

School of Chemical and Biomedical Engineering



CB4094

**Can you afford a house?
Let me guide you!**

Name/Matric No. : O Kah Hao (U2022308E)

Yong Zane (U2020123B)

Group : B22/B, Group 14

1. Background

For most, buying a home represents a significant milestone in life. However, in Singapore there are many legal steps and processes associated with purchasing a property which sometimes may be daunting to the uninitiated. Before making any decision to buy a home, to avoid wasting time on doing research on the houses you are not able to buy you should first check on the eligibility to buy that specific property.

For our case, we are assuming that there are a pair of fresh-graduate Bioengineering couples holding Singapore citizenship who are planning to buy their dream house and start a family. Nevertheless, they are unsure about the current market price for resale flats and they are worried that they are not able to pay off their loan payment as housing prices can be unaffordable for them to buy another house in Singapore. Moreover, in Singapore, there is a Minimum Occupation Period which is a period of time where you are required to physically occupy that flat before you sell it in an open market. This couple is not willing to spend any unnecessary cash and increase their financial burden.

This couple are still quite busy with their internship and both of them do not have any extra time to go around searching for a flat. But from what they know, they decided to apply for a resale flat because it is more straightforward and they could not bear to wait for another five years for a Build-To-Order. After interviewing them, we know that both of them are working in Geylang with a starting salary of 3500sgd each. We also know that one of their parents stays in Kallang which is a town located beside Geylang. Since they are planning to have a kid, a 3 room flat type would be the most suitable for them

After doing some research on houses in Geylang, they have a rough idea of their ideal flat but with their financial condition they wish to get a 3 room flat type with similar floor area and remaining lease but with a cheaper price. Hence, with the condition stated above, we are about to find the best resale flat with the data we have and build a regression model to predict whether the expected resale flat they want will have a cheaper price. With the data we had would we be able to find the same type of resale flat with a better price? What if the couple may not be able to make the purchase this year in 2022, can we predict the price for resale flat price of 2023, and how reliable is our prediction?

2. Objectives

1. To predict the resale price of the same type of flat.
2. To predict future prices of the resale flat within a specific town in 2022 and beyond.
3. Determine if the predicted sales price is affordable for the couple.

3. Setup/Techniques used

The methodology to solving the problem would be to analyse the variables of interest to the problem. Under “importing relevant datasets”, the variables were selected and stored in a new dataframe which would be used for analysis. [cell 2]

The new dataframe, trial, is unsuitable for analysis so the variable data types were extracted and changed. Variables that were not needed for our analysis were dropped from the dataframe to have a cleaner dataset. From the data set we had, we changed the remaining lease (object) into totalMonths (float) which is a numerical data set allowing us to draw a conclusion in a later stage using linear regression with other numerical data. This can be seen in “Manipulating the dataset of interest”. [cell 5-6]

To narrow the scope of the study to a specific town, the distribution of flats in each town were shown to give the user a better idea of where resale flats are available and this gives them an opportunity to select the town they desire which is done in “Town selection and filter”. [cell 7-8]

Under “Creating cat plots”, various count plots for flat type, flat model and a heat map of flat type and flat model were shown to give users a better visualisation of the type of resale flats available in that town. This sets an expectation for the users and helps to quickly identify the probability that that flat type is more available than the others and if their desired flat type is not available, they can choose other flat types or look into other towns. [cell 10-11]

To determine the affordability of the house and to predict its future price, the potential predictors (floor area and total months) for resale price were identified by looking at the covariance of the predictors against resale price. A high covariance (>0.6) between the sale price and the predictors determines that both will be useful in predicting sales price hence a multivariate regression was used to predict the sales price. [cell 22]

To create the regression model, a dataset with just the predictors and the response (sales price) was chosen and this dataset was randomly split into test and train sets which were used to determine the model’s usefulness in predicting the price. The explained variance of the regression model was rather high (>0.8) for the test and train set which determined it was very

useful and reliable in predicting sales price in that town. Furthermore, the predicted price obtained from using the regression model was compared to the actual 2022 data.

To test the regression model's ability to predict 2022 data, a set of assumed 2022 data was created, by assuming a 4% growth in resale price from 2021's resale price, and then tested to find its explained variance. The obtained explained variance was > 0.8 which indicated that the regression model was useful in predicting 2022's resale price. [cell 26-30]

4. Results

To determine if the couple can pay for the resale flat, the couple's gross salary, grants provided and loan repayment will be used to determine the affordability of the resale flat.

Assumptions

1. The couple gross monthly income = \$7000 based on median salary from Graduate Employment Survey (GES) 2021
 - a. Ordinary account monthly contribution is 23%
 - b. The salary does not increase
2. The flat they are looking at is a 3 room flat
3. The grant amount received by the couple = \$95,000
 - a. Family grant - \$50,000 (both are Singaporean)
 - b. Enhanced Housing grant - \$25,000
 - c. Proximity grant - \$20,000 (the couple will stay within 4km of the parents)
4. The couple will be taking up a HDB loan to pay the remaining fees.
 - a. The loan interest is 2.6% per annum
 - b. The repayment plan is 25 years
5. The couple is planning to make the resale flat purchase in 2022.
6. The couple is interested in purchasing a house in Geylang.

