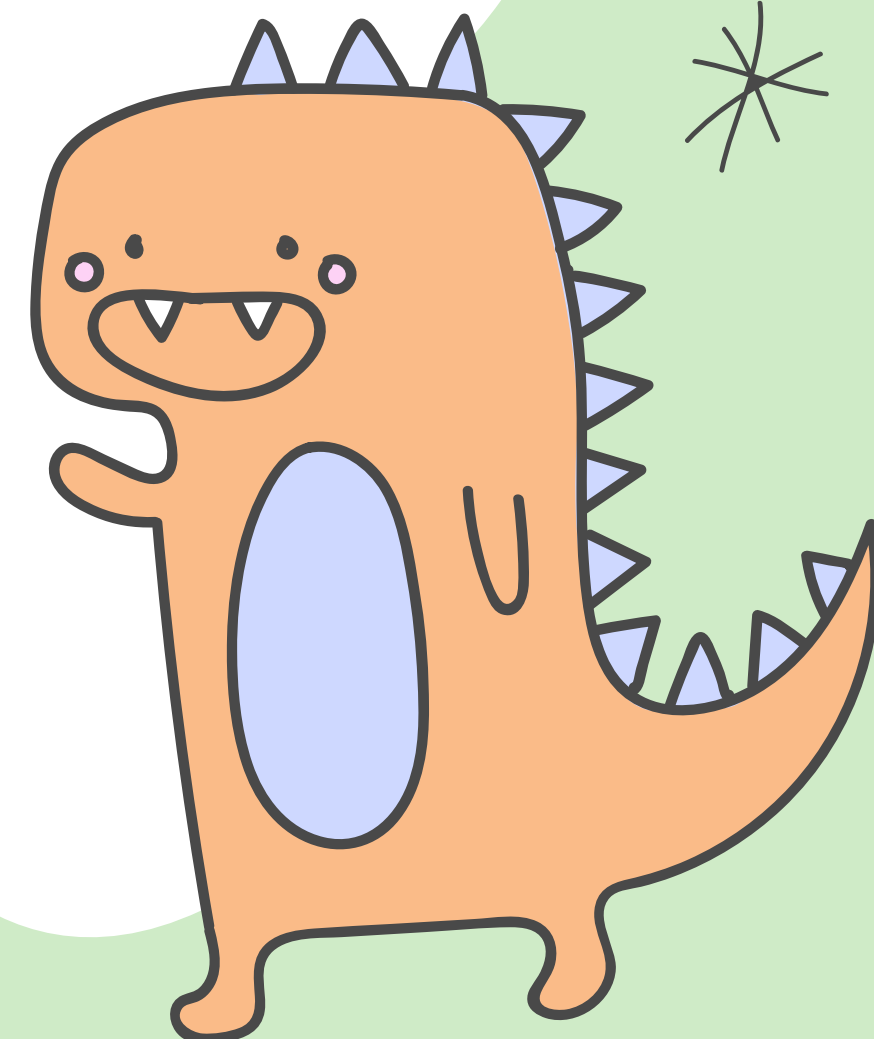
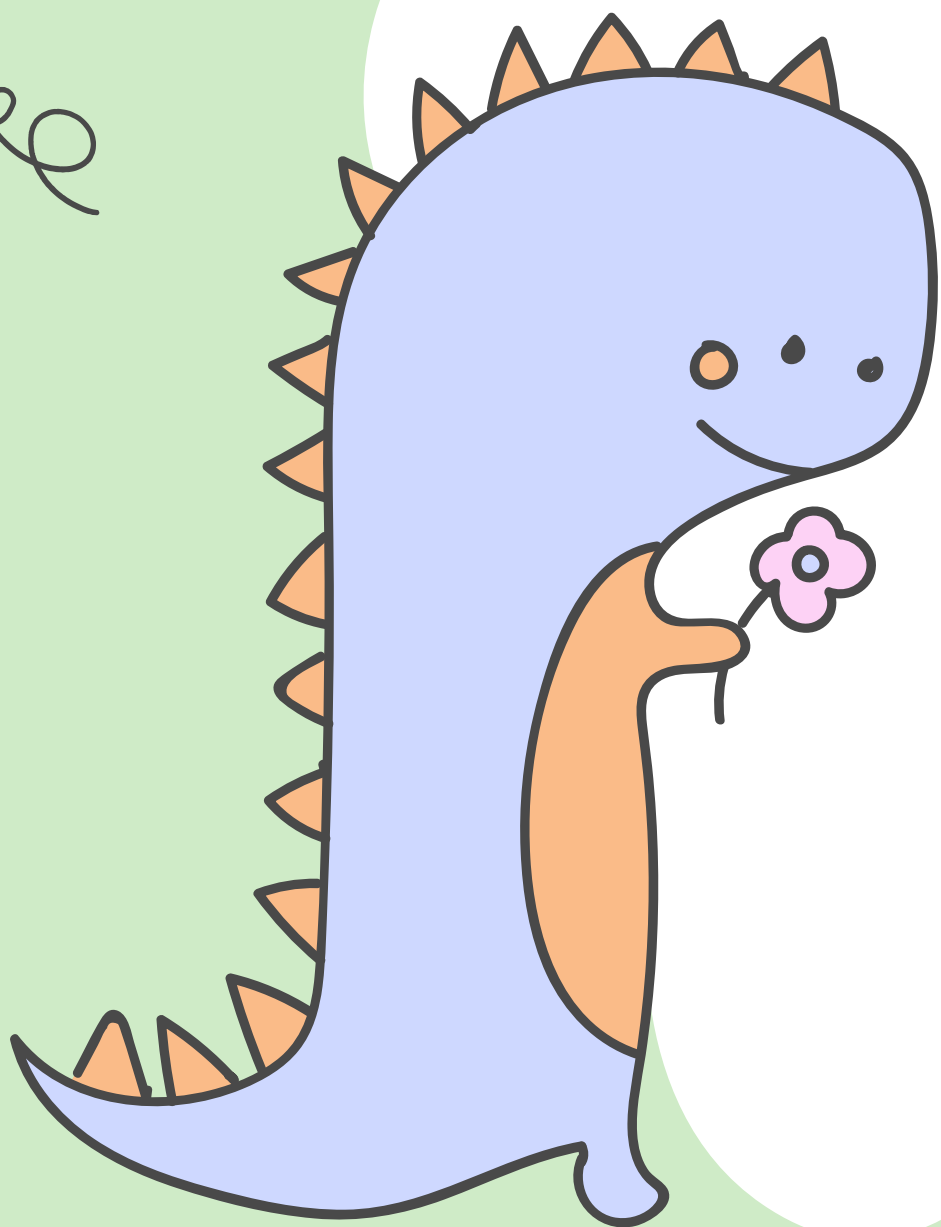
