Project 2

HDB Resale Prices

vivalamuffin

TABLE OF CONTENTS

01
Context

Who are we?

02
MARKET ANALYSIS

Understanding the market

03
Problem

What is the issue?

04
Research findings

Trends

05
Recommendations
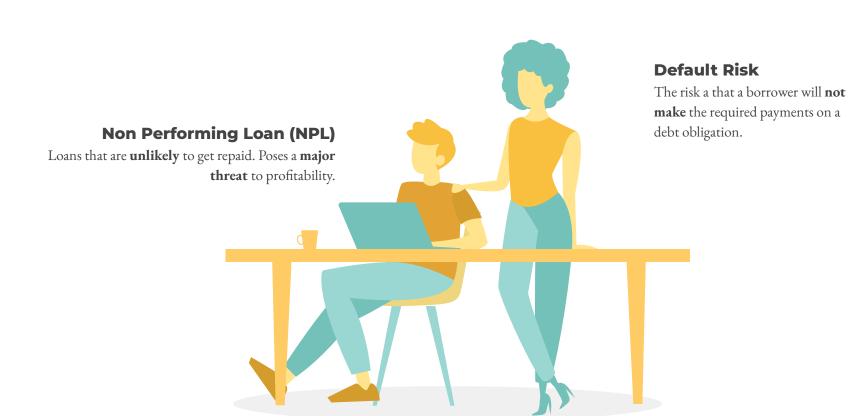
Our solution

06
Future work

Limitations

Default Risk

The risk a that a borrower will **not make** the required payments on a debt obligation.

Non Performing Loan (NPL)

Loans that are **unlikely** to get repaid. Poses a **major threat** to profitability.

- HDB's information on resale process is **comprehensive** and has a detailed guide on the procedure and documentation required.

- There is a wealth of information and resources available to **multiple stakeholders.**

- Conclusion: the process is **easy to follow,** and information is **freely and readily** available.

40%

Of Singaporeans are facing mortgage stress.

3.85%

Fixed home loan rates have spiked to levels not seen in years.

$32 million

NPL amount of top 3 SG banks (2022).

Property price and income are two **critical factors** that lenders consider while assessing loans. **Easy access** to income, but **not to** resale price.

Both factors have a **strong positive correlation** and will give a **holistic indication** of affordability.

They are thus ideal to create a **simple metric** that can be an **early-stage risk assessment** tool.

Relation of Income vs Resale Price

Correlation = 0.94

Average Resale Price   Average yearly income

Source: Income statistics from - https://blog.seedly.sg/monthly-household-income-singapore-comparison-district-where-you-stay/
Source: What FIs look for - https://www.homeloanexperts.com.au/home-loan-articles/7-factors-that-affect-your-home-loan-eligibility/

"Is it possible to predict HDB resale prices using indicators such as flat attributes and location? How can a predictor help as an early warning risk indicator to Banks?"

Problem Statement

Research findings

Average price of resale flat

$449 162

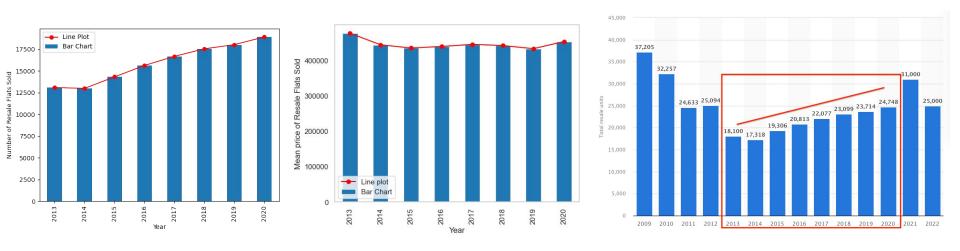## Cheapest resale flat

$150 000

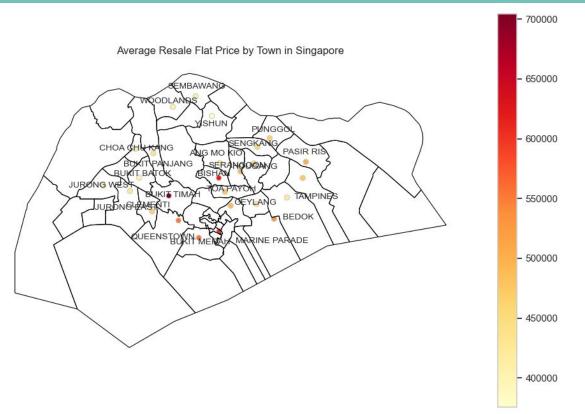## Most expensive resale flat

$1 258 000

Is there a consistent upward trend in the pricing of residential flats over the years?

