Available online at www.HEFJournal.org

Journal of
Human, Earth, and Future

ISSN: 2785-2997

Vol. 5, No. 3, September, 2024

Time Series Clustering Analysis for Increases Food Commodity Prices in Indonesia Based on K-Means Method

M. Fariz Fadillah Mardianto [1*], N. Ramadhan Al Akhwal Siregar [1], Steven Soewignjo [1], F. Friska Rahmana Putri [1], Hadi Prayogi [1], Citra Imama [1], Dita Amelia [1], Sediono [1], Deshinta Arrova Dewi [2]

[1] Department of Mathematics, Faculty of Science and Technology, Airlangga University, Surabaya, Indonesia.

[2] Faculty of Information Technology, International University (INTI), Nilai, Malaysia.

Received 05 March 2024; Revised 23 July 2024; Accepted 04 August 2024; Published 01 September 2024

Abstract

The global food crisis is perceived to have a significant impact on the national food sector. Time series clustering, a potent data mining technique, is employed to decipher and interpret intricate temporal patterns. Dynamic Time Warping (DTW), a measure that currently appears to be the most relevant, is predicated on the distance between sequences of elements. This paper explores the application of DTW in data mining algorithms to cluster commodity prices in Indonesia, aiming for enhanced accuracy based on time series movement. The clustering algorithm employs the K-Means method, necessitating a comprehensive description of the groups it forms. The analysis results reveal time series clustering for commodity prices using K-Means. Optimal results are achieved with five clusters, based on the commodity price trend. Influencing factors include seasonal variations and government policies related to consumer demand. It is imperative for the government to establish a robust market monitoring system to track commodity price fluctuations in real-time, thereby facilitating the design of effective price stabilization policies. The insights gleaned from this study can guide decision-makers in implementing targeted interventions to stabilize prices, bolster food security, and ensure sustainable economic growth.

Keywords: Food Commodity Prices; Dynamic Time Wraping; K-Means; Sustainable Economy Growth; Time Series Clustering.

1. Introduction

The issue of the 2023 global economic recession will make Indonesia faced with problems in achieving prosperity for its citizens. This issue was reinforced by the World Bank in its report, which predicts the possibility of a global economic recession in 2023 [1]. In this study, a time series clustering approach is used to predict food commodity prices in Indonesia, while the DTW approach is used to group provinces in Indonesia based on their economic growth as an effort to prevent and address economic recessions, especially those related to food commodity prices in the country. The high food and energy commodity prices influenced the supply chain to trigger inflation in several sectors; at the same time, Indonesia's inflation rate reached 5.71% year on year (YOY) [2]. Efforts to control inflation by several central banks around the world have generally increased the benchmark interest rate [3].

According to the Central Bureau of Statistics (BPS), there are three global phenomena that have triggered a spike in food commodity prices. First, climate change is affecting global food production. Second, the conditions after the

* Corresponding author: m.fariz.fadillah.m@fst.unair.ac.id

http://dx.doi.org/10. 28991/HEF-2024-05-03-02

COVID-19 pandemic. Third, the war between Russia and Ukraine, which further exacerbated food and energy scarcity. Some previous studies about the topic provided some information. According to research by Malau et al. [4], climate phenomena such as El Nino Southern Oscillation (ENSO), which consists of El Nino and La Nina, have the potential to impact on food commodity conditions. El Nino, which is followed by a decrease in rainfall, has a greater impact on rice and corn production. La Nina, which is accompanied by increased rainfall, has a greater impact on soybean production. The impact of the spread of COVID-19 has been felt by all countries, and its effects are widespread, bringing hard times and threatening food security that depends on world oil prices for millions of people [5]. The condition of food security in Indonesia is exacerbated by the impact of the Russia-Ukraine war, which has disrupted the availability of fertilizers, which are part of the food production process [6]. Therefore, the problem of food security is very urgent to overcome because the fulfillment of food needs is very important and strategic to maintain the sovereignty of the country, though not depending on food. Thus, the fulfillment of food needs is very important and strategic to maintain state sovereignty, though not depending on food imports from developed countries [7].

To realize strong, sustainable, balanced, and inclusive global economic growth, a high-level conference was held by the heads of state through the G20 Meeting. One of the expected achievements is the creation of national food security through coordinated stimulus on a large scale [8]. Issues related to food security, cases of hunger, improved nutrition, and sustainable agriculture have also been formulated in the number two Sustainable Development Goals (SDGs). The National Strategic Food Price Information Center (PIHPS) in Indonesia states that several national strategic food commodities that contribute significantly to the formation of volatile inflation figures are rice, shallots, garlic, cayenne pepper, granulated sugar, beef, chicken, and cooking oil.

Time series data analysis has become an increasingly critical aspect of statistical research and data-driven decision-making in various fields. With the proliferation of data collection technologies, temporal data sequences have become prevalent in domains such as finance, healthcare, environmental sciences, and beyond. Extracting meaningful patterns and structures from these time-varying data streams is essential for making informed decisions, generating accurate predictions, and gaining valuable insights into underlying processes [9].

Time series clustering, as a powerful data mining technique, has emerged as a vital tool to explore and interpret complex temporal patterns. The objective of time series clustering is to discover patterns in an unlabeled time series dataset by systematically grouping data into cohesive clusters, where objects within the same cluster exhibit maximum similarities, while minimizing similarities between objects belonging to different clusters. Clustering methods group similar time series together, providing a foundation for better understanding the dynamics and similarities among various sequences [9]. Among the wide array of distance metrics available for time series comparison, Dynamic Time Warping (DTW) has garnered significant attention for its capability to handle temporal distortions and variations, enabling the comparison of sequences with differing lengths and shifts in the temporal domain.