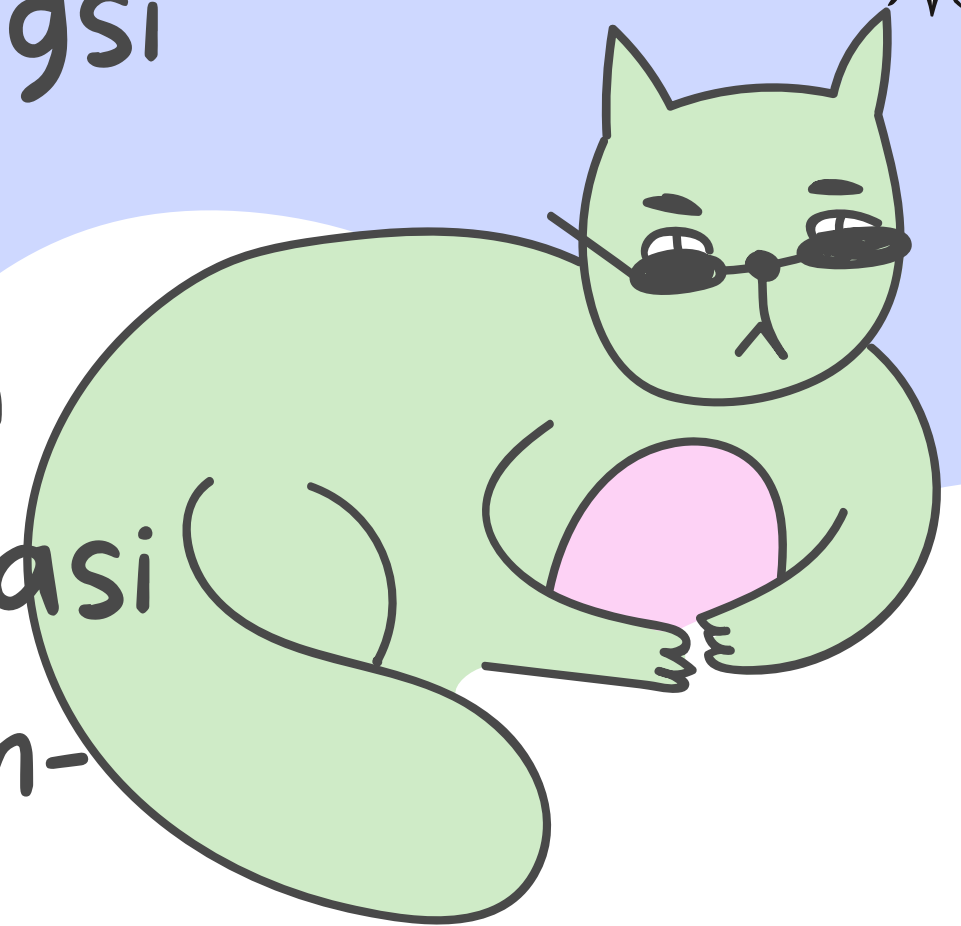
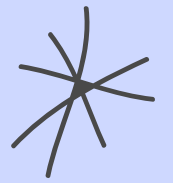


