Vol. 07, No. 01, January 2023, pp. 19-30

INTERNATIONAL JOURNAL OF INNOVATION IN ENTERPRISE SYSTEM

Journal homepage: https://ijies.sie.telkomuniversity.ac.id
e-ISSN: 2580-3050, DOI: https://doi.org/10.25124/ijies.v7i01.204

Decision-Making Process for Tourism Potential Segmentation: A Case Study Analysis

Fandi Achmad[1], Irsyad Thalib Abdillah[2*], Husni Amani[3]

[1]School of Industrial and System Engineering
Telkom University, Bandung, Indonesia
fandiac.achmad@gmail.com

[2]School of Industrial and System Engineering
Telkom University, Bandung, Indonesia
irsyadthalib@student.telkomuniversity.ac.id

[3]School of Industrial and System Engineering
Telkom University, Bandung, Indonesia
husni@telkomuniversity.ac.id

[*]irsyadthalib@student.telkomuniversity.ac.id

ARTICLE INFO

Article history:
Received 05 January 2023
Accepted 28 January 2023
Published 31 January 2023

ABSTRACT

The tourism industry is one of the critical sectors driving regional economic development, creating jobs, and preserving cultural and natural heritage. The Rembang Regency Government is trying to design a strategy to increase various tourism potentials that support the tourism industry. Stakeholders cannot optimally manage various tourism potentials in Rembang Regency because the support system facilities still need to be improved. Potential tourism segmentation based on facility support systems (SSF) can assist the government in carrying out strategies to increase the potential of the tourism industry. The main objective of this research is to assist stakeholders in the decision-making process and generate policy advice through tourist mapping results, reducing large-scale tourism industry problems to a scale that can be understood and acted upon in successful tourism industry marketing. Methods in data collection using literature study and observation with accidental sampling to determine as many as 20 tours. The K-Means method was chosen to classify tourism potential in Rembang Regency, Central Java, Indonesia. The results showed three potential tourism clusters in Rembang Regency based on Support System and Facilities (SSF) covering telecommunication, power source, transportation, waste management, location, clean water source, supporting industry, spatial, hospitality, safety, and security. The results of this study can be used in designing decision-making strategies by stakeholders to improve the tourism industry. Increasing the potential of each tourist cluster by stakeholders has its advantages and disadvantages.

Keywords:
Tourism industry;
Segmentation tourism; K-Means; Decision-Making;
Support system and facilities

1. INTRODUCTION

The tourism industry is an important sector that drives economic development and has a strategic position in increasing foreign exchange in several countries [1], [2], [3]. The tourism industry has become a significant contributor to the global economy, with international tourism receipts exceeding $1.5 trillion in 2019 [4]. The industry provides employment,

income opportunities, and cultural and economic benefits to destinations worldwide [5] because the tourism industry is a complex industry that includes various sectors such as accommodation, transportation, and entertainment. Industrial growth and development are influenced by various internal and external factors, including economic conditions, cultural values, and political stability [6], [7], [8].

Effective management of the tourism industry can bring many benefits to a country, including increased economic growth and development, job creation, and preservation of cultural and natural heritage [1], [3], [6], [9], [10]. Successful tourism potential management requires a comprehensive understanding of the various factors influencing the industry, including economic, social, cultural, and environmental considerations [11], [12]. Development and implementation of policy strategies that effectively balance the needs and interests of various stakeholders, including tourists, local communities, and governments [6], [8], [13]. This has prompted several regions in Indonesia to start looking at developing tourism areas as one of the leading sectors for increasing the regional economy [2], [7], [14], [15], one of which is Rembang Regency, Central Java, Indonesia [3], [16], [17], [18]. The position map of Rembang Regency, Central Java, Indonesia, is shown in Figure 1.

