

DIRECTIONS:

STEP 1: If you haven't done so already, install PostgreSQL and load the Rockbuster database using the instructions in the Exercise. Then download your Achievement 3 project brief (PDF) to get an idea of what each Exercise will cover.

STEP 2: Compare and contrast spreadsheets and databases by following the steps below:

- Download the Rockbuster "actor.csv" file and open it in Excel.
- Drawing on what you've learned in previous Achievements, use the appropriate functions in Excel to count all the actors whose first name is "Ed." Write down the result in a text document.

The screenshot shows an Excel spreadsheet with columns A through H. Column A is 'actor_id', B is 'first_name', C is 'last_name', D is 'last_update', E is empty, F is 'ANSWER', G is 'Total "ED" count', and H is empty. Rows 1 and 2 are visible. Row 1 contains headers. Row 2 contains data for actor_id 1, first_name Penelope, last_name Guinness, last_update 47:57.6. The formula bar shows '=COUNTIF(B2:B201,B4)'.

	A	B	C	D	E	F	G	H
1	actor_id	first_name	last_name	last_update		ANSWER		
2	1	Penelope	Guinness	47:57.6		STEP 2	Total "ED" count	3
3	2	Nick	Wahlberg	47:57.6				

- Launch pgAdmin 4, open the Query Tool, copy-paste the SQL statement below into the Query Editor, and execute it.

```
SELECT COUNT(*)
FROM actor
WHERE first_name = 'Ed'
```

The screenshot shows the pgAdmin 4 Query Editor interface. The query editor contains the SQL statement: SELECT COUNT(*) FROM actor WHERE first_name = 'Ed'. The query has been executed, and the result is shown in the Data Output tab. The result is a single row with the value 3.

	count
1	3

ANSWER:

The Excel dataset counts three (3) "Ed" names which is the same result as SQL. However, to arrive at this result, I first filtered the "Ed" on the 2nd column and then highlighted the matching name to count the total but immediately realized that it will not calculate the count. So, I recount the use of the formula "COUNTIF" to give an exact measure. On the other note, SQL was straightforward and did all of the steps I took within seconds.

STEP 3: To answer the next set of questions, you'll be pasting 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;
```

COLUMN NAMES:

- Payment ID
- Customer ID
- Staff ID
- Rental ID
- Amount
- Payment Date

The screenshot shows the pgAdmin 4 interface with the query editor. The query is: `SELECT * FROM payment LIMIT 10;`. The results are displayed in a table with 10 rows and 7 columns.

	payment_id [PK] integer	customer_id smallint	staff_id smallint	rental_id integer	amount numeric (5,2)	payment_date timestamp without time zone
1	17503	341	2	1520	7.99	2007-02-15 22:25:46.996577
2	17504	341	1	1778	1.99	2007-02-16 17:23:14.996577
3	17505	341	1	1849	7.99	2007-02-16 22:41:45.996577
4	17506	341	2	2829	2.99	2007-02-19 19:39:56.996577
5	17507	341	2	3130	7.99	2007-02-20 17:31:48.996577
6	17508	341	1	3382	5.99	2007-02-21 12:33:49.996577
7	17509	342	2	2190	5.99	2007-02-17 23:58:17.996577
8	17510	342	1	2914	5.99	2007-02-20 02:11:44.996577
9	17511	342	1	3081	2.99	2007-02-20 13:57:39.996577
10	17512	343	2	1547	4.99	2007-02-16 00:10:50.996577

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

