

Alumno: Daniel Izquierdo Bonilla

Pregunta 1

1. Captura o recogida de datos: ¿Qué datos internos/externos se recogieron? ¿De quién? ¿Cuál era el objetivo de cada una de ellos?

Se utilizaron los siguientes datos:

- Búsquedas realizadas.
- Dispositivos utilizados.
- Día preferido para el visionado.
- Tiempo empleado en la plataforma y en cada uno de los tipos de contenidos de la misma.
- Comprobar cuanto tiempo de dican a ver el capítulo, si lo dejan a medias etc.
- Las valoraciones de los consumidores.
- Preferencias en base a la zona geográfica.
- Información de sus RRSS.

2. Almacenamiento: ¿Cómo se guardarían los datos?

Todos los datos son guardados tanto de forma física, ya sean discos HDD, SSD o en la nube en distintos servidores.

Dichos datos se almacenarían en formato CSV.

3. Pre-procesado: Combinación de datos, selección, filtros...

Todos los datos obtenidos han de ser organizados, en este caso se utilizan filtros como por ejemplo: de edad, región, géneros más vistos, tanto actores como directores favoritos etc.

4. Análisis: ¿Cuáles son los resultados y las conclusiones?

El resultado de estos análisis es poder realizar una inversión a prácticamente riesgo cero, ya que gracias a todos los datos obtenidos el producto que se ofrece

5. Visualización: ¿Qué visualizaciones podrían ser interesantes para tomar decisiones?

Poder tener distintos gráficos que visualizar para comprimir toda la información de forma que sea sencilla la visualización de los mismos permitiendo que el tiempo destinado a la toma de decisiones sea menor y la decisión tomada sea la más óptima.

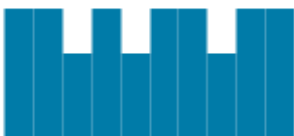
6. Publicación (o distribución). ¿Se ejecutó alguna estrategia adicional relacionada con la ciencia de datos para maximizar el impacto o la distribución? Podéis pensar alguna más.

Crearon 10 trailers con el fin de acercar la mayor audiencia posible a la serie, haciendo que cada uno de los trailers se asemejase al uso que hacían de la plataforma.

Pregunta 2

1. Buscad y seleccionad un portal que proporcione datos abiertos (una administración, una empresa, etc.).

En <https://www.kaggle.com/datasets> podemos acceder a multitud de CSV's como por ejemplo el de la siguiente imagen:

About this file					
Screentime - App Ranking					
Date	Rank 1	Rank 2	Rank 3		
Date	Rank 1	Rank 2	Rank 3		
	Whatsapp	81%	Instagram	30%	WeChat
	Instagram	11%	Zoom	15%	Chrome
	Other (2)	7%	Other (15)	56%	Other (14)
08/26/2022	Whatsapp	Instagram	Netflix		
08/27/2022	Whatsapp	Instagram	Chrome		
08/28/2022	Whatsapp	Instagram	WeChat		
08/29/2022	Whatsapp	Chrome	WeChat		
08/30/2022	Whatsapp	LinkedIn	Chrome		
08/31/2022	Whatsapp	Zoom	WeChat		
09/01/2022	Instagram	Whatsapp	Spotify		
09/02/2022	Whatsapp	WeChat	Instagram		
09/03/2022	Facebook	YouTube	Instagram		

2. Explorad el contenido disponible. ¿Qué tipo de conjuntos de datos proporcionan? ¿Os parece que el contenido es fácilmente accesible (es decir, es fácil de utilizar)?

Podemos filtrar las categorías de los distintos conjuntos de datos, visualizar en gráficos los datos que nos proporcionan, descargar los distintos csv almacenados en la página etc.

Todo esto de una forma sencilla ya que únicamente requeriremos de una cuenta en Kaggle.

3. Seleccionad uno de los conjuntos de datos disponibles del portal seleccionado, explicad su contenido concreto y pensad una aplicación práctica en la que puedan ser útiles.

La imagen mostrada anteriormente nos permite ver de forma sencilla las 3 apps más utilizadas y en qué porcentaje, además de poder filtrar los datos.

Todos estos datos pueden ser usados por ejemplo para decidir qué aplicación es mejor para anunciar un producto.