

The background of the slide features a dark blue gradient. Overlaid on it are several abstract elements: a circular inset in the bottom-left corner showing a close-up of a printed circuit board; a large, semi-transparent white circle in the center; and three overlapping triangles in the top-left corner, colored blue, green, and light blue.

Asmiyeni Islamiati

Marketing Campaign Customer Analysis using Machine Learning

DATASET



INTRODUCTION





Overview

Sebuah perusahaan dapat berkembang dengan pesat saat mengetahui perilaku customer personality nya, sehingga dapat memberikan layanan serta manfaat lebih baik kepada customers yang berpotensi menjadi loyal customers. Dengan mengolah data historical marketing campaign guna menaikkan performa dan menyasar customers yang tepat agar dapat bertransaksi di platform perusahaan, dari insight data tersebut fokus kita adalah membuat sebuah model prediksi kluster sehingga memudahkan perusahaan dalam membuat keputusan



OBJECTIVE

- 01 Melakukan eksplorasi data historis marketing campaign untuk memahami distribusi, outliers, dan pola-pola penting serta visualisasi data untuk mengidentifikasi tren dan hubungan antar variabel.
- 02 Membuat dan memilih fitur-fitur relevan, menggunakan teknik transformasi dan standarisasi data, serta metode statistik untuk mengevaluasi pentingnya setiap fitur.
- 03 Mengimplementasikan algoritma clustering seperti K-Means dan mengevaluasi hasil clustering dengan metrik seperti silhouette score untuk memastikan kualitas dan akurasi klaster yang dihasilkan.



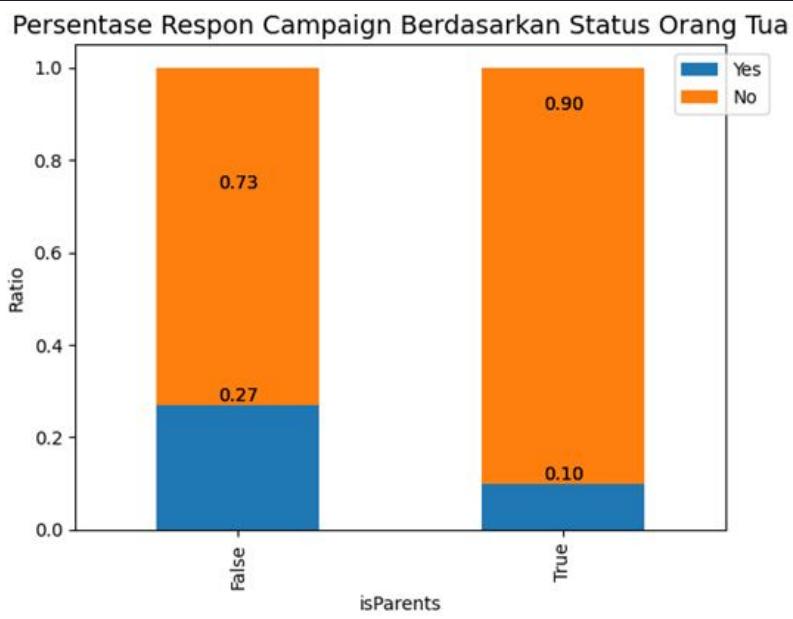
Eksploration Data Analysis & Insight

Di bawah terdapat beberapa variabel tambahan yang telah dilakukan dalam proses feature engineering.

