

Homework

Final Project - Stage 1



1. Descriptive Statistics

data.info()

```
<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
```

data.nunique()

```
ID                    2240
Year_Birth            59
Education             5
Marital_Status        8
Income               1974
Kidhome                3
Teenhome               3
Dt_Customer           663
Recency              100
MntWines             776
MntFruits            158
MntMeatProducts      558
MntFishProducts      182
MntSweetProducts     177
MntGoldProds         213
NumDealsPurchases    15
NumWebPurchases      15
NumCatalogPurchases  14
NumStorePurchases    14
NumWebVisitsMonth    16
AcceptedCmp3          2
AcceptedCmp4          2
AcceptedCmp5          2
AcceptedCmp1          2
AcceptedCmp2          2
Complain              2
Z_CostContact         1
Z_Revenue             1
Response              2
dtype: int64
```

data.duplicated()

duplicated rows percentage			
0	182	8.12	
	Education	Marital_Status	Income
1501	2n Cycle	Divorced	31163.0
1107	2n Cycle	Married	15315.0
1171	2n Cycle	Married	20130.0
383	2n Cycle	Married	35688.0
2015	2n Cycle	Married	37284.0
...
1218	PhD	Together	68682.0
2062	PhD	Together	70038.0
1630	PhD	Together	73059.0
1959	PhD	Widow	56551.0
2217	PhD	Widow	82032.0

182 rows x 3 columns

1. Descriptive Statistics

data.describe()

a. Numerical Features

	count	mean	std	min	10%	20%	30%	40%	50%	60%	70%	80%	90%	max
Year_Birth	2240.0	1968.805804	11.984069	1893.0	1952.0	1957.0	1962.0	1966.0	1970.0	1973.0	1976.0	1979.0	1984.0	1996.0
Income	2216.0	52247.251354	25173.076661	1730.0	24117.5	32011.0	38198.5	44529.0	51381.5	58482.0	65247.5	71819.0	79844.0	666666.0
Kidhome	2240.0	0.444196	0.538398	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	2.0
Teenhome	2240.0	0.506250	0.544538	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	1.0	2.0
Recency	2240.0	49.109375	28.962453	0.0	9.0	19.0	29.0	39.0	49.0	59.0	69.0	79.0	89.0	99.0
MntWines	2240.0	303.935714	336.597393	0.0	6.0	16.0	34.0	81.0	173.5	284.4	418.6	581.2	822.1	1493.0
MntFruits	2240.0	26.302232	39.773434	0.0	0.0	1.0	2.0	4.0	8.0	15.0	25.0	44.0	83.0	199.0
MntMeatProducts	2240.0	166.950000	225.715373	0.0	7.0	12.0	20.0	35.0	67.0	108.4	177.0	298.4	499.0	1725.0
MntFishProducts	2240.0	37.525446	54.628979	0.0	0.0	2.0	3.0	7.0	12.0	20.0	37.0	65.0	120.0	259.0
MntSweetProducts	2240.0	27.062946	41.280498	0.0	0.0	1.0	2.0	5.0	8.0	14.0	26.0	44.2	89.0	263.0
MntGoldProds	2240.0	44.021875	52.167439	0.0	3.0	6.0	11.0	17.0	24.0	34.0	46.0	73.0	122.0	362.0
NumDealsPurchases	2240.0	2.325000	1.932238	0.0	1.0	1.0	1.0	1.0	2.0	2.0	3.0	3.0	5.0	15.0
NumWebPurchases	2240.0	4.084821	2.778714	0.0	1.0	2.0	2.0	3.0	4.0	4.0	5.0	6.0	8.0	27.0
NumCatalogPurchases	2240.0	2.662054	2.923101	0.0	0.0	0.0	1.0	1.0	2.0	2.0	4.0	5.0	7.0	28.0
NumStorePurchases	2240.0	5.790179	3.250958	0.0	2.0	3.0	3.0	4.0	5.0	6.0	7.0	9.0	11.0	13.0
NumWebVisitsMonth	2240.0	5.316518	2.426645	0.0	2.0	3.0	4.0	5.0	6.0	6.0	7.0	7.0	8.0	20.0
AcceptedCmp1	2240.0	0.064286	0.245316	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
AcceptedCmp2	2240.0	0.013393	0.114976	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
AcceptedCmp3	2240.0	0.072768	0.259813	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
AcceptedCmp4	2240.0	0.074554	0.262728	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
AcceptedCmp5	2240.0	0.072768	0.259813	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
Complain	2240.0	0.009375	0.096391	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
Response	2240.0	0.149107	0.356274	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0

b. Categorical Features

	count	unique	top	freq
Education	2240	5	Graduation	1127
Marital_Status	2240	8	Married	864

c. Date time Features

	count	unique	top	freq	first	last
Dt_Customer	2240	663	2012-08-31	12	2012-07-30	2014-06-29

1. Descriptive Statistics

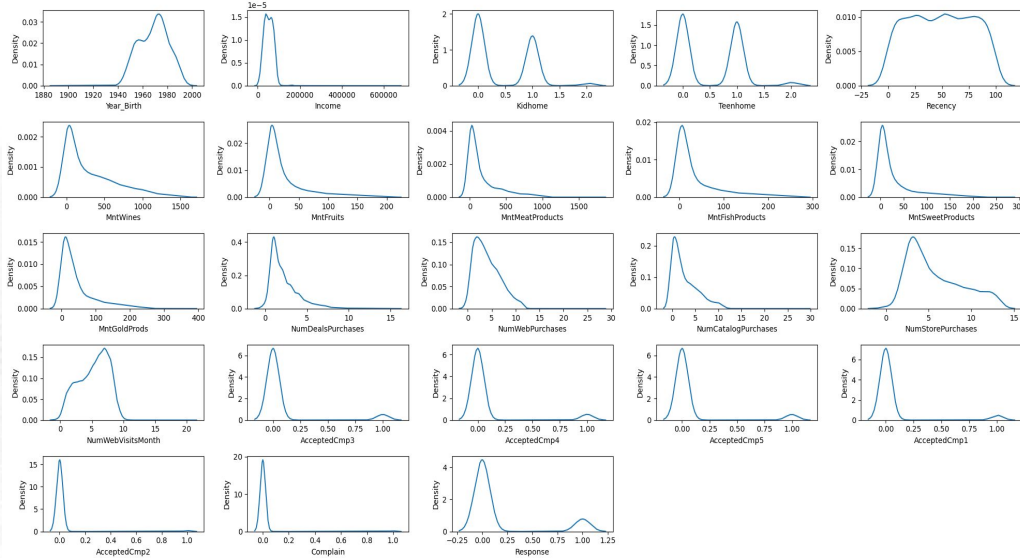
1. **Apakah ada kolom dengan tipe data kurang sesuai, atau nama kolom dan isinya kurang sesuai?**
 Feature **Dt_Customer** merupakan tanggal registrasi pelanggan dengan tipe data object. Tipe data ini tidak sesuai, sehingga perlu diubah menjadi tipe Date Time.

1. **Apakah ada kolom yang memiliki nilai kosong? Jika ada, apa saja?**
 Feature **Income** mempunyai nilai kosong karena hanya berjumlah 2216 rows.

