

Pemanfaatan *K-Means Clustering* untuk Optimalisasi Penjualan Produk Roti Berdasarkan Data Penjualan Harian

Irwan¹, Adi Panca Pamungkas², Wiwin Handoko³

^{1,2}Mahasiswa Program Studi Sistem Informasi, Universitas Royal

³Dosen Program Studi Sistem Informasi, Universitas Royal

¹ irwananakdesa@gmail.com; ² or44203@gmail.com; ³ win.van.handoko@gmail.com

* Email Koresponden

INFO ARTIKEL

ABSTRAK

Penjualan produk roti menjadi salah satu aspek penting dalam bisnis roti yang dipengaruhi oleh fluktuasi permintaan yang tidak dapat diprediksi dengan mudah. Pengelolaan penjualan yang efisien memerlukan pemahaman yang mendalam mengenai pola penjualan produk. Penelitian ini bertujuan untuk mengoptimalkan penjualan produk roti dengan menggunakan algoritma *K-Means Clustering* untuk menganalisis kinerja penjualan harian di Toko Roti Amin. Data yang digunakan mencakup jumlah penjualan dan frekuensi transaksi produk roti sebanyak 356 data selama 2 minggu dalam bulan desember 2024. Hasil pengujian menunjukkan bahwa produk roti dapat dikelompokkan ke dalam tiga *cluster*, yaitu 129 data dalam “Penjualan Bagus”, 28 data pada *cluster* “Penjualan Sedang”, dan 199 data *cluster* “Penjualan Rendah”. Temuan ini membantu pengusaha roti dalam pengelolaan stok, perencanaan produksi, dan strategi pemasaran yang lebih tepat sasaran. Meskipun terdapat beberapa keterbatasan dalam penggunaan *K-Means Clustering*, seperti ketergantungan pada pemilihan *centroid* awal, penelitian ini membuktikan bahwa penerapan teknik ini dapat meningkatkan pengelolaan inventaris dan memaksimalkan keuntungan dalam bisnis roti.

Kata Kunci: Penjualan, Produk Roti, *K-Means Clustering*, *Cluster*.



This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.
Copyright ©2024 by Author. Published by
PT Beranda Teknologi Academia

ABSTRACT

Bread product sales have become an important aspect of the bakery business, influenced by fluctuations in demand that are not easily predictable. Efficient sales management requires a deep understanding of sales patterns. This study aims to optimize bread product sales by using the K-Means Clustering algorithm to analyze daily sales performance at Toko Roti Amin. The data used includes sales volume and transaction frequency for bread products, consisting of 356 data points. The results show that the bread products can be grouped into three clusters: 129 data in the “Good Sales” cluster, 28 data in the “Moderate Sales” cluster, and 199 data in the “Low Sales” cluster. These findings assist bakery owners in managing stock, production planning, and more targeted marketing strategies. Although there are limitations in using K-Means Clustering, such as dependence on the initial centroid selection, this study proves that applying this technique can enhance inventory management and maximize profit in the bakery business.

Keywords: Sales, Bread Products, K-Means Clustering, Clusters.

PENDAHULUAN

Penjualan produk adalah aspek penting dalam dunia bisnis, terutama dalam industri makanan yang sangat bergantung pada tren permintaan konsumen yang sering berubah. Penjualan produk adalah kegiatan yang sangat krusial dalam dunia bisnis, karena merupakan indikator utama keberhasilan suatu usaha. Setiap produk yang dipasarkan memiliki karakteristiknya sendiri, dan pemahaman terhadap pola penjualan produk menjadi sangat penting untuk menentukan strategi bisnis yang tepat. Terdapat banyak sekali macam produk makanan yang saat ini dijual dan diedarkan oleh para pebisnis, salah satunya adalah produk roti.

