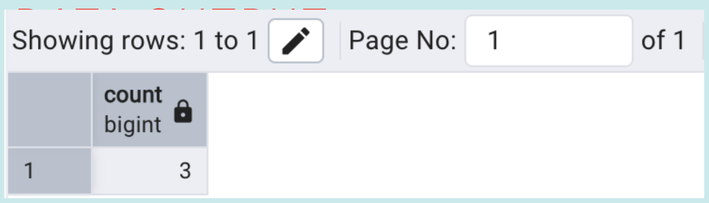
DATABASES & SQL FOR ANALYSTS

Rockbuster Stealth

STEP 2

Launch pgAdmin 4, open the Query Tool, copy-paste the following SQL statement into the Query Editor, and execute it. This statement will count all the instances of an actor with the first name “Ed” in the “actor” table. Copy the result from the Data Output window into your text document. Does your answer match the result from your earlier Excel count?

```
SELECT COUNT(*)
  FROM ACTOR
  WHERE FIRST_NAME = 'ED'
```



Showing rows: 1 to 1 Page No: 1 of 1

	count bigint
1	3

Did you find it easier to use Excel or the SQL statement and database to count the number of “Eds”? Explain your answer in your text document.

I guess for this exact task both options work. Excited to learn more about SQL soon.

STEP 3

To answer the next set of questions, you’ll paste the queries provided into the Query Editor in pgAdmin 4. Note down your answers in your running text document.

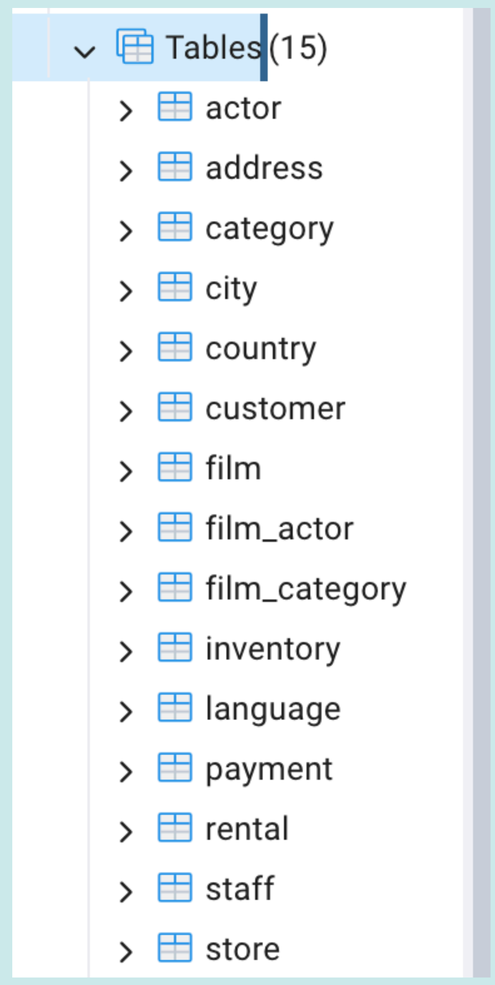
Execute the following query and list the names of the columns in the payment table.

```
SELECT * FROM PAYMENT LIMIT 10;
```

payment_id
customer_id
staff_id
rental_id
amount
payment_date

```
SELECT * FROM INFORMATION_SCHEMA.TABLES
 WHERE TABLE_SCHEMA = 'PUBLIC'
 AND TABLE_TYPE = 'BASE TABLE'
```

"actor"
"store"
"address"
"category"
"city"
"country"
"customer"
"film_actor"
"film_category"
"inventory"
"language"
"rental"
"staff"
"payment"
“film"



Within the pgAdmin 4 console, can you think of another way to list all the table names in the database instead of the SQL statement above?

• In pgAdmin 4 - Rockbuster dropdown menu -> Schemas dropdown menu -> Tables dropdown menu – Here it lists all the names of the names. You could also click Tables and navigate to Properties or Statistics at the top to display the names.

Analyze the rental duration distribution. How many days are most films rented for?

