The purpose of this slide is to fulfill the requirements for DSI 39, Project 2.

Group Name: The Colony
Members:
● Jackie Seah
● Li Cheng
● Mei Qi
● Yvonne Lim

Role:
CPFB Data Science team

Audience:
CPFB Senior Management

Problem Statement:
With rising resale flat prices and record high prices such as the latest 4 room flat in Tiong Bahru at S$1.5mil. How can CPF board assist potential homeowners to plan their housing finances better?

Analysis and Prediction of HDB Resale Flat Price

CPFB Senior Management Meeting
15 Sep 2023

Meeting Agenda

HDB Resale Prices are at a Record High

THE STRAITS TIMES — SINGAPORE

HDB resale prices jumped 12.7% in 2021, record 259 million-dollar flats sold

Record of 259 million-dollar HDB resale flats in 2021!

Source:
https://www.straitstimes.com/singapore/housing/hdb-resale-prices-jumped-127-in-2021-record-259-million-dollar-flats-sold

today — Singapore World Big Read Adulting 101 Gen Y Speaks Gen Z Speaks Voices Commentary 8 Days More

Record 21 4-room HDB resale flats sold for over S$1m in Q1 2023; analysts expect trend to continue

BY LOW YOUJIN

More than 20 4room HDB resale flats sold over S$1 million!

Source:
todayonline.com/singapore/record-4-room-resale-flats-2143666

HDB Resale Flat Landscape

56%

of Singaporeans who are in their 20s are *worried that they cannot afford a home.*

40%

of Singaporeans are worried that they will *not be able to pay their monthly instalments for their mortgage.*

$3.9 Bn

CPF monies are used by homeowners to pay their housing loan instalments in 2022.

Available data on HDB resale prices are limited

1. Toa Payoh

Size of flat	Price
3-room	S$330,000 - S$800,000
4-room	S$530,000 - S$1,100,000
5-room	S$700,000 - S$1,300,000

2. Queenstown

Size of flat	Price
3-room	S$300,000 - S$740,000
4-room	S$600,000 - S$950,000
5-room	S$680,000 - S$1,400,000

Estimated HDB resale prices found online varied significantly from $500k - $700k !

OFFER TO PURCHASE REAL ESTATE

How can our agency assist potential homeowners to plan their housing finances better?

Knowing the price of the HDB resale flat that they intend to purchase, aspiring homeowners can better plan their finances

AMOUNT OF CPF SAVINGS YOU CAN USE

HOUSING LOAN AMOUNT YOU CAN GET

Project Objective

Fast and High Accuracy

Intuitive & Easy to Use

We aim to deploy a HDB resale price calculator that is <u>reliable and accurate</u> ($r^2 > 0.8$, RMSE < $50k) which is <u>intuitive and easy</u> (minimal number of interpretable features) for home owners to use.

RMSE is the measure of the differences between values (sample or population values) predicted by a model or an estimator and the values observed

Coefficient of determination, R^2 is the proportion of the variation in the dependent variable that is predictable from the independent variable(s)

HOW DO WE GET A USER-FRIENDLY & ACCURATE CALCULATOR?

Methodology for Price Prediction

- **Understanding the Dataset**
  - 76 potential features* to predict resale price
  - Data period: Mar 2012 - Apr 2021

- **Principle of Feature* Selection**
  - High correlation with resale price (e.g., larger floor area leads to higher price)
  - Literature Review & Domain expertise
 - High variability (i.e., feature values that changes with the resale price)

