

Interview Question: Data Analysis and Manipulation

You are provided with CSV files containing movie and rating data. The task is to perform various insights and analyses on this dataset. The process involves:

Task Overview:

1. Data Import:

- Set up a PostgreSQL database. (You can setup a free PostgreSQL instance from [Render](#))
- Create tables to store movie and rating data from the CSV files
(You can download the CSV files from [here](#))
- Import the CSV data into the respective tables in the PostgreSQL database.

2. Insights and Analysis:

- Use any scripting language of your choice (e.g., Python, Javascript, etc.) to perform the following insights:

a. **Top 5 Movie Titles:** Sort and print the top 5 movie titles based on the following criteria:

- Duration
- Year of Release
- Average rating (consider movies with minimum 5 ratings)
- Number of ratings given

b. **Number of Unique Raters:** Determine and print the count of unique rater IDs.

c. **Top 5 Rater IDs:** Sort and print the top 5 rater IDs based on:

- Most movies rated
- Highest Average rating given (consider raters with min 5 ratings)

d. **Top Rated Movie:**

- Find and print the top-rated movies by:
 - Director 'Michael Bay',
 - 'Comedy',
 - In the year 2013
 - In India (consider movies with a minimum of 5 ratings).

e. **Favorite Movie Genre of Rater ID 1040:** Determine and print the favorite movie genre for the rater with ID 1040 (defined as the genre of the movie the rater has rated most often).

f. **Highest Average Rating for a Movie Genre by Rater ID 1040:** Find and print the highest average rating for a movie genre given by the rater with ID 1040 (consider genres with a minimum of 5 ratings).

g. **Year with Second-Highest Number of Action Movies:** Identify and print the year with the second-highest number of action movies from the USA that received an average rating of 6.5 or higher and had a runtime of less than 120 minutes.

h. **Count of Movies with High Ratings:** Count and print the number of movies that have received at least five reviews with a rating of 7 or higher.

3. Language and Tools:

- You can use any scripting language and appropriate tools/libraries to accomplish the tasks efficiently.

4. Submission:

- Please present your solution by uploading it to GitHub and share the repository link with us.
- Additionally, include any assumptions made during the process and comments to explain the logic of your code.

Note: Focus on readability, efficiency, and clarity in your solution. The goal is to assess your ability to handle data manipulation tasks, database querying, and problem-solving skills.