Untitled Report

Analysis of Users Engagement Drop

The report describes the analysis done to understand the reasons behind the drop in Yammer's user activity during July- August 2014.

Problem: There is a drop in Yammer's user activity. Investigate is needed to find out the possible reasons behind the decline and come up with solution proposals if possible (Yammer is a social network for communicating with coworkers. Individuals share documents, updates, and ideas by posting them in groups).

The Solutions:
The following hypothesis are checked to find out possible reasons leading to the engagment dip:
1. *Users Activity by Country:* The trend will indicate if there are holidays, summer vocations and/or Internet connectivity issues that have caused the problem.
2. *User Activity by Month:* The trend will show if there is sudden influx of new users enrollment due to an marketing event and users are not using the account after signing up
3. *User Activity by Device or Device Type:* The trend will indicate if the engagement dip is related to devices like computers, phones or tablets
4. *User Activity by Activity Type*: The trend will indicates if the engagement dip is related to specific activity

First, confirming that indeed engagement has dropped. From last few days of July thr first week of Aug.

Query 1: Problem in The Engagement Dip
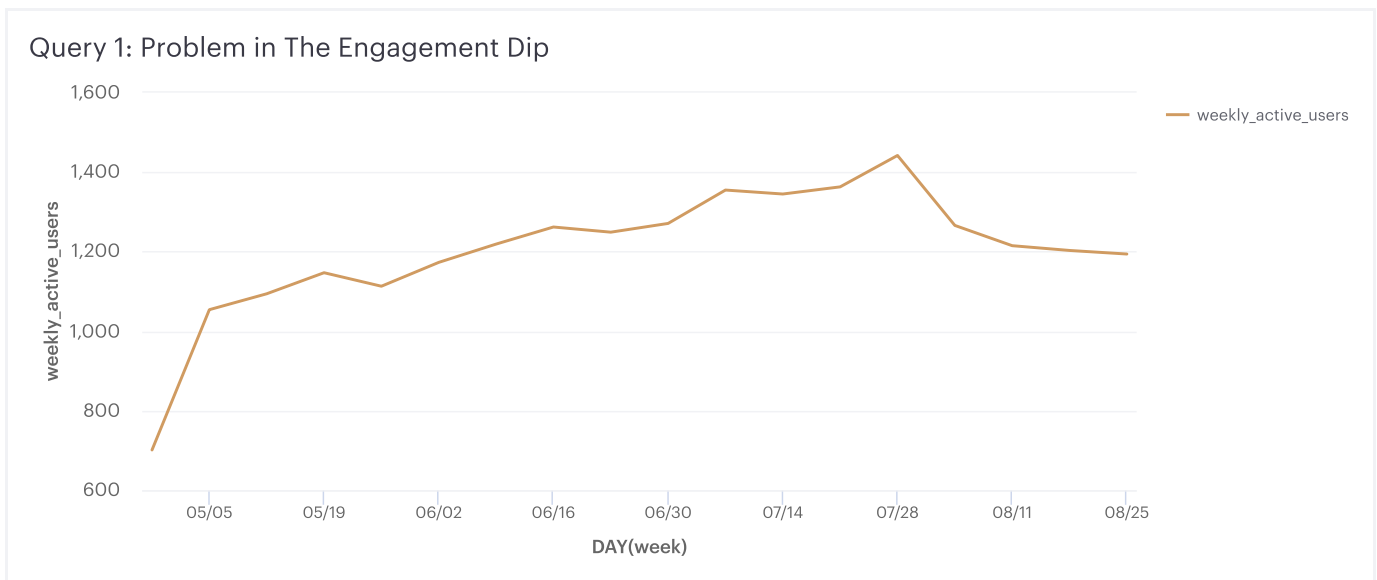
Let's examine user activity by country to see if there is a trend that indicates if there are holidays or internet connectivity issues in some part of the world.
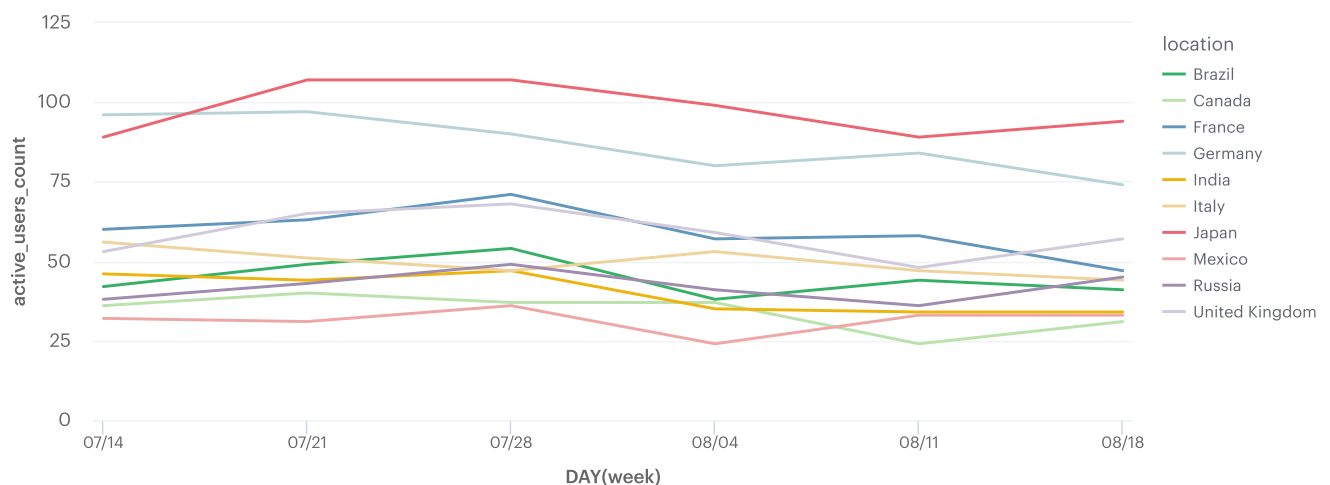
Query 2: Active Users By Country - Top 10
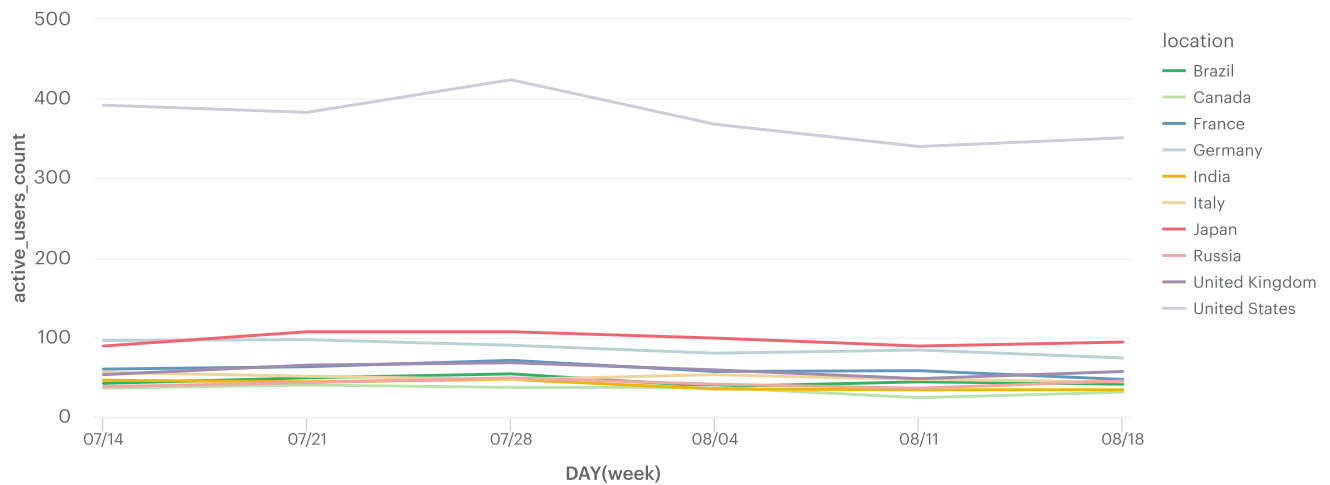
United States has a big share of active users and clearly there is a fall on user activity in period of 28-Jul to 04-Aug.

Let's remove United States from the query and better examine if there is similar pattern among other countries.

## Query 3: Users By Country (Exec. US) - Top 10

Result: All countries except Italy are displaying similar pattern like U.S in user activity for the same period. This rules-out the hypothesis of major festive or holidays season and/or internet connectivity which might have brought down the number of users.

