**HDB Resale Affordability Prediction Using Time Series Regression in R**

**Aparna Sudarshan Parab**

**Basani Shruthi**

**Thet Myat Noe**

**Koh Chin Weng**

School of Computing and Information Systems

Singapore Management University

Singapore

**Abstract**

The Singapore housing market is experiencing a boom due to monetary policies both within and outside the country; housing affordability has thus become a key concern for the Singaporeans. This paper will attempt to thus predict housing prices and affordability to address these concerns, particularly for the HDB resale market who despite a global economic downturn due to the federal reserve’s tightening starting March 2022, has managed to see growth till the time this paper was written. The Autoregressive integrated moving average (ARIMA) model will be used to assess the affordability of housing by predicting the future value of Singapore’s ratio of household assets and liabilities. The ARIMA model will also be used to predict the future value of the 3 most popular HDB flat types, particularly the 3, 4 and 5 room units. Lastly, we will also perform a linear regression to assess the ability of the Singapore Overnight Rate Average (SORA) in controlling housing prices, as a tool the government uses to cool the housing market.

**Keywords**

Singapore housing, HDB resale, affordability, ARIMA, Linear Regression.

## 1. Introduction

The Singapore Housing and Development Board (HDB) flat resale prices have been on a general upward trend in recent years, particularly from 2020, mainly due to the delay in built-to-order (BTO) flats as supply chain issues plagued the construction of new flats during the COVID-19 pandemic (Ahmad 2021).

Chart, line chart

Description automatically generated

Figure 1. Price Index of HDB Resale Flats (hdb.gov.sg 2022)

Looking at historical Price index of HDB Resale Flats in Figure 1, it shows a general upward trend with housing market prices for HDB resale flats, which remain unphased even during the global financial crisis of 2007 – 2008, seeing only a slight dip in 2008 Q4 and rebounding immediately after. The only substantial decline in the price index occurred in 2013 when the Singapore Government introduced cooling measures such as a higher ABSD as well as a tighter loan to value limit (mas.gov.sg 2013). This bucks the trend of the private housing sector as shown in Figure 2. Which is seen to not only be impacted by the financial crisis of 2007 – 2008 but also the cooling measures implemented by the government in 2013.

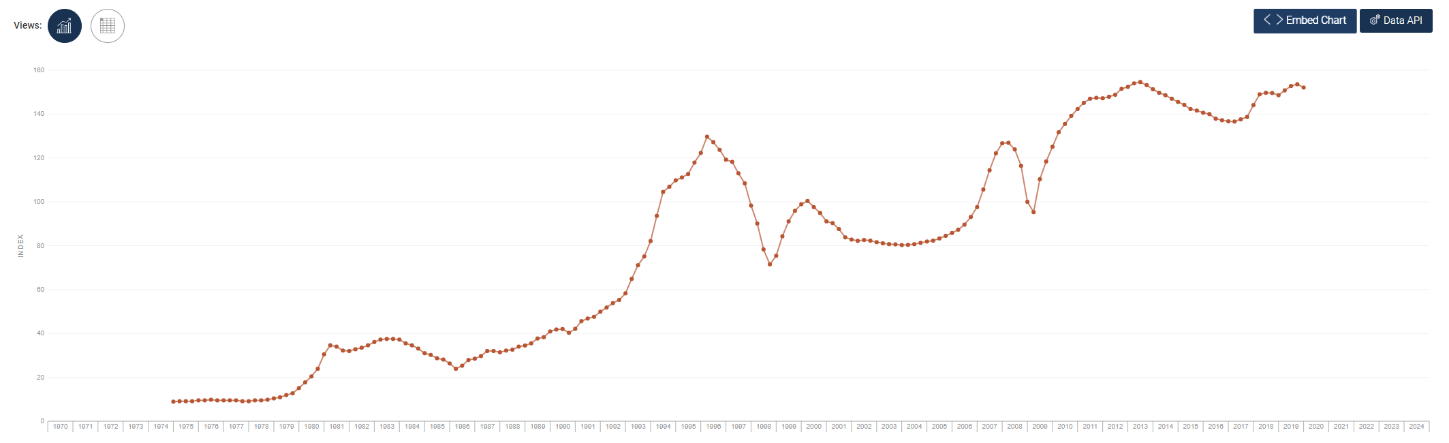


Figure 2. Private Residential Property Price Index (data.gov.sg 2020)

