

SQL 经典花式题:

Session	Timestamp	Action
1	14:00	View-Photo
1	14:01	view-photo
1	14:02	view-photo
2	14:05	View-product
2	14:06	view-product

问: 每个 session
连续 view photo 最多是多少

思路: 用 row_number over() 将 Timestamp 排开来 partition by session
用 row_number over() 将 Timestamp 排开来 partition by session, action
然后再对 2 个新的列相减

例

id	timestamp	action	rk_all	rk_event	rk_diff
1	14:00	photo	1	1	0
1	14:01	photo	2	2	0
1	14:02	photo	3	3	0
1	14:03	product	4	1	3
1	14:04	photo	5	4	1
1	14:05	photo	6	5	1

红-黄

3-4 只加 1
但 3-5 加 2.

现在红-黄和 action 列一一对应, 只要中间 product 导致 photo 不一致, 差值就会被改变

group by session, action 红-黄 (rk-diff)
→ 取 max(count) group by session

```

SELECT *
, ROW_NUMBER() OVER (PARTITION BY Session
ORDER BY Timestamp) AS rk_all
, ROW_NUMBER() OVER (PARTITION BY
Session, Action ORDER BY Timestamp) AS rk_event
, ROW_NUMBER() OVER (PARTITION BY Session
ORDER BY Timestamp) -
ROW_NUMBER() OVER (PARTITION BY
Session, Action ORDER BY Timestamp) AS rk_diff
INTO temp
FROM table;
SELECT Session, MAX(total_count) AS max_count
FROM (
SELECT Session, Action, rk_diff, COUNT(1) AS
total_count
FROM temp
WHERE Action = 'View Photo'
GROUP BY 1, 2, 3)
GROUP BY 1;
    
```

然后就刷到了 LC 180 题

id	num
1	1
2	1
3	1
4	2
5	2

→ 标签: Logs
找出至少 3 次连续出现的 num

常规做法 from Logs L1

Logs L2

Logs L3

where L1.id = L2.id - 1 = L3.id - 2

L1.num = L2.num = L3.num

select distinct num

但是可以将其考虑成只有一个 session 的方式

id	num	rk	rknum
1	1	1	1
2	1	2	2
3	1	3	3
4	2	4	1
5	2	5	2

rk = row_number() over
(order by id)

rknum
= row_number() over
(partition by num, order by id)

↑
将 id 看成之前的
time_stamp, num 看成 action
1 - photo 2 - view product.

```

SELECT *
, ROW_NUMBER() OVER (PARTITION BY Session
ORDER BY Timestamp) AS rk_all
, ROW_NUMBER() OVER (PARTITION BY
Session, Action ORDER BY Timestamp) AS rk_event
, ROW_NUMBER() OVER (PARTITION BY Session
ORDER BY Timestamp) -
ROW_NUMBER() OVER (PARTITION BY
Session, Action ORDER BY Timestamp) AS rk_diff
INTO temp
FROM table;
SELECT Session, MAX(total_count) AS max_count
FROM (
SELECT Session, Action, rk_diff, COUNT(1) AS
total_count
FROM temp
WHERE Action = 'View Photo'
GROUP BY 1,2,3)
GROUP BY 1;

```



思路: 对每个 session
 先全体 row-number()
 再对每个 action / num
 row-number()
 然后作差再 group by.

然后就刷到了 LC 180 题

id	num
1	1
2	1
3	1
4	2
5	2

→ 标题 logs

找出至少 3 次连续出现的 num??

常规做法 from Logs L1

logs L2

logs L3

where L1.id = L2.id - 1 = L3.id - 2

L1.num = L2.num = L3.num

select distinct num

)

但是可以将其考虑成只有一个 session 的方式

id	num	rk	rk-num	rk-diff
1	1	1	1	0
2	1	2	2	0
3	1	3	3	0
4	2	4	4	3
5	2	5	5	1

id 就是
time-stamp

num 就是
是 action
1 — photo
2 — product

rk = row-number() over
(order by id)

rk-diff
= rk - rk-num

rk-num
= row-number() over
(partition by num, order by i

现在 rk-diff 和 num 同步了 group by num, rk-diff
 select distinct num
 Having count(*) >= 3