The number of resale flats being sold increases over the years but the average price remains about the same.
This could be due to an increased amount of supply of resale flats.
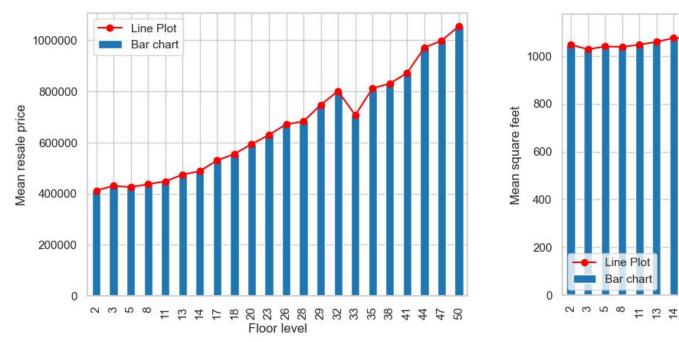
Credit: https://www.statista.com/statistics/1035048/singapore-total-number-hdb-resale-units/

Resale flats in the central part of Singapore are more expensive
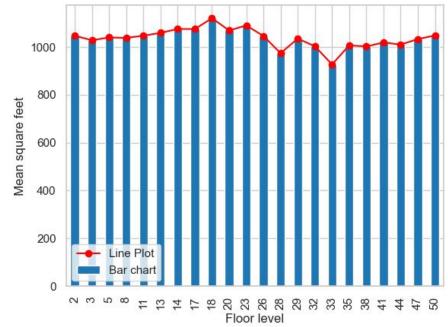

Average Resale Flat Price by Town in Singapore

The central area enjoys a premium price due to it's proximity to the central business district and the abundance of amenities like shopping malls.
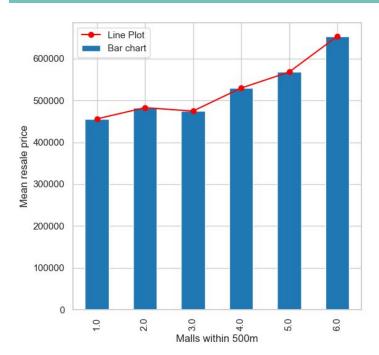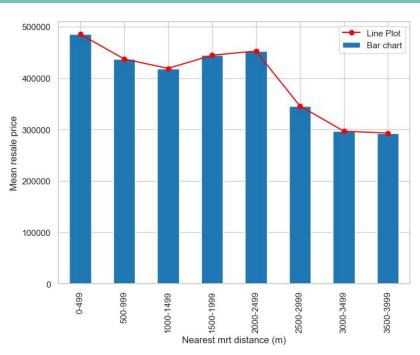
The higher the floor, the more expensive the flat

Flats on the higher floors do not necessarily have more area.
Flats on higher floors offer better views and more privacy.

The more amenities that are nearby, the more expsive the flat is

The more malls within 500m, the more expensive the flat.
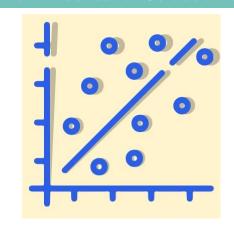The highest average resale price of a flat is found when an mrt is within 500m.

