

Clustering Pasar Tradisional di Wilayah Kabupaten Tangerang Berdasarkan Harga Bahan Pokok dengan Metode *K-means*

Syahid Abdullah ^{a,1,*}, Bangkit Dwi Nuryanto ^{b,2}, Riad Sahara ^{c,3}, M. Ikhwani Saputra ^{d,4}, Cian Romadhona Hassolthine ^{e,5}

^{a,b,c,d,e} Program Studi PJJ Informatika Universitas Siber Asia, Kampus Menara, Jl. RM. Harsono, Ragunan, Jakarta Selatan, DKI Jakarta 12550

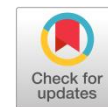
¹syahidabdullah@lecturer.unsia.ac.id; ²dwi.bank90@gmail.com; ³riadsahara@lecturer.unsia.ac.id; ⁴muhammadikhwani@lecturer.unsia.ac.id;

⁵cianhassolthine@lecturer.unsia.ac.id

* Penulis Korespondensi

ABSTRAK

Clustering adalah metode yang digunakan untuk mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya. Salah satu metode *clustering* yang populer digunakan adalah *K-means*. Dalam penelitian ini, *K-means* akan digunakan untuk *clustering* pasar tradisional di Wilayah Kabupaten Tangerang, Banten berdasarkan harga bahan pokok. Dengan melakukan analisis *clustering* menggunakan metode *K-means*, dapat diidentifikasi pola harga bahan pokok yang mungkin tidak terlihat secara langsung. Dari percobaan yang sudah dilakukan, terlihat bahwa harga bahan pokok di 19 pasar tradisional di Kabupaten Tangerang memang cenderung tidak variatif. Jumlah *cluster* yang dihasilkan adalah 2 *cluster*, di mana harga bahan pokok pada pasar tradisional yang ada pada *cluster* pertama cenderung lebih tinggi dibandingkan pada *cluster* kedua. Ada 3 pasar yang ada pada *cluster* pertama, yaitu Pasar Balaraja Sentiong, Karawaci, dan Kelapa Dua. Hal tersebut diduga terjadi karena letak pasar yang strategis dan jumlah penduduk yang padat sehingga membutuhkan pasokan yang lebih banyak dan daya beli masyarakat.

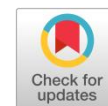


KATA KUNCI

Clustering
Harga Bahan Pokok
K-means
Kabupaten Tangerang
Pasar Tradisional

ABSTRACT

Clustering is a method used to group data that has characteristic similarities between it and the others. One of the most popular clustering methods used is *K-means*. In this study, *K-means* will be used to cluster the traditional market in the Kabupaten Tangerang, Banten based on the prices of staple. By performing clustering analysis using the *K-means* method, it is possible to identify the price patterns of commodities that may not be visible directly. From the experiments already carried out, it is seen that the price of staples in the 19 traditional markets in Tangerang district tends to be unvariable. The number of clusters produced is 2, where the value of staples on the traditional market that exists on the first cluster tends to be higher than on the second cluster. There are 3 markets in the first cluster, namely Balaraja Sentiong, Karawaci, and Kelapa Dua Market. This is believed to be due to the strategic market placement and the density of the population, which require more supply and purchasing power from the community.



KEYWORD

Clustering
K-means Method
Kabupaten Tangerang
Staple Price
Traditional Market



This is an open-access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license

1. Pendahuluan

Clustering adalah metode yang digunakan untuk mengelompokkan data yang mempunyai kemiripan karakteristik antara data satu dengan data lainnya yang telah diperoleh [1]. Ciri khasnya adalah mempunyai sifat tanpa arahan (*unsupervised*), maksudnya adalah teknik ini diterapkan tanpa perlunya data latih dan tidak memerlukan target output. Salah satu metode *clustering* yang populer digunakan adalah *K-means*. Konsep utama dari metode *K-means* adalah untuk mengelompokkan sejumlah data ke dalam kelompok-kelompok yang saling berhubungan berdasarkan kesamaan atribut tertentu [2].

