IT8701
Introduction to Programming for Data Science
CA 1 Assignment

Singapore
Public Housing Market
Analysis

Eugene Chua
1048186J

Introduction

The Housing & Development Board (HDB for short) is Singapore's public housing authority.

Established in 1960 during Singapore's housing crisis, today, more than 1 million flats have been completed across the island.

HDB flats house 80% of Singapore's resident population, of which about 90% own their home.

Introduction

In the Jupyter notebook accompanying this slide deck, we seek to understand Singapore's public housing market in 5 parts:

1. Public Housing Market Overview
2. Town Growth: Mature & New Estates
3. Town Profiles: Mature Vs. New Estates
4. HDB Resale Price Index
5. HDB Resale Flat Prices

Libraries & Datasets

Libraries

- Numpy
- Matplotlib

Datasets & URLs

Dwelling Units Under HDB's Management
https://beta.data.gov.sg/collections/165/view

HDB Resale Price Index
https://beta.data.gov.sg/collections/152/view

Resale Transaction by Flat Type
https://beta.data.gov.sg/collections/167/view

Part 1
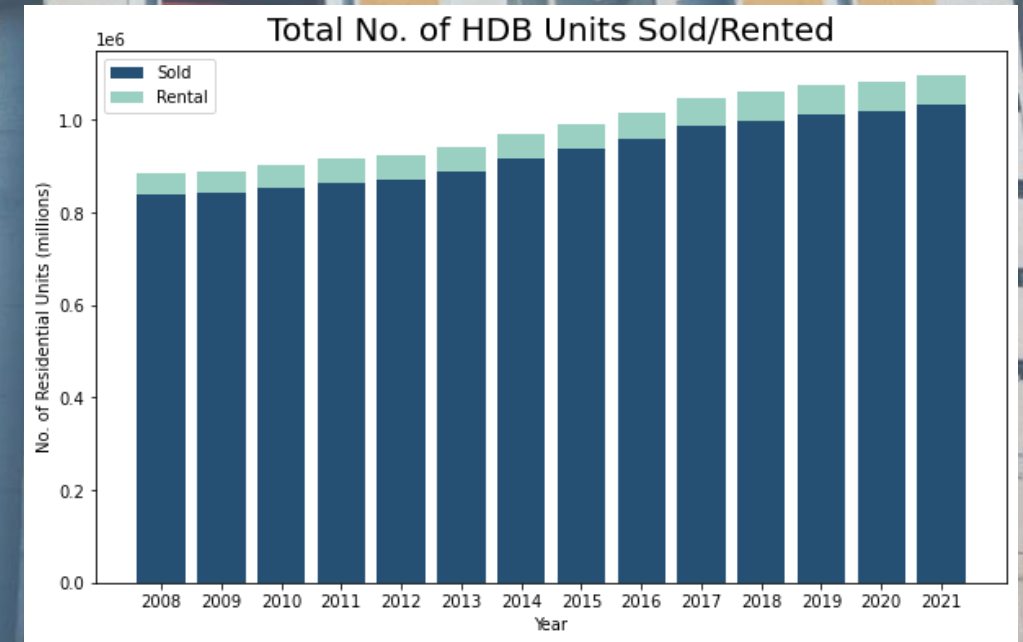## Public Housing Market Overview

Dataset:
Dwelling Units Under HDB's Management

Methodology:
- This dataset contains all HDB units that were sold or rented from 2008 to 2021.
- I used numpy to load the CSV file. Dataset was split into sold and rental units.
- Using a for loop, I summed up the number of HDB units for each year and stored them separately - one list each for sold and rented units.
- Lastly, I plotted them using `plt.bar` to get an overview of the yearly trend. To stack the bar chart, I specified that units sold be placed below rented units.

Insights:
As of 2021, HDB sold a total of 1,032,699 units (2008: 837,883), marking a 194,816 increase from 2008 figures.

