

Finding the Best Home or the Best Investment

Benedict WEE Yu Rui, Russell YAP Song Chen, Lucas LEONG Li Heng

Abstract—As potential home owners, we are very interested in the pricing situation of resale flats as it is the first major investment most Singaporeans will have to make. However, datasets from data.gov.sg are not very useful as it is lacking in terms of geographical information. Although there are good visualizations from OneMap, it does not allow us to compare homes against one another, making it very difficult to make decisions on which house is a better choice at a glance. Our visualization aims to find out if there is a pattern in the proximity of a home to transport amenities such as bus stops and mrt stations and see if it is true that people are willing to pay a premium for a house with the exact same specifications in terms of size with the only difference being its location. It will also allow users to intuitively compare houses in terms of its age, distance to bus stops, Mrt stations and its price per sqm. Finally, our visualization also compares the change in prices of resale flats between the last financial crisis in 2008 to date so that users can tell which type of flats are the most popular and have the ability to appreciate over time. We hope that our data visualization will help anyone have a better view of the resale market and to make quick decisions on where are the locations which have the potential to be a good investment.

Index Terms—Resale flats, housing, HDB, investment, Property market.

1 INTRODUCTION

Over the years, HDB prices have been on the rise ever since the prices of HDB were changed from being pure construction cost to market prices in 1993^[1]. As a result of this change, prices of resale flats and new flats entered in a vicious circle, rising 50% in just 6 months of 1993 and tripled to 1996. This move closed the price gap between small and large flat types and hub pricing have never been the same again. What are the possible factors that contributed to this increase in price of the resale flats? Other than intangible factors such as "feng shui" which is a Chinese philosophical system of harmonizing everyone with the surrounding environment, other factors which are very important in the price of a house was its proximity to MRT stations, bus stops, shopping centres, famous schools and etc...With housing prices being a hot topic that most undergraduates are talking about as we approach that final period called graduation, we wanted to find a way to visualize this information easily. After looking at various implementations, we realized that these visualizations were very superficial and only provided basic information. It was very hard to make comparisons between different houses as you had to extract the information onto excel or paper manually and then compare it. As a result, many people tend to buy houses based on gut feel and skimp on the research portion as it is simply too difficult to do so.

We want to help new potential home-owners have a better insight into the housing market so that their first big investment will be an intelligent one which will not only give them a roof over their heads, but some additional income as well should they decide to sell off their homes to upgrade after the Minimum Occupancy Period (MOP) of 5 years.

2 MOTIVATION AND OBJECTIVES

The main motivation for this project was the fact that current applications which attempt to help users make a better decision about their future home are very superficial. It looks cool and interactive but it does not really help a person make a better choice on which house is a good investment.

The reason why we say this is because we have always heard the old saying, you should buy a house at a good location, but what exactly is a good location? An article taken from HOME&DÉCOR^[2], a website which specialized in information related to homes gave a very detailed list of steps on the information a person should look for before making a purchase of a new home? Everyone talks about it, but no one has actually properly visualized it and attempted to make it easier for people to understand this pricing factor known as its location. Thus, we sought to find a dataset which can help us create visualization which would help us achieve the goal of helping users make a better decision on which house they should get.

Unfortunately, most datasets are not useful in providing us with any insights unless a good amount of data cleaning and wrangling is done. For example, transaction data taken from data.gov.sg only gives you the block number and road name with no postal code. How is that useful? Therefore, one of our primary goals is to come up with a visualization which would provide users with a good visual representation about the resale market in Singapore and give people access to a better data set than the ones currently out there. On top of that, we want to create an easy-to-use visualization tool that can help people analyze the historical flat data so that users can see which flats in Singapore would be the most value for money so that users can actually buy a home which is worth its investment. We also felt that it would be interesting to explore the patterns in the resale flat prices and see what factors really affect the prices of HDBs and find out how much of a premium people attach to amenities such as proximity to public transport, age of the lease, schools and etc... Our tool aims to help people gain the following insights:

- Which areas are under-valued and worth investing in.
- How much have housing prices grown since the last recession in 2008 till now.
- Which flat types are the most active in the resale market and which types are a better for investment purposes.