To test if there is spike in new user enrollment, let's look at activated users by month.
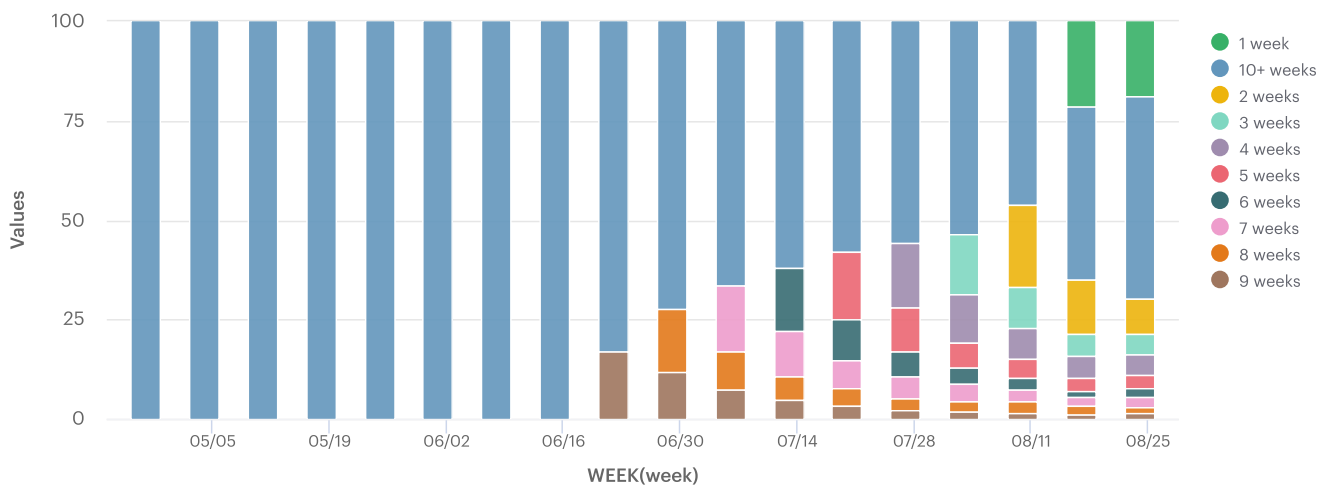
Query 4: Activated Users by Month
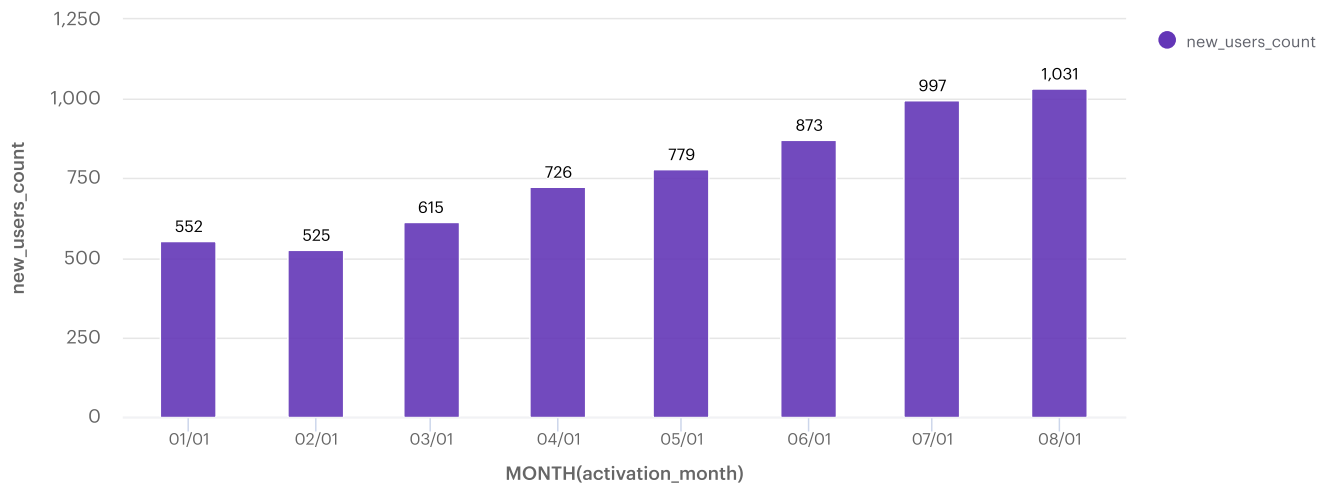
Result: There is clear evidence that there is steadily increase in enrollment from January to August so we can rule out possible spike in enrollment.

For curious reason, is there a difference in engagement activity from existing old users vs new users?

Query 5: Engagement by User Account Age

From the cohort users based on when they signed up the product, there is a decrease in engagement among new and old user across time. And this makes sense as new users would use the account more often but eventually wears off overtime. This also suggests that there is a very specific event happened in 28-Jul to 04-Aug that is causing the engagement dip rather than users quitting the platform. There could be an issue in accessing the Yammer platform, logging events, or broken product features.

I'll drill down into engagement by device type.

Query 6: Active Users by Device Type
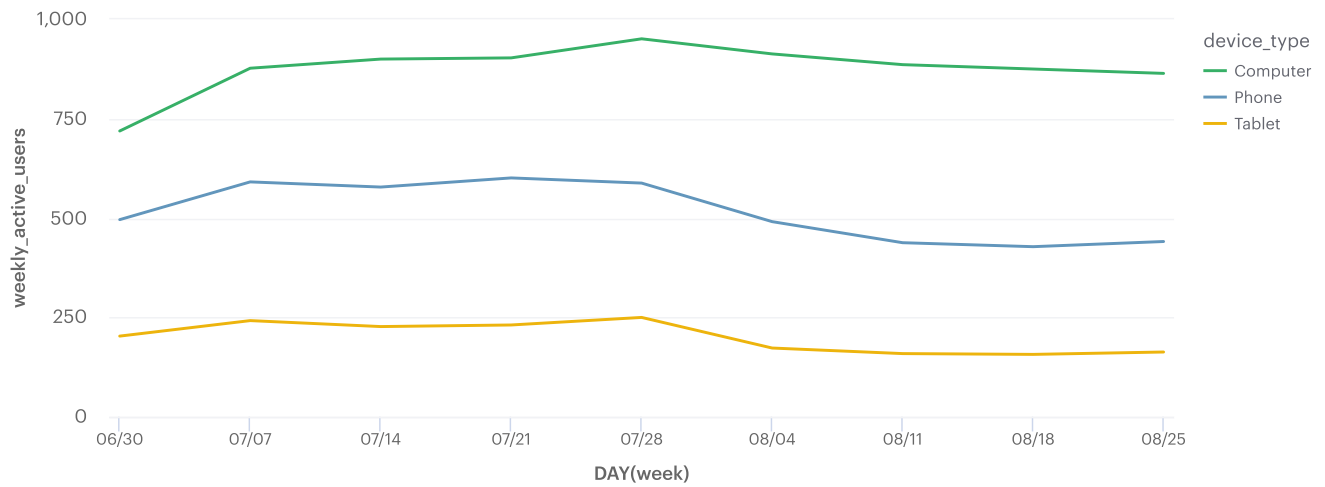
Active Users Per Device Type

week	device_type			
	Computer	Phone	Tablet	Totals
2014-06-30T00:00:00.000Z	719	496	203	**1418**
2014-07-07T00:00:00.000Z	877	591	242	**1710**
2014-07-14T00:00:00.000Z	900	578	227	**1705**
2014-07-21T00:00:00.000Z	903	601	231	**1735**
2014-07-28T00:00:00.000Z	951	588	250	**1789**
2014-08-04T00:00:00.000Z	913	491	173	**1577**
2014-08-11T00:00:00.000Z	886	438	159	**1483**
2014-08-18T00:00:00.000Z	875	428	157	**1460**
2014-08-25T00:00:00.000Z	864	441	163	**1468**
Totals	**7888**	**4652**	**1805**	**14345**

Percentage Change in Active User by Device Type

week	device_type			
	Computer	Phone	Tablet	Totals
2014-06-30T00:00:00.000Z	0	0	0	**0**
2014-07-07T00:00:00.000Z	18	16	16	**50**
2014-07-14T00:00:00.000Z	2	-2	-6	**-6**
2014-07-21T00:00:00.000Z	0	3	1	**4**
2014-07-28T00:00:00.000Z	5	-2	7	**10**
2014-08-04T00:00:00.000Z	-4	-19	-44	**-67**
2014-08-11T00:00:00.000Z	-3	-12	-8	**-23**
2014-08-18T00:00:00.000Z	-1	-2	-1	**-4**
2014-08-25T00:00:00.000Z	-1	2	3	**4**
Totals	**16**	**-16**	**-32**	**-32**

Focusing on July 28 to Aug 04,
The drop in active user by device type is Computer: -4 , Phone: -19 , Tablet: -44

It looks like engagement is down across all devices, but the problem is localized in phone and tablet category. Maybe there're product/feature changes related to phone and tablet introduced on July 28 that leads to the dip in user activity. An discussion with PM is needed to confirm this.

