

K-MEANS IN CLUSTERING THE SATISFACTION LEVEL OF CIKUNDUL HOT WATER

DUDI H GUSTIAN¹, MUHAMAD MUSLIH², DESHINTA ARROVA DEWI³, VENTI SITI⁴,
RINA RUSTIANA⁵, PURNAMAWATI⁶

¹Lecturer, Nusa Putra University, Faculty of Engineering, Computer and Design, West Java, Indonesia and
Research Fellow, INTI International University, Nilai, Malaysia

^{2,4,5,6}Lecturer. Nusa Putra University, Faculty of Engineering, Computer and Design, West Java, Indonesia

³Senior Lecturer. INTI International University, Faculty of Data Science and Information Technology,
Negeri Sembilan, Malaysia

Corresponding author: ¹dudih@nusaputra.ac.id

E-mail: ¹dudih@nusaputra.ac.id, ²muslih@nusaputra.ac.id, ³deshinta.ad@newinti.edu.my,

⁴venti.siti.vauziah_si19@nusaputra.ac.id, ⁵rina.rustiana_si19@nusaputra.ac.id,

⁶punama_si19@nusaputra.ac.id

ABSTRACT

Sukabumi is a small administrative city in West Java with a land area of 4,800.231 Ha. This city, on the other hand, serves as the economic hub of the Sukabumi Regency, a much larger region. This city is home to a number of well-known tourist attractions that generate revenue for the region through both artificial and natural tourism. Cukundul Hot Spring, with its hot water acting as a magnet, is one of the natural tourist attractions that benefit Sukabumi City's local administration. However, because of the effects of PPKM implementation, the number of visitors to this tour has significantly decreased since the Covid 19 pandemic. As a result, the regional authority of Sukabumi City has seen a decrease in the amount of money it receives from these tourist attractions. The purpose of this paper is to examine the current situation and suggest some improvements. K-Means Clustering can provide a general overview of the parameter mapping investigated in this study based on visitor sample data from approximately 35 visitors. Although this study is not robust enough to serve as a benchmark due to the small sample size, it can serve as an example for managers to improve visitor services and for local governments to better manage budgets for maintenance and development, ensuring that this tour remains a favorite of tourists from within and outside Sukabumi.

Keywords: *West Java, Sukabumi City, K-Means Clustering, Visitor*

1. INTRODUCTION

Sukabumi City is a city in West Java Province, Indonesia with the smallest area in West Java. In the national urban system, this city has been designated as a Regional Activity Center (PKW). The position of the City of Sukabumi is located in the south-central part of West Java, located at the foot of Mount Gede and Mount Pangrango which are 584 m above sea level and 120 Km from the National Capital (Jakarta) and with an area of 4,800.231 Ha. Where for the whole area of Sukabumi City is adjacent to the territory of Sukabumi Regency with the details that in the North it is bordered by Cisaat District and Sukabumi District, Sukabumi Regency.

South side with Nyalindung District, Sukabumi Regency, West side with Cisaat District, Sukabumi Regency [1].

The city of Sukabumi has many natural tourist areas that are in great demand, especially by tourists who come from outside the city, where there are lots of tourist objects that can be visited, both those that are well managed and those that are not well organized. Various types of tourism owned by the City of Sukabumi include the natural tourist attraction of the Salabintana nature tourism park and the Halimun Cottage which is in the North of Sukabumi City, the Cikundul hot spring which is in Lembursitu Sukabumi City [2].

The city of Sukabumi has several tourist objects consisting of natural tourism, cultural tourism and man-made tourism, but the lack of information about these tourism objects makes these objects less well-known and not many people know about them.

Table 1: Tourism Object and Types in Sukabumi City [3].

No	Attractions	Types of Tourism
1	Alun-Alun digital Sukabumi City and Masque Agung Sukabumi City	Artificial
2	Bandros Atta	Artificial
3	Pasim Prana Swimming Pool	Artificial
4	Sport Garden Swimming Pool	Artificial
5	Mochi Kaswari Lampion	Artificial
6	Sukabumi Fantasi Home	Artificial
7	Santasea Water park	Nature
8	Cikundul Hot Springs	Nature
9	3D Cikundul Hot Springs	Artificial
10	Toserba Selamart Culinary	Artificial
11	Sukabumi Beauty Villa	Artificial
12	Love Padlocks of Sukabumi	Artificial
13	Museum of Prabu Siliwangi	Culture
14	Museum of Pegadaian	Culture
15	Ki Pahare Museum	Culture