DTW is a distance determination method that evaluates the similarity or difference between a pair of time series data. This method determines the optimal alignment between two-time series by stretching or compressing certain segments of the sequences. The ability of DTW to capture nonlinear alignments between sequences has demonstrated promising results in various domains, from anomaly detection and pattern recognition to forecasting and event analysis. Chen et al. [10] explored urban functional areas by utilizing social media data at the building level and employing DTW distance with k-medoids clustering. Their research showcased improved spatial understanding and granularity in uncovering spatial patterns and user interactions [11], but they faced challenges related to computational complexity and parameter tuning. Izakian et al. [12] explored fuzzy clustering with DTW distance, seeking to unravel complex time series patterns with soft cluster assignments. Their journey uncovered the robustness of DTW in handling temporal variations and enhancing cluster interpretability, but they encountered the challenge of parameter tuning while balancing granularity and uncertainty in diverse domains. In their quest for enhanced time series clustering, Huang et al. [13] presented an innovative method incorporating smooth subspace clustering. Their journey unveiled the advantages of Dynamic Time Warping (DTW) distance, showing Time Series k means outperformed traditional k-means in handling nonlinear patterns and reducing sensitivity to outliers. Time Series K-Means automatically assigns weights to timestamps based on the significance of the time interval during the batching process. K-Means Clustering is a clustering method that is influenced by initialization and requires determining the number of clusters [14].

Despite the existing body of research on the economic impacts of global phenomena and recessions, a critical literature gap persists in the specific context of Indonesia. Previous studies have acknowledged the global nature of challenges, such as climate change, post-COVID-19 conditions, and geopolitical conflicts like the Russia-Ukraine war, affecting food commodity prices. However, the unique economic landscape of Indonesia, with its diverse provinces and complex socio-economic factors, demands a more granular examination. Existing research, while providing valuable insights, often lacks a focused analysis on how these global phenomena directly influence Indonesia's economic dynamics, particularly in the realm of food commodity prices. Furthermore, the utilization of time series clustering and Dynamic Time Warping (DTW) in the economic analysis of Indonesia represents an unexplored frontier in the literature. While time series analysis has gained traction in various domains, its application to predict food commodity prices and group provinces based on economic growth is a distinctive feature of this study. Thus, the research gap lies in the absence of a comprehensive investigation into the intricacies of Indonesia's economic challenges, specifically tied to global events, and the unique methodologies employed to address these issues in this study.

The introduction highlights the unique use of the K-Means clustering algorithm and the specific focus on Indonesia's economic landscape. The literature review surveys prior research on time series clustering, identifying gaps, and setting the stage for the current study. The methodology details the research design, justifying the choice of K-Means, and outlining data collection and analysis. The results section presents key findings from the time series clustering analysis, discussing clusters and their implications for Indonesia's economic challenges. In summary, the article provides a well-structured exploration, contributing valuable insights into the economic dynamics of rising food commodity prices in Indonesia.

2. Literature Review

2.1. Dynamic Time Warping (DTW)

Time series data cluster analysis has been extensively developed. For time series data, a variety of clustering algorithms can be used, such as hierarchical algorithms, grid-based algorithms, and boundary detection [15]. Instead of merely utilizing linear distance calculations like Euclidean, Manhattan, Canberra, and others, distance calculations employing Dynamic Time Warping (DTW), sometimes referred to as non-linear sequence alignment, are more realistic for usage in pattern comparisons. Under certain limitations, this approach is a well-known method for determining the best alignment between two supplied (time-dependent) sequences. K-Means is the clustering algorithm used in this study by DTW to cluster time series data [16, 17]. The only difference between this K-Means method and the standard one is how the closeness between members is determined.

Let there be sequences by Y and Z represented time series data respectively:

$$\mathrm{Y} = y_1, y_2, \dots, y_i, \dots., y_n \tag{1}$$

$$\mathrm{Z} = z_1, z_2, \dots, z_j, \dots., z_m \tag{2}$$

Prior to obtaining the best warping path, the distance between two points is represented by the elements of an n×m distance matrix, which is made up of $D(i,j) = d(y_i, z_j)$. By using the reiteration, the cumulative distance matrix is as follows:

$$M_{DTW} = d(y_i, z_j) + min \begin{cases} M_{DTW}(i-1, j-1) \\ M_{DTW}(i-1, j) \\ M_{DTW}(i, j-1) \end{cases} \tag{3}$$

Where $\mathrm{M_{DTW}}$ is cumulative distance, $d(y_i, z_j)$ is the minimum value of the adjacent elements:

$$\mathrm{DTW(y,z) = min}\Sigma_{\mathrm{k=1}}^{\mathrm{c}} \omega_{\mathrm{k}} \tag{4}$$

2.2. DTW Barycenter Averaging

The DTW Barycenter Average (DBA) is a technique that builds upon DTW. DBA is an averaging method that minimizes the squared distance (DTW) to the average sequence by iteratively adjusting an initial (potentially arbitrary) average sequence [18]. The calculation of each average sequence coordinate as the barycenter of the corresponding coordinates of the sequence set is the fundamental idea of DBA. Rather, one or more coordinates from each sequence contribute to the updating of each average coordinate. Furthermore, by calculating the barycenter of this set of coordinates, the partial sum for each coordinate in the mean series can be minimized. In order to reduce the total Sum Within Groups of Squares (WGSS), each coordinate will therefore minimize its portion of the WGSS. Following the computation of each barycenter, the revised mean sequence is ascertained..

