Продажи магазина напитков

-- BEGINNER QUESTIONS --

1) Наиболее покупаемые продукты (топ 10)

```
SELECT
    Product,
    COUNT(*) AS Num_by_products
FROM beverage
GROUP BY Product
ORDER BY num_by_products DESC
LIMIT 10;
```

	product text	num_by_products bigint
1	Hohes C Orange	337114
2	Granini Apple	319497
3	Cranberry Juice	319328
4	Passion Fruit Juice	318923
5	Tomato Juice	318773
6	Rauch Multivitam...	318538
7	Mango Juice	317764
8	San Pellegrino	256068
9	Vittel	255790
10	Volvic Touch	248967

2) Количество покупок по категориям

```
SELECT
    Category,
    COUNT(*) AS Num_by_categories
FROM beverage
GROUP BY Category
ORDER BY num_by_categories DESC;
```

	category text	num_by_categories bigint
1	Alcoholic Beverages	2251625
2	Water	2250217
3	Juices	2249937
4	Soft Drinks	2248131

3) Наибольшее количество заказов по регионам (топ 5)

```
SELECT
    Region,
    COUNT(DISTINCT Order_ID) AS Num_by_regions
FROM beverage
GROUP BY Region
ORDER BY num_by_regions DESC
LIMIT 5;
```

	region text	num_by_regions bigint
1	Hamburg	201457
2	Rheinland-Pfalz	192687
3	Niedersachsen	192298
4	Saarland	191273
5	Sachsen	190813

4) Количество заказов по типам клиентов

```
SELECT
    Customer_Type,
    COUNT(DISTINCT Order_ID) AS Num_of_customer_type
FROM beverage
GROUP BY Customer_Type;
```

	customer_type text	num_of_customer_type bigint
1	B2B	1068808
2	B2C	1931192

-- INTERMEDIATE QUESTIONS --

1) Затраты и количество купленных товаров по регионам

```
SELECT
    Region,
    COUNT(Quantity) AS Quantity,
    ROUND(SUM(Total_Price::numeric), 2) AS Total_Price
FROM beverage
GROUP BY Region
ORDER BY Total_Price DESC;
```

	region text	quantity bigint	total_price numeric
1	Hamburg	604054	82470771.65
2	Hessen	547225	78400110.48
3	Saarland	573596	78390587.50
4	Rheinland-Pfalz	577967	75838677.38
5	Mecklenburg-Vorpommern	544936	75517247.15
6	Thüringen	562554	75324865.73
7	Berlin	547405	74567927.13

2) Затраты и количество купленных товаров по категориям

```
SELECT
    Category,
    COUNT(Quantity) AS Quantity,
    ROUND(SUM(Total_Price::numeric), 2) AS Total_Price
FROM beverage
GROUP BY Category
ORDER BY Total_Price DESC;
```

	category text	quantity bigint	total_price numeric
1	Alcoholic Beverages	2251625	911797918.77
2	Juices	2249937	133167848.64
3	Soft Drinks	2248131	82802542.53
4	Water	2250217	48912852.46

3) Продажи по дате (по дням)

```
SELECT
    Order_Date,
    ROUND(SUM(Total_Price::numeric), 2) AS Sales
FROM beverage
GROUP BY Order_Date
ORDER BY Order_Date;
```

	order_date date	sales numeric
1	2021-01-01	987429.55
2	2021-01-02	1067312.00
3	2021-01-03	964076.39
4	2021-01-04	1033108.48
5	2021-01-05	1083223.79
6	2021-01-06	923362.86
7	2021-01-07	1027440.71

4) Продажи по дате (по месяцам)

```
SELECT DATE_PART('year', Order_Date) AS date_year,
    DATE_PART('month', Order_Date) AS date_month,
    ROUND(SUM(Total_Price::numeric), 2) AS Sales
FROM beverage
GROUP BY date_month, date_year
ORDER BY date_year, date_month;
```

	date_year double precision	date_month double precision	sales numeric
1	2021	1	32425429.01
2	2021	2	29110169.08
3	2021	3	32957436.51
4	2021	4	31142058.46
5	2021	5	31855490.10
6	2021	6	30993287.15
7	2021	7	32112059.31

