



40.011 Data Business Analytics

Spring Term 2023

Client: The Right Move SG

Forecasting Property Prices

Team #6

Chang Wei Cher, Nicklas	1005524
Tong Zhen Chung	1006097
Kwa Yu Liang	1006176
Lim Cheng Yi, Fano	1006373
Chong Melvin	1006375

Advisor: Dr. Douglas Streeter Rolph

Executive Summary

The Right Move SG (the client) is a real estate consultancy firm that aims to utilise data to make more informed decisions. Thus, this project seeks to fully leverage available data to predict and compare properties that yield a significant return on investment. This project is significant as it results in more accurate and precise consultations that significantly increases the client's reputation and broadens its clientele. Additionally, the methodology provides insight into properties that are worth investing more effort into for a deeper analysis.

Currently, the client utilises a qualitative approach to compare property pricing and data trends, relying on intuition to reach a conclusion. While this approach may serve well for now, intuition may eventually fail as the property market is chaotic and volatile in nature. Hence, a Regression Analysis approach is used to generate data-supported predictions, which are presented in a clear and approachable user interface via Excel. Since the extraction of online data is usually a labour-intensive process, web-scraping was utilised to reduce the effort needed to manually extract data from selected websites.

This project also enables the client to quantify the effect of the variables. One such discovery is that the future pricing of most properties in District 15 is minimally affected by both distance to the nearest MRT and proximity to "good" schools. Conversely, the volume of units sold in a quarter has a significant effect on future property pricing.

Nevertheless, the results should not be blindly relied on due to several limitations. Firstly, the model fitted well for some properties but poorly for others, which might indicate that the model is unable to capture the true nature of pricing of those properties. Further analysis on the variables used in the model shows that the intercept is usually the most significant factor contributing to future predictions, implying the model might have missed other significant variables.

Several points of improvements can be implemented regarding the methodology of this project. As this project limits its scope to only District 15, one suggestion is to expand the scope of the project to include other districts, to provide more properties for comparison. Another point of improvement is through the discovery and addition of potentially significant variables into the prediction model, but a caveat to this approach is that it may lead to the model satisfying only the historical data but not future predictions.

With this model as a reference, The Right Move SG will be able to further embody their aim of being a "data-driven" real estate consultancy firm, placing itself at the forefront of the property consultancy market. As for the clientele of the company, they will receive data-grounded feedback that other consultancies are not able to provide, enabling them to make well informed decisions when purchasing properties.

1. Introduction

1.1 Company

The Right Move SG (the client) is a Singapore based company which aims to offer data-driven real estate consultancy for property buyers and investors. With 13 years of combined knowledge and experience, they have sold a total of \$200M worth of properties.

1.2 Problem Definition

Predicting property market prices is challenging due to the dynamic nature of the market, influenced by various factors such as interest rates, economic conditions, supply and demand, location, and property-specific features. The COVID-19 pandemic has also impacted the market, leading to skyrocketing prices in Singapore. Making informed investment decisions requires a deep understanding of the market and the ability to factor in multiple variables.

1.3 Project Objectives

This project aims to predict property prices accurately by using past transaction records to identify factors that affect property prices. The Price per Square Feet (PSF) is a common unit of measurement in the real estate market, and this project will focus on predicting it.

1.4 Problem Scope

As there are thousands of properties in Singapore with millions of transacted records, this project's scope will be limited to properties in District 15 of Singapore that are sold under the 99-year lease. District 15 was chosen by the client as it has a high density of condominiums. It is also a highly desirable location for buyers due to its proximity to East Coast Park, excellent public transportation with the upcoming Thompson-East Coast line (TEL), renowned schools, vibrant food scene, and convenient access to the CBD and Changi Airport.

2. Methodology and Tools

2.1 Methodology Overview

The Project is broken down into four phases; data collection, data cleaning, modelling, which a multiple linear regression model is used, and lastly results presentation, presented using an excel sheet as requested by the client. A schematic diagram of the data pipeline and how the four phases are linked is provided in the *Appendix A01*.

2.2 Data collection

Together with the client, predictor variables that could be used to predict the PSF were identified. Listed below is the final list of variables that data was collected on, and their sources.

<u>Predictor/Response Variable</u>	<u>Source</u>
Price per Square Feet (PSF) [Response Variable]	Edgeprop.sg
Predictor Variables:	
Property Price Index (PPI)	Data.gov.sg
Date of transaction	Edgeprop.sg
Number of bedrooms	Edgeprop.sg
Floor level (which level the unit is on)	Edgeprop.sg
Nearest MRT station	Onemap.sg
Distance to MRT station	Onemap.sg
Primary schools within 1 km	Onemap.sg

The URA website was used initially to provide the names of properties listed in District 15 and resale transaction records for the past five years.