HDB resale flat prices soared 10.3% in 2022, 12.7% in 2021 and the fourth quarter of 2022 was the 11th consecutive quarter that HDB resale price index has increased (Chew 2023). This steep rise in HDB resale price trends in recent years present an opportunity to analyze and predict future HDB price trends which may be useful for potential house buyers and property investors in planning the best time and funds needed to buy a flat.

## 1.1 Objectives

The objective of the paper is to predict the future HDB residential resale price based on historical price data, as well as to predict if Singaporeans would still be able to afford housing when looking at their household mortgage and liabilities to asset ratio. Lastly, this paper will also assess the SORA rate’s ability of acting as a tool for the Singapore government to cool the housing market.

## 2. Literature Review

With the advent of measures taken during the pandemic by the Singapore government to ease the housing costs via the means of introducing a low interest rate environment during the pandemic and other forms of stimulus such as the resilience budget, which drew up to 17 billion of Singapore’s reserves; the largest stimulus package in the country’s history (Tham 2020). Furthermore, the monetary policies of the Singapore government coupled with the federal reserve’s stimulus to the global economy (through the means of the world reserve currency) has contributed to the rise in inflation and housing prices (Richard and Marc 2001). Hence from Milton’s literature, we shall assess the SORA rates, a monetary policy tool the Singapore government can utilize to cool the booming housing market.

Inflation, as mentioned previously is on the rise in Singapore, hence instead of taking the typical approach of assessing the debt-to-income ratio used by financial institutions to provide mortgage loans, we shall employ the use of the household assets, liabilities, and mortgage loan amounts to assess affordability. This is because apart from rising home prices, the general cost of living in Singapore is of paramount concern to its citizens as well at present (Aqil 2022), hence the need for us to assess the household not only by income, but by the total assets and liabilities the household possesses on top of mortgage loans.

The property market in Singapore has been one of the most attractive investment options for both locals and foreigners, seeing a steady upward trend in prices over the past few decades (Burgos 2022). This begs the question on the longevity of the trend, since prices were on the rise for 31 months (about 2 and a half years) in January and have since stabilized in the month of February, albeit analysts remaining bullish on prices.

To conclude, the aforementioned variables that influence price and affordability will be further investigated in the later part of the paper. The study will also assess other variables that could impact the cost of housing, such as the flat type, region, etc., to determine other potential variables that will impact the study's endeavors.

## 3. Methodology

Following the paper’s endeavor of prediction, the Autoregressive integrated moving average (ARIMA) model will be utilized for the study of how time series variables will perform in the future. The ARIMA (p, d, q) model is a linear model that fit to handle stochastic series; “p” refers to the autoregressive model AR and “q” the moving average model MA, lastly, “d” refers to the number of series difference (Chen et al. 2008). Prior to the use of the ARIMA model, we will perform the Box – Cox transformation on the data using R’s BoxCox.lambda function, this is to ensure that the usual assumptions for a linear model like the ARIMA model would hold (Li 2015); the Box – Cox transformation would essentially transform the data to allow the ARIMA model to deal with data with a non-constant variance (Seyyed et al. 2016).

The ARIMA modeling would then be conducted with R’s auto.arima; the auto.arima function returns the best ARIMA model according to the lowest AIC and BIC values. Post the use of the auto.arima function, we utilize the lambda generated with the previous Box – Cox transformation to transform the data and to predict its future value (Hyndman 2015). The auto.arima function uses the KPSS unit root test to determine the value of “d” and selects the “p”, “q” and “c” values by minimizing AICc.

