PROFESSIONAL

Demographic Filtering

WhiteHat Jr
Live Online Coding for Kids

INSTRUCTIONS:

Goal of the Project:

In class 139, we performed demographic-filtering on the data.

In this project we are going to understand the data that we got in the last project (article's data), analyse it and then perform demographic filtering on it!

****Note - This is continuation of Project C-138. Complete that project before this project.**

Getting Started:

Open the Colab in which you did the project of C-138 and re-run all the cells.

Specific Tasks to complete the Project:

1. In **shared_article.csv**, we have 2 types of articles:
    - **CONTENT SHARED -** Available for user
    - **CONTENT REMOVED -** Removed from the platform and not available for user
2. Remove all article rows that have type as **CONTENT REMOVED**.
3. We have the following types of events in **user_interactions.csv**:
    - **VIEW**
    - **LIKE**
    - **BOOKMARK**
    - **FOLLOW**
    - **COMMENT CREATED**
4. Count the number of total events (Sum of Views, Likes, Bookmarks, Follows and Comments Created) for all the articles and create a new column in **shared_article.csv** data.

Note - It might take some time (up to 10-15 minutes) to process all the data by Google Colab.

5. Sort the DataFrame with the total events column that we just created in descending order.
6. Top 10 articles can now be considered as the demographically filtered recommendations.

Submitting the Project:

1. Copy the link to the Google Colab in the Student Dashboard.

Hints:

1. To apply a condition to the dataframe, the following syntax can be used:

```
 total_views = df2[(df2["contentId"] == df1_row["contentId"]) & (df2["eventType"] == "VIEW")].shape[0]
```

REMEMBER... Try your best, that's more important than being correct.
After submitting your project your teacher will send you feedback on your work.

———— xxx ———— xxx ———— xxx ———— xxx ———— xxx ————