Sukabumi city is one of the small cities where local revenue is very dependent on optimizing existing land. Therefore, the local government continues to strive for local revenue for development and services to the community. A real example of this is the improvement in the quality of health at the puskesmas, especially the quality of services and infrastructure. From the infrastructure side, accelerating development, such as the Ahmad Yani sidewalk, includes the people's taxes paid. According to the Mayor of Sukabumi, that in 2021 the city of Sukabumi will receive Regional Original Income (PAD) worth IDR 1.243 trillion and this according to him is an achievement that exceeds the target set, which is around 104 percent [4]. One source of local revenue for the city of Sukabumi comes from the tourism sector. where this source is one of the mainstays of the city of Sukabumi in order to build a more advanced and prosperous city. This condition can be seen in Figure 2 about the total regional income of the city of Sukabumi from the

tourism sector in 2020 – 2021. In Figure 2, it can be seen that in 2020 semester 1 received income of around 72,906,500, then for semester 2 it has increased to 121,879,071 . However, at the beginning of 2021, the city of Sukabumi's income from the tourism sector experienced a sharp decline, namely only 640,000 and rose drastically to 149,355,500. This is understandable because at the end of 2020 and early 2021, Indonesia globally is experiencing a pandemic which of course also has an impact on the tourism sector, as a result of lowering visitor levels which also has an impact on the income of tourism managers.

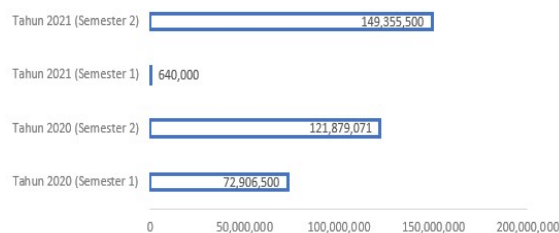


Figure 1: Total Regional Income from The Tourism Sector [5]

One of the tourist attractions that is quite sought after by the public is the Cikundul hot spring. Which is one of the tourist objects in the city of Sukabumi which is located at Jalan Proklamasi Number 242, Cikundul Village, Lembursitu District. This tourist spot has attractive natural scenery and is located next to the Cimandiri river which is also a favorite location for tourists to do rafting. At this tourist location there is also one public hot spring pool and one public cold water pool and interestingly there are also several private hot spring bath rooms and also lodging. Basically, this tourist spot is quite interesting with the facilities offered to visitors at a cost that is relatively affordable to the public [6].



Figure 2: One of the Favorite Areas in Cikundul Hot Spring [7]

Therefore, this bathhouse is quite crowded with visitors every day, but according to the head of the Sukabumi City Disporapar Tourism Sector, he admits that the PAD target from the tourism sector has never increased significantly every year. In 2017, the revenue from this hot spring was only IDR 177 million, meaning that it only increased slightly in 2018, which was more than IDR 203 million. [8].

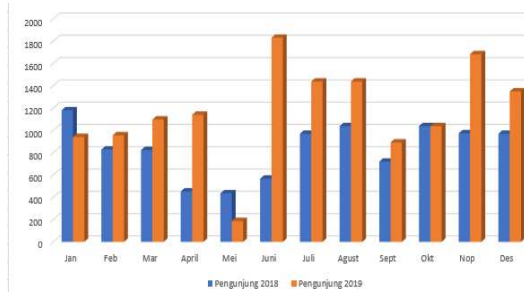


Figure 3: Number of Visitors to Cikundul Hot Spring Tourism in 2018-2019

Based on Figure 3 above, you can see the growth in the number of visitors coming to the Cikundul Sukabumi Hot Spring Tourism. In 2018 there were 9,824 people and in 2019 it increased to 13,995 people or an increase of 43%. where the highest visitors are in June and November for 2019 and January, October to December for 2018. Especially for 2019 it happened like that because it was during the school holidays and national holidays, so not only the local people came, but from outside Sukabumi [9].

However, the results of field observations revealed that in 2020, the number of visitors to these hot springs had decreased before the Covid 19 pandemic. This condition is shown in Figure 4 below.

