HDB Flat Resale Price Analysis

File created on: 8/24/2024 7:19:58 PM

HDB Flat Resale Price Analysis

Which factors could affect the resale price of the HDB flats?

MOTIVATION:

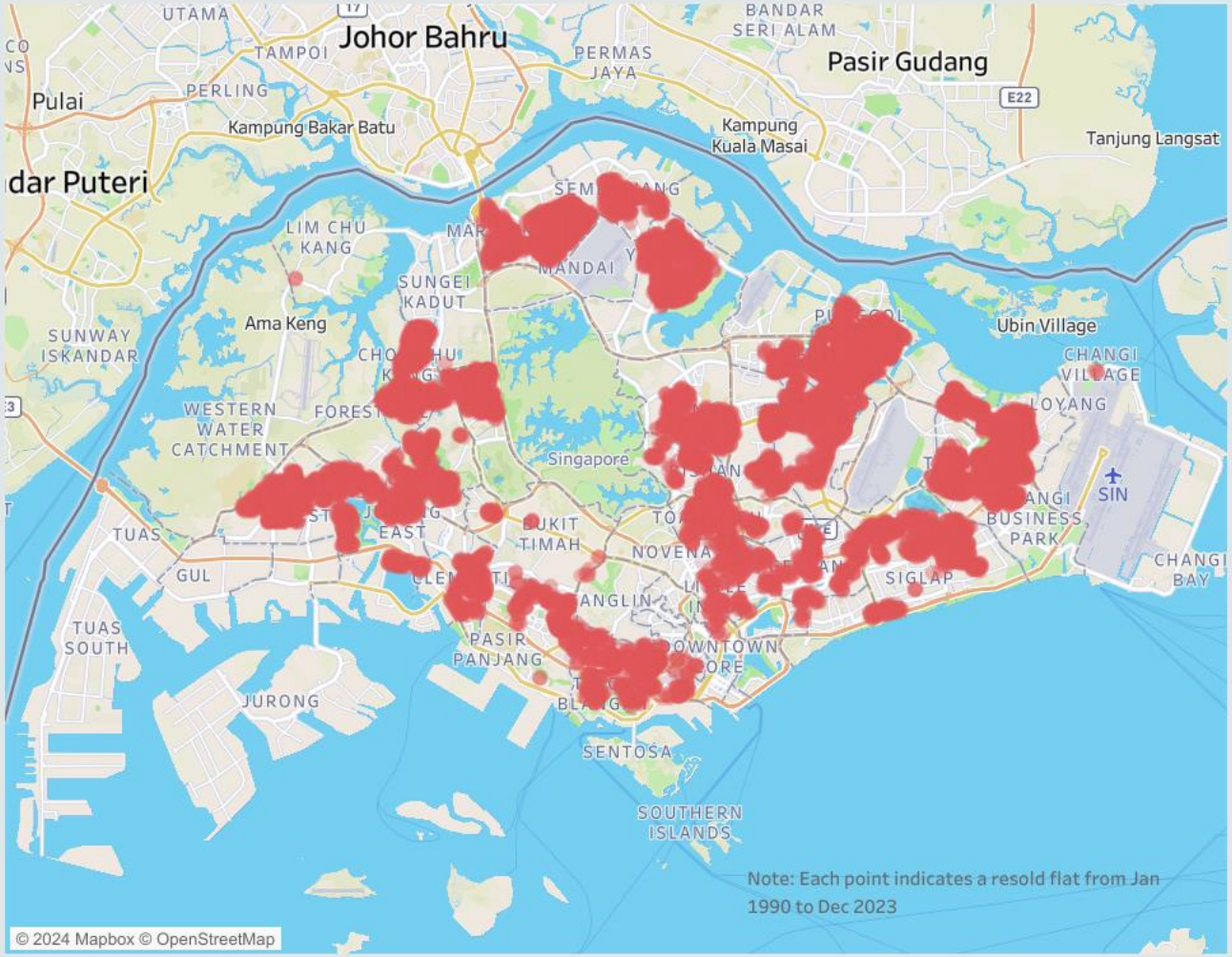
An HDB flat is a public housing unit in Singapore developed and managed by the Housing and Development Board (HDB), offering affordable homeownership to Singaporean citizens. Compared to Build-To-Order (BTO) flats, resale flats offer immediate availability, a variety of locations with established amenities, and greater flexibility in terms of unit type. However, a significant challenge is the rising prices in the resale market, which can be considerably higher than new BTO flats and may limit affordability for some buyers.

OBJECTIVE:

To study the factors that could affect the resale price of the HDB flats based on the data available.