We will also utilize the linear regression model to predict the value of a variable based on the value of another variable (ibm.com, 2023). This is done to investigate the impacts of variables such as the SORA rates on the price per square meter (PSM) by utilizing the independent variable PSM to find linear relationships with other variables.

## 4. Data Collection

Variables used in this project are obtained from data.gov.sg, ura.gov.sg, eservices.mas.gov.sg, singstat.gov.sg and consolidated in one data file (csv format).

Graphical user interface, text

Description automatically generated

Table 1. Dataset description

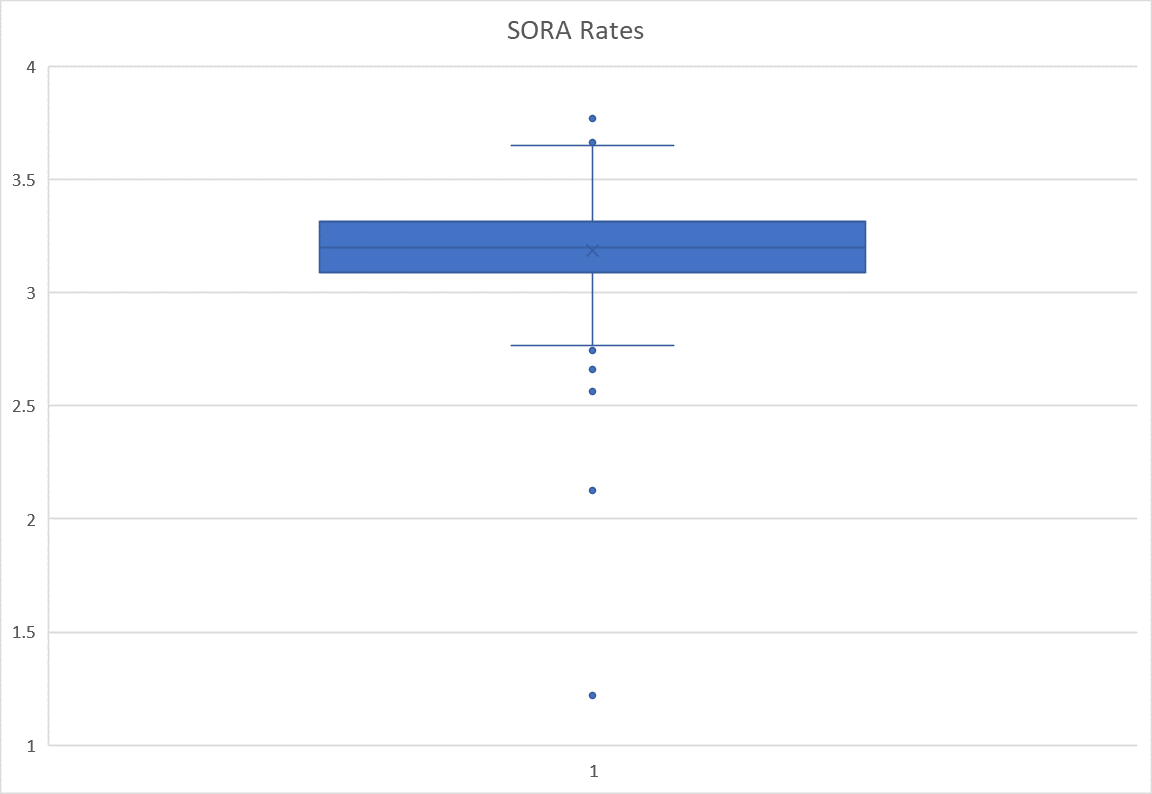
With the R libraries used as follow:

Table

Description automatically generated

Table 2: R libraries used in Analysis

We would also like to mention that when looking at the PSM, the data is aggregated through the means of median; PSM is used heavily as a variable to predict future price action or identifying plausible relationships with other variables. The aggregation is done as we understand that with the constraints such as the lack of sample size and other formatting constraints of the paper, that we are unable to take the region or town of the HDB resale transaction into account, the reason would be discussed later in section 5.1.   
  
