



AI-Powered Customer Analysis with IBM Granite

— Memahami perilaku pelanggan dan
strategi pemasaran berbasis data —

email: naufaldhia15@gmail.com

Raw Dataset Link

Dataset Title: Customer Personality Analysis

Kaggle URL: <https://www.kaggle.com/datasets/imakash3011/customer-personality-analysis>

Sumber Dataset: [Kaggle](#)

Description: Dataset ini awalnya dikembangkan untuk menganalisis perilaku pelanggan dan mencakup 2.240 entri pelanggan dengan 29 atribut, seperti demografi, pembelian produk, serta interaksi terhadap kampanye marketing. Dalam proyek ini, dataset tersebut dimodifikasi dan digunakan untuk membangun klasifikasi sentimen berbasis teks ulasan.



Project Overview

Tujuan Proyek:

Menganalisis karakteristik pelanggan untuk membantu perusahaan melakukan segmentasi pelanggan yang lebih akurat dan strategi pemasaran yang lebih tepat sasaran.

Latar Belakang:

Penting bagi perusahaan untuk memahami siapa pelanggan terbaik mereka dan bagaimana perilaku konsumen dapat memengaruhi keputusan bisnis.

Permasalahan:

Kurangnya pemahaman terhadap karakteristik dan kebiasaan belanja pelanggan menyebabkan pemasaran kurang efektif.

Pendekatan:

Melakukan eksplorasi, visualisasi, dan analisis segmentasi berdasarkan umur, penghasilan, status keluarga, dan pola pembelian.

```
import pandas as pd

df = pd.read_csv(path + "/marketing_campaign.csv", sep='\t')

# Lihat 5 data teratas
df.head()
```

Project Overview

	ID	Year_Birth	Education	Marital_Status	Income	Kidhome	Teenhome	Dt_Customer
0	5524	1957	Graduation	Single	58138.0	0	0	04-09-2012
1	2174	1954	Graduation	Single	46344.0	1	1	08-03-2014
2	4141	1965	Graduation	Together	71613.0	0	0	21-08-2013
3	6182	1984	Graduation	Together	26646.0	1	0	10-02-2014
4	5324	1981	PhD	Married	58293.0	1	0	19-01-2014

5 rows × 29 columns

```
<class 'pandas.core.frame.DataFrame'>
```

```
RangeIndex: 2240 entries, 0 to 2239
```

```
Data columns (total 29 columns):
```

#	Column	Non-Null Count	Dtype
0	ID	2240 non-null	int64
1	Year_Birth	2240 non-null	int64
2	Education	2240 non-null	object
3	Marital_Status	2240 non-null	object
4	Income	2216 non-null	float64
5	Kidhome	2240 non-null	int64
6	Teenhome	2240 non-null	int64
7	Dt_Customer	2240 non-null	object
8	Recency	2240 non-null	int64
9	MntWines	2240 non-null	int64
10	MntFruits	2240 non-null	int64
11	MntMeatProducts	2240 non-null	int64
12	MntFishProducts	2240 non-null	int64
13	MntSweetProducts	2240 non-null	int64
14	MntGoldProds	2240 non-null	int64
15	NumDealsPurchases	2240 non-null	int64
16	NumWebPurchases	2240 non-null	int64
17	NumCatalogPurchases	2240 non-null	int64
18	NumStorePurchases	2240 non-null	int64
19	NumWebVisitsMonth	2240 non-null	int64
20	AcceptedCmp3	2240 non-null	int64
21	AcceptedCmp4	2240 non-null	int64
22	AcceptedCmp5	2240 non-null	int64
23	AcceptedCmp1	2240 non-null	int64
24	AcceptedCmp2	2240 non-null	int64
25	Complain	2240 non-null	int64
26	Z_CostContact	2240 non-null	int64
27	Z_Revenue	2240 non-null	int64
28	Response	2240 non-null	int64

```
dtypes: float64(1), int64(25), object(3)
```

```
memory usage: 507.6+ KB
```