1. **Apakah ada kolom yang memiliki nilai summary agak aneh? (min/mean/median/max/unique/top/freq)**
 - Feature **ID** merupakan identifikasi pelanggan. Berdasarkan analisis jumlah nilai unik pada feature tersebut, diketahui bahwa jumlahnya sama dengan jumlah baris dataset (2240), sehingga tidak memungkinkan untuk mengamati riwayat perjalanan pelanggan.
 - Feature **Z_CostContact** dan **Z_Revenue** hanya memiliki satu data unik, maka keduanya tidak akan memberikan hasil analisis yang signifikan.
 - Feature **Dt_Customer**, pelanggan paling terakhir melakukan registrasi di 29 Juni 2014, maka dengan asumsi saat ini adalah tahun 2014, ada keanehan pada Feature **Year Birth** dimana tahun lahir tertua ada di tahun 1893 atau usia pelanggan 121 tahun. Hal ini merupakan hal yang kurang masuk akal. Diduga terdapat kesalahan input tahun lahir oleh pelanggan/kesalahan pencatatan oleh sistem.
 - Feature **Income** memiliki keanehan karena memiliki nilai maksimum mencapai ratusan ribu (666.666), sedangkan nilai ukuran pemusatan dan penyebarannya hanya mencapai puluhan ribu. Diduga nilai ini merupakan outlier yang disebabkan karena kesalahan input atau pencatatan oleh sistem.
 - Feature **MntFishProducts**, **MntFruits**, **MntGoldProds**, **MntMeatProducts**, **MntSweetProducts**, **MntWines** memiliki keanehan dilihat dari nilai maksimumnya yang jauh dari nilai ukuran pemusatan atau ukuran penyebaran lainnya. Sehingga diduga terdapat nilai outlier pada feature tersebut.
 - Feature **Marital Status** memiliki keanehan karena memiliki 8 nilai unik. Maka pada tahap selanjutnya akan dianalisis lebih dalam setiap nilai pada feature ini.

2. Univariate Analysis

a. Numerical Features



Berdasarkan kdeplot di samping, diketahui distribusi dari feature dan target sebagai berikut:

1. Distribusi Normal:

Recency

1. Distribusi Left-Skewed (median > mean):

- Year_Birth

1. Distribusi Right-Skewed (mean > median)

- Income

- MntFishProducts, MntFruits, MntGoldProds, MntMeatProducts, MntSweetProducts, MntWines

- NumCatalogPurchases, NumDealsPurchases, NumStorePurchases, NumWebPurchases, NumWebVisitsMonth

- AcceptedCmp1, AcceptedCmp2, AcceptedCmp3, AcceptedCmp4, AcceptedCmp5 (didominasi dengan value 0)

- Response

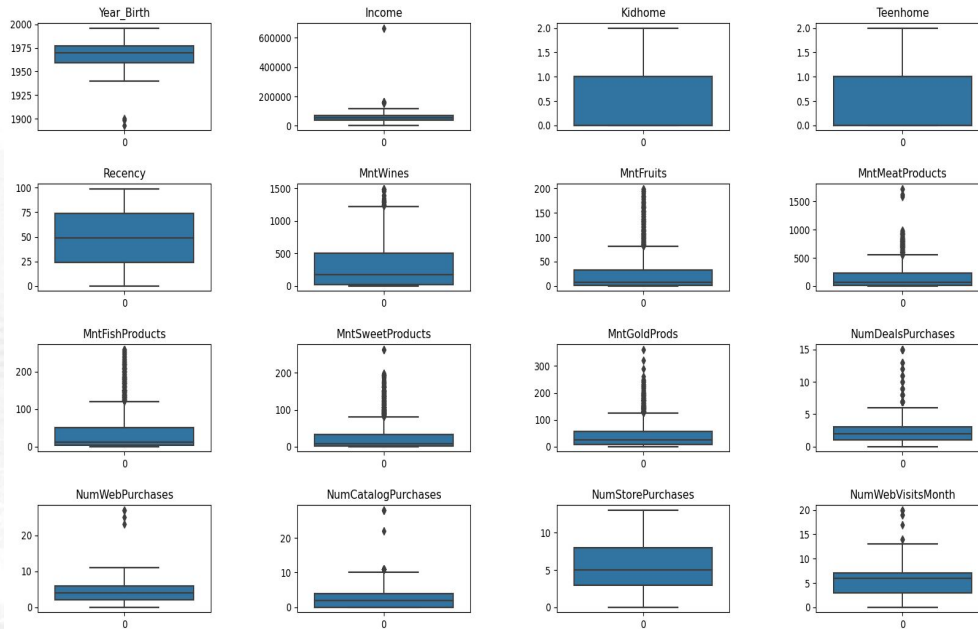
1. Distribusi Bimodal

- Kidhome

- Teenhome

2. Univariate Analysis

a. Numerical Features

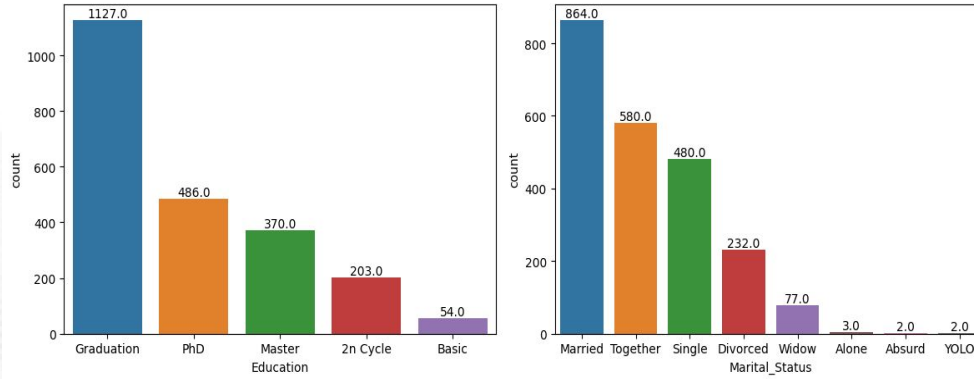


Berdasarkan boxplot di samping, dapat dilihat feature yang mengandung nilai outlier sebagai berikut:

1. Year_Birth memiliki outlier yaitu pada nilai minimumnya (di bawah tahun 1900)
2. Income memiliki nilai outlier yang cukup ekstrem yaitu pada nilai maksimumnya.
3. MntFishProducts, MntFruits, MntGoldProds, MntMeatProducts, MntSweetProducts, MntWines memiliki nilai outlier yang cukup banyak yaitu pada bagian atas boxplot (nilai tinggi)
4. NumCatalogPurchases, NumDealsPurchases, NumStorePurchases, NumWebPurchases, NumWebVisitsMonth memiliki nilai outlier yang cukup banyak yaitu pada bagian atas boxplot (nilai tinggi)

2. Univariate Analysis

b. Categorical Features

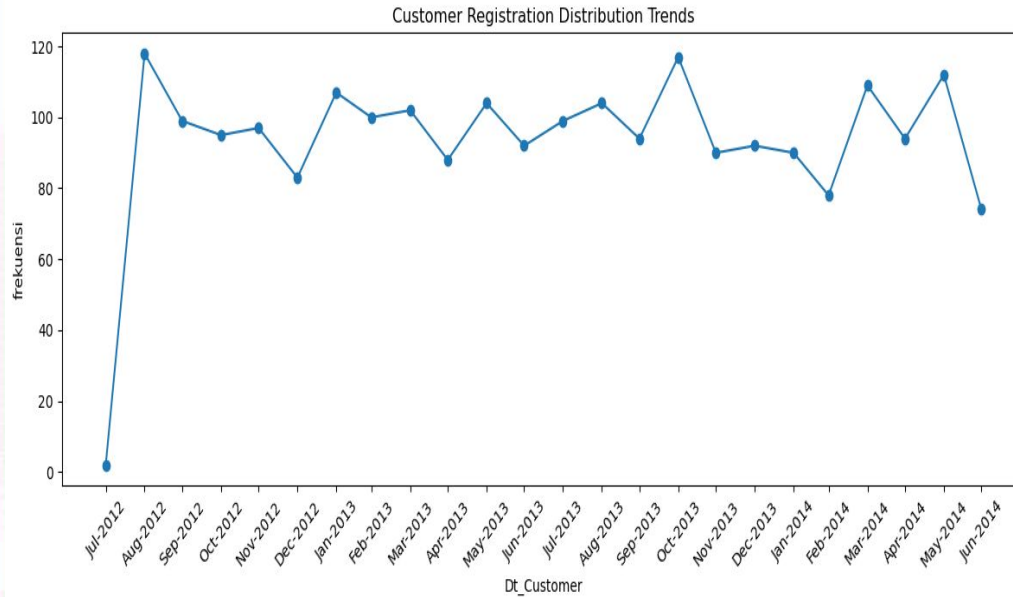


Berdasarkan grafik di samping, diketahui bahwa:

1. Pelanggan yang mempunyai level pendidikan Graduation adalah yang terbanyak, dilanjutkan dengan PnD, Master, 2n Cycle, dan Basic.
2. Pelanggan dengan status perkawinannya Married adalah yang terbanyak, dilanjutkan dengan Together, Single, Divorced, Widow, Absurd, dan YOLO.
3. Level Pendidikan Basic dan 2n Cycle perlu dipahami lebih lanjut definisinya.
4. Status Perkawinan Alone, Absurd, dan YOLO dianggap sebagai outlier yang tidak terdefinisi.

