



Week 2 & 3 - Spreadsheet and statistics

Milestone 1 and 2

By Ligasyah Arnanda Primadana



Table of Content

1. Case
2. Business Background
3. Business Problem
4. Business Understanding
5. Data Cleaning
6. Outliers Handling
7. Descriptive Statistics
8. Interpretations



Case

To serve as a data analyst for TokoBli, the preeminent e-commerce enterprise in Indonesia. The platform offers a diverse array of products, including men's and women's fashion, electronics, and home living items. On every twin date, such as October 10th, November 11th, and December 12th, a promotional campaign is launched featuring significant flash sales and complimentary shipping for all purchases. During these promotional periods, the volume of transactions can surge up to tenfold compared to regular days.

In the course of the last three campaign periods, the team experimented with three distinct campaign strategies. During the October 10, 2023 (10/10) campaign, scenario A was employed. For the November 11, 2023 (11/11) campaign, scenario B was implemented, and scenario C was utilized for the December 12, 2023 (12/12) campaign.

To enhance the effectiveness of future campaigns, the business team has commissioned the data analyst team to assess and extract insights from the previous campaigns, each characterized by a different scenario.





Its time to analyze

The Datas





Business Background



Work as Data Analyst

On Toko Bli



Ceo Demand about

Improvement of
Future Twin Date
Campaign



The Vision

From the data we can
find our customer
behavior on 3
campaign



Consumer Analysis

Our consumer have
certain pattern so we
can do more effective
campaign based on
that data



Delivered and
recomendataion

We as Data Analysis
find certain
customer behavior





Business Problem



Data Analyst

Need to do some coordination with marketing team for the details of the campaign and data engineer to follow up the data such as Status



The Matrics

There are some matrices need to be analyze for improvement future campaign such as :

1. the campaign budget for discounts,
2. the number of transactions,
3. the number of customers,
4. the number of products sold, and the amount of revenue earned during the campaign period.



Main Objective

This data analysis will help us achieve optimal revenue and transactions while minimizing the campaign budget.





Business Understanding

Q: What is the Problem Statement for TokoBli?

The Main objective is to systematically evaluate the effectiveness of the three campaign strategies by analyzing key performance metrics within a defined timeframe to identify the optimal approach for maximizing revenue and transactions while minimizing costs for the next month campaign



Business Understanding 2

Q: What business metrics do you think are best to answer the problems identified in point A?

By systematically analyzing these metrics—campaign budget for discounts, number of transactions, number of customers, number of products sold, and amount of revenue earned—the data analyst team can derive valuable insights into the effectiveness of each campaign strategy.

For the instance, Customer purchase behavior on our three campaign is most important metrics to be analyze.





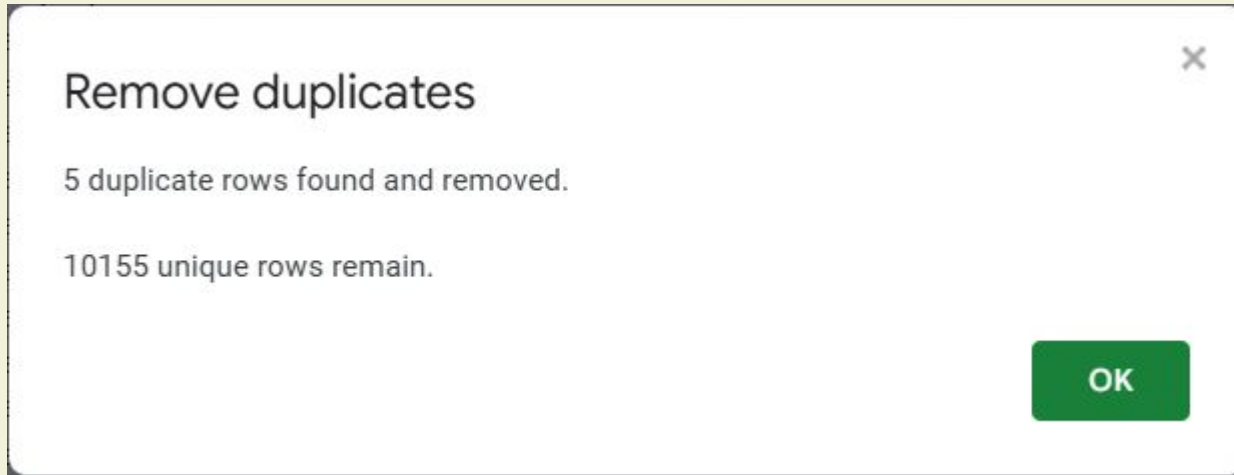
Data Cleaning

Clean up the data



Data Cleaning - Remove Duplicates

Removing duplicate using methods on spreadsheets Data > Data Cleanup > Remove Duplicates





Data Cleaning - Missing Data

For the first step to find the missing data is using methods `=countblank()` on every row. Here the result:

There are 4 column have a missing values (Status, Price, QTY and Discount).

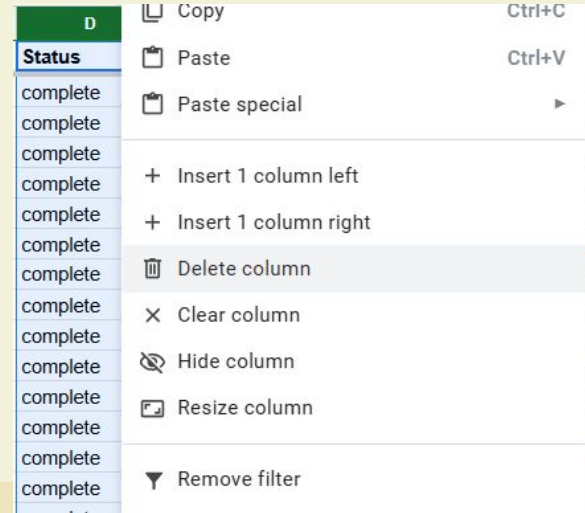
Status	Created At	Campaign	SKU	Price	QTY	Discount
10101	0	0		0	1	1

After using countblank i deleted the formula for easier to analysis purpose.



Data Cleaning - Missing Data

For the Status Column there are more 70% of missing data so the i remove the column because of there will be hard to analyze the data with that much of missing value. But for the business problem us as data analyst need to coordinate with the data engineer to follow up the data.





Data Cleaning - Missing Data (Price, QTY, and Discount)

There are one of each column from price, QTY and Discount missing value. My methods to fill this blank data is with inputting the value based on this formula :

- $\text{Price} = (\text{Total Revenue} + \text{Discount} - \text{Shipping Cost}) / (\text{QTY})$
- $\text{QTY} = (\text{Total Revenue} + \text{Discount} - \text{Shipping Cost}) / (\text{Price})$
- $\text{Discount} = (\text{Price} * \text{QTY}) + \text{Shipping Cost} - \text{Total Revenue}$

Here the results :

Transaction	Product ID	Customer ID	Created At	Campaign	SKU	Price	QTY	Discount	Shipping Cost	Total Revenue
ID-147	725300	84580	2023-10-10	Campaign 10/10	SADPAR59F6C59E02C06	2,500,000	1	0	0	2,500,000

Transaction	Product ID	Customer ID	Created At	Campaign	SKU	Price	QTY	Discount	Shipping Cost	Total Revenue
ID-152	725305	84581	2023-10-10	Campaign 10/10	HALBUK5A013F87DB79F	1,760,000	1	0	0	1,760,000

Transaction	Product ID	Customer ID	Created At	Campaign	SKU	Price	QTY	Discount	Shipping Cost	Total Revenue
ID-177	725721	44913	2023-10-10	Campaign 10/10	WOFDEA59CE2178B2F7E	1,990,000	1	199,000	0	1,791,000



Data Cleaning - Unnecessary Blank Space

There are several unused row, i remove it for better experience on analysis

10156				
10157				
10158				
10159				

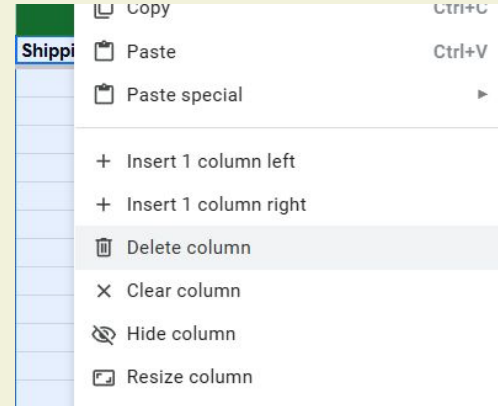
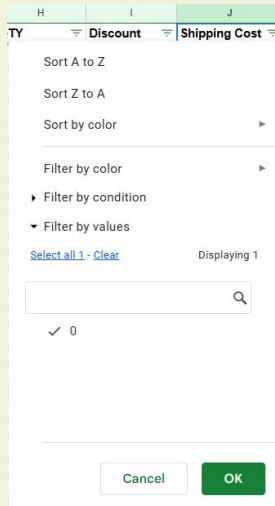




Data Cleaning - Shipping Cost Column Remove

In Shipping Cost there are only one values, so for the effectiveness in data analysis processing i removed it. Here the process:

The reason why shipping cost only zero because on the three campaign all the shipping cost is zero so not really necessary to analyze.





Data Cleaning - Change the Data Type

For the analysis there are need to using correctly data type. There are column Price, QTY, discount and Total Revenue changed into numbers. What not using currency for price, discount and total revenue? Is for easier process analysis in excel but for the report will be present using value currency. Here the Process:

23		Default...	-	10	+	B	I	÷	A	🔍	📄	📊	📈	📉	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	📏	📐	
----	--	------------	---	----	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	--



Data Cleaning - Decrease Decimal

For the easier to read i also decrease the decimal. I am not using decimal for the easier to understanding too. Here the Process:

	D	E	F	G	H	I	J	K
	Created At	Campaign	SKU	Price	QTY	Discount	Shipping Cost	Total Revenue
21	2023-10-10	Campaign 10/10	BAGLUX5A099F4E3D5CF	2,100,000	5	0	0	10,500,000
168	2023-10-10	Campaign 10/10	SOGAMB5A311E53D6BE9	4,200,000	1	0	0	4,200,000
168	2023-10-10	Campaign 10/10	SOGAMB5A09BA2372A3B	1,200,000	1	0	0	1,200,000
168	2023-10-10	Campaign 10/10	BAGGAR5A0A925948525	3,970,000	1	0	0	3,970,000
168	2023-10-10	Campaign 10/10	KN059D64DAF36939	270,000	2	0	0	540,000
168	2023-10-10	Campaign 10/10	PS05A43482AF3C58	200,000	1	0	0	200,000
6715	2023-10-10	Campaign 10/10	MEFTHE59F6D761A69C1-40	7,990,000	1	0	0	7,990,000
8616	2023-10-10	Campaign 10/10	WOFCHA5A267D9B0F85	8,990,000	1	0	0	8,990,000
8616	2023-10-10	Campaign 10/10	WOFDEA59CE228149430-36	4,490,000	1	0	0	4,490,000
8765	2023-10-10	Campaign 10/10	BAGDOV5A099F4EF3FFA	3,700,000	1	0	0	3,700,000
5960	2023-10-10	Campaign 10/10	SOGAMB5A2E671470128	900,000	2	0	0	1,800,000
5960	2023-10-10	Campaign 10/10	SOGAMB5A2E67122360F	900,000	2	0	0	1,800,000
5960	2023-10-10	Campaign 10/10	SOGAMB5A2E6713BB9AD	900,000	2	0	0	1,800,000
5960	2023-10-10	Campaign 10/10	SOGMAY5A2F7FC4D8DCA	2,200,000	1	0	0	2,200,000
5960	2023-10-10	Campaign 10/10	SOGMAY5A2F7FD65351	4,800,000	1	0	0	4,800,000
5960	2023-10-10	Campaign 10/10	SOGSND5A43754748823	900,000	1	0	0	900,000
5960	2023-10-10	Campaign 10/10	SOGSND5A4375431C7D4	1,700,000	1	0	0	1,700,000
5960	2023-10-10	Campaign 10/10	SOGSND5A437540871D2	2,200,000	1	0	0	2,200,000
5960	2023-10-10	Campaign 10/10	SOGAMB5A2E6716F1D40	900,000	2	0	0	1,800,000
5960	2023-10-10	Campaign 10/10	SOGAMB5A2E671507BC4	900,000	1	0	0	900,000
6593	2023-10-10	Campaign 10/10	BAGLUX5A099F4E3D5CF	2,100,000	1	210,000	0	1,890,000
7684	2023-10-10	Campaign 10/10	MEFTHE59F6D780CDF16-43	5,990,000	1	599,000	0	5,391,000
7684	2023-10-10	Campaign 10/10	BAGKEM5A38AE901D18E	4,450,000	1	445,000	0	4,005,000



Descriptive Statistics

- a. Removing Outliers
- b. Descriptive Statistics
- c. Business Context





Removing The Outliers

For removing the outliers i using the formula of $(Q3 + (IQR * 1.5))$ for upper inner fence and $(Q1 - (IQR * 1.5))$ For Lower Inner Fence. But there are problem where the value of Lower inner fence is minus, so for further analysis the result will be rounding into 0 for Lower Inner Fence. Here the Calculation:

Q1	Q3	IQR	Upper Inner Fence	Lower Inner Fence
2,241,000	6,987,500	4,746,500	14,107,250	-4,878,750
Rounding			14,107,250	0



How to remove the outliers

For removing the outliers i will using the filters after find the calculation number of the Upper inner fence. Why the lower inner fence not filtered because of the are no value under the zero. Here the process:

There are 12 Row getting removed on this process

Discount | Total Revenue

Sort A to Z
Sort Z to A
Sort by color

Filter by color

Filter by condition

Greater than

14113500

Filter by values

Select all 1.2K - Clear

Displaying 200

45,000
95,000
135,000
164,000

Cancel OK

J	
Total Revenue	
78,400,000.00	
69,900,000.00	
64,000,000.00	
50,190,000.00	
48,090,000.00	
16,590,000.00	
74,250,000.00	
74,250,000.00	
21,000,000.00	
83,556,000.00	
65,100,000.00	
69,800,000.00	

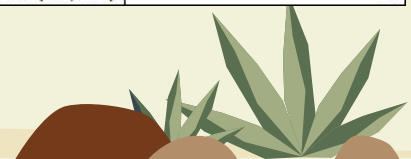


Descriptive Statistics and interpretation

Here the statistical descriptive for QTY after removing total revenue outliers Using XL Miner :

1. For the central tendency is in normal distribution. Its possible to analysis further
2. Fro max and min value is quite big difference we need to find where the maximum categories usually effective on campaign

QTY	
Mean	1
Standard Error	0
Median	1
Mode	1
Standard Deviation	1
Sample Variance	1
Kurtosis	442
Skewness	17
Range	29
Minimum	1
Maximum	30
Sum	12,808
Count	10,139
Largest(1)	30
Smallest(1)	1
Confidence Level(95,0%)	0





Descriptive Statistics and interpretation

Here the statistical descriptive for Discount after removing total revenue outliers Using XL Miner :

1. For the median and mode of discount is 0, is the users only want to using free shipping for the campaign? Need to ask the customer what their biggest drive in every campaign

Discount	
Mean	81,980
Standard Error	1,859
Median	0
Mode	0
Standard Deviation	187,143
Sample Variance	35,022,331,990
Kurtosis	5
Skewness	3
Range	900,000
Minimum	0
Maximum	900,000
Sum	831,198,000
Count	10,139
Largest(1)	900,000
Smallest(1)	0
Confidence Level(95,0%)	3,643



