1. Background

Yammer (background):

Yammer is a social network for communicating with coworkers. Individuals share documents, updates, and ideas by posting them in groups.

Yammer has a centralized Analytics team, which sits in the Engineering organization. Their primary goal is to drive better product and business decisions using data. They do this partially by providing tools and education that make other teams within Yammer more effective at using data to make better decisions. They also perform ad-hoc analysis to support specific decisions.

Components of this Analysis:

1. Problem
2. Plan (Brainstorming Potential Causes)
3. Data
4. Analysis
5. Conclusion (Recommendation)
6. Further Quesitons

2. Analysis

2.1 Problem:

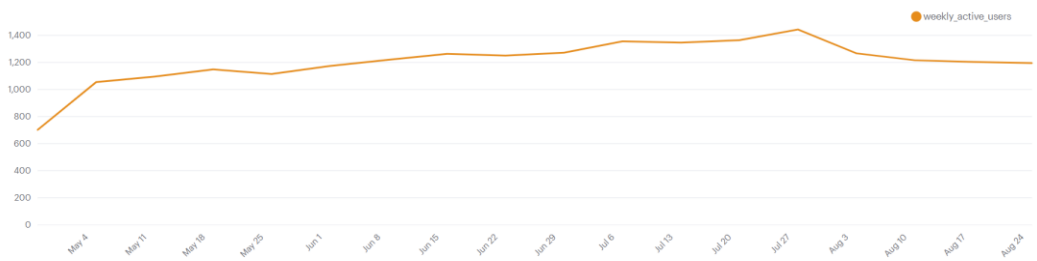
Tuesday morning, September 2, 2014.

what you think about the latest activity on the user engagement dashboards.

Engagement definition:

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."



2.2 Plan:

Step 1:

Define engagement: user made some type of server call by interacting with the product (shown in the data as events of type "engagement")

Step2:

Think about some possible causes of the drop in the engagement.

Hypothesis 1: Technical issue. (ask engineers)

- Broken traffic (the log was broken; infrastructure like data pipeline were down)
- Broken feature (some features of the app were broken, but they were hard to detect)
- Internet issue (maybe internet provider shut down their service)
- Search crawler change (somehow effect employee to interact with the event)

Hypothesis 2: Non-technical issue

- Holiday (break)

2.3 Data:

Table 1, Users: descriptive information about the users' accounts

Table 2, Events: an event is an action that a user has taken on Yammer. These events include login events, messaging events, search events, events logged as users progress through a signup funnel, events around received emails. (specific action the user took)

For event_type:

- Signup_flow
- engagement

For event_name

- Create_user: User is added to Yammer's database during signup process
- Enter_email: User begins the signup process by entering her email address
- Enter_info: User enters her name and personal information during signup process
- Complete_signup: User completes the entire signup/authentication process
- Home_page: User loads the home page

- Like_message
- Login
- Search_autocomplete: User selects a search result from the autocomplete list
- Search_run
- Search_click_result_X: User clicks search result X on the results page, where X is a number from 1 through 10
- Send_message
- View_inbox

Table 3, Email Events: This table contains events specific to the sending of emails. It is similar in structure to the events table above.

For action:

- Sent_weekly_digest: the user was delivered a digest email showing relevant conversations from the previous day.
- Email_open:
- Email_clickthrough: the user clicked a link in the email.
- Sent_reengagement_email

Table 4, Rollup Periods:

The final table is a lookup table that is used to create rolling time periods.

2.4 Analysis (dig in):

Step 0:

0.1 Scope the time interval for analysis: July to August (2014)

Step 1: Event-level Aggregate

Get more information about the details of engagement.

1. What events are categorized as engagement

2. See what's the difference in engagement between months
3. Count the number of occurrences of each 'engagement' event over month, and their change over month;

   Dive deep to see what happened in August. Identified events that drop. (Drop as compared to the last month, July)

What I found:

I saw that 'home_page' is the one that drops significantly. 'like_message' and 'view_inbox' follow.

Questions: What cause the drops in these categories?

Hypothesis:

1. **Technical issue: Broken feature**
2. **Drops in number of users.**

Step 2: User-level Aggregate

1. See if the drop in engagement was due to a drop in the number of users or the number of engagements per user.

   What I found:

   I found that number of users that participate engagement event decreases from 3058 to 2795 (July to Aug.) 8.6%

   The number of engagements per user decreases from 30 to 26 (July to Aug.)

   The number of events also drops from 91830 to 73793 (July to Aug.) (We know that this is due to 'home_page', 'like_message', 'view_inbox'…)

   Now, I have a question. Why are there drops in these categories?

Question 1: Which group of users are the main contributor to the drop in engagement in August?

Idea:

We can group by device, event_name, and location. But I think choosing location would be bias, as the majority works in the US. Instead, here I chose "device" as the group by criterion.

Method 1 (by device, one of the feature in the dataset):

Consider divide users into groups according to device type. Check which group of these users drop.

What I found:

After I run the query, I found that the number of users who use iphone 5 drop significantly in August. Other phone users and tablets users also show the similar pattern.

As compared to the above categories, the number of laptop and computer users did not show any significant drop.

Thus, I conclude that the drop in the number of users is mainly from those who use mobile devices. (Over 50%)

Question 2: Why mobile device?

Hypothesis:

1. **(Technical Aspect) There are some problems with the mobile devices' API.**
2. **(Non-technical aspect) There are some specific interactions with mobile device users that cause the drop.**

What I would do:

For case 1: I will go talk with engineer to check whether there is a technical problem happening in the backend.

For case 2: I will check with the data to see what are the interaction that causing the drop.

Question 3: What are the possible interactions users can have?

Idea:

Look at the data at hand, I saw that there is a Table of "email" (table 3). Email service could be possible interaction that users can take on device. I decided to investigate it.

What I did:

I check for email service and see how each action (sent_weekly_digest, sent_reengagement, email open, email_clickthrough) change over month.

What I found:

I found that there are 4 actions in email service, and only the number of "email_clickthrough" shows a drop (-729, 27% decrease) in August. Other action all showed an increasing pattern in August.

Conclusion:

Email clickthrough rate maybe the main contributor of the drop in users' engagement.

Question 4: Is the CLT rate somehow related to mobile device users?

Hypothesis:

Decrease in CLT rate & decrease in the number of mobile device user, these 2 things are correlated

What I did:

I checked for the change in "click_through" for each type of device (mobile device, laptop…) from July to August.

What I found:

The decrease in "click_through" rate mostly happened on mobile devices.

Conclusion:

It looks like my hypothesis is more convincing right now.

Question 5: Which type of emails account for the drop in action like "click_through"?

Define the domain:

There are 2 types of emails that sent to our users: “sent_weekly_digest”, and “sent_reengagement”

What I did:

Check for counts of “email_open” and “click_through” for each type of emails.

What I found:

The decrease in actions only happened in “sent_weekly_digest”.

Conclusion:

Users have less interests in weekly digested emails than in reengagement email (well, we confirm it right now).

2.5 Conclusion:

As we define the engagement as server calls made by our user to interact with the Yammer app, I found that there are 2 metrics that account for the drop in user engagement from July to August, 2014: one is the number of users that make server call action (event_type as “engagement”), and second is the email clickthrough rate. Considering those whose action are viewed as engagement, we can see that there is a sharp decrease in number of mobile device users. It accounts for over 50% of the total decrease in the drop of engagement. In addition, email clickthrough rate decrease by 27% from July to August, and we can find that the decrease is also mainly in mobile device. It seems that both metrics are somehow related to mobile device. Furthermore, I found that although users have increasing number in opening emails from both weekly digest type and reengagement type. The click through rate increased in reengagement email, but decrease in weekly digest emails.

3. Recommendation:

My recommendation is to immediately take a look into the weekly digest emails specially for mobile device like phones and tablets. I think it is very likely that there is a technical problem there, which makes our users difficult to interact with the email links.

4. Further Analysis:

- (location) Consider group users by location, and see if there are anything that we did not notice before.
- (cohort analysis) See if the change is attributed to a short user life-cycle