We have however taken the median instead of the mean of the PSM when aggregating the data to mitigate the loss of accountability for the effects of extreme values due to random factors that are unlikely to persist over time, as substantiated by the literature from Galton’s reversion to mediocrity. That said however, the solution provided is not foolproof as pointed out by (Pal 2015). Nonetheless, we hope this paper serves as the basis for further literature development.   
  
To note, the daily SORA rates will be aggregated as well by calculating the monthly mean as the distribution unlike PSM, is fairly tame with few outliers as shown in the box plot below



## 5. Results and Discussion

## 5.1 Descriptive Statistics

Chart, histogram

Description automatically generated

Figure 3. Median HDB resale price per square meter

(disclaimer: low sample size for “Multi-Generation” and “1 Room” flats)

As shown in Figure 3. above, unsurprisingly the smaller flats cost more PSM when looking at the regular 1 to 5 room flats. Executive flats trace the price of the 4 and 5 room flats and all three HDB flat types have observed a similar trend since 2006 and it is worthwhile to note that these price PSM converge in 2019 – 2020. The overall trend since 2006 for all flat types remains on an upward trend, with prices PSM more than doubling in 16 years.

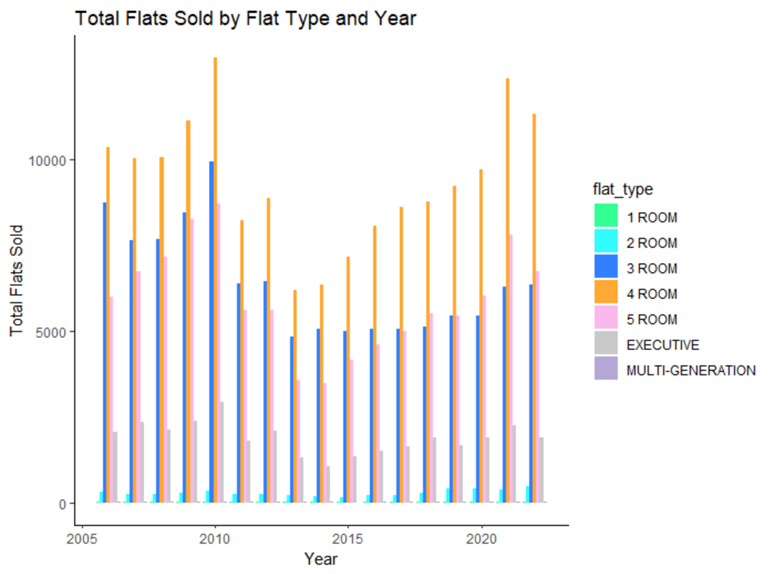


Figure 4. Total flats sold by flat types

(disclaimer: low sample size for “Multi-Generation” and “1 Room” flats)

Since 2005, the top 3 bestselling types of flats remains to be the 4, 3 and 5 room flats respectively. It can also be observed that while the price index mentioned above indicated little to no change during the financial crisis, it is observed that the overall number of transactions has dipped in 2010, plausibly from the aftermath of the financial crisis. Indicating sticky home prices even during a challenging economic scene. On the contrary, after the introduction of the cooling measures by the government in 2013, transactions for resale HDB have been on the rise since, reviving the previous momentum the HDB resale market has enjoyed.

Chart, box and whisker chart

Description automatically generated

Figure 5. Price per square meter by flat type boxplot

(disclaimer: low sample size for “Multi-Generation” and “1 Room” flats)

Looking at the boxplot, we zoom into the median prices and we can see that while the median PSM is relatively moderately leveled for 3 – 5 rooms and executive flats, multigenerational flats as well as 1 room flats are seeing higher medians and less outliers (this could be due to the small sample size) Lastly, it can be observed that the 4 Room flats typically have a higher bound for the upper outliers with the 3 – 5 room flats seeing about the same levels for their outliers; 5 room flats however have a larger upper and lower bound range as compared to the 3 room flats.

