1.   Introduction

In Singapore, a cosmopolitan city-state in Southeast Asia, food is central to the national identity and is an important cultural and social activity. Singaporeans are obsessed with eating, and take pride in the fact that Singapore is a gastronomic heaven with both delicious local cuisine and world-class international fares. Internationally, Singapore is also well-known for its food, especially its local hawker cuisine, and food is a major tourist attraction in the country.

This project seeks to use location data to determine the neighborhood in Singapore with the best food. This information has several useful applications for different target audiences:

- Tourists who are foodies can find places to stay that will give them easy access to the best food, or find neighborhoods that they can visit for a one-stop food trial.
- Considering the importance of food nationally, Singaporeans can use the information as an additional metric when they are considering where to move and stay.

2.   Data

- I obtained the list of neighborhoods in Singapore from the Wikipedia page: https://en.wikipedia.org/wiki/Planning_Areas_of_Singapore
- The latitude and longitude coordinates of each neighborhood were collected manually using Google search.
- With this information, I use the Foursquare API to search for food venues near each of the neighborhoods, which gives me the following information:
    - ID
    - Name
    - Latitude
    - Longitude
    - Category
- For each of these food venues, I use the API to get further information on the location, namely:
    - Overall rating
    - Number of tips
    - Number of likes

3.   Methodology

3.1.   Web Scrapping and Data Collection

I first scrapped the Wikipedia page for the list of neighborhoods in Singapore, and cleaned the data which gave me the following DataFrame:

	Neighborhood	Region	Population
0	Ang Mo Kio	North-East	163,950
1	Bedok	East	279,380
2	Bishan	Central	88,010
3	Boon Lay	West	30
4	Bukit Batok	West	153,740

I collected the latitude and longitude coordinates of each neighborhood manually using Google search. I compiled the information in a CSV file hosted on my GitHub repository and imported the file. Combining the two DataFrame together, I get the following data:

	Neighborhood	Region	Population	Latitude	Longitude
0	Ang Mo Kio	North-East	163,950	1.3691	103.8454
1	Bedok	East	279,380	1.3236	103.9273
2	Bishan	Central	88,010	1.3526	103.8352
3	Boon Lay	West	30	1.3386	103.7058
4	Bukit Batok	West	153,740	1.3590	103.7637

3.2. Getting Information on Food Venues

I used the Foursquare API to make a venue search with the search term 'food' to get the list of food places within each neighborhood. For each of these venues, I extracted the following details from the result:

- ID
- Name
- Latitude
- Longitude
- Category

I cleaned the data by removing results whose categories are not food-related. I get the following DataFrame of shape (874,8):

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	ID	Venue	Venue Latitude	Venue Longitude	Category
0	Ang Mo Kio	1.3691	103.8454	4c657dabe2ee9c749ce32d2f	Ang Mo Kio Central Market & Food Centre	1.372037	103.846366	Food Court
1	Ang Mo Kio	1.3691	103.8454	4d85cc1d509137042a3b9e5b	AMK 722 Food House	1.371989	103.847495	Food Court
2	Ang Mo Kio	1.3691	103.8454	4b691f1ff964a520ba992be3	Food Loft	1.372554	103.847913	Food Court
3	Ang Mo Kio	1.3691	103.8454	4bc08969b492d13a345ba460	AMK Food House	1.368245	103.844001	Coffee Shop
4	Ang Mo Kio	1.3691	103.8454	4b149e47f964a52023a523e3	S11 Food House	1.371501	103.847783	Food Court

For each of these venues, I then used the Foursquare API get the details of the venue. I extracted the following details from the result:

- Rating
- Tip Count
- Likes

This gives me the following data:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	ID	Venue	Venue Latitude	Venue Longitude	Category	Rating	Tip Count	Likes
0	Ang Mo Kio	1.3691	103.8454	4c657dabe2ee9c749ce32d2f	Ang Mo Kio Central Market & Food Centre	1.372037	103.846366	Food Court	6.8	162	99
1	Ang Mo Kio	1.3691	103.8454	4d85cc1d509137042a3b9e5b	AMK 722 Food House	1.371989	103.847495	Food Court	5.7	16	11
2	Ang Mo Kio	1.3691	103.8454	4b691f1ff964a520ba992be3	Food Loft	1.372554	103.847913	Food Court	4.9	30	5
3	Ang Mo Kio	1.3691	103.8454	4bc08969b492d13a345ba460	AMK Food House	1.368245	103.844001	Coffee Shop	5.5	14	2
4	Ang Mo Kio	1.3691	103.8454	4b149e47f964a52023a523e3	S11 Food House	1.371501	103.847783	Food Court	5.7	82	56

3.3. Data Wrangling

For venues without a rating, tip count or likes, I replace the empty values with the mean values of the other venues in the same category. I normalize the rating, tip count and likes values using simple feature scaling such that they range from 0 to 1.