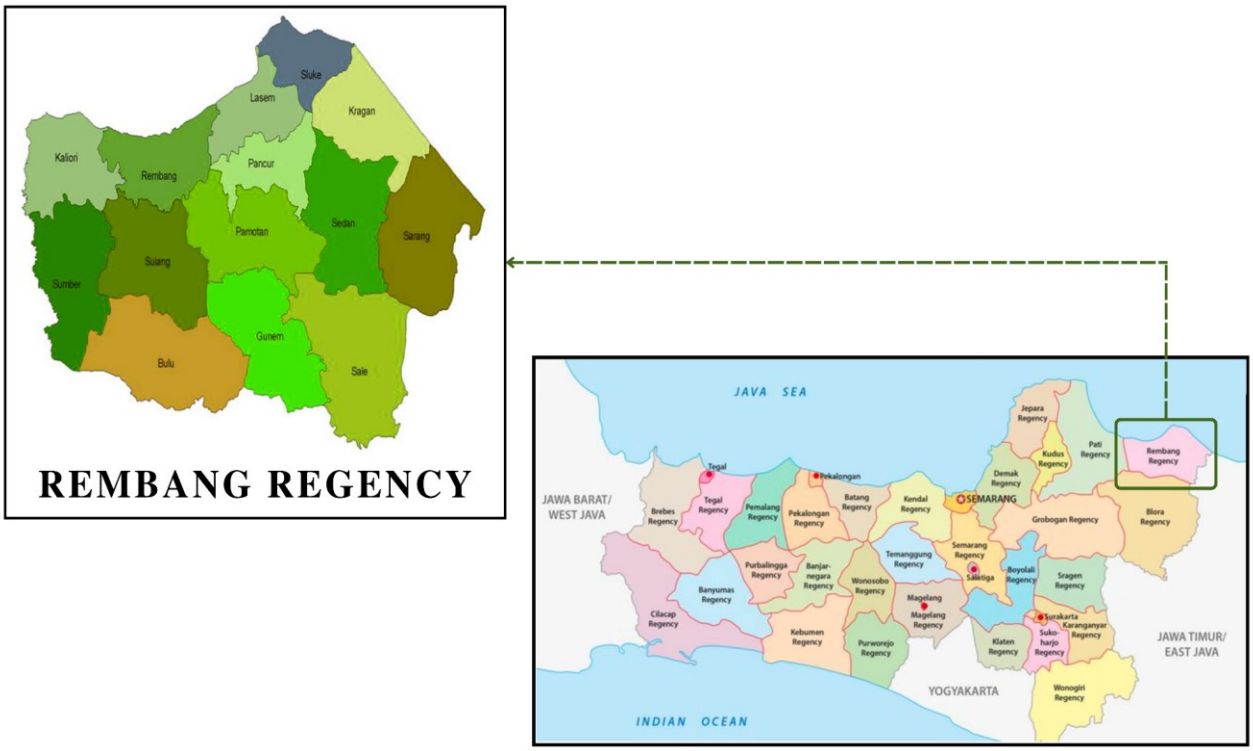


Figure 1 - Map of Rembang Regency, Central Java, Indonesia

Figure 1 is a map of Rembang Regency is one area that has a lot of tourism potential to support tourism industry activities [3], [16], [17], [19]. During the Covid-19 pandemic, a decrease in tourists occurred in Rembang Regency due to restrictions on activities by the Government [20]. The decrease in the number of tourists resulted in a decrease in regional income, which also impacted the hotel sector. Hotels as a means of accommodation have a significant role in the development of the tourism industry because they function as a place to stay for tourists who come while they are on tour. The Government of Rembang Regency, Central Java, Indonesia, is trying to increase the potential for sustainable tourism [3], [19]. This is in the Regional Regulation (PERDA) of Rembang Regency Number 12 of 2019, which states the strategies of the Rembang Regency government to improve the tourism industry's performance [20]. Given this, the tourism industry is one of the crucial sectors of economic growth in Rembang [21]. Tourism development in Rembang Regency still needs to be optimally developed, so a lot of tourism potential can be developed to become a tourist attraction to support tourism industry activities. Tourism potential in Rembang Regency, Central Java, Indonesia, is expected to be mapped and utilized to support tourism industry activities. Tourism potential in Rembang Regency, Central Java, Indonesia, is expected to be mapped based on support system facilities so that it can be used to support tourism industry activities. One of the factors for tourists visiting tourist attractions is an adequate support system for tourism activities [1], [3]. One strategy that helps with this is regarding grouping or clustering tourism potential, which aims to answer questions such as designing a strategy for increasing the tourism industry through segmentation of tourism potential to produce policy advice to stakeholders, reducing large-scale tourism industry problems to a scale that can be understood and followed up by stakeholders through mapping of tourism potential in Rembang Regency. Besides that, it can help

stakeholders in the decision-making process for successful tourism industry marketing. Various regions have successfully implemented strategies to increase the tourism industry through clustering [22]. Cluster formation can help the industry to increase competitiveness. The accumulation of various sectors, including tourism with similarities and related activities, will limit the resulting economic externalities and decrease costs incorporated in clusters. The advantages resulting from cluster formation include ease of capital, access to suppliers and specific service inputs, and the transfer of information and knowledge. [23].

The tourism industry is a concept that needs to be understood for analysis and as a decision-making material [24], [25]. The definition of tourism that is developing in the world is very diverse, multidimensional, and closely related to the scientific background of its originators [25]. These definitions can be grouped into three categories: those that see tourism from the demand side, the supply side, and those that have combined the demand and supply sides [25]. According to [20], the tourism industry is a collection of interrelated businesses that produce goods and/or services to meet the needs of tourists in organizing tourism. The tourism industry is an industry that is related to several other sectors because tourism is a combination of phenomena and interrelationships between sectors and tourists, business suppliers, governments and tourist destinations, and the community of tourist areas [26], [27].

Tourism potential is everything in tourism that is in good condition that is real and intangible which is worked on, arranged, and presented in such a way that it can be helpful or utilized, manifested as the capability of the factors and elements needed or determining tourism development, whether in the form of atmosphere, events, objects or services or services [28]. One of the main attractions of tourism is the opportunity it gives people to experience new cultures, traditions, and ways of life [28].

Segmentation can be interpreted as dividing a market into different groups. Each of these groups consists of consumers with the same or nearly identical characteristics [29]. Segmentation also has an essential role for several reasons, namely [30]:
• Segmentation allows companies to focus more on allocating their resources.
• Segmentation can provide an overview for the company to determine which segments to serve.
• Segmentation can provide a clearer picture of the competition map and determine the company's position.
• Segmentation is the basis for determining the components of the strategy

This study uses the K-means clustering method to map tourism potential based on the Support System Facilities (SSF) variable [3], where no research classifies tourism potential based on its support system facilities. The purpose of segmentation in Rembang Regency is to identify potential and suitable tourist objects to be developed according to the priorities of available facilities. Besides that, it can help stakeholders in the decision-making process to successfully market the tourism industry.

2. METHOD

2.1. Data Driven Decision Making

Decision-making is one of the main tools in any organization to define a goal, moving from a process based on experience and intuition to an increasingly sophisticated data analysis process [31], [32]. In the era of information and technology, data analysis is very important for predicting product and service business scenarios [33]. Thus, big data has become an effective decision-making tool. The digital age has made vast amounts of data available to businesses, making it a key asset for better decision-making across all fields and sectors [33], [34]. Processing these large volumes of data through the advanced analytics of the Big Data architecture helps improve and automate decision-making in the day-to-day business processes of companies [33]. For good decision-making, it is essential that apart from being right, it must be done on time and at minimum cost [33], [35]. With big data tools, crosslinking information from multiple sources is possible, which helps get accurate and diverse information for decision-making.

2.2. Support System and Facilities (SSF)

Tourism is one of the complementary industries that can provide rapid economic growth in terms of providing employment, income, and living costs and activating other production sectors in tourist-receiving countries [8], [36]. According to [3], [19], Supporting Systems and Facilities are services directly or indirectly developed to support and encourage tourism potential. Facility Support System affects the long-term success of the tourism industry because tourists' need for carrying capacity and facilities reflects tourism development in the area [8], [37]. Supporting systems and facilities in tourism are related to individual motivation to visit a tourist destination [8], [38]. Tourism demand focuses on the quantity of any product or service that tourists are willing to buy at a given price and time [38].

*irsyadthalib@student.telkomuniversity.ac.id

2.3. Data Visualization

Data Visualization is a part of statistics that discusses methods of presenting data to make it interesting and informative. In data visualization, there are several data sizes, including data center, data distribution, slope, and slope sizes [39]. Data are presented in tables, graphs, diagrams, and concentration measures such as mode and median. The main goal of data visualization is to communicate information clearly and effectively through graphics [39]. According to [40], data visualization is something to show a narrative or picture so that they can understand something more deeply.

2.4. K-Means Clustering

KK-means clustering is a method of partition clustering or prototype-based data objects that aims to group several clusters so that all members in a cluster are close to the appropriate prototype [41], [42]. K-means clustering starts by selecting k-cluster centers randomly or centroids. Each point in the dataset is assigned to the nearest centroid. Then recalculate the positions of all cluster centers until there are no points that can move from one cluster to another [41], [43], [44]. The stages in the K-means algorithm process [41] are shown in Table 1 and Figure 2.