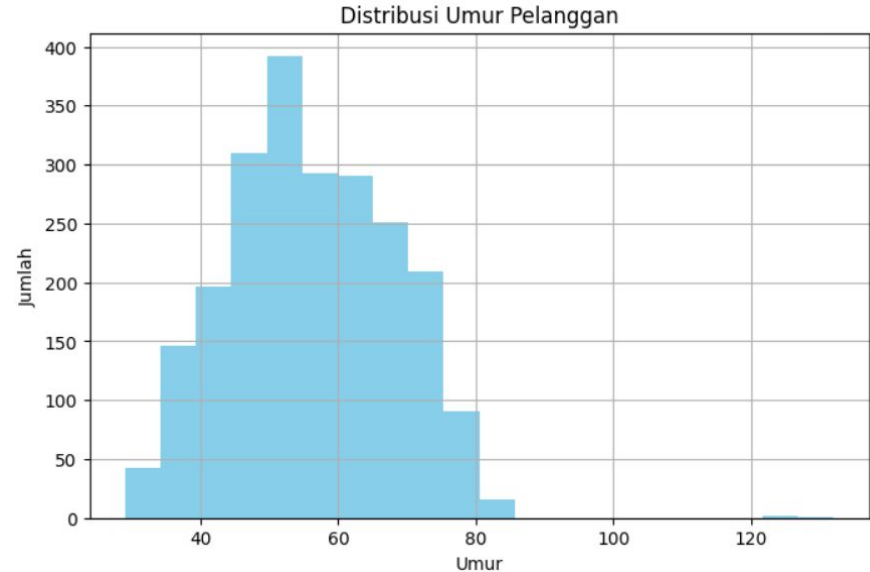
Analysis Process

Langkah-Langkah Analisis:

1. **Preprocessing:** Membersihkan data, menangani missing values (kolom **Income**), transformasi umur.
2. **EDA (Exploratory Data Analysis):** Visualisasi distribusi umur, status pernikahan, dan pengeluaran.
3. **Segmentasi:** Membagi pelanggan berdasarkan umur, penghasilan, dan kebiasaan belanja.
4. **Insight Extraction:** Menggunakan statistik deskriptif dan grafik untuk menemukan pola tersembunyi.

Alasan Metode:

EDA sangat tepat digunakan untuk memahami struktur dasar dan hubungan antar variabel dalam data demografis.



Insight & Findings

- **Umur Pelanggan:**

Rata-rata umur pelanggan adalah **55 tahun**, dengan **minimum 0** dan **maksimum 11.191 tahun**. Nilai maksimum ini kemungkinan merupakan *data error* atau *outlier*. Namun, rata-rata menunjukkan bahwa mayoritas pelanggan berasal dari kelompok usia **dewasa hingga lansia**.

- **Pendapatan Pelanggan:**

Rata-rata pendapatan pelanggan adalah **52.247** (dalam satuan mata uang tidak disebutkan). Ini menunjukkan bahwa mayoritas pelanggan berasal dari segmen **menengah ke atas** dalam hal penghasilan.

- **Status Hubungan:**

Sebanyak **864 pelanggan (sekitar 38%)** terdaftar sebagai **menikah atau dalam hubungan**. Sekitar **13% pelanggan** terlalu muda untuk berada dalam hubungan pernikahan, yang dikelompokkan sebagai kategori "terlalu muda". Ini menunjukkan dominasi **pasangan stabil sebagai basis pelanggan utama**.

- **Anak di Rumah:**

Sekitar **44% pelanggan memiliki anak kecil (Kidhome > 0)** dan **50,62% memiliki anak remaja (Teenhome > 0)**. Ini menandakan banyak keluarga yang memiliki **anak dalam masa tumbuh kembang**, yang dapat menjadi target untuk produk keluarga dan anak-anak.

