ROCKBUSTER STEALTH DATA ANALYTICS PROJECT
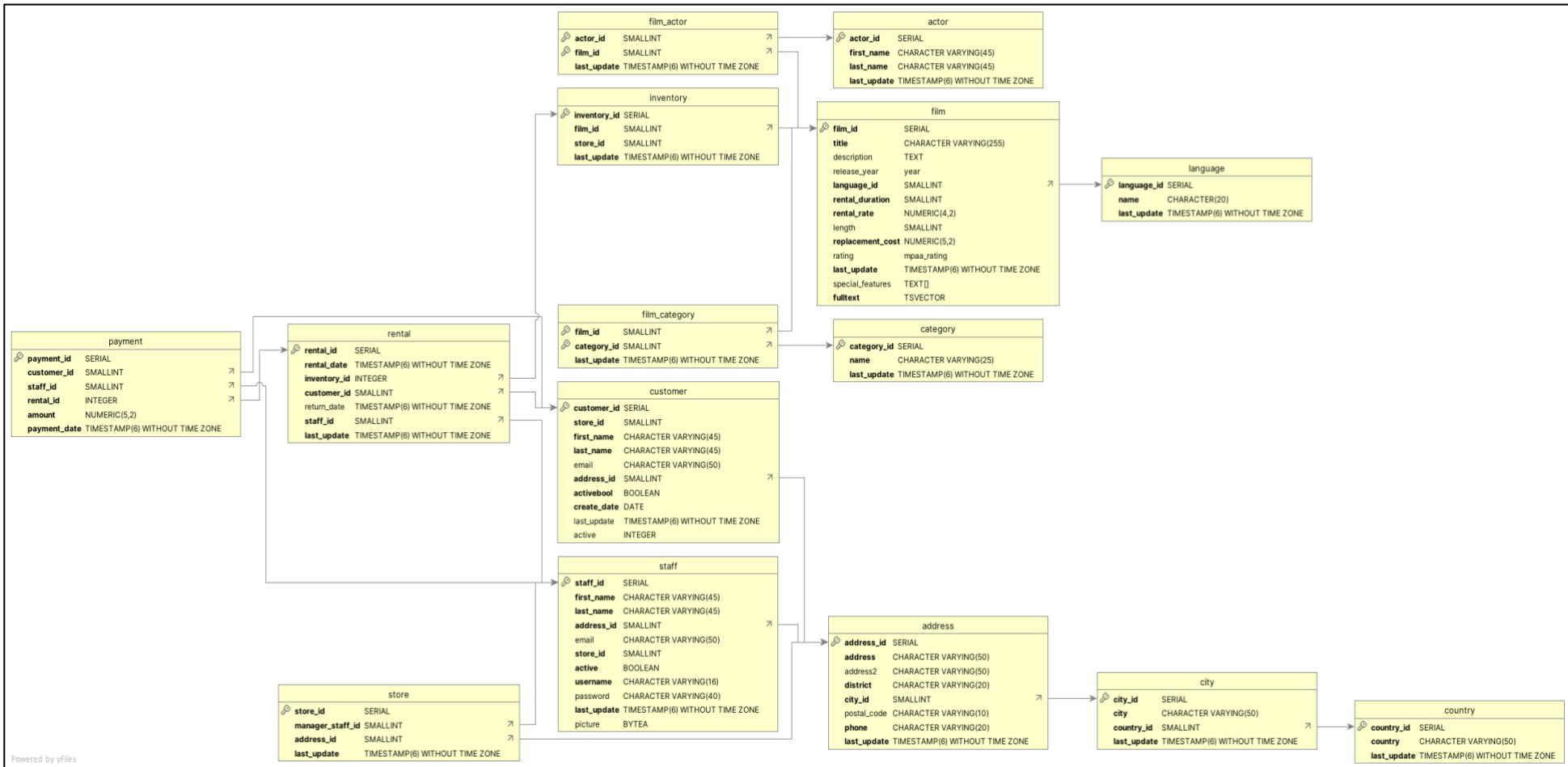
DATA DICTIONARY

By: Vijay Kasturi
Date: 21-Nov-2024

Table of Contents

1.Entity Relationship Model

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

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

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	

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	

store

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

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

2. Data Dictionary

The Rockbuster database has a Snowflake Schema. The schema consists of 1 or more Fact Tables connected to Dimension & Sub-dimension Tables.

