BİLGETÜRK
BilgeTürk Uluslararası Sosyal Araştırmalar Dergisi
Ülkü Ocakları Eğitim ve Kültür Vakfı
2025, Cilt: 9, Sayı: 2, 115-134-178
Geliş Tarihi: 08.12.2025; Kabul Tarihi: 30.12.2025

K-MEANS CLUSTERING ANALYSIS: EXAMINATION OF LOGISTICS PERFORMANCE INDEX (LPI) VALUES USING R SOFTWARE

Ali Şimşek[1*]

Abstract: This study addresses a key limitation in Logistics Performance Index (LPI) research. Countries are typically grouped by geographic region or economic bloc, and these classifications may not reflect true similarities in logistics performance. Traditional groupings may therefore fail to capture underlying patterns in logistics capability. A data-driven country classification based on LPI scores is proposed using k-means clustering. This approach offers a methodological contribution that can inform weighting, ranking, and forecasting studies in literature. The analysis applies k-means clustering in the R environment to the World Bank's 2023 overall LPI scores for 139 countries and groups countries by logistics performance level. The elbow criterion indicates that the 139 countries can be partitioned into three clusters. These clusters are labeled "countries with limited logistics infrastructure and service capacity," "rising economies with developing logistics systems," and "advanced and globally competitive logistics hubs." Silhouette analysis favors a more parsimonious structure and supports a two-cluster solution. The two clusters are "countries with low-to-medium logistics performance" and "countries with high logistics performance." Overall, the results show that LPI-based groupings differ from conventional geographic or economic blocs and provide a more methodologically coherent segmentation of countries in terms of logistics performance.

Keywords: Logistics Performance Index (LPI), Logistics Performance, Country Classification, K-Means Clustering, Cluster Analysis, R Programming.

K-ORTALAMALAR KÜMELEME ANALİZİ: R YAZILIMI KULLANILARAK LOJİSTİK PERFORMANS ENDEKSİ (LPE) DEĞERLERİNİN İNCELENMESİ

Özet: Bu çalışma, Lojistik Performans Endeksi (LPE) araştırmalarında öne çıkan önemli bir sınırlılığı ele almaktadır. Ülkeler genellikle coğrafi bölge veya ekonomik bloklara göre gruplandırılmakta; ancak bu sınıflandırmalar, ülkelerin lojistik performans bakımından gerçek benzerliklerini yeterince yansıtmayabilmektedir. Bu nedenle geleneksel gruplandırmalar, lojistik kapasitedeki temel örüntüleri ortaya koymada yetersiz kalabilmektedir. Bu çalışmada, LPE puanlarına dayalı veri odaklı bir ülke sınıflandırması k-ortalamalar (k-means) kümeleme yöntemiyle önerilmektedir. Bu yaklaşım, alanyazında ağırlıklandırma, sıralama ve kestirim çalışmalarına temel oluşturabilecek yöntemsel bir katkı sunmaktadır. Analizde, Dünya Bankası'nın 2023 yılına ait 139 ülke için genel LPE puanları kullanılmış; R ortamında k-ortalamalar kümeleme analizi uygulanarak ülkeler lojistik performans düzeylerine göre gruplandırılmıştır. Dirsek ölçütü, 139 ülkenin üç kümeye ayrılabileceğini göstermektedir. Bu kümeler "lojistik altyapısı ve hizmet kapasitesi sınırlı ülkeler", "lojistik sistemleri gelişmekte olan yükselen ekonomiler" ve "ileri düzeyde ve küresel ölçekte rekabetçi lojistik merkezler" olarak adlandırılmıştır. Buna karşılık, siluet analizi daha yalın bir yapıyı desteklemekte ve iki kümeli bir çözüm önermektedir. İki küme "düşük ve orta düzey lojistik performansa sahip ülkeler" ile "yüksek lojistik performansa sahip ülkeler" şeklinde tanımlanmaktadır. Bulgular, LPE'ye dayalı ülke gruplandırmalarının geleneksel coğrafi veya ekonomik blok sınıflandırmalarından ayrıştığını ve lojistik performans açısından daha yöntemsel tutarlılığa sahip bir sınıflama sunduğunu göstermektedir.

Anahtar Kelimeler: Lojistik Performans Endeksi (LPE), Lojistik Performans, Ülke Sınıflandırması, K-Ortalamalar Kümeleme, Kümeleme Analizi, R Programlama.

[1]**Adres:** Isparta Uygulamalı Bilimler Üniversitesi, Isparta Meslek Yüksekokulu, Isparta, Türkiye.

[*] **Sorumlu Yazar:** alisimsek@isparta.edu.tr

Atıf: Şimşek, A. (2025). K-Means Clusterıng Analysıs: Examınatıon Of Logıstıcs Performance Index (Lpı) Values Usıng R Software. BilgeTürk Uluslararası Sosyal Araştırmalar Dergisi, 9(2): 148-178.

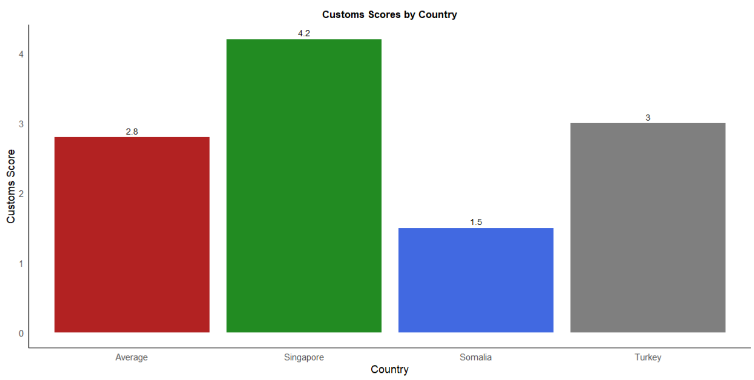
1. INTRODUCTION

The adoption of any action plan or policy, including infrastructure investments and revisions to national regulations and laws, requires a thorough and comprehensive assessment of countries' logistics performance. Such an assessment enables the development of a clear and holistic set of parameters for measuring the quality of logistics performance. Greater emphasis has therefore been placed on the Logistics Performance Index (LPI) as a tool for improving the effectiveness and efficiency of logistics systems (Ju et al., 2024). LPI is defined as a benchmarking tool designed to identify the opportunities and challenges that a given region faces in terms of trade logistics and to reveal how the region's logistics performance can be improved. The index is based on a global-scale study conducted among express cargo carriers and freight forwarders (İnaç et al., 2024). The World Bank's LPI evaluates countries across six components. The six components considered in this index are listed as follows (Arvis et al., 2023):

(a) Efficiency of customs and border management clearance,
(b) Quality of trade- and transport-related infrastructure,
(c) Ease of arranging competitively priced international shipments,
(d) Competence and quality of logistics services,
(e) Ability to track and trace consignments,
(f) Frequency of on-time shipments.

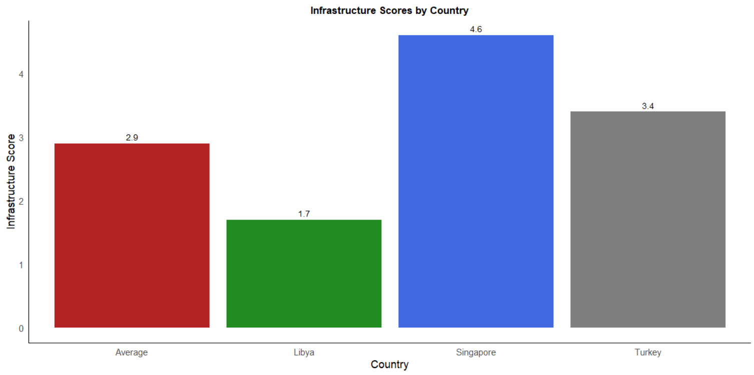
LPI is structured along two main dimensions: input indicators and outcome indicators. The input indicators consist of the customs, infrastructure, and service quality components. These three input components define policy instruments and structural elements that shape a country's logistics environment. The outcome indicators consist of the timeliness, international shipments, and tracking & tracing components. These three outcome components reveal how the logistics system performs in terms of time, cost, and reliability (Arvis et al., 2023). The LPI and its six dimensions help countries identify strengths and weaknesses in logistics processes and benchmark their performance against other countries (Dışkaya & Bozkurt, 2025). In this context, a review of the relevant literature shows that LPI data have been employed across a wide range of research domains. These domains include multi-criteria decision making (MCDM) methods (Martí et al., 2014; Martí et al., 2017; Çemberci et al., 2015; Ulutaş and Karaköy, 2019; Işık et al., 2020; Mercangöz et al., 2020; Yalçın and Ayvaz, 2020; Alnıpak, 2024; Orhan, 2019; Orakçı, 2024; Kamacı, 2025; Ju et al., 2024), efficiency and productivity analyses (Puertas et al., 2014; Acar, 2021a, 2021b; Güdelek et al., 2024), forecasting models (Cansız and Ünsalan, 2020; Son et al., 2020; Babayigit et al., 2023), content analysis (Tümtürk, 2024), and policy and strategy development approaches (Goçer et al., 2022; İnaç et al., 2024). In addition, methodological improvement proposals aimed at the structural refinement of LPI have been put forward by Beysenbaev and Dus (2020). In this regard, the proposal of a new country distribution constitutes one of the original contributions of the present study. Another original contribution is that this country distribution is analyzed in the R software using the k-means clustering method.

The customs index covers customs clearance procedures, including import and export agencies and various services at border crossings. These processes constitute about one-third of the import or export time on average and are influenced by the efficiency of the agency managers and service providers involved in the process (Faria et al., 2015). Figure 1 presents the 2023 customs index results.



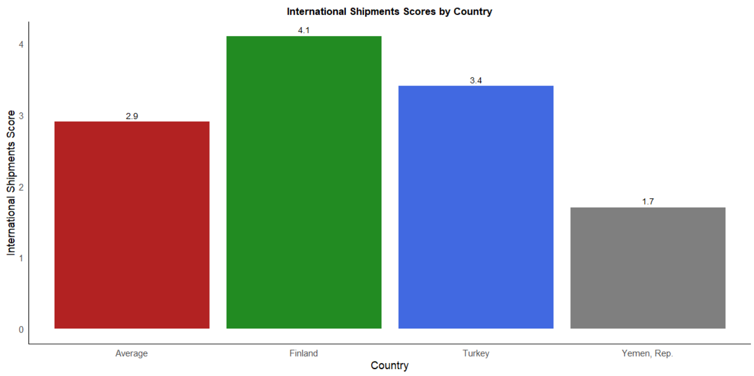
Source: World Bank (2025).
Figure 1. Customs Scores by Country

According to Figure 1, the average value among 139 countries is 2.8, whereas the highest value is 4.2 in Singapore and the lowest is 1.5 in Somalia. Türkiye's score is 3.0, which is slightly above the world average. The gap between the lowest- and highest-ranked countries is 2.7 points. Türkiye is positioned closer to the upper half of this range. Its 49th-place ranking indicates that Türkiye's performance in customs procedures is at an upper–middle level globally, and that there remains potential for improvement relative to the highest performance level. The infrastructure index encompasses elements of transport infrastructure and information and communication technologies (ICT) related to physical transport conditions (Faria et al., 2015). Figure 2 shows the 2023 infrastructure scores by country.