# Project - Data Analysis



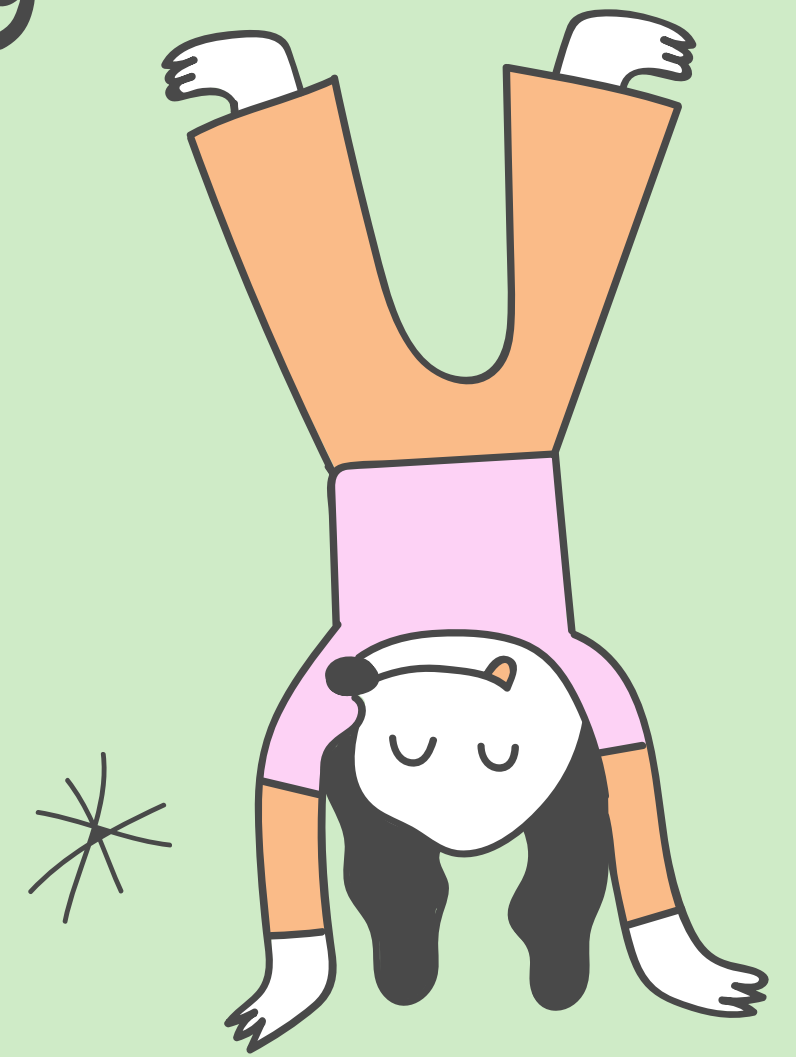
# Definisi

IMDb Movie Award adalah penghargaan tahunan yang diberikan oleh IMDb (Internet Movie Database) kepada film-film terbaik yang dirilis dalam periode waktu tertentu. IMDb adalah sebuah situs web yang berfungsi sebagai database film, televisi, dan acara TV. Penghargaan ini diberikan sebagai bentuk pengakuan dan apresiasi terhadap kualitas dan kesuksesan film-film tersebut.



