

Marketing Campaign

DATALICIOUS PRESENTATION 2023

Rakamin Academy
Data Science Batch 32

READ MORE



COMPANY BAGROUND

Lottomart is a supermarket retail that sell various type of products such as Fish, Meat, Fruits, Sweet Products, Wines, and Gold Products. For the last 6 months, Marketing Team has conducted a campaign in the form of giving discount vouchers to all customers via Broadcast Messages.



MEET THE TEAM



NUR IMAM
MASRI



SITI HAJJAH
MARDIAH



PRASIDYA
BAGASKARA



RISKIYATUL
HASANAH



M. RAYHAN
AZZINDANI



ASTUTI
RAHMAWATI



M. HARWIN
PRAYOGA



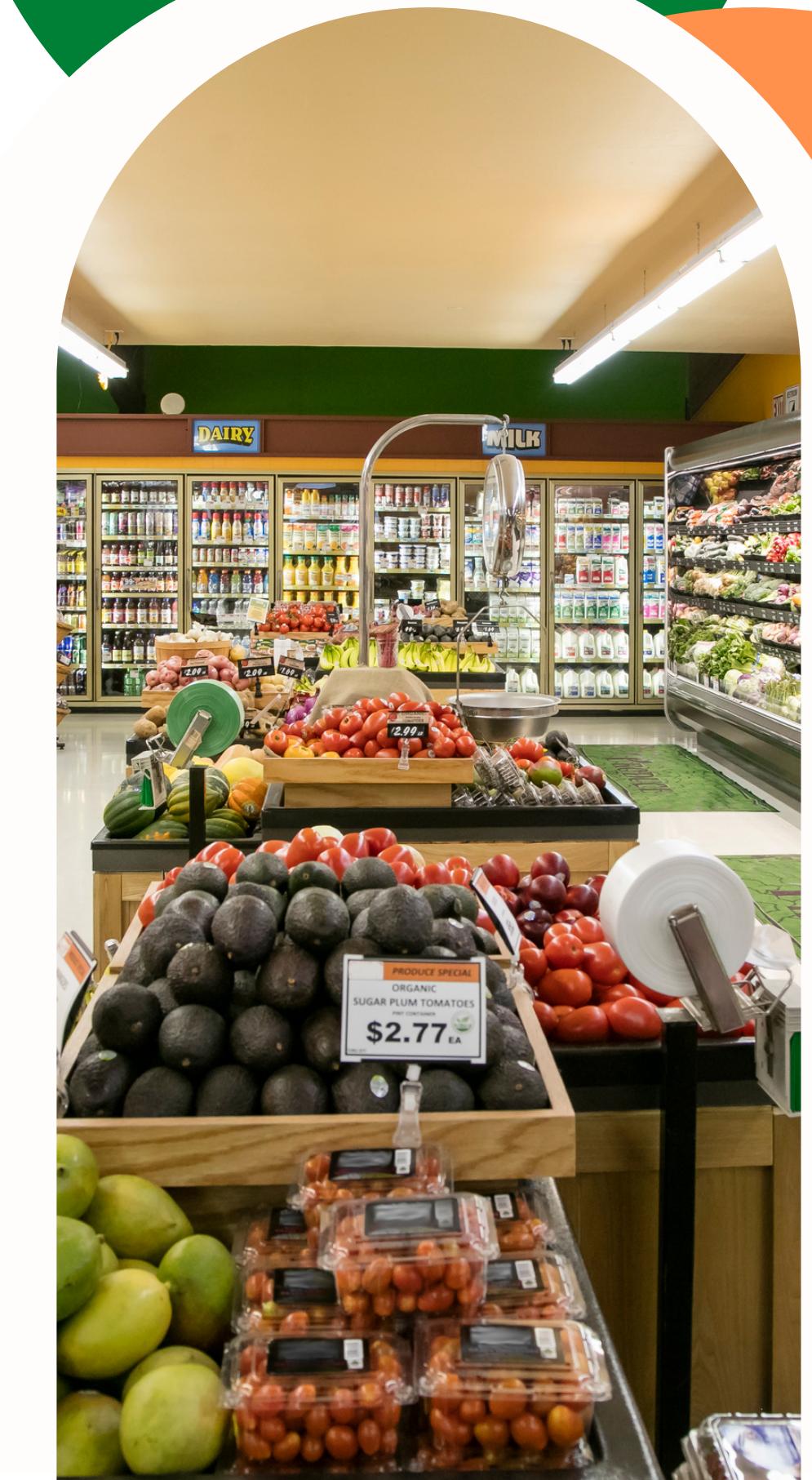
CHRISTINE



M. IFZAL
ASRIL



JOHANES
CRISTIAN



OUTLINE

**Business
Understanding**

**Exploratory Data
Analysis**

**Data
Preprocessing**

**Modelling &
Evaluation**

**Business
Recomendation**



Datalicious - Lotto Mart

Business Understanding

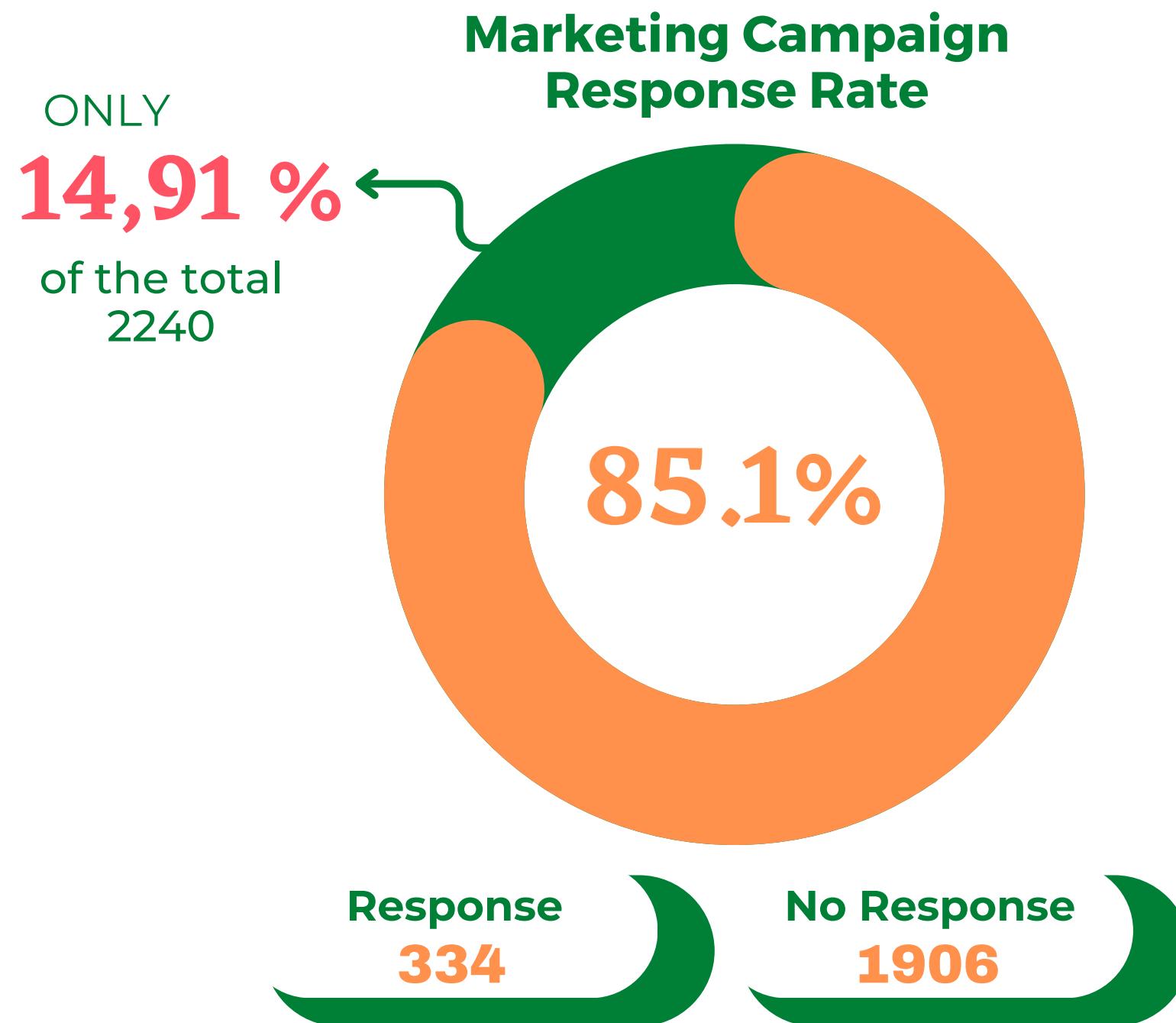
➤ Problem Statement

➤ Goals, Objective & Business Metrics

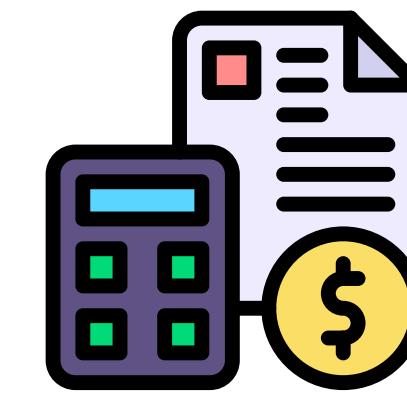


PROBLEM STATEMENT

1. Lower Response Rate



2. Inefficient Cost



3. Profit Isn't Maximum



GOALS, OBJECTIVE & BUSINESS METRICS

Goals



Increasing response rate and minimizing marketing cost for each customers, so it can **boost profit** for the next marketing campaign

Objective



1. Create a **classification model** to predict which **customer groups** will **respond** for the next marketing campaign
2. Create a **clustering model** to make it easier for companies **to determine the right target of customers**

Business Metrics



1. Response Rate
2. Net Profit Margin
3. Return of Investment

Datalicious - Lotto Mart

Exploratory Data Analysis

› Dataset Information

› Preliminary Insight



DATA UNDERSTANDING

Accepted/Responses
Campaign

AcceptedCmp1
AcceptedCmp2
AcceptedCmp3
AcceptedCmp4
AcceptedCmp5
Complain
Response (target)

