E-Commerce Data Analysis Using SQL

Insights into Profitability, Returns, Customer Behavior & Time Trends

Presented by:

Samuel Mati

Data Analyst | Python | SQL | Power BI | AI-Driven Insights

Date: August 2025

Preview the Dataset

```
USE hermira;

-- Show all records
SELECT * FROM ecomerce;
```

Which products are consistently unprofitable?

```
WITH product_profit AS (
    SELECT Product_Name, SUM(Profit) AS total_profit
    FROM ecomerce
    GROUP BY Product_Name
)
SELECT Product_Name, total_profit
FROM product_profit
WHERE total_profit < 0
ORDER BY total_profit ASC
LIMIT 10;
```

Which product categories are unprofitable overall?

```
SELECT Category, ROUND(SUM(profit)) AS total_profit
FROM ecomerce
GROUP BY Category
ORDER BY SUM(profit) DESC;
```

Which regions generate the most and least profit?

```
SELECT Region, ROUND(SUM(profit)) AS total_profit
FROM ecomerce
GROUP BY Region
ORDER BY total_profit DESC;
```

Do discounts negatively impact profit margins?

```
SELECT
    discount,
    ROUND(SUM(profit), 2) AS total_profit,
    ROUND(SUM(profit) / NULLIF(SUM(sales), 0), 2) AS profit_margin
FROM ecomerce
GROUP BY discount
ORDER BY profit_margin ASC;
```

Which shipping modes lead to the most delivery delays?

```
SELECT
   `Ship Mode`,
   AVG(DATEDIFF(`Ship Date`,`Order Date`)) AS avg_shipping_delay
FROM ecomerce
GROUP BY `Ship Mode`
ORDER BY avg_shipping_delay DESC;
```

What percentage of products are returned, and how does this affect profits?

```
SELECT
    discount,
    ROUND(SUM(profit), 2) AS total_profit,
    ROUND(SUM(profit) / NULLIF(SUM(sales), 0), 2) AS profit_margin
FROM ecomerce
GROUP BY discount
ORDER BY profit_margin ASC;
```

Which shipping modes lead to the most delivery delays?

```
-- Return Rate (%) and Count
SELECT
  Returned,
  COUNT(*) AS total_returns,
  ROUND(100.0 * COUNT(*) / (SELECT COUNT(*) FROM ecomerce), 2) AS percent_of_orders
FROM ecomerce
GROUP BY Returned
ORDER BY percent_of_orders DESC;

-- How Returns Affect Profits
SELECT
   Returned,
   ROUND(SUM(profit), 2) AS total_profit,
   ROUND(AVG(profit), 2) AS avg_profit_per_order,
   COUNT(*) AS num_orders
FROM ecomerce
GROUP BY Returned
ORDER BY total_profit;
```

Which customer segments or categories have the highest return rates?

```
-- Return Rate by Customer Segment
SELECT
  Segment,
  COUNT(*) AS total_orders,
  SUM(Returned) AS returned_orders,
  ROUND(100 * SUM(Returned) / COUNT(*), 2) AS return_rate
FROM ecomerce
GROUP BY Segment
ORDER BY return_rate DESC;
-- Return Rate by Product Category
SELECT
  Category,
  COUNT(*) AS total_orders,
  SUM(Returned) AS returned_orders,
  ROUND(100 * SUM(Returned) / COUNT(*), 2) AS return_rate
FROM ecomerce
GROUP BY Category
ORDER BY return_rate DESC;
```

Who are the top 10 customers by purchases or lifetime value?

```
-- By Total Purchases
SELECT
    `Customer Name`,
    COUNT(*) AS total_purchases
FROM ecomerce
GROUP BY `Customer Name`
ORDER BY total_purchases DESC
LIMIT 10;


-- By Total Sales
SELECT
    `Customer Name`,
    ROUND(SUM(Sales), 2) AS total_sales,
    ROUND(SUM(Profit), 2) AS total_profit
FROM ecomerce
GROUP BY `Customer Name`
ORDER BY total_sales DESC
LIMIT 10;
```

Which customer segments are the most and least profitable?

```sql
-- Most Profitable Segments
SELECT
  Segment,
  ROUND(SUM(Profit), 2) AS total_profit
FROM ecomerce
GROUP BY Segment
ORDER BY total_profit DESC;

-- Least Profitable Segments
SELECT
  Segment,
  ROUND(SUM(Profit), 2) AS total_profit
FROM ecomerce
GROUP BY Segment
ORDER BY total_profit ASC;
```

What are the monthly and quarterly sales & profit trends?

```
-- Monthly Trends
SELECT
  DATE_FORMAT(`Order Date`, '%Y-%M') AS Month,
  ROUND(SUM(Sales), 2) AS total_sales,
  ROUND(SUM(Profit), 2) AS total_profit
FROM ecomerce
GROUP BY Month
ORDER BY total_sales DESC;

-- Quarterly Trends
SELECT
  CONCAT(YEAR(`Order Date`), '-Q', QUARTER(`Order Date`)) AS quarter,
  ROUND(SUM(Sales), 2) AS total_sales,
  ROUND(SUM(Profit), 2) AS total_profit
FROM ecomerce
GROUP BY quarter
ORDER BY total_sales DESC;
```

Dynamic KPIs for 2024 Consumer Segment

```
WITH Dynamic_KPIs AS(
SELECT
    YEAR(`Order Date`) AS Year,
    Segment,
    ROUND(SUM(Sales), 2) AS total_sales,
    ROUND(SUM(Profit), 2) AS total_profit,
    ROUND(SUM(Profit)/ NULLIF(SUM(Sales),0) *100, 2) AS profit_margin_rate,
    ROUND(SUM(Returned)/COUNT(*) *100, 2) AS return_rate,
    ROUND(SUM(CASE
            WHEN DATEDIFF(`Ship Date`, `Order Date`) > 0 THEN 1
            ELSE 0
            END ) / COUNT(*) *100, 2) AS delivery_delay_rate
FROM ecomerce
GROUP BY YEAR(`Order Date`), Segment
)
SELECT * FROM Dynamic_KPIs WHERE year = 2024 AND Segment = 'Consumer';
```

Thanks for reading!

What else would you analyze if this were your dataset?

Drop a comment or DM if you'd like access to the dataset or full code.

Let’s connect!!

Samuel Mati

https://sam-analyst.vercel.app/