



VISUALIZACIÓN DE INFORMACIÓN

Hemos decidido realizar un trabajo de investigación sobre el tema de la producción cinematográfica. Hemos llevado a cabo una búsqueda exhaustiva sobre el tema y hemos encontrado un conjunto de datos en Kaggle (<https://www.kaggle.com/rounakbanik/the-movies-dataset/data>) obtenidos a partir del recurso IMDb (<http://www.imdb.com>). Después de refinarlo y seleccionar las 1.000 películas con mayor presupuesto, hemos creado el fichero “IMDb (dataset)”.

Aprovechando los datos que contiene este fichero, nos gustaría narrar de forma visual las siguientes historias:

- a) Representad de forma geográfica (que incluya también una representación cuantitativa por presupuestos) los países productores de las películas.
- b) Evaluad de forma interactiva y dinámica cómo ha sido la evolución cuantitativa de los géneros Drama, Aventura, Comedia y Acción durante el periodo 1995-2005. Utilizad para ello, por ejemplo, la herramienta Flourish (<https://flourish.studio>).
- c) Identificad cuál ha sido la película más taquillera y cuál fue la décima que más recaudó. ¿Cómo quedan el resto de las películas que forman el top-ten de la taquilla?
- d) Describid, de forma estática, la evolución cuantitativa de los géneros cinematográficos durante el periodo 1995-2005.
- e) Comparad, respecto a la película *The Lord of the Rings: The Two Towers*, los ingresos que obtuvieron las veinte películas más taquilleras.
- f) Mostrad si existe algún tipo de correlación (positiva o negativa) entre el presupuesto y los ingresos de las cien películas más taquilleras. Y ¿entre los ingresos recaudados y la puntuación media obtenida en las votaciones de los espectadores? y ¿entre el número de visitantes de la página de la película en IMDb (popularidad) y la puntuación media obtenida en las votaciones de los espectadores?
- g) Analizad de forma visual el ROI (retorno de la inversión) de las películas estrenadas en 1990.
- h) Describid cómo es la distribución del conjunto total de películas utilizando como criterio el presupuesto de las mismas.
- i) En cada uno de los años del periodo 1995-2005, mostrad la relación cuantitativa que mantienen los géneros cinematográficos.

j) Representad (en el mínimo espacio posible y de forma interactiva) los ingresos de las películas de los años 1990 y 1991 de forma que se muestre la estructura jerárquica entre el género (el primero que aparece en la celda), el año de estreno y la valoración de los espectadores (películas mejor valoradas (con una valoración superior a 6); películas peor valoradas (con una valoración igual o inferior a 6)).

k) Identificad al menos tres historias más que puedan ser explicadas con los datos contenidos en el informe. Representar de forma adecuada esos datos.