

**KLASTERISASI DAN PREDIKSI POLA PENJUALAN
DI CV. CIPTA USAHA SELARAS DENGAN
ALGORITMA K-MEANS DAN REGRESI
LINEAR SEDERHANA**

SKRIPSI

Oleh:
Tia Aulia
202110715185



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

**KLASTERISASI DAN PREDIKSI POLA PENJUALAN
DI CV. CIPTA USAHA SELARAS DENGAN
ALGORITMA *K-MEANS* DAN REGRESI
LINEAR SEDERHANA**

SKRIPSI

Oleh:
Tia Aulia
202110715185



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

LEMBAR PERSETUJUAN PEMBIMBING

Judul Skripsi : Klasterisasi dan Prediksi Pola Penjualan di CV.
Cipta Usaha Selaras dengan Algoritma K-Means
dan Regresi Linear Sederhana

Nama Mahasiswa : Tia Aulia

Nomor Pokok Mahasiswa : 202110715185

Program Studi/Fakultas : Informatika/Illmu Komputer

Tanggal Lulus Ujian Skripsi : 28 Juli 2025

Jakarta, 04 Agustus 2025
MENYETUJUI,

Pembimbing I



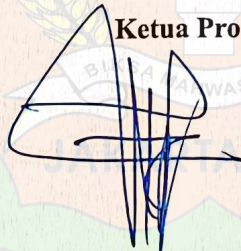
Wowon Priatna, S.T., M.TI
NIDN: 0429118007

Pembimbing II



Muhammad Yasir, S.Si., M.Kom
NIDN: 0317129009

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 Skripsi : Klasterisasi dan Prediksi Pola Penjualan di CV. Cipta
Usaha Selaras dengan Algoritma K-Means dan Regresi
Linear Sederhana
Nama Mahasiswa : Tia Aulia
Nomor Pokok Mahasiswa : 202110715185
Program Studi/Fakultas : Informatika / Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 28 Juli 2025

Jakarta, 04 Agustus 2025

MENGESAHKAN,

Ketua Tim Penguji : Rafika Sari, S.Si., M.Si
NIDN: 0329098902

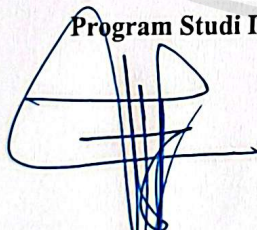
Penguji I : Rasim, S.T., M.Kom
NIDN: 0415027301

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

MENGETAHUI,

Ketua

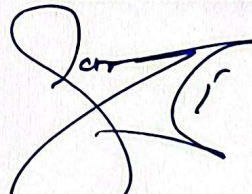
Program Studi Informatika



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

Dekan

Fakultas Ilmu Komputer



Dr. Robertus Suraji, SS., MA
NIP: 1904406



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Tia Aulia
NPM : 202110715185
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Skripsi : Klasterisasi dan Prediksi Pola Penjualan di CV. Cipta
Usaha Selaras dengan Algoritma K-Means dan Regresi
Linear Sederhana

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 mana pun.

Bekasi, 04 Agustus 2025
Penulis,



Tia Aulia

ABSTRAK

Tia Aulia, 202110715185. Klasterisasi dan Prediksi Pola Penjualan di CV. Cipta Usaha Selaras dengan Algoritma K-Means dan Regresi Linear Sederhana, Bekasi. Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya, 2025.

Dalam era digital dan persaingan bisnis yang semakin kompetitif, pemanfaatan data menjadi hal krusial untuk mendukung pengambilan keputusan yang lebih tepat sasaran. Salah satu tantangan yang dihadapi CV. Cipta Usaha Selaras adalah kesulitan dalam memahami pola pembelian pelanggan dan memperkirakan nilai penjualan secara akurat. Penelitian ini bertujuan untuk mengelompokkan pelanggan berdasarkan perilaku pembelian serta memprediksi total pembelian menggunakan pendekatan algoritma K-Means dan Regresi Linear Sederhana. Data yang digunakan merupakan data transaksi tahun 2024 yang mencakup frekuensi pembelian (Qty PO) dan total nilai transaksi. Proses analisis dilakukan berdasarkan tahapan CRISP-DM, dimulai dari pemahaman bisnis hingga visualisasi hasil. Hasil klasterisasi menunjukkan bahwa pelanggan dapat dibagi ke dalam tiga segmen dengan karakteristik yang berbeda, dan evaluasi menggunakan nilai Silhouette Score rata-rata sebesar 0,7913 mengindikasikan bahwa segmentasi yang dihasilkan tergolong cukup optimal. Sementara itu, model prediksi menghasilkan persamaan regresi linier dengan nilai koefisien determinasi (R^2) sebesar 0,6910, nilai MAE sebesar Rp213 juta, dan nilai MSE sebesar Rp206 triliun. Meskipun terdapat beberapa data dengan hasil prediksi yang berbeda cukup jauh dari nilai aslinya, model ini tetap menunjukkan performa yang cukup baik sebagai acuan awal untuk estimasi penjualan. Hasil penelitian ini diharapkan dapat digunakan sebagai dasar dalam penyusunan strategi pemasaran yang lebih terarah dan berbasis data, serta menjadi fondasi awal untuk penerapan analisis pelanggan yang berkelanjutan..

