Prediction of Under Valued Properties in Singapore

IBM Data Science Capstone Project

Ghazi Salahuddin

12th October 2019

- Singapore, an island city-state off southern Malaysia, is a global financial center with a tropical climate and multicultural population.

- Its colonial core centers on the Padang, a cricket field since the 1830s and now flanked by grand buildings such as City Hall, with its 18 Corinthian columns.

- In Singapore's circa-1820 Chinatown stands the red-and-gold Buddha Tooth Relic Temple, said to house one of Buddha's teeth.

Area
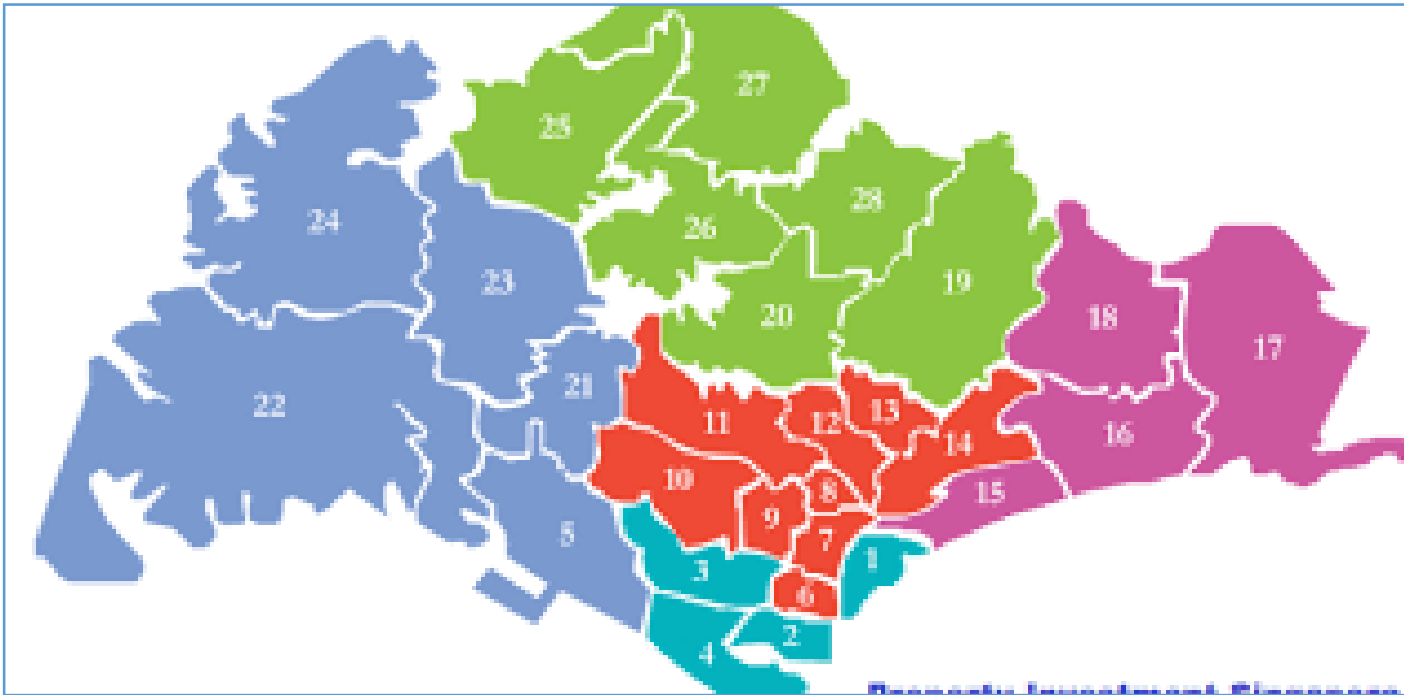- Total — 725.1 km² (280.0 sq mi)[3] (176th)

Population
- 2018 estimate — 5,638,700[4][Note 1] (113th)
- Density — 7,804/km² (20,212.3/sq mi) (2nd)

Singapore is divided into 28 Districts

Business Requirement

1. An investment firm is exploring options to invest within Singapore due to its central location in Asia.

2. Investor did not apply any area restriction so the analysis will be applied on all 28 districts of Singapore for predicting most attractive investment options.

