

### Critical difference between Count() & Sum()

- 1) Count() - Count all NULL & NOT NULL values
- 2) Sum() - count 1's only or aggregate

### Worked on: Consecutive Streak

- 1) Count(case when active>0 then 1 else 0 end)
- 2) Sum(case when active >0 then 1 else 0 end)

Sum() - gives 1's summation only, Output - 2 only

Count() - gives 10 → which is the row count

```
1  with revised_data as (SELECT
2    user_id,
3    dt,
4    LAG(dt) OVER (PARTITION BY user_id ORDER BY dt) AS prev_dt,
5    (julianday(dt) - julianday(LAG(dt) OVER (PARTITION BY user_id ORDER BY dt))) AS gap_days
6    FROM trans_data
7  ),
8  flag as(
9    select
10   *,
11   case when gap_days >= 2 then 'active' else 0 end as active_2d
12   from revised_data
13 )
14 select
15   sum(case when active_2d = 'active' then 1 else 0 end) as active_users,
16   count(case when active_2d = 'active' then 1 else 0 end) as active
17   from flag
```

	active_users	active
1	2	10

Revised logic → where instead of 0 : NULL is taken

```

1   with revised_data as (SELECT
2       user_id,
3       dt,
4       LAG(dt) OVER (PARTITION BY user_id ORDER BY dt) AS prev_dt,
5       (julianday(dt) - julianday(LAG(dt) OVER (PARTITION BY user_id ORDER BY dt))) AS gap_days
6   FROM trans_data
7   ),
8   flag as(
9       select
10          *,
11          case when gap_days >= 2 then 'active' else 0 end as active_2d
12      from revised_data
13   )
14   select
15      sum(case when active_2d = 'active' then 1 else 0 end) as active_users,
16      count(case when active_2d = 'active' then 1 else null end) as active
17     from flag

```

	active_users	active
1	2	2