Rockbuster Stealth

Data Dictionary

Prepared for CareerFoundry

By

Innocent Bayai

17 June 2024
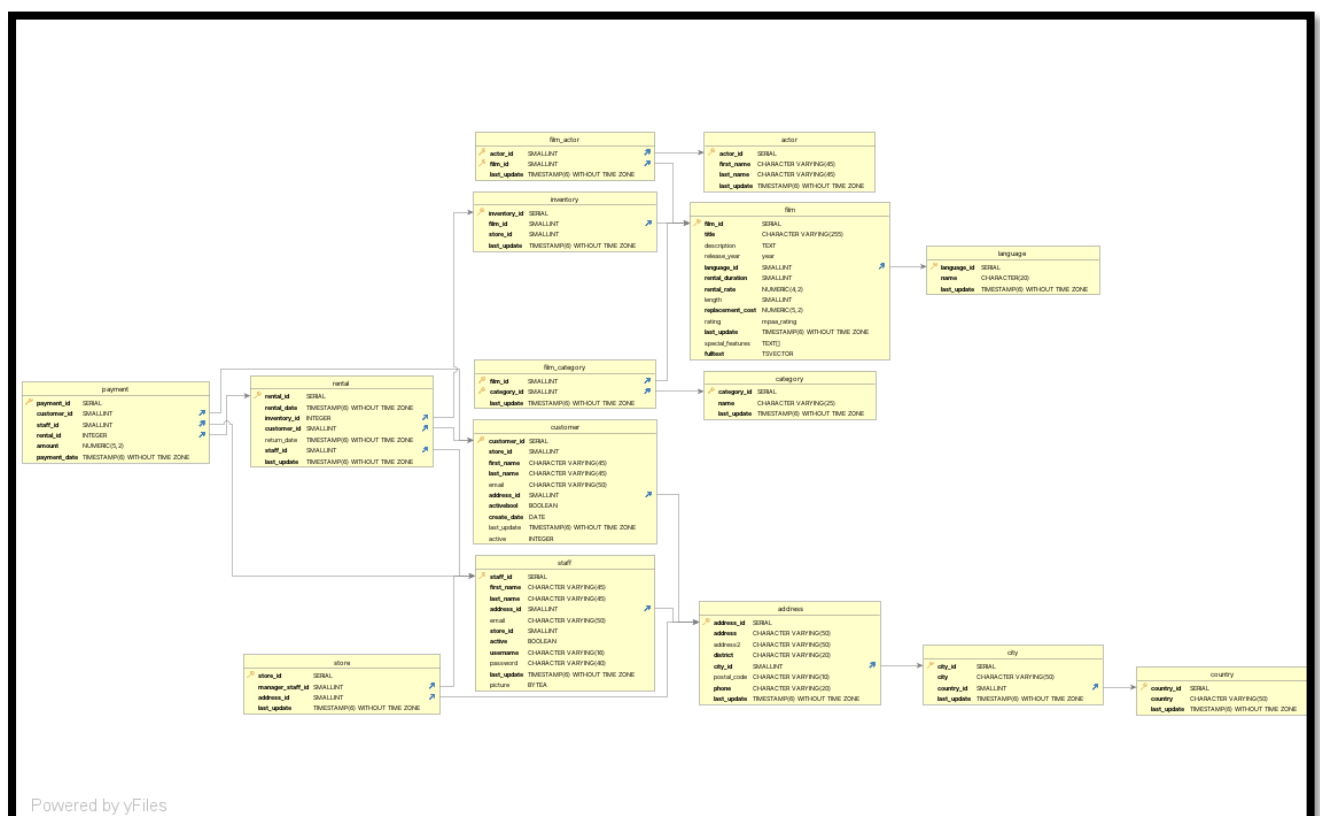
Contents

Overview: Rockbuster Stealth

Rockbuster Stealth LLC is a movie rental company that used to have stores around the world. Facing stiff competition from streaming services such as Netflix and Amazon Prime, the Rockbuster Stealth management team is planning to use its existing movie licenses to launch an online video rental service in order to stay competitive.

The dataset provided contains information about Rockbuster's film inventory, film categories, ratings, customers, and payments, among other things. This data dictionary attempts to provide the structure of the data and the connections between different tables through primary and foreign keys. An Entity Relationship Diagram (ERD) presented hereunder provides the summarized version of all tables (fact and dimension) as well as the keys. The next section provides details of the tables, that is the different columns and keys.

Figure 1: ERD for Rockbuster Stealth LLC

The Rockbuster database takes the Snowfake schema although the schema has a complex structure. The complexity comes from the fact that there is no single and central fact table linked to dimension tables. Rather, one fact table has dimension tables that in turn have their own dimension tables. Details of the content of the tables are presented in the next section.

Fact and dimension tables for the Rockbuster database

Fact Tables

Payment

Columns	Key	Data type	Description
payment_id	PK	SERIAL	A unique payment identifier.
customer_id	FK	SMALLINT	A unique small integer used to identify a customer.
staff_id	FK	SMALLINT	A unique small integer used to identify Rockbuster's individual staff.
rental_id	FK	INTEGER	A unique integer (without a decimal) used to identify a video rental transaction.
amount		NUMERIC(5,2)	Monetary value comprising a max of 5 digits, and 2 decimals.
payment_date		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for payment date (dd-mm-yy) and time (hh:mm:ss).

