Exploring relations between household income and purchasing properties in Melbourne

Introduction

The research question put forth for investigation is: "**Is there a relation between household income and purchasing properties across Melbourne?".** This question aims to examine Melbourne's inclusiveness and liability since housing is not only a necessity but also can reflect the level of liability and quality of life. Within any society, the way people live in communities, housing accessibilities, and the affordability of houses are critical. Therefore, it is worthy to investigate the potential relationship between household income and consumer's purchasing behaviours. In this report, [1]Melbourne City is chosen as our sample set, since it is the city with the second largest population within Australia, which has seen great influxes of demand for housing.

Data pre-processing

The first dataset **"Personal income in Australia"** published by "Australian Bureau of Statistics"[2] is used to explore the personal household. It includes the total numbers of earners and the median/sum/mean amount of money they earn, as well as the median ages of earners across all different areas in Australia from year 2011 to 2018 and divided as different area levels. The second dataset **"MELBOURNE_HOUSE_PRICES"** downloaded from Kaggle[3], it records the housing transaction with property location, date, housing characteristics, price, etc from 2016-18. In order to compare the housing cost and household income for each area, several data pre-processing steps have taken place to align the data dimension (i.e. year and suburb) for both datasets.

Firstly, since the income dataset only covers from the year 2016 to 2018, the yr 2016-2018 VIC income data in 'table1.4' is extracted. Furthermore, the **'groupby'** function is used to combine the suburbs with more detailed locations (like Ballarat – North and – South) into its main suburb. Clustering is used to gain basic understanding about income level across the suburbs, the average of median and mean income data from

---

1 Committee for Economic Development of Australia (CEDA) (2017) Housing Australia, Melbourne, CEDA., p. 16.

2 https://www.kaggle.com/anthonypino/melbourne-housing-market

3 https://www.abs.gov.au/statistics/labour/earnings-and-work-hours/personal-income-australia/latest-release

2016-18 is calculated using the '**mean'** function. The clustering are visualised in following figure 3 and 4. It is found that the mean income for most suburbs is concentrated from 40k to 60k, several suburbs have higher than 70k income levels. However, almost all suburbs have 35k and 60k median income levels. The difference might be caused by the relatively large income gap existing in some suburbs. Therefore, the mean income is chosen as our main assessment metrics to investigate how income levels affect purchasing behaviour, considering the group of households with high income.

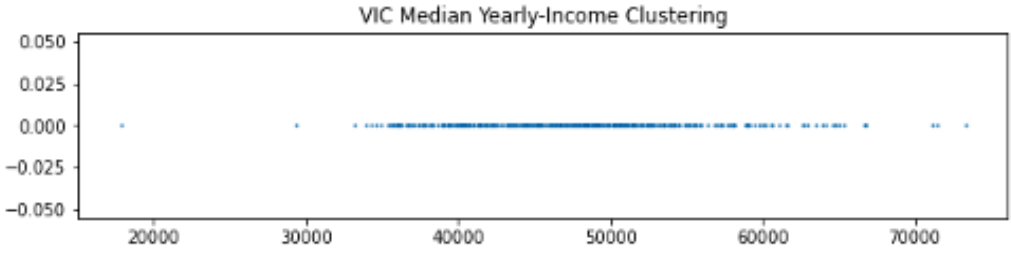

Figure 1: VIC Mean Yearly-Income Clustering

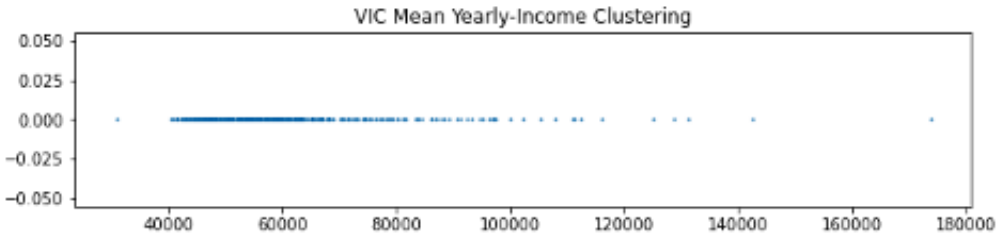

Figure 2: VIC Median Yearly-Income Clustering

Secondly, as for the housing dataset, two sub-datasets are merged to get a more complete picture. The un-relevant columns are dropped and the data is sorted by year-month order for readability. Next, the data without prices cannot support our analysis sufficiently and therefore are deleted, as 'Prices' is the main metric. Moreover, a similar **'Groupby'** suburb operation is taken for housing data. More importantly, the histogram is used to learn the housing price distribution, which is shown as figure 1. From this basic exploration, we found that there are several houses with extremely high prices, which is far from the mean and the histogram shows skewed right. To avoid the further exploration affected by those values, we deleted the data with the z-score of price is higher than 3 as following:

```
: removed_house_full = house_full[(np.abs(stats.zscore(house_full["Price"])) < 3)]
  removed_house_full.head()
```

After removing the outliers, the price of housing data follows the distribution in Figure 2.

