YAMMER

PRODUCT ANALYSIS PROJECT REPORT

Name: Alsatwar Sravan Kumar

alsatwarsravan@gmail.com

Table of Contents

1. INTRODUCTION

Analytics uses data and math to answer business questions, discover relationships, predict unknown outcomes, and automate decisions. Yammer, a social networking platform, faces product development problems; the team focuses on answering specific questions to discover novel relationships, solve the cases, and validate their testing results.

This project aims to solve those questions; Yammer specifically asked to solve three case problems related to their product.

- Drop-in User Engagement
- Understanding Search Functionality
- Validating A/B Test results

Yammer provided users data spanning four months to analyze the problem. This study used SQL programming, MySQL Workbench, and MS Excel. SQL (Structured Query Language) asks questions about data; MySQL runs the code, MS Excel, to draw the graphs and charts.

This report follows the pattern of a question, answer, validation and recommendation; At the start, it presents with clear problem statement with graphs/charts, generates the hypothesis of the problem, validates the answers using charts/graphs, concludes with a solution and recommends further steps to be taken.

1.1 About Yammer

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. 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. Yammer analysts are continuously trained to consider each project's value; they seek to maximize the return on their time. Analysts typically opt for less precise solutions to problems if it means investing substantially less time as well. This includes high-level decision making, like choosing which projects to prioritize. It also influences the way analysts think about metrics. Product decisions are always evaluated against core engagement, retention, and growth metrics in addition to product-specific usage metrics.

1.2 Problem Cases

1. A Drop in User Engagement
    - Engagement Dip – Figure out the source of the problem.
2. Understanding Search Functionality
    - Product Team thinks about revamping search – the job is to figure out whether the team should change it and what should be changed.
3. Validating A/B Test
    - A new feature tests off the charts. Your job is to determine the validity of the experiment.

1.3 DATA DESCRIPTION

This study's raw data are four-month customer data of Yammer; it is not the actual data, but it was similar to that, but it is not real for privacy and security reasons. The dataset is categorized into four tables. This dataset consisted of primary key *user_id* connecting all the tables. The figures below show the details of each table –

Table 1: Users

This table includes one row per user, with descriptive information about that user's account

user_id:	A unique ID per user. Can be joined to user_id in either of the other tables.
created_at:	The time the user was created (first signed up)
state:	The state of the user (active or pending)
activated_at:	The time the user was activated, if they are active
company_id:	The ID of the user's company
language:	The chosen language of the user

Table 2: Events

This table includes one row per event, where 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.

user_id:	The ID of the user logging the event. Can be joined to user_id in either of the other tables.
occurred_at:	The time the event occurred.

event_type:	The general event type. There are two values in this dataset: "*signup_flow*," which refers to anything occurring during the user's authentication process, and "engagement," which refers to general product usage after the user has signed up for the first time.
event_name:	The specific action the user took. Possible values include: *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*: User likes another user's message login: User logs in to Yammer *search_autocomplete*: User selects a search result from the autocomplete list *search_run*: User runs a search query and is taken to the search results page *search_click_result_X*: User clicks search result X on the results page, where X is a number from 1 through 10. *send_message*: User posts a message *view_inbox*: User views messages in her inbox
location:	The country from which the event was logged (collected through IP address).
device:	The type of device used to log the event.

Table 3: Email Events

This table contains events specific to the sending of emails. It is similar in structure to the events table above.

user_id:	The ID of the user to whom the event relates. Can be joined to user_id in either of the other tables.
occurred_at:	The time the event occurred.
action:	The name of the event that occurred. "sent_weekly_digest" means that the user was delivered a digest email showing relevant conversations from the previous day. "email_open" means that the user opened the email. "email_clickthrough" means that the user clicked a link in the email.

Table 4: Experiments

This table shows groups of users are sorted into for experiments. There should be one row per user, per investigation (a user should not be in both the test and control groups in a given experiment).

user_id:	The ID of the user logging the event. Can be joined to *user_id* in either of the other tables.
occurred_at:	The time the user was treated in that particular group.
experiment:	The name of the experiment. This indicates what changed in the product during the experiment.
experiment_group:	The group into which the user was sorted. "*test_group*" is the new version of the feature; "*control_group*" is the old version.
location:	The country where the user was located when sorted into a group (collected through IP address).
device:	The type of device used to log the event.

1.4 TOOLS USED

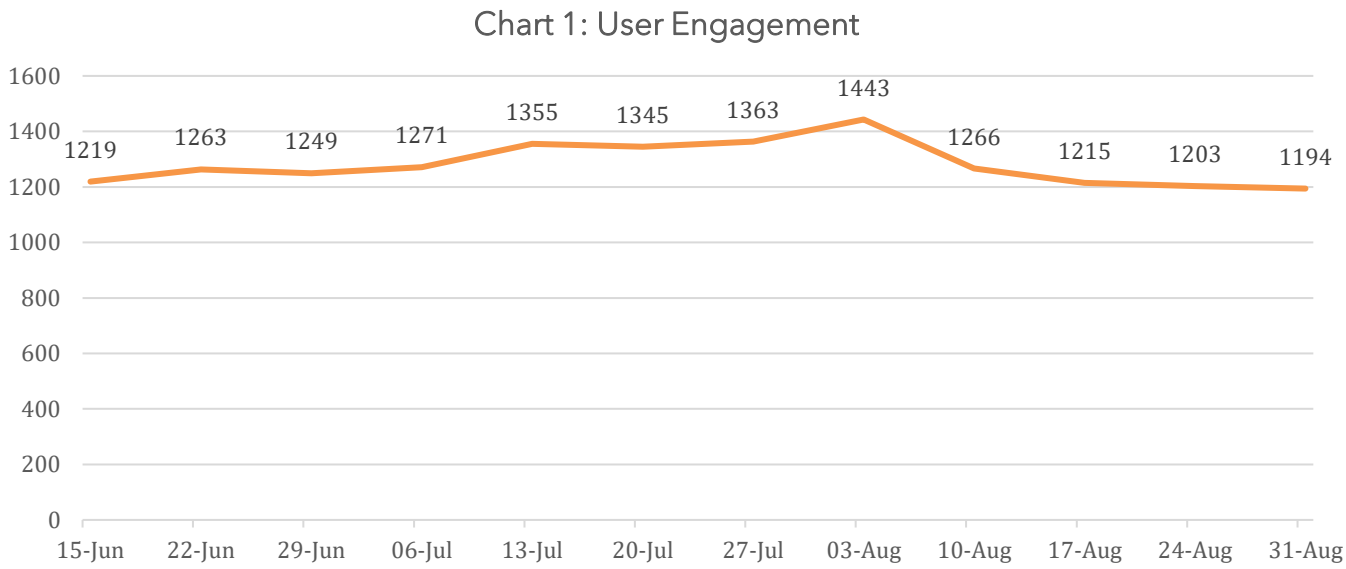
1. **MySQL Workbench 8.0 CE**: This tool was used for data querying using SQL, where required data for analysis was extracted.
2. **MS Excel:** This was used to draw graphs and charts.

2. BODY

This part solely focuses on answering the big questions of Yammer.

2.1 INVESTIGATING A DROP IN USER ENGAGEMENT

2.1.1 The Problem Statement



The above chart shows the number of engaged users each week. Yammer defines engagement as having made some 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."

