**Trend Analysis:**

**1. Average movie runtime change over the years:**

**a. For every year what is the average run-time and plot it on the graph.**

SELECT

    [StartYear] AS Year,

    AVG([runtimeMinutes]) AS AverageRuntime

FROM [dbo].[TitleSCDDim]

WHERE [StartYear] IS NOT NULL AND [runtimeMinutes] IS NOT NULL

GROUP BY [StartYear]

ORDER BY [StartYear];

**2. Is there a correlation between the average movie rating and the year of release:**

**a. The goal is to identify in which year we had good movies.**

SELECT

    t.[StartYear] AS Year,

    AVG(r.[averageRating]) AS AverageRating

FROM [dbo].[RatingFact] r

JOIN [dbo].[TitleSCDDim] t ON r.[Title\_FK] = t.[Title\_SK]

WHERE t.[StartYear] IS NOT NULL

GROUP BY t.[StartYear]

ORDER BY t.[StartYear];

**3. Visualization of the distribution of movie releases per year**

**a. No. of movies released every year**

SELECT

    [StartYear] AS Year,

    COUNT(\*) AS TotalMovies

FROM [dbo].[TitleSCDDim]

WHERE [StartYear] IS NOT NULL

GROUP BY [StartYear]

ORDER BY [StartYear];

**Genre Analysis:**

**1. Which genres have seen the most significant ratings in popularity over the past decade?**

**a. The goal is to identify the popularity of the genres.**

SELECT

    g.[Genres],

    AVG(r.[averageRating]) AS OverallAverageRating

FROM [dbo].[MovieGenreBri] mg

JOIN [dbo].[GenreDim] g ON mg.[Genre\_FK] = g.[Genre\_SK]

JOIN [dbo].[TitleSCDDim] t ON mg.[Title\_FK] = t.[Title\_SK]

JOIN [dbo].[RatingFact] r ON t.[Title\_SK] = r.[Title\_FK]

WHERE t.[StartYear] >= YEAR(GETDATE()) - 10

GROUP BY g.[Genres]

ORDER BY OverallAverageRating ASC;

**2. Show the top 5 genres as compared to gross earnings for the 9 box office movies.**

SELECT TOP 5

    g.[Genres],

    SUM(CAST(f.[Gross] AS DECIMAL(18,2))) AS TotalGross

FROM [dbo].[MovieGenreBri] mg

JOIN [dbo].[GenreDim] g ON mg.[Genre\_FK] = g.[Genre\_SK]

JOIN [dbo].[BoxOfcFact] f ON mg.[Title\_FK] = f.[Title\_SK]

GROUP BY g.[Genres]

ORDER BY TotalGross DESC;

**Performance Metrics:**

**1. Correlation between the movie's runtime and its average rating**

**a. The goal is to understand the impact on the ratings due to the movie's runtime.**

SELECT

    t.[runtimeMinutes],

    AVG(r.[averageRating]) AS AverageRating

FROM [dbo].[RatingFact] r

JOIN [dbo].[TitleSCDDim] t ON r.[Title\_FK] = t.[Title\_SK]

WHERE t.[runtimeMinutes] IS NOT NULL

GROUP BY t.[runtimeMinutes];

**2. Correlation between the movie's runtime and its average gross**

**a. The goal is to understand the impact on the gross earnings due to the movie's runtime.**

SELECT

    t.[runtimeMinutes],

    AVG(CAST(f.[Gross] AS DECIMAL(18,2))) AS AverageGross

FROM [dbo].[BoxOfcFact] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

WHERE t.[runtimeMinutes] IS NOT NULL

GROUP BY t.[runtimeMinutes];

**3. Relationship between the number of votes and the average rating of movies.**

**a. The goal is to identify the correlation between votes and rating**

SELECT

    SUM(r.[numVotes]) AS TotalVotes,

    AVG(r.[averageRating]) AS AverageRating

FROM [dbo].[RatingFact1] r

GROUP BY r.[averageRating]

ORDER BY r.[averageRating];

**Director Success Metrics:**

**1. Identify directors with the most films rated above 7. Sort them in descending order**

**a. The goal is to identify which directors have made films that were received well by the audience**

SELECT TOP 10

    n.[primaryName] AS DirectorName,

    COUNT(\*) AS NumberOfMovies

FROM [dbo].[NameSCDDim] n

JOIN [dbo].[TitlePrincipalFact3] tp ON n.[Name\_SK] = tp.[Name\_FK]

JOIN [dbo].[RatingFact1] r ON tp.[Title\_FK] = r.[Title\_FK]

