Juliana Lee

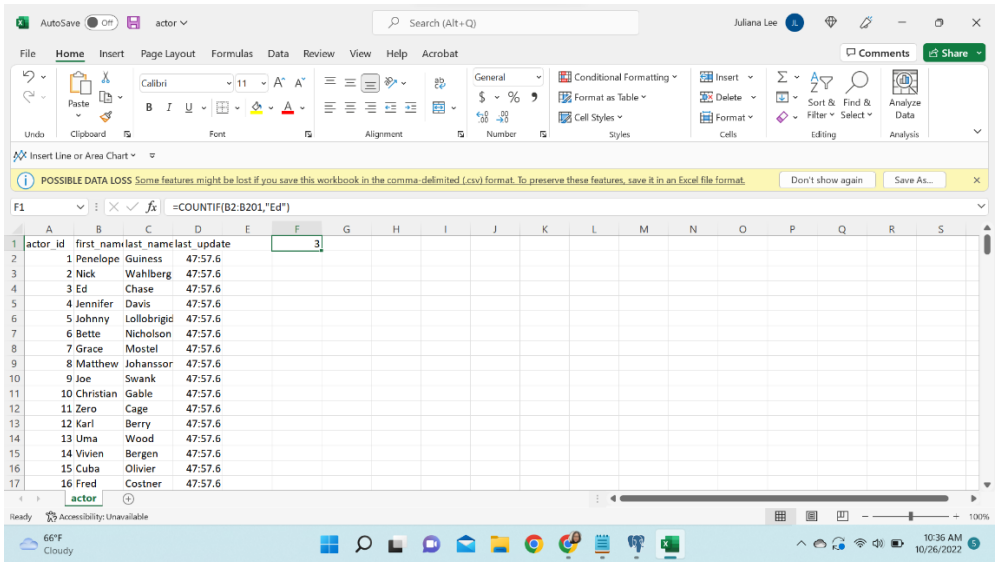
Task 3.1: Intro to Relational Databases

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:

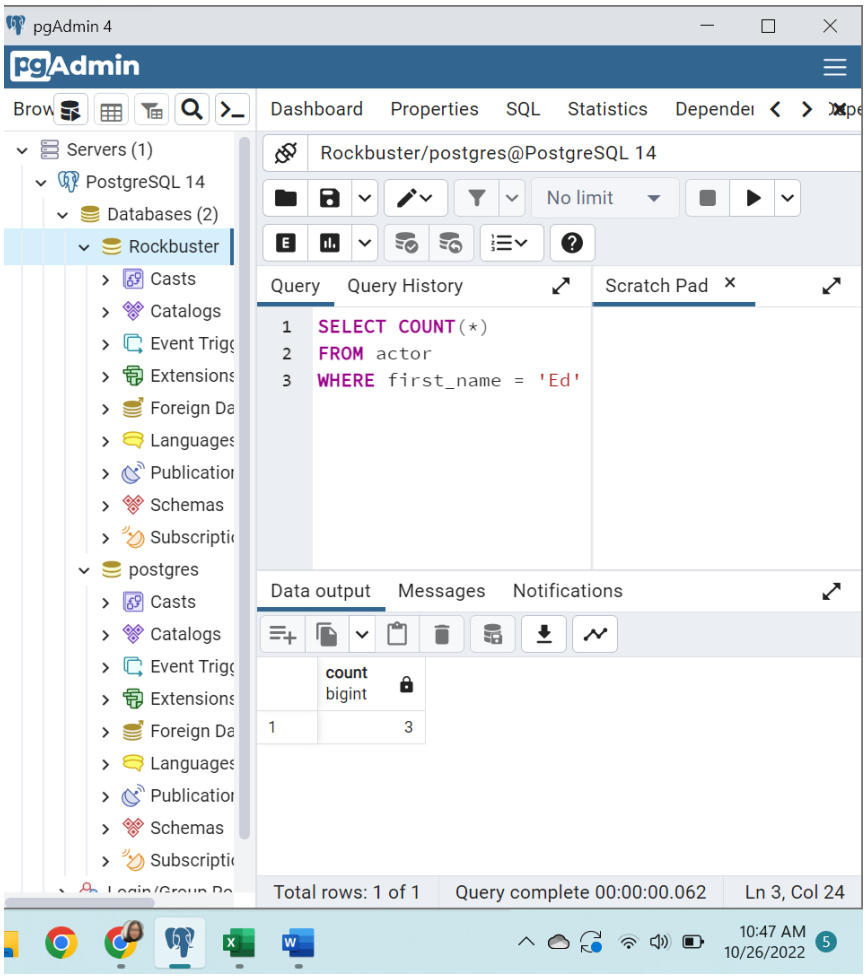
- Download the Rockbuster "actor.csv" file and open it in Excel.
- Drawing on what you've learned in previous Achievements, use the appropriate functions in Excel to count all the actors whose first name is "Ed." Write down the result in a text document.
  - 