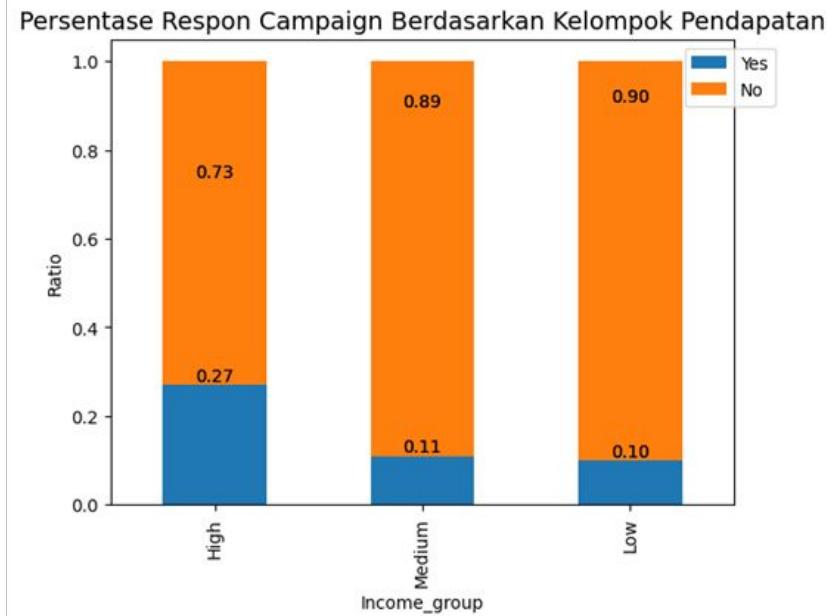
- Variabel **Conversion_Rate**, yang merupakan hasil bagi antara '**Response**' terhadap kampanye pemasaran dengan jumlah kunjungan ke situs web dalam sebulan '**NumWebVisitsMonth**'. Variabel ini memberikan gambaran tentang efektivitas kampanye dalam menghasilkan respons dari pengunjung situs.
- variabel **Age** yang diekstrak dari tahun kelahiran pelanggan. usia termuda pelanggan adalah 28 tahun, sementara usia tertua adalah 131 tahun. Untuk mempermudah analisis, kita membagi usia pelanggan ke dalam tiga kategori: "muda" (di bawah 40 tahun), "dewasa" (antara 40 dan 65 tahun), dan "tua" (di atas 65 tahun).

Conversion_Rate	Age	Children	isParents	Age_group	Income_group
0.142857	67	0	False	Lansia	Medium
0.000000	70	2	True	Lansia	Medium
0.000000	59	0	False	Dewasa	High
0.000000	40	1	True	Dewasa	Low
0.000000	43	1	True	Dewasa	Medium

- Variabel **Children** didapatkan dari jumlah anak remaja **Teenhome** dan jumlah anak kecil **Kidhome**. Variabel **isParents** menandakan apakah seorang pelanggan memiliki anak atau tidak.
- variabel **Income_group**, yang mengkategorikan tingkat pendapatan pelanggan menjadi tiga kelompok: "rendah", "menengah", dan "tinggi". Proses ini dilakukan dengan membagi rentang pendapatan dari yang terendah (1730000.0) hingga tertinggi (666666000.0) ke dalam tiga kategori berdasarkan quartile.



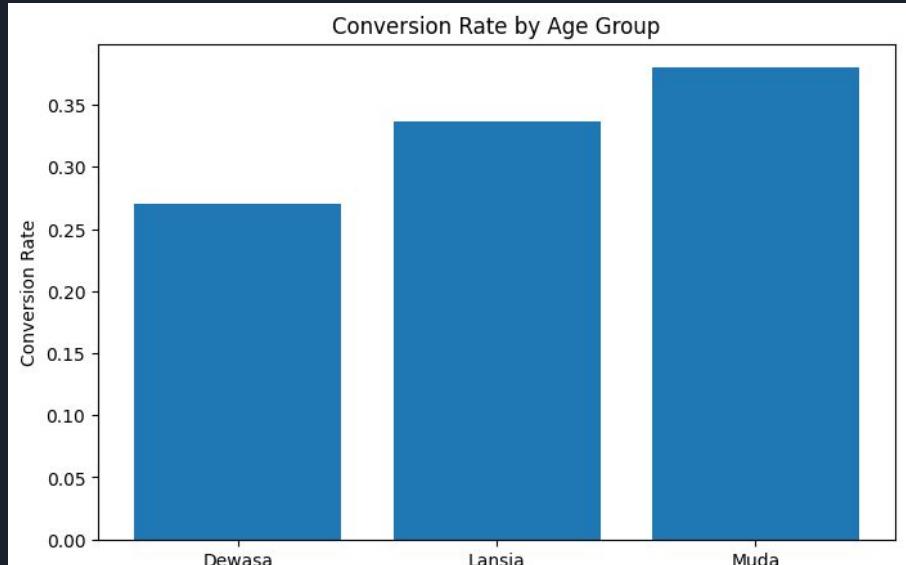
Kehadiran anak-anak di rumah mempengaruhi keputusan pelanggan untuk merespons kampanye pemasaran. Secara keseluruhan data, customer yang merespon campaign **didominasi** oleh mereka yang tidak memiliki anak (27%). Karena pola ini menunjukkan tingkat respons yang lebih tinggi, pesan pemasaran dapat disesuaikan bagi segmen ini seperti pesan yang menekankan pada kebebasan, gaya hidup aktif, atau fleksibilitas dapat lebih menarik bagi pelanggan yang bukan orangtua.



