

Qahtani, Haitham Al; Sankar, Jayendira P.

Article

The cluster analysis in the aluminium industry with K-means method: an application for Bahrain

Cogent Business & Management

Provided in Cooperation with:

Taylor & Francis Group

Suggested Citation: Qahtani, Haitham Al; Sankar, Jayendira P. (2024) : The cluster analysis in the aluminium industry with K-means method: an application for Bahrain, Cogent Business & Management, ISSN 2331-1975, Taylor & Francis, Abingdon, Vol. 11, Iss. 1, pp. 1-19, <https://doi.org/10.1080/23311975.2024.2361475>

This Version is available at:

<https://hdl.handle.net/10419/326310>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>

The cluster analysis in the aluminium industry with K-means method: an application for Bahrain

Haitham Al Qahtani & Jayendira P. Sankar

To cite this article: Haitham Al Qahtani & Jayendira P. Sankar (2024) The cluster analysis in the aluminium industry with K-means method: an application for Bahrain, Cogent Business & Management, 11:1, 2361475, DOI: [10.1080/23311975.2024.2361475](https://doi.org/10.1080/23311975.2024.2361475)

To link to this article: <https://doi.org/10.1080/23311975.2024.2361475>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 04 Jun 2024.



Submit your article to this journal [↗](#)



Article views: 953



View related articles [↗](#)





View Crossmark data [↗](#)



Citing articles: 1 View citing articles [↗](#)

The cluster analysis in the aluminium industry with K-means method: an application for Bahrain

Haitham Al Qahtani^a  and Jayendira P. Sankar^b 

^aUniversity of Technology Bahrain, Salmabad, Bahrain; ^bCollege of Administrative and Financial Sciences, University of Technology Bahrain, Salmabad, Bahrain

ABSTRACT

This study examines the utilization of the K-means clustering method to analyze Bahrain's aluminum industry. In addition, this study emphasizes the importance of clustering in understanding productivity, quality, and competitiveness within the sector. Data collection involved rigorous cleaning of diverse sources to ensure accuracy. By employing the K-means algorithm, this study successfully identified distinct clusters within the dataset, offering insights into industry dynamics. In addition, it proposes a roadmap for cluster development, providing actionable recommendations for stakeholders to enhance competitiveness and sustainability. Overall, this research advances knowledge of clustering techniques and informs strategic decision-making in Bahrain's aluminum industry.

ARTICLE HISTORY

Received 12 March 2024
Revised 1 May 2024
Accepted 22 May 2024

KEYWORDS

Bahrain aluminium industry; K-means cluster analysis; gap analysis; assessment of linkages; road map; positioning

REVIEWING EDITOR



Diego Corrales-Garay,
Universidad Rey Juan
Carlos, Spain

SUBJECTS

Business, Management,
and Accounting;
Production, Operations,
and Information
Management;
Management of
Technology and
Innovation; Development
Studies; Economics and
Development; Industry
and Industrial Studies

1. Introduction

Clustering analysis has become very important through the influential work of Porter across ten different countries (Kim et al., 2023). In addition, Kalicanin and Gavric (2014) mentioned that clustering aims to group objectives based on factors influencing competitive advantage. Productivity is a crucial factor in any industry, with the optimized use of resources, such as labor, materials, and machines to generate goods and services effectively (Jain et al., 2016; Prakash et al., 2017). Moreover, in the context of the aluminum industry in Bahrain, it is one of the key drivers of productivity and economic growth. Importantly, clustering analysis studies in the aluminum industry in Bahrain will gain prominence in understanding the dynamics that affect a nation's competitive strength. Furthermore, clustering involves grouping firms, innovation centers, and information hubs to foster collaboration and explore novel market strategies, methods, and products (Ye et al., 2021; Zhao et al., 2023). Therefore, this study investigates the impact of clustering within Bahrain's aluminum industry and explores its implications on productivity and product quality.

CONTACT Jayendira P. Sankar  jpsankar@utb.edu.bh  College of Administrative and Financial Sciences, University of Technology Bahrain, Salmabad, Bahrain.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Traditional manufacturing strategies involve multiple machines, and extensive labor leads to tooling expenditure, material handling expenditure, space requirements, and high production costs (Heragu & Ekren, 2009; Pereira et al., 2019). In contrast, owing to heavy competition and rapid technological growth, manufacturing industries must adopt innovative strategies to control costs and increase overall productivity (Huang, 2023; Hwang & Kim, 2022). Therefore, the central goal of this study is to use an analysis-based approach to minimize waste and enhance production quality and value addition. In addition, this study involved a meticulous examination of the current manufacturing process, identifying the challenges, and implementing strategies for streamline operations. Moreover, this research conducted a study in the aluminum industry in Bahrain, focusing on producing high-quality outputs.

Predominantly, the context of the aluminum industry in Bahrain emphasizes the relevance of clustering studies in collaborative environments and fostering economic connections. Furthermore, clustering in the aluminum industry involves grouping firms that produce similar products and share inputs, such as labor and technology, horizontally and vertically (Madsen et al., 2003; Yang & Gu, 2021). Moreover, the critical characteristics of clusters in Bahrain's aluminum industry include diverse participants, geographical proximity, and economic interdependence regarding economic status and activities. Regarding the cluster models, Karaca (2018) and Vargas-Hernández et al. (2020) state that the pure cluster model emphasizes the benefits of geographical proximity, the industrial complex model identifies stable relationships based on commercial ties, and the social network model focuses on economic activities and social integration of roles within networks and institutions.

Numerous studies have utilized clustering techniques across diverse fields, showcasing their effectiveness and versatility. Despite variations in objectives and datasets, clustering principles remain consistent in education (Liu, 2022), healthcare (Yang et al., 2023), image processing (Mittal et al., 2022), social network analysis (Sharma, 2011), financial services (Cook & Pandit, 2012), environmental science (Koo et al., 2023), manufacturing (Unterberger & Müller, 2021), transportation and logistics (Rivera et al., 2016), agriculture (Dureti et al., 2023), and so on have all benefited from clustering in uncovering insights and aiding decision-making. Interestingly, studies across different domains often adopt similar clustering approaches. Holý et al. (2017) used hierarchical clustering in retail to analyze consumer preferences, while Schulz et al. (2020) employed k-means clustering in biomedical research for disease subtype classification. Despite varied contexts, these studies share a common thread in leveraging clustering for data exploration and pattern recognition. Further, this suggests potential for clustering techniques in indices and policy evaluation, where they can extract insights from complex datasets encompassing economic, social, and policy-related variables. Therefore, this K-mean clustering application framework study represents a novel approach to enhancing decision-making and policy formulation in these domains.

The study of the clusters in the Bahrain aluminum industry offers economic advantages, such as supply chain cost, reducing transportation, creating a large market at a lower price, enhancing access to a qualified workforce, promoting cooperation, encouraging competition, and fostering specialization. Various clustering analysis methods have been proposed. The location quotient (LQ) method, input-output approaches, system dynamics approaches, dynamic models, and the Porter Diamond Model have been utilized in clustering studies to assess economic development, competitiveness, and competition between clusters. Clustering analysis of Bahrain's aluminum industry is a multifaceted exploration of economic interdependence, geographical proximity, and collaborative advantages. This study aims to provide insights into how clustering can contribute to increased productivity, enhanced product quality, and overall competitiveness in the dynamic landscape of the aluminum industry in Bahrain.

First, it aims to address the theoretical gap in clustering the aluminum industry in Bahrain using the K-means method, which has not been done before in any part of the research. Second, there is a methodological gap in using the K-means method for specific clustering of the aluminum industry in Bahrain and its rationale. Third, the practical gap can be filled with practical recommendations for policymakers to implement clustering initiatives effectively in Bahrain's aluminum industry. Fourth, there is a contextual gap in no earlier studies within Bahrain's broader economic development goals, policy frameworks, or industry dynamics. The remainder of this paper is organized as follows. [Section 2](#) reviews the relevant literature. [Section 3](#) describes the research methodology and [Section 4](#) presents the results and findings.

Section 5 presents the discussion and implications of this study. Finally, Section 6 concludes the paper and presents the limitations and avenues for future research.

2. Review of literature

Clustering is a holistic approach that emphasizes collaboration, shared resources, and interconnectedness to enhance competitiveness (Pereira et al., 2023; Zhao et al., 2023). The characteristics of industrial clusters include physical closeness, value creation through interconnected production, and a shared business environment (Kim et al., 2023). Therefore, the dynamic interplay between cooperation and competition is pivotal for fostering innovation within industrial clusters (Bagherzadeh et al., 2022; König, 2023). The literature review was a thorough review of the literature between 2001 and 2023. To ensure the reliability of the literature, all articles were retrieved from the Scopus database.

A study of Brazilian industrial clustering on intellectual property (IP) usage in manufacturing organizations in developing countries highlighted the importance of footwear firms in combating counterfeit and pirated goods threats and the growing importance of IP protection in global competitiveness (Cavalheiro & Brandao, 2017). Kadokawa (2013) revealed that the location decision of new manufacturing plants with advantages in Japan's industrial clustering is a key location factor, particularly in high-tech industries. Similarly, Libaers and Meyer (2011) point out the impact of industrial clustering on small technology-based firms, with distinctions between serial and non-serial innovators, which have more effective implications for economic development.

Guo and Guo (2011) exposed Chinese industrial clusters through knowledge-spanning mechanisms that synthesize learning opportunities and absorptive capacity perspectives to understand technological learning behavior across cluster types, cognitive subgroups, and innovation types. Chen (2011) revealed the competitive advantage of production systems, such as Taiwan's industrial clustering, highlighting the role of lead firms' relational capabilities fostered by cluster embeddedness in effectively governing suppliers. Additionally, Park et al. (2009) mentioned that inter-industrial knowledge flows are crucial for developing knowledge clusters and a national innovation system of Korean industrial clusters using two types of knowledge flows, embodied and disembodied.