Customer Information

ID
Year_Birth
Education
Marital
Kidhome
Teenhome
Income
DtCustomer
Recency

Z_CostContact
Z_Revenue

Sales Product Type

MntFishProducts
MntMeatProducts
MntFruits
MntSweetProduct
MntWines
MntGoldProds

Number of Purchases
per Type

NumDealsPurchases
NumCatalogPurchas
NumStorePurchases
NumWebPurchases
NumWebVisitsMonth

Dataset memiliki 29 kolom dan 2240 baris
Kolom income memiliki 2216 nilai non-null,dan 24 nilai null (1.07% data null)
Tidak terdapat data duplikat

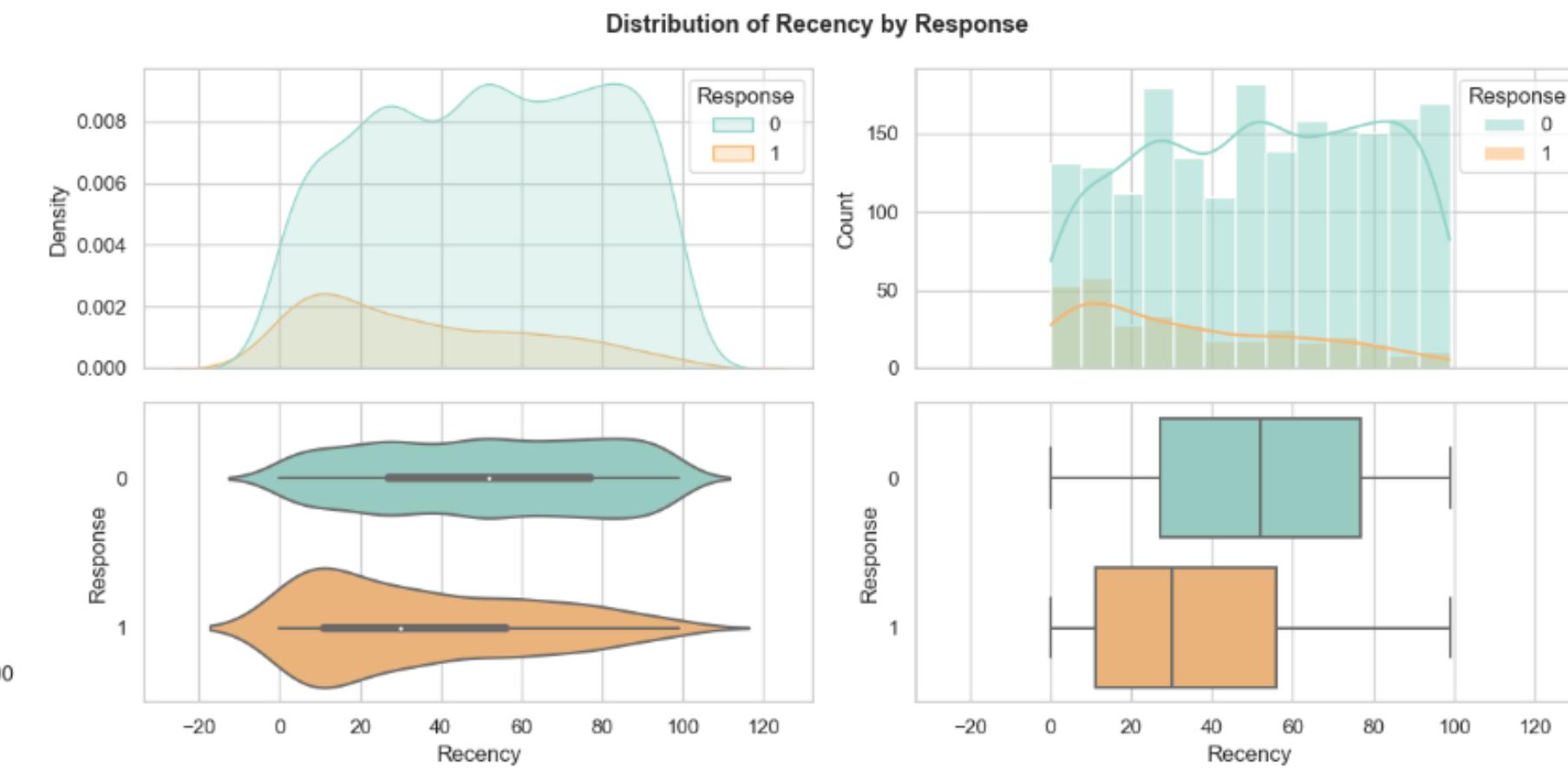


Business Insights



Dari visualisasi Income, dapat dilihat bahwa customer yang merespon terbanyak berasal dari customer dengan pendapatan > \$75000

Recency = 0-20 days



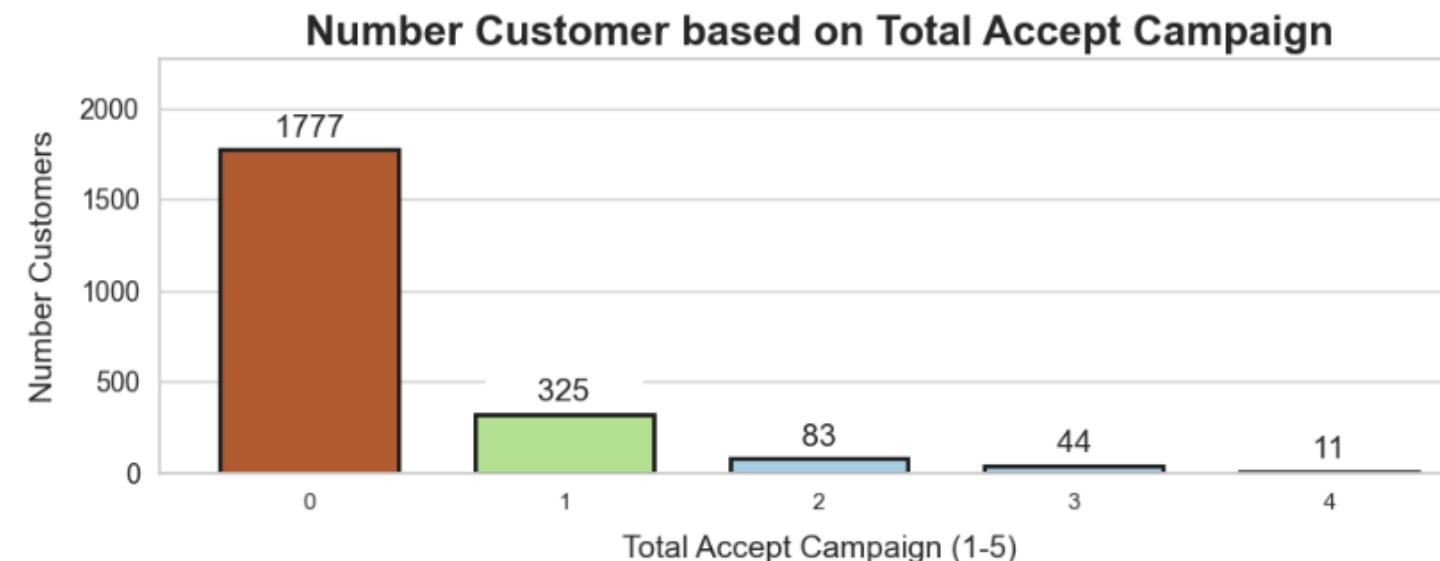
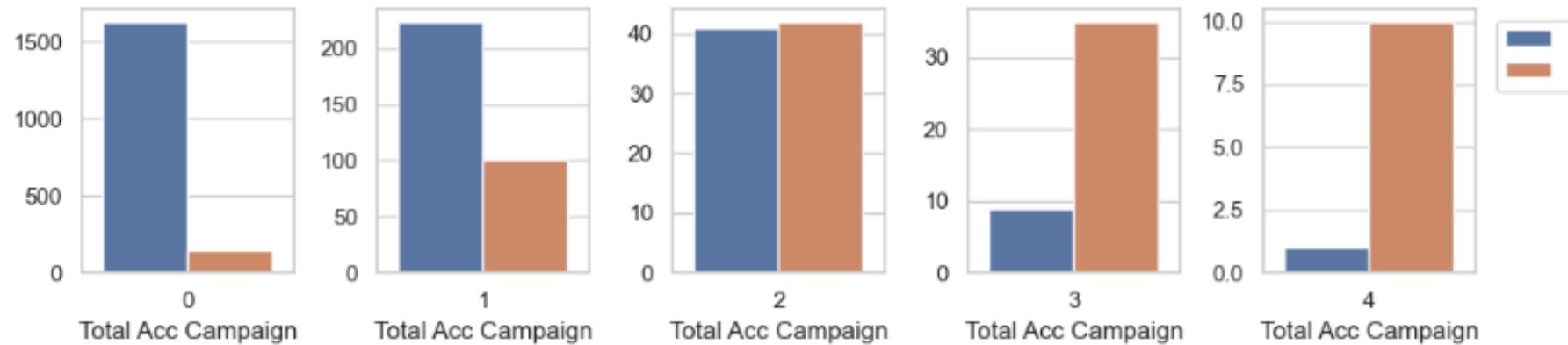
Dari visualisasi Recency, dapat dilihat bahwa customer yang merespon terbanyak berasal dari customer dengan Recency yang rendah

