## Base de données

Projet de Base de données – Spotify API https://github.com/AmokraneMancer/project\_bdd\_spotify\_api

Réalisé par :

MANCER Mohammed Amokrane

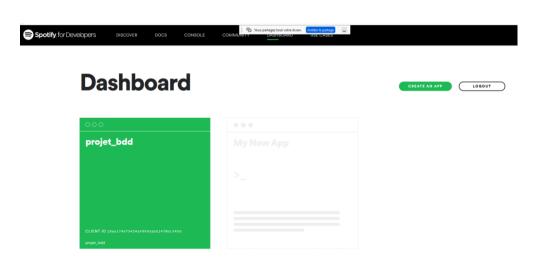
**SADDEDINE Hassina** 

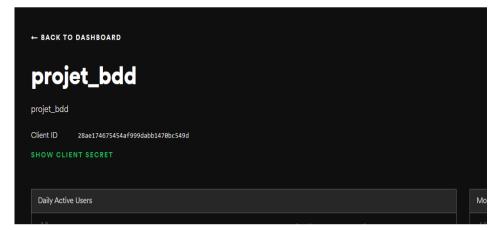
# Présentation de projet

Le projet consiste à récupérer les données à partir d'une API publique et puis de les stocker dans une base de données relationnelle.

# L'API de Spotify

## Créer une application

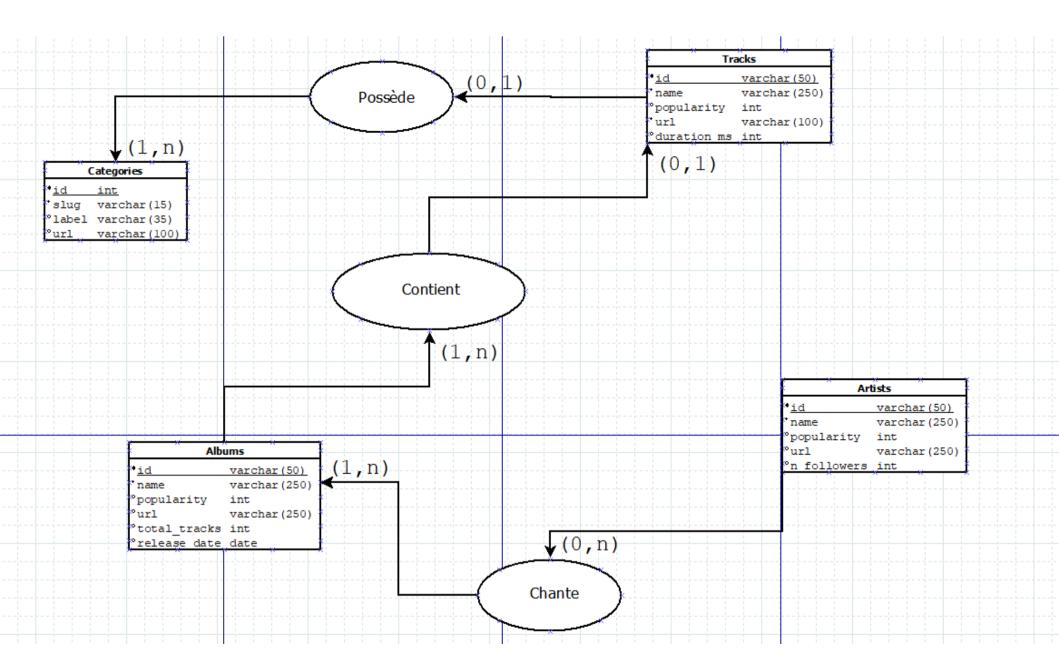




## Génération d'une clé (Token) pour se connecter à l'api

```
Entrée [1]: import pyodbc
            import requests
            import pandas as pd
            import json
Entrée [2]: CLIENT ID = '28ae174675454af999dabb1470bc549d'
            CLIENT SECRET = '7f05016f535e4260b7d3dcf771f14ecc'
            AUTH URL = 'https://accounts.spotify.com/api/token'
            # POST
            auth response = requests.post(AUTH URL, {
                'grant type': 'client credentials',
                'client id': CLIENT ID,
                'client secret': CLIENT SECRET,
            })
            # convert the response to JSON
            auth response data = auth response.json()
            # save the access token
            access token = auth response data['access token']
Entrée [3]: access token
  Out[3]: 'BQDG3uPTyM9f0d8HGVGuZHVMvqnFNPdvZ-9v8Cq5i1qdyvsGhlov clU8zdgNcTOGxnnc-mjVskjfCGphgY'
Entrée [4]: headers = {
                'Authorization': 'Bearer {token}'.format(token=access token)
```

## Schéma entité-relation



## Schéma relationnel

### Schéma initial non optimisé :

Tracks (id, name, popularity, url, duration\_ms)

Artists (id, name, popularity, url, n\_followers)

Albums (id, name, popularity, url, total\_tracks, release\_date)

Categories (id, slug, label, url)

Possède (#id track, #id cat)

Contient (#id track, #id album)

Chante (#id\_artist, #id\_album)

### Schéma optimisé:

Tracks (id, name, popularity, url, duration\_ms, #id cat, #id album)

