Resale HDB Price Analysis – Does proximity to expressways impact value?

GovTech DSAID Quantitative Strategy Technical Assessment

Joshua Gei

Situation

Some forum posters have complained that the value of their HDB flats suffer because they are near expressways, which are very noisy.

Others say expressway proximity is good, due to the unblocked view (at least for higher floors).

Based on this feedback, HDB has tasked your team to analyse whether there is merit to either view, based on transaction prices for resale HDB flats in recent years.

Key Outcomes

- Assess and evaluate the **value** of expressway proximity on HDB resale prices
- Assess the relative importance of various factors for HDB flats near expressways
  - e.g. is **elevation** is more important for flats close to expressways?
- Provide high-level recommendations and next steps to HDB for further study on this topic

Executive Summary – 3 main takeaways

1. On an aggregate level across all flats, the impact of Expressway proximity is at parity

2. Proximity to certain specific Expressways have more significant impact than others

3. Higher floors add even more value for properties close to Expressways

Overview and Methodology

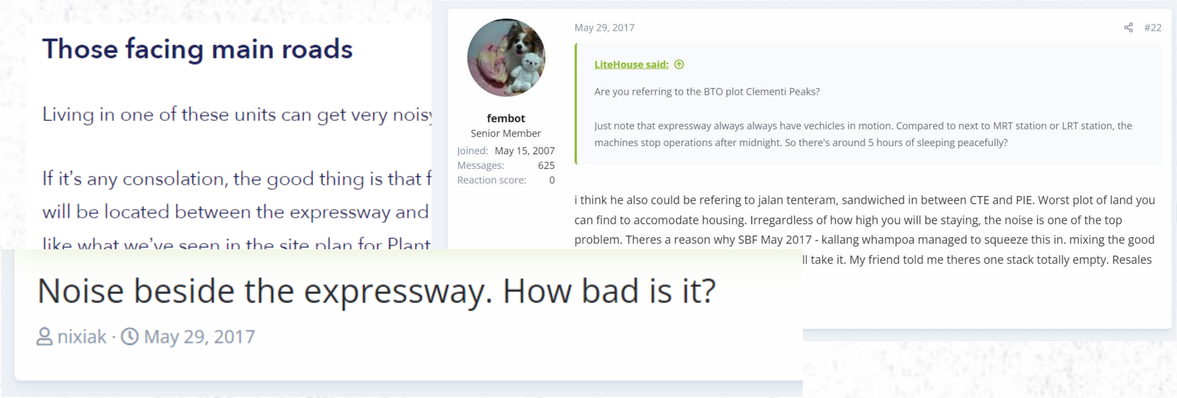
Overview: HDB Flats near Expressways: Pros and Cons

Utilized qualitative data from forums, articles, and reports to build a profile of the ground sentiment prior to deep-diving into the data

Pros:

Unobstructed view (for higher floors)

Tend to have multi-storey carparks between expressway and residential blocks (i.e. enables easier parking)



Those facing main roads

Living in one of these units can get very noisy

If it's any consolation, the good thing is that f
will be located between the expressway and
like what we've seen in the site plan for Plant

May 29, 2017 #22

fembot
Senior Member
Joined: May 15, 2007
Messages: 625
Reaction score: 0

LiteHouse said:

Are you referring to the BTO plot Clementi Peaks?

Just note that expressway always always have vechicles in motion. Compared to next to MRT station or LRT station, the machines stop operations after midnight. So there's around 5 hours of sleeping peacefully?

i think he also could be refering to jalan tenteram, sandwiched in between CTE and PIE. Worst plot of land you can find to accomodate housing. Irregardless of how high you will be staying, the noise is one of the top problem. Theres a reason why SBF May 2017 - kallang whampoa managed to squeeze this in. mixing the good
ll take it. My friend told me theres one stack totally empty. Resales

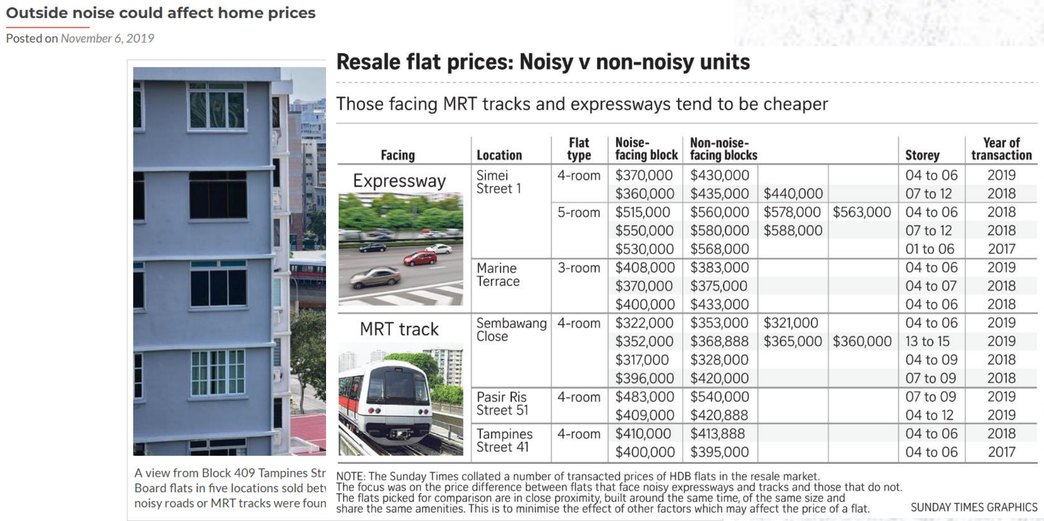
Noise beside the expressway. How bad is it?

nixiak · May 29, 2017

Cons:

Noisy & dusty, especially for lower floors and units facing the expressway due to traffic

Expressways worse than MRT stations/main roads (because of bikes making 'znggg' sounds, compared to constant drone of white noise)



Outside noise could affect home prices

Posted on *November 6, 2019*

A view from Block 409 Tampines Str
Board flats in five locations sold betv
noisy roads or MRT tracks were foun

Resale flat prices: Noisy v non-noisy units

Those facing MRT tracks and expressways tend to be cheaper

Expressway

MRT track

Facing	Location	Flat type	Noise-facing block	Non-noise-facing blocks			Storey	Year of transaction
Expressway	Simei Street 1	4-room	$370,000	$430,000			04 to 06	2019
			$360,000	$435,000	$440,000		07 to 12	2018
		5-room	$515,000	$560,000	$578,000	$563,000	04 to 06	2018
			$550,000	$580,000	$588,000		07 to 12	2018
			$530,000	$568,000			01 to 06	2017
	Marine Terrace	3-room	$408,000	$383,000			04 to 06	2019
			$370,000	$375,000			04 to 07	2018
			$400,000	$433,000			04 to 06	2018
MRT track	Sembawang Close	4-room	$322,000	$353,000	$321,000		04 to 06	2019
			$352,000	$368,888	$365,000	$360,000	13 to 15	2019
			$317,000	$328,000			04 to 09	2018
			$396,000	$420,000			07 to 09	2018
	Pasir Ris Street 51	4-room	$483,000	$540,000			07 to 09	2019
			$409,000	$420,888			04 to 12	2019
	Tampines Street 41	4-room	$410,000	$413,888			04 to 06	2018
			$400,000	$395,000			04 to 06	2017

NOTE: The Sunday Times collated a number of transacted prices of HDB flats in the resale market. The focus was on the price difference between flats that face noisy expressways and tracks and those that do not. The flats picked for comparison are in close proximity, built around the same time, of the same size and share the same amenities. This is to minimise the effect of other factors which may affect the price of a flat.

