

# Taziev Timur

1. Сделать сессионизацию по таблице events, с разрывом сессии через 30 минут или после смены канала привлечения.

```
SELECT
  ID,
  Datetime,
  Medium,
  SUM(session_start) OVER (
    PARTITION BY ID
    ORDER BY Datetime) AS session_id
FROM
  (SELECT
    ID,
    Datetime,
    Medium,
    if( Medium != neighbor(Medium, -1)
      OR Datetime - neighbor(Datetime, -1) > 30 * 60, 1, 0) AS session_start
  FROM events
  ORDER BY 1, 2)
```

	ID	Datetime	Medium	session_id
120	++++kfcnQ...	2022-06-17 11:43:42	referral	3
121	++++kfcnQ...	2022-06-17 11:43:43	referral	3
122	+++24L3pT...	2022-06-17 07:58:31	cpc	1
123	+++24L3pT...	2022-06-17 07:58:32	cpc	1
124	+++24L3pT...	2022-06-17 07:58:32	cpc	1
125	+++24L3pT...	2022-06-17 07:58:46	cpc	1
126	+++24L3pT...	2022-06-18 06:13:29	cpc	2
127	+++24L3pT...	2022-06-18 06:13:30	cpc	2
128	+++24L3pT...	2022-06-18 06:13:30	cpc	2
129	+++24L3pT...	2022-06-18 06:13:30	cpc	2
130	+++24L3pT...	2022-06-18 06:13:31	cpc	2
131	+++24L3pT...	2022-06-18 06:13:32	cpc	2
132	+++24L3pT...	2022-06-18 06:13:33	cpc	2
133	+++24L3pT...	2022-06-18 06:13:33	cpc	2
134	+++24L3pT...	2022-06-18 06:13:33	cpc	2
135	+++24L3pT...	2022-06-18 06:13:33	cpc	2
136	+++24L3pT...	2022-06-18 06:13:33	cpc	2
137	+++24L3pT...	2022-06-18 06:13:34	cpc	2
138	+++24L3pT...	2022-06-18 06:13:36	cpc	2
139	+++24L3pT...	2022-06-18 06:13:36	cpc	2
140	+++24L3pT...	2022-06-18 06:13:36	cpc	2
141	+++24L3pT...	2022-06-18 06:13:36	cpc	2
142	+++24L3pT...	2022-06-18 06:13:36	cpc	2
143	+++24L3pT...	2022-06-18 06:13:36	cpc	2
144	+++24L3pT...	2022-06-18 06:13:37	cpc	2
145	+++24L3pT...	2022-06-18 06:13:37	cpc	2
146	+++24L3pT...	2022-06-18 06:13:37	cpc	2
147	+++24L3pT...	2022-06-18 06:15:53	cpc	2
148	+++24L3pT...	2022-06-19 03:47:17	organic	3
149	+++24L3pT...	2022-06-19 03:47:20	organic	3
150	+++24L3pT...	2022-06-19 03:47:20	organic	3
151	+++24L3pT...	2022-06-19 03:47:20	organic	3
152	+++24L3pT...	2022-06-19 03:47:25	organic	3

2. Посчитать месячный ретеншн в таблице orders по когортам с помощью функции retention

```
SELECT
    SUM(User_Month_Retention[1]) as Active_1_month,
    SUM(User_Month_Retention[2]) as Active_2_month,
    SUM(User_Month_Retention[3]) as Active_3_month
FROM
    (WITH toMonth(ActionDate) as active_month
     SELECT
         Client_id,
         retention(
             active_month = '6',
             active_month = '7',
             active_month = '8') AS User_Month_Retention
     FROM orders
     GROUP BY Client_id
     ORDER BY Client_id)
```

	Active_1_month	Active_2_month	Active_3_month
1	5	3	0

3. Посчитать месячный ретеншн в таблице orders по когортам не используя функцию retention.

```
SELECT
    Client_id,
    groupUniqArray(toMonth(ActionDate)) as retention_month
FROM orders
GROUP BY Client_id
```

	Client_id	retention_month
1	1	0x0607
2	2	0x0607
3	3	0x07
4	4	0x07
5	5	0x06
6	6	0x06
7	7	0x0607
8	8	0x07
9	9	0x07
10	10	0x07
11	11	0x07
12	12	0x07
13	13	0x08

4. Сделать массив из повторенных 5 раз цифр 1, затем 5 цифр 2 и тд до 5 цифр 10(1,1,1,1,2,2,2,2, ..., 10,10,10,10,10), не перечисляя все 50 цифр, а воспользовавшись функциями arrayResize и arrayFlatten.

