



# Campaign Performance and Marketing Effectiveness Analysis Case Study: TokoBli E-commerce

Venny Amilia Deslaweny

Data Analyst



[www.linkedin.com/in/vennydeslaweny](https://www.linkedin.com/in/vennydeslaweny)



[www.linkedin.com/in/vennydeslaweny](https://www.linkedin.com/in/vennydeslaweny)



[ve.200501@gmail.com](mailto:ve.200501@gmail.com)



# Executive Summary

This analysis evaluates the effectiveness of marketing campaigns and customer purchasing behavior in TokoBli's e-commerce data.

Key findings:

- Certain campaigns generated higher revenue and better efficiency.
- Customer purchasing behavior varies across product categories.
- Discount strategy influenced sales performance but requires optimization.
- Targeted promotions can improve campaign ROI and customer retention.

# TL Sim 1



# 1. Business Understanding

## Business Question

- A. **Problem Statement,** TokoBli aims to evaluate the effectiveness of three major campaign strategies (10/10, 11/11, and 12/12) to identify which campaign is the most optimal in increasing revenue and budget efficiency, while also considering the number of transactions, new customers, and total products sold.

## Insight

**Business Problem Summary,** TokoBli needs to determine which campaign among 10/10, 11/11, and 12/12 is the most effective in maximizing total revenue, optimizing budget efficiency, and acquiring new customers as a foundation for data-driven budget allocation in future campaigns.

[Use this Data Set Analysis](#)

# 1. Business Understanding

## Business Question

B. identify and select the most relevant business metrics from the available data that best address the problem statement (e.g. average evenue per campaign, total transactions, etc)

## Insight

The main step after defining the problem statement is to determine the most relevant business metrics to answer analytical questions and support strategic decision-making within the company. The selected metrics are measurable, relevant, and based on actual business outcomes, making them suitable for quantitatively and objectively assessing the effectiveness of each campaign strategy.

The combination of **outcome metrics** (such as *Total Revenue* and *Transactions*) and **efficiency driver metrics** (such as *Revenue-to-Budget Ratio* and *Average Discount*) ensures that the analysis remains comprehensive covering both performance results and efficiency aspects.

The insights derived from these metrics will serve as the foundation for developing strategic recommendations for TokoBli's management, particularly in:

- Allocating promotional budgets more effectively, and
- Designing future campaign strategies that are more data-driven (*data-driven marketing optimization*).

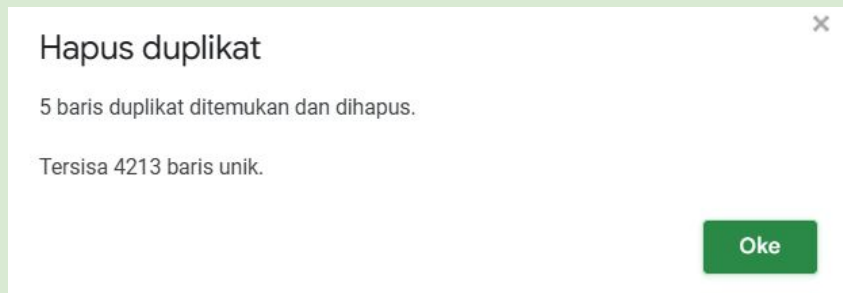
## 2. Clean Up The Data

### Business Question

A. Check and handle duplicate values

### Insight

- Duplicate and rename dataset, highlight all cells in the duplicated dataset, Data Cleaning:
  - Removed duplicate records
  - Handled missing values
  - Standardized data types



## 2. Clean Up The Data

### Business Question

#### B. Handle Missing Data

### Insight

Highlight all cells, click ribbon menu Data > Create Filter, check all columns one by one and found

1. In **Transaction** column found 5 blank rows, delete 5 rows
2. In **Status** found 4187 blank rows (more than 50% of the total rows are missing),so we delete the column instead
3. In **Price** column found 1 blank rows, decided to delete the 1 rows since we do not know the price
4. In **QTY** column found 1 blank rows, decided to delete the 1 rows since we do not know the qty
5. In **Discount** column found 1 blank rows, decided to delete the 1 rows since we do not know the discount

## 2. Clean Up The Data

### Business Question

C. Check and handle, remove duplicat, and ensure correct column data type

### Insight

Notice there are strange numbers in **Prince** and **Discount**

1. Highlight all cells in **Prince** and **Discount** columns
2. Reformat data type by clicking ribbon menu Format > Number > Automatic since it will change to the majority of data type in each column