How we dealt with missing values

Missing values	Number of missing values
Mall_Nearest_Distance	829
Mall_Within_500m	92789
Mall_Within_1km	25426
Mall_Within_2km	1940
Hawker_Within_500m	97390
Hawker_Within_1km	60868
Hawker_Within_2km	29202

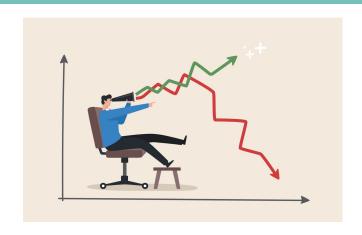We used the K Nearest Neighbor method to impute the missing data and chose k to be 5. The data we chose to impute are:
- Hawker_Within_1km
- Mall_Nearest_Distance
- Mall_Within_1km

Feature Selection & Model Performance

Feature Selection

Model Performance

Feature Indicator

Model Indicator

Model Indicator

R-Value:

R^2-Value:

RMSE Value:

-1 to 1

0 - 100%

Price Prediction

RMSE - Root Mean Square Value

Feature Selection

Model Performance

R-Value

R^2-**Value**

RMSE Value

Strength of
Linear Relationship

Variance
Explainability
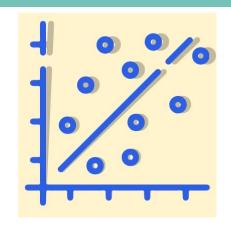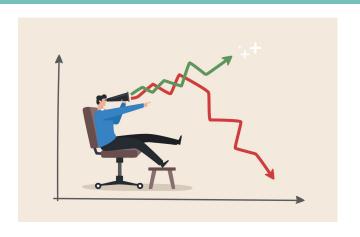
Accuracy of
Prediction

Feature Selection
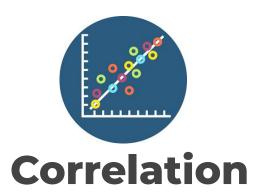
R-Value

Strength of
Linear Relationship

Model Performance

R^2-Value

Variance
Explainability

RMSE Value

Accuracy of
Prediction

Domain Research

What do buyers and domain experts say about this subject?

Correlation
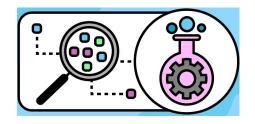
What is the strength of the linear relationship between a factor and the resale price?

Engineering

Based on Correlation and Domain Research, can we create features that show a stronger relationship?
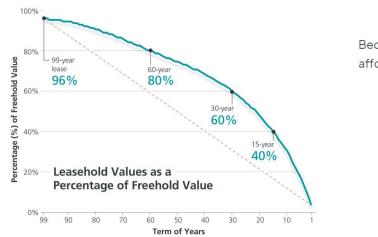
PropertyGuru Group, said that while being close to nature reserves and nation parks may be an advantage, it may not be as critical as other factors such as proximity to schools, shopping centres and MRT stations.

Flats in mature estates such as Ang Mo Kio, Bedok, and Queenstown tend to command higher prices. The proximity of

If you're eyeing a rare type of resale flat, such as an HDB maisonette or jumbo flat, be prepared to fork out a higher COV. As these flat types are no longer in production, they're highly in demand because of their rarity.

https://www.straitstimes.com/singapore/resale-flats-with-views-of-greenery-fetch-higher-price-study
https://dollarbackmortgage.com/blog/rise-in-hdb-prices-factors/
https://www.propertyguru.com.sg/property-guides/hdb-valuation-sales-12882

Homes located on higher storeys tend to command higher premiums. In fact, it's not uncommon to see a price difference of thousands of dollars based on the floor level alone. As such, you can sell (or

Leasehold Values as a Percentage of Freehold Value

Because the PSF is directly affected by the size of the property, it is not the most accurate indicator of affordability. It's like comparing apples and oranges; the playing field is not exactly equal. Though you

https://www.propertyguru.com.sg/property-guides/high-floor-vs-low-floor-unit-which-is-better-45449
https://ohmyhome.com/en-sg/blog/psf-doesnt-tell-you-enough-heres-what-else-you-can-focus-on/#:~:text=In%20contrast%2C%20with%20a%20larger,hence%20the%20higher%20overall%20quantum.

