Journal of Intelligent Decision Support System (IDSS) 6 (3) (2023), pp. 112-120

Published by: Institute of Computer Science (IOCS)

Journal of Intelligent Decision Support System (IDSS)

Journal homepage: www.idss.iocspublisher.org

Optimization of k-means algorithm in grouping data using the statistical gap method

Alfiansyah Hasibuan[1], Djubir R.E. Kembuan[2], Christine Takarina Meitty Manoppo [3], Medi Hermanto Tinambunan[4]

[1,4] Informatics Engineering Study Program, Faculty of Engineering, Universitas Negeri Manado, Manado, Indonesia
[2] Building Engineering Education Study Program, Faculty of Engineering, Universitas Negeri Manado, Manado, Indonesia
[3] Information and Communication Technology Education Study Program, Faculty of Engineering, Universitas Negeri Manado, Manado, Indonesia

Article Info

Article history:
Received Aug 19, 2023
Revised Aug 22, 2023
Accepted Aug 23, 2023

Keywords:
Algorithm
clustering
EuStockMarkets
K-Means
Statistical gaps

ABSTRACT

In this study, we study the core concepts of the K-Means algorithm, explore its algorithmic framework, computation steps, and practical applications. Using data that is used as a basic need to perform calculations from the k-means algorithm optimization method. Using data taken from the r studio dataset with the EuStockMarkets dataset. The purpose of this study is to optimize the k-means algorithm and cluster the clustering process from a dataset, minimizing the objective function that has been set in the clustering process. The tools used are R Studio. Based on the results of this study, profiling of each group formed can be carried out. Based on the grouping results that have been carried out, the grouping results are 75.7% the accuracy of the statistical Gap method in optimizing clusters from existing datasets and the results of 92.9% are obtained from the results of minimizing the object functions in the dataset from grouping with k-means. The smaller the percentage in this grouping process the better it is in optimizing the clusters from the dataset. The author applies the k-means clustering algorithm to minimize objects for grouping from the EuStockMarkets dataset which consists of 4 variables. And the author uses the Statistical Gap method to optimize the clusters from the dataset.

Corresponding Author:

Alfiansyah Hasibuan,
Informatics Engineering Study Program, Faculty of Engineering,
Universitas Negeri Manado,
Jl. Kampus Unima, Tonsaru, Kec. Tondano Sel., Kabupaten Minahasa, Sulawesi Utara, 95618, Indonesia.
Email: alfiansyahhasibuan@unima.ac.id

Introduction

Optimization, a basic concept that transcends disciplinary boundaries, serves as the driving force behind this teaching. In essence, optimization involves the process of finding the best solution from a set of feasible alternatives, with the goal of maximizing the benefits that are often limited and objectives vary. Optimization provides tools and methodologies for navigating complex decision-making scenarios. (Abualigah et al., 2021). Whether it's designing aerodynamic vehicles, managing supply chains, allocating financial portfolios or refining machine learning algorithms, optimization principles play a critical role in shaping modern advancements. (Wang et al., 2020).

Optimization problems take many forms, from linear programming, in which a linear objective function is optimized under linear constraints, to more complex nonlinear, dynamic, and stochastic optimization challenges.(Alhaj et al., 2020). As computing power continues to soar, innovative algorithms and techniques have emerged to address these complex problems, enabling exploration of vast solution spaces and discovery of optimal or near optimal solutions.(G. Zhang et al., 2018). Clustering, a basic technique in analyzing data is emerging as a powerful method for determining the inherent relationships between data points. In essence, clustering seeks to group together data points that are similar while separating those that are different, providing a prism through which the complexity of hidden data can be explored and understood..(Arunkumar et al., 2019). Imagine a large set of data points representing many different individuals, products or events. These points have many attributes that collectively define their uniqueness (Kumar et al., 2014). The clustering step consists of partitioning these seemingly chaotic blocks into consistent subsets, each of which consists of data points with the same characteristics. (Arinik et al., 2020). The process is like arranging scattered stars into constellations to paint a meaningful picture of the night sky together.(Dhuhita, 2015)

