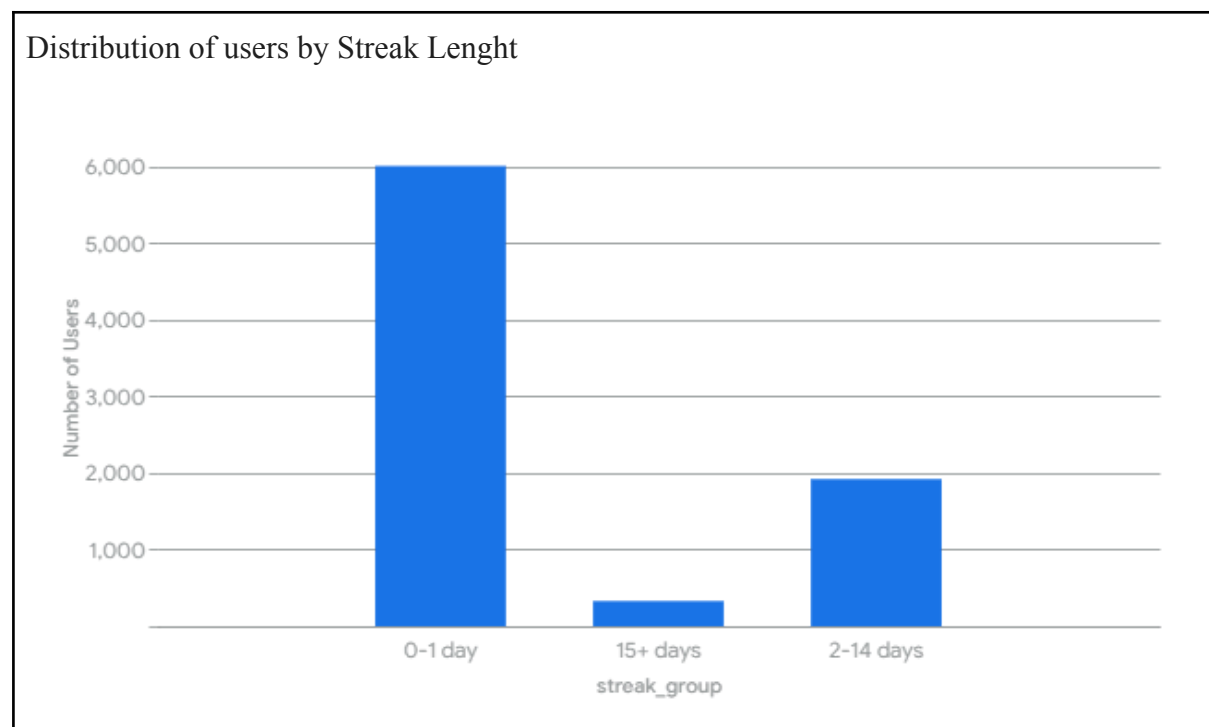


Assignment 3.4 MBAN1

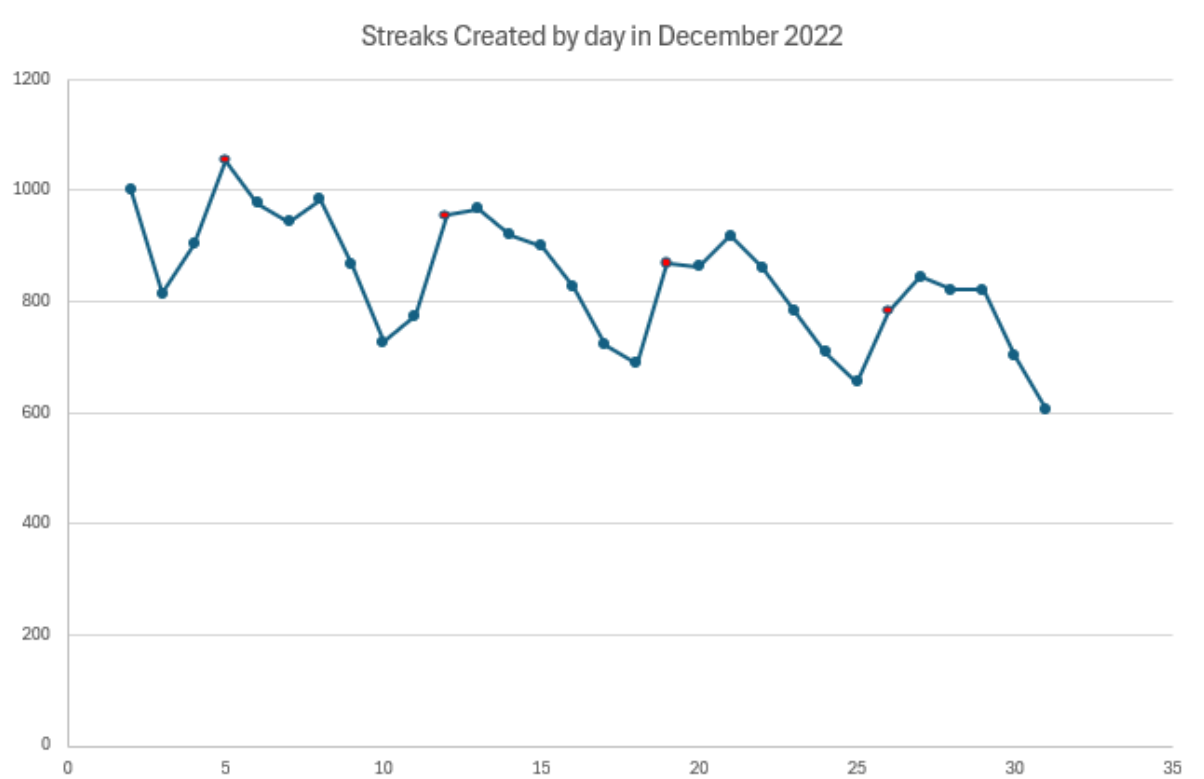
Lucas Sotkovszki | November 26th 2024

The online learning platform currently shows a concerning trend of low user engagement, this is indicated by the low number of users with Streaks, which serves as a key indicator of user retention. The data shows that a significant part of our user base has a maximum streak length of 0, indicating inconsistent use of the platform.

To understand better the different users of our platform, 3 groups were identified. Of those whose maximum streak is 0 or 1, 72.75% of the users, those who have between 2 days and 2 weeks of usage, and those with more than 15 days of consecutive usage of the platform are 3.98%. The median streak length is 0 days, which supports that the platform is failing to engage the users on consecutive days.



In addition, a distinct weekly pattern is observed, with a peak in streak creation of distinct users at the beginning of each week, suggesting that Mondays serve as a trigger for user engagement. We can see that there are a lot of streaks created on Mondays, represented by red dots in the chart, and peaking on Tuesdays and Wednesdays, when it starts to decline. This is important to understand consumer behavior and how they interact with our product. Another insight from this data is that the graph goes downwards, indicating that in general there are fewer and fewer streaks created during December 2022.



Business Problem and Hypothesis

The frequency of streaks of 0 or 1 day means a significant challenge to the long-term growth of our business. Low engagement means reduced user retention and potentially lower revenue.

Low user engagement may stem from a combination of factors, including a lack of perceived value in consistent platform use, usability challenges hindering seamless interaction, difficulties in integrating platform usage into daily routines, and content that is not adequately aligned with user needs or learning preferences.

Recommendations

Once we identify the users that have the least and the most engagement with our platform, we need to identify who these customers are with different personas to know how to target them. In terms of the users with a lot of streaks, we need to know what makes them go back to the app every day and what features should be improved to improve their overall experience.

