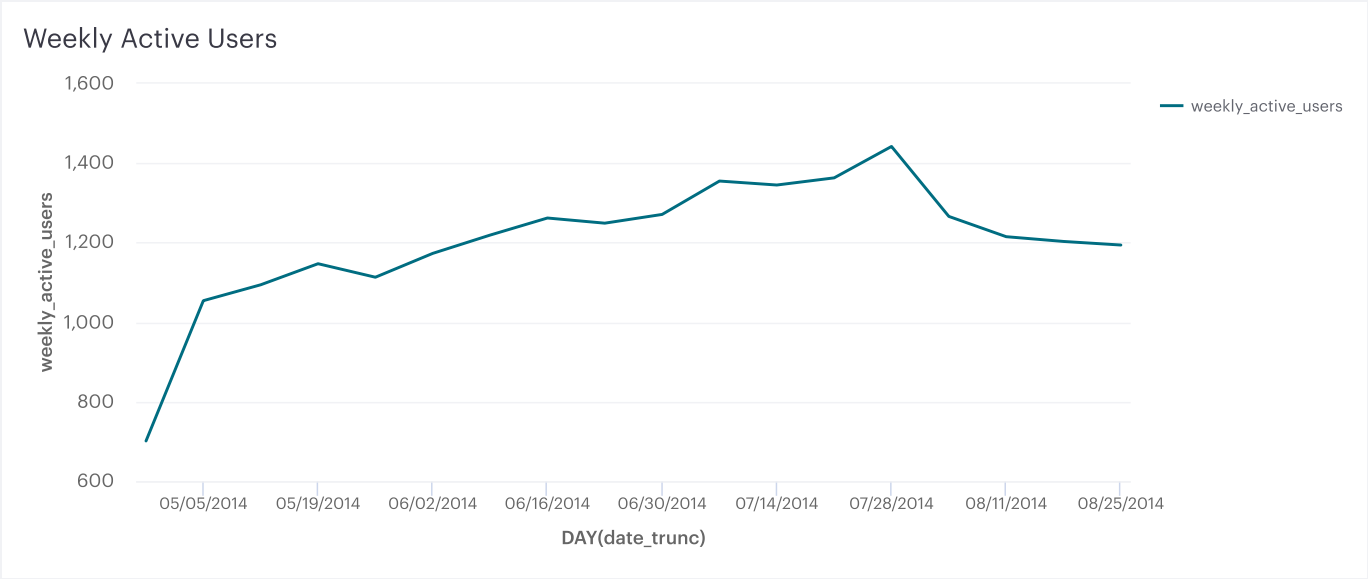
Untitled Report

SQL_YAMMER_CASE_STUDY

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:



Query:

```
SELECT DATE_TRUNC('week', e.occurred_at),
       COUNT(DISTINCT e.user_id) AS weekly_active_users
  FROM tutorial.yammer_events e
 WHERE e.event_type = 'engagement'
   AND e.event_name = 'login'
 GROUP BY 1
 ORDER BY 1
```

Possible Causes

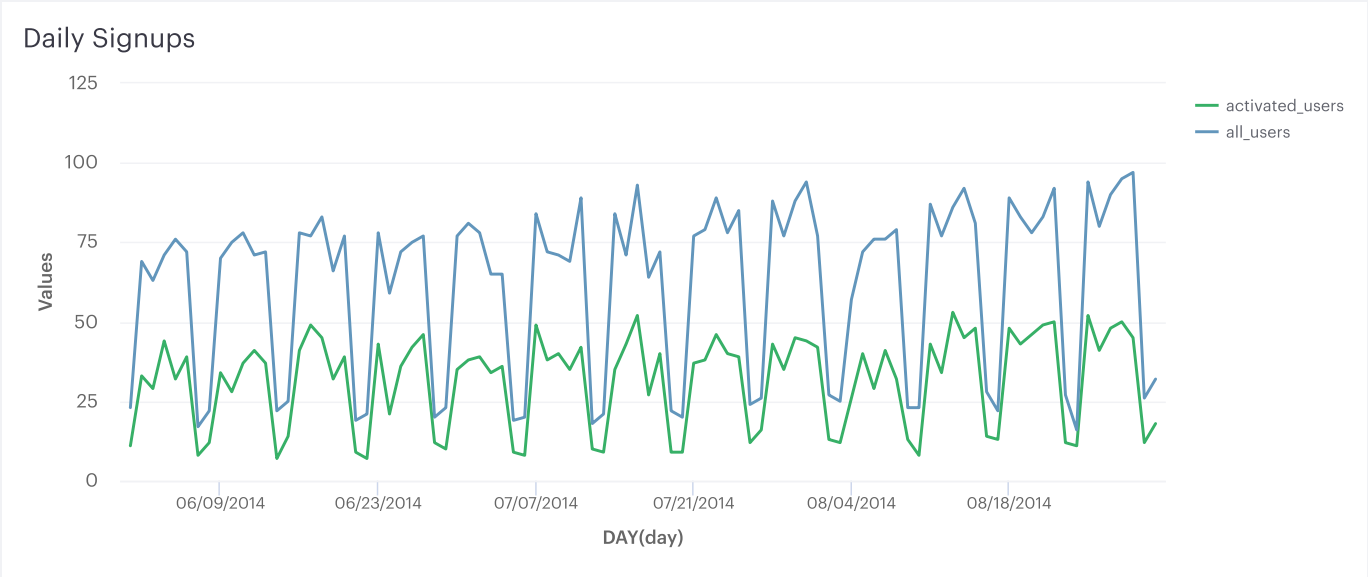
Merely by watching the line, three possible causes jump into my mind that could be the reason of this drop:

1. **Signing Up Rate Decrease**: the increase of weekly active users is actually a byproduct of the increase of new users. If signing up rate is dropping off, then the number of weekly active users should be. As for the reasons of this dropping off, it is totally possible that in the past few months, big marketing events happened which attracted lots of customers and resulted in a more than average increase, it is also likely that in August many companies enter their vacation month, both of them will make this drop looks like abnormal. To check if this is the case, information from marketing department is needed.
2. **Loss of Key Account**: if one of our heavy buyers (usually big companies) stopped using our service, their will be a significant drop.
3. **Function Problem**: there may be some function problem of our product that disappointed our customers and drive them away.

Hopefully our data would give us a hint.

Solving the case

Daily Signups

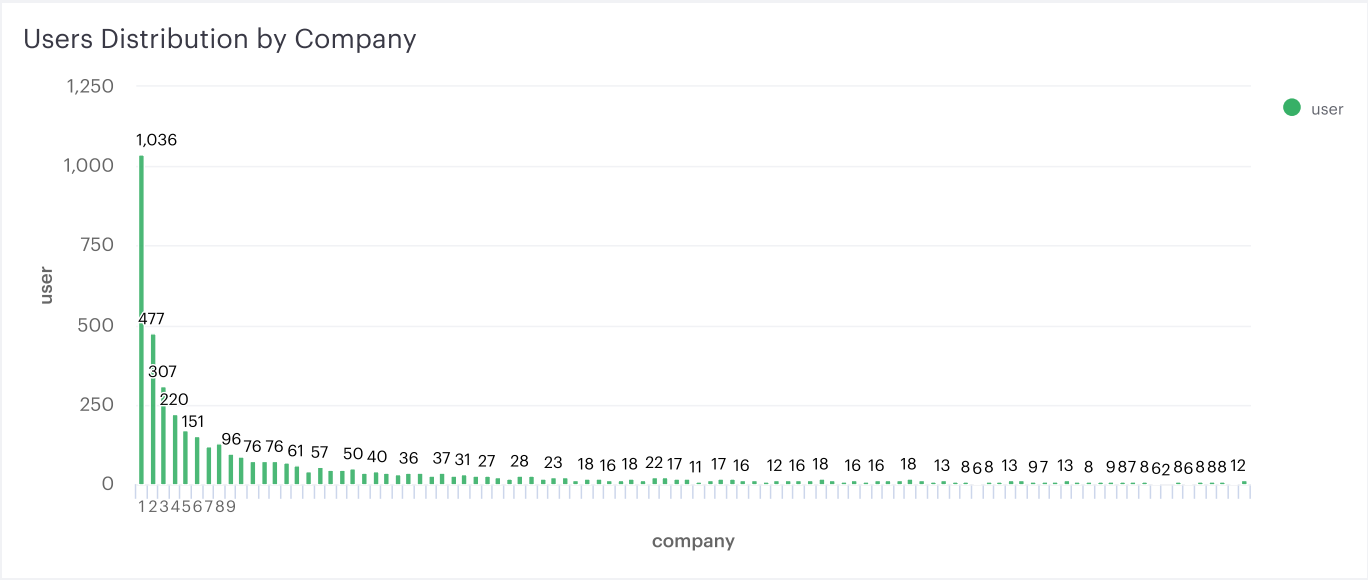


Query:

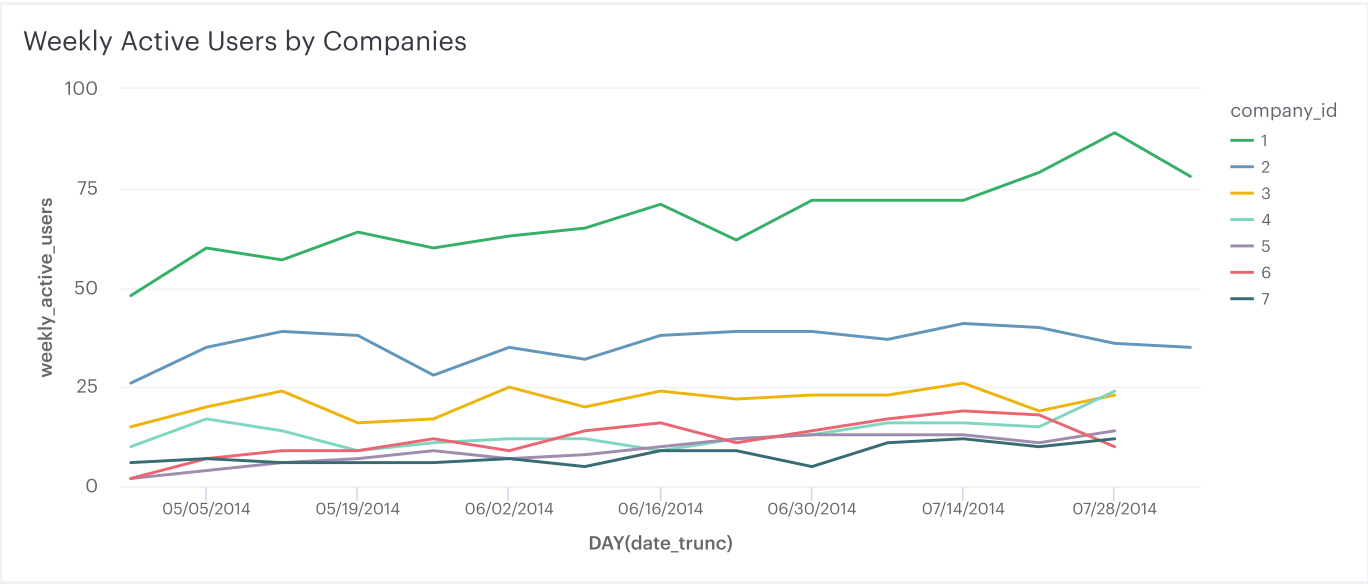
```
SELECT DATE_TRUNC('day',created_at) AS day,
       COUNT(*) AS all_users,
       COUNT(CASE WHEN activated_at IS NOT NULL THEN u.user_id ELSE NULL END) AS activated_users
  FROM tutorial.yammer_users u
 WHERE created_at >= '2014-06-01'
   AND created_at < '2014-09-01'
 GROUP BY 1
 ORDER BY 1
```

Read from the chart, it seems there is nothing abnormal with the increase rate of new users. So the drop must come from our regular users. Next, I am going to check my second hypothesis, that some big companies quit using our product.

Loss of Key Account



This distribution tells us that Yammer does have several key accounts. Since the drop of weekly active users is at a level of 200 around, the following analysis will focus only on the first 7 companies which each has over 100 Yammer users



