

**OPTIMASI PROMOSI VAKSIN MELALUI SEGMENTASI PELANGGAN DENGAN
METODE *CLUSTERING* DI PT. XYZ**



NAFA AZZAHRA

2110512050

**S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
2025**

PERNYATAAN ORISINALITAS

PERNYATAAN ORISINALITAS

Skripsi ini merupakan hasil karya sendiri serta semua sumber informasi yang dikutip maupun yang dirujuk telah saya nyatakan benar.

Nama : Nafa Azzahra

NIM : 2110512050

Tanggal : 08 Juli 2025

Bilamana dikemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia dituntut dan diproses sesuai dengan ketentuan berlaku.

Jakarta, 08 Juli 2025



Nafa Azzahra

PERNYATAAN PERSETUJUAN

PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Pembangunan Nasional “Veteran” Jakarta, saya bertanda tangan di bawah ini :

Nama : Nafa Azzahra
NIM : 2110512050
Fakultas : Ilmu Komputer
Program Studi : S1 Sistem Informasi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Pembangunan Nasional “Veteran” Jakarta Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalti Free Right*) atas karya ilmiah saya yang berjudul:

Optimasi Promosi Vaksin Melalui Segmentasi Pelanggan

Dengan Metode *Clustering* Di PT. XYZ

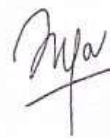
Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti ini Universitas Pembangunan Nasional “Veteran” Jakarta berhak menyimpan, mengalih data/formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat dan mempublikasikan Tugas Akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Dengan demikian pernyataan ini saya buat dengan sebenar-benarnya.

Dibuat di : Jakarta

Pada tanggal : 8 Juli 2025

Yang Menyatakan



Nafa Azzahra

LEMBAR PENGESAHAN

LEMBAR PENGESAHAN

Judul : Optimasi Promosi Vaksin Melalui Segmentasi Pelanggan
Dengan Metode Clustering Di PT. XYZ
Nama : Nafa Azzahra
NIM : 2110512050
Program Studi : S1 Sistem Informasi

Disetujui oleh :

Penguji 1:
I Wayan Widi Pradnyana, S.Kom., MTI.



Penguji 2:
Ruth Mariana Bunga Wadu, S.Kom., M.MSI.



Pembimbing 1:
Nur Hafifah Matondang, S.Kom., MM., MTI.

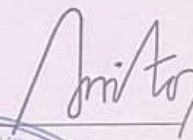
Pembimbing 2:
Octaviano Pratama, S.Kom., M.Kom.



Diketahui oleh :

Koordinator Program Studi:
Anita Muliawati, S.Kom., MTI.
NIP. 197005212021212002

Dekan Fakultas Ilmu Komputer
Prof. Dr. Ir. Supriyanto, S.T., M.Sc., IPM
NIP. 197605082003121002





Tanggal Ujian Tugas Akhir :
3 Juli 2025

**PERNYATAAN MENGENAI SKRIPSI DAN
SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA**

Dengan ini saya menyatakan bahwa skripsi dengan judul “Optimasi Promosi Vaksin Melalui Segmentasi Pelanggan Dengan Metode *Clustering* Di PT. XYZ” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir proposal ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Universitas Pembangunan Nasional “Veteran” Jakarta.

