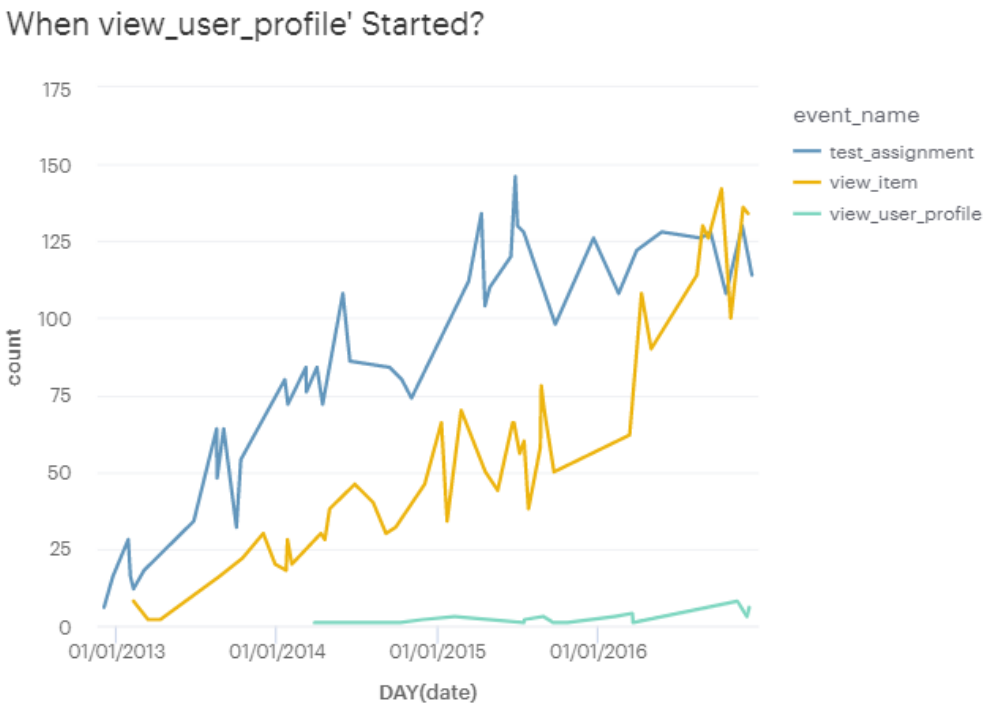
Identifying Unreliable Data + Nulls

Exercise 1: Using any methods you like determine if you can trust this events table.

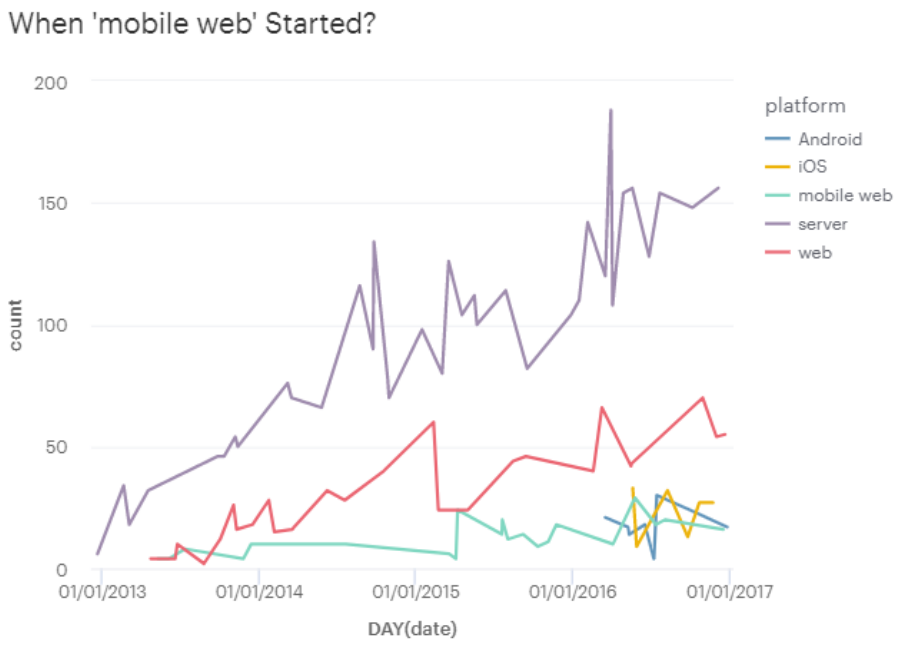
```
select date (event_time) as date,
    count(*) as rows
from dsv1069.events_201701
group by 1
limit 5
```

