3.1 Introduction 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

Done

Step 2

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

- Download the Rockbuster "actor.csv" file and open it in Excel.

  Done

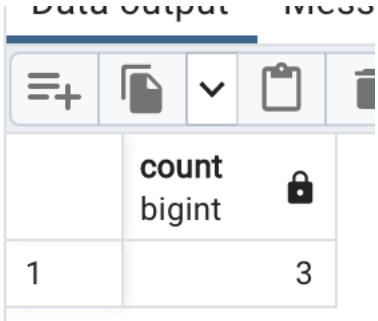
- Drawing on what you've learned in previous Achievements, use the appropriate functions in Excel to count all the actors whose first name is "Ed." Write down the result in a text document.

  3 cells

- 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.



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;
```

Data output | Messages | Notifications

	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'
```

Data output | Messages | Notifications

	table_catalog name	table_schema name	table_name name	table_type character varying	self_referencing_column_name name	reference character
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]
13	Rockbuster	public	staff	BASE TABLE	[null]	[null]
14	Rockbuster	public	payment	BASE TABLE	[null]	[null]
15	Rockbuster	public	film	BASE TABLE	[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?
- 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
```

	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

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.

  Structured, because we can categorize each data into different specific variable

- 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.

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	49293310005 75420	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