



**PROPOSAL**

**SM 4290 – FINAL YEAR PROJECT**

**TITLE:**

**SPATIAL CONCENTRATION ANALYSIS: FACTORS AFFECTING BRUNEI'S HOUSE PRICES**

**RESEARCHER:**

Atikah Farhain Binti Haji Yahya

NO. I.D: 18B9024

Mathematics

**SUPERVISOR:**

Dr Haziq Jamil

Assistant Professor in Statistics

Mathematical Sciences, Faculty of Science

**SUBMISSION DATE:**

15<sup>th</sup> April, 2022

Universiti Brunei Darussalam

Tungku Link BE1410

Negara Brunei Darussalam

## **Abstract**

Real estate ownership becomes more of a challenge year by year due to aspects such as inflation, increase in unemployment, scarcity of land ownership amongst newer successors and so on (Girouard et al., 2006). The perspective this research paper takes upon is to analyse house prices in Brunei. A geospatial analysis heat map distinguishes the collective price points of housing areas in Brunei, where warmer colors indicate high costs and dissipating tones mean housing of lower costs. There is also a relation between the houses' characteristic variables to its cost. This includes the district it is in, the distance from its residential location to surrounding crucial amenities, the size of the house, including the number of bedrooms and bathrooms/toilets. In order to find how each parameter contributes to deciding the finalised price, a relationship between the cost and its criteria is put through linear regression through R to determine its correlation. Moreover, this report also explores the commission value of real estate dealers by comparing houses sold by agents with ones that are sold by individual private sellers.

**Keywords:** Real estate, house pricing, linear regression, geospatial analysis, heat map

## **ELEMENTS OF RESEACH PROPOSAL**

Abstract

1. Title
2. Introduction
  - a. General Introduction
  - b. Study Problem
3. Literature Review
4. Research Methods
  - a. Subjects
    - i. Inclusion/Exclusion Criteria
    - ii. Sampling
  - b. Data collection
    - i. Variables
    - ii. Measures/Instrument
    - iii. Procedure
  - c. Statistical considerations
    - i. Sample size
    - ii. Data analysis
5. Ethical Considerations
6. Work Plan
7. Summary

## 1. Title

Spatial concentration analysis: Factors affecting Brunei's house prices

## 2. Introduction

### I. General Background

Modernisation of time calls for the social competitiveness towards achieving the best fulfillment in life. Although fulfillment is assimilated in many materialistic possessions, a proper home acts as one of the biggest indicators of one's success. Interesting enough, over many academically written papers published, it has been indefinite how the word "home" came across writings. For example, Coolen & Meesters (2012) would describe it as an infrastructure that contains the bare minimum of what is considered to be labelled as a house. An enclosed structure with enough space and possessions to sit on, lie down and perform necessary actions in pursuit of survival. In contrast to Blunt & Varley (2004) where they believe the concept of home is not based upon any structural root, but is related to how one personally feels. However, a most suitable description of a home in today's modernised times is when Noor Hasharina Hassan (2017) highlighted a strong social concept around owning a home inclusive of one's social status as she integrated it as one's identity altogether.

With that retrospect, it paves way to how the infrastructure of houses in Brunei are curated. Aspects such as race, culture, lifestyles, occupation and time may factor in the prospect of what makes a house desirable (Hite, 2010). However, this report is interested in exploring is the relationship between characteristics of a house and how that influences its selling price. Other aspects that might affect the cost of a house is its accessibilities to crucial amenities in its mukim. Another additional information that we are interested to venture into is the rate of commissions real estate agents earn based on similar housing prospect with used houses' price points.

### II. Study Problem

Purpose:

The purpose of this study is to be able to explain what factors affect the house pricing in Brunei.

Aim:

The aim of this research project is to analyse the trends in housing prices in Brunei and to understand what characteristic adds significant cost to said real estate.

Objectives:

1. Visualise geospatial map of housing prices.
2. Study factors affecting house prices based on the relationship between variables.
3. Justify the highest significant criteria of a house that adds value to the cost.
4. Deducing the commission of real estate dealers compared to individual private home sellers.

