E-COMMERCE COMPANY CASE STUDY

BY. SURAJIT BAL

INTRODUCING

SURAJIT BAL

Data Analyst with a proven track record of analyzing complex data sets, identifying trends, and providing actionable insights to drive business decisions. Skilled in statistical analysis, data visualization, and data mining techniques. Proficient in SQL, Python, and Excel. Strong problem-solving abilities and a keen attention to detail.

OVERVIEW

The project involves leveraging extensive databases within a dynamic e-commerce company to extract actionable insights. The objective is to inform key departments—such as marketing and supply chain—on strategic decisions that optimize operations, improve customer satisfaction, and enhance sales performance. This initiative aims to utilize data analysis to drive forward the company's business strategies effectively.

OBJECTIVES

* Customer Insights: Understanding our customer base to tailor marketing strategies.

* Product Analysis: Evaluating product performance to inform stock and sales strategies.

* Sales Optimization: Analyzing sales data to identify trends, opportunities, and areas for improvement.

* Inventory Management: Managing stock levels to ensure product availability while minimizing excess inventory.

DATASET DETAILS

* **CUSTOMERS DATASET:** INFORMATION ON CUSTOMERS, INCLUDING CUSTOMER_ID, NAME, AND LOCATION.

* **PRODUCTS DATASET:** DATA ON PRODUCT INVENTORY, INCLUDING PRODUCT_ID, NAME, CATEGORY, AND PRICE.

* **ORDERS DATASET:** RECORDS OF ORDERS PLACED TO CUSTOMERS, INCLUDING ORDER_ID, ORDER_DATE, CUSTOMER_ID, AND TOTAL_AMOUNT.

* **ORDERDETAILS DATASET:** RECORDS OF ORDERED PRODUCTS, INCLUDING ORDER_ID, PRODUCT_ID, QUANTITY, AND PRICE_PER_UNIT.

DATA ANALYSIS IN MYSQL

Problem statement:
Identify the top 3 cities with the highest number of customers to determine key markets for targeted marketing and logistic optimization.

```sql
select location, count(customer_id) number_of_customers
from Customers
Group by 1 Order by 2 desc limit 3;
```

location	number_of_custom...
Delhi	16
Chennai	15
Jaipur	11

Problem statement

Determine the distribution of customers by the number of orders placed. This insight will help in segmenting customers into one-time buyers, occasional shoppers, and regular customers for tailored marketing strategies.

```sql
With CTE as (select customer_id, count(order_id) NumberOfOrders
from Orders
Group by 1 Order by 2 asc)
select NumberOfOrders, count(customer_id) as CustomerCount
from CTE
Group By 1 Order by 1 asc;
```

NumberOfOrde...	CustomerCount
1	26
2	26
3	18
4	6
5	6
6	1
8	1

Problem statement

Identify products where the average purchase quantity per order is 2 but with a high total revenue, suggesting premium product trends.

```sql
select Product_id, avg(quantity) AvgQuantity, sum(price_per_unit * quantity) TotalRevenue
from OrderDetails
Group by 1 Having avg(quantity) = 2 Order by 3 desc;
```

Product_id	AvgQuantity	TotalRevenue
1	2.0000	1620000
8	2.0000	390000

Problem statement

For each product category, calculate the unique number of customers purchasing from it. This will help understand which categories have wider appeal across the customer base.

```
select p.category category, count(distinct(o.customer_id)) unique_customers
from Products p INNER JOIN OrderDetails od using(product_id)
INNER JOIN orders o using(order_id)
Group by 1 Order by 2 desc;
```

	category	unique_custom...
▶	Electronics	79
	Wearable Tech	61
	Photography	45

Problem statement

Analyze the month-on-month percentage change in total sales to identify growth trends.

```sql
with CTE as (select date_format(order_date, "%Y-%m") Month, sum(total_amount) TotalSales
from Orders Group by 1)
select *,
round(((TotalSales - lag(TotalSales) over(order by Month))/lag(TotalSales) over(order by Month))*100,2) PercentChange
from CTE;
```

Month	TotalSales	PercentChange
2023-03	789000	NULL
2023-04	1704000	115.97
2023-05	1582000	-7.16
2023-06	1040000	-34.26
2023-07	2568000	146.92
2023-08	1800000	-29.91
2023-09	2927000	62.61
2023-10	1497000	-48.86
2023-11	1151000	-23.11
2023-12	2774000	141.01
2024-01	1555000	-43.94
2024-02	396000	-74.53