2.1 FACT TABLES

2.1.1 Payment Table

Column	Data Type	Description
payment_id	SERIAL	Unique identifier of Payment transaction
customer_id	SMALLINT	Unique identifier of Customer
staff_id	SMALLINT	Unique identifier of Staff who handled the transaction
rental_id	INTEGER	Unique identifier of Rental transaction
amount	NUMERIC(5,2)	Total amount of the Transaction
payment_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of the Payment transaction

2.1.2 Rental Table

Column	Data Type	Description
rental_id	SERIAL	Unique identifier of Rental transaction
rental_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of the Rental transaction
inventory_id	INTEGER	Unique identifier of Inventory Unit
customer_id	SMALLINT	Unique identifier of Customer
return_date	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of the Return transaction
staff_id	SMALLINT	Unique identifier of Staff who handled the transaction
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2 DIMENSION TABLES

2.2.1 Store Table

Column	Data Type	Description
store_id	SERIAL	Unique identifier of Store
manager_staff_id	TIMESTAMP(6) WITHOUT TIME ZONE	Unique identifier of the Staff who manages the store
address_id	INTEGER	Unique identifier of the Address record
last_update	SMALLINT	Date & Time of Last update to record

2.2.2 Film Actor Table

Column	Data Type	Description
actor_id	SMALLINT	Unique identifier of the Actor
film_id	SMALLINT	Unique identifier of the Film
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2.3 Inventory Table

Column	Data Type	Description
inventory_id	SERIAL	Unique identifier of Inventory item
film_id	SMALLINT	Unique identifier of the Film
store_id	SMALLINT	Unique identifier of Store
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2.4 Film Category Table

Column	Data Type	Description
film_id	SMALLINT	Unique identifier of the Film
category_id	SMALLINT	Unique identifier of the Category
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2.5 Customer Table

Column	Data Type	Description
customer_id	SERIAL	Unique identifier of Customer
store_id	SMALLINT	Unique identifier of Store
first_name	CHARACTER VARYING(45)	First Name of Customer
last_name	CHARACTER VARYING(45)	Last Name of Customer
email	CHARACTER VARYING(50)	Email of Customer
address_id	SMALLINT	Unique identifier of the Address record
activebool	BOOLEAN	Flag to indicate if Customer is Active (True / False)
create_date	DATE	Date of record creation
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record
active	INTEGER	Integer equivalent of True / False for Customer Active flag

2.2.6 Staff Table

Column	Data Type	Description
staff_id	SERIAL	Unique identifier of the Staff
first_name	CHARACTER VARYING(45)	First Name of Staff
last_name	CHARACTER VARYING(45)	Last Name of Staff
address_id	SMALLINT	Unique identifier of the Address record
email	CHARACTER VARYING(50)	Email of Staff

store_id	SMALLINT	Unique identifier of Store
active	BOOLEAN	Active Flag
username	CHARACTER VARYING(16)	Username of Staff
password	CHARACTER VARYING(40)	Password of Staff
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record
picture	BYTEA	Picture of Staff

2.2.7 Actor Table

Column	Data Type	Description
actor_id	SERIAL	Unique identifier of the 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 & Time of Last update to record

2.2.8 Film Table

Column	Data Type	Description
film_id	SERIAL	Unique identifier of the Film
title	CHARACTER VARYING(255)	Title of Film
description	TEXT	Description of Film
release_year	year	Year of Release
language_id	SMALLINT	Unique identifier of Language
rental_duration	SMALLINT	Duration of Rental
rental_rate	NUMERIC(4,2)	Rental Rate
length	SMALLINT	Length of Film
replacement_cost	NUMERIC(5,2)	Cost of Replacement of Film
rating	mpaa_rating	Age Rating of Film
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record
special_features	TEXT[]	Special Features of Film
fulltext	TSVECTOR	Text of Film

2.2.9 Category Table

Column	Data Type	Description
category_id	SERIAL	Unique identifier of the Category
name	CHARACTER VARYING(25)	Name of Category
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2.10 Language Table

Column	Data Type	Description
language_id	SERIAL	Unique identifier of Language
name	CHARACTER(20)	Name of Language
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2.11 Address Table

Column	Data Type	Description
address_id	SERIAL	Unique identifier of the Address record
address	CHARACTER VARYING(50)	Address Line 1
address2	CHARACTER VARYING(50)	Address Line 2
district	CHARACTER VARYING(20)	Name of District
city_id	SMALLINT	Unique identifier of City
postal_code	CHARACTER VARYING(10)	Postal Code
phone	CHARACTER VARYING(20)	Phone Number
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2.12 City Table

Column	Data Type	Description
city_id	SERIAL	Unique identifier of City
city	CHARACTER VARYING(50)	Name of City
country_id	SMALLINT	Unique identifier of Country
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

2.2.13 Country Table

Column	Data Type	Description
country_id	SERIAL	Unique identifier of Country
country	CHARACTER VARYING(50)	Name of Country
last_update	TIMESTAMP(6) WITHOUT TIME ZONE	Date & Time of Last update to record

------ End of Document ------