Artists (<u>id</u>, name, popularity, url, n\_followers)

Albums (<u>id</u>, name, popularity, url, total\_tracks, release\_date)

Categories (id, slug, label, url)

Possède (#id\_track, #id\_cat)

Contient (#id\_track, #id\_album)

Chante (#id\_artist, #id\_album)

1FN OK

**2FN OK** 

**3FN OK** 

## Récupérer des catégories via 'browse'

## **Get a List of Categories**

Get a list of categories used to tag items in Spotify (on, for example, the Spotify player's "Browse" tab).

**Endpoint** 

GET https://api.spotify.com/v1/browse/categories

#### Out[56]:

	href	icons	id	name
0	https://api.spotify.com/v1/browse/categories/t	[{'height': 275, 'url': 'https://t.scdn.co/med	toplists	Top Lists
1	https://api.spotify.com/v1/browse/categories/h	[{'height': None, 'url': 'https://t.scdn.co/im	holidays	Happy Holidays
2	https://api.spotify.com/v1/browse/categories/2020	[{'height': None, 'url': 'https://t.scdn.co/im	2020	2020 Wrapped
3	https://api.spotify.com/v1/browse/categories/a	[{'height': None, 'url': 'https://t.scdn.co/im	at_home	At Home
4	https://api.spotify.com/v1/browse/categories/w	[{'height': None, 'url': 'https://t.scdn.co/im	wellness	Wellness
5	https://api.spotify.com/v1/browse/categories/r	[{'height': None, 'url': 'https://t.scdn.co/im	radar	RADAR
6	https://api.spotify.com/v1/browse/categories/mood	[{'height': 274, 'url': 'https://t.scdn.co/med	mood	Mood
7	https://api.spotify.com/v1/browse/categories/pop	[{'height': 274, 'url': 'https://t.scdn.co/med	pop	Pop
8	https://api.spotify.com/v1/browse/categories/h	[{'height': 274, 'url': 'https://t.scdn.co/med	hiphop	Hip Hop
9	https://api.spotify.com/v1/browse/categories/e	[{'height': 274, 'url': 'https://t.scdn.co/med	edm_dance	Dance/Electronic
10	https://api.spotify.com/v1/browse/categories/i	[{"height": None, 'url": 'https://t.scdn.co/im	indie_alt	Indie
11	https://api.spotify.com/v1/browse/categories/a	[{"height": None, 'url": 'https://t.scdn.co/im	alternative	Alternative