Problem statement

Examine how the average order value changes month-on-month. Insights can guide pricing and promotional strategies to enhance order value.

```
with CTE as (select date_format(order_date, "%Y-%m") Month, avg(total_amount) AvgOrderValue
from Orders Group by 1)
select Month, AvgOrderValue,
round((AvgOrderValue - lag(AvgOrderValue) over(order by Month)),2) ChangeInValue
from CTE Order by 3 desc;
```

Month	AvgOrderValue	ChangeInValue
2023-12	132095.2381	36178.57
2023-04	81142.8571	20450.55
2023-06	104000.0000	16111.11
2023-08	112500.0000	13730.77
2023-11	95916.6667	12750.00
2023-09	121958.3333	9458.33
2023-05	87888.8889	6746.03
2024-01	129583.3333	-2511.90
2023-07	98769.2308	-5230.77
2023-10	83166.6667	-38791.67
2024-02	44000.0000	-85583.33
2023-03	60692.3077	NULL

Problem statement

Based on sales data, identify products with the fastest turnover rates, suggesting high demand and the need for frequent restocking.

```sql
select product_id, count(quantity) SalesFrequency
from OrderDetails
Group by 1 Order by 2 desc limit 5;
```

product_id	SalesFrequency
7	78
3	68
4	68
2	67
8	65

Problem statement

List products purchased by less than 40% of the customer base, indicating potential mismatches between inventory and customer interest.

```sql
with CTE as (select count(distinct(customer_id)) TotalCustomer
from Customers ),
CTE2 as (select p.Product_id, count(distinct(o.customer_id)) UniqueCustomerCount
from Products p INNER JOIN OrderDetails od using(product_id)
INNER JOIN Orders o using(order_id) INNER JOIN customers c using(customer_id)
Group by 1)
Select CTE2.Product_id, P.Name, CTE2.UniqueCustomerCount
from CTE JOIN CTE2 on 1=1 JOIN products p on CTE2.Product_id = p.product_id
where (CTE2.UniqueCustomerCount/ CTE.TotalCustomer) < 0.40;
```

Product_id	Name	UniqueCustomerCo...
1	Smartphone 6"	36
8	Wireless Earbuds	38

Problem statement

Evaluate the month-on-month growth rate in the customer base to understand the effectiveness of marketing campaigns and market expansion efforts.

```
with CTE as (select customer_id, min(order_date) first_date from Orders
Group by customer_id)
select date_format(first_date, "%Y-%m") FirstPurchaseMonth,
count(distinct customer_id) TotalNewCustomers
from CTE Group by 1 Order by 1;
```

	FirstPurchaseMon...	TotalNewCustom...
▶	2023-03	11
	2023-04	18
	2023-05	11
	2023-06	8
	2023-07	11
	2023-08	9
	2023-09	5
	2023-10	3
	2023-11	1
	2023-12	4
	2024-01	2
	2024-02	1

Problem statement

Identify the months with the highest sales volume, aiding in planning for stock levels, marketing efforts, and staffing in anticipation of peak demand periods.

```
select date_format(order_date, "%Y-%m") Month, sum(total_amount) TotalSales
from Orders
Group by 1 Order by 2 desc limit 3;
```

Month	TotalSales
2023-09	2927000
2023-12	2774000
2023-07	2568000

BUSINESS INSIGHTS

- The cities to be focused on as part of the marketing strategies are Delhi, Chennai, and Jaipur.
- The downward growth trend in the customer base implies that the marketing campaign by the company is not very effective.
- As the number of orders increases, the customer count decreases.
- The customer category in which the company experienced the most activity is the "Occasional shoppers".

- Product 1 exhibits the highest total revenue with an average purchase quantity of two.
- The electronics products need more focus as they are in high demand among the customers.
- During February 2024, sales experienced the largest decline.
- According to the sales trend analysis, sales from March 2023 to August 2023 fluctuated with no clear trend.
- The month of December 2023 has the highest change in the average order value.

- Product ID 7 has the highest turnover rate and needs to be restocked frequently.
- Poor visibility on the platform might cause certain products to have a purchase rate below 40% of the total customer base. Implementing targeted marketing campaigns to raise awareness and interest may help improve the sales of these underperforming products.
- The analysis suggests that the months of September and December will require major restocking of products and increased staff.

THANK YOU!

+91-89810-44773

balsurajit93@gmail.com

https://github.com/Surajit-1108

https://www.linkedin.com/in/surajitbal/