

Exploratory Data Analysis of Sustainability Performance Metrics in Automotive MSMEs

Dipak Patil* , Geeta Rao 

MIT Art, Design and Technology University, Pune, Maharashtra, India

Corresponding Email: dipak0508dipak@gmail.com

Article Info	Abstract
Received 30/05/2026	<p>Micro, Small, and Medium Enterprises (MSMEs) play a crucial role in India's manufacturing sector, yet they face numerous obstacles to sustainability and Environmental, Social, and Governance (ESG) practices. In this study, the sustainability performance of 127 automotive components MSMEs of Pune, Chakan, Satara, Nashik, and Kolhapur has been evaluated. The goal is to identify Key Performance Indicators (KPIs) that have an impact on the performance of an organization, including economic, environmental, and social indicators. A mixed-method approach of structured questionnaires and interviews was used. To find patterns in sustainability and important KPIs, Exploratory Data Analysis (EDA), Principal Component Analysis (PCA), and clustering techniques were used. Results show that Customer Retention (CR), Resource Efficiency (RE), and Training & Development (TD) are the most significant factors impacting sustainability performance. Pune- and Chakan-based firms showed marginally better sustainability practices, and there were significant variations between different levels in the hierarchy of employees. The findings offer practical implications for MSMEs, policy makers, and industry stakeholders to reinforce the implementation of ESG, enhance the performance of the organizations, and help sustainable industrial development.</p>
Revised 23/10/2026	
Accepted 29/10/2026	

Keywords: Automotive Components, Clustering Techniques, Exploratory Data Analysis, MSMEs, Sustainability Assessment.

1. Introduction

1.1. Background and Context

The Indian economy is dominated by Micro, Small, and Medium Enterprises (MSMEs), which contribute more than 30 percent to the national GDP and offer employment to over 110 million individuals. Their prevalence in manufacturing and service industries highlights their importance in promoting industrial growth and socio-economic development. Maharashtra, as one of the most industrialized states in India, is a region of a high concentration of MSMEs, especially in the automotive component industry, which is a vital node in the national and global supply chains. Over the past few years, the increasing focus on sustainability in global value chains has turned compliance with Environmental, Social, and Governance (ESG) standards into a voluntary approach to a strategic requirement. In the case of MSMEs, it is no longer a choice but a necessity to align operations with these sustainability imperatives in order to remain competitive, have market access, and be resilient in the face of a changing industrial ecosystem. However, the implementation of sustainability by MSMEs in India is threatened with overwhelming challenges. Although they are highly strategic

entities, these enterprises usually have little situational awareness, purchase flexibility, or technical facilities to apply regulated sustainability models. Most recent research findings point out that MSMEs are aware of sustainability, but there is little application of structured benchmarking tools to organizational performance in economic, environmental, as well as social aspects [1]. The automotive industry in particular has come under pressure to adopt sustainable supply chain management (SSCM), mainly from regulators and investors. The factors have now been identified as government regulation, environmental certifications, and customer expectations as major contributors to the adoption of sustainability in this industry [2]. There is, however, uneven implementation among smaller suppliers.

Within the framework of manufacturing MSMEs, sustainability barriers lie within the scope of outdated operations, shortage of skilled labor, low-level collaboration networks, and low institutional support. In a study dedicated to the automobile MSMEs in India, 22 barriers to sustainable manufacturing have been identified, with operational and regulatory limitations seen as the main obstacle [3]. More importantly, the MSME Manufacturing Business Facilitator (MSME-MBF) Index,

which measures the infrastructure, access to credit, availability of skilled labor, and ease of doing business in a state, puts Maharashtra at the top in India [4]. It is an interesting region to explore sustainability in practice in data-plentiful places such as Pune, Chakan, Satara, Nashik, and Kolhapur in Maharashtra. The sophistication of sustainability requires a shift from intuition-based decision-making to data-informed decision-making. Sadly, data analytics (in particular, the exploratory data analysis (EDA)) has not reached its peak even at the level of Indian MSMEs. Current research and policy instruments are frequently unable to offer these in detail, taking organizational role, firm size, or even geographical differences into consideration; there is a strong need for accessible information.

1.2. Motivation for Study

Even though the focus of MSMEs in sustainability discourse, the majority of businesses still work without the systemic monitoring of the Key Performance Indicators (KPIs). Though big companies can deploy dashboards and predictive systems as a means of observing ESG objectives, the average Indian MSME firm is more prone to informal judgment, ad hoc reporting, or simple financial statements. Such a lack of structured sustainability monitoring is a serious hindrance to evidence-based policy intervention, benchmarking, as well as continuous improvement [5]. Although the prospect of high-level technologies like predictive analytics and the real-time monitoring of information have already been proven feasible in Industry 4.0 situations in the global market, the reality is that Indian MSMEs have resource and capability limitations in implementing such systems. In a recent study, predictive analytics and machine learning have been categorized as the highest-impact Industry 4.0 technologies in terms of sustainable performance; however, adoption is low among MSMEs that lack strategic direction and technical infrastructure [6].

Besides, the application of EDA and visual analytics is almost non-existent in the MSME sustainability literature. Today, we see the definite advantages of big data and visual modeling in project performance and decision support, but microenterprises in India hardly use it [5]. The literature available on the market is largely devoted to the studies of big corporations or overall digitalization and ignores peculiar structural and organizational obstacles generated in small-scale manufacturing. This leaves one big hole: what regional, sector-specific knowledge can be available in terms of knowledge about sustainability performance, particularly when that performance is disaggregated by size of company, location, or type of actor (e.g. CEO versus worker views). MSMEs are, therefore, bereft of the evidence base in terms of evidence localization as well as the instrumentation to foster sustained propagation of sustainability.

1.3. Research Gap, Aim, Novelty and Contributions

While there have been a number of studies examining frameworks and approaches to sustainability assessment for Small and Medium Enterprises (SMEs), the majority of these studies have examined either conceptual models, large organizations, or specific sustainability indicators. There is scant literature that has examined the sustainability performance of automotive Micro, Small and Medium

Enterprises (MSMEs) in India from a holistic perspective using a data-driven analytical approach. Moreover, few studies combine EDA, Principal Component Analysis (PCA), and clustering to select key sustainability indicators (KPIs) that are effective and to analyze differences by enterprise size, geography, and hierarchy.

The purpose of this research is to assess the sustainability performance of automotive MSMEs of Maharashtra by employing data-driven exploratory analytics and to determine the most significant economic, environmental, and social KPIs to influence organizational performance.

The novelty of the present study is the application of EDA and clustering techniques on a primary dataset collected from automotive MSMEs to decipher the hidden patterns and performance drivers for sustainability. The analysis takes into account regional differences and organizational size, besides the employee hierarchy, unlike previous studies, to give a multidimensional understanding of sustainability practices in MSMEs.

This study has contributed significantly through:

- Creating a detailed sustainability assessment framework based on 22 organizational performance indicators for automotive MSMEs.
- The use of EDA and clustering techniques to identify the most influential sustainability KPIs.
- Cross-industry comparison of Sustainability performance at different levels of industry organization, enterprise size, and industrial region.
- Delivering practical recommendations to facilitate ESG implementation, policymaking, and sustainability improvement of automotive MSMEs.

2. Literature Review

There have been predominant uses of conceptual frameworks in the measurement of sustainability performance in SMEs, including the Global Reporting Initiative (GRI), Triple Bottom Line (TBL), and the Balanced Scorecard. Although these methods provide a basis for measuring aspects of sustainability, they may turn out to be idealistic within the market of MSMEs as a result of limited resources and limited ability to implement. As an example, Vitale et al. [7] observed the strategic application of the sustainability performance management models within European SMEs and raised the question of the systematic approach that could not be applied to most small enterprises, which is also the problem reflected in Indian MSMEs. Malesios et al. [8] developed a criteria-based framework for measuring SME sustainability performance by combining economic, social, and environmental indicators and noted that performance monitoring models are used inconveniently in emerging economies. Sustainability Balanced Scorecard is a theory with potential, but its use in reality, especially in SMEs in the Netherlands, was very low, as reported by Chehimi et al. [9], and the latter is likely to be repeated in India. Focusing on the specificity of sectors in sustainable manufacturing, Chourasiya et al. [10] provided a structural model of a sector-specific framework, establishing it

in the Indian textile MSMEs and indicating its replicability to different fields such as the automotive industry. Singh et al. [11] have found the KPIs of sustainable manufacturing in global SMEs and initially presented the list of KPIs concerning the efficiency of resources and minimization of waste, as well as training, yet without breaking it down into regions.

Recently, Gidage and Bhide [12] investigated the connection of intellectual capital, green innovation, and sustainability within creative industry MSMEs in India, which implies that knowledge assets and innovation are a crucial element of the outcome of ESG. Das et al. [13] described that network connectivity and government intervention were both important factors in determining the level of sustainability in Indian MSMEs, especially through indirect policy levers and shared infrastructure. In measuring sustainability, corporate measures have been the center of discussion, but Illahi and Mir [14] pointed out how such measurements could be transformed towards regional localities, such as metropolitan transport policy, to become a proxy for localized sustainability measures. Providing cross-country evidence in India and Japan, Laskar et al. [15] linked corporate sustainability performance with financial performance, stressing that structured performance monitoring will create strategic value. Babber and Mittal [16] suggested an integrated lean, agile, innovative manufacturing system that could help improve sustainability in Indian MSMEs and stressed that adaptive process enhancements should be built into the model. Nonetheless, few of these studies reflect data-driven exploratory procedures like clustering to bring out hidden patterns in the performance sustainability of MSMEs. Missing are the studies that look specifically into the major hubs of MSMEs in Maharashtra, such as Pune, Chakan, and Kolhapur. In addition to that, there is still a lack of comparative studies that investigate role-wise (e.g., manager vs. worker) and firm-size-based sustainability perception variations. Though there has been partial progress in the definition of KPIs and frameworks, integration of EDA, unsupervised learning, and visualization of actionable sustainability insights remain in their early stages in the Indian MSMEs. The paper aims to fill that gap by evaluating the actual survey-based data with contemporary data analysis methods to arrive at regionally based, role-sensitive regional insights that can be used in strategic policy and business decisions.

