



### Ejercicio3: Análisis de Datos de Artistas y música

El objetivo de esta práctica es proporcionar una comprensión práctica y aplicada del análisis de datos utilizando Python y Pandas. A través del análisis de un conjunto de datos de artistas musicales (fichero `artistas.csv`), vamos a manipular, explorar y visualizar datos para obtener conclusiones sobre los mismos.

El conjunto de datos incluye información sobre artistas musicales, como nombres, países de origen, etiquetas (géneros y características), y estadísticas de oyentes y reproducciones de plataformas como MusicBrainz y Last.fm.

Columnas del fichero:

- **mbid**: Identificador único de MusicBrainz para el artista.
- **artist\_mb**: Nombre del artista según MusicBrainz.
- **artist\_lastfm**: Nombre del artista según Last.fm.
- **country\_mb**: País de origen del artista según MusicBrainz.
- **country\_lastfm**: País de origen del artista según Last.fm.
- **tags\_mb**: Etiquetas asociadas al artista según MusicBrainz.
- **tags\_lastfm**: Etiquetas asociadas al artista según Last.fm.
- **listeners\_lastfm**: Número de oyentes en Last.fm.
- **scrobbles\_lastfm**: Número de reproducciones en Last.fm.
- **ambiguous\_artist**: Indica si el artista aparece más de una vez en el listado.



Realizar los siguientes cálculos:

0. Limpieza y Preparación de Datos:

- Eliminar filas con datos faltantes.
- Eliminar filas con datos en idiomas que no podemos interpretar

1. Análisis Descriptivo Básico:

- Contar el número total de artistas.
- Calcular la media, mediana y desviación estándar del número de oyentes y reproducciones (scrobbles) en Last.fm.
- Identificar los países con más artistas representados y representarlos en un gráfico.

2. Análisis de Etiquetas:

- Determinar las etiquetas más comunes y realizar un análisis de frecuencia. (Expandir valores y posteriormente contar)
- Explorar las relaciones entre etiquetas y países (por ejemplo, qué etiquetas son más populares en ciertos países).

3. Análisis de Duplicados y consistencia:

- Verificar si hay artistas duplicados y analizar las posibles causas de duplicación.
- Verificar inconsistencias entre artistas según MusicBrainz y Last.fm



#### Visualización de Datos:

4. Crear gráficos de barras o gráficos de tarta para visualizar la distribución de artistas por país.
  5. Utilizar gráficos de dispersión para visualizar la relación entre el número de oyentes y el número de reproducciones.
  6. Obtener los artistas de España, ver cual es el primer artista de España según sus reproducciones y extraer información sobre las etiquetas más comunes.
- Realizar un gráfico de tarta con la representación de los 10 artistas más escuchados de España
7. ¿Qué otros análisis o tratamiento de datos se te ocurre que podemos extraer? Desarrolla y expón tu idea.