To get the coordinates of the average sequence at iteration $i + 1$, let's take the set of sequences to be averaged ($S = S_1 . \dots . S_N$), the average sequence at iteration $C = < C_1 . \dots . C_T >$, and the update of C at iteration $C' = < C'_1 . \dots . C_T >$. Furthermore, each average sequence coordinate is defined in an arbitrary vector space E, which is often a Euclidean space:

$$\forall t \in [1.T]. C_t \in E \tag{5}$$

We take into consideration the function assoc, which associates each average sequence coordinate with one or more S sequence coordinates. This function is calculated between C and every S sequence during the DTW computation. Next, we define the t^th coordinate of the average sequence C'_t as follows:

$$C_t' = barycenter\ (assoc\ (Ct)) \tag{6}$$

where,

$$barycenter\ \{X_1.,,,.X_\alpha\} = \frac{X_1 + ,,, + X_\alpha}{\alpha} \tag{7}$$

(The vector addition is the addition of X_i.).

2.3. K-Means Clustering

The goal of K-Means Cluster Analysis, is to divide up existing objects into one or more clusters or groups of objects with similar characteristics are grouped together in one cluster, and objects with dissimilar characteristics are grouped together in a different cluster. Minimizing the objective function, which essentially aims to maximize variance across clusters and decrease variation inside one cluster, is the aim of the clustering process. The K-Means technique aims to combine all available data into a single group, where each group's data possesses unique attributes compared to the other groups data [19].

However, K-Means can run into some difficulties when applied to high-dimensional data. The main problem is "*Curse of Dimensionality*" [20]. In a high dimensional space, all data points tend to be far from each other and therefore are almost evenly distributed. This can reduce the effectiveness of K-Means clusters, because the distance between data points and cluster centroids may be similar for all clusters [21]. This results in clusters that are less clear and may not represent well the patterns present in the data. In this study, we will compare the effectiveness of the K-Means clustering method with the use of Principal Component Analysis (PCA) to overcome the "Curse of Dimensionality".

2.4. Clustering Performance Evaluation

The rand index and silhouette score are used to pick the optimum cluster criterion based on performance evaluation. The similarity of the outcomes from two distinct clustering techniques can be compared using the rand index [22]. Frequently represented as R, the rand index is determined by:

$$R = \frac{a + b}{\binom{n}{2}} \tag{8}$$

a: The frequency with which two items are part of the same clusters using two different clustering techniques.

b: The frequency with which two items are part of different clusters using two different clustering techniques.

The rand index always takes on a value between 0 and 1 where:

0: Denotes a disagreement between two clustering techniques regarding the clustering of any pair of elements.

1: Shows that every element pair's clustering is exactly agreed upon by two clustering techniques.

Each cluster is represented by a silhouette score. Silhouette score is a method used to see cluster density. All groupings are shown by plotting all the silhouettes into a single diagram, it is possible to compare the quality of the clusters [23].

$$s(i) = \frac{a(i) - b(i)}{max\ (a(i).b(i))} \tag{9}$$

It is simple to determine whether a cluster has the ideal value for a certain cluster based on the silhouette score values. A score of less than zero for a given point indicates that it can be mapped to a cluster other than the one that contains the data point [23].

3. Method

3.1. Algorithm of K-Means

At this classification stage, as many as 10 types of existing food commodities will be combined using the K-Means method with the following steps:

1. Determine the value or amount of k.
2. Initialize cluster center k by randomizing data objects.

3. Determine the distance between each given object and the center to assign each object to its matching cluster center. Next, choose the center that is closest to the object.
4. Compute the center of the new cluster by the average of all the members contained in each cluster.
5. If no N objects change membership, clustering is complete. If not, repeat steps 3 to 5.

3.2. Algorithm of Dynamic Time Warping

1. To ensure consistent intervals, normalize the data to be analyzed and reorganize it on a weekly basis. 2.
2. Choose point K to be the medoid that was initially grouped.
3. Compute the DTW distance matrix successively from the remaining sample points to the K medoid.
4. Using the DTW distance, set the sample point to the closest medoid.
5. Determine the new medoid by computing the sample points' absolute minimum error distance for each class cluster.
6. When the medoid of each cluster type stops changing, go back to step 3 and end the process.

Figure 1 shows the flowchart of the research methodology through which the objectives of this study were achieved.