Dibandingkan dengan metode *clustering* lainnya, *K-means* memiliki beberapa kelebihan. Di antaranya adalah metode ini relatif sederhana dan mudah diimplementasikan, serta efisien jika digunakan untuk data

dengan jumlah besar. *K-means* sendiri sudah banyak digunakan untuk berbagai penelitian, mulai dari pengelompokan penjualan barang [3], pengelompokan data rekam medis pasien [4], hingga pengelompokan kasus Covid-19 [5]. Secara umum, penelitian-penelitian tersebut dilakukan untuk mengetahui pola dari kasus yang sedang dihadapi sehingga bisa memunculkan solusi untuk mengatasi permasalahan dalam kasus tersebut. Adapun dalam penelitian ini, *K-means* akan digunakan untuk *clustering* pasar tradisional di Wilayah Kabupaten Tangerang, Banten berdasarkan harga bahan pokok..

Harga bahan pokok memiliki dampak yang signifikan terhadap daya beli konsumen dan kesejahteraan masyarakat. Pasar tradisional masih menjadi pilihan banyak warga di Indonesia meskipun banyak muncul swalayan/mini market alasannya sederhana yaitu harga barang. Harga barang yang sudah masuk swalayan sudah pasti akan lebih mahal karena ada pajak dari setiap penjualan produk. Apalagi di masa pandemi yang baru saja terjadi, yang banyak berdampak pada semua lini ekonomi membuat harga kebutuhan semakin terasa mahal. Kabupaten Tangerang, sebagai salah satu daerah padat penduduk di Indonesia, memiliki sejumlah pasar tradisional yang menyediakan berbagai macam bahan pokok untuk memenuhi kebutuhan sehari-hari. Namun, fluktuasi harga bahan pokok sering kali menjadi perhatian utama bagi pemerintah, pelaku bisnis, dan masyarakat umum. Fluktuasi harga yang tidak terkendali dapat menyebabkan ketidakstabilan ekonomi dan berdampak negatif pada kesejahteraan masyarakat. Oleh karena itu, *clustering* menggunakan metode *K-means* dalam penelitian ini menjadi sangat penting yang bertujuan untuk menganalisis pola harga bahan pokok.

Dalam penelitian ini kami mengambil sample data harga beberapa kebutuhan pokok dipasar wilayah Kabupaten Tangerang. Setidaknya ada 19 pasar tradisional dan ada 5 bahan kebutuhan pokok yang akan diteliti dengan metode *clustering*. Dengan melakukan analisis *clustering* menggunakan metode *K-means*, dapat diidentifikasi pola harga bahan pokok yang mungkin tidak terlihat secara langsung. Kelompok pasar tradisional yang terbentuk dapat memberikan wawasan tentang keragaman harga di pasar wilayah Kabupaten Tangerang. Informasi ini akan bermanfaat bagi pemerintah dalam merumuskan kebijakan harga, pedagang dalam mengoptimalkan pasokan dan penawaran, serta konsumen dalam mempersiapkan anggaran dan memilih tempat pembelian yang sesuai.

2. Tinjauan Pustaka

2.1. Clustering

Clustering adalah metode yang digunakan untuk membagi rangkaian data menjadi beberapa kelompok berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya [2]. *Cluster* adalah sekelompok atau sekumpulan objek-objek data yang similar satu sama lain dalam *cluster* yang sama dan dissimilar terhadap objek-objek yang berbeda *cluster* [6]. Objek akan dikelompokkan ke dalam satu atau lebih *cluster* sehingga objek-objek yang berada dalam satu *cluster* akan mempunyai kesamaan yang tinggi antara satu dengan yang lainnya.

