

# Desentrañando el Éxito del Anime: Un Análisis Exploratorio de Datos

Este análisis exploratorio de datos se ha realizado sobre un conjunto de datos que contiene información detallada sobre una gran variedad de animes.

## 1. Objetivo del Análisis

El objetivo de este análisis es ir más allá de las listas de ranking de MyAnimeList y examinar la calidad de los datos, identificar patrones y relaciones entre las variables, y obtener información que permita un análisis más profundo.

Se busca comprender qué elementos hacen que ciertas series resuenen con la audiencia y alcancen la cima de la popularidad.

## 2. Metodología

### 2.1. Limpieza y Preprocesamiento de Datos

Se llevó a cabo un proceso exhaustivo de limpieza y preprocesamiento del dataset de anime para asegurar la calidad y coherencia de los datos. Las etapas clave incluyeron:

- **Carga e Inspección Inicial:** Se cargó el dataset con pandas y se realizó una inspección inicial para comprender su estructura, **identificando la presencia de valores nulos y registros duplicados**.
- **Eliminación de Duplicados:** Se eliminaron filas duplicadas para evitar sesgos en el análisis.
- **Manejo de Valores Faltantes:**
  - Imputación con mediana o moda para valores numéricos.
  - Relleno con "unknown" en columnas categóricas.
  - Uso de la columna `anime_url` para inferir valores faltantes en `english_name`.
  - Eliminación de la columna `demographics` debido a su alto porcentaje de valores nulos.
- **Ingeniería de Características:** Creación de `premiered_season` y `premiered_year` a partir de `premiered`.
- **Transformación de Datos:** Conversión de `duration` a un formato unificado de minutos y segundos.
- **Verificación Post-Limpieza:** Análisis posterior para asegurar la correcta gestión de los valores nulos y la ausencia de duplicados.
- **Guardado del Dataset Limpio:** Se generó un nuevo archivo CSV con los datos procesados.

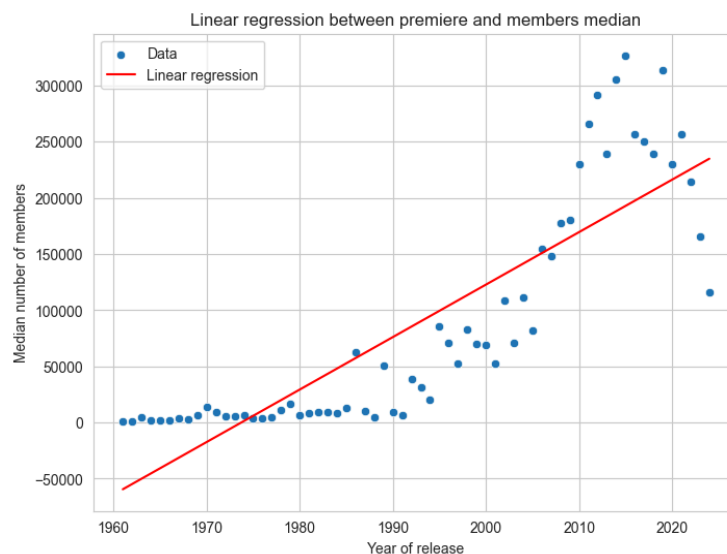
## 2.2. Análisis de Distribución y Pruebas de Hipótesis.

Para comprender mejor la distribución de los miembros en el dataset de anime, se realizaron análisis de la tendencia temporal, la variabilidad entre estudios y la normalidad de los datos.

### 2.2.1. Tendencia Temporal y Prueba de Hipótesis sobre la Cantidad de Miembros a lo largo del Tiempo

La figura muestra un gráfico de dispersión que representa la relación entre el año de estreno y la cantidad de miembros, con una línea de regresión que indica una tendencia creciente, lo que respalda la hipótesis inicial de que los animes más recientes tienden a tener una mayor cantidad de miembros.

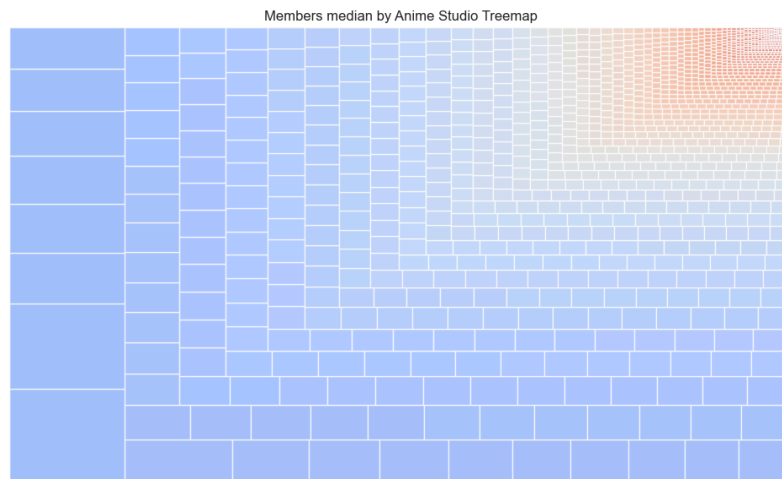
Esta tendencia se confirma con el cálculo del coeficiente de correlación de Spearman ( $\rho = 0.93$ ), elegido por su capacidad para evaluar relaciones monótonas sin asumir linealidad o normalidad. El valor de 0.93 indica una fuerte asociación positiva entre el año de estreno y la cantidad de miembros.