Edgeprop.sg is a paid web service that our client has access to, providing data past 2018 which could not be exported from the URA website. However, Edgeprop.sg and onemap.sg, do not provide any easy way to export the data, therefore web scraping was used in this part of the project to collect the data. Python was used with BeautifulSoup, Selenium, and Pandas libraries for web scraping.

2.3 Data cleaning

After data collection, three datasets existed. One contained the data for PPI, the second contained the attributes of a property and the third contained the past resale transaction data. In this phase, the three-dataset merged into a single dataset that will be used by Multiple Linear Regression model.

Firstly, certain properties had no transaction after 2017 and 2018. After clarifying with the client, these properties (namely *Katong Park Towers* and *Seaside Residences*) have since went En bloc and are no longer for sale. Since they can no longer be bought or sold, they are removed from the dataset.

Next, the transactions dataset is grouped into quarters as the PPI data is given in quarters. PSF is aggregated by taking the mean, while counting the number of transactions in that quarter gave the volume of transactions for that quarter.

However, since aggregating the number of bedrooms (for example, an average of 2.5 bedrooms sold in a quarter), and the floor height does not make any sense these variables were removed. Besides that, after grouping by quarters, certain condominiums have too little data for any meaningful analysis, and hence is removed from the dataset too (*The Red House* with 4 records, and *Seaside Park* with 7 records).

Next, one-hot encoding is used to convert categorical variables into binary:

- For the covid column, if the transacted record is between 2020 and 2022, it gets a value of 1 and 0 otherwise.
- For the school column, if the property is within 1 km of Tao Nan School or Kong Hua School, it gets a value of 1 and 0 otherwise.
- For the MRT column, if the property is within 350m (calculated based on estimated 5 minutes walking distance) of an MRT station it gets a value of 1, 0 otherwise. In

addition, if the MRT station is part of the TEL, it can only get a 1 if the transaction happened after 2014 Q4 as TEL was only announced in August 2014.

Lastly, the dataset is split into training dataset, and testing dataset (used for evaluation of the model). The most recent 20% of the transactions of each condominium/apartment is used as the testing dataset, while the oldest 80% is used for “training” the model.

2.4 Modelling

As linear regression approach is used, the problem of unit roots had to be avoided. The unit roots problem states that if absolute values are used in the regression, adjusted r^2 values will tend towards 1 even though the model does not have good predictive ability.

To address this issue, the data was transformed from absolute values into percentage change and natural logarithmic forms instead:

- PSF to $\Delta\text{PSF}(\%) = \frac{\text{PSF}_t - \text{PSF}_{t-1}}{\text{PSF}_{t-1}}$
- PPI to $\Delta\text{PPI}(\%) = \frac{\text{PPI}_t - \text{PPI}_{t-1}}{\text{PPI}_{t-1}}$
- Vol to $\ln(\text{Vol})$
- Age_Q to $\ln(\text{Age}_Q)$

A simple linear regression of involving $\Delta\text{PSF}(\%)$ against $\Delta\text{PPI}(\%)$ was first ran on all unique properties, and subsequently on District 15 using the averaged data from condominiums located in said district. This also includes condominiums that were previously filtered out due to insufficient data during the data cleaning process.

After running the regression, a common observation was that the residual plots were not randomly scattered, as shown in *A02*. As such, it was hypothesised that the predictor variables and response variables were not linearly related. Therefore, these variations in the predictor variables were investigated:

- Non-linear variables
 - $\Delta PPI(\%)^2, \ln(Vol), \ln(Age_Q)$
- Interaction terms

○ $\Delta PPI(\%) * \ln(Vol)$	○ $\Delta PPI(\%) * MRT$
○ $\Delta PPI(\%) * \ln(Age_Q)$	○ $\ln(Vol) * \ln(Age_Q)$
○ $\Delta PPI(\%) * Schools$	○ $\ln(Vol) * MRT$
○ $\Delta PPI(\%) * Covid$	○ $\ln(Age_Q) * MRT$

As the problem of unit roots was already addressed by transforming the original predictor variables to non-level forms, and the model inputs were carefully selected through consulting with the client, it is now safe to strive for high adjusted r^2 values for our models.

The best model for each condominium was selected based off the highest adjusted r^2 from 16383 combinations of equations formed from the predictor variables and its variations, with $\Delta PPI(\%)$ included in all the equations. This is based off the client's intuition that PPI is the underlying cause affecting condominiums pricing.

