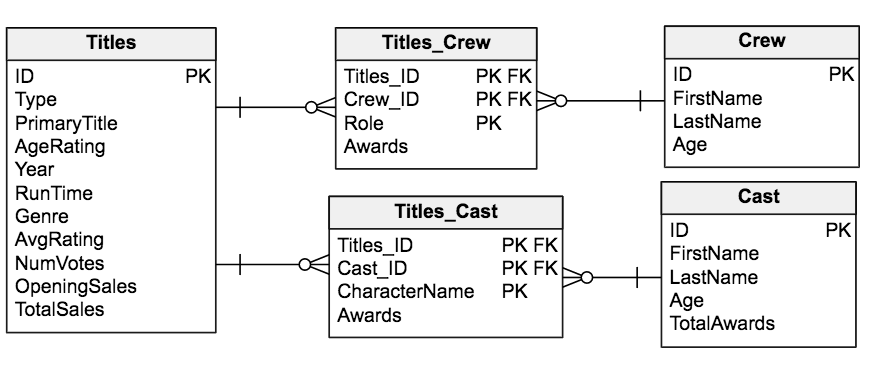
**SQL Analysis – Movie Database**

Consider the following schema for a movie database:



*Reference: loosely based on http://www.imdb.com/interfaces/*

**Data dictionary:**

**Titles**

* ID (INT) - unique identifier of the movie title
* Type (VARCHAR) – the type/format of the title (e.g. movie, short, etc)
* PrimaryTitle (VARCHAR) – the title used by the filmmakers on promotional materials at the point of release
* AgeRating (VARCHAR) – The rating of the movie (e.g. G, PG, PG-13, etc.)
* Year (INT) – year of the release of the title
* RunTime (INT) – runtime of the title, in minutes
* Genre (VARCHAR) – primary genre associated with the title (e.g. Comedy, Drama, etc.)
* AvgRating (DECIMAL) – Star score out of five stars
* NumVotes (INT) – Number of votes used in calculation of AvgRating
* OpeningSales (BIGINT) – Revenue generated during the first week of a Title's release, in dollars
* TotalSales (BIGINT) – Total revenue generated from the Title's release, in dollars

**Cast**

* ID (INT) – unique identifier of the cast member
* FirstName (VARCHAR) – first name of the cast member
* LastName (VARCHAR) – last name of the cast member
* Age (INT) – current age in years of the cast member
* TotalAwards (INT) – total number of awards awarded to the cast member

**Crew**

* ID (INT) – unique identifier of the crew member
* FirstName (VARCHAR) – first name of the crew member
* LastName (VARCHAR) – last name of the crew member
* Age (INT) – current age in years of the crew member

**Titles\_Cast**

* Titles\_ID (INT) - unique identifier of the movie title
* Cast\_ID (INT) – unique identifier of the crew member
* CharacterName (VARCHAR) – name of character played by the cast member in the title
* Awards (INT) – number of awards received by that cast member for playing that character in that title

**Titles\_Crew**

* Titles\_ID (INT) - unique identifier of the movie title
* Crew\_ID (INT) – unique identifier of the crew member
* Role (VARCHAR) – role of the crew member in the title (e.g. producer, writer, etc.)
* Awards (INT) – number of awards received by that crew member for playing that role in that title

**Write SQL queries to answer the following questions.**

***Question 10:***

Return the average of the AvgRating and the count of all of the titles in the database by genre.

SELECT Genre, AVG(AvgRating), COUNT(\*)

FROM Titles

GROUP BY Genre;

***Question 11:***

Look up the ID number for the crew member named Steven Spielberg. Assume there is only one crew member named Steven Spielberg.

SELECT ID

FROM Crew

WHERE FirstName='Steven' AND LastName='Spielberg';

***Question 12:***

Assuming Steven Spielberg's crew ID number is 1126, return the number of times he has been credited with playing each type of Role (and rename this column to RoleCount), along with the name of the Role.

SELECT Role, COUNT(\*) RoleCount

FROM Titles\_Crew

WHERE Crew\_ID=1126

GROUP BY Role;

***Question 13:***

For a given title, Steven Spielberg may play multiple crew roles. Assume that if he receives an award for any role associated with a given title, he considers himself to be "award-winning." Of all of the titles for which Steven has been a crew member, in what percent of these titles is he "award-winning." If you use more than one query, explain how to combine the results.

To find how many award winning titles Steven Spielberg has participated in:

SELECT COUNT(\*) FROM (

SELECT Titles\_ID

FROM Titles\_Crew

WHERE Awards>0 AND Crew\_ID=1126

GROUP BY Titles\_ID

);

Total number of titles Steven Speilberg has participated in:

SELECT COUNT(\*) FROM (

SELECT Titles\_ID

FROM Titles\_Crew

WHERE Crew\_ID=1126

GROUP BY Titles\_ID

);

then divide award winning titles by total titles to get percent

***Question 14:***

Return movie titles, along with the total number of cast awards associated with each title. Be sure to order the results from most number of awards to least number of awards. If you use more than one query, explain how to combine the results.

SELECT Titles.PrimaryTitle, SUM(Titles\_Crew.Awards)

FROM Titles

JOIN Titles\_Crew

ON Titles.ID=Titles\_Crew.Titles\_ID

GROUP BY Titles.PrimaryTitle;

***Question 15:***

Meryl Streep has won many awards and made a lot money for the movie industry in her lifetime. Return the following summary metrics for all titles where she was a cast member by year: Average TotalSales, Total TotalSales, Average OpeningSales, Total OpeningSales, Total Awards – be sure to calculate the total awards only for Meryl Streep using the Titles\_Cast table. If you use more than one query, explain how to combine the results.

SELECT

Titles.PrimaryTitle, AVG(Titles.TotalSales),

SUM(Titles.TotalSales), AVG(Titles.OpeningSales),

SUM(Titles.OpeningSales), SUM(Titles\_Crew.Awards)

FROM Cast

JOIN Titles\_Cast ON Titles\_Cast.Cast\_ID=Cast.ID

JOIN Titles ON Titles.ID=Titles\_Cast.Titles\_ID

WHERE Cast.FirstName='Meryl' AND Cast.LastName='Streep'

GROUP BY Titles.PrimaryTitle

***Question 16:***

Tim Burton and Johnny Depp work together a lot. Return all of the films where Johnny Depp was a cast member and Tim Burton was a crew member. Sort these by AvgRating from highest to lowest. If you use more than one query, explain how to combine the results.

SELECT

Titles.PrimaryTitle, AVG(Titles.AvgRating) as Rating

FROM Crew

JOIN Titles\_Crew ON Titles\_Crew.Crew\_ID=Crew.ID

JOIN Titles ON Titles.ID=Titles\_Crew.Titles\_ID

JOIN Titles\_Cast ON Titles\_Cast.Titles\_ID=Titles.ID

JOIN Cast ON Cast.ID=Titles\_Cast.Cast\_ID

WHERE Cast.FirstName='Johnny' AND Cast.LastName='Depp'

AND Crew.FirstName='Tim' AND Crew.LastName='Burton'

GROUP BY Titles.PrimaryTitle

ORDER BY Rating DESC