2. Univariate Analysis

c. Date Time Features



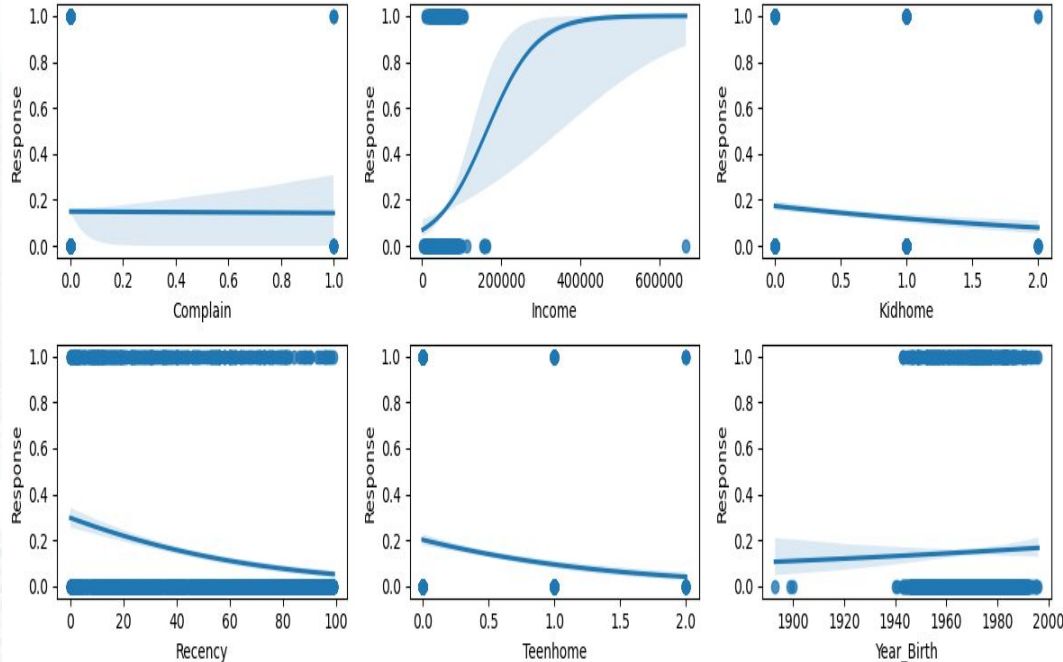
Berdasarkan grafik di samping, diketahui bahwa:

1. Sepanjang Juli-2012 hingga Juni-2014, jumlah customer yang mendaftar terbesar terjadi pada Agustus-2012 dan Oktober-2013 mencapai 120 customer baru.
2. Sedangkan Juli-2012 dan Juni-2014 adalah yang terendah.

3. Multivariate Analysis

a. Numerical Features

['Complain', 'Income', 'Kidhome', 'Recency', 'Teenhome', 'Year_Birth'] vs Response



Berdasarkan Regression Plot disamping:

1. Income:

Korelasi positif yang kuat dengan Response, menunjukkan bahwa semakin tinggi pendapatan, semakin tinggi kemungkinan pelanggan merespon.

2. Jumlah Anak (Teenhome, Kidhome):

Korelasi negatif dengan Response, menunjukkan bahwa semakin sedikit jumlah anak, semakin tinggi kemungkinan pelanggan merespon.

3. Recency:

Korelasi negatif dengan Response, menunjukkan bahwa semakin lama sejak pelanggan terakhir berinteraksi, semakin tinggi kemungkinan pelanggan merespon.

4. Year_Birth:

Korelasi negatif dengan Response, menunjukkan bahwa semakin muda usia pelanggan, semakin tinggi kemungkinan pelanggan merespon.

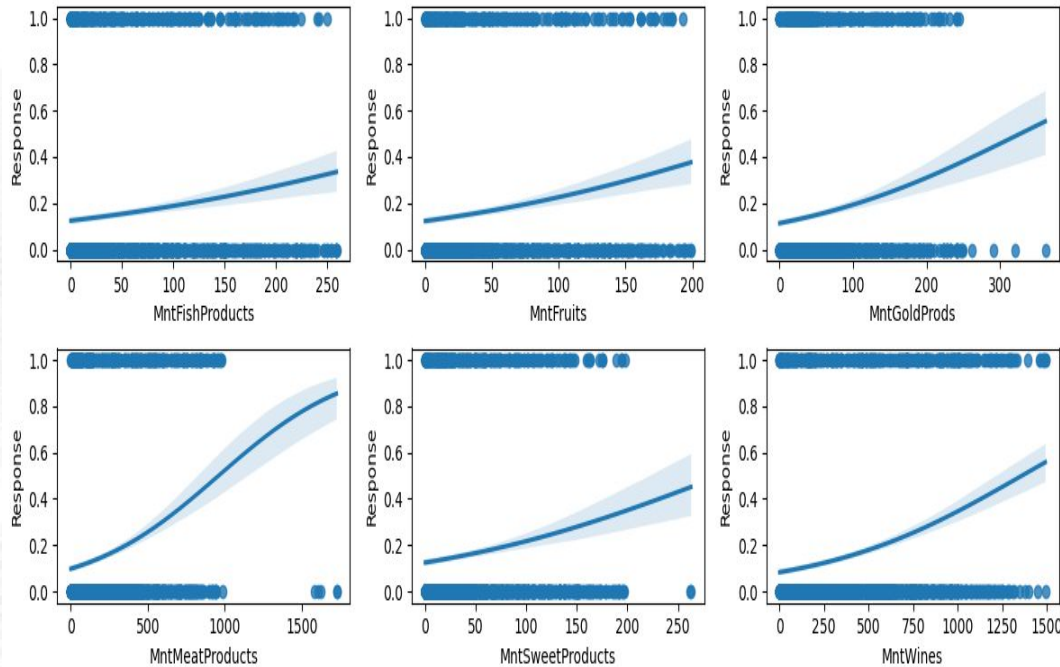
5. Complain:

Tidak terlihat adanya korelasi yang signifikan terhadap Response, ditandai dengan garis regresi yang cenderung lurus.

