

**PENERAPAN *K-MEANS CLUSTERING* DALAM
PENGELOMPOKAN PREFERENSI PELANGGAN
TERHADAP ALBUM K-POP DAN *MERCHANDISE*
BERDASARKAN DATA PENJUALAN *THJMINE*
*STORE***

SKRIPSI

**Oleh:
Aditiya Dwi Cahyo
202110715076**



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2025**

LEMBAR PERSETUJUAN PEMBIMBING

Judul Tugas Akhir : Penerapan *K-Means Clustering* Dalam
Pengelompokan Preferensi Pelanggan
Terhadap Album K-Pop dan *Merchandise*
Berdasarkan Data Penjualan *THJMine Store*.
Nama Mahasiswa : Aditiya Dwi Cahyo
Nomor Pokok Mahasiswa : 202110715076
Program Studi/Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian Tugas Akhir : 28 Juli 2025

Jakarta, 09 Juli 2025

MENYETUJUI,

Pembimbing I



Wowon Priatna, S.T., M.TI.

NIDN: 0429118007

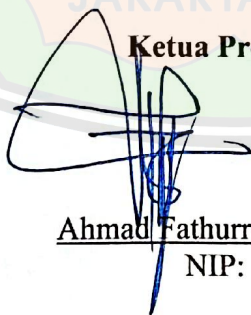
Pembimbing II



Agus Hidayat, S.Kom., M.Si.

NIDN: 0304088303

Ketua Program Studi



Ahmad Fathurrozi, S.E., M.M.S.I.

NIP: 2012486

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya

2025

LEMBAR PENGESAHAN

Judul Tugas Akhir : Penerapan *K-Means Clustering* Dalam
Pengelompokan Preferensi Pelanggan
Terhadap Album K-Pop dan *Merchandise*
Berdasarkan Data Penjualan *THJMine Store*

Nama Mahasiswa : Aditiya Dwi Cahyo

Nomor Pokok Mahasiswa : 202110715076

Program Studi/Fakultas : Informatika / Ilmu Komputer

Tanggal Lulus Ujian Tugas Akhir : 28 Juli 2025

Jakarta, 04 Agustus 2025

MENGESAHKAN,

Ketua Tim Penguji : Dr. Herlawati, S.Si., M.M., M.Kom
NIDN: 0311097302

Penguji I : Siti Setiawati, S.Pd., M.Pd
NIDN: 0313107904

Penguji II : Wowon Priatna, S.T., M.TI
NIDN: 0429118007

MENGETAHUI,

Ketua

Program Studi Informatika

Ahmad Fathurrozi, S.E., M.M.S.I.
NIP: 2012486

Dekan

Fakultas Ilmu Komputer

Dr. Robertus Suraji, S.S., M.A.
NIP: 1904406



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini:

Nama : Aditiya Dwi Cahyo
NPM : 202110715076
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Penerapan *K-Means Clustering* Dalam Pengelompokan
Preferensi Pelanggan Terhadap Album K-Pop dan
Merchandise Berdasrkan Data Penjualan *THJMINE Store*

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 04 Agustus 2025

Penulis



Aditiya Dwi Cahyo

ABSTRAK

Aditiya, 202110715076. Penerapan *K-Means Clustering* Dalam Pengelompokan Preferensi Pelanggan Terhadap Album K-Pop dan *Merchandise* Berdasarkan Data Penjualan *THJMine Store*. Jakarta: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya. 2025.

Pada beberapa tahun terakhir, industri musik global mengalami lonjakan signifikan, salah satunya oleh peningkatan popularitas diberbagai belahan dunia termasuk di Indonesia. Peningkatan ini tidak hanya pada musik digital saja, tetapi juga dalam pembelian produk fisik seperti album dan *merchandise*. Fenomena ini membuka peluang besar bagi pelaku bisnis ritel, termasuk pada toko *THJMine Store* yang menjual produk album dan *merchandise* K-Pop. Namun, ditengah antusias pasar yang tinggi, toko *THJMine Store* menghadapi permasalahan dalam pengelolaan stok serta menentukan strategi promosi yang tepat. Permasalahan utama yang dihadapi adalah belum adanya segmentasi pelanggan yang jelas serta kesulitan dalam promosi strategi yang tepat. Hal ini menyebabkan terjadinya stok berlebih dan strategi pemasaran yang kurang tepat sasaran. Berdasarkan permasalahan tersebut, penelitian ini bertujuan untuk mengatasi permasalahan tersebut dengan menerapkan algoritma *K-Means Clustering* dalam pendekatan *CRISP-DM* serta menghasilkan *cluster* berdasarkan pola pembelian yang dapat merepresentasikan segmentasi pelanggan. Penelitian ini menggunakan data transaksi yang digunakan sebanyak 710 total pembelian dari 225 total pelanggan atas 110 total produk periode Juli 2022 hingga Januari 2025, dengan menggunakan algoritma *K-Means Clustering* dan pendekatan *CRISP-DM* serta menggunakan evaluasi *Davies Bouldin Index* (DBI). Berdasarkan dari hasil pengelompokan preferensi pelanggan dengan algoritma *K-Means Clustering* berhasil membentuk tiga *cluster* preferensi pelanggan, yaitu *cluster 0*, *cluster 1*, dan *cluster 2*. Hasil evaluasi menunjukkan skor DBI sebesar 0.6342, yang menandakan kualitas *cluster* yang baik. Pengelompokan preferensi pelanggan ini dapat membantu *THJMine Store* dalam memahami segmentasi pelanggan, mengembangkan strategi pemasaran yang lebih tepat dan meningkatkan efisiensi pengelolaan stok berdasarkan data.

