STUDENTS ROOM RENTAL PRICES IN SINGAPORE

[Document subtitle]

ABSTRACT

A full report consisting of all of the following components: Introduction where I discuss the business problem and who would be interested in this project. Data where I describe the data that will be used to solve the problem and the source of the data. Methodology section which represents the main component of the report where I discuss and describe any exploratory data analysis that I did, any inferential statistical testing that I performed, if any, and what machine learnings I used and why. Results section where I discuss the results. Discussion section where I discuss any observations I noted and any recommendations I can make based on the results. Conclusion section where I conclude the report.

Boukhalkhal, Mehdi

Capstone Project

Contents

Introduction:

The reputation of Singapore is well established. Boiling, lively and multicultural, it also benefits from an ideal environment, planted on the South East Asia region, making it one of the most attractive cities in the world.

We can wonder about what makes a country special, which makes it worth a detour. For Singapore, it is the art of living, but also so many elements that have known how to create, over the years, an exciting country: its cultural and artistic life, its cuisine, its architecture, its history, etc. Singapore is a country with a unique, pleasant and magnificent style, distinguished and elegant. It is also a growing country, which develops and builds.

However, Singapore is also connected with situation that market is highly competitive and cost of leaving is one of the highest in the region. In fact, fresh graduates and some international students who don’t own a house in Singapore face the problem of high rental and housing prices.

Data Description:

We need to explore, segment and cluster the neighborhood in Singapore. Source of our data will be:

1- Information about neighborhoods in Singapore provided by Singapore Government (Open source rental price index)
2- The geospatial .json file that divides the Singapore into several towns that aids the choropleth plot.
3- Foursquare API to search and explore available restaurant and supermarket in a town. The count of supermarket and count of affordable meal option will be used to determine a town is favorable.

Methodology

The data obtained from open source is saved in my Github repository for reference purposes.

The rental data contains quarter, town, flat type and median rent information of each town in Singapore.

	quarter	town	flat_type	median_rent
0	2005-Q2	ANG MO KIO	1-RM	na
1	2005-Q2	ANG MO KIO	2-RM	na
2	2005-Q2	ANG MO KIO	3-RM	800
3	2005-Q2	ANG MO KIO	4-RM	950
4	2005-Q2	ANG MO KIO	5-RM	-

After loading the dataset in and from the head of dataset we do already see some missing values in the median rent column, which will require cleaning and handling NA values. These rows will be dropped as they are insignificant and do not provide useful information for our analysis.

Notice that old statistics may not be relevant in our study as the rental has increased significantly over the years. Hence, we filter data and only uses information from year 2017 onwards such that the rental price will be closer to the current market price. The non-numerical values in median rent column are removed. Now, that our dataframe is ready and clean for further processing.