Part 1
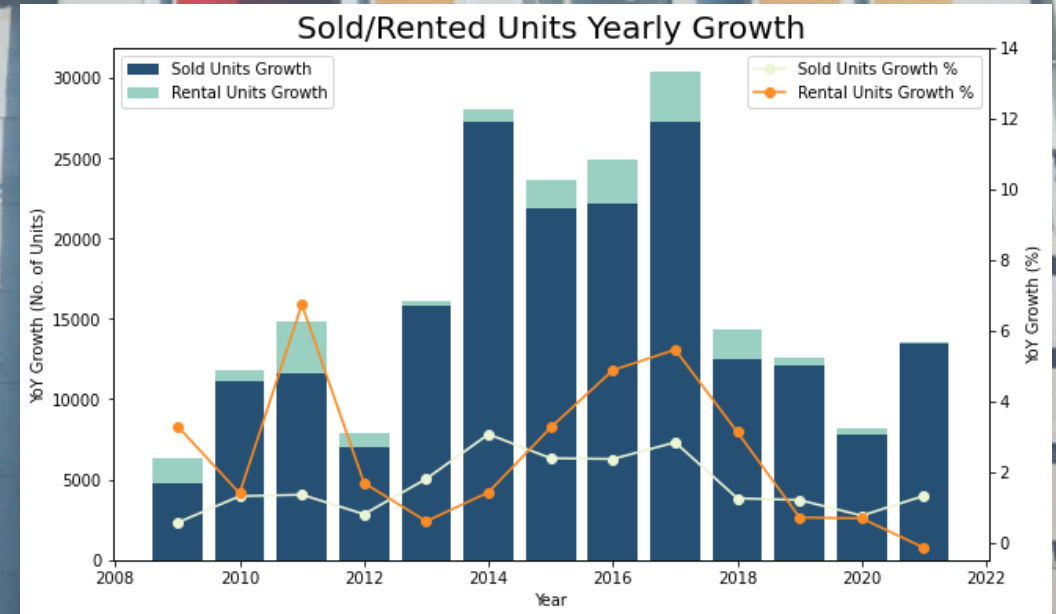## Public Housing Market Overview

Dataset:
Dwelling Units Under HDB's Management

Methodology:
- This chart expands on the previous chart, this chart goes into more detail.
- I looped through the summed figures to extract the yearly changes and the yearly growth rates.
- In addition to a stacked bar chart, I layered a line chart over to present a clearer view of yearly changes.
- To make sense of the chart, I added a second y-axis using `ax1.twinx()`.

Insights:
- The growth of units sold was stable, ranging between 0.6% and 3.1%.
- The rental market was more volatile. From 2013 to 2017, rental units' yearly growth rate increased steadily before declining from 2018 onwards.