3. Multivariate Analysis

a. Numerical Features

['MntFishProducts', 'MntFruits', 'MntGoldProds', 'MntMeatProducts', 'MntSweetProducts', 'MntWines'] vs Response

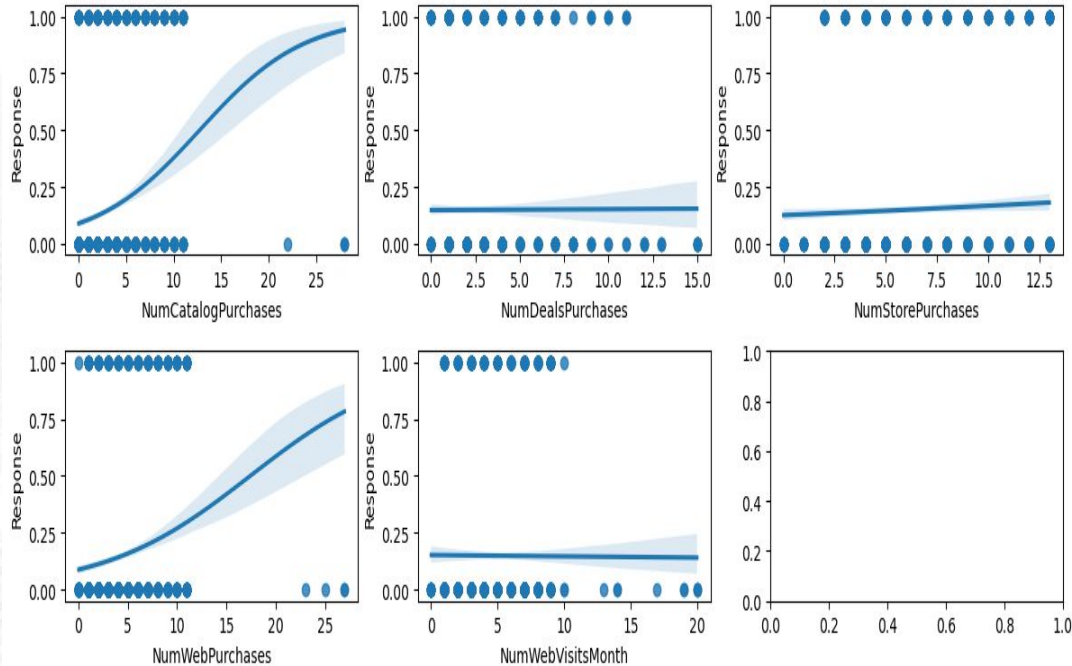


Berdasarkan regression plot di samping, diketahui bahwa semua feature amount spent product selama 2 tahun memiliki korelasi positif terhadap response dimana spending untuk produk Meat menunjukkan pola korelasi yang lebih kuat dibandingkan dengan produk lainnya.

3. Multivariate Analysis

a. Numerical Features

['NumCatalogPurchases', 'NumDealsPurchases', 'NumStorePurchases', 'NumWebPurchases', 'NumWebVisitsMonth'] vs Response

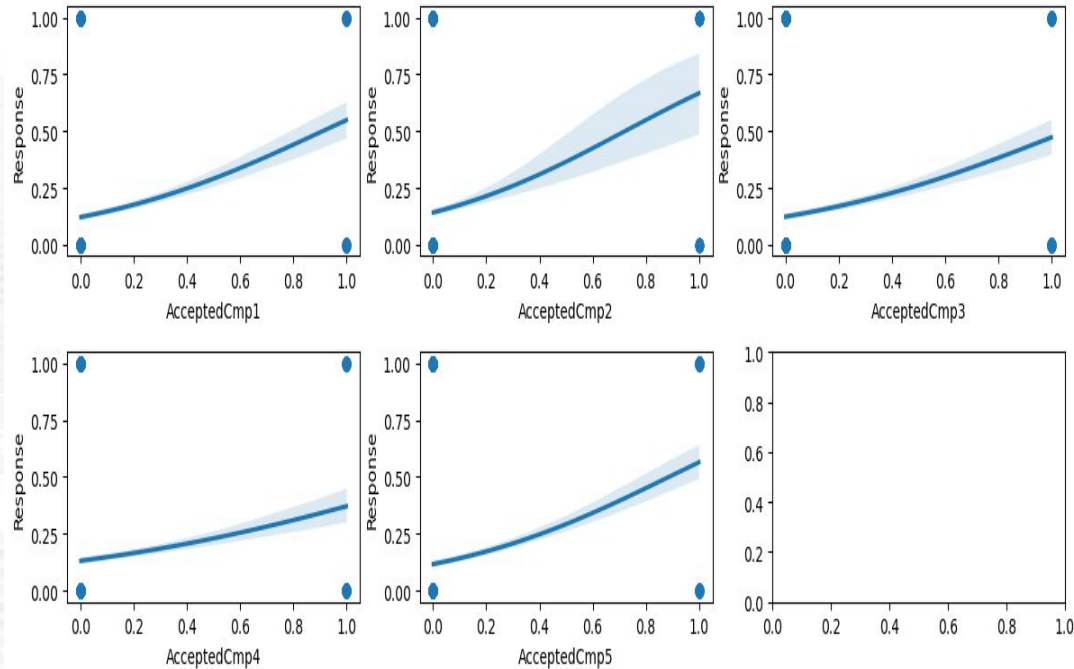


Berdasarkan regression plot di samping, diketahui bahwa pembelian yang memiliki korelasi yang kuat terhadap Response adalah pembelian melalui Katalog dan Web. Sementara pembelian dengan diskon (Deals) atau melalui Store tidak menunjukkan adanya korelasi yang signifikan terhadap Response.

3. Multivariate Analysis

a. Numerical Features

['AcceptedCmp1', 'AcceptedCmp2', 'AcceptedCmp3', 'AcceptedCmp4', 'AcceptedCmp5'] vs Response

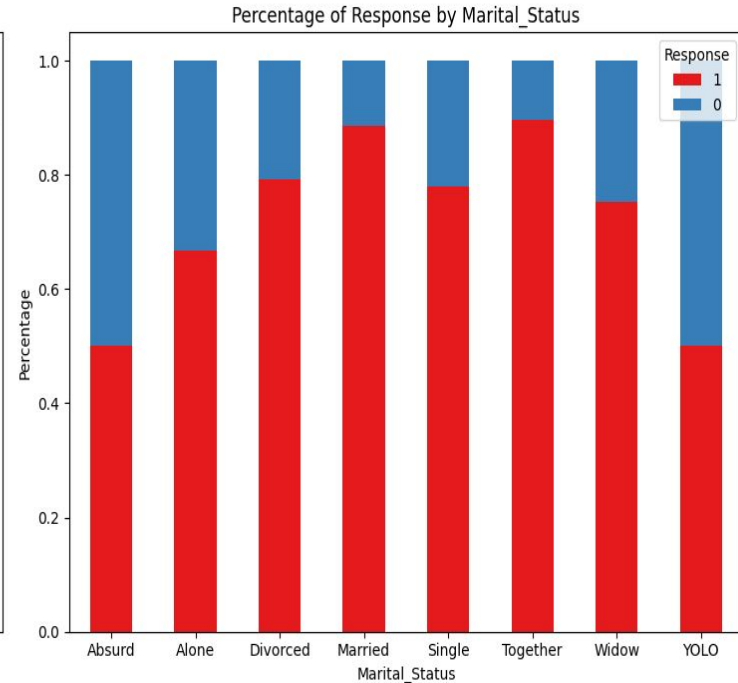
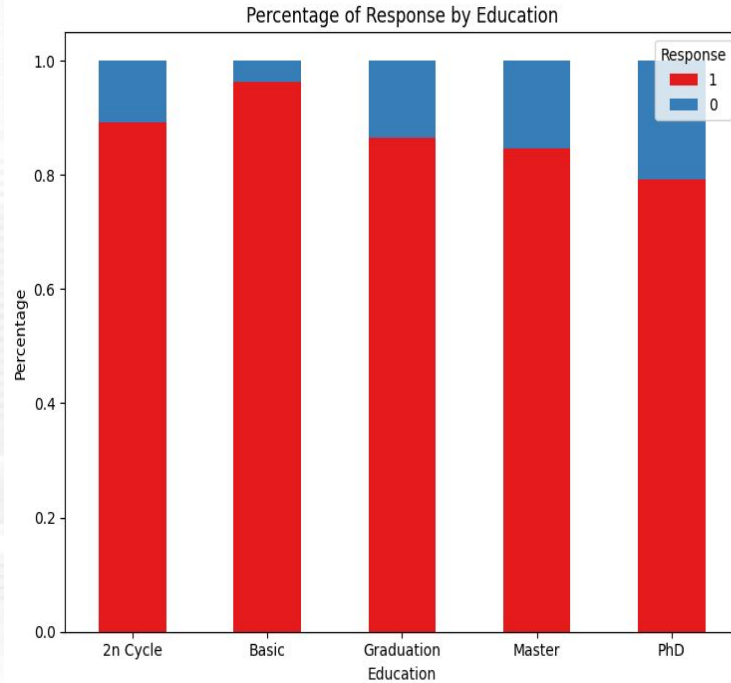


Berdasarkan regression plot di samping, diketahui bahwa kelima feature campaign memiliki korelasi positif terhadap Response.