Kata Kunci: *Klasterisasi, K-Means, Regresi Linear Sederhana, Prediksi Penjualan, Segmentasi Pelanggan.*

ABSTRACT

Tia Aulia, 202110715185. *Customer Clustering and Sales Pattern Prediction at CV. Cipta Usaha Selaras Using K-Means Algorithm and Simple Linear Regression.*
Bekasi: Faculty of Computer Science, University Of Bhayangkara Jakarta Raya, 2025.

In the digital era and increasingly competitive business environment, data utilization has become crucial to support more accurate and targeted decision-making. One of the challenges faced by CV. Cipta Usaha Selaras is the difficulty in understanding customer purchasing patterns and accurately forecasting sales values. This study aims to group customers based on their purchasing behavior and predict total purchases using the K-Means algorithm and Simple Linear Regression. The data used consists of 2024 transaction records, including purchase frequency (Qty PO) and total transaction value. The analysis process follows the CRISP-DM stages, starting from business understanding to result visualization. The clustering results indicate that customers can be divided into three distinct segments, with an average Silhouette Score of 0.7913, suggesting that the segmentation is reasonably optimal. Meanwhile, the prediction model produces a linear regression equation with a coefficient of determination (R^2) of 0.6910, a Mean Absolute Error (MAE) of Rp213 million, and a Mean Squared Error (MSE) of Rp206 trillion. Although some data points showed predictions that significantly differed from the actual values, the model still demonstrates a reasonably good performance as an initial basis for sales estimation. The findings of this research are expected to serve as a reference for developing more targeted, data-driven marketing strategies and as a foundation for ongoing customer analytics implementation.

Keywords: *Clustering, K-Means, Simple Linear Regression, Sales Prediction, Customer Segmentation.*

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK

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

Nama : Tia Aulia
NPM : 202110715185
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 :

“Klasterisasi dan Prediksi Pola Penjualan di CV. Cipta Usaha Selaras dengan Algoritma K-Means dan Regresi Linear Sederhana”

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 : Bekasi
Pada tanggal : 04 Agustus 2025

Yang Menyatakan



Tia Aulia

KATA PENGANTAR

Segala puji dan syukur ke hadirat Allah SWT atas limpahan rahmat, karunia, dan hidayah-Nya sehingga tugas akhir dengan judul: **“Klasterisasi dan Prediksi Pola Penjualan di CV. Cipta Usaha Selaras dengan Algoritma K-Means dan Regresi Linear Sederhana”** dapat diselesaikan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Bhayangkara Jakarta Raya. Ucapan terima kasih disampaikan kepada semua pihak yang telah memberikan bimbingan, dukungan, dan doa selama proses penyusunan skripsi ini. Secara khusus, ucapan terimakasih yang tulus diberikan kepada:

1. Bapak Irjen Polisi (Purn) Prof. Dr. Drs. H. Bambang Karsono, S.H., M.M., Ph.D., D.Crim.(Hc) 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.T.I., selaku Dosen Pembimbing I, atas bimbingan, arahan, dan dukungan yang konsisten selama proses penyusunan skripsi ini.
5. Bapak Muhammad Yasir, S.Si., M.Kom., selaku Dosen Pembimbing II, atas bimbingan, arahan, dan motivasi yang konsisten diberikan selama penyusunan skripsi ini.
6. Bapak Hendarman Lubis, S.Kom., M.Kom., selaku Dosen Pembimbing Akademik, atas pendampingan dan dukungannya selama masa perkuliahan.