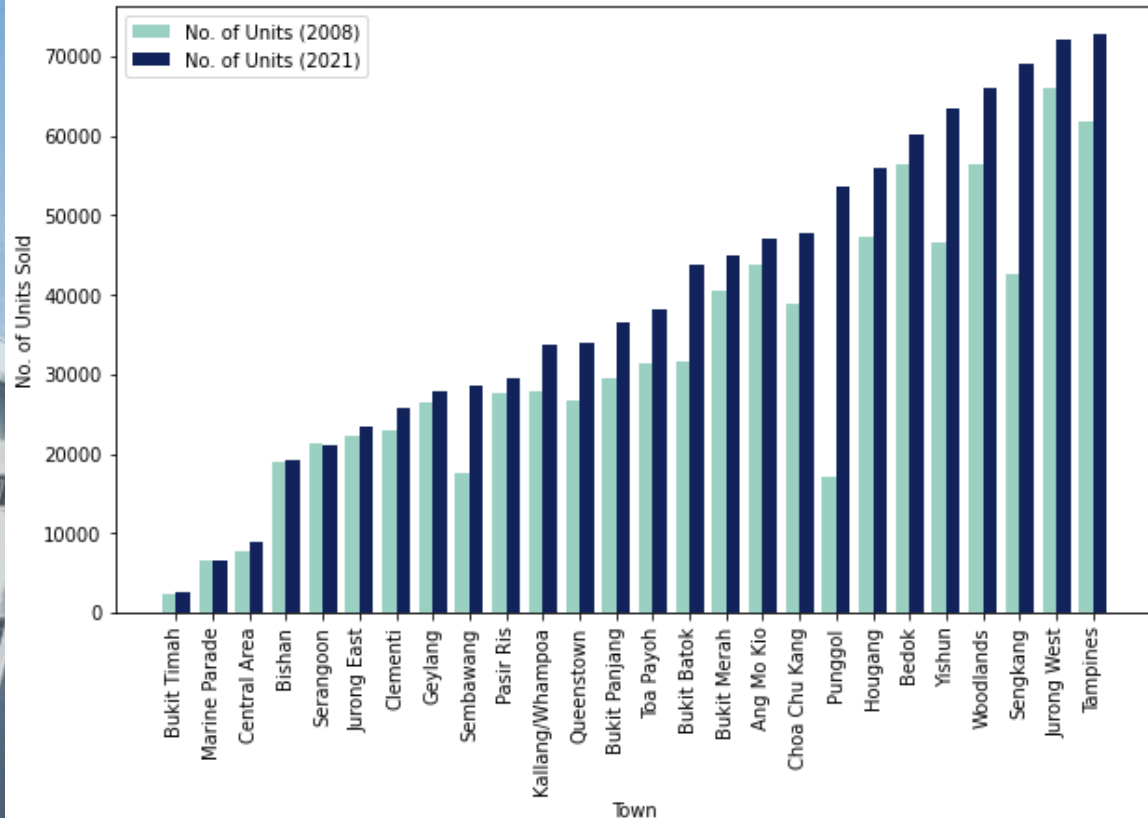# Town Growth: Mature & New Estates

HDB Units Sold - 2008 vs 2021

Dataset:
Dwelling Units Under HDB's Management

Methodology:
- For each town, I stored the yearly no. of units sold in a master dictionary.
- Using HDB's estate classification, I grouped them into mature and new estates; each has its own dictionary.
- Next, I extracted the following data for each town: (i) units sold in 2008 (ii) units sold in 2021 (iii) 2008 vs 2021 change (iv) growth rate; these are stored in a summary dictionary.
- Estates were sorted by number of units sold.
- Lastly, using the summary dictionary, I plotted a bar chart comparing 2008 and 2021 figures, side by side, for each town.
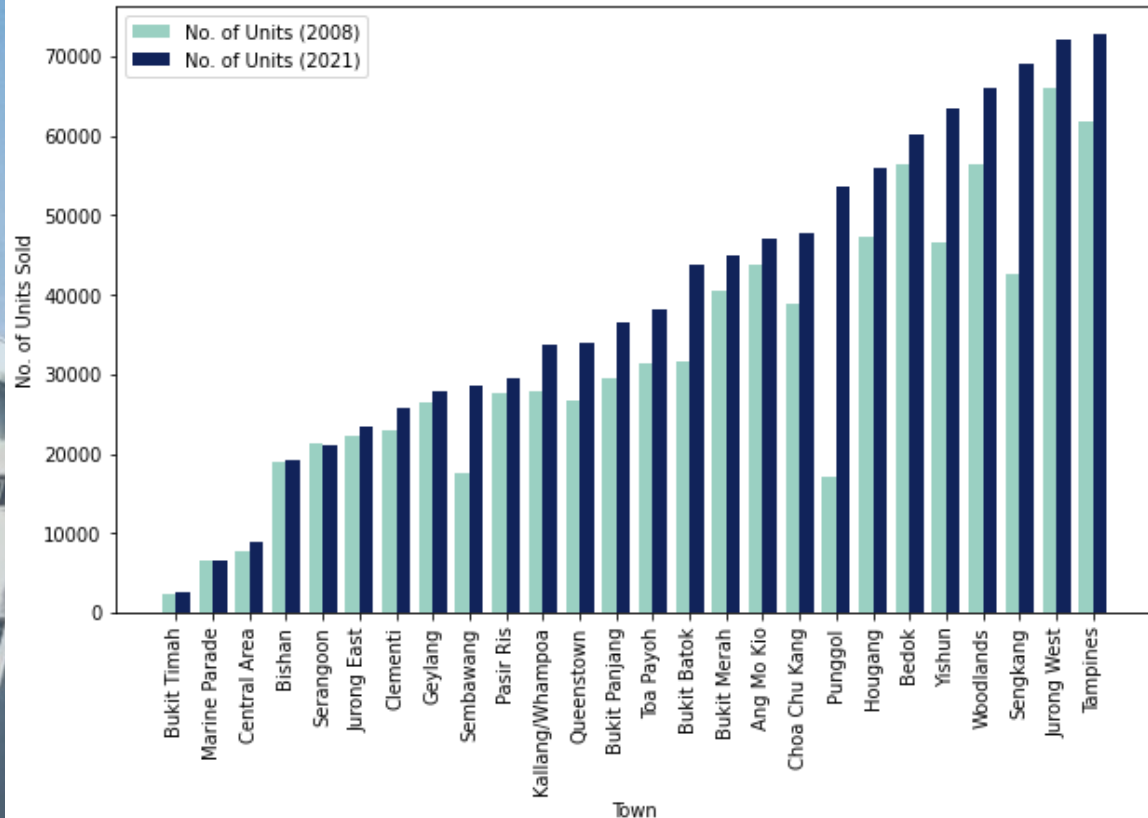- To achieve the final visual, I explicitly defined the number of x-axis labels and the bars' widths.

(continued)

Town Growth: Mature & New Estates

HDB Units Sold - 2008 vs 2021

Insights:
- Over the 13-year period, Jurong West and Tampines (both mature estates) were the largest.
- By the end of 2021, the latter pipped the former to the top.
- Estates that grew the most were all new estates:
  - (i) Punggol: +213.1%
  - (ii) Sengkang: +62.4%
  - (iii) Sembawang: +61.3%
  - (iv) Bukit Batok: +38.4%
  - (v) Yishun: +36.3%
- Of the 194,816 new HDB units sold during the 2008-2021 period, a third of them (62,983 units) were in Punggol and Sengkang.
- Collectively, 102,874 units were sold in the top five fastest growing estates - more than half of the total sold this period.

Part 3
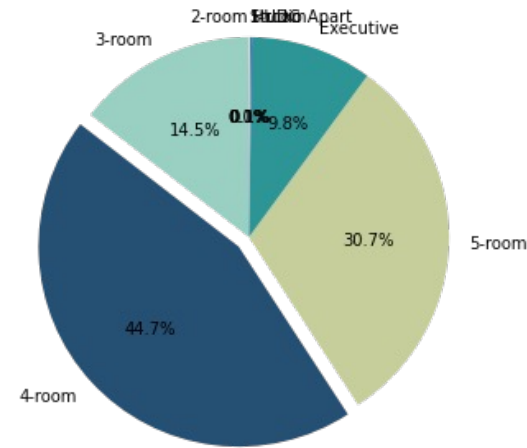## Town Profiles: Mature Vs. New Estates

