**Customer Behaviour Analysis**

**Name: Saurabh Shivdas Sonawane**

**Date: September 2025**

**⮚Introduction:**

The purpose of this project is to study the customer behaviour.

Nowadays, online shopping has become an integral part of everyday life. Millions of people order products through the ecommerce platforms ( Flipkart, Amazon, Meesho etc) Therefore it is essential for e-commerce platforms to understand their customers purchasing behaviour to recommend relevant products, improve user experience, and strengthen customer loyalty.

**⮚Problem Statement:**

The dataset contained customer details, product information, orders and sales records. However, the raw data did not provide answers to important business questions such as:

1) who the high-spending customers are

2) most popular products

3) The customers who are repeated buyers

**⮚Objective:**

To analyze the customer purchasing behaviour, product sales pattern and sales trends using SQL to derive actionable insights.

**⮚Methodology:**

The analysis was done using SQL. Queries were wrtitten to join multiple tables and filter relevant information. we loaded the data from each Excel file in SQL using the following Query

|  |
| --- |
| ⬩ The following query creates the brands table and loads data into it  create table brands  (brand\_id int ,  brand\_name varchar(50),  primary key(brand\_id));  load data infile "C:/ProgramData/MySQL/MySQL Server 8.0/Uploads/brands.csv"  into table brands  fields terminated by ',' enclosed by '"'lines terminated by '\n'  ignore 1 rows; |

Here, we created a table named brands and loaded the data from file name brands.csv. similarly ,we created the other tables and load their data using similar commands.

**⬩Key Buisness Questions Answered Using SQL:**

1) Repeat Buyers: To identify customers who buy more than once and customers who buy only once

2) Customer spending groups: Partitioning customers according to their spending behaviours (High,Medium,Low)

3)Top products: products with highest sales

4) Product sales categories: Partitoning the products according to their demand (High,Average,Low) Using the count of no of times product orderd

5) Discount Analysis: Discount Availed by each customers

6)Average spending and No. of orders: Total purchase, No. of orders and

average spending per customer

7)Brand analysis: popular products for each brand

**⮚ Findings:**

**1)Top products:**

**⬩**Query:

|  |
| --- |
| **⬩**This Query finds top 5 products with highest sales  with product\_sales as (  select p.product\_name,count(ot.product\_id) as totall\_sales  from orders o join order\_item ot on o.order\_id=ot.order\_id  join products p on p.product\_id=ot.product\_id  group by p.product\_name),  ranked\_products as(  select product\_name,  totall\_sales,  rank() over(  order by totall\_sales desc)  as product\_rank  from product\_sales)  select \* from ranked\_products  where product\_rank<=5  order by product\_rank asc; |

**⬩**Output:

|  |  |  |
| --- | --- | --- |
| product\_name | total\_sales | product\_rank |
| Electra Townie Original 21D - 2016 | 193 | 1 |
| Electra Cruiser 1 (24-Inch) - 2016 | 193 | 1 |
| Electra Townie Original 7D EQ - 2016 | 185 | 3 |
| Electra Girl's Hawaii 1 (16-inch) - 2015/2016 | 180 | 4 |
| Surly Ice Cream Truck Frameset - 2016 | 110 | 5 |

**⬩**Insights: A few products consistently generated the highest sales volume, showing a clear product preferance among customers.

**2) Products Sales Categories:**

**⬩**Query:

|  |
| --- |
| **⬩**Query divides product into 3 category according to demand(High, Avg, Low)  select p.product\_name,  count(ot.product\_id) as no\_of\_times\_ordered,  case when count(ot.product\_id)>=0 and count(ot.product\_id)<40 then 'Low Demand'  when count(ot.product\_id)>=40 and count(ot.product\_id)<100 then 'Average Demand'  else 'High Demand'  end as product\_demand  from orders o join order\_item ot on o.order\_id=ot.order\_id  join products p on p.product\_id=ot.product\_id  group by p.product\_id; |

**⬩**Output:

|  |  |  |
| --- | --- | --- |
| Product\_name | No\_of\_time\_orderd | Product\_Demand |
| Ritchey Timberwolf Frameset - 2016 | 77 | Average Demand |
| Surly Wednesday Frameset - 2016 | 86 | Average Demand |
| Surly Troll Frameset - 2017 | 29 | Low Demand |
| Surly Ice Cream Truck Frameset - 2016 | 110 | High Demand |
| … | … | … |

**⬩**Insights: Products categorized as High Demand dominated the overall sales, while Low Demand products contributed minimally

**3)Average Spending & Orders:**

**⬩** Query:

|  |
| --- |
| **⬩**This Query finds total purchase, no of orders placed and average purchase of each customer  select concat(c.first\_name,' ',c.last\_name) as customer\_name,  sum(ot.quantity\*ot.list\_price\*(1-ot.discount)) as totall\_order\_price,  count(ot.order\_id) as order\_count,  avg(ot.quantity\*ot.list\_price\*(1-ot.discount)) as avg\_order\_price  from customer c join orders o on c.customer\_id=o.customer\_id  join order\_item ot on o.order\_id=ot.order\_id  group by c.customer\_id order by avg\_order\_price desc; |

**⬩**Output:

|  |  |  |  |
| --- | --- | --- | --- |
| Customer\_name | Total\_order\_price | Order\_count | Avg\_order\_price |
| Hae Ramirez | 12090 | 1 | 12090 |
| Marilyn Frank | 9499.999993 | 1 | 9499.999993 |
| Ester Acevedo | 7999.99997 | 1 | 7999.99997 |
| … | … | … | … |

**⬩**Insights: On average, each customer placed multiple orders, with total spending per customer varying across groups. Frequent buyers also had higher average order values

**4)Repeat Buyers:**

**⬩** Query:

|  |
| --- |
| **⬩**This Query shows no. of orders placed by each customer and whether the customer is Repeat buyer or not  select concat(c.first\_name,' ' ,c.last\_name) as customer\_name,  count(o.customer\_id) as order\_count,  case when count(o.customer\_id)>1 then 'Repeat\_Order'  when count(o.customer\_id)=1 then 'Single\_Buyer'  else 'NO\_Purchase'  end as Buyer\_Type  from customer c join orders o on c.customer\_id=o.customer\_id  group by c.customer\_id; |

**⬩**Output:

|  |  |  |
| --- | --- | --- |
| Customer\_name | orders\_placed | Buyer\_type |
| Debra Burks | 3 | Repeat\_Buyer |
| Shae Hickman | 1 | Single\_Buyer |
| Theo Reese | 2 | Repeat\_Buyer |
| … | … | … |

**⬩**Insights: A significant share of revenue came from repeat buyers, while a smaller portion of customers purchased only once. This highlights the importance of customer retention.

**⮚Conclusion:** This project unfolded like a journey into the minds of customers—revealing what they buy, how much they spends, and what truly drives their loyalty. The analysis showed that repeat customers and high spenders were the real backbone of revenue, while discounts often acted as the spark that encouraged bigger purchases. We also discovered how certain products naturally pulled more attention, becoming the “must-haves” of the store, while others faded quietly in the background. Popular brands continued to win trust, reminding us that brand value shapes customer choices as much as price. Together, these insights highlight a powerful story: with SQL, raw and scattered data can be transformed into a clear narrative of customer behavior. For businesses, it means focusing not just on sales, but on relationships-retaining loyal buyers, crafting smarter product strategies, and ultimately driving growth with purpose.