

# Internship Project Report on E-commerce Project

## 01 Executive Summary

This report presents the findings from an internship project focused on analyzing customer data for an e-commerce company. The primary objective was to extract actionable insights from the dataset, which includes customer demographics, spending habits, and interactions with marketing campaigns. This analysis aimed to enhance understanding of customer behavior, evaluate the effectiveness of marketing campaigns, and propose strategies for improving business performance. The insights derived from this project have significant implications for the company's marketing strategies and overall operations.

## 02 Introduction

### Background

The rapid growth of e-commerce has transformed the retail landscape, providing companies with vast amounts of customer data. This project was initiated to leverage this data to gain a deeper understanding of customer behavior and to optimize marketing efforts.

### Objectives

1. Analyze customer demographics and spending patterns.
2. Evaluate the effectiveness of marketing campaigns.
3. Provide actionable insights for enhancing customer engagement and business performance.

### Scope

The analysis covers various aspects of customer data, including demographic information, purchase behavior, and responses to marketing campaigns. The report outlines the methodologies used, presents the findings, and offers recommendations based on the insights gained.

## 03 Description of the Organization

### Company Overview

The organization is a prominent e-commerce platform known for its diverse product range and customer-centric approach. It operates globally and serves a wide customer base with a focus on providing a seamless shopping experience.

## Products and Services

The platform offers a vast array of products, including electronics, fashion, home goods, and more. Additionally, it provides personalized recommendations and engages customers through various marketing campaigns.

## Organizational Structure

The company is structured into several key departments, including Marketing, Sales, Customer Service, and Data Analytics. Each department plays a crucial role in the company's operations and strategic initiatives.

## Market Position and Competitiveness

The company holds a strong market position due to its extensive product offerings, competitive pricing, and excellent customer service. It continuously strives to enhance its market competitiveness through data-driven strategies and customer insights.

# 04 Experiential Learning

## Data Collection and Preparation

The dataset used for this project includes the following key variables:

- Customer ID
- Birth year
- Education
- Marital status
- Income
- Household composition
- Enrollment date
- Recency of purchases
- Complaints
- Campaign acceptances
- Spending categories

Data cleaning and preparation involved handling missing values, outliers, and ensuring data consistency to maintain the quality of the analysis.

## Analytical Tools and Techniques

Various analytical tools and techniques were employed to analyze the data:

- **Descriptive Analytics:** Used to summarize and describe the main features of the dataset, providing a clear overview of customer demographics and spending patterns.
- **Predictive Analytics:** Machine learning models were applied to predict future customer behavior and campaign success, enabling the company to make data-Driven decisions.

- **Segmentation Analysis:** Customers were segmented based on demographics and spending patterns to identify distinct groups with similar characteristics.

## Descriptive Analytics

Descriptive analytics provided insights into customer demographics, such as age distribution, education levels, marital status, and income brackets. Spending patterns were analyzed to identify high-value customers and popular product categories.

## Predictive Analytics

Predictive models, such as regression analysis and classification algorithms, were used to forecast future customer behavior and determine the likelihood of campaign success. These models helped in understanding the factors that influence customer purchases and campaign responses.

## Segmentation Analysis

Segmentation analysis identified key customer segments, such as frequent buyers, high spenders, and those responsive to specific marketing campaigns. This information was crucial for tailoring marketing strategies to different customer groups.