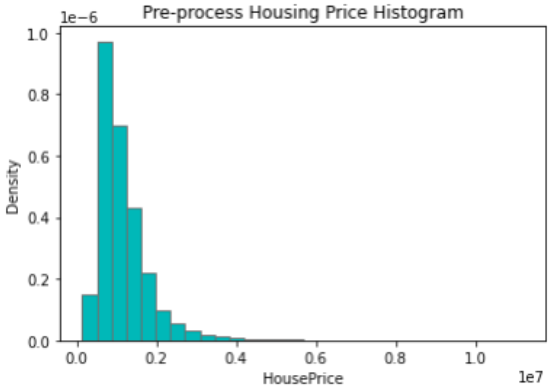

Figure 3: pre-process housing price

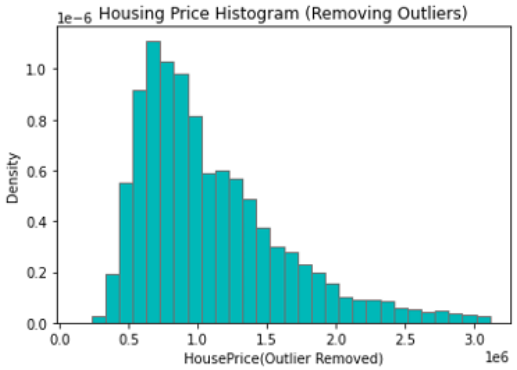

Figure 4: housing price (removing outliers)

Data Exploration

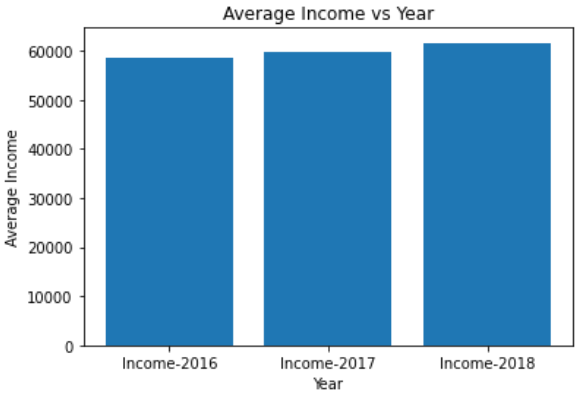

Figure 5: Average income vs Year

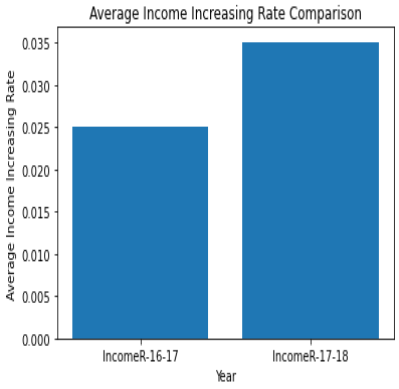

Figure 6: Average income IR Comparison

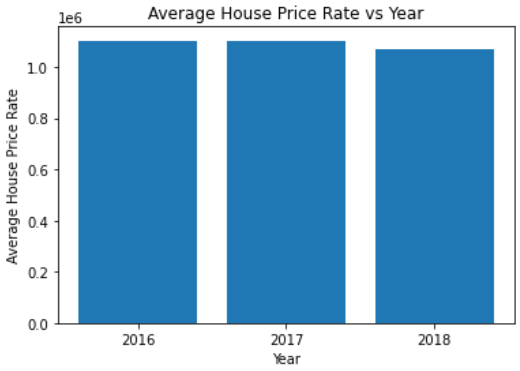

Figure 7: Average Housing Price Increasing Rate vs. Year

From *figure 5* and *figure 6* we can discover that despite the yearly average income increase not significantly high from 2016 to 2018, the average income rate increase is high. From figure 7 we can observe the house price is decreasing, and the average

house price rate is decreasing too. This is a good indication, due to the income rate increasing and house price rate decreasing, households are more affordable for a house in 2018 than 2016. This means people who are not able to purchase a house in 2016, they might be able to consume a house in 2018, and this could potentially increase their liveability.