I'll look at activity type and email action log to further narrowing down the problem

Query 7 - Percentage Change in Active Users per Activity Type

week	event_name						
	email_clickthrough	email_open	home_page	like_message	login	send_message	sent_reengagement_
2014-06-30T00:00:00.000Z	0	0	0	0	0	0	
2014-07-07T00:00:00.000Z	31	29	16	17	15	18	
2014-07-14T00:00:00.000Z	-2	2	-2	1	0	0	
2014-07-21T00:00:00.000Z	-4	-3	0	0	1	-1	
2014-07-28T00:00:00.000Z	8	12	7	3	5	6	
2014-08-04T00:00:00.000Z	-45	-3	-14	-9	-13	-12	
2014-08-11T00:00:00.000Z	0	0	-5	-8	-4	-8	
2014-08-18T00:00:00.000Z	11	4	0	0	0	1	
2014-08-25T00:00:00.000Z	0	6	0	-5	0	0	
Totals	**-1**	**47**	**2**	**-1**	**4**	**4**	

Again, focusing on the period of 28-Jul to 04-Aug, users have decreased in all activities.

Combining all evidences, I have found the issue:
  1. Email clickthroughs are down since the end of July but number of user opening email is still up. This suggests an issue in the email that being sent. Even though users were opening the email but with the invalid email link, it failed to direct users to Yammer platform.
  2. The email issue is greater in mobile and tablet users that suggest an issue with the app

These findings should be discussed with PM for validation and remedy actions. In addition recommendations are
  ● Investigate the emails that have gone out since 28-Jul
  ● Talk to PM to find out what recent changes were made to the app

SQL QUERIES:
Query 1:
SELECT DATE_TRUNC('week', occurred_at) AS Week,
    COUNT(DISTINCT user_id) AS weekly_active_users
FROM tutorial.yammer_events
WHERE event_type = 'engagement'
 AND event_name = 'login'
GROUP BY Week
ORDER BY Week
Query 2:
/*active users count by top 10 countries during the engagement dip period*/
SELECT Date_Trunc('week', occurred_at) AS Week, location,

```
COUNT(DISTINCT user_id) AS Active_users_count
FROM tutorial.yammer_events
WHERE event_type = 'engagement'
 AND event_name = 'login'
 AND occurred_at >= '2014-07-14'
 AND occurred_at < '2014-08-25'
 AND location IN
  (/*get top 10 countries based on active users count*/
  ( SELECT location
   FROM tutorial.yammer_events
   WHERE event_type = 'engagement'
   GROUP BY 1
   ORDER BY COUNT(DISTINCT user_id) DESC LIMIT 10)
GROUP BY 1, 2
ORDER BY 1, 2)
```

Query 3:
```
/*excluding USA : active users count by top 10 countries during the engagement dip period 07/14 to 8/25 */
SELECT Date_Trunc('week', occurred_at) AS Week, location,
    COUNT(DISTINCT user_id) AS Active_Users_count
FROM tutorial.yammer_events
WHERE event_type = 'engagement'
 AND event_name = 'login'
 AND occurred_at >= '2014-07-14'
 AND occurred_at < '2014-08-25'
 AND location IN
  (SELECT location
   FROM tutorial.yammer_events
   WHERE location <> 'United States'
    AND event_type = 'engagement'
   GROUP BY location
   ORDER BY Count(DISTINCT user_id) DESC LIMIT 10)
GROUP BY 1, 2
ORDER BY 1, 2)
```

Query 4:
```
SELECT Date_TRUNC('month', activated_at) AS Activation_Month,
    COUNT(user_id) AS New_Users_Count
FROM tutorial.yammer_users
WHERE activated_at IS NOT NULL
 AND Created_at >= '2014-01-01'
GROUP BY 1
ORDER BY 1
```

