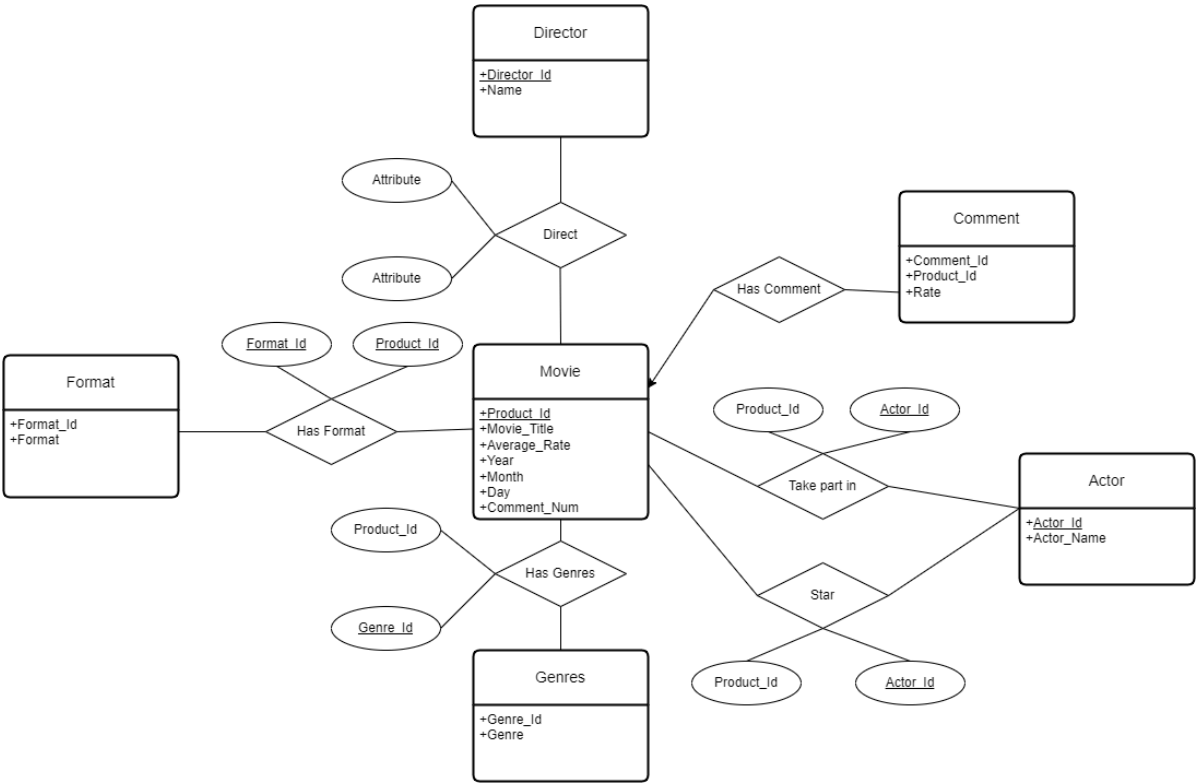
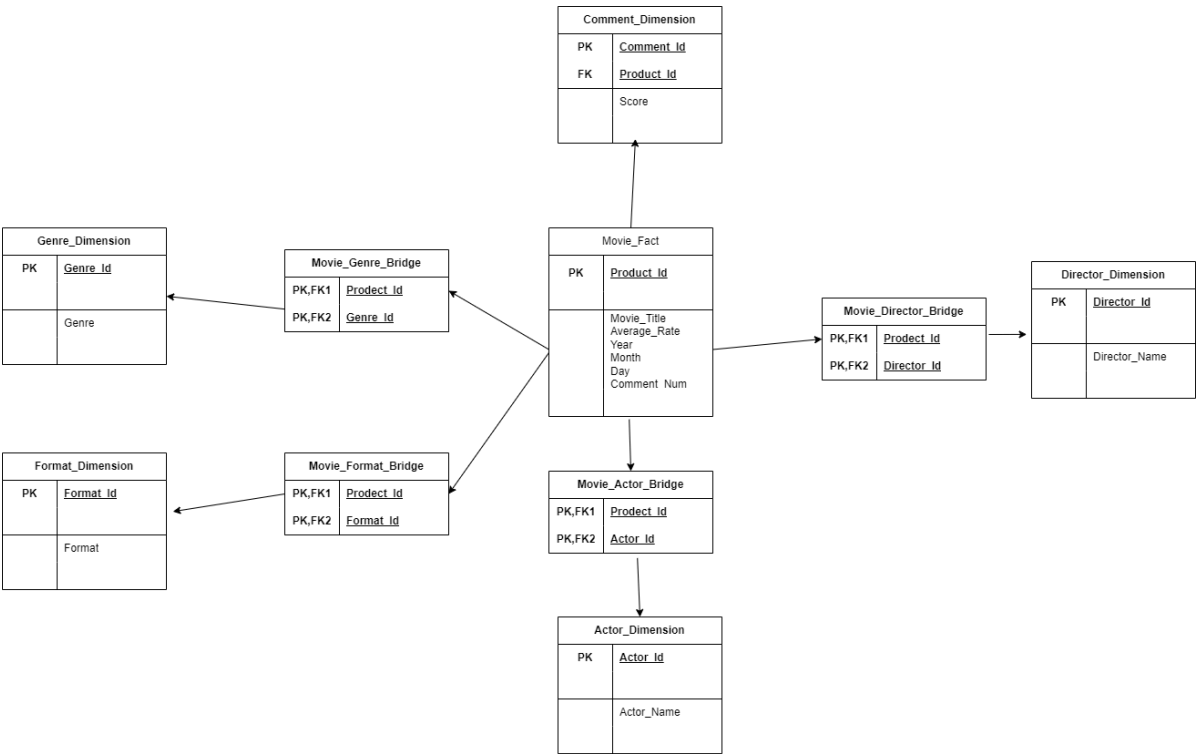


