Analysis of Rear Differential Component Clustering in Transmission Systems Using Hierarchical Cluster Analysis with and without Procurement Strategy Matrix Variables

Saowalak Sombunsook[1], Kittiwat Sirikasemsuk[1,*] and Kanogkan Leerojanaprapa[2]
[1] Department of Industrial Engineering, School of Engineering, King Mongkut's Institute of Technology Ladkrabang, Lat Krabang, Lat Krabang, Bangkok, 10520, Thailand
[2] Department of Statistics, School of Science, King Mongkut's Institute of Technology Ladkrabang, Lat Krabang, Lat Krabang, Bangkok, 10520, Thailand
***Corresponding Author E-mail:** kittiwat.sirikasemsuk@gmail.com
Received: Jan 29, 2025; **Revised:** Apr 28, 2025; **Accepted:** Apr 29, 2025

Abstract

The automotive industry has faced significant challenges due to the large number of Tier 2 suppliers for Rear Differential components, with 17 suppliers providing 32 different parts. This situation has resulted in increased production costs and more complex supply chain management. This study aimed to analyze the clustering of Rear Differential components in transmission systems using Hierarchical Cluster Analysis, with the goal of supporting cost reduction in the automotive industry. Two clustering models were compared: Model 1, which excluded procurement strategy matrix variables (Special Requirements, Raw Material Grade, Raw Material Type, Manufacturing Process, Tier 2 Supplier Information, and Company Location), and Model 2, which incorporated an additional variable related to the Procurement Strategy Matrix. The decision criteria for determining the optimal number of clusters were based on four key factors: 1) Product design, 2) Characteristics, 3) Materials, and 4) Manufacturing. The clustering results for both models revealed the same optimal number of 13 clusters; however, the similarity matrix between the clusters differed. Furthermore, the number of members within each cluster varied. Based on the criteria for determining the optimal number of clusters, Model 2, which included the Procurement Strategy Matrix variable, demonstrated superior clustering efficiency compared to Model 1. Ultimately, this research identified 13 optimal clusters, reducing the number of Tier 2 suppliers from 17 to 13, representing a 23.53% reduction.

Keywords: Hierarchical clustering, Supplier selection, Cost reduction

1. Introduction

This research examined a pickup assembly company from a case study in the eastern region of Thailand. The study focused on the transmission system, specifically the Rear differential, a commonly used component in general vehicle models. This resulted in variations in component parts in terms of design. The Rear differential accounted for the highest procurement cost, representing 37% of short-lot order expenses, and had a high production volume in 2023.

The findings revealed that the company procured rear differentials from a Tier 1 supplier, who, in turn, sourced 32 parts of components from 17 Tier 2 suppliers. When considering suppliers providing fewer than two component types per supplier, the Tier 1 supplier had to identify alternative suppliers. This led to higher expenses, extended negotiation times, and affected the automaker's cost management and competitiveness.

This research emphasized the importance of supply chain management and aimed to classify Rear differential components using the Hierarchical Cluster Analysis technique. The classification process utilized two models: Model 1, which excluded six procurement strategy matrix variables, and Model 2, which included these six variables along with an additional procurement strategy matrix variable, totaling seven variables.

The seven variables selected for Model 2 were chosen based on their relevance to procurement decision making and their direct or indirect impact on engineering design criteria, namely functionality, material compatibility, and manufacturing process. These variables are commonly used in strategic procurement frameworks and include factors such as cost impact, supply risk, supplier availability, volume flexibility, technical complexity, and lead time. Each of these variables influences how components can be grouped effectively to support engineering efficiency and supplier consolidation. By incorporating these variables, Model 2 provided a more comprehensive and realistic classification of parts, aligning procurement strategies with technical specifications and design constraints.

Once the Rear differential components with similar characteristics were grouped, four key engineering design criteria were applied to evaluate the appropriateness of grouping related components within each cluster. Subsequently, existing manufacturers with production capabilities for each group were identified to create a revised procurement list. These capable manufacturers proposed prices for

DOI: 10.55003/ETH.420207 ISSN: 2985-1688 (Online)

components within their respective groups (based on the same order quantities), serving as a guideline for supplier selection and cost reduction for the company.

1.1 Theory of Cluster Analysis

The study of cluster analysis using the Hierarchical Cluster Analysis technique is a widely used method in data analysis and statistics. This technique is applied to divide cases (referring to items or organizations) into subgroups of two or more.

The basic assumptions of the hierarchical clustering method are as follows [1],[2]:

1) When dividing cases, the number of cases should not be too large (ideally fewer than 200).

2) It is not necessary to know the number of groups beforehand.

3) It is not necessary to know which case belongs to which group initially.

4) The variables used for grouping must include more than one variable, which can be qualitative, quantitative, or a variable with a single value.

Selection of Variables or Factors Influencing Case Differentiation. When selecting variables or factors expected to influence the differences between cases, qualitative variables can be used. These variables are assigned only two possible values: 0 and 1. A variable of interest is assigned a value of 1, while a variable not of interest is assigned a value of 0, as shown in Eq. (1) [2].

$$X_i = \begin{cases} 1 & \textit{if the characteristic is of interest} \\ 0 & \textit{if the characteristic is of interest} \end{cases} \quad (1)$$

When X_i It is the variable that influences the differentiation between cases.

The data used for clustering has varying ranges for each variable. Therefore, the data is transformed to standardize the values of each variable, ensuring consistency. This is done using standardization, as shown in Eq. (2) [3].

$$z_i = \frac{x_i - \bar{x}}{S} \quad (2)$$

When

Z_i : The standardized value of the i-th observation.