Business Insights



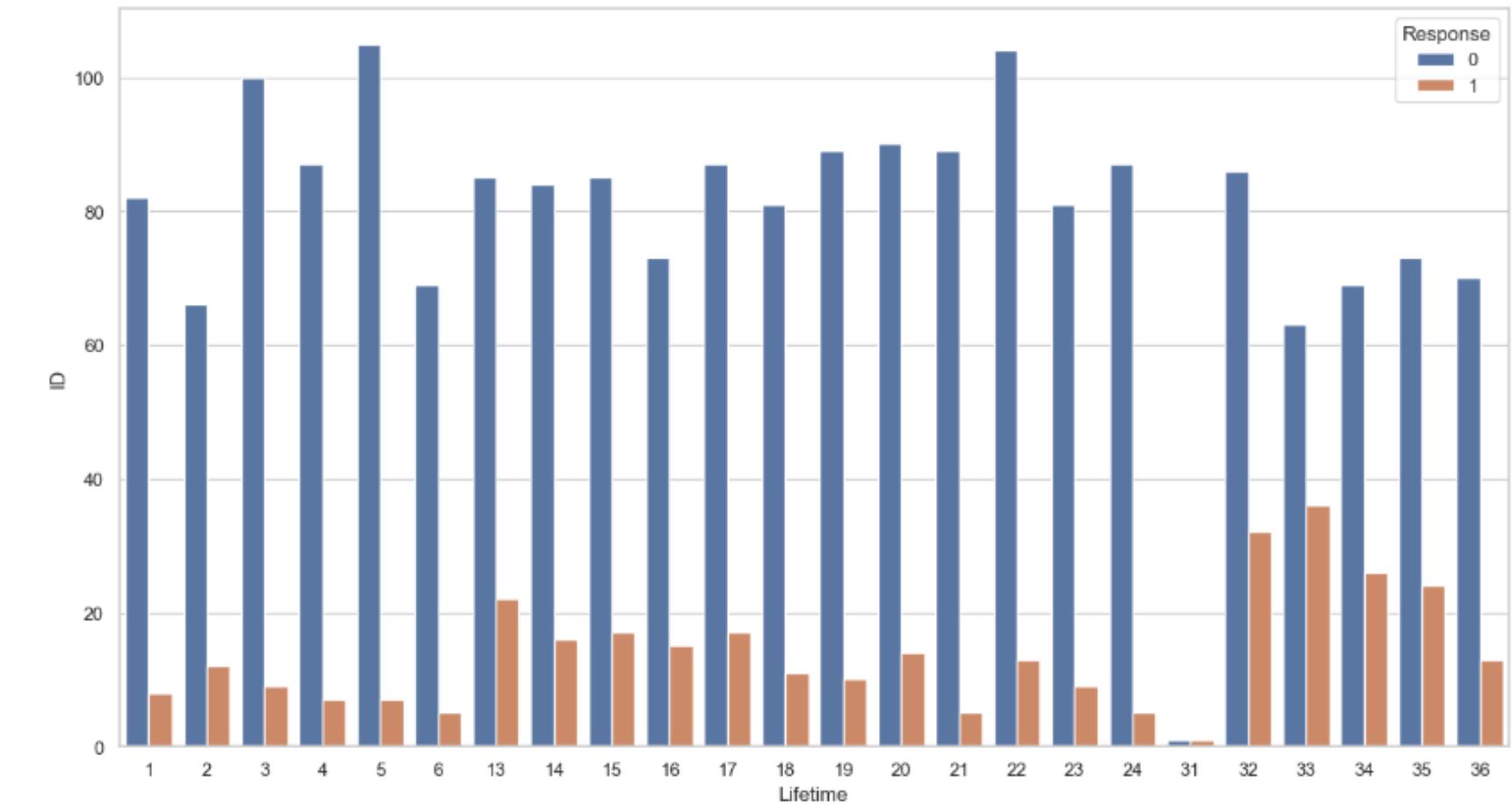
Total Accepted Campaign by Customer >=2 (from 5 campaigns)

Number Customer based on Total Accept Campaign



Paling banyak pada Lima Campaign kita adalah 0 (tidak pernah merespon), namun ada yang sedikit berpotensi pada, hanya sekali (1) atau dua kali (2) merespon masing-masing 325 dan 83 Customers

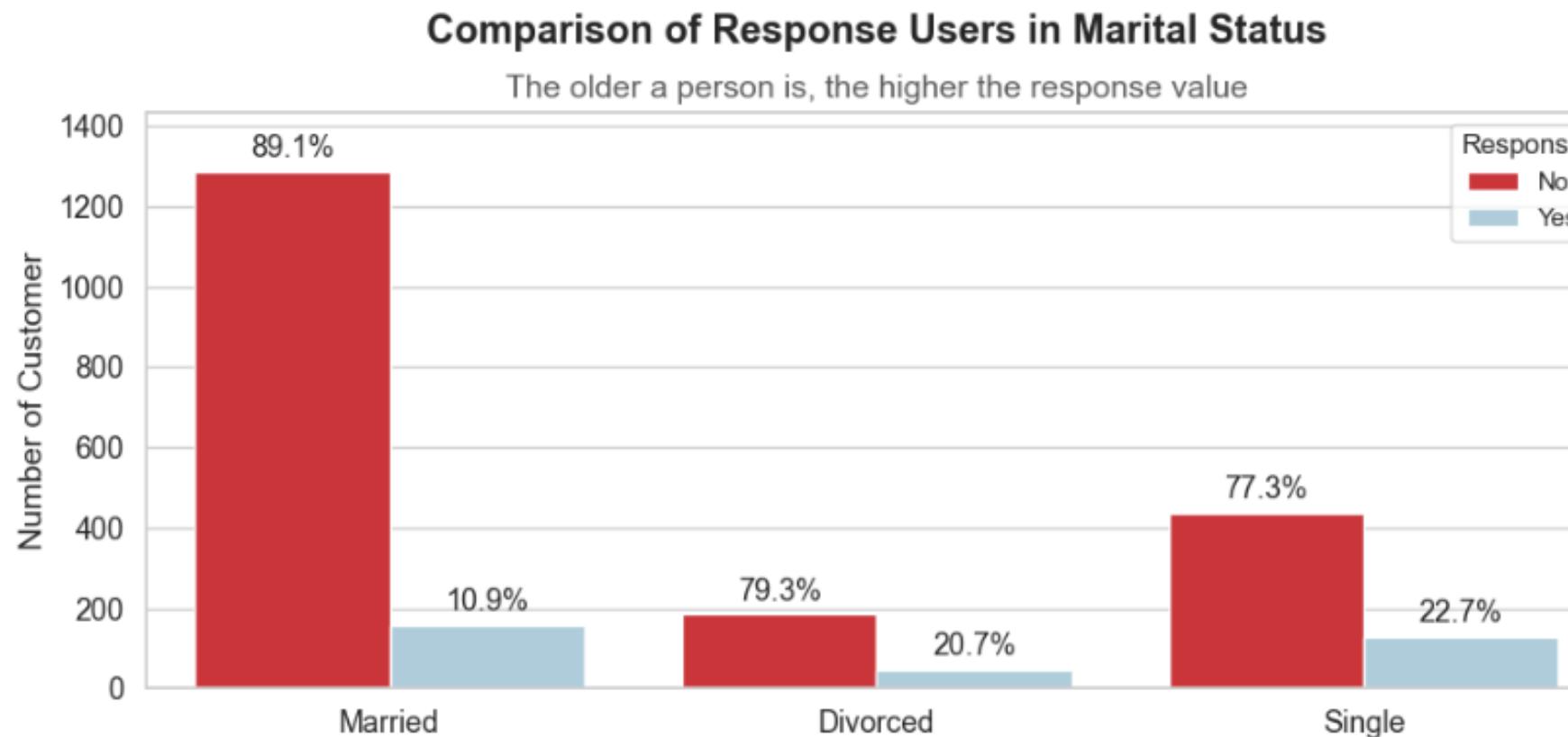
Lifetime minimal 32 - 35 months



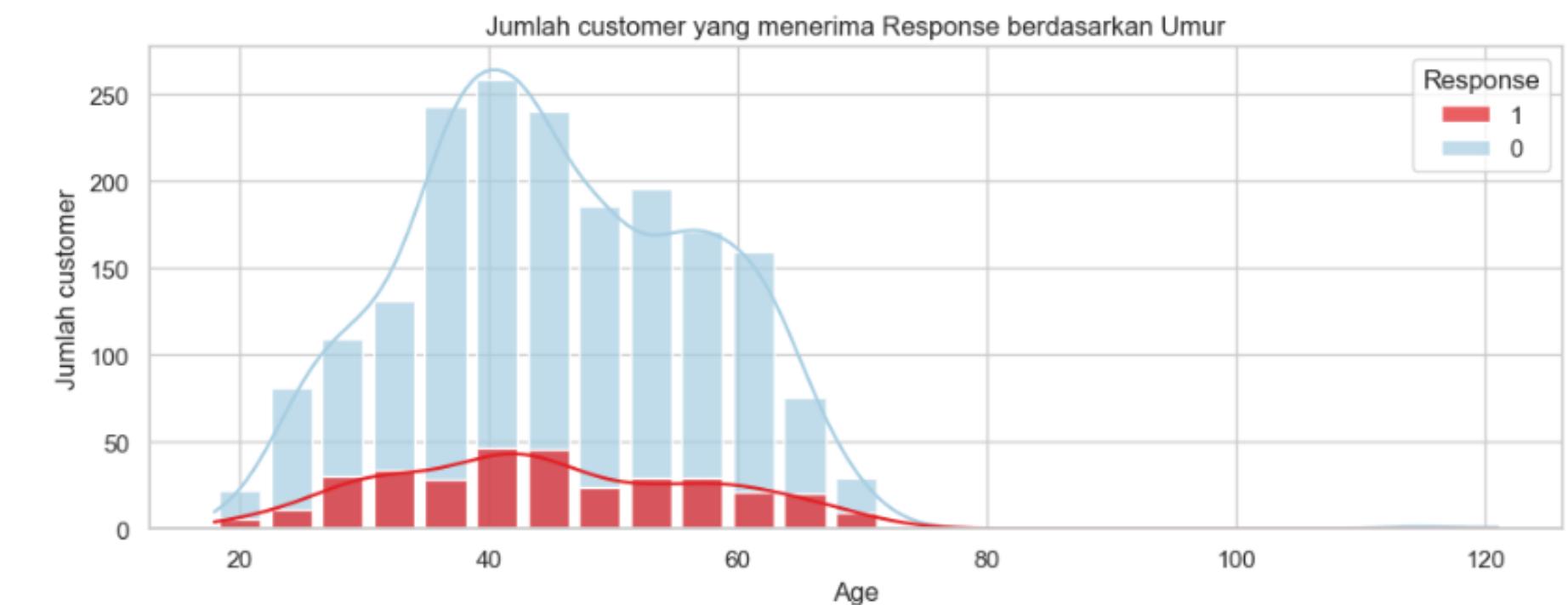
Customer yang memiliki lifetime yang tinggi cenderung memberi respon