2.5 Analysis

This method resulted in a significant improvement in the adjusted r^2 scores as compared to the initial linear regression model that only involves $\Delta PPI(\%)$ against $\Delta PSF(\%)$, especially with Mandarin Gardens (0.18 to 0.70) and Sanctuary Green (-0.08 to 0.65). The improvement of other condominiums is attached in *A03* while respective coefficients of the best models are displayed in *A04*. Unfortunately, coefficients of predictor variables are not suitable to be used for a direct interpretation due to the inconsistencies in units. Therefore, to understand the economic significance of each predictor variable, the coefficients are multiplied by the estimated values of the variables. This shows how many units each variable contributes to a one unit increase in $\Delta PSF(\%)$ in *A05*.

For the prediction of future property prices, the coefficients and future predictor variables are inputted into the model. To estimate the value of future predictor variables, the law of expectations is used. From $E(\Delta PSF(\%)) = \alpha + B_1 E(\Delta PPI(\%)) + \dots + \varepsilon$, future predictor variables are estimated with its average value, except for the age of the condominium, as its actual age is known.

The regression model outputs the value rate of growth of property prices, $\Delta PSF(\%)$. Future pricing can be calculated by taking the current property price, PSF , and increasing it by the growth rate $\Delta PSF(\%)$. This is accomplished by the following expression:

$$PSF_t = PSF_{t-1} * (1 + \Delta PSF_{t-1}(\%))$$

Afterwards, the 95% confidence interval is constructed for the predicted PSF in order to give a range of pricing on what future pricing may take. This is accomplished by substituting the lower and upper bounds of $\Delta PSF(\%)$ into the previous equation:

$$CI = PSF_{t-1} * (1 + \Delta PSF(\%)_{t-1} \pm 1.96 * \sigma_e).$$

Lastly, relevant metrics such as profits, risk, and property score were calculated for easy apple-to-apple comparison of different condominiums, definitions as follows:

$$\mathbf{profits} = \frac{PSF_t - PSF_{latest}}{PSF_{latest}}, \quad \mathbf{risk} = \text{Width of CI} * \frac{RMSE}{\text{No. of quarters with records}}$$

$$\mathbf{property score} = \frac{\mathbf{Profits}}{\mathbf{Risk}}$$

These results are summarized in the image on the next page, with invalid results filtered out, by deeming profits above 0.5 and below -0.5 as unrealistic.



Summary of results

3. Main results

The final deliverable requested by the client was an Excel dashboard that is user friendly and has a dynamic interface shown in *Appendix A06*. The client has requested for it is to be designed so that properties to be compared at a single glance. With minimal input, the dashboard will update with relevant key information that the client wishes to inquire. It is also designed for easy future expansion in various fields.

3.1 Required input

Meetings with client revealed several crucial inputs that they desire while the following are what we implemented. First clients would need to input the year intended to forecast and which quarter of that year, shown in *Appendix A06A*. Next, there would be two drop-down menus on the main sheet for the client to select the properties within to view, as shown in *appendix A07B*.

3.2 Information displayed

After entering all the required inputs, information regarding predicted PSF, Upper and Lower bound (based on a 95% confidence interval) is displayed. Key information like nearest MRT station and its distance to said nearest MRT, indication to any specified “good” schools, profit, risk as well as property score will also be displayed. This key information is displayed side-by-side allowing the client to compare properties at one glance effectively as shown in *appendix A06C*. Additionally, the client requested for formulas used for the derivation of profits and risk

to be shown at the left side for a quick re-cap for both themselves and explaining to their customers.

3.3 Possibility of upscaling

Additionally, the worksheet titled “Forecasting” was included for the client to update more data in the future or when there are better prediction models available. The only columns that will require updating are column C and D (growth rate and RMSE). Other relevant values will be revised automatically upon revision of growth rate and RMSE, as shown in *Appendix A07*. The current model is programmed to generate up to 10 years’ worth of prediction as specified by the client, but further predictions can be easily generated by adding more quarters in the worksheet.

Lastly, the Excel is designed to allow client to update the drop-down menus to accommodate for more projects, MRT, schools, or updated volume of transaction and RMSE, as shown in *Appendix A08*. The main dashboard will automatically reflect these updates, reducing the hassle of intensive modification.

With this structure, the process of connecting the front and back-end systems is streamlined. Any manipulation of data on the back-end database stemming from selections made on the front-end is now efficient and optimised. The client can easily compare properties on an apple-to-apple basis while being able to update the database without in-depth proficiency in any software.

4. Assumptions, Limitations and Contribution

4.1 Assumptions

The first assumption was the definition of “near MRT” to be within 350m from any stations, which approximates a 5-minute walk from any property to the nearest MRT. The second assumption is when the period where the COVID-19 pandemic became significant. Due to the pandemic's uncertain start and end dates, the group decided between 2020-2022 as heavily

affected by the COVID-19 pandemic as clear timeframe is needed to investigate the relationship between Covid-affected years and the PSF growth rate. The third assumption was defining of “Good” schools to be either Tao Nan School or Kong Hwa School. These assumptions are provided by our client as their intuition states that buyers will seek out properties nearer to these schools for their child’s future primary education. In addition, future PPI growth rate was calculated based on average growth rate from historical PPI data as predicting it is infeasible for the duration of this project. Lastly, it is assumed that there will not be any erratic behaviour from property seller, which could affect short-term PSF trend.

