Milestone 2

Group Members

Alden Mo

Ao Shen

Yunzhe Yu

Questions

1. ### Show most viewed channels in descending order

   This gives us a view of their favorite channels regardless of subscription or not for a recommendation

   Group by, COUNT

2. ### Show the most viewed channels that you aren't subscribed to

   Subscribed channels automatically will show up on the feed, by knowing their most watched unsubscribed channels we can recommend those videos.

   Minus

3. ### Get the day with the most videos watched

   This lets us look at the day with the most videos watched, as it’s most videos and not minutes it gives us a look at (most likely) their favorite types of short form content

   Max, Subquery

4. ### Get the day with the most minutes of video watched

   This allows us to determine the single day that the user spent the most time on the site watching videos, allowing us to personally analyze the videos watched on that day

   Sum, Subquery

5. Show the average amount of time spent watching videos every day of the week

   This allows us to determine the most common day the person watches a video to determine their usual schedule which can be used to feed short form content for days with less watch time and long form content for days with longer watch time.

   Sum, Group by weekdays

6. What are the most common channels/genres to comment on

   This will let us determine a user's engagement with specific content, as commenting usually shows a level of engagement more than watching the video.

   Order by, Group by

7. Find the day with the greatest or least amount of videos watched vs. searches

   This will help us to find the difference between videos watched and searches.

   Assuming that there is a significant difference between the number of videos watched and the number of searches performed on a given day, we can draw some conclusions about the most desired content and topics at that time.

   Max(), Min(), Group by

8. Determine the most common (non-determiner) word for watched videos and/or comments

   This allows for grouping by keywords for the user which can be used with the aid of machine learning of human analysis to determine their most common keywords.

   substr, group by, count

9. Amount of watch time per genre

   We can use this result to find which genre is most popular, and acconding this, user can get more related genre tweets.

   count(*), max(), group by

10. Change in genres and/or watch time around 2020 summer (start of the pandemic)

    This question can be used to determine any changes in sentiment during the pandemic for the user represented in the data, this allows us to determine whether or not to bias the data after 2020 more heavily than data before as the pandemic created record watch times.

    where, count, intersect

11. What are the most common sentiments expressed in YouTube video comments

    The question will use the knowledge of text analysis to identify the sentiments expressed such as positive or negative. It can help YouTubers and marketers know more about the users and adjust their content and algorithm.

    Group by

Teamwork:

We collaborated as a team to brainstorm and come up with a list of more than 10 questions. We then analyzed each question's relevance and narrowed down the list to ten questions that would provide meaningful insights and demonstrate our learning from the course CNIT 372. Each team member contributed a lot to the process, bringing in their expertise in data analysis, visualization, and text analysis to shape the questions.