DATA DICTIONARY

Farhan Ahmad

CareerFoundry | Data Analytics Immersion 3.10
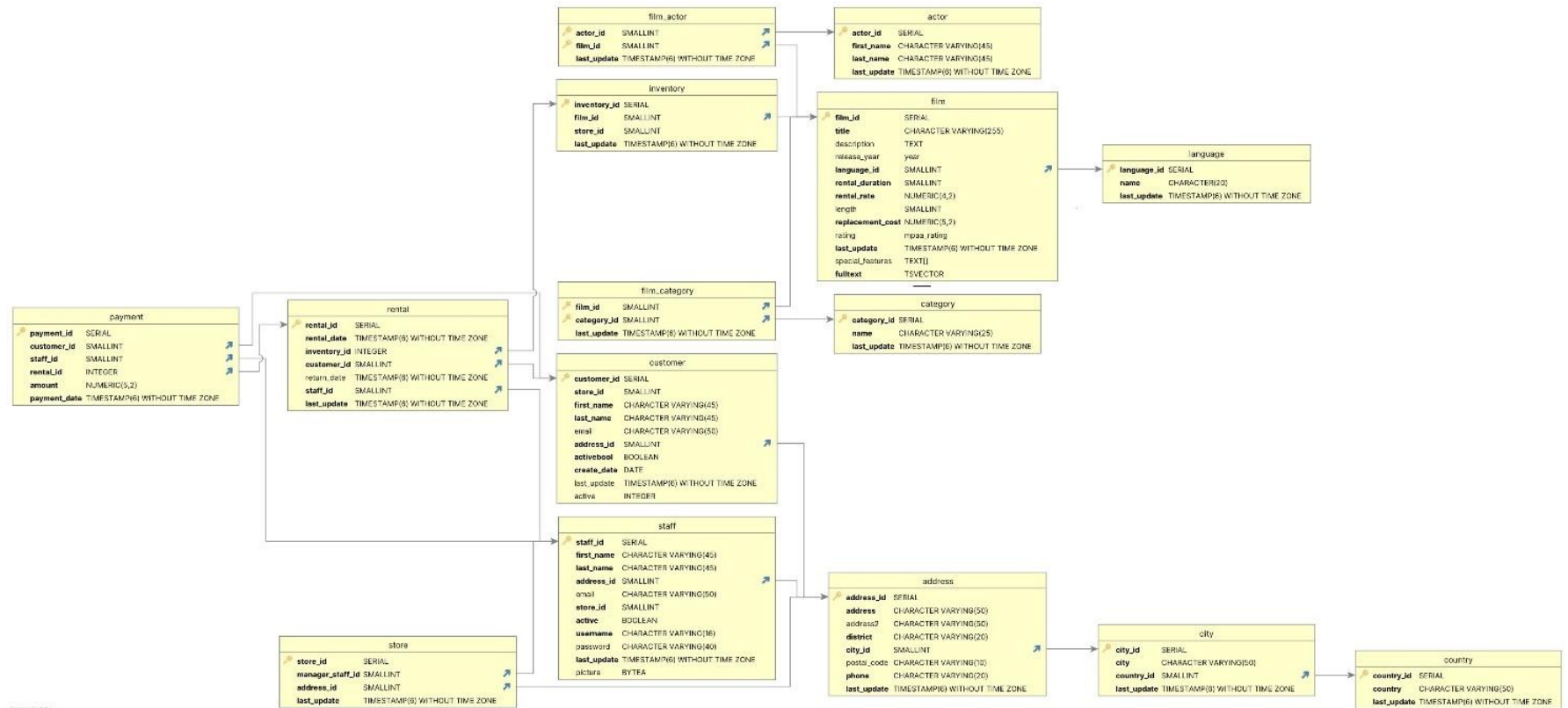
Rockbuster Stealth Data Dictionary

Contents

1. Introduction

This is a data dictionary for the Rockbuster Stealth PostgreSQL database.

The database contains 15 tables, including 2 Fact Tables and 13 Dimension Tables.