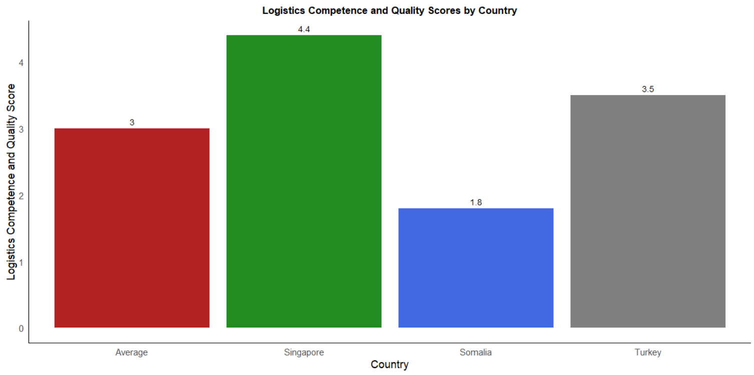
Source: World Bank (2025).
Figure 2. Infrastructure Scores by Country

Figure 2 shows that the mean infrastructure score across the 139 countries is 2.9. The lowest score is 1.7 for Libya, while the highest is 4.6 for Singapore. Türkiye records an infrastructure score of 3.4, which is approximately 0.5 points above the global average. Based on this value, Türkiye ranks 43rd among the 139 countries. This value indicates that Türkiye holds an upper–middle position globally in terms of infrastructure quality, and that a significant gap remains compared to the highest performance level. The international shipments index analyzes the management of the flow of goods and evaluates the ability to organize shipments efficiently, particularly with respect to deliveries and competitive costs (Arvis et al., 2023). In this context, timeliness and flexibility are considered fundamental elements (Faria et al., 2015). Figure 3 shows the 2023 international shipments scores by country.



Source: World Bank (2025).
Figure 3. International Shipments Scores by Country

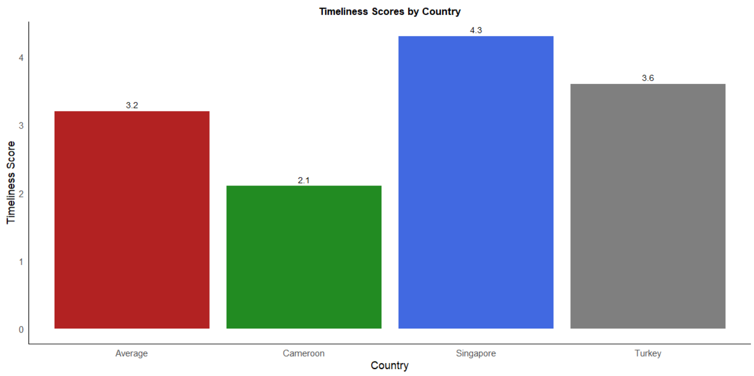
According to Figure 3, the average value among 139 countries is 2.9, whereas the highest score is 4.1 in Finland and the lowest score is 1.7 in Yemen. Türkiye's score is 3.4, which is above the average. Türkiye is ranked 33rd in this indicator. This suggests that Türkiye occupies a relatively strong position globally in arranging international shipments, albeit still below the top performance level. The logistics quality and competence index covers logistics service providers. These services are carried out via road, rail, and air transport. For any group of countries, regardless of their logistics performance level, a lack of competition among firms can pave the way for corruption at border crossings. Such corruption can prevent the emergence of new competitors that could operate more efficiently in international operations (Faria et al., 2015). Figure 4 shows the 2023 logistics competence and quality scores by country.



Source: World Bank (2025).
Figure 4. Logistics Competence and Quality Scores by Country

According to Figure 4, the average value among 139 countries is 3.0, whereas the highest score is 4.4 in Singapore and the lowest is 1.8 in Somalia. Türkiye's logistics competence and quality score is 3.5, which is above the average. Türkiye is ranked 40th on this indicator. This indicates that Türkiye's performance in the competence and quality of logistics services is at an upper–middle level globally, with room for further improvement relative to the highest performance level.

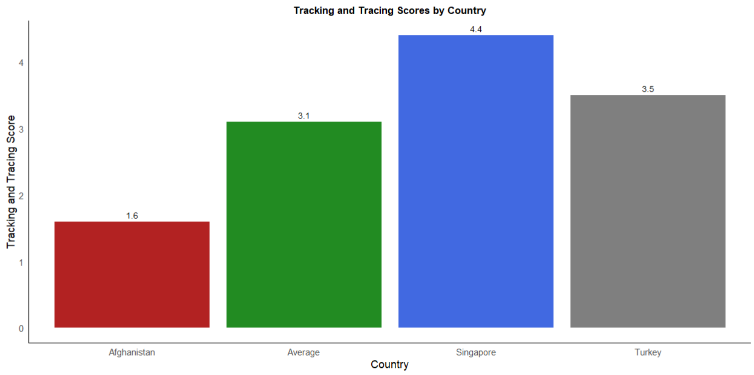
The timeliness index reflects the impact of timeliness and reliability in the trading system on logistics performance. A lack of timeliness and reliability increases costs and lowers competitiveness, thereby potentially restricting trade (Faria et al., 2015). Figure 5 shows the 2023 timeliness scores by country.



Source: World Bank (2025).
Figure 5. Timeliness Scores by Country

According to Figure 5, the average value among 139 countries is 3.2, whereas the highest score is 4.3 in Singapore and the lowest is 2.1 in Cameroon. Türkiye's timeliness score is 3.6, which is above the average. Türkiye is ranked 39th on this indicator. This indicates that Türkiye has a moderately high performance globally in terms of on-time deliveries, and that there is still room for improvement relative to the highest performance level.

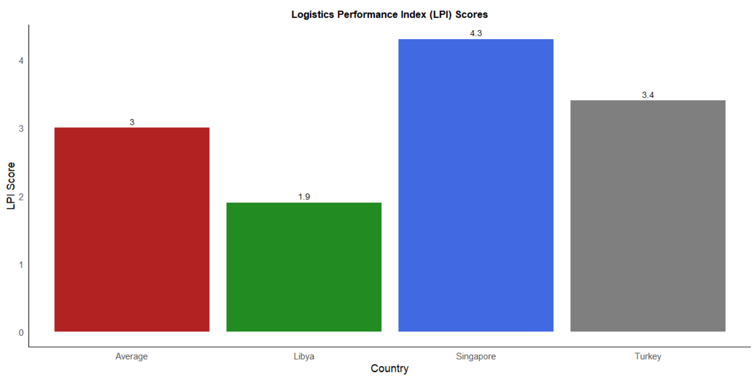
The tracking and tracing index evaluates the management of logistics flows from the point of origin to the destination. The necessity of shortening transit time makes this process critical (Faria et al., 2015). Figure 6 shows the 2023 tracking and tracing scores by country.



Source: World Bank (2025).
Figure 6. Tracking and Tracing Scores by Country

According to Figure 6, the average value among 139 countries is 3.1, whereas the highest score is 4.4 in Singapore and the lowest is 1.6 in Afghanistan. Türkiye's tracking and tracing score is 3.5, which is 0.4 points above the average. Türkiye is ranked 37th on this indicator. This result indicates that, in terms of the traceability of shipments, Türkiye has an upper–middle level performance globally, and that there still remains room for improvement relative to the highest performance level.

Scores for each performance component are generated on a country basis. Items are classified according to the LPI indicators, with each item referring to one or more of these indicators. The total LPI score is calculated as the average of the scores obtained for each criterion (Göçer et al., 2022). Accordingly, Figure 7 presents the 2023 LPI scores.



Source: World Bank (2025).
Figure 7. Logistics Performance Index (LPI) Scores

According to Figure 7, the average value among 139 countries is 3.0, whereas the highest score is 4.3 in Singapore and the lowest is 1.9 in Libya. Türkiye's LPI score is 3.4, which is above the average. Türkiye is ranked 38th on this index. This indicates that Türkiye holds an upper–middle position on a global scale in terms of overall logistics performance, and that there still remains room for improvement relative to the highest performance level.

This study aims to classify countries' logistics performance using the k-means clustering method based on LPI data. In line with this aim, it proposes an alternative country grouping to traditional classifications based on geographic or economic blocs. In addition, it provides a more consistent and comparable classification structure for LPI-based analyses. In this way, the study not only makes a methodological contribution to the logistics performance literature but also establishes an analytical basis for policy makers and practitioners to define country groups more accurately. In the Introduction section, fundamental information on the LPI and its subcomponents is presented and the theoretical framework of the study is outlined. The Methodology section explains the k-means clustering analysis and details the solution steps implemented in the R software. In the Results section, the clusters obtained from the analysis process described in the Methodology section are presented in detail and the findings are discussed. In the Conclusion, Discussion and Recommendations section, the findings are evaluated in a broader perspective, similarities and differences with previous studies are highlighted, and suggestions are developed for future research and practical applications.

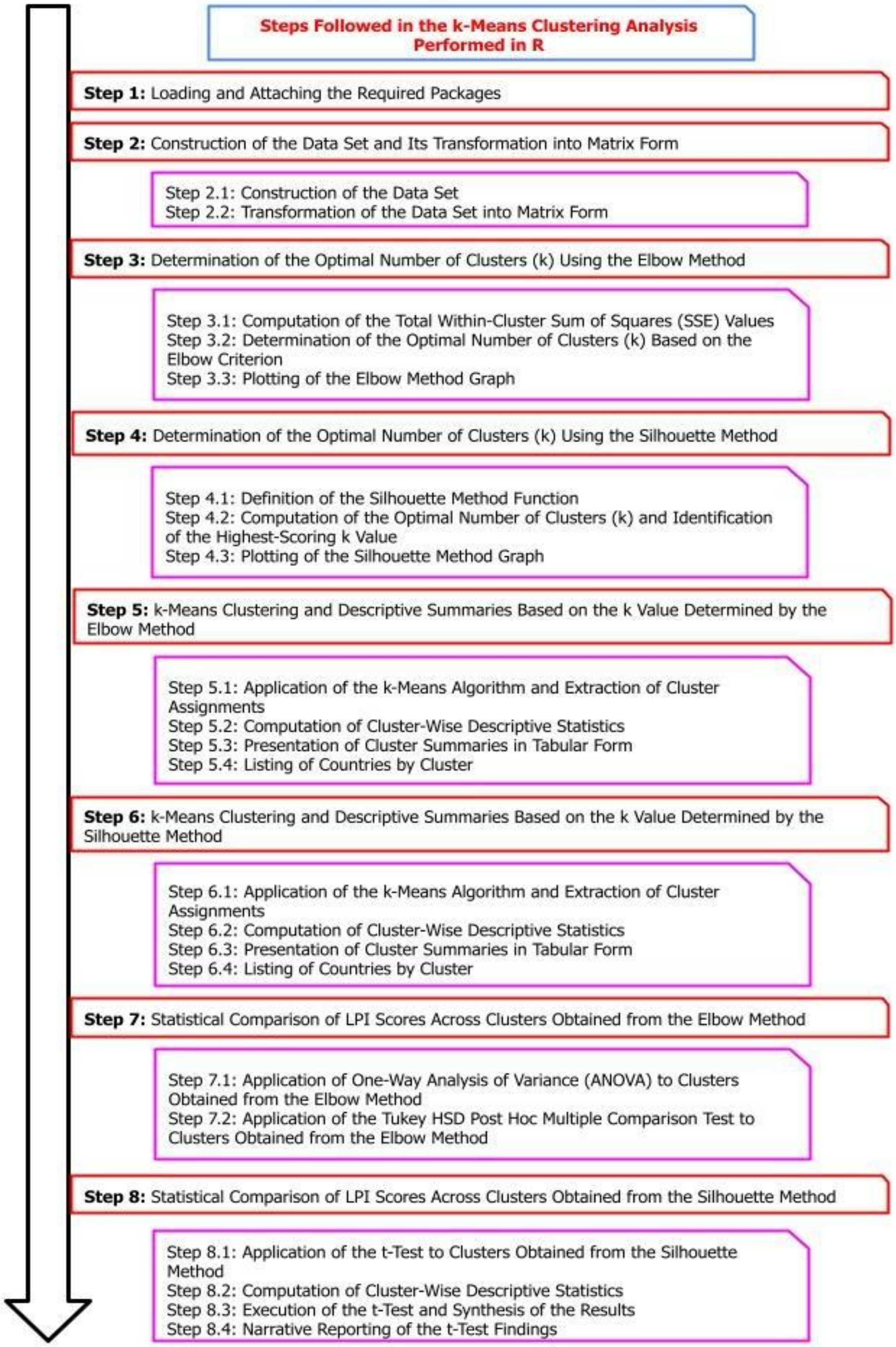
2. METHODOLOGY

This section describes the k-means clustering method and the steps followed in the k-means clustering analysis performed using R software.

