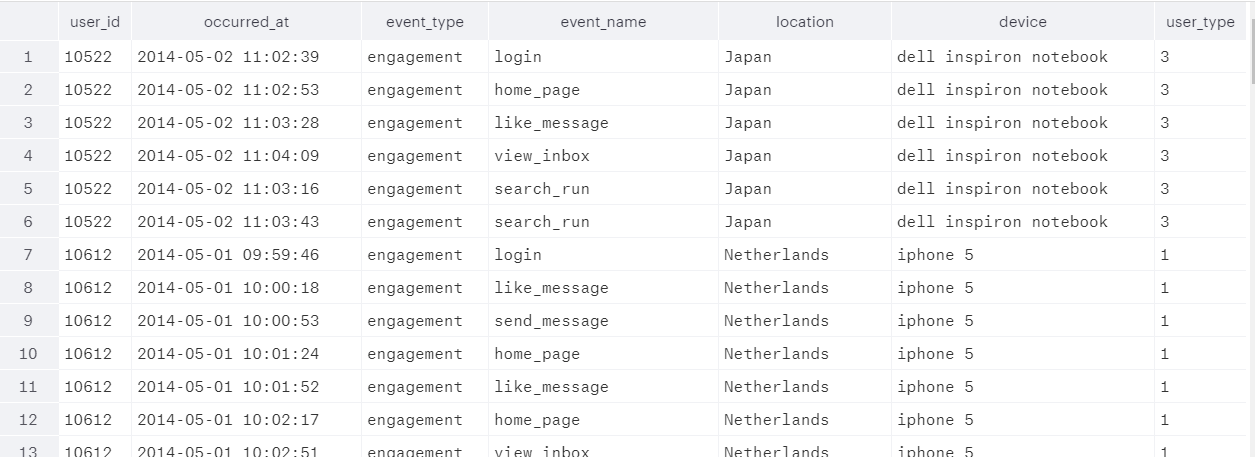
Yammer is a social network for communicating with coworkers. Individuals share documents, updates, and ideas by posting them in groups. Yammer is free to use indefinitely, but companies must pay license fees if they want access to administrative controls, including integration with user management systems like ActiveDirectory.

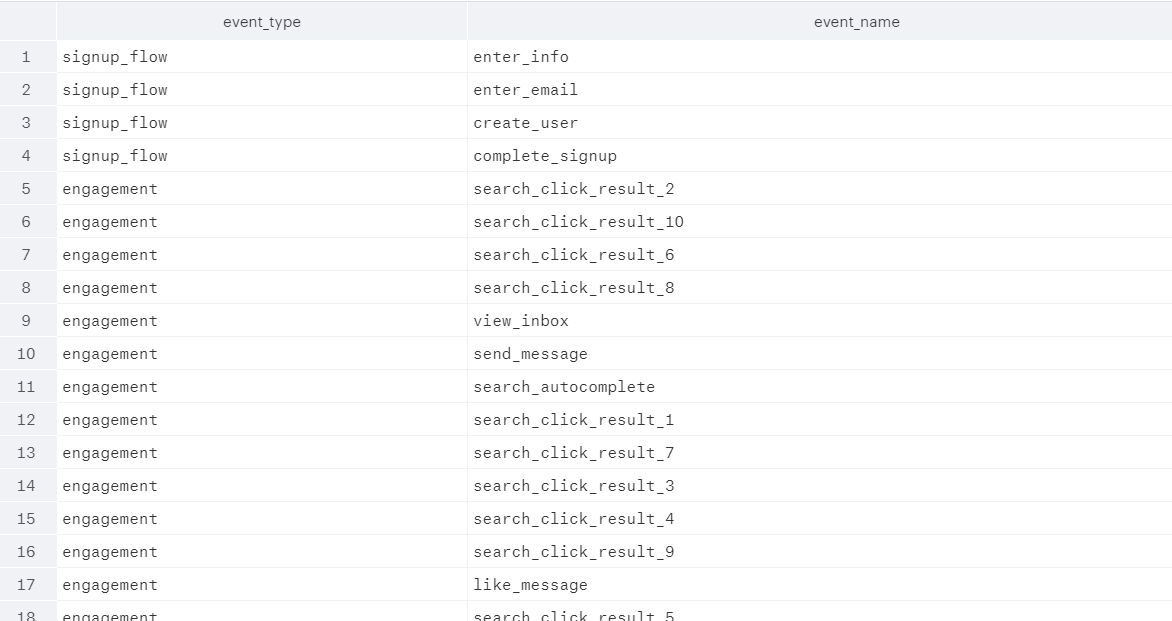
Overview of the table.