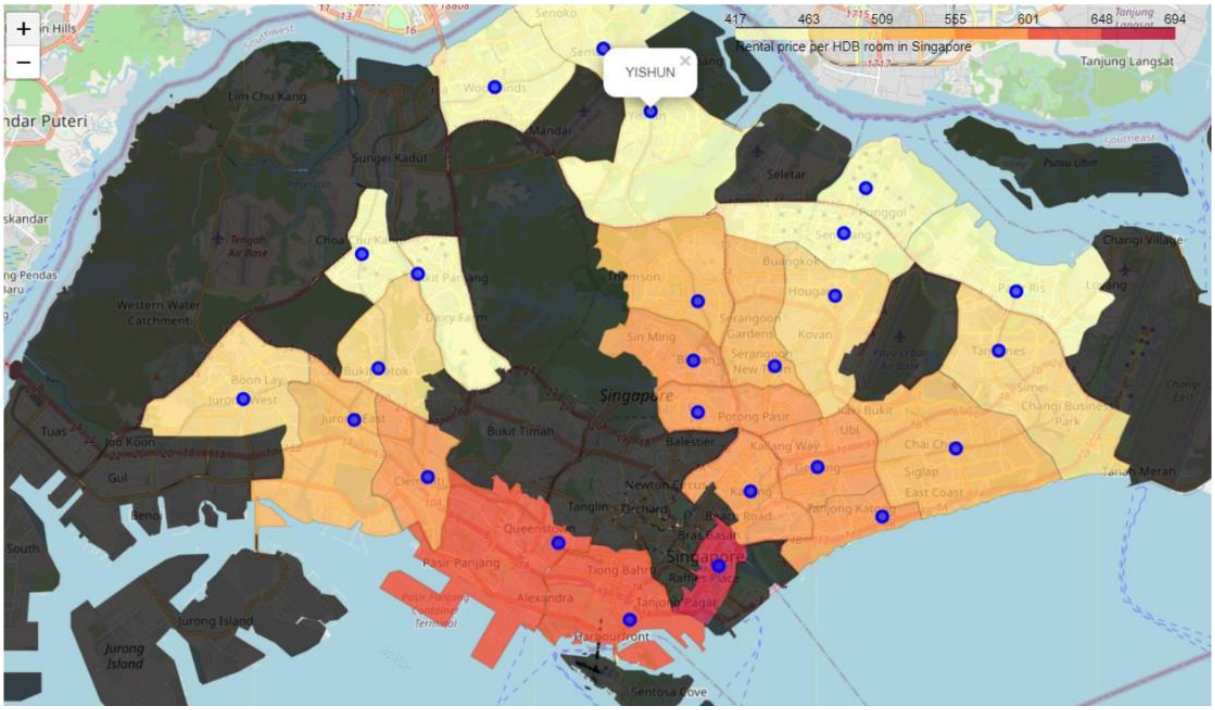
Individual students are more concerned at a single room price. The data comes with HDB flat type that suggests the number of rooms available in the flat, and this information will be used to define the average price per room in that particular flat.

flat_type	number_of_room	price_per_room
5-RM	5	448.792857
4-RM	4	524.608333
3-RM	3	580.769231
2-RM	2	741.000000

We observe that HDB flat of higher number of rooms will always result in a cheaper price. Hence, it is advisable for students to group together and rent HDB flat of more rooms.

Now, we proceed to group the room type to define the average rental price in each town. Notice, that we have missing geolocation information from the dataframe loaded from Singapore_median_rent.csv and we only have the town name as an identifier. To plot a marker on the choropeth map, we then use the geolocator library to obtain the latitude and longitude information of each town.

	town	price_per_room
0	SEMBAWANG	416.979167
1	PUNGGOL	420.366667
2	CHOA CHU KANG	423.142857
3	WOODLANDS	426.342593
4	SENGKANG	429.903846



We are now able to visualize that the highest rental price town mostly locate in the South region of Singapore. The more outskirt from the central it is, the cheaper the rental price will be.

Now we proceed to explore the restaurant and supermarket in all **25 towns**.

After utilizing the Foursquare API to explore the town, we can produce merged table of the number of supermarkets located in town.

	town	price_per_room	latitude	longitude	number_of_supermarket
0	SEMBAWANG	416.979167	1.449093	103.820055	9
1	PUNGGOL	420.366667	1.405258	103.902330	17
2	CHOA CHU KANG	423.142857	1.384749	103.744534	12
3	WOODLANDS	426.342593	1.436897	103.786216	12
4	SENGKANG	429.903846	1.391432	103.895314	20

After sorting the table with number of supermarkets, we can observe that although Sembawang is one of the cheapest towns to be resided in, the number of supermarkets surrounding suggests that perhaps Sembawang has limited accessibility to groceries shopping. Sengkang, which shares the same price range with Sembawang, may be a better alternative given its higher number of supermarkets.