Kata Kunci: *Clustering, K-Means, Preferensi Pelanggan, Album K-Pop, CRISP-DM, Pola Penjualan, Data Mining, Davies Bouldin Index.*

ABSTRACT

Aditiya, 202110715076. *Application of K-Means Clustering In Grouping Customer Preferences For K-Pop Albums And Merchandise Based On THJMINE Store Sales Data.* Jakarta: Faculty of Computer Science. Bhayangkara Jakarta Raya University. 2025.

In recent years, the global music industry has experienced significant growth, partly due to the increasing popularity of music worldwide, including in Indonesia. This growth is not limited to digital music but also includes purchases of physical products such as album and merchandise. This phenomenon has opened up great opportunities for retail businesses, including THJMINE Store, which sells K-Pop albums and merchandise. However, amid high market enthusiasm, THJMINE Store faces challenges in inventory management and determining the right promotional strategy. The main issues faced are the lack of clear customer segmentation and difficulties in implementing effective promotional strategies. This has led to excess inventory and marketing strategies that are not on target. Based on these issues, this study aims to address them by applying the K-Means Clustering algorithm within the CRISP-DM approach and generating clusters based on purchasing patterns that can represent customer segmentation. This study uses transaction data consisting of 710 total purchases from 225 total customers for 110 total products sales from July 2022 to January 2025, utilizing the K-means Clustering algorithm and CRISP-DM approach, with evaluation using the Davies Bouldin Index (DBI). Based on the results of customer preference clusters were successfully formed: cluster 0, cluster 1, and cluster 2. The evaluation results showed a DBI score of 0.6342, indicating good cluster quality. This customer preference clustering can assist These THJMINE Store understanding customer segmentation, developing more targeted marketing strategies, and improve inventory management efficiency based on data.

Keyword: *Clustering, K-Means, Customer Preferences, K-Pop Albums, CRISP-DM, Sales Patterns, Data Mining, Davies Bouldin Index.*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

Sebagai sivitas akademik universitas bhayangkara jakarta raya, saya yang bertanda tangan dibawah ini:

Nama : Aditiya Dwi Cahyo
NPM : 202110715076
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

Penerapan *K-Means Clustering* Dalam Pengelompokan Preferensi Pelanggan Terhadap Album K-Pop dan *Merchandise* Berdasarkan Data Penjualan *THJMine Store*

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 04 Agustus 2025

Yang Menyatakan


Aditiya Dwi Cahyo

KATA PENGANTAR

Alhamdulillah, segala puji bagi Allah Subhanahu wa Ta'ala yang telah melimpahkan rahmat dan karunia-Nya, dan dapat menyelesaikan penelitian skripsi yang berjudul “Penerapan *K-Means Clustering* Dalam Pengelompokan Preferensi Pelanggan Terhadap Album K-Pop dan *Merchandise* Berdasarkan Data Penjualan *THJMine Store*.”

Penyusunan skripsi ini sebagai salah satu tahapan dalam proses penyelesaian tugas akhir guna memperoleh gelar Sarjana pada Program Studi Informatika, jenjang Strata-1, Universitas Bhayangkara Jakarta Raya. Penyusunan skripsi ini tentunya tidak terlepas dari dukungan dan bantuan berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terimakasih kepada:

1. Bapak Irjen Pol (Purn.) Prof. Dr. Drs. H. Bambang Karsono, S.H., M.M., Ph.D., D.Crim (Honoris Causa) selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Bapak Dr. Robertus Suraji, S.S., M.A. selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Ketua Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya.
4. Bapak Wowon Priatna, S.T., M.TI. selaku Dosen Pembimbing I pada penelitian ini.
5. Bapak Agus Hidayat, S. Kom., M.Si. selaku Dosen Pembimbing II, sekaligus Dosen Pembimbing Akademik.