SUNDAY TIMES GRAPHICS

No consensus on whether proximity to Expressways is good or bad; though sentiment appears to lean negative (but worthwhile noting that forums are known to overrepresent negative comments[1])

Methodology – Datasets used and Preprocessing

Datasets used:

1. HDB resale dataset

- Cleaned & processed (corrected for inflation & feature engineering)
- Geocoded (with OneMap API)
- Filtered for data from **2018 onwards,** since the study is focused on 'recent years'

2. National Map Line dataset

- Cleaned & processed
- Compared with HDB resale dataset to obtain the nearest Expressway to each HDB flat, and the corresponding distance

Primary metric for comparison used in this study will be resale price/sqm – since this is roughly the same across the primary flat types (3/4/5-room – refer to Annex), hence enables us to use a wider set of data for analysis as opposed to e.g. only one flat type.

Sample row in processed dataframe. The key features used for this study are highlighted in light blue:

latitude	longitude	address	town	flat_type	flat_model	floor_area_sqm	remaining_lease	lease_commence_date	resale_date	storey_range	resale_price_with_inf	resale_price/sqm	nearest_line	distance
1.377259	103.748292	818B CHOA CHU KANG AVE 1	CHOA CHU KANG	4 ROOM	Model A	92.0	93.833333	2017	2022-12-01	19 TO 21	580000.000000	6304.347826	KRANJI EXPRESSWAY	1.223560

Using OneMap API

From HDB Dataset

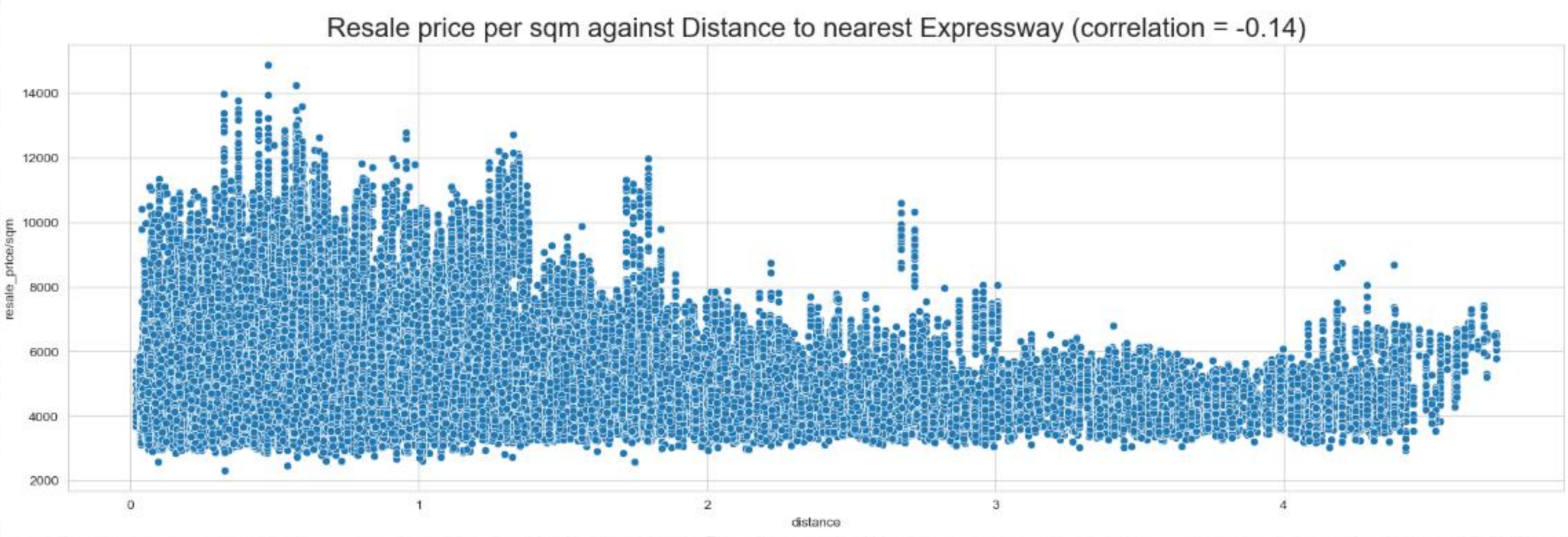
From National Map Line

Details of data preprocessing can be found in the attached notebooks

Observation 1:

On an aggregate level, the impact of Expressway proximity is at parity

On an aggregate level, the correlation between distance to nearest expressway and HDB resale price is minimal



- **Correlation of -0.14** implies there is no clear one-sided impact that proximity to expressway has on resale prices across all resale flats

- An observation here is that property 3-4km away or further do not reach as high prices, but that's likely an effect of being in the heartlands, whereas central areas tends to be closer to some expressway or another (and centrality is known to be a big factor for property prices in SG)

NEW SLIDE (updated post-submission)

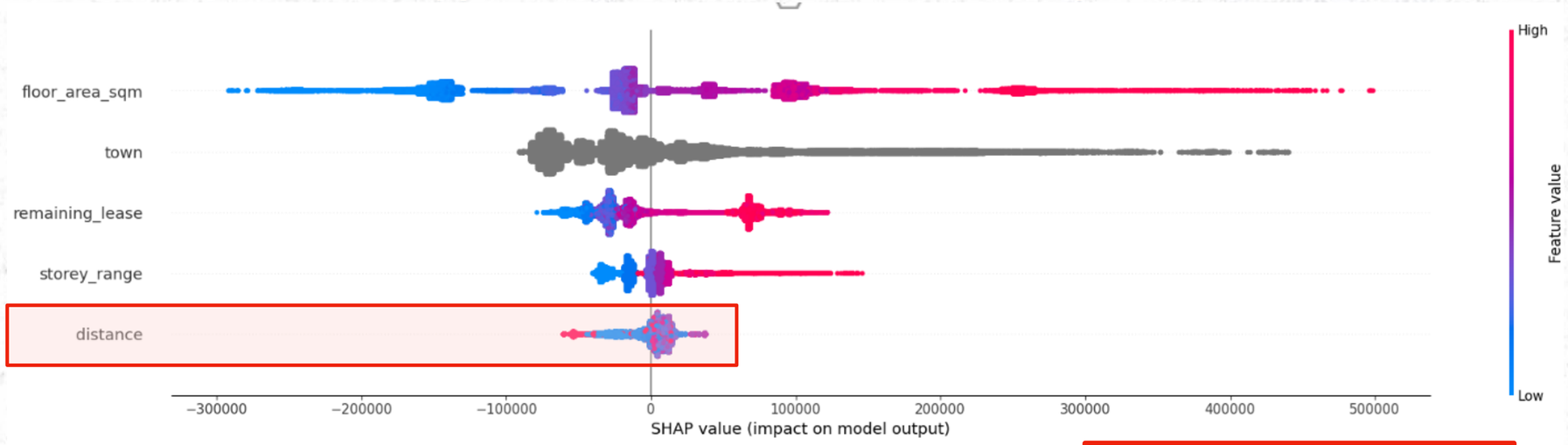
Applying SHAP further shows that the impact of Expressway Proximity is generally minimal compared to other factors

Used a regression model to predict resale prices, and applied SHAP to show the relative importance of each factor:

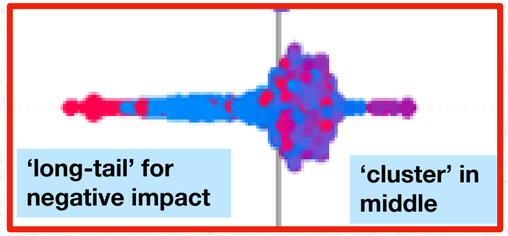
Model:
- CatBoostRegressor (R^2 of 0.88)

Features used:
- floor_area_sqm
- town *(categorical feature)*
- remaining_lease
- storey_range
- distance to nearest expressway

Output variable:
- resale_price



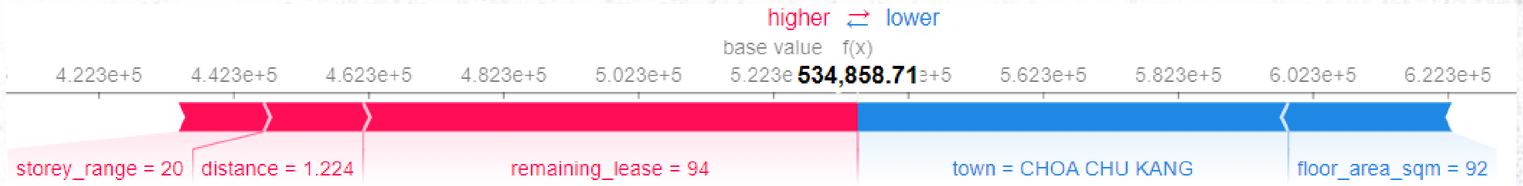
- The magnitude for each feature in the figure above shows their relative importance (i.e. floor area is most important, then town, etc.)
- Expressway proximity is a small factor (negligible for most flats, shown by the cluster in the middle)
- However, the **'long-tail'** on the negative side shows there could be some cases where it has a larger impact



What is SHAP?