### 2.2.2. Variabilidad entre Estudios y Prueba de Hipótesis sobre la Cantidad de Miembros según el Estudio de Animación

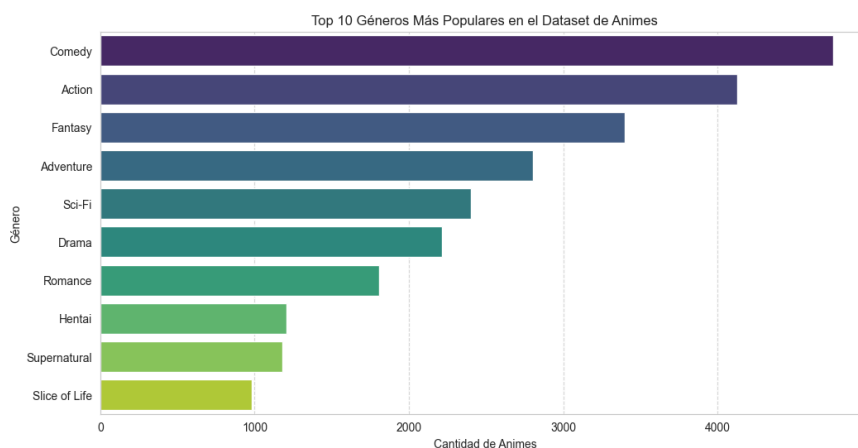
La figura muestra un treemap que visualiza la distribución de la cantidad de miembros por estudio de animación. El tamaño de cada rectángulo es proporcional a la cantidad media de miembros del estudio, lo que permite identificar fácilmente los estudios con mayor y menor cantidad de miembros.

Para determinar si existen diferencias significativas en la cantidad media de miembros entre los diferentes estudios de animación, se formuló la hipótesis nula de que no existen diferencias significativas. Se realizó una prueba ANOVA, adecuada para comparar medias de una variable continua entre múltiples grupos categóricos. Los resultados ( $p < 0.05$ ) **permiten rechazar la hipótesis nula**, confirmando diferencias significativas en la cantidad media de miembros según el estudio de animación.



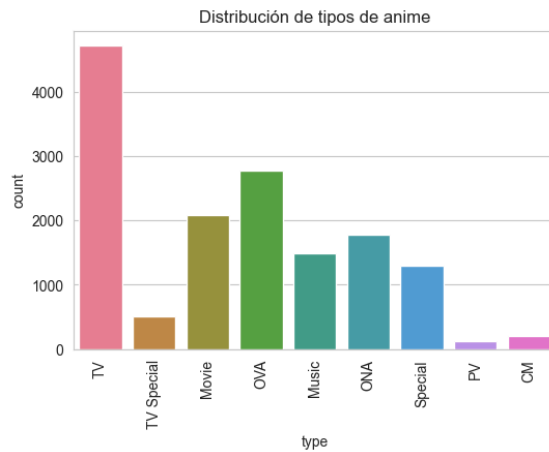
### 3. Principales Hallazgos

- Géneros Populares: **Un análisis de frecuencia de la columna 'genres' en el top 15,000 reveló que** Comedia, Acción, Fantasía y Aventura son los más frecuentes.

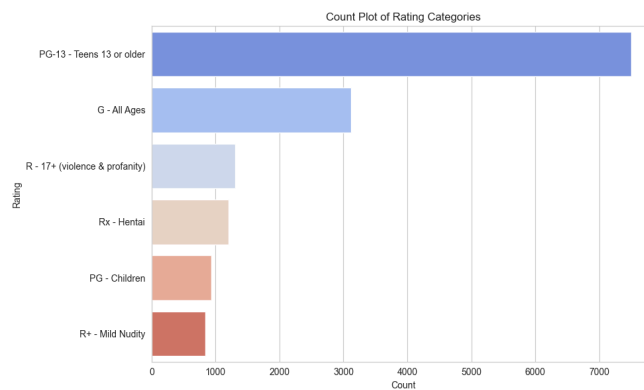


- **El análisis de correlación mostró que** animes con mayor número de miembros y favoritos tienden a estar mejor posicionados en el ranking.

