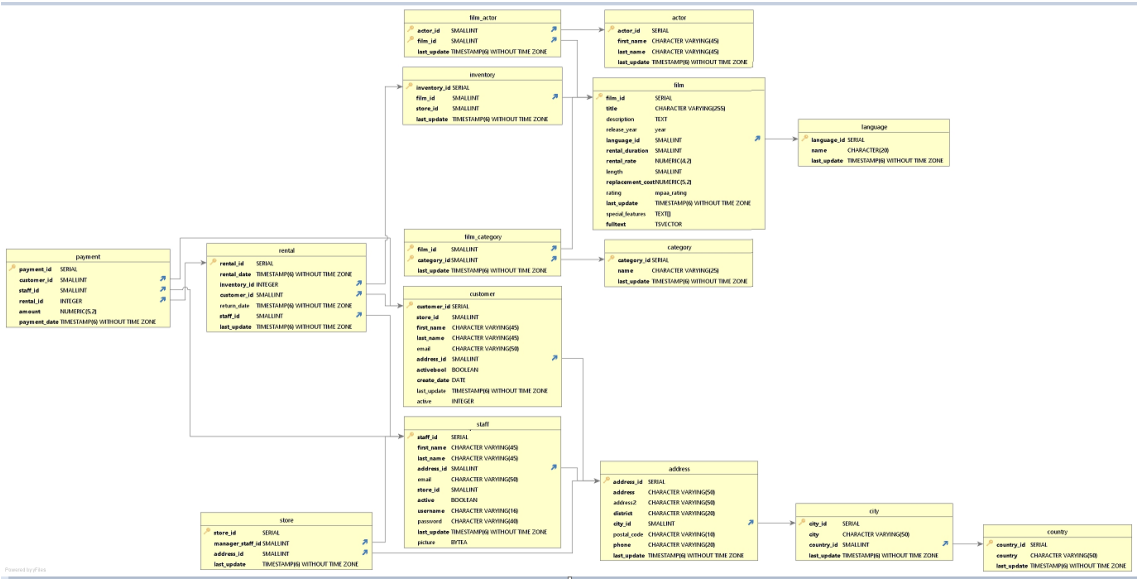
3.2: Data Storage & Structure

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 do
- cument.



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´s a snowflake shcema as the fact table is surrounded by dimensión 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.

- 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.

Fact Table

Rental

Columns	Data Type	Description
rental_id	SERIAL	Rental identification key
rental_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date of the rental (begining)
inventory_id	INTEGER	Inventory identification key
customer_id	SMALLINT	Customer identification key
return_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date of the rental (end)
staff_id	SMALLINT	Staff identification key
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

Dimension Tables

film

Columns	Data Type	Description
film_id	SERIAL	Film identification key
title	CHARACTER VARYING(25)	Title of the film
description	TEXT	Summary of the film
release_year	YEAR	Release year of the film
language_id	SMALLINT	Language identification key
rental_duration	SMALLINT	Calculation between rental_date and return_date
rental_rate	NUMERIC(4.2)	Amount to calculate price
length	SMALLINT	Duration of the film
replacement_cost	NUMERIC(5.2)	Price to pay in case of lost
rating	mpaa_rating	Film rating
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry
special_features	TEXT[]	Special features of the film (extras)
fulltext	TSVECTOR	

film_actor

Columns	Data Type	Description
actor_id	SMALLINT	Actor identification key
film_id	SMALLINT	Film identification key
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

actor

Columns	Data Type	Description
actor_id	SERIAL	Actor identification key
first_name	CHARACTER VARYING(45)	Name of the actor
last_name	CHARACTER VARYING(45)	Last name of the actor
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

film_category

Columns	Data Type	Description
film_id	SMALLINT	Film identification key
category_id	SMALLINT	Category identification key
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

category

Columns	Data Type	Description
category_id	SERIAL	Category identification key
name	CHARACTER VARYING(25)	Name of the category
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

language

Columns	Data Type	Description
language_id	SERIAL	Language identification key
name	CHARACTER(20)	Name of the language
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

store

Columns	Data Type	Description
store_id	SERIAL	Store identification key
manager_staff_id	SMALLINT	Manager identification key (there is no table for m
adress_id	SMALLINT	Address identification key
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

staff

Columns	Data Type	Description
staff_id	SERIAL	Staff identification key
first_name	CHARACTER VARYING(45)	Name of the person (Staff)
last_name	CHARACTER VARYING(45)	Last name of the person (Staff)
address_id	SMALLINT	Address identification key
email	CHARACTER VARYING(50)	Email of the person (Staff)
store_id	SMALLINT	Store identification key
active	BOOLEAN	Active/No active (contract of the staff)
username	CHARACTER VARYING(16)	username on the system
password	CHARACTER VARYING(40)	password on the system
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry
picture	BYTEA	Picture/No picture (contrpic of the staff)

address

Columns	Data Type	Description
adress_id	SERIAL	Address identification key
adress	CHARACTER VARYING(50)	Address description (1)
adress2	CHARACTER VARYING(50)	Address description (2)
district	CHARACTER VARYING(20)	District of the address
city_id	SMALLINT	City of the address
postal_code	CHARACTER VARYING(10)	Postal code of the address
phone	CHARACTER VARYING(20)	Phone of the address
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

city

Columns	Data Type	Description
city_id	SERIAL	City identification key
city	CHARACTER VARYING(50)	City name
country_id	SMALLINT	Country identification key
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

country

Columns	Data Type	Description
country_id	SERIAL	Country identification key
country	CHARACTER VARYING(50)	Country name
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

customer

Columns	Data Type	Description
customer_id	SERIAL	Customer identification key
store_id	SMALLINT	Store identification key
first_name	CHARACTER VARYING(45)	Name of the person (Customer)
last_name	CHARACTER VARYING(45)	Last name of the person (Customer)
email	CHARACTER VARYING(50)	Email of the person (Customer)
adress_id	SMALLINT	Address of the person (Customer)
activebool	BOOLEAN	Active/No active (contract with the store)
create_date	DATE	Date when the customer started
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry
active	INTEGER	

payment

Columns	Data Type	Description
payment_id	SERIAL	Payment identification key
customer_id	SMALLINT	Customer identification key
staff_id	SMALLINT	Staff identification key
rental_id	INTEGER	Rental identification key
amount	NUMERIC(5.2)	Price of the rental
payment_date	TIMESTAMP(6) WITHOUT TIME ZONE	Payment date of the rental#

inventory

Columns	Data Type	Description
inventory_id	SERIAL	Inventory identification key
film_id	SMALLINT	Film identification key
store_id	SMALLINT	Store identification key
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last update of the data entry

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?
  Rental → inventory → film → actor
- What language are the majority of movies in the collection?
  inventory → film → language

Step 5. Save and upload your file: