Credit Card Transactions

Data Analysis Case Study

By

GrowData Skills

DATA UNDERSTANDING

There is only one table

credit_card_transcations: contains information about credit card usage

This is just as small part of table

transaction_id	city	transaction_date	card_type	exp_type	gender	amount
1	Delhi	29-10-2014	Gold	Bills	F	82475
2	Greater Mumbai	22-08-2014	Platinum	Bills	F	32555
3	Bengaluru	27-08-2014	Silver	Bills	F	101738
4	Greater Mumbai	12-04-2014	Signature	Bills	F	123424
5	Bengaluru	05-05-2015	Gold	Bills	F	171574
6	Delhi	08-09-2014	Silver	Bills	F	100036
7	Delhi	24-02-2015	Gold	Bills	F	143250
8	Greater Mumbai	26-06-2014	Platinum	Bills	F	150980
9	Delhi	28-03-2014	Silver	Bills	F	192247
10	Delhi	01-09-2014	Platinum	Bills	F	67932
11	Delhi	22-06-2014	Platinum	Bills	F	280061
12	Greater Mumbai	07-12-2013	Signature	Bills	F	278036
13	Greater Mumbai	07-08-2014	Gold	Bills	F	19226
14	Delhi	27-04-2014	Signature	Bills	F	254359
15	Greater Mumbai	15-08-2014	Signature	Bills	F	302834
16	Greater Mumbai	28-11-2014	Platinum	Bills	F	647116
17	Greater Mumbai	14-06-2014	Signature	Bills	F	421878
18	Greater Mumbai	30-03-2015	Gold	Bills	F	986379

- Write a query to print top 5 cities with highest spends and their percentage contribution of total credit card spends

```sql
SELECT
    city,
    SUM(amount) as total_spend,
    ROUND (
      100.0*SUM(amount)/
      (SELECT SUM(amount)
       FROM credit_card_transcations),2)
        AS pect_of_total_spend
FROM credit_card_transcations
GROUP BY city
ORDER BY total_spend DESC
LIMIT 5;
```

Result

	city	total_spend	pect_of_total_spend
▶	Greater Mumbai	576751476	14.15
	Bengaluru	572326739	14.05
	Ahmedabad	567794310	13.93
	Delhi	556929212	13.67
	Kolkata	115466943	2.83

- Write a query to print highest spend month and amount spent in that month for each card type

```sql
WITH cte_1 AS (
    SELECT
        card_type,
        MONTH(transaction_date) AS transaction_month,
        YEAR(transaction_date) AS transaction_year,
        SUM(amount) AS monthly_spend
    FROM credit_card_transcations
    GROUP BY card_type,
        YEAR(transaction_date),
        MONTH(transaction_date)
    ORDER BY monthly_spend DESC
)
,cte_2 AS (
    SELECT
    *,
    RANK() OVER(PARTITION BY card_type
                ORDER BY monthly_spend desc) as rk
    FROM cte_1
)
SELECT
    card_type,
    transaction_month,
    transaction_year,
    monthly_spend
FROM cte_2
WHERE rk = 1;
```

Result

	card_type	transaction_month	transaction_year	monthly_spend
▶	Gold	1	2015	55455064
	Platinum	8	2014	57936507
	Signature	12	2013	58799522
	Silver	3	2015	59723549

- Write a query to print the transaction details(all columns from the table) for each card type when--it reaches a cumulative of 1000000 total spends

```sql
WITH cte_1 AS (
    SELECT
        *,
        SUM(amount)
        OVER(PARTITION BY card_type ORDER BY transaction_date, transaction_id)
            AS cumm_spend
    FROM credit_card_transcations
    ORDER BY card_type, cumm_spend
)
,cte_2 AS (
    SELECT
        *,
        ROW_NUMBER()
        OVER(PARTITION BY card_type ORDER BY cumm_spend) AS rk
    FROM cte_1
    WHERE cumm_spend >= 1000000
)
SELECT
    transaction_id,
    city,
    transaction_date,
    card_type,
    exp_type,
    gender,
    amount FROM cte_2
WHERE rk =1
```

Result

transaction_id	city	transaction_date	card_type	exp_type	gender	amount
1522	Delhi	2013-10-04	Gold	Food	M	281924
191	Ahmedabad	2013-10-05	Platinum	Bills	F	612572
73	Delhi	2013-10-04	Signature	Bills	F	550782
7565	Bengaluru	2013-10-04	Silver	Food	F	205179