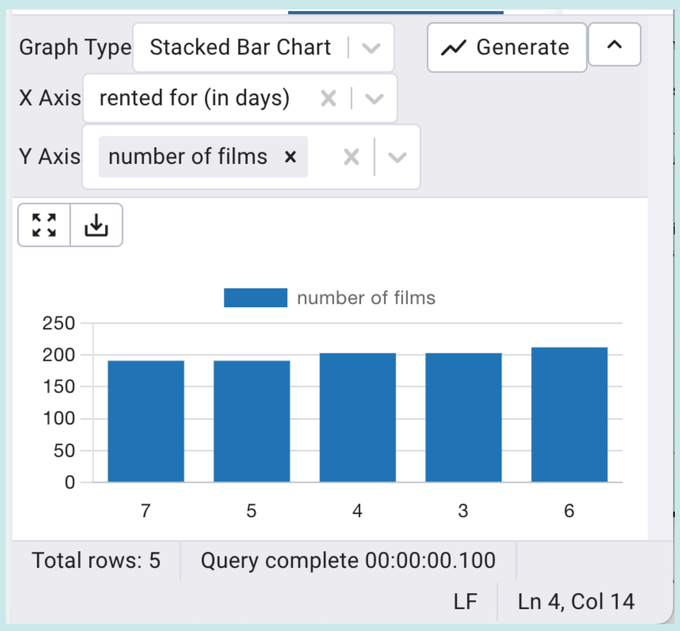
```
SELECT RENTAL_DURATION AS "RENTED FOR (IN DAYS)", COUNT(*) AS "NUMBER OF FILMS"
  FROM FILM
  GROUP BY 1
  ORDER BY 2
```

DATA OUTPUT

Showing rows: 1 to 5 Page No: 1 of 1

	rented for (in days) smallint	number of films bigint
1	7	191
2	5	191
3	4	203
4	3	203
5	6	212

Using the graph visualizer, we can identify that most films are rented for 6 days. (212 films)



STEP 4

Consider who in Rockbuster Stealth might want to use an OLAP or OLTP system for their data needs; for example, the sales department, which is interested in sales trends, would likely use an OLAP system. Describe at least two situations for each type of system.

OLAP vs. OLTP at Rockbuster Stealth

OLAP (Online Analytical Processing)

Purpose: Supports complex analysis and decision-making, typically used for trend analysis, reporting, and business intelligence.
Example 1: Sales Department
Situation: Analyzing monthly or yearly sales trends to understand which movies contribute most to revenue across different time periods and regions.
Usage: OLAP systems allow multi-dimensional analysis to compare sales by genre, time, or geography.
Example 2: Management Team

- Situation: Evaluating customer lifetime value and identifying profitable customer segments.
- Usage: OLAP systems help managers generate reports and dashboards for strategic decision-making using aggregated historical data.

OLTP (Online Transaction Processing)

Purpose: Supports day-to-day operations with real-time, fast, and reliable transaction processing.

Example 1: Customer Service Department

- Situation: Processing customer video rentals, returns, and payments in real-time.
- Usage: OLTP systems ensure that each transaction is immediately recorded, maintaining accurate inventory and payment records.

Example 2: Inventory Management Team

- Situation: Updating available movie stock whenever a customer rents or returns a video.
- Usage: OLTP systems support real-time inventory updates to avoid overbooking or errors in availability.

Step 5

Rockbuster Stealth has received an invoice for its new video collection licenses.

Take a moment to familiarize yourself with the data in the invoice, then note down the answers to the questions below.

- Does the invoice contain structured or unstructured data? Write an explanation for your answer.

The invoice contains structured data.

Explanation:

Structured data is highly organized and easily searchable within relational databases.

The invoice includes clearly defined fields such as Invoice Number, Customer Name, Address, Item Number, Quantity, Description, Price, and Payment Information.

These can all be categorized into database tables with specific columns.

- Organize and store the information on the invoice in a database. Step one will be to create a table in the text document you've started (you can insert a table if you're using MS Word or Google Docs, for example). Make sure your table contains columns with the appropriate labels, as well as the values from the invoice in each column. You're focusing, here, on a high-level structuring of your data.

Each piece of data can be assigned to a specific column, such as description, customer name, or account number. Each row represents a complete invoice entry or record, similar to how data is organized in a database table.

Here's a high-level table structure you can use:

Invoice_Number

Customer_Name

Customer_Address

Item_Code

Quantity

Description

Unit_Price

Subtotal

Company_Name

Company_Address

Account_Name

Account_Number

Invoice_Number	Customer_Name	Customer_Address	Item_Code
2019001	Mr. Timothy Walker	40 Sheila Sparks, NV	1

Quantity	*Description*	*Unit_Price*	*Subtotal*
1	New Video Collection Licensing	730$	730$

Company_Name	Company_Address	Account_Name	Account_Number
Oaklanders	4826 Norma Avenue, Anderson, TX	Miko Santo 4929 3310 0057 5422	4929 3310 0057 5422