SARA AKBAR
EBRAHIM - 22967

Table of Contents

Business Problem 3
Business Entities 4
ERD 6
Pipeline 7

Business Problem

Letterboxd is a social networking platform for film enthusiasts. It allows users to track the films they've watched, create, and share reviews, explore recommendations, and connect with other film lovers.

Target Users: Movie buffs, cinephiles, casual film viewers interested in sharing their opinions and exploring new films.

Core functionalities:

- User accounts and profiles
- Film database with detailed information (title, synopsis, cast, crew)
- Ability to track watched films and create watchlists.
- Writing and sharing reviews with ratings
- Following other users and discovering new content through recommendations and user activity feeds.
- Discussions and community features (e.g., commenting on reviews, creating lists).

Value proposition for users:

- Organize and track their film experiences.
- Express their opinions and engage with other film enthusiasts.
- Discover new movies based on recommendations and reviews.
- Connect with like-minded people who share their passion for film.

Value proposition for Letterboxd:

- Subscription model offering premium features.
- Advertising and targeted marketing to film studios and related businesses.
- Data insights and analytics (anonymized data) about user preferences and film trends.
- Partnerships with film festivals or distributors for promotional opportunities.

Business Entities

The database consists of the following core entities:

Member

Represents a registered user of the system with the following attributes:

User ID (Primary Key), Username (Unique), Name, Email (Unique), Avatar URL, Date of Birth, Bio, Gender, Country, State (Optional), Joined Date, Subscription Status.

Film

Represents a movie or film within the system with the following attributes:

Film ID (Primary Key), Name, Poster URL, Overview, Original Language, Director, Runtime, Studio ID, Country, Release Year, Rating, Top 250 Position.

Studio

Represents a film production studio with the following attributes:

Studio ID (Primary Key), Name, Location, Founded Year, Website URL.

Genre

Represents a film genre category with the following attributes:

Genre ID (Primary Key), Genre Name (Unique).

Review

Represents a user-written review for a film with the following attributes:

Review ID (Primary Key), User ID (Foreign Key), Film ID (Foreign Key), Rating, Rewatch Flag, Review Text, Spoiler Flag, Watched Date, Date Created, Like Count.

Watchlist

Represents a user's list of films they intend to watch with the following attributes:

Watchlist ID (Primary Key), User ID (Foreign Key), Film ID (Foreign Key), Date Added.

List

Represents a user-created list of films (e.g., favorites, to-watch later with the following attributes:

List ID (Primary Key), User ID (Foreign Key), List Name, Description, Film Count, Date Created, Published Date (Optional), Ranked Flag (Optional).

ListFilm

Connects Films to Lists (Many-to-Many relationship with the following attributes:

List ID (Foreign Key), Film ID (Foreign Key) (Primary Key combination), Position (Order within the list), Description (Optional).

Follow

Represents a user following another user with the following attributes:

Follower ID (Foreign Key), Following ID (Foreign Key) (Primary Key combination).

LogEntry:

Represents a user's diary entry about a film with the following attributes:

Diary Entry ID (Primary Key), User ID (Foreign Key), Film ID (Foreign Key), Review_id, Date, Watched Date, Rating, Rewatch Flag, Liked Flag.

FilmGenre

Connects Films to Genres (Many-to-Many relationship with the following attributes:

Film ID (Foreign Key), Genre ID (Foreign Key) (Primary Key combination).

SubscriptionPlan

Represents a subscription plan offered to users with the following attributes:

Plan ID (Primary Key), Plan Name, Description, Price, Benefits.

Subscription

Represents a user's subscription to a plan with the following attributes:

Subscription ID (Primary Key), User ID (Foreign Key), Plan ID (Foreign Key).