Central to clustering lies in the concept of similarity - a metric that measures how similar or different two data points are. Data points that show a higher degree of similarity tend to be grouped together, forming groups that combine common characteristics. (George Seif, 2018) Clustering not only allows summarizing large data sets, but also uncovers hidden relationships and anomalies that could otherwise go unnoticed. (Hamaker et al., 2020). Clustering, the process of classifying data points into groups with similar characteristics, is at the heart of many fields including machine learning, data analysis, and pattern recognition. (Oyewole & Thopil, 2023) Among the many clustering techniques, the K-Means algorithm has become the foundation because of its simplicity, efficiency and effectiveness in determining the inherent data structure.(Arinik et al., 2020). The K-Means algorithm works on the principle of partitioning the data set into a predetermined number of clusters, each of which is marked with a central point called a centroid. (Žalik, 2008) By iteratively optimizing centroid positions and assigning data points to the nearest centroid, K-Means seeks to minimize intra-cluster variance, resulting in clusters that are internally cohesive and distinct from each other. (Pulkit Sharma, 2023) The name K-Means algorithm refers to the underlying process that divides the data into K distinct clusters, in order to minimize the squared Euclidean distance between each data point and a defined cluster center. This iteration optimization process includes interlacing tuning and determination steps, where data points are moved to clusters based on proximity, and centroids are recalculated to reflect the mean of points in each cluster.(Al-jabery et al., 2020)

One of the most attractive properties of the K-Means algorithm is its scalability – its efficiency allows it to be applied to data sets of all sizes and dimensions. This flexibility, together with the ability to interpret and reveal hidden structures, makes K-Means an invaluable tool for tasks such as customer segmentation, image compression, and anomaly detection. (J. Zhang et al., 2018)

In this study, we study the core concepts of the K-Means algorithm, explore its algorithmic framework, computation steps, and practical applications. Using data that is used as a basic need to perform calculations from the k-means algorithm optimization method. The purpose of this study is to optimize the k-means algorithm and group in the clustering process from a dataset, minimizing the objective function that has been set in the clustering process.

Method

In the article about Optimizing the K-Means Algorithm in Grouping Data Using the Statistical Gap Method, the method used is the k-means algorithm to minimize object functions in the dataset, and the statistical gap method to optimize clusters from the dataset.

The most popular unsupervised machine learning approach for dividing a given data set into a set of k groups, or k clusters, is called "K-means clustering," where k is the number of groupings the analyst has predetermined. It organizes things into groups in order to maximize intra-class similarity—the degree of similarity between objects within the same cluster—and minimize inter-class similarity—the degree of similarity between objects from different clusters. Each cluster in k-means clustering is represented by its centroid, or the mean of the points allocated to the cluster, as its center.(Al-jabery et al., 2020)

This research is to analyze the performance of K-means in clustering. The data used in this study are datasets sourced from R Studio, namely the USArrests dataset, and the tools used are R Studio.

Table 1. EuStockMarkets dataset sourced from R Studio dataset

No		DAX	SMI	CAC	FTSE
1	1,991,496	1628.75	1678.1	1772.8	2443.6
2	1,991,500	1613.63	1688.5	1750.5	2460.2
3	1,991,504	1606.51	1678.6	1718.0	2448.2
4	1,991,508	1621.04	1684.1	1708.1	2470.4
5	1,991,512	1618.16	1686.6	1723.1	2484.7
6	1,991,496	1628.75	1678.1	1772.8	2443.6
.	.	.	.	.	.

Source R Studio dataset

The clustering stages used in this study are:

Grouping Distance Measure

This calculation is known as the difference or distance matrix. The magnitude of the gaps is a crucial component of clustering. It specifies the formula used to determine how similar two elements (x, y) are, which will have an impact on the cluster's structure.(Wierzchoń & Kłopotek, 2018)

The choice of spacing size is an important step in grouping. It defines how the similarity of two elements (x, y) is calculated and it will affect the shape of the cluster.

The classic method for measuring distances is the Euclidean and Manhattan distance, which is defined as follows:

Euclidean distance (Pulkit Sharma, 2020):

$$\boldsymbol{d_{euc}(x, y) = \sqrt{\sum_{i=1}^{n}(x_1 - y_1)^2}} \quad (1)$$

Manhattan distance (Linear & Survival, n.d.):

$$\boldsymbol{d_{man}(x, y) = \sum_{i=1}^{n}|(x_1 - y_1|} \quad (2)$$

Where :

x and *y* are two vectors of length *n*.