## Catégories après nettoyage

#### t[78]:

	url	slug	label
0	https://api.spotify.com/v1/browse/categories/t	toplists	Top Lists
1	https://api.spotify.com/v1/browse/categories/h	holidays	Happy Holidays
2	https://api.spotify.com/v1/browse/categories/2020	2020	2020 Wrapped
3	https://api.spotify.com/v1/browse/categories/a	at_home	At Home
4	https://api.spotify.com/v1/browse/categories/w	wellness	Wellness
5	https://api.spotify.com/v1/browse/categories/r	radar	RADAR
6	https://api.spotify.com/v1/browse/categories/mood	mood	Mood
7	https://api.spotify.com/v1/browse/categories/pop	pop	Pop
8	https://api.spotify.com/v1/browse/categories/h	hiphop	Hip Hop
9	https://api.spotify.com/v1/browse/categories/e	edm_dance	Dance/Electronic
10	https://api.spotify.com/v1/browse/categories/i	indie_alt	Indie

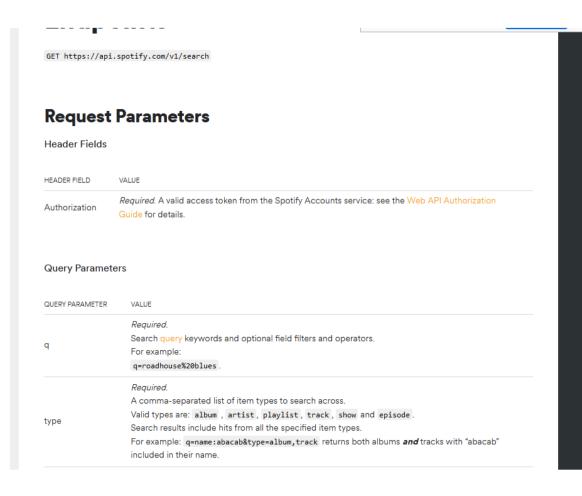
## Insertion dans la base de donnés

```
con = pyodbc.connect('Driver={SQL Server};Server=DESKTOP-0G5VHDF\SQLEXPRESS;Database=Spotify;Trusted_Connection=yes;'

cursor = con.cursor()
  # Insert Dataframe into SQL Server:
  id = 0
  for index, row in categories.iterrows():
      cursor.execute("INSERT INTO categories (id,slug,label,url) values(?,?,?,?)", id, row.slug, row.label, row.url)
      id = id + 1
  con.commit()
  cursor.close()
```

```
| Solect | From dbo.categories; | Solect | From dbo.categories; | Solect | From dbo.categories; | Solect | Sole
```

## Récupérer les chansons et albums de chaque catégorie



## **Get an Album**

Get Spotify catalog information for a single album.

### **Endpoint**

GET https://api.spotify.com/v1/albums/{id}

```
[6]: # charger les catégories à partir de la base de données
    con = pyodbc.connect('Driver={SQL Server}; Server=DESKTOP-OG5VHDF\SQLEXPRESS; Database=Spotify; Trusted_Connection=yes;')
    req = 'select slug from categories'
    cursor = con.cursor()
    cursor.execute(req)
    categories_list = []
    for row in cursor.fetchall():
        categories_list.append(row[0])
    categories_list = {}
    for i in range(len(categories_list)):
        category_name = categories_list[i]
        response_list[category_name] = []
        SEARCH_URL = 'https://api.spotify.com/v1/search?q={}&type=track&limit=50'.format(category_name)
        r = requests.get(SEARCH_URL, headers=headers)
        response_list[category_name].append(r.json())
```

duration_ms	evolicit	external_ids	external_urls	href	id	is_local	name	popularity	preview_url	track n
udiduoii_iiis	Oxplicit	external_lus	external_uns	11161	Iu	13_IOCAI	Haille	popularity	preview_uii	uack_III
200221	True	{'isrc': 'USSM12004501'}	{'spotify': 'https://open.spotify.com /track/6E	https://api.spotify.com /v1/tracks /6EDO9iiTtwNv	6EDO9iiTtwNv6waLwa1UUq	False	POPSTAR (feat. Drake)	90	https://p.scdn.co/mp3- preview /f06dde2517250354	
166560	True	{'isrc': 'USQX91900309'}	{'spotify': 'https://open.spotify.com/track/6u	https://api.spotify.com /v1/tracks /6uFn47ACjqYk	6uFn47ACjqYkc0jADwEdj1	False	Pop Out (feat. Lil Tjay)	84	https://p.scdn.co/mp3- preview /505ae166c11669c6	
227478	True	{'isrc': 'USAT22003620'}	{'spotify': 'https://open.spotify.com /track/2M	https://api.spotify.com /v1/tracks /2MbdDtCv5LUV	2MbdDtCv5LUVjYy9RuGTgC	False	WHATS POPPIN (feat. DaBaby, Tory Lanez & Lil W	88	https://p.scdn.co/mp3- preview /3748f897bbea6068	
220537	True	{'isrc': 'USEP41915005'}	{'spotify': 'https://open.spotify.com /track/4G	https://api.spotify.com /v1/tracks /4GssR27i.leam	4GssB27iJeqmfGxS94Tfij	False	Popular Monster	79	https://p.scdn.co/mp3- preview /18707ef9fd54e5d0	<b>~</b>