Table 1 - Basic K-means algorithm

Basic K-Means Algorithm
1: Select *K* points as initial centroids
2: Repeat
3: Form *K* clusters by assigning each point to its closest centroid
4: Recompute the centroid of each cluster
5: until Centroids do not change.

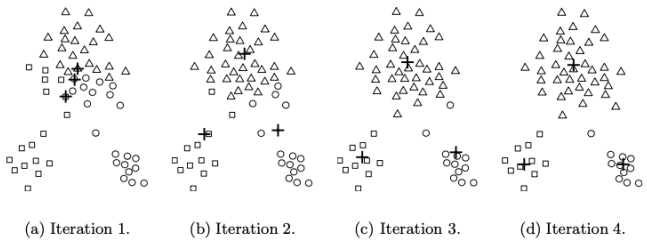


Figure 2 - Basic Concepts and Algorithm K-means

Table 1 describes the stages of using the K-means method so that the results of these stages are visualized, as shown in Figure 2. To evaluate the quality of clustering, the sum of the squared error (SSE) formula, also known as scatter, is used. Consider data whose measure of proximity is the Euclidean distance. Using notation in formula one is defined as follows [41].

$$\mathrm{SSE} = \sum_{i=1}^{K} \sum_{x \in C_i} dist(c_i, x)^2 \tag{1}$$

Dist. is the standard Euclidean distance (L2) between two objects in Euclidean space. With this assumption, the centroid that minimizes the SSE of the cluster is the mean. Using the notation in formula 1, the centroid (mean) of the i-th cluster is defined by Equation 2 [41].

$$\mathbf{c}_i = \frac{1}{m_i} \sum_{\mathbf{x} \in C_i} \mathbf{x} \tag{2}$$

This study uses a type of quantitative research. Quantitative research aims to determine and explain the clustering results using K-Means in evaluating Support System Facilities for determining segmentation of tourism potential [45], [46]. In the data collection process, the method used is to use field surveys and literature studies. The field survey method was carried out through structured observations of 20 tourist attractions in Rembang Regency as samples in April - July 2022. Observations were carried out through interviews with stakeholders, managers of tourist attractions, and the Rembang district tourism office. Data collection in this study used questionnaires distributed to SMEs around tourist areas in Rembang Regency, Central Java, Indonesia. SMEs were chosen because an SME does not move places and has knowledge of the development of tourism around it.

The sampling carried out in this study used the accidental sampling method. Accidental sampling in this study facilitates the research process because there are no particular or complicated criteria in sampling. The researcher also deliberately selects samples from a population where the required information can be obtained easily so that the majority of samples

are selected because they are in the location required by the researchers [47]. Accidental sampling in this study also provides easy access, so it becomes the primary consideration for researchers in sampling [47], [48]. Support Systems and Facilities (SSF) are services developed either directly or indirectly to support and encourage tourism potential [1], [3], [19]. In previous research, the SSF variable will affect the success of the tourism industry in the long run because tourists' needs for carrying capacity and facilities reflect tourism development in an area [1], [3], [19], [49], [50]. The research variables used in this study can be explained in Table 2.

Table 2 - Variable Support System and Facilities

Notation	Variable	Scale Data	Reference
TL	Telecommunication	Average Internet Speed	[1], [3], [19], [49], [50]
PS	Power Source	Total power source	
TP	Transportation	Distance of tourist attractions from city center	
WM	Waste Management	Number of trash cans	
LT	Location	Area of tourism	
CW	Clean Water Source	Number of clean water sources at tourist sites	
SI	Supporting Industry	Number of business areas in tourist area	
ST	Spatial	Number of supporting tours around tourist attractions	
HT	Hospitality	The number of public facilities in the tourist area	
SS	Safety and Security	Total availability of security posts and medical equipment	

Table 2 describes the variables used in segmenting tourism potential in this study. The evaluation process for each tourism potential in Rembang Regency, Central Java, Indonesia, uses questions on variables in Table 2. Segmentation of tourism potential in Rembang Regency using the K-means clustering method, where this method can group prototype-based data objects that aim to group several clusters so that all members in a cluster are close to the appropriate prototype [41], [42]. The steps in this study are shown in Figure 3.

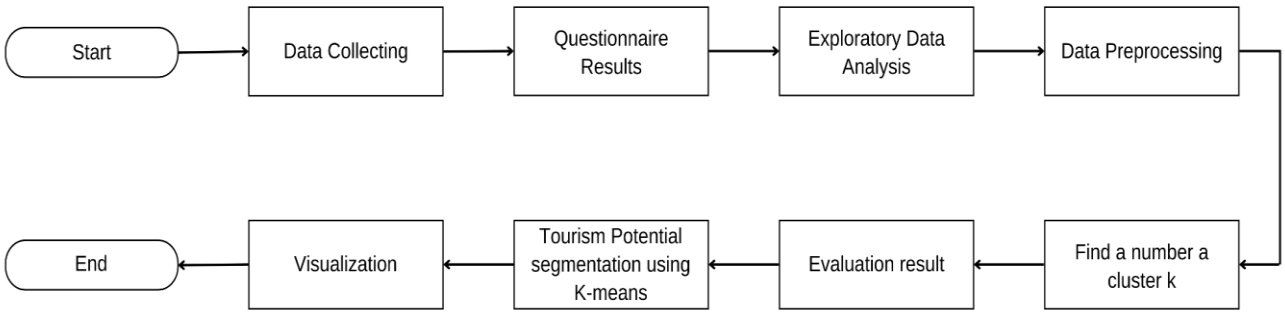


Figure 3 - Research Stages

Figure 3 explains the flow of stages in this study, starting from data collecting, questionnaire results, exploratory data analysis, data preprocessing, data processing with finding a number a cluster k, evaluation results, tourism potential segmentation using k-means, and visualization.

3. RESULT AND DISCUSSION

Rembang Regency is the easternmost Regency in Central Java Province and is located on the north coast of Central Java. In the north, the Rembang area is in direct contact with the Rembang bay (Java Sea), which makes the Rembang district have various tourism potentials that support tourism industry activities [3], [16], [17]. Segmenting tourism potential in Rembang Regency, Central Java, Indonesia, can help stakeholders decide on designing strategies to increase the tourism industry in Rembang Regency, Central Java, Indonesia. Because of an increase in the tourism industry can increase the regional economy [51]. The research was carried out by dividing tourist attractions in Rembang Regency into several clusters so that there will be various recommendations that are differentiated based on clusters of tourist attractions as a consideration in making decisions for the Rembang Regency tourism office to increase the potential that exists in each of these tourist attractions. The following are the results of each of the stages in this study.

