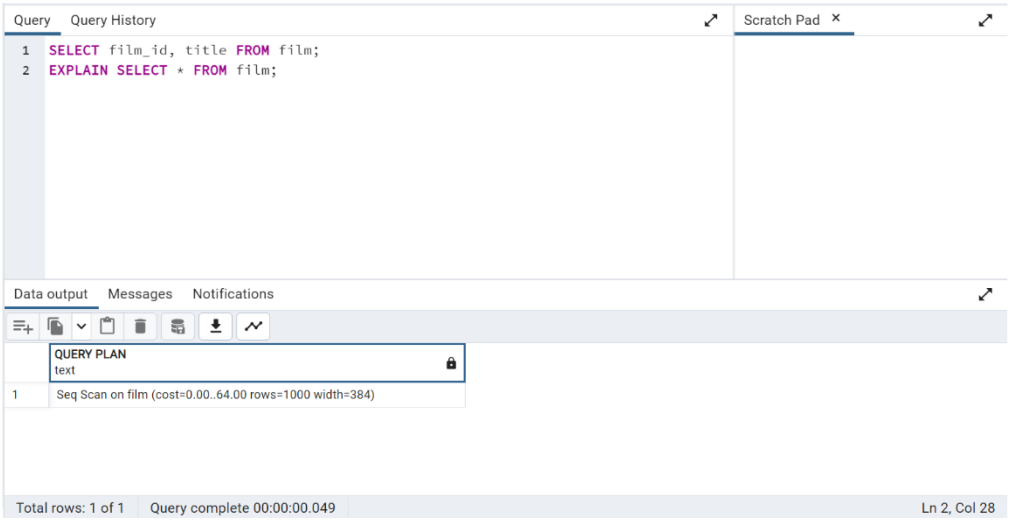
Juliana Lee

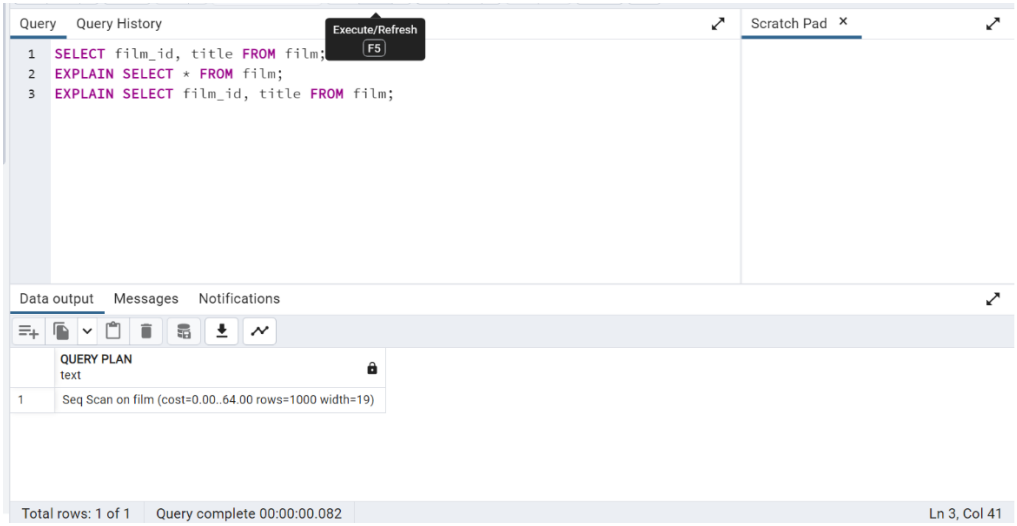
Task 3.4: Database Querying in SQL

1a) New query: SELECT film_id, title FROM film;

1b) Original:

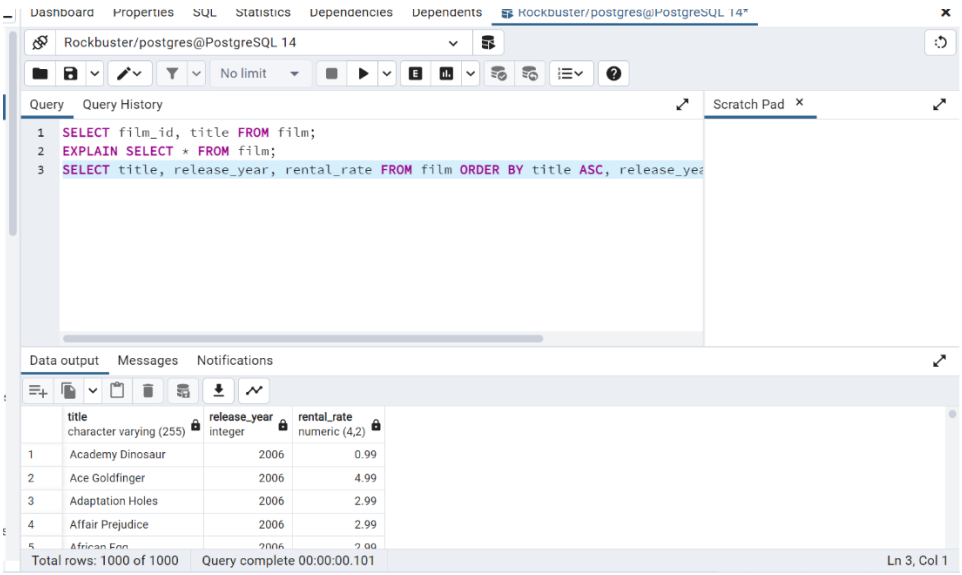


Revised:



The cost of both the original and revised query is the same (cost=0.00…64.00). The width value of the original is 384 and the width value of the revised is 19. The revised query targets the specific columns, processing the result faster. Therefore, the revised query would be more cost efficient than the original query.

Juliana Lee



2b) Extract the data output of your query into a CSV file for the film collection department to analyze in Excel. To do this, click the button "Save results to file": ✅

3a) Average rental rate SQL query: SELECT rating, AVG(rental_rate) AS average_rental_rate FROM film GROUP BY rating;

3b) Min and Max SQL query: SELECT rating, MAX(rental_rate) AS maximum_rental_rate, MIN(rental_rate) AS minimum_rental_rate FROM film GROUP BY rating;

4a) Outline:

1. **Extract: Collect all the necessary data from the Android app (data engineers or the app team)**
2. **Transform: Convert the extracted data to standardized data types to determine table and column structures and to find the keys to build relationships with tables. (data engineers/data analysts)**
3. **Load: Load transformed data into new database. (data engineers)**

4b) Problems: If you start analyzing data before it's been loaded into the data warehouse, you are looking at incomplete data. Incomplete data hinders the ability to see relationships because it has not been filtered out. Therefore, the information is not yet meaningful. Also, analyzing data before it's been filtered out would be inefficient and difficult.