Untuk membuktikan hasil dari semua Regression Plot, akan dibuat pula Correlation Heatmap untuk mengukur besar korelasi semua feature terhadap Response.

3. Multivariate Analysis

b. Categorical Features



3. Multivariate Analysis

b. Categorical Features

Insight Response Campaign Berdasarkan Tingkat Pendidikan:

1. **Tingkat Response Tinggi pada Tingkat Pendidikan Tinggi:**
Pelanggan dengan tingkat pendidikan PhD (20.78%) memiliki tingkat Response Campaign tertinggi, diikuti oleh Master (15.41%) dan Graduation (13.48%). Ini menunjukkan bahwa tingkat pendidikan tinggi berkorelasi positif dengan kemungkinan merespon Campaign.
1. **Varian Response pada Tingkat Pendidikan Rendah:**
Tingkat Response lebih bervariasi pada tingkat pendidikan rendah, di mana Basic (3.70%) menunjukkan Response lebih rendah dibandingkan dengan 2n Cycle (10.84%) dan Graduation (13.49%). Hal ini mungkin menunjukkan bahwa pelanggan dengan tingkat pendidikan lebih rendah cenderung memiliki tingkat Response yang lebih rendah terhadap Campaign.
1. **Pentingnya Pendidikan dalam Pengaruh Response:**
Pelanggan dengan tingkat pendidikan tinggi, seperti PhD dan Master, cenderung lebih Responsif terhadap Campaign. Ini dapat menjadi informasi kunci dalam merancang Campaign yang lebih efektif dan menargetkan kelompok pelanggan yang lebih cenderung merespon berdasarkan tingkat pendidikan mereka.

Insight Response Campaign Berdasarkan Status Perkawinan:

1. **Perbedaan Signifikan pada Tingkat Response:**
Terdapat perbedaan signifikan dalam tingkat Response Campaign antara status perkawinan. Pelanggan yang "Married" memiliki tingkat Response yang lebih rendah (11.34%) dibandingkan dengan pelanggan yang "Divorced" (20.69%), "Single" (22.08%), dan "Together" (10.35%).
1. **Tingkat Response Tinggi pada Status Perkawinan "Single" dan "Divorced":**
Pelanggan yang berstatus "Single" dan "Divorced" menunjukkan tingkat Response Campaign yang lebih tinggi, masing-masing sebesar 22.08% dan 20.69%. Ini menandakan bahwa status perkawinan ini dapat menjadi faktor penting dalam menentukan Response positif terhadap Campaign.
1. **Pentingnya Personalisasi Campaign untuk Setiap Status Perkawinan:**
Perusahaan dapat mempertimbangkan strategi pemasaran yang lebih personal dan disesuaikan dengan masing-masing status perkawinan. Status perkawinan seperti "Married" mungkin memerlukan pendekatan yang berbeda untuk meningkatkan Response Campaign.

3. Multivariate Analysis

c. Date Time Features

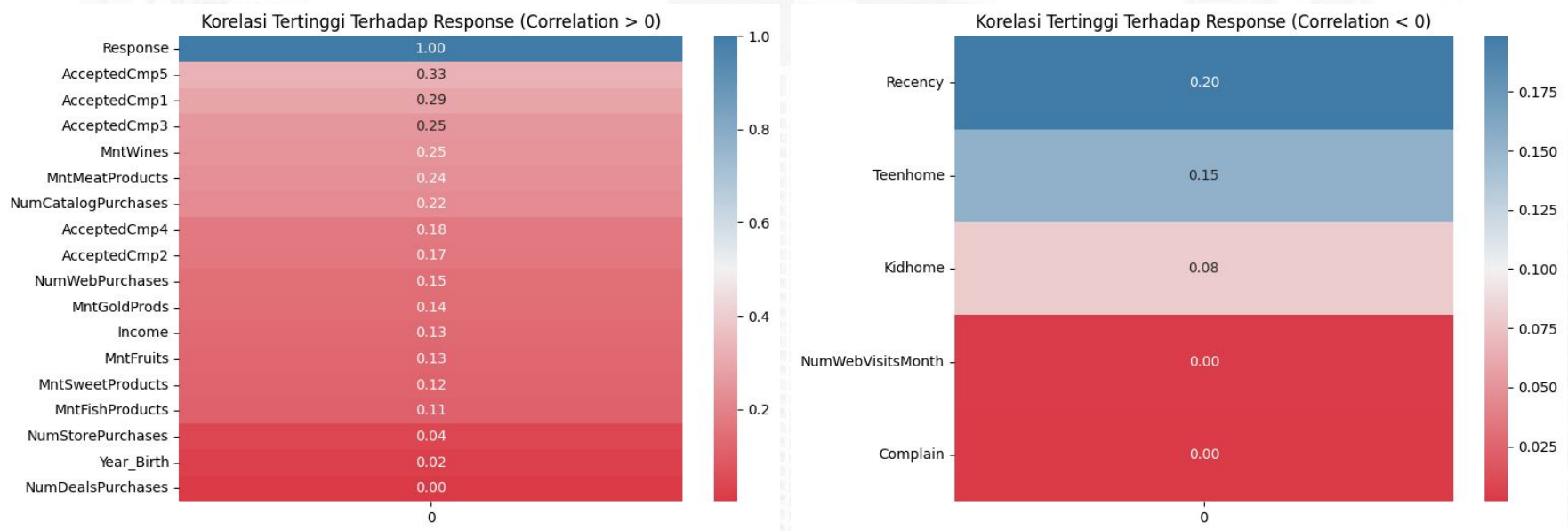


Berdasarkan grafik di samping, diketahui bahwa:

1. Dalam rentang waktu Juli 2012 hingga Juni 2014, lebih banyak pelanggan yang mendaftar namun tidak merespon campaign dibandingkan dengan yang merespon campaign.
2. Jumlah pelanggan yang mendaftar dan merespon campaign mencapai puncaknya pada Agustus 2012 - September 2012, yaitu kurang lebih 30 pelanggan per bulan, namun cenderung menurun hingga Juni 2014.