Change the decimal separator in the **Total Revenue**

1. Highlight all cells in **Total Revenue** columns
2. Reformat data type by clicking ribbon menu Format > Number > Automatic since it will change to the majority of data type in each column

### Reasoning

- Since we want to calculate the amount of Rupiah spent on the product, it is essential to change the data type in a format that is easily readable and apprehensible



# 3. Descriptive Analytics

## Business Question

- A. Create Statistical measurement to know the data distribution by using XL Miner (QTY)

## Insight

QTY					
1	QTY				
1					
2	Mean	1.289211	Mean (Average): 1.29. The majority of transactions involve only one product per transaction, as indicated by both the median and mode values of 1. This means that most customers purchase only a single item per transaction, resulting in a relatively low purchase volume per customer.		
1	Standard Error	0.018929			
1	Median	1			
1	Mode	1			
1	Standard Deviation	1.227898			
1	Sample Variance	1.507732			
2	Kurtosis	299.8942			
2	Skewness	14.79453	A very high positive skewness indicates that the distribution is heavily skewed to the right, meaning most transactions fall on the lower end (1-2 QTY), with only a few transactions involving very large purchase quantities.		
1	Range	30			
1	Minimum	1	Most transactions involve a very small number of products, with a minimum of 1 product per transaction; however, there are a few extreme transactions (outliers) where customers purchased up to 31 products at once.		
1	Maximum	31			
1	Sum	5425	A total of 5,425 products were successfully sold. There were 4,208 individual transactions recorded during the campaign period.		
1	Count	4208			
1					
1					
2					

# 3. Descriptive Analytics

## Business Question

- A. Create Statistical measurement to know the data distribution by using XL Miner (DISCOUNT)

## Insight

1	Discount		
2	599000	Discount	
3	0		
4	44000	Mean	83263.37928
5	0	Standard Error	2921.793696
6	0	Median	0
7	0	Mode	0
8			
9	0	Standard Deviation	189534.1244
10	0	Sample Variance	35923184328
11	432000	Kurtosis	5.267116662
12			
13	370000	Skewness	2.468335294
14			
15	299000	Range	900000
16	0	Minimum	0
17	0	Maximum	900000
18	0	Sum	350372300
19	0	Count	4208
20	0		
21	90000		
22	0		
23	0		

The average discount per transaction is around Rp83,263, suggesting that overall, the discount value offered remains moderate.

The distribution of discount values is highly dispersed, there is a large variation between transactions with no discount and those with substantial discounts.

The high positive skewness indicates a strongly right-skewed distribution, most transactions had no discount, but a few involved large discounts (outliers).

The difference between the smallest and largest discount reaches Rp900,000, indicating the presence of extreme promotions on a small portion of products.

Extra Miles

# 3. Descriptive Analytics

## Business Question

- A. Create Statistical measurement to know the data distribution by using XL Miner (TOTAL REVENUE)

## Insight

	A	B	C	D	E	F
1	Total Revenue					
2	5391000	Total Revenue				
3	3990000					
4	836000	Mean	4696327.322	Average revenue per transaction ~ Rp4.7 million. Indicates healthy sales level overall.		
5	3990000	Standard Error	63199.30981			
6	4320000	Median	4230000			
7	8990000	Mode	7200000			
8	3990000	Standard Deviation	4100169.383	High variability, large differences between small and big transactions.		
9	2990000	Sample Variance	16811388969681			
10	8208000	Kurtosis	152.8963992			
11	7030000	Skewness	9.466341377	Strong positive skew → few transactions generate extremely high revenue.		
12	2691000	Range	83511000			
13	2000000	Minimum	45000			
14	7400000	Maximum	83556000			
15	6990000	Sum	19766841700	Total revenue across all campaigns ~ Rp19.77 billion.		
16	2300000	Count	4209			
17	490000					
18	8990000					
19	1710000					
20	3050000					

Extra Miles

# 3. Descriptive Analytics

## Business Question

B. Identity and Remove outliers base on “Total Revenue” Q1, Q3, IQR, Upper Inner Fence, Lower Inner Fence

## Insight

Q1	Q3	IQR	Upper Inner Fence	Lower Inner Fence
2280000	6950000	4670000	15122500	-24184375

We will remove outlier in “Total Revenue” Column.

Reasoning : we believe that it has direct relationship to all campaign

Based on the IQR method, we **delete the rows** that have totalNumofPurchase **above 12** and below LIF. But negative value in NumofPurchase is irrational, therefore we treat the LIF as 0. We prefer to use the inner fence (1.5 IQR) than the outer fence (3 IQR) as it already satisfies the outlier removal. Total of 2 rows removed.