With the advancements made in recent years, a number of components based upon empirical studies have started to focus on the significance of quantitative approaches to facilitate the development of sustainability on the platform of micro and small businesses. As an example, Haileslassie et al. [17] have done a thorough evaluation of the typologies of dry land farms in India based on hierarchical clustering, wherein the economic, social, and environmental sustainability pillars come out to be expressed differently in terms of each type of farm, with economic sustainability being behind in most instances. Agasty et al. [18] conducted research in which they developed a Sustainability Innovation Index on Indian MSMEs using multivariate statistical tools and concluded that productivity, institutional support, and cost parameters are the factors that drive innovation towards sustainable practices. Inefficiencies in operations among MSMEs are also disadvantageous in achieving sustainability, as a study by Siddhu et al. [19]

revealed the importance of maintenance management in Indian SMEs and the low rate of sustenance of lean manufacturing programs because of the lack of data-based execution plans. As a policymaker would suggest, Verma [20] noted that MSMEs hold great potential to realize the green development agenda in India because of fragmented data and institutional support mismatch, specifically with SDG 8. Roy et al. [21] took the sustainability debate to the urban Indian scene, disclosing that although the SDG oversight systems are in place at the city levels, regional differences cannot be overemphasized, and sometimes the environmental achievements do not correlate well with the socioeconomic benefits unless data integration tools are employed through correlation matrixing and clustering.

Considering closely relevant frequent MSMEs in the agricultural sector, Sharma et al. [22] used clustering analysis to cluster typologies of farms in Punjab, where they recommended context-based interventions that regard both the structural and operational disparities in the size of enterprises, cropping density, and sources of income. In the meantime, Pandya and Kumar [6] determined how Reality 4.0 technologies such as AI, IoT, and Big Data can influence sustainability in Indian service-sector MSMEs. They mentioned the paradox of high potential and low adoption, which is mostly caused by the capability and infrastructure gradient. Coupled with this is the exposure of Kadaba et al. [23] on how a mixture of government schemes (e.g., ONDC, PMMY, Startup India) and digital platforms can support inclusive and sustainable growth of MSMEs, provided that they are well-combined and monitored using digital analytics. Last of all, Herman and Zsido slice [24] suggested a Financial Sustainability Index based on groups of SMEs in Europe, which demonstrates that both aspects of balanced financial performance and liquidity are strongly associated with sustainability outcomes, which generates a useful, SME-based blueprint applicable in the Indian MSME context.

Although a number of frameworks such as GRI and TBL have been applied in the assessment of sustainability in SMEs, their practical use at the regional and sectoral level by Indian MSMEs is, however, minimal. The majority of studies are not quantitative, region-based (e.g., automotive MSMEs of Maharashtra), and do not consider role-wise and size-wise variation in the perception of sustainability. In addition, there is a lack of using EDA, clustering, and dimensionality reduction to acquire policy-relevant insights. This paper tries to fill these gaps by using contemporary data analytics techniques on a primary data set obtained regarding MSMEs in Maharashtra.

The sustainability indicators used in the current study were not created from scratch but rather were assembled from known sustainability assessment frameworks and recent empirical studies relating to MSMEs. The economic dimension included indicators that are commonly used in SME sustainability assessment studies, such as revenue growth, cost management, innovation investment, customer retention and organization performance. Environmental aspects such as resource efficiency, energy efficiency, water, waste management and carbon footprint were taken from Environmental, Social and Governance (ESG) reporting principles, the Global Reporting

Initiative (GRI) and sustainability performance measurement frameworks. The social dimension was developed after some research on organizational sustainability and MSME competitiveness. Also, the themes that emerged from the semi-structured CEO interviews were based on documented operational barriers commonly observed in recent literature on MSMEs or micro-businesses. The instruments used for data collection were content-validated and practically relevant with this literature-driven approach.

3. Methodology

The study follows a mixed-methods research design to assess the sustainability performance of MSMEs in the automotive components sector in Maharashtra. The method combines both quantitative data collected by structured questionnaires and qualitative information collected by interviews in order to achieve both statistical rigor and contextual interpretation. EDA methods are used to reveal the patterns, determine the most important sustainability indicators, and come up with recommendations that can be implemented by policymakers and stakeholders in the industry. The methodology also describes the sources of data, segmentation of respondents, and the composition of variables in economic, environmental, and social aspects. Before analysis, the data is preprocessed to fill in the gaps and provide consistency, which will serve as the basis of clustering in the later stages of the investigation.

3.1. EDA Overview

The first analytical step was to use Exploratory Data Analysis (EDA) in order to gain insights into the nature, quality, and structure of the collected data prior to using multivariate statistical techniques. EDA offers a structured approach for summarizing the data, determining the distribution of the data, detecting the outliers, analyzing the relationship between the variables, and determining whether the data is suitable for further analysis in the form of Principal Component Analysis (PCA) or clustering. The present study analysed the responses obtained from 127 automotive Micro, Small and Medium Enterprises (MSMEs) of Maharashtra, India. The data set included 22 variables, such as enterprise characteristics, economic, environmental, and social sustainability indicators, and the combined performance of the organization. The statistical properties of the data were checked by means of descriptive statistics, frequency analysis, distribution visualization, correlation analysis, and outlier identification. Prior to the development of models, missing values, inconsistent records, and extreme observations were investigated for analytical robustness. The insights gained from EDA were used for feature interpretation, validation of data quality, and as a statistical basis for dimensionality reduction (via PCA) and for enterprise performance segmentation (via clustering analysis).

3.2. Data Collection Strategy

A mixed-method research approach was used in order to get a holistic picture of the sustainability performance of the automotive component MSMEs. The primary quantitative data were obtained from 127 MSMEs spread across various

manufacturing clusters of Pune, Chakan, Satara, Kolhapur, and Nashik, with varying numbers of employees, turnover, and various attributes of MSMEs. These industrial regions were chosen because they represent the key automotive manufacturing ecosystem in Maharashtra and have a wide range of MSME suppliers.

The quantitative survey instrument was created based on a thorough review of sustainability assessment frameworks such as the Environmental, Social, and Governance (ESG) indicators, the Global Reporting Initiative (GRI) guidelines, the Triple Bottom Line (TBL) principles, and sustainability performance measurement studies included in recent literature. The questionnaire consisted of 22 organizational variables under three categories: economic, environmental, and social sustainability. The perception-based variables were analyzed on a 5-point Likert scale from 1 = Very Low to 5 = Very High. The questionnaire was pilot-tested with representatives of 10 MSMEs before its full-scale deployment for the purpose of assessing the clarity of the questions, content validity, and consistency of the responses. The comments from the pilot study led to some minor changes in the wording and ordering of some of the questions. Then the final survey took place between January and May 2025 via Google Forms, with 122 full online responses, complemented by a further 5 responses from validated surveys conducted directly in industry, which brought the total number of enterprises completing the survey to 127.

In order to complement the quantitative findings, 5 Chief Executive Officers (CEOs) from various enterprise sizes and geographical areas were interviewed in a semi-structured fashion. The interviews delved into the sustainability strategies organizations are pursuing, implementation hurdles, employee engagement, and sustainability's long-term ESG goals. The qualitative findings were thematic and used to complement the quantitative results but not as a standalone analysis.

The sampling was done using a stratified sampling approach to gain a cross-sectional representation of the sustainability perceptions of various organizational levels. Respondents comprised Workers/Operators (20), Supervisors (28), Middle Managers (36), Senior Managers/Chief Technology Officers (38), and Chief Executive Officers (5). This hierarchical representation allowed the comparison of the sustainability awareness and priorities of the organizations at the operational and strategic decision-making levels.

Informed consent was obtained from all subjects before the data were collected. The responses of individual respondents were anonymized and kept confidential and the study was carried out solely for academic reasons and within the accepted norms of research ethics and data privacy.

3.3. Dataset Composition

The final analytical dataset comprised 127 observations and 22 variables, of which 21 variables were organizational characteristics, and 1 was a sustainability performance indicator. Three variables were categorical (Company, Location, and Employee Size), and the rest were quantitative organizational attributes and sustainability indicators. These

included financial data, including annual turnover, as well as economic, environmental, and social sustainability data, which were gathered on a 5-point Likert scale. **Organizational Performance** represents a composite sustainability performance index derived from multiple organizational indicators rather than an individual questionnaire item measured on the five-point Likert scale. While each sustainability indicator was evaluated using a five-point Likert scale, the aggregated Organizational Performance index provides an overall measure of enterprise sustainability and therefore is not constrained by the original response scale.