Person

Represents people who work on/in Films

Person_id(Primary Key), name, bio, image_url

CastCrew

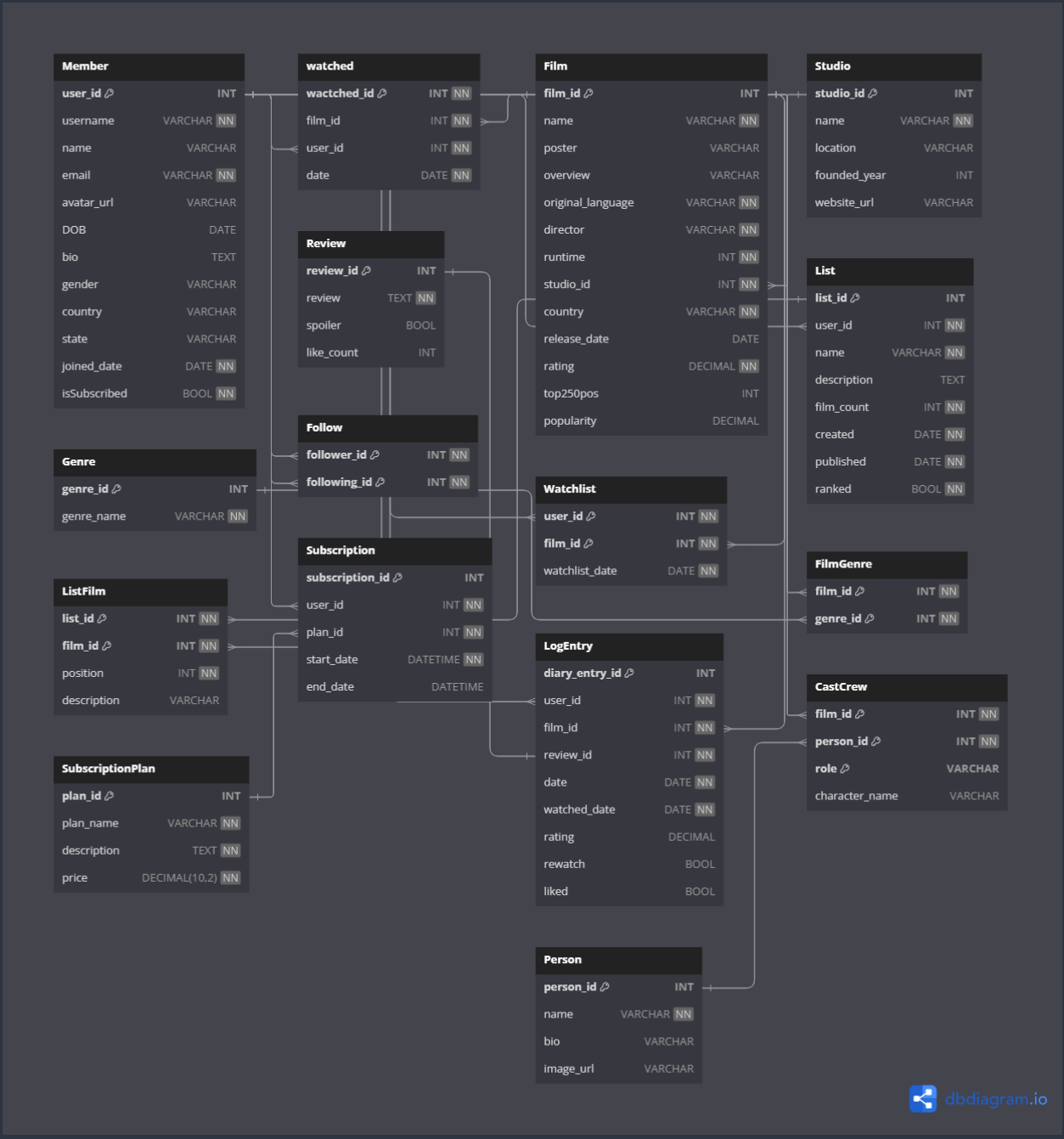
Represents cast and crew of film

Film_id, Person_id, role (Combined PK), character_name

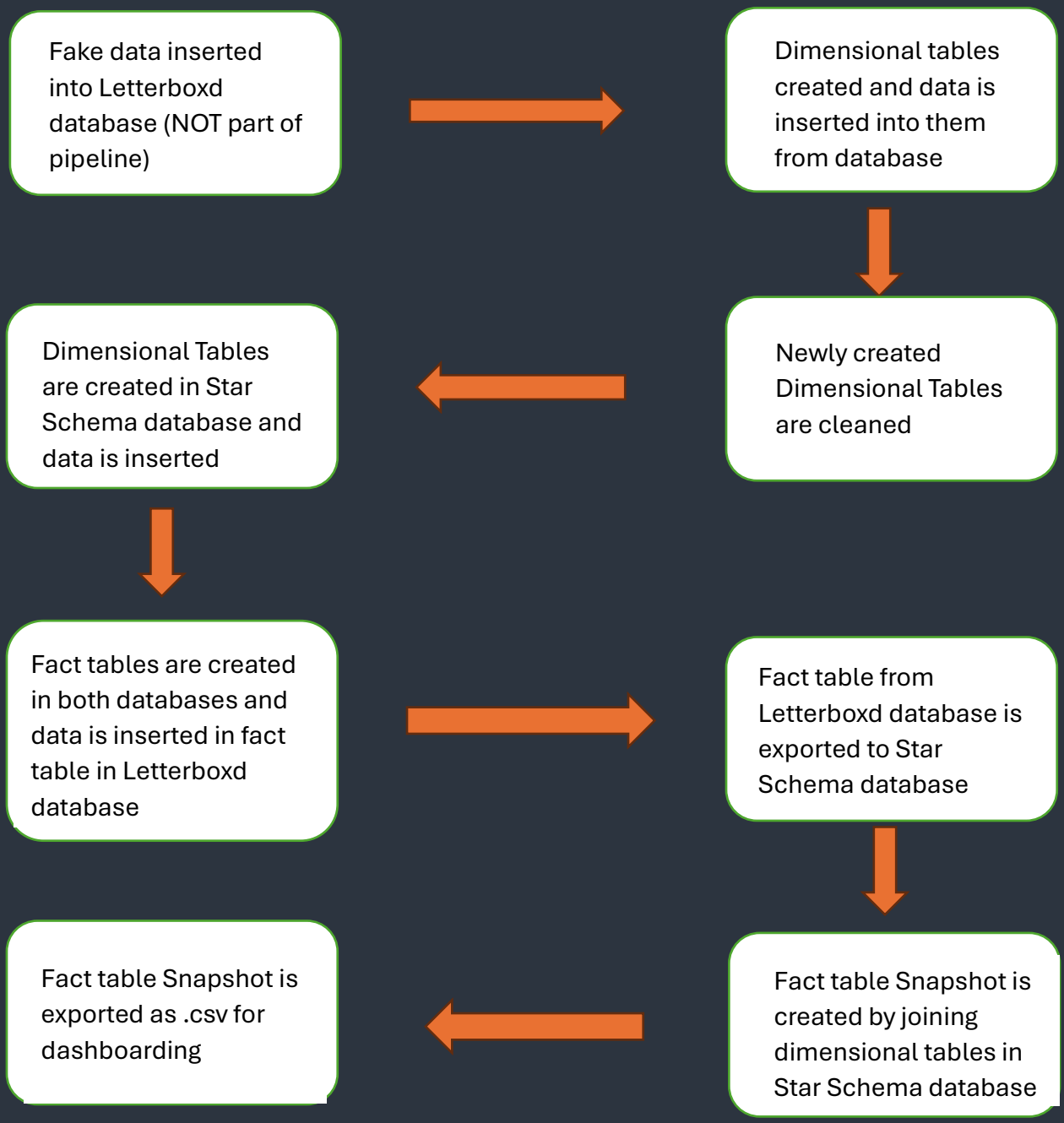
Watched

Record of user watching a film: Watched_id, film_id, user_id, date

ERD


Member
user_id INT
username VARCHAR NN
name VARCHAR
email VARCHAR NN
avatar_url VARCHAR
DOB DATE
bio TEXT
gender VARCHAR
country VARCHAR
state VARCHAR
joined_date DATE NN
isSubscribed BOOL NN
watched
wactched_id INT NN
film_id INT NN
user_id INT NN
date DATE NN
Review
review_id INT
review TEXT NN
spoiler BOOL
like_count INT
Follow
follower_id INT NN
following_id INT NN
Genre
genre_id INT
genre_name VARCHAR NN
Subscription
subscription_id INT
user_id INT NN
plan_id INT NN
start_date DATETIME NN
end_date DATETIME
ListFilm
list_id INT NN
film_id INT NN
position INT NN
description VARCHAR
SubscriptionPlan
plan_id INT
plan_name VARCHAR NN