X_i : The original value of the i-th observation.

$\bar{x}$: The mean of the variable.

S : The standard deviation of the variable.

The process begins by defining one cluster as equal to the number of data points. Then, the similarity or distance between data points is considered, using the Euclidean distance formula to calculate the distance between the centroid of a cluster and all data points within that cluster. This is done using Eq. (3), which calculates the distance between the centroid and the data points. the procedure continues by merging two clusters by measuring the distance between their centroids (Centroid Clustering). The distance of the new combined cluster is then calculated. This process is repeated until only one cluster remains. Once the clustering is completed, no further changes will occur, and the number of clusters is evaluated to determine the most appropriate number of clusters to use. This method ensures that the data is grouped in the most meaningful way, allowing for the selection of the optimal number of clusters based on the resulting structure [1].

$$\text{Distance } (c_i, p) = \sqrt{\sum (c_{i_n} - p_n)^2} \quad (3)$$

The centroid of group j is defined as: $c_i = (c_{i_1}, c_{i_2}, \ldots, c_{i_n})$ The data points in group j are represented as: $p = (p_1, p_2, \ldots, p_n)$

Hierarchical clustering is particularly suitable for this study, which involves real automotive components, due to the relatively small sample size of 32 components. This sample size is similar to those used in several other studies. Additionally, this technique does not require prior knowledge of the number of groups or their characteristics, making it ideal for researchers' analysis. Its step-by-step grouping process allows researchers to examine the natural structure of the data, ensuring that the resulting clusters are meaningful and aligned with both engineering design criteria and strategic perspectives

1.2 Theory of the Procurement Strategy Matrix

The Procurement Strategy Matrix is a strategic activity of a company, involving the management of various processes abroad to meet the company's needs for procuring goods and services essential for product manufacturing (direct) or organizational operations (indirect). It is divided into four components as follows [4],[5]:

1) High Importance to the Organization but Low Financial Value or Impact (High Level & Low Value)

This group is referred to as Engineering Items or items with a Limited Number of Suppliers, used in specialized production processes.

2) High Importance to the Organization and High Financial Value or Impact (High Level & High Value)

This group is referred to as Unique Items, components contributing to the final product, or items that involve Integration with Suppliers.

3) Low Importance to the Organization but High Financial Value or Impact (Low Level & High Value)

This group is referred to as Leverage Items or Basic Production Materials.

4) Low Importance to the Organization and Low Financial Value or Impact (Low Level & Low Value)

This group is referred to as Routine Purchases or items with General Standard Specifications.

1.3 Literature review

The literature review for Cluster Analysis was presented as follows:

S. Plungsri and K. PuntusavaseK [6] demonstrated about stated that product size clustering was conducted using three datasets, each containing 437 data points. The analysis determined the maximum number of products that could be stacked on a pallet. The results showed that costs could be reduced by 2.9 million baht (13.43%), storage space could be reduced by 2.19 million cubic inches (48.65%), and transportation volume could be increased by 5,200 boxes per round (86.7%).

N. Wiroonsri, et al. [7] conducted a Nonhierarchical Cluster Analysis using the K-means method to segment 588 pet food customers based on transaction data. The analysis identified eight customer groups and created 15 new variables derived from the transaction data. These variables were utilized to classify customers and develop sales strategies.

Z.H. Che [8] proposed a mathematical optimization model for clustering and supplier selection. The clustering results were then used to identify suitable supplier combinations using the Analytic Hierarchy Process (AHP). The study analyzed data from seven components, with ten operational variables and ten suppliers. The objective was to select one supplier for each component from the ten available options.

A.T. Almaktoom, et al. [9] addressed the measurement of service levels in supply chains and the development of robust optimization methods for system design. Results from a case study demonstrated that implementing the proposed robust optimization approach enabled the system to achieve a 90% service level while minimizing performance costs and the impacts of uncertainty.

D. Marini and J.R. Corney [10] proposed the use of a Selection Matrix to support design processes involving manufacturing methods that produced shapes close to the desired final form. The primary decision-making criteria for supplier selection and ranking were established based on expert opinions, which focused on factors such as product characteristics, geometry, production processes, and materials. This approach was aimed at reducing raw material usage and machinery requirements while aligning with production volume and material demands.

A. Srirattanapraphan and W. Songpan [11] stated that user clustering based on actual usage behavior was conducted to develop a method for optimizing the use of data within the registration system. The users were grouped into six categories using hierarchical cluster analysis.

U. Moonpen and S. Mungsing [12] stated that the optimization variables for clustering were used to analyze the necessary characteristics of travel packages. The clustering methods employed included the K-means clustering method, hierarchical clustering, random clustering, and the DBSCAN clustering method.

A. Araveeporn and J. Promsanga, [13] stated that the criterion used to compare the performance of clustering was the average difference in data between groups for the K-means clustering method applied in three different ways.

W. Khalid, Z. N. Lee Herbert-Hansen Nadja [14] showed the framework, based on k-means clustering, aids international location decisions by providing objective, quantitative outputs. It is fast, flexible, and allows customizable indicators to suit decision-makers' preferences.

After reviewing the literature, the researchers found that previous studies [3–13] focused on selecting appropriate clustering techniques to identify suitable groups for companies to develop or create strategies tailored to each group. For this reason, the researchers applied clustering to the Rear differential components in the transmission system of a case study company, a pickup truck assembly plant in the eastern region. The focus was on sustainability and alignment with the company's values, creating key strategies to enhance competitive advantage in supplier selection [15]. This approach integrates engineering design considerations with the Procurement Strategy Matrix, differing from the majority of past research, which typically focused on selecting suppliers based on a single variable. This clustering approach reflects the changes in the industrial world.

In reviewing previous research, it was found that some studies, such as [8], used K-means clustering to categorize suppliers based on product type, customer demand, production cost, product quality, and production time. However, it did not consider design factors, which are crucial for capturing the complexity of actual production processes. Moreover, K-means clustering produces fixed groupings based on a predefined number of clusters, which differs from Hierarchical Cluster Analysis (HCA). HCA is more suitable when the objective is to explore relationships among data points flexibly, providing a more detailed structure known as a dendrogram. Other research [2] incorporated technical data but did not integrate procurement strategies comprehensively. This study aims to fill these gaps by combining both engineering and procurement data to create clusters that not only support efficient supplier selection but also align with the strategic goals of the company.

This research involves variables where the exact number of groups is unknown and aims to segment based on qualitative variables. Therefore, the Hierarchical Clustering technique was chosen to group the Rear Differential components (Rear Assy, differential) by establishing appropriate criteria for clustering. The analysis focused on the similarity of data from variables that demonstrate the differences between each variable, which influences the clustering results.

2. Research Methodology

The research methodology involved studying component parts of the Rear Differential produced by Tier 2 suppliers using Hierarchical Cluster Analysis with the SPSS statistical software. The analysis involved inputting two models of variables, totaling seven types of variables, and creating 18 new sets of data, each with 32 types. The data was presented in tabular form. The detailed steps for applying the methodology with the program were as follows:

1) Data Collection and Variable Definition (see section 2.1)

2) Management of Qualitative Data and Creation of New Variables (see section 2.2)

3) Data Transformation into Standardized Values and Condition Checking (see section 2.3)

4) Clustering of Components Using Hierarchical Clustering Technique (see section 2.4)

5) Results of Clustering Rear Differential Components for Both Models (see section 3.1)

6) Comparison of Clustering Results for Both Models Using Distance Metrics (see section 3.2)

7) Results of Suitable Clustering for Rear Differential Components (see section 3.3)

8) Comparison of Suppliers Before and After Clustering, Including Order Costs (see section 3.4)

9) Conclusion (see section 4)

2.1 Data Collection and Variable Definition

The researcher collected raw data of Rear Differential Component Parts based on the design principles of 32 component parts from 17 Tier 2 suppliers. The data was gathered using the Procurement Strategy Matrix to evaluate components according to their importance and market complexity, as shown in **Figure 1.**

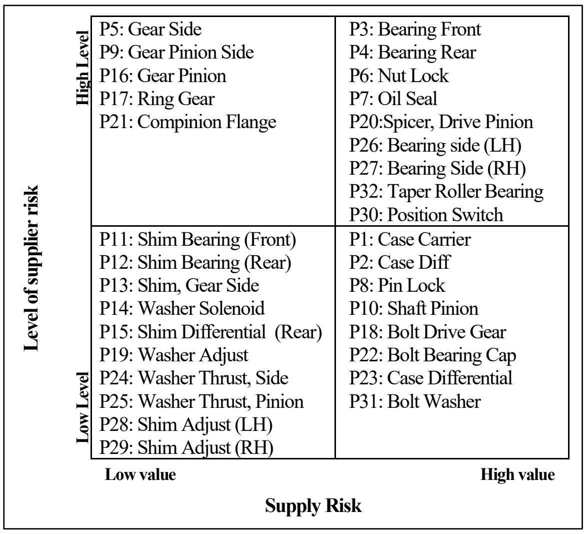

Figure 1 Rear Differential Components from the Procurement Strategy Matrix

The raw data collected was then used to create clustering variables for **Model 1**, which did not include the Procurement Strategy Matrix variables. The six variables used for clustering in this model were: Special Requirements, Raw Material Grade, Raw Material Type, Manufacturing Process, Tier 2 Supplier Information, and Company Location

Subsequently, the components data derived from the Procurement Strategy Matrix was used to create clustering variables for **Model 2**, as shown in **Table 1.**

2.2 Management of Qualitative Data and Creation of New Variables

The researcher managed the qualitative data by converting it into quantitative form. Variables of interest were assigned a value of 1, while variables of no interest were assigned a value of 0. The new variables were created as follows:

Model 1 created a total of 14 new variables, as shown in data sets 1–14.

Model 2 created a total of 18 new variables, adding the procurement strategy variables, with data sets 15-18. Each model generated 32 data points per new variable set, as shown in **Table 2.**

Table 1 Defining the Variable Models for Clustering Rear Differential Components

Variables of Model 1	Variables of Model 2
(1) Specific Requirement	(1) Specific Requirement
(2) Material Type	(2) Material Type
(3) Material Category	(3) Material Category
(4) Forming Process	(4) Forming Process
(5) Specific Data for Tier 2 Suppliers	(5) Specific Data for Tier 2 Suppliers
(6) Geographical Location for Tier 2 Suppliers	(6) Geographical Location for Tier 2 Suppliers
	(7) Procurement Strategy Matrix

2.3 Data Transformation into Standardized Values and Condition Checking

To handle non-numeric variables, the Binary Data technique was used to convert the data into a format suitable for clustering. The data was then standardized to reduce discrepancies caused by different measurement units. This transformation involved adjusting the values of all variables so that the minimum value became 0 and the maximum value became 1. The new variables were stored in the raw data file, preparing the data for clustering the Rear Differential components.

This research examined the basic conditions of the variables used for clustering by comparing the standardized data transformation for Model 1 and Model 2. It was found that Model 2 showed a greater difference in the stamping process variable than Model 1, while the Special Material variable in Model 2 had a lower value compared to Model 1.

