Singapore Visitors Venue Recommendation

By Kanupuri Bharadwaja

Date: 14/04/2020

Introduction: Business Problem

The goal of this project is to collect best visited/reviewed places using Foursquare API and provide an accurate recommendation. We will be utilizing data retrieved from Singapore open data sources and FourSquare API venue recommendations.

If a person is planning to visit Singapore as a Tourist and looking for reasonable accommodation.

If the user wants to receive a venue recommendation where he can stay - places like Best Hotels or rent an apartment according to his budget with nearby places of interest or search category option.

With clear segmenting and cluster analysis I would like to present a comparison table of all possible town venues.

Data Requirements, Collection

We will download the data from data.gov.sg and utilize the Median Rental prices

- Singapore Median Rental Prices by town.
- Popular Food venues in the vicinity. (Sample category selection)

Singapore Towns and median residential rental prices.

Data retrieved from Singapore https://data.gov.sg website. We will use the most recent recorded rental prices from this data source (Q4 2019) being the most relevant price available at this time.

Note: While this demo makes use of Food Venue Category, Other possible categories can also be used for the same implementation such as checking categories like:

- Outdoors and Recreation
- Nightlife

As FourSquare API only allows 50 free venue queries, I would like to limit the scope of this search.

Data Cleaning and Preparation

Singapore Towns List with median residential rental prices.

The source data contains median rental prices of Singapore HDB units from 2005 up to 4th quarter of 2019.

Data Cleanup and re-grouping. The retrieved table contains some un-wanted entries and needs some cleanup.

The following tasks will be performed:

- Drop/ignore cells with missing data.
- Use most current data records.
- Fix data types.

We will use the Geocode library to retrieve the coordinates (latitude and longitude of each town center. For this exercise, I just used the MRT stations as the center points of each evaluated town.
The town coordinates will be used in retrieval of Foursquare API location data.

Performing Segmentation and Clustering Analysis in Singapore

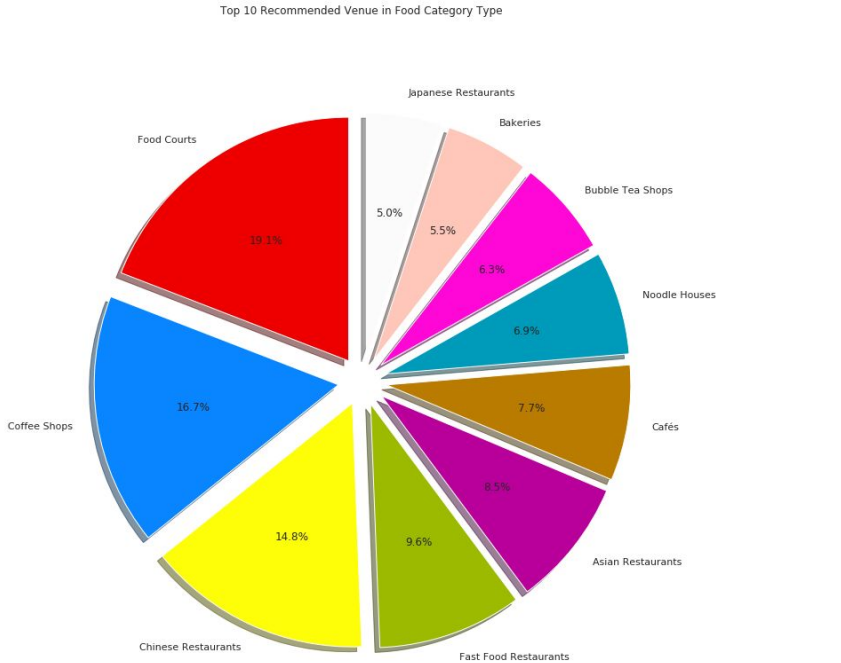
Retrieving FourSquare Places of interest.

We will be using the Foursquare API - **explore** function to get the most common venue categories in each neighborhood and then segment the neighborhoods into clusters.