3. The overall purpose is to predict underpriced properties based on multiple parameters and avoid over-priced venues.

Scope of Analysis and Required Data soruces

1 The property prices based on three major factors will be considered:

1 Property Tenure

2 Distance from City Area

1 Proximity with Major shopping/restaurant areas

2 Step by Step aprroach will be applied

- Gather above mentioned three factor and current housing prices

- Derive housing prices formulae based on these three factors

- Shortlist top twenty undervalued locations (outliers) for further investigation and top twenty overvalued ones for rejection.

2 3 Data Sources used for required data

1. Recent property price in various neighbors using **webscraping from websites**

2. Estates' locations and driving distances from city central area using GoogleMap API

3. Proximity with major shopping areas using FourSquare API

Data Collection – Webscrapping

condominium-sale transactions data is collected using Web scrapping from *https://www.squarefoot.com.sg/latest-transactions/sale/residential/condominium*.
In total 704 condominium-sale transactions, records were collected from this source on the date of 12th October 2019.

CONDOMINIUM

Note: Transactions in the last 1 month
Source: URA, *https://www.squarefoot.com.sg*

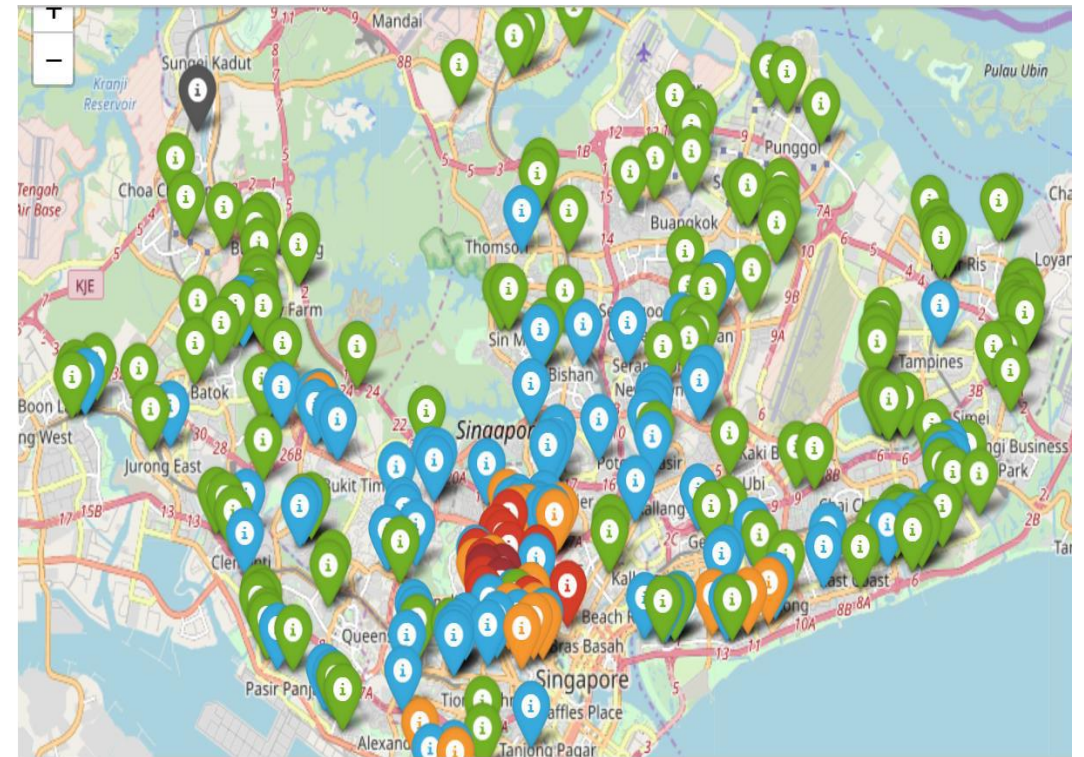| DATE	DISTRICT	PROJECT NAME	ADDRESS	PROPERTY TYPE	TENURE	AREA (SQ FT)	TYPE OF AREA	PRICE ($PSF)	PRICE ($)
6 OCT 2019	21	MAYFAIR MODERN		CONDOMINIUM	99 YRS FROM 2018	1,044	STRATA	2,165	2,261,000
6 OCT 2019	9	HAUS ON HANDY		CONDOMINIUM	99 YRS FROM 2018	743	STRATA	2,865	2,128,000
6 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	1,033	STRATA	1,328	1,372,000
6 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	1,238	STRATA	1,282	1,587,000
6 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	915	STRATA	1,271	1,163,000
6 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	592	STRATA	1,372	812,000
5 OCT 2019	28	PARC BOTANNIA		CONDOMINIUM	99 YRS FROM 2016	980	STRATA	1,401	1,372,000
5 OCT 2019	13	THE TRE VER		CONDOMINIUM	99 YRS FROM 2018	495	STRATA	1,686	835,000
5 OCT 2019	15	AMBER PARK		CONDOMINIUM	FREEHOLD	743	STRATA	2,449	1,819,260
5 OCT 2019	21	MAYFAIR MODERN		CONDOMINIUM	99 YRS FROM 2018	624	STRATA	2,162	1,350,000
5 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	1,033	STRATA	1,264	1,306,000
5 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	678	STRATA	1,377	934,000
5 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	915	STRATA	1,248	1,142,000
5 OCT 2019	18	TREASURE AT TAMPINES		CONDOMINIUM	99 YRS FROM 2018	592	STRATA	1,387	821,000