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

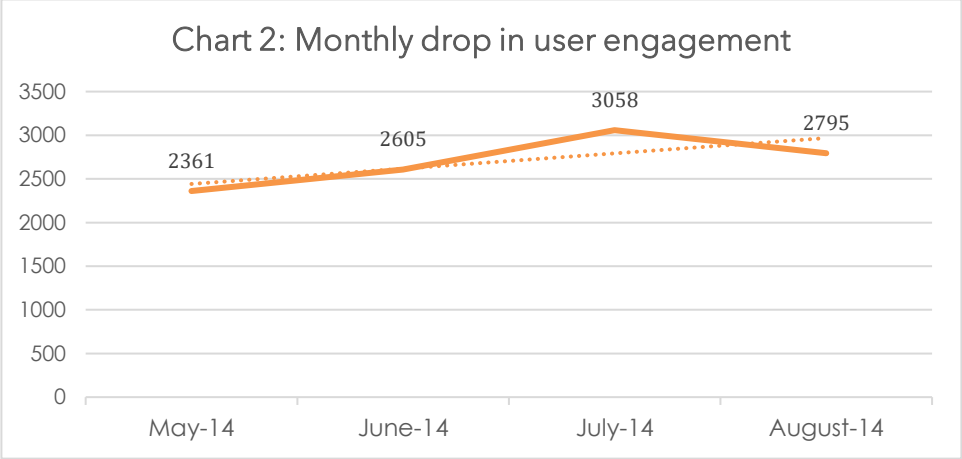
2.1.2 Potential Causes

1. **UI Design Problem:** There may be a glitch in the new design; Yammer may have released a product update in August, which a significant number of users did not like.
2. **CTA related issues:** There may be troubleshooting in any of the CTAs in the new version or bugs in the older version.

3. **Vacation or Festival Holidays:** August may be simply a month when many users go on vacation.
4. **Marketing Event**: When big promotions are launched, it's expected that a company will attract lower quality customers because of the discounts provided. Therefore, it's a possibility that the dip in August is the aftereffect of a large marketing event in July.
5. **Data is not being appropriately tracked:** It's possible that the data pipelines are not working, and thus, the information is not being followed. This is the problem caused internally and would require the help of data engineers to fix it.

2.1.3 Analysis

- Let us find the dip in engagement with monthly aggregates.



- We can see that the number of users declined from 3058 in July to 2795 in August, a drop of 8.6%. Now let us find specifically which event has attributed mainly to the dip.

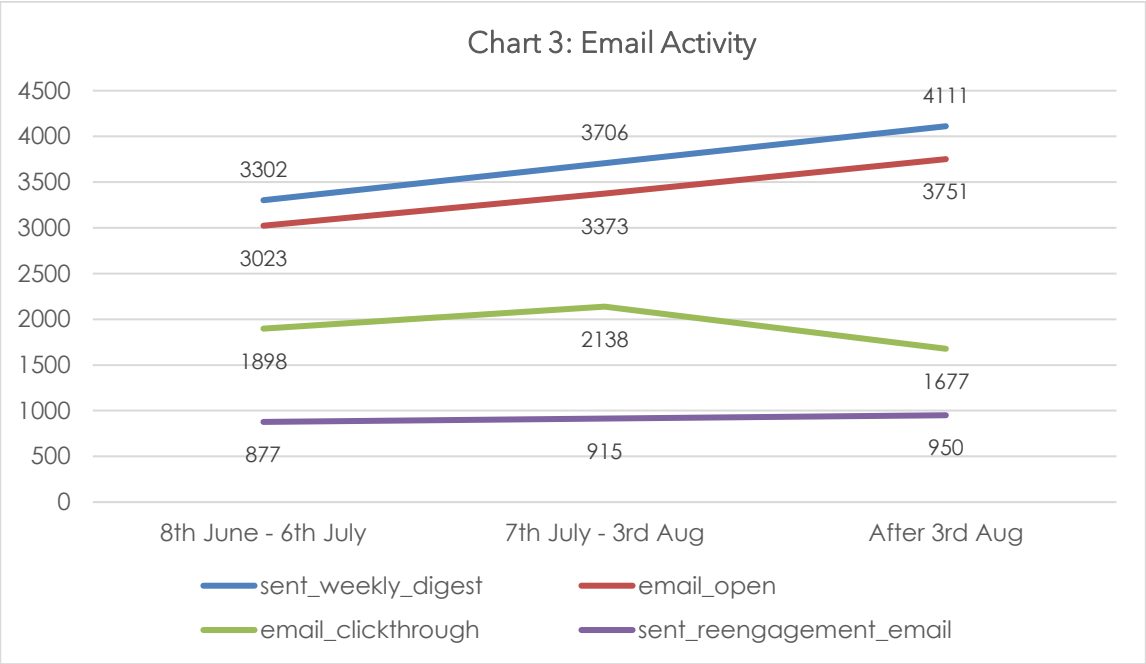
Table-5

Event_Type	Month	Number of Users (August)	Change in users
home_page	Aug-14	21603	-5162
like_message	Aug-14	13332	-3359
view_inbox	Aug-14	13011	-3077
send_message	Aug-14	7324	-2039
login	Aug-14	9271	-1752

- The above table depicts a vast drop in 'home_page' visitors where Yammer lost more than 5000 users and a significant drop in home_page, like_message, view_inbox *send_message,*

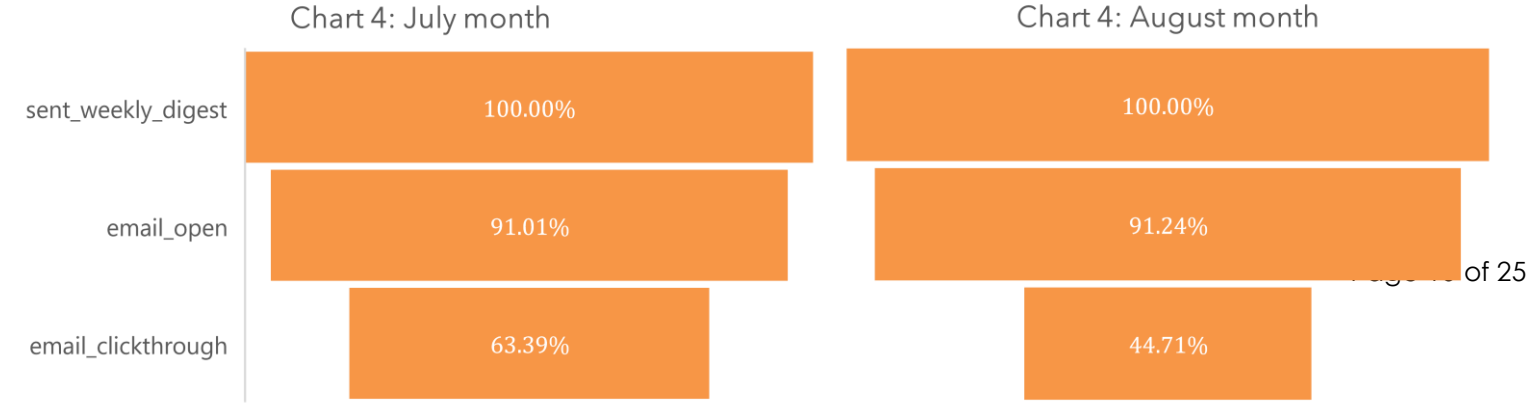
login pages where users are rarely visiting in the August month. It is skeptical to know whether the fall was due to domination of competitor's marketing event or bug seeing into Yammer website, which might be due to glitch in UI Design, CTA button breaks, etc.