Descriptive Statistics and interpretation

Here the statistical descriptive for total revenue after removing total revenue outliers
Using XL Miner :

1. From the data the mean and median is below the mode, it means the curve is negative skew. for the descriptive analysis need to be careful to use
2. The score below 1 it means the skew is negative
3. The Minimum and Maximum Data is having quite big difference, for the further analysis need to difference between the categories of product.
4. Sum of the Total Revenue be benchmark for next campaign total revenue.
5. Number of People involved in this campaign, it can be benchmark for next campaign

Total Revenue	
Mean	4,563,700.66
Standard Error	26,537.59
Median	4,275,000.00
Mode	7,200,000.00
Standard Deviation	2,672,139.05
Sample Variance	7,140,330,000,000.00
Kurtosis	-0.50
Skewness	0.37
Range	13,895,000.00
Minimum	45,000.00
Maximum	13,940,000.00
Sum	46,271,361,000.00
Count	10,139.00
Largest(1)	13,940,000.00
Smallest(1)	45,000.00
Confidence Level(95,0%)	52,018.93





Adds on Idea for Pivot Table

There are further probability of improvement for next campaign based on categorization. We need to analyze more base on categorization. Here the data:

1. The big four of success categories of our campaign are on beauty, health, men fashion and superstore based on QTY
2. Based on Total Revenue Men fashion are the best from others categories.
3. Books based on revenue and QTY not have a good transaction on our campaign.

Category Name	SUM of Total Re	SUM of QTY	SUM of Discoun
Appliances	1,976,429,000	344	61,761,000
Beauty & Groom	6,094,399,800	1,752	126,671,200
Books	105,977,000	22	1,043,000
Computing	782,882,000	171	17,108,000
Entertainment	341,599,000	61	5,511,000
Health & Sports	4,207,256,000	1,745	65,614,000
Home & Living	3,469,498,000	901	67,619,000
Kids & Baby	1,369,652,000	454	38,198,000
Men's Fashion	14,693,964,000	2,664	126,036,000
Mobiles & Tablet	2,720,082,000	755	44,848,000
Others	12,500,000	3	0
School & Educat	299,166,000	106	5,079,000
Soghaat	2,108,225,000	821	44,109,000
Superstore	4,679,307,200	2,443	152,194,800
Women's Fashic	4,125,550,000	730	78,380,000
Grand Total	46,986,487,000	12,972	834,172,000





Conclusions and recommendations

After looking the data so far we as data analyst need to doing several coordination with others team for better insight.

1. We as data analyst need to have meeting with the Marketing teams, Customers handling and Data Engineer team to solve several problem.
2. Data Engineer need to complete status column to get better understanding
3. We need to discuss with marketing to find what their intention on every campaign and need to analysis more about every categorization analysis
4. We need to create more detail assessment to our customer for get their impression on every campaign we conduct so far.
5. From the insight in pivot tables, we need to utilize the strategies with marketing need to emphasize more the successful categories or enhancing the less transaction categories.



Week 2 & 3 - Spreadsheet and statistics

Milestone 2

By Ligasyah Arnanda Primadana

The Data for Milestone 2

Difference between the three campaign



Values	Campaign 10/10	Campaign 11/11	Campaign 12/12
Total Transactions	3329	3399	3414
Total Customers	1498	1567	1447
Total Products Sold	4,211	4,115	4,485
Total Revenue	14,853,951,100	16,622,135,800	14,801,325,100
Total Discount Campaign Budget	275,925,900	302,636,200	252,834,900

Here The Difference

1. Total Transactions: Increased steadily across all campaigns, indicating effective outreach.
2. Total Customers: Grew initially but declined in Campaign 12/12, suggesting retention issues.
3. Total Products Sold: Fluctuated, with a notable rebound in Campaign 12/12, highlighting effective promotions.
4. Total Revenue: Peaked in Campaign 11/11 but dropped in Campaign 12/12, indicating potential pricing strategy concerns.
5. Total Discount Campaign Budget: Increased in Campaign 11/11 but decreased in Campaign 12/12, warranting a review of discount effectiveness.

Overall, while transactions and product sales show positive trends, customer retention and revenue stability require further attention.



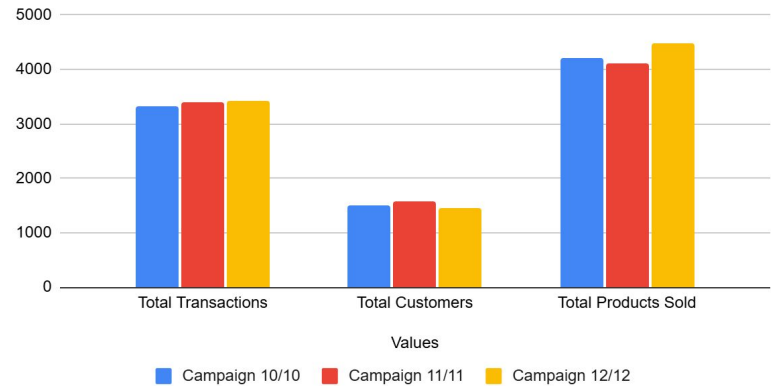
Comparison Between Total Transaction, Customer and Product Sold

The comparison of the three campaigns in terms of total transactions, total customers, and total products sold reveals the following insights:

1. Total Transactions: Campaign 12/12 shows the highest number of transactions, followed closely by Campaign 11/11, indicating effective engagement strategies in the latter campaigns.
2. Total Customers: Campaign 11/11 attracted the most customers, while Campaign 12/12 experienced a decline, suggesting potential challenges in retaining customer interest.
3. Total Products Sold: Campaign 12/12 leads in product sales, reflecting successful promotional efforts, despite the drop in customer numbers.

