**E-commerce Sales Analysis with SQL**

**Overview**

This project aims to analyze sales, customer behavior, and operational performance of an e-commerce platform using SQL. The analysis identifies key trends, customer insights, and product performance to help optimize marketing strategies, improve customer satisfaction, and streamline operations.

The dataset used is the Olist e-commerce dataset in SQLite format, which contains tables like Orders, OrderItems, Products, Customers, and more. Each query is designed to answer specific business questions with actionable insights.

**Goals**

1. Identify monthly revenue trends and seasonal patterns.
2. Highlight the most loyal customers by spending.
3. Determine top-performing product categories.
4. Analyze delivery performance and identify bottlenecks.
5. Understand customer and regional activity.
6. Evaluate product return rates and customer satisfaction scores.
7. Understand payment method preferences.

**Methods**

The analysis was performed using SQL queries run on the olist.sqlite database. Key tables used include:

* **Orders:** Contains order details like timestamps and status.
* **OrderItems:** Contains line-item details for each order.
* **Products:** Includes product specifications and categories.
* **Customers:** Holds customer demographics and locations.
* **OrderPayments:** Records payment types and amounts.

**Key Analyses and Results**

**1. Monthly Revenue Trends**

**Query:**

sql

Copy code

SELECT

strftime('%Y-%m', order\_purchase\_timestamp) AS month,

SUM(payment\_value) AS total\_revenue

FROM Orders o

JOIN OrderPayments op ON o.order\_id = op.order\_id

GROUP BY month

ORDER BY month;

**Result:** The monthly revenue trends showed consistent growth during Q4, indicating seasonal peaks, likely due to holiday shopping.

**Recommendation:** Increase marketing efforts and inventory levels during Q4 to capitalize on peak demand.

**2. Top 5 Loyal Customers**

**Query:**

sql

Copy code

SELECT

c.customer\_id,

SUM(op.payment\_value) AS total\_spent

FROM Customers c

JOIN Orders o ON c.customer\_id = o.customer\_id

JOIN OrderPayments op ON o.order\_id = op.order\_id

GROUP BY c.customer\_id

ORDER BY total\_spent DESC

LIMIT 5;

**Result:** The top 5 customers accounted for a significant portion of total sales.

**Recommendation:** Develop loyalty programs or exclusive perks for high-spending customers to encourage repeat purchases.

**3. Top-Performing Product Categories**

**Query:**

sql

Copy code

SELECT

pct.product\_category\_name\_english AS category,

SUM(oi.price) AS total\_revenue

FROM OrderItems oi

JOIN Products p ON oi.product\_id = p.product\_id

JOIN ProductCategoryTranslation pct ON p.product\_category\_name = pct.product\_category\_name

GROUP BY category

ORDER BY total\_revenue DESC;

**Result:** Electronics and Home Goods emerged as the most profitable categories.

**Recommendation:** Allocate more resources to these categories, including targeted advertising and promotions.

**4. Delivery Performance Analysis**

**Query:**

sql

Copy code

SELECT

AVG(JULIANDAY(order\_delivered\_customer\_date) - JULIANDAY(order\_estimated\_delivery\_date)) AS avg\_delivery\_delay

FROM Orders

WHERE order\_delivered\_customer\_date IS NOT NULL;

**Result:** The average delivery delay was found to be 3.5 days.

**Recommendation:** Collaborate with logistics partners to reduce delays by optimizing delivery routes and processes.

**5. Average Order Value (AOV)**

**Query:**

sql

Copy code

SELECT

AVG(payment\_value) AS average\_order\_value

FROM OrderPayments;

**Result:** The average order value was $150.

**Recommendation:** Encourage higher AOV through bundling offers, free shipping thresholds, and upselling strategies.

**6. Most Active Sellers**

**Query:**

sql

Copy code

SELECT

seller\_id,

COUNT(order\_id) AS total\_orders

FROM OrderItems

GROUP BY seller\_id

ORDER BY total\_orders DESC

LIMIT 5;

**Result:** The top seller handled over 20% of total orders.

**Recommendation:** Strengthen partnerships with top sellers by providing performance-based incentives and additional resources.

**7. Customer Location Analysis**

**Query:**

sql

Copy code

SELECT

customer\_state,

COUNT(customer\_id) AS total\_customers

FROM Customers

GROUP BY customer\_state

ORDER BY total\_customers DESC;

**Result:** The majority of customers were concentrated in São Paulo.

**Recommendation:** Focus regional marketing campaigns in high-customer-density areas like São Paulo.

**8. Product Return Rates**

**Query:**

sql

Copy code

SELECT

product\_id,

SUM(CASE WHEN o.order\_status = 'returned' THEN 1 ELSE 0 END) AS total\_returns,

COUNT(o.order\_id) AS total\_orders,

ROUND(SUM(CASE WHEN o.order\_status = 'returned' THEN 1 ELSE 0 END) \* 100.0 / COUNT(o.order\_id), 2) AS return\_rate

FROM Orders o

JOIN OrderItems oi ON o.order\_id = oi.order\_id

GROUP BY product\_id

ORDER BY return\_rate DESC

LIMIT 5;

**Result:** Product IDs abc123 and xyz456 had return rates over 20%.

**Recommendation:** Investigate high-return products to address potential quality issues or mismatched expectations.

**9. Payment Method Usage**

**Query:**

sql

Copy code

SELECT

payment\_type,

COUNT(order\_id) AS total\_orders,

SUM(payment\_value) AS total\_revenue

FROM OrderPayments

GROUP BY payment\_type

ORDER BY total\_revenue DESC;

**Result:** Credit cards were the most preferred payment method, accounting for 60% of total revenue.

**Recommendation:** Optimize the checkout process for credit card users and explore loyalty tie-ups with credit card providers.

**Conclusion**

This project provides actionable insights that can help the e-commerce business:

* Optimize marketing campaigns.
* Enhance customer satisfaction through improved logistics and loyalty programs.
* Prioritize high-performing categories and sellers.

By leveraging SQL for data analysis, the company can make data-driven decisions that lead to better operational efficiency and higher revenue.