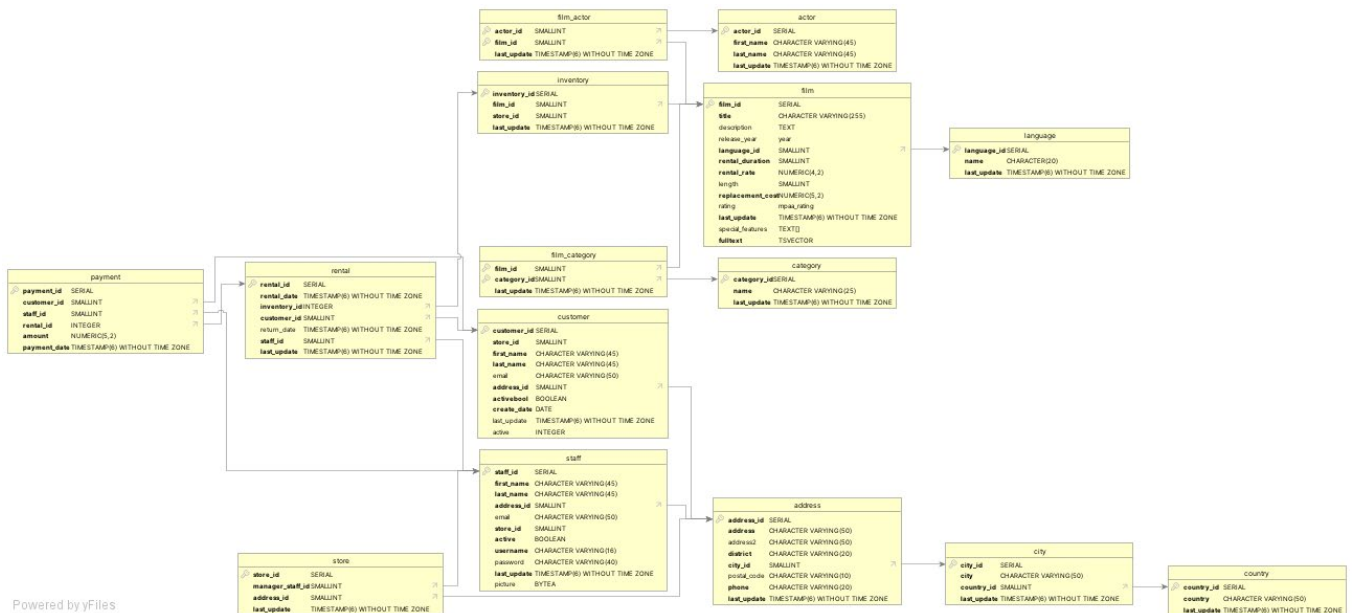


Data Immersion Ach 03.02

Ryan Wick – 01/09/2024

- **Directions**
- **Step 1. Create your Answers document:**
 - Create a new text document and call it “Answers 3.2.” You’ll save a copy of your ERD, data dictionary, and written answers in this document.
- **Step 2. Extract the ERD:**
 - Download and install [DbVisualizer](#) or [Lucidchart](#) (if you haven’t already done so).
 - Extract the ERD from the Rockbuster database and save it as an image (PNG or JPEG) using the instructions in the Exercise.
 - Copy-paste the ERD into your answers document.



- **Tip!**

If the ERD image you save is blurry, try zooming in on the ERD in DbVisualizer or Lucidchart, then download it again.
- **What to do if DbVisualizer doesn't work?**

In some cases there is a bug in DbVisualizer where the connection between it and PgAdmin breaks and DbVisualizer can't access the database. Since no solution to this bug is known for now, you can use Lucidchart as an alternative tool.

