# Recomendación de películas

En esta era de la conectividad constante, las plataformas ofrecen una amplia gama de programas, películas y documentales. El precio accesible, el amplio catálogo de contenidos, la flexibilidad de horarios, la ausencia de parones publicitarios y la comodidad de estar en tu propia casa son algunas de las ventajas del streaming respecto a la televisión. En ocasiones, el espectador encuentra dentro de la vorágine del día un espacio de desconexión y lo destina a ver una película, pero seleccionar que mirar puede representar gran parte del tiempo disponible. Por tal motivo, sería ventajoso contar con un análisis que permita al espectador guiarse en su elección de forma rápida según sus preferencias. Tambien puede ser utilizado por las empresas cinematográficas.

**Grupo 15:**

* Korb Melisa
* Machetti Natalia

**Objetivos**

Objetivo general:

Recomendar películas por genero según sus características y la opinión del público. A partir de dicha recomendación recopilar cuales son las características de las películas más aceptadas por el público para contribuir al ámbito cinematográfico actual.

Objetivos específicos:

***Objetivo específico 1:*** analizar la aceptación de películas por género, director y país de origen según vistas y likes

*Objetivo específico 2:* determinar si existe relación, y cómo se comporta, entre la cantidad de vistas y la puntuación

*Objetivo específico 3:* verificar si existe relación entre la cantidad de vistas de las películas y su duración

**Selección de Dataset**

El Dataset seleccionado fue extraído de <https://www.kaggle.com/datasets/ky1338/10000-movies-letterboxd-data>. Los datos se recopilaron en una lista de Letterboxd. Esta lista, diseñada como una ruleta de películas, cuenta con 10 002 entradas de diversos géneros y décadas. Los datos fueron tomados el 8/4/2025.

Organización de información en el Dataset original:

De cada una de las películas enlistadas se encuentra información acerca de los actores, director, productora, tiempo de duracion, cantidad de vistas y puntuación.

**Librerías**

Las librerías utilizadas para explorar y visualizar los datos del Dataset son:

* pandas
* numpy
* matplotlib
* seaborn

**Exploración y visualización de datos:**

* Caracterización del DataFrame
  + Cantidad de filas (observaciones) y columnas (features)
  + Tipo de variables
  + Valores faltantes (identificación)
* Feature engineering
  + Selección de features: Eliminación y cambio de nombres de columnas
  + Combinación de variables
* Distribución de variables continuas
  + Gráficos: histograma
  + Numérico: media aritmética, mediana, desvios, rango y quartiles
* Outliers
  + Identificación: Rango intercuartílico
* Frecuencia de categorías
  + Gráficos: barras
  + Numérico: tabla de frecuencia

Para cumplir el *objetivo 1* se realizó

* Comparación de conjunto de datos
  + Gráfico: boxplots y gráficos de barra

Para cumplir con los *objetivos 2 y 3* se realizó:

* Evaluación de relaciones: correlación
  + Numéricos: análisis de correlación
  + Gráfico: Scatterplots

**Conclusión**

A partir del análisis del porcentaje entre likes y vistas podemos observar que de las vistas totales solo entre el 12 y 24% recibe Likes.

Respecto al genero la mayor cantidad de películas disponibles son de drama, seguidas por comedia y horror. Dentro de los directores encontramos en primer lugar en cantidad de producciones disponibles a Chuck Jones (Warner Bros) y el país con mayor cantidad de películas disponibles es Estados Unidos, seguido en significativa menor medida por Inglaterra.

Otro análisis merecen las estrellas, ya que la mayoría de los espectadores califican a las películas con 4, evidentemente el público reserva el 5 solo para algunas ocasiones.

O*bjetivo 1:* el género más visto es comedia-aventura y en segundo lugar drama, una combinación de estos dos puestos es la que tiene la mayor cantidad de Likes, comedia-drama.

Los directores más vistos son Quentin Tarantino y Christopher Nolan, pero no son los que reciben más Likes, y en primer lugar de Likes se ubica el Diector Damien Chazelle.

Las películas de USA-UK, USA-Alemania y Corea del Sur son las más vistas, pero sorprendentemente las películas de Corea del Sur son las que más Likes reciben.

*Objetivos 2 y 3:* La correlación entre películas más vistas y likes permite observar que efectivamente las películas más populares son las preferidas por los espectadores. Mientras que la cantidad de vistas no está condicionada por la duración.