Universidad Francisco Marroquín

Data Wrangling

Catedrático: Juan Carlos Girón

Auxiliar: José Josué

Laboratorio #3

**Instrucciones:** Utilizando la librería de dplyr en R, conteste a las siguientes preguntas sobre las tablas de IMDB.

<https://relational.fel.cvut.cz/dataset/IMDb>

1. Información general sobre la base de datos:
   1. ¿Cuántas películas existen en la base de datos?
   2. ¿Cuántos directores?
2. ¿Cuál es el número promedio de géneros por director?

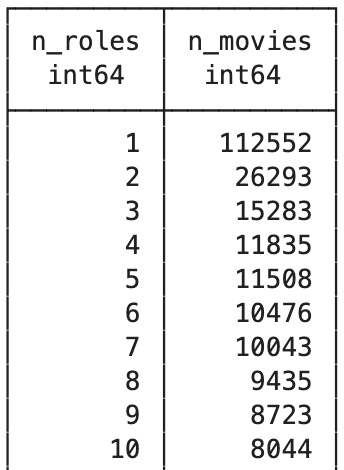
*Nota: utilizar director\_genres*

1. Genere un nuevo reporte por “Role” con la siguiente información:
   1. Número de películas
   2. Número de actores
   3. Número de actrices
   4. Número de directores
2. Genere un nuevo reporte con la siguiente información:
   1. Información del director (ID, nombre, apellido)
   2. Número de películas que ha dirigido
   3. Número de actores con los que ha trabajado
   4. Género más común de sus películas

*Nota: esta table deberá tener una única fila por director. En total, la tabla deberá registrar la misma cantidad de filas que su respuesta al inciso 1b.*

1. Encuentre la distribución de “Roles” por las siguientes dimensiones:
   1. Película
   2. Director

*Nota: el resultado deberá ser una tabla o una gráfica similar al ejemplo debajo:*

**