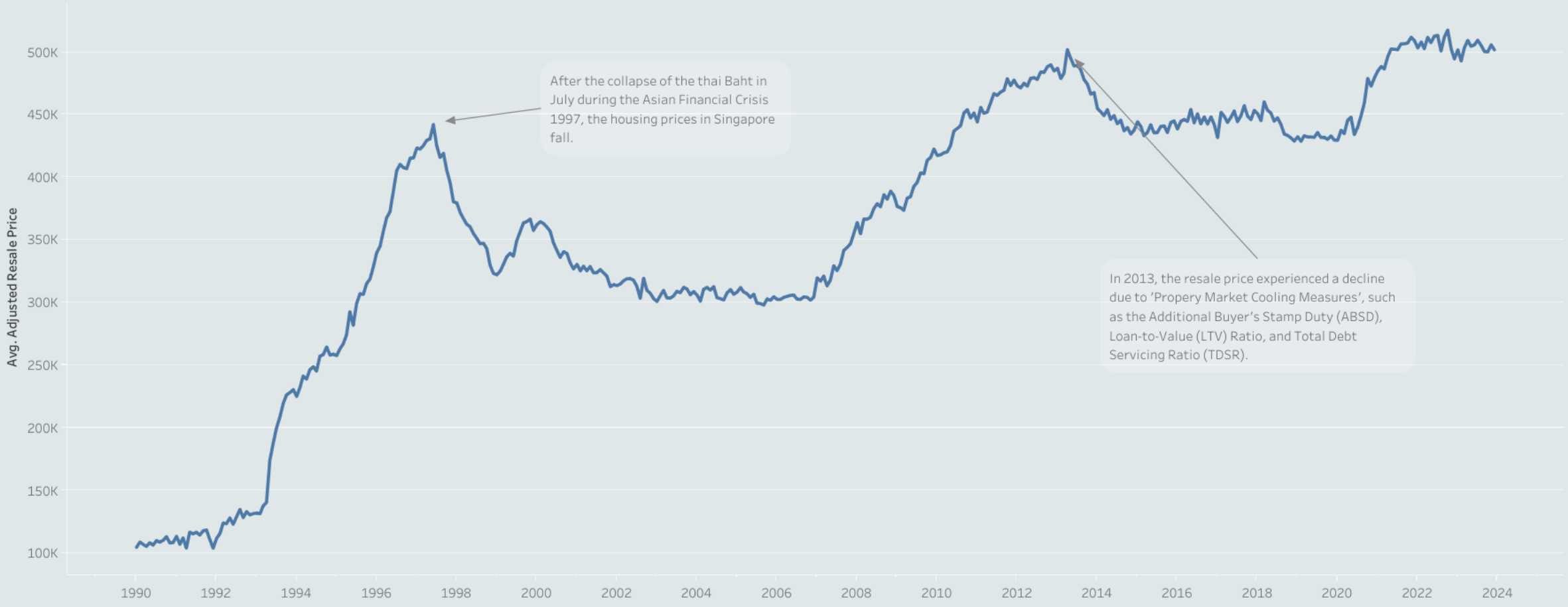
SCOPE:

- The resale flat price dataset is obtained from data.gov.sg.
- The dataset covers the flat resale transactions from Jan 1990 to Dec 2023.
- The resale price in this analysis is adjusted with CPI index to minimize the effect of inflation.



HDB Flat Resale Price Analysis

An Overview of Average Adjusted Resale Price from Jan 1990 to Dec 2023



Generally, it shows an **upward trend** from Jan 1990 to Dec 2023, although there are two major drops happened in 1997 and 2013.

HDB Flat Resale Price Analysis

Adjusted Resale Price against Floor Area

To start off, we check the **linear relationship** among the numerical variables.

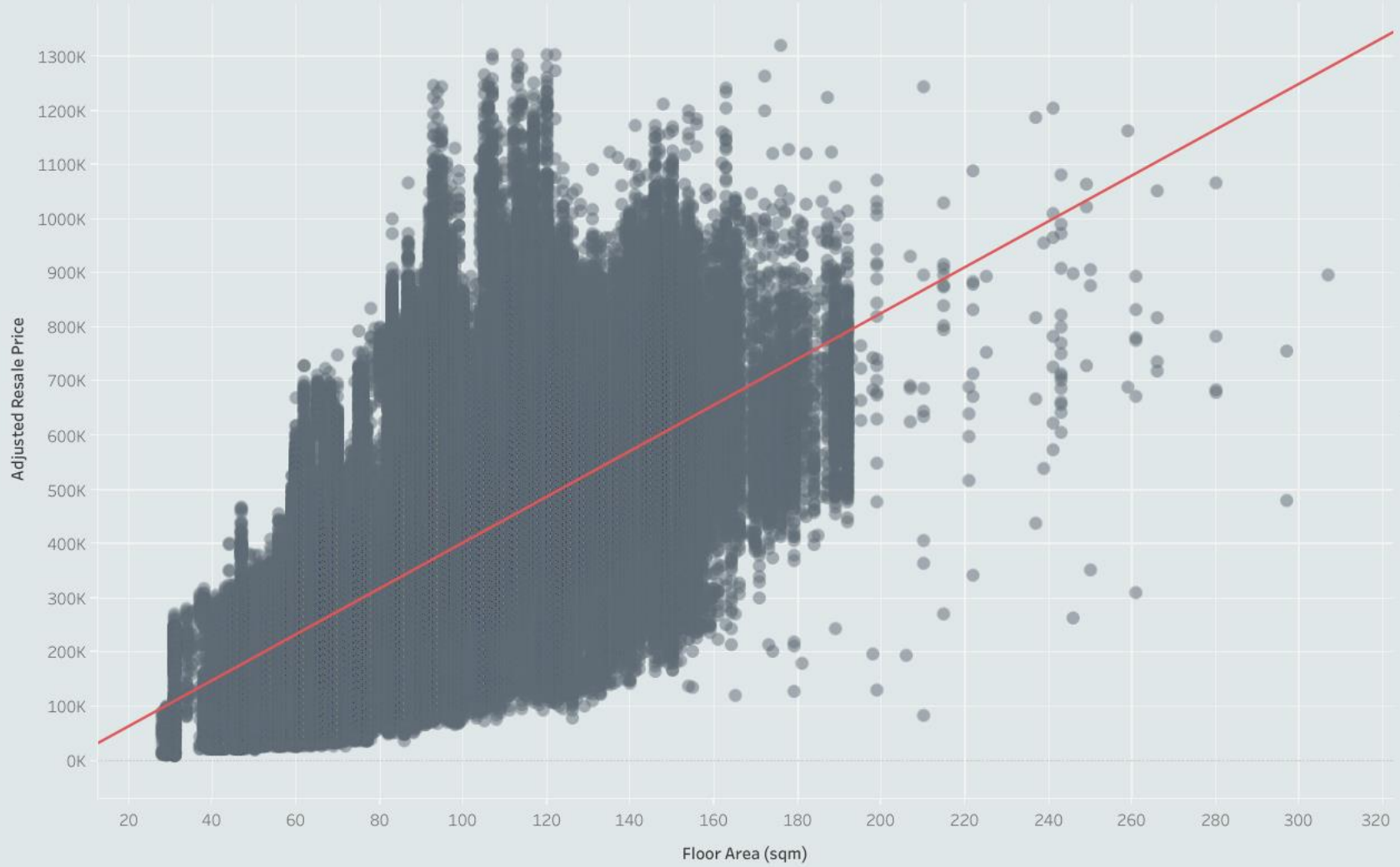
Among the four numerical variables, only floor area shows linear relationship with adjusted resale price.

