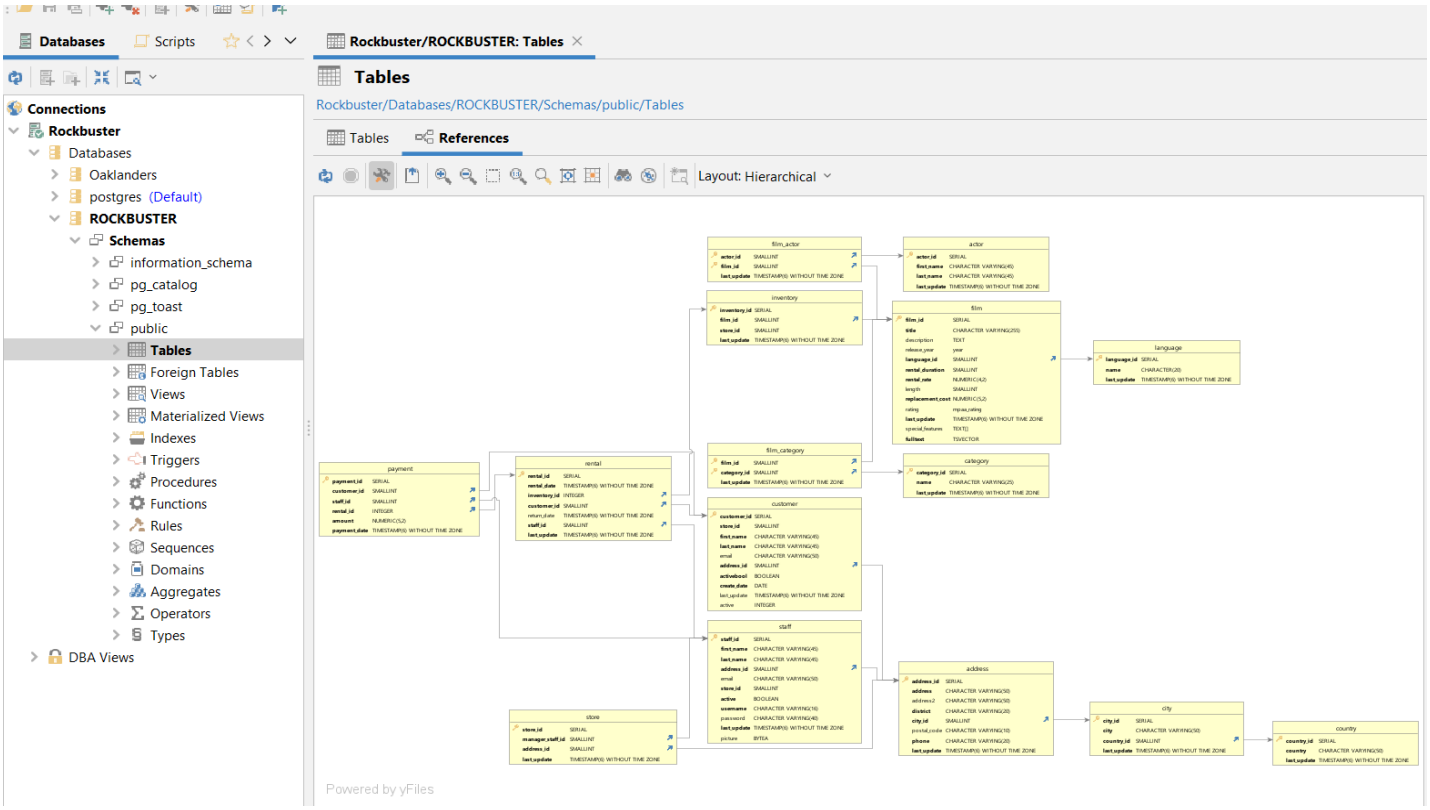
3.2: Data Storage & Structure

- **Download and install DbVisualizer or Lucidchart (I have done so already).**
- **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.**



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.**
  - ❖ The ERD of the Rockbuster database has a snowflake schema because apart from the fact table and dimension tables it's also further categorized into the subdimension 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 Table

Rental		
Column	**Data type**	**Description**
rental_id	SERIAL	Number assigned to rental
rental_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date of rental
Inventory_id	INTEGER	The number assigned to the item on the table
Customer_id	SMALLINT	Number assigned to customer
Return_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of return
Staff_id	SMALLINT	Number assigned to staff
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update

Dimension Tables

Payment		
Column	**Data Type**	**Description**
payment_id	SERIAL	Number assigned to payment
customer_id	SMALLINT	Number assigned to customers
staff_id	SMALLINT	Number assigned to staff
rental_id	INTEGER	Number assigned to rental
amount	NUMERIC(5,2)	Amount paid
payment_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date of payment