2.2. Algoritma K-means

Metode *K-means* adalah salah satu metode yang paling populer dalam analisis data dan pengelompokan (*clustering*). Tujuan utama dari metode *K-means* adalah untuk mengelompokkan sejumlah data ke dalam kelompok-kelompok yang saling berhubungan berdasarkan kesamaan atribut tertentu [2]. Proses *K-means* dimulai dengan menentukan jumlah kelompok atau centroid yang akan dibentuk. Centroid adalah titik pusat dari setiap kelompok dan digunakan untuk mewakili kelompok tersebut. Selanjutnya, *K-means* memperbarui posisi centroid secara iteratif hingga konvergensi. Iterasi dilakukan dengan meminimalkan jumlah kesalahan kuadrat antara setiap data dengan centroid kelompok yang sesuai.

3. Metodologi Penelitian

Penelitian ini dimulai dengan pengumpulan data. Data yang digunakan dalam penelitian ini diambil dari data Dinas Perdagangan dan Perindustrian Kabupaten Tangerang mengenai harga kebutuhan bahan pokok di pasar tradisional. Setelah data diperoleh, dilakukan *clustering* menggunakan algoritme *K-means*. Adapun proses dasar algoritma *k-means*:

1. Menentukan jumlah klaster yang ingin dibentuk dan tentukan pusat kluster. Pada tahapan ini, digunakan metode Silhouette untuk memperoleh jumlah *cluster* yang paling optimum [1].

- Penelitian ini menggunakan jarak euclidean untuk menghitung jarak setiap data ke pusat *cluster*.

$$d(i, k) = \sqrt{\sum_i^m (Cij + Ckj)^2}$$

- Kelompokkan data ke dalam *cluster* dengan jarak yang paling pendek dengan persamaan

$$\min \sum_k^i - aik - = \sqrt{\sum_i^m (Cij + Ckj)^2}$$

- Hitung pusat kluster yang baru dengan persamaan

$$Ckj = \sum_k^i \frac{Xij}{p}$$

Dengan : $X_{ij} \in$ Kluster ke k p = banyaknya anggota kluster ke - k

- Mengulangi langkah ke-2 sampai dengan ke-4 sehingga sudah tidak ada lagi data yang berpindah ke kluster yang lain.

4. Hasil dan Pembahasan

4.1. Hasil

Data diambil dari data Dinas Perdagangan dan Perindustrian Kabupaten Tangerang mengenai harga kebutuhan bahan pokok di pasar tradisional. Berikut tabel harga kebutuhan bahan pokok di pasar tradisional di daerah Kabupaten Tangerang per tanggal 10 Juli 2022:

Table 1. Harga bahan pokok di pasar tradisional Kabupaten Tangerang per tanggal 10 Juli 2022

Pasar	Harga Bahan Pokok (dalam rupiah)				
	Beras /kg	Minyak goreng /liter	Telur /kg	Gula pasir /kg	Bawang merah /kg
Pasar Kutabumi	12500	21000	29000	14000	63500
Pasar Sepatan	12600	21300	29100	14300	63550
Pasar Mauk	12500	21100	29050	14500	63500
Pasar Kemiri	12450	21400	29100	14350	63400
Pasar Kronjo	12400	21450	29500	13900	63600
Pasar Cituis	13000	21300	29050	14300	63400
Pasar Kampung Melayu	12700	21000	29000	14250	63500
Pasar Balaraja Sentiong	12500	22000	29100	14500	63550
Pasar Segitiga Balaraja	12600	21500	29200	14050	63500
Pasar Tigaraksa	12450	21350	29100	14200	63800
Pasar Cisoka	12500	21500	28950	14300	63500
Pasar Kemis	12600	21450	29150	14300	63450
Pasar Bayam	12450	21500	28900	14500	63500
Pasar Karawaci	12800	21700	29500	14600	63800
Pasar Kelapa Dua	12850	21650	29450	14550	63700
Pasar Bojong Nangka	12500	21500	29300	14300	63600
Pasar Curug	12400	21450	29100	14300	63450
Pasar Korelet	12500	21500	29050	14550	63550
Pasar Pelangi Rajeg	12450	21650	29000	14500	63550