Chart, histogram

Description automatically generated

Figure 6. SORA Rates

The SORA Rates have been shown to be low during times of quantitative easing, post–the covid 19 pandemic and financial crisis of 2007–2008. That however has not been shown to have directly influenced the prices of the resale HDB market when comparing Figure 8. and Figure 1. although it is supposed to, we will thus investigate this phenomenon to ascertain the monetary policy tool’s ability to curb housing prices.

Chart, histogram

Description automatically generated

Figure 7. PSM by flat types by region

(disclaimer: low sample size for “Multi-Generation” and “1 Room” flats)

The Central region, unsurprisingly, has the highest median price compared to all other regions, except for the 2-room flats. It can also be observed that the bulk of the multi-generational flats are sold in the North region, which typically enjoys a lesser price premium as observed by the median price of the other flat types when compared to other regions, plausibly contributing to the erratic movement in the overall sales PSM shown above.

Chart, line chart, histogram

Description automatically generated

Figure 8. PSM by region

The PSM for housing sees a similar trend regardless of region in the past since 2006 and apart from the central region, there is no large distinctions in PSM between the regions. We will thus disregard the region as a variable for the rest of the study when assessing price and affordability considering our inability to cater to the analysis of regional data by flat types which was previously identified as a variable to be used in the predictive model.

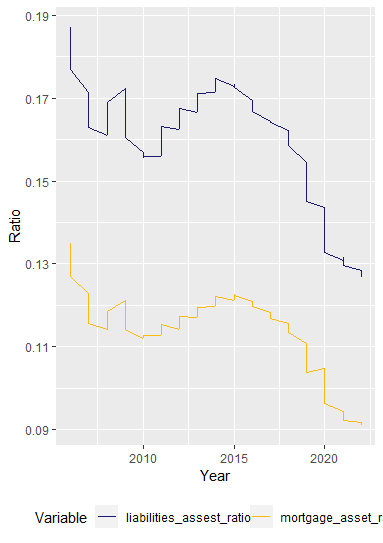


Figure 9. Household financial ratios

We have created a ratio with the Singaporean household mortgage and liabilities divided by household assets as shown in Figure 10., to act as a proxy for affordability; both ratios are on the decline since 2015, indicating that the Singaporean households should possess a higher ability to afford housing, whether or not that will remain to be the case will be explored during the predictive analysis section of the paper.

Hypotheses:

The following hypotheses are thus derived from our observations in the descriptive statistics section:

1. H0: Prices of public resale housing do not have a linear relationship with the SORA Rates

H1: Prices of public resale housing have a linear relationship with the SORA Rates

1. H0: Singaporeans are not at risk of being able to afford public housing

H1: Singaporeans are at risk of being able to afford public housing

1. H0: The prices of HDB Resale flats will not be on an upward trajectory in the event of a lack of cooling measures.

H1: The prices of HDB Resale flats will be on an upward trajectory in the event of a lack of cooling measures

## 5.2 Inferential Statistics

5.2.1 Hypothesis 1

We first run a correlation test to test for multicollinearity between the SORA Rates and PSM, and it can be observed with the correlation matrix (Figure 10.) below that the two variables involved do not suffer from multicollinearity.