Then to analyze in determining the distance based on correlation by subtracting the correlation coefficient using the following method (Greenacre, Michael (Universitat Pompeu Fabra, n.d.).

Pearson correlation distance(Edelmann et al., 2021):

$$\boldsymbol{d_{cor}(x, y) = 1 - \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}} \quad (3)$$

Spearman correlation distance:

$$d_{spear}(x, y) = 1 - \frac{\sum_{i=1}^{n}(x'_i - \bar{x}')(y'_i - \bar{y}')}{\sqrt{\sum_{i=1}^{n}(x'_i - \bar{x}')^2 \sum_{i=1}^{n}(y'_i - \bar{y}')^2}} \quad (4)$$

Where :

x'_i = rank (x_i) and y'_i = rank y_i

The Spearman correlation method calculates the correlation between rank variable *x* and rank variable *y*.

Additionally, computations are done to locate clusters so that the overall variation inside the cluster is kept to a minimum. The default approach is Hartigan-Wong's (1979) algorithm, which sums the squared distances of the Euclidean distance between objects and their respective centroids to define the total within-cluster variance. The Hartigan-Wong algorithm's equation:(Kencana, 2020)

$$W(C_k) = \sum_{x_i \in C_k}(x_i - \mu_k)^2 \quad (5)$$

Where:

x_i is data points belonging to the cluster C_k
μ_k is the average value of the points assigned to the cluster C_k

The defined formula of the total within-cluster variation is as follows:
$tot.withiness = \sum_{k=1}^{k} W(C_k) = \sum_{k=1}^{k} \sum_{x_i \in c_k} (x_i - \mu_k)^2$.. (6)

The in-cluster total of squares measures the compactness of the smallest possible clustering.
Stages of the k-means algorithm

1. Determine the number of clusters (K) to be created (by the analyst)
2. Pick k random items from the data collection to serve as the first cluster averages or centers.
3. Based on the Euclidean distance between the item and the center of mass, assign each observation to the closest centroid.
4. Calculate the new mean of all the data points in the cluster to update the cluster centroid for each of the k clusters. The average of all variables for observations in the k-th cluster is contained in a vector of length p, where p is the total number of variables..
5. minimize the sum of squares' overall value iteratively (Equation 6). To put it another way, keep repeating steps 3 and 4 until the cluster assignments stop shifting or the allotted number of iterations has been used. The R programming language sets the maximum number of iterations to 10 by default..

Determining Optimal Clusters

The analyst determines the number of clusters to be used, preferably using the optimal number of clusters. To assist the analysis process, the following is a method for determining optimal clusters, namely:

Gap Statistic Method

For observation data and reference data, the total intracluster variation is calculated using different k values. The gap statistics for a given k are defined as follows:(Mohajer et al., 2011)

$Gap_n(k) = E_n^* \log(W_{(k)}) - \log(W_{(k)})$.. (7)
Where:
E_n^* sample size n from the reference distribution
$log(W_{(k)})$ calculate the average

The steps of the Gap Statistics method are as follows:

1. Observed data clusters, varying the number of clusters from $k= k_1.....k_{max}$ and calculate accordingly W_k
2. Generate reference datasets B and each cluster with the number of clusters varying from $k= k_1.....k_{max}$ Calculate the approximate statistical gap presented in the calculation of formula number 7
3. Then calculate with the formula $\overline{w} = (1/B)\sum_b \log(W_{kb}^*)$, sthen calculate the standard deviation with the formula $sd_k = \sqrt{(1/B)\sum_b(\log(W_{kb}^*) - \overline{w})^2}$ and then determine the result using the formula $s_k = sd_k\ x\ \sqrt{1 + 1/B}$
4. Choose the number of clusters as the smallest using the formula$Gap_{(k)} \geq Gap(k+1) - s_{k-1}$

Results and Discussions

The data used for analysis is sourced from the R Studio dataset. Before clustering, data normalization is first performed to assimilate the units to standard values and no longer follow units of measure, and check multicollinearity to avoid strong correlations between two or more groups of variables.