Dataset:
Dwelling Units Under HDB's Management

Methodology:
- This section examines the unique profiles of mature and new estates.
- I plotted 4 pie charts for this section: (i) new estates in 2008. (ii) new estates in 2021, (iii) mature estates in 2008, and (iv) mature estates in 2021.
- I wrote a helper function to return me a dictionary containing units sold for each flat type by year and estate type.
- Starting with the new estates, I initiated a subplot for 1 row and 2 columns. Instead of `plt.pie`, I used `axs[i].pie` to plot the profile of new estates in 2008. I did the same for 2021 so that I could place them side by side.
- The same process was followed for mature estates.
- To achieve this final visual, I used the `explode` argument in `axs.pie` to make the 4-room flats' slice protrude.
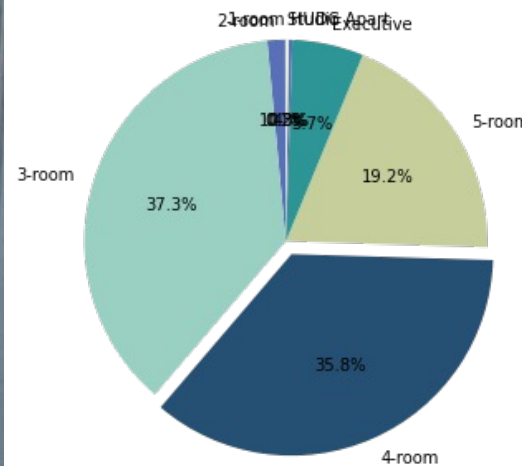
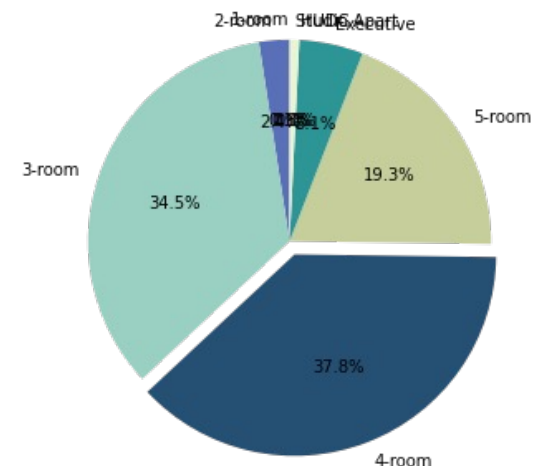*(continued)*
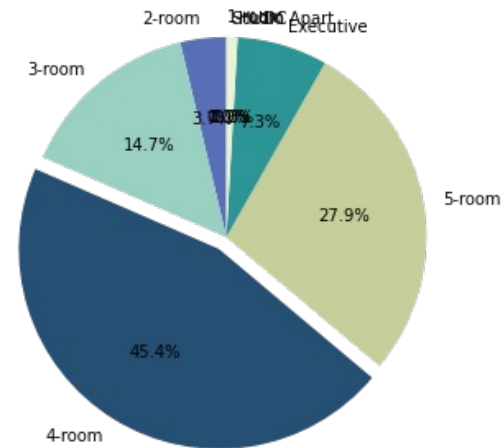
Part 3
## Town Profiles: Mature Vs. New Estates

