

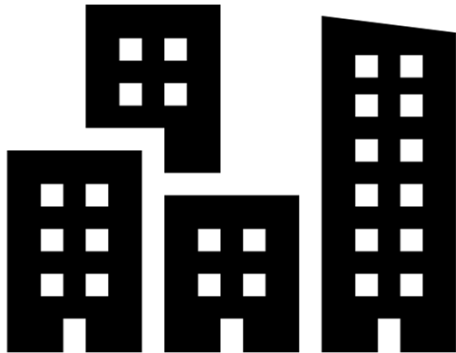


SINGAPORE HDB FLAT RESALE PRICE PREDICTION

04 MARCH 2020

BY LEE GANG

THE PROBLEM



- Property valuation reports – costly & time-consuming
- 80% of Singapore residents stays in HDB Flats (public housing apartments).
- Valuation report for resale transactions of HDB Flats can only be obtained AFTER price confirmation between Buyer and Seller.
- **Buyer Impact:** Potentially selling below appraised value → loss of earnings
- **Seller Impact:** If Buying Price > Appraised Price, can only top up difference in cash!
- **Current Situation:** Rely on property agent, compare similar property recently transacted (if any)
- **The Solution:** Get estimates of property price → **Quick, Accurate and Free?**

THE SOLUTION?

Explained **97%** of variance in the data

Average prediction error of **2.7%** of average price



**XGBOOST
PREDICTIVE MODEL**

66% of all predictions have **<4%** error.

95% of all predictions have **<9%** error.

For Nerds: { 'MAE': \$122psm, 'RMSE': \$207psm, 'R²': 0.97 }

PREDICTION ON TEN MOST RECENT TRANSACTIONS

	ACTUAL	PREDICTED										
	resale_price	pred_resale_price	price_sqm	pred_price	town	flat_type	flat_model	floor_area_sqm	street_name	month	remaining_lease	
0	426000.0	424631.0	4580.645161	4565.9233	PUNGGOL	4 ROOM	Premium Apartment	93.0	PUNGGOL CENTRAL	1.0	89.500000	
1	600000.0	597799.0	5084.745763	5066.0900	CLEMENTI	5 ROOM	Improved	118.0	CLEMENTI WEST ST 2	1.0	60.083333	
2	433888.0	431828.0	3337.600000	3321.7534	WOODLANDS	5 ROOM	Improved	130.0	WOODLANDS ST 31	1.0	71.916667	
3	375588.0	374900.0	5523.352941	5513.2334	PUNGGOL	3 ROOM	Premium Apartment	68.0	PUNGGOL DRIVE	1.0	95.166667	
4	428000.0	435935.0	4652.173913	4738.4280	YISHUN	4 ROOM	Model A	92.0	YISHUN AVE 9	1.0	94.750000	
5	710000.0	688264.0	4797.297297	4650.4340	SERANGOON	EXECUTIVE	Maisonette	148.0	SERANGOON NORTH AVE 4	1.0	71.750000	
6	570000.0	573616.0	5135.135135	5167.7150	SENGKANG	5 ROOM	Premium Apartment	111.0	ANCHORVALE CRES	1.0	94.916667	
7	850000.0	849683.0	6910.569106	6907.9927	BUKIT TIMAH	5 ROOM	Standard	123.0	FARRER RD	1.0	53.250000	
8	465000.0	466943.0	5054.347826	5075.4727	JURONG EAST	4 ROOM	Model A	92.0	TEBAN GDNS RD	1.0	92.666667	
9	420000.0	420164.0	4565.217391	4566.9976	SENGKANG	4 ROOM	Model A	92.0	FERNVALE RD	1.0	94.500000	



Data



Model



Evaluate

HOW DID WE GET THERE?

THE DATA



DATA.GOV.SG

HDB Flats Resale Transaction Data (Jan 2017- Jan 2020)

FOURSQUARE

Nearby Venues for Each HDB Resale Transaction Record
(Within 500m Radius)

THE MODELS

LASSO REGRESSION

- Multivariate Linear Regression with L1 Regularisation
- Simple and Easy to Interpret Results

EXTREME GRADIENT BOOSTED TREES (XGBOOST)

- Fast and Scalable Implementation of Gradient Boosted Decision Trees
- One of the Top Machine Learning Algorithms Used Widely in Competitions and Industry Applications



THE DIAGNOSIS

LASSO REGRESSION

- 81% variance in data explained
- \$406psm (9%) average absolute error
- 95% predictions have <30% error