- Learning about competitor's marketing event, let us check whether Yammer re-engagement emails can attract their users, where they send *sent_weekly_digest* email; means that the user was delivered a digest email showing relevant conversations from the previous day; *sent_reengagement_email* where email is sent one more time when a user does not respond to the previously shipped email. The below chart shows email activity and Click-Through Rates (CTRs)



- Yammer has continuously increased the count of sending weekly digest emails and re-engagement emails, but there is a decrement of email_clickthroughs in August from 2138 to 1677, a drop of 21.56 %

- When we look at the users funnel below, there is a huge drop in the last step due to bug in CTA buttons or UI Design.



- Now let us validate this hypothesis by looking at email click-throughs segmented by devices or device types.
- Therefore, a table has been built to provide data of several distinct users visited at least once in July and August; the last column of the table shows the change in users' percentage from the previous month.

Device	Before August 3	After August 3	Change in %
macbook pro	789	793	0.51%
lenovo thinkpad	549	523	-4.74%
macbook air	393	355	-9.67%
iphone 5	419	315	-24.82%
dell inspiron notebook	270	269	-0.37%
samsung galaxy s4	314	251	-20.06%
nexus 5	225	191	-15.11%
iphone 5s	256	186	-27.34%
acer aspire notebook	135	151	11.85%
asus chromebook	139	145	4.32%
ipad air	168	135	-19.64%
dell inspiron desktop	133	133	0.00%
iphone 4s	180	130	-27.78%
hp pavilion desktop	139	126	-9.35%
nexus 7	136	91	-33.09%
ipad mini	110	86	-21.82%

Table 6: *Table only shows the devices where each device has an average of users greater than 90

- Table 6 depicts a drop in most of the devices. When we categorize devices into mobiles, desktops, and tablets, we notice a significant drop of more than 20% in most of them; by categorizing these devices, we can get the exact solution to the problem.
- Let us build another table where we segment these devices into mobiles, desktops, and tablets.

Table 7			
Device Type	**Jun-14**	**Jul-14**	**Change in %**
desktop	1966	2191	11.44%
mobile	1502	1644	9.45%
tablet	588	636	8.16%
	Jul-14	**Aug-14**	**Change in %**
desktop	2191	2172	-0.87%
mobile	1644	1256	-23.60%
tablet	636	465	-26.89%

- Table 7 indicates a massive drop in user's engagement who use mobile and tablet because we observe a more remarkable change greater than 20% in mobile and tablet users. It may be due to a problem in UI Design.
- Finally, we find a precise reason for the drop in user engagement.

2.1.4 Conclusions:

- The drop in engagement was attributed mainly to a decline in five engagement events (home_page, like_message, view_inbox, send_message, and login).
- After we took an aggregated look at the emails table, we noticed a significant decrease in click-through rates from July to August even though there was an increase in the number of emails opened.
- By segmenting the click-through rates by device type (mobile, tablet, and laptop), we noticed that the drop in click-through rates was attributed to mobile and tablet devices.

2.1.5 Recommendations:

- Yammer product team immediately has to take a more in-depth look into the emails specifically for mobile devices and tablets. There may be a technical problem, making it difficult for users to click the email, or a UI Design problem, where the content and layout of the email are not enticing users to click.
- A good first step would be to see what changes have been made from July to August and working backward.

Table 8		
Company	**July**	**August**
amazon	11.54%	-31.03%
nokia	9.30%	-34.04%
htc	27.27%	-41.67%
microsoft	31.67%	-37.97%

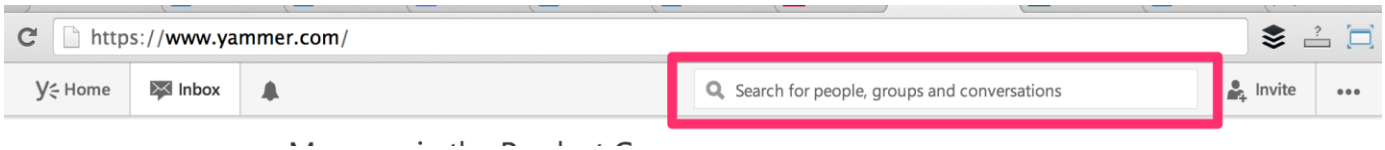
- Looking at table 8, the Yammer marketing team has successfully acquired new customers using the latest released devices (since amazon had released kindle fire in 2013, fire phone in 2011, Nokia Lumia 635 in Feb 2014, etc.). But, Yammer has lost many of these users due to the poor impression of their product. Therefore, the product team and engineering team has to focus on the technical glitches of the latest devices.