However, more important is to get feedback from those users that are not coming back. It is crucial to know why they are not engaging with the app, and once we understand why we can improve the user experience. For instance, the streaks of “dopamine” that the users are supposed to feel when they return may not be rewarding enough, and the streaks design should be changed.

Moreover, by the weekly pattern, streak creation declines as the week goes by, therefore increasing the number of notifications or reminders during those days might increase the engagement users have with the platform.

Appendix



I am a business analytics for structure data masters student I need u to help me with the logic of an assignment and the approach I should take and help me brainstorm. I want you to act as a professional business analyst, so be extremely critical with your approaches. In the file there is some numerical findings I got from SQL and Excel.

Users_with_streaks

CSV

Show code <>



Analysis complete ▾

I'll analyze the data you provided and offer a breakdown of the logic, potential approaches, and areas to brainstorm to help you understand the assignment better. I'll approach this as a professional business analyst, providing constructive criticism and ensuring the chosen methods are robust and relevant to your task.

It seems like the dataset needs some cleaning and further exploration to understand the underlying patterns and relationships. Let me refine the data for better analysis.

Based on the provided information, here's a breakdown of the logic, a potential approach, and brainstorming points to help you with your assignment:

Understanding the Data

The dataset seems to contain information about user streaks, possibly in the context of a mobile app, online game, or learning platform. Here's what the key columns likely represent:

- `user_id` : A unique identifier for each user.
- `max_streak_length` : The longest consecutive days a user engaged in a specific activity.
- `Unnamed: 5` : This column's meaning isn't clear from the provided data. It might represent a specific day within a streak, a category of streaks, or some other metric related to user engagement.