Clustering in the British broadcasting industry's impact on firm growth and entry survival rates has significant positive clustering on firm growth and entry survival areas akin to high-tech industries, such as computing and biotechnology (Cook et al., 2001). Likewise, the connection between emotions and product shapes in automobiles, sofas, and kettles uses four fundamental dimensions: trend, emotion, complexity, potency, and significant shape features (SSFs) (Hsiao & Chen, 2006). In addition, the success of spinoffs in the global fashion design industry, their role in clustering design firms in selected cities, and insignificant effects from migration flows and localization economies (Wenting, 2008). Correspondingly, to identify the sub-sectors of Turkish textile patents, the study of patent count data and fuzzy-based clustering to detect technology trends, such as classic or dated, through k-means clustering (Dereli & Durmuşoğlu, 2009).

A study on timely sales identification represents the hybrid model by merging the fuzzy neural network and k-means cluster, which increases the forecast accuracy in small regions. In addition, this study reiterates that the hybrid model is more effective than the alternative (Chang et al., 2009). In addition, the risk identification and perception in fisheries systems in Faroes, Iceland, Greece, and the UK through clustering highlighted the supporting social science theories and the subjective nature of risk (Tingley et al., 2010). Another thing to remember is that Clustering occupational accidents in the Italian wood-processing industry revealed that K-means clustering identifies the standard sequence of events leading to accidents and devises preventive measures (Palamara et al., 2011). Similarly, a case-mix project in the Netherlands on healthcare products for hospitals clustered the product structure with cost homogeneity (Westerdijk et al., 2012).

A study by Jun and Park (2013) on Apple's technological innovation by analyzing patent applications using clustering of patents identifies vacant technology domains and social network analysis for the future. Furthermore, clustering regional businesses in the northern Spanish region for principle components analysis provided superior cluster solutions (Argüelles et al., 2014). Boix et al. (2015) studied the spatial patterns of the location and co-location of clusters in 16 European countries and revealed a

highly clustered creative belt from southern England and southern Germany. Similarly, it brings out innovation from multiple sectors in Korean industrial convergence by clustering technology-driven new value generators, service-integrated social business generators, policy-driven environmental enhancers, and technology enhancers (Geum et al., 2016).

Zhou et al. (2017) showed insightful clustering of monthly household electricity consumption patterns in Jiangsu province in China using a fuzzy c-mean clustering model for energy efficiency strategies in the power industry. Fuzzy c-mean clustering was used to evaluate supplier sustainability in the plastic pipe industry; it offers managerial insights into outsourcing decisions and resilience strategies to minimize costs and maximize sustainability during disruptions (Jabbarzadeh et al., 2018). Additionally, the combination of the length, recency, frequency, monetary, and periodicity (LRFMP) model and K-means clustering in Iranian fintech companies is supported by marketing intelligence and strategic decision-making in the business-to-business (B2B) fintech industry (Sheikh et al., 2019). Likewise, using latent Dirichlet allocation and K-means clustering for university selection and visualizing cooperation networks on blockchain has brought significant challenges, such as selecting research teams, evaluating university competitors, and accessing technology (Ran et al., 2020).

In connection with the Fourth Industrial Revolution, finding the relationship between green economics and digital technology development in Ukraine through clustering analysis revealed the alignment of digitalization efforts with national innovation strategies and mitigating environmental risks (Plantec et al., 2021). Specifically, Edina et al. (2021) revealed Industry 4.0 investments and their impact on Hungarian food companies' business performance using K-means clustering to find process innovation problems in the increasing use of Industry 4.0 tools. In addition, a study on the influential factors affecting China's electricity industry emissions (CEEI) using K-means clustering reveals the aiding CEEI control and promotion of low-carbon transformation (He et al., 2022). Correspondingly, industrial clusters in China's new energy enterprises revealed enhanced profitability by reducing production inefficiency and new energy industry development (Sun et al., 2022).

Clustering is vital in advancing sustainable development goals in healthcare, energy, transportation, and manufacturing (Oyewole & Thopil, 2023). In addition, a study of China's low-carbon development of the power sector for neutrality goals and carbon peaking using K-means clustering revealed emission reduction by aiding provincial-level emission reduction strategies for global green transformation (Wang et al., 2023). Another thing to remember is a study by Singhal et al. (2023), which categorizes online fashion consumers based on their perceptions and relationships with brands through social media using K-means cluster analysis, emphasizing a pioneering approach to understanding consumer behavior in online fashion.

Table 1 represents the overview of reviewed sources, including author with year, focus of study, methodology, and key findings. Clustering emphasizes shared resources and collaboration, which fosters competitiveness through interconnected production and leveraging physical proximity in industrial clusters. Additionally, the interplay between competition and cooperation drives innovation. Further, literature reviews spanning 2001–2023, notably on Japanese manufacturing and Brazilian IP usage, underscore the impact of clustering on global competitiveness. Moreover, studies on Taiwanese and Chinese clusters highlight knowledge exchange roles, whereas British broadcasting and fashion design showcase diverse applications. In addition, techniques, such as fuzzy clustering enhance energy efficiency and supplier sustainability analysis. Furthermore, in the context of Industry 4.0, clustering aids in emission control and aligns with sustainability goals. Crucially, clustering advances sectors, such as low-carbon transitions, shaping consumer behavior, energy, and healthcare. Thus, no specific literature on clustering analysis in the aluminum industry using K-means in Bahrain shows a literature gap. Therefore, the purpose of this study is to fill the literature gap and provide solid recommendations for the effective clustering of the aluminum industrial sector in Bahrain.

3. Methodology

Many studies have used a methodology similar to that used in this study, but these data mining methods are still the most commonly used in the natural sciences. Furthermore, clustering algorithms were

Table 1. Overview of reviewed sources.

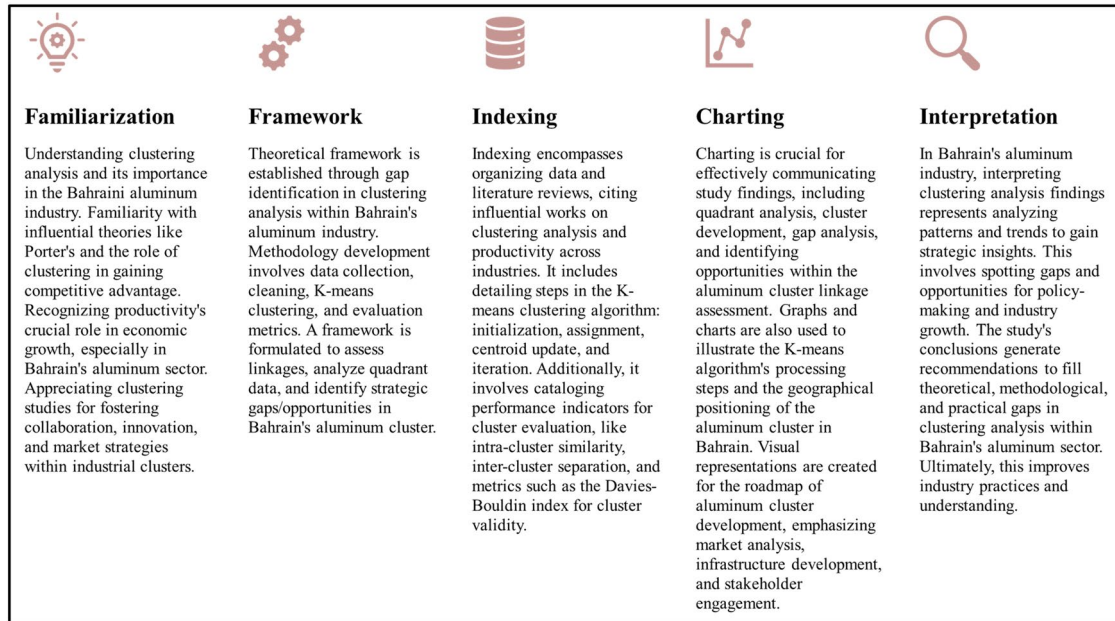
Author and year	Focus of study	Methodology	Key findings
Argüelles et al. (2014)	Clustering regional businesses in northern Spain	Principle components analysis	Superior cluster solutions for regional businesses
Boix et al. (2015)	Spatial patterns of clusters in European countries	Geographic analysis	Identification of highly clustered creative belt
Cavalheiro and Brandao (2017)	IP usage in Brazilian manufacturing organizations	Qualitative analysis	Importance of IP protection in combating counterfeit goods, enhancing global competitiveness
Chang et al. (2009)	Hybrid model for timely sales identification	Integration of fuzzy neural network and k-means cluster	Increased forecast accuracy in small regions
Chen (2011)	Competitive advantage of Taiwan's industrial clustering	Case study analysis	Role of lead firms' relational capabilities in governing suppliers
Cook et al. (2001)	Impact of clustering on firm growth in British broadcasting industry	Statistical analysis	Significant positive clustering effects on firm growth and entry survival rates
Dereli and Durmuşoğlu (2009)	Identification of sub-sectors of Turkish textile patents	Fuzzy-based clustering	Detection of technology trends through k-means clustering
Edina et al. (2021)	Impact of Industry 4.0 investments on Hungarian food companies	K-means clustering	Identification of process innovation problems
Geum et al. (2016)	Innovation in Korean industrial convergence	Clustering of technology-driven value generators	Identification of innovation sectors in convergence
Guo and Guo (2011)	Knowledge-spanning mechanisms in Chinese industrial clusters	Synthesis of learning opportunities	Understanding technological learning behavior across cluster types
He et al. (2022)	Factors affecting China's electricity industry emissions	K-means clustering	Aid in emission control and low-carbon transformation
Hsiao and Chen (2006)	Connection between emotions and product shapes in various industries	Conceptual analysis	Identifying fundamental dimensions and shape features
Jabbarzadeh et al. (2018)	Evaluation of supplier sustainability in plastic pipe industry	Fuzzy c-mean clustering	Managerial insights for outsourcing decisions and resilience strategies
Jun and Park (2013)	Analysis of Apple's technological innovation using patent clustering	Patent analysis and social network analysis	Identification of vacant technology domains and future trends
Kadokawa (2013)	Location decision of new manufacturing plants in Japan's industrial clusters	Case study analysis	Key location factor, particularly in high-tech industries
Libaers and Meyer (2011)	Impact of industrial clustering on small technology-based firms	Literature review	Differential implications for economic development between serial and non-serial innovators
Oyewole and Thopil (2023)	Role of clustering in advancing sustainable development goals	Conceptual analysis	Contribution to sustainability in various sectors
Palamara et al. (2011)	Clustering occupational accidents in Italian wood-processing industry	K-means clustering	Identifying standard sequence of events leading to accidents
Park et al. (2009)	Inter-industrial knowledge flows in Korean industrial clusters	Network analysis	Crucial for developing knowledge clusters and national innovation systems
Plantec et al. (2021)	Relationship between green economics and digital technology in Ukraine	Clustering analysis	Alignment with national innovation strategies and environmental risk mitigation
Ran et al. (2020)	Clustering for university selection and cooperation networks on blockchain	Latent Dirichlet allocation and K-means clustering	Challenges in research team selection and technology access
Sheikh et al. (2019)	Clustering of Iranian fintech companies	LRFP model and K-means clustering	Support for strategic decision-making in B2B fintech industry
Singhal et al. (2023)	Categorization of online fashion consumers based on social media perceptions	K-means cluster analysis	Understanding consumer behavior in online fashion
Sun et al. (2022)	Profitability enhancement in China's new energy enterprises	Statistical analysis	Reduction of production inefficiency and industry development
Tingley et al. (2010)	Risk identification in fisheries systems through clustering	Qualitative analysis	Highlighting subjective nature of risk and supporting social science theories
Wang et al. (2023)	Low-carbon development of China's power sector	K-means clustering	Support for provincial-level emission reduction strategies
Wenting (2008)	Role of spinoffs in clustering design firms in global fashion industry	Case study analysis	Insignificant effects from migration flows and localization economies