2.2 UNDERSTANDING SEARCH FUNCTIONALITY

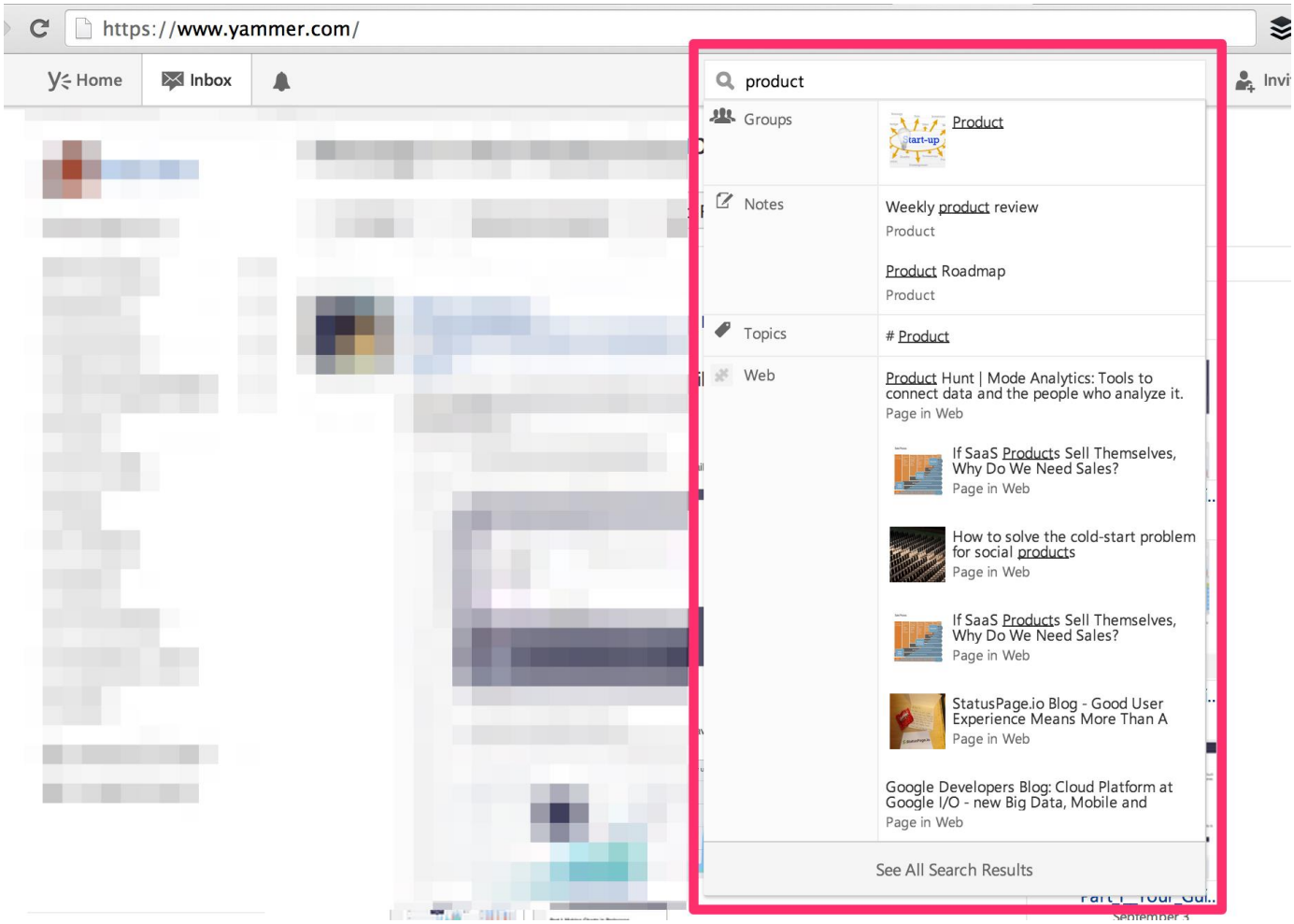
2.2.1 The Problem Statement

The product team determines priorities for the next development cycle and is considering improving the site's search functionality. It currently works as follows:

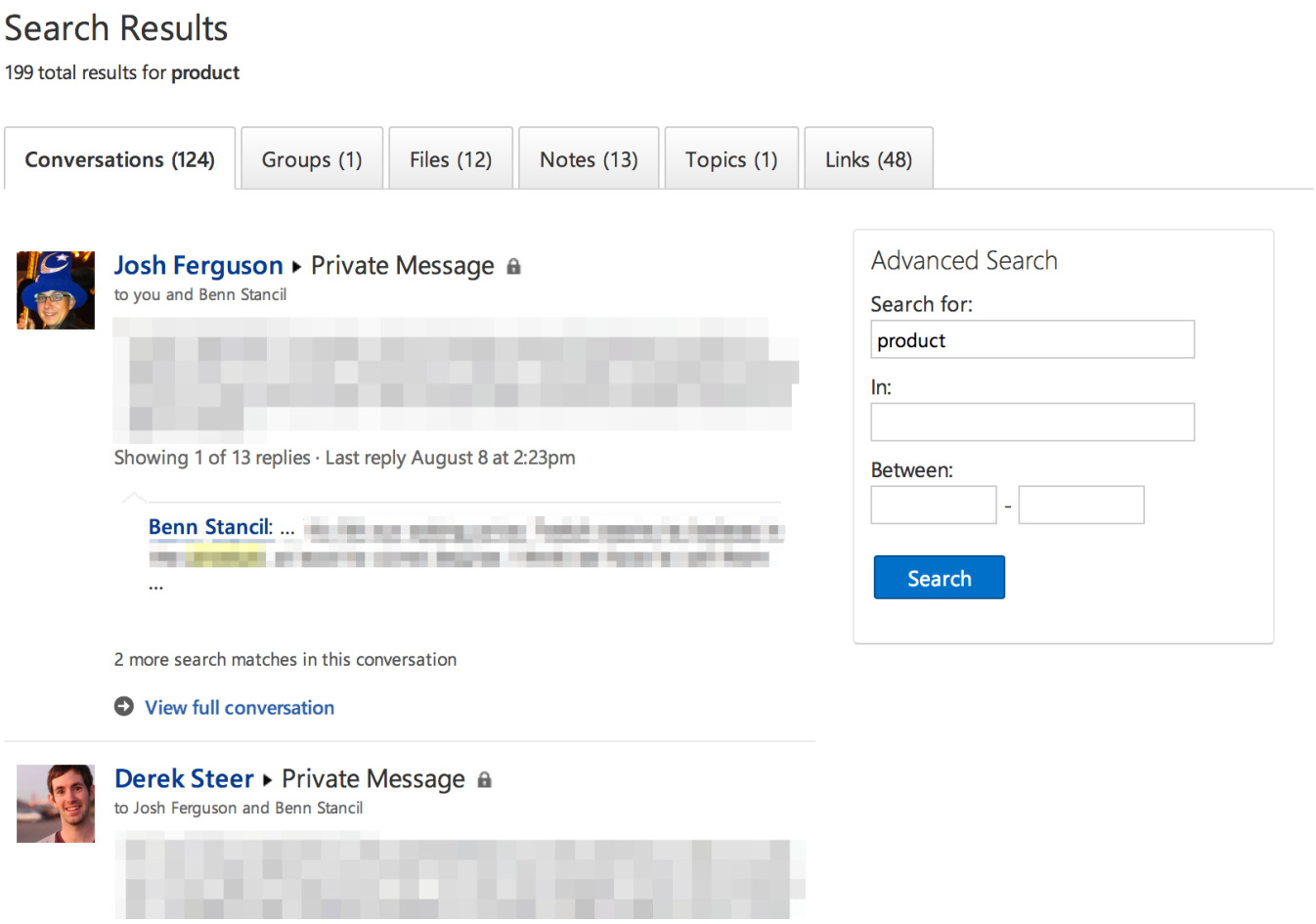
- There is a search box in the header the persists on every page of the website. It prompts users to search for people, groups, and conversations.



- A dropdown list with the most relevant results appears when a user begins to type in the search box. The results are separated by category (people, conversations, files, etc.). There is also an option to view all results.



- When the user hits enter or selects "view all results" from the dropdown, she is taken to a results page. Results are separated by tabs for different categories (people, conversations, etc.). Each tab is ordered by relevance and chronology (more recent posts surface higher).