select \* from tutorial.yammer\_events;



select distinct event\_type, event\_name from tutorial.yammer\_events

order by event\_type desc;



3 Cases that I can analyze

[A Drop in Engagement](https://community.modeanalytics.com/sql/tutorial/a-drop-in-user-engagement/)

Engagement dips—you figure out the source of the problem.

[Understanding Search](https://community.modeanalytics.com/sql/tutorial/understanding-search-functionality/)

The product team is thinking about revamping search. Your job is to figure out whether they should change it at all, and if so, what should be changed.

[The Best A/B Test Ever](https://community.modeanalytics.com/sql/tutorial/validating-ab-test-results/)

A new feature tests off the charts. Your job is to determine the validity of the experiment.

**A Drop in User Engagement**

Weekly Active Users

select date\_trunc('week', occurred\_at),

count(distinct user\_id) as weekly\_active\_users

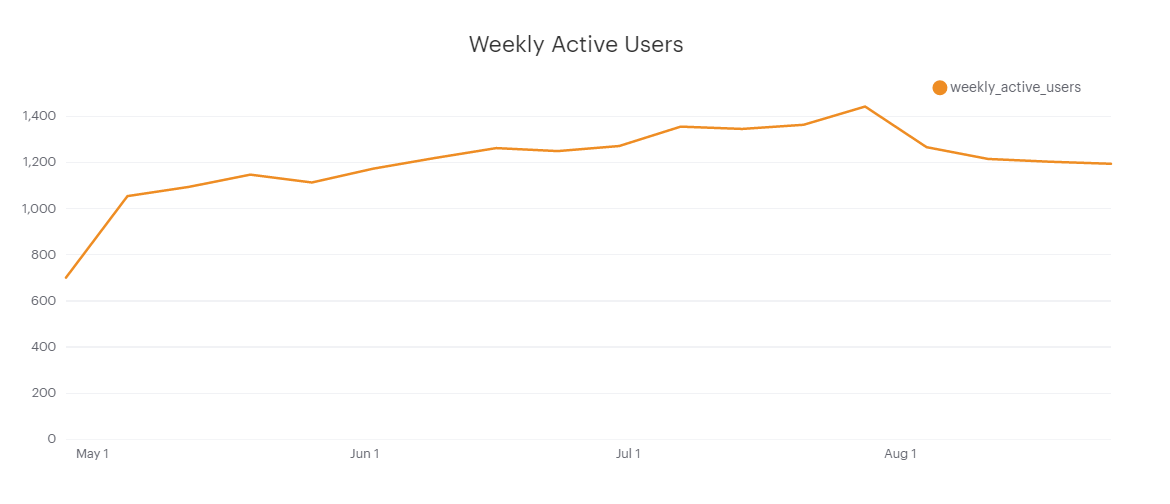
from tutorial.yammer\_events

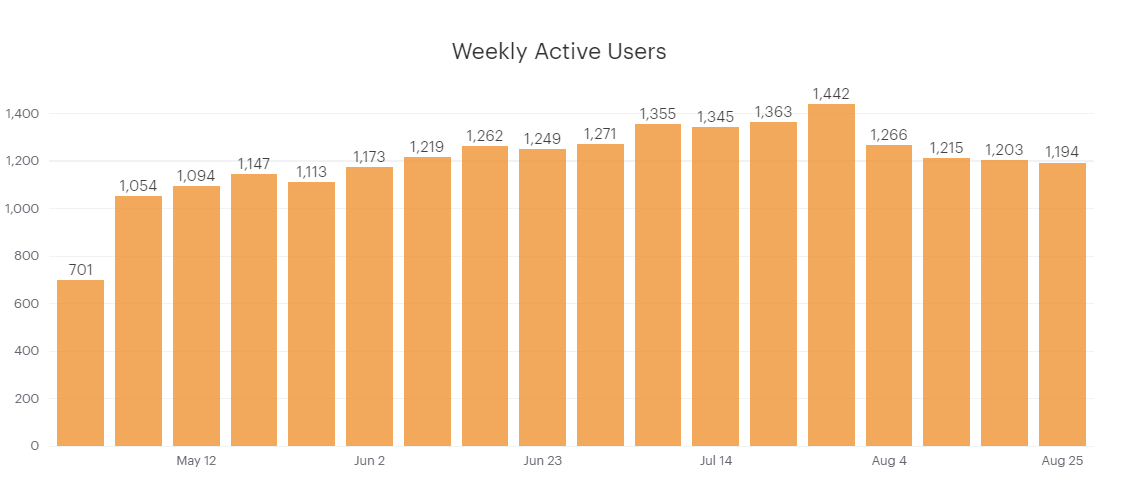
where event\_type='engagement'

and event\_name='login'

group by 1

order by 1;





The above chart shows the number of engaged users each week. So what factors are driving the dip in retention shown in the chart above?

Hypothesis: to better understand and analyze the issues, I made some hypothesis.

**Holiday**: It’s likely that people using a work application like Yammer might engage at a lower rate on holidays. If one country has much lower engagement than others, it’s possible that this is the cause.

**Broken feature:** It is possible that something in the application is broken, and therefore impossible for people to use. This is a little harder to pinpoint because different parts of the application would show differently in the metrics. For example, if something in the signup flow broke, preventing new users from joining Yammer, growth would also be down. If a Mobile app was unstable and crashed, engagement would be down for only that device type.