Business Insights

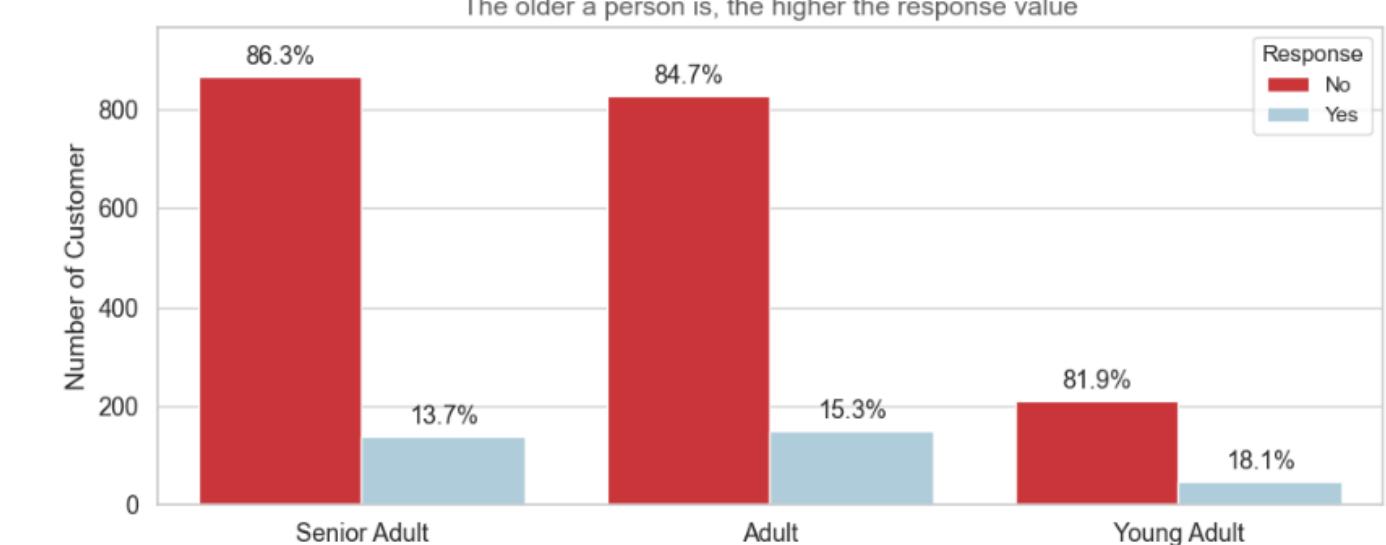
**Marital Status =
Married > Single > Divorced**



**Age &
Age Group (dominated by
Senior Adult group)**



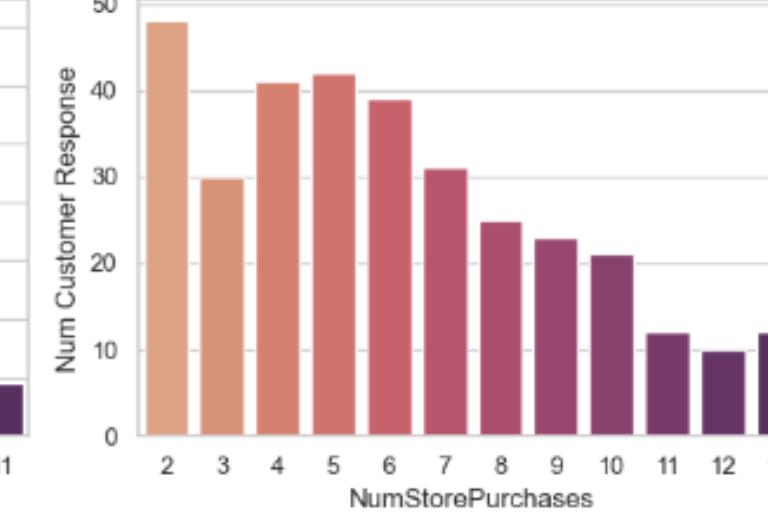
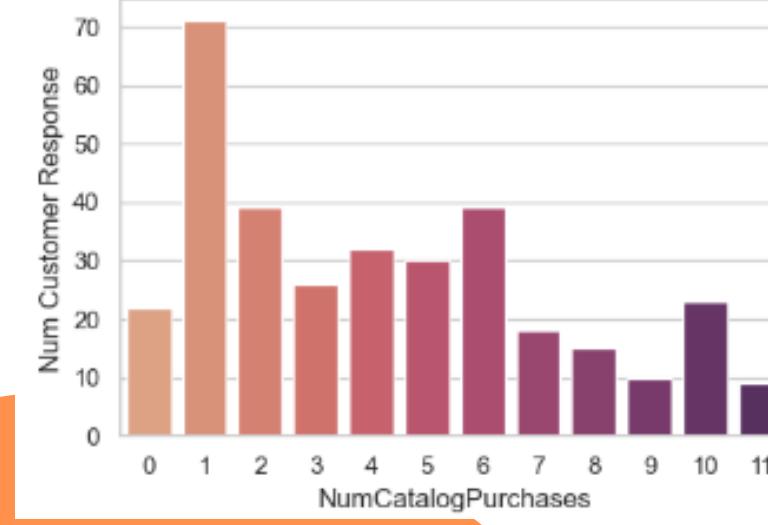
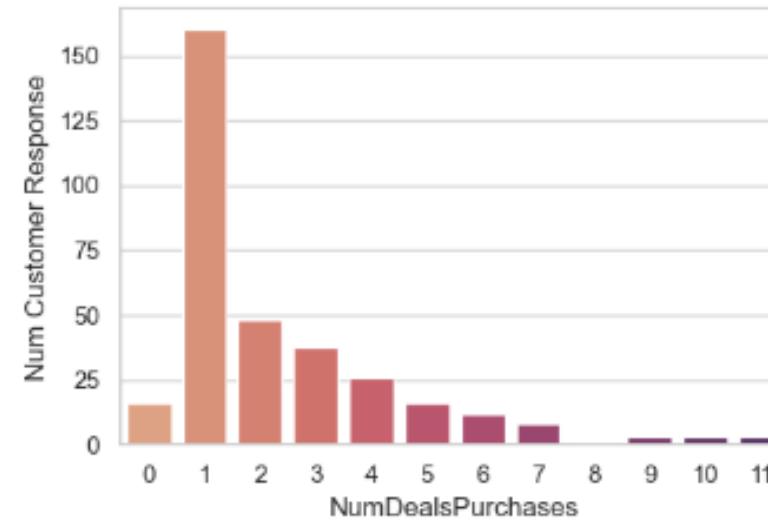
Comparison of Response Users in Age Group



Business Insights



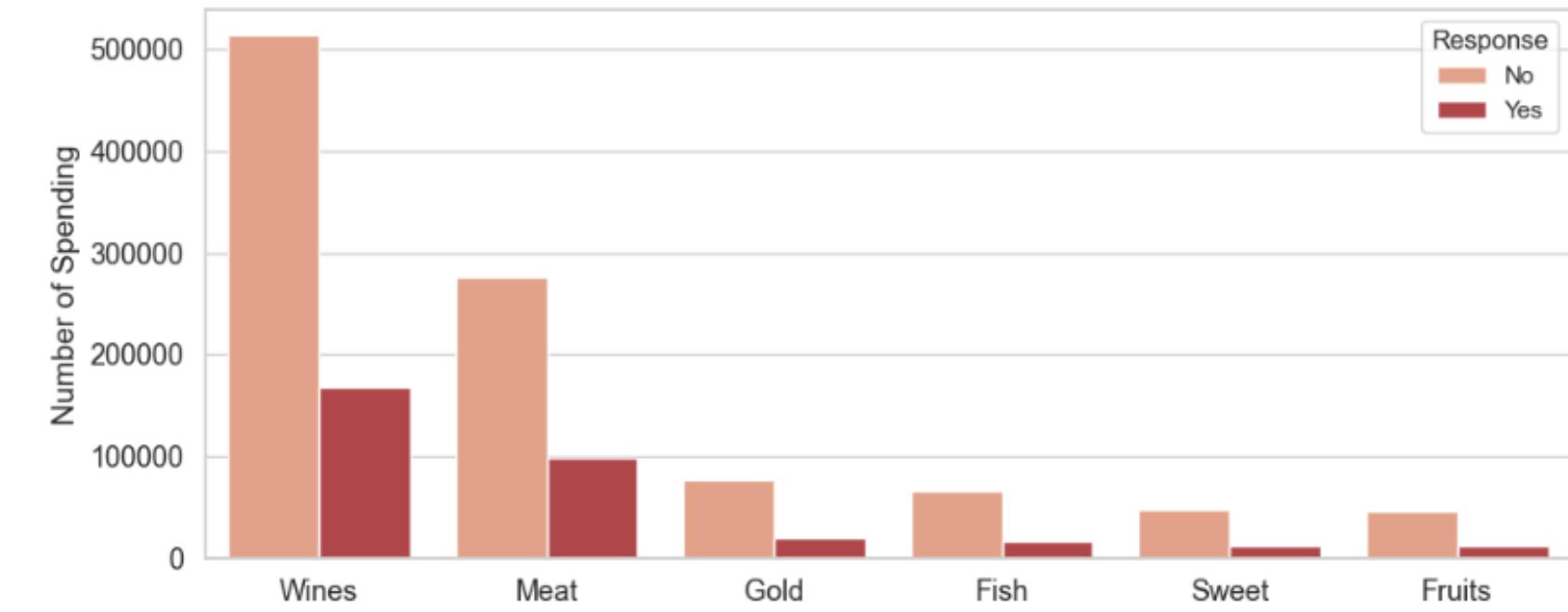
**Responded Customer Purchase Types
(the fewer purchases, the more likely it is)**



**Types of Purchased Products
(customers tend to buy Wines and Meat)**



Comparison of Response Users in Spending Product
Response customers mostly use Wines and Meat Product