*Note:
Feature = Potential factor that will affect the HDB resale price
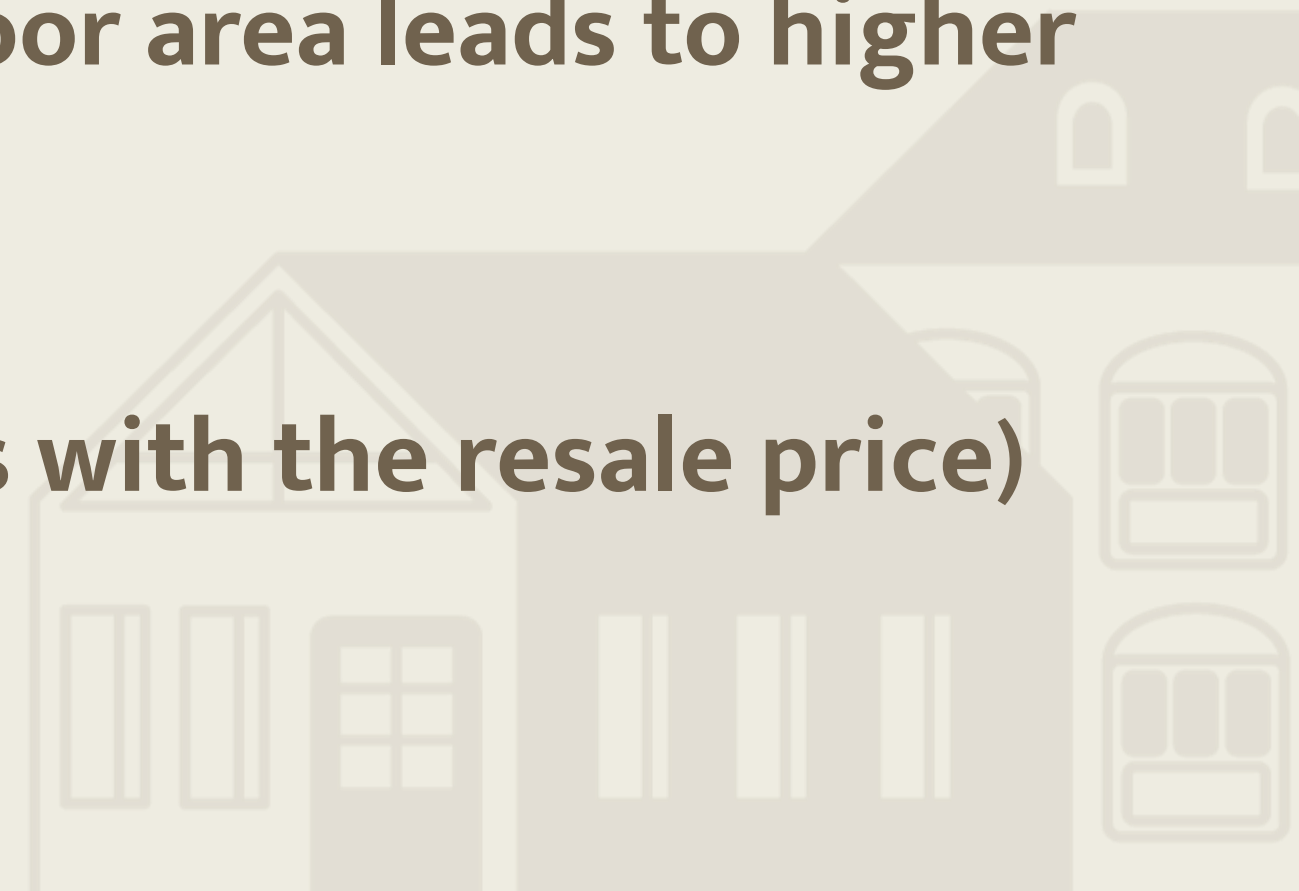
