

NOMBRE: Benjamín Farías Valdés

N.ALUMNO: 17642531



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

IIC2026 — Visualización de Información — 2' 2020

Entrega 1

1. Abstracción de Tareas

- **Descubrir y Disfrutar la Tendencia y Distribución** de la proporción entre protagonistas de sexo masculino y femenino a través del tiempo. Estas tareas corresponden a las principales, ya que permiten validar o contrastar las hipótesis planteadas sobre la proporción de protagonistas masculinos y femeninos a lo largo de la historia. Además, le proporciona entretenimiento al grupo de amigos al poder conversar y aprender sobre el tema. Se necesitan 2 atributos del dataset: **el año de la película y el sexo biológico del protagonista**. Además, se requiere **derivar la cantidad de protagonistas tanto femeninos como masculinos por cada año**, de forma que finalmente se pueda obtener la **razón (por año) entre protagonistas según su sexo biológico (cantidad de protagonistas femeninos / cantidad de protagonistas masculinos)**.
- **Comparar y Resumir la Distribución** de protagonistas según sexo. Esta tarea entrega información complementaria, contrastando las distribuciones de los protagonistas según sexo por separado, antes de juntarlas en un solo atributo de proporción. Los datos necesarios para esta parte son los mismos que la anterior, pero **utilizando sólo la cantidad de protagonistas femeninos y masculinos por año**.

NOMBRE: Benjamín Farías Valdés

N.ALUMNO: 17642531



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN

IIC2026 — Visualización de Información — 2' 2020

Entrega 1

2. Abstracción de Datos

- **title:**
 - **Categorico**, porque corresponde a un nombre.
 - **Llave**, ya que no existen títulos iguales entre películas distintas.
- **site:**
 - **Categorico**, porque corresponde a un enlace (texto).
 - **Llave**, ya que no existen enlaces iguales entre películas distintas.
- **country:**
 - **Categorico**, porque corresponde a un nombre de país.
 - **Valor**, ya que puede ser compartido entre películas distintas.
- **yearReleased:**
 - **Ordenado**, porque cuenta con un orden temporal intrínseco.
 - **Cuantitativo**, ya que se pueden calcular diferencias exactas entre los años.
 - **Secuencial**, ya que a partir de un cierto año mínimo se empezaron a crear películas biográficas.
 - **NO Cíclico**, ya que los años incrementan indefinidamente.
 - **Valor**, ya que puede ser compartido entre películas distintas.
- **boxOffice:**
 - **Ordenado**, porque la cantidad de dinero cuenta con un orden numérico intrínseco.
 - **Cuantitativo**, ya que la cantidad de dinero cuenta con operaciones aritméticas exactas.
 - **Secuencial**, ya que toma valores positivos.
 - **NO Cíclico**, ya que sus valores incrementan indefinidamente.

- **Valor**, ya que puede ser compartido entre películas distintas.
- **director:**
 - **Categorico**, porque corresponde a un nombre.
 - **Valor**, ya que puede ser compartido entre películas distintas.
- **subject:**
 - **Categorico**, porque corresponde a un nombre.
 - **Valor**, ya que puede ser compartido entre películas distintas con el mismo protagonista.
- **typeOfSubject:**
 - **Categorico**, porque corresponde a una descripción.
 - **Valor**, ya que puede ser compartido entre películas distintas con protagonistas similares.
- **raceKnown:**
 - **Categorico**, porque corresponde un booleano.
 - **Valor**, ya que puede ser compartido entre películas distintas.
- **subjectRace:**
 - **Categorico**, porque corresponde un nombre de etnicidad.
 - **Valor**, ya que puede ser compartido entre películas distintas con protagonistas de la misma raza.
- **personOfColor:**
 - **Categorico**, porque corresponde un booleano.
 - **Valor**, ya que puede ser compartido entre películas distintas.
- **subjectSex:**
 - **Categorico**, porque corresponde a un valor binario.
 - **Valor**, ya que puede ser compartido entre películas distintas con protagonistas del mismo sexo.
- **leadActorActress:**
 - **Categorico**, porque corresponde a un nombre.
 - **Valor**, ya que puede ser compartido entre películas distintas con los mismos actores/actrices.
- **numberOfMales:**
 - **Ordenado**, porque representa una cantidad numérica.
 - **Cuantitativo**, ya que la cantidad de protagonistas cuenta con operaciones aritméticas exactas para comparación.
 - **Secuencial**, ya que toma valores positivos.
 - **NO Cíclico**, ya que sus valores incrementan indefinidamente.
 - **Valor**, ya que puede ser compartido entre años distintos.
- **numberOfFemales:**
 - **Ordenado**, porque representa una cantidad numérica.

- **Cuantitativo**, ya que la cantidad de protagonistas cuenta con operaciones aritméticas exactas para comparación.
- **Secuencial**, ya que toma valores positivos.
- **NO Cíclico**, ya que sus valores incrementan indefinidamente.
- **Valor**, ya que puede ser compartido entre años distintos.

■ **ratioOfSexByYear:**

- **Ordenado**, porque representa una proporción numérica.
- **Cuantitativo**, ya que la proporción cuenta con operaciones aritméticas exactas.
- **Secuencial**, ya que toma valores positivos de 0 a 1.
- **NO Cíclico**, ya que sus valores no pueden salirse del intervalo, y tampoco se da la vuelta al superar el máximo.
- **Valor**, ya que puede ser compartido entre años distintos.

El dataset es de **tipo tabla**, ya que sus datos están organizados en filas y sus atributos en columnas. Su comportamiento será **estático**, ya que los datos derivados no serán guardados en el dataset, sino que serán utilizados exclusivamente para generar la visualización de forma inmediata.

Finalmente, cada película corresponde a un dato de tipo **ítem**, mientras que las columnas del dataset corresponden a los **atributos**.