3. Multivariate Analysis

Berikut ini merupakan correlation heatmap untuk menganalisis korelasi antara feature dengan target (Responsee)



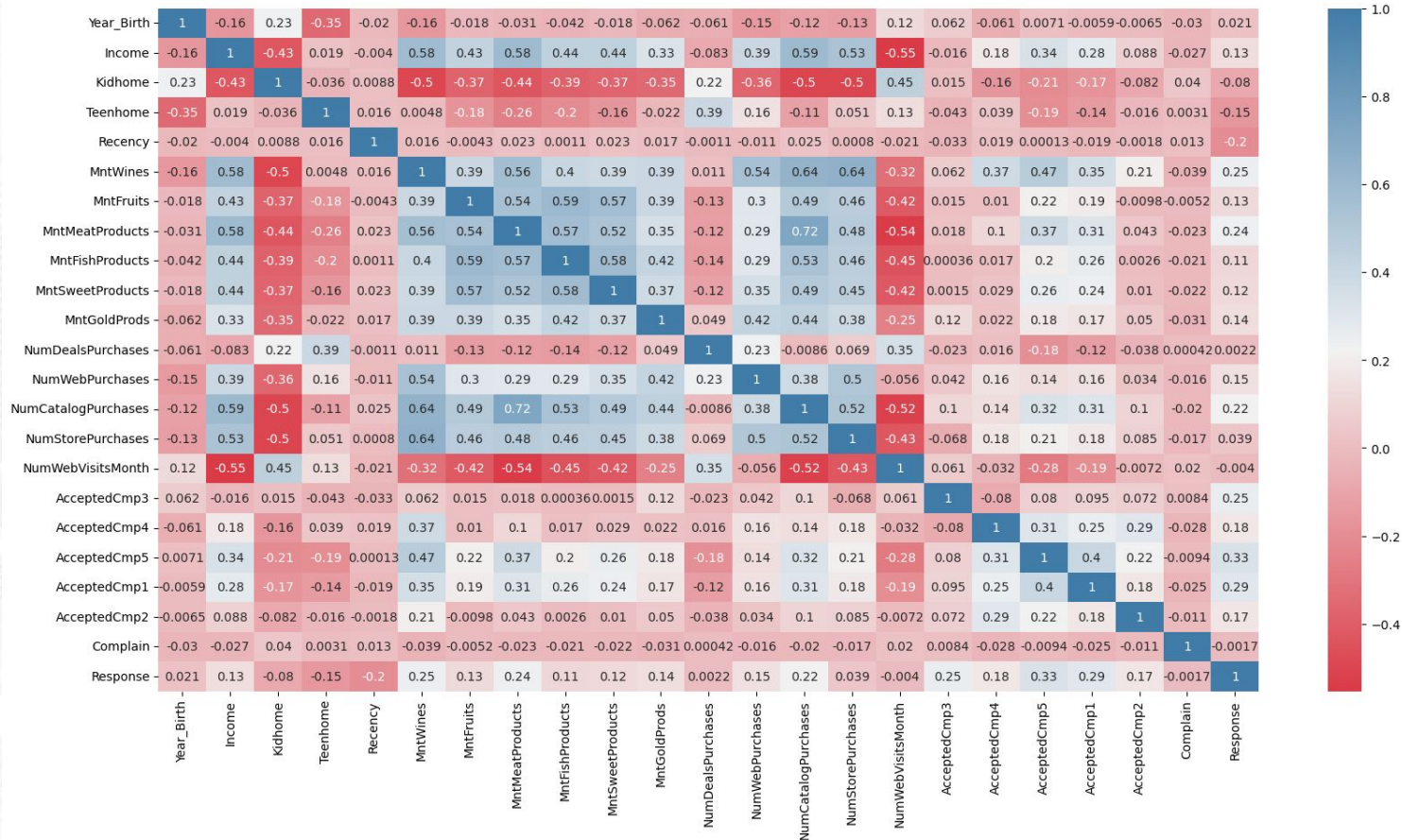
Dari hasil korelasi di atas, beberapa poin penting dapat diidentifikasi:

Feature yang termasuk dalam campaign (AcceptedCmp1, AcceptedCmp2, dst.), feature yang termasuk dalam produk spending seperti (MntWines, MntMeatProducts, dst.), feature yang termasuk pada channel penjualan seperti (NumCatalogPurchases, NumWebPurchases, dst.) mempunyai korelasi positif terhadap target.

Sedangkan hanya beberapa feature yang mempunyai nilai korelasi negatif terhadap target yaitu Recency dan jumlah anak (Teenhome dan Kidhome).

3. Multivariate Analysis

Berikut ini merupakan correlation heatmap untuk menganalisis korelasi antara feature dan target serta korelasi antar feature.



3. Multivariate Analysis

Dari hasil korelasi di atas, beberapa poin penting dapat diidentifikasi:

1. Korelasi Antara Feature:

Korelasi antara beberapa pasang feature cukup tinggi, yang bisa menunjukkan adanya multicollinearity. Contohnya, "MntMeatProducts" dengan "NumCatalogPurchases" (0.72), "MntWines" dengan "NumCatalogPurchases" dan "NumStorePurchases" (0.64), "Income" dan beberapa feature pengeluaran makanan ("MntWines", "MntFruits", "MntMeatProducts")

1. Rekomendasi untuk Proses Modeling:

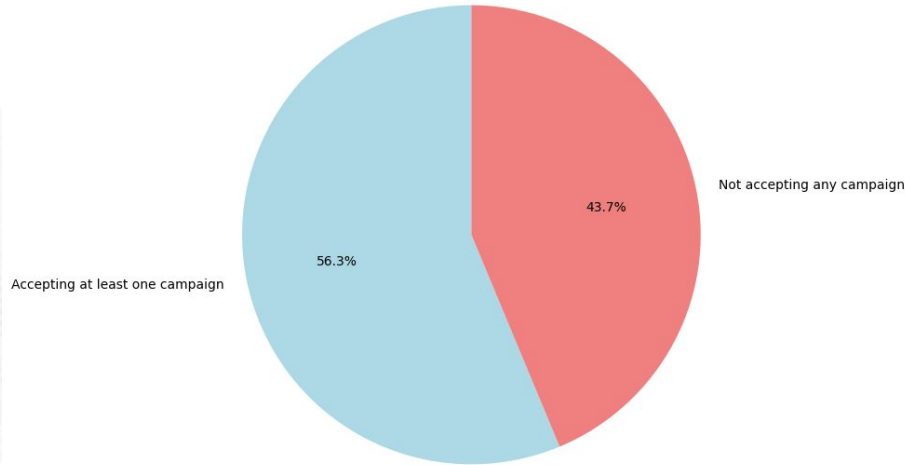
Kami akan melakukan analisis lebih lanjut untuk memahami apakah feature-feature yang berkorelasi tinggi tersebut memang memiliki signifikansi dalam konteks domain atau jika ada feature yang dapat dihilangkan.

Penerapan PCA bisa menjadi solusi untuk mengatasi multicollinearity dan mereduksi dimensi. Namun, sebelumnya, kami akan memastikan untuk melakukan standarisasi atau normalisasi pada data.

Kami akan juga akan menganalisa feature-feature dengan korelasi tinggi terhadap target ("Response") karena mungkin memiliki kontribusi yang signifikan pada model.

4. Business Insight

Percentage of Yes Response with and without Accepting Campaigns



Dari grafik disamping, kita dapat mengambil beberapa insight yang dapat berguna dalam konteks bisnis:

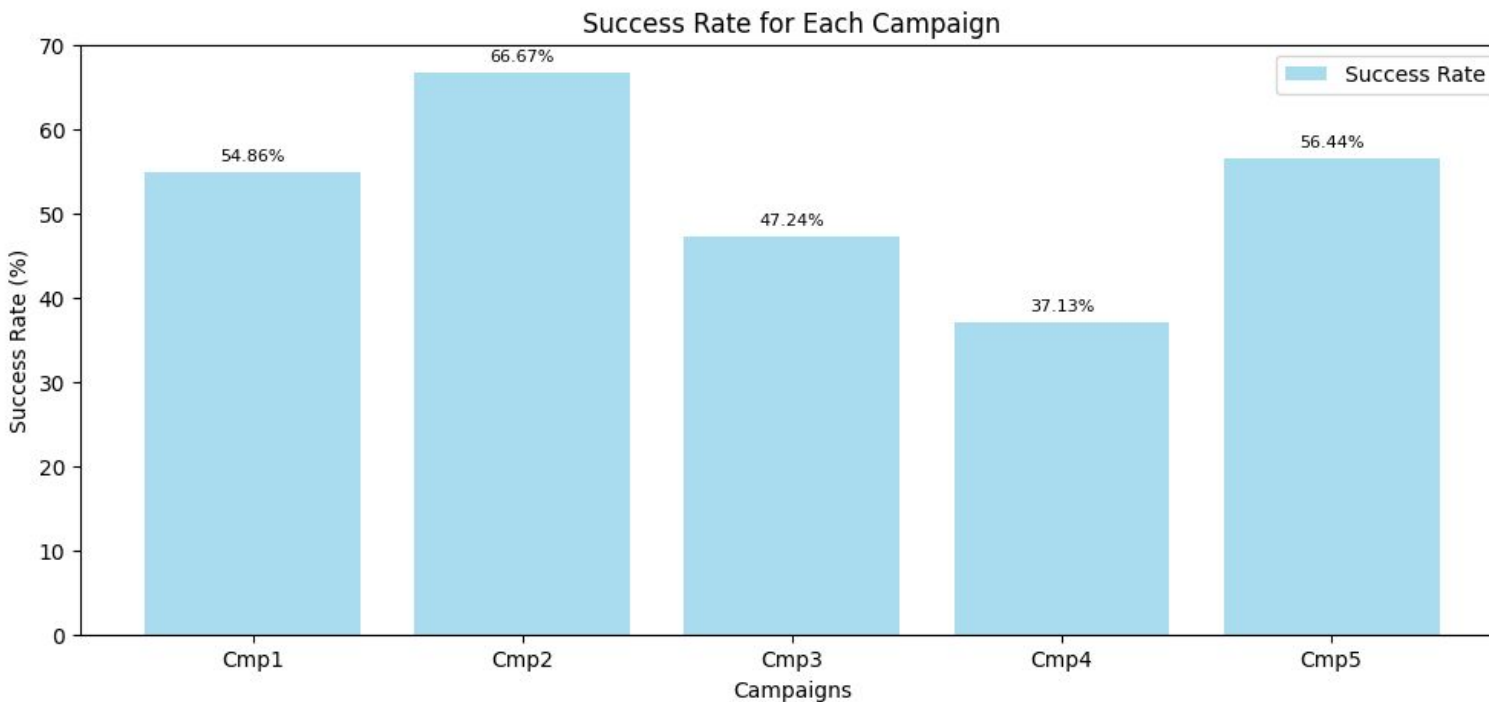
1. Pentingnya Campaign

Lebih dari setengah dari responden yang memberikan Yes Response telah menerima setidaknya satu Campaign (56.29%). Hal ini menunjukkan bahwa Campaign memainkan peran penting dalam meraih Response positif dari pelanggan.

2. Peluang Meningkatkan Kesuksesan Campaign

Terdapat potensi untuk meningkatkan kesuksesan Campaign karena masih ada sekitar 43.71% responden yang memberikan Yes Responsee tanpa menerima Campaign apa pun. Analisis lebih lanjut dapat dilakukan untuk memahami alasan di balik keputusan ini dan untuk memperbaiki strategi Campaign agar lebih menarik bagi pelanggan.

4. Business Insight



4. Business Insight

Dari grafik diatas, kita dapat mengambil beberapa insight sebagai berikut:

1. Efektivitas Pada Campaign Tertentu

Pemahaman tentang keberhasilan Campaign tertentu dapat membantu bisnis mengidentifikasi strategi yang efektif dan jenis Campaign yang lebih disukai oleh pelanggan. Misalnya, **Campaign 2 (Cmp2)** memiliki tingkat kesuksesan yang tinggi (**66.67%**), sehingga strategi dari Campaign 2 dapat dijadikan acuan untuk strategi Campaign di masa mendatang.

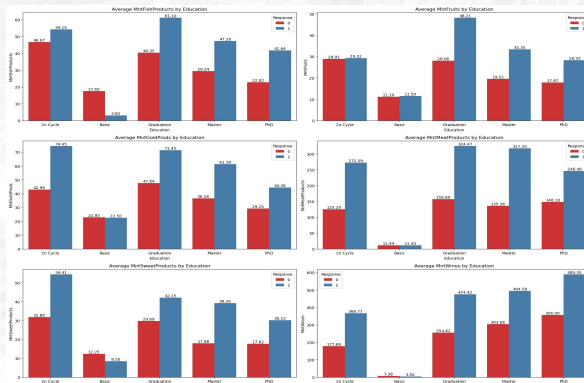
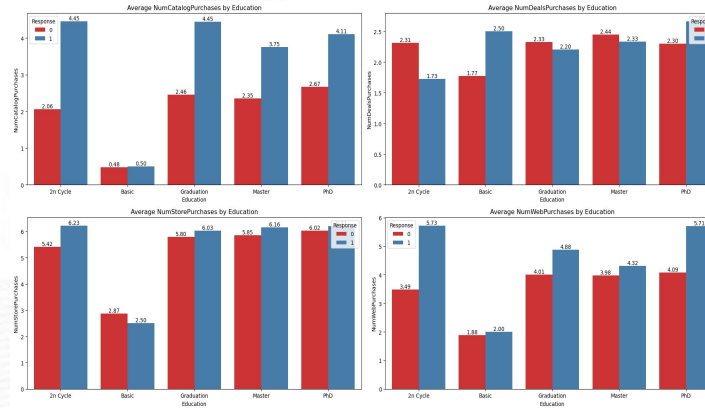
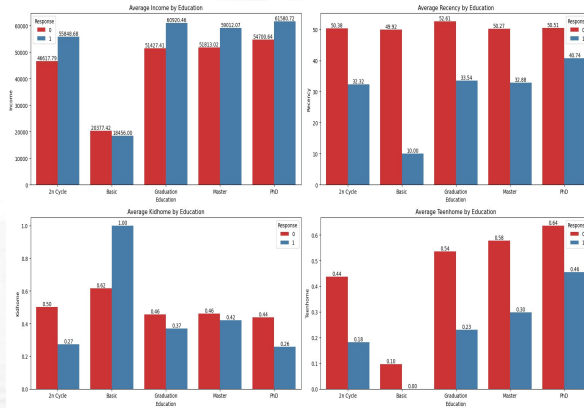
2. Potensi Peningkatan Respon Keseluruhan

Meskipun total Success Rate pada seluruh Campaign dan responden saat ini adalah 14.91%, peluang untuk meningkatkannya terbuka lebar. Analisis lebih lanjut, termasuk penargetan yang lebih baik, pesan Campaign yang lebih efektif, dan pemahaman lebih mendalam tentang preferensi pelanggan, dapat membantu meningkatkan tingkat respon keseluruhan.

