

Databases & SQL for Analysts

Task 3.2: Data Storage & Structure

Ngawang Dhundup

[LINK NAME:](#)

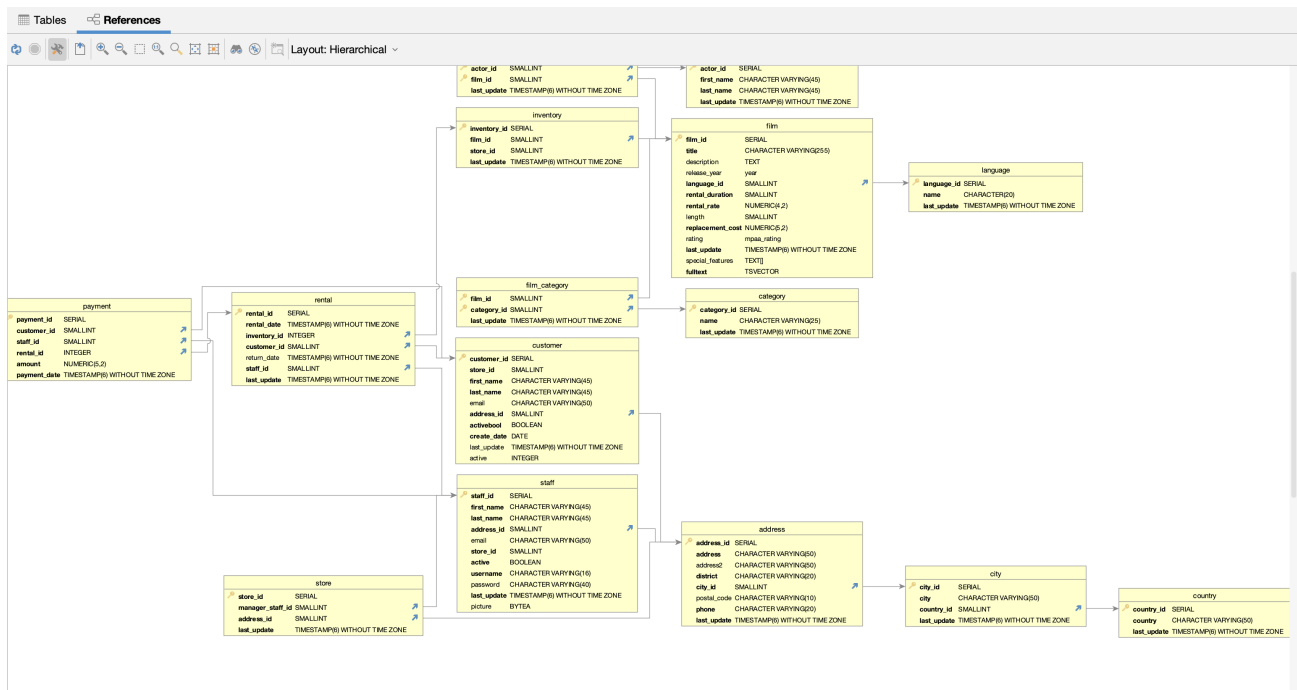
Directions:

Step 1. Create your Answers document:

Create a new text document and call it “Answers 3.2.” You'll save a copy of your ERD, data dictionary, and written answers in this document.

Step 2. Extract the ERD:

- Download and install [DbVisualizer](#) or [Lucidchart](#) (if you haven't already done so).
- Extract the ERD from the Rockbuster database and save it as an image (PNG or JPEG) using the instructions in the Exercise.
- Copy-paste the 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? Write a brief explanation for your answer.
- The Rockbuster database has a snowflake schema because the main table is surrounded by other sub dimension tables
- 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 Table:

Rental

Key	Columns	Data Type	Description
Key	rental_id	SERIAL	The unique id assigned to each movie rental
	rental_date	TIMESTAMP(6) WITHOUT TIME ZONE	The date and time of when it was rented out
	inventory_id	Integer	The id assigned for record keeping
	customer_id	SMALLINT	Customer id to whom the movie was issued
	return_date	TIMESTAMP(6) WITHOUT TIME ZONE	The date and time of return
	staff_id	SMALLINT	The id assigned to each staff
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Dimension Tables:

Inventory

Key	Columns	Data Type	Description
Key	inventory_id	SERIAL	The id assigned for record keeping
	film_id	SMALLINT	The unique id assigned to each movie
	store_id	SMALLINT	The unique id assigned to each store
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Customer

Key	Columns	Data Type	Description
Key	customer_id	SERIAL	Customer id to whom the movie was issued
	store_id	SMALLINT	The unique id 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 whether customer is active or not
	create_date	DATE	Creation date
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status
	active	INTEGER	Status of customer

Staff

Key	Columns	Data Type	Description
Key	staff_id	SERIAL	The id 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(50)	Email address of staff
	store_id	SMALLINT	The unique id assigned to each store
	active	BOOLEAN	Status whether customer is active or not
	username	CHARACTER VARYING(16)	Username of staff
	password	CHARACTER VARYING(40)	Password of staff
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status
	picture	BYTEA	Picture of staff