(Continued)

Table 1. Continued.

Author and year	Focus of study	Methodology	Key findings
Westerdijk et al. (2012)	Case-mix project in the Netherlands healthcare industry	Statistical analysis	Clustered product structure with cost homogeneity
Zhou et al. (2017)	Clustering of household electricity consumption patterns	Fuzzy c-mean clustering	Insights for energy efficiency strategies in the power industry

Source: Self-made by the author(s) based on the available data.

**Figure 1.** Five key steps of framework analysis.

Source: Self-made by the author(s) based on the available data.

used to group data with common properties. Clustering is a well-known method used in economics to ensure the correctness of clustering according to the number of employees and local units, and the K-means clustering method was used in the original data mining. Additionally, data relevant to the aluminum manufacturing industry in Bahrain on the geographical distribution of manufacturing units, workforce demographics, sales figures, and production quantities were collected. Specifically, this study utilized reliable and accurate data from databases, industry reports, and government publications. The data were then cleaned to remove and fix incomplete, duplicate, incorrectly formatted, corrupted, and incorrect data because they were collected from multiple sources. Furthermore, this study used the classification of statistical regional units in Bahrain to define the geographical region for analysis and to concentrate on aluminum manufacturing activities in Bahrain.

Moreover, the datasets were selected to highlight the uniqueness of Bahrain's aluminum manufacturing sector, such as geographical distribution, market demand, and production volume. In addition, the parameters were fixed for the number of clusters (K value), and the distance metric was used to measure the similarity between the data points. The following steps were followed to ensure the correct clustering of the aluminum manufacturing sector in Bahrain.

1. Bahrain statistics for each region published by the institutions and the number of local units in the sector were obtained.
2. An algorithm was developed for the k-means method (data-mining method).
3. The solution to the written algorithm determines the clusters to be set up.

The Davies-Bouldin Index (DBI) metric, introduced by Davies and Bouldin (1979), serves as a tool for evaluating clusters. It assesses the internal validity of clustering by analyzing how effectively it has been executed by computing various statistical features derived from the dataset. In essence, the DBI value should ideally approach zero or be close to non-negative values, indicating a higher quality of the obtained cluster and facilitating the judgment of its goodness.

$$d(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Figure 1 represents the five key steps of framework analysis with familiarization framework, indexing, charting, and interpretation. The study implemented strategies to optimize effectiveness and overcome potential limitations. Firstly, enhancing data collection methods is crucial, utilizing both traditional approaches and advanced data mining techniques to comprehensively analyze the aluminium manufacturing sector in Bahrain, focusing on aspects like geographical distribution, market demand, and production volume. Secondly, implementing differential privacy offers a rigorous approach to safeguarding data privacy while enabling organizations to analyze sensitive data. Additionally, ensuring the robustness of K-means clustering through methods, such as intra-cluster and inter-cluster similarity, cluster validation, and determining the optimal number of clusters (k) is essential for accurate partitioning of datasets. Moreover, mitigating bias in clustering analysis by employing techniques, such as K-means clustering, evaluation metrics, and Euclidean distance helps ensure the accurate partitioning of datasets into distinct clusters. By incorporating multiple clustering algorithms and validation techniques, potential limitations are mitigated, enhancing the reliability of study outcomes and facilitating informed decision-making processes based on valuable insights extracted from the data.

3.1. Data

Data was gathered using a combination of methodologies commonly employed in natural sciences research. Specifically, data mining methods, particularly clustering algorithms, were utilized to analyze and group data with shared characteristics. The primary focus was on the aluminum manufacturing industry in Bahrain. Further, the dataset encompasses various facets of the aluminum manufacturing sector, including but not limited to the geographical distribution of manufacturing units, workforce demographics, sales figures, and production quantities. In addition, these variables were chosen to characterize Bahrain's aluminum manufacturing landscape comprehensively. Also, the data was sourced from diverse outlets, including databases, industry reports, and government publications. These repositories ensured the reliability and accuracy of the data used in the study.

4. Clustering analysis and results

In this study, K-means clustering, proposed by MacQueen in 1967 (MacQueen, 1967), was employed to determine the effectiveness and efficiency of partitioning the datasets. In addition, the algorithm iteratively assigns each data point to one of k clusters based on the similarity of their attributes. This study continues until convergence when the clusters exhibit maximum intra-cluster and minimum inter-cluster similarities. Second, data mining is critical to modern data analysis to extract valuable insights from large datasets. Furthermore, clustering analysis facilitates the grouping of objects with characteristics similar to those of a fundamental data-mining technique. Hence, this study aims to identify patterns and relationships within datasets that provide valuable information for decision-making processes. Therefore, this study utilized the K-means clustering method as a nonhierarchical approach to partition the dataset into distinct clusters.

According to Kangalli et al. (2014), five methods are employed in stacker clustering: single link, full link, average link, ward method, and central method, all based on different principles. Tan et al. (2022) used the divisor hierarchical method to form clusters by dividing the large clusters into small clusters based on their similarity. In addition, Swarndepp and Pandya (2016), in partition-type clustering, such as K-means, iteratively adjust cluster centers until points are closest, forming distinct clusters based on the

minimum distance. Further, Nidhi and Patel (2016) stated that in density-based clustering, clusters are formed based on data density and distance based on higher-density areas as clusters in the database. Additionally, Du and Wu (2022) reported that in grid-based clustering, the performance depends on the frame count, and it is fast accessibility owing to the low processing load. Moreover, Gergely and Vargha (2021), in the model-based clustering mathematical model for data, assumed that its distribution aligns with specific probability theories in the data space.

First, to achieve optimal partitioning, the K-means algorithm successfully partitions the dataset into k clusters, where k is a user-defined parameter. Subsequently, the algorithm converges to an optimal solution by iteratively updating the cluster centroids and assigning data points to the nearest centroid. Second, to achieve intra-cluster similarity, the data points exhibited high homogeneity within each cluster, indicating that objects with similar characteristics were effectively grouped. In addition, this high intra-cluster similarity suggests that the clusters were internally cohesive and representative of distinct patterns within the data. Third, to achieve inter-cluster separation, the cluster demonstrated minimal overlap, indicating clear boundaries between the different groups of data points. Therefore, this separation is crucial for accurately distinguishing distinct patterns and associations in a dataset. Fourth, to achieve cluster validation, the validity of the obtained clusters was assessed using metrics, such as the Davies-Bouldin index. Additionally, these metrics evaluate the compactness and separation of clusters and provide quantitative measures of clustering quality.

4.1. Processing steps of K-means algorithms

Step 1: The initial cluster centers are established using two distinct methods. The first method involves selecting k random points from the objects to serve as initial cluster centers. Alternatively, the second method involves determining the center points by calculating the average of all objects.

Step 2: The distance to the center points of each selected object is computed. Based on the resulting distances, all objects are assigned to the cluster closest to them among k clusters.

Step 3: The new center points of the resulting clusters are updated to the average values of all objects within each cluster.

Step 4: Continuously iterate Steps 2 and 3 until there are no further changes in the center points.