	town	price_per_room	latitude	longitude	number_of_supermarket
21	KALLANG	588.162037	1.310759	103.866262	37
16	GEYLANG	558.754630	1.318186	103.887056	27
7	YISHUN	461.898148	1.429384	103.835028	26
11	TAMPINES	506.712963	1.354653	103.943571	24
24	DOWNTOWN CORE	693.680556	1.287475	103.856033	22
19	BISHAN	562.175926	1.351452	103.848250	21
8	HOUGANG	479.657407	1.371904	103.892725	21
4	SENGKANG	429.903846	1.391432	103.895314	20
10	JURONG WEST	495.444444	1.339636	103.707339	18
14	ANG MO KIO	524.078704	1.370073	103.849516	18
12	BEDOK	510.949074	1.323976	103.930216	18
1	PUNGGOL	420.366667	1.405258	103.902330	17
13	JURONG EAST	514.055556	1.333115	103.742297	16
9	BUKIT BATOK	481.259259	1.349057	103.749591	15
17	TOA PAYOH	560.652778	1.335391	103.849741	15
5	BUKIT PANJANG	432.689815	1.378629	103.762136	14
20	CLEMENTI	574.555556	1.315100	103.765231	14
18	MARINE PARADE	560.714286	1.302689	103.907395	13
3	WOODLANDS	426.342593	1.436897	103.786216	12
2	CHOA CHU KANG	423.142857	1.384749	103.744534	12
15	SERANGOON	533.217593	1.349862	103.873729	10
6	PASIR RIS	454.041667	1.373031	103.949255	9
23	BUKIT MERAH	641.784091	1.270439	103.828318	9
0	SEMBAWANG	416.979167	1.449093	103.820055	9
22	QUEENSTOWN	639.722222	1.294623	103.806045	6

Repeat the query with restaurant/food using Foursquare API. Instead of looking at the number of restaurants, we are more interested at the variety of restaurant and how affordable it is. As we have limited premium call to examine the venues, for each town we only take a sample of 10 restaurant id.

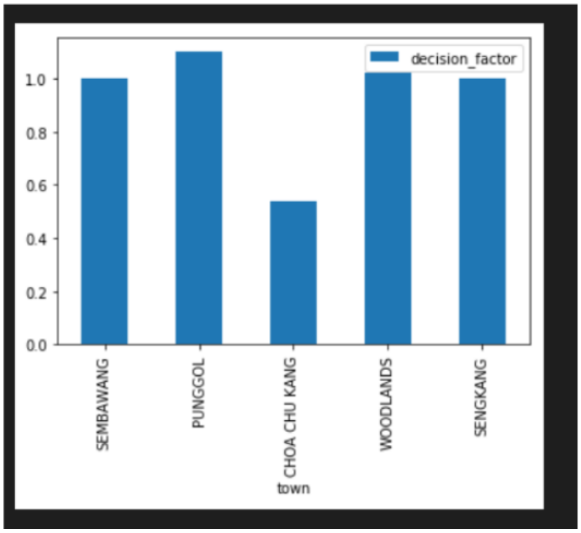
Also, due to limited premium call, we only able to examine up to 5 towns of similar rental price range, which are SEMBAWANG, PUNGGOL, CHUA CHO KANG, WOODLANDS and SENGKANG. Assuming the weightage of price, variety and rating is equal, we can determine the food worthiness of each town by taking their averages.

	town	price_per_room	latitude	longitude	number_of_supermarket	food_worthiness
0	SEMBAWANG	416.979167	1.449093	103.820055	9	2.986111
1	PUNGGOL	420.366667	1.405258	103.902330	17	2.723810
2	CHOA CHU KANG	423.142857	1.384749	103.744534	12	2.677778
3	WOODLANDS	426.342593	1.436897	103.786216	12	2.880952
4	SENGKANG	429.903846	1.391432	103.895314	20	2.566667