### Insertion des chansons et albums

```
.....
    . .
     create table tracks (
                                                                   create table albums (

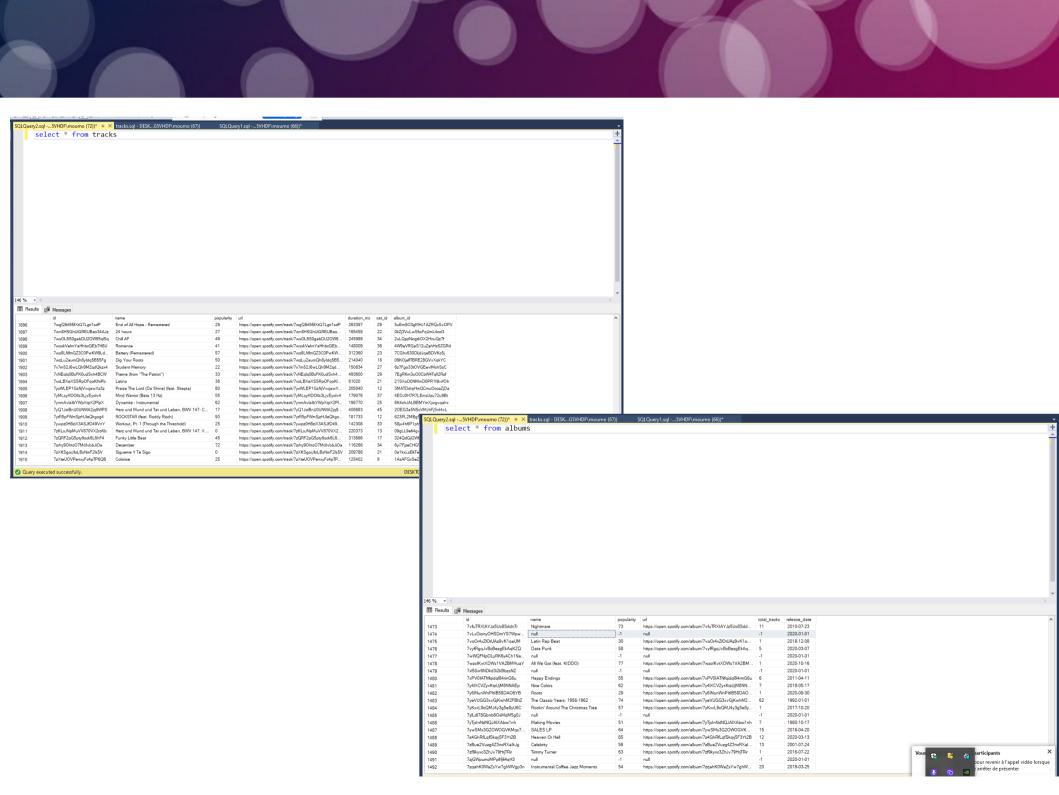
→id varchar(50).

→id varchar(50) primary key,

→ name varchar(250).
      ---*popularity int,
                                                                      *name varchar(250),
      wurl varchar(250).
                                                                      ⇒popularity int.
      \longrightarrow id_album varchar(50),
      duration ms int,
                                                                      →url varchar(250),
      -->cat id int,

→total tracks int.

      FOREIGN KEY (cat id) REFERENCES categories(id),
                                                                      ⇒release date date
      );
for index, row in tracks.iterrows():
    trv:
        album = row.album
        id album = album['id']
        url = row.external urls
        url = url['spotify']
        # l'album
        ALBUM URL = 'https://api.spotify.com/v1/albums/{}'.format(id album)
        r = requests.get(ALBUM URL, headers=headers)
        json albums = r.json()
        album name = json albums['name'] if 'name' in json albums.keys() else 'null'
        album popularity = json albums['popularity'] if 'popularity' in json albums.keys() else -1
        album t tracks = json albums['total tracks'] if 'total tracks' in json albums.keys() else -1
        album date = json albums['release date'] if 'release date' in json albums.keys() else '2020-01-01'
        album url = json albums['external urls']['spotify'] if 'external urls' in json albums.keys() else 'null'
        cursor.execute("INSERT INTO tracks (id , name, popularity, url, duration ms, cat id, album id) values(?,?
                       row.id, row.label, row.popularity, url, row.duration ms, cat id, id album)
        cursor.execute("INSERT INTO albums (id ,name,popularity,url, total tracks, release date) values(?,?,?,?,?
                       id album, album name, album popularity, album url, album t tracks, album date )
```