According to Liberti et al. (2014) Euclidean distance formula:

$$p = (p_1, p_2, \dots, p_n) \text{ ve } q = (q_1, q_2, \dots, q_n) \quad (1)$$

$$\sqrt{\sum_i^n (p_i - q_i)^2} = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2}$$

At this stage, Ergun et al. (2020) stated that the main drawback of the K-means algorithm is its inability to automatically determine the optimal number of clusters (k). To address this issue, experimentation with various k values is essential for successful clustering. Equation (2) is suggested as a practical method to determine the appropriate number of clusters.

$$k = \sqrt{\frac{n}{2}} \quad (2)$$

In K-means clustering, which aims to partition data into k clusters, the most common evaluation criterion is the quadratic error criterion, also known as the sum of squared errors (SSE). This metric assesses clustering quality by measuring the sum of squared distances between each data point and its assigned cluster centroid (Yuan & Yang, 2019). A lower SSE value indicates a better clustering performance, with clusters more tightly packed around their centroids.

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} \text{dist}^2(m_i, x) \quad (3)$$

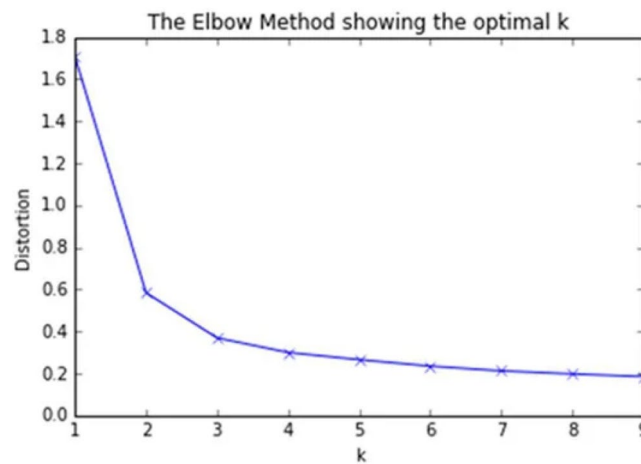


Figure 2. The elbow method showing the optimal K.
Source: Self-made by the author(s) based on the available data.

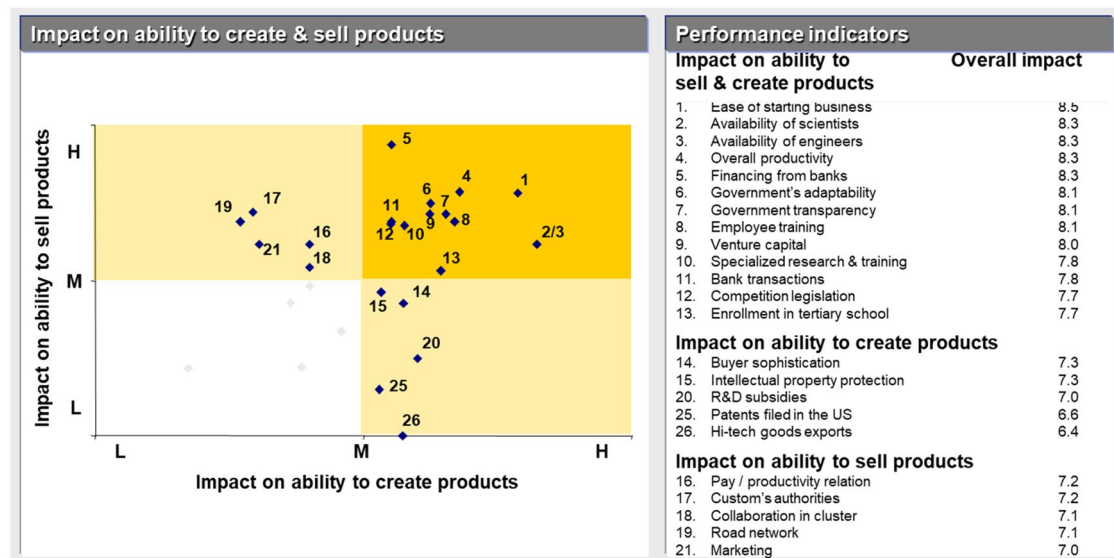


Figure 3. Aluminium cluster and assessment of linkages in Bahrain.
Source: Self-made by the author(s) based on the available data.

The optimal value of K can be more accurately determined by plotting the K-SSE curve and identifying the inflection point. As depicted in Figure 2, a distinct inflection point is clearly visible.

4.2. Aluminium cluster and assessment of linkages in Bahrain

Assessing the linkages within Bahrain's aluminum cluster involves examining the connections and interactions between the different entities involved in the country's production, distribution, and consumption of aluminum and aluminum products. In addition, these linkages can be analyzed from various perspectives, such as supply chain, industrial, value chain, market, policy and regulatory, technological, social, and economic linkages. Moreover, this analysis aims to inform strategic decisions related to investment, policy formulation, infrastructure development, and market positioning to foster sustainable growth and competitiveness of the cluster.

In Figure 3, all four quadrants represent the aluminum cluster and the assessment of linkages in Bahrain. Quadrant one, with a high impact on the ability to create products and high impact on the

ability to sell products, ranges between 8.5 and 7.7 overall impact (ease of starting business, availability of scientists, availability of engineers, overall productivity, financing from banks, governments adaptability, government transparency, employee training, venture capital, specialized research and training, bank transactions, competition legislation, and enrollment in tertiary school). quadrant two, with a high impact on the ability to sell products and a low impact on the ability to create products, ranges between 7.3 and 6.4 for overall impact (pay and productivity relation, customs authorities, collaboration in cluster, road network, and marketing). Quadrant three, with a high impact on the ability to create a product and a low impact on the ability to sell products, ranges between 7.3 and 7.0 overall impact (buyer sophistication, intellectual property protection, R&D subsidies, patents filed in the US, and high-tech goods exports). Hence, in quadrant four, no performance indicators have a low impact on the ability to create and sell products. Therefore, this cluster analysis in the aluminum industry using the K-means method represents a systematic approach to precisely identify Bahrain's most attractive economic development fields.

4.3. Aluminium cluster development in Bahrain

Bahrain's efforts in aluminum cluster development aim to attract investments, promote innovation, and create employment opportunities within the sector. In addition, the government has proactively supported businesses in the aluminum industry through various incentives, infrastructure development, and regulatory frameworks. Moreover, Bahrain benefits from its strategic location, well-established infrastructure, and skilled workforce, making it an attractive destination for companies looking to establish or expand their presence in the aluminum sector.

In Figure 4, all four quadrants represent aluminum cluster development in Bahrain. Quadrant one has a large gap and a high overall impact (productivity and intellectual property protection). quadrant two has a high overall impact and a low gap (buyer sophistication, pay, and productivity relations, employee training, customs authorities, and collaboration cluster). At the same time, in quadrant two, a few performance indicators slightly touch the middle (availability of engineers and marketing). quadrant three has a large gap size and a low overall impact (R&D subsidies, competition advantage, or unique products). Hence, quadrant four has no performance indicator for a low gap size and a low overall impact. Therefore, the cluster analysis in the aluminum industry with the K-means method represents the gap analysis criteria, including human resources (training engineers), encouragement (R&D subsidies and IP protection), and factors strengthening cluster integration (infrastructure) in Bahrain.

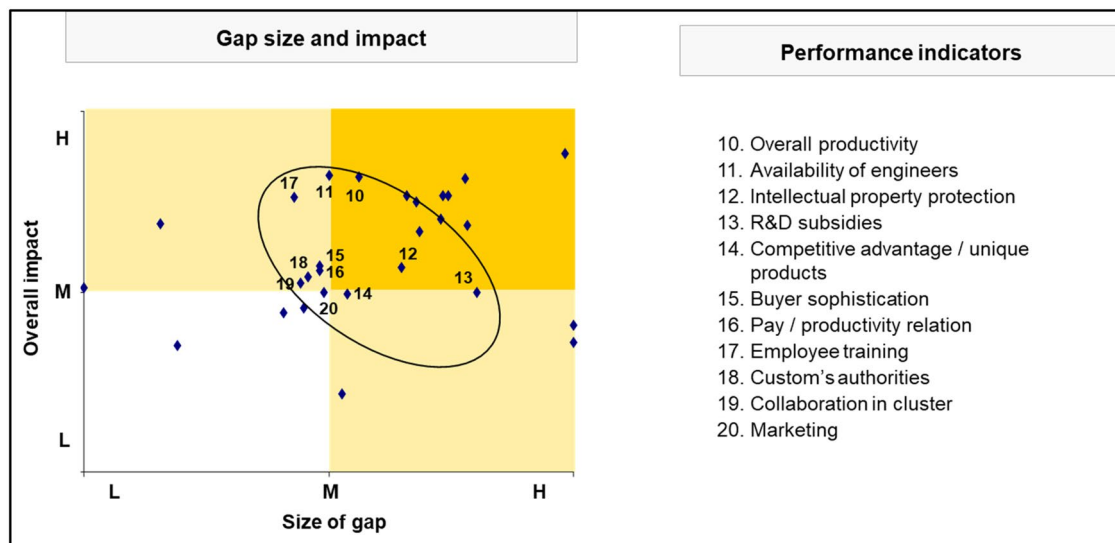


Figure 4. Aluminium cluster development in Bahrain.

Source: Self-made by the author(s) based on the available data.