Jakarta, 08 Juli 2025



Nafa Azzahra
2110512050

ABSTRAK

Dalam dunia bisnis, suatu perusahaan harus dapat memahami serta mengetahui karakteristik pelanggan sebagai hal utama sehingga diperlukan analisis segmentasi pelanggan untuk mengelompokkan pelanggan agar perusahaan dapat mengetahui serta menganalisis kebutuhan masing-masing kelompok pelanggan. PT. XYZ merupakan perusahaan yang memiliki produk berupa platform di bidang kesehatan anak untuk memantau pertumbuhan serta perkembangan anak. Dalam menjalankan bisnisnya, PT. XYZ menghadapi tantangan berupa ketidakstabilan dalam layanan pemesanan vaksin yang terjadi karena kurangnya *brand awareness* dan metode promosi yang tidak dilakukan berdasarkan data. Penelitian ini membahas mengenai segmentasi pelanggan dengan menggunakan metode *k-means clustering* melalui 3 variabel yang ditentukan berdasarkan RFM model *recency*, *frequency* dan *monetary*. Data yang digunakan merupakan data pemesanan vaksin bulan Januari 2023 hingga Mei 2024, yang diperoleh langsung dari PT. XYZ. Hasil analisis menghasilkan nilai K optimal yaitu 4 berdasarkan *elbow method*. Hasil evaluasi model menggunakan metode *silhouette score* sebesar 0.379610 dan *davies bouldin index* sebesar 0.930230. Visualisasi dikembangkan pada streamlit dengan bantuan variabel lainnya untuk melihat karakteristik pelanggan. Penelitian ini berkontribusi terhadap pengelompokkan segmentasi pelanggan untuk memahami karakteristik pelanggan sehingga perusahaan dapat melakukan promosi dengan lebih terarah dan tepat sasaran.

Kata Kunci: *K-Means Clustering*, Promosi, Segmentasi Pelanggan, Visualisasi Data

ABSTRACT

In the business world, a company must be able to understand and know the characteristics of customers as the main thing so that customer segmentation analysis is needed to group customers so that the company can know and analyze the needs of each customer group. PT. XYZ is a company that has a product in the form of a platform in the field of children's health to monitor the growth and development of children. In running its business, PT. XYZ faces challenges in the form of instability in vaccine ordering services that occur due to lack of brand awareness and promotional methods that are not carried out based on data. This study discusses customer segmentation using the k-means clustering method through 3 variables determined based on the RFM model recency, frequency and monetary. The data used is vaccine ordering data from January 2023 to May 2024, which was obtained directly from PT. XYZ. The results of the analysis produce an optimal K value of 4 based on the elbow method. The results of the model evaluation using the silhouette score method are 0.379610 and the Davies Bouldin index is 0.930230. Visualization is developed on streamlit with the help of other variables to see customer characteristics. This research contributes to customer segmentation grouping to understand customer characteristics so that companies can carry out promotions in a more targeted and targeted manner.

Keywords: *K-Means Clustering*, Promotion, Customer Segmentation, Data Visualization

**OPTIMASI PROMOSI VAKSIN MELALUI SEGMENTASI PELANGGAN DENGAN
METODE *CLUSTERING* DI PT. XYZ**

NAFA AZZAHRA

2110512050

Proposal Skripsi
sebagai salah satu syarat untuk melaksanakan
penelitian oleh mahasiswa pada
Program Studi S1 Sistem Informasi

**S1 SISTEM INFORMASI
FAKULTAS ILMU KOMPUTER
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAKARTA
2025**

KATA PENGANTAR

Puji serta syukur penulis panjatkan kepada yang paling utama Tuhan Yang Maha Esa atas segala anugerah serta karunia yang telah diberikan, sehingga penulisan skripsi ini dapat selesai dengan tepat waktu sebagai syarat untuk kelulusan pada Program Studi S1 Sistem Informasi.

Dalam proses pengerjaan skripsi ini, penulis telah menerima bimbingan, bantuan, doa serta dukungan dari banyak pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih sebanyak banyaknya kepada :

1. Tuhan Yang Maha Esa atas segala nikmat, anugerah dan karunia-Nya.
2. Ayah dan Bunda yang sudah memberikan *support* dalam bentuk apapun dan doa yang tiada henti.
3. Prof. Dr. Ir. Supriyanto, M.Sc., IPM., selaku Dekan Fakultas Ilmu Komputer UPN "Veteran" Jakarta.
4. Ibu Anita Muliawati, S.Kom., MTI., selaku Ketua Program Studi S1 Sistem Informasi UPN "Veteran" Jakarta.
5. Ibu Nur Hafifah Matondang, S.Kom., MM., M.TI. selaku dosen pembimbing 1 yang sudah meluangkan waktunya untuk memberikan bimbingan, saran serta motivasi selama penulisan skripsi ini.
6. Bapak M. Octaviano Pratama, S.Kom., M.Kom selaku dosen pembimbing 2 penulis yang sudah memberikan bimbingan terkait penulisan skripsi ini.
7. Bima, Azizah dan Yeremi yang sudah memberikan waktu, dukungan dan semangat serta menemani penulis dalam pengerjaan skripsi.
8. Teman – teman penulis yang tidak bisa disebutkan satu persatu.