4.2 Limitations

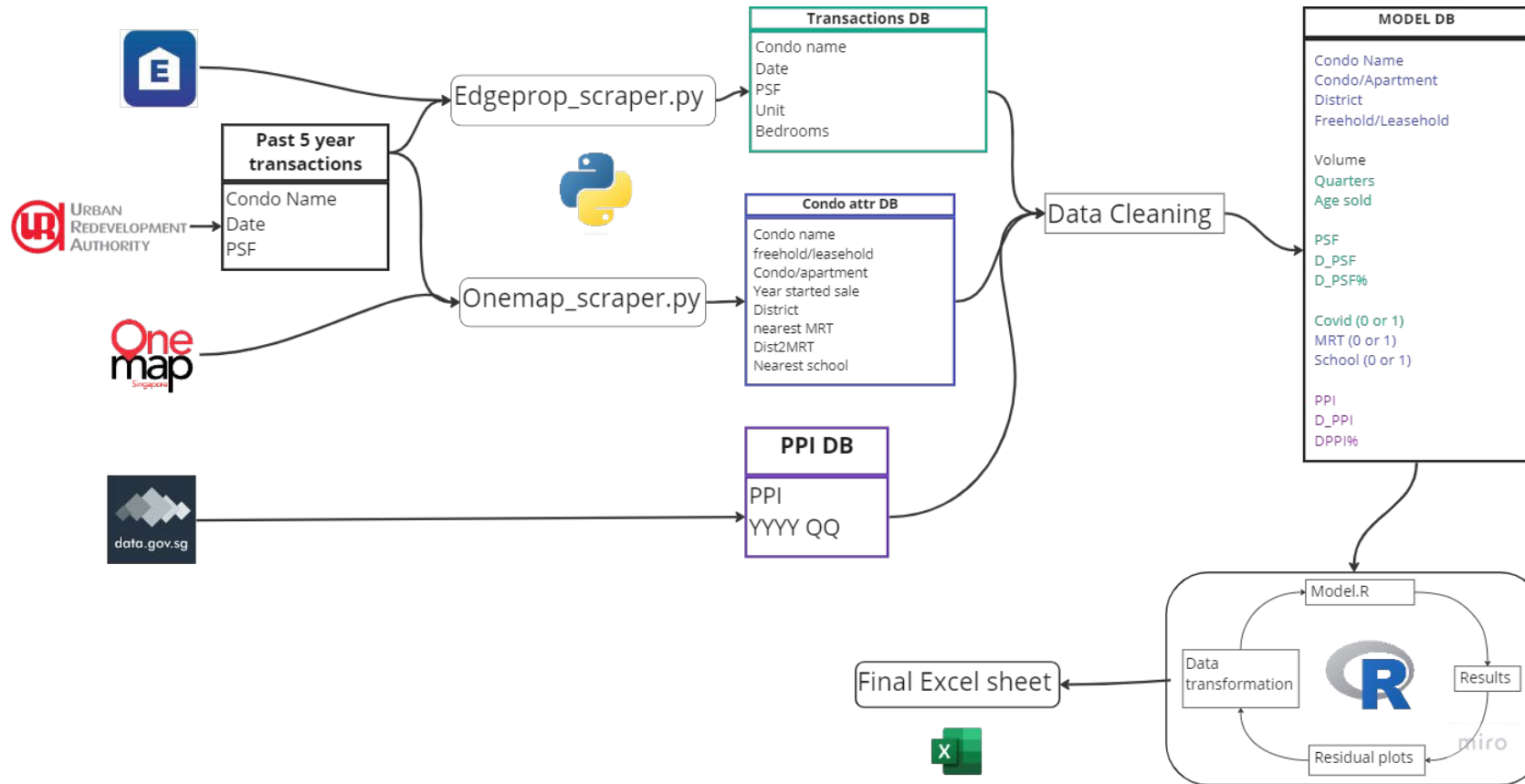
One limitation was due to the scope of this project which focuses only on District 15. Certain key variables that improve the model remains undiscovered. Another limitation was the exclusion of unit-specific predictor variables. Since PPI data was published in quarterly basis, transactions had to be grouped by quarters as well. Unit-specific predictor variables such as number of bedrooms and floor level had to be excluded because taking a mean would not be sensible. This resulted in losing potential significant variables. Furthermore, accuracy of predicted PSF is also limited by the model as a linear model was assumed to be the best fit in forecasting. Lastly, assumed best model was obtained based on best adjusted r^2 values, which may inaccurately reflect reality as high adjusted r^2 values may be due to overfitting the data. Due to the nature of the model, the true impact and success can only be seen in subsequent years when the property goes on sale.

