Case Study from Mode Analytics: "Yammer Analysis: Drop in User Engagement"

Date used from: https://community.modeanalytics.com/sql/tutorial/a-drop-in-user-engagement/

The problem

You show up to work Tuesday morning, September 2, 2014. The head of the Product team walks over to your desk and asks you what you think about the latest activity on the user engagement dashboards. You fire them up, and something immediately jumps out.

The above chart shows the number of engaged users each week. Yammer defines engagement as having made some type of server call by interacting with the product (shown in the data as events of type "engagement"). Any point in this chart can be interpreted as "the number of users who logged at least one engagement event during the week starting on that date."

You are responsible for determining what caused the dip at the end of the chart shown above and, if appropriate, recommending solutions for the problem.

Getting oriented

Before you even touch the data, come up with a list of possible causes for the dip in retention shown in the chart above. Make a list and determine the order in which you will check them. Make sure to note how you will test each hypothesis. Think carefully about the criteria you use to order them and write down the criteria as well.

Initial Comments

Possible causes include:

Vacation: There is a 15% user dropoff the first week of August. Given the time of year, it could be that there are overall fewer users on the system than normal. We can look at total user count by week to understand the relative % vs the absolute numbers.

Systems Issue: There might be an error with the user interface, back-end system, etc. preventing users from using the system as normal. If there is data relative to error codes or similar, we can look to leverage and analyze.

Traffic differences: understanding if there might have been a change in customer usage (e.g. major client stopped using Yammer), or the like. Analyze total usage trends and traffic volume.

Digging in

Once you have an ordered list of possible problems, it's time to investigate.

For this problem, you will need to use four tables. The tables names and column definitions are listed below—click a table name to view information about that table. Note: this data is fake and was generated for the purpose of this case study. It is similar in structure to Yammer's actual data, but for privacy and security reasons it is not real.

Initial Analysis

Confirming that indeed, engagement has dropped. Looks like the first week of August shows about a 10% drop-off of engagement events.

SQL Code:
-- Engagement event counts by week
SELECT DATE_TRUNC('week', occurred_at) as week_date, count(event_name)
FROM tutorial.yammer_events events
WHERE event_type = 'engagement'
GROUP BY DATE_TRUNC('week', occurred_at)
ORDER BY DATE_TRUNC('week', occurred_at) asc
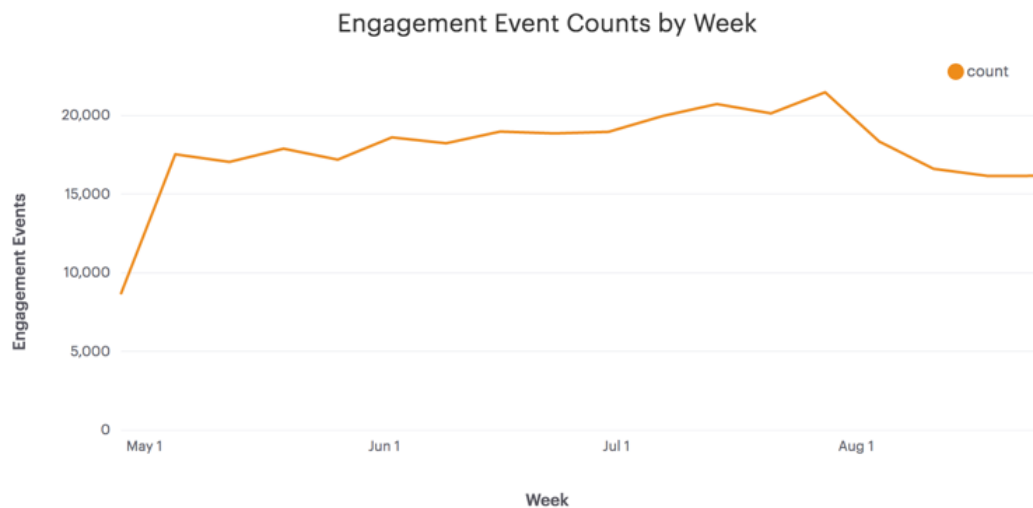
Hypothesis #1: summer holidays and vacation are causing a dip in user engagement.