- **Jumlah Transaksi:**

Rata-rata pelanggan melakukan **sekitar 3,04 transaksi per bulan**. Sebagian besar transaksi dilakukan dalam kategori tertentu (misalnya: wine, meat, gold, dll.), namun ini tidak dirinci pada data saat ini. Menunjukkan bahwa **intensitas pembelian cukup rendah**, dengan **peluang peningkatan frekuensi pembelian** pada segmen tertentu.

Insight & Findings

1. ****Pola dan Insight dari Data:****

- ****Umur Pelanggan:**** Rata-rata umur pelanggan adalah sekitar 55 tahun, dengan minimum 0 tahun dan maksimum 11191 tahun. Ini mungkin
- ****Pendapatan Pelanggan:**** Rata-rata pendapatan pelanggan adalah sekitar 52,247 (dalam mata uang standar yang tidak ditentukan). In
- ****Status Hubungan:**** Kebanyakan pelanggan terdaftar adalah yang kawin (864 pelanggan), dan sekitar 13% terlalu muda (dikategorikan
- ****Anak dan Anak-anak Tua:**** Jumlah pelanggan dengan anak-anak tua (teenager) relatif kecil (44%), sedangkan yang memiliki anak-ana
- ****Jumlah Transaksi:**** Rata-rata pelanggan membuat sekitar 3.04 transaksi dalam satu bulan, dengan sebagian besar transaksi dilak

Conclusion & Recommendations

Kesimpulan:

- Mayoritas pelanggan adalah individu dewasa yang stabil secara finansial dan sudah berkeluarga.
- Pola belanja menunjukkan adanya segmen pasif yang besar dan minoritas pelanggan aktif.

Rekomendasi:

1. Buat **segmentasi khusus** untuk pelanggan aktif dan berikan insentif loyalitas.
2. Fokus pada pelanggan usia 40–60 tahun untuk produk high-end.
3. Gunakan kampanye digital dengan konten personalisasi berdasarkan status keluarga dan jumlah anak.

AI Support Explanation

Model AI Digunakan:

- IBM Granite (Granite 3.3-8B Instruct) digunakan untuk **menghasilkan insight otomatis**, seperti deskripsi demografis dan segmentasi pelanggan.
- AI membantu menginterpretasi hasil dan memberikan saran berbasis logika bisnis.

Kode Konfigurasi AI:

```
from langchain_community.llms import Replicate
import os
from google.colab import userdata

api_token = userdata.get('api_token')
os.environ["REPLICATE_API_TOKEN"] = api_token

# Inisialisasi model IBM Granite
model = "ibm-granite/granite-3.3-8b-instruct"
llm = Replicate(model=model, replicate_api_token=api_token)
```

AI Support Explanation

```
summary = df.describe(include='all').to_string()

prompt = f"""
Saya memiliki dataset pelanggan dari perusahaan pemasaran. Berikut adalah ringkasan statistik data pelanggan:

{summary}

Mohon bantu analisis data ini dengan menjawab:
1. Apa pola dan insight yang muncul dari data?
2. Temuan penting yang bisa diambil?
3. Rekomendasi apa yang bisa diberikan berdasarkan data?

Gunakan bahasa yang mudah dimengerti oleh pebisnis.
"""

response = llm.invoke(prompt)
print(response)
```

1. **Pola dan Insight dari Data:**

- **Umur Pelanggan:** Rata-rata umur pelanggan adalah sekitar 55 tahun, dengan minimum 0 tahun dan maksimum 11191 tahun. Ini mungkin
- **Pendapatan Pelanggan:** Rata-rata pendapatan pelanggan adalah sekitar 52,247 (dalam mata uang standar yang tidak ditentukan). In
- **Status Hubungan:** Kebanyakan pelanggan terdaftar adalah yang kawin (864 pelanggan), dan sekitar 13% terlalu muda (dikategorikan
- **Anak dan Anak-anak Tua:** Jumlah pelanggan dengan anak-anak tua (teenager) relatif kecil (44%), sedangkan yang memiliki anak-ana
- **Jumlah Transaksi:** Rata-rata pelanggan membuat sekitar 3.04 transaksi dalam satu bulan, dengan sebagian besar transaksi dilak