The k-means clustering method is defined as one of the best-known clustering techniques and has long been present in the literature (Hartigan and Wong, 1979). Wu et al. (2008) evaluated this clustering method as one of the top ten clustering algorithms in data mining. The term “k-means” was first introduced by MacQueen (1967). The k-means clustering method is defined as a clustering algorithm in which the number of clusters, k, is specified in advance by the user as an input parameter (Simovici, 2022).

The k-means clustering procedure starts with computing the histogram of the density values. First, the density distribution of the data set is calculated, and k initial centroids, representing the number of clusters to be searched for, are randomly selected from the density range. Then, each data point is assigned to the cluster associated with the nearest centroid by considering the Euclidean distance between its own density value and the density values of the centroids. After each iteration, the new centroid of each cluster is updated by taking the arithmetic mean of the density values of the data points assigned to that cluster.

The cluster assignment and centroid updating steps are repeated until the cluster labels of the data points no longer change between successive iterations. In this way, the convergence of the k-means clustering method is ensured, and the final clustering result is obtained (Moftah et al., 2014). In this context, a review of the relevant literature indicates that k-means clustering analysis has been employed in a wide range of fields. These studies include dimension reduction and clustering-based classification approaches (Şengöz and Özdemir, 2016; Rençber, 2019; Onan, 2018; Özyer, 2024; Ikotun et al., 2023), analyses of real estate price dynamics (Hepşen et al., 2015), theoretical and methodological improvements to the k-means algorithm (Pollard, 1981; Zha et al., 2001; Likas et al., 2003; Modha and Spangler, 2003; de Hoon et al., 2004; Kanungo et al., 2004; Khan and Ahmad, 2004; Pham et al., 2005; Çelik, 2009; Niknam and Amiri, 2010; Na et al., 2010; Wang and Song, 2011; Syakur et al., 2018), general methodological evaluations of clustering methods (Ahmed et al., 2020), macroeconomic and regional development analyses (Yılmaz and Kaya, 2005; Yılmaz and Temurlenk, 2005; Keskin, 2018; Özdemir and Yorulmaz, 2025), medical data and image analysis (İşler and Narin, 2012; Yücebaş and Kınacı, 2016; İlkin et al., 2020; Mutlu and Gül, 2023; Zhang et al., 2019), analyses of gender and governance (Ülgen and Arda Özalp, 2017; Gündoğdu and Aytekin, 2020; Zorlutuna, 2024), traffic, spatial risk, and logistics analyses (Selvi and Çağlar, 2017; Win et al., 2021; Ergün, 2025; Günher et al., 2025), and consumer profiling and marketing-oriented studies (İnce et al., 2013). K-means clustering was selected for this study for several reasons. The method is computationally efficient for datasets of this size (Kanungo et al., 2004). The approach yields clearly interpretable, non-overlapping clusters that are well suited to policy-oriented recommendations (Ahmed et al., 2020). Strong performance is also observed with continuous numerical data such as LPI scores, for which Euclidean distance provides a meaningful measure of similarity (Wu et al., 2008). In addition, k-means supports objective validation of cluster solutions through established techniques, including the elbow method and the silhouette coefficient, both of which are used in this analysis. Hierarchical clustering and density-based approaches could also be applied, but k-means is the most appropriate choice given the structure of the data and the study objective of producing distinct and actionable country groupings. Widespread use across disciplines in literature, together with these methodological advantages, further supports the suitability of k-means clustering for the present study. The steps followed in the k-means clustering analysis conducted in R software are presented in Figure 8.



Source: Created by the author.
Figure 8. Steps Followed in the k-means Clustering Analysis Performed in R

3. RESULTS

This section describes the procedures used for the k-means clustering analysis conducted in R. Step 1 installs the required R packages and loads them into the working session using the **install.packages()** and **library()** commands. The **ggplot2 package** supports visualization of LPI patterns and clustering results, **dplyr** is used for data preparation and transformation, **grid** and **gridExtra** are used to arrange multiple plots within a single figure, and **stringr** supports text processing such as cleaning country names and labels. These libraries make their functions available for the subsequent analysis steps, as illustrated in Figure 9.

```
install.packages(
  c("ggplot2",
          "dplyr",
          "gridExtra",
          "stringr")
)
```

```
library(ggplot2)
library(dplyr)
library(gridExtra)
library(grid)
library(stringr)
```

Source: Created by the author.
Figure 9. Loading and Attaching the Required Packages

In Step 2.1, the dataset to be used in the clustering analysis is created in the R environment. In the code, a character vector named Country, which contains the names of the countries, is first defined, followed by a numeric vector named **LPI_Score**, which contains the LPI scores corresponding to each country. These two vectors are then combined using the **data.frame()** function to construct a data frame named **lpi_df**. In this data frame, the rows represent the countries, and the columns represent the variables **Country** and **LPI_Score**. In this way, the raw dataset to be used in the analysis is structured (see Fig. 10).

```
Country <- c("Singapore", "Finland", "Denmark", "Germany", "Netherlands", "Switzerland",
"Austria", "Belgium", "Canada", "Hong Kong SAR, China", "Sweden", "United Arab Emirates",
"France", "Japan", "Spain", "Taiwan, China", "Korea, Rep.", "United States", "Australia", "China",
"Greece", "Italy", "Norway", "South Africa", "United Kingdom", "Estonia", "Iceland", "Ireland",
"Israel", "Luxembourg", "Malaysia", "New Zealand", "Poland", "Bahrain", "Latvia", "Qatar",
"Thailand", "India", "Lithuania", "Portugal", "Saudi Arabia", "Türkiye", "Croatia", "Czech
Republic", "Malta", "Oman", "Philippines", "Slovak Republic", "Slovenia", "Viet Nam", "Brazil",
"Bulgaria", "Cyprus", "Hungary", "Kuwait", "Romania", "Botswana", "Egypt, Arab Rep.", "North
Macedonia", "Bosnia and Herzegovina", "Chile", "Indonesia", "Peru", "Uruguay", "Antigua and
Barbuda", "Benin", "Colombia", "Costa Rica", "Honduras", "Mexico", "Namibia", "Argentina",
"Montenegro", "Rwanda", "Serbia", "Solomon Islands", "Sri Lanka", "Bahamas, The", "Belarus",
"Djibouti", "El Salvador", "Fiji", "Kazakhstan", "Papua New Guinea", "Paraguay", "Ukraine",
"Bangladesh", "Congo, Rep.", "Dominican Republic", "Guatemala", "Guinea-Bissau", "Mali",
"Nigeria", "Russian Federation", "Uzbekistan", "Albania", "Algeria", "Armenia", "Bhutan",
"Central African Republic", "Congo, Dem. Rep.", "Ghana", "Grenada", "Guinea", "Jamaica",
"Mauritius", "Moldova", "Mongolia", "Nicaragua", "Tajikistan", "Togo", "Trinidad and Tobago",
"Zimbabwe", "Bolivia", "Cambodia", "Gabon", "Guyana", "Iraq", "Lao PDR", "Liberia", "Sudan",
"Burkina Faso", "Gambia, The", "Iran, Islamic Rep.", "Kyrgyz Republic", "Madagascar",
"Mauritania", "Syrian Arab Republic", "Venezuela, RB", "Cuba", "Yemen, Rep.", "Angola",
"Cameroon", "Haiti", "Somalia", "Afghanistan", "Libya")

LPI_Score <- c(4.3, 4.2, 4.1, 4.1, 4.1, 4.0, 4.0, 4.0, 4.0, 4.0, 4.0, 4.0, 4.0, 4.0, 3.9, 3.9, 3.9,
3.9, 3.9, 3.8, 3.8, 3.8, 3.7, 3.7, 3.7, 3.7, 3.7, 3.7, 3.7, 3.7, 3.7, 3.7, 3.6, 3.6, 3.6, 3.6, 3.6, 3.6,
3.6, 3.6, 3.6, 3.6, 3.5, 3.5, 3.5, 3.5, 3.5, 3.4, 3.4, 3.4, 3.4, 3.4, 3.3, 3.3, 3.3, 3.3, 3.3, 3.3, 3.3,
3.3, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.2, 3.1, 3.1, 3.1, 3.1, 3.0, 3.0, 3.0, 3.0, 3.0, 3.0, 2.9, 2.9, 2.9,
2.9, 2.9, 2.9, 2.9, 2.8, 2.8, 2.8, 2.8, 2.8, 2.8, 2.7, 2.7, 2.7, 2.6, 2.6, 2.6, 2.6, 2.6, 2.6, 2.6, 2.6,
2.6, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.4, 2.3,
2.3, 2.3, 2.3, 2.3, 2.3, 2.3, 2.3, 2.3, 2.3, 2.2, 2.2, 2.2, 2.1, 2.1, 2.1, 2.0, 1.9, 1.9)

lpi_df <- data.frame(
 Country = Country,
 LPI_Score = LPI_Score
)
```

Source: Created by the author.
Figure 10. Construction of the Data Set

In Step 2.2, the **LPI_Score** variable in the constructed data frame **(lpi_df)** is converted into matrix form (see Fig. 11).

```
lpi_mat <- as.matrix(lpi_df$LPI_Score)
```

Source: Created by the author.
Figure 11. Transformation of the Data Set into Matrix Form

In Step 3.1, within the scope of the Elbow method, the within-cluster sum of squares (SSE) values are computed for different numbers of clusters (see Fig. 12).

```
cat("Elbow method is being applied...\n")
sse <- numeric()
for (k in 1:10) {
  set.seed(42)
  km <- kmeans(lpi_mat, centers = k, nstart = 20)
  centers <- km$centers
  clusters <- km$cluster
  total_sse <- 0
  for (i in 1:length(lpi_mat)) {
    center <- centers[clusters[i]]
    total_sse <- total_sse + (lpi_mat[i] - center)^2
  }
  sse[k] <- total_sse
}
cat("SSE values:", paste(round(sse, 2), collapse = ", "), "\n")
```

Source: Created by the author.
Figure 12. Computation of the Total Within-Cluster Sum of Squares (SSE) Values

In Step 3.2, the optimal number of clusters, k, is automatically computed using the Elbow method (see Fig. 13).

```
find_elbow_point <- function(sse) {
  k_vals <- 1:length(sse)
  p1 <- c(k_vals[1], sse[1])
  p2 <- c(k_vals[length(k_vals)], sse[length(sse)])

  distances <- sapply(1:length(k_vals), function(i) {
    numerator <- abs((p2[2] - p1[2]) * k_vals[i] - (p2[1] - p1[1]) * sse[i] +
                p2[1] * p1[2] - p2[2] * p1[1])
    denominator <- sqrt((p2[2] - p1[2])^2 + (p2[1] - p1[1])^2)
    numerator / denominator
  })

  which.max(distances)
}

elbow_k <- find_elbow_point(sse)
cat(paste0("Optimal k according to Elbow point (based on SSE): ", elbow_k, "\n"))
```

Source: Created by the author.
Figure 13. Determination of the Optimal Number of Clusters (k) Based on the Elbow Criterion

