

Final Project

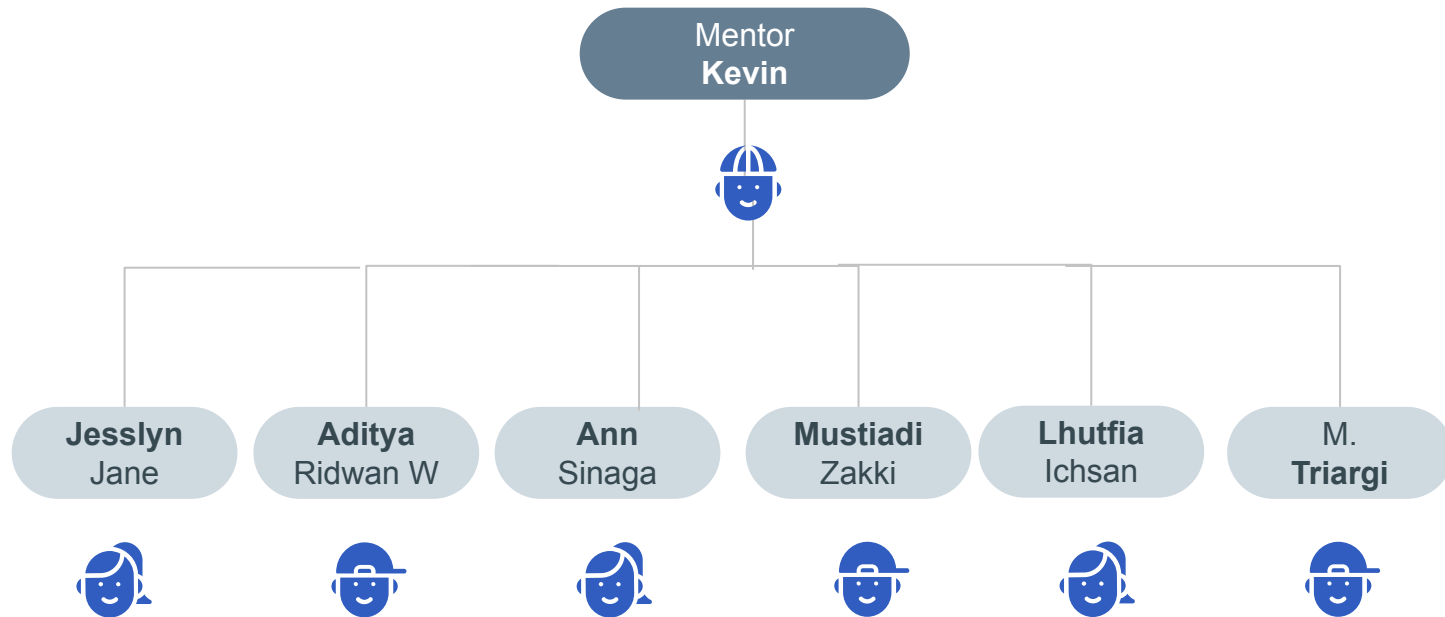
Marketing Campaign
Halcyon Group

 **Rakamin**
Academy


HALCYON



Introducing our great data science members:



Halcyon

~ Happy, Prosperous, Calm and Peaceful ~

Table of contents

01

**Business
Understanding**

02

**Exploratory Data
Analysis**

03

**Data
Preprocessing**

04

**Modelling &
Evaluation**

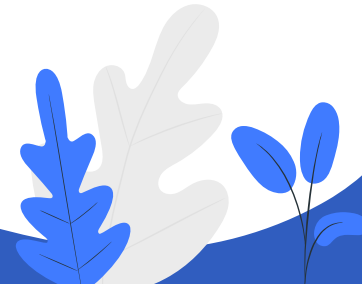
05

**Business
Recommendation**

06

**Business
Simulation**

01 Business Understanding



Company Background

Problems

Problem yang dihadapi perusahaan:

- **Tingkat acceptance** marketing campaign yang kecil sekitar 14.91%
- **Cost yang dikeluarkan** perusahaan dalam melakukan marketing campaign tidak efisien
- **Revenue tidak sebanding** dengan cost yang dikeluarkan



Goal and Objective



Goal

Meningkatkan **response rate** dari marketing campaign, sehingga **profit** serta **efisiensi marketing cost** dapat lebih optimal.



Objective

Membuat suatu model prediksi yang dapat mengklasifikasikan **pengguna yang tepat** untuk meningkatkan efektivitas marketing campaign.