First, I will explore the users table and understand how many active vs pending users we have. There are 19,066 total users, with 9,381 active users.
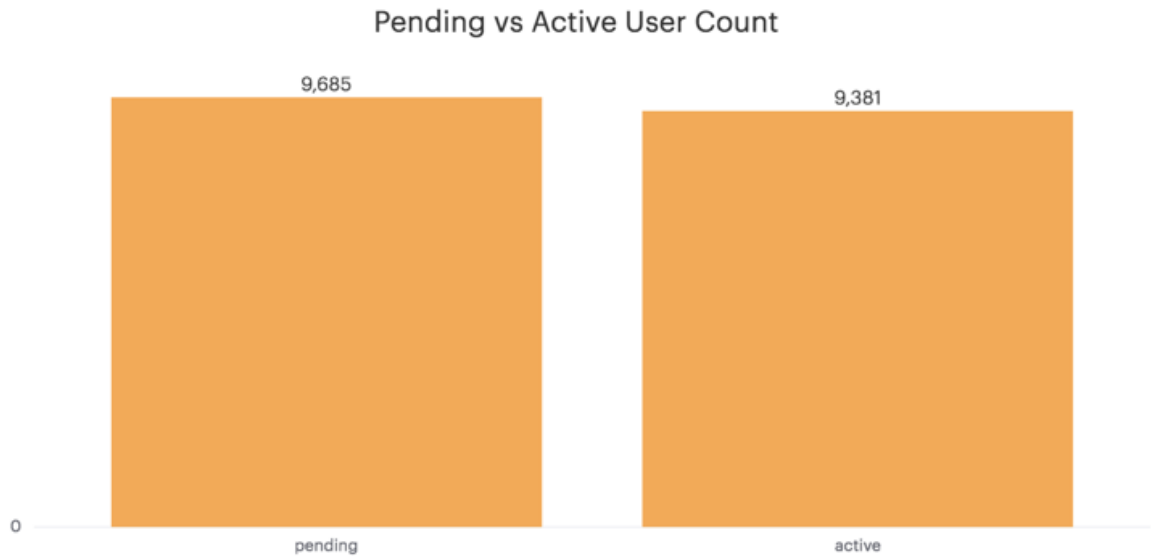
SQL code:
-- Understanding the total Pending vs Active User Count
SELECT state, count(user_id) as User_Count
FROM tutorial.yammer_users
GROUP BY state
ORDER BY count(user_id) desc

Pending vs Active User Count

Next, let's take a look at where the active users live. It appears that most of the users live in the United States, meaning that August would be a summer holiday month (vs, say, Australia which would be more likely in Jan/Feb).

SQL Code:
-- Confirming that engagement events only arise from active users
SELECT events.location, count(users.user_id)
FROM tutorial.yammer_events events
JOIN tutorial.yammer_users users
ON events.user_id = users.user_id
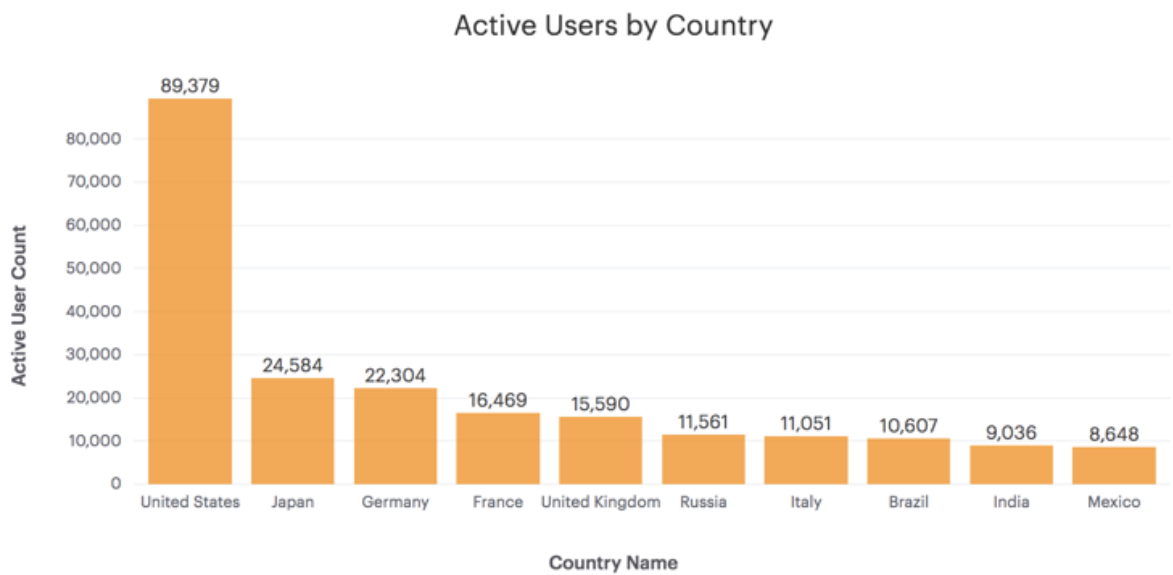WHERE state = 'active' AND event_type = 'engagement'
GROUP BY location
ORDER BY count(users.user_id) desc
LIMIT 10

