----------SUMMARY ----------

This exercise is part of the requirements of the Springboard Data Science Career Track program. As one of the curated tutorials and exercises at Mode Analytics website, it is meant to not only demonstrate critical analytical thinking ability using SQL syntax, but also allows one understand the meaning behind what is being measured. This and other case studies on that website were drawn from suggestions and interviews of Analytics managers.

Although the requirement of the Springboard program is to complete at least one of the cases, I will try to do more as time permits. You can skip over to the cases, and try reading an overview of Yammer as a company, and then my solution.

By following my thought process, anyone should be able to replicate the same analysis, and possible get a head start into any direction you need. Thank you for reading.

-------CASE STUDY----------

Investigating a Drop in User Engagement 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:

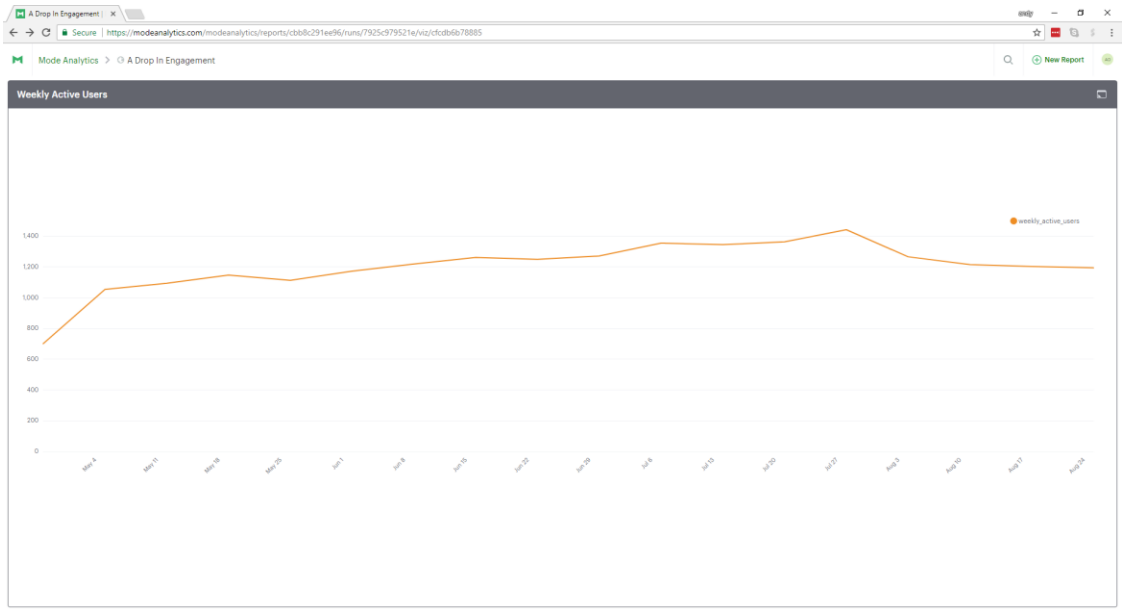

Figure 1: Same chart can be found at https://community.modeanalytics.com/sql/tutorial/a-drop-in-user-engagement/

The chart above 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.

---------- HYPOTHESIS AND ASSUMPTIONS----------

The chart above is that of weekly active users on the Y-axis against a duration of time from April 28, 2014 to August 25, 2014. It has some time series plot characteristics, because although but variables (X and Y changes), we are more concerned with how one of them (Y-axis) changes with respect to the other (X-axis). We can see that the pattern of the line plot is one has the weekly active users progressively increases over time, hit a peak on July 28, 2014 and then dips substantially before finally leveling out. In this case, the weekly active users is the dependent variable and time the independent.

