|  |  |
| --- | --- |
| **User Table** |  |
| user\_id | pk |
| user\_name |  |
| first\_name |  |
| last\_name |  |
| address |  |
| city |  |
| state |  |

|  |  |
| --- | --- |
| **Movie Table** |  |
| movie\_id | pk |
| movie\_name |  |
| movie\_studio |  |
| movie\_score |  |

|  |
| --- |
| **Events Table** |
| user\_id |
| movie\_id |
| date\_watched |

The primary key for the events table is the concatenation of user\_id, movie\_id, date\_watched. A user could have watched one or more movies, one or more times. A movie could be watched by one or more users, one or more times.

Instructions: Write a SQL query that would provide the answer for each of the following questions:

1. Prepare a list of user\_names with the number of movies that they have watched.
2. Now adjust that query to include only a list of user\_names that have watched more than 5 distinct movies.
3. Using the output from question number 2, create a column using the rank function to order the list by # of distinct movies watched by user\_name.
4. For each user\_name, provide the latest date that they watched any movie and the name of the movie they watched.
5. How would you include a column for the average number of movies per user, and then find each users variance from the mean of the population?
6. Prepare a list of the top 10 users with the highest average movie score (this is the average movie score of the movies they have watched). Include only the first and last name as a single field for each user.
7. For every month, give me the movie studio with the most watched film and how many times it was watched in the month.
8. For the month of December, what is the top watched movie by state (the movie with the most amount of views)?
9. Include a list of all movies and the # of times it has been watched by month in 2019. Also include 2 additional columns showing the Year-over-Year % change by month and the Month-over-Month % change by month.