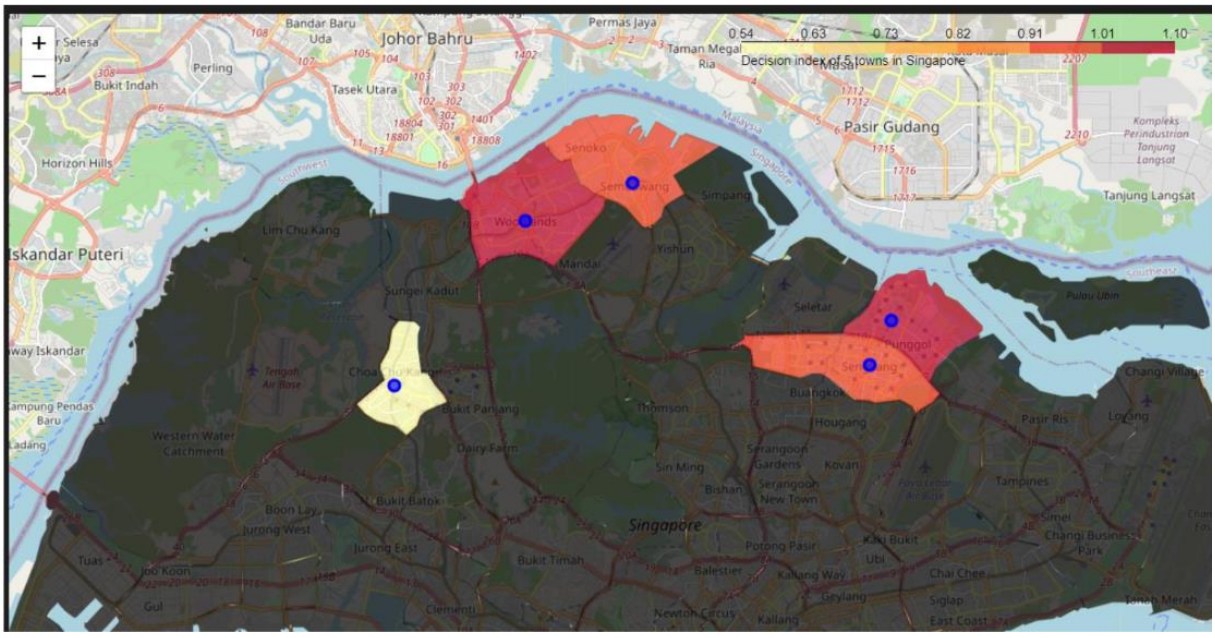
Result

We are now ready to compare the 5 towns with the number of supermarkets and food worthiness index. To generate the decision factor, we normalised both components and take the sum of it.

	town	price_per_room	latitude	longitude	number_of_supermarket	food_worthiness	decision_factor
1	PUNGGOL	420.366667	1.405258	103.902330	17	2.723810	1.101918
3	WOODLANDS	426.342593	1.436897	103.786216	12	2.880952	1.022018
0	SEMBAWANG	416.979167	1.449093	103.820055	9	2.986111	1.000000
4	SENGKANG	429.903846	1.391432	103.895314	20	2.566667	1.000000
2	CHOA CHU KANG	423.142857	1.384749	103.744534	12	2.677778	0.537628



Hence, the winner of the 5 towns who have rental price range of $415 - $430 is Punggol, who have high number of supermarket and food worthiness.



Discussion

To narrow down the result, we have chosen to only analyses 5 towns thoroughly, which are Punggol,Sengkang, Woodlands, Chua Cho Kang, and Sembawang. These 5 towns are the those who have the lowest rental price among all the towns in Singapore. If permitted by the number of premium calls from foursquare API, more towns can be studied with more data so that the analysis is more in-depth and well supported.

I have decided to use the number of supermarkets and restaurant rating with price because student who has financial constraints, these are the main factors that I would consider if I were to rent an affordable place in somewhere in Singapore. Of course, more factors can be considered such as the accessibility to public transport. Yet, in Singapore undeniably the infrastructure is already well-established, and the only concern is whether it is accessible to the student's work place. However, notice that the choice of town is highly dependent on where the student studies at or where the student is going to work at. To enable a more detailed comparison, the preferred location should be taken into account.

Conclusion

If rental price were the top consideration factor, the analysis will be narrowed down to the 5 chosen towns as performed in our methodology sections. For this reason, we can focus at analyzing the other factors to make a better decision in choosing a place to stay.