Rental

Columns	Key	Data type	Description
rental_id	PK	SERIAL	A unique identifier for a rental transaction.
rental_date		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for video rental starting date (dd-mm-yy) and time (hh:mm:ss).
inventory_id	FK	INTEGER	A unique integer (without decimals) used to identify Rockbuster's rental inventory.
customer_id	FK	SMALLINT	A unique small integer used to identify a video rental customer.
return_date		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for video rental return date (dd-mm-yy) and time (hh:mm:ss).
staff_id	FK	SMALLINT	A unique small integer used to identify Rockbuster's individual staff.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Dimension Tables

Inventory

Columns	Key	Data type	Description
inventory_id	PK	SERIAL	A unique integer (without decimals) used to identify Rockbuster's rental inventory.
film_id	FK	SMALLINT	A unique small integer used to identify each film at Rockbuster.
store_id	FK	SMALLINT	A unique small integer used to identify an individual store.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Film

Columns	Key	Data type	Description
film_id	PK	SERIAL	A unique small integer used to identify each film at Rockbuster.
title		CHARACTER VARYING(255)	This is a film title with a fixed length of 255 characters.
description		TEXT	A film's summary, no length limit.
release_year		year	The year a film was released.
language_id	FK	SMALLINT	A unique identifier for film language.
rental_duration		SMALLINT	The length of a video rental (in days).
rental_rate		NUMERIC(4,2)	Amount charged for rented videos with a maximum of 4 integers and 2 decimals.
length		SMALLINT	Length of film in minutes.
replacement_cost		NUMERIC(5,2)	Amount charged to customer to replace lost rented videos. The amount consists of 5 integers and 2 decimals.
rating		mpaa_rating	Film rating introduced by Movie Picture Association
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).
special_features		TEXT[[	Extra videos related to the film available for audience
fultext		TSVECTOR	Text-searchable data where we can use a word as an index

Language

Columns	Key	Data type	Description
language_id	PK	SERIAL	A unique identifier for film language.
name		CHARACTER(20)	The name of the language used in film.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Film actor

Columns	Key	Data type	Description
actor_id	PK	SMALLINT	A unique identifier for an individual actor.
film_id	PK	SMALLINT	A unique small integer used to identify each film at Rockbuster
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Actor

Columns	Key	Data type	Description

actor_id	PK	SERIAL	A unique identifier for an individual actor.
first_name		CHARACTER VARYING(45)	Actor's first name, character limit of 45.
last_name		CHARACTER VARYING(45)	Actor's last name, character limit of 45.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Film category

Columns	Key	Data type	Description
film_id	PK	SMALLINT	A unique small integer used to identify each film at Rockbuster
category_id	PK	SMALLINT	A unique identifier for film category in this database.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Category

Columns	Key	Data type	Description
category_id	PK	SERIAL	A unique identifier for film category in this database.
name		CHARACTER VARYING(25)	The name of the category with maximum of 25 characters.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Customer

Columns	Key	Data type	Description
customer_id	PK	SERIAL	A unique small integer used to identify a video rental customer.
store_id		SMALLINT	A unique small integer used to identify an individual store.
first_name		CHARACTER VARYING(45)	Customer's first name, max 45 characters
last_name		CHARACTER VARYING(45)	Customer's last name, max 45 characters
email		CHARACTER VARYING(50)	Customer's email. Max 50 characters
address_id		SMALLINT	Customer's address
activebool		BOOLEAN	An indication of whether a customer is active or not.
create_date		DATE	The date when the customer account was created with the format yy-mm-dd.

last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).
active		INTEGER	Integer, a number to indicate whether the customer account is active or not.

Staff

Columns	Key	Data type	Description
staff_id	PK	SERIAL	A unique small integer used to identify Rockbuster's individual staff.
first_name		CHARACTER VARYING(45)	Rockbuster employee's first name, max 45 characters
last_name		CHARACTER VARYING(45)	Rockbuster employee's last name, max 45 characters
address_id	FK	SMALLINT	Employee's address.
email		CHARACTER VARYING(50)	Employee's email with a character limit of 50.
store_id	FK	SMALLINT	A unique small integer used to identify an individual store.
active		BOOLEAN	An indication of whether a staff is active or not.
username		CHARACTER VARYING(16)	Rockbuster employee's username with a maximum of 16 characters.
password		CHARACTER VARYING(40)	Rockbuster's employee's password used to log in to internal account, maximum of 40 characters.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).
picture		BYTEA	Shows Rockbuster's employee's picture.

Store

Columns	Key	Data type	Description
store_id	PK	SERIAL	A unique small integer used to identify an individual store.
manager_staf f_id		SMALLINT	A unique identifier for a manager staff.
address_id		SMALLINT	A unique address for a store.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Address

Columns	Key	Data type	Description
adress_id		SERIAL	An identifier for a specific address.
address		CHARACTER VARYING(50)	Any address in this Rockbuster's database

		CHARACTER VARYING(50)	This is line 2 of the address.
address2		CHARACTER VARYING(50)	This is line 2 of the address.
district		CHARACTER VARYING(20)	Identifies a district as part of address.
city_id		SMALLINT	Unique identifier for city stated in an address.
postal_code		CHARACTER VARYING(10)	Captures a postal code in an address with a maximum 10 characters.
phone		CHARACTER VARYING(20)	Phone number associated with an address with a maximum of 20 characters.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

City

Columns	Key	Data type	Description
city_id		SERIAL	Unique identifier for city stated in an address.
city		CHARACTER VARYING(50)	City name in the address in this database.
country_id		SMALLINT	A unique identifier for country.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).

Country

Columns	Key	Data type	Description
country_id		SERIAL	A unique identifier for country.
country		CHARACTER VARYING(50)	Name of a country as captured in an address.
last_update		TIMESTAMP(6) WITHOUT TIME ZONE	The 6 digit notation for last record update date (dd-mm-yy) and time (hh:mm:ss).