3 RELATED WORKS

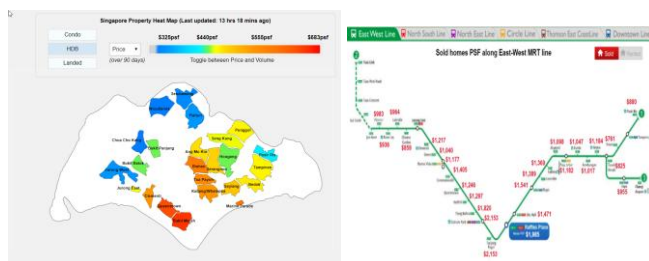
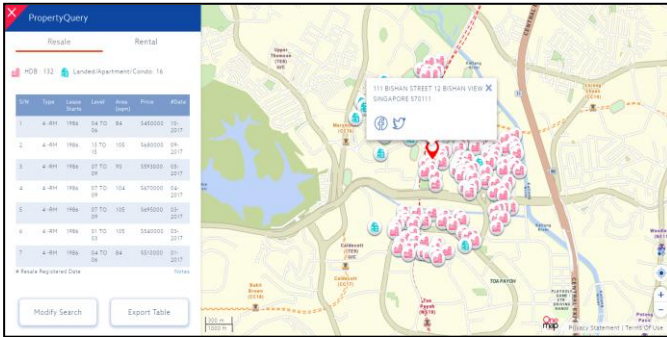


Figure 1: Figure on the left shows a heat map of the properties in Singapore based on its price per square feet (psf) while the figure on the right shows the average psf of houses located along MRT stations.

Figure.1 above shows 2 visualizations were implementations that try to give people a better visualization of the housing market situation in Singapore. However, despite the fact that these two visualizations show a good aggregate view of the Singapore market, there are some issues. For the image on the left which is done by the Singapore Real Estate Exchange (SREX)^[3], I am unable to see a map view of the

various houses located in the area. Thus, it makes it difficult to visualize where the houses are located and I would have to use another tool like Google Maps to actually see where the house is located on the map which is an extra step. For the image on the right which is also done by SREX [4], it gives users a good visualization of which MRT stations are more valuable at a glance. However, in order to zoom in to houses in the area, a user will have to go to the table located at the bottom (not shown in screenshot) and click on it. Then the user will be redirected onto another page which displays the various houses in a table form. Once again, there is no map view, and the user would then have to use a map tool in order to see where



exactly the house is located at which is an additional step.

Figure2: Screenshot of a query for a resale flat in OneMap property query.

Figure.2 above shows an implementation done by OneMap[5]. It has a good search function which allows you to search for any house by its name, address or postal code. Once you select a house, all its past transactions and current pricing details are shown, allowing you to know what is the price activity for the selected house. Although this is good, there is no way for a person to do comparisons between houses. Also, although you can see the various amenities around the area, the distance to those amenities is not highlighted and the distance is not shown, so it makes it difficult to easily see how close the chosen house is to these amenities.

Thus, we set out to combine the benefits of each visualization into one consolidated one so that it is easy for a user to get all the information they need to make comparisons and evaluate the value of a resale flat.

4 VISUALIZATION APPROACH

To plan out how we were going to carry out this visualization, we decided to do a 4-step agile process. (1) Data Collation, Exploration and Preparation; (2) Sketch a rough model of the visualization tool; (3) Create Visualization tool; (4) Evaluate usefulness of the tool

If the tool fails the usefulness test in step 4, we would go back to step 1 and repeat the process so that we can continuously improve our visualization.

4.1 Data Collation, exploration and preparation

4.1.1 Resale price data preparation

As mentioned earlier, the data came in the following format, 1) month, 2) town, 3) flat_type, 4) block, 5) street_name, 6) storey_range, 7) floor_area, 8) flat_model, 9) lease_commence_date, 10) resale_price. Since the dataset was lacking postal code information, we had to concatenate the block and street name that was provided in the dataset and use it as a parameter to query the OneMap APIs so that we could get the longitude, latitude and postal code.

The first issue we had was the 250 calls per min limit that was imposed by OneMap. As such, we had to shard our data into sizes of 250 lines and call it every 1 minute.