Active Users by Country
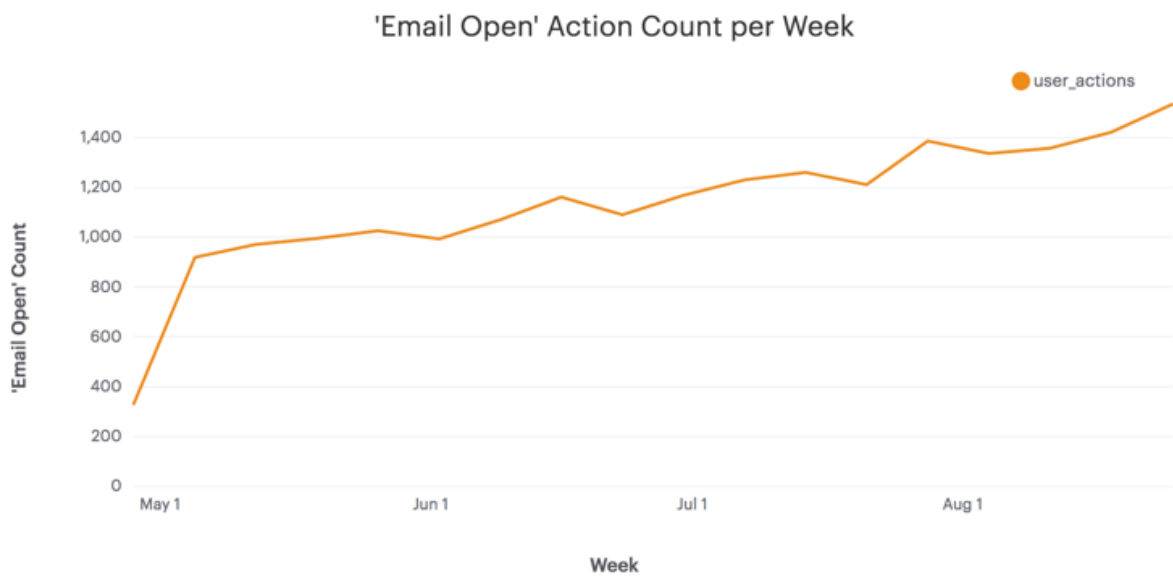

Next, let's look at user activity by week. Including email open.

SQL Code:
```
-- Email Open Action count by Week
SELECT DATE_TRUNC('week', occurred_at) as week_date, count(user_id) as
user_actions
FROM tutorial.yammer_emails emails
WHERE action = 'email_open'
GROUP BY DATE_TRUNC('week', occurred_at)
ORDER BY DATE_TRUNC('week', occurred_at) asc
```

'Email Open' Action Count per Week

Hmmm… it appears that email open activity is on the rise, this might disprove the hypothesis that a large group of users took vacation in August.

Hypothesis #2: there may have been a systems issue during the early August timeframe.