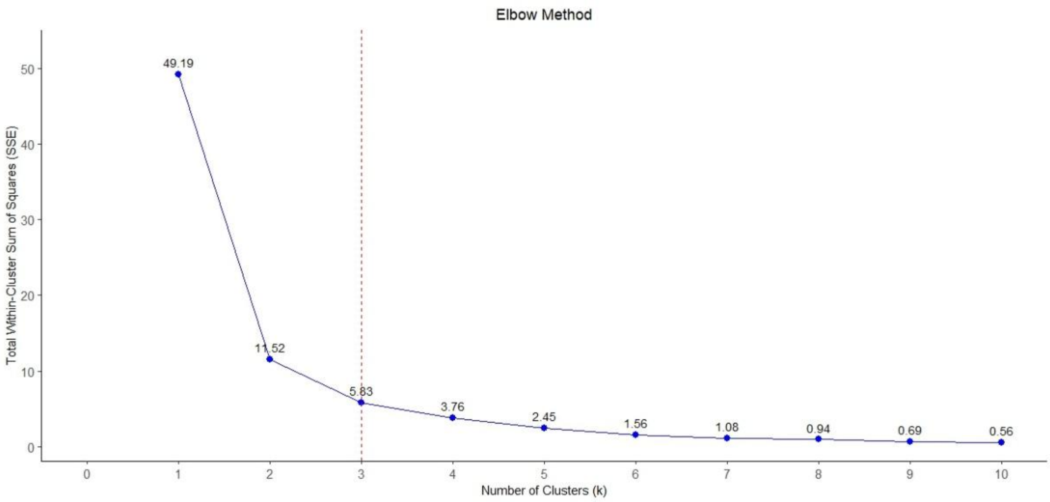
Using the code presented in Step 3.2, the SSE values obtained from the Elbow method indicate that the reduction in error slows down markedly after k = 3. Therefore, it is concluded that grouping the countries into three clusters represents the most appropriate solution in terms of their logistics performance index. In Step 3.3, the SSE–k plot, which visually presents the results of the Elbow method and facilitates their interpretation, is generated (see Fig. 14).

```
elbow_df <- data.frame(
  k  = 1:10,
  SSE = sse
)

ggplot(elbow_df, aes(x = k, y = SSE)) +
  geom_line(color = "blue") +
  geom_point(size = 2, color = "blue") +
  geom_text(aes(label = round(SSE, 2)), vjust = -0.8, size = 3,
          family = "Trebuchet MS") +
  scale_x_continuous(breaks = 0:10, limits = c(0, 10)) +
  scale_y_continuous(expand = expansion(mult = c(0.05, 0.12))) +
  geom_vline(xintercept = elbow_k, linetype = "dashed", color = "red") +
  theme(
    text             = element_text(family = "Trebuchet MS", size = 9),
    axis.title        = element_text(size = 9),
    axis.text          = element_text(size = 9),
    axis.line          = element_line(color = "black"),
    panel.background   = element_blank(),
    panel.grid          = element_blank(),
    panel.border        = element_blank(),
    axis.line.x.bottom = element_line(color = "black"),
    axis.line.y.left   = element_line(color = "black"),
    axis.line.x.top    = element_blank(),
    axis.line.y.right  = element_blank(),
    plot.title          = element_text(size = 11, family = "Trebuchet MS", hjust = 0.5)
  ) +
  labs(
    title = "Elbow Method",
    x = "Number of Clusters (k)",
    y = "Total Within-Cluster Sum of Squares (SSE)"
  )
```

Source: Created by the author.
Figure 14. Plotting of the Elbow Method Graph

In Step 3.3, the final form of the SSE–k plot, which visually presents the results of the Elbow method and facilitates their interpretation, is shown in Figure 15.



Source: Created by the author.
Figure 15. Plotting of the Elbow Method Graph

In Step 4.1, a custom function for the Silhouette method is defined. The purpose of this function is to compute the average silhouette coefficient for a given dataset and a specified number of clusters, k (see Fig. 16).

```
cat("\n Silhouette method is being applied...\n")

silhouette_custom <- function(data, k) {
  km <- kmeans(data, centers = k, nstart = 20)
  cluster <- km$cluster
  dist_mat <- as.matrix(dist(data))
  n <- nrow(data)
  sil_values <- numeric(n)

  for (i in 1:n) {
    own_cluster <- cluster[i]
    own_indices <- which(cluster == own_cluster)
    other_indices <- which(cluster != own_cluster)

    a_i <- if (length(own_indices) > 1) {
      mean(dist_mat[i, own_indices[own_indices != i]])
    } else {
      0
    }

    b_i <- Inf
    for (c in unique(cluster[other_indices])) {
      c_indices <- which(cluster == c)
      b_c <- mean(dist_mat[i, c_indices])
      if (b_c < b_i) b_i <- b_c
    }

    sil_values[i] <- (b_i - a_i) / max(a_i, b_i)
  }

  mean(sil_values)
}
```

Source: Created by the author.
Figure 16. Definition of the Silhouette Method Function

In Step 4.2, using the Silhouette method, the number of clusters with the highest silhouette coefficient is automatically identified among the tested cluster counts (see Fig. 17).

```
silhouette_scores <- sapply(2:10, function(k) silhouette_custom(lpi_mat, k))
cat("Silhouette scores:\n")
print(round(silhouette_scores, 4))

sil_df <- data.frame(
  k = 2:10,
  Silhouette = silhouette_scores
)

silhouette_k <- sil_df$k[which.max(silhouette_scores)]
cat(paste0("Optimal k according to Silhouette method: ", silhouette_k, "\n"))
```

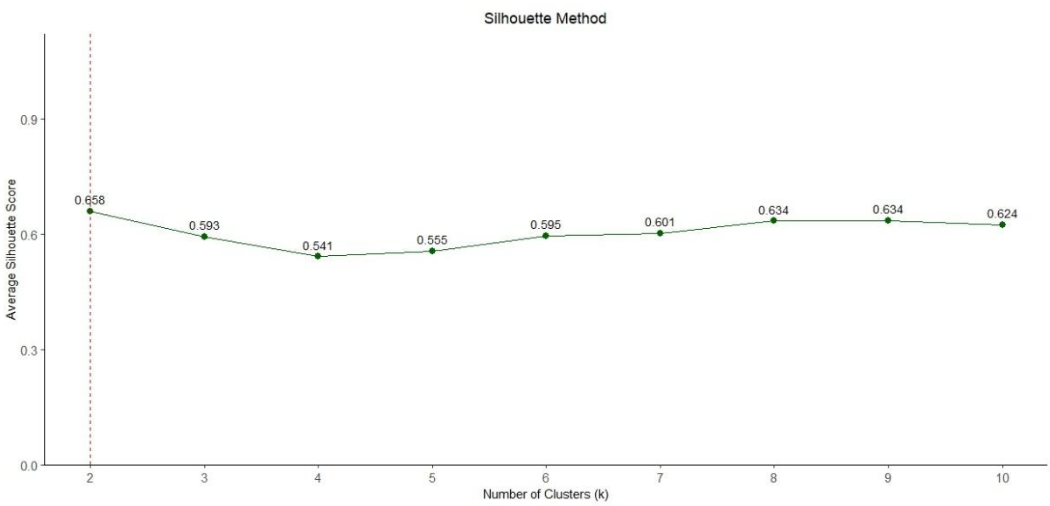
Source: Created by the author.
Figure 17. Computation of the Optimal Number of Clusters (k) and Identification of the Highest-Scoring k Value

Using the Silhouette method, the average silhouette coefficients are calculated for the tested numbers of clusters in the range $2 \leq k \leq 10$. The highest average value is obtained as 0.6582 for $k = 2$. For the other k values, the coefficients remain lower. This result indicates that, based on LPI scores, the countries are most distinctly separated under a two-cluster solution. Furthermore, it is demonstrated that the optimal number of clusters according to the Silhouette method is k = 2. In Step 4.3, a plot illustrating the variation of Silhouette scores with k and the selected optimal k value is generated (see Fig. 18).

```
ggplot(sil_df, aes(x = k, y = Silhouette)) +
  geom_line(color = "darkgreen") +
  geom_point(size = 2, color = "darkgreen") +
  geom_text(aes(label = round(Silhouette, 3)), vjust = -0.8, size = 3,
          family = "Trebuchet MS") +
  scale_x_continuous(breaks = 2:10, limits = c(2, 10)) +
  scale_y_continuous(limits = c(0, 1), expand = expansion(mult = c(0, 0.12))) +
  geom_vline(xintercept = silhouette_k, linetype = "dashed", color = "red") +
  theme(
    text              = element_text(family = "Trebuchet MS", size = 9),
    axis.title        = element_text(size = 9),
    axis.text         = element_text(size = 9),
    axis.line         = element_line(color = "black"),
    panel.background  = element_blank(),
    panel.grid        = element_blank(),
    panel.border      = element_blank(),
    axis.line.x.bottom = element_line(color = "black"),
    axis.line.y.left  = element_line(color = "black"),
    axis.line.x.top   = element_blank(),
    axis.line.y.right = element_blank(),
    plot.title        = element_text(size = 11, family = "Trebuchet MS", hjust = 0.5)
  ) +
  labs(
    title = "Silhouette Method",
    x = "Number of Clusters (k)",
    y = "Average Silhouette Score"
  )
```

Source: Created by the author.
Figure 18. Plotting of the Silhouette Method Graph

In Step 4.3, the results obtained from the Silhouette method are visualized in a graph (see Fig. 19).



Source: Created by the author.
Figure 19. Plotting of the Silhouette Method Graph

According to Figure 19, the horizontal axis represents the number of clusters k (ranging from 2 to 10), while the vertical axis shows the average silhouette score computed for each k. The green line and points indicate how the silhouette scores change as the number of clusters increases. Next to each point, the corresponding score value is displayed as a label. The red dashed vertical line on the left marks the optimal number of clusters (k = 2), at which the average silhouette score reaches its maximum. In this way, the number of clusters that yields the best separation between clusters is presented visually. In Step 5.1, the k-means clustering algorithm is implemented using the optimal number of clusters determined by the Elbow method. Then, the information on which cluster each country belongs to is added to the dataset (see Fig. 20).

```
set.seed(42)
k_model_elbow <- kmeans(lpi_df$LPI_Score, centers = elbow_k, nstart = 25)
lpi_df$Cluster_Elbow <- as.factor(k_model_elbow$cluster)
```

Source: Created by the author.
Figure 20. Application of the k-means Algorithm and Extraction of Cluster Assignments

In Step 5.2, descriptive statistics are computed for the clusters obtained according to the Elbow method (see Fig. 21).

```
summary_elbow <- lpi_df %>%
  group_by(`Cluster` = Cluster_Elbow) %>%
  summarise(
    `Number of Countries` = n(),
    Mean = round(mean(LPI_Score), 5),
    `Std. Dev.` = round(sd(LPI_Score), 5),
    Minimum = round(min(LPI_Score), 5),
    Maximum = round(max(LPI_Score), 5),
    .groups = "drop"
  )

summary_elbow
```

Source: Created by the author.
Figure 21. Computation of Cluster-Wise Descriptive Statistics

In Step 5.3, the descriptive statistics for the clusters are presented as a formatted table that can be directly used in the article (see Fig. 22).

```
my_theme <- ttheme_minimal(
  core = list(
    fg_params = list(fontfamily = "Trebuchet MS", fontsize = 10,
                     col = "black", hjust = 0.5, x = 0.5),
    bg_params = list(fill = "white", col = "black")
  ),
  colhead = list(
    fg_params = list(fontface = "bold", fontfamily = "Trebuchet MS",
                     fontsize = 10, col = "black", hjust = 0.5, x = 0.5),
    bg_params = list(fill = "white", col = "black")
  )
)

g_elbow <- tableGrob(summary_elbow, rows = NULL, theme = my_theme)
g_elbow$heights <- unit(rep(0.6, nrow(g_elbow)), "cm")
g_elbow$widths  <- unit(c(1.5, 4.0, 1.8, 1.8, 1.8, 1.8), "cm")

grid.newpage()
grid.draw(g_elbow)
```

Source: Created by the author.
Figure 22. Presentation of Cluster Summaries in Tabular Form

In Step 5.3, the descriptive statistics for the clusters, formatted in a way that can be directly used in the article, are presented in Table 1.

Table 1. Presentation of Cluster Summaries in Tabular Form

Cluster	Number of Countries	Mean	Std. Dev.	Minimum	Maximum
1	61	2.44262	0.19703	1.9	2.7
2	41	3.10000	0.20616	2.8	3.4
3	37	3.80541	0.22354	3.5	4.3

Source: Created by the author.

According to Table 1, the countries are divided into three clusters based on their LPI scores. The first cluster comprises 61 countries. In this cluster, the mean LPI score is 2.44, with low values observed in the range 1.9–2.7. The second cluster includes 41 countries. In this cluster, a medium level of LPI is observed, with an average score of 3.10 and values ranging from 2.8 to 3.4. The third cluster consists of 37 countries. In this cluster, the highest mean LPI score is found, at 3.81, with values in the range 3.5–4.3. The standard deviation values are relatively low within the clusters. This indicates that each cluster is internally homogeneous. In Step 5.4, the countries are listed according to the clusters obtained using the Elbow method (see Fig. 23).

```
cat("\nCountries in each cluster (Elbow-based k):\n")

countries_elbow <- lpi_df %>%
  arrange(Cluster_Elbow, desc(LPI_Score)) %>%
  select(Cluster_Elbow, Country, LPI_Score)

print(countries_elbow)
```

Source: Created by the author.
Figure 23. Listing of Countries by Cluster

The list of countries according to the clusters obtained using the Elbow method is presented in Table 2.