Produk roti merupakan salah satu jenis makanan yang banyak diminati oleh konsumen, dengan berbagai varian rasa dan jenis yang disesuaikan dengan preferensi pasar. Dalam bisnis roti, keberagaman produk yang ditawarkan dan fluktuasi permintaan yang terjadi setiap hari menjadi tantangan tersendiri. Penjualannya seringkali dipengaruhi oleh faktor musiman, waktu tertentu dalam sehari, atau bahkan tren konsumsi yang dapat berubah seiring waktu. Oleh karena itu, pengelolaan penjualan roti memerlukan pemahaman yang mendalam mengenai produk mana yang memiliki kinerja terbaik, serta produk yang kurang diminati oleh konsumen. Dengan informasi ini, pengusaha roti dapat melakukan perencanaan produksi dan pemasaran yang lebih efektif, sehingga dapat memaksimalkan keuntungan dan mengurangi pemborosan.

Dalam usaha untuk mengoptimalkan penjualan produk roti, penting untuk mengidentifikasi pola-pola penjualan yang mungkin tersembunyi dalam data [1]. Namun, proses analisis manual terhadap penjualan harian seringkali memakan waktu dan tidak efektif dalam mengungkapkan informasi yang relevan [2]. Oleh karena itu, penggunaan metode analisis berbasis data, seperti *K-Means Clustering*, dapat menjadi solusi yang tepat. Algoritma *K-Means Clustering* dapat mengelompokkan produk roti berdasarkan kinerja penjualannya, dengan tujuan untuk mengetahui produk mana yang lebih laris, yang memerlukan stok lebih banyak, atau bahkan yang perlu dikurangi atau dihentikan produksinya. Dengan memanfaatkan teknik ini, pengusaha roti dapat membuat keputusan yang lebih informasional dan strategis, sehingga dapat mengoptimalkan inventaris, meminimalkan kerugian, dan meningkatkan kepuasan konsumen.

Namun, permasalahannya adalah meskipun sudah ada pengusaha roti yang sudah memahami pentingnya pemantauan penjualan dan pengelolaan produk, masih banyak yang masih kesulitan dalam mengidentifikasi pola penjualan yang tepat dan membuat keputusan berbasis data, salah satunya adalah pengusaha Toko Roti Amin. Pihak Toko Roti Amin masih bergantung pada metode analisis tradisional yang mengandalkan pengamatan manual atau perkiraan dapat menyebabkan ketidaktepatan dalam perencanaan produksi dan distribusi, yang berujung pada pemborosan bahan baku atau kekurangan stok produk yang diminati. Selain itu, keberagaman produk roti dan fluktuasi permintaan yang tidak dapat diprediksi dengan mudah membuat sulit bagi pengusaha untuk mengelola inventaris dengan efisien. Oleh karena itu, dibutuhkan pendekatan yang lebih canggih untuk menggali pola tersembunyi dalam data penjualan harian, sehingga pengusaha dapat mengambil keputusan yang lebih tepat dan berbasis data.

Pendekatan yang bisa dilakukan salah satunya adalah penerapan *clustering*. *Clustering* adalah proses mengelompokkan data ke dalam klaster atau kategori berdasarkan tingkat kesamaan antara satu data dengan data lainnya [3][4]. Salah satu teknik dalam *clustering* adalah algoritma *K-Means Clustering*. *K-Means Clustering* adalah salah satu metode yang terdapat pada cabang ilmu *data mining* yang digunakan untuk mengelompokkan data berdasarkan kemiripan atau kedekatan fitur-fitur tertentu. Algoritma *K-Means Clustering* banyak diterapkan dalam segmentasi pelanggan, pengelompokan dokumen, analisis gambar, dan berbagai bidang lainnya karena keunggulannya dalam kecepatan dan skalabilitas, meskipun kinerjanya dapat dipengaruhi oleh pemilihan awal *centroid* [5][6][7]. Dalam konteks penjualan produk roti, *K-Means Clustering* dapat diterapkan untuk mengelompokkan produk berdasarkan kinerja penjualannya, seperti frekuensi terjual dan volume penjualan. Dengan mengelompokkan produk ke dalam beberapa kategori, pengusaha roti dapat lebih

mudah mengidentifikasi produk mana yang termasuk dalam kategori produk unggulan, produk yang harus lebih dipromosikan, atau produk yang kurang diminati dan perlu pengurangan stok.