5) Прибыль по дате (по годам)

```
SELECT
    DATE_PART('year', Order_Date) AS Sales_by_Year,
    ROUND(SUM(Total_Price::numeric), 2) AS Sales
FROM beverage
GROUP BY Sales_by_Year
ORDER BY Sales_by_Year
LIMIT 10;
```

	sales_by_year double precision	sales numeric
1	2021	380046647.03
2	2022	394859779.78
3	2023	401774735.59

-- МЕТРИЧЕСКИЕ ПОКАЗАТЕЛИ --

1) DAU: число покупателей в день (за последние полгода)

```
WITH d AS(
    SELECT
        Order_Date,
        COUNT(DISTINCT Customer_ID) AS cust_day
    FROM beverage
    WHERE Order_Date >= '2023-07-01'
    GROUP BY Order_Date
)

SELECT
     ROUND(AVG(cust_day)::numeric, 2) AS DAU
FROM d;
```

	dau numeric
1	2401.24

2) WAU: число покупателей в неделю (за последние полгода)

```
WITH w AS(
    SELECT
        DATE_PART('year', Order_Date) AS year_num,
        DATE_PART('week', Order_Date) AS week_num,
        COUNT(DISTINCT Customer_ID) AS cust_week
    FROM beverage
    WHERE Order_Date >= '2023-07-01'
    GROUP BY year_num, week_num
)

SELECT
     ROUND(AVG(cust_week)::numeric, 2) AS WAU
FROM w;
```

	wau numeric
1	8361.48

3) MAU: число покупателей в месяц (за последние полгода)

```
WITH m AS(
```

```
    SELECT
        DATE_PART('year', Order_Date) AS year_num,
        DATE_PART('month', Order_Date) AS month_num,
        COUNT(DISTINCT Customer_ID) AS cust_month
    FROM beverage
    WHERE Order_Date >= '2023-07-01'
    GROUP BY year_num, month_num
)

SELECT
     ROUND(AVG(cust_month)::numeric, 2) AS MAU
FROM m;
```

	mau numeric
1	9998.00

4) Sticky_factor: степень лояльности и частота взаимодействия клиентов (за последние полгода)

```
WITH d AS(
    SELECT
        Order_Date,
        COUNT(DISTINCT Customer_ID) AS cust_day
    FROM beverage
    WHERE Order_Date >= '2023-07-01'
    GROUP BY Order_Date
),
m AS(
    SELECT
        DATE_PART('year', Order_Date) AS year_num,
        DATE_PART('month', Order_Date) AS month_num,
        COUNT(DISTINCT Customer_ID) AS cust_month
    FROM beverage
    WHERE Order_Date >= '2023-07-01'
    GROUP BY year_num, month_num
)

SELECT
```

```
    ROUND(AVG(cust_day) / AVG(cust_month) * 100, 2) AS sticky_factor
FROM d, m;
```

	sticky_factor numeric
1	24.02

5) Lifetime

```
WITH f as( -- День первой покупки
    SELECT DISTINCT Customer_ID,
        Order_Date,
        FIRST_VALUE(Order_Date) OVER (PARTITION BY Customer_ID
ORDER BY Order_Date) AS First_Order
    FROM beverage
    WHERE Order_Date >= '2023-07-01'
    ORDER BY Order_Date
),
df as ( -- Разница между датой текущей покупки и датой первой покупки
    SELECT
      Order_Date - First_Order AS diff,
      COUNT(DISTINCT Customer_ID) AS count_cust
    FROM f
    GROUP BY diff
),
ret as( -- Доля оставшихся клиентов
    SELECT
        diff,
      ROUND(count_cust * 1.0 / FIRST_VALUE(count_cust) OVER (ORDER BY
diff), 4) AS retention
    FROM df
)

SELECT
  SUM(retention) AS lifetime
FROM ret;
```

	lifetime numeric
1	43.9427