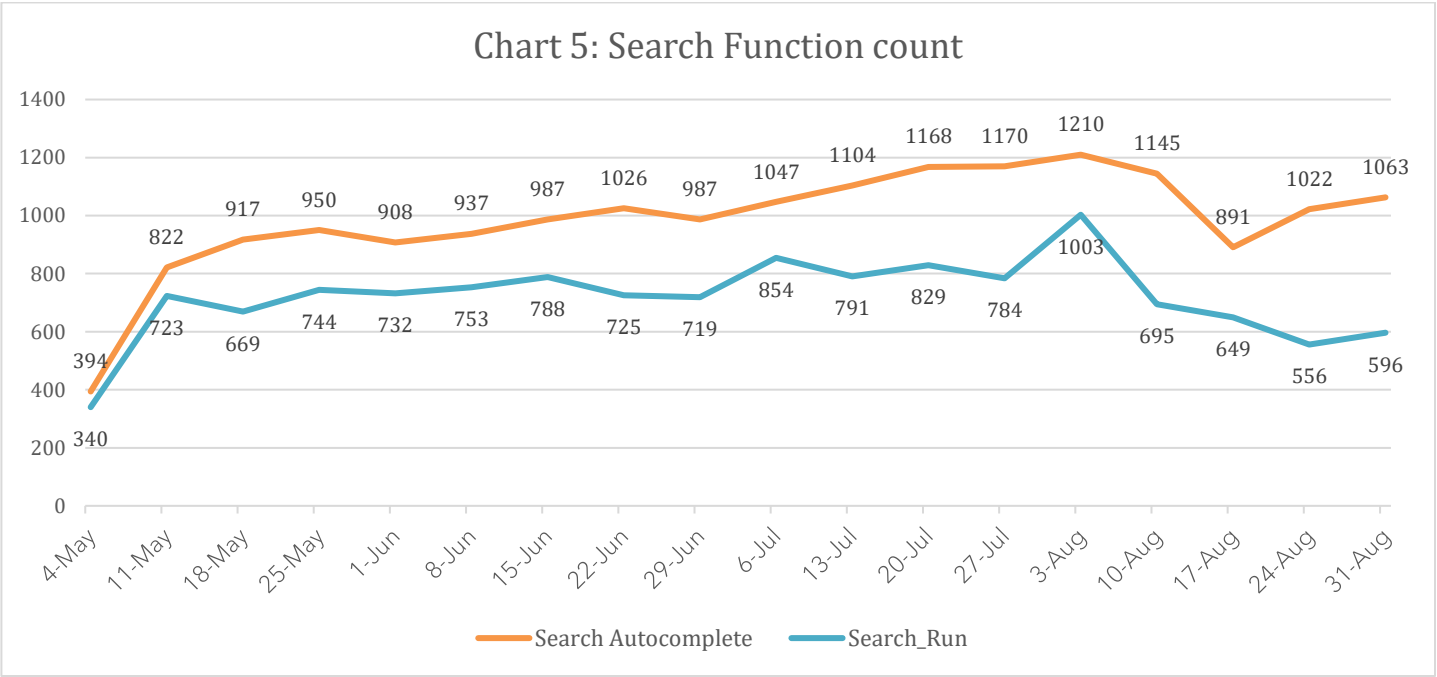
- The search results page also has an "advanced search" box that allows the user to search again within a specific Yammer group or date range.
- *Before tackling the search, the product team wants to ensure that the engineering team's time will be well-spent in doing so. After all, each new feature comes at the expense of some other potential feature(s). The product team is most interested in determining whether they should even work on search in the first place and, if so, how they should modify it*.

2.2.2 Hypothesis

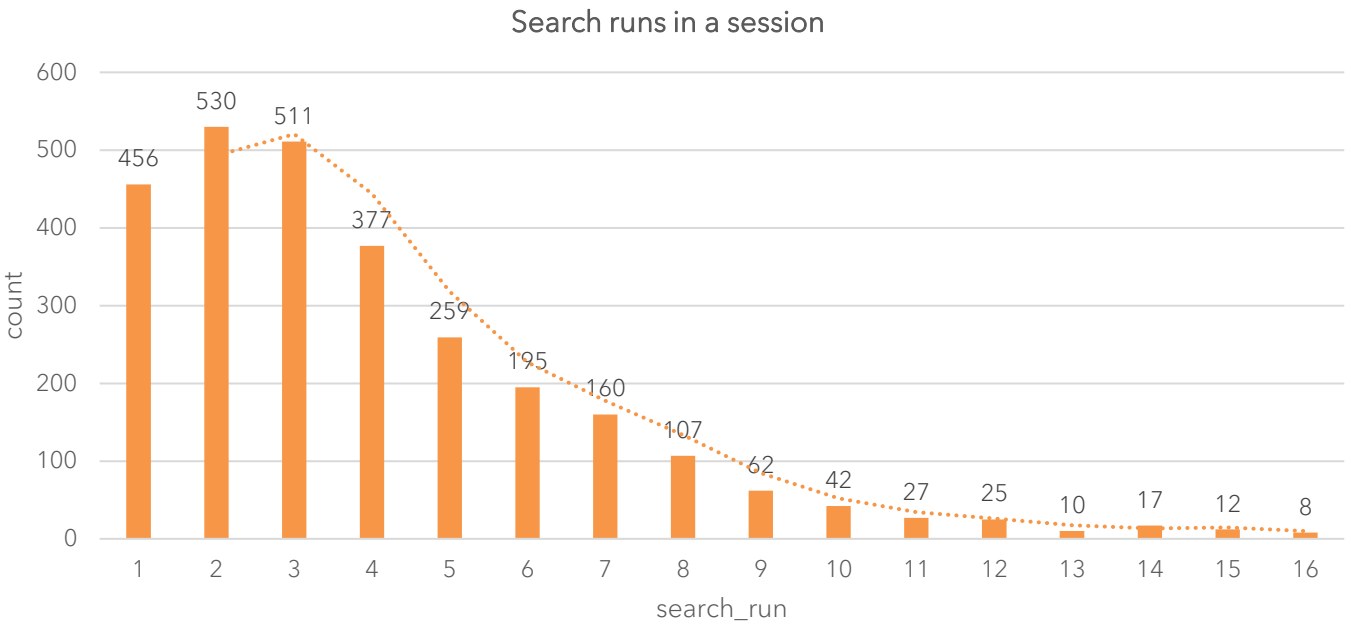
1. **Search use**: The first thing to understand is whether anyone even uses search at all
2. **Search frequency:** If users search a lot, they're likely getting value out of the feature - with a significant exception. If users repeatedly search within a short timeframe, they're likely refining their terms because they could not find what they wanted initially.
3. **Click-throughs:** If a user clicks many links in the search results, likely, she doesn't have a great experience. However, the inverse is not necessarily true—clicking only one result does *not* imply success. If the user clicks through one result, then refines her search, that's certainly not a great experience, so search frequency is probably a better way to understand that piece of the puzzle. Click-throughs are, however, very useful in determining whether search rankings are good. If users frequently click low results or scroll to additional pages, the ranking algorithm should probably be adjusted.
4. **Search rank algorithm:** If a user runs a query and clicks the bottom five results most of the time or does not click any of the results, likely, search rank algorithm is not doing well.
5. **Autocomplete Click-throughs:** The autocomplete feature is undoubtedly part of the equation, though its success should be measured separately to understand its role.

2.2.3 Analysis

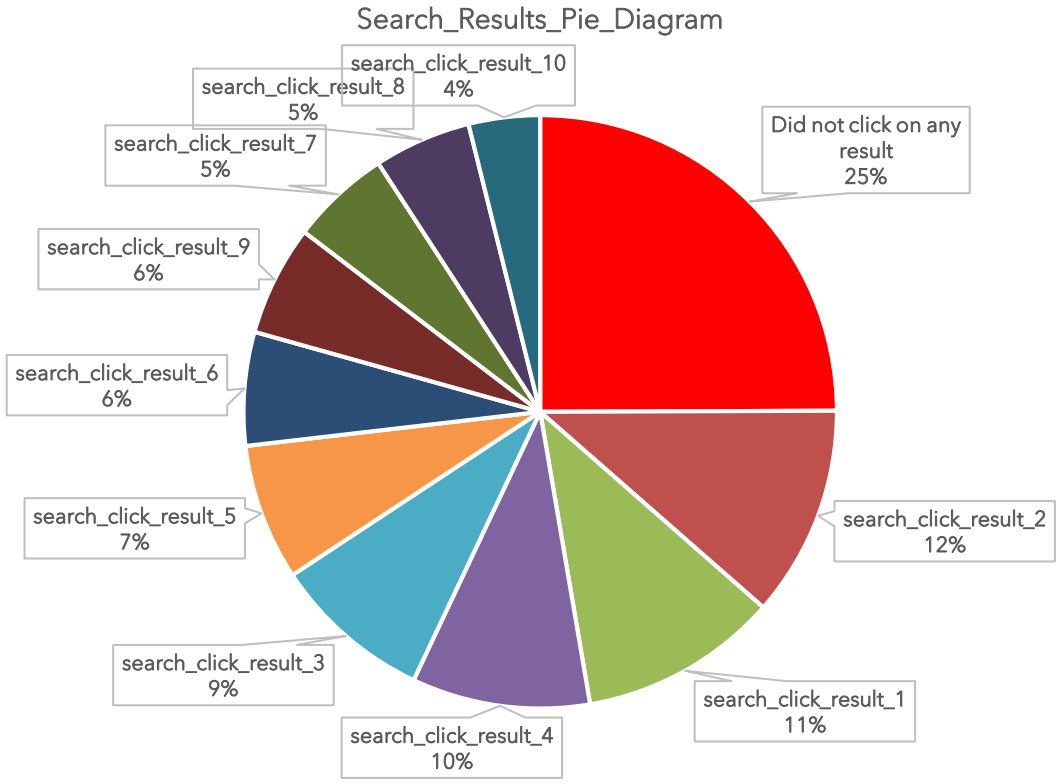
The criteria above suggest that understanding search on a session by session basis will be necessary for this problem. So before seeking to understand whether the search is good or bad, it would be wise to define a session for this problem, both practically and in terms of the data. For the following solution, a session is defined as a string of events logged by a user without an 8-minute break between two events. So, if a user goes 8 minutes without logging an event, the session is ended, and her next engagement will be considered a new session.



