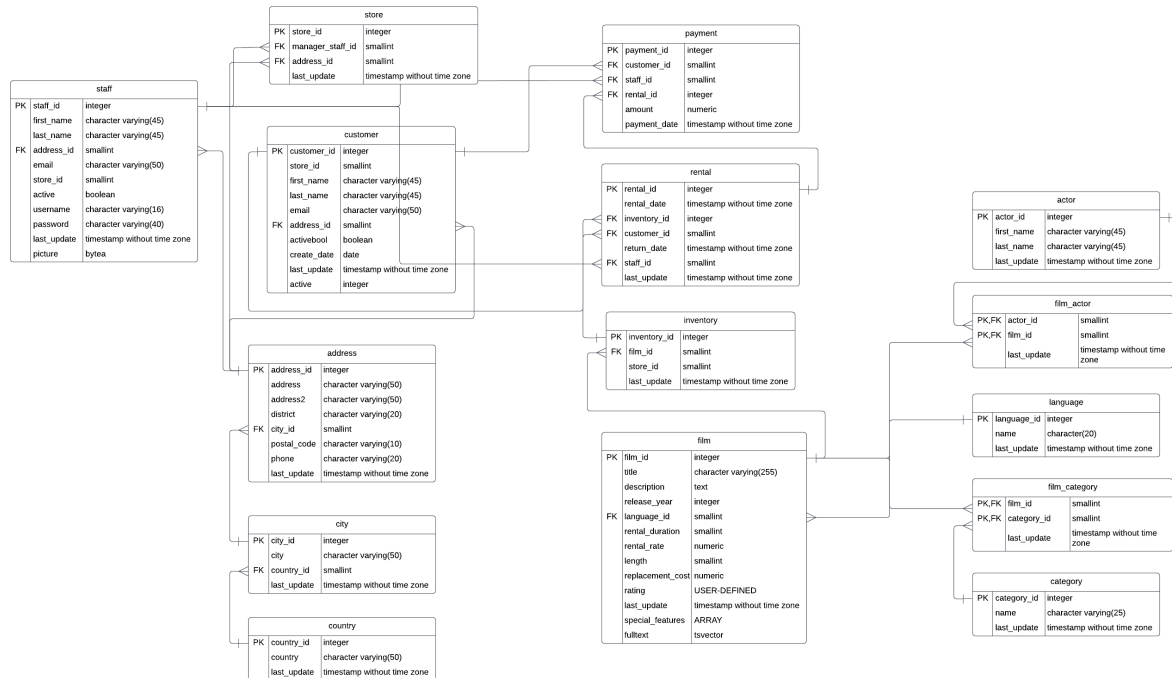# 3.2: Data Storage & Structure

Step 1. Done

Step 2. Extract the ERD:

Step 3. The schema is a Snowflake because there is no central part and tables are interlinked.

- List of all the fact tables and all the dimension tables in the schema.

Fact Tables are:- Rental

Columns	Data Type	Description
rental_id	SERIAL	The unique id assigned to each movie rental
rental_date	TIMESTAMP(6) WITHOUT TIME ZONE	Rental date and time
inventory_id	Integer	The id assigned for record keeping
customer_id	SMALLINT	Customer id to whom the movie was issued
return_date	TIMESTAMP(6) WITHOUT TIME ZONE	The date and time of return
staff_id	SMALLINT	The id assigned to each staff

Dimensional tables are:-

Inventory

Columns	Data Type	Description
inventory_id	SERIAL	The id assigned for record keeping
film_id	SMALLINT	The unique id assigned to each movie
store_id	SMALLINT	The unique id assigned to each store
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Customer

Columns	Data Type	Description
customer_id	SERIAL	Customer id to whom the movie was issued
store_id	SMALLINT	The unique id assigned to each store
first_name	CHARACTER VARYING(45)	First name of customer
last_name	CHARACTER VARYING(45)	Last name of customer
email	CHARACTER VARYING(50)	Email address of customer
address_id	SMALLINT	Postal address of customer
activebool	BOOLEAN	Status whether customer is active or not
create_date	DATE	Creation date
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Staff

Columns	Data Type	Description
staff_id	SERIAL	The id assigned to each staff
first_name	CHARACTER VARYING(45)	First name of staff
last_name	CHARACTER VARYING(45)	Last name of staff
address_id	SMALLINT	Postal address of staff
email	CHARACTER VARYING(50)	Email address of staff
store_id	SMALLINT	The unique id assigned to each store
active	BOOLEAN	Status whether customer is active or not
username	CHARACTER VARYING(16)	Username of staff
password	CHARACTER VARYING(40)	Password of staff
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Payment

Columns	Data Type	Description
payment_id	SERIAL	The id assigned for record keeping
customer_id	SMALLINT	Customer id to whom the movie was issued
staff_id	SMALLINT	The unique id assigned to each store
rental_id	INTEGER	The unique id assigned to each movie rental
amount	NUMERIC(5,2)	The amount paid by customer

Store

Columns	Data Type	Description
store_id	SERIAL	The unique id assigned to each store
manager_staff_id	SMALLINT	ID of Manager
address_id	SMALLINT	Postal address of store
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Film Category

Columns	Data Type	Description
film_id	SMALLINT	The unique id assigned to each movie
category_id	SMALLINT	The unique id assigned to each movie category
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Film Actor

Columns	Data Type	Description
actor_id	SMALLINT	The unique id assigned to an actor
film_id	SMALLINT	The unique id assigned to each movie
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Actor

Columns	Data Type	Description
actor_id	SERIAL	The unique id assigned to actor
first_name	CHARACTER VARYING(45)	First name of actor
last_name	CHARACTER VARYING(45)	Last name of actor
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Film

Columns	Data Type	Description
film_id	SERIAL	The unique id assigned to each movie
title	CHARACTER VARYING(255)	Title of the movie
description	TEXT	What is the movie about
release_year	year	Year of release
language_id	SMALLINT	ID assigned to language(s) movie is available in
rental_duration	SMALLINT	Length of time for which movie is on rent
rental_rate	NUMERIC(4,2)	Price of rent
length	SMALLINT	Duration of movie
replacement_cost	NUMERIC(5,2)	Cost of replacement in the case of loss or theft
rating	mpaa-rating	Movie rating
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status
special_features	TEXT[]	If there are any special features available (HD, subtitles etc)
fulltext	TSVECTOR	Key words

Category

Columns	Data Type	Description
category_id	SERIAL	The unique id assigned to each movie category
name	CHARACTER VARYING(25)	Name of category
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Address

Columns	Data Type	Description
address_id	SERIAL	Postal address of store
address	CHARACTER VARYING(50)	Postal address of store
address2	CHARACTER VARYING(50)	Postal address of store
district	CHARACTER VARYING(20)	Time of last updated status
city_id	SMALLINT	City
postal_code	CHARACTER VARYING(10)	Postal Code
Phone	CHARACTER VARYING(20)	Phone number
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Language

Columns	Data Type	Description
language_id	SERIAL	ID assigned to language(s) movie is available in
name	CHARACTER(20)	Name of language
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

City

Columns	Data Type	Description
city_id	SERIAL	Id assigned to City
name	CHARACTER VARYING(50)	Name of City
country_id	SMALLINT	Country code
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Country

Columns	Data Type	Description
country_id	SERIAL	Country code
country	CHARACTER VARYING(50)	Name of Country
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Step 4. Find information:

- Which actors brought Rockbuster the most revenue? – Film_actor, Film and Rental

- What language are the majority of movies in the collection? - Film and Language