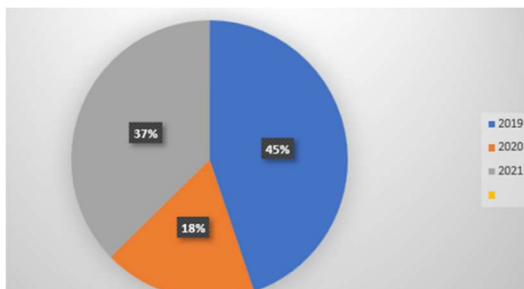


Figure 4: Data on Number of Visitors for the Last 3 Years

In Figure 4 above, it appears that the number of visitors in 2019 was recorded at around 3000 visitors, in 2020 it has decreased to around 1250 and in 2021 it has increased to 2500 visitors. This is because in 2019 there was still no Covid pandemic, while for 2020 the Covid pandemic has hit Indonesia, especially in Sukabumi, resulting in a decrease in visitor levels due to the pppm policy but in 2021 it has increased due to the relaxation of the pppm policy because the covid pandemic has started to fall and is under control. However, what the local government needs to pay attention to, especially hot spring managers, is the level of visitor satisfaction which will have a direct impact on the number of visitors. Please note that managing tourism objects is not an easy job because it requires hard work, cooperation, and commitment between local government, managers, and the community as visitors. Even though it is good, good and complete with tourist attractions, if it is not managed properly, it will be damaged and not properly maintained because it is a regional asset that is a source of income.

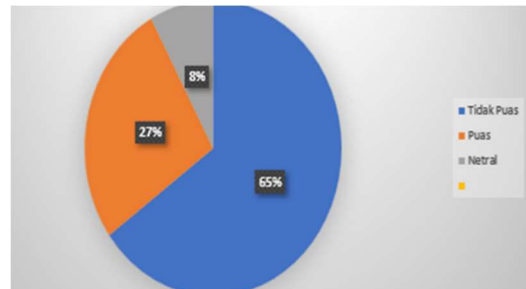


Figure 5: Graph of Visitor Satisfaction Level at Cikundul Hot Spring

Figure 5 above shows a graph of the level of visitor satisfaction at Cikundul hot springs in 2022. Where this data was taken through the observations of researchers who directly took samples from the total number of visitors who came on a certain day with a total sample taken directly of around 60 visitors. From the graph, there are problems even though this tourist spot is quite popular, in terms of the comfort felt by visitors, it is still quite low. From the sample data taken, around 65% of the visitors were dissatisfied with the facilities provided, especially the maintenance of bathing assets which belong to the local government. This is of course a record for the management to improve service for the better for visitors.

2. LITERATURE REVIEW

This study uses the K-Means method because according to previous research, this method can obtain show that the frequency of tourist vacations [10], would be further analyzed to determine suitable media promotion based on the marketing mix method [11], the results of measuring parameters Execution Time on the algorithm K-Means, which is 0.00199 seconds and the process is much faster [12].

This study provides a solution in mapping the satisfaction level of visitors to the Cikundul hot spring, through mapping and distributing the satisfaction levels presented. Where it will be seen which indicators have the highest, moderate, and less influence so that the best service model is obtained in order to find out which parameters are weak in the service so that they can be improved in order to achieve service that satisfies visitors.

This study uses the K-Means method because according to previous research, this method can obtain show that the frequency of tourist vacations [10], would be further analyzed to determine suitable media promotion based on the marketing mix method [11], the results of measuring parameters Execution Time on the algorithm K-Means, which is 0.00199 seconds, and the process is much faster [12].

This study provides a solution in mapping the satisfaction level of visitors to the Cikundul hot spring, through mapping and distributing the satisfaction levels presented. Where it will be seen which indicators have the highest, moderate, and less influence, so that the best service model is obtained in order to find out which parameters are weak in the service so that they can be improved in order to achieve service that satisfies visitors.

The benefits of this research for the management of the Cikundul hot spring are being able to provide significant input in order to improve quality and service for the better. This of course allows the local government to be more serious and pay attention to assets that have been the regional income, through allocating budgets to carry out routine maintenance both in terms of facilities and welfare for employees at these tourist sites. So that income in terms of tourism can continue to be stable and even good according to the desired target.

Eben Haezer Wisanta (2021). Conducted research with the title "K-Means Algorithm Analysis for Service Satisfaction Clustering: Pekanbaru Public Service Mall. Researchers used research sample data from satisfaction data obtained by distributing questionnaires to the ends of the Public Service Mall and 206 questionnaires were collected. The variable studied is the level of satisfaction of each public service with a total of 68 variables. This study uses k-means as a clustering algorithm which results in cluster C1 having the lowest value, namely $C1 = 2.052381$, where in cluster C1 there are as many as 22 public services that require attention in serving the community. used as a reference by the head of the Pekanbaru Public Service Mall as a reference to determine community satisfaction with services and can improve services according to the needs of the people of Pekanbaru city [13].

Teuku Muhammad Dista (2022). Conducted research with the title "Clustering Mall Visitors Using the K-Means and Particle Swarm Optimization Methods". Researchers used research sample data from visitors who visited the Cikundul hot springs. This study uses the K-Means and Particle Swarm Optimization methods which produce the results of the K-Means method which are proven by the Davies Bouldin Index method. This study has concluded that customers who have high-income levels and high expenditure scores are targets with the highest priority level for malls [14].