A picture containing chart

Description automatically generated

Figure 10. Correlation Matrix

The linear regression has highlighted that the SORA rate has a linear relationship with the PSM, however, the R Squared is low at a value of 0.121 and the correlation is relatively weak at -0.348. Thus, showing that the SORA rate is not an effective tool for the government to use when implementing cooling measures in the future for HDB resale prices

5.2.1 Hypothesis 2

|  |  |
| --- | --- |
| image | image |
| Figure11. ARIMA for Mortgage to Assets ratio | Figure 12. ARIMA for Liabilities to Assets ratio |

From the ARIMA models created, it is observed that based on the confidence interval of 95%, the forecast of the Mortgage to Assets ratio will flatline, while the Liabilities to Assets ratio will decline. This indicates that the typical Singaporean household will maintain the present level of affordability for their mortgage liabilities but have a higher propensity of lowering their total household liabilities in relation to their household assets. It is however also important to note that the forecast may vary according to the 95% confidence intervals as shown in the charts above.

To conclude hypothesis 2, the Lljung box test was also done, with a resulting pvalue of >0.05 for the 2 ARIMA models, indicating that both models disallow us to reject the test’s null hypothesis of the presence of autocorrelation, thus validating the ARIMA models. The Mortgage to Asset ratio also has a lower BIC of –1990 than the Liabilities to Asset ratio –1886, making it the better model of the 2.

5.2.3 Hypothesis 3

First, to predict the prices of the top 3 most transacted HDB resale flat types (3 Room, 4 Room, 5 Room) as shown in descriptive session, ARIMA modelling of price (per square meter) time series and forecast function was used.

Times series data was first decomposed to check trend and seasonality.

|  |  |  |
| --- | --- | --- |
|  |  |  |
| Figure 13: ARIMA for 3 room HDB | Figure 14: ARIMA for 4 room HDB | Figure 15: ARIMA for 5 room HDB |

From the ARIMA models created, it is observed that based on the confidence interval of 95%, the future value of 3 and 4 room resale HDB’s PSM could rise above the range of 8000 or fall below that of 3000 or 4000 after 8 years respectively; the forecasted PSM for both the 3 and 4 room resale HDBs are shown to rise in the near term, eventually tapering off and flattening respectively.

The 5 room flats however show a different trend, with the forecast predicting that the PSM will rise in the next 8 years with the lower bound showing that PSM will remain relatively at the present level and the upper bound could eventually even rise to 10000 PSM

To conclude hypothesis 3, the Lljung box test was also done, with a resulting pvalue of >0.05 for all 3 ARIMA models, indicating that all the models disallow us to reject the test’s null hypothesis of the presence of autocorrelation, thus validating the ARIMA models

## 5.3 Proposed Improvements

The models having passed their respective validation tests are proof that the paper has indeed accomplished what it has set out to do, however, as mentioned in the methodology section of the paper, the regression models were all created with aggregated data, due to the limitations in resources available, further literature should consider this study as a base for development, especially on the subject of location of the HDB flats, (i.e. Region or Town) as those variables have been shown to influence the prices of the HDB resale flats.

## 6. Conclusion

To conclude, absent of variables outside of the scope of study, our models have shown that HDB Resale prices will continue to rise, or decline based on their flat types. Singaporeans might also face affordability issues for HDB resale flats within a 95% confidence interval, but the forecast shows that HDB resale flats will remain affordable as substantiated by the graph in Figure 10. Lastly, the study has also shown that the government should stick to utilizing cooling measures to control the HDB resale prices as there is little correlation and a small R squared between the SORA rates and PSM.

## References (12 font)

Ahmad, Z., IN FOCUS: What is pushing HDB resale prices higher?, Available: https://www.channelnewsasia.com/singapore/focus-what-pushing-hdb-resale-prices-higher-2029296, Jul 10, 2021, Accessed on Mar 09, 2023

Aqil, H. M., Government will do 'utmost' to help Singaporeans cope with cost of living concerns: Lawrence Wong, Available: <https://www.channelnewsasia.com/singapore/gst-inflation-cost-living-budget-lawrence-wong-2512791>, Feb 23, 2022, Accessed on Mar 18, 2023.

Burgos, J., Singapore Housing Market Defies Global Downturn And Curbs Amid Soaring Rents, Available: https://www.forbes.com/sites/jonathanburgos/2022/10/31/singapore-housing-market-defies-global-downturn-and-curbs-amid-soaring-rents/?sh=3e31e7101869, Oct 31, 2022, Accessed on Mar 18, 2023.