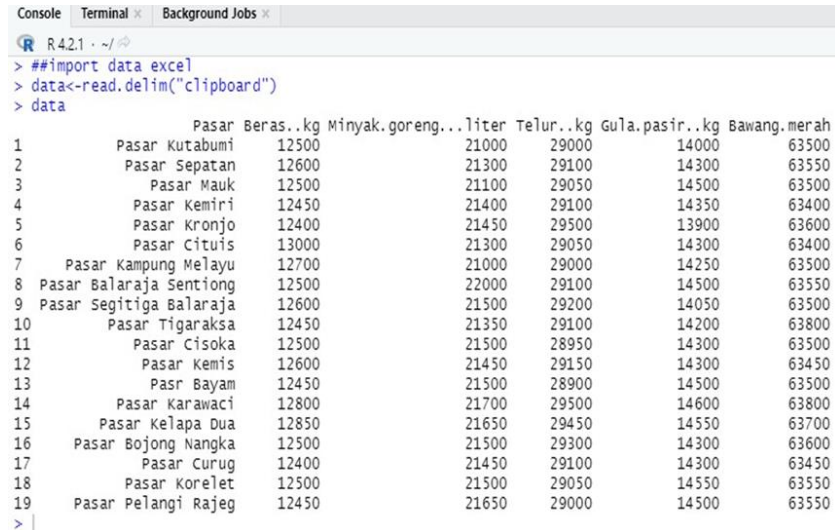
Proses *clustering* menggunakan bantuan tool RStudio yang memperoleh hasil sebagai berikut:

1. Memasukan/import data ke dalam RStudio.

```
1 ##import data excel
2 data<-read.delim("clipboard")
3 data
4
```

Gambar 1. Sourcecode untuk import data

Hasilnya:

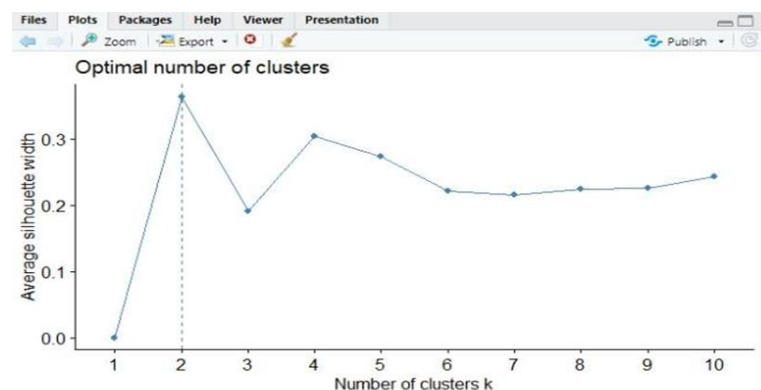


```
> ##import data excel
> data<-read.delim("clipboard")
> data
```

	Pasar	Beras..kg	Minyak.goreng...liter	Telur..kg	Gula.pasir..kg	Bawang.merah
1	Pasar Kutabumi	12500	21000	29000	14000	63500
2	Pasar Sepatan	12600	21300	29100	14300	63550
3	Pasar Mauk	12500	21100	29050	14500	63500
4	Pasar Kemiri	12450	21400	29100	14350	63400
5	Pasar Kronjo	12400	21450	29500	13900	63600
6	Pasar Cituis	13000	21300	29050	14300	63400
7	Pasar Kampung Melayu	12700	21000	29000	14250	63500
8	Pasar Balaraja Sentiong	12500	22000	29100	14500	63550
9	Pasar Segitiga Balaraja	12600	21500	29200	14050	63500
10	Pasar Tigaraksa	12450	21350	29100	14200	63800
11	Pasar Cisoka	12500	21500	28950	14300	63500
12	Pasar Kemis	12600	21450	29150	14300	63450
13	Pasar Bayam	12450	21500	28900	14500	63500
14	Pasar Karawaci	12800	21700	29500	14600	63800
15	Pasar Kelapa Dua	12850	21650	29450	14550	63700
16	Pasar Bojong Nangka	12500	21500	29300	14300	63600
17	Pasar Curug	12400	21450	29100	14300	63450
18	Pasar Korelet	12500	21500	29050	14550	63550
19	Pasar Pelangi Rajeg	12450	21650	29000	14500	63550