WHERE r.[averageRating] >= 8 AND tp.[category] = 'Director'

GROUP BY n.[primaryName]

ORDER BY NumberOfMovies DESC;

**2. Determine the directors who have directed the highest number of films overall and their respective gross earnings trends.**

**a. The goal is to identify if there are any notable directors who consistently contribute to highgrossing movies**

SELECT

    n.[primaryName] AS DirectorName,

    COUNT(tp.[Title\_FK]) AS TotalMovies,

    SUM(CAST(f.[Gross] AS DECIMAL(18,2))) AS TotalGross

FROM [dbo].[NameSCDDim] n

JOIN [dbo].[TitlePrincipalFact3] tp ON n.[Name\_SK] = tp.[Name\_FK]

JOIN [dbo].[BoxOfcFact2] f ON tp.[Title\_FK] = f.[Title\_SK]

WHERE tp.[category] = 'Director'

GROUP BY n.[primaryName]

ORDER BY TotalMovies DESC, TotalGross DESC;

**Actor and Actress Film Records:**

**1. List the top 10 actors/actresses with the most films rated between 4 and 7.**

**a. The goal is to identify the popular actors/actresses**

SELECT TOP 10

    n.[primaryName],

    COUNT(Distinct(tp.[Title\_FK])) AS NumberOfMovies,

    AVG(r.[averageRating]) AS AverageOfAverageRating

FROM [dbo].[NameSCDDim] n

JOIN [dbo].[TitlePrincipalFact3] tp ON n.[Name\_SK] = tp.[Name\_FK]

JOIN [dbo].[RatingFact1] r ON tp.[Title\_FK] = r.[Title\_FK]

WHERE r.[averageRating] BETWEEN 4 AND 7 AND (tp.[category] = 'Actor' OR tp.[category] = 'Actress')

GROUP BY n.[primaryName]

ORDER BY NumberOfMovies DESC, AverageOfAverageRating DESC;

**2. Compare the top 5 actors and actresses based on movie ratings.**

**a. The goal is to compare the top 5 actor and actresses**

SELECT TOP 5

    n.[primaryName],

    AVG(r.[averageRating]) AS AverageRating

FROM [dbo].[NameSCDDim] n

JOIN [dbo].[TitlePrincipalFact3] tp ON n.[Name\_SK] = tp.[Name\_FK]

JOIN [dbo].[RatingFact1] r ON tp.[Title\_FK] = r.[Title\_FK]

WHERE tp.[category] = 'Actor'

GROUP BY n.[primaryName]

ORDER BY AverageRating DESC;

SELECT TOP 5

    n.[primaryName],

    AVG(r.[averageRating]) AS AverageRating

FROM [dbo].[NameSCDDim] n

JOIN [dbo].[TitlePrincipalFact3] tp ON n.[Name\_SK] = tp.[Name\_FK]

JOIN [dbo].[RatingFact1] r ON tp.[Title\_FK] = r.[Title\_FK]

WHERE tp.[category] = 'Actress'

GROUP BY n.[primaryName]

ORDER BY AverageRating DESC;

**Seasonal Analysis:**

**1. How does movie performance vary across different seasons of the year? (based on gross earnings)**

**a. The goal is to identify which season has an impact on the revenue.**

SELECT

    b.[Season],

    SUM(CAST(f.[Gross] AS DECIMAL(18,2))) AS TotalGross

FROM [dbo].[BoxOfcDateDim] b

JOIN [dbo].[BoxOfcFact2] f ON b.[BoxOfcDate\_SK] = f.[BoxOfcDate\_SK]

GROUP BY b.[Season];

**2. Top 3 movies based on the season (spring, summer, fall)**

**a. Spring - Jan to April**

**b. Summer - May to August**

**c. Fall - September to December**

WITH SeasonalGross AS (

    SELECT

        b.[Season],

        t.[PrimaryTitle],

        SUM(CAST(f.[Gross] AS DECIMAL(18,2))) AS TotalGross

    FROM [dbo].[BoxOfcDateDim] b

    JOIN [dbo].[BoxOfcFact2] f ON b.[BoxOfcDate\_SK] = f.[BoxOfcDate\_SK]

    JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

    GROUP BY b.[Season], t.[PrimaryTitle]

),

RankedMovies AS (

    SELECT

        Season,

        PrimaryTitle,

        TotalGross,

        ROW\_NUMBER() OVER (PARTITION BY Season ORDER BY TotalGross DESC) AS RowNum

    FROM SeasonalGross

)

