**-----------SUMMARY ----------**

This exercise is part of the requirements of the [Springboard Data Science Career Track program](https://www.springboard.com/workshops/data-science-career-track/). As one of the curated tutorials and exercises at [Mode Analytics](https://modeanalytics.com/) 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](https://community.modeanalytics.com/sql/tutorial/sql-business-analytics-training/); as time permits, I will try to do more. 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:

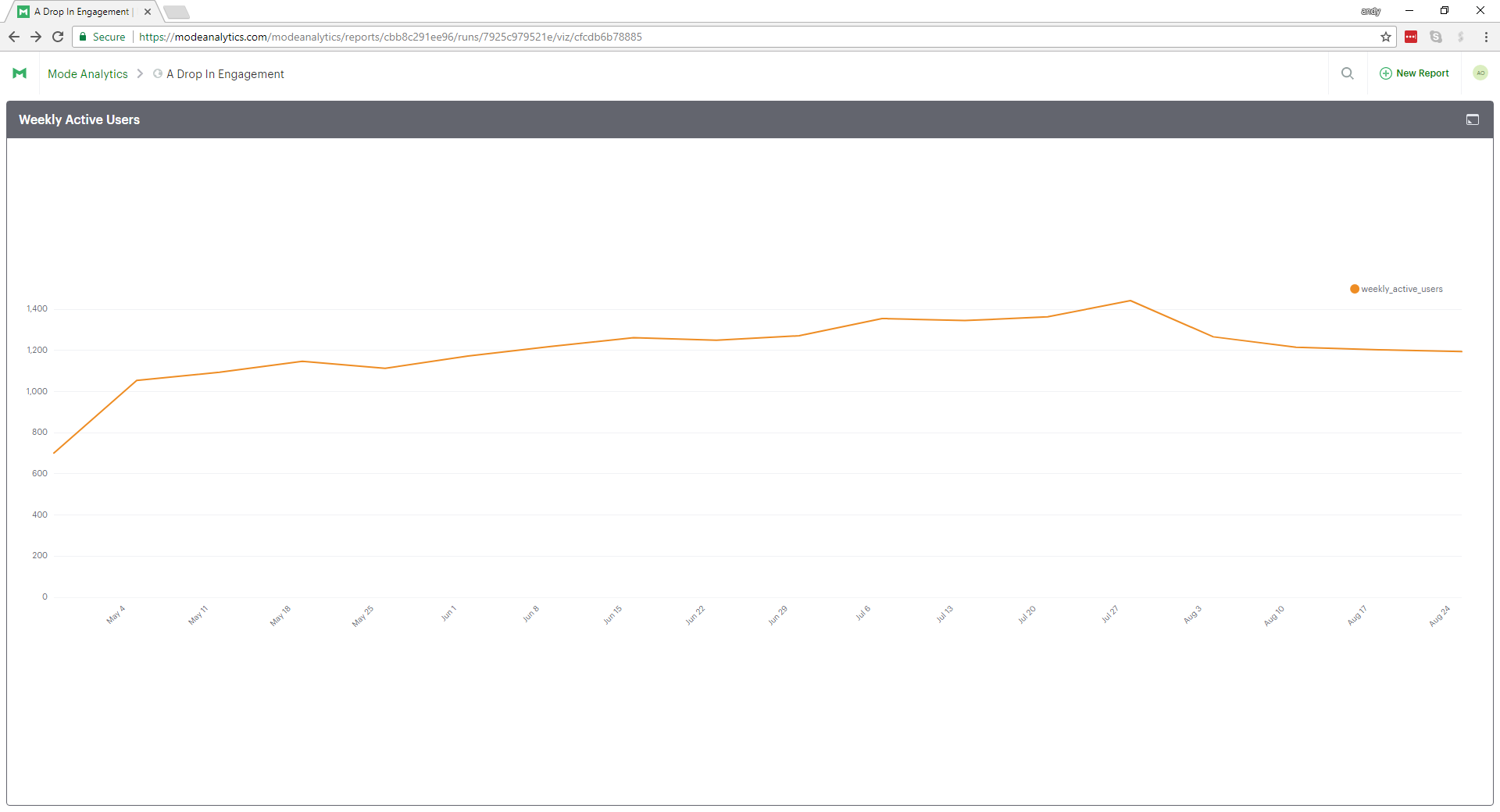


Figure 1: Same chart can be found at <https://community.modeanalytics.com/sql/tutorial/a-drop-in-user-engagement/>

The above chart 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, a way to look that chart is in terms of what is affecting the users. It can either be something internal or external. Remember the user count is the dependent variable. It can be somewhat thought of has being composite, because either the user is thinking something, doing or did something and/or made or makes a decision. So, we need to drill down into the user itself to get a better picture of the issues. From the internal and external factors that could be affecting, 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 they’re not getting their money’s worth or see the service does not repeatedly meet expectation.
2. **License issues:** remember that Yammy 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?
3. **Technical 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 Yammy meeting the service level agreement based on promised support?
4. **Vacation:** could it be that some employees of the company engaging with the product are on vacation? This is reasonable because summer is when most people take time off for vacations because the weather is warmer. So, usage could somewhat reduce. However, it may not justify why user count remained steady after the dip because the numbers should be back up when people return from vacation.
5. **Change in workforce:** some companies hire lot of temporary workers during some seasons to keep up with peak times, and let the workers go out of season. So were temporary workers let go after a while? Did some companies downsize to try balancing their books?

---------- Pending Updates---------------