The question now is: what is the likely cause of the change of the product usage knowing that the product (Yammer) is a social network that meets some requirements? These requirements encompass functionality and features together with how the product is supported. They are:

- It is used for communicating with coworkers for sharing documents, updates, and ideas by allowing users post 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 Active Directory.
- The proposed solution for this issue does not have to be overly precise, so has not to hamper productivity of analysts.
- Causes of the dip on that chart should be significant enough to cause substantial impact on overall usage of the product.
- Evaluation and analysis are based on core engagement, retention and growth in addition to product-specific usage metrics

Having said all that, the chart should be looked at in terms of what is affecting the users. It could be Internal or external influence. Internal factors being something in the line of the users’ thought, actions present or past, or decision made by them. External could cover something imposed on the users by either their companies or other external factors. Remember the user count is the dependent variable, and should be seen as composite, so, drilling down on the user itself can provide a better picture of the issues. I hypothesize the following:

1. **User experience and/or satisfaction: t**his is about the first thing I would look at. Are the users pleased with the service they are getting or being put off by it? Many a times, people try a product for a while, and leave when they see the service does not repeatedly meet expectation. So, user experience is key here.

2. **Technical or functional issues**: since the product is a social network for sharing information in groups, are there technical or functional usage issues? Are parts of the features provided by the product broken and not allowing usage? Have they been identified and resolved? Is the product meeting the service level agreement based on promised support?
3. **Vacation and/or change in workforce:** could it be that some employees of the company engaging with the product are on vacation? Usage could drop during that period. Some companies hire temporary workers to fill that need up, pending when permanent staff resume. This however, cannot justify why user count remained steady after the dip because if that was the case, the numbers should be back up when people return from vacation or remain steady with the temps filling in. Are some of the companies downsizing?
4. **Marketing and user engagement:** This is similar to way Facebook send you a message that someone poked you or
5. **License issues:** remember that the product requires one get a license when administrative controls, and additional features are needed. So, is the license affordable or are companies moving their employees to a competitor with about same service level and features?

To start off, it is good to note that it makes sense to limit analysis of the timing to shortly before and after the drop in engagement, since trends outside those boundaries are consistent.

User experience or satisfaction can most likely be gauged by significant changes in users' activities on the Yammers platform. We can see that in terms of the **"event names."** Looking at the two plots on the next page, we see that all activities carried out by the users had insignificant changes except for five events in descending order:

- Home page
- Like message
- View inbox
- Login
- Send message

For the group bar plot, we are looking at relative significant changes the counts of each events per week. The line graph is a much better representation showing trending of each events per week. Homepage engagement had the highest significant drop after the week ending 07/28/2014. The other four are shown to be related to accessing messages. It logical to think users need to get to the Home page first before getting access to login to check messages probably because the login prompt is on the Home page. So, lot of people are either not getting to the Homepage as before compared to the other events or aren't getting good content or not happy with the way the platform works, so stopped using the service. This could also be related to technical or functional issue. The Home page or some of the messaging features are probably broken causing frustration for users?

Figure 4 show that users most affected are users in English speaking countries. Of the English speaking countries, Figure 5 shows users in the United States had the biggest drop far exceeding those of other countries. Users in Japan