**Broken tracking code:** It’s possible that the code that logs events is, itself, broken. If you see a drop to absolutely zero events of a certain type and you rule out a broken feature, then this is a possibility.

**Traffic anomalies from bots:** Most major website see a lot of activity from bots. A change in the product or infrastructure that might make it harder for bots to interact with the site could decrease engagement (assuming bots look like real users). This is tricky to determine because you have to identify bot-like behavior through patterns or specific events.

**Traffic shutdown to your site:** It is possible for internet service providers to block your site. This is pretty rare for professional applications, but nevertheless possible.

**Marketing event:** A Super Bowl ad, for example, might cause a massive spike in sign-ups for the product. But users who enter through one-time marketing blitzes often retain at lower rates than users who are referred by friends, for example. Because the chart uses a rolling 7-day period, this will register as high engagement for one week, then almost certainly look like a big drop in engagement the following week. Most often, the best way to determine this is to simply ask someone in the Marketing department if anything big happened recently.

**Bad data:** There are lots of ways to log bad data. For example, most large web apps separate their QA data from production data. One way or another, QA data can make its way into the production database. This is not likely to be the problem in this particular case, as it would likely show up as additional data logged from very few users.

**Search crawler changes:** For a website that receives a lot of traffic, changes in the way search engines index them could cause big swings in traffic.

For the following solution, a session is defined as a string of events logged by a user without a 10-minute break between any two events. So if a user goes 10 minutes without logging an event, the session is ended and her next engagement will be considered a new session.

