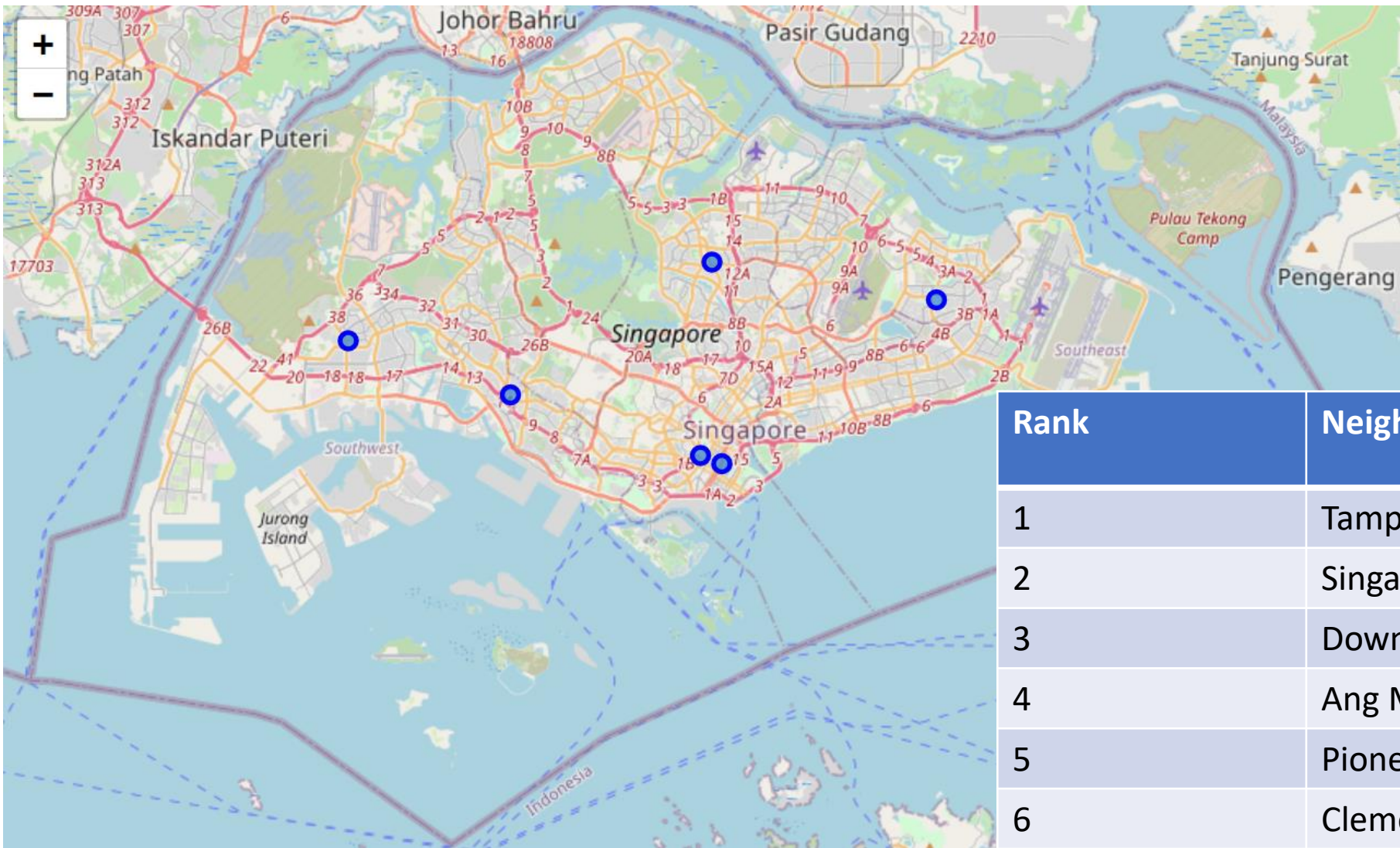
Exploratory Data Analysis

Top five most common venues in Singapore

Rank	Venue Category	Number of venue
1	Chinese Restaurant	84
2	Coffee Shop	81
3	Food Court	71
4	Japanese Restaurant	70
5	Asian Restaurant	59

