The Best Housing Estate in Singapore

Applied Data Science Capstone Project (Week 1)

Steven Sie

Table of Contents

Introduction

As the one of the most advanced country in South East Asia, Singapore has attracted talents from around the region. Malaysia is the closest country in terms of geographical location to Singapore. There is a large number of Malaysian working in Singapore and commuting between the two countries daily. Many have settled as Permanent Resident (PR) in Singapore.

Up to 80% of Singapore's population live in Housing and Development Board (HDB) flats. HDB flats are the mainstream public housing flats in Singapore. The HDB flats have an advantage over the private housing for being relatively more affordably. PRs are not permitted to buy new HDB flats in Singapore, as the HDB flats are reserved for the citizens only. PRs could only buy the HDB flats from the resale market.

Residential properties close to the Mass Rapid Transit (MRT) are popular among home buyers as MRT is the main form of transportation in the island city. Living near to MRT also means more choices of entertainment, retails, eateries and other amenities. There are HDB flats near to most of the MRT stations. Home buyers typically look at businesses and amenities around the station when decide the choice of housing.

This project sets out to assist the many PRs from Malaysia to look for ideal housing estate near to the MRT station. The main considerations in the selection of housing estate are the proximity to the MRT station, types of business near the station and the price of housing flats.

Data

The data used in the project was taken from three sources, i.e. the MRT stations information from the land transport authority of Singapore, the Foursquare location data, and the resale market price data from the housing authority of the Singapore. The analysis was conducted on Jupyter notebook using Python 3.

MRT station data

In order to analyse the businesses and amenities around the MRT stations, the location data of the stations was sourced from the official data from the authority. The train station data was downloaded from the Singapore Land Transport Authority Datamall (1) webpage. It is in Esri shapefile (SHP) format. It contains the name and the geographical information of all the MRT stations.

The .shp and .dbf files were read into the notebook. See the Diagram 1.

	stn_names	stn_codes	coordinates
0	EUNOS MRT STATION	EW7	[35782.955299999565, 33560.077600000426]
1	CHINESE GARDEN MRT STATION	EW25	[16790.746600000188, 36056.30189999938]
2	KHATIB MRT STATION	NS14	[27962.310800000094, 44352.56799999997]
3	KRANJI MRT STATION	NS7	[20081.697399999946, 45214.54790000059]
4	REDHILL MRT STATION	EW18	[26163.47800000012, 30218.819599999115]
...	...	...	...
180	TELOK BLANGAH MRT STATION	CC28	[25376.846900000237, 28138.96670000069]
181	TELOK AYER MRT STATION	DT18	[29667.621600000188, 29414.58669999987]
182	NEWTON MRT STATION	NS21	[28519.36160000041, 32735.207200000063]
183	CANBERRA MRT STATION	NS12	[27598.621771013826, 47200.24193103178]
184	WOODLANDS NORTH MRT STATION	TE1	[22699.99319899935, 47770.399814307006]

185 rows × 3 columns

Diagram 1

Foursquare location data

The next source of data was retrieved via the Foursquare API. The API was used to call the top 50 popular venues at within 400m radius of each MRT station. This was done using the 'explore' function of requesting URL. The data extracted is as shown in Diagram 2.

	stn_names	stn_names latitude	stn_names longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	EUNOS MRT STATION	1.319778	103.903252	福春叻沙卤面 Laksa & Lor Mee	1.320254	103.904132	Noodle House
1	EUNOS MRT STATION	1.319778	103.903252	GYMM BOXX	1.317912	103.901094	Gym
2	EUNOS MRT STATION	1.319778	103.903252	Mr Bean	1.319703	103.902479	Breakfast Spot
3	EUNOS MRT STATION	1.319778	103.903252	香港街珍達記	1.316703	103.901415	Chinese Restaurant
4	EUNOS MRT STATION	1.319778	103.903252	LiHO	1.319729	103.902512	Bubble Tea Shop

Diagram 2

Resale market price data

To understand the latest price of HDB housing, the median resale prices of flats was sourced from the official data from the authority. The csv file was downloaded from the Data.gov.sg (2) webpage. Generally, the town and the MRT train station have the similar name, which enable the analysis of price at various MRT stations.

The into the notebook. See the Diagram 3. It registered the resale median resale price of different types of flat in various towns and the period of sales.

	quarter	town	flat_type	price
0	2007-Q2	Ang Mo Kio	1-room	na
1	2007-Q2	Ang Mo Kio	2-room	-
2	2007-Q2	Ang Mo Kio	3-room	172000
3	2007-Q2	Ang Mo Kio	4-room	260000
4	2007-Q2	Ang Mo Kio	5-room	372000
...	...	...	...	...
7951	2019-Q4	YISHUN	2-ROOM	-
7952	2019-Q4	YISHUN	3-ROOM	265000
7953	2019-Q4	YISHUN	4-ROOM	370000
7954	2019-Q4	YISHUN	5-ROOM	470000
7955	2019-Q4	YISHUN	EXEC	565000

7956 rows × 4 columns

Diagram 3

Methodology

<text>
<text>

Result

<text>
<text>

Discussion

<text>
<text>

Conclusion

<text>
<text>

Bibliography

1. Land Transport Authority Datamall. [Online]
https://www.mytransport.sg/content/dam/datamall/datasets/Geospatial/TrainStation.zip.

2. Data.gov.sg. [Online] https://data.gov.sg/dataset/44b852a9-e7f8-4381-b896-e7c809da0f9c/download.