MSME Clusterization Using K-Means Clustering in Garut Regency, Indonesia

Dini Turipanam Alamanda
Faculty of Economics Universitas Garut

Eti Kusmiati*
Faculty of Economics Universitas Garut

Diqy Fakhrun Shiddieq
Faculty of Economics Universitas Garut

Fikri Fahru Roji
Faculty of Economics Universitas Garut

ABSTRACT
Clusters are groups of activities consisting of core industries, related industries, supporting industries, and supporting and related economic activities, which are interrelated and mutually supportive. This study aims to cluster micro, small and medium enterprises (MSMEs) in Garut Regency based on their similarity of attributes attached to the factors. These aspects include a taxpayer identification ownership, business license ownership, total net worth, total capital, number of employees, total business volume, product certification, product license, and brand. Clusterization has an important urgency for the development of MSMEs. Clustering can help group MSMEs into groups that are easier to analyze and make the right decisions to increase the efficiency and competitiveness of MSMEs. In addition, the application of clustering for MSMEs can have positive implications such as identifying the characteristics of MSMEs, developing appropriate marketing strategies, improving product or service quality, collaboration and information exchange, and access to resources. This can help MSMEs to increase their efficiency and competitiveness in the market. After data screening, the sample of MSMEs was analyzed using K-means clustering analysis and the help of RapidMiner 5.3 software. There were six iterations conducted before the data position value was stable, with the results revealing three MSME clusters. Cluster 2 had the highest percentage of MSMEs, with 84%. Cluster 1 had a higher percentage of MSMEs with ownership of business license numbers than Clusters 2 and 3. Cluster 3 had a larger and more significant number of MSMEs and a larger total net worth and ownership of product and brand licenses than the other clusters, while Cluster 2 showed an intermediate profile in almost all factors. The study results provide an overview of the local government in adopting cluster formation policies in Garut Regency.

Keywords: Clusterization, K-Means Clustering, MSME.

1. INTRODUCTION
Micro, small, and medium enterprises (MSMEs) play an important role in absorbing more than 97% of the workforce in Indonesia, which impacts 57% of the gross domestic product (GDP) (Indonesia Central Bureau of Statistics, 2021). The existence of MSMEs in the structure of the Indonesian economy occupies a very important position. Therefore, efforts
are needed to develop it more comprehensively. One of the efforts to develop and empower MSMEs is the cluster system approach.

MSME clustering is considered one of the strategies for developing MSMEs (Ganau & Rodríguez-Pose, 2018; Kim & Hwang, 2019). Clusters are developed in a management system, where similar business entities in one area form a group in which collaboration, synergy, and unity can be created, ultimately leading to strength (Oshima & Toma, 2023; Kiryluk-Dryjska & Wieckowska, 2020). Another viewpoint explains that clusters are geographical concentrations of companies and institutions that interact with each other in a particular field, marked by a series of related industries and other entities that are important for business continuity (Romanova et al., 2019). A clustering process is a form of internal integration and collaboration that can provide resilience and synergistic effects in countering competitive pressures (Idrissi et al., 2019). Furthermore, clusters require collaboration between the actors within the cluster, which is, in part, a strategy to gain a competitive edge (Cottineau & Arcute, 2020).

Initially, the idea of clusters was closely related to geographical proximity-based businesses (Ganau & Rodríguez-Pose, 2018; Cottineau & Arcute, 2020). The cluster concept expanded with an emphasis on value creation, economic efficiency (Porter, 2000), and corporate collaboration (Villa & Taurino, 2018; Sultan et al., 2020). Furthermore, Porter (2000) emphasized clusters in the value chain concept to produce one type of product, while the close distance between business groups is considered added value due to agglomeration. In comparison, Lan et al. (2019) proposed an MSME industrial cluster based on a wireless network.

Several studies related to MSME clustering were carried out by Felzensztein et al. (2019), which indicated that cluster policies benefit businesses that are joined to benefit company internationalization. Idrissi et al. (2019) defined the strengths and weaknesses in both clusters (Morocco and Italy), covering three dimensions of activity: human and material resources, activities, processes, and strategies. McPhillips (2020) pointed out that there needs to be a comprehensive understanding of the potential role of clusters as innovation intermediaries in the context of economic transition. Ikram et al. (2018) stated that the provision of business-friendly facilities, as a synergistic benefit of supply chain vertical and horizontal (cluster) integration, boosts the overall competitiveness of MSMEs. In contrast, Saha et al. (2018) examined the concept of clusters and specialization as an effective strategy that utilizes entrepreneurial development and creates value based on knowledge and innovation.