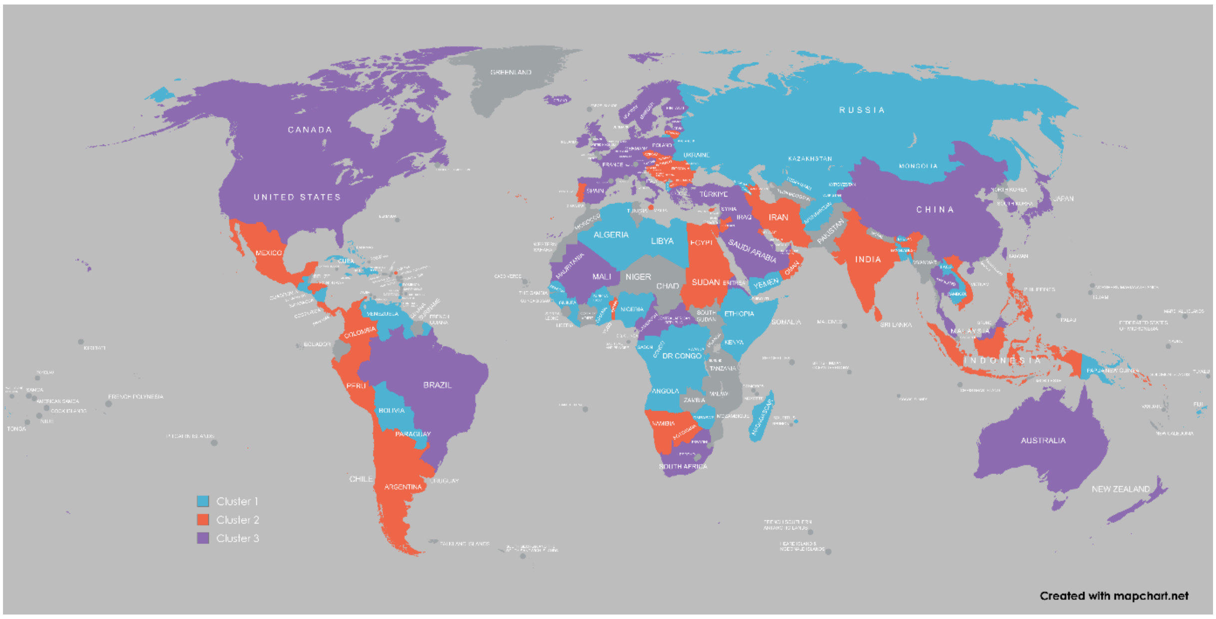
Table 2. Listing of Countries by Cluster

No.	Country	Cluster	No.	Country	Cluster	No.	Country	Cluster
1	Bahamas, The	1	**22**	Bhutan	1	**42**	Lao Pdr	1
2	Belarus	1	**23**	Central African Rep.	1	**43**	Liberia	1
3	Djibouti	1	**24**	Congo, Dem. Rep.	1	**44**	Sudan	1
4	El Salvador	1	**25**	Ghana	1	**45**	Burkina Faso	1
5	Georgia	1	**26**	Grenada	1	**46**	Fiji	1
6	Kazakhstan	1	**27**	Guinea	1	**47**	Gambia, The	1
7	Papua New Guinea	1	**28**	Jamaica	1	**48**	Iran, Islamic Rep.	1
8	Paraguay	1	**29**	Mauritius	1	**49**	Kyrgyz Republic	1
9	Ukraine	1	**30**	Moldova	1	**50**	Madagascar	1
10	Bangladesh	1	**31**	Mongolia	1	**51**	Mauritania	1
11	Congo, Rep.	1	**32**	Nicaragua	1	**52**	Syrian Arab Rep.	1
12	Dominican Republic	1	**33**	Tajikistan	1	**53**	Venezuela, Rb	1
13	Guatemala	1	**34**	Togo	1	**54**	Cuba	1
14	Guinea-Bissau	1	**35**	Trinidad and Tobago	1	**55**	Yemen, Rep.	1
15	Mali	1	**36**	Zimbabwe	1	**56**	Angola	1
16	Nigeria	1	**37**	Bolivia	1	**57**	Cameroon	1
17	Russian Federation	1	**38**	Cambodia	1	**58**	Haiti	1

No.	Country	Cluster	No.	Country	Cluster	No.	Country	Cluster
18	Uzbekistan	1	**39**	Gabon	1	**59**	Somalia	1
19	Albania	1	**40**	Guyana	1	**60**	Afghanistan	1
20	Algeria	1	**41**	Iraq	1	**61**	Libya	1
21	Armenia	1						
62	India	2	**76**	Bulgaria	2	**90**	Antigua and Bar.	2
63	Lithuania	2	**77**	Cyprus	2	**91**	Benin	2
64	Portugal	2	**78**	Hungary	2	**92**	Colombia	2
65	Saudi Arabia	2	**79**	Kuwait	2	**93**	Costa Rica	2
66	Türkiye	2	**80**	Romania	2	**94**	Honduras	2
67	Croatia	2	**81**	Botswana	2	**95**	Mexico	2
68	Czech Republic	2	**82**	Egypt, Arab Rep.	2	**96**	Namibia	2
69	Malta	2	**83**	North Macedonia	2	**97**	Argentina	2
70	Oman	2	**84**	Panama	2	**98**	Montenegro	2
71	Philippines	2	**85**	Bosnia and Herzeg.	2	**99**	Rwanda	2
72	Slovak Republic	2	**86**	Chile	2	**100**	Serbia	2
73	Slovenia	2	**87**	Indonesia	2	**101**	Solomon Islands	2
74	Vietnam	2	**88**	Peru	2	**102**	Sri Lanka	2
75	Brazil	2	**89**	Uruguay	2			
103	Singapore	3	**116**	Japan	3	**128**	Estonia	3
104	Finland	3	**117**	Spain	3	**129**	Iceland	3
105	Denmark	3	**118**	Taiwan, China	3	**130**	Ireland	3
106	Germany	3	**119**	Korea, Rep.	3	**131**	Israel	3
107	Netherlands	3	**120**	United States	3	**132**	Luxembourg	3
108	Switzerland	3	**121**	Australia	3	**133**	Malaysia	3
109	Austria	3	**122**	China	3	**134**	New Zealand	3
110	Belgium	3	**123**	Greece	3	**135**	Poland	3
111	Canada	3	**124**	Italy	3	**136**	Bahrain	3
112	Hong Kong, China	3	**125**	Norway	3	**137**	Latvia	3
113	Sweden	3	**126**	South Africa	3	**138**	Qatar	3
114	United Arab Emir.	3	**127**	United Kingdom	3	**139**	Thailand	3
115	France	3						

Source: Created by the author.

The visual representation of the countries according to the clusters obtained using the Elbow method is shown in Figure 24.



Source: The data were calculated by the author, and the visualization was produced by the author using the MapChart (2025) platform.
Figure 24. Listing of Countries by Cluster

In Step 6.1, the k-means clustering algorithm is implemented using the optimal number of clusters determined by the Silhouette method, and the new cluster assignments are added to the dataset (see Fig. 25).

```
set.seed(42)
k_model_sil <- kmeans(
  data.frame(LPI_Score = lpi_df$LPI_Score),
  centers = silhouette_k,
  nstart  = 25
)

lpi_df$Cluster_Silhouette <- as.factor(k_model_sil$cluster)
```

Source: Created by the author.
Figure 25. Application of the k-means Algorithm and Extraction of Cluster Assignments

In Step 6.2, descriptive statistics are computed for the clusters obtained according to the Silhouette method (see Fig. 26).

```
summary_sil <- lpi_df %>%
  group_by(`Cluster` = Cluster_Silhouette) %>%
  summarise(
    `Number of Countries` = n(),
    Mean = round(mean(LPI_Score), 5),
    `Std. Dev.` = round(sd(LPI_Score), 5),
    Minimum = round(min(LPI_Score), 5),
    Maximum = round(max(LPI_Score), 5),
    .groups = "drop"
  )

summary_sil
```

Source: Created by the author.
Figure 26. Computation of Cluster-Wise Descriptive Statistics

In Step 6.3, the cluster summaries obtained according to the Silhouette method are visualized in tabular form (see Fig. 27). In this way, summary statistics are obtained for the Silhouette-based clusters, including the number of countries, mean, standard deviation, minimum, and maximum LPI values. These statistics are presented in a table format that can be directly used in the article.

```
g_sil <- tableGrobs(summary_sil, rows = NULL, theme = my_theme)
g_sil$heights <- unit(rep(0.6, nrow(g_sil)), "cm")
g_sil$widths  <- unit(c(1.5, 4.0, 1.8, 1.8, 1.8, 1.8), "cm")

grid.newpage()
grid.draw(g_sil)
```

Source: Created by the author.
Figure 27. Presentation of Cluster Summaries in Tabular Form

In Step 6.3, the descriptive statistics for the clusters, formatted in a way that can be directly used in the article, are presented in Table 3.

Table 3. Presentation of Cluster Summaries in Tabular Form

Cluster	Number of Countries	Mean	Std. Dev.	Minimum	Maximum
1	79	2.54557	0.26008	1.9	3.0
2	60	3.59667	0.32519	3.1	4.3

Source: Created by the author.

According to Table 3, the Silhouette-based clustering divides the countries into two groups. The first cluster comprises 79 countries. In this cluster, the mean LPI score is 2.55, with values in the range 1.9–3.0, which represents a relatively low to medium level of logistics performance. The second cluster includes 60 countries, with a mean LPI score of 3.60 and values ranging from 3.1 to 4.3, indicating a higher level of logistics performance. The fact that the standard deviation values are limited in both clusters shows that the clusters exhibit a relatively homogeneous structure internally. In Step 6.4, the countries are listed according to the clusters obtained using the Silhouette method (see Fig. 28).

```
cat("\nCountries in each cluster (Silhouette-based k):\n")

countries_sil <- lpi_df %>%
  arrange(Cluster_Silhouette, desc(LPI_Score)) %>%
  select(Cluster_Silhouette, Country, LPI_Score)

print(countries_sil)
```

Source: Created by the author.
Figure 28. Listing of Countries by Cluster

The list of countries according to the clusters obtained using the Silhouette method is presented in Table 4.

