Data Mining Assignment- Postgre SQL

TLP term- 2

October 2023

Instructions: For this assignment, it is assumed that you have already set up and populated the sakila database on your system. This fictional database is designed to mimic a DVD rental store, where you can perform various queries like renting films, tracking rentals, managing inventory, and handling customer information.The following questions should be addressed by utilizing suitable queries using the PostgreSQL command line with the sakila database.

1. ***(2 marks)*** The film table stores information about movies. By appropriately querying the film table, find the top 10 most popular films of all time, based on the number of rentals. Your output should have 2 columns containing the film_name and rental_duration respectively.
2. ***(4 marks)*** Which are the top 3 most popular films among customers who have also rented the film TEQUILA PAST? Your output should have 4 columns- film_id, film_name, rental_count (in descending order) and rating.
3. ***(6 marks)*** Calculate the total revenue generated by each film category (ex: Action, Drama, Sports, etc) and list the category_name along with their total_revenue, as well as the average_revenue_per_film in each category. Your output shall thus have 3 columns.
4. ***(4 marks)*** Use a nested select query to find the first name and the last name of customers who paid more than 10 dollars in any transaction. You have to look at the payment table and then match the customer_id with that in the customer table. How many such customers are there?
5. ***(4 marks)*** Determine the top 10 customers who have rented the most films. Your output should include their total rental count and the total amount they've spent on rentals.
6. ***(5 marks)*** Now, request a solution from a large language model (LLM) like ChatGPT or BARD for each of these problems. After obtaining the solutions, compare them with your solutions.
   (*Advice*: LLMs do not always generate an appropriate code required. Therefore, it is recommended to solve the questions in the order provided, first by yourself, and then prompt the LLM).