Sinta Maria Sinaga, Jaya Tata Hardinata, M. Fauzan (2021). Conducting research with the title "Implementation of Data Mining Clustering Level of Consumer Satisfaction with Go-Jek Services, level of public satisfaction with Gojek services". This study uses research sample data from customers. The variables examined here are ratings in the application, rating customer comments. This research uses the Data Mining Clustering method which is based on the results of a study of 120 Gojek consumers in Pematangsiantar City who are satisfied as many as 43 consumers and 77 consumers are dissatisfied [15].

Briliananda Widhi Nugraha (2021). Conducted research with the title "Application of the K-Means Method for Grouping Graduate User Satisfaction Levels in the ITN Malang Career Center Tracer Study". This study uses student sample data. The method used is the K-Means method. This research has concluded that the results of this

research are products in the form of website applications, products have features, namely the system on the website can provide alumni quality analysis based on graduate user tracer study data and to determine competency data, based on the process of comparing manual calculations with programs found levels accuracy of 57.5%, based on the results of user testing, the percentage of respondents obtained is 60% Very Good and 40% Good [16].

Mohammad Arsyad Fathurrohman (2021). Conducted this research with the title "Determining Promotion Media based on Customer Characteristics using Clustering Analysis and Marketing Mix Strategy". This study uses customer sample data. The method used is the K-means Algorithm method. The research has concluded that those results would be further analyzed to determine suitable promotion media based on the marketing mix method. The results of the study are expected to help businesses achieve customer satisfaction [17].

Widya Juli Mawaddah, Indra Gunawan, Ika Purnama Sari (2022). Conducted this research with the title "Implementation of Data Mining Algorithm for Clustering of Palm Oil Harvested Data". This study took sample data at the Marihat Unit PPKS office. The data used is data on oil palm yields in 2020 as many as 100 data. This research resulted in an analysis of yields with very good scores of 66 items, harvest data with good grades of 32 items, and harvest data with fair scores of 2 items, based on Net and Gross for each region. Based on this, it can be concluded that the K-Means Algorithm can be used to cluster oil palm yields at PPKS Marihat [18].

Lily Wulandari (2022). Conducted this research with the title "Algorithm Analysis of KMeans and Fuzzy C-Means for Clustering Countries Based on Economy and Health". This research uses K-Means and Fuzzy C-Means algorithms. This research has concluded that The K-Means Algorithm process produces 32 developed countries and 135 developing countries, while the Fuzzy Algorithm process C-Means produces 33 developed countries and 134 developing countries. Results analysis of performance testing with the Davies Bouldin Index parameter on the K-Means algorithm gets the smallest (better) value, namely equal to 0.6606398 DB. Meanwhile, the test results with Silhouette parameters The coefficients on Fuzzy C-Means are the greater the discovery (the better) and get a value of 0.896 S. On the other hand, sufficient testing significant in this research is the result of measuring the execution parameter Time on

the K-Meansi algorithm, which is 0.00199 seconds and the process much faster [19].

2.1 Framework Thinking

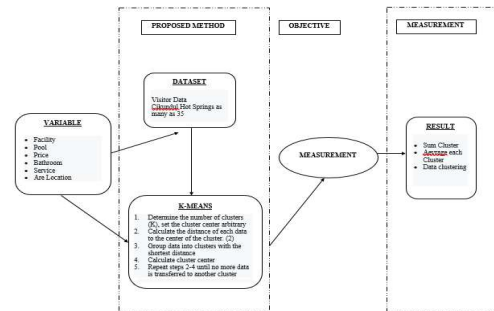


Figure 6: Framework Thinking in This Research.

3. METHODOLOGY

3.1 Framework Thinking

3.1.1 Research Tools

The research tool is a computer with general desktop computer specifications. This study uses research tools in the form of hardware and software, namely:

1. Hardware
 - Processor
 - Monitor with resolution 14 inch HD 16:9 (1366 x 768)
 - Large RAM memory 2GB
2. Software
 - Windows operating system
 - Microsoft Excel
 - Rapidminer 5

3.1.2 Research Tools

Research materials that will be used by the authors include the results of surveys and observations that have been carried out. The research materials include:

1. Data Source

The source of data in this study is visitors. Visitors are people who come to visit a place or country which consists of many people with different purposes different. In this stage the researcher collected data on the Analysis of Visitor Satisfaction Levels at Cikundul Hot Springs, hot

spring data were taken from 2019-2022 with a total data of 35 respondents.

