

Dataset: Lista Top Canciones por año

Angel Vicente Sanjuán Martín
03 de Abril de 2021

Descripción

Nombre del dataset : **listsongs.csv**

Nombre del esquema : **Esquema_listsong.ods**

Son el conjunto de datos obtenido como parte de la actividad, la lista de las 100/50 (dependiendo del año) mejores canciones según la página [el portal de música](#) desde los años 2009 a 2020

Contenido

Para cada registro se recogen los siguientes atributos

- **Year:** Año al que pertenece la entrada en lista.
- **Ranking:** Puesto obtenido dentro del año.
- **Song:** Título de la canción.
- **Player:** Cantante que interpreta la canción.
- **WeeksList:** Número de semanas que la canción ha estado en las listas de el portal de musica.
- **Url:** Enlace de la canción en la página el portal de música

Agradecimientos

Los datos han sido obtenidos de la página [El portal de Música](#), un sito perfecto para extraer este contenido gracias a la facilidad y disponibilidad de su contenido.

El lenguaje de programación usado ha sido Python y las técnica de programación para obtener la información han sido Web Scraping, el cual facilita enormemente la extracción de la información alijada en las paginas html.

Inspiración

Elegí este contexto ya que soy un gran aficionado a la música en general y no me imagino un solo día sin escucharla,

Todos los datos obtenidos en el Dataset, pueden ser utilizados por personas que quieran conocer los éxitos de los años ya pasados y una sencilla forma de escucharlos.

Licencia

La licencia escogida para la publicación de este conjunto de datos ha sido GNU GENERAL PUBLIC LICENSE Version 3 .

Es una licencia de derechos de autor ampliamente usada en el mundo del software libre y código abierto, lo que hace es permitir que los usuarios finales (personas, organizaciones, etc) tengan libertad de usar, estudiar, compartir (copiar) y modificar el software. Su objetivo es doble: recalcar que el software cubierto por esta licencia es libre y, por otro lado, protegerlo (mediante una práctica

conocida como copyleft) de intentos de apropiación que restrinjan esas libertades a nuevos usuarios cada vez que la obra (software, plugin, template) es distribuida, modificada o mejorada.

Código fuente y dataset

Tanto el código fuente escrito para la extracción de datos como el dataset generado, así como el esquema de los datos, pueden ser accedidos a través de este enlace :

<https://github.com/avsanjuan/scrapingTopSongs>

Acceso a la publicación del dataset en zenodo a través del siguiente enlace:

<https://zenodo.org/record/4666540>

Recursos

1. <https://www.python.org/>
2. Mitchel, R. (2018). Web Scraping with Python 2e: O'Reilly Media,
3. <https://github.com/>
4. <https://zenodo.org/>