

Cafe market share using satellite image data and Google Database in Malang City

Abdul Wahid Hasyim^{1,*}, Eddi Basuki Kurniawan¹, Wulan Dwi Purnamasari¹

¹ Urban and Regional Planning Engineering Department, Universitas Brawijaya, Malang, 65145, Indonesia

*awahidhs@ub.ac.id

Received 13-04-2022; accepted 29-04-2022

Abstract. Cafes in Malang City have a very large quantity. Almost every main road corridor in Malang City has a small, medium, or large cafe. In a competitive situation, location factors can be critical, making it very important. Therefore, a study is needed to analyze the market share of cafes in Malang City from its spatial elements. This study aims to determine the spatial factors that affect the market share of cafes based on sample data in the form of upper-middle-class cafes in Malang City. The variables used to measure the market share of the café sample spatially are proximity to universities, building density, road hierarchy, rating, and several competitors. This study uses Google Satellite Image Data sources and remote sensing methods for processing spatial data and analyzes it using Multiple Linear Regression Analysis. Based on the analysis results, it is known that all variables are positive. The variables that have the most influence on the market share of the café sample are building density, university proximity, and road hierarchy.

Keywords: Café, Market Share, Google Database, Remote Sensing

1. Introduction

Malang City is a city located in East Java, Indonesia. The city is located 90 km south of Surabaya and is the second-largest city in East Java after Surabaya, and is one of the largest cities in Indonesia by population. In addition, Malang City is also considered a city located in a strategic area. The Malang City area has much potential in terms of physical condition, location, and function. It will impact the economic growth of Malang City itself [1]. The strategic form is natural and artificial tourism facilities, attracting tourists to visit Malang City.

Learning facilities and facilities in Malang City have caused Malang City to be nicknamed the City of Education, which has a service scale up to the national level. So that many students from outside Malang City attend education, especially education at public or private universities [2]. The high number of people entering Malang City due to several factors causes the total population of Malang City to increase rapidly. Of course, this event is the main trigger for increasing economic activity in various sub-sectors in Malang City [3]. Previous researchers have predicted the growth of the city of Malang with an Artificial Neural Network (ANN) model using MOLUSCE, a QGIS plugin. The study aims to describe the town's growth in Malang City over 24 years and to predict the city's future development

using the ANN model for 2027. The results show that the built area will increase by 11.79% from 2015 to 2027. In 2027, the built area will cover the city by 73.21% of the city area [4].

Spatially, one of the strongest impacts is the amount of land that has the potential to be used for business activity. One manifestation of this impact is the emergence of businesses in the cafe and restaurant sector. Cafes in Malang City have a very large quantity. Almost every main road corridor in Malang City has a small, medium, or large cafe. Based on GDP Malang City (2021) data, Café and Restaurant is one of the sectors contributing the highest GRDP in Malang City. So this sector should get better attention from several stakeholders, especially the government and private parties.

The location of each cafe will also determine the amount of opportunity for the sale of beverage products provided. If the cafe's location is close to the center of the youth crowd, it is considered to have the potential to dominate the market share [5]. Existing the distribution of cafes in Malang tends to be closer to the center of the educational area. Such as Soekarno Hatta street, close to Brawijaya University, Sigura-Gura Corridor street, close to ITN University, and Jalan Ijen corridor, close to Malang State University. It causes these cafes to have very tight competitiveness. Every cafe often has to innovate to produce more attractive and unique visual cafes to win the so-called market share [6].

Location factors can be critical in a competitive situation, making it very important [7]. So that business actors will correctly determine the potential location for establishing a café business and not fall in its market share. Previous studies also found that the business area is closely related to the success of the business [8]. This study aims to determine the spatial factors that affect the market share of cafes based on sample data in the form of upper-middle-class cafes in Malang City.

2. Material and Methods

The research location of this research is in Malang City. Geographically, Malang City is located in the southern part of the eastern part of Java Island and has an area of 145.28 km². Malang city is situated between 440-667 meters above sea level. Malang City is located in the middle of Malang Regency, which is situated at 112.06°-112.07° east longitude and 7.06°-8.02° south latitude.

