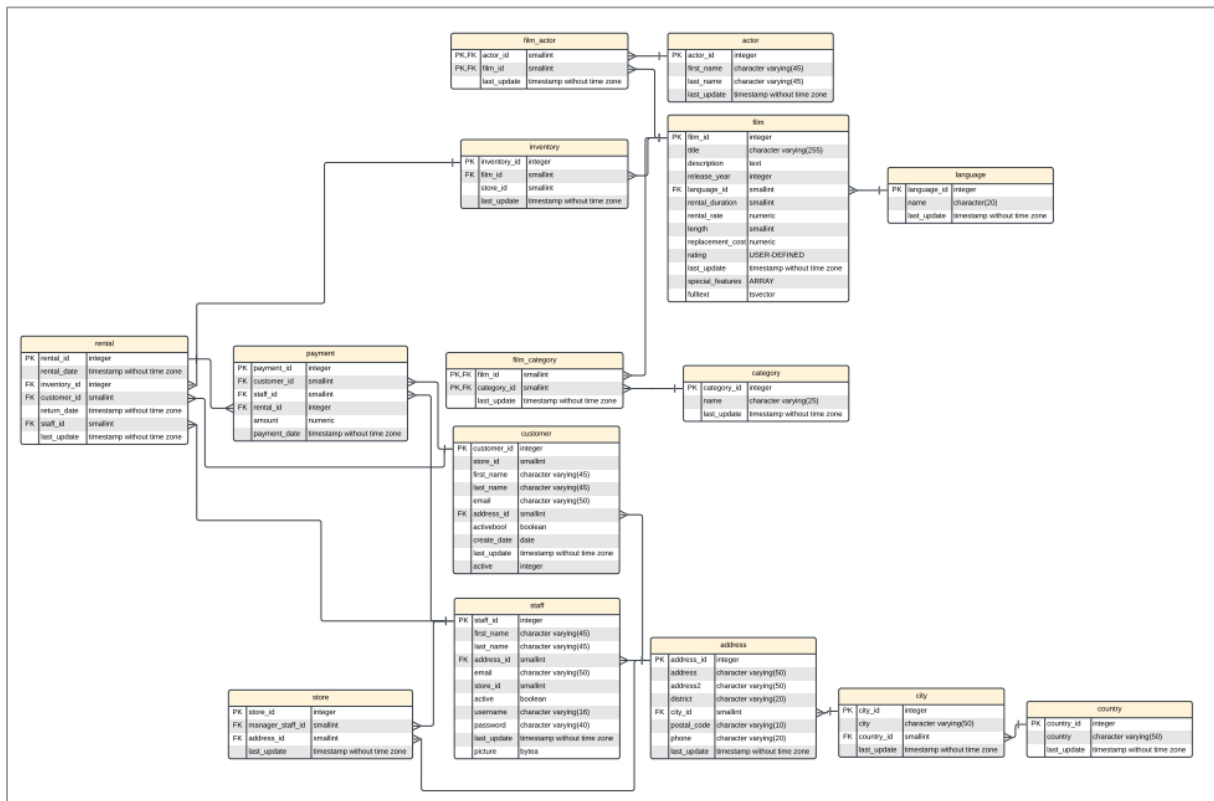


Exercise 3.2 - Data Storage & Structure

Step 2. Extract the ERD:

- Copy-paste the (Lucidchart) ERD into your answers document:



Step 3. Create the first draft of a data dictionary:

- Take a moment to examine your ERD. Does the Rockbuster database have a snowflake schema or a star schema?
 - It is a snowflake schema, as there is no central part in which all segments are connected. The fact table have multiple branches and are connected to multiple tables, creating subdimensions, typical to snowflake.
- List all the fact tables and all the dimension tables in the schema. For each table, list every column and its data type, and write a brief description of the column. To get an idea of what this should look like, check out these [example fact and dimension tables](#).

Fact Tables		
- Rental		
Columns	Data Type	Description
rental_id	Serial	ID number provided for the individual rent event.
rental_date	TIMESTAMP(6) WITHOUT TIMEZONE	Date in which the rent occurred.
inventory_id	INTEGER	ID number of the product contained in the rental.
customer_id	SMALLINT	ID number for the client.
return_date	TIMESTAMP(6) WITHOUT TIMEZONE	Date which the return should occur.
staff_id	SMALLINT	ID number of the staff member that worked on the event
last_update	TIMESTAMP(6) WITHOUT TIMEZONE	Last update on the event.