HDB Resale Statistics

(Q1 2023)

TOWNS	1-ROOM	2-ROOM	3-ROOM	4-ROOM	5-ROOM	EXECUTIVE
ANG MO KIO	-	*	$380,400	$520,000	$697,500	*
BEDOK	-	*	$367,000	$518,000	$684,000	*
BISHAN	-	-	*	$690,000	$880,000	*
BUKIT BATOK	-	$340,000	$385,000	$590,000	$758,000	*
BUKIT MERAH	*	*	$455,000	$838,000	$875,400	*
BUKIT PANJANG	-	-	$388,000	$493,000	$608,500	*
BUKIT TIMAH	-	-	*	*	*	*
CENTRAL	-	*	*	$894,000	*	-
CHOA CHU KANG	-	*	*	$495,000	$603,000	$730,000
CLEMENTI	-	*	$390,400	$597,500	*	*
GEYLANG	-	*	$332,500	$668,000	*	*
HOUGANG	-	*	$382,900	$525,000	$670,000	*
JURONG EAST	-	*	$365,000	$492,000	*	*
JURONG WEST	-	*	$345,000	$476,000	$589,000	$727,500
KALLANG/WHAMPOA	-	*	$383,000	$778,900	$865,000	*
MARINE PARADE	-	-	$425,000	*	*	-

Estate

Does the town area matter?

Floor

Does a higher floor matter?

Model

What is considered valuable?

Convenience

Do these amenities fall within 1km?
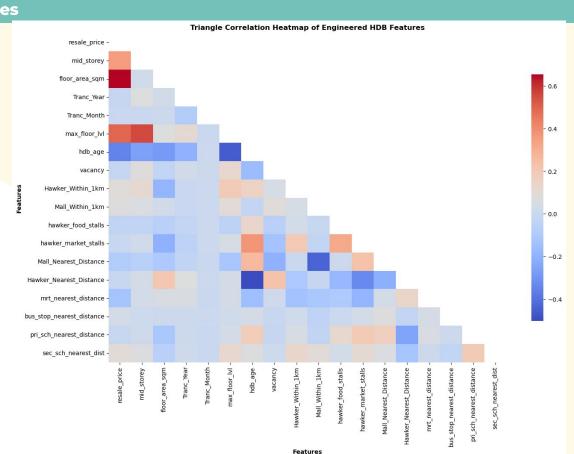
Inflation

Do prices increase year on year?

Distance

How far are these amenities?

Observation:

- HDB age and floor area big influences on price.
- Many possible variables and factors
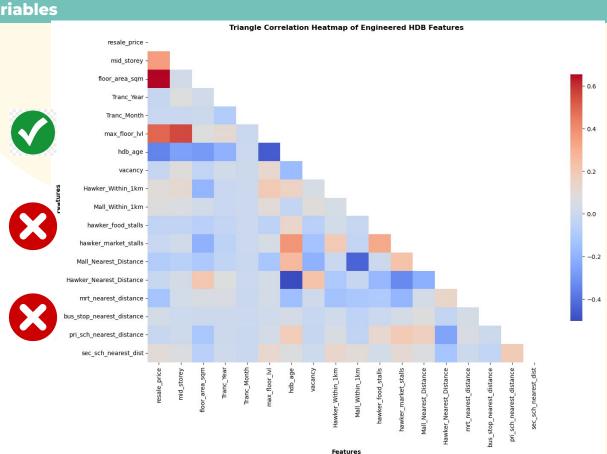- Many single variables have poor correlations with resale prices

Triangle Correlation Heatmap of Engineered HDB Features

Evaluation:

Floor Level

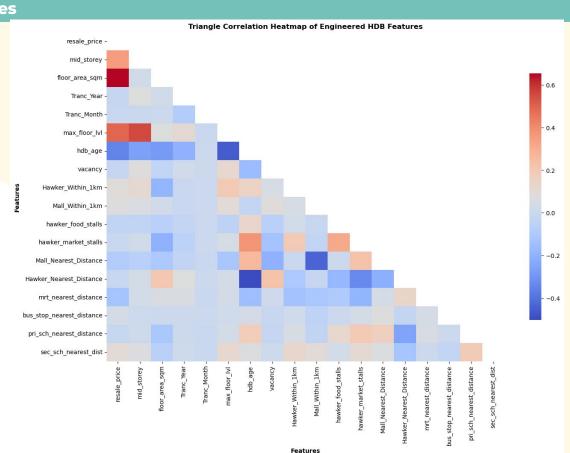Amenity Distance

Convenience

Triangle Correlation Heatmap of Engineered HDB Features

Insight:

- Include HDB Block Height in modelling

- Group similar features to reveal stronger relationships

Triangle Correlation Heatmap of Engineered HDB Features
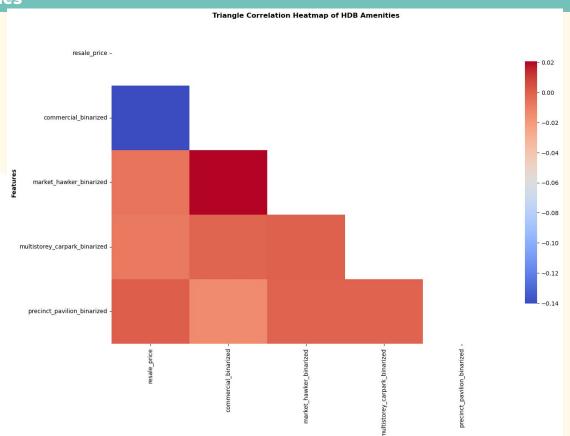
Observation:

- Individual Amenities do not contribute significantly to price

- Having access to 1 amenity does not raise prices

Triangle Correlation Heatmap of HDB Amenities

Primary School

Is proximity to a good school a selling point?

Mature Estate

Do mature estates command a higher price?

Town Region

Does the town region affect demand?

HDB Model Type

What about these rarer types?

Floor Proportion

How high is their floor?

Corridor Length
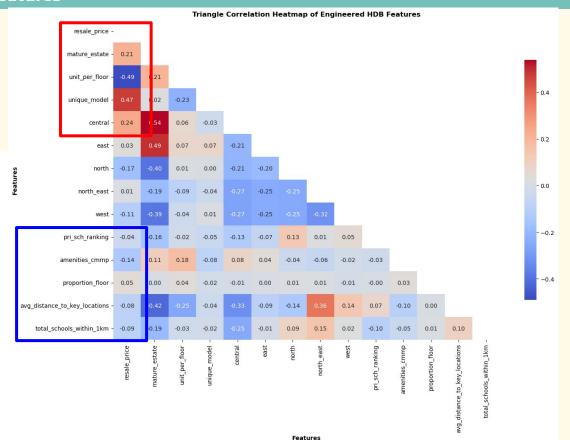
Does it affect price?

Insight:
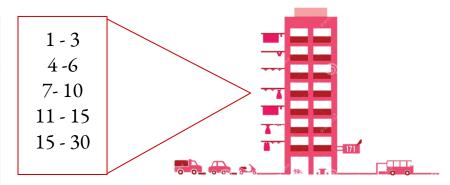
- Many hidden relationships are significant
- e.g Units per Floor

- Certain variables that seemed obvious are not strong e.g School Ranking

Triangle Correlation Heatmap of Engineered HDB Features
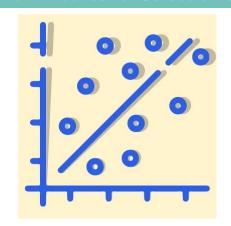
Total Number of Features