3.1.3 Data Collection Methods

1. Observation

In this case direct observation of the research object to be examined is the Cikundul hot spring tour. And analyze the problems that exist in the Cikundul hot spring tourist attraction. Observations were carried out in May 2022 at the Cikundul hot spring, the results of observations of the number of visitors in 2019 were recorded at around 3000 visitors, in 2020 it decreased to around 1250 and in 2021 it rose to 2500 visitors. This is because in 2019 there was still no Covid pandemic, while for 2020 the Covid pandemic has hit Indonesia, especially in Sukabumi, resulting in a decrease in visitor levels due to the PPKM policy but in 2021 it has increased due to the relaxation of the PPKM policy.

a. Interview

This interview was conducted by asking a number of questions to the tour administrators of Cikundul Hot Springs. The questions asked could be in the form of oral or written form in the form of a questionnaire.

b. Questionnaire

In the data collection process, the researcher submitted a Likert scale questionnaire to visitors to the Cikundul hot spring tourism object. The research was used to measure visitors' opinions of the services and facilities of the Cikundul hot spring tourist attraction. This scale is used to complete a questionnaire that requires respondents to indicate their level of agreement with a series of questions. Usually, the questions used for research are called research variables and are specified specifically. To collect data, the results obtained were 35 visitors who filled out the questionnaire.

c. Library Studies

In this method what is done is by looking for materials that support the definition of problems through books, the internet, which are related to the object of research. This is done to strengthen research.

3.2 Methods

The goal of k-means clustering, like other clustering methods, is to get groups of data by maximizing the similarity of characteristics within

clusters and maximizing the differences between clusters. The Kmeans clustering algorithm groups data based on the distance between the data and the cluster centroid points obtained through an iterative process. The analysis needs to determine the number of K as input to the algorithm. In the realm of machine learning, the k-means clustering algorithm is included in the type of unsupervised learning.

In general, the K-Means Cluster Analysis method uses the following algorithm [20].

1. Determine k as the number of clusters formed. To determine the number of clusters k is carried out with several considerations such as theoretical and conceptual considerations that might be proposed to determine how many clusters.
2. Generate k Centroid (cluster center point) randomly. The determination of the initial centroid is carried out randomly from the k clusters of available objects, then to calculate the centroid of the next I-th cluster.
3. Calculate the distance of each object to each centroid of each cluster. To calculate the distance between objects and the centroid the author uses Euclidian Distance.
4. Allocate each object to the nearest centroid. To allocate objects into each cluster during iterations, in general, it can be done in two ways, namely by hard Kmeans, where strictly speaking each object is declared as a member of the cluster by measuring its proximity to the center point of the cluster, another way can be done by fuzzy Cmeans.
5. Perform iterations, then determine the position of the new centroid using the equation.
6. Repeat step 3 if the new centroid position is not the same. Convergence checking is done by comparing the group assignment matrix in the previous iteration with the group assignment matrix in the current iteration. If the results are the same, the k-means cluster analysis algorithm has converged, but if it is different, it has not yet converged, so further literacy is needed.

To measure the distance between the data and the cluster center, Euclidian distance is used, then the distance matrix will be obtained as follows: Euclidian distance formula:

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \tag{1}$$

x = cluster center
y = data

example with M(a,b), where a is the total number of visitors, and b is reduced to the number of visitor satisfaction to make the calculation method easier.

4. RESULT AND DISCUSSION

4.1 First Literacy

Then calculate the distance from the remaining data samples to the cluster center, for

Table 2: Sample Data on Satisfaction Questionnaire Results from Passengers.