These results indicated that Model 2 exhibited more similarity in terms of special materials than in the stamping process. The verification of these basic conditions was shown in **Table 3.**

Table 2 New Variables Derived from Model 1 and Model 2 Variables

Data Set	New Variables	Clustering Variables	
		Variable of Interest = 1	Variable of No Interest = 0
Data Set 1	Define Manufacturer Country	Manufacturer in Thailand	Not a Manufacturer in Thailand
Data Set 2	Define Stamping Type	Stamping	Not Stamping
Data Set 3	Define Casting Type	Casting	Not Casting
Data Set 4	Define Forging Type	Forging	Not Forging
Data Set 5	Define Completed Part Type	Completed Part	Not Completed Part
Data Set 6	Define Tier 3 Manufacturer Country	Manufacturer in Thailand	Not a Manufacturer in Thailand
Data Set 7	Define Material Type	Steel	Not Steel
Data Set 8	Define Material Characteristics	Long Steel Bar	Not Long Steel Bar
Data Set 9	Define Special Material	Special Material	Not Special Material
Data Set 10	Define Precision Requirement for Turning	Requires Precision	Does Not Require Precision
Data Set 11	Define Need for Property Improvement	Requires Property Improvement	Does Not Requires Property Improvement
Data Set 12	Define Surface Coating Requirement	Requires Surface Coating	Does Not Requires Surface Coating
Data Set 13	Define Specific Knowledge Requirement	Manufacturer Knowledge	Not Manufacturer Knowledge
Data Set 14	Define Workpiece Complexity	Complex Workpiece	Not Complex Workpiece
Data Set 15	Define Engineering Parts	Engineering Parts	Not Engineering Parts
Data Set 16	Define Specific Identity	Specific Identity	Not Specific Identity
Data Set 17	Define Basic Manufacturing	Basic Manufacturing	Not Basic Manufacturing
Data Set 18	Define Normal Purchase	Normal Purchase	Not Normal Purchase

Table 3 Comparison of the Standardized Values Using Two Model Variables

Standardized Values	Variables of Model 1		Variables of Model 2	
	Mean	Std. Deviation	Mean	Std. Deviation
Manufacturer Country	0.72	0.457	0.72	0.457
Forging Type	0.25	0.440	0.25	0.440
Casting Type	0.09	0.296	0.09	0.296
Stamping Type	**0.34**	**0.483**	**0.34**	**0.483**
Completed Parts	0.66	0.483	0.66	0.483
Tier 3 Supplier Country	0.13	0.336	0.13	0.336
Raw Material Type	0.94	0.246	0.94	0.246
Raw Material Characteristics	0.47	0.507	0.47	0.507
Special Material	**0.28**	**0.457**	**0.28**	**0.457**
Machining Precision	0.97	0.177	0.97	0.177
Property Improvement (Heat Treatment)	0.16	0.369	0.16	0.369
Coating/Painting Process	0.28	0.457	0.28	0.457
Special Requirements	0.09	0.296	0.09	0.296
Part Complexity	0.38	0.492	0.38	0.492
High Level & Low Value Parts	-	-	0.16	0.369
High Level & High Value Parts	-	-	0.25	0.440
Low Level & High Value Parts	-	-	0.28	0.457
Low Level & Low Value Parts	-	-	0.31	0.471

2.4 Clustering of Components Using Hierarchical Clustering Technique

The number of groups to be divided was defined as k groups, where the number of parts equaled the number of groups and clusters. In this study, there were a total of 32 types, which correspond to 32 clusters

The clustering process began with 2 clusters and increased progressively. Random representatives for clustering were designated as R1, R2, ..., Rk, and the results of each cluster were labeled C1, C2, ..., Ck. The members of each cluster, representing rear axle components, were denoted as P1, P2, ..., Pk.

As the number of clusters decreased, the distance or similarity between data points was recalculated, and this process was repeated until the members of each cluster no longer changed. The clustering process was based on the proximity of data points, where the closest data points were connected and grouped together. Subsequently, the next closest data points were connected, and the process continued.

The clustering sequence was illustrated using a dendrogram, showing the hierarchical results for Model 1 and Model 2, as depicted in **Figures 2–3.**

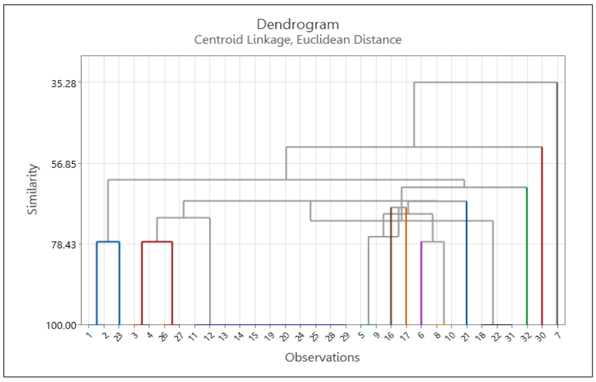

Figure 2 Dendrogram of Clustering by the variables of Model 1

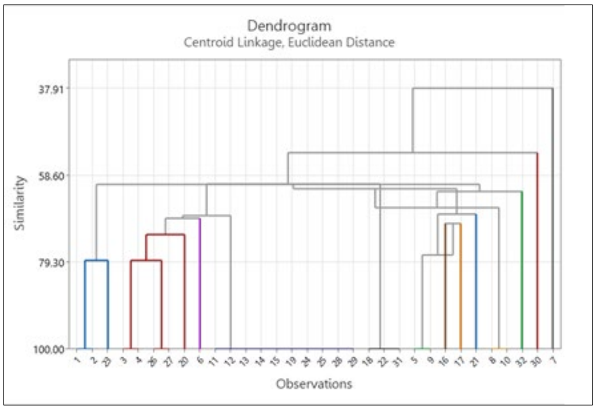

Figure 3 Dendrogram of Clustering by the variables of Model 2

3. Research results and analysis

The analysis of clustering results using the variables from Model 1 and Model 2 was summarized as follows:

3.1 Results of Clustering Rear Differential Components for Both Models

The analysis focuses on evaluating the appropriateness of the cluster using engineering-based design criteria and examining the stability of the clustering results. The discussion includes the use of different models, selected variables, and the optimal number of clusters (k).