### **3. Literature Review**

House pricing studies has been done in various approaches where results were based off of factors that may affect certain areas and countries. Clark et al. (2006) believed that a nuclear plant in California poses detrimental hazards that would bid house prices at significantly lower price points. Another similar research studies the house pricing in Warsaw near close proximity of the Chopin Airport might cost less due to high noise levels that proves to be a nuisance (Cellmer, Belej, & Konowalczyk, 2019). However, prices can stagger up due to rapid urbanization and fast-changing structural reconditioning that serves to be the case in Guangzhou, China (Chen, Zhuang, & Zhang, 2020). These kinds of studies are predominantly carried out using geospatial analysis and regression models that help rationalise the relationship between house prices and their respective explanatory variable. However, in this report, it looks into the relationship of house prices and its level of accessibilities in a mukim, much similar to a research study made in Dailan City, China (Yang et al., 2018).

Another area of interest in this report is the heat map designed to visualise Brunei's areas of high housing bids. Komagome-Towne (2016) proved that heat maps of single-family housing prices can help predict

the future of house pricings in the area over the years to come. Similar to a study done by Barreca, Fregonara and Ronaldo (2021) where they use geospatial heat map to structure implications that distinguishes the relationship between house prices and the energy performance. What this report wants to visualise is the understanding of what variables are highly influencing the house pricings.

#### **4. Research Methods**

##### **I. Subjects**

###### **a. Inclusion/exclusion criteria**

It is fundamental to ensure certain criteria should be inclusive in this research in order to fulfill the objectives of this study. In hinge sight, important criteria should be:

- Residential accommodation either newly build or used
- Proposed houses that have not been built
- Houses which are under construction and in progress to completion

However, the exclusion criteria should be:

- Advertised houses that does not have critical information in helping with the research
- Houses or apartments that are for rent or that lies on leased land
- Any land on sale as there are no similar parameters to compare
- Government housing that does not come with a selling price

###### **b. Sampling**

The sampling system works using a non-systematic sampling, where unlike conventional methods of running all respective data through a population probability with respect to the inclusion and exclusion criteria, this method chooses to take samples of an area to represent a bigger part of the data. According to Huelin et al. (2015), a distinctive feature of a systematic sampling is its minimization of bias information

to be able to represent a more comprehensive, broad-based findings, while the non-systematic sampling finds being informational as more of a priority, and does not cover the overall view of a population.

Specifically, the kind of sampling method that will be executed is a multi-stage sampling. Similar like creating a census block, a small number of house data will be taken in a specific mukim, which will then represent the overall data.

## **II. Data Collection**

### **a. Variables**

Important variables to consider in producing the right results should include: district of where the house is located, the specific town it is in, the land size, the type of house, number of floors the house has and the number of bedrooms and bathrooms it contains. These are readily available information that could be found on the Instagram posts. However, it is interesting to consider additional data such as a house's distance and the crucial amenities found nearby. Although this information is sometimes provided by real estate dealers, it is usually in an unreasonable measurement i.e. estimated minutes. Instead, we are more interested in the distance when it is seen in kilometers.

### **b. Measures/Instruments**

A special statistical software is used throughout this research called 'R'. Its extensive and rich features are capable of computing the data and provide findings that could reach the objectives and aim of this research proposal. According to Lüdecke et al. (2020), features in R such as parameters, provides a user-friendly yet comprehensive method in standardizing estimates while streamlining the overall results of statistical models. This proves to be more than sufficient in tackling data for both regression and spatial autocorrelation analysis.

### **c. Procedures**

Below establishes the procedure of how to appropriate findings for this research:

1. Manual scraping: the data is manually scraped on Instagram individually. This type of scraping is preferred in comparison to scraping using a software (i.e. Python or R) because majority of the information is placed on the photos instead of as text form. Due to time and resources constraints, it is a much-preferred method.
2. Data cleaning: during data collection, there will be missing data that might not be available in all house postings. Therefore, a further clean-up is performed where repeated data are eliminated (multiplicity and duplicates) and some missing data is retrieved through a process called imputation. There will also be further data cleaning by manually estimating the distance between houses and crucial amenities nearby. This is executed through individually tracing the route of the house and nearby amenities such as mosques, schools and/or commercial areas. This will be recorded in kilometers.
3. Construct mathematical model: The construction of mathematical model is based on the relationship between certain variables with one another. In this report, it is fundamental to distinguish a model that can relate house pricing and the criteria that the house attain.
4. Data analysis: The mathematical model will particularly be used in deducing the data analysis that will be explained further in the section below under b. Data analysis.