3. Segmentasi Pelanggan:

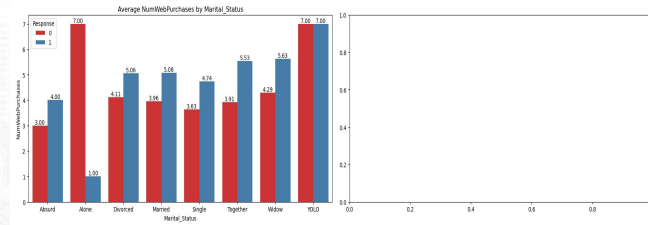
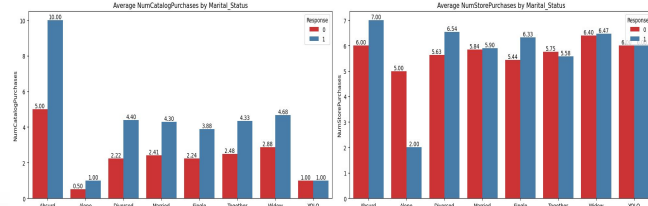
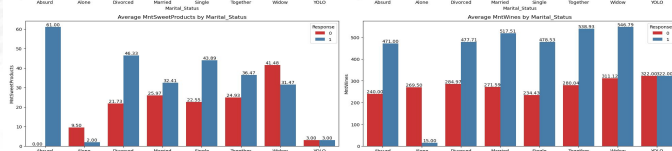
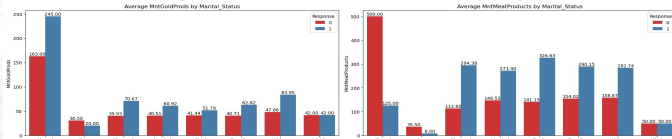
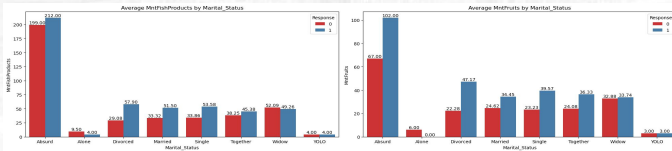
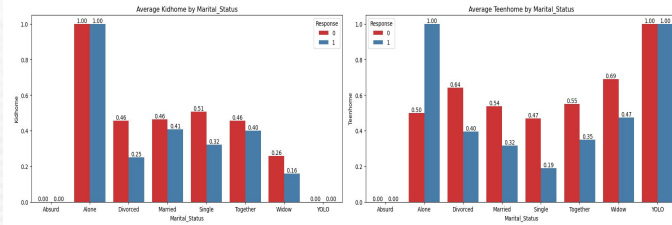
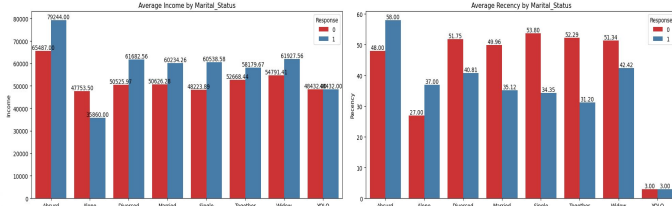
Penting untuk melakukan segmentasi pelanggan berdasarkan preferensi dan kebiasaan mereka terkait acceptance pada Campaign. Dengan memahami kelompok pelanggan yang merespon Campaign dengan baik dan kelompok yang tidak, untuk kedepannya dapat menyusun strategi yang lebih terarah dan personalisasi Campaign sesuai dengan setiap segmen.

4. Business Insight



1. Pelanggan yang memberikan respons terhadap kampanye memiliki rata-rata pendapatan yang lebih tinggi, periode pembelian yang lebih baru, dan jumlah anak yang lebih sedikit.
2. Pelanggan yang merespons kampanye menunjukkan tingkat pengeluaran yang lebih tinggi untuk berbagai kategori produk seperti wine, meat, fish, fruits, sweets, dan gold.
3. Pelanggan yang merespons kampanye cenderung memiliki rata-rata nilai pembelian yang lebih tinggi di seluruh saluran pembelian, termasuk melalui katalog, situs web, dan toko fisik.

4. Business Insight



1. Pelanggan yang merespons kampanye menunjukkan rata-rata pendapatan yang lebih tinggi, periode pembelian yang lebih baru, dan jumlah anak yang lebih sedikit.
2. Pelanggan yang memberikan respons terhadap kampanye menunjukkan tingkat pengeluaran yang lebih tinggi untuk berbagai kategori produk seperti wine, meat, fish, fruits, sweets, dan gold.
3. Pelanggan yang merespons kampanye cenderung memiliki rata-rata nilai pembelian yang lebih tinggi di seluruh saluran pembelian, termasuk melalui katalog, situs web, dan toko fisik.

Summary - Business Insight

Berdasarkan analisis mendalam terhadap data Campaign dan karakteristik pelanggan, kami menyarankan beberapa langkah strategis untuk meningkatkan efektivitas Campaign dan memaksimalkan keuntungan bisnis:

1. Segmentasi Pelanggan Berdasarkan Respons Campaign:

Melakukan segmentasi pelanggan berdasarkan respons Campaign dapat membantu dalam menyesuaikan strategi pemasaran. Fokuskan upaya pada kelompok pelanggan yang telah menunjukkan respons positif, seperti tingkat pendidikan Graduation, PhD, dan Master, serta status pernikahan Single, Married, dan Divorced.

2. Personalisasi Pesan dan Penawaran:

Personalisasi pesan dan penawaran Campaign untuk setiap kelompok pelanggan yang telah diidentifikasi dapat meningkatkan keterlibatan. Berdasarkan karakteristik unik dari setiap kelompok, buatlah pesan yang relevan dan tawarkan insentif yang sesuai dengan preferensi mereka.

3. Penargetan Tingkat Pendidikan Tinggi:

Tingkat pendidikan tinggi seperti Graduation, PhD, dan Master memiliki potensi besar untuk respons Campaign. Fokuskan penawaran khusus, informasi produk, dan keuntungan tambahan pada kelompok ini untuk memaksimalkan partisipasi.

4. Optimalkan Pengeluaran Pelanggan yang Merespon:

Pelanggan yang merespon Campaign memiliki kecenderungan pengeluaran yang lebih tinggi pada berbagai kategori produk. Optimalisasi persediaan dan promosi pada produk-produk yang paling diminati oleh kelompok pelanggan ini dapat meningkatkan nilai transaksi.

5. Memperkuat Campaign dengan Data Pembelian dan Channel:


Analisis menunjukkan bahwa pelanggan yang merespon Campaign memiliki rata-rata pembelian yang lebih tinggi di berbagai saluran seperti catalog, web, dan toko fisik. Penguatan Campaign dengan peningkatan ketersediaan produk melalui saluran ini dapat meningkatkan aksesibilitas produk bagi pelanggan.





6. Monitoring secara Realtime dan Analisis Reaksi Pelanggan:


Melakukan monitoring secara real time terhadap respons pelanggan contohnya dengan menggunakan BI Tools untuk Dashboarding, dan melakukan analisis lebih lanjut terhadap perubahan tren dan preferensi. Keterlibatan yang berkelanjutan dan penyesuaian cepat terhadap dinamika pasar dapat menjadi kunci kesuksesan jangka panjang.




5. GIT

<https://github.com/hilmanman92/market-insider/tree/master>


 market-insider Public




 Pin  Unwatch 1  Fork 0  Star 0



 master had recent pushes 9 minutes ago [Compare & pull request](#)

 master  2 branches  0 tags [Go to file](#) [Add file](#) [Code](#)

This branch is 3 commits ahead, 1 commit behind main. [Contribute](#)

 hilmanman92 penambahan summary pada Ddesc. Statistics dan summary EDA pada RE... 378190d 10 minutes ago 3 commits

 README.md	penambahan summary pada Ddesc. Statistics dan summary EDA pada R...	10 minutes ago
 market insider - marketing campaign...	penambahan summary pada desc. statistics, eda, dan business insights	19 minutes ago
 marketing_campaign.csv	add project to github	13 hours ago






 README.md 

Summary Descriptive Statistics

- Apakah ada kolom dengan tipe data kurang sesuai, atau nama kolom dan isinya kurang sesuai?
 - Feature Dt_Customer merupakan tanggal registrasi pelanggan dengan tipe data object. Tipe data ini tidak sesuai, sehingga perlu diubah menjadi tipe Date Time.
- Apakah ada kolom yang memiliki nilai kosong? Jika ada, apa saja?
 - Feature Income mempunyai nilai kosong karena hanya berjumlah 2216 rows.
- Apakah ada kolom yang memiliki nilai summary agak aneh? (min/mean/median/max/unique/top/freq)

About

Final Project of Market Insider

-  Readme
-  Activity
-  0 stars
-  1 watching
-  0 forks

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)