Insights:
- From 2008 to 2021, HDB sold 194,816 new residential units.
- Of these, 74.2% (144,580 units) were sold in new estates and 25.8% (50,236 units) in mature estates.
- A key distinction between mature and new is their proportions of 3-room flats.
- As of 2021, 3-room flats made up more than a third of flats sold in mature estates compared to just 14.7% in new estates.
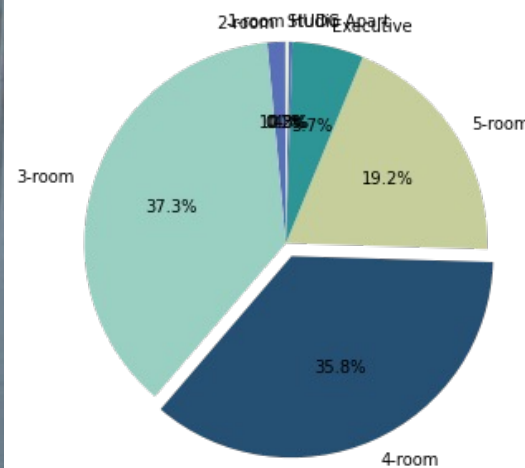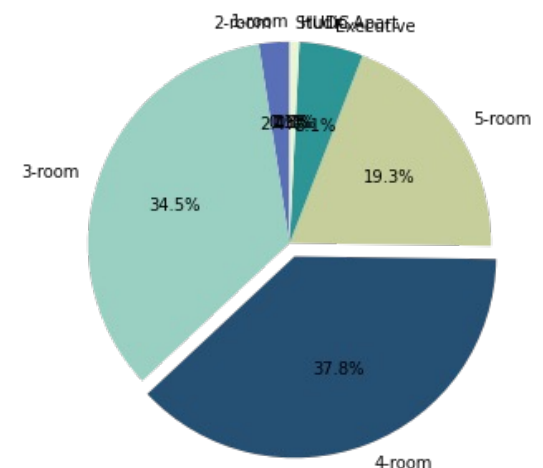- 4-room flats are trending upwards for both estate types.
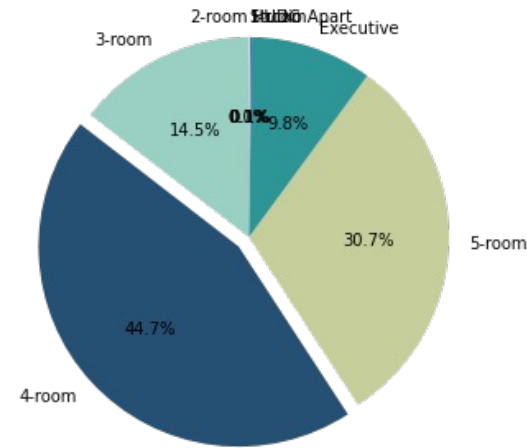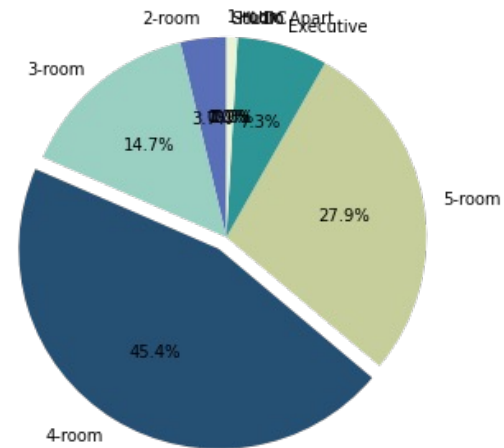
HDB Resale Price Index
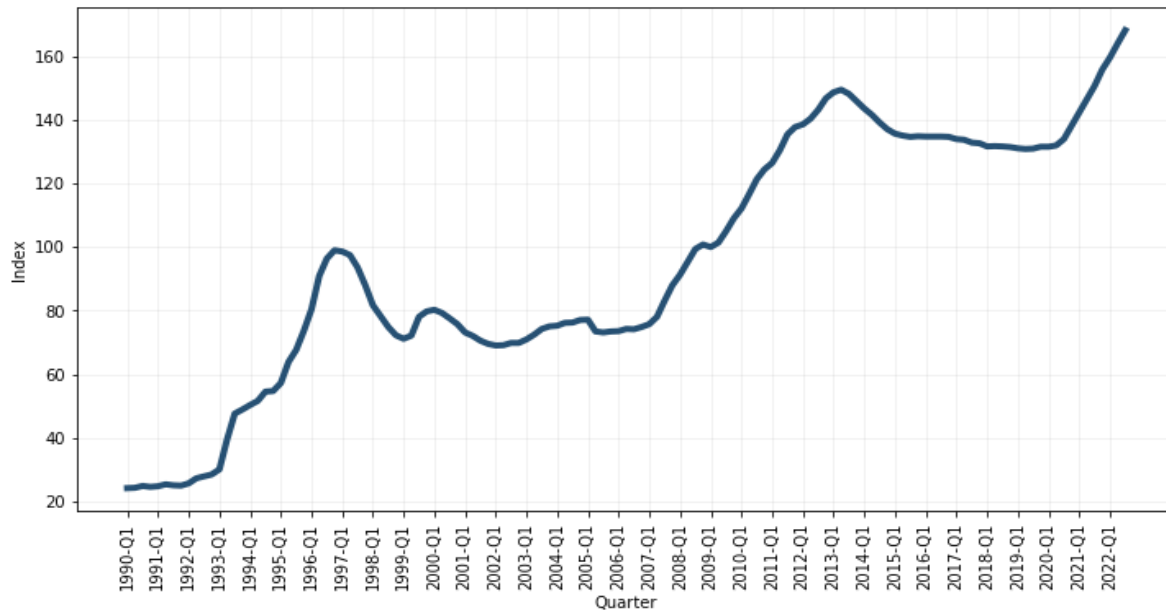
Dataset:
HDB Resale Price Index

Methodology:
- I used numpy to load the CSV file and then plotted a line chart.
- To make the chart more readable, I thickened the line width and added a grid.
- The original ticks in the x-axis were overcrowded. Hence I explicitly defined the number of ticks to be shown.
- Additionally, I rotated the x-labels 90 degrees to declutter the x-axis.

Insights:
- The resale market has broadly trended upwards since 1990.
- There were 3 periods of rapid increase: (i) 1992-1997, (ii) 2007-2014, and (iii) 2020 to present.
- There were 2 periods of major downswings: (i) 1997 Asian Financial Crisis, (ii) 2013-2019 due to cooling measures.