4.3 Contribution

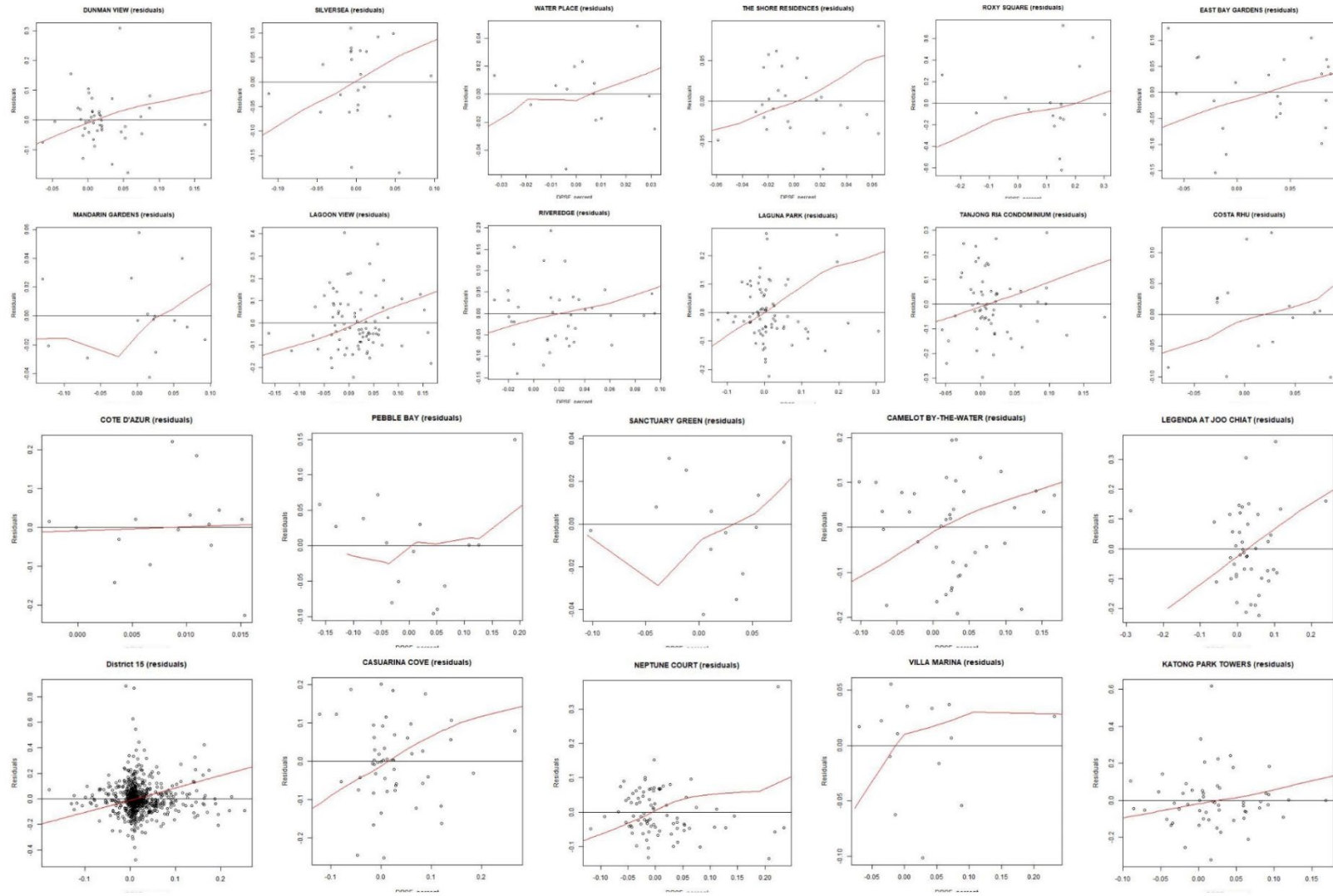
The Client can now make a data-driven forecast on PSF, instead of relying on guesswork and intuition. This would drastically improve the quality of consultation they are able to offer, distinguishing themselves from other firms. Additionally, insights of how heavily each variables impact future PSF is known, assisting them in identifying crucial traits when looking for profitable properties.

