

## SQL Queries

### 1) What are the top 5 brands by receipts scanned among users 21 and over?

```
SELECT brand,
       COUNT(receipt_id) AS receipt_count
FROM   fetchrewards.transactions
JOIN   fetchrewards.products
       ON products.barcode = transactions.barcode -- Matching records based on barcode
JOIN   fetchrewards.users
       ON users.ID = transactions.USER_ID -- Matching records based on userID
WHERE  EXTRACT(YEAR FROM AGE(BIRTH_DATE)) >= 21 -- Filters users with age 21 or more
       AND brand IS NOT NULL -- Ensures brand is not null
GROUP BY
       brand -- Groups the results by brand
ORDER BY
       receipt_count DESC -- Sorts the result in descending order
LIMIT 5;
```

#### Output :

BRAND	RECEIPT_COUNT
DOVE	21
TRESEMMÉ	19
CHEX MIX	19
SUAVE	19
ST. IVES	18

### 2) Top 5 brands by sales among users that have had their account for at least six months

```
SELECT
  brand,
  SUM(final_sale) AS total_sales
FROM
  transactions
JOIN
  products ON products.barcode = transactions.barcode
JOIN
  users ON users.ID = transactions.user_id
WHERE
  users.created_date <= CURRENT_DATE - INTERVAL '6 months'
  AND final_sale IS NOT NULL
  AND brand IS NOT NULL
GROUP BY
  brand
ORDER BY
  total_sales DESC
LIMIT 5;
```

**OutPut :**

BRAND	TOTAL_SALES
EQUATE	6429.88
CVS	6107.09
DOVE	3772.68
PEPSI	3745.84
NATURE VALLEY	3324.12

**3 .Percentage of sales in the Health & Wellness category by generation**

```
WITH generation_sales AS (  
  SELECT  
    CASE  
      WHEN EXTRACT(YEAR FROM AGE(users.birth_date)) BETWEEN 18 AND 24  
    THEN 'Gen Z'  
      WHEN EXTRACT(YEAR FROM AGE(users.birth_date)) BETWEEN 25 AND 40  
    THEN 'Millennials'  
      WHEN EXTRACT(YEAR FROM AGE(users.birth_date)) BETWEEN 41 AND 56  
    THEN 'Gen X'  
      WHEN EXTRACT(YEAR FROM AGE(users.birth_date)) >= 57 THEN 'Baby Boomers'  
      ELSE 'Unknown'  
    END AS generation,  
    SUM(final_sale) AS total_sales  
  FROM  
    transactions  
  JOIN  
    products ON products.barcode = transactions.barcode  
  JOIN  
    users ON users.ID = transactions.user_id  
  WHERE  
    products.category = 'Health & Wellness'  
    AND final_sale IS NOT NULL  
  GROUP BY  
    generation  
)  
total_sales AS (  

```

```

SELECT
    SUM(final_sale) AS grand_total
FROM
    transactions
JOIN
    products ON products.barcode = transactions.barcode
WHERE
    products.category = 'Health & Wellness'
)
SELECT
    generation,
    total_sales,
    ROUND((total_sales / (SELECT grand_total FROM total_sales)) * 100, 2) AS percentage
FROM
    generation_sales;

```

### **OutPut :**

GENERATION	CATEGORY_SALES	PERCENTAGE
Baby Boomers	0.0	0.0%
Gen X	0.0	0.0%
Millennials	0.0	0.0%

## **1. Who are Fetch's power users?**

**Assumption: Power users are defined as users who have scanned the highest number of receipts.**

```

SELECT
    USER_ID,
    COUNT(RECEIPT_ID) AS RECEIPT_COUNT
FROM
    TRANSACTION_TAKEHOME
GROUP BY
    USER_ID
ORDER BY
    RECEIPT_COUNT DESC
LIMIT 10;

```

**Output:**

USER_ID	RECEIPT_COUNT	TOTAL_SPENDING
643059f0838dd2651fb27f50	4	75.99
62ffec490d9dbaff18c0a999	6	52.28
5f4c9055e81e6f162e3f6fa8	2	37.96
5d191765c8b1ba28e74e8463	2	34.96
6351760a3a4a3534d9393ecd	4	27.74
64dd9170516348066e7c4006	4	26.52
62c09104baa38d1a1f6c260e	6	20.28
61a58ac49c135b462ccddd1c	6	19.92
6661ed1e7c0469953bfc76c4	4	18.60
5b441360be53340f289b0795	4	18.32

**2. Which is the leading brand in the Dips & Salsa category?**

**Assumption:** The leading brand is the one with the highest total sales in the "Dips & Salsa" category.

```
SELECT
  P. BRAND,
  SUM (T. FINAL_SALE) AS TOTAL_SALES
FROM
  TRANSACTION_TAKEHOME T
JOIN
  PRODUCTS_TAKEHOME P
ON
  T. BARCODE = P. BARCODE
WHERE
  P. CATEGORY_2 = 'Dips & Salsa'
GROUP BY
  P. BRAND
ORDER BY
  TOTAL_SALES DESC
LIMIT 1.
```

**OutPut :**

BRAND	TOTAL_SALES
MARKETSIDE	165,280.06

### 3. At what percent has Fetch grown year over year?

**Assumption:** Growth is based on the count of new users created year over year.

```
WITH UserCounts AS (
  SELECT
    YEAR(CREATED_DATE) AS Year,
    COUNT(ID) AS UserCount
  FROM
    USER_TAKEHOME
  GROUP BY
    YEAR(CREATED_DATE)
),
Growth AS (
  SELECT
    Year,
    UserCount,
    LAG(UserCount) OVER (ORDER BY Year) AS PreviousYearCount,
```

```

        ((UserCount - LAG(UserCount) OVER (ORDER BY Year)) / LAG(UserCount) OVER
        (ORDER BY Year)) * 100 AS GrowthRate
    FROM
        UserCounts
)
SELECT
    Year,
    UserCount,
    PreviousYearCount,
    GrowthRate
FROM
    Growth
ORDER BY
    Year.

```

## **Output:**

Calculate Fetch's year-over-year (YoY) growth rate based on the total number of transactions each year.

1. Extracted the year from the PURCHASE\_DATE field in the transactions table and count the total number of transactions (TOTAL\_TRANSACTIONS) for each year.
2. Perform a self-join on the yearly transactions table to compare each year's transactions with the previous year (TRANSACTION\_YEAR = PREVIOUS\_YEAR + 1).
3. Calculate the YoY growth percentage using the formula:  

$$\text{YoY Growth} = \frac{\text{Current Year Transactions} - \text{Previous Year Transactions}}{\text{Previous Year Transactions}} \times 100$$

$$\text{YoY Growth} = \frac{\text{Current Year Transactions} - \text{Previous Year Transactions}}{\text{Previous Year Transactions}} \times 100$$
4. Filter out years without data for the previous year (yt2.TOTAL\_TRANSACTIONS IS NOT NULL).

Note:

The query returned no results because there were no yearly transaction records with sufficient data to calculate year-over-year growth. This is due to missing or incomplete transaction dates in the dataset.