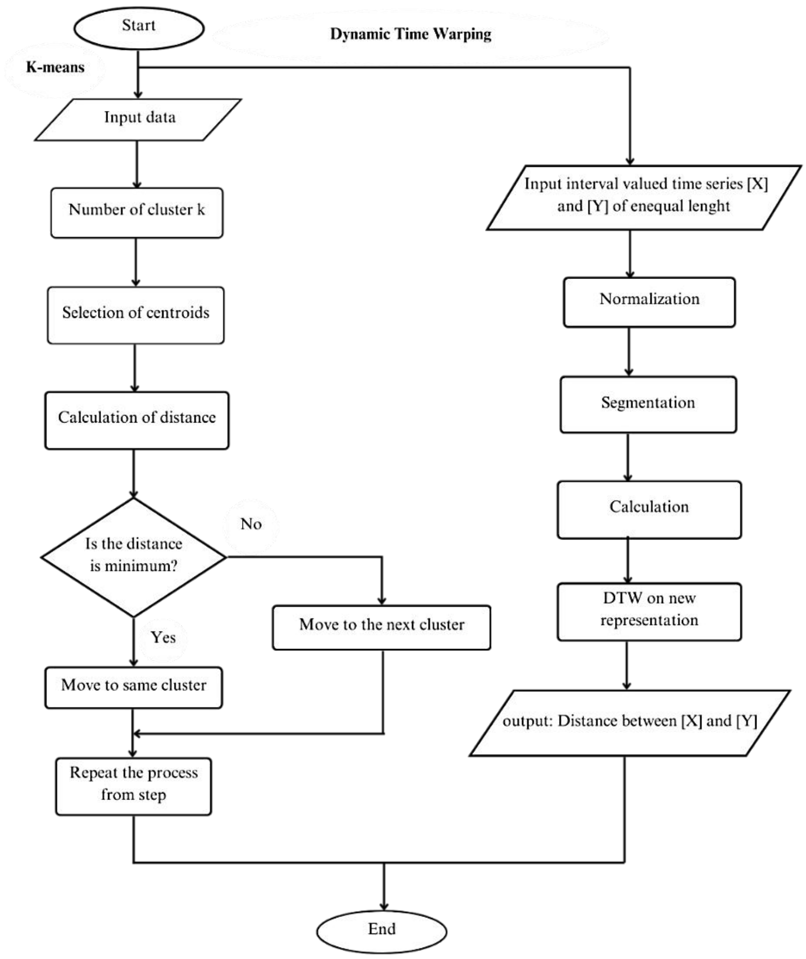

Figure 1. Flowchart of the methodology

4. Result and Discussion

4.1. Time Series Plot

The data used in this study consisted of prices for 10 commodities, namely sugar, cooking oil, beef, chicken meat, onion, garlic, red chili, cayenne pepper, chicken eggs, and rice prices. The observation period was carried out on a weekly basis from March 2020 to February 2023, with a total of 157 weeks of observations. This data is obtained from the National PIHPS page. Figure 2 shows the time series plot for every commodity.

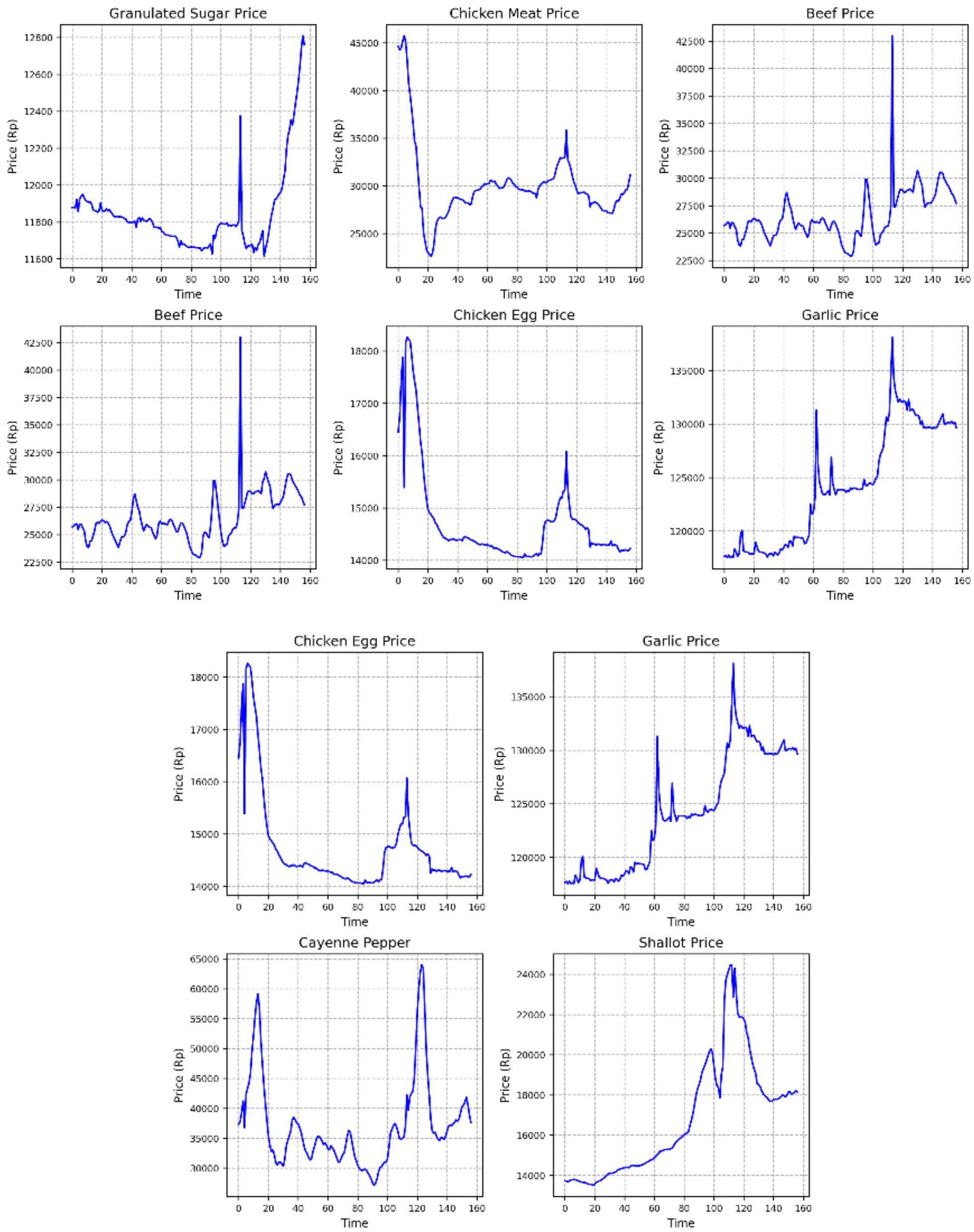

Figure 2. Time Series Plot for Commodities Price in Indonesia

4.2. Min-Max Normalization

Before time series clustering analysis is performed, the data needs to be normalized to scale the data into a uniform range, which is [0, 1]. Normalization is carried out with the aim that each commodity has the same range of values. This is important to do to prevent the dominance of one commodity that has a larger scale than other commodities. By normalizing, each commodity will make a balanced contribution to clustering analysis. The normalization method used is min-max normalization with each value in the data converted into a range of [0, 1] by maintaining the relative proportion between values. This normalization process can be carried out using the following formula:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{10}$$

Where x is the original value in the data and x' is the normalized value. By applying this normalization, each commodity will have a value range between 0 and 1, thus allowing for a more balanced and objective clustering analysis.

