Credit Card Transaction Analysis

SQL- CASE STUDY

Tools used- Excel, MySQL
Dataset source- Ankit Bansal

Sweta Sah

Credit_card_transaction_analysis

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

```
SELECT city, SUM(amount) AS total_amount, SUM(amount)*100 / (SELECT SUM(amount) FROM credit_card_transaction_analysis) AS percentage_contribution

FROM credit_card_transaction_analysis

GROUP BY city

ORDER BY total_amount DESC

LIMIT 5;
```

city	total_amount	percentage_contribution
Greater Mumbai	576751476	14.1540
Bengaluru	572326739	14.0454
Ahmedabad	567794310	13.9342
Delhi	556929212	13.6675
Kolkata	115466943	2.8337

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

```
-- USING CTE

WITH c AS (

SELECT card_type, year, month, SUM(amount) AS amount

FROM credit_card_transaction_analysis

GROUP BY Card_type, year, month

ORDER BY amount DESC),

t AS (

SELECT *, dense_rank() OVER(PARTITION BY card_type ORDER BY amount DESC) AS r

FROM c)

SELECT * FROM t

WHERE r=1;
```

card_type	year	month	amount	r
Gold	2015	1	55455064	1
Platinum	2014	8	57936507	1
Signature	2013	12	58799522	1
Silver	2015	3	59723549	1

-- USING SUBQUERY

```
SELECT t.card_type, t.year, t.month, t.amount
FROM
(SELECT c.*, dense_rank() OVER(PARTITION BY c.card_type ORDER BY c.amount DESC) AS r
FROM(
SELECT card_type, year, month, SUM(amount) AS amount
FROM credit_card_transaction_analysis
GROUP BY Card_type, year, month) AS c) AS t
WHERE t.r=1;
```

card_type	year	month	amount
Gold	2015	1	55455064
Platinum	2014	8	57936507
Signature	2013	12	58799522
Silver	2015	3	59723549

3)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 (We should have 4 rows in the o/p one for each card type)

```
SELECT t.*
FROM
(SELECT c.*, DENSE_RANK() OVER(PARTITION BY c.card_type ORDER BY c.cs DESC) AS r
FROM
(SELECT *, SUM(amount) OVER(PARTITION BY card_type ORDER BY year, month ROWS BETWEEN
UNBOUNDED PRECEDING AND CURRENT ROW) AS cs
FROM credit_card_transaction_analysis) c
WHERE c.cs <= 1000000) t
WHERE t.r= 1;
```

Index	City	date	Month	Year	Card_type	Exp_type	Gender	Amount	cs	r
23296	Jaipur	2013-10-12	10	2013	Gold	Bills	M	171529	886361	1
13027	Ahmedabad	2013-10-28	10	2013	Platinum	Entertainment	M	97720	879850	1
19171	Periyasemur	2013-10-31	10	2013	Signature	Fuel	M	84988	819542	1
23766	Pune	2013-10-19	10	2013	Silver	Bills	M	222712	892296	1

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

```
SELECT city

FROM

(SELECT city, SUM(amount) AS spend, SUM(amount)*100 / (SELECT SUM(amount) FROM
credit_card_transaction_analysis) AS percentage_spend

FROM credit_card_transaction_analysis

WHERE Card_type= "Gold"

GROUP BY city

ORDER BY percentage_spend ASC

LIMIT 1) t;
```

	city
▶	Dhamtari

5)Write a query to print 3 columns: city, highest_expense_type , lowest_expense_type (example format : Delhi , Bills, Fuel)

```
WITH t AS (SELECT c.*,

row_number() OVER(PARTITION BY c.city ORDER BY c.amount) AS ar,

row_number() OVER(PARTITION BY c.city ORDER BY c.amount DESC) AS dr

FROM

(SELECT city, exp_type, SUM(amount) AS amount

FROM credit_card_transaction_analysis

GROUP BY city, Exp_type

ORDER BY city, amount) AS c)


SELECT t.city,

MAX(CASE WHEN t.ar=1 THEN t.exp_type END) AS lowest_expense_type,

MAX(CASE WHEN t.dr=1 THEN t.exp_type END) AS highest_expense_type

FROM t

GROUP BY t.city;
```

city	lowest_expense_type	highest_expense_type
Achalpur	Entertainment	Grocery
Adilabad	Food	Bills
Adityapur	Grocery	Food
Adoni	Entertainment	Bills
Adoor	Bills	Fuel
Afzalpur	Food	Fuel

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

```
WITH t AS (SELECT exp_type, SUM(CASE WHEN gender='F' THEN amount END) AS f_spend, SUM(amount) AS total_spend

FROM credit_card_transaction_analysis

GROUP BY exp_type

ORDER BY total_spend)


SELECT t.exp_type, ROUND((t.f_spend/t.total_spend)*100, 2) AS f_percentage_contribution

FROM t;
```

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

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

```
WITH t AS (SELECT *,

LAG(amount) OVER(PARTITION BY card_type, exp_type ORDER BY year, amount) AS previous_amount

FROM

(SELECT card_type, exp_type, year, SUM(amount) AS amount

FROM credit_card_transaction_analysis

WHERE (month=1 AND YEAR= 2014) OR (month=12 AND YEAR=2013)

GROUP BY card_type, exp_type, year

ORDER BY year, amount) c)


SELECT t.card_type, t.exp_type, ROUND((t.amount-t.previous_amount)*100/t.previous_amount, 1) AS mom_growth
```

```
FROM t
ORDER BY mom_growth DESC
LIMIT 1;
```

	card_type	exp_type	mom_growth
▶	Gold	Travel	87.9

8)Which city has highest total spend to total no of transactions ratio during weekends

```
SELECT city, ROUND(SUM(amount)/COUNT(AMOUNT), 2) AS ratio
FROM credit_card_transaction_analysis
WHERE DAYOFWEEK(date) IN (1, 7)
GROUP BY city
ORDER BY ratio DESC
LIMIT 1;
```

	city	ratio
▶	Sonepur	299905.00

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

```
WITH cte1 AS (SELECT city, date,
   SUM(t) OVER (PARTITION BY city ORDER BY date ROWS BETWEEN UNBOUNDED PRECEDING AND
CURRENT ROW) AS cumulative_count
   FROM (
   SELECT city, date, COUNT(amount) AS t
        FROM credit_card_transaction_analysis
        GROUP BY city, date
        ORDER BY city, date) subquery),
cte2 AS (
   SELECT city, MIN(date) AS date_reach_500
   FROM cte1
   WHERE cumulative_count >= 500
   GROUP BY city),
cte3 AS (
   SELECT city, MIN(date) AS first_transaction_date
```

```
    FROM credit_card_transaction_analysis
    GROUP BY city)
SELECT cr.city, DATEDIFF(cr.date_reach_500, cf.first_transaction_date) AS days_to_reach_500
FROM cte2 cr
JOIN cte3 cf
ON cr.city = cf.city
ORDER BY days_to_reach_500
LIMIT 1;
```

	city	days_to_reach_500
▶	Bengaluru	81