Skripsi ini masih sangat jauh dari kata sempurna. Penulis sangat terbuka terhadap masukan yang membangun guna perbaikan di masa mendatang. Penulis berharap agar skripsi ini dapat berkontribusi memberikan dampak yang baik bagi banyak orang.

Jakarta, 1 Juli 2025



Penulis

DAFTAR ISI

PERNYATAAN ORISINALITAS.....	ii
PERNYATAAN PERSETUJUAN.....	iii
LEMBAR PENGESAHAN.....	iv
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR GAMBAR.....	xii
DAFTAR TABEL.....	xiv
DAFTAR SIMBOL.....	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Tujuan Penelitian.....	3
1.4. Manfaat Penelitian.....	4
1.5. Ruang Lingkup Penelitian.....	4
1.6. Luaran yang Diharapkan	4
1.7. Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Promosi.....	6
2.2. Segmentasi Pelanggan.....	6
2.3. <i>Data Mining</i>	7
2.4. <i>Machine Learning</i>	8
2.5. RFM Model	9
2.6. <i>Clustering</i>	9
2.7. Metode K-Means <i>Clustering</i>	10
2.8. Metode <i>Elbow</i>	10
2.9. <i>Davies-Bouldin Index (DBI)</i>	11
2.10. <i>Silhouette Score</i>	12
2.11. Python.....	12
2.12. Streamlit	13
2.13. Penelitian Terdahulu	13
BAB III METODE PENELITIAN	17

3.1.	Tahapan Penelitian	17
3.2.	Identifikasi Masalah	18
3.3.	Studi Literatur	18
3.4.	Pemahaman Data.....	18
3.5.	<i>Data Preprocessing</i>	18
3.5.1.	<i>Data Cleaning</i>	18
3.5.2.	<i>Data Transformation</i>	19
3.5.3.	<i>Data Selection</i>	19
3.6.	Penerapan Algoritma K-Means	19
3.7.	Evaluasi Model.....	20
3.8.	Analisis dan Visualisasi Hasil <i>Clustering</i>	20
3.9.	Alat Bantu Penelitian	21
3.9.1.	Perangkat Keras (<i>Hardware</i>)	21
3.9.2.	Perangkat Lunak (<i>Software</i>).....	21
3.10.	Jadwal Rencana Penelitian	21
BAB IV	HASIL DAN PEMBAHASAN.....	23
4.1	Pemahaman Data.....	23
4.2	Data Preprocessing.....	23
4.2.1	<i>Data Cleaning</i>	24
4.2.2	<i>Data Transformation</i>	25
4.2.3	<i>Data Selection</i>	26
4.3	Penerapan Algoritma K-Means	28
4.3.1	Elbow Method	28
4.3.2	Inisialisasi Centroid Awal	29
4.3.3	Proses K-Means <i>Clustering</i>	30
4.4	Evaluasi Model.....	31
4.5	Analisis dan Visualisasi Hasil	33
4.5.1	Analisis Cluster	33
4.5.2	Visualisasi Hasil	34
4.6	Tampilan Halaman Streamlit	41
BAB V	PENUTUP	47
5.1	Kesimpulan.....	47
5.2	Saran.....	47