The satellite image data used in this study is Google Satellite imagery with a zoom specification of level 22 and has natural color. Satellite image data is processed radiometrically (Colour Correction) to provide an overview of the existing surface of Malang City and slightly processed geometrically to provide a higher level of accuracy. The main purpose of using Google Satellite Image Data is to make precise and accurate variable data needed in this study, such as building density and the distance of the café to the university.

The sample used in this study is a middle-class cafe and above. It is indicated by several factors such as rating, number of reviews, and the cafe's physical condition, which can be found on the Google Maps database. Databases on google maps, especially cafes in Malang City, will be extracted into shapefiles that can be inputted into GIS software. Then the data will be selected according to the middle and upper cafe criteria. Table 1 is the criteria for middle and upper cafes.

Table 1. Criteria for upper-middle café

Criteria		Penilaian
1.	Review	Number of middle and upper cafe reviews > 500 reviews
2.	Rating	Café has a rating value of >4
3	Physical condition cafe	The Middle and upper cafe has several facilities (such as Café parking area, wifi, toilets, waiters, etc.) and modern interior/exterior design)

The variables used in this study are the factors that are thought to be the determinants of market share in the spatial element based on the theoretical studies that have been carried out. Table 2 shows the variables used in the study.

Table 2. Research variable

	Variables	Categories	Required Data
1.	X1 : Distance of cafe to university	4 : Close to 3 Universities 3 : Close to 2 Universities 2 : Close to 1 University 1 : Far from university	<ul style="list-style-type: none"> • SHP Data Selected Café Location • Buffer 1 km per location point • University Area SHP
2.	X2 : Building Density	4 : 75%-100% 3 : 50%-75% 2 : 25%-50% 1 : 0%-25%	<ul style="list-style-type: none"> • SHP Data for the selected Café Location • Buffer 1 km per location point • Malang City Building SHP
3	X3 : Road Hierarchy	4: Arteries 3: Collector 2: Collector 1: Neighborhood Road	<ul style="list-style-type: none"> • SHP Data for the selected Café Location • SHP of Malang City Road Agency
4	X4 : <i>rating</i>	-	Café Rating Data on Google Map Database
5	X5 : Number of Rivals	4: No Rivals 3: There is 1 Rival 2: There are 2 Rivals 1: More than 3 Rivals	<ul style="list-style-type: none"> • SHP Data for the selected Café Location • Buffer 1 km per location point
6	Y : Number of Reviews (Assumed As Number of visitors)	-	Café Review Data on Google Map Database

To determine the effect of the independent variable on the dependent variable, an analytical method is used that can describe the equation model. The analytical method used is Multiple Linear Regression Analysis. The classical assumptions were first tested on the compiled data at the multiple linear regression analysis stages. Suppose the data has passed the classical assumption test. In that case, the data is feasible to be processed to produce an equation model. The resulting equation model will also undergo a model feasibility test process so that the model is possible to use.

3. Result and Discussion

Several data acquisition scenarios have been designed according to research needs. Data compiled using several methods, GIS and Remote Sensing, will be rearranged to form a data variable. These variables will later be reprocessed using multiple linear regression analysis to find a model of the influence of variables on the market share of cafes in Malang City.

3.1. Distribution of Malang City Café

Data on the distribution of cafes in Malang City is obtained by extracting the google map database into a file with SHP format that can be inputted into GIS software. The extracted cafe location data is

Based on the results of identification using GIS software, it is obtained that the scoring results of the categorization of the proximity of the university to each cafe sample are obtained. Most of the sample cafes in Malang (15 cafes) are included in the category close to 1 university or a score of 2. Meanwhile, there is only one university, namely UB Coffee, for those in the category close to 3 universities or more.

3.3. Building Density

Building density is measured based on the number of buildings contained in a one km buffer from each café point (Figure 3).