The second issue was that our dataset gave addresses in short-hand. One example is, Serangoon Gdns instead of Serangoon Gardens. Thus, we had to find all the various short forms and use a script to

change to words accordingly so that the OneMap API would not give us an error. The third issue, after retrieving the data, we realized that there were many redundant entries which were not what we wanted, such as schools, carparks, elderly homes and etc... Thus, we had to use filter code to filter off the entries which did not match. There were 60 different key words (shown in Table 1 on the right) that we had to use in order to come up with the final cleaned dataset. To check the validity of the the cleaned dataset, we compared the results. After which, we had to convert our csv file into a GeoJSON format so that our data could be mapped onto our visualization.

4.1.2 MRT and bus stop data preparation

For bus stop and MRT, the dataset was taken from the LTA website. This dataset was given in a SPH format, after which, an online converter was used to convert it into a csv file. Proj4J was then used to convert the x and y coordinates provided into EPSG:4326 which is the GeoLat format used by Singapore.

4.1.3 Subzone data preparation

For subzones, we used togeojson.js to convert the KML file was retrieved from data.gov.sg into a GeoJSON file

4.2 Visualization Development

After completing the first stage, we started to plan rough visualizations on paper and on tableau just to see how it would look and we used these sketches to decide on whether we would choose to further develop the idea into the actual visualization.

4.2.1 Resale price over time visualization

Initially, we wanted to come up with a visualization that shows the how the average resale price has changed over the years in accordance with the Resale Price Index (RPI). The main idea was to help people understand the events that have an effect on the prices of resale flats. This visualization would be done on a dual-axis bar chart using D3 with annotations on the chart to let users know the major events that happened during the periods of time when the resale prices fluctuated. However, after completing a mock visualization on tableau, we realized that it was not very useful and insightful. Thus, we felt that this visualization should not be included and this was removed.

4.2.2 Interactive Heat map

The next visualization that we wanted to do was an interactive heat map which shows the various bus stops, MRT stations and subzones layered with the resale flats. The idea that we were working on was: what are the location factors of a HDB which would directly affect its valuation. We also wanted to do our visualization by including transportation amenities, shopping centres and schools. However, due to a lack of time, we had to prioritise which ones to focus on for our implementation. Thus we decided to use the distance to transportation amenities, age of the resale flat and resale price as our basis of comparison. The initial hypothesis that we had was: The closer the house is to a transport amenity be it bus or an MRT station and the younger the age of the house, the more valuable the house is going to be.

Table 1. List of Filter Words

SCHOOL	CAR PARK	DBS	CENTRE
OCBC	CHILD	LTD	POST
LEARN	CLINIC	KINDERGARTEN	TOWN
KID	STUDENT	HOME	MDIS
UOB	PCF	MONTESSORI	PUBLIC
OFFICE	SERVICE	INSTITUTION	PRO-TEACH
PROFESSOR	LITTLE	ANIMAL	VET
FAMILYLAND	FRIENDS	FOOD	MY FIRST
MKT	YO:HA	COMMUNITY	CARE
HUB	FSC	HSBC	INDUSTRIAL
ACADEMY	MALL	SQUARE	CHURCH
EDUCATION	BUS	MARKET	NASRY
GROW	TEMPLE	STAMFORD	CLUB
BASC	STANDARD	ARMY	MRT
TECH	HUT	CHRISTIAN	PLAYLAND

To help visualize this, we decided to use a radar chart to visualize the 4 variables, Distance to MRT Station, Distance to bus stop, Price per sq metre (psm) and Number of years left on the lease. For both psm and number of years left on the lease, the size of the radar would be directly related. So the greater the psm or the number of years left, the greater the value will be on the radar chart. However, for the two distance values, it will be done in reverse. The shorter the distance of the house to the MRT station or bus stop, the greater the values on the radar chart.



Figure3: Screenshot of radar chart showing the 4 variables we want to compare with 2 four-room flats being shown for comparison

Looking at Figure.4 above, we can clearly see that the prices of both houses being compared is around the same. Thus to make a decision between these 2 houses, one should look at the other 3 factors. Since the house in red is much closer to the MRT station and only slightly further from a bus stop as compared to the house in yellow. The obvious choice would be to go for the house in red. We also had to take note that there could be multiple entries for each house because for the same postal code, there could be multiple types such as 2-room, 3-room, different storey ranges... On top of that, we should also display some of the past transactions so that users can see an overview of all the transactions for that address. Thus, we decided to display all of the entries in a sorted list grouped by flat-type, followed by date of the sale as shown below in Figure.4