Business Metrics



Response Rate

Rasio jumlah customer yang merespon dibandingkan dengan total impresi campaign

*asumsi: online campaign

$$= \frac{\text{Total response}}{\text{Total Campaign Impression}}$$

Business Metrics

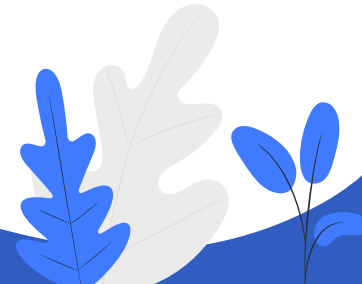


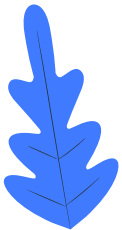
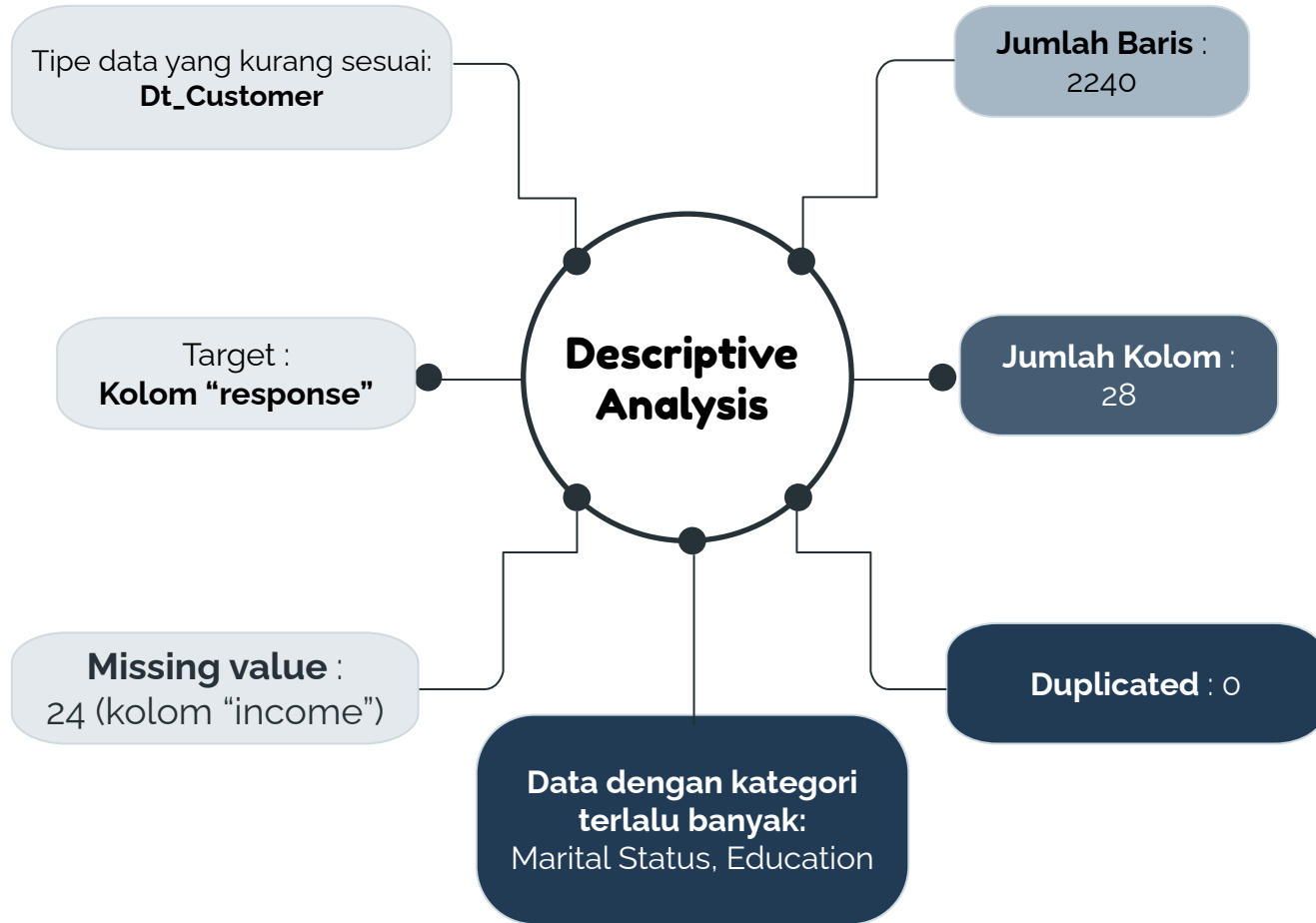
Revenue Rate

Rasio jumlah keuntungan yang diperoleh perusahaan berdasarkan total response customer yang didapat

$$= \frac{((\text{Revenue per response} \times \text{total responses}) - (\text{Cost} \times \text{Total Campaign}))}{\text{Total Revenue}}$$