No	Age	Sex	Facility	Pool	Price	Bathroom	Service
1	16	PEREMPUAN	6.8	6.8	2	4	5.2
2	38	LAKI-LAKI	6.8	5.8	2	4	5.2
3	23	PEREMPUAN	6.8	6.8	3.6	4	5.2
4	25	PEREMPUAN	6.4	6.8	3.6	4	5.2
5	19	LAKI-LAKI	6.4	6.8	3.2	4	5.2
6	21	PEREMPUAN	6.4	6.8	2	4	5.2
7	47	LAKI-LAKI	6.4	6.8	2	4	5.2
8	30	LAKI-LAKI	6.8	6.8	2	4	5.2
9	16	PEREMPUAN	6.4	6.8	2	4	5.2
10	15	PEREMPUAN	6.4	6.8	2	4	5.2
11	28	PEREMPUAN	6.4	6.8	2	4	5.2
12	40	PEREMPUAN	6.4	6.8	2	4	5.2
13	50	PEREMPUAN	6.4	6.8	2	4	5.2
14	35	LAKI-LAKI	6.4	6.8	2	4	5.2
15	27	LAKI-LAKI	6.4	6.8	2	4	5.2
16	32	LAKI-LAKI	6.4	6.8	2	4	5.2
17	17	PEREMPUAN	6.4	6.8	2	4.4	5.2
18	17	PEREMPUAN	6.4	6.8	2	4	5.2
19	36	PEREMPUAN	6.4	6.8	2	4	4.8
20	23	PEREMPUAN	6.4	6.8	2	4	5.2
21	34	PEREMPUAN	6.4	6.8	2	4	5.2
22	36	PEREMPUAN	6.4	6	2	4	5.2
23	17	LAKI-LAKI	6.4	6.8	2	4	5.2
24	18	LAKI-LAKI	6.4	6.8	2	4	5.2
25	25	LAKI-LAKI	6.8	6.4	2	4	5.2
26	19	LAKI-LAKI	6.4	5.6	2	4	5.2
27	19	PEREMPUAN	6.4	6.8	2	4	4.8
28	21	PEREMPUAN	6.4	6.8	2	4	4.8
29	21	PEREMPUAN	6	6.8	2	4	5.2
30	38	PEREMPUAN	6	6.8	2	4	5.2
31	36	PEREMPUAN	6.4	6.8	2	4.4	5.2
32	27	PEREMPUAN	6.4	6.4	2	4.4	5.2
33	23	LAKI-LAKI	6.4	6	2	4	5.2
34	33	LAKI-LAKI	6.8	6.4	2	4	5.2
35	20	LAKI-LAKI	6.4	6.4	2	4	5.2

Table 2 above shows the results of processing the satisfaction questionnaire from visitors to the Cikundul hot spring, where around 35 visitors were taken with a composition of 14 men and 21 women. The parameters measured in this study were facilities, pools, prices, bathrooms, and services. The value of each parameter is measured by means of an average and then used as an analysis

using the K-Means method. In table 1 above it also appears that facilities, pools, and services have a fairly high average value for dissatisfaction, this is in line with the observations of the research team who directly saw conditions in the field. Of course, this is an evaluation material for the manager to improve service to consumers.

In table 3 below, each parameter is analyzed for the clustering process, where for parameters belonging to the high category, C1 is enough with 17, medium 11 and low 7. This means that for this first literacy there are quite a lot of parameters in the high category, although this has not yet become a the main benchmark in the

clustering process because the centroid value is still changing and is the initial centroid and is shown in table 3 below.

Table 3: Distribution of Clusterization of Each Parameter in The First Iteration.

Facility	Pool	Price	Bathroom	Service	Local Area	Cluster
6.8	6.8	2	4	5.2	4.4	3
6.8	5.8	2	4	5.2	4.4	1
6.8	6.8	3.6	4	5.2	4.4	1&3
6.4	6.8	3.6	4	5.2	4.4	2
6.4	6.8	3.2	4	5.2	4.4	2&3
6.4	6.8	2	4	5.2	4.4	1
6.4	6.8	2	4	5.2	4.4	3
6.8	6.8	2	4	5.2	4.4	2
6.4	6.8	2	4	5.2	4.4	3
6.4	6.8	2	4	5.2	4.4	1
6.4	6.8	2	4	5.2	4.4	1,2,3
6.4	6.8	2	4	5.2	4.4	1
6.4	6.8	2	4	5.2	4.4	3
6.4	6.8	2	4	5.2	4.4	2
6.4	6.8	2	4	5.2	4.4	1&3
6.4	6.8	2	4	5.2	4.4	2
6.4	6.8	2	4.4	5.2	4.4	2&3
6.4	6.8	2	4	5.2	4.4	1
6.4	6.8	2	4	4.8	4.4	1&3
6.4	6.8	2	4	5.2	4.4	2
6.4	6.8	2	4	5.2	4.4	3
6.4	6	2	4	5.2	4.4	1&2
6.4	6.8	2	4	5.2	4.4	1,2,3
6.4	6.8	2	4	5.2	4.4	1,2,3
6.8	6.4	2	4	5.2	4.4	3
6.4	5.6	2	4	5.2	4.4	1
6.4	6.8	2	4	4.8	4.4	1&2
6.4	6.8	2	4	4.8	4.4	2
6	6.8	2	4	5.2	4.4	2&3
6	6.8	2	4	5.2	4	1
6.4	6.8	2	4.4	5.2	4.4	1&3
6.4	6.4	2	4.4	5.2	4.8	2
6.4	6	2	4	5.2	4.4	3
6.8	6.4	2	4	5.2	4.4	2
6.4	6.4	2	4	5.2	4.4	1,2,3

Table 4: Initial Centroid Value.