HDB Resale Price Index

Quarterly Return % Distribution
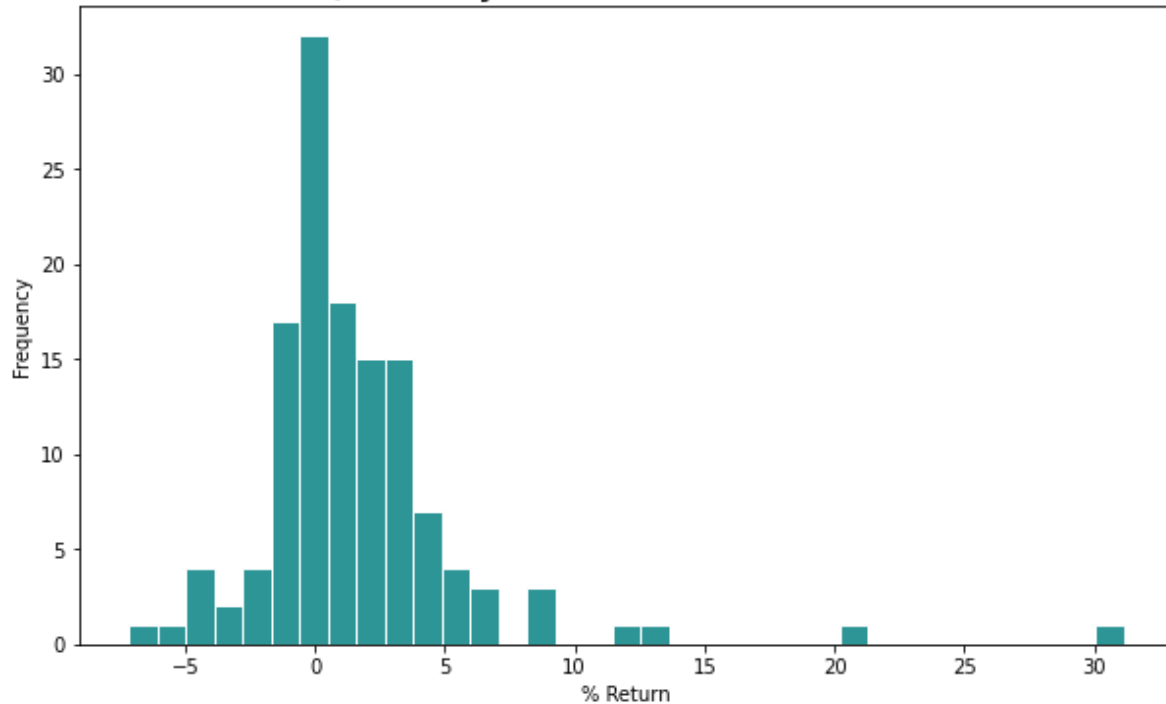
Dataset:
HDB Resale Price Index

Methodology:
- Following on from the previous chart, I plotted a histogram to visualize the quarterly return distribution.
- I used a for loop to calculate the percentage change for each quarter and stored them into a list for plotting.

Insights:
- The returns distribution is positively skewed.
- Most of the quarterly returns cluster around the 0% to 2% range.
- While there are periods of downtrends, quarterly price decreases are not as intense as price increases during uptrends.
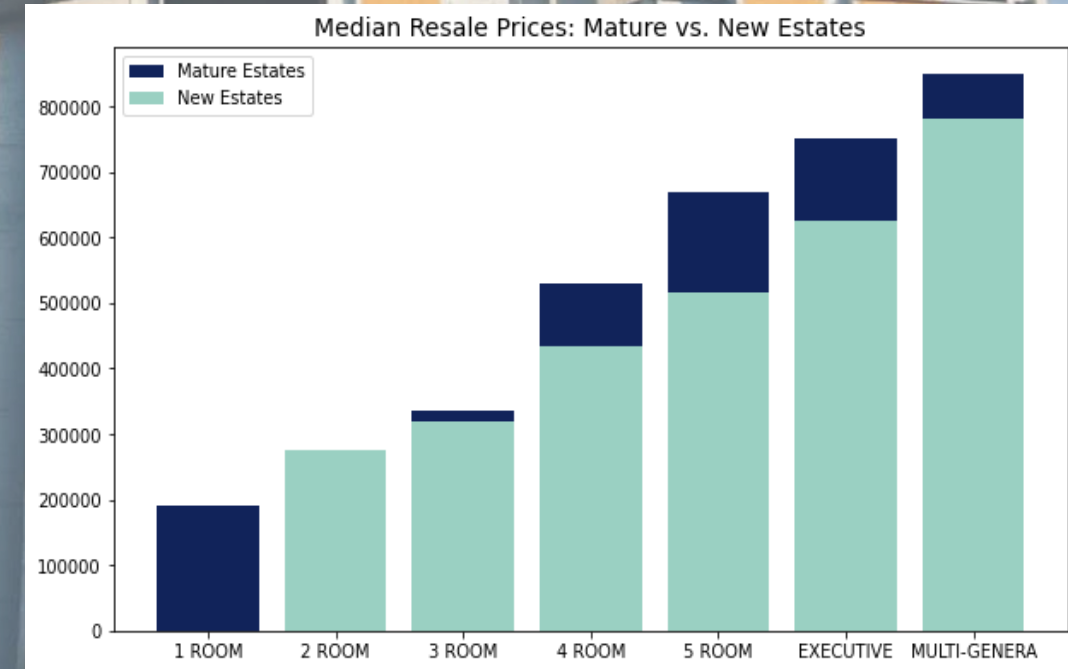
Part 5
HDB Resale Flat Prices