Industry	Attract-iveness	Bahrain market	US market	Overall		Comments
Metal window and door mfg	●	○	○	○	▶	These two segments are large and growing both in GCC and the US
Fabricated structural metal mfg	●	○	○	○		
Ornament & architectural metal work	●	○	○	○		
Motor vehicle transmission & power train	●	◐	◐	◐		
All other motor vehicle parts mfg	●	◐	◐	◐		
Motor vehrical body parts	●	○	○	○	▶	GCC could be interesting, however the US imports are decreasing
Sheet metal work mfg	●	○	○	○		
Engine parts mfg	●	○	○	○		
Kitchen utilities, pot, and pan mfg	●	○	○	○		
Construction machinery mfg	●	○	○	○		
Other aircraft parts & auxiliary equ. mfg	●	◐	○	◐	▶	One of the most attractive areas for both the GCC and the US
Travel trailer & camper mfg	●	◐	◐	◐		
Prepress services mfg	●	○	○	○		
Farm machinery & equipment mfg	●	○	○	○		
Motor vehicle steering & suspension comp. mfg	●	◐	◐	◐		
Manufactured home (mobile home) mfg	●	○	○	○	▶	Fairly attractive in both GCC and the US
<div> <div>● High segment attractiveness</div> <div> <div>◐ Export market attractiveness</div> <div>◐ High</div> <div>◐ Medium</div> <div>○ Low</div> </div> </div>					<div>Conclusion</div> <div> <div>▶ Target</div> <div>▶ Fallback</div> <div>◐ Not feasible</div> </div>	

Figure 5. Identifying opportunities from gap analysis.
Source: Self-made by the author(s) based on the available data.

4.4. Identifying opportunities from gap analysis

Gap analysis is a valuable tool for identifying organizational improvements and growth opportunities. Further identifying opportunities from a gap analysis includes identifying performance gaps, prioritizing critical gaps, understanding causes, exploring the competitive landscape, customer feedback analysis, technological advancements, talent and skill development, market trends and changes, innovative solutions, and risk assessment. Therefore, organizations can uncover opportunities for growth, innovation, and competitive advantage by systematically analyzing performance gaps and taking proactive steps to address them.

In Figure 5, identifying opportunities through gap analysis shows a list of high-segment attractiveness industries. Further, the overall (Bahrain and US markets) high export market attractiveness in motor vehicle transmission and power train and all other motor vehicle parts manufacturing indicates that these two segments are large and growing both in Bahrain and US markets, which Bahrain needs to target. In addition, the export market in Bahrain is highly attractive. Still, the low export market attractiveness in the US market, which represents the medium export market attractiveness in other aircraft parts and auxiliary equipment manufacturing, which represents Bahrain, could be interesting. Still, US imports are decreasing, showing a fallback.

Additionally, both (Bahrain and the US markets) have high market attractiveness in travel trailer and camper manufacturing. This segment is one of the most attractive areas for Bahrain and the US market, which Bahrain needs to target. Finally, both (Bahrain and the US markets) have high market attractiveness in motor vehicle steering and suspension component manufacturing, representing that this segment fairly attracts both Bahrain and the US market, which Bahrain needs to target. Therefore, the cluster analysis in the aluminum industry with the K-means method represents opportunities in motor vehicle parts and travel trailer manufacturing, which are the most attractive areas in the aluminum value chain in Bahrain.

4.5. Road map for the aluminium cluster in Bahrain

Creating a roadmap for an aluminum cluster in Bahrain involves several key steps and considerations: market analysis, stakeholder engagement, infrastructure, development, investment promotion, regulatory framework, research and development, skill development, cluster networking, sustainability initiatives, marketing and promotion, and monitoring and evaluation. Hence, through these steps and actively involving relevant stakeholders, Bahrain can develop a thriving aluminum cluster that contributes significantly to the country's economic growth and industrial diversification.

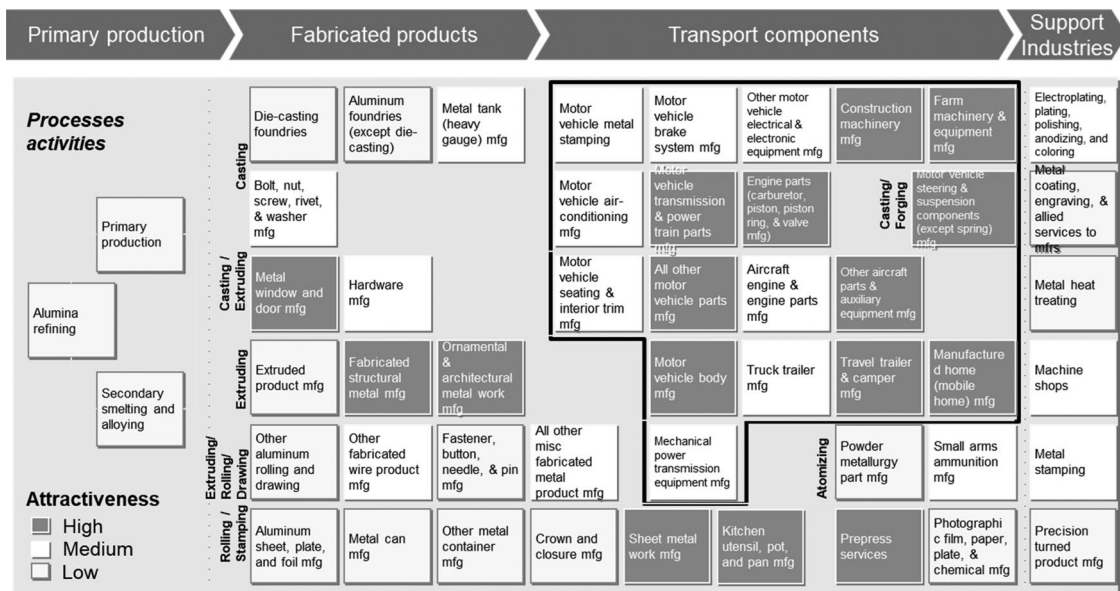


Figure 6. A road map for the aluminium cluster in Bahrain.

Source: Self-made by the author(s) based on the available data.

In Figure 6, they identify a road map for the aluminum cluster in Bahrain. Primary production, including processes and activities, is less attractive (primary production, alumina refining, secondary smelting, and alloying). In addition, fabricated products (casting, casting/forging, casting/extruding, extruding, extruding/rolling/drawing/automizing, and rolling/stamping) are highly attractive for (metal window and door manufacturing, fabricated structural metal manufacturing, ornamental and architectural metal work manufacturing, sheet metal work manufacturing, kitchen utensil, pot and pan manufacturing, and prepress services). Furthermore, transport components (casting, casting/forging, casting/extruding, extruding, and extruding/rolling/drawing/automizing) have high attractiveness in (construction machinery manufacturing, farm machinery and equipment manufacturing, motor vehicle transmission and power train parts manufacturing, engine parts (carburetor, piston, piston ring, and value manufacturing), motor vehicle steering and suspension components (except spring) manufacturing, all other vehicle parts manufacturing, other aircraft parts and auxiliary equipment manufacturing, motor vehicle body manufacturing, travel trailer and camper manufacturing, and manufacturing home (mobile home) manufacturing). Additionally, no specific industry has high attractiveness in the support industries (casting, casting/forging, casting/extruding, extruding, extruding/rolling/rolling/drawing/automizing, and rolling/stamping). Moreover, in the fabricated products, there were six industry categories with medium attractiveness and 11 with low attractiveness. In the transport component, there were eight industry categories with medium attractiveness and none with low attractiveness. In the support industries, there are three with medium attractiveness and three with low attractiveness.

4.6. Positioning of the cluster in Bahrain

Cluster-related businesses must properly combine a future product portfolio with market regions and value chain coverage.

Figure 7 shows the positions of the cluster in Bahrain by considering the key aspects and deducing the targeted positioning with coverage of the value chain (smelter, foundry, single parts, components, and modules), segments (automotive, machinery, and construction goods), and regions (TR/WE, Asia, GCC, and others). Further, the key aspects being considered, including profitability and key success factors, need to be understood in detail. Market and competitive intelligence require current match/gap to determine the feasibility of developing the opportunities and smart strategies to develop the desired positions to be developed. Moreover, the strategic positioning of new aluminum industries in

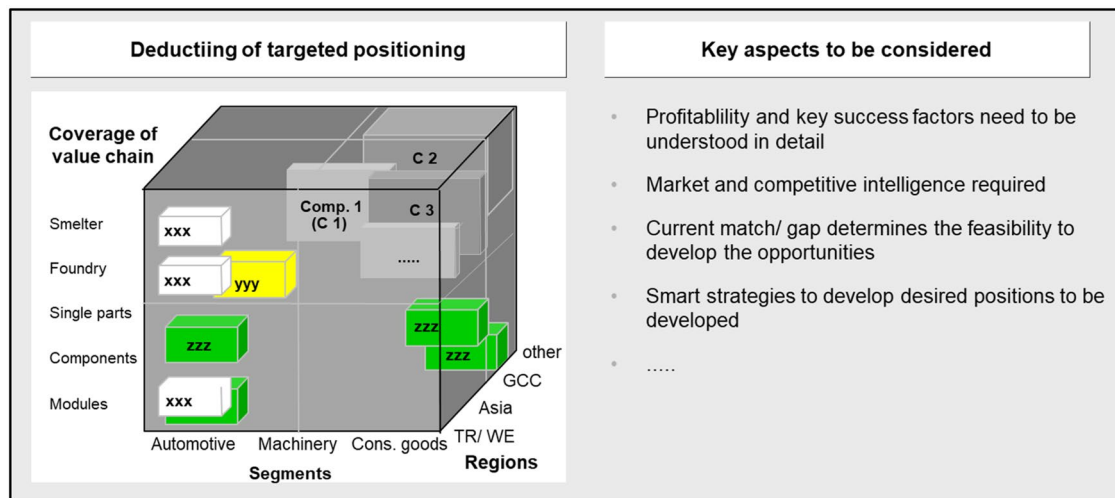


Figure 7. Positioning of the cluster in Bahrain.
 Source: Self-made by the author(s) based on the available data.

Bahrain should be fine-tuned with respect to value chain coverage, customer segments, and regional focus.