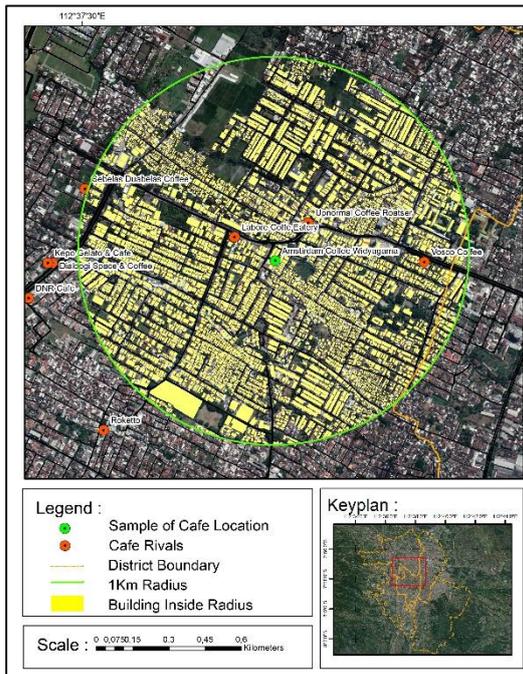


Figure 3. Sample map of building density around Café Amsterdam Coffee

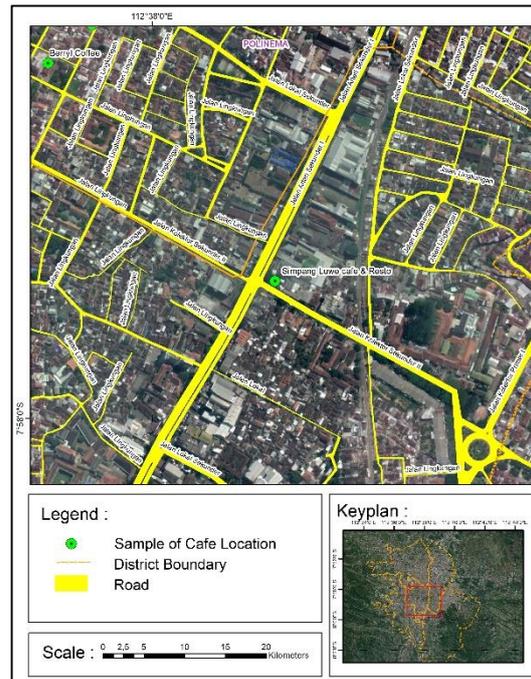


Figure 4. Street hierarchy map at Simpang Luwe Café & Resto as a sample

Based on the results of the calculation of building density in a one km buffer in each sample of cafes in Malang City, it was found that most of them were at a building density of 75%-100% or with a score of 4. or a score of 1 only 1 cafe, namely Classic Chocolate.

3.4. Path Hierarchy

Road hierarchy is also assessed as a factor that can determine market share from a spatial point of view. It is related to the ease of accessibility of visitors to the cafe location. The higher the road class, the easier it is for cafe visitors to visit the cafe location. The determination of the road hierarchy score is based on the road class in front of the cafe.

The identification of the road hierarchy in each sample of cafes in Malang (Figure 4) found that most of them had collector class road access or a score of 3. The number of cafes with collector class road access was 13 cafes, while only five cafes with arterial road class accessibility.

3.5. Rating

The rating of the sample cafe in Malang City was mostly has a rating between 4.3 to 5. The number of cafes with this rating is 25 Café. The highest rating are Amsterdam Coffee, Beryl Coffee, Starbuck Blimbing, Dialogi Space & Coffee, Starbuck Ijen Street, and Starbucks Malang Town Square, each of which has a rating of 4.7. Meanwhile, the lowest Malang cafe sample rating with a rating value of 4.3

is Cafe Litchi, Social Palace Café, DNR Café, Upnormal Coffee Roatser, Kepo Gelato & Café, and Simpang Luwe cafe & Resto.

3.6. Number of Rivals

The number of café rivals is based on the location of the selected sample rival café in a one km buffer as a measure of café proximity (Figure 5). The existence of other cafes in the same class at locations below one km from the main cafe location will also affect the market share of the café [9]. So the variable number of competitors is essential for this study.

