

### Task 3.1

#### Step 2


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

- 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.

actor_id	first_name	last_name	last_update
3	Ed	Chase	47:57.6
136	Ed	Mansfield	47:57.6
179	Ed	Guinness	47:57.6

Total count: 3

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

	count bigint	
1	3	

Result: 3; I would say SQL is easier, since it is less steps just one query language, for excel, I need filter the first name, and find Ed, and then either count manually myself or look at the spreadsheet to find the answer.

#### 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; customer\_id; staff\_id; rental\_id; amount; payment\_date;

	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'
```

Actor; store; address; category; city; country; customer; film\_actor; film\_category; inventory, language; rental; staff; payment; file

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

Total rows: 15 of 15

Query complete 00:00:00.062

Ln 4, Col 1

- 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?

Go to the schemas and open the tables, all the tables are listed there.

- 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

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 – Total Orders or Sales of a FI clothing store for number of years, manipulate data to look at YTD figures for each year

OLTP - Record of transactions across the ATM network or credit card payment at a point in time

Step 5: Does the invoice contain structured or unstructured data? Write an explanation for your answer.

Semi-structured data since it falls somewhere between structured and unstructured data in that it retains a certain level of organization while maintaining a greater degree of flexibility than a tabular format.

Invoice Number	Payment description	Price	Quantity	Payee	Payee's Account name	Payee's Account Number
2019001	New Video Collection Licensing	\$730	1	Oaklanders Sound Stutdio	Miko Santo	4929331000575422

Bonus: I did not use the excel or tableau, but the SQL's own visualization