5. Discussion and implications

This study aims to perform cluster analysis in the aluminum industry using the K-means method in Bahrain. The results show the effectiveness and efficiency of K-means clustering, which achieves optimal partitioning with high intra-cluster similarity and minimal inter-cluster overlap and cluster validity using the Davies-Bouldin index scant with the study by Thakare and Bagal (2015). In addition, a DBI value of 0.918 indicates the clustering quality of a dataset. A lower DBI suggests better clustering, with more distinct clusters and denser packing within each cluster. Therefore, a DBI of 0.918 suggests relatively good clustering performance, though its interpretation may vary depending on context and the dataset. Furthermore, the study presents the processing steps of K-means algorithms (initialization, assignment, centroid update, and iteration), and the Euclidean distance was used to compute the similarity between data points and cluster centroids based on the study by Ikotun et al. (2023). Likewise, this study explored the number of clusters (K) and suggested a practical method for determining K, which is consistent with the study by Yuan and Yang (2019).

Performance indicators are decided based on selecting metrics that effectively evaluate those aspects. Further, several performance indicators, such as the effectiveness and efficiency of partitioning, optimization of the k-mean algorithm, assessment of linkages in Bahrain's aluminium cluster, cluster development and opportunities, roadmap for cluster development, and cluster positioning, were used. Also, performance indicators are based on objectives and aspects being evaluated in the study. Additionally, the study discovered the aluminum cluster and linkage assessment in Bahrain from various perspectives, such as supply chain, market, and policy, and quadrant analysis identified areas with a high impact on product creation and sales guiding strategic decisions, scant the study by Hong et al. (2014). Likewise, the study revealed the aluminum cluster development in Bahrain and its efforts to attract investment, promote innovation, and create employment opportunities in the aluminum sector (Hadhri et al., 2016), and quadrant analysis highlighted gaps in factors, such as R&D subsidies and IP protection for strategic development scant the study by Novillo-Villegas et al. (2022). The study also identified opportunities from gap analysis in high-segment attractiveness industries, such as motor vehicle parts and travel trailer manufacturing with market attractiveness for potential growth areas that scant the study by Hanawalt and Rouse (2017).

In addition, the study spotted a road map of aluminum cluster development with key steps, including market analysis, stakeholder engagement, infrastructure development, and skills enhancement

through a systematic approach involving stakeholders that can lead to the growth and competitiveness of Bahrain's aluminum cluster, which scant the study by Prebanić and Vukomanović (2023). Further, the study displayed the positioning of the aluminum cluster in Bahrain, including future product portfolios, market regions, and value chain coverage, including strategic positioning aligned with profitability, market intelligence, and regional focus, which scant the study by OECD (2019). This study provides insights into applying clustering techniques, particularly K-means, in various contexts, aiding decision-making processes and strategic planning for cluster development in Bahrain's aluminum industry.

Based on empirical and detailed research, this study proposes countermeasures for cluster analysis in the aluminum industry using the K-means method in Bahrain. In addition, rich resources and complex network structures in the cluster area can provide better resources and information for enterprises. First, to address the theoretical gap, this study provides actionable insights and recommendations that will inform the strategic decision-making process, promote innovation, and drive the sustainable growth of the aluminum sector in Bahrain. Hence, the feasibility of comprehensive research on Bahrain's aluminum sector is vital for identifying growth opportunities and fostering innovation by analyzing market trends, competitors, and regulations to pinpoint areas for innovation and growth, considering emerging technologies and sustainability. Therefore, these tailored recommendations will promote sustainable growth and align with industry objectives, involving stakeholder collaboration. Presenting findings and recommendations will facilitate collective strategy development. Continuous monitoring and adaptation will drive growth and innovation in Bahrain's aluminum sector.

Second, data collection involved gathering and clearing relevant metrics to address the methodological gap through clustering analysis, which was performed using a developed K-means algorithm evaluated through intra-cluster similarity and the Davies-Bouldin index. Third, the study addressed the practical gap by analyzing factors, such as manufacturing distribution and sales, and aimed to offer insights for industry enhancement that will guide strategic decisions and support Bahrain's economic growth. Hence, combining sustainability and technology is reshaping Bahrain's industries, driven by global initiatives like the SDGs through integrating eco-conscious practices to fulfill social responsibilities and bring cost savings and competitive advantages by technologies, such as AI, IoT, blockchain, and 3D printing to enhance efficiency and customer experiences, offering significant competitive edges when combined with sustainable principles. Also, embracing sustainability and technology can spur economic growth, attract investment, foster entrepreneurship, and create jobs, necessitating government policies and public-private partnerships for a prosperous future.

Fourth, this study addressed the contextual gap by applying K-means clustering to understand Bahrain's aluminum manufacturing sector by determining optimal cluster numbers, evaluating clustering quality, and examining linkages within the industry, aiming to inform strategic decision-making and investments. Finally, the study addressed the literature gap by exploring data mining methodologies, mainly clustering algorithms, in analyzing Bahrain's aluminum sector. It also covered data collection, cleaning, K-means clustering, algorithm steps, and evaluation metrics, and assessed the industry cluster, proposing a development roadmap for strategic positioning and growth opportunities.

6. Conclusion

This study thoroughly examines clustering analysis in Bahrain's aluminum industry, utilizing the K-means method to address gaps in theory, methodology, practice, context, and literature. Furthermore, the study emphasizes the significance of clustering in boosting productivity, product quality, and competitiveness within the sector and fostering collaboration, innovation, and economic connections. Additionally, the study achieved a DBI value of 0.918, indicating the clustering quality of the dataset. This lower DBI suggests better clustering, with more distinct clusters and denser packing within each cluster. Also, by applying the K-means algorithm, the study successfully identified distinct clusters within the dataset, offering insights into the industry's structure and dynamics and suggesting growth and strategic development opportunities. Also, the study proposes a roadmap for cluster development, providing actionable recommendations for policymakers, industry stakeholders, and researchers to

enhance competitiveness and sustainability. Overall, it contributes to existing knowledge by applying advanced clustering techniques and offers valuable insights for shaping strategic decisions and fostering sustainable development in Bahrain's aluminum sector amid evolving technological and market landscapes.

6.1. Limitations and future research

The study on Bahrain's aluminium industry presents limitations and avenues for future research. First, this study's focus on Bahrain's aluminium industry restricts the generalizability of findings to broader regional or sectoral contexts. This limitation highlights the need for further research to explore clustering dynamics across diverse industries and regions. Second, the absence of dynamic industry factors, such as technological advancements and global market trends in the analysis limits the study's comprehensiveness. Future research should aim to incorporate these factors to provide a more nuanced understanding of industry evolution and clustering dynamics. Third, data limitations and biases raise concerns about transparency and reliability. A thorough examination of data limitations is necessary to ensure the credibility of findings and interpretations.

In future research. First, comparative analyses with industries or sectors in other countries or within Bahrain itself could yield valuable insights into universal clustering trends and variations. Exploring similarities and differences across contexts can enrich understanding and inform policy decisions. Second, longitudinal studies to examine the impact of external factors, such as economic conditions and trade policies, on industry clustering dynamics could provide a more comprehensive understanding. Long-term analysis allows for the observation of trends and patterns over time, offering insights into the resilience and adaptability of industries. Third, employing alternative clustering methods to complement existing approaches and enhance the accuracy and reliability of results. Experimenting with different methodologies can offer new perspectives and improve the robustness of findings. Finally, enhancing data collection processes can contribute to the reliability and validity of research outcomes. Employing advanced data collection techniques and ensuring data quality is essential for producing credible insights and conclusions. Addressing these limitations and pursuing suggested future research directions will enable scholars to advance their understanding of industry clustering dynamics, ultimately leading to more robust and applicable findings with broader implications.

Ethical approval

Ethical approval for this study was obtained from the Ethics Committee of the Research Center at the University of Technology Bahrain. This study adhered to the ethical principles outlined in the IRB-HSBS Informed Consent Guidelines of the University of Technology Bahrain.

Consent for publication

Not applicable.

Author contributions

HAQ designed the study and was responsible for the quality assurance of the study results. JPS was responsible for the quality assurance of the interpretation of the results. HAQ was mainly responsible for analyzing the data for this manuscript, and JPS drafted the manuscript in close collaboration with HAQ. Both authors participated in editing, reading, and approving the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

About the authors

Dr. **Haitham Al Qahtani**, a highly educated professional, holds a PhD in Chemical Engineering from the prestigious University of Tufts in Boston, a Master of Energy Engineering from the esteemed University of Leeds in the UK, and a Bachelor of Petroleum Engineering from the renowned University of Southern California in the USA. His academic journey has equipped him with a deep understanding of his field, which he now applies in his role as the Vice President of Academic Affairs at the University of Technology Bahrain. His wealth of academic and professional experience is evident in his previous roles as a project leader during the establishment of the British College of Bahrain, Group Strategic Planning & Project Developer at the Yusuf Bin Ahmed Kanoo Group of Companies, and Vice President for knowledge-based services at the Economic Development Board.

Dr. **Jayendira P. Sankar** received his PhD from the University of Madras, India. He worked for ten years in India and in Bahrain for seven years. Overall, seventeen years of academic experience, Senior Fellow HEA-UK and Academic Fellow CIP D-UK with many publications in various high-impact journals and conferences. The current research interests include Development Economics: economic growth, poverty, inequality, market failure, fiscal, economic and social conditions, healthcare, education, employment sector, industrial and social infrastructure; Human Resource Management: employee welfare, work-family balance, consumer behaviour, retail marketing, corporate social responsibility, sustainability, and teaching-learning and assessments.

Funding

No funding was received for conducting this study.

ORCID

Haitham Al Qahtani  <http://orcid.org/0009-0009-0817-4746>

Jayendira P. Sankar  <http://orcid.org/0000-0001-8435-2123>

Data availability statement

The data that support the findings of this study are available from the corresponding author JPS upon reasonable request.

