

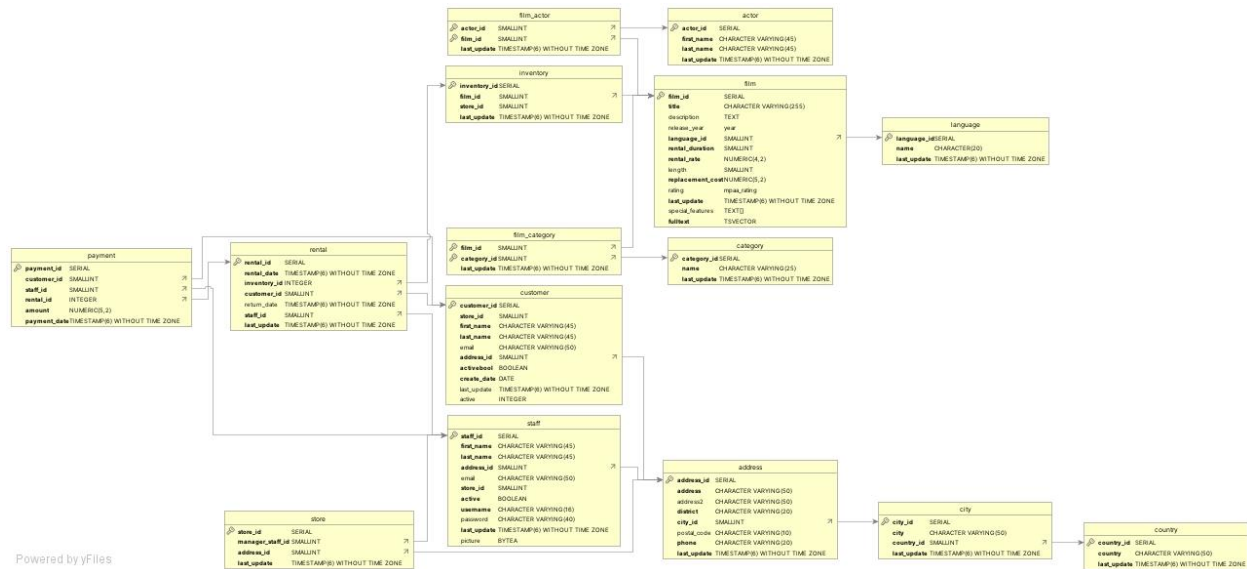
Career Foundry

Data Analytics Immersion

A3.E2

Kendra Jackson

Step 2



Step 3

1. Snowflake Schema. Dimensions have sub-dimension connections.

Fact Table

Rental

Column	Data Type	Description
Rental_id	Serial	Unique identification number assigned to each rental
Rental_date	Timestamp(6) without time zone	Date of rental
Inventory_id	Integer	Identification number assigned for record keeping of stock
Customer_id	Smallint	Identification number assigned to each customer when rental issued
Return_date	Timestamp(6) without time zone	Date of rental return

Staff_id	Smallint	Identification number assigned to each staff member
Last_update	Timestamp(6) without time zone	Date? of last update or time in 00:00:00 format?

Payment

Column	Data Type	Description
Payment_id	Serial	Unique identification number assigned to each payment
Customer_id	Smallint	Identification number assigned to each customer
Staff_id	Smallint	Identification number assigned to each staff (likely those involved in the transaction)
Rental_id	Integer	Unique identification number assigned to each rental
Amount	Numberis (5,2)	Amount paid for rental by customer
Payment_date	Timestamp(6) without time zone	Date of payment

Dimension Table (entities involved in event)

Inventory

Column	Data Type	Description
Inventory_id	Serial	Identification number assigned for record keeping of stock
Film_id	Smallint	Identification number assigned to each film
Store_id	Smallint	Identification number assigned to each store
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Staff

Column	Data Type	Description
Staff_id	Serial	Identification number assigned to each staff member
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	Identification number assigned to each store
Active	Boolean	Status of whether an employee is still active or not (current employee vs previous employee)
Username	Character varying (16)	Staff username
Password	Character varying (40)	Password staff use in conjunction with username to login
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?
Picture	Bytea	Picture of staff member

Customer

Column	Data Type	Description
Customer_id	Serial	Identification number assigned to each customer
Store_id	Smallint	Identification number 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 of whether an customer is still active or not
Create_Date	Date	Date of when a customer's account was created
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?
Active	Integer	Status of customer

Film Catgeory

Column	Data Type	Description
Film_id	Smallint	Unique identification number assigned to each film
Film_Category	Smallint	Number assigned to each category of film
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Film Actor

Column	Data Type	Description
Actor_id	Smallint	Identification number assigned to each actor
Film_id	Smallint	Identification number assigned to each film
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Store

Column	Data Type	Description
Store_id	Serial	Unique identification number assigned to each store
Manager_staff_id	Smallint	Identification number assigned to manager
Address_id	Smallint	Postal address of store
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Actor

Column	Data Type	Description
Actor_id	Serial	Identification number assigned to each 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	Date? Of last update or time in 00:00:00 format?

Film

Column	Data Type	Description
Film_id	Serial	Unique identification number assigned to each film
Title	Character Varying (255)	Title of film
Description	Text	Description of what film is about
Release_year	Year	The year the film was released
Language_id	Smallint	Identification number assigned to the language the film is available in
Rental_duration	Smallint	Length of the time that film was rented out for
Rental_rate	Numeric (4,2)	Cost to rent film
Length	Smallint	Length of film in minutes

Replacement_cost	Numeric (5,2)	Cost to replace film if lost, damaged, stolen
Rating	Mpaa_rating	Rating of film
Last_Update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?
Special_features	Text []	Any special features included in film (CC, sign language)
Fulltext	TSVECTOR	Keywords associated with film

Language

Column	Data Type	Description
Language_id	Serial	Identification number assigned to the language the film is available in
Name	Character (20)	Name of language
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Category

Column	Data Type	Description
Catgeory_id	Serial	Identification number assigned to the category the film falls under
Name	Character Varying (20)	Name of category
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Address

Column	Data Type	Description
Address_id	Serial	Unique identification number assigned to each store address
Address	Character varying (50)	Address of Store
Address2	Character varying (50)	Additional address of Store
District	Character varying (20)	District of store
City_id	Smallint	Identification number assigned to city of store
Postal_code	Character varying (10)	Postal code of city of store
Phone	Character varying (20)	Store phone number
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

City

Column	Data Type	Description
City_id	Serial	Identification number assigned to the each city
City	Character Varying (50)	Name of city
Country_id	Smallint	Identification number assigned to each country
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Country

Column	Data Type	Description
Country_id	Serial	Identification number assigned to each country
Name	Character Varying (50)	Name of country
Last_update	Timestamp(6) without time zone	Date? Of last update or time in 00:00:00 format?

Step 4

1. Which actors brought Rockbuster the most revenue
 - a. Payment, Rental, Inventory, Film_Actor, Actor
2. What language are the majority of movies in the collection
 - a. Language and film