Joining business entities in a cluster can help businesses, especially MSMEs in developing countries (Alamanda et al., 2022). Thus, policymakers must anticipate problems that will occur before developing clusters so that they can take the necessary steps to overcome them. Much information is needed to process and identify cluster development strategies for MSMEs, which will enable the redefinition of policies in the field of cluster development.

In the context of SMEs, clustering can have several urgencies, including (1) Facilitate Data Analysis. By clustering, very complex and diverse MSME data can be grouped into simpler groups. This allows easier data analysis because each group can be analyzed separately and with more focus; (2) Facilitate Decision Making. With clustering, decision-making in MSME development becomes easier because each group can identify its characteristics and take steps according to these characteristics; (3) Increasing the Efficiency of MSME Development. By clustering, SMEs can identify groups that have the same characteristics
and needs. This can increase efficiency in MSME development because development can be carried out in a more focused and more effective manner; (4) Increasing MSME Competitiveness. By clustering, MSMEs can be grouped into groups that have the same characteristics as other MSMEs. This can assist MSMEs in strengthening their competitiveness by identifying and adopting best practices from groups with similar characteristics. The application of clustering for MSMEs can have several implications, including: by clustering, the characteristics of MSMEs can be identified more clearly based on the group, which allows MSMEs to know their strengths and weaknesses in competition in the market. Clusterization also has implications for developing appropriate marketing strategies, improving product or service quality, increasing collaboration and information exchange between MSME groups, and increasing access to resources.

Based on these objectives, factor analysis techniques are necessary to achieve the best results when determining MSME clusters. Thus, this research is aimed at categorizing MSME clustering in the context of business entities in the same class and category. In other words, MSME clusterization is formed not in the context of clusters based on geographical proximity but in the proximity of attributes.

2. LITERATURE REVIEW

2.1. Cluster

Smirnyagin (2022) used the term of “cluster” to represent the concentration of firms in a territorial space, as well as to indicate the concentrating of production. Porter (2000) stated that a company's success depended heavily on its economic environment's competitiveness, determining the basic conditions (shared resources) and competition within the cluster.

Porter (1990) was the main promoter of economic clusters. He defined a cluster as a group of related companies and geographically neighboring related organizations operating in a certain area, characterized by common activities, and complementing each other. Clusters are a combination of mutual interests in a geographical area of interconnected business structures, representatives of local authorities, and scientific institutions within a certain area that provide enterprises and organizations from various industries with modern machinery and equipment, new technologies to create conditions for improving production efficiency, and product competitiveness in domestic and foreign markets (Liu et al., 2019). Kaplinsky & Morris (2019) defined clusters as an agglomeration of economic agents within one economic sector that are geographically adjacent and interact with each other. According to Derlukiewicz et al. (2020), a cluster is a geographically limited concentration of related similar and complementary businesses with active channels for business transactions, communication, and dialogue that face common threats and share common opportunities. Clusters are the concentration of similar industries in a geographic area complemented by core and supporting industries (Kim & Hwang, 2019).

The definition of a cluster has expanded from a simple idea to encompass a broad and complex meaning. At its inception, the definition of clusters was mostly related to business development based on geographical proximity (Ganau & Rodriguez-Pose, 2018; Cottineau & Arcaute, 2020; Shiposha, 2020). However, as the concept developed, it expanded to include the important social infrastructures, information exchanges, and collaborations between companies (Villa & Taurino, 2018; Sultan et al., 2020). Porter (2000) indicates that there is no definite limit regarding geographical proximity between business units in a cluster. The geographical criteria referred to lies more in whether economic efficiency over
existing physical distances manifests itself in various profitable business activities. Andersson et al. (2004) described two types of clusters. First, they determined that clusters are spatial groupings of similar and related companies and industries, including regional or local groups. In comparison, the second cluster type is considered a functionally defined system of interrelated actions. Hence, functionally related systems are less constrained in a narrowly defined area.