3.1. Data Collecting

In this study, data collection was carried out using observation and a literature study. The field survey method was carried out through structured observations of 20 tourist attractions in Rembang Regency as samples in April - July 2022, shown

in Table 3. The selection of tourist objects in Rembang Regency, Central Java, Indonesia, used accidental sampling. Observations were made through interviews with several stakeholders, IKM around tourist areas, managers of tourist attractions, and the Rembang district tourism office. The tourist objects used in this study are shown in Table 3.

Table 3 - Research Tourism Object in Rembang Regency

No	Tourism	Type of Tourism
1	Nasi Gandul Soto Sapi	Culinary
2	Warung Nasi Gandul Gambiran Pak Wandi	
3	Warung Makan Bu Tin	
4	Nasi Tahu Serepeh Bu Slamet	
5	Rumah Makan Cemoro Sewu	
6	Ayam Geprek Sambel Korek	
7	Caruban	Natural
8	Air Terjun Pasucen	
9	Watu Layar	
10	Pantai Nyamplung Indah	
11	Wisata Gunung Kajar	
12	Air Terjun Kedung Grenjeng	
13	Gunung Argopuro Lasem	
14	Pulau Gede Rembang	
15	Alun-Alun Rembang	Artificial
16	Jembatan Merah Hutan Mangrove	
17	Waduk Panohan	
18	Situs Perahu Kuno Punjulharjo	Culture
19	Vihara Ratanavana Arama	
20	Petilasan Sunan Bonang	

Table 3 describes the tourist objects in Rembang Regency, Central Java, Indonesia, used in this study's collection. These tourism objects represent several tourism potentials, such as culinary tourism, natural tourism, artificial tourism, and cultural tourism.

3.2. Exploratory Data Analysis

Data processing in this study used python. The results of the questionnaire data obtained were carried out first by exploring the existing data by entering data to display the results of the existing data recap. Data exploration is carried out to analyze and investigate the data set, summarizing the main characteristics of the data to the visualization method. This can help find data patterns, spot anomalies, hypothesis testing, and assumptions [41], [52]. With the data exploration carried out, an overview of the data that has been obtained can be seen.

3.3. Data Preprocessing

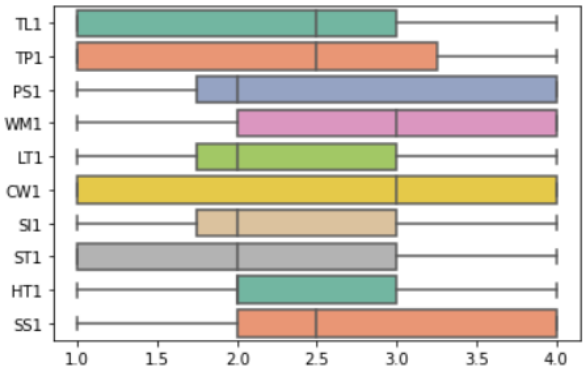
Next is to do data preprocessing by doing several steps. Data preprocessing is converting raw data into a form that is easier to understand. This process is needed to correct errors in raw data, which is often incomplete and has an irregular format [41], [53]. The first step in data preprocessing is checking for missing values in the data in each variable. The results of checking the missing value are shown in Figure 4a, and the results of outlier data analysis using a boxplot are shown in Figure 4b.

Figure 4a explains that all the support system and facilities variables used in this study have a value of 0 for the missing value, thus indicating no missing value in all research variables [41], [53]. In addition, dtype indicates the type of data used, namely in the form of an integer. If there is no missing value in the data, we proceed to the following data preprocessing step. The next step is to check outlier data. Checking outlier data using box plots. Box Plot is a summary of data visualization presented graphically by describing sample distributions such as skewness, measures of central tendency, and measures of the spread or diversity of observational data [41]. Figure 4b explains that based on the results of checking outlier data using a box plot, there are no outliers, so the data used is feasible for data processing.

```
[86] df.dtypes
     df.isnull().sum()

     TL1    0
     TP1    0
     PS1    0
     WM1    0
     LT1    0
     CW1    0
     SI1    0
     ST1    0
     HT1    0
     SS1    0
     dtype: int64
```

(a)



(b)

Figure 4 - (a) Missing Value (b) Results of Box Plot Analysis

3.4. Data Processing (Find a number a cluster k)

After the data has been preprocessed to ensure no missing values and outlier data, the next step is to process the data by conducting model tuning. Model tuning is a process that can be carried out after obtaining the dataset, which starts with determining the goals to be achieved by selecting the algorithm using examples of tuning models such as classification, clustering, segmentation, or statistical tests [54]. In the k-means clustering algorithm, there are k parameter values so that before a subjective decision is made on the value of k, model tuning can be done, for example, with the Elbow method [41], [55]. The Elbow method is a method for determining the correct number of clusters through the percentage of comparison results between the number of clusters that will form an angle at a point [56]. The results of the analysis using the elbow method are presented in Figure 5.

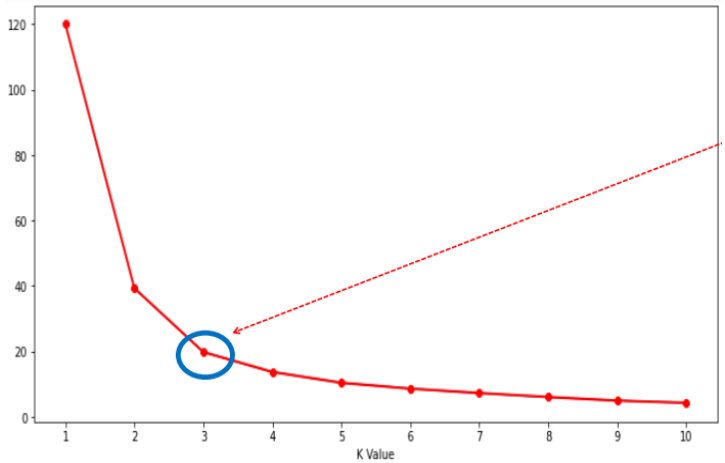


Figure 5 - Elbow Method Analysis Results

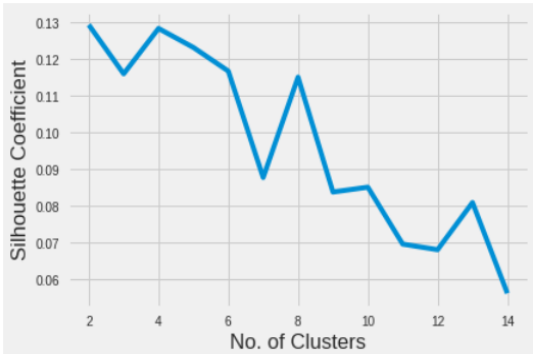
Figure 5 shows that based on the elbow method analysis of the data used, it is found that at point (k = 3), the inertia value has decreased, which has started to decrease. To strengthen the notion of this point as the optimal K value, it can be compared with clustering quality using the silhouette coefficient method. This study uses the silhouette coefficient method to determine the number of clusters and model evaluation. It aims to measure the close relationship between objects in a cluster, and the separation method measures how far a cluster is separated from other clusters [41], [57]. The results of the analysis using the silhouette coefficient method can be seen in Figure 6.

```
from sklearn.metrics import silhouette_score
km = KMeans(n_clusters=3, random_state=42, )
km.fit(std_atr)
score = silhouette_score(std_atr,km.labels_)
print('Silhouette Score: %.3f' % score)

Silhouette Score: 0.115
```