SHAP (SHapley Additive exPlanations) is a method to explain the output of a machine learning model by assigning each feature an importance value for a **particular local prediction**

Example of a local instance interpretability (i.e. one resale flat transaction)



address	818B CHOA CHU KANG AVE 1
town	CHOA CHU KANG
flat_type	4 ROOM
flat_model	Model A
floor_area_sqm	92.0
remaining_lease	94
lease_commence_date	2017
storey_range	20
resale_price_with_inf	580000.0
resale_price/sqm	6304.347826
nearest_line	KRANJI EXPRESSWAY
resale_date	2022-12-01 00:00:00
distance	1.22356

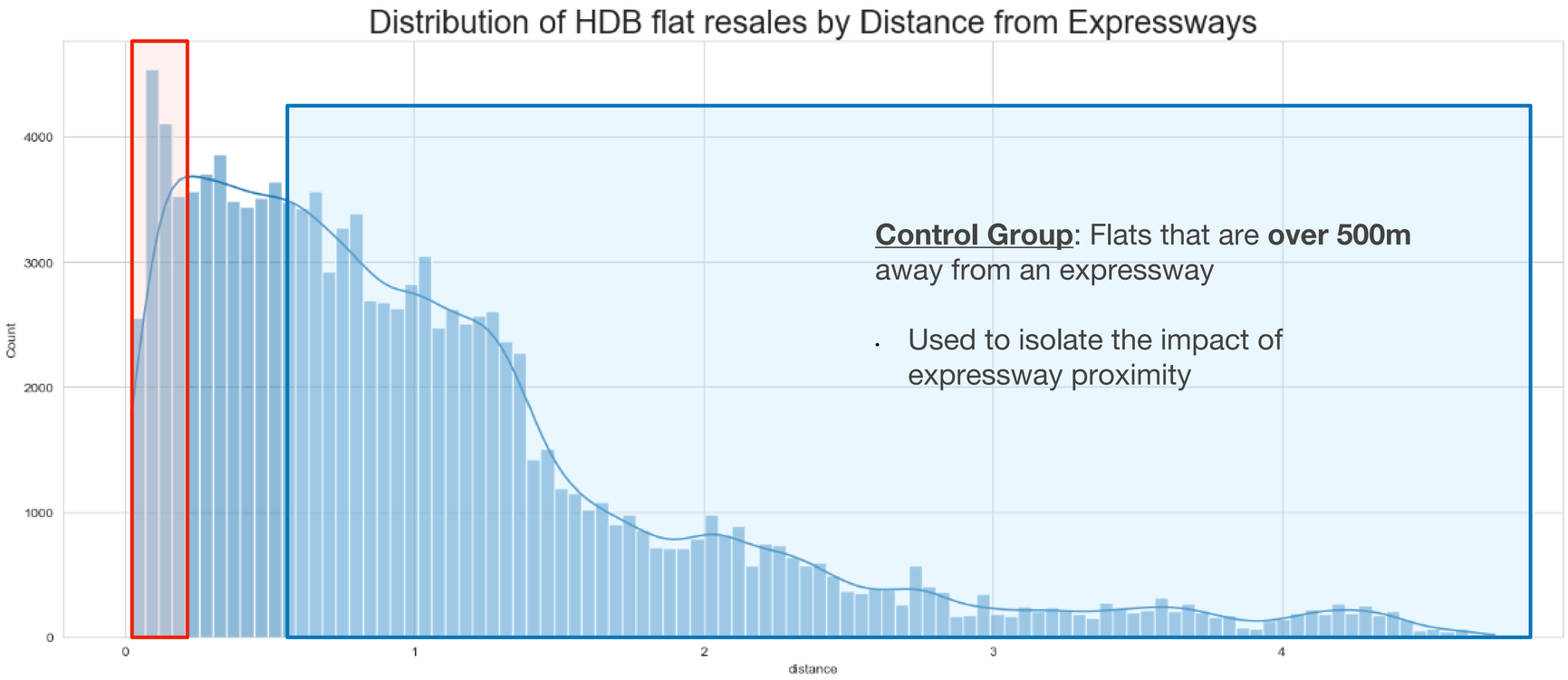
Hence, more granular analysis on the most-affected HDB units could yield more conclusive and specific results

Further Analysis performed will include splitting the resale data into two groups:

1. A **target group** that is very close to an Expressway and heavily impacted by both the pros and cons of being close

2. A **control group** that is sufficiently far away from the Expressway, and assumed to be unimpacted by proximity

(i.e. the impact of being 500m away from an Expressway and 4km away is the same – both are negligible)



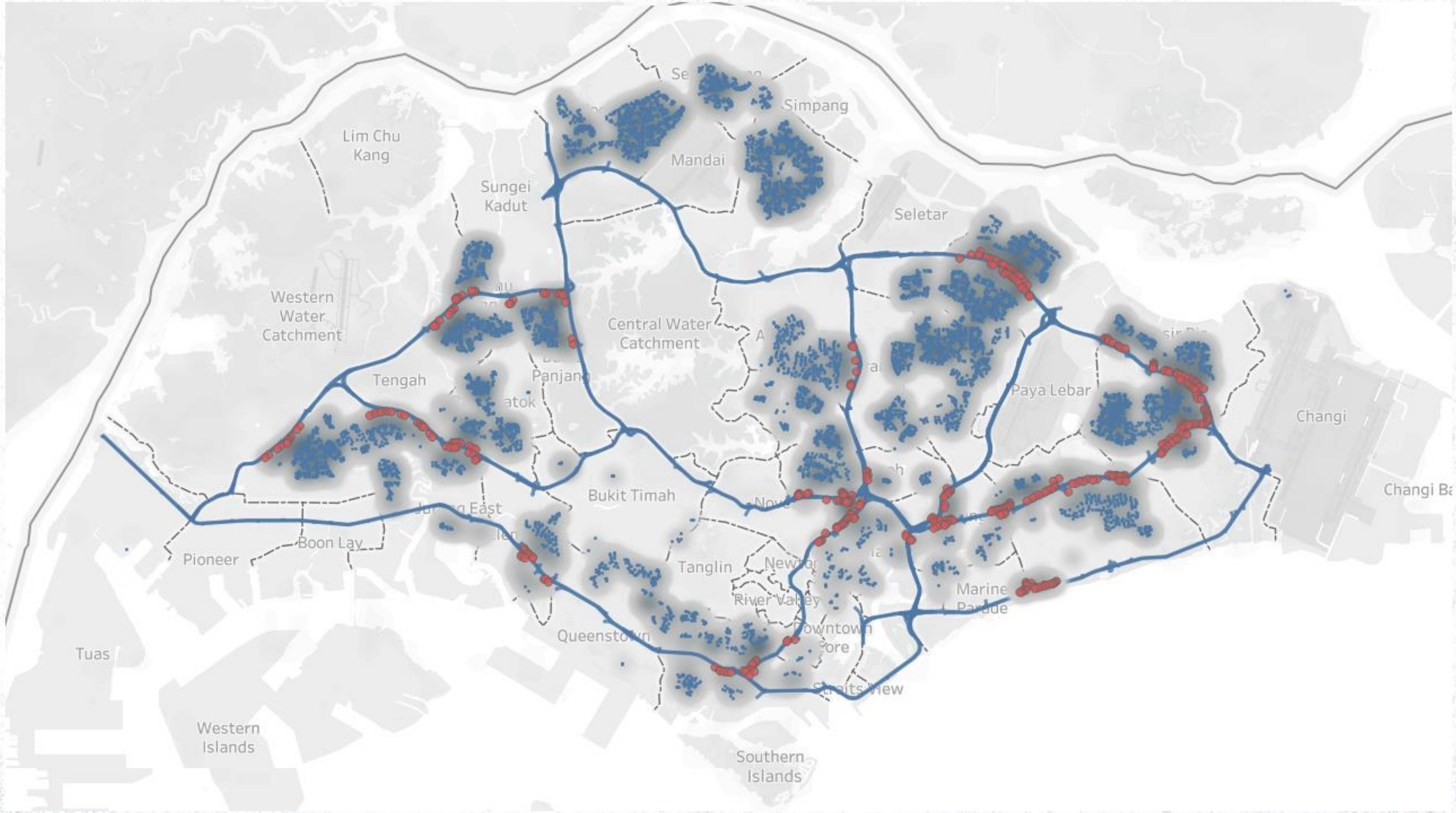
Target Group: Flats that are within **100m** from an expressway