Appendix

A01: Project pipeline



A02: Final residual plots



A03: Improved adjusted r^2 values

Project Name	Adjusted R² (simple linear regression)	Adjusted R² (final)
Camelot By-The-Water	0.106503	0.146
Casuarina Cove	0.299067	0.299067
Costa Rhu	-0.02069	0.116697
Cote D'azur	-0.07438	-0.07438
Dunman View	0.170694	0.170694
East Bay Gardens	-0.04762	0.234514
Katong Park Towers	0.093518	0.093518
Lagoon View	0.111507	0.149911
Laguna Park	0.260551	0.297036
Legenda At Joo Chiat	0.185077	0.194734
Mandarin Gardens	0.181225	0.697347
Neptune Court	0.286786	0.425416
Pebble Bay	-0.03195	0.22846
Riveredge	0.024785	0.118896
Roxy Square	0.008117	0.013567
Sanctuary Green	-0.08139	0.652081
Silversea	-0.03069	0.071824
Tanjong Ria Condominium	0.088759	0.088759
The Shore Residences	0.06199	0.214428
Villa Marina	0.069411	0.307519
Water Place	-0.06842	0.051857
District 15	0.103376	0.114566

A04: Coefficients of final models (based off highest adjusted r^2 score)

Project Name	(Intercept)	Δ PPI(%)	ln(Age_Q)	MRT	Δ PPI(%)*ln(Age_Q)	Δ PPI(%)*MRT	ln(Vol)	Δ PPI(%)*ln(Vol)	Covid	Δ PPI(%)*Covid	ln(Vol)*ln(Age_Q)
Camelot By-The-Water	-0.41	38.93	0.10	-0.05	-9.34	7.78	0	0	0	0	0
Casuarina Cove	0.00	1.67	0	0	0	0	0	0	0	0	0
Costa Rhu	-0.08	1.64	0	0	0	0	0.08	-2.35	0	0	0
Cote D'azur	0.01	-0.28	0	0	0	0	0	0	0	0	0
Dunman View	0.00	1.03	0	0	0	0	0	0	0	0	0
East Bay Gardens	-0.04	2.11	0	0	0	0	0.11	-2.10	0	0	0
Katong Park Towers	0.01	0.99	0	0	0	0	0	0	0	0	0
Lagoon View	-0.02	0.91	0	0	0	0	0.04	0	0	0	0
Laguna Park	0.11	-8.78	-0.02	0	1.92	0	-0.01	0.79	0	0	0
Legenda At Joo Chiat	0.03	1.79	0	0	0	0	-0.04	0	0	0	0
Mandarin Gardens	-18.16	0.23	3.58	0	0	0	6.93	5.50	0.03	-10.83	-1.37
Neptune Court	0.93	-0.53	-0.21	0	0	0	-0.58	0.86	0	0	0.13
Pebble Bay	31.84	592.49	-6.73	0	-130.46	0	-18.29	7.74	-0.02	19.20	3.85
Riveredge	-0.02	0.28	0	0	0	0	0.04	0	0	0	0
Roxy Square	-0.56	-1.69	0.19	0	0	0	0	0	0	0	0
Sanctuary Green	-12.71	-0.80	2.76	0	0	0	3.94	0	-0.15	0	-0.84
Silversea	-0.01	-9.44	0	0	0	0	0.00	6.02	0	0	0
Tanjong Ria Condominium	0.00	1.11	0	0	0	0	0	0	0	0	0
The Shore Residences	-0.12	-70.84	0.03	0	19.47	0	0	0	0.03	-7.21	0
Villa Marina	-10.08	454.99	2.19	0	-98.87	0	11.45	3.81	0.06	10.91	-2.47
Water Place	-1.40	-3.26	0.33	0	0	0	-0.06	2.74	0	0	0
District 15	0.02	-0.86	0.00	0.00	0.34	-0.90	0.00	0.34	0	0	0

