SQL 100 Days Challenge – Day 50 Reflection

Topic: E-commerce Customer, Orders & Product Analytics

Dataset: Customers, Orders, OrderDetails

Milestone Note:
Today marks **Day 50** — halfway through this 100-day SQL journey! 🚀
Crossing this milestone feels special, not just because of the number, but because I can clearly see the transformation in my ability to solve tough problems with confidence.

Practice Experience:

- The first 6 questions were a warm-up with familiar concepts: revenue aggregation, category ranking, repeat purchase detection, monthly trends, and return ratios.
- From **Question 7 onwards, the difficulty increased** — cross-category buyers, country-wise contribution, product popularity using ROW_NUMBER(), and CLV calculation.
- The **Bonus Challenge (Churn Prediction)** was the toughest and most exciting — it combined multiple business rules (last order >90 days, <2 orders in last 12 months) into a query that felt like building a predictive rule-based model entirely in SQL.

Key Learnings:

1. **Revenue Insights:** Customer-level revenue, orders, and AOV analysis.
2. **Category Rankings:** Electronics vs Furniture revenue breakdown with % contributions.
3. **Monthly Trends:** Used LAG() + moving averages to capture growth and volatility.
4. **Return Ratios:** Pivoted data to classify customers into high, medium, and low return rates.
5. **Cross-Category Buyers:** Combined STRING_AGG() with HAVING to detect multi-category shoppers.
6. **Country Analysis:** Revenue shares and rankings across geographies.
7. **Product Popularity:** ROW_NUMBER() to rank top-selling product per category.
8. **CLV (Customer Lifetime Value):** Normalized revenue by tenure to rank long-term value.
9. **Churn Prediction:** Advanced CASE WHEN + date logic to flag **"At Risk" customers**.

Insights:

- Certain customers generated revenue across multiple categories, boosting retention.
- Furniture contributed fewer orders but higher AOV compared to Electronics.
- Return ratios highlighted "high return risk" customers that could impact profitability.

- CLV rankings showed that long-term customers with steady purchases are more valuable than occasional big spenders.
- Churn risk analysis flagged customers inactive for >90 days with low engagement — mimicking real CRM churn models.

Skills Reinforced:

- Window functions (LAG, ROW_NUMBER, RANK)
- Complex CASE WHEN logic for segmentation
- Date handling (DATEDIFF, moving averages, inactivity detection)
- Pivoting + classification with STRING_AGG()
- Applying SQL for **predictive-style business rules**

Personal Note:
Reaching Day 50 feels like a huge win 🎉. The questions were tough, especially the bonus churn prediction, but they showcased how SQL can directly support **customer retention strategies in e-commerce**. What once felt advanced now feels achievable, and I can't wait to carry this momentum into the next 50 days.

Next Steps:

- Apply churn prediction logic on larger datasets.
- Build CLV cohorts to compare customer groups.
- Extend monthly revenue trends into full time-series forecasting prep.