Tingkat pendapatan customers berhubungan dengan keputusan mereka untuk merespons kampanye pemasaran. Oleh karena itu **direkomendasikan** untuk mengutamakan campaign pada customers dengan latar belakang pekerjaan yang berpenghasilan tinggi di atas 68522000.

Diagram "Conversion Rate by Age Group di samping menunjukkan adanya tingkat konversi pelanggan yang berbeda berdasarkan kelompok usia. Dari diagram tersebut, dapat dilihat bahwa:

- **Kelompok Dewasa** memiliki tingkat konversi di antara 0,25 dan 0,3. Hal tersebut menunjukkan bahwa konversi dalam kelompok usia ini cukup stabil namun relatif lebih rendah dibandingkan kelompok lainnya.
- **Kelompok Muda** menunjukkan tingkat konversi yang lebih tinggi, berada di atas 0,35 namun di bawah 0,4. Hal ini menandakan bahwa pelanggan muda lebih cenderung melakukan konversi dibandingkan kelompok usia lainnya.
- **Kelompok Lansia** memiliki tingkat konversi yang berkisar antara 0,3 dan 0,35, sedikit lebih tinggi dibandingkan kelompok dewasa namun lebih rendah dibandingkan kelompok muda.





Feature Selection

Features:

```
Index(['ID', 'Year_Birth', 'Education', 'Income',
'Kidhome', 'Teenhome', 'Dt_Customer', 'Recency',
'MntCoke', 'MntFruits', 'MntMeatProducts',
'MntFishProducts', 'MntSweetProducts', 'MntGoldProds',
    'NumDealsPurchases', 'NumWebPurchases',
'NumCatalogPurchases', 'NumStorePurchases',
    'NumWebVisitsMonth', 'AcceptedCmp3', 'AcceptedCmp4',
'AcceptedCmp5', 'AcceptedCmp1',
    'AcceptedCmp2', 'Complain', 'Z_CostContact',
'Z_Revenue', 'Response', 'Conversion_Rate',
    'Total_Day_Membership', 'Total_Years_Membership',
'Age', 'Children', 'isParents', 'Age_group',
    'Income_group', 'total_spent', 'total_purchases',
'Total_Accepted_Cmp', 'Marital_Status_Bertunangan',
    'Marital_Status_Cerai', 'Marital_Status_Duda',
'Marital_Status_Janda', 'Marital_Status_Lajang',
    'Marital_Status_Menikah'],
      dtype='object')
```

Analisis segmentasi customer akan menggunakan **LRFM**, yakni:

- Length (L) : Mengukur durasi pelanggan sejak pertama kali berbelanja dalam sebuah market
- Recency (R) : Mengukur waktu sejak transaksi terakhir pelanggan
- Frequency (F): Mengukur seberapa sering pelanggan melakukan transaksi dalam periode waktu tertentu.
- Monetary (M): Mengukur jumlah uang yang dihabiskan pelanggan selama periode waktu tertentu.

Fitur yang sesuai dalam metode segmentasi tersebut adalah:

L : Total_Day_Membership

R : Recency

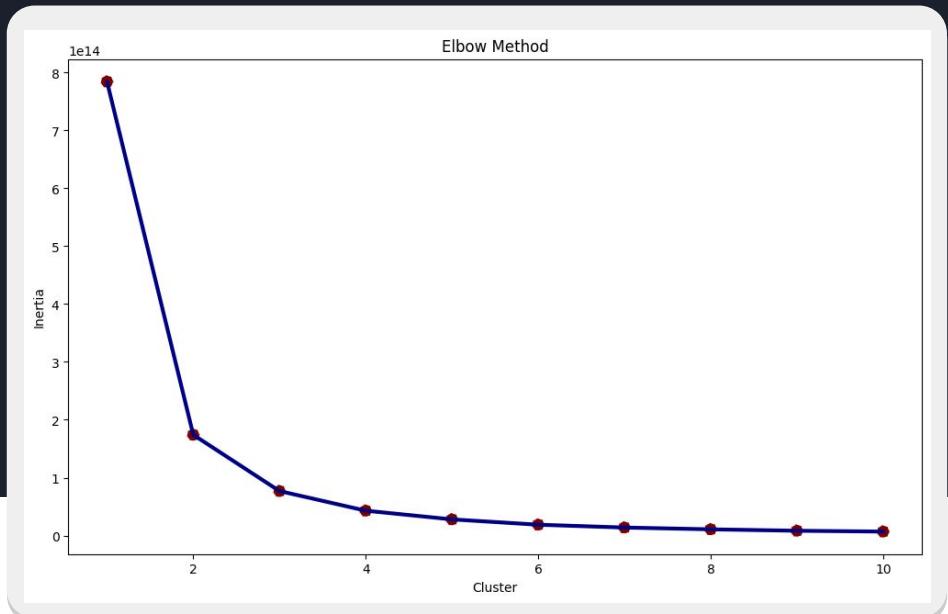
F : total_purchases

M : total_spent

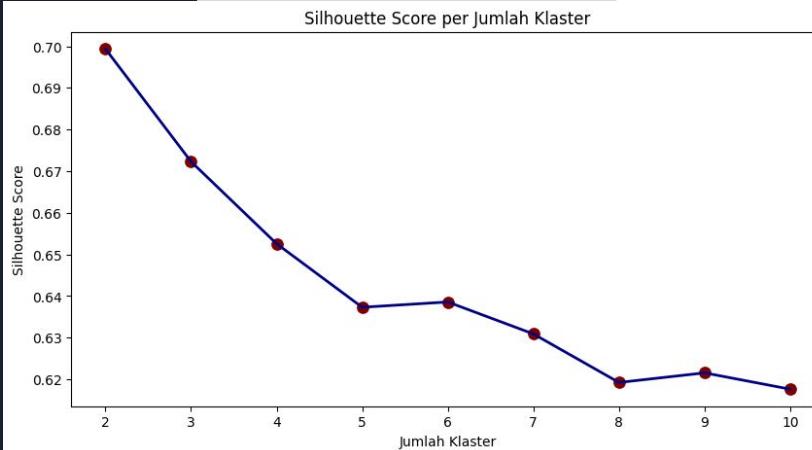


Clustering using Machine Learning

Setelah memetakan nilai-nilai inertia terhadap jumlah cluster, kita mengamati penurunan tajam pada inertia dari 1 ke 2 cluster. Penurunan signifikan kemudian berlanjut tetapi mulai melambat setelah 3 cluster. Dari nilai inertia yang didapatkan, terlihat jelas bahwa jumlah cluster yang optimal adalah 3.



Berdasarkan grafik Silhouette Score, jumlah cluster optimal adalah 3 dengan nilai Silhouette Score sebesar 0,67. Nilai Silhouette Score ini mengindikasikan bahwa tiga cluster memberikan pemisahan yang cukup baik di antara data, dengan jarak yang jelas antara cluster dan tingkat kepadatan yang tinggi di dalam cluster.



Dengan demikian, menggunakan tiga cluster akan memungkinkan analisis segmentasi yang lebih akurat dan pengelompokan pelanggan yang lebih efektif, membantu dalam strategi pemasaran yang lebih tepat sasaran.

