

Overall status

Status-All the analysis required to be done has been completed.

Firstly, we understood the table structure and the relationship between records and attributes in the IMDB database. Then, we gathered parameters required for analysis such as nconst, tconst which are further used in analysis 1 and analysis 2. Then, we developed queries as per requirement for the given problem and visualized the result and the required analysis has been done by visualizing.

Objective:

The aim of this project is to understand the IMDB database schema and perform analysis on the given data and visualize the results.

File Descriptions

Analysis1a.sql:

This file consists of sql query for computing the number of movies an actor/actress acted in their career span per year.

Op_analysis1a.txt:

This file contains the output generated by analysis1a.sql in text format.

Analysis1b.sql:

This file consists of sql query for dividing the span of an actor/actress career into number of movies acted per decade.

Op_analysis1b.txt:

This file contains the output generated by analysis1b.sql in text format.

Analysis2.sql

This file consists of sql query for computing the number of movies all the actors/actresses acted in the given span of years and finding the actor and actress with most movies in the range and finding the actors and actresses who acted in 3 movies in the range of years.

Op_analysis2.txt

This file contains the output generated by analysis2.sql in text format.

Analysis 1a:

Approach:

Firstly, we have identified the required attribute i.e., nconst from table name_basics for the actors given in the parameter file and the actors of our choice. Then, we filtered the movies in which these actors acted from title_principals table. Then, with the movie's id i.e., tconst we filtered the start year of all the movies. Then, we grouped the start years to get the count of movies per year. Then, the result is ordered by start year to get the career span of actors.

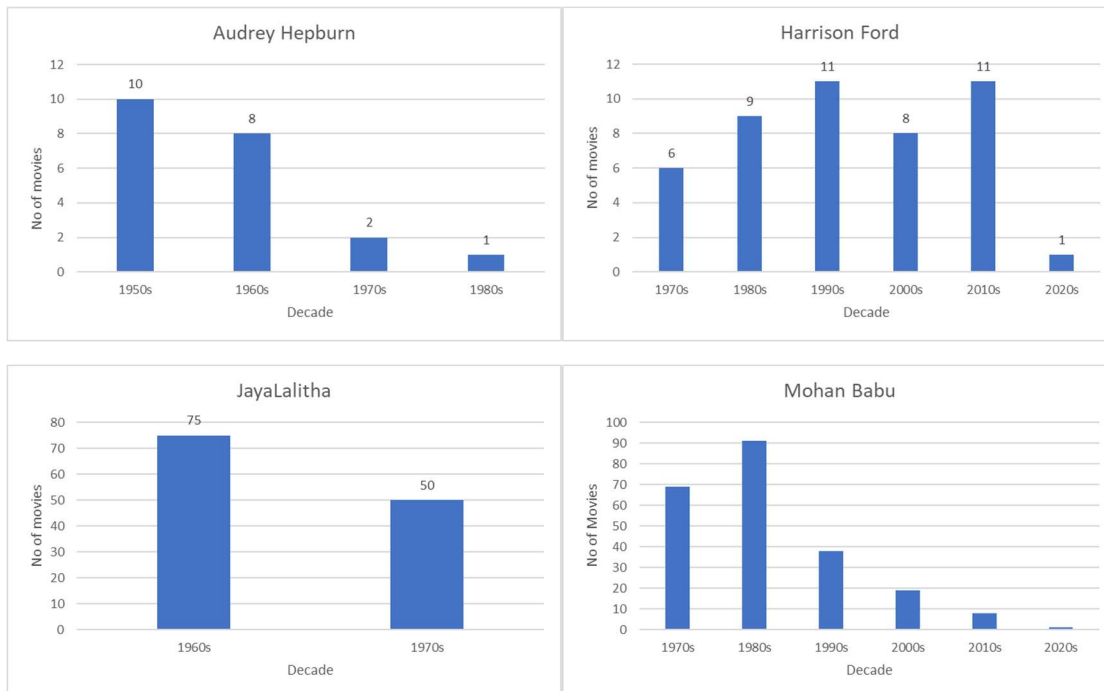
Analysis 1b:

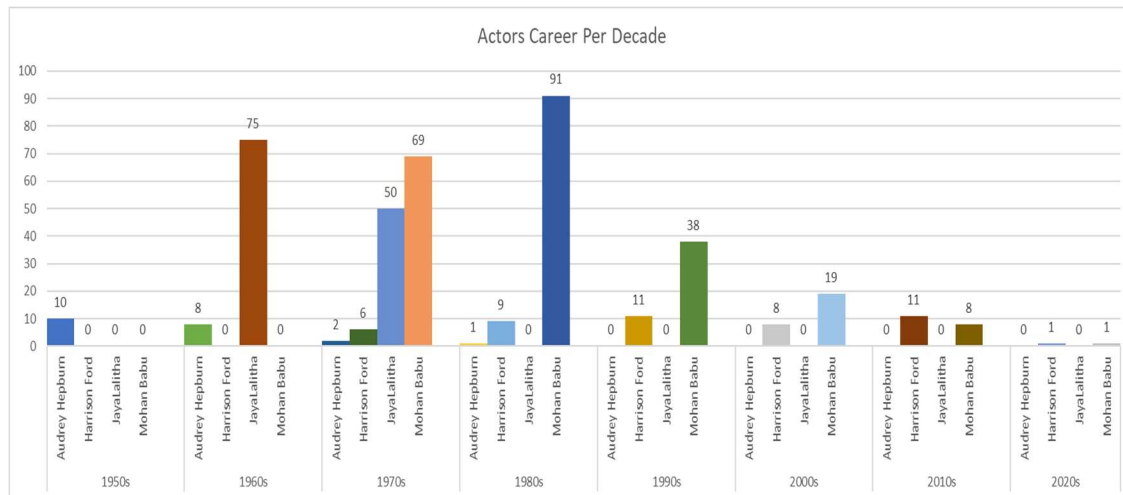
Approach:

In addition to the query done in analysis 1a, we have added a new column that classifies the year based on the decade it was made using case statements. Then the result is grouped by decade to get count of movies made by an actor in each decade.

Analysis:

For this analysis, the two actors given are Harrison Ford and Audrey Hepburn and the actors we have chosen are Mohan babu and jaya lalitha. We have plotted the careers of these four actors individually and then compared their careers in each decade with each other in another graph. All the plottings are done in excel from the data obtained from our query.





Inferences:

- Indian actors on an average acted in a greater number of movies when compared to their American actors.

Analysis 1c:

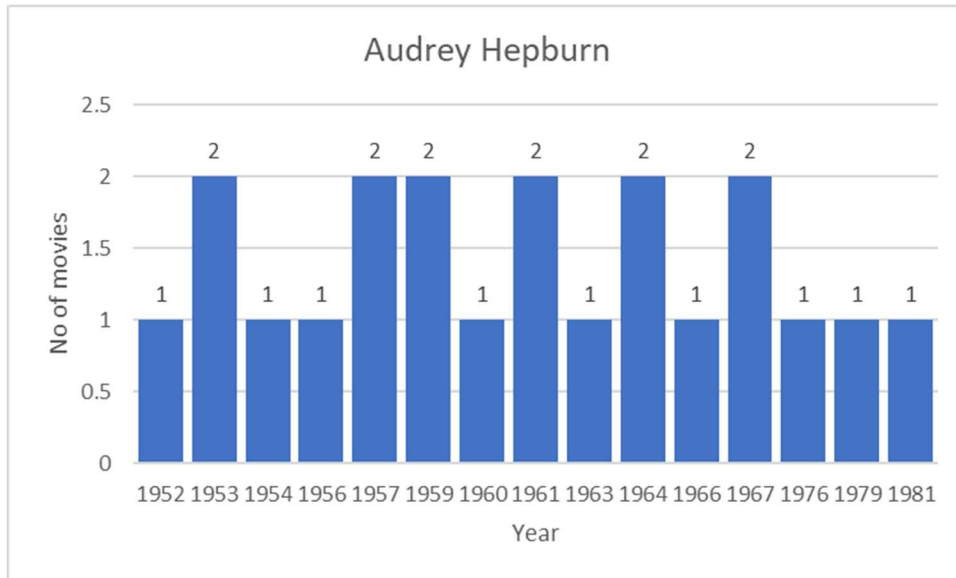
Approach:

In analysis 1c, we are asked to understand the career peaks and troughs of our 4 actors and draw inferences from them. Below are our findings for each actor and actress.