- **Step 3. Create the first draft of a data dictionary:**
- Take a moment to examine your ERD. Does the Rockbuster database have a snowflake schema or a star schema? Write a brief explanation for your answer.
 - It is a star schema based both on its complexity and utilization of dimension and sub dimension tables.
- List all the fact tables and all the dimension tables in the schema. For each table, list every column and its data type, and write a brief description of the column. To get an idea of what this should look like, check out these [example fact and dimension tables](#).
- **Fact Tables:**
- **rental**

| Columns | Data Type | Description |
|--------------|--------------------------------|---|
| Rental_id | Serial | Unique identifier for each rental transaction. |
| Rental_date | Timestamp(6) Without Time Zone | Date and time when the rental was initiated. |
| Inventory_id | Integer | Unique identifier for the specific item being rented (e.g., video copy, game unit). |
| Customer_id | Small Integer | Unique identifier for the customer renting the item. |
| Return_date | Timestamp(6) Without Time Zone | Date and time when the rental was returned. |
| Staff_id | Small Integer | Unique identifier for the staff member who processed the rental. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- **Payment**

| Columns | Data Type | Description |
|--------------|--------------------------------|---|
| Payment_id | Serial | Unique identifier for each payment transaction. |
| Customer_id | Small Integer | Unique identifier for the customer who made the payment. |
| Staff_id | Small Integer | Unique identifier for the staff member who processed the payment. |
| Rental_id | Integer | Unique identifier for the rental associated with the payment. |
| Amount | Numeric(5,2) | Amount of the payment. |
| Payment_date | Timestamp(6) Without Time Zone | Date and time when the payment was made. |

- **Dimension Tables:**

- **film_category**

| Columns | Data Type | Description |
|-------------|--------------------------------|---|
| Film_id | Small Integer | Unique identifier for a specific film. |
| Category_id | Small Integer | Unique identifier for a specific film category (e.g., Action, Comedy, Documentary). |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- **category**

| Columns | Data Type | Description |
|-------------|--------------------------------|--|
| Category_id | Serial | Unique identifier for each film category. |
| Name | Character Varying (25) | Name of the film category (e.g., "Action", "Comedy", "Documentary"). |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- **film**

| Columns | Data Type | Description |
|------------------|--------------------------------|---|
| Film_id | Serial | Unique identifier for each film. |
| Title | Character Varying(255) | Title of the film. |
| Description | Text | Brief synopsis or description of the film. |
| Release_year | year | Year the film was released. |
| Language_id | Small Integer | Unique identifier for the language of the film. |
| Rental_duration | Small Integer | Number of days a customer can rent the film. |
| Rental_rate | Numeric(4,2) | Price to rent the film for the specified duration. |
| Length | Small Integer | Running time of the film in minutes. |
| Replacement_cost | Numeric(5,2) | Cost to replace a damaged or lost copy of the film. |
| Rating | Mpaa_rating | Maturity rating of the film (e.g., G, PG, R). |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |
| Special_features | Text[] | Special features included with the film (e.g., "Deleted Scenes", "Trailers"). |
| Fulltext | Tsvector | Generated index for full-text search on film attributes. |

-

- **actor**

| Columns | Data Type | Description |
|-------------|--------------------------------|--|
| Actor_id | Serial | Unique identifier for each actor. |
| First_name | Character Varying(45) | First name of the actor. |
| Last_name | Character Varying(45) | Last name of the actor. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- **film_actor**

| Columns | Data Type | Description |
|-------------|--------------------------------|--|
| Actor_id | Small Integer | Unique identifier for the actor. |
| Film_id | Small Integer | Unique identifier for the film. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- **inventory**

| Columns | Data Type | Description |
|--------------|--------------------------------|--|
| Inventory_id | Serial | Unique identifier for each copy of a film in a specific store. |
| Film_id | Small Integer | Unique identifier for the film. |
| Store_id | Small Integer | Unique identifier for the store. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- **language**

| Columns | Data Type | Description |
|-------------|--------------------------------|--|
| Language_id | Serial | Unique identifier for each language. |
| Name | Character(20) | Name of the language (e.g., "English", "Spanish", "French"). |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- **City**