Insight of Cluster

Cluster 0: Low Value Customer

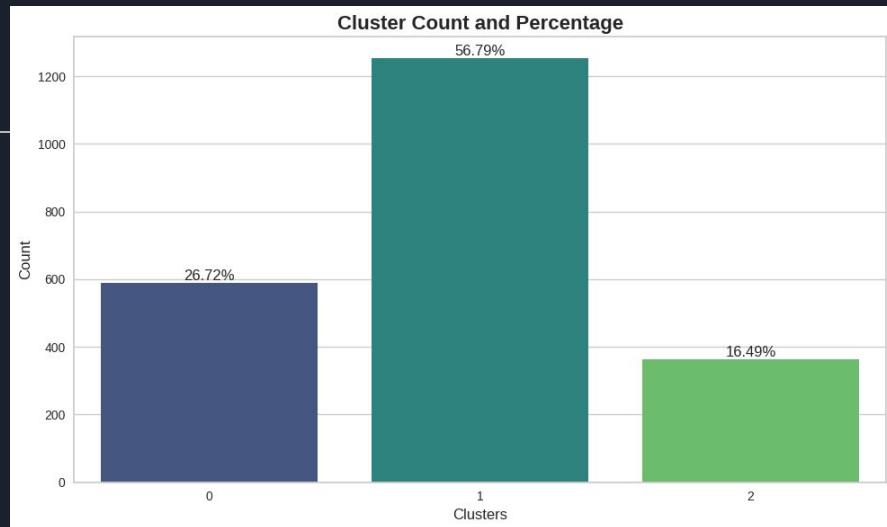
590 Pelanggan (26.72%)

Cluster 1: High Value Customer

1254 Pelanggan (56.79%)

Cluster 2: Medium Value Customers

364 Pelanggan (16.49%)





Cluster 0: Low Value Customer

590 Pelanggan (26.72%)

Cluster ini terdiri dari pelanggan yang menunjukkan nilai terendah berdasarkan hasil analisis clustering. Pelanggan dalam cluster ini memiliki masa keanggotaan yang lama dan cukup sering kembali, namun memiliki jumlah pembelian yang sedikit, frekuensi yang rendah, dan pengeluaran total yang minimal. Strategi yang dapat diambil untuk kelompok ini termasuk program loyalitas atau diskon untuk mendorong peningkatan pembelian

Cluster 1: High Value Customer

1254 Pelanggan (56.79%)

Cluster ini berisi pelanggan yang sangat bernilai, dengan masa keanggotaan yang panjang, sering kembali, melakukan repeat order, serta mengeluarkan uang dalam jumlah besar. Untuk kelompok ini, penawaran eksklusif, layanan prioritas, dan program penghargaan dapat membantu mempertahankan dan meningkatkan nilai mereka lebih lanjut.

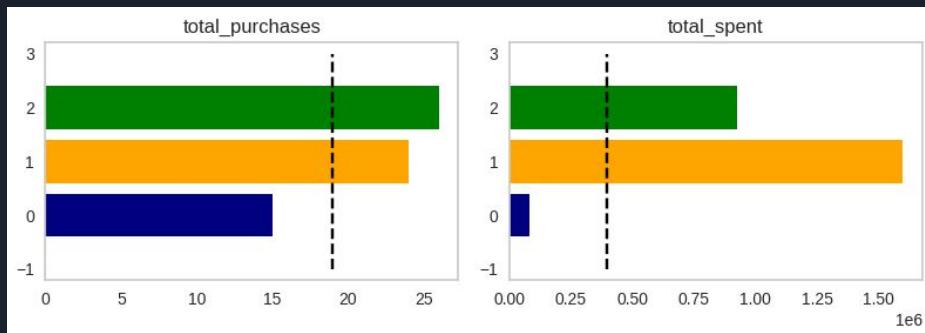
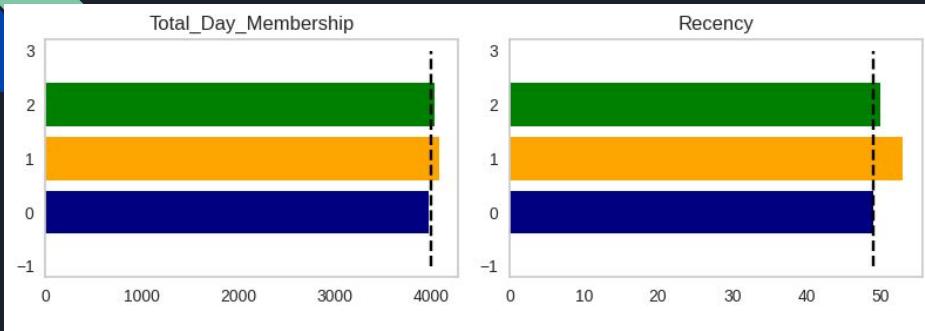
Cluster 2: Medium Value Customers

364 Pelanggan (16.49%)

Pelanggan dalam cluster ini berada di antara dua kelompok sebelumnya. Mereka menunjukkan aktivitas pembelian yang stabil dengan pengeluaran yang moderat. Pelanggan dalam cluster ini juga memiliki masa keanggotaan yang panjang dan sering kembali, namun mereka mirip dengan Cluster 1 dalam hal perilaku pembelian dan pengeluaran.