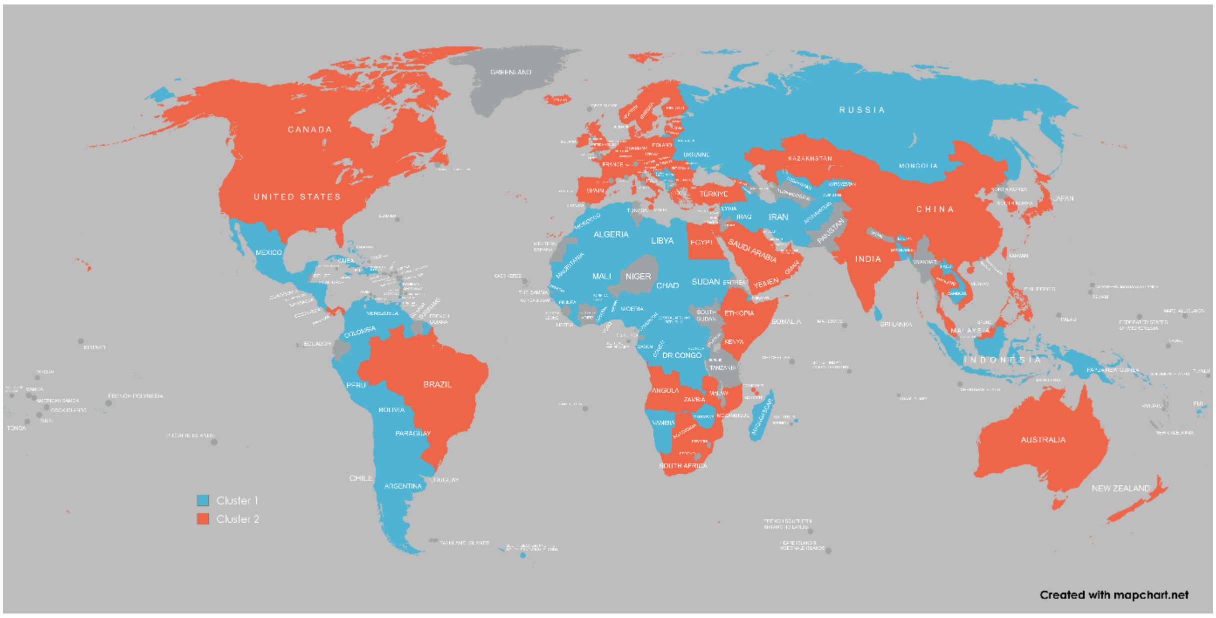
Table 4. Listing of Countries by Cluster

No.	Country	Cluster	No.	Country	Cluster	No.	Country	Cluster
1	Bosnia and Herzegovina	1	**28**	Bangladesh	1	**54**	Zimbabwe	1
2	Chile	1	**29**	Congo, Rep.	1	**55**	Bolivia	1
3	Indonesia	1	**30**	Dominican Rep.	1	**56**	Cambodia	1
4	Peru	1	**31**	Guatemala	1	**57**	Gabon	1
5	Uruguay	1	**32**	Guinea-Bissau	1	**58**	Guyana	1
6	Antigua and Barbuda	1	**33**	Mali	1	**59**	Iraq	1
7	Benin	1	**34**	Nigeria	1	**60**	Lao Pdr	1
8	Colombia	1	**35**	Russian Federation	1	**61**	Liberia	1
9	Costa Rica	1	**36**	Uzbekistan	1	**62**	Sudan	1
10	Honduras	1	**37**	Albania	1	**63**	Burkina Faso	1
11	Mexico	1	**38**	Algeria	1	**64**	Fiji	1
12	Namibia	1	**39**	Armenia	1	**65**	Gambia, The	1
13	Argentina	1	**40**	Bhutan	1	**66**	Iran, Islamic Rep.	1
14	Montenegro	1	**41**	Cent. African Rep.	1	**67**	Kyrgyz Republic	1
15	Rwanda	1	**42**	Congo, Dem. Rep.	1	**68**	Madagascar	1
16	Serbia	1	**43**	Ghana	1	**69**	Mauritania	1
17	Solomon Islands	1	**44**	Grenada	1	**70**	Syrian Arab Rep.	1
18	Sri Lanka	1	**45**	Guinea	1	**71**	Venezuela, Rb	1
19	Bahamas, The	1	**46**	Jamaica	1	**72**	Cuba	1
20	Belarus	1	**47**	Mauritius	1	**73**	Yemen, Rep.	1
21	Djibouti	1	**48**	Moldova	1	**74**	Angola	1
22	El Salvador	1	**49**	Mongolia	1	**75**	Cameroon	1
23	Georgia	1	**50**	Nicaragua	1	**76**	Haiti	1
24	Kazakhstan	1	**51**	Tajikistan	1	**77**	Somalia	1
25	Papua New Guinea	1	**52**	Togo	1	**78**	Afghanistan	1
26	Paraguay	1	**53**	Trinidad and Tobago	1	**79**	Libya	1
27	Ukraine	1						
80	Singapore	2	**101**	Italy	2	**121**	Türkiye	2
81	Finland	2	**102**	Norway	2	**122**	Croatia	2
82	Denmark	2	**103**	South Africa	2	**123**	Czech Republic	2
83	Germany	2	**104**	United Kingdom	2	**124**	Malta	2
84	Netherlands	2	**105**	Estonia	2	**125**	Oman	2
85	Switzerland	2	**106**	Iceland	2	**126**	Philippines	2
86	Austria	2	**107**	Ireland	2	**127**	Slovak Republic	2
87	Belgium	2	**108**	Israel	2	**128**	Slovenia	2
88	Canada	2	**109**	Luxembourg	2	**129**	Vietnam	2
89	Hong Kong Sar, China	2	**110**	Malaysia	2	**130**	Brazil	2
90	Sweden	2	**111**	New Zealand	2	**131**	Bulgaria	2
91	United Arab Emirates	2	**112**	Poland	2	**132**	Cyprus	2
92	France	2	**113**	Bahrain	2	**133**	Hungary	2
93	Japan	2	**114**	Latvia	2	**134**	Kuwait	2
94	Spain	2	**115**	Qatar	2	**135**	Romania	2
95	Taiwan, China	2	**116**	Thailand	2	**136**	Botswana	2
96	Korea, Rep.	2	**117**	India	2	**137**	Egypt, Arab Rep.	2
97	United States	2	**118**	Lithuania	2	**138**	North Macedonia	2
98	Australia	2	**119**	Portugal	2	**139**	Panama	2

No.	Country	Cluster	No.	Country	Cluster	No.	Country	Cluster
99	China	2	**120**	Saudi Arabia	2			
100	Greece	2						

Source: Created by the author.

The visual representation of the countries according to the clusters obtained using the Silhouette method is shown in Figure 29.



Source: The data were calculated by the author, and the visualization was produced by the author using the MapChart (2025) platform.

Figure 29. Listing of Countries by Cluster

In Step 7.1, it is tested whether the mean LPI scores differ statistically across the clusters obtained using the Elbow method by employing a one-way analysis of variance (ANOVA). In the code shown in Figure 30, the ANOVA model is first specified with the expression **aov(LPI_Score ~ Cluster_Elbow, data = lpi_df)**.

Then, the summary results are extracted to obtain an ANOVA table that includes the term, sum of squares, degrees of freedom, mean square, F statistic, and p-value. In the final part, this table is formatted with a custom theme and rendered as a graphical output. In this way, the significance of the differences in LPI scores across clusters is reported in the form of an ANOVA table.

```
cat("\n ANOVA test (for Elbow clusters) is being applied...\n")

anova_model   <- aov(LPI_Score ~ Cluster_Elbow, data = lpi_df)
anova_summary <- summary(anova_model)
anova_df      <- as.data.frame(anova_summary[[1]])
anova_df$Term <- rownames(anova_df)
anova_df <- anova_df[, c("Term", setdiff(names(anova_df), "Term"))]
anova_df$Term <- anova_df$Term %>%
  str_trim() %>%
  recode(
    "Cluster_Elbow" = "Countries",
    "Residuals"     = "Error Variance"
  )
colnames(anova_df) <- c("Terms", "Sum of Squares", "df",
                        "Mean Square", "F value", "p")
anova_df <- anova_df %>%
  mutate(across(where(is.numeric), ~ round(., 4)))
anova_df[["p"]] <- ifelse(
  is.na(anova_df[["p"]]),
  "",
  ifelse(anova_df[["p"]] == 0, "0.000", format(anova_df[["p"]], digits = 3))
)
anova_df[["F value"]] <- ifelse(is.na(anova_df[["F value"]]),
                                "", anova_df[["F value"]])
my_theme_anova <- ttheme_minimal(
  core = list(
    fg_params = list(fontfamily = "Arial", fontsize = 10,
                     col = "black", hjust = 0.5, x = 0.5),
    bg_params = list(fill = "white", col = "black")
  ),
  colhead = list(
    fg_params = list(fontface = "bold", fontfamily = "Arial",
                     fontsize = 10, col = "black", hjust = 0.5, x = 0.5),
    bg_params = list(fill = "white", col = "black")
  )
)
anova_table <- tableGrobs(anova_df, rows = NULL, theme = my_theme_anova)
anova_table$heights <- unit(rep(0.6, nrow(anova_table)), "cm")
anova_table$widths  <- unit(c(3.0, 3.0, 3.0, 3.0, 2.0, 2.0), "cm")

grid.newpage()
grid.draw(anova_table)
```

Source: Created by the author.

Figure 30. Application of One-Way Analysis of Variance (ANOVA) to Clusters Obtained from the Elbow Method

The ANOVA table reporting the significance of the differences in LPI scores is presented in Table 5.

Table 5. Application of One-Way Analysis of Variance (ANOVA) to Clusters Obtained from the Elbow Method

Terms	Sum of Squares	df	Mean Square	F value	p
Countries	2	43.3618	21.6809	505.929	0.000
Error Variance	136	5.8281	0.0429		

Source: Created by the author.

The ANOVA results show that the mean LPI scores differ significantly across clusters ($F_{(2, 136)}$ = 505.929, $p < 0.001$).

The fact that the sum of squares for countries (clusters) is much larger than the error variance indicates that a substantial proportion of the total variability in LPI scores is attributable to between-cluster differences. In Step 7.2, Tukey's Honest Significant Difference (HSD) multiple comparison test is applied to examine the pairwise differences between the clusters obtained using the Elbow method.

The resulting outputs are visualized in tabular form. In this way, the cluster differences found to be significant by ANOVA are statistically decomposed to reveal which specific pairs of clusters account for these differences (see Fig. 31).

```
cat("\n Tukey HSD post-hoc test (Elbow clusters) is being applied...\n")

tukey_result <- TukeyHSD(anova_model)

tukey_df <- as.data.frame(tukey_result$Cluster_Elbow)
tukey_df$Comparison <- rownames(tukey_df)

tukey_df <- tukey_df[, c("Comparison", "diff", "lwr", "upr", "p adj")]
colnames(tukey_df) <- c("Cluster Comparison", "Difference",
                        "Lower CI", "Upper CI", "p")

tukey_df <- tukey_df %>%
  mutate(across(where(is.numeric), ~ round(., 4)))

tukey_df$p <- ifelse(
  is.na(tukey_df$p),
  "",
  ifelse(tukey_df$p == 0, "0.000", format(tukey_df$p, digits = 3))
)

my_theme_tukey <- ttheme_minimal(
  core = list(
    fg_params = list(fontfamily = "Arial", fontsize = 10,
                     col = "black", hjust = 0.5),
    bg_params = list(fill = "white", col = "black")
  ),
  colhead = list(
    fg_params = list(fontface = "bold", fontfamily = "Arial", fontsize = 10,
                     col = "black", hjust = 0.5),
    bg_params = list(fill = "white", col = "black")
  )
)

tukey_table <- tableGrobs(tukey_df, rows = NULL, theme = my_theme_tukey)
tukey_table$heights <- unit(rep(0.6, nrow(tukey_table)), "cm")
tukey_table$widths  <- unit(c(4.2, 2.6, 3.1, 3.1, 3.0), "cm")

grid.newpage()
grid.draw(tukey_table)
```

Source: Created by the author.
Figure 31. Application of the Tukey HSD Post Hoc Multiple Comparison Test to Clusters Obtained from the Elbow Method

The results on pairwise differences between the clusters obtained using the Elbow method are presented in Table 6.

Table 6. Application of the Tukey HSD Post Hoc Multiple Comparison Test to Clusters Obtained from the Elbow Method

Cluster Comparison	Difference	Lower CI	Upper CI	p
2–1	0.6574	0.5583	0.7564	0.000
3–1	1.3628	1.2606	1.4650	0.000
3–2	0.7054	0.5942	0.8166	0.000

Source: Created by the author.