Overall, while Campaign 12/12 excels in transactions and product sales, the decline in customer numbers highlights the need for improved retention strategies moving forward.

Comparison between every campaign in Total Transaction, Total Customer and Total Product Sold





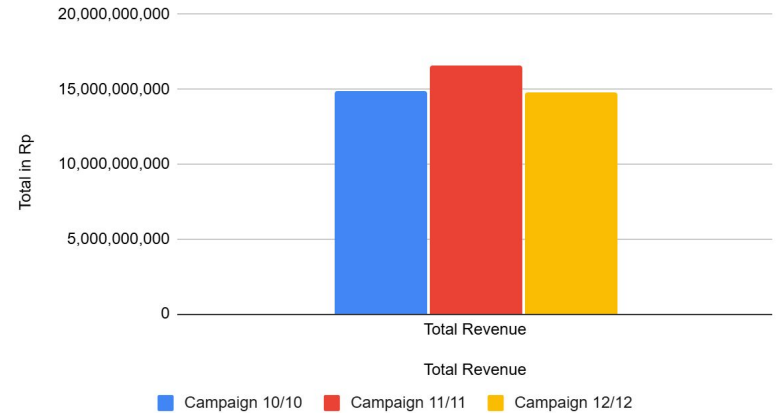
Comparison Between Campaign in Total Revenue

The comparison of total revenue across the three campaigns provides the following insights:

1. Campaign 11/11: This campaign generated the highest revenue, indicating successful marketing strategies and consumer engagement during this period.
2. Campaign 10/10: Revenue is slightly lower than Campaign 11/11, suggesting a solid performance but not as impactful as the subsequent campaign.
3. Campaign 12/12: Revenue decreased compared to Campaign 11/11, indicating potential challenges in maintaining sales momentum and consumer interest.

In summary, while Campaign 11/11 stands out in revenue generation, the decline in Campaign 12/12 highlights the need for strategies to sustain revenue growth in future campaigns.

Comparison between every campaign in total revenue





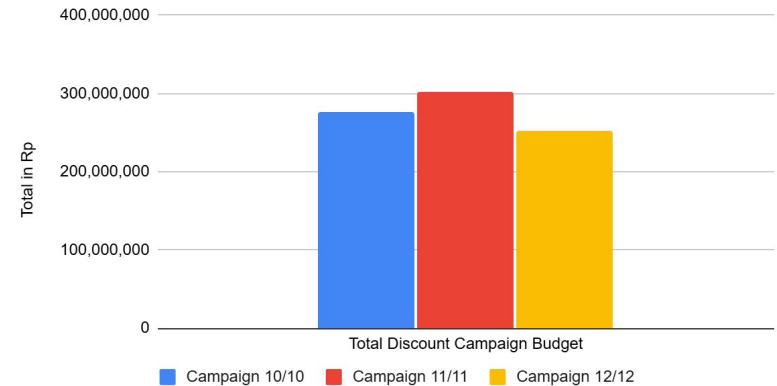
Comparison Between Campaign in Total Discount Campaign Budget

The analysis of the total discount campaign budget across the three campaigns reveals the following insights:

1. Campaign 11/11: This campaign allocated the highest budget for discounts, indicating a strong emphasis on promotional strategies to drive sales.
2. Campaign 10/10: The budget was lower than Campaign 11/11, suggesting a more conservative approach to discounts during this period.
3. Campaign 12/12: The discount budget decreased compared to Campaign 11/11, which may have contributed to the decline in revenue and customer numbers observed in this campaign.

In summary, while Campaign 11/11 effectively utilized a larger discount budget to boost sales, the reduction in Campaign 12/12 highlights the need for careful budget management to sustain promotional effectiveness in future campaigns.

Comparison between every campaign in Total Discount Campaign Budget





Ratio revenue and discount campaign budget between the three campaign

Values	Campaign 10/10	Campaign 11/11	Campaign 12/12
Total Revenue	14,853,951,100	16,622,135,800	14,801,325,100
Total Discount Campaign Budget	275,925,900	302,636,200	252,834,900
Ratio Revenue and Campaing Budget	53.83	54.92	58.54

Here the Insight

1. Campaign 10/10: The ratio of approximately 53.83 indicates that for every unit spent on discounts, revenue was about 53.83 times that amount, reflecting a reasonably effective use of the discount budget.
2. Campaign 11/11: The ratio increased to around 54.92, suggesting a slight improvement in the effectiveness of discount spending compared to Campaign 10/10. This indicates that promotional strategies were more successful in generating revenue.
3. Campaign 12/12: The ratio rose significantly to approximately 58.54, marking the highest effectiveness among the three campaigns. This suggests that, despite a decrease in total revenue, the campaign generated more revenue per unit of discount spent, likely due to more targeted promotions or appealing product offerings.