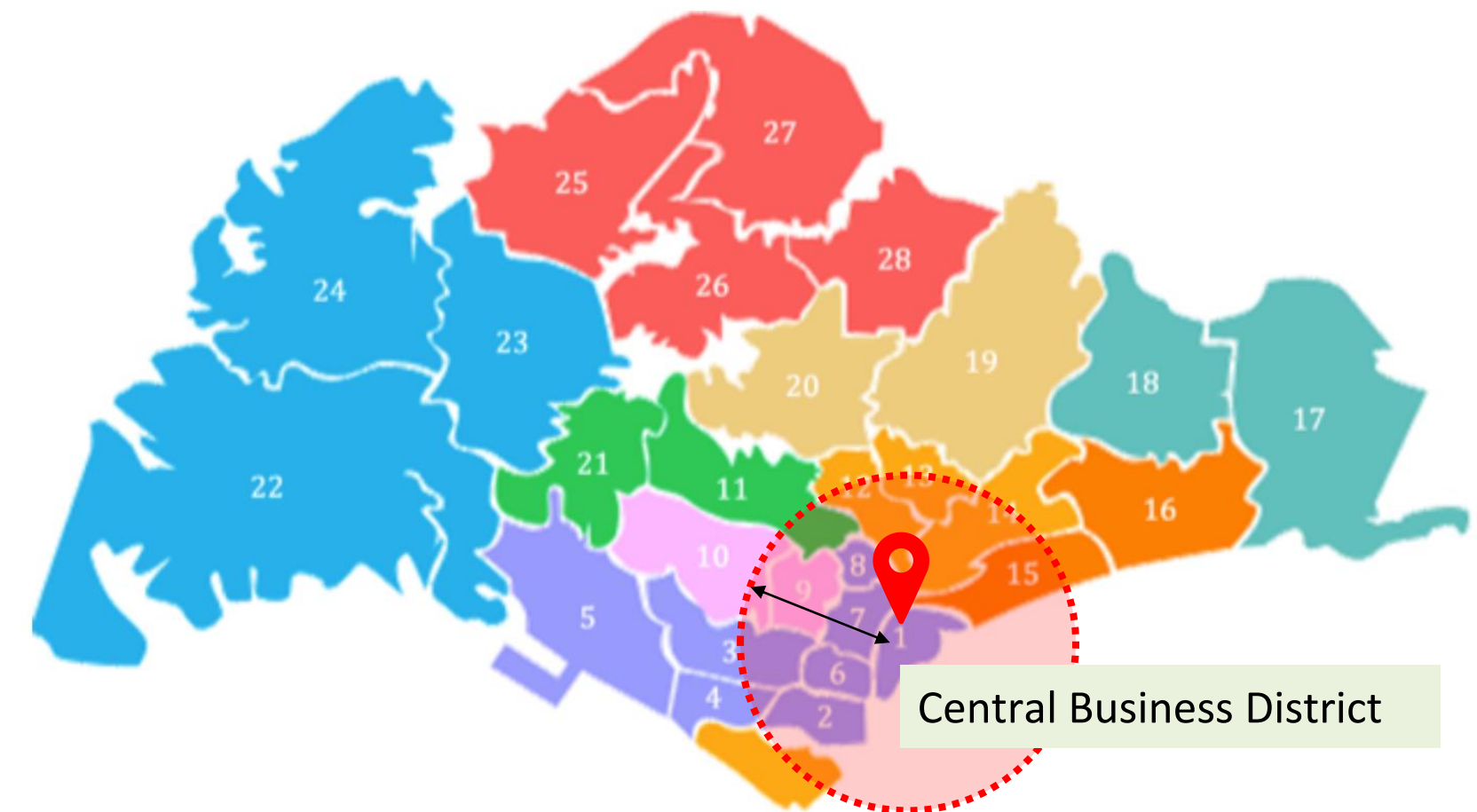# Main categories of features affecting resale price