- First, let us look at how users use the autocomplete feature and search run. Looking at the above chart, users are keen to use the autocomplete feature than actually running a search query that takes them to the results page.
- Users prefer not to go to the search results page first; let us dig deeper to understand search run functionality.
- When users run full searches and unable to get the appropriate results, they tend to re-run the query with another keyword; when the search runs multiple times in a session, it indicates that search engine algorithm is not working well or users who like querying and doing it more frequently. The below figure depicts the exact scenario where users run full searches more than ten times a session.

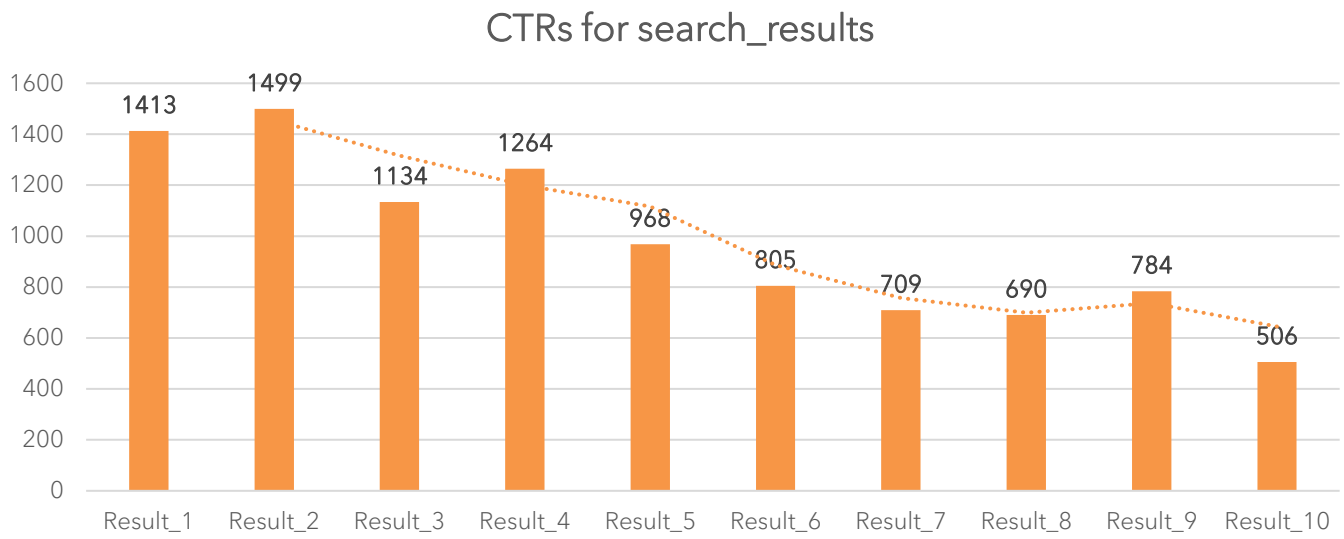


- Yammer search engine shows top 10 results on its page where users click one of the search results, but looking at the below figure, it is likely that search isn't performing well since 25% of users do not click on any of the results.



- There is a drop in 25% of users at the first stage, but let us look into the 75% users. Are they able to

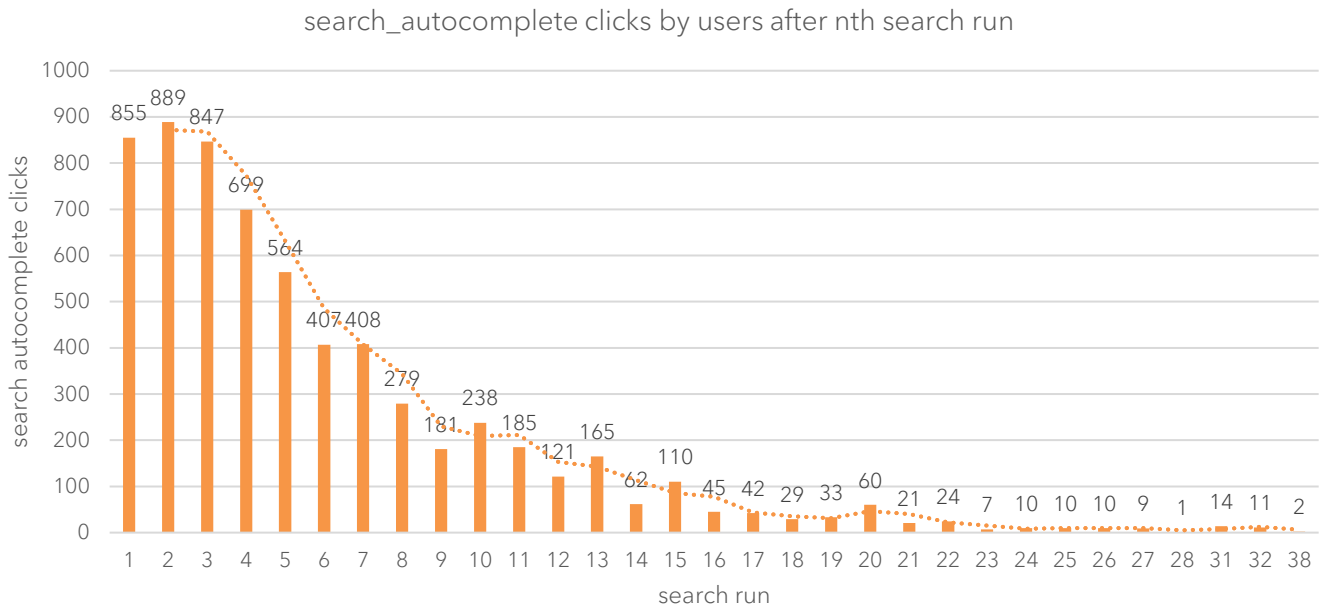
  get relevant results chronically? Let's analyze it by looking into CTRs for the top 10 search results.