Conclusion

Overall, the increasing trend in the ratio from Campaign 10/10 to Campaign 12/12 indicates an improvement in the efficiency of discount strategies, highlighting the need to analyze the successful elements of these campaigns for future marketing efforts.



Ratio revenue and discount campaign budget between the three campaign

Campaign 10/10: The ratio is the lowest among the three campaigns, indicating that while revenue was generated, the effectiveness of the discount budget was not as high as in subsequent campaigns.

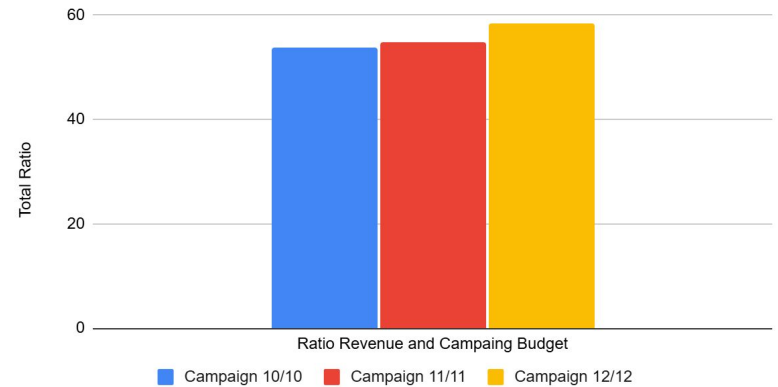
Campaign 11/11: The ratio shows a slight increase, suggesting improved effectiveness in utilizing the discount budget to generate revenue compared to Campaign 10/10.

Campaign 12/12: The ratio reaches its highest point, indicating the most effective use of the discount budget. This suggests that the strategies employed during this campaign were particularly successful in generating revenue relative to the amount spent on discounts.

Conclusion

Overall, the increasing trend in the ratio from Campaign 10/10 to Campaign 12/12 demonstrates a significant improvement in the efficiency of discount spending. This indicates that the promotional strategies became more effective over time, culminating in the highest ratio in Campaign 12/12. Future campaigns should analyze these successful elements to maintain and enhance revenue generation.

Comparison between every campaign in Ratio Revenue and Campaign Budget



The Highest But also Maintain the Discount



Category Name	Campaign 10/10				Campaign 11/11				Campaign 12/12			
	Total Transac	Total Custom	Total Products	Total Revenue	Total Transac	Total Customer	Total Products	Total Revenue	Total Transac	Total Customer	Total Products	Total Revenue
Appliances	84	79	87	556,784,000	105	84	118	701,922,000	106	94	112	85,901,000
Beauty & Grooming	624	319	822	2,935,095,500	239	126	285	1,010,663,800	434	254	329	1,339,742,000
Books	3	2	3	5,970,000	3	3	3	14,106,000	16	12	112	417,368,000
Computing	54	46	62	285,063,000	38	27	41	157,050,000	63	57	68	340,769,000
Entertainment	9	9	10	58,632,000	13	11	13	72,540,000	27	26	28	140,527,000
Health & Sports	288	143	401	1,082,422,000	294	82	409	1,074,092,000	612	135	935	2,050,742,000
Home & Living	283	184	312	1,151,061,000	235	160	261	980,455,000	297	189	329	1,339,742,000
Kids & Baby	160	101	182	563,593,000	149	79	160	388,691,000	103	70	112	417,368,000
Men's Fashion	614	430	641	3,371,160,000	1347	877	1,431	8,246,126,000	543	360	582	3,006,878,000
Mobiles & Tablets	273	203	300	1,006,570,000	160	123	175	637,921,000	253	194	270	997,191,000
Others	2	2	3	12,500,000	0	0	0	0	0	0	0	0
School & Education	53	32	57	172,862,000	11	8	11	33,202,000	35	25	39	95,602,000
Soghaat	325	131	414	1,117,046,600	93	56	115	362,828,000	218	98	292	628,350,400
Superstore	325	119	678	1,126,062,000	560	139	939	2,040,230,000	384	177	796	1,280,959,200
Women's Fashion	232	159	239	1,409,130,000	152	95	154	902,309,000	323	190	338	1,815,902,000

Based on the data The Men Fashion is the Highest Category:

- **Sales Performance:** Men's Fashion consistently ranks as the top category in terms of transactions, customers, products sold, and revenue across all three campaigns. This indicates strong consumer demand and effective marketing strategies for this category.
- **Discount Budget Efficiency:** Despite generating high sales figures, the discount campaign budget for Men's Fashion remains relatively low compared to other categories. This suggests that the category's success is not solely reliant on discounts, indicating effective pricing strategies and brand loyalty.
- **Market Positioning:** The strong performance of Men's Fashion across all metrics highlights its importance in the overall product mix. The ability to attract a significant number of customers while maintaining a lower discount budget suggests that this category may benefit from targeted marketing efforts to further enhance its market position.

Conclusion

In summary, Men's Fashion emerges as the best-selling product category across the three campaign periods, excelling in transactions, customers, products sold, and revenue while maintaining a lower discount campaign budget. This indicates a successful balance of effective marketing and pricing strategies, warranting continued focus in future campaigns.