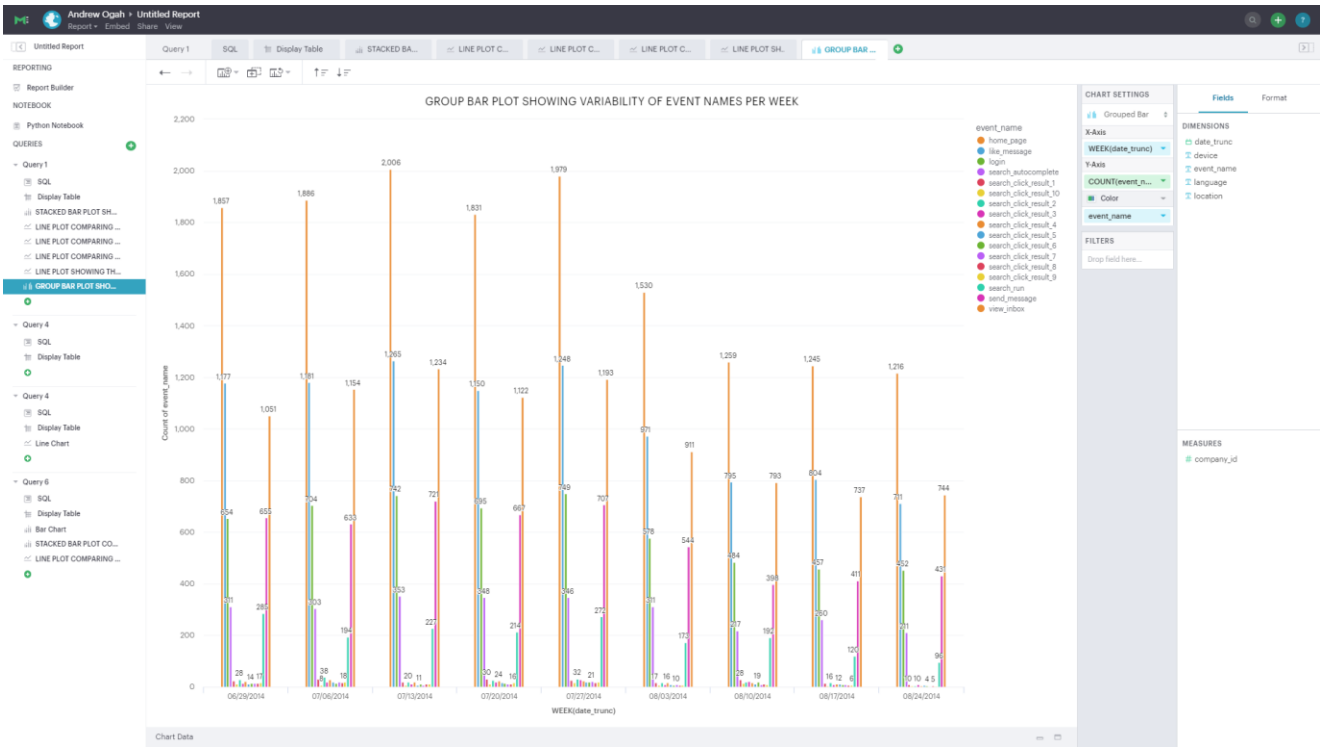

Figure 2: Group bar plot showing drop in the user activities after week 07/28/2014 in terms of the event names

