Exercise 3.1 - Intro to Relational Databases

Step 1

Install PostgresSQL and load Rocbuster database.

Step 2

Compare the spreadsheets and databases by following the steps below:
o   Download the Rockbuster "actor.csv" file and open it in Excel.
o   Use the appropriate functions in Excel to count all the actors whose first name is "Ed." Write down the result.

	A	B	C	D	E	F	G
1	actor_ ▼	first_nam ▼	last_name ▼	last_update ▼		Actors whose first name is "Ed."	= 3
2	1	Penelope	Guiness	26.05.2013 14:47			
3	2	Nick	Wahlberg	26.05.2013 14:47			
4	3	Ed	Chase	26.05.2013 14:47			
5	4	Jennifer	Davis	26.05.2013 14:47			
6	5	Johnny	Lollobrigida	26.05.2013 14:47			
7	6	Bette	Nicholson	26.05.2013 14:47			
8	7	Grace	Mostel	26.05.2013 14:47			
9	8	Matthew	Johansson	26.05.2013 14:47			
10	9	Joe	Swank	26.05.2013 14:47			
11	10	Christian	Gable	26.05.2013 14:47			
12	11	Zero	Cage	26.05.2013 14:47			
13	12	Karl	Berry	26.05.2013 14:47			
14	13	Uma	Wood	26.05.2013 14:47			

o   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'
```
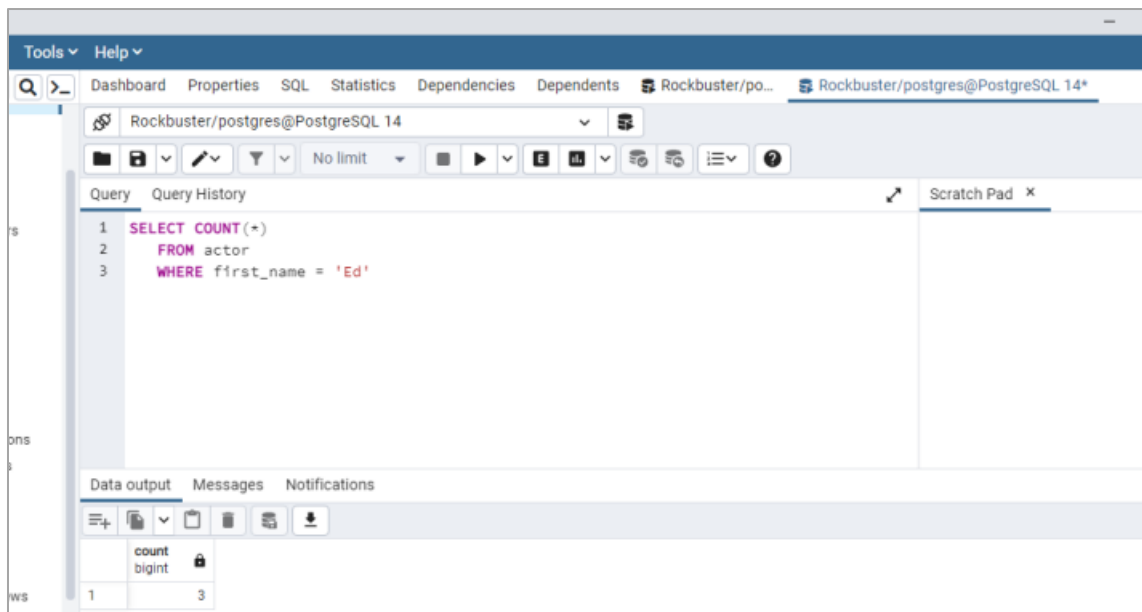
Was it easier to use Excel or the SQL statement and database to count the number of "Eds"?

→ At the moment it's easier for me to use a pivot table in excel to find the count of a word (for example Ed) because I am familiar with how to use pivot tables. I can see how the SQL statement can be easy to use once I understand how to formal the statement. It's also easier to read the results in Excel because it clearly states what value you are looking at (the names of people) whereas in SQL it just gives a value for your output.

Step 3

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

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

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


➢ 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'
```