This web site provided list of last 30 days property transactions in Singapore.

Data Cleaning & Plotting

- Multiple unwanted columns are removed from scrapped data.
- Property physical locations (Lat and Long)are calculated using GoogleMaps API.
- Location are plotted on a map using Folium.
- Distance of property from central locations are calculated.

Then, locations are mapped on the map with prices highest to lowest are color coded (Dark-red, Red, Orange, Blue, Green, and Grey respectively) for visualizing.

	Project Name	Tenure	Price($psf)	Longitude	Latitude	DistanceWithCentral(km)
0	WESTCOVE CONDOMINIUM	99 Yrs FROM 1995	875.0	103.767842	1.299290	9.461205
1	WEST BAY CONDOMINIUM	99 Yrs FROM 1991	811.0	103.766799	1.300242	9.595547
2	WATERFRONT WAVES	99 Yrs FROM 2007	986.0	103.931683	1.339192	10.892026
3	WATERFRONT KEY	99 Yrs FROM 2007	1203.4	103.932551	1.338767	10.944877
4	WATERFRONT ISLE	99 Yrs FROM 2009	1212.0	103.929216	1.340098	10.726281
5	WATERFRONT GOLD	99 Yrs FROM 2009	1192.0	103.930007	1.339721	10.773228
6	WATERBANK AT DAKOTA	99 Yrs FROM 2009	1650.0	103.889377	1.308146	5.068365
7	WATER PLACE	99 Yrs FROM 1998	1332.0	103.873747	1.298078	3.003435
8	VOLARI	Freehold	2054.0	103.834276	1.316262	4.137248
9	VIZ AT HOLLAND	Freehold	1314.0	103.803880	1.309969	6.061709

Location map with color-coded psf information

Properties
PSF Price Distribution
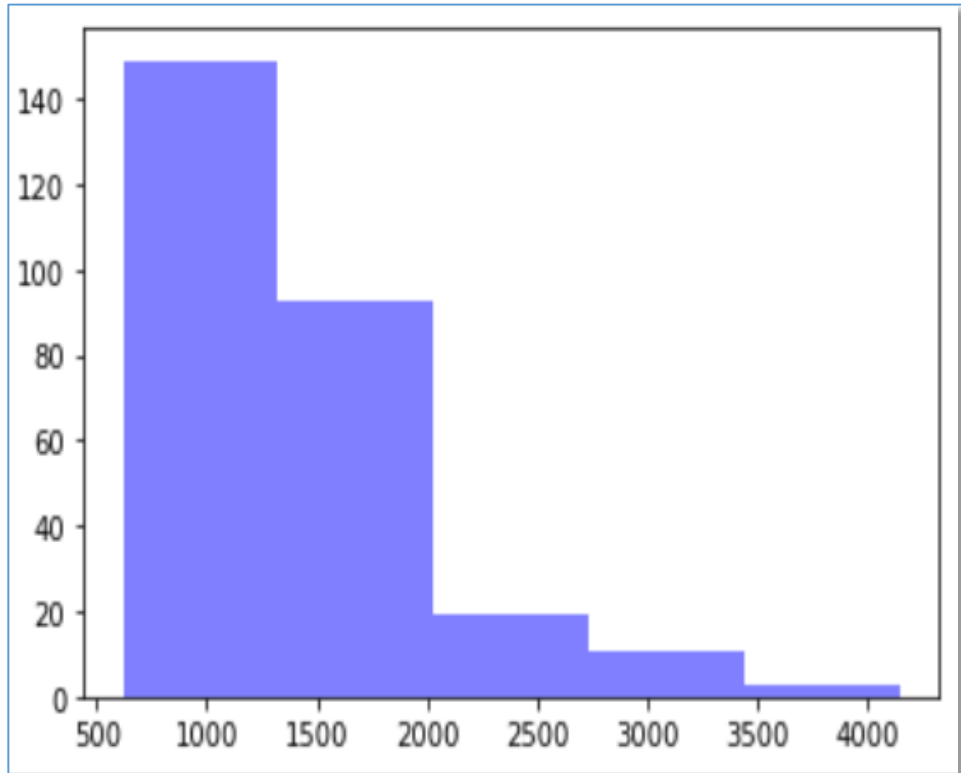
Over 300 properties with psf prices range from $500 to S$4,000.

