SQL 100 Days Challenge – Day 53 Reflection

Topic: Customer Orders & Payment Analytics

Dataset: Customers, Orders, Payments

Practice Experience:

- Today's set of questions was a **bit tough**, but I was able to apply logic and solve them step by step.
- The practice covered a wide range of **customer, order, and payment insights** that are highly relevant in real-world e-commerce analytics.
- The **Bonus Challenge** was focused on **query optimization for large datasets**, which pushed me to think beyond query writing and into performance tuning.

Key Learnings:

1. **Revenue by Country:** Aggregating confirmed order revenue per country.
2. **Customer Lifetime Value:** Ranking customers by total spend.
3. **Cancellation Rate:** Calculating cancellation % with CASE WHEN + NULLIF.
4. **Monthly Sales Trend:** Using LAG() to compare revenue month over month.
5. **Top Payment Modes:** Identifying most common modes (Card, UPI, PayPal, etc.).
6. **Repeat Customers:** Detecting customers with >2 completed orders.
7. **Order Timelines:** Leveraging LAG() to track previous orders, revenue change, and days between orders.
8. **Revenue by Tier:** Measuring average spend per loyalty tier (Gold, Silver, Bronze).
9. **Late Payments:** Detecting delays between order date and payment date, categorized by duration.
10. **High-Value Orders:** Using subqueries to filter orders above the global average order value.
11. **Bonus – Query Optimization:**

- Designed an efficient query to calculate **monthly revenue for 2023**.
- Suggested creating a **composite index** (OrderDate, Status) with included columns (TotalAmount, OrderID) for performance improvement on large datasets.

Insights:

- India and USA customers contributed strong revenue, with Gold-tier customers spending the most on average.
- Cancellation rates were higher in some regions, affecting order reliability.
- Monthly sales showed growth with certain spikes around March–April.
- Card was the dominant payment method, followed by PayPal.
- Late payments were mostly minor delays, but categorizing them gave useful operational insights.
- High-value customers (like Emma and Alice) stood out as key contributors.

Skills Reinforced:

- Window functions (LAG) for timelines and comparisons.
- Subqueries for average benchmarks.
- Advanced CASE WHEN classifications.
- Indexing strategies for optimization in large datasets.
- Combining analytics + database performance thinking.

Personal Note:
Today's practice was a good reminder that **SQL isn't only about writing correct queries — it's also about writing queries that scale**. Applying indexing strategies for optimization felt like a step toward real-world database management. The mix of analytics and performance tuning made today's session both challenging and rewarding.

Next Steps:

- Explore indexing strategies for multi-join queries.
- Build customer churn predictions based on cancellation + late payment behavior.
- Simulate performance tests on large synthetic datasets.