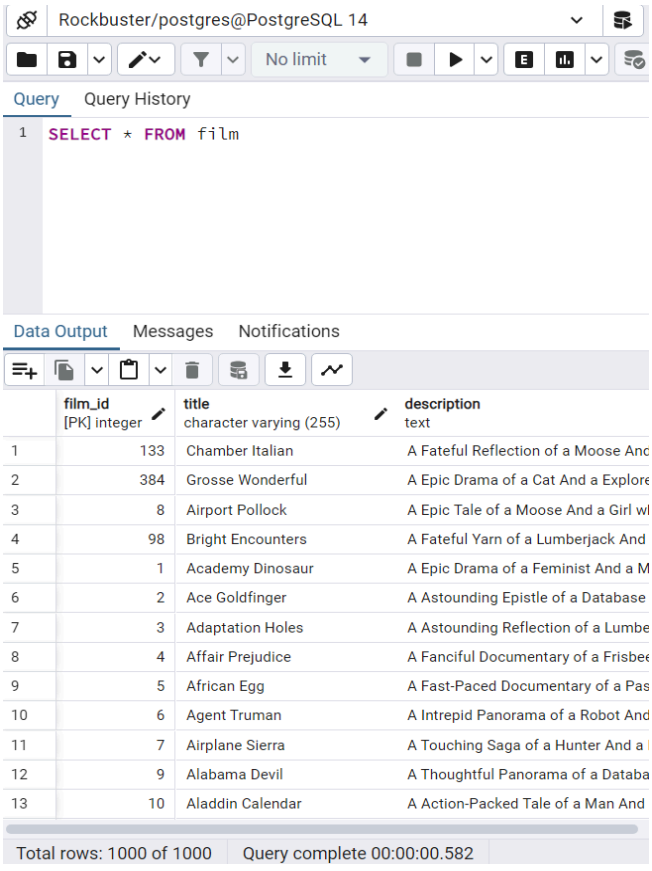
3.4 - Database Querying in SQL
Katherine Lecce

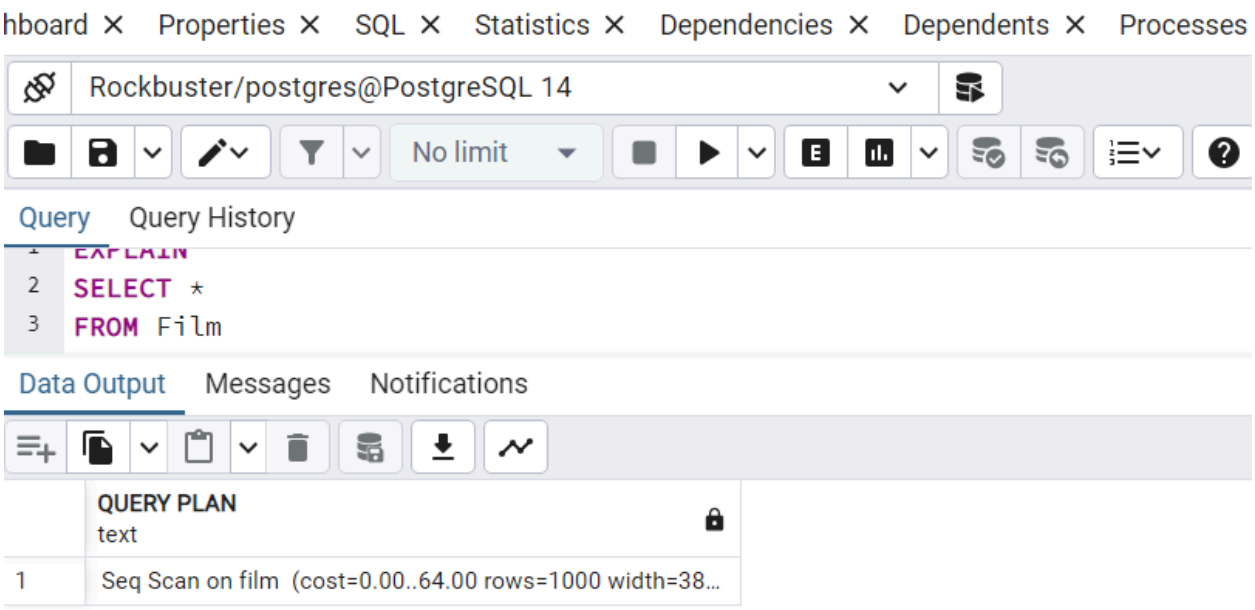
Directions:

1. 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?** We can use Query Optimization to write an ‘Explain’ query this function will help us find the best way to execute the system to find the best way to write the query for the column film_id and the best query plan to use to retrieve the data from the ‘film’ table. The ‘Query Plan’ output gives us information about dimensions of the query and potential errors that might occur.