SELECT

    Season,

    PrimaryTitle,

    TotalGross

FROM RankedMovies

WHERE RowNum <= 3

ORDER BY Season, TotalGross DESC;

**Release Regions:**

**1. Identify movies that have had the widest release across multiple regions.**

**a. The goal is to identify which movie was released in most regions.**

SELECT TOP 10

    t.[PrimaryTitle],

    COUNT(Distinct(mr.[Region\_FK])) AS NumberOfRegions

FROM [dbo].[MovieRegionBri1] mr

JOIN [dbo].[TitleSCDDim] t ON mr.[Title\_FK] = t.[Title\_SK]

GROUP BY t.[PrimaryTitle]

ORDER BY NumberOfRegions DESC;

**Matrix:**

**Create a dashboard for the 9 box office movies that shows movie-wise:**

**1. No. of directors**

SELECT

    f.[Title\_SK],

    t.[PrimaryTitle],

    COUNT(DISTINCT p.[Person\_SK]) AS NumberOfDirectors

FROM [dbo].[BoxOfcFact2] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

JOIN [dbo].[TitlePrincipalFact3] tp ON t.[Title\_SK] = tp.[Title\_FK]

JOIN [dbo].[PersonDim] p ON tp.[Person\_FK] = p.[Person\_SK] AND p.[category] = 'Director'

GROUP BY f.[Title\_SK], t.[PrimaryTitle];

**2. No. of actors/actresses**

SELECT

    f.[Title\_SK],

    t.[PrimaryTitle],

    COUNT(DISTINCT p.[Person\_SK]) AS NumberOfActorsActresses

FROM [dbo].[BoxOfcFact2] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

JOIN [dbo].[TitlePrincipalFact3] tp ON t.[Title\_SK] = tp.[Title\_FK]

JOIN [dbo].[PersonDim] p ON tp.[Person\_FK] = p.[Person\_SK]

WHERE p.[category] IN ('Actor', 'Actress')

GROUP BY f.[Title\_SK], t.[PrimaryTitle];

**3. No. of writers**

SELECT

    f.[Title\_SK],

    t.[PrimaryTitle],

    COUNT(DISTINCT p.[Person\_SK]) AS NumberOfWriters

FROM [dbo].[BoxOfcFact2] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

JOIN [dbo].[TitlePrincipalFact3] tp ON t.[Title\_SK] = tp.[Title\_FK]

JOIN [dbo].[PersonDim] p ON tp.[Person\_FK] = p.[Person\_SK]

WHERE p.[category] = 'Writer'

GROUP BY f.[Title\_SK], t.[PrimaryTitle];

**4. Total worldwide gross earnings**

SELECT

    f.[Title\_SK],

    t.[PrimaryTitle],

    SUM(CAST(f.[Gross] AS DECIMAL(18,2))) AS TotalGross

FROM [dbo].[BoxOfcFact2] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

GROUP BY f.[Title\_SK], t.[PrimaryTitle]

ORDER BY TotalGross DESC;

**5. Average ratings**

SELECT

    f.[Title\_SK],

    t.[PrimaryTitle],

    AVG(r.[averageRating]) AS AverageOfAverageRating

FROM [dbo].[BoxOfcFact2] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

JOIN [dbo].[RatingFact1] r ON t.[Title\_SK] = r.[Title\_FK]

GROUP BY f.[Title\_SK], t.[PrimaryTitle]

ORDER BY AverageOfAverageRating DESC;

**6. No. of regions, it was released**

SELECT

    f.[Title\_SK],

    t.[PrimaryTitle],

    COUNT(DISTINCT mr.[Region\_FK]) AS NumberOfRegions

FROM [dbo].[BoxOfcFact2] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

JOIN [dbo].[MovieRegionBri1] mr ON t.[Title\_SK] = mr.[Title\_FK]

GROUP BY f.[Title\_SK], t.[PrimaryTitle]

ORDER BY NumberOfRegions DESC;

**7. No. of genres**

SELECT

    f.[Title\_SK],

    t.[PrimaryTitle],

    COUNT(DISTINCT mg.[Genre\_FK]) AS NumberOfGenres

FROM [dbo].[BoxOfcFact2] f

JOIN [dbo].[TitleSCDDim] t ON f.[Title\_SK] = t.[Title\_SK]

JOIN [dbo].[MovieGenreBri1] mg ON t.[Title\_SK] = mg.[Title\_FK]

GROUP BY f.[Title\_SK], t.[PrimaryTitle]

ORDER BY NumberOfGenres DESC;