Clusters have two key elements: companies in interconnected clusters located in a near proximity that is easily recognized as an industrial area (Kiryluk-Dryjska & Wieckowska, 2020). Another opinion explains three basic characteristics, namely: (1) communality/similarity/togetherness/unity, i.e., businesses operate in fields that are similar or related to one another with a shared market focus or a range of joint activities; (2) concentration, i.e., there are groupings of businesses that can and do interact; and (3) connectivity, i.e., there are interrelated/dependent organizations with different types of relationships (Batz et al., 2018; Yigit Ozkan et al., 2020).

### 2.2. MSME Development Strategy through Clusterization

The performance of MSME-based clusters in developing countries has been the focus of researchers (Lan et al., 2019). Humphrey and Schmitz (2002) state that the cluster concept drives economic growth in developing countries. Literature that analyzes clusters confirms that these groupings in developing countries mostly consist of small and medium enterprises (Kumar et al., 2020). In developing countries these small to medium-scale enterprises operate vertically within close social bonds and production chains and are not fully integrated (Telizhenko et al., 2019). Factors, such as limited access to information and distrust of legal contracts, were also identified as obstacles to cluster improvement (Yigit Ozkan et al., 2020). On the other hand, industrial activity in developing countries tends to be concentrated in certain locations, especially certain metropolises and capital cities, due to the need for more infrastructure in suburban areas (Sultan et al., 2020).

Even though there are challenges as described, some researchers consider that clusters are a strategy that can be used to improve the performance of MSMEs (Kuo et al., 2022). The cluster directs industrial/company collaboration with other institutions beneficial for competition, including providing raw materials and infrastructure. Batz et al. (2018) state that MSMEs tend to join clusters that aim to accelerate the innovation process and increase their possibilities to access external knowledge and resources. MSMEs can optimize their business interactions within the cluster through the supply chain, such as customers, suppliers, and producers (Kayvanfar et al., 2019). Clusters make MSMEs more competitive, and cause them to grow in the value chain, develop managerial capabilities and worker skills, and generate efficiency (Kaplinsky & Morris, 2019). In other words, the impact of competition within the cluster causes an increase in company productivity through innovation and the expansion and strengthening of companies within the cluster itself. McPhillips (2020) states that the role of clusters is to increase openness and collaboration not only for member companies but for all participants in the regional innovation ecosystem. Meanwhile, Idrissi et al. (2019) state that clusters could offer innovative products of higher quality in terms of technology components and ultimately cost less.

### 3. METHODS
The purpose of this study is to group Micro small to medium-sized enterprises (MSMEs) into clusters. This research uses quantitative methods. The data comes from the Garut Regency Cooperative and MSME Office and lists 25,688 units. Data processing began with screening the completeness of the MSME data. This data is then used as an attribute that is analyzed, which includes ownership of a taxpayer identification number, ownership of a business license number, total net worth, total capital, number of employees, total business volume, ownership of product certification, ownership of product licenses, and brands.

The initial screening showed that 8,919 MSME units had complete data according to the research attributes and were ready for further analysis. The cluster formation process was carried out using the K-means clustering algorithm, which groups data based on the cluster center point (centroid) closest to the data. Objects were grouped into one or more clusters so that objects in one cluster would have a high similarity between one object and another (Viloria & Lezama, 2019). K-means clustering is one of the algorithms in determining the classification of objects based on the attributes/features of these objects into K clusters/partitions. K is a positive number representing the number of groups/clusters/partitions to the object. Data partitioning is done by finding the minimum distance value between the data and the centroid values set randomly or with the initial set of centroids. The centroid value can also be determined based on the successive K objects.

The following seven steps are part of the process of applying K-means clustering: (1) normalizing data; (2) determining the number of clusters to be formed; (3) generating random values for the initial cluster center (centroid); (4) calculating the distance of the input data to each centroid; (5) grouping each data based on its proximity to the centroid with the smallest distance; (6) updating new centroid values; and (7) repeating until no cluster members change (Cicea et al., 2019).

After the screening, which aims to get data that is ready to be processed (data normalization), the next step was to classify data into numbers (codes) to facilitate the clustering process (codification). The codification process was carried out based on the grouping of SMEs into certain attributes. The grouping of MSMEs in Garut Regency was carried out by considering the nine attributes described in Table 1.