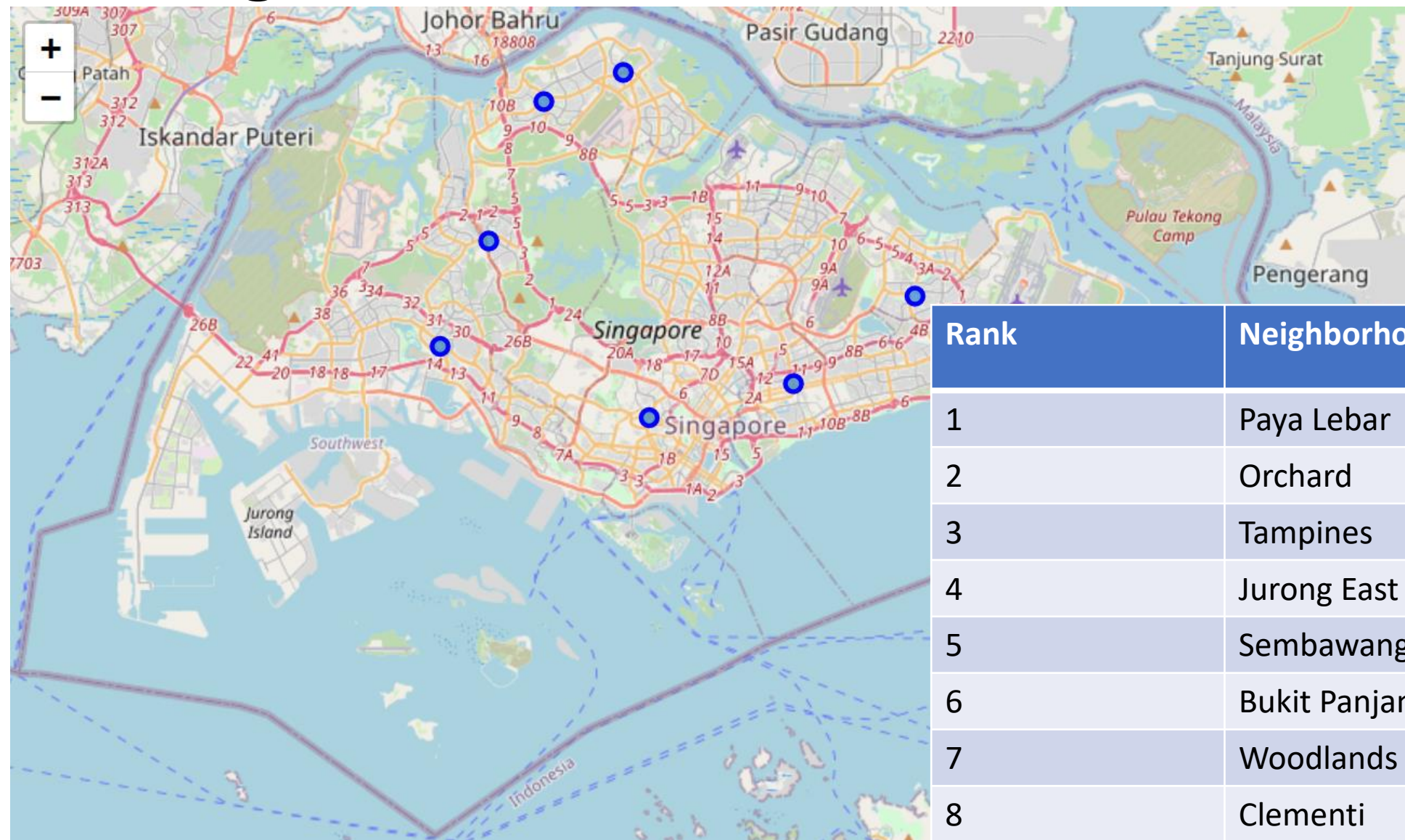
These are all in the food and beverage industry. Singapore has a large selection of dining options