4.3. K-Means Clustering with Principal Component Analysis (PCA)

This study our focus is on how effectively the Principal Component analysis (PCA) solves the Curse of Dimensionality in comparison to the K-Means clustering method. Moreover, this study employs PCA to mitigate the

Curse of Dimensionality problem by reducing the dimensionality of the data. This technique accomplishes dimensionality reduction by projecting data into lower-dimensional spaces while preserving the maximum variation in the data. In the context of clustering, PCA aids in compressing data, thereby enhancing the clarity of cluster patterns.

Before proceeding to the clusterization, we will estimate the optimal cluster count based on changes in the elbow. In this research, the Elbow Method was employed for determining the optimal number of clusters, ranging from 2 to 9, using the K-Means clustering algorithm. the choice of evaluating the optimal number of clusters within the range of 2 to 9 was influenced by the inherent nature of the dataset, where the commodities under investigation totalled 10, thus this research don't include cluster 10 because it will have no meaning in this cluster analysis. The elbow of each cluster count, ranging from 2 to 9, is depicted in Figure 3 as follows:

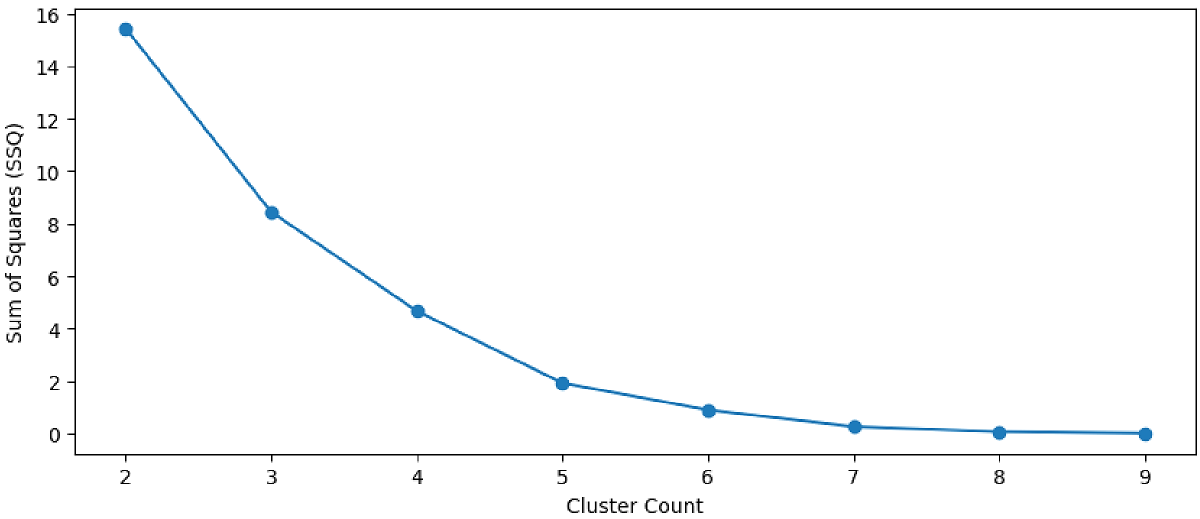

Figure 3. Elbow Method for K-Means

In accordance with Figure 2, the optimal cluster count can be ascertained by identifying the sharpest turn or elbow. Figure 2 suggests that the data may align well with either 3 or 5 clusters. To establish the most suitable cluster count, this study conducts a comparative analysis of data distribution using 3 and 5 clusters. The performance of K-Means clusters between 3 and 5 clusters is depicted in Figure 4.

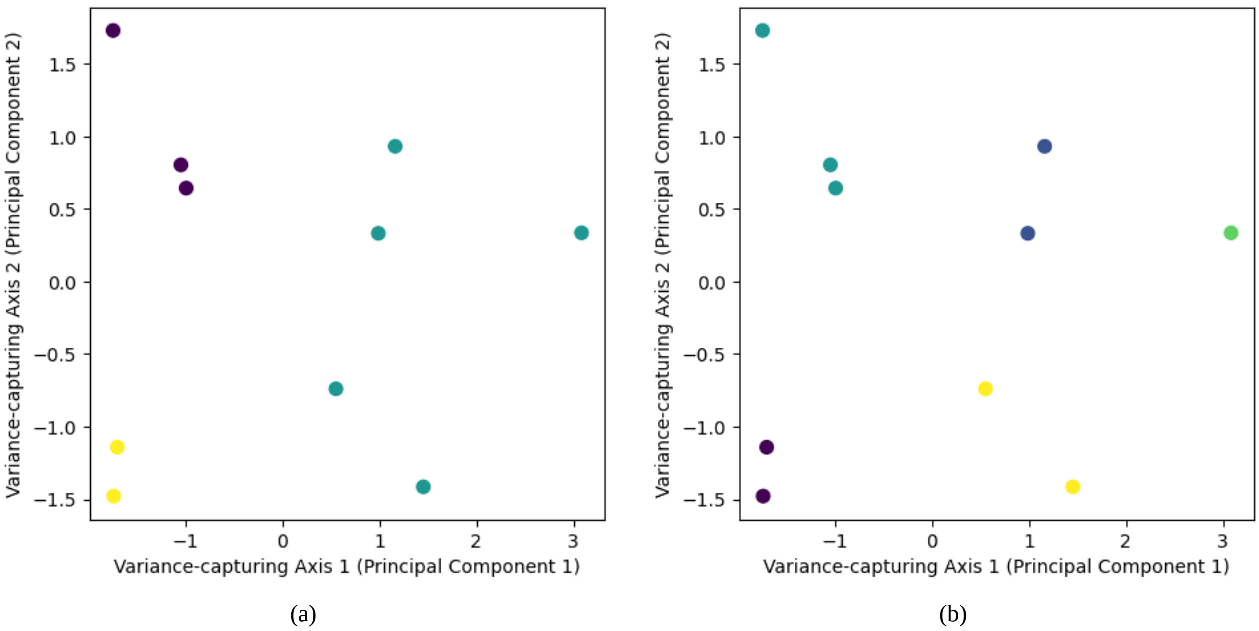

Figure 4. K-Means Clusters Performance; (a) 3 Clusters; (b) 5 Clusters

According to Figure 4, a cluster count of 5 demonstrates a good fit for data, effectively distinguishing data points based on their centroids. Consequently, K-Means will be conducted with 5 clusters. To enhance the visualization of the clustering result with 5 clusters, this study plots the time series data based on their respective clusters. The resulting plot is presented in Figure 5.