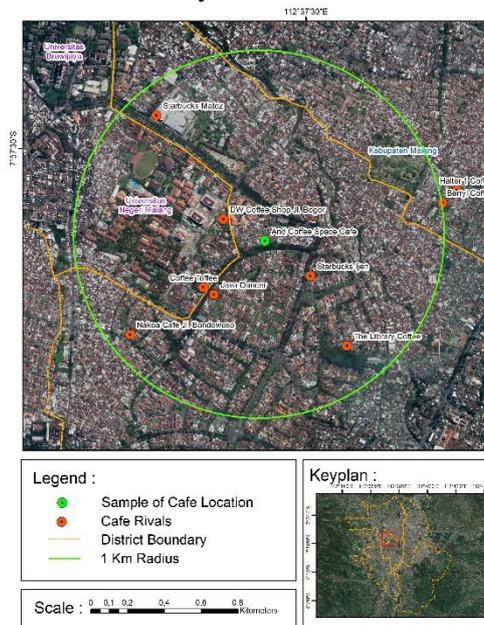


Figure 5. Rival café location map within one km radius

Identifying the competitor's number in each cafe sample found that most of the cafe samples were in category 1. It had many competitors or was said to be close to each other in a one km buffer. It raises the assumption that most café samples compete with each other by location.

3.7. Regression

Multiple Linear Regression Analysis aims to calculate the value of the influence of each independent variable on the dependent variable. However, it is necessary to test the classical assumptions on the compilation of research variable data before that. The following are the results of the classical assumption test carried out on the compilation of data in this study.

a. Autocorrelation

Table 3. Research variable

Model	R	Std. Error of the Estimate	Durbin-Watson
1	0.890 ^a	0.28577263	2.055

Based on the table 3 model summary, it can be seen that the Durbin Watson value is 2.055. This figure is compared with the Durbin Watson table with the limit values of dL and dU on the number of samples $n = 33$ and the number of variables or $n = 5$. The result is avoiding the problem of autocorrelation.

b. *Multicollinearity Test*

Table 4. Coefficients

<i>Model</i>	<i>Sig.</i>	<i>Collinearity Statistics</i>	
		<i>Tolerance</i>	<i>VIF</i>
1 (Constant)	0.000		
X1	0.002	0.376	6.002
X2	0.000	0.307	2.205
X3	0.045	0.147	6.878
X4	0.000	0.139	8.062
X5	0.001	0.454	8.213

Table 4 shows that the value of VIF (variance inflation factors) on all predictor variables is less than 10. The independent variables used avoid multicollinearity problems.

c. *Normality Test*

A normality test determines whether the data used has been normally distributed. Normally, the significance value is less than 5% or 0.05.

Table 5. Kolmogorov Smirnov

		<i>Unstandardized Residual</i>
N		33
Normal Parameters ^{a,b}	Mean	0.0809
	Std. Deviation	0.30216
Kolmogorov-Smirnov Z		1.349
Asymp. Sig. (2-tailed)		0.054

Based on the Kolmogorov Smirnov test on research data using SPSS [10], it can be seen that the research data is normal distributed. It is indicated by the sig value exceeding 0.05 or 5%, namely 0.054. A sig value of more than 0.05 means that it is not significant or means that the data is relatively the same as the average, so it is normal.

d. *Heteroscedasticity Test*

Table 6. Glesjer

<i>Model</i>	<i>Standardized Coefficients</i>		<i>t</i>	<i>Sig.</i>
	<i>Beta</i>			
1 (Constant)			3.012	0.003
X1	-0.053		-0.584	0.560
X2	0.036		0.294	0.769
X3	-0.010		-0.055	0.956
X4	-0.127		-0.696	0.487
X5	0.125		1.247	0.214

Based on table 6, it can be concluded that the research data does not experience heteroscedasticity problems. It is because the sig value of each independent variable exceeds 0.05, which means it is not significant. So that all variables do not experience symptoms of heteroscedasticity.

e. Model

The data used does not experience problems in the classical assumption test. The data is ready to be carried out with multiple linear regression analysis. The model results from the multiple linear regression analysis processes are as follows.

$$Y = 160 + (0,276) X1 + (0,476) X2 + (0,281) X3 + (0,003) X4 + (0,100) X5$$

Information :

Y = Number of Visitors

X1 = Distance of cafe to university

X2 = Building Density

X3 = Path Hierarchy

X4 = Rating

X5 = Number of Rivals

To determine whether the model is feasible to predict the number of cafe visitors as a win in market share, the model has carried out a feasibility test with three stages: model reliability test, regression coefficient test, and coefficient of determination test. The following are the results of the model's feasibility test

a. Model Reliability Test

Table 7. Anova

	<i>Model</i>	<i>Sum of Squares</i>	<i>df</i>	<i>Mean Square</i>	<i>F</i>	<i>Sig.</i>
1	Regression	50.483	7	7.212	82.837	0.000 ^a
	Residual	17.848	205	0.087		
	Total	68.331	212			