## **V. Statistical considerations**

### **a. Sample size**

There will be a total of 200 house data that will be scraped from Instagram. According to the most recent records of Brunei's number of housing released by Ministry of Finance and Economy under Brunei Darussalam Population and Housing Consensus (2016), there is a total of 75,000 units of houses estimated in Brunei. Table 1 shows the total number of estimated houses according to district:



Table 1: Number of housing unit according to district

District	Bandar Muara	Belait	Tutong	Temburong
Total Number of Housing Unit	51,000	13,000	9,000	2,000

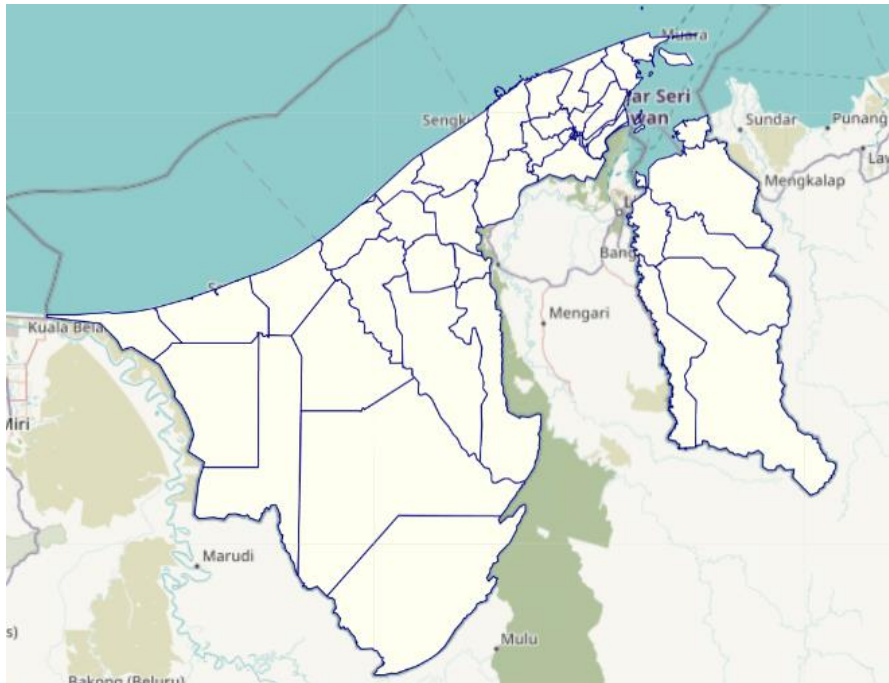
Therefore, the probability of inclusion based on the number of samples is  $\frac{200}{75000}$  or 26.67%.

### **b. Data analysis**

Various data analysis will be exercised in order to achieve the goal of this research. Such methods include:

1. Geospatial autocorrelation analysis:

A geospatial heat map can give an overall glance of where Brunei's high-cost housing areas are located. Additionally, densely built houses can also be reflected using geospatial heat map. Figure 1 shows a map of Brunei with its mukim highlighted. In this report, the geospatial autocorrelation analysis will look into highlighting how each mukim behaves in accordance to their distance with each critical amenity. According to b Mohamad @ Masri, b Nawawi and b Sipan (2016), a location of a settlement takes priority of various factors and one of it is its accessibility. The more accessibility it has towards amenities, the more desired an area becomes. What this report focuses on is whether house prices play into each mukim's accessibility.



*Figure 1 - A map of Brunei highlighting all mukim for geospatial heat map*

## 2. Linear regression:

Once a mathematical model has been distinguished, it is required for explanatory variables to respond with dependent variables in order to show the relationship between them. This is a good tool to see how criteria of houses can influence the prices, and how much commissions could be earned by real estate dealers when compared to private individual house sellers.

## **5. Ethical Considerations**

This section holds a strong significance in researches involving protecting the identity of people involved in gathering the data to support findings in this report (Connelley, 2014). However, data collection required for this study can be accessed by the public and is not limited to certain groups of viewers. In a sense, it is not required for any ethical considerations as there are no integrity to be protected for this specific method of research.

## **6. Work Plan**

This research will require proper allocation of time to invest in each section of the report. Given in figure 2 the full schedule of how each task is sequenced over a period of two academic semesters in Universiti Brunei Darussalam between January of 2022 to November at the end of the year.