Datalicious - Lotto Mart

Data Preprocessing

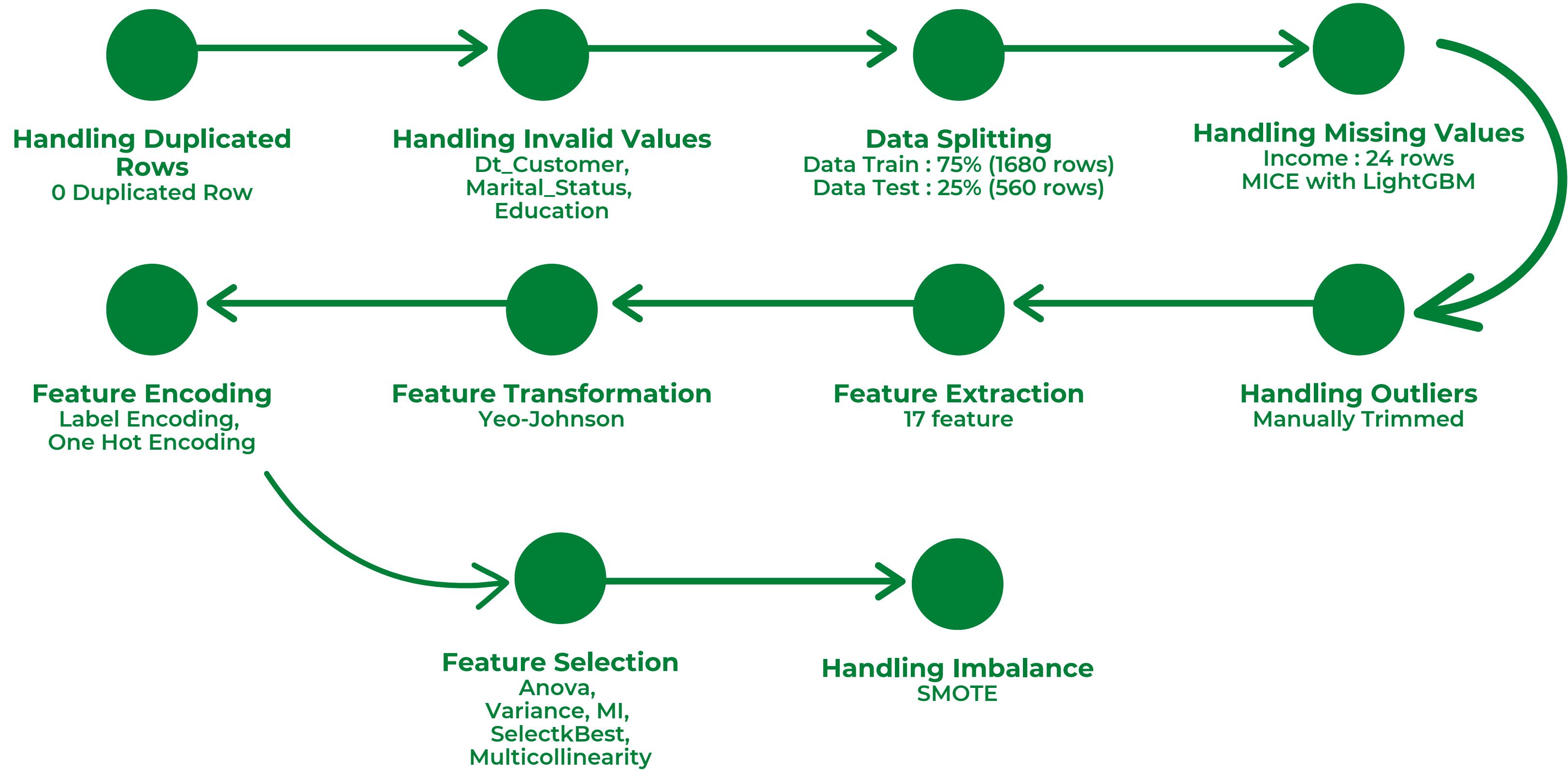
- Handling Missing, Invalid
- Feature Transform & Engineering
- Feature Selection
- Handling Imbalance Data



DATA PRE-PROCESSING

CLASSIFICATION

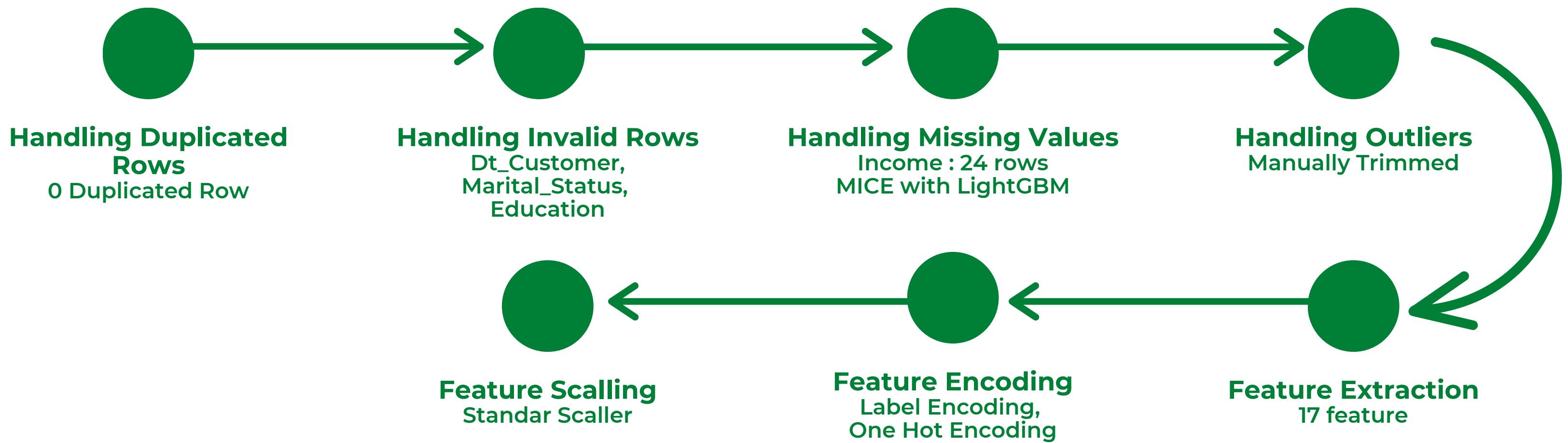
DATA LICIOUS



DATA PRE-PROCESSING

CLUSTERING

DATALLICIOUS



Datalicious - Lotto Mart

Modeling and Evaluation

- Parameter Evaluasi Model
- Machine Learning Techniques
- Model Comparison
- Feature Importance



Modeling and Evaluation

Parameter Evaluasi Model

DATALICIOUS

Positive = Response Campaign
Negative = Tidak Response Campaign

False Positive

Main target to be reduced

False Negative

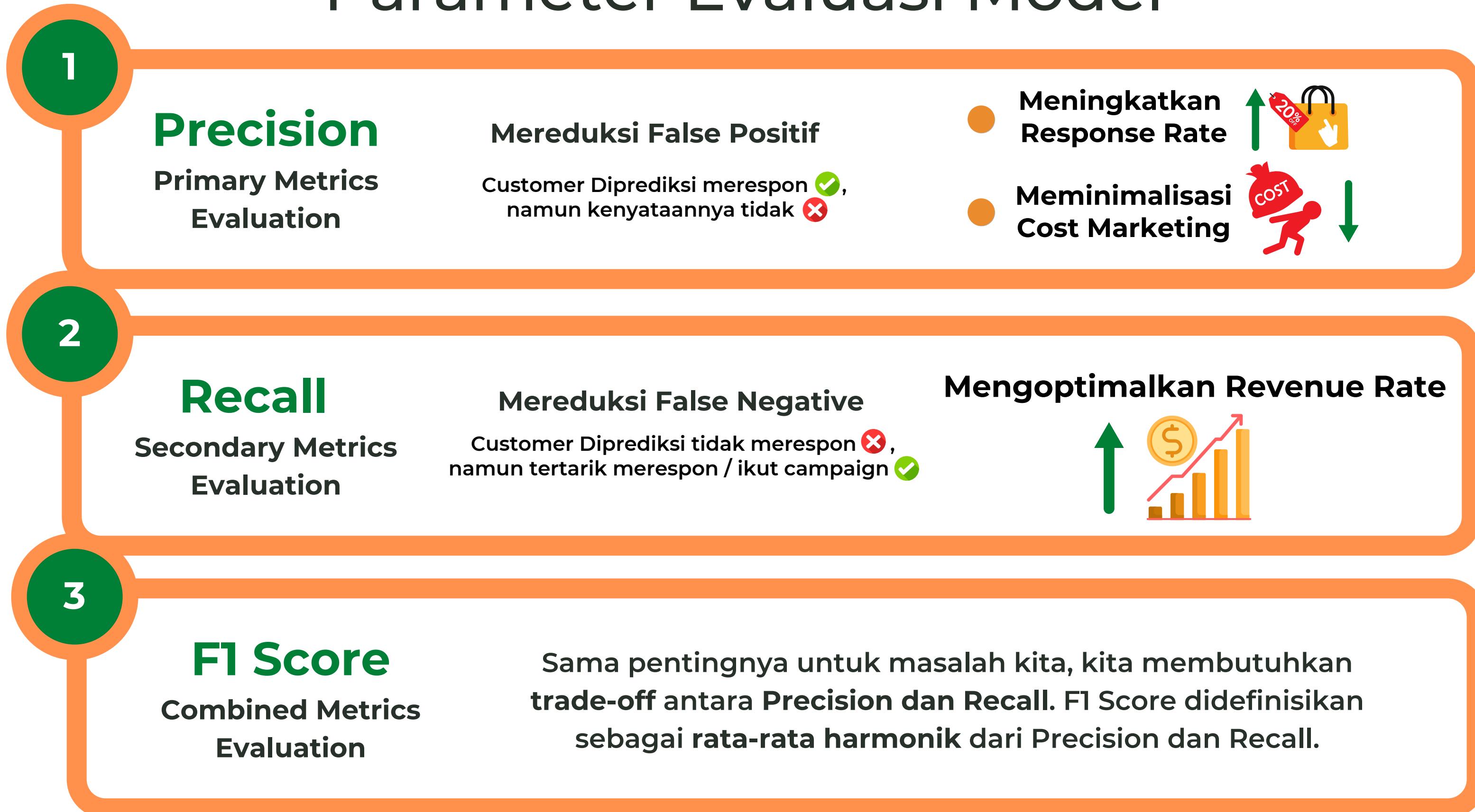
Second target to be reduced



Modeling and Evaluation

Parameter Evaluasi Model

DATA LICOUS



Modeling and Evaluation

Machine Learning Techniques

-
- 1. Decision Tree
 - 2. Random Forest 
 - 3. Logistic Regression
 - 4. Gaussian Naive Bayes
 - 5. K-Nearest Neighbor
 - 6. MLP Classifier (Neural Network)
 - 7. Adaboost Classifier
 - 8. XGBoost Classifier
 - 9. Gradient Boosting Classifier
 - 10. Support Vector Machine