- 5798 number of resales
- 394 unique flats

(i.e. sufficiently large sample size)

NEW SLIDE (updated post-submission)

Visualization of Target flats (close to Expressway)



Grey – All HDB resale flats (2018-2022)

Red – Target Flats within 100m of an Expressway

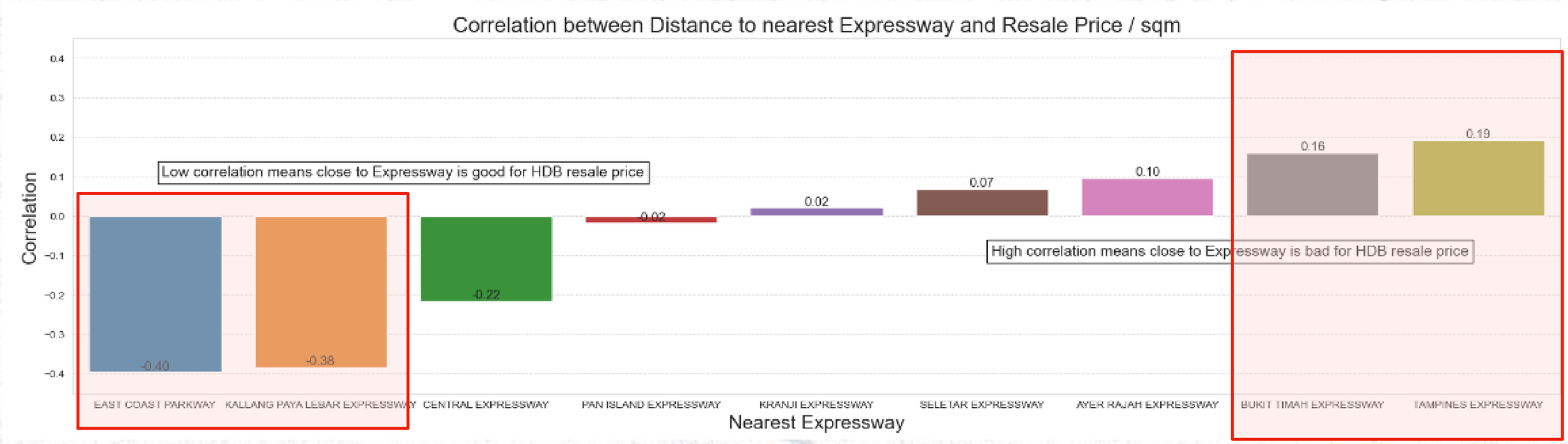
Blue – Control Flats at least 500m away from an Expressway

Observation 2:

Proximity to certain specific Expressways have more significant impact than others

Proximity to ECP and KPE appears to generally improve price, whereas proximity to TPE / AYE worsens price

In spite of having little correlation between Expressway proximity and resale price on aggregate, there are **more pronounced results for specific expressways**, as seen in the graph below:

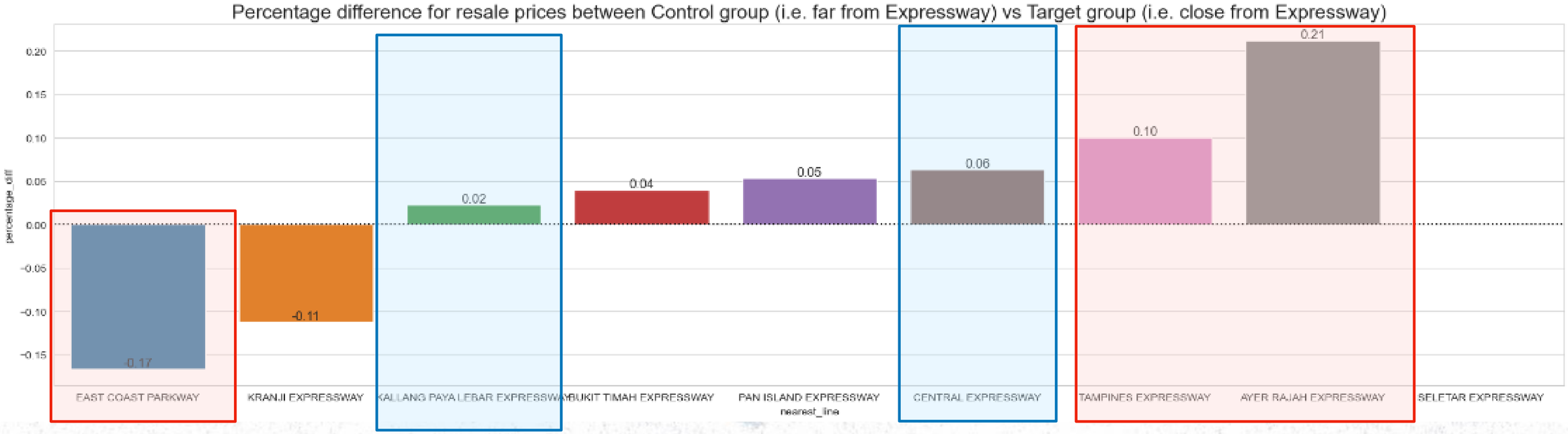


This could be due to perhaps the view being more valuable for some properties (e.g. Marine Parade properties near ECP have a more valuable view), and perhaps the noise from TPE is much worse for residents compared to other Expressways.

Further analysis is required to isolate the precise impact, but this gives us a high-level view that proximity affects different expressways differently

The difference is up to ~20% between properties close and far away from these specific Expressways

Using the **control group** of resale flats far away from Expressways, and comparing their price to the **target group** of resale flats close to Expressways (positive value means far away is more expensive), to get the percentage difference:



(highlighted in blue) Worthwhile noting that KPE, MCE, and CTE **have long stretches of underground tunnels** that were reflected in the National Map Line dataset.

Assuming that underground expressways have minimal impact on people's standard of living in HDBs, then possibly the **impact of Expressway proximity is understated here**; as these unaffected flats dilute the impact of proximity on affected flats.