Properties
Nearby restaurants are identified.

	DistanceWithCentral(km)	low_lease_year	Yr_From	freehold	number_of_nearby_restaurant
0	9.461205	1	1995	0	68
1	9.595547	1	1991	0	62
2	10.892026	1	2007	0	25
3	10.944877	1	2007	0	27
4	10.726281	1	2009	0	37
5	10.773228	1	2009	0	35
6	5.068365	1	2009	0	94
7	3.003435	1	1998	0	61
8	4.137248	0	2019	1	25
9	6.061709	0	2019	1	83

Multiple parameters are idetnfied to identify under valued and over values peoperties.

Distance from Center (KM)
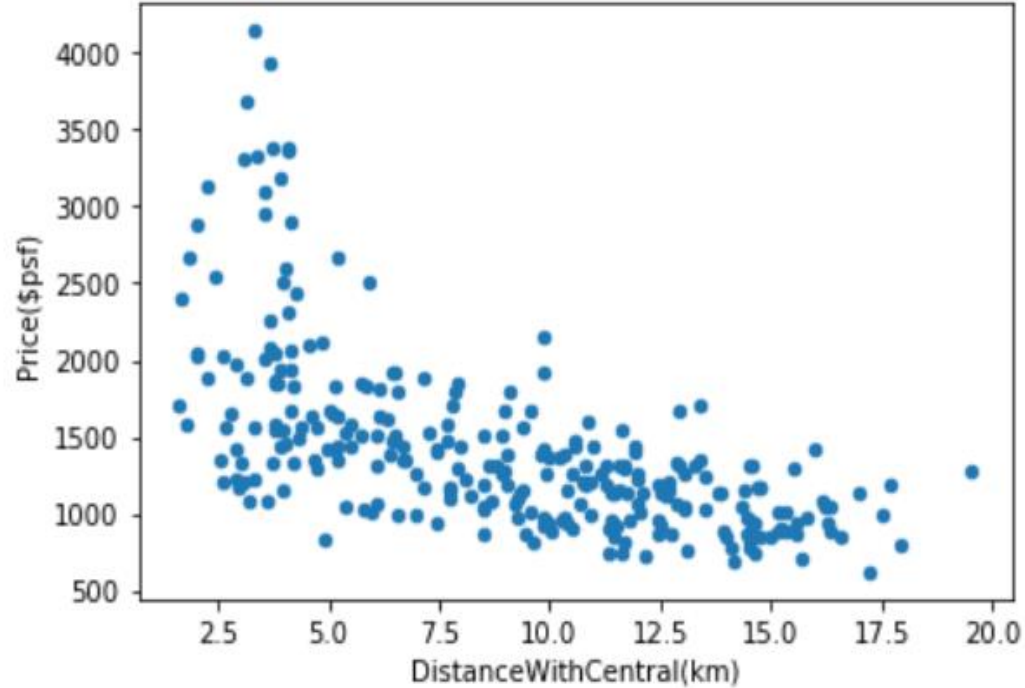
Low Lease Year
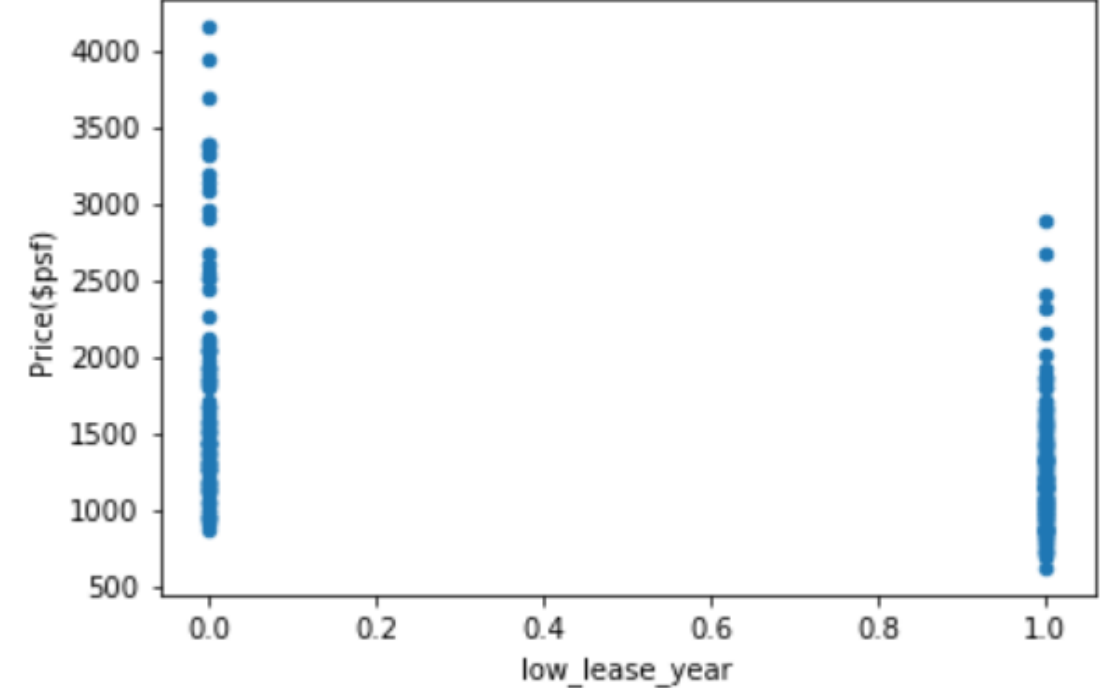
Analysis of Individual relationship between parameters