7. Bapak Kusmayadi, selaku Direktur Utama CV. Cipta Usaha Selaras, atas kesempatan dan data yang telah diberikan untuk kepentingan penelitian ini.
8. Bapak Wakhid Khaironi, selaku Koordinator Produksi di CV. Cipta Usaha Selaras, atas bantuan dan kerja samanya dalam penyediaan data serta penjelasan yang mendukung kelancaran proses penelitian.
9. Kedua orang tua tercinta, atas kasih sayang, doa, dukungan moral, dan semangat yang tiada henti dalam setiap langkah, serta menjadi sumber kekuatan selama proses studi dan penyusunan skripsi ini.
10. Zacky Achmad Sholeh, selaku partner skripsi yang telah banyak membantu dalam proses diskusi, berbagi referensi, dan memberi dukungan secara teknis maupun moral sepanjang perjalanan penyusunan tugas akhir ini.
11. Kepada rekan-rekan seperjuangan, khususnya “Cemara” dan “Princess A5”, atas kebersamaan, semangat, dan dukungan moral yang sangat berarti selama masa perkuliahan hingga penyusunan skripsi ini.
12. Kepada *Treasure*, yang melalui karya dan semangatnya menjadi hiburan sekaligus motivasi dalam menjalani masa-masa pengerjaan skripsi.

Segala bentuk kritik dan saran yang membangun sangat diharapkan untuk menyempurnakan karya ini. Semoga hasil penelitian ini dapat memberikan manfaat bagi pihak-pihak yang membutuhkan dan menjadi kontribusi nyata dalam pengembangan ilmu pengetahuan.

Jakarta, 04 Agustus 2025
Hormat Saya



Tia Aulia
202110715185

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR.....	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	3
1.3 Rumusan Masalah	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	4
1.6 Batasan Masalah.....	4
1.7 Sistematika Penelitian	5
BAB II TINJAUAN PUSTAKA.....	7
2.1 <i>State of The Art</i>	7
2.2 <i>Machine Learning</i>	11
2.3 <i>Data Mining</i>	12
2.4 <i>Cross Industry Standard Process for Data Mining (CRISP-DM)</i>	12
2.5 Klasterisasi	14
2.6 Algoritma <i>K-Means</i>	15
2.7 Evaluasi Klasterisasi (<i>Silhouette Score</i>).....	17
2.8 Regresi Linear Sederhana.....	18
2.9 Evaluasi Model Prediksi.....	20
2.10 Prediksi Penjualan	22

2.11	Data Penelitian	24
2.12	<i>Tools</i> Pemograman	24
BAB III METODOLOGI PENELITIAN		26
3.1	Obyek dan Waktu Penelitian	26
3.1.1	Sejarah Perusahaan.....	26
3.1.2	Struktur Organisasi.....	27
3.2	Kerangka Penelitian	29
3.3	Uraian Metodologi	30
3.3.1	Identifikasi Masalah	30
3.3.2	Rumusan Masalah	30
3.3.3	Metode Pengumpulan Data	30
3.3.4	Tahapan <i>CRISP-DM</i>	31
3.4	Metode Analisis.....	39
BAB IV HASIL DAN PEMBAHASAN.....		41
4.1.	<i>Business Understanding</i>	41
4.2	<i>Data Understanding</i>	41
4.3	<i>Data Preparation</i>	43
4.3.1	Pembersihan Data.....	43
4.3.2	Transformasi Data	44
4.3.3	Pengolahan Data per Pelanggan.....	45
4.3.4.	<i>Normalisasi Data</i>	46
4.4	<i>Modeling</i>	48
4.4.1	<i>K-Means Modeling</i>	48
4.4.2	Hasil Klaster.....	52
4.4.3	Regresi Linear Sederhana <i>Modeling</i>	54
4.4.4	Hasil Prediksi Total Pembelian Berdasarkan Frekuensi	58
4.5	<i>Evaluation</i>	59
4.5.1	Evaluasi Model <i>K-Means</i>	60
4.5.2	Evaluasi Model Regresi Linear.....	62
4.6	<i>Deployment</i>	65
4.6.1	Visualisasi Model <i>K-Means</i>	65

4.6.2	Visualisasi Model Linear Regresi Sederhana.....	67
4.7	Pembahasan	69
BAB V PENUTUP.....		75
5.1	Kesimpulan.....	75
5.2	Saran.....	76
DAFTAR PUSTAKA.....		78
LAMPIRAN.....		81