Business Recommendation



- Cluster 0: Pelanggan dengan keanggotaan panjang (4000 hari) dan baru-baru ini aktif (recency 50), namun dengan sedikit pembelian (kurang dari 20) dan pengeluaran sangat sedikit (kurang dari 0,25). Mereka dianggap tidak cukup menguntungkan.
- Cluster 1: Pelanggan dengan keanggotaan sangat panjang (lebih dari 4000 hari) dan baru-baru ini aktif (recency lebih dari 50), melakukan banyak pembelian (20-25) dan pengeluaran tinggi (lebih dari 1,5). Mereka sangat berharga dan menguntungkan.
- Cluster 2: Pelanggan dengan keanggotaan panjang (4000 hari) dan baru-baru ini aktif (recency lebih dari 50), melakukan banyak pembelian (lebih dari 25) dan pengeluaran signifikan (1). Mereka aktif dan berpotensi menguntungkan, meskipun tidak sebesar Cluster 1.



Cluster 0: Low Value Customer

Dengan total pengeluaran sebesar Rp. 190.705.000, diperlukan strategi untuk lebih memfokuskan kampanye dalam meningkatkan frekuensi pembelian dan nilai transaksi setiap pelanggan. Hal tersebut dapat membuat pelanggan lebih meningkatkan total pengeluarannya. Strategi ini meliputi:

1. Menyajikan program loyalitas yang menggiatkan bagi pelanggan dalam kluster ini untuk meningkatkan retensi dan mendorong pembelian berulang.
 2. Meningkatkan nilai transaksi dengan strategi up selling dan cross-selling, yang dapat dilakukan melalui rekomendasi produk terkait atau penawaran bundle.
 3. Memberikan informasi yang lebih detail mengenai produk atau layanan pelanggan sebagai upaya peningkatan kesadaran dan minat pelanggan dalam melakukan pembelian.
-

Cluster 1: High Value Customer

Dengan Total Pengeluaran sebesar

Rp. 599.195.000, direkomendasikan untuk lebih berfokus untuk mempertahankan loyalitas pelanggan.

Hal tersebut dapat dilakukan dengan cara:

1. Menawarkan pengalaman pelanggan yang eksklusif (VIP), seperti layanan pelanggan 24/7 atau program pengiriman prioritas (member) untuk mempertahankan loyalitas mereka.
2. Menawarkan paket premium (bundling) atau produk eksklusif yang sesuai dengan kebutuhan dan preferensi mereka.
3. Memanfaatkan data pelanggan untuk menyesuaikan penawaran dan komunikasi, sehingga mereka merasa dihargai dan diperhatikan.

Cluster 2: Medium Value Customer

Total Pengeluaran sebesar Rp. 538.350.000

menunjukkan bahwa ada potensi untuk meningkatkan nilai transaksi dengan lebih memfokuskan kampanye pada kualitas layanan dan menawarkan insentif yang sesuai. Hal tersebut dapat direalisasikan dengan:

1. Mendorong peningkatan nilai transaksi dengan menawarkan insentif (affiliate) untuk pembelian dalam jumlah besar atau untuk produk tambahan.
2. Menyajikan konten yang bernilai dan relevan kepada pelanggan untuk memperkuat keterlibatan dan memperkuat koneksi mereka.
3. Memberikan penawaran program pengembalian yang mudah dan transparan untuk meningkatkan kepercayaan dan kenyamanan pelanggan dalam berbelanja.

Thank you!

Find me more on:

linkedin.com/in/asmiyenislamiati

github.com/txtmeyen

instagram.com/asmiyenislamiati