Blk 395 BUKIT BATOK WEST AVENUE 5, Singapore 650395

Nearest MRT : 841.23 metres (BUKIT GOMBAK MRT)

Nearest BUS : 207.31 metres (Stop 43649)

Lease Commence Date : 2004 (86 years left)

Flat Model : Model A

Resale Price :

Date	Type	Level	Area(sq.m)	Price
01-2017	4 ROOM	07 TO 09	92	\$408,000
05-2017	4 ROOM	13 TO 15	90	\$430,000
06-2017	4 ROOM	07 TO 09	92	\$415,888
07-2017	4 ROOM	19 TO 21	92	\$460,000

Figure.4: Screenshot of a table showing the list of prices for a resale flat at postal code 65039

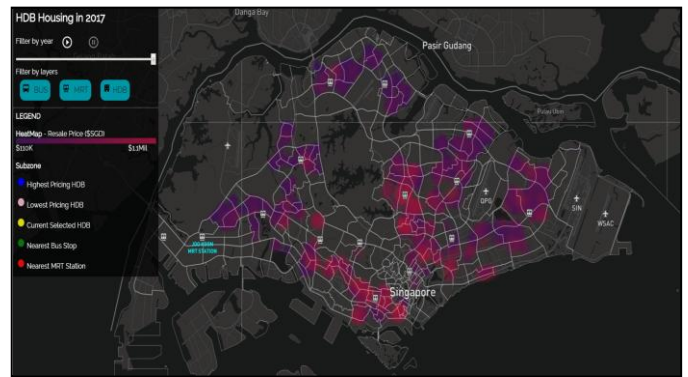


Figure5: Screenshot of our interactive map at the heat map view

When a user first enters our heat map visualization, he will see an overview of the map of Singapore. The subzones are clearly shown and a heat map of the resale prices is shown. The higher the prices at the subzone, the hotter the zone will be.

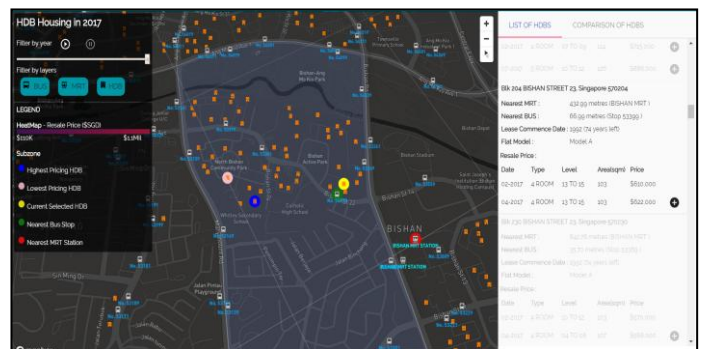


Figure.6: Screenshot of our interactive map at the zoomed-in view

Putting it altogether, the way a user would use our map is as such: First, they would look at the overall view as shown in Figure.5 to see where are the areas which are not too expensive based on the colour. Next, a user would select a subzone by clicking on it and the user would be shown a zoomed in view as shown in Figure.6 below of the chosen subzone with the list of resale flats on the right side of the visualization. A person can scroll between the results to see the different prices and attributes of resale flats within the subzone. After which, if a user wanted to make comparisons, the user can click the + button in the table displayed to add the selected resale flat to the radar chart in the Comparison of HDBs tab as shown below in Figure.7. A user can freely add and remove HDB flats and visualize them on the radar chart for quick and easy comparisons. When a house is selected on the map, its own position will be highlighted in yellow. The most expensive house in the area would be highlighted in blue while the cheapest house in the area will be highlighted dark pink. The closest bus stop is highlighted in green while the nearest MRT is highlighted in red.