1. One of the easiest things to check is growth, both because it’s easy to measure and because most companies (Yammer included) track this closely already. We can see nothing really changed about the growth rate—it continues to be high during the week, low on weekends:

SELECT DATE\_TRUNC('day',created\_at) AS day,

COUNT(\*) AS all\_users,

COUNT(CASE WHEN activated\_at IS NOT NULL THEN u.user\_id ELSE NULL END) AS activated\_users

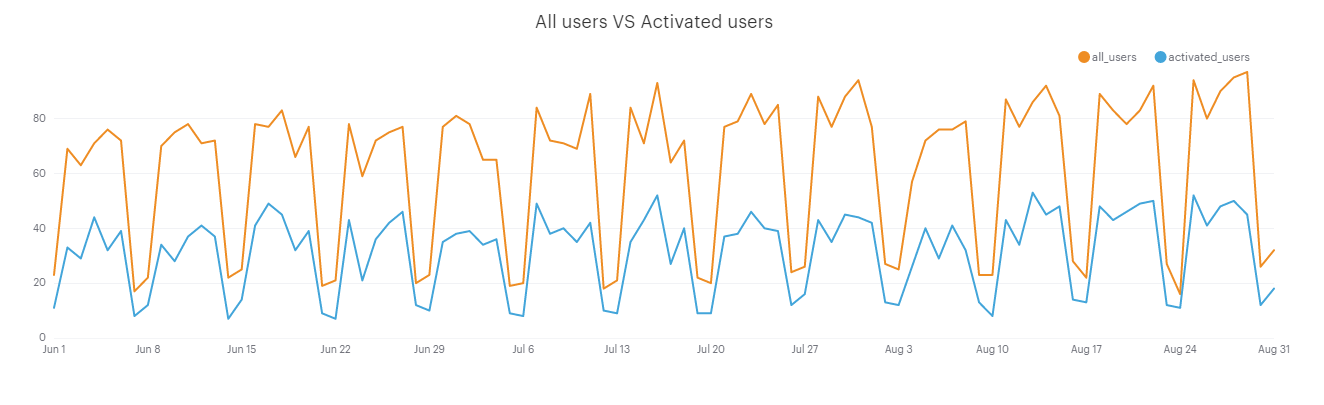
FROM tutorial.yammer\_users u

WHERE created\_at >= '2014-06-01'

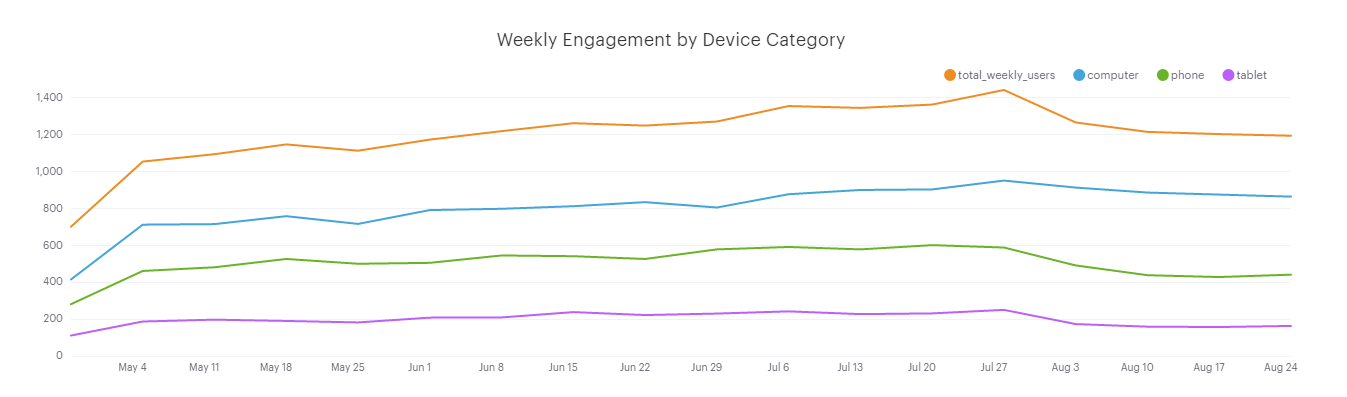
AND created\_at < '2014-09-01'