Figure 6 - Silhouette Method Analysis Results

Figure 6 shows that the analysis using the silhouette coefficient method obtained the selected silhouette score of 0.115. The silhouette score results are the selected scores from all the silhouette coefficient clusters [57]. Next is to choose the number of clusters (n-clusters) from the results of the silhouette score of each n-cluster. The results of the analysis using the silhouette coefficient method for each number of clusters can be seen in Figures 7(a) and 7(b).



(a)

```
For n_clusters=2, Silhouette Coefficient = 0.12934815545156297
For n_clusters=3, Silhouette Coefficient = 0.11598534087614129
For n_clusters=4, Silhouette Coefficient = 0.12835946470244777
For n_clusters=5, Silhouette Coefficient = 0.12317493386727858
For n_clusters=6, Silhouette Coefficient = 0.1167314674432404
For n_clusters=7, Silhouette Coefficient = 0.08769606664459982
For n_clusters=8, Silhouette Coefficient = 0.11508353822694566
For n_clusters=9, Silhouette Coefficient = 0.08375180749452925
For n_clusters=10, Silhouette Coefficient = 0.08507047792868512
For n_clusters=11, Silhouette Coefficient = 0.06954638658550315
For n_clusters=12, Silhouette Coefficient = 0.06802301717525455
For n_clusters=13, Silhouette Coefficient = 0.08088962347858182
For n_clusters=14, Silhouette Coefficient = 0.056138634847212765
```

(b)

Figure 7 - (a) Silhouette Coefficient Method Analysis (b) Result Value of Each n-cluster

Figure 7(a) shows the results of the silhouette coefficient graph that each number of clusters has a different silhouette coefficient value. The silhouette coefficient value is used to select the number of clusters (n-clusters). More exact values for each (n-cluster) can be seen in Figure 7(b). The silhouette coefficient value that corresponds to the silhouette score is (n-cluster) = 3 because it has a silhouette score value of 0.115. After getting the number of selected clusters, the next step is to evaluate the tuning model by checking the positive or negative silhouette score values for the objects in the three selected clusters. The results of the evaluation of the tuning model are shown in Figure 8.

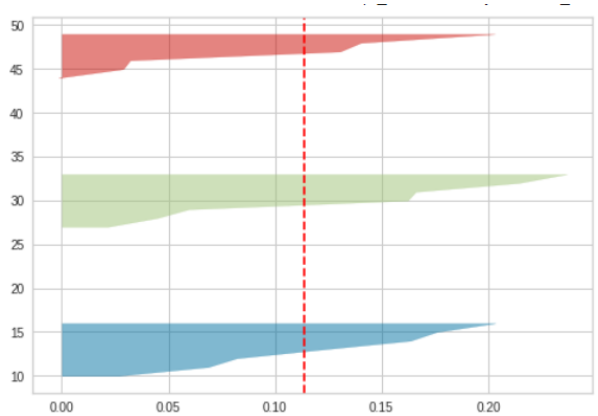


Figure 8 - Results of the Evaluation of the Tuning Model

3.5. Evaluation Result

The results of model evaluation by visualizing the tuning model show that all data objects have a positive silhouette score. It indicates that the objects have been grouped into the appropriate configuration, namely 3 clusters [57]. Based on data processing results, segmentation analysis of Rembang tourism objects can be visualized, as shown in Figure 9.

Figure 9 displays the visualization results of tourism potential cluster segmentation based on support systems and facilities variables. Based on the visualization results of cluster segmentation, the object points are included in three clusters. Namely, cluster 0 has a red color, cluster 1 has a blue color, and cluster 2 has a green color. The centroid values explain the differences regarding the three clusters for each research variable. The centroid is the center value (center) of a cluster (Suppose we set k=3; then centroids C1, C2, and C3 will be formed randomly. This stage will calculate the distance of each data to the centroid that has been made using Euclidean distance [57]. The centroid values for each variable affect the assessment of the clusters formed. The centroid values for the variables in each cluster are described in Table 4.

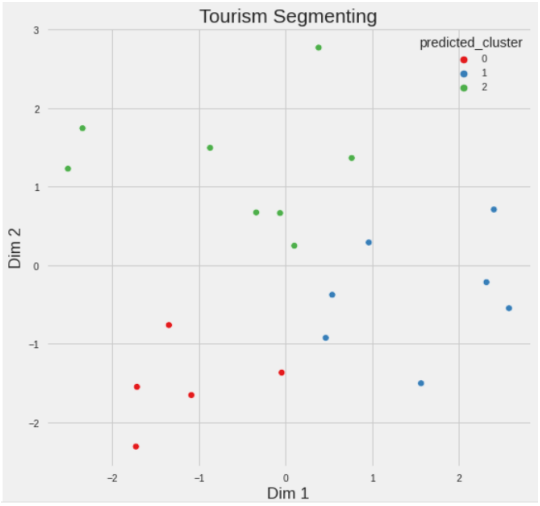


Figure 9 - Silhouette Method Analysis Results

Table 4 – Centroid Value of each Variable

	Centroids									
Cluster	**TL**	**TP**	**PS**	**WM**	**LT**	**CW**	**SI**	**ST**	**HT**	**SS**
0	-0,209	0,639	0,101	-0,994	-0,399	-0,477	0,641	0,016	-0,240	0,738
1	0,838	-0,611	0,186	-0.229	-0,399	-0,119	-0,674	-0,477	-0,756	-0,998
2	-0,471	-0,021	-0,064	0,918	0,599	0,447	0,025	0,345	0,748	0,195

Table 4 shows the results of the centroid value of each variable. The centroid values for each variable affect the assessment of the clusters formed. Table 4 shows the results of the centroid values for each variable. The centroid value of each variable affects the assessment of the clusters formed. In cluster 0, the most significant centroid values are safety and security, supporting industry, and transportation, with values of 0.738, 0.642, and 0.639. In cluster 1, the largest centroid values are telecommunications and power sources, with values of 0.838 and 0.186. Whereas in cluster 2, the largest centroid values are waste management, hospitality, location, clean water source, and spatial, with values of 0.918, 0.748, 0.599, 0.447, and 0.345.