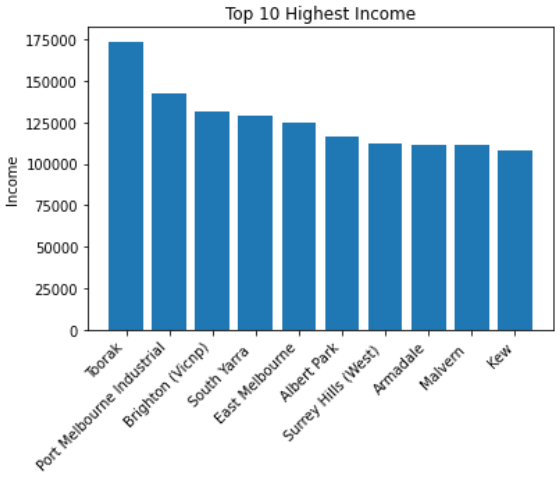

Figure 8: Top 10 highest income suburb

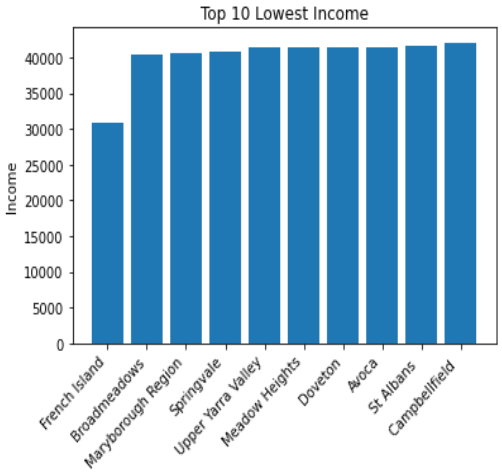

Figure 9: Top 10 lowest income

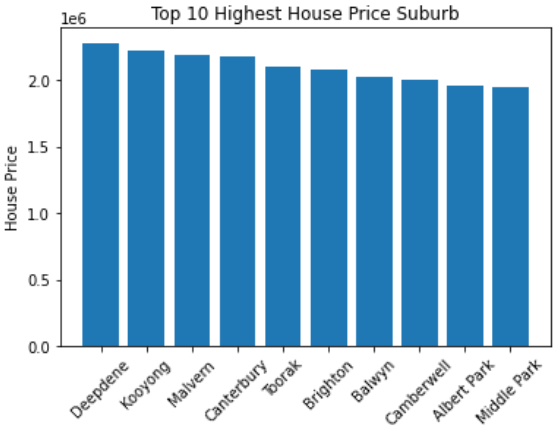

Figure 10: Top 10 highest house price suburb

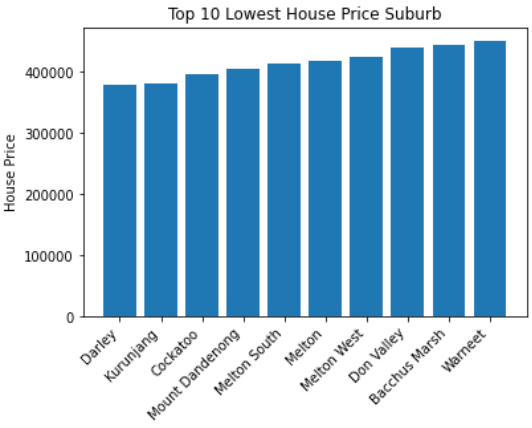

Figure 11: Top 10 lowest housing price suburb

When comparing *figure 9* and *figure 11,* we can observe that none of the top 10 lowest income suburbs in the top 10 lowest house price Suburb. This shows people who live-in low-income suburbs do not mean they just only can purchase a very cheap house. on the other hand, *figure 8* and *figure 10* shows that people who live-in high-income suburbs do not mean they have to spend more on the house. This can potentially show that Melbourne has great inclusiveness. Even if households earn less, they are still able to purchase a relatively expensive house.