Proses *K-Means Clustering* dimulai dengan menentukan jumlah kelompok yang diinginkan (K) [8], kemudian data akan dibagi ke dalam kelompok-kelompok tersebut berdasarkan kedekatan antar data. Setiap kelompok akan memiliki karakteristik atau pola penjualan yang lebih homogen, yang memudahkan pengusaha dalam mengambil keputusan terkait stok produk dan strategi pemasaran. Dengan demikian, penerapan algoritma *K-Means Clustering* dapat menjadi solusi yang efektif dalam mengoptimalkan penjualan produk [9], serta memberikan wawasan yang lebih tajam untuk perencanaan bisnis yang lebih baik.

Penelitian ini menggunakan kriteria seperti jumlah terjual dan frekuensi penjualan. Jumlah terjual adalah banyaknya produk yang terjual, dan frekuensi penjualan adalah jumlah transaksi yang dilakukan. Hasil pengujian akan memperlihatkan produk mana yang lebih laku atau banyak terjual dan yang tidak. Hasil *clustering* akan berjumlah 3 *cluster*, dengan label “Penjualan Bagus”, “Penjualan Sedang”, dan “Penjualan Rendah”. Tujuannya untuk menganalisis hasil penjualan untuk masing-masing produk dengan transaksi per-hari.

METODE

Penelitian yang digunakan menggunakan pendekatan dalam *data mining* dan *machine learning*. Pendekatan *data mining* dan *machine learning* memungkinkan pembuatan prediksi dengan memanfaatkan data historis sebagai dasar analisis [10]. *Data mining* adalah proses mengekstraksi atau menyaring informasi dari kumpulan data berukuran besar melalui berbagai tahapan untuk memperoleh wawasan yang bernilai dari data tersebut [11]. Sedangkan *machine learning* adalah teknik pemrograman komputer yang memanfaatkan data historis untuk melatih model, dengan tujuan meningkatkan kinerja dalam mengekstraksi informasi yang berguna dari suatu kumpulan data [12]. Algoritma dalam *data mining* dan *machine learning* yang digunakan adalah *K-Means Clustering*.

K-Means Clustering adalah metode pengelompokan data yang tidak bersifat hierarkis, yang mengelompokkan data menjadi satu atau lebih *cluster*. Algoritma K-means merupakan teknik pengelompokan berbasis partisi [13]. Berikut langkah-langkah dalam algoritma K-Means.

1. Menyusun data sesuai dengan kebutuhan sebagai sampel;
2. Menentukan jumlah kluster yang diinginkan;
3. Menetapkan nilai *centroid* atau titik pusat setiap kluster. Untuk mengukur jarak antara setiap data terhadap *centroid* yang memiliki kategori kondisi paling dekat atau serupa, digunakan model Euclidean dengan rumus sebagai berikut [14]:

$$d_{ij} = \sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + \dots + (x_{ki} - x_{kj})^2} \quad (1)$$

Keterangan :

d_{ij} = Jarak antara data ke-i dan titik pusat (*cluster j*).

x_{ki} = Nilai atribut ke-k dari data ke-i.

x_{kj} = Nilai atribut ke-k dari data ke-j.

Pada dasarnya, walaupun K-Means Clustering mampu mengelompokkan data ke dalam *cluster-cluster*, terdapat beberapa kelemahan yang dimiliki algoritma ini, yaitu sebagai berikut:

1. Jika jumlah data relatif kecil, penentuan *cluster* awal dapat dilakukan dengan mudah.
2. Jumlah *cluster*, yaitu K, harus ditentukan terlebih dahulu sebelum proses perhitungan dimulai.
3. Tidak mungkin mengetahui *cluster* yang sebenarnya hanya dengan menggunakan data yang sama, namun jika data dimasukkan dengan cara berbeda, hasil *cluster* yang berbeda mungkin akan muncul, terutama jika jumlah datanya sedikit.
4. Kontribusi setiap atribut dalam proses pengelompokan tidak dapat diketahui karena diasumsikan bahwa setiap atribut memiliki bobot yang setara [15].