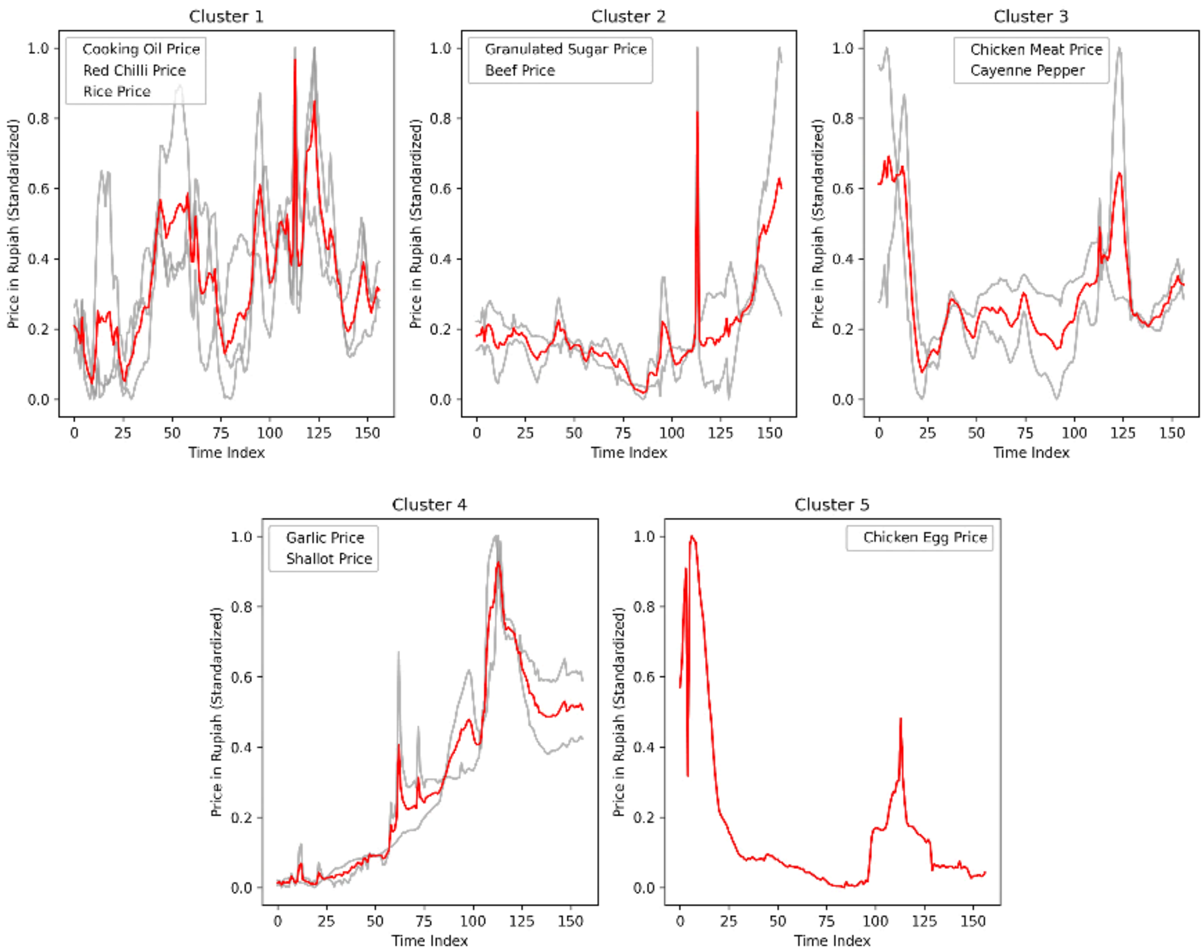

Figure 5. Time Series Cluster with 5 Clusters

According to Figure 5, the K-Means for time series clustering with 5 clusters exhibits a strong performance in distinguishing commodities based on their data patterns. Each commodity price within a cluster shares similar data patterns, with the red line representing the average price of commodities within that cluster. The distribution of the clusters can be visualized in Figure 6 as follows:

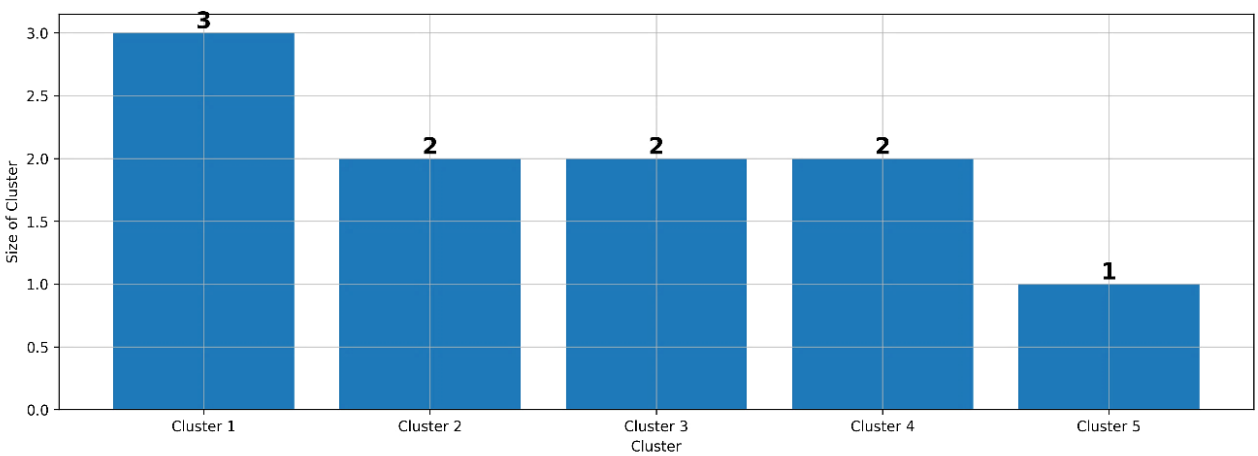

Figure 6. Cluster Distribution

The distribution of clusters, visualized in Figure 6, provides valuable insights into the arrangement of commodities based on similar data patterns. This distribution serves as a precursor to the subsequent analysis, where the resultant clusters are detailed in a table along with the corresponding commodity names. Table 1 offers a comprehensive overview of how commodities are grouped within each cluster, linking the visual representation of clusters to specific results and associated commodities.

Table 1. K-Means Cluster with PCA Mapping

Cluster	1	2	3	4	5
Commodity	Cooking Oil, Red Chili, and Rice	Granulated Sugar and Beef	Chicken Meat, Cayenne Pepper	Garlic and Shallot	Chicken Egg

In this study employing time series clustering analysis, we identified distinct patterns within the temporal evolution of commodity prices. A K-means clustering algorithm, with a reproducible seed for consistency, revealed five discernible clusters. Cluster 1 encapsulates a time series pertaining to Cooking Oil, Red Chili, and Rice, suggesting potential interconnected market dynamics or shared economic influences. Cluster 2 consists of a time series associated with Granulated Sugar and Beef, indicating parallel trends possibly influenced by shared demand and supply factors. Cluster 3 showcases synchronized patterns in Chicken Meat and Cayenne Pepper prices, potentially linked to common agricultural conditions or consumer preferences. Cluster 4 comprises a time series concerning Garlic and Shallot, indicating congruent price dynamics potentially influenced by agricultural or market forces. Lastly, Cluster 5 encompasses a time series related to Chicken Egg prices, implying shared factors impacting the poultry industry or seasonal variations. This clustering approach provides a structured methodology for discerning and interpreting temporal patterns within diverse commodity markets, shedding light on potential interdependencies and shared economic influences.

5. Conclusion

Based on the outcomes of the cluster analysis, commodities grouped in Clusters 1 and 4, specifically onions, garlic, sugar, chicken eggs, and rice, exhibit a notable stability in prices throughout the year. This resilience is attributed to their limited susceptibility to seasonal variations and consistent availability in the market. Conversely, commodities classified in Clusters 2 and 3, including chicken, cooking oil, red chili, cayenne pepper, and beef, demonstrate price fluctuations driven by factors such as seasonality, weather conditions, government policies, and shifts in consumer demand. Through the application of the K-Means method, this study attempts to group the time series data of food commodity prices into clusters that represent certain patterns of increase. The results of the analysis are expected to provide a deeper understanding of the dynamics of the food market in Indonesia. To address these dynamic price trends, it is imperative for the government to institute a robust real-time market monitoring system. Such a system can inform the development of effective price stabilization policies, encompassing interventions like strategic commodity stock management to mitigate the impact of excessively high or low prices. The importance of this research lies in its ability to identify patterns of price increases that may be difficult to understand directly by human observers. The implications of the results of this study may be useful for relevant parties in developing economic policies that are more appropriate and responsive to changes in food commodity prices in Indonesia. A comprehensive strategy involving collaboration among farmers, producers, traders, and consumers is essential for the successful management of commodity price fluctuations.

6. Declarations

6.1. Author Contributions

Conceptualization, M.F.F.M., N.R.A.A.S., S.S., F.F.R.P., H.P., C.I., D.A, S., and D.A.D.; methodology, M.F.F.M., N.R.A.A.S., S.S., F.F.R.P., H.P., C.I., D.A, S., and D.A.D.; writing—original draft preparation, M.F.F.M., N.R.A.A.S., S.S., F.F.R.P., H.P., C.I., D.A, S., and D.A.D.; writing—review and editing, M.F.F.M., N.R.A.A.S, S.S., F.F.R.P., H.P., C.I., D.A., S., and D.A.D. All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

The authors would like to Universitas Airlangga (Unair) for supporting this research under Airlangga Research Fund (ARF) with the number of research 1902/UN3.1.8/PT/2023. This research is supported by Universitas Airlangga Indonesia and INTI International University Malaysia who contributed to the research and preparation of this scientific article. Big appreciation to Pusat Informasi Harga Pangan Strategis (PIHPS) or National Strategic Food Price Information Center by Bank Indonesia (BI) for the data availability.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7. References

[1] World Bank. (2022). Is a Global Recession Imminent? World Bank, Washington, D.C., United States. Available online: https://www.worldbank.org/en/research/brief/global-recession (accessed on May 2024).

[2] BPS-Statistics Indonesia. (2022). The year-on-year (y-on-y) inflation in October 2022 was 5.71 percent. The highest inflation saw in Tanjung Selor at 9.11 percent. Badan Pusat Statistik (BPS - Statistics Indonesia), Jakarta, Indonesia. Available online: https://www.bps.go.id/en/pressrelease/2022/11/01/1866/the-year-on-year--y-on-y--inflation-in-october-2022-was-5-71-percent--the-highest-inflation-saw-in-tanjung-selor-at-9-11-percent-.html (accessed on May 2024).