From the scatterplot, it seems that **the larger the floor area, the higher the adjusted resale price**. This will be the hypothesis to be tested.

Surprisingly, remaining lease, distance to train station, and storey range have very weak linear dependence with the adjusted resale price.

HDB Flat Resale Price Analysis

Adjusted Resale Price against Floor Area



The larger the floor area, the higher the adjusted resale price.

To test the hypothesis above, a **linear regression** is conducted.

The result shows that the floor area can only explain **44%** of variance in adjusted resale price.

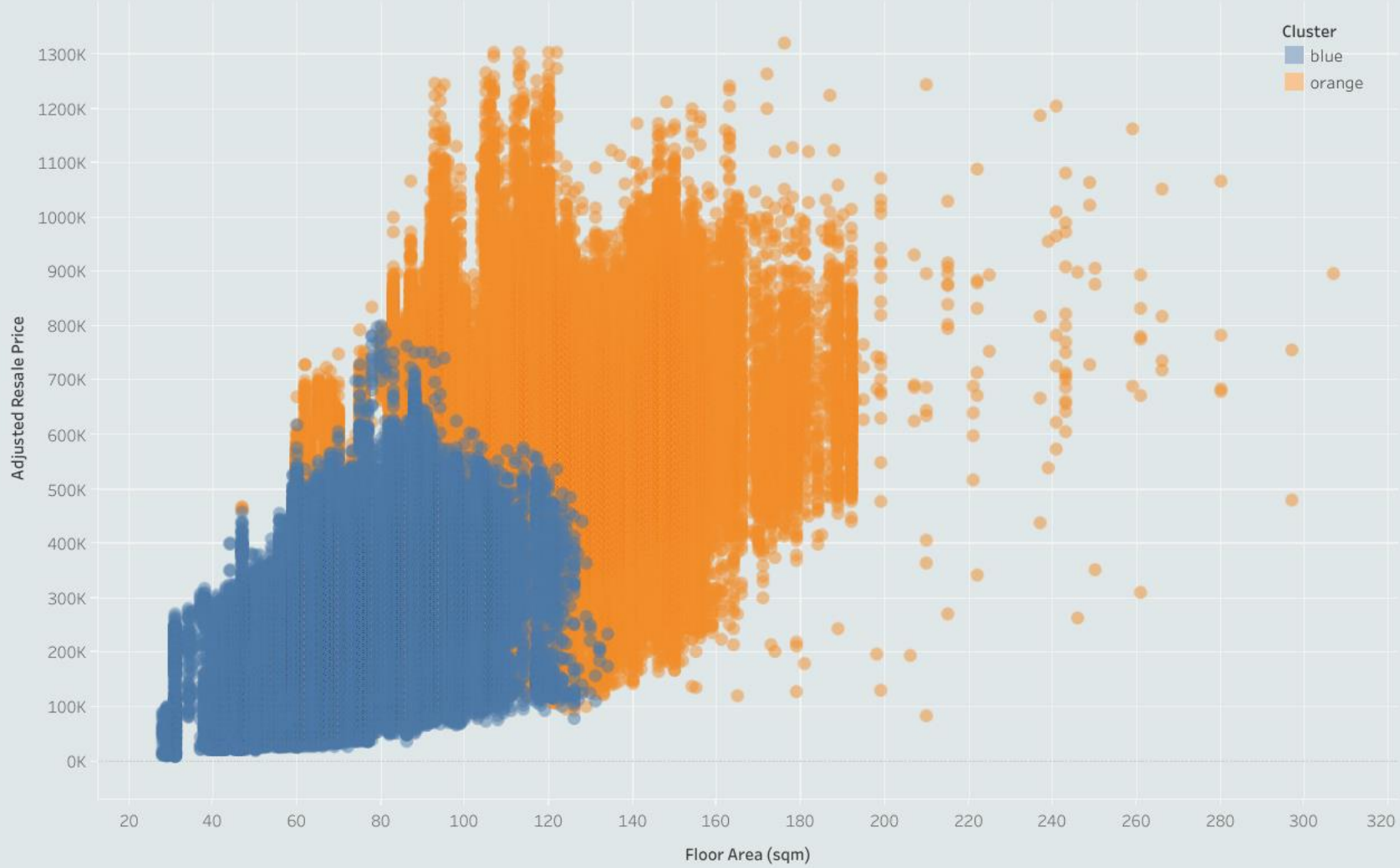
Also, there are many points far apart from the regression line. For instance, when the floor area is between 90 to 120 sqm, there are many data points that indicate high resale price.

This means that the relationship is **not fully linear**.

Since the adjusted resale price cannot be entirely explained by floor area, let's apprach cluster analysis to discover any underlying patterns within the data.