description TEXT NN
price DECIMAL(10,2) NN
Film
film_id INT
name VARCHAR NN
poster VARCHAR
overview VARCHAR
original_language VARCHAR NN
director VARCHAR NN
runtime INT NN
studio_id INT NN
country VARCHAR NN
release_date DATE
rating DECIMAL NN
top250pos INT
popularity DECIMAL
Watchlist
user_id INT NN
film_id INT NN
watchlist_date DATE NN
LogEntry
diary_entry_id INT
user_id INT NN
film_id INT NN
review_id INT NN
date DATE NN
watched_date DATE NN
rating DECIMAL
rewatch BOOL
liked BOOL
Person
person_id INT
name VARCHAR NN
bio VARCHAR
image_url VARCHAR
Studio
studio_id INT
name VARCHAR NN
location VARCHAR
founded_year INT
website_url VARCHAR
List
list_id INT
user_id INT NN
name VARCHAR NN
description TEXT
film_count INT NN
created DATE NN
published DATE NN
ranked BOOL NN
FilmGenre
film_id INT NN
genre_id INT NN
CastCrew
film_id INT NN
person_id INT NN
role VARCHAR
character_name VARCHAR
dbdiagram.io

Pipeline



NOTE: Dimensional Queries are not part of pipeline (shown separately) before fact table creation