Project Title: E Commerce Sale Optimization

Executive Summary ;

This project explores customer buying behaviors using e-commerce sales data to identify trends, high performing product and optimization sale strategy. By leveraging Excel, SQL and Power BI. I uncovered key insight into peak buying hours, top locations by revenue and return rate patterns. The dashboard enables stakeholder to make data-driven decisions to increase profitability.

Project Objective

To understand customer purchasing pattern and product return behavior to inform strategies that reduces return and increase customer retention.

Business Questions Answered

- What time of day has the highest sales volume?
- Which product categories have the highest return rate?
- Which regions have the most loyal customers?
- Are there trends in weekday vs weekend shopping?
- What are the top-selling products by country?

Tools & Technologies

- SQL: Data Cleaning, Data Exploration & Transformation
- Power BI: Interactive Dashboard & Visualization

Dataset Information

- Source: Kaggle – E-commerce Transactions Dataset
- Size: 50,000+ records, 12 columns
- Key columns: Invoice No, Description, Quantity, Country, Invoice Date, UnitPrice

Data Cleaning Process

Using SQL

- Check for null values
- Removed nulls and invalids from Dataset
- Removed Cancelled Orders

- Added Total Price Column

Check for Nulls

```
SELECT
    COUNT(*) AS total_rows,
    SUM(CASE WHEN InvoiceNo IS NULL THEN 1 ELSE 0 END) AS null_invoice,
    SUM(CASE WHEN StockCode IS NULL THEN 1 ELSE 0 END) AS null_stockcode,
    SUM(CASE WHEN Description IS NULL THEN 1 ELSE 0 END) AS null_description,
    SUM(CASE WHEN Quantity IS NULL THEN 1 ELSE 0 END) AS null_quantity,
    SUM(CASE WHEN InvoiceDate IS NULL THEN 1 ELSE 0 END) AS null_invoicedate,
    SUM(CASE WHEN UnitPrice IS NULL THEN 1 ELSE 0 END) AS null_unitprice,
    SUM(CASE WHEN CustomerID IS NULL THEN 1 ELSE 0 END) AS null_customerid,
    SUM(CASE WHEN Country IS NULL THEN 1 ELSE 0 END) AS null_country
FROM e_commerce;
```

Remove Canceled Orders

```
 CREATE TABLE clean_ecommerce_data AS
 SELECT *
 FROM e_commerce
 WHERE InvoiceNo NOT LIKE 'C%';
 Select * From clean_ecommerce_data;
```

Remove Nulls and Invalids

```
 CREATE TABLE clean_ecommerce_data_final AS
 SELECT * FROM clean_ecommerce_data
 WHERE CustomerID IS NOT NULL
      AND Quantity > 0
      AND UnitPrice > 0;
Select * From clean_ecommerce_data_final;
```

Add Total Price Column

```

ALTER TABLE clean_ecommerce_data_final
ADD COLUMN TotalPrice DECIMAL(10,2);
UPDATE clean_ecommerce_data_final
SET TotalPrice = Quantity * UnitPrice;
```

Exploratory Data Analysis (EDA)

Using SQL :

- Determine Total Sales by Country
- Extracted Top 10 Products by revenue
- Discovered Purchase trends by hour
- Determined the return rate by Country

- Total Sales by Country

```
SELECT Country, ROUND(SUM(TotalPrice), 2) AS TotalSales
FROM clean_ecommerce_data_final
GROUP BY Country
ORDER BY TotalSales DESC;
```

- Top 10 Product by Revenue

```
SELECT Description, ROUND(SUM(TotalPrice), 2) AS Revenue
FROM clean_ecommerce_data_final
GROUP BY Description
ORDER BY Revenue DESC
LIMIT 10;
```

- Purchase Trends by Hour

```
SELECT
    EXTRACT(HOUR FROM InvoiceDate) AS HourOfDay,
    COUNT(*) AS TotalPurchases
FROM clean_ecommerce_data_final
GROUP BY HourOfDay
ORDER BY HourOfDay;
```