A flat's floor size positively correlates with resale price; the flat model also shows varying median resale prices

($) '000,000

Resale price against floor area (in sqm)

Distribution of resale price for 4 room flat type
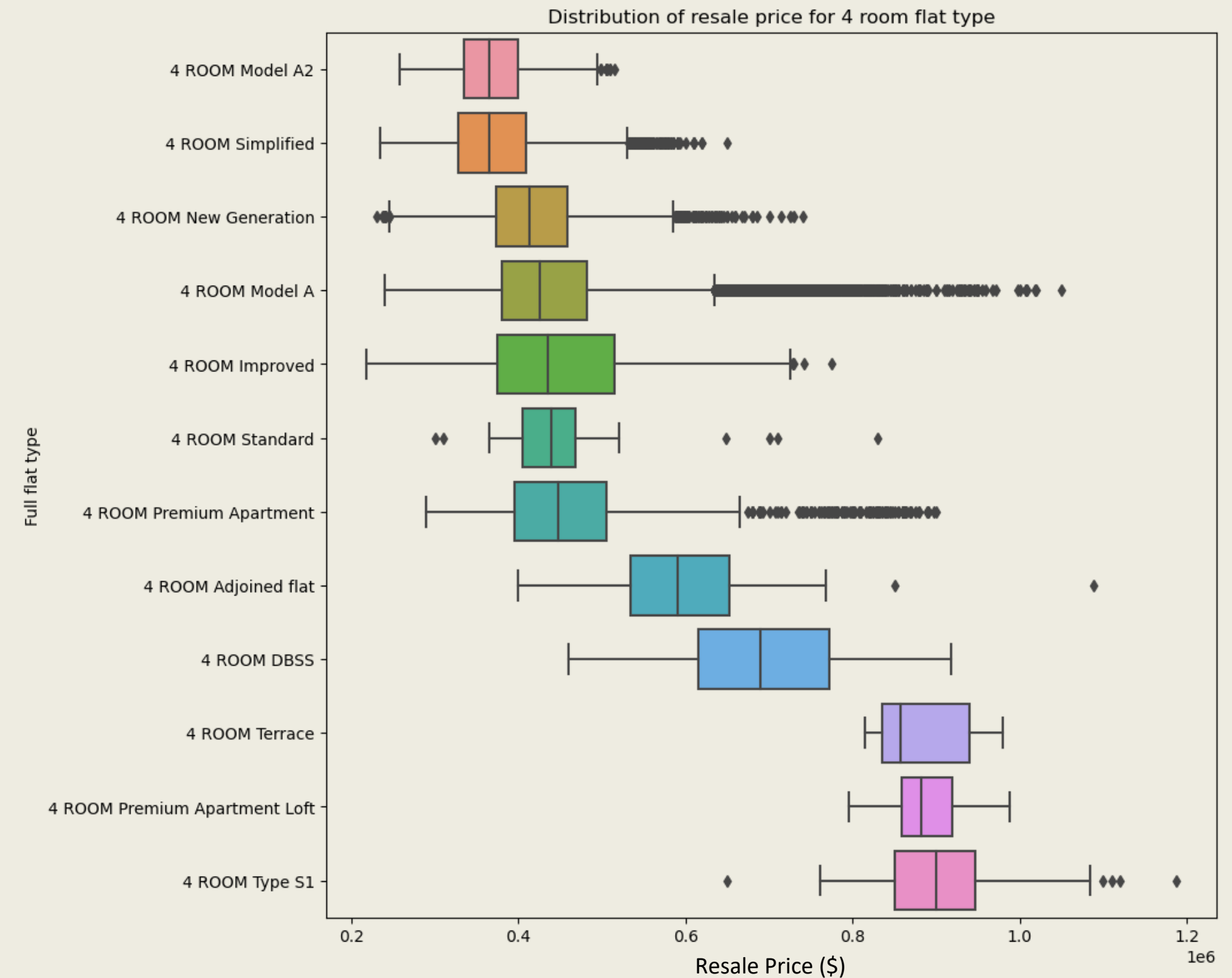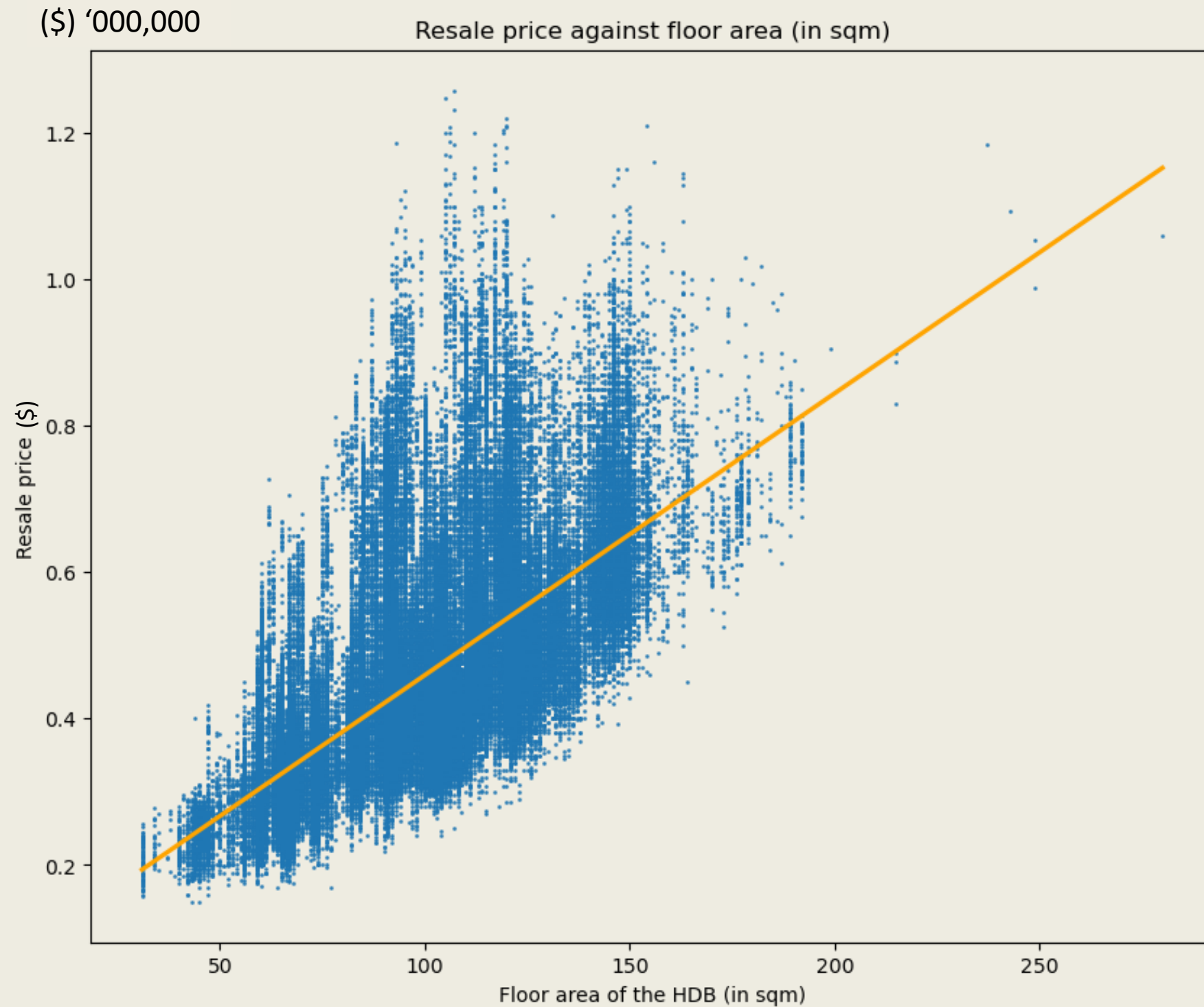
Postal information and proximities of flats to Central Business District (CBD) were surfaced as potential factors influencing property prices

