

1.

Film table

The screenshot shows a database query editor interface. At the top is a toolbar with icons for file operations, filters, and execution. Below the toolbar are three tabs: 'Query', 'Query History', and 'Scratch Pad'. The 'Query' tab is active and contains a SQL query. Below the query editor are three tabs: 'Data Output', 'Messages', and 'Notifications'. The 'Data Output' tab is active and shows a table with columns from the film table. The table has 8 columns: title, release_year, language_id, rental_duration, rental_rate, length, replacement_cost, and rating. Each column has a data type and a lock icon. The 'title' column is highlighted in blue.

```
1 SELECT title,  
2 release_year,  
3 language_id,  
4 rental_duration,  
5 rental_rate,  
6 length,  
7 replacement_cost,  
8 rating,  
9 count(*)  
10 FROM film  
11 GROUP BY title.
```

title	release_year	language_id	rental_duration	rental_rate	length	replacement_cost	rating
character varying	integer	smallint	smallint	numeric (4,2)	smallint	numeric (5,2)	mpaa_rating

Customer table

The screenshot shows a SQL IDE interface. The top pane displays a SQL query with line numbers 3 through 13. The query is as follows:

```
3 last_name,  
4 email,  
5 count(*)  
6 FROM customer  
7 GROUP BY store_id,  
8 first_name,  
9 last_name,  
10 email  
11 HAVING COUNT(*) > 1  
12  
13
```

The bottom pane is titled "Data Output" and shows the results of the query. The results are organized into columns with their data types and constraints:

store_id	first_name	last_name	email	count
smallint	character varying	character varying	character varying	bigint

Below the results, a status bar indicates "Total rows: 0 of 0" and "Query complete 00:00:00.066". The cursor is positioned at "Ln 10, Col 6".

No duplicate data was found in either query. If there had been, we could have fixed this by either:

1. Creating a view with unique records
2. Deleting the duplicate records

As an analyst, it's possible we will not have the permissions to do either of these tasks, so we would likely run a query using the GROUP BY and DISTINCT functions.

Checking for non-uniform data:

I chose to use the SELECT DISTINCT function to do a check through the data. For these tables, I checked each column individually to view the range of entries in each column. (Samples below)

2.

QueryQuery History

1SELECT store_id,
2first_name,
3last_name,
4email
5FROM customer
6GROUP BY store_id,
7first_name,
8last_name,
9email
10
11

Scratch Pad x

Data OutputMessagesNotifications

store_id
smallint

first_name
character varying

last_name
character varying

email
character varying

1	2	Bruce	Schwarz	bruce.schwarz@sakilacustomer.org
2	1	Robin	Hayes	robin.hayes@sakilacustomer.org
3	2	Vickie	Brewer	vickie.brewer@sakilacustomer.org
4	2	Raymond	Mcwhorter	raymond.mcwhorter@sakilacustomer.org
5	1	Natalie	Meyer	natalie.meyer@sakilacustomer.org
6	2	Alberto	Hennings	alberto.hennings@sakilacustomer.org

QueryQuery History

1SELECT title,
2release_year,
3language_id,
4rental_duration,
5rating
6FROM film
7GROUP BY title,
8release_year,
9language_id,
10rental_duration,
11rating

Scratch Pad x

Data OutputMessagesNotifications

title
character varying

release_year
integer

language_id
smallint

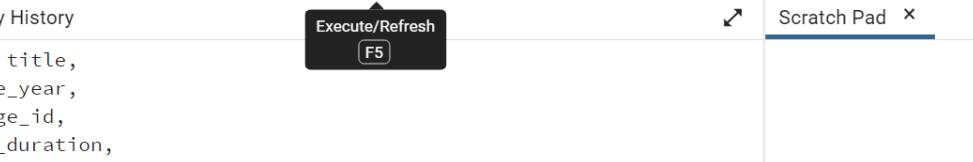
rental_duration
smallint

rating
mpaa_rating

1	Spirit Flintstones	2006	1	7	R
2	Virginian Pluto	2006	1	5	R
3	Rings Heartbreakers	2006	1	5	G
4	Amistad Midsummer	2006	1	6	G
5	Airport Pollock	2006	1	6	R
6	Pure Runner	2006	1	3	NC-17

Total rows: 1000 of 1000Query complete 00:00:00.160Ln 11, Col 7

3.missing data



The screenshot displays the DBeaver SQL editor interface. The top toolbar includes icons for file operations, query execution, and settings. The main editor area shows a SQL query:

```
1 SELECT title,  
2 release_year,  
3 language_id,  
4 rental_duration,  
5 rating  
6 FROM film  
7 WHERE title IS null  
8  
9
```

A tooltip "Execute/Refresh (F5)" is visible over the execution button. The right sidebar shows a "Scratch Pad". The bottom toolbar includes icons for data output, messages, and notifications. The bottom status bar displays column information:

title	release_year	language_id	rental_duration	rating
character varying	integer	smallint	smallint	mpaa_rating

Query

Query History

Scratch Pad

x

```

1 SELECT customer_id,
2 store_id,
3 first_name,
4 last_name,
5 email
6 FROM customer
7 WHERE store_id IS null
8
9

```

Data Output

Messages

Notifications

≡+

📄

▼

📋

▼

🗑️

🗄️

⬇️

📈

customer_id	store_id	first_name	last_name	email
[PK] integer	smallint	character varying	character varying	character varying

Total rows: 0 of 0

Query complete 00:00:00.123

Ln 7, Col 15

Summarize your data: Use SQL to calculate descriptive statistics for both the film table and the customer table. For numerical columns, this means finding the minimum, maximum, and average values. For non-numerical columns, calculate the mode value. Copy-paste your SQL queries and their outputs into your answers document.