Dimension Tables

- Payment

Columns	Data Type	Description
payment_id	SERIAL	ID number for the transaction
customer_id	SMALLINT	ID number for the client.
staff_id	SMALLINT	ID number of the staff member that worked on the event
rental_id	INTEGER	ID number provided for the individual rent event
amount	NUMERIC (5,2)	Total monetary value of the payment
payment_date	TIMESTAMP(6) WIHOUT TIME ZONE	Date in which the payment was done

- Customer

Columns	Data Type	Description
customer_id	SERIAL	ID number for the client.
store_id	SMALLINT	ID number for the store
first_name	CHARACTER VARYING(45)	First name of the client
last_name	CHARACTER VARYING(45)	Last name of the client
email	CHARACTER VARYING(50)	Email address of the client
address_id	SMALLINT	ID number for the client address
activebool	BOOLEAN	Client is active or inactive option
create_date	DATE	Date which the customer was created on the system
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event
active	INETEGER	Client is active or inactive option

- Staff

Columns	Data Type	Description
staff_id	SERIAL	ID number of the staff member that worked on the event
first_name	CHARACTER VARYING(45)	First name of the staff member
last_name	CHARACTER VARYING(45)	Last name of the staff member
address_id	SMALLINT	Address of the staff member
email	CHARACTER VARYING(50)	Email of the staff member
store_id	SMALLINT	ID number for the store
active	BOOLEAN	Staff is active or inactive option
username	CHARACTER VARYING(16)) Staff login information
password	CHARACTER VARYING(40)	Staff password information
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.
picture	BYTEA	Staff is active or inactive option

- Film

Columns	Data Type	Description
film_id	SERIAL	ID number for the movie
title	CHARACTER VARYING(25)	Text, Title of the movie
description	TEXT	Text, description of the movie
release_year	year	Year in which the movie was released
language_id	SMALLINT	Languages which the movie is available
rental_duration	SMALLINT	Duration of the rental for this movie
rental_rate	NUMERIC(4,2)	Cost of the rental
length	SMALLINT	Length of the movie
replacement_cost	NUMERIC(5,2)	Cost of a replacement unity
rating	mpaa_rating	Rating of the movie according to the MPAA ruling
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.
special_feature	TEXT	Special features available in the movie
fulltext	TSVECTOR	Keywords, metatags

- Film_category		
Columns	Data Type	Description
film_id	SERIAL	ID number for the movie
category_id	SMALLINT	ID number for the category of the movie
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- Category		
Columns	Data Type	Description
category_id	SERIAL	ID number for the category of the movie
name	CHARACTER VARYING(25)	Name of the movie category
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- Film_actor		
Columns	Data Type	Description
actor_id	SMALLINT	ID number for the actor on the movie
film_id	SMALLINT	ID number for the movie
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- Inventory		
Columns	Data Type	Description
inventory_id	SERIAL	ID number of the product contained in the rental.
film_id	SMALLINT	ID number for the movie
store_id	SMALLINT	ID number of the store in which the event happened
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- Actor		
Columns	Data Type	Description
actor_id	SERIAL	ID number for the actor on the movie
first_name	CHARACTER VARYING(45)	Text, first name of the actor
last_name	CHARACTER VARYING(45)	Text, last name of the actor
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- Language		
Columns	Data Type	Description
language_id	SERIAL	ID for the language of the movie
name	CHARACTER(20)	Name of the language
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event

- Store		
Columns	Data Type	Description
store_id	SERIAL	ID number for the store.
manager_staff_id	SMALLINT	ID number of the manager
address_id	SMALLINT	Address of the store
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- Address		
Columns	Data Type	Description
address_id	SERIAL	ID number for the address
address	CHARACTER VARYING(50)	Address part 1
address2	CHARACTER VARYING(50)	Address part 2
district	CHARACTER VARYING(20)	District of the address
city_id	SMALLINT	City where the address is
postal_code	CHARACTER VARYING(10)	Postal code of the address
phone	CHARACTER VARYING(20)	Phone number registered with address
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- City		
Columns	Data Type	Description
city_id	CHARACTER VARYING(50)	ID number for the City
country_id	SMALLINT	ID number for the Country
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- Country		
Columns	Data Type	Description
country_id	SERIAL	ID number for the country
country	CHARACTER VARYING(50)	Country name
last_update	TIMESTAMP(6) WIHOUT TIME ZONE	Last update on the event.

- If a column name doesn't tell you enough to write a description, you can also view the tables in pgAdmin 4. The SQL syntax for selecting a table is `SELECT * FROM table_name`. So `SELECT * FROM film` would return the film table, for example.

Step 4. Find information:

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?
→Film_actor and Film and Rental.
- What language are the majority of movies in the collection?
→Film and Language