DAFTAR PUSTAKA	49
DAFTAR RIWAYAT HIDUP.....	51
LAMPIRAN.....	52

DAFTAR GAMBAR

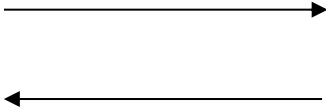
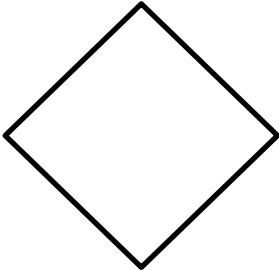
Gambar 1. 1 Cohort analysis pelanggan pemesanan vaksin	2
Gambar 2. 1 The 7PS marketing mix	6
Gambar 2. 2 Tahapan data mining	8
Gambar 2. 3 Area machine learning.....	8
Gambar 2. 4 Flowchart algoritma k-means clustering	10
Gambar 2. 5 Gambar titik elbow	11
Gambar 3. 1 Diagram alir tahap penelitian	17
Gambar 4. 1 Pengecekan missing values dan data duplicate	24
Gambar 4. 2 Data setelah dilakukan penanganan missing values.....	25
Gambar 4. 3 Data setelah dilakukan transformasi.....	26
Gambar 4. 4 Pemilihan variabel berdasarkan RFM	27
Gambar 4. 5 Data setelah dilakukan preprocessing	27
Gambar 4. 6 Elbow method.....	28
Gambar 4. 7 Hasil k-means clustering	30
Gambar 4. 8 Hasil evaluasi model	31
Gambar 4. 9 Visualisasi hasil klasterisasi dengan PCA.....	32
Gambar 4. 10 Analisis variabel RFM.....	33
Gambar 4. 11 Visualisasi hasil top 10 kota dengan pelanggan terbanyak tahun 2023	35
Gambar 4. 12 Visualisasi hasil top 10 kota dengan pelanggan terbanyak tahun 2024	35
Gambar 4. 13 Visualisasi distribusi cluster berdasarkan bulan pemesanan tahun 2023	36
Gambar 4. 14 Visualisasi distribusi cluster berdasarkan bulan pemesanan tahun 2024	37
Gambar 4. 15 Visualisasi distribusi jenis vaksin teratas 2023	37
Gambar 4. 16 Visualisasi distribusi jenis vaksin teratas 2023	38
Gambar 4. 17 Distribusi group age per-cluster 2023	38
Gambar 4. 18 Distribusi group age per-cluster 2024	39
Gambar 4. 19 Distribusi child gender tahun 2024	40
Gambar 4. 20 Distribusi child gender tahun 2024	40
Gambar 4. 21 Distribusi cluster berdasarkan kategori promo 2023.....	41
Gambar 4. 22 Distribusi cluster berdasarkan kategori promo 2024.....	41
Gambar 4. 23 Tampilan awal streamlit	42
Gambar 4. 24 Tampilan halaman hasil clustering 1	42

Gambar 4. 25 Tampilan halaman hasil clustering 2	43
Gambar 4. 26 Tampilan halaman analisis cluster 2	43
Gambar 4. 27 Tampilan halaman analisis cluster 3	44
Gambar 4. 28 Tampilan halaman visualisasi data 1	44
Gambar 4. 29 Tampilan halaman visualisasi data 2	45
Gambar 4. 30 Tampilan halaman data cluster 1	45
Gambar 4. 31 Tampilan halaman data cluster 2	46

DAFTAR TABEL

Tabel 2. 1 Ringkasan penelitian terdahulu	13
Tabel 3. 1 Jadwal rencana penelitian.....	21
Tabel 4. 1 Struktur dataset.....	23
Tabel 4. 2 Nilai sum of squared errors (SSE)	28
Tabel 4. 3 Kode python inisialisasi centroid	29
Tabel 4. 4 Hasil inisialisasi centroid	29
Tabel 4. 5 Kode python k-means clustering.....	30
Tabel 4. 6 Distribusi data per-cluster	31
Tabel 4. 7 Tabel karakteristik berdasarkan RFM.....	34

DAFTAR SIMBOL

Simbol	Nama
	Terminator Simbol untuk merepresentasikan awal dan akhir suatu proses.
	Flow Simbol yang digunakan untuk menggambarkan arah suatu proses yang sedang berjalan.
	Process Simbol yang menunjukkan proses yang dilakukan oleh komputer.
	Decision Simbol yang digunakan untuk memutuskan sesuatu proses berdasarkan kondisi tertentu.

DAFTAR LAMPIRAN

Lampiran 1. Surat Permohonan Izin Riset	52
Lampiran 3. Source Code Python.....	53
Lampiran 4. Source Code Streamlit	61
Lampiran 5. Hasil Turnitin	75