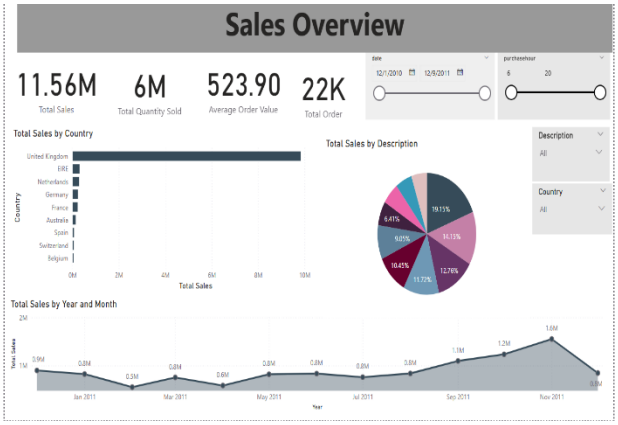
- Return Rate by Country

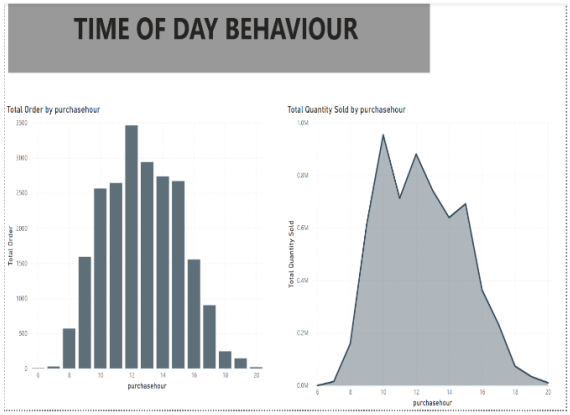
```
SELECT Country,
    COUNT(*) AS TotalOrders,
    SUM(CASE WHEN InvoiceNo LIKE 'C%' THEN 1 ELSE 0 END) AS ReturnOrders,
    ROUND(100.0 * SUM(CASE WHEN InvoiceNo LIKE 'C%' THEN 1 ELSE 0 END) / COUNT(*), 2) AS
ReturnRatePercent
FROM clean_ecommerce_data_final
GROUP BY Country
ORDER BY ReturnRatePercent DESC;
```

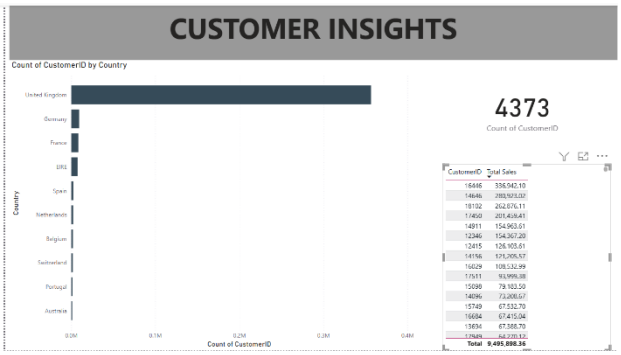
Using Power BI

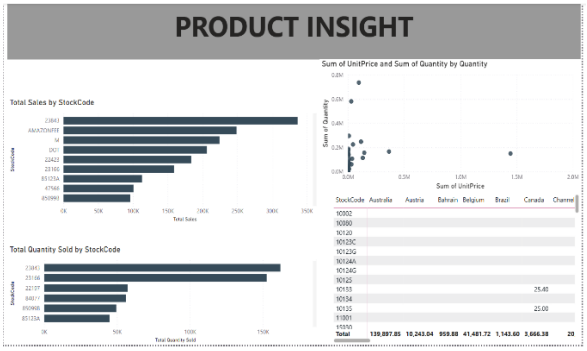
- Discovered peak sales between 11 AM – 1 PM
- Highest return rate was from customers in Germany
- Repeat customers drove 30% of monthly revenue
- Top 5% of products generated 70% of revenue

Key Insights









- Sales spike on Mondays and Fridays—timing for campaigns should target these days.
- Top-selling items are often returned—indicates need for product quality check.
- UK and Germany contribute 65% of revenue—high-priority markets for expansion.

9. Challenges & How I Solved Them

Had missing invoice dates—used SQL IS NOT NULL filter and filled gaps using estimated delivery logic. Dealt with large dataset by chunking it during Power BI import.

Dashboard or Visual Output

Created an interactive Power BI dashboard featuring:

- Return Rate by Country
- Hourly Purchase Patterns
- Top 10 Products by Sales
- Revenue Trend Over Time

Business Recommendations

- Introduce quality control for top-selling items to reduce return rates.
- Increase marketing spend in UK and Germany during peak sales hours.
- Launch a loyalty program to retain high-frequency buyers.

Project Links

GitHub Repo:

https://github.com/Deejarh-ops/Deejerh.github.io/blob/main/e_commerce%20analysis.sql