3.1: Intro to Relational Databases

1. **Step 1**
   Install PostgresSQL and load Rocbuster database.
2. **Step 2**
   Compare and contrast spreadsheets and databases by following the steps below:
   - Download the Rockbuster "actor.csv" file and open it in Excel.
   - Use the appropriate functions in Excel to count all the actors whose first name is "Ed." Write down the result.

     =COUNTIF(Table1[#All];"Ed") or locate the specific column to avoid the last_name being "Ed" also calculated **column =COUNTIF(B:B;"Ed")**

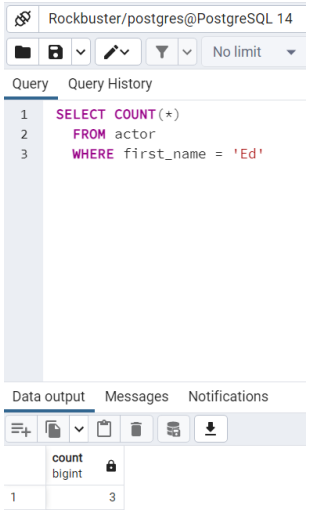
	A	B	C	D	E	F	G
1	actor_id	first_name	last_name	last_update		How many "Ed"	3
2	1	Penelope	Guiness	2013-05-26 14:47:57.62			
3	2	Nick	Wahlberg	2013-05-26 14:47:57.62			
4	3	Ed	Chase	2013-05-26 14:47:57.62			
5	4	Jennifer	Davis	2013-05-26 14:47:57.62			
6	5	Johnny	Lollobrigida	2013-05-26 14:47:57.62			

   - Launch pgAdmin 4, open the Query Tool, copy-paste the SQL statement 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.

     Easier using SQL because we can just hit the name of the table without having to specify further the numerical column location (A, B, C, etc.) which or this we should know previously how the table looks like. Using SQL statement, we just have to know the variable name.



3. **Step 3**
   - Execute the following query and list the names of the columns in the payment table.

```
SELECT * FROM payment LIMIT 10;
```

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

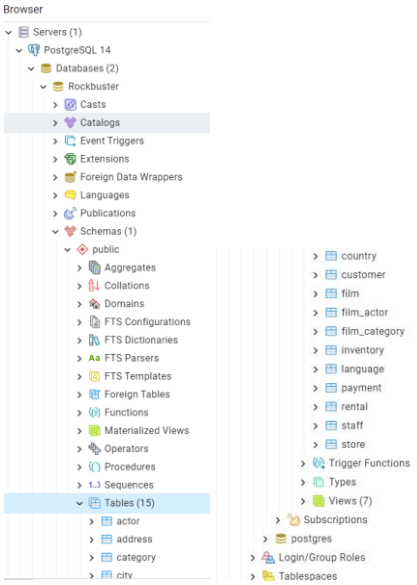
```
SELECT * FROM information_schema.tables
WHERE table_schema = 'public'
AND table_type = 'BASE TABLE'
```

Resulting in 15 Tables

	table_catalog name	table_schema name	table_name name	table_type character varying
1	Rockbuster	public	actor	BASE TABLE
2	Rockbuster	public	store	BASE TABLE
3	Rockbuster	public	address	BASE TABLE
4	Rockbuster	public	category	BASE TABLE
5	Rockbuster	public	city	BASE TABLE
6	Rockbuster	public	country	BASE TABLE
7	Rockbuster	public	customer	BASE TABLE
8	Rockbuster	public	film_actor	BASE TABLE
9	Rockbuster	public	film_catego...	BASE TABLE
10	Rockbuster	public	inventory	BASE TABLE
11	Rockbuster	public	language	BASE TABLE
12	Rockbuster	public	rental	BASE TABLE
13	Rockbuster	public	staff	BASE TABLE
14	Rockbuster	public	payment	BASE TABLE
15	Rockbuster	public	film	BASE TABLE

Total rows: 15 of 15 Query complete 00:00:00.107

   - 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. Just access the available table on the GUI.



- 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
```

6 days

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

4. **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 – Business & marketing team: fetching data on orders and new customers resulting from digital engagement.
- OLTP – Finance team: manipulating data on the order transactions and employees expenses from the business trip/event.

5. **Step 5**

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

- Does the invoice contain structured or unstructured data? Write an explanation for your answer.
  Structured, because we can categorize each data into different specific variable.
- Organize and store the information on the invoice in a database.

Transaction Table					
Invoice Number	Item	Quantity	Description	Price	Currency
2019001	001	01	New Video Collection Licensing	730	$

Merchant Table						
Merchant Name	Account Name	Account Name	Address	City	State	State Abbreviation
Oaklanders Sound Studio	Miko Santo	4929331000575420	4826 Norma Avenue	Anderson	Texas	TX

Customer Table						
First Name	Last Name	Gender	Address	City	State	State Abbreviation
Timothy	Walkers	Male	40 Sheila La	Sparks	Nevada	NV