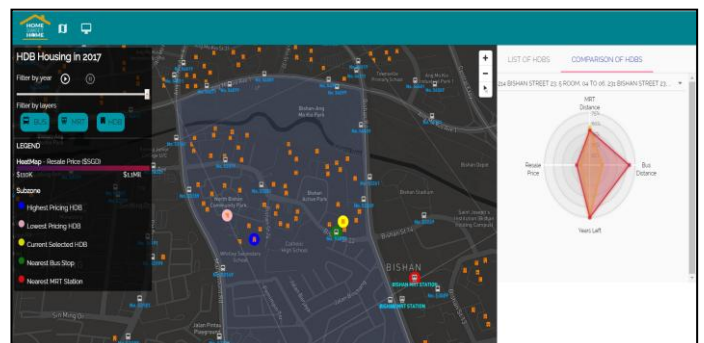


Figure7: Screenshot of our interactive map with radar chart comparison

4.2.3 Bubble plot of change in resale prices

This visualization aims to give users an idea of how much have the prices of resale flats grown since the last economic downturn which was in 2008 due to Global Financial Crisis^[8] to July 2017. The visualization shown in Figure.8 aims to show users the overall performance of the resale flats market. A user is also able to change the grouping from an overall view into a detailed view.

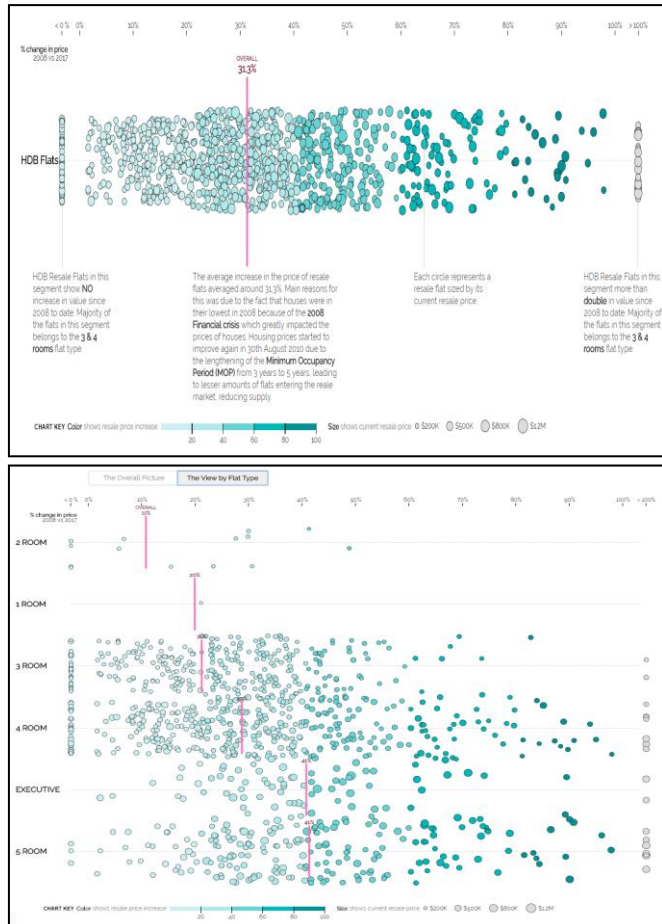


Figure.8: Screenshot of our bubble plot of the change in resale prices. The top-most one is the overall view while the bottom one is the view by flat-type

By doing so, a user will be able to see the outliers at a glance and also where is the main resale activity happening. The circles represent a resale flat sized by its current resale price and if a user hovers over a circle on the bubble plot, the user will be able to see the details of the house, such as the address, % increase in price, resale price in (2017) and the resale price in (2008). The line in pink shows the average % increase in resale price for each chart. Values which are on the extreme right end are houses which have more than doubled in value, while those on the extreme left are those that have dropped in value. Eye balling data, we did not notice any visible pattern in where these houses were located so we have no clear reason at this point of time of the reason for these anomalies. One possible reason is that for the houses that dropped in value, their values before the 2008 financial crisis could have been extremely overpriced at that point of time, thus resulting in the prices of the resale flats not being able to catch up to its value after the drop in price.

5 KEY FINDINGS & INSIGHTS

This section will discuss the following key findings and insights that our group was able to obtain from our visualization tool:

5.1 Heat map insight - West is Best

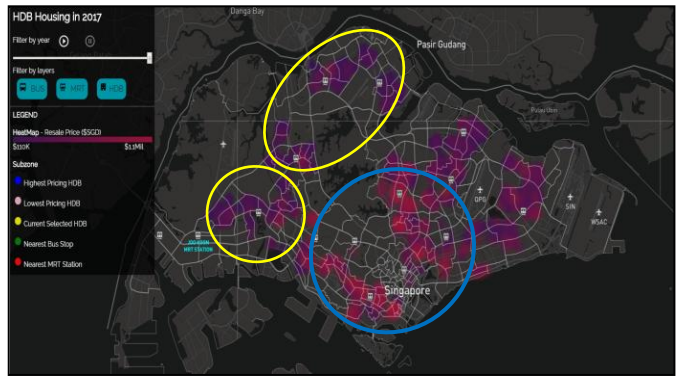


Figure.9: Screenshot of heat map with annotations

The North and Western Areas of Singapore seem like good investments based on the heat map on Figure.9 above.