The couple have identified a house in Geylang with a floor area of 67 and remaining lease of 55 years.

The predicted sale price determined by the multivariate regression model is:

Predicted sale price = $5111.387(\text{floor_area_sqm}) + 523.013(\text{totalMonths}) - 381448.463$

Predicted sale price = \$306,203

The couples gross monthly income = \$7000

The couple's ordinary account (OA) monthly CPF contribution = $0.23(\$7000) = \1610

Item	Deduction (SGD)	Addition (SGD)
Predicted resale price	N/A	+ 306,203
Grants	- 95,000	N/A
Interest	N/A	+199,031.95
Total amount payable	N/A	410,234.95
Total amount payable per month	N/A	1367.4498
Ordinary account monthly contribution	-1610	N/A
Total amount payable per month (in cash)		-243

The couple would not be required to use cash for the loan repayment.

Using 2022 data available in the dataset, we can determine the error based on actual 2022 data.

Actual resale (SGD)	Predicted resale price (SGD)	Error (%)
345000	306203	11.24

The high R^2 and the relative low error determines that the regression model is quite successful in predicting the resale price for 2022.

5. Conclusion

Based on the results obtained above, it is determined that the regression model is very useful in predicting the resale price for 2022. However, the prediction of the resale price of 2023 and 2024 has decreased in accuracy which can be seen from the explained variance value. Hence, for the couple who wants to delay their plan to buy a house, they will have to take the risk of not having the best price.

From the predicted price obtained from the regression model, the monthly repayment was calculated and it is determined that the couple would be able to afford to flat with the OA monthly contribution without needing any top up payments. This code is not limited to only to Bioengineering fresh-graduates but also whoever wants to buy resale flats given that they provide information such as monthly gross income and their eligibility of their grant.

Furthermore, the R^2 of the regression model for 2023/2024 decreases but it is still relatively high which indicates that the regression model would be able to predict the resale price in 2023/2024 but its limitation would be that the error in predicting resale price will increase.

There are limitations to what can be done for this project and if worked upon, this model and report would have a higher degree of accuracy to determine the couple's ability to afford a resale flat.

Firstly, there are limitations in predicting future resale price using historical data as there could be other factors such as house condition, the safety of the neighborhood, distance from the nearest MRT and distance to city center that would influence resale price but these predictors were not included in the data set. This implies that future work would be required to investigate these factors that were not listed and its influence to resale price.

Next, the codes used for the determination. Based on our limited knowledge in coding, we are unable to use complex code and concepts which could better analyse the code such as implementing object oriented (OO) programming to improve efficiency of the code and call on methods from classes based on the objects created and using loops could help to improve efficiency of the notebook which would greatly improve the time taken to compute the price.

Lastly, our solution would be limited to our assumptions made as they ignore other real world factors such as inflation, recession, disruptive policies that could affect the housing market and the resale price of the houses. It ignores other policies and grants made available to different classes of people which would affect the conclusion. Other users following this model may not be placed in a similar situation which affects the conclusion drawn and this data set should be updated to get a better prediction of the resale price.

References

Flatportal. HDB. (n.d.). Retrieved April 2, 2022, from <https://homes.hdb.gov.sg/home/calculator/budget>

First-timer applicants. HDB. (n.d.). Retrieved April 2, 2022, from <https://www.hdb.gov.sg/residential/buying-a-flat/resale/financing/cpf-housing-grants/firsttimer-applicants>

Living with/ near parents or child. HDB. (n.d.). Retrieved April 2, 2022, from <https://www.hdb.gov.sg/residential/buying-a-flat/resale/financing/cpf-housing-grants/living-with-near-parents-or-child>

Living with/ near parents or child. HDB. (n.d.). Retrieved April 2, 2022, from <https://www.hdb.gov.sg/residential/buying-a-flat/resale/financing/cpf-housing-grants/living-with-near-parents-or-child>

CPF: What are the CPF allocation rates from 1 January 2022? Central Provident Fund Board (CPF Board). (n.d.). Retrieved April 2, 2022, from <https://www.cpf.gov.sg/member/faq/growing-your-savings/saving-as-an-employee/what-are-the-allocation-rates-from-1-january-2022>

Flatportal. HDB. (n.d.). Retrieved April 2, 2022, from <https://homes.hdb.gov.sg/home/calculator/budget>

Graduate Employment Survey - MOE.GOV.SG. (n.d.). Retrieved April 2, 2022, from <https://www.moe.gov.sg/-/media/files/post-secondary/ges-2021/web-publication-ntu-ges-2021.pdf>

Eligibility. HDB. (n.d.). Retrieved April 2, 2022, from <https://www.hdb.gov.sg/residential/selling-a-flat/eligibility>