

Anexo IV. Actualización de datos mediante la API de Spotify

A continuación se presenta el código utilizado para tanto obtener los datos ausentes de las variables 'popularity' y 'time_signature' como para actualizarlos en aquellas tuplas de las cuales sí conocíamos dichos valores, con el fin de trabajar con datos relevantes:

El código usa la librería spotipy, una librería en Python que facilita el uso de la API de Spotify, la principal fuente de datos de nuestro estudio.

Atención: este documento contiene información personal para la autenticación del programa en Spotify. Bajo la asunción de que este programa se presenta con fines académicos se restringe su publicación o reparto. Los únicos usuarios autorizados para compartir esta información son los integrantes del grupo y las docentes de la asignatura Proyecto II, Integración de datos 2020 - 2021.

```
#Abrimos Los paquetes necesarios
```

```
import spotipy  
  
from spotipy.oauth2 import SpotifyClientCredentials  
  
import pandas as pd
```

```
#Abrimos el archivo que tenemos que actualizar, y inicializamos la API
```

```
output = pd.read_csv('outputpop.csv')  
  
clientID, clientSec = '0d2beefe0f0846eea477a529f194cad5',  
                    '5fdcae7eaa6e420d932ecb01a9506bd8'  
  
ccm = SpotifyClientCredentials(client_id=clientID, client_secret=clientSec)  
  
sp = spotipy.Spotify(client_credentials_manager=ccm)
```

```
#CÓDIGO PARA POPULARITY
```

```
try:
```

```
    for i in range(len(output)+1):
```

```
        iden = output['id'][i] #Identificador de la canción
```

```
        output['popularity'][i] = sp.track(f'spotify:track:{iden}')['popularity']
```

```
#Actualizamos el archivo
```

```
except: #Si hay algún error, guardamos el archivo, y nos quedamos con el último índice, para no perdernos

    print(i)

output.to_csv('outputpop.csv', index = False)

#CÓDIGO PARA TIME_SIGNATURE

bool_series = pd.isnull(output['time_signature']) #En este caso solo recorreremos los nulos

for i, row in output[bool_series].iterrows():

    iden = row['id']

    output['time_signature'][i] =
    sp.audio_features(f'spotify:track:{iden}')[0]['time_signature'] #Los actualizamos

output.to_csv('outputpop.csv', index = False)
```