Lease Year From

Freehold Property

Lease Year From
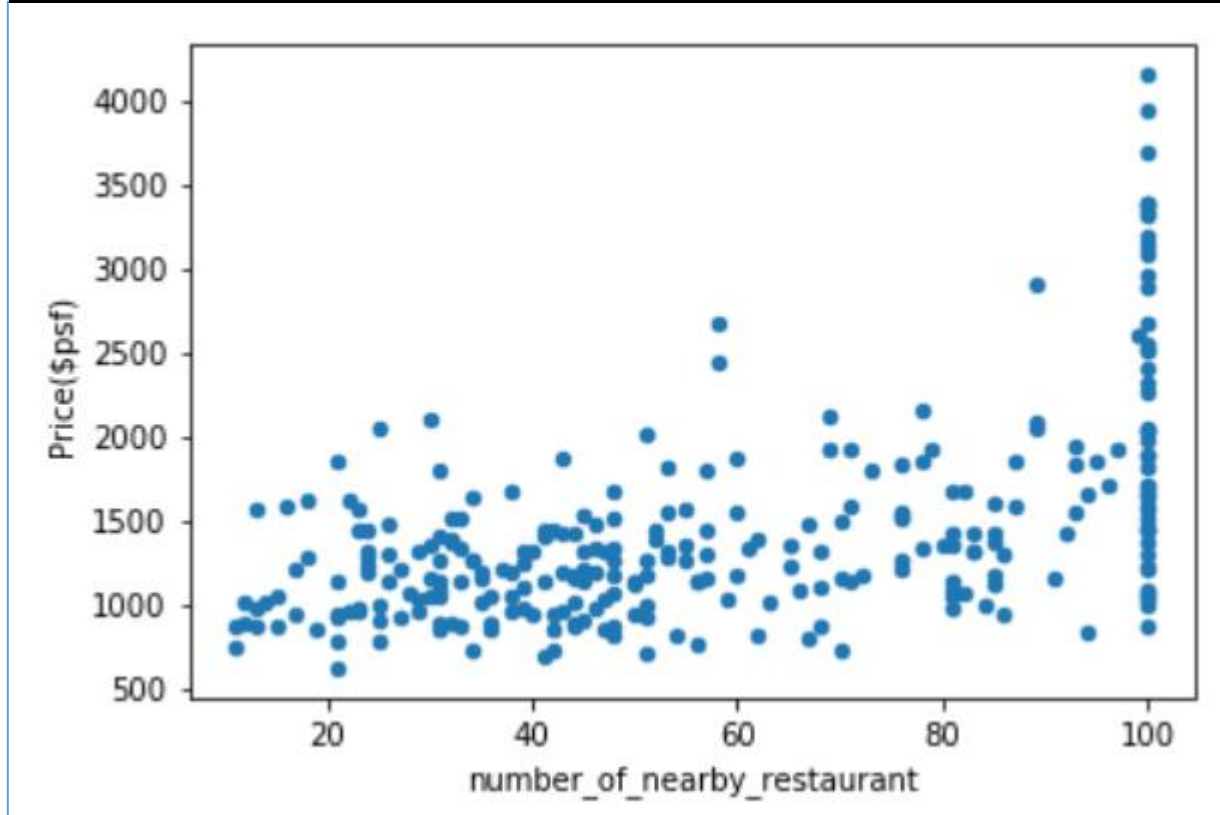
- Multiple Parameters are analyzed and price model is created to dietnify under values and overvalues properties.