5. Для массива из прошлой задачи вывести каждый третий элемент

```
SELECT
    range(1, 11) as numbers,

    arrayResize(
        arraySort(
            arrayFlatten(
                arrayMap(x-> arrayResize([11], 6, numbers[x]), numbers))), 50) as result,

    arrayResize(
        arrayMap(x-> arrayElement(result, x*3), arrayEnumerate(result)), 16) as res_3
```

	numbers	result	res_3
1	0x010203040506...	0x0101010101010202020203030303040...	0x0102020303040505060607080809090A

6. Задача из прошлого ДЗ, но теперь ее нужно решить проще с помощью функций clickhouse: за каждый день вывести канал привлечения с максимальным GMV, и собственно сам GMV этого канала

```
WITH
  T1 AS (SELECT
    ActionDate,
    Medium,
    sum(NetSales) AS GMV
  FROM orders
  GROUP BY ActionDate, Medium),
  T2 AS (SELECT
    ActionDate,
    argMax(Medium,GMV) AS MAX_MEDIUM
  FROM T1
  GROUP BY ActionDate)
SELECT
  ActionDate,
  T2.MAX_MEDIUM AS MAX_MEDIUM,
  T1.GMV AS GMV
FROM
  T2 JOIN T1
  ON T2.MAX_MEDIUM = T1.Medium
  AND T2.ActionDate = T1.ActionDate
ORDER BY ActionDate
```



Миша Попов

23:40 8 окт.

GMV - это SUM(NetSales), а не SUM(NetSalse \* Qty)

	ActionDate	max_medium	GMV
1	2022-06-02	Organic	1290
2	2022-06-04	Advertising	12126.6669921875
3	2022-06-05	Advertising	3031.666748046875
4	2022-06-08	Advertising	3031.666748046875
5	2022-06-20	Advertising	2563.60009765625
6	2022-06-21	Advertising	1281.800048828125
7	2022-06-26	Advertising	6103.33349609375
8	2022-06-27	Advertising	2580
9	2022-06-28	Referral	4813.33349609375
10	2022-06-29	Referral	2406.666748046875
11	2022-06-30	Advertising	11375
12	2022-07-02	Advertising	6932.5
13	2022-07-04	Advertising	4621.66650390625
14	2022-07-05	Advertising	6932.499755859375
15	2022-07-06	Advertising	2310.833251953125
16	2022-07-07	Advertising	2250.791748046875
17	2022-07-08	Advertising	2250.791748046875
18	2022-07-09	Advertising	4526.625
19	2022-07-10	Advertising	2275.833251953125
20	2022-07-11	Advertising	4441.66650390625
21	2022-07-12	Advertising	3433.313232421875
22	2022-07-13	Advertising	8663.3330078125
23	2022-07-14	Advertising	2165.833251953125
24	2022-07-15	Advertising	5265.8134765625
25	2022-07-16	Advertising	3998.33349609375
26	2022-07-17	Referral	3998.33349609375
27	2022-07-18	Advertising	7996.6669921875
28	2022-07-19	Advertising	1999.166748046875
29	2022-07-20	Referral	5997.5

30	2022-07-21	Organic	1999.166748046875
31	2022-07-22	Referral	3673.2333984375
32	2022-07-23	Advertising	9995.833984375
33	2022-07-24	Referral	1836.61669921875
34	2022-07-26	Organic	3998.33349609375
35	2022-07-27	Organic	1999.166748046875
36	2022-07-31	Organic	5245
37	2022-08-28	Referral	3120
38	2022-08-29	Referral	3120

7. Найти клиентов (последовательность шагов), которые купили 'Coffee Columbia', и затем в течении 3 дней купили также 'Coffee Brasil'

```
SELECT
    o1.Client_id
FROM
    (SELECT
        Client_id,
        ActionDate
    FROM orders
    WHERE ItemName = 'Coffee Columbia') o1

JOIN
    orders o2
ON o1.Client_id = o2.Client_id
WHERE
    o2.ItemName = 'Coffee Brasil'
    AND o2.ActionDate <= o1.ActionDate + INTERVAL 3 DAY
```

	Client_id
1	13