**Задача 2.1**

SELECT event\_name, COUNT(event\_name) AS number\_of

FROM(

SELECT event\_name, min(event\_timestamp), user\_pseudo\_id

FROM(

SELECT event\_name, event\_timestamp, user\_pseudo\_id FROM test\_table\_001 WHERE event\_name in ('gdpr\_group\_1', 'gdpr\_group\_2', 'gdpr\_group\_3', 'gdpr\_group\_4', 'gdpr\_group\_5', 'gdpr\_group\_6')

UNION ALL

SELECT event\_name, (event\_timestamp), user\_pseudo\_id FROM test\_table\_002 WHERE event\_name in ('gdpr\_group\_1', 'gdpr\_group\_2', 'gdpr\_group\_3', 'gdpr\_group\_4', 'gdpr\_group\_5', 'gdpr\_group\_6')

UNION ALL

SELECT event\_name, (event\_timestamp), user\_pseudo\_id FROM test\_table\_003 WHERE event\_name in ('gdpr\_group\_1', 'gdpr\_group\_2', 'gdpr\_group\_3', 'gdpr\_group\_4', 'gdpr\_group\_5', 'gdpr\_group\_6')

UNION ALL

SELECT event\_name, (event\_timestamp), user\_pseudo\_id FROM test\_table\_004 WHERE event\_name in ('gdpr\_group\_1', 'gdpr\_group\_2', 'gdpr\_group\_3', 'gdpr\_group\_4', 'gdpr\_group\_5', 'gdpr\_group\_6')

UNION ALL

SELECT event\_name, (event\_timestamp), user\_pseudo\_id FROM test\_table\_005 WHERE event\_name in ('gdpr\_group\_1', 'gdpr\_group\_2', 'gdpr\_group\_3', 'gdpr\_group\_4', 'gdpr\_group\_5', 'gdpr\_group\_6')

UNION ALL

SELECT event\_name, (event\_timestamp), user\_pseudo\_id FROM test\_table\_006 WHERE event\_name in ('gdpr\_group\_1', 'gdpr\_group\_2', 'gdpr\_group\_3', 'gdpr\_group\_4', 'gdpr\_group\_5', 'gdpr\_group\_6')

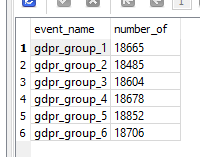
)

GROUP BY user\_pseudo\_id

)

GROUP BY event\_name

**Результат:**



**Задача 2.2**

SELECT event\_date,

COUNT(DISTINCT user\_pseudo\_id),

COUNT(user\_pseudo\_id),

COUNT(user\_pseudo\_id) / COUNT(DISTINCT user\_pseudo\_id) AS mean

FROM (

SELECT p0.user\_pseudo\_id,

p0.event\_date,

p0.event\_timestamp,

p0.event\_name,

p1.event\_name,

p1.event\_timestamp,

(CAST(p1.event\_timestamp AS int) - CAST(p0.event\_timestamp AS int))/1000000 AS delta

FROM (

SELECT \*

FROM (

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_001

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_002

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_003

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_004

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_005

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_006

)

) AS p0

JOIN (

SELECT \*

FROM (

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_001

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_002

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_003

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_004

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_005

UNION ALL

SELECT event\_date, event\_timestamp, event\_name, user\_pseudo\_id FROM test\_table\_006

)

) AS p1 on p0.user\_pseudo\_id=p1.user\_pseudo\_id

WHERE p0.event\_name='first\_open' AND

p1.event\_name='video\_show' AND

(CAST(p1.event\_timestamp AS int) - CAST(p0.event\_timestamp AS int))/1000000 < 3600 AND (CAST(p1.event\_timestamp AS int) - CAST(p0.event\_timestamp AS int))/1000000 > 0

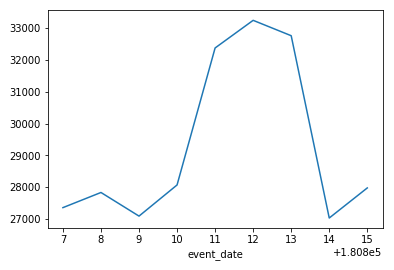
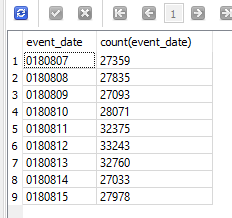
)

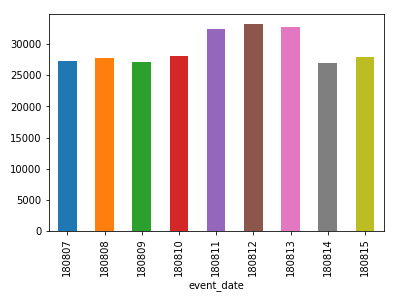
GROUP BY event\_date

**Результат:**

\*\*\*вопрос поставлен неоднозначно (не указано, какое значение необходимо найти сумма\среднее\медиана)

1. Сумма событий по дням (lineplot приувеличивает визуально реальную разницу)





1. Среднее\Медианное значение событий по дням недели