No	Facility	Pool	Price	Bathroom	Service	Local Area	Cluster
C1	6.8	6.8	3.6	4	5.2	4.4	High
C2	6.4	6.4	2	4.4	5.2	4.8	Medium
C3	6.4	6.8	2	4	5.2	4.4	Low
C1	6.8	6.8	3.6	4	5.2	4.4	High

4.2 Third Literacy

After the literacy process is carried out, the clusterization process stops at literacy three. This is shown by the centroid value that does not change or is the same as the previous centroid value. For parameters included in the high group category, there are 21 parameter data, moderate 11 parameter data, and 3 low parameter data. From this it can be

taken from the analysis that even though the service is lacking in terms of parameters, especially facilities, pools, and services, in general visitors feel satisfied with relatively low costs, and the location of the area is not too far away because it is in Sukabumi City and the most unique thing is the bathroom which is indeed widely available in this hot spring. This is shown in Table 4 below.

Table 5: The Distribution of Clusterization for Each Parameter in the Third Iteration.

Facility	Pool	Price	Bathroom	Service	Local Area	Group
6.8	6.8	2	4	5.2	4.4	3
6.8	5.8	2	4	5.2	4.4	3
6.8	6.8	3.6	4	5.2	4.4	2&3
6.4	6.8	3.6	4	5.2	4.4	1,2,3
6.4	6.8	3.2	4	5.2	4.4	1&2
6.4	6.8	2	4	5.2	4.4	3
6.4	6.8	2	4	5.2	4.4	3
6.8	6.8	2	4	5.2	4.4	2&3
6.4	6.8	2	4	5.2	4.4	2
6.4	6.8	2	4	5.2	4.4	1,2,3
6.4	6.8	2	4	5.2	4.4	1&3
6.4	6.8	2	4	5.2	4.4	1&3
6.4	6.8	2	4	5.2	4.4	2
6.4	6.8	2	4	5.2	4.4	2&3
6.4	6.8	2	4	5.2	4.4	2&3
6.4	6.8	2	4	5.2	4.4	1&3
6.4	6.8	2	4.4	5.2	4.4	1
6.4	6.8	2	4	5.2	4.4	1,2,3
6.4	6.8	2	4	4.8	4.4	2&3
6.4	6.8	2	4	5.2	4.4	2&3
6.4	6.8	2	4	5.2	4.4	1&3
6.4	6	2	4	5.2	4.4	1&3
6.4	6.8	2	4	5.2	4.4	1,2,3
6.4	6.8	2	4	5.2	4.4	1,2,3
6.8	6.4	2	4	5.2	4.4	2
6.4	5.6	2	4	5.2	4.4	3
6.4	6.8	2	4	4.8	4.4	3
6.4	6.8	2	4	4.8	4.4	1,2,3
6	6.8	2	4	5.2	4.4	1&2
6	6.8	2	4	5.2	4	1,2,3
6.4	6.8	2	4.4	5.2	4.4	3
6.4	6.4	2	4.4	5.2	4.8	3
6.4	6	2	4	5.2	4.4	2
6.8	6.4	2	4	5.2	4.4	1,2,3
6.4	6.4	2	4	5.2	4.4	1,2,3

After analyzing with the K-Means method, the centroid value is obtained which does not change from the previous centroid value. Which produces the third literacy with the centroid value in Figure 8 above. For the average value of each cluster obtained for C1 for location and service around 31.60, C2 around 6.11 and C3 around 6.04. This shows that this hot spring is still one of the main destinations for the

community even though there are many other baths, especially in Sukabumi City because they are still in the high or good category in the eyes of the community.

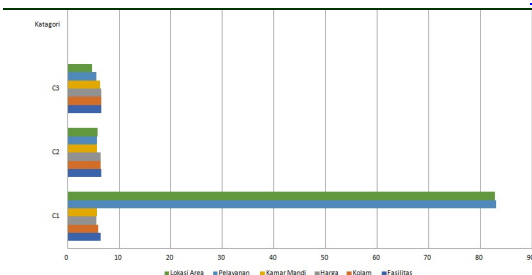


Figure 6: Framework Thinking in This Research.

5. CONCLUSION

From the results of the analysis with K-Means clusterization, a mapping picture of the level of service satisfaction provided by visitors to the Cikundul hot spring is obtained. Where even the level of satisfaction is the weak point of this tourist spot, especially in the parameters of the bathroom, price, pool, and facilities. But the parameter values that are included in C1 are the most about 21 data from a total of 35 data taken, even though C2 and C3 have fairly even centroid values. This illustrates that the shortcomings of these four things which are the shortcomings of this tour are still covered by the location which is quite close and the service which is quite friendly from the management. But with so many parameters that reduce the positive assessment earlier, Of course, it makes the management, especially the Sukabumi city government, pay attention and be serious in managing this tourist spot because it is an asset and income for the local government. So that if it is not managed properly and seriously, it is feared that this tourist attraction will be abandoned by visitors because of the proliferation of quite good bathing attractions in the city of Sukabumi, especially those managed by private parties with complete facilities, and of course, the income from these baths will decrease.

