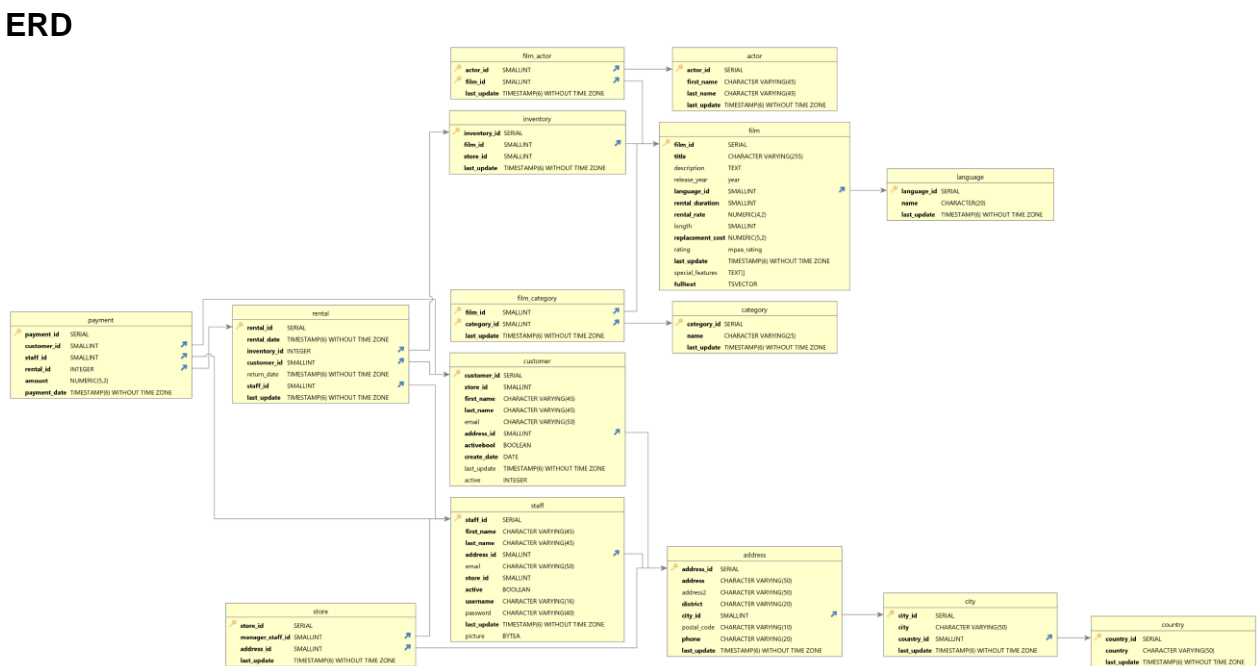


## 3.2 Data Storage & Structure

### 1. ERD



### 2. Does the Rockbuster database have a snowflake schema or a star schema?

Snowflake – because there is a *centralized fact table* (=> rental) which is connected to dimensions.

### 3. Fact table

Rental

Column	Data Type	Description
rental_id	Serial	Number assigned to rental
rental_date	TIMESTAMP(6) WITHOUT TIMEZONE	Date of rental
inventory_id	INT	Number assigned to item in the table
customer_id	SMALLINT	Number assigned to customer
return_date	TIMESTAMP(6) WITHOUT TIMEZONE	Date rental was returned
staff_id	SMALLINT	Number assigned to staff member
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## 4. Dimension tables

payment

Column	Data Type	Description
payment_id	Serial	Number assigned to rental
customer_id	SMALLINT	Number assigned to customer
staff_id	SMALLINT	Number assigned to staff member
rental_id	INT	Number assigned to rental
amount	NUMERIC(5, 2)	Amount paid
Payment_date	TIMESTAMP(6) WITHOUT TIMEZONE	Date of payment

store

Column	Data Type	Description
store_id	Serial	Number assigned to store
Manager_staff_id	SMALLINT	Number assigned to store manager
address_update	SMALLINT	Number assigned to address
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

film\_actor

Column	Data Type	Description
actor_id	SMALLINT	Number assigned to actor
film_id	SMALLINT	Number assigned to film
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

inventory

Column	Data Type	Description
inventory_id	Serial	Number assigned to item
film_id	SMALLINT	Number assigned to film
store_id	SMALLINT	Number assigned to store
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## film\_category

Column	Data Type	Description
film_id	SMALLINT	Number assigned to film
film_category	SMALLINT	Number assigned to category
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## customer

Column	Data Type	Description
customer_id	Serial	Number assigned to customer
store_id	SMALLINT	Number assigned to store
first_name	CHARACTERVARYING(50)	First name of customer
last_name	CHARACTERVARYING(45)	Last name of customer
email	CHARACTERVARYING(50)	Email of customer
address_id	SMALLINT	Number assigned to customer address
activebool	BOOLEAN	Customer active status
create_date	DATE	Date when created
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update
active	INT	Customer active status

## staff

Column	Data Type	Description
staff_id	Serial	Number assigned to staff member
first_name	CHARACTERVARYING(50)	First name of staff member
last_name	CHARACTERVARYING(45)	Last name of staff member
email	CHARACTERVARYING(50)	Email of staff member
address_id	SMALLINT	Number assigned to staff members address
store_id	SMALLINT	Number assigned to store
username	CHARACTERVARYING(16)	Username of staff member

password	CHARACTERVARYING(40)	Password of staff member
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update
active	BOOLEAN	Staff member active status
picture	BYETA	Picture of staff member

## actor

Column	Data Type	Description
actor_id	Serial	Number assigned to actor
first_name	CHARACTERVARYING(45)	First name of actor
last_name	CHARACTERVARYING(45)	Last name of actor
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## film

Column	Data Type	Description
film_id	Serial	Number assigned to film
title	CHARACTERVARYING(45)	Title of film
description	TEXT	Description of film
release_year	year	Release year of film
Language_id	SMALLINT	Number assigned to language
rental_duration	SMALLINT	Length of film rental
rental_rate	NUMERIC(4, 2)	Price of film rental
length	SMALLINT	Length of film
replacement_cost	NUMERIC(5, 2)	Cost of replacement of the film
rating	mpaa_rating	Film rating
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update
special_features	TEXT []	Special features of the film
fulltext	TSECTOR	Keywords of the film

## language

Column	Data Type	Description
language_id	Serial	Number assigned to language
name	CHARACTERVARYING(20)	Name of language
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## category

Column	Data Type	Description
category_id	Serial	Number assigned to category
name	CHARACTERVARYING(20)	Name of category
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## address

Column	Data Type	Description
address_id	Serial	Number assigned to address
address	CHARACTERVARYING(50)	Street address
address2	CHARACTERVARYING(50)	Supplementary street address
district	CHARACTERVARYING(20)	District
city_id	SMALLINT	Number assigned to city
postal_code	CHARACTERVARYING(10)	Postal code
phone	CHARACTERVARYING(20)	Phone number
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## city

Column	Data Type	Description
city_id	Serial	Number assigned to city
city	CHARACTERVARYING(20)	Name of city
country_id	SMALLINT	Number assigned to country
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

## country

Column	Data Type	Description
country_id	Serial	Number assigned to country
country	CHARACTERVARYING(20)	Name of country
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Date of last update

5. Which tables are needed for the following business questions?
- Which actors brought Rockbuster the most revenue?  
Actor, film\_actor, film, inventory, rental, payment
  - What language are the majority of movies in the collection?  
Language, film