References

- Argüelles, M., Benavides, C., & Fernández, I. (2014). A new approach to the identification of regional clusters: Hierarchical clustering on principal components. *Applied Economics*, 46(21), 1–19. <https://doi.org/10.1080/00036846.2014.904491>
- Bagherzadeh, M., Ghaderi, M., & Fernandez, A. S. (2022). Coopetition for innovation – The more, the better? An empirical study based on preference disaggregation analysis. *European Journal of Operational Research*, 297(2), 695–708. <https://doi.org/10.1016/j.ejor.2021.06.010>
- Boix, R., Hervás-Oliver, J. L., & De Miguel-Molina, B. (2015). Micro-geographies of creative industries clusters in Europe: From hot spots to assemblages. *Papers in Regional Science*, 94(4), 753–773. <https://doi.org/10.1111/pirs.12094>
- Cavalheiro, G. M. D. C., & Brandao, M. (2017). Assessing the IP portfolio of industrial clusters: The case of the Brazilian footwear industry. *Journal of Manufacturing Technology Management*, 28(8), 994–1010. <https://doi.org/10.1108/JMTM-10-2016-0137>
- Chang, P. C., Liu, C. H., & Fan, C. Y. (2009). Data clustering and fuzzy neural network for sales forecasting: A case study in printed circuit board industry. *Knowledge-Based Systems*, 22(5), 344–355. <https://doi.org/10.1016/j.knsys.2009.02.005>
- Chen, L. C. (2011). The governance and evolution of local production networks in a cluster: The case of Taiwan's machine tool industry. *GeoJournal*, 76(6), 605–622. <https://doi.org/10.1007/s10708-009-9317-2>
- Cook, G. A. S., & Pandit, N. R. (2012). Clustering and the location of multinational enterprises: An exploration of financial services in London. In *The regional economics of knowledge and talent: Local advantage in a global context* (pp. 281–299). Edward Elgar. <https://doi.org/10.4337/9781781953549.00019>
- Cook, G. A. S., Pandit, N. R., & Swann, G. M. P. (2001). The dynamics of industrial clustering in British broadcasting. *Information Economics and Policy*, 13(3), 351–375. [https://doi.org/10.1016/S0167-6245\(01\)00041-5](https://doi.org/10.1016/S0167-6245(01)00041-5)
- Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1(2), 224–227. <https://doi.org/10.1109/TPAMI.1979.4766909>

- Dereli, T., & Durmuşoğlu, A. (2009). Classifying technology patents to identify trends: Applying a fuzzy-based clustering approach in the Turkish textile industry. *Technology in Society*, 31(3), 263–272. <https://doi.org/10.1016/j.tech-soc.2009.06.007>
- Du, M., & Wu, F. (2022). Grid-based clustering using boundary detection. *Entropy*, 24(11), 1606. <https://doi.org/10.3390/e24111606>
- Dureti, G. G., Tabe-Ojong, M. P., & Owusu-Sekyere, E. (2023). The new normal? Cluster farming and smallholder commercialization in Ethiopia. *Agricultural Economics*, 54(6), 900–920. <https://doi.org/10.1111/agec.12790>
- Edina, E., József, P., Miklós, N., & Judit, O. (2021). The role of Industry 4.0 technologies in the innovation activities of food manufacturing companies. *Statistikai Szemle*, 99(10), 978–996. <https://doi.org/10.20311/stat2021.10.hu0978>
- Ergun, M., Uyguçgil, H., & Atalik, Ö. (2020). Creating a geodemographic classification model within geo-marketing: The case of Eskişehir province. *Bulletin of Geography. Socio-Economic Series*, 47(47), 45–61. <https://doi.org/10.2478/bog-2020-0003>
- Gergely, B., & Vargha, A. (2021). How to use model-based cluster analysis efficiently in person-oriented research. *Journal for Person-Oriented Research*, 7(1), 22–35. <https://doi.org/10.17505/jpor.2021.23449>
- Geum, Y., Kim, M. S., & Lee, S. (2016). How industrial convergence happens: A taxonomical approach based on empirical evidences. *Technological Forecasting and Social Change*, 107, 112–120. <https://doi.org/10.1016/j.techfore.2016.03.020>
- Guo, B., & Guo, J. J. (2011). Patterns of technological learning within the knowledge systems of industrial clusters in emerging economies: Evidence from China. *Technovation*, 31(2–3), 87–104. <https://doi.org/10.1016/j.technovation.2010.10.006>
- Hadhri, W., Arvanitis, R., & M'Henni, H. (2016). Determinants of innovation activities in small and open economies: the Lebanese business sector. *Journal of Innovation Economics & Management*, 21(3), 77–107. <https://doi.org/10.3917/jie.021.0077>
- Hanawalt, E., & Rouse, W. (2017). Assessing location attractiveness for manufacturing automobiles. *Journal of Industrial Engineering and Management*, 10(5), 817–852. <https://doi.org/10.3926/jiem.2321>
- He, Y., Xing, Y., Zeng, X., Ji, Y., Hou, H., Zhang, Y., & Zhu, Z. (2022). Factors influencing carbon emissions from China's electricity industry: Analysis using the combination of LMDI and K-means clustering. *Environmental Impact Assessment Review*, 93, 106724. <https://doi.org/10.1016/j.eiar.2021.106724>
- Heragu, S. S., & Ekren, B. (2009). Materials handling system design. In *Environmentally conscious materials handling* (pp. 1–26). John Wiley & Sons. <https://doi.org/10.1002/9780470432730.ch1>
- Holý, V., Sokol, O., & Černý, M. (2017). Clustering retail products based on customer behaviour. *Applied Soft Computing*, 60, 752–762. <https://doi.org/10.1016/j.asoc.2017.02.004>
- Hong, S. J., Kwon, I. W. G., & Li, J. (2014). Assessing the perception of supply chain risk and partnerships using importance-performance analysis model: A case study of SMEs in China And Korea. *Supply Chain Forum*, 15(2), 110–125. <https://doi.org/10.1080/16258312.2014.11517344>
- Hsiao, K. A., & Chen, L. L. (2006). Fundamental dimensions of affective responses to product shapes. *International Journal of Industrial Ergonomics*, 36(6), 553–564. <https://doi.org/10.1016/j.ergon.2005.11.009>
- Huang, X. (2023). The roles of competition on innovation efficiency and firm performance: Evidence from the Chinese manufacturing industry. *European Research on Management and Business Economics*, 29(1), 100201. <https://doi.org/10.1016/j.iedeen.2022.100201>
- Hwang, W. S., & Kim, H. S. (2022). Does the adoption of emerging technologies improve technical efficiency? Evidence from Korean manufacturing SMEs. *Small Business Economics*, 59(2), 627–643. <https://doi.org/10.1007/s11187-021-00554-w>
- Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, 622, 178–210. <https://doi.org/10.1016/j.ins.2022.11.139>
- Jabbarzadeh, A., Fahimnia, B., & Sabouhi, F. (2018). Resilient and sustainable supply chain design: Sustainability analysis under disruption risks. *International Journal of Production Research*, 56(17), 5945–5968. <https://doi.org/10.1080/00207543.2018.1461950>
- Jain, R., Gupta, S., Meena, M. L., & Dangayach, G. S. (2016). Optimisation of labour productivity using work measurement techniques. *International Journal of Productivity and Quality Management*, 19(4), 485–510. <https://doi.org/10.1504/IJPM.2016.1000353>
- Jun, S., & Park, S. S. (2013). Examining technological innovation of Apple using patent analysis. *Industrial Management & Data Systems*, 113(6), 890–907. <https://doi.org/10.1108/IMDS-01-2013-0032>
- Kadokawa, K. (2013). A search for an industrial cluster in Japanese manufacturing sector: Evidence from a location survey. *GeoJournal*, 78(1), 85–101. <https://doi.org/10.1007/s10708-011-9433-7>
- Kalicanin, D., & Gavric, O. (2014). The importance of clusters as drivers of competitive advantage of companies. *Ekonomika Preduzeća*, 62(3–4), 164–172. <https://doi.org/10.5937/ekopre1404164K>
- Kangalli, S. G., Uyar, U., & Buyrokoğlu, S. (2014). OECD Ülkelerinde Ekonomik Özgürlük: Bir Kümeleme Analizi. *Uluslararası Alanya İşletme Fakültesi Dergisi*, 6(3), 95–109.

- Karaca, Z. (2018). The cluster analysis in the manufacturing industry with K-mean method: An application for Turkey. *Eurasian Journal of Economics and Finance*, 6(3), 1–12. <https://doi.org/10.15604/ejef.2018.06.03.001>
- Kim, D. H., Kim, S., & Lee, J. S. (2023). The rise and fall of industrial clusters: Experience from the resilient transformation in South Korea. *The Annals of Regional Science*, 71(2), 1–23. <https://doi.org/10.1007/s00168-022-01170-6>
- Kim, H., Hwang, S. J., & Yoon, W. (2023). Industry cluster, organizational diversity, and innovation. *International Journal of Innovation Studies*, 7(3), 187–195. <https://doi.org/10.1016/j.ijis.2023.03.002>
- König, T. (2023). Between collaboration and competition: Co-located clusters of different industries in one region—The context of Tuttlingen’s medical engineering and metal processing industries. *Regional Science Policy & Practice*, 15(2), 288–325. <https://doi.org/10.1111/rsp3.12581>
- Koo, G. P. Y., Zheng, H., Aik, J. C. L., Tan, B. Y. Q., Sharma, V. K., Sia, C. H., Ong, M. E. H., & Ho, A. F. W. (2023). Clustering of environmental parameters and the risk of acute ischaemic stroke. *International Journal of Environmental Research and Public Health*, 20(6), 4979. <https://doi.org/10.3390/ijerph20064979>
- Libaers, D., & Meyer, M. (2011). Highly innovative small technology firms, industrial clusters and firm internationalization. *Research Policy*, 40(10), 1426–1437. <https://doi.org/10.1016/j.respol.2011.06.005>
- Liberti, L., Lavor, C., Maculan, N., & Mucherino, A. (2014). Euclidean distance geometry and applications. *SIAM Review*, 56(1), 3–69. <https://doi.org/10.1137/120875909>
- Liu, R. (2022). Data analysis of educational evaluation using K-means clustering method. *Computational Intelligence and Neuroscience*, 2022, 3762431. <https://doi.org/10.1155/2022/3762431>
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297).
- Madsen, E. S., Smith, V., & Dilling-Hansen, M. (2003). Industrial clusters, firm location and productivity – Some empirical evidence for Danish firms. In *Working Paper 03-26* (03-26, Issue May 2014). https://pure.au.dk/ws/files/32304323/03-26_esmvs.pdf
- Mittal, H., Pandey, A. C., Saraswat, M., Kumar, S., Pal, R., & Modwel, G. (2022). A comprehensive survey of image segmentation: Clustering methods, performance parameters, and benchmark datasets. *Multimedia Tools and Applications*, 81(24), 35001–35026. <https://doi.org/10.1007/s11042-021-10594-9>
- Nidhi, & Patel, K. A. (2016). An efficient and scalable density-based clustering algorithm for normalize data. *Procedia Computer Science*, 92, 136–141. <https://doi.org/10.1016/j.procs.2016.07.336>
- Novillo-Villegas, S., Ayala-Andrade, R., Lopez-Cox, J. P., Salazar-Oyaneder, J., & Acosta-Vargas, P. (2022). A roadmap for innovation capacity in developing countries. *Sustainability*, 14(11), 6686. <https://doi.org/10.3390/su14116686>
- OECD (2019). Measuring distortions in international markets: The aluminium value chain. In *OECD trade policy papers* (pp. 1–121). OECD Publishing. <https://doi.org/10.1787/c82911ab-en>
- Oyewole, G. J., & Thopil, G. A. (2023). Data clustering: Application and trends. *Artificial Intelligence Review*, 56(7), 6439–6475. <https://doi.org/10.1007/s10462-022-10325-y>
- Palamara, F., Piglione, F., & Piccinini, N. (2011). Self-organizing map and clustering algorithms for the analysis of occupational accident databases. *Safety Science*, 49(8–9), 1215–1230. <https://doi.org/10.1016/j.ssci.2011.04.003>
- Park, J., Lee, H., & Park, Y. (2009). Disembodied knowledge flows among industrial clusters: A patent analysis of the Korean manufacturing sector. *Technology in Society*, 31(1), 73–84. <https://doi.org/10.1016/j.techsoc.2008.10.011>
- Pereira, D., Leitão, J., Oliveira, T., & Peirone, D. (2023). Proposing a holistic research framework for university strategic alliances in sustainable entrepreneurship. *Heliyon*, 9(5), e16087. <https://doi.org/10.1016/j.heliyon.2023.e16087>
- Pereira, T., Kennedy, J. V., & Potgieter, J. (2019). A comparison of traditional manufacturing vs additive manufacturing, the best method for the job. *Procedia Manufacturing*, 30, 11–18. <https://doi.org/10.1016/j.promfg.2019.02.003>
- Plantec, Q., Le Masson, P., & Weil, B. (2021). Impact of knowledge search practices on the originality of inventions: A study in the oil & gas industry through dynamic patent analysis. *Technological Forecasting and Social Change*, 168, 120782. <https://doi.org/10.1016/j.techfore.2021.120782>
- Prakash, A., Jha, S. K., Prasad, K. D., & Singh, A. K. (2017). Productivity, quality and business performance: An empirical study. *International Journal of Productivity and Performance Management*, 66(1), 78–91. <https://doi.org/10.1108/IJPPM-03-2015-0041>
- Prebanić, K. R., & Vukomanović, M. (2023). Exploring stakeholder engagement process as the success factor for infrastructure projects. *Buildings*, 13(7), 1785. <https://doi.org/10.3390/buildings13071785>
- Ran, C., Song, K., & Yang, L. (2020). An improved solution for partner selection of industry-university cooperation. *Technology Analysis & Strategic Management*, 32(12), 1478–1493. <https://doi.org/10.1080/09537325.2020.1786044>
- Rivera, L., Gligor, D., & Sheffi, Y. (2016). The benefits of logistics clustering. *International Journal of Physical Distribution & Logistics Management*, 46(3), 242–268. <https://doi.org/10.1108/IJPDLM-10-2014-0243>
- Schulz, M. A., Chapman-Rounds, M., Verma, M., Bzdok, D., & Georgatzis, K. (2020). Inferring disease subtypes from clusters in explanation space. *Scientific Reports*, 10(1), 12900. <https://doi.org/10.1038/s41598-020-68858-7>
- Sharma, S. (2011). Improved BSP clustering algorithm for social network analysis. *Bonfring International Journal of Software Engineering and Soft Computing*, 1(1), 15–20. <https://doi.org/10.9756/BIJSESC.1003>
- Sheikh, A., Ghanbarpour, T., & Gholamiangonabadi, D. (2019). A preliminary study of fintech industry: A two-stage clustering analysis for customer segmentation in the B2B setting. *Journal of Business-to-Business Marketing*, 26(2), 197–207. <https://doi.org/10.1080/1051712X.2019.1603420>

- Singhal, V., Singh, A. B., Ahuja, V., & Gera, R. (2023). Consumer segmentation in the fashion industry using social media: An empirical analysis. *Journal of Information and Organizational Sciences*, 47(2), 399–419. <https://doi.org/10.31341/jios.47.2.9>
- Sun, J., Fan, P., Wang, K., & Yu, Z. (2022). Research on the impact of the industrial cluster effect on the profits of new energy enterprises in China: Based on the Moran's I index and the fixed-effect panel stochastic frontier model. *Sustainability*, 14(21), 14499. <https://doi.org/10.3390/su142114499>
- Swarndhepp, S. J., & Pandya, S. (2016). An overview of partitioning algorithms in clustering techniques. *International Journal of Advanced Research in Computer Engineering & Technology*, 5(6), 1943–1946.
- Tan, A. J. J., Chong, C. Y., & Aleti, A. (2022). E-SC4R: Explaining software clustering for remodularisation. *Journal of Systems and Software*, 186, 111162. <https://doi.org/10.1016/j.jss.2021.111162>
- Thakare, Y., & Bagal, S. (2015). Performance evaluation of K-means clustering algorithm with various distance metrics. *International Journal of Computer Applications*, 110(11), 12–16. <https://doi.org/10.5120/19360-0929>
- Tingley, D., Ásmundsson, J., Borodzicz, E., Conides, A., Drakeford, B., Rúnar Eðvarðsson, I., Holm, D., Kapiris, K., Kuikka, S., & Mortensen, B. (2010). Risk identification and perception in the fisheries sector: Comparisons between the Faroes, Greece, Iceland and UK. *Marine Policy*, 34(6), 1249–1260. <https://doi.org/10.1016/j.marpol.2010.05.002>
- Unterberger, P., & Müller, J. M. (2021). Clustering and classification of manufacturing enterprises regarding their Industry 4.0 reshoring incentives. *Procedia Computer Science*, 180, 696–705. <https://doi.org/10.1016/j.procs.2021.01.292>
- Vargas-Hernández, J. G., Vargas Gonzalez, O. C., & Morones Servín, J. S. (2020). A theoretical approach to the concept of the cluster. *Problems of Management in the 21st Century*, 15(1), 56–67. <https://doi.org/10.33225/10.33225/pmc/20.15.56>
- Wang, W., Tang, Q., & Gao, B. (2023). Exploration of CO₂ emission reduction pathways: identification of influencing factors of CO₂ emission and CO₂ emission reduction potential of power industry. *Clean Technologies and Environmental Policy*, 25(5), 1–15. <https://doi.org/10.1007/s10098-022-02456-1>
- Wenting, R. (2008). Spinoff dynamics and the spatial formation of the fashion design industry, 1858–2005. *Journal of Economic Geography*, 8(5), 593–614. <https://doi.org/10.1093/jeg/lbn030>
- Westerdijk, M., Zuurbier, J., Ludwig, M., & Prins, S. (2012). Defining care products to finance health care in the Netherlands. *The European Journal of Health Economics*, 13(2), 203–221. <https://doi.org/10.1007/s10198-011-0302-6>
- Yang, F., & Gu, S. (2021). Industry 4.0, a revolution that requires technology and national strategies. *Complex & Intelligent Systems*, 7(3), 1311–1325. <https://doi.org/10.1007/s40747-020-00267-9>
- Yang, W. C., Lai, J. P., Liu, Y. H., Lin, Y. L., Hou, H. P., & Pai, P. F. (2023). Using medical data and clustering techniques for a smart healthcare system. *Electronics*, 13(1), 140. <https://doi.org/10.3390/electronics13010140>
- Ye, D., Zheng, L., & He, P. (2021). Industry cluster innovation upgrading and knowledge evolution: A simulation analysis based on small-world networks. *SAGE Open*, 11(3), 215824402110316. <https://doi.org/10.1177/21582440211031604>
- Yuan, C., & Yang, H. (2019). Research on K-value selection method of K-means clustering algorithm. *J*, 2(2), 226–235. <https://doi.org/10.3390/j2020016>
- Zhao, L., Liang, Y., & Tu, H. (2023). How do clusters drive firm performance in the regional innovation system? A causal complexity analysis in Chinese strategic emerging industries. *Systems*, 11(5), 229. <https://doi.org/10.3390/systems11050229>
- Zhou, K., Yang, S., & Shao, Z. (2017). Household monthly electricity consumption pattern mining: A fuzzy clustering-based model and a case study. *Journal of Cleaner Production*, 141, 900–908. <https://doi.org/10.1016/j.jclepro.2016.09.165>