Men's Fashion on Maintain the Discount

Key Findings

Revenue Generation:

Men's Fashion leads in revenue across all campaigns, with figures of 3,371,160,000 in Campaign 10/10, 8,246,126,000 in Campaign 11/11, and 3,006,878,000 in Campaign 12/12. Other strong categories include Beauty & Grooming and Health & Sports, especially in Campaign 11/11. Discount Campaign Budget:

Men's Fashion maintains a low discount budget relative to its revenue, with budgets of 33,760,000 in Campaign 10/10 and 59,184,000 in Campaign 11/11. Categories like Health & Sports have higher budgets but lower revenue. Ratio of Revenue to Discount Budget:

Men's Fashion shows a high revenue-to-discount ratio, particularly 100 in Campaign 10/10 and 139 in Campaign 11/11, indicating effective use of discounts. Other categories, such as Beauty & Grooming, also perform well but not as strongly as Men's Fashion.

Conclusion
Table demonstrates that Men's Fashion excels in revenue generation and discount efficiency. Its high revenue-to-discount ratio suggests effective marketing strategies, making it a key focus for future campaigns.

Discount Campaign Budget				
Category Name	Campaign			Total
	Campaign 10/10	Campaign 11/11	Campaign 12/12	
Appliances	17,366,000	28,898,000	15,497,000	61,761,000
Beauty & Grooming	69,487,500	24,408,200	32,145,500	126,041,200
Books	0	424,000	619,000	1,043,000
Computing	6,417,000	4,010,000	6,681,000	17,108,000
Entertainment	1,548,000	1,830,000	2,133,000	5,511,000
Health & Sports	19,658,000	22,498,000	23,458,000	65,614,000
Home & Living	17,219,000	27,745,000	22,655,000	67,619,000
Kids & Baby	8,397,000	20,599,000	9,202,000	38,198,000
Men's Fashion	33,760,000	59,184,000	33,092,000	126,036,000
Mobiles & Tablets	16,440,000	14,269,000	14,139,000	44,848,000
Others	0	0	0	0
School & Education	1,648,000	588,000	2,843,000	5,079,000
Soghaat	22,007,400	7,442,000	14,659,600	44,109,000
Superstore	43,308,000	60,780,000	45,762,800	149,850,800
Women's Fashion	18,670,000	29,961,000	29,948,000	78,579,000



Ratio Revenue and Discount Campaign Budget Per Campaign

Key Findings

- Highest Ratio:
 1. Men's Fashion achieved the highest ratio of 100 in Campaign 10/10, indicating strong revenue generation relative to its discount budget.
- Other Notable Ratios:
 1. Beauty & Grooming had a ratio of 59 in Campaign 12/12, showing effective use of discounts.
 2. Health & Sports recorded a ratio of 55 in Campaign 10/10, indicating good revenue relative to its discount spending.
- Comparison Across Campaigns:
 1. Men's Fashion consistently outperforms other categories, while the ratios for others vary, suggesting different levels of discount effectiveness.

Insights

- Men's Fashion demonstrates strong brand loyalty and effective discount strategies, making it a key focus for future campaigns.
- Beauty & Grooming and Health & Sports have potential for improvement by analyzing their discount strategies to enhance their revenue-to-budget ratios.

Conclusion

Men's Fashion leads in the ratio of revenue to discount campaign budget across the three periods, particularly in Campaign 10/10, highlighting its effectiveness in maximizing revenue from discounts.

Category Name	Ratio Revenue and Discount Campaign Budget Per Campaign				
	Campaign 10/10	Campaign 11/11	Campaign 12/12	Total	
Appliances	32	24	46		103
Beauty & Grooming	42	41	59		142
Books	0	33	139		172
Computing	44	39	51		135
Entertainment	38	40	66		143
Health & Sports	55	48	87		190
Home & Living	67	35	59		161
Kids & Baby	67	19	45		131
Men's Fashion	100	139	91		330
Mobiles & Tablets	61	45	71		176
Others	0	0	0		0
School & Education	105	56	34		195
Soghaat	51	49	43		142
Superstore	26	34	28		88
Women's Fashion	75	30	61		166



Overall Insight

1. **Men's Fashion Dominance:** This category consistently led in transactions, customers, products sold, and revenue, showcasing effective marketing and pricing strategies.
2. **Customer Engagement Fluctuations:** While transactions increased, customer numbers declined in Campaign 12/12, indicating potential retention issues.
3. **Discount Strategy Effectiveness:** Categories like Men's Fashion effectively utilized their discount budgets, while others showed inefficiencies with high budgets but low revenue.
4. **Revenue Variability:** Revenue peaked in Campaign 11/11 but dropped in Campaign 12/12, suggesting a need for reassessment of promotional strategies.
5. **High Revenue-to-Discount Ratios:** Men's Fashion had the highest efficiency in generating revenue relative to discounts, indicating strong consumer demand.



Business Recommendations

1. Invest in Men's Fashion: Continue promoting this category and consider expanding product lines.
2. Enhance Customer Retention: Develop targeted marketing campaigns and loyalty programs to retain customers.
3. Optimize Discounts: Review discount strategies for categories with high budgets but low revenue to improve efficiency.
4. Monitor Trends: Stay updated on market trends and consumer preferences to adapt strategies accordingly.
5. Leverage Data Analytics: Use data to monitor performance and inform decision-making.
6. Implement A/B Testing: Test promotional strategies to identify the most effective approaches.