Since the clustering results did not provide statistical values or hypothesis test results, the verification was conducted based on four key engineering-based design criteria for analyzing suitable clusters, which include:

1) Material used in production.
2) Usage Characteristics.
3) Design Characteristics.
4) Manufacturing process.

These criteria were used as a framework to assess the consistency of cluster membership. The verification process was conducted repeatedly until the cluster memberships remained unchanged and stable.

When $k = 13$ was applied, the clustering results from Model 1 and Model 2 differed. Specifically, component P20 in cluster C7 could not be clearly separated based on the variables used in Model 1. Although the cluster members shared similar attributes in terms of material and manufacturing processes, the design criterion did not distinctly differentiate the components in Model 1, leading to a lack of clarity in the grouping.

In the same random (R12), Model 2 met all four evaluation criteria with a 100% pass rate across all groups. This led to a more stable clustering result, where group members remained unchanged, and their characteristics could be identified according to the defined criteria. The use of Model 2 variables improved the accuracy of the clustering process by grouping components with identical or similar characteristics. By applying the appropriate clustering criteria and repeating the process until stability was achieved, the group structures were confirmed to be consistent. For confirmation of the optimal cluster, where $k = 14$ was applied, the results did not show a significantly clearer structure compared to $k = 13$. This confirmed that the optimal number of clusters was $k = 13$, considering the stability of the grouping and the effectiveness in classifying the components according to the specified criteria.

This indicated that the variables used in Model 2 had a significant impact on the clustering process. The inclusion of Model 2 variables allowed for more accurate grouping, as members with similar or identical characteristics were grouped. By applying the engineering-based design criteria and iterating until stable results were achieved, it was ensured that the members within each cluster remained unchanged. Details of the evaluation results are presented in **Tables 4–5**.

Table 4 Example of Clustering and Results of Component Members Using Model 1 Variables When k = 13

Random Grouping			Criteria for Appropriate Grouping Analysis				
Randomizations (Times)	Number of Groups (k)	Cluster Members (Parts)	Criteria 1: Material	Criterial 2: Characteristics	Criterial 3: Design	Criterial 4: Manufacturing	Ratio Passing criteria
R12 (k = 13)	C1	P1, P2, P23	✓	✓	✓	✓	98%
	C2	P3, P4, P26, P27	✓	✓	✓	✓	
	C3	P5, P9	✓	✓	✓	✓	
	C4	P6	✓	✓	✓	✓	
	C5	P7	✓	✓	✓	✓	
	C6	P8, P10	✓	✓	✓	✓	
	C7	P11, P12, P13 P14, P15, P19 **P20,** P24, P25, P28, P29	✓	✓	✗	✓	
	C8	P16	✓	✓	✓	✓	
	C9	P17	✓	✓	✓	✓	
	C10	P18, P22, P31	✓	✓	✓	✓	
	C11	P21	✓	✓	✓	✓	
	C12	P30	✓	✓	✓	✓	
	C13	P32	✓	✓	✓	✓	

Note: In particular, component P20 within cluster C7 could not be separated based on the variables in Model 1.

Table 5 Example of Clustering and Results of Component Members Using Model 2 Variables When k = 13

Random Grouping			Criteria for Appropriate Grouping Analysis				
Randomizations (Times)	Number of Groups (k)	Cluster Members (Parts)	Criteria 1: Material	Criterial 2: Characteristics	Criterial 3: Design	Criterial 4: Manufacturing	Ratio Passing criteria
R12 (k = 13)	C1	P1, P2, P23	✓	✓	✓	✓	100%
	C2	P3, P4, **P20**, P26, P27	✓	✓	✓	✓	
	C3	P5, P9	✓	✓	✓	✓	
	C4	P6	✓	✓	✓	✓	
	C5	P7	✓	✓	✓	✓	
	C6	P8, P10	✓	✓	✓	✓	
	C7	P11- P15, P19, P24, P25, P28, P29	✓	✓	✓	✓	
	C8	P16	✓	✓	✓	✓	
	C9	P17	✓	✓	✓	✓	
	C10	P18, P22, P31	✓	✓	✓	✓	
	C11	P21	✓	✓	✓	✓	
	C12	P30	✓	✓	✓	✓	
	C13	P32	✓	✓	✓	✓	

3.2 Comparison of Clustering Results for Both Models Using Distance Metrics

Hierarchical clustering analysis was performed to compare the two models with different sets of variables. Model 1 used 14 variables, while Model 2 included 18 variables. The clustering process was conducted with k = 13, based on the similarity or dissimilarity of the centroids (Centroid Clustering), the Euclidean distance between the centroid and all data points in that group was calculated. The distance for the new group is computed. This process continues until only one group remains, and once the grouping was complete, no further changes will occur.