 **Random Forest** 
Best Fit Models



Modeling and Evaluation

Model Comparison

Model (Test)	Accuracy	Precision	Recall	F1 Score	Cross Val F1 (k=5)	ROC AUC	Cross Val ROC AUC (k=5)
Random Forest	0.909	0.786	0.530	0.633	0.519	0.905	0.896
XGBoost Classifier	0.889	0.652	0.542	0.592	0.525	0.895	0.887
Logistic Regression	0.861	0.523	0.675	0.589	0.510	0.900	0.885
Support Vector Machine	0.859	0.520	0.614	0.564	0.448	0.895	0.884
K-Nearest Neighbors	0.845	0.482	0.639	0.549	0.400	0.842	0.810
Adaboost Classifier	0.855	0.511	0.578	0.542	0.539	0.892	0.885
MLP Classifier	0.846	0.483	0.518	0.500	0.558	0.830	0.887
Gradient Boosting Classifier	0.846	0.482	0.494	0.488	0.487	0.725	0.752
Naive Bayes	0.830	0.441	0.542	0.486	0.451	0.838	0.824
Decision Tree	0.838	0.456	0.494	0.474	0.488	0.699	0.718

Nilai Precision, Recall dan F1 Score pada hasil evaluasi yang paling baik dihasilkan oleh

Random Forest, XGBoost Classifier, dan Logistic Regression



Modeling and Evaluation

Model Comparison

Models	Precision (Train)	Precision (Test)	Recall (Train)	Recall (Test)	F1 Score (Train)	F1 Score (Test)	Total Diff
Logistic Regression (HT)	0.755	0.528	0.660	0.675	0.704	0.593	0.323
Random Forest (HT3) ✓	0.867	0.702	0.584	0.482	0.698	0.571	0.394
Decision Tree (HT2)	0.719	0.441	0.574	0.542	0.639	0.486	0.463
XGBoost Classifier (HT2)	0.826	0.595	0.699	0.566	0.757	0.580	0.541
K-Nearest Neighbors (HT)	0.763	0.465	0.871	0.639	0.814	0.538	0.806
XGBoost Classifier (HT)	0.976	0.710	0.920	0.530	0.947	0.607	0.996
Random Forest (HT2)	0.972	0.724	0.912	0.506	0.941	0.596	0.999
Random Forest (HT1)	0.997	0.772	0.993	0.530	0.995	0.629	1.054
Decision Tree (HT)	1.000	0.484	0.990	0.554	0.995	0.517	1.430

Setelah dilakukan Hyperparameter Tuning

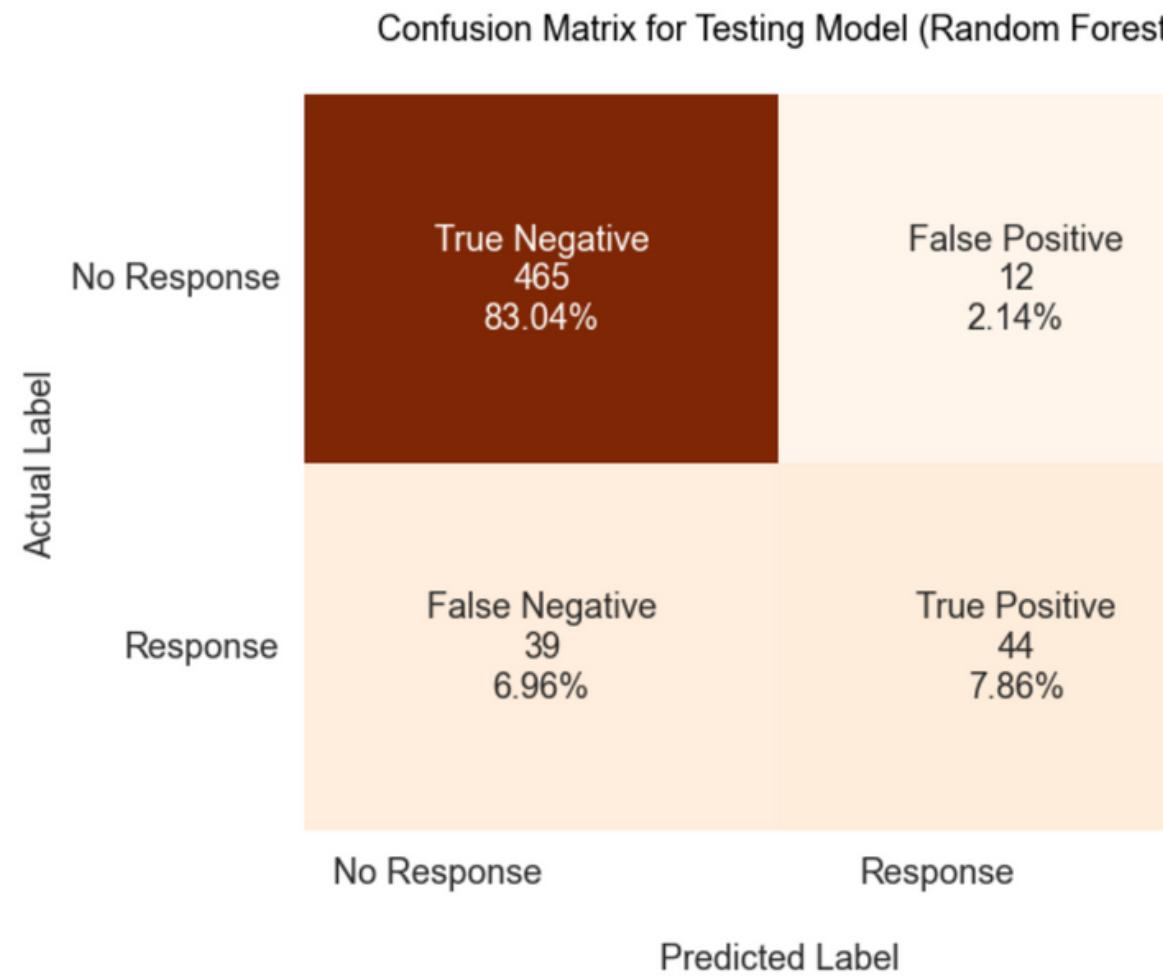
Nilai Total Diff (Gap) pada dataset train dan test yang paling stabil dengan metrics yang lumayan tinggi adalah pada Random Forest



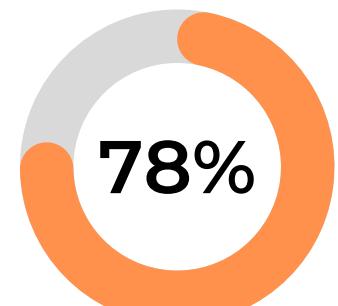
Modeling and Evaluation

Model Selection

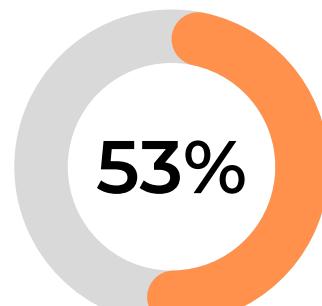
DATALLICIOUS



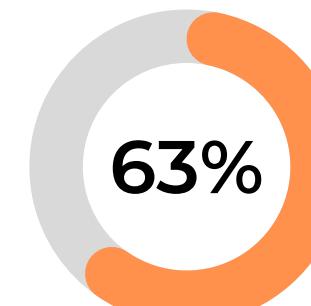
Precision



Recall



F1 Score



True Positif	Diprediksi merespon ✓	benar, merespon ✓
True Negative	Diprediksi tidak merespon ✗	benar, tidak merespon ✗
False Positif	Diprediksi merespon ✓	salah, tidak merespon ✗
False Negative	Diprediksi tidak merespon ✗	salah, merespon ✓