### Table 1. Classification of Research Variables

<table>
<thead>
<tr>
<th>No</th>
<th>Variable (Attribute)</th>
<th>Criteria and Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ownership of taxpayer identity (V1)</td>
<td>Available [1], Not Available [2]</td>
</tr>
<tr>
<td>2</td>
<td>Ownership of a business license (V2)</td>
<td>Available [1], Not Available [2]</td>
</tr>
<tr>
<td>3</td>
<td>Total net worth (IDR) (V3)</td>
<td>0 – 49 M [1], 50-99 M [2], 100 – 149 M [3], 150 – 199 M [4], 200-249 M [5], 250 – 299 [6], 300 – 349 M [7], 350 – 399 M [8], 400-449 M [9], 450-499 M [10], &gt;500 M [11]</td>
</tr>
<tr>
<td>4</td>
<td>Amount of capital (IDR) (V4)</td>
<td>0 – 49 M [1], 50-99 M [2], 100 – 149 M [3], 150 – 199 M [4], 200-249 M [5], 250 – 299 [6], 300 – 349 M [7], 350 – 399 M [8], 400-449 M [9]</td>
</tr>
<tr>
<td>5</td>
<td>Number of employees (V5)</td>
<td>0-4 [1], 5-9 [2], 10-14 [3], 15-19 [4], 20-24 [5], 25-29 [6], 30-35 [7], &gt; 35 [8]</td>
</tr>
<tr>
<td>6</td>
<td>Total business volume (V6)</td>
<td>0 – 49 M [1], 50-99 M [2], 100 – 149 M [3], 150 – 199 M [4], 200-249 M [5], 250 – 299 [6]</td>
</tr>
<tr>
<td>7</td>
<td>Ownership of product certification (V7)</td>
<td>Available [1], Not Available [2]</td>
</tr>
</tbody>
</table>
4. RESULTS

In this study, the repetition was carried out in six iteration stages to get the data position value not to change, which means the value of each cluster was stable and had not experienced much change. Based on data processing using the K-means algorithm, the following results were obtained: (1) clusters that formed three clusters of 8,919 business units; (2) the number of MSMEs that joined in Cluster 1 included 457 business units (5% of the total MSMEs). In Cluster 2, as many as 7,535 business units (84% of total MSMEs) were included, and Cluster 3 represented 927 business units (10% of the total SMEs). Table 2 describes the cluster center for each variable studied.

Table 2. Cluster Center in Each Variable

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Cluster 1</th>
<th>Cluster 2</th>
<th>Cluster 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Centroid Score</td>
<td>Number of MSME</td>
<td>Centroid Score</td>
</tr>
<tr>
<td>V1</td>
<td>1.7046</td>
<td>779</td>
<td>1.9973</td>
</tr>
<tr>
<td>V2</td>
<td>1.6937</td>
<td>774</td>
<td>1.1979</td>
</tr>
<tr>
<td>V3</td>
<td>1.1269</td>
<td>515</td>
<td>1.5160</td>
</tr>
<tr>
<td>V4</td>
<td>1.0919</td>
<td>499</td>
<td>1.2495</td>
</tr>
<tr>
<td>V5</td>
<td>1.8796</td>
<td>859</td>
<td>1.9988</td>
</tr>
<tr>
<td>V6</td>
<td>1.3961</td>
<td>638</td>
<td>1.9989</td>
</tr>
<tr>
<td>V7</td>
<td>1.0656</td>
<td>487</td>
<td>1.9979</td>
</tr>
<tr>
<td>V8</td>
<td>1.1794</td>
<td>539</td>
<td>1.2319</td>
</tr>
<tr>
<td>V9</td>
<td>1.0810</td>
<td>494</td>
<td>1.0389</td>
</tr>
</tbody>
</table>

Centroid values indicate member similarity to the cluster or proximity object to the mean value in the cluster. For example, in Table 2 above, the n-value centroid in Cluster 1 is 1.7046 (in the variable ownership of a taxpayer identification number). That means each MSME that joins Cluster 1 has a value not far from its centroid value (gathered close to the value centroid). In other words, MSMEs that join Cluster 1 are those that have similarities in the attributes that are built.