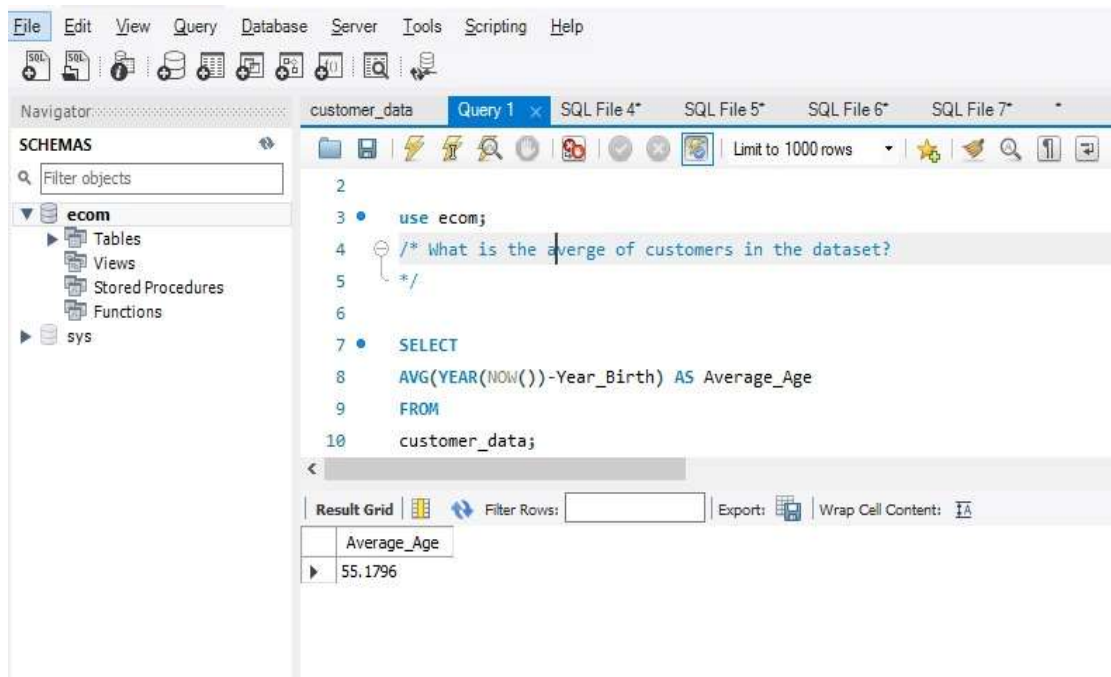
## Findings and Insights

- **Customer Demographics:** The analysis revealed that the majority of customers are aged between 25–45 years, with a significant portion holding a college degree. Income levels varied, with a notable segment earning above the median income.
- **Spending Patterns:** High-value customers were identified, contributing significantly to total sales. Popular product categories included electronics and fashion.
- **Campaign Effectiveness:** Certain marketing campaigns were more successful in driving customer engagement and purchases. Personalized campaigns showed higher acceptance rates compared to generic ones.

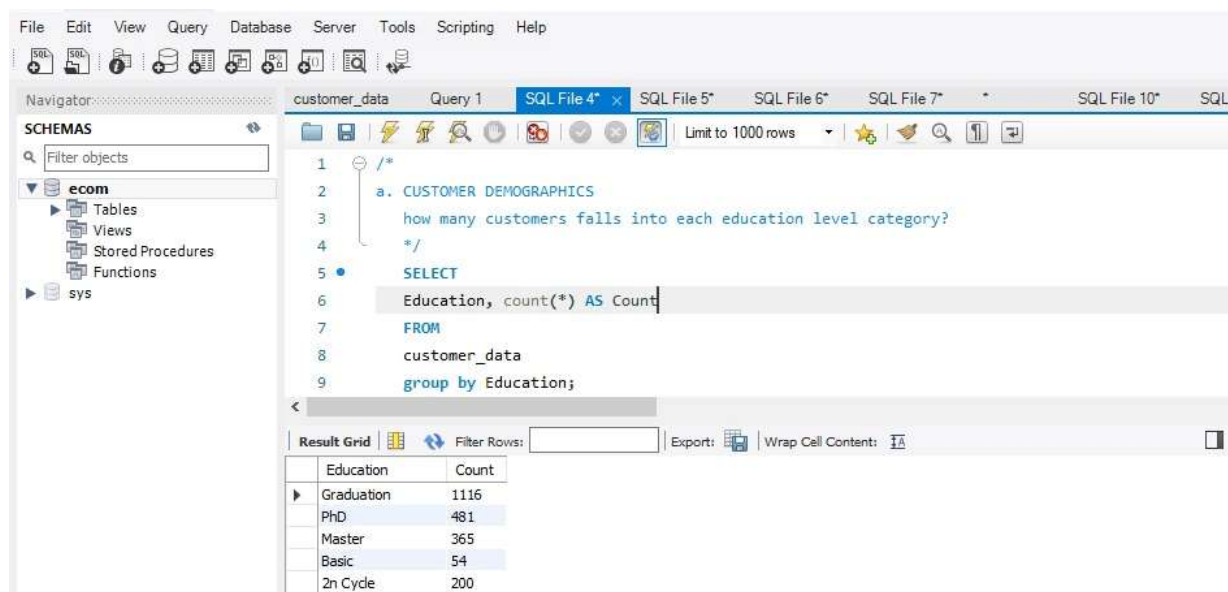
## I found the following insights

### 1. Customer Demographics

- i. What is the average age of customers in the dataset?



ii. How many customers fall into each education level category?



iii. What is the distribution of marital statuses among the customers?