HDB Flat Resale Price Analysis

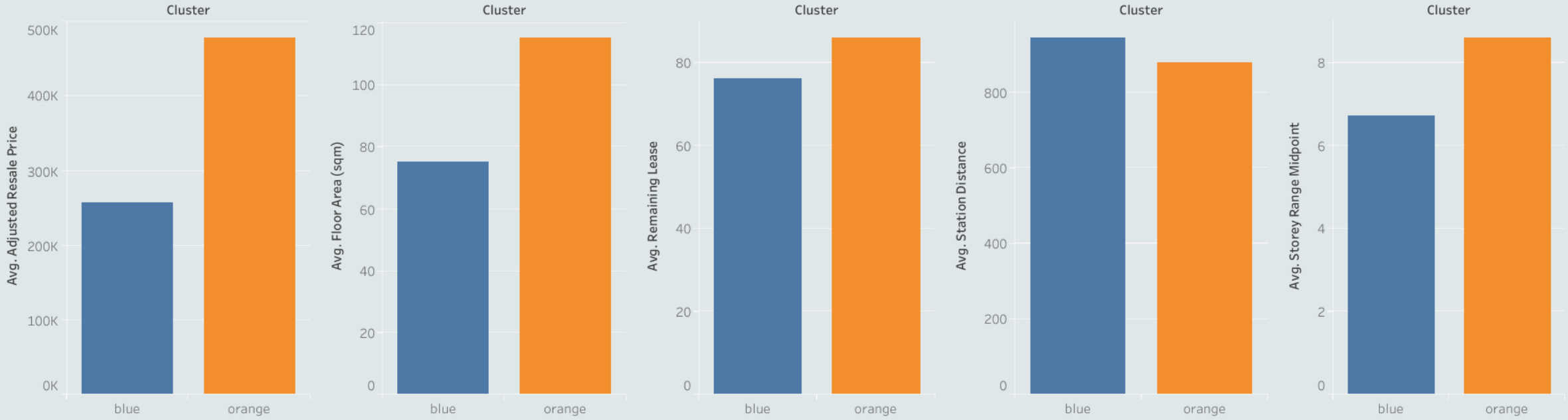
Adjusted Resale Price against Floor Area



A **cluster analysis** is to group similar data points into clusters. We can then compare the numerical metrics of different clusters to uncover new patterns.

Our cluster analysis yielded **two distinct clusters** of data points, which are represented in blue and orange colors.

HDB Flat Resale Price Analysis



From the first bar chart, **orange cluster** has much **higher average adjusted resale price** than blue cluster.

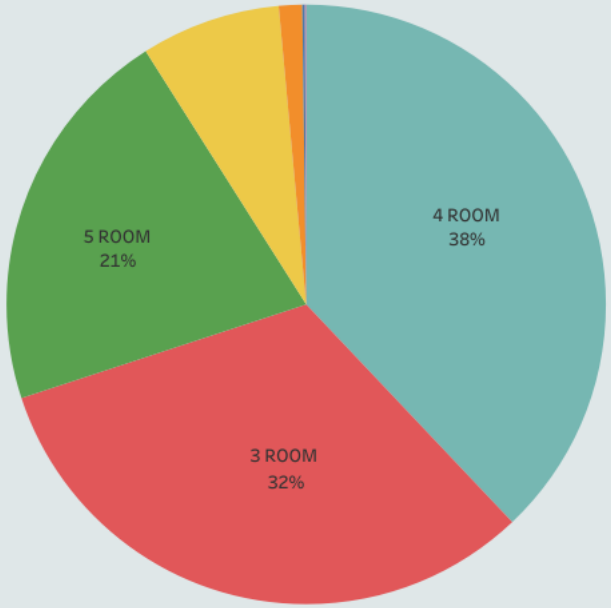
At the same time, orange cluster has:
- **larger average floor area**
- slightly longer average remaining lease
- slightly shorter average distance to train station
- slightly higher average storey range

In short, orange cluster has the better conditions, and thus has higher adjusted resale price.

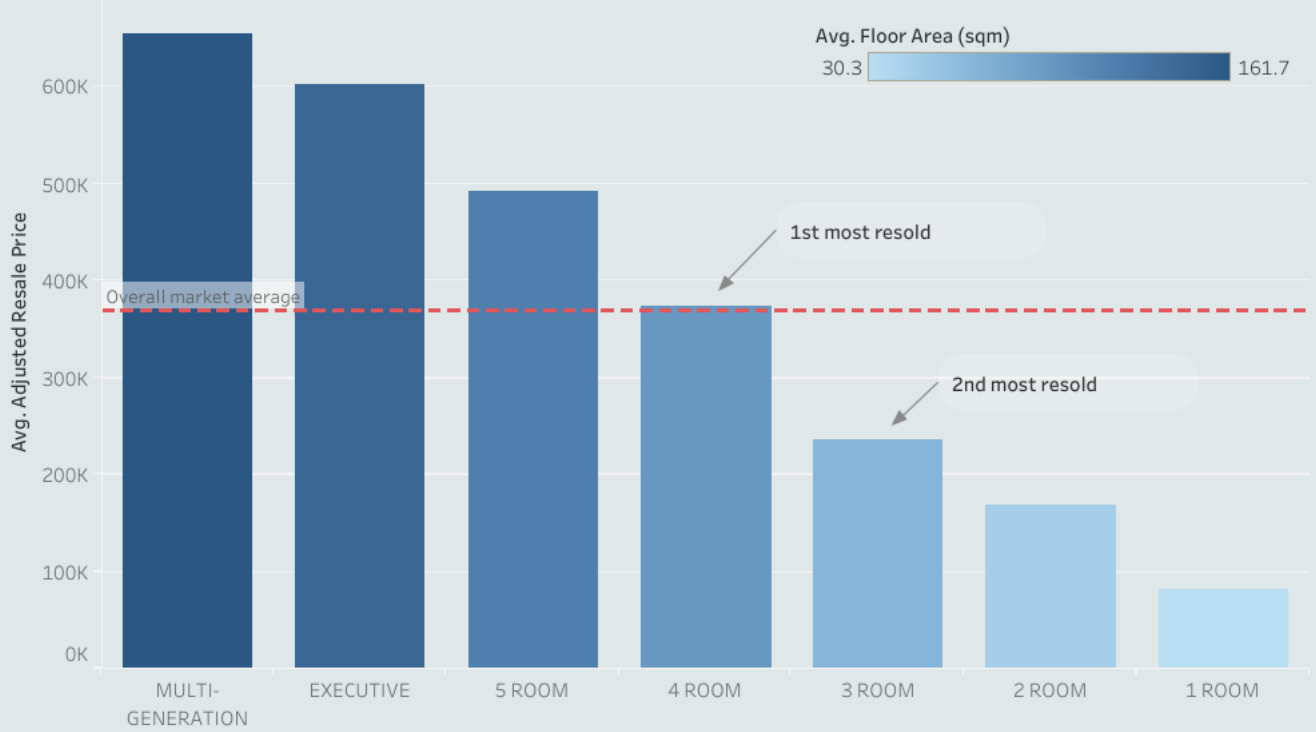
This is aligned with our common expectations, as well as our hypothesis.

