Step 1

EXPLAIN

SELECT *

FROM film

Seq Scan on film (cost=0.00..64.00 rows=1000 width=662)

EXPLAIN

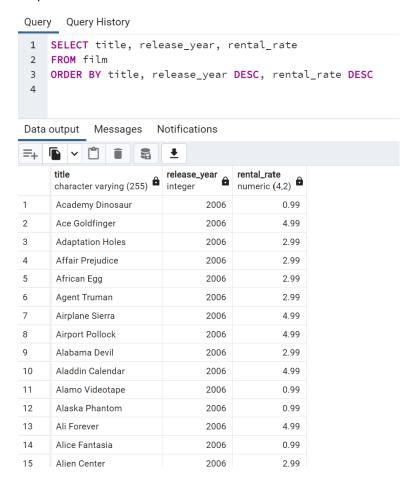
SELECT film_id, title

FROM film

• Seq Scan on film (cost=0.00..64.00 rows=1000 width=520)

The queries seem to have similar costs, but it is still better to choose the query that gives the specific information that you need.

Step 2



Step 3

Query Query History

- 1 SELECT rating, AVG(rental_rate) AS average_rental_rate,
- 2 MIN(rental_duration) AS minimum_rental_duration,
- 3 MAX(rental_duration) AS maximum_rental_duration
- 4 FROM film
- 5 **GROUP BY** rating

Data output Messages Notifications					
	rating mpaa_rating	average_rental_rate numeric	minimum_rental_duration smallint	maximum_rental_duration smallint	
1	R	2.9387179487179487	3	7	
2	NC-17	2.970952380952381	3	7	
3	PG	3.0518556701030928	3	7	
4	G	2.888876404494382	3	7	
5	PG-13	3.034843049327354	3	7	

Step 4

In the ETL process, a data engineer would extract data from multiple data sources, convert it into another format, and then load the data into a new database. If data analysis is done before this process happens, there is an increased chance of working with dirty data, and it would cost much more time and money.