The screenshot shows the SQL Server Enterprise Manager interface. The left pane displays the 'ecom' schema with tables, views, stored procedures, and functions. The right pane shows a query window with the following SQL code:

```
1  /* Customer Demographics
2     what is the distribution of marital status among the customers?
3  */
4  select
5     Marital_Status, count(*) AS count
6  from
7     customer_data
8  group by Marital_Status;
```

Below the query window, the 'Result Grid' shows the following data:

Marital_Status	count
Single	471
Together	573
Married	857
Divorced	232
Widow	76
Alone	3
Absurd	2
YOLO	2

2. Income and

Household Information

i. What is the average household income of the customers?

The screenshot shows the SQL Server Enterprise Manager interface. The left pane displays the 'ecom' schema with tables, views, stored procedures, and functions. The right pane shows a query window with the following SQL code:

```
1  /*
2     INCOME AND HOUSEHOLD INFORMATION
3     whts is the average household income of the customers
4  */
5  select
6     avg(Income) as Average_Income
7  from
8     customer_data;
```

Below the query window, the 'Result Grid' shows the following data:

Average_Income
52247.2514

ii. How many customers have children and teenagers in their households?

The screenshot shows the SQL Server Enterprise Manager interface. The left pane displays the 'ecom' schema with tables, views, stored procedures, and functions. The central pane shows a query titled 'INCOME AND HOUSEHOLD INFORMATION' with the following SQL code:

```
/*
how many customers have children and teenagers in their household?
*/
select
Kidhome, Teenhome, count(*) as count
from
customer_data
group by Kidhome, Teenhome;
```

The right pane shows the 'Result Grid' with the following data:

Kidhome	Teenhome	count
0	0	633
1	1	369
1	0	497
0	1	620
0	2	30
1	2	21
2	0	17
2	1	29

### 3. Purchase Behavior

i. What is the average number of days since the last purchase?

The screenshot shows the SQL Server Enterprise Manager interface. The left pane displays the 'ecom' schema with tables, views, stored procedures, and functions. The central pane shows a query titled 'PURCHASE BEHAVIOUR' with the following SQL code:

```
/*
what is the average number of days since the last purchase
*/
select
avg(Recency) as Average_Recency
from
customer_data;
```

The right pane shows the 'Result Grid' with the following data:

Average_Recency
49.01263537906137

ii. How much on average do customers spend on each product category?

The screenshot shows the SQL Project interface with a query editor and a result grid. The query is titled "PURCHASE BEHAVIOUR" and asks "how much on average do customers spend on each spend on each category". The query uses the `customer_data` table and calculates average spending for six product categories: MntWines, MntFruits, MntmeatProducts, MntFishProducts, MntSweetProducts, and MntGoldProds.

```
1  /* PURCHASE BEHAVIOUR
2  |
3  |  how much on average do customers spend on each spend on each category
4  |  */
5  •  select
6  |    avg(MntWines) as avg_wine_spending,
7  |    avg(MntFruits) as avg_fruit_spending,
8  |    avg(MntmeatProducts) as avg_meat_spending,
9  |    avg(MntFishProducts) as avg_fish_spending,
10 |    avg(MntSweetProducts) as avg_sweet_spending,
11 |    avg(MntGoldProds) as avg_gold_spending
12 |  from
13 |    customer_data;
```

The result grid shows the following data:

avg_wine_spending	avg_fruit_spending	avg_meat_spending	avg_fish_spending	avg_sweet_spending	avg_gold_spending
305.09160649819495	26.3560	166.99593862815885	37.6376	27.0289	43.9653

#### 4. Campaign Response:

i. How many customers accepted each campaign offer?

The screenshot shows the SQL Project interface with a query editor and a result grid. The query is titled "CAMPAIGN RESPONSE" and asks "how many customers accepted each campaign offer?". The query uses the `customer_data` table and calculates the number of customers who accepted five different campaign offers: AcceptedCmp1, AcceptedCmp2, AcceptedCmp3, AcceptedCmp4, and AcceptedCmp5.

```
1  /*
2  |  CAMPAIGN RESPONSE
3  |  how many customers accepted each campaign offer?
4  |  */
5  •  select
6  |    sum(AcceptedCmp1) as Campaign1_Accepted,
7  |    sum(AcceptedCmp2) as Campaign2_Accepted,
8  |    sum(AcceptedCmp3) as Campaign3_Accepted,
9  |    sum(AcceptedCmp4) as Campaign4_Accepted,
10 |    sum(AcceptedCmp5) as Campaign5_Accepted
11 |  from
12 |    customer_data;
```

The result grid shows the following data:

Campaign1_Accepted	Campaign2_Accepted	Campaign3_Accepted	Campaign4_Accepted	Campaign5_Accepted
142	30	163	164	162



ii. What is the overall response rate for the last campaign?

The screenshot shows the SQL Developer interface with a query window titled 'SQL File 14'. The query is as follows:

```
1  /*
2  CAMPAIGN RESPONSE
3  what is the overall response rate for the last campaign?
4  */
5  select
6  Response, count(*) as Response_count
7  from
8  customer_data
9  group by Response;
```

The result grid shows the following data:

Response	Response_count
1	333
0	1883

## 5. Online and Offline Purchases

i. How many purchases are made through the website, catalog, and in-store?

The screenshot shows the SQL Developer interface with a query window titled 'SQL File 15'. The query is as follows:

```
1  /*
2  ONLINE AND OFFLINE PURCHASES
3  how many purchases are made through website,catalog and instore?
4  */
5  select
6  NumWebPurchases, NumCatalogPurchases, NumStorePurchases
7  from
8  customer_data;
9
```

The result grid shows the following data:

	NumWebPurchases	NumCatalogPurchases	NumStorePurchases
8	10	4	
1	1	2	
8	2	10	
2	0	4	
5	3	6	
6	4	10	
7	3	7	
4	0	4	
3	0	2	



ii. What is the average number of web visits per month?

The screenshot shows the SQL Enterprise Manager interface. The left pane displays the 'ecom' database schema, with 'customer\_data' selected. The central pane contains the following SQL query:

```
1 /*
2  ONLINE AND OFFLINE PURCHASES
3  what is the average number of web vists per month?
4  */
5  select
6  avg(NumWebVisitsMonth) as avg_web_visits
7  from
8  customer_data;
```

The bottom pane shows the 'Result Grid' with the following data:

avg_web_visits
5.319043321299639

## 6. Complaints and Engagement

i. How many customers have complained in the last 2 years?

The screenshot shows the SQL Enterprise Manager interface. The left pane displays the 'ecom' database schema, with 'customer\_data' selected. The central pane contains the following SQL query:

```
1 /*
2  COMPLAINTS AND ENGAGEMENT
3  how many customers have complained in the last 2 years
4  */
5  select
6  count(*) as complaint_count
7  from
8  customer_data
9  where
10 complain=1;
```

The bottom pane shows the 'Result Grid' with the following data:

complaint_count
21

- ii. What is the overall engagement rate (accepted any campaign or responded) among customers?

The screenshot shows the SQL Studio interface with a query editor and a result grid. The query is as follows:

```
1  /*  
2  COMPLAINS AND ENGAGEMENT  
3  What is the overall engagement rate (accepted any campaign or responded) among customers?  
4  */  
5  select  
6  (sum(AcceptedCmp1) + sum(AcceptedCmp2) + sum(AcceptedCmp3) + sum(AcceptedCmp4) + sum(AcceptedCmp5)  
7  as Total_Engaged_Customer  
8  from  
9  customer_data;
```

The result grid shows the following data:

Total_Engaged_Customer
661

## 7. Discounts and Deals

- i. How many deals are made with discounts?

The screenshot shows the SQL Studio interface with a query editor and a result grid. The query is as follows:

```
1  /* DISCOUNTS AND DEALS  
2  How many deals are with discounts  
3  */  
4  select  
5  sum(NumDealsPurchases) as Total_Deals,  
6  avg(NumDealsPurchases) as Avg_deals,  
7  count(*) as Total_Customers  
8  from  
9  customer_data;
```

The result grid shows the following data:

Total_Deals	Avg_deals	Total_Customers
5149	2.3236	2216

ii. What is the average number of deals and purchases made by customers?

