Newsfeed Analysis Project Write-Up

Author: Avinash Bisram

Sections:

Overview: Hello! The purpose of this write-up is to explain the background, goals, deliverables, and insights gained from the *Newsfeed Analysis in Tableau* project. This write-up is meant to accompany the final Tableau Dashboard which **can be found here** (a screenshot of the final view can also be seen in the *Deliverables* section).

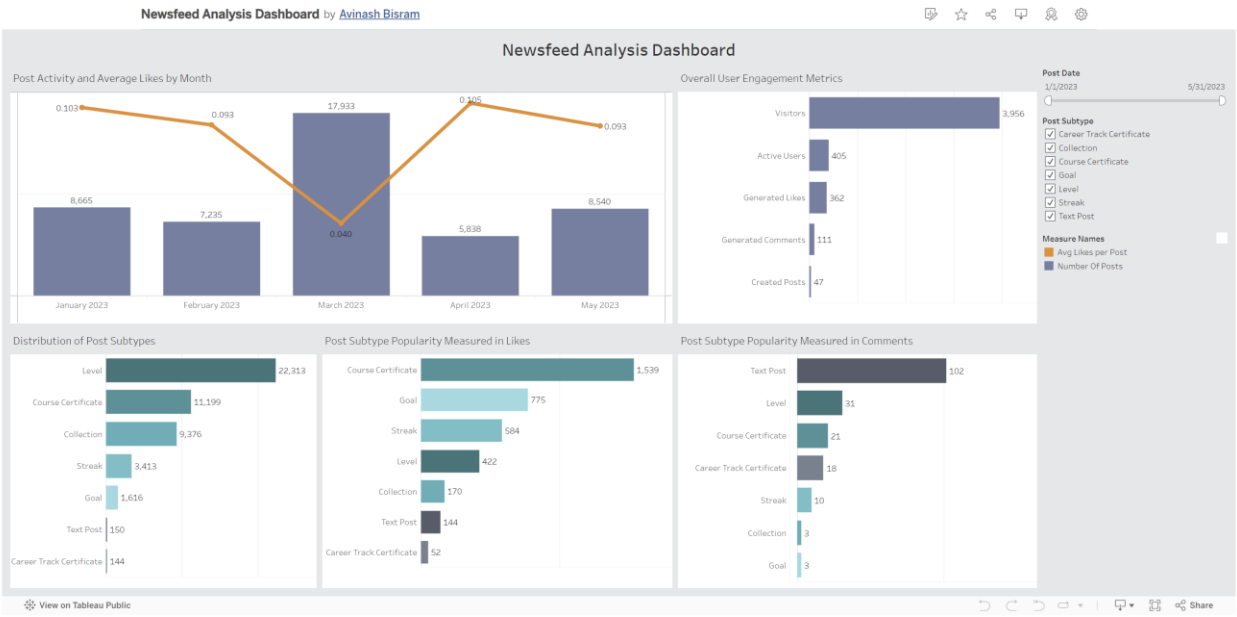
Background: This project is one of many that can be found on **365DataScience.com**. The dataset used was a subset of real user-interaction data relating to the site's Newsfeed service (see the *Dataset* section for more information). As paraphrased directly from the site, the Newsfeed is a part of the website which constantly updates with new content, showcasing the latest news and achievements from users. It offers a mix of automated and manual posts which showcase student milestones, the completion of a course or career track, extending one's learning streak, among other things.

Dataset: The dataset used to construct the final dashboard was injected directly into the blank Tableau packaged workbook available at the start of the project. The first table consisted of 824 rows and 6 fields (post date, post subtype, post type, numberOfPosts, numberOfLikes, numberOfComments). The second table consisted of 1 row and 5 fields (activeUserCount, commentedCount, likedCount, postedCount, visitedCount).

Goals: The goal of this project was to create visualizations capable of answering several Key Performance Questions outlined by the 365DataScience team and gain crucial information about how users are interacting with the Newsfeed. Specifically, this project was aimed to identify which aspects of the Newsfeed are beneficial and which ones are detracting from the overall user experience. Furthermore, after creating the dashboard, I was tasked with expanding on insights and trends to offer suggestions on how to enhance user interaction with the Newsfeed service.

Deliverables:

The main deliverable for this project was a Tableau dashboard created using the data provided. An image of the final view can be seen below.



The dashboard consisted of **5 visualizations** and **2 filters**. **4 calculated fields** were also made to facilitate creating these visuals.