GROUP BY 1

ORDER BY 1



2. The understanding that the problem is localized to older users leads us to believe that the problem probably isn’t related to a one-time spike from marketing traffic or something that is affecting new traffic to the site like being blocked or changing rank on search engines. Now let’s take a look at various device types to see if the problem is localized to any particular product.



select date\_trunc('week',occurred\_at),count(distinct user\_id) as total\_weekly\_users,

count(distinct case when device in ('macbook pro','lenovo thinkpad','macbook air','dell inspiron notebook',

'asus chromebook','dell inspiron desktop','acer aspire notebook','hp pavilion desktop','acer aspire desktop','mac mini')

then user\_id else null end) as computer,

count(distinct case when device in ('iphone 5','samsung galaxy s4','nexus 5','iphone 5s','iphone 4s','nokia lumia 635',

'htc one','samsung galaxy note','amazon fire phone')

then user\_id else null end) as phone,

count(distinct case when device in ('ipad air','nexus 7','ipad mini','nexus 10','kindle fire','windows surface',

'samsumg galaxy tablet')

then user\_id else null end) as tablet

from tutorial.yammer\_events

where event\_type='engagement'

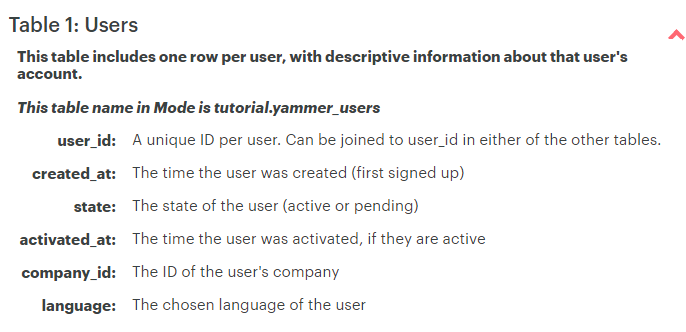
and event\_name='login'