The results revealed that adding more variables in Model 2 increased the distance between clusters, and the centroids of the variables showed greater differences, as illustrated in **Figure 4.**

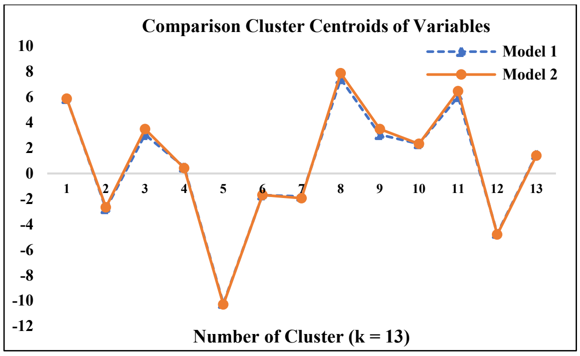

Figure 4 Comparison Cluster Centroids of Variables (when k = 13)

The comparison of the distances between cluster centroids revealed that Model 2 had higher centroid values than Model 1, indicating a broader and more diverse data distribution. The addition of more variables in Model 2 resulted in more distinct clusters and an increase in the number of clusters (k) enhances the clarity of the grouping process. The increased distance between the centroids suggests that the separation between clusters is better, with greater differentiation between them, which helps make the data grouping process clearer and more efficient. This enhanced the accuracy in distinguishing the members of the differential carrier assembly components, as shown in **Figure 5.** For example, the summation of Distances between Cluster centroids in Model 1 was 78.63, and Model 2 was 86.417, or Cluster C5 shows that Model 1 was 105.42 and Model 2 was 109.87

The larger distance between cluster centroids in Model 2 confirms that the grouping process has improved, contributing to higher accuracy in distinguishing the members of Rear different components.

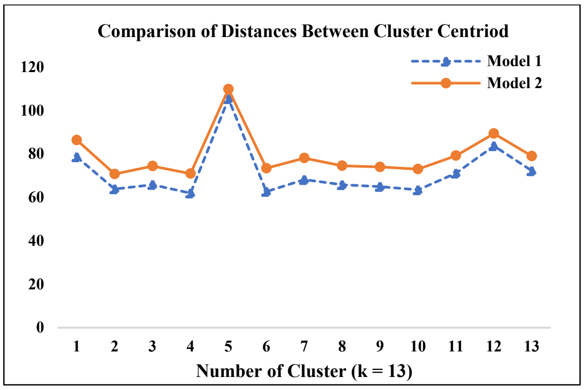

Figure 5 Comparison of Distances between Cluster Centroids (when k = 13)

3.3 Results of Suitable Clustering for Rear Differential Components

This research focused on examining the clustering analysis process by using appropriate clustering criteria and repeating the process until stable results were achieved. This meant that the members within each group no longer changed. After random trials, it was found that the most suitable clustering occurred when $k = 13$, which represented the highest level of alignment with the defined criteria. For the results of clustering the Rear differential components using variables from both models, it was found that there were 13 groups, but there were differences in the group members, particularly with $\boldsymbol{P20}$, which showed different results. The results were shown in **Table 6**.

Table 6 Comparison of Cluster Members Based on Variables from Model 1 and Model 2 When k = 13

Clustering	Clustered Using Model 1 Variables		Clustered Using Model 2 Variables	
	Cluster Membership	Components Name	Cluster Membership	Components Name
C1	3	P1, P2, P23	3	P1, P2, P23
C2	**4**	P3, P4, P26, P27	**5**	P3, P4, **P20,** P26, P27
C3	2	P5, P9	2	P5, P9
C4	1	P6	1	P6
C5	1	P7	1	P7
C6	2	P8, P10	2	P8, P10
C7	**11**	P11, P12, P13, P14, P15, P19, **P20,** P24, P25, P28, P29	**10**	P11, P12, P13, P14, P15, P19, P24, P25, P28, P29
C8	1	P16	1	P16
C9	1	P17	1	P17
C10	3	P18,22,31	3	P18,22,31
C11	1	P21	1	P21
C12	1	P30	1	P30
C13	1	P32	1	P32

Based on the analysis from the R13 sampling round (k = 14), it was found that part $P20$ could not be distinctly separated from group C7. The classification using Model 1, which includes six variables (Special Requirements, Raw Material Grade, Raw Material Type, Manufacturing Process, Tier 2 Suppliers, and Company Location) along with considerations of differences in Design Criteria, yielded ambiguous results. The grouping could not be clearly distinguished, thereby presenting challenges in selecting suitable suppliers for group C7. This was due to the possibility that no single supplier could handle the full range of part types within this group.

Therefore, although the use of variables in Model 1 contributes to technical consistency within groups in terms of materials and manufacturing processes, further consideration of design aspects and supplier capabilities is necessary to ensure optimal suitability.

This research demonstrated that the use of clustering variables in Model 2, combined with the Procurement Strategy Matrix, enhanced the decision-making process for selecting the appropriate clusters. This approach focused on carefully considering the relationships within each group. Additionally, reducing the changes in group membership directly impacted on the significant reduction in analysis time, which was a key factor in making the clustering process more accurate and efficient.

3.4 Comparison of Suppliers Before and After Clustering, Including the Cost Approximation

The analysis of the Hierarchical Cluster Analysis results allowed for the identification of parts that could be sourced from the same supplier. By considering the increase in production volume within groups C1, C2, and C7, and selecting suppliers with strong potential based on the company's cost reduction policy, potential cost savings per purchase order for each part were achieved to offer more competitive prices.

The cost reduction strategy was driven by:

1) A reduction in administrative handling costs per supplier.

2) A reduction in transportation and logistics costs, and

3) Additional savings from the consideration of hidden costs, such as testing and validation expenses when switching suppliers or modifying parts.