3.6. Evaluation Result

The next step is to identify the members of each Rembang tourist attraction cluster. The mapping results of tourism mapping in Rembang Regency are grouped based on the support system and facility (SSF), including telecommunication, power source, transportation, waste management, location, clean water source, supporting industry, spatial, hospitality, safety, and security. It shows that the best cluster means the system and support facilities are better than the other clusters. A detailed explanation of the tourist clusters in Rembang Regency based on the centroid value of each variable is presented in Table 5.

Table 5 – Cluster Value

Cluster	Cluster Value	Cluster Name
0	Moderate (-0,184)	Popular Tourism
1	Worst (-3,239)	Local Tourism
2	Best (2,271)	Best Tourism

Table 5 shows that each cluster has its name. Cluster 0 has a cluster value of -0.184 in a moderate position, where each tourism potential in cluster 0 is called a famous tourism cluster. Cluster 1 has a cluster value of -3.239 in the worst position, where every tourism potential in cluster 1 is called a local tourism cluster. While cluster 2 has a cluster value of 2.271 in the best position, every tourism potential in cluster 2 is named best tourism. The results of mapping tourism potential in Rembang Regency, Central Java, Indonesia, are shown in Table 6.

There are many challenges identified in this study. This analysis was carried out in Rembang Regency, Central Java, Indonesia, as a snapshot analysis where segments are identified at specific regional points. However, the results of this

cluster need further research if applied to areas with different characteristics because after the segment changes, the fundamental analysis must be repeated, and the strategy must be adapted in such a way [33], [57].

Table 6 – Tourism Cluster in Rembang Regency

Cluster Name	Tourism
Local Tourism	Warung Makan Bu Tin
	Air Terjun Pasucen
	Watu Layar
	Pulau Gede Rembang
	Waduk Panohan
	Warung Nasi Gandul Gambiran Pak Wandi
	Nasi Tahu Srepeh Bu Slamet
Popular Tourism	Nasi Gandul Soto Sapi
	Air Terjun Kedung Grenjeng
	Gunung Argopuro Lasem
	Jembatan Merah Hutan Mangrove
	Ayam Geprek Sambel Korek
Best Tourism	Vihara Ratanavana Arama
	Petilasan Sunan Bonang
	Pantai Caruban
	Pantai Nyamplung Indah
	Wisata Gunung Kajar
	Alun-Alun Rembang
	Situs Perahu Kuno Punjulharjo
	Rumah Makan Cemoro Sewu

Based on the centroid values of the research variables studied, there are several recommendations from this study for strategies to increase the potential of tourist attractions based on the existing clusters of tourist attractions in Rembang Regency, namely:

1. Local Tourism: If stakeholders want to increase the tourism potential included in this cluster, stakeholders can focus on eight sections, namely Transportation, Waste Management, Location, Clean Water Sources, Supporting Industries, Spatial Planning, Hospitality, and Safety and Safety because these variables still have negative centroid values, so they are still considered inadequate. Strategies to improve transportation by providing solutions to the problem of the distance of tourist attractions from the city center. Strategy to improve waste management by increasing the number of trash cans. Strategy to increase the location by improving the tourism area. The strategy for increasing clean water sources is to increase the number of clean water sources at tourist sites. The strategy is to increase the Supporting Industry by increasing the number of business areas in tourist areas. The strategy to improve Spatial Planning is to increase the number of tourist supports around tourist attractions. The strategy is to increase Hospitality by increasing the number of public facilities in tourist areas. As well as a strategy to improve safety and security by increasing the availability of security posts and medical equipment.

2. Popular Tourism: If stakeholders want to increase the tourism potential included in this cluster, stakeholders can focus on five sections, namely Telecommunications, Waste Management, Location, Clean Water Sources, and Hospitality, because these variables still have centroid values negative, so it is still considered inadequate. Strategy to improve Telecommunications by improving Average Internet Speed. Strategy to improve waste management by increasing the number of trash cans. Strategy to increase the location by improving the tourism area. The strategy for increasing clean water sources is to increase the number of clean water sources at tourist sites. The strategy is to increase Hospitality by increasing the number of public facilities in tourist areas.

3. Best Tourism: If stakeholders want to increase the tourism potential included in this cluster, stakeholders can focus on three parts: Telecommunication, Power Sources, and Transportation. Because these variables still have negative centroid values, they are still considered inadequate. Strategy to improve Telecommunication by improving Average

Internet Speed. Strategy to increase Power Source by increasing Total power source. As well as strategies to improve transportation by providing solutions to the problem of the distance of tourist attractions from the city center.

4. CONCLUSION

The strategy for developing tourism potential in Rembang Regency can be carried out using the tourism industry's support system and Facilities (SSF). The development of tourism potential is more effective if we know the characteristics of the supporting infrastructure for potential tourist attractions to produce a segmentation of tourism potential in Rembang Regency, Central Java, Indonesia. In increasing the potential of a sustainable tourism industry in Rembang Regency, it can be concluded that there are 3 clusters, namely cluster Local Tourism with the tour of Warung Makan Bu Tin, Air Terjun Pasucen, Watu Layar, Pulau Gede Rembang, Waduk Panohan, Warung Nasi Gandul Gambiran Pak Wandi, Nasi Tahu Srepeh Bu Slamet. Cluster Popular Tourism consists of Nasi Gandul Soto Sapi, Air Terjun Kedung Grenjeng, Gunung Argopuro Lasem, Jembatan Merah, Hutan Mangrove, Ayam Geprek Sambel Korek, and Cluster Best Tourism consists of Vihara Ratanavana Arama, Petilasan Sunan Bonang, Pantai Caruban, Pantai Nyamplung Indah, Wisata Gunung Kajar, Alun-Alun Rembang, Situs Perahu Kuno Punjulharjo, and Rumah Makan Cemoro Sewu.

This research has provided policy directions to stakeholders in Rembang Regency so they can be executed in developing tourism potential. This study has limitations that can be improved in future studies. The time interval for collecting different data can be too long due to the peak of the Covid-19 pandemic, so we need help collecting data. Thus, future research is expected to obtain data at the right time of collection. In future research, we may consider all potential tourist objects in the Rembang Regency to obtain an overall mapping of tourism potential.

