

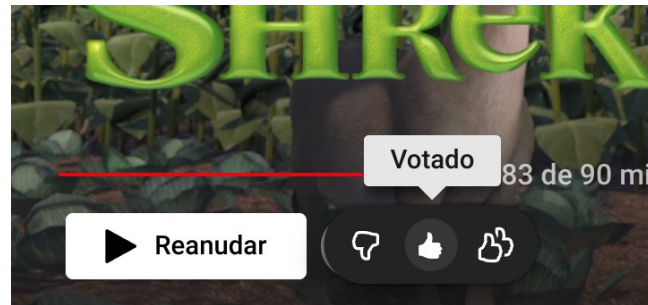
# Calificación de películas

Equipo 6  
Emmanuel Peto Gutiérrez  
Rodrigo Fernando Velázquez Cruz

# Definición del problema

Se requiere un sistema que ayude a un usuario a elegir una o varias películas que posiblemente le podrían gustar. Se tiene como información:

- Las calificaciones que este usuario le ha dado a las películas que ya ha visto.
- Las calificaciones que otros usuarios le han dado a otras películas.



# Objetivo

Analizar los datos que se tienen sobre las películas para (posiblemente) elegir ver una. Algunos ejemplos de la información que se puede obtener:

- Calificación promedio de una película.
- Cantidad de películas de un género específico.
- Moda de edades para una película.

PELÍCULAS POPULARES



El Stand de los  
Besos  
45 %



Batman vs  
Superman: El Origen  
27 %



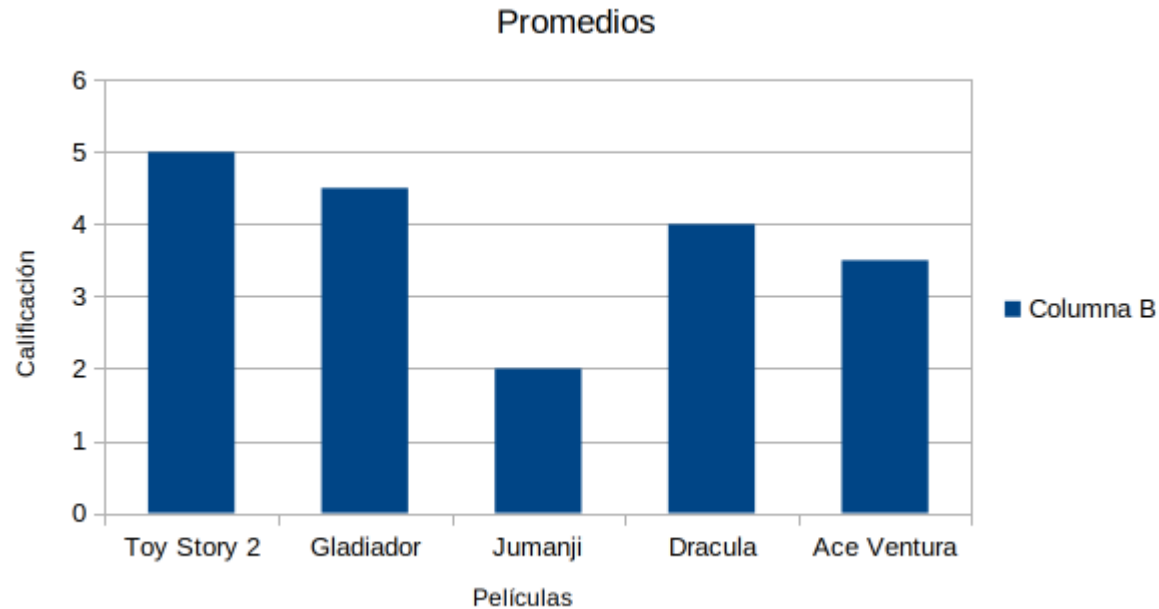
Mi Vida a los 17  
95 %



¡Huye!  
99 %

# Presentación de los datos

- Se mostrarán gráficas de barras de la consulta realizada.



# Extensión de la BD

Estrategias para extender las columnas.

- Separar el título de la película de su año.
- Inventar datos sobre los usuarios (edad, nombre, apellido).
- Hacer “join” de las diferentes tablas.



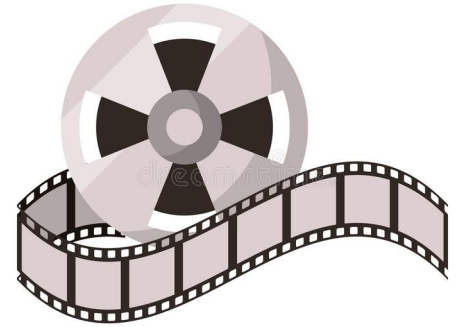
# Base de datos a usar

Se usarán los datos de “MovieLens”.

- 27,000,000 ratings
- 1,100,000 etiquetas
- 58,000 películas
- 280,000 usuarios

(posiblemente solo se utilice un subconjunto de estos datos)

Enlace: <https://grouplens.org/datasets/movielens/latest/>



# Base de datos a usar

grouplens

[about](#)

[datasets](#)

[publications](#)

[blog](#)

## MovieLens Latest Datasets

These datasets will change over time, and are not appropriate for reporting research results. We will keep the download links stable for automated downloads. We will not archive or make available previously released versions.

*Small:* 100,000 ratings and 3,600 tag applications applied to 9,000 movies by 600 users. Last updated 9/2018.

- [README.html](#)
- [ml-latest-small.zip](#) (size: 1 MB)

*Full:* 27,000,000 ratings and 1,100,000 tag applications applied to 58,000 movies by 280,000 users. Includes tag genome data with 14 million relevance scores across 1,100 tags. Last updated 9/2018.

## Datasets

[MovieLens](#)

[WikiLens](#)

[Book-Crossing](#)

[Book Genome Dataset](#)

[Jester](#)

[EachMovie](#)

# Adicional

- Hacer un sistema de recomendaciones basado en coeficientes de correlación de Pearson.
- (solo si nos da tiempo)