	table_catalog name	table_schema name	table_name name	table_type character varying	self_referencing_column_name name	reference_generation character varying
1	Rockbuster	public	actor	BASE TABLE	[null]	[null]
2	Rockbuster	public	store	BASE TABLE	[null]	[null]
3	Rockbuster	public	address	BASE TABLE	[null]	[null]
4	Rockbuster	public	category	BASE TABLE	[null]	[null]
5	Rockbuster	public	city	BASE TABLE	[null]	[null]
6	Rockbuster	public	country	BASE TABLE	[null]	[null]
7	Rockbuster	public	customer	BASE TABLE	[null]	[null]
8	Rockbuster	public	film_actor	BASE TABLE	[null]	[null]
9	Rockbuster	public	film_catego…	BASE TABLE	[null]	[null]
10	Rockbuster	public	inventory	BASE TABLE	[null]	[null]
11	Rockbuster	public	language	BASE TABLE	[null]	[null]
12	Rockbuster	public	rental	BASE TABLE	[null]	[null]

Total rows: 15 of 15   Query complete 00:00:00.071

➢ Payment
➢ Film
➢ Actor
➢ Address
➢ Category
➢ City
➢ Country
➢ Customer
➢ Film_actor
➢ Film_category
➢ Inventory
➢ Language
➢ Rental
➢ Staff o Store

o 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?

→ Under the Browser on the left side of the screen, when you select for Schemas, you can see that there are 15 Table options with the above categories listed. See screenshot below.
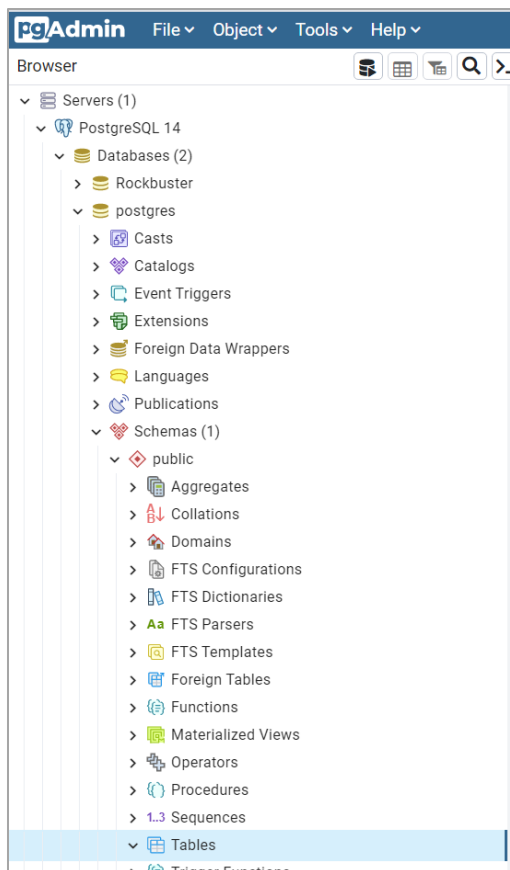
```
SELECT rental_duration AS "rented for (in days)", COUNT(*) AS "number of films"
  FROM film
  GROUP BY 1
  ORDER BY 2
```

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

→ Most films are rented between 3-7 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

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 – Budgeting and forecasting team, financial reporting team
→OLTP-- Retail sales team, and financial transaction team

Step 5

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

Does the invoice contain structured or unstructured data?
→ Yes, Structured. Because all the informations are clearly specified and could be categorised into different variables.

Organize and store the information on the invoice in a database.

Overall Invoice Information								
Invoice_id	Address	Company	Item #	Qty	Description	Price ($)	Account Name	Acct #
2019001	4826 Norma, Ave Anderson, TX	Oaklanders Sound Studio	001	01	New video collection licensing	730	Miko Santo	4929331000575422

Invoice Breakdown							
Customer							
Title	First Name	Last Lame	Gender	Address	City	State	State abbreviation
Mr	Timothy	Walker	Male	Sheila La	Sparks	Nevada	NV
Transaction							
Invoice No.	Item #	Qty	Description		Currency	Price	
2019001	001	01	New video collection licensing		US Dollars	730	
Merchant							
Name	Address	City	State	State Abbreviation	Account Name	Account No.	
Oaklanders	4826 Norma Ave	Anderson	Texas	TX	Mike Santo	4929331000575420	