group by 1

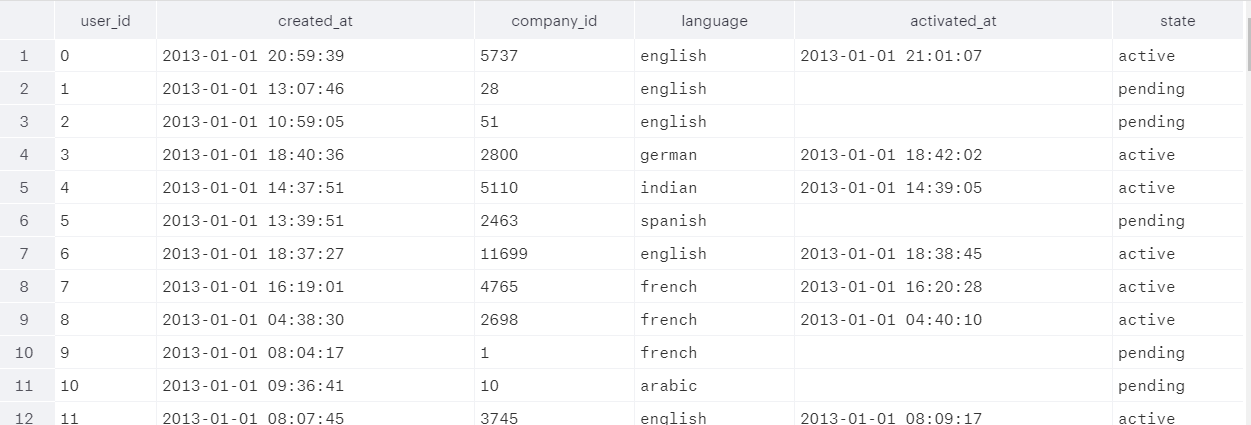
order by 1

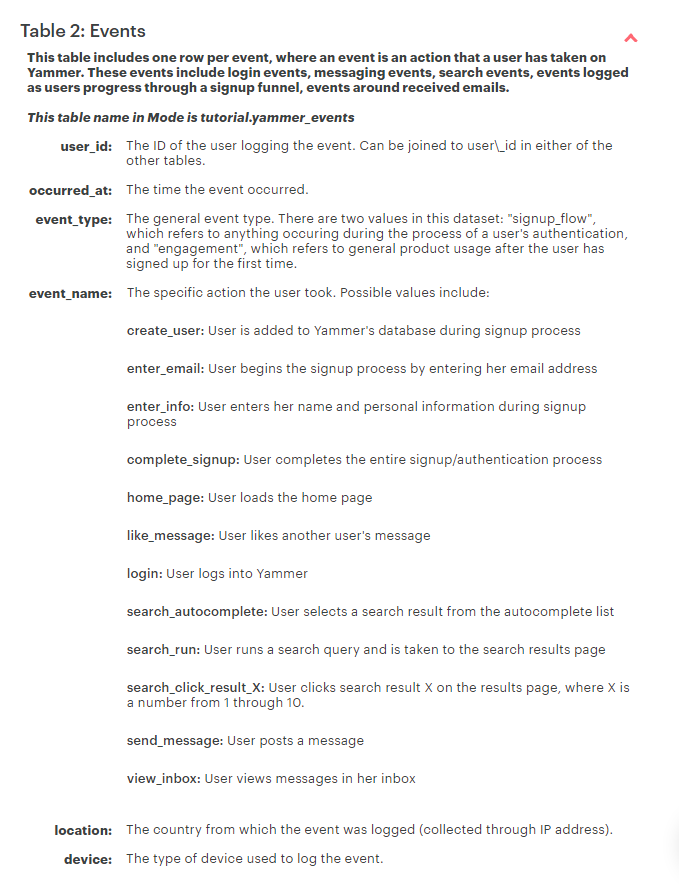
Reference:

Tables that I used to analyze.