HASIL DAN PEMBAHASAN

Tahapan awal pada proses ini adalah memanggil dataset yang sudah dimasukkan ke dalam *jupyter notebook*. Data yang digunakan berupa data penjualan perhari di Toko Roti Amin selama 2 minggu terakhir bulan desember 2024.

[58]:

	Tanggal Transaksi	Nama Produk	Jumlah Terjual	Frekuensi Penjualan	Harga Produk (Rp)	Pendapatan Penjualan (Rp)
0	2024-12-16	Banana Strudel	20	9	30000	600000
1	2024-12-16	Bolu Gulung Mocca	3	3	35000	105000
2	2024-12-16	Bolu Gulung Pandan	5	6	35000	175000
3	2024-12-16	Bolu Mentega	9	4	25000	225000
4	2024-12-16	Bolu Mentega Hias	8	2	27000	216000
...	...	...	...	...	...	...
351	2024-12-31	Roti Sobek Srikaya	8	8	12000	96000
352	2024-12-31	Roti Sosis	18	4	3000	54000
353	2024-12-31	Roti Tawar	17	12	17000	289000
354	2024-12-31	Selai Srikaya	17	12	11000	187000
355	2024-12-31	Sobek Keju	16	6	12000	192000

356 rows × 6 columns

Gambar 1. Data

Setelah itu, dilakukan pengujian *cluster*, dengan tampilan *cluster* dan nilai *centroid* yang juga memuat persebaran data sebagai berikut.

[65]:

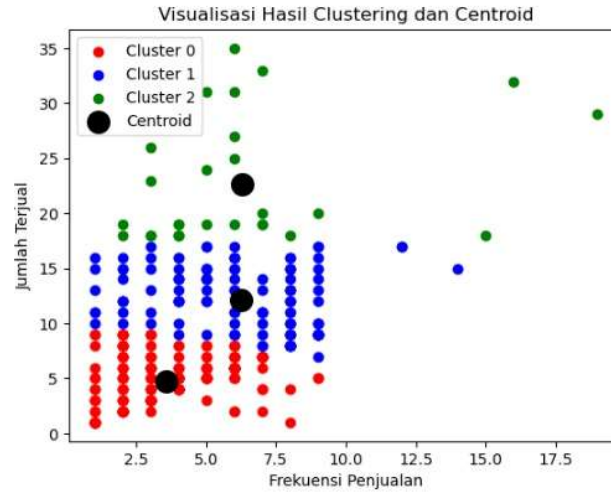
```
#membuat dataframe hasil kluster agar bisa dilihat pada dataset
df['Cluster'] = y_prediksi
df
```

[65]:

	Tanggal Transaksi	Nama Produk	Jumlah Terjual	Frekuensi Penjualan	Harga Produk (Rp)	Pendapatan Penjualan (Rp)	Cluster
0	2024-12-16	Banana Strudel	20	9	30000	600000	2
1	2024-12-16	Bolu Gulung Mocca	3	3	35000	105000	0
2	2024-12-16	Bolu Gulung Pandan	5	6	35000	175000	0
3	2024-12-16	Bolu Mentega	9	4	25000	225000	1
4	2024-12-16	Bolu Mentega Hias	8	2	27000	216000	0
...	...	...	...	...	...	...	...
351	2024-12-31	Roti Sobek Srikaya	8	8	12000	96000	1
352	2024-12-31	Roti Sosis	18	4	3000	54000	2
353	2024-12-31	Roti Tawar	17	12	17000	289000	1
354	2024-12-31	Selai Srikaya	17	12	11000	187000	1
355	2024-12-31	Sobek Keju	16	6	12000	192000	1

356 rows × 7 columns

Gambar 2. Tampilan *cluster*



Gambar 3. Persebaran data dan *centroid*

Dapat dilihat pada gambar 2, bahwa nama untuk tiap *cluster* masih dilabeli dengan angka 0, 1, dan 2. Maka dari itu nama tersebut perlu diubah agar mudah untuk dibaca, menjadi *cluster* “Penjualan Bagus”, “Penjualan Sedang”, dan “Penjualan Rendah”.