Table 2 also shows that of all the attributes formed, Cluster 2 occupies the most positions in collecting the number of SMEs that joined. Most MSME members join. Cluster 2 has closeness in the attributes of business volume, while the most slightly joined are MSMEs that have closeness in terms of the ownership of product brands. Figure 1 below shows the results of clustering based on the variables formed using RapidMiner 5.3 tools.
Figure 1. MSME Clusterization based on Attribute
To conduct further analysis of the joining MSMEs for each cluster, analysis was carried out through the cluster performance graph. Figure 2 shows that three clusters formed with nine attributes. In Cluster 1, the joined MSMEs have ownership attributes of product certification, product license ownership, and more brand ownership than other attributes. However, the incorporated SMEs have the lowest total income compared to the MSMEs that joined in Clusters 2 and 3. In the second and third clusters, the MSMEs joined have the same attribute, namely the amount of net worth. The difference is that the net worth owned by each MSME in the third cluster is higher than in the second cluster.

The third cluster of MSMEs has a higher net worth and more employee ownership than MSMEs that join Cluster 2 and Cluster 1, but Cluster 3 has the fewest business permits. The second cluster is characterized by MSMEs having a high net worth (although not as high as Cluster 3) but ownership of business licenses and sales volume were low.

From the explanation above, 457 MSMEs joined in Cluster 1 business (5% of total MSMEs), characterized by ownership of product certification, product license ownership, and brand ownership with the lowest sales volume. MSMEs that joined Cluster 2 included 7,535 business units (84% of the total UMKM) characterized by a high net worth; however, ownership of business licenses and business volume were low. While MSMEs showed joining in Cluster 3, as many as 927 business units (10% of the total MSMEs) were characterized by ownership of the highest net worth among other clusters but low business license ownership.

The MSMEs incorporated in Cluster 1 already had an awareness of the importance of licensing, including product permits (such as halal certificates and safety and food certificates). However, this proves that MSME awareness of product licensing could be much higher. It was proven that the number of incorporated business units were only 5% of the total MSMEs. Of the nine attributes studied, business volume had the lowest position. Thus, the policy or strategy carried out in Cluster 1 must be related to handling the problem of low sales volume. Policies or strategies that can be carried out by SMEs or cluster managers include a promotion, price policy, innovation, improving product quality,
and others (Suryanarayana, 2023). The government, as a regulator, can help MSMEs in Cluster 1 with promotions and product marketing policies, coaching and training facilitation for quality improvement products, and others.

Cluster 2 has the most members with 84% of the total MSMEs. Cluster 2 is characterized by a high net worth but low ownership of business licenses and low business volume. One common characteristic of MSMEs in Indonesia is that they often do not have business licenses. That was illustrated in MSMEs that are members of Cluster 2. Strategies that The government could carry out strategies regarding the large number of MSMEs that do not have business licenses by authorizing a gift incentive. The government, could offer incentives to MSMEs incorporated in a cluster through grants, ease of access to funding, tax holidays, and other promotions. In addition, the cluster manager could perform coaching for MSME members to improve awareness of the importance of having a business license. Another strategy that could be applied to Cluster 2 is net worth optimization. Even though the net worth of the companies are high, this is accompanied by low business volume, leading to the possibility that the wealth may be “idle” or not used for business. Idle wealth, for example, might include ownership of a warehouse or a machine that is not optimally used, funds (either cash or savings) that are not used for business, and others. UMKM, with other members in the cluster, could conduct training on business financial management.

Cluster 3 is characterized by the highest net worth ownership among other clusters and low business license ownership. The number of business units joins as much as 10% of the total MSMEs. This condition describes all of the MSMEs in Indonesia, most of which experience difficulties in capital and low awareness of business legality. Besides raising awareness of the importance of business legality, another strategy can be carried out in the same way as Cluster 2, namely the provision of government incentives.