Dataset:
Resale Flat Prices (from Jan 2017 onwards)

Methodology:
- Here I examine the relationship between estate and resale prices.
- I split the dataset into 2: one each for mature and new estates.
- I then wrote a helper function to loop through the unique flat types and stored their respective resale prices in a dictionary.
- Next, I looped through the dictionaries to obtain the median resale price for each flat type and stored them in an empty list.
- This process was done for both mature and new estates.
- Lastly, I plotted a bar chart of the median resale price for each flat type in both estate types.

Insights:
- Almost every flat type in mature estates, commands a price premium over those in new estates.
- There is not much difference between the median 3-room flat sold in both estate types.
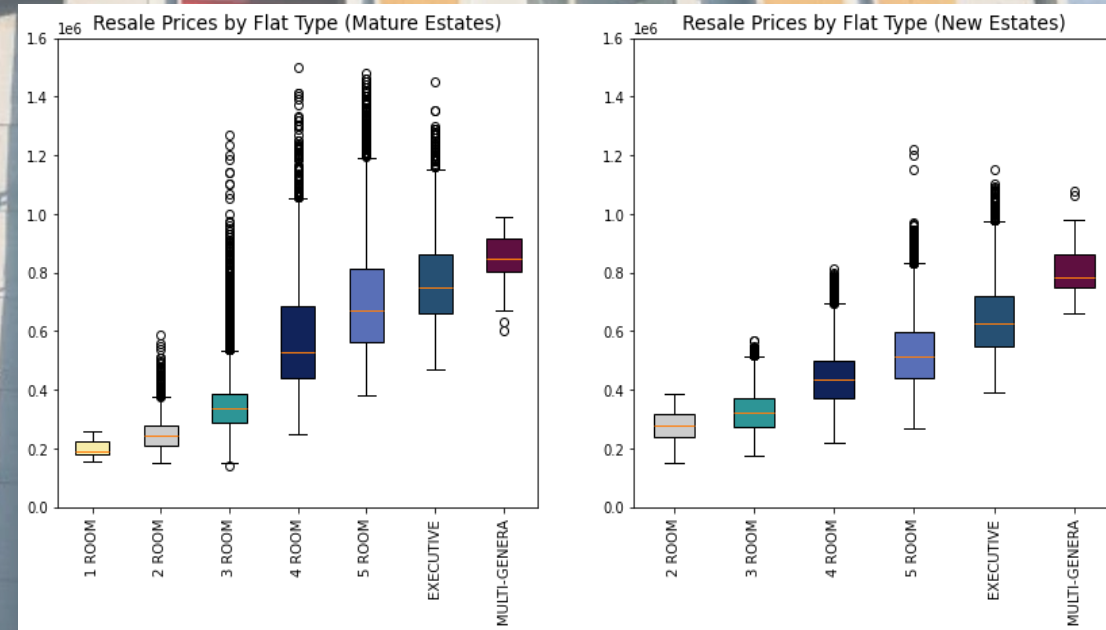
Part 5
HDB Resale Flat Prices

Dataset:
Resale Flat Prices (from Jan 2017 onwards)

Methodology:
- In this sub-section, I use boxplots to do a side by side comparison of resale prices in mature and new estates.
- Data is already organized in dictionaries from the previous slide. I created a subplot of 1 row and 2 columns and plotted 2 boxplots using `axs[i].boxplot`.
- Next, I looped through pre-specified colours palettes and applied them to each flat type.
- To ensure comparability between the chart, I've used `axs[i].set_ylim` to set a common y-axis limit for both of them.