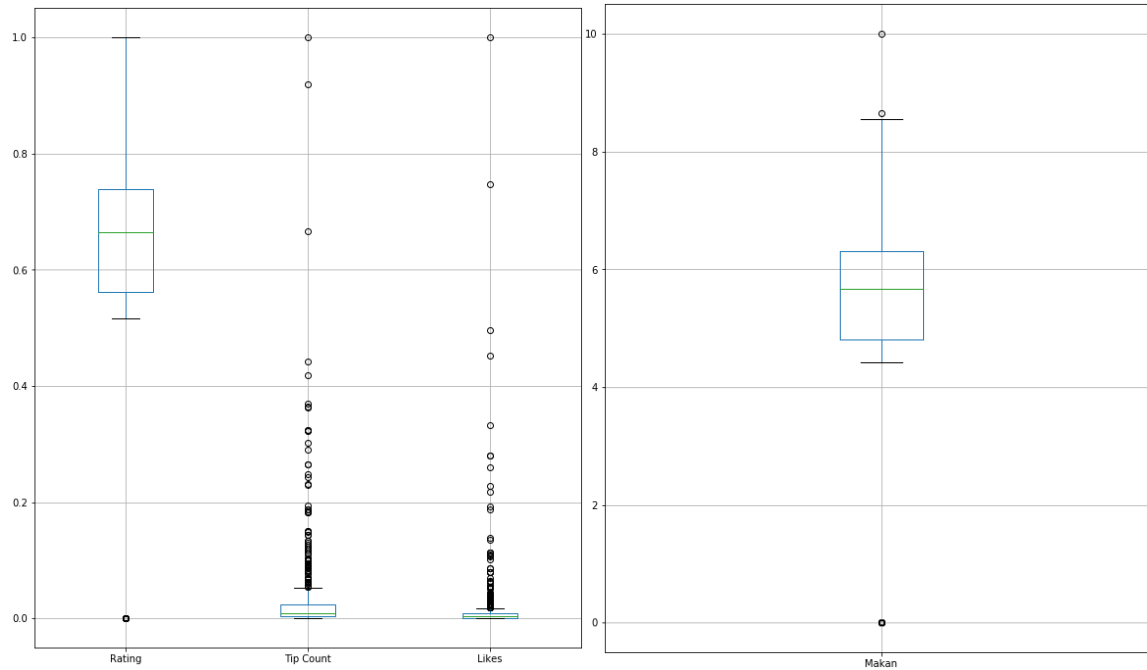The following shows the normalized values:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	ID	Venue	Venue Latitude	Venue Longitude	Category	Rating	Tip Count	Likes
0	Ang Mo Kio	1.3691	103.8454	4c657dabe2ee9c749ce32d2f	Ang Mo Kio Central Market & Food Centre	1.372037	103.846366	Food Court	0.764045	0.364865	0.107492
1	Ang Mo Kio	1.3691	103.8454	4d85cc1d509137042a3b9e5b	AMK 722 Food House	1.371989	103.847495	Food Court	0.640449	0.036036	0.011944
2	Ang Mo Kio	1.3691	103.8454	4b691f1ff964a520ba992be3	Food Loft	1.372554	103.847913	Food Court	0.550562	0.067568	0.005429
3	Ang Mo Kio	1.3691	103.8454	4bc08969b492d13a345ba460	AMK Food House	1.368245	103.844001	Coffee Shop	0.617978	0.031532	0.002172
4	Ang Mo Kio	1.3691	103.8454	4b149e47f964a52023a523e3	S11 Food House	1.371501	103.847783	Food Court	0.640449	0.184685	0.060803

3.4. Scoring Venues and Neighborhoods

I give each venue a 'Makan' score (makan is the Malay word for eating, as well as a commonly used word in the local English creole, Singlish).

The Makan score is tabulated using the following formula: rating + (tip count + like) * 1/4.

- Given that most of tip count and likes are closer to zero (see figure below), the formula reduces the extent of the penalty for venues that have no tips or likes.
- After experimenting with various formulae, the above formula gives a reasonably good outcome with a decent spread in the scores and few outliers (see figure below).
- Venues with a high number of tips and likes (i.e. close to normalized value of 1), will get up to a maximum 50% boost on their overall score.