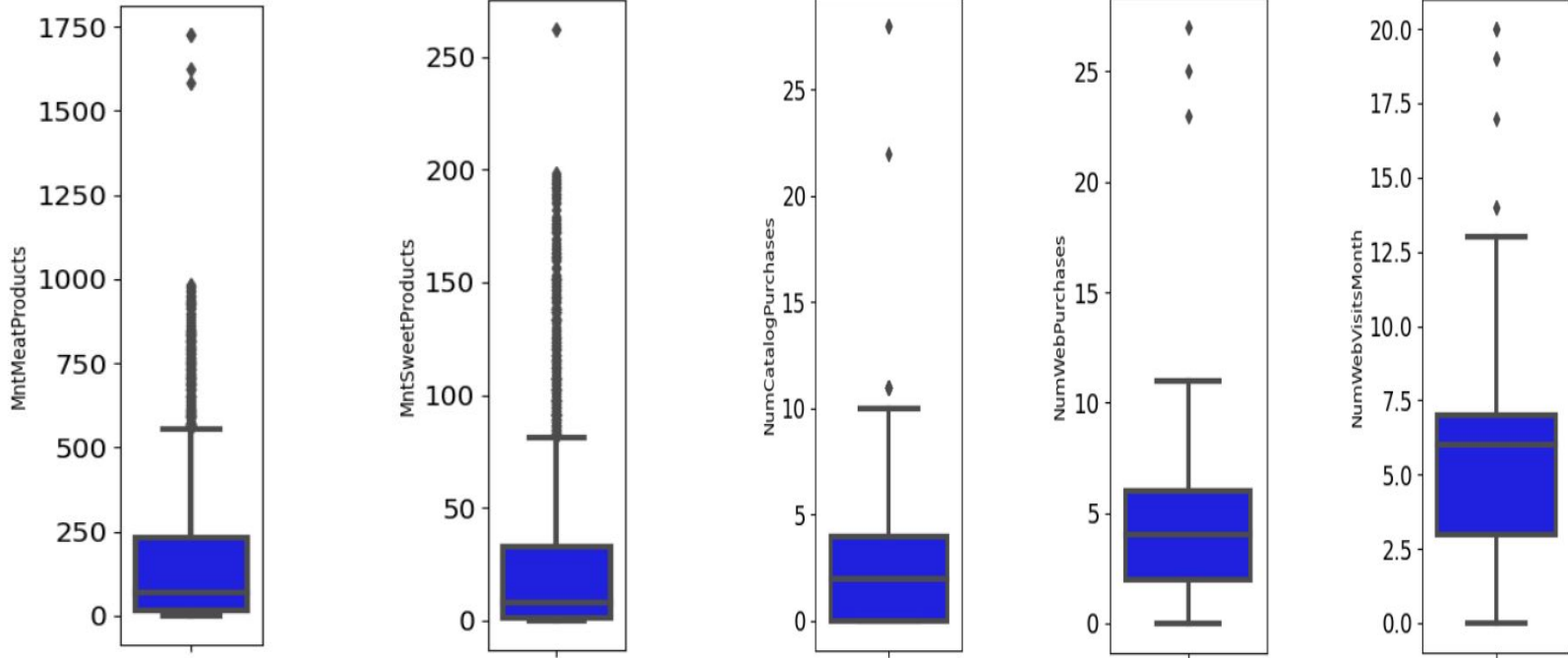
02 Exploratory Data Analysis





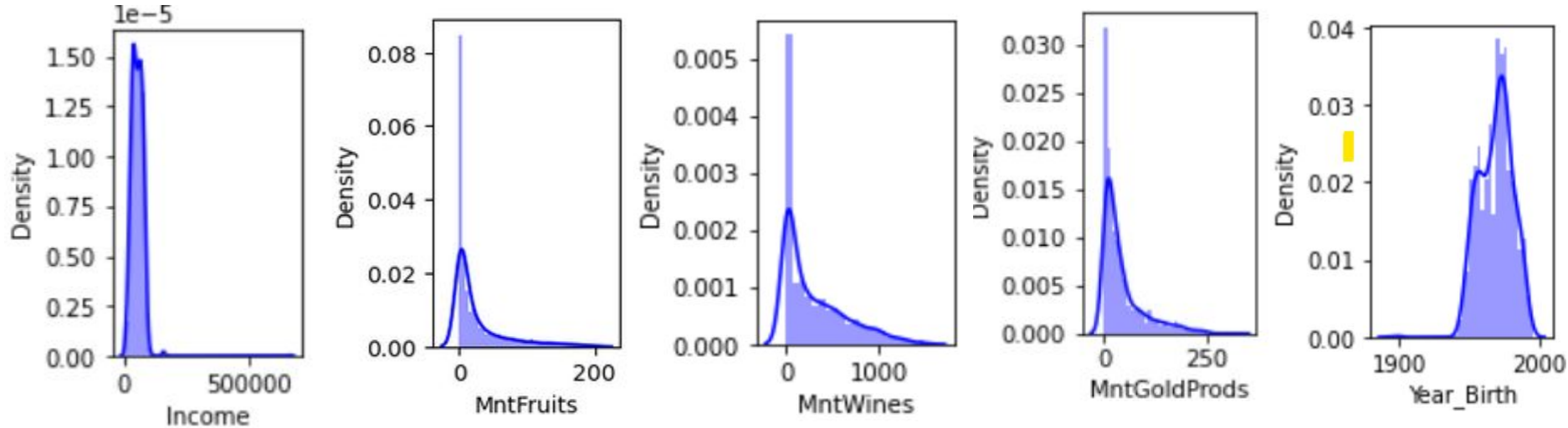
Univariate Analysis

Beberapa Kolom dengan data outlier



Univariate Analysis

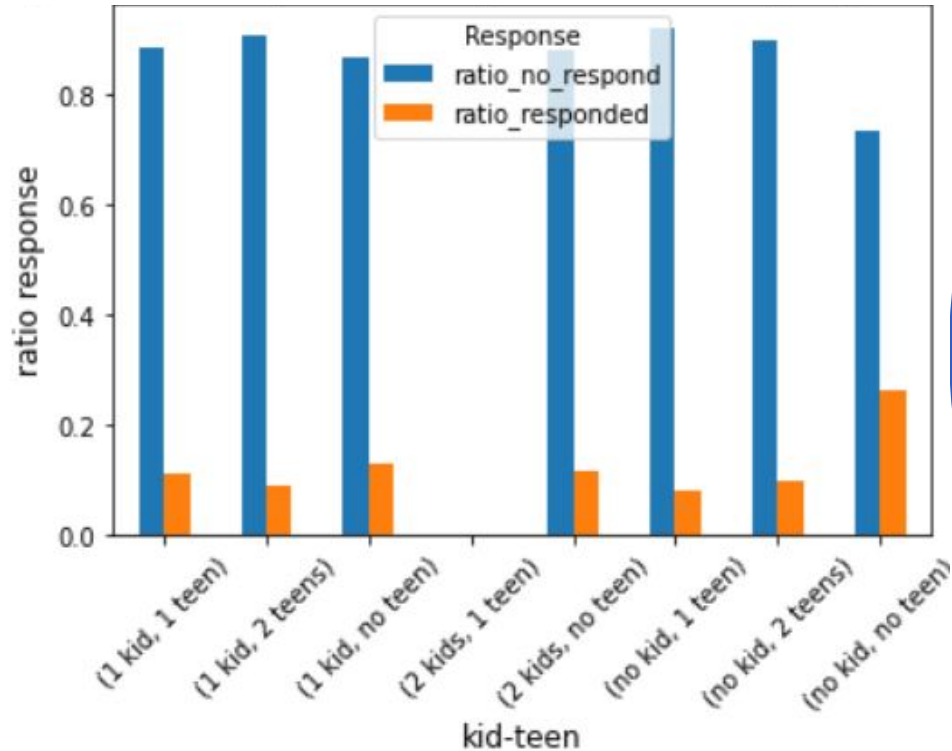
Beberapa kolom yang sebarannya **Skewed**



Positively Skewed : Income, Wines, Fruit, Meat, Fish, Sweet Product, Gold