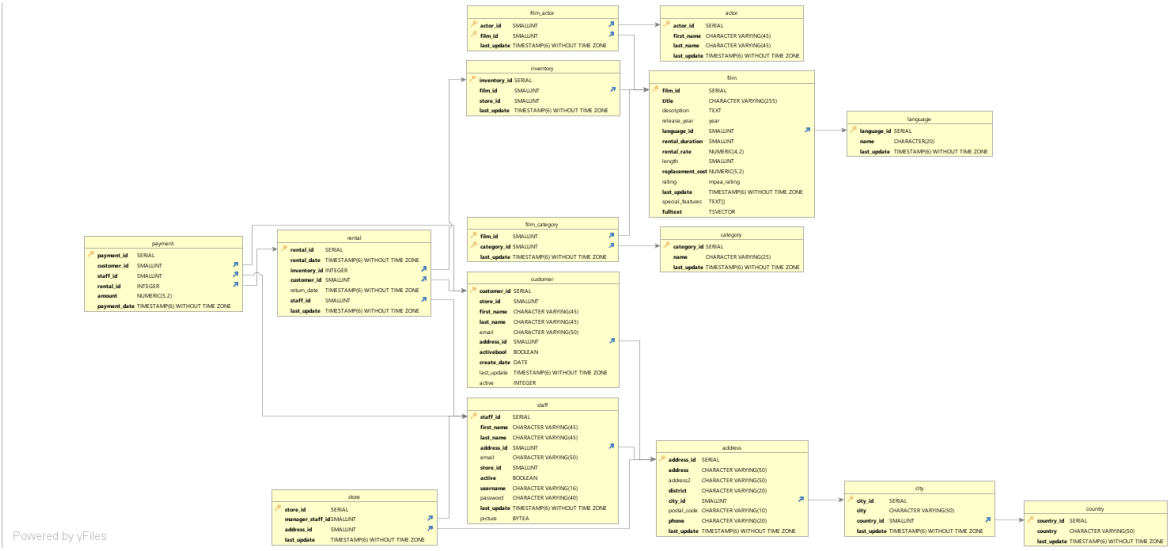
3.2: Data Storage and Structure

It's a snowflake schema because there are several dimensions and sub-dimensions generated from the other tables. In terms of address, for instance, the country is a dimension and the city is a sub-dimension. E.g Rental (as this an event that happened) which will be in a fact table. It relates to the following dimension tables which store data of the entities involved: payment for this rental, the film category, the customer who rented the film, the staff who assisted in the process, the store where rentals took place, and the address, city, and country of the store where the rental took place.



Fact table

Table	Column	Data Types	Description
Rental	Rental_id	Integer	A unique identifier for rental
	Rental_date	Timestamp without a time zone	The date and time when the rental is made
	Inventory_id	Integer	A unique identifier for inventory
	Customer_id	Smallint	A unique identifier for customer
	Return_date	Timestamp without a time zone	The date and time when the return is made
	Staff_id	Smallint	A unique identifier for the staff
	Last_update	Timestamp without a time zone	The date and time when the last update is made

DImension table

Table	Column	Data Types	Description
Actor	Actor_id	Integer	A unique identifier for the actor
	First_name	Character varying	First name of the actor

	Last_name	**Character varying**	**Last name of the actor**
	Last_update	Timestamp without a time zone	The date and time when the last update was made
Address	Address_id	Integer	A unique identifier for the address
	Address	Character varying	The first line of an address
	Address_2	Character varying	An optional second line of an address
	District	Character varying	The region of an address
	City_id	Smallint	The city of an address
	Postal_code	Character varying	The postal code for the address
	Phone	Character varying	The phone number for the address
	Last_update	Timestamp without a time zone	The date and time when the last update is made
City	City_id	Integer	A unique identifier for the city
	City	Character varying	Name of the city
	Country_id	Smallint	A unique identifier for the country
	Last_update	Timestamp without time zone	The date and time when the last update was made
Country	Country_id	Integer	A unique identifier for the country
	Country	Character varying	Name of country
	Last_update	Timestamp without time zone	A unique identifier for customer
Customer	Customer_id	Integer	A unique identifier for the customer
	Store_id	Character varying	First name of the customer
	First_name	Character varying	First name of the customer
	Last_name	Character varying	Last name of the customer
	Email	Character varying	Email of the customer
	Address_id	Smallint	A unique identifier for the address
	Activebool	Boolean	Status whether the customer is active or not
	Create_date	Date	The date when the customer is created in the system

	Last_update	Timestamp without time zone	The date and time when the last update is made
	Active	Integer	Number
Film	Film_id	Integer	A unique identifier for film
	Title	Character varying	Title of the film
	Description	Text	Description of the film
	Release_year	Integer	Year of the film release
	Language_id	Smallint	A unique identifier for the language
	Rental_duration	Smallint	The duration of the rental
	Rental_rate	Numeric	The date and time when the rental is made
	Length	Smallint	The length of the film
	Replacement_cost	Numeric	The cost of replacing the film
	Rating	USER-DEFINED	The rating of the film
Payment	Payment_id	Integer	A unique identifier for payments
	Customer_id	Smallint	A unique identifier for customer
	Staff_id	Smallint	A unique identifier for the staff
	Rental_id	Intger	A unique identifier for rental
	Amount	Numeric	Amount paid by the customer
	Payment_date	Timestamp without time zone	The date and time when a payment is paid.

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?

- Looking into payment, rental, inventory, film, and film_actor tables will be a good start.

What language is the majority of movies in the collection?

- These questions can be answered by looking into films and language tables and the inventory.