Potential Approaches

Given the limited information about the dataset and your specific assignment goals, here are a few general approaches you could consider:

Ask Gemini



(which you haven't provided yet).

Brainstorming Points

- **Clarify the meaning of** `Unnamed: 5` : Understanding this column is crucial for meaningful analysis. Can you provide more context or data associated with it?
- **Gather additional data:** Depending on your analysis goals, you might need more information about the users, such as demographics, app usage patterns, or dates of activity.
- **Define specific research questions:** What are you trying to achieve with this analysis? Are you looking for patterns, making predictions, or testing hypotheses?
- **Consider the business context:** How could the insights from your analysis be used to improve the app, game, or platform? For example, could you identify factors that contribute to longer streaks and use this knowledge to increase user engagement?

Critical Considerations

- **Data Cleaning:** Ensure the data is accurate and complete before conducting any analysis.

Ask Gemini



the first columns is the `user_id`, and the second question is the number of days that user has been on a streak in a platform like dulingo for example

Show code <>



Analysis complete ▾

I'll reload the data with the provided headers and then proceed with the analysis.

Now that the data has been processed, I'll start the analysis by examining the distribution of streak lengths. I'll also look for any outliers or other patterns that may be relevant to the business.

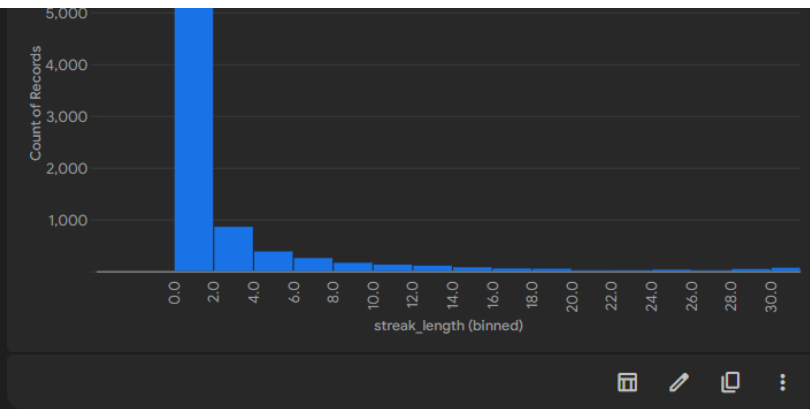
Here's a summary of my findings from the data:

Summary Statistics:

- The minimum streak length is 0 days, while the maximum is 31 days. This is

Ask Gemini





Based on my analysis, here are some key considerations for the business based on the streak patterns:

1. **User Segmentation:** Segment users based on their streak lengths to tailor interventions and communications. For instance, users with long streaks might benefit from loyalty programs or exclusive content, while users with shorter streaks might need reminders or motivational messages to encourage continued engagement.
2. **Gamification:** Incorporate elements of gamification, such as badges,

Ask Gemini



i did this code confirm that this returns the number of streaks created by day during the month

```
SELECT  
day(ss.streak_created), count(distinct user_id),  
MAX(streak_length) AS max_streak_length...
```



Analysis complete

You're absolutely correct! Here's why that code effectively returns the number of streaks created by day during the month:

- **streaks_new table:** This temporary table, created earlier in your code, contains the calculated `streak_length` for each user and day.
- **user_streaks_sql table:** This appears to be your source data table containing information about user streaks.
- **INNER JOIN :** This combines `streaks_new` with `user_streaks_sql` using the common column `user_id`, ensuring you're working with users who have streak data.
- **DAY(ss.streak_created) :** This extracts the day of the month from the `streak_created` column.

Ask Gemini

- `GROUP BY DAY(ss.streak_created)` : This groups the results by the day of the month, so you get aggregated data for each day.
- `COUNT(DISTINCT user_id)` : This counts the number of *unique* users who created a streak on each day.
- `MAX(streak_length) AS max_streak_length` : This calculates the maximum streak length observed for each day.
- `HAVING max_streak_length > 0` : This filters out any days where the maximum streak length was 0, effectively showing you only the days with streaks created.
- `ORDER BY day(ss.streak_created) ASC` : This orders the results in ascending order of the day of the month, giving you a chronological view.

Therefore, this code snippet accurately returns the number of streaks created each day during the month by counting distinct users with streaks and filtering out days with no streaks.

Sources and related content ^ ⓘ

Users_with_streaks

Ask Gemini