A05: Economic significance

Project Name	Intercept	ΔPPI(%)	ln(Age_Q)	MRT	ΔPPI(%)*ln(Age_Q)	ΔPPI(%)*MRT	ln(Vol)	ΔPPI(%)*ln(Vol)	Covid	ΔPPI(%)*Covid	ln(Vol)*ln(Age_Q)
Camelot By-The-Water	-5.20	10.06	6.20	-0.60	-11.47	2.01	0	0	0	0	0
Casuarina Cove	0	1.02	0	0	0	0	0	0	0	0	0
Costa Rhu	-2.96	0.88	0	0	0	0	5.53	-2.46	0	0	0
Cote D'azur	1.55	-0.55	0	0	0	0	0	0	0	0	0
Dunman View	0	0.95	0	0	0	0	0	0	0	0	0
East Bay Gardens	-0.70	0.90	0	0	0	0	1.39	-0.59	0	0	0
Katong Park Towers	0.50	0	0	0	0	0	0	0	0	0	0
Lagoon View	-0.85	0	0	0	0	0	1.60	0	0	0	0
Laguna Park	-35.37	13.77	37.20	0	-15.88	0	3.10	-1.83	0	0	0
Legenda At Joo Chiat	1.25	1.27	0	0	0	0	-1.51	0	0	0	0
Mandarin Gardens	-84.27	0	85.52	0	0	0	55.58	0.70	0	0	-56.55
Neptune Court	49.39	0	-57.50	0	0	0	-55.14	0.58	0	0	63.87
Pebble Bay	-59.69	-16.82	60.88	0	17.88	0	52.17	0	0	0	-53.09
Riveredge	-0.60	0	0	0	0	0	1.51	0	0	0	0
Roxy Square	-1.74	0	2.81	0	0	0	0	0	0	0	0
Sanctuary Green	-49.38	0	50.68	0	0	0	26.82	0	0	0	-27.07
Silversea	0.75	9.29	0	0	0	0	0	-8.89	0	0	0
Tanjong Ria Condominium	0	0.83	0	0	0	0	0	0	0	0	0
The Shore Residences	-2.02	-8.41	1.86	0	9.57	0	0	0	0	0	0
Villa Marina	21.89	-16.01	-22.80	0	16.68	0	-40.28	0	0	0	41.74
Water Place	-29.64	-1.06	32.25	0	0	0	-1.95	1.41	0	0	0
District 15	1.92	-0.72	-1.89	0	1.37	-0.75	0	0	0	0	0

A06: Excel dashboard interface

Year to forecast: 2026
Q to forecast: 2

Select Condo 1: EAST BAY GARDENS
 Lower Bound: \$1,814.44 | Predicted PSF: \$1,972.97 | Upper Bound: \$2,353.05
 Distance to MRT: 332 m
 Near School (1km): No
 Profit: 1.0343
 Risk: 0.7680
 Profit/Risk: 1.3467

Select Condo 2: COTE D'AZUR
 Lower Bound: \$1,602.07 | Predicted PSF: \$1,822.83 | Upper Bound: \$2,077.57
 Distance to MRT: 302 m
 Near School (1km): Tao Nan
 Profit: 0.1282
 Risk: 0.1479
 Profit/Risk: 0.8667

