Part 1: CMS Data Exercise

Here at Clipboard Health, we provide staffing to long-term care facilities. The Centers for Medicaid and Medicare Services publish a quarterly report containing daily staffing data for all registered nursing homes in the U.S. This dataset is called Payroll Based Journal (PBJ) Daily Nurse Staffing and can be found here. You can review the data dictionary to better understand the data that's available, but a few notes pertaining to this exercise:

- Nursing homes are staffed by both employees and contractors. Employees work for the nursing home full time, whereas contractors work for the nursing home on a temporary basis. Most nursing homes are staffed using some combination of full time employees and contractors, but the proportions within each nursing home can vary substantially. All workers from Clipboard Health (and our competitors) are classified as contractors.
- Clipboard Health is a nationwide staffing platform, operating in all 50 states.
- There are a variety of other datasets that can be joined to this dataset, all found under the nursing home data section on CMS' website.
- As mentioned previously, the PBJ data is separated by quarter. For this exercise, please focus on the most recent quarter available (2024Q2).

Can you please use the PBJ data and any other CMS data that you see fit to make a few recommendations to the Clipboard Health sales leadership team? There are no right or wrong answers but a few tips:

- Be sure to include the supporting data (charts, tables, etc.) for any recommendation you make. The purpose of this exercise is largely to see how well you can quickly learn and analyze new datasets, as well as communicate the learnings and recommendations from those analyses.
- Feel free to use any analytical tools that you have at your disposal.
- We do a lot of writing at Clipboard Health (we believe that writing is thinking), so each recommendation should include some writing around how you analyzed the data, what you uncovered, and the specifics of what you'd recommend the sales team do based on your findings.
- There is no minimum number of recommendations we'd like you to make, but please include no more than 5.

Part 2: SQL Test

Instructions: Please write SQL queries for each of the following questions. You may assume that all tables follow typical database conventions unless otherwise specified.

Tables:

Assume you have the following tables in your database:

1. **Sales**
    - `sales_id` (INT)

    - `customer_id` (INT)
    - `product_id` (INT)
    - `sale_date` (DATE)
    - `quantity` (INT)
    - `total_amount` (DECIMAL)
2. **Customers**
    - `customer_id` (INT)
    - `customer_name` (VARCHAR)
    - `sales_region` (VARCHAR)
    - `sign_up_date` (DATE)
3. **Products**
    - `product_id` (INT)
    - `product_name` (VARCHAR)
    - `category` (VARCHAR)
    - `price` (DECIMAL)

Questions:

1. Write a query to return the `customer_name`, `product_name`, and `total_amount` for each sale in the last 30 days.
2. Write a query to find the total revenue generated by each product category in the last year. The output should include the product category and the total revenue for that category.
3. Write a query to return all customers who made purchases in 2023 and are located in the "West" region.
4. Write a query to display the total number of sales, total quantity sold, and total revenue for each customer. The result should include the `customer_name`, total sales, total quantity, and total revenue.
5. Write a query to find the top 3 customers (by total revenue) in the year 2023.
6. Write a query to rank products by their total sales quantity in 2023. The result should include the `product_name`, total quantity sold, and rank.
7. Write a query that categorizes customers into "New" (if they signed up in the last 6 months) or "Existing" based on their `sign_up_date`. Include the `customer_name`, `region`, and category in the result.
8. Write a query to return the month and year along with the total sales for each month for the last 12 months.
9. Write a query to return the product categories that generated more than $50,000 in revenue during the last 6 months.
10. Write a query to check for any sales where the `total_amount` doesn't match the expected value (i.e., `quantity` * `price`).