- Top 20 Under Valued Properties are identified

- Top 20 Over Valued properties are identified.

- Both undervalues and overvalued are plotted on a map

Top 20 Under Valued Properties

Project Name	Project Name	Tenure	Price($psf)	Longitude	Latitude	DistanceWithCentral(km)	low_lease_year	freehold	Yr_From	number_of_nearby_restaurant	PriceDifferentPercent
KANDIS RESIDENCE	KANDIS RESIDENCE	99 Yrs FROM 2016	1278.857143	103.838741	1.459187	19.531974	1	0	2016	24	-0.553447
TWENTYONE ANGULLIA PARK	TWENTYONE ANGULLIA PARK	Freehold	4146	103.83083	1.304666	3.305955	0	1	2019	100	-0.491164
3 ORCHARD BY-THE-PARK	3 ORCHARD BY-THE-PARK	Freehold	3938	103.826118	1.304599	3.682151	0	1	2019	100	-0.469679
THE MARQ ON PATERSON HILL	THE MARQ ON PATERSON HILL	Freehold	3690	103.830257	1.301872	3.136714	0	1	2019	100	-0.425696
THE JOVELL	THE JOVELL	99 Yrs FROM 2018	1286.875	103.966701	1.362074	15.53681	1	0	2018	18	-0.402709
THE LAKEFRONT RESIDENCES	THE LAKEFRONT RESIDENCES	99 Yrs FROM 2010	1419	103.721147	1.343864	15.971925	1	0	2010	44	-0.390081
NOUVEL 18	NOUVEL 18	Freehold	3384	103.829453	1.312404	4.059904	0	1	2019	100	-0.389158
THE NASSIM	THE NASSIM	Freehold	3372	103.82388	1.307686	4.09427	0	1	2019	100	-0.38756
SYMPHONY SUITES	SYMPHONY SUITES	99 Yrs FROM 2014	1135	103.845801	1.436482	16.982261	1	0	2014	21	-0.384601
PARC ROSEWOOD	PARC ROSEWOOD	99 Yrs FROM 2011	1185	103.786742	1.429247	17.695395	1	0	2011	38	-0.375956
BISHOPSGATE RESIDENCES	BISHOPSGATE RESIDENCES	Freehold	3335	103.82586	1.299887	3.391567	0	1	2019	100	-0.368875
HILLTOPS	HILLTOPS	Freehold	3309	103.839904	1.308442	3.085894	0	1	2019	100	-0.358703
MAYFAIR MODERN	MAYFAIR MODERN	99 Yrs FROM 2018	2150.2	103.782804	1.339202	9.83515	1	0	2018	78	-0.355352
ARDMORE PARK	ARDMORE PARK	Freehold	3189	103.830589	1.311391	3.894594	0	1	2019	100	-0.348882
CUSCADEN RESERVE	CUSCADEN RESERVE	99 Years Leasehold	3392	103.825885	1.305391	3.758971	0	0	2019	100	-0.337042
DUCHESS RESIDENCES	DUCHESS RESIDENCES	999 Yrs FROM 1875	1877	103.802963	1.325245	7.124541	1	0	1875	43	-0.336368
MEYERHOUSE	MEYERHOUSE	Freehold	2673	103.895211	1.299551	5.21836	0	1	2019	58	-0.328583
HAUS ON HANDY	HAUS ON HANDY	99 Yrs FROM 2018	2877	103.847918	1.301023	2.028145	1	0	2018	100	-0.327388
THE CLAYMORE	THE CLAYMORE	Freehold	3090	103.832349	1.309124	3.576628	0	1	2019	100	-0.322214
8 SAINT THOMAS	8 SAINT THOMAS	Freehold	3140.25	103.837581	1.297821	2.239844	0	1	2019	100	-0.309038