Neighborhoods with Most Gyms



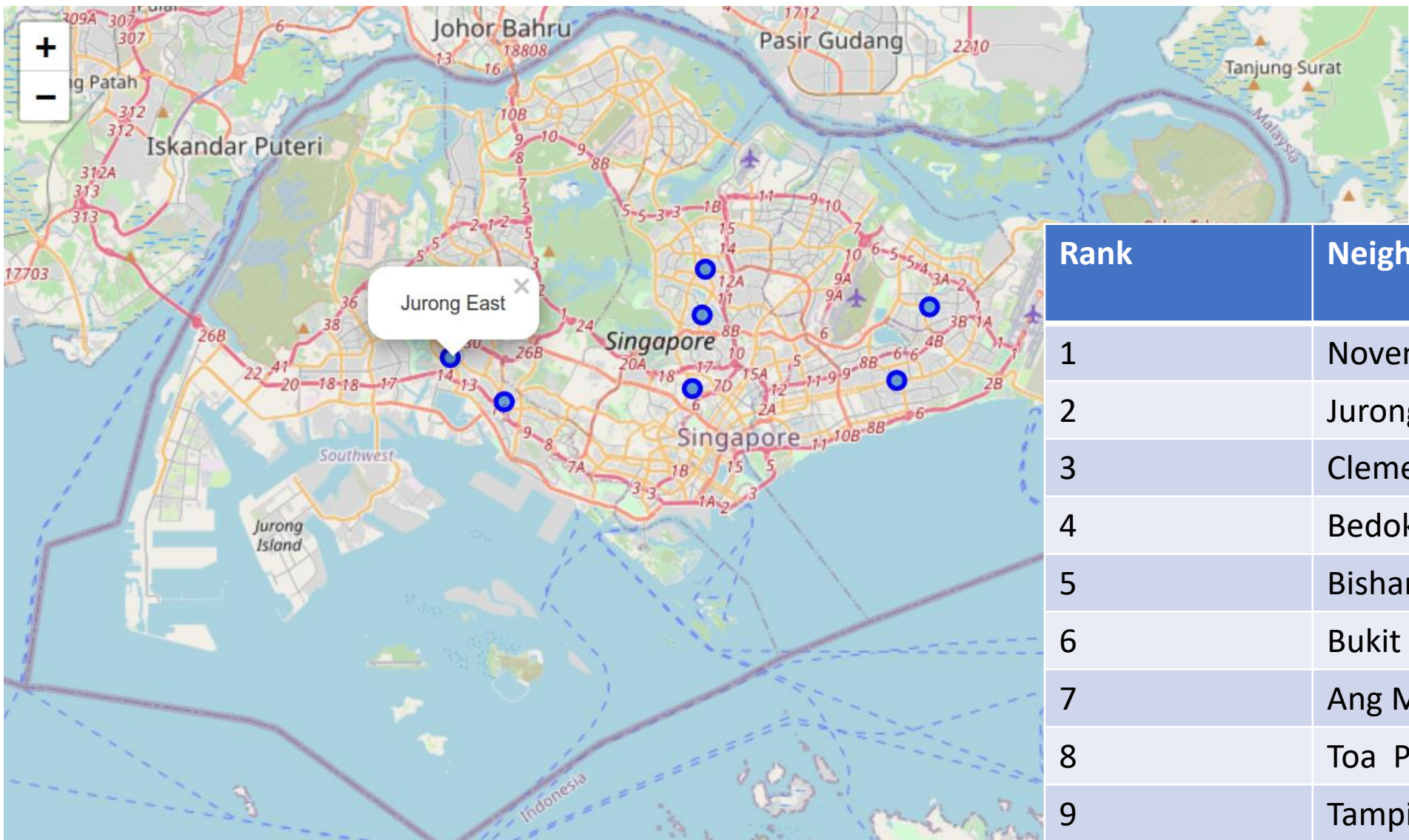
Rank	Neighborhood	Number of Gyms
1	Tampines	4
2	Singapore River	3
3	Downtown Core	3
4	Ang Mo Kio	2
5	Pioneer	2
6	Clementi	2

Neighborhoods with Most Malls



Rank	Neighborhood	Number of Malls
1	Paya Lebar	6
2	Orchard	4
3	Tampines	3
4	Jurong East	3
5	Sembawang	2
6	Bukit Panjang	2
7	Woodlands	2
8	Clementi	2