HDB Flat Resale Price Analysis

Percentage of Flat Types Resold



Average Adjusted Resale Price and Floor Area across Flat Type



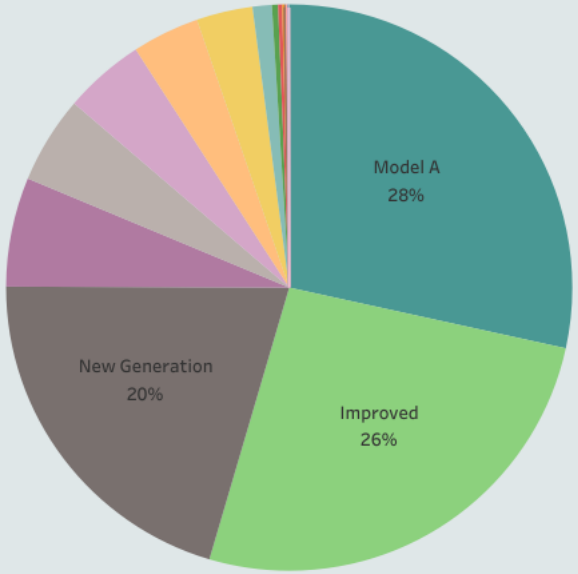
4 room and 3 room are the top 2 most common flat types in the resale market. However, their resale prices are around or below the overall market average.

Multi-generation and **executive** flat types have the highest average adjusted resale price.

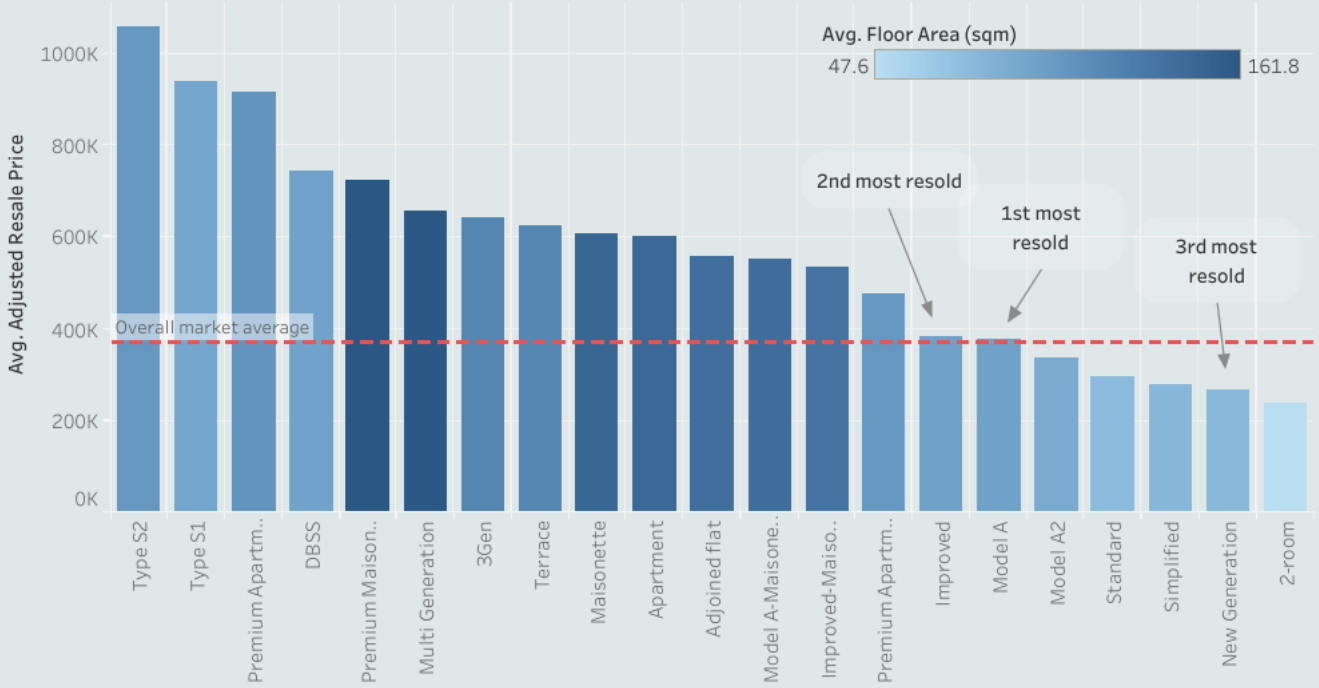
From the color gradient of the bar chart, flat types with larger average floor area have higher average adjusted resale price. This again aligns with our hypothesis.

