SQL 100 Days Challenge – Day 57 Reflection

Topic: E-commerce Logistics & Customer Experience Analytics

Dataset:

- **Customers** (CustomerID, Name, Country, JoinDate)
- **Orders** (OrderID, CustomerID, OrderDate, TotalAmount, Status)
- **Shipments** (ShipmentID, OrderID, ShippedDate, ExpectedDelivery, ActualDelivery)
- **Feedback** (FeedbackID, CustomerID, OrderID, Rating, Comment)

Practice Experience

- Today's focus was on **logistics and delivery analytics**, along with customer feedback insights.
- Many questions were **confusing at first**, mainly because they required handling **date differences (days count)** and combining shipment timelines with feedback ratings.
- Once I carefully broke queries into steps, the logic became clearer.

Key Learnings from Queries

1. **Delivery Delays:** Measured average delivery delays using DATEDIFF().
2. **Top Spenders:** Identified top 3 customers by spending across completed orders.
3. **Late Deliveries by Country:** Calculated percentages with CASE WHEN.
4. **Undelivered Orders:** Isolated shipments with ActualDelivery IS NULL.
5. **Repeat Feedback:** Found customers providing multiple reviews.
6. **Running Spend Totals:** Used SUM() OVER to build customer-wise spend progress.
7. **Most Frequent Delay:** Built a mode function for delivery delays.
8. **Rating vs Delays:** Correlated customer ratings with shipment delay averages.
9. **Quick Reorders:** Detected consecutive orders within 30 days using LAG().
10. **Maximum Delay Case:** Highlighted the customer/order with the largest delay.
11. **Bonus – At-Risk Customers:** Combined multiple signals (≥2 late deliveries + avg rating ≤ 3) to classify risk levels (High/Medium/Low).

Insights

- **Alice** stood out as a high spender with multiple orders and consistent positive feedback.
- **Germany and India** showed higher late delivery percentages compared to other countries.
- Ratings clearly dropped as delivery delays increased, reinforcing the **strong link between operations and customer satisfaction**.
- At-risk customers (e.g., Eva) highlighted how late deliveries + low ratings can be used for **early churn detection**.

Skills Reinforced

- Advanced date handling with DATEDIFF
- Window functions (LAG, SUM OVER)
- Multi-condition risk modeling with CASE WHEN
- Building correlation queries (ratings vs delays)
- Thinking in terms of **business metrics**: churn risk, delivery performance, and satisfaction

Personal Note

Today's practice was **tricky because of day calculations**, but it helped me strengthen my approach to time-based queries.

I learned that **delivery timelines and customer ratings are deeply connected**, and SQL can surface these patterns effectively.

Breaking problems step by step — especially on confusing queries — gave me confidence to handle even more complex scenarios.