Table 2. EuStockMarkets dataset sourced from R Studio normalized dataset

No	DAX	SMI	CAC	FTSE
1	-0.8314094	-1,021104	-0.7841071	-1,148792
2	-0.8453475	-1,014851	-0.8225346	-1,131796
3	-0.8519110	-1,020804	-0.8785387	-1,144083
4	-0.8385167	-1,017497	-0.8955984	-1,121353
5	-0.8411716	-1,015993	-0.8697504	-1,106712
6	-0.8481315	-1,025013	-0.8849146	-1,125039
.	.	.	.	.

The dataset above has been normalized using the R Studio application. To measure the similarity between observations and forming clusters, they use the distance metric. So, values with a high range will have a greater influence on grouping, therefore data with that range must be standardized.

Table 3. Data set for multicollinearity test results with a value of k = 4

cluster	DAX	SMI	CAC	FTSE
1	1.2193818	1.2503905	1.0224511	1.27259064
2	-0.6988512	-0.7445033	-0.5786131	-0.80182658
3	2.6381025	2.5360648	2.9035569	2.36748703
4	-0.1494421	-0.0803198	-0.2811928	0.02222111

The dataset has been tested for multicollinearity to relate linearity between the independent variables in multiple regression. The multicollinearity test is intended to see the relationship/correlation between each variable. A good regression model will not have a correlation between independent variables. The correlation value exceeds 0.8, this indicates the presence of multicollinearity. Based on the results of the multicollinearity test analysis in Table 3, can be concluded that there is multicollinearity in the research data, then cluster analysis is carried out using the K-Means method.

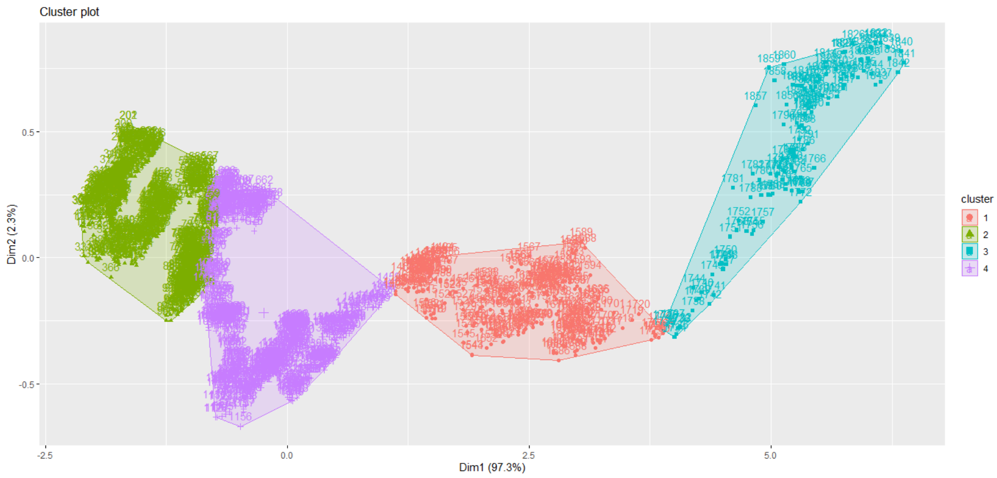

Figure 1. Grouping vector

Cluster sizes of k=4 and 269, 834, 132, and 625 are the outcomes. For the four groups and the four variables (DAX, SMI, CAC, FTSE), we examine cluster centers (means). Author also receives categorization for each result that is shown. Since the algorithm's number of clusters (k) must be established before it can be used, it is frequently beneficial to experiment with a range of k values and look for variations in the outcomes. The same procedure can be used to generate results for clusters 4, 5, 6, and 7, as indicated in the image:

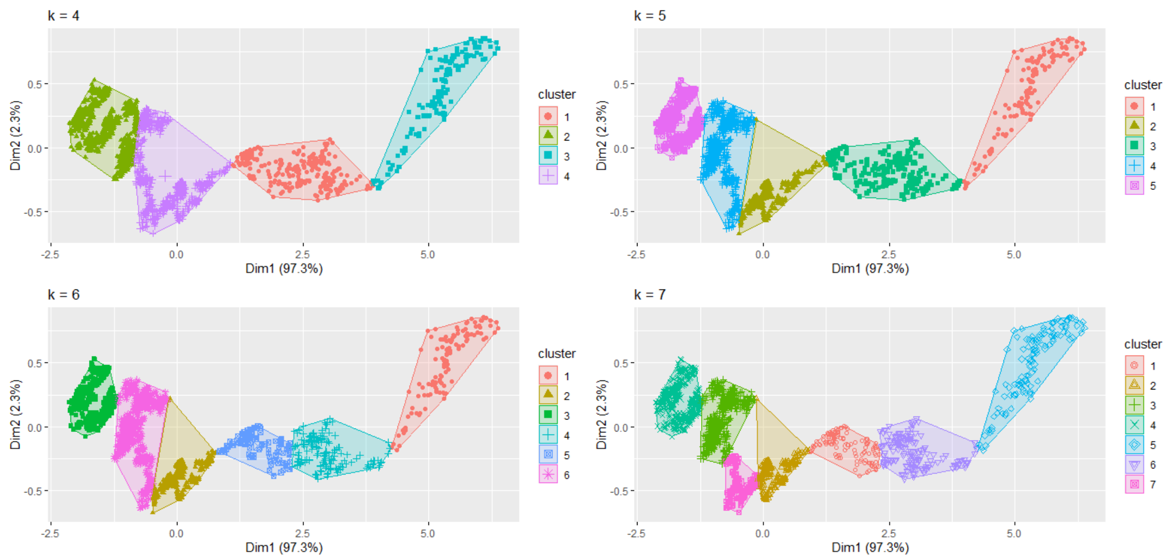

Figure 2. The results of the number of clusters k = 4 to k = 7

The optimal number of clusters is not revealed by the aesthetic appearance of this clustering, even though it does inform us where actual dilineations do or do not occur, such as cluster 4 in a graph with k = 5, k = 6, and k = 7 between clusters. This study incorporates the Statistical Gap approach from the Determining Optimal Clusters method to optimize clusters. To enhance the results of clustering visualization, use the statistical gap method. This method, such with the K-Means method, can be used with any clustering algorithm. The statistic contrasts the overall within-cluster variance for k values that deviate from what would be predicted in a reference distribution with zero data and no discernible clustering.

Table 4. Calculation results using Gap Statistics with R Studio

	logW	E.logW	gap	SE.sim
[1,]	6,879992	7,235699	0,355707	0,00779
[2,]	6,244525	6,632501	0,387976	0,008883
[3,]	6,035270	6,333887	0,298617	0,007273
[4,]	5,695372	6,156071	0,460699	0,007127
[5,]	5,457508	6,040222	0,582714	0,005859
[6,]	5,333695	5,957969	0,624274	0,006197
[7,]	5,277589	5,900392	0,622803	0,006484
[8,]	5,160304	5,856121	0,695816	0,00673
[9,]	5,092778	5,816264	0,723486	0,006367
[10,]	4,998546	5,777439	0,778892	0,00616

It can be seen from table 4 above, explaining that the average Gap value has been sorted from the largest value to the smallest value. From the table above it can also be seen that the highest Statistical Gap value is 0.778892 and the lowest Statistical Gap value is 0.355707 and for the average gap the highest is 6.879992 and the lowest is 4.998546 from the list of tables shown in this article. From the above values, it can be concluded that in obtaining optimal clusters based on the Statistical Gap method, it can also be seen from the difference in Gap(k)-Gap(k+1)-sk+1 which is referred to as the SE.sim value. This SE.sim value is a necessary value in obtaining the optimal number of clusters using the statistical gap method where optimal clusters are obtained if the SE.sim value ≥ 0.

To make it easier, the results can be seen in the image below where the SE.sim value and the number of clusters are displayed graphically.