## Récupérer et insérer les artistes

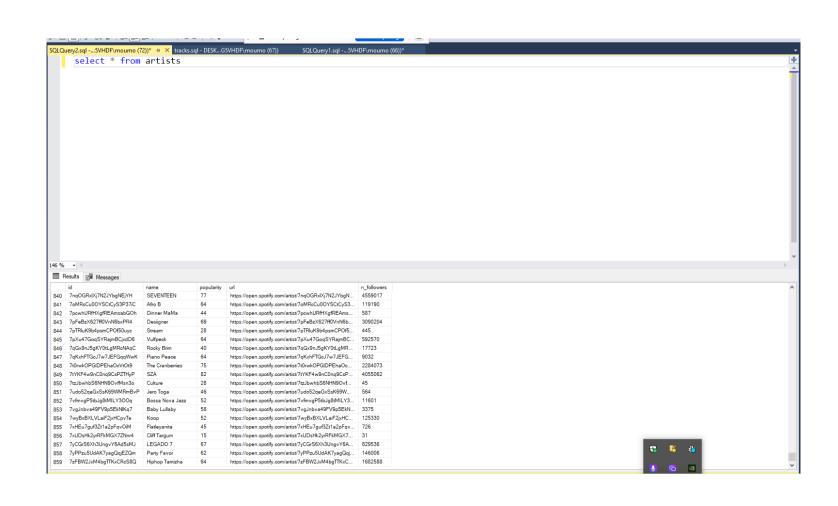
### **Get an Artist**

Get Spotify catalog information for a single artist identified by their unique Spotify ID.

#### **Endpoint**

GET https://api.spotify.com/v1/artists/{id}

```
cursor = con.cursor()
album by artist = []
for album id in albums list:
   ALBUM URL = 'https://api.spotify.com/v1/albums/{}'.format(album id)
    r = requests.get(ALBUM URL, headers=headers)
   if 'artists' in r.json().keys():
       artists = r.ison()['artists']
   else:
   for ar in artists:
            artist id = ar['id']
           ARTIST URL = 'https://api.spotify.com/v1/artists/{}'.format(artist id)
            artist json = requests.get(ARTIST URL, headers=headers).json()
            artist name = artist json['name']
            artist popularity = artist json['popularity']
            artist url = artist json['external urls']['spotify']
            n followers = artist json['followers']['total']
            album by artist.append((album id, artist id))
            cursor.execute("INSERT INTO artists (id , name, popularity, url, n followers) values(?,?,?,?,?)",
                               artist id, artist name, artist popularity, artist url, n followers)
       except Exception as ex:
            sqlstate = ex.args[0]
            if sqlstate == '23000':
                pass
con.commit()
cursor.close()
```



## Insérer dans la table chante (album\_by\_artist)