HDB Flat Resale Price Analysis

Percentage of Flat Models Resold



Average Adjusted Resale Price and Floor Area across Flat Model



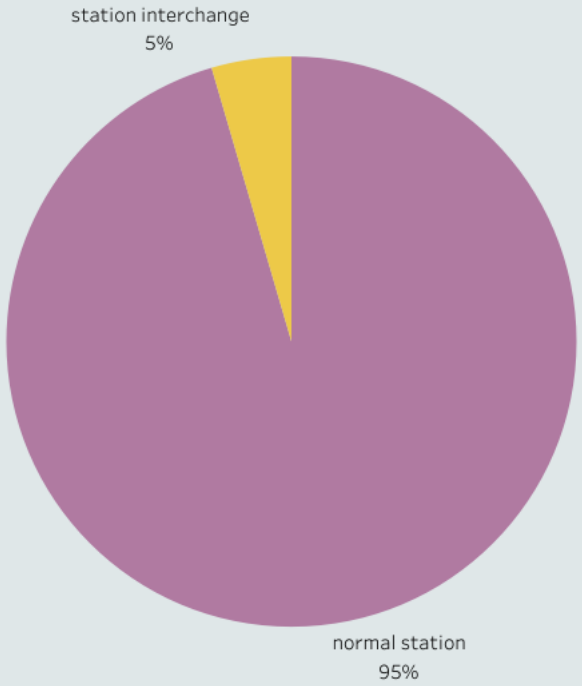
Model A, Improved, and New Generation are the top 3 most common flat models in resale market.

The hypothesis about floor area and adjusted resale price also applies here, but with a few exceptions. For example, the top four highest average adjusted resale price models do not have the largest floor area.

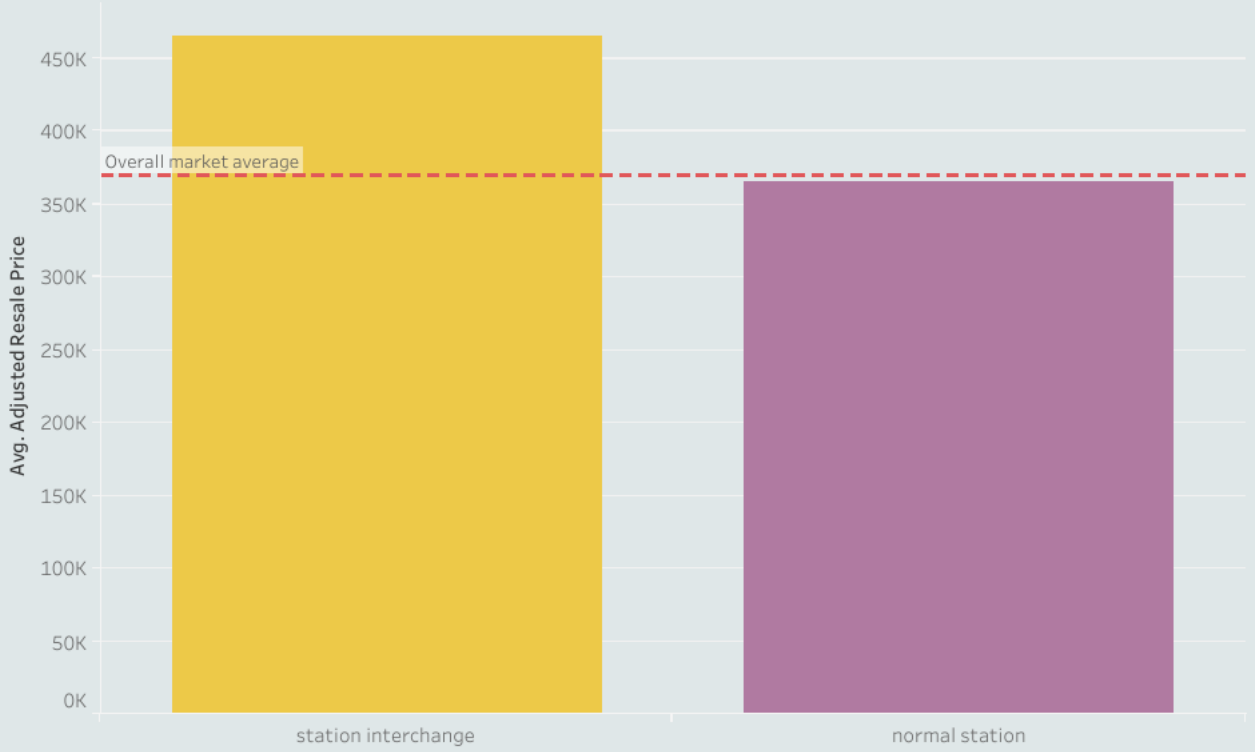
This could due to the transactions of these models happened only in more recent years, at the same time the market price has increased over the years.