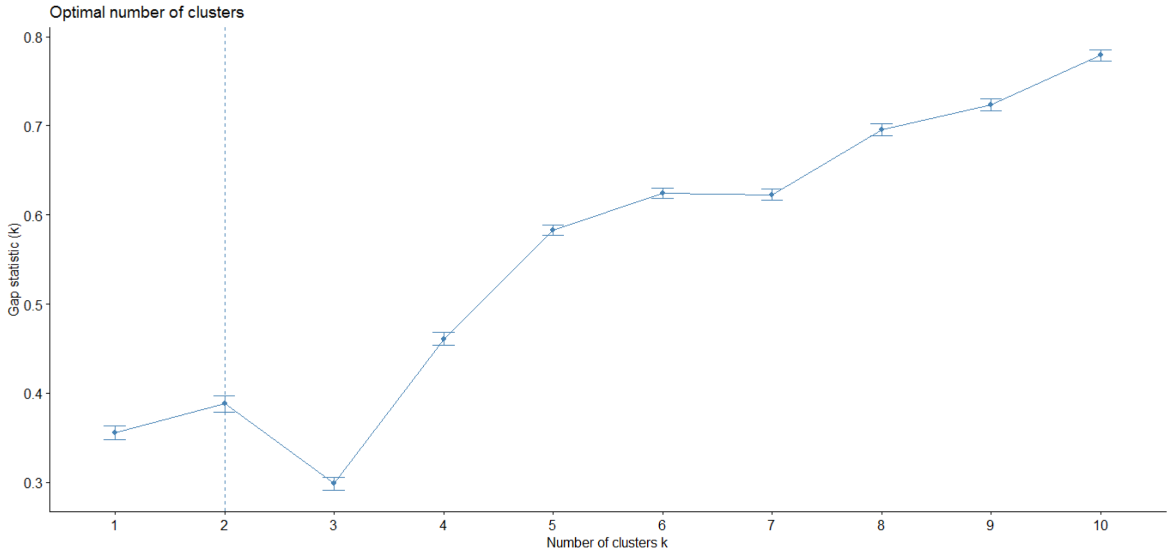

Figure 3. Graphical gap statistical Method

The value of B is the number of samples counted, and it has been obtained that the number of k = 2 is the most optimal for cluster formation. Thus, compared with the previous method, it can be concluded that the optimal value of k to form a cluster is 2. In accordance with calculations using Gap Statistics that the value of k = 2 is the best cluster.

Table 5. Cluster results using the statistical Gap method, finding that grouping optimization at k = 2

cluster	DAX	SMI	CAC	FTSE
1	-0,45944	-0,455771	-0,447269	-0,4440773
2	1,703984	1,690390	1,658859	1,647021

The cluster uses k-means with 2 clusters on the 1465th row, 395 is the best cluster. The results above are the number of random sets to be selected. in accordance with some previous references, the nstart value that is often used is 25. Then from the output of the k-means results with k = 2, cluster 1 is formed with 1465 datasets, cluster 2 with 395 datasets.