These factors were carefully evaluated to ensure that the potential supplier would achieve sustainable cost benefits without compromising quality or production stability. Furthermore, it was observed that some of the existing suppliers within the same clusters exhibited historical operational risks related to logistics and transportation (such as frequent delivery delays and customs clearance issues), exchange rate volatility (affecting the stability of procurement costs for imported components), and slow claims processing in cases of product defects. In response to these risks, the researcher evaluated and selected potential suppliers within each cluster. These suppliers demonstrated higher potential for long-term strategic partnerships, offering improved operational reliability and lower risk profiles compared to the existing suppliers.

As a result of this approach, Group C1 achieved a 9% cost reduction, while Groups C2 and C7 saw a 5% reduction. In total, the cost reduction for the differential parts across these three groups was 7%. For the other groups (C3, C4, C5, C6, C8, C9, C10, C11, C12, C13), only one supplier was involved in each group. The results indicated that grouping parts with multiple suppliers led to more significant cost reductions. The comparison of potential suppliers and projected costs was shown in **Table 7**

The researcher analyzed the reduction in the number of suppliers and found that grouping similar parts together helped reduce supply chain management from 17 suppliers to 13 suppliers (a 23.53% reduction). This resulted in a reduction in the cost of ordering Rear Assy, differential from 1,579 million baht to 1,527 million baht (a 3.3% reduction). This was based on an order of 140,000 Rear Assy, differential in the year 2023. As a result, the company was able to save 52 million baht annually on the total order volume for the year. The results were shown in **Table 8.**

Table 7 Selection of Potential Suppliers from the Clustering Results Using 2 Model Variables (k = 13)

Cluster	Before Clustering		After Clustering		Cost Reduction Ratio
	Original Parts Supplier	Original Cost (Baht)	Potential Supplier	Forecasted Cost (Baht)	
C1: 3 parts	SC1 (P1) SC2 (P2, P23)	2,707	SC1	2,463	-9%
C2: 5 parts	SC3 (P3, P4) SC13 (P20) SC15 (P26, P27)	1,579	SC3	1,494	-5%
C7: 10 parts	SC8 (P11, P12, P13, P14, P15, P28, P29) SC12 (P19, P24, P25)	338	SC8	321	-5%
Total	7 Suppliers	4,624 THB	3 Suppliers	4,278 THB	

Table 8 Results of Analyzing the Reduction in the Number of Suppliers and the Cost of Purchasing Rear Assy, differential in 2023

Criteria	Before Clustering	After Clustering	Reduction Ratio
Number of Suppliers	17	13	23.53%
Cost of Ordering in 2023	1,579 million baht	1,527 million baht	3.30%

4. Conclusion

The findings of this study highlighted the increasing pressure on the automotive industry to enhance efficiency and reduce costs in response to intensifying competition and evolving market demands. A fundamental aspect of addressing this challenge involved the analysis and management of transmission system components, which played a pivotal role in improving operational efficiency and optimizing supply chain management. Through the application of Hierarchical Cluster Analysis techniques, this study achieved cost reduction by systematically clustering differential gear components.

The clustering variables, which captured product characteristics, production processes, and design constraints from Model 1, were integrated with the Procurement Strategy Matrix variables from Model 2. A comparative analysis of the clustering results for rear differential components revealed that while the total number of clusters remained unchanged, the inclusion of Model 2 variables, when combined with the Procurement Strategy Matrix, resulted in a greater differentiation of data between clusters. This refinement enhanced the accuracy of cluster identification by grouping rear differential components based on intrinsic relationships within each cluster, thereby ensuring alignment with the appropriate clustering criteria.

The study's findings on cluster analysis demonstrated that establishing an optimal cluster configuration of comprising 13 groups, based on four key criteria related to design and development functions, ensured that the resulting clusters aligned with engineering logic and practical applications in component design, this clustering approach resulted in a reduction in the number of Tier 2 suppliers from 17 to 13, representing a 23.53% decrease. Moreover, existing suppliers expanded their capabilities to manufacture up to three additional types of components. This strategic reallocation of production enabled the company to prioritize suppliers with higher manufacturing potential, thereby enhancing their capacity to accommodate a broader range of parts. Furthermore, the analysis also contributed to the company's cost reduction strategy, as evidenced by a 3.3% decrease in the cost of Rear Assy, differential in the 2023 order.

From a theoretical perspective, this study contributes to the broader discourse on supply chain optimization and component clustering by demonstrating the synergistic benefits of integrating Design-level and strategic-level variables. The approach reinforces theoretical frameworks on modular production systems and supplier consolidation, providing empirical evidence that supports the alignment of design constraints with procurement strategies. Despite its contributions, the study has certain limitations. The research was conducted within a single automotive manufacturer and focused on a specific component group, which may limit the generalizability of the findings to industries with differing supply chain structures. Future research should explore the applicability of this approach across diverse sectors and consider reflecting market variations. Additionally, this research provided valuable recommendations for procurement strategy development. The variables utilized in this study, particularly those related to design and the integration of the Procurement Strategy Matrix, offered a structured approach for enhancing flexibility in response to fluctuating market demands while simultaneously reducing the time required for analysis.

This study encountered several challenges that affected both the development and implementation of the proposed model. One of the primary limitations was the restricted access to technical data from the supplier, which impacted the precision of the model and the ability to control costs effectively. Determining the optimal cluster (k) required numerous iterations to ensure that the clustering structure was both stable and meaningful. In addition, the evaluation of clusters and the comparison of component costs with potential suppliers proved to be complex. These challenges directly affected the efficiency of decision-making processes and the strategic management of resources, emphasizing the importance of comprehensive data integration and cross-functional collaboration for future implementation.