**Total rows after outlier removal 4198**

Extra Miles

# 3. Descriptive Analytics

## Business Question

C. Top 3 business insight

## Insight

**1. Most transactions are of low value**, yet they generate substantial total revenue thanks to a small group of **high-value customers (high spenders)**.

**Business implication:** TokoBli has significant potential to increase revenue if it can extend the purchasing behavior of these high spenders to regular customers through **upselling or bundle offer strategies**.

**2. The discount strategy is efficient but not yet optimal in driving transaction volume.**

**Business implication:** This approach helps maintain profit margins, but promotional effectiveness could be improved through more targeted discounts, for example, **personalized offers for new customers** or **underperforming products**.

**3. Campaign performance is strong but uneven, total revenue is driven by a small segment of customers and large transactions.**

**Business implication:** For long-term growth, TokoBli should focus on **revenue stability** by expanding its base of active customers rather than relying solely on big spenders during campaign periods.

Extra Miles

# TL Sim 2

LINK SPREADSHEET



# 4. EDA & Statistical Measurement

## Business Question

- A. Determine whether there are different total transaction, total customers, total products sold, total revenue, and total discount campaign budget among the three campaign period

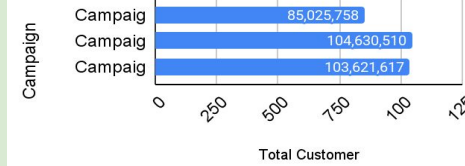
## Insight

We use the **total SUM** of amount spent on each product in pivot table and plot it into a bar chart and found that the five tables beside, we can conclude that:

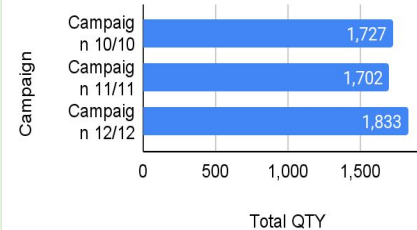
- The **11/11 campaign** had the highest number of **transactions** and **customers**, but also the largest discount budget, resulting in lower margin efficiency.
- The **12/12 campaign** was more efficient in terms of **revenue** compared to its **discount budget**, as it achieved almost the same total revenue as the 11/11 campaign but with a much smaller discount, indicating higher efficiency. Has the highest **quantity (QTY)** in this campaign.
- The **10/10 campaign** showed moderate performance and can be used as a baseline reference.

## Extra Miles

Total Customer vs Campaign

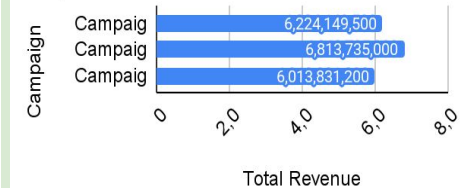


Total QTY vs Campaign

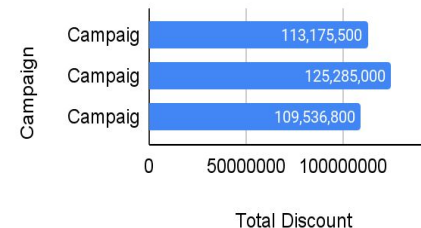


Total Revenue vs Campaign

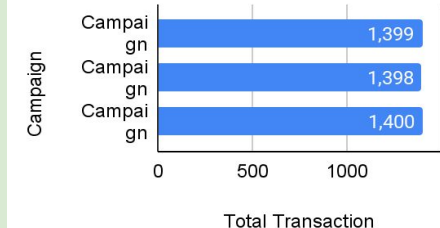
in Rp.



Total Discount vs Campaign



Total Transaction vs



# 4. EDA & Statistical Measurement

## Business Question

B. Evaluate differences in the ratio of the Revenue to discount campaign budget for each campaign period.

## Insight

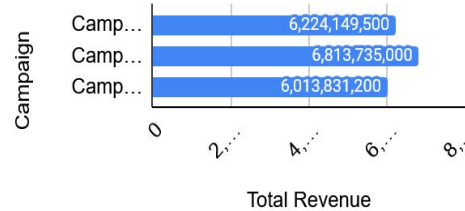
We use the **PivotTable Analyze → Fields, Items & Sets → Calculated Field → Name (Ratio on Revenue to Discount) → Formula** (= 'Total Revenue' / Discount ) and found that the one tables beside on bottom, we can conclude that:

- The 3rd campaign is the most efficient (every Rp 1 spent on discount generates Rp 54.90 in revenue).
- Campaigns 1 and 2 are still inefficient because the high discount amounts do not correspond proportionally to the revenue generated.