[69]:	Tanggal Transaksi	Nama Produk	Jumlah Terjual	Frekuensi Penjualan	Harga Produk (Rp)	Pendapatan Penjualan (Rp)	Cluster
0	2024-12-16	Banana Strudel	20	9	30000	600000	Penjualan Sedang
1	2024-12-16	Bolu Gulung Mocca	3	3	35000	105000	Penjualan Rendah
2	2024-12-16	Bolu Gulung Pandan	5	6	35000	175000	Penjualan Rendah
3	2024-12-16	Bolu Mentega	9	4	25000	225000	Penjualan Bagus
4	2024-12-16	Bolu Mentega Hias	8	2	27000	216000	Penjualan Rendah
...	...	...	...	...	...	...	...
351	2024-12-31	Roti Sobek Srikaya	8	8	12000	96000	Penjualan Bagus
352	2024-12-31	Roti Sosis	18	4	3000	54000	Penjualan Sedang
353	2024-12-31	Roti Tawar	17	12	17000	289000	Penjualan Bagus
354	2024-12-31	Selai Srikaya	17	12	11000	187000	Penjualan Bagus
355	2024-12-31	Sobek Keju	16	6	12000	192000	Penjualan Bagus

356 rows × 7 columns

Gambar 4. Perubahan nama *cluster*

Berdasarkan hasil di atas, dapat dianalisis bahwa terdapat 129 data yang merupakan anggota *cluster* “Penjualan Bagus”, 28 data pada *cluster* “Penjualan Sedang”, dan pada *cluster* “Penjualan Rendah” terdapat 199 data.

KESIMPULAN

Berdasarkan hasil penelitian yang dilakukan, dapat disimpulkan bahwa penerapan algoritma *K-Means Clustering* efektif dalam mengoptimalkan penjualan produk roti dengan menganalisis kinerja penjualan harian. Melalui pengelompokan produk roti berdasarkan frekuensi dan jumlah penjualan, tiga *cluster* yang berbeda berhasil terbentuk, yaitu *cluster* “Penjualan Bagus”, “Penjualan Sedang”, dan “Penjualan Rendah”. Dari 356 data yang digunakan, terdapat 129 data *cluster* “Penjualan Bagus”, 28 data dalam anggota *cluster* “Penjualan Sedang”, dan 199 yang masuk ke dalam *cluster* “Penjualan Rendah”. Pengelompokan ini memberikan wawasan yang jelas mengenai produk mana yang memiliki

performa penjualan terbaik, produk yang membutuhkan lebih banyak perhatian dalam promosi, dan produk yang perlu dikurangi stoknya.

Hasil *clustering* ini dapat dijadikan dasar bagi pengusaha roti untuk melakukan perencanaan produksi dan pemasaran yang lebih tepat sasaran, mengurangi pemborosan, dan meningkatkan efisiensi dalam pengelolaan inventaris. Meskipun *K-Means Clustering* dapat memberikan hasil yang baik dalam pengelompokan produk, terdapat beberapa keterbatasan seperti ketergantungan pada pemilihan awal *centroid* dan jumlah data yang digunakan. Oleh karena itu, penelitian lebih lanjut diperlukan untuk meningkatkan akurasi dan efisiensi penggunaan algoritma ini dalam bisnis roti dan sektor lainnya.