Query 5:
```
SELECT DATE_TRUNC('week',z.occurred_at) AS "week",
   AVG(z.age_at_event) AS "Average age during week",
   COUNT(DISTINCT CASE WHEN z.user_age > 70 THEN z.user_id ELSE NULL END) AS "10+ weeks",
   COUNT(DISTINCT CASE WHEN z.user_age < 70 AND z.user_age >= 63 THEN z.user_id ELSE NULL END) AS "9
weeks",
   COUNT(DISTINCT CASE WHEN z.user_age < 63 AND z.user_age >= 56 THEN z.user_id ELSE NULL END) AS "8
weeks",
```

```
    COUNT(DISTINCT CASE WHEN z.user_age < 56 AND z.user_age >= 49 THEN z.user_id ELSE NULL END) AS "7
weeks",
    COUNT(DISTINCT CASE WHEN z.user_age < 49 AND z.user_age >= 42 THEN z.user_id ELSE NULL END) AS "6
weeks",
    COUNT(DISTINCT CASE WHEN z.user_age < 42 AND z.user_age >= 35 THEN z.user_id ELSE NULL END) AS "5
weeks",
    COUNT(DISTINCT CASE WHEN z.user_age < 35 AND z.user_age >= 28 THEN z.user_id ELSE NULL END) AS "4
weeks",
    COUNT(DISTINCT CASE WHEN z.user_age < 28 AND z.user_age >= 21 THEN z.user_id ELSE NULL END) AS "3 weeks",
    COUNT(DISTINCT CASE WHEN z.user_age < 21 AND z.user_age >= 14 THEN z.user_id ELSE NULL END) AS "2 weeks",
    COUNT(DISTINCT CASE WHEN z.user_age < 14 AND z.user_age >= 7 THEN z.user_id ELSE NULL END) AS "1 week",
    COUNT(DISTINCT CASE WHEN z.user_age < 7 THEN z.user_id ELSE NULL END) AS "Less than a week"
 FROM (
   SELECT e.occurred_at,
      u.user_id,
      DATE_TRUNC('week',u.activated_at) AS activation_week,
      EXTRACT('day' FROM e.occurred_at - u.activated_at) AS age_at_event,
      EXTRACT('day' FROM '2014-09-01'::TIMESTAMP - u.activated_at) AS user_age
    FROM tutorial.yammer_users u
    JOIN tutorial.yammer_events e
    ON e.user_id = u.user_id
    AND e.event_type = 'engagement'
    AND e.event_name = 'login'
    AND e.occurred_at >= '2014-05-01'
    AND e.occurred_at < '2014-09-01'
   WHERE u.activated_at IS NOT NULL
   ) z
 GROUP BY 1
 ORDER BY 1
```

Query 6:

```
SELECT a.*,
lag(weekly_active_users) OVER (PARTITION BY device_type ORDER BY week, device_type) AS rows,
weekly_active_users - lag(weekly_active_users) OVER (PARTITION BY device_type ORDER BY week, device_type) as
diff,
(weekly_active_users - lag(weekly_active_users) OVER( PARTITION BY device_type ORDER BY week, device_type)) *
100/ weekly_active_users AS pct_change
FROM
 (SELECT
   DATE_TRUNC('week', occurred_at) as week,
   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 'Computer'
   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 'Phone'
   WHEN device IN ('ipad air','nexus 7','ipad mini','nexus 10','kindle fire','windows surface',
   'samsumg galaxy tablet') THEN 'Tablet' ELSE NULL END AS device_type,
   COUNT(DISTINCT user_id) AS weekly_active_users
   FROM tutorial.yammer_events
   WHERE event_type = 'engagement'
   AND event_name = 'login'
```

```
    AND occurred_at >= '2014-07-01'
    AND occurred_at <'2014-09-01'
    GROUP BY 1,2) a
 Query 7:
/* active users per activity type*/
SELECT a.*,
lag(weekly_active_users) OVER (PARTITION BY event_name ORDER BY week, event_name),
 weekly_active_users - lag(weekly_active_users) OVER (PARTITION BY event_name ORDER BY week, event_name) AS
diff,
(weekly_active_users - lag(weekly_active_users) OVER (PARTITION BY event_name ORDER BY week, event_name))
*100/ weekly_active_users AS pct_change
FROM
(SELECT event_name,
DATE_TRUNC('week', occurred_at) AS week,
COUNT(DISTINCT user_id) AS weekly_active_users
FROM tutorial.yammer_events
WHERE event_type = 'engagement'
AND event_name NOT LIKE '%search%'
AND occurred_at >= '2014-07-01'
AND occurred_at < '2014-09-01'
group by 1,2
UNION ALL
SELECT action,
DATE_TRUNC('week', occurred_at) As week,
COUNT(DISTINCT user_id)
FROM tutorial.yammer_emails
WHERE occurred_at >='2014-07-01'
AND occurred_at < '2014-09-01'
group by 1,2)a
```