Disclaimer

The authors whose names are written certify that they have no conflict of interest

Acknowledgement

This research was supported by Telkom University and the Enterprise System and Solution Lab, who have helped provide the time and opportunity to collect data in Rembang Regency, Central Java, Indonesia. We want to thank the Rembang Regency Tourism Office and managers of tourist attractions in Rembang Regency for providing information related to the data needed in the research. We do not forget to thank "Rumah Merah Lasem" and “Lutfia Septiningrum”, which have provided all the facilities and infrastructure for data collection.

REFERENCES

[1] L. J. Lickorish, and C. L. Jenkins, “Introduction to tourism,” *Routledge*. 2007. DOI: doi.org/10.1016/j.tourman.2015.10.006
[2] D. C. Bojanic, and M. Lo, “A comparison of the moderating effect of tourism reliance on the economic development for islands and other countries,” *Tourism Management*, vol. 53, pp. 207-214. 2016. DOI: https://doi.org/10.1016/j.tourman.2015.10.006
[3] L. Septiningrum, A.F. Rizana, R.P. Soesanto, and A.A. Rumanti, “Success Factor for Elevating Indonesian Tourism Area Potential,” *IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. 2022.
[4] World Tourism Organization, “Tourism Highlights, 2020 Edition,” 2020. Retrieved from https://www.e-unwto.org/doi/pdf/10.18111/9789284420853 (Accessed Nov. 28, 2022)
[5] P. Kotler, J. T. Bowen, and J.C. Makens, “Marketing for Hospitality and Tourism (7th ed.),” *Pearson*. 2017.
[6] M. A. Camilleri, “The Tourism Industry: An Overview,” *Tourism, Hospitality and Event Management*, pp. 3–27, 2018.
[7] Y. Sulistyadi, “*Buku Indikator Perencanaan Pengembangan Pariwisata Berkelanjutan,”* Bandar Lampung: Anugrah Utama Raharja, 2019.
[8] G. Lohmann, and A. P. Netto, “Tourism theory: Concepts, models, and systems,” *Cabi*. 2016.
[9] M. A. Kashem, M. Shamsuddoha, T. Nasir, and A. A Chowdhury, “The role of artificial intelligence and blockchain technologies in sustainable tourism in the Middle East,” *Worldwide Hospitality and Tourism Themes*, (ahead-of-print). 2022.
[10] A. Japutra, and R. Situmorang, “The repercussions and challenges of COVID-19 in the hotel industry: Potential strategies from a case study of Indonesia,” *International Journal of Hospitality Management*, vol. 95, no. 102890. 2021. DOI: https://doi.org/10.1016/j.ijhm.2021.102890
[11] J. A. Folgado-Fernández, A. M. Campón-Cerro, and J. M. Hernández-Mogollón, “Potential of olive oil tourism in promoting local quality food products: A case study of the region of Extremadura, Spain,” *Heliyon*, vol. 5, no. 10, e02653. 2019.
[12] J. Abbas, R. Mubeen, P. T. Iorember, S. Raza, and G. Mamirkulova, “Exploring the impact of COVID-19 on tourism: transformational potential and implications for a sustainable recovery of the travel and leisure industry,” *Current Research in Behavioral Sciences*, vol. 2, 100033. 2021.
[13] F. Serravalle, A. Ferraris, D. Vrontis, A. Thrassou, and M. Christofi, “Augmented reality in the tourism industry: A multi-stakeholder analysis of museums,” *Tourism Management Perspectives*, vol. 32, no. January 2019. DOI: 10.1016/j.tourman.2015.10.006
[14] H. H. Adinugraha, M. Sartika, and A. Kadarningsih, “Desa wisata halal: konsep dan implementasinya di Indonesia,” Human Falah, vol. 5, no. 1, pp. 28-48. 2018.
[15] M. Melinda, “Permasalahan Pengembangan Potensi Pariwisata,” 2019.
[16] G. K. J. Joandani, R. Pribadi, and C. A. Suryono, “Kajian potensi pengembangan ekowisata sebagai upaya konservasi mangrove di Desa Pasar Banggi, Kabupaten Rembang,” *Journal of marine Research*, vol. 8, no. 1, pp. 117-126. 2019.
[17] R. Yusantiar, and D. W. Soewardi Koen, “Perancangan Identitas Visual untuk Promosi Pariwisata Kabupaten Rembang,” *ANDHARUPA: Jurnal Desain Komunikasi Visual & Multimedia*, vol. 4, no. 02, pp. 207-220. 2018.
[18] Badan Pusat Statistik, “Konsep dan Definisi Statistik Kunjungan Wisatawan Mancanegara”. [Online]. Available: https://www.bps.go.id/subject/16/pariwisata.html (Accessed Des. 01, 2022).

[19] F. Achmad, L. Septiningrum, Y. Prambudia, A. A. Rumanti, and I. I. Wiratmadja, "Tourism Industry Development Model through Support System Facilities, Stakeholder Participation, and Environmental Dynamism: Dealing with Post-Pandemic Covid-19 Approach". *International Journal of Innovation, Management and Technology,* 2023.

[20] DISPORA Jawa Tengah, Buku Pariwisata Jawa Tengah Dalam Angka 2020. Semarang, 2022. Accessed: Apr. 08, 2022.

[21] Pemerintah Kabupaten Rembang. "PEMKAB REMBANG TERUS DORONG DESA KEMBANGKAN POTENSI WISATA". [online]. Available: https://rembangkab.go.id/berita/pemkab-rembang-terus-dorong-desa-kembangkan-potensi-wisata/ (Accessed Des, 04 2022).

[22] H. Sasana, H. Nurcahyanto, and I. Novitaningtyas, "The development strategy of world heritage tourism in Indonesia," *African Journal of Hospitality, Tourism and Leisure*. vol. 8 no. 5, pp. 1-14. 2019.

[23] E. N. Homer, A. D. Wicaksono, and F. Usman, "Penentuan jenis klaster industri di kawasan industri arar kabupaten sorong berdasarkan metode delphi dan analytical hierarchi process (ahp)," *The Indonesian Green Technology Journal*. Vol. 5 no. 1, pp. 16-23. 2016.

[24] J. M. Meis, F. M. Enzinger, K. L. Martz, and J. A. Neal, "Malignant peripheral nerve sheath tumors (malignant schwannomas) in children," *The American journal of surgical pathology*, vol. *16*, no. 7, pp. 694-707. 1992.