DAFTAR PUSTAKA

- [1] P. Primadona and R. Fauzi, "Penerapan *Data mining* Pada Penjualan Produk Elektronik," *Comput. Sci. Ind. Eng.*, vol. 9, no. 4, 2023, doi: 10.33884/comasiejournal.v9i4.7712.
- [2] F. Y. Silpa *et al.*, "Strategi Optimalisasi Manajemen Inventaris Dan Penjualan Di Toko SRC FIO," vol. 02, no. 01, pp. 323–332, 2024.
- [3] A. Aditya, I. Jovian, and B. N. Sari, "Implementasi *K-Means Clustering* Ujian Nasional Sekolah Menengah Pertama di Indonesia Tahun 2018/2019," *J. Media Inform. Budidarma*, vol. 4, no. 1, p. 51, 2020, doi: 10.30865/mib.v4i1.1784.
- [4] Y. Hartati, S. Defit, and G. W. Nurcahyo, "Klasterisasi Bibit Terbaik Menggunakan Algoritma K-Means dalam Meningkatkan Penjualan," *J. Inform. Ekon. Bisnis*, vol. 3, pp. 4–10, 2021, doi: 10.37034/infeb.v3i1.56.
- [5] N. Noviyanto, "Penerapan *Data mining* dalam Mengelompokkan Jumlah Kematian Penderita COVID-19 Berdasarkan Negara di Benua Asia," *Paradig. - J. Komput. dan Inform.*, vol. 22, no. 2, pp. 183–188, 2020, doi: 10.31294/p.v22i2.8808.
- [6] A. A. Aldino, D. Darwis, A. T. Prastowo, and C. Sujana, "Implementation of K-Means Algorithm for *Clustering* Corn Planting Feasibility Area in South Lampung Regency," *J. Phys. Conf. Ser.*, vol. 1751, no. 1, 2021, doi: 10.1088/1742-6596/1751/1/012038.
- [7] N. Hendrastuty, "Penerapan *Data mining* Menggunakan Algoritma *K-Means Clustering* Dalam Evaluasi Hasil Pembelajaran Siswa," *J. Ilm. Inform. Dan Ilmu Komput.*, vol. 3, no. 1, pp. 46–56, 2024, [Online]. Available: <https://doi.org/10.58602/jima-ilkom.v3i1.26>
- [8] T. Hidayat, "Klasifikasi Data Jamaah Umroh Menggunakan Metode *K-Means Clustering*," *J. Sistim Inf. dan Teknol.*, vol. 4, pp. 19–24, 2022, doi: 10.37034/jsisfotek.v4i1.115.
- [9] I. Syafrinal and E. L. Febrianti, "Penerapan Algoritma K-Means Pada Aplikasi *Data mining* Untuk Menentukan Pola Penjualan (Studi Kasus: Zahra Mart)," *J. Digit*, vol. 13, no. 1, p. 31, 2023, doi: 10.51920/jd.v13i1.320.
- [10] W. Handoko and A. Nasution, "Implementation of *Data mining* By Using K-Means To Classify Marriage Age," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 9, no. 2, pp. 161–166, 2023, doi: 10.33330/jurteksi.v9i2.2045.
- [11] S. Asyuti and A. A. Setyawan, "*Data mining* Dalam Penggunaan Presensi Karyawan Denga Cluster Means," *J. Ilm. Sains Teknol. Dan Inf.*, vol. 1, no. 1, pp. 01–10, 2023, [Online]. Available: <https://jurnal.alimspublishing.co.id/index.php/JITI/article/download/6/6>
- [12] A. D. Sidik and A. Ansawarman, "Prediksi Jumlah Kendaraan Bermotor Menggunakan *Machine learning*," *Formosa J. Multidiscip. Res.*, vol. 1, no. 3, pp. 559–568, 2022, doi: 10.55927/fjmr.v1i3.745.
- [13] P. Alkhairi and A. P. Windarto, "Penerapan K-Means *Cluster* pada Daerah Potensi Pertanian Karet Produktif di Sumatera Utara," *Semin. Nas. Teknol. Komput. Sains*, pp. 762–767, 2019, [Online]. Available: <https://prosiding.seminar-id.com/index.php/sainteks/article/download/228/223>
- [14] M. Ali, "Penerapan Algoritma K-Medoids untuk Klustering Bakery dan Cake yang Laris," *J.*



- Sains dan Teknol. Inf.*, vol. 2, no. 4, pp. 130–135, 2023, doi: 10.47065/jussi.v2i4.4604.
- [15] M. A. Sembiring, “Penerapan Metode Algoritma *K-Means Clustering* Untuk Pemetaan Penyebaran Penyakit Demam Berdarah Dengue (Dbd),” *J. Sci. Soc. Res.*, vol. 4, no. 3, p. 336, 2021, doi: 10.54314/jssr.v4i3.712.