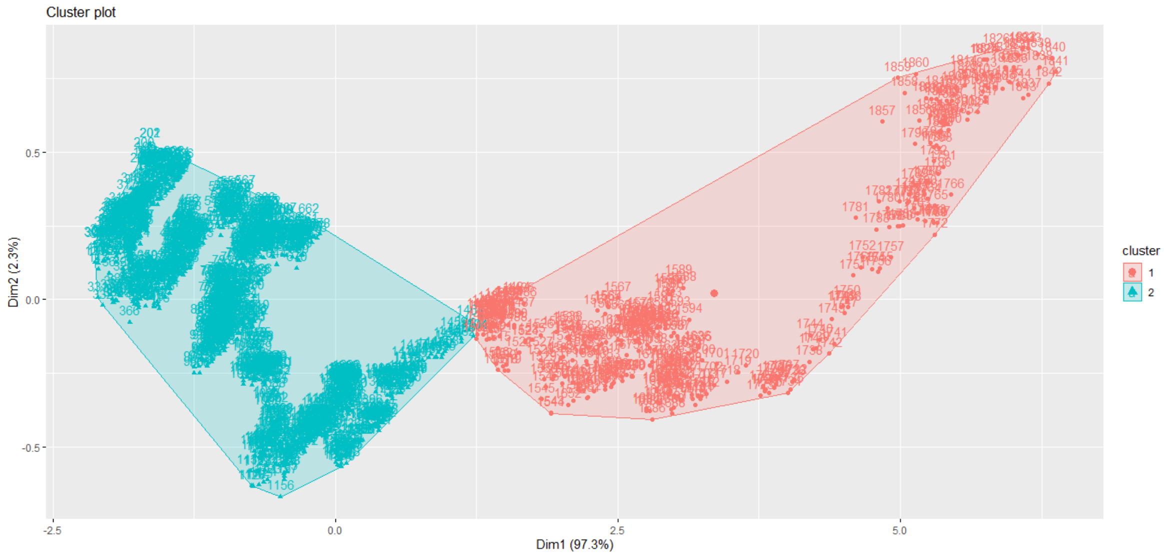

Figure 6. Visual results of k-means clustering using the Gap Statistics method

From the calculations that have been done using the R Studio application, the data grouping is obtained then the author can extract the clusters and add our initial data to do some descriptive statistics at the cluster level. From the figure it can be seen that the statistical gap method minimizes the objective function that has been set in the clustering process.

Table 5. Optimization final results using gap statistical Method

NO	Gap Statistik	k-means
1	(between_SS / total_SS = 75.7 %)	(between_SS / total_SS = 92.9 %)

Based on the grouping results that have been carried out, the grouping results are 75.7% the accuracy of the statistical Gap method in optimizing clusters from existing datasets and the results of 92.9% are obtained from the results of minimizing the object functions in the dataset from grouping with k-means. The smaller the percentage in this grouping process the better it is in optimizing the clusters from the dataset.

Conclusions

From the results of the analysis that has been carried out above, it can be concluded that applying the k-means grouping algorithm to minimize objects for grouping from the EuStockMarkets dataset which consists of 4 variables. The author also uses the Statistical Gap method to optimize clusters from the dataset. After considering the experiments that have been carried out more carefully and carefully, we determine that k = 2 clusters are optimal and provide the most meaningful insight into the segmentation of the dataset. From the calculations that have been done, it is found that by applying the Statistical Gap method in clustering, the grouping results are 75.7%, the accuracy of the Gap statistical method in optimizing clusters from existing datasets and the results of 92.9% are obtained from the results of minimizing the object functions in the dataset. from clustering with k-means. The smaller the percentage in this grouping process the better it is in optimizing the clusters from the dataset. The clustering of the K-means algorithm is a very simple and fast clustering method. Moreover, it can efficiently handle very large data sets. However, there are some drawbacks to the k-means algorithmic approach. One of the potential weaknesses of the clustering K-means algorithm is that the researcher must first determine the number of groups from the dataset. Therefore, suggestions for future researchers, as it’s known that the k-means algrotima still has deficiencies in grouping, therefore it’s hoped that future researchers can use other algorithms in clustering data. In addition, it’s hoped that future researchers can optimize the k-means algorithm in determining groups from datasets using other methods, so that a comparison of the results of each method is used to see which one is more efficient in clustering datasets.

References

Abualigah, L., Diabat, A., Mirjalili, S., Abd Elaziz, M., & Gandomi, A. H. (2021). The Arithmetic Optimization Algorithm. *Computer Methods in Applied Mechanics and Engineering*, *376*, 113609. https://doi.org/10.1016/j.cma.2020.113609

Al-jabery, K. K., Obafemi-Ajayi, T., Olbricht, G. R., & Wunsch II, D. C. (2020). Clustering algorithms. In *Computational Learning Approaches to Data Analytics in Biomedical Applications*. https://doi.org/10.1016/b978-0-12-814482-4.00003-6

Alhaj, Y. A., Al-qaness, M. A. A., Dahou, A., Abd Elaziz, M., Zhao, D., & Xiang, J. (2020). Effects of Light Stemming on Feature Extraction and Selection for Arabic Documents Classification. In *Studies in Computational Intelligence* (Vol. 874). https://doi.org/10.1007/978-3-030-34614-0_4

Arinik, N., Figueiredo, R., & Labatut, V. (2020). Multiple partitioning of multiplex signed networks: Application to European parliament votes. *Social Networks*, *60*, 83–102. https://doi.org/10.1016/j.socnet.2019.02.001

Arunkumar, N., Mohammed, M. A., Abd Ghani, M. K., Ibrahim, D. A., Abdulhay, E., Ramirez-Gonzalez, G., & de Albuquerque, V. H. C. (2019). K-Means clustering and neural network for object detecting and identifying abnormality of brain tumor. *Soft Computing*, *23*(19), 9083–9096.

https://doi.org/10.1007/s00500-018-3618-7

Dhuhita, W. M. P. (2015). Clustering Metode K-Means Untuk Menentukan Status Gizi Balita. *Jurnal Informatika*, *15*(2), 160–174.