# Masalah Solusi Optimum

aktris dengan pengalaman berakting di beragam genre film dengan rating tinggi merupakan kandidat yang paling cocok. Tim casting membutuhkan daftar berisi 5 rekomendasi aktris



# Data Collecting

- Actors
- Directors
- Directors\_Genres
- Movies
- Movies\_Directors
- Movies\_Genres
- Rolest

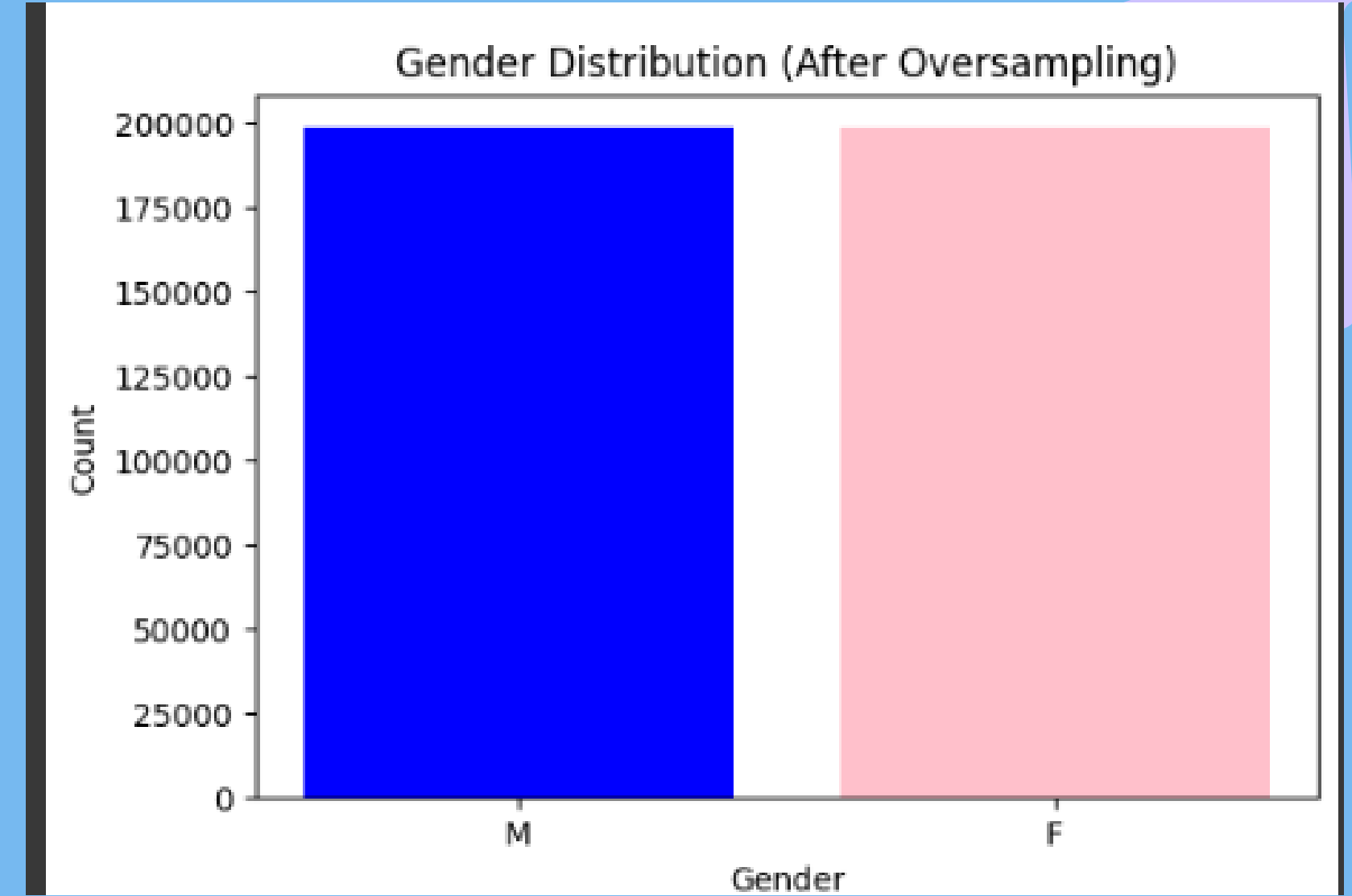
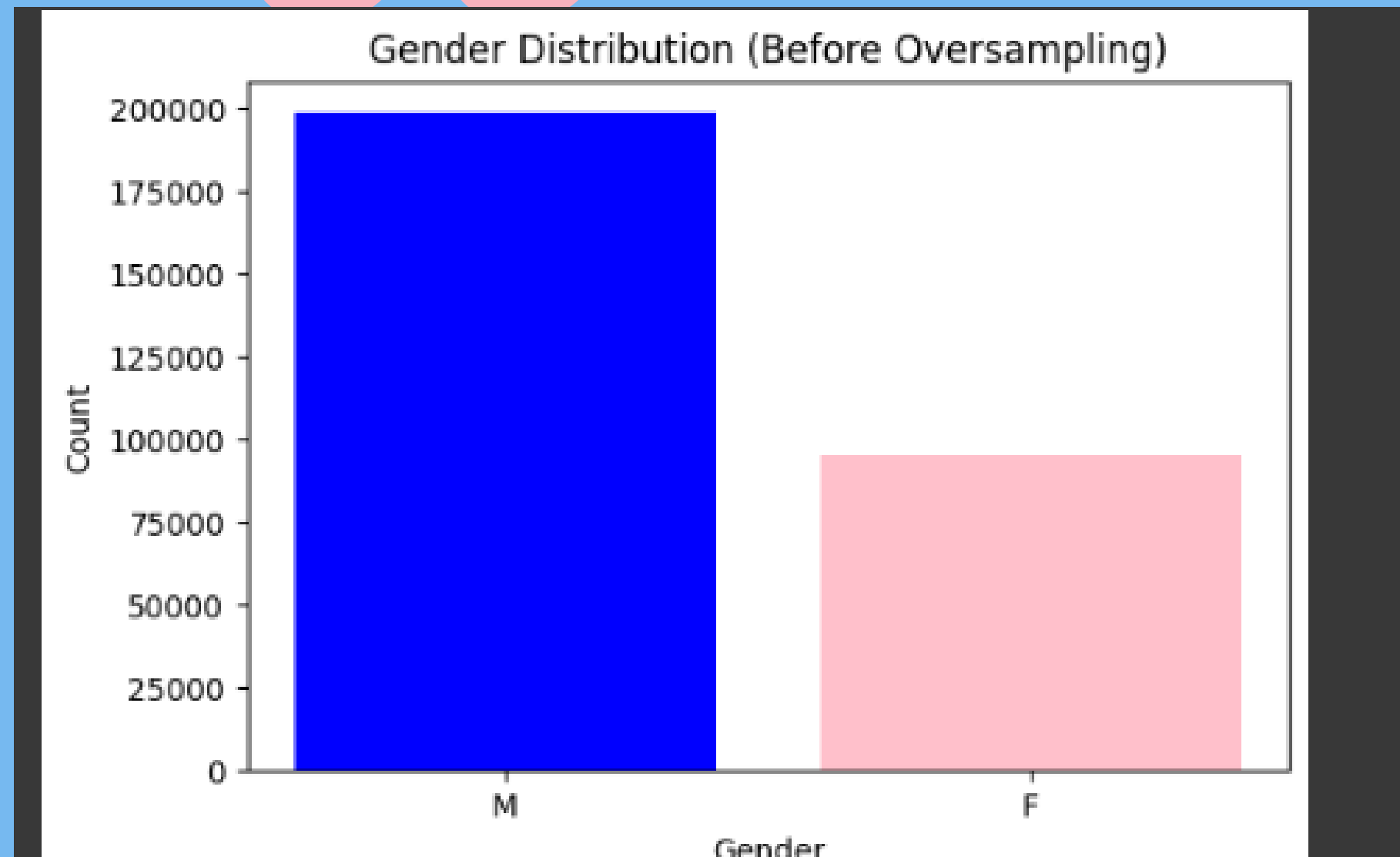