Product, Deals Purchase, Web Purchases, Catalog Purchases, Store Purchases, Web Visit Months

Negatively Skewed : Year Birth

Multivariate Analysis

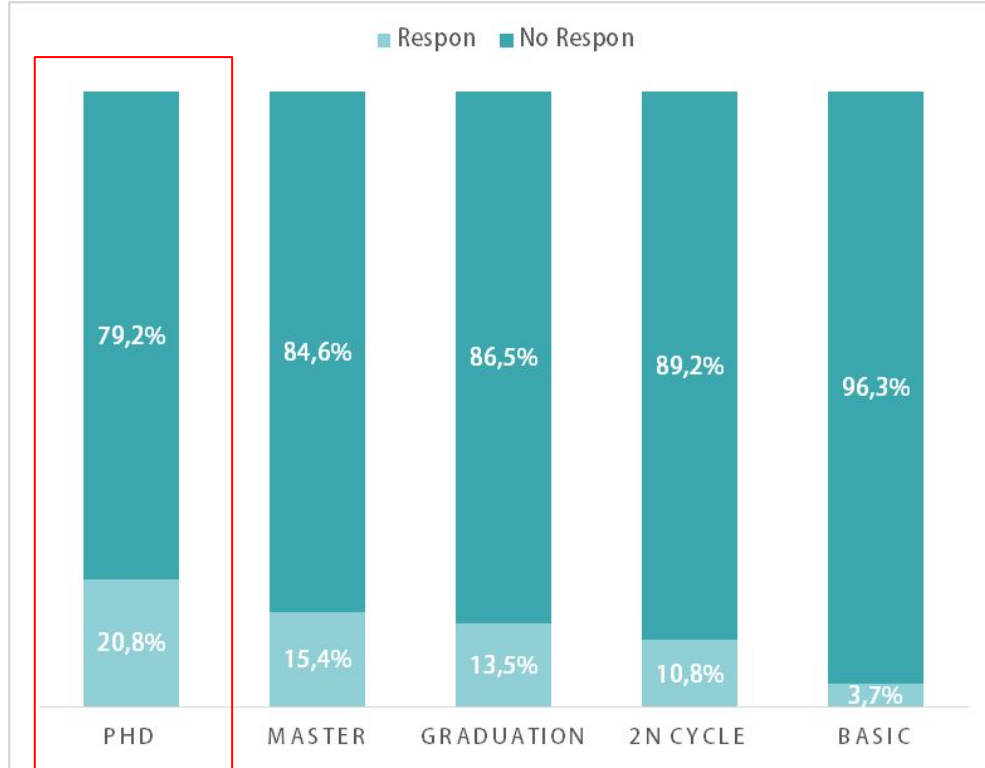


Korelasi antara kepemilikan anak dan remaja terhadap respon

Customer yang tidak mempunyai anak kecil dan anak remaja cenderung memiliki tingkat respon yang lebih tinggi