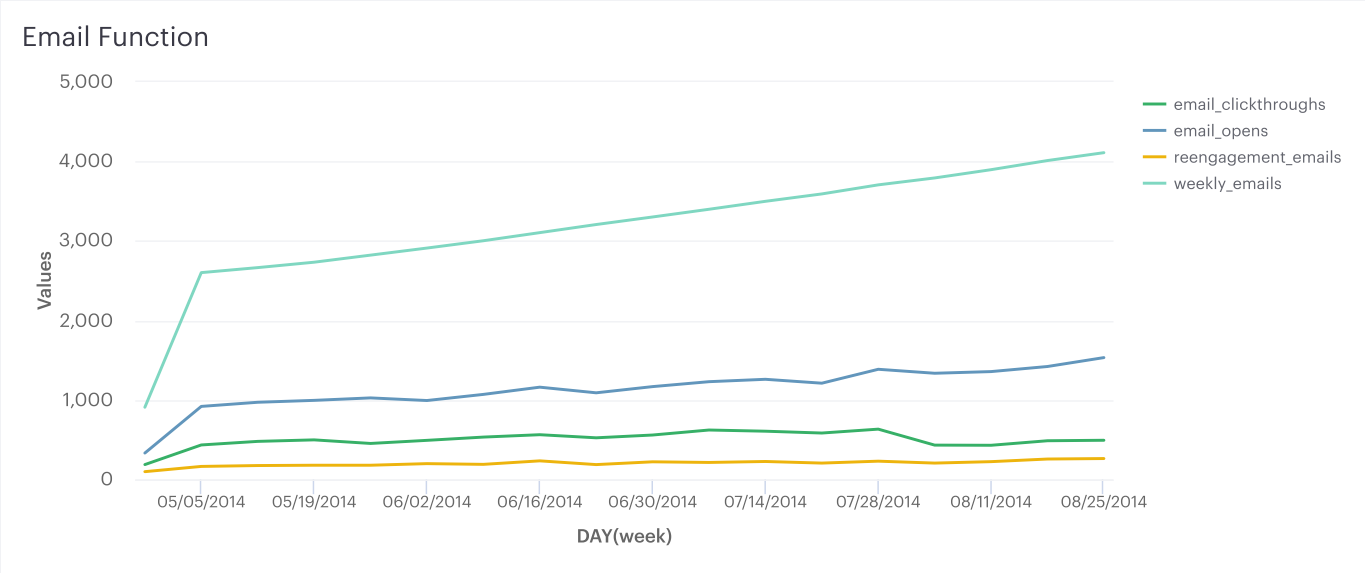
Query:

```
SELECT u.company_id,
    DATE_TRUNC('week', e.occurred_at),
    COUNT(DISTINCT e.user_id) AS weekly_active_users
FROM tutorial.yammer_events e
   LEFT JOIN tutorial.yammer_users u
   ON e.user_id = u.user_id
     WHERE e.event_type = 'engagement'
      AND e.event_name = 'login'
      AND u.company_id IN (1,2,3,4,5,6,7)
 GROUP BY 1,2
 ORDER BY 2
```

It is pretty clear from the chart above that Yammer is losing customers from key accounts. For company 1, which is the key account with most subscribers, the active users are dropping. For company from 3 to 7, we are facing a problem of missing data. Without any clues why there is no records of this 5 companies after date 07/28, it is totally possible that these companies stopped using our product, which is a important reason that causes the active users drop in general.

Email Function

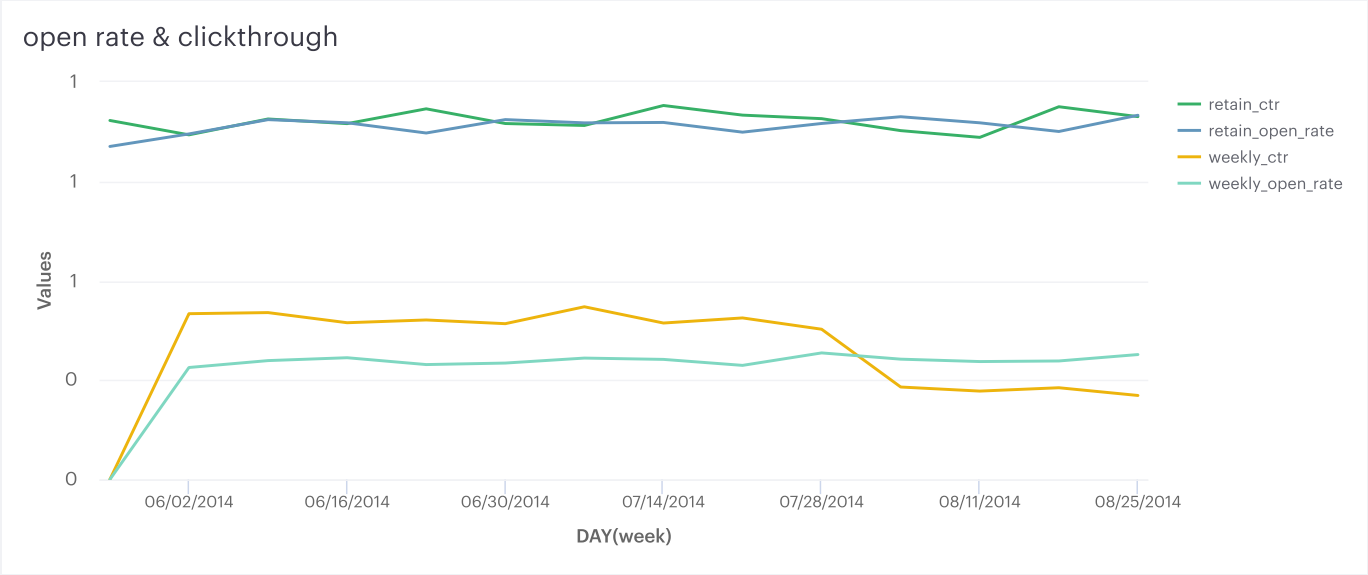
Besides loss of key account, could our function be a problem as well? Let's check our email function at first since this is the most commonly used one.



