

Databases & SQL for Analysts

3.1: Intro to Relational Databases

STEP 2

count all the actors whose first name is “Ed.” : 3

The screenshot shows a Microsoft Excel spreadsheet titled "actor (1) - Excel". The ribbon menu is visible at the top, with "Daten" (Data) selected. The main area displays a table of data from the "actor" table, with columns labeled A through F. The data shows three rows where the first name is "Ed":

	A	B	C	D	E	F
1	actor_id	first_name	last_name	last_update		
4	3	Ed	Chase	47:57.6		
38	136	Ed	Mansfield	47:57.6		
80	179	Ed	Guiness	47:57.6		
02						
02						

The screenshot shows the pgAdmin 4 interface. The top navigation bar includes links for Dashboard, Properties, SQL, Statistics, Dependencies, Dependents, Processes, and the current connection, ROCKBUSTER/postgres@PostgreSQL 18*. Below the navigation is a toolbar with various icons for database management. The main area has two tabs: 'Query' (selected) and 'Scratch Pad'. The 'Query' tab contains the following SQL code:

```
1 SELECT COUNT(*)  
2 FROM actor  
3 WHERE first_name = 'Ed'  
4
```

The 'Data Output' tab shows the results of the query:

	count	bigint
1	3	

Below the table, the status bar indicates "Showing rows: 1 to 1" and "Page No: 1 of 1".

Does your answer match the result from your earlier Excel count?

Yes, it matches with our results in excel: **there is 3 rows**

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.

In this case, it was easy for both Excel and SQL, for me, but it will be easier to find the counts with SQL than with Excel for a similar case if there is a larger number of data

STEP 3

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

The screenshot shows the pgAdmin 4 interface with the 'SQL' tab selected. Below the tabs, there is a table structure for the 'payment' table. The columns are listed as follows:

	payment_id [PK] integer	customer_id smallint	staff_id smallint	rental_id integer	amount numeric (5,2)	payment_date timestamp without time zone
--	----------------------------	-------------------------	----------------------	----------------------	-------------------------	---

Under the “table_name” column, what are the names of the tables that are available in the Rockbuster database? (List all names.)

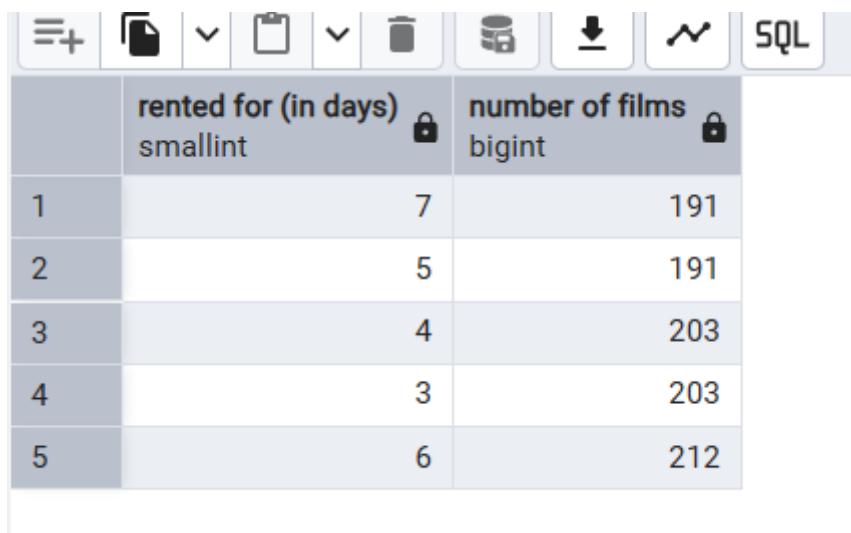
The screenshot shows the pgAdmin 4 interface with the 'table_name' column highlighted. The table lists various tables in the database:

table_name
name
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?

- > 1..3 Sequences (13)
- >  Tables (15)
 - >  actor
 - >  address
 - >  category
 - >  city
 - >  country
 - >  customer
 - >  film
 - >  film_actor
 - >  film_category
 - >  inventory
 - >  language
 - >  payment
 - >  rental
 - >  staff
 - >  store
- >  Trigger Functions
- >  Types

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



The screenshot shows a database interface with various icons at the top. Below is a table with two columns: 'rented for (in days)' (smallint) and 'number of films' (bigint). The data is as follows:

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

The majority of films are rented for 6 days, with 212 films recorded. Rentals of 3 and 4 days follow closely behind, with 203 films each.

STEP 4

Describe at least two situations for each type of system.

OLTP (Customer Service): Record a movie return instantly to update the customer's account.

OLAP (Sales Department): Analyze monthly rentals by film category to identify consumption trends.

As a data analyst we will use the OLAP System to our analysis

STEP 5

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

- **unstructured data.** The information is presented in free-form layout, not in a predefined schema like rows and 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.

Item Code	Item Description	Quantity	Unit Price	Subtotal	Total Amount Due	Payment Account Name	Payment Account Number
001	New Video Collection Licensing	1	\$730	\$730	\$730	Miko Santo	4929 3310 0057 5422
Invoice Number	Recipient Name	Recipient Address	Sender Name	Sender Address			
2019001	Mr. Timothy Walker	40 Sheila La, Sparks, NV	Oaklanders Sound Studio	4826 Norma Avenue, Anderson, TX			

Bonus Task