Top 20 Over Valued properties

Project Name	Project Name	Tenure	Price($psf)	Longitude	Latitude	DistanceWithCentral(km)	low_lease_year	freehold	Yr_From	number_of_nearby_restaurant	PriceDifferentPercent
THE PEARL @ MOUNT FABER	THE PEARL @ MOUNT FABER	99 Yrs FROM 2002	1213	103.825626	1.272414	3.087742	1	0	2002	76	0.437333
SELETAR SPRINGS CONDOMINIUM	SELETAR SPRINGS CONDOMINIUM	99 Yrs FROM 1997	730	103.880424	1.388929	12.153288	1	0	1997	42	0.452201
VILLA MARINA	VILLA MARINA	99 Yrs FROM 1995	973	103.928762	1.313282	9.248241	1	0	1995	81	0.453831
GRANDEUR 8	GRANDEUR 8	99 Yrs FROM 2002	937.333333	103.848436	1.37691	10.389008	1	0	2002	86	0.473083
NOUVELLE PARK	NOUVELLE PARK	Freehold	942	103.878876	1.368738	9.964798	0	1	2019	40	0.526685
WESTCOVE CONDOMINIUM	WESTCOVE CONDOMINIUM	99 Yrs FROM 1995	875	103.767842	1.29929	9.461205	1	0	1995	68	0.529695
TANJONG RIA CONDOMINIUM	TANJONG RIA CONDOMINIUM	99 Yrs FROM 1993	1090	103.87593	1.29818	3.214245	1	0	1993	66	0.540775
HORIZON TOWERS	HORIZON TOWERS	99 Yrs FROM 1979	1211	103.834371	1.299345	2.610705	1	0	1979	100	0.54356
CHERRYHILL	CHERRYHILL	Freehold	1126	103.879992	1.351115	8.180725	0	1	2019	85	0.563604
DUNMAN VIEW	DUNMAN VIEW	99 Yrs FROM 1997	1061	103.898439	1.311275	6.107121	1	0	1997	100	0.590171
WEST BAY CONDOMINIUM	WEST BAY CONDOMINIUM	99 Yrs FROM 1991	811	103.766799	1.300242	9.595547	1	0	1991	62	0.600426
SUITES @ EASTCOAST	SUITES @ EASTCOAST	Freehold	929	103.933791	1.315858	9.873363	0	1	2019	51	0.611966
CENTRAL GROVE	CENTRAL GROVE	99 Yrs FROM 1997	1050	103.885773	1.317336	5.397597	1	0	1997	100	0.644962
SIMSVILLE	SIMSVILLE	99 Yrs FROM 1994	1032	103.890117	1.318294	5.821762	1	0	1994	100	0.648001
KERRISDALE	KERRISDALE	99 Yrs FROM 1998	1083	103.859324	1.314626	3.609284	1	0	1998	100	0.688798
TERESA VILLE	TERESA VILLE	Freehold	1168	103.82545	1.277691	2.93622	0	1	2019	72	0.705894
CITY SQUARE RESIDENCES	CITY SQUARE RESIDENCES	Freehold	1219	103.858816	1.311909	3.304211	0	1	2019	100	0.730707
LE CRESCENDO	LE CRESCENDO	Freehold	992	103.890345	1.332034	6.948355	0	1	2019	84	0.83995
MALVERN SPRINGS	MALVERN SPRINGS	Freehold	1000	103.905195	1.307356	6.574975	0	1	2019	100	0.925165
QUEENS	QUEENS	99 Yrs FROM 1998	833	103.808441	1.294263	4.929774	1	0	1998	94	1.070684

