

3.4: Database Querying in SQL

1. Refining Your Query:

- `SELECT film_id, title`
`FROM film`
- Compare the cost:

The image shows two side-by-side screenshots of a database client interface, likely pgAdmin, comparing the execution of two SQL queries. Both queries are executed against a database named 'Rockbuster/postgres@PostgreSQL 14'.

Left Screenshot:

- Query:**

```
1 EXPLAIN
2 SELECT film_id, title
3 FROM film
```
- Query Plan:** Seq Scan on film (cost=0.00..64.00 rows=1000 width=19)

Right Screenshot:

- Query:**

```
1 EXPLAIN
2 SELECT *
3 FROM film
```
- Query Plan:** Seq Scan on film (cost=0.00..64.00 rows=1000 width=384)

The results show that the cost of running the two queries is the same. However, the second query only targets the intended columns, hence the width of the processed data is shorter than the first. The difference would be magnified when large amounts of data are processed thereby shortening the cost. Optimizing the query can include adding a LIMIT or grouping the data according to the desired outcome.

2. Ordering the Data:

The image shows a screenshot of a database client interface displaying a query with an ORDER BY clause and its results.

Query:

```
1 SELECT title, release_year, rental_rate
2 FROM film
3 ORDER BY title ASC, release_year DESC, rental_rate DESC
```

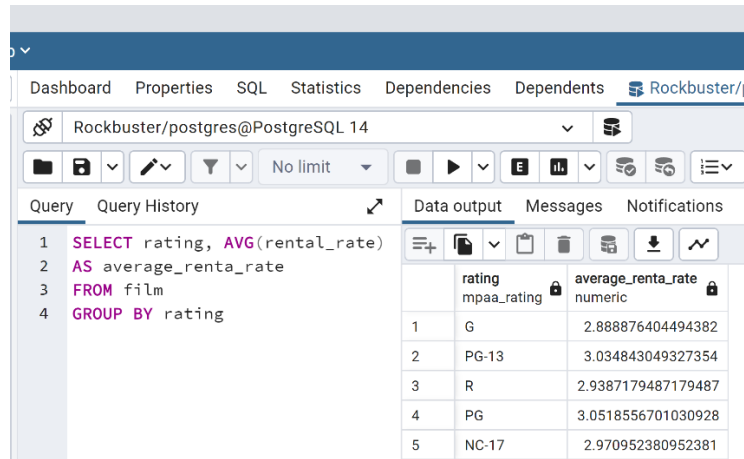
Data output:

	title	release_year	rental_rate
1	Academy Dinosaur	2006	0.99
2	Ace Guldfinger	2006	4.99
3	Adaptation Holes	2006	2.99
4	Affair Prejudice	2006	2.99
5	African Egg	2006	2.99
6	Agent Trueman	2006	2.99
7	Airplane Sierra	2006	4.99
8	Airport Pollock	2006	4.99
9	Alabamas Devil	2006	2.99
10	Aladdin Calendar	2006	4.99
11	Alamo Videolape	2006	0.99
12	Alaska Phantom	2006	0.99
13	All Forever	2006	4.99
14	Alice Fantasia	2006	0.99
15	Allen Center	2006	2.99

Total rows: 1000 of 1000 Query complete 00:00:00.110 Ln 3, Col 55

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?



The screenshot shows a SQL IDE interface with a query editor on the left and a results pane on the right. The query is as follows:

```

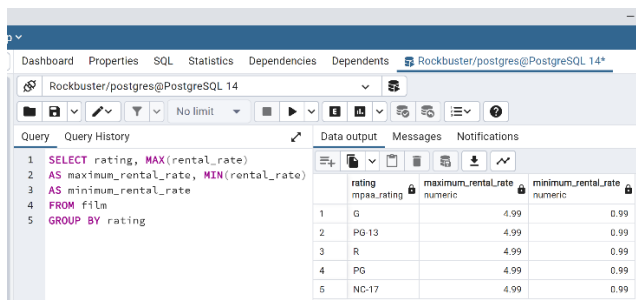
1 SELECT rating, AVG(rental_rate)
2 AS average_rental_rate
3 FROM film
4 GROUP BY rating

```

The results pane displays a table with two columns: 'rating' (mpaa_rating) and 'average_rental_rate' (numeric). The data is as follows:

rating	average_rental_rate
1	G
2	PG-13
3	R
4	PG
5	NC-17

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



The screenshot shows a SQL IDE interface with a query editor on the left and a results pane on the right. The query is as follows:

```

1 SELECT rating, MAX(rental_rate)
2 AS maximum_rental_rate, MIN(rental_rate)
3 AS minimum_rental_rate
4 FROM film
5 GROUP BY rating

```

The results pane displays a table with three columns: 'rating' (mpaa_rating), 'maximum_rental_rate' (numeric), and 'minimum_rental_rate' (numeric). The data is as follows:

rating	maximum_rental_rate	minimum_rental_rate
1	G	4.99
2	PG-13	0.99
3	R	4.99
4	PG	0.99
5	NC-17	4.99

4. **Database Migration Procedure:** Carried out by Data Engineers.

- Extract: the necessary information from the Android app
 - Transform: the extracted data is converted into another format.
 - Load: At this point the transformed data is inserted or loaded into the new database.
- b. Foreseeable problems include formatting issues and inclusion of irrelevant data.