5. Acknowledgments

The authors sincerely appreciate the support and insights provided by a pickup assembly company in eastern Thailand, which served as the case study for this research. Special thanks go to the employees across various departments for their invaluable cooperation. Additionally, Ms. Saowalak Sombunsook gratefully acknowledges the Monozukuri Degree Scholarship from the Graduate School of Engineering, King Mongkut's Institute of Technology Ladkrabang, for its essential support throughout this study.

6. AI Usage Statement

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

7. Conflict of Interest Statement

The authors declare that there are no conflicts of interest regarding the publication of this paper. The authors did not receive scholarships or financial support for the research, providing complete transparency about potential conflicts of interest related to funding.

8. Author Contributions

Ms. Saowalak Sombunsook led data collection, variable definition, and qualitative data management. She performed Hierarchical Cluster Analysis, interpreted clustering results, validated experimental results, visualized data, and contributed to manuscript preparation, editing, and review. Dr. Kittiwat Sirikasemsuk provided methodological guidance, conceptualized the research, and contributed to manuscript preparation and review. Dr. Kanogkan Leerojanaprapa reviewed the methodology, provided statistical expertise, and assisted in interpreting data analysis. She also contributed to manuscript reviews.

9. References

[1] K. Vanichbuncha, "Cluster Analysis," in *Statistics for research*, 14th ed., Bangkok, Thailand: 3 Lada

Limited Partnership, 2023, ch. 11, sec. 11.3, pp. 216–220.

[2] S. Sinthongsomboon, "Cluster Analysis," in *Multivariate Analysis by using SPSS and Minitab*, 1st ed., Bangkok, Thailand, Chamchuree products co., ltd., 2016, ch. 6, sec. 6.5, pp. 141–144.

[3] A. Rasri, W. Choeichom., P. Chimpalee and N. Kunjiratanachot, "Clustering Information Exposure Behavior the Welfare State Needs Among Older Adults in Thailand." *Journal of Science Ladkrabang*, vol. 33, no. 2, pp. 77–89, 2024.

[4] A. Amilcar, F. A. Andreia A. and D. F. Luis, "Application of a purchasing portfolio model to define medicine purchasing strategies: An empirical study," *Socio-Economic Planning Sciences*, vol. 84,. 2022, Art. no. 101318, doi: 10.1016/j.seps.2022.101318.

[5] M. Rahiminia, J. Razmi, S. S.Farahani and A. Sabbaghnia, "Cluster-based supplier segmentation: a sustainable data-driven approach," *Modern Supply Chain Research and Applications*, vol. 5, no. 3, pp. 209–228, 2023, doi: 10.1108/MSCRA-05-2023-0017.

[6] S. Plungsri and K. Puntusavase, "K-means Clustering for Grouping Product Size for Reducing Cost Packaging," *The Journal of Industrial Technology*, vol. 16, no. 2, pp. 30–44, 2020.

[7] N. Wiroonsri, L. Chokratprapa, N. Srisamarn and P. Dechpichai1, "Clustering Customers Using Their In-Depth Buying Behavior: A Pet Food Manufacturing Company Case Study," *Journal of Science Ladkrabang*, vol. 31, no. 1, pp. 103–119, 2022.

[8] Z. H. Che, "Clustering and selecting suppliers based on simulated annealing algorithms," *Computers & Mathematics with Applications*, vol. 63, no. 1, pp. 228–238. 2012, doi: 10.1016/j.camwa.2011.11.014.

[9] A. T. Almaktoom, K. K. Krishnan, P. Wang and S. Alsobhi, "Cost efficient robust global supply chain system design under uncertainty," *The International Journal of Advanced Manufacturing Technology*, vol. 85, pp. 853–868, 2016, doi: 10.1007/s00170-015-7965-6.

[10] D. Marini and J. R. Corney, "Process selection methodology for near net shape manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 106, pp. 1967–1987, 2020, doi: 10.1007/s00170-019-04561-w.

[11] A. Srirattanapraphan and W. Songpan, "Guidelines for Using Student Personal Information by Clustering Based on User Behavior by Hierarchical Clustering Techniques: A Case Study of the Registration System at Khon Kaen University," in *The Proceedings of 7th National Conference, Suan Sunandha Rajabhat University*, 2024, pp. 585–595. [Online] Available: http://www.journalgrad.ssru.ac.th/index.php/miniconference/article/view/4703/3298

[12] U. Moonpen and S. Mungsing, "Clustering Algorithm Optimization Model for Essential Attribute Analysis of Tour Package Forms Inbound Tourism Market in Thailand," *The Journal of King Mongkut's University of Technology North Bangkok*, vol. 30, no. 4, pp. 656–667, 2019, doi: 10.14416/j.kmutnb.2020.04.006.

[13] A. Araveeporn and J. Promsanga, "Comparing K-Mean Clustering Methods of DNA in Brain Tumors for High-Dimensional Data," *Journal of Science Ladkrabang*, vol. 32, no. 2, pp. 67–79, 2023.

[14] W. Khalid and Z. N. Lee Herbert-Hansen, "Using k-means clustering in international location decision." *Journal of Global Operations and Strategic Sourcing*, vol. 11, no. 3, pp. 274–300, 2018, doi: 10.1108/JGOSS-11-2017-0056.

[15] S. C. Park and J. H. Lee, "Supplier selection and stepwise benchmarking: a new hybrid model using DEA and AHP based on cluster analysis," *Journal of the Operational Research Society*, vol. 69, no. 3, pp. 449–466, 2017, doi: 10.1057/s41274-017-0203-x.