Under-Valued (Green) and Over-Valued (Red) Properties Map
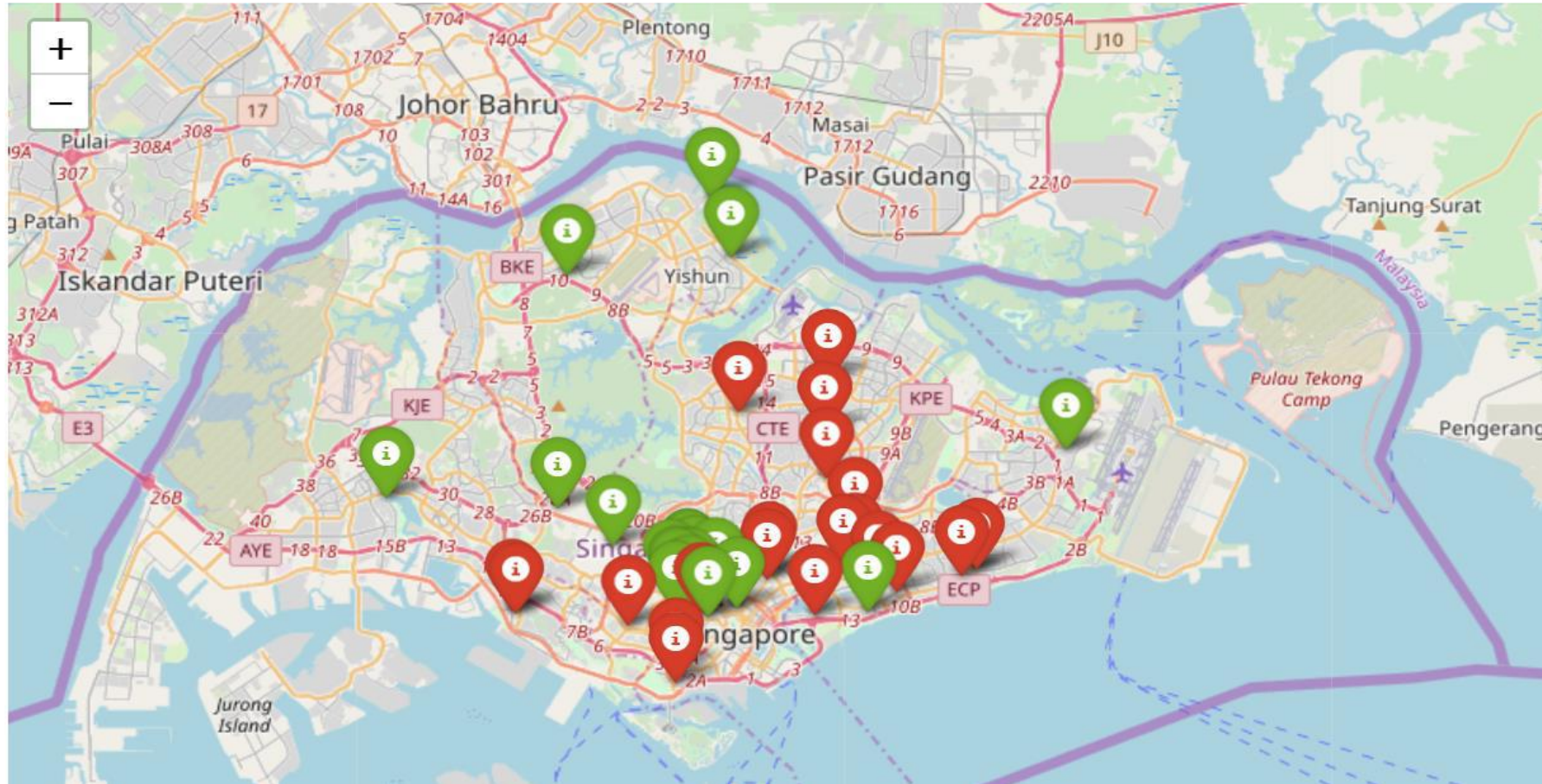
One cluster of under values properties is identified within center, while overvalued properties are found scattered across the Singapore rather than gathered in specific district.

- According to the coefficients, Distance from Central (-243.61) is the highest negative contributor and number of nearby restaurants (138.93) is the highest positive contributor.

- However, the distance parameter contributes about double than the restaurant parameter.

- Freehold type (-80.42) is more preference than lease hold type (-201.67).

- The number of remaining lease year (26.75) positively affect the price as well.