Data for the questions below:
PostgreSQL
Host: pg-technical-assessment-v2.cvqh5ju3daog.us-east-1.rds.amazonaws.com
Port: 5432
Database: postgres
Username: postgres
Password: xjcFvv8iy2WRR3H

SQL Challenges:

Given the following Tables in the database above:

1. ***orders*:**
   This table represents individual orders from an eCommerce store
   Key fields:
   a. id | int
   b. created_at | timestamp
   c. customer_id | int
2. ***orders_line_items:***
   This table contains every item in each order - often orders will contain multiple line items
   Key fields:
   a. order_id | int
   b. sku | varchar
   c. product | varchar
   d. product_type | varchar (category of the product)
   e. quantity | numeric
   f. price | numeric

You can assume that id and order_id are unique order identifiers and customer_id is a unique customer identifier. Please join ***orders*** and ***orders_line_items*** table using source_id and order_id.

1. Write a query that would find the total number of customers who have placed at least 3 product orders. Where a product order is any that has product_amount_post_refund > 0.
2. Write a query that would find, of that group, how many ordered shampoo on their *first*

order. You can assume that any product with the text “Shampoo” in the name, is a shampoo.

3. Which products are most correlated with improved retention? (highest likelihood of of customer making another purchase)
4. Show monthly retention of different monthly acquisition cohorts
    a. Assume the first order for each customer is the acquisition cohort (month) of the customer
    b. Show how likely users (by cohort) are to buy in their 2nd month, 3rd month etc.
5. In your own words (no data-pull necessary), how would you figure out how to price each of our products, in order to maximize company profits?