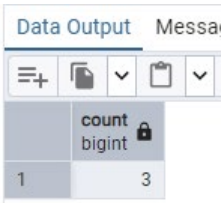
Data Immersion Ach 03.01

Ryan Wick – 01/08/2024

- **Directions**
- **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.
- **Step 2**
    - Compare and contrast spreadsheets and databases by following these steps:
    - Download the Rockbuster "actor.csv" file and open it in Excel. 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.
        - **ANSWER:** Ed = 3
    - Launch pgAdmin 4, open the Query Tool, copy-paste the following SQL statement into the Query Editor, and execute it. This statement will count all the instances of an actor with the first name "Ed" in the "actor" table. Copy the result from the Data Output window into your text document. Does your answer match the result from your earlier Excel count?
        - **ANSWER:**



    - SELECT COUNT(*)
    - FROM actor
    - WHERE first_name = 'Ed'
    - **Tip!**
        - Check out this video for a brief walkthrough of how to use the Query Tool in pgAdmin 4.

- Did you find it easier to use Excel or the SQL statement and database to count the number of “Eds”? Explain your answer in your text document.
    - **ANSWER:** Honestly it was easier for Excel both because it’s a simple task but also I have more familiarity with it due to our previous exercises.

- **Step 3**
    - To answer the next set of questions, you’ll paste 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;
    - Under the “table_name” column, what are the names of the tables that are available in the Rockbuster database? (List all names.)
        - **ANSWER:** 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'
    - 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?
        - **ANSWER:** You could navigate to the “Schemas” node and expand it. From there you’d just have to notate which ones were listed as “public”.
    - Analyze the rental duration distribution. How many days are most films rented for?
        - **ANSWER:** 6 days at 212 films with 4 & 3 days tying for second.
    - SELECT rental_duration AS "rented for (in days)", COUNT(*) AS "number of films"
    - FROM film
    - GROUP BY 1
    - ORDER BY 2

- **Step 4**
    - Consider 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 two situations for each type of system.
        - **ANSWER:**
            - **OLAP**
                - A situation that would best utilize an OLAP system would be something like a marketing department. Using this type of system, they can perform much larger and complex queries without much issue.
                - **Example:** The marketing team wants to know just how well their different marketing campaigns went across different geolocations. They could perform a complex query and analysis of the information to see which areas had higher or lower engagement along with possible consideration of the types of campaigns run in each area.
            - **OLTP**
                - A situation I believe would best be suited with an OLTP would be for any type of customer service department. Since they will be constantly assisting customers on the go, they need the information to be as up to date as possible. Also, since they are likely just doing simply transactions the OLTP system should be able to handle the volume much better.
                - **Example:** As customers call into the help desk line the representative could use the OLTP system to quickly pull the customer's information, rental history, etc. to save the customer time from having to provide all that background information.
- **Step 5**
    - Rockbuster Stealth has received an invoice for its new video collection licenses.
    - Figure 25.
    - Take a moment to familiarize yourself with the data in the invoice, then note down the answers to the questions below.
    - Does the invoice contain structured or unstructured data?
        - **ANSWER:** Structured

- 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.
    - **ANSWER:**

Invoice:	Name:	Address:	Item:	Qty:	Description:	Price:	Sub Total	Account Name:	Account N.O.:
2019001	Mr. Timothy Walker	40 Sheila LA Sparks, NV	1	1	New Video Collection Licensing	$730	$730	Miko Santo	4929 3310 0057 5422

- **Step 6**
    - Save the text document containing your answers as a PDF and upload it here for your tutor to review. Don't hesitate to contact your tutor or mentor if you have any questions!