Store		
Column	**Date Type**	**Description**
store_id	SERIAL	Number assigned to store
manager-staff-id	SMALLINT	Number assigned to staff that are managers
address_id	SMALLINT	Number assigned to store address
Last_update	TIMESTAMP(6) WITHOUT ZONE	Date and time of update

Film_actor		
Column	**Date Type**	**Description**
actor_id	SMALLINT	Number assigned to the actor
film_id	SMALLINT	Number assigned to the film
Last_update	TIMESTAMP(6) WITHOUT ZONE	Date and time of update

Inventory		
Column	**Date Type**	**Description**
Inventory_id	SERIAL	Number assigned to items
Film_id	SMALLINT	Number assigned to film
Store_id	SMALLINT	Number assigned to store
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update

Film_Category		
Column	**Data Type**	**Description**
film_id	SMALLINT	Number assigned to film
Category_id	SMALLINT	Number assigned to category/genre
Last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date and time of update

Customer		
Column	**Data Type**	**Description**
customer_id	SERIAL	Number assigned to customer
Store_id	SMALLINT	Number assigned to store
First_name	CHARACTER VARYING(45)	Customer's first name
Last_name	CHARACTER VARYING(45)	Customer's last name
email	CHARACTER VARYING(50)	Customer's email
Address_id	SMALLINT	Number assigned to address
activebool	BOOLEAN	Customer's active status
Create_date	DATE	Date created
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update
active	INTEGER	Customer's active status

Staff		
Column	**Data Type**	**Description**
Staff_id	SERIAL	Number assigned for staff
First_name	CHARACTER VARYING(45)	Staff's first name
Last_name	CHARACTER VARYING(45)	Staff's last name
Address_id	SMALLINT	Number assigned for staff address
email	CHARACTER VARYING(50)	Staff's email
Store_id	SMALLINT	Number assigned to store
active	BOOLEAN	Staffs' active status
username	CHARACTER VARYING(16)	Staffs' username
password	CHARACTER VARYING(40)	Staffs' password
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update
picture	BYTEA	Staffs' picture

Actor		
Column	**Data Type**	**Description**
Actor_id	SERIAL	Number assigned to actor
First_name	CHARACTER VARYING(45)	Actor's first name
Last_name	CHARACTER VARYING(45)	Actors' last name
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update

Film		
Column	**Data Type**	**Description**
Film_id	SERIAL	Number assigned film
title	CHARACTER VARYING(255)	Film title
description	TEXT	Film description
Release_year	YEAR	Year film was released
Language_id	SMALLINT	Number assigned to language
Rental_duration	SMALLINT	How long the films are rented for
Rental_date	NUMERIC(4,2)	Date of rental
length	SMALLINT	Length of film
Replacement_cost	NUMERIC(5,2)	The amount it cost to replace a film.
rating	mpaa_rating	Film rating
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update
Special_features	TEXT[]	Special features included in the film
fulltext	TSVECTOR	Keywords associated with the film

Category		
Column	**Data Type**	**Description**
Category_id	SERIAL	Number assigned to category
name	CHARACTER VARYING(25)	Name of category
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of last update

Address		
Column	**Data Type**	**Description**
address_id	SERIAL	Number assigned to address
address	CHARACTER VARYING(50)	Street address
Address2	CHARACTER VARYING(50)	Supplementary address
district	CHARACTER VARYING(20)	District
City_id	SMALLINT	Number assigns to city
Postal_code	CHARACTER VARYING(10)	Postal code
phone	CHARACTER VARYING(20)	Phone number
Last _update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update

Language		
Column	**Data Type**	**Description**
Language_id	SERIAL	Number assigned to language
name	CHARACTER(20)	Name of language
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update

City		
Column	**Data Type**	**Description**
City_id	SERIAL	Number assigned to city
city	CHARACTER VARYING(50)	City's name
Country_id	SMALLINT	Number assigned to country
Last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date and time of update

Country		
Column	**Data Type**	**Description**
Country_id	SERIAL	Number assigned to country
country	CHARACTER VARYING(50)	Country name
Last_update	TIMESTAMP(6)WITHOUT TIME ZONE	Date and time of update

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?**
    - I would need the actor table to get each actor's name, then the film_actor table to see which actor is associated with each film and the payment table to see the amount paid for each film over the years.
- **What language are the majority of movies in the collection?**
    - I would need the language table to see the name of each language and the film table to see which language is associated with each film