Caveat: *the small value for PIE does not mean that staying near it is noiseless or not dusty at all!*

Simply means that on aggregate, the good and bad effects displace each other, or the high and low flat prices displace each other

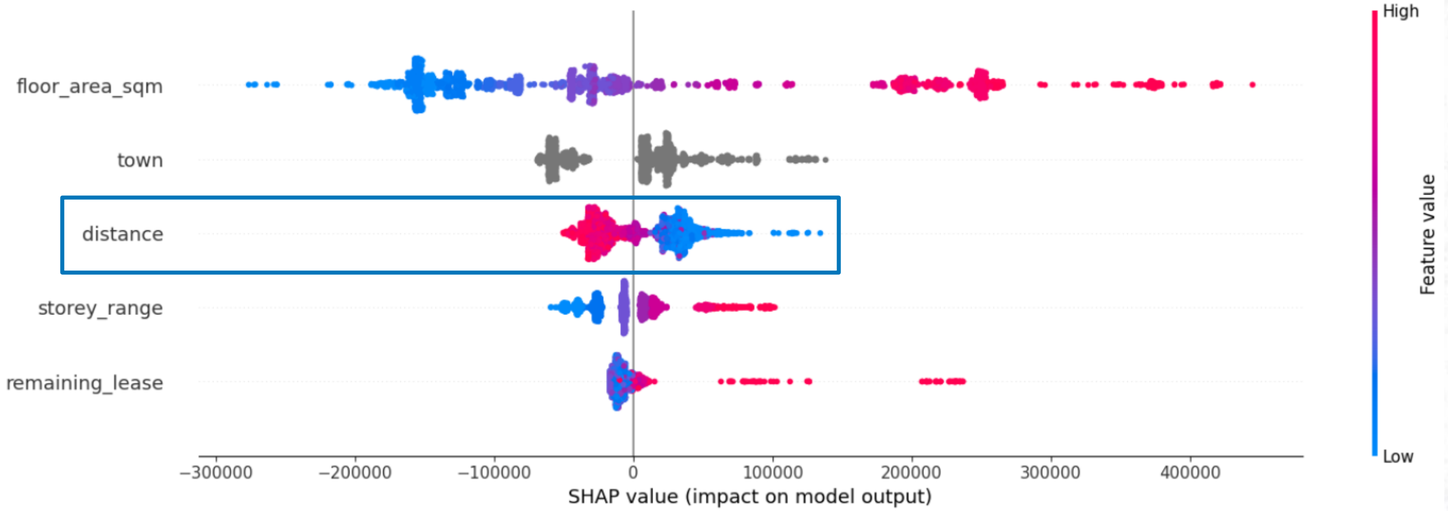
NEW SLIDE (updated post-submission)

Visualization of SHAP values for flats close to 1) ECP, 2) TPE

Trained CatBoost models on datasets of HDB resales where their nearest expressway is either ECP or TPE

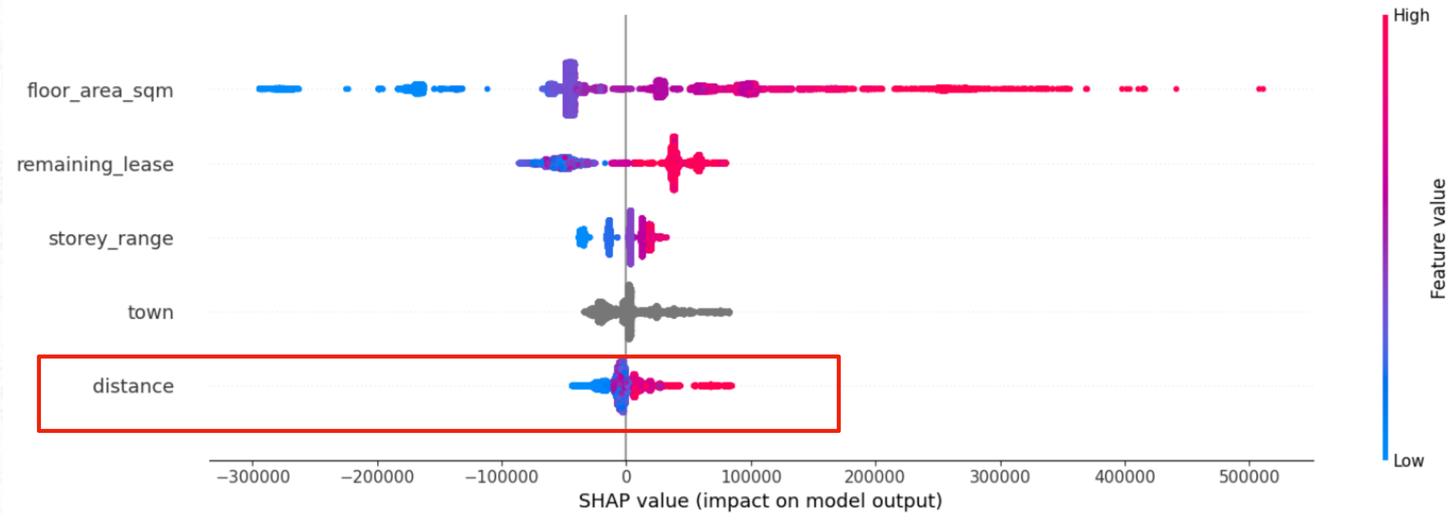
Near ECP (close = better price)

- Model has an R^2 squared of 0.96
- Observe the 'distance' plot (boxed in blue) - a **shorter distance** from the ECP (denoted by the blue dots), has a **positive impact on resale price**, and vice versa
- Also - interesting to note that distance to ECP plays a has a larger impact on flat resale prices than flat height and remaining lease



Near TPE (close = worse price)

- Model has an R^2 squared of 0.83
- Observe the 'distance' plot (boxed in red) - a **shorter distance** from the ECP (denoted by the blue dots), has a **negative impact on resale price**, and vice versa



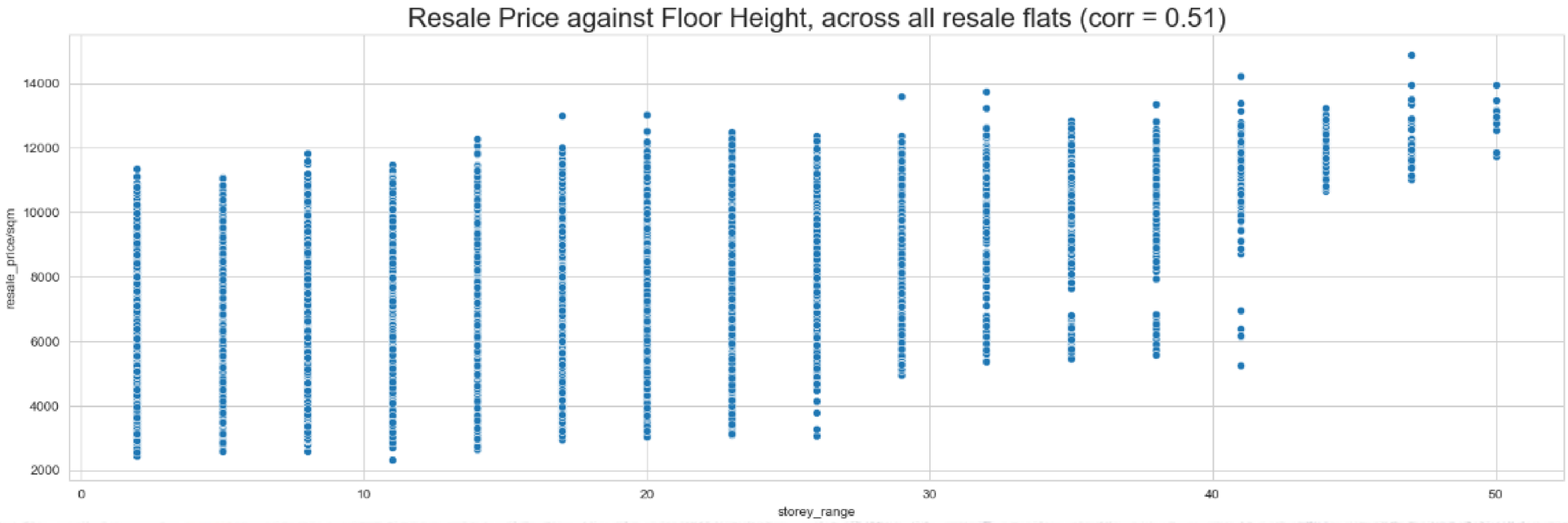
Observation 3:

Higher floors have a larger impact for properties close to Expressways

Units on higher floors are generally more valued, regardless of their proximity to Expressways

Average correlation across all properties:

0.51



There is a **strong positive correlation (0.51)** between resale price and floor height, regardless of their proximity to expressways

Additionally, compared to resale flats on the 5th floor and below, resale flats on the
- 10th floor and above are **21%** more expensive
- 12th floor and above are **31%** more expensive
- 15th floor and above are **45%** more expensive

(specific to this dataset, which across all resale transactions from 2018 to 2022)

However, higher floors are even more valued when units are close to some Expressways

Table of correlation between unit floor and resale price

Target group and sample size

Control group and sample size

Expressway	Correlation	Flat_count	Correlation_close	Flat_count_close	Flat_count_close_%	Correlation_far	Flat_count_far
EAST COAST PARKWAY	0.462341	2009	0.752408	370	18.4	0.395328	1201
CENTRAL EXPRESSWAY	0.574478	14636	0.633848	710	4.9	0.591164	10604
KRANJI EXPRESSWAY	0.410866	8639	0.609003	476	5.5	0.364753	5573
KALLANG PAYA LEBAR EXPRESSWAY	0.434233	6456	0.422628	168	2.6	0.390648	5388
PAN ISLAND EXPRESSWAY	0.457851	28698	0.385412	2161	7.5	0.486735	16269
TAMPINES EXPRESSWAY	0.259997	26716	0.239302	1497	5.6	0.247856	17288
BUKIT TIMAH EXPRESSWAY	0.321453	3777	0.168880	103	2.7	0.313444	2413
AYER RAJAH EXPRESSWAY	0.657257	11664	0.063783	313	2.7	0.690083	7407
SELETAR EXPRESSWAY	0.286981	18965	NaN	0	0.0	0.288127	18405

Average correlation across all properties:

0.51

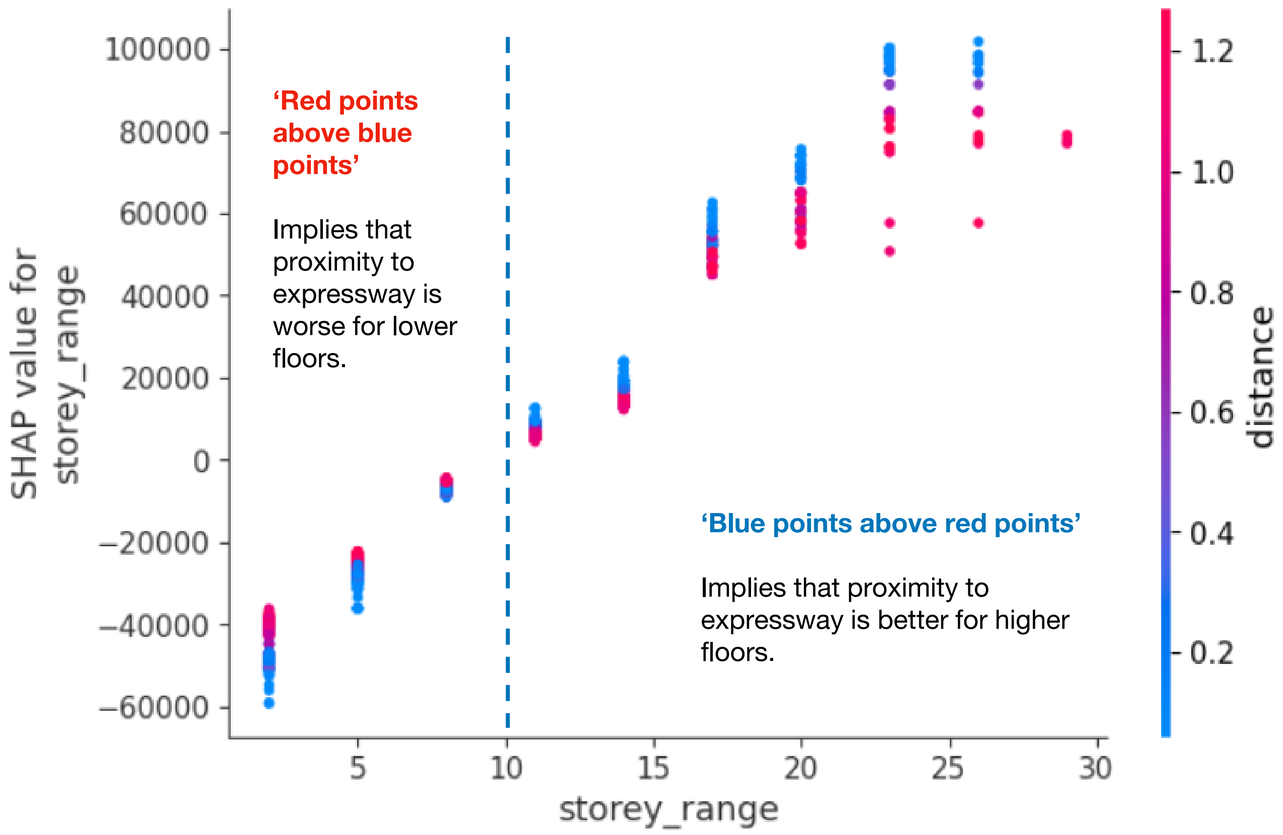
Any correlation higher than 0.51 in the 'Correlation_close' column implies that **close proximity to Expressways increases the relative importance of high floors** compared to on average

- **(highlighted in blue)** This is true for ECP, CTE, and KJE. For ECP, this supports the hypothesis that the higher prices near expressway could be due to the view, as higher floors are generally associated with better views
- **(highlighted in red)** For PIE and TPE, the correlation is lower; though the difference in correlation between the target and control groups are not too significant; hence it might not just be due to the proximity to Expressway
- For KPE, AYE, and BTE, the sample sizes for 'close' flats are quite small (<3% of all flats). Also, the bulk of AYE flats are all within the 11-20 floor range, so there is less variability in prices due to height

NEW SLIDE (updated post-submission)

The outsized value of higher floors or units close to certain expressways can be illustrated via a SHAP dependence plot

SHAP dependence plot



Applied a CatBoost regression model to predict resale prices of flats where the **closest expressway is East Coast Parkway (ECP)**, achieved an R^2 of 0.96.

The SHAP dependence plot shows the relationship between the unit storey level and the distance to expressway, and the impact of the unit storey level on the resale price.

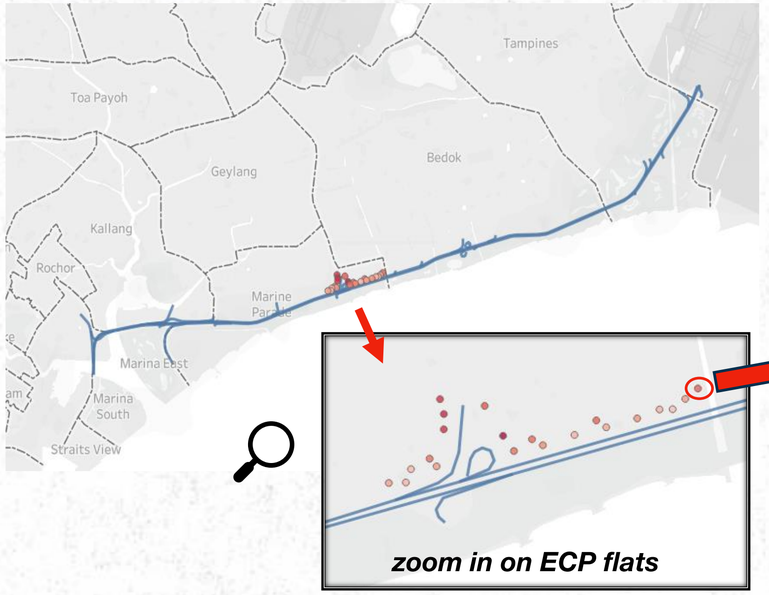
Interpreting the plot:

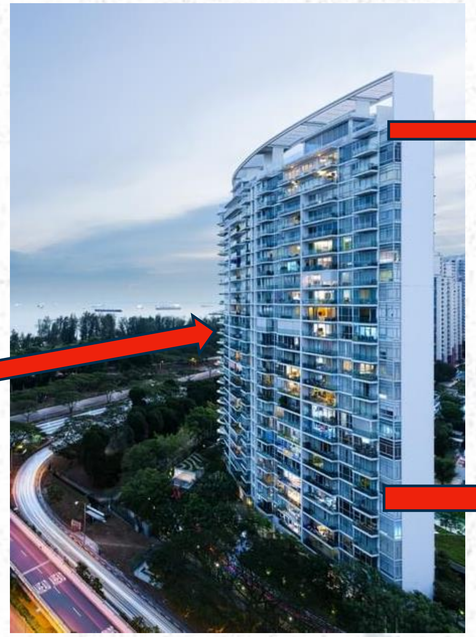
1. Higher storeys result in positive impact on resale price, as expected
2. (*left of dotted line*) For lower stories, the negative impact of storey height on price is worse if a unit is close to the Expressway
3. (*right of dotted line*) **For higher stories, the positive impact of storey height on price is better if a unit is close to the Expressway**
4. The 'swap' where this occurs is approximately at the ~10th floor mark – this could imply that for units close to ECP, the benefits of expressway proximity outweighs the negative impacts from the 10th floor upwards

NEW SLIDE (updated post-submission)

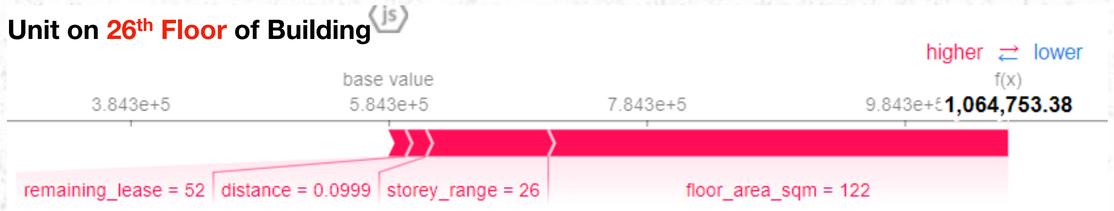
SHAP for a single flat further validates this hypothesis - i.e. high floors = proximity good, low floors = proximity bad

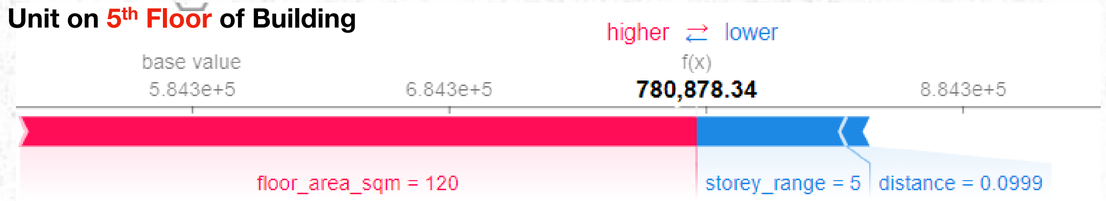
For a specific expressway (ECP), zoom into a particular flat





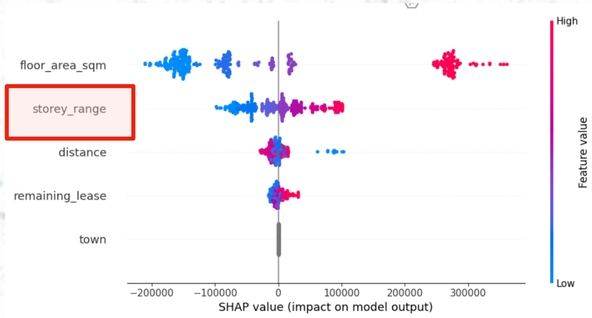
Same building, flat type and floor area, but units on different floors





Interesting to note that all flats that are within 100m from ECP are located in a small cluster in Marine Parade.

Using SHAP on a regression model for these flats, it can be seen that flat height is now the most important factor after floor area in predicting resale price.



- All variables kept constant (same unit, similar resale date, similar floor area, etc.), and only tangible difference is the storey the unit is on (lvl 26 vs lvl 5)
- Note that **distance is a plus factor for the lvl 26 unit, but a minus factor for the lvl 5 unit**

Recommendations and Further Study

Recommendations

This remains a high-level analysis, and significant further study (next slide) will have to be taken before taking any significant steps. Nonetheless, we will propose some immediate follow-ups that HDB management could consider:

Improving conditions for impacted residents:

- HDB could consider working with the relevant Town Councils to glean on-the-ground feedback from residents in units close to the Expressways that are most severely impacted (e.g. AYE/TPE)

- If this justifies the severity that residents are impacted, HDB could consider working with relevant agencies and work out best modalities to improve sound/noise-proofing

- This could involve working with BCA to identify any potential partnerships with built-environment firms specializing in noise-pollution dampening (e.g. see right) that could alleviate the impact of noise/dust on residents

Publicity and Communications

- To combat the negative sentiment of impact of Expressway proximity (from the forums/articles in slide 5), HDB could consider publicizing some data-driven reports that **on aggregate, proximity does not impact prices negatively**

Limitations

Some important features lacking in this study:

Direction the unit is Facing:

- Inward Facing vs Externally Facing the expressway - not captured in the data, but is a big factor on the exposure to noise and dust (especially from the qualitative feedback on forums)

Underground portions of Expressways

- These are quite length segments especially for e.g. KPE/CTE, so we could have diluted our study quite significantly (assuming that underground expressways result in less/no noise and dust)

Additionally, we did not account for time-series – e.g. seasonality, economic downturns

In general, it is tricky to isolate the impact of just Expressway proximity alone, considering the myriad of factors that affect resale flat prices

The intent of HDB's request is also important – if it is for residents' well-being – it is also possible that the resale prices are inefficiently priced, hence on-the-ground surveys could be more useful

Further Study

Some possible avenues of study are:

- **Relative weight of various factors**, as well as their co-dependency, or using some explainable model (e.g. **SHAP/LIME**) to compare the impact of Expressway proximity against other factors

- **Compare against the impact of other contributors of noise:**
  - Proximity to MRT tracks
  - Proximity to main roads in general

- **Impact of new construction of expressways**
  - How are home values affected when new construction is announced, building and launched?

Thank You

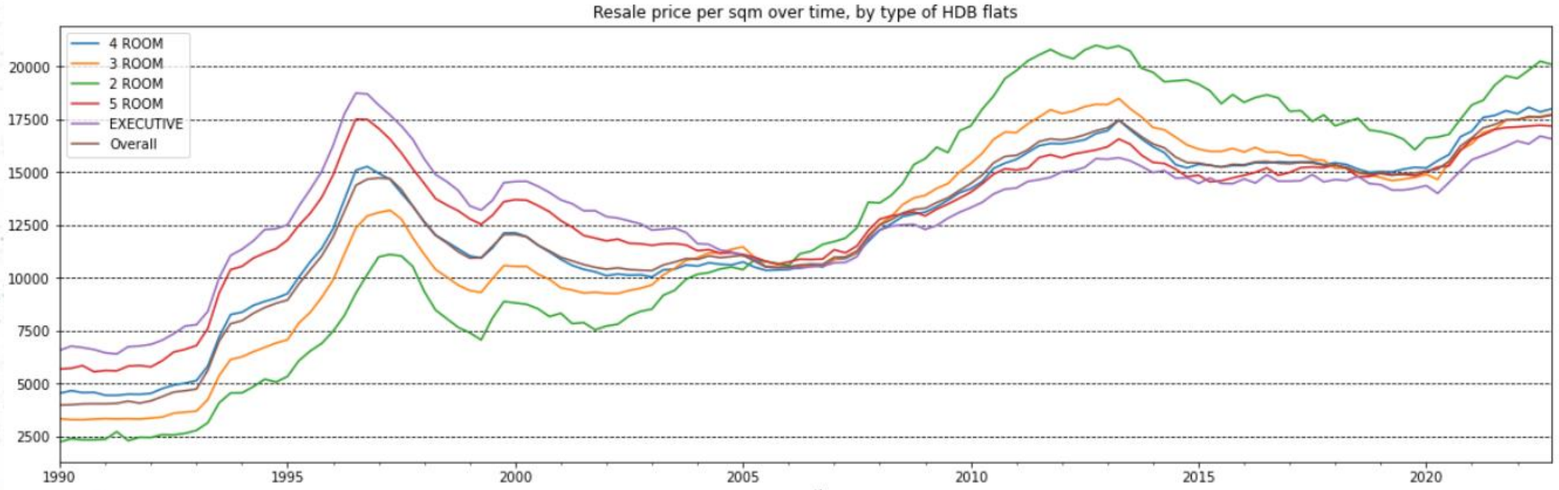
Annex

Annex A: Justifying the use of resale price/sqm as a basis for comparison, instead of just sticking to one flat type