# Data Cleaning.



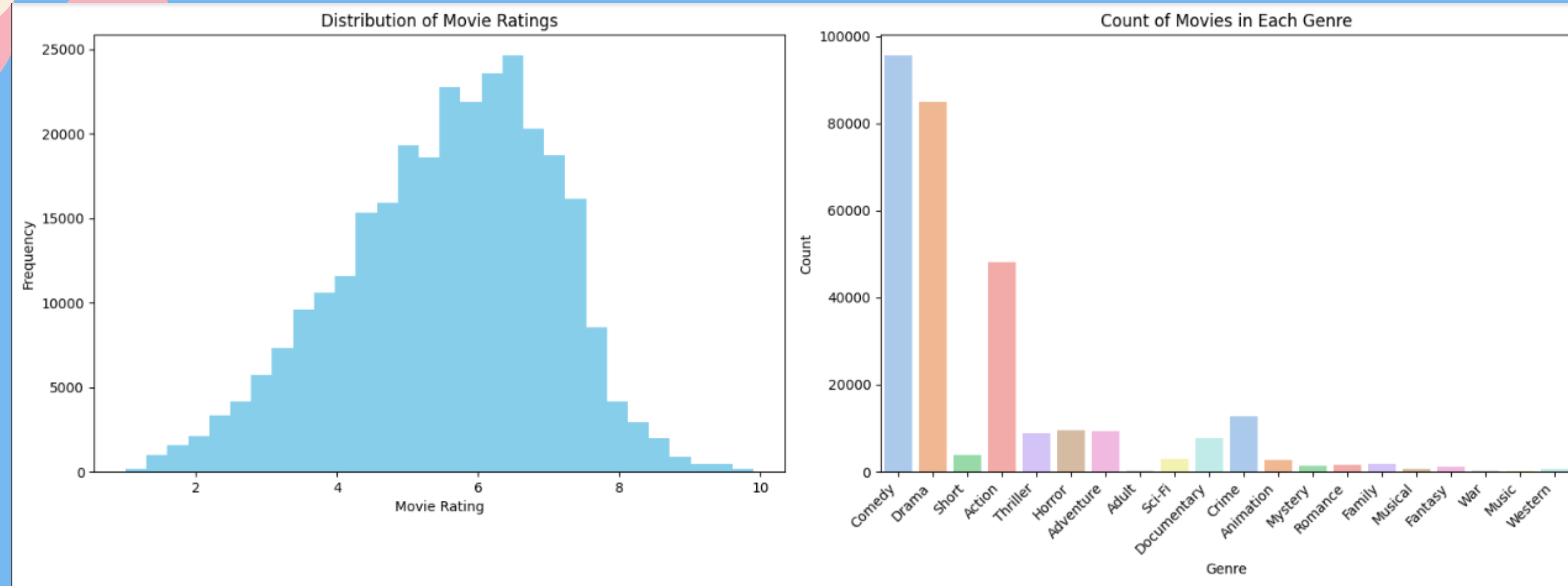
- Menbuang table yang tidak diperlukan
- mengubah nama table
- membuat full name untuk actors
- ambil data hanya di range tahun 1990-1999
- cek missing value