A: Inputting of year and quarter to forecast

B: Two drop-box for easily selection of property

C: Dynamic cells displaying key information

A07: Scalable backend worksheet to predict future PSF, UB and LB

Project Name	Quarters from 2023Q1	R hat (growth rate)	RMSE	Predicted PSF	UB	LB
CAMELOT BY-THE-WATER	27	0.064095641205038	=Update!\$E\$2	=E28*(1+C29)	=E29*(1+C29+(1.96*D29))	=E29*(1+C29-(1.96*D29))
CAMELOT BY-THE-WATER	28	0.0631214445169035	=Update!\$E\$2	=E29*(1+C30)	=E30*(1+C30+(1.96*D30))	=E30*(1+C30-(1.96*D30))
CAMELOT BY-THE-WATER	29	0.0621540367110036	=Update!\$E\$2	=E30*(1+C31)	=E31*(1+C31+(1.96*D31))	=E31*(1+C31-(1.96*D31))
CAMELOT BY-THE-WATER	30	0.061193323822438	=Update!\$E\$2	=E31*(1+C32)	=E32*(1+C32+(1.96*D32))	=E32*(1+C32-(1.96*D32))
CAMELOT BY-THE-WATER	31	0.0602392138237566	=Update!\$E\$2	=E32*(1+C33)	=E33*(1+C33+(1.96*D33))	=E33*(1+C33-(1.96*D33))
CAMELOT BY-THE-WATER	32	0.0592916165720614	=Update!\$E\$2	=E33*(1+C34)	=E34*(1+C34+(1.96*D34))	=E34*(1+C34-(1.96*D34))
CAMELOT BY-THE-WATER	33	0.0583504437578961	=Update!\$E\$2	=E34*(1+C35)	=E35*(1+C35+(1.96*D35))	=E35*(1+C35-(1.96*D35))
CAMELOT BY-THE-WATER	34	0.0574156088558628	=Update!\$E\$2	=E35*(1+C36)	=E36*(1+C36+(1.96*D36))	=E36*(1+C36-(1.96*D36))
CAMELOT BY-THE-WATER	35	0.0564870270768845	=Update!\$E\$2	=E36*(1+C37)	=E37*(1+C37+(1.96*D37))	=E37*(1+C37-(1.96*D37))
CAMELOT BY-THE-WATER	36	0.0555646153220564	=Update!\$E\$2	=E37*(1+C38)	=E38*(1+C38+(1.96*D38))	=E38*(1+C38-(1.96*D38))
CAMELOT BY-THE-WATER	37	0.0546482921380216	=Update!\$E\$2	=E38*(1+C39)	=E39*(1+C39+(1.96*D39))	=E39*(1+C39-(1.96*D39))
CAMELOT BY-THE-WATER	38	0.0537379776738072	=Update!\$E\$2	=E39*(1+C40)	=E40*(1+C40+(1.96*D40))	=E40*(1+C40-(1.96*D40))
CAMELOT BY-THE-WATER	39	0.0528335936390682	=Update!\$E\$2	=E40*(1+C41)	=E41*(1+C41+(1.96*D41))	=E41*(1+C41-(1.96*D41))
CAMELOT BY-THE-WATER	40	0.0519350632636817	=Update!\$E\$2	=E41*(1+C42)	=E42*(1+C42+(1.96*D42))	=E42*(1+C42-(1.96*D42))
CAMELOT BY-THE-WATER	41	0.0510423112586413	=Update!\$E\$2	=E42*(1+C43)	=E43*(1+C43+(1.96*D43))	=E43*(1+C43-(1.96*D43))
CAMELOT BY-THE-WATER	42	0.050155263778199	=Update!\$E\$2	=E43*(1+C44)	=E44*(1+C44+(1.96*D44))	=E44*(1+C44-(1.96*D44))
CAMELOT BY-THE-WATER	43	0.0492738483832134	=Update!\$E\$2	=E44*(1+C45)	=E45*(1+C45+(1.96*D45))	=E45*(1+C45-(1.96*D45))
CASUARINA COVE	0	=Update!D6	=Update!\$E\$5	=Update!C6	=E46*(1+C46+(1.96*D46))	=E46*(1+C46-(1.96*D46))
CASUARINA COVE	1	0.0365221179999584	=Update!\$E\$5	=E46*(1+C47)	=E47*(1+C47+(1.96*D47))	=E47*(1+C47-(1.96*D47))
CASUARINA COVE	2	0.0365221179999584	=Update!\$E\$5	=E47*(1+C48)	=E48*(1+C48+(1.96*D48))	=E48*(1+C48-(1.96*D48))
CASUARINA COVE	3	0.0365221179999584	=Update!\$E\$5	=E48*(1+C49)	=E49*(1+C49+(1.96*D49))	=E49*(1+C49-(1.96*D49))
CASUARINA COVE	4	0.0365221179999584	=Update!\$E\$5	=E49*(1+C50)	=E50*(1+C50+(1.96*D50))	=E50*(1+C50-(1.96*D50))
CASUARINA COVE	5	0.0365221179999584	=Update!\$E\$5	=E50*(1+C51)	=E51*(1+C51+(1.96*D51))	=E51*(1+C51-(1.96*D51))
CASUARINA COVE	6	0.0365221179999584	=Update!\$E\$5	=E51*(1+C52)	=E52*(1+C52+(1.96*D52))	=E52*(1+C52-(1.96*D52))
CASUARINA COVE	7	0.0365221179999584	=Update!\$E\$5	=E52*(1+C53)	=E53*(1+C53+(1.96*D53))	=E53*(1+C53-(1.96*D53))
CASUARINA COVE	8	0.0365221179999584	=Update!\$E\$5	=E53*(1+C54)	=E54*(1+C54+(1.96*D54))	=E54*(1+C54-(1.96*D54))

A08: Backend worksheet for dashboard information and calculation

1
 2
 3
 4
 5

A	B	C	D	E	F
Project Name	MRT dist	Schools	Volume of Transaction	RMSE	
CAMELOT BY-THE-WATER	207 m		0	124	0.198368
CASUARINA COVE	107 m		0	155	0.058404
COSTA RHU	522 m		0	139	0.08541
COTE D'AZUR	302 m	Tao Nan		109	0.033903
DUNMAN VIEW	957 m	Kong Hwa Primary School		190	0.037612
EAST BAY GARDENS	332 m		0	56	0.079846
LAGOON VIEW	225 m		0	310	0.108972
LAGUNA PARK	124 m		0	414	0.099567
LEGENDA AT JOO CHIAT	802 m		0	126	0.064895
MANDARIN GARDENS	597 m		0	107	0.201354
NEPTUNE COURT	736 m		0	646	0.087579
PEBBLE BAY	148 m		0	87	0.071546
RIVEREDGE	581 m		0	134	0.191121
ROXY SQUARE	231 m		0	25	0.058245
SANCTUARY GREEN	249 m		0	98	0.360822
SILVERSEA	480 m	Tao Nan		130	0.049812
TANJONG RIA CONDOMINIUM	108 m		0	223	0.111199
THE SHORE RESIDENCES	534 m	Tao Nan		155	0.051132
VILLA MARINA	269 m		0	91	0.039165
WATER PLACE	139 m		0	88	0.11381
					0.048506

6

Final
Forecasting
Update
Dropbox
+