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:

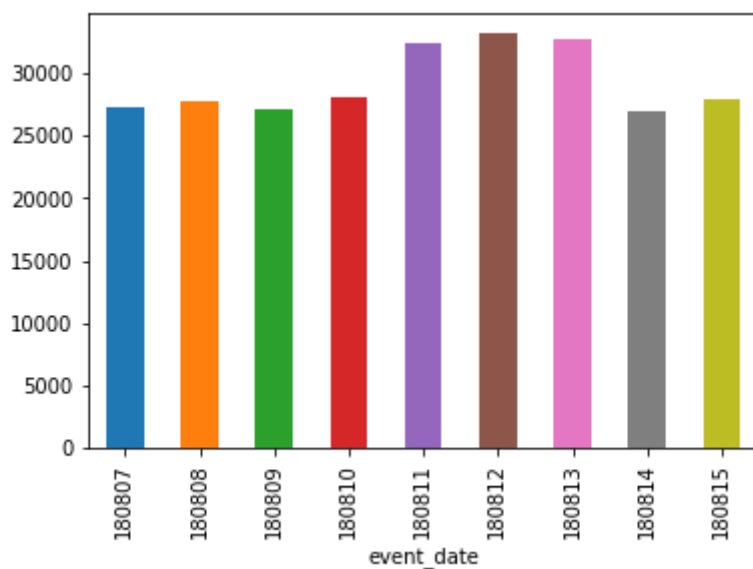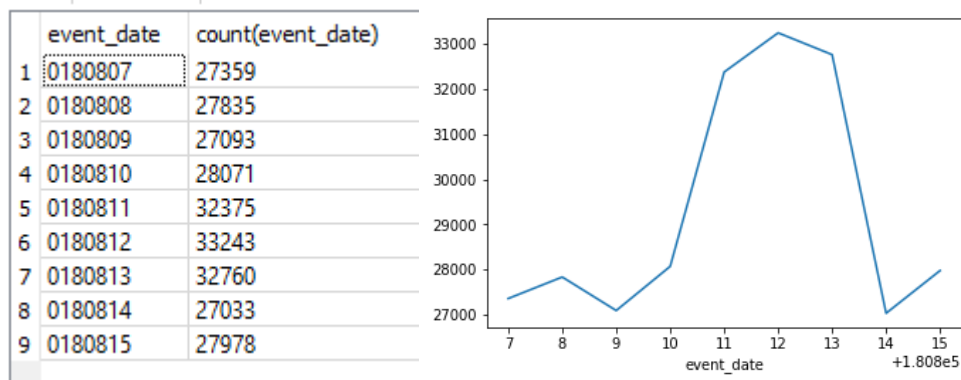|	event_name	number_of
1	gdpr_group_1	18665
2	gdpr_group_2	18485
3	gdpr_group_3	18604
4	gdpr_group_4	18678
5	gdpr_group_5	18852
6	gdpr_group_6	18706

Задача 2.2

```sql
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)

	event_date	count(event_date)
1	0180807	27359
2	0180808	27835
3	0180809	27093
4	0180810	28071
5	0180811	32375
6	0180812	33243
7	0180813	32760
8	0180814	27033
9	0180815	27978

2) Среднее\Медианное значение событий по дням недели