XGBOOST

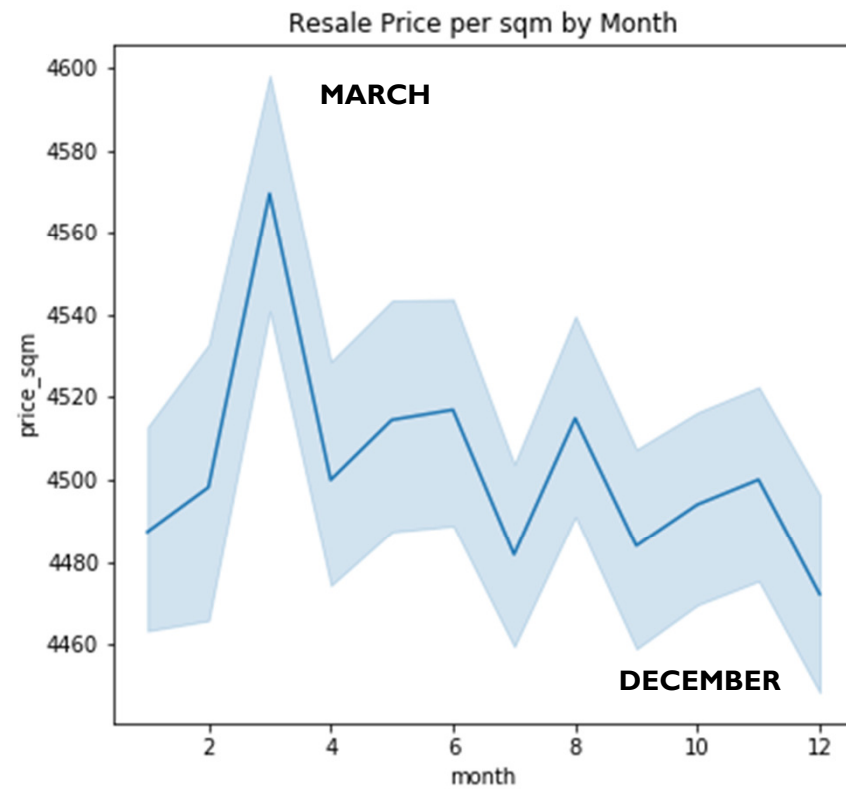
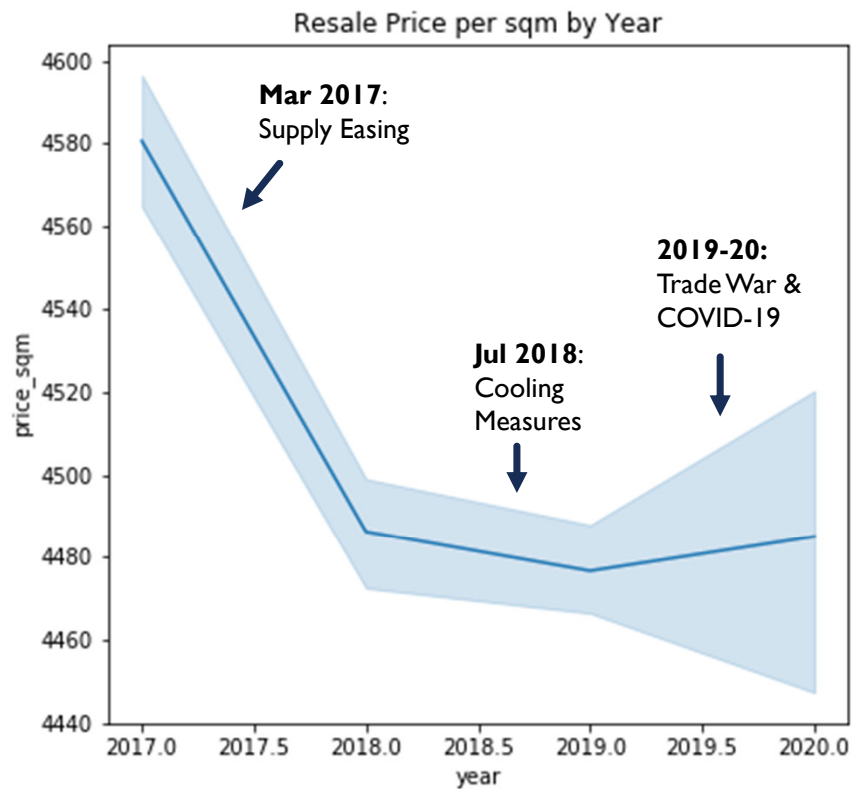
- 97% variance in data explained
- \$122psm (2.7%) average absolute error
- 95% predictions have <9% error



