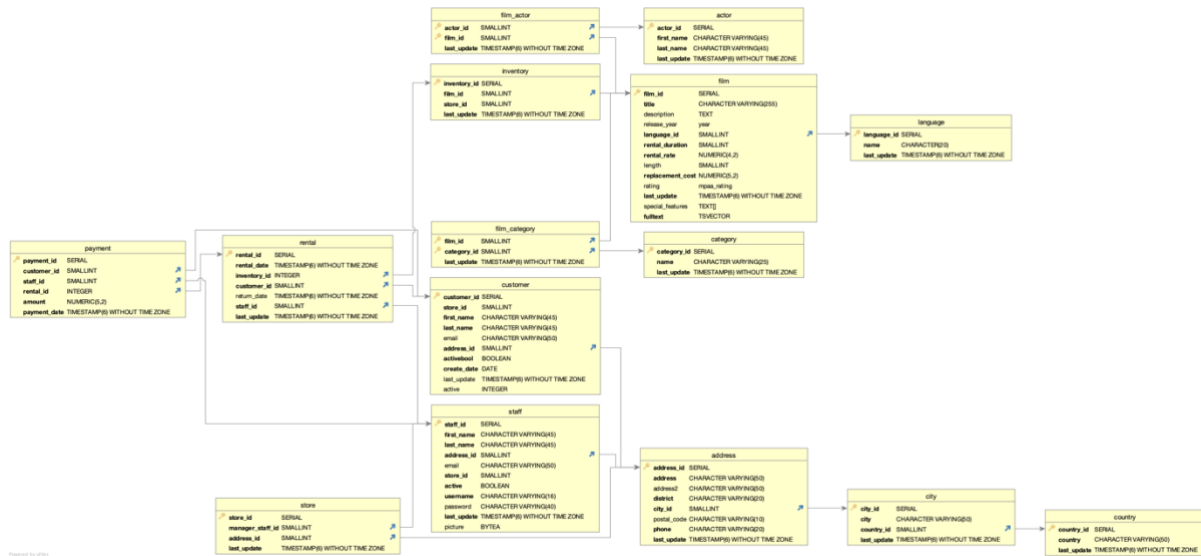# 3.2: DATA STORAGE & STRUCTURE

STEP 2-ROCKBUSTER ERD

From the look of it, the Rockbuster ERD has a snowflake schema. The dimension tables are further divided into sub-dimension tables making the schema more spread out.

STEP 3-DATA DICTIONARY

FACT TABLES

Table	Column	Data Type	Description
payment	payment_id	SERIAL	A unique payment ID
	customer_id	SMALLINT	A unique customer ID
	staff_id	SMALLINT	A unique staff ID
	rental_id	INTEGER	A unique ID
	amount	NUMERIC(5,2)	Amount paid for the rental
	payment_date	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date when the payment was made
rental	rental_id	INTEGER	A unique ID for rental
	rental_date	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date when the rental was made
	inventory_id	INTEGER	A unique ID for inventory
	customer_id	SMALLINT	A unique customer ID
	return_date	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date when the rental was returned
	staff_id	SMALLINT	A unique staff ID
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update

DIMENSION TABLES

Table	Column	Data Type	Description
store	store_id	SERIAL	A unique store ID
	manager_staff_id	SMALLINT	A unique store manager ID
	address_id	SMALLINT	Store address ID
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
customer	customer_id	SERIAL	A unique customer ID
	store_id	SMALLINT	A unique store ID
	first_name	CHARACTER VARYING(45)	First name of the customer
	last_name	CHARACTER VARYING(45)	Last name of the customer
	email	CHARACTER VARYING(50)	Customer's email
	address_id	SMALLINT	Customer's address
	activebool	BOOLEAN	Is the customer active? YES or NO
	create_date	DATE	Date the customer was registered
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
	active	INTEGER	Could be number of days the account/customer has been active?
staff	staff_id	SERIAL	A unique staff ID
	first_name	CHARACTER VARYING(45)	First name of the staff employee
	last_name	CHARACTER VARYING(45)	Last name of the staff employee
	address_id	SMALLINT	Employee's Address
	email	CHARACTER VARYING(50)	Employee's email
	store_id	SMALLINT	A unique store ID
	active	BOOLEAN	Is the employee still active? YES or NO
	username	CHARACTER VARYING(16)	Employee's system username
	password	CHARACTER VARYING(40)	Employee's system password
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
	picture	BYTEA	Employee's profile picture
address	address_id	SERIAL	A unique address ID
	address	CHARACTER VARYING(50)	Address
	address2	CHARACTER VARYING(50)	Address
	district	CHARACTER VARYING(20)	District
	city_id	SMALLINT	A unique city ID

Table	Column	Data Type	Description
	postal_code	CHARACTER VARYING(10)	Postal Code
	phone	CHARACTER VARYING(20)	Phone number
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
city	city_id	SERIAL	A unique city ID
	city	CHARACTER VARYING(50)	City
	country_id	SMALLINT	A unique country ID
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
country	country_id	SERIAL	A unique country ID
	country	CHARACTER VARYING(50)	Country
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
film_actor	actor_id	SMALLINT	A unique actor ID
	film_id	SMALLINT	A unique film ID
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
actor	actor_id	SERIAL	A unique actor ID
	first_name	CHARACTER VARYING(45)	First name of an actor
	last_name	CHARACTER VARYING(45)	Last name of an actor
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
film	film_id	SERIAL	A unique film ID
	title	CHARACTER VARYING(255)	Title of the movie
	description	TEXT	Text description of the movie
	release_year	year	Year the movie was released
	language_id	SMALLINT	A unique language ID
	rental_duration	SMALLINT	How long was the film rented for?
	rental_rate	NUMERIC(4,2)	How often was the film rented?
	length	SMALLINT	Length of the movie
	replacement_cost	NUMERIC(5,2)	Cost of replacing the film?
	rating	mpaa_rating	Film rating
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
	special_features	TEXT	Special features of the film
	fulltext	TSVECTOR	Used for fulltext search
language	language_id	SERIAL	A unique language ID
	name	CHARACTER(20)	Name of the language
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
inventory	inventory_id	SERIAL	A unique ID for inventory

Table	Column	Data Type	Description
	film_id	SMALLINT	A unique film ID
	store_id	SMALLINT	A unique store ID where the film is stored
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
film_category	film_id	SMALLINT	A unique film ID
	category_id	SMALLINT	A unique category ID
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update
category	category_id	SERIAL	A unique category ID
	name	CHARACTER VARYING(25)	Name of the category
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time and date of last update

---

STEP 4

Use your data dictionary to figure out which tables you'd need to answer the questions below:

1. Which actors brought Rockbuster the most revenue?
   a. I would investigate tables: actor, film_actor, film, inventory, rental and payment.

2. What language are the majority of movies in the collection?
   a. I would investigate tables: language, film, inventory