**MINI PROJECT**

**Databricks Notebook:**

<https://databricks-prod-cloudfront.cloud.databricks.com/public/4027ec902e239c93eaaa8714f173bcfc/5318821742706204/4195284667863997/5056357010360280/latest.html>

The above link is for the Databricks notebook.

**Project Submission:**

(1) Has actually user activity or engagement dropped recently and if so, how serious or significant is it?

**Solution:**

Yes, user engagement had dropped recently. It dropped from July month. I support my answer by proving in the form of graphs plotted by DATE, WEEK and MONTH accumulated by the user activity/engagement.

1. In the graph that I plotted first, I grouped the user activity by DATE. The graph plotted has a lot of ups and downs. So to get a better view of the user activity, I plotted the same graph by grouping WEEK wise.

A screenshot of a cell phone

Description automatically generated

1. In weekly view of the graph, we can see that from April 27th  2014 (4,185) to May 4th 2014 (7,787) user activity has increased by 3,602. After that day, there are many ups and downs. But, from July 27th 2014 (8,291) to August 10th 2014 (5,922) user activity has decreased by 2,369. This indicates that recently the user activity has decreased. To get an even more concise view, I plotted the graph grouping by MONTH.

A screenshot of a computer

Description automatically generated

1. In the monthly view of the data we can clearly see that from June 30th 2014 (36,055) to July 31st 2014 (26,846) the user activity has fallen drastically.

A screenshot of a cell phone

Description automatically generated

Hereby, we can conclude that user activity has dropped drastically from the above graph. This could mean that there is something which is not working as expected and that need to be rectified.

(2) Think about possible reasons (at least three) for drop in activity, i.e., develop some hypotheses that you can later test if/as necessary in a future analysis. Investigate each of these potential reasons by conducting analysis using the relevant data, writing SQL queries, and generating related visualizations.

**Solution:**

Below are some of my investigations on what might be the potential reasons for a drop in user activity. Each of these investigations are accompanied by some visualizations that will summarize the analysis done and what the company can do to improve the user activity.

1. One possible reason for drop in the user activity can be attributed to some of the recent changes done on the application. This could be because of some underling bugs which might have snuck past testing and the web/mobile application can become non compatible with latest changes. So I wrote a SQL query to prove my hypothesis and plotted the graph for the user activity data grouping by devices. Below is the graph

A screenshot of a cell phone

Description automatically generated

From above graph we can see that there is a drop in user activity who uses mobile phones. The reason for that could be because of a problem in the mobile application or the type of actions that is mostly performed on phones. This lead me to my next hypothesis which has to do with emails.

1. The next possible reason for drop in user activity might be the emails. I thought that the reason for considering this is because a company tries to engage users with their product using emails. One way to bring the users back is to send digest emails to users. So I wrote an SQL query to group by email actions and plot the relevant graph which is shown below. A screenshot of a cell phone

   Description automatically generated

We can clearly see from the above graph that the email clickthrough’s have dropped in the recent past. The email clickthrough’s will indicate that the user is interacting with the product by clicking on the links. There might be multiple reasons for the user not clicking the links in the email. One reason might be because the user is not motivated enough to click the links because of an outdated look and feel of the email page or the links are not properly displayed. Another reason could be because the product might be lacking the sense of freshness for old users to engage with it appropriately. This lead me to my next hypothesis.

1. From the above hypothesis, I have compelling reason to believe that old users might not motivated enough to use the product. I wrote a query to group the activity by the age of the users(the difference between sign up and engagement date) using the product. I have plotted the below graph to prove my hypothesis.

A screenshot of a cell phone

Description automatically generated

The above graph clearly displays a significant drop in the user activity for users who are using the product for more than 3 months. This indicates that the old users might not be liking the product/application or they found an alternative solution for it.

From the above 3 hypothesis, we can see that the drop has started from July 27th 2014. So I guess there might be some changes applied to the company’s web or mobile application in July of 2014. If I were an employee of the company, I would revisit those changes and fix the issues if there are any that were introduced.

1. One more assumption I had in mind is that user activity might depend on the day when the user interacts with the product. So I wrote a SQL query to check my hypothesis by grouping the user activity based on weekdays and weekends. I plotted a graph based on the data I got from executing the query.

A screenshot of a cell phone

Description automatically generated

The above graph indicates that the user activity is very less on Saturday and Sunday. So I plotted the user activity data to differentiate between weekdays and weekends to reduce the clutter in the pie chart above.

A screenshot of a cell phone

Description automatically generated

From the graph shown, we can deduce that the users use the product on weekdays more than weekends. This might be because they use it mostly during their work hours. The drop in the user activity on weekdays started on July 27th. Organizations normally acquire licenses for using third party applications. So one or more organizations using the product could have not renewed the license or opted for some other application. The product might no longer be helpful for the organization growth.

**Here are some of my recommendations based on the hypothesis done above**

* The major factor for the drop in user activity can be attributed to the lack of newness in the product. There is a significant drop in the user activity for people who have used the product for more than 3 months. I would suggest for the company to come up with some exciting new features that will attract the user base.
* The users are sending less messages using the product. This might indicate they are using an alternative messaging application to send and receive messages. The company could revisit the licensing and provide the license for lesser amount of money or provide discount for user retention.
* Looking at the drop in user activity, phones look like a popular platform used by majority of the users because there is a significant drop in phone activity. This gives an opportunity for the company to divert some of their resources to build a robust and efficient phone application that is efficient and has great user experience.