[25] I. G. B. R. Utama, "Pengantar Industri Pariwisata," *Deepublish*. 2015.

[26] D. T. Anggarini, "Upaya pemulihan industri pariwisata dalam situasi pandemi Covid-19," *Jurnal Pariwisata*, vol. 8, no. 1, pp. 22-31. 2021.

[27] I. Azizuddin, and M. H. Ainul Yaqin, "Industri pariwisata halal: pendorong inovasi untuk halal preneurs". *Jurnal Ilmiah Ekonomi Islam*, vol. 8, no. 1, pp. 106-116. 2022.

[28] Y. Heryati, "Potensi Pengembangan Obyek Wisata Pantai Tapandullu di Kabupaten Mamuju," *GROWTH Jurnal Ilmiah Ekonomi Pembangunan*, vol. 1 no. 1, pp. 56-74. 2019.

[29] N. K. Ahmadi, "Analisis Segmentasi Terhadap Keputusan Pembelian Produk Eiger di Bandar Lampung," *Jurnal Manajemen Magister Darmajaya*, vol. 3 no. 01, pp. 75-95. 2017.

[30] P. Kotler, and A. Gary, "Prinsip-Prinsip Pemasaran. Edisi 13. Jilid 1," *Jakarta. Erlangga*. 2012.

[31] N. Elgendy, and A. Elragal, "Big data analytics in support of the decision-making process". *Procedia Computer Science*, vol. 100, pp. 1071-1084. 2016.

[32] S. Agarwal, K. Narayanan, M. Sinha, R. Gupta, S. Eswaran, and T. Mukherjee, "Decision Support Framework for Big Data Analytics," *IEEE World Congress on Services (SERVICES)* (pp. 53-54). IEEE. 2018, July.

[33] E. F. Zineb, R. A. F. A. L. I. A. Najat, and A. B. Jaafar, "An Intelligent Approach for Data Analysis and Decision Making in Big Data: A Case Study on E-commerce Industry," *International Journal of Advanced Computer Science and Applications*, vol. 12 no. 7. 2021.

[34] F. Sari, "Metode dalam pengambilan keputusan". *Deepublish*. 2018.

[35] J. Lu, Z. Yan, J. Han, and G. Zhang, "Data-driven decision-making (d 3 m): Framework, methodology, and directions," *IEEE transactions on emerging topics in computational intelligence*, vol. 3, no. 4, pp. 286-296. 2019.

[36] D. Tangian, B. D. Polii, and Mengko, "Analysis of Potential Tourism Attractions in Manado City," *International journal of Tourism and hospitality in Asia Pasific (IJTHAP)*, vol. *3*, no. 1, pp. 102-109. 2020.

[37] D. Lee-Ross, and C. Lashley, "*Entrepreneurship and small business management in the hospitality industry," Routledge*. 2010.

[38] S. J. Page, "*Tourism management*," *Routledge*. 2014.

[39] C. N. Knaflic, "Storytelling with data: A data visualization guide for business professionals," *John Wiley & Sons*. 2015.

[40] V. Friedman, "Data visualization and infographics in: Graphics," *Monday inspiration*, vol. 14, 2008.

[41] P. T. Tan, M. Steinbach, and V. Kumar, "Introduction to data mining". *Pearson Education India*. 2016.

[42] E. Patel, and D. Kushwaha, "Clustering cloud workloads: K-means vs gaussian mixture model," *Procedia computer science*, vol. 171, pp. 158-167. 2020.

[43] Y. Hozumi, R. Wang, C. Yin, and G. Wei, "UMAP-assisted K -means clustering of large-scale SARS-CoV-2 mutation datasets," *Computers in Biology and Medicine*, vol. 131, pp. 1–14. 2021. DOI: https://doi.org/10.1016/j.compbiomed.2021.104264

[44] J. E. Z. Gbadoubissa, A. A. A. Ari, and A. M. Gueroui, "Efficient k-means based clustering scheme for mobile networks cell sites management," *Journal of King Saud University-Computer and Information Sciences*, vol. 32, no. 9, pp. 1063-1070. 2020.

[45] N. Wakhidah, "Clustering menggunakan k-means algorithm". *Jurnal Tranformatika*, vol. 8, no. 1, pp. 33-39. 2010.

[46] Y. Wicaksono, "SEGMENTASI POTENSI DAERAH WISATA DI KOTA BIMA MENGGUNAKAN K-MEANS," *Indonesian Journal of Informatics and Research*, vol. 2, no. 2, pp. 26-32. 2021.

[47] D. Sugiyono, "Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D," 2013.

[48] Etikan, I., & Bala, K. (2017). Sampling and sampling methods. Biometrics & Biostatistics International Journal, 5(6), 00149.

[49] S. R. P. Wulung, "a Spatial Approach for the Development of Sustainable Tourism Destinations," *J. Sustain. Tour. Res*., vol. 3, no. 2, pp. 68–73, 2021.

[50] V. Owiyo and J. M. Mulwa, "Safety and security in tourism destinations: Its moderating role in the destination competitiveness determinants and destination competitiveness nexus in western Kenya circuit," *Int. J. Res. Manag. Bus. Stud*., vol. 5, no. 3, pp. 30–33, 2018.

[51] F. M. Rafif, S. Wulandari, and B. H. Sagita, "Positioning Strategy Design of Online Travel Agent Applications Based on Perceptual Mapping," *International Journal of Innovation in Enterprise System*, vol. *6*, no. 01, pp. 74-84. 2022.

[52] S. Idreos, O. Papaemmanouil, and S. Chaudhuri, "Overview of data exploration techniques," *In Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data* (pp. 277-281). 2015, May.

[53] H. Yang, "Data preprocessing," *Pennsylvania State University*: *Citeseer.* 2018.

[54] H. Fan, L. Zheng, C. Yan, and Y. Yang, "Unsupervised person re-identification: Clustering and fine-tuning," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 14, no. 4, pp. 1-18. 2018.

[55] I. Indrajani, "Analisis dan Penerapan Metode Tuning pada Basis Data Funding," *ComTech: Computer, Mathematics and Engineering Applications*, vol. 6, no. 1, pp. 143-153. 2015.

[56] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, "Integration k-means clustering method and elbow method for identification of the best customer profile cluster," *In IOP conference series: materials science and engineering*. vol. 336, no. 1, p. 012017. IOP Publishing. 2018, April.

[57] K. R. Shahapure, and C. Nicholas, "Cluster quality analysis using silhouette score," *In 2020 IEEE 7th International Conference on Data Science and Advanced Analytics (DSAA)* (pp. 747-748). 2020.