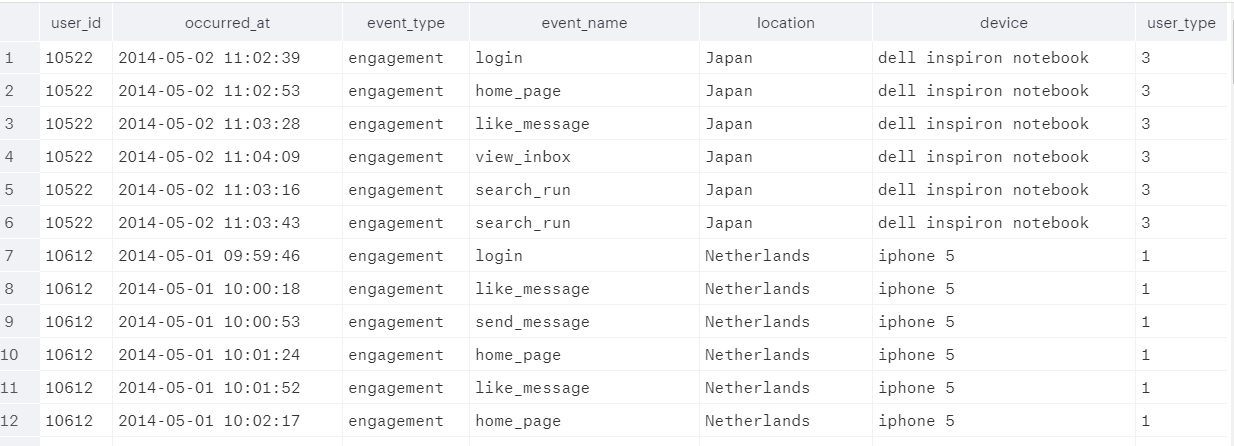


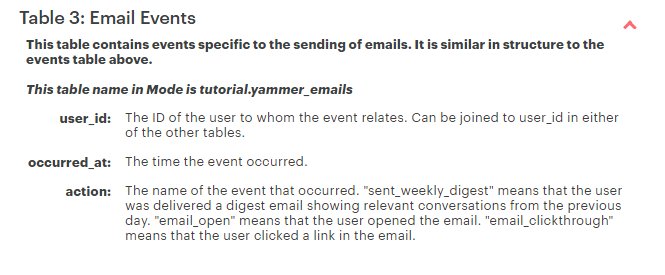
select \* from tutorial.yammer\_users;



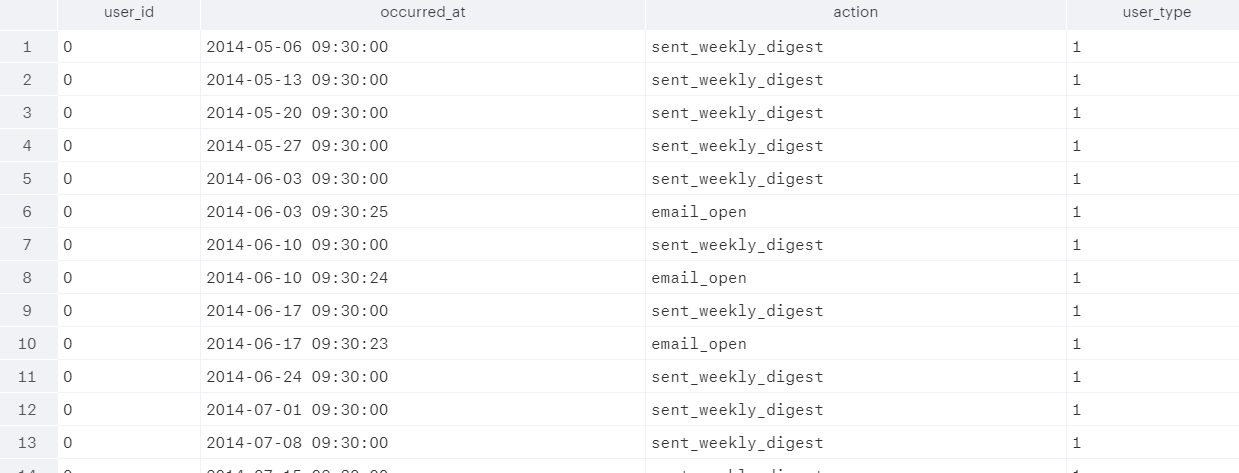


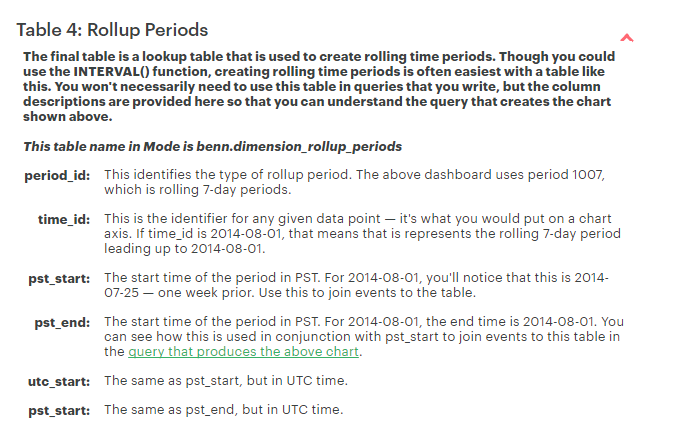
select \* from tutorial.yammer\_events;





select \* from tutorial.yammer\_emails;





select \* from benn.dimension\_rollup\_periods;