Multivariate Analysis

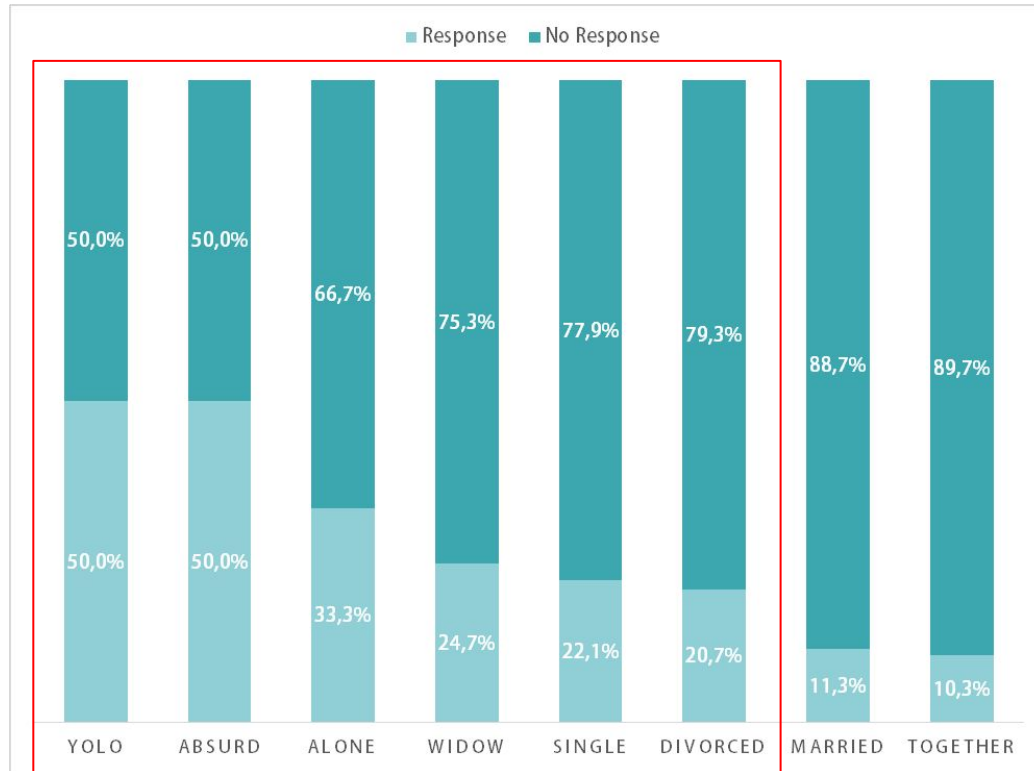
Persentase Respon Terhadap Education



Customer dengan tingkat education PhD memiliki tingkat respon yang lebih tinggi

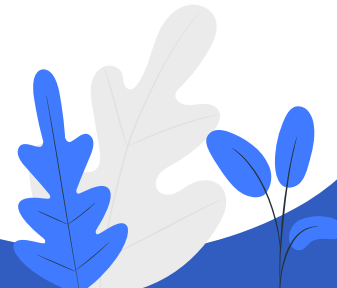
Multivariate Analysis

Persentase Respon Terhadap Marital Status



Customer dengan marital status yang cenderung hidup sendiri, memiliki tingkat respon yang lebih tinggi

02 Data Pre Processing



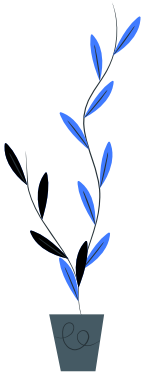
Data Cleansing



Missing Value

< 3% missing value

Handling : **Drop**



Data Type

Dt_Customer

Handling: **convert to date**



Outlier

Handling : **Z - Score**

Baris yang dihapus
hanya 12 %



Feature Encoding

Marital Status

Absurd, Yolo, Alone, Divorced, Single : 0
Together, Married: 1

Education

Basic: 0
Graduation: 1
Master, 2nd Cycle: 2
PhD: 3



Feature Transformation



Log Transformation : Untuk Memperkecil Range



Kolom yang Ditransformasi :