Gambar 2. Tampilan data yang sudah di-import

2. Memanggil *library* yang akan digunakan dalam proses *clustering* dan memindahkan kolom pasar bergeser ke kolom pertama.
 - *library cluster*: *library* yang digunakan untuk melakukan analisis *cluster*
 - *library factoextra*: *library* yang digunakan untuk mengekstrak dan memvisualisasikan hasil dari analisis multivariat.
 - *library tidyverse*: *library* yang digunakan untuk melakukan pengolahan data.
3. Menentukan jumlah *cluster* yang akan dibuat menggunakan metode *silhouette*. Hasilnya *cluster* optimal yang dapat dibuat adalah 2 *cluster*.



Gambar 3. Jumlah *cluster* optimal yang diperoleh adalah 2 *cluster*

4. Proses *clustering* menggunakan *K-means*. Hasilnya:

```
> final=kmeans(data1,2)
> final
K-means clustering with 2 clusters of sizes 3, 16

Cluster means:
Beras..kg Minyak.goreng...liter Telur..kg Gula.pasir..kg Bawang.merah
1 12716.67 21783.33 29350.00 14550.0 63683.33
2 12537.50 21371.88 29096.88 14287.5 63521.88

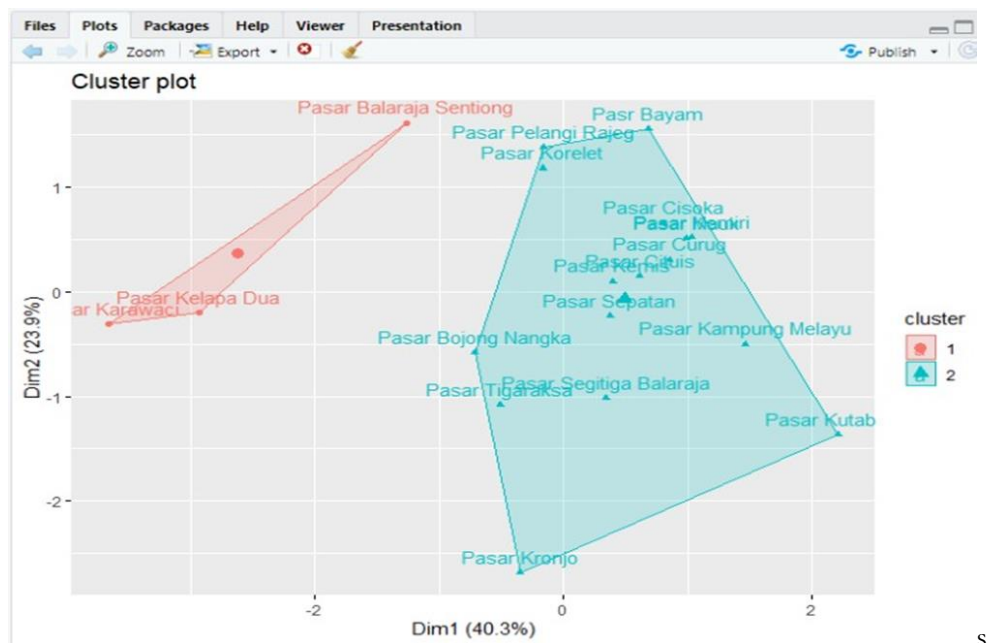
Clustering vector:
Pasar Kutabumi 2 Pasar Sepatan 2 Pasar Mauk 2 Pasar Kemiri 2
Pasar Kronjo 2 Pasar Cituis 2 Pasar Kampung Melayu 2 Pasar Balaraja Sentiong 1
Pasar Segitiga Balaraja 2 Pasar Tigaraksa 2 Pasar Cisoka 2 Pasar Kemis 2
Pasr Bayam 2 Pasar Karawaci 1 Pasar Kelapa Dua 1 Pasar Bojong Nangka 2
Pasar Curug 2 Pasar Korelet 2 Pasar Pelangi Rajeg 2

within cluster sum of squares by cluster:
[1] 275000 1822031
(between_SS / total_SS = 30.3 %)

Available components:
[1] "cluster" "centers" "totss" "withinss" "tot.withinss" "betweenss"
[7] "size" "iter" "ifault"
```