The screenshot shows the SQL Server Enterprise Manager interface. The left pane displays the 'ecom' database schema with tables, columns, indexes, foreign keys, triggers, views, stored procedures, and functions. The right pane shows a query window with the following SQL code:

```
1  /* DISCOUNTS AND DEALS
2  whats is the average number of deals and purchases made by customer?
3  */
4  select
5  avg(NumWebPurchases + NumCatalogPurchases + NumStorePurchases) as AVG_Total_Purchases
6  from
7  customer_data;
```

The bottom pane shows the 'Result Grid' with the following data:

AVG_Total_Purchases
12.55731046931408

## 8. Campaign Effectiveness

i. What is the conversion rate for each campaign?

The screenshot shows the SQL Server Enterprise Manager interface. The left pane displays the 'ecom' database schema. The right pane shows a query window with the following SQL code:

```
3  what is the conversion rate for each campaign
4  */
5  select
6  Campaign, sum(Response) as accepted, count(*) as Total
7  from(
8  select 'Campaign1' as Campaign, Response from customer_data where AcceptedCmp1=1
9  union all
10 select 'Campaign2' as Campaign, Response from customer_data where AcceptedCmp2=1
11 union all
12 select 'Campaign3' as Campaign, Response from customer_data where AcceptedCmp3=1
13 union all
14 select 'Campaign4' as Campaign, Response from customer_data where AcceptedCmp4=1
15 union all
16 select 'Campaign5' as Campaign, Response from customer_data where AcceptedCmp5=1
17 union all
18 select 'LastCampaign' as Campaign, Response from customer_data where Response=1
19 ) As CampaignData
20 group by Campaign;
```

The bottom pane shows the 'Result Grid' with the following data:

Campaign	accepted	Total
Campaign1	79	142
Campaign2	20	30
Campaign3	77	163
Campaign4	62	164
Campaign5	91	162
LastCampaign	333	333

ii. Which campaign has the highest acceptance rate?

The screenshot shows a SQL query in SQL File 22. The query calculates the acceptance rate for each campaign by summing responses and counting total entries. The results are displayed in a grid below the query editor.

```
3  what is the conversion rate for each campaign
4  */
5  Select
6  Campaign, sum(Response) as accepted, count(*) as Total
7  From(
8  Select 'Campaign1' as Campaign, Response from customer_data where AcceptedCmp1=1
9  Union all
10 Select 'Campaign2' as Campaign, Response from customer_data where AcceptedCmp2=1
11 Union all
12 Select 'Campaign3' as Campaign, Response from customer_data where AcceptedCmp3=1
13 Union all
14 Select 'Campaign4' as Campaign, Response from customer_data where AcceptedCmp4=1
15 Union all
16 Select 'Campaign5' as Campaign, Response from customer_data where AcceptedCmp5=1
17 Union all
18 select 'LastCampaign' as Campaign, Response from customer_data where Response=1
19 ) As CampaignData
20 group by Campaign;
```

Campaign	accepted	Total
Campaign1	79	142
Campaign2	20	30
Campaign3	77	163
Campaign4	62	164
Campaign5	91	162
LastCampaign	333	333

## 9. Web Engagement and Purchases

i. What is the correlation between the number of web visits and web purchases?