Insights:
- The median resale price in mature estates was higher across all flat types (other than 2-room flats).
- There are price premiums for the median 4-room (+SGD95,000) and 5-room (+SGD155,000) flats in mature estates.
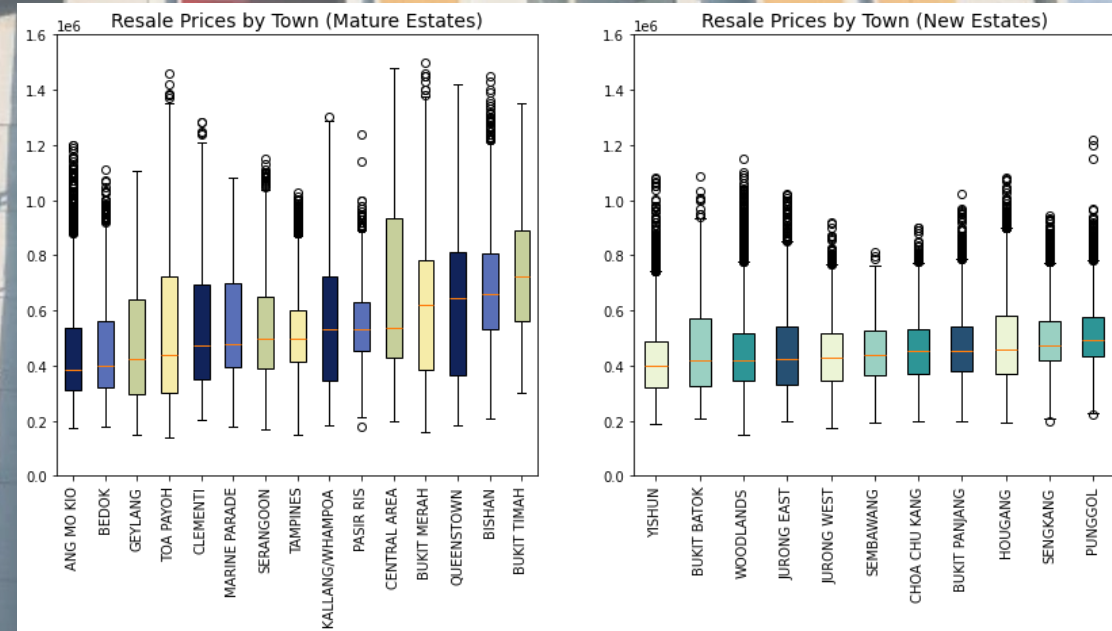
Part 5
HDB Resale Flat Prices

Dataset:
Resale Flat Prices (from Jan 2017 onwards)

Methodology:
- In this sub-section, I split the mature and new estates into individual towns to understand why resale flats in mature estates have greater variance.
- Using the dictionaries I have on hand, I sort them by town, instead of flat type, to give me their respective median flat price.
- The process is the same as the previous slide, but applied to towns.
- One additional step I made was to sort the chart by ascending median prices. To do so, I used the `sorted` function and had to specify the `key` with which to sort the dictionaries.
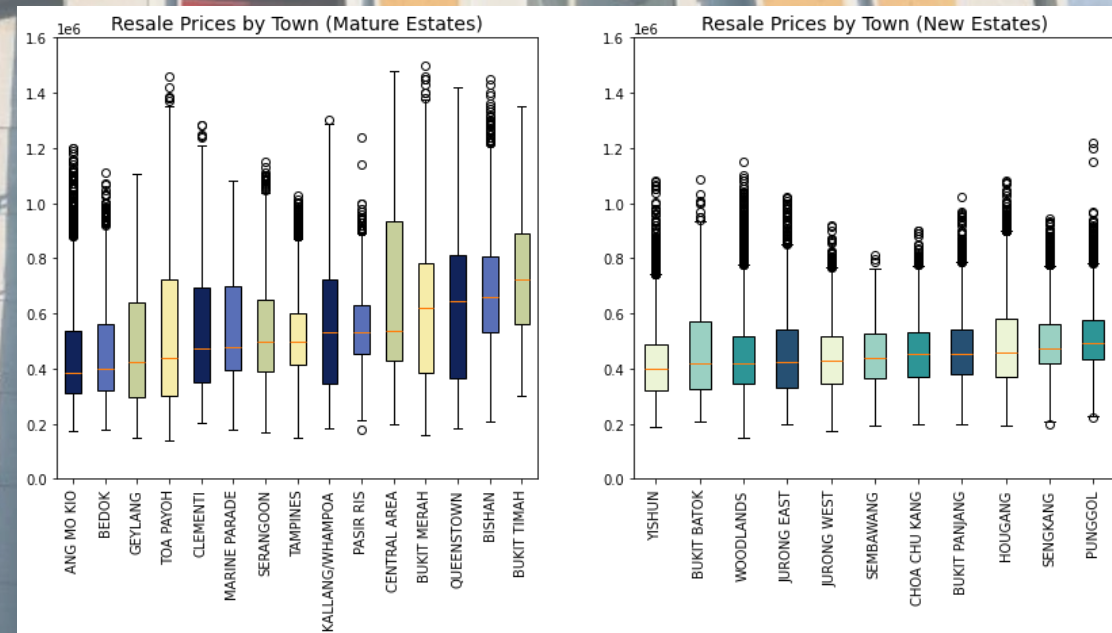
(continued)

Part 5
HDB Resale Flat Prices

Insights:

- There isn't much variance across towns in new estates. The majority of median prices hover around the SGD440,000 range and the standard deviation of these prices is ~SGD27,000.
- In mature estates, the average median resale price is ~SGD520,000 while standard deviation is ~SGD97,000. This is explains the box plots' long bodies.
- Mature estates are more dissimilar as a group than the new estates. This is driven by the individual towns within it having vastly different flat type profiles.
- New estates are all located at a farther distance away from the centre of Singapore. For mature estates, the proximity to central Singapore differs from town to town.
- Hence, the median resale price for mature estates falls along a much wider spectrum than those in new estates resulting in the greater variance in resale prices.

End Of Presentation