

Day 20 – Popular Posts

Problem Statement:

The column '*perc_viewed*' in the table '*post_views*' denotes the percentage of the session duration time the user spent viewing a post. Using it, calculate the **total time** that each post was viewed by users. Output post ID and the total viewing time in seconds, but only for posts with a total viewing time of over 5 seconds.

Stratascratch Link: https://platform.stratascratch.com/coding/2073-popular-posts?code_type=5

Input:

Table1: User_sessions

session_id integer	user_id character varying (10)	session_starttime timestamp without time zone	session_endtime timestamp without time zone	platform character varying (20)
1	U1	2020-01-01 12:14:28	2020-01-01 12:16:08	Windows
2	U1	2020-01-01 18:23:50	2020-01-01 18:24:00	Windows
3	U1	2020-01-01 08:15:00	2020-01-01 08:20:00	IPhone
4	U2	2020-01-01 10:53:10	2020-01-01 10:53:30	IPhone
5	U2	2020-01-01 18:25:14	2020-01-01 18:27:53	IPhone
6	U2	2020-01-01 11:28:13	2020-01-01 11:31:33	Windows
7	U3	2020-01-01 06:46:20	2020-01-01 06:58:13	Android
8	U3	2020-01-01 10:53:10	2020-01-01 10:53:50	Android
9	U3	2020-01-01 13:13:13	2020-01-01 13:34:34	Windows
10	U4	2020-01-01 08:12:00	2020-01-01 12:23:11	Windows
11	U4	2020-01-01 21:54:03	2020-01-01 21:54:04	IPad

Table2: Post_views

session_id integer	post_id integer	perc_viewed double precision
1	1	2
1	2	4
1	3	1
2	1	20
2	2	10
2	3	10
2	4	21
3	2	1
3	4	1
4	2	50
4	3	10
6	2	2
8	2	5
8	3	2.5

ExpectedOutput:

post_id integer	total_viewtime numeric
4	5.1
2	24.0

SOLUTION BREAKDOWN

1. Extract session_duration in SECONDS from session_starttime and session_endtime
2. Find view_time_per_session_post using perc_viewed column and session_duration
3. Sum the view_time and group by post_id
4. Filter for those records with total_view_time more than 5 seconds

Draft

```
select us.session_id,  
       pv.post_id,  
       DATEDIFF(SECOND,session_starttime,session_endtime) as  
session_duration,  
       pv.perc_viewed,  
       DATEDIFF(SECOND,session_starttime,session_endtime) *  
pv.perc_viewed/100 as view_time_per_session_post  
from post_views pv left join user_sessions us on  
pv.session_id=us.session_id
```

	session_id	post_id	session_duration	perc_viewed	view_time_per_session_post
1	1	1	100	2	2
2	1	2	100	4	4
3	1	3	100	1	1
4	2	1	10	20	2
5	2	2	10	10	1
6	2	3	10	10	1
7	2	4	10	21	2.1
8	3	2	300	1	3
9	3	4	300	1	3
10	4	2	20	50	10
11	4	3	20	10	2
12	6	2	200	2	4
13	8	2	40	5	2
14	8	3	40	2.5	1

Final Solution

```
SELECT *
FROM (
    SELECT pv.post_id,
           SUM(DATEDIFF(SECOND,session_starttime,session_endtime) *
pv.perc_viewed/100) AS total_viewtime
    FROM post_views pv
         LEFT JOIN user_sessions us ON pv.session_id=us.session_id
    GROUP BY pv.post_id
) sq
WHERE total_viewtime > 5
```

OUTPUT

	post_id	total_viewtime
1	2	24
2	4	5.1