Chen, P., Yuan, H., and Shu, X, Forecasting Crime Using the ARIMA Model, Proceedings - 5th International Conference on Fuzzy Systems and Knowledge Discovery, FSKD 2008. 5. 627 - 630. 10.1109/FSKD.2008.222.

Li, P., Box-Cox Transformations: An Overview, Available: https://www.researchgate.net/profile/Chandra-Swami-Telgo/post/Whar\_is\_the\_alternative\_for\_non-normally\_distributed\_data\_in\_mixed\_design/attachment/59d64d6d79197b80779a6fc6/AS%3A488088414756864%401493380754638/download/Box-Cox+Transformations\_An+Overview.pdf, Apr 11, 2015, Accessed on Mar 18, 2023.

Chew, H., HDB resale flat prices up 10.3% in 2022, slower than 12.7% increase in 2021, Available: https://www.channelnewsasia.com/singapore/hdb-flat-resale-price-index-rpi-2022-sales-volume-bto-3179506, Jan 1, 2023, Accessed on Mar 09, 2023

data.gov.sg, Available: https://data.gov.sg/dataset/private-residential-property-price-index-by-type-of-property, Jun 9, 2020, Accessed on Mar 17, 2023

hdb.gov.sg, Available: https://www.hdb.gov.sg/residential/selling-a-flat/overview/resale-statistics, Jul 20, 2022,el Accessed on Mar 09, 2023

Hyndman, R. J., Automatic algorithms foe time series forecasting, Available: Google-Oct2015-part1.pdf (robjhyndman.com), 2015, Accessed on Mar 18, 2023.

Ibm.com, Linear Regression, Available: https://www.ibm.com/topics/linear-regression#:~:text=Resources-,What%20is%20linear%20regression%3F,is%20called%20the%20independent%20variable, Accessed on Mar 18, 2023.

Liew, I., HDB resale prices rise 2.3% in Q4, slowest increase in 2022, Available: https://www.straitstimes.com/singapore/housing/hdb-resale-prices-rise-23-in-q4-slowest-increase-in-2022, Feb 1, 2023, Accessed on Mar 09, 2023

mas.gov.sg, Available: <https://www.mas.gov.sg/news/media-releases/2013/additional-measures-to-ensure-a-stable-and-sustainable-property-market>, Jan 11, 2013, Accessed on Mar 17, 2023

Pal, M., Mean Reversion Framework. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=2607963, May 20, 2015, Accessed on Mar 18.

Richard C. K. B., and Marc D. W., Inflation Is Always and Everywhere a Monetary Phenomenon: Richmond vs. Houston in 1864, *The American Economic Review*, vol. 91, no. 5, pp. 1621–30, 2001.

Seyyed, M. T. N., Mahboubeh, N. and Ebrahaim, A. G., A Novel DoS and DDoS Attacks Detection Algorithm Using ARIMA Time Series Model and Chaotic System in Computer Networks, *IEEE Communications Letters*. 20. 1-1. 10.1109/LCOMM.2016.2517622.

Tham, Y., Covid-19 aid: $17 billion to be drawn from reserves for stimulus measures, Available: https://www.straitstimes.com/politics/covid-19-aid-17-billion-to-be-drawn-from-reserves-for-stimulus-measures, Mar 26, 2020, Accessed on Mar 17, 2023.

**Appendix**

**Shiny App:**

Shiny app developed for this project can be used to check the Popular flat type sold per region and year, boxplot of HDB resale prices in each region for all 5 flat types per year, correlation matrix of variable that could influence HDB resale price and linear regression of Resale price vs selected variables.

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated

Chart, bar chart

Description automatically generated

Chart

Description automatically generated

Chart, histogram

Description automatically generated

Graphical user interface

Description automatically generated