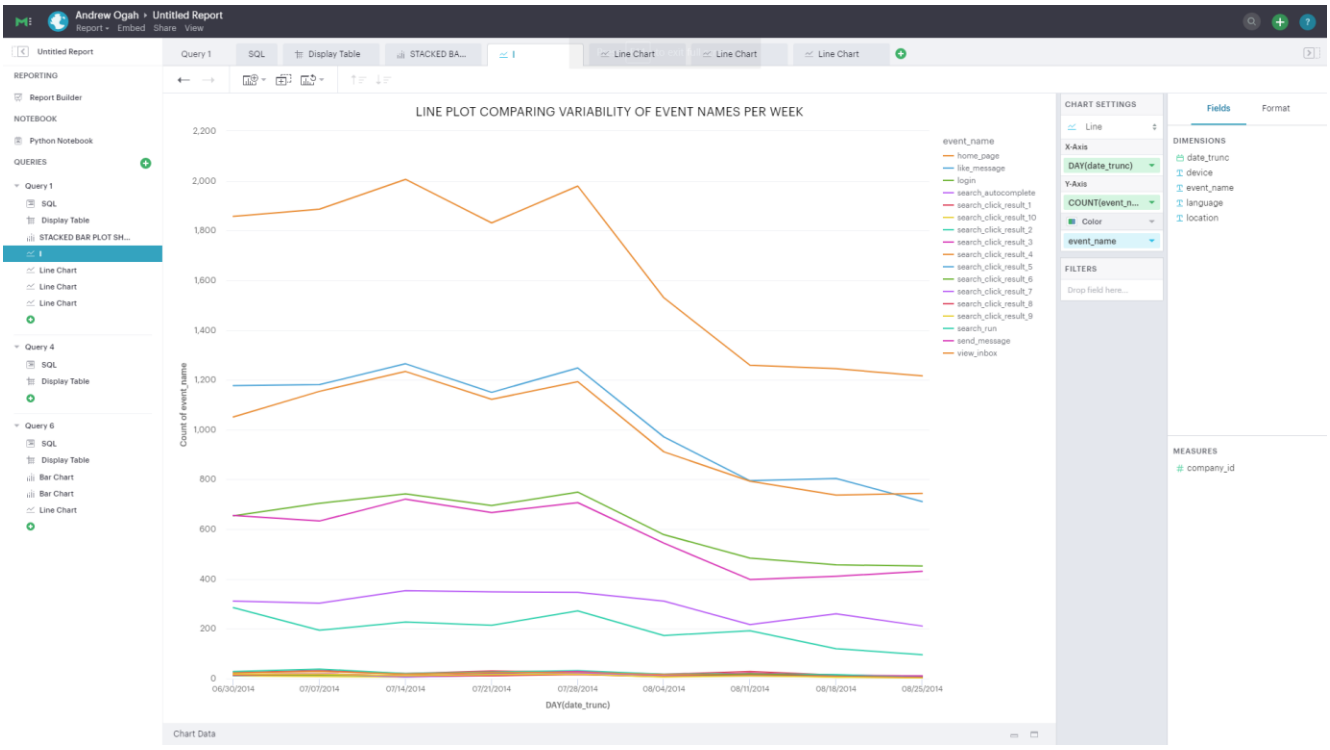

Figure 3: Line plot showing drop in the user activities after week 07/28/2014 in terms of the event names

and Germany also had some drops about the same time as shown in the zoomed in line plot in Figure 6, but we only looking at causes significant enough to cause substantial impact on overall usage of the product and a solution that is not overly precise that will hamper my time as an analyst. This means the usage drop in the US should be enough for this analysis.

Figure 4: Line plot showing drop in usage based on language of communication after week 07/28/2014

From these plots, we now know that users in the US had significant drop in Home page visits, and also issues accessing messaging services of Yammer platform.