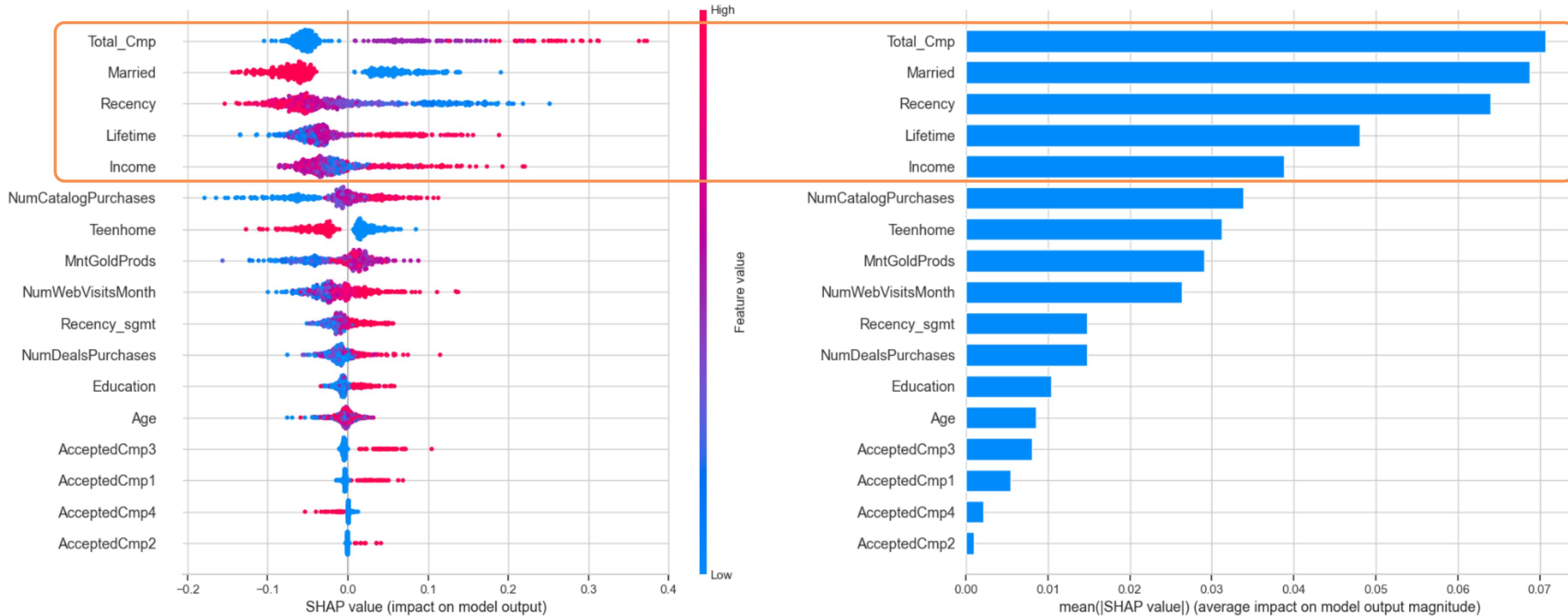


Modeling and Evaluation

Feature Importance



DATA LICIOUS



Positive = Response Campaign



Negative = Tidak Response Campaign

05 / 10

Datalicious - Lotto Mart

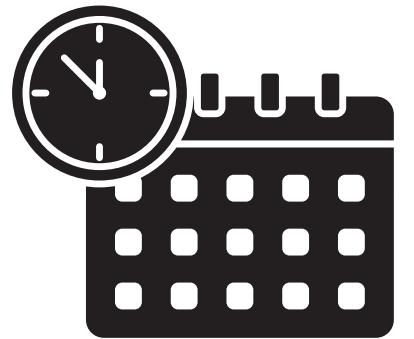
Modeling and Evaluation

- Clustering with RFM Analysis
- Clustering Results



Modeling and Evaluation

Clustering with RFM Analysis



Recency

The freshness of customer activity



Resources : the number of days since the last purchase



Frequency

The frequency of the customer transactions



Resources : the number of transactions made from various types of purchases



Monetary

The intention of customer to spend



Resources : the amount spent on various products

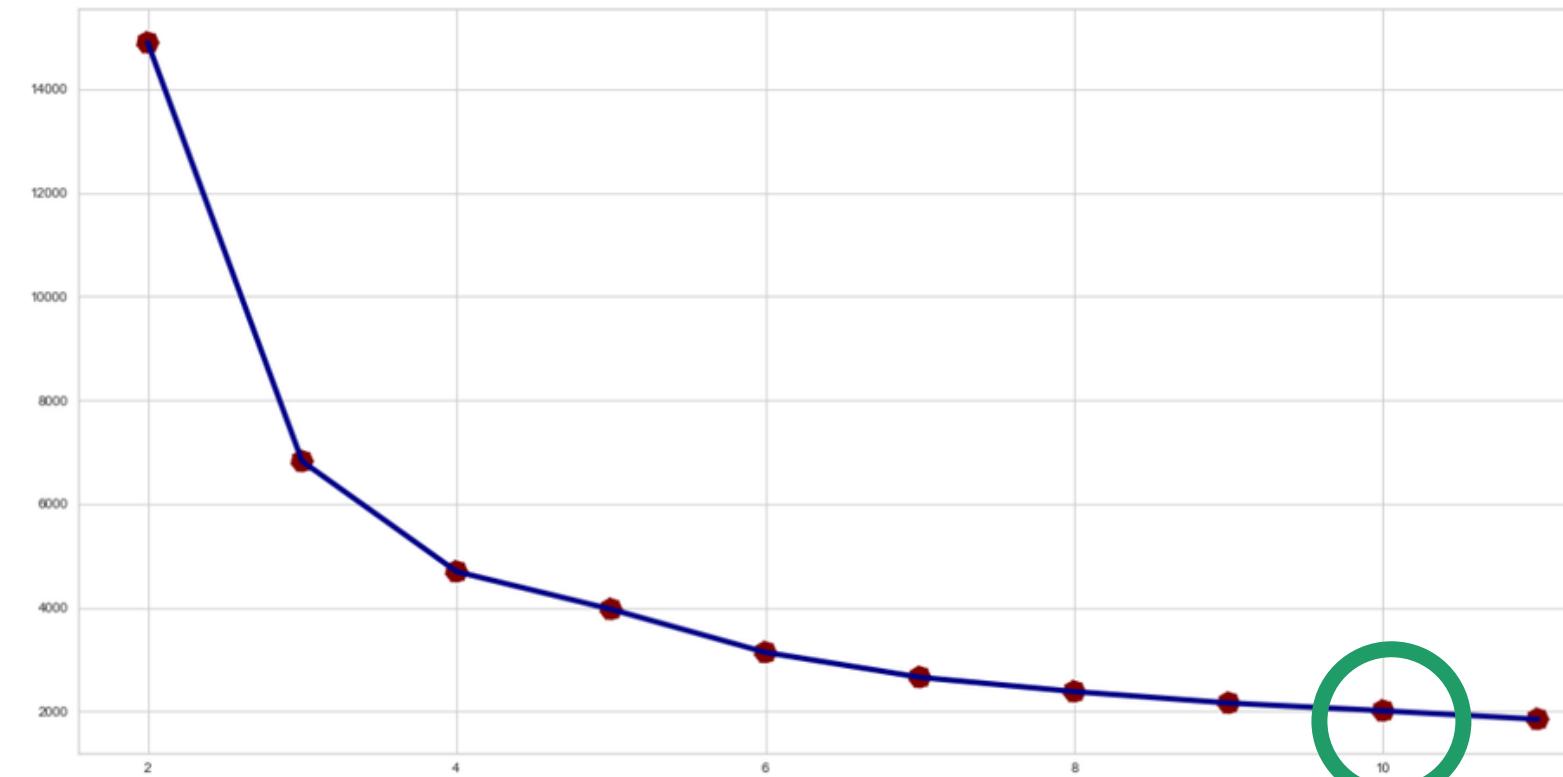
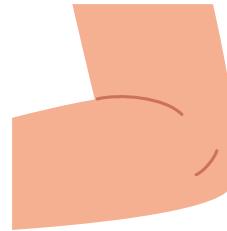


Modeling and Evaluation

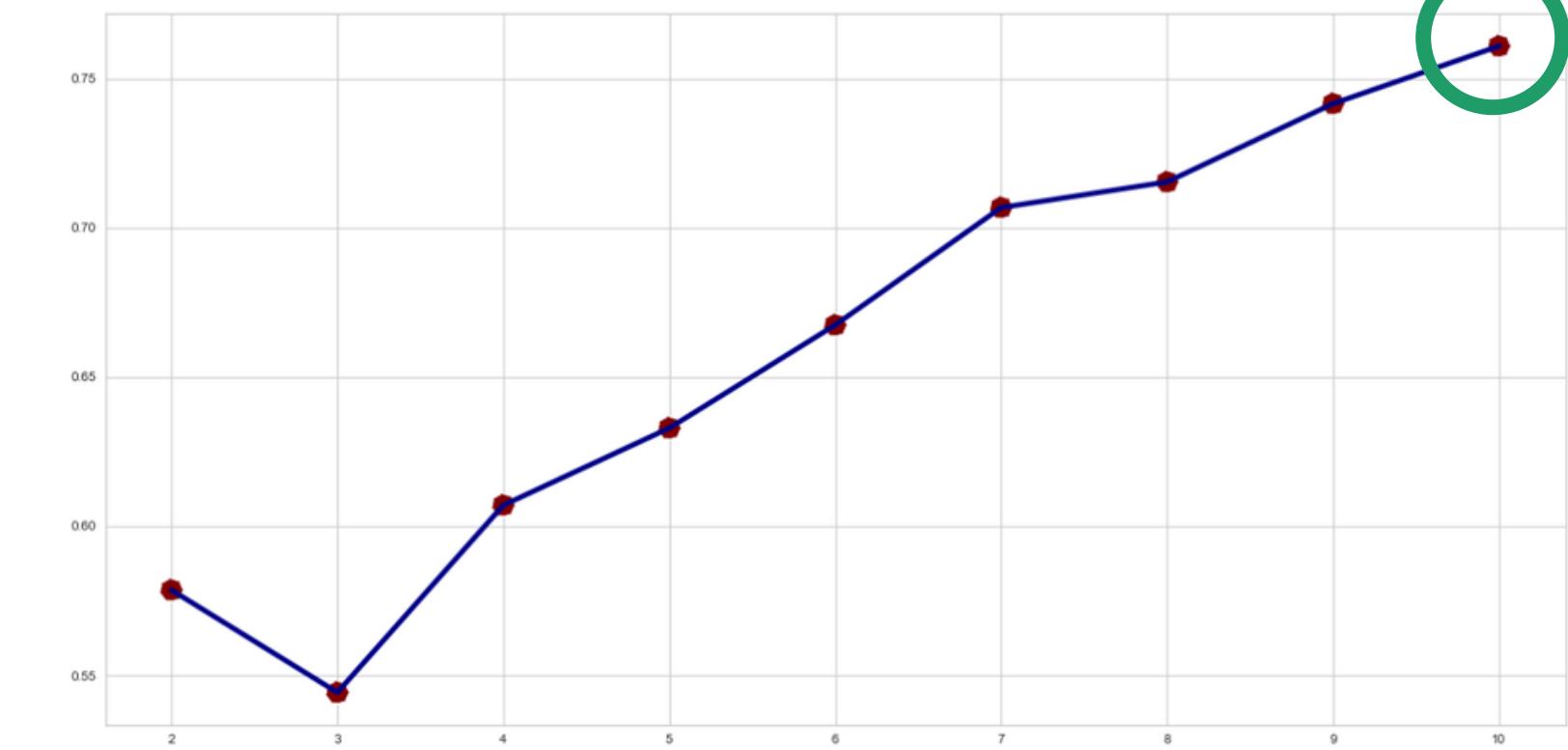
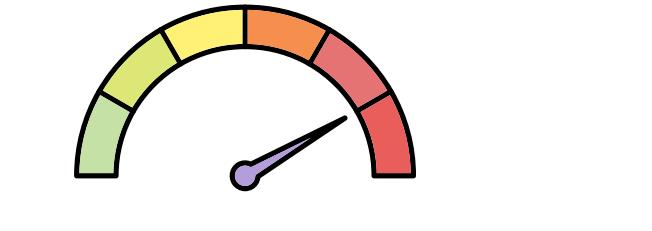
Clustering with RFM Analysis