Payment

Key	Columns	Data Type	Description
Key	payment_id	SERIAL	The id assigned for record keeping
	customer_id	SMALLINT	Customer id to whom the movie was issued
	staff_id	SMALLINT	The unique id assigned to each store
	rental_id	INTEGER	The unique id assigned to each movie rental
	amount	NUMERIC(5,2)	The amount paid by customer
	payment_date	TIMESTAMP(6) WITHOUT TIME ZONE	Time of payment

Store

Key	Columns	Data Type	Description
Key	store_id	SERIAL	The unique id assigned to each store
	manager_staff_id	SMALLINT	ID of Manager
	address_id	SMALLINT	Postal address of store
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Film Category

Key	Columns	Data Type	Description
Key	film_id	SMALLINT	The unique id assigned to each movie
Key	category_id	SMALLINT	The unique id assigned to each movie category
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Film Actor

Key	Columns	Data Type	Description
Key	actor_id	SMALLINT	The unique id assigned to an actor
Key	film_id	SMALLINT	The unique id assigned to each movie
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Actor

Key	Columns	Data Type	Description
Key	actor_id	SERIAL	The unique id assigned to an 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

Key	Columns	Data Type	Description
Key	film_id	SERIAL	The unique id assigned to each movie
	title	CHARACTER VARYING(255)	Title of the movie
	description	TEXT	What is the movie about
	release_year	year	Year of release
	language_id	SMALLINT	ID assigned to language(s) movie is available in
	rental_duration	SMALLINT	Length of time for which movie is on rent
	rental_rate	NUMERIC(4,2)	Price of rent
	length	SMALLINT	Duration of movie
	replacement_cost	NUMERIC(5,2)	Cost of replacement in the case of loss or theft
	rating	mpaa-rating	Movie rating
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status
	special_features	TEXT[]	If there are any special features available (HD, subtitles etc)
	fulltext	TSVECTOR	Key words

Category

Key	Columns	Data Type	Description
Key	category_id	SERIAL	The unique id assigned to each movie category
	name	CHARACTER VARYING(25)	Name of category
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Address

Key	Columns	Data Type	Description
Key	address_id	SERIAL	Postal address of store
	address	CHARACTER VARYING(50)	Postal address of store
	address2	CHARACTER VARYING(50)	Postal address of store
	district	CHARACTER VARYING(20)	Time of last updated status
	city_id	SMALLINT	City
	postal_code	CHARACTER VARYING(10)	Postal Code
	Phone	CHARACTER VARYING(20)	Phone number
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Language

Key	Columns	Data Type	Description
Key	language_id	SERIAL	ID assigned to language(s) movie is available in
	name	CHARACTER(20)	Name of language
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

City

Key	Columns	Data Type	Description
Key	city_id	SERIAL	Id assigned to City
	name	CHARACTER VARYING(50)	Name of City
	country_id	SMALLINT	Country code
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Country

Key	Columns	Data Type	Description
Key	country_id	SERIAL	Country code
	country	CHARACTER VARYING(50)	Name of Country
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time of last updated status

Keys

Key	Columns	Data Type	Description	Links to table
Key	rental_id	SERIAL	The unique id assigned to each movie rental	Inventory, Customer, Staff
Key	inventory_id	SERIAL	The id assigned for record keeping	Film
Key	customer_id	SERIAL	Customer id to whom the movie was issued	Address
Key	staff_id	SERIAL	The id assigned to each staff	Address
Key	payment_id	SERIAL	The id assigned for record keeping	Customer, Staff, Rental
Key	store_id	SERIAL	The unique id assigned to each store	Staff, Address
Key	category_id	SMALLINT	The unique id assigned to each movie category	Film_category
Key	actor_id	SMALLINT	The unique id assigned to an actor	Film_actor, Actor
Key	film_id	SMALLINT	The unique id assigned to each movie	Film, Film_category, Film_actor, Inventory,
Key	actor_id	SERIAL	The unique id assigned to an actor	Film_actor, Actor
Key	film_id	SERIAL	The unique id assigned to each movie	Film, Film_category, Film_actor, Inventory,
Key	category_id	SERIAL	The unique id assigned to each movie category	Film_category
Key	address_id	SERIAL	Postal address of store	Staff, Store, Customer
Key	language_id	SERIAL	ID assigned to language(s) movie is available in	Film
Key	city_id	SERIAL	Id assigned to City	Address
Key	country_id	SERIAL	Country code	City

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](#), [Film](#), [Rental](#), [Actor](#)
- What language are the majority of movies in the collection?
- [Film](#), [Language](#)

Step 5. Save and upload your file:

Save your “Answers 3.2” document as a PDF and upload it here for your tutor to review. Make sure to keep a copy for yourself! You’ll be finalizing your data dictionary at the end of this Achievement.