- This figure depicts the sawtooth structure moving from result_1 to result_10. The user's result CTRs do not decrease chronologically; instead, we see an increase in clicks for result four, and result_9 is unlikely to visit for a good search rank algorithm.

	Search_Clicks	Constituted (%)
First four results	5310	54.34%
Results sixth to tenth	4462	45.66%

- The first four results do not even constitute 60% of clicks; if search had performed well, then the first four results would weigh 70% of the total.
- This figure concludes the poor performance of the Yammer search rank algorithm
- When a user performs a search run, the search engine filters the content and shows the top 10 results. After that, the search engine looks into previous queries, filters its related content, and displays it on the search autocomplete page. Let's look at the performance of the search autocomplete feature.



- The above figure shows the number of clicks on one of the search autocomplete page results after the user completes his nth full search.
- Yammer's search isn't able to optimize the user's large content; when the user runs the search query multiple times, we see the decrement in search autocomplete clicks.
- The search engine is not looking into extra parameters (like the type of content watched continuously, categorizing users into various segments, etc.), improving its SEO quality.

2.2.4 Conclusions & Recommendations

- Users are keen to use an autocomplete feature than actually running a search query, but search engines cannot perform well when users perform more full search runs.
- The search engine algorithm isn't performing well since 25% do not click on any of the results' results. Yammer has to work on its search functionality to engage more users.
- The search rank algorithm also has to be improved since the first four results cannot constitute at least 70% of clicks.
- Yammer has to work on its search functionality indeed.
- Most importantly, Yammer has to focus on improving its search autocomplete feature because new users mostly prefer to click to explore the product.

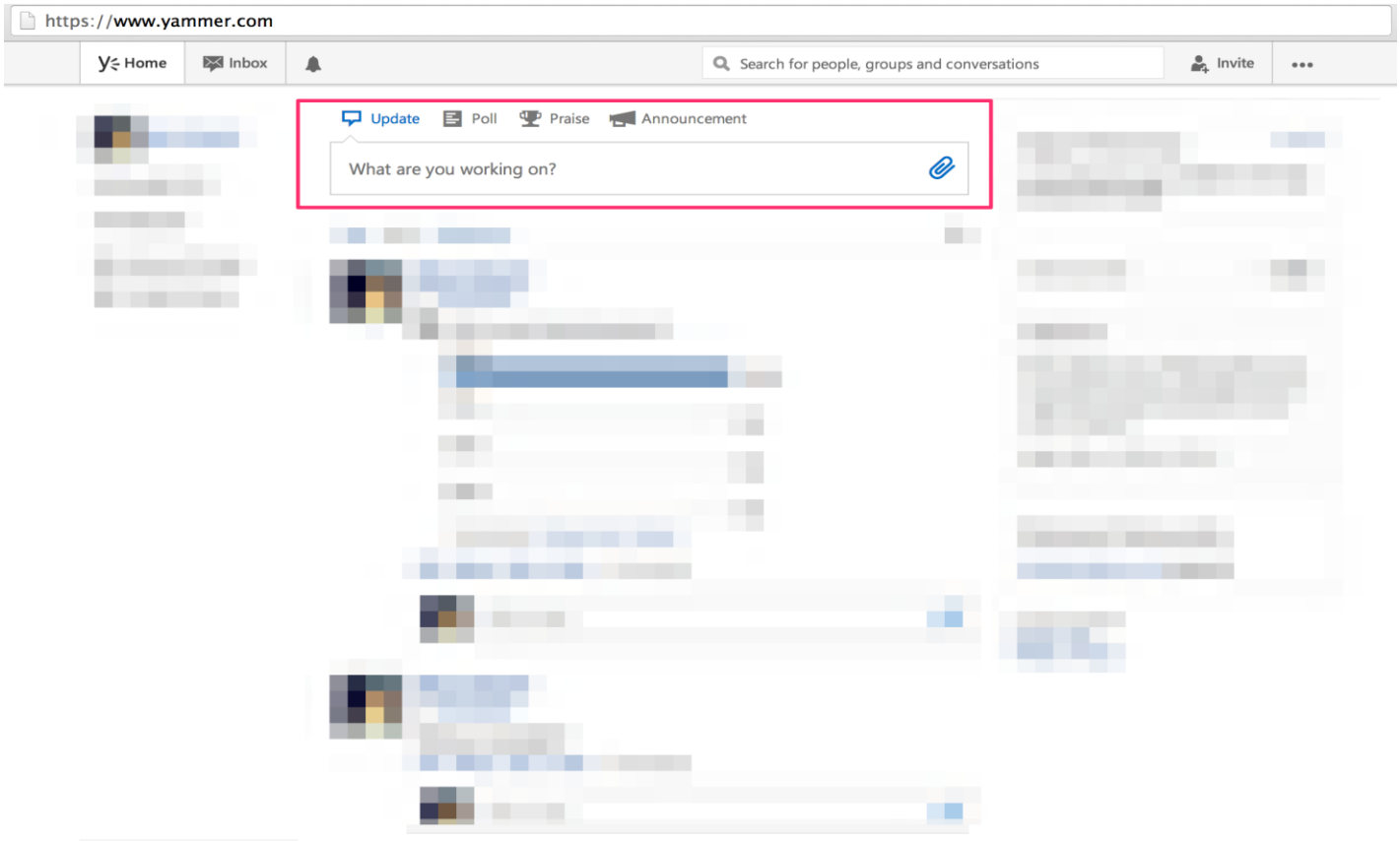
2.3 A/B TESTING RESULTS VALIDATION

2.3.1 The Problem Statement

Yammer not only develops new features but is continuously looking for ways to improving existing ones. Like many software companies, Yammer frequently tests these features before releasing them to all of their customers. These A/B tests help analysts and product managers better understand a feature's effect on user behavior and the overall user experience.

This case focuses on improving Yammer's core "publisher"—the module at the top of a Yammer feed where users type their messages. To test this feature, the product team ran an A/B test from June 1 through June 30. During this period, some users who logged into Yammer were shown the old version of the publisher (the "control group"), while other users were shown the new version (the "treatment group").

Problem: Check the results of the A/B Test and recommend the next following steps to be taken



2.3.2 Analysis

A/B tests can alter user behavior in a lot of ways, and sometimes these changes are unexpected. Before analyzing the data, it's essential to know how the test was conducted.

Users are divided into two groups, ' test group' and 'control group,' where the newer version was shown to the test group and the older version to the control group. The data of the two groups are given in the *yammer_experiments* table.

The below table shows info about two cohorts

Experiment_Group	Users
control_group	1746
test_group	849

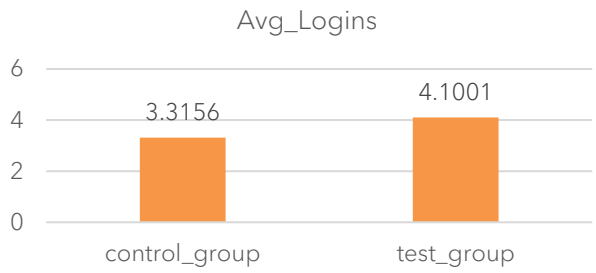
- To quantify the rise or drop of metric value, 'rate uplift' and 'rate difference' are calculated
- To ensure these results are not achieved by chance, statistical significance tests are conducted to reject the null hypothesis (to accept results are less probable to occur).
- P-value is considered as a statistical significance test and should be less than 0.005 for the test group.
- Statistical value' standard deviation' and t_stat two-tailed test were also considered
- A two-tailed test was also considered since there is a possibility of a dip in test group values.
- To know more about statistical significance tests, go through the appendix

To test the results, we have to dig into two to three metrics to make sure outcomes are positive

The metrics used to analyze the test was

- **Login frequency**

At first, Yammer has to make sure users are getting value out of new versions; looking into login frequency where numbers have gone up on an average login per user.