Based on table 7, it can be seen that the sig value is 0.000, which means it is smaller than the significance level value of 5% or 0.05. So it can be concluded that the regression model is feasible to explain how much influence all independent variables have on the number of visitors.

b. Regression Coefficient Test

Based on table 4, it can be seen that in the sig column, all independent variables have a probability value of less than a significance value of 5% or 0.05. It means that all independent variables have a significant influence with a 95% confidence level on the dependent variable, namely the number of visitors. Hence, all variables free deserves to be made accessible.

c. Coefficient of Determination Test

Table 8. Model Summary

<i>Model</i>	<i>R</i>	<i>R Square</i>	<i>Adjusted R Square</i>
1	0.860 ^a	0.739	0.730

Based on table 8, it can be concluded that the proportion of the influence of all independent variables is 0.739 or 73.9% on the dependent variable, namely the number of visitors. It shows in the model summary table in the R Square column.

4. Conclusions

The most influential variable in the number of visitors is the density of buildings, with a coefficient value of 0.476. Building density is indicated as the intensity of population activity. The denser a location is, the higher the intensity of its activities. It will also have the potential to bring in café visitors.

While the variable that has the least effectiveness is the café rating, with a coefficient value of 0.003. The Café rating as a value for the quality of products and services is only known when potential visitors look at the google map and not as a factor related to location.

The second most influential variable is the road hierarchy, with a coefficient value of 0.281. The road hierarchy indicates the ease of accessibility to the cafe location. The higher the class or street hierarchy, the higher the potential for the café to be visited.

So for a cafe to win market share in an area, the spatial aspects that need to be prioritized are the density of buildings and the road hierarchy. Because these two variables have the highest positive influence value in this study. It should be emphasized that the variables used in this study are spatial aspects. At the same time, the winning of market share is not determined from the spatial aspect but by many factors such as marketing strategy, product and service quality, and others.

References

- [1] A. W. Hasyim, A. B. Pandiangan, and W. Sasongko, "Perbedaan Harga Pasar Lahan Dan Ketetapan Harga Lahan Oleh Pemerintah (NJOP) Di Kecamatan Sidoarjo," *Urban Reg. Planning. Brawijaya Univ.*, 2020.
- [2] Pramadani, "Pengaruh Materialisme, Prestise dan Kelompok Referensi Terhadap Keputusan Pembelian Konsumen Pada Kafe Mewah Di Kota Malang," *J. Ekon. dan Bisnis Univ. Brawijaya*, 2014.
- [3] A. W. Hasyim and H. I. . Gusti, "Faktor-faktor yang Mempengaruhi Pertumbuhan Lahan di Kabupaten Manggarai Timur," *Urban Reg. Planning. Brawijaya Univ.*, 2019.
- [4] A. Nugroho, A. Hasyim, and F. Usman, "Urban Growth Modelling of Malang City using Artificial Neural Network Based on Multi-temporal Remote Sensing," *Civ. Environ. Sci.*, vol. 001, no. 02, pp. 052–061, 2018, doi: 10.21776/ub.civense.2018.00102.2.
- [5] R. F. Lusch and M. D. Patrick, *Retail Management*. Ohio : South Western Publishing Co., 1990.
- [6] J. Griffin, *Customer Loyalty : Menumbuhkan dan Mempertahankan Kesetiaan Pelanggan*. Jakarta: Erlangga, 2005.
- [7] H. Hani, *Dasar-dasar Manajemen Produksi dan Operasi*. Yogyakarta : BPFE, 2000.
- [8] N. Indarti, "Business Location and Success: The Case of Internet Café Business in Indonesia," *Gadjah Mada Int. J. Bus.*, vol. 6, no. 2, p. 171, 2004, doi: 10.22146/gamaijb.5543.
- [9] M. Suyanto, *Marketing Strategy Top Brand Indonesia*. Yogyakarta: Andi Offset., 2007.
- [10] I. Ghozali, *Aplikasi Analisis Multivariate dengan Program SPSS*. Semarang: Universitas Diponegoro, 2005.