Conclusion

Focusing on high-performing categories, improving customer retention, and optimizing discount strategies will drive sustainable growth in future campaigns.



Difference between Campaign 10/10 and Campaign 11/11 Total Revenue

t-Test: Two-Sample Assuming Unequal Variances		
	Campaign 10/10	Campaign 11/11
Mean	4461985.91	4890301.79
Variance	6994044361764	6957590470647
Observations	3329	3399
Hypothesized Mean Difference	0	
df	6722	
t Stat	-6.65	
P(T<=t) one-tail	0	
t Critical one-tail	1.65	
P(T<=t) two-tail	0	
t Critical two-tail	1.96	

Given that the p-value is 0, which is less than the typical significance level of 0.05, we can conclude that there is a statistically significant difference between the total revenues of Campaign 10/10 and Campaign 11/11. The negative t-statistic indicates that the mean revenue for Campaign 10/10 is significantly lower than that for Campaign 11/11.

Difference between Campaign 10/10 and Campaign 12/12 Total Revenue



t-Test: Two-Sample Assuming Unequal Variances		
	Campaign 10/10	Campaign 12/12
Mean	4461985.91	4335478.94
Variance	6994044361764	7300189181921
Observations	3329	3414
Hypothesized Mean Difference	0	
df	6741	
t Stat	1.94	
P(T<=t) one-tail	0.026	
t Critical one-tail	1.65	
P(T<=t) two-tail	0.052	
t Critical two-tail	1.96	

For a two-tailed test, the p-value is 0.052, which is slightly above the typical significance level of 0.05. This suggests that there is not a statistically significant difference between the total revenues of Campaign 10/10 and Campaign 12/12 at the 5% significance level. However, it is very close, indicating a potential trend that might be significant at a slightly higher significance level.

Difference between Campaign 11/11 and Campaign 12/12 Total Revenue



t-Test: Two-Sample Assuming Unequal Variances		
	Campaign 11/11	Campaign 12/12
Mean	4890301.79	4335478.94
Variance	6957590470647	7300189181921
Observations	3399	3414
Hypothesized Mean Difference	0	
df	6808	
t Stat	8.58	
P(T<=t) one-tail	0	
t Critical one-tail	1.65	
P(T<=t) two-tail	0	
t Critical two-tail	1.96	

Given that the p-value is 0 for both one-tailed and two-tailed tests, which is significantly less than the conventional significance level of 0.05, we can conclude that there is a statistically significant difference between the total revenues of Campaign 11/11 and Campaign 12/12. The positive t-statistic of 8.58 indicates that the mean revenue for Campaign 11/11 is substantially higher than that for Campaign 12/12.



Insights and business recommendations

Insights

Performance Hierarchy: Campaign 11/11 emerged as the most successful, significantly outperforming both Campaign 10/10 and Campaign 12/12 in terms of total revenue. This indicates that the strategies implemented during Campaign 11/11 were particularly effective.

Comparative Analysis: While Campaign 10/10 and Campaign 12/12 did not show a statistically significant difference, the close p-value suggests that there may be underlying factors affecting their performance. Campaign 10/10 had a slightly higher revenue than Campaign 12/12, indicating that elements from Campaign 10/10 could be leveraged to enhance Campaign 12/12.

Strategic Reevaluation: The stark contrast in performance between Campaign 11/11 and Campaign 12/12 highlights the need for a thorough reevaluation of the strategies used in Campaign 12/12. Learning from the successful elements of Campaign 11/11 could provide valuable insights for improving future campaigns.

These combined insights emphasize the importance of analyzing successful strategies and addressing areas of underperformance to optimize overall campaign effectiveness.



Business recommendation

Business Recommendations

- Analyze Successful Strategies:

Conduct a detailed analysis of the strategies used in Campaign 11/11 to identify key factors contributing to its success. Focus on marketing channels, messaging, and target audience engagement.

Optimize Campaign 12/12:

- Review and enhance the strategies employed in Campaign 12/12. Incorporate successful elements from Campaign 11/11 to improve revenue generation and overall effectiveness.

Leverage Insights from Campaign 10/10:

- Investigate the aspects of Campaign 10/10 that may have contributed to its performance. Consider integrating these elements into Campaign 12/12 to boost its results.



Business recommendation

- Implement A/B Testing:

Utilize A/B testing for future campaigns to experiment with different strategies and identify the most effective approaches. This will allow for data-driven decision-making and continuous improvement.

- Establish Performance Metrics:

Develop clear performance metrics and KPIs for all campaigns. Regularly monitor these metrics to assess effectiveness and make timely adjustments as needed.

- Foster Continuous Learning:

Create a culture of continuous learning within the marketing team. Encourage sharing of insights and best practices from each campaign to enhance collective knowledge and strategy development.

- Engage with Target Audience:

Enhance engagement with the target audience through surveys and feedback mechanisms. Understanding customer preferences can inform future campaign strategies and improve overall satisfaction.

By implementing these recommendations, the organization can enhance its marketing effectiveness, drive improved revenue outcomes, and ensure sustained growth in future campaigns.



Thank you