本项目的三个数据库都围绕以下ER图建表



三个数据库都存储了所有的实体和联系的信息。其中，MySQL还作为Hive的Metadata数据库，用来存储hive中所有的分区、列等相关信息。

Star Schema



1. 核心表：Fact 表

Movie_Fact 是事实表，存储了关于电影的核心度和相关信息：

- 主键 (PK)
 - **Product_Id**：每部电影的唯一标识。
- 度量字段
 - **Average_Rate**：电影的平均评分。
 - **Year**、**Month**、**Day**：电影的日期字段，用于时间维度分析。
 - **Comment_Num**：评论数量，表示每部电影的用户交互情况。
- 描述字段
 - **Movie_Title**：电影标题，用于标识和展示。

2. 维度表：Dimension 表

维度表用于描述事实表中的业务实体，支持多维度分析。

(1) Genre_Dimension:

- 存储电影的类型信息。
- 字段：
 - **Genre_Id** (PK)：类型唯一标识。
 - **Genre**：类型名称，如“动作”、“喜剧”。

(2) Format_Dimension:

- 描述电影的格式信息。
- 字段：
 - **Format_Id** (PK)：格式唯一标识。
 - **Format**：格式名称，如“Blue-ray”、“DVD”。

(3) Director_Dimension:

- 描述电影导演的信息。
- 字段：
 - **Director_Id** (PK)：导演唯一标识。
 - **Director_Name**：导演的名字。

(4) Actor_Dimension:

- 描述电影演员的信息。
- 字段：
 - Actor_Id (PK) : 演员唯一标识。
 - Actor_Name : 演员名字。

(5) Comment_Dimension:

- 描述与电影相关的用户评论信息。
- 字段：
 - Comment_Id (PK) : 评论唯一标识。
 - Product_Id (FK) : 对应电影的标识。
 - Score : 评论评分。

3. 桥表: Bridge 表

桥表用于连接事实表和维度表，支持多对多关系。

(1) Movie_Genre_Bridge:

- 描述每部电影和其对应的类型之间的关系。
- 主键和外键：
 - Product_Id (PK, FK1) : 对应电影。
 - Genre_Id (PK, FK2) : 对应类型。

(2) Movie_Format_Bridge:

- 描述每部电影和其对应的格式之间的关系。
- 主键和外键：
 - Product_Id (PK, FK1) : 对应电影。
 - Format_Id (PK, FK2) : 对应格式。

(3) Movie_Actor_Bridge:

- 描述每部电影和其对应演员之间的关系。
- 主键和外键:

- `Product_Id` (PK, FK1) : 对应电影。
- `Actor_Id` (PK, FK2) : 对应演员。

(4) Movie_Director_Bridge:

- 描述每部电影和其对应导演之间的关系。
- 主键和外键：
 - `Product_Id` (PK, FK1) : 对应电影。
 - `Director_Id` (PK, FK2) : 对应导演。