DATALLICIOUS

Elbow Method



Silhouette Score



Customers will be divided into 10 clusters with the average silhouette score is 0.76



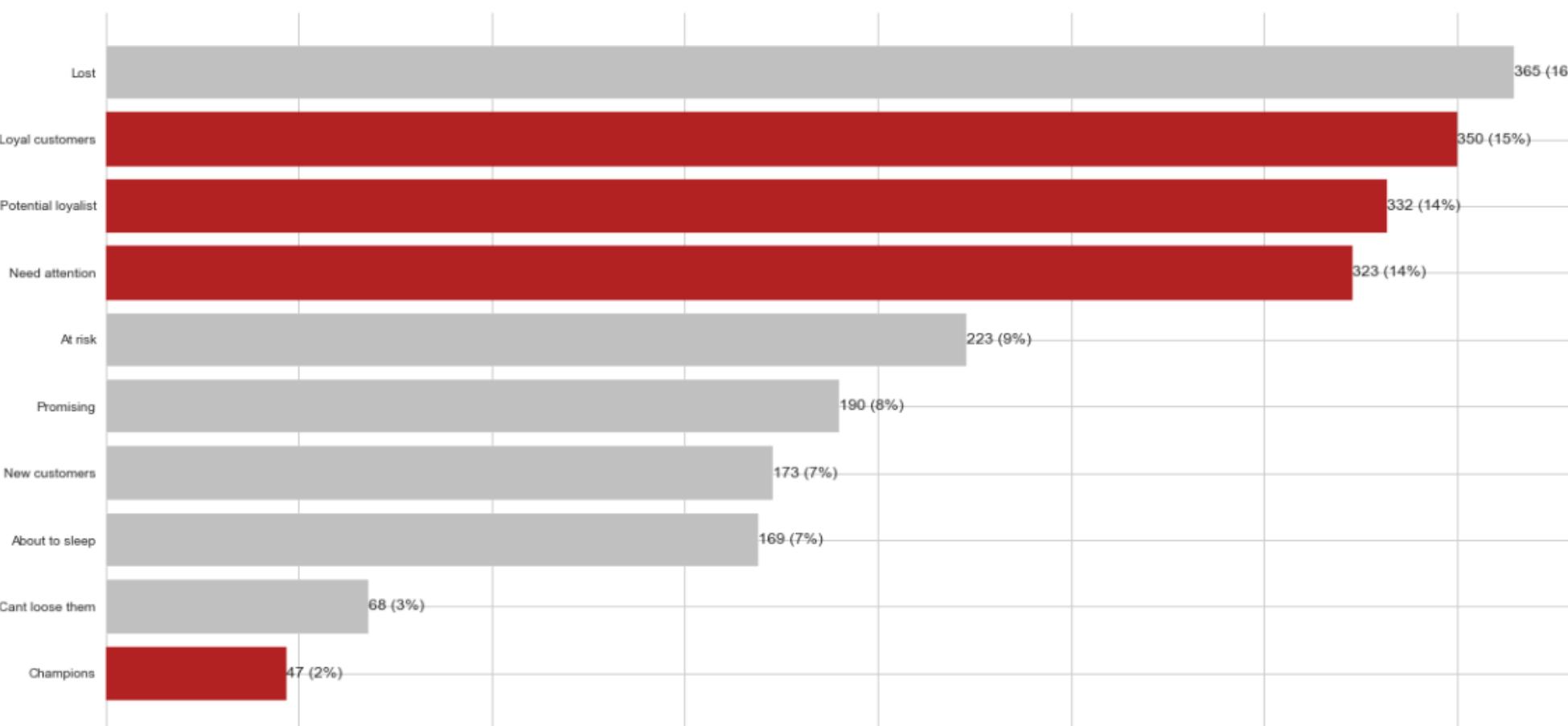
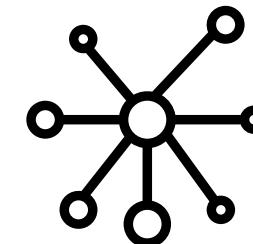
Modeling and Evaluation



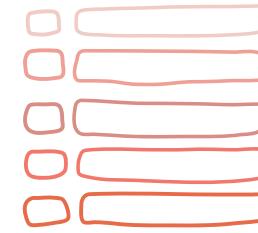
Clustering Results

DATA LICKIOUS

Cluster Distribution



Cluster Type



Segment	Recency	Frequency	Monetary
About to sleep	medium	low	low
At risk	high	medium	high
Cant loose them	high	high	high
Champions	low	high	medium
Lost	high	medium	medium
Loyal customers	medium	high	high
Need attention	medium	medium	medium
New customers	low	low	low
Potential loyalist	low	medium	medium
Promising	medium	low	low

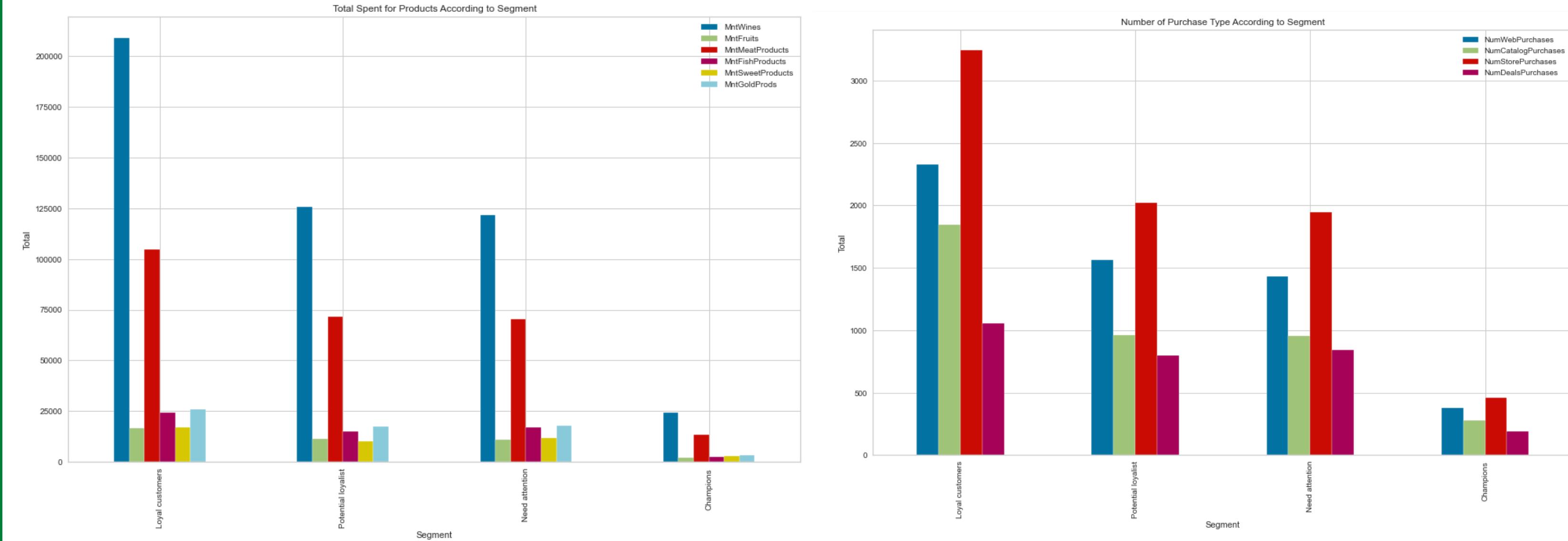
In supporting goals to get business recommendations, we focus on 4 customer segments with a total of 1052 customers.



Modeling and Evaluation

Clustering Results

DATALLICIOUS



The most purchased product is wines and The most frequent type of purchase is store purchases



Datalicious - Lotto Mart

Business Recommendation

- > Recommendation
- > Business Impact
- > Simulation



Business Recommendations



Discount/Flash Sale

price reduction with a certain period of time

especially **for customers who have never response a campaign or 1x response the campaign only**

Goals:

to **increase** customers shopping/purchasing interest



Vouchers/Rewards

especially **for customers who have response the campaign at least 2 campaigns**

customers will get coupons and can be exchanged for discounts or other benefits such as points, free shipping, free gifts, etc.

Goals:

maintain customer's interest and retention



Promo Bundling/Special Offer

can implement **for all customers**

combining the main product with supporting products for more economical price for customers such as Gold Products + Wines, Wines + Meats, etc.

Goals:

encourage customers to buy more products

Business Recommendations





Business Simulation

Business metrics are quantifiable measures used to track business processes to judge the performance level of your business.

Response Rate

ROI

Net Profit Margin

09 / 10



Response Rate

14,91%

before model



63,66 %

True Positive (TP)= 44
False Positive (FP)= 12

response rate = TP/(TP+FP)

78,57%

after model

Net Profit Margin / NPM

-82,90 %

Before Modelling

148,19

65,29 %

After Modelling



Total Response (1=accept response)	334			Total Campaign (TP+FP)	56
Total Campaign (all customers)	2239			Total Cost (\$)	168
Total Cost (\$)	= Total campaign*cost = 2239 * 3	6717		Total Revenue (\$)	484
Total Revenue (\$)	= Total response * revenue = 334 * 11	3674	a customer	Total Profit (\$)	316
Total Profit (\$)	= Total Revenue - Total Cost = 3674 - 6717	-3043	Revenue (\$)	NPM (%)	65.29
NPM (%)	= (Total Profit/ Total Revenue) *100	-82.83	Cost (\$)		

Return of Investment

Average Order Size (AOS)	= Total Revenue/Total Customer = $3674/2239$	1.64
Average Order Frequency (AOF)	= Total Order/Total Customer = $2239/2239$	1
Average Customer Value (ACV)	= AOS/AOF	1.64
Average Customer Lifetime (ACL)	= first order date-last order date	2
Customer Lifetime Value (CLV) (\$)	= ACL*ACV	3.28



ROI > 1.0x

Before Modelling		
number of new customers	12	
CAC	= Total Marketing Cost/Number of new Customers = $6717 / 12$	559.75
CLV		3.28
ROI	= CLV : CAC	0.006

After Modelling		
number of new customers	56	
CAC	= Total Marketing Cost/Number of new Customers = $(TP * cost) / 56$ = $(44 * 3) / 56$	2.36
CLV		3.28
ROI	= CLV : CAC	1.39

Thank You

END OF SLIDE >