- 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'
```



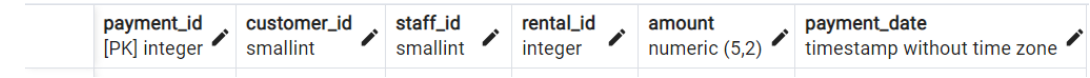
- Copy the result that tells you the number of times the first name "Ed" appears in the "actor" table from the **Data Output** window into your text document from step 2b. Check that your answer matches your answer from step 2a. Was it easier to use Excel or the SQL statement and database to count the number of "Eds"? Provide an explanation for your answer in the same text document.
    - **Both Excel and SQL yielded the same result, 3 counts of "Ed". It was easy to use both; however, using the SQL statement is ultimately easier. The reason why it was easy to use Excel is because the data set is not too big. However, with a bigger data set the better choice would be to use SQL.**

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;
```

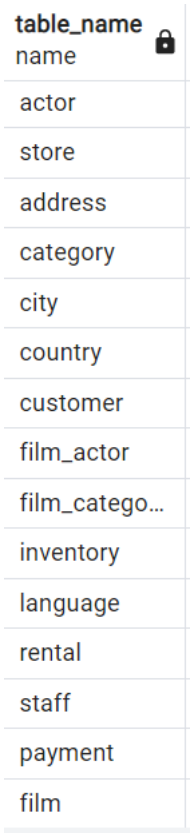


payment_id [PK] integer	customer_id smallint	staff_id smallint	rental_id integer	amount numeric (5,2)	payment_date timestamp without time zone

1. **Payment_id**
2. **Customer_id**
3. **Staff_id**
4. **Rental_id**
5. **Amount**
6. **Payment_date**