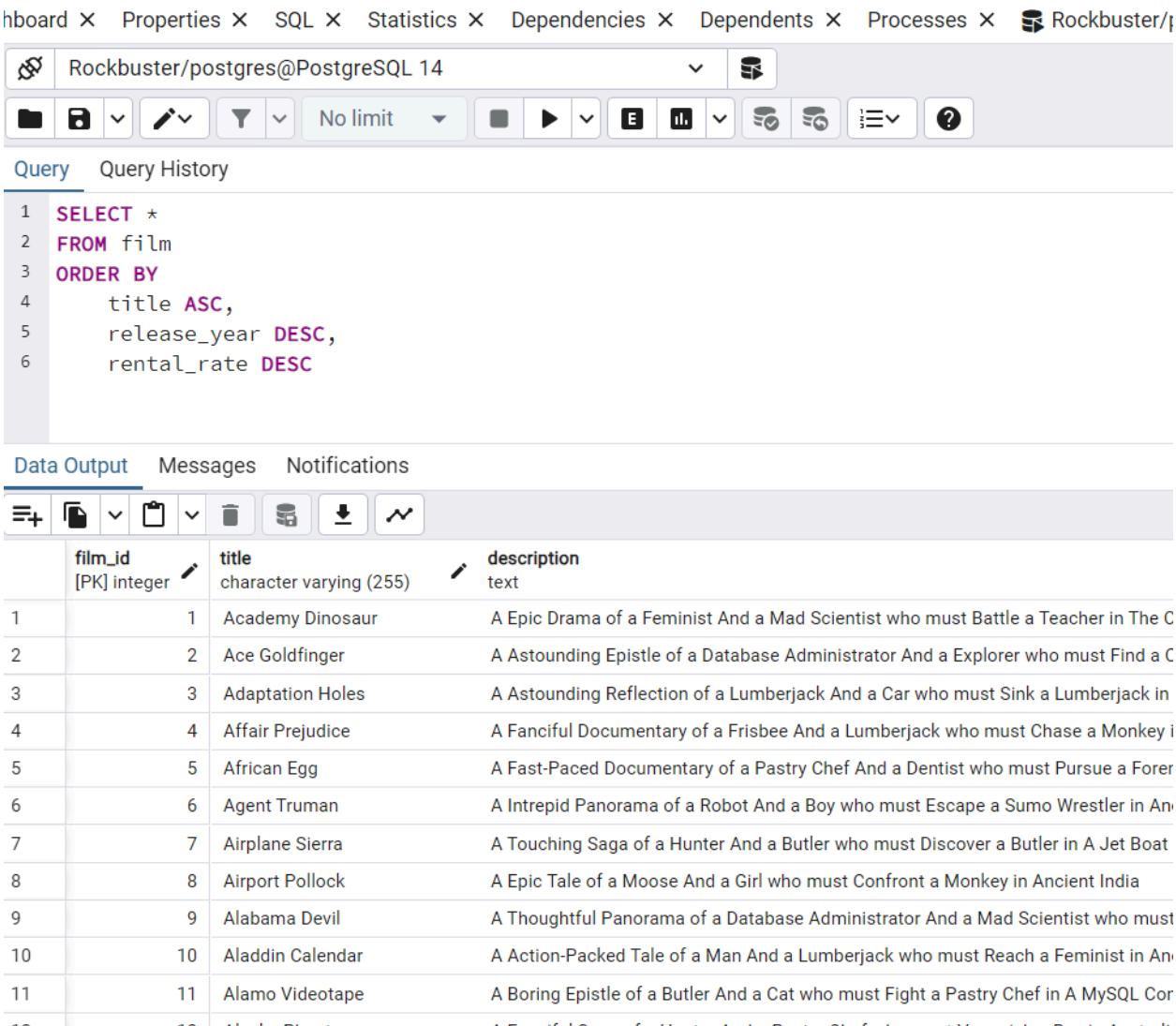
Original Query:
→



New Query



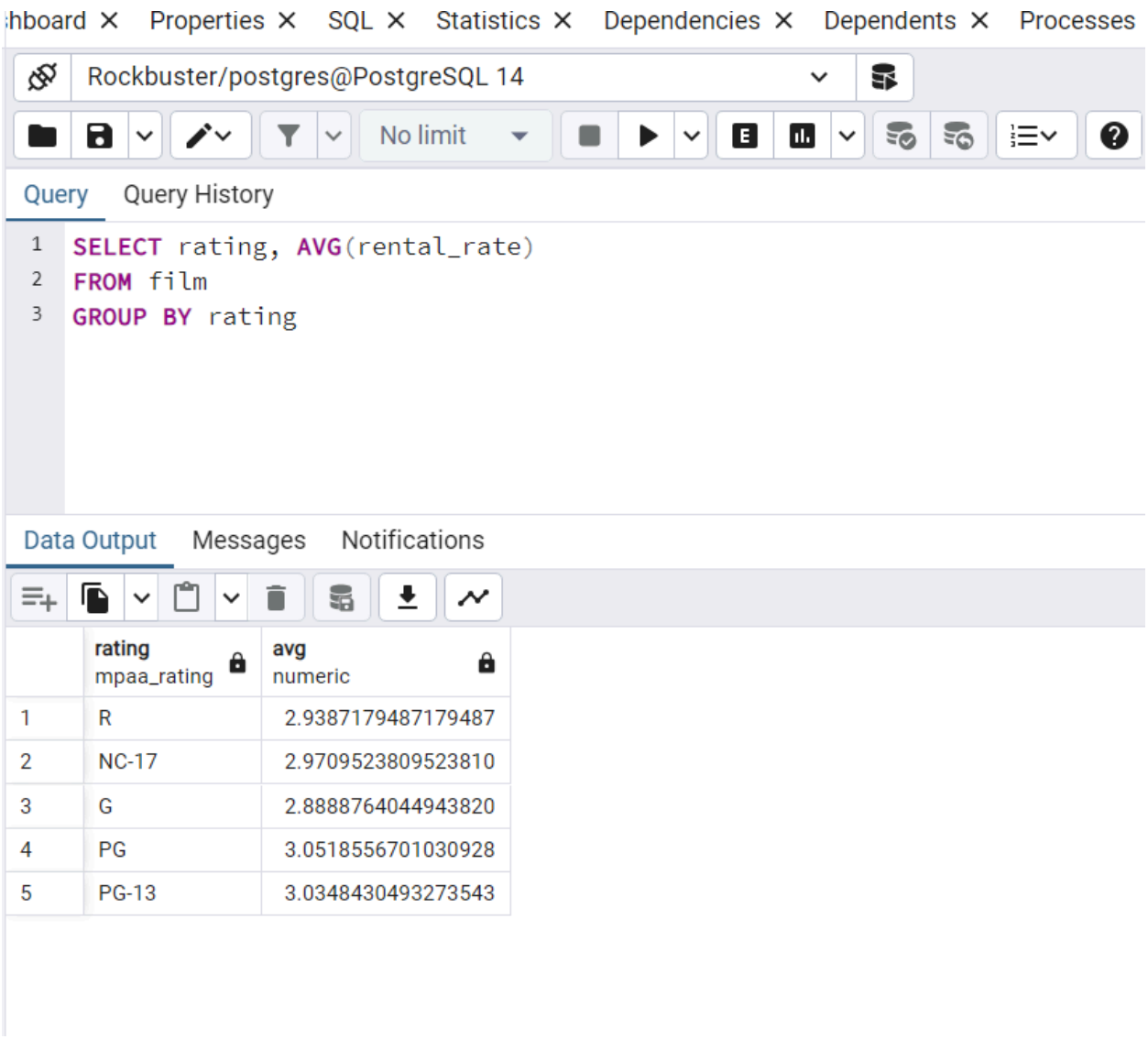
2. Ordering the Data



	film_id [PK] integer	title character varying (255)	description text
1	1	Academy Dinosaur	A Epic Drama of a Feminist And a Mad Scientist who must Battle a Teacher in The C
2	2	Ace Goldfinger	A Astounding Epistle of a Database Administrator And a Explorer who must Find a C
3	3	Adaptation Holes	A Astounding Reflection of a Lumberjack And a Car who must Sink a Lumberjack in
4	4	Affair Prejudice	A Fanciful Documentary of a Frisbee And a Lumberjack who must Chase a Monkey i
5	5	African Egg	A Fast-Paced Documentary of a Pastry Chef And a Dentist who must Pursue a Forer
6	6	Agent Truman	A Intrepid Panorama of a Robot And a Boy who must Escape a Sumo Wrestler in An
7	7	Airplane Sierra	A Touching Saga of a Hunter And a Butler who must Discover a Butler in A Jet Boat
8	8	Airport Pollock	A Epic Tale of a Moose And a Girl who must Confront a Monkey in Ancient India
9	9	Alabama Devil	A Thoughtful Panorama of a Database Administrator And a Mad Scientist who must
10	10	Aladdin Calendar	A Action-Packed Tale of a Man And a Lumberjack who must Reach a Feminist in An
11	11	Alamo Videotape	A Boring Epistle of a Butler And a Cat who must Fight a Pastry Chef in A MySQL Cor