| Columns | Data Type | Description |
|-------------|--------------------------------|--|
| City_id | Serial | Unique identifier for each city. |
| City | Character Varying(50) | Name of the city. |
| Country_id | Small Integer | Unique identifier for the country where the city is located. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- country

| Columns | Data Type | Description |
|-------------|--------------------------------|--|
| Country_id | Small Integer | Unique identifier for each country. |
| Country | Character Varying(50) | Name of the country. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- store

| Columns | Data Type | Description |
|------------------|--------------------------------|--|
| Store_id | Serial | Unique identifier for each store. |
| Manager_staff_id | Small Integer | Unique identifier for the staff member managing the store. |
| Address_id | Small Integer | Unique identifier for the store's address. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- staff

| Columns | Data Type | Description |
|-------------|--------------------------------|---|
| Staff_id | Serial | Unique identifier for each staff member. |
| First_name | Character Varying(45) | First name of the staff member. |
| Last_name | Character Varying(45) | Last name of the staff member. |
| Address_id | Small Integer | Unique identifier for the staff member's address. |
| email | Character Varying(50) | Email address of the staff member. |
| Store_id | Small Integer | Unique identifier for the store where the staff member is primarily assigned. |
| active | Boolean | Indicates whether the staff member is currently active or not. |
| username | Character Varying(16) | Username for system login. |
| password | Character Varying(40) | Password for system login. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |
| Picture | Bytea | Stores the staff member's photo or image. |

- **Customer**

| Columns | Data Type | Description |
|-------------|--------------------------------|---|
| Customer_id | Serial | Unique identifier for each customer. |
| Store_id | Small Integer | Unique identifier for the store where the customer is primarily associated. |
| First_name | Character Varying(45) | First name of the customer. |
| Last_name | Character Varying(45) | Last name of the customer. |
| email | Character Varying(50) | Email address of the customer. |
| Address_id | Small Integer | Unique identifier for the customer's address. |
| Activebool | Boolean | Indicates whether the customer is an active member or not. |
| Create_date | Date | Date when the customer account was created. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |
| Active | Integer | Some type of number showing status of customers. |

- **address**

| Columns | Data Type | Description |
|-------------|--------------------------------|--|
| Address_id | Serial | Unique identifier for each address. |
| address | Character Varying(50) | Street address (e.g., "123 Main St"). |
| Address2 | Character Varying(50) | Optional second address line (e.g., "Apt. 2B"). |
| district | Character Varying(20) | Neighborhood or district within the city. |
| City_id | Small Integer | Unique identifier for the city where the address is located. |
| Postal_code | Character Varying(10) | Postal code or ZIP code. |
| phone | Character Varying(20) | Phone number associated with the address. |
| Last_update | Timestamp(6) Without Time Zone | Date and time when the record was last modified. |

- If a column name doesn't tell you enough to write a description, you can also view the tables in pgAdmin 4. The SQL syntax for selecting a table is `SELECT * FROM table_name`. So `SELECT * FROM film` would return the film table, for example.

- **Data dictionary example**

This document is the first draft of a data dictionary that you'll refine later in this Achievement. Take a look at this sample [data dictionary for Adventure Works](#) to see what a finalized version looks like. Don't be alarmed by the sheer size of this example! Your first draft isn't supposed to look like this. Exploring the various tables and the information they contain will give you a rough idea of what we're looking for in this task.

- **Step 4. Find information:**

- Now that your data dictionary and ERD are ready to use, your manager has given you a list of business questions to answer. Use your data dictionary to figure out which tables you'd need to answer the questions below:

- Which actors brought Rockbuster the most revenue?
 - [Payment, rental, inventory, film, film_actor, actor](#)
- What language are the majority of movies in the collection?
 - [Language, film](#)

- **Step 5. Save and upload your file:**

- Save your "Answers 3.2" document as a PDF and upload it here for your tutor to review. Make sure to keep a copy for yourself! You'll be finalizing your data dictionary at the end of this Achievement.