Experiment_Group	Total_Logins	Avg_Logins	Rate_Difference	Rate_Lift	StdDev
control_group	5789	3.3156	0	0	2.577
test_group	3481	4.1001	0.7845	0.2366	3.3091

Experiment_Group	t_stat(two tailed)	p_value
control_group	0	1
test_group	6.0705	0

- Users have responded positively to the newly added feature since the average number of logins per user for the test group is higher than the control group, and the probability of occurring this event is 'zero' (which proves the event did not happen by chance)

- **Engagement of users**

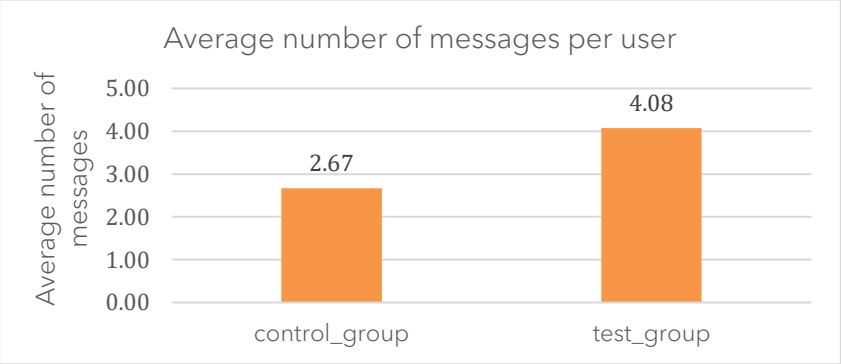
  Second, users log in on more days (days engaged the specific number of days customers use Yammer). If this metric were flat and logins were up, it might imply that people were logging in and logging out in quick succession, which could mean the new feature introduced a login bug. But both metrics are up, so it appears that the problem with this test isn't cherry-picking metrics–things look right across the board.

Experiment_Group	Users	Total Users	Total_Logins	Avg_Logins	Rate_Difference	Rate_Lift
control_group	1746	2595	5296	3.0332	0	0
test_group	849	2595	3057	3.6007	0.5675	0.1871

Experiment_Group	StdDev	t_stat(two_tailed)	p_value
control_group	2.1512	0.0027	1.000000000
test_group	2.697	6.1337	0.000000084

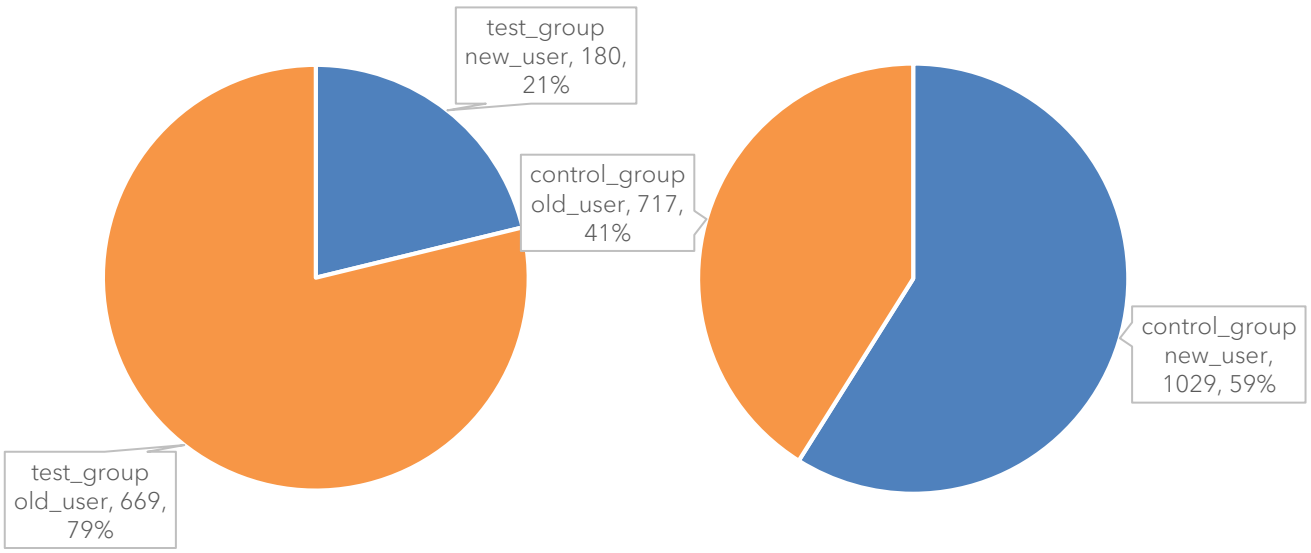
- **Number of Messages**

  Yammer has introduced new feature to increase user's messaging rate, looking at the figures number of messages of test group were higher than control group.



Experiment Group	Total messages	Avg number of messages	Rate Difference	Rate_lift
control_group	4660	2.669	0	0
test_group	3460	4.075	1.406	0.527
Experiment Group	**stdev**	**t_stat(two tailed)**	**p_value**	
control_group	3.5576	0	1	
test_group	4.7648	7.628	0	

- **Group segmentation**
  - The test, however, does suffer from a methodological error. The test lumps new users and existing users into the same group and measures the number of messages they post during the testing window.
  - This means that a user who signed up in January would be considered the same way as a user who signed up a day before the test ended, even though the second user has much less time to post messages. It would make more sense to consider new and existing users separately. This makes comparing magnitudes more appropriate, but it also lets you test for novelty effects.
  - Users familiar with Yammer might try out a new feature just because it's new, temporarily boosting their overall engagement. The feature isn't "new for new users," so they're much less likely to use it just because it's different.
  - Looking at the below figures revealed the fatal problem – all new users were treated in the control group.



2.3.3 Conclusions & Recommendations

- Introducing a new feature by Yammer has increased user's engagement rate during the month of June, which indicates adding this new feature can improve Yammer's product's overall engagement rate. Also, statistical significance tests proved that the probability of these events occurring is improbable.
- Overall, the test results are still strong. The above result should validate that change across different cohorts and fix the logging error that treated all new users in one group.

3. APPENDIX

3.1 A/B Testing Terminology

1. *average*: The average number of messages per user in that treatment group (total/users).
2. *rate_difference*: The difference in posting rates between treatment groups (test group average - control group average).
3. *rate_lift*: The percent difference in posting rates between treatment groups ((test group average/control group average) - 1).
4. *stdev*: The standard deviation of messages posted per user for users in the treatment group.
5. *t_stat*: A test statistic for calculating if the average of the treatment group is statistically different from the control group's average. It is calculated using the averages and standard deviations of the treatment and control groups.
6. *p_value*: Used to determine the test's statistical significance.

   The test above, which compares average posting rates between groups, uses a simple Student's t-test for determining statistical significance. For testing on averages, t-tests are common, though other, more advanced statistical techniques are sometimes used. Furthermore, the test above uses a two-tailed test because the treatment group could perform better or worse than the control group.