

# Business Analyst Test Case

## SQL Task:

Please answer the following two SQL tasks under the conditions below. The table details you need are shown in the following image:

Orders	Vouchers	Categories	OrdersDetail	Products
<div><div>id</div><div>customer_id</div><div>created_at</div><div>status_id</div><div>source</div><div>vendor_id</div><div>payment_status</div></div> <div><div>bigint</div><div>bigint</div><div>datetime</div><div>tinyint</div><div>varchar</div><div>bigint</div><div>varchar</div></div>	<div><div>id</div><div>amount</div><div>order_id</div></div> <div><div>bigint</div><div>bigint</div><div>bigint</div></div>	<div><div>id</div><div>title</div><div>vendor_id</div><div>position</div><div>active</div><div>root_id</div><div>parent_id</div><div>lft</div><div>rgt</div><div>level</div></div> <div><div>bigint</div><div>bigint</div><div>bigint</div><div>smallint</div><div>tinyint(1)</div><div>int</div><div>int</div><div>int</div><div>int</div><div>int</div></div>	<div><div>id</div><div>order_id</div><div>product_id</div><div>price</div><div>discount</div></div> <div><div>bigint</div><div>bigint</div><div>bigint</div><div>bigint</div><div>bigint</div></div>	<div><div>id</div><div>vendor_id</div><div>title</div><div>category_id</div></div> <div><div>bigint</div><div>bigint</div><div>bigint</div><div>bigint</div></div>

- An order is valid when:
  - Orders.status\_id is not 128
  - Orders.source does not contain "FOOD"
  - Orders.payment\_status is "completed" or "manually\_completed"
- Each product belongs to a leaf category (level 3), which rolls up to a middle category (level 2), which in turn rolls up to a root category (level 1).
- Some orders may have no discounts and/or no vouchers.
- Since you do not have access to our database, please write your queries in plain text format.
- Provide answers with different variations if possible, at least one without CTEs or window functions.
- Indicate which variation is the most optimized and why.

1. Write a query to calculate the monthly 30 days retention of customers (how many of the customers returned in 30 days or less) for the past 12 months. Expected output columns:

Year	Month	Count of Users	Count of Returned Users	Retention (%)
2025	January	...		

2. Write a query that returns, for each vendor and category, the monthly totals for the last 12 months. Expected output columns:

year_month	orders_count	sales_value	discount_value	voucher_value	category_level_1	category_level_2	vendor_id
202505	...						

---

## Visualization and Analytics Task:

You have been given a dataset containing transactional and customer data (see attached file). Explore and visualize the data to deliver the best possible business insights. There are no strict instructions—use your judgment to clean, analyze, and present the data in a way that helps decision-makers understand key patterns and opportunities.

- Bonus points if you can automate any part of your process using R or Python.

## Deliverables:

- **Report:** Describe your methodology for solving each task, specifying all steps and their purposes.
- **Queries & Codes:** Submit well-documented scripts and queries. Ensure the code is clean, readable, and includes comments explaining your logic.
- **Visualizations:** Including plots or charts that support your findings is a plus. Create interactive dashboards that effectively visualize your insights.
- **Use of AI:** If you use any AI tools or assistance (e.g., ChatGPT, code generators), please disclose where, how, and why you used them.