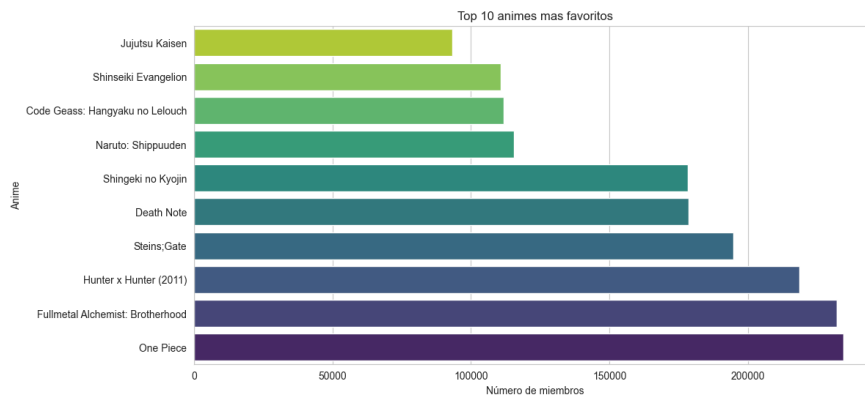
- Crecimiento de la audiencia: **Se observó en el análisis de tendencia temporal**, los animes lanzados después del 2000 tienen un mayor número de consumidores.
- **Un análisis de la distribución del tipo de anime entre los mejor valorados indicó que** observamos una alta frecuencia de animes de tipo 'TV'.



- El gráfico muestra la proporción de animes en cada categoría de clasificación de edad, revelando la predominancia de contenido clasificado como **PG-13 (Teens or Older)** y **G (All Ages)**.



- El gráfico de los 20 animes con mayor número de miembros destaca a **One Piece**, **Fullmetal Alchemist**, y **Hunter x Hunter** entre los más populares.



## 5. Conclusiones y Recomendaciones

El crecimiento constante de la audiencia de anime a lo largo del tiempo, evidenciado por la mayor cantidad de miembros en títulos recientes, subraya la continua relevancia del medio. La alta popularidad de géneros como Comedia, Acción, Drama, Fantasía y Aventura sugiere que estos siguen siendo pilares para atraer espectadores. La marcada diferencia en la audiencia entre estudios de animación indica que existen factores clave en sus estrategias que influyen en su éxito.

Se recomienda realizar análisis más profundos, como regresiones y estudios comparativos de estudios, para identificar los factores específicos que impulsan la popularidad del anime y las mejores prácticas en la industria.

# Diccionario:

Significado de cada columna en el dataset de anime que hemos proporcionado.

columnas\_anime = {

"anime\_id": "Identificador único del anime en MyAnimeList (MAL ID).",

"anime\_url": "Enlace URL a la página del anime en MyAnimeList.",

"image\_url": "Enlace URL de la imagen principal o de portada del anime (formato JPEG).",

"name": "Título oficial del anime.",

"english\_name": "Título oficial del anime en inglés, si está disponible.",

"japanese\_names": "Título oficial del anime en japonés, si está disponible.",

"score": "Puntuación o calificación promedio del anime en MyAnimeList (escala de 1 a 10; mayor puntuación indica mejor valoración).",

"genres": "Lista de géneros asociados al anime, separados por comas (ejemplo: Acción, Comedia, Fantasía).",

"themes": "Lista de temas asociados al anime, separados por comas (ejemplo: Psicológico, Viajes en el tiempo).",

"synopsis": "Breve resumen o descripción de la trama del anime.",

"type": "Tipo de anime (ejemplo: TV, Película, OVA, ONA, Especial, Música).",

"episodes": "Número de episodios de la serie de anime (para series de TV, OVAs, etc.).",

"producers": "Lista de compañías productoras involucradas en la creación del anime, separadas por comas.",

"studios": "Lista de estudios de animación responsables de la animación del anime, separados por comas.",

"source": "Material original del anime (ejemplo: Manga, Original, Novela ligera, Videojuego).",

"duration": "Duración de cada episodio en minutos (para series de TV, OVAs, etc.).",

"rating": "Clasificación por edad o de contenido del anime (ejemplo: PG-13, R+, G).",

"rank": "Posición del anime en el ranking basado en su puntuación (un rango menor indica una mejor posición). Este dataset se centra en los 15,000 animes mejor clasificados según este ranking.",

**"popularity":** "Ranking de popularidad del anime en MyAnimeList (un rango menor indica mayor popularidad).",

**"favorites":** "Número de usuarios que han marcado este anime como favorito en MyAnimeList.",

**"scored\_by":** "Número de usuarios que han contribuido a la puntuación del anime.",

**"members":** "Número de usuarios que ya han visualizado el anime o planean hacerlo.",

**"premiered\_season":** "Estación del año en la que se estrenó el anime.",

**"premiered\_year":** "Año de estreno del anime."

}