The government and cluster management must actively remind MSMEs of the importance of business legality. The results show that this is an attribute that needs to be completed by MSMEs. Lack of awareness and high costs are reasons many businesses to not apply for licenses. The absence of business permits (legality) is one of the reasons for the low number of MSMEs accessing funding from official financing institutions such as banks (Konstantynova, 2019) as well as the low amount of MSME products to enter the international market (exports). The low sales volume of MSMEs is caused by the poor quality of human resources. The low quality of human resources in SMEs in Indonesia impacts MSMEs' slow response to the market. Low innovation in MSMEs affects the competitiveness and overall performance of MSMEs (Kumar et al., 2020).

5. DISCUSSION

MSMEs have a fairly large role in Indonesia by providing a sizeable contribution to the economy and absorbing labor. Structurally, the number of MSMEs in Indonesia is the largest compared to medium and large businesses. Anton et al. (2015) identified five characteristics of MSME in Indonesia that make a business important for the country's economic development. First, MSMEs in Indonesia are owned by local people and employ millions of domestic workers. Second, SMEs are very common in rural areas, and their business is based on agriculture, so they are important for rural economic development. Third, SMEs are labor-intensive, with many young and low-educated people involved in the business. Fourth, SMEs in Indonesia obtain their finances from personal savings. Fifth,
this business produces simple consumer goods. They cater to the domestic market and target low-income consumers.

Other researchers explained that the characteristics of MSMEs in Indonesia, in general, include the following. The quality of human resources is low, and most businesses do not have licenses or meet other legal requirements. The amount of wealth and capital is limited, the number of workers employed are limited, and the type of product is not fixed. Often there is not a permanent place of business, financial management is not performed, they do not have access to banking, and most do not have a separation between personal and corporate wealth (Jaswadi et al., 2015; Fitriati et al., 2020).

In general, the problems of MSMEs in Indonesia are limited funds sources, low-quality human resources, and others factors that impact performance. Policymakers must map out a business strategy. One strategy that can be implemented is the formation of clusters (Goerzen, 2018). The formation of clusters is capable of improving MSME business performance by providing supply chains from upstream to downstream and encouraging MSME competitiveness (Saha et al., 2018).

Therefore, joining MSMEs with clusters can accelerate innovation (Rojas-Lema et al., 2021). For example, an innovation that can be done in Cluster 1 is product quality improvement to encourage improvement in sales. Clusters encourage a collaborative atmosphere with increased synergy between the government, cluster managers, and MSMEs (Villa & Taurino, 2018) 2018). With regular collaboration among stakeholders, it is possible to facilitate the process of fostering MSMEs in various business operations, such as access to financial institutions, and obtaining business permits (Yigit Ozkan et al., 2020).

Some of the advantages of MSMEs joining clusters include that clusters create synergies, enabling the involvement of various resources from other sectors, such as education, science, finances, agriculture, defense, and health (Goerzen, 2018). In addition, the development of SMEs in clusters can be done through cooperation with large companies through contracts, deepening specialization, developing production niches, and expanding access to financial resources. The joining of MSMEs in clusters also increases innovation in production due to the intensification of the company's cooperation with research organizations, and this, in turn, contributes to growth (Batz et al., 2015). Developing innovation management within the company through introducing new products and cooperation will improve quality and competitiveness. A better understanding of suppliers’ needs and requirements and consumers' quality and technical characteristics of products and components will lead to significant time savings (Cicea et al., 2019).

The development of MSMEs in clusters will eventually improve regional and state investment attractiveness (Kaplinsky & Morris, 2019). It will create a higher level of investor confidence in the network structure rather than in individual companies. Clusters also speed up the pace of creation and new business development. New companies have opportunities to grow under more favorable conditions because of their relationship with cluster partners. On a more macro basis, clusters can create closed production cycles on domestic value chains through the identification and creation of eye groups or missing links in production, which contribute to the development of substitute imports (Shiposha, 2020).

6. CONCLUSION
The research findings show that MSMEs in Garut Regency can be formed into three clusters. Each cluster combines MSMEs with attributes that are close to each other. The formed MSME clusters have implications for innovation, collaboration, and synergy among stakeholders, which impact MSMEs' value creation and performance. Follow-up findings show that one of the most dominant characteristics of MSMEs is the low awareness of MSMEs in owning business permits. These findings reinforce the findings of previous studies. The practical implications of these findings are that the government, as a regulator, must facilitate and make it convenient to encourage MSMEs to obtain business licenses.

REFERENCES