MntWines

MntFruits

MntMeatProducts

MntFishProducts

MntSweetProducts MntGoldProds

NumDealsPurchases

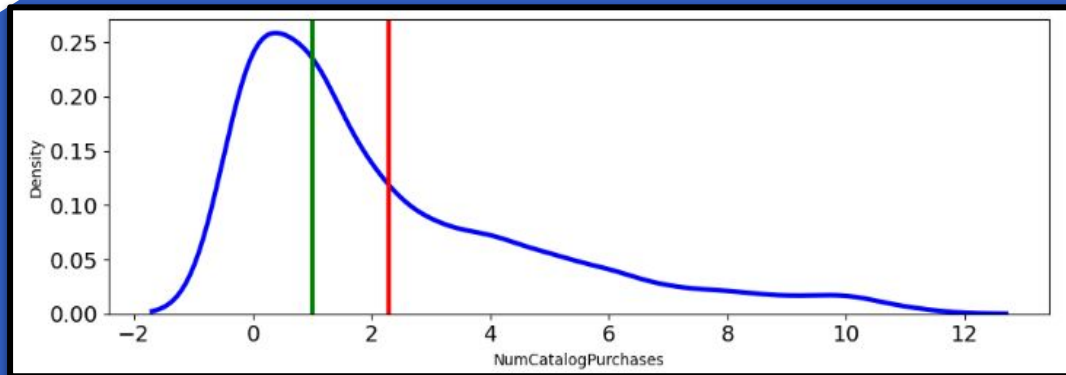
NumWebPurchases

NumCatalogPurchases

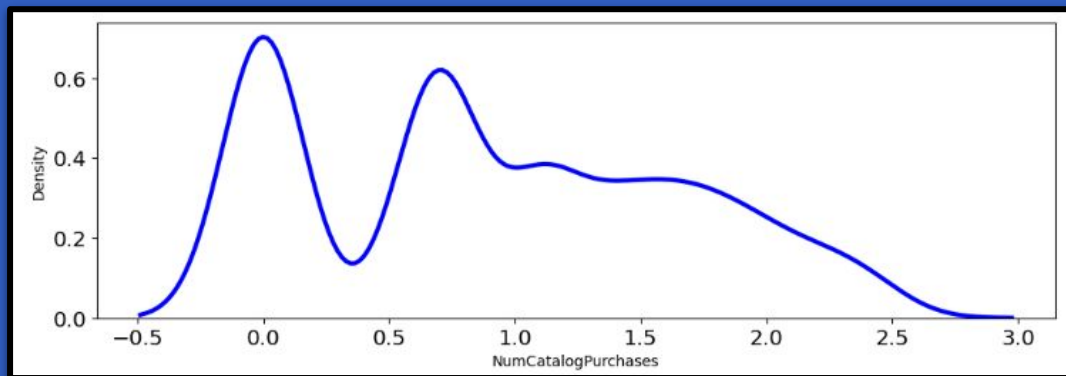
Log Transformation

Range data pada kolom besar menjadi kecil

Before



After



Feature Extraction

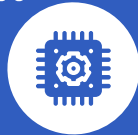
Total Dependents

Summary of : Marital
Status + Kid Home +
Teen Home



Total Accepted Campaign

Summary of : Accepted
Campaign 1 - 5



Total Spent

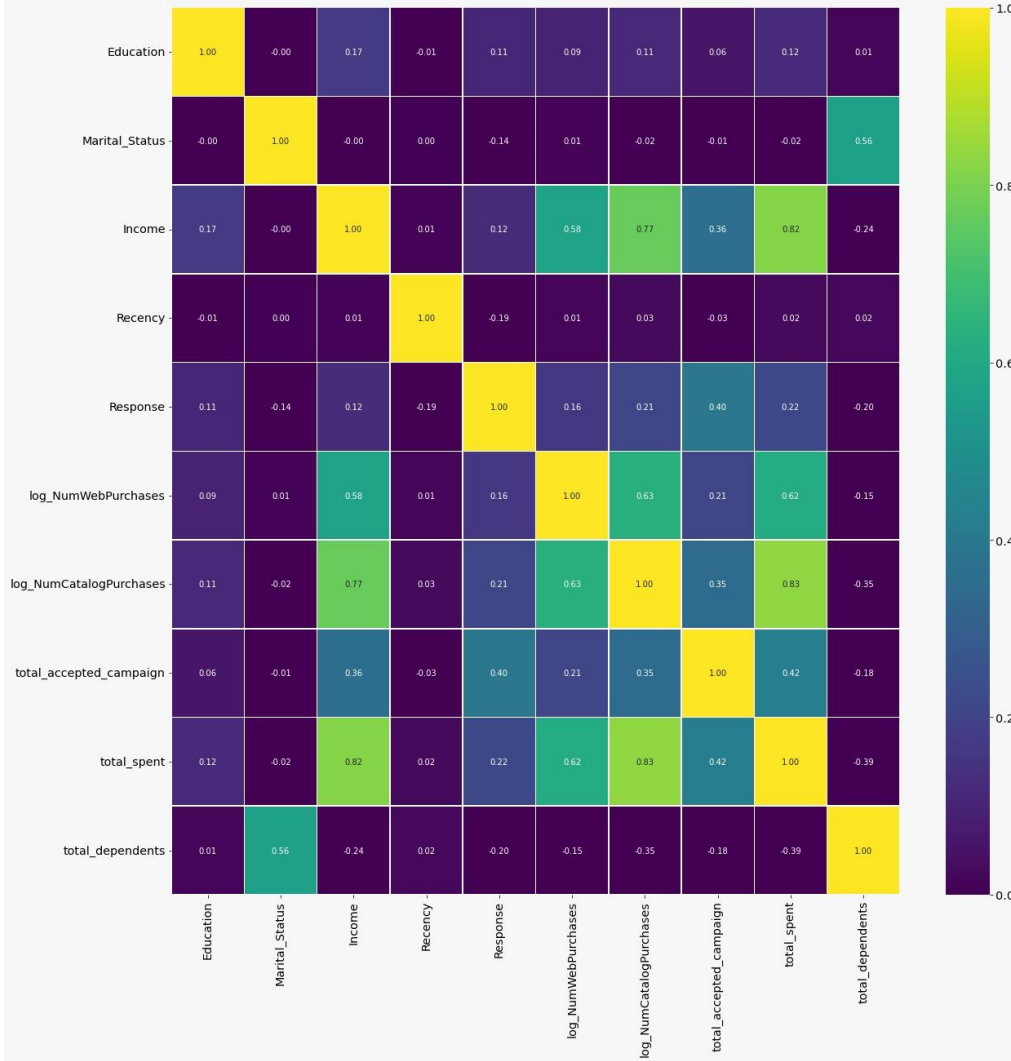
Summary of : Wine +
Sweet Product + Gold +
Fish + Meat + Fruit