[3] Punhani, A., Faujdar, N., Mishra, K. K., & Subramanian, M. (2022). Binning-Based Silhouette Approach to Find the Optimal Cluster Using K-Means. IEEE Access, 10, 115025–115032. doi:10.1109/ACCESS.2022.3215568.

[4] Malau, L. R. E., Rambe, K. R., Ulya, N. A., & Purba, A. G. (2023). The impact of climate change on food crop production in Indonesia. Jurnal Penelitian Pertanian Terapan, 23(1), 34–46. doi:10.25181/jppt.v23i1.2418.

[5] Nugraheni, M.Sc., R. D., & Inayah, I. (2022). Impact of the COVID-19 pandemic on world oil and food prices: VECM analysis. Jurnal Ekonomi Dan Pembangunan, 30(1), 15–29. doi:10.14203/jep.30.1.2022.15-29.

[6] Rachmawati, R. A., Saragih, H. J. R., Yanca, I. G. S. K., & Widodo, P. (2022). Analysis of the Threat of Rising Food Prices in Indonesia Due to the Russian-Ukrainian War During the Covid-19 Pandemic. Jurnal Kewarganegaraan, 7(1), 1247.

[7] Ariani, M., & Suryana, A. (2023). Center for Behavioral and Circular Economics Research. National Research and Innovation Agency Jalan Gatot Subroto, 21(1), 1–20. doi:10.21082/akp.v21n1.2023.1-20.

[8] Petitjean, F., Ketterlin, A., & Gançarski, P. (2011). A global averaging method for dynamic time warping, with applications to clustering. Pattern Recognition, 44(3), 678–693. doi:10.1016/j.patcog.2010.09.013.

[9] Maharaj, E. A., D'Urso, P., & Caiado, J. (2019). Time series clustering and classification. Chapman and Hall/CRC, London, United Kingdom. doi:10.1201/9780429058264.

[10] Chen, Y. T., & Witten, D. M. (2022). Selective inference for k-means clustering. Journal of Machine Learning Research (JMLR), 24(152), 1-41.

[11] Ezugwu, A. E., Ikotun, A. M., Oyelade, O. O., Abualigah, L., Agushaka, J. O., Eke, C. I., & Akinyelu, A. A. (2022). A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. Engineering Applications of Artificial Intelligence, 110. doi:10.1016/j.engappai.2022.104743.

[12] Izakian, H., Pedrycz, W., & Jamal, I. (2015). Fuzzy clustering of time series data using dynamic time warping distance. Engineering Applications of Artificial Intelligence, 39, 235–244. doi:10.1016/j.engappai.2014.12.015.

[13] Huang, X., Li, Z., Lu, J., Wang, S., Wei, H., & Chen, B. (2020). Time-series clustering for home dwell time during COVID-19: What can we learn from it? ISPRS International Journal of Geo-Information, 9(11), 675. doi:10.3390/ijgi9110675.

[14] Alaziz, S. N., Alshowiman, A. A., Albayati, B., El-Bagoury, A. al A. H., & Shafik, W. (2023). Clustering of COVID-19 Multi-Time Series-Based K-Means and PCA With Forecasting. International Journal of Data Warehousing and Mining, 19(3), 1–25. doi:10.4018/IJDWM.317374.

[15] Oviedo, E. H. S., Travé-Massuyès, L., Subias, A., Pavlov, M., & Alonso, C. (2023). DTW k-means clustering for fault detection in photovoltaic modules. XI International Congress of Mechanical Engineering, Mechatronics and Automation 2023, 1-3.

[16] Hu, H., Liu, J., Zhang, X., & Fang, M. (2023). An Effective and Adaptable K-means Algorithm for Big Data Cluster Analysis. Pattern Recognition, 139, 109404. doi:10.1016/j.patcog.2023.109404.

[17] Bagirov, A. M., Aliguliyev, R. M., & Sultanova, N. (2023). Finding compact and well-separated clusters: Clustering using silhouette coefficients. Pattern Recognition, 135. doi:10.1016/j.patcog.2022.109144.

[18] Esquivias, M. A., Jayadi, A., Shafiai, S., Abd Rashid, I. M., Borhanordin, A. H., Agusti, K. S., & Yahwidya, L. (2023). The Nexus between Food Security and Investment, Exports, Infrastructure, and Human Capital Development. Journal of Human, Earth, and Future, 4(2), 221–240. doi:10.28991/HEF-2023-04-02-07.

[19] Stabilitas (2023). Tujuh Program Unggulan GNPIP 2023 Jadi Kunci Stabilitas Harga. Stabilitas, Jakarta, Indonesia. Available online: https://www.stabilitas.id/tujuh-program-unggulan-gnpip-2023-jadi-kunci-stabilitas-harga/ (accessed on June 2024).

[20] Jain, A. K. (2010). Data clustering: 50 years beyond K-means. Pattern Recognition Letters, 31(8), 651–666. doi:10.1016/j.patrec.2009.09.011.

[21] Wang, G., Xing, C., Wang, J. S., Wang, H. Y., & Liu, J. X. (2022). Clustering Validity Evaluation Method Based on Two Typical Clustering Algorithms. IAENG International Journal of Computer Science, 49(3), 871.

[22] Ashari, I. F., Dwi Nugroho, E., Baraku, R., Novri Yanda, I., & Liwardana, R. (2023). Analysis of Elbow, Silhouette, Davies-Bouldin, Calinski-Harabasz, and Rand-Index Evaluation on K-Means Algorithm for Classifying Flood-Affected Areas in Jakarta. Journal of Applied Informatics and Computing, 7(1), 89–97. doi:10.30871/jaic.v7i1.4947.

[23] Lenssen, L., & Schubert, E. (2024). Medoid Silhouette clustering with automatic cluster number selection. Information Systems, 120, 102290. doi:10.1016/j.is.2023.102290.