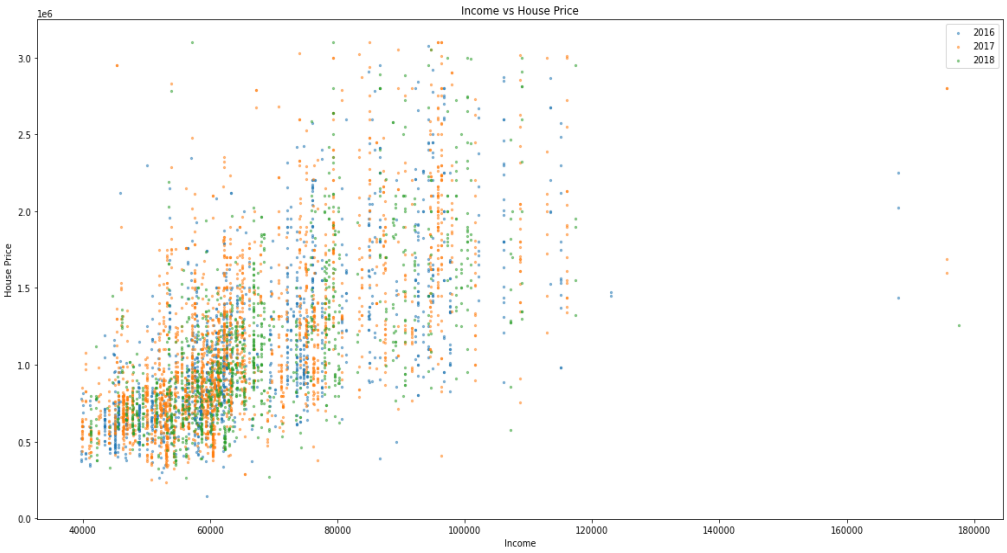

Figure12: Income vs House Price Scatter Plots

As the scatter plots represented in figure 12, we can find that for most of the suburb, the yearly income level and housing price is concentrated around 40k-61k level and 0.4 to 1.2e+06 respectively, which is consistent with the data in table 1. Basically, the trend can be viewed that most people choose to purchase a house that matches their income level. Moreover, there are few interesting patterns found. For yr2016, the people with relatively low-income level purchased the high housing prices, it might be related to the decreasing interest rate in yr2016[4], which encourages people to purchase a property. On the other hand, for the people with a 100k-120k income level, the housing prices of the suburb is relatively high. Also, the widespread plots within high income levels, reflecting the high-income people might purchase the houses covers the various prices. The house is not only a necessity but can be regarded as an investment.

Period	2016	2017	2018
Mean Yearly Income	58543.50	59678.54	61670.54
Mean Housing Price	0.95e+06	1.06e+06	1.04e+06

Table 1: mean yearly income from yr 2016-18

[4] (Arnold, 2016) (ceda, 2017)

Limitation and improvement

The limitations caused by 'housing price' dataset cannot be overlooked. First of all, the finding cannot reflect the current trend efficiently, since the period only convers year 2016-18, only presenting the previous trend and difficult to simulate the prediction model. In addition, the result cannot represent the entire Melbourne housing market circumstances completely, as it only contains the property for sale and only house data is extracted for analysis. Besides, the situations facing different positions across the suburb were not considered in our analysis, the sub-suburbs with detailed locations were simply grouped as a whole suburb. Especially for suburbs with wide fields, the finding is lacking accuracy.

In order to improve this investigation, the other factors such as tax, house loans should be taken into consideration to learn more about the financial stress bearing level of the house ownership. When purchasing a property, the accessibility to house loans could potentially be a significant consideration, particularly for the people with low-income level. Apart from that, given that the younger generation are more likely to rent a property than purchasing a house, other property types such as apartment and renting should be taken into account in the future analysis for complete investigation around inclusiveness and liveability.

Conclusion

Housing is important for most of us, to some extent housing is one of the necessities, it can help us increase happiness. Overall, after investigation. We discovered that household income does have significant impacts on house affordability. And it is clearly stated that people with lower income most likely will purchase a lower house price. However, the people who earn less still have the option to purchase a better house via home loan. This shows Melbourne has decent inclusiveness. Because people living in Melbourne with less income are still able to purchase a house even though it is very expensive.

Based on this question, the main stakeholders are the households, local government, immigrants who are interested in moving to Melbourne, and housing agencies. The findings would be of interest to the local government, as they may require this information to understand what kinds of benefits package or taxation regulation could

be taken to increase disposable income for citizens, thus increasing liveability and boosting the housing market.

References

Arnold, P. (2016, 01 18). *Lenders slash home loan rates in 2016, with just one catch*. Retrieved from RateCity: https://www.ratecity.com.au/home-loans/articles/lenders-slash-home-loan-rates-in-2016-with-just-one-catch

ceda. (2017). *POPULATION HOUSING AUSTRALIA.* Australia : ceda.

Personal Income in Australia. (2020, 12 16). Retrieved from Australian Bureau of Statistics: https://www.abs.gov.au/statistics/labour/earnings-and-work-hours/personal-income-australia/latest-release#data-download

Pino, T. (2018, 05 22). *Melbourne Housing market*. Retrieved from Kaggle: https://www.kaggle.com/anthonypino/melbourne-housing-market