Features	Number
Numerical Features	26
Categorical Features	2
Total	28

1 - 3
4 -6
7- 10
11 - 15
15 - 30

Feature Selection

R-Value

Strength of
Linear Relationship

Model Performance

R^2-Value

Variance
Explainability

RMSE Value

Accuracy of
Prediction

Linear Regression Model

Ridge Regression Model

Lasso Regression Model

Test Scores

R^2-Value: 93.211%

RMSE: $37,342

Test Scores

R^2-Value: 93.208%

RMSE: $38,349

Test Scores

R^2-Value: 93.206%

RMSE: $37,353

Linear Regression Model

Ridge Regression Model

Lasso Regression Model

Test Scores

R^2-Value: 93.211%

RMSE: $37,342

Test Scores

R^2-Value: 93.208%

RMSE: $38,349

Test Scores

R^2-Value: 93.206%

RMSE: $37,353

Linear Regression Model

Ridge Regression Model

Lasso Regression Model

Test Scores

R^2-Value: 93.211%

RMSE: $37,342

Test Scores

R^2-Value: 93.208%

RMSE: $38,349

Test Scores

R^2-Value: 93.206%
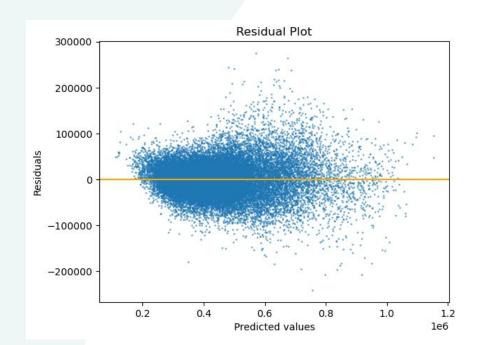
RMSE: $37,353

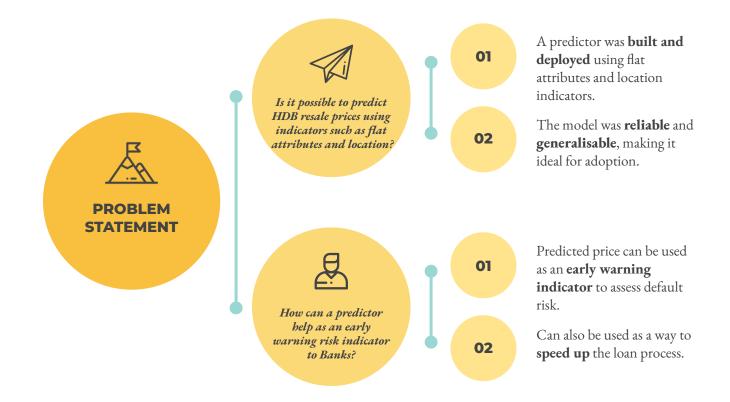# Linear Model Residual Plot

Residual Plot

```
count      30127.000000
mean          -5.903119
std        37342.479250
min      -241464.828125
25%       -22458.968750
50%        -1128.328125
75%        20821.945312
max       275192.078125
Name: resale_price, dtype: float64
```

PROBLEM STATEMENT

Is it possible to predict HDB resale prices using indicators such as flat attributes and location?

01 A predictor was **built and deployed** using flat attributes and location indicators.

02 The model was **reliable** and **generalisable**, making it ideal for adoption.

How can a predictor help as an early warning risk indicator to Banks?

01 Predicted price can be used as an **early warning indicator** to assess default risk.

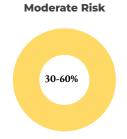**02** Can also be used as a way to **speed up** the loan process.

Model Demonstration

Resale Price-to-Income (RPI) Ratio

RPI Ratio = (Predicted Resale Price/Applicant's Annual Income) ×100

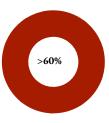## Interpretation

- The higher the RPI ratio, the less likely the applicant is to default on payments.

- Thresholds can be customised to suit the needs of the bank.

Example Thresholds

Low Risk

<30%

Moderate Risk

30-60%

High Risk

>60%

Recommendations and Impact

PROFIT DRIVEN

 # 1

The RPI Ratio can be used as an **early stage risk assessment** tool, thus **reducing Default Risk.**

 # 2

Banks can **customise** their interest rates based on the RPI Ratio.

SERVICE DRIVEN

 # 3

The model can be provided as a **white-label app** to banks to include in their **services portfolio.**

 # 4

Using predicted resale price can **speed up** the loan process and **personalise** each loan.

Future Scope and Developments

- The model can also include prices for private properties (like condominiums) to give a more comprehensive outlook of SGs property market.

- The app can include a comparison of predicted price vs average price (say by region) to give an indication of heating/cooling of the market.

- The model can be adjusted for inflation to make predictions more realistic.

WHAT DO YOU MEAN YOU "SKIPPED" LAST MONTH'S PAYMENT?