

Task 3.2: Data Storage and Structure

Step 1. Create your Answers document.

Step 2. Extract the ERD.



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

The Rockbuster database has a snowflake schema because the fact table is surrounded by multiple sub-dimension tables.

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
Customer_id	SMALLINT	Identification number assigned to each customer when the rental was issued
Return_date	TIMESTAMP (6) WITHOUT TIME ZONE	Date and time of return

Staff_id	SMALLINT	Identification number assigned to each staff
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last update

Dimension tables:

Inventory

Column	Data type	Description
Inventory_id	SERIAL	Unique identification number assigned for record keeping
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	Time of last updated status

Staff

Column	Data Type	Description
Staff_id	SERIAL	Unique identification number 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 (45)	Email address of staff
Store_id	SMALLINT	Identification number assigned to each store
Active	BOOLEAN	Whether the status is active or not
Username	CHARACTER VARYING (16)	Username of staff

Password	CHARACTER VARYING (40)	Password that staff used with username
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status
Picture	BYTEA	Picture of staff member

Customer

Column	Data Type	Description
Customer_id	SERIAL	Unique 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	Whether status of customer is active or not
Create_date	DATE	Date of when customer's account was created
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status
Active	INTEGER	Status of customer

Film category

Column	Data type	Description
Film_id	SMALLINT	Unique identification number assigned to each film
Film_category	SMALLINT	Unique identification number assigned to each category

Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status
--------------------	------------------------------------	-----------------------------

Film Actor

Column	Data type	Description
Actor_id	SMALLINT	Unique identification number assigned to each actor
Film_id	SMALLINT	Unique identification number assigned to each film
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status

Store

Column	Data type	Description
Store_id	SERIAL	Unique identification number assigned to each store
Manager_staff_id	SMALLINT	Identification number of manager of store
Address_id	SMALLINT	Postal address of store
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status

Payment

Column	Data type	Description
Payment_id	SERIAL	Unique Identification number assigned for payment
Customer_id	SMALLINT	Identification number assigned to each customer

Staff_id	SMALLINT	Identification number assigned to each staff
Rental_id	INTEGER	Identification number assigned to each rental
Amount	NUMERIC (5,2)	Amount paid by customer
Payment_date	TIMESTAMP (6) WITHOUT TIME ZONE	Time of payment

Actor

Column	Data type	Description
Actor_id	SERIAL	Unique 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	Time of last updated status

Film

Column	Data Type	Description
Film_id	SERIAL	Unique identification number assigned to each film
Title	CHARACTER VARYING (255)	Title of film
Description	TEXT	What the film is about
Release_year	Year	The year that the film was released
Language_id	SMALLINT	Identification number assigned to the language(s) the film is available in
Rental_duration	SMALLINT	Length of time that film was rented out for

Rental_rate	NUMERIC (4,2)	Total revenue of each film rental
Length	SMALLINT	Length of film in minutes
Replacement_cost	NUMERIC (5,2)	Cost to replace film if lost or defective
Rating	Mpaa_rating	Rating of film
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status
Special_features	TEXT[]	Any special feature included in the film (ex. Sign language)
Fulltext	TSVECTOR	Keywords

Language

Column	Data type	Description
Language_id	SERIAL	Unique identification number assigned to each language
Name	CHARACTER (20)	Name of language
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status

Category

Column	Data type	Description
Category_id	SERIAL	Unique identification number assigned to each category
Name	CHARACTER VARYING (25)	Name of category
Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status

Address

Column	Data type	Description
--------	-----------	-------------

Address_id	SERIAL	Unique identification number assigned to each address of store
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 THE ZONE	Time of last updated status

City

Column	Data type	Description
City_id	SERIAL	Unique identification number assigned to 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	Time of last updated status

Country

Column	Data type	Description
Country_id	SERIAL	Unique identification number assigned to each country
Name	CHARACTER VARYING (50)	Name of country

Juliana Lee

Last_update	TIMESTAMP (6) WITHOUT TIME ZONE	Time of last updated status
--------------------	------------------------------------	-----------------------------

Step 4. Find information.

The film actor and film tables are needed to find out which actors brought Rockbuster the most revenue.

The film and language tables are needed to find out what language most films are in.