The database can be downloaded here: Rockbuster Database

2. Rockbuster Stealth Entity Relationship Diagram

film_actor

actor_id	SMALLINT
film_id	SMALLINT
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

actor

actor_id	SERIAL
first_name	CHARACTER VARYING(45)
last_name	CHARACTER VARYING(45)
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

inventory

inventory_id	SERIAL
film_id	SMALLINT
store_id	SMALLINT
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

film

film_id	SERIAL
title	CHARACTER VARYING(255)
description	TEXT
release_year	year
language_id	SMALLINT
rental_duration	SMALLINT
rental_rate	NUMERIC(4,2)
length	SMALLINT
replacement_cost	NUMERIC(5,2)
rating	mpaa_rating
last_update	TIMESTAMP(6) WITHOUT TIME ZONE
special_features	TEXT[]
fulltext	TSVECTOR

language

language_id	SERIAL
name	CHARACTER(20)
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

payment

payment_id	SERIAL
customer_id	SMALLINT
staff_id	SMALLINT
rental_id	INTEGER
amount	NUMERIC(5,2)
payment_date	TIMESTAMP(6) WITHOUT TIME ZONE

rental

rental_id	SERIAL
rental_date	TIMESTAMP(6) WITHOUT TIME ZONE
inventory_id	INTEGER
customer_id	SMALLINT
return_date	TIMESTAMP(6) WITHOUT TIME ZONE
staff_id	SMALLINT
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

film_category

film_id	SMALLINT
category_id	SMALLINT
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

category

category_id	SERIAL
name	CHARACTER VARYING(25)
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

customer

customer_id	SERIAL
store_id	SMALLINT
first_name	CHARACTER VARYING(45)
last_name	CHARACTER VARYING(45)
email	CHARACTER VARYING(50)
address_id	SMALLINT
activebool	BOOLEAN
create_date	DATE
last_update	TIMESTAMP(6) WITHOUT TIME ZONE
active	INTEGER

staff

staff_id	SERIAL
first_name	CHARACTER VARYING(45)
last_name	CHARACTER VARYING(45)
address_id	SMALLINT
email	CHARACTER VARYING(50)
store_id	SMALLINT
active	BOOLEAN
username	CHARACTER VARYING(16)
password	CHARACTER VARYING(40)
last_update	TIMESTAMP(6) WITHOUT TIME ZONE
picture	BYTEA

address

address_id	SERIAL
address	CHARACTER VARYING(50)
address2	CHARACTER VARYING(50)
district	CHARACTER VARYING(20)
city_id	SMALLINT
postal_code	CHARACTER VARYING(10)
phone	CHARACTER VARYING(20)
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

city

city_id	SERIAL
city	CHARACTER VARYING(50)
country_id	SMALLINT
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

country

country_id	SERIAL
country	CHARACTER VARYING(50)
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

store

store_id	SERIAL
manager_staff_id	SMALLINT
address_id	SMALLINT
last_update	TIMESTAMP(6) WITHOUT TIME ZONE

3. Legend

KEY	Columns
🔑	Primary Key
↗	Foreign Key

4. Fact Tables

4.1 payment

KEY	Columns	Data Type	Description
🔑	payment_id	SERIAL	Primary key; unique identifier associated with each payment
↗	customer_id	SMALLINT	Foreign key; unique identifier associated with each customer
↗	staff_id	SMALLINT	Foreign key; unique identifier associated with staff members
↗	rental_id	INTEGER	Foreign key; unique identifier associated with each rental
	amount	NUMERIC(5,2)	Amount of payment. 5 total digits, 2 decimals
	payment_date	TIMESTAMP(6) WITHOUT TIME ZONE	Time of payment (time zone information not included)

4.2 rental

KEY	Columns	Data Type	Description
🔑	rental_id	SERIAL	Primary key; unique identifier for rentals
	rental_date	TIMESTAMP (6) WITHOUT TIME ZONE	Date and time of rental (does not include time zone information)
↗	inventory_id	INTEGER	Foreign key; unique identifier for inventory
↗	customer_id	SMALLINT	Foreign key; unique identifier assigned to customers

	return_date	TIMESTAMP(6) WITHOUT TIME ZONE	6-digit return timestamp (does not include time zone information)
↗	staff_id	SMALLINT	Foreign Key; unique identifier assigned to staff
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last time the table was updated (time zone information not included)

5. Dimension Tables

5.1 store

Key	Columns	Data Type	Description
🔑	store_id	SERIAL	Primary key; unique identifier assigned to each store
↗	manager_staff_id	SMALLINT	Foreign key; unique identifier assigned to each manager
↗	address_id	SMALLINT	Foreign key; unique identifier assigned to each address
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone information not included)

5.2 film_actor

Key	Columns	Data Type	Description
🔑 ↗	actor_id	SMALLINT	Primary key; foreign key; unique identifier assigned to each actor
🔑 ↗	film_id	SMALLINT	Primary key; foreign key; unique identifier assigned to each film
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone information not included)

5.3 inventory

Key	Columns	Data Type	Description
🔑	inventory_id	SERIAL	Primary key; unique identifier assigned to each inventory item
↗	film_id	SMALLINT	Foreign key; unique identifier assigned to each film
	store_id	SMALLINT	Unique identifier assigned to each store
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time when table was last updated (time zone information not included)