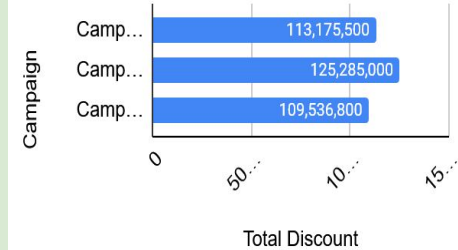
## Extra Miles

Total Revenue vs Campaign

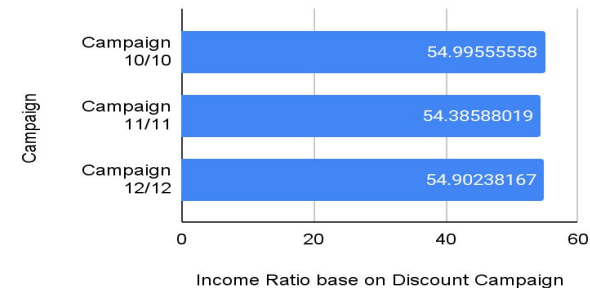
in Rp.



Total Discount vs Campaign



Income Ratio base on Revenue to





# 4. EDA & Statistical Measurement

## Business Question

C. Identify which product category had the highest transaction, customer, product sold, and revenue, while also having the lowest discount campaign budget for eight campaign period

## Insight

We use the **PivotTable → Row + Category Product → Value + Transaction + Customer + QTY + Revenue + Discount → Filter + Campaign (one by one 10/10, 11/11, 12/12) → Tabel Campaign**, we can conclude that :

**All campaigns show the same pattern:**

- **The top-performing categories are not** the ones receiving the lowest discount.
- Zero-discount categories are typically those with **lower performance**, not the leading categories.

## Extra Miles

Campaign	Transaction	Customer	Revenue	QTY	Discount
10/10	Beauty & Grooming	Men's Fashion	Men's Fashion	Beauty & Grooming	Books
					Others
11/11	Men's Fashion	Men's Fashion	Men's Fashion	Men's Fashion	Entertainment
12/12	Health & Sports	Health & Sports	Men's Fashion	Health & Sports	Books

From the table on the top, we can see that although on Identify which product category,

### Campaign 10/10

The lowest discount = 0, but the category with zero discount is not the top-performing category.

Therefore, Campaign 10/10 does not meet the criteria (because the best-performing category ≠ the category with the lowest discount).

### Campaign 11/11

Again, the best-performing category is **Men's Fashion**, but the category with zero discount is **Entertainment**, which is not the top performer.

Thus, Campaign 11/11 does not meet the criteria.

### Campaign 12/12

The lowest discount is given to **Books**, but the best-performing categories are **Health & Sports** and **Men's Fashion**.

So Campaign 12/12 also does not meet the criteria. There is **no single product category** across all three campaigns that:

- ✓ Has the highest transactions
- ✓ Has the highest customers
- ✓ Has the highest QTY
- ✓ Has the highest revenue
- ✓ While also receiving the lowest campaign discount

# 4. EDA & Statistical Measurement

## Business Question

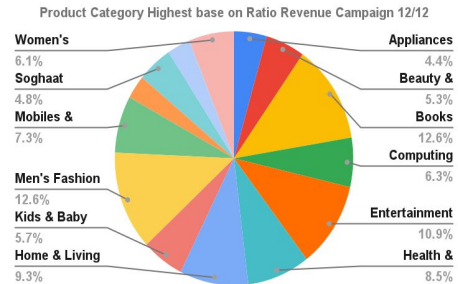
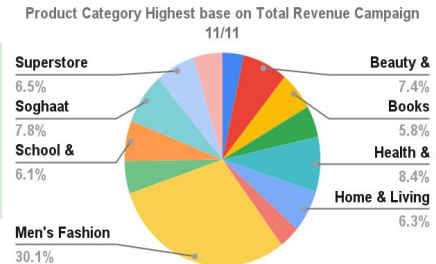
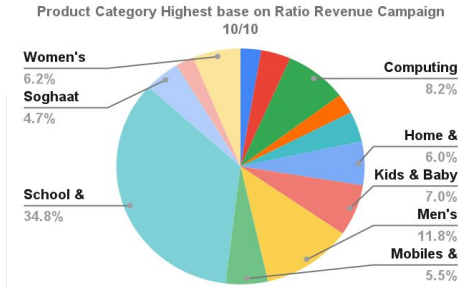
D. Find the product category with the highest ratio of revenue to discount campaign budget in each campaign period.