Postal District	Postal Sector (1st 2 digits of postal codes)	General Location
01	01, 02, 03, 04, 05, 06	Raffles Place, Cecil, Marina, People's Park
02	07, 08	Anson, Tanjong Pagar
03	14, 15, 16	Queenstown, Tiong Bahru
04	09, 10	Telok Blangah, Harbourfront
05	11, 12, 13	Pasir Panjang, Hong Leong Garden, Clementi New Town

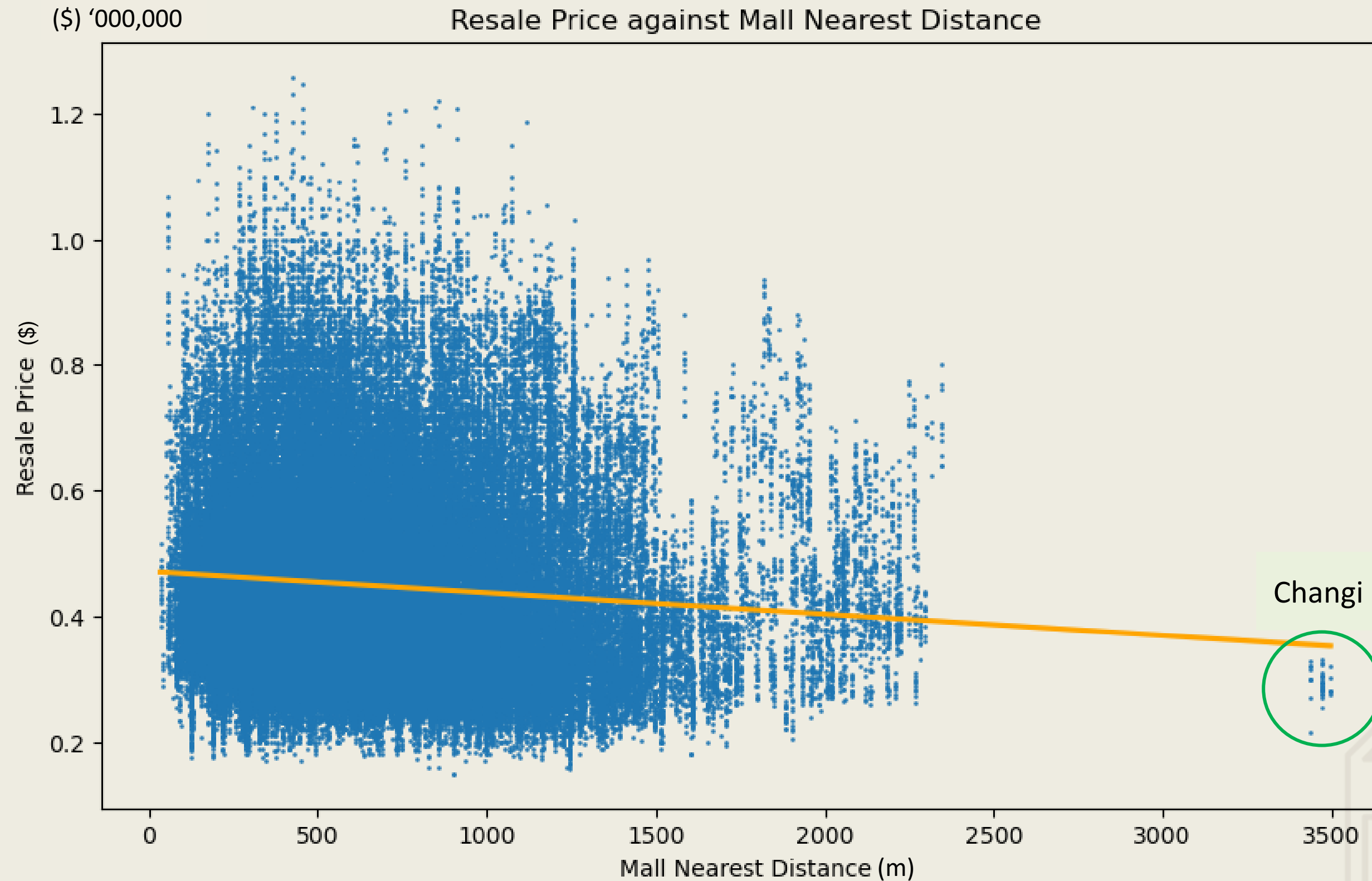**Singapore District Code Map Demarcation**

Central Business District

Source:
https://dr.ntu.edu.sg/handle/10356/147773
https://www.mingproperty.sg/singapore-district-code/
https://blog.moneysmart.sg/property/rent-singapore-cost-guide/

Flats located near the Mall have a higher resale price

($) '000,000

Resale Price against Mall Nearest Distance

Resale Price ($)

Mall Nearest Distance (m)

Changi

Potential homeowners looking for relatively cheaper resale flats, but want to have a mall nearby, should avoid Changi

Flats located near the MRT have a higher resale price

($) '000,000

Resale price against MRT Nearest Distance

Resale Price ($)