5.4 film category

Key	Columns	Data Type	Description
🔑 ↗	film_id	SMALLINT	Primary key; foreign key; unique identifier assigned to each film
🔑 ↗	category_id	SMALLINT	Primary key; foreign key; unique identifier assigned to each category
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	6-digit timestamp when table was last updated (time zone information not included)

5.5 customer

Key	Columns	Data Type	Description
🔑	customer_id	SERIAL	Primary key; unique identifier assigned to each customer
	store_id	SMALLINT	Unique identifier assigned to each store
	first_name	CHARACTER VARYING(45)	Customer first name
	last_name	CHARACTER VARYING(45)	Customer last name
	email	CHARACTER VARYING(50)	Customer email
↗	address_id	SMALLINT	Foreign key; unique identifier assigned to each address

	activebool	BOOLEAN	Active logic function
	create_date	DATE	Date when account was created
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Last time table was updated (time zone information not included)
	active	INTEGER	Status of whether account is currently active or not

5.6 staff

Key	Columns	Data Type	Description
🔑	staff_id	SERIAL	Primary key; unique identifier assigned to each staff member
	first_name	CHARACTER VARYING(45)	First name of each staff member
	last_name	CHARACTER VARYING(45)	Last name of each staff member
↗	address_id	SMALLINT	Foreign key; unique identifier assigned to each address
	email	CHARACTER VARYING(50)	Email address of each staff member
	store_id	SMALLINT	Unique identifier assigned to the store where the staff member works
	active	BOOLEAN	Active boolean function
	username	CHARACTER VARYING(16)	Username of the staff member
	password	CHARACTER VARYING(40)	Password of the staff member
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time when the table was last updated (time zone information not included)
	picture	BYTEA	Photo of the staff member

5.7 actor

Key	Columns	Data Type	Description
🔑	actor_id	SERIAL	Primary key; unique identifier assigned to each actor
	first_name	CHARACTER VARYING(45)	First name of the actor
	last_name	CHARACTER VARYING(45)	Last name of the actor
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time table was last updated (time zone information not included)

5.8 film

Key	Columns	Data Type	Description
🔑	film_id	SERIAL	Primary key; unique identifier assigned to each film
	title	CHARACTER VARYING(255)	Title of the film
	description	TEXT	Description of the film
	release_year	YEAR	Year of the film's release
↗	language_id	SMALLINT	Foreign key; unique identifier assigned to each language
	rental_duration	SMALLINT	Length of the rental duration for the film in days
	rental_rate	NUMERIC(4,2)	Rental rate for the film. 4 digits, 2 for decimal (ex. 10.00)
	length	SMALLINT	Length of the film in minutes
	replacement_cost	NUMERIC(5,2)	Cost of replacing the film. 5-digit number, 2 decimal places (ex. 100.00)
	rating	MPAA_RATING	Film rating
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone information not included)
	special_features	TEXT[]	Special features associated with the film

Key	Columns	Data Type	Description
	fulltext	TSVECTOR	Vector of terms for full text search in a database.

5.9 category

Key	Columns	Data Type	Description
🔑	category_id	SERIAL	Primary key; unique identifier assigned to each movie category/genre
	name	CHARACTER VARYING(25)	Name of the category
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone information not included)

5.10 address

Key	Columns	Data Type	Description
🔑	address_id	SERIAL	Primary key; unique identifier assigned to each address
	address	CHARACTER VARYING(50)	Address information for a customer, staff member, or store
	address2	CHARACTER VARYING(50)	Additional address information for a customer, store, or staff member
	district	CHARACTER VARYING(20)	District information for a customer, store, or staff member
↗	city_id	SMALLINT	Foreign key; unique identifier assigned to each city in the database
	postal_code	CHARACTER VARYING(10)	Postal code for a customer, store, or staff member
	phone	CHARACTER VARYING(20)	Phone number for a customer, store, or staff member
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone

Key	Columns	Data Type	Description
			information not included)

5.11 language

Key	Columns	Data Type	Description
🔑	language_id	SERIAL	Primary key; unique identifier assigned to each language
	name	CHARACTER(20)	Name of the language
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone information not included)

5.12 city

Key	Columns	Data Type	Description
🔑	city_id	SERIAL	Primary key; unique identifier assigned to each city in the database
	city	CHARACTER VARYING(50)	City name
↗	country_id	SMALLINT	Foreign key; unique identifier assigned to each country in the database
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone information not included)

5.13 country

Key	Columns	Data Type	Description
🔑	country_id	SERIAL	Primary key; unique identifier assigned to each country in the database
	country	CHARACTER VARYING(50)	country name
	last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Time the table was last updated (time zone information not included)