The role-wise distribution of key sustainability performance indicators, including Customer Retention, Resource Efficiency, and Training Development, is presented in Fig. 1.

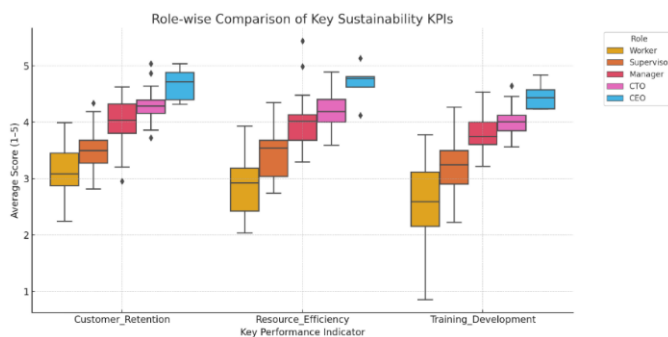


Figure 1. Role-wise boxplot comparison of key sustainability KPIs.

The dataset was analyzed after examining it to ensure completeness, consistency, and statistical validity. Descriptive statistics, frequency distributions, data-type verifications, and summary statistics were performed to assess data integrity through the Python data analysis environment, which includes `info()` and `describe()` functions. Measures of central tendency and dispersion were used to analyze the continuous variables such as Revenue (Cr) and Organizational Performance, while the ordinal sustainability indicators were summarized using frequency distributions and descriptive statistics. The revenue of the enterprises varied from 2.25 crore to 230.87 crore, indicating that there was a significant variation among the enterprises participating in the study. On the other hand, Resource Efficiency and Cost Management achieved an average score of 3.437 and 3.386, respectively, suggesting moderate sustainability performance among the 50 companies surveyed. The mean score of the composite organizational performance score was 3.68, indicating the overall level of sustainability maturity of the sampled MSMEs before the further multivariate analysis.

3.4. Data Types and Initial Distribution

The data set was systematically pre-processed before the multivariate statistical analysis to make the data reliable and consistent for the analysis. This was ensured by verifying the types of variables and categorizing them as continuous,

categorical, or ordinal. Considering standard practice found in sustainability assessment studies, the Likert-scale sustainability indicators were analyzed descriptively as ordinal variables and then used in correlation analysis, Principal Component Analysis (PCA), and clustering.

To explore and understand the distributional properties of the variables, descriptive statistics, frequency distributions, Kernel Density Estimation (KDE), boxplots, and correlation matrices were used. Special focus was placed on aspects related to sustainability, such as Energy Efficiency, Water Usage, Carbon Footprint, Customer Retention, and Innovation Investment. There was moderate central tendency around the median scales on the Likert scale, with most variables displaying the typical moderate and realistic organizational responses to sustainability survey questions.

The organizational performance variable is a combination of various sustainability indicators that include the economic, environmental, and social aspects. The composite organizational performance scores were linearly normalized to a common 1-5 scale using min-max normalization, so that they are consistent with the individual sustainability indicators and can be meaningfully compared across all variables. This change ensures the relative position of enterprises and the proportional differences remain the same as before this change, while the composite index is still aligned with the original measurement scale used for all the questions in the questionnaire. The composite organizational performance scores were linearly normalized to a five-point index for visualization only to keep the original indicators on the Likert scale. The normalization equation used by Geron [25] is the following min-max normalization equation.

$$OP_{\text{norm}} = 1 + 4 \left(\frac{OP - OP_{\min}}{OP_{\max} - OP_{\min}} \right) \quad (1)$$

where:

OP_{norm} = normalized organizational performance score (1–5 scale)

OP = original composite organizational performance score

$OP_{\min}$ = minimum observed composite score (= 2.00)

$OP_{\max}$ = maximum observed composite score (= 7.72)

Exploratory data analysis was conducted to detect any possible outliers, inconsistencies, or missing observations. The data were checked for missing values and judged to have little impact on the overall data. Outstanding observations on enterprise revenue and organizational performance were kept since they reflected true organizational differences, not measurement errors. The preprocessing phase guaranteed that the data set met the assumptions needed for the later visualization, correlation analysis, dimensionality reduction, and enterprise clustering. A sample of the processed data used for the later statistical analysis is shown in Table 1.

Table 1. Preview of 5 rows of the dataset.

Company	Location	Employees	Revenue_Cr	Resource_Efficiency	Organizational_Performance
Company_1	Pune	51-100	85.95	2.0	2.69
Company_2	Pune	1-10	84.85	4.0	2.82
Company_3	Satara	101-200	7.95	3.0	4.81
Company_4	Pune	51-100	230.87	4.0	7.17
Company_5	Nashik	51-100	97.98	3.0	3.09

4. Results and Discussion

4.1. Univariate Analysis

To provide background context on the distribution of sustainability performance indicators and their associated numerical characteristics, a univariate statistical test was conducted. Kernel Density Estimation (KDE) histograms were plotted against the selected variables, namely *Revenue_Cr*, *Resource_Efficiency*, *Cost_Management*, *Innovation_Investment*, *Customer_Retention*, and *CSR_Activities*, as shown in Fig. 2. Such variables were chosen because they had significance in highlighting important economic, social, and environmental aspects of sustainability. Distribution of *Revenue_Cr* was found to be very right-skewed, depicting the presence of a small number of companies with relatively large annual revenues. This implies a non-homogeneous sample of same-economy enterprises with micro enterprises, small, and medium-sized enterprises, and some of them may possess features similar to larger enterprises. The indicators *Resource_Efficiency* and *Cost_Management* showed a relative distribution with an approximately equal level of individual items throughout the period of the Likert scale (2 to 5), which means that firms perform in those areas on a very wide scale. On the contrary, the pattern on *Innovation_Investment* was left-skewed, with most of the distribution falling on the lower end of the rating scale. It can be taken to mean that there are companies in the data set that dedicate relatively restricted investment to innovation-related operations. The probability density plot is provided by the following equation that approximates probability density function using kernels (usually Gaussian) [26].

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (2)$$

The *Customer_Retention* indicator has high concentration around the upper limits (4 to 5) of the scale, indicating that the majority of the participating organizations regard the topic of customer loyalty and satisfaction as an enforced strategic issue. Meanwhile, the *CSR_Activities* variable showed even greater dispersion of answers, which means that there is a significant variation in the engagement in social and community activities in firms. These are the preliminary results that offer significant directional pointers to further multivariate and correlation analysis. Moreover, the stated distributional properties are the basis of outlier identification and feature transformation in the further phases of data pre-processing.

Boxplots were created to study the distribution structure and determine the possible outliers of the critical performance indicators of sustainability and organizational

variables, which were as follows: *Revenue_Cr*, *Resource_Efficiency*, *Cost_Management*, *Innovation_Investment*, *Customer_Retention*, and *CSR_Activities*. The boxplots showed clear trends of these variables (Fig. 3). As the value of the *Revenue_Cr* variable in the higher range had few extreme scores, it was established that there were substantial outliers within a small group of firms with very high annual revenues. This portrays the variability in the scale of businesses among the sampled MSMEs, which could signify a blend involving large-scale operations as well as established medium-sized businesses. By contrast, variables like *Resource_Efficiency* and *Cost_Management* had relatively small interquartile ranges and no significant outliers, meaning most organizations perform similarly in these areas. The following box plot equation shows the distribution of data in terms of median, quartiles, and limits of outliers [27].

$$IQR = Q3 - Q1 \quad (3)$$

$$Lower\ Bound = Q1 - 1.5 \times IQR \quad (4)$$

$$Upper\ Bound = Q3 + 1.5 \times IQR \quad (5)$$

In the case of the *Innovation_Investment* measurement variable, the boxplot indicated that the distribution was left-skewed, although it appeared balanced, with some outliers on both the lower and upper ends, indicating that some companies spend very little on innovative projects. The same pattern was witnessed in the *CSR_Activities* entrant, where a few companies had no significant action recorded; hence, they became the low-value outliers. These results indicate the lack of homogeneity of sustainability adoption, especially in those areas where strategic vision and long-term planning tend to be dominant. More importantly, there were no marked outliers in *Customer_Retention*, which implies that an active and constant practice of customer engagement is present throughout the firms. In general, the boxplot visualization allowed gaining a good impression of the distribution and skewness of the important indicators and contributed to pinpointing anomalies that could be the subject of future contextually specific investigation or niche interventions.

Figs. 2 and 3 are complementary in the statistical information they provide, although they include the same variables. The Kernel Density Estimation (KDE) histograms show the distributional characteristics, skewness, and concentration of observations, while the boxplots summarize the median, interquartile range, dispersion, and possible outliers. All of these visualizations can be interpreted together to get a complete picture of data distribution before multivariate analysis.