Neighborhoods with Most Coffee Shops

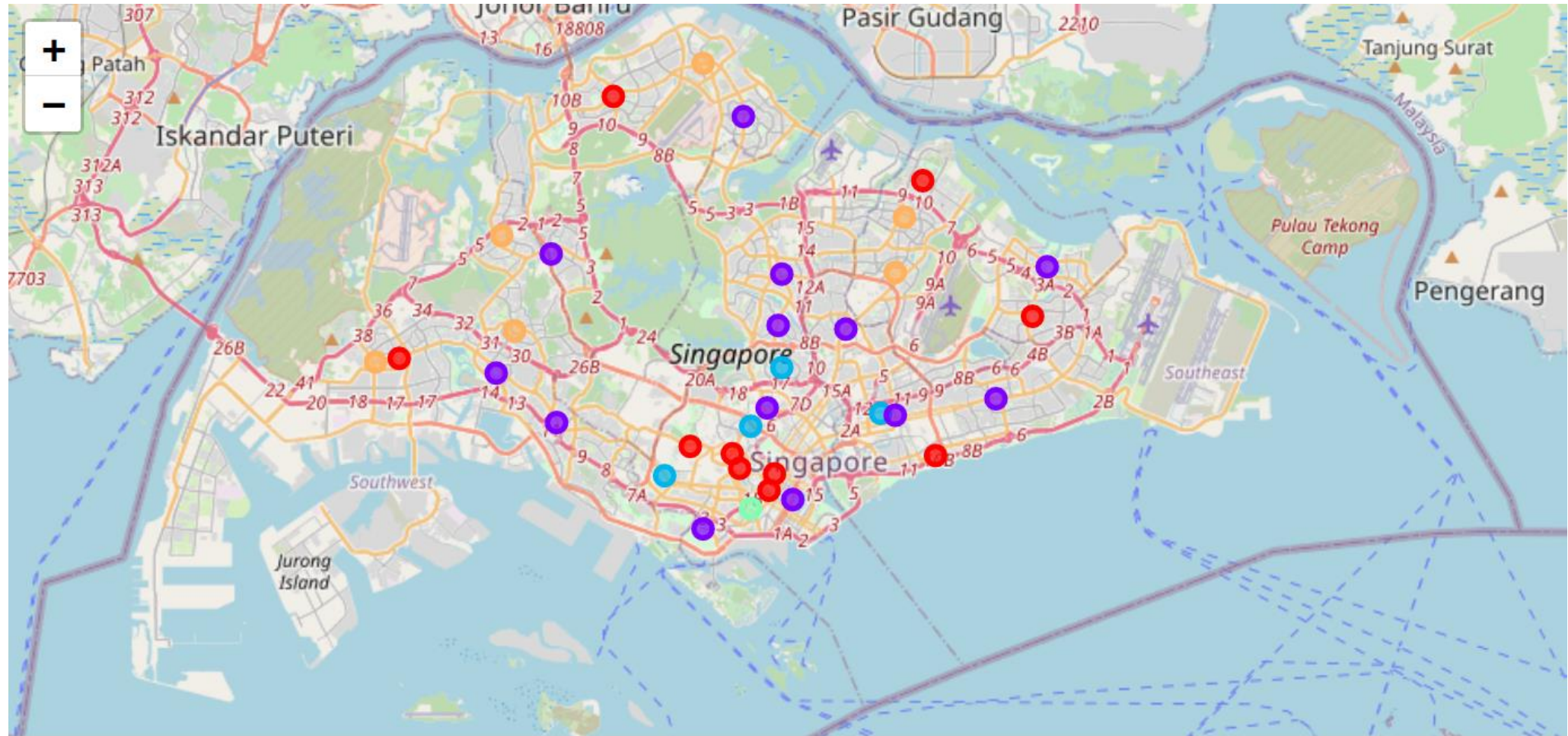


Rank	Neighborhood	Number of Coffee Shops
1	Novena	6
2	Jurong East	6
3	Clementi	6
4	Bedok	4
5	Bishan	4
6	Bukit Batok	4
7	Ang Mo Kio	4
8	Toa Payoh	4
9	Tampines	4

Machine Learning Methodology

- Use k-means clustering algorithm to group the neighborhoods into similar groups
- First, remove the outlier neighborhoods that are not residential district, resulting in 34 remaining neighborhoods.
- Grouping into five clusters results in a more even distribution of neighborhoods over the clusters.

Results



Results

Top five most common venues (taking into account ties) along with the number of such venues in each cluster.

Cluster	Top Five Most Common Venues
0	Japanese Restaurant (41), Café (22), Hotel (21), Chinese Restaurant (20), Bakery (20)
1	Coffee Shop (43), Food Court (38), Chinese Restaurant (30), Asian Restaurant (24), Café (23)
2	Chinese Restaurant (21), Noodle House (14), Food Court (12), Seafood Restaurant (6), Asian Restaurant (5), Coffee Shop (5)
3	Spanish Restaurant (1), Seafood Restaurant (1), Nail Salon (1), Hotel (1), Grocery Store (1), Gastropub (1), Farmers Market (1), Dim Sum Restaurant (1), Coffee Shop (1), Café (1), Bakery (1)
4	Coffee Shop (13), Fast Food Restaurant (12), Bus Station (12), Chinese Restaurant (11), Food Court (11)

Results

- The five clusters of neighborhoods have distinct flavors

Cluster	Neighborhoods	
0	Boon Lay, Marine Parade, Museum, Orchard, Punggol, River valley, Singapore River, Tampines, Tanglin, Woodlands	Exciting
1	Ang Mo Kio, Bedok, Bishan, Bukit Merah, Bukit Panjang, Clementi, Downtown Core, Jurong East, Novena, Pasir Ris, Paya Lebar, Serangoon, Yishun	Regular
2	Geylang, Newton, Queenstown, Toa Payoh	Central and ungentrified
3	Outram	Unique
4	Bukit Batok, Choa Chu Kang, Hougang, Pioneer, Sembawang, Senkang	Far from central and ungentrified

Conclusion

- Divided 34 Singaporean neighborhoods into 5 clusters.
- K-mean clustering worked well in identifying clusters with distinct flavors
- Potential tenants can narrow down their choice of neighborhoods based on preferences and budget