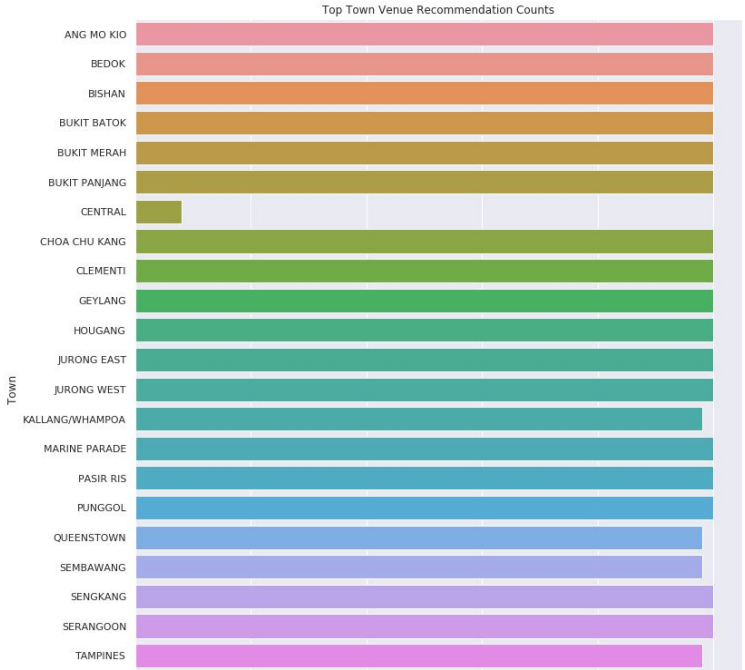
We will use the *k*-means clustering algorithm for analysis.

Finally, we will use the Folium library to visualize the recommended neighborhoods and their emerging clusters.

Exploratory Data Analysis - 1



Exploratory Data Analysis - 2



Modeling

Clustering Neighborhoods

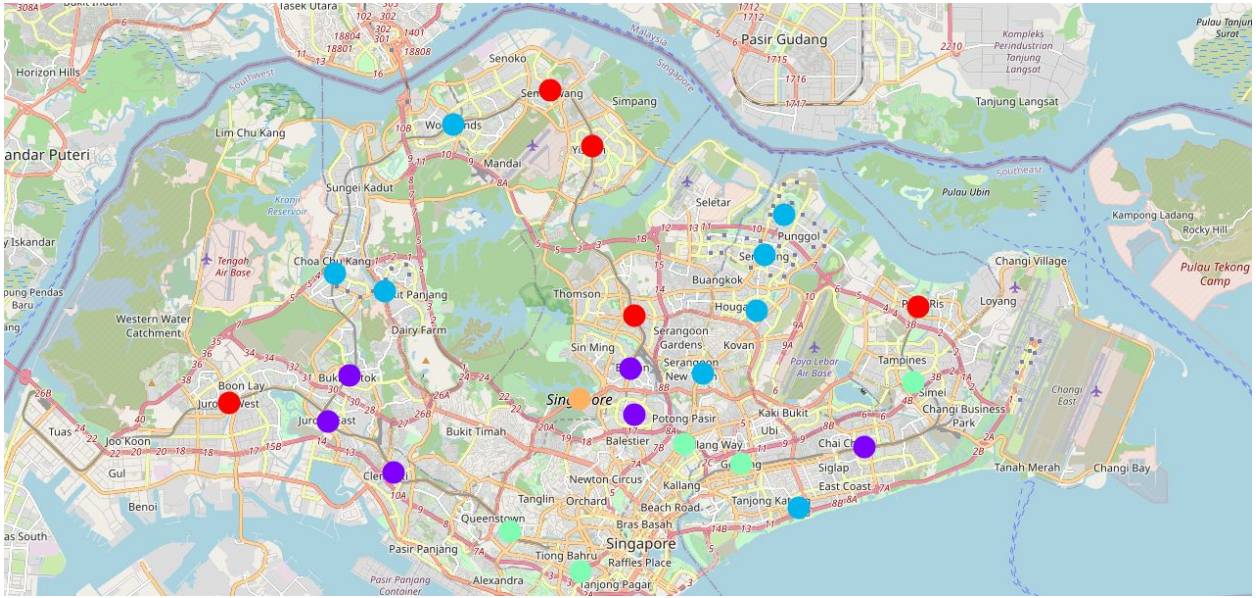
Run *k*-means to cluster the Towns into 5 clusters.

In [323]:
```
# set number of clusters
kclusters = 5
sg_grouped_clustering = sg_grouped.drop('Town', 1)
# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=1).fit(sg_grouped_clustering)

# check cluster labels generated for each row in the dataframe
print(kmeans.labels_[0:10])
print(len(kmeans.labels_))
```

```
[0 1 1 1 3 2 4 2 1 3]
25
```

Evaluation and Results



Conclusion and Future Directions

In this notebook, we've analyzed the best town venues and provided recommendations based on the Best Food venue category.

The generated clusters from our results shows that there are very good and interesting places located in areas where the median rents are cheaper. This kind of results may be very interesting for travelers who are travelling on a limited budget.

Results show that most popular food venues among Singaporeans, residents and visitors are **Food Courts, Coffee Shops, Chinese Restaurants and Fast Food Restaurants**.

Thank you.

References

https://data.gov.sg
https://www.foursquare.com