Query:

```
SELECT DATE_TRUNC('week', occurred_at) AS week,
       COUNT(CASE WHEN e.action = 'sent_weekly_digest' THEN e.user_id ELSE NULL END) AS weekly_emails,
       COUNT(CASE WHEN e.action = 'sent_reengagement_email' THEN e.user_id ELSE NULL END) AS reengagement_emails,
       COUNT(CASE WHEN e.action = 'email_open' THEN e.user_id ELSE NULL END) AS email_opens,
       COUNT(CASE WHEN e.action = 'email_clickthrough' THEN e.user_id ELSE NULL END) AS email_clickthroughs
  FROM tutorial.yammer_emails e
 GROUP BY 1
 ORDER BY 1
```

It is worth noting that the click through rate drops dramatically within the time window of active users' drop. Let's look into more details:



Query:

```
SELECT week,
       weekly_opens/CASE WHEN weekly_emails = 0 THEN 1 ELSE weekly_emails END::FLOAT AS weekly_open_rate,
       weekly_ctr/CASE WHEN weekly_opens = 0 THEN 1 ELSE weekly_opens END::FLOAT AS weekly_ctr,
       retain_opens/CASE WHEN retain_emails = 0 THEN 1 ELSE retain_emails END::FLOAT AS retain_open_rate,
       retain_ctr/CASE WHEN retain_opens = 0 THEN 1 ELSE retain_opens END::FLOAT AS retain_ctr
  FROM (
SELECT DATE_TRUNC('week',e1.occurred_at) AS week,
       COUNT(CASE WHEN e1.action = 'sent_weekly_digest' THEN e1.user_id ELSE NULL END) AS weekly_emails,
       COUNT(CASE WHEN e1.action = 'sent_weekly_digest' THEN e2.user_id ELSE NULL END) AS weekly_opens,
       COUNT(CASE WHEN e1.action = 'sent_weekly_digest' THEN e3.user_id ELSE NULL END) AS weekly_ctr,
       COUNT(CASE WHEN e1.action = 'sent_reengagement_email' THEN e1.user_id ELSE NULL END) AS retain_emails,
       COUNT(CASE WHEN e1.action = 'sent_reengagement_email' THEN e2.user_id ELSE NULL END) AS retain_opens,
       COUNT(CASE WHEN e1.action = 'sent_reengagement_email' THEN e3.user_id ELSE NULL END) AS retain_ctr
  FROM tutorial.yammer_emails e1
  LEFT JOIN tutorial.yammer_emails e2
    ON e2.occurred_at >= e1.occurred_at
   AND e2.occurred_at < e1.occurred_at + INTERVAL '5 MINUTE'
   AND e2.user_id = e1.user_id
   AND e2.action = 'email_open'
  LEFT JOIN tutorial.yammer_emails e3
    ON e3.occurred_at >= e2.occurred_at
   AND e3.occurred_at < e2.occurred_at + INTERVAL '5 MINUTE'
   AND e3.user_id = e2.user_id
   AND e3.action = 'email_clickthrough'
 WHERE e1.occurred_at >= '2014-06-01'
   AND e1.occurred_at < '2014-09-01'
   AND e1.action IN ('sent_weekly_digest','sent_reengagement_email')
 GROUP BY 1
       ) a
 ORDER BY 1
```

Now Yammer team can be pretty sure that there are something wrong with the email function, which we should definitely look into.

Conclusion

Based on the analysis above, the email function faculty and loss of key accounts are the two potential cause that I found for this weekly active users drop. This two causes may be related to each other, perhaps those key accounts quit because they were disappointed of our email functions. To solve this, product team should be involved in to detect and update our email features. Marketing team need to work together as well finding out why we are losing key accounts.