# Data Visualization



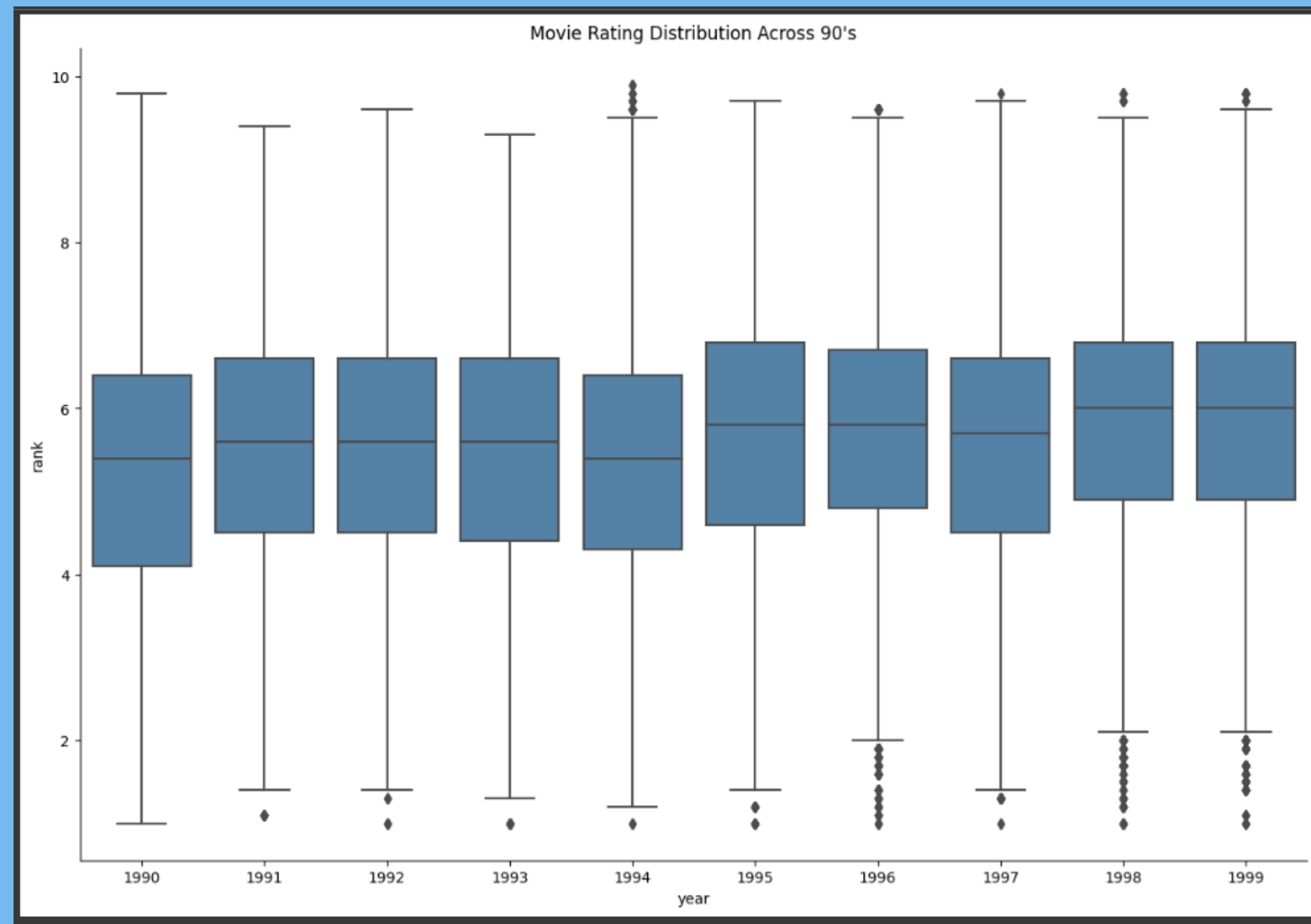
★ untuk membuat distribusi kelas menjadi lebih seimbang sehingga model pembelajaran mesin yang dihasilkan dapat mengenali dan mempelajari pola dari kelas minoritas dengan lebih baik.

# Data Visualization



Distribusi rating film banyak terpusat pada rating 4-5 dan sangat sedikit pada nilai rating 8 ke atas. Hal ini dapat menunjang asumsi analisis dalam menggunakan rating film sebagai variabel penilaian karena jumlah film dengan nilai rating tinggi sedikit jumlahnya. Dengan begitu, akan mudah dibedakan sutradara dengan rating film tinggi dengan yang rata-rata.