Notice that despite some zones being in close proximity to an MRT station, the zones are not hot spots. These zones in particular are those located in the northern and western sides of Singapore and these 2 zones are circled in yellow in Figure.5 above. These zones fall in line with the reported top 10 most affordable areas for HDB resale flats in the written by Rajagopal.^[6] These areas were: Sengkang, Yishun, Tai Seng, Khatib, Woodlands, Bukit Gombak, Sembawang, Yew Tee, Admiralty and Chua Chu Kang.

For the hotspots, as shown in Figure.9 above highlighted in blue, notice that these areas are in close proximity to the central area of Singapore which is where our Central Business District(CBD) is. Based on the article written by Teng^[7], this pattern is in line with what was reported in the article. This areas were: Redhill, Bras Basah, Tiong Bahru, Bugis, Telok Ayer, City Hall, Common Wealth, Labrador Park, Outram Park and Tanjong Pagar.

The reason the resale flats in these area are so expensive is precisely because of their close proximity to the CBD. As the CBD is viewed as the place with the most activity in terms of business and recreation, people attach a premium to being close to it and having easy access to public transport is an added premium that people are willing to pay as well. What does this mean for the areas which are located far away from the CBD? This means that the premium attached to being near an MRT has not been realized yet.

With the recent announcement by the URA of the Jurong Lake District Project, it is signal that the Western side of Singapore is going to get a big upgrade very soon with the Western Corridor of Singapore being postured to become the second CBD of Singapore as stated by the URA^[9]. The URA also stated that this project will focus on making public transport the key mode of transportation within the area with the expansion of the bus network service within the area and roads built for the sole purpose of public transportation^[10]. Thus, the Western side of Singapore has the highest chance of appreciating in value in the nearer future as compared to the Northern side of Singapore. However, this will take time to develop, so older resale-flats will not be the best candidates as an investment due to the shorter life-span it will have by the time the developments are complete.

Therefore, resale flats in the Western region of Singapore, particularly the Jurong Lake areas which are also located close to MRT Stations and Bus Stops and with a young age make very good investments for people who are thinking of getting a resale flat.

5.2 Bubble plot insight - Optimal or Safe?

If a user is looking for a resale flat which would give him the greatest yield, they should look for either executive flats or 5-room flats as these flat types have the highest average increase in price of around 41% and 43% respectively. However, a safer investment would be 3-room or 4-room resale-flats as these flats have the

highest volume as shown in Figure.10 below. Thus this means that resale flats of these types are more liquid as compared to the other

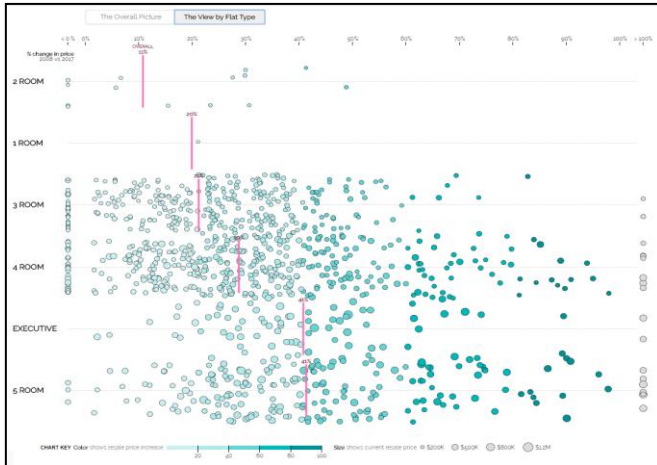


Figure.10 Screenshot of heat map of % change in resale prices

types. If a user were to buy a 3 or 4-room flat, they will have a higher chance of finding a willing buyer in the future as compared to the other flat types. So it all depends on the risk appetite of a user, the greater the risk appetite, the higher up the value-chain they should go. The 1 and 2-room flats are not great investments due to their nature of being for elderly folk or singles, thus it is not in high demand.