HDB Flat Resale Price Analysis

Percentage of Interchange Flags Resold



Average Adjusted Resale Price across Interchange Flag

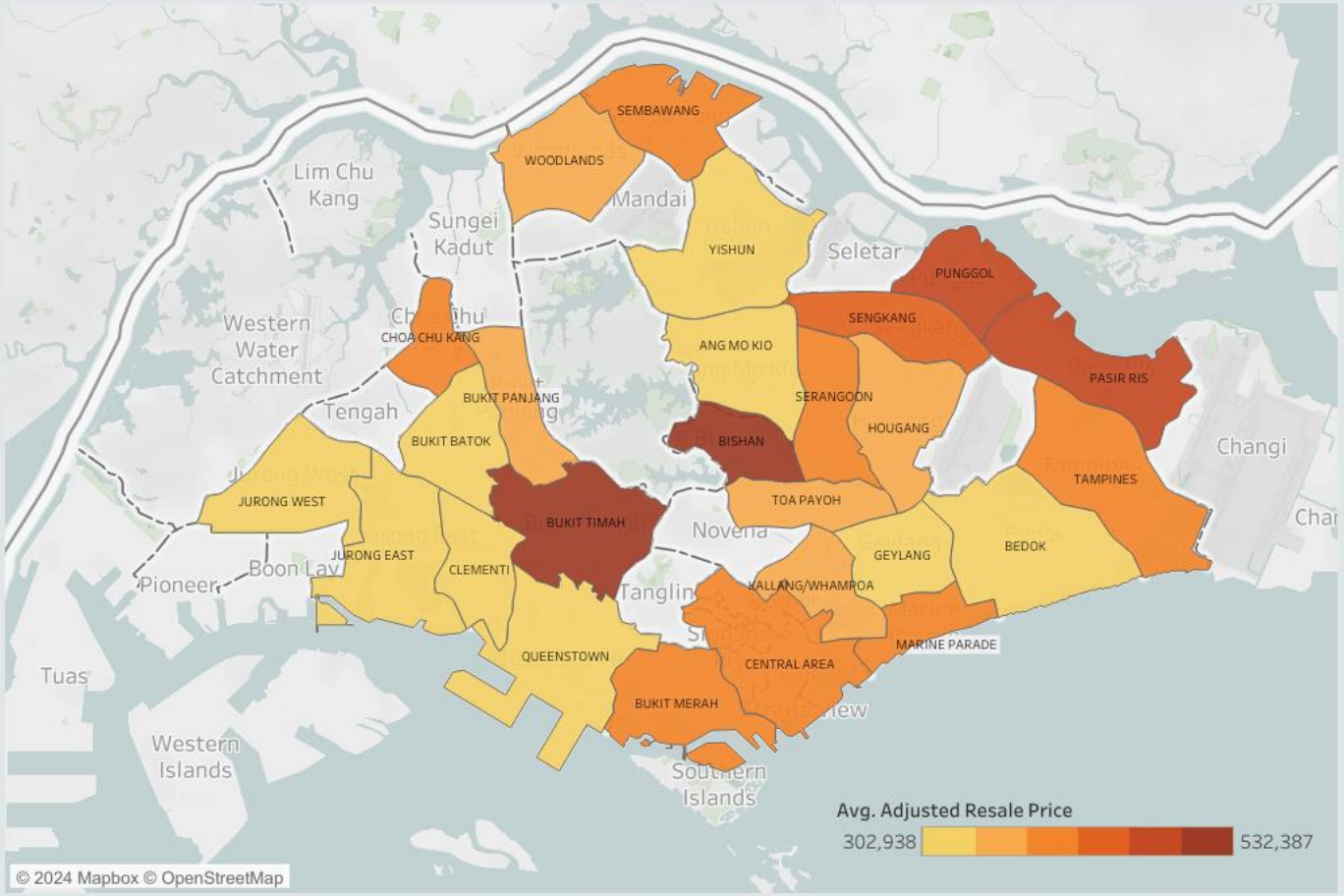


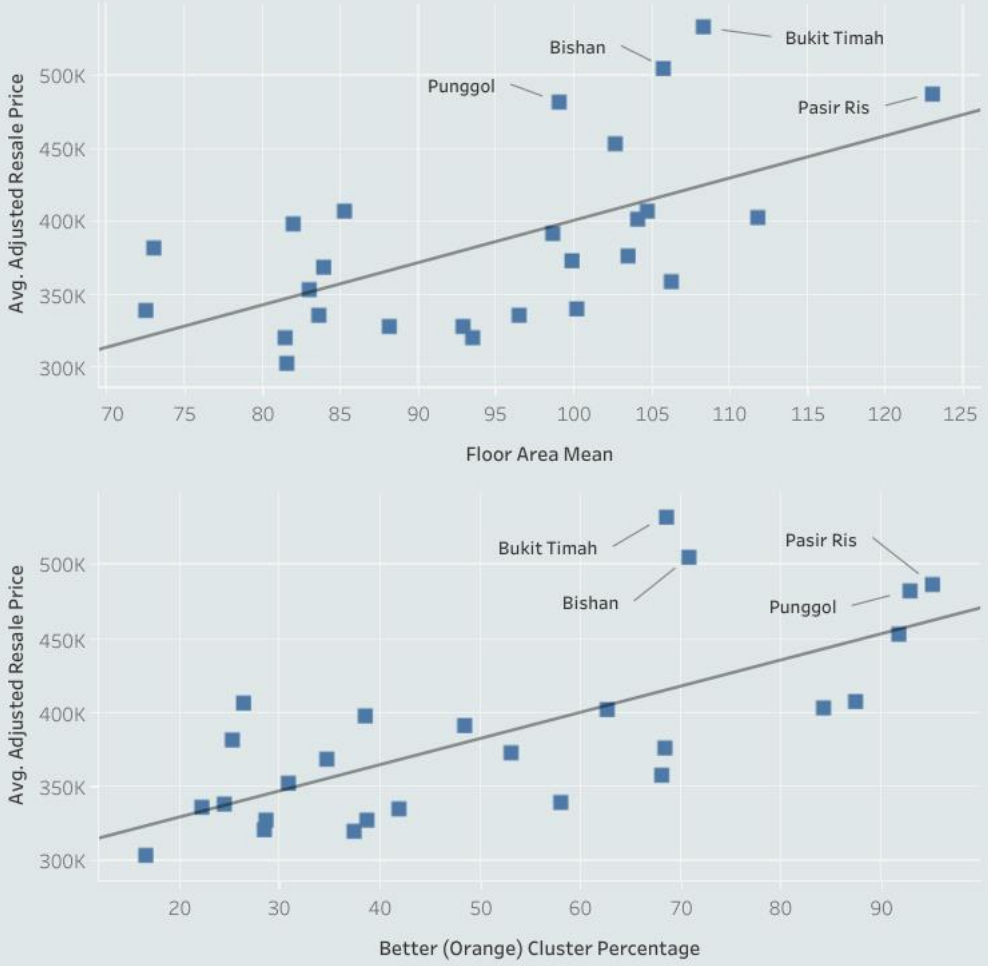
The nearest train station of 95% of resold flats is normal station, while the remaining 5% is station interchange.

However, the average adjusted resale price of **flats with station interchange** as nearest station is higher than the other group, and the overall market average as well. This could be due to the convenience of station interchange that brings commuters to different train travel lines.