Edelmann, D., Móri, T. F., & Székely, G. J. (2021). On relationships between the Pearson and the distance correlation coefficients. *Statistics and Probability Letters*, *169*, 108960. https://doi.org/10.1016/j.spl.2020.108960

George Seif. (2018). *The 5 Clustering Algorithms Data Scientists Need to Know*. Towards Data Science. https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68

Greenacre, Michael (Universitat Pompeu Fabra, B. (n.d.). *Chapter 6 Measures of distance and correlation between variables*. 1–10.

Hamaker, E. L., Mulder, J. D., & van IJzendoorn, M. H. (2020). Description, prediction and causation: Methodological challenges of studying child and adolescent development. *Developmental Cognitive Neuroscience*, *46*(January), 100867. https://doi.org/10.1016/j.dcn.2020.100867

Kencana, E. N. (2020). Sains Data dengan R : Klasterisasi Menggunakan K-Means Clustering Sains Data dengan R : Klasterisasi Menggunakan K-Means Clustering. *ResearchGate*, *June*. https://doi.org/10.13140/RG.2.2.29495.34721

Kumar, V., Chhabra, J. K., & Kumar, D. (2014). Parameter adaptive harmony search algorithm for unimodal and multimodal optimization problems. *Journal of Computational Science*, *5*(2), 144–155. https://doi.org/10.1016/j.jocs.2013.12.001

Linear, T., & Survival, A. (n.d.). *Manhattan distance Mathematica domain of a scalar*. 219–233. https://doi.org/10.1016/B978-0-12-409520-5.50020-5

Mohajer, M., Englmeier, K.-H., & Schmid, V. J. (2011). *A comparison of Gap statistic definitions with and without logarithm function*. *096*. http://arxiv.org/abs/1103.4767

Oyewole, G. J., & Thopil, G. A. (2023). Data clustering: application and trends. In *Artificial Intelligence Review* (Vol. 56, Issue 7). Springer Netherlands. https://doi.org/10.1007/s10462-022-10325-y

Pulkit Sharma. (2020). 4 Types of Distance Metrics in Machine Learning. *Analytics Vidhya*. https://www.analyticsvidhya.com/blog/2020/02/4-types-of-distance-metrics-in-machine-learning/

Pulkit Sharma. (2023). *The Ultimate Guide to K-Means Clustering: Definition, Methods and Applications*. https://www.analyticsvidhya.com/blog/2019/08/comprehensive-guide-k-means-clustering/

Wang, C., Koh, J. M., Yu, T., Xie, N. G., & Cheong, K. H. (2020). Material and shape optimization of bi-directional functionally graded plates by GIGA and an improved multi-objective particle swarm optimization algorithm. *Computer Methods in Applied Mechanics and Engineering*, *366*, 113017. https://doi.org/10.1016/j.cma.2020.113017

Wierzchoń, S. T., & Kłopotek, M. A. (2018). Cluster Analysis. In *Studies in Big Data* (Vol. 34). https://doi.org/10.1007/978-3-319-69308-8_2

Žalik, K. R. (2008). An efficient k′-means clustering algorithm. *Pattern Recognition Letters*, *29*(9), 1385–1391. https://doi.org/10.1016/j.patrec.2008.02.014

Zhang, G., Zhang, C., & Zhang, H. (2018). Improved K-means algorithm based on density Canopy. *Knowledge-Based Systems*, *145*, 289–297. https://doi.org/10.1016/j.knosys.2018.01.031

Zhang, J., Xiao, M., Gao, L., & Pan, Q. (2018). Queuing search algorithm: A novel metaheuristic algorithm for solving engineering optimization problems. *Applied Mathematical Modelling*, *63*, 464–490. https://doi.org/10.1016/j.apm.2018.06.036