## Insight

We use the **PivotTable → Row + Category Product → Value + Rasio Pendapatan → Filter + Campaign (one by one 10/10, 11/11, 12/12) → Tabel Campaign**, we can conclude that :

- In the **10/10 Campaign**, the highest product category is **School & Education**, with **34.8%**.
- In the **11/11 Campaign**, the highest product category is **Men's Fashion**, with **30.1%**.
- In the **12/12 Campaign**, the highest product category includes **two products with the same value: Men's Fashion and Books**, each at **12.6%**.

## Extra Miles



# 4. EDA & Statistical Measurement

## Business Question

E. Summarize overall insights and provide actionable business recommendations based on all EDA and statistical findings.

## Insight

### Key Insights

- ❖ Campaign performance is not solely dependent on the size of the discount, but also on product and customer segmentation strategies.
- ❖ High value products are better suited for campaigns with limited yet high-impact promotions (such as bundling or cashback).
- ❖ Everyday essential products are well suited for volume based discounts to increase transaction frequency.

## Extra Miles

### Business Recommendations

- ❖ Optimize campaigns using a hybrid approach: apply low to medium discount levels and combine them with bundling offers to maintain healthy margins.
- ❖ Prioritize categories with a high revenue-to-discount ratio (such as Home & Living) for upcoming campaigns.
- ❖ Enhance promotion personalization based on purchase history to improve discount budget efficiency.
- ❖ Use EDA results to allocate campaign budgets more effectively—focusing more on segments that consistently deliver high ROI.

# 5. Hypothesis Testing

## Business Question

A. How would you proceed to analyze? What are your next steps?

## Insight

- By looking at the Campaign Sheet, we must separate the data
- We may filter based on the campaign and copy paste the data to another sheet
- After that we can start to formulate the hypothesis
- In this case we formulate the hypothesis as follows

**H<sub>0</sub>** :  $\text{MntWines}_X = \text{MntWines}_Y$

>>there is no statistically significant difference between the X and Y campaign in Wine revenue

**H<sub>a</sub>** :  $\text{MntWines}_X \neq \text{MntWines}_Y$

>>there is a statistically significant difference between the X and Y campaign in Wine revenue.

# 5. Hypothesis Testing

## Business Question

B. How would you assess the difference in terms of statistics?

## Insight

- We use the t-test as our statistical testing since **we do not know the population variance**
- We will use **t-test for two sample assuming unequal variances** since this is the safest way for t-test since we do not know how much difference the variance of MntWines revenue are between the 2 campaign and we have two samples
- We will use **5% level of significance** and compare it to the p-value obtained from the XLminer summary output in **two tail setting**
- We will interpret the p value to conclude whether we reject or accept the null hypothesis

# 5. Hypothesis Testing

## Business Question

C. What is/are your insight(s) on this matter?

## Insight

- 1) Based on the calculations p-value is above 0.05 (0.35), which means that we accept the null hypothesis
- 2) Meaning, **there is no statistically significant difference** in the Wine revenue between X and Y campaign. It means that the **Y campaign does not bring significant difference of revenue in Wine products.**

t-Test: Two-Sample Assuming Unequal Variances		
	MntWines_X	MntWines_Y
Mean	310.8711152	297.7097902
Variance	114146.3832	112467.1091
Observations	1094	1144
Hypothesized Mean Difference	0	
df	2229	
t Stat	0.9245436357	
P(T<=t) one-tail	0.1776517185	
t Critical one-tail	1.645537464	
P(T<=t) two-tail	0.3553034371	
t Critical two-tail	1.961028778	

# 5. Hypothesis Testing

## Business Question

D. What recommendations would you suggest to the company based on your analysis of the claim?

## Insight

Considerings:

1. Y campaign cost more than the X campaign
2. We have conduct hypothesis testing to test the claim
- 3. We found out that the claim is false**

The suggestion (business recommendation):

1. We suggest that A-Mart Supply may:
  - a. Create new ad campaign that is different from the existing campaign
  - b. Use the insights / information based on the data to create targeted campaign
  - c. etc

# Thank You



[www.linkedin.com/in/vennydeslaweny](https://www.linkedin.com/in/vennydeslaweny)