Apart from 2-room flats (a negligible percentage of all resales), the resale price / sqm of other flat types are roughly equal.

```
# real increase in property prices across the board - using price/sqm, by flat types. Data has been corrected for inflati
df_resale_price = pd.DataFrame()
for i in df['flat_type'].unique():
    df_resale_price[i] = df[df['flat_type'] == i].groupby(df["month"]).mean(numeric_only=True)['resale_price/sqm']
df_resale_price['Overall'] = df.groupby(df["month"]).mean(numeric_only=True)['resale_price/sqm']
df_resale_price = df_resale_price.drop(['1 ROOM','MULTI-GENERATION'], axis=1)

# resample by quarter so it's smoother and removes anomalous months
df_resale_price.index = pd.to_datetime(df_resale_price.index)
df_resale_price = df_resale_price.resample('QS').sum()
df_resale_price.plot(title = 'Resale price per sqm over time, by type of HDB flats', figsize=(20,6)).grid(color='black',
```



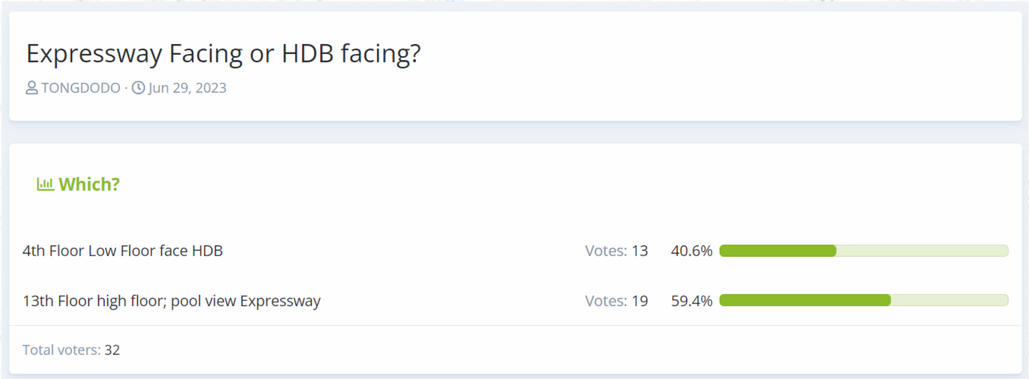
Annex B

Let’s follow the journey of Mr Tan

Avid HDB enthusiast

Wants to own a home

Worried about the implications of a home close to the expressway



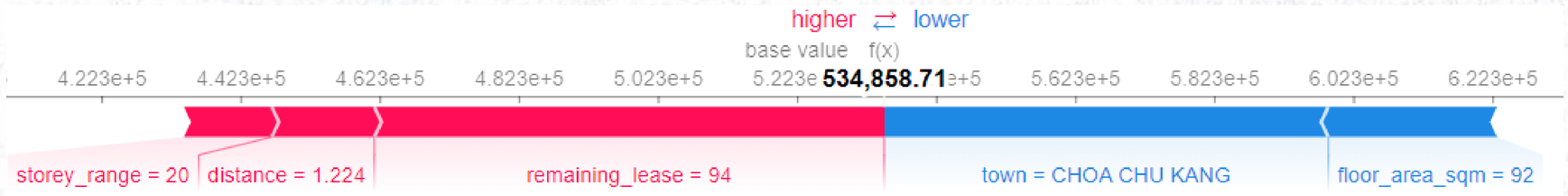
Recommendation – try to make an informed choice. See which expressway the home is close to, and see how much 1) the proximity, and 2) the floor height, impacts the price

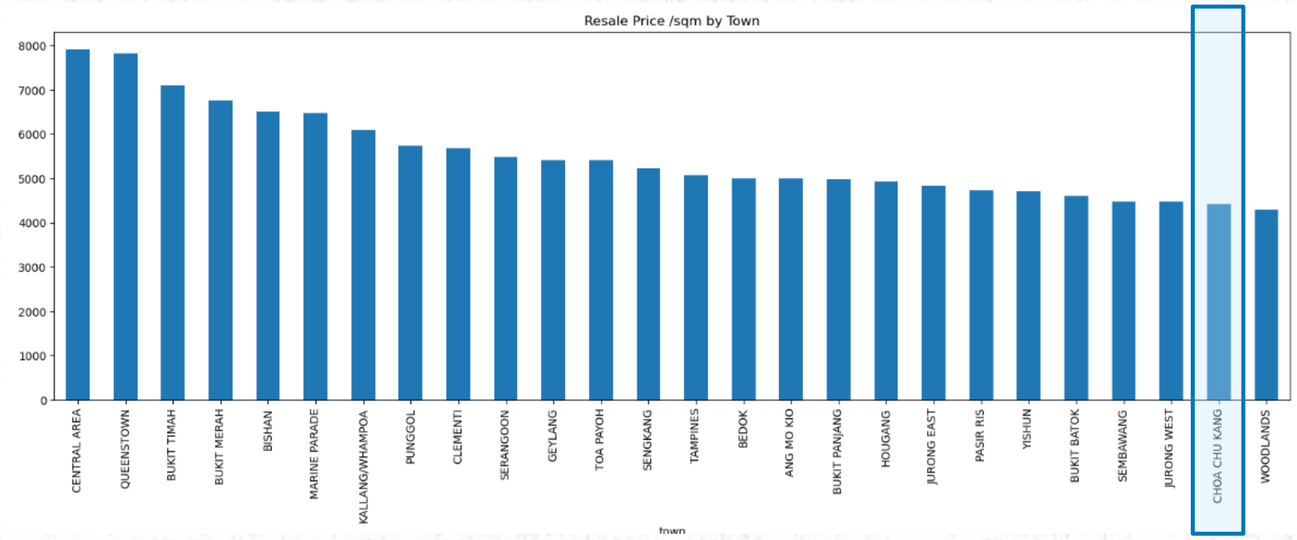
NEW SLIDE (updated post-submission)

Annex C: SHAP for local interpretability – the impact of individual factors can be broken down

Local Interpretability

	0
latitude	1.377259
longitude	103.748292
address	818B CHOA CHU KANG AVE 1
town	CHOA CHU KANG
flat_type	4 ROOM
flat_model	Model A
floor_area_sqm	92.0
remaining_lease	94
lease_commence_date	2017
storey_range	20
resale_price_with_inf	580000.0
resale_price/sqm	6304.347826
nearest_line	KRANJI EXPRESSWAY
resale_date	2022-12-01 00:00:00
distance	1.22356





Summary statistics for whole dataset:

	mean
latitude	1.367827
longitude	103.842464
floor_area_sqm	97.642431
remaining_lease	74.795155
lease_commence_date	1995.950247
storey_range	8.769200
resale_price_with_inf	522265.891716
resale_price/sqm	5389.582787
distance	1.072392