Short descriptions of the 5 visualizations are as follows:

- “Post Activity and Average Likes by Month” (Top Left): Dual-axis chart consisting of a bar chart representing the overall number of posts per month and a line chart representing the average likes per post per month.
- “Overall User Engagement Metrics” (Top Right): Horizontal bar chart illustrating the static engagement metrics present in the second data table (not intended to change with the filters as per the project’s instructions).
- “Distribution of Post Subtypes” (Bottom Left): Horizontal bar chart illustrating the frequency of post subtypes (the various automated and manual posts that can appear in the Newsfeed) within the selected time frame.
- “Post Subtype Popularity Measured in Likes” (Bottom Middle): Horizontal bar chart visualizing the frequency of likes per each post subtype (hovering over each bar expresses the frequency as a percentage of all posts of that subtype – in other words a likes-to-posts ratio).
- “Post Subtype Popularity Measured in Comments” (Bottom Right): Horizontal bar chart visualizing the frequency of comments per each post subtype (hovering over each bar expresses a comments-to-posts metric similar to as stated above).

Key Performance Questions:

The three main key performance questions this dashboard sought to answer were:

- Regarding average likes per post, which post subtype is the least favored?
    - As can be seen by filtering each post subtype and analyzing the dual-axis chart or hovering over the bars in the Popularity Measured by Like bar-chart, the Collection subtype was the least favored as quantified by average likes per post (0.0181). This was closely followed by the Level subtype with a value of 0.0189.
- In absolute terms, which types of posts are most common, and which are the rarest on the newsfeed?
    - As can be seen in the Distribution of Post Subtypes chart, Level posts are by far the most common at a little over 22,000 within the given timeframe. This is slightly under ALL THE OTHER subtypes combined. The rarest type of post is the Career Track Certificates with 144 total posts followed closely by the Text Posts at 150.
- What percentage of visitors to the newsfeed actively interacted with the feature?
    - This can be calculated using the Overall User Engagement Metrics chart in the top right. Dividing the number of Active Users by total Visitors in the time frame (405 divided by 3956) gives a value of just over 10%.

Additional Insights and Recommendations:

Looking at the dashboard as a whole, we can take away a few key insights about the health of 365DataScience's Newsfeed service. We can see in the dual-axis chart that the average likes per month is not necessarily tied to the quantity of posts. Specifically, March 2023 saw a drastic uptick in posts while the average likes ratio actually decreased quite significantly. What does this tell us? Well, if the goal is to increase user engagement, simply increasing the quantity of posts being generated and displayed in the newsfeed probably isn't the best course of action.

More importantly, analyzing the distribution of the different post subtypes as well as the popularity of those posts measured by likes (and to a lesser degree comments) paints a pretty clear picture. The newsfeed seems to be dominated by Level posts with Collection posts being third in the overall distribution. These counts dwarf those of the other types such as Streak, Goal, Text Post, and Career Track Certificate. However, if one hovers over the bars in the popularity by likes visualization, we see a nearly complete reverse in the average likes per post. Text Posts have close to a 1-to-1 ratio of receiving likes (0.96 to be exact) which is followed by Goal (0.4796), Career Track Certificate (0.3611), and Streak (0.1711). As stated earlier, the prominent post types of Level and Collection both have between a 0.01 and 0.02 likes-per-post value. This ratio was not originally intended for this project, but I decided to add it in as I deemed it an important KPI to maximize the user experience. Not only do we want more users engaging with the Newsfeed, but we want those engagements to be meaningful and identifying the types of posts that have high interaction ratios compared to low ones are vital for this.

Level and Collection posts dominate the newsfeed while receiving very insignificant levels of interaction. One potential solution is to make those post types more enticing to users. However, seeing just how much more they are present in the newsfeed compared to others, a better solution would probably be to reduce their overall quantity. This can be done by allowing users to filter the

types of posts they want to see in their newsfeed (thereby increasing the visibility of posts they may want to interact with) as well as potentially aggregating these types of posts to show once per week (encompassing all of the Level and Collection achievements obtained during the week).

Furthermore, if a desired threshold was identified for the ideal interaction ratio per post, specific steps could be taken for each subtype to increase the interaction they bring. The text posts are by far the most successful and they also happen to be the only manual post types. Perhaps giving the users, the option to present other post types as manual could add some of that personality that other users seem to enjoy interacting with. As 365DataScience is largely an education site above all, completing courses and obtained course certificates are a big deal for users. Perhaps rather than automatically posting that a user has obtained one of these certificates, the site can allow users to add their thoughts in the form of text to those same posts (can double as reviews for each course). This would stimulate user engagement to posts outside of the typical text posts (usually questions and answers). After restructuring the newsfeed with a focus on quality over quantity, the 365DataScience team could explore the possibility of incentivizing user engagement with the service by potentially offering xp or coin rewards (the main "currencies" used by the site).

Lastly, it is paramount to choose the correct Key Performance Indicators (KPIs) to track when working towards a certain strategy. Similar to adding "interaction ratio" to measure the overall "quality" of post types, additional metrics that may be useful to track include the number of users visiting the newsfeed per month and the amount of time they spend there. As can be seen in the image to the right, navigating to the newsfeed involves opening a drop-down menu which is completely separate from those that most users likely spend the large majority of their time in (Learn, Projects, and Home). It would be interesting to understand how many users are actually viewing the newsfeed over time as finding ways to highlight this feature should also increase passive engagement.