DAFTAR TABEL

Tabel 2. 1 <i>State of The Art</i>	7
Tabel 3. 1 Formasi Karyawan Pelaksana Produksi di CV. Cipta Usaha Selaras...	28
Tabel 3. 2 Detail Atribut Utama Dataset	33
Tabel 3. 3 Data Setelah Melewati Tahap Preparation	35
Tabel 3. 4 Data Ternormalisasi.....	36
Tabel 4. 1 Struktur Dataset Penjualan CV. Cipta Usaha Selaras Tahun 2024	42
Tabel 4. 2 Jumlah Missing Values pada Setiap Kolom Dataset	43
Tabel 4. 3 Hasil Transformasi Atribut Total Harga	44
Tabel 4. 4 Hasil Konversi Atribut Tanggal.....	44
Tabel 4. 5 Hasil Data Tahap Preparation.....	45
Tabel 4. 6 Hasil Normalisasi Fitur Klasterisasi.....	47
Tabel 4. 7 Centroid Awal Berdasarkan Karakteristik Pelanggan	49
Tabel 4. 8 Hasil Iterasi Pertama Berdasarkan Jarak ke <i>Centroid</i> Awal	50
Tabel 4. 9 Hasil Pembaruan <i>Centroid</i> pada Iterasi Kedua	51
Tabel 4. 10 Hasil Evaluasi Klaster pada Iterasi Kedua	52
Tabel 4. 11 Hasil Normalisasi Akhir Klasterisasi Pelanggan.....	53
Tabel 4. 12 Hasil Frekuensi Pembelian dari Setiap Klaster	53
Tabel 4. 13 Frekuensi Pembelian dan Nilai Transaksi Pelanggan	56
Tabel 4. 14 Komponen Perhitungan Regresi Linear	56
Tabel 4. 15 Hasil Prediksi Total Pembelian Berdasarkan Frekuensi Transaksi	59
Tabel 4. 16 Rata-Rata Nilai <i>Silhouette Score</i> tiap Klaster	60
Tabel 4. 17 Hasil Evaluasi Model Regresi Linear Sederhana	64
Tabel 4. 18 Hasil Klasterisasi <i>K-Means</i>	69
Tabel 4. 19 Rekomendasi Pendekatan Lanjutan dan Pelanggan Utama per Klaster	70
Tabel 4. 20 Evaluasi Kualitas Klaster Berdasarkan <i>Silhouette Score</i>	71
Tabel 4. 21 Hasil Prediksi Total Pembelian Berdasarkan Frekuensi.....	72

DAFTAR GAMBAR

Gambar 2. 1 Jenis <i>Machine Learning</i>	11
Gambar 2. 2 Tahapan <i>CRISP-DM</i>	13
Gambar 2. 3 Metode Klasterisasi	15
Gambar 3. 1 Logo Cv Cipta Usaha Selaras.....	26
Gambar 3. 2 Struktur Organisasi CV. Cipta Usaha Selaras	27
Gambar 3. 3 Kerangka Penelitian	29
Gambar 3. 4 Tahapan <i>Business Understanding</i>	32
Gambar 3. 5 Tahapan <i>Data Understanding</i>	33
Gambar 3. 6 Tahapan <i>Data Preparation</i>	34
Gambar 3. 7 Tahapan <i>Modeling</i>	36
Gambar 3. 8 Alur Evaluasi Model Klasterisasi dan Prediksi	39
Gambar 4. 1 Top 3 Pelanggan Berdasarkan Total Pembelian per Segmentasi.....	66
Gambar 4. 2 Top 3 Pelanggan Berdasarkan Frekuensi Pembelian per Segmentasi	67
Gambar 4. 3 Grafik prediksi total pembelian dari frekuensi transaksi	68

DAFTAR LAMPIRAN

Plagiarisme

Biodata Mahasiswa

Kartu Konsultasi Bimbingan

Kartu Bimbingan Pembimbing 1

Kartu Bimbingan Pembimbing 2

Surat Rekomendasi Pembimbing 1

Surat Rekomendasi Pembimbing 2

Hasil Klasterisasi K-Means

Hasil Prediksi Pembelian

Dataset

Surat Permohonan Izin Penelitian

Surat Balasan Izin Penelitian CV. Cipta Usaha Selaras

