
Module 1 Final Project:

A short analysis on 2019

European metal festivals 🤘

Anthony Blanc -Nguyen • 01.08.2020

The subject

Due to the recent COVID-19  pandemic, most public gatherings have been banned by public authorities to prevent the spread of the virus.

All music festivals throughout Europe have chosen to cancel in order to protect the festival-goers.


The dataset is focused on the previous edition of 2019 and gathers info on the festivals lineups and social media presence with Twitter.

Main challenges & Strengths

Challenges

- Find an interesting subject
- Draw a problematic to answer
- Check available data on the subject

Strengths

- Coding is not an issue
- Rubber ducks are everywhere 

Data sources

Web scraping

The original table was retrieved from the following link:

[https://www.festival-alarm.com/us/Categories/Metal-festivals/\(year\)/2019](https://www.festival-alarm.com/us/Categories/Metal-festivals/(year)/2019)

Basic techniques of data wrangling using BeautifulSoup and regex to recreate the original table.

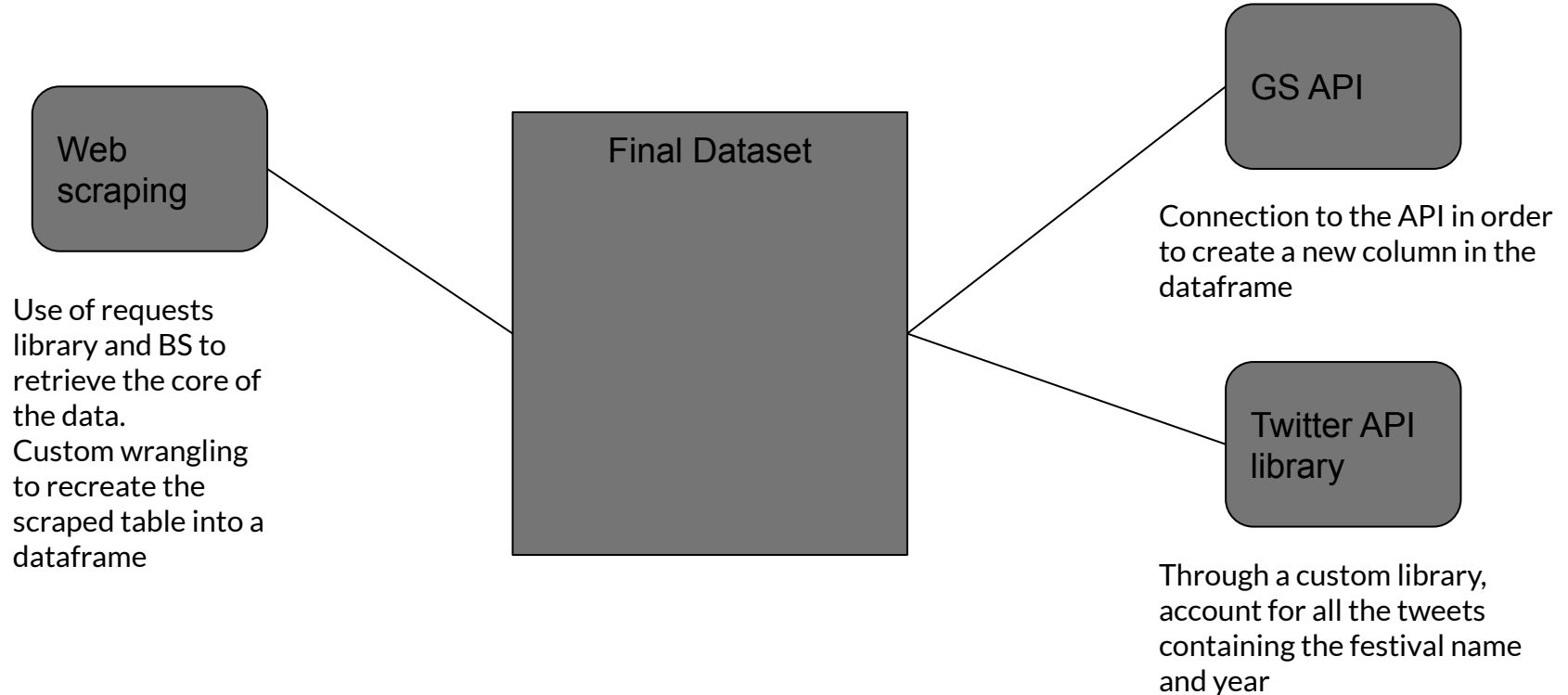
Google sheet API

For the sake of the exercise, a small database was created to pair each festival with all the bands who attended. The GS API served to retrieve the data from this private source.

Twitter API (public API & library)

Last addition was to create a “social presence” column to account for the number of tweets related to the festival.

Workflow



Main insights

Festival lineups

You can make queries on the table to see which festivals a particular band attended.

Social media presence

A new column accounts for the number of tweets.

An unclear objective

Though answering to the challenge, the difficulty to gather consistent data was the main issue I ran into. However, some use cases can be derived from the data gathered:

- Plan your summer festivals schedule
- Estimate the festival hype

Epilogue

1. Finding a subject with consistent data can be quite a challenge
2. A table on a website is the worst kind of data
3. Sadly, no festival this summer



