

## Rendu de la partie modélisation

Noms : Desbles Loann & Billard Ethan  
Groupe 4

### 1.Schéma Entité/Association Finalisé

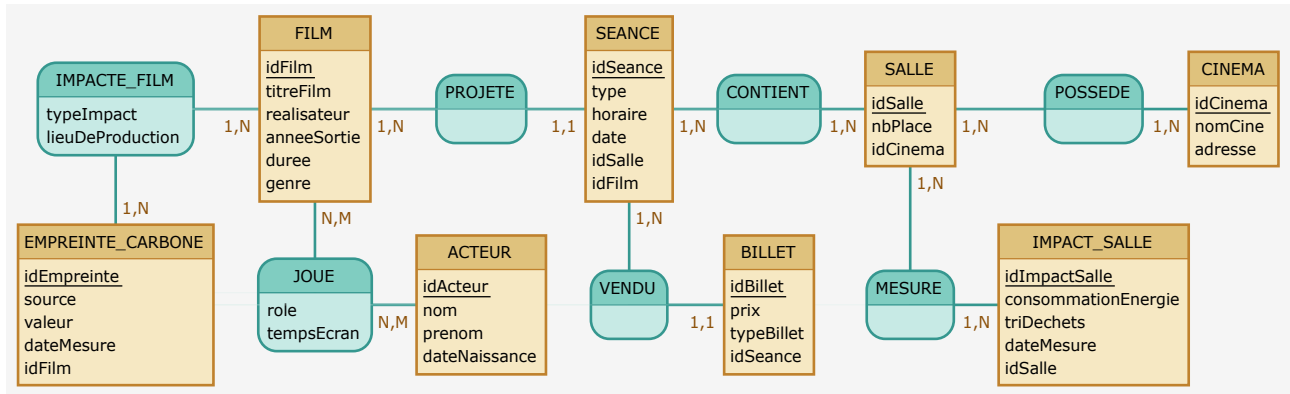


Schéma E/A respectant la 3FN

Table	Clés Primaires & Clés Étrangères
ACTEUR	id_Acteur (PK)
BILLET	id_Billet (PK), id_Seance (FK → SEANCE)
CINEMA	id_Cinema (PK)
EMPREINTE_CARBONE	id_Empreinte (PK), id_Film (FK → FILM)
FILM	id_Film (PK)
IMPACTE_FILM	( id_Film , id_Empreinte ) (PK, FK → FILM, FK → EMPREINTE_CARBONE)
IMPACT_SALLE	id_Impact_Salle (PK), id_Salle (FK → SALLE)
JOUE	( id_Film , id_Acteur ) (PK, FK → FILM, FK → ACTEUR)
SALLE	id_Salle (PK), id_Cinema (FK → CINEMA)
SEANCE	id_Seance (PK), id_Film (FK → FILM), id_Salle (FK → SALLE)

FK: Clé étrangère, PK: Clé primaire

### 2.Guide de l'utilisateur :

1. Création des tables avec CreateTable.sql qui se trouve dans le dossier SQL
2. Implémentation des tables avec les fichiers csv respectifs à chaque table dans le dossier CSV, puis aller dans le README.txt qui est lié à un git, le séparateur est une virgule.
3. Les requêtes, elles, se trouvent dans Requetes.sql dans le dossier SQL

### 3.Explication des requêtes

1. Calculer la somme des recettes générées par les billets vendus par séance
2. Liste les acteurs ayant joué dans un seul film.
3. Calcule la somme des recettes par séance.
4. Compte le nombre de salles pratiquant le tri des déchets.
5. Liste les cinémas possédant au moins trois salles.
6. Calcule la consommation énergétique totale par cinéma.
7. Calcule la moyenne de l'empreinte carbone.
8. Affiche le film avec l'empreinte carbone la plus élevée.

Ces requêtes s'avèrent être utiles pour identifier l'impact écologique causé par le cinéma en tant qu'industrie et les cinémas en tant qu'infrastructure, ces données peuvent permettre dans le futur de mettre en place des stratégies d'atténuation de pollution.

### 4.Difficultés rencontrés

#### 1.Partie SQL :

- Problème dû à la longueur de certains varchar rencontrés dans la base de données que nous avons extraites
- Erreur lors de l'insertion des données dû aux contraintes des clés étrangères ce qui nous a amené à modifier nos scripts python
- Soucis avec des dates générées aléatoirement avec des dates du 30 février qui ne sont pas sensé exister

#### 2.Partie extraction :

- Trouver des DataSet complet
- Lecture des données dans différent format : json dans cellule csv en format string
- Problèmes d'occurrence morte ou mal formatée dans les DataSet faisant planter l'extraction
- Volume de données important, la complexité des algo de traitement doit être pris en compte (utilisation d'arbre binaire)
- Maintenir la base de données, les occurrence mortes étant supprimées il est nécessaire de croiser les tables pour retirer les occurrences présentes que dans une table générant des erreurs d'import dans PgAdmin
- Génération cohérente de données afin de compléter les tables