6 LIMITATIONS

I do not dare say that we have found all the factors that make up the intangible factor called location. However, I believe that our visualization is a step in the right direction towards the creation of a visualization that will help people get a better overview of the resale market. There are many more variables that should be taken into account, such as proximity to schools, proximity to shopping centres, people's perceived value of the location, the racial distribution in the neighbourhood and many more. Thus, all I can say is that this visualization is still a work in progress but it does help to show the tangible effects of location factors on the value of a resale flat.

An analytical model could also be explored to find the exact relationship between the various variables that we have discussed and come up with a more accurate radar chart other than just the pure values in comparison to one another.

7 CONCLUSION

Housing prices is going to be one of the most important topics that any person will be talking about as the years go by. And having a good visualization tool would greatly help the average folk have a better understanding of the housing market so that they can make well-informed decisions about what resale flat they should go for. A user could always go for a home which is very accessible and connected at this point of time and pay a high premium. Alternatively, a user could go for a more long-term approach and buy a resale flat in a location that is undervalued now but possesses a huge potential to grow in value.

Our recommendation to such users is as follows: Choose a house in the western side of Singapore, particularly the areas close to the Jurong Lake District as this area is primed to become the next CBD. This area will become much more developed and the future benefits would be very worthwhile if you have the time-frame to wait for it. For the type of flat, you can either choose something an executive type or 5-room flat if you are riskier or a 3 or 4-room flat if you are safer. However, we recommend that you choose a size that fits your future needs. For example, if you require a 5-room flat in the future, you should choose to get one now if you can as it will be relatively under-valued as this point of time and when the time comes to

actually sell it if you want to upgrade, the gain in value that you may have gotten should be able to cover the overall gain the market and help you upgrade with lesser worry about the amount of money required.

8 ACKNOWLEDGEMENT

We would like to thank Professor Kam Tin Seong for regularly giving us feedback about our visualization and helping us to come up with ways to counter the performance issues when dealing with a large visualization such as ours.

REFERENCES

- [1] T. HDB history and floor plan evolution 1930s – 2010s. Retrieved November 24, 2017, from <http://www.teoalida.com/singapore/hdbfloorplans/>
- [2] PROPERTY: Before you buy your dream resale HDB flat. Retrieved November 23, 2017, from <http://www.homeanddecor.com.sg/blogs/property-you-buy-your-dream-resale-hdb-flat>
- [3] Singapore Property Heat Map. Retrieved from <https://www.srx.com.sg/heat-map>
- [4] Latest Home Prices Near MRT Stations. Retrieved from <https://www.srx.com.sg/mrt-home-prices>
- [5] OneMap PropertyQuery. Retrieved from <https://www.onemap.sg/main/v2/propertyquery>
- [6] Rajagopal, S. B. (2017, Jul 13). 10 Most Affordable Areas For HDB Resale Flats. Retrieved November 23, 2017, from <https://mustsharenews.com/most-affordable-hdb/>
- [7] Teng, K. Y. (2017, Nov 15). 10 Most Expensive Areas To Buy An HDB Resale Flat. Retrieved November 23, 2017, from <https://mustsharenews.com/most-expensive-hdb/>
- [8] Global Property Guide (2009, Nov 24). Massive recovery in Singapore prompts housing bubble fears. Retrieved November 23, 2017, from <https://www.globalpropertyguide.com/Asia/singapore/Price-History-Archive/Massive-recovery-in-Singapore-prompts-housing-bubble-fears-326>
- [9] Au-Yong, R. (2017, August 25). More jobs, new homes and car-lite push unveiled in Jurong Lake District proposals. Retrieved November 23, 2017, from <http://www.straitstimes.com/singapore/more-jobs-new-homes-and-car-lite-push-unveiled-in-jurong-lake-district-proposals>
- [10] Phua, R. (2017, August 25). Jurong Lake District to create more than 100,000 jobs, 20,000 homes. Retrieved November 23, 2017, from <http://www.channelnewsasia.com/news/singapore/jurong-lake-district-to-create-more-than-100-000-jobs-20-000-9155708>