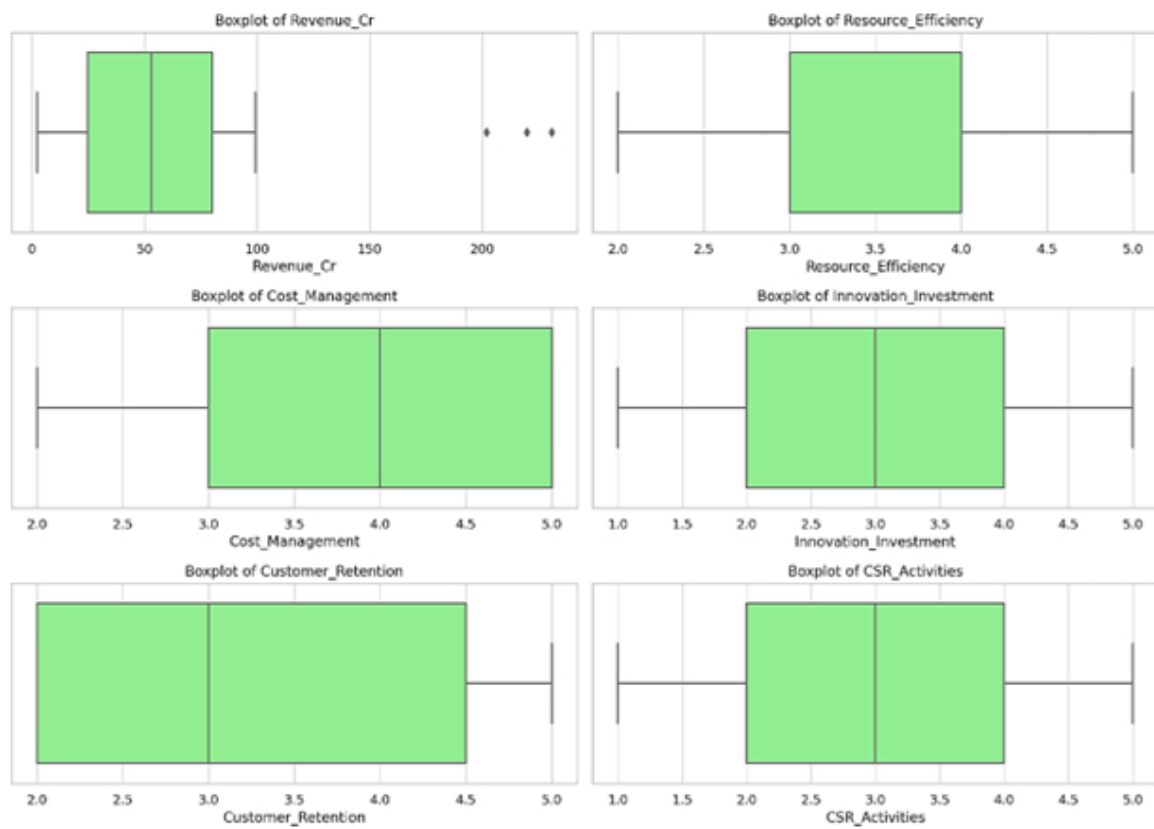


Figure 3. Univariate boxplots and outlier

detection.

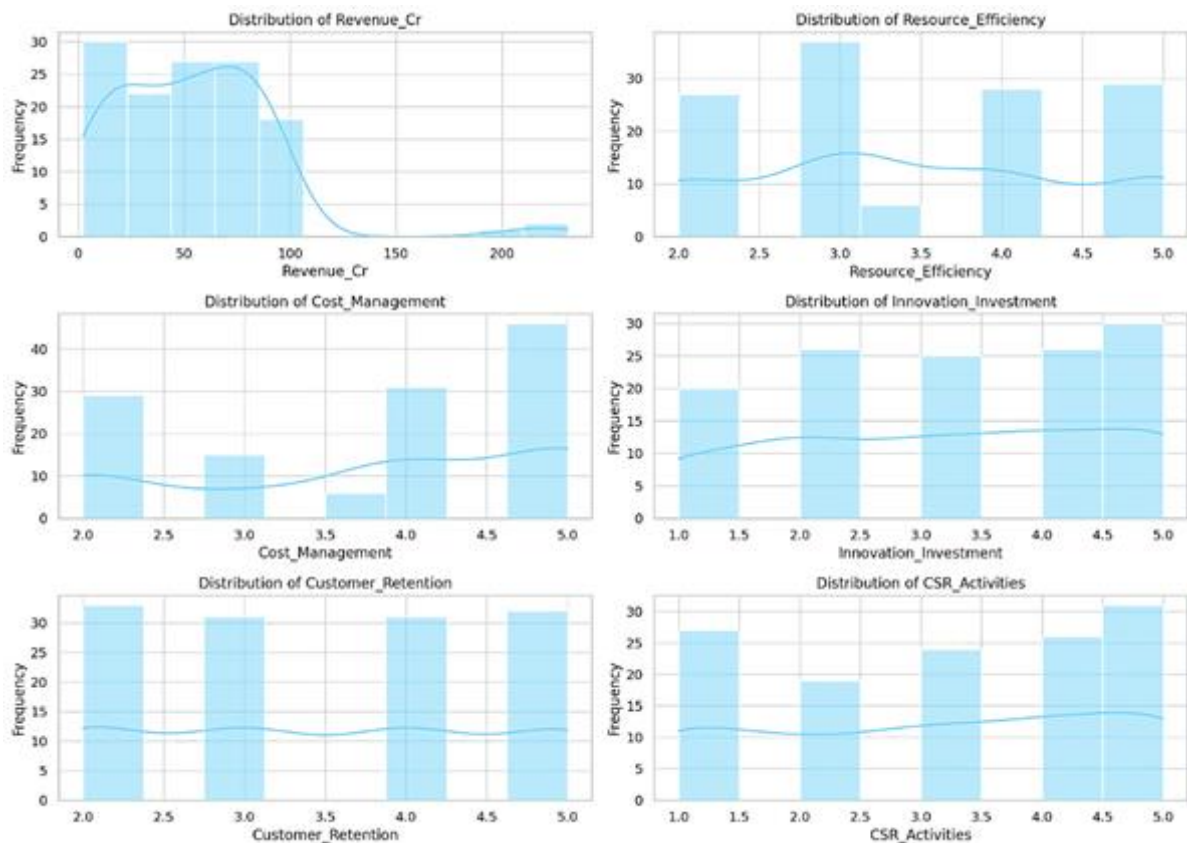


Figure 2. Kernel Density Estimation (KDE) histograms of key numeric features.

The observed distribution patterns show that implementation of sustainability practices varies within the automotive MSMEs, with variations in the level of organizational maturity and resource availability. The relatively low interquartile ranges of the operational practices Resource Efficiency and Cost Management indicate that these practices are now reasonably standardized by the surveyed firms, while the higher interquartile ranges for Innovation Investment and Corporate Social Responsibility (CSR) Activities suggest that long-term sustainability investments are not consistently implemented [28]-[29]. In line with this, some of the recent sustainability studies showed that operational efficiency practices are adopted earlier than strategic sustainability investments due to the financial and technological barriers that hinder MSMEs. From a managerial point of view, the variability identified indicates the need for a shift in focus from solely performance efficiency to innovation capability, employee development, and CSR awareness in the implementation of policy. Moreover, the existence of outliers among enterprises in terms of revenue supports the assumption of the enterprises' heterogeneous nature and hence the use of multivariate techniques for extraction of latent patterns of sustainability.

4.2. Bivariate Analysis

A bivariate analysis was used to determine whether there is any relationship among sustainability indicators and the general performance of an organization by way of the Pearson correlation coefficient. The degree and direction of the association between the set of numerical variables included in the dataset was visually presented by developing a heatmap, as illustrated in Fig. 4. It was found that there is a strong positive correlation between Resource Efficiency and Cost Management ($r=0.65$), implying that cost-conscious organizations also become effective in resource use. This correlation makes sense and fits in sustainable operational practice since conservation of resources leads to direct cost savings. In a similar way, there was a significant relationship found between Employee Wellbeing and Training & Development ($r=0.26-0.2$). This observation highlights the importance of worker-focused programs to promote workplace satisfaction and productivity and confirms the notion that an investment in professional development and education is one of the contributors to improved well-being in organizations. A rather large number of variables were found to express a moderate positive correlation with Organization Performance, among others: Energy Efficiency, Customer Retention, and Sustainable Practices (with $r = 0.30-0.50$). These variables seem to have the potential to drive these sustainability-related performances and should consequently be examined statistically by the application of predictive or inferential modeling. Interestingly, the revenue variable was observed to have a very low correlation with most dimensions of sustainability. This observation shows that the improvement of financial turnover does not necessarily have an effect on the improvement of sustainability outcomes,

especially when it comes to the subject of MSMEs in the automotive components industry. This note proves not only the hypothesis about the fact that sustainability maturity is not merely an economic scale but also a strategic prioritization. The correlation heatmap provided below in the formula is a measure of linear correlation between two variables.

$$r_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

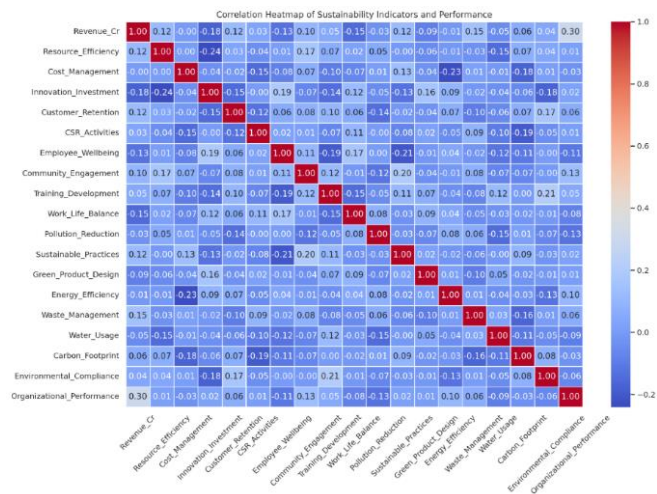


Figure 4. Bivariate Analysis – Correlation Heatmap

4.3. Bubble Plot Analysis

To get a deeper insight into the relationships between the leading sustainability indicators and organizational performance, the bubble plots were used as visual representations of the relationships between the selected pairs of variables with strong or moderate correlation according to the Pearson correlation table. This was goal-oriented and aimed at establishing whether the paired indicators' pattern could support or prove the basic assumptions regarding performance prediction and strategic alignment in MSMEs. Fig. 5 displays four bubble graphs showing how the variables of the following pairs are related: (i) Resource Efficiency with Cost Management, (ii) Employee Wellbeing with Training & Development, (iii) Customer Retention with Organizational Performance, and (iv) Energy Efficiency with Organizational Performance. Every plot gives an idea about how the association between the variables of interest is nature and orientation-wise. The equation of the bubble plot given below shows the relationship between two numeric variables for each point [30].