6. Segenap Bapak dan Ibu Dosen Fakultas Ilmu Komputer yang telah memberikan ilmu, arahan, serta motivasi selama perkuliahan berlangsung.
7. Bapak Jarman dan Ibu Supartini, selaku orang tua yang senantiasa memberikan dukungan penuh, baik secara ekonomi, moral, dan spiritual.
8. Septi Lestari, S.E., selaku partner dan *support system* yang selalu memberikan dukungan dalam proses penyusunan penelitian.
9. Terimakasih kepada teman-teman MGG saya atas semangat, motivasi, tawa, dan dukungan yang luar biasa selama masa perkuliahan dalam penyusunan skripsi ini.
10. Terimakasih kepada seluruh alam semesta dan elemen-elemen serta makhluk-makhluk yang selalu mengilhami *positive vibes* dalam penyusunan hingga selesai.

Penulisan skripsi ini masih terdapat banyak kekurangan, baik dari segi penyampaian, penulisan, maupun substansi penelitian. Oleh karena itu, sangat terbuka terhadap kritik dan saran yang bersifat membangun demi perbaikan dan penyempurnaan tugas akhir ini ke depannya.

Jakarta 05 Mei 2025
Penulis



Aditiya Dwi Cahyo
202110715076

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PENGESAHAN	ii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK.....	vi
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Batasan Masalah.....	5
1.5 Tujuan Penelitian.....	5
1.6 Manfaat Penelitian	6
1.7 Sistematika Penulisan	7
BAB II LANDASAN TEORI	9
2.1 Tinjauan Pustaka	9
2.2 Pengelompokan Preferensi Pelanggan	12
2.3 Data Mining.....	13
2.4 Cross Industry Standard Process For Data Mining (CRISP-DM).....	14
2.5 Algoritma K-Means Clustering.....	16
2.6 Davies Bouldin Index (DBI).....	18

2.7	Pengertian <i>Python</i>	19
2.8	<i>Google Interactive Notebook (Google Colab)</i>	20
BAB III METODOLOGI PENELITIAN		21
3.1	Objek Penelitian dan Tempat Penelitian	21
3.2	Kerangka Penelitian	23
3.3	<i>Business Understanding</i>	23
3.4	<i>Data Understanding</i>	24
3.5	<i>Data Preparation</i>	25
3.6	<i>Modelling</i>	25
3.7	Evaluasi	26
3.8	<i>Deployment</i>	26
3.9	Metode Pengumpulan Data	27
BAB IV HASIL DAN PEMBAHASAN		31
4.1	Hasil Penelitian	31
4.2	<i>Business Understanding</i>	31
4.3	<i>Data Understanding</i>	33
4.4	<i>Data Preparation</i>	34
4.4.1	<i>Data Cleaning</i>	34
4.4.2	<i>Data Transformation</i>	39
4.5	<i>Modelling</i>	53
4.5.1	<i>Elbow Method</i>	55
4.5.2	<i>K-Means Modelling</i>	57
4.5.3	Hasil <i>Cluster</i>	75
4.6	<i>Evaluation</i>	77
4.7	Pembahasan	82
BAB V PENUTUP		84
5.1	Kesimpulan	84

5.2	Saran.....	85
DAFTAR PUSTAKA.....		87
LAMPIRAN.....		90



DAFTAR TABEL

Tabel 2.1 <i>State of the art</i>	9
Tabel 3.1 Hasil Wawancara	28
Tabel 4.1 Dataset Sebelum Dilakukan Remove Irrelevant Data	37
Tabel 4.2 Hasil Remove Irrelevant Data	38
Tabel 4.3 Atribut Baru “Harga Satuan” di Dataset THJMINE Store	39
Tabel 4.4 Hasil Klasifikasi 3 Kategori.....	41
Tabel 4.5 Hasil Aggregrasi Berdasarkan Kategori di Dataframe	50
Tabel 4.6 Dataframe Baru Normalisasi Data.....	53
Tabel 4.7 Data Training	54
Tabel 4.8 Data Testing	54
Tabel 4.9 Penjelasan Script Elbow Method.....	56
Tabel 4.10 Menentukan Centroid Awal	57
Tabel 4.11 Hasil Iterasi Ke-1	59
Tabel 4.12 Centroid Iterasi Ke-2	60
Tabel 4.13 Hasil Perhitungan Iterasi Ke-2.....	62
Tabel 4.14 Centroid Iterasi Ke-3	63
Tabel 4.15 Hasil Perhitungan Iterasi Ke-3.....	64
Tabel 4.16 Centroid Iterasi Ke-4	65
Tabel 4.17 Hasil Perhitungan Iterasi Ke-4.....	67
Tabel 4. 18 Centroid Iterasi Ke-5	68
Tabel 4.19 Hasil Perhitungan Iterasi Ke-5.....	69
Tabel 4.20 Centroid Iterasi Ke-6	70
Tabel 4.21 Hasil Perhitungan Iterasi Ke-6.....	72
Tabel 4.22 Centroid Pada Iterasi Ke-7.....	73
Tabel 4.23 Jarak Antara Titik Data Dan Centroid Cluster	78
Tabel 4.24 Hasil Perhitungan SSB	80