Total Revenue

Total Accepted
Campaign * Revenue





Feature Selection

Semakin banyak feature akan semakin memberatkan Machine Learning

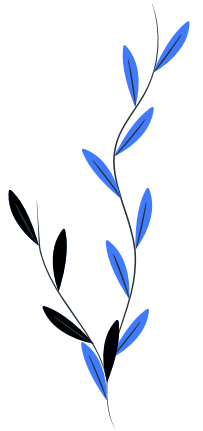
34 → **9 Features**



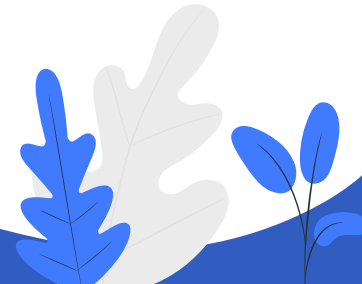
Class Imbalance

Degree Imbalance tergolong Moderate. Handling : **Oversampling**

Jumlah sampel yang dipelajari Machine lebih banyak apabila menggunakan oversampling.



04 Modelling and Evaluation



Parameter Evaluasi Model

Precision as Primary Parameter Evaluation

- Meningkatkan Response Rate
- Mereduksi False Positif (Customer yang diprediksi akan merespon, namun kenyataannya tidak)

Recall as Secondary Parameter Evaluation

- Mengoptimalkan Revenue Rate
- Mereduksi False Negative (Customer yang diprediksi tidak mengikuti campaign, namun pada kenyataannya berkeinginan ikut campaign)

Modelling

	Train Precision	Test Precision	Train Recall	Test Recall
Decision Tree	1,00	0,44	0,98	0,62
Random Forest	0,99	0,55	0,99	0,62
XGBoost	0,90	0,51	0,85	0,55
AdaBoost	0,80	0,43	0,73	0,54
KNN	0,77	0,29	0,73	0,38



Hypertuning Result

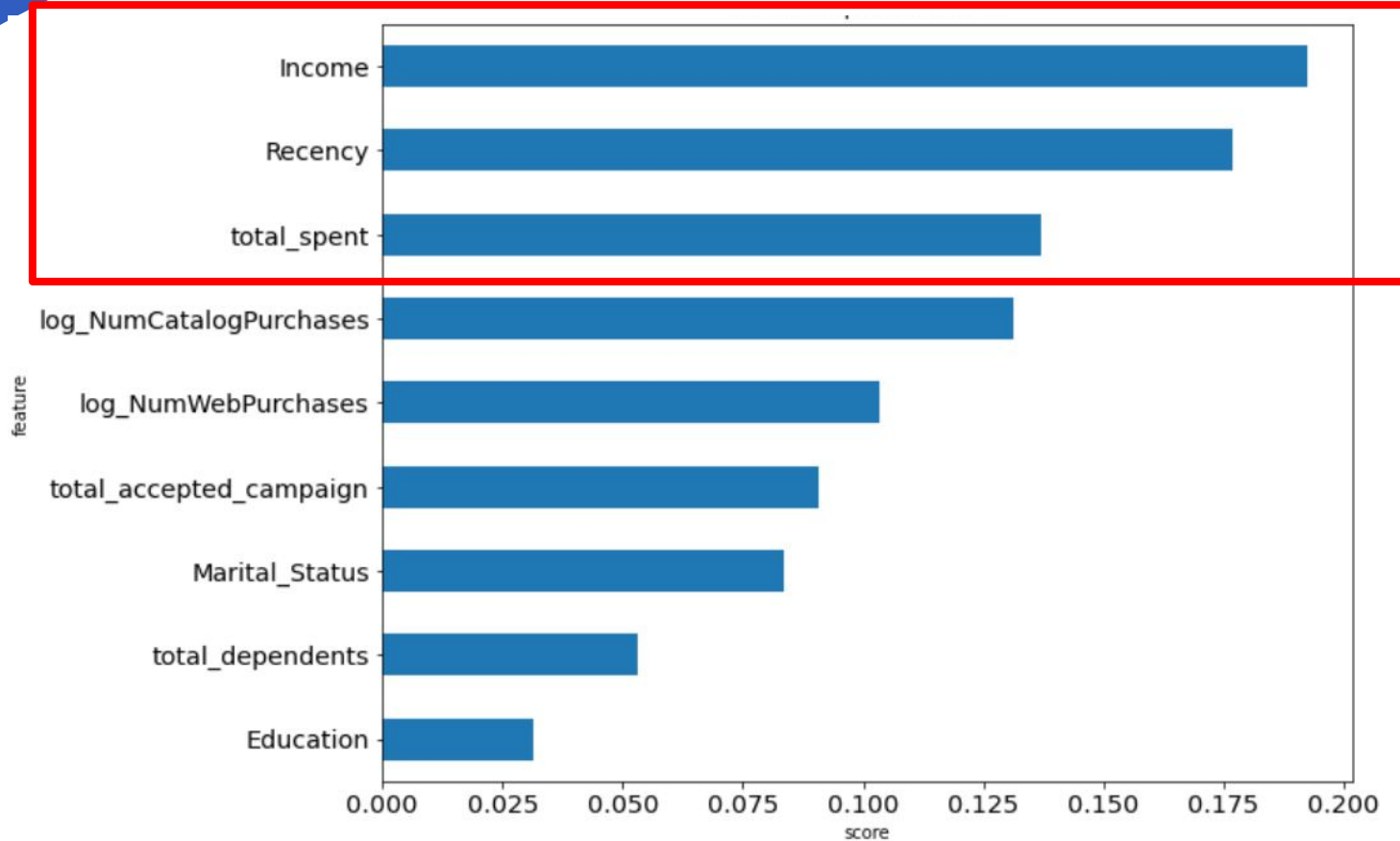
	Train Precision	Test Precision	Train Recall	Test Recall
Random Forest	0,84	0,58	0,34	0,33
XGBoost	0,83	0,58	0,27	0,27
Decision Tree	0,80	0,51	0,29	0,31