HDB Flat Resale Price Analysis

Average Adjusted Resale Price across Towns



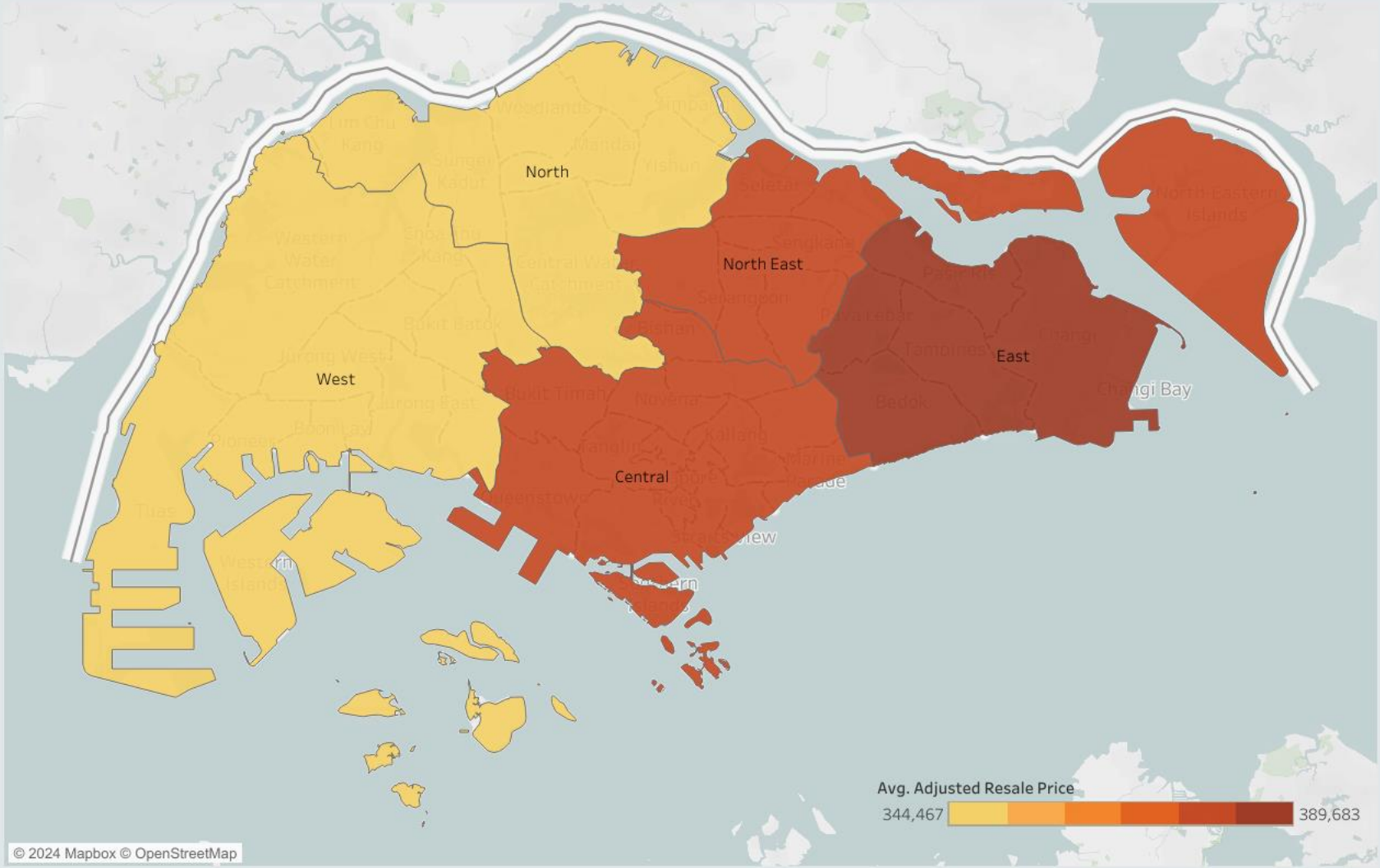


Bukit Timah and **Bishan** have the highest average adjusted resale price, followed by **Pasir Ris** and **Punggol**.

This is mainly due to their **higher average floor area** and **percentage of better cluster**.

HDB Flat Resale Price Analysis

Average Adjusted Resale Price across Regions



The average adjusted resale price **increases from the west side to the east side**.

However, the difference of the highest and lowest price average is only about $45,000, which is **not much significant**.

HDB Flat Resale Price Analysis

CONCLUSIONS:

• Resale price of a flat is positively affected by its **floor area**.
• Longer remaining lease, shorter distance to train station, and higher storey range tend to increase the resale price.
• **Multi-generation** and **executive** flat types are the most expensive flat types.
• **Type S1**, **Type S2**, and **premium apartment loft** are the most expensive flat models.
• A **station interchange** would increase the resale price of its nearby flats.
• **Bukit Timah** and **Bishan** have the highest average adjusted resale price, followed by **Pasir Ris** and **Punggol**.
• The average adjusted resale price **increases from the west side to the east side**. However, the difference is not much significant.

LIMITATIONS OF THIS CASE STUDY:

• The number of variables in this analysis is limited. The resale price could be affected by many other factors such as flat condition, the supply and demand of Built-to-Order flats (BTOs), distance to other amenities such as shopping mall, loan interest rates, and more. Overlooking these influential variables could lead to an incomplete understanding of the market dynamics.
• There are 1,890 duplicates found in the dataset. Since there is no transaction ID for each record, it's hard to classify them as errors. Therefore, these duplicates are kept to use.
• There are 137 flats failed to extract their coordinates from OneMap API. This might be due to the demolition of the blocks under the Selective En bloc Redevelopment Scheme (SERS) program. For this reason, their coordinates are obtained manually from Google Map using their postal codes. If the postal code is not available and Google Map cannot recognize the address, the coordinates of the nearest block based on block number will be used. Although this would be less accurate, it at least shows the approximate locations of these flats.

NEXT STEPS:

• Source and integrate the data of other factors into the main dataset and analyze their impacts towards resale price.
• As floor area alone cannot fully explain the variation in resale price, multivariate analysis should be performed to gain a deeper understanding of the effects of each variable towards the resale price.