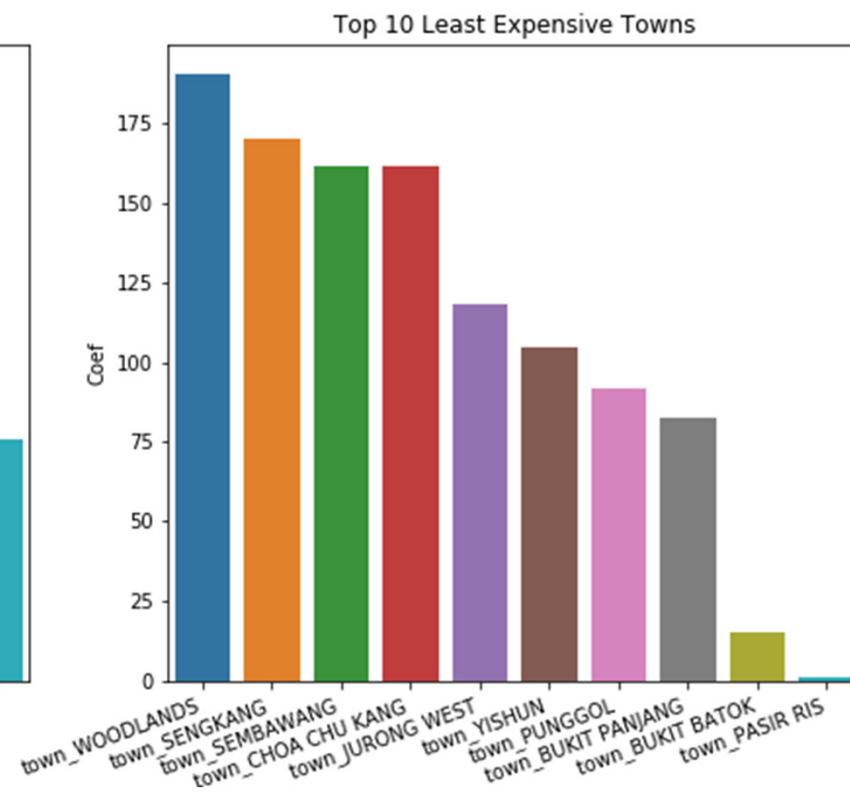
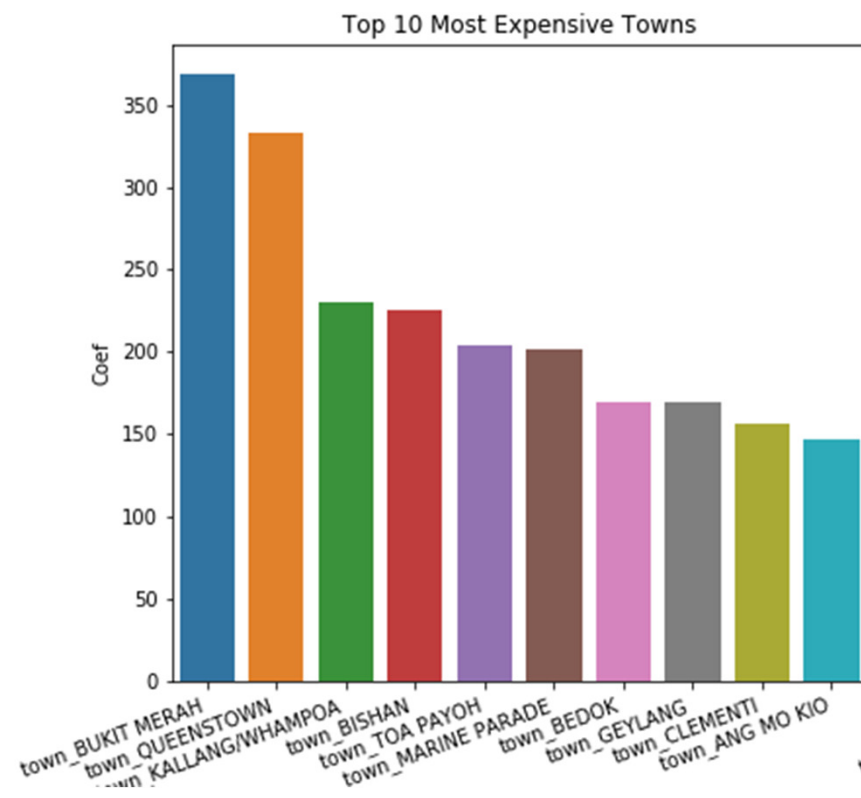


OTHER INTERESTING FINDINGS

RESALE PRICE IN DECLINE OVER THE YEARS, MARCH SEES THE HIGHEST PEAK IN PRICES.



TOP 10 MOST EXPENSIVE & LEAST EXPENSIVE LOCATIONS





THANK YOU

LEEGANGI990@GMAIL.COM

LINKS:

[FULL REPORT](#)

[JUPYTER NOTEBOOK CODES](#)