- 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'
```

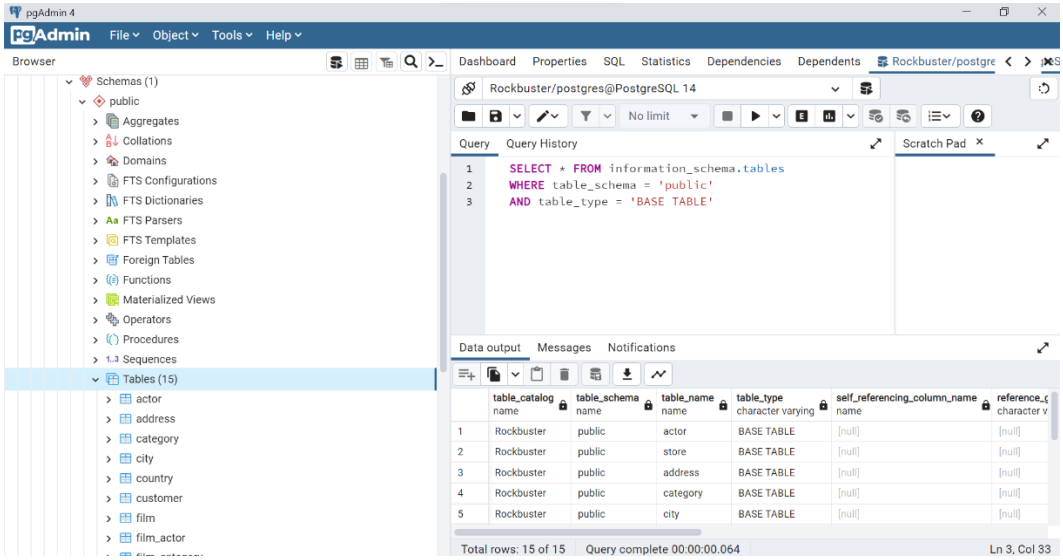


table_name name
actor
store
address
category
city
country
customer
film_actor
film_catego...
inventory
language
rental
staff
payment
film

1. **Actor**
2. **Store**
3. **Address**
4. **Category**
5. **City**
6. **Country**
7. **Customer**
8. **Film actor**
9. **Film category**
10. **Inventory**
11. **Language**
12. **Rental**
13. **Staff**
14. **Payment**
15. **Film**

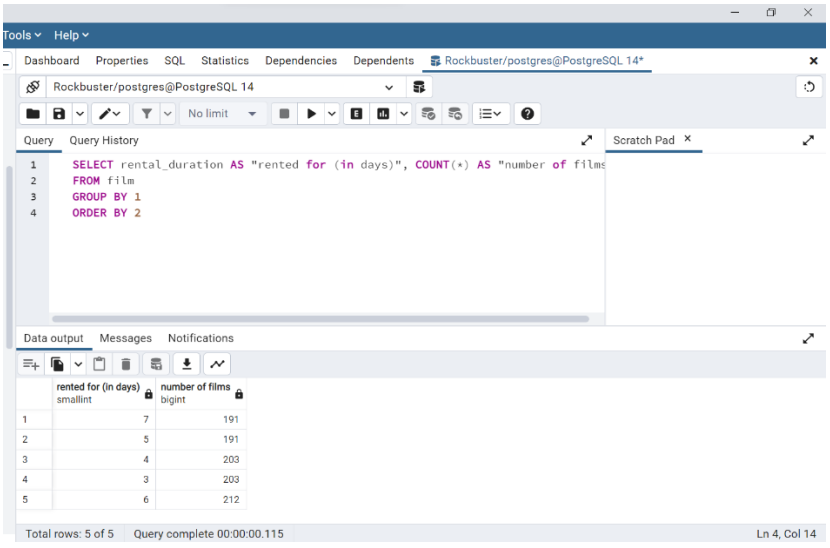
Juliana Lee

- 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?
    - **Yes, first look at the *Browser Section*, then look at the *Schema*, lastly press the *Table* and all 15 table names will be shown.**



- 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
```



Most films are rented for 6 days (212 films).

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

1. **Customer Service department: interested in reviews**
2. **Marketing department: interested in current market trends**

OLTP

1. **Finance department: record and update rental purchases**
2. **Inventory department: record movies rented out and returned and update records in system**

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 contains structured data. The invoice is predefined data that can be categorized into different variables such as invoice number, item number, quantity, description, etc.**
- 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.

Juliana Lee

Merchant Information					
Merchant Name	**Account Holder Name**	**Account number**	**Address**	**City**	**State**
Oaklanders Sound Studio	Mike Santo	4929331000575420	4826 Norma Ave.	Anderson	TX

Customer Information						
Title	**First Name**	**Last Name**	**Gender**	**Address**	**City**	**State**
Mr.	Timothy	Walker	Male	40 Sheila LA	Sparks	NV

Transaction					
Invoice Number	**Item**	**Quantity**	**Description**	**Currency**	**Price**
2019001	001	01	New Video Collection Licensing	USD ($)	730