MRT Nearest Distance (m)
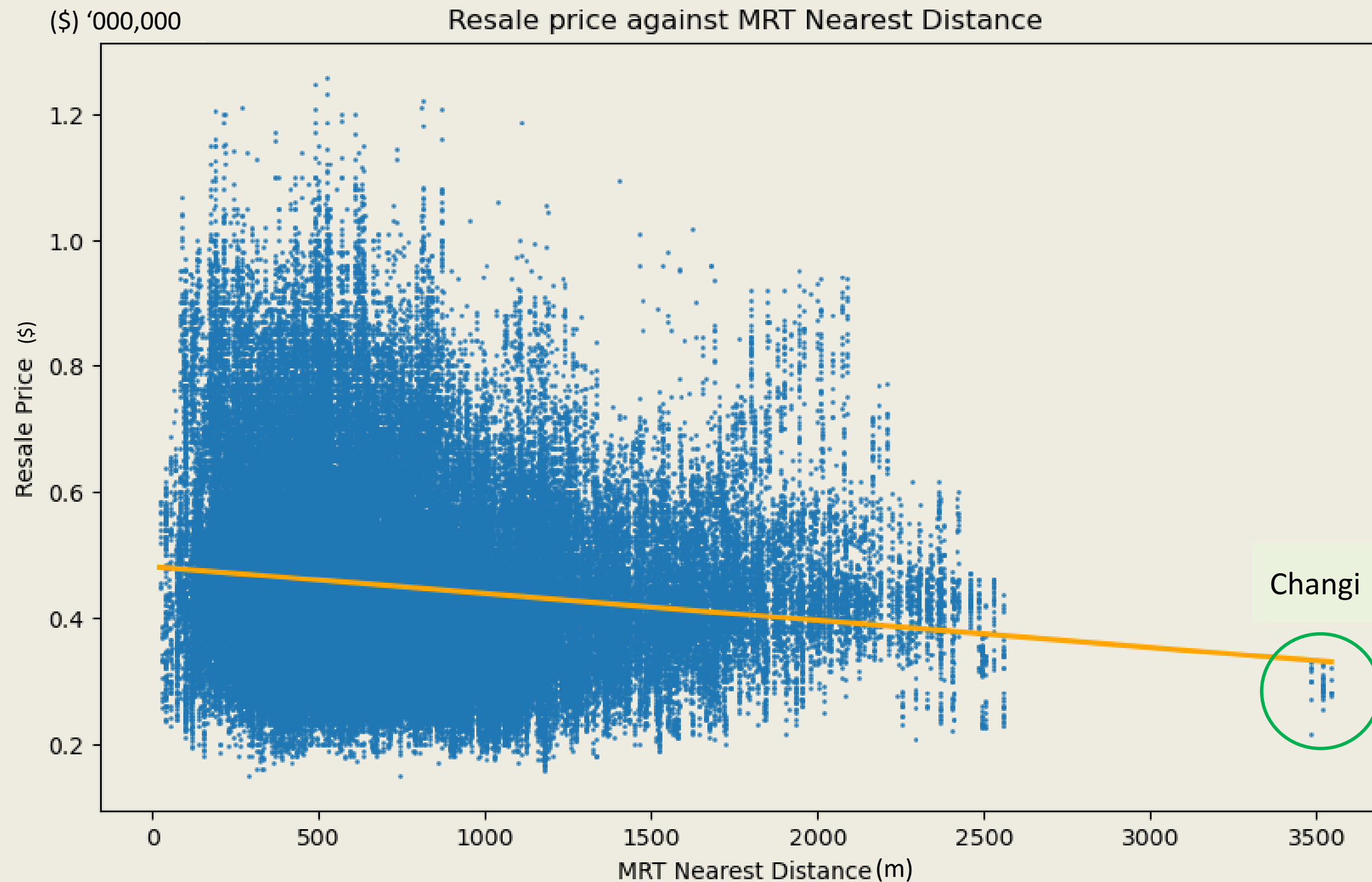
Changi

Potential homeowners looking for relatively cheaper resale flats, but want to have a MRT nearby, should avoid Changi
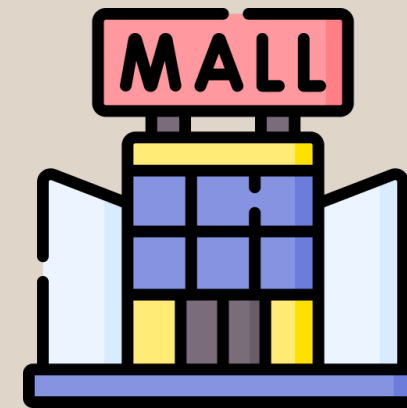
9 Final Features Affecting Resale Price

