

EJERCICIO 1 [10%]

Justificad brevemente vuestra selección, ya sea por motivos personales o profesionales.

El proyecto de creación de la visualización de datos se realizará sobre el análisis de la oferta de diferentes plataformas de *streaming* (Netflix, Hulu, Prime Video y Disney +) en base a los diferentes criterios del usuario. Este proyecto está motivado por la necesidad que existe actualmente de elegir entre las cada vez más numerosas plataformas de *streaming* que existen, ya que cada una suele centrarse en un tipo de contenidos, un público concreto, o unos determinados géneros.

La visualización que se creará pretenderá ser una herramienta para un usuario que quiera elegir la plataforma que más le convenga a conveniencia de sus preferencias en lo referente a géneros, tipo de contenidos (serie o película), países originarios de los títulos, puntuación de los títulos en IMDB o público al que van dirigidos los contenidos.

EJERCICIO 2 [10%]

La relevancia del conjunto de datos en su contexto. ¿Son datos actuales? ¿Tratan un tema importante por algún colectivo concreto? ¿Se ha tenido en cuenta la perspectiva de género?

El conjunto de datos que se usará para crear la visualización se corresponde con datos actuales que tienen en cuenta los títulos añadidos a estas plataformas de *streaming* hasta agosto de 2021.

Como se ha comentado anteriormente, una herramienta que ayude al usuario a elegir que plataforma se adecúa más a sus gustos y preferencias es de gran utilidad si tenemos en cuenta que abundan las plataformas de *streaming*, cada una con un catálogo propio y un precio mensual de suscripción considerable. Permitir al usuario un análisis visual sencillo y rápido que le permita identificar cual de estas plataformas le resulta más conveniente le permitirá realizar la suscripción a un menor número de plataformas.

La perspectiva de género no tiene cabida en este tipo de datos, ya que solo se tienen en cuenta datos sobre los contenidos disponibles en cada plataforma.

EJERCICIO 3 [25%]

La complejidad (medida, variables disponibles, tipos de datos, etc.). ¿Tiene del orden de centenares o miles de registros? ¿Tiene del orden de decenas de variables? ¿Combina datos categóricos y cuantitativos? ¿Incluye otros tipos de datos? Evitad los conjuntos excesivamente simples.

Los datos que se tratarán contarán con variables cualitativas como el tipo de contenido (película o serie), el género (acción, drama, romance, aventura, etc.) de los contenidos, el público al que van dirigidos (todos los públicos, mayores de 13, mayores de 16, mayores de 18, etc.), el país donde se produjo el contenido; como también variables cuantitativas como puedan ser el año de la producción, la puntuación que posee en *IMDB*, la puntuación que posee en *Rotten Tomatoes*. Además, podemos considerar que los datos cuentan con una serie de variables de texto, como puedan ser el título o el director que también pueden ser interesantes de analizar.

EJERCICIO 4 [25%]

La originalidad. No repetid los conjuntos de datos clásicos. Podéis, por ejemplo, combinar o mejorar visualizaciones existentes. ¿hay otras visualizaciones basadas en este conjunto de datos? ¿Es una evolución o actualización de un conjunto anterior? ¿Habéis enriquecido un conjunto de datos ya existente?

En la práctica anterior realicé una de las visualizaciones en base a los títulos disponibles en Netflix. Esto motivó la creación de una nueva visualización enriquecida que permitiera el análisis de los contenidos disponibles en diferentes plataformas de *streaming* para así poder comparar cuál resulta la mejor

opción en base a unas preferencias. Dado que actualmente estas plataformas disponen de contenidos cinematográficos y televisivos, los datos que se usarán para la visualización provienen de dos conjuntos de datos diferentes disponibles en Kaggle. Ambos conjuntos contienen datos de las plataformas a considerar (Netflix, Hulu, Prime Video y Disney +):

- Películas disponibles en la plataforma. Datos disponibles en:
<https://www.kaggle.com/ruchi798/movies-on-netflix-prime-video-hulu-and-disney>
- Series disponibles en la plataforma. Datos disponibles en:
<https://www.kaggle.com/ruchi798/tv-shows-on-netflix-prime-video-hulu-and-disney>

Se combinarán ambos conjuntos de datos para poder realizar una comparación efectiva que tenga en cuenta los dos tipos de títulos más consumidos: las películas y las series. De esta forma, se creará una visualización original, pues las visualizaciones existentes tienen en cuenta únicamente uno de estos dos tipos de contenidos.

EJERCICIO 5 [30%]

Las cuestiones que responderéis con la visualización de datos, ¿Tienen en cuenta los puntos anteriores? ¿Están bien planteadas? ¿Son adecuadas por el conjunto de datos elegido? (30%)

La principal pregunta a contestar mediante esta visualización es: *¿cuál es la mejor plataforma para este usuario?*

A su vez, la respuesta a esta pregunta se desglosa en contestar otra serie de preguntas:

- ¿Qué plataforma contiene más títulos?
- ¿Dónde se encuentran los títulos con mejor puntuación?
- ¿Cuáles son los géneros, públicos y países de origen de los títulos de cada plataforma?
- ¿Cuándo se produjeron los contenidos de cada plataforma?
- ¿Qué plataforma contiene títulos relacionados con un director concreto?

La visualización que se creará pretenderá dar respuesta a todas estas preguntas.