Table 6 shows that the mean LPI scores differ significantly in all pairwise comparisons between clusters (all p-values < 0.001). The mean LPI score of the second cluster is 0.66 points higher than that of the first cluster, with a confidence interval ranging from 0.56 to 0.76. The third cluster exhibits an average LPI level that is approximately 1.36 points higher than the first cluster and 0.71 points higher than the second cluster. In both comparisons, the confidence intervals do not include zero. Accordingly, the clusters display a hierarchical structure in terms of LPI performance, such that Cluster 1 < Cluster 2 < Cluster 3. In Step 8.1, the statistical significance of the difference in mean LPI scores between the two clusters obtained using the Silhouette method is evaluated by means of an independent two-sample t-test. In this way, the significance of the difference between the LPI means of the two Silhouette-based clusters is quantitatively assessed using the t-test (see Fig. 32).

```
cat("\n Two-sample t-test for Silhouette clusters is being applied...\n")

if (length(levels(lpi_df$Cluster_Silhouette)) == 2) {

  # Run the independent-samples t-test
  ttest_sil <- t.test(LPI_Score ~ Cluster_Silhouette, data = lpi_df)
  print(ttest_sil)   # raw output from t.test (optional)

  # Scalars for each cluster
  n1    <- sil_stats$n[sil_stats$Cluster == "1"]
  n2    <- sil_stats$n[sil_stats$Cluster == "2"]
  mean1 <- sil_stats$Mean[sil_stats$Cluster == "1"]
  mean2 <- sil_stats$Mean[sil_stats$Cluster == "2"]
  sd1   <- sil_stats$SD[sil_stats$Cluster == "1"]
  sd2   <- sil_stats$SD[sil_stats$Cluster == "2"]

  # t-test parameters
  t_val <- as.numeric(ttest_sil$statistic)
  df_t  <- as.numeric(ttest_sil$parameter)
  p_val <- as.numeric(ttest_sil$p.value)
  ci    <- as.numeric(ttest_sil$conf.int)   # length-2 numeric vector
}
```

Source: Created by the author.
Figure 32. Application of the t-Test to Clusters Obtained from the Silhouette Method

In Step 8.2, basic descriptive statistics are computed for the clusters obtained according to the Silhouette method. In this way, a summary statistical profile of the LPI level is obtained for each Silhouette-based cluster prior to the t-test (see Fig. 33).

```
sil_stats <- lpi_df %>%
  group_by(Cluster_Silhouette) %>%
  summarise(
    Cluster = first(as.character(Cluster_Silhouette)),
    n       = n(),
    Mean    = mean(LPI_Score),
    SD      = sd(LPI_Score),
    .groups = "drop"
  ) %>%
  select(Cluster, n, Mean, SD)
```

Source: Created by the author.
Figure 33. Computation of Cluster-Wise Descriptive Statistics

In Step 8.3, two procedures are carried out jointly. First, the descriptive statistics for the clusters obtained according to the Silhouette method (n, mean LPI, standard deviation) are rounded to four decimal places and organized within ttest_desc_df. These summaries are then rendered as a cluster-based summary table using a custom theme. Next, using the previously computed t-test results, the mean difference between clusters (Difference), the lower and upper confidence interval bounds for this difference (Lower CI, Upper CI), the t value, degrees of freedom, and p-value are converted into a one-row summary table (ttest_summary_df). This summary table is visualized as a formatted t-test results table.

In summary, for the Silhouette-based clusters, both the descriptive statistics and the t-test findings are reported as two separate tables that can be directly used in the article (see Fig. 34).

```
ttest_desc_df <- sil_stats %>%
  mutate(
    Mean = round(Mean, 4),
    SD   = round(SD,   4)
  )
print(ttest_desc_df)
my_theme_ttest_desc <- ttheme_minimal(
  core = list(
    fg_params = list(fontfamily = "Arial", fontsize = 10,
                     col = "black", hjust = 0.5),
    bg_params = list(fill = "white", col = "black")
  ),
  colhead = list(
    fg_params = list(fontface = "bold", fontfamily = "Arial",
                    fontsize = 10, col = "black", hjust = 0.5),
    bg_params = list(fill = "white", col = "black")
  )
)
ttest_desc_table <- tableGrob(ttest_desc_df, rows = NULL, theme = my_theme_ttest_desc)
ttest_desc_table$heights <- unit(rep(0.6, nrow(ttest_desc_table)), "cm")
ttest_desc_table$widths  <- unit(c(3.0, 2.0, 3.0, 3.0), "cm")
grid.newpage()
grid.draw(ttest_desc_table)

diff_21  <- mean2 - mean1       # difference of means (2 - 1)
lower_21 <- ci[2]               # CI for Cluster 2 - Cluster 1
upper_21 <- ci[1]
# p for table: if < 0.001 show "0.000", else 3 decimals
p_for_table <- if (p_val < 0.001) {
  "0.000"
} else {
  formatC(p_val, format = "f", digits = 3)
}
ttest_summary_df <- data.frame(
  `Cluster Comparison` = "2-1",
  Difference = round(diff_21, 4),
  `Lower CI` = round(lower_21, 4),
  `Upper CI` = round(upper_21, 4),
  t  = round(t_val, 4),
  df = round(df_t, 2),
  p  = p_for_table,
  check.names = FALSE
)
print(ttest_summary_df)

my_theme_ttest <- ttheme_minimal(
  core = list(
    fg_params = list(fontfamily = "Arial", fontsize = 10,
                    col = "black", hjust = 0.5),
    bg_params = list(fill = "white", col = "black")
  ),
  colhead = list(
    fg_params = list(fontface = "bold", fontfamily = "Arial",
                    fontsize = 10, col = "black", hjust = 0.5),
    bg_params = list(fill = "white", col = "black")
  )
)
ttest_table <- tableGrob(ttest_summary_df, rows = NULL, theme = my_theme_ttest)
ttest_table$heights <- unit(rep(0.6, nrow(ttest_table)), "cm")
ttest_table$widths  <- unit(c(4.2, 2.6, 3.0, 3.0, 2.0, 2.0, 2.4), "cm")
grid.newpage()
grid.draw(ttest_table)
```

Source: Created by the author.
Figure 34. Execution of the t-Test and Synthesis of the Results

The descriptive statistics for the Silhouette-based clusters are presented in Table 7.

Table 7. Execution of the t-Test and Synthesis of the Results

Cluster Comparison	Difference	Lower CI	Upper CI	t	df	p
2–1	1.0511	0.9497	1.1525	-20.5399	110.52	0.000

Source: Created by the author.

In Step 8.4, the narrative report of the independent-samples t-test conducted for the Silhouette clusters is automatically generated. The code used to obtain this report is shown in Figure 35.

```
p_text <- if (p_val < 0.001) {
  "< .001"
} else {
  paste0("= ", formatC(p_val, format = "f", digits = 3))
}

sig_text <- if (p_val < 0.05) "significant" else "non-significant"

report_text <- sprintf(
  paste(
    "An independent-samples t-test was conducted to compare LPI scores ",
    "between the two Silhouette clusters. ",
    "Cluster 1 (n = %d) had a mean LPI score of %.2f (SD = %.2f), ",
    "whereas Cluster 2 (n = %d) had a mean of %.2f (SD = %.2f). ",
    "The difference between the two clusters was %s, ",
    "t(%.2f) = %.2f, p %s, ",
    "95%% CI for the mean difference (Cluster 1 – Cluster 2) [%.2f, %.2f].",
    sep = ""
  ),
  n1, mean1, sd1,
  n2, mean2, sd2,
  sig_text,
  df_t, t_val, p_text,
  ci[1], ci[2]
)

cat("\nT-test report:\n")
cat(report_text, "\n\n")

} else {
  cat("Silhouette method did not result in exactly 2 clusters; two-sample t-test is not appropriate.\n")
}
```

Source: Created by the author.
Figure 35. Narrative Reporting of the t-Test Findings

An independent-samples t-test was conducted to compare LPI scores between the two Silhouette clusters. Cluster 1 (n = 79) had a mean LPI score of 2.55 (sd = 0.26), whereas Cluster 2 (n = 60) had a mean of 3.60 (sd = 0.33). The difference between the two clusters was significant, $t_{(110.52)} = -20.54$, p < .001, 95% CI for the mean difference (Cluster 1 – Cluster 2) [-1.15, -0.95].

4. CONCLUSION, DISCUSSION, AND RECOMMENDATIONS

This study aims to group countries using the k-means clustering method based on the 2023 LPI data. The analysis conducted in the R software environment indicates that, according to the elbow method, 139 countries can be partitioned into three clusters. These clusters are labelled as “countries with limited logistics infrastructure and service capacity”, “rising economies with developing logistics systems” and “advanced and globally competitive logistics hubs”. According to the silhouette method, a more parsimonious structure emerges, in which countries are grouped into two main clusters, namely “countries with low and medium logistics performance” and “countries with high logistics performance”. These results provide, for LPI-based country classifications, a structure that differs from traditional geographic or economic blocs and reveal logistics performance levels in a more distinctive manner.

The clusters obtained in the study indicate differentiated priority areas for policy makers and practitioners. Countries in the first cluster, described as “countries with limited logistics infrastructure and service capacity”, should prioritize investments aimed at strengthening physical infrastructure and simplifying basic customs procedures. The second cluster, “rising economies with developing logistics systems”, would benefit from policies that promote more efficient use of existing infrastructure, expansion of digital logistics solutions and enhanced cross-border logistics integration. The third cluster, “advanced and globally competitive logistics hubs”, needs to focus on innovative logistics practices, green logistics policies and high value-added logistics services in order to maintain its competitive advantage.

Logistics performance depends not only on infrastructure and policy but also on the human and managerial capabilities within logistics organizations. Coaching-based leadership represents a shift away from traditional managerial control toward coaching and mentoring behaviors that support employee development and enhance performance (Özdemir, 2025a). Workplace mindfulness refers to sustained, nonjudgmental awareness of present-moment experience at work and can promote more deliberate responses to workplace demands (Özdemir, 2025b). Mindful leadership extends these principles to leadership practice by incorporating mindfulness to strengthen decision making and support well-being under competitive and uncertain conditions (Özdemir, 2023).

This study has several limitations. The analysis relies solely on the 2023 LPI scores published by the World Bank. Country groupings are derived exclusively by means of the k-means clustering method, and the appropriate number of clusters is determined using the elbow and silhouette techniques. Conducting all analyses within the R software environment constrains methodological diversity.

Future research could examine how country clusters evolve over time and how countries transition between clusters by using multi-year LPI data. Additional studies could also compare the clustering results obtained using different methods. Alternative approaches such as hierarchical clustering and fuzzy clustering could be employed to evaluate the robustness of the country classification structure proposed in this study in greater detail.

Ethics Committee Approval
N/A

Peer-review
Externally peer-reviewed.

Conflict of Interest
The authors have no conflicts of interest to declare.

Funding
The authors declared that this study has received no financial support.

REFERENCES

Acar, M. F. (2021a). Lojistik etkinlik: Türkiye ve OECD. Avrupa Bilim ve Teknoloji Dergisi, (23), 512–517.

Acar, M. F. (2021b). Lojistik performans indeks: Türkiye–Avrupa Birliği karşılaştırması. International Journal of Advanced Engineering and Pure Sciences, 33(3), 422–428.

Ahmed, M., Seraj, R., & Islam, S. M. S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. Electronics, 9(8), 1295.

Alnıpak, S. (2024). AHS–COCOSO yöntemi ile APEC ülkelerinin lojistik performanslarının değerlendirilmesi. Tarsus Üniversitesi Uygulamalı Bilimler Fakültesi Dergisi, 4(1), 13–26.

Arvis, J.-F., Ojala, L., Shepherd, B., Ulybina, D., & Wiederer, C. (2023). Connecting to compete 2023: Trade logistics in an uncertain global economy. The Logistics Performance Index and its indicators. World Bank.

Babayigit, B., Gürbüz, F., & Denizhan, B. (2023). Logistics performance index estimating with artificial intelligence. International Journal of Shipping and Transport Logistics, 16(3–4), 360–371.

Beysenbaev, R., & Dus, Y. (2020). Proposals for improving the Logistics Performance Index. The Asian Journal of Shipping and Logistics, 36(1), 34–42.

Cansız, Ö. F., & Ünsalan, K. (2020). Yapay zekâ ve istatistiksel yöntemler ile küresel ticarette rekabet ölçütü olan lojistik performans indeksine (LPI) etken parametrelerin ülke bazlı incelenmesi ve tahmin modellerinin geliştirilmesi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 32(2), 571–582.

Çelik, T. (2009). Unsupervised change detection in satellite images using principal component analysis and k-means clustering. IEEE Geoscience and Remote Sensing Letters, 6(4), 772–776.

Çemberci, M., Civelek, M. E., & Canbolat, N. (2015). The moderator effect of global competitiveness index on dimensions of logistics performance index. Procedia - Social and Behavioral Sciences, 195, 1514–1524.