In analyzing the other email events, it appears that while emails sent and opened increased consistently, the "user_clickthrough" event seemed to dip during the same timeframe as the dip in engagement.
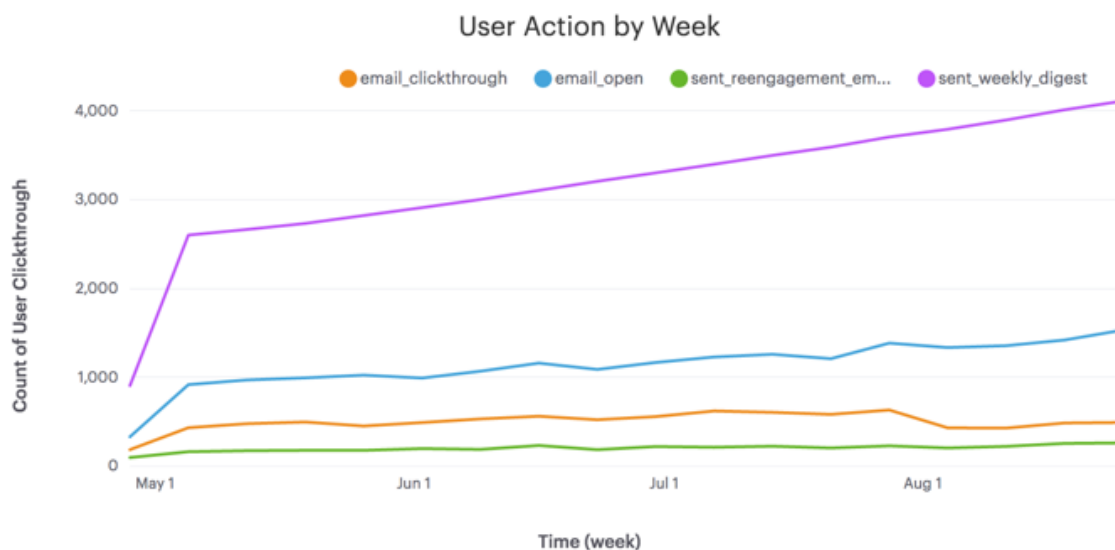
SQL Code:
```
-- User Action by Week
SELECT DATE_TRUNC('week', occurred_at) as week_date, action, count(user_id) as user_actions
FROM tutorial.yammer_emails emails
GROUP BY DATE_TRUNC('week', occurred_at), action
ORDER BY DATE_TRUNC('week', occurred_at) asc
```
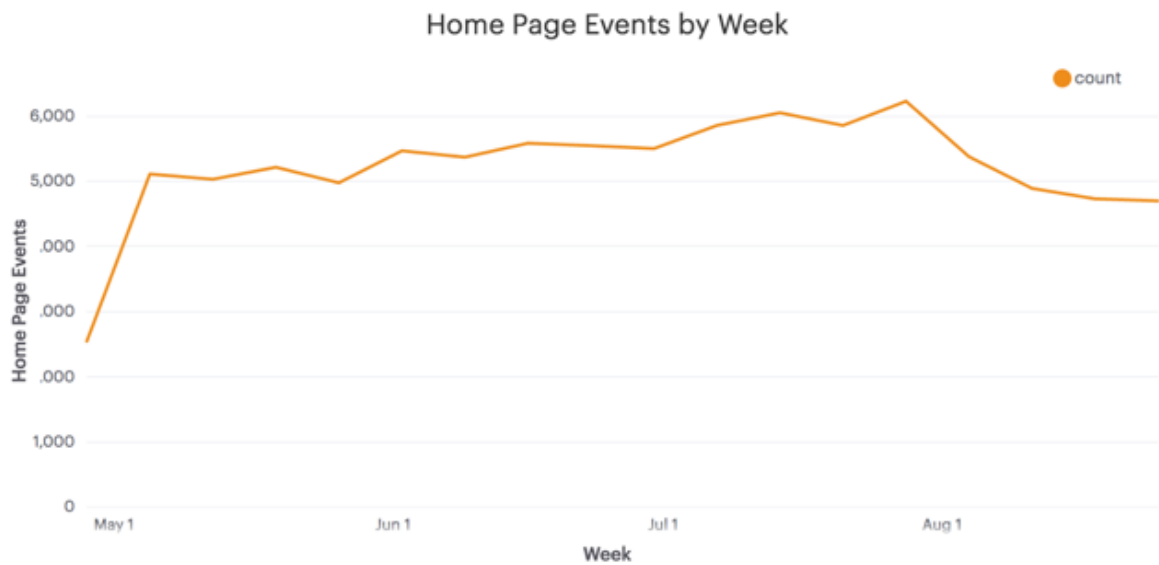

User Action by Week

Count of User Clickthrough vs Time

In analyzing specific events closer, it appears that the home page events follow a similar trend as click through. I wonder if there may be something related to the email link associated with the home page.

