Jake Elliot

Data Immersion
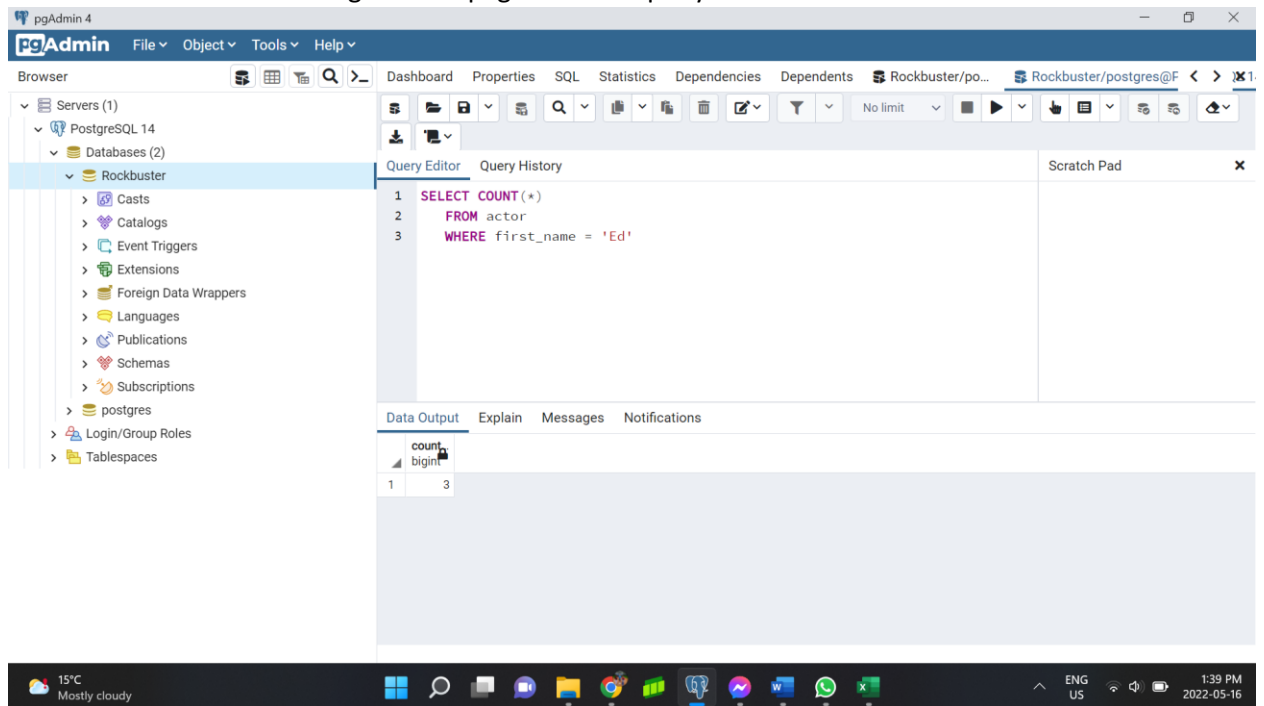
May 14, 2022

Exercise 3.1

Step 2

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.

    a.   There are 3 'Ed's in the Excel sheet.
    b.   Here is a screenshot of the PgAdmin 4 page with the query.

    c.   I found Excel easier, but that's because I'm more used to it.

Step 3

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

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

| 17503 | 341 | 2 | 1520 | 7.99 | "2007-02-15 22:25:46.996577" |
| 17504 | 341 | 1 | 1778 | 1.99 | "2007-02-16 17:23:14.996577" |

17505	341	1	1849	7.99	"2007-02-16 22:41:45.996577"
17506	341	2	2829	2.99	"2007-02-19 19:39:56.996577"
17507	341	2	3130	7.99	"2007-02-20 17:31:48.996577"
17508	341	1	3382	5.99	"2007-02-21 12:33:49.996577"
17509	342	2	2190	5.99	"2007-02-17 23:58:17.996577"
17510	342	1	2914	5.99	"2007-02-20 02:11:44.996577"
17511	342	1	3081	2.99	"2007-02-20 13:57:39.996577"
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"

"film"

- This list can also be found on the tab, under Schemas → Tables

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

Most films are rented for 6 days.

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.

    A. OLAP
        1. The Marketing department would fall under the OLAP category as they analyze the data to determine who to market to.
        2. Consumer Insights would like to be able to see and analyze what consumers are watching, where trends are leading.
    B. OLTP
        1. Content department would likely fall under both categories, but I'll place them here as they'll want to update the movie selection and the information.
        2. Accounting would need to update the transactions as they appear, thus in OLTP.

Step 5

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

    1. Structured vs. Unstructured Data
        a. I would consider this Structured Data as the information on the invoice would follow certain restrictions, such as for an address, the price, an invoice number, 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.

2. Database

Category	Information
Studio Name	Oaklanders Sound Studio
Invoice Number	2019001
Name	Mr. Timothy Walker
Address	40 Sheila La Sparks, NV
Item #	001
QTY	01
Description	New Video Collection Licensing
Price	$730
Sub Total	$730
Make Your Payment To	Oaklanders
Address	4826 Norma Avenue, Anderson, TX
Account Name	Miko Santo
Account #	4929 3310 0057 5422