Model paling optimal : **Random Forest**

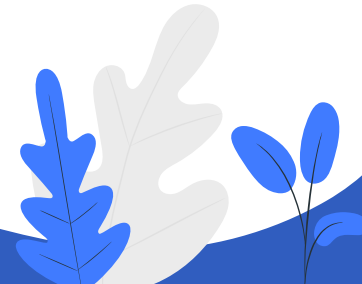
- **Precision & recall tinggi**
- **Gap** antara train dan test **rendah**



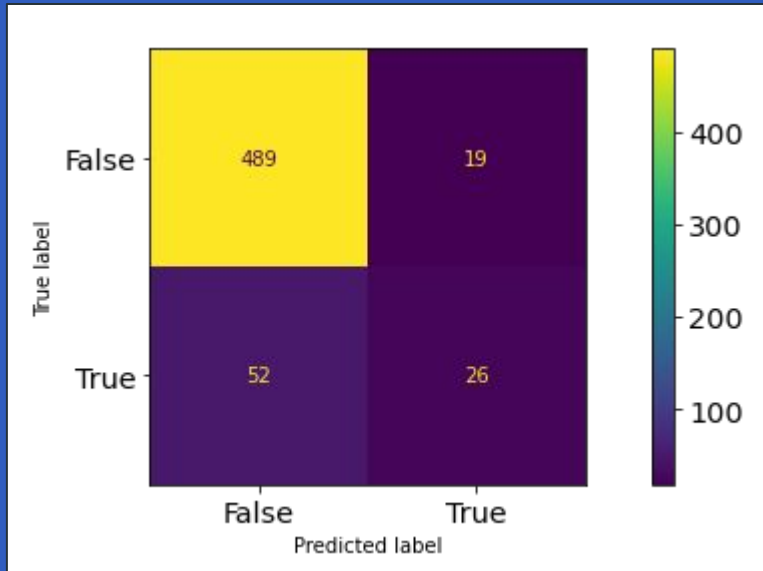
Feature Importance



05 Business Simulation



Confusion Matrix



26

True positif : Diprediksi merespon, dan itu benar

489

True Negative : Diprediksi tidak merespon, dan itu benar

19

False Positif : Diprediksi merespon, dan itu salah

52

False Negative : Diprediksi tidak merespon, dan itu salah

Response Rate

39,4 %



14,9 %

Before Model

54,3 %

After Model



Revenue Rate

153,9 %



-104,1 %

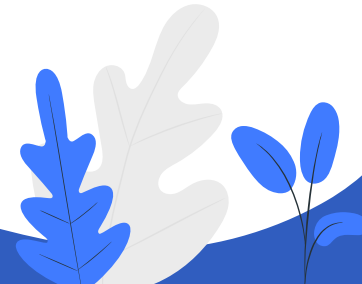
Before Model :
Loss 104,1%

49,8 %

After Model :
Profit 49,8 %



06 Business Recommendation



Business Recommendation

Income

Memberikan voucher belanja untuk customer dengan income >70000



Recency

Memberikan voucher belanja yang berlaku tidak lebih dari 2 minggu sejak voucher diterima.



Total Spent

Memberlakukan strategi up-selling/cross-selling dan memberi voucher khusus untuk minimal pembelian



Membuat promosi bulanan di tanggal cantik dan memasang e-advertising





Thanks

Taehyung makan bakwan
Thank you epry wan ❤️