Analysis:

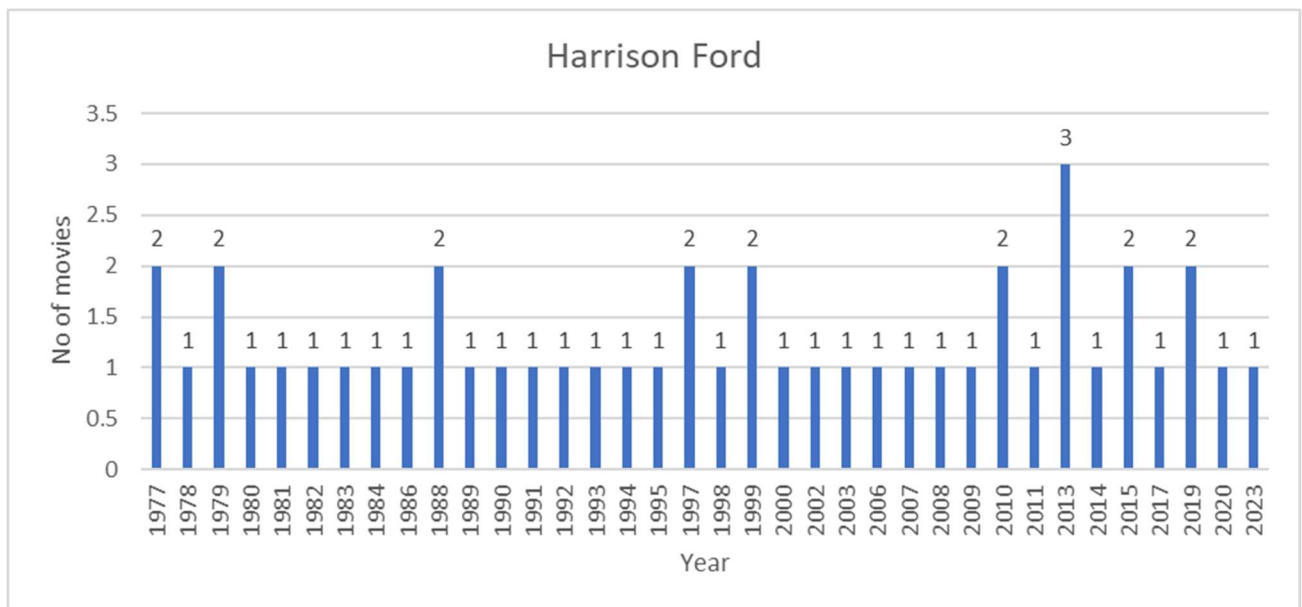
Audrey Hepburn:

- Audrey Hepburn started her acting career at the age of 23 and continued to act in movies consistently each year.
- She had a fair movie career. There were not many peaks in her career. At most she acted in 2 movies in a year.
- She had a career break in 1967 for a span of 9 years where she was away from movies. This period coincides with her divorce from her first marriage with Mel Ferrer in 1968 and her second childbirth in 1967.
- After her personal life was stable, she started acting in a few movies, however these movies did not receive any commercial success. Then she completely stopped acting in movies in 1981, then had her second divorce from Andrea Dotti in 1982 and started working on humanitarian projects for UNICEF.



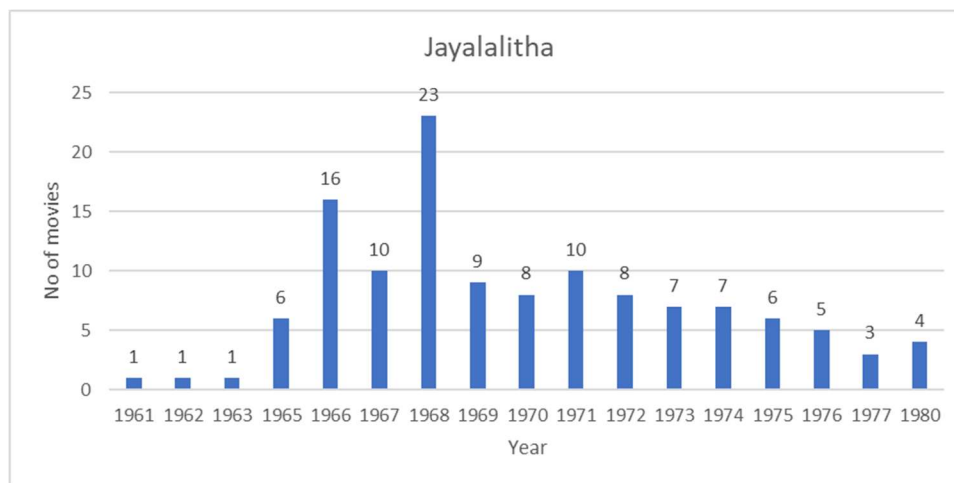
Harrison Ford:

- Harrison Ford started acting in movies in his mid 30's which is unusual for actors in Hollywood. Before working as an actor, Ford worked as a technical associate in production studios as a voice artist for unnamed roles in movies and television.
- Even Ford has a flat career with not many peaks in his career. He has performed consistently in his movies with at least a movie in a year.
- Ford had few career breaks in the 2000's due to health issues related to his old age. He came back to the movies after every break.
- He is now 80 and is still acting in movies which show his love towards the art of movies and film making.



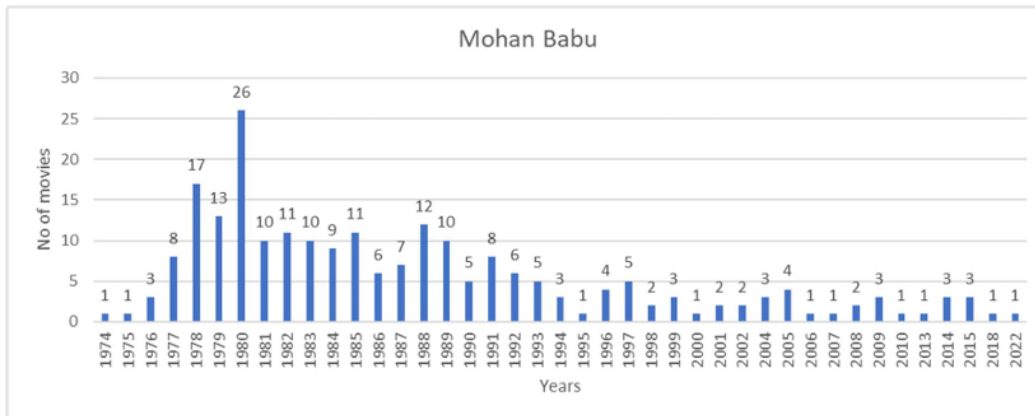
Jayalalitha:

- Jayalalitha started acting at the age of 13 as a child artist in tamil film industry. She moved to mainstream actress roles at the age of 16 which was a practice back then in 1965.
- She had a very successful career once she started performing as lead actress in movies. She was the leading actress in her times and had a peak career from 1965 to 1977.
- During this time, she was associated with her co actor MGR in his political activism and took a career break in 1977 to help him in his political endeavors.
- After 3 years she came back to acting in movies. However, she completely withdrew from acting after formally joining MGR's political party.
- Jayalalitha later went on to become Chief Minister of Tamilnadu, India and served in that role for 14 years in a total of 6 terms.



Mohan Babu:

- Mohan babu has an extensive acting career ranging from 1974 to 2022 in the Indian regional film industry, acting in more than 250 movies in his career.
- He had his peek career from 1978 to early 1990's. He started acting in antagonist roles and then switched to protagonist roles.
- Babu moved into politics from 1995 and we can see a dip in movies since then. He joined a political party and help a legislature role in Indian parliament from 1996 to 1997.
- He acts sparingly now only on guest appearances in movies.



Analysis 2:

Approach:

In analysis 2, we are asked to find the number of movies in which an actor/actress had acted in the given range of years. To achieve this, we have joined tables name_basics, title_principals and title_basics based on required parameters to get the number of movies an actor/actress acted by using group by function. These results are then filtered to get the actor and actress with most movies and the list of actors and actresses that acted in 3 movies in the given span of years. The result is filtered based on year.

Result:

PRIMARYNAME	YEAR	COUNT	REMARK
Tom Byron	1983	258	1.Actor with most movies
Sharon Mitchell	1982	117	2.Actress with most movies

Results of actors and actresses with 3 movies in our range are included in the output file included in the zip file.