Çınaroğlu, S. (2021). Türkiye'de iller düzeyinde sağlık personeli dağılımı ve daha etkin politika ihtiyacı. Hacettepe Sağlık İdaresi Dergisi, 24(2), 235–254.

de Hoon, M. J. L., Imoto, S., Nolan, J., & Miyano, S. (2004). Open source clustering software. Bioinformatics, 20(9), 1453–1454.

Dışkaya, S., & Bozkurt, A. A. (2025). A bibliometric analysis of studies on logistics performance measurement. Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 53, 1–24.

Ergün, M. (2025). Efficiency in international logistics: Trade, emissions and the case of Türkiye. Başkent Üniversitesi Ticari Bilimler Fakültesi Dergisi, 9(2), 198–220.

Faria, R. N., de Souza, C. S., & Vieira, J. G. V. (2015). Evaluation of logistic performance indexes of Brazil in the international trade. RAM. Revista de Administração Mackenzie, 16(1), 213–235.

Filiz, M. (2023). OECD ülkelerinde sağlık hizmetleri arz ve talebi üzerinde bir değerlendirme. Adıyaman Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 43, 633–659.

Goçer, A., Özpeynirci, Ö., & Semiz, M. (2022). Logistics performance index-driven policy development: An application to Turkey. Transport Policy, 124, 20–32.

Güdelek, M., Dursunkaya, E., Kaya, E., Palamutoğlu, M., Akşahin, T., & Nebati, E. E. (2024). Veri zarflama analizi ile lojistik etkinlik ölçümü. Toplum, Ekonomi ve Yönetim Dergisi, 5(3), 548–567.

Gündoğdu, H. G., & Aytekin, A. (2020). Yönetişim göstergeleri bağlamında ülkelerin kümeleme analizi ve ARAS ile değerlendirilmesi. Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 66, 301–318.

Günher, E., Fidan, M., & Akbayır, Ö. (2025). SOM ve k-ortalama kümeleme algoritmaları kullanarak vagon tamire tutma verilerinin incelenmesi. Demiryolu Mühendisliği, 21, 168–177.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. Journal of the Royal Statistical Society: Series C (Applied Statistics), 28(1), 100–108.

Hepşen, A., Aydın, O., & Vatandaş, O. (2015). K-ortalama algoritması ile kümelenmiş konut fiyatlarının fonksiyonel veri analizi: İstanbul örneği. Finans Politik & Ekonomik Yorumlar, 52(604), 75–85.

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. Information Sciences, 622, 178–210.

İlkin, S., Aytar, O., Gençtürk, T. H., & Şahin, S. (2020). Use of K-means clustering algorithm for lesion segmentation in dermoscopic images. GU Journal of Science, Part C: Design and Technology, 8(1), 182–191.

İnaç, H., Ayözen, Y. E., Yelshibayev, R., & Issayeva, G. (2025). An application of logistics performance index-driven policy development to Turkey and Kazakhstan. Journal of the Knowledge Economy, 16, 4898–4917.

İnce, H., İmamoğlu, S. Z., & Keskin, H. (2013). Öz-düzenlemeli harita ağları ile k-ortalama kümeleme analizinin karşılaştırılması: Tüketici profilleme örneği. Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 28(4), 723–731.

Işık, Ö., Aydın, Y., & Koşaroğlu, Ş. M. (2020). The assessment of the logistics performance index of CEE countries with the new combination of SV and MABAC methods. LogForum, 16(4), 549–559.

İşler, Y., & Narin, A. (2012). WEKA yazılımında k-ortalama algoritması kullanılarak konjestif kalp yetmezliği hastalarının teşhisi. SDÜ Teknik Bilimler Dergisi, 2(4), 21–29.

Ju, M., Mirović, I., Petrović, V., Erceg, Ž., & Stević, Ž. (2024). A novel approach for the assessment of logistics performance index of EU countries. Economics, 18, Article 20220074. https://doi.org/10.1515/econ-2022-0074.

Kamacı, K. (2025). Orta Koridor ülkelerinin lojistik performansının LOPCOW–ağırlıklı TOPSIS yaklaşımıyla değerlendirilmesi. Tarsus Üniversitesi Uygulamalı Bilimler Fakültesi Dergisi, 5(2), 180–197.

Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2004). A local search approximation algorithm for k-means clustering. Computational Geometry, 28, 89–112.

Keskin, M. E. (2018). A regional analysis of the socio-economical properties of the Turkey cities. Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 32(4), 1135–1153.

Khan, S. S., & Ahmad, A. (2004). Cluster center initialization algorithm for k-means clustering. Pattern Recognition Letters, 25, 1293–1302.

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. Pattern Recognition, 36(2), 451–461.

MapChart. (2025, October 29). World map: Microstates. https://www.mapchart.net/detworld.html

Martí, L., Martín, J. C., & Puertas, R. (2017). A DEASBM model to evaluate the efficiency of the logistics performance index among 36 countries. Transportation Research Part A: Policy and Practice, 108, 1–11.

Martí, L., Puertas, R., & García, L. (2014). The importance of the Logistics Performance Index in international trade. Empirical Economics, 47(2), 525–537.

Mercangöz, B. A., Yıldırım, B., & Kuzu Yıldırım, S. (2020). Time period based COPRAS-G method: Application on the Logistics Performance Index. LogForum, 16(2), 239–250.

Modha, D. S., & Spangler, W. S. (2003). Feature weighting in k-means clustering. Machine Learning, 52, 217–237.

Moftah, H. M., Azar, A. T., Al-Shammari, E. T., Ghali, N. I., Hassanien, A. E., & Shoman, M. (2014). Adaptive k-means clustering algorithm for MR breast image segmentation. Neural Computing and Applications, 24, 1917–1928.

Mutlu, F., & Gül, S. (2023). Improving the performance of EM and K-means algorithms for breast lesion segmentation. Anatolian Current Medical Journal, 5(4), 492–497.

Niknam, T., & Amiri, B. (2010). An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis. Applied Soft Computing, 10, 183–197.

Orakçı, E. (2024). EATWOS, OCRA ve REF III teknikleriyle ülkelerin lojistik performans indeksine dayalı etkinliklerinin incelenmesi. İktisadi İdari ve Siyasal Araştırmalar Dergisi, 9(25), 590–611.

Orhan, M. (2019). Türkiye ile Avrupa Birliği ülkelerinin lojistik performanslarının entropi ağırlıklı EDAS yöntemiyle karşılaştırılması. Avrupa Bilim ve Teknoloji Dergisi, (17), 1222–1238.

Özdemir, K. (2023). Mindful leadership. In Ö. E. Arslan (Ed.), International research in social, human, and administrative sciences XVIII (pp. 1–25). Eğitim Yayınevi.

Özdemir, K. (2025a). Coaching-based leadership: A scale adaptation study. Sosyal Mucit Academic Review, 6(2), 292–313.

Özdemir, K. (2025b). Workplace mindfulness: A scale adaptation study. Journal of Economics Business and Political Researches, 10(27), 439–453.

Özdemir, U., & Yorulmaz, Ö. (2025). Türkiye’deki illerin insani gelişmişlik endeksine göre farklı kümeleme teknikleri ile sınıflandırılması. EKOIST Journal of Econometrics and Statistics, 42, 175–197.

Özyer, S. T. (2024). Bölümleyici kümeleme için doğru merkezi noktaların tayini. DUJE (Dicle University Journal of Engineering), 15(2), 277–284.

Pham, D. T., Dimov, S. S., & Nguyen, C. D. (2005). Selection of K in K-means clustering. Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 219, 103–119.

Pollard, D. (1981). Strong consistency of k-means clustering. The Annals of Statistics, 9(1), 135–140.

Puertas, R., Martí, L., & García, L. (2014). Logistics performance and export competitiveness: European experience. Empirical Economics, 48, 947–964.

Rezaei, J., van Roekel, W., & Tavasszy, L. (2018). Measuring the relative importance of the logistics performance index indicators using Best Worst Method. Transport Policy, 68, 158–169.

Selvi, H. Z., & Çağlar, B. (2017). Çok değişkenli haritalama için kümeleme yöntemlerinin kullanılması. Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi, 6(2), 415–429.

Şengöz, N., & Özdemir, G. (2016). Temel bileşenler analizi ve k-ortalama kümeleme yönteminin birlikte kullanımı: Bir örnek uygulama. Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 8(15), 85–94.

Shi, N., Liu, X., & Guan, Y. (2010). Research on k-means clustering algorithm: An improved k-means clustering algorithm. In Proceedings of the Third International Symposium on Intelligent Information Technology and Security Informatics (pp. 63–67). IEEE.

Simovici, D. A. (2022). Clustering: Theoretical and practical aspects. World Scientific Publishing.

Son, G. W., Cho, H. S., & Moon, H. C. (2020). The determinants of Korea's export using global Logistics Performance Index (LPI). Journal of Maritime Policy Studies, 35(2), 103–132.

Süslü, D. İ., Atalay, K. D., & Derya, T. (2025). Bütünleşik AHP ve k-ortalama kümeleme tabanlı yedek parça talep tahmini: Sağlık ekipmanları alanında bir uygulama. Black Sea Journal of Engineering and Science, 8(3), 660–671.

Syakur, M. A., Khotimah, B. K., Rochman, E. M. S., & Satoto, B. D. (2018). Integration k-means clustering method and elbow method for identification of the best customer profile cluster. IOP Conference Series: Materials Science and Engineering, 336, 012017. https://doi.org/10.1088/1757-899X/336/1/012017.

Tümtürk, A. (2024). Türkiye'nin lojistik performansına verdiği önemin değerlendirilmesi: Dernek yayınları üzerinde içerik analizi uygulaması. Journal of Transportation and Logistics, 9(2), 280–297.

Ülgen, G., & Arda Özalp, L. F. (2017). Refah rejimleri sınıflandırma çalışmaları: Cinsiyet boyutları. Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi, 39(2), 637–656.

Ulutaş, A., & Karaköy, Ç. (2019). An analysis of the logistics performance index of EU countries with an integrated MCDM model. Economics and Business Review, 5(4), 49–67.

Wang, H., & Song, M. (2011). Ckmeans.1d.dp: Optimal k-means clustering in one dimension by dynamic programming. The R Journal, 3(2), 29–33.

Win, T. K., Watanabe, D., & Hyodo, T. (2021). Data analysis of high-capacity vehicles by machine learning for sustainable logistics in Japan. Toros University FEASS Journal of Social Sciences, 8(Special Issue), 51–69.

World Bank. (2025). Logistics Performance Index (LPI): International LPI. (Retrieved: October 29, 2025), from https://lpi.worldbank.org/international/global.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z.-H., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. Knowledge and Information Systems, 14, 1–37.

Yalçın, B., & Ayvaz, B. (2020). Çok kriterli karar verme teknikleri ile lojistik performansın değerlendirilmesi. İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 19(38), 117–138.

Yılmaz, Ö., & Kaya, V. (2005). Genişleme sürecindeki Avrupa Birliği: Ekonomik performansa dayalı kümeleme analizi. Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 5(1), 361-376.

Yılmaz, Ö., & Temurlenk, M. S. (2005). Türkiye'deki istatistik bölgelerin kişi başına düşen gelir açısından hiyerarşik ve hiyerarşik olmayan kümeleme analizi ile değerlendirilmesi: 1965–2001. Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi, 19(2), 75-92.

Yücebaş, S. C., & Kınacı, A. C. (2016). K-ortalama kümelerinin sınıf bilgisi olarak karar ağacı oluşturmada kullanılması ve glokom çoklu sınıflandırılmasında başarına etkisi. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 4, 747–755.

Zha, H., Ding, C., Gu, M., & Simon, H. (2002). Spectral relaxation for K-means clustering. Advances in Neural Information Processing Systems, 14, 1057–1064.

Zhang, Y., Hato, T., Dagher, P. C., Nichols, E. L., Smith, C. J., Dunn, K. W., & Howard, S. S. (2019). Automatic segmentation of intravital fluorescence microscopy images by k-means clustering of FLIM phasors. Optics Letters, 44(16), 3928–3931.

Zorlutuna, Ş. (2024). Türkiye'de illerin toplumsal cinsiyet eşitliğine göre sınıflandırılması. Akademik Araştırmalar ve Çalışmalar Dergisi, 16(30), 19–35.