I normalize the scores to get a 'Makan' score out of 10. I then calculate the aggregate 'Makan' score for each neighborhood, and determine their ranking.
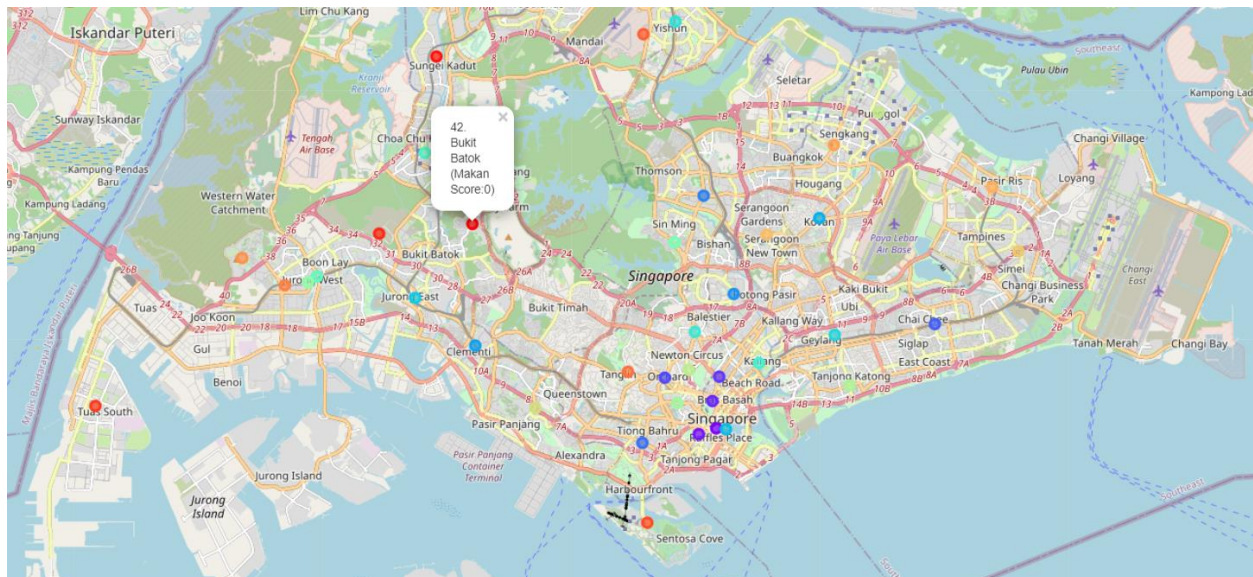
4.   Results

The following table shows the Makan scores for the individual venues:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	ID	Venue	Venue Latitude	Venue Longitude	Category	Rating	Tip Count	Likes	Makan
0	Ang Mo Kio	1.3691	103.8454	4c657dabe2ee9c749ce32d2f	Ang Mo Kio Central Market & Food Centre	1.372037	103.846366	Food Court	0.764045	0.364865	0.107492	6.606489
1	Ang Mo Kio	1.3691	103.8454	4d85cc1d509137042a3b9e5b	AMK 722 Food House	1.371989	103.847495	Food Court	0.640449	0.036036	0.011944	5.468544
2	Ang Mo Kio	1.3691	103.8454	4b691f1ff964a520ba992be3	Food Loft	1.372554	103.847913	Food Court	0.550562	0.067568	0.005429	4.701023
3	Ang Mo Kio	1.3691	103.8454	4bc08969b492d13a345ba460	AMK Food House	1.368245	103.844001	Coffee Shop	0.617978	0.031532	0.002172	5.275926
4	Ang Mo Kio	1.3691	103.8454	4b149e47f964a52023a523e3	S11 Food House	1.371501	103.847783	Food Court	0.640449	0.184685	0.060803	5.491593

The following table shows the top ten neighborhoods based on the aggregate makan score:

	Neighborhood	Latitude	Longitude	Makan
0	Singapore River	1.2870	103.8500	233.632583
1	Outram	1.2849	103.8439	230.998949
2	Museum	1.2966	103.8485	219.333125
3	Rochor	1.3051	103.8509	218.745046
4	Orchard	1.3048	103.8318	212.049354
5	Bedok	1.3236	103.9273	198.531635
6	Bukit Merah	1.2819	103.8239	195.687392
7	Ang Mo Kio	1.3691	103.8454	186.343704
8	Toa Payoh	1.3343	103.8563	178.139644
9	Clementi	1.3162	103.7649	174.715102

I created a map to visualize the neighborhoods and provide information on their ranking and Makan scores. The colors of the labels following the visible light spectrum, and are in order of their ranking, with purple being the highest ranked and red being the lowest ranked. When the user clicks on the label, the name, rank and Makan score of the neighborhood is displayed.

5.   Discussion

Based on the results, we see that the best places for food are mostly in the Downtown area at the southern end of the island. In fact, the top five locations, namely Singapore River, Outram, Museum, Rochor and Orchard, are all located around the same area in Downtown. The Downtown area is without a doubt the best place for both local and tourist foodies. However, granted that the location is at the heart of the city, the housing prices and hotel rates are much higher in the area. Instead of staying there, individuals could choose to visit the area for a one-stop food trial instead.

A major limitation of the project is the lack of details on the venues from the Foursquare API. Of the 874 venues, 186 of them have no rating, 412 of them have no tips and 361 of them have no likes. Data wrangling reduces the numbers to 180, 33 and 43 respectively, but this is not the most accurate as the information is extrapolated from other venues. The project will therefore benefit from more extensive information, and can perhaps be better applied to other regions in the world where Foursquare is more widely used,

5.   Conclusion

In this project, I used data analysis to study and determine the best neighborhoods in Singapore for food. Other projects that can be considered as a further extension include:

- Analyzing the best neighborhood to visit for different categories of food.
- Ranking the food venues within each neighborhood and visualizing the data to produce localized food maps.
- Clustering and classifying neighborhoods based on the types of food offerings available within each neighborhood.