REFERENCES:

- [1] <http://perkotaan.bpiw.pu.go.id/n/kota-otonom/23>
- [2] R. A. Devison, L. Tyesta, And R. Herawati, "Duties And Functions Of The Sukabumi City Youth, Sports, Tourism And Creative Economy Office In The Management Of Tourist Attractions", *Diponegoro Law Journal*, Vol. 6, No. 1, 2017.
- [3] Sukabumi City Youth, Sports, and Tourism Office, Sukabumi City BAPPEDA, Filed Survey, 2019.
- [4] <https://www.republika.co.id/berita/r588rt380/pendapatan-asli-daerah-pemkot-sukabumi-2021-lampai-target#:~:text=REPUBLIKA.CO.ID%2C%20SUKABUMI,1%2C243%20triliun%20pada%20ahun%202021>
- [5] <https://opendata.sukabumikota.go.id/ui/datasheet/Jumlah%20Pendapatan%20Daerah%20dari%200Pariwisata%202020-2021/60>
- [6] <https://diporapar.sukabumikota.go.id/site/detail?id=75> 12 December 2022.
- [7] <https://sukabuminow.com/genjot-pad-sektor-pariwisata-pemkot-sukabumi-akan-optimalkan-destinasi-yang-satu-ini/> 18 December 2022.
- [8] D. Hariani, and S. G. Wibawa, "The Effect of Marketing Mix on Tourist Satisfaction in Cikundul Hot Spring Tourism, Sukabumi, West Java", *Travelling: Journal of Travel Tourism, Destinations, and Hospitality*, Vol. 3, No. 1, 2020, pp. 61-76.
- [9] M. E. Yildirim, M. Kaya, and I. FurkanInc, "Case Study: Unsupervised Approach for Tourist Profile Analysis by K-means Clustering in Turkey", *Journal of Internet Computing and Services (JICS)*, Vol. 23, No. 1, 2022, pp. 11-17.
- [10] E. H. Wisanta, and Y. N. Marlim, "Analysis of the K-means Algorithm for Service Satisfaction Clustering: Pekanbaru Public Service Mall", *National Seminar on Informatics (SENATIKA), Proceedings of SENATIKA*, 2021.
- [11] T. M. Dista, and F. F. Abdulloh, "Clustering Mall Visitors Using the K-Means and Particle Swarm Optimization Methods", *Jurnal Media Informatika Budidarma*, Vol. 6, No. 3, 2022, pp. 1339-1348.
- [12] S. M. Sinaga, J. T. Hardinata, and M. Fauzan, "Implementation of data Mining Clustering Level of Consumer Satisfaction with Go-Jek Services", *KESATRIA: Journal of Information System Implementation (Computer & Management)*, Vol. 2, No. 2, 2021, pp. 118-124.
- [13] B. W. Nugraha, A. Mahmudi, and F. S. Wahyuni, "Application of the K-Means Method for Grouping the Level of Graduate User Satisfaction in the Tracker Study of the ITN Malang Career Center", *JATI (Journal of*

- Informatics Engineering Students*), Vol. 5, No. 2, 2021.
- [14] M. A. Fathurrohman, J. Afifah, R. Fariza, and A. U. Khasanah, "Determining Promotion Media Based on Customer Characteristics Using Clustering Analysis and Marketing Mix Strategy", *Proceedings of the Second Asia Pacific International Conference on Industrial Engineering and Operations Management, Surakarta, Indonesia, September 14-16, 2021*.
- [15] W. J. Mawaddah, I. Gunawan, and I. P. Sari, "Implementation of Data Mining Algorithms for Clustering Oil Palm Crop Data", *JOMLAI: Journal of Machine Learning and Artificial Intelligence*, Vol. 1, No. 1, 2022, pp. 43-54.
- [16] L. Wulandari, and B. O. Yogantara, "Algorithm Analysis of K-Means and Fuzzy C-means for Clustering Countries based on Economy and Health", *Exatra Factor*, Vol. 15, No. 2, 2022, pp. 109-116.
- [17] R. Muliono, and Z. Sembiring, "Data Mining Clustering Uses the K-Means Algorithm for Clustering the Tridarma Level of Lecturer Teaching", *CESS (Journal of Computer Engineering System and Science)*, Vol. 4, No. 2, 2019.