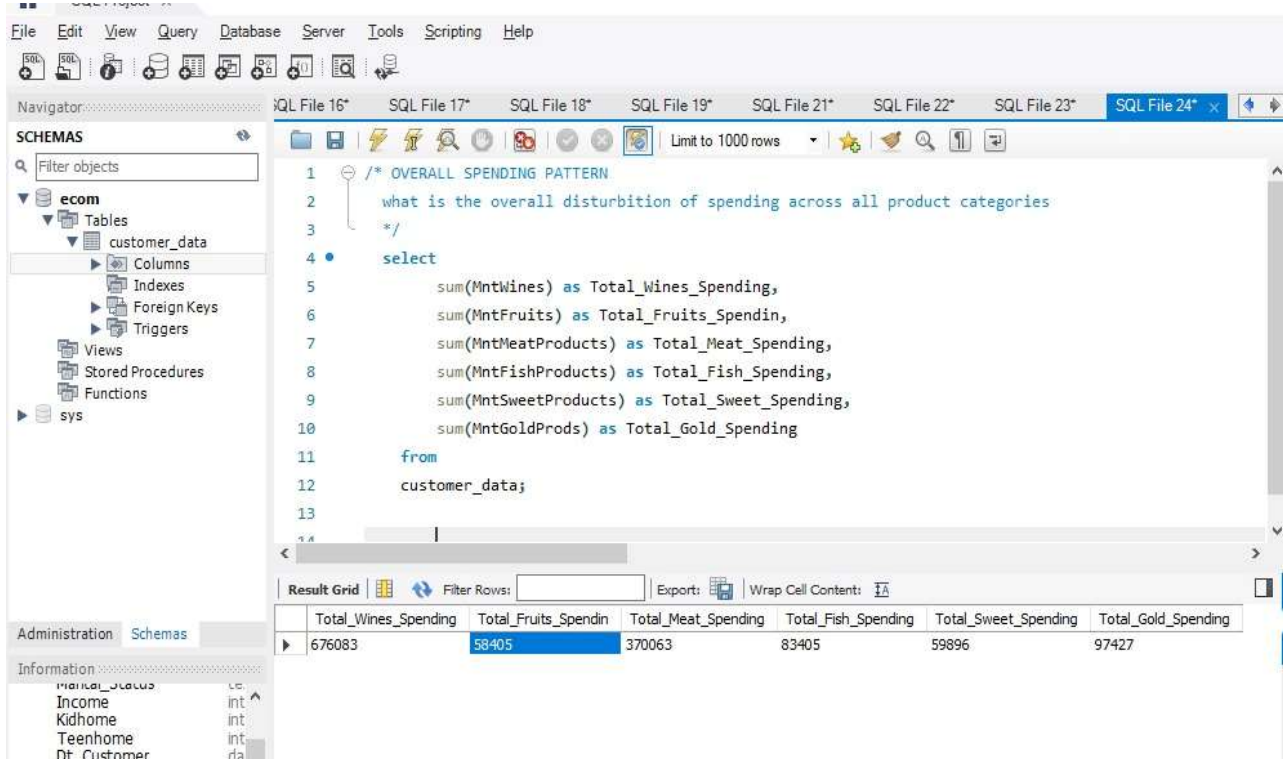
The screenshot shows a SQL query in SQL File 23. The query selects the number of web visits per month and the number of web purchases from the customer\_data table. The results are displayed in a grid below the query editor.

```
1  /*
2  WEB ENGAGEMENT AND PURCHASES
3  what is the correlation between the no of web visits and web purchases
4  */
5  Select
6  NumWebVisitsMonth, NumWebPurchases
7  from
8  customer_data;
```

NumWebVisitsMonth	NumWebPurchases
7	8
5	1
4	8
6	2
5	5
6	6
6	7
8	4
9	3
20	1
8	2
2	3
6	6
8	1
3	7

## 10. Overall Spending Patterns

i. What is the overall distribution of spending across all product categories?



The screenshot shows the SQL Server Enterprise Manager interface. The left pane displays the 'SCHEMAS' tree with 'ecom' expanded, showing 'Tables', 'Columns', 'Indexes', 'Foreign Keys', 'Triggers', 'Views', 'Stored Procedures', 'Functions', and 'sys'. The right pane shows a query in 'SQL File 24' with the following SQL code:

```
1  /* OVERALL SPENDING PATTERN
2  what is the overall disturbition of spending across all product categories
3  */
4  select
5      sum(MntWines) as Total_Wines_Spending,
6      sum(MntFruits) as Total_Fruits_Spendin,
7      sum(MntMeatProducts) as Total_Meat_Spending,
8      sum(MntFishProducts) as Total_Fish_Spending,
9      sum(MntSweetProducts) as Total_Sweet_Spending,
10     sum(MntGoldProds) as Total_Gold_Spending
11  from
12     customer_data;
13
```

Below the query editor, the 'Result Grid' is displayed with the following data:

Total_Wines_Spending	Total_Fruits_Spendin	Total_Meat_Spending	Total_Fish_Spending	Total_Sweet_Spending	Total_Gold_Spending
676083	58405	370063	83405	59896	97427

## 05 Internship Outcomes and Conclusion

### Key Outcomes

1. **Enhanced Customer Understanding:** Gained deep insights in to customer demographics and behavior, which can inform future marketing and sales strategies.
2. **Improved Campaign Strategies:** Identified effective marketing tactics and areas for improvement, allowing the company to optimize its campaigns.
3. **Actionable Recommendations:** Provided strategic recommendations to the marketing team to enhance customer engagement and increase sales.

### Conclusion

The internship provided a comprehensive understanding of data analytics in an e-commerce context. The insights derived from the analysis can significantly contribute to the

company's marketing strategies and overall business performance. By leveraging customer data, the company can better understand its customers, tailor its marketing efforts, and ultimately drive growth and success.