DAFTAR GAMBAR

Gambar 1.1 Grafik Penjualan Album K-Pop 2013-2023	1
Gambar 1.2 Hasil Survei Merchandise Yang Dimiliki Penggemar	2
Gambar 2.1 Tahapan <i>CRISP-DM</i>	14
Gambar 2.2 Logo <i>Python</i>	19
Gambar 2.3 Logo <i>Google Interactive Notebook (Google Colab)</i>	20
Gambar 3.1 <i>THJMINE Store</i> pada platfrom <i>Shopee</i>	21
Gambar 3.2 Maps <i>THJMINE Store</i>	22
Gambar 3.3 Kerangka Penelitian Penerapan Algoritma <i>K-Means Clustering</i> Dalam Pengelompokan Preferensi Pelanggan Terhadap Album K-Pop dan <i>Merchandise</i> Berdasarkan Data Penjualan <i>THJMINE Store</i>	23
Gambar 3.4 Dataset <i>THJMINE Store</i>	24
Gambar 4.1 Dataset Penjualan <i>THJMINE Store</i>	33
Gambar 4.2 <i>Checking Missing Values</i>	35
Gambar 4.3 Mengecek Nama Kolom Dan Tipe Data.....	36
Gambar 4.4 Melakukan <i>Remove Irrelevant Data</i>	38
Gambar 4.5 Membuat Atribut Baru Bernama “harga satuan”	40
Gambar 4.6 Klasifikasi 3 Kategori	41
Gambar 4.7 <i>Script</i> Visualisasi Jumlah Penjualan Per Item Album	43
Gambar 4.8 Visualisasi Jumlah Penjualan Per Item Album	43
Gambar 4.9 <i>Script</i> Visualisasi Jumlah Penjualan Per Item <i>Photocard</i>	44
Gambar 4.10 Visualisasi Jumlah Penjualan Per Item <i>Photocard</i>	45
Gambar 4.11 <i>Script</i> Visualisasi Jumlah Penjualan Per Item <i>Merchandise</i>	46
Gambar 4.12 Hasil Visualisasi Jumlah Penjualan Per Item <i>Merchandise</i>	46
Gambar 4.13 <i>Script</i> Visualisasi <i>Pie Chart</i> Jumlah Per Kategori	47
Gambar 4.14 Hasil Visualisasi <i>Pie Chart</i> Penjualan Per Kategori	48
Gambar 4.15 Penambahan Fitur Baru Dan Penggabungan <i>Dataframe</i>	49
Gambar 4.16 Distribusi Data	50
Gambar 4.17 Distribusi Data album, <i>merchandise</i> , <i>photocard</i> , <i>total_item</i> dan <i>total_belanja</i>	51

Gambar 4.18 Visualisasi <i>Elbow Method</i>	56
Gambar 4.19 Kurva <i>Elbow</i> Jumlah <i>Cluster</i> Optimal.....	56
Gambar 4.20 Melakukan <i>Modelling K-Means Clustering</i>	74
Gambar 4.21 <i>Centroid</i> Akhir	74
Gambar 4.22 Rata-Rata <i>Cluster</i>	75
Gambar 4.23 Membuat Kolom Baru	76
Gambar 4.24 Hasil Dataset Dengan Label	77
Gambar 4.25 Untuk Melihat <i>Score DBI</i>	81
Gambar 4.26 Hasil Evaluasi <i>DBI</i>	82



DAFTAR LAMPIRAN

- Lampiran 1** Hasil Cek Plagiarisme
- Lampiran 2** Biodata Mahasiswa
- Lampiran 3** Kartu Konsultasi Bimbingan Pembimbing I
- Lampiran 4** Kartu Konsultasi Bimbingan Pembimbing II
- Lampiran 5** Rekomendasi Pembimbing I
- Lampiran 6** Rekomendasi Pembimbing II
- Lampiran 7** Surat Permohonan Izin Penelitian
- Lampiran 8** Surat Balasan Penelitian
- Lampiran 9** Hasil Wawancara
- Lampiran 10** Dataset Penjualan toko *THJMINE Store*

