



Présentation type rappel des données

1. Présentation et Analyse des données fournis :

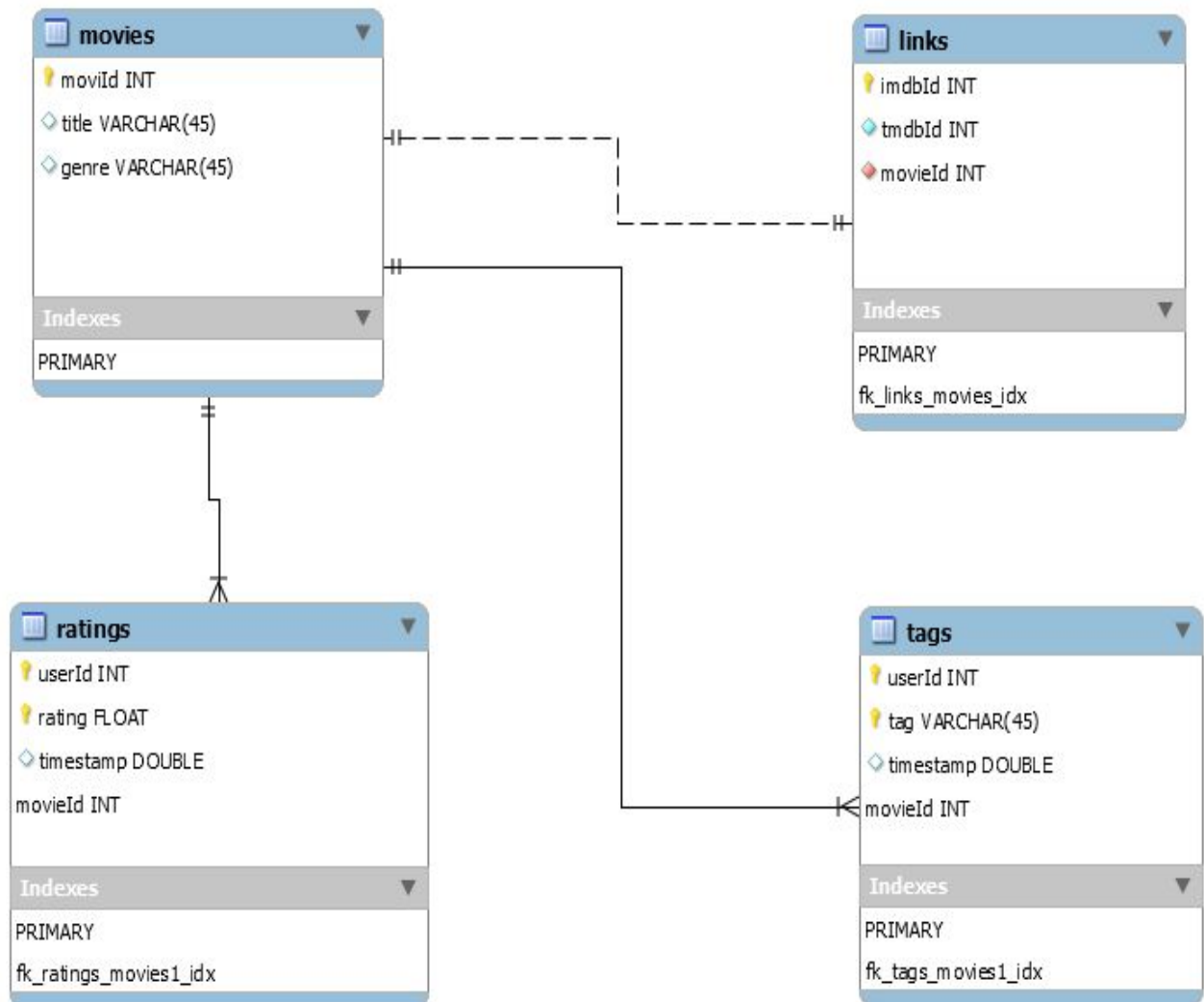
Nous avons reçu de la part de notre client quatres fichiers de type csv notamment **Movies.csv**, **Tags.csv**, **Links.cs**, **ratings.csv**, le tableau ci dessous décrit les fichiers fournies

Fichier	Attributs	Description
Movies	movielfd, title, genre	Contient un listing de film, classé par le titre et le genre.
Tags	userId, movielfd, tag, timestamp	Contient les tags des utilisateurs identifiés pour des films en fonction de l'heure.
Links	movielfd, imdbId, tmbdId	contient des liens numériques servant à compléter des liens url pour trouver des films sur des sites spécialisés.
ratings	userId, movielfd, ratings, timestamp	contient les votes données sur des films par certains utilisateurs identifiés, ainsi que la date et l'heure auxquelles l'évaluation a eu lieu.



Présentation type rappel des données

2. Schéma relationnel de la base movies



3. Présentation améliorée des données :

- Fusionner le tableau links avec le tableau movie.ID
- Pour le tableau tags, on a plusieurs tags pour un film et on le même tag pour différents films. Il faudrait une nouvelle table pour regrouper les différentes clés primaires.
- Formater le titre movies en séparant le titre de l'année.