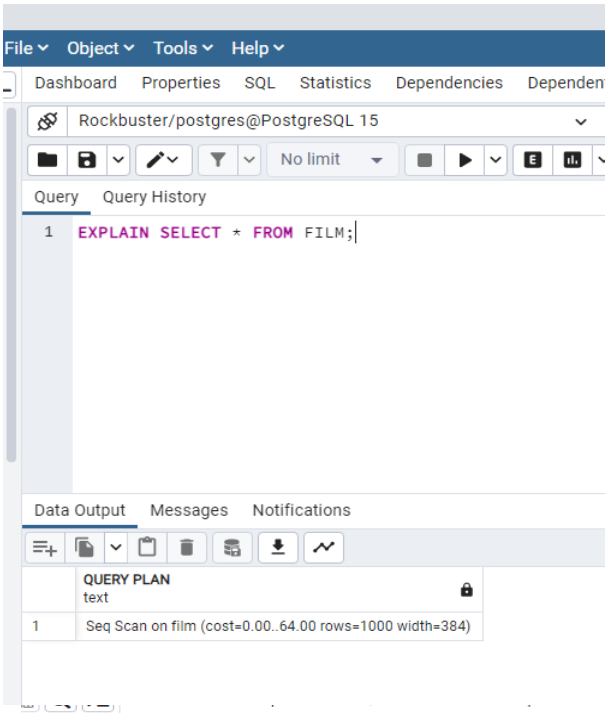
3.4: Database Querying in SQL

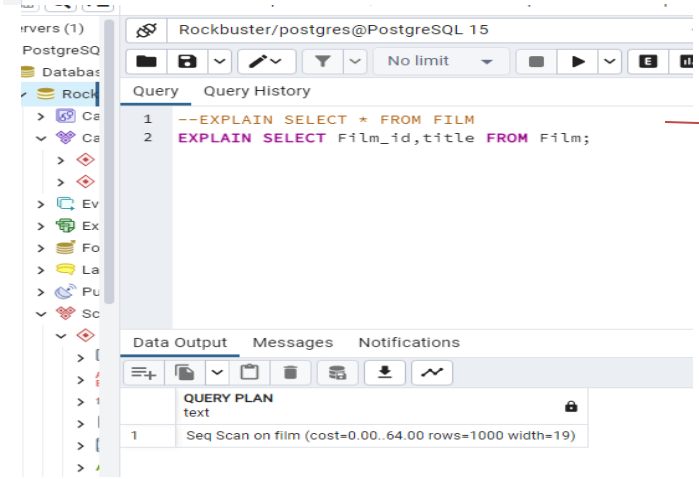
Refining Your Query: You need to get some data from the “film” table and decide to use the query SELECT * FROM film.

- You realize that only the “film_id” and “title” columns are needed.Write a new query that selects only those 2 columns.
- Compare the cost of the original query and the revised query, and write a few sentences explaining the comparison. Can you suggest any ways to optimize this query?



Select * Query Statement:

Seq Scan on film (cost=0.00..64.00 rows=1000 width=388)



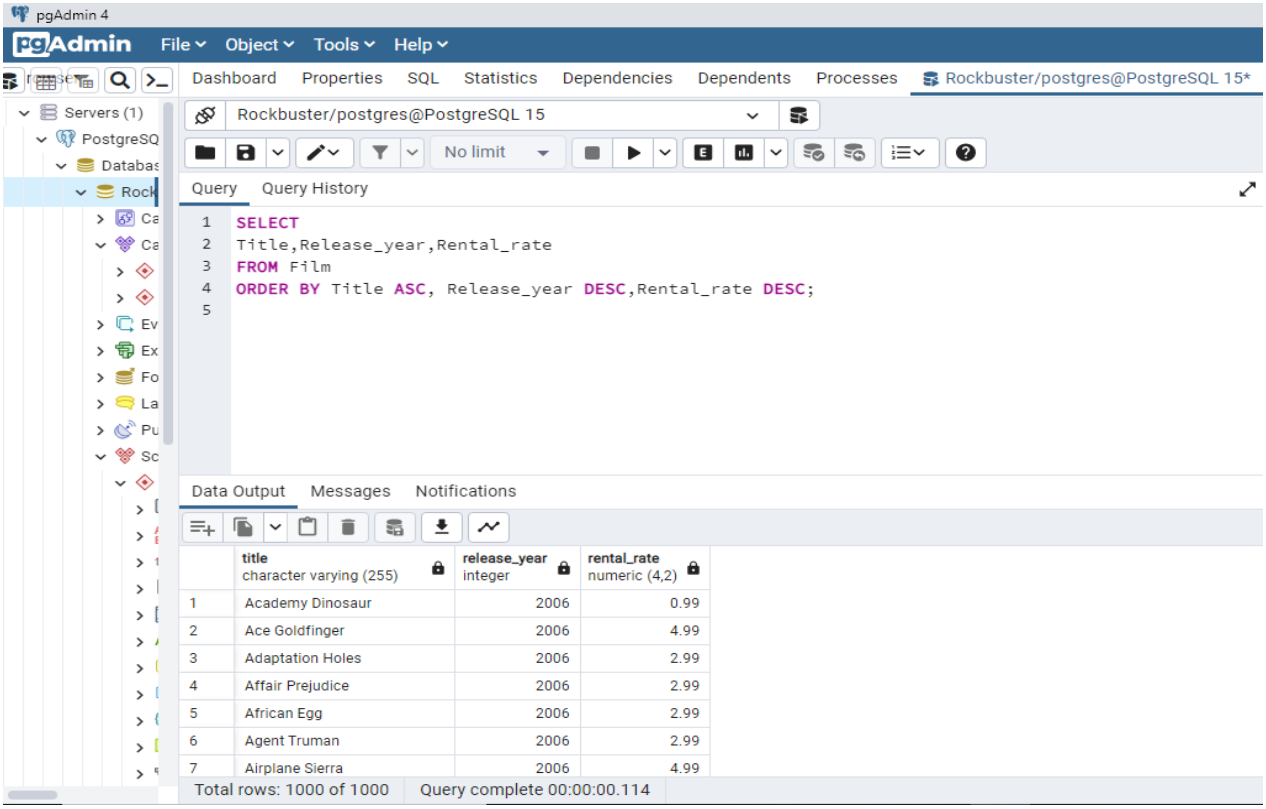
Select film_id,title Query Statement:

Seq Scan on film (cost=0.00..64.00 rows=1000 width=19)

Answers: According to me, the cost of First query is more expensive than second one. Because the first one takes more time in checking all the rows and also width 384. The second one is only showing the result of two columns, which means less work or burden by RDMS ,faster result and cost efficient.

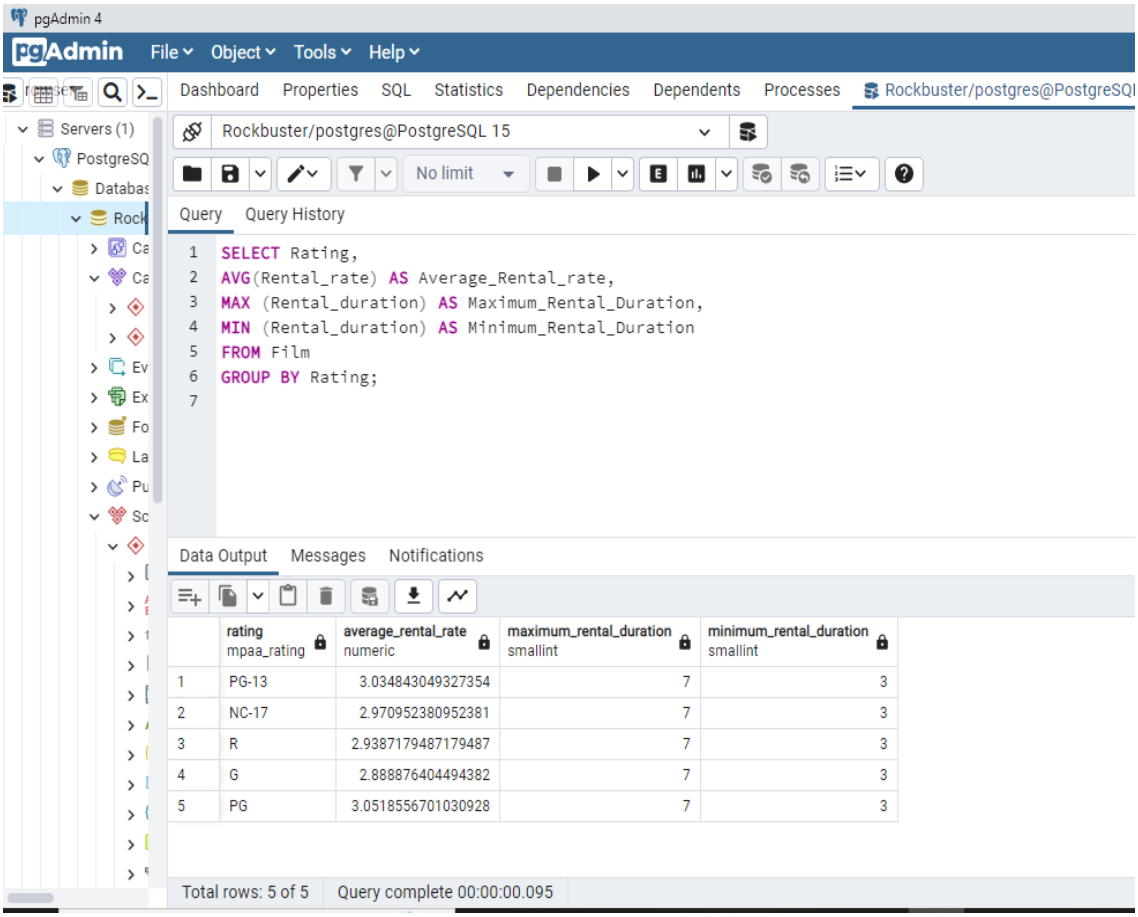
2. Ordering the Data:

- In the pgAdmin Query Tool, run a query that selects every film from the “film” table, with the movies sorted by title from A to Z, then by most recent release year, and then by highest to lowest rental rate.
- 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”:



3. Grouping Data: The strategy department has asked you the questions below. Write a SQL query to retrieve the correct answers, then extract your results as a csv file.

- What is the average rental rate for each rating category?
- What are the minimum and maximum rental durations for each rating category?



4. Database Migration: Your team has decided to use an external tool to collect data on user behavior in the new Rockbuster Android app. Data collected from this new source will need to be loaded into the data warehouse before you can analyze it.

- Can you outline the procedure for migrating the data and who will be responsible for it?
  Answers:
  The procedure for migrating the Data is done by ELT Process.
  1.Extract: First step is to collecting the data from Data source. Here we will collect data from new Rockbuster Android App.
  2.Transform: In this step, the extracted data is converted into another Format. Here may be user behavior experience converted into rating system, User age, telephone number etc.
  3.Load: at this point, the transformed data is loaded into new Database.

  This all task is performed by Data Engineers.

- What problems do you foresee if you start analyzing the data before it's been loaded into the data warehouse?
  - The Quality of Data like unorganized data, Clarity, too much Information etc.
  - Time consuming
  - Changing the format of Data like Behavioral data into numerical format etc.