RESEARCH PROJECT TIMELINE														
Project details				Timeline: January - November 2022										
No.	Activity	Estimated duration	Actual duration	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
1	Identify research area of interest	1 week	3 days	■										
2	Background readings and literature	On-going	On-going	■	■	■	■	■	■	■	■	■	■	■
3	Formulate aims and objectives	1 week	1 week	■										
4	Significance and rationale of research	1 week	1 week	■										
5	Literature review	8 weeks	2 weeks	■	■									
6	Research design (methodology)	4 weeks			■	■								
7	Submission of research proposal	15/4/22	15/4/22				■							
8	Conduct data scraping	6 months	On-going	■	■	■	■	■	■					
9	Application of mathematical models and regression	3 months	NIL						■	■	■			
10	Discussion and analysis	3 month	NIL								■	■	■	■
11	Conclusion	1 week	NIL										■	
12	Submission and presentation of research project	NIL	NIL											■

Figure 2 - Project Timeline between January to November 2022

## 7. Summary

Brunei’s housing prices can be justified through independent variables along with its explanatory variables and conducting regression models to distinguish the relationship of house prices and factors affecting it. The data is gathered through the process of scraping Instagram throughout a span of six months and used to develop the mathematical model that can plot Brunei’s house pricings in a geospatial heat map. Ultimately, once the relationship is distinguished between the factors affecting and the house prices, it is important to see which variable greatly influenced the prices. Moreover, it is possible to see the commission rates of real estate agents by the relationship of house prices sold by agents (newly build houses) and used houses sold by individual private sellers.

## **8. References**

- b Mohamad @ Masri, M., b Nawawi, A., & b Sipan, I. (2016). Review of Building, Locational, Neighbourhood Qualities Affecting . *Procedia - Social and Behavioral Sciences*, 234, 452-460.
- Blunt, A., & Varley, A. (2004). Geographies of Home. *Cultural geographies*, 3-6.
- Cellmer, R., Belej, M., & Konowalczyk, J. (2019). Impact of a Vicinity of Airport on the Prices of. *International Journal of Geo-Information*, 8 (11), 471.
- Chen, S., Zhuang, D., & Zhang, H. (2020). GIS-Based Spatial Autocorrelation Analysis of Housing Prices Oriented towards a View of Spatiotemporal Homogeneity and Nonstationarity: A Case Study of Guangzhou, China. *Complexity*.
- Clark, D., Michelbrink, L., Allison, T., & Metz, W. (2006). Nuclear Power Plants and Residential Housing Prices. *Growth and Change*, Volume 28, Issue 4, p. 496-519.
- Connelley, L. M. (2014). Ethical considerations in research studies. *MedSurg Nursing*, Vol. 23, Issue 1.
- Coolen, H., & Meesters, J. (2012). Editorial special issue: house, home and dwelling. *Journal of Housing and the Built Environment*, 1-10.
- Girouard, N., Kennedy, M., den Noord, P., & André, C. (2006). *Recent House Price Developments: The Role of Fundamentals*. Paris, France: OECD Economics Department Working Papers.
- Hassan, N. H. (2017). *Housing Matters: The Value of Home Ownership* . Gadong, Brunei Darussalam: Institute of Asian Studies, Universiti Brunei Darussalam.
- Hite, D. (2010). *House Price Impacts of Racial*.. Cincinnati: University of Cincinnati Economics Working Papers Series.
- Huelin, R., Iheanacho, I., Payne, K., & Sandman, K. (2015). *What's in a Name? Systematic and Non-Systematic Literature Reviews, And Why The Distinction Matters?* Los Angeles: PRISMA-P Group.
- Komagome-Towne, A. (2016). *Models and visualizations for housing price prediction*. Pomona: California State Polytechnic University.
- Lüdecke, D., Ben-Shachar, M. S., Patil, I., & Makowski, D. (2020). Extracting, Computing and Exploring the Parameters of. *The Journal Of Open Source Software*, 5(53), 2445.
- Yang, J., Bao, Y., Zhang, Y., Li, X., & Ge, Q. (2018). Impact of Accessibility on Housing Prices in Dalian City of China Based on a Geographically Weighted Regression Model. *Chinese Geographical Science*, 28, 505-515.