# Data Visualization

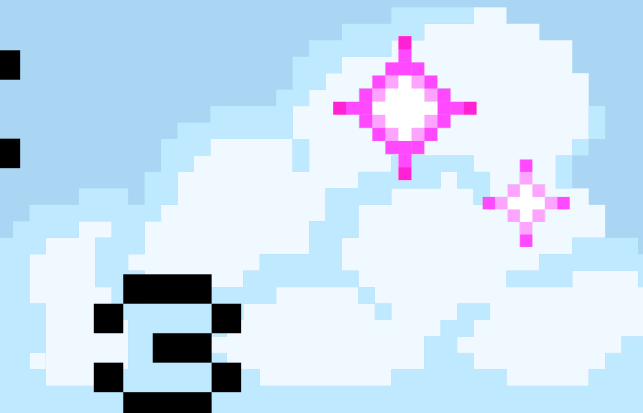


Berdasarkan boxplot di atas, tidak ditemukan satu periode waktu dengan kualitas film yang tinggi secara mayoritas. Namun dapat dilihat bahwa film dengan rating tertinggi selama 1 dekade diproduksi pada tahun 1990..





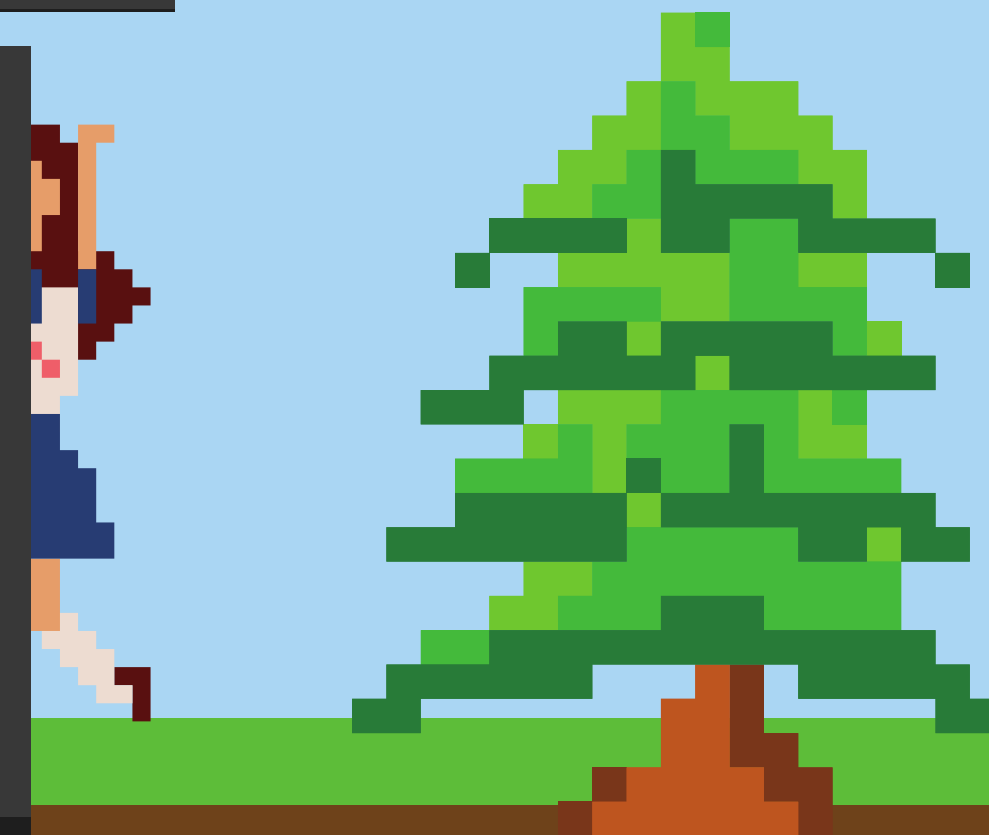
# SISTEM ANALISIS PERTANYAAN NOMOR 3



	director_name	gender	movie_count	movie_rating_mean	career_duration	num_genres
140980	Seymour Cassel	M	44	5.390909	9	11
121128	Paul (II) Guilfoyle	M	30	6.020000	9	10
111710	Miranda (I) Richardson	F	22	5.940909	9	10
76159	John Hurt	M	27	5.929630	9	10
65626	James Earl Jones	M	33	5.903030	9	10

daftar berisi 5 rekomendasi aktris:

- Seymour Cassel
- Paul (II) Guilfoyle
- Miranda (I) Richardson
- John Hurt
- James Earl Jones







MAKASIH YA BANG