SQL Code:
```
-- Event Type by Week
SELECT DATE_TRUNC('week', occurred_at) as week_date, event_name,
count(user_id)
FROM tutorial.yammer_events events
WHERE event_type = 'engagement'
GROUP BY DATE_TRUNC('week', occurred_at), event_name
ORDER BY DATE_TRUNC('week', occurred_at) asc
```
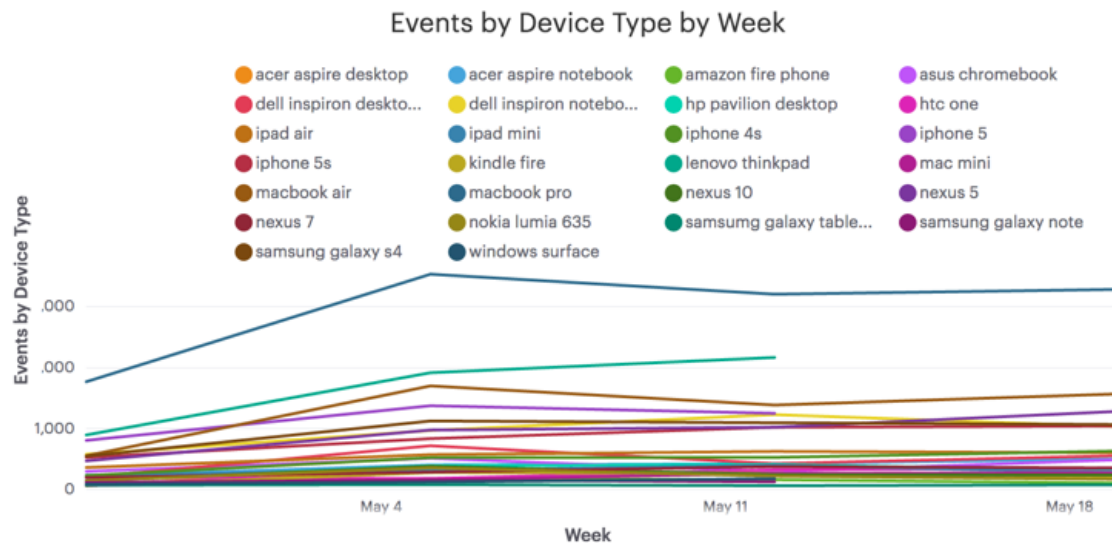
Home Page Events by Week

There may also be something associated with device type, as it appears that the mobile
devices may be having more of an impact in the recent weeks.

SQL Code:

```
-- Events by Device Type by Week
SELECT DATE_TRUNC('week', occurred_at) as week_date, device, count(user_id)
FROM tutorial.yammer_events events
GROUP BY DATE_TRUNC('week', occurred_at), device
ORDER BY DATE_TRUNC('week', occurred_at) asc
```

Events by Device Type by Week

Legend:
- acer aspire desktop
- acer aspire notebook
- amazon fire phone
- asus chromebook
- dell inspiron deskto...
- dell inspiron notebo...
- hp pavilion desktop
- htc one
- ipad air
- ipad mini
- iphone 4s
- iphone 5
- iphone 5s
- kindle fire
- lenovo thinkpad
- mac mini
- macbook air
- macbook pro
- nexus 10
- nexus 5
- nexus 7
- nokia lumia 635
- samsumg galaxy table...
- samsung galaxy note
- samsung galaxy s4
- windows surface

Making a recommendation

Start to work your way through your list of hypotheses in order to determine the source of the drop in engagement. As you explore, make sure to save your work. It may be helpful to start with the code that produces the above query, which you can find by clicking the link in the footer of the chart and navigating to the "query" tab.

Answers:

- Do the answers to any of your original hypotheses lead you to further questions?
    - yes, it appears that there may be an issue with the email blasts that are being sent, and possibly how they are read on the mobile app. I would want to know if there have been any product updates in the early August timeframe.
- If so, what are they and how will you test them?
    - N/A, would deal with talking with the product team.
- If they are questions that you can't answer using data alone, how would you go about answering them (hypothetically, assuming you actually worked at this company)?
    - it appears that the issue lies in the email clickthrough rates and potentially the mobile app. Ask the head of product if there have been any email template changes, or product changes on the mobile app.
- What seems like the most likely cause of the engagement dip?
    - it appears that the issue lies in the email clickthrough rates and potentially the mobile app.
- What, if anything, should the company do in response?
    - Investigate the email template and links. Same for the mobile app. I can't say that it is 100% the cause, but likely we will find something that could

be improved (e.g. layout, UX) or there may be something recently
implemented that needs to be fixed.