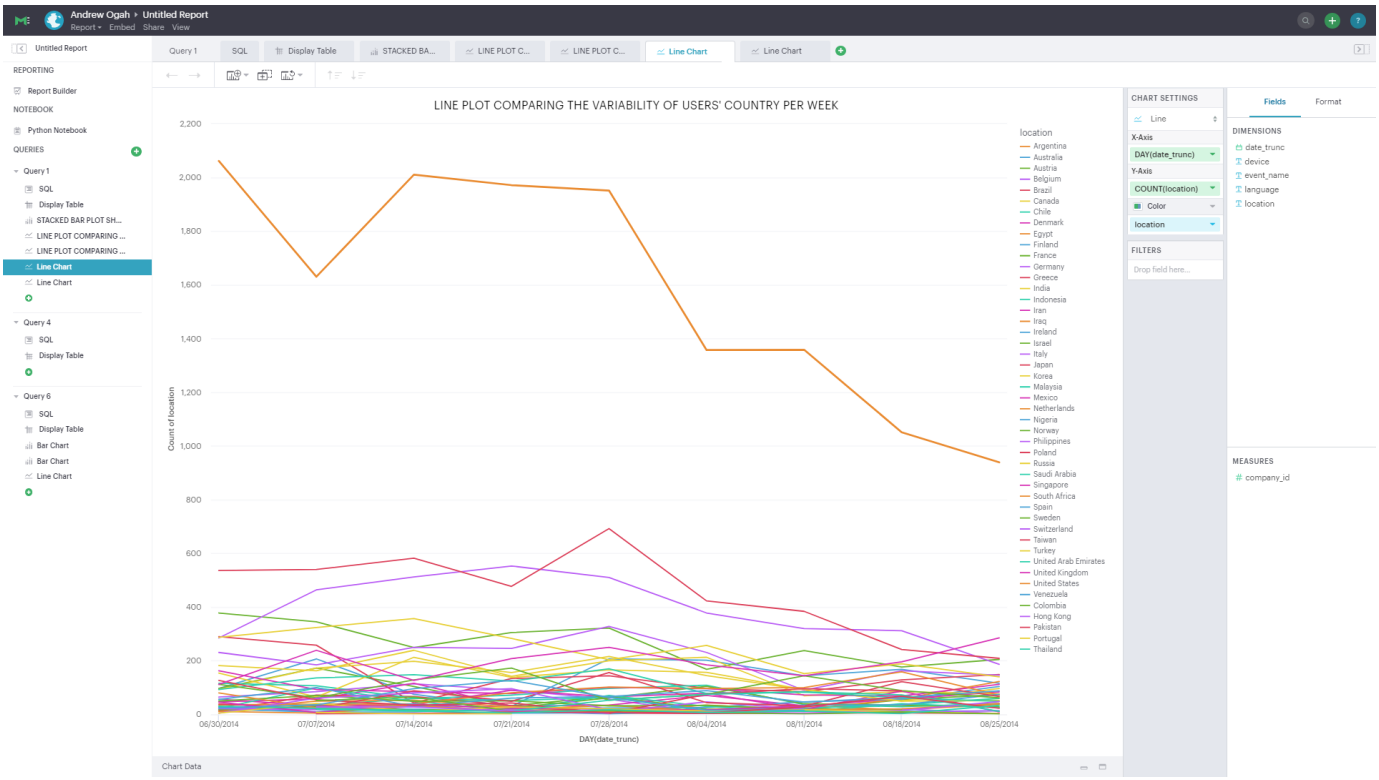

Figure 5: Line plot showing drop in usage based on country user was based out of after week 07/28/2014