Gambar 4. Hasil *clustering* menggunakan *K-means*

5. Visualisasi hasil *clustering*. Hasilnya:



Gambar 5. Hasil visualisasi menggunakan *K-means*

4.2. Pembahasan

Dari percobaan yang sudah dilakukan, terlihat bahwa harga bahan pokok di 19 pasar tradisional di Kabupaten Tangerang memang cenderung tidak variatif. Jumlah *cluster* yang dihasilkan adalah 2 *cluster*, di mana 3 pasar tradisional, yaitu Pasar Balaraja Sentiong, Karawaci, dan Kelapa Dua berada pada *cluster* pertama, sedangkan sisanya berada pada *cluster* kedua. Harga bahan pokok pada *cluster* pertama cenderung lebih tinggi daripada pasar yang berada di *cluster* kedua.

5. Penutup

5.1. Kesimpulan

Dari hasil penelitian di atas, dapat disimpulkan bahwa sebaran harga kebutuhan pokok di pasar tradisional wilayah Kabupaten Tangerang tergolong merata, bisa dilihat dari *cluster* yang dihasilkan dalam proses *clustering*. Ada 3 pasar yang memiliki harga kebutuhan pokok tertinggi, yaitu Pasar Balaraja Sentiong, Karawaci, dan Kelapa Dua. Hal tersebut kemungkinan karena letak pasar yang strategis, jumlah penduduk yang padat sehingga membutuhkan pasokan yang lebih banyak dan daya beli masyarakat.

5.2. Saran

Hasil penelitian ini dapat menjadi dasar untuk mengembangkan kebijakan yang lebih baik dalam mengatur harga bahan pokok, meningkatkan ketersediaan pasokan, dan memberikan manfaat bagi konsumen dan pedagang di pasar tradisional.

Daftar Pustaka

- [1] D. A. I. C. Dewi dan D. A. K. Pramita, "Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma *Clustering* K-Medoids dalam Pengelompokan Produksi Kerajinan Bali". *Matrix : Jurnal Manajemen Teknologi Dan Informatika*, 9(3), 102-109, 2019. doi:10.31940/matrix.v9i3.1662.
- [2] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction* (2nd ed.). Springer.
- [3] S. Aulia. "Klasterisasi Pola Penjualan Pestisida Menggunakan Metode *K-means Clustering* (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja)." *Djtechno: Jurnal Teknologi Informasi* 1.1 (2020): 1-5.
- [4] R. Ordila, *et al.* "Penerapan Data Mining Untuk Pengelompokan Data Rekam Medis Pasien Berdasarkan Jenis Penyakit Dengan Algoritma *Clustering* (Studi Kasus: Poli Klinik Pt. Inecda)." *Jurnal Ilmu Komputer* 9.2 (2020): 148-153.
- [5] Nabila, Zulfa, *et al.* "Analisis Data Mining Untuk *Clustering* Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma *K-means*." *Jurnal Teknologi Dan Sistem Informasi* 2.2 (2021): 100-108.
- [6] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.