Exercise 2: Using any methods you like, determine if you can you trust this events table. (HINT: When did we start recording events on mobile)

```
select date (event_time) as date,
       event_name,
       count(*)
from dsv1069.events_ex2
group by 1,2
```




```
select date (event_time) as date,
    platform,
    count(*)
from dsv1069.events_ex2
group by 1, 2
```



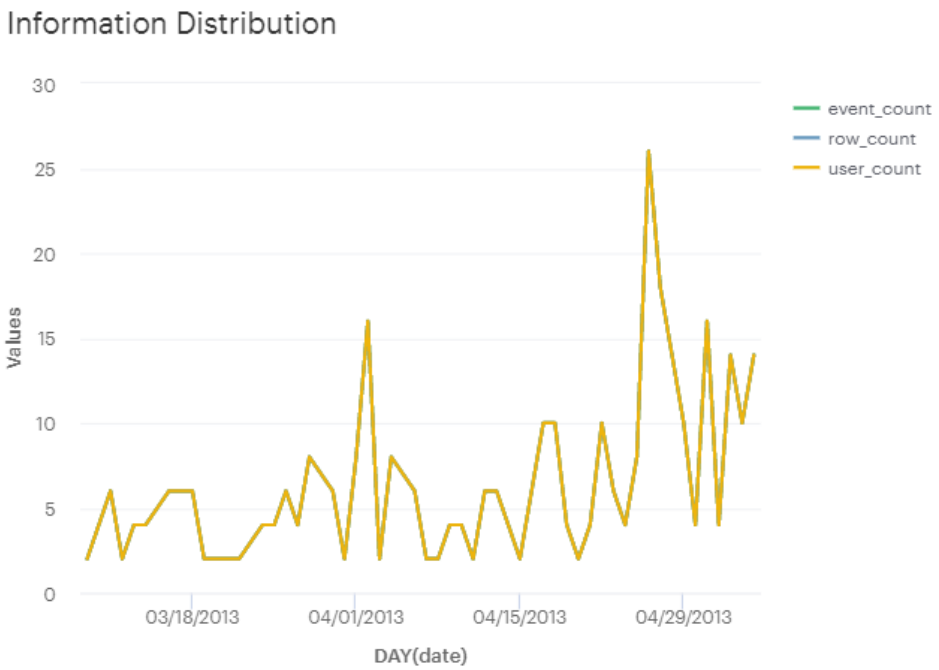
Exercise 3: Imagine that you need to count item views by day. You found this table item_views_by_category_temp - should you use it to answer your question? NO.

```
1.  select date(event_time) as date,
2.      count(*) as row_count
3.  from dsv1069.events
4.  group by 1
```

Exercise 4: Using any methods you like, decide if this table is ready to be used as a source of truth. YES.

```
1.  select date(event_time) as date,
2.      platform,
3.      count(*) as row_count,
```

```
4.      count(event_id) as event_count,
5.      count(user_id) as user_count
6.  from dsv1069.events
7.  group by 1, 2
```



Exercise 5: Is this the right way to join orders to users?

```
1.  select count(*)
2.  from dsv1069.orders o
3.  join dsv1069.users u
4.   on o.user_id = coalesce(u.parent_user_id,u.id)
```

	count
1	47310

Mode Report Link: https://app.mode.com/sum14/reports/85fe20bab776