FLAT ATTRIBUTES

LOCATION

AMENITIES

ACCESSIBILITY

- Year of HDB flat transaction
- Floor area of the HDB flat
- Age of the HDB flat
- Flat model and type of the HDB flat
- Storey of the HDB

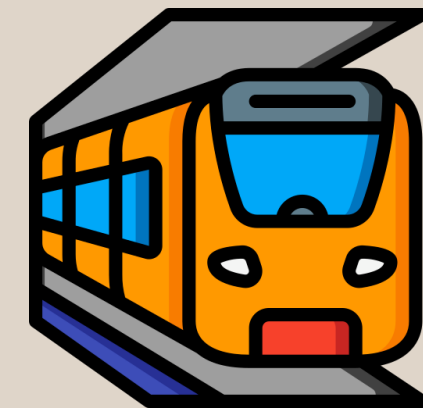- Postal sector of the HDB flat
- Distance of HDB flat from CBD area

- Distance of HDB flat from nearest mall

- Distance of HDB flat from nearest MRT

Linear Regression Model chosen as it has the best performance and faster runtime than Lasso

Regression Model	Train R^2 Score (closer to 1 is better)	Test R^2 Score (closer to 1 is better)	Cross Validation Score (closer to 1 is better)	RMSE (lower is better)
Linear	0.89	0.89	0.89	48,132
Lasso	0.89	0.89	0.89	48,132
Ridge	0.89	0.89	0.89	48,135

- Train Score ≈ Test Score: No sign of underfitting and overfitting
- Consistent Cross Val score: Model reliability

Using linear regression model, **89%** of the variations in resale price can be explained by our selected features, with **±$48,000** price difference.

RMSE is the measure of the differences between values predicted by a model and the actual values observed

Coefficient of determination, R^2 is the proportion of the variation in the dependent variable that is predictable from the independent variable(s)

SCAN HERE FOR CALCULATOR

HDB RESALE PRICE

CALCULATOR DEMO

https://dsi39-project2-colony.streamlit.app/

Summary

- **Problem: The price of HDB resale flats are rising and many are concerned about the finances. To help people better plan for it, there is demand for a HDB resale price calculator.**

- **Result:**
  - **9 key features are selected.**
  - **A linear regression model built, that is reliable and shows consistent performance across different datasets.**
  - **A working web app that is accessible via internet devices**

Moving Forward

IMPROVE PREDICTION SCORE

- **Explore other factors (e.g. impact of cooling measures on resale price)**
- **Explore other machine learning models**

USER TESTING

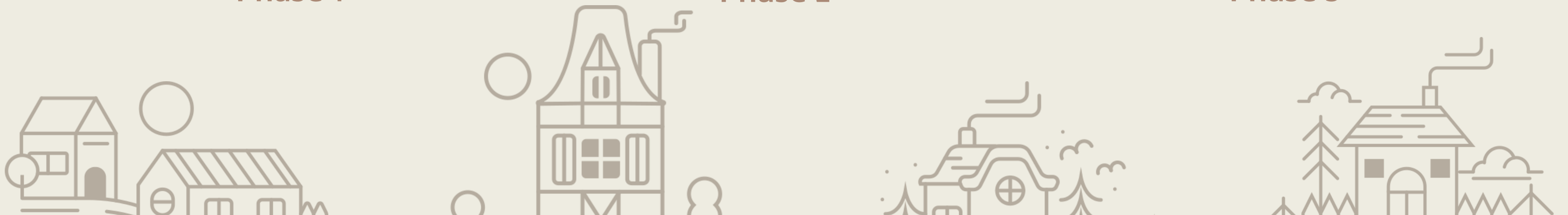- **Test and improve app usability with focus groups**

PUBLIC LAUNCH

- **Campaign launch to educate housing finance planning with calculation.**

Phase 1

Phase 2

Phase 3

Thank You

Interesting find: Resale flat price is not affected by nearby primary schools

($) '000,000

Price to Primary School Nearest Distance