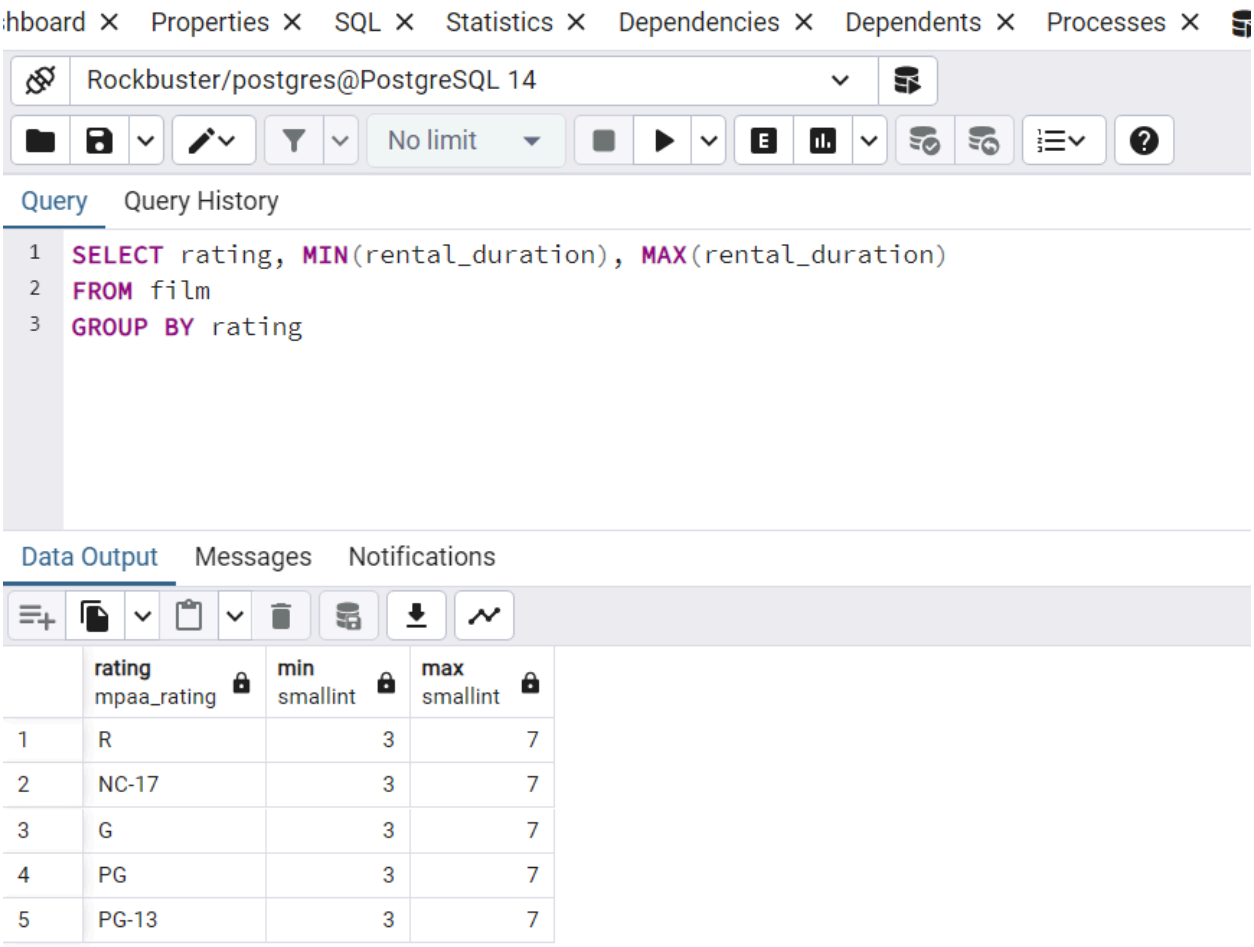
3. Group Data

Average rental rate for each rating category:



	rating mpaa_rating	avg numeric
1	R	2.9387179487179487
2	NC-17	2.9709523809523810
3	G	2.8888764044943820
4	PG	3.0518556701030928
5	PG-13	3.0348430493273543

Minimum and maximum rental durations for each rating category



	rating mpaa_rating	min smallint	max smallint
1	R	3	7
2	NC-17	3	7
3	G	3	7
4	PG	3	7
5	PG-13	3	7

4. Database Migration

1. **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?**

1. Extract: Extract the data from the the external data from the new source for the behavior analysis from customers → Done by the Data Engineers
2. Transform: Transform the external data by cleaning, validating, formatting, calculating, and summarizing any data points from the external data that is needed for the analysis.
3. Load: After data is transformed and uploaded to the data warehouse, the data will be formatted in the appropriate format.

→ **What problems do you foresee if you start analyzing the data before it's been loaded into the data warehouse?** Without properly transforming the data beforehand, the project will take a lot longer and the data will be very disorganized. This is the first issue; it would also pose the risk of not being the correct format making it hard to use or unusable all together. This can also pose the risk of the data being incomplete and inaccurate which will result in wrong decision making, loss of profit, and overall poor business performance throughout the app.