Mintra Sojiphan 6220422057

1. Create dataset

```sql
-- Create new table "dataset" and converts a string representation of date to a DATE object.

CREATE OR REPLACE TABLE `churn-334305.supermarket.dataset`
  AS SELECT PARSE_DATE("%Y%m%d", SAFE_CAST(SHOP_DATE AS STRING)) AS SHOP_DATE, CUST_CODE
  FROM `churn-334305.supermarket.data`
  WHERE CUST_CODE IS NOT NULL
  ORDER BY SHOP_DATE;
```

dataset

SCHEMA | DETAILS | PREVIEW

Row	SHOP_DATE	CUST_CODE
1	2006-04-10	CUST0000167145
2	2006-04-10	CUST0000774171
3	2006-04-10	CUST0000083013
4	2006-04-10	CUST0000047700
5	2006-04-10	CUST0000135397
6	2006-04-10	CUST0000414458
7	2006-04-10	CUST0000325292

2. Classify customer status

```sql
WITH temp1 AS (
  SELECT CUST_CODE,
    FORMAT_DATE('%Y-%m', DATE(SHOP_DATE)) SHOP_MONTH,
    DATE_DIFF(DATE(SHOP_DATE), '2000-01-01', MONTH) pos,
    DATE_DIFF(DATE(MIN(SHOP_DATE) OVER(PARTITION BY CUST_CODE)), '2000-01-01', MONTH) first_pos
  FROM `churn-334305.supermarket.dataset`
), temp2 AS (
  SELECT *, pos = first_pos AS new_cus
  FROM temp1
), temp3 AS (
  SELECT *, LAST_VALUE(new_cus) OVER(win) OR pos - 1 = LAST_VALUE(pos) OVER(win) AS repeat_cus
  FROM temp2
  WINDOW win AS (PARTITION BY CUST_CODE ORDER BY pos RANGE BETWEEN 1 PRECEDING AND 1 PRECEDING)
),
  temp4 AS (
  SELECT *,LAST_VALUE(new_cus) OVER(win) IS NULL OR pos = LAST_VALUE(pos) OVER(win) AS churned_cus
  FROM temp3
  WINDOW win AS (PARTITION BY CUST_CODE ORDER BY pos RANGE BETWEEN 1 PRECEDING AND 1 PRECEDING)
),
  temp5 AS (
  SELECT *, pos-2 = LAST_VALUE(pos) OVER(win) AS reactivate_cus
  FROM temp4
  WINDOW win AS (PARTITION BY CUST_CODE ORDER BY pos RANGE BETWEEN 2 PRECEDING AND 1 PRECEDING)
),
  final AS (
  SELECT SHOP_MONTH, CUST_CODE,
  CASE
   WHEN new_cus = true THEN 'New'
   WHEN repeat_cus = true THEN 'Repeated'
   WHEN churned_cus = true THEN 'Churned'
   WHEN reactivate_cus = true THEN 'Reactivated' END AS status,
  FROM temp5
)
```

3. Result

Code

```sql
SELECT SHOP_MONTH, status,
  if (status='Churned',-1*COUNT(DISTINCT CUST_CODE),COUNT(DISTINCT CUST_CODE)) COUNT
FROM final
WHERE status IS NOT NULL
GROUP BY SHOP_MONTH, status
```

Preview

Job information | Results | JSON | Execution detail

Row	SHOP_MONTH	status	COUNT
1	2006-06	Repeated	787
2	2007-10	Repeated	889
3	2007-01	Churned	-306
4	2007-05	Repeated	885
5	2007-09	Churned	-375
6	2007-09	Repeated	889
7	2006-09	Repeated	931
8	2006-10	New	255
9	2007-01	New	175
10	2006-07	Churned	-181
11	2008-04	Churned	-663
12	2008-03	Repeated	870

4. Visualization