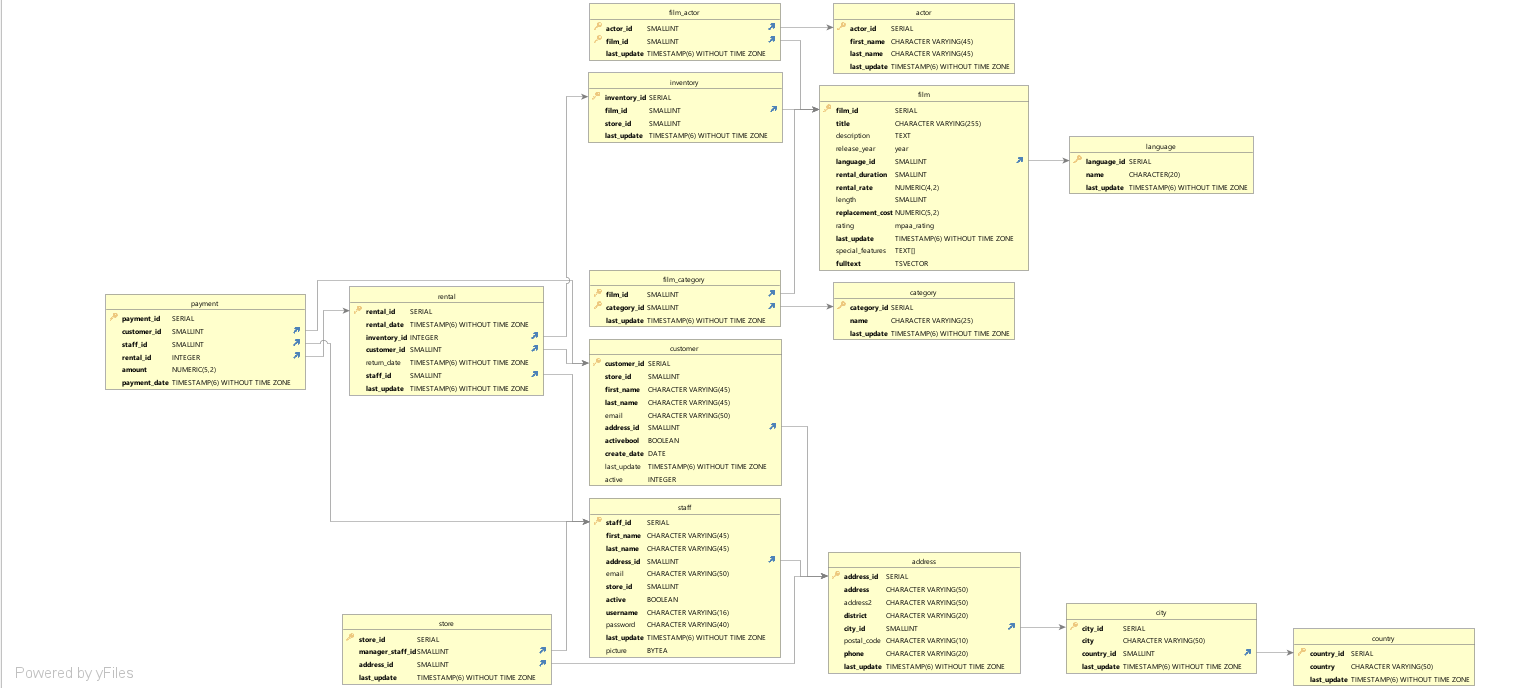
**3.2: Data Storage and Structure**

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.