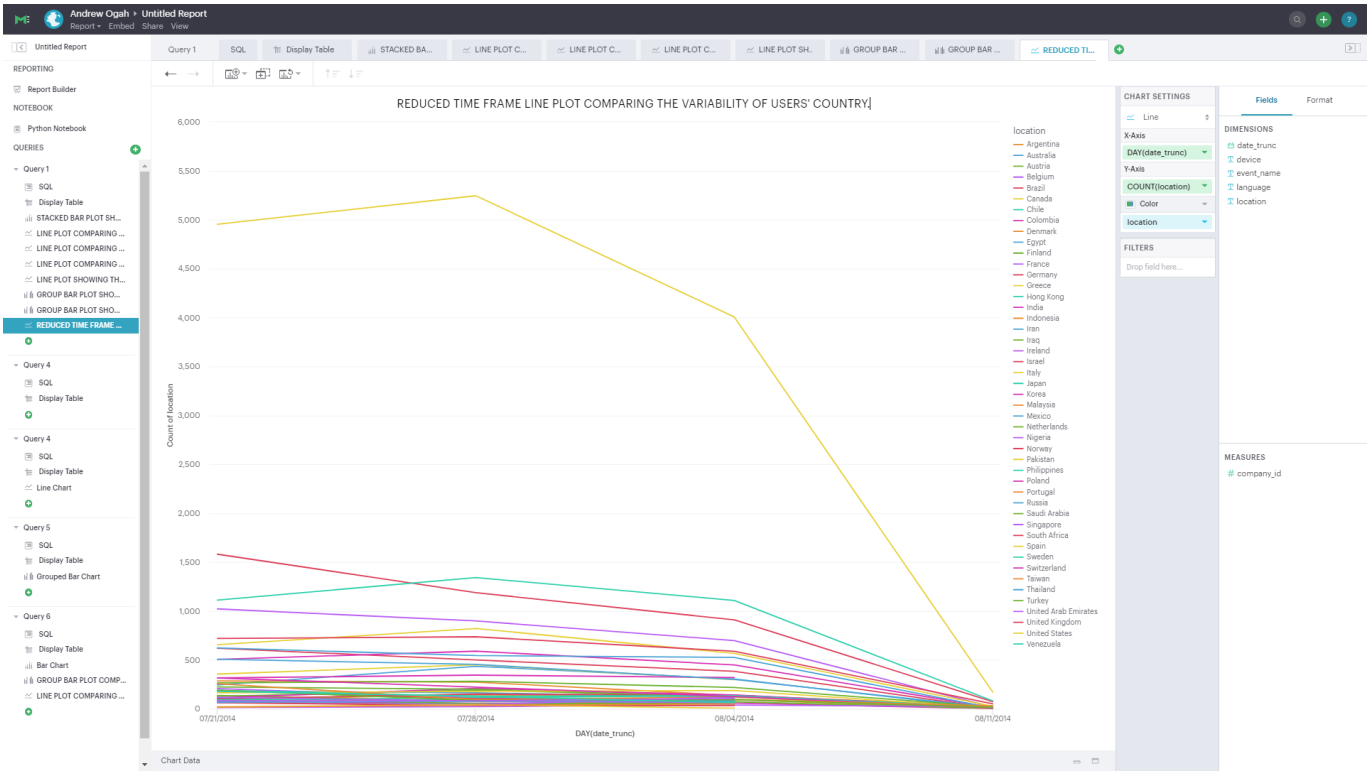

Figure 6: Zoomed in line plot of drop in usage based on country user was based out of after week 07/28/2014

Across all devices used to access the Yammer platform, there were varying dip in usage on all devices used to access the Yammer platform with that of Lenovo ThinkPad the most significant shortly after 07/28/2014. But there was significant drop across board after 08/03/2014 as shown in Figure 7 below.

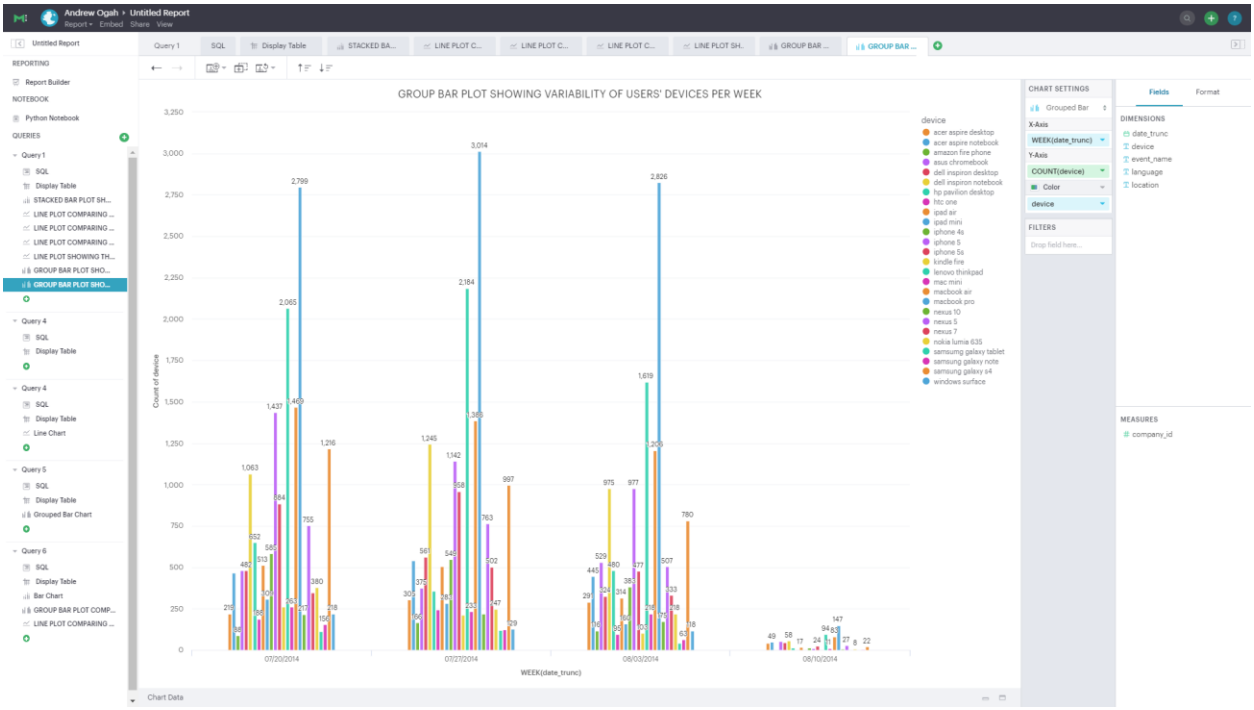

Figure 7: Group bar plot showing variability of users' devices per week.

Now, looking at the chart below in Figure 8, we see the likely cause of most of the issues going on. After the week ending on 07/20/2014, Yammer stopped sending reengagement emails to it users, but increased weekly digest mailed out. And, users' engagements with the platform dropped significantly the week after, with a subsequent further significant drop in the following weeks.

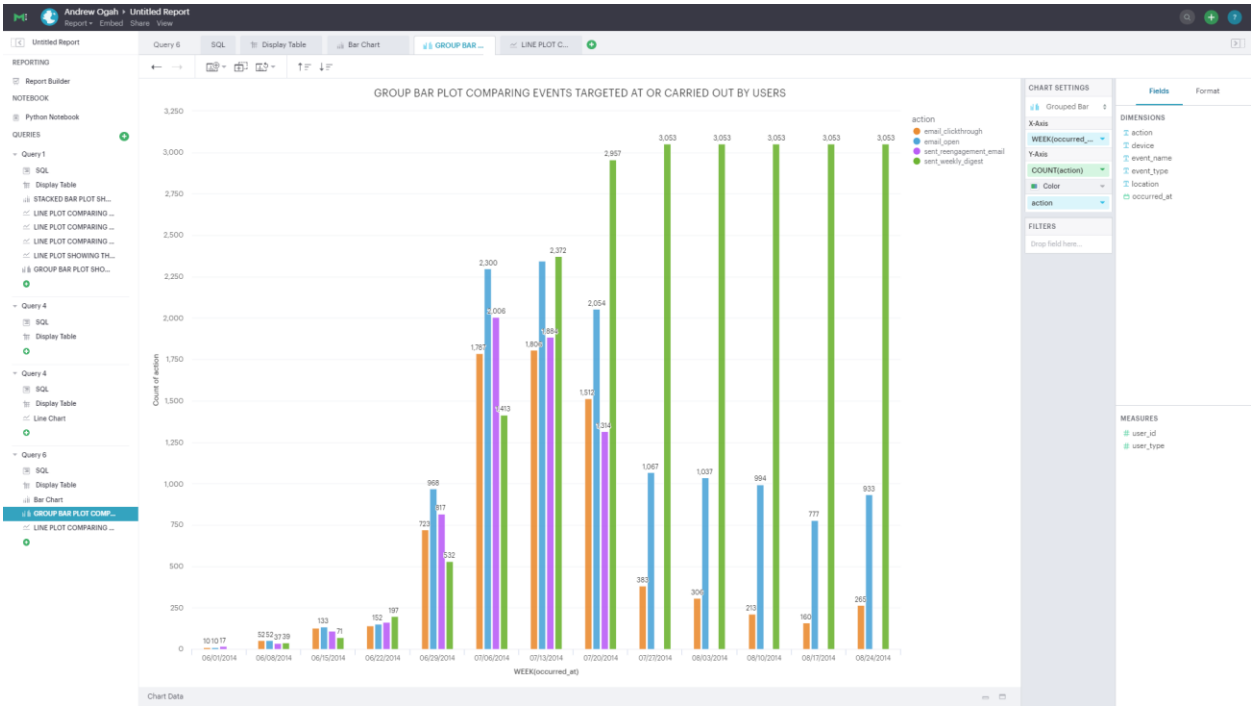

Figure 8: Group bar plot showing events targeted at or carried out by users after week 07/28/2014

So basically we now understand that:

- Most significant drop in Yammer usage was in the US
- This issue started shortly after Yammer stopped sending out reengagement emails

My recommendation will be to advise the head of the Product team to reintroduce and improve reengagement of users much like all other social media platform like Facebook does, with a primary focus on the users in the US. Reengagement is basically drawing attention of users to your platform. If you've ever received message from Facebook out of the blue when you've been away from the platform for some time asking you if you know someone or you've been poked and you follow the link to login only to notice it is all garbage or from LinkedIn that someone is viewing your profile, then you know what reengagement is. This is basically drawing people to access the platform and keep them on it.