```
SELECT * FROM information_schema.tables
WHERE table_schema = 'public'
AND table_type = 'BASE TABLE'
```

TABLE NAMES:

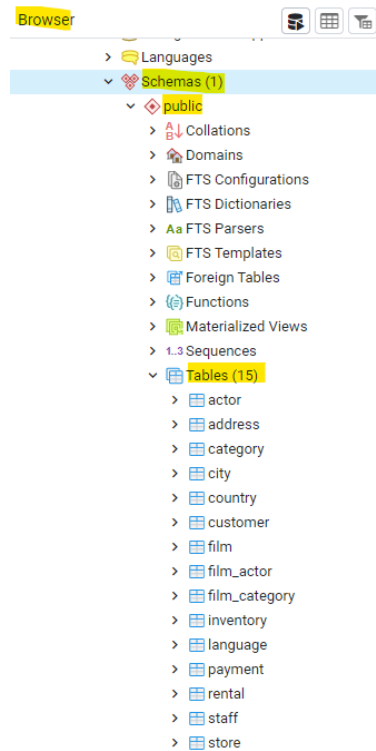
- staff
- rental
- category
- city
- film_category
- store
- country
- film
- actor
- address
- language
- film_actor
- inventory
- customer
- payment

The screenshot shows the pgAdmin 4 interface with the query editor. The query is: `SELECT * FROM information_schema.tables WHERE table_schema = 'public' AND table_type = 'BASE TABLE'`. The results are displayed in a table with 11 rows and 8 columns.

	table_catalog	table_schema	table_name	table_type	self_referencing_column_name	reference_generation	user_defined
1	Rockbuster	public	staff	BASE TABLE	[null]	[null]	[null]
2	Rockbuster	public	category	BASE TABLE	[null]	[null]	[null]
3	Rockbuster	public	film_category	BASE TABLE	[null]	[null]	[null]
4	Rockbuster	public	country	BASE TABLE	[null]	[null]	[null]
5	Rockbuster	public	actor	BASE TABLE	[null]	[null]	[null]
6	Rockbuster	public	language	BASE TABLE	[null]	[null]	[null]
7	Rockbuster	public	inventory	BASE TABLE	[null]	[null]	[null]
8	Rockbuster	public	payment	BASE TABLE	[null]	[null]	[null]
9	Rockbuster	public	rental	BASE TABLE	[null]	[null]	[null]
10	Rockbuster	public	city	BASE TABLE	[null]	[null]	[null]
11	Rockbuster	public	store	BASE TABLE	[null]	[null]	[null]

- 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?

Referring to the "Browser" under "*Schemas*" then select "*public*" and click on "*Tables (15)*" to view all table 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
```

Rockbuster/postgres@PostgreSQL 10 (x86) ▾

Query Editor Query History

```
1 SELECT rental_duration AS "rented for (in days)", COUNT(*) AS "number of films"
2 FROM film
3 GROUP BY 1
4 ORDER BY 2
```

Data Output Explain Messages Notifications

	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 film rentals are between 4 to 6 days.

STEP 4: Think about 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 2 situations for each type of system.

OLAP – departments involved with the decision-making on business operations like marketing, sales, upper management, and most execs will likely need access to foresee records of the historical trends and possible company-wide issues.

OLTP – employees who are mostly dealing with customer-based daily transaction activities like storefront and warehouse. They will need access to maintain customer order updates (create/update accounts, order cancellations, etc.).

STEP 5: Rockbuster Stealth has received an invoice for the licenses for its new video collection.

Take a moment to familiarize yourself with 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 is structured data but may be in a different form. The details on the form show transactional records for a client tracking variables like invoice number, address, item, item quantity, item description, price, total price, account name, and account number.

➤ *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.*

MERCHANT

MERCHANT	ADDRESS	CITY	STATE	ACCOUNT NAME	ACCOUNT NO.
Oaklanders Sound Studio	4826 Norma Avenue	Anderson	TX	Mike Santo	4929331000575422

TRANSACTION

INVOICE NO.	ITEM CODE	ITEM QTY	DESCRIPTION	CURRENCY	PRICE
2019001	001	01	New Video Collection Licensing	USD	730

CUSTOMER

TITLE	FIRST NAME	LAST NAME	GENDER	ADDRESS	CITY	STATE
MR.	TIMOTHY	WALKER	M	40 Sheila Lane	Sparks	NV