- Write a query to find city which had lowest percentage spend for gold card type

```sql
WITH cte_1 AS (
    SELECT
        city,
        card_type,
        SUM(amount) AS card_wise_spend
    FROM credit_card_transcations
    GROUP BY city, card_type
)
,cte_2 AS (
    SELECT
        city,
        SUM(amount) AS city_wise_spend
    FROM credit_card_transcations
    GROUP BY city
)
SELECT
    cte_1.city AS city
FROM cte_1
INNER JOIN cte_2
    ON cte_1.city = cte_2.city
WHERE card_type = "Gold"
ORDER BY card_wise_spend/city_wise_spend
LIMIT 1;
```

Result

	city
▶	Dhamtari

- Write a query to print 3 columns: city, highest_expense_type, lowest_expense_type

```sql
WITH cte_1 AS (
    SELECT
        city,
        exp_type,
        SUM(amount) AS total_spend
    FROM credit_card_transcations
    GROUP BY city, exp_type
)
,cte_2 AS (
    SELECT
        *,
        RANK() OVER(PARTITION BY city ORDER BY total_spend DESC)
            AS highest_spend_rk,
        RANK() OVER(PARTITION BY city ORDER BY total_spend ASC)
            AS lowest_spend_rk
    FROM cte_1
)
SELECT
    city,
    MAX(case when highest_spend_rk = 1 THEN exp_type END )
        AS "highest_expense_type",
    MAX(case when lowest_spend_rk = 1 THEN exp_type END )
        AS "lowest_expense_type"
FROM cte_2
GROUP BY city;
```

Result

city	highest_expense_type	lowest_expense_type
Achalpur	Grocery	Entertainment
Adilabad	Bills	Food
Adityapur	Food	Grocery
Adoni	Bills	Entertainment
Adoor	Fuel	Bills
Afzalpur	Fuel	Food
Agartala	Grocery	Food
Agra	Bills	Grocery
Ahmedabad	Bills	Grocery
Ahmednagar	Fuel	Grocery
Aizawl	Food	Grocery

- Write a query to find percentage contribution of spends by females for each expense type

```sql
SELECT
    exp_type,
    ROUND (
       100.0*
       SUM (CASE WHEN gender = "F" THEN amount
END       )/SUM(amount),2)
         AS contribution_by_females
FROM credit_card_transcations
GROUP BY exp_type;
```

Result

	exp_type	contribution_by_females
▶	Bills	63.95
	Food	54.91
	Entertainment	49.37
	Grocery	50.91
	Fuel	49.71
	Travel	51.13

- Which card and expense type combination saw highest month over month growth in Jan-2014

```sql
WITH cte_1 AS (
    SELECT
        card_type,
        exp_type,
        Month(transaction_date) AS trans_month,
        YEAR(transaction_date) AS trans_year,
        SUM(amount) AS spend
    FROM credit_card_transcations
        WHERE ( Month(transaction_date) = 1
                AND YEAR(transaction_date) = 2014 )
            OR ( Month(transaction_date) = 12
                 AND YEAR(transaction_date) = 2013 )
    GROUP BY card_type, exp_type,
                 Month(transaction_date),
                 YEAR(transaction_date)
)
,cte_2 AS (
    SELECT *,
    LAG(spend,1) OVER(PARTITION BY card_type, exp_type
                      ORDER BY trans_year )
        AS prev_month_spend
FROM cte_1
)
SELECT
    *,
    ROUND(100.0*(spend - prev_month_spend)/prev_month_spend,2)
        AS mom_growth_pct
FROM cte_2
ORDER BY mom_growth_pct desc
LIMIT 1;
```

Result

	card_type	exp_type	trans_month	trans_year	spend	prev_month_spend	mom_growth_pct
▶	Gold	Travel	1	2014	2092554	1113534	87.92

- During weekends which city has highest total spend to total no of transactions ratio

```sql
SELECT
    city,
    SUM(amount)/count(transaction_id)
        AS total_spend_to_total_trans_ratio
FROM credit_card_transcations
WHERE DAYNAME(transaction_date) IN ("Saturday","Sunday")
GROUP BY city
ORDER BY total_spend_to_total_trans_ratio DESC
LIMIT 1;
```

Result

	city	total_spend_to_total_trans_ratio
▶	Sonepur	299905.0000

- Which city took least number of days to reach its 500th transaction after the first transaction in that city

```sql
WITH cte_1 AS (
    SELECT
        *,
        ROW_NUMBER()
        OVER(PARTITION BY city ORDER BY transaction_date ASC)
            AS trans_rank
    FROM credit_card_transcations
)
SELECT
    city,
    DATEDIFF(MAX(transaction_date), MIN(transaction_date))
        AS days_diff
FROM cte_1
WHERE trans_rank IN (1,500)
GROUP BY city
HAVING COUNT(city) = 2
ORDER BY days_diff
LIMIT 1;
```

Result

	city	days_diff
▶	Bengaluru	81