$$(x_i, y_i) \quad (7)$$

$$i = 1, 2, 3, \dots, n$$

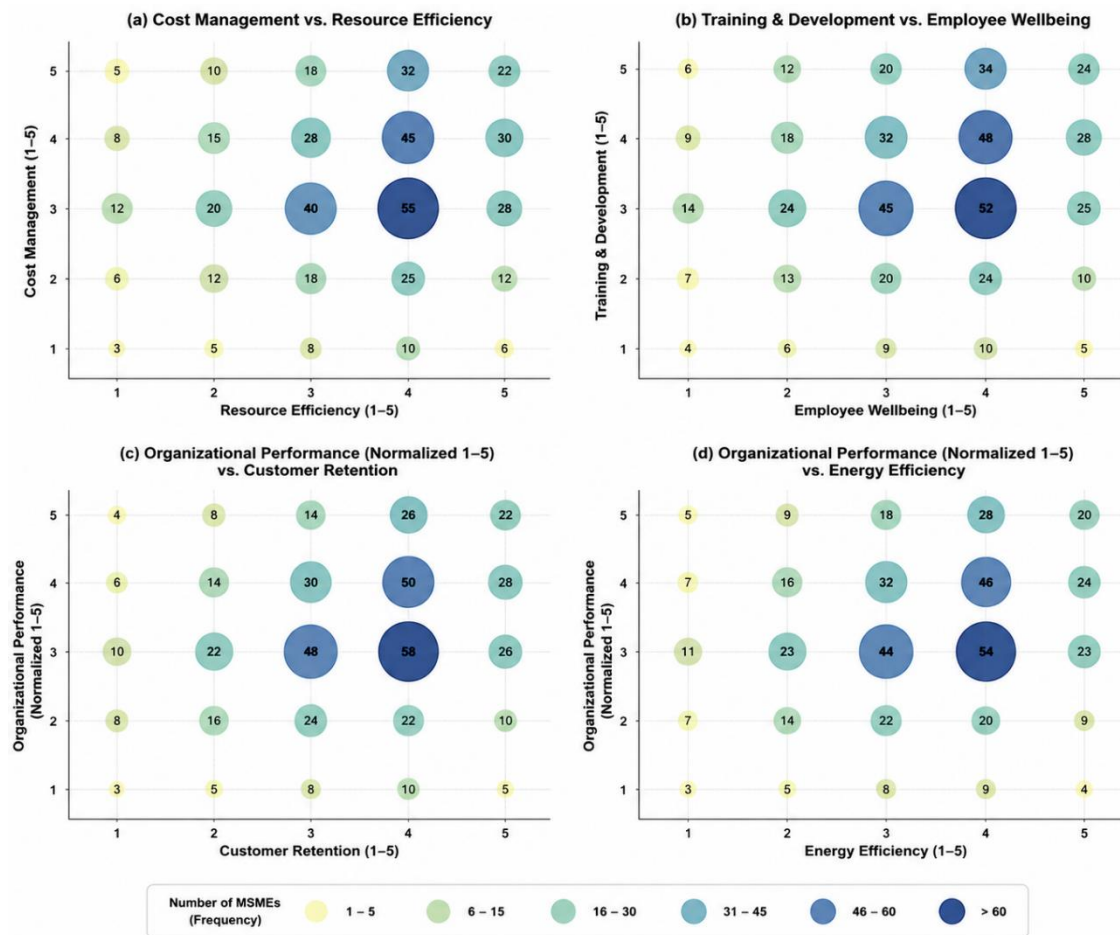


Figure 5. Bubble plot visualization.

Frequency bubble plots were used to analyze the relationships between the principal sustainability indicators, as an alternative to traditional scatter plots. Most of the variables were measured on five-point Likert scales, which meant that many enterprises had the same combination of responses and thus observations overlapped. The bubble plots overcome this by summarizing the same coordinate pair and using proportional bubble size to indicate how many times that value occurred. The visualization thus retains the original ordinal responses and offers a better insight into the concentration of the responses and the association patterns between sustainability indicators.

Fig. 5 shows the relationships between the principal sustainability indicators in a pairwise manner by the use of frequency bubble plots, where the size of the bubbles represents the number of MSMEs that answered a specific combination of the indicators with a similar Likert scale value. This visualization is especially useful to reduce the overplotting that often occurs with data represented in an ordinal manner on a survey. The Cost Management versus Resource Efficiency bubble plot shows a strong cluster of enterprises around the moderate-to-high level of responses (scores 3-4), with persons who reported higher levels of resource efficiency overall tending to report higher levels of cost management. The bigger bubbles seen in these response categories indicate that there is a strong link between operational resource optimization and financial efficiency, indicating the implementation of lean

manufacturing practices and efficient use of resources in automotive MSMEs.

Likewise, the Training & Development versus Employee Wellbeing bubble plot shows that most of the enterprises are grouped around the higher ratings on the Likert scale, indicating a positive relationship between Workforce Development and Employee Wellbeing. More respondents in the upper-middle categories scored high, as indicated by the number of larger bubbles, suggesting that organizations with a focus on employee training and organizational commitment. The results highlight the importance of ongoing skill development and employee engagement practices in enhancing the social sustainability aspect of MSMEs.

A clear positive trend is also seen between Customer Retention and the Normalized Organizational Performance Index. It can be seen that the bigger bubbles are located around customer retention scores of 4 - 5 and normalized organizational performance scores of ~ 3 - 5, suggesting that the companies with better customer retention scores tend to perform better in their organizations. This finding confirms previous studies which show that customer-oriented business practices can play a major role in the long-term sustainability and competitiveness of the organization. Similarly, the Energy Efficiency versus Normalized Organizational Performance bubble plot shows a

moderately positive correlation; energy-efficient practices tend to be more efficient across the board, and conversely, organizations that are more efficient with energy tend to perform better [31]. While the spread is wider across the values than the other variable pairs, the majority of the observations lie in the higher performance area, indicating that there is a positive relationship between energy management and sustainable development of the organization, especially for automotive manufacturing that is resource intensive.

4.4. Group-wise Insights and Segment Analysis

To explore the trends in sustainability practices further, a segment-wise analysis was carried out by dividing the companies into segments in terms of location and employee size. To specify the systemic trends and regional or organizational differences, the means of the primary sustainability indicators, namely, Organizational Performance, Energy Efficiency, Customer Retention, and Resource Efficiency, were computed within every subgroup.

4.4.1. Insights by location

Significant geographical variation was observed in terms of sustainability performance of the surveyed firms. Table 2 shows the average sustainability indicator scores of the firms based on location. The average results for organizational performance and energy efficiency were significantly better for firms in Pune and Chakan, suggesting stronger alignment with the sustainability practices in these industrial belts. This could be due to better infrastructure, availability of skilled manpower, or dynamic regionalization policy in these regions. Conversely, those organizations which have their base in Satara and Kolhapur have shown lower performance rates as measured, implying that there are possible areas where capacity-building activities could be done, or specific sustainability training programs could be carried out.

Table 2. Average Sustainability Indicator Scores by Location.

Location	Organizational Performance	Energy Efficiency	Customer Retention	Resource Efficiency
Chakan	3.92	3.63	3.62	3.65
Kolhapur	3.63	3.48	3.37	3.52
Nashik	3.76	3.41	3.49	3.45
Pune	3.89	3.57	3.56	3.69
Satara	3.59	3.38	3.33	3.38

Note: Pune and Chakan reported the highest sustainability performance, while Kolhapur and Satara exhibited slightly lower scores across key indicators.

4.4.2. Insights by employee size

The employee size was used as a proxy variable for the capacity and maturity of the organization. The impact of organizational size on sustainability performance was examined through grouping of firms according to employee size as shown in Table 3. Firms with over 200 employees recorded the highest mean across all sustainability indicators under consideration, with an average of 3.72 and an energy efficiency score of 3.79. This tendency indicates that larger companies tend to institutionalize changes towards living in a sustainable way; it can be explained perhaps by the availability of resources,

formal governance approaches, and strategies based on compliance.

Enterprises with mid-sized companies, especially companies with 11-50 employees, scored low on resource efficiency and customer retention (mean scores 3.25 and 3.27, respectively). Interestingly, small business establishments with 1-10 employees had comparatively high-performance scores and lean operational models/localized market positioning, although their energy efficiency score was relatively low (3.17), presumably due to the lack of technological investments. This knowledge can be regarded as the basis for custom policy solutions and benchmarking in the MSME segment of the automobile component sector of an industry characterized by the heterogeneity of sustainability adoption by firms of different sizes and locations.

Table 3. Average Sustainability Indicator Scores by Employee Size.

Employee Size	Organizational Performance	Energy Efficiency	Customer Retention	Resource Efficiency
1-10	3.75	3.17	3.65	3.33
11-50	3.38	3.43	3.25	3.27
51-100	3.80	3.43	3.28	3.47
101-200	3.73	3.52	3.61	3.63
200+	3.72	3.79	3.72	3.74

Note: Larger organizations (200+) consistently scored higher in sustainability indicators, whereas mid-sized firms (11-50) showed comparatively lower performance.

The findings also show a perception gap between frontline employees and senior management, suggesting that sustainability efforts are not uniformly experienced across the organization. While frontline employees think about sustainability from an operational standpoint, focusing on conditions in the workplace and day-to-day operations, senior management generally has more awareness of strategic sustainability objectives, regulatory compliance requirements, and long-term benefits of the organization. Such a gap can be due to a lack of employee engagement in sustainability planning, a lack of communication of the organization's sustainability goals, and a lack of access to sustainability-related training. Recent studies have shown similar findings: for successful sustainability implementation, in addition to managerial commitment, employees must actively participate, skills must be continually developed, and communication must be transparent. Hence, automotive MSMEs should consider supplementing strategic sustainability planning with formal employee engagement programs, periodic training sessions, and bottom-up feedback mechanisms to boost the alignment of the company and enhance the effectiveness of the ESG implementation.

The group-wise analysis shows that regional characteristics and organizational capacity impact sustainability performance. Companies with a strong sustainability score were consistently found in existing industrial clusters, indicating that industrial ecosystems, the accessibility of infrastructure, and the maturity of supply chains have a positive impact on ESG implementation [32]. Likewise, there was a tendency for larger companies to perform better in sustainability due to their higher financial

resources, formal governance structure, and sustainability initiatives. Similar findings have been noted in recent research on MSME sustainability, which shows that enterprise size and capability to organize still play a role in determining the level of maturity. The findings underscore the need for policy initiatives and capacity building at the regional level, targeting especially small businesses, which are not located within a major industrial village.

4.5. Visual Analysis of Top-Performing vs. Underperforming Companies

A comparative analysis of the normalized organizational performance index was subsequently carried out to examine in more detail the spread of sustainability performance in the enterprises surveyed. The organizational performance indicator is a composite sustainability index built out of several dimensions of economic, environmental, and social performance and was then normalized on a common 1-5 scale using the min-max normalization technique, to be consistent with the original Likert-scale indicators. The normalized organizational performance index is shown in Fig. 6 as the top and bottom ten performing companies. The top-performing enterprises scored within the range of 4.2-5.0 on normalized scores, which shows that they performed well on a number of factors related to sustainability, such as customer retention, resource efficiency, employee development, energy efficiency, etc. This usually involved more advanced integration of sustainability practices into operational and strategic decisions, balancing economic competitiveness with environmental responsibility.

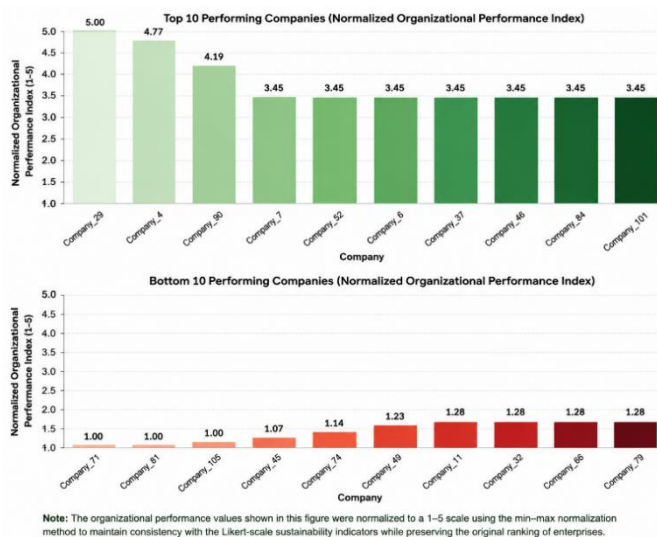


Figure 6. Top 10 and Bottom 10 performing companies based on the normalized organizational performance index (1–5 scale).

The lowest 10 companies, on the other hand, had normalized organizational performance scores ranging from 1.0 to 2.5, indicating relatively low levels of sustainability efforts. These organizations tended to have poorer scores for corporate social responsibility (CSR) activities, employee wellbeing, investment in innovation, and optimization of resources, which

suggests areas to improve organizational sustainability capabilities. While some of the less successful companies were micro and small enterprises, analysis indicates that it is the effective incorporation of sustainability practices rather than enterprise size that impacts organizational performance. This discovery underscores the need for strategic planning, ongoing employee training, innovation, and resource management to enhance the long-term sustainability of organizations.

The comparative findings also show a clear distinction between organizations with organized sustainability practices and those that practice sustainability in a relatively limited manner. The successful companies consistently scored the highest in customer engagement, operational effectiveness, and workforce development, while the less successful companies did not perform as well in innovation, employee capability development, and sustainability governance. These results are supported by recent research that has found that SMEs that implement structured sustainability management systems also tend to attain a higher degree of maturity and better long-term business performance than those that only implement sustainability systems for regulatory compliance purposes [33]. Therefore, sustainability cannot be considered a standalone environmental initiative, but rather an integral part of the organizational strategy, which can help MSMEs improve their competitiveness, operational resilience, and sustainable growth.

4.6. Top Sustainability KPIs Influencing Organizational Performance

The EDA, based on responses from 127 sales companies in the automobile components industry, identified three prevailing sustainability indicators (which fall under one of the sustainability pillars). These indicators, namely Customer Retention (Economic), Resource Efficiency (Environmental), and Training & Development (Social), are summarized in Table 4.

Customer Retention was found to be the best indicator of the performance question, with a correlation coefficient of about 0.52 with the measure of organizational performance. There was also a consistent effort by high-performance companies that showed higher customer retention, most of which was characterized by dependable service, effective relationship management, as well as sustainable products. It is worth noting that the organizations in Pune and Chakan were doing better in this aspect than those in Satara and Kolhapur, which may be due to some geographic and infrastructural benefits. Customer retention can be described as cost-saving and revenue stability, which are the economic implications of customer retention. This is one of the aspects that make retention a strategically important tool of MSMEs.

The second most influential KPI was Resource Efficiency, correlated with ~0.50. It depicts the capacity of an organization to maximize energy, water, and substance usage. The analysis revealed that the leading firms, especially those with lean manufacturing procedures and monitoring programs, were in a better position to control prices and ensure environmental issues. This element of the environment is especially important when considering that MSMEs are experiencing growing

regulatory and market pressure to conform to the rules of sustainability.

Training & Development, an important social indicator, ranked third with a correlation of ~ 0.46 . Companies that invested in employee skill development not only reported better workforce morale and retention rates but also performed better in innovation and operational efficiency. However,

even inasmuch as small firms required a structured HR system, they did not have it in order, which resulted in poor training and little internal capacity-building. These results prove once again the mutual dependence of the economic, environmental, and social pillars in ensuring organizational success. The special attention to these KPIs helps MSMEs to create resilient and high-performing businesses that align with long-term sustainability principles.

Table 4. Top three sustainability KPIs influencing organizational performance.

Rank	KPI	Sustainability Indicator Type	Correlation with Performance	Key EDA Insights	Challenges Identified	Recommended Actions
1	Customer Retention	Economic	~ 0.52	Strong upward trend in performance Clustering of high-retention firms High in large firms Positively aligned with cost management	Poor retention in firms <50 employees Weaker performance in Kolhapur & Satara	Customer feedback systems Long-term SLAs Leverage CSR to build brand trust
2	Resource Efficiency	Environmental	~ 0.50	High among Chakan and Pune-based units Correlated with Employee Wellbeing	High variation in mid-tier firms Lack of formal waste/process tracking	Lean manufacturing Energy/water audits Technology upgrades for monitoring
3	Training & Development	Social	~ 0.46	Stronger in top-performing companies	Underprioritized in small enterprises Lack of formal HR frameworks	Incentivized training Strategic workforce planning Modernization of internal HR practices

The identified KPIs provide an overall consensus that sustainability of an organization needs a balance between three aspects: economic, environmental, and social. Customer Retention boosts organizational learning and innovation capacity, Resource Efficiency promotes competitiveness and lower environmental impact of operations, and Training and Development improves long-term business resilience [33]. These indicators are not stand-alone indicators but rather indicators that support each other for better organizational performance. Recent literature on sustainability also highlights that successful MSMEs combine customer orientation with resource optimization, technological capability, and human resource development into a single sustainability strategy. Hence, the design of sustainability assessment frameworks, ESG capacity-building programs, and targeted support initiatives for automotive MSMEs should be geared towards these KPIs.

The findings above prove the point that sustainability cannot be treated as a unitary concept, but it is a multidimensional construct; that is, operational excellence, customer-centricity, and human capital development determine sustainability. Allowing the concentration of efforts on the three most effective KPIs such as Customer Retention, Resource Efficiency, and Training and Development, MSMEs will help organize the performance of the organization at the highest level and meet global environmental sustainability goals.

4.7. Clustering Analysis

An unsupervised K-means clustering analysis was conducted using the principal sustainability indicators to further explore the heterogeneity of sustainability performance of the participating automotive MSMEs. The goal of this analysis was to discover enterprises with similar sustainability characteristics and to investigate the presence of naturally occurring groups of enterprises along their performance. In the distance calculation process, the contribution of each sustainability variable was made equal by standardizing all the variables prior to clustering.

For this, the optimal number of clusters was selected using the Elbow Method and Silhouette Score as shown in Fig. 7. The Elbow curve shows that as the number of clusters increases, the within-cluster error continues to decrease but at a decreasing rate, suggesting that there is diminishing return with an increase in the number of clusters after 2 clusters. Similarly, the Silhouette Score remains fairly consistent starting around $k = 4$, indicating that this number of clusters offers a good balance between cluster compactness and separation. Four clusters were chosen as the best solution for further analysis based on the complementary validation criteria.

The resulting cluster distributions are shown in Fig. 8, using enterprises as visualization objects in four dimensions of sustainability: economic dimension, social dimension, environmental dimension, and sustainability index. The red star is the centroid of each respective cluster and each colour is one cluster. The clustering results show that enterprises are clearly separated with respect to their sustainability maturity levels.

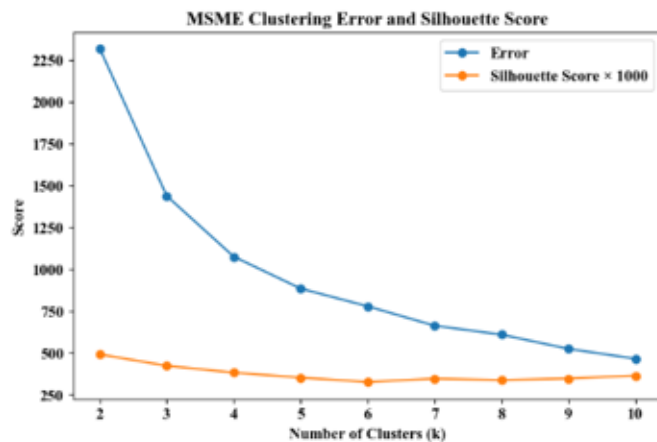


Figure 7. Optimal Number of Clusters using the Elbow method and silhouette score.

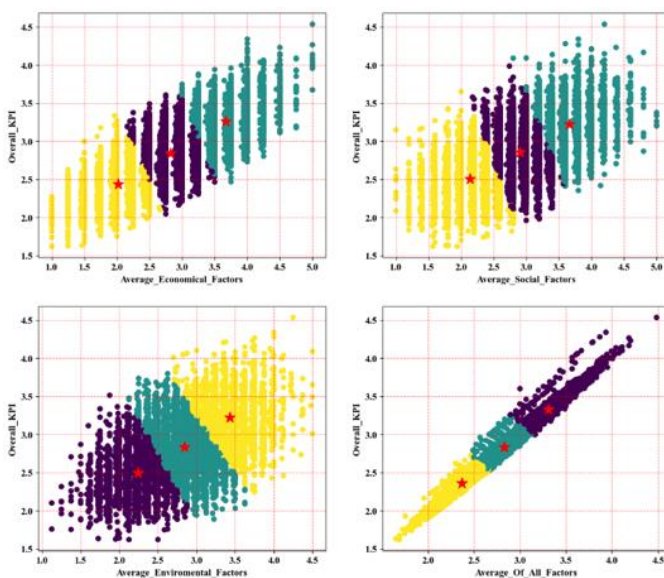


Figure 8. K-means clustering results showing enterprise distribution across economic, social, environmental, and overall sustainability dimensions. Red stars indicate cluster centroids.

In terms of economic sustainability, companies in the top-performing cluster show consistently higher organizational KPI values, which depict better economic management, investments in innovation, customer retention, and cost optimization. In contrast, companies in the least favorable cluster tend to have lower sustainability scores, which could mean that there are areas for improvement when it comes to resource allocation and operational efficiency.

Similar patterns emerge in the social sustainability clustering. The higher employee wellbeing, workforce development, and organizational engagement indicators cluster around higher performance organizations. This observation indicates that investments in human capital can play a significant part in an organization's sustainability, thereby reinforcing earlier findings on the importance of employee-centered sustainability approaches.

Clustering by environmental sustainability indicators demonstrates higher dispersion among enterprises, indicating variations in their approaches to energy efficiency, waste management, and utilization of environmental resources. Based on this finding, it is clear that the occurrence of environmental sustainability adoption is not homogeneous among automotive MSMEs, but rather has maturity levels of environmental sustainability adoption that vary significantly between organizations.

Last but not least, the clusters obtained by applying the overall sustainability factors demonstrate the best separation between enterprises, meaning that the overall sustainability factors have a better discriminatory power than each sustainability dimension. The relatively small clusters and good centroid distances indicate a good separation of the organizational sustainability performance levels by multivariate sustainability indicators.

4.8. Qualitative Insights from CEO Interviews

To add to the quantitative findings, thematic analysis was done for the semi-structured interviews with five Chief Executive Officers of various automotive MSMEs. The interviews revealed that low budget, limited technological capability, lack of workforce skills, and growing customer demands for ESG compliance are all key factors in the implementation of sustainability. Most of the participants pointed out that resource efficiency and customer retention receive the closest attention because of the immediate operational and financial advantages, although carbon footprint reduction and advanced sustainability reporting are often pushed to the back of the line because of the lack of resources.

The interviews also showed that employees' awareness of sustainability practices varies widely throughout organizational levels. Sustainability efforts were mostly seen as an investment in the business by senior management, and mostly related to operational responsibility by frontline staff. This finding confirms the quantitative findings of the role-wise comparison in Fig. 1, which consistently show that senior managers attributed higher sustainability scores than operational personnel.

Moreover, several CEOs pointed out that government incentives, technical guidance, and digital monitoring tools could significantly help with sustainability adoption among MSMEs. These qualitative observations complement the quantitative results and help to explain the practical difficulties of implementing sustainability in the resource-limited manufacturing businesses.

5. Conclusion

This study examined the sustainability performance of MSMEs in the automotive component sector in Maharashtra using a data-driven analytical framework. Using EDA and clustering methods on survey data from 127 firms, the study identified significant insights with implications for policy and industry practice.

The analysis showed that Customer Retention was the most significant economic Key Performance Indicator (KPI), which

has a strong positive correlation with overall organizational performance. Resource Efficiency proved to be highly connected with environmental sustainability and cost management, whereas Training and Development, as the social dimension of sustainability, was strongly related to employee well-being and innovation performance. Geographical and organizational heterogeneity was also observed, with the firms in Pune and Chakan doing better as compared to those in Kolhapur and Satara. Likewise, bigger companies (200+ workers) always scored higher in terms of sustainability, which indicated greater resource capability and strategic management, and smaller companies showed lower results in such aspects as CSR and Energy Efficiency. Role-based comparisons also demonstrated the existence of perceptual gaps, with frontline workers scoring lower on sustainability KPIs than senior management. The additional clustering analysis demonstrated that automotive MSMEs exhibit distinct sustainability maturity profiles rather than forming a homogeneous population. The identified clusters provide a practical basis for designing targeted sustainability interventions and benchmarking organizations according to their overall sustainability performance.

Overall, the research highlights the potential of data visualization and EDA as effective means of closing the gap between sustainability knowledge and managerial decision-making. The results recommend that MSMEs should implement systematic KPI monitoring tools, reinforce employee training programs, and improve customer engagement and resource management practices. The resulting analytical framework illustrates a replicable, evidence-based method of sustainability assessment in developing economies. Future research can include machine learning models in predictive sustainability analysis and expand the study area regionally to facilitate cross-sectoral and inter-state comparison.

Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this manuscript.

Use of LLM and AI

The authors used ChatGPT (OpenAI) to improve language clarity, grammar, and readability during manuscript preparation. The authors reviewed and approved the final content and take full responsibility for the accuracy and integrity of the manuscript.

Research Funding

The authors received no external financial funding for the research, authorship, or publication of this article.

Data Availability

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

Author Contribution Statement

Dipak Patil proposed the research problem, developed the theoretical framework, performed the computations, and prepared the original draft of the manuscript.

Dr. Geeta Rao verified the analytical methods, supervised the research work and findings, reviewed the results, and contributed to editing and finalizing the manuscript.

Both authors discussed the results, contributed to the interpretation of findings, and approved the final version of the manuscript for publication.

References

- [1] A. Gani, M. Asjad, and F. Talib, "Prioritization and Ranking of indicators of sustainable manufacturing in Indian MSMEs using fuzzy AHP approach," *Materials Today: Proceedings*, vol. 46, pt. 15, pp. 6631–6637, 2021. doi: <https://doi.org/10.1016/j.matpr.2021.04.101>.
- [2] D. Mathivathanan, V. Agarwal, K. Mathiyazhagan, T. Saikouk, and A. Appolloni, "Modeling the pressures for sustainability adoption in the Indian automotive context," *Journal of Cleaner Production*, vol. 342, Art. no. 130972, 2022. doi: <https://doi.org/10.1016/j.jclepro.2022.130972>.
- [3] N. Virmani, S. Bera, and R. Kumar, "Identification and testing of barriers to sustainable manufacturing in the automobile industry: a focus on Indian MSMEs," *Benchmarking: An International Journal*, vol. 28, no. 3, pp. 857–880, 2021. doi: <https://doi.org/10.1108/BIJ-08-2020-0413>.
- [4] S. B. Aji, N. Hanani, Suhartini, R. Asmara, and A. S. H. Wahyuningtyas, "Measuring the competitiveness of local specialty food agro-industrial MSMEs in Indonesia: Critical Factor Index and Data Envelopment Analysis approach," *Transactions of the Chinese Society of Agricultural Machinery*, vol. 56, no. 4, 2025. doi: <https://doi.org/10.62321/issn.1000-1298.2025.4.6>.
- [5] S. K. Mangla, R. Raut, V. S. Narwane, Zuopeng (Justin) Zhang, and P. Priyadarshinee, "Mediating effect of big data analytics on project performance of small and medium enterprises," *Journal of Enterprise Information Management*, vol. 34, no. 1, pp. 168–198, 2021. doi: <https://doi.org/10.1108/JEIM-12-2019-0394>.
- [6] D. Pandya and G. Kumar, "Industry 4.0 technologies for sustainable performance in Indian manufacturing MSMEs," in *International Conference on Industrial Engineering and Operations Management*, Istanbul, Turkey, Mar. 7–10, 2022, pp. 3532–3542. doi: <https://doi.org/10.46254/AN12.20220642>.
- [7] G. Vitale, A. Rosignuolo, and A. Riccaboni, "Sustainability performance measurement systems: A systematic literature review and research agenda for (a better) future," *Measuring Business Excellence*, vol. 29, no. 3, pp. 609–634, 2025. doi: <https://doi.org/10.1108/MBE-09-2024-0155>.
- [8] C. Malesios, D. De, A. Moursellas, P. K. Dey, and K. Evangelinos, "Sustainability performance analysis of small and medium sized enterprises: Criteria, methods and framework," *Socio-Economic Planning Sciences*, vol. 75, Art. no. 100993, 2021. doi: <https://doi.org/10.1016/j.seps.2020.100993>.
- [9] M. Chehimi, "Adopting the sustainability balanced scorecard in small and medium-sized enterprises: Strategic integration of environmental management or symbolic compliance?," *Journal of Environmental Management*, vol. 398, Art. no. 128421, 2026. doi: <https://doi.org/10.1016/j.jenvman.2025.128421>.
- [10] R. Chourasiya, S. Pandey, and R. K. Malviya, "Developing a framework to analyse the effect of sustainable manufacturing adoption in Indian textile industries," *Cleaner Logistics and Supply Chain*, vol. 4, Art. no. 100045, 2022. doi: <https://doi.org/10.1016/j.clscn.2022.100045>.
- [11] S. Singh, E. U. Olugu, and S. N. Musa, "Development of key performance measures for sustainable manufacturing in global SMEs," in *Designing and Implementing Global Supply Chain Management*, S. Joshi and R. Joshi, Eds. IGI Global, 2016, pp. 210–218. doi: <https://doi.org/10.4018/978-1-4666-9720-1.ch011>.
- [12] M. Gidage and S. Bhide, "Exploring the nexus between intellectual capital, green innovation, sustainability and financial performance in creative industry MSMEs," *Journal of Enterprising Communities: People and Places in the Global Economy*, vol. 19, no. 3, pp. 457–484, 2025. doi: <https://doi.org/10.1108/JEC-07-2024-0134>.
- [13] M. Das, K. Rangarajan, and G. Dutta, "Network and government intervention influencing sustainability and business growth of SMEs: a study with Indian MSMEs," *International Journal of Enterprise Network Management*, vol. 12, no. 2, pp. 131–152, 2021. doi: <https://doi.org/10.1504/IJENM.2021.116436>.
- [14] U. Illahi and M. S. Mir, "Assessment of transport sustainability using a hybrid approach: A comparison of four metropolitan cities of India," *Case*

- Studies on Transport Policy*, vol. 9, no. 2, pp. 703–714, 2021. doi: <https://doi.org/10.1016/j.cstp.2021.03.008>.
- [15] N. Laskar, T. K. Chakraborty, and S. G. Maji, "Corporate sustainability performance and financial performance: Empirical evidence from Japan and India," *Management and Labour Studies*, vol. 42, no. 2, pp. 88–106, 2017. doi: <https://doi.org/10.1177/0258042X17707659>.
- [16] G. Babber and A. Mittal, "Achieving sustainability through the integration of lean, agile, and innovative systems: Implications for Indian micro small medium enterprises (MSMEs)," *Journal of Science and Technology Policy Management*, vol. 16, no. 2, pp. 365–400, 2025. doi: <https://doi.org/10.1108/JSTPM-05-2023-0087>.
- [17] A. Hailelassie et al., "Empirical evaluation of sustainability of divergent farms in the dryland farming systems of India," *Ecological Indicators*, vol. 60, pp. 710–723, 2016. doi: <https://doi.org/10.1016/j.ecolind.2015.08.014>.
- [18] S. Agasty, F. Tarannum, and S. A. Narula, "Sustainability innovation index for micro, small and medium enterprises and their support ecosystems based on an empirical study in India," *Journal of Cleaner Production*, vol. 415, Art. no. 137793, 2023. doi: <https://doi.org/10.1016/j.jclepro.2023.137793>.
- [19] S. S. Sidhu, K. Singh, and I. P. S. Ahuja, "Role of maintenance practices in Indian SMEs: A literature review," in *Proceedings of the 1st International Conference on Intelligent Communication and Computational Research (ICICCR-2020)*, 2020. doi: <https://doi.org/10.2139/ssrn.3564871>.
- [20] T. L. Verma and D. K. Nema, "Role of micro, small and medium enterprises (MSMEs) in achieving sustainable development goals," *International Journal for Research in Engineering Application & Management*, vol. 4, no. 12, pp. 575–582, Mar. 2019. doi: <https://doi.org/10.61336/jmsr/26-02-09>.
- [21] A. Roy, N. Garai, and J. K. Biswas, "Exploration of urban sustainability in India through the lens of sustainable development goals," *Discover Sustainability*, vol. 4, no. 1, Art. no. 41, 2023. doi: <https://doi.org/10.1007/s43621-023-00158-2>.
- [22] K. Sharma, S. S. Walia, J. Kaur, R. Bhagat, and J. Layek, "Farming system typology construction for the adoption of new technologies in north-west India," *The Indian Journal of Agricultural Sciences*, vol. 95, no. 3, pp. 355–360, 2025. doi: <https://doi.org/10.56093/ijas.v95i3.162807>.
- [23] K. M., Mahesh P. S. Aithal, and Sharma K. R. S., "Government initiatives and digital innovation for Atma Nirbhar MSMEs/SMEs: To achieve sustainable and inclusive economic growth," *International Journal of Management, Technology, and Social Sciences*, vol. 8, no. 1, pp. 68–82, 2023. doi: <https://doi.org/10.47992/IJMTS.2581.6012.0256>.
- [24] E. Herman and K. E. Zsido, "The financial sustainability of retail food SMEs based on financial equilibrium and financial performance," *Mathematics*, vol. 11, no. 15, Art. no. 3410, 2023. doi: <https://doi.org/10.3390/math11153410>.
- [25] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras & TensorFlow*, 3rd ed. Sebastopol, CA, USA: O'Reilly Media, 2022. [Online]. Available: <https://www.oreilly.com/library/view/hands-on-machine-learning/9781098125967/>.
- [26] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Routledge, 2018. doi: <https://doi.org/10.1201/9781315140919>.
- [27] W. S. Cleveland, *Visualizing Data*. Hobart Press, 1993. [Online]. Available: <https://dl.acm.org/doi/abs/10.5555/529269>.
- [28] J. W. Tukey, *Exploratory Data Analysis*. Reading, MA: Addison-Wesley, 1977. doi: https://doi.org/10.1007/978-3-031-20719-8_2.
- [29] S. Kannan and N. Gambetta, "Technology-driven sustainability in small and medium-sized enterprises: A systematic literature review," *Journal of Small Business Strategy*, vol. 35, no. 1, pp. 129–157, 2025. doi: <https://doi.org/10.53703/001c.126636>.
- [30] A. Martins, M. C. Branco, P. N. Melo, and C. Machado, "Sustainability in small and medium-sized enterprises: A systematic literature review and future research agenda," *Sustainability*, vol. 14, no. 11, Art. no. 6493, 2022. doi: <https://doi.org/10.3390/su14116493>.
- [31] A. S. Shaik, A. Nazrul, S. M. Alshibani, V. Agarwal, and A. Papa, "Environmental and economical sustainability and stakeholder satisfaction in SMEs: Critical technological success factors of big data analytics," *Technological Forecasting and Social Change*, vol. 204, Art. no. 123397, 2024. doi: <https://doi.org/10.1016/j.techfore.2024.123397>.
- [32] J. A. Clemente-Almendros, M. Vallej García, and M. Blanco-Hernández, "SMEs and sustainable practices: Identifying key factors from Spanish evidence," *SAGE Open*, vol. 15, no. 1, pp. 1–19, 2025. doi: <https://doi.org/10.1177/21582440251329631>.
- [33] J. Vázquez, S. Aguirre, E. Puertas, G. Bruno, P. C. Priarone, and L. Settineri, "A sustainability maturity model for micro, small and medium-sized enterprises